Universidade do Minho
Escola de Engenharia

Raman Choudhary

# Classical, quantum, and post-quantum aspects of Kochen-Specker contextuality scenarios

Doctoral Programme in Informatics at the Universities of Minho, Aveiro, and Porto (MAP-i)

February 2025

Classical, quantum, and post-quantum aspects of Kochen-Specker contextuality scenarios
Raman Choudhary
UMinho | 2012

Universidade do Minho
Escola de Engenharia

Raman Choudhary

# Classical, quantum, and post-quantum aspects of Kochen-Specker contextuality scenarios

Doctoral Programme in Informatics at the Universities of Minho, Aveiro, and Porto (MAP-i)




Supervised by
Rui Soares Barbosa
Luís Paulo Santos
Shane Mansfield

February 2025

# Copyright and Terms of Use for Third Party Work

This dissertation reports on academic work that can be used by third parties as long as the internationally accepted standards and good practices are respected concerning copyright and related rights.
This work can thereafter be used under the terms established in the license below.
Readers needing authorization conditions not provided for in the indicated licensing should contact the author through the RepositóriUM of the University of Minho.

## License granted to users of this work:

*[Caso o autor pretenda usar uma das licenças Creative Commons, deve escolher e deixar apenas um dos seguintes ícones e respetivo lettering e URL, eliminando o texto em itálico que se lhe segue. Contudo, é possível optar por outro tipo de licença, devendo, nesse caso, ser incluída a informação necessária adaptando devidamente esta minuta]*



**CC BY**

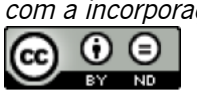
`https://creativecommons.org/licenses/by/4.0/` *[Esta licença permite que outros distribuam, remixem, adaptem e criem a partir do seu trabalho, mesmo para fins comerciais, desde que lhe atribuam o devido crédito pela criação original. É a licença mais flexível de todas as licenças disponíveis. É recomendada para maximizar a disseminação e uso dos materiais licenciados.]*



**CC BY-SA**

`https://creativecommons.org/licenses/by-sa/4.0/` *[Esta licença permite que outros remisturem, adaptem e criem a partir do seu trabalho, mesmo para fins comerciais, desde que lhe atribuam o devido crédito e que licenciem as novas criações ao abrigo de termos idênticos. Esta licença costuma ser comparada com as licenças de software livre e de código aberto «copyleft». Todos os trabalhos novos baseados no seu terão a mesma licença, portanto quaisquer trabalhos derivados também permitirão o uso*

*comercial. Esta é a licença usada pela Wikipédia e é recomendada para materiais que seriam beneficiados com a incorporação de conteúdos da Wikipédia e de outros projetos com licenciamento semelhante.]*



**CC BY-ND**

`https://creativecommons.org/licenses/by-nd/4.0/` *[Esta licença permite que outras pessoas usem o seu trabalho para qualquer fim, incluindo para fins comerciais. Contudo, o trabalho, na forma adaptada, não poderá ser partilhado com outras pessoas e têm que lhe ser atribuídos os devidos créditos.]*



**CC BY-NC**

`https://creativecommons.org/licenses/by-nc/4.0/` *[Esta licença permite que outros remisturem, adaptem e criem a partir do seu trabalho para fins não comerciais, e embora os novos trabalhos tenham de lhe atribuir o devido crédito e não possam ser usados para fins comerciais, eles não têm de licenciar esses trabalhos derivados ao abrigo dos mesmos termos.]*



**CC BY-NC-SA**

`https://creativecommons.org/licenses/by-nc-sa/4.0/` *[Esta licença permite que outros remisturem, adaptem e criem a partir do seu trabalho para fins não comerciais, desde que lhe atribuam a si o devido crédito e que licenciem as novas criações ao abrigo de termos idênticos.]*



**CC BY-NC-ND**

`https://creativecommons.org/licenses/by-nc-nd/4.0/` *[Esta é a mais restritiva das nossas seis licenças principais, só permitindo que outros façam download dos seus trabalhos e os compartilhem desde que lhe sejam atribuídos a si os devidos créditos, mas sem que possam alterá- los de nenhuma forma ou utilizá-los para fins comerciais.]*

# Acknowledgements

The work that went into this thesis happened over a span of three and a half years, which is a significant amount of time in anyone's life. Throughout this period, I repeatedly realized that for any project to come to fruition, very many things, even the ones outside the project, have to go right. So, the fact that I am at the finishing stage of my thesis summons many acknowledgments.

Firstly, I would first like to thank my PhD supervisor, Rui Soares Barbosa, who has been immensely kind and helpful right from day zero, when I was back home in India and we were applying for the PhD scholarship, to my first day in Portugal when he came to pick me up at the airport, and further throughout the final stages of this thesis. Throughout this period, his range and depth of knowledge have been invaluable to me. He gave me complete freedom to explore the topics of my liking which has built in me a lot of confidence to pursue my own projects and paths in the years to come. I am also thankful to him for the critical reading of this thesis and his gracious call-outs to my embarrassing mistakes. Thanks are also due to Adán. It is his works in contextuality that inspired me to pursue this field. And then, when I got to meet him for the first time, our meeting at the lunch table at INL became a seed to many future collaborations and discussions that continue up to this day. His child-like curiosity and disciplined work ethic are very contagious and inspiring. I am also thankful to Rafa, Laurens, Som, Lorenzo, along with all the collaborators who contributed constructively to my knowledge of quantum foundations. Thanks are also due to my co-supervisors Luís Paulo Santos and Shane Mansfield.

Zooming out a bit, a big thanks goes out to all the QLOC and Rossier group members who have always kept a healthy atmosphere in the office space. More than anything I am very fortunate to have formed deep friendships with so many of you. A blanket acknowledgment for all would be unfair, so, proceeding in the alphabetical ordering of names, I would like to thank each one of you. Starting from Angelos, with whom I share quite a similar sense of humor. Our knack for 'punny jokes' has only increased with time, yours more so than mine. I will never forget our whimsical participation in the half-marathon, and its aftermath; Then come Anita and Davide who have been a glue in keeping the people of our group together by consistently organising numerous events at their house and outside; Then we have Antonio

whose sharp wit and hilarious life stories have been very fun to take part in. Along with those our draining jiu-jitsu training sessions have been amazing; Then we have João who has been an amazing friend for all his wholesomeness. Our decision to consistently go for bouldering sessions has been the best; Then comes Laurens. I still remember all the fun discussions we had during our hikes and visits to different conferences. Equally amazing have been the gym sessions with you; Next comes my compatriot Manisha, who is the most unique of all. Your child-like energy and unabashed expressiveness has been a constant source of energy and fun to everyone; Michael is another person I am very thankful to because he helped me a lot during my first month of stay in Portugal and towards the end also to extract documents for the submission of this thesis; Next is Rafa who has been an amazing housemate and friend since day one. I am always filled with laughter thinking about your quirky beat-boxing drills. Thanks are also due to Alexandra, Som, Lorenzo, Filipa and Luisa.

Most importantly, I would also like to thank my lovely mom, Amaresh, and my father, Dheerandra, for keeping themselves healthy throughout this time and therefore letting me work peacefully so far away from them. I would also like to thank my sister, Naina, for visiting our parents' home from time to time when I could not. I am also thankful to my aunts, Kamlesh mosi and Manju mosi for displaying the valor that you did in the hardest times of your lives. Your courageous characters and calm demeanours are a blessing to this world. I also thank everyone in our extended family who has showered me with love and support from time to time. Special ones being Ipshita, Rahul, Parul didi, Rupal didi, Shekhar, Ayansh, and Vihaan.

Last but not least, I am grateful to FCT- Fundação para a Ciência e a Tecnologia- for providing the financial support through the PhD Grant SFRH/BD/151452/2021. I am also grateful to the Digital Horizon Europe project FoQaCiA, Foundations of Quantum Computational Advantage, GA no. 101070558 and FCT exploratory research project KaleidosQope, 2023.13603.PEX for funding crucial visits to conferences. I am also indebted to International Iberian Nanotechnology Laboratory (INL), Braga for hosting me throughout this research period.






# Statement of Integrity

I hereby declare having conducted this academic work with integrity.

I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

University of Minho, Braga, February 2025



Raman Choudhary

# Abstract

Bell nonlocality and Kochen-Specker (KS) contextuality are quantitative markers of non-classicality. They capture the non-intuitive aspects of quantum theory. The theory-independent formulation of these features informs us that any theory that agrees with the predictions of quantum theory manifests such non-classicality. Moreover, it unifies the two features where Bell non-locality can be seen as a special case of KS contextuality. In this thesis, we utilize this unified framework to independently analyze certain classical, quantum and post-quantum aspects of KS contextuality scenarios.

At the classical level we introduce a general method for lifting noncontextuality inequalities, deriving facets of noncontextual polytopes for more complex scenarios from known facets of simpler sub-scenarios. With our method we can produce facet-defining noncontextuality inequalities in all scenarios with contextual correlations, and we present examples of facet-defining noncontextuality inequalities for scenarios where no examples were known. This work strictly extends Pironio's result titled 'Lifting Bell inequalities' to lifting arbitrary KS noncontextuality inequalities.

At the quantum level, we analyse state-dependent contextuality aspects of multiqubit Pauli operators. In particular, we study realizability of $n$-cycle scenarios by $m$-qubit Pauli operators and then fully characterise these operators' ability to witness contextuality in $n$-cycle scenarios. We then provide a sufficient, but not necessary, condition for $m$-qubit Pauli operators to witness contextuality in arbitrary KS contextuality scenarios.

At the post-quantum level, we analyse the exclusivity principle's, dubbed E-principle, ability to detect post-quantum correlations. Oftentimes, a given correlation satisfies E-principle but two or more of its independent copies jointly violate the principle guaranteeing its post-quantumness. This is known as the activation effect of E-principle. We refine the conventional method for studying such activation effects, which involves finding certain cliques inside exclusivity graphs, to finding cliques inside an edge-colored exclusivity multigraph. This allows us to draw a novel connection to Ramsey theory that provides a powerful tool for ruling out E-principle violations. We exploit this connection, drawing on Ramsey-theoretic results to study violations of the E-principle by multiple copies of n-cycle PR boxes.

**Keywords.** Bell non-locality, contextuality, inequality lifting, quantum foundations.

# Resumo

A não-localidade de Bell e a contextualidade de Kochen-Specker (KS) são marcadores quantitativos de não-classicalidadeda teoria quântica. Nesta tese, utilizamos umenquadramento unificado para analisarcertos aspectos clássicos, quânticos e pós-quânticos de cenários de contextualidade de KS.

Ao nível clássico, introduzimos um método geral para elevar desigualdades de não-contextualidade, derivando facetas de polítopos de não-contextualidade de cenários mais complexos a partir de facetas conhecidas de sub-cenários mais simples. Com o nosso método, conseguimos produzir desigualdades de não-contextualidade que definem facetas em qualquer cenário que admita correlações contextuais, e apresentamos exemplos de desigualdades de não-contextualidade definidoras de facetas para cenários onde não se conheciam tais exemplos. Este trabalho estende estritamente o resultado de Pironio intitulado 'Lifting Bell inequalities' para a elevação de desigualdades arbitrárias de não-contextualidade de KS.

Ao nível quântico, analisamos aspectos de contextualidade dependente do estado testemunhada por operadores de Pauli de múltiplos qubits. Em particular, estudamos a realizabilidade de cenários de $n$-ciclo por operadores de Pauli de $m$ qubits e obtemos uma caracterização completa da capacidade de estes operadores testemunharem contextualidade em cenários de $n$-ciclo. Fornecemos também uma condição suficiente, mas não necessária, para que operadores de Pauli de $m$ qubits testemunhem contextualidade em cenários arbitrários de contextualidade de KS.

Ao nível pós-quântico, analisamos a capacidade do princípio de exclusividade, denominado princípio-E, para detectar correlações pós-quânticas. Frequentemente, uma dada correlação satisfaz o princípio-E, mas a combinação de duas ou mais cópias independentes viola este princípio, garantindo que a correlação é pós-quântica. Este fenómeno é conhecido como o efeito de activação do princípio-E. Refinamos o método convencional para estudar tais efeitos de ativação, que envolve encontrar certos cliques dentro de grafos de exclusividade, para a pesquisa de cliques dentro de um multigrafo com uma coloração de arestas. Isto permite-nos estabelecer uma nova ligação com a teoria de Ramsey, que fornece uma ferramenta poderosa para descartar violações do princípio-E. Exploramos esta conexão, recorrendo a resultados da teoria de Ramsey para estudar violações do princípio-E por múltiplas cópias de caixas PR de $n$-ciclo.

**Palavras-chave.** Contextualidade, elevação de desigualdade, fundamentos Quânticos, não localidade de Bell.

# Contents


**1 Introduction** **2**

1.1 Motivation . . . . . 2

1.2 Contributions . . . . . 4

**2 Background** **7**

2.1 Conceptual background . . . . . 7

2.1.1 KS contextuality and Bell non-locality . . . . . 7

2.1.2 Bell non-locality is a special case of KS contextuality . . . . . 11

2.1.3 Physical setting of KS contextuality scenarios . . . . . 13

2.2 Mathematical background . . . . . 15

2.2.1 Representing KS contextuality scenarios . . . . . 15

2.2.2 Correlations . . . . . 17

2.2.3 Convex polytopes . . . . . 18

2.2.4 No-disturbance and the noncontextual polytopes . . . . . 20

2.3 Fundamental KS scenarios . . . . . 22

2.3.1 Vorob'ev's theorem . . . . . 22

2.3.2 Cycle scenarios . . . . . 23

**3 Lifting noncontextuality inequalities** **25**

3.1 Overview . . . . . 27

3.2 Visualising lifting . . . . . 28

3.3 Lifting: from Bell to KS . . . . . 32

3.4 Lifting noncontextuality inequalities . . . . . 33

3.4.1 Notation . . . . . 33

3.4.2 Proof outline . . . . . 34

3.4.3 Measurement lifting . . . . . 35

3.4.4 Outcome lifting . . . 47
3.5 Applications . . . 52
3.5.1 Lifting the CHSH inequality . . . 53
3.5.2 Lifting the KCBS inequality . . . 54
3.5.3 Further remarks . . . 55
3.6 Discussion . . . 56

**4 Contextuality with Pauli observables in cycle scenarios 57**
4.1 Overview . . . 59
4.2 Preliminaries . . . 60
4.3 Constraints on faithful realizations . . . 61
4.3.1 Forbidden edge Paulis . . . 61
4.3.2 Commutativity conditions . . . 62
4.4 Impossible faithful realizations . . . 62
4.5 Possible faithful realizations . . . 65
4.5.1 m-cycle with m-qubit Paulis . . . 66
4.5.2 (m+2)-cycle with m-qubit Paulis. . . . 67
4.5.3 Cycle realizations from Path graph realizations . . . 67
4.5.4 2m or a 2m-1 sized cycle with m-qubit Paulis . . . 71
4.6 Contextuality with Pauli cycles . . . 72
4.6.1 4-cycle scenario . . . 72
4.6.2 5-cycle scenario . . . 74
4.6.3 n > 4 cycle scenarios . . . 75
4.7 General compatibility graphs . . . 76
4.7.1 Implication of 4-cycle contextuality . . . 77
4.7.2 Two induced 5-cycles . . . 77
4.8 Discussion . . . 80

**5 Exclusivity principle, Ramsey theory, and n-cycle PR boxes 82**
5.1 Overview . . . 84
5.2 Preliminaries . . . 85
5.2.1 Exclusivity graphs . . . 85
5.2.2 The exclusivity principle . . . 86

5.2.3 Joint exclusivity graph of independent boxes . . . 87
5.3 Refining the joint exclusivity graph . . . 88
5.3.1 Joint multigraph of independent boxes . . . 88
5.3.2 Monochromatic projection of $\Gamma^k$ . . . 89
5.3.3 Ramsey-theoretic constraints . . . 93
5.4 Exclusivity graph of n-cycle PR boxes . . . 94
5.4.1 Violation condition . . . 95
5.5 Result highlights . . . 95
5.5.1 k = 1 . . . 96
5.5.2 k = 2 . . . 96
5.5.3 k = 3 . . . 97
5.5.4 k > 3 . . . 98
5.6 Proofs . . . 98
5.6.1 Exclusivity graphs of n-cycle PR boxes . . . 98
5.6.2 Impossibility of extending a trivial $K_{2^k}$ to $K_{2^k+1}$ . . . 102
5.6.3 Activation effects of $n$-cycle PR boxes . . . 103
5.6.4 Exploiting Ramsey theory . . . 106
5.7 Discussion . . . 108

**6 Summary and future work** **110**

6.1 Summary . . . 110
6.2 Future work . . . 111
6.3 Other work . . . 113

# List of Figures


1 Two KS contextuality witnessing arrangements of qubit Pauli operators (a) Magic square, also called Peres-Mermin square: Operators in each row and column multiply to produce $+I$ (green), except the last column where it is $-I$ (red) (b) Magic Pentagram, also called Mermin star: Operators in each edge of the star multiply to produce $+I$ (green) except one edge it is $-I$ (red). . . . . 8

2 CHSH scenario: A preparation source sending particles to two space-like separated parties Alice and Bob who make local measurements on the received systems. . . . . 10

3 Sequential CHSH scenario: Alice and Bob making time-like separated compatible ideal measurements . . . . . 14

4 Categorising KS scenarios based on implementations . . . . . 15

5 Compatibility graphs (a) bi-partite graph representing a Bell scenario (b) Pentagon, a non $n$-partite graph, representing KCBS scenario, the simplest non-trivial non-Bell-like scenario. 16

6 Two equivalent representations of the same polytope (a) V-representation where knowing the vertices (black dots) is enough to define any point, via convex combinations, of the polytope, (b) minimal H-representation where an intersection of the facet-defining hyperplanes exactly captures the polytope. . . . . 19

7 A path graph depicting compatibility relations between three measurements $A, B, C$. . 23

8 An $n$-cycle compatibility graph with measurements $A_0, A_1, \ldots, A_{n-1}$. . . . . 23

9 Depiction of measurement lifting in simple scenarios with dichotomic measurements. (a) Case I measurement lifting, from the scenario **S** with two incompatible measurements to the scenario **T** with three pairwise-incompatible measurements. The shaded line and directional arrow on the left represent a facet-defining inequality saturated by vertices $v_{10}$ and $v_{11}$, while the shaded plane and arrow on the right represent the lifted inequality, saturated by all extensions of $v_{11}$ and $v_{10}$ to $T$. (b) Case II measurement lifting, from the scenario with a single measurement to the scenario with two compatible measurements. The facet-defining inequality saturated only by $v_0$, depicted on the left, is lifted to a facet-defining inequality saturated by both extensions of $v_0$ ($v_{00}$ and $v_{01}$) plus one extension of $v_1$ ($v_{11}$), depicted on the right. . . . . 30

10 (a) Compatibility graph of the initial scenario with the node labelled '4' representing a measurement added to the CHSH scenario. Bi-coloured nodes represent dichotomic measurements and edges represents compatibility between measurements. (b) Compatibility graph of scenario obtained from (a) by adding a new measurement (node labelled '5') compatible with all the measurements effectively contributing to inequality (3.27). . 47

11 Compatibility graph of the antiheptagon ($\bar{C}_7$) scenario with bi-coloured nodes representing dichotomic measurements. A facet-defining NC inequality is obtained for this scenario via sequential measurement lifting of the CHSH inequality from the sub-scenario on measurements $\{0, 1, 2, 3\}$. . . . . 53

12 (a) Compatibility graph of the initial KCBS $5$-cycle scenario with five dichotomic measurements, represented by bi-coloured nodes. (b) Compatibility graph of the $5$-cycle scenario with trichotomic measurements, represented by tri-coloured nodes. Inequality (3.38) is a facet-defining NC inequality for this extended scenario, obtained via sequential outcome lifting of the KCBS inequality. . . . . 54

13 A faithful 3-qubit Pauli realization of a 5-cycle scenario: each Pauli corresponding to the node of the 5-cycle is a 'cycle Pauli' while the operator $XXI$ being the product of two commuting cycle Paulis, $XII$ and $IXI$, is an 'edge Pauli'. . . . . 61

14 In a faithful 2-qubit Pauli realization of a 7-cycle (in green) we introduce two edge Paulis (in pink) $L_0 = P_0P_1$ ; $L_3 = P_3P_4$ and some more edges allowed by the four conditions obtained in Sec.4.3.2. . . . . 63

15 An arbitrary, faithful $m$-qubit Pauli realization of an $n$-cycle (green) with the corresponding edge Paulis $L_i = P_i P_{i\oplus 1}$ (pink), the dotted lines refers to the Paulis (including edge Paulis) not drawn but are part of the cycle. . . . . 64

16 Realization of $m+2$ sized cycle by $m$-qubit Paulis (a) $m=3$ (b) $m=4$ (c) $m=5$. . 68

17 Converting a path graph into a cycle by addition of a qubit (a) a 3-qubit path graph realization of $H_4$ (b) Adding a qubit and assigning anti-commuting Pauli operators, $Y$ and $Z$, to the first and last nodes and $I$ to the rest to get a 4-qubit realization of $C_4$. . . 68

18 Concatenating a faithful $m$-qubit Pauli realization of path graph $H_l$ with a faithful $m'$-qubit Pauli realization of $H_{l'}$, to obtain a $(m+m')$-qubit Pauli realization of $H_{l+l'-2}$. This works by introducing $m'$-qubit Identity operators in the first $l-2$ operators of the realization of path graph $H_l$ while introducing the Pauli realization for $H_{l'}$ from the $(l-1)$th node in $H_l$. In the additional $l'-2$ nodes, one then fixes $P_l$ in the first $m$ qubits. . . . . 70

19 A faithful 3-qubit Pauli realization of a $H_8$. . . . . 71

20 Two 5-cycles joined together at a node: a graph achievable only with $m \geq 4$ qubit Pauli operators. No such arrangement of Paulis can ever produce contextuality. . . . . 78

21 Two 5-cycles joined together: a graph achievable when $m \geq 3$ qubit Pauli operators. Such a graph can't produce contextuality for any realization of Pauli operators. . . . . 78

22 Two 5-cycles sharing two edges $\{P_1, P_5\}, \{P_4, P_5\}$: a graph achievable $\forall\ m \geq 2$: (a) represents the generic $m$-qubit Paulis and respective context label on the edges. (b) exemplifies one realization with Paulis that violates a noncontextuality inequality of the scenario. . . . . 79

23 Exclusivity graph of the PR box . . . . . 86

24 An example illustrating multicolor product of two graphs $G_1$ and $G_2$. The multigraph obtained can be thought to have multiple layers of nodes where each layer is isomorphic to $G_1$ (resp. $G_2$) with one of the colors, while the graph obtained by picking one node from each layer and the edges of the other color is necessarily isomorphic to $G_2$ (resp. $G_1$) as shown in the middle (resp. rightmost) representation above. . . . . 89

25 The joint exclusivity multigraph constructed in Figure 24 and its monochromatic projections. 90

26 An edge between non-adjacent nodes in a cycle divides it into two smaller cycles. If the original cycle is odd, a chord necessarily leads to (a) an even and an odd cycle. If the original cycle is even, a chord can divide the cycle into (b) two even cycles or (c) two odd cycles. . . . . 92

27 Exclusivity graphs of the $n$-cycle PR boxes: (a) when $n$ is even: Möbius ladder graph ($M_{2n}$) (b) when $n$ is odd: circular ladder graph, also popularly known as Prism graph ($Y_n$). Drawn in red are the smallest odd cycles for the two cases where $e, f, t$ are the events used in Theorem 18 to illustrate the construction of these cycles. . . . . 99

28 Interchanging nodes of the middle context (a) both consecutive criss-cross connections become vertical (b) a cross connection and a vertical one beneath it turn into vertical and cross connections respectively. . . . . 99

29 Reference configuration for the exclusivity graph of PR-box. Positively correlated events of a given context, when the joint outcome $++$ is on the left, connect in a cross-like manner to the context under it while negatively correlated events of a given context, when the joint outcome $+-$ is on the left, arrange in a vertical manner to the context under it. . . . . 100

30 Edge bifurcation of a $K_5$ exclusivity graph into two edge-wise disjoint $C_5$ graphs. The outer one denoted in red and the inner one in grey. . . . . 102

31 Constructing a $K_5$ using two independent $C_5$ cycles via multicolor product: (a) events coming from PR-box whose exclusivity graph is colored in red (b) events coming from PR-box whose exclusivity graph is colored in grey (c) joint events from the two copies leading to a $K_5$. . . . . 104

32 Impossibility of extending a $C_5$ in (a) with one edge supported by $n$-cycle PR-box whose exclusivity graph is colored grey while the rest of the edges are supported by the $n$-cycle PR box whose exclusivity graph is colored red, to (b) a $K_5$ for all $n \geq 6$ cycle scenarios since it is inconsistent with Corollary 17. . . . . 104

33 Impossibility of extending a $C_5$, with two edges supported by events from an $n$-cycle PR box whose exclusivity graph is colored grey while the other edges are supported by an $n$-cycle PR box whose exclusivity graph is colored red, to a $K_5$: (a) when the two grey edges are adjacent, the extension leads to violation of Corollary 17 as shown in (b) with nodes $\{u, w, x\}$; (c) when the two grey edges are not adjacent, the extension leads to violation of Corollary 19 due to a $C_5$ with nodes $\{u, v, x, w, y\}$ as shown in (d). . . . . 105

34 Two edge-colored $K_4$ where each monochromatic sub graph is a bi-partite graph. . . . . 107

# List of Tables

1 Correlation table for the PR box . . . . . 85

2 Relationship between $k$ and $n$ for the activation effects of E-principle with $n$-cycle PR boxes. For a given $k$ number of independent copies on the left: on the right, red marks that the $n$-cycle PR box does not violate the E-principle, blue indicates that there is a violation, while gray indicates no knowledge of violation hence an open problem for future work. Results known prior to this work were only those in the first column ($n = 4$). . . . 97

# Chapter 1

# Introduction

## 1.1 Motivation

Quantum theory (QT) has arguably been the most precisely tested theory in the history of science. Originally developed more than a century ago to explain the phenomenon of black body radiation, the theory soon became a tool for studying the atomic world. Since then, quantum theory's predictions keep being experimentally verified and the theory now undergirds most of modern physics with the exception of the theory of gravity.

Despite such success, QT has kept physicists and philosophers equally puzzled. This is because although the theory successfully plays its role as predictive machinery, it remains completely silent on what its mathematical elements physically mean. Debates since its inception have been going on to sort out this puzzle. Einstein's inhibition on quantum theory's probabilistic nature is captured by the popular quote (from his letter to Max Born in 1926) "I am at all events convinced that *He* does not play dice". His famous work along with Boris Podolsky and Nathan Rosen in 1935 attempted to understand the implications quantum theory had on reality. The title of the work is representative of the metaphysical challenges that quantum theory presented: "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" There are many examples of debates, discussions and rebuttals that scientists and philosophers engaged in at the time, but long story short, the same debates still continue.

We still don't know how to conceptualize the reality that underlies quantum phenomena but we have made progress in understanding what it will not look like. This has been unveiled by various foundational no-go theorems and their experimental verifications. These no-go theorems from Bell [1], Bell-Kochen-Specker [2, 3], and Spekkens [4] prove the inconsistency between our classically tutored intuitions and any theory that agrees with the predictions of quantum theory, forcing physicists and philosophers alike to overhaul our stance on the nature of Nature.

Many physicists look away from such challenges by dismissing them as "outside the scope of physics"

and go about their day to day business with the "shut up and calculate" approach. The field of quantum foundations, on the other hand, braves these challenges with the slogan "shut up and contemplate". How these contrasting – operationalist vs. realist – approaches have led us to arrive at modern physics is wonderfully outlined in Ref. [5]. One motivation for the latter approach is that QT is most likely not the end of the road, hence, one day when it breaks down knowing the principles that underpin its verified predictions can help build a more general theory. Furthermore, it is evident that chasing after these challenges has, as a spin-off, given us ground-breaking applications like quantum random number generator (QRNG), quantum key distribution (QKD) with fundamental guarantees of randomness and security respectively, thanks to the above-mentioned no-go theorems.

As a matter of fact, finding new applications powered by non-classical features of quantum theory is now one of the central motivations, rather than a peripheral one, to investigate quantum foundations. This proposition is further bolstered by the fact that different quantum computation models demonstrably require non-classicality of quantum theory to attain universality, e.g. KS contextuality has been identified as a necessary condition for universal quantum computation via magic state distillation [6]. Similarly, in the measurement based quantum computation (MBQC) circuit, with mod 2 linear classical processing, contextuality has been proven to be a necessary condition to compute any non-linear Boolean function [7]. In Ref. [8, 9], quantum shallow circuits have been proven to perform certain computations way more efficiently than their classical counterparts, where quantum contextuality seems to be powering this advantage. This means that a deeper understanding of the above-mentioned no-go theorems is necessary for the development of future technologies based on quantum theory.

Each of these no-go theorems assumes a definition of classicality based on which a (classical) model derives all possible (classical) statistics for a given experimental situation formally called a 'scenario'. For Bell's theorem, the definition of classicality is local causality i.e. in each round of statistics generation given a preparation, the probability of each outcome of a given measurement depends upon a common cause and not on any other space-like separated measurement; for Kochen-Specker's theorem it is non-contextual outcome-determinism i.e. outcome of each measurement is pre-determined, in each round of the experiment, irrespective of which other compatible measurements (also called context) are being performed with it; for Spekkens' theorem it is the idea that operationally indistinguishable procedures must correspond to the same reality i.e. must attain the same representation in the ontological model. Quantum predictions violate each of these assumptions.

Additionally, these notions are quite connected: Fine's theorem tells us that a deterministic locally causal model produces the same set of statistics as a probabilistic locally causal model. The former

is as if the KS notion of classicality has been applied to space-like separated measurements, a natural setting for Bell scenarios. Therefore, KS models already capture locally causal statistics as a special case. Spekkens' notion of classicality, on the other hand, is the most general of all as it subsumes the other two, and furthermore applies to preparation and transformation procedures.

To better understand nature it is imperative to gain more insights about these non-classical features. The same understanding is also crucial to build more general theories in the face of any observed limitations of quantum theory. More practically, one also hopes that such knowledge can lead to a more systematic approach to utilising and harnessing these non-classical features as resources.

## 1.2 Contributions

This thesis is concerned with understanding different facets of *KS scenarios*. With differing motivations, we independently analyzed different classical, quantum and post-quantum aspects of these scenarios. We now outline the main contributions of this thesis.

- **Classical aspects.** We analysed the geometric connection between the classical polytopes of two scenarios where one scenario is an extension of the other via addition of measurements and/or outcomes. The connection is called *lifting* where facet-inducing inequalities of the classical polytope of the smaller scenario help derive facet-inducing inequalities of the classical polytope of the bigger scenario. The upshot of this result is that once the facets of a simpler classical polytope have been fully or partially identified, one need not start from scratch when searching for the facets of a more complex classical polytope. This result non-trivially extends lifting of Bell inequalities shown by Pironio in Ref. [10]. The non-triviality is established due to the fact that there are more ways to extend an arbitrary KS scenario to another one, compared to extending a Bell scenario to another Bell scenario. Furthermore, enabled by Vorob'ev's theorem [11], which guarantees that any contextuality-witnessing scenario contains an induced $n$-cycle sub-scenario (for some $n \geq 4$), and by the complete characterisation of the noncontextual polytopes for all such cycle scenarios [12], our lifting results allow us to derive many facet-inducing inequalities for any arbitrary KS contextuality-witnessing scenario. One may instead concentrate efforts on finding the facet-defining inequalities that are absent in or do not arise from simpler scenarios. This is important since finding all the noncontextuality inequalities for an arbitrary scenario encoded as a compatibility graph is an NP-complete problem [13], and our method can help prevent redundancies in obtaining facet inequalities. This work has been presented in Chapter 3, and has been published as Phys. Rev. A

109, 052216 [14].

- **Quantum aspects.** We analysed state-dependent contextuality relative to multi-qubit Pauli observables. These operators play an important role in various models of quantum computation like MBQC, PBC, etc. With our analysis we proved that for $m$-qubit Pauli operators to violate some non-contextuality inequality of an arbitrary scenario, it is sufficient for the scenario to have an induced 4-cycle as its sub-scenario. We further showed that the same is not a necessary condition by providing an example of a scenario that doesn't have any induced 4-cycle as its sub-scenario. These general results followed from our dedicated study on qubit Pauli violation of $n$-cycle noncontextuality inequalities (for $n \geq 4$), and Vorob'ev's theorem. Exploiting the complete characterisation for $n$-cycle scenarios in Ref. [15], we proved that $m$-qubit Pauli operators cannot violate any $n$-cycle non-contextuality inequality when $n > 4$ i.e. whenever the $n$-cycle scenario is a non-Bell like scenario. On the other hand, for $n = 4$, every qubit Pauli realisation of the scenario violates some 4-cycle noncontextuality inequality maximally i.e. it reaches the quantum bound. These results hold independently of $m$, the number of qubits. A crucial spin-off to this analysis was our results on the faithful[1] $m$-qubit Pauli realisations of $n$-cycle scenarios where we proved a no-go theorem that precludes faithful Pauli realisations of cycle scenarios above a certain cycle size, as a function of $m$. Additionally, we provided a few different constructions that realise certain cycle scenarios permissible by our no-go theorem, each with a different linear dependence on $m$. Overall, with this analysis we moved away from the conventional state-independent analyses of qubit Pauli operators [16, 17, 18, 19, 20] to the less studied state-dependent contextuality aspects of these operators. This work is presented in Chapter 4, and available as the pre-print arXiv:2502.03451 [quant-ph] [21].

- **Post-quantum aspects.** We extended the study of the Exclusivity principle, dubbed E-principle, for detecting post-quantumness of correlations to the broader domain of KS contextuality scenarios, advancing the analyses from Ref. [22] which were restricted to Bell scenarios. The E-principle states that if a set of events are pairwise exclusive, then the sum of probabilities of such events should always be less than or equal to 1. Since quantum theory satisfies this principle, any correlation that violates it is necessarily post-quantum. Although equivalent to the no-signalling condition for any bipartite Bell scenario, this equivalence breaks down for two independent copies of the same scenario. Such violation of the E-principle due to multiple copies, known as its *activation effect*,

[1] A faithful realization is one where the corresponding quantum operators have exactly the same compatibility structure as demanded by the scenario

was studied in Ref. [22] for the nonlocal extremal boxes of $(2, m, 2)$[2], $(2, 2, d)$, and $(3, 2, 2)$ Bell scenarios where the authors mapped the problem of exhibiting activation effects to finding certain cliques inside joint exclusivity graphs. We refined this joint exclusivity structure to be an edge-colored exclusivity multigraph. This allowed us to draw a novel connection to Ramsey theory [23, 24], which guarantees the existence of certain monochromatic subgraphs in sufficiently large edge-colored cliques, providing a powerful tool for ruling out E-principle violations. This general link was then applied to study E-principle activation effects for the extremal contextual correlations of $n$-cycle scenarios [15], called $n$-cycle PR boxes. The exploitation of known Ramsey theory results allowed us to generate a plethora of results about the activation effects of $n$-cycle PR boxes. In particular, we showed that for the usual ($n = 4$) PR box, the known E-principle violation of $\frac{5}{4}$ is the maximal achievable using two copies, and that this same upper bound applies to two copies of the KCBS ($n = 5$) PR box. We also proved that $n \geq 6$-cycle PR boxes, unlike the extremal boxes of the aforementioned Bell scenarios, do not exhibit activation effects with two or three copies, providing the first such example of correlations in the literature. Finally, for any number of independent copies $k$, we establish a lower bound on $n$ above which $n$-cycle PR boxes do not exhibit activation effects with $k$ copies. These are cases where the size of violation-producing cliques shoots up exponentially (in the number of independent copies), hence rendering brute-force searches completely impractical, highlighting the strength of our established connection. This work is presented in Chapter 5, and available as the pre-print arXiv:2411.09773 [quant-ph] [25].

**Other Work.** Apart from the three works presented above, there is an ongoing work, in collaboration with Adán Cabello, Emmanuel Zambrini, and Junior R. Gonzales-Ureta, that didn't make it to this thesis. In this work we provide a novel method to compute extremal vertices of the no-signalling polytope of Bell scenarios and explore some applications of these correlations. More details on this work are presented in Chapter 6.

**Outline of the thesis.** This thesis begins with Chapter 2 that provides a background review of the basic concepts underpinning the results in the thesis. We then present our three contributions, described above, in detail in chapters 3-5. We finish with a summary and future research directions in Chapter 6.

[2] The notation $(n, m, d)$ is such that $n$ represents the number of parties, $m$ denotes the number of measurements per party, and $d$ denotes the number of outcomes of each local measurement in the scenario.

# Chapter 2

# Background

In this chapter we do a small background review of the concepts and definitions used throughout the rest of this thesis. Sec. 2.1 presents nuances about KS contextuality and Bell non-locality. To do so, we start with the definition of KS contextuality, Bell nonlocality, exhibited by quantum theory, using canonical examples in Sec. 2.1.1, and then present a theory-independent unification of the two in Sec. 2.1.2. The same section explains the sense in which KS contextuality can be seen as a generalisation of Bell nonlocality. In Sec. 2.1.3, we explain the physical setting in which KS contextuality tests are valid. Next, we move to the formal description of KS scenarios in Sec. 2.2 where we describe the compatibility graph representation of KS scenarios in Sec. 2.2.1. We then explain different classes of correlations for a KS scenario in Sec. 2.2.2. In Sec. 2.2.3, we briefly review concepts in convex polytope theory to then focus on the correlation sets that form convex polytopes in Sec. 2.2.4. Finally, with Sec. 2.3, we end the chapter by first introducing the theorem that ascribes cycle scenarios as the most fundamental KS scenarios in Sec. 2.3.1, and then precisely define cycle scenarios in Sec. 2.3.2. These scenarios are heavily exploited in this thesis.

## 2.1 Conceptual background

One aim of this section is to introduce KS contextuality and Bell non-locality to the reader and then highlight how the latter can be seen as a special case of former in a theory-independent manner. The other aim is to highlight the physical setting in which the KS contextuality tests are logically justified.

### 2.1.1 KS contextuality and Bell non-locality

**KS contextuality.** The KS theorem proves a logical inconsistency between quantum theory and classical theories. More specifically, the theorem [3] shows that a 'classical' assignment of outcomes to the projectors in a set of 117 projectors, associated to a Hilbert space of dimension three, is impossible. The notion of classicality being *non-contextual outcome determinism* (NCOD) where the assigned outcome to every

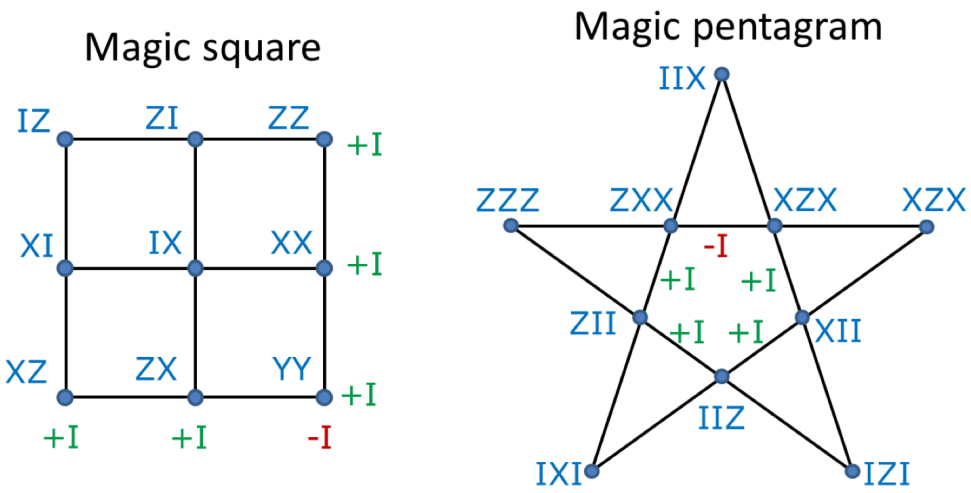



Figure 1: Two KS contextuality witnessing arrangements of qubit Pauli operators (a) Magic square, also called Peres-Mermin square: Operators in each row and column multiply to produce $+I$ (green), except the last column where it is $-I$ (red) (b) Magic Pentagram, also called Mermin star: Operators in each edge of the star multiply to produce $+I$ (green) except one edge it is $-I$ (red).

projector is assumed to pre-exist – even before the measurement process – without it having to depend upon any other compatible measurement. This inconsistency proven by KS theorem could be fixed only when the assigned outcome to at least one observable depends upon the context i.e. what other compatible observables are measured with it, breaking the very way we reason about measurement outcomes in classical physical theories. This violation of classicality is called KS contextuality. A much simplified proof of the same inconsistency was later shown using a set of 18 projectors in the four dimensional Hilbert Space [26]. Further simpler, and hence more popular, proofs utilising qubit Pauli operators also capture the same inconsistency e.g. Peres-Mermin magic square and Mermin star [18, 19, 20].

To illustrate KS contextuality, let's look at the famous Peres-Mermin square. It is composed of nine two qubit Pauli operators arranged in a 3x3 array as shown in Fig. 1(a). Each row as well as a column contains mutually commuting operators. For each given row multiplying the operators gives $I$. The same is true for operators in each column except the last one where operators multiply to give $-I$, highlighted with red color in Fig. 1(a). Quantum theory dictates that in any joint measurement of three mutually compatible Pauli operators $A, B, C$ where $ABC = I$ (or $ABC = -I$), if the outcome for $A$ is $a$, for $B$ is $b$, and for $C$ is $c$, then it must be that $abc = 1$ (or $abc = -1$). Let us now consider a classical model for these quantum measurements. This model simultaneously assigns outcomes to all Pauli measurements in the scenario in accordance with the NCOD assumption. The assigned outcomes must be in $\{+1, -1\}$ since these are the possible outcomes of qubit Pauli measurements. More precisely, for such a classical model

to be run-by-run consistent with the joint measurement outcomes of the operators $A$, $B$, and $C$, it must be that the assigned outcomes to these observables are the corresponding eigenvalues. So, if $v(A) = a$, $v(B) = b$, and $v(C) = c$, where $v(A)$ is the assignment to measurement $A$ in a given run, it must hold that $v(A)v(B)v(C) = 1$ (or $v(A)v(B)v(C) = -1$). This leads to the following constraints on the classical model:

$$v(IZ)v(ZI)v(ZZ) = 1 \qquad v(IZ)v(XI)v(XZ) = 1$$
$$v(XI)v(IX)v(XX) = 1 \qquad v(ZI)v(IX)v(ZX) = 1$$
$$v(XZ)v(ZX)v(YY) = 1 \qquad v(ZZ)v(XX)v(YY) = -1$$

Clearly, each outcome assigned to a given operator appears twice on the left hand side of the equations above. Therefore, multiplying all equations above leads to the contradiction $1 = -1$, suggesting that the NCOD model is inconsistent with the algebraic structure underpinning quantum theory. A similar inconsistency can be shown for a set of ten three-qubit Pauli operators, see Fig. 1(b), famously called a Mermin star.

**Bell non-locality.** Bell's theorem provides another proof of non-classicality of quantum theory where the notion of classicality is called *local causality*. This theorem concerns itself with space-like separated parties, each receiving a system (from a common source) and making measurement on it in each run of the experiment. The assumption of local causality imposes that in each such run a common cause, commonly denoted by $\lambda$, *solely* determines the probability of outcome of each space-like separated measurement, also called a local measurement. More imaginatively, one can think of $\lambda$ as an entity that captures all the past interactions between the systems coming out from the source in a given run. In other words, in every run of the experiment, the probability of outcome of each local measurement can depend on the common cause $\lambda$ only, and not on any other space-like separated measurement. Since the parties are making space-like separated measurements the only way these measurement outputs can have correlations is via such common causes $\lambda$. This is the best we can imagine, with our classical intuition, so that the correlations observed between space-like separated parties are consistent with the principles of special relativity. But surprisingly, quantum theory violates this intuitive picture of such an experiment, i.e. there are Bell scenarios where correlations predicted by quantum theory do not admit such a locally causal description.

The simplest scenario, depicting such a quantum-classical gap, is the CHSH scenario [27] which involves two space-like separated parties each with two dichotomic measurements. In the CHSH scenario, the two parties are conventionally called Alice and Bob and their respective measurement sets are labelled as $\{A_1, A_2\}$ and $\{B_1, B_2\}$ while the outcome of the measurement of each party is labelled as $a, b$ re-

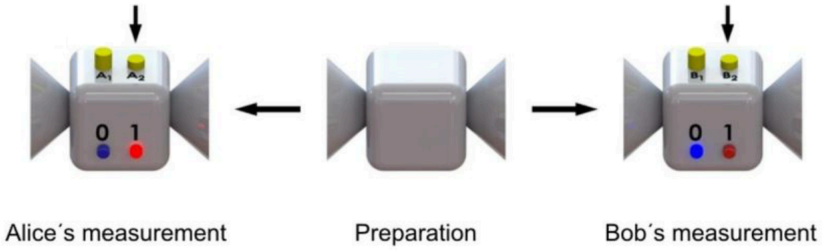



Figure 2: CHSH scenario: A preparation source sending particles to two space-like separated parties Alice and Bob who make local measurements on the received systems.

spectively, where $a, b \in \{+1, -1\}$. The above wordy description of the assumption of local causality can be faithfully and succinctly captured mathematically by the following factorisation of the joint probability of Alice's and Bob's measurements, given the common cause $\lambda$:

$$p(a, b|A_i, B_j, \lambda) = p(a|A_i, \lambda)p(b|B_j, \lambda) \; \forall \; i, j \in \{1, 2\} \tag{2.1}$$

where $p(a|A_i, \lambda), p(b|B_j, \lambda) \in [0, 1]$. The observed joint probability distribution is obtained by summing over all possible values of the common cause:

$$p(a, b|A_i, B_j) = \sum_{\lambda} p(\lambda)p(a|A_i, \lambda)p(b|B_j, \lambda) \tag{2.2}$$

As will be later explained in Sec. 2.2.4, for any arbitrary Bell scenario, the set of correlations consistent with local causality forms a convex polytope. The set of quantum correlations is known to strictly contain this classical polytope inside it. Therefore, any quantum correlation outside this classical polytope exhibits Bell nonlocality. Furthermore, the description of local causality doesn't necessarily require that $p(a|A_i, \lambda)$ , $p(b|B_j, \lambda) \in \{0, 1\}$, i.e. the common cause doesn't need to deterministically determine the outcomes of local measurements. Hence, it could seem that the set of correlations consistent with a probabilistic (i.e., factorisable) locally causal model would be strictly bigger than a deterministic locally causal model. Fine's theorem [28, 29] nullifies this expectation. The theorem states that any probabilistic locally causal model can be refined to a deterministic locally causal model. In other words, for a given Bell scenario, the set of classical correlations is the same irrespective of whether one chooses a deterministic or a probabilistic locally causal model.

We now come back to the case of CHSH scenario to demonstrate violation of local causality by quantum theory. A canonical way to capture this gap is by the following quantity:

$$S \equiv \langle A_1 B_1 \rangle + \langle A_1 B_2 \rangle + \langle A_2 B_1 \rangle - \langle A_2 B_2 \rangle \tag{2.3}$$

where $\langle A_i B_j \rangle = \sum_{ab} abP(ab|A_i B_j)$. For a locally causal model, we denote this quantity as $S_{LC} = \sum_\lambda p(\lambda) S_\lambda$, where:

$$S_\lambda = \langle A_1 B_1 \rangle_\lambda + \langle A_1 B_2 \rangle_\lambda + \langle A_2 B_1 \rangle_\lambda - \langle A_2 B_2 \rangle_\lambda \tag{2.4}$$

Due to Eqn. (2.1), $\langle A_i B_j \rangle_\lambda = \langle A_i \rangle_\lambda \langle B_j \rangle_\lambda$. This allows us to write $S_\lambda$ as follows:

$$S_\lambda = \langle A_1 \rangle_\lambda (\langle B_1 \rangle_\lambda + \langle B_2 \rangle_\lambda) + \langle A_2 \rangle_\lambda (\langle B_1 \rangle_\lambda - \langle B_2 \rangle_\lambda) \tag{2.5}$$

Without loss of generality, thanks to Fine's theorem, we can further impose the assumption of determinism. This means that $\langle A_i \rangle_\lambda, \langle B_j \rangle_\lambda \in \{+1, -1\}\ \forall i, j \in \{0, 1\}$ which further implies that either $\langle B_1 \rangle_\lambda = \langle B_2 \rangle_\lambda$ or $\langle B_1 \rangle_\lambda = -\langle B_2 \rangle_\lambda$. This further means that $S_\lambda = \pm 2$. Inputting this value in $S_{LC}$ means that:

$$-2 \leq S_{LC} \leq 2 \tag{2.6}$$

In words, a locally causal model predicts the quantity $S$ to remain between $-2$ and $2$.

Quantum theory, on the other hand, violates this bound using the two-qubit entangled state $|\psi\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}$ shared between Alice and Bob with the following local quantum measurements: $A_1 = {\sigma_x}^1$, $A_2 = \sigma_y$ and $B_1 = \frac{-\sigma_x - \sigma_y}{\sqrt{2}}$, $B_2 = \frac{\sigma_y - \sigma_x}{\sqrt{2}}$. With these measurements $\langle S \rangle$ attains the maximum quantum value $S_{QM} = 2\sqrt{2} > S_{LC}$. This marks the non-classicality of quantum theory via Bell non-locality.

We remind the reader that in any claim of Bell non-locality there is a crucial, usually understated, assumption in deriving the classical bound. This assumption entails that the measurement choices of different parties, in each round, are independent of the common cause $\lambda$ and vice versa. It is called measurement independence and is expressed, for the CHSH case, as $p(A, B|\lambda) = p(A, B)$, or equivalently via Bayes' theorem, called parameter independence and expressed as $p(\lambda|A, B) = p(\lambda)$ where variables $A$ and $B$ vary over the choices of measurements of Alice and Bob respectively. Furthermore, the equivalence between the two types of locally causal models proven by Fine's theorem allows us to see local causality and non-contextual outcome determinism in a unified way as we now explain.

### 2.1.2 Bell non-locality is a special case of KS contextuality

As will become clear later, this thesis contains a lot of results that extend some already known results within Bell non-locality literature to the more general setting of KS contextuality. Before engaging with those results, it is imperative to understand the sense in which KS contextuality captures Bell non-locality

[1] $\sigma_x$ denotes the $X$-Pauli measurement

as a special case. A good starting point for this is to examine the notions of classicality concerning KS contextuality and Bell non-locality.

The notion of classicality introduced by Kochen-Specker, as explained in Sec. 2.1.1, is called non-contextual outcome determinism (NCOD) whereby, in each run of the experiment, outcomes are assigned to measurements irrespective of the context in which any measurement is performed. Bell's local causality, on the other hand, involves space-like separated measurements whereby, in each run of the experiment, probabilities of outcomes of each measurement are determined solely by the common cause and not by any other space-like separated measurement. To see how the former absorbs the latter, let us focus on the term *context* that comes naturally in discussions of KS contextuality, as described in Sec. 2.1.1.

**Meaning of context.** Given a set of measurements, say $\mathcal{X}$, and a set $\mathcal{C}$ of its compatible subsets, a context of a given measurement $A \in \mathcal{X}$ denotes what other measurements in $\mathcal{X}$ can be performed with it i.e. each set in $\mathcal{C}$ containing $A$ defines a context for $A$. In any classical theory, all possible measurements can be simultaneously performed on a system, hence, a reader privy only to such theories, may not find the notion of context very meaningful at the moment. To them, set $\mathcal{C}$ need not even be separately defined because all measurements in $\mathcal{X}$ must be compatible. To such a reader's surprise, we highlight that there are more general theories than quantum theory that don't share this feature with classical theories. It is in these theories where the notion of context becomes meaningful (non-trivial), e.g. quantum theory allows for a set of three measurements, denoted by Hermitian operators $A, B, C$, such that $[A, B] = [A, C] = 0$ while $[B, C] \neq 0$. Three such measurements are impossible to be performed simultaneously although $A$ can be performed with either $B$ or $C$, illustrating two contexts for it, $\{A, B\}$ and $\{A, C\}$. It is for such general theories that each compatible subset of $\mathcal{X}$ in $\mathcal{C}$ is called a context, without referring to a specific measurement in $\mathcal{X}$, because each such subset denotes a context for some measurement in $\mathcal{X}$.

**Unification.** In case of Bell scenarios, therefore, a context would consist of a set of pairwise space-like separated measurements. We learnt, in Sec. 2.1.1, that Fine's theorem allows the common cause $\lambda$ (in each run) to determine the probabilities of space-like separated measurements deterministically. But this exactly amounts to assigning outcomes to space-like separated measurements without caring about the contexts. Thus, the assumption of NCOD includes the assumption of local causality as a special case. Violation of local causality can therefore be seen as a violation of NCOD, legitimising it as an instance of KS contextuality. This allows us to use the phrase 'KS contextuality scenarios' more generally, with Bell scenarios as a special case when each context contains space-like separated measurements. This unification further implies that a lot of our general results apply de facto to Bell scenarios as a special case.

**KS contextuality redefined.** With this unification we can define what KS contextuality means in a bit more formal, but still intuitive, fashion. We do so by defining its negation: in each run of the experiment, a variable $\lambda$ determines the values of all the measurements of a given scenario non-contextually, which in case of Bell scenarios becomes the common cause. After all runs of the experiment, an NCOD model defines a probability distribution $p(\lambda)$. This distribution is equivalent to a joint probability distribution (JPD) over the space of joint outcomes of all measurements in the scenario. Now, if a given statistics corresponding to the contexts of the given scenario can be explained by this JPD via marginalisation, it marks a proof of non-contextuality of the same statistics whereas if the given statistics cannot be explained by any such JPD via marginalisation, it is KS contextual.

### 2.1.3 Physical setting of KS contextuality scenarios

Until now we have seen examples of how KS contextuality and Bell non-locality are exhibited by quantum theory, and that the latter can in fact be seen as a special case of the former. We also saw how the physical setting of a Bell scenario looks like, but we haven't yet discussed the physical setting for non-Bell-like KS contextuality scenarios, e.g. the Peres-Mermin square as introduced in Sec. 2.1.1. We remark to the reader that by 'physical setting', we mean the setting in which it is justified to claim non-classicality via KS contextuality. This will become more clear now.

**The setting.** For non-Bell-like KS scenarios, within every run of the experiment, a source prepares a system on which a set of compatible time-like separated measurements, i.e. the ones belonging to a context, are sequentially performed. But unlike a Bell scenario where space-like separation guarantees no-signalling between any two subsets of parties, compatible time-like separated measurements on a given system can signal/disturb[2] each other. This is a problematic feature to have. Problematic because if the statistics of any measurement depends upon the ones done prior to it, in other words it depends upon the context, then one loses the warrant to impose the assumption of NCOD in constructing a classical model for that experimental setting. This is because in an NCOD (classical) model, variable $\lambda$, associated to the preparation in the given run, is supposed to *solely* determine the outcomes (and hence the overall statistics) of each of the measurements in the sequence. Thus, disturbance at the operational level invalidates comparison with an NCOD assuming model and hence it is an unsuitable situation to claim KS contextuality. This invites introduction to the idea of ideal measurements because no-disturbance, in time-like separated sequential measurements, can be guaranteed if the compatible measurements are ideal, the meaning of which we now explain.

[2] Disturbance means that from the statistics of a measurement in the sequence, one can detect the presence of some measurement performed before it.

**Ideal measurements.** A measurement is ideal[3] (or sharp), as defined in [30, 31, 32], if (i) it yields the same outcome when performed consecutive times, (ii) it only disturbs measurements that are incompatible with it, (iii) all its coarse grainings have realisations satisfying (i) and (ii). In quantum theory, an ideal measurement is a projective measurement, denoted by a Hermitian operator, or equivalently by a set of projectors summing to identity, also dubbed as a projection valued measure (PVM). More generally, arbitrary measurements in quantum theory are represented by positive operator valued measures (POVMs), which includes ideal as well as non-ideal measurements.

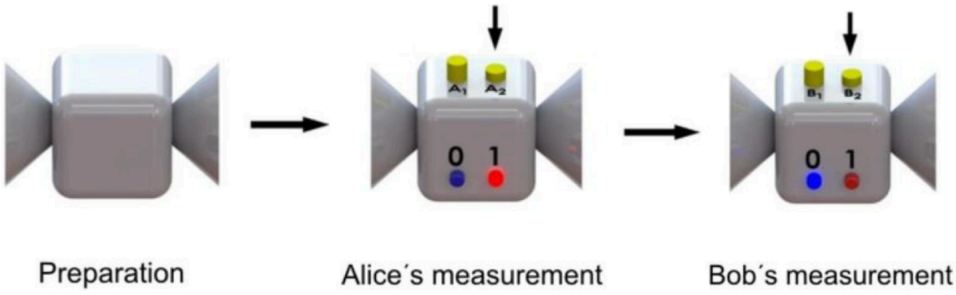



Figure 3: Sequential CHSH scenario: Alice and Bob making time-like separated compatible ideal measurements

**Categorising KS scenarios (based on implementation).** In case of Bell scenarios, space-like separation is enough to guarantee no-disturbance (no-signalling), irrespective of whether the measurements are ideal or not. Overall, the assumption of NCOD justifiably applies to KS contextuality scenarios which are either Bell scenarios or ones that consist of sequential compatible ideal measurements. Note that one can take the contexts in a Bell scenario and implement (or simulate) them via sequential compatible ideal measurements as well, but because NCOD applies to both situations alike, such a model predicts the same set of (classical) correlations in the two implementations. This lets us divide KS scenarios into two types, based on their implementation (also summed up by Fig. 4):

1. non-sequential KS scenarios: standard Bell scenarios with space-like separated measurements.
2. sequential KS scenarios: time-like separated compatible ideal measurements. This category can be further divided into two sub-types:
   (a) Bell-like sequential KS scenarios, see Fig. 3, depicting a sequential CHSH scenario (contrast it with the implementation in Fig. 2).
   (b) non Bell-like KS scenarios, e.g. the famous KCBS scenario.

[3] Note that this is an operational i.e. a theory independent notion of ideal measurements.

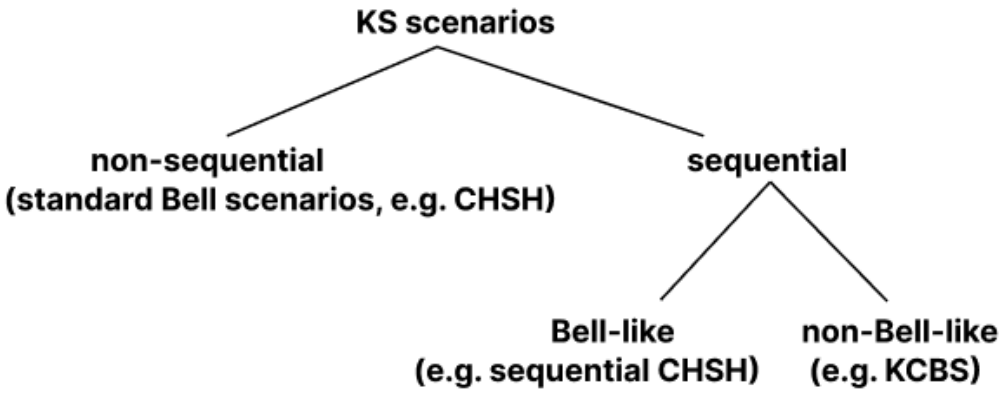



Figure 4: Categorising KS scenarios based on implementations

**Quantum realisation.** In a quantum-mechanical realisation of a KS scenario, the measurements are taken to be quantum observables. Since for sequential KS scenarios, these measurements need to be ideal, a quantum mechanical realisation corresponds to one with projective measurements (PVMs). Compatibility now corresponds to commutativity of the projective measurements. For Bell scenarios, on the other hand, one does not need to impose that the measurements must be ideal, and so in quantum theory they could be represented by non-ideal (unsharp) measurements (POVMs). But, thanks to Naimark's dilation theorem [33, 34], all quantum-mechanical correlations in Bell scenarios are also achievable by ideal measurements which are space-like separated. Therefore, the set of quantum correlations for a Bell scenario is contained in the set of quantum correlations for its corresponding sequential Bell-like KS contextuality scenario, being identical when restricting to finite-dimensional quantum systems [35]. This observation justifies regarding quantum nonlocality as a special case of quantum contextuality [36, 37, 38] and has proved very useful in understanding the principles that shape quantum Bell nonlocal correlations [39, 40].

## 2.2 Mathematical background

Having provided the definition and intuition for KS contextuality to the reader, we now move onto the formal description of these scenarios and the associated correlation sets and their geometries.

### 2.2.1 Representing KS contextuality scenarios

As indicated in the previous section, claiming non-classicality requires a precise description of the setting in which such an analysis is made. Moreover, the description of this setting must not depend on the very theories it is supposed to compare/test. A *scenario* describes such a theory-independent setting.

Formally, a *KS contextuality scenario*, **S**, henceforth called scenario for brevity, is given by a finite set $\mathcal{X}_{\mathbf{S}}$ of measurements, a symmetric reflexive relation on $\mathcal{X}_{\mathbf{S}}$ indicating *compatibility*, and a finite set $\mathcal{O}_{\mathbf{S},A}$ of outcomes for each measurement $A \in \mathcal{X}_{\mathbf{S}}$. Graphs turn out to be very handy to depict KS scenarios.

**Compatibility graph.** We can further think of the set $\mathcal{X}_{\mathbf{S}}$ equipped with the compatibility relation as a graph, known as the *compatibility graph* of scenario **S**. A *measurement context* is a clique in the compatibility graph, i.e. a set of pairwise-compatible measurements. It denotes the set of measurements whose joint statistics are operationally accessible. We write $\mathcal{C}_{\mathbf{S}} = \{C_1, \ldots, C_n\}$ for the set of maximal contexts, fixing an enumeration. By maximal we mean that for all $C_i$ and $C_j$ in $\mathcal{C}_{\mathbf{S}}$, $C_i \subseteq C_j$ implies $C_i = C_j$. We omit the subscript **S** whenever discussing a single scenario. Note that one might imagine a more general situation where set $\mathcal{X}_{\mathbf{S}}$ equipped with the compatibility relation is represented by a compatibility hypergraph, instead of a graph, where contexts are represented by its hyperedges, but in this thesis we restrict ourselves to compatibility graphs which cover all KS scenarios where pairwise compatibility of a set of measurements always implies their joint measurability. This includes all Bell scenarios as well as those KS scenarios that are realizable by theories, like quantum theory, where ideal measurements follow the compatibility rule mentioned above.

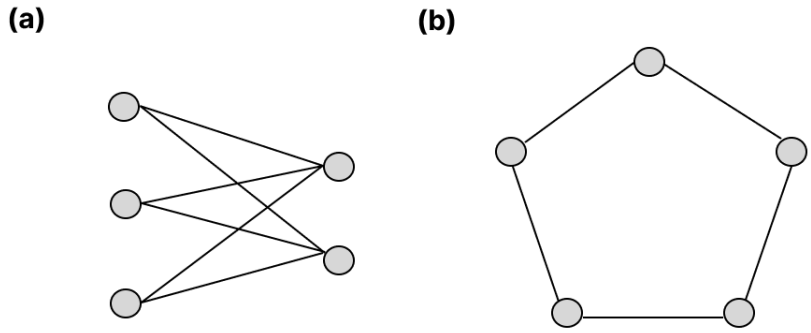



Figure 5: Compatibility graphs (a) bi-partite graph representing a Bell scenario (b) Pentagon, a non $n$-partite graph, representing KCBS scenario, the simplest non-trivial non-Bell-like scenario.

In a Bell scenario, measurements are assigned to parties where any two measurements are compatible if and only if they belong to different parties, the compatibility graph of these scenarios undertake the special form of complete $n$-partite graphs where $n$ denotes the number of parties in the scenario. An $n$-partite graph is one where the set of nodes can be divided into $n$ disjoint sets (called partitions) such that no vertices within the same partition are adjacent. So, the only possible edges between nodes are across different partitions. A complete $n$-partite graph is an $n$-partite graph where nodes within each partition connect with all the nodes in other partitions, see Fig. 5(a).

### 2.2.2 Correlations

Now that we know the operational conditions (physical setting) to test for KS contextuality as well the way to represent KS scenarios, we focus on different class of statistics that arise depending on the theory dictating the experiment. A *correlation* (a.k.a. empirical model or behaviour) on a scenario **S** collects specific outcome statistics for that scenario. It is a family comprising a probability distribution on the joint outcomes for each maximal context in $\mathcal{C}$. For each $C_i$, we fix an enumeration of the set of joint outcomes $\mathcal{O}_{C_i} = \prod_{A \in C_i} \mathcal{O}_A = \{s_i^1, s_i^2, \ldots, s_i^{m_i}\}$. A correlation can then be represented as a vector

$$p = [p(s_1^1), \ldots, p(s_1^{m_1}), \ldots, p(s_n^1), \ldots, p(s_n^{m_n})],$$

where the component $p(s_i^j)$, often written more explicitly as $p(s_i^j \mid C_i)$, represents the probability of observing the $j$-th joint outcome $s_i^j$ upon jointly performing the measurements of the $i$-th context $C_i$. We have $p \in \mathbb{R}^d$, where $d = \sum_{C_i \in \mathcal{C}} |\mathcal{O}_{C_i}|$ is the dimension of the vector space wherein all possible correlations on the given scenario live. The fact that they determine a probability distribution $p_i$ over each maximal context $C_i$ means that the components must all be non-negative, i.e. $p(s_i^j) \geq 0$ for all $i, j$, and in addition satisfy normalisation for each context $C_i$,

$$\sum_{j=1}^{m_i} p(s_i^j) = 1.$$

Moreover, we assume that the marginal probability distributions on the outcomes of each observable (or set of observables) is independent of the context, i.e. of what other measurements are performed together. This condition is called *no-disturbance* – or, in Bell scenarios, *no-signalling* – and it prevents using the statistics for achieving communication between observers that are performing compatible measurements. Correlations obtainable in quantum mechanics always satisfy this condition. More precisely, the no-disturbance condition can be stated as follows: for any pair of maximal contexts $C_i$ and $C_{i'}$, the corresponding distributions $p_i$ and $p_{i'}$ determine the same marginal distribution over the intersection $C_i \cap C_{i'}$, i.e. for each (partial) joint outcome $t \in \mathcal{O}_{C_i \cap C_{i'}}$, it holds that

$$\sum_{j \in E_i(t)} p(s_i^j) \;=\; \sum_{j' \in E_{i'}(t)} p(s_{i'}^{j'}), \tag{2.7}$$

where the index $j$ on the left runs over the set $E_i(t) = \{j \mid s_i^j|_{C_i \cap C_{i'}} = t\}$ indexing those joint outcomes in $\mathcal{O}_{C_i}$ that extend $t$, and analogously for $j'$. One may think of either side of Eq. (2.7) as determining the probability $p(t)$ of obtaining outcome $t$ in the partial context $C_i \cap C_{i'}$; no-disturbance guarantees it is uniquely well defined. Notice that the no-disturbance assumption justifies considering only the maximal contexts in $\mathcal{C}$.

**Heirarchy of correlations.** Let us now see the some classes of correlations that are studied within KS contextuality. For a given scenario, let $L_{ND}$ denote the subset of $\mathbb{R}^d$ consisting of all correlation vectors that satisfy no-disturbance. Similarly, write $L_Q$ and $L_{NC}$ for the sets of quantum correlations and of noncontextual correlations (to be defined more precisely in Sec. 2.2.4), respectively. There is a sequence of inclusions

$$L_{NC} \subseteq L_Q \subseteq L_{ND} \tag{2.8}$$

and moreover these inclusions are strict for any non-trivial scenario. However, it holds that the affine dimension of all three sets is the same,

$$\dim(L_{NC}) = \dim(L_Q) = \dim(L_{ND}).$$

These correlations define different geometries inside the affine space of correlations. $L_{NC}$ and $L_{ND}$ are convex polytopes whereas $L_Q$ is a convex set that is not a polytope. To grasp the results in this thesis, the reader doesn't require more knowledge of the quantum set than what's captured in Eqn. 2.8, but a significant amount of work in this thesis expects basic understanding of convex polytopes, we now move to providing a basic review of these important geometric objects.

### 2.2.3 Convex polytopes

In this section we recall some basic definitions of polytope theory that will be useful later on. For brevity, in the context of this thesis, the term polytope always means convex polytope. For more details, we refer the reader to [41]. Before getting to the definition of convex polytopes, we introduce some preliminaries.

An *affine combination* of a set of points $p_1, \ldots, p_k \in \mathbb{R}^d$ is a linear combination whose coefficients add up to $1$, i.e. $\sum_{i=1}^k \alpha_i p_i$ for some $\alpha_i \in \mathbb{R}$ with $\sum_i \alpha_i = 1$. A *convex combination* is an affine combination with the additional condition that all coefficients be non-negative, i.e. $\alpha_i \geq 0$. The *affine span* (*convex hull*) of a set of points is the set of all their affine (convex) combinations.

A set of points $p_0, \ldots, p_k \in \mathbb{R}^d$ is *affinely independent* when $\sum_{i=0}^k \alpha_i p_i = 0$ with $\alpha_i \in \mathbb{R}$ satisfying $\sum_i \alpha_i = 0$ implies that $\alpha_i = 0$ for all $i$. This is equivalent to the vectors $p_1 - p_0, \ldots, p_k - p_0$ being linearly independent. It is also equivalent to saying that no element of the set is an affine combination of the others.

A set of $k+1$ affinely independent points spans a set of dimension $k$. More generally, the affine span of a set of points has dimension $k$ if and only if the maximum number of affinely independent points in that set is $k+1$. The affine span of any such $k+1$ points equals the affine span of the whole set, and those $k+1$ points are called an *affine basis* for the set. The dimension $\dim(S)$ of an arbitrary subset $S$

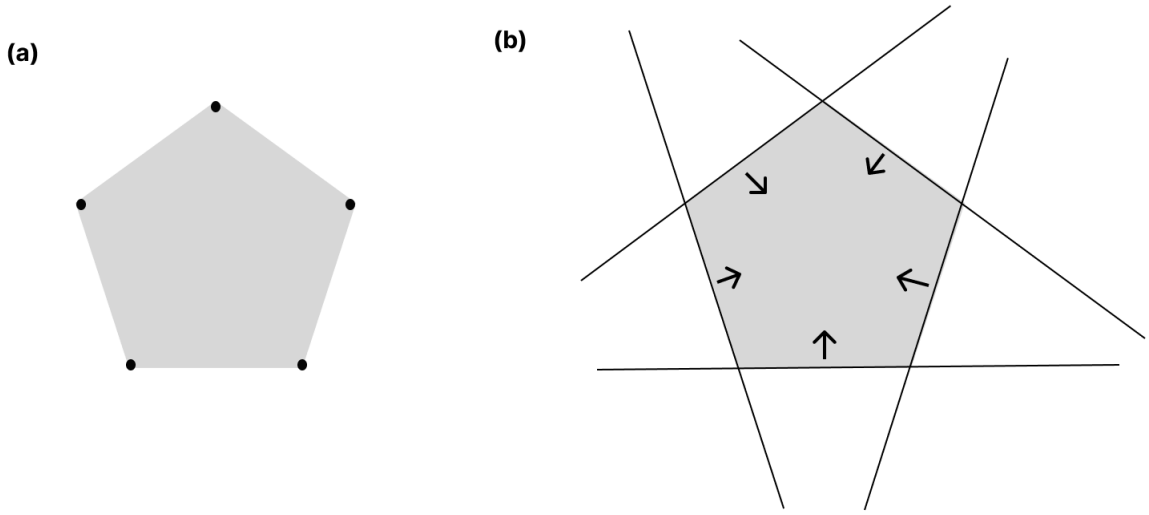



Figure 6: Two equivalent representations of the same polytope (a) V-representation where knowing the vertices (black dots) is enough to define any point, via convex combinations, of the polytope, (b) minimal H-representation where an intersection of the facet-defining hyperplanes exactly captures the polytope.

of $\mathbb{R}^d$ is defined to be the dimension of its affine span.

A convex polytope is a subset $P \subseteq \mathbb{R}^d$ that is a convex hull of a finite set of points, its *vertices*. This set of vertices constitutes the *V-representation* of the polytope, see Fig. 2.2.3 (a). A *linear inequality* on $\mathbb{R}^d$ is a predicate $I(p) \equiv b \cdot p \geq \beta$ on a variable vector $p \in \mathbb{R}^d$, determined by a vector $b \in \mathbb{R}^d$ of coefficients and a bound $\beta \in \mathbb{R}$. It defines a *half-space* of $\mathbb{R}^d$: $\{p \in \mathbb{R}^d \mid I(p)\} = \{p \in \mathbb{R}^d \mid b \cdot p \geq \beta\}$. A linear inequality is a *valid inequality* for a polytope $P \subseteq \mathbb{R}^d$ whenever $P$ lies entirely within the half-space described by it, that is, when

$$P \cap \{p \in \mathbb{R}^d \mid b \cdot p \geq \beta\} = P.$$

Since any point in $P$ can be written as a convex combination of its vertices, it is sufficient to test an inequality on the vertices to check for its validity.

Given a valid inequality $b \cdot x \geq \beta$ for a polytope $P$, the set $F = \{p \in P \mid b \cdot p = \beta\}$ of points of $P$ that saturate the inequality is called the *face of $P$ supported by the inequality*. Each face of a convex polytope $P$ is a convex polytope itself. A face $F$ is called *proper* when it is neither the empty face nor the whole polytope itself, i.e. when $F \neq \varnothing$ and $F \neq P$. Clearly, for a proper face $F$ of $P$, $0 \leq \dim(F) < \dim(P)$. The minimal proper faces, of dimension $0$, are the vertices of $P$. A maximal proper face, of dimension $\dim(P) - 1$, is called a *facet* of $P$. According to the Minkowski–Weyl theorem [42, 43, 41], a convex polytope can equivalently be described as an intersection of finitely many half-spaces. This defines an

H-representation of the polytope,

$$P = \bigcap_i \{p \in \mathbb{R}^d \mid b^{(i)} \cdot p \geq \beta^{(i)}\}.$$

The set of inequalities supporting the facets of $P$ gives a minimal complete description of the polytope as an H-representation, Fig. 2.2.3 (b). In fact, any valid inequality of $P$ can be derived from the facet-defining inequalities.

By the remarks above about the dimension of any subset of $\mathbb{R}^d$, the maximum number of affinely independent points in a polytope $P$ is $\dim(P) + 1$. One can always draw such $\dim(P) + 1$ affinely independent points from the vertices of $P$. Similarly, any face $F$ of $P$ contains $\dim(F) + 1$ affinely independent vertices of $P$, which of course saturate the inequality supporting $F$. In particular, a facet contains a set of $\dim(P)$ affinely independent vertices of $P$. This means that one can always choose an affine basis of $P$ comprising $\dim(P) + 1$ affinely independent vertices in such a way that $\dim(P)$ of those vertices belong to a chosen facet $F$, and the remaining vertex can be chosen arbitrarily among the vertices not belonging to $F$. This finishes the relevant recapitulation of convex polytope theory necessary to grasp the main results in this thesis.

### 2.2.4 No-disturbance and the noncontextual polytopes

It is imperative to understand the no-disturbance and the noncontextual polytopes to get a geometric understanding of proofs of KS contextuality and different correlation sets. For any given scenario, the no-disturbance polytope ($L_{ND}$) constitutes the biggest possible set of correlations that are eligible to be tested for KS contextuality, while the noncontextual polytope ($L_{NC}$) is the set of all non-contextual correlations inside $L_{ND}$.

**No-disturbance polytope.** Due to the no-disturbance condition defined in Eqn. (2.7) along with the positivity and normalization conditions for probabilities, one naturally attains the H-representation of the no-disturbance polytope. This is easy to see, just the positivity and normalisation conditions on probabilities lead to a simplex which contains correlations that may not all satisfy the no-disturbance conditions, but since each no-disturbance condition (a linear constraint) maps to a hyperplane that intersects this simplex, we obtain a polytope all of whose points satisfy the no-disturbance conditions.

**Non-contextual polytope.** The non-contextual polytope is naturally defined in terms of its V-representation with the non-contextual value assignments as its vertices. A *noncontextual deterministic assignment* is a joint assignment of outcomes to all the measurements in a scenario, $\lambda \in \mathcal{O}_{\mathcal{X}} = \prod_{A \in \mathcal{X}} \mathcal{O}_A$, which

one may think of as a *hidden variable*. It determines a deterministic correlation

$$v_\lambda = [v_\lambda(s_1^1), \ldots, v_\lambda(s_1^{m_1}), \ldots, v_\lambda(s_n^1), \ldots, v_\lambda(s_n^{m_n})],$$

where $v_\lambda(s_i^j) \in \{0, 1\}$ indicates whether $\lambda$ assigns the joint outcome $s_i^j$ to the context $C_i$,

$$v_\lambda(s_i^j) = \begin{cases} 1 & \text{if } \lambda|_{C_i} = s_i^j \\ 0 & \text{otherwise.} \end{cases}$$

Note that over each context $C_i$ such a vector $v_\lambda$ has precisely one component valued $1$ and the remaining $0$. A correlation $p$ is said to be *noncontextual (NC)* if it can be described as a convex combination of such deterministic correlations, i.e. if

$$p = \sum_{\lambda \in \mathcal{O}_\mathcal{X}} \alpha_\lambda v_\lambda \tag{2.9}$$

for some $\alpha_\lambda \geq 0$ with $\sum_\lambda \alpha_\lambda = 1$. This means that each component satisfies

$$p(s_i^j) = \sum_\lambda \alpha_\lambda v_\lambda(s_i^j).$$

One may think of the coefficients $\alpha_\lambda$ as defining a probability distribution on the hidden-variable space of global assignments. Eq. (2.9) means that the set of NC correlations is the convex hull of a finite number of vectors $\{v_\lambda\}_{\lambda \in \mathcal{O}_\mathcal{X}}$.

**Complementary representations.** Although the above mentioned representations of the two polytopes follow naturally from the corresponding definitions, it is the complementary representation, for each polytope, that is more useful in so far as obtaining or identifying KS contextual correlations is concerned with further foundational and computational benefits, as we now explain:

(a) The V-representation of the no-disturbance polytope directly provides KS contextual correlations, along with non-contextual ones. These correlations are also commonly called extremal boxes. The famous PR-box is one such example. The contextual vertices have helped us immensely in learning about quantum theory, e.g. it is known that these correlations, at least for Bell scenarios, can't be realised by quantum theory [44], therefore, any principle proposed to single out quantum theory must necessarily detect these correlations as post-quantum [22, 45]. More generally, such post-quantum correlations allow for a thinking space to ask interesting questions like why quantum theory is not more non-local (contextual) [46], what are the physical principles that underlie quantum physics [47, 48, 49, 50, 51].

(b) The H-representation of the noncontextual polytope, on the other hand, provides noncontextuality inequalities for that scenario, marking a boundary between the noncontextual and contextual

regions. The violation of any such inequality [52, 53, 54] by a correlation witnesses contextuality. A minimal H-representation is given by the facet-defining inequalities. Although for restricted classes of polytopes, such as simplicial polytopes, there are known polynomial-time algorithms for facet enumeration (the problem of computing the H-representation given the V-representation) [55, 56, 57], characterising all the facets of the noncontextuality polytope for an arbitrary scenario is a notoriously hard problem to solve (NP-complete [13]). Identifying whether a given correlation is contextual amounts to detecting its membership inside the non-contextual polytope. In general, detecting membership is more efficient given the H-representation than given the V-representation. Therefore, to test contextuality for a given correlation it is advantageous to know the H-representation of the non-contextual polytope.

## 2.3 Fundamental KS scenarios

Cycle scenarios make consistent appearances in this thesis. In this section we first explain why they are the most fundamental KS scenarios and then define them precisely.

### 2.3.1 Vorob'ev's theorem

Vorob'ev's theorem attributes foundational stature to $n$-cycle scenarios which we define more precisely in Sec. 2.3.2. This theorem captures compatibility structures that can never produce contextual correlations. Originally introduced in the context of game theory [58, 59], it recently gained prominence in the KS contextuality literature [60, 61, 62, 63]. Vorob'ev's theorem states that if the compatibility graph of a scenario is a chordal graph (every induced cycle has size 3), then all statistics corresponding to such a scenario are non-contextual. In other words, corresponding to any observed statistics for such scenarios, there exists a JPD over all outcomes of all measurements that reproduces these statistics by marginalisation. Therefore, for each such scenario $L_{NC} = L_{ND}$. Eqn. 2.8 further implies that $L_{NC} = L_{ND} = L_Q$ holds for each such scenario, marking them as trivial KS scenarios. An alternative formulation of Vorob'ev's theorem is that an arbitrary KS scenario can witness contextuality if and only if the scenario has at least one $n$-cycle scenario ($n \geq 4$) as its sub-scenario.

To illustrate, consider a simple scenario whose compatibility graph is a path graph of size three ($H_3$), as shown in Fig. 7. Any observed statistics for this scenario corresponds to contexts $\{A, B\}$ and $\{B, C\}$, and can be denoted as $p_{AB}$, $p_{BC}$. The JPD $p_{ABC} = \frac{p_{AB}p_{BC}}{p_B}$ (where we define $p_{ABC} = 0$, whenever $p_B = 0$), due to the no-disturbance condition, always marginalises to the observed statistics $p_{AB}$, $p_{BC}$ of

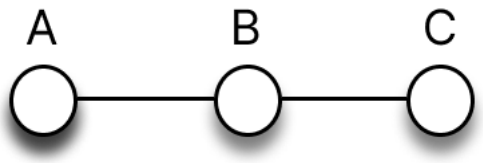



Figure 7: A path graph depicting compatibility relations between three measurements $A, B, C$.

the scenario. In a similar fashion, one can always come up with such a JPDs for more complicated chordal graphs that reproduce any statistics of the corresponding scenarios via marginalisation. This theorem's contrapositive then tells us that non-chordality (existence of induced cycles of size four or more) of the compatibility graph is necessary for contextuality. In Ref. [32], the authors proved that chordality is also a sufficient condition to witness contextuality. In fact, the authors showed that quantum theory is enough to extract contextuality from any scenario with a non-chordal compatibility graph.

### 2.3.2 Cycle scenarios

Due to the foundational stature, attributed by Vorob'ev's theorem, cycle scenarios make consistent appearances in this thesis. More formally, these scenarios are called $n$-cycle scenarios, where $n$ denotes the number of measurements. They are called so because their compatibility graph is isomorphic to a cycle graph $C_n$, see Fig. 8 where measurements $A_0, \ldots, A_{n-1}$, are cyclically (modulo $n$) compatible. Among the cycle scenarios that are interesting for KS contextuality except for $n = 4$, which is the

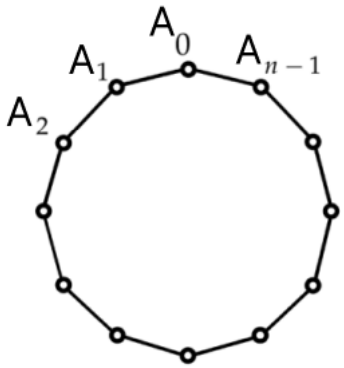



Figure 8: An $n$-cycle compatibility graph with measurements $A_0, A_1, \ldots, A_{n-1}$.

CHSH scenario, no other $n$-cycle scenario is a Bell scenario. The simplest non-Bell-like scenario is the famous KCBS scenario, represented by a 5-cycle. In Ref. [12], the authors characterised the $n$-cycle scenarios at the classical, quantum, and no-disturbance levels. For any given $n \geq 4$: they derived the minimal H-representation of the non-contextual polytope, provided the quantum states and observables

that maximally violate the facet-inducing inequalities of the non-contextual polytope, and further derived the V-representation of the no-disturbance polytope.

**n-cycle inequalities.** The minimal H-representation of the noncontextual polytope of $n$-cycle scenarios is given by the following inequalities:

$$\sum_{i=0}^{n-1} \gamma_i \langle A_i A_{i\oplus 1} \rangle \overset{NCHV}{\leq} n-2 \tag{2.10}$$

where $\gamma_i \in \{-1,+1\}$ with $\prod_{i=0}^{n-1} \gamma_i = -1$ and $\{A_i\}_{i=0}^{n-1}$ are the measurements labelling each vertex. These inequalities are exploited in Chapter 3 to derive facet-defining inequalities for more complex scenarios. Furthermore, in Chapter 4 we fully characterise the ability of multi-qubit Pauli operators to violate $n$-cycle inequalities.

**n-cycle PR box.** It is further useful the know the non-classical vertices of the no-disturbance polytope of $n$-cycle scenarios. These turn out to be a natural generalization of the PR box correlation from the CHSH scenario ($n = 4$), and we therefore refer to them as $n$-cycle PR boxes. These boxes are defined as follows:

1. For each context $\{A_i, A_{i\oplus 1}\}$, the joint outcomes $(a_i, a_{i\oplus 1})$ in the support of the respective probability distribution satisfy either $a_i \cdot a_{i\oplus 1} = 1$ or $a_i \cdot a_{i\oplus 1} = -1$, where $a_i$ denotes the outcome of measurement $A_i$. In other words, for any given context, the only permissible joint outcomes are either the correlated pairs $\{++, --\}$ or the anti-correlated pairs $\{+-, -+\}$.

2. For each context, the probability distribution is uniform: each of the two permissible joint outcomes occurs with the same probability, $\frac{1}{2}$ [4].

3. The choices between correlation and anti-correlation for each context obey a global constraint: the number of contexts exhibiting anti-correlated outcomes ($a_i \cdot a_{i\oplus 1} = -1$) is odd. Consequently, when $n$ is even, an odd number of contexts exhibit correlated outcomes, while when $n$ is odd, an even number of contexts exhibit correlated outcomes.

In Chapter 5, we will exploit these correlations to study the activation effects of the exclusivity principle.

[4] Such uniform probabilities are actually forced upon us by no-disturbance once the support of the distributions is fixed (by the other two items). The fact that the possibilistic support fixes the values of the probabilities uniquely is actually a feature of any extremal correlation in any scenario [64].

# Chapter 3

# Lifting noncontextuality inequalities

**Chapter Summary.** In this chapter, we introduce a method for producing facet-defining NC inequalities in arbitrary KS contextuality scenarios. This is based on – and strictly extends – the lifting techniques used by Pironio [10] for Bell scenarios. *Lifting* is a commonly used technique in convex polyhedral theory to derive facet-defining inequalities of a polytope in $\mathbb{R}^n$ from facet-defining inequalities of a related polytope in $\mathbb{R}^m$ where $m < n$. We introduce a general method for lifting noncontextuality inequalities, deriving facets of noncontextual polytopes for more complex scenarios from known facets of simpler sub-scenarios. The method further allows us to identify facet-defining NC inequalities for *all* scenarios which admit contextual correlations and thus provides a key to explore an infinity of as-of-yet unexplored scenarios. This is ensured by Vorob'ev's theorem [11] which guarantees that any contextuality-witnessing scenario contains an induced $n$-cycle sub-scenario (for some $n \geq 4$), and by the complete characterisation of the noncontextual polytopes for all such cycle scenarios [12]. Overall, our results shed light on the structure of noncontextuality polytopes and the relationship between such polytopes across different scenarios.

**Motivation.** A minimal set of conditions for deciding whether measurement statistics are contextual is provided by the set of inequalities that support the facets of the noncontextual polytope. However, all known facet-defining NC inequalities are either for Bell scenarios [65] or refer to cyclic [52] or state-independent [53, 66, 67] contextuality scenarios. Characterising all the facet-defining inequalities of such polytopes is a notoriously hard problem (NP-complete [68]) to solve in general. As such, there are few fully characterised scenarios, for which all facet-defining NC inequalities are known: non-Bell scenarios include two-outcome $n$-cycle ($n \geq 5$) scenarios [12], Bell scenarios include the $(2, 2, 2)$ CHSH scenario [65], and various classes generalising it such as $(2, m', 2)$ [69, 70], $(n', 2, 2)$ [71], and $(2, 2, k')$ [72] scenarios, where $(n', m', k')$ stands for $n'$ parties, each with $m'$ measurement settings, each with $k'$ possible outcomes.

Despite the demoralising hardness of characterising arbitrary scenarios, some non-trivial work has been done linking simpler Bell scenarios with more complex ones. In Ref. [10], Pironio proposed a method

to derive (some of the) facet-defining Bell inequalities of complex Bell polytopes starting from known inequalities of simpler Bell polytopes. It employed the idea of *lifting*, a commonly used technique in convex polyhedral theory to derive facet-defining inequalities of a polytope in $\mathbb{R}^n$ from facet-defining inequalities of a related polytope in $\mathbb{R}^m$ where $m < n$. The upshot is that once the facets of a simpler polytope have been fully or partially identified, one need not start from scratch when searching for the facets of a more complex polytope. One may instead concentrate efforts on finding the facet-defining inequalities that are absent in or do not arise from simpler scenarios.

Aiming to foster facet characterisation for general Bell scenarios, Pironio showed that any facet-defining inequality of an arbitrary Bell polytope can be lifted to one or more facet-defining inequalities of any more complex Bell polytope, where by 'more complex' we mean a Bell scenario with more parties, more local measurements for a party, or more outcomes for a measurement (or a combination of all three) than the original scenario. Building on this work, in Ref. [73] Pironio characterised Bell scenarios whose only facets are given by liftings of the CHSH inequality. These include e.g. the bipartite scenarios where one party has a binary choice of dichotomic measurements, irrespective of the number of measurement settings and outcomes for the other party.

Non-Bell-like KS contextuality scenarios, on the other hand, have not received as much attention in terms of facet characterisation. This work aims to address this gap.

**Outline of the chapter.** With increasing intensity, we explain the details of our lifting method in this chapter, starting from a non-technical introduction in Sec. 3.1 where we help build some intuition by providing a concise, high-level summary of our main results devoid of any mathematical statements to then add more refinements to it in Secs. 3.2 and 3.3, to finally provide the full technical details in Sec. 3.4. Sec. 3.2 helps the reader in visualizing the idea of lifting using examples in one, two and three dimensions. In Sec. 3.3, we explain why our generalisation, of Pironio's liftings for Bell inequalities, is non-trivial. Sec. 3.4 introduces the relevant notations, for the proofs to follow later, in Sec. 3.4.1: the common logical underpinnings of the two lifting proofs in Sec. 3.4.2; the measurement lifting results in Sec. 3.4.3; and the outcome lifting results in Sec. 3.4.4. We then apply our lifting results to extract facet-defining noncontextuality inequalities for scenarios for which no such inequalities were hitherto known in Sec. 3.5. We finally close the chapter with a discussion on the future prospects of our results in Sec. 3.6.

## 3.1 Overview

We now give a concise overview of our lifting technique to later refine it and help the reader develop intuition for our results.

A scenario **S** can be extended to a more complex scenario **T** by adding more measurements and/or more outcomes. We focus on one such step at a time: either adding *one* measurement or adding *one* outcome for an already existing measurement. An arbitrary extension can be seen as a sequence of such single-step extensions. We fix an initial facet-defining NC inequality for **S** which we aim to lift to (one or more) facet-defining NC inequalities for **T**. We achieve this in both situations, but the specifics differ somewhat.

**Measurement lifting.** When **T** is obtained from **S** by adjoining a new measurement $A$, the method of lifting depends both on the compatibility relations between $A$ and the pre-existing measurements and on the facet-defining inequality being lifted, namely on the set of measurements that *effectively contribute* to that inequality. We distinguish two cases:

I. If $A$ is incompatible with some measurement that contributes to the initial inequality, then the inequality lifts unchanged to a facet-defining inequality for the scenario **T**. The new measurement $A$ is traced out and does not effectively contribute to the lifted inequality, thus it need not be performed for testing the inequality.

II. If $A$ is compatible with all the measurements that contribute to the initial inequality, then the inequality lifts to a facet-defining inequality of **T** for each outcome $a_k$ of $A$. Each such lifted inequality can be tested by first measuring $A$, post-selecting on obtaining the outcome $a_k$, and then testing the initial inequality.

**Outcome lifting.** When **T** is obtained from **S** by adding a new outcome $a_0$ for an already existing measurement $A$, the original inequality lifts to a facet-defining inequality on **T** for each choice of (pre-existing) outcome $a_k \neq a_0$ of $A$. The choice indicates the outcome $a_k$ of $A$ with which $a_0$ is to be 'clubbed together'. When testing the inequality, any occurrence of the new outcome $a_0$ for $A$ is treated as if it were an $a_k$ in the original scenario. An exception is that the initial inequality cannot itself be the result of case II measurement lifting with the *same* choice of outcome $a_k$: intuitively, that would entail post-selecting on two different outcomes for $A$.

## 3.2 Visualising lifting

Before diving into the detailed description of our results and proofs in Sec. 3.4, we offer an accessible, intuitive sketch of the main ingredients of our lifting method. This section uses small – trivial yet visualisable – examples to convey the flavour of the various forms of lifting.

A scenario **S** can be extended to a larger scenario **T** through the addition of new measurements and/or outcomes. In each case, **S** is a sub-scenario of **T**. It is natural to enquire whether the knowledge of a facet-defining NC inequality for the sub-scenario **S** gives us some information about facet-defining NC inequalities for **T**. Lifting answers this question in the affirmative in that it is a method to transform each facet-defining inequality of the noncontextual polytope $S$ into one or more facet-defining inequalities of the polytope $T$.

The noncontextual polytope for the simplest contextuality-witnessing scenario (i.e. a scenario for which $L_{NC} \neq L_{ND}$) is $8$-dimensional. This makes it impossible to visualise lifting starting from this polytope. Nonetheless, there are some simple visualisable NC polytopes – albeit for scenarios unable to witness contextuality – which aptly capture the idea of lifting. We use these examples to provide some intuition for the idea behind our lifting method. Despite being thoroughly uninteresting from the point of view of contextuality, these examples are indeed special cases of our lifting results which 'contain all the germs of generality'.

**Measurement lifting**: For measurement lifting, we consider two cases, as outlined in Sec. 3.1.

**Case I** applies when the new measurement added to **S** is not simultaneously compatible with all the measurements effectively contributing to the facet-defining inequality being lifted, i.e. it is incompatible with at least one such measurement; see Sec. 3.4.3 for details including the meaning of 'effectively contributing'. We illustrate this case with an example that serves as a proxy for all such situations.

Consider **S** the scenario with two dichotomic measurements, $A$ and $B$, that are incompatible with each other. Its NC polytope $S$ is two-dimensional (embedded in a four-dimensional ambient vector space) and is shown on the left-hand side of Figure 9(a). Its four vertices correspond to the deterministic assignments – $00$, $01$, $10$, $11$ – to measurements $A$ and $B$, in that order. Observe that the fact that $A$ and $B$ are incompatible induces an affine dependency among the vertices of the polytope,

$$v_{00} - v_{01} + v_{11} - v_{10} = 0, \tag{3.1}$$

which would not hold were $A$ and $B$ compatible; cf. the simplex on the right-hand side of Figure 9(b). This kind of affine relation induced by incompatibility plays a crucial role in our proofs; see for example Eq. (3.11).

Since $S$ is a two-dimensional polytope, its facets are one-dimensional. One of its facet-defining inequalities is depicted in the figure; it supports the facet containing the vertices $v_{10}$ and $v_{11}$.

Now consider extending **S** to **T** by introducing a new measurement $C$ incompatible with both $A$ and $B$. The NC polytope $T$ is three-dimensional (embedded in a six-dimensional ambient vector space), and is shown on the right-hand side of Figure 9(a). Notice that each outcome of $C$ determines an extension of each deterministic assignment in **S** (hence, vertex of $S$) to a deterministic assignment in **T** (vertex of $T$); e.g. $v_{010}$ and $v_{011}$ in $T$ are the two extensions of $v_{01}$ in $S$. One can think of this as captured by a polytope projection from $T$ to $S$ which 'forgets' the outcome of $C$.

The polytope $T$ has a facet with vertices $v_{100}$, $v_{101}$, $v_{110}$, and $v_{111}$, whose supporting inequality is depicted on the right of Figure 9(a). These four vertices of $T$ are precisely the extensions of the vertices $v_{01}$ and $v_{11}$ of $S$, which saturate the initial inequality depicted on the left. This inequality shown on the right of Figure 9(a) is thus the *lifting* of the inequality on the left. One can similarly obtain three other facet-defining inequalities of $T$ from the remaining three facets of $S$.

In general, case I measurement lifting maps a facet $F$ of $S$ to the facet of $T$ whose set of vertices is exactly the set of all extensions of vertices in the original facet $F$. As we see in Sec. 3.4.3, the explicit form of the lifted inequality turns out to be, in a sense, the same as that of the initial inequality, since the outcome of the new measurement is ignored and thus 'traced out'.

**Case II** applies when the newly added measurement is compatible with all the preexisting measurements that effectively contribute to the facet-defining inequality being lifted.

To visualise lifting in this case, we take **S** to be the scenario with a single dichotomic measurement, say $A$. Its NC polytope $S$ is a one-dimensional line segment or $1$-simplex (embedded in a two-dimensional ambient vector space), shown on the left-hand side of Figure 9(b). It has only two facet-defining inequalities. One of them, saturated only by the vertex $v_0$, is depicted in the figure.

We extend this scenario to **T** by adding a dichotomic measurement $B$ compatible with $A$. The resulting NC polytope $T$ is a three-dimensional tetrahedron or $3$-simplex (embedded in a four-dimensional ambient vector space), shown on the right-hand side of Figure 9(b).

The facet-defining inequality depicted in the figure is one of two possible liftings of the initial inequality on the left of Figure 9(b). It is saturated by the vertices $v_{00}$, $v_{01}$, and $v_{11}$. The first two are all the extensions of $v_0$, the vertex that saturates the initial inequality, while the third is one of the two possible extensions of $v_1$, the other, non-saturating vertex of $S$. The vertex of $T$ corresponding to the other extension, $v_{10}$, does not saturate the inequality shown on the right of Figure 9(b). However, a different lifting of the same initial inequality shown on the left of Figure 9(b) would support the facet containing the vertices $v_{00}$, $v_{01}$, and

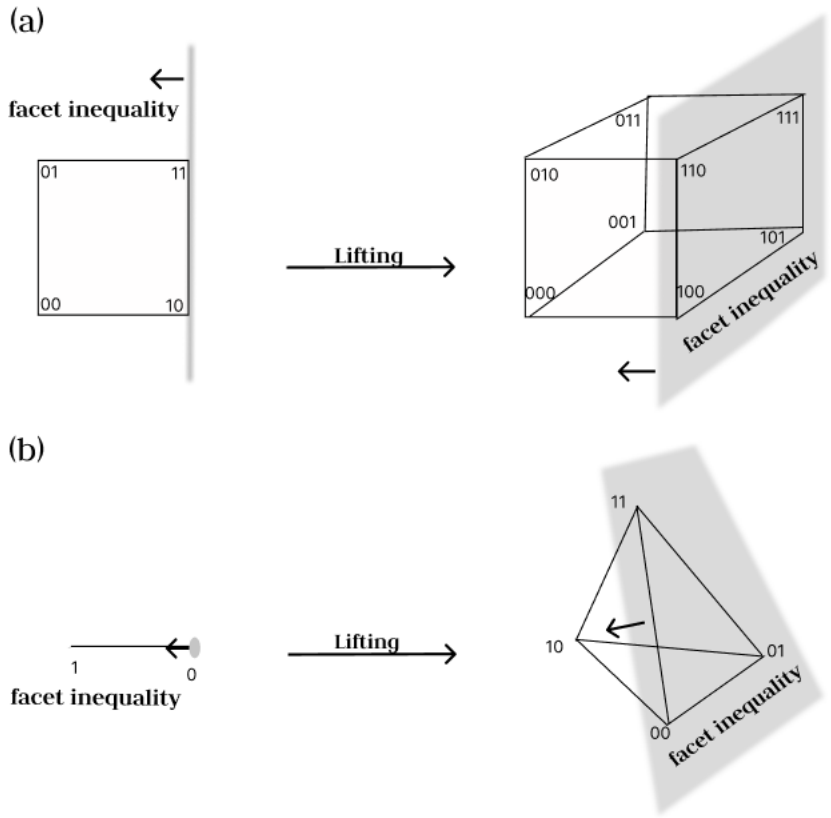



Figure 9: Depiction of measurement lifting in simple scenarios with dichotomic measurements. (a) Case I measurement lifting, from the scenario **S** with two incompatible measurements to the scenario **T** with three pairwise-incompatible measurements. The shaded line and directional arrow on the left represent a facet-defining inequality saturated by vertices $v_{10}$ and $v_{11}$, while the shaded plane and arrow on the right represent the lifted inequality, saturated by all extensions of $v_{11}$ and $v_{10}$ to $T$. (b) Case II measurement lifting, from the scenario with a single measurement to the scenario with two compatible measurements. The facet-defining inequality saturated only by $v_0$, depicted on the left, is lifted to a facet-defining inequality saturated by both extensions of $v_0$ ($v_{00}$ and $v_{01}$) plus one extension of $v_1$ ($v_{11}$), depicted on the right.

$v_{10}$.

In general, case II measurement lifting maps a facet $F$ of $S$ to a facet of $T$ whose set of vertices consists of:

(i) every possible extension of the vertices in the original facet $F$, and

(ii) every possible extension of the remaining vertices of $S$ (not in the original facet $F$) except those that assign a chosen fixed outcome to the new measurement (outcome $0$ for $B$ in the example depicted above). Varying the choice of this fixed outcome for the new measurement yields different liftings to $T$ of the same facet-defining inequality of $S$.

Notice that in both cases I and II above, the lifted inequality is effectively testing the original one, i.e. it simply represents a re-expression of the original inequality in the language of the larger polytope $T$. Moreover, while in case I each facet of $S$ yields a single facet of $T$, in case II, it yields one facet for each possible outcome of the newly added measurement.

**Outcome lifting**: In the case of outcome lifting, we consider extending a scenario **S** to **T** by adding an extra outcome $a_0$ to an existing measurement $A$. Here, unlike in the case of measurement lifting, there is no (immediate) concept of extension of assignments, and thus of vertices of $S$ to vertices of $T$. In fact, the vertices of $S$ could be seen as a strict subset of those of polytope $T$.

These vertices behave the same way with respect to an outcome-lifted inequality in $T$ as they do with respect to the initial facet-defining inequality of $S$, either saturating both or falling short by the same amount. However, there are more vertices in the larger polytope $T$, namely those that assign the new outcome $a_0$ to the measurement $A$. Each such vertex behaves exactly as the vertex obtained by substituting $a_k$ for the outcome of $A$ (while leaving the rest of the assignment intact), for some fixed choice of outcome $a_k$ of $A$ already present in the initial scenario **S**.

In other words, a facet $F$ of $S$ is lifted to a facet of $T$ whose set of vertices consists of: the vertices in $F$ plus the vertices obtained from a vertex in $F$ by replacing (the fixed) outcome $a_k$ by $a_0$ for the measurement $A$.

Similar to case II measurement lifting, there is an element of choice involved. Indeed, for each choice of $k$, one obtains a (possibly different) facet-defining inequality of $T$. There is one exception, though: choosing $k$ fails to yield a facet of $T$ when the initial facet of $S$ is itself obtainable from a sub-scenario of **S** via case II measurement lifting by fixing the choice of outcome $a_k$ when adding measurement $A$; in other words, when the facet-defining inequality being lifted is such that $A$ is compatible with all the measurements effectively contributing to it and all its non-saturating vertices assign outcome $a_k$ to $A$.

One may think of each such choice of $k$ as determining a polytope projection $T \to S$ which performs a coarse-graining of the outcomes of $A$ by clubbing together $a_0$ and $a_k$ into the same outcome. The lifted inequality then corresponds to 'tracing out' along this identification, much as in case I measurement lifting. This means that the form of the lifted inequality is such that it does not distinguish between the vertices with $a_k$ and $a_0$ as outcomes of $A$, while the remaining vertices behave as they did with respect to the original inequality. Indeed, in the new scenario, the lifted inequality is still in effect testing the original inequality. That is why the new outcome needs to be 'clubbed together' with – and thus made indistinguishable from – some pre-existing outcome, so that the 'effective' number of outcomes is kept the same.

## 3.3 Lifting: from Bell to KS

We now comment on the comparison with the work of Pironio in Ref. [10] for the case of Bell scenarios, a source of inspiration for the present work.

From an arbitrary contextuality scenario, the simplest (single-step) extensions are of two types:

(a) adding a new measurement, and

(b) adding a new outcome for an already existing measurement.

Lifting across Bell scenarios, as in Ref. [10], means that the extended scenario must also be a Bell scenario. In such extensions, type (a) can be further divided into two subcases: adding a new party with a single measurement setting, or adding a new measurement to an already existing party. When adding a new party with a single measurement, this new measurement is necessarily compatible with all the preexisting measurements (as they belong to different parties). Hence, this situation is always captured by case II measurement lifting (see Secs. 3.1 and 3.2). On the other hand, when adding a new measurement to an already existing party, the new measurement is necessarily incompatible with all other measurements of that same party. Thus, if the initial Bell inequality being lifted effectively involves this party, this situation falls within case I measurement lifting; otherwise, it falls within case II.

On a more technical note, the following distinction is relevant in establishing the Bell lifting results, and needs to be revised for our generalisation. When adding a new party, the new measurement is appended to all previously existing maximal contexts. By contrast, when adding a new measurement to an already existing party, the previous maximal contexts are preserved intact but new maximal contexts are introduced, namely those involving the newly added measurement. When generalising to arbitrary

contextuality scenarios, one no longer has a notion of parties. Consequently, this neat separation is no longer possible, and we are forced to consider situations in which a mix of new maximal contexts and extensions of previous maximal contexts arises. This is reflected in the general form taken by a correlation vector on the extended scenario, shown in Eq. (3.2) ahead.

## 3.4 Lifting noncontextuality inequalities

We now move on to presenting our results in detail. After fixing some notation in Sec. 3.4.1, we sketch in Sec. 3.4.2 the structure common to our proofs of measurement and outcome lifting. We then delve into each of the forms of lifting in detail, in Secs. 3.4.3 and 3.4.4, respectively.

### 3.4.1 Notation

As before, when extending one scenario to another, the initial scenario is always denoted by **S** and its extension by **T**. The respective noncontextual polytopes are denoted by $S$ and $T$. The dimensions of these polytopes are denoted by $d_S$ and $d_T$, respectively. The newly added measurement (for measurement lifting) or the existing measurement to which a new outcome is added (for outcome lifting) is denoted by $A$.

We use the letter $p$ to refer to a general vector in the space of correlations for **S**, i.e. the ambient vector space of the polytope $S$, and we similarly use $q$ for **T**. We reserve the letters $u, v, w$ for vertices of $S$ or $T$. Recall from Sec 2.2.4 that these vertices are of the form $v_\lambda$ corresponding to global assignments $\lambda$ of outcomes to all measurements. To simplify the presentation, from now on we often omit the distinction between a vertex and its underlying assignment: e.g. we say that a vertex $v$ assigns a value to a measurement, and write $v|_C$ for the joint outcome assignment over a context $C$ determined by the vertex $v$ (the restriction of the underlying global assignment to the measurements in $C$).

Recall that in specifying correlations on a scenario we only consider the maximal contexts, since statistics for their subsets can be recovered by marginalisation unambiguously due to no-disturbance. We use the term *partial context* to refer to a proper subset of a maximal context, whenever such emphasis is required.

Non-full-dimensional convex polytopes satisfy non-trivial affine equality constraints. In the case of NC polytopes, these correspond to the no-disturbance and normalisation conditions. A valid inequality $b \cdot p \geq \beta$ for a polytope remains valid if one adds to it any linear combination of the equality constraints. In the case of NC polytopes, one can always add a multiple of a normalisation condition to rewrite any NC

inequality so that its lower bound is zero. Therefore, without loss of generality, an arbitrary NC inequality will be represented in the form $b \cdot p \geq 0$. Moreover, at various points, we make use of the freedom afforded by the (linear) no-disturbance conditions to pick particularly convenient vectors of coefficients.

We typically denote by $B_P$ an affine basis of interest for a polytope $P$, whose elements are taken from the vertices of the polytope. We make crucial use of the fact that $\dim(P)$ vertices of $B_P$ can be chosen from any given facet. When lifting a facet-defining inequality of $S$, we consider such an affine basis $B_S = \{u_i\}_{i=1}^{d_S+1}$. For slight convenience, we enumerate $B_S$ in a way that lists $d_S$ saturating vertices of the inequality (i.e. vertices in its supported facet) as the first $d_S$ elements, with the last element being the non-saturating one.

### 3.4.2 Proof outline

Before burdening the reader with the mathematical details of the proofs of our results, we outline the logical structure underlying them. The proofs of measurement lifting and outcome lifting follow a common script.

Let $I_T$ be the inequality claimed to be the lifting of inequality $I_S$. To prove this claim we proceed as follows:

1. We pick an affine basis $B_S = \{v_1, \ldots, v_{d_S+1}\}$ of vertices of $S$ chosen so that its first $d_S$ points saturate the initial inequality $I_S$.

2. We use it to construct an over-complete affine basis $B_T$ of vertices of $T$, partitioned as $B_T = \bigcup_j V_j$, where each component $V_j = \{v_1^{(j)}, \ldots, v_{d_S+1}^{(j)}\}$ consists of $d_s + 1$ affinely independent vertices of $T$. Each $v_i^{(j)}$ is built from the corresponding vertex $v_i$ of $B_S$. The specific construction depends on the type of lifting under consideration, but the common aspect is that $v_i^{(j)}$ is an 'extension' of $v_i$ in that it is mapped to $v_i$ under the appropriate polytope projection $T \to S$.

3. By construction, the first $d_S$ vertices of each partition component $V_j$ saturate the inequality $I_T$. Moreover, there is (at least) one partition component $V_k$ where the last vertex does not saturate $I_T$.

4. For all other partition components, i.e. for all $j \neq k$, either the last element $v_{d_S+1}^{(j)}$ saturates the inequality $I_T$, or one can affinely eliminate it from $B_T$ by expressing it as an affine combination of the remaining vertices of $V_j$ and the vertices of $V_k$, essentially using Eq. (3.1). This leaves only one remaining non-saturating vertex within $B_T$, namely $v_{d_S+1}^{(k)}$, the last vertex in $V_k$.

5. This is enough to establish the claim. Since (the reduced) $B_T$ affinely spans $T$, it contains a set of

$d_T + 1$ affinely independent vertices. From the previous step, it has only one non-saturating vertex. It must therefore contain a set of $d_T$ affinely independent vertices that saturate $I_T$, proving that this inequality is facet-defining.

### 3.4.3 Measurement lifting

We first consider lifting under measurement extensions. Let **S** be extended to **T** by the addition of a new measurement $A$. Write $J \subseteq \mathcal{X}_{\mathbf{S}}$ for the set of measurements in **S** that are compatible with $A$, i.e. the neighbourhood of $A$ in the compatibility graph of **T**.

#### The extended ambient space

For any (not necessarily maximal) context, i.e. set of pairwise compatible measurements, $U$ contained in $J$, the set $U \cup \{A\}$ is a context of **T**. If $U$ is maximal within $J$ then this results in a maximal context. Thus, $A$ may contribute to maximal contexts of **T** in two ways:

1. create maximal contexts by being compatible with maximal contexts in $\mathcal{C}_{\mathbf{S}}$: for each context maximal within $J$ that is also maximal within $\mathcal{X}_{\mathbf{S}}$ (i.e. belongs to $\mathcal{C}_{\mathbf{S}}$), appending $A$ to it gives a (maximal) context in $\mathcal{C}_{\mathbf{T}}$.

2. create maximal contexts by being compatible with partial contexts in **S**: for each context maximal within $J$ that is a partial context with respect to $\mathcal{X}_{\mathbf{S}}$ (i.e. does not belong to $\mathcal{C}_{\mathbf{S}}$), appending $A$ to it gives a (maximal) context in $\mathcal{C}_{\mathbf{T}}$.

Let $\mathcal{C}_{\mathsf{mc}+A}$ denote the set of all contexts in $\mathcal{C}_{\mathbf{T}}$ arising as in item 1 above, where 'mc' reminds us that these contexts are extensions of maximal contexts of **S**. Likewise, write $\mathcal{C}_{\mathsf{pc}+A}$ for the set of contexts in $\mathcal{C}_{\mathbf{T}}$ arising as in item 2, where 'pc' indicates that they are extensions of partial contexts of **S**. Notice that the case covered by item 2 also implies the existence of maximal contexts in $\mathcal{C}_{\mathbf{S}}$ that are not fully compatible with $A$, and thus remain maximal contexts in $\mathcal{C}_{\mathbf{T}}$. We write $\mathcal{C}_{\mathsf{old}}$ for the set of all such contexts. We can therefore write $\mathcal{C}_{\mathbf{T}}$ as a disjoint union of three mutually exclusive sets:

$$\mathcal{C}_{\mathbf{T}} = \mathcal{C}_{\mathsf{old}} \cup \mathcal{C}_{\mathsf{mc}+A} \cup \mathcal{C}_{\mathsf{pc}+A}.$$

This induces a direct sum decomposition of the vector space of correlations for **T**, which is the ambient space of the polytope $T$. An arbitrary vector $q$ in this space can be written as

$$q = [q_{\mathsf{old}}, q_{\mathsf{mc}+A}, q_{\mathsf{pc}+A}], \tag{3.2}$$

where $q_{\text{old}}$ represents the components indexed by joint outcomes of contexts in $\mathcal{C}_{\text{old}}$ and likewise for $q_{\text{mc}+A}$ and $q_{\text{pc}+A}$.

### Effectively contributing measurements

We need to consider two different forms of measurement lifting, depending on the (compatibility behaviour of the) new measurement and on the initial inequality being lifted. The distinction hinges on the notion of an *effectively contributing measurement to an inequality*, which we define by its negation.

Consider an inequality $b \cdot p \geq 0$ over the space of correlations for **S**. In short, a measurement $M \in \mathcal{X}_{\mathbf{S}}$ is said *not* to contribute effectively to the inequality when the latter is insensitive to the outcome of $M$, i.e. its left-hand side remains invariant under any change of outcome for $M$, rendering this measurement's role irrelevant to the inequality.

More formally, recall from Sec. 2.2.4 that the vertices of $S$ are deterministic noncontextual correlations $v_\lambda$, determined by global assignments $\lambda$ of outcomes to all measurements in $\mathcal{X}_{\mathbf{S}}$. Write $\lambda[M \mapsto m_k]$ for the assignment obtained from $\lambda$ by changing $M$'s outcome to $m_k$,

$$\lambda[M \mapsto m_k](B) = \begin{cases} m_k & \text{if } B = M \\ \lambda(B) & \text{otherwise.} \end{cases}$$

The condition is then that for any $\lambda$ and any outcome $m_k$ of $M$, one has $b \cdot v_\lambda = b \cdot v_{\lambda[M \mapsto m_k]}$. In terms of the vector of coefficients $b$, this means that, using the equality constraints (viz. no-disturbance), $b$ can be brought into a form that is symmetric under any change of outcome for $M$, i.e. such that

$$b(m_k\, s \mid M\, U) = b(m_j\, s \mid M\, U) \tag{3.3}$$

for all $m_k$, $m_j$ outcomes of $M$, and all maximal contexts $\{M\} \cup U$ and outcome assignments to the remaining measurements $s \in \mathcal{O}_U$. With the coefficient vector brought to such a symmetric form, one can unambiguously write $b(s \mid U)$ for the unique coefficient in Eq. (3.3), without reference to the measurement $M$. One could, in fact, effectively regard the inequality as being defined over the sub-scenario that excludes the measurement $M$.

We assume throughout that the starting inequality on **S** has been put in such an *effective* form, eliminating *all* the measurements not effectively contributing to it. Note that this means that there is a single $b(s \mid U)$ whenever $U$ is a partial context of **S** that is maximal within the set of effectively contributing measurements, even if $U$ belongs to more than one maximal context of **S**, when combined with different sets of non-effectively-contributing measurements.

**The form of the lifted inequalities**

Let $b \cdot p \geq 0$ be a facet-defining NC inequality for the scenario **S**. What we aim, effectively, is to test this inequality in the extended scenario **T**. This is the premise of lifting: choosing a vector of coefficients $b'$ in such a way that the inequality $b' \cdot q \geq 0$ over the space of correlations for **T** remains in essence equivalent to $b \cdot p \geq 0$, but also is a facet-defining NC inequality for **T**. In other words, the lifted inequality $b' \cdot q \geq 0$ still effectively tests (reduced) statistics on the sub-scenario **S** but given statistics for **T**. This requires tweaking with the role of the added measurement $A$ in the expression $b' \cdot q$ by a judicious choice of the coefficients $b'$.

To this end, we must consider two different constructions of coefficient vectors built from $b$, which provide an appropriate answer in different circumstances. In summary (to be expanded below):

I. the coefficient vector $b_{I,tr(A),0}$ traces $A$ out in the $\mathcal{C}_{\mathsf{mc}+A}$ components whereas it zeroes out each component in $\mathcal{C}_{\mathsf{pc}+A}$, leaving $\mathcal{C}_{\mathsf{old}}$ ones as they appear in the original inequality;

II. the coefficient vector $b_{0,a_k,a_k}$ enforces only the conditional statistics in $b' \cdot q$: it fixes an outcome $a_k$ of $A$ within the contexts in $\mathcal{C}_{\mathsf{mc}+A} \cup \mathcal{C}_{\mathsf{pc}+A}$ (i.e. the contexts containing $A$), treating the corresponding components like the original coefficient vector $b$, while it zeroes out all components where the outcome of $A$ is not equal to $a_k$, including all components in the $\mathcal{C}_{\mathsf{old}}$ part.

The subscripts '$I, tr(A), 0$' and '$0, a_k, a_k$' above were chosen to be suggestive of the action of the coefficient vector on each of the three components of the space of correlations as defined in Eq. (3.2) in comparison to that of the initial vector of coefficients $b$: '$I$' for leaving the action unchanged from $b$, '$0$' for picking coefficients equal to zero, '$tr(A)$' for tracing out the outcomes of the new measurement $A$, '$a_k$' for selecting on a chosen outcome $a_k$ for $A$.

Depending on the compatibility relations between the newly added measurement $A$ and the preexisting measurements in $\mathcal{X}_{\mathbf{S}}$, one of the above options gives a vector of coefficients for a lifted inequality. We now explain when to use which form and describe the constructions in more detail. Bear in mind once again that, in each case, these will only give *one* possible form for the inequality supporting the lifted facet: since the noncontextual polytopes are not full-dimensional, there is not a unique form (i.e. vector of coefficients) for the inequality supporting each facet.

**Case I** is used to obtain a lifted inequality when the measurement $A$ is not compatible with all the measurements effectively contributing to the initial inequality $b \cdot p \geq 0$. In this case, the lifted inequality takes the form

$$b_{I,tr(A),0} \cdot q \geq 0. \tag{3.4}$$

where we build from $b$ the new vector of coefficients $b_{I,tr(A),0}$ as follows:

- over a context $C$ in $\mathcal{C}_{\mathsf{old}}$, we pick the same coefficients as in the initial inequality thus leaving $b$ unchanged: for each $s \in \mathcal{O}_C$,

$$b_{I,tr(A),0}(s \mid C) = b(s \mid C); \tag{3.5}$$

- over a context $\{A\} \sqcup C$ in $\mathcal{C}_{\mathsf{mc}+A}$, we trace out measurement $A$ by assigning coefficients taken from $b$ irrespective of $A$'s outcome: for each $s \in \mathcal{O}_C$ and $a \in \mathcal{O}_A$,

$$b_{I,tr(A),0}(a\, s \mid A\, C) = b(s \mid C); \tag{3.6}$$

- a context $\{A\} \sqcup U$ in $\mathcal{C}_{\mathsf{pc}+A}$ is ignored completely by picking the coefficients to be zero: for each $s \in \mathcal{O}_U$ and $a \in \mathcal{O}_A$,

$$b_{I,tr(A),0}(a\, s \mid A\, U) = 0.$$

To better understand this construction, note that any non-disturbing correlation $q$ on **T** determines a correlation $q|_{\mathbf{S}}$ on the sub-scenario **S** by marginalisation, forgetting the outcome of $A$ by summing over all the possibilities. Concretely, given $C \in \mathcal{C}_{\mathbf{S}}$, either it is in $\mathcal{C}_{\mathsf{old}}$ and then $q|_{\mathbf{S}}(s \mid C) = q(s \mid C)$, or it is wholly compatible with $A$ (so that $\{A\} \sqcup C \in \mathcal{C}_{\mathsf{mc}+A}$) and then $q|_{\mathbf{S}}(s \mid C) = \sum_{a \in O_A} q(a\, s \mid A\, C)$. The inequality determined by $b_{I,tr(A),0}$ effectively tests for the initial inequality $b \cdot p \geq 0$ over such marginal statistics. That is, for all non-disturbing correlations $q$ on **T** we have that

$$b_{I,tr(A),0} \cdot q \;=\; b \cdot q|_S. \tag{3.7}$$

Later, within the proof of Theorem 3, it will become clear why this construction (case I) is not applicable when $A$ is compatible with all the 'effectively contributing measurements'. Despite still yielding a valid inequality, it does not yield a facet-defining one. For those situations, we need to consider the other case that we now describe.

**Case II** is used to obtain a lifted inequality whenever $A$ is compatible with all the effectively contributing measurements for the initial inequality $b \cdot p \geq 0$. Recall that we assume that the vector $b$ is in the effective form described in Sec. 3.4.3. The explicit form of the lifted inequality is then:

$$b_{0,a_k,a_k} \cdot q \geq 0. \tag{3.8}$$

where we build from $b$ the new vector of coefficients $b_{0,a_k,a_k}$ as follows:

- a context $C$ in $\mathcal{C}_{\text{old}}$ is ignored: for each $s \in \mathcal{O}_C$,

$$b_{0,a_k,a_k}(s \mid C) = 0; \tag{3.9}$$

- over a context $\{A\} \sqcup U$ in $\mathcal{C}_{\text{pc}+A} \cup \mathcal{C}_{\text{mc}+A}$, we pick the coefficient from the initial $b$ when the outcome for $A$ is the fixed $a_k$ and ignore the event otherwise: for each $s \in \mathcal{O}_U$ and $a \in \mathcal{O}_A$,

$$b_{0,a_k,a_k}(a\, s \mid A\, U) = \begin{cases} b(s \mid U) & \text{if } a = a_k, \\ 0 & \text{if } a \neq a_k, \end{cases} \tag{3.10}$$

  where $b(s \mid U)$ is well defined due to $b$ being in effective form: see Eq. (3.3) and subsequent explanation, noting that any extension of the context $U$ in **S** consists solely of measurements incompatible with $A$, which are therefore not effectively contributing to the inequality $b \cdot p \geq 0$.

**Recovering the initial inequality**

Since we provide an 'if and only if' condition in the measurement lifting theorem (Theorem 3), it is also worth noticing that the constructions of the new (lifted) inequalities over **T** can be inverted, i.e. one can recover the initial inequality given inequalities (3.4) and (3.8). In going back from scenario **T** to scenario **S**, the contexts in $\mathcal{C}_{\text{old}}$ remain, those in $\mathcal{C}_{\text{pc}+A}$ are eliminated, and the measurement $A$ is dropped from each context in $\mathcal{C}_{\text{mc}+A}$, yielding the set of contexts we denote by $\mathcal{C}_{\text{mc}}$.

Given the coefficient vector $b_{I,tr(A),0}$, one can easily recover the vector $b$: for contexts in $\mathcal{C}_{\text{old}}$, the components are unchanged; for a context $C \in \mathcal{C}_{\text{mc}}$, the components can be read off from the components of $b_{I,tr(A),0}$ on the corresponding context $C \cup \{A\} \in \mathcal{C}_{\text{mc}+A}$, since $b(s \mid C) = b_{I,tr(A),0}(a\, s \mid A\, C)$ for any $a \in \mathcal{O}_A$.

The second case is a little more subtle. Given the vector $b_{0,a_k,a_k}$, one may not necessarily recover the original coefficient vector $b$, but one can recover some $b'$ that determines an equivalent inequality. Note that $b_{0,a_k,a_k}$ has non-zero coefficients only in the components corresponding to $\mathcal{C}_{\text{mc}+A} \cup \mathcal{C}_{\text{pc}+A}$ For each context $U \cup \{A\} \in \mathcal{C}_{\text{mc}+A} \cup \mathcal{C}_{\text{pc}+A}$, pick some maximal context $C_U$ in $\mathcal{C}_{\mathbf{S}}$ that extends $U$, and let $R_U = C_U \setminus U$ be the set of added measurements. Note that if $U \cup \{A\}$ comes from $\mathcal{C}_{\text{mc}+A}$, then $U$ is already itself maximal in **S**, being in $\mathcal{C}_{\text{mc}}$, and so $R_U = \varnothing$; if it comes from $\mathcal{C}_{\text{pc}+A}$, then $U$ is a partial context of **S** which may thus be extended to some maximal context in $\mathcal{C}_{\text{old}}$ (in this case the choice of $C_U$ is not necessarily unique). Now, for any $s \in \mathcal{O}_U$ and $r \in \mathcal{O}_{R_U}$, set $b'(s\, r \mid U\, R_U) = b_{0,a_k,a_k}(a_k\, s \mid A\, U)$. The components in $b'$ corresponding to contexts not of the form $C_U$ for any $U$ are set to zero. Note that the assignment $U \mapsto C_U$ is injective: if $U_1, U_2$ are both included in the same maximal context and

$U_i \cup \{A\} \in \mathcal{C}_{\mathsf{pc}+A}$, then $U_1 \cup U_2 \cup \{A\}$ is itself a context including both $U_i \cup \{A\}$, which forces $U_1 = U_2$ by maximality of the contexts in $\mathcal{C}_{\mathsf{pc}+A}$.

**Validity of the lifted inequalities**

Before proving our main results regarding their facet-defining nature, we prove the validity of these newly defined inequalities for the polytope $T$.

Recall from Sec 6 that it is sufficient to test the inequality on vertices of $T$. Over each maximal context $C$, a vertex $v$ has precisely one non-zero component, equal to $1$, that of the joint outcome $v|_C \in \mathcal{O}_C$. Thus, evaluating an inequality on a vertex amounts to adding up the coefficients corresponding to those components, a single surviving one for each maximal context.

Moreover, note that any vertex $v$ of $T$ determines a vertex $u$ of $S$ by restricting the global assignment to the measurements in $\mathcal{X}_{\mathbf{S}}$.

**Proposition 1.** *Let the scenario* **T** *be an extension of the scenario* **S** *by the addition of a single measurement $A$. Then $b \cdot p \geq 0$ is a valid inequality for $S$ if and only if $b_{I,tr(A),0} \cdot q \geq 0$ is a valid inequality for $T$.*

*Proof.* This essentially amounts to Eq. (3.7), but we now show it explicitly in terms of the vertices.

Let $v$ be a vertex of $T$, and $u = v|_{\mathbf{S}}$ be the corresponding vertex of $S$ obtained by restriction of the global outcome assignment. The non-zero coefficients in $b_{I,tr(A),0}$ are all in components relative to contexts in $\mathcal{C}_{\mathsf{old}} \cup \mathcal{C}_{\mathsf{mc}+A}$. When evaluating $b_{I,tr(A),0} \cdot v$, we have: over a context $C \in \mathcal{C}_{\mathsf{old}}$ the surviving coefficient corresponds to $v|_C = u|_C$, being given as $b(u|_C \mid C)$ by Eq. (3.5); over a context $\{A\} \cup C \in \mathcal{C}_{\mathsf{mc}+A}$, the surviving coefficient corresponds to $v|\{A\} \cup C$, being given as $b_{I,tr(A),0}(v|_{\{A\}\cup C} \mid A\, C) = b(u|_C \mid C)$ by Eq. (3.6).

Since each context $C \in \mathcal{C}_{\mathbf{S}}$ is either in $\mathcal{C}_{\mathsf{old}}$ or extends to one in $\mathcal{C}_{\mathsf{mc}+A}$, the above cases exhaust all the contexts of **S** and we obtain that $b_{I,tr(A),0} \cdot v \;=\; b \cdot u$, an instance of Eq. (3.7).

The forward direction of the result then follows from the validity of the initial inequality for $S$, which implies $b \cdot u \geq 0$. Conversely, any vertex $u$ of $S$ can be extended to some vertex $v$ of $T$ by picking any outcome for $A$. Validity of the final inequality gives $b_{I,tr(A),0} \cdot v \geq 0$, hence the same instance of Eq. (3.7) implies $b \cdot u \geq 0$. □

**Proposition 2.** *Let the scenario* **T** *be an extension of the scenario* **S** *by the addition of a single measurement $A$, and let $b \cdot p \geq 0$ be an inequality over* **S** *(written in effective form). If $A$ is compatible with all the measurements effectively contributing to the inequality $b \cdot p \geq 0$, then $b \cdot p \geq 0$ is a valid inequality for $S$ if and only if $b_{0,a_k,a_k} \cdot q \geq 0$ is a valid inequality for $T$.*

*Proof.* Let $v$ be a vertex of $T$, and $u = v|_{\mathbf{S}}$ be the corresponding restriction to a vertex of $S$. The only non-zero coefficients in $b_{0,a_k,a_k}$ correspond to contexts in $\mathcal{C}_{\mathsf{mc}+A} \cup \mathcal{C}_{\mathsf{pc}+A}$.

If the vertex $v$ assigns outcome $a_k$ to $A$, then the surviving coefficient at each maximal context $\{A\}\cup U$ is $b_{0,a_k,a_k}(v|_{\{A\}\cup U} \mid A\,U) = b(u|_U \mid U)$ by Eq. (3.10). We thus obtain $b_{0,a_k,a_k} \cdot v \;=\; b \cdot u$, and so $b_{0,a_k,a_k} \cdot v \geq 0$ by validity of the initial inequality for $S$.

If $v$ assigns any other outcome to $A$, then all the surviving coefficients are zero by Eqs. (3.9) and (3.10). Thus, the vertext saturates the inequality, $b_{0,a_k,a_k} \cdot v = 0$.

Conversely, given a vertex $u$ of $S$, it can be extended to a vertex $v$ of $T$ by assigning outcome $a_k$ to $A$, and $b \cdot u \geq 0$ follows from $b_{0,a_k,a_k} \cdot v \;=\; b \cdot u$ and validity of the final inequality. □

### Facetness of the lifted inequalities

We now move to the core of the measurement lifting result: to show that the constructed inequalities for $T$ are facet-defining. For both cases, the proof follows the general plan outlined in Sec. 3.4.2.

Given the initial inequality $b \cdot p \geq 0$ which supports a facet of $S$, we pick

$$B_S = \{v_1, \ldots, v_{d_S+1}\}$$

an affine basis for $S$ made up of vertices, chosen so that the first $d_S$ vertices saturate the inequality. Any vertex of $S$ extends to a vertex of $T$ by fixing an outcome for the new measurement $A$. For each outcome $a_k \in \mathcal{O}_A$, the corresponding extension of vertex $v_i$ is denoted $v_i^{(k)}$. Thus, the set

$$V_k = \{v_1^{(k)}, \ldots, v_{d_S+1}^{(k)}\}$$

contains the vertices of $T$ that are extensions of those in $B_S$ by fixing the outcome $a_k$ for $A$.

There is a one-to-one correspondence between vertices of $S$ and vertices of $T$ that have a fixed outcome $a_k$ for $A$. This correspondence preserves and reflects affine dependencies. From the fact that $B_S$ is an affine basis for $S$, we can therefore conclude that each set $V_k$ is affinely independent and is, moreover, an affine basis for all the vertices of $T$ that assign outcome $a_k$ to $A$. That is, any vertex that assigns outcome $a_k$ to $A$ is an affine combination of the vertices in $V_k$. Consequently, the union $B_T = \bigcup_k V_k$ affinely spans the polytope $T$. In other words, it is an overcomplete affine basis for $T$, where 'overcomplete' indicates that its elements are not necessarily affinely independent, and so some of them could be discarded with the remaining set still affinely spanning $T$.

**Theorem 3.** *Let the scenario* **T** *be an extension of the scenario* **S** *by the addition of a single measurement* $A$*, and let* $b \cdot p \geq 0$ *be an inequality over* **S** *(written in effective form). Then,*

*I. if $A$ is incompatible with some measurement that effectively contributes to the inequality, then $b\cdot p \geq 0$ supports a facet of $S$ if and only if $b_{I,tr(A),0} \cdot q \geq 0$ supports a facet of $T$;*

*II. if $A$ is compatible with all the measurements that effectively contribute to the inequality, then for any outcome $a_k$ of $A$, $b \cdot p \geq 0$ supports a facet of $S$ if and only if $b_{0,a_k,a_k} \cdot q \geq 0$ supports a facet of $T$.*

*Proof.* We establish each case separately.

**Case I.** First, assume that $b \cdot p \geq 0$ supports a facet of $S$. We have to show that the inequality $b_{I,tr(A),0} \cdot q \geq 0$ is saturated by $d_T$ affinely independent vertices and no more. Already from the proof of Proposition 1, or even from Eq. (3.7), we know that $b_{I,tr(A),0} \cdot v_i^{(k)} = b \cdot v_i$ for all $v_i \in B_S$ and $a_k \in \mathcal{O}_A$. Hence, in each $V_k$, all but the last element saturate the inequality $b_{I,tr(A),0} \cdot q \geq 0$.

Fixing some $k$, we show next that for all $l \neq k$, the (non-saturating) vertex $v_{d_S+1}^{(l)}$ can be written as an affine combination of $V_l \cup V_k \setminus \{v_{d_S+1}^{(l)}\}$. This implies that the vertex can be dropped from $B_T$ without affecting its affine span, i.e., the fact that it spans the whole polytope $T$. Doing so for all $l \neq k$, the only non-saturating point that remains within $B_T$ is $v_{d_S+1}^{(k)}$. Consequently, among any set of $d_T + 1$ affinely independent points in (the reduced) $B_T$, $d_T$ of them must saturate the inequality $b_{I,tr(A),0} \cdot q \geq 0$, implying that the inequality is facet-defining for $T$.

In establishing the above-mentioned affine elimination, the key idea is to use an affine relation akin to Eq. (3.1) with respect to the measurement $A$ and some effectively contributing measurement incompatible with $A$, and then expanding on the affine bases $V_k$ and $V_l$. We now see this in detail.

From the assumption, there is a measurement $M$ that effectively contributes to the initial inequality and is not compatible with $A$. Let $w^{(k)}$ be a vertex of $T$ obtained from $v_{d_S+1}^{(k)}$ by changing the outcome of the measurement $M$ to something other than what it is for $v_{d_S+1}^{(k)}$, keeping the rest of the measurement outcomes the same. Likewise, perform the same change of outcome of $M$ to obtain $w^{(l)}$ from $v_{d_S+1}^{(l)}$. Considering the four vectors $v_{d_S+1}^{(k)}, w^{(k)}, v_{d_S+1}^{(l)}, w^{(l)}$, we establish that the following equation holds:

$$w^{(k)} - v_{d_S+1}^{(k)} = w^{(l)} - v_{d_S+1}^{(l)}. \tag{3.11}$$

This has the same form as Eq. (3.1), and holds for much the same reason, deriving from the incompatibility between the measurements $M$ and $A$. It follows from the way $w^{(k)}$ and $w^{(l)}$ are constructed from $v_{d_S+1}^{(k)}$ and $v_{d_S+1}^{(l)}$, respectively: changing the outcome of $M$ means that no change is made to the components corresponding to contexts in $\mathcal{C}_{\mathsf{mc}+A} \cup \mathcal{C}_{\mathsf{pc}+A}$, since $M$ cannot belong to any such context, incompatible as it is with $A$, i.e.,

$$v_{d_S+1}^{(k)}(\mathsf{mc}{+}A, \mathsf{pc}{+}A) = w^{(k)}(\mathsf{mc}{+}A, \mathsf{pc}{+}A), \tag{3.12}$$

$$v_{d_S+1}^{(l)}(\mathsf{mc}{+}A, \mathsf{pc}{+}A) = w^{(l)}(\mathsf{mc}{+}A, \mathsf{pc}{+}A); \tag{3.13}$$

moreover, the fact that $v_{d_S+1}^{(k)}$ and $v_{d_S+1}^{(l)}$ only differ on the outcome they assign to $A$ means that

$$v_{d_S+1}^{(k)}(\mathsf{old}) = v_{d_S+1}^{(l)}(\mathsf{old}), \tag{3.14}$$

and for the same reason

$$w^{(k)}(\mathsf{old}) = w^{(l)}(\mathsf{old}). \tag{3.15}$$

Eq. (3.11) can then be justified as follows: the difference between the two terms on the left-hand side cancels out the components over $\mathcal{C}_{\mathsf{mc}+A} \cup \mathcal{C}_{\mathsf{pc}+A}$ due to Eq. (3.12), and similarly on the right-hand side due to Eq. (3.13); for the components over $\mathcal{C}_{\mathsf{old}}$, the equation follows by subtracting Eq. (3.14) from Eq. (3.15).

Now, due to the symmetry between $V_k$ and $V_l$ and the symmetric construction of $w^{(k)}$ and $w^{(l)}$, we can expand these vectors in the respective affine basis ($V_k$ and $V_l$) with the same coefficients, i.e.,

$$w^{(k)} = \sum_{i=1}^{d_S+1} \lambda_i v_i^{(k)}, \tag{3.16}$$

$$w^{(l)} = \sum_{i=1}^{d_S+1} \lambda_i v_i^{(l)}, \tag{3.17}$$

where $\sum_i \lambda_i = 1$. This results in

$$w^{(k)} - v_{d_S+1}^{(k)} = \sum_{i=1}^{d_S} \lambda_i v_i^{(k)} + (\lambda_{d_S+1} - 1) v_{d_S+1}^{(k)} \tag{3.18}$$

$$w^{(l)} - v_{d_S+1}^{(l)} = \sum_{i=1}^{d_S} \lambda_i v_i^{(l)} + (\lambda_{d_S+1} - 1) v_{d_S+1}^{(l)}. \tag{3.19}$$

Given Eq. (3.11), we can equate the right-hand sides, obtaining

$$\sum_{i=1}^{d_S} \lambda_i v_i^{(k)} + (\lambda_{d_S+1} - 1) v_{d_S+1}^{(k)} = \sum_{i=1}^{d_S} \lambda_i v_i^{(l)} + (\lambda_{d_S+1} - 1) v_{d_S+1}^{(l)}. \tag{3.20}$$

Provided $\lambda_{d_S+1} \neq 1$, we can rearrange this as

$$v_{d_S+1}^{(l)} = v_{d_S+1}^{(k)} - \sum_{i=1}^{d_S} \frac{\lambda_i v_i^{(k)}}{1 - \lambda_{d_S+1}} + \sum_{i=1}^{d_S} \frac{\lambda_i v_i^{(l)}}{1 - \lambda_{d_S+1}}, \tag{3.21}$$

exhibitting $v_{d_S+1}^{(l)}$ as an affine combination of $V_k \cup V_l \setminus \{v_{d_S+1}^{(l)}\}$, as desired.

The last step requires that $\lambda_{d_S+1} \neq 1$. A value of $\lambda_{d_S+1} = 1$ precludes $v_{d_S+1}^{(l)}$ being expressed as an affine combination of $V_k \cup V_l \setminus \{v_{d_S+1}^{(l)}\}$, thwarting the proof strategy. One can work around this problem by picking a different non-saturating vertex of the original inequality as the last element $v_{d_S+1}$ of $B_S$, or by picking a different outcome of $M$ to which its value changes in going from $v_{d_S+1}^{(k)}$ to $w^{(k)}$. We defer

a careful analysis to Sec. 3.4.3, where we show that it is always possible to make a choice that avoids the troublesome $\lambda_{d_S+1} = 1$ whenever the measurement $M$ effectively contributes to the initial inequality and is incompatible with $A$.

**Case II.** We now consider the case when $A$ is compatible with all the effectively contributing measurements to the initial inequality $b \cdot p \geq 0$. Fix an outcome $a_k$ of $A$ and consider the inequality $b_{0,a_k,a_k} \cdot q \geq 0$. In this case, we have that for all $l \neq k$ every vector in $V_l$ saturates the inequality, since as we have seen the left-hand side evaluates to zero. Only in $V_k$ does the $(d_S + 1)$-th vertex not saturate the inequality. The vertex $v^{(k)}_{d_S+1}$ is thus the only element of $B_T$ that does not saturate the inequality. Since $B_T$ affinely spans the polytope $T$, there must be $d_T + 1$ affinely independent vertices within it. As only one element of $B_T$ does not saturate the inequality, there is a set of $d_T$ affinely independent vertices that do. Hence, $b_{0,a_k,a_k} \cdot q \geq 0$ represents a facet-defining inequality for $T$.

**Converse.** This completes our proof of measurement lifting in one direction. Now we go about proving the converse statements, i.e. that if the inequality (3.4) or (3.8) is facet-defining for $T$, then the original inequality $b \cdot p \geq 0$ is facet-defining for $S$. Notice that earlier (in Sec. 3.4.3) we already showed how to recover $b$ from the given inequalities (3.4) or (3.8).

Given $b_{I,tr(A),0} \cdot q \geq 0$ or $b_{0,a_k,a_k} \cdot q \geq 0$ facet-defining, let $F$ be the set of vertices that saturate it, and partition it into two sets $F^{(k)}$ and $F^{(\neg k)}$: those that assign $a_k$ to $A$ and those that do not. A consequence of the assumed facet-defining nature of the inequality is that for any non-saturating vertex $v$ (i.e. any $v \notin F$), the set $F \cup \{v\}$ affinely spans the whole polytope $T$. Given the specific inequalities under consideration, one can always choose such a $v$ that assigns outcome $a_k$ to $A$.
Now, let $w$ be any vertex that assigns $a_k$ to $A$. Expand it as an affine combination of $F \cup \{v\}$:

$$w = \lambda v + \sum_{v_i \in F^{(k)}} \lambda_i v_i + \sum_{v_j \in F^{(\neg k)}} \lambda_j v_j. \tag{3.22}$$

Then, for all vertices that appear in the equation, switch the outcome of $A$ to $a_k$. This transformation preserves the equality because it is a linear map on the space of correlations: for contexts in $\mathcal{C}_{\mathsf{mc}+A} \cup \mathcal{C}_{\mathsf{pc}+A}$, each component of the form $p(a_k\, s \mid A\, U)$ takes a value given by the linear expression $\sum_j p(a_j\, s \mid A\, U)$ and the remaining components become $0$; for contexts in $\mathcal{C}_{\mathsf{old}}$, the components are unchanged. This yields

$$w = \lambda v + \sum_{v_i \in F^{(k)}} \lambda_i v_i + \sum_{v_j \in F^{(\neg k)}} \lambda_j v_j[A \mapsto a_k]. \tag{3.23}$$

Notice that the transformation leaves the vertices in $F^{(k)}$ as well as $v$ and $w$ unchanged, while those in $F^{(\neg k)}$ get mapped to vertices in $F^{(k)}$. Hence, Eq. (3.23) shows that $w$ can be written as an affine combination of vertices in $F^{(k)} \cup \{v\}$.

Given how $w$ was specified, we have established that $F^{(k)} \cup \{v\}$ affinely spans the set of all vertices of $T$ that assign $a_k$ to $A$. This set determines a subpolytope of $T$ that is isomorphic to $S$ by the correspondence that forgets the outcome of $A$. The vertices of $S$ obtained in this way from $F^{(k)}$ saturate the inequality over **S** determined by $b$. We have shown that it is enough to add a single non-saturating vertex, namely $v|_{\mathbf{S}}$, in order to affinely span the whole of $S$. Thus the inequality determines a face of co-dimension $1$, i.e. a facet, of $S$. □

**The curious case of $\lambda_{d_S+1} = 1$**

We now discharge the issue left not fully resolved in the proof of case I above, on why one can always avoid a situation when $\lambda_{d_S+1} = 1$. We do this by examining what such a situation implies about the original inequality being lifted.

Observe that Eqs. (3.16) and (3.17) actually arise at the level of the original scenario **S** already. Let $m$ be the outcome of $M$ used in the change from $v^{(k)}_{d_S+1}$ to $w^{(k)}$ in the proof of case I. That is, adopting the notation from Sec. 3.4.3, one has $w^{(k)} = v^{(k)}_{d_S+1}[M \mapsto m]$. At the level of **S**, one may consider a similar change $w = v_{d_S+1}[M \mapsto m]$, and expanding $w$ as an affine combination of $B_S$ gives

$$w = \sum_{i=1}^{d_S+1} \lambda_i v_i. \tag{3.24}$$

The case $\lambda_{d_S+1} = 1$ implies that

$$w = v_{d_S+1} + \sum_{i=1}^{d_S} \lambda_i v_i, \tag{3.25}$$

where $\sum_{i=1}^{d_S} \lambda_i = 0$. Under the initial inequality, i.e. multiplying the vector of coefficients $b$, this yields

$$b \cdot w \;=\; b \cdot v_{d_S+1} + \sum_{i=1}^{d_S} \lambda_i b \cdot v_i \;=\; b \cdot v_{d_S+1}, \tag{3.26}$$

where the last equality follows from the fact that all the vertices $\{v_i\}_{i=1}^{d_S}$ saturate the inequality $b \cdot p \geq 0$. In other words, the vertex $w$ attains the same value for the inequality as the vertex $v_{d_S+1}$.

Now, note that, in picking the affine basis $B_S$, the last vertex $v_{d_S+1}$ could be chosen arbitrarily to be any non-saturating vertex of the inequality. Moreover, the outcome $m$ could also be freely chosen. The fact that the measurement $M$ effectively contributes to the inequality (see Sec 3.4.3) means that there exists a vertex $v$ and an outcome $m$ of $M$ such that $b \cdot v \neq b \cdot v[M \mapsto m]$. Without loss of generality, such $v$ can always be chosen to be non-saturating, because one of $v$ or $v[M \mapsto m]$ must not saturate the inequality, as they attain different values. This guarantees that it is possible to pick a non-saturating vertex $v_{d_S+1}$ and an outcome $m$ of $M$ in such a way that the problematic $\lambda_{d_S+1} = 1$ does not occur.

In turn, if the measurement $M$ does not effectively contribute to the inequality, then $\lambda_{d_S+1} = 1$ for all choices of $v_{d_S+1}$ and outcome $m$. We illustrate such a situation through a (paradigmatic) example where the newly added measurement is compatible with all the measurements that effectively contribute to the initial inequality, and check that the case I construction indeed fails to lift it to a facet-defining inequality.

Consider the scenario given by the compatibility graph shown in Figure 10(a). The following is a facet-defining NC inequality to which measurement '4' does not effectively contribute:

$$\begin{aligned} p(010|014) + p(011|014) + p(100|014) + p(101|014) \\ + p(01|12) + p(10|12) + p(01|23) + p(10|23) - p(01|30) - p(10|30) \geq 0. \end{aligned} \quad (3.27)$$

In fact, this inequality is a lifting of the CHSH inequality (with the classical bound set to zero) over the induced sub-scenario with measurement set $\{0, 1, 2, 3\}$ through the addition of measurement '4'. Notice how the measurement '4' is being 'traced out', and so the inequality is still, in essence, the CHSH inequality:

$$\begin{aligned} I_{\mathsf{CHSH}} = p(01|01) + p(10|01) + p(01|12) + p(10|12) \\ + p(01|23) + p(10|23) - p(01|30) - p(10|30) \geq 0. \end{aligned} \quad (3.28)$$

Now, consider the extension of the scenario in Figure 10(a) to the scenario in Figure 10(b). The newly added measurement '5' is compatible with all the measurements that effectively contribute to the inequality (3.27). This indicates that, in order to lift the inequality, one must use case II. Indeed, doing so gives rise to two facet-defining inequalities, one for each outcome $a_k \in \mathcal{O}_5 = \{0, 1\}$:

$$\begin{aligned} I_{a_k} = p(10a_k|015) + p(01a_k|015) + p(10a_k|125) + p(01a_k|125) + p(10a_k|235) + p(01a_k|235) \\ - p(01a_k|305) - p(10a_k|305) \geq 0. \end{aligned}$$

Had we tried to apply case I, which would amount to tracing measurement '5' out, we would have obtained the inequality

$$I_{\mathsf{CHSH}} = \sum_k I_{a_k} \geq 0,$$

which is still, in essence, the CHSH inequality on the induced sub-scenario with measurement set $\{0, 1, 2, 3\}$. This is a valid but not a facet-defining NC inequality for the scenario in Figure 10(b).

The case just described exemplifies what happens in general: in situations that fall within the scope of case II, the construction of case I leads to a valid but non-facet-defining inequality, which can be written as a linear combination of the lifted inequalities obtained through case II.

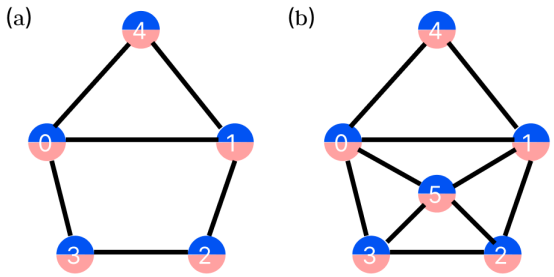



Figure 10: (a) Compatibility graph of the initial scenario with the node labelled '4' representing a measurement added to the CHSH scenario. Bi-coloured nodes represent dichotomic measurements and edges represents compatibility between measurements. (b) Compatibility graph of scenario obtained from (a) by adding a new measurement (node labelled '5') compatible with all the measurements effectively contributing to inequality (3.27).

### 3.4.4 Outcome lifting

We now move on to study lifting under outcome extensions. Let **S** be extended to **T** by the addition of a new outcome $a_0$ for an existing measurement $A \in \mathcal{X}_{\mathbf{S}}$. That is, in **S** the measurement $A$ has outcome set $\mathcal{O}_{\mathbf{S},A} = \{a_1, \ldots, a_m\}$, which is extended in **T** to $\mathcal{O}_{\mathbf{T},A} = \{a_0\} \cup \mathcal{O}_{\mathbf{S},A} = \{a_0, a_1, \ldots, a_m\}$.

The maximal contexts of **T** are the same as those of **S**. The ambient vector space of correlations for **T** has components matching those for **S** and additional ones referring to the new outcome $a_0$. Namely, for each maximal context containing $A$, $\{A\} \cup U \in \mathcal{C}_{\mathbf{S}}$, and each joint outcome for the rest of the measurements, $s \in \mathcal{O}_U$, a vector $q$ on the space of correlations for **T** has an additional component $q(a_0\, s \mid A\, U)$.

To exhibit lifting, we must first map an inequality over **S** into one over **T**. For this, we 'club together' the outcome $a_0$ with some other pre-existing outcome $a_k$ in $\mathcal{O}_{\mathbf{S},A}$, treating the new outcome as if it were $a_k$. More explicitly, let

$$b \cdot p \geq 0 \tag{3.29}$$

be the initial facet-defining inequality for **S**. By 'clubbing together' with $a_k$ we mean that wherever a probability $p(a_k\, s \mid A\, U)$ (for some $\{A\} \cup U \in \mathcal{C}_{\mathbf{S}}$ and $s \in \mathcal{O}_C$) appears in the inequality (3.29) we replace it by $q(a_0\, s \mid A\, U) + q(a_k\, s \mid A\, U)$ in the new inequality over **T**.

Let us carefully define the new inequality by giving its vector of coefficients. From the vector of coefficients $b$ of the original inequality over **S**, we construct two new vectors over the vector space of correlations for **T**:

- the vector $b_*$ is equal to the original vector $b$ padded with zeros on the new components, those

that correspond to observing outcome $a_0$ for measurement $A$ (more precisely, to joint outcome assignments for maximal contexts containing the measurement $A$ which assign outcome $a_0$ to $A$).

- the vector $b_{a_k}$ is non-zero only on the new components (events assigning outcome $a_0$ to $A$), taking the same coefficient as $b$ does on the event obtained by changing $A$'s outcome to $a_k$: that is, for each $\{A\} \cup U$ a maximal context and $s \in \mathcal{O}_U$,

$$b_{a_k}(a_0\, s \mid A\, U) \;=\; b(a_k\, s \mid A\, U),$$

and $b_{a_k}$ is zero at every other component.

The new vector of coefficients is $b_* + b_{a_k}$, yielding the following inequality over the space of correlations for **T**:

$$b_* \cdot q + b_{a_k} \cdot q \geq 0. \tag{3.30}$$

Much like a case I measurement-lifted inequality could be understood through the marginalisation map that takes each correlation $q$ on **T** to the correlation $q|_{\mathbf{S}}$ on **S**, so one way to understand this inequality is through a *coarse-graining* map from correlations on **T** to correlations on **S**. Given a correlation $q$ on **T**, define a correlation $q_{k\equiv 0}$ on **S** by 'clubbing together' outcomes $a_0$ and $a_k$ for $A$, as follows:

$$q_{k\equiv 0}(a_k\, s \mid A\, U) = q(a_k\, s \mid A\, U) + q(a_0\, s \mid A\, U)$$

for $\{A\} \cup U \in \mathcal{C}_{\mathbf{S}}$ and $s \in \mathcal{O}_U$, and $q_{k\equiv 0}(s \mid C) = q(s \mid C)$ in all other cases. Then, the lifted inequality (3.30) effectively tests for the original inequality (3.29) on the coarse-grained statistics. That is, analogously to Eq. (3.7): for all correlations $q$ on **T**,

$$b_* \cdot q + b_{a_k} \cdot q \;=\; b \cdot q_{k\equiv 0}.$$

The mapping $q \mapsto q_{k\equiv 0}$ clearly preserves deterministic non-contextual correlations, as it can be seen to work at the level of global outcome assignments. These correlations are the vertices of the noncontextual polytopes. The mapping is therefore a map of polytopes $T \to S$. This observation is the key to prove the validity of the lifted inequality (3.30), which we now explicitly state (cf. Propositions 1 and 2, the analogous statements for measurement lifting).

**Proposition 4.** *Let the scenario* **T** *be an extension of the scenario* **S** *by the addition of a single outcome* $a_0$ *to an existing measurement* $A$*. If* $b \cdot p \geq 0$ *is a valid inequality for* $S$*, then* $(b_* + b_{a_k}) \cdot q \geq 0$ *is a valid inequality for* $T$*.*

*Proof.* Let $v$ be any vertex of the polytope $T$, corresponding to an global outcome assigment $\lambda \in \mathcal{O}_{\mathcal{X}}$. If it assigns to $A$ an outcome other than $a_0$, then the second term in $b_* \cdot v + b_{a_k} \cdot v$ vanishes and the first reduces to $b \cdot v_{k\equiv 0}$ where $v_{k\equiv 0}$ is the vertex of $S$ with precisely the same underlying global outcome assignment $\lambda$ as $v$. If it assigns outcome $a_0$ to $A$, then the first term in $b_* \cdot v + b_{a_k} \cdot v$ vanishes and the second reduces to $b \cdot v_{k\equiv 0}$, where now $v_{k\equiv 0}$ is the vertex of $S$ whose underlying global outcome assignment is $\lambda[A \mapsto a_k]$, i.e. it assigns outcome $a_k$ to $A$ while retaining the same outcomes as $\lambda$ for the other measurements. The result then follows from the validity of $b \cdot p \geq 0$. □

While the inequality (3.30) is always valid for $T$, it is not always a lifting of the initial inequality (3.29), as it may fail to be facet-defining. We now describe a (general) situation in which it fails to lift.

Suppose that the initial inequality (3.29) is such that all non-zero coefficients in $b$ correspond to joint outcomes that assign the fixed outcome $a_k$ for $A$, i.e. only components of the form $b(a_k\, s \mid A\, U)$ may be non-zero. This renders it an inequality of the form (3.8), that is, $b = c_{0,a_k,a_k}$ for some vector of coefficients $c$ over the sub-scenario **R** obtained from **S** by dropping the measurement $A$. So, the initial inequality over **S** is itself a lifting from the smaller scenario **R** via case II measurement lifting. If one then chooses to club $a_0$ with $a_k$, one ends up getting that $b_* = c_{0,a_k,a_k}$ (where now the right-hand side represents a case II measurement lifting of the inequality determined by $c$ directly from **R** to **T**), and $b_{a_k} = c_{0,a_0,a_0}$ (where the right-hand side also represents a case II measurement lifting directly from **R** to **T**). Consequently, the candidate lifted inequality (3.30) would read

$$c_{0,a_k,a_k} \cdot q + c_{0,a_0,a_0} \cdot q \geq 0.$$

This is clearly not a facet-defining inequality for $T$, since by Theorem 3 each of the two summands is itself a facet-defining inequality, obtained via case II measurement lifting from **R** to **T**. In our outcome lifting result, we must therefore exclude such cases. Note that such an inequality determined by $b = c_{0,a_k,a_k}$ (case II measurement lifted from **R** to **S**) can be lifted to **T** if we choose to club $a_0$ with a different outcome $a_j \neq a_k$. In that case, however, the vector $b_{a_j}$ is equal to zero and the lifted inequality is simply $c_{0,a_k,a_k} \cdot q \geq 0$, which can also be obtained via case II measurement lifting directly from **R** to **T**.

With the proviso of excluding the situation discussed in the previous paragraph, we now prove the facet-defining nature of the noncontextuality inequality (3.29) in all other situations (cf. Theorem 3 for measurement lifting).

**Theorem 5.** *Let the scenario* **T** *be an extension of the scenario* **S** *by the addition of a single outcome* $a_0$ *to an existing measurement* $A$*. If* $b \cdot p \geq 0$ *is a facet-defining NC inequality for* **S***, then for any pre-existing*

*outcome $a_k$ of $A$ for which $b \neq c_{0,a_k,a_k}$, the inequality $(b_* + b_{a_k}) \cdot q \geq 0$ is a facet-defining NC inequality for* **T**.

*Remark.* The condition that $b$ not be of the form $c_{0,a_k,a_k}$, i.e. not be obtainable by case II measurement lifting when adding measurement $A$ with choice of outcome $a_k$, admits an equivalent formulation in terms of the properties of the inequality being lifted (or of the facet it defines). The case being excluded is when the measurement $A$ is compatible with all the measurements effectively contributing to the inequality $b \cdot p \geq 0$ *and*, moreover, all the vertices of $S$ that do not saturate the inequality assign outcome $a_k$ to $A$ (i.e. $b \cdot v = 0$ for all $v$ not assigning $a_k$ to $A$).

*Proof.* As in the proof for measurement lifting (Theorem 5), the strategy is to use an affine basis for $S$ to define a set of vertices spanning $T$, from which one can affinely eliminate all but one vertex not saturating the lifted inequality. We start by picking a set of affinely independent vertices of $S$,

$$B_S = \{v_1, \ldots, v_{d_s+1}\},$$

in such a way that the first $d_S$ vectors saturate the initial facet-defining inequality $b \cdot p \geq 0$. Recall that the last vertex, $v_{d_s+1}$, may be chosen arbitrarily from among the vertices of $S$ not saturating the inequality.

We now form two sets of vertices of $T$ from $B_S$:

- $V^* = \{v_1^*, \ldots, v_{d_S+1}^*\}$, where each $v_i^*$ is the vertex of $T$ corresponding to the same underlying global outcome assignment as $v_i$ (note that even though $v_i$ and $v_i^*$ have the same underlying outcome assignment, they are different vectors living in vector spaces with different dimensions: $v_i^*$ is obtained from $v_i$ by padding it with zeroes in the new components);
- $V' = \{v_1', \ldots, v_{d_S+1}'\}$, where each $v_i'$ is the vertex of $T$ whose underlying assignment is obtained from that of $v_i$ by substituting the outcome $a_0$ for the outcome $a_k$ if the latter is the outcome that $v_i$ assigns to $A$

Observe that if $v_i$ assigns an outcome other than $a_k$ to $A$, then $v_i' = v_i^*$.

The union of these two sets $B_T = V^* \cup V'$ affinely spans the whole polytope $T$, i.e. it is a (potentially over-complete) affine basis for $T$. Both $v_i^*$ and $v_i'$ are mapped to $v_i$ under the coarse-graining $T \to S$, i.e. $(v_i^*)_{k \equiv 0} = (v_i')_{k \equiv 0} = v_i$. Consequently, the last vertex in each of $V^*$ and $V'$ is the only one that does not saturate the inequality $(b_* + b_{a_k}) \cdot q \geq 0$.

We divide the proof into two cases:

1. the non-saturating vertex $v_{d_S+1}$ in $B_S$ can be picked so that it does not assign outcome $a_k$ to $A$.

2. every vertex of $S$ not saturating the initial inequality assigns outcome $a_k$ to $A$.

The first case is quickly discharged. We have $v^*_{d_S+1} = v'_{d_S+1}$, i.e. the last vertex within $V^*$ and the last vertex within $V'$ are exactly the same. Hence, $B_T$ contains only one vertex not saturating the inequality. It must therefore contain $d_T$ affinely independent points that saturate it, proving that it defines a facet of $T$.

In the second case, the last vertices of $V^*$ and $V'$ are different, differing only on the outcome each assigns to the measurement $A$, respectively $a_k$ and $a_0$. We show that $v'_{d_S+1}$, the last element of $V'$, can be eliminated as an affine combination of the remaining vertices in $B_T$, using a similar approach as for measurement lifting.

By the remark below the theorem statement, in this case, there exists a measurement $M$ which is incompatible with $A$ and which effectively contributes to the initial inequality. Let $w$ be a vertex of $S$ obtained from $v_{d_S+1}$ by changing the outcome of $M$ to a different value. It can be expanded in the affine basis $B_S$ as

$$w \;=\; \sum_{i=1}^{d_S+1} \lambda_i v_i, \tag{3.31}$$

with $\sum_{i=1}^{d_S+1} \lambda_i = 1$. For the same reason discussed in Sec. 3.4.3, the vertex $w$ (more precisely, the outcome it assigns to $M$) may be chosen so that $\lambda_{d_S+1} \neq 1$. In that case, we would have

$$w \;=\; v_{d_S+1} + \sum_{i=1}^{d_S} \lambda_i v_i, \tag{3.32}$$

with $\sum_{i=1}^{d_S} \lambda_i = 0$, and applying $b$,

$$b \cdot w \;=\; b \cdot v_{d_S+1} + \sum_{i=1}^{d_S} b \cdot \lambda_i v_i \;=\; b \cdot v_{d_S+1}, \tag{3.33}$$

meaning that $b$ does not distinguish between the different outcomes that $v_{d_S+1}$ and $w$ assign to $M$. This can be avoided because $M$ effectively contributes to the inequality.

Now, let $w^*$ and $w'$ be the vertices of $T$ built from $w$ like $v^*_i$ and $v'_i$ are built from $v_i$, respectively by considering the same global outcome assignment and the one where $A$'s outcome is changed to $a_0$. Analogously to Eq. (3.11) in the measurement lifting proof, we have the following equality

$$w^* - v^*_{d_S+1} = w' - v'_{d_S+1}. \tag{3.34}$$

To see this explicitly, divide the maximal contexts in $\mathcal{C}_\mathbf{T}$ into two mutually exclusive subsets – those without $A$ and those with $A$ – and consider the corresponding direct sum decomposition of the vector space into two summands. The vectors $v^*_{d_S+1}$ and $v'_{d_S+1}$ (similarly, $w^*$ and $w'$) only differ in the second summand,

as their underlying global assignments differ only for $A$. On the other hand, since $M$ is incompatible with $A$, it only appears in maximal contexts not containing $A$, and so $v^*_{d_S+1}$ and $w^*$ (similarly, $v'_{d_S+1}$ and $w'$) only differ in the first summand. In summary, we can write these four vectors as

$$v^*_{d_S+1} = [u_{m_0}, u_{a_k}] \qquad w^* = [u_{m_1}, u_{a_k}]$$
$$v'_{d_S+1} = [u_{m_0}, u_{a_0}] \qquad w' = [u_{m_1}, u_{a_0}]$$

for some $u_{m_0}, u_{m_1}$ in the first summand and $u_{a_0}, u_{a_k}$ in the second, where the indices are suggestive of the differing outcomes for measurements $M$ and $A$. Eq. (3.34) is now evident from inspection.

Again as before, the vertices $w^*$ and $w'$ are affine combinations of those in $V^*$ and $V'$, respectively, with the same coefficients as in Eq. (3.31):

$$w^* = \sum_{i=1}^{d_S+1} \lambda_i v^*_i, \qquad w' = \sum_{i=1}^{d_S+1} \lambda_i v'_i,$$

Putting these expansions into Eq. (3.34) yields

$$\sum_{i=1}^{d_S} \lambda_i v^*_i + (\lambda_{d_S+1} - 1) v^*_{d_S+1} = \sum_{i=1}^{d_S} \lambda_i v'_i + (\lambda_{d_S+1} - 1) v'_{d_S+1},$$

and given that $\lambda_{d_S+1} \neq 1$,

$$v'_{d_S+1} = v^*_{d_S+1} - \sum_{i=1}^{d_S} \frac{\lambda_i v^*_i}{1 - \lambda_{d_S+1}} + \sum_{i=1}^{d_S} \frac{\lambda_i v'_i}{1 - \lambda_{d_S+1}}.$$

This exhibits $v'_{d_S+1}$ as an affine combination of the remaining elements of $B_T$, showing it is redundant in $B_T$. In other words, the set $B_T \setminus \{v'_{d_S+1}\}$ still affinely spans $T$. But it now contains only one vertex that does not saturate the inequality $(b_* + b_{a_k}) \cdot q \geq 0$. Necessarily, then, among $d_T + 1$ affinely independent vertices in this set, $d_T$ of them must saturate the inequality, showing that it supports a facet of $T$. □

## 3.5 Applications

To demonstrate its power, we apply our lifting technique to extract facet-defining noncontextuality inequalities of scenarios for which no such facet inequalities had been hitherto described. To do so, we pick two well-known inequalities, the Clauser–Horne–Shimony–Holt (CHSH) and the Klyachko–Can–Binicioğlu–Shumovsky (KCBS) inequalities, as our starting points. In Sec. 3.5.1, we sequentially measurement lift the former to extract a facet-defining NC inequality for the scenario described by the anti-heptagon $\bar{C}_7$, the complement of the 7-cycle, with binary outcomes. In Sec. 3.5.2, we sequentially outcome lift the latter to obtain a facet-defining NC inequality for the 5-cycle scenario with three outcomes per measurement.

### 3.5.1 Lifting the CHSH inequality

In our first example, we apply sequential measurement lifting to the CHSH inequality, defined on the $4$-cycle scenario, and derive a facet-defining inequality for the scenario with compatibility graph $\bar{C}_7$, the complement of the 7-cycle, shown in Figure 11. All measurements are dichotomic in this example, taking values in the outcome set $\{0, 1\}$.

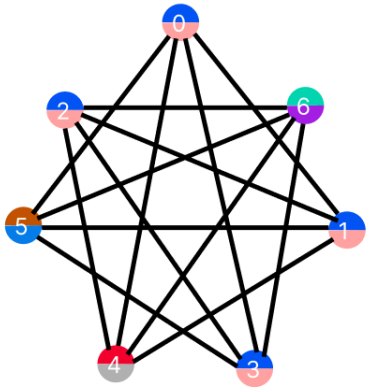



Figure 11: Compatibility graph of the antiheptagon ($\bar{C}_7$) scenario with bi-coloured nodes representing dichotomic measurements. A facet-defining NC inequality is obtained for this scenario via sequential measurement lifting of the CHSH inequality from the sub-scenario on measurements $\{0, 1, 2, 3\}$.

We start from the sub-scenario given by the induced subgraph of $\bar{C}_7$ with node set $\{0, 1, 2, 3\}$ as per Figure 11, a $4$-cycle. The initial facet-defining inequality (CHSH) is

$$p(00|01) + p(11|01) + p(00|12) + p(11|12) + p(00|23) + p(11|23) + p(01|30) + p(10|30) \leq 3. \quad (3.35)$$

We now add measurement '4' to this sub-scenario, with the compatibilities induced from $\bar{C}_7$. This leads to a new scenario with maximal contexts $\{0, 1, 4\}$, $\{1, 2, 4\}$, $\{2, 3\}$, $\{0, 3\}$. The way in which the new measurement is affixed to the initial scenario, namely the fact that it is not compatible with measurement '3', means that case I of measurement lifting applies. Hence, we simply need to trace out the added measurement. Inequality (3.35) then turns out to be facet-defining for this scenario as well. Note that the term $p(00|01)$ can be written as $p(000|014)+p(001|014)$ in terms of probabilities over maximal contexts of the extended scenario. We then similarly add measurement '5'. This yields a new scenario with maximal contexts $\{0, 1, 4\}$, $\{1, 2, 4\}$, $\{2, 3\}$, $\{0, 3, 5\}$, $\{0, 1, 5\}$. Again, case I of measurement lifting applies since the added '5' is incompatible with measurement '2', and the inequality (3.35) remains facet-defining for this scenario. Finally, to get to the anti-heptagon, we add measurement '6', arriving at the scenario

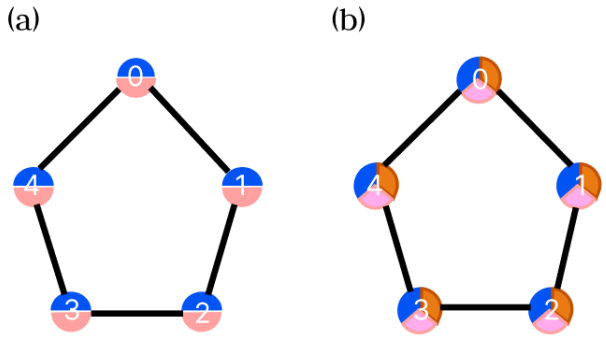



Figure 12: (a) Compatibility graph of the initial KCBS $5$-cycle scenario with five dichotomic measurements, represented by bi-coloured nodes. (b) Compatibility graph of the $5$-cycle scenario with trichotomic measurements, represented by tri-coloured nodes. Inequality (3.38) is a facet-defining NC inequality for this extended scenario, obtained via sequential outcome lifting of the KCBS inequality.

with compatibility graph $\bar{C}_7$. Its maximal contexts are $\{0,1,4\}$, $\{1,2,4\}$, $\{0,3,5\}$, $\{0,1,5\}$, $\{2,3,6\}$, $\{2,4,6\}$, $\{3,5,6\}$. Once more, case I applies since the added '6' is incompatible with measurements '0' and '1', and therefore inequality (3.35) is facet-defining for the anti-heptagon scenario. Note that the expression in (3.35) can of course be expanded in terms of probabilities over maximal contexts of the anti-heptagon scenario. It is the first-ever identified facet-defining NC inequality for this scenario.

Note that it is relatively straightforward to see that (3.35) must be a valid inequality for the scenario $\bar{C}_7$, as for any other scenario that contains the initial $4$-cycle as a sub-scenario (even if not an induced one). The crucial point is that, in this case, it supports a facet and not a lower-dimensional face.

### 3.5.2 Lifting the KCBS inequality

In the second example, we apply sequential outcome lifting to the KCBS inequality, defined on the two-outcome $5$-cycle scenario, and obtain a facet-defining NC inequality for the scenario with the same compatibility graph, $C_5$, but now with each measurement having three allowed outcomes; see Figure 12. The starting facet-defining inequality (KCBS) is

$$p(01|01) + p(01|12) + p(01|23) + p(01|34) + p(01|40) \leq 2. \tag{3.36}$$

We first add a new outcome, labelled 2, to measurement '0'. We choose to club this new outcome with the pre-existing outcome labelled 1. Then according to our lifting method, in the inequality (3.36), where outcome 1 for measurement '0' appears in the context $\{0,4\}$, the term $p(01|40)$ has to be replaced by

$p(01|40) + p(02|40)$ to obtain the new facet-defining inequality:

$$p(01|01) \; + \; p(01|12) \; + \; p(01|23) \; + \; p(01|34) \; + \; p(01|40) \; + \; \boldsymbol{p(02|40)} \quad \leq \quad 2. \quad (3.37)$$

Likewise, we add a new outcome, labelled 2, to measurement '1', and club it with its old outcome labelled 1. Tweaking the term $p(01|01)$ in the inequality (3.37) as before, we obtain a new facet-defining inequality:

$$p(01|01) + \boldsymbol{p(02|01)} + p(01|12) + p(01|23) + p(01|34) + p(01|40) + p(02|40) \; \leq \; 2.$$

We continue like this by adding a new outcome, labelled 2, to the remaining measurements '2', '3', and '4', and clubbing it with the respective outcome labelled 1. In the end, we arrive at the following facet-defining inequality for the trichotomic $5$-cycle scenario:

$$p(01|01) + \boldsymbol{p(02|01)} + p(01|12) + \boldsymbol{p(02|12)} + p(01|23) + \boldsymbol{p(02|23)} + p(01|34) + \boldsymbol{p(02|34)} + p(01|40) + \boldsymbol{p(02|40)} \leq 2, \quad (3.38)$$

where the new terms not present in the initial inequality are highlighted.

There was an element of choice in the way the above lifting of the KCBS inequality was obtained. We could just as well have chosen to club the new outcome with the outcome labelled 0 rather than that labelled 1. Indeed, we could have made a different choice for each measurement. Each such sequence of choices determines a different facet-defining inequality for the final scenario, each corresponding to a different coarse-graining of outcomes. Thus, our method yields not one but several new facet-defining inequalities for the three-outcome $5$-cycle scenario as sequential outcome liftings of the KCBS inequality. We conclude this section with some more general remarks about applying our lifting technique.

### 3.5.3 Further remarks

Any contextuality-witnessing scenario must contain an $n$-cycle (with $n \geq 4$) as an induced sub-scenario. This follows from Vorob'ev's theorem [11]. Combining this fact with the results of Ref. [12], which completely characterises the facets for dichotomic cycle scenarios, leads us to conclude that we can provide via lifting many facet-defining NC inequalities for any scenario that can witness contextuality. Notice that, in the two examples detailed above, we did in fact lift two cycle inequalities, respectively CHSH and KCBS.

We also highlight an observation about applying our technique to sequentially lift an inequality from an initial scenario **S** to a final scenario **T**. When measurement lifting through the addition of two measurements, say $A$ and $B$, the lifted inequality might depend on whether one first adds $A$ and then $B$ or vice versa. This means that, in order to obtain all the facet-defining inequalities of $T$ that are lifted from $S$, one must consider all the possible orderings to reach scenario **T** by extending **S**.

## 3.6 Discussion

In this chapter we presented the extension of the method introduced by Pironio for lifting facet-defining Bell inequalities to arbitrary contextuality scenarios described by compatibility graphs. The method allows us to produce facet-defining noncontextuality inequalities in all scenarios that admit contextual correlations. For most of these scenarios, no such inequalities had been previously described.

Our results have been presented in the form of two main lifting results titled 'measurement lifting', and 'outcome lifting'. The former involved two scenarios where one is an arbitrary extension of the other via the addition of a measurement, whereas the latter involved liftings across two scenarios where one is an arbitrary extension of the other via an addition of outcome to some existing measurement. An arbitrary extension can then be seen as a sequential application of these two forms of liftings. Some nuanced remarks on this are made in Sec. 3.5.3.

Apart from the main results, other noteworthy highlights of this work are the following: the new inequality, in any case, is a mere re-expression of the old inequality in the language of the bigger polytope, thus keeping the classical, quantum and no-disturbance bounds unchanged; the methods of lifting depend upon how the scenario is extended with respect to the measurements that are effectively contributing in a given non-contextuality inequality of the initial scenario. Moreover, the same factor determines whether a facet-defining non-contextuality inequality of the smaller polytope lifts to one or multiple such inequalities of the bigger polytope.

Later in Chapter 6, we will discuss the potential future directions cropping out from the contributions in this work.

# Chapter 4

# Contextuality with Pauli observables in cycle scenarios

**Chapter Summary.** KS contextuality is known to be a resource for quantum computational advantage in a variety of settings. Such results often boil down to contextuality relative to multi-qubit Pauli observables. Spurred by this observation, in this chapter, we initiate a systematic study of (state-dependent) Pauli contextuality, by focusing on $n$-cycle scenarios, the most fundamental KS scenarios. We consider the realizability of these scenarios by $m$-qubit Pauli observables, and the potential for witnessing contextuality therein. First, we show that the maximum size of a cycle realizable strictly by $m$-qubit Pauli observables is upper bounded by $3m$. We then provide some constructions for which the size of faithfully realizable cycles, by $m$-qubit Pauli operators, varies linearly as a function of $m$. We then prove that no $n$-cycle Pauli realization for $n > 4$ can witness contextuality (on any quantum state), whereas for $n = 4$ every Pauli realization exhibits contextuality, attaining the quantum bound for all noncontextuality inequalities on some pure state. Finally, we discuss the effect of cycles on the presence of contextuality in arbitrary Pauli scenarios, in light of Vorob'ev's theorem.

**Motivation.** It is natural to expect the quintessentially non-classical features of quantum theory to play a central role as a resource for quantum computational advantage. Indeed, numerous such links have been unveiled recently. And many of these results boil down to the presence of contextuality witnessed with Pauli observables. For example, Howard et al. [6] showed that contextuality is a necessary resource for universal quantum computation via magic state distillation, while Raussendorf [7] proved that the measurement-based quantum computation model (with $\mathbb{Z}_2$-linear classical control) cannot compute nonlinear Boolean functions without the presence of contextuality. Both of these results involve contextuality with respect to (qudit or qubit) Pauli observables only. This is perhaps unsurprising. After all, Pauli measurements play a central role in many models of quantum computation, in both near-term and fault-tolerant settings.

Pauli-based computation (PBC), a recently proposed quantum universal model, involves sequential measurements of adaptively chosen commuting Pauli observables on an initially prepared magic state [74]. On the other hand, hybrid algorithms like variational quantum eigensolvers (VQEs), seen as strong candidates for witnessing practical quantum advantage in NISQ devices, also involve measurements of Pauli observables in each iteration. Focusing on the measurement part of VQE, Kirby and Love [16] defined a (state-independent) contextuality test on VQE instances, and subsequently [17] showed that the classical simulation of noncontextual VQEs is only an NP-complete problem, i.e. it is classically hard, as opposed to being quantumly hard as represented by the QMA complexity class.

The contextuality test for VQE instances proposed in [16] concerns precisely such state-independent contextuality. The instances considered in that work are given by Hamiltonians expressed in Pauli decomposition, i.e. as a linear combination of operators from the $m$-qubit Pauli group $\mathcal{P}_m$. Starting from the set $S \subseteq \mathcal{P}_m$ of Paulis appearing in such a decomposition, it considers its closure $\bar{S}$ under products of commuting elements.[1]The instance is deemed to be contextual if $\bar{S}$ exhibits state-independent strong contextuality (like the above-mentioned square and star). The authors also provide a necessary and sufficient condition on the set $S$ for this to happen.

The initial motivation for the present work comes from noticing the following caveat in this analysis. The VQE algorithm performs only the measurements in $S$, not those in $\bar{S}$. As such, in a classical simulation, it suffices to reproduce the statistics of these measurements on certain quantum states, regardless of whether they could be consistently extended to encompass $\bar{S}$. A more refined classicality analysis would thus focus on the (weak) contextuality of the correlations realizable by the measurements of $S$ alone on some quantum state. While strong contextuality relative to $m$-qubit Pauli observables has been extensively studied – from the state-independent proofs mentioned above to state-dependent arguments such as GHZ, or more generally, All-versus-Nothing arguments [75, 76] – the *weaker* notion of probabilistic contextuality has received comparatively little attention, even though it does play a role in many of the applications.
In this work, we initiate a systematic study of the following question, prompted by the discussion above: *given a set $S$ of $m$-qubit Pauli operators, when does it directly witness contextuality?*

One way to approach this question is to take the compatibility graph of Pauli operators in $S$ and treat it as an abstract measurement scenario. One can then derive all the noncontextuality inequalities for this scenario and test whether the given Pauli realisation can violate any of them. A major bottleneck, however, is that finding all the noncontextuality inequalities for an arbitrary scenario encoded as a compatibility graph is an NP-complete problem. Hence, as a way to get a handle on the general problem, we start

[1] $\bar{S}$ is the partial Abelian group (in the same sense of partial Boolean algebra) generated by $S$ *within* $\mathcal{P}_m$; we refer to upcoming work for more about this perspective.

by considering the most fundamental KS scenarios, the $n$-cycle scenarios, which are explained in Sec. 2.3.

**Outline of the Chapter.** We begin by providing an overview of our analyses in Sec. 4.1. We then introduce some mathematical background and definitions needed to follow the rest of the chapter in Sec. 4.2. In Sec. 4.3, we highlight the constraints on the faithful qubit Pauli cycle realizations. We then exploit these constraints to prove our no-go theorem in Sec. 4.4. In Sec. 4.5, we provide constructions that guarantee qubit Pauli realizations in accordance with the no-go theorem. Next, in Sec. 4.6, we present our complete characterization of qubit Pauli contextuality for the most fundamental scenarios. Using this characterization, we then move on to qubit Pauli realizations of arbitrary scenarios in Sec. 4.7. Finally, with Sec. 4.8, we close the chapter.

## 4.1 Overview

We now present a concise overview of our analysis of KS contextuality from the purview of multi-qubit Pauli operators.

To fully characterize the contextuality of $m$-qubit Pauli operators in $n$-cycle scenarios, we first analyze the faithful realizability of these scenarios by such operators. In doing so, we prove a no-go theorem that provides the cycle size, as a function of $m$, above which no cycle can be faithfully realized by $m$-qubit Pauli operators. The proof of this theorem rests on our clever introduction of new Pauli operators, called 'edge Paulis', defined more precisely in the next section, one for each edge of a given $m$-qubit Pauli realization of a cycle scenario, see Figure 15. Running subgraph isomorphism tests then inform us that for a given $m$, all Pauli cycle realizations rendered possible by our no-go theorem aren't always realizable. This shifts our focus towards identifying the possible faithful realizations.

Consequently, we provide constructions that guarantee faithful Pauli realization of all cycles up to a size that is a linear function of $m$. In providing these constructions, we prove supporting results on faithful realizations of path graphs, by $m$-qubit Pauli operators, and their conversion to faithful realizations of cycle scenarios. We then move towards the contextuality tests of the Pauli realizable cycle scenarios. Our introduction of edge Pauli operators turns out to be immensely useful in fully characterizing contextuality of qubit Pauli operators for $n$-cycle scenarios. This is due to two reasons (i) The form of the $n$-cycle non-contextuality inequalities is such that their quantum version can be written only terms of the edge-Pauli operators, i.e. as a sum of expectation values of the edge Pauli operators (ii) The compatibility constraints imposed by the edge Pauli operators simplify a lot of calculations. We then leverage this complete char-

acterization and Vorob'ev's theorem to show that an induced 4-cycle subscenario is sufficient, but not necessary, for an arbitrary $m$-qubit Pauli realization of an arbitrary KS scenario to exhibit contextuality.

## 4.2 Preliminaries

**Pauli group.** An $m$-qubit Pauli operator means an operator composed of tensor product of $m$ single qubit operators: the identity operator and/or the following Pauli operators:

$$X \equiv \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, Y \equiv \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, Z \equiv \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

The Pauli group on $M$ qubits (under matrix multiplication) is the set

$$\mathcal{P}_{\mathbf{m}} = \left\{ \begin{pmatrix} \pm 1 \\ \pm i \end{pmatrix} \sigma_{\alpha_1} \otimes \ldots \otimes \sigma_{\alpha_m} \right\} \tag{4.1}$$

where $\sigma_{\alpha_j} \in \{X, Y, Z, I\} \; \forall \; j$. Because we are concerned with Hermitian qubit Pauli operators, we can only make use of $\pm 1$ multiples of $m$-qubit Pauli operators in $\mathcal{P}_{\mathbf{m}}$. Among the many nice properties of operators in $\mathcal{P}_{\mathbf{m}}$, there are a few that will be consistently used in this chapter:

- For any two Pauli operators $P, Q \in \mathcal{P}_{\mathbf{m}}$, either $[P, Q] = 0$ (commute) or $\{P, Q\} = 0$ (anti-commute).
- The square of any Hermitian Pauli operator is Identity, i.e. $\forall \; P \in \mathcal{P}_{\mathbf{m}}$ such that $P^\dagger = P$, $P^2 = I$.
- The product of two Hermitian Pauli operators is a Hermitian Pauli operator. More precisely for $P, Q \in \mathcal{P}_{\mathbf{m}}$, if $P^\dagger = P$, $Q^\dagger = Q$, and $[P, Q] = 0$, then $(PQ)^\dagger = PQ$ where $PQ \in \mathcal{P}_{\mathbf{m}}$.
- Any maximal stabilizer[2] subgroup of $\mathcal{P}_{\mathbf{m}}$ is generated by $m$ pairwise commuting Pauli operators.

More properties of the Pauli group along with the proofs of the above-mentioned ones can be found in Ref. [77].

**Faithful realization.** A $m$-qubit Pauli realization of a cycle scenario is called faithful iff there is a bijective map from the nodes of the cycle scenario to the $m$-qubit Paulis such that the Pauli operators have the exact same compatibility relations as in the cycle scenario, see Figure 13.

**Cycle Pauli.** For a given $m$-qubit Pauli realization of an $n$-cycle scenario, each Pauli corresponding to a node in the cycle scenario is called a cycle Pauli, see Figure 13.

[2] An abelian subgroup of $\mathcal{P}_{\mathbf{m}}$ that does not have $-I$ is a stabilizer subgroup.

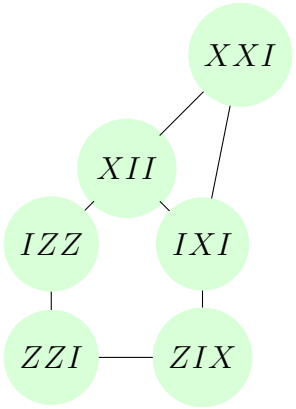



Figure 13: A faithful 3-qubit Pauli realization of a 5-cycle scenario: each Pauli corresponding to the node of the 5-cycle is a 'cycle Pauli' while the operator $XXI$ being the product of two commuting cycle Paulis, $XII$ and $IXI$, is an 'edge Pauli'.

**Edge Pauli.** A Pauli that is a product of two commuting cycle Paulis is called an edge Pauli, see Figure 13.

In the rest of the text, given an arbitrary $n$-cycle scenario, the cycle Pauli for the i$^{\text{th}}$ node will be represented by $P_i$ while the i$^{\text{th}}$ edge Pauli by $L_i = P_i P_{i\oplus 1}$.

## 4.3 Constraints on faithful realizations

Before discussing our results on qubit Pauli contextuality and faithful realization of cycle scenarios by Pauli operators, we describe the structural constraints enforced on any faithful $m$-qubit Pauli realization of a cycle scenario. These constraints are a consequence of the above-mentioned properties of the operators in the Pauli group $\mathcal{P}_{\mathbf{m}}$ as well as the faithfulness condition on the realization. Note that these constraints drive all the analyses to be followed later on in the chapter.

We divide up these constraints into two types (i) Forbidden edge Paulis (ii) Commutativity constraints. For a given $m$-qubit Pauli realization for a cycle scenario, the first constraint specifies the Pauli operators that can never be edge Pauli operators. While the second constraint defines the commutativity relations within the set of edge Paulis and the ones across the edge and cycle Pauli operators.

### 4.3.1 Forbidden edge Paulis

The following two are the forbidden properties of edge Pauli operators, given a faithful Pauli realization of a cycle scenario:

1. An edge Pauli $L_i = P_i P_{i\oplus 1}$ cannot be a multiple of any of the cycle Paulis constituting it, i.e.

$L_i \neq kP_i$ (or $L_i \neq kP_{i\oplus 1}$). This is because if $L_i = kP_i$, then $P_{i\oplus 1} = kI$. This would mean that $P_{i\oplus 1}$ will have to commute with all cycle Paulis in the cycle, which is forbidden because of the faithfulness condition.

2. An edge Pauli $L_i = P_i P_{i\oplus 1}$ cannot be a multiple of any of the cycle Paulis not constituting it, i.e. $L_i \neq kP_l$ where $l \in \{0, 1, 2, 3..., n-1\} \setminus \{i, i \oplus 1\}$. This is because $L_i = kP_l$ along with the fact that $[L_i, P_i] = 0$ implies that $[P_i, P_l] = 0$. In other words, the cycle Pauli $P_i$ must commute with a non-neighboring Pauli $P_l$. This is forbidden because of the faithfulness condition.

In conclusion, for a given $m$-qubit Pauli realization of a cycle scenario, an edge Pauli cannot be a multiple of any cycle Pauli.

### 4.3.2 Commutativity conditions

The following commutativity relations prove to be very fruitful for the results to follow later. First three conditions define the commutative properties within the set of edge Pauli operators while the last one defines the commutation relation between edge and cycle Pauli operators:

1. No edge Pauli commutes with its nearest neighboring edge Paulis, i.e. $\{L_i, L_{i\oplus 1}\} = 0$ and $\{L_i, L_{i\oplus(n-1)}\} = 0$.

2. No edge Pauli commutes with its next-nearest edge Pauli operator except for when $n = 4$, i.e. for all $n > 5$ cycle Pauli realizations $\{L_i, L_{i\oplus 2}\} = 0$ and $\{L_i, L_{i\oplus(n-2)}\} = 0$.

3. Other than its nearest and next-nearest edge Pauli, an edge Pauli commutes with all other edge Paulis, i.e. $[L_i, L_j] = 0$ only when $i \oplus 3 \leq j \leq i \oplus (n-3)$.

4. An edge Pauli commutes with all cycle Paulis except for the neighbours of the cycle Paulis constituting it, i.e. $[L_i, P_l] = 0$ only when $l \neq i \oplus (n-1)$ and $i \oplus 2$.

With these results we are now well equipped to begin our analysis involving faithful multi-qubit Pauli realizations of cycle scenarios.

## 4.4 Impossible faithful realizations

Using the above-mentioned constraints, we can derive upper bounds on the size of the cycle scenario that is faithfully realizable by $m$-qubit Pauli operators. This question is motivated by our computational

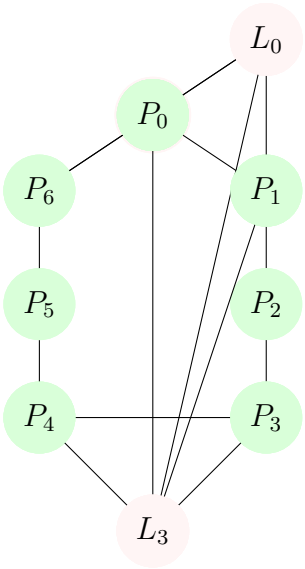



Figure 14: In a faithful 2-qubit Pauli realization of a 7-cycle (in green) we introduce two edge Paulis (in pink) $L_0 = P_0P_1$ ; $L_3 = P_3P_4$ and some more edges allowed by the four conditions obtained in Sec.4.3.2.

searches where we ran subgraph isomorphism tests to see which cycles are realizable by 2,3-qubit Pauli operators. We found that all cycles from size 4 to 6 are realizable by 2-qubit Paulis, whereas for 3-qubit Paulis all cycles from size 4 to 9, except, surprisingly, the 8-cycle are realizable. We now explain as to why a 7-cycle (or bigger) realization by 2-qubit Pauli operators is impossible. Figure 14 will serve as an illustration for this proof.

**Impossibility of realizing a 7-cycle scenario with 2-qubit Paulis.** Assume that the cycle Paulis in Figure 14 are 2-qubit Pauli operators. We now use the fact that a maximal stabilizer subgroup for $m$ qubits has $m$ independent generators to show that such a cycle is impossible to realize. For that consider two independent commuting Paulis $\{P_0,\ L_0\}$ and the maximal subgroup, say $\mathcal{R}$, generated by them i.e. $\mathcal{R} \equiv \langle P_0, L_0 \rangle$. Since $L_3$ commutes with the generators in this maximal subgroup, hence $L_3 \in \mathcal{R}$ i.e. $L_3 = {P_0}^{\alpha}{L_0}^{\beta}$ where $\alpha, \beta \in \{0, 1\}$. Consider the following possibilities:

- $\alpha = 0, \beta = 1 \implies L_3 = L_0$: This is forbidden because $[L_3, L_4] \neq 0$, but $[L_0, L_4] = 0$. These relations follow from conditions 2 and 3 respectively in Sec. 4.3.2.
- $\alpha = 1, \beta = 0 \implies L_3 = P_0$: This is forbidden because if it is true, then $P_1 = I$ which breaks the faithfulness of the realization.
- $\alpha = 0, \beta = 0 \implies L_3 = I$: This is forbidden because if it is true, then $P_3 = P_4$ which breaks the faithfulness of the realization.

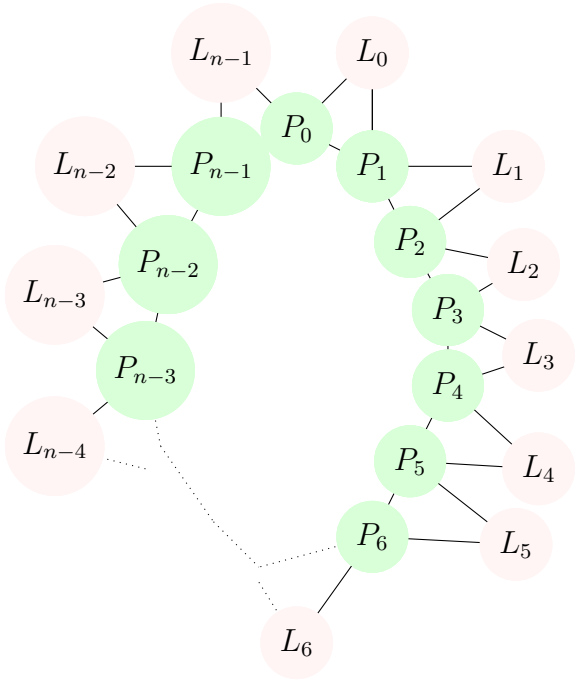



Figure 15: An arbitrary, faithful $m$-qubit Pauli realization of an $n$-cycle (green) with the corresponding edge Paulis $L_i = P_i P_{i\oplus 1}$ (pink), the dotted lines refers to the Paulis (including edge Paulis) not drawn but are part of the cycle.

- $\alpha = 1, \beta = 1 \implies L_3 = P_0 L_0 = P_0 P_0 P_1 = P_1$: This is forbidden since it enforces the edges $(P_3, P_1)$ and $(P_4, P_1)$, hence violating the faithfulness of the realization.

This exhausts all the possibilities and hence such a cycle realization is prohibited. It can be trivially seen that the same argument will hold for any two-qubit Pauli realization of an $n$-cycle, $\forall\ n \geq 7$. We now generalize this result in the form of the following theorem.

**Theorem 6.** *For a given $m$, there exists no faithful $m$-qubit Pauli realization of an $n$-cycle scenario such that $n > 3m$.*

*Proof.* Consider an arbitrary $m$-qubit Pauli realization of an $n$-cycle as shown in Figure 15, including the edge Pauli operators. Now, collect together the following Pauli operators:

$$S \equiv \{P_0, L_0, L_3, L_6, \ldots, L_{n-3}\}$$

. By the constraints mentioned in Sec. 4.3.2, all Paulis in $S$ pairwise commute. More precisely, this is because:

1. The edge Paulis in $S$ all pairwise commute due to condition 3 in Sec. 4.3.2.

2. Each $L_j \in S$ commutes with $P_0$ via the last condition highlighted in Sec. 4.3.2.

We now prove that a certain subset of the Pauli operators in $S$ are also independent[3]: First pick all the elements up to $L_i$ starting from $P_0$ from the realization (in the clockwise sense of Figure 15) from the set $S$ i.e. $S_i \equiv \{P_0, L_0, L_3, L_6, \ldots, L_{i-3}, L_i\}$. Consider the edge Pauli $L_{i+1}$. Via conditions 3 and 4 in Sec. 4.3.2- $L_{i+1}$ commutes with every element in $S_{i-1} = \{P_0, L_0, L_3, L_6, \ldots, L_{i-3}\}$. By construction $L_i$ also commutes with every element in $S_{i-1}$. Additionally, via condition 1 in Sec. 4.3.2, $[L_i, L_{i+1}] \neq 0$. Therefore, $L_i \notin \langle P_0, L_0, L_3, L_6, \ldots, L_{i-3}\rangle$. This is because, if it did then $[L_i, L_{i+1}] = 0$ which we know can not hold true (via condition 1 of Sec. 4.3.2). The above argument holds for any $L_i \in S$, except for the last Pauli operator i.e. $L_{n-3}$ . For this case, the Pauli $L_{(n-3)+1}$ does not commute with $P_0, L_0$ due to condition 4 in Sec. 4.3.2 and hence $L_{n-3}$ cannot be guaranteed to be independent of the rest of the elements in $S$. So, we throw away this element from $S$. This proves that all Paulis in $S \setminus \{L_{n-3}\}$ are independent. We now update $S$:

$$S = \{P_0, L_0, L_3, L_6, \ldots, L_{3(\lceil \frac{n}{3} \rceil - 2)}\}$$

We know that for any given $m$, any maximal subgroup is generated by $m$ independent, pairwise commuting Paulis. We can further write, $S = \{P_0\} \cup \{L_{3j}\}_{j=0}^{\lceil \frac{n}{3} \rceil - 2}$ and note that it has $\lceil \frac{n}{3} \rceil$ operators. Since $|S| \leq m$. This means that $\lceil \frac{n}{3} \rceil \leq m$, hence $n \leq 3m$. □

Having obtained $3m$ as the upper bound, we further ran subgraph isomorphism tests to test the tightness of this bound. Unlike for $m = 2, 3$, when $m = 4$, $3m = 12$ is not a tight upper bound. For this case, all cycles from size 4 to 9 are faithfully realizable. Similarly, when $m = 5$, $3m = 15$ is not a tight upper bound, but all cycles from size 4 to 12 are realizable.

## 4.5 Possible faithful realizations

We now focus our attention on some constructions that provide faithful realizations of Pauli operators over cycle scenarios. We will now provide a few constructions that provide faithful realizations, in accordance with the no-go theorem. But before doing so, we highlight a trivial but useful fact via the following proposition.

**Proposition 7.** *If an $n$-cycle is realizable by $m$-qubit Pauli operators, it is also realizable by $(m+1)$-qubit Pauli operators.*

[3] independent means that no Pauli operator can be written as a product of other commuting Paulis.

*Proof.* Given a $m$-qubit Pauli realization of an $n$-cycle, one can append a new qubit to each such Pauli operator and assign operator $I$ or the same Pauli operator corresponding to the added operator. This produces a $m+1$-qubit Pauli realization for the $n$-cycle. □

### 4.5.1 m-cycle with m-qubit Paulis

Before presenting our construction, we highlight a relevant construction from Ref. [78]. In this work, the authors provided a general construction to realize any compatibility graph via qudit Pauli operators. Their construction applied to the cycle scenarios and qubit Pauli operators guarantees a realization of $m$-cycle using $m$ qubit Pauli operators. Below we present the construction in Ref. [78], to then provide ours which is an improvement over same:

$$\begin{aligned} XIII\ldots III \\ IXII\ldots III \\ ZIXI\ldots III \\ ZZIX\ldots III \\ ZZZI\ldots III \\ \ldots \\ ZZZZ\ldots IXI \\ IZZZ\ldots ZIX \end{aligned} \tag{4.2}$$

Here, starting from the first $m$-qubit Pauli $XIII\ldots III$, and then shifting the first single qubit Pauli operator $X$ to the second qubit, we keep adding a new single qubit $Z$ operator from the left, at the first qubit, while fixing the first $m-1$ Pauli operators of the previous $m$-qubit Pauli to its right. This ensures that a given Pauli is only compatible with the Pauli operator right before it, while being incompatible with the rest. Following this strategy one reaches a point with a path realization of $m-1$, $m$-qubit Pauli operators with the last Pauli being $ZZZZ\ldots IXI$. At this point, the pattern of appending single qubit $Z$ Pauli operator from the left is broken, and in the final step, a single qubit $I$ is introduced at the left while using the first $m-1$ Pauli operators of the previous $m$-qubit Pauli to its right. This ensures that this $m$-qubit Pauli operator is compatible with the previous one and the first $m$-qubit Pauli operators only.

Overall, this construction from Ref. [78] guarantees an $m$-qubit Pauli realization of $m$-cycle. For a visual proof, one can track down how the single qubit $X$ operator in the construction shifts, from first qubit to the $m$th qubit, while the $Z$ operator is appended except at the last Pauli.

### 4.5.2 (m+2)-cycle with m-qubit Paulis.

Our improvement on the previous construction gurantees an $m+2$ sized cycle using $m$-qubit Pauli operators. Basically, we intervene at the final step of that construction, where instead of appending a single qubit $I$ operator at the first qubit, we continue the usual appending of the single qubit $Z$ operator at the first qubit to attain $ZZZZ\ldots ZIX$ and then do the same to attain $ZZZZ\ldots ZZI$. Now, we apply the last step in the previous construction and obtain $IZZZ\ldots ZZZ$, an $m$-qubit Pauli operator that commutes with the first and the Pauli operator right before it while anti-commuting with the rest. Overall, the construction is as follows:

$$\begin{aligned} XIII\ldots III \\ IXII\ldots III \\ ZIXI\ldots III \\ ZZIX\ldots III \\ ZZZI\ldots III \\ \ldots \\ ZZZZ\ldots IXI \\ ZZZZ\ldots ZIX \\ ZZZZ\ldots ZZI \\ IZZZ\ldots ZZZ \end{aligned} \tag{4.3}$$

Figure 16 shows specific examples of this construction. Inductively, Proposition 7 and this construction guarantees faithful realizations of all cycles up to size $m+2$ by $m$-qubit Pauli operators. This is because, by the above construction, single ($m=1$) qubit Paulis can realize a 3-cycle, two ($m=2$) qubit Pauli operators can realize a 4-cycle, similarly three ($m=3$) qubit Pauli operators realize a 5-cycle and so on, and Proposition 7 guarantees that $m$-qubit Pauli operators realize all cycles realizable by $k$-qubits $\forall\ 1 \leq k \leq m-1$.

### 4.5.3 Cycle realizations from Path graph realizations

Before presenting our next construction with which we almost double the number of cycles we get compared to the previous constructions, we need a few results about faithful Pauli realizations of path graphs and their conversion to faithful cycle realizations. We capture these results in the form of theorems and corollaries.

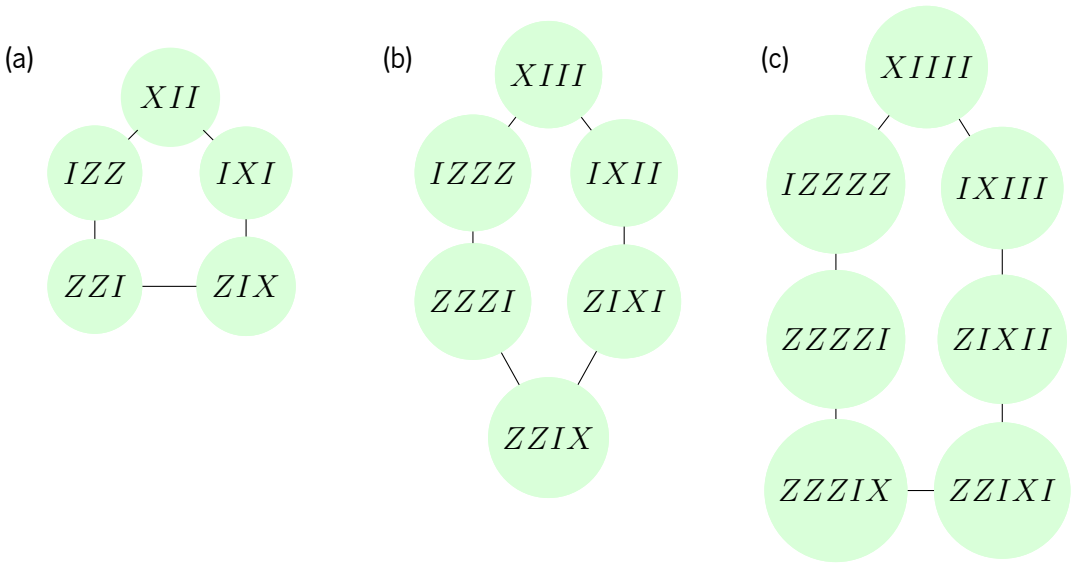



Figure 16: Realization of $m+2$ sized cycle by $m$-qubit Paulis (a) $m=3$ (b) $m=4$ (c) $m=5$.

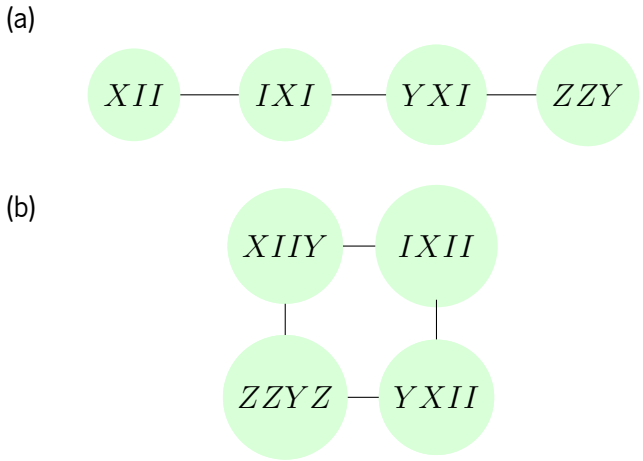



Figure 17: Converting a path graph into a cycle by addition of a qubit (a) a 3-qubit path graph realization of $H_4$ (b) Adding a qubit and assigning anti-commuting Pauli operators, $Y$ and $Z$, to the first and last nodes and $I$ to the rest to get a 4-qubit realization of $C_4$.

**Theorem 8.** *If a path graph with $l$ nodes is faithfully realizable by $m$-qubit Pauli operators, then one can always extend it to a faithful realisation of $l$-cycle using $(m+1)$-qubit Pauli operators.*

*Proof.* Let $H_l$ be a path graph with $l$ nodes where each node is some $m$-qubit Pauli operator. Furthermore, let the $m$-qubit Pauli operator on the $i$th node of graph $H_l$ be denoted as $P_i$. Since it is a faithful realization, Paulis $P_1$ and $P_l$ anti-commute. Recall that this means there are odd number of qubits corresponding to which Pauli operators across $P_1$ and $P_l$ anticommute. Consider the following $l$, $(m+1)$-qubit Pauli operators: $XP_1, IP_2, \ldots, IP_{l-1}, YP_l$. In words, corresponding to the $(m+1)$th qubit we introduce Pauli $X$ for $P_1$, Pauli $Y$ for $P_l$ while appending $I$ for all $P_i$ where $i \in [2, l-1]$. This leads to a $l$-cycle faithful realization by these Pauli operators because it preserves all commutation and anti-commutation relations as in $H_l$ except between $XP_1$ and $YP_l$ which now commute due to even number of anti-commutating single qubit Pauli operators across these two $(m+1)$-qubit Pauli operators, see Figure 17 for an example. □

**Corollary 9.** *If a path graph with $l$ nodes is faithfully realizable by $m$-qubit Pauli operators, then one can always construct a faithful realization of $k$-cycle using $(m+1)$-qubit Pauli operators, $\forall\ 3 \leq k \leq l$.*

*Proof.* Any $m$-qubit faithful realization of a path graph with $l$ nodes contains a faithful $m$-qubit realization of each of its sub-paths with $3$ to $l$ nodes. Applying Theorem 8 to each such sub path proves our claim. □

In Section 4.4, we mentioned that for $m = 3$, one can faithfully realize all cycles from size 4 to 9 except the 8-cycle. Via proposition 7, 4-qubit Pauli operators realize all the cycles realised by 3-qubit Pauli operators. Additionally, we guarantee that 4-qubit Pauli operators also faithfully realize the 8-cycle. This is because the 3-qubit Pauli realization corresponding to $C_9$ has an induced path graph with 8 nodes, which via Theorem 8 guarantees our claim. More generally, Theorem 8 guarantees that if the maximum size of the faithfully realized cycle by a given $m$-qubit Pauli operators is $k$ ($\leq 3m$), then all cycles from size $3$ to $k$ will be necessarily faithfully realized by $(m+1)$-qubit Pauli operators. We now present a result that gives us the power to combine two faithful Pauli realizations of path graphs. We capture this precisely in the following theorem.

**Theorem 10.** *Given a faithful $m$-qubit Pauli realization of a path graph $H_l$, and a faithful $m'$-qubit Pauli realization of path graph $H_{l'}$, one can construct a faithful $(m+m')$-qubit Pauli realization of path graph $H_{l+l'-2}$.*

*Proof.* The construction works as follows, see Figure 18 for illustration. Consider $P_1, P_2, \ldots, P_l$, in the same order, as the $m$-qubit Pauli realization of $H_l$ and $Q_1, Q_2, \ldots, Q_{l'}$, in the same order, as the $m'$-qubit Pauli realization of $H_{l'}$. To each of the first $l-2$, $m$-qubit Pauli operators in the path $H_l$, append

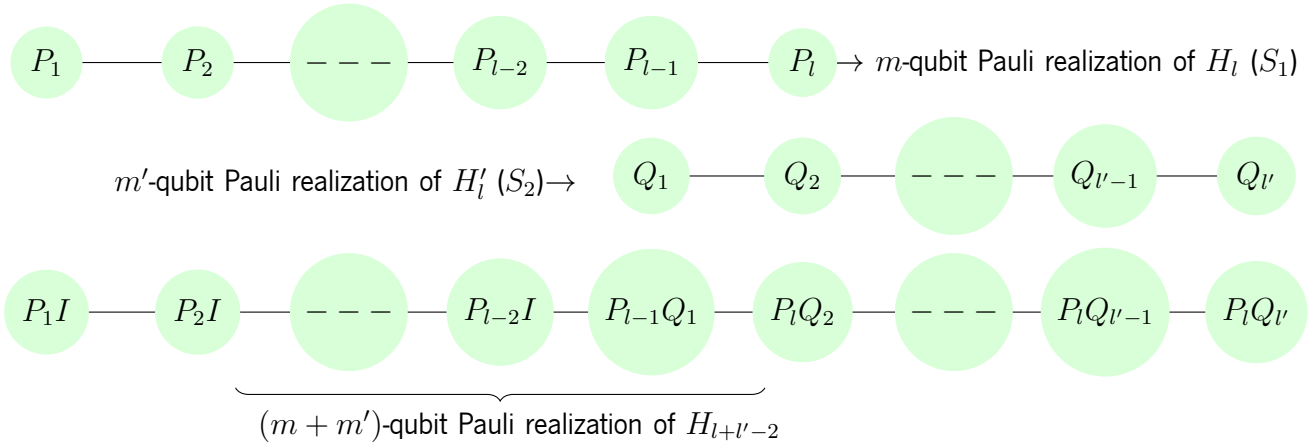



Figure 18: Concatenating a faithful $m$-qubit Pauli realization of path graph $H_l$ with a faithful $m'$-qubit Pauli realization of $H_{l'}$, to obtain a $(m+m')$-qubit Pauli realization of $H_{l+l'-2}$. This works by introducing $m'$-qubit Identity operators in the first $l-2$ operators of the realization of path graph $H_l$ while introducing the Pauli realization for $H_{l'}$ from the $(l-1)$th node in $H_l$. In the additional $l'-2$ nodes, one then fixes $P_l$ in the first $m$ qubits.

a $m'$-qubit $I$ operator, while to the last two operators $P_{l-1}$ and $P_l$, append $Q_1$, $Q_2$ respectively. This construction, let's call it $S_1$, produces a $(m+m')$-qubit faithful realization of path graph $H_l$. Faithful because the appended $m'$-qubit Pauli operators are all pairwise compatible and hence don't perturb the compatiblity relations associated to the initial $m$-qubit Pauli realization of $H_l$. Now consider $l'-2$ copies of the $m$-qubit Pauli $P_l$ and append $Q_3, Q_4, \ldots, Q_{l'}$ respectively to each of the copies. This construction, let's call it $S_2$, produces a faithful $(m+m')$-qubit Pauli realization of $H_{l'-2}$. We can now stack $S_2$ underneath $S_1$ and show that this leads to a faithful $(m+m')$-qubit Pauli realization of path graph $H_{l+l'-2}$: Anti-commutativity across each node in $S_2$ with the first $l-2$ nodes in $S_1$ is ensured because, for each such node in $S_1$, the $m$-qubit Pauli operator (corresponding to first $m$ qubits) anti-commutes with the corresponding operators of every node in $S_2$, while the remaining $m'$-qubit Pauli operator commutes with the corresponding operators of every node in $S_2$. On the other hand, for the last two nodes in $S_1$, the $m$-qubit Pauli operator corresponding to the first $m$ qubits commutes with the corresponding $m$-qubit Pauli operator for every node in $S_2$. This means that the commutativity relationship of the last two nodes in $S_1$ with all of the nodes in $S_2$ is decided by the last $m'$ qubits. All the Pauli operators corresponding to these qubits from $S_1$ to $S_2$ are $Q_1, Q_2, \ldots, Q_{l'}$ which is isomorphic to $H_{l'}$. This suggests that $S_2$ stacked to $S_1$ leads to a faithful $(m+m')$-qubit Pauli realization of $H_{l+l'-2}$. □

**Corollary 11.** *Given a faithful $m$-qubit Pauli realization of a path graph $H_l$, there exists a faithful $(m+1)$-*

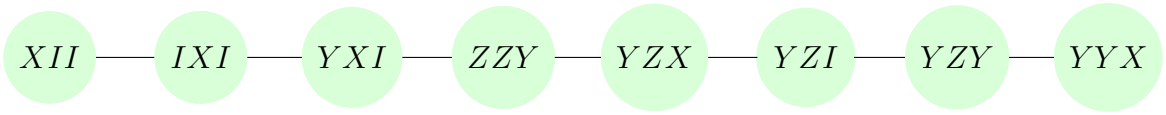



Figure 19: A faithful 3-qubit Pauli realization of a $H_8$.

*qubit Pauli realization of path graph $H_{l+1}$.*

*Proof.* This follows from Theorem 10 where we select $m' = 1$ which gives a faithful realization of $H_3$ (i.e. $l' = 3$). One example of this realization is $X - I - Y$. With the faithful $m$-qubit Pauli realization of a path graph $H_l$, it will produce a faithful $(m+1)$-qubit Pauli realization of path graph $H_{l+1}$. □

### 4.5.4 2m or a 2m-1 sized cycle with m-qubit Paulis

We are now ready to provide our construction. Given a faithful $m$-qubit Pauli realization of a path graph $H_l$ and repeated application of Theorem 10 for $p$ $(\geq 0)$ rounds, where we keep using the faithful $m'$-qubit Pauli realization of path graph $H_{l'}$ in each round, gives us a faithful $(m + pm')$-qubit Pauli realization of path graph $H_{l+p(l'-2)}$. This allows us to construct faithful realizations of cycles for arbitrary number of qubits. Consider the case when $m = m' = 3$ and $l = l' = 8$. Figure 19 shows a faithful realization of $H_8$ with three qubits. Starting from this realization, consider $p$ repeated applications of Theorem 10. This provides a faithful $M$-qubit Pauli realization of path graph $H_L$, where $M = 3 + 3p$ and $L = 8 + 6p$. Based on this construction, we now define three cycle constructions as follows:

1. Application of Theorem 8 to this construction gives us a $L$-cycle using $M' = M + 1$ qubits. Note that $L = 2M'$.

2. Application of corollary 11 to this construction gives us a faithful $(M+1)$-qubit path realization of $H_{L+1}$. Further applying Theorem 8 gives us a faithful $M'$-qubit realization of $(L+1)$-cycle where $M' = M + 2$. Note that $L + 1 = 2M' - 1$.

3. Application of Theorem 10 such that $m' = 2$, and $l' = 2$, gives us a faithful $(M+2)$-qubit Pauli realization of path graph $H_{L+3}$. Further applying Theorem 8 gives us a faithful $M'$-qubit path realization of $(L+3)$-cycle where $M' = M + 3$. Note that $L + 3 = 2M' - 1$.

Note that since $M = 3(1 + p)$, the cycle realization in construction (1) implies that with number of qubits $M' \equiv 1$ (mod 3) we can faithfully realize a cycle of size $2M'$ while constructions (2) and (3) imply that with number of qubits $M' \equiv 0$ (mod 3) or $M' \equiv 2$ (mod 3), we can faithfully realize a cycle of size

$2M'-1$. This means that a faithful Pauli realization of a cycle of size $2M'-1$ or $2M'$ is guaranteed with $M'$ qubits. Corollary 9 further implies that with $M'+1$ qubits all cycles from size 3 up to size $2M'-1$ or $2M'$ are faithfully realizable. We can further rephrase this result as follows. With $M'$ qubits all cycles from size 3 up to size $2M'-3$ or $2M'-2$ can be faithfully realized by Pauli operators depending on whether $M \not\equiv 1$ (mod 3) or $M \equiv 1$ (mod 3) respectively. This leaves open the question of whether there exists a faithful realization of the cycle of size $2M'-2$ when $M \not\equiv 1$ (mod 3) and $2M'-1$ when $M \equiv 1$ (mod 3). This finishes the descriptions of our constructions.

## 4.6 Contextuality with Pauli cycles

We now focus on the contextuality witnessing aspects of the Pauli realizations of $n$-cycle scenarios. Luckily, the commutation properties in Section 4.3.2 turn out to be enough for this characterization. The first step in every proof below takes inspiration from Tsirelson's original idea of squaring the operator to derive quantum bound for the CHSH inequality. We dedicate these proofs to his contribution to the field of quantum foundations.

For this analysis, we make use of the complete characterization of the non-contextual polytopes of $n$-cycle scenarios, see Section 8 for more details. The general form of the $n$-cycle noncontextuality inequalities is:

$$\sum_{i=0}^{n-1} \gamma_i \langle A_i A_{i\oplus 1}\rangle \overset{NCHV}{\leq} n-2$$

where $\gamma_i \in \{-1,+1\}$ with $\prod_{i=0}^{n-1}\gamma_i = -1$ and $\{A_i\}_{i=0}^{n-1}$ are the measurements labelling each vertex. We begin our analysis with the simplest cycle scenario, i.e. the 4-cycle scenario.

### 4.6.1 4-cycle scenario

We start off by constructing operators corresponding to the inequalities of the 4-cycle:

$$\Gamma = \sum_{i=0}^{3} \gamma_i \langle A_i A_{i\oplus 1}\rangle \overset{NCHV}{\leq} 2$$

The relevant quantum operator form of the LHS in terms of $m$-qubit Pauli operators is as follows:

$$\langle \Gamma \rangle = \gamma_0 \langle P_0 P_1\rangle + \gamma_1 \langle P_1 P_2\rangle + \gamma_2 \langle P_2 P_3\rangle + \gamma_3 \langle P_3 P_4\rangle$$

Remember that we introduced the edge Paulis previously, and so the operator $\Gamma$ can be written in terms of them as follows:

$$\Gamma = \gamma_0 L_0 + \gamma_1 L_1 + \gamma_2 L_2 + \gamma_3 L_3$$

The condition for contextuality of the Paulis then becomes:

$$\langle \Gamma \rangle > 2$$

So, if we can find some that $|\Psi\rangle$ that gives the expectation value above 2 means those statistics don't have a non-contextual hidden variable explanation. We square the operator obtained above:

$$\Gamma^2 = (\gamma_0 L_0 + \gamma_1 L_1 + \gamma_2 L_2 + \gamma_3 L_3)^2$$

For brevity, we write $\Gamma$ as:

$$\Gamma = a + b$$

where $a = \gamma_0 L_0 + \gamma_1 L_1, b = \gamma_2 L_2 + \gamma_3 L_3$.

Therefore $\Gamma^2 = (a+b)^2 = a^2 + b^2 + ab + ba$. Here,

$$a^2 = \gamma_0^2 L_0^2 + \gamma_1^2 L_1^2 + \gamma_0\gamma_1\{L_0, L_1\}$$

$$b^2 = \gamma_2^2 L_2^2 + \gamma_3^2 L_3^2 + \gamma_2\gamma_3\{L_2, L_3\}$$

$$ab = \gamma_0\gamma_2\{L_0, L_2\} + \gamma_0\gamma_3\{L_0, L_3\}$$

$$ba = \gamma_1\gamma_2\{L_1, L_2\} + \gamma_1\gamma_3\{L_1, L_3\}$$

Using ${\gamma_i}^2 = 1$ and the edge Pauli commutation relations presented in Section 4.3.2: $\{L_0, L_1\} = \{L_0, L_3\} = 0 = \{L_2, L_3\} = \{L_2, L_1\}$ and $[L_0, L_2] = 0 = [L_1, L_3]$, implies that:

$$a^2 = 2I$$

$$b^2 = 2I$$

$$ab + ba = 2\gamma_0\gamma_2(L_0L_2) + 2\gamma_1\gamma_3(L_1L_3)$$

$$\Gamma^2 = 4I + 2\gamma_0\gamma_2 L_0L_2 + 2\gamma_1\gamma_3 L_1L_3$$

We can further simplify this expression by noting that the commutation relations of the 4-cycle Paulis follow the following property:

$$L_1L_3 = P_1P_2P_3P_0 = -P_0P_1P_2P_3 = -L_0L_2$$

This simplifies $\Gamma^2$ as follows:

$$\Gamma^2 = 4I + 2L_1L_3(\gamma_0\gamma_2 - \gamma_1\gamma_3)$$

We know that there are only odd number of indices $i$ s.t. $\gamma_i = -1$, hence, for the 4-cycle, two cases exist: (i) only one index $i$ s.t. $\gamma_i = -1$ (ii) three indices $i$ s.t.$\gamma_i = -1$.

If (i) is true, then:

$$\Gamma^2 = 4I + 4L_1L_3 = 4(I + L_1L_3) \qquad \text{(where } \gamma_1 \text{ or } \gamma_3 \text{ is -1)}$$

$$\Gamma^2 = 4I - 4L_1L_3 = 4(I - L_1L_3) \qquad \text{(where } \gamma_0 \text{ or } \gamma_2 \text{ is -1)}$$

If (ii) is true, then:

$$\Gamma^2 = 4I + 4L_1L_3 = 4(I + L_1L_3) \qquad \text{(where } \gamma_1 \text{ or } \gamma_3 \text{ is 1)}$$

$$\Gamma^2 = 4I - 4L_1L_3 = 4(I - L_1L_3) \qquad \text{(where } \gamma_0 \text{ or } \gamma_2 \text{ is -1)}$$

We get equivalent conditions from both possibilities. Also, note that $L_1L_3$ produces some Pauli operator in $\mathcal{P}_m$. Therefore, for every case the maximum eigenvalue of $\Gamma^2$ is 8. The maximum eigenvalue of $\Gamma$ is $\sqrt{8}$ i.e. $+2\sqrt{2}$ or $-2\sqrt{2}$. When $\Gamma = -2\sqrt{2}$, for a given inequality, then $\Gamma = 2\sqrt{2}$ for a valid inequality which is a negative multiple of the other one. Moreover, the state corresponding to this maximal violation is an eigenstate of the Pauli $L_1L_3 = -P_0P_1P_2P_3$. Therefore, each Pauli realization of a 4-cycle maximally violates some 4-cycle NC inequality.

### 4.6.2 5-cycle scenario

In the 5-cycle case the quantum operator ($\Gamma$) appears as follows:

$$\Gamma = \gamma_0L_0 + \gamma_1L_1 + \gamma_2L_2 + \gamma_3L_3 + \gamma_4L_4$$

Squaring the operator, gives:

$$\Gamma^2 = 5I + \sum_{\substack{i=0,k=0 \\ i\neq k}}^{4} \gamma_i\gamma_k\{L_i, L_k\}$$

We know from the commutation constraints highlighted in Section 4.3.2 that for $k = 5$, all pair of distinct edge Paulis anti-commute, hence:

$$\Gamma^2 = 5I$$

This means that the maximum eigenvalue of $\Gamma$ is $\sqrt{5} < 3$. Therefore, no $m$-qubit Pauli realization for a 5-cycle produces a contextual behaviour, for any $m$.

### 4.6.3 n > 4 cycle scenarios

Turns out that we can capture all cases above the 4-cycle scenario in a general manner. We capture this general situation in the form of a theorem.

**Theorem 12.** *No $m$-qubit Pauli realization can witness contextuality in an $n > 4$-cycle scenario.*

*Proof.* From the specific examples that we saw above we can notice that once we square the operator $\Gamma$, the only terms that survive among the anti-commutators are the ones where the edge Paulis commute with each other. For any $n \geq 5$ cycle:

$$\Gamma = \sum_{i=0}^{n-1} \gamma_i L_i$$

The condition for contextuality is $\langle \Gamma \rangle_{|\Psi\rangle} > n - 2$.

$$\Gamma^2 = kI + \sum_{\substack{i=0,k=0 \\ i \neq k}}^{n-1} \gamma_i \gamma_k \{L_i, L_k\}$$

As noted in the 5-cycle case, within the summation on the r.h.s the anti-commuting pairs don't contribute whereas each commuting term appears as $2\gamma_i \gamma_k L_i L_k$. This means that:

$$\Gamma^2 = kI + \sum_{\substack{i=0,k=0 \\ i \neq k}}^{n-1} 2\gamma_i \gamma_k L_i L_k$$

Now we only need the number of such surviving terms to finish our proof. Using Figure 15 as an illustration, we count as follows:

1. $L_0, L_1, L_2$ each commutes with $n - 5$ other edge paulis .

2. For $L_3$ we need to avoid redundancy and discount any commutation with Paulis in 1. We have $(n - 5) - 1 = n - 6$ commutations i.e. we excluded one with $L_0$.

3. For $L_4$ we need to exclude the commutation with $L_0$ & $L_1$. Hence $(n - 5) - 2 = n - 7$ such relations.

4. We keep going like this, we reach $L_{n-4}$ where only 1 commutation relation needs to be counted i.e. with $L_{n-1}$.

5. Beyond that every edge Pauli commutation combination has already been counted for in the steps above.

This means that the total unique counts that contribute in the r.h.s (summation part) of $\Gamma^2$ above are:

$$3(n-5)+(n-6)+(n-7)+(n-6)+\ldots+1 = 2(n-5)+\frac{(n-5)(n-4)}{2}$$

Clearly, this sum only makes sense for $n \geq 5$. Now, we try to derive an upper bound on the maximum eigenvalue of $\Gamma^2$ operator defined above:

$$\langle \Gamma^2 \rangle_{|\Psi\rangle} = n + \sum_{i \neq k} 2\gamma_i \gamma_k \langle L_i L_k \rangle_{|\Psi\rangle}$$

If we try to compute the algebraic upper bound of r.h.s above by noting that each term in the summation is $\leq 1$ (can't all together be 1 since all the Paulis across terms don't pairwise commute), Therefore:

$$\langle \Gamma^2 \rangle_{|\Psi\rangle} < n + 2\{2(n-5)+\frac{(n-5)(n-4)}{2}\} = n^2 - 4n$$

Therefore,

$$0 \leq \text{eig.val.}(\Gamma^2) < n^2 - 4n$$

where eig.val.($\Gamma^2$) $\equiv$ any eigenvalue of operator $\Gamma^2$. This also means that:

$$-\sqrt{n^2-4n} < \text{eig.val.}(\Gamma) < \sqrt{n^2-4n}$$

For non-contextuality it must be that for all states $|\Psi\rangle$ in $\mathcal{H}_2^{\otimes m}$:

$$\langle \Gamma \rangle_{|\Psi\rangle} \leq (n-2)$$

Hence, condition for non-contextuality means that:

$$\sqrt{n^2-4n} < n-2$$

which always holds $\forall\, n \geq 5$. Hence $\langle \Gamma \rangle_{|\Psi\rangle} < n-2$ ($\forall\, n \geq 5$ and $|\Psi\rangle$). This means that the statistics obtained from any faithful $m$-qubit Pauli realization of an $n$-cycle, lies strictly inside the classical (NC) polytope ($\forall\, n \geq 5$). □

## 4.7 General compatibility graphs

Having fully characterized Pauli contextuality in $n$-cycle scenarios, we now move our discussion towards more general scenarios. Owing to Vorob'ev's theorem, as explained in Section 2.3.1, and our result that only the 4-cycle scenario produces contextual correlations by qubit Pauli operators, we now answer whether the presence of an induced 4-cycle is a necessary and sufficient condition for Pauli contextuality in arbitrary KS scenarios.

### 4.7.1 Implication of 4-cycle contextuality

The proofs above imply that if an arbitrary compatibility graph has at least one 4-cycle as its induced subgraph, then there always exists a quantum state that produces statistics indescribable by any noncontextual model. We can prove this by the following argument: Imagine that a graph $G$ has an induced 4-cycle and assume that the statistics corresponding to Pauli realizations of $G$ are all non-contextual. This means that a joint probability distribution (JPD) over the outcomes of measurements realizing $G$ explains all the statistics one can observe over any sub scenario of $G$, by marginalizing over the rest of the graph. But from our result, on 4-cycle Pauli contextuality above, we know that we can find some quantum state $|\Psi\rangle$ for every Pauli realization such that it violates one of the 4-cycle NC inequalities, or its lifted version. This means that for each 4-cycle Pauli realization, there always exists statistics that cannot be explained by a JPD over $G$. Hence, the initial assumption is contradicted and graph $G$ can produce contextuality [4] via qubit Pauli operators.

This raises a natural question: Given an arbitrary Pauli compatibility graph, with at least one induced $n$-cycle ($n \geq 5$), can it witness contextuality? For graphs where there is exactly one such cycle, it's not complicated to construct a JPD, but for generic graphs one has to be careful. The same question further raises another important question: What is the precise role of induced cycles within compatibility graphs in producing contextuality? We know that presence of cycles ensures that one can always find some quantum observables such that the whole graph produces at least as much contextuality as any induced cycle [32] whereas for the current case of the Pauli compatibility graph we know that the induced Pauli cycles independently can't produce contextuality. But this still doesn't imply that such a Pauli graph cannot exhibit contextuality. Therefore, we now move on to study Pauli compatibility graphs that are composed combinations of two 5-cycles.

### 4.7.2 Two induced 5-cycles

We consider three cases where two 5-cycles conjoin together differently.

**Case I.** Only one node is common as depicted in Figure 20. Clearly, such a Pauli-graph always has a non-contextual model. This is because a JPD ($p_T$) over it exists:

$$p_T = \frac{p_{C_1} p_{C_2}}{p_{IIIX}}$$

where $p_{C_i}$ is a JPD for $i^{th}$ cycle and $p_{IIIX}$ is probability for outcomes of measurement $IIIX$. Due to no-

[4] More than anything, this proof underlines that adding more Paulis in addition to the 4-cycle in no way nullifies (or decreases) the contextuality witnessing effects of the 4-cycle.

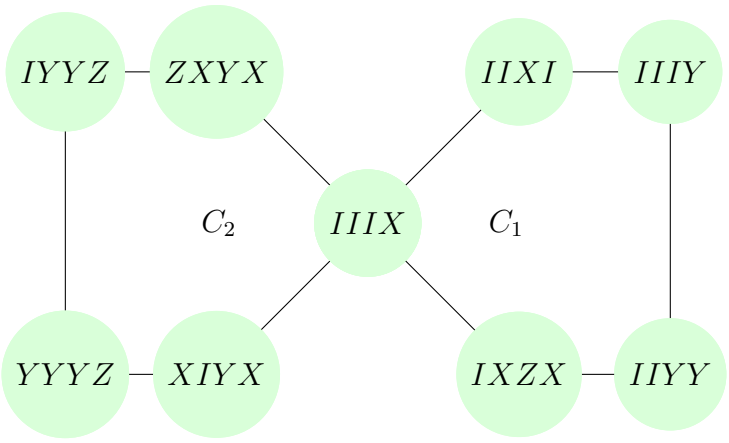



Figure 20: Two 5-cycles joined together at a node: a graph achievable only with $m \geq 4$ qubit Pauli operators. No such arrangement of Paulis can ever produce contextuality.

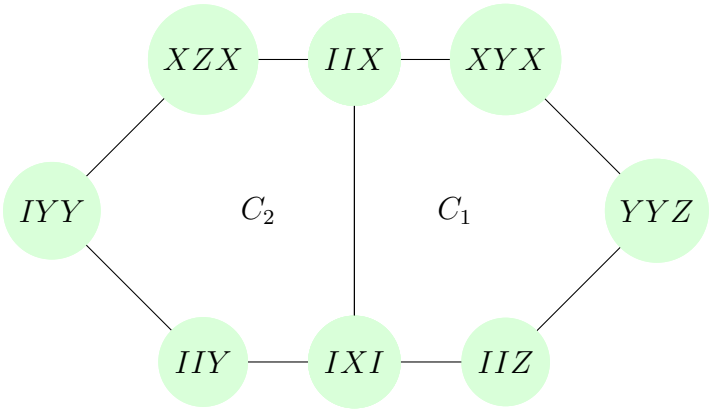



Figure 21: Two 5-cycles joined together: a graph achievable when $m \geq 3$ qubit Pauli operators. Such a graph can't produce contextuality for any realization of Pauli operators.

disturbance condition the common measurement $IIX$ will always admit a unique probability distribution.

**Case II.** Two cycles share an edge as depicted in Figure 21. This Pauli-graph always possesses a non-contextual model due to the existence of a JPD ($p_T$) over it:

$$p_T = \frac{p_{C_1} p_{C_2}}{p_{\{IIX, IXI\}}}$$

where $p_{C_i}$ is a JPD for $i^{th}$ cycle and $p_{\{IIX,IXI\}}$ is JPD over the common context $\{IIX, IXI\}$. Due to the no-disturbance condition, this common context will always admit a unique probability distribution.

**Case III.** The two 5-cycles conjoin together with two edges in common. This case becomes non-trivial because now there is no guarantee that the JPDs over the two cycles give the same marginals for the unmeasurable correlation, see $\{P_1, P_4\}$ in Figure 22: Due to the no-disturbance condition the JPDs over the two cycles are constrained to give the same overlap for the common contexts $\{P_1, P_5\}$ and $\{P_4, P_5\}$

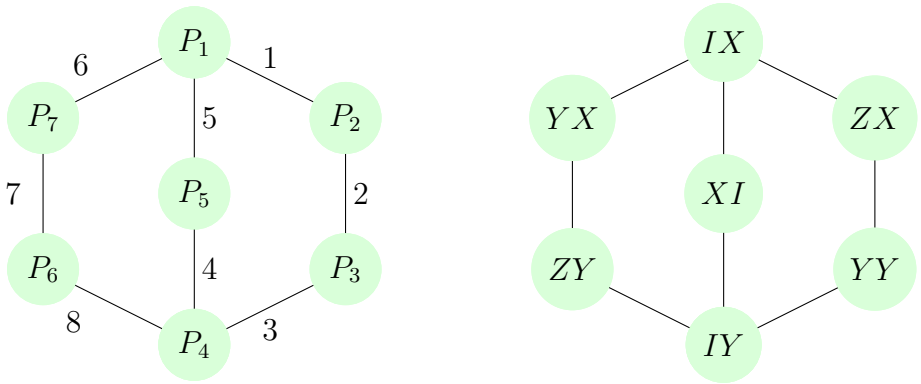



Figure 22: Two 5-cycles sharing two edges $\{P_1, P_5\}, \{P_4, P_5\}$: a graph achievable $\forall\ m \geq 2$: (a) represents the generic $m$-qubit Paulis and respective context label on the edges. (b) exemplifies one realization with Paulis that violates a noncontextuality inequality of the scenario.

in Figure 22. But since $\{P_1, P_4\}$ is not a context, the same doesn't hold for it.

To test whether Pauli operators produce contextuality we first obtained all the noncontextuality inequalities corresponding to this graph, using the PORTA software package [79]. Then, using the qubit Pauli operators illustrated in Figure 22(b), we observed a violation of the following noncontextuality inequality:

$$-p^1_{--} + p^2_{-+} + p^3_{-+} - p^4_{--} + (p^5_{-+} + p^5_{+-} + 2p^5_{--}) - p^6_{--} + (p^7_{-+} + p^7_{+-} + 2p^7_{--}) + (p^8_{+-} - p^8_{-+} + p^8_{--}) \leq 3 \quad (4.4)$$

Here $p^i_{ab}$ represents probability of joint outcome $ab$ on measuring the $i^{th}$ context. By translating the projectors of contexts to Pauli operators, we can translate this inequality into an operator inequality, as follows:

$$-(P_1P_2 + P_2P_3 + P_3P_4 + P_4P_5 + P_1P_7) + P_4P_6 - (P_4 + P_5 + P_6 + P_7) \leq 4I \quad (4.5)$$

Using the Pauli realization in Figure 22(b), the l.h.s above becomes:

$$-(ZI + XZ + YI + XY + YI) + ZI - (IY + XI + ZY + YX)$$

Now, if we check for the maximum eigenvalue of this expression, it turns out to be $4.2716 > 4$. Hence, a violation of the inequality. The corresponding state is: $(0.2787 - 0.5952i, -0.2787 - 0.3342i, -0.4092 + 0.1482i, -0.4352 + 0.0000i)^T$ in the computational basis. This implies that the distribution for the context $\{P_1, P_4\}$ obtained from the graph, via marginalisation, won't be unique.

Therefore, these conjoined cycles produce contextuality. This illustrates that the quantum violation of the NC inequalities for a given scenario (compatibility graph) doesn't necessarily accompany the violation of some induced cycle NC inequality within the graph. So, it seems that the fundamental role of an induced cycle in a compatibility graph of a fixed set of quantum measurements is only to preclude application of Vorob'ev's theorem to the graph. The chosen set of quantum measurements then may or may not produce contextuality.

Another way to understand this is in terms of the geometrical approach to NC correlations i.e. the correlation polytopes. In Chapter 3, we showed that a facet-defining inequality for a KS scenario remains a facet-defining inequality for any scenario that is an extension of it. Therefore, for scenarios that can be seen as extensions of cycle scenarios, the $n$-cycle noncontextuality inequalities are just a proper subset of all facet-inducing inequalities of the scenario. The same holds for the case presented in case III above. Pauli contextuality observed is then just the violation of a noncontextuality inequality outside this proper subset.

## 4.8 Discussion

In this chapter we analyzed the contextuality-witnessing aspects of multi-qubit Pauli operators. In doing so, we also analyzed the faithful realizability of these operators for cycle scenarios.

Our analyses exploit our creative introduction of 'edge Pauli' operators over $m$-qubit Pauli realization of cycle scenarios. These Pauli operators provide commutativity constraints additional to the ones coming from the faithfullness constraint on the cycle Pauli operators. From deriving our no-go theorem to fully characterizing the cycle scenarios, the most exploited constraint of edge Pauli operators is that each edge Pauli commutes with every other edge Pauli except for its nearest and next-nearest neighbors. An advantage of introducing these operators is that the corresponding quantum operator for an $n$-cycle inequality is the sum of expectation values of edge Pauli operators. This further allowed us to use Tsirelson's technique, of squaring the quantum operator for the inequality, and exploit the fact that qubit Paulis either commute or anti-commute, to provide the complete characterization of Pauli contextuality for cycle scenarios.

In extending our qubit-Pauli contextuality analysis beyond cycle scenarios we studied scenarios composed of different configurations of 5-cycles. We found that in one such scenario, where the overlapping parts of two 5-cycles are two common contexts, see Figure 22, there is a Pauli realization that violates a non-contextuality inequality that is not a cycle inequality. This suggests that the role of an induced cycle sub-scenario in an arbitrary situation is simply to prevent a joint probability distribution from existing for

the whole compatibility graph as dictated by Vorob'ev's theorem, highlighted in Sec. 2.3.1. Later in Sec. 6.2, we will discuss the potential future directions cropping out from this work.

# Chapter 5

# Exclusivity principle, Ramsey theory, and n-cycle PR boxes

**Chapter Summary.** The exclusivity principle (E-principle) states that the sum of probabilities of pairwise exclusive events cannot exceed 1. In this chapter, we present the work that extends the study of the E-principle for detecting post-quantumness of correlations to the broader domain of KS contextuality scenarios, advancing the analyses from Ref. [22] which were restricted to Bell scenarios. Since quantum theory satisfies this principle, any correlation that violates it is necessarily post-quantum. Although equivalent to the no-signalling condition for any bipartite Bell scenario, this equivalence breaks down for two independent copies of the same scenario. Such violation of the E-principle due to multiple copies, known as its *activation effect*, was studied in Ref. [22] for the nonlocal extremal boxes of $(2, m, 2)$, $(2, 2, d)$, and $(3, 2, 2)$ Bell scenarios. The authors mapped the problem of exhibiting activation effects to finding certain cliques inside joint exclusivity graphs. With this work, we refine the joint exclusivity structure to be an edge-colored exclusivity multigraph. This allows us to draw a novel connection to Ramsey theory, which guarantees the existence of certain monochromatic subgraphs in sufficiently large edge-colored cliques, providing a powerful tool for ruling out E-principle violations. We exploit the power of this connection to study activation effects of E-principle for the contextual extremal boxes of $n$-cycle scenarios, dubbed $n$-cycle PR boxes.

**Motivation.** Quantum theory, unlike the theory of special relativity, lacks a description where simple physical principles guide its mathematical machinery. Identifying such principles is one of the major goals in quantum foundations research, pursued under programmes aiming at 'reconstructions of quantum theory'. Such efforts seek to enhance our understanding of quantum theory and diminish the mystery around its interpretations.

Several principles have been proposed to characterize quantum correlations, such as information

causality, macroscopic locality, non-trivial communication complexity, or no advantage for non-local computation. However, Gallego et al. [80] challenged the effectiveness of these principles in singling out quantum correlations by providing examples of post-quantum correlations in tripartite Bell scenarios satisfying each of them. This revealed a key shortcoming in their bipartite formulations, highlighting the need for multipartite versions.

Subsequently, Fritz et al. [22] introduced a new intrinsically multipartite principle, the Local Orthogonality principle, and explored its ability to detect post-quantum correlations in certain Bell scenarios. This principle also applies to Kochen–Specker (KS) contextuality scenarios more broadly, where it is commonly known as the exclusivity principle (or E-principle). It states that the sum of probabilities of a set of pairwise exclusive events [1] must be less than or equal to $1$.

The E-principle was used to witness post-quantumness of extremal correlations in the $(2, 2, d)$, $(2, m, 2)$, and $(3, 2, 2)$ Bell scenarios [22]. The method made use of the *activation effect* of the E-principle, where a correlation may independently satisfy the E-principle but the joint consideration of multiple independent copies of that correlation 'activates' a violation. The simplest example is that of Popescu–Rohrlich (PR) boxes: while a single PR box satisfies the E-principle, joint events coming from two independent PR boxes violate it. Since quantum correlations satisfy the E-principle and thus so do independent products of quantum correlations, this violation witnesses the post-quantumness of the PR box itself. That is, the fact that two PR boxes violate the E-principle certifies that a single PR box is not quantum realizable. Similarly, it turns out that for any non-local extremal correlation of the $(2, 2, d)$, $(2, m, 2)$, and $(3, 2, 2)$ Bell scenarios, two copies are enough to detect post-quantumness via the E-principle [22].

Besides certifying post-quantumness of the aforementioned Bell non-local correlations, the E-principle has provided other interesting insights, which we briefly review. Cabello [81] used the E-principle to derive the quantum maximum of the famous KCBS non-contextuality inequality. Cabello et al. [82] showed that the exclusivity graphs of non-contextuality inequalities violated by quantum theory necessarily contain certain induced subgraphs, and further provided numerical evidence that the quantum maxima of the non-contextuality inequalities corresponding to those basic exclusivity graphs can be singled out by the E-principle. Amaral et al. [83] showed that for self-complementary exclusivity graphs, the E-principle singles out the quantum set of correlations, known as the theta body of the graph [2]. This result was later generalized by Cabello [84] to arbitrary exclusivity graphs. Furthermore, it is known that any theory that

[1] An event here refers to the joint outcome of a set of compatible measurements. Two events are deemed exclusive when they assign different outcomes to a common measurement, e.g. the events $[++|AB]$ and $[--|AC]$ differ on the common measurement $A$.

[2] Note that the theta body of an exclusivity graph contains all quantum correlations consistent with that graph and not just within a given scenario. That is, it contains quantum correlations coming from all possible scenarios that produce events realizing the given exclusivity graph.

follows Specker's principle immediately satisfies the E-principle, including quantum and classical theories [85]. But the converse does not hold: the set of almost quantum correlations is inconsistent with Specker's principle [86] yet satisfies the E-principle [87].

Despite the multitude of contributions of E-principle to studies related to KS scenarios, its analysis for testing the post-quantumness of correlations beyond Bell scenarios is still missing in the literature. This work fills this gap by analysing the contextual extremal vertices of the no-disturbance polytopes of $n$-cycle KS scenarios.

**Outline of the chapter.** With increasing intensity, we explain the details of our established connection of Ramsey theory and E-principle violations in this chapter, starting from an overview of this connection in Sec. 5.1. We then introduce the mathematical background required to study activation effects of the E-principle in Sec. 5.2. In Sec. 5.3, we explain our refinement of the conventional method to study activation effects. We then move on to describe $n$-cycle PR boxes, in Sec. 5.4, which are the correlations to which we apply our Ramsey theory results as highlighted in Sec. 5.5. We then provide the proofs of all these results in Sec. 5.6. We finally close the chapter with a discussion in Sec. 5.7.

## 5.1 Overview

We now present a concise overview of the connection we established between Ramsey theory [23, 24] and activation effects of E-principle.

A systematic approach to study activation effects of E-principle uses graph theory where an exclusivity graph is built whose nodes represent possible events of a correlation and whose edges represent exclusivity between such events. The activation setup comprising multiple independent correlations is then captured by constructing the joint exclusivity graph as the OR product of the individual exclusivity graphs. By the very definition of E-principle, its violation manifests as the existence of certain cliques inside this joint exclusivity graph. Ref. [22] exploited computational tools to find such cliques, but while that method worked well for relatively small graphs, in general the problem of finding maximal cliques is known to be NP-hard [88].

Given the complexity of the clique-finding problem, our discovery of the previously unknown connection between the activation effect of the E-principle and Ramsey theory proves to be highly useful, see Sec. 5.5. We set up this connection by introducing a refinement to the above defined method of analysing activation effects, which is by replacing the joint exclusivity graph by an edge-colored multigraph. More specifically, a unique color is assigned to the exclusivity graph of each independent correlation, allowing one to preserve

| | $++$ | $+-$ | $-+$ | $--$ |
|---|---|---|---|---|
| $A_0B_0$ | $\frac{1}{2}$ | 0 | 0 | $\frac{1}{2}$ |
| $A_0B_1$ | 0 | $\frac{1}{2}$ | $\frac{1}{2}$ | 0 |
| $A_1B_0$ | $\frac{1}{2}$ | 0 | 0 | $\frac{1}{2}$ |
| $A_1B_1$ | $\frac{1}{2}$ | 0 | 0 | $\frac{1}{2}$ |

Table 1: Correlation table for the PR box

the information about which individual copies are responsible for determining exclusivity between two joint events. By contrast, in the conventional approach, this detailed information is erased by the OR product. The problem of finding violations of the E-principle then becomes that of finding edge-colored cliques inside this multigraph. Such edge-colored cliques are precisely the objects of study of Ramsey theory when applied to graphs.

Ramsey theory investigates an intriguing phenomenon about how order emerges from randomness: as combinatorial structures grow larger, certain patterns become inevitable, as expressed by the popular catchphrase 'infinite chaos is impossible'. When applied to graph theory, it studies the minimal size an edge-colored complete graph must be to force the appearance of specific monochromatic substructures regardless of the particular coloring of the edges. A classic illustration involves a complete graph whose edges can be colored either red or green: once the graph contains at least six vertices, it must contain either a red triangle or a green triangle.

## 5.2 Preliminaries

To precisely understand the power of Ramsey theory in determining activation effects of E-principle, it is important to familiarize oneself with some basic definitions and concepts. We now present the same.

### 5.2.1 Exclusivity graphs

Given a correlation, one constructs its *exclusivity graph* of possible events. The nodes of this graph represent joint outcome events over maximal contexts that occur with non-zero probability, e.g. the event $[++|A_0B_0]$ for the PR box. Two vertices are connected by an edge whenever the two events are *exclusive*, meaning that there is a common measurement to which they assign different outcomes; for example, the events $[++|A_0B_0]$ and $[-+|A_0B_1]$ are exclusive since they differ on the common measurement $A_0$. Figure 23 depicts the exclusivity graph of the PR box, see table 1, which is an extremal correlation of the

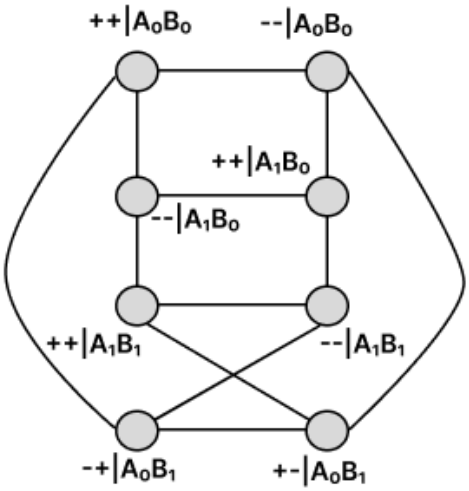



Figure 23: Exclusivity graph of the PR box

$(2, 2, 2)$ Bell scenario, i.e. a vertex of its polytope of no-signalling correlations. One may think of exclusivity graphs of correlations as being vertex-weighted: the weight of a node is the probability assigned by the correlation to the event it represents. In the case of the PR box, all nodes are assigned the same weight, $\frac{1}{2}$. Such uniform weighting is typical of the class of exclusivity graphs studied in this text, namely those arising from extremal correlations in cycle scenarios ($n$-cycle PR boxes) and their products. Consequently, we mostly focus on the structure of the underlying unweighted graphs.

### 5.2.2 The exclusivity principle

The *exclusivity principle* (or *E-principle*) states that the sum of probabilities of a set of mutually exclusive events must be less than or equal to $1$.

To test this principle on a given correlation, it suffices to consider its (vertex-weighted) exclusivity graph. Since edges in this graph represent exclusivity, its cliques (fully connected subsets of nodes) correspond to sets of mutually exclusive events, each determining an E-principle inequality. This renders cliques crucial structures for studying the exclusivity principle using graph theory. Verifying whether a correlation satisfies the exclusivity principle then amounts to checking the inequalities corresponding to all the maximal cliques of its exclusivity graph. This quickly becomes intractable, since it requires enumerating all the maximal cliques of a graph, which is NP-hard [3]. Still, various software packages are available which can be used to find cliques within modestly sized graphs.

[3] More precisely, the problem of deciding whether a graph has a clique of a given size is NP-complete [88]. Listing all maximal cliques clearly allows one to solve this decision problem. Moreover, since there are graphs with exponentially many maximal cliques, in the worst case their enumeration requires exponential time.

Violations of the E-principle exhibit what are known as *activation effects*, where correlations that individually satisfy the E-principle may violate it when combined. In particular, there exist correlations that satisfy the E-principle on their own, yet lead to a violation when multiple independent copies are considered together. The PR box provides a concrete example: while a single PR box satisfies the E-principle, two independent PR boxes (i.e. the correlation obtained as the product of two PR boxes) violate it.

In graph-theoretic terms, the violation is witnessed by a clique in the joint exclusivity graph of the two (or more) independent correlations, even though the individual exclusivity graphs admit no violation-witnessing cliques.

### 5.2.3 Joint exclusivity graph of independent boxes

To study activation effects, we must consider the product of correlations that captures their independent parallel composition. Here the word 'independent' evokes an analogy with independent random variables in probability theory, which are jointly distributed according to the product measure. The situation we have in mind is that of the two (or more) measurement scenarios being run in parallel and producing outcomes independently of each other.

To keep the analysis at the level of exclusivity graphs, we construct the joint exclusivity graph from the exclusivity graphs of the individual correlations. This involves specifying how independent events combine into joint events, and defining the exclusivity relation between joint events from the individual ones, as well as the probabilistic weights assigned to joint events. In short, a joint event is represented by a tuple comprising one event drawn from each individual graph, and two joint events are deemed exclusive if their component events are exclusive in at least one of the individual graphs. The probabilistic independence of the correlations is reflected in that the probability of a joint event is the product of the probabilities of its constituent events.

This idea is formally captured through the *OR product*, also known as the *co-normal product*, of graphs. Let $G_i$ be the exclusivity graph of the $i$th correlation for $i \in \{1, \ldots, k\}$. The joint exclusivity graph of these $k$ independent correlations will be denoted $J^k(G_1, \ldots, G_k)$, or $J^k$ for short leaving the graphs implicit. It is defined as follows. The vertex set of $J^k$ is

$$V(J^k) = V(G_1) \times \cdots \times V(G_k), \tag{5.1}$$

where '$\times$' denotes the Cartesian product. That is, a node of $J^k$ is a $k$-tuple $u = (u_1, \ldots, u_k)$ where $u_i \in V(G_i)$, i.e. each $u_i$ is a node in $G_i$. The edge set of $J^k$ is defined as

$$E(J^k) = \{(u, v) \mid \exists\, i \in \{1, \ldots, k\}.\ (u_i, v_i) \in E(G_i)\} \tag{5.2}$$

Finally, we define the vertex-weighting on $J^k$: owing to the independence of the correlations, the probability $P(u)$ of a joint event $u = (u_1, \ldots, u_k)$ is given by the product

$$P(u) = P(u_1) \cdots P(u_k). \tag{5.3}$$

Note that the OR product of graphs does not retain information about which individual graph(s) gives rise to each edge in the product graph. As a result, $J^k$ fails to encode which of the $k$ original exclusivity graphs determine exclusivity between a given pair of joint events.

## 5.3 Refining the joint exclusivity graph

We introduce a fine-grained version of the joint exclusivity graph, using colors to track the source of exclusivity between pairs of joint events. We then show how each monochromatic subgraph inside this new structure relates to the corresponding exclusivity graph in $\{G_1, \ldots, G_k\}$, and we explain how Ramsey theory constrains these monochromatic subgraphs, and so the original exclusivity graphs as well.

### 5.3.1 Joint multigraph of independent boxes

We improve upon the construction of the joint exclusivity graph $J^k$ to encode which individual graphs contribute to exclusivity between joint events. To this end, we introduce a refined construction, the *multicolor product* of graphs, producing a joint exclusivity structure $\Gamma^k$ that is an edge-colored multigraph. The key idea is that we assign a unique color $i$ to the edges arising from each original graph $G_i$. A pair of nodes may thus be connected by multiple edges of different colors, with each color appearing at most once and indicating that the two joint events are exclusive due to the corresponding component. Figure 24 illustrates this construction with a toy example.

Formally, given exclusivity graphs $G_1, \ldots, G_k$, the joint exclusivity multigraph is given by their multicolor product $\Gamma^k = \Gamma^k(G_1, \ldots, G_k)$, defined as follows. The vertex set $V(\Gamma^k)$ is exactly the same as that of $J^k$:

$$V(\Gamma^k) = V(G_1) \times \cdots \times V(G_k). \tag{5.4}$$

The crucial difference lies in the edge structure. Between any two nodes $u, v \in V(\Gamma^k)$, representing joint events, there is a separate edge for each individual graph $G_i$ which has an edge between the corresponding components $u_i$ and $v_i$. Hence, the edge set $E(\Gamma^k)$ is partitioned into color classes:

$$E(\Gamma^k) = \bigcup_{i=1}^{k} E_i(\Gamma^k) \tag{5.5}$$

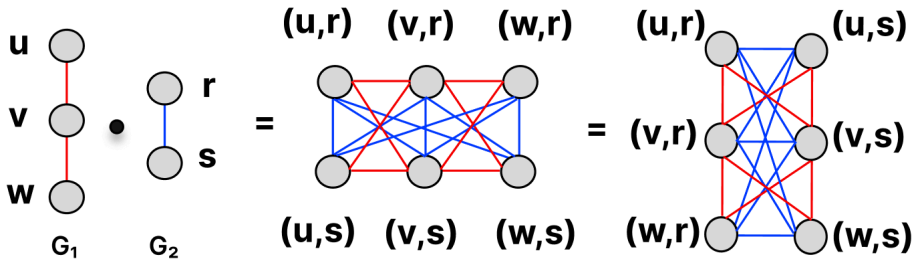



Figure 24: An example illustrating multicolor product of two graphs $G_1$ and $G_2$. The multigraph obtained can be thought to have multiple layers of nodes where each layer is isomorphic to $G_1$ (resp. $G_2$) with one of the colors, while the graph obtained by picking one node from each layer and the edges of the other color is necessarily isomorphic to $G_2$ (resp. $G_1$) as shown in the middle (resp. rightmost) representation above.

where each $E_i(\Gamma^k)$ is the set of edges with color $i$, which are derived from the $i$th individual graph $G_i$:

$$E_i(\Gamma^k) = \{(u, v)_i \mid (u_i, v_i) \in E(G_i)\}. \tag{5.6}$$

Here, $(u, v)_i$ denotes an edge of color $i$ between vertices $u$ and $v$. The relationship to the usual joint exclusivity graph $J^k$ is given by

$$E(J^k) = \{(u, v) \mid \exists i \in \{1, \ldots, k\}. (u, v)_i \in E(\Gamma^k)\}. \tag{5.7}$$

Of course, the vertex-weighting indicating probability of joint events remains as for $J^k$, given by Eqn. 5.3.

To test the E-principle on $\Gamma^k$, one seeks sets of nodes such that every pair of nodes in the set is connected by at least one colored edge. Such a set of nodes forms an edge-colored clique in $\Gamma_k$, with the possibility of more than one edge (of different colors) between each pair of nodes. The corresponding events are pairwise exclusive, hence form a clique in $J^k$. Thus, as before, a violation of the E-principle occurs when the sum of probabilities of the events in such an (edge-colored) clique exceeds $1$.

We remark that edge-colored multigraphs have been previously employed in a similar setting but with a different intention in Refs. [89, 90, 91], where nodes represent events contributing to a given Bell inequality and edges denote exclusivity, with the color indicating the party from which such exclusivity arises.

### 5.3.2 Monochromatic projection of $\Gamma^k$

In Sec. 5.3.3, we will see how Ramsey theory necessitates the existence of certain monochromatic subgraphs inside sufficiently large cliques of $\Gamma^k$. To facilitate that analysis and to draw meaningful conclusions from it, it is crucial to formally define the monochromatic projections of $\Gamma^k$ and to study the connections between such projections and the individual exclusivity graphs $G_i$.

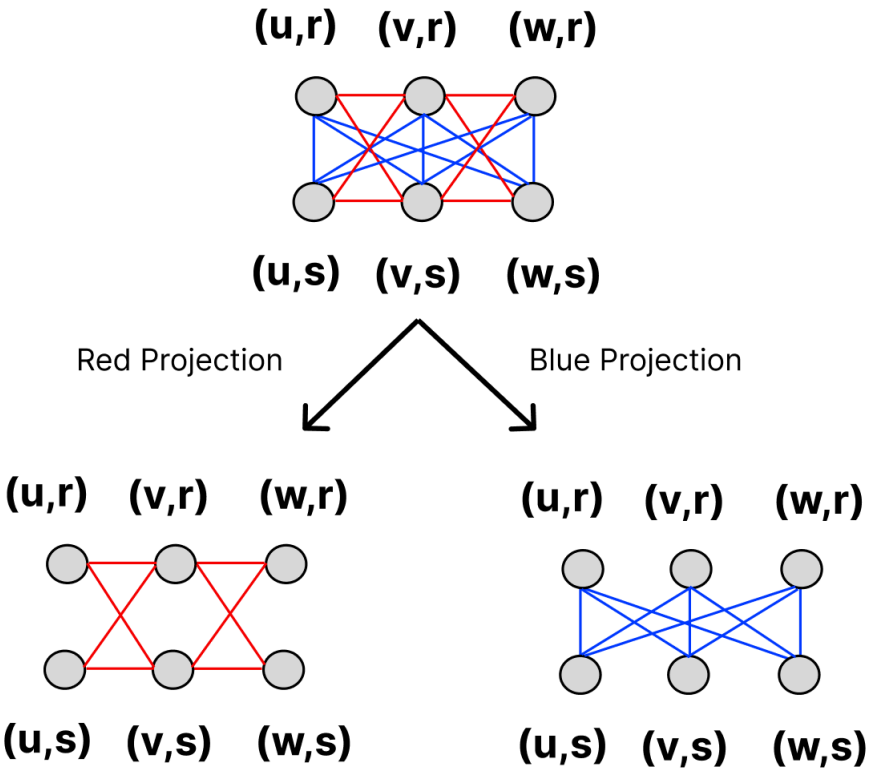



Figure 25: The joint exclusivity multigraph constructed in Figure 24 and its monochromatic projections.

For each color $i \in \{1, \ldots, k\}$, the *monochromatic projection* of color $i$ is defined as $H_i(\Gamma^k) = (V(\Gamma^k), E_i(\Gamma^k))$, i.e. $H_i(\Gamma^k)$ is the subgraph of $\Gamma^k$ with the same set of nodes but containing only the edges of color $i$. Thus, it tracks only the exclusivity relations between joint events that are derived from the exclusivity graph $G_i$ of the $i$th correlation. Figure 25 illustrates the two monochromatic projections of the toy example multigraph from Figure 24.

To follow the proofs of our results, it is imperative to highlight the precise relationship between the monochromatic projections $H_i(\Gamma^k)$ and the original exclusivity graphs $G_i$.

The first observation is that one can partition the nodes of $\Gamma^k$, and hence of $H_i(\Gamma^k)$, according to their $i$th component, i.e. by the equivalence relation whereby $u, u' \in V(\Gamma^k)$ are deemed equivalent if and only if $u_i = u'_i$. For each node $v \in V(G_i)$, write

$$V_v^i = \{u \in V(\Gamma^k) \mid u_i = v\} \tag{5.8}$$

for the set of joint events (nodes of $\Gamma^k$) whose $i$th component is the fixed event $v$. In other words, $V_v^i = V(G_1) \times \cdots \times V(G_{i-1}) \times \{v\} \times V(G_{i+1}) \times \cdots \times V(G_k)$. These sets partition $V(\Gamma^k)$:

$$V(\Gamma^k) = \bigcup_{v \in V(G_i)} V_v^i. \tag{5.9}$$

Moreover, all the nodes in each of these sets are neighborhood-equivalent in $H_i(\Gamma^k)$, the monochro-

matic projection of color $i$: if $(v, w) \in E(G_i)$ then every node in the class $V_v^i$ is connected to every node in the class $V_w^i$ by an edge of color $i$ in $\Gamma^k$, while if $(v, w) \notin E(G_i)$ then no nodes of $V_v^i$ are connected to any nodes of $V_w^i$ by an edge of color $i$ in $\Gamma^k$. In the example from Figure 25, the equivalence classes with respect to the red graph $G_1$ are $V_u^1 = \{(u, r), (u, s)\}$, $V_v^1 = \{(v, r), (v, s)\}$, and $V_w^1 = \{(w, r), (w, s)\}$, while the equivalence classes with respect to the blue graph $G_2$ are $V_r^2 = \{(u, r), (v, r), (w, r)\}$ and $V_s^2 = \{(u, s), (v, s), (w, s)\}$.

The following proposition is an immediate consequence of the structure just observed.

**Proposition 13.** *The monochromatic projection $H_i(\Gamma^k)$ contains a subgraph isomorphic to $G_i$.*

*Proof.* For each $v \in V(G_i)$, pick a node $\tilde{v} \in V(\Gamma^k)$ from the corresponding equivalence class $V_v^i$, i.e. such that $\tilde{v}_i = v$. The induced subgraph of $H_i(\Gamma^k)$ determined by these nodes is isomorphic to $G_i$ since $(\tilde{v}, \tilde{w})_i \in E(\Gamma^k)$ if and only if $(v, w) \in E(G_i)$. □

The central observation that underlies most of our results is expressed by the following proposition relating monochromatic odd cycles in $\Gamma^k$ with odd cycles in the corresponding original graph.

**Proposition 14.** *For each odd cycle in $H_i(\Gamma^k)$, there is an odd cycle in $G_i$ of at most the same length.*

*Proof.* Let $C$ be an odd cycle in $H_i(\Gamma^k)$. If all the nodes in $C$ belong to different equivalence classes, then their $i$th components all differ and thus form a cycle in $G_i$. Otherwise, let $v$ and $v'$ be two nodes in the same equivalence class, i.e. such that $v_i = v'_i$. We show that, in that case, there exists a strictly smaller odd cycle in $H_i(\Gamma^k)$. Fig.26(c) illustrates the basic idea of the proof. Since nodes in the same equivalence class cannot be adjacent in $H_i(\Gamma^k)$, $v$ and $v'$ cannot be consecutive nodes in the cycle $C$. Let $u$ and $w$ be the two neighboring nodes of $v$ in the cycle $C$. Being in the same equivalence class as $v$, the node $v'$ must also be connected to both $u$ and $w$. Consider two paths starting from $v'$ and going in opposite directions along the cycle $C$: the path from $v'$ to $u$ and the path $v'$ to $w$, both avoiding the node $v$ (see Figure 26(a)). Since $C$ is an odd cycle, then one of these paths must necessarily have even length. Adjoining to it the edge $(u, v')$ or the edge $(w, v')$, we obtain an odd cycle. This cycle has strictly smaller length than $C$, for it is formed by a proper subset of its nodes having dropped at least the node $v$.

This process can be repeated until one reaches an odd cycle each of whose nodes belongs to a different equivalence class. The fact that the length strictly decreases at each step guarantees termination, with the limit case being a cycle of length $3$, where each node must belong to a different equivalence class. □

Note that a similar proof strategy fails for even cycles because drawing an edge between two nodes in an even cycle leads to both partitions having even length or both having odd length, as shown respectively

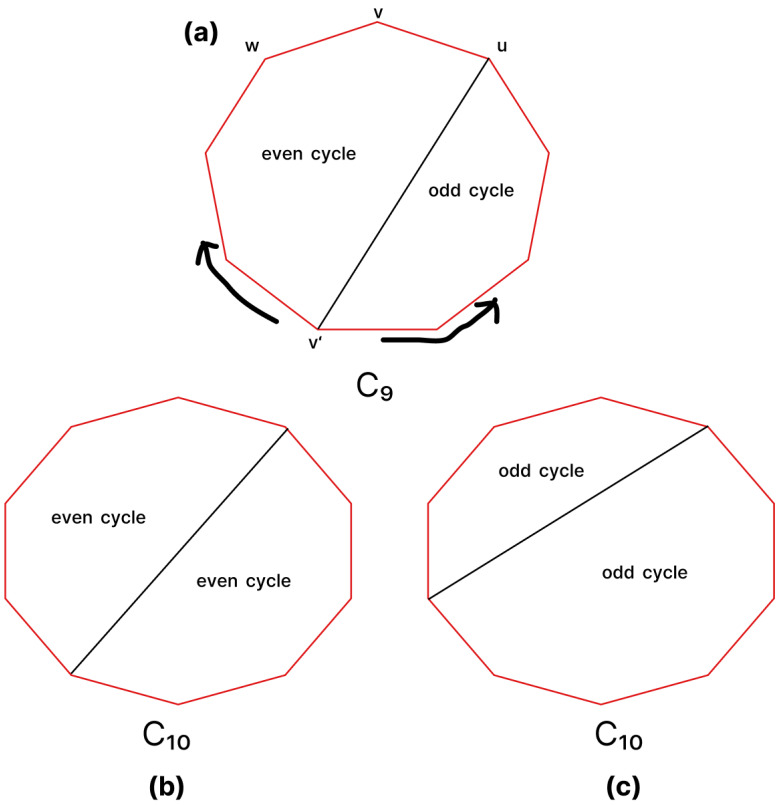



Figure 26: An edge between non-adjacent nodes in a cycle divides it into two smaller cycles. If the original cycle is odd, a chord necessarily leads to (a) an even and an odd cycle. If the original cycle is even, a chord can divide the cycle into (b) two even cycles or (c) two odd cycles.

in Figure 26(b) and Figure 26(c). By contrast, in the case of an odd cycle, one partition is guaranteed to have odd length, as shown in Figure 26(a), guaranteeing that the iterative process eventually reaches an odd cycle each of whose nodes belongs to a unique equivalence class, which in turn implies the existence of a cycle of the same length in $G_i$.

**Corollary 15.** *The length of the smallest odd cycle in $H_i(\Gamma^k)$ is equal to the length of the smallest odd cycle in $G_i$.*

*Proof.* By Proposition 14, the length of the smallest odd cycle in $G_i$ is at most that of the smallest odd cycle of $H_i(G)$. On the other hand, by Proposition 13, $H_i(\Gamma^k)$ contains a cycle that is isomorphic to the smallest odd cycle of $G_i$. □

Notice that in establishing the above relationships between $H_i(\Gamma^k)$ and $G_i$ we made no assumptions about the exclusivity graphs $G_1, \ldots, G_k$. Thus, the results hold in full generality, for any correlation in any measurement scenario. Later, we make use of this fully general connection to explore the specific case of $n$-cycle PR boxes. We now explain how violations of the E-principle (or lack thereof) for a joint exclusivity graph can be studied through the lens of Ramsey theory.

### 5.3.3 Ramsey-theoretic constraints

Ramsey theory is a branch of combinatorics that studies the emergence of regularity in arbitrary partitions of (large enough) combinatorial structures. A typical question considers a structure that is cut into a set number of pieces and asks how large the structure must be to guarantee that, regardless of the partition, at least one of the pieces contains a certain pattern.

The prototypical Ramsey-theoretic problem problem – which, along with minor variations, underlies our results – can be formulated in the language of graph theory. It considers colorings of the edges of a complete graph and asks for the minimal size that guarantees the presence of specific monochromatic substructures. More precisely, given $k$ graphs $S_1, \ldots, S_k$, where each $S_i$ is associated with a color $i \in \{1, \ldots, k\}$, the Ramsey number $R(S_1, \ldots, S_k)$ denotes the minimum size of a complete graph such that any $k$ edge coloring contains a monochromatic $S_i$ in color $i$ for some $i$. A simple example is $R(K_3, K_3) = 6$: as already mentioned, any way of coloring the edges of $K_6$ must produce a monochromatic triangle, but the same is not true for $K_5$. A more involved example is $R(K_3, K_3, K_4) = 30$ [92]: if three colors are available, then $K_{30}$ is the smallest complete graph such that any edge coloring contains a $K_3$ of color $1$ or $2$ or a $K_4$ of color $3$. Reading it the other way round, once we know a certain Ramsey number

$R(S_1, \ldots, S_k)$, we can conclude that for any $k$ edge-colored clique of size at least $R(S_1, \ldots, S_k)$ there is at least one color $i \in \{1, \ldots, k\}$ whose monochromatic projection contains $S_i$.

Since a multigraph allows for more edges than a graph with the same nodes, Ramsey-theoretic constraints $(S_1, \ldots, S_k)$ that apply to any $k$ edge coloring of a complete graph also apply to a $k$ edge coloring of a multigraph of the same size. Therefore, for our analysis, we do not necessarily need new Ramsey-theoretic results specific for multigraphs in order to squeeze out conclusions applicable to multigraphs like $\Gamma^k$. In fact, the Ramsey theory of multigraphs is not sufficiently developed in the literature. Our work is therefore based only on Ramsey-theoretic results established for ordinary graphs.

In our analysis, this theoretical framework becomes powerful because it constrains the possible structures within joint exclusivity graphs, allowing one to rule out violations of the exclusivity principle. As established earlier in Sec. 5.3.1, a violation of the E-principle corresponds to certain cliques inside the edge-colored multigraph $\Gamma^k$. Ramsey-theoretic results force monochromatic substructures to be present inside the cliques being sought for violation. That is, the existence of violation-producing cliques within $\Gamma^k$ implies certain constraints on the monochromatic projections $H_i(\Gamma^k)$. Guaranteeing the absence of the required substructures within the corresponding projections immediately precludes the existence of any violation-producing clique in $\Gamma^k$, ruling out the possibility of violating the E-principle. This forms the logical basis of our results and also establishes a general connection between Ramsey theory and activation effects of the E-principle. Concretely, in our proofs, we use a few Ramsey-theoretic results that necessitate the presence of odd cycles in the monochromatic projections $H_i(\Gamma^k)$. Proposition 14 then implies a direct constraint on the original exclusivity graph $G_i$, which we can explicitly rule out.

Notably, while this approach is powerful for proving impossibility results, establishing the existence of E-principle violations through Ramsey theory is more challenging. To prove impossibility, it suffices to show inconsistency between the monochromatic projections and the sought-after clique with just one known Ramsey-theoretic result. On the other hand, merely being consistent with Ramsey-theoretic constraints does not in itself guarantee the existence of a violation.

## 5.4 Exclusivity graph of n-cycle PR boxes

In Sec. 2.3.2, we explained the most fundamental KS scenarios called $n$-cycle scenarios. For the purpose of this work, we are interested in the non-classical vertices of the no-disturbance polytopes of $n$-cycle scenarios, also explained in Sec. 2.3.2, called $n$-cycle PR boxes.

### 5.4.1 Violation condition

In the case of (products of) $n$-cycle PR boxes, the problem of E-principle violation admits an especially simple description. This is due to the fact that in any $n$-cycle PR box each possible event occurs with the same probability, $\frac{1}{2}$. Consequently, when dealing with $k$ independent copies of such boxes, by Eqn. 5.3 each event in the joint exclusivity graph occurs with probability $\frac{1}{2^k}$.

Given a clique of size $l$ in such a joint exclusivity graph, the condition for it to violate the E-principle then becomes:

$$\frac{l}{2^k} > 1. \tag{5.10}$$

Therefore, the existence of a clique of size $2^k + 1$ inside the joint exclusivity graph of $k$ independent $n$-cycle PR boxes guarantees violation of the E-principle.

## 5.5 Result highlights

So far, we have established a connection between Ramsey theory and activation effects of the E-principle by refining the conventional method for studying the latter. This refinement consists in introducing edge colors in the joint exclusivity graph of independent correlations to track which components contribute to the exclusivity between any pair of joint events. Witnessing violations of the E-principle then amounts to finding cliques within the resulting edge-colored multigraph where the probabilities of the corresponding events sum to more than $1$. Ramsey theory forces the existence of certain monochromatic structures within sufficiently large edge-colored cliques, and can thus be leveraged to study activation effects.

We now exploit the power of this general connection in the specific setting of $n$-cycle PR boxes. In this section, we highlight the key results and the methods used to obtain them, leaving detailed proofs for Sec. 5.6.

We fix the number of copies $k$ for which we explore activation of the E-principle by $n$-cycle PR boxes. Recall that for $k$ independent copies of an $n$-cycle PR box, the condition for violating the E-principle is the existence of a $K_{2^k+1}$ within the edge-colored joint exclusivity multigraph $\Gamma^k$. In Sec. 5.6.2 we show that one can easily construct cliques of size $2^k$ within $\Gamma^k$, but that these specific cliques can never be extended to one of size $2^k + 1$, underlining the fact that finding a $K_{2^k+1}$ is not a straightforward task.

Table 2 summmarizes the results we obtained relating $k$ and $n$, which we detail in the following subsections. Note that a violation of the E-principle with $k$ copies guarantees a violation with $k + 1$ copies, as highlighted in Remark 20.

### 5.5.1 k = 1

For every $n$-cycle scenario, as for any bipartite Bell scenario, it is known that all the extremal vertices of the no-disturbance polytope satisfy the E-principle (with a single copy). Since each possible event in an $n$-cycle PR box occurs with probability $\frac{1}{2}$, violating the E-principle would require the existence of a $K_3$ within the exclusivity graph of these extremal correlations. Corollary 17 highlights the impossibility of such a $K_3$. Therefore, activation effects become relevant for these scenarios.

### 5.5.2 k = 2

We show in Theorem 21 that, except for the cases $n = 4$ and $n = 5$, no other $n$-cycle PR box violates the E-principle with two copies.

Recall that a $K_5$ within the joint exclusivity graph $\Gamma^2$ is required to exhibit a violation of the E-principle with two copies. The proof of this theorem exploits the fact that the edges of a $K_5$ graph can be divided into two disjoint copies of $C_5$, i.e. two cycles of length $5$; see Figure 30. For $n = 4$ or $n = 5$, each of these $5$-cycles can be realized independently within the exclusivity graph of a single $n$-cycle PR box. However, it is not possible to find a $C_5$ within the exclusivity graph of an $n$-cycle PR box when $n \geq 6$, as shown in Corollary 19. Our proof goes further to show that there is no other way to come up with a $K_5$ within $\Gamma^2$ for $n \geq 6$.

$K_5$ turns out to be a small enough clique that it can be explored analytically on a case-by-case basis to reach our results about $k = 2$ for any $n$. Still, a known Ramsey-theoretic result highlights an interesting fact about the $n = 4, 5$ cases, while another less well-known Ramsey result corroborates the result we obtained for $n \geq 6$, providing a neat alternative perspective.

The first of these is the classical Ramsey theory example $R(C_3, C_3) = 6$. It implies that if a $K_6$ exists within $\Gamma^2$, one of the monochromatic projections – and so, by Corollary 15, also the exclusivity graph of a single $n$-cycle PR box – must contain a $C_3$ cycle. In turn, Corollary 17 tells us that this is impossible for any $n$-cycle scenario. This means that two copies of $n = 4, 5$-cycle PR boxes can only violate the E-principle minimally, i.e. with a $K_5$. This aligns with the computational findings in Ref. [22], where the search for E-principle violations in the case $k = 2$ and $n = 4$ (two copies of the CHSH PR box) turned up only cliques of size $5$. This Ramsey-theoretic result explains why.

To corroborate our result for two copies of $n \geq 6$-cycle PR boxes one can invoke a result by Erdös et al. [93], from one of the first papers that studied Ramsey numbers for more than two colors. In this work, the authors define a new type of Ramsey number, $R(\leq [C])$, better explained in the proof of Theorem 24

| k | n |
|---|---|
| 1 | 4 5 6 7 8 ... |
| 2 | 4 5 6 7 8 ... |
| 3 | 4 5 6 7 8 ... |
| 4 | 4 5 6 7 8 ... 17 18 19 ... |
| l ($\geq 5$) | 4 5 6 7 8 ... 17 18 19 ... $2^l + 1$ $2^l + 2$ ... |

Table 2: Relationship between $k$ and $n$ for the activation effects of E-principle with $n$-cycle PR boxes. For a given $k$ number of independent copies on the left: on the right, red marks that the $n$-cycle PR box does not violate the E-principle, blue indicates that there is a violation, while gray indicates no knowledge of violation hence an open problem for future work. Results known prior to this work were only those in the first column ($n = 4$).

ahead. A particular instance is $R(\leq (C_5, C_3)) = 5$; it means that any two edge-colored $K_5$ must contain a monochromatic odd cycle of length $3$ or $5$. Corollaries 17 and 19 inform us that this is impossible for $n \geq 6$-cycle PR boxes, corroborating the result we obtain in Theorem 21. This makes $n \geq 6$-cycle boxes the first reported instances in the E-principle literature of correlations that do not violate the E-principle with only two copies.

### 5.5.3 k = 3

In Theorem 24, we show that, for $n \geq 6$, $n$-cycle PR boxes do not violate the E-principle with three copies. On the other hand, for $n = 4, 5$, the known two-copy violation implies a three-copy one, as per Remark 20. This completely characterizes the case $k = 3$ of three copies of $n$-cycle PR boxes.

Recall that a $K_9$ within $\Gamma^3$ is required to exhibit a violation of the E-principle with three copies. In contrast with the previous $k = 2$ setting, a case-by-case analysis becomes impractical. Our proof relies on a Ramsey-theoretic result to rule out the existence of a $K_9$ within $\Gamma^3$. A result from Ref. [93] asserts that $R(\leq (C_5, C_5, C_5)) = 9$, which means that any three edge-coloring of $K_9$ contains a monochromatic odd cycle of length $3$ or $5$. A $K_9$ within $\Gamma_3$ would then imply the existence of an odd cycle of length $3$ or $5$ in at least one of its monochromatic projections, and consequently, by Corollary 15 on the exclusivity graph of an individual $n$-cycle PR box. This, in turn, is impossible for $n \geq 6$ due to Corollaries 17 and 19.

### 5.5.4 k > 3

Finally, we obtain a general result applicable to an arbitrary number of copies $k$, bounding the size $n$ of $n$-cycle PR boxes that may witness a $k$-copy activation effect. Specifically, in Theorem 25 we show that for $n \geq 2^k + 2$, $n$-cycle PR boxes do not violate the E-principle with $k$ copies. Since we have already fully characterized the cases $k = 2, 3$, this result becomes relevant for $k \geq 4$. This leaves open the range $6 \leq n \leq 2^k + 1$, for which we do not know whether $k$ copies suffice to violate the E-principle.

Recall that a $K_{2^k+1}$ within $\Gamma^k$ is required to exhibit violation of E-principle with $k$ copies. Our proof applies a result mentioned in the introduction of Ref. [94], a recent work that explored multi-colored Ramsey numbers for odd cycles. It states that a $k$ edge-colored $K_{2^k+1}$ must contain a monochromatic odd cycle. We then leverage Theorem 18, which provides the size of the smallest odd cycle in the exclusivity graph of any $n$-cycle PR box, to obtain the bound.

## 5.6 Proofs

We now provide the proofs of the results highlighted in the previous section.

### 5.6.1 Exclusivity graphs of n-cycle PR boxes

We start with the characterization of the exclusivity graphs of $n$-cycle PR boxes. As mentioned in Section 5.4, these turn out to be well-studied graphs: the exclusivity graph of an $n$-cycle PR box takes the form of a Möbius ladder graph ($M_{2n}$) when $n$ is even and of a Prism graph ($Y_n$) when $n$ is odd. Figure 27 illustrates these two situations. Note that the standard notation for Möbius ladder graph, $M_{2n}$, indicates its order (total number of vertices) in the subscript, while the standard notation for the circular ladder graph, $Y_n$, indicates only half of its order.

**Theorem 16.** *The exclusivity graph of an $n$-cycle PR box is isomorphic to the Möbius ladder graph $M_{2n}$ when $n$ is even and to the circular ladder graph $Y_n$ when $n$ is odd.*

*Proof.* Consider a specific starting configuration of the exclusivity graph of an $n$-cycle PR box. The configuration starts from the possible events of first context at the top and go down to the final context at the bottom in the cyclic order of contexts. For a given context, if the possible events are perfectly correlated then the node on the left will denote the outcome $++$ while the one on the right denotes outcome $--$ ; while when the possible events are anti-correlated, the node on the left will denote outcome $+-$ while

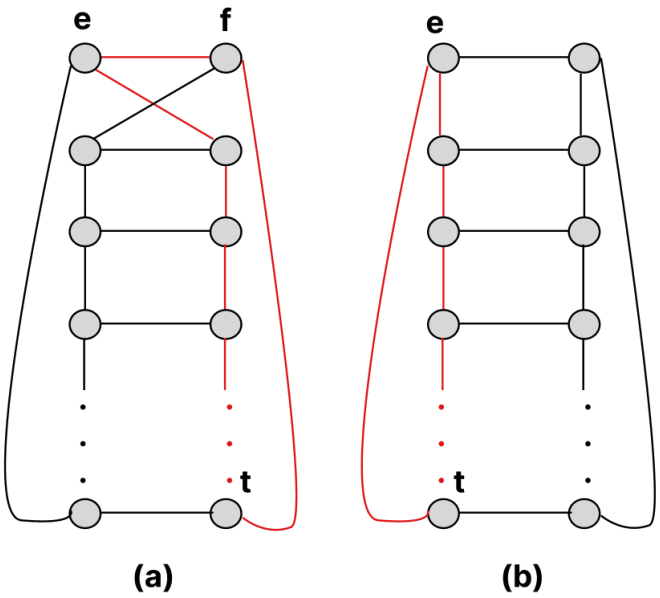



Figure 27: Exclusivity graphs of the $n$-cycle PR boxes: (a) when $n$ is even: Möbius ladder graph ($M_{2n}$) (b) when $n$ is odd: circular ladder graph, also popularly known as Prism graph ($Y_n$). Drawn in red are the smallest odd cycles for the two cases where $e, f, t$ are the events used in Theorem 18 to illustrate the construction of these cycles.

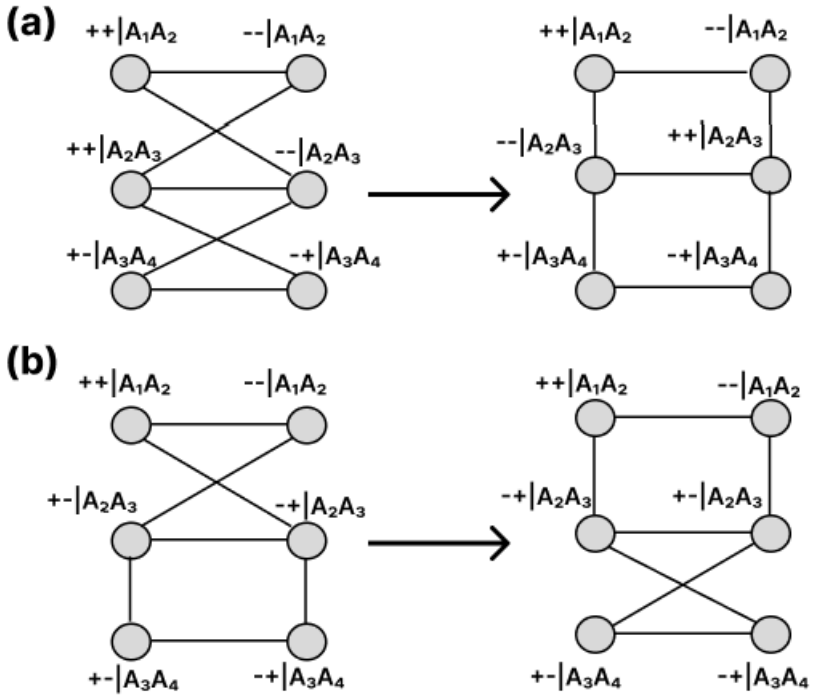



Figure 28: Interchanging nodes of the middle context (a) both consecutive criss-cross connections become vertical (b) a cross connection and a vertical one beneath it turn into vertical and cross connections respectively.

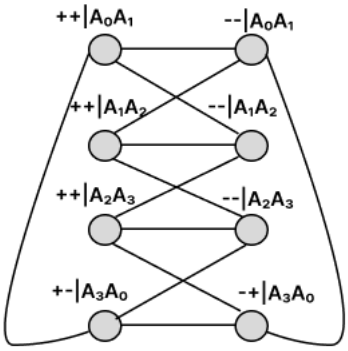



Figure 29: Reference configuration for the exclusivity graph of PR-box. Positively correlated events of a given context, when the joint outcome $++$ is on the left, connect in a cross-like manner to the context under it while negatively correlated events of a given context, when the joint outcome $+-$ is on the left, arrange in a vertical manner to the context under it.

the one on right will denote outcome $-+$. Figure 29 illustrates this reference configuration for a 4-cycle PR-box. Once this arrangement is fixed we know how nodes of a given context connect with the nodes of the context underneath it. Note that since the ordering of the contexts is cyclic, a given context has a measurement in common only to a context before and after it. Hence, no exclusivity relation exists between any nodes beyond the adjacent contexts in the cycle. Within this configuration, the nodes corresponding to perfectly correlated events of a given context connect to the nodes of the context (cyclically) underneath it in a cross like manner irrespective of whether the context has correlated or uncorrelated events. On the other hand, when the nodes, of a given context, correspond to anti-correlated events they connect with the context (cyclically) beneath it in a vertical manner irrespective of whether the context has correlated or uncorrelated events.

With this initial configuration of the exclusivity graph in hand, we now try to convert the cross connections into vertical ones by interchanging the nodes within contexts, in a sequential manner, to redraw it – hence exhibit isomorphism – to the Möbius ladder graph and the circular ladder graph for $n$ even and odd cases respectively. Note the obvious but important fact that each cross connection is followed either by a cross connection or a vertical one in this configuration, as shown on the left sides of Figure 28(a) and Figure 28(b).

Let us start from the very first pair of contexts leading to a cross connection in this initial configuration. To convert this cross connection into a vertical one, we interchange the nodes in the (cyclically) bottom context (i.e. events of the bottom context in the pair) while keeping the nodes of the first context as they were. This means that, if before this interchange, there was a cross connection right underneath the first cross connection (l.h.s of Figure 28(a)) then after the interchange there are only vertical connections across

the three contexts as is shown in the r.h.s of Figure 28(a). On the other hand if before the interchange there was a vertical connection beneath the first cross connection (l.h.s of Figure 28(b)) then after the interchange this vertical connection converts into a cross connection as shown in the r.h.s of Figure 28(b).

Overall, in the first case we get rid of the two cross connections, converting them into vertical ones, while in the latter case it is as if the first cross connection has been transported (cyclically) downwards into the graph. Note that when $n$ is even there will be odd number of cross connections in the initial configuration (since by definition there are odd positively correlated contexts) while when $n$ is odd there are even number of cross connections in the initial configuration. Repeating the interchange process described above sequentially leads to having only one cross connection when $n$ is even while all cross connections will be transformed into vertical ones, when $n$ is odd. Overall, this leads to a Möbius ladder graph ($M_{2n}$) when $n$ is even, see Figure 27(a), while a circular ladder graph ($Y_n$) (also known as Prism graph), when $n$ is odd, see Figure 27(b). □

A single copy of an $n$-cycle PR box satisfies E-principle. This is because for an $n$-cycle PR box violation of E-principle requires existence of $K_3$ inside the exclusivity graph. This is not possible as captured by the following Corollary of the above theorem.

**Corollary 17.** *The exclusivity graph of an $n$-cycle PR-box does not contain a triangle ($K_3$) for any $n \geq 4$.*

*Proof.* As can be seen in Figure 27, in each graph $M_{2n}$ and $Y_n$ (for any $n$), the nodes in the neighbourhood of a given node are not connected with each other. Therefore, there is no $K_3$ possible inside these graphs. □

Now that we know how the exclusivity graphs of $n$-cycle scenarios look like we are ready to define the most relevant property of these graphs for our proofs later. The property is the size of the smallest odd cycle inside these graphs. The following theorem captures that.

**Theorem 18.** *The smallest odd cycle in the exclusivity graph of an $n$-cycle PR box has length $n+1$ when $n$ is even and length $n$ when $n$ is odd.*

*Proof.* We first show an explicit construction of the respective odd cycles mentioned in the theorem statement and then provide a unified proof for the two cases exhibiting why there can't be a smaller odd cycle in the two cases.

Explicit construction of respective odd cycles: pick a path from the first context to the last where only one node is picked from each context. Say you pick event $e$ from the first context and end up at some event $t$ in the last context from top to bottom. Now, to close this path into a cycle we need to go back

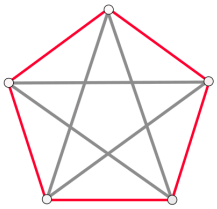


Figure 30: Edge bifurcation of a $K_5$ exclusivity graph into two edge-wise disjoint $C_5$ graphs. The outer one denoted in red and the inner one in grey.

from event $t$ to the first context. When $n$ is even we can return back to event $f \neq e$ in the first context (because event $t$ is connected to $f$), hence giving an odd cycle of size $n+1$, see Figure 27(a). In the case when $n$ is odd, event $t$ straight away connects with event $e$ and hence we get an odd cycle of size $n$, see Figure 27(b).

Now we show that a smaller odd cycle can't exist in the two cases. Note that for a smaller odd cycle to exist we can't use events from all the contexts. More explicitly, when $n$ is even i.e. the graph is $M_{2n}$, a smaller odd cycle would have a size less than $n+1$, which means the biggest such cycle can only be of size $n-1$. When $n$ is odd, the biggest such odd cycle would be of size $n-2$. Since there are $n$ contexts, we cannot pick at least one event from each context to construct such odd cycles. This means that such an odd cycle has to live inside a sub graph utilizing less than $n$ contexts in both the cases. This further implies that such a sub graph would either be a ladder graph or a sub graph of a ladder graph for both cases. It is well known that a ladder graph (hence any of its sub graph as well) is a bi-partite graph. Furthermore, it was first shown by König that a graph is bi-partite if and only if it contains no odd cycle [95, 96]. This proves that an odd cycle less than $n+1$ for $M_{2n}$ when $n$ is even and less than $n$ for $Y_n$ when $n$ is odd can't exist. This finishes our proof. □

**Corollary 19.** *The exclusivity graph of an n-cycle PR-box contains a* $5$*-cycle* $C_5$ *only if* $n = 4$ *or* $n = 5$.

*Proof.* From Theorem 18, when $n = 4$ or $n = 5$, the smallest odd cycle has size 5. For $n \geq 6$, it is always larger than 5. □

### 5.6.2 Impossibility of extending a trivial $K_{2^k}$ to $K_{2^k+1}$

Before we move on to studying activation effects of $n$-cycle PR boxes we eliminate a trivial possibility of violation of E-principle. We know that for an $n$-cycle PR box, the condition of violation of E-principle is the existence of $K_{2^k+1}$ inside the $k$ edge-colored joint exclusivity multigraph ($\Gamma^k$). Here we show that there

always exists a $K_{2^k}$ inside the $k$ edge-colored $\Gamma^k$ that can't be extended to a $K_{2^k+1}$ supporting the fact that finding a $K_{2^k+1}$ isn't a straightforward task.

It is easy to visualize that OR product of two complete graphs $K_n$ and $K_m$ is a complete graph $K_{m+n}$. This implies that inside $\Gamma^k$ exists sub graph $K_{2^k}$ which can be seen as an OR product of $k$, $K_2$ graphs. Such a $K_{2^k}$, let's call it $T$, can never be extended to a $K_{2^k+1}$ inside $\Gamma^k$. We now prove this claim: Let us see if some node $v$ exists in $V(\Gamma^k) \setminus V(T)$ such that $v$ and $V(T)$ together produce a clique of size $2^k+1$. Let us say that $v$ is exclusive to an event $e \in V(T)$ due to its $lth$ component i.e. $(e_l, v_l) \in E(G)$. For each event in $V(T)$ the $lth$ component is either $e_l$ or $e_l^\perp$. In fact, half of the events in $V(T)$ have $e_l$ as their $lth$ component while others have $e_l^\perp$. If $(e_l, v_l) \in E(G)$ and we know, by construction that $(e_l, e_l^\perp) \in E(G)$, then $(e_l^\perp, v_l) \notin E(G)$ because there can't be any $K_3$ in $G$ as shown in Corollary 17. Therefore, a connection due to the $lth$ component would divide the set of events in $V(T)$ into two equal subsets, say $A$ and $B$ each of size $2^{k-1}$, where all events in $A$ are connected with $v$ due to the $lth$ component being $e_l$ while $v$ can only connect with the nodes in $B$ because of components other than the $lth$ one. Now, say $v$ connects with a one node in $B$ due to the $mth$ component where $m \neq l$. Due to this the nodes within $B$ now get divided into two equal disjoint subsets with each subset having size $2^{k-2}$. Progressing this way, for each new component of $v$, we are left with a subset of nodes which are unconnected to $v$ due the current component under consideration with half the size of the previous set. Therefore, when we have utilized all the components of $v$, we will be left with a subset of size $2^{k-k} = 1$ that will remain disconnected from $v$ suggesting that we can never form a $K_{2^k+1}$ using $v$ and $V(T)$. Therefore, such a $K_{2^k}$ can't be extended to $K_{2^k+1}$.

We now state a trivial consequence of the fact that an OR product of $K_m, K_n$ is a $K_{m+n}$ in the form of a remark.

*Remark* 20. Given a clique of size $s$ in the joint exclusivity graph of $k$ copies $\Gamma^k$, one can find a clique of size $2s$ inside $\Gamma^{k+1}$.

This holds because one can do an OR product of $K_s$ inside $\Gamma^k$ with a $K_2$ inside the $(k+1)$th copy of the original graph. This produces a $K_{2s}$ inside $\Gamma^{k+1}$.

### 5.6.3 Activation effects of $n$-cycle PR boxes

With the trivial case out of our way we are now ready to provide our results concerning violation of E-principle for different $k$. Simplest case being $k = 2$ where a $K_5$ is desired inside $\Gamma^2$ to witness violation of E-principle.

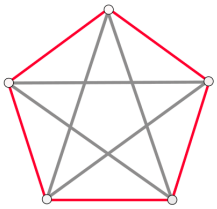


Figure 31: Constructing a $K_5$ using two independent $C_5$ cycles via multicolor product: (a) events coming from PR-box whose exclusivity graph is colored in red (b) events coming from PR-box whose exclusivity graph is colored in grey (c) joint events from the two copies leading to a $K_5$.

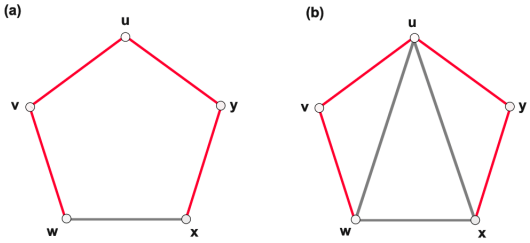



Figure 32: Impossibility of extending a $C_5$ in (a) with one edge supported by $n$-cycle PR-box whose exclusivity graph is colored grey while the rest of the edges are supported by the $n$-cycle PR box whose exclusivity graph is colored red, to (b) a $K_5$ for all $n \geq 6$ cycle scenarios since it is inconsistent with Corollary 17.

**Theorem 21.** *The joint exclusivity graph of two independent $n$-cycle PR-boxes contains a $K_5$ when $n = 4$ or $n = 5$. That is, two copies activate violation of the E-principle for $4$- and $5$-cycle PR boxes.*

*Proof.* We start with a nice observation that all the edges of a $K_5$ can be divided into two disjoint subsets. The outer edge-induced graph, denoted with red in Figure 30, denotes a $C_5$ while the inner one, denoted with grey in Figure 30, is also isomorphic to $C_5$. The independence of events at each node of $K_5$ along with this edge based bifurcation of $K_5$, means that one can use Corollary 19 to ensure that the exclusivity in the outer and the inner $C_5$ graphs is realized independently by two different copies of the $n$-cycle PR-boxes when $n = 4, 5$. $C_5$ in Figure 31(a) denotes events $\{e_i\}_{i=1}^5$ coming from one copy while events $\{f_i\}_{i=1}^5$ in Figure 31(b) are coming from the other PR-box. One can then construct joint independent events $\{(e_i, f_i)\}$ whose exclusivity graph is a $K_5$ as denoted in Figure 31(c). This means E-principle will be violated for $n = 4, 5$ by just using two copies. Notice that Corollary 19 implies that such a construction of $K_5$ won't work for $n \geq 6$. We now show that there is also no other way to obtain a $K_5$ with two copies when $n \geq 6$. □

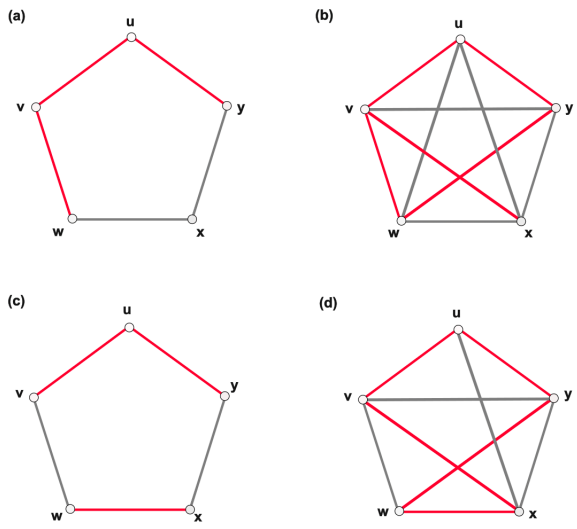



Figure 33: Impossibility of extending a $C_5$, with two edges supported by events from an $n$-cycle PR box whose exclusivity graph is colored grey while the other edges are supported by an $n$-cycle PR box whose exclusivity graph is colored red, to a $K_5$: (a) when the two grey edges are adjacent, the extension leads to violation of Corollary 17 as shown in (b) with nodes $\{u, w, x\}$; (c) when the two grey edges are not adjacent, the extension leads to violation of Corollary 19 due to a $C_5$ with nodes $\{u, v, x, w, y\}$ as shown in (d).

**Theorem 22.** *The joint exclusivity graph of two independent $n$-cycle PR-boxes does not contain a $K_5$ for $n \geq 6$. That is, two copies do not activate a violation of the E-principle for $n \geq 6$-cycle PR boxes.*

*Proof.* If a $K_5$ were to exist for two $n \geq 6$ PR-boxes, thanks to Corollary 19, a $C_5$ inside the joint exclusivity structure $\Gamma^2$ cannot have all its edges supported by events from just one PR box. In other words, all edges cannot be assigned the same color. This leads to two cases:

1. where a $C_5$ has exactly one edge that is colored grey while the rest are colored red, see Figure 32(a).

2. where a $C_5$ has exactly two edges that are colored grey while the rest are colored red. This leads to two situations based on whether the two grey edges are adjacent or not in a $C_5$, see Figure 33(a),(c).

Consider case 1, see Figure 32(a), with just one edge with color grey. Corollary 17 implies that in the $C_3$ composed of events $\{u, v, w\}$ the edge connecting events $u$ and $w$ must be colored grey. The same holds true for the $C_3$ composed of $\{u, x, y\}$ where the edge corresponding to events $\{u, x\}$ is also forced to be of color grey, see Figure 32(b). This leads to a $C_3$ graph $\{u, w, x\}$ where all edges are of grey color. This $C_3$ violates Corollary 17, hence case 1 is impossible. With a similar sequential extension of graphs via Corollary 17 as above we can show that case 2 is also impossible. Figure 33 covers both situations

mentioned in case 2 above. Starting from the sub case represented by Figure 33(a), using Corollary 17 to sequentially extend each of the graphs with vertex sets $\{u, v, y\}$, $\{w, x, y\}$, $\{u, v, w\}$, $\{x, v, w\}$, and $\{u, v, x\}$ into a $C_3$ leads to a $K_5$ depicted in Figure 33(b) where thee $C_3$ composed of events $\{u, w, x\}$ violates Corollary 17.

Starting from the sub case represented in Figure 33(c), using Corollary 17 to sequentially extend each of the graphs with vertex sets $\{u, v, y\}$, $\{v, x, y\}$, $\{u, v, x\}$, and $\{w, v, y\}$ into a $C_3$ leads to a $C_5$ composed of $\{u, v, x, w, y\}$ with all its edges having color red, see Figure 33(d), violating Corollary 19. This exhausts the possibility of violation of E-principle with two copies of $n$-cycle PR boxes $\forall\ n \geq 6$. □

**Corollary 23.** *For $n \geq 6$, the joint exclusivity graph of a pair of $n$-cycle PR-boxes does not contain a $K_l$ for any $l \geq 6$.*

*Proof.* For a pair of n-cycle PR-boxes of a given $n \geq 6$, a $K_l$ exclusivity graph for $l \geq 6$ would contain a $K_5$ as its induced sub graph but the latter can't exist due to Theorem 21, hence, proving the corollary. □

### 5.6.4 Exploiting Ramsey theory

Until now we have seen that $n = 4, 5$ cycle PR boxes show activation effects as soon as more than one copy is considered while $n \geq 6$ cycle PR boxes don't when two copies are taken into account. We now see how our new found connection with Ramsey theory can help us say more about activation effects of $n$-cycle PR boxes.

The first result follows from the canonical Ramsey number $R(K_3, K_3) = 6$. For $n$-cycle PR boxes this means that for two copies if a $K_6$ exists inside $\Gamma^2$, it will have to have at least one monochromatic $K_3$ as sub graph. This isn't possible, thanks to Corollary 17, for any $n$. For $n = 4, 5$ this means that we can't violate E-principle beyond $K_5$ i.e. $R(K_3, K_3) = 6$ upper bounds violation of E-principle in $\Gamma^2$. This explains why in Ref. [22] authors didn't find a clique bigger than $K_5$ for two copies of CHSH ($n = 4$ cycle) PR box.

We now utilize known Ramsey theory results to extract useful results for $n \geq 6$ cycle PR boxes. Our first result follows from one of the many small but significant results proven in [93]. This work from Paul Erdös et. al. is one of the first works in the Ramsey theory literature that considered going beyond two colorings of the edges. The authors considered three and four colorings of edges where the structures assigned to the colors consisted of complete graphs, complete bipartite graphs, paths, and cycles. The relevant result for us is when, for three possible colorings of edges, each color is assigned an odd cycle. The authors define 'cycle Ramsey number' $R(\leq [C]) = p$ (where $[C] \equiv (C_{2d_1+1}, C_{2d_2+1}, C_{2d_3+1})$

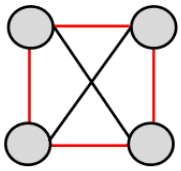


Figure 34: Two edge-colored $K_4$ where each monochromatic sub graph is a bi-partite graph.

which is to be read as color $i$ being assigned a cycle of size $2d_i + 1$) which denotes the smallest integer $p$ such that for any arbitrary three coloring of $K_p$ there exists at least one $i$ that contains an odd cycle of size less than or equal to $2d_i + 1$.

**Theorem 24.** *For $n \geq 6$, the joint exclusivity graph of three independent $n$-cycle PR boxes does not contain a $K_9$. That is, three copies cannot activate a violation of the E-principle for $n \geq 6$-cycle PR boxes.*

*Proof.* $R(\leq (C_5, C_5, C_5)) = 9$ obtained in [93] translated in the language of our work means that for a $K_9$ to exist in a joint exclusivity graph constructed from three statistically independent copies of $n$-cycle PR boxes (for any $n \geq 4$), an individual copy itself must have a $C_5$ or a $C_3$ cycle as its sub graph. But from Theorem 18, we know that for any $n \geq 6$ cycle PR box, the smallest odd cycle has size more than 5 implying that a $K_9$ can't exist in such cases. This finishes the proof of Theorem 24. This marks the impossibility of violating E-principle using three copies of $n \geq 6$-cycle PR boxes. □

This theorem provides the first example of extremal non-classical correlations for which three copies aren't enough to witness violation of E-principle. Earlier, for the scenarios explored, in [22], two copies were always enough to witness E-principle's violation.

Another result from Ref. [94] exploring edge coloring of $K_{2^k+1}$ graphs with $k$ colors in the context of Ramsey theory helps us infer the class of $n$-cycle scenarios whose PR boxes can't violate E-principle for a given $k$, where $k \geq 4$. The result states that any $k$ edge-colored clique of size $2^k + 1$ must have a monochromatic odd cycle. For the ease of reader we explain why that must be the case. It is easy to show that for complete graph $K_{2^k}$ there exists a $k$ coloring of edges such that each color comprises a bipartite graph. Before we provide an explicit construction for this we show that such $k$ edge colorings exist for $K_m$ only if $m \leq 2^k$.

Consider a $k$ edge coloring of $K_m$ where each coloring corresponds to a bi-partite graph. For each vertex of this graph $K_m$, imagine writing a binary vector of length $k$ where the $ith$ coordinate determines which side of the bi-partition of color $i$ this vertex lies. All vertices of $K_m$ must receive distinct binary vector because if two nodes receive the same binary vector it means there is no edge that connects the

two nodes. Since there can be at most $2^k$ such distinct labels, this means that $m \leq 2^k$. Therefore for $m > 2^k$ no such coloring exists. One explicit example of the above coloring when $m = 2^k$ is as follows: Take a two edge-colored $K_4$ such that each monochromatic projection is a bi-partite graph, see Figure 34. Now take two copies of this graph and connect the two $K_4$ graphs via a third color. This gives a $K_8$ all of whose monochromatic projections are bi-partite. Following the same idea one can progressively build a $k$ edge-colored $K_{2^k}$, each of whose $k$ monochromatic projections is bi-partite. Following theorem captures this result's consequence for $n \geq 6$ cycle PR boxes.

**Theorem 25.** *For $n \geq 2^k + 2$, the joint exclusivity graph of $k$ independent copies of an $n$-cycle PR box does not contain a $K_{2^k+1}$. That is, $k$ copies do not activate a violation of the E-principle for $n \geq 2^k + 2$-cycle PR boxes.*

*Proof.* The proof follows from a simple result provided in Ref. [94]. The result states that any time we have a $k$ edge-colored $K_{2^k+1}$ there then has to be at least one monochromatic graph that is not bi-partite i.e. every $k$ edge coloring of $K_{2^k+1}$ contains an odd monochromatic cycle. A direct implication of this result is that for $K_{2^k+1}$ to exist in $\Gamma^k$ at least one monochromatic projection must have an odd cycle. But we know from Theorem 18 that the smallest odd cycle for $n \geq 2^k + 2$ is of size $2^k + 3$ which is bigger than $2^k + 1$ the order of clique that is necessary for violation. Hence, no $n \geq 2^k + 2$ PR box can violate the E-principle for $k \geq 4$ copies. This finishes the proof. □

## 5.7 Discussion

In this chapter we presented our refinement of the conventional method for probing activation effects of the exclusivity principle. We then applied this refinement to explore activation effects for the $n$-cycle PR boxes, generating a wealth of results highlighted in Section 5.5.

Our refinement amounts to preserving in the joint exclusivity structure the information as to which components determine the exclusivity between each pair of joint events, something not taken into account in the conventional approach. We do so by building a joint exclusivity structure consisting of an edge-colored multigraph, where we assign a unique color to each component and use it to mark the edges (denoting exclusivity between joint events) determined by the exclusivity graph of that component (namely, by the exclusivity between the individual events at that component), as explained in Section 5.3.1.

The problem of exhibiting activation of the E-principle then maps to finding specific cliques within this joint edge-colored exclusivity multigraph. This is where Ramsey theory comes into play, as it necessitates the presence of certain monochromatic structures inside edge-colored cliques, which in turn translates

to constraints on the exclusivity graphs of the individual components. If these constraints are impossible to satisfy, then no clique of the required size exists in the joint exclusivity multigraph, ensuring that the E-principle is not violated in the given situation. On the other hand, however, satisfying the constraints from a given Ramsey result does not guarantee the existence of the corresponding clique. As such, Ramsey theory plays an asymmetric role in studying activation effects of the E-principle. Luckily, in our investigation of $n$-cycle PR boxes, we found relevant results in the Ramsey theory literature that let us draw conclusions about activation effects of the E-principle, both excluding violations of the E-principle with certain numbers of copies and bounding the amount of existing violations (as in the case $n = 4, 5$ with $k = 2, 3$).

We hope that the unexpected bridge established between Ramsey theory and quantum foundations encourages fruitful collaboration and two-way interaction between mathematicians and physicists in these fields.

Later in Chapter 6, Sec. 6.2, we will discuss the potential future directions cropping out from the contributions in this work.

# Chapter 6

# Summary and future work

## 6.1 Summary

This thesis presented our independent analyses of KS scenarios at three levels: classical, quantum, and post-quantum where each work is a substantial extension or refinement of the past analyses along the corresponding lines of research. At the classical level, as discussed in Chapter 3, we extended Pironio's [10] results on 'lifting Bell inequalities' to lifting arbitrary KS noncontextuality inequalities. At the quantum level, as discussed in Chapter 4, we dedicatedly analysed state-dependent contextuality of qubit Pauli operators, extending the traditional analyses of these operators beyond state-independent proofs of KS contextuality. At the post-quantum level, as discussed in Chapter 5, we refined the analysis by Fritz et. al. [22], of detecting post-quantumness of extremal Bell non-local correlations by the Exclusivity principle, to beyond Bell-like scenarios by establishing this problem's general connection with Ramsey theory. A common thread across each of these contributions is the class of $n$-cycle scenarios, which are the most fundamental scenarios for witnessing KS contextuality, thanks to Vorob'ev's theorem. The characterisation of these scenarios at the classical, quantum, and post-quantum levels, in Ref. [15], is heavily exploited in this thesis.

For our lifting results, the complete availability of facet-defining inequalities of these scenarios, thanks to Vorob'ev's theorem, allows us to derive many facet-defining inequalities of the non-contextual polytopes of arbitrary KS contextuality scenarios. For our analysis at the quantum level, we fully characterised the qubit-Pauli contextuality for $n$-cycle scenarios to conclude that no $n$-cycle noncontextuality inequality can be violated by these operators when $n \geq 5$ while for $n = 4$ every qubit Pauli realization of the scenario can attain the quantum bound. Building on this result and Vorob'ev's theorem we then showed that the presence of an induced 4-cycle is a sufficient but not a necessary condition for qubit Pauli operators to witness contextuality in arbitrary KS scenarios. For our analysis at the post-quantum level, we exploited the power of our Ramsey theory connection with E-principle's activation effects on the contextual vertices of

the $n$-cycle no-disturbance polytope. We call these vertices $n$-cycle PR boxes, due to their straightforward generalization from the famous PR box ($n = 4$). Using Ramsey theory results we then showed that $n \geq 6$ cycle PR boxes are the first examples of correlations, in the literature, that don't activate violation of E-principle using two copies. Another Ramsey theory result let us identify a class of $n$-cycle PR boxes which cannot exhibit activation effects for a given number of copies, say $k$, where the clique to be searched for for the violation is exponential in $k$, something intractable via computational searches.

More results and details related to each work are presented in the corresponding chapters. We now discuss the potential research directions for future works.

## 6.2 Future work

There are many possible future directions of research that spring from the contributions in this thesis. We list them below.

**Lifting noncontextuality inequalities.** Presented in Chapter 3, this work invites investigations in the spirit of Ref. [73] to identify new contextuality scenarios whose noncontextual polytope can be fully characterised via liftings, e.g. where every facet-defining NC inequality is obtained via lifting from some particularly simple class of noncontextuality inequalities such as cycle inequalities [12]. Complementarily, one may focus on identifying facet-defining NC inequalities that cannot be obtained via lifting from any sub-scenario. As such, our work can be seen as a step towards complete characterisations of noncontextual polytopes beyond the few known scenarios [12].

Another aspect that is worth clarifying, in the wake of the comments in Sec. 3.5.3, is how the various forms of lifting compose. This would help avoid redundancy, trimming the search space for inequalities lifted from sub-scenarios. Future applications of our method may include finding facet-defining inequalities in different kinds of scenarios such as Bell scenarios with sequential compatible measurements [97, 98, 99], extended Bell scenarios [100], or scenarios incorporating causality [101, 102, 103].

On a more practical note, contextuality has been linked to quantum advantage in a number of information-processing tasks. One prominent example is that of communication complexity protocols: it is shown in Ref. [104] that a violation of a NC inequality witnesses quantum advantage in a suitably designed protocol. New NC inequalities obtained by lifting may therefore capture quantum advantage in new, suitably related communication complexity protocols. Similar analyses can be carried out for other information-processing tasks whose success probability is given by a linear functional that determines a NC inequality.

Another aspect that can be explored for newly derived NC inequalities is self testing. An inequality

is self testing when a maximal quantum violation certifies a unique quantum realization (up to global isometries). It has been studied both for state-dependent [105] and state-independent [106] inequalities. An open question is whether the self-testing property is preserved under liftings. Furthermore, Ref. [107] showed that any state-independent violation of a NC inequality offers quantum advantage in a specific, suitably designed distributed computing task. Deriving new such inequalities using our lifting method might therefore provide new distributed computing protocols exhibiting quantum advantage.

Another practically relevant aspect that lifting might facilitate is in the vertex enumeration of the no-disturbance polytope of the bigger scenario. By changing the bound of the lifted inequality one may find subpolytopes of the no-disturbance polytope that are relatively easier to solve for vertex enumeration. Such techniques may be developed using inspiration from polytope slicing technique [108] used to find facet-defining inequalities of bigger polytopes.

**Qubit Pauli contextuality.** Presented in Chapter 4, this work inspires similar investigations for analysing KS contextuality from the purview of qudit Pauli operators. This is especially relevant for odd-prime dimensional qudit Pauli operators since they can only exhibit state-dependent contextuality. In other words, these operators cannot produce state-independent contextuality. In Ref. [6], the authors showed that no scenario realizable by odd-prime qudit Pauli operators has a non-contextuality inequality that can be violated by a strict subset of states denoted as $\mathcal{P}_{SIM}$ in the paper. It seems interesting to extend this analysis, in the spirit of our work, to identify whether there are any KS scenarios where these operators cannot witness any contextuality at all. This would mean that applications corresponding to those KS scenarios will give no quantum advantage with odd dimensional qudit Pauli observables. A good starting point to investigate this might be to first fully characterize $n$-cycle scenarios beyond dichotomic measurements and then study contextuality of qudit Pauli operators on these scenarios to further conclude results on more general ones, in light of Vorob'ev's theorem.

**Exclusivity principle, Ramsey theory, and $n$-cycle PR boxes.** Presented in Chapter 5, this work invites investigations for whether Ramsey-theoretic results can be used to explore E-principle activation effects for correlations on more general KS scenarios. An advantage of considering $n$-cycle PR boxes was that all possible events occur with the same probability. Because of this, the problem of finding violations of the E-principle with a given number of copies maps to finding cliques of a certain size, with no consideration given to the weights of the nodes (i.e. the probabilities of the corresponding joint events) due to their uniformity. The same property does not hold in general for (extremal boxes of) arbitrary scenarios, wherein witnessing violations also depends on the weights of the nodes. One could still, in principle, use Ramsey-flavoured results for vertex-weighted graphs. The main obstacle is that finding Ramsey numbers

is very hard: after 100 years of research, there are only partial results, mostly consisting of upper or lower bounds for Ramsey numbers, going only up to four colors [109]. It would be interesting to investigate how these known results can be utilized in exploring violations of the E-principle in more general scenarios, as well as other questions in physics that can be reformulated as clique-finding problems.

Another natural question is whether for every $n$-cycle PR box there exists a number of copies $k$ that activates a violation of the E-principle? We conjecture that the answer to this question is yes. We have already seen that two copies suffice when $n = 4, 5$, but not when $n \geq 6$. In fact, we further showed that even three copies are not enough to activate a violation when $n \geq 6$.

Let us provide some evidence in support of the conjecture. Cabello [84, Lemma 4] showed that for a given exclusivity graph, its theta body $TH(G)$ coincides with the largest set $\mathcal{P}$ of probability assignments (vertex-weightings) such that: (i) the independent product of any two assignments in the set (i.e. $p \otimes p'$ for $p, p' \in \mathcal{P}$) satisfies the E-principle, and (ii) the independent product of any number of copies of an assignment in the set (i.e. $p^{\otimes k}$ for $p \in \mathcal{P}$) satisfies the E-principle [1]. Note that $TH(G)$ collects the probability assignments arising from all the quantum correlations consistent with that graph, not just from a fixed scenario; that is, it contains quantum correlations coming from all possible scenarios that produce events realizing the given exclusivity graph. Using the Lovász number characterization of $M_{2n}$ and $Y_n$ from Refs. [15, 110], one can easily establish that any $n$-cycle PR box lies outside the theta body of its exclusivity graph. However, this is not quite enough to conclusively answer our question: in light of the above-mentioned result, it is conceivable that a correlation outside the theta body satisfies the E-principle with any number of copies but sidesteps the result due to condition (i), by violating the E-principle when considered jointly with another correlation from $TH(G)$. This seems unlikely, especially given the highly symmetric nature of $n$-cycle scenarios. Still, whether this is possible is left as an open problem for now.

We conclude hoping that the unexpected bridge established between Ramsey theory and quantum foundations encourages fruitful collaboration and two-way interaction between mathematicians and physicists in these fields.

## 6.3 Other work

**Bipartite extremal no-signalling non-PR boxes and their applications.** This is an ongoing work, in collaboration with Adán Cabello, Emmanuel Zambrini, and Junior R. Gonzales-Ureta, that didn't make it to this thesis. It is centered specifically around Bell scenarios, more precisely, around the extremal

[1] Note that the the assignment $p \otimes p'$ is a probability assignment for the OR-product of the graph $G$ with itself, while $p^{\otimes k}$ is a probability assignment for the OR-product of $k$ copies of $G$.

no-signalling (ENS) boxes of bi-partite Bell scenarios. ENS boxes are nonlocal correlations that correspond to vertices of the no-signalling polytope of a Bell scenario. It is known that no ENS box can be realized by quantum theory [44]. In other words, according to quantum theory, ENS boxes are nonphysical. Nonetheless, knowing ENS boxes becomes crucial for addressing a number of diverse problems: testing physical principles proposed to single out quantum theory; generating new Bell inequalities [111, 112]; determining the type of non-locality exhibited by certain quantum correlations [113]; determining the monogamy properties of physical "fully nonlocal" correlations, among others. The only scenarios for which all ENS boxes are known are $(N, 2; N, 2)$[2], $(2, d; 2, d)$, $(d, 2; d, d)$, and $(2, 2; 2, 2; 2, 2)$. In $(N, 2; N, 2)$ and $(2, d; 2, d)$, all ENS boxes can be derived from PR boxes. The conventional methods to derive ENS boxes for a given Bell scenario rely on (1) solving vertex enumeration for the no-signalling polytope, (2) finding a no-signalling behaviour maximally violating a facet-defining inequality of the local polytope. Each of these methods has its own limitations: (1) becomes practically infeasible for bigger polytopes while (2) fails to capture all ENS boxes because not all such boxes maximally violate a facet-defining inequality of the local polytope.

In this work, we provided a novel method that reliably generates all ENS boxes for a given scenario, even including those which do not maximally violate any facet inequality of the local polytope. The method is based on the polytope slicing technique whose intuition is similar to one provided in Ref. [108]: one translates the bound of a known facet to focus on a small section of the no-signalling polytope to then either apply vertex enumeration on the resulting slice of the polytope or use NPA hierarchy with post-processing methods to recover the ENS boxes. We applied this method to identify all the ENS boxes for some scenarios: $(2, 3; 3, 2)$, $(3, 3; 3, 2)$, and $(2, 3; 3, 3)$. The new boxes include the first examples of ENS boxes which, in the bipartite case, are not derived from PR boxes. We are using these non-PR boxes (at the time of writing this thesis) to address some open problems like determining monogamy properties of certain "fully nonlocal" correlations, detecting post-quantumness of these correlations using E-principle, among a few others. My main role has been in analysing E-principle's role in detecting post-quantumness of these non-PR boxes. These boxes deviate from PR boxes due to having non-uniform probability distributions for contexts unlike a PR box where the probability distribution corresponding to every context is uniform. Due to this non-uniformity of probability distributions, the problem of finding (or negating) E-principle violation producing cliques via Ramsey theory becomes harder because the weights of nodes are not uniform and we found no relevant work in the Ramsey theory literature to help with our post-quantumness analysis for these boxes. Nonetheless, using exhaustive computational search we found that, just like for the PR box,

[2] the notation $(m_A, d_A; m_B, d_B; m_C, d_C)$ is such that $m_A$ denotes the total number of measurements of party $A$, while $d_A$ denotes outcomes of any measurement for party $A$.

two copies of each of these non-PR boxes we found are enough to activate violation of E-principle. We hope to have this work out by the time this thesis is defended.

# Bibliography


[1] John S Bell. On the einstein podolsky rosen paradox. *Physics Physique Fizika*, 1(3):195, 1964.

[2] John S Bell. On the problem of hidden variables in quantum mechanics. *Reviews of Modern physics*, 38(3):447, 1966.

[3] Simon Kochen and Ernst P Specker. The problem of hidden variables in quantum mechanics. *Ernst Specker Selecta*, pages 235–263, 1990.

[4] Robert W Spekkens. Contextuality for preparations, transformations, and unsharp measurements. *Physical Review A—Atomic, Molecular, and Optical Physics*, 71(5):052108, 2005.

[5] Lucien Hardy and Robert Spekkens. Why physics needs quantum foundations. *arXiv preprint arXiv:1003.5008*, 2010.

[6] Mark Howard, Joel Wallman, Victor Veitch, and Joseph Emerson. Contextuality supplies the 'magic'for quantum computation. *Nature*, 510(7505):351–355, 2014.

[7] Robert Raussendorf. Contextuality in measurement-based quantum computation. *Physical Review A—Atomic, Molecular, and Optical Physics*, 88(2):022322, 2013.

[8] Sergey Bravyi, David Gosset, and Robert König. Quantum advantage with shallow circuits. *Science*, 362(6412):308–311, 2018.

[9] Sergey Bravyi, David Gosset, Robert König, and Marco Tomamichel. Quantum advantage with noisy shallow circuits. *Nature Physics*, 16(10):1040–1045, 2020.

[10] Stefano Pironio. Lifting Bell inequalities. *J. Math. Phys.*, 46(6):062112, 2005.

[11] N. N. Vorob'ev. Consistent families of measures and their extensions. *Theory Probab. Its Appl.*, 7(2):147–163, 1962.

[12] Mateus Araújo, Marco Túlio Quintino, Costantino Budroni, Marcelo Terra Cunha, and Adán Cabello. All noncontextuality inequalities for the $n$-cycle scenario. *Phys. Rev. A*, 88:022118, Aug 2013.

[13] David Avis, Hiroshi Imai, Tsuyoshi Ito, and Yuuya Sasaki. Two-party bell inequalities derived from combinatorics via triangular elimination. *J. Phys. A*, 38(50):10971, 2005.

[14] Raman Choudhary, Rui Soares Barbosa, and Adán Cabello. Lifting noncontextuality inequalities. *Physical Review A*, 109(5):052216, 2024.

[15] Mateus Araújo, Marco Túlio Quintino, Costantino Budroni, Marcelo Terra Cunha, and Adán Cabello. All noncontextuality inequalities for the n-cycle scenario. *Physical Review A—Atomic, Molecular, and Optical Physics*, 88(2):022118, 2013.

[16] William M Kirby and Peter J Love. Contextuality test of the nonclassicality of variational quantum eigensolvers. *Physical Review Letters*, 123(20):200501, 2019.

[17] William M Kirby and Peter J Love. Classical simulation of noncontextual Pauli hamiltonians. *Physical Review A*, 102(3):032418, 2020.

[18] N David Mermin. Simple unified form for the major no-hidden-variables theorems. *Physical review letters*, 65(27):3373, 1990.

[19] Asher Peres. Incompatible results of quantum measurements. *Physics Letters A*, 151(3-4):107–108, 1990.

[20] N David Mermin. Hidden variables and the two theorems of john bell. *Reviews of Modern Physics*, 65(3):803, 1993.

[21] Raman Choudhary and Rui Soares Barbosa. Contextuality with pauli observables in cycle scenarios. *arXiv preprint arXiv:2502.03451*, 2025.

[22] Tobias Fritz, Ana Belén Sainz, Remigiusz Augusiak, J Bohr Brask, Rafael Chaves, Anthony Leverrier, and Antonio Acín. Local orthogonality as a multipartite principle for quantum correlations. *Nature communications*, 4(1):2263, 2013.

[23] Ronald L Graham, Bruce L Rothschild, and Joel H Spencer. *Ramsey theory*, volume 20. John Wiley & Sons, 1991.

[24] Yusheng Li and Qizhong Lin. *Elementary methods of graph Ramsey theory*. Springer, 2022.

[25] Raman Choudhary and Rui Soares Barbosa. Exclusivity principle, ramsey theory, and $n$-cycle pr boxes. *arXiv preprint arXiv:2411.09773*, 2024.

[26] Adán Cabello, José M Estebaranz, and Guillermo García-Alcaine. Bell-Kochen-Specker theorem: A proof with 18 vectors. *Physics Letters A*, 212(4):183–187, 1996.

[27] John F Clauser, Michael A Horne, Abner Shimony, and Richard A Holt. Proposed experiment to test local hidden-variable theories. *Physical review letters*, 23(15):880, 1969.

[28] Arthur Fine. Hidden variables, joint probability, and the bell inequalities. *Physical Review Letters*, 48(5):291, 1982.

[29] Arthur Fine. Joint distributions, quantum correlations, and commuting observables. *Journal of Mathematical Physics*, 23(7):1306–1310, 1982.

[30] Giulio Chiribella and Xiao Yuan. Measurement sharpness cuts nonlocality and contextuality in every physical theory. *arXiv preprint arXiv:1404.3348*, 2014.

[31] Giulio Chiribella and Xiao Yuan. Bridging the gap between general probabilistic theories and the device-independent framework for nonlocality and contextuality. *Information and Computation*, 250:15–49, 2016.

[32] Zhen-Peng Xu and Adán Cabello. Necessary and sufficient condition for contextuality from incompatibility. *Physical Review A*, 99(2):020103, 2019.

[33] MA Naimark. About second-kind self-adjoint extensions of symmetrical operator. *Izv. Akad. Nauk USSR, Ser. Mat*, 4:53–104, 1940.

[34] Roberto Beneduci. Notes on naimark's dilation theorem. In *Journal of Physics: Conference Series*, volume 1638, page 012006. IOP Publishing, 2020.

[35] William Slofstra. The set of quantum correlations is not closed. In *Forum of Mathematics, Pi*, volume 7, page e1. Cambridge University Press, 2019.

[36] Adán Cabello, Simone Severini, and Andreas Winter. Graph-theoretic approach to quantum correlations. *Phys. Rev. Lett.*, 112:040401, 2014.

[37] Samson Abramsky and Adam Brandenburger. The sheaf-theoretic structure of non-locality and contextuality. *New J. Phys.*, 13:113036, 2011.

[38] Antonio Acín, Tobias Fritz, Anthony Leverrier, and Ana Belén Sainz. A combinatorial approach to nonlocality and contextuality. *Commun. Math. Phys.*, 334:533–628, 2015.

[39] Adán Cabello. Simple explanation of the quantum violation of a fundamental inequality. *Phys. Rev. Lett.*, 110:060402, 2013.

[40] Adán Cabello. Quantum correlations from simple assumptions. *Phys. Rev. A*, 100:032120, 2019.

[41] Günter M. Ziegler. *Lectures on Polytopes*. Springer, New York, NY, 2012.

[42] Hermann Minkowski. Allgemeine Lehrsätze über die konvexen Polyeder. *Nachr. Ges. Wiss. zu Gö"ttingen*, pages 198–219, 1897.

[43] H. Weyl. Elementare Theorie der konvexen Polyeder. *Comment. Math. Helvetici*, 7:290–306, 1934.

[44] Ravishankar Ramanathan, Jan Tuziemski, Michał Horodecki, and Paweł Horodecki. No quantum realization of extremal no-signaling boxes. *Physical Review Letters*, 117(5):050401, 2016.

[45] Lucas Pollyceno, Rafael Chaves, and Rafael Rabelo. Information causality in multipartite scenarios. *Physical Review A*, 107(4):042203, 2023.

[46] Sandu Popescu. Why isn't nature more non-local? *Nature Physics*, 2(8):507–508, 2006.

[47] Gilles Brassard, Harry Buhrman, Noah Linden, André Allan Méthot, Alain Tapp, and Falk Unger. Limit on nonlocality in any world in which communication complexity is not trivial. *Physical Review Letters*, 96(25):250401, 2006.

[48] Nicolas Brunner and Paul Skrzypczyk. Nonlocality distillation and postquantum theories with trivial communication complexity. *Physical review letters*, 102(16):160403, 2009.

[49] Wim Van Dam. Implausible consequences of superstrong nonlocality. *Natural Computing*, 12:9–12, 2013.

[50] Marcin Pawłowski, Tomasz Paterek, Dagomir Kaszlikowski, Valerio Scarani, Andreas Winter, and Marek Żukowski. Information causality as a physical principle. *Nature*, 461(7267):1101–1104, 2009.

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

[52] Alexander A. Klyachko, M. Ali Can, Sinem Binicioğlu, and Alexander S. Shumovsky. Simple test for hidden variables in spin-1 systems. *Phys. Rev. Lett.*, 101:020403, 2008.

[53] Adan Cabello. Experimentally testable state-independent quantum contextuality. *Phys. Rev. Lett.*, 101:210401, 2008.

[54] Costantino Budroni, Adán Cabello, Otfried Gühne, Matthias Kleinmann, and JanÅke Larsson. Kochen-Specker contextuality. *Rev. Mod. Phys.*, 94:045007, Dec 2022.

[55] David Avis and Komei Fukuda. A pivoting algorithm for convex hulls and vertex enumeration of arrangements and polyhedra. In *Proceedings of the Seventh Annual Symposium on Computational Geometry*, pages 98–104, 1991.

[56] Donald R. Chand and Sham S. Kapur. An algorithm for convex polytopes. *J. ACM*, 17(1):78–86, jan 1970.

[57] Garret Swart. Finding the convex hull facet by facet. *J. Algorithms*, 6(1):17–48, 1985.

[58] N N Vorob'ev. Markov measures and markov extensions. *Theory of Probability & Its Applications*, 8(4):420–429, 1963.

[59] N N Vorob'ev. Coalition games. *Theory of Probability & Its Applications*, 12(2):251–266, 1967.

[60] Costantino Budroni and Giovanni Morchio. The extension problem for partial boolean structures in quantum mechanics. *Journal of mathematical physics*, 51(12):122205, 2010.

[61] Costantino Budroni and Adan Cabello. Bell inequalities from variable-elimination methods. *Journal of Physics A: Mathematical and Theoretical*, 45(38):385304, 2012.

[62] Ravishankar Ramanathan, Akihito Soeda, Paweł Kurzyński, and Dagomir Kaszlikowski. Generalized monogamy of contextual inequalities from the no-disturbance principle. *Physical review letters*, 109(5):050404, 2012.

[63] Rui Soares Barbosa. On monogamy of non-locality and macroscopic averages: examples and preliminary results. *arXiv preprint arXiv:1412.8541*, 2014.

[64] Samson Abramsky, Rui Soares Barbosa, Kohei Kishida, Raymond Lal, and Shane Mansfield. Possibilities determine the combinatorial structure of probability polytopes. *Journal of Mathematical Psychology*, 74:58–65, 2016.

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

[66] Piotr Badziąg, Ingemar Bengtsson, Adán Cabello, and Itamar Pitowsky. Universality of state-independent violation of correlation inequalities for noncontextual theories. *Phys. Rev. Lett.*, 103:050401, 2009.

[67] Sixia Yu and C. H. Oh. State-independent proof of Kochen-Specker theorem with 13 rays. *Phys. Rev. Lett.*, 108:030402, 2012.

[68] David Avis, Hiroshi Imai, Tsuyoshi Ito, and Yuuya Sasaki. Two-party Bell inequalities derived from combinatorics via triangular elimination. *J. Phys. A*, 38(50):10971, nov 2005.

[69] Philip M. Pearle. Hidden-variable example based upon data rejection. *Phys. Rev. D*, 2:1418–1425, Oct 1970.

[70] Samuel L. Braunstein and Carlton M. Caves. Wringing out better Bell inequalities. *Ann. Phys.*, 202:22–56, 1990.

[71] George Svetlichny. Distinguishing three-body from two-body nonseparability by a bell-type inequality. *Phys. Rev. D*, 35:3066–3069, May 1987.

[72] Daniel Collins, Nicolas Gisin, Noah Linden, Serge Massar, and Sandu Popescu. Bell inequalities for arbitrarily high-dimensional systems. *Phys. Rev. Lett.*, 88:040404, Jan 2002.

[73] Stefano Pironio. All Clauser–Horne–Shimony–Holt polytopes. *J. Phys. A*, 47(42):424020, oct 2014.

[74] Sergey Bravyi, Graeme Smith, and John A Smolin. Trading classical and quantum computational resources. *Physical Review X*, 6(2):021043, 2016.

[75] Samson Abramsky, Rui Soares Barbosa, Kohei Kishida, Raymond Lal, and Shane Mansfield. Contextuality, cohomology and paradox. In Stephan Kreutzer, editor, *24th EACSL Annual Conference on Computer Science Logic (CSL 2015)*, volume 41 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 211–228. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2015.

[76] Samson Abramsky, Rui Soares Barbosa, Giovanni Carù, and Simon Perdrix. A complete characterization of all-versus-nothing arguments for stabilizer states. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 375(2106):20160385, 2017.

[77] Michael A Nielsen and Isaac L Chuang. *Quantum computation and quantum information*. Cambridge university press, 2010.

[78] Samson Abramsky, Şerban-Ion Cercelescu, and Carmen-Maria Constantin. Commutation groups and state-independent contextuality. In *9th International Conference on Formal Structures for Computation and Deduction (FSCD 2024)*. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2024.

[79] Thomas Christof and Andreas Loebel. Porta `http://porta.zib.de/`, 2015.

[80] Rodrigo Gallego, Lars Erik Würflinger, Antonio Acín, and Miguel Navascués. Quantum correlations require multipartite information principles. *Physical review letters*, 107(21):210403, 2011.

[81] Adán Cabello. Simple explanation of the quantum violation of a fundamental inequality. *Physical review letters*, 110(6):060402, 2013.

[82] Adán Cabello, Lars Eirik Danielsen, Antonio J López-Tarrida, and José R Portillo. Basic exclusivity graphs in quantum correlations. *Physical Review A—Atomic, Molecular, and Optical Physics*, 88(3):032104, 2013.

[83] Barbara Amaral, Marcelo Terra Cunha, and Adán Cabello. Exclusivity principle forbids sets of correlations larger than the quantum set. *Physical Review A*, 89(3):030101, 2014.

[84] Adán Cabello. Quantum correlations from simple assumptions. *Physical Review A*, 100(3):032120, 2019.

[85] Ernst Specker. Die Logik nicht gleichzeitig entscheidbarer Aussagen. *Ernst Specker Selecta*, pages 175–182, 1990.

[86] Tomáš Gonda, Ravi Kunjwal, David Schmid, Elie Wolfe, and Ana Belén Sainz. Almost quantum correlations are inconsistent with specker's principle. *Quantum*, 2:87, 2018.

[87] Miguel Navascués, Yelena Guryanova, Matty J Hoban, and Antonio Acín. Almost quantum correlations. *Nature communications*, 6(1):6288, 2015.

[88] Richard M Karp. On the computational complexity of combinatorial problems. *Networks*, 5(1):45–68, 1975.

[89] Rafael Rabelo, Cristhiano Duarte, Antonio J López-Tarrida, Marcelo Terra Cunha, and Adán Cabello. Multigraph approach to quantum non-locality. *Journal of Physics A: Mathematical and Theoretical*, 47(42):424021, 2014.

[90] Lina Vandré and Marcelo Terra Cunha. Quantum sets of the multicolored-graph approach to contextuality. *Physical Review A*, 106(6):062210, 2022.

[91] Lucas EA Porto, Rafael Rabelo, Marcelo Terra Cunha, and Adán Cabello. The quantum maxima for the basic graphs of exclusivity are not reachable in bell scenarios. *Philosophical Transactions of the Royal Society A*, 382(2268):20230006, 2024.

[92] Michael Codish, Michael Frank, Avraham Itzhakov, and Alice Miller. Computing the Ramsey number R(4, 3, 3) using abstraction and symmetry breaking. *Constraints*, 21:375–393, 2016.

[93] P Erdös, RJ Faudree, CC Rousseau, and RH Schelp. Generalized Ramsey theory for multiple colors. *Journal of Combinatorial Theory, Series B*, 20(3):250–264, 1976.

[94] A Nicholas Day and J Robert Johnson. Multicolour Ramsey numbers of odd cycles. *Journal of Combinatorial Theory, Series B*, 124:56–63, 2017.

[95] Dénes König. Theory of finite and infinite graphs. In *Theory of Finite and Infinite Graphs*, pages 45–421. Springer, 1990.

[96] Kheemeng Koh, Fengming Dong, and Eng Guan Tay. *Introduction to graph theory: H3 mathematics*. World Scientific Publishing Company, 2007.

[97] Adán Cabello. Proposal for revealing quantum nonlocality via local contextuality. *Phys. Rev. Lett.*, 104:220401, 2010.

[98] Bi-Heng Liu, Xiao-Min Hu, Jiang-Shan Chen, Yun-Feng Huang, Yong-Jian Han, Chuan-Feng Li, Guang-Can Guo, and Adán Cabello. Nonlocality from local contextuality. *Phys. Rev. Lett.*, 117:220402, 2016.

[99] Junior R. Gonzales-Ureta, Ana Predojević, and Adán Cabello. Optimal and tight Bell inequalities for state-independent contextuality sets. *Phys. Rev. Res.*, 5:L012035, Mar 2023.

[100] Peng Xue, Lei Xiao, G. Ruffolo, A. Mazzari, T. Temistocles, M. Terra Cunha, and R. Rabelo. Synchronous observation of Bell nonlocality and state-dependent contextuality. *Phys. Rev. Lett.*, 130:040201, Jan 2023.

[101] Tobias Fritz. Beyond Bell's theorem II: Scenarios with arbitrary causal structure. *Commun. Math. Phys.*, 341(2):391–434, 2016.

[102] Stefano Gogioso and Nicola Pinzani. The sheaf-theoretic structure of definite causality. In Chris Heunen and Miriam Backens, editors, *18th International Conference on Quantum Physics and Logic (QPL 2021)*, volume 343 of *Electronic Proceedings in Theoretical Computer Science*, pages 301–324. Open Publishing Association, 2021.

[103] Samson Abramsky, Rui Soares Barbosa, and Amy Searle. Combining contextuality and causality: a game semantics approach. *Phil. Trans. R. Soc. A*, 382:20230002, 2024.

[104] Shashank Gupta, Debashis Saha, Zhen-Peng Xu, Adán Cabello, and Archan S Majumdar. Quantum contextuality provides communication complexity advantage. *Physical Review Letters*, 130(8):080802, 2023.

[105] Kishor Bharti, Maharshi Ray, Antonios Varvitsiotis, Naqueeb Ahmad Warsi, Adán Cabello, and Leong-Chuan Kwek. Robust self-testing of quantum systems via noncontextuality inequalities. *Physical review letters*, 122(25):250403, 2019.

[106] Zhen-Peng Xu, Debashis Saha, Kishor Bharti, and Adán Cabello. State-independent certification of quantum observables. *arXiv preprint arXiv:2309.05735*, 2023.

[107] Debashis Saha, Paweł Horodecki, and Marcin Pawłowski. State independent contextuality advances one-way communication. *New Journal of Physics*, 21(9):093057, 2019.

[108] José Jesus and Emmanuel Zambrini Cruzeiro. Tight bell inequalities from polytope slices. *Physical Review A*, 108(5):052220, 2023.

[109] Stanisław Radziszowski. Small Ramsey numbers. *The Electronic Journal of Combinatorics*, pages DS1–Jan, 2012.

[110] Kishor Bharti, Maharshi Ray, Zhen-Peng Xu, Masahito Hayashi, Leong-Chuan Kwek, and Adán Cabello. Graph-theoretic approach for self-testing in bell scenarios. *PRX Quantum*, 3(3):030344, 2022.

[111] Nicolas Brunner, Valerio Scarani, and Nicolas Gisin. Bell-type inequalities for nonlocal resources. *Journal of mathematical physics*, 47(11), 2006.

[112] Thomas Cope and Roger Colbeck. Bell inequalities from no-signaling distributions. *Phys. Rev. A*, 100:022114, Aug 2019.

[113] Yuan Liu, Ho Yiu Chung, Emmanuel Zambrini Cruzeiro, Junior R Gonzales-Ureta, Ravishankar Ramanathan, and Adán Cabello. Equivalence between face nonsignaling correlations, full nonlocality, all-versus-nothing proofs, and pseudotelepathy. *Physical Review Research*, 6(4):L042035, 2024.