UNIVERSIDADE ESTADUAL DE CAMPINAS

Instituto de Física Gleb Wataghin

LUCAS DA SILVA POLLYCENO

Multipartite Bell nonlocality and communication problems

Não-localidade de Bell multipartida e problemas de comunicação

Campinas
2024

Lucas da Silva Pollyceno

Multipartite Bell nonlocality and communication problems

Não-localidade de Bell multipartida e problemas de comunicação

Thesis presented to the Institute of Physics Gleb Wataghin of the University of Campinas in partial fulfillment of the requirements for the degree of Doctor of Science, in in the area of Physics.

Tese apresentada ao Instituto de Física Gleb Wataghin da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor em Ciências, na área de Física.

Supervisor: Prof. Dr. Rafael Luiz da Silva Rabelo

Este exemplar corresponde à versão final da Tese defendida pelo aluno Lucas da Silva Pollyceno e orientado pelo Prof. Dr. Rafael Luiz da Silva Rabelo.

Campinas
2024

Ficha catalográfica
Universidade Estadual de Campinas (UNICAMP)
Biblioteca do Instituto de Física Gleb Wataghin
Lucimeire de Oliveira Silva da Rocha - CRB 8-9174

Pollyceno, Lucas da Silva, 1996-
P766m
Multipartite Bell nonlocality and communication problems / Lucas da Silva Pollyceno. – Campinas, SP : [s.n.], 2024.

Orientador(es): Rafael Luiz da Silva Rabelo.
Tese (doutorado) – Universidade Estadual de Campinas (UNICAMP), Instituto de Física Gleb Wataghin.

1. Não-localidade de Bell. 2. Causalidade da informação. 3. Problemas de comunicação. 4. Correlações quântica. I. Rabelo, Rafael Luiz da Silva, 1986-. II. Universidade Estadual de Campinas (UNICAMP). Instituto de Física Gleb Wataghin. III. Título.

Informações complementares

Título em outro idioma: Não-localidade de Bell multipartida e problemas de comunicação
Palavras-chave em inglês:
Bell nonlocality
Information causality
Communication problems
Quantum correlations
Área de concentração: Física
Titulação: Doutor em Ciências
Banca examinadora:
Rafael Luiz da Silva Rabelo [Orientador]
Pawel Kurzynski
Renato Moreira Ângelo
Marco Túlio Coelho Quintino
Marcus Vinicius Segantini Bonança
Data de defesa: 17-12-2024
Programa de Pós-Graduação: Física

Identificação e informações acadêmicas do(a) aluno(a)
- ORCID do autor: https://orcid.org/0000-0002-9847-1753
- Currículo Lattes do autor: http://lattes.cnpq.br/8917319799683392

UNICAMP

IFGW INSTITUTO DE FÍSICA GLEB WATAGHIN

MEMBROS DA COMISSÃO EXAMINADORA DA TESE DE DOUTORADO DO ALUNO **LUCAS DA SILVA POLLYCENO – RA 230338** APRESENTADA E APROVADA AO INSTITUTO DE FÍSICA GLEB WATAGHIN, DA UNIVERSIDADE ESTADUAL DE CAMPINAS, EM 17/12/2024.

COMISSÃO JULGADORA:

- Prof. Dr. Rafael Luiz da Silva Rabelo (IFGW/ UNICAMP) - Presidente e Orientador

- Prof. Dr. Pawel Kurzynski (Adam Mickiewicz University)

- Prof. Dr. Renato Moreira Angelo (Departamento de Física - UFPR)

- Prof. Dr. Marco Túlio Coelho Quintino (Paris-Sorbonne Université)

- Prof. Dr. Marcus Vinicius Segantini Bonança (IFGW/ UNICAMP)

OBS.: Ata da defesa com as respectivas assinaturas dos membros encontra-se no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa da Unidade.

CAMPINAS
2024

Acknowledgements

I want to begin by expressing my deepest gratitude to my professor and advisor, Rafael, without whom this thesis's conception, development, and completion would not have been possible. I am deeply grateful for having walked this path and for your trust in me from our first contact in 2018. I thank you for the stimulating freedom and autonomy you provided, which, coupled with active mentorship, were crucial not only for the execution of this project but also for my mental well-being. I will carry every conversation, advice, and lesson with me.

I extend my heartfelt thanks to my wife and life partner, Rafa, my best friend and greatest love. Thank you for everything: for sharing life with me in Campinas, for every weekend, every day out, and every trip, for your unwavering support, and for enduring the six months we were apart while I was in Poland. I am deeply grateful for every moment we have shared, every plant we grew together, every film and song, and every conversation and laugh. Thank you for knowing me so deeply, allowing me to know you so profoundly, and being such an integral part of my life.

Agradeço aos meus pais, Marilda e Izaque, por absolutamente tudo. Obrigado por seu amor, por sempre acreditarem em mim, por terem me incentivado aos estudos desde pequeno, e por, mesmo diante das dificuldades, me proporcionarem as condições para que eu pudesse viver o que vivi. Agradeço profundamente seu apoio incondicional durante todas as fases de minha formação, em especial durante a graduação e durante minha vinda para Campinas no mestrado. Obrigado por serem meu porto seguro em todo esse tempo.

I want to thank my friends in Curitiba, with whom, despite life's setbacks, I was able to maintain, rebuild, and deepen ties. To João, for whom I hold genuine admiration: thank you for always taking time to see me during my visits to Curitiba and for your trips to Campinas. I am immensely grateful that physics brought us together as friends. To Eduardo, thank you for your sincere, easygoing, and unconditional friendship since our early childhood. You hold a special place in my heart and were undoubtedly fundamental at the beginning of my academic journey. To Kalel, thank you for hosting me in Germany, visiting me in Poland, sharing trips, engaging in countless conversations, and exchanging long audio messages about physics and beyond. To Carlos, thank you for keeping our video calls alive and for all your support with data science. I also express my heartfelt gratitude to my childhood friends and brothers, Junior and Maylon. I am truly fortunate to have you in my life.

To all the friends I made in Campinas, my heartfelt thanks. To Marcelo, in particular, thank you for your companionship, lightness, and countless conversations in our office. Life gifted me with a brother, someone from whom I have learned so much and will carry with me forever. Salud! To Zé, I am deeply honored by your trust in me early on and profoundly grateful for the deep friendship and partnership we built during this time. To my friend and housemate João, who has been a significant part of my life since my first week in Campinas. Your scientific spirit has always inspired me, and I look up to you as an example of humanity and resilience.

I am grateful to the friends and colleagues I made in the MathFoundQ research group. A special thanks to Carlos, Pedro, Lucas, Arthur, Guilherme, Clara, Denis, Tassius, Gabriel Ruffolo, Vitor, and Gabriela, with whom I enjoyed working, learning, traveling, playing football, and sharing laughter. A special mention goes to Carlos for your friendship, for sharing joys and challenges since the beginning of my graduate studies, for your collaboration, and for reviewing this thesis. To Pedro, from whom I learned so much about physics, science, and commitment over the years. I also extend my thanks to Professor Marcelo for his invaluable insights and his inspiring scientific dedication and engagement. Finally, I am deeply grateful to Professor Rafael Chaves for his essential collaboration throughout my academic journey.

I am grateful for all the people I had the chance to meet during my period abroad. I would like to thank Marcin for the opportunity to work at ICTQT. To Pedro and Tushita, for all your friendship and support since I reached Poland. I am really glad to have had the chance of working with you and to have you as friends. To Amanda, Anubhav, Chitra, Giovanni and Andre, for all your friendship, kindness, and collaboration.

I also express my deep gratitude to the many teachers who have been part of this journey as sources of encouragement and inspiration illuminating my path. To my teachers from Colégio Estadual Anibal Khury Neto: Margherita (Chemistry), Jane (Mathematics), Marleide (Science), Guido (Physics), Leslie (History), Iselci (Geography), Arruda (History), Edilson (Sociology), and Wagner (Arts). To the teachers from Colégio Acesso: André (Physics), Marlus (Literature), Herlon (Geography), Renato (English), and Julio Cezar (History). To all my professors from the Physics undergraduate program at UFPR, especially Professors Alexandre, Renato, Rodrigo Mossanek, Cyro, Marcio Bettega, and Fabio Zanetti. I also take this opportunity to thank the Pró-Reitoria de Pertencimento e Políticas de Permanência Estudantil (formerly PRAE) at UFPR, which was essential for my academic continuity during my undergraduate studies. To all my graduate professors at IFGW, particularly Professors Ricardo Doretto and Marcos Cesar. To the UNICAMP staff for their professionalism, especially all the professionals at Building D and the IFGW Graduate Office, for their support.

I would like to thank the "Ministry of Science, Technology and Innovation" and the “National Council for Scientific and Technological Development – CNPq” (process 140549/2021-8) for their financial support. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) – Finance Code 001", during my research period abroad under the CAPES/PrInt program.

"Eu cheguei de muito longe
E a viagem foi tão longa
E na minha caminhada
Obstáculos na estrada
Mas enfim aqui estou"

– Erasmo Carlos

Abstract

Bell nonlocality, one of the most remarkable features of quantum mechanics, fundamentally challenges the classical notion of *local causality* by revealing strong forms of correlations between space-like separated systems. This phenomenon is not only central to the foundational understanding of quantum theory but also emerged as a critical resource for quantum information science, particularly in areas such as cryptographic protocols. Despite its significance, the quantum theory presents an intrinsic apparent limitation on the strength of its achievable nonlocality, which, currently, is not entirely understood in terms of an operational physical description. The present thesis addresses open questions concerning quantum Bell nonlocality, particularly its role in multipartite systems and its interplay with communication tasks. A key focus is on *device-independent principles*, which aim to characterize quantum correlations by ruling out implausible consequences identified in certain superstrong nonlocal correlations. The principle of *Information Causality* ($\mathcal{IC}$) is central to this investigation, offering a compelling framework for understanding quantum limits on nonlocality. While quantum correlations inherently respect $\mathcal{IC}$, the principle's original formulation has been shown to be insufficient for scenarios involving multiple parties. This work advances the field by employing a systematic framework establishing more suitable operational multipartite formulations for $\mathcal{IC}$. Furthermore, the thesis explores the connection between $\mathcal{IC}$ and the *monogamy of Bell inequality violations*, demonstrating how the multipartite framework naturally recovers this property and ensuring implications for device-independent quantum key distribution. The thesis proceeds with the following structure: (i) reviews the foundational background of Bell nonlocality; (ii) examines its role in quantum communication; (iii) reviews the literature about $\mathcal{IC}$ and introduces the multipartite $\mathcal{IC}$ formulation; and finally (iv) discusses its implications for monogamy relations and cryptographic applications.

Resumo

A não-localidade de Bell, uma das características mais notáveis da mecânica quântica, desafia fundamentalmente a noção clássica de *causalidade local* ao revelar fortes formas de correlações entre sistemas separados espacialmente. Esse fenômeno não é apenas central para a compreensão dos fundamentos da teoria quântica, mas também emergiu como um recurso crucial para a ciência da informação quântica, especialmente em áreas como protocolos criptográficos. Apesar de sua importância, a teoria quântica apresenta uma aparente limitação intrínseca em suas correlações não-locais, que, atualmente, não é completamente compreendida em termos de uma descrição física operacional. A presente tese aborda questões em aberto relacionadas a não-localidade de Bell quântica, em particular seu papel em sistemas multipartidos e sua relação com tarefas de comunicação. Um foco central está nos *princípios independentes de dispositivos*, que buscam caracterizar correlações quânticas ao eliminar consequências implausíveis identificadas em certas correlações não-locais *super-fortes*. O princípio da *Causalidade da Informação* ($\mathcal{IC}$) ocupa uma posição central nesta investigação, oferecendo um quadro promissor para compreender os limites quânticos da não-localidade. Embora as correlações quânticas respeitem intrinsecamente $\mathcal{IC}$, a formulação original do princípio mostrou-se insuficiente para cenários envolvendo múltiplas partes. Este trabalho avança o campo ao empregar uma abordagem sistemática que estabelece formulações operacionais multipartidas mais adequadas para $\mathcal{IC}$. Ademais, a tese explora a conexão entre $\mathcal{IC}$ e a *monogamia de violações de desigualdades de Bell*, demonstrando como a versão multipartida naturalmente recupera essa propriedade, o que resulta em implicações em distribuição de chaves quânticas seguras em um contexto independente de dispositivos. A tese segue a seguinte estrutura: (i) revisa os fundamentos conceituais da não-localidade de Bell; (ii) analisa seu papel em tarefas de comunicação quântica; (iii) revisa a literatura sobre $\mathcal{IC}$ e introduz novas formulações multipartidas de $\mathcal{IC}$; e, por fim, (iv) discute suas implicações para relações de monogamia e aplicações criptográficas.

Contents

Introduction

Quantum mechanics has long been recognized as one of the most profound and successful scientific theories of the 20th century. Its predictive power and experimental confirmation have revolutionized our understanding of the microscopic world. Among the theory's distinctive features is a strong form of correlation, which is attested thanks to the work of John S. Bell in 1964 [1]. More specifically, Bell theoretically demonstrated that under a set of reasonable assumptions, any physical theory relying on the classical hypothesis of *local causality* would fail to reproduce certain statistical correlations predicted by quantum mechanics. Such incompatibility, as quantified by violations of Bell inequalities, is commonly referred to as *Bell nonlocality*, and has been systematically subjected to rigorous experimental scrutiny over the decades [2, 3, 4, 5, 6, 7, 8]. Bell nonlocality is originally related to the intense early foundational debate on the interpretation of the inherently probabilistic nature of quantum theory, such as *Einstein, Podolsky, and Rosen paradox* (EPR) [9, 10]. The significance of these achievements was recognized with the 2022 Nobel Prize in Physics to Alain Aspect, John F. Clauser, and Anton Zeilinger for their pioneering contributions to this field.

Beyond the remarkable contribution to the scientific understanding of nature, Bell nonlocality has been identified as an essential resource underpinning quantum advantages in important information processing tasks to quantum information science. These include applications in security in *cryptographic protocols* [11], and even enhancements in communication performance, such as *communication complexity problems* [12]. These advancements have driven scientists to explore whether nonlocality could serve as a *fundamental axiom of quantum mechanics*. In fact, despite the recognized success of quantum theory, since its early years, some physicists, such as John von Neumann [13], have been committed to identifying what the primitive physical notions from which such mathematical formalism emerge[1] [14]. Thus, is Bell nonlocality the defining characteristic of quantum mechanics?

In 1994 [15], Sandu Popescu and Daniel Rohrlich have theoretically demonstrated a hypothetical even stronger form of correlations, fulfilling the requirements of the Bell-type experiment but exceeding the limits of quantum correlations. The latter is captured by the so-called Tsirelson's bound on quantum violations of Bell inequalities [16]. The result negatively answers the original question; however, their pioneering work offers an alternative framework that embeds the quest for primitive physical notions of quantum theory. *i.e., do all nonlocal correlations beyond the quantum boundary have a correspondence in nature?* In this sense, whether coherent reasons exist for certain nonlocal correlations remaining unobserved in experiments, this fact would naturally provide new insights regarding the laws of quantum mechanics by describing the quantum nonlocality boundary. Indeed, the close significance of Bell nonlocality in information processing tasks was shown to lead to *implausible consequences* for such a superstrong form of nonlocality [17]. This question defines a promising research direction, often referred to as *device-independent principles*[2]. Some examples are *Nontrivial communication complexity* [17, 18], *Macroscopic locality* [19], and *Information causality* [20].

The principle of *Information causality* ($\mathcal{IC}$) stands out as one of the most compelling

[1] As it happens with Special Relativity, for example.

[2] The name will be clarified in the main text

candidates for explaining the quantum boundary on nonlocality. Introduced in [20], the principle proposes a suggestive limitation on communication scenarios involving a sender and a receiver. Essentially, $\mathcal{IC}$ asserts that the accessible information to a receiver cannot exceed the information effectively transmitted by the sender. While quantum correlations inherently comply with $\mathcal{IC}$, the principle, notably, rules out all of those nonlocal correlations surpassing the Tsirelson bound for maximal quantum violations of Bell inequalities [20]. Nonetheless, whether $\mathcal{IC}$ excludes all stronger-than-quantum correlations remains unclear [21], mainly for the particular challenge of properly defining $\mathcal{IC}$ operationally. Subsequent refinements [22, 23] have improved precision but were shown insufficient to characterize quantum correlations fully. Recent advancements, in particular, suggest that quantum correlations require multipartite formulated principles [24]. This fact is also supported by later developments demonstrating the limitation of the original bipartite formulation for $\mathcal{IC}$ in excluding certain superstrong nonlocal correlations involving three parties systems [25].

The present thesis builds upon this investigative framework, addressing key challenges characterizing quantum correlations through device-independent principles. Specifically, we first investigate the role of Bell nonlocality as a resource for information processing and address more suitable approaches to witness the strength of nonlocal correlations in communication instances, such as the *Random Access Codes* [26, 27, 28]. In light of this, we extend the study of the implausible consequences of certain nonlocal correlations in communication scenarios involving multiple parties, using the lens of the Information Causality principle [29]. We employ a general geometric information-theoretical framework [22] and systematically investigate novel multipartite operational descriptions for $\mathcal{IC}$ statement. On the other hand, while it remains open whether $\mathcal{IC}$ may fully single out the set of quantum correlations, mainly because of the high computational complexity, we alternatively proceed to investigate the interplay of properties ensured by $\mathcal{IC}$ and the ones observed in quantum mechanics. In this direction, we demonstrate that $\mathcal{IC}$ recovers the strong form of the so-called *monogamy of Bell inequalities violations* [30]. *i.e.*, we have demonstrated that the introduced operational forms identifying $\mathcal{IC}$-statement ensure that two parties achieving the maximum quantum violation cannot establish any correlation with any other possible extra part. Interestingly, the solid cryptographic significance of such monogamous relation enables stating a clear connection of $\mathcal{IC}$ principle and secrecy for cryptographic keys without addressing the mathematical formalism of quantum theory. More importantly, in this case, $\mathcal{IC}$ ensures security of cryptographic quantum key distribution, even when considering hypothetical eavesdropping that could potentially breakthrough the laws of quantum mechanics [30].

The thesis is organized as follows:

Chapter 1 introduces the crucial background of the thesis. It provides a conceptual overview of Bell's nonlocality in quantum information science. It discusses critical milestones in theoretical investigations, culminating in recognizing Bell nonlocality's key role in understanding quantum theory predictions.

Chapter 2 discusses Bell's nonlocality within the communication paradigm, focusing on its utility as a resource in different information processing tasks. This Chapter reviews the device-independent framework, emphasizing its significance in cryptographic protocols, random access codes, and communication complexity problems. Additionally, the Chapter addresses the potential of supra-quantum nonlocal correlations leading to physical implausibilities.

Chapter 3 delves into the information causality principle. The Chapter reviews the original formulation of $\mathcal{IC}$ and its subsequent refinements, addressing its successes and limitations. It also introduces a novel multipartite formulation of $\mathcal{IC}$, leveraging advanced geometric tools such as *Shannon's entropic cone*, and examines its potential in excluding certain supra-quantum correlations, contrasting with the original bipartite frameworks.

Chapter 4 establishes a connection between $\mathcal{IC}$ and the monogamy of Bell inequality violations. This Chapter first addresses some misunderstandings in the literature, demonstrating that the original bipartite framework cannot recover such monogamy relations. Then, it shows how the novel multipartite framework reveals monogamy relations. It also discusses the implications of these findings for practical applications, such as secure device-independent quantum key distribution.

In summary, the thesis addresses open questions on the understanding of quantum correlations within the Bell nonlocality paradigm, laying the groundwork for future advancements in more complex scenarios. The chapters are followed by the appendix section, where some concepts and theoretical proofs are better developed.

Chapter 1

Preliminaries

Although quantum theory is the most accurate physical theory ever developed for measuring physical quantities, understanding why the quantum world behaves as it does continues to drive numerous scientific investigations. In this context, a relevant question arises: Is it possible to establish quantum theory through fundamental physical postulates, similar to the approach used in special relativity? Answering this question is notably challenging. Specifically, might these mathematical objects, such as *Hilbert spaces* or *quantum states*, be interpreted merely as a predictive tool within the quantum theory, or do they represent a physical attribute of the object it describes? In this context, it becomes convenient to conduct this analysis from a generalized standpoint, avoiding any assumptions about the system. This approach, known as *device-independent* (DI) analysis, is further elaborated upon in [31, 32, 33].

The DI framework identifies physical properties exclusively from experimental data. Thus, an experiment can be seen as a *black box*, as receiving an input, x, representing the choice of measurement available to an observer in the lab, and returning a result, a, corresponding to possible experimental results. In such a setting, the most comprehensive description is achieved through probability distributions, specifically the conditional probabilities of possible results a given the possible measurements x, denoted as $p(a|x)$. Thus, in a scenario with m possible measurements, each yielding r potential outcomes, a complete description must encompass all probability distributions of potential outcomes for each measurement choice, $p(a|x)$. These probabilities can then be organized into a vector, where each distribution corresponds to a component within a vector space of dimension $d = rm$,

$$\mathbf{p} = \Big(p(a_1|x_1), p(a_2|x_1), \ldots, p(a_r|x_1), p(a_1|x_2), \ldots, p(a_r|x_m)\Big) \in \mathbb{R}^d. \tag{1.1}$$

We refer to the vector $\mathbf{p}$ as the *behavior* of the black box in Fig.1.

More generally, we are interested in the possibly observed statistics in experiments that parallelly happen in different spatially separated laboratories. In the device-independent framework, each of the n laboratories is represented by distinct black

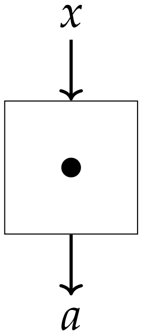


Figure 1 – Black box.

boxes (Fig.1), and the scenario is frequently specified by the tuple (n, m, r). The vector, $\mathbf{p}$, specifies the experiment whose components are joint probability distributions of the outcomes, conditional on each possible measurement choice in each laboratory, $p(a, b, \ldots | x, y, \ldots)$. Fig.2 illustrates the case $n = 2$. For simplicity, we will discuss the forthcoming concepts considering the bipartite case; however, these ideas can be readily generalized to any number of parties n. Once these are probability distributions, the components of $\mathbf{p}$ must satisfy the normalization conditions $\sum\limits_{a,b} p(a, b|x, y) = 1$ and non-negativity $p(a, b|x, y) \geqslant 0$. Geometrically, these constraints define a region in $\mathbb{R}^d$ to which $\mathbf{p}$ belongs, defined by the set of behaviors $\mathbf{p}$ that meet the probabilistic requirements,

$$\mathcal{P} = \left\{ \mathbf{p} \in \mathbb{R}^d \mid p(a, b|x, y) \geqslant 0 \; \forall a, b, x, y; \; \sum_{a,b} p(a, b|x, y) = 1 \; \forall x, y \right\}. \tag{1.2}$$

In this context, mentioning some terminologies from convex geometry becomes helpful. The appendix A briefly introduces some fundamental concepts that are part of the DI framework vocabulary. For further details, we address *Ref.* [34].

Historically, we say that operations in the first laboratory are performed by Alice and in the second by Bob. In general, when parties perform such an experiment, the probability distribution $p(a, b|x, y)$ generally cannot be obtained through the marginal descriptions of Alice and Bob,

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

This implies that their experiment outcomes are not statistically independent of each other. In fact, the most general marginal description that each laboratory can individually write is of the form,

$$p_A(a|x, y) = \sum_b p(a, b|x, y), \tag{1.3a}$$

$$p_B(b|x, y) = \sum_a p(a, b|x, y), \tag{1.3b}$$

It evaluates the probability of obtaining a result of a given the choice of measurement

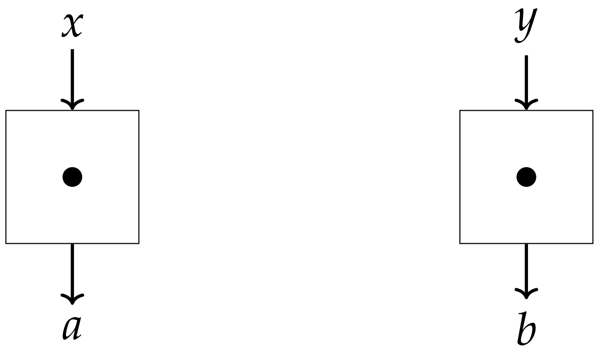


Figure 2 – Bipartite Bell scenario.

of x while also accounting for a possible choice made by Bob in his laboratory. It can be achieved through signaling from Bob to Alice, similarly to $p_B(b|x,y)$. Interestingly, however, if parties may freely communicate along the experiment, they can produce any statistics belonging to the set of correlations $\mathcal{P}$, independently of the physical systems that the parties manipulate. It can be easily seen from Bayes' rule in the form $p(a,b|x,y) = p(a|x,y)p(b|x,y,a)$. Consequently, in this context, we require that such experiments be arranged so that no information leakage happens from one laboratory to another. Such requirement geometrically constraints the set of possible correlations $\mathcal{P}$, which defines the set of *nonsignaling correlations*. Scenarios where no-communicating laboratories conduct parallel experiments are commonly referred to as *Bell-type experiments* or simply *Bell scenarios* due to the seminal work of John S. Bell in 1964 [1]. A Bell scenario is defined by the number of parties n, each capable of performing m possible measurements, for each of which r possible outcomes can be obtained.

1.1 No-signaling correlations

The core idea behind the device-independent approach is to identify the possible statistics observed in a Bell experiment, depending on different physical systems that parties can manipulate in their laboratory. It is convenient, therefore, to assume that their experiment arrangement prevents any possible signaling among the labs. Consequently, any observed correlation in the experiment results cannot yield from the possible communication from one laboratory to another. This crucial extra assumption is called the *nonsignaling* principle. Physically, it might be ensured when the systems operated by Alice and Bob in Fig.2 are spacelike separated, preventing communication between the two parties, at least until the experiment concludes. This argument can be further formalized using special relativity theory. A *measurement event* in a given laboratory is represented by $x|a$ ($y|b$) and defined as Alice's (Bob's) action that begins with the choice of measurement x (y) and ends with obtaining outcome a (b). Two measurement events, $x|a$ and $y|b$, are said to be *spacelike separated* if neither event intersects the light cone of the other. A sketch of the spacetime diagram for this case is shown in Fig.3. Under this assumption, there is no possibility of subluminal communication between Alice and Bob during the experiment, ensuring that the observed correlations do not arise from any information exchange between the parties.

This nonsignaling assumption restricts the black boxes' behavior $\mathbf{p}$. Alice's marginal description in Eq.(1.3), for example, under nonsignaling constraint, cannot depend on the choices made by Bob in his laboratory and vice versa. Thus, the *non-signaling* principle translates as the following constraints on the marginal descriptions:

$$p_A(a|x,y) = p_A(a|x), \quad \forall\, x,y,a; \tag{1.4a}$$

$$p_B(b|x,y) = p_B(b|y), \quad \forall\, x,y,b. \tag{1.4b}$$

Consequently, a behavior $\mathbf{p}$ in a Bell scenario with $n \geqslant 2$ is said to be *non-signaling* when respecting the constraints (1.4). In terms of the components of $\mathbf{p}$, the nonsignaling

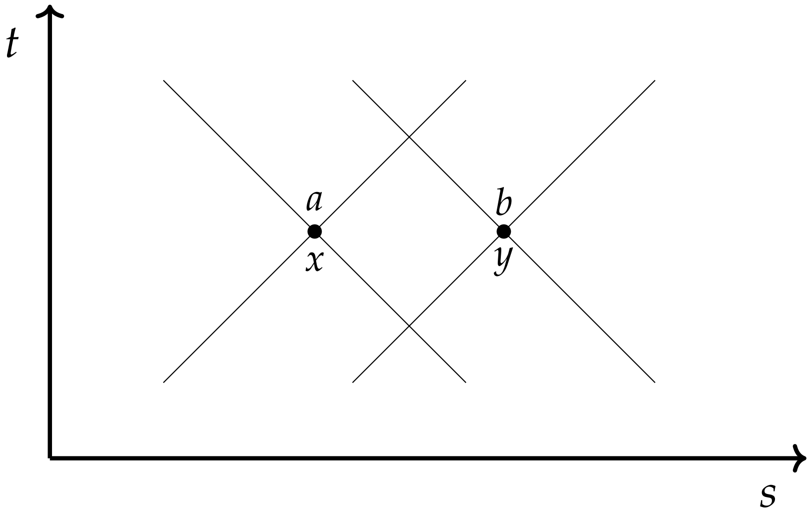


Figure 3 – Spacelike separated measurement events of Alice and Bob.

constraints are read as:

$$\sum_b p(a,b|x,y) = \sum_b p(a,b|x,y'), \quad \forall\, a,x,y,y'; \tag{1.5a}$$

$$\sum_a p(a,b|x,y) = \sum_a p(a,b|x',y), \quad \forall\, b,y,x,x'. \tag{1.5b}$$

Geometrically, the non-signaling constraints define a region for $\mathbf{p}$ in $\mathbb{R}^d$, given by

$$\mathcal{P}_{NS} = \left\{ \mathbf{p} \in \mathcal{P} \;\middle|\; \begin{aligned} &\sum_b p(a,b|x,y) = \sum_b p(a,b|x,y'), \quad \forall\, a,x,y,y', \\ &\sum_a p(a,b|x,y) = \sum_a p(a,b|x',y), \quad \forall\, b,y,x,x' \end{aligned} \right\}. \tag{1.6}$$

It is straightforward to show that the set $\mathcal{P}_{NS}$ is convex. For two behaviors $\mathbf{p}_1, \mathbf{p}_2 \in \mathcal{P}_{NS}$, with their convex combination $\mathbf{p} = \alpha\mathbf{p}_1 + (1-\alpha)\mathbf{p}_2$, marginalizing over a yields,

$$\begin{aligned} \sum_a p(a,b|x,y) &= \alpha \sum_a p_1(a,b|x,y) + (1-\alpha) \sum_a p_2(a,b|x,y); \\ &= \alpha p_{1_B}(b|y) + (1-\alpha) p_{2_B}(b|y) = p_B(b|y). \end{aligned}$$

Marginalizing over b gives $p_A(a|x)$. Thus, $\mathbf{p} \in \mathcal{P}_{NS}$, confirming that the set (1.6) is convex. Moreover, by definition (A.7), $\mathcal{P}_{NS}$ is a polytope, commonly referred to as the *non-signaling polytope*.

1.2 Local correlations

By arranging space-like separated laboratories in a Bell experiment, the parties physically ensure their possibly observed correlations cannot yield from communication among the labs. However, there still might be the case that such correlation is a result

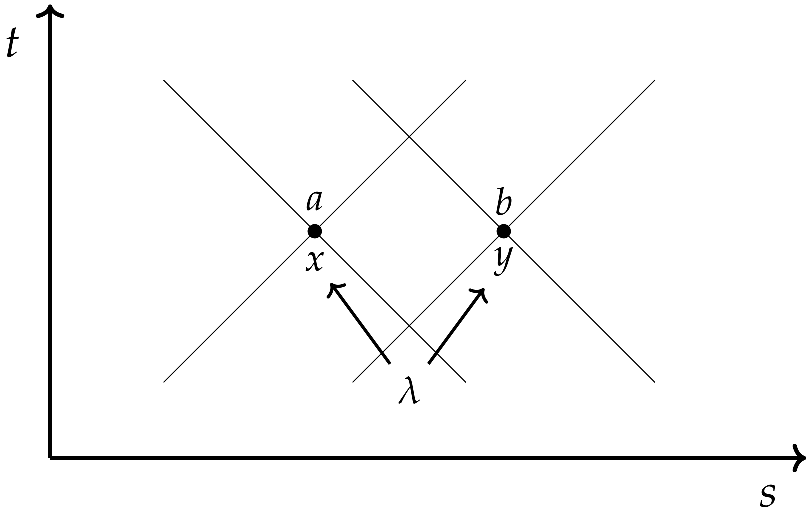


Figure 4 – Representation of a common hidden variable for Alice and Bob's measurement events.

of unobserved factors that somehow affect and correlate the experiments. Consider a scenario where Alice and Bob each toss a coin within their respective laboratories as an illustrative example. In this setup, the experiment begins with a machine that produces pairs of coins and distributes one to each participant. The parties have no prior knowledge regarding the machine's operation. However, they observe a strong correlation in their outcomes after tossing the coins across multiple rounds. In every round, their results are identical, such that they consistently record (coin$_A$, coin$_B$) = (heads, heads) or (tails, tails). Upon meticulous analysis of their devices, Alice and Bob discover that the coins are biased. Moreover, they identify that the machine consistently outputs pairs of coins with identical biases —half the time biased towards heads and the other half towards tails. Consequently, the apparent correlation is fully explained when the machine's behavior is taken into account. Thus, it becomes crucial to identify whether the performed experiment is possibly affected by aspects that are not initially accounted for or are even unobservable. Once their experimental arrangement ensures nonsignaling among the labs, any possible *causally local* explanation for the observed correlation might be condensed in terms of an extra variable λ. Back to the space-time diagram, we may understand *causally local* explanations as those following a sequence of causal factors with no superluminal signals. Consequently, λ relies on the common light cone of both experiments (Fig.4). Thus, whether all pre-factors of such joint observation can be recognized in λ, the joint description may factorize as,

$$p(a,b|x,y,\lambda) = p_A(a|x,\lambda)p_B(b|y,\lambda). \tag{1.7}$$

In this sense, it means that variable λ identifies the reason for observing the experimental result a, and any residual indeterminacies concerning the results have no relation with what happens in Bob's laboratory, such that the joint distribution factorizes. Consequently, whether there exists a probability distribution $p(a,b,\lambda|x,y)$, fulfilling the local factorization condition (1.7), from which the observed experimental statistics $p(a,b|x,y)$ can be obtained as

$$p(a,b|x,y) = \int_\Lambda p(\lambda)p_A(a|x,\lambda)p_B(b|y,\lambda)d\lambda, \tag{1.8}$$

we say that $p(a,b|x,y)$ is compatible with common past pre-stated correlation model, and it is said a *local correlation*. Historically, in this case, $p(a,b|x,y)$ is said as admitting a *local-hidden-variable* model (1.8). Naturally, experiments specified by $p(a,b|x,y)$, which do not admit the local model (1.8) are said *nonlocal correlations*. Notice that the local model (1.8) implicitly accounts that Alice and Bob can freely make their measurement choices, *i.e.*, $p(x,y|\lambda) = p(x,y)$, which equivalently means $p(\lambda|x,y) = p(\lambda)$.

Geometrically, the locality hypothesis in (1.8) constraints the permitted region for $\mathbf{p}$ in $\mathbb{R}^d$. Consider, for example the simplest case where $a,b,x,y \in \{0,1\}$ and the following quantity is considered,

$$E(x,y) = p(a = b|x,y) - p(a \neq b|x,y). \tag{1.9}$$

If a given behavior admits a hidden variable model, the probabilities in (1.9) can be decomposed as in (1.8),

$$p(a = b|x,y) = \int_\lambda p(\lambda)[p_A(a = 0|x,\lambda)p_B(b = 0|y,\lambda) + p_A(a = 1|x,\lambda)p_B(b = 1|y,\lambda)]d\lambda; \tag{1.10a}$$

$$p(a \neq b|x,y) = \int_\lambda p(\lambda)[p_A(a = 0|x,\lambda)p_B(b = 1|y,\lambda) + p_A(a = 1|x,\lambda)p_B(b = 0|y,\lambda)]d\lambda. \tag{1.10b}$$

Since $p(a = b|x,y) = p(0,0|x,y) + p(1,1|x,y)$ and $p(a \neq b|x,y) = p(1,0|x,y) + p(0,1|x,y)$, we consider the following quantity:

$$E(0,0) + E(0,1) + E(1,0) - E(1,1) = \beta_{CHSH}. \tag{1.11}$$

Following (1.9) and (1.10), we have:

$$\begin{aligned}&\beta_{CHSH} = \\ &2\int_\Lambda p(\lambda)\Big([2p_A(a = 0|x = 0,\lambda) - 1][p_B(b = 0|y = 0,\lambda) + p_B(b = 0|y = 1,\lambda) - 1] + \\ &\qquad + [p_A(a = 0|x = 1,\lambda) - 1][p_B(b = 0|y = 0,\lambda) - p_B(b = 0|y = 1,\lambda)]\Big).\end{aligned} \tag{1.12}$$

Since all probabilistic terms fall within the interval $[0,1]$, the terms in the brackets can vary between $[-1,1]$. Consequently, the maximum value the expression in parentheses can attain is 1, implying that the expression in (1.11) is bounded above:

$$\beta_{CHSH} \leqslant 2. \tag{1.13}$$

This inequality is known as the *CHSH inequality* due to the seminal work [2]. It entails a necessary and sufficient condition for a given behavior $\mathbf{p}$ to satisfy the locality condition in (1.8). Each Bell scenario entails a finite number of inequalities that must be satisfied for a given behavior to be considered local. Such inequalities are known as *Bell inequalities* and define the region of local behaviors in $\mathbb{R}^d$:

$$\mathcal{P}_L = \left\{ \mathbf{p} \in \mathcal{P} \mid \mathbf{a}_i^T \mathbf{p} \leqslant \beta_i \; i = 1,2,...,n \right\}. \tag{1.14}$$

In this case, $\mathbf{a}_i^T \mathbf{p} \leqslant \beta_i$ represents the possible Bell inequalities within this scenario, which is specified by the vector of real coeficients, $\mathbf{a}_i^T$, and bounds, β_i. By definition (1.8), the set of local behaviors $\mathcal{P}_L$ is convex, and as (A.7), it is also a convex polytope, referred to as the *local polytope*.

It is straightforward to show that local behaviors naturally address the nonsignaling condition (1.5). In fact, marginalizing (1.8) over b and using the normalization condition, we obtain (1.4a),

$$\begin{aligned} p_A(a|x,y) &= \sum_b \int_\Lambda p(\lambda) p_A(a|x,\lambda) p_B(b|y,\lambda) d\lambda \\ &= \int_\Lambda p(\lambda) p_A(a|x,\lambda) \left(\sum_b p_B(b|y,\lambda) \right) d\lambda \\ &= \int_\Lambda p(\lambda) p_A(a|x,\lambda) d\lambda = p_A(a|x). \end{aligned} \tag{1.15}$$

Similarly, marginalizing (1.8) over a yields (1.4b). In this context, Alice and Bob's marginal descriptions of their outputs a and b depend only on their respective choices x and y, as in (1.4). On the other hand, the converse is not true. As it should become clear in the next subsection, there are non-signaling correlations that do not admit a local model as (1.8). Thus, it becomes evident that the set of local behaviors is contained within the non-signaling polytope, *i.e.*, $\mathcal{P}_L \subset \mathcal{P}_{NS}$.

1.3 Polytope (2,2,2)

As detailed in (A.8), a polytope can alternatively be defined as the convex hull of a finite number of points. In this approach, a polytope $\mathcal{P}$ can be fully characterized by its extremal points[1], since any point within $\mathcal{P}$ can be obtained as a convex combination of these. Particularly for subsequent discussions, non-signaling extreme behaviors within the Bell scenario $(2,2,2)$ are of particular interest.

Given the description of a polytope in terms of linear constraints, as in (1.6) and (1.14), existing algorithms such as *PANDA* [35] can identify all vertices of a polytope. For the bipartite, non-signaling polytope in (1.6) with two inputs and two outputs, there are 24 vertices. Among these, 16 correspond to deterministic local behaviors that are also extreme points of the local polytope in (1.14), and the remaining 8 correspond to nonlocal behaviors. These behaviors are comprehensively characterized in [36]. The 16 local vertices are those for which the components can be expressed as:

$$p_L^{\mu\nu\sigma\tau}(a,b|x,y) = \begin{cases} 1 & \text{if } a = \mu x \oplus \nu,\ b = \sigma y \oplus \tau, \\ 0 & \text{otherwise,} \end{cases} \tag{1.16}$$

where $\mu, \nu, \sigma, \tau \in \{0,1\}$. The 8 nonlocal extreme points are those for which the components can be written as follows:

[1] That is, the vertices of the polytope.

$$p_{NL}^{\mu\nu\sigma}(a,b|x,y) = \begin{cases} 1/2 & \text{if } a \oplus b = x \cdot y \oplus \mu x \oplus \nu y \oplus \sigma, \\ 0 & \text{otherwise,} \end{cases} \tag{1.17}$$

where $\mu, \nu, \sigma \in \{0,1\}$. These are referred to as *PR boxes* after the work of Sandu Popescu and Daniel Rohrlich [15]. The violation of the CHSH inequality might immediately witness their nonlocal nature (1.13), which achieves the value $\beta_{CHSH} = 4$.

1.4 Quantum correlations

As detailed in Appendix B, in quantum mechanics, given a state ρ and a set of POVM elements $\{M_{a|x}\}$ associated with the results a of a measurement process, the probability of obtaining an outcome a given measurement x is determined by the *Born rule*:

$$p(a|x) = \text{Tr}\left(\rho M_{a|x}\right). \tag{1.18}$$

When considering Bell scenarios, the joint probability distributions $p(a,b|x,y)$ for the bipartite case reads as:

$$p(a,b|x,y) = \text{Tr}\left(\rho M_{a|x} \otimes M_{b|y}\right). \tag{1.19}$$

Here, ρ represents a composite quantum state shared between the parties, where each party measurement procedure is described in terms of the measurement operator set, $\{M_{a|x}\}$, for each result a when Alice measures x, and $\{M_{b|y}\}$ for each result b when Bob measures y. Consequently, $M_{a|x} \otimes M_{b|y}$ forms the joint measurement elements, which, as we have discussed in the Appendix B, can be projective measurements or POVMs. The behaviors, $\mathbf{p}$, with components, $p(a,b|x,y)$, that can be described via Born's rule (1.19), are referred to as *quantum behaviors*. In this sense, there exists a quantum state, ρ, and measurement operators, $M_{a|x}$, and $M_{b|y}$, such that Born's rule recovers the probability $p(a,b|x,y)$. Thus, Born's rule delineates a region in $\mathbb{R}^d$ for behaviors achievable by quantum theory, namely the *set of quantum correlations*, $\mathcal{P}_Q$. The set $\mathcal{P}_Q$ is convex [37, 38, 39]; however, it is not a polytope, as it cannot be characterized by the convex hull of a finite set of points [40].

By marginalizing equation (1.19), one can recover equation (1.4), since:

$$\begin{aligned} p_A(a|x,y) &= \sum_b p(a,b|x,y) = \sum_b \text{Tr}\left(\rho M_{a|x} \otimes M_{b|y}\right) \\ &= \text{Tr}\left(\rho M_{a|x} \otimes \mathbb{1}\right) = p_A(a|x); \end{aligned} \tag{1.20a}$$

$$\begin{aligned} p_B(b|x,y) &= \sum_a p(a,b|x,y) = \sum_a \text{Tr}\left(\rho M_{a|x} \otimes M_{b|y}\right) \\ &= \text{Tr}\left(\rho \mathbb{1} \otimes M_{b|y}\right) = p_B(b|y). \end{aligned} \tag{1.20b}$$

Consequently, quantum correlations in Bell experiments are also non-signaling. However, fulfilling the non-signaling condition alone does not suffice to classify a behavior as quantum; as shown in [15], there exist non-signaling correlations that quantum theory cannot reproduce according to Born's rule (1.19), *i.e.*, $\mathcal{P}_Q \subset \mathcal{P}_{NS}$. More importantly, however, as established by the seminal *Bell's theorem* in 1964 [1], there are quantum correlations as (1.19) that no local model as (1.8) can replicate. This fact becomes evident in the bipartite Bell scenario with dichotomic inputs and outputs. In the quantum case, Alice and Bob share a bipartite quantum system, upon which they each perform projective measurements associated to dichotomic observables namely A_0, A_1 for Alice and B_0, B_1 for Bob. With a and b as the eigenvalues of Alice's and Bob's measurement operators, the expected value of the joint measurement operator can be expressed in terms of the bipartite behavior as follows:

$$\langle A_x \otimes B_x \rangle = \sum_{a,b} ab\, p(a,b|x,y). \tag{1.21}$$

By choosing operators A_x and B_y with spectrum $\{-1,1\}$, we obtain:

$$\langle A_x \otimes B_x \rangle = p(-1,-1|x,y) + p(1,1|x,y) - p(1,-1|x,y) - p(-1,1|x,y). \tag{1.22}$$

Consequently, $\langle A_x \otimes B_x \rangle = E(x,y)$ in (1.9), yielding the *CHSH inequality* in terms of operators A_x and B_y:

$$\langle A_0 \otimes B_0 \rangle + \langle A_0 \otimes B_1 \rangle + \langle A_1 \otimes B_0 \rangle - \langle A_1 \otimes B_1 \rangle = \beta_{CHSH} \leqslant 2. \tag{1.23}$$

Since the CHSH inequality provides a necessary and sufficient condition for a behavior to satisfy the locality assumption in (1.8), a violation of this criterion implies that the behavior is necessarily non-local. Suppose Alice and Bob share the following quantum state:

$$|\phi^+\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}, \tag{1.24}$$

and measurements are performed using the following observables:

$$\begin{aligned} A_0 &= \sigma_x, & A_1 &= \sigma_z; \\ B_0 &= \frac{\sigma_x + \sigma_z}{\sqrt{2}}, & B_1 &= \frac{\sigma_x - \sigma_z}{\sqrt{2}}. \end{aligned} \tag{1.25}$$

By evaluating the expected values in the inequality (1.23), we obtain:

$$\beta_{CHSH} = 2\sqrt{2}. \tag{1.26}$$

That is referred to as Tsirelson's bound, which is the maximum achievable value for quantum mechanics of the CHSH inequality [16]. Therefore, this behavior derived from quantum theory violates Bell's inequality (1.23) and thus cannot be explained by any local model (1.8). Consequently, we may write the following hierarchy among the sets of correlations $\mathcal{P}_L \subset \mathcal{P}_Q \subset \mathcal{P}_{NS}$, which we depict in Fig.5. It is important to stress the close connection between the Bell nonlocality phenomenon and *entanglement* of quantum states. In fact, *separable states*, (B.7), in the form $\rho_{AB} = \sum_i p(i)\rho_A^i \otimes \rho_B^i$ are always compatible with the local model (1.8). Hence, entanglement becomes necessary for violating Bell's inequalities through quantum mechanics. On the other hand, entanglement is not a sufficient condition to observe Bell inequality violations. *i.e.*, despite any *pure* entangled state violating a Bell inequality [41, 42], there are *mixed* entangled states

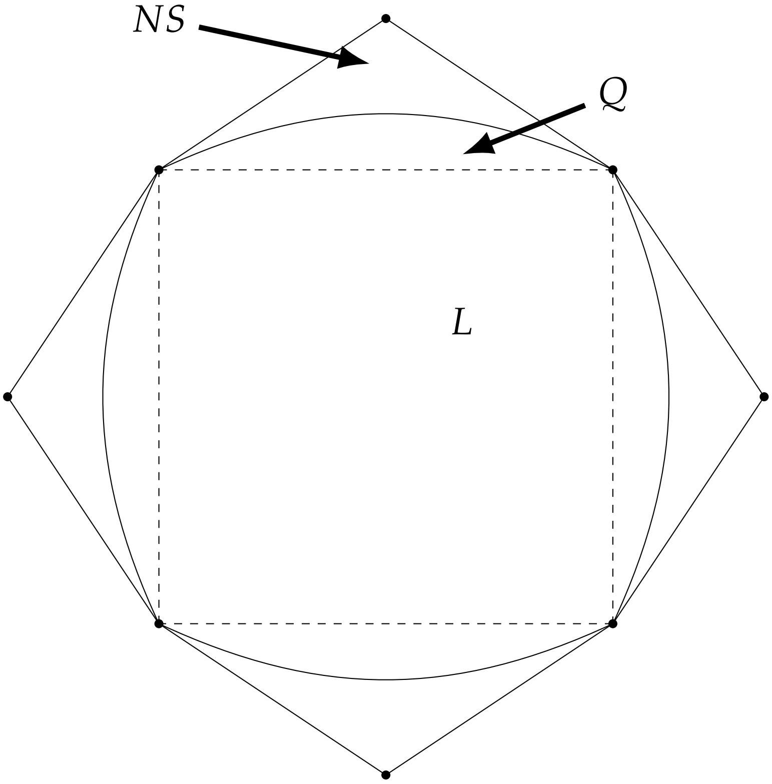


Figure 5 – A pictorial representation illustrating the hierarchy among the sets of nonsignaling, quantum, and local behaviors.

which do not violate any Bell inequality [42].

1.5 Multiple parties scenarios

As previously mentioned, the notions of nonsignaling (1.5), quantum (1.19), and local (1.19) correlations for multipartite scenarios follow a straight generalization of the presented bipartite definitions. For systematic studies on multipartite scenarios we address to the *Ref.* [43, 44]. In fact, the local model (1.8) may be written as,

$$p(a,b,\cdots,c|x,y,\cdots,z) = \int_{\Lambda} p(\lambda)p_A(a|x,\lambda)p_B(b|y,\lambda)\cdots p_C(c|z,\lambda)d\lambda. \qquad (1.27)$$

Nevertheless, the geometrical structure of the set of correlations in this case exhibits a significant increase of complexity even for the most straightforward cases [45, 46]. In fact, the Bell scenario with three parties and dichotomic inputs and outputs, for example, presents 53856 Bell inequalities bounding the local set, while the bipartite equivalent exhibits only 24. Because of the symmetry under the possible relabelings, they can be separated into 46 equivalence classes. It included the previous CHSH inequality (1.13) from the simpler scenario, but also new features as inequalities never violated with quantum correlations [47]. The presence of CHSH in this bigger scenario stresses the fact that even when only two of the three parties are nonlocally correlated, and the third remaining one completely uncorrelated, the produced statistics will be incompatible with the multipartite local model (1.27). This observation motivated the introduction of the *genuinely multipartite nonlocality* notion [48], which addresses the problem of identifying nonlocality, which is a result of correlations including all parties. In the case of three parties, we say the correlation $p(a,b,c|x,y,z)$ is *genuinely multipartite nonlocal* whether it does not admit a model as,

$$\begin{aligned} p(a,b,c|x,y,z) = q_1 \int & p\left(\lambda_1\right) p_A\left(a|x,\lambda_1\right) p_{BC}\left(b,c|y,z,\lambda_1\right) d\lambda_1 \\ + q_2 \int & p\left(\lambda_2\right) p_B\left(b|y,\lambda_2\right) p_{AC}\left(a,c|x,z,\lambda_2\right) d\lambda_2 \\ + \cdots & \\ + q_3 \int & p\left(\lambda_3\right) p_C\left(c|z,\lambda_3\right) p_{AB}\left(a,b|x,y,\lambda_3\right) d\lambda_3. \end{aligned} \tag{1.28}$$

One example of Bell inequality witnessing nonlocal feature according to this definition is the commonly referred to *Svetlichny's inequality* [48], which is written for three parties with dichotomic inputs and outputs scenario as

$$S_3 \equiv \sum_{xyz} (-1)^{xy+xz+yz} E(x,y,z) \leqslant 4,$$

where $E(x,y,z) = \sum_{a,b,c} (-1)^{a+b+c} p(a,b,c|x,y,z)$. The maximum quantum violation in this case is $S_3 = 4\sqrt{2}$, which might be achieved with the *GHZ* state [49] for three qubits, with suitable measurements,

$$|\mathrm{GHZ}\rangle_3 = \frac{1}{\sqrt{2}}\left(|000\rangle + |111\rangle\right). \tag{1.29}$$

Naturally, the set of nonsignaling correlations also follows such an increase in complexity. The tripartite dichotomic scenario is fully characterized by 53856 extremal points, which are grouped in 46 classes of equivalence, each of them respectively associated with the maximum permitted nonsignaling value of each of the Bell inequalities of the local polytope [44]. Three of them present genuinely tripartite nonlocality by violating Svetlichny's inequality (1.29), and are of the form:

$$p_{XYZ}(a,b,c|x,y,z) = \begin{cases} 1/4 & \text{if} \quad a \oplus b \oplus c = xyz; \\ 0 & \text{else;} \end{cases} \tag{1.30}$$

$$p_{(X+Y)Z}(a,b,c|x,y,z) = \begin{cases} 1/4 & \text{if} \quad a \oplus b \oplus c = xz \oplus yz; \\ 0 & \text{else;} \end{cases} \tag{1.31}$$

$$p_{XY+YZ+XZ}(a,b,c|x,y,z) = \begin{cases} 1/4 & \text{if} \quad a \oplus b \oplus c = xy \oplus yz \oplus xz; \\ 0 & \text{else;} \end{cases} \tag{1.32}$$

The extremal (1.32) achieves the maximum nonsignaling value for (1.29) of $S_3 = 8$. Despite the evident richer structure of the multiple parties scenarios, most nonlocal extremal points may also be understood in terms of the bipartite PR box correlation (1.17). The correlations (1.30), (1.31), and (1.31) can always be simulated when parties can share copies of PR boxes among the parties [36]. In fact, the same reference shows that in any bipartite scenario, all extremal nonsignaling correlations can be achieved by sufficient copies of PR boxes and local operations. Such property also follows a wide range of correlations in multipartite scenarios. However, there are those correlations that are not accomplished with such strategy [50]. In the tripartite dichotomic scenario, for example, one of the 46 classes may not be simulated in terms of PR boxes, which we may write as the correlation $p(a, b, c|x, y, z)$ fulfilling the property [44],

$$\begin{aligned} a_0 \oplus b_1 &= 0, \\ b_0 \oplus c_1 &= 0, \\ c_0 \oplus a_1 &= 0, \\ a_0 \oplus b_0 \oplus c_0 &= 0, \\ a_1 \oplus b_1 \oplus c_1 &= 1. \end{aligned} \tag{1.33}$$

In the subsequent Chapters, we shall discuss how these multipartite correlations are closely related to enhancement in communication task performance when nonlocal correlations assist parties.

1.6 Discussion

Bell nonlocality is historically closely related to the search of physicists for a better understanding of the intrinsic probabilistic feature of quantum mechanics. In the earliest, while some people argued in favor of some *incompleteness* on the theory [9], which could explain why we are not able to predict with certainty the result of quantum experiments, others already suggested that quantum mechanics expresses intrinsic characteristics of certain physical systems. *i.e.*, there exist indeterminacy in nature [10]. The long debate could only be discussed in terms of experimental scrutiny after the work of John S. Bell in 1964 [1]. In fact, under some reasonable set of assumptions, the so-called *Bell's theorem*, (1.26), shows that quantum theory predicts a strong form of correlation, which the statistics could never be explained in terms of extra *local-hidden-variables* as (1.8). Experimental arrangements specifically addressing Bell's theorem were further developed [2, 3], followed by even more sophisticated setups addressing certain experimental *loopholes* [4, 5]. The first *loophole-free* Bell test was only recently demonstrated in 2015 with the series of papers [6, 7, 8]. It is important to mention that the main characters, Alain Aspect, John F. Clauser, and Anton Zeilinger, involved with such experimental demonstration, received The Nobel Prize in Physics in 2022 "for experiments with entangled photons, establishing the violation of Bell inequalities and pioneering quantum information science."

As we shall discuss in the next Chapter, however, beyond the foundational discussion around the interpretation of quantum theory predictions, later developments demonstrated that Bell nonlocality is closely related to the advantage of applications of quantum resources over the classical ones. Some examples are *self-testing of quantum states* [51], security on *cryptographic* protocols [11], and enhancement of performance in *communication complexity problems* [12]. In this sense, we may ask whether nonlocality is the distinctive feature that identifies quantum theory, from which its mathematical formalism could emerge. Equivalently, may quantum nonlocality be considered an Axiom? That question was negatively answered by Sandu Popescu and Daniel Rohrlich in 1994 [15]. In fact, as we discussed, there exists a limit on quantum nonlocality (Eq.

(1.26)), and there are nonlocal correlations fulfilling all assumptions from Bell scenarios that could never be reproduced by following the quantum mechanics laws. Thus, the question about the features making sense of quantum theory's laws might be slightly rephrased, and we may ask whether all nonlocal correlations have a correspondence in nature. In this direction, if there are consistent reasons for certain nonlocal correlations to be *never* experimentally observed, we may 'make sense' of quantum mechanics laws by rationalizing the limit on quantum mechanics nonlocality. This query is precisely the research line that asks for *device-independent principles*. Some examples are the principle of *Non-trivial communication complexity* [17, 18], *Macroscopic locality* [19], and *Information causality* [20]. The latter is of particular relevance for the present thesis.

The next Chapter is devoted to discussing the close relationship between Bell's nonlocality and its advantages in communication problems.

Chapter 2

Device-Independent approach and communication problems

Bell nonlocality is, in fact, one of the cornerstones of quantum information science since it has been identified as an important tool related to quantum advantage in information processing problems [11, 12, 51]. Interestingly, one of the most exciting features that motivate employing *Bell tests* in such context is its device-independent design [52]. In the context of *cryptographic key distribution*, for example, while standard quantum frameworks, such as *BB84 protocol* [53, 54], become unsafe in the DI approach, *Ekert91 protocol* [11] incorporates *Bell test*. It avoids the need for keys to leave the laboratories, and Bell's theorem ensures the key has no pre-existence prior to the measurements, ensuring the key security. Further developments have shown an even closer connection between nonlocality and cryptography security, where the degree of nonlocality implies security enhancement for the so-called *CHSH protocol* [52]. Interestingly, at the same time, nonlocality has been studied as a resource enhancing communication performance [12, 28, 20]. In this Chapter, we review some of the main literature relating nonlocality as a resource for communication and present some contributions advancing to bigger and more general scenarios.

2.1 Device-independent quantum key distribution

The challenge of hiding secret messages is far from new, as it has long been central to safeguarding critical information such as state or military secrets. Nowadays, much of our online activity is secured by cryptographic protocols, like the widely used Rivest–Shamir–Adleman cryptosystem (RSA) [55]. Nonetheless, current cryptographic schemes demand continuous vigilance, as advancements that could potentially compromise these systems are always a concern. In fact, while one of the standard current cryptographic systems (RSA) is based on the inefficiency of current computational technology in solving specific mathematical problems, quantum computation offers, in principle, alternative ways to efficiently solve it (say, the prime number factoring problem [56]). Given the uncertainty about the adversary's computational power, alternative ways of designing cryptographic schemes based on the laws of physics have been developed, regardless of the computational power of an intercepting part.

The crucial part of modern cryptographic strategies relies on distributing safe cryptographic keys among the parties. For instance, the one-time pad encryption scheme, first proposed by Vernam cipher in 1920 [57], was proven information-theoretically secure by Shannon in 1949 [58]. In this scheme, the message is represented as a bit string M, while the two communicating parties, Alice and Bob, share identical bit-string cryptographic keys, S_A and S_B, where $M, S_A, S_B \in \{0,1\}^n$. The encryption strategy relies on chiper the message by binary adding with the key as $C_i = M_i \oplus S_{Ai}$. Thus, Shannon proved that even though the adversary has complete access to the ciphertext, C, it provides no information about the original message under certain assumptions for the keys. For our purpose, we will focus only on one crucial requirement: the adversaries must have no access to the keys. Equivalently, we are interested in the problem of how the parties may secretly *distribute cryptographic keys*.

In general terms, we may understand any mechanism hiding the communicated message by *key*. Here, we reduce our scope to bit-string keys. Thus, Alice and Bob must generate precisely the same key in two separate places, ideally avoiding information leakage to an eventual eavesdropper. Remarkably, quantum mechanics predicts several helpful phenomena in this context, such as *inherent randomness* and the *no-cloning theorem*. Indeed, the first *quantum key distribution* protocol was proposed by Charles Bennett and Gilles Brassard in 1984, the so-called *BB84 protocol* [53], which encodes classical bits into qubits. Essentially, the protocol takes advantage of the fact that quantum states cannot be cloned [59] or measured without disturbance. Thus, by transmitting keys through quantum states, parties can always detect any information leakage caused by potential eavesdroppers. Subsequent developments include the *Ekert 91* protocol [11], *Entanglement-based BB84* [54], and *Six-state* protocol [60], quickly followed by the primary experimental demonstrations of QKD [61, 62, 63, 64, 65]

Although the QKD framework is, undoubtedly, a milestone for quantum information science - where the laws of quantum mechanics play a crucial role in security assurance - standard QKD protocols rely on the precise characterization of the devices. For example, the *Entanglement-based* BB84 protocol [54] assumes that Alice and Bob initially share a maximally entangled EPR pair. However, as later demonstrated in *Ref.* [52], when no hypotheses are made on the dimension of the state, the BB84 protocol statistics can be achieved by separable states, for which no security can be certified. In this direction, despite the often natural assumption about the full control of the physical systems by experimentalists, this level of control is particularly challenging in quantum mechanics [66, 67, 68]. Consequently, it becomes important to question whether security may be extracted for scenarios involving untrusted devices. *i.e.*, *device-independent* (DI) scenarios.

The first DI proposal for cryptographic key distribution was the *Ekert 91* protocol [11], which insightfully explores Bell's theorem. While the standard QKD approaches address interference from potential eavesdroppers, Ekert's protocol enables the keys to remain within the laboratories during the experiment. Moreover, through Bell's theorem, the protocol certifies that the keys have no causal relations with any variable prior to the experiment, consequently ensuring that no information about them is accessible before the experiment. To do so, in Ekert's setup, Alice and Bob share a maximally entangled pair of qubits $|\Psi^-\rangle_{AB} = (|01\rangle_{AB} - |10\rangle_{AB})/\sqrt{2}$ and, in each run of the experiment, perform the respective measurements, A_x, and B_y, specified as:

$$\begin{aligned} A_1 &= \sigma_z, & B_1 &= \sigma_z, \\ A_2 &= \sigma_x, & B_2 &= \frac{1}{\sqrt{2}}(\sigma_z - \sigma_x), \\ A_3 &= \frac{1}{\sqrt{2}}(\sigma_z + \sigma_x), & B_3 &= \frac{1}{\sqrt{2}}(\sigma_z + \sigma_x). \end{aligned} \tag{2.1}$$

In rounds where they measure measurement settings, (A_1, B_1) and (A_3, B_3), Alice and Bob observe perfectly anti-correlated results. In the other rounds—(A_1, B_3), (A_1, B_2), (A_2, B_3), and (A_2, B_2)—they test the CHSH inequality (1.13). Violating this inequality assures that the outputs do not pre-exist and that the results are indeed anti-correlated for measurements in identical directions, providing a secure basis for establishing the secret key. Later developments have discussed that security for Ekert's protocol actually relies on entanglement witnessed by Bell inequalities [54, 69], and did not offer any new insight beyond BB84 protocol. In fact, the maximum CHSH violation achieving $S = 2\sqrt{2}$ certifies Alice and Bob share the maximally entangled state $|\Psi^-\rangle_{AB}$ [51], which necessarily cannot be entangled with any other part. More interestingly, however, the Ekert 91 protocol works without any extra hypothesis about sources and devices and requires solely the Bell inequality violation. This fact highlights the DI feature of the protocol, stating a major difference with respect to the standard QKD ones.

For Ekert 91, Bell's nonlocality ensures the quantum nature of sources and devices. Nevertheless, the DIQKD framework goes one step further and asks whether security can be proved only from Bell's theorem, independent of quantum theory's formal structure [52, 70]. The DIQKD scenario can be formalized as follows: in their respective laboratories, Alice and Bob control parts of a composite system on which measurements will be performed. The shared systems might be communicated between the parties prior, or, in the worst case, might even be distributed by an eavesdropper, Eve, since no assumptions are made about the devices. Once the subsystems are received, the laboratories are assumed to be entirely closed, and no information leakage occurs during the experiments. In this case, the statistics of Alice and Bob are specified by correlations from a Bell scenario $p(a, b|x, y)$, and, by construction, respects the no-signaling conditions (1.4). The lists with the experiment outputs, a and b, build up the raw key. According to Ekert's intuition, if $p(a, b|x, y)$ violates a Bell inequality, a and b are not known by the potential distributing part, Eve, since it did not pre-exist before the measurements. Consequently, in the DI framework, no security can be extracted for correlations compatible with the description in terms of local classical model (1.8), $p(a, b|x, y) = \sum_\lambda p_\lambda p(a|x, \lambda)p(b|y, \lambda)$. Otherwise, Eve may simply keep a copy of λ, having a complete description of the keys.

The first scheme providing secrecy based on physical principle was proposed in *Ref.* [71], where security is proven assuming only the no-signaling principle. As previously mentioned, there are no-signaling correlations never achieved by quantum theory [15]. Interestingly, however, even assuming eavesdroppers who could hypothetically break the laws of quantum mechanics, secrecy can be extracted as long as the parties cannot signal superluminally. Subsequent works immediately followed the steps of [71], extending the theoretical framework to prove DI security in different contexts [52, 70, 72, 73]. *Ref.* [52], in particular, introduced the so-called *CHSH protocol*, which, as the name suggests, relies on the violation of CHSH inequality (1.13). In that context, the parties have a, b, x and $y \in \{0, 1\}$, aiming to optimize the CHSH violation, which we may re-write as:

$$\beta(\mathcal{A}, \mathcal{B}) \equiv \frac{1}{4}\sum_{x,y} p(a \oplus b = xy|x, y) \leqslant \frac{3}{4}. \tag{2.2}$$

Thus, the parties can establish the keys with the outcomes a and b. After the measurements, Bob publicly announces his measurement choices y, enabling Alice to adjust the key by flipping $a \to a \oplus 1$ when $x = y = 1$, and simply retaining the data otherwise. Notice that when parties share a PR-box, $\beta(\mathcal{A}, \mathcal{B}) = 1$, ensuring that their keys perfectly agree.

To prove security, the authors describe the optimal Eve individual attack specified by the tripartite distribution $p(a,b,e|x,y,z)$, where the cardinality of inputs and outputs for Eve does not necessarily need to be restricted. In this case, we may assume the correlations between Alice and Bob are distributed by Eve,

$$\begin{aligned} p(a,b|x,y) &= \sum_e p(a,b,e|x,y,z) \\ &= \sum_e p(e|z)p(a,b|x,y,z,e), \end{aligned} \tag{2.3}$$

and her optimal strategy is preparing the extremal points associated with the marginal bipartite scenario involving Alice and Bob [70]. Aware of Alice and Bob's intent to implement the CHSH protocol, Eve's most effective strategy is preparing those extremal points that maximize the CHSH value (2.2) between Alice and Bob. When Eve sends deterministic correlations, she has complete knowledge of Alice and Bob's data. In contrast, she has no information about their outcomes when preparing a PR-box[1]. Thus, Eve's strategy is specified by:

$$p(a,b|x,y) = \sum_{e_a,e_b} p(e_a,e_b)p_{e_a,e_b}(a,b|x,y), \tag{2.4}$$

where $e_a, e_b \in \{0,1,2\}$ specify the extremal points, $p_{e_a,e_b}(a,b|x,y)$, which are:

$$\begin{aligned} L_1 &= \{a = \alpha x \oplus \beta, b = \gamma y \oplus \delta | \alpha = \gamma, \delta = \beta \oplus \gamma\}; \\ L_2 &= \{a = \alpha x \oplus \beta, b = \gamma y \oplus \delta | \alpha = \gamma \oplus 1, \delta = \beta\}; \\ L_3 &= \{a \oplus b = xy\}, \end{aligned} \tag{2.5}$$

where $\alpha, \beta, \gamma, \delta \in 0, 1$. A critical result from [52] demonstrates that with this protocol, without loss of generality, Eve can choose $p(e_a, e_b)$ as a convex mixture of a uniform distribution among the deterministic points and a PR-box. Thus, $p(a,b|x,y)$ can be expressed as:

$$p(a,b|x,y) = \frac{p_{NL}}{2}\delta_{a\oplus b,xy} + p_L \sum_{(e_a,e_b)\neq(2,2)} \frac{1}{8}\delta_{a,D_{e_a}(x)}\delta_{b,F_{e_b}(y)}, \tag{2.6}$$

where p_{NL} controls Alice and Bob nonlocality, and respects $p_{NL} + p_L = 1$. The deterministics functions D_{e_a} and F_{e_b} denote conviniently the distributions in (2.5).

The foundational Csiszár-Körner theorem guarantees security [75], which establishes that Alice and Bob can always distill a secret cryptographic key against an *individual eavesdropper attack* whenever their data admit an advantage in terms of Shannon's mutual information, *i.e.*,

$$I(A:B) > I(E:B), \tag{2.7}$$

[1] As we shall see in the next chapters, PR-box correlations imply no correlation with any other part [74].

where,

$$I(A:B) \equiv \sum_{a\in A, b\in B} p(a,b) \log\left(\frac{p(a,b)}{p(a)p(b)}\right). \tag{2.8}$$

In other words, whenever the mutual information between Alice and Bob's set of keys exceeds that between Bob's keys and Eve's set of decodings, Alice and Bob can securely establish cryptographic keys. Therefore, based on the strategy described by Eq. (2.6), whenever the parties are limited by no-signaling principle, security is always ensured by condition (2.7) as $p_{NL} \gtrsim 0.318$. Even more importantly, the quantum achievable value for (2.6) is $p_{NL} \lesssim 0.414$ [52]. Consequently, even when Alice and Bob are limited to producing correlations described by quantum mechanics, they can guarantee security against a no-signaling Eve, which is not necessarily limited by quantum physics. This result naturally prompts whether DI secrecy may always be extracted for arbitrarily small Bell inequality violations. In fact, by improving the analysis with classical *pre* and *postprocessing* protocols, such as *advantage distillation* [76], privacy can be ensured as long as $p_{NL} \gtrsim 0.093$ [70]. DI security against no-signaling collective attacks was later proved [72, 73], as well as, more recently, security proof for arbitrarily small nonlocality [77], and experimental implementation [78, 79].

As briefly highlighted in this subsection, the DI framework has initiated an advantageous program where the theoretical proof does not rely on assumptions about the sources and measurement apparatus. Interestingly, the no-signaling principle permits supra-quantum correlations, meaning that DIQKD security proofs could remain valid even after a hypothetical breakthrough in quantum mechanics laws. As we will discuss in the last Chapter of the present thesis, alternative DI principles ensure secrecy independently of quantum formalism [30]. Remarkably, since DI security relies on Bell inequalities violations, one significant challenge for practical DIQKD is highly connected to the current technological barrier to efficiently accomplishing loophole-free Bell tests, such as [6, 7, 8].

2.2 Nonlocality-assisted communication

As briefly introduced in the previous section, quantum mechanics allows for expressive enhancements in the context of cryptographic key distribution. Beyond QKD, quantum resources may also improve communication itself. In particular, despite *Holevo's bound* stating that a quantum bit encodes at most the same amount of information as a classical bit [80], *dense coding*, for instance, doubles the capacity of parties sending information when they dispose of entanglement and communicate through quantum states [81]. Moreover, sending quantum states instead of classical ones also allows communication advantage when the scenario addresses a particular task, such as the random access codes (RAC) [82, 26, 83]. Interestingly, all cases where quantum resources are available admit an advantage with respect to the classical counterpart. Say, *classical communication assisted by quantum correlations*[2] [84, 85], *quantum communication* [68, 67], or even *quantum communication assisted by quantum correlations* [81, 54, 86]. The interconversion among these different kinds of resources is still far from completely clear. However, recent developments have introduced computational tools allowing a more systematic analysis

[2] Frequently referred to as Entanglement-assisted communication.

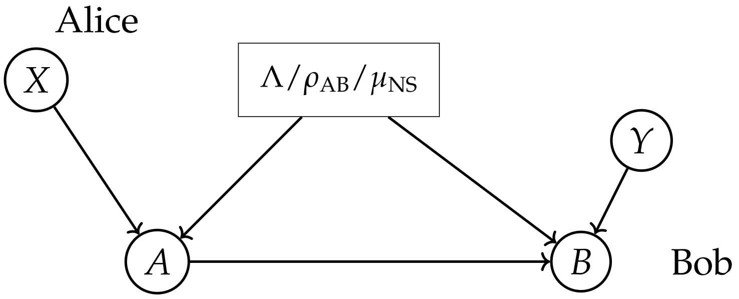


Figure 6 – Communication scenario where $x \in [n_X] = 0, 1, \cdots, n_X - 1$, $y \in [n_Y] = 0, 1, \cdots, n_Y - 1$, and $b \in [n_B] = 0, 1, \cdots, n_B - 1$. The symbol a represents the classical systems sent by the transmitting part, and $\lambda/\rho_{AB}/\mu_{\mathrm{NS}}$ represents the possible systems correlating the parties.

[86, 87]. For instance, in *Ref.* [88], we describe an equivalence between the classical and quantum communication scenario whenever the parties dispose of entanglement as a free resource. In this case, everything possible by sending $d-$dimensional quantum systems is reproduced by sending two $d-$dimensional classical systems. Intuitively, entanglement phenomena play a crucial role in communication, similar to the Bell and QKD scenarios. In fact, while there are entangled states useless for communication, as long as parties may access higher-dimensional entangled states, it always provides communication advantage [88].

More generally, a communication scenario might be formalized as instances of a *prepare and measure* (PM) scenario. In this case, a sender (Alice) prepares a system to be communicated based on an input $x \in [n_X] := \{0, \cdots, n_X - 1\}$, and a receiver (Bob) who will measure the received systems based on an input $y \in [n_Y]$. The experiment is fully described by the conditional probability distributions $p(b|x,y)$, where $b \in [n_B]$ denotes the receiver's measurement outputs. The communication scenario is specified by the tuple (n_X, n_Y, n_B) and is depicted in Fig.6 for the possible different cases where parties are correlated by *classical* (Λ), *quantum* (ρ_{AB}), or even more general *no-signaling* resources (μ_{NS}). In spite of the several mentioned possibilities, for this thesis, we will limit our discussion to classical communication scenarios. Analogously to Bell scenarios, the classical model reads as [68],

$$p(b|x,y) = \sum_a \sum_\lambda p(\lambda) p(a|x,\lambda) p(b|a,y,\lambda), \tag{2.9}$$

where, $a \in [n_A]$ The set of correlations admitting the classical model (2.9) is also a polytope; consequently, its characterization is equivalently written in terms of Bell-type inequalities [68],

$$\sum_{b,x,y} C_{b,x,y} p(b|x,y) \leqslant \beta_C, \tag{2.10}$$

which are obtained via facet enumeration algorithms (PANDA being an example of [35]). In this case, we may read this set of inequalities as characterizing the boundaries of the communication performance achievable by the resource Λ. In turn, the correlations achieved by quantum resources are described as [86, 87],

$$p(b|x,y) = \sum_{a} \mathrm{Tr}[\rho_{AB}(M_{a|x} \otimes N_{b|a,y})], \quad \text{where } \rho_{AB} \in \mathcal{L}(\mathbb{C}^D \otimes \mathbb{C}^D), \tag{2.11}$$

In fact, there are quantum correlations as (2.11) that do not admit a classical model (2.9), which is witnessed by the violation of the Bell-type inequality (2.10). Quantum bounds on the classicality witnesses (2.10) may efficiently be achieved through semidefinite programming optimization methods, which we refer to as β_Q [86]. Not surprisingly, just as in Bell scenarios, correlations assisted with more general no-signaling resources may lead to even stronger nonclassicality, surpassing the quantum violations of (2.10) [17, 28, 20]. Consequently, we may generically define correlations assisted by no-signaling resources μ_{NS} as probabilities distribution $p(b|x,y)$ for which there exists a correspondent $(n_X, n_Y \cdot n_{Y'}; n_A, n_B)$-Bell scenario specified by $p_{\mathrm{Bell}}(a,b|x,y,y')$, such that

$$p(b|x,y) = \sum_{a} p_{\mathrm{Bell}}(a,b|x,y,y'=a). \tag{2.12}$$

Beyond witnessing the advantageous implications of nonclassical resources, likewise the standard Bell inequalities, the inequalities (2.10) define different communication tasks, which are closely related to several information processing problems, such as *semi-device independent dimension certification* [68, 89], *self-testing states* [90, 91], as well as *semi-device independent cryptographic keys distribution* [92]. Of particular relevance, one class of the Bell-type inequalities (2.10) addresses the random access codes tasks [82, 26, 83], which has special significance for the present thesis.

2.2.1 Random Access Codes

Formally, a random access code (RAC) addresses a communication task where the input for the sending part is a n-dit string of size n, $x = x_0, x_1, ..., x_{n-1}$, while the receiving part needs to decode the message in order to access one of the n initial dits randomly [26, 27]. Thereby, Alice encodes her input-dits into a message of m dits, such that $m < n$, and $y \in 0, 1, ..., n-1$ indicates to Bob which dit he should attempt to guess. In the literature, a RAC scenario is frequently denoted by $n \mapsto m$. The figure of merit, in this case, is the success probability:

$$p_s \equiv \frac{1}{nd^n} \sum_{i}^{n} p(x_i = b_y|x, y = i), \tag{2.13}$$

which sets up one Bell-type inequality (2.10). When parties are limited to $d = 2$ and classical resources, their success probability (2.13) is limited by [93]:

$$p_s \leqslant \frac{1}{2}\left(1 + \frac{m}{n}\right). \tag{2.14}$$

Interestingly, for the particular $(n_X = 4, n_Y = 2, n_B = 2)$ PM scenario, the RAC success probability precisely defines one of the facets of the classical polytope defined from (2.9) [92].

As mentioned for the general PM scenarios, quantum communication is advantageous over classical one in RAC task [27, 93, 94]. Particularly significant, however, is that even when Alice and Bob are limited to classical communication, sharing nonclassical correlations significantly enhances their performance. For the $n \mapsto 1$ binary RAC where $d = 2$, for instance, entanglement-assisted correlations as (2.11) surpasses the classical bound (2.14) up to the optimal value of [85]:

$$p_s \leqslant \frac{1}{2}\left(1 + \frac{1}{\sqrt{n}}\right). \tag{2.15}$$

Remarkably, the entanglement-assisted RAC bound (2.15) also surpasses the optimal performances of RAC with quantum communication [85].

When parties may access even more general no-signaling correlations as (2.12), however, Alice and Bob may outperform the classical and quantum bounds up to the algebraic bound $p_s = 1$, then trivializing RAC tasks. To illustrate, consider first the simplest non-trivial $2 \mapsto 1$ binary RAC, where Alice and Bob dispose of a PR-box (1.17). Here we distinguish variables from the Bell scenario as $a', b', x', y' \in \{0, 1\}$, then $a' \oplus b' = x' \cdot y'$. The optimal protocol in this case consists of Alice using as input $x' = x_0 \oplus x_1$ in her part of the no-signaling resource and coding the message $a = a' \oplus x_0$. Thus, Bob can perfectly decode the message by inputting on his part, $y' = y$, and producing $b = a \oplus b'$ as an outcome. In this case, the PR-box property ensures $b_y = x_y$, which naturally saturates the success probability (2.13) as $p_s = 1$. In fact, a straight interconversion exists between nonlocality and advantage in RAC. For example, achieving $p_s = 1$ for the $2 \mapsto 1$ binary RAC means that parties necessarily share a PR-box [28, 95]. Remarkably, in the case of dits, there always exists a nonlocal correlation trivializing the $2 \mapsto 1$ RAC [96]. Interestingly, the strong consequences of nonlocal correlations also appear in different other information processing problems. For example, the communication complexity problems also become trivial employing PR boxes [97, 17]. We shall devote our attention in the subsequent subsection to discussing the implausibility of such strong consequences of extremal nonlocal correlations, which compose one of the building blocks for the developments of *device-independent operational principles* [18, 20].

More recent developments have investigated which classes of nonlocal correlations accomplish this straight interconversion, showing that for binary variables, $(n-1)$ identical copies of PR-boxes are necessary and sufficient to trivialize any $2^n \mapsto 1$ RAC [98]. Consequently, in this case, we may write:

$$p_s \leqslant 1. \tag{2.16}$$

To prove this equivalence, the authors recursively concatenate the $(2^n - 1)$ $2 \mapsto 1$ protocols. Such procedures have been explored in different contexts, such as the entanglement-assisted RACs [85], (2.15), and in the DI principles [20, 29]. For further details, we address the reader to Chapter 3, where we provide a detailed multipartite generalization to the concatenation procedure from which the standard one follows as a particular case.

This section will address a more straightforward strategy, configuring a more suitable use of PR boxes. In fact, despite the significance of the interconversion between RAC and nonlocal correlations, the inherent structure of the RAC task limits the potentiality of the communication power of those correlations. This becomes clear considering a simple $3 \mapsto 1$ binary RAC example, where the parties execute the protocol specified in Fig.7. In this case, they dispose of two PR-boxes, which we distinguish with the upper

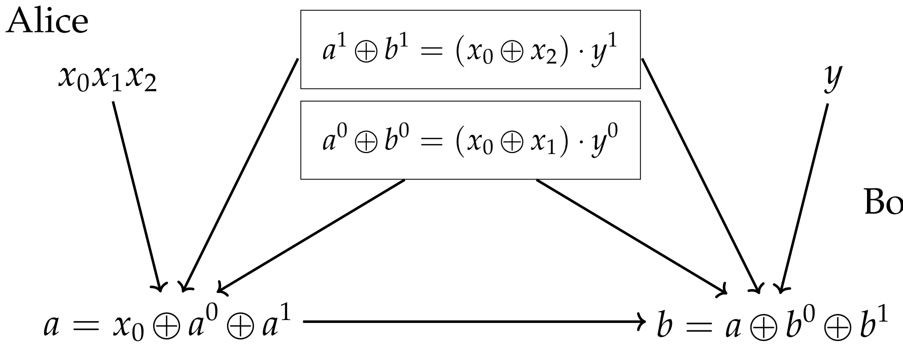

Alice
$x_0x_1x_2$
$a^1 \oplus b^1 = (x_0 \oplus x_2) \cdot y^1$
$a^0 \oplus b^0 = (x_0 \oplus x_1) \cdot y^0$
y
Bob
$a = x_0 \oplus a^0 \oplus a^1$
$b = a \oplus b^0 \oplus b^1$

Figure 7 – Encoding strategy for 3 ↦ 1 binary RAC, where parties share two PR boxes specified as $a^j \oplus b^j = x^j \cdot y^j$. Alice encodes the initial bits in the box, as $x^j = x_0 \oplus x^{j+1}$, and builds up the message $a = x_0 \oplus a^0 \oplus a^1$. Bob's decoding strategy consists of producing the output $b = a \oplus b^0 \oplus b^1$.

indexes, *i.e.*, $a^j \oplus b^j = x^j \cdot y^j$, which yields the output function:

$$b_{y^0y^1} = (x_0 \oplus x_1)y^0 \oplus (x_0 \oplus x_2)y^1 \oplus x_0. \tag{2.17}$$

Notice that for the cases where Bob inputs on his boxes $(y^0y^1) = (00), (10), (01)$, He perfectly recovers one of the initial bits, thus winning the RAC task with certainty. Nevertheless, the boxes inputs (y^0y^1) allow for even more potentiality since Bob can even choose to access a function of the initial bits $b_{11} = x_0 \oplus x_1 \oplus x_2$. The success probability (2.13) never captures the last case. The straight extension of the protocol can always trivialize any $n \mapsto 1$ binary RAC by using $n-1$ PR-boxes with the following strategy:

$$\begin{aligned} x^i &= x_{i+1} \oplus x_0 \quad \forall\, i \in \{0, \cdots, n-2\}; \\ a &= x_0 \oplus \bigoplus_{j}^{n-2} a^j; \\ b &= a \oplus \bigoplus_{j}^{n-2} b^j; \\ &\begin{cases} y^j = 0 \quad \forall j \text{ if } y = 0; \\ y^j = \delta_{j,y} \text{ if } y \geqslant 1. \end{cases} \end{aligned} \tag{2.18}$$

Similarly to the particular case, (2.17), the strategy (2.18) provides the potential access to Bob to the multiple functions of the initial bits as,

$$b_{y^0y^1,\cdots,y^{n-1}} = \left(\bigoplus_{i=0}^{n-2} (x_0 \oplus x_{i+1}) y^i \right) \oplus x_0, \tag{2.19}$$

which are never accessed when parties are limited to winning the RAC task. In this case, it is necessary to address a more general figure of merit than (2.13), which we may read as

$$p_s \equiv \frac{1}{nd^n}\sum_{i}^{n} p(f_j(x_0,\cdots,x_{n-1}) = b_y|x_0,\cdots,x_{n-1},y=i), \tag{2.20}$$

where $f_j(x_0,\cdots,x_{n-1}) : \{0,1\}^n \mapsto \{0,1\}$ addresses to the different boolean functions of the initial bit string.

The RAC task is originally closely related to relevant problems as *finite automata machine* in the context of computation [82, 99, 26, 27]. However, as we shall see, this convenient slight change of perspective better highlights the weirdness of certain nonlocal correlations, which enables significant advances in our understanding of the correlations observed in nature [29, 30, 23]. Remarkably, recent developments have investigated slight modifications of RAC tasks with different application contexts [100, 101, 102, 103]. *Ref.* [101], for example, systematically studies the generalized version of the RAC task in (2.20) for several different instances, such as classical, quantum, and supra-quantum regime. Of particular relevance, the authors investigate the success probability of Bob rightly computing the value of a given boolean function, $f_j(x_0,\cdots,x_{k-1})$, for any subset of k bits from the initial bit-string (where $k \leqslant n$). Especially meaningful, in the latter case, for all functions f_j, it is always possible to trivialize the task (2.20) using $n \mapsto 1$ RAC, with the support of PR-boxes [101]. In this direction, notice that whether PR-boxes are a free resource available to the parties, the protocol (2.18) allows the parties parallelly win with certainty all the generalized tasks in (2.20) in the same round, by sending a single bit of message. In fact, (2.18) allows $p_s = 1$ in $n \mapsto 1$ for all integer n. Consequently, if Alice has initially n bits, there exists a double exponential, but finite, number of boolean functions of n variables, which may be directly encoded in a $2^{2^n} \mapsto 1$ with strategy (2.18).

The slight modification on the RACs task in (2.20) highlights the similarities of RACs and its cousin communication task, the *communication complexity problems.*

2.2.2 Communication complexity problems

The *Comunication complexity* (CC) scenario can be stated as the problem of determining the minimum amount of information that parties need to exchange to correctly compute the value of a function that jointly depends on the parties' initial set of data. In the particular case of two parties, Alice and Bob have initially their respective set of data, $\mathbf{x} \in \{0,1\}^n$ and $\mathbf{y} \in \{0,1\}^n$, and need to compute the joint function $f(\mathbf{x},\mathbf{y})$, where for integers n and m we have $f : \{0,1\}^n \times \{0,1\}^n \mapsto \{0,1\}^m$. Hence, while RACs address multiple potential computations in each round, the parties focus on rightly performing a particular computation in the CC scenario. Similarly to RACs, CC problems are closely related to several applications in computer science, such as *finite automata machines*, *Turing machines*, and *decision tree computation*, for which we address *Ref.* [104, 105] for further detail. Quantum mechanics may offer an advantage in CC problems, first shown by considering parties sending quantum states [106] and later sharing entangled states [107]. Moreover, nonlocality is closely related to advantage on the CC problem since violating Bell inequalities was shown to be a necessary and sufficient condition for quantum strategies outperforming classical ones [12, 108].

More specifically, different functions f imply different communication complexities. For example, there is no quantum CC advantage when parties need to compute the inner product function $f(\mathbf{x},\mathbf{y}) = \bigoplus_{j=0}^{n-1} x_j y_j$, even when they may explore an unlimited amount of entangled qubits [109]. In this case, parties have necessarily to communicate n bits. Nevertheless, when parties may access PR-boxes, $a^j \oplus b^j = x^j \cdot y^j$, the task may be trivially solved with a single bit message [97, 17]. Indeed, Alice and Bob can encode each of their n bits into n different PR-boxes, as $x^j = x_j$ and $y^j = y_j$. Thus, Alice may simply signal to Bob, $a = \bigoplus_{j=0}^{n-1} a^j$, enabling him to decode as

$$\begin{aligned} b &= \left(\bigoplus_{j=0}^{n-1} a^j\right) \oplus \left(\bigoplus_{j=0}^{n-1} b^j\right), \\ &= \bigoplus_{j=0}^{n-1} \left(a^j \oplus b^j\right), \\ &= \bigoplus_{j=0}^{n-1} x_j \cdot y_j. \end{aligned} \tag{2.21}$$

Interestingly, however, any boolean function $f : \{0,1\}^n \times \{0,1\}^n \mapsto \{0,1\}$ may be written as a sum of products of a finite number of polynomials, $P_j(\mathbf{x})$ and $Q_j(\mathbf{y})$. *i.e.*, $f(\mathbf{x},\mathbf{y}) = \bigoplus_j P_j(\mathbf{x}) \cdot Q_j(\mathbf{y})$. Consequently, with the support of a finite number of PR-boxes, Alice and Bob can compute the value of any function $f(\mathbf{x},\mathbf{y})$ with a single bit message [17]. *i.e.*, when PR boxes support parties, the communication complexity problem becomes trivial. The result is straightforwardly extended to multipartite scenarios when parties may access the multipartite equivalent correlation, such as (1.30).

As stressed by Wim van Dam in the original work [17], the result presents no apparent conflict with any physical notion. Nevertheless, the task trivialization for any possible distributed function strongly opposes the current comprehension that certain problems are inherently more computationally demanding than others. Similarly to the concept of different *complexity classes* in computer science, a similar notion emerges within the context of CC [110]. In fact, CC problems are closely related to the depth of circuits problems [105]. Consequently, we may interpret the result (2.21) as an *implausible consequence* of supposing the existence of PR box correlations in nature. Building upon this perspective, further developments have shown that CCP is also trivialized for a substantial variety of nonlocal correlations beyond the quantum boundaries [18]. As previously observed, quantum theory presents limited nonlocality (1.26). In this sense, such implausible consequence in (2.21) offers a rational explanation for certain super-strong nonlocal correlations to be possibly never observed in nature. This approach is followed in other DI principle proposals. Of particular relevance, in the subsequent Chapters, we shall deeply investigate one such proposal, the so-called *Information Causality principle* [20], which identifies implausible consequences for all nonlocal correlation beyond the Tsirelson's bound (1.26).

Chapter 3

Information causality in multiple parties scenarios

One of the most significant device-independent (DI) principles presenting substantial advancements in the direction of singling out the set of quantum correlations from the more nonlocal ones is the information causality principle ($\mathcal{IC}$) [20]. The principle establishes a reasonable limitation on communication scenarios composed of a sender and a receiver, which, as will soon become clarified, aims to prevent implausible consequences arising from the hypothetical existence of some supra-quantum resources.

In simple terms, the amount of available information to the receiver concerning the sender's initial data cannot exceed the amount of information effectively transmitted by the sender. Quantum correlations are known to always respect $\mathcal{IC}$, whereas the principle prohibits many stronger-than-quantum correlations [20, 21]. Perhaps the most remarkable achievement of $\mathcal{IC}$ is the recovery of the *Tsirelson's* bound (*Ref.* [16]) on the maximum quantum violation of the Clauser-Horne-Shimony-Holt (CHSH) inequality. However, it remains unclear whether *all* stronger-than-quantum correlations violate it.

Despite its intuitive appeal, one of the main challenges lies in providing an operational mathematical formulation of the principle. A sufficient criterion for the violation of $\mathcal{IC}$ was introduced in Ref. [20], but it has since proven the existence of non-quantum correlations never violating the principle [21]. Since then, other techniques have been developed, establishing refined criteria for $\mathcal{IC}$ [22, 23, 23]; but, so far, none of them have proven robust enough to precisely characterize the set of quantum correlations, even in the simplest scenario.

Further developments have revealed that any informational principle describing the quantum correlations set must be genuinely multipartite [24]. This observation is complemented by the latter findings, showing that the bipartite formulation of the original $\mathcal{IC}$ proposal is insufficient to exclude specifics extremal tripartite stronger-than-quantum correlations [25]. Hence, it emerges the necessity for a genuinely multipartite formulation of $\mathcal{IC}$, in order to be a valid and precise operational principle for quantum theory.

In this chapter, we review the primitives and the current literature concerning the $\mathcal{IC}$ principle and present our recent advancements for a novel multipartite reformulation for $\mathcal{IC}$ (published in [29]). In that regard, we explore the general Shannon's entropic cone approach — a sophisticated geometric approach widely applied in classical information theory but still underexplored in quantum information science. Significantly, the covered developments compose a crucial part of the results presented in the subsequent chapter (summarized in [30]), where we bridge foundational discussions of DI principles with practical applications, such as cryptographic key distributions.

3.1 $\mathcal{IC}$ statement

As extensively discussed in Chapter 1, a close relation exists between nonlocality and advantage in communication tasks. Remarkably, we have shown that when parties can access PR boxes as a free resource, they can trivialize RAC tasks, even in the more general form where the receiving part aims to access some post-processed information from the initial data sent. Similar to other proposals, the implausibility of this result lies in its subtle interpretation. Achieving a success probability of $p = 1$ implies that, even after receiving the message, the receiver (Bob) has potential access to the entire data set of the sender (Alice). Indeed, by construction, the choice of which part of the data set to access depends solely on Bob's input, y, independent of Alice's message. Hence, the Information Causality ($\mathcal{IC}$) principle regards the causal effect of Alice's transmitted message on the final available information of Bob about her set of data $\mathbf{x}$. Before Alice's transmission, Bob has no access to $\mathbf{x}$, which is changed by the communicated message. Therefore, $\mathcal{IC}$ postulates that the message causal effect on the receiver potential information cannot exceed the own message amount of encoded information[1]. In other words, the amount of information that the message itself contains limits the influence a message has. With this in mind, we are ready to present the formal statement of $\mathcal{IC}$, first introduced in [20]:

> *"Bob's potentially available information gain about the initial n bits of Alice, considering all their possible local as well as pre-established shared resources, cannot be greater than the number of bits m, sent by Alice."*

Equivalently, as later clarified by the authors in [111], it is implicitly assumed that the communication occurs through a single use of a classical channel with a capacity of m bits. As we will explore in the subsequent sections and chapters, this subtlety in the definition is crucial to the recent significant advancements in our understanding of $\mathcal{IC}$ [112, 29, 23, 113, 30]. Therefore, we rephrase the original statement in the following more precise form:

> *"The available information gain of the receiver about the initial data of the sender, considering all their possible local as well as pre-established shared resources, cannot be greater than the ammount of effectively received information, through a single use of a classical channel of capacity C."*

3.2 Bipartite operational criteria for $\mathcal{IC}$

One of the primary challenges related to $\mathcal{IC}$ lies in establishing well-defined operational criteria that can effectively capture all violations of the $\mathcal{IC}$ statement. To address

[1] For the sake of clarity, this counter-intuitive feature is best illustrated by a simple example. Consider Alice, who wishes to send the synopsis of a book containing n pages via e-mail instead of sending bits to Bob. For unspecified reasons, her e-mail service limits her to sending information equivalent to a single page. In this case, aside from bothering Bob with n separate e-mails, her best strategy is to condense portions of the book into one page. The RAC example with $p = 1$ is analogous to Bob being able to choose which part of the entire book to read, even after receiving only Alice's e-mail.

this, it is necessary first to formally define the $\mathcal{IC}$ scenario, as described in [20], where we follow the second statement outlined in the previous section. In this scenario, the sender, Alice ($\mathcal{A}$), receives bit string, $\mathbf{x} \equiv \{X_1, \ldots, X_n\} \in \{0,1\}^n$, of length n, sampled uniformly at random. $\mathcal{A}$ then encodes $\mathbf{x}$ into a classical message M of m bits, where $m < n$. This message M is transmitted through a noisy classical channel with capacity $C \leqslant m$, resulting in Bob, $\mathcal{B}$, effectively receiving M'. $\mathcal{B}$, in turn, decodes the message M' to produce a guess G_i about a randomly selected bit of X_i from $\mathcal{A}$ data, where $i \in \{1, \ldots, n\}$. To perform the task, $\mathcal{A}$ and $\mathcal{B}$ may utilize a composite system, ρ_{AB}, which states a no-signaling correlation between them.

The desired informational measure to compute Bob's available information should account for Alice's initial data, $\mathbf{x}$, as well as all Bob's local resources, the received message, M', and his part of the shared system, ρ_B. That is, $I(\mathbf{x} : M', \rho_B)$. The natural quantity capturing this is the standard mutual information measure. However, its form is theory-dependent[2]. Regardless of the particular form of I, the authors in [20] present the following set of properties, which are sufficient for any underlying theory to respect $\mathcal{IC}$:

1. Consistency: Whenever the random variables denote classical systems, I reduces to the classical mutual information;
2. Data processing: $I(A : B') \leqslant I(A : B) \quad \text{when} \quad B \longrightarrow B'$.
   i.e., local operations on B cannot increase its knowledge about A;
3. Chain-rule: $I(A : B|D) = I(A : B, D) - I(A : D)$;
4. Symmetry: $I(A : B) = I(B : A)$;
5. Non-negativity: $I(A : B) \geqslant 0$.

Naturally, these five conditions together imply the $\mathcal{IC}$ statement [20]:

$$I(\mathbf{x} : M', \rho_B) \leqslant C \leqslant m. \tag{3.1}$$

We leave the formal proof for the upper bound on $I(\mathbf{x} : M', \rho_B)$ to the subsequent sections, where we discuss the extension to multipartite scenarios [29].

In both classical and quantum cases, the mutual information measures do satisfy the five axioms, and consequently, $\mathcal{IC}$ always holds for systems described by quantum and classical theory. The scenario can be described in terms of quantum theory, where Alice's data set is written as orthogonal states $|\vec{x}\rangle \in \mathcal{H}^{2^n}$, and now ρ_{AB} describes a quantum state. Initially, this state is given by,

$$\left(\frac{1}{2^n} \sum_{\vec{x} \in \{0,1\}^n} |\vec{x}\rangle\langle\vec{x}| \right) \otimes \rho_{AB},$$

and quantum measurement $\{E_{M|\vec{x}}\}_M$ specify the encoding protocol of Alice. Besides, it is important to stress the generality of the result, which implies that $\mathcal{IC}$ always holds

[2] For instance, in both classical and quantum cases, this is accomplished using mutual information measures, in terms of Shannon and von Neumann entropies, respectively. However, the only assumption made about ρ_{AB} is that it respects the no-signaling condition.

for every theory satisfying the five presented properties. Notably, further research has revealed that the five required properties can be reduced to only two [114]. While the rationale for adopting these specific properties is subject to debate, subsequent works have proposed physically motivated entropy measures that respect them [115] and have identified which classes of general probabilistic theories are consistent with $\mathcal{IC}$ [116]. Nevertheless, whether other probabilistic theories with correspondence in nature satisfy these conditions beyond quantum and classical ones is unclear.

Alternatively, instead of asking about the physical appeal of $\mathcal{IC}$ in nature, one might ask whether the principle uniquely identifies the set of quantum correlations. To address this, it is worth noting that, by construction, everything in the $\mathcal{IC}$ scenario is classical except for the shared resource. By further analyzing the five required properties alongside (3.1), the authors in [20] derived the following necessary condition for $\mathcal{IC}$ in the case of uncorrelated initial bits[3]:

$$\sum_{i=1}^{n} I(X_i : G_i) \leqslant C \leqslant |m|, \tag{3.2}$$

which is a theory-independent criterion. The particular form of (3.2), in terms of noisy communication, first appeared in [112] and has the formal proof provided in [29]. As we will further demonstrate, this formulation indeed represents the most appropriate way of expressing $\mathcal{IC}$ criteria.

3.3 Information-theoretic implications of quantum causal structures

As previously mentioned, further developments have been trying to investigate appropriate operational informational criteria for $\mathcal{IC}$ [115, 116]. Remarkably, the framework introduced in [22] offers a general method to derive information-theoretic criteria for a given set of quantum variables when they are assumed to follow certain causal relations among them. In Appendix C, we provide a detailed description of the approach. In a nutshell, the method is described by the three-step algorithm: *(i) Characterizing the Shannon's cone.* It consits of enumerating all elementary inequalities that entropies of n variables must respect (they are *strong subadditivity* $H(A,C) + H(B,C) \geqslant H(A,B,C) + H(C)$ and *weak monotonicity* $H(A,B) + H(A,C) \geqslant H(B) + H(C)$). *(ii) Causal description.* Specifying all causal relations in terms of conditional mutual information (*i.e.*, $I(A : B|C) = 0$) and *data processing* inequalities. *(iii) Marginalization.* Removing from the description all entropic terms involving not jointly observable variables. Ultimately, the algorithm provides a set of entropic inequalities, which summarizes all information-theoretic criteria for the n causal-related variables. In that sense, information principles such as information causality are nothing else than entropic constraints arising from imposing a quantum description on a given causal structure.

The geometric-entropic framework is a typical standard computational instrument in the context of standard information theory [117]. Interestingly, it has a wide range of applications in several science research fields, for example, Bayesian statistics [118], or

[3] In Section 3.5, we generalize the theoretical proof to multipartite scenarios. Since the results in this section follow as a special case of the multipartite scenario, the formal proof is addressed in subsequent sections.

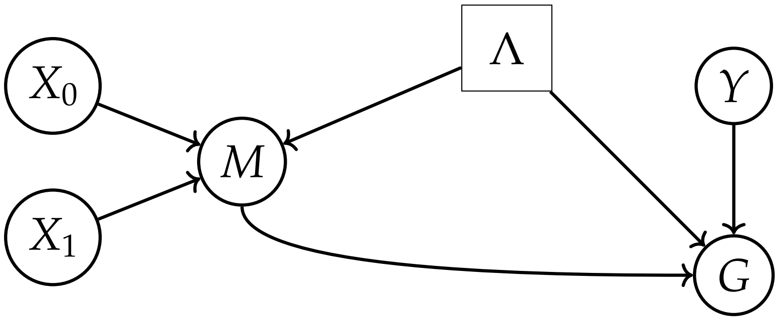


Figure 8 – Causal structure for the standard $\mathcal{IC}$ scenario.

even collective social behavior [119]. The approach has recently received attention in the quantum information field by the broad developments in the context of *marginal problems* [120, 121, 122, 22, 123, 124], which are closely related to Bell nonlocality [125] and Contextuality paradigms [126]. Indeed, the geometric treatment in the classical scenario provides general nonclassicality witnesses, equivalent to those provided by standard Bell inequalities violation [125]. In the context of quantum causality, in turn, the pioneering work of *Ref.* [22] generalizes the entropic framework for sets of random variables, including quantum systems, with remarkable applications to quantum networks scenarios [22, 127, 128]. Moreover, recent attempts have even extended the formalism to analyze causal relations in the context of generalized probabilistic theories.

In particular, *Ref.* [22] provides the general entropic description to the $\mathcal{IC}$ scenario for the noiseless case (*i.e.*, $M = M'$), whose quantum causal structure is depicted in Fig.8. Assuming $n = 2$, for simplicity, the set of random variables follows as $\{X_1, X_2, M, G_1, G_2, \rho_{\mathcal{A},\mathcal{B}}\}$. In this case, the only relevant causal relation is:

$$I(X_1, X_2 : \rho) = 0, \tag{3.3a}$$

which states the independence between the initial bits of $\mathcal{A}$ and the correlating resource. The remaining causal relations are encoded in terms of data processing inequalities. Following Appendix C, and denoting, respectively, $\mathcal{A}$ and $\mathcal{B}$ subsystems by ρ_A and ρ_B, the sets of jointly observable variables is written as

$$\begin{aligned} S_1 &= \{X_1, X_2, \rho_{\mathcal{A}}, \rho_{\mathcal{B}}\}, \\ S_2 &= \{X_1, X_2, M, \rho_{\mathcal{B}}\}, \\ S_3 &= \{X_1, X_2, M, G_1\}, \\ S_4 &= \{X_1, X_2, M, G_2\}. \end{aligned} \tag{3.4}$$

This fact implies that the only relevant data processing inequalities are of the form:

$$I(X_1, X_2, \rho_{\mathcal{B}} : \rho_{\mathcal{A}}) \geqslant I(X_1, X_2, \rho_{\mathcal{B}} : M_1), \tag{3.5a}$$

$$I(X_1, X_2 : \rho_{\mathcal{A}}, \rho_{\mathcal{B}}) \geqslant I(X_1, X_2, : M_1, G_j). \tag{3.5b}$$

By performing the Fourier-Motzkin elimination for the most general marginal scenario, $\mathcal{M} = \{\{X_1, X_2, M, G_1\}, \{X_1, X_2, M, G_1\}\}$, it achieves a set of 176 non-equivalent entropic inequalities, capturing all possibly derivable criteria within the communication $\mathcal{IC}$ scenario in Fig.8, holding for quantum theory. Of particular relevance, the authors present the following criterion as the most general criterion for $\mathcal{IC}$ when $n = 2$:

$$I(X_1 : M, G_1) + I(X_2 : M, G_2) + I(X_1 : X_2|M, G_2) \leqslant H(M) + I(X_1 : X_2), \tag{3.6}$$

from which (3.2) follows as a particular case for the noiseless scenario. Interestingly, by following the five properties mentioned in the previous section, the authors provide a general form of (3.6) for an arbitrary number of bits,

$$\begin{aligned}\sum_{i=1}^{n} I(X_i : G_i, M) + \sum_{i=2}^{n} I(X_1 : X_i|G_i, M)\\ \leqslant H(M) + \sum_{i=2}^{n} H(X_i) - H(X_1, ..., X_n).\end{aligned} \tag{3.7}$$

By the same arguments as (3.2), Eq. (3.7) establishes a necessary condition for $\mathcal{IC}$ that always holds for quantum mechanics. Note that, differently from (3.2), (3.7) accounts for possible correlations within the initial bits, and the first term considers a more general decoding strategy available to Bob. Additionally, here we introduce a simple refinement of (3.7) to noisy communication scenarios, which, as we will see further, implies the following most robust information criterion for $\mathcal{IC}$ in bipartite scenarios:

$$\begin{aligned}\sum_{i=1}^{n} I(X_i : G_i, M') + \sum_{i=2}^{n} I(X_1 : X_i|G_i, M')\\ \leqslant C + \sum_{i=2}^{n} H(X_i) - H(X_1, ..., X_n).\end{aligned} \tag{3.8}$$

To see that, first, we only need to recall the chain rule and data processing inequality, leading to the respective relations:

$$I(X_1 : X_i|G_i, M') = I(X_1 : X_i, G_i, M') - I(X_1 : G_i, M'), \tag{3.9a}$$

$$I(X_1 : X_i, G_i, M') \leqslant I(X_1 : X_i, \rho_B, M'), \tag{3.9b}$$

$$I(X_i : G_i, M') \leqslant I(X_i : \rho_B, M'). \tag{3.9c}$$

It allows us to rewrite the left-hand side of (3.7) as:

$$\sum_{i=1}^{n} I(X_i : G_i, M') + \sum_{i=2}^{n} I(X_1 : X_i|G_i, M') \leqslant I(X_1 : \rho_B, M') + \sum_{i=2}^{n} I(X_1 : X_i, \rho_B, M'). \tag{3.10}$$

By employing chain-rule once more on $I(X_1 : X_i, \rho_B, M')$, we return to the original shape in (3.7), but in the theory-dependent form,

$$\sum_{i=1}^{n} I(X_i : G_i, M') + \sum_{i=2}^{n} I(X_1 : X_i|G_i, M') \leqslant \sum_{i=1}^{n} I(X_i : \rho_B, M') + \sum_{i=2}^{n} I(X_1 : X_i|\rho_B, M'). \tag{3.11}$$

Assuming the quantum case, we may write the conditional mutual information term explicitly,

$$I(X_1 : X_i|\rho_B, M') = H(X_1, \rho_B, M') + H(X_i, \rho_B, M') - H(X_1, X_i, \rho_B, M') - H(\rho_B, M'), \tag{3.12}$$

It allows us to explore the following entropic relation provided in [22], which is also required in the formal proof of (3.7):

$$H(X_1, X_i, \rho_B, M') \geqslant H(X_1, \cdots, X_n, \rho_B, M') + (n-1)H(X_1, \rho_B, M'). \tag{3.13}$$

Thus, with (3.12) and (3.13) we have in (3.11),

$$\begin{aligned}\sum_{i=1}^{n} I(X_i : G_i, M') + \sum_{i=2}^{n} I(X_1 : X_i|G_i, M') \leqslant \\ \sum_{i=1}^{n} I(X_i : \rho_B, M') + \sum_{i=2}^{n} [H(X_i, \rho_B, M') - H(\rho_B, M')] \\ - H(X_1, \cdots, X_n, \rho_B, M'),\end{aligned} \tag{3.14}$$

that can be simplified by invoking the non-negativity of entropic information measure as follows:

$$\begin{aligned}\sum_{i=1}^{n} I(X_i : G_i, M') + \sum_{i=2}^{n} I(X_1 : X_i|G_i, M') \leqslant \\ \sum_{i=1}^{n} I(X_i : \rho_B, M') + \sum_{i=1}^{n} H(X_i) - H(X_1, \cdots, X_n),\end{aligned} \tag{3.15}$$

At this point, we may simply recall the first mentioned result in the context of the original paper in Eq. (3.1), where we finally achieve the theory-independent criterion in (3.8).

Notice, however, that the term addressing the initial bits follows from an upper bound on the conditional term $I(X_1 : X_i|G_i, M')$. Therefore, for completely uncorrelated initial bits, (3.8) simplifies to:

$$\sum_{i=1}^{n} I(X_i : G_i, M') \leqslant C. \tag{3.16}$$

All current attempts in the literature are limited to this case, and we shall see in the next section that (3.8) sets up the most robust current bipartite criterion, sufficient to witness $\mathcal{IC}$ violations.

Remarkably, *Ref.* [128] characterizes the same scenario exploring non-Shannon-type inequalities. Their computation yields 265 inequalities, among which only 52 are achieved without non-Shannon extra constraints. The non-Shannon-type inequalities significantly enhance the characterization of Shannon's cone. However, the significance of the new set of inequalities still requires a more profound analysis.

3.4 Current understanding of $\mathcal{IC}$

As extensively discussed in the previous section, there are compelling reasons to believe that $\mathcal{IC}$ should hold in Nature. In the present section, we discuss the consequences of assuming $\mathcal{IC}$ by showing which kinds of correlations are forbidden under the current information-theoretic criteria. In Section 3.1, we motivated the $\mathcal{IC}$ statement by highlighting the implications of Popescu-Rohrlich (PR) boxes in the context of Random Access Codes (RACs). Specifically, the protocol (2.18) from subsection 2.2.1 trivializes the RAC $n \mapsto 1$ using PR boxes, and clearly violates the original criterion (3.2), *i.e.*, $I(X_1 : G_1) = \cdots = I(X_{2^n} : G_{2^n}) = 1$. More interestingly, however, in *Ref.* [20] the authors demonstrate for the noiseless case that such a protocol achieves a violation of the inequality (3.2) whenever their correlations respect $E_I^2 + E_{II}^2 > 1$, where $E_j = 2P_j - 1$ is defined in terms of the conditional probabilities $p(a, b|x, y)$ as

$$P_I = \frac{1}{2}[p(a \oplus b = 0|0,0) + p(a \oplus b = 0|1,0)]; \tag{3.17a}$$

$$P_{II} = \frac{1}{2}[p(a \oplus b = 0|0,1) + p(a \oplus b = 1|1,1)]. \tag{3.17b}$$

Consequently, it defines another necessary condition for $\mathcal{IC}$ to hold for non-signaling correlations, which is given by [20]:

$$E_I^2 + E_{II}^2 \leqslant 1. \tag{3.18}$$

In the Appendix F we provide a ganeralized formal proof of (3.18) in multipartite scenarios, which naturally includes the proof for the bipartite case. This constraint, (3.18), is equivalent to the bipartite quadratic Bell inequality, the so-called Uffink's inequality [129]. Interestingly, however, as will be demonstrated further, this equivalence does not hold in multipartite scenarios.

Of particular relevance, this mapping from the $\mathcal{IC}$ inequality (3.2) to Uffink's inequality (3.18) establishes the most significant result related to $\mathcal{IC}$. Indeed, it proves that any correlation beyond Tsirelson's limit for the Clauser-Horne-Shimony-Holt (CHSH) inequality [2] will violate the information causality principle and thus witness its incompatibility with quantum theory. More precisely, as proven by Tsirelson [16], the classically valid CHSH inequality

$$\text{CHSH} = \langle A_0B_0\rangle + \langle A_0B_1\rangle + \langle A_1B_0\rangle - \langle A_1B_1\rangle \leqslant 2, \tag{3.19}$$

achieves a maximum value in quantum theory of $\text{CHSH}_Q = 2\sqrt{2}$. A PR box leads to $\text{CHSH}_{NS} = 4$. A direct analysis of (3.18) shows that any distribution achieving $\text{CHSH} > \text{CHSH}_Q$ violates Uffink's inequality [20] and thus has its post-quantum nature witnessed by the $\mathcal{IC}$ principle.

Later developments have also compared $\mathcal{IC}$ with other DI principles, showing $\mathcal{IC}$ as the proposal that better approximates the set of quantum correlations. *Ref.* [96], in particular, compares $\mathcal{IC}$ to the *macroscopic locality* principle introduced in *Ref.* [19]. In this case, they introduce the RACs protocol for d dimensional variables (see Section 2.2.1), generalizing the concatenation procedure from [20], to investigate the following

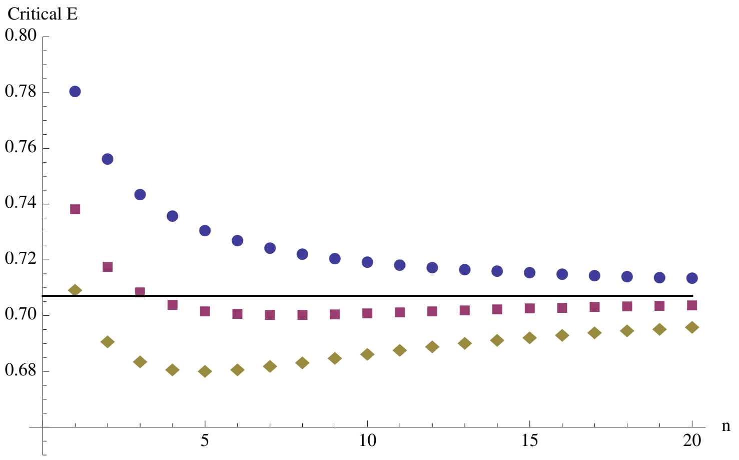


Figure 9 – Results from [96]. Critical E for which (3.2) ceases to be violated, as a function of n, and for different values of d ($d = 2$ circles, $d = 5$ squares, $d = 10$ diamonds). The solid line corresponds to $1/\sqrt{2}$, which, according to numerical calculations under numerical precision, coincides with macroscopic locality bound for $d \leqslant 5$.

isotropic box,

$$\mathrm{P}_E(a,b|x,y) \equiv E\mathrm{PR}_d(a,b|x,y) + (1-E)\frac{1}{d^2}, \tag{3.20}$$

where E controls the nonlocal behavior of P_E, and PR_d denotes the PR-box version to $(d2dd)$-Bell scenarios,

$$\mathrm{PR}_d(a,b|x,y) \equiv \begin{cases} 1/d & \text{if } \; x\cdot y = (b-a)\mathrm{mod} d \\ 0 & \text{otherwise.} \end{cases} \tag{3.21}$$

The results of this study are summarized in Fig.9 that shows the critical values of E as a function of the number of copies n, for which the correlation (3.20) respects the criterion (3.2) in the noiseless communication scenario. In accordance with the original result for the simplest scenario, $d = 2$, when n increases, E asymptotically approximates to the quantum boundary, $E_T = 1/\sqrt{2} \approx 0.707$, which also coincides with the bound implied the *macroscopic locality*. For $d > 2$, however, E_T no longer defines the maximum value for quantum correlations [96]. At the same time, it continues to determine the range for isotropic correlations respecting macroscopic locality [4]. Naturally, the authors prove the inequivalence between $\mathcal{IC}$ and macroscopic locality by showing that if the correlations observed in nature are all those satisfying the macroscopic locality principle, $\mathcal{IC}$ would be violated.

In the same direction, further proposals introduce a physically motivated slight modification on the definitions of the quantum correlations that outer approximate quantum set, named *almost quantum* set [130]. Remarkably, the almost quantum set can be efficiently characterized by means of semidefinite programming techniques, and, for bipartite scenarios, the almost quantum set coincides with the set Q_{1+AB} of the

[4] It might be computationally verified through the Q_1 set [19], via the NPA hierarchy [40], introduced in the Appendix E.

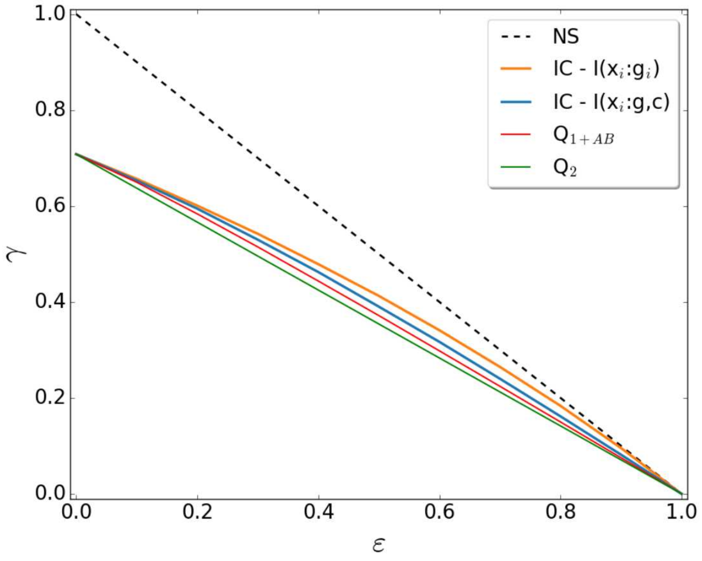


Figure 10 – Non-signaling (NS) polytope slice given by Eq.(3.22). Every dot above these curves violates the respective criterion represented. The black dashed line describes the NS edge. The green and red solid lines describe the outer approximations of the quantum set, given respectively by level 2 and $1 + AB$ of NPA hierarchy [40] (the latter implies the edge for almost quantum). The limits defined by criteria (3.18) and (3.7) are respectively depicted with the blue and orange solid lines.

NPA hierarchy [40]. Intriguingly, it was shown that all DI-principles[5], except $\mathcal{IC}$, allow correlations beyond the almost quantum set, and consequently are unable to recover the set of quantum correlations completely. In this sense, $\mathcal{IC}$ remains the strongest DI principle in prohibiting correlations beyond the quantum boundary.

Despite the remarkable achievements, such as the emergence of the Tsirelson's bound from (3.2), it remains unclear whether all correlations beyond the quantum set may violate $\mathcal{IC}$. In this regard, while Tsirelson's bound, $\text{CHSH}_Q = 2\sqrt{2}$, refers to a specific region of the no-signaling set, a more comprehensive analysis in [21] investigates the $\mathcal{IC}$ criterion (3.18) for different slices of the non-signaling polytope. Of particular relevance, they present the following slice,

$$p_{\gamma,\varepsilon}(a,b|x,y) = \gamma p_{\text{PR}}(a,b|x,y) + \varepsilon p_{\text{L}}(a,b|x,y) + (1-\gamma-\varepsilon)\frac{1}{4}, \tag{3.22}$$

where $p_{\text{PR}}(a,b|x,y) = \delta_{a\oplus b,xy}/2$, $p_{\text{L}}(a,b|x,y) = \delta_{a,0}\delta_{b,0}$, and different edges are depicted in Fig.10. This analysis shows a clear discrepancy between the quantum boundary and the $\mathcal{IC}$ condition (3.18), which leaves the question of whether $\mathcal{IC}$ may completely single out the set of quantum correlations open. It is important to stress that criterion (3.18) assumes the particular concatenation protocol (see Appendix F), and it might still be the case that more suitable communication strategies would more efficiently reveal $\mathcal{IC}$ violations by correlations undetected by (3.18). In this direction, further

[5] At that moment, they were Non-trivial Communication Complexity [18], No Advantage for Nonlocal Computation [131], Information Causality [20], Macroscopic Locality [19], and Local Orthogonality [132].

developments have improved (3.18) boundary by employing nonlocality distillation procedures [133], which, however, was not sufficient to close the gap between the criterion (3.18) and the quantum set. In particular, more recent developments have shown that certain device-independent (DI) principles fail to identify the quantum set in specific regions of the no-signaling polytope, named as *quantum voids* [134, 135]. These are regions where the quantum edge coincides with the boundary of the classical polytope.

Alternatively, while looking for a suitable communication protocol consists of a high-complexity problem, as previously discussed, latter refinements propose more suitable criteria to better witness $\mathcal{IC}$ violation [115, 116, 22, 23]. Notably, Chaves *et al.*, in *Ref.* [22] introduces a systematic general framework to derive more suitable criteria, such as (3.6). In this regard, recent studies demonstrate that incorporating noisy communication into the $\mathcal{IC}$ scenario provides more suitable information criteria, significantly reducing the complexity of detecting $\mathcal{IC}$ violation. Indeed, *Ref.* [112] shows that the noisy criterion (3.2), for the simplest scenarios with only two initial bits for Alice, enhances every bound on no-signaling correlations achieved through the concatenation procedure. Specifically, this approach recovers Tsirelson's bound and improves the critical values in Fig.9. This is accomplished by assessing criterion (3.2) for different noisy communication channels. In this sense, evaluating scenarios with poor transmission channels enables more efficient witnessing of the nonlocality assistance on the communication, and consequently $\mathcal{IC}$ violation. In fact, all results of *Ref.* [112] are achieved by taking the limit[6] $C \to 0$. This effect is illustrated in Fig.10, where the boundary verified by (3.18) is completely recovered through the noisy criterion (3.2) for $n = 2$. Interestingly, here we show that (3.8) provides even stronger constraints. Fig 10, presents (3.8) for the simplest case with two initial bits for Alice, and we verify that the new inequality is significantly more powerful than the multiple copies criterion (3.18), and the original (3.2), across the entire region (3.22).

Despite remaining unclear whether $\mathcal{IC}$ may fully distinguish the set of quantum correlations, the noisy framework has already enabled significant advancements on the $\mathcal{IC}$ understanding [136, 23, 113, 30]. More fundamentally, *Ref.* [136] shows that $\mathcal{IC}$ offers some meaning for the composition rule of quantum mechanics when evaluated in the context of general probabilistic theories. Moreover, as we shall see in the next Chapter, $\mathcal{IC}$ also recovers the monogamy of Bell inequalities violation implied by quantum theory [30], which states straight connections of $\mathcal{IC}$ with security on cryptographic key distribution. Interestingly, recent discoveries also propose efficient $\mathcal{IC}$-based techniques to obtain polynomial inequalities approximating the set of quantum correlations for arbitrary Bell scenarios [137, 113]. Naturally, the recent breakthrough of the noisy framework suggests the potential for further enhancement by combining the concatenation protocol and noisy communication frameworks. However, our current preliminary analyses show no improvement in this direction.

Due to the inherently bipartite structure of $\mathcal{IC}$ formulation, it is essential to stress that further developments show that, in order to describe the whole set of quantum nonlocal correlations correctly, quantum mechanics requires an intrinsic multipartite structure to information principles [24]. More specifically, consider correlations from a Bell scenario with three parties specified by $p(a, b, c|x, y, z)$. To test whether the correlation satisfies any bipartite operational principles, $p(a, b, c|x, y, z)$ must be *locally* post-processed into

[6] Obviously, it necessarily requires $C > 0$.

an effectively bipartite correlation $\tilde{p}(a', b'|x', y')$. This process, commonly referred to as *wiring* in the literature, is crucial for such an analysis. In this context, *Ref.* [24] proves the existence of tripartite supra-quantum correlations for which the wiring procedure invariably results in a classical bipartite effective correlation $\tilde{p}(a', b'|x', y')$. Consequently, no bipartite formulated principle can witness such a post-quantumness as the case of $\mathcal{IC}$ and non-trivial communication complexity [18]. This finding is further supported by subsequent research, which identifies classes of extremal tripartite correlations that exceed quantum limits but do not violate any bipartite principle via wiring [25]. Interestingly, this conclusion aligns with more recent works showing that no bipartite nonlocal causal theory can explain quantum mechanics correlations [138]. Thus, the need for multipartite formulations of the information causality principle becomes evident. This observation sets the stage for the next section, where we explore whether $\mathcal{IC}$ can rule out implausible consequences arising from multipartite correlations beyond the quantum set.

3.5 Multipartite $\mathcal{IC}$

Our first goal is to introduce a natural generalization of the bipartite $\mathcal{IC}$ framework to accommodate the multipartite scenario. For that, we will closely follow the previously mentioned informational-geometric approach (see Appendix C). Specifically, we consider a particular class of quantum causal structures that naturally generalize the known bipartite scenario: Consider N parts, among which $N-1$ are senders in possession of their respective bit-strings $\mathbf{x}^k = (X_1^k, X_2^k, \cdots, X_n^k)$, where $k \in \{1, 2, \cdots, N-1\}$. Each sender encodes a classical message M_k of size $|M_k| < n$ to the N^{th}-part, the receiver who has to compute one out of n possible bits functions $f_j(X_j^1, X_j^2, \cdots, X_j^{N-1})$, by producing the guess G_j, where $j \in \{1, \cdots, n\}$. The receiver, in turn, receives $\mathbf{M'} = (M'_1, M'_2, \cdots, M'_{N-1})$, which denotes all messages reaching after passing through a classical noisy channel of capacity $C_k \leqslant |M_k|$. This scenario is illustrated as a directed acyclic

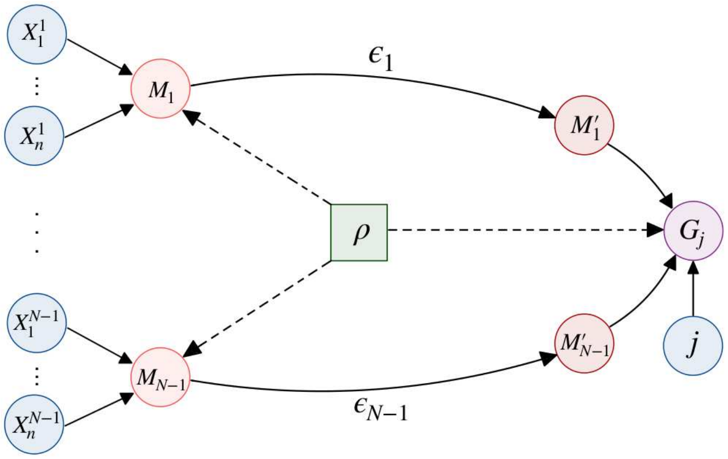


Figure 11 – Quantum causal structure, described as a DAG, associated with the multipartite information causality scenario. ϵ_i over the arrow of the message M_i denotes the effect of the classical noisy channel through which the respective message is sent.

graph (DAG) in Fig. 11, where ϵ_k denotes the effect of the noisy channel through which the respective message is sent.

For the sake of completeness, the $\mathcal{IC}$ statement may be straightly extended for this multipartite scenario as follows:

> *"The receiver's available information about the initial bits of each $N-1$ senders, considering all their possible local as well as pre-established shared resources, cannot be greater than the number of effectively received bits through a single use of a classical channel of capacity C_k."*

Nevertheless, as it will become evident shortly, the same extension does not happen for the operational bipartite criteria presented in Section 3.2.

To illustrate, consider the tripartite scenario, such that Alice and Bob have just two initial uncorrelated bits and that the communication task of Charlie is to compute two specific functions $f_1 = x_1^1 \oplus x_1^2$ and $f_2 = x_2^1 \oplus x_2^2$. The communication task is trivialized when the parties share the following extremal tripartite non-signaling (post-quantum)

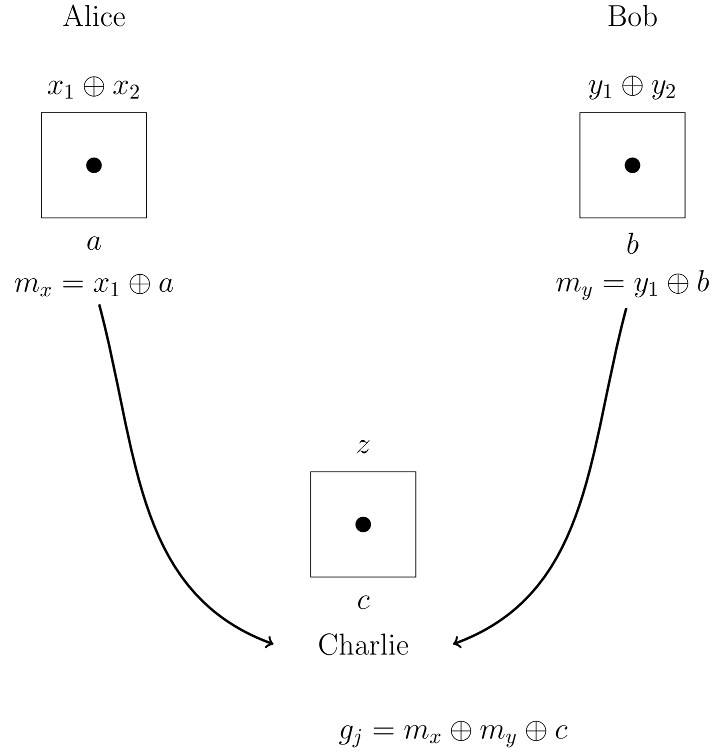


Figure 12 – The communication protocol is performed by Alice, Bob, and Charlie, who share a non-signaling resource. Alice (Bob) receives initially two bits $\{x_1, x_2\}$ ($\{y_1, y_2\}$) and perform her local measurements as $x = x_1 \oplus x_2$ ($y = y_1 \oplus y_2$). After obtaining her outputs a (b), encodes the message as $m_x = a \oplus x_1$ ($m_y = b \oplus y_1$). Charlie inputs on his side $z = 0$ if he wants to compute f_1, and $z = 1$ if he wants to compute f_2. After receiving the messages, Charlie computes his guess by following $g_j = m_x \oplus m_y \oplus c$.

correlation [44],

$$p(a,b,c|x,y,z) = \begin{cases} 1/4 & \text{if} \quad a \oplus b \oplus c = xz \oplus yz; \\ 0 & \text{else,} \end{cases} \tag{3.23}$$

where $a, b, c, x, y, z \in \{0, 1\}$. The parties perform the protocol detailed in Fig. 12 to achieve it. Charlie can always perfectly compute each function in each run since $g_1 = x_1^1 \oplus x_1^2$ and $g_2 = x_2^1 \oplus x_2^2$. In other words, similar to the usual $\mathcal{IC}$ scenario, Charlie has potential access to the four bits of Alice and Bob but receives only two bits of information communicated by them. Note, however, that all the bipartite criteria fail to witness such a clear violation of the principle. For example, by simply applying (3.2) for the respective bipartitions Alice and Charlie and Bob and Charlie, this protocol yields for all informational terms $I(X_i^k : G_j) = 0$, which implies no violation of (3.2). This fact, together with the previously mentioned requirements for multipartite formulation of informational principles [24, 25], motivates the search for more suitable operational criteria for $\mathcal{IC}$.

As detailed in the Appendix C, the entropic-geometric approach introduced in [22] offers a general framework to derive information-theoretic constraints, given the causal relations among the variables. Thus, we employ such a method in order to derive suitable constraints for the multipartite scenario in Fig.11. Analogous to the discussions in Section 3.3, it is imperative to consider all causal relations defining the quantum causal structure in Fig.11. For simplicity, consider $N = 3, n = 2$, and the noiseless case where $M_k = M'_k$ (we will see soon that we might easily extend to the noisy case). In this scenario, we identify the following set of random variables: $\{X_1^1, X_2^1, X_1^2, X_2^2, M_1, M_2, G_1, G_2, \rho\}$, for which the causal relations translate as:

$$I(X_1^1, X_2^1, X_1^2, X_2^2 : \rho) = 0, \tag{3.24a}$$

$$I(X_1^1 : X_1^2, X_2^2) = 0, \tag{3.24b}$$

$$I(X_2^1 : X_1^2, X_2^2) = 0. \tag{3.24c}$$

In the first constraint, we assume the independence among the sender's initial bits and the resource correlating the parties. In addition, (3.24a) together with (3.24b) and (3.24c) state that, initially, the senders are the only parties accessing their data[7]. In a fully classical context, the additional conditional independence completing the causal structure characterization is $I(X_1^1, X_2^1, X_1^2, X_2^2 : G_0, G_1 | M_1, M_2, \rho) = 0$. However, in the quantum case, Charlie's guesses, G_j, may lack a joint description with his subsystem, $\rho_C = \mathrm{Tr}_{AB}(\rho)$, thereby preventing its inclusion. Hence, we must encode the remaining causal relations regarding the data processing inequalities. Denoting, respectively, Alice

[7] *i.e.*, Alice and Bob sets of data have no correlation.

and Bob subsystems by ρ_A and ρ_B, the sets of jointly observable variables is written as

$$\begin{aligned}
S_1 &= \{X_1^1, X_2^1, X_1^2, X_2^2, \rho_A, \rho_B, \rho_C\}, \\
S_2 &= \{X_1^1, X_2^1, X_1^2, X_2^2, M_1, \rho_B, \rho_C\}, \\
S_3 &= \{X_1^1, X_2^1, X_1^2, X_2^2, M_2, \rho_A, \rho_C\}, \\
S_4 &= \{X_1^1, X_2^1, X_1^2, X_2^2, M_1, M_2, \rho_C\}, \\
S_5 &= \{X_1^1, X_2^1, X_1^2, X_2^2, M_1, M_2, G_1\}, \\
S_6 &= \{X_1^1, X_2^1, X_1^2, X_2^2, M_1, M_2, G_2\}.
\end{aligned} \tag{3.25}$$

This formulation implies that the only relevant data processing inequalities are of the forms:

$$I(X_1^1, X_2^1, X_1^2, X_2^2, \rho_B, \rho_C : \rho_A) \geqslant I(X_1^1, X_2^1, X_1^2, X_2^2, \rho_B, \rho_C : M_1), \tag{3.26a}$$

$$I(X_1^1, X_2^1, X_1^2, X_2^2, \rho_A, \rho_C : \rho_B) \geqslant I(X_1^1, X_2^1, X_1^2, X_2^2, \rho_A, \rho_C : M_2), \tag{3.26b}$$

$$I(X_1^1, X_2^1, X_1^2, X_2^2, \rho_C : \rho_A, \rho_B) \geqslant I(X_1^1, X_2^1, X_1^2, X_2^2, \rho_C : M_1, M_2), \tag{3.26c}$$

$$I(X_1^1, X_2^1, X_1^2, X_2^2 : \rho_A, \rho_B, \rho_C) \geqslant I(X_1^1, X_2^1, X_1^2, X_2^2, \rho_C : M_1, M_2, G_j). \tag{3.26d}$$

By performing the Fourier-Motzkin elimination for the most general marginal scenario, $\mathcal{M} = \{\{X_1^1, X_2^1, X_1^2, X_2^2, M_1, M_2, G_1\}, \{X_1^1, X_2^1, X_1^2, X_2^2, M_1, M_2, G_1\}\}$, the procedure yields several entropic inequalities, capturing the most general form of the informational and causal constraints within the communication scenario, holding for quantum theory. Of particular relevance, the following criterion achieved by this procedure reflects the supra-quantum feature provided by the correlation (3.23) in the protocol of Fig.11:

$$\mathcal{I} = I(X_1^1 : X_1^2, G_1) + I(X_2^1 : X_2^2, G_2) + I(X_1^2 : X_1^1, G_1) + I(X_2^2 : X_2^1, G_2) \leqslant H(M_1, M_2). \tag{3.27}$$

Indeed, (3.23) maximally violates (3.27), since in this case $\mathcal{I} = 4$, while $H(M_1, M_2) = 2$. For computational details concerning (3.27), the reader is referred to the Appendix D.

3.5.1 Analytical generalization

It is noteworthy that the Eq. (3.27) can be generalized to scenarios with an arbitrary number of parties N, bits n, and noisy communication specified by C_k. In that case, we analyze the quantity $I(\mathbf{x}^k : \mathbf{x}^1, \cdots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \cdots, \mathbf{x}^{N-1}, \mathbf{M}', c)$, measuring the knowledge of the entire network about the data set of the part k. The main foundational components of this proof are identical to the original axioms assumed in [20]:

1. Consistency: Whenever the random variables denote classical systems, I reduces to the classical mutual information;

2. Data processing:

$$I(A : B') \leqslant I(A : B), \quad \text{when} \quad B \longrightarrow B'; \tag{3.28}$$

3. Chain-rule:

$$I(A:B|D) = I(A:B,D) - I(A:D); \tag{3.29}$$

4. Symmetry: $I(A:B) = I(B:A)$;

5. Non-negativity: $I(A:B) \geqslant 0$.

In that sense, analogously to [20], the parameter I is independent of any underlying physical theory. Similarly to Bell inequalities, I depends only on the party's classical variables. Therefore, without the need to specify the form of I, it holds for every informational measure holding the five mentioned properties[8]:

$$\sum_{k=1}^{N-1}\sum_{i=1}^{n} I(X_i^k : X_i^1,\ldots,X_i^{k-1},X_i^{k+1},\ldots,X_i^{N-1},\mathbf{M}',G_i) \\ \leqslant \sum_{k=1}^{N-1} C_k + \sum_{k=1}^{N-1}\sum_{i=1}^{n} I(X_{i+1}^k,\ldots,X_n^k : X_i^k), \tag{3.30}$$

Proof. By applying the chain rule (3.29) two times, we obtain:

$$\begin{aligned} I(\mathbf{x}^k &: \mathbf{x}^1,\cdots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\cdots,\mathbf{x}^{N-1},\mathbf{M}',c) \\ &= I(X_1^k : \mathbf{x}^1,\cdots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\cdots,\mathbf{x}^{N-1},\mathbf{M}',c) \\ &\quad + I(X_2^k,\cdots,X_n^k : \mathbf{x}^1,\cdots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\cdots,\mathbf{x}^{N-1},\mathbf{M}',c|X_1^k), \end{aligned} \tag{3.31}$$

$$\begin{aligned} &= I(X_1^k : \mathbf{x}^1,\cdots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\cdots,\mathbf{x}^{N-1},\mathbf{M}',c) \\ &\quad + I(X_2^k,\cdots,X_n^k : \mathbf{x}^1,\cdots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\cdots,\mathbf{x}^{N-1},\mathbf{M}',c,X_1^k) \\ &\quad\quad - I(X_2^k,\cdots,X_n^k : X_1^k). \end{aligned} \tag{3.32}$$

From the data processing (3.28), we have

$$I(X_2^k,\cdots,X_n^k : \mathbf{x}^1,\cdots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\cdots,\mathbf{x}^{N-1},\mathbf{M}',c,X_1^k) \geqslant \\ I(X_2^k,\cdots,X_n^k : \mathbf{x}^1,\cdots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\cdots,\mathbf{x}^{N-1},\mathbf{M}',c). \tag{3.33}$$

Furthermore, by applying the chain rule in the first term in the right-hand side of (3.31), and using strong subadditivity, $I(A:B|C) \geqslant 0$, we obtain:

$$\begin{aligned} I(X_1^k &: \mathbf{x}^1,\cdots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\cdots,\mathbf{x}^{N-1},\mathbf{M}',c) \\ &= I(X_1^k : X_2^1,\cdots,X_n^1|X_1^1,\mathbf{x}^2,\cdots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\cdots,\mathbf{x}^{N-1},\mathbf{M}',c) \\ &\quad + I(X_1^k : X_1^1,\mathbf{x}^2,\cdots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\cdots,\mathbf{x}^{N-1},\mathbf{M}',c) \\ &\quad\quad \geqslant I(X_1^k : X_1^1,\mathbf{x}^2,\cdots,\mathbf{x}^{k-1},\mathbf{x}^{k+1},\cdots,\mathbf{x}^{N-1},\mathbf{M}',c). \end{aligned} \tag{3.34}$$

[8] It is important to stress the result of [114], showing the set of five axioms might be reduced to only two.

Therefore, revisiting Eq. (3.31), we write:

$$\begin{aligned}
&I(\mathbf{x}^k : \mathbf{x}^1, \cdots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \cdots, \mathbf{x}^{N-1}, \mathbf{M}', c) \\
&\qquad \geqslant I(X_1^k : X_1^1, \mathbf{x}^2, \cdots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \cdots, \mathbf{x}^{N-1}, \mathbf{M}', c) \\
&\qquad\quad + I(X_2^k, \cdots, X_n^k : \mathbf{x}^1, \cdots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \cdots, \mathbf{x}^{N-1}, \mathbf{M}', c) \\
&\qquad\qquad - I(X_2^k, \cdots, X_n^k : X_1^k).
\end{aligned} \tag{3.35}$$

Similarly to (3.34), we can employ the chain rule and strong subadditivity $N-3$ times in the first right-hand side term in (3.35) in order to highlight only the first bit X_1^k of each bit-string $\mathbf{x}^k$:

$$\begin{aligned}
&I(\mathbf{x}^k : \mathbf{x}^1, \cdots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \cdots, \mathbf{x}^{N-1}, \mathbf{M}', c) \\
&\qquad \geqslant I(X_1^k : X_1^1, X_1^2, \cdots, X_1^{k-1}, X_1^{k+1}, \cdots, X_1^{N-1}, \mathbf{M}', c) \\
&\qquad\quad + I(X_2^k, \cdots, X_n^k : \mathbf{x}^1, \cdots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \cdots, \mathbf{x}^{N-1}, \mathbf{M}', c) \\
&\qquad\qquad - I(X_2^k, \cdots, X_n^k : X_1^k).
\end{aligned} \tag{3.36}$$

Notice that the right-hand side third term in (3.36), apart of X_1^k from the bit-string x^k, is precisely $I(\mathbf{x}^k : \mathbf{x}^1, \cdots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \cdots, \mathbf{x}^{N-1}, \mathbf{M}', c)$. Therefore, by performing the same steps $n-1$ times, we achieve:

$$\begin{aligned}
&I(\mathbf{x}^k : \mathbf{x}^1, \cdots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \cdots, \mathbf{x}^{N-1}, \mathbf{M}', c) \\
&\geqslant \sum_i^n I(X_i^k : X_i^1, X_i^2, \cdots, X_i^{k-1}, X_i^{k+1}, \cdots, X_i^{N-1}, \mathbf{M}', c) - \sum_i^n I(X_{i+1}^k, \cdots, X_n^k : X_i^k).
\end{aligned} \tag{3.37}$$

Given that $\mathbf{M}'$ are classical variables, we use the data processing inequality (3.28) to refine the above inequality as,

$$\begin{aligned}
&I(\mathbf{x}^k : \mathbf{x}^1, \cdots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \cdots, \mathbf{x}^{N-1}, \mathbf{M}', c) \\
&\quad \geqslant \sum_{i=1}^n I(X_i^k : X_i^1, X_i^2, \cdots, X_i^{k-1}, X_i^{k+1}, \cdots, X_i^{N-1}, \mathbf{M}', G_i) - \sum_{i=1}^n I(X_{i+1}^k, \cdots, X_n^k : X_i^k).
\end{aligned} \tag{3.38}$$

The next step is then decomposing $\mathbf{M}'$ into $M_1', \cdots, M_k', \cdots, M_{N-1}'$, and deriving the upper bound for $I(\mathbf{x}^k : \mathbf{x}^1, \cdots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \cdots, \mathbf{x}^{N-1}, \mathbf{M}', c)$ by applying the chain rule. In this case,

$$\begin{aligned}
&I(\mathbf{x}^k : \mathbf{x}^1, \cdots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \cdots, \mathbf{x}^{N-1}, M_1', \cdots, M_k', \cdots, M_{N-1}', c) \qquad (3.39) \\
&\quad = I(\mathbf{x}^k : M_k' | \mathbf{x}^1, \cdots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \cdots, \mathbf{x}^{N-1}, M_1', \cdots, M_{k-1}', M_{k+1}', \cdots, M_{N-1}', c) \\
&\qquad + I(\mathbf{x}^k : \mathbf{x}^1, \cdots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \cdots, \mathbf{x}^{N-1}, M_1', \cdots, M_{k-1}', M_{k+1}', \cdots, M_{N-1}', c).
\end{aligned}$$

The second term on the right-hand side vanishes due to the no-signaling assumption. Applying the chain rule to the remaining term and using the non-negativity of mutual information $I(A : B) \geqslant 0$, we get

$$I(\mathbf{x}^k : \mathbf{x}^1, \cdots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \cdots, \mathbf{x}^{N-1}, M'_1, \cdots, M'_k, \cdots, M'_{N-1}, c) \\ \leqslant I(M'_k : \mathbf{x}^1, \cdots, \mathbf{x}^k, \cdots, \mathbf{x}^{N-1}, M'_1, \cdots, M'_{k-1}, M'_{k+1}, \cdots, M'_{N-1}, c). \quad (3.40)$$

Applying the data processing inequality (3.28) again, we know that including M_k on the right-hand side can only increase the mutual information,

$$I(\mathbf{x}^k : \mathbf{x}^1, \cdots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \cdots, \mathbf{x}^{N-1}, M'_1, \cdots, M'_k, \cdots, M'_{N-1}, c) \\ \leqslant I(M'_k : \mathbf{x}^1, \cdots, \mathbf{x}^k, \cdots, \mathbf{x}^{N-1}, M'_1, \cdots, M'_{k-1}, M'_{k+1}, \cdots, M'_{N-1}, c, M_k). \quad (3.41)$$

From the causal structure depicted in Figure 11, it is clear that M_k shields M'_k from all other variables $\mathbf{V}$, such that $I(M'_k : \mathbf{V}|M_k) = I(M'_k : \mathbf{V}, M_k) - I(M'_k : M_k) = 0$. Thus, we simplify the inequality (3.41) as,

$$I(\mathbf{x}^k : \mathbf{x}^1, \cdots, \mathbf{x}^{k-1}, \mathbf{x}^{k+1}, \cdots, \mathbf{x}^{N-1}, M'_1, \cdots, M'_k, \cdots, M'_{N-1}, c) \\ \leqslant I(M'_k : M_k) = C_k. \quad (3.42)$$

Finally, by combining (3.38) and (3.42), and subsequently summing over all k, we derive a necessary condition (3.30) ensuring $\mathcal{IC}$:

$$\sum_{k=1}^{N-1} \sum_{i=1}^{n} I(X_i^k : X_i^1, \ldots, X_i^{k-1}, X_i^{k+1}, \ldots, X_i^{N-1}, \mathbf{M}', G_i) \\ \leqslant \sum_{k=1}^{N-1} C_k + \sum_{k=1}^{N-1} \sum_{i=1}^{n} I(X_{i+1}^k, \ldots, X_n^k : X_i^k).$$

□

The derived criterion (3.30) holds for any number of parties N, for an arbitrary number of bits n, with communication through a classical channels characterized by the capacities C_k. Notice that (3.30) builds up a more stringent constraint than (3.27), even when $N = 3$, $n = 2$ with noiseless communication ($M_k = M'_k$), since it includes the message into account for the receiver's available information. Nevertheless, naturally, (3.27) trivially follows from (3.30) as $I(A : B, C) \geqslant I(A : B)$. Moreover, it is important to stress that criterion (3.30) is written in the most general form for $\mathcal{IC}$, wherein the upper bound on the receiver's available information includes noisy communication description. In this case, a given resource explored by the parties violates $\mathcal{IC}$ whenever there exists a protocol and noisy communication channels characterized by C_k, such that (3.30) is violated. Indeed, previous developments in *Ref.* [112] have shown that such framework consists of the strongest form of $\mathcal{IC}$ in bounding the set of no-signaling correlations, and we shall see in the next section that such claim also holds for the multipartite case.

The multipartite formulation expressed in (3.30) can be violated by the multipartite extension of the post-quantum correlation (3.23), given by

$$p(a_1, a_2, \cdots, a_N | x_1, x_2, \cdots, x_N) = \begin{cases} 1/2^{N-1} & \text{if} \quad \bigoplus_{k=1}^{N} a_k = \bigoplus_{k=1}^{N-1} x_k x_N; \\ 0 & \text{else.} \end{cases} \tag{3.43}$$

Considering $n = 2$, $f_j = X_j^1 \oplus X_j^2 \oplus \cdots \oplus X_j^{N-1}$, alongside the direct extension of the protocol described in Fig. 12 for the multipartite case, we see that the communication task is trivialized, implying the maximal violation of the multipartite $\mathcal{IC}$ inequality (3.30).

3.5.2 Concatenation procedure

As previously discussed, the first proposal for the information causality criterion, (3.2), witnessed the post-quantum nature of all non-signaling correlations beyond Tsirelson's bound [20]. For that, however, it was essential to consider a concatenation procedure involving many copies of the correlation under test. Here, we show how such concatenation can be constructed for the tripartite scenario and generalize it to arbitrary multipartite scenarios.

Similarly to the bipartite scenario, the success probability for the protocol in Fig.(12) can be connected to the probability of the resource shared between the parts, more specifically to the probability $p(a \oplus b \oplus c = xz \oplus yz | x, y, z)$. The probabilities of Charlie correctly computing the values of $x_1 \oplus y_1$ and $x_2 \oplus y_2$ are, respectively,

$$\begin{aligned} P_I = \frac{1}{4}[&p(a \oplus b \oplus c = 0|0,0,0) \\ &+p(a \oplus b \oplus c = 0|0,1,0) \\ &+p(a \oplus b \oplus c = 0|1,0,0) \\ &+p(a \oplus b \oplus c = 0|1,1,0)]; \end{aligned} \tag{3.44a}$$

$$\begin{aligned} P_{II} = \frac{1}{4}[&p(a \oplus b \oplus c = 0|0,0,1) \\ &+p(a \oplus b \oplus c = 1|0,1,1) \\ &+p(a \oplus b \oplus c = 1|1,0,1) \\ &+p(a \oplus b \oplus c = 0|1,1,1)]. \end{aligned} \tag{3.44b}$$

In the particular case where the parties share the correlation described by (3.23), we obtain $P_I = P_{II} = 1$.

In Fig. 13, we specify the concatenation procedure for the tripartite communication protocol of Fig. 12. In this case, Alice and Bob initially receive the respective bit-strings $\mathbf{x}$ and $\mathbf{y}$ of length $n = 2^K$ and share $2^K - 1$ identical copies of binary-input/binary-output non-signaling boxes with Charlie. The success probability that Charlie produces a guess g_j correctly is given by (see appendix F)

$$p(g_j = x_j \oplus y_j) = \frac{1}{2}(1 + E_I^{K-r} E_{II}^r), \tag{3.45}$$

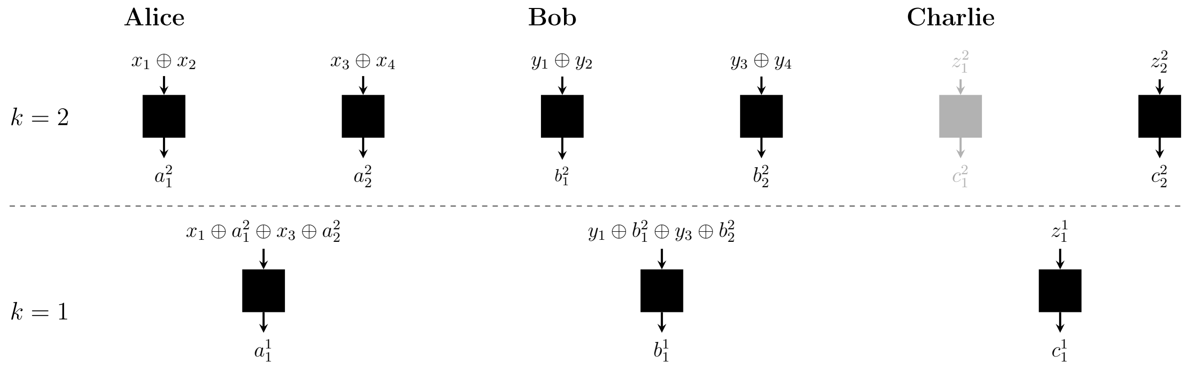


Figure 13 – Concatenation performed by Alice, Bob, and Charlie of the protocol in Fig.12. Alice and Bob initially receive $n = 2^K$ bits and Charlie receives a K-bitstring $\{z_1, z_2, ..., z_K\}$, which indicates which pair $x_j \oplus y_j$ he is interested in, $j = \sum_{l=1}^{K} z_l 2^{l-1}$. Thus, Alice and Bob encode their bits in pairs, following the protocol in Fig.12. In this case, instead of sending each respective message, they encode pairs of these in other identical NS boxes with the same strategy. So, both Alice and Bob perform this procedure until one message remains. Alice and Bob are then allowed to send these one-bit messages to Charlie, who receives the message, and to each NS box, performs the decoding protocol just as Fig.12. In a given concatenation level k, Charlie recovers the sum of Alice and Bob's messages previously encoded in the current box, which is associated with a subsequent higher level $k+1$ NS-box. The picture shows a particular case with $n = 4$, where a_i^k, b_i^k, and c_i^k represent the output of the box i in the level k to Alice, Bob, and Charlie, respectively. In the level $k = 1$, Charlie recovers the messages associated with the box $i = 1$ of the level $k = 2$ and can recover $x_3 \oplus y_3$ or $x_4 \oplus y_4$, depending on z_2^2.

where r denotes the number of times that Charlie measures $z = 1$ in the K levels of the concatenation code displayed in Fig.13 and $E_i = 2P_i - 1$ (see Eq.(3.44)). By considering this success probability, we show in Appendix F.1 that $\mathcal{IC}$ is always violated when $E_I^2 + E_{II}^2 > 1$. In other words, when combined with a concatenation procedure and multiple copies of the behavior under test, the tripartite information causality inequality (3.27) leads to a generalization of the bipartite inequality (3.18), given by

$$E_I^2 + E_{II}^2 \leqslant 1. \tag{3.46}$$

Analogous to (3.27), the multiple copies criterion (3.46) is maximally violated by the behavior (3.23) since, for this case, $E_I = E_{II} = 1$. Moreover, for isotropic correlations described by a visibility parameter E and such that $E_I = E_{II} = E$, the tripartite multiple copies inequality is violated when $E > 1/\sqrt{2}$, which is the same bound obtained by [20] for the bipartite scenario. However, for the tripartite scenario, the Navascués-Pironio-Acin (NPA) hierarchy [40] implies that for any $E \geqslant 1/2$ the corresponding correlation will have a post-quantum nature. That is, the tripartite information causality, at least with the specific concatenation considered here, cannot recover the quantum bound.

As previously mentioned, the bipartite version of (3.46) is precisely the quadratic inequality obtained by Uffink [129]. However, for more than two parts, such equivalence no longer holds. For the tripartite scenario, the Uffink inequality reads as

$$(C_{001} + C_{010} + C_{100} - C_{111})^2 + (C_{110} + C_{101} + C_{011} - C_{000})^2 \leqslant 16, \quad (3.47)$$

where $C_{xyz} = \sum_{a,b,c} (-1)^{a+b+c} p(a,b,c|x,y,z)$. Indeed, it is impossible to alternate between the inequalities (3.46) and (3.47) merely by changing labels. Even more importantly, as we will show in the next section, there are post-quantum correlations violating the multiple copies inequality (3.46) that do not violate the tripartite Uffink inequality (3.47) (as well as all the inequalities that are obtained from it by relabelling of parties, measurements, and outcomes). Interestingly, it has been recently introduced in *Ref.* [139] a broadly applicable framework that proves *self-test* in multipartite scenarios. In particular, they prove that the maximum quantum bound for the multipartite version of the Uffink inequality (3.47) is only achieved when parties share a $N \geqslant 3$ qubit GHZ state $|GHZ_N\rangle = (|0\rangle^{\otimes N} + e^{i\phi_N}|1\rangle^{\otimes N})/\sqrt{2}$ and performing maximally anti-commuting projective measurements for each qubit, $A_j = \sigma_x$ and $A_j = \sigma_y$. Therefore, a natural promising avenue for further research may involve investigating the quadratic multiple copies inequality (3.46) within the framework of self-testing.

3.5.3 Numerical Tests

More importantly, to understand the strength of the criteria derived, we considered the following slice of the non-signaling set,

$$p(a,b,c|x,y,z) = \gamma p_{45} + \epsilon p_D + (1-\gamma-\epsilon)p_W, \quad (3.48)$$

where $\gamma, \epsilon \in [0,1]$, $p_{45}(a,b,c|x,y,z)$ is the distribution we defined in section 3.5 in (3.23), $p_D(a,b,c|x,y,z) = \delta_{a,0}\delta_{b,0}\delta_{c,0}$ and $p_W(a,b,c|x,y,z) = 1/8$. Thus, we obtained Fig. 14, which highlights that (3.46) excludes an even broader range of supra-quantum correlation than (3.27). In addition, despite the gap between the quantum set and the presented criteria, we enforced that the derived bound follows from the particular communication protocol depicted in Fig.13. Therefore, it does not preclude the possibility of identifying more suitable protocols that could single out the quantum set for this slice of the non-signaling set or rule out post-quantum extremal correlations.

Fig.14 also presents the edge implied by (3.30) for the same slice in (3.48) for $N = 3$ parties, $n = 2$ initial bits. In this case, we considered that all communication is made through a binary symmetric channel that flips the bit with probability ϵ. In this case, we followed the results from [112] to obtain the curve and considered $\epsilon \to 1/2$. The results clearly indicate that our multiple copies criterion as expressed in (3.46) is in complete agreement with the noisy channel approach, even when applied to the simplest binary symmetric noisy channel. The codes related to the Fig. 14 are available in [140].

In the context of the tripartite scenario involving binary-input/binary-output, there exist 53856 non-signaling extremal correlations that are classified into 46 different equivalence classes, among which 45 are supra-quantum ones [44]. A significant result from [25] states that class 4 of these could never have its post-quantumness detected by principles with a strict bipartite formulation, just as those in (3.2) and (3.7). Thus, we also checked the ability of (3.46) to exclude correlations from class 4, and even more generally, we tested all the 45 supra-quantum extremal distributions of the non-

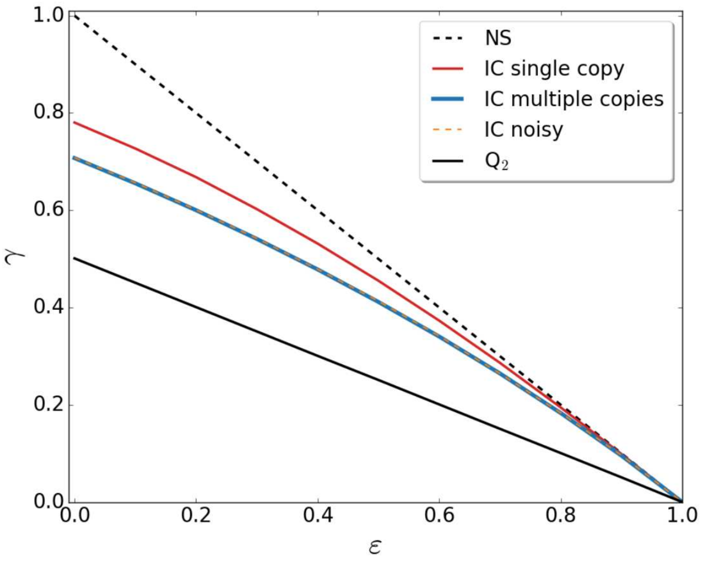


Figure 14 – Non-signaling (NS) polytope slice given by Eq.(3.48). Every dot above these curves represents a correlation that violates the respective criterion represented. The black dashed and solid lines describe the NS and quantum edges, respectively (the last was computed with the level 2 of NPA hierarchy [40]). The single and multiple copies limits defined by the criteria (3.27) and (3.46) are respectively depicted with the red and blue solid lines. Finally, the orange dashed line describes the edge defined by the noisy channel criterion (3.30).

signaling set. In Table 1, we highlight those classes where we could find a violation of (3.46). Even though class 4 does not violate (3.46), we stress that this result is limited to a specific protocol performed among the parties. As it always happens with applications of information causality, it is an open problem whether other protocols would imply a violation of (3.27) for class 4 and other of supra-quantum correlations. Furthermore, Table 1 contains the same analysis for the tripartite Uffink inequality (3.47). From these results, it is clear that there is no equivalence between the multiple copies $\mathcal{IC}$ inequality (3.46) and the Uffink result (3.47) since there exist extremal non-signaling correlations that respect one constraint while violating the other. The codes related to the results from Table 1 are available in [141].

Table 1 – Classes of non-signaling extremal correlations defined in [44] that violate (3.46) or (3.47)

Inequality	Extremal boxes
(3.46)	35, 37, 38, 40, 41, 42, 43, 44, 45
(3.47)	21, 22, 30, 34, 36, 39, 41, 44, 46

3.6 Discussion

This Chapter introduces a novel multipartite communication task in which the previous $\mathcal{IC}$ formulation fails to detect nonlocal advantage. By employing the quantum causal structure formalism, we have proposed a new criterion ensuring $\mathcal{IC}$ in such a new context and proved its truthfulness for the whole set of quantum correlations for any number of parts. Furthermore, we have demonstrated that our model accommodates the concatenation approach from [20], enabling us to derive even stronger constraints for the multipartite non-signaling correlations set. In that case, our multipartite inequality has proven to be stronger than the multipartite Uffink's inequality from [129], which contrasts with the earlier bipartite result from *Ref.* [20]. In addition, our findings align with the recent noisy channel approach from [112], which allows a broad range of analyses for such a multipartite context.

Although the present multipartite criteria do not specifically single out class 4 correlations defined in [44], we emphasize that our results are limited by one specific protocol, which is optimal to Eq.(3.43). However, it only ensures that it is optimal for some non-signaling correlations. Thus, searching for better protocols for different correlations may yield more substantial results, representing one of the main interesting further directions. Furthermore, the analysis of non-dichotomy scenarios, or cases where the sender's initial bits are correlated, may also produce interesting results, as previously analyzed in [96]. Moreover, our findings pave the way for a new class of non-sequential multipartite RACs, where multiple parts exchange messages with the task of computing a boolean function of the senders' initial bits. The figure of merit, in this case, is to calculate the success probability concerning the receiver to accurately compute such a function. Thus, investigating these new thresholds for such a probability of success may have important implications for quantum information processing.

Chapter 4

Monogamy of nonlocality from Information causality

4.1 Monogamy of CHSH inequalities

Consider the simplest non-trivial Bell inequality, the Clauser-Horne-Shimony-Holt (CHSH) inequality, in the dichotomous input-output Bell scenario, expressed as follows:

$$\beta(\mathcal{A},\mathcal{B}) = \frac{1}{4}\sum_{x,y} p(a \oplus b = xy|x,y) \leqslant \frac{3}{4}. \tag{4.1}$$

Here, we denote the parties as $\mathcal{A}$ and $\mathcal{B}$. For the input and output variables we have a, b, x, $y \in \{0,1\}$ and the symbol $\oplus$ denotes sum modulo 2. The value $3/4$ corresponds to the maximum classical bound, which can be violated by quantum and more general no-signaling correlations. While the maximum violation of (4.1) attainable through quantum theory, referred to as Tsirelson's bound [16], is $\beta_Q = \frac{1}{2}(1+\frac{1}{\sqrt{2}}) \approx 0.8535$, no-signaling principle allows correlations achieving $\beta_{\text{NS}} = 1$. Correlations violating Bell inequalities exhibit several nonclassical features, with several practical applications in information processing problems [11, 142, 17, 71, 52]. However, even at theoretical level, nonlocality cannot be considered a free resource. Indeed, when multiple parties are included, monogamy relations emerge.

In essence, *monogamy of nonlocality* asserts that if two parties, $\mathcal{A}$ and $\mathcal{B}$, share nonlocal correlations, the amount of nonlocality that either of them may share with a potential third party, $\mathcal{E}$, is limited. An example of this is the monogamy relation of CHSH inequalities implied by the no-signaling condition (1.5) [74, 143] :

$$\beta(\mathcal{A},\mathcal{B}) + \beta(\mathcal{B},\mathcal{E}) \leqslant \frac{3}{2}. \tag{4.2}$$

Correlations that satisfy (4.2) are not necessarily no-signaling [144], but all no-signaling correlations do respect it. It is important to note, however, that even though (4.2) is weaker than no-signaling condition, when $\beta(\mathcal{A},\mathcal{B}) = \beta_{NS} = 1$, it implies that $\mathcal{B}$ (and $\mathcal{A}$) must be completely uncorrelated with the third party $\mathcal{E}$, such that, $\beta(\mathcal{B},\mathcal{E}) = 1/2$. Interestingly, *Ref.* [74] presents a general form of the $\mathcal{NS}$-monogamy relation for general Bell's inequalities, encompassing arbitrary number of parties, measurement settings, and outcomes.

Of particular importance, within the realm of quantum mechanics, is the following quadratic relation established in [145]:

$$\left(\beta(\mathcal{A},\mathcal{B}) - \frac{1}{2}\right)^2 + \left(\beta(\mathcal{B},\mathcal{E}) - \frac{1}{2}\right)^2 \leqslant \frac{1}{8}. \tag{4.3}$$

This relation is tighter than the no-signaling one, and, similarly, when $\mathcal{A},\mathcal{B}$ observe the maximum quantum value of $\beta(\mathcal{A},\mathcal{B}) = \beta_Q = (1 + 1/\sqrt{2})/2$, then $\beta(\mathcal{B},\mathcal{E})$ must be $1/2$.

In more general terms, monogamy of nonlocality implies that if the two parties $\mathcal{A},\mathcal{B}$ observe nonlocal correlations, such that $\beta(\mathcal{A},\mathcal{B}) > 3/4$, then the strength of their correlations with $\mathcal{E}$, as measured by the value of the CHSH functional $\beta(\mathcal{B},\mathcal{E})$ (or $\beta(\mathcal{A},\mathcal{E})$) remains limited. Thus, we might read monogamy relations in the following form [144]:

$$\beta(\mathcal{B},\mathcal{E}) \leqslant f_T^M\left(\beta(\mathcal{A},\mathcal{B})\right), \tag{4.4}$$

where $f_T^M : [1/2, 1] \rightarrow [0, 1]$ is the function describing the monogamy implied by a specific principle T, such as $\mathcal{IC}$. While (4.3) holds for quantum theory, we are interested in whether a non-trivial [1] monogamy relation of the form (4.4) can be derived via a DI physical principle, without invoking the abstract Hilbert space formalism. As we shall see, monogamy relations of the form (4.4) are cryptographically significant as they can be used to ensure security in DIQKD protocols, against adversaries restricted by the nonlocal theory T [144].

4.2 Optimal CHSH values under the $\mathcal{IC}$ constraint

In Chapter 3 we introduced how information-theoretic constraints bound the set of correlations in Bell scenarios. In this section we discuss how those constraints, such as $\mathcal{IC}$, relate to simultaneous violations of Bell inequalities. The monogamy relations between the values of two Bell inequalities, say $\beta(\mathcal{A},\mathcal{B})$ and $\beta(\mathcal{B},\mathcal{E})$, as implied by $\mathcal{IC}$, can be framed as a maximization problem. Specifically, we seek to determine the maximum value for $\beta(\mathcal{B},\mathcal{E})$, as permitted by a given information-theoretic criterion[2], given a specific value of $\beta(\mathcal{A},\mathcal{B})$.

As previously discussed, all current formulations of $\mathcal{IC}$ are highly dependent on the specific protocol in use. In this regard, to test (3.2) for example, it is necessary to specify the protocol and the resources available to the involved parties. This requirement introduces significant complexity to the analysis. In principle, it suggests that we may need to assess (3.2) across the entire non-signaling polytope [3] to determine whether $\mathcal{IC}$ can imply some monogamy relation such as (4.4), or even recover quantum monogamy, as expressed in (4.3). Nevertheless, as we show in [30], it is enough to analyze a significantly smaller region of $\mathcal{NS}$ set, which we state as the following lemma:

Lemma 1. *To find the maximum CHSH value between $\mathcal{B}$ and $\mathcal{E}$, $\beta(\mathcal{B},\mathcal{E})$, permitted by information causality, when $\mathcal{A}$ and $\mathcal{B}$ witness a CHSH value, $\beta(\mathcal{A},\mathcal{B})$, it suffices to consider tripartite no-signaling correlations $p(a,b,e|x,y,z)$ of the form,*

$$p(a,b,e|x,y,z) = \alpha\frac{1}{4}\delta_{a\oplus b,xy} + \gamma\frac{1}{4}\delta_{e\oplus b,zy} + (1-\alpha-\gamma)1/8, \tag{4.5}$$

where $\alpha, \gamma \in [0,1]$, and $\alpha + \gamma \leqslant 1$.

[1] i.e., tighter than (4.2).

[2] Such as those outlined in Chapter 3

[3] The convex polytope of tripartite no-signaling correlations has 53856 extremal points [44].

Indeed, this parametrization encompasses all aspects associated with (4.2) and (4.3). As a result, the problem is reduced to find the optimal value of γ, as implied by any formulation of $\mathcal{IC}$ that satisfies the data processing relation, for a given value of α. Thus, analyzing (4.5) is sufficient to address the monogamy relations implied by $\mathcal{IC}$.

Proof. Consider a tripartite Bell scenario, where parties $\mathcal{A}, \mathcal{B}, \mathcal{E}$ have binary input-output $x, y, z, a, b, e \in \{0,1\}$. In this context, we observe that the Bell expressions $\beta(\mathcal{A}, \mathcal{B})$ and $\beta(\mathcal{B}, \mathcal{E})$ are individually maximized by their respective PR-boxes $\mathrm{PR}_{\mathcal{AB}} \otimes L_{\mathcal{E}}$ and $\mathrm{PR}_{\mathcal{BE}} \otimes L_{\mathcal{A}}$. Here, $L_{\mathcal{A}}$ and $L_{\mathcal{E}}$ denote some local distributions for $\mathcal{A}$ and $\mathcal{E}$, respectively, such that:

$$p_{\mathrm{PR}_{\mathcal{AB}} \otimes L_{\mathcal{E}}}(a,b,e|x,y,z) = \frac{1}{2}\delta_{a \oplus b, xy} \cdot p_{L_{\mathcal{E}}}(e|z), \tag{4.6}$$

$$p_{\mathrm{PR}_{\mathcal{BE}} \otimes L_{\mathcal{A}}}(a,b,e|x,y,z) = \frac{1}{2}\delta_{b \oplus e, yz} \cdot p_{L_{\mathcal{A}}}(a|x). \tag{4.7}$$

The product structure of these correlations follows from the corresponding $\mathcal{NS}$ -monogamy relation (4.2), which states that tripartite distributions must *factorize* whenever any two of the three parties share a PR-box. Moreover, in this Bell scenario, no other extremal nonlocal tripartite no-signaling box achieves the maximal violation of the CHSH inequalities (see Table 4 in *Ref.* [44]). Hence, any of those contributions, but PR-box, can be effectively ignored. Consequently, without loss generality, we can restrict our analysis to correlations $\mathbb{P}^{(\alpha,\gamma)}_{\mathcal{ABE}}$ given by a convex combination of correlations $\mathrm{PR}_{\mathcal{AB}} \otimes L_{\mathcal{E}}$ and $\mathrm{PR}_{\mathcal{BE}} \otimes L_{\mathcal{A}}$ and a white noise distribution $W_{\mathcal{ABE}}$, where $p_{W_{\mathcal{ABE}}}(a,b,e|x,y,z) = 1/8, \quad \forall a,b,e,x,y,z$, such that,

$$\mathbb{P}^{(\alpha,\gamma)}_{\mathcal{ABE}} = \alpha \mathrm{PR}_{\mathcal{AB}} \otimes L_{\mathcal{E}} + \gamma \mathrm{PR}_{\mathcal{BE}} \otimes L_{\mathcal{A}} + (1-\alpha-\gamma) W_{\mathcal{ABE}} \tag{4.8}$$

where $\alpha, \gamma \geqslant 0$, $\alpha + \gamma \leqslant 1$. Consequently, the optimal CHSH values, $\beta(\mathcal{A}, \mathcal{B})$ and $\beta(\mathcal{B}, \mathcal{E})$, for a given pair (α, γ) in (4.8) are determined by a combination of marginal local functions $L_{\mathcal{A}}$ and $L_{\mathcal{E}}$.

We recall that our objective is to achieve optimal CHSH values, under Shannon's mutual information functional constraints, denoted by $\mathcal{I}(P) \leqslant m$, where P is the joint probability distribution of all observed variables in the DAG (see section 3.4 of Chapter 3), and m is a number specified by marginals of P describing the communication. It is evident that P is protocol-dependent, based on the communication strategy employed by the parties. However, we consistently assume a particular protocol in all cases, which renders the functional depending only on the correlation $\mathbb{P}^{(\alpha,\gamma)}_{\mathcal{ABE}}$, *i.e.*, $\mathcal{I}(\mathbb{P}^{(\alpha,\gamma)}_{\mathcal{ABE}})$[4], and m is then a fixed number. Particular examples of such constraints were discussed in the previous Chapter in terms of information-theoretic constraints in Eq. (3.2),(3.8), (3.30). Indeed, such additional constraint implies a particular form for the marginal distributions $L_{\mathcal{A}}$ and $L_{\mathcal{E}}$ in Eq. (4.8). To illustrate this, consider that the distribution $\bar{\mathbb{P}}^{(\alpha,\gamma)}_{\mathcal{ABE}}$, which is obtained from the original $\mathbb{P}^{(\alpha,\gamma)}_{\mathcal{ABE}}$, by flipping all outputs. In this case

[4] The precise mathematical notation for $\mathcal{I}(\mathbb{P}^{(\alpha,\gamma)}\mathcal{ABE})$ should account for the fixed protocol under consideration; however, for simplicity, we have omitted this index, as the same protocol is assumed for each functional throughout this work.

we have:

$$\bar{\mathbb{P}}^{(\alpha,\gamma)}_{\mathcal{ABE}} = \alpha \mathrm{PR}_{\mathcal{AB}} \otimes \bar{L}_{\mathcal{E}} + \gamma \mathrm{PR}_{\mathcal{BE}} \otimes \bar{L}_{\mathcal{A}} + (1-\alpha-\gamma)\bar{W}_{\mathcal{ABE}}, \tag{4.9}$$

where

$$\begin{aligned} p_{\bar{L}_{\mathcal{E}}}(e|z) &= p_{L_{\mathcal{E}}}(e \oplus 1|z), \\ p_{\bar{L}_{\mathcal{A}}}(a|x) &= p_{L_{\mathcal{A}}}(a \oplus 1|x), \\ p_{\bar{W}_{\mathcal{ABE}}}(a,b,e|x,y,z) &= p_{W_{\mathcal{ABE}}}(a \oplus 1, b \oplus 1, e \oplus 1|x,y,z). \end{aligned} \tag{4.10}$$

Note, however, that the values of the CHSH expressions, $\beta(\mathcal{A},\mathcal{B})$ and $\beta(\mathcal{B},\mathcal{E})$ remain unaltered when all outcomes are *simultaneously* flipped, and they depend solely on the coefficients (α,γ). In parallel, flipping the outputs constitutes a simple local post-processing operation, which, according the data processing inequality (3.28) implies:

$$\mathcal{I}(\bar{\mathbb{P}}^{(\alpha,\gamma)}_{\mathcal{ABE}}) \leqslant \mathcal{I}(\mathbb{P}^{(\alpha,\gamma)}_{\mathcal{ABE}}). \tag{4.11}$$

Accordingly, as the upper bounds on $\mathcal{I}$ depends exclusively on the fixed communication protocol, the interconversion between $\bar{\mathbb{P}}^{(\alpha,\gamma)}_{\mathcal{ABE}}$ and $\mathbb{P}^{(\alpha,\gamma)}_{\mathcal{ABE}}$ does not produce a violation of the information-theoretic constraints. Hence, from the *convexity of Shannon's mutual information*, the convex combination of $\bar{\mathbb{P}}^{(\alpha,\gamma)}_{\mathcal{ABE}}$ and $\mathbb{P}^{(\alpha,\gamma)}_{\mathcal{ABE}}$ also satisfies the bounds on the mutual information functional, $\mathcal{I}$. Specifically:

$$\mathcal{I}\left(\delta\mathbb{P}^{(\alpha,\gamma)}_{\mathcal{ABE}} + (1-\delta)\bar{\mathbb{P}}^{(\alpha,\gamma)}_{\mathcal{ABE}}\right) \leqslant \delta\mathcal{I}\left(\mathbb{P}^{(\alpha,\gamma)}_{\mathcal{ABE}}\right) + (1-\delta)\mathcal{I}\left(\bar{\mathbb{P}}^{(\alpha,\gamma)}_{\mathcal{ABE}}\right), \tag{4.12}$$

where $\delta \in [0,1]$. In particular, for $\delta = 1/2$ we have

$$\mathbb{P}^{*(\alpha,\gamma)}_{\mathcal{ABE}} \equiv \frac{\mathbb{P}^{(\alpha,\gamma)}_{\mathcal{ABE}} + \bar{\mathbb{P}}^{(\alpha,\gamma)}_{\mathcal{ABE}}}{2} = \alpha \mathrm{PR}_{\mathcal{AB}} \otimes \mathrm{W}_{\mathcal{E}} + \gamma \mathrm{PR}_{\mathcal{BE}} \otimes \mathrm{W}_{\mathcal{A}} + (1-\alpha-\gamma)W_{\mathcal{ABE}}, \tag{4.13}$$

where $\mathrm{W}_{\mathcal{A}}$ and $\mathrm{W}_{\mathcal{E}}$ are uniform distribution, and the components of $\mathbb{P}^{*(\alpha,\gamma)}_{\mathcal{ABE}}$ are written explicitly in (4.5). Here, we used the following facts

$$\begin{aligned} \frac{1}{2}(L_{\mathcal{A}} + \bar{L}_{\mathcal{A}}) &= \mathrm{W}_{\mathcal{B}}, \\ \frac{1}{2}(L_{\mathcal{E}} + \bar{L}_{\mathcal{E}}) &= \mathrm{W}_{\mathcal{E}}, \\ W_{\mathcal{ABE}} &= \bar{W}_{\mathcal{ABE}} \end{aligned}$$

Notice that, since the values of the CHSH expressions, $\beta(\mathcal{A},\mathcal{B})$ and $\beta(\mathcal{B},\mathcal{E})$ are the same for $\mathbb{P}^{(\alpha,\gamma)}_{\mathcal{ABE}}$ and $\bar{\mathbb{P}}^{(\alpha,\gamma)}_{\mathcal{ABE}}$, the same applies to $\mathbb{P}^{*(\alpha,\gamma)}_{\mathcal{ABE}}$. Moreover, the form of (4.13) is independent of the marginals $L_{\mathcal{A}}$ and $L_{\mathcal{E}}$. Consequently, since $\mathbb{P}^{(\alpha,\gamma)}_{\mathcal{ABE}}$ in Eq. (4.8) has marginals that achieve the maximum CHSH values in $\beta(\mathcal{A},\mathcal{B})$ and $\beta(\mathcal{B},\mathcal{E})$, as allowed by some information-theoretic constraint, we can always transform this distribution to $\mathbb{P}^{*(\alpha,\gamma)}_{\mathcal{ABE}}$, maintaining the same CHSH values, without producing violations of the information-theoretic constraints of interest. Therefore, $\mathbb{P}^{*(\alpha,\gamma)}_{\mathcal{ABE}}$ is sufficient to determine the maximum CHSH value between $\mathcal{B}$ and $\mathcal{E}$, $\beta(\mathcal{B},\mathcal{E})$, as permitted by information causality in the forms in Eq. (3.2),(3.8), (3.30), when $\mathcal{A}$ and $\mathcal{B}$ witness a CHSH value, $\beta(\mathcal{A},\mathcal{B})$. This concludes the proof.

□

4.3 Bipartite $\mathcal{IC}$ does not imply any monogamy of nonlocality

As it is clear from Chapter 3, the original $\mathcal{IC}$ formulation has an inherent bipartite structure (see DAG in Fig. 8). Such particular feature adds extra problems to apply the approach to Bell's scenarios involving no more than two parties. For the tripartite scenario, for example, the correlation $p(a,b,e|x,y,z)$ must be locally post-processed into an effectively bipartite correlation $p_{eff}(a',b'|x',y')$. Such a processing is typically referred to as wiring, and Fig.15 depicts such a procedure. The parties are assumed to dispose of the tripartite resource and two of them form a composite box (for instance $\mathcal{A}$ and $\mathcal{B}' \equiv (\mathcal{B},\mathcal{E})$), where they can signal to build their effective in/output. A wiring is then specified by the choice of bipartition and by the functions

$$x = F_1(x'), \qquad a' = F_2(a), \tag{4.14}$$

$$y = F_3(y',z,e), \qquad z = F_4(y'), \qquad b' = F_5(b,e),$$

where, $F_i : \{0,1\}^n \to \{0,1\},\ \forall i \in \{1,2,3,4,5\}$. Thus, we say that $p(a,b,e|x,y,z)$ violates $\mathcal{IC}$ if there exists some wiring procedure that produces an effective $p_{eff}(a',b'|x',y')$ that violates some of the bipartite criteria (3.2),(3.8). Indeed, such approaches have been explored in many works studying $\mathcal{IC}$ and multipartite Bell scenarios [25, 24, 146]. Of particular relevance, the authors of [147, 148, 149] claimed having derived the quantum monogamy property in (4.3) from bipartite formulation from IC, by employing the

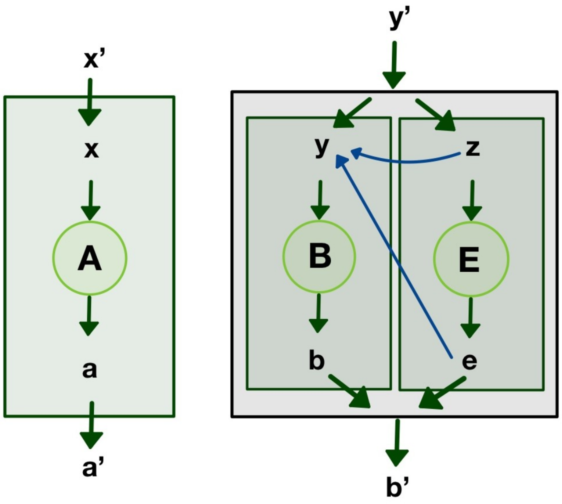


Figure 15 – Wiring procedure that takes tripartite correlations $p(a,b,e|x,y,z)$ and produces a bipartite effective one $p_{eff}(a',b'|x',y')$.

presented procedure. Contrary to these claims, we find that neither the original (3.2), nor the most general $\mathcal{IC}$ (3.8) criteria imply stronger-than-no-signaling monogamy relations, for *any* wiring of the form (4.14).

In the case of (3.8), for each pair of parameters (α, γ), we consider all possible wirings as (4.14), producing effective bipartite distributions. Then, for each $p_{eff}(a', b'|x', y')$, we consider the standard protocol presented Eq. (4.14), and the parties communicating through a binary symmetric channel that flips the message bit with probability $\epsilon \in [0, 1]$. As discussed in [112], however, the scenario with binary symmetric channel enhances all results from the noiseless case, when $\epsilon \to 1/2$. Fig.16 presents the slice (4.5) with the various edges implied by the different constraints of interest. In this case, under computational precision, the bipartite $\mathcal{IC}$ criteria (3.2) and (3.8) retrieve the Tsirelson's bounds of $\beta(\mathcal{B}, \mathcal{E})$, when $\beta(\mathcal{A}, \mathcal{B}) \in [1/2, (1 - 1/\sqrt{2})/2]$. However, for $\beta(\mathcal{A}, \mathcal{B}) \in [(1 - 1/\sqrt{2})/2, \beta_Q]$ the monogamy relation implied by these criteria coincides with the no-signaling monogamy relation (4.2). In particular, even when $\mathcal{A}$ and $\mathcal{B}$ observe the maximum quantum violation, $\beta(\mathcal{A}, \mathcal{B}) = \beta_Q$, the bipartite criteria (3.2),(3.8) implies no bound on $\beta(\mathcal{B}, \mathcal{E})$ stronger than no-signaling one. Thus, we conclude that the original bipartite formulation for $\mathcal{IC}$ fail to yield non-trivial monogamy relations beyond no-signaling.

As we shall see in the next section, however, the multipartite formulation of $\mathcal{IC}$ does imply non-trivial monogamy relations.

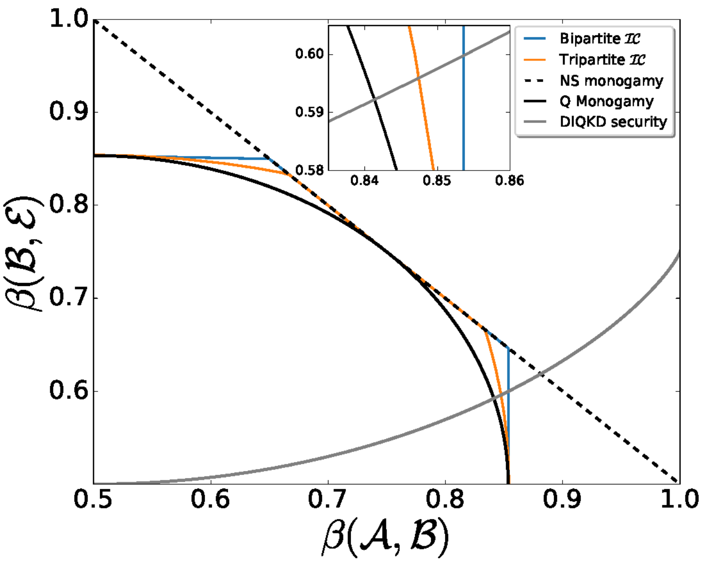


Figure 16 – Relations between the CHSH values for $\mathcal{B}$ and $\mathcal{E}$ as a function of CHSH values for $\mathcal{A}$ and $\mathcal{B}$ implied by the different monogamy relations discussed in this manuscript. The dashed and solid black lines denote non-signaling and quantum monogamy relations from (4.2) and (4.3), respectively. The boundary implied by the bipartite criterion for $\mathcal{IC}$ (3.8) after the wiring process is represented by the blue solid line. The orange line, in turn, denotes the monogamy relation implied by the tripartite criterion for $\mathcal{IC}$ in (4.15). Finally, the solid gray line exhibits the security condition (4.31) for DIQKD through the CHSH protocol. The critical value of security from (4.15) is $\beta(\mathcal{A}, \mathcal{B}) \geqslant 0.8471$.

4.4 Monogamy from tripartite $\mathcal{IC}$

For multipartite $\mathcal{IC}$, we considered the simplest tripartite scenario, previously presented in Fig.11. In this case, $n = 2$, the parties perform the protocol introduced in Section 3.5, communicating independently through a binary symmetric channel. The classical channels are specified by $\epsilon_1, \epsilon_2 \in [0, 1]$, which denote the flipping probability of each message, respectively, *i.e.* $p(M_1' = M_1 \oplus 1 | M_1) = \epsilon_1$ and $p(M_2' = M_2 \oplus 1 | M_1) = \epsilon_2$. For this particular scenario, the criterion in (3.30) translates into

$$\begin{aligned} I(X_1^1 : X_1^2, M_1', M_2', G_1) + I(X_2^1 : X_2^2, M_1', M_2', G_2) \\ + I(X_1^2 : X_1^1, M_1', M_2', G_1) + I(X_2^2 : X_2^1, M_1', M_2', G_2) \\ \leqslant 2 - h(1 - \epsilon_1) - h(1 - \epsilon_2), \end{aligned} \tag{4.15}$$

where $h(p) = -p \log p - (1 - p) \log(1 - p)$. Fig.16 illustrates our main result, where tripartite $\mathcal{IC}$ imposes a non-trivial bound on $\beta(\mathcal{A}, \mathcal{E})$. Despite Eq. (4.15) does not fully recover quantum monogamy as in Eq. (4.3), it does imply a monogamy relation similar to (4.4). Specifically, the criterion (4.15) entails tighter bounds on $\beta(\mathcal{B}, \mathcal{E})$ than no-signaling monogamy relations for the range $\beta(\mathcal{A}, \mathcal{B}) \in [0.8333, \beta_Q]$. Notably, when $\mathcal{A}$ and $\mathcal{B}$ maximally violate the CHSH inequality with $\beta(\mathcal{A}, \mathcal{B}) = \beta_Q$, (4.15) is violated for all distribution $p(a, b, e|x, y, z)$ giving $\beta(\mathcal{B}, \mathcal{E}) > 1/2$, up to computational precision. Consequently, when $\mathcal{A}$ and $\mathcal{B}$ witness the maximum quantum violation of the CHSH inequality ($\beta(\mathcal{A}, \mathcal{B}) = \beta_Q$), information causality imposes that $\mathcal{B}$ must be entirely uncorrelated with any third party $\mathcal{E}$, such that the CHSH value between $\mathcal{B}$ and $\mathcal{E}$ must be $\beta(\mathcal{B}, \mathcal{E}) = 1/2$, thereby recovering the quantum monogamy (4.3).

To build up the boundary in Fig.16, for each pair (α, γ), we examined violations of (4.15) for different values of ϵ_1 and ϵ_2 within the whole range $[0, 1]$. The codes related to Fig.16 are publicly available in [150]. It is noteworthy that, as with all other results associated with $\mathcal{IC}$, the bounds depicted in Fig. 16 are highly dependent on the specific protocol in Fig.12. Consequently, it remains an unsolved question whether the tripartite formulation for $\mathcal{IC}$ can fully capture quantum monogamy, as expressed in (4.3), for alternative protocols or diverse noisy channels. Nevertheless, as we demonstrate in the next section, the bound imposed by (4.15) is enough to guarantee the information theoretic security of DIQKD protocols.

4.5 Security in QKD protocol from $\mathcal{IC}$

In addition to its foundational significance in recovering key properties of quantum mechanics, such as monogamy of Bell inequalities violations, our findings reveal that $\mathcal{IC}$ also holds cryptographic relevance. This is due to the close relation of monogamy relations, such as (4.4), and security of device-independent quantum key distribution (DIQKD) protocols. Indeed, as we shall see, a specific nonlocal theory T that enforces a monogamy relation as (4.4), ensures security against adversaries constrained by the theory T [144].

4.5.1 Security and guessing probability

In that context, we focus on the CHSH-based protocol, where the parties share a large number of states and conduct an experiment, aimed at optimizing the violation of the CHSH inequality (4.1), and the key is established from the outcomes a and $b \in \{0,1\}$ from the Bell experiment. For the agreement procedure, Bob then publicly announces his measurement choice y to Alice, and they agree on the key by flipping $a \to a \oplus 1$ when $x = y = 1$ while retaining the data otherwise. This protocol has been proven secure against no-signaling individual attacks in [52], and even collective attacks [73]. Several other developments have still been explored in this context, such as the recent security proof for arbitrarily small nonlocality [77], and experimental implementation [78, 79]. As previously discussed in Chapter 3, in turn, $\mathcal{IC}$ generalizes the concept of no-signaling to scenarios with communication. Thus, a natural question is to ask whether $\mathcal{IC}$ can also imply security on cryptographic key distribution protocols, or even if it may enhance the security obtained from the no-signaling principle.

A foundational result regarding quantum key distribution (QKD) security was presented in [75], establishing that $\mathcal{A}$ and $\mathcal{B}$ can always distill a secret cryptographic key against an eavesdropper $\mathcal{E}$, when they have an advantage in terms of mutual information, *i.e.*,

$$I(A:B) > I(E:B). \tag{4.16}$$

In other words, whenever the mutual information between $\mathcal{A}$ and $\mathcal{B}$ set of keys exceeds that between $\mathcal{B}$ keys and $\mathcal{E}$'s set of decodings, $\mathcal{A}$ and $\mathcal{B}$ can securely establish cryptographic keys. Thus, in order to infer security from the monogamy result presented in the last section, we need first relate the security criterion (4.16) with the respective Bell inequality violations (see [144, 151]). As we will soon clarify, in the CHSH protocol, Bell inequalities violations are closely related to guessing probabilities of $\mathcal{A}$ and $\mathcal{E}$[5]. In that regard, we may simply re-write (4.16) in terms of Shannon's entropy:

$$H(B) - h(P_A) > H(B) - \sum_{\varepsilon} p(\varepsilon) h\left(p(b=0|\varepsilon)\right). \tag{4.17}$$

Here, $h(\cdot)$ is the binary entropy, and P_A denotes the probability of $\mathcal{A}$ correctly guessing $\mathcal{B}$'s outcome. On the left-hand side we invoke Fano's inequality [152] and the fact that $\mathcal{A}$ and $\mathcal{B}$ have binary outcomes, *i.e.* $I(A:B) \geqslant H(B) - h(P_A)$. Note, therefore, that (4.17) is only a sufficient condition to ensure (4.16). On the right-hand side, we express the conditional entropy $H(B|E)$ explicitly in terms of probability distribution. As previously emphasized in [151], we remember that, in principle, the cardinality of $\mathcal{E}$'s output might be non-binary. Indeed, in the DIQKD scenario, Eve is bounded solely by DI principles, but no boundaries are assumed for her computational power. In the worst case, $\mathcal{E}$ may have even a more precise description of $\mathcal{B}$'s outputs b, provided by her outputs ε, *i.e.*, the distribution $p(b|\varepsilon)$. Consequently, $\mathcal{E}$'s best strategy for guessing $\mathcal{B}$'s output, when receiving the outcome ε, is:

$$g_{\varepsilon} = \arg\left(\max_{b}\; p(b|\varepsilon)\right), \tag{4.18}$$

[5] *i.e.*, the sake of completeness, "guessing probabilities" refer to the success probability of one party producing a correct guess of the other party's output.

where,

$$p(g_\varepsilon) = \max_b p(b|\varepsilon). \tag{4.19}$$

Therefore, $\mathcal{E}$'s guessing probability, P_E, is expressed as the weighted sum:

$$P_E = \sum_\varepsilon p(\varepsilon)p(g_\varepsilon). \tag{4.20}$$

Hence, Eq. (4.17) can be rewritten. Since $h\left(p(b=0|\varepsilon)\right) = h\left(p(b=1|\varepsilon)\right)$, from (4.19), we may write $h\left(p(b=0|\varepsilon)\right) = h\left(p(g_\varepsilon)\right)$. Thus, Eq. (4.17) reduces to:

$$h(P_A) < \sum_\varepsilon p(\varepsilon)h\left(p(g_\varepsilon)\right). \tag{4.21}$$

The maximum value of $h(p(g_\varepsilon))$ is achieved when $p(g_\varepsilon) = 1/2$. Let p denote the sum of all $p(g_\varepsilon)$ such that $p(g_\varepsilon) = 1/2$ by p. We then have:

$$p \leqslant \sum_\varepsilon p(\varepsilon)h(p(g_\varepsilon)). \tag{4.22}$$

Similarly, for (4.20), denoting the sum of all $p(g_\varepsilon)$ taking the value 1 by q, we have $P_E \geqslant p/2 + q$. From normalization, $p + q \leqslant 1$, and if follows that for P_E, we can write:

$$P_E \geqslant 1 - p\,\frac{1}{2}. \tag{4.23}$$

Consequently, (4.21), (4.22), and (4.23) together provide a sufficient condition to satisfy (4.16), yielding a security criterion in terms of the guessing probabilities of $\mathcal{A}$ and $\mathcal{E}$:

$$P_E < 1 - \frac{1}{2}\,h(P_A). \tag{4.24}$$

4.5.2 Security from $\mathcal{IC}$

At this point, we are ready to relate guessing probability and Bell functionals from Eq. (4.1) with guessing probabilities. From the CHSH protocol, $\mathcal{B}$'s output is equal to $b = a \oplus xy$ with probability $\beta(\mathcal{A}, \mathcal{B})$. Since $\mathcal{A}$ has access to $\mathcal{B}$'s inputs y, she can guess $\mathcal{B}$'s output with success probability P_A of:

$$P_A = \beta(\mathcal{A}, \mathcal{B}). \tag{4.25}$$

In the case of $\mathcal{E}$, to maximize P_E, she intercepts $\mathcal{B}$'s input y and explores it to produce a correct guess about $\mathcal{B}$'s output. More generally, $\mathcal{E}$ may also have her own input z, producing the guess $g_{y',z}$, where y' represents possible incorrect eavesdropping by $\mathcal{E}$. In the ideal case, $y' = y$. Her guessing strategy is then specified by $p(g_{y',z}) =$

$\max_b p(b|y',z)^6$, where

$$P_E = \sum_{y',z} = p(g_{y',z})p(y',z) \leqslant \max_{y',z}(p(g_{y',z})). \tag{4.26}$$

Thus, we now ask what CHSH values, $\beta(\mathcal{E},\mathcal{B})$, $\mathcal{E}$ can achieve, through such a mechanism. In this case, we denote by e and z, the respective $\mathcal{E}$'s input and output in the Bell experiment. Since, in a Bell experiment, $\mathcal{E}$does not have access to $\mathcal{B}$'s inputs, her best strategy consists of simply producing a guess as the most likely $g_{y',z}$. That is,

$$G = \arg\left(\max_{y',z}\; p(g_{y',z})\right). \tag{4.27}$$

In this scenario, $\mathcal{E}$ aims to maximize $p(e \oplus b = yz)$. Therefore, when G takes the values $g_{0,0}$, $g_{0,1}$, and $g_{1,0}$, her best strategy is to output $e = G$. Conversely, when $G = g_{1,1}$, she outputs $e = G \oplus 1$. In this case, the correlations in the CHSH functional, $\beta(\mathcal{E},\mathcal{B})$, may relate to $p(g_{y,z})$. For instance, when $G = g_{0,0}$ we have:

$$p(e \oplus b = 0|y = 0, z = 0) = p(g_{y'=0,z=0}), \tag{4.28a}$$
$$p(e \oplus b = 0|y = 0, z = 1) = p(g_{y'=0,z=0}), \tag{4.28b}$$
$$p(e \oplus b = 0|y = 1, z = 0) = p(g_{y'=0,z=0}), \tag{4.28c}$$
$$p(e \oplus b = 1|y = 1, z = 1) = 1 - p(g_{y'=0,z=0}), \tag{4.28d}$$

which together with (4.26), leads to,

$$\beta(\mathcal{E},\mathcal{B}) = \frac{1}{4}(1 + 2p(g_{y'=0,z=0})) \geqslant \frac{1}{4}(1 + 2P_E). \tag{4.29}$$

It is straightforward to verify that (4.29) is general, as it holds for all G in (4.27).

With this treatment, we can finally connect the monogamy of Bell inequality violations with secrecy in cryptographic protocols. Indeed from the general monogamy in Eq. (4.4), we can rewrite (4.29) as:

$$P_E \leqslant 2f_T^M\left(\beta(\mathcal{A},\mathcal{B})\right) - \frac{1}{2}. \tag{4.30}$$

Thus, by combining (4.24), (4.25), and (4.30), we derive a sufficient condition for the Csiszár and Körner security criterion in (4.16):

$$h\left(\beta(\mathcal{A},\mathcal{B})\right) < 3 - 4f_T^M\left(\beta(\mathcal{A},\mathcal{B})\right). \tag{4.31}$$

This condition imposes threshold values of $\beta(\mathcal{A},\mathcal{B})$ sufficient for security, which can be determined by examining the specific form of f_T^M in (4.31), for a given nonlocal theory T. For instance, correlations that satisfy the no-signaling condition obey the *linear* monogamy relation (4.2). In this case, the threshold value of $\beta(\mathcal{A},\mathcal{B})$ for secure DIQKD, with no-signaling monogamy (4.2), via (4.31), turns out to be ≈ 0.881, which is not realizable with quantum theory [153].

[6] More precisely, $p(g_{y',z}) = \sum_{\varepsilon} p(\varepsilon)p(g_{\varepsilon,y',z}) = \sum_{\varepsilon} p(\varepsilon)\max_b\left(p(b|\varepsilon,y',z)\right)$.

Of particular relevance, however, for the quantum quadratic monogamy relation (4.3) applied to (4.31), the threshold value of $\beta(\mathcal{A},\mathcal{B})$, for secure DIQKD is ≈ 0.841 [153]. In contrast to the quantum bound, $\beta(\mathcal{A},\mathcal{B}) = \beta_Q = \frac{1}{2}(1+\frac{1}{\sqrt{2}})$, $\mathcal{A}$ and $\mathcal{B}$ can establish a secret key, based only on the quantum monogamy (4.3). It is important to note, however, that the parties being constrained by (4.3) do not necessarily imply they are restricted to quantum mechanics. Therefore, by assuming (4.3), and by means of a quantum mechanics experiment, $\mathcal{A}$ and $\mathcal{B}$ ensure security even against potentially supra-quantum attacks.

While (4.3) holds within quantum theory, our interest lies in whether the monogamy relation derived from $\mathcal{IC}$ principle may ensure security through the criterion (4.31). As presented in Section 4.4, $\mathcal{IC}$ implies bounds of the form (4.4), which are tighter than (4.2). In Fig. 16, the solid gray line represents the security constraint (4.31). In fact, tripartite formulation of $\mathcal{IC}$ (Eq. (4.15)) ensures (4.31) for $\beta(\mathcal{A},\mathcal{B}) \geqslant 0.8471$. Consequently, we conclude that $\mathcal{IC}$ ensures security for DIQKD protocols, whenever $\mathcal{A}$ and $\mathcal{B}$ witness the CHSH value in the realizable range of quantum correlations, $\beta(\mathcal{A},\mathcal{B}) \in [0.8471, \beta_Q]$.

As extensively discussed in prior sections of the present thesis, it remains an open question whether $\mathcal{IC}$ may completely single out the set of quantum correlations. This fact highlights the significance of the cryptographic security based on the tripartite $\mathcal{IC}$. In that case, we have that through an experiment bounded by quantum mechanics, Alice and Bob ensure security against a potentially supra-quantum eavesdropping. Consequently, regardless of whether $\mathcal{IC}$ fully characterizes the set of quantum correlations, the theoretical security proof based on the $\mathcal{IC}$ principle holds, even in the face of a hypothetical breakthrough beyond quantum mechanics.

Interestingly, concerning the bipartite formulation of $\mathcal{IC}$, it is straight to conclude that there are no means to guarantee security from (4.31). Under computational precision, there exists no bound on $\beta(\mathcal{B},\mathcal{E})$ imposed by bipartite $\mathcal{IC}$for $\beta(\mathcal{A},\mathcal{B}) \leqslant \beta_Q$. Consequently, there is no CHSH value in the helm of quantum mechanics, $\beta(\mathcal{A},\mathcal{B}) \leqslant \beta_Q$, for which bipartite $\mathcal{IC}$ ensures (4.31).

4.6 Discussion

We establish a clear connection between the concept of information causality ($\mathcal{IC}$) and the monogamy of Bell inequality violations. In that context, we first describe Eq. (4.5) as the region on the NS polytope wherein all optimal values of Bell inequalities, subject to the $\mathcal{IC}$ constraint, are found. This result is particularly significant, as it applies not only to $\mathcal{IC}$-related expressions but also to any informational function that respects the data processing inequality. Thus, we demonstrate that in the standard $\mathcal{IC}$ bipartite scenario (for the standard protocol) no monogamy relation between CHSH expressions can be derived. This consequence stands in contrast to our main finding, where we introduce a novel multipartite $\mathcal{IC}$ framework that reveals the existence of a monogamy relation. As a noteworthy consequence of these results, we prove that $\mathcal{IC}$ guarantees security against non-signaling attacks for DIQKD protocols.

The presented framework opens up several promising avenues for further research. The result that identifies the optimal slice for analyzing Bell functional values under the $\mathcal{IC}$ constraint has so far only been applied to the CHSH expression (Eq.(4.1)). It remains unclear, however, whether one can extend this result to encompass other Bell inequalities. It is also worth observing the protocol-dependent structure inherent in every result within the context of $\mathcal{IC}$. In this regard, a critical open question is whether the gap between $\mathcal{IC}$ and quantum monogamy, as illustrated in Fig.16, could be narrowed by assigning different encoding and decoding protocols to the parties. If so, the security proof for DIQKD could be significantly strengthened, leading to improved threshold values for security. Furthermore, our analysis has focused exclusively on the noisy binary symmetric channel, which leaves open the possibility of considering different more suitable channels, depending on the specific context. Naturally, expanding the investigation to address alternative types of eavesdropping, such as collective or coherent attacks, represents a natural and important extension of this work. Finally, as monogamy relations are also closely tied to quantum random number generation certification (QRNG), another compelling direction for future research would be to investigate the potential connection between $\mathcal{IC}$ and QRNG in detail.

Chapter 5

Final remarks

In this thesis, we have investigated fundamental questions at the intersection of quantum mechanics, nonlocality, and communication. By investigating Bell nonlocality and its significance in information processing, this work advances the understanding of quantum correlations. In particular, it addresses significant open problems in literature by introducing novel frameworks and formulating new operational criteria for the *Information Causality* principle. These contributions shed light on the interplay between nonlocality and communication, offering insights into quantum correlations and their limitations. Within this context, a central focus of this thesis was the development of a multipartite framework for $\mathcal{IC}$, which approaches the limitations of its original bipartite formulation. In this regard, we employ the systematic Shannon's entropic cone framework [22]. These findings are particularly significant, as they enable the derivation of more substantial constraints on the non-signaling correlations set, surpassing the previously established multipartite inequalities [129]. This result aligns with recent developments on noisy channel approaches [112], suggesting broader applicability of our findings to multipartite contexts. Additionally, we investigated the connection between $\mathcal{IC}$ and the monogamy of Bell inequality violations. While the bipartite framework of $\mathcal{IC}$ fails to recover monogamy relations, we demonstrated that the introduced multipartite framework naturally ensures such monogamy properties, recovering the strong form of monogamy relation implied within quantum theory. This result has implications in device-independent quantum key distribution, as it establishes $\mathcal{IC}$ as a fundamental principle guaranteeing security against non-signaling attacks. The ability to link $\mathcal{IC}$ to monogamy relations also opens promising avenues for exploring its implications in related areas, such as quantum random number generation certification (QRNG) and other cryptographic applications.

As it happens with all known formulations for $\mathcal{IC}$, the protocol-dependent structure of the principle leaves open the possibility of more suitable protocols that better highlight $\mathcal{IC}$ violation. For instance, while the proposed multipartite criteria provide robust constraints, they are optimal only for specific protocols and certain classes of non-signaling correlations. Identifying alternative communication tasks for which other classes of non-signaling correlation are optimal could also significantly improve the presented results. In the context of monogamous relations and $\mathcal{IC}$, the current work primarily addresses the CHSH expression and the noisy binary symmetric channel, leaving open the possibility of exploring alternative Bell inequalities and channels. Extending the security results to more complex forms of eavesdropping, such as collective or coherent attacks, is naturally an interesting further direction. Addressing these extensions could refine the security proofs for DIQKD protocols and improve threshold values for practical implementations.

Lastly, the broader question of whether $\mathcal{IC}$ or related device-independent principles can fully characterize the set of quantum correlations remains unresolved. Nevertheless, the introduced monogamous relation via $\mathcal{IC}$ highlights the practical significance of the principle. *i.e.*, whether $\mathcal{IC}$ fits as a holding principle of nature, parties may always ensure security by means of an experiment with quantum systems without introducing the Hilbert spaces formalism of quantum theory. Consequently, even if in the future we negatively answer whether $\mathcal{IC}$ fully singles out the set of quantum correlations, the theoretical proofs still hold, which would ensure security even in the hypothetical scenario where the eavesdropper could breakthrough the laws of quantum mechanics.

Bibliography

[1] BELL, J. S. On the einstein podolsky rosen paradox. *Physics Physique Fizika*, American Physical Society, v. 1, p. 195–200, Nov 1964. Disponível em: <https://link.aps.org/doi/10.1103/PhysicsPhysiqueFizika.1.195>.

[2] CLAUSER, J. F. et al. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, American Physical Society, v. 23, p. 880–884, Oct 1969. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.23.880>.

[3] FREEDMAN, S. J.; CLAUSER, J. F. Experimental test of local hidden-variable theories. *Phys. Rev. Lett.*, American Physical Society, v. 28, p. 938–941, Apr 1972. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.28.938>.

[4] ASPECT, A.; DALIBARD, J.; ROGER, G. Experimental test of bell's inequalities using time-varying analyzers. *Phys. Rev. Lett.*, American Physical Society, v. 49, p. 1804–1807, Dec 1982. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.49.1804>.

[5] ASPECT, A.; GRANGIER, P.; ROGER, G. Experimental realization of einstein-podolsky-rosen-bohm gedankenexperiment: A new violation of bell's inequalities. *Phys. Rev. Lett.*, American Physical Society, v. 49, p. 91–94, Jul 1982. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.49.91>.

[6] HENSEN, B. et al. Experimental loophole-free violation of a bell inequality using entangled electron spins separated by 1.3 km. *Nature*, v. 526, p. 682–686, 08 2015.

[7] SHALM, L. K. et al. Strong loophole-free test of local realism. *Phys. Rev. Lett.*, American Physical Society, v. 115, p. 250402, Dec 2015. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.115.250402>.

[8] GIUSTINA, M. et al. Significant-loophole-free test of bell's theorem with entangled photons. *Phys. Rev. Lett.*, American Physical Society, v. 115, p. 250401, Dec 2015. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.115.250401>.

[9] EINSTEIN, A.; PODOLSKY, B.; ROSEN, N. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, American Physical Society, v. 47, p. 777–780, May 1935. Disponível em: <https://link.aps.org/doi/10.1103/PhysRev.47.777>.

[10] BOHR, N. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, American Physical Society, v. 48, p. 696–702, Oct 1935. Disponível em: <https://link.aps.org/doi/10.1103/PhysRev.48.696>.

[11] EKERT, A. K. Quantum cryptography based on bell's theorem. *Phys. Rev. Lett.*, American Physical Society, v. 67, p. 661–663, Aug 1991. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.67.661>.

[12] BRUKNER, C. et al. Bell's inequalities and quantum communication complexity. *Phys. Rev. Lett.*, American Physical Society, v. 92, p. 127901, Mar 2004. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.92.127901>.

[13] NEUMANN, J. von. *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, 1955. (Goldstine Printed Materials). ISBN 9780691028934. Disponível em: <https://books.google.com.br/books?id=JLyCo3RO4qUC>.

[14] D'ARIANO, G. M.; CHIRIBELLA, G.; PERINOTTI, P. *Quantum Theory from First Principles: An Informational Approach*. [S.l.]: Cambridge University Press, 2017.

[15] POPESCU, S.; ROHRLICH, D. Quantum nonlocality as an axiom. *Foundations of Physics*, v. 24, n. 3, p. 379–385, 1994. Disponível em: <https://doi.org/10.1007/BF02058098>.

[16] CIREL'SON, B. S. Quantum generalizations of bell's inequality. *Letters in Mathematical Physics*, v. 4, p. 93–100, 1980. Disponível em: <https://doi.org/10.1007/BF00417500>.

[17] DAM, W. van. *Implausible Consequences of Superstrong Nonlocality*. 2005.

[18] BRASSARD, G. et al. Limit on nonlocality in any world in which communication complexity is not trivial. *Phys. Rev. Lett.*, American Physical Society, v. 96, p. 250401, Jun 2006. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.96.250401>.

[19] NAVASCUéS, M.; WUNDERLICH, H. A glance beyond the quantum model. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, The Royal Society, v. 466, n. 2115, p. 881–890, Nov 2009. ISSN 1471-2946. Disponível em: <http://dx.doi.org/10.1098/rspa.2009.0453>.

[20] PAWŁOWSKI, M. et al. Information causality as a physical principle. *Nature*, Springer Science and Business Media LLC, v. 461, n. 7267, p. 1101–1104, Oct 2009. ISSN 1476-4687. Disponível em: <http://dx.doi.org/10.1038/nature08400>.

[21] ALLCOCK, J. et al. Recovering part of the boundary between quantum and nonquantum correlations from information causality. *Phys. Rev. A*, American Physical Society, v. 80, p. 040103, Oct 2009. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevA.80.040103>.

[22] CHAVES, R.; MAJENZ, C.; GROSS, D. Information–theoretic implications of quantum causal structures. *Nature Communications*, Springer Science and Business Media LLC, v. 6, n. 1, Jan 2015. ISSN 2041-1723. Disponível em: <http://dx.doi.org/10.1038/ncomms6766>.

[23] YU, B.; SCARANI, V. *Information causality beyond the random access code model*. 2022. Disponível em: <https://arxiv.org/abs/2201.08986>.

[24] GALLEGO, R. et al. Quantum correlations require multipartite information principles. *Phys. Rev. Lett.*, American Physical Society, v. 107, p. 210403, Nov 2011. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.107.210403>.

[25] YANG, T. H. et al. Information-causality and extremal tripartite correlations. *New Journal of Physics*, IOP Publishing, v. 14, n. 1, p. 013061, jan 2012. Disponível em: <https://doi.org/10.1088/1367-2630/14/1/013061>.

[26] AMBAINIS, A. et al. Dense quantum coding and a lower bound for 1-way quantum automata. *Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing*, Association for Computing Machinery, New York, NY, USA, p. 376–383, 1999. Disponível em: <https://doi.org/10.1145/301250.301347>.

[27] AMBAINIS, A. et al. Dense quantum coding and quantum finite automata. *J. ACM*, Association for Computing Machinery, New York, NY, USA, v. 49, n. 4, p. 496–511, jul 2002. ISSN 0004-5411. Disponível em: <https://doi.org/10.1145/581771.581773>.

[28] WOLF, S.; WULLSCHLEGER, J. Oblivious transfer and quantum non-locality. *Proceedings. International Symposium on Information Theory, 2005. ISIT 2005.*, p. 1745–1748, 2005. Disponível em: <https://api.semanticscholar.org/CorpusID:8960223>.

[29] POLLYCENO, L.; CHAVES, R.; RABELO, R. Information causality in multipartite scenarios. *Phys. Rev. A*, American Physical Society, v. 107, p. 042203, Apr 2023. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevA.107.042203>.

[30] POLLYCENO, L. et al. Monogamy of nonlocality from multipartite information causality. *arXiv:2405.20115*, 2024. Disponível em: <https://arxiv.org/abs/2405.20115>.

[31] BANCAL, J.-D. *On the device-independent approach to quantum physics : advances in quantum nonlocality and multipartite entanglement detection*. Tese (Doutorado) — Université de Genève, 05/04 2012. Disponível em: <https://nbn-resolving.org/urn:nbn:ch:unige-217102>.

[32] SCARANI, V. *The device-independent outlook on quantum physics (lecture notes on the power of Bell's theorem)*. 2015.

[33] BRUNNER, N. et al. Bell nonlocality. *Reviews of Modern Physics*, American Physical Society (APS), v. 86, n. 2, p. 419–478, Apr 2014. ISSN 1539-0756. Disponível em: <http://dx.doi.org/10.1103/RevModPhys.86.419>.

[34] BOYD, S.; VANDENBERGHE, L. *Convex Optimization*. Cambridge,: Cambridge University Press, 2004.

[35] LÖRWALD, S.; REINELT, G. Panda: a software for polyhedral transformations. *EURO Journal on Computational Optimization*, Springer Berlin Heidelberg, p. 1–12, 2015. ISSN 2192-4406. Disponível em: <http://dx.doi.org/10.1007/s13675-015-0040-0>.

[36] BARRETT, J. et al. Nonlocal correlations as an information-theoretic resource. *Phys. Rev. A*, American Physical Society, v. 71, p. 022101, Feb 2005. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevA.71.022101>.

[37] WERNER, R. F.; WOLF, M. M. Bell inequalities and entanglement. *arXiv:quant-ph/0107093*, 2001. Disponível em: <https://arxiv.org/abs/quant-ph/0107093>.

[38] PITOWSKY, I. *Quantum Probability — Quantum Logic*. 1. ed. [S.l.]: Springer-Verlag Berlin Heidelberg, 1989. v. 321.

[39] QUINTINO, M. T. C. *Black box correlations: locality, noncontextuality, and convex polytopes*. Dissertação (Mestrado) — UFMG, 2012. Disponível em: <http://hdl.handle.net/1843/BUOS-A46HJC>.

[40] NAVASCUéS, M.; PIRONIO, S.; ACíN, A. A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. *New Journal of Physics*, IOP Publishing, v. 10, n. 7, p. 073013, Jul 2008. ISSN 1367-2630. Disponível em: <http://dx.doi.org/10.1088/1367-2630/10/7/073013>.

[41] GISIN, N. Bell's inequality holds for all non-product states. *Physics Letters A*, v. 154, n. 5, p. 201–202, 1991. ISSN 0375-9601. Disponível em: <https://www.sciencedirect.com/science/article/pii/037596019190805I>.

[42] POPESCU, S.; ROHRLICH, D. Generic quantum nonlocality. *Physics Letters A*, v. 166, n. 5, p. 293–297, 1992. ISSN 0375-9601. Disponível em: <https://www.sciencedirect.com/science/article/pii/037596019290711T>.

[43] SCARANI, V. *Bell Nonlocality*. Oxford University Press, 2019. (Oxford Graduate Texts). ISBN 9780198788416. Disponível em: <https://books.google.com.br/books?id=39ShDwAAQBAJ>.

[44] PIRONIO, S.; BANCAL, J.-D.; SCARANI, V. Extremal correlations of the tripartite no-signaling polytope. *Journal of Physics A: Mathematical and Theoretical*, IOP Publishing, v. 44, n. 6, p. 065303, 2011. ISSN 1751-8121. Disponível em: <http://dx.doi.org/10.1088/1751-8113/44/6/065303>.

[45] PITOWSKY, I.; SVOZIL, K. Optimal tests of quantum nonlocality. *Phys. Rev. A*, American Physical Society, v. 64, p. 014102, Jun 2001. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevA.64.014102>.

[46] ŚLIWA, C. Symmetries of the bell correlation inequalities. *Physics Letters A*, v. 317, n. 3, p. 165–168, 2003. ISSN 0375-9601. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0375960103011150>.

[47] ALMEIDA, M. L. et al. Guess your neighbor's input: A multipartite nonlocal game with no quantum advantage. *Physical Review Letters*, American Physical Society (APS), v. 104, n. 23, 2010. ISSN 1079-7114. Disponível em: <http://dx.doi.org/10.1103/PhysRevLett.104.230404>.

[48] SVETLICHNY, G. Distinguishing three-body from two-body nonseparability by a bell-type inequality. *Phys. Rev. D*, American Physical Society, v. 35, p. 3066–3069, May 1987. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevD.35.3066>.

[49] GREENBERGER, D. M.; HORNE, M. A.; ZEILINGER, A. Going beyond bell's theorem. In: ____. *Bell's Theorem, Quantum Theory and Conceptions of the Universe*. Dordrecht: Springer Netherlands, 1989. p. 69–72.

[50] BARRETT, J.; PIRONIO, S. Popescu-rohrlich correlations as a unit of nonlocality. *Phys. Rev. Lett.*, American Physical Society, v. 95, p. 140401, Sep 2005. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.95.140401>.

[51] MAYERS, D.; YAO, A. Quantum cryptography with imperfect apparatus. In: *Proceedings 39th Annual Symposium on Foundations of Computer Science (Cat. No.98CB36280)*. [S.l.: s.n.], 1998. p. 503–509.

[52] ACÍN, A.; GISIN, N.; MASANES, L. From bell's theorem to secure quantum key distribution. *Phys. Rev. Lett.*, American Physical Society, v. 97, p. 120405, Sep 2006. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.97.120405>.

[53] BENNETT, C.; BRASSARD, G. Quantum cryptography: Public key distribution and coin tossing. *Theoretical Computer Science - TCS*, v. 560, p. 175–179, 01 1984.

[54] BENNETT, C. H.; BRASSARD, G.; MERMIN, N. D. Quantum cryptography without bell's theorem. *Phys. Rev. Lett.*, American Physical Society, v. 68, p. 557–559, Feb 1992. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.68.557>.

[55] RIVEST, R. L.; SHAMIR, A.; ADLEMAN, L. M. A method for obtaining digital signatures and public-key cryptosystems. *Commun. ACM*, v. 26, p. 96–99, 1978. Disponível em: <https://api.semanticscholar.org/CorpusID:30798417>.

[56] SHOR, P. W. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, v. 26, n. 5, p. 1484–1509, 1997.

[57] VERNAM, G. S. Cipher printing telegraph systems: For secret wire and radio telegraphic communications. *Journal of the A.I.E.E.*, v. 45, n. 2, p. 109–115, 1926.

[58] SHANNON, C. E. Communication theory of secrecy systems. *The Bell System Technical Journal*, v. 28, n. 4, p. 656–715, 1949.

[59] WOOTTERS, W. K.; ZUREK, W. H. A single quantum cannot be cloned. *Nature*, v. 299, p. 802–803, 1982. Disponível em: <https://www.nature.com/articles/299802a0>.

[60] BRUSS, D. Optimal eavesdropping in quantum cryptography with six states. *Phys. Rev. Lett.*, American Physical Society, v. 81, p. 3018–3021, Oct 1998. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.81.3018>.

[61] BENNETT, C. H. et al. Experimental quantum cryptography. *Journal of Cryptology*, v. 5, p. 1432–1378, 1992.

[62] BUTTLER, W. T. et al. Daylight quantum key distribution over 1.6 km. *Phys. Rev. Lett.*, American Physical Society, v. 84, p. 5652–5655, Jun 2000. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.84.5652>.

[63] GISIN, N. et al. Quantum cryptography. *Rev. Mod. Phys.*, American Physical Society, v. 74, p. 145–195, Mar 2002. Disponível em: <https://link.aps.org/doi/10.1103/RevModPhys.74.145>.

[64] HUGHES, R. J. et al. Practical free-space quantum key distribution over 10 km in daylight and at night. *New Journal of Physics*, v. 4, n. 1, p. 43, jul 2002. Disponível em: <https://dx.doi.org/10.1088/1367-2630/4/1/343>.

[65] GOBBY, C.; YUAN, Z. L.; SHIELDS, A. J. Quantum key distribution over 122 km of standard telecom fiber. *Applied Physics Letters*, AIP Publishing, v. 84, n. 19, p. 3762–3764, maio 2004. ISSN 1077-3118. Disponível em: <http://dx.doi.org/10.1063/1.1738173>.

[66] BRUNNER, N. et al. Testing the dimension of hilbert spaces. *Phys. Rev. Lett.*, American Physical Society, v. 100, p. 210503, May 2008. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.100.210503>.

[67] WEHNER, S.; CHRISTANDL, M.; DOHERTY, A. C. Lower bound on the dimension of a quantum system given measured data. *Phys. Rev. A*, American Physical Society, v. 78, p. 062112, Dec 2008. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevA.78.062112>.

[68] GALLEGO, R. et al. Device-independent tests of classical and quantum dimensions. *Phys. Rev. Lett.*, American Physical Society, v. 105, p. 230501, Nov 2010. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.105.230501>.

[69] CURTY, M.; LEWENSTEIN, M.; LÜTKENHAUS, N. Entanglement as a precondition for secure quantum key distribution. *Phys. Rev. Lett.*, American Physical Society, v. 92, p. 217903, May 2004. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.92.217903>.

[70] SCARANI, V. et al. Secrecy extraction from no-signaling correlations. *Phys. Rev. A*, American Physical Society, v. 74, p. 042339, Oct 2006. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevA.74.042339>.

[71] BARRETT, J.; HARDY, L.; KENT, A. No signaling and quantum key distribution. *Phys. Rev. Lett.*, American Physical Society, v. 95, p. 010503, Jun 2005. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.95.010503>.

[72] ACÍN, A. et al. Device-independent security of quantum cryptography against collective attacks. *Phys. Rev. Lett.*, American Physical Society, v. 98, p. 230501, Jun 2007. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.98.230501>.

[73] PIRONIO, S. et al. Device-independent quantum key distribution secure against collective attacks. *New Journal of Physics*, v. 11, n. 4, p. 045021, apr 2009. Disponível em: <https://dx.doi.org/10.1088/1367-2630/11/4/045021>.

[74] PAWŁOWSKI, M.; BRUKNER, i. c. v. Monogamy of bell's inequality violations in nonsignaling theories. *Phys. Rev. Lett.*, American Physical Society, v. 102, p. 030403, Jan 2009. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.102.030403>.

[75] CSISZAR, I.; KORNER, J. Broadcast channels with confidential messages. *IEEE Transactions on Information Theory*, v. 24, n. 3, p. 339–348, 1978.

[76] MAURER, U. Secret key agreement by public discussion from common information. *IEEE Transactions on Information Theory*, v. 39, n. 3, p. 733–742, 1993.

[77] WOOLTORTON, L.; BROWN, P.; COLBECK, R. Device-independent quantum key distribution with arbitrarily small nonlocality. *Phys. Rev. Lett.*, American Physical Society, v. 132, p. 210802, May 2024. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.132.210802>.

[78] ZHANG, W. et al. A device-independent quantum key distribution system for distant users. *Nature*, Springer Science and Business Media LLC, v. 607, n. 7920, p. 687–691, jul. 2022. ISSN 1476-4687. Disponível em: <http://dx.doi.org/10.1038/s41586-022-04891-y>.

[79] NADLINGER, D. P. et al. Experimental quantum key distribution certified by bell's theorem. *Nature*, Springer Science and Business Media LLC, v. 607, n. 7920, p. 682–686, jul. 2022. ISSN 1476-4687. Disponível em: <http://dx.doi.org/10.1038/s41586-022-04941-5>.

[80] HOLEVO, A. S. Bounds for the quantity of information transmitted by a quantum communication channel. *Problemy Peredachi Informatsii*, American Physical Society, v. 9, p. 3–11, Nov 1973. Disponível em: <https://api.semanticscholar.org/CorpusID:118312737>.

[81] BENNETT, C. H.; WIESNER, S. J. Communication via one- and two-particle operators on einstein-podolsky-rosen states. *Phys. Rev. Lett.*, American Physical Society, v. 69, p. 2881–2884, Nov 1992. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.69.2881>.

[82] WIESNER, S. Conjugate coding. *SIGACT News*, Association for Computing Machinery, New York, NY, USA, v. 15, n. 1, p. 78–88, 1983. ISSN 0163-5700. Disponível em: <https://doi.org/10.1145/1008908.1008920>.

[83] NAYAK, A. Optimal lower bounds for quantum automata and random access codes. In: *40th Annual Symposium on Foundations of Computer Science (Cat. No.99CB37039)*. [S.l.: s.n.], 1999. p. 369–376.

[84] GAVINSKY, D. On the role of shared entanglement. *arXiv:quant-ph/0604052*, 2006. Disponível em: <https://arxiv.org/abs/quant-ph/0604052>.

[85] PAWLOWSKI, M.; ZUKOWSKI, M. Entanglement-assisted random access codes. *Phys. Rev. A*, American Physical Society, v. 81, p. 042326, Apr 2010. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevA.81.042326>.

[86] TAVAKOLI, A. et al. Correlations in entanglement-assisted prepare-and-measure scenarios. *PRX Quantum*, American Physical Society, v. 2, p. 040357, Dec 2021. Disponível em: <https://link.aps.org/doi/10.1103/PRXQuantum.2.040357>.

[87] PAUWELS, J. et al. Entanglement in prepare-and-measure scenarios: many questions, a few answers. *New Journal of Physics*, IOP Publishing, v. 24, n. 6, p. 063015, jun 2022. Disponível em: <https://dx.doi.org/10.1088/1367-2630/ac724a>.

[88] VIEIRA, C. et al. Interplays between classical and quantum entanglement-assisted communication scenarios. *New Journal of Physics*, IOP Publishing, v. 25, n. 11, p. 113004, nov. 2023. ISSN 1367-2630. Disponível em: <http://dx.doi.org/10.1088/1367-2630/ad0526>.

[89] BOWLES, J.; QUINTINO, M. T.; BRUNNER, N. Certifying the dimension of classical and quantum systems in a prepare-and-measure scenario with independent devices. *Phys. Rev. Lett.*, American Physical Society, v. 112, p. 140407, Apr 2014. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.112.140407>.

[90] TAVAKOLI, A. et al. Self-testing quantum states and measurements in the prepare-and-measure scenario. *Phys. Rev. A*, American Physical Society, v. 98, p. 062307, Dec 2018. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevA.98.062307>.

[91] MIKLIN, N.; OSZMANIEC, M. A universal scheme for robust self-testing in the prepare-and-measure scenario. *Quantum*, Verein zur Forderung des Open Access Publizierens in den Quantenwissenschaften, v. 5, p. 424, abr. 2021. ISSN 2521-327X. Disponível em: <http://dx.doi.org/10.22331/q-2021-04-06-424>.

[92] PAWŁOWSKI, M.; BRUNNER, N. Semi-device-independent security of one-way quantum key distribution. *Phys. Rev. A*, American Physical Society, v. 84, p. 010302, Jul 2011. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevA.84.010302>.

[93] AMBAINIS, A. et al. Quantum random access codes with shared randomness. *arXiv:quant-ph/0810.2937*, 2009.

[94] HAYASHI, M. et al. (4,1)-quantum random access coding does not exist—one qubit is not enough to recover one of four bits. *New Journal of Physics*, IOP Publishing, v. 8, n. 8, p. 129–129, aug 2006. Disponível em: <https://doi.org/10.1088/1367-2630/8/8/129>.

[95] GRUDKA, A. et al. When are popescu-rohrlich boxes and random access codes equivalent? *Phys. Rev. Lett.*, American Physical Society, v. 113, p. 100401, Sep 2014. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.113.100401>.

[96] CAVALCANTI, D.; SALLES, A.; SCARANI, V. Macroscopically local correlations can violate information causality. *Nature Communications*, Springer Science and Business Media LLC, v. 1, n. 1, Dec 2010. ISSN 2041-1723. Disponível em: <http://dx.doi.org/10.1038/ncomms1138>.

[97] DAM, W. van. *Nonlocality and Communication Complexity*. Tese (Doutorado) — University of Oxford, Department of Physics, 2000. Disponível em: <https://sites.cs.ucsb.edu/~vandam/oxford_thesis.pdf>.

[98] CHATURVEDI, A.; PAWLOWSKI, M.; HORODECKI, K. Random access codes and nonlocal resources. *Phys. Rev. A*, American Physical Society, v. 96, p. 022125, Aug 2017. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevA.96.022125>.

[99] AMBAINIS, A.; FREIVALDS, R. 1-way quantum finite automata: strengths, weaknesses and generalizations. *arXiv:quant-ph/9802062*, 1998. Disponível em: <https://arxiv.org/abs/quant-ph/9802062>.

[100] TANASESCU, A.; ILIESCU, V.-F.; POPESCU, P. G. Optimal entanglement-assisted almost-random access codes. *Phys. Rev. A*, American Physical Society, v. 101, p. 042309, Apr 2020. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevA.101.042309>.

[101] DORIGUELLO, J. F.; MONTANARO, A. Quantum random access codes for boolean functions. *Quantum*, Verein zur Forderung des Open Access Publizierens in den Quantenwissenschaften, v. 5, p. 402, mar. 2021. ISSN 2521-327X. Disponível em: <http://dx.doi.org/10.22331/q-2021-03-07-402>.

[102] ALVES, G. P.; GIGENA, N.; KANIEWSKI, J. m. k. Biased random access codes. *Phys. Rev. A*, American Physical Society, v. 108, p. 042608, Oct 2023. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevA.108.042608>.

[103] M, V. et al. Mutually unbiased balanced functions and generalized random access codes. *Phys. Rev. A*, American Physical Society, v. 104, p. 012420, Jul 2021. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevA.104.012420>.

[104] KUSHILEVITZ, E.; NISAN, N. *Communication complexity*. USA: Cambridge University Press, 1996. ISBN 0521560675.

[105] KUSHILEVITZ, E. Communication complexity. In: ZELKOWITZ, M. V. (Ed.). *Advances in Computers*. Elsevier, 1997. v. 44, p. 331–360. ISSN 0065-2458. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0065245808603423>.

[106] YAO, A. C.-C. Quantum circuit complexity. In: *IEEE Annual Symposium on Foundations of Computer Science*. [s.n.], 1993. Disponível em: <https://api.semanticscholar.org/CorpusID:2870099>.

[107] CLEVE, R.; BUHRMAN, H. Substituting quantum entanglement for communication. *Phys. Rev. A*, American Physical Society, v. 56, p. 1201–1204, Aug 1997. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevA.56.1201>.

[108] BUHRMAN, H. et al. Nonlocality and communication complexity. *Rev. Mod. Phys.*, American Physical Society, v. 82, p. 665–698, Mar 2010. Disponível em: <https://link.aps.org/doi/10.1103/RevModPhys.82.665>.

[109] CLEVE, R. et al. Quantum entanglement and the communication complexity of the inner product function. *arXiv:quant-ph/9708019*, 1998. Disponível em: <https://arxiv.org/abs/quant-ph/9708019>.

[110] BABAI, L.; FRANKL, P.; SIMON, J. Complexity classes in communication complexity theory. In: *27th Annual Symposium on Foundations of Computer Science (sfcs 1986)*. [S.l.: s.n.], 1986. p. 337–347.

[111] PAWŁOWSKI, M.; SCARANI, V. *Quantum Theory Informational Foundations and Foils: Information Causality*. [S.l.]: Springer Netherlands, Dordrecht, ISBN, 2016. 423 - 438 p.

[112] MIKLIN, N.; PAWŁOWSKI, M. Information causality without concatenation. *Phys. Rev. Lett.*, American Physical Society, v. 126, p. 220403, Jun 2021. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.126.220403>.

[113] JAIN, P.; GACHECHILADZE, M.; MIKLIN, N. *Information causality as a tool for bounding the set of quantum correlations*. 2024. Disponível em: <https://arxiv.org/abs/2308.02478>.

[114] AL-SAFI, S. W.; SHORT, A. J. Information causality from an entropic and a probabilistic perspective. *Physical Review A*, v. 84, n. 6, p. 042323, 2011.

[115] SHORT, A. J.; WEHNER, S. Entropy in general physical theories. *New Journal of Physics*, v. 12, n. 3, p. 033023, mar 2010. Disponível em: <https://dx.doi.org/10.1088/1367-2630/12/3/033023>.

[116] BARNUM, H. et al. Entropy and information causality in general probabilistic theories. *New Journal of Physics*, v. 12, n. 3, p. 033024, mar 2010. Disponível em: <https://dx.doi.org/10.1088/1367-2630/12/3/033024>.

[117] YEUNG, R. W. *Information Theory and Network Coding*. [S.l.]: Springer Science & Business Media, 2008.

[118] JAYNES, E. T. Information theory and statistical mechanics. *Phys. Rev.*, American Physical Society, v. 106, p. 620–630, May 1957. Disponível em: <https://link.aps.org/doi/10.1103/PhysRev.106.620>.

[119] SALGE, C.; POLANI, D. Digested Information as an Information Theoretic Motivation for Social Interaction. *Journal of Artificial Societies and Social Simulation*, v. 14, n. 1, p. 1–5, 2011. Disponível em: <https://ideas.repec.org/a/jas/jasssj/2010-27-2.html>.

[120] FRITZ, T.; CHAVES, R. Entropic inequalities and marginal problems. *IEEE Transactions on Information Theory*, Institute of Electrical and Electronics Engineers (IEEE), v. 59, n. 2, p. 803–817, Feb 2013. ISSN 1557-9654. Disponível em: <http://dx.doi.org/10.1109/TIT.2012.2222863>.

[121] CHAVES, R.; FRITZ, T. Entropic approach to local realism and noncontextuality. *Phys. Rev. A*, American Physical Society, v. 85, p. 032113, Mar 2012. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevA.85.032113>.

[122] CHAVES, R.; LUFT, L.; GROSS, D. Causal structures from entropic information: geometry and novel scenarios. *New Journal of Physics*, IOP Publishing, v. 16, n. 4, p. 043001, Apr 2014. ISSN 1367-2630. Disponível em: <http://dx.doi.org/10.1088/1367-2630/16/4/043001>.

[123] WEILENMANN, M.; COLBECK, R. Inability of the entropy vector method to certify nonclassicality in linelike causal structures. *Phys. Rev. A*, American Physical Society, v. 94, p. 042112, Oct 2016. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevA.94.042112>.

[124] CHAVES, R.; BUDRONI, C. Entropic nonsignaling correlations. *Phys. Rev. Lett.*, American Physical Society, v. 116, p. 240501, Jun 2016. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.116.240501>.

[125] BRAUNSTEIN, S. L.; CAVES, C. M. Information-theoretic bell inequalities. *Phys. Rev. Lett.*, American Physical Society, v. 61, p. 662–665, Aug 1988. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.61.662>.

[126] CHAVES, R. Entropic inequalities as a necessary and sufficient condition to noncontextuality and locality. *Phys. Rev. A*, American Physical Society, v. 87, p. 022102, Feb 2013. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevA.87.022102>.

[127] WEILENMANN, M.; COLBECK, R. Analysing causal structures with entropy. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, The Royal Society, v. 473, n. 2207, p. 20170483, nov. 2017. ISSN 1471-2946. Disponível em: <http://dx.doi.org/10.1098/rspa.2017.0483>.

[128] WEILENMANN, M.; COLBECK, R. Non-shannon inequalities in the entropy vector approach to causal structures. *Quantum*, Verein zur Forderung des Open Access Publizierens in den Quantenwissenschaften, v. 2, p. 57, mar. 2018. ISSN 2521-327X. Disponível em: <http://dx.doi.org/10.22331/q-2018-03-14-57>.

[129] UFFINK, J. Quadratic bell inequalities as tests for multipartite entanglement. *Phys. Rev. Lett.*, American Physical Society, v. 88, p. 230406, May 2002. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.88.230406>.

[130] NAVASCUÉS, M. et al. Almost quantum correlations. *Nature communications*, Nature Publishing Group, v. 6, n. 1, p. 1–7, 2015. Disponível em: <https://doi.org/10.1038%2Fncomms7288>.

[131] LINDEN, N. et al. Quantum nonlocality and beyond: Limits from nonlocal computation. *Phys. Rev. Lett.*, American Physical Society, v. 99, p. 180502, Oct 2007. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.99.180502>.

[132] FRITZ, T. et al. Local orthogonality as a multipartite principle for quantum correlations. *Nature Communications*, Springer Science and Business Media LLC, v. 4, n. 1, 2013. ISSN 2041-1723. Disponível em: <http://dx.doi.org/10.1038/ncomms3263>.

[133] ALLCOCK, J. et al. Closed sets of nonlocal correlations. *Phys. Rev. A*, American Physical Society, v. 80, p. 062107, Dec 2009. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevA.80.062107>.

[134] RAI, A. et al. Geometry of the quantum set on no-signaling faces. *Phys. Rev. A*, American Physical Society, v. 99, p. 032106, Mar 2019. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevA.99.032106>.

[135] BRITO, S. G. A. et al. Nonlocality distillation and quantum voids. *Phys. Rev. A*, American Physical Society, v. 100, p. 012102, Jul 2019. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevA.100.012102>.

[136] PATRA, R. K. et al. Principle of information causality rationalizes quantum composition. *Phys. Rev. Lett.*, American Physical Society, v. 130, p. 110202, Mar 2023. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.130.110202>.

[137] GACHECHILADZE, M. et al. Quantum bell inequalities from information causality – tight for macroscopic locality. *Quantum*, Verein zur Forderung des Open Access Publizierens in den Quantenwissenschaften, v. 6, p. 717, maio 2022. ISSN 2521-327X. Disponível em: <http://dx.doi.org/10.22331/q-2022-05-24-717>.

[138] COITEUX-ROY, X.; WOLFE, E.; RENOU, M.-O. No bipartite-nonlocal causal theory can explain nature's correlations. *Phys. Rev. Lett.*, American Physical Society, v. 127, p. 200401, Nov 2021. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.127.200401>.

[139] PANWAR, E.; PANDYA, P.; WIEśNIAK, M. An elegant scheme of self-testing for multipartite bell inequalities. *npj Quantum Information*, Springer Science and Business Media LLC, v. 9, n. 1, jul. 2023. ISSN 2056-6387. Disponível em: <http://dx.doi.org/10.1038/s41534-023-00735-3>.

[140] POLLYCENO, L. Code concerning the figure 14. *https://github.com/Pollyceno/TripartiteSlices*, 2022.

[141] POLLYCENO, L. Code concerning the results from the table 1. *https://github.com/Pollyceno/TripartiteViolation*, 10 2022.

[142] TONER, B. F.; BACON, D. Communication cost of simulating bell correlations. *Phys. Rev. Lett.*, American Physical Society, v. 91, p. 187904, Oct 2003. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevLett.91.187904>.

[143] TONER, B. Monogamy of non-local quantum correlations. *Proc. R. Soc. A*, v. 465, p. 59–69, 2009.

[144] PAWŁOWSKI, M. Security proof for cryptographic protocols based only on the monogamy of bell's inequality violations. *Phys. Rev. A*, American Physical Society, v. 82, p. 032313, Sep 2010. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevA.82.032313>.

[145] TONER, B.; VERSTRAETE, F. *Monogamy of Bell correlations and Tsirelson's bound*. 2006.

[146] XIANG, Y.; REN, W. Bound on genuine multipartite correlations from the principle of information causality. *Quantum Info. Comput.*, Rinton Press, Incorporated, Paramus, NJ, v. 11, n. 11–12, p. 948–956, nov 2011. ISSN 1533-7146.

[147] HSU, L.-Y. Multipartite information causality. *Phys. Rev. A*, American Physical Society, v. 85, p. 032115, Mar 2012. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevA.85.032115>.

[148] ADLAM, E. Tsirelson's bound and the quantum monogamy bound from global determinism. *arXiv:2011.08284v1*, 2021.

[149] ADLAM, E. Tsirelson's bound and the quantum monogamy bound from global determinism. *arXiv:2011.08284v2*, 2021.

[150] POLLYCENO, L. Code concerning the figure 16. *https://github.com/Pollyceno/TripartiteSlices*, 2022.

[151] HWANG, W.-Y.; GITTSOVICH, O. Comment on "security proof for cryptographic protocols based only on the monogamy of bell's inequality violations". *Phys. Rev. A*, American Physical Society, v. 85, p. 046301, Apr 2012. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevA.85.046301>.

[152] FANO, R. M. *Transmission of Information*. [S.l.]: M.I.T. Press, 1968.

[153] PAWŁOWSKI, M. Reply to "comment on 'security proof for cryptographic protocols based only on the monogamy of bell's inequality violations"'. *Phys. Rev. A*, American Physical Society, v. 85, p. 046302, Apr 2012. Disponível em: <https://link.aps.org/doi/10.1103/PhysRevA.85.046302>.

[154] NIELSEN, M. A.; CHUANG, I. L. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. [S.l.]: Cambridge University Press, 2010.

[155] PERES, A. *Quantum Theory: Concepts and Methods*. [S.l.]: Springer Netherlands, 2002. (57).

[156] COHEN-TANNOUDJI, C.; DIU, B.; FRANK, L. *Quantum Mechanics*. [S.l.]: Wiley, 1978. v. 1.

[157] WILLIAMS, H. P. Fourier's method of linear programming and its dual. *Mathematical Association of America*, v. 93, n. 9, p. 681–695, 1986. Disponível em: <http://eprints.lse.ac.uk/id/eprint/31577>.

[158] NAVASCUéS, M. et al. A physical approach to tsirelson's problem. *Foundations of Physics*, Springer Science and Business Media LLC, v. 42, n. 8, p. 985–995, Mar 2012. ISSN 1572-9516. Disponível em: <http://dx.doi.org/10.1007/s10701-012-9641-0>.

[159] TSIRELSON, B. Tsirelson's comments. *http://www.tau.ac.il/ tsirel/Research/bellopalg/-main.html*, . Disponível em: <http://www.tau.ac.il/~tsirel/Research/bellopalg/main.html>.

[160] SLOFSTRA, W. Tsirelson's problem and an embedding theorem for groups arising from non-local games. *Journal of the American Mathematical Society*, American Mathematical Society (AMS), v. 33, n. 1, p. 1–56, Sep 2019. ISSN 1088-6834. Disponível em: <http://dx.doi.org/10.1090/jams/929>.

Appendices

Chapter A

Convex Geometry

A set $\mathcal{C} \subseteq \mathbb{R}^n$ is termed convex if any line segment joining two points within $\mathcal{C}$ is entirely contained within $\mathcal{C}$. Formally, for any $x_1, x_2 \in \mathcal{C}$,

$$\alpha x_1 + (1 - \alpha)x_2 \in \mathcal{C}, \quad \alpha \in [0,1]. \tag{A.1}$$

A combination of the form $\theta_1 x_1 + ... + \theta_k x_k$ is referred to as a convex combination of points $x_1, ..., x_k$, and can be conceptualized as a mixture of these points with respective weights $\theta_i \geqslant 0$, such that $\theta_1 + ... + \theta_k = 1$. Figure 17 illustrates examples of convex and non-convex sets.

The smallest convex set that contains any arbitrary set $\mathcal{C}$ is termed the convex hull of $\mathcal{C}$, denoted by $\overline{\mathcal{C}}$. Formally, the convex hull is defined as the set of all convex combinations of elements within $\mathcal{C}$,

$$\overline{\mathcal{C}} = \left\{ \sum_i \theta_i x_i \in \mathbb{R}^n \mid x_i \in C,\ \theta_i \geqslant 0,\ \sum_i \theta_i = 1 \right\}. \tag{A.2}$$

By definition, a convex hull is intrinsically convex. This concept is visually represented in Fig.18 through various convex hull examples.

A.1 Cones

If for every $x \in \mathcal{C}$ and $\theta \geqslant 0$, it follows that $\theta x \in \mathcal{C}$, then $\mathcal{C}$ forms a *cone*. If $\mathcal{C}$ additionally satisfies the convexity condition as per Eq.(A.1), the set constitutes a *convex cone*. This implies that for any two points $x_1, x_2 \in \mathcal{C}$,

$$\theta_1 x_1 + \theta_2 x_2 \in \mathcal{C}, \quad \text{where} \quad \theta_1, \theta_2 \geqslant 0. \tag{A.3}$$

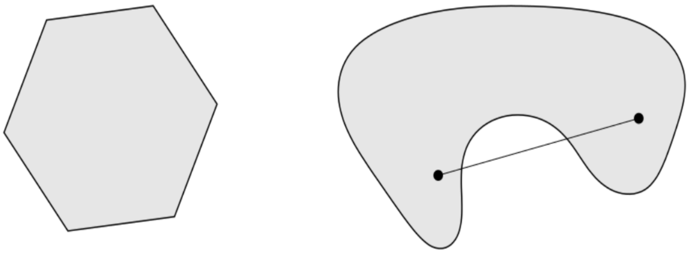


Figure 17 – Figure taken from [34]. On the left is a sketch of a convex set; on the right is an example of a non-convex set.

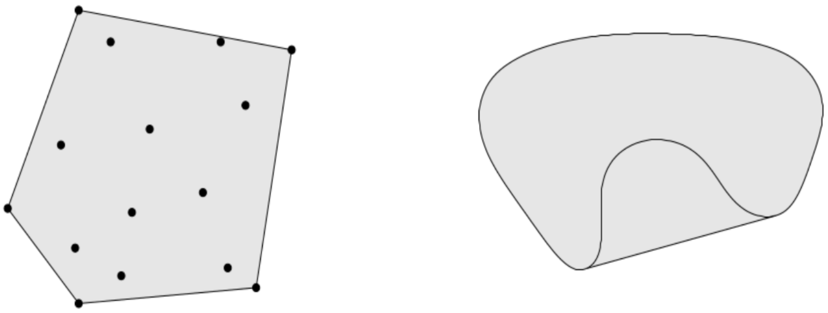


Figure 18 – Figure adapted from [34] with examples of convex hulls. On the left, we observe the smallest convex set containing a finite number of points, highlighted by black lines. On the right, we have the smallest convex set encompassing the non-convex set illustrated in Fig.17.

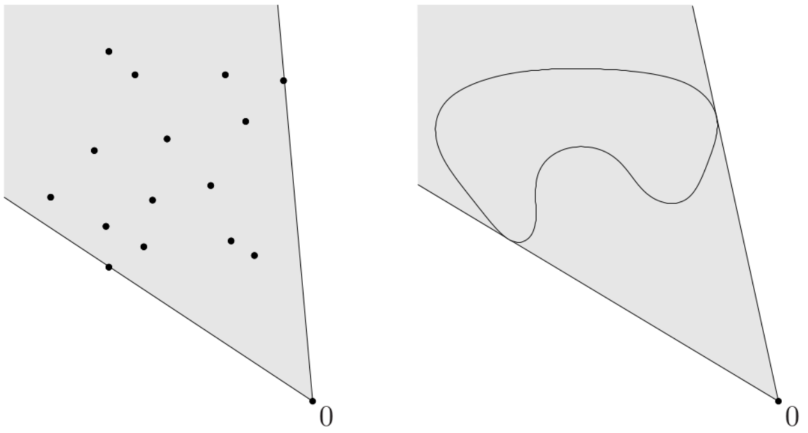


Figure 19 – Figure adapted from [34] depicting examples of conic hulls. On the left is the smallest convex cone containing a finite set of points, highlighted by black lines. The smallest convex cone encompassing the non-convex set from Fig.17 is displayed on the right.

Analogous to the previous case, the expression $\theta_1 x_1 + \cdots + \theta_k x_k$, with $\theta_1, \ldots, \theta_k \geqslant 0$, is termed a *conic combination* of the points $x_1, \ldots, x_k$. A necessary and sufficient condition for a set C to be a convex cone is for it to contain all conic combinations of its elements.

The *conic hull* of a set C is the collection of all conic combinations of points in C (see Fig.19),

$$\left\{ \sum_{i}^{k} \theta_i x_i \in \mathbb{R}^n \mid x_i \in C,\ \theta_1, \ldots, \theta_k \geqslant 0 \right\}. \tag{A.4}$$

A.2 Hyperplanes

A set of points x having a constant inner product $b \in \mathbb{R}$ relative to a vector $a \in \mathbb{R}^n$ defines a *hyperplane*,

$$\{x \in \mathbb{R}^n \mid a^T x = b\}. \tag{A.5}$$

The value b can be interpreted as an offset from the origin. Condition (A.5) divides the vector space $\mathbb{R}^n$ into two *half-spaces*, where a closed half-space represents the solution set of linear inequalities of the form,

$$\{x \in \mathbb{R}^n \mid a^T x \leqslant b\}. \tag{A.6}$$

Evidently, the boundary of (A.6) is the hyperplane defined in (A.5).

A.3 Polytope

A *polytope* can be defined as the solution set of a finite number of linear equalities and inequalities:

$$\mathcal{P} = \{x \in \mathbb{R}^n \mid a_i^T x \leqslant b_i,\ i = 1, \ldots, m,\ c_j^T x = d_j,\ j = 1, \ldots, p\}. \tag{A.7}$$

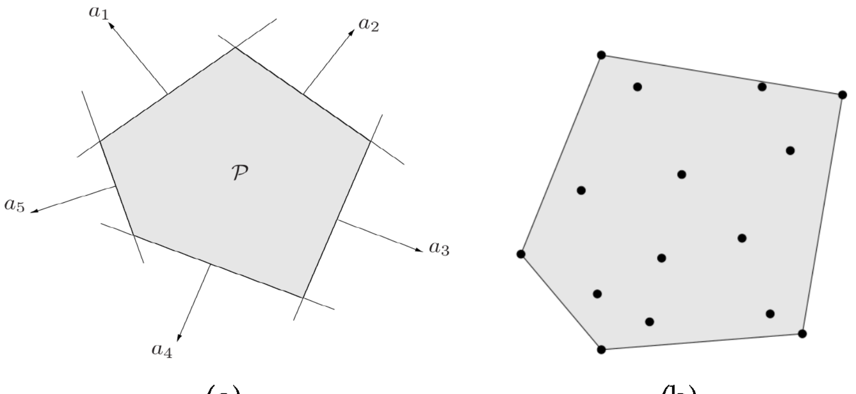
Thus, according to definitions (A.6) and (A.5), a polytope can be seen as the intersection of a finite number of half-spaces and hyperplanes. Alternatively, a polytope can also be defined from the convex hull of a finite set of points, i.e.,

$$\left\{\sum_i^k \theta_i x_i \in \mathbb{R}^n \mid \theta_1, \ldots, \theta_k \geqslant 0,\ \theta_1 + \ldots + \theta_k = 1\right\}. \tag{A.8}$$

Fig.20 presents examples of polytopes corresponding to both definitions.

Returning to Eq.(1.1), based on the convexity condition (A.1) from the previous section, we observe that the *set of behaviors* in Eq.(1.2) is convex. When considering the convex combination of two points $\mathbf{p}_1, \mathbf{p}_2 \in \mathcal{P}$, we have

$$\mathbf{p} = \alpha \mathbf{p}_1 + (1 - \alpha)\mathbf{p}_2, \quad \alpha \in [0, 1].$$



(a) (b)

Figure 20 – Figure adapted from [34] with examples of polytopes. On the left, we see the intersection of a finite number of half-spaces. On the right, we have the convex hull of a finite number of points, highlighted by black lines.

In this case, each component of **p** clearly satisfies the positivity condition. For a given component of **p**, we obtain $p(a|x) = \alpha p_1(a|x) + (1-\alpha)p_2(a|x)$. Summing over all possible outcomes a and using the normalization of p_1 and p_2, we have

$$\begin{aligned}\sum_a p(a|x) &= \sum_a \left[\alpha p_1(a|x) + (1-\alpha)p_2(a|x)\right],\\ &= \alpha \sum_a p_1(a|x) + (1-\alpha)\sum_a p_2(a|x),\\ &= \alpha + (1-\alpha) = 1.\end{aligned}$$

Thus, $\mathbf{p} \in \mathcal{P}$ confirms that the set of behaviors in Eq.(1.2) is convex. Furthermore, since the number of linear constraints defining $\mathcal{P}$ depends on the number of possible measurements and outcomes, by definition in Eq.(A.8), we conclude that the set in Eq.(1.2) is a polytope.

Chapter B

Notions of quantum theory used in this thesis

In this chapter, we will provide a concise overview of essential concepts within quantum theory that will be fundamental for the discussions in the subsequent chapters. For further details, the reader may refer to the references [154, 155, 156], which serve as the basis for this discussion.

B.1 A few definitions

As is well known, quantum theory is a probabilistic framework built upon mathematical postulates that facilitate the formulation of predictions. Although the theory does not delineate the physical laws governing the quantum realm, it associates mathematical objects with critical concepts necessary for physical description, such as state, measurement, interaction, and evolution.

In this work, we examine the non-local aspects of quantum theory, which typically considers only the two most fundamental elements: state and measurement.

The first postulate of quantum theory describes the state. It associates a physical system of interest with a Hilbert space in which an operator acting within this complex vector space fully describes the system's state.

Definition 1. *Quantum state*[1]*: It is an operator ρ that acts on the Hilbert space $\mathcal{H}$ and satisfies the following properties:*

(i) $\rho \geqslant 0$;

(ii) $\operatorname{Tr} \rho = 1$.

The operator ρ can be expressed in terms of the following convex combination:

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|, \tag{B.1}$$

where $|\psi_i\rangle\langle\psi_i|$ are projectors, $|\psi_i\rangle$ is a normalized vector belonging to $\mathcal{H}$, and p_i satisfies $\sum_i p_i = 1$ and $p_i \geqslant 0$. States described by a single projector, $\rho = |\psi\rangle\langle\psi|$, are known as *pure states*, while those represented in equation (B.1) are referred to as *mixed states*.

From this foundation, we can advance to another area of interest within the quantum formulation, namely measurement. The measurement postulate establishes a connection between another class of operators acting in $\mathcal{H}$ and the execution of an experiment involving the physical system in question. In a laboratory setting, the experimental procedure can be read as implementing a measurement choice x followed by acquiring

[1] Or density operator.

a result a. Accordingly, in the framework of quantum theory, each result a is associated with an operator $M_{a|x}$, such that given the state ρ and $M_{a|x}$, one can determine the probability of obtaining the outcome a in a realization of the experiment. Thus, we define quantum measurement as follows:

Definition 2. *Quantum measurement: It is defined by a set of m operators associated with m possible outcomes a, $\{M_{a|x}\}$, acting in $\mathcal{H}$, such that:*

(i) $M_{a|x} \geqslant 0 \quad \forall\, a;$

(ii) $\sum_a M_{a|x} = \mathbb{1}.$

Operators $M_{a|x}$ that satisfy the above properties are known as elements of a POVM[2] and the complete set $\{M_{a|x}\}$ is referred to as a POVM. In particular, there exists a class of POVMs, commonly referred to as *projective measurements*,

Definition 3. *Projective measurement: It is defined by a set of m operators associated with m possible outcomes a, $\{\Pi_{a|x}\}$, acting in $\mathcal{H}$, which are orthogonal projectors, satisfying:*

(i) $\Pi_{a|x}\Pi_{a'|x} = \delta_{a,a'}\Pi_{a|x};$

(ii) $\sum_a \Pi_{a|x} = \mathbb{1}.$

With such elements in hand, quantum theory allows us to obtain the probability distribution associated with the possible outcomes. Therefore, in a measurement process described by the POVM $\{M_{a|x}\}$, where the physical system is described by the state ρ, the probability of obtaining a result a when the measurement x is performed is given by the *Born rule*:

$$p(a|x) = \mathrm{Tr}(M_{a|x}\rho). \tag{B.2}$$

B.2 Composite Systems

Typically, we may wish to describe quantum systems with multiple degrees of freedom or even a system comprising more than one physical system. In such cases, the Hilbert space associated with the global system is given by the tensor product of the Hilbert spaces of the subsystems. In the bipartite case, the global Hilbert space $\mathcal{H}_{AB}$ is given by:

$$\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B, \tag{B.3}$$

and for extension to larger systems, one simply adds the respective Hilbert space using the tensor product. In this context, definition 1 remains valid, and any operator ρ_{AB} in $\mathcal{H}_{AB}$ that is positive and has unit trace is a permitted quantum state for the system.

[2] Acronym for 'positive operator-valued measure'.

In this scenario, if each subsystem A and B admits respective measurement choices x and y, with possible outcomes a and b, the jointly associated measurement operators defined in 2 are given by $\{M_{a|x} \otimes M_{b|y}\}$, with $M_{a|x}$ acting in $\mathcal{H}_A$ and $M_{b|y}$ acting in $\mathcal{H}_B$. The Born rule for the joint probability of obtaining outcomes a and b, when measurements x and y are performed, is then given by:

$$p(a,b|x,y) = \mathrm{Tr}(M_{a|x} \otimes M_{b|y}\, \rho_{AB}). \tag{B.4}$$

Furthermore, when dealing with multipartite systems, the description of only a portion of the known global system is given by the *reduced state*:

Definition 4. *The reduced state of a bipartite state ρ_{AB} for system A is given by:*

$$\rho_A = \mathrm{Tr}_B\, \rho_{AB}, \tag{B.5}$$

and similarly for subsystem B. The operator ρ_A acts in $\mathcal{H}_A$ and follows definition 1, thereby referred to as the quantum state of the system A.

Once the statistics regarding the outcomes of each subsystem can be obtained from ρ_A, when we wish to disregard subsystem B completely, the reduced state provides the optimal marginal description concerning subsystem A, and vice versa. Nevertheless, an important characteristic of quantum systems is that even when in possession of the marginal descriptions of each subsystem, it is not always possible to reconstruct the state of the global system; that is to say, in general,

$$\rho_{AB} \neq \rho_A \otimes \rho_B. \tag{B.6}$$

States allowing the factorized form, $\rho_A \otimes \rho_B$, are attributed to a description of completely uncorrelated systems. However, the expression in (B.6) may not merely stem from the existence of correlations between the subsystems. To illustrate this, let us consider a scenario in which the states ρ_A^i and ρ_B^i are prepared by two devices, where $i \in \{1, 2, ..., n\}$. Suppose a source randomly generates the indices i with probability $p(i)$ and communicates this information to both subsystems A and B. In that case, the preparation of the states may be correlated such that the description of the global state is represented by[3]

$$\rho_{AB} = \sum_i p(i) \rho_A^i \otimes \rho_B^i. \tag{B.7}$$

States that conform to such a description are referred to as *separable states*.

Nevertheless, as we will further explore, quantum theory permits composite states ρ_{AB} that cannot be accounted for by classical correlations. States that do not allow the decomposition in (B.7) are designated as *entangled states*.

[3] It is important to note that this representation is more general than $\rho_{AB} = \rho_A \otimes \rho_B$.

Chapter C

Informational-theoretic constraints for quantum causal structures

Generally, in information theory, we typically use information measures to investigate correlations, uncertainties, and information regarding random variables. These quantities, in turn, respect information-theoretic constraints. In classical information theory, these constraints can be summarized in terms of the non-negativity of Shannon information measures. *i.e.*, for any set of 3 variables X, Y, and Z [117]:

$$H(X) \geqslant 0; \tag{C.1a}$$
$$H(X,Y) \geqslant 0; \tag{C.1b}$$
$$H(X|Z) \geqslant 0; \tag{C.1c}$$
$$I(X:Y) \geqslant 0; \tag{C.1d}$$
$$I(X:Y|Z) \geqslant 0. \tag{C.1e}$$

These inequalities represent all possible insights obtainable through classical information theory regarding a set of random variables. This set of inequalities is referred to as the *basic inequalities*. It constitutes the fundamental set of *information inequalities* from which all *Shannon-type inequalities* can be derived [117]. For a set of n random variables, $X_1, X_2, ..., X_n$, to incorporate all restrictions implied by the basic inequalities, (C.1), it suffices to consider the set of *elemental inequalities*[1] defined by:

$$H(X_i|X_{\mathbb{N}-i}) \geqslant 0, \quad \text{where } i \in \mathbb{N}; \tag{C.2a}$$
$$I(X_i : X_j|X_K) \geqslant 0, \quad \text{where } i \neq j \text{ e } K \subset \mathbb{N} - \{i,j\}. \tag{C.2b}$$

The set in (C.2) is indeed minimal and significantly reduces the number of inequalities to consider. The proof is found in Section 14.6 in [117]. The number of elementary inequalities m in (C.2), for a set of n random variables, is given by:

$$m = n + \left(\begin{array}{c} n \\ 2 \end{array}\right) 2^{n-2}. \tag{C.3}$$

Geometrically, for a set of random variables $S = X_1, ..., X_n$, one can associate the vector $\mathbf{H} \in \mathbb{R}^{2^n-1}$, whose components are joint entropies for all possible subsets of S. For example, for $S = X_1, X_2$, we have:

$$\mathbf{H} = \left[\begin{array}{c} H(X_1) \\ H(X_2) \\ H(X_1, X_2) \end{array}\right].$$

A vector $\mathbf{H}$ is called *entropic* if there exists a joint probability distribution, $p(x_1, ..., x_n)$, that allows obtaining $\mathbf{H}$. This notion enables the definition of a region in $\mathbb{R}^{2^n-1}$, in which

[1] From which all basic inequalities can be derived.

entropic vectors lie:

$$\Gamma_n^* = \{\mathbf{H} \in \mathbb{R}^{2^n-1} | \mathbf{H} \quad \text{is} \quad \text{entropic}\}.$$

Shannon information measures can always be expressed in what is known as their *canonical form*, in terms of joint entropies, *i.e.*,

1. $H(A) \geqslant 0$;
2. $H(A|B) = H(A,B) - H(B) \geqslant 0$;
3. $I(A:B|C) = H(A,C) + H(B,C) - H(A,B,C) - H(C) \geqslant 0$.

Thus, in this geometric perspective, elementary inequalities (C.2) can be written as $\mathbf{G} \cdot \mathbf{H} \geqslant \mathbf{0}$, where $\mathbf{H}$ is the entropic vector and $\mathbf{G}$ is a matrix whose rows represent each of the inequalities, with elements corresponding to the respective coefficients for each component of the entropic vector. In this manner, the basic inequalities define a region in the non-negative orthant of $\mathbb{R}^{2^n-1}$ known as the *Shannon cone*,

$$\Gamma_n = \{\mathbf{H} \in \mathbb{R}^{2^n-1} | \mathbf{G} \cdot \mathbf{H} \geqslant 0\}.$$

Since the entropy functions of n random variables respect the basic inequalities, Γ_n^* is contained within Γ_n. Therefore, the basic inequalities serve as a necessary but not sufficient condition for a vector $\mathbf{H} \in \mathbb{R}^{2^n-1}$ to be entropic, thus:

$$\Gamma_n^* \subset \Gamma_n.$$

Moreover, the relationships among the random variables of interest $S = X_1, ..., X_n$ can also be represented geometrically. For instance, if the variables X_i and X_j are independent, then $p(x_i, x_j) = p(x_i)p(x_j)$ implies that

$$H(X_i, X_j) = H(X_i) + H(X_j). \tag{C.4}$$

Consequently, independence between the two variables forms a hyperplane, $I \cdot \mathbf{H} = 0$, which restricts the entropic vector in $\mathbb{R}^{2^n-1}$. More generally, the variables $S = X_1, ..., X_n$ follow causal relationships that define the causal structure of S and must be considered in the geometric description. A causal structure is described by the relation:

$$p(x_1, x_2, ..., x_n) = \prod_j p(x_j | pa_j), \tag{C.5}$$

where the set PA_j is referred to as the *Markovian parents* of X_j and includes all variables in S that exert some causal influence over X_j. Similar to (C.4), causal relationships define hyperplanes, $I \cdot \mathbf{H} = 0$, in $\mathbb{R}^{2^n-1}$, thereby constraining the entropic vector $\mathbf{H}$. The region defined by the intersection of all hyperplanes defined by the causal relationships, $\mathbf{I} \cdot \mathbf{H} = 0$, and the Shannon cone defines the restricted Shannon cone, $\Gamma_n \cap L_C$, which is a polytope:

$$\Gamma_n \cap L_C = \{\mathbf{H} \in \mathbb{R}^{2^n-1} | \mathbf{G} \cdot \mathbf{H} \geqslant 0, \mathbf{I} \cdot \mathbf{H} = 0\}.$$

In general, not all variables can be observed simultaneously for a given set of random variables $S = X_1, ..., X_n$. In this case, the *marginal scenario* of S is a collection of subsets, $\mathcal{M} = M_1, ..., M_{|\mathcal{M}|}$, in which $M_i \subseteq S$ for which one has access to the probability distribution $p(\mathbf{x}_{Mi})$. In the context of the entropic-geometric description, obtaining the marginal scenario of a set of variables entails excluding components of the entropic vector that cannot be part of the problem's description. This is equivalent to projecting the Shannon cone (or restricted Shannon cone) onto the subspace of observable variables. Computationally, this projection is achieved via the Fourier-Motzkin algorithm, which eliminates variables from a set of inequalities [157]. In this way, we obtain the entropic description of causal structures:

$$\Gamma_{\mathcal{M}} = \Pi_{\mathcal{M}}(\Gamma_n \cap L_C). \tag{C.6}$$

Chapter D

Computational details

In this appendix, we discuss the computational details for deriving the criterion in (3.27) within the tripartite scenario. The scenario under consideration involves 9 random variables, resulting in an entropic vector with $2^9 = 512$ components. This dimensionality makes direct elimination via the Fourier-Motzkin method impractical, as the problem scales doubly exponentially. An alternative approach to circumvent this issue is to consider a restricted set of inequalities for elimination—specifically, those capable of detecting the non-classicality of the protocol illustrated in Fig.12. For this purpose, we need to examine the entropy of the message from the proposed communication task.

As shown, the protocol utilizing the non-classical resource in (3.23) accomplishes the task with an entropy of $H(M_x, M_y) = 2$. However, a message with entropy $H(M_x, M_y) = 4$ would be required in a setting with only classical resources. This observation allows us to reduce the number of inequalities that define the restricted Shannon cone $\Gamma_n \cap L_C$ for this problem. The approach involves randomly removing an inequality from the set and then running a linear program to minimize $H(M_x, M_y)$, subject to the remaining constraints in $\Gamma_n \cap L_C$ and also constrained by the equality conditions for all components of the entropic vector implied by the protocol with box (3.23), except for those involving M_x and M_y since our objective is to minimize $H(M_x, M_y)$. If the minimized result in this procedure exceeds $H(M_x, M_y) = 2$, the initially removed inequality does not restrict $H(M_x, M_y)$ to be below 2. It thus does not detect the non-classicality of the protocol with (3.23), allowing it to be excluded. Conversely, the removed inequality is retained if minimization yields a result less than or equal to $H(M_x, M_y) = 2$. This process is repeated until no further inequalities can be excluded.

Through this method, we obtain a significantly reduced set of inequalities, allowing for Fourier-Motzkin elimination in the marginal scenario $\mathcal{M} = \{\{X_0, X_1, Y_0, Y_1, M_x, M_y, G_0\}, \{X_0, X_1, Y_0, Y_1, M_x, M_y, G_1\}\}$. Following elimination, we derive a set of inequalities that, by construction, are violated by the protocol shown in Fig.12. This procedure can be repeated to obtain diverse non-trivial constraints, thus achieving the criterion in (3.27).

Chapter E

NPA hierarchy

In this section, we discuss a convergent hierarchy of semidefinite programs that model approximations of the set of quantum behaviors.

This method, known as the *NPA hierarchy*, is named after the original work of Miguel Navascues, Stefano Pironio, and Antonio Acin in 2008 [40]. In essence, the method explores the following lemma:

Lemma 2. *Let $\mathcal{F} = F_1, F_2, \ldots, F_n$ be a set of linear operators acting on a Hilbert space $\mathcal{H}$. For any state ρ on $\mathcal{H}$, the Hermitian matrix $\Gamma(\rho, \mathcal{F})$ is positive semidefinite ($\Gamma \geq 0$), with entries given by*

$$\Gamma_{ij} = \mathrm{Tr}(\rho F_i^\dagger F_j). \tag{E.1}$$

Now, consider the following set of linear operators, $\mathcal{F}_1 = \left\{ \mathbb{1}, {M_{a|x}}^{a,x}, {M_{b|y}}^{b,y} \right\}$, and a bipartite Bell scenario, as discussed in Chapter 1, in which Alice and Bob's measurement processes are represented by the POVMs $\{M_{a|x}\}_{a,x}$ and $\{M_{b|y}\}_{b,y}$, respectively. In this context, given $\mathcal{F}_1$ and the state ρ, we can calculate the entries of the matrix $\Gamma^{(1)}$ using (E.1).

As discussed in Chapter 1, in this scenario, quantum behaviors are those components that can be derived using the Born rule, $p(a,b|x,y) = \mathrm{Tr}\left(\rho M_{a|x} \otimes M_{b|y}\right)$, defining the set of quantum behaviors $\mathcal{P}_Q$. An alternative way to define the quantum set is to assume that the behavior components are given by

$$p(a,b|x,y) = \mathrm{Tr}\left(\rho M'_{a|x} M'_{b|y}\right), \tag{E.2}$$

where $\rho, M'_{a|x}$, and $M'_{b|y} \in \mathcal{H}_A \otimes \mathcal{H}_B$, imposing commutativity between all measurements of Alice $M'_{a|x}$ and all measurements of Bob $M'_{b|y}$. The set of behaviors that can be obtained according to (E.2) is denoted by $\mathcal{P}_{Q'}$. It is clear that $\mathcal{P}_Q \subseteq \mathcal{P}_{Q'}$, since when $M'_{a|x} = M_{a|x} \otimes \mathbb{1}$ and $M'_{b|y} = \mathbb{1} \otimes M_{b|y}$, we retrieve the Born rule (1.19), with the joint measurement operators given by $M_{a|x} \otimes M_{b|y}$, while still preserving the condition $[M'_{a|x}, M'_{b|y}] = 0$. The equivalence between the two sets $\mathcal{P}_Q$ and $\mathcal{P}_{Q'}$ was an open question for several decades and became known as the *Tsirelson problem* [158, 159]. Only recently was a proof discovered showing that these sets are not equivalent [160]. Consequently, under the definition given by (E.2), it becomes evident that some elements $\Gamma_{i,j}$ will correspond to components of the behavior vector $\mathbf{p}$ in the device-independent approach, that is,

$$\Gamma^{(1)}_{i,j} = \mathrm{Tr}(\rho M_{a|x} M_{b|y}) = p(a,b|x,y). \tag{E.3}$$

In other cases, however, we may encounter indeterminate terms,

$$\Gamma^{(1)}_{i,j} = \mathrm{Tr}(\rho M_{a|x} M_{a'|x'}). \tag{E.4}$$

Lemma 2 ensures that if the behavior $\mathbf{p}$ is quantum, then values exist for the indeterminate entries (E.4) such that $\Gamma^{(1)}$ is positive semidefinite. Therefore, $\Gamma^{(1)} \geq 0$ constitutes a necessary condition for a behavior $\mathbf{p}$ to belong to the set $\mathcal{P}_{Q'}$. Behaviors satisfying $\Gamma^{(1)} \geq 0$ define a set $\mathcal{P}_{Q_1}$, which defines an approximation of the set $\mathcal{P}_{Q'}$ and, consequently, is also an approximation of the quantum set $\mathcal{P}_Q$, that is,

$$\mathcal{P}_Q \subseteq \mathcal{P}_{Q_1}. \tag{E.5}$$

$\mathcal{P}Q_1$ constitutes the first level of the NPA hierarchy. The second level of the hierarchy is reached by considering the following set of operators, $\mathcal{F}_2 = \mathcal{F}_1 \cup \{\{M_{a|x}M_{a'|x'}\}, \{M_{b|y} M_{b'|y'}\}, \{M_{a|x}M_{b|y}\}\}$. The behaviors $\mathbf{p}$ for which $\Gamma^{(2)} \geq 0$ define the set $\mathcal{P}_{Q_2}$, such that

$$\mathcal{P}_Q \subseteq \mathcal{P}_{Q_2} \subseteq \mathcal{P}_{Q_1}. \tag{E.6}$$

At the n-th level of the hierarchy, the matrix $\Gamma^{(n)}$ yields the set of behaviors $\mathcal{P}_{Q_n}$ such that $\Gamma^{(n)} \geq 0$. Thus,

$$\mathcal{P}_Q \subseteq \mathcal{P}_{Q_n} \subseteq \cdots \subseteq \mathcal{P}_{Q_1}. \tag{E.7}$$

A crucial important result from [40] is that, in the limit as $n \to \infty$, we have

$$\lim_{n\to\infty} \mathcal{P}_{Q_n} = \mathcal{P}_{Q'}. \tag{E.8}$$

Chapter F

Concatenation in a multipartite communication task

Here we extend the tripartite communication task from section 3.5 to a general multipartite scenario. Thus, consider N parts, among which $N-1$ are senders that initially have their respective bit-strings $\mathbf{x}^k = (X_1^k, X_2^k, \cdots, X_n^k)$, where $k \in \{1, 2, \cdots, N-1\}$. Each sender encodes a classical message M_k of size $m < n$ to the N^{th}-part, the receiver. This last one needs rightly compute one of n possible initial bits functions $f_j(X_j^1, X_j^2, \cdots, X_j^{N-1})$, by producing the guess G_j, where $j \in \{1, \cdots, n\}$. Just as in the main text, in addition to the classical messages, non-signaling correlations are allowed among all N parts.

Now consider a little more particular case, where $n = 2$ and $f_j = X_j^1 \oplus X_j^2 \oplus \cdots \oplus X_j^{N-1}$. Just as in the previously described tripartite scenario, we find such a particular multipartite communication task is trivialized by a generalization of the correlation (3.23) for the $(N, 2, 2)$ Bell scenario, *i.e.*

$$p(a_1, a_2, \cdots, a_N | x_1, x_2, \cdots, x_N) = \begin{cases} 1/2^{N-1} & \text{if} \quad \bigoplus_{k=1}^{N} a_k = \bigoplus_{k=1}^{N-1} x_k x_N; \\ 0 & \text{else.} \end{cases} \tag{F.1}$$

where a_k and x_k respectively denote the output and input of the part k. To see this, consider that the N parts perform the strategy depicted in Fig.12. That is, each sender performs the encoding $x_k = X_1^k \oplus X_2^k$ and $M_k = X_1^k \oplus a_k$, and the receiver computes the guess $G_j = \bigoplus_{k=1}^{N-1} M_k \oplus a_N$. In this case, by considering (3.43) we find

$$\begin{aligned} G_j &= \bigoplus_{k=1}^{N-1} (X_1^k \oplus a_k) \oplus a_N; \\ &= \left(\bigoplus_{k=1}^{N-1} X_1^k \right) \oplus \left(\bigoplus_{k=1}^{N} a_k \right) \\ &= \left(\bigoplus_{k=1}^{N-1} X_1^k \right) \oplus \left(\bigoplus_{k=1}^{N-1} x_k x_N \right) \\ &= \left(\bigoplus_{k=1}^{N-1} X_1^k \right) \oplus \left(\bigoplus_{k=1}^{N-1} (X_1^k \oplus X_2^k) x_N \right). \end{aligned} \tag{F.2}$$

Therefore, if the receiver chooses his measurement as $x_N = j$, when $j = 0$ we have $G_0 = X_1^1 \oplus X_1^2 \oplus \cdots \oplus X_1^{N-1}$, and for $j = 1$ we obtain $G_1 = X_2^1 \oplus X_2^2 \oplus \cdots \oplus X_2^{N-1}$. *i.e.*, the receiver always computes the functions perfectly and trivializes the communication task. It is clear that the task success is related to the probability of the non-signaling boxes working just as (3.43), *i.e.*, $p(a_0 \oplus a_1 \oplus \cdots \oplus a_{N-1} = x_0 x_{N-1} \oplus x_1 x_{N-1} \oplus \cdots \oplus x_{N-2} x_{N-1} | x_0, x_1, \cdots, x_{N-1})$. Thus, the probabilities that the receiver computes the function values f_1 and f_2 correctly are, respectively, given by

$$P_I = \frac{1}{2^{N-1}} \left[\sum_{x_1,...,x_{N-1}} p\left(\bigoplus_{k=1}^{N} a_k = \bigoplus_{k=1}^{N-1} x_k x_N | x_1, ..., x_{N-1}, x_N = 0 \right) \right]; \tag{F.3a}$$

$$P_{II} = \frac{1}{2^{N-1}} \left[\sum_{x_1,...,x_{N-1}} p\left(\bigoplus_{k=1}^{N} a_k = \bigoplus_{k=1}^{N-1} x_k x_N | x_1, ..., x_{N-1}, x_N = 1 \right) \right]. \tag{F.3b}$$

When the parts share (3.43), we have $P_I = P_{II} = 1$. However, by introducing a parameter $E \in [0,1]$, we can investigate other non-signaling behaviors by means of the following probability of success:

$$p\left(\bigoplus_{k=1}^{N} a_k = \bigoplus_{k=1}^{N-1} x_k x_N \right) = \frac{1}{2}(1+E). \tag{F.4}$$

The perfect correlations of behavior (3.43) are retrieved when $E = 1$, and uniform probabilities are retrieved when $E = 0$.

From this example, one can see that the concatenation approach, depicted in Fig. 13, can also be employed in this multipartite scenario. This is due to the fact that, to complete the task, it is sufficient for the receiver to know only $\bigoplus_{k=1}^{N-1} M_k$, instead of each message M_k. For instance, when $n = 4$, the senders can divide their bits into two pairs and perform the encoding just as in the previous strategy. Now, if instead of sending their respective messages, M_k^0 and M_k^1, the parts encode them in a third NS-box (3.43) by employing (F.2), the receiver is able to recover perfectly one of the functions $\bigoplus_{k=1}^{N-1} M_k^{i=0,1}$. This allows the parts to perform the same decoding one more time, resulting in perfect access by the receiver to one of the functions $f_0 = X_0^1 \oplus X_0^2 \oplus \cdots \oplus X_0^{N-1}$, $f_1 = X_1^1 \oplus X_1^2 \oplus \cdots \oplus X_1^{N-1}$, $f_2 = X_2^1 \oplus X_2^2 \oplus \cdots \oplus X_2^{N-1}$, or $f_3 = X_3^1 \oplus X_3^2 \oplus \cdots \oplus X_3^{N-1}$.

In the most general scenario, the receivers have, initially, $n = 2^K$ bits, share $n-1$ perfect copies of the non-signaling resource (3.43), and the senders and the receiver perform the strategy just as depicted in Fig. 13. Here, for each part k, we denote the output and input of the box i of the level l by $a_k^{i,l}$ and $x_k^{i,l}$, respectively. Thus, we may write the guess produced by the receiver as:

$$G_j = \left(\bigoplus_{k=0}^{N-1} M_k \right) \oplus \left(\bigoplus_{l=0}^{K-1} a_N^{i_l,l} \right), \tag{F.5}$$

where the box i_l is defined in terms of the box measured in the previous level, $i_l = 2i_{l-1} + z_l + 1$, when $l \geqslant 1$. In this case, the receiver performs measurements in K boxes, one in each level, among which $(K-r)$ are to $z_N^{i,l} = 0$ and r to $z_N^{i,l} = 1$, where $r = z_0 + z_1 + \cdots + z_{K-1}$. Just as in the single copy scenario, the task success is directly related to the probability that the $n-1$ non-signaling boxes behave as (3.43), *i.e.*, (F.4).

Thus, when $E < 1$, for each box, there exists a probability that the receiver output $a_{N-1}^{i,l}$ is wrong and the property $\bigoplus_{k=1}^{N-1} a_k^{i,l} = \bigoplus_{k=1}^{N-2} x_k^{i,l} x_N^{i,l}$ does not hold. However, if an even number of mistakes is produced in the outputs of the receiver, then they all cancel each other and the produced guess with (F.5) will be correct. Therefore, the success probability for the multipartite task with concatenation is equal to the probability that the receiver produces an even number of wrong outputs, *i.e.*:

$$p\left(G_j = \bigoplus_{k=0}^{N-2} X_j^k\right) = Q_{\text{even}}^{(K-r)}(P_I) \cdot Q_{\text{even}}^{(r)}(P_{II}) + Q_{\text{odd}}^{(K-r)}(P_I) \cdot Q_{\text{odd}}^{(r)}(P_{II}), \tag{F.6}$$

where P_I and P_{II} are defined in (F.3) and $Q_{\text{even}}^{(s)}(P)$ and $Q_{\text{odd}}^{(s)}(P)$ are given by

$$Q_{\text{even}}^{s}(P) = \sum_{j=0}^{\lfloor\frac{s}{2}\rfloor} \binom{s}{2j} (1-P)^{2j} P^{s-2j} = \frac{1}{2}(1 + (2P-1)^s); \tag{F.7a}$$

$$Q_{\text{odd}}^{s}(P) = \sum_{j=0}^{\lfloor\frac{s-1}{2}\rfloor} \binom{s}{2j+1} (1-P)^{2j+1} P^{s-2j-1} = \frac{1}{2}(1 - (2P-1)^s). \tag{F.7b}$$

These describe the probabilities of the receiver producing an even and an odd number of mistakes, respectively, after s measurements; P denotes the probability of obtaining the right output in a NS-box.

By inserting (F.7) in (F.6) and considering the bias from (F.4) in the probabilities from (F.3), we find the communication task success probability

$$p\left(G_j = \bigoplus_{k=0}^{N-2} X_j^k\right) = \frac{1}{2}(1 + E_I^{K-r} E_{II}^r), \tag{F.8}$$

where $E_i = 2P_i - 1$.

F.1 Multiple copies inequality

Here we prove the multipartite generalization of the multiple copies criterion (3.46), firstly derived in Ref. [20] for a strict bipartite scenario.

First of all, we need to prove a simplified lower bound for (3.30). So, rewriting the left-hand side summation argument in (3.30), we have

$$I(X_i^k : X_i^1, \cdots, X_i^{k-1}, X_i^{k+1}, \cdots, X_i^{N-1}, G_i) = \tag{F.9}$$

$$\begin{aligned} & H(X_i^k) - H(X_i^k | X_i^1, X_i^2 \cdots, X_i^{k-1}, X_i^{k+1}, \cdots, X_i^{N-1}, G_i) \\ =& 1 - H(X_i^k \oplus X_i^1 | X_i^1, X_i^2, \cdots, X_i^{k-1}, X_i^{k+1}, \cdots, X_i^{N-1}, G_i) \\ \geqslant& 1 - H(X_i^k \oplus X_i^1 | X_i^2, \cdots, X_i^{k-1}, X_i^{k+1}, \cdots, X_i^{N-1}, G_i). \end{aligned} \tag{F.10}$$

Here, we particularized to the case where every bit X_i^k is associated with a uniform distribution, $H(X_i^k) = 1$. Further, we considered the fact that $H(A|B,C) = H(A \oplus B|B,C)$, because knowing B results in the same uncertainty about A and $A \oplus B$, and $H(A \oplus B|B,C) \geqslant H(A \oplus B|C)$, *i.e.*, to remove the conditioning in B does not increase the uncertainty of $A \oplus B$. This same argument can be applied $N-2$ times in order to move every conditioned random variable in the right-hand side of (F.10):

$$I(X_i^k : X_i^1, \cdots, X_i^{k-1}, X_i^{k+1}, \cdots, X_i^{N-1}, G_i) \geqslant 1 - H(X_i^1 \oplus X_i^2 \oplus \cdots \oplus X_i^{N-1} \oplus G_i). \quad \text{(F.11)}$$

However, from the communication task, when $X_i^1 \oplus X_i^2 \oplus \cdots \oplus X_i^{N-1} \oplus G_i = 0$, we necessarily have $G_i = X_i^1 \oplus X_i^2 \oplus \cdots \oplus X_i^{N-1}$. Thus, the probability $p(X_i^1 \oplus X_i^2 \oplus \cdots \oplus X_i^{N-1} \oplus G_i = 0)$ is exactly the success probability of the receiver, $p(G_i = X_i^1 \oplus X_i^2 \oplus \cdots \oplus X_i^{N-1})$, while $p(X_i^1 \oplus X_i^2 \oplus \cdots \oplus X_i^{N-1} \oplus G_i = 1)$ is the complementary part. Therefore, the right-hand side term from (F.11) can be written in terms of the binary entropy, which in (3.30) finally yields:

$$(N-1)\sum_i^n (1 - h(p(G_i = X_i^1 \oplus X_i^2 \oplus \cdots \oplus X_i^{N-1}))) \leqslant \mathcal{I} \leqslant H(M_1, \cdots, M_{N-1}). \quad \text{(F.12)}$$

Notice that we considered the fact that the left-hand side has no dependence on the index k. Furthermore, the rightmost term in (3.30) does not appear in (F.12), because we are assuming a uniform distribution for every initial bit X_i^k.

At this point, we particularize our description to the concatenation strategy earlier described in appendix F. Here we rewrite the left-hand side summation in (F.12) in terms of the number of instances r where the receiver performed measurement $x_{n_j}^k = 1$, and substitute the concatenation success probability (F.8):

$$\begin{aligned} (N-1)\sum_i^n (1 - h(p(G_i = X_i^1 \oplus X_i^2 \oplus \cdots \oplus X_i^{N-1}))) = & \qquad \text{(F.13)} \\ (N-1)\sum_r^K \binom{K}{r}\left[1 - h\left(\frac{1 + E_I^{K-r}E_{II}^r}{2}\right)\right] & \\ \geqslant \frac{(N-1)}{2\ln 2}\sum_r^K \binom{K}{r}(E_I^2)^{N-r}(E_{II}^2)^r & \\ = \frac{(N-1)}{2\ln 2}(E_I^2 + E_{II}^2)^K, & \qquad \text{(F.14)} \end{aligned}$$

where we considered $1 - h\left(\frac{1+y}{2}\right) \geqslant \frac{y^2}{2\ln 2}$ and $E_i = 2P_i - 1$, from (F.3). After performing such encoding, each sender sends only a single bit message. Thus, $H(M_1, \cdots, M_{N-1})$ in (F.12) is always fixed in $N-1$, necessarily. Therefore, with (F.12) and (F.14), we find that when $E_I^2 + E_{II}^2 > 1$, the new proposed criterion (3.30) can always be violated by some concatenation protocol with K levels. Thus, we finally conclude the proof for the previously mentioned criterion in (3.46):

$$E_I^2 + E_{II}^2 \leqslant 1. \quad \text{(F.15)}$$

Chapter G

List of Publications

The content of this Thesis is based on results developed in the following papers:

Information causality in multipartite scenarios,
Lucas Pollyceno, Rafael Chaves, and Rafael Rabelo
Physical Review A, 107, (4), 042203, 2023.

Monogamy of nonlocality from multipartite information causality
Lucas Pollyceno, Anubhav Chaturvedi, Chithra Raj, Pedro R. Dieguez, Marcin Pawłowski
arXiv:2405.20115

Throughout this period, the author also contributed to the works:

Witnessing Nonclassicality in a Causal Structure with Three Observable Variables,
Pedro Lauand, Davide Poderini, Ranieri Nery, George Moreno, Lucas Pollyceno, Rafael Rabelo, Rafael Chaves
PRX Quantum, 4, (2), 020311, 2023.

Interplays between classical and quantum entanglement-assisted communication scenarios,
Carlos Vieira, Carlos de Gois, Lucas Pollyceno, and Rafael Rabelo
New J. Phys., 25, 113004, 2023.