Logical and Topological Contextuality in Quantum Mechanics and Beyond

Giovanni Carú

Linacre College
University of Oxford

Supervisor: Samson Abramsky

A thesis submitted for the degree of
Doctor of Philosophy in Computer Science

Hilary 2019

Ai miei genitori.

Acknowledgements

I would like to thank my supervisor Samson Abramsky for his guidance and encouragement throughout these years at Oxford. Working with him has been a great pleasure and a true honour. Special credit goes to my collaborators Samson Abramsky, Rui Soares Barbosa, Nadish de Silva, Kohei Kishida, Shane Mansfield and Simon Perdrix, who have contributed to some of the work presented in this thesis. I gratefully acknowledge generous financial support from the EPSRC Doctoral Training Partnership and the Oxford–Google DeepMind Graduate Scholarship. I'm also thankful to the Quantum Group in Oxford's Computer Science Department, the Simons Institute for the Theory of Computing at U.C. Berkeley, the Departments of Mathematics and Computer Science of Tulane University, and all of their members for providing such exciting and stimulating work environments.

I have been blessed with many fantastic friends throughout my life, and I can't thank them enough for their support during my studies. I am particularly grateful to: my good friend and officemate Rui Soares Barbosa, who kindly proof-read this thesis and provided me with great guidance, help and support throughout this project; the rest of the "sheaf team": Kohei Kishida, Linde Wester, Sivert Aasnaess and particularly Nadish de Silva and Shane Mansfield, with whom Samson, Rui and I shared numerous inspiring discussions as well as many wonderful memories, including three beautiful and productive months in Berkeley; my friends and coworkers in New Orleans: Mike Mislove, who helped me a lot during my term at Tulane, as well as Bert Lindenhovius and Diego Euribe, the best companions I could have asked for during my american adventure; my longtime buddy Rocco Meli, who joined me in Oxford after many unforgettable years together in Lugano and Lausanne, and my great friend Vladimir Olivero, who probably helped me more than anyone by making me laugh and keeping me sane, despite his horrible taste in football; all of my friends at Linacre college, particularly Nyree Manoukian, Shahé Gregorian, Ashley Tsai, and Karrar Ali, who taught me more about life than he probably realises; my flatmates Temitope Ajileye, Matthew Lau (along with rest of the MFoCS class of 2015), and Andrea Geraci, one of the best people I've ever met.

I would also like to thank my teachers and mentors, without whom I would have never made it to this point: Fabiola Gianotti, a constant source of immense admiration, for her kindness and encouragement; Kathryn Hess–Bellwald, to whom I owe my passion for Topology; Eva

Bayer–Fluckiger, Jacques Thévénaz and Justin Young for their support in my Oxford application. I am especially grateful to my high school teacher Vannia Della Casa, who made me fall in love with mathematics and pushed me towards my goals with infinite patience and moving affection.

The most heartfelt thanks goes to my girlfriend Danaïs, who's always stood by my side despite the distance, and even occasionally crossed entire continents just so that we could be together. Thank you for showing me the true face of happiness every day, for your unconditional love and support, and for putting up with my many flaws and fears.

Finally, the biggest thanks of all goes to my family, for their constant love and passionate encouragement. To my grandparents: nonna Jaja and nonno Bruno, who taught me the value of friendship and fostered my ambition to become a scientist; to nonna Titti and nonno Franco, who would have been proud to see this and, in a way, I'm sure they did. I am also very thankful to Hanne-Lore and the late Cav. Carlo Lucchini, my "nonni bis", whose kindness and generousness will never be matched;

Dulcis in fundo, I would like to thank my parents, the true heroes behind everything I have ever achieved. This is for you. And since you probably won't understand a word of it, let me simply say that it's the greatest thing I've ever done.

Abstract

The main subjects of this thesis are non-locality and contextuality, two fundamental features of quantum mechanics that constitute valuable resources for quantum computation. Our analysis is based on Abramsky & Brandenburger's sheaf theoretic framework, which captures both these phenomena in a unified treatment and in a very general setting. This high-level description transcends quantum physics and allows to precisely characterise the notion of contextuality as the apparent paradox realised by data being *locally consistent* but *globally inconsistent*. More specifically, we aim to develop a deeper understanding of so-called *logical* forms of contextuality, i.e. situations where the phenomenon can be witnessed using purely logical arguments, disregarding probabilities.

The sheaf theoretic description of logical contextuality has recently inspired the development of a topological treatment of the phenomenon based on sheaf cohomology. In this thesis, we embark on a detailed analysis of the cohomology of contextuality, exposing key shortcomings in the current methods, and introducing an (almost) complete cohomological characterisation of logical forms of contextuality. More specifically, we show that, in its current formulation, sheaf cohomology does not constitute a complete invariant for contextuality, not even in its strongest forms, and that higher cohomology groups cannot be used to study the phenomenon. Then, we solve these issues by introducing a novel construction, which derives refined versions of the presheaves describing empirical models to expose their deeper extendability properties, resulting in a sheaf cohomological invariant which is applicable to the vast majority of empirical models, and conjectured to work universally.

We propose a general theory of *contextual semantics* using the language of valuation algebras. In particular, we give a general definition of contextual behaviour as a fundamental gap between local agreement and global disagreement of information sources. Not only does this formalism aptly capture and generalise the known instances of contextuality beyond quantum theory, but it also provides inspiration for further applications of the phenomenon, and paves the way for the transfer of results and techniques between different fields. We give a prime example of this potential by developing faster algorithms to detect contextuality based on mainstream methods of *generic inference*.

Finally, we turn our attention back to instances of contextuality in quantum physics, and study strong contextuality in multi-qubit states.

We give a complete combinatorial characterisation of *All-vs-Nothing* proofs of strong contextuality in stabiliser quantum mechanics. This allows to produce the complete list of all stabiliser states exhibiting this kind of contextuality, which consitututes an important resource in certain models of quantum computation. Then, we extend our search for strongly contextual behaviour beyond stabiliser states, and identify the minimum quantum resources needed to realise strong non-locality. Additional results include a partial classification of strongly non-local models comprised of three-qubit states and local projective measurements, and the introduction of a new infinite family of strongly non-local three-qubit states.

Contents

Chapter I. Introduction 1
1. General motivation 1
2. Logical and topological aspects of contextuality 3
3. Contributions 5
4. Statement of collaboration 6
5. Outline of the thesis 6

Chapter II. Background: the sheaf theoretic structure of contextuality 9
1. Overview 9
2. A basic example 10
2.1. Measurement scenario 10
2.2. Empirical model 11
2.3. Contextuality 12
3. Measurement scenarios 16
3.1. Bell-type scenarios 17
3.2. Measurement scenarios and simplicial complexes 18
4. Empirical models 19
4.1. Sheaf theory 19
4.2. Events 20
4.3. Event distributions 21
4.4. Empirical models 22
4.5. Possibilistic models 23
4.6. Bundle diagrams 24
4.7. Modelisation of experimental data 25
5. Non-locality and contextuality 26
5.1. A hierarchy of contextuality 27
5.2. All-vs-Nothing arguments 29
5.3. Contextuality in bundle diagrams 30
5.4. Vorob'ev's theorem 31
6. The cohomology of non-locality and contextuality 32
6.1. Sheaf cohomology 33
6.2. Cohomology obstructions 35

Chapter III. The cohomology of contextuality: extensions and limitations 39
1. Overview 39
2. False negatives in cohomology 41
3. A false negative for strong contextuality 43

4. Extension to higher cohomology groups 46
4.1. A hierarchy of cohomological obstructions 48
4.2. Higher cohomology groups cannot be used to study contextuality 50
5. An alternative description of the first cohomology group 51
5.1. The connecting homomorphisms 51
5.2. $\mathcal{F}$-torsors 52
5.3. $\mathcal{F}$-torsors and contextuality 52
Discussion 56

Chapter IV. Towards a complete cohomological invariant for non-locality and contextuality 59
1. Overview 59
2. Line models and scenarios 60
2.1. A motivating example 60
2.2. Line scenarios 61
2.3. Line models 64
2.4. Interpretation and examples 66
3. The contextuality of line models 68
4. Cyclic models and their properties 70
4.1. Paths and cycles 70
4.2. Fundamental properties of paths and cycles 72
5. The cohomology of cyclic models: a complete invariant 75
5.1. Preliminaries 75
5.2. A complete cohomological invariant for contextuality in cyclic models 76
5.3. Examples 83
6. Extending the invariant to general models 87
6.1. Examples 89
7. Probabilistic line models 94
7.1. The contextuality of probabilistic line models 96
Discussion 98

Chapter V. Contextuality and generic inference: a theory of disagreement 99
1. Overview 99
2. Valuation algebras and generic inference 100
2.1. Valuation prealgebras 100
2.2. Valuation algebras 101
2.3. Information algebras 102
2.4. Inference problems 102
3. Basic examples 103
4. Advanced examples: language and models 108
4.1. Propositional logic 109
4.2. General languages 111
4.3. Examples 115
5. Valuation algebras and sheaf theory 118
5.1. Examples 118
6. A theory of disagreement 120

6.1. Defining disagreement 120
6.2. Examples 121
7. Disagreement and contextuality 126
8. Disagreement, complete disagreement and inference problems 127
8.1. Ordered valuation algebras 128
8.2. A general construction for composition 129
8.3. Detecting disagreement is an inference problem 132
8.4. Complete disagreement 133
9. Disagreement and possibilistic forms of contextuality 134
9.1. Detecting logical and strong contextuality is an inference problem 135
10. Inference algorithms for contextuality 136
10.1. Current algorithms and their complexity 136
10.2. Inference algorithms 138
10.3. The fusion algorithm 139
10.4. The complexity of the fusion algorithm 139
10.5. A fusion algorithm for contextuality 142
10.6. Message passing schemes and the collect algorithm 147
10.7. The complexity of the collect algorithm 149
10.8. A collect algorithm for contextuality 150
Discussion 153

Chapter VI. A complete characterisation of All-vs-Nothing arguments for stabiliser states 155
1. Overview 155
2. Mermin's proof 157
3. The stabiliser world 158
3.1. Stabiliser subgroups induce XOR theories 160
4. Characterising AvN arguments 161
4.1. AvN triples 161
4.2. Graph states 162
4.3. The AvN triple theorem and its consequences 163
5. Applications 167
5.1. Counting AvN triples 167
5.2. Generating AvN triples 169
Discussion 170

Chapter VII. Minimum quantum resources for strong non-locality 173
1. Overview 173
2. Background 175
2.1. Quantum models and scenarios 175
2.2. SLOCC classes of three-qubit states 177
3. Two-qubit states are not strongly non-local 178
4. W-SLOCC states are not strongly non-local 179
5. Strong non-locality in the SLOCC class of GHZ 181
5.1. The n-partite GHZ state and local equatorial measurements 181
5.2. Balanced GHZ-SLOCC states and local equatorial measurements 182

5.3. Further restrictions 184
6. A family of strongly non-local three-qubit models 186
Discussion 188

Chapter VIII. Conclusion 191

Bibliography 195

List of Symbols 207

Index 209

CHAPTER I

Introduction

1. General motivation

The advent of quantum physics marks one of the greatest scientific revolutions in human history. Sparked from the crisis of classical mechanics in the early 1900s, quantum mechanics developed into what is arguably the most complete physical theory known to date, and found far-reaching applications to countless areas of modern science. Quantum chemistry, quantum optics, light-emitting diodes, superconducting magnets, the laser, transistors and semiconductors, the microprocessor, medical and research imaging such as electron microscopy and MRI are only some of the fields and inventions spawned or heavily influenced by quantum mechanics. Despite this tremendous success, the advancement of the theory has been traditionally met with controversy and discomfort. Although this trait is shared by many revolutionary processes in history, the dispute over quantum theory, and particularly its foundations, stands out as being especially intricate and difficult to resolve.

The reason is to be found in the theory's highly counterintuitive character, which ignited a profound physical and philosophical debate, originally animated by some of the greatest physicists of the 20th century – Planck, Bohr, Heisenberg, Einstein, Schrödinger, Von Neumann, Dirac, Pauli, Bell to name a few notable examples – and still extremely lively to this day. This is because quantum mechanics challenged the foundations of physics more than any other theory before. In fact, one could even say that some aspects of the theory question the very idea of *physical reality*, and force us to reject our fundamental perceptions of the world we live in.

Non-locality and contextuality lie at the heart of this disconcerting proposition. When non-locality was first identified by Einstein, Podolsky and Rosen in their famous *EPR paradox* [**EPR35**], it was treated as an obscure aspect of quantum physics, a paradoxical trait that threatened the foundations of the theory itself. Indeed, non-locality violates what was then a widely accepted *criterion of reality*, which essentially requires a physical theory to assign well-defined predetermined values to every physical quantity. In fact, the existence of non-local behavior was simply deemed impossible. According to [**EPR35**], the only plausible explanation to what Einstein originally defined as a "spooky action at a distance" (non-locality) was the existence of underlying hidden elements of reality – or *hidden variables* – which would explain the illusory faster-than light exchanges of information observed by EPR's thought experiment. As a result, the article concluded that quantum physics is fundamentally *incomplete*.

This claim had to be abandoned after the formulation of classic no-go theorems by Bell [**Bel64**] and Kochen-Specker [**KS67**]. Indeed, these results show that non-locality and the more general concept of contextuality are unavoidable aspects of any theory which agrees with the predictions of quantum mechanics, elevating these phenomena from a potential 'bug' of quantum physics to fundamental physical features. The EPR paradox turned out not to be a paradox at all. It only showed how astonishingly unsuitable our intuition and classical conception of physics are when dealing with Nature's most complex interactions.

Although Bell's definite answer to the question of non-locality somewhat revived what was then a stagnant debate on quantum foundations, the general scientific community largely ignored the issue and focused on using the undoubted and unparalleled power of prediction of the new theory to develop novel fields and applications. Albeit a full understanding of the foundations of quantum physics was far from being achieved (and it still is) quantum theory was simply too good for physicists to care about its completeness and its most controversial aspects.

In the 1980s, the birth of quantum computation utterly changed this attitude. Computer scientists started to consider the peculiarities of quantum mechanics as valuable *resources* to break through the limits of classical computation, information and cryptography. The study of quantum foundations gained new relevance, and it now represents a well-established research field in physics, computer science, and mathematics.

While the hardware technology for quantum computers is still at an early stage, numerous potential applications of quantum mechanics to information processing have been identified on a theoretical scale: remarkable new algorithms, such as Shor's algorithm to quickly factorise large numbers [**Sho99**] and Grover's algorithm to efficiently search in unsorted databases [**Gro96**], cryptographic schemes such as quantum key distribution protocols [**BB14, Eke91**] and device-independent quantum cryptography [**MY98, VV14, MS16**], and novel perspectives on computational complexity [**Wat09**]. Realising this potential would revolutionise computation and information theory, and the benefits for the scientific community and the general public would be enormous.

While it is widely acknowledged that quantum computation offers significant advantages over classical computation, a full grasp of what the specific aspects of quantum theory enabling these advantages are is far from achieved. Many quantum computer scientists identify the reason for this in a fundamental lack of knowledge concerning the structure of quantum physics. In essence, 'quantum computers operate in a manner so different from classical computers that our techniques for designing algorithms and our intuitions for understanding the process of computation no longer work' [**Sho03**]. For this reason, fundamental research on the most non-classical aspects of quantum physics, such as entanglement, non-locality, contextuality and superposition, is necessary for the development of the field.

Recent work by Raussendorf [**Rau13**] and Howard et al. [**HWVE14**] identified contextuality as essential ingredients of quantum computation, showing that these phenomena are the pivotal source of power in the mainstream paradigms of measurement based quantum computation [**GC99, RB01b, RB01a, KLM01, RBB03, Nie03, Leu04, RBB03, BBD$^+$09, Rau13**] and magic state distillation [**BK05, Kni05, CAB12**]. These results suggest that understanding contextuality is not only of fundamental importance to the foundations of physics, as clarified above, but it is the key to potentially inaugurate a new era of computation.

2. Logical and topological aspects of contextuality

In 2011, Abramsky and Brandenburger introduced an abstract mathematical framework based on sheaf theory to describe non-locality and contextuality, thereby providing a common general theory for the study of these phenomena, which had been carried out in a rather concrete, example-driven fashion until then [**AB11a**].[1] This high-level description showed that contextuality is not a feature specific to quantum mechanics, but rather a general mathematical property. As such, it can be independently applied to other areas of computer science not necessarily related to quantum theory. This constitutes a remarkable observation, which motivates further research on contextual behaviour outside of quantum physics.

In their work, Abramsky and Brandenburger presented a hierarchy of different strengths of contextuality:

$$\text{Probabilistic contextuality} \; < \; \text{Logical contextuality} \; < \; \text{Strong contextuality},$$

all of which arise naturally in quantum mechanics. While probabilistic contextuality largely remains studied in relation to quantum theory, logical forms of contextuality (i.e. logical and strong) have been particularly prolific in the establishment of connections with other fields [**Abr14b**], with notable examples in relational databases [**Abr13a, Bar15a**], constraint satisfaction problems [**AGK13, ABdSZ17**], natural language semantics [**AS14**], and logical paradoxes [**ABK$^+$15, Kis16b, dS17**]. Contrary to the probabilistic case, logical and strong contextuality can be witnessed at the level of *possibilities*, thus exposing a deeper structure that turns out to be abundantly observable across different fields. For this reason, although probabilistic contextuality will play a substantial role in our study, this thesis will be mostly concerned with logical forms of contextuality.

The sheaf theoretic description of contextuality exposes the phenomenon's intrinsic nature as a *fundamental discrepancy between local consistency and global inconsistency* [**ABK$^+$15**], which finds a compelling illustration in a famous artwork by M. C. Escher portraying Penrose & Penrose's never-ending staircase [**PP58**] (Figure I.1). If one focuses on a *local* portion of the staircase, the picture appears to be a perfectly consistent description of reality. This is showed in Figure I.2, where the piece is split in four parts, each giving a faithful representation of a portion of a real staircase. The local consistency of Escher's staircase resides in that the four parts of Figure I.2 are compatible with each other: any local picture can be 'glued' to any other adjacent part to obtain a larger figure which is still perfectly consistent. However, once we glue everything together, the paradoxical aspect of the never-ending staircase immediately emerges, thus resulting in a *globally inconsistent* picture. This is the essence of contextuality in the sheaf theoretic framework: the impossibility of extending a locally compatible family of sections of a space to a globally consistent picture.

The viewpoint offered by Escher's lithograph suggests that contextuality has a *spatial* connotation. Following this idea, we are interested in understanding it from a purely

[1] It shall be mentioned that other general approaches to contextuality have appeared in recent years: Spekken's contextuality of preparations and unsharp measurements [**Spe05**], the graph-theoretic approach of Cabello, Severini and Winter [**CSW14**], and the combinatorial approach of Acín et al. [**AFLS15**].

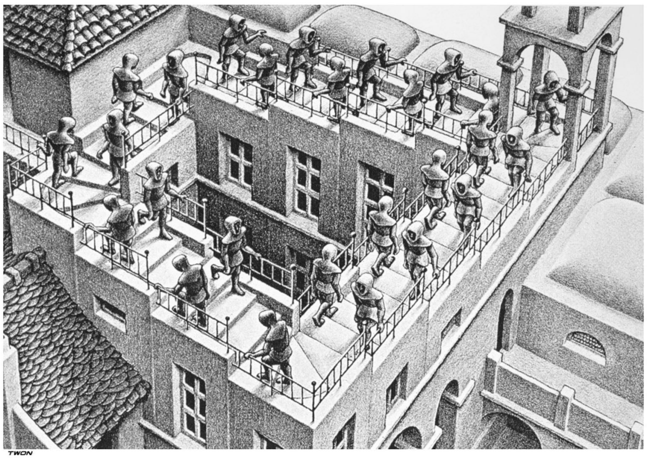


FIGURE I.1. M.C. Escher, detail from *Klimmen en dalen (Ascending and descending)*, 1960. Litograph, 35.6 cm $\times$ 28.6 cm

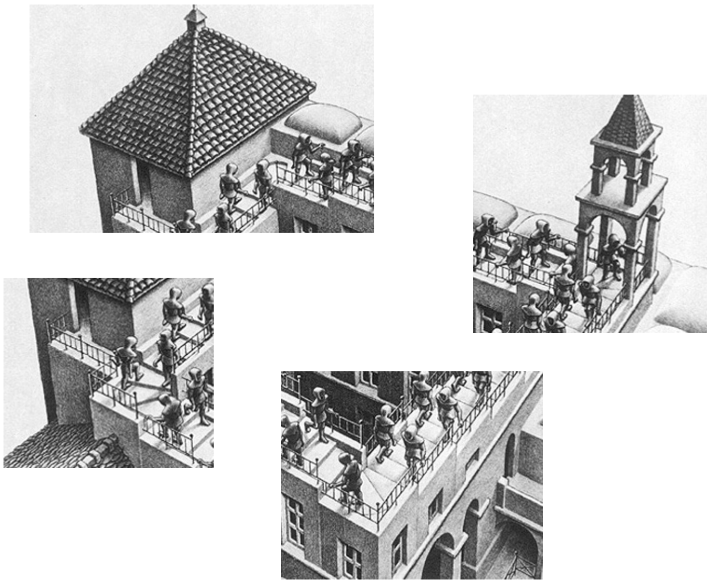


FIGURE I.2. Local consistency: individual parts of the staircase are self consistent and pairwise compatible.

topological standpoint. In particular, it is natural to ask whether classic tools of topology can be used to characterise the phenomenon. This hypothesis finds support in the work of R. Penrose himself, who developed topological methods based on *cohomology theory* to study *impossible figures* [**Pen92**] such as the never-ending staircase we just used to describe the quintessential nature of contextuality. Moreover, similar techniques have been extensively used to study the general problem of extending local properties to global ones in sheaf theory.

This thesis aims at achieving such a complete characterisation of logical forms of contextuality using cohomology theory, following the works of Abramsky et al. on sheaf cohomology [**AMB12, ABK$^+$15**]. Once such a topological description is established, we aim at extending the idea of local consistency vs global inconsistency to a higher level, with the intent of developing a general theory of *contextual semantics*, able to capture contextual behaviour in a variety of fields, including the aforementioned examples of contextuality beyond quantum mechanics. Our goal is to use this general theory to transfer methods and results from one field to the other, and to take advantage of these connections to better understand contextuality, both in quantum theory and beyond.

3. Contributions

Here, we outline the main contributions of this thesis:

- We analyse the limits of the current sheaf cohomological methods for contextuality. In particular, we characterise the structure of *false negatives*, and show that the current cohomological description does not constitute a complete invariant for strong contextuality, not even under strong symmetry and connectedness assumptions on the measurement scenario, disproving a previous conjecture of [**AMB12**]. We extend the theory of cohomological obstructions to higher cohomology groups, giving a definite answer to speculations on their usefulness to resolve the issue of false negatives: although higher obstructions do provide more information on the topological structure of the model, they cannot be employed to detect contextuality. We also introduce an alternative description of the cohomology obstructions using $\mathcal{F}$-torsors. This is presented in Chapter III, whose content has been published in [**Car17**].
- An (almost) complete sheaf cohomological invariant for logical and strong contextuality is introduced. The invariant is applicable to the vast majority of empirical models, including all the models appeared in the literature, and it is conjectured to be valid in general. The issue of false negatives is solved by introducing the novel constructions of *line* models and scenarios, which expose the deeper local extendability properties of the presheaves describing empirical models. The power of the invariant is demonstrated in a large number of examples, which include all the instances of false negatives known to date. This is the subject of Chapter IV, which has been presented at the 15th International Conference on Quantum Physics and Logic (QPL 2018). A pre-print is available at [**Car18**].
- A general definition of contextual behaviour is introduced in the language of valuation algebras [**She89, SSS$^+$90**]. This novel description naturally specialises to all the instances of contextual behaviour observed so far, both within and beyond quantum physics. Moreover, it extends the scope for contextuality to a variety of other domains and allows to translate theorems, methods and algorithms from one field to the other.

  New algorithms for the detection of logical and strong contextuality are developed using the connection established in the previous paragraph. Such algorithms are based on mainstream methods of *generic inference* [**She89, SS91,**

Koh03, Pou08], and are proven to outperform the current algorithms, especially in the case of logical contextuality. This is presented in Chapter V, whose content has been developed in collaboration with Samson Abramsky, and partially published in [**AC19**].

- A complete characterisation of All-vs-Nothing (AvN) arguments for stabiliser states is introduced. This is achieved thanks to the proof of the AvN triple conjecture, formulated in [**Abr14a, ABK$^+$15**]. This result allows to produce an exhaustive list of strongly contextual multi-qubit stabiliser states, which constitute fundamental resources for measurement based quantum computation [**GC99, RB01b, RB01a, KLM01, RBB03, Nie03, Leu04, RBB03, BBD$^+$09, Rau13**]. Moreover, it leads to other interesting structural results, such as the reducibility of every AvN argument to Mermin's proof of the strong contextuality of the GHZ model [**Mer90b, Mer90a**]. This is the subject of Chapter VI. The results have been presented in a joint paper with Samson Abramsky, Rui Soares Barbosa and Simon Perdrix, published in [**ABCP17**].
- We identify the minimum quantum resources needed to realise strong non-locality, and provide a partial classification of three-qubit states giving rise to this phenomenon. We show that no two-qubit system, with any finite number of local measurements, is sufficient. For three-qubit systems, we show that strong non-locality can only be realised in the GHZ SLOCC class, and with equatorial measurements. Within this class, we identify an infinite family of states not LU-equivalent to the GHZ state that realise strong non-locality with finitely many measurements. States in this class feature decreasing entanglement between one qubit and the other two, which has to be compensated by an increasing number of local measurements on the latter. This is the subject of Chapter VII, whose results have been published in [**ABC$^+$17**], in collaboration with Samson Abramsky, Rui Soares Barbosa, Nadish de Silva, Kohei Kishida and Shane Mansfield.

4. Statement of collaboration

The contents of Chapters III and IV have been completely and independently developed by the author. Chapter V is a joint collaboration with Samson Abramsky. The results of Chapter VI are based on preliminary findings of the author's Master's dissertation [**Car15**] and a joint paper with Samson Abramsky presented as a poster at the 13th International Conference on Quantum Physics and Logic (QPL2016) [**AC16**]; the final content of the Chapter is due to joint work with Samson Abramsky, Rui Soares Barbosa and Simon Perdrix. Chapter VII is the result of a collaboration with Samson Abramsky, Rui Soares Barbosa, Nadish de Silva, Kohei Kishida and Shane Mansfield.

5. Outline of the thesis

Chapter II presents some background on the sheaf-theoretic framework for non-locality and contextuality, with particular attention to sheaf cohomology. In Chapter III, we analyse the limits of the current cohomological framework for contextuality. The complete cohomology invariant for non-locality and contextuality is presented in Chapter IV. Chapter V introduces contextuality in valuation algebraic terms, and presents new algorithms to detect it. The complete characterisation of All-vs-Nothing arguments for stabiliser states is presented in Chapter VI. Chapter VII establishes the minimum quantum resources needed to realise strong non-locality. Finally, conclusions and possible future research directions are discussed in Chapter VIII.

CHAPTER II

Background: the sheaf theoretic structure of contextuality

Summary

This chapter presents Abramsky & Brandenburger's sheaf–theoretic description of non-locality and contextuality. This high-level mathematical framework allows to study contextuality independently of quantum mechanics, and sets the ground for a topological analysis of these highly non-classical phenomena. In particular, it allows the development of methods to detect contextuality based on sheaf cohomology.

1. Overview

Although non-locality and contextuality have been traditionally studied in the context of quantum mechanics, it is important to remark that the content of Bell's [**Bel64**] and Kochen–Specker's [**Bel66, KS67**] theorems applies not just to quantum physics, but to *any* theory that matches its predictions. In other words, no physical theory which agrees with quantum mechanics can be local or non-contextual. For this reason, it is desirable to describe these peculiar phenomena at an appropriate level of abstraction, without presupposing quantum physics.

Instead of dealing with the typical elements of a quantum setting, such as states and observables, our main subject of study will be *empirical models*, abstract structures which embody the empirical results of an ideal experiment, regardless of its physical implementation or theoretical interpretation.

As mentioned in the introduction, non-locality and contextuality can be elegantly thought of as a fundamental discrepancy between local consistency and global inconsistency in geometrical figures. This rather heuristic definition finds a compelling theoretical counterpart in the language of *sheaf theory* [**AB11a**], a powerful high-level mathematical framework, suited to study the extendability of local features to global ones.

In this chapter, we will review the sheaf–theoretic structure of non-locality and contextuality. Particular attention will be given to the line of research involving the use of cohomology theory to detect contextuality [**AMB12, ABK$^+$15, Car15, Car17, Car18, ORBR17, OTR18, Aas18**].

Outline of the chapter. In Section 2, we start by informally introducing the main concepts through the basic concrete example of Bell's model. In Section 3, a general description of measurement scenarios is introduced, and the alternative viewpoint based on abstract simplicial complexes is presented. We define empirical models abstractly using the language of sheaf theory in Section 4.4. Particular attention will be devoted to possibilistic empirical models and their topological representation as bundle diagrams. Section 5 defines non-locality and contextuality and introduces a hierarchy of different strengths of these phenomena. Finally, in Section 6 we introduce sheaf cohomology as a powerful method to detect contextuality.

2. A basic example

2.1. Measurement scenario. The best way to introduce the concept of empirical model is to examine a concrete example. Consider an ideal experimental setting, where two experimenters, Alice and Bob, perform measurements on a physical system. Suppose Alice has two measurements a_1 and a_2 at her disposal. She can choose which one to carry out, but she cannot perform them both simultaneously. Similarly, Bob can choose between measurements b_1 and b_2. Furthermore, suppose that all these measurements are dichotomic, i.e. they produce an outcome $o_A, o_B \in \{0, 1\}$.

At each run of the experiment, Alice and Bob choose a measurement to perform, and record the outcome observed. Each possible choice of joint measurements is called a **measurement context** or simply a **context**. In this particular scenario, the contexts are

$$\{a_1, b_1\}, \{a_1, b_2\}, \{a_2, b_1\}, \{a_2, b_2\}.$$

This measurement scenario is referred to as a $(2, 2, 2)$ scenario, to indicate that there are 2 parties, each with 2 possible measurements, and 2 outcomes for each measurement.

The structure of this simple scenario can be effectively represented by a graph, as shown in Figure II.1.

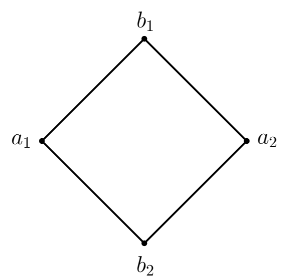


FIGURE II.1. A graphical representation of the measurement structure of the $(2, 2, 2)$ scenario. Each vertex represents a measurement, while edges correspond to contexts.

Notice that such a representation is the same for all $(2, 2, l)$ scenarios, as it does not reflect the fact that each measurement is dichotomic. In order to add this element into the picture, we introduce a *fibre* above each vertex, which represents the possible outcomes for the corresponding measurement, as shown on the left–hand diagram of Figure II.2. The result is a *bundle*-like picture, which will be used extensively throughout this thesis to represent scenarios and empirical models alike, and will be presented in more detail in Section 4.6.

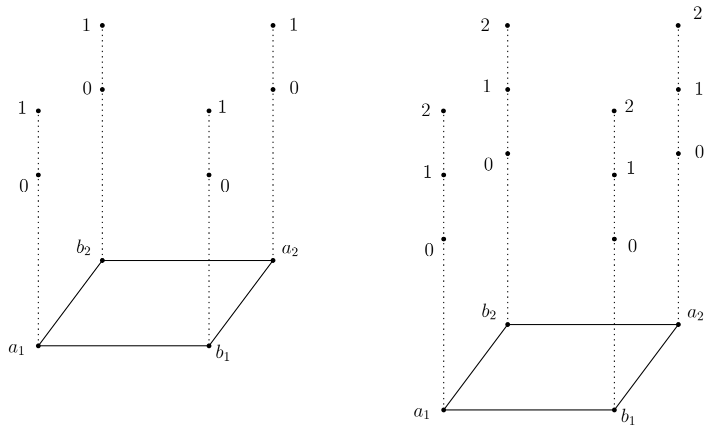


FIGURE II.2. On the left hand side, a bundle diagram representing a $(2, 2, 2)$ scenario. On the right, a $(2, 2, 3)$ scenario.

2.2. Empirical model. At each run of the experiment, Alice and Bob register an **event**, i.e. an assignment of outcomes to each of the measurements they have elected to perform. Such an event can be represented by an element of the function set $\{0, 1\}^C$, where C is the context determined by Alice's and Bob's choices of measurement. For instance, the situation where Alice chooses to perform a_1 and observes outcome 0, and Bob chooses to perform b_2 and obtains outcome 1 corresponds to the following event:

$$\{a_1 \mapsto 0, b_2 \mapsto 1\}.$$

By collecting the information on the joint outcomes of each run of the experiment, one obtains a probability distribution over each event at any particular context $\{a_i, b_j\}$, $i, j = 1, 2$. More formally, the statistics of the experiment can be summarised by a collection of distributions of the form

$$\mathsf{Prob}(o_A, o_B \mid a_i, b_j)$$

which express the probability of Alice and Bob obtaining outcomes o_A and o_B when choosing measurement a_i and b_j respectively. These probabilities intuitively constitute what we call an **empirical model**, although the formal definition of this concept, given in Section 4.4, is more general and allows to account for a much greater class of scenarios.

TABLE II.1. A representation of a general empirical model over a $(2, 2, 2)$ scenario as a probability table.

A	B	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
a_1	b_1	$\mathsf{Prob}(0,0 \mid a_1,b_1)$	$\mathsf{Prob}(1,0 \mid a_1,b_1)$	$\mathsf{Prob}(0,1 \mid a_1,b_1)$	$\mathsf{Prob}(1,1 \mid a_1,b_1)$
a_1	b_2	$\mathsf{Prob}(0,0 \mid a_1,b_2)$	$\mathsf{Prob}(1,0 \mid a_1,b_2)$	$\mathsf{Prob}(0,1 \mid a_1,b_2)$	$\mathsf{Prob}(1,1 \mid a_1,b_2)$
a_2	b_1	$\mathsf{Prob}(0,0 \mid a_2,b_1)$	$\mathsf{Prob}(1,0 \mid a_2,b_1)$	$\mathsf{Prob}(0,1 \mid a_2,b_1)$	$\mathsf{Prob}(1,1 \mid a_2,b_1)$
a_2	b_2	$\mathsf{Prob}(0,0 \mid a_2,b_2)$	$\mathsf{Prob}(1,0 \mid a_2,b_2)$	$\mathsf{Prob}(0,1 \mid a_2,b_2)$	$\mathsf{Prob}(1,1 \mid a_2,b_2)$

It is often convenient to present an empirical model as a **probability table**, such as Table II.1. Each row of the table represents a context, while each column corresponds to a joint outcome. The entries of the table are the empirically observed probabilities that constitute the empirical model.

Examples. To clarify the concepts introduced thus far, we present two key examples of empirical models that will be thoroughly studied throughout this thesis. First, consider the empirical model described in Table II.2.

TABLE II.2. Bell's empirical model.

A	B	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
a_1	b_1	$1/2$	0	0	$1/2$
a_1	b_2	$3/8$	$1/8$	$1/8$	$3/8$
a_2	b_1	$3/8$	$1/8$	$1/8$	$3/8$
a_2	b_2	$1/8$	$3/8$	$3/8$	$1/8$

This model is the key element of the CHSH proof [**CHSH69, Bel87**] of Bell's theorem [**Bel64**], and constitutes, as we shall see in the next section, a prime example of contextual behaviour in quantum mechanics. The probabilities in the table are obtained by interpreting the measurement labels as particular single-qubit projective measurements which have $+1$ eigenvectors separated by a $\pi/3$ angle in the XY-plane of the Bloch sphere, and applying them to the **Bell state**:

$$\left|\Phi^+\right\rangle := \frac{|00\rangle + |11\rangle}{\sqrt{2}}.$$

As mentioned in the introduction, our aim is to study non-locality and contextuality independently of quantum physics. For this reason, we will deal with many instances of empirical models that do not arise from quantum mechanics. For instance, the Popescu–Rohrlich box model [**PR94, Ras85, KT85**], displayed in Table II.3, cannot be realised by any choice of quantum state and observables.

TABLE II.3. The PR-box model.

A	B	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
a_1	b_1	$1/2$	0	0	$1/2$
a_1	b_2	$1/2$	0	0	$1/2$
a_2	b_1	$1/2$	0	0	$1/2$
a_2	b_2	0	$1/2$	$1/2$	0

2.3. Contextuality. From the point of view of classical physics, we are naturally led to believe that when Alice and Bob perform their measurements and observe the corresponding outcomes, they are simply looking at a portion of a predetermined assignment of outcomes to *all* of the measurements, which is completely independent of their choice. In particular, at any given run of a $(2, 2, 2)$ scenario, we assume that a_1, b_1, a_2, b_2 all have well-defined values, e.g.

$$\{a_1 \mapsto 0, b_1 \mapsto 0, a_2 \mapsto 1, b_2 \mapsto 1\}, \tag{II.1}$$

even though Alice and Bob can only observe two of them, say

$$\{a_1 \mapsto 0, b_2 \mapsto 1\}.$$

With this assumption, we expect the probabilities observed over many runs of the experiment to be simply generated by a *global probability distribution* over all global assignments similar to (II.1), as shown with an example in Table II.4.

TABLE II.4. A probability distribution over global assignments on the left gives rise to an empirical model on the right.

$a_1b_1a_2b_2$	Prob.
0000	$1/8$
0011	$1/4$
0100	$1/4$
1000	$1/4$
1010	$1/8$
Other	0

$\longrightarrow$

A	B	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
a_1	b_1	$3/8$	$3/8$	$1/4$	0
a_1	b_2	$3/8$	$3/8$	$1/4$	0
a_2	b_1	$3/8$	$3/8$	$1/4$	0
a_2	b_2	$5/8$	$1/8$	0	$1/4$

Given an empirical model, it is natural to ask ourselves what is the global distribution underlying the empirically observed probabilities. However, as it turns out, there exist models where such a global distribution cannot be found. This phenomenon is called *non-locality* or, more generally, *contextuality*.[1]

A proof of contextuality. Let us give a first example of a contextuality proof based on the concept of *logical Bell inequality* [**AH12**], a formal counterpart of the notion of *Bell inequality* [**CHSH69**], which is widely studied in the quantum literature [**Tsi80, Fin82, ADR82, KT85**].

Suppose we have N propositional formulae $\varphi_1, \ldots, \varphi_N$. We think of the Boolean variables appearing in each formula as empirically testable quantities. Thus, each φ_i corresponds to a certain statement on the results of an experiment involving these quantities. Given a probability distribution for the outcomes of the experiment, it is possible to assign a probability p_i to each formula φ_i representing its likelihood to be satisfied by the experiment. Let $\Phi := \bigwedge_{i=1}^{N} \varphi_i$ and $P := \mathsf{Prob}(\Phi)$. Then,

(II.2)
$$1 - P = \mathsf{Prob}(\neg\Phi) = \mathsf{Prob}\left(\bigvee_{i_1}^{N} \neg\varphi_i\right) \leq \sum_{i=1}^{N} \mathsf{Prob}(\neg\varphi) = \sum_{i=1}^{N}(1 - p_i) = N - \sum_{i=1}^{N} p_i.$$

[1] The precise definition of these two concepts will be given in Section 5. Until then, it is sufficient to know that non-locality is a special case of contextuality, where the experimental setting in question is multipartite as in the example we have just seen.

Now, suppose φ_i cannot be all satisfied at the same time, then $P = 0$. Hence, inequality (II.2) becomes

$$\sum_{i=1}^{N} p_i \leq N - 1, \tag{II.3}$$

and constitutes an example of a **logical Bell inequality**. Let us consider again Bell's model presented in Table II.2. In fact, it is sufficient to focus only on a subset of the entries, highlighted in Table II.5.

TABLE II.5. Some entries of the Bell's model

A	B	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
a_1	b_1	$1/2$			$1/2$
a_1	b_2	$3/8$			$3/8$
a_2	b_1	$3/8$			$3/8$
a_2	b_2		$3/8$	$3/8$	

If we interpret the experimental outcomes 1, 0 as true and false respectively, each event can be characterised by a propositional formula. For instance, the top left entry, which corresponds to the event $\{a_1 \mapsto 0, b_1 \mapsto 0\}$, can be represented by the formula $\varphi : \neg a_1 \wedge \neg b_1$. Following the same idea, one can associate to each row of Table II.5 a formula describing the events in question:

$$\begin{aligned}
\varphi_1 &: (\neg a_1 \wedge \neg b_1) \vee (a_1 \wedge b_1) \equiv a_1 \Leftrightarrow b_1 \\
\varphi_2 &: (\neg a_1 \wedge \neg b_2) \vee (a_1 \wedge b_2) \equiv a_1 \Leftrightarrow b_2 \\
\varphi_3 &: (\neg a_2 \wedge \neg b_1) \vee (a_2 \wedge b_1) \equiv a_2 \Leftrightarrow b_1 \\
\varphi_4 &: (a_2 \wedge \neg b_2) \vee (\neg a_2 \wedge b_2) \equiv a_2 \oplus b_2
\end{aligned}$$

It is straightforward to see that these formulas are jointly contradictory, in fact

$$a_1 \overset{\varphi_1}{\Longleftrightarrow} b_1 \overset{\varphi_3}{\Longleftrightarrow} a_2 \overset{\varphi_4}{\Longleftrightarrow} \neg b_2 \overset{\varphi_2}{\Longleftrightarrow} \neg a_1.$$

The pairs of events highlighted in Table II.5 are mutually exclusive, thus the probability assigned to each formula is given by the sum of the two probabilities of the corresponding row, e.g. $p_1 = 1/2 + 1/2 = 1$. Therefore,

$$\sum_{i=1}^{4} p_i = 1 + \frac{6}{8} + \frac{6}{8} + \frac{6}{8} = \frac{13}{4} > 3,$$

which is a violation of (II.3).

How is this possible? Each formula φ_i involves only a portion of the Boolean variables in $\{a_1, b_1, a_2, b_2\}$, while Φ contains them all. The invalid step in the argument resides in the assignment of a probability to Φ. Indeed, such an assignment can be made only if there is a *global assignment of probabilities* to all of the variables simultaneously which yields the empirically observed probabilities. The very fact that Bell's model violates inequality (II.3) indicates that such an assignment does not exist, and thus the model is contextual.

Contextuality and quantum computation. Despite being originally considered an obscure, paradoxical phenomenon that threatened the foundations of quantum theory, contextuality has recently gained great relevance as a key resource for quantum computation [**Rau13, HWVE14, ABM17**]. In order to give a taste of how contextuality can be used to achieve faster computation, let us present a simple example based on [**AB09**].

Suppose we have a classical computer, which is only capable of doing addition modulo 2. Of course, such a computer is very limited, and far from being classically universal. Our goal is to show that this simple computer can be promoted to classical universality by granting it access to a contextual resource.

Consider the empirical model partially displayed in Table II.6, based on a $(3, 2, 2)$-scenario.[2]

TABLE II.6. Four rows of the GHZ model.

A	B	C	000	001	010	011	100	101	110	111
a_1	b_1	c_1	$1/4$	0	0	$1/4$	0	$1/4$	$1/4$	0
a_1	b_2	c_2	0	$1/4$	$1/4$	0	$1/4$	0	0	$1/4$
a_2	b_1	c_2	0	$1/4$	$1/4$	0	$1/4$	0	0	$1/4$
a_2	b_2	c_1	0	$1/4$	$1/4$	0	$1/4$	0	0	$1/4$

This model, known as the **Greenberger-Horne-Zeilinger (GHZ) model** [**GHZ89, GHSZ90**], is obtained by applying Pauli measurements X and Y to the **Greenberger-Horne-Zeilinger (GHZ) state**

$$|\mathsf{GHZ}\rangle := \frac{|000\rangle + |111\rangle}{\sqrt{2}},$$

and will be studied more in detail in Section 5.2.

We will now show that, thanks to the GHZ model, the classical computer limited to addition modulo 2 is capable of computing the OR function. Note that this is enough to achieve universality, since negation is already available to the classical computer (it is just an addition $\oplus 1$), and these two gates are sufficient to compute any Boolean function. We interpret bits as instructions for the three parties Alice, Bob and Charlie about which measurements to choose. We interpet 0 as an instruction to perform their first respective measurement, i.e.

$$0 \mapsto a_1, b_1, c_1$$

and 1 for their second measurement, i.e.

$$1 \mapsto a_2, b_2, c_2.$$

Now, given two input bits i_1, i_2, we let the classical computer calculate $i_1 \oplus i_2$. Then, the choices of measurement for Alice, Bob and Charlie are determined by $\langle i_1, i_2, i_1 \oplus i_2 \rangle$. For instance, if the input bits are $i_1 = 0$ and $i_2 = 1$, the parties will choose to perform measurements a_1, b_2 and c_2 respectively.

[2]The complete information on the empirical model can be displayed in a table with 8 rows. For our purposes, we will only need the ones presented in Table II.6.

After the corresponding measurements are performed, the classical computer outputs the sum $o_A \oplus o_B \oplus o_C$ of the three outcomes obtained. By looking at Table II.6, one can see that the possible events of context $\{a_1, b_1, c_1\}$, corresponding to the input bits $\langle 0, 0\rangle$, are those whose sum of outcomes modulo 2 is 0. Therefore, the output of $\langle 0, 0\rangle$ will be 0, regardless of the individual outcomes observed. For the other contexts, which correspond to the remaining possible pairs of inputs, the possible events are those whose sum of outcomes modulo 2 is 1, hence their output is 1. We conclude that the function computed is indeed $\mathsf{OR}(i_1, i_2)$.

This simple example shows how the limited computational power of a simple machine can be improved by quantum resources. While this particular instance is of little practical interest, it represents the starting point for a much more profound connection between contextuality and the computational power of a particular model for quantum computation, observed in [**Rau13**]. Indeed, the example we just presented is a nothing but a simple *measurement based quantum computation* (MBQC), i.e. a process which consists of preparing an entangled resource state – in this case the GHZ state – and performing on it single qubit measurements selected by a classical linear co-processor. The key finding of [**Rau13**], subsequently refined in [**ABM17**], is that *all* MBQCs which compute a non-linear Boolean function with sufficiently high probability are contextual, and that the probability of success increases with the amount of contextuality in the computation.[3] Since the MBQC model with suitable resource states achieves quantum universality [**GE07, VdNMDB06**], this result strongly suggests that contextuality is the key element of quantum theory which enables quantum computers to outperform their classical counterparts, an observation supported by the earlier work of Howard et al. [**HWVE14**], which established the importance of contextuality in the magic state distillation model for quantum computing [**BK05, Kni05, CAB12**]. Although quantum computation will not be directly investigated in this thesis, the crucial role played by contextuality in this area constitutes a major motivation for a formal understanding of this phenomenon.

3. Measurement scenarios

Guided by the example of Section 2, we now introduce a general definition of measurement scenario that captures the structure of experimental settings in their most general sense.

Definition II.1. A **measurement scenario** is a triple $\langle X, \mathcal{M}, (O_m)_{m\in X}\rangle$, comprised of

- A finite **set of measurements** X.
- A **measurement cover** $\mathcal{M} \subseteq \mathcal{P}(X)$, whose elements are called **contexts**.
- A finite **set of outcomes** O_m for each measurement $m \in X$.

Two or more measurements are said to be **compatible** if they are contained in a context. The collection $\mathcal{M}$ is a cover, i.e. such that

$$\bigcup_{C\in\mathcal{M}} C = X,$$

[3]Contextuality in empirical models can be quantified via the so-called *contextual fraction*, which will be reviewed in Section 10.1.1 of Chapter V.

and it is assumed to be an **antichain**, i.e. such that, for all $C, C' \in \mathcal{M}$, if $C \subseteq C'$, then $C = C'$.

A measurement scenario $\langle X, \mathcal{M}, (O_m) \rangle$ is said to be **quantum realisable** if one can associate to each measurement label in X a quantum measurement in the same Hilbert space such that measurements in the same context commute.

The set of measurements contains the labels of all the measurements considered in the experiment. The measurement cover contains the contexts, i.e. the maximal sets of jointly performable measurements. The antichain condition guarantees their maximality. Finally, each set of outcomes O_m represents the possible outcomes that measurement m can produce. Note that, in many scenarios, the set of outcomes is the same for all measurements. In this case, we shall denote by O the unique set of outcomes.

Example II.2.

- In the example of Section 2, the scenario is determined by the following elements:
  - $X = \{a_1, b_1, a_2, b_2\}$.
  - $\mathcal{M} = \{\{a_1, b_1\}, \{a_1, b_2\}, \{a_2, b_1\}, \{a_2, b_2\}\}$.
  - $O = \{0, 1\}$.
- The $(3, 2, 2)$ scenario underlying the GHZ model of Table II.6 is described as follows:
  - $X = \{a_1, b_1, c_1, a_2, b_2, c_2\}$.
  - $\mathcal{M} = \{\{a_i, b_j, c_k\} \mid i, j, k = 1, 2\}$.
  - $O = \{0, 1\}$.
- An abstract experimental setting can be described by any specification of measurements and contexts, even in the absence of a clear physical interpretation as in the concrete examples described above. For instance:
  - $X = \{a, b, c, d, e, f, g\}$.
  - $\mathcal{M} = \{\{a, b, c\}, \{a, b, d\}, \{a, c, d\}, \{b, c, d\}, \{b, e, f\}, \{e, g\}\}$,
  - $O = \{0, 1, 2, 3\}$.

  is a valid scenario

3.1. Bell-type scenarios. Among the infinitely many kinds of scenarios we can define, **Bell-type scenarios** deserve particular attention. These scenarios are a general version of the experimental setting introduced in Section 2. The common feature of this class of scenarios is their multi-partite character.

Definition II.3. A scenario $\langle X, \mathcal{M}, (O_m) \rangle$ is said to be Bell-type if

- The measurement set X can be partitioned into subsets $\{X_i\}_{i \in I}$, where I labels different 'parts' of the system, and X_i represents the measurements that can be carried out at part i.
- The cover $\mathcal{M}$ consists of the contexts of the form $\{x_i\}_{i \in I}$, where $x_i \in X_i$ for all $i \in I$. This corresponds to performing one and only one measurement for each part of the system.

An important subclass of Bell-type scenarios are (n, k, l) **scenarios**, where $n, k, l \in \mathbb{N}$. An (n, k, l) scenario is a Bell-type scenario where n parties have k measurements available, each having l possible outcomes.

3.2. Measurement scenarios and simplicial complexes. Measurement scenarios find an elegant representation in terms of **abstract simplicial complexes** [**AB11b, Bar14, Bar15a**], a purely combinatorial description with a straightforward geometrical interpretation.

Definition II.4. An (abstract) **simplicial complex** on a set of vertices V is a collection $\Sigma \subseteq \mathcal{P}_{\mathsf{fin}}(V)$ such that

- $\emptyset \in \Sigma$.
- For all $v \in V$, $\{v\} \in \Sigma$.
- Σ is downward closed: for all $\sigma \in \Sigma$ and $\tau \subseteq V$, if $\tau \subseteq \sigma$, then $\tau \in \Sigma$. This is equivalent to saying that $\Sigma = \downarrow_{\subseteq}\Sigma$, where

  $$\downarrow_{\subseteq}\Sigma := \{\tau \in \mathcal{P}_{\mathsf{fin}}(V) \mid \exists \sigma \in \Sigma : \tau \subseteq \sigma\}$$

  denotes the *downward closure* of Σ.

The elements of V are called the **vertices** of Σ. In general, the set of vertices of a simplicial complex Σ is denoted by $V(\Sigma)$. Subsets $\sigma \in \Sigma$ are called **faces** or **simplices**. Maximal faces under inclusion are called **facets**, and we denote by $\max \Sigma$ the set of facets of Σ. The **dimension of a face** σ is defined by $\dim(\sigma) := |\sigma| - 1$. The **dimension** of Σ, denoted $\dim \Sigma$ is the maximum among the dimensions of its faces. For all $q \geq 0$, we denote by Σ^q the set of q–simplices:

$$\Sigma^q := \{\sigma \in \mathcal{P}_{\mathsf{fin}}(X) \mid \dim(\sigma) = q\},$$

so that $\Sigma = \bigcup_{q=0}^{\dim \Sigma} \Sigma^q$.

Given a measurement scenario $\langle X, \mathcal{M}, (O_m) \rangle$, one can associate to it a simplicial complex Σ with vertices in X defined by

$$\Sigma := \downarrow_{\subseteq}\mathcal{M}.$$

Conversely, every simplicial complex can be interpreted as the basis for a measurement scenario. Indeed, given a simplicial complex Σ, one can define $X := V(\Sigma)$, and $\mathcal{M} := \max \Sigma$. It is then sufficient to specify the outcome sets $(O_x)_{x \in V(\Sigma)}$ for each vertex to obtain a well-defined scenario $\left\langle V(\Sigma), \max \Sigma, (O_x)_{x \in V(\Sigma)} \right\rangle$.

Example II.5.

- We have already seen an example of a simplicial representation of a measurement scenario in Figure II.1, where the simplicial complex describing a $(2, 2, l)$ scenario is pictured. The formal definition of the complex is

  $$\Sigma = \{\emptyset, \{a_1\}, \{b_1\}, \{a_2\}, \{b_2\}, \{a_1, b_1\}, \{a_1, b_2\}, \{a_2, b_1\}, \{a_2, b_2\}\}.$$

- The $(3, 2, 2)$ scenario onto which the GHZ model of Table II.6 is defined, corresponds to the simplicial complex shown in Figure II.3.
- The definition of measurement scenario is extremely flexible and accomodates all sorts of measurement compatibility structure. For instance, the simplicial complex pictured in Figure II.4 corresponds to the abstract scenario introduced at the end of Example II.2.

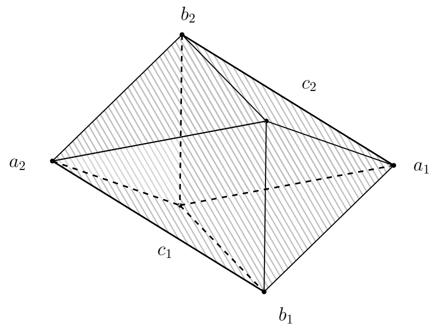


FIGURE II.3. Simplicial complex representation of a $(3, 2, l)$ scenario: a hollow octahedron.

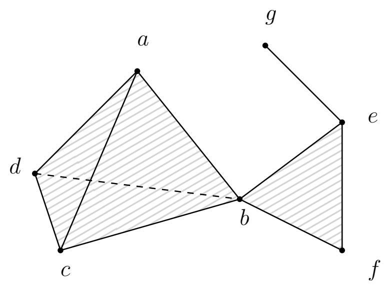


FIGURE II.4. A general simplicial complex which can be interpreted as a measurement scenario.

This viewpoint can be further extended, and most of the theory presented in the rest of this chapter can be equivalently formulated in terms of simplicial complexes, simplicial maps, simplicial quotients and fibrations [**Bar15a**]. Although this framework will not be discussed any further in this thesis, we will often use simplicial complexes to graphically represent scenarios and empirical models in order to visually investigate their contextual properties. In particular, we will largely take advantage of the **bundle diagram** representation, which will be presented in Section 4.6.

4. Empirical models

In the previous section, we presented a general description of experimental settings. Now, we introduce an abstract characterisation of the probabilistic results of such experiments. The main concepts will be defined in the language of **sheaf theory**.

4.1. Sheaf theory. In simplest terms, a (pre)**sheaf** is a mathematical object suited to track locally defined data associated to the open sets of a topological space. The data can be restricted to smaller subsets and 'glued' together to build bigger structures. In the area of contextuality, the data in question consist of probability distributions that characterise the empirical results of an experiment. Due to the peculiar structure of measurement scenarios, such data can only be defined locally to each context, and we are left to determine whether these local data can be consistently and coherently merged into a global probability distribution that explains them classically.

DEFINITION II.6. Let X be a topological space, and $\mathcal{C}$ a category. A $\mathcal{C}$-valued **presheaf** on X is a functor

$$F : \mathbf{Open}(X)^{op} \longrightarrow \mathcal{C},$$

where $\mathbf{Open}(X)$ denotes the poset category generated by the open sets of X, ordered by inclusion.

- Given any open set $U \subseteq X$, the elements of $F(U)$ are called **local sections**, or simply **sections** at U. Elements of $F(X)$ are called **global sections**.
- For each pair of open sets $U \subseteq U'$ of X, the map

$$\rho_U^{U'} := F(U \subseteq U') : F(U') \longrightarrow F(U)$$

  is called a **restriction map**. If $s \in F(U')$, its restriction $\rho_U^{U'}(s)$ to U' is often denoted $s|_U$, in analogy with function restriction.
- Two sections $s \in F(U)$, $s' \in F(U')$ are said to be **compatible** if

$$s|_{U \cap U'} = s'|_{U \cap U'}.$$

  By extension, a family $\{s_i \in F(U_i)\}_{i \in I}$ of sections of F is said to be **compatible** if its members are pairwise compatible.
- A presheaf $S : \mathbf{Open}(X) \to \mathbf{Set}$ is said to be a **subpresheaf** of a presheaf $F : \mathbf{Open}(X) \to \mathbf{Set}$ if $S(U) \subseteq F(U)$ for all $U \in \mathbf{Open}(X)$, and they share the same restriction maps.

DEFINITION II.7. A **sheaf** on X is a presheaf F on X that satisfies the following property: given an open cover $\mathcal{U} = \{U_i\}_{i \in I}$ of X, and a compatible family $\{s_i \in F(U_i)\}_{i \in I}$, there exists a unique global section $g \in F(X)$ such that $g|_{U_i} = s_i$ for all $i \in I$.

The sheaf condition says that pairwise consistent local data can always be glued to form global data in a unique way.

4.2. Events. Let us fix a measurement scenario $\langle X, \mathcal{M}, (O_m) \rangle$. An **event** occurs when a set of compatible measurements is performed, and their outcomes are observed. Formally, if $U \subseteq X$ is a set of compatible measurements, an **event over** U is described by a tuple in $\prod_{m \in U} O_m$. If all the measurements share the same outcome set O, this reduces to an element of O^U. Events occur locally inside a context, and thus can be effectively described by a sheaf.

DEFINITION II.8. The **sheaf of events** of a scenario $\langle X, \mathcal{M}, (O_m) \rangle$ is the sheaf

$$\mathcal{E} : \mathcal{P}(X)^{op} \longrightarrow \mathbf{Set}$$

on the set X, seen as a discrete topological space, defined by

- For all $U \subseteq X$,

$$\mathcal{E}(U) := \prod_{m \in U} O_m.$$

- For all $U \subseteq U' \subseteq X$, the corresponding restriction map of $\mathcal{E}$ is given by cartesian projection

$$\mathcal{E}(U \subseteq U') := \pi_{\mathcal{E}(U)} : \prod_{m \in U'} O_m \longrightarrow \prod_{m \in U} O_m :: \langle o_m \rangle_{m \in U'} \longmapsto \langle o_m \rangle_{m \in U}$$

In the case of a unique outcome set O, the definition of $\mathcal{E}$ reduces to $\mathcal{E}(U) := O^U$, with restriction maps coinciding with function restriction

$$\mathcal{E}(U \subseteq U') : O^{U'} \longrightarrow O^U :: s \longmapsto s|_U$$

It is quite simple to see that $\mathcal{E}$ does satisfy the sheaf condition.

4.3. Event distributions. In order to define empirical models, we will need to describe the concept of *distribution* at an appropriate level of generality. In particular, we will relax the notion of *probability*, and allow distributions to be defined over an arbitrary semiring.

DEFINITION II.9. Let R be a semiring. An **R-distribution** on a set S is a function $d : S \to R$ such that its support

$$\mathsf{supp}(d) := \{s \in S \mid d(s) \neq 0\}$$

is finite and $\sum_{s \in S} d(s) = 1$. The **R-distribution functor**

$$\mathcal{D}_R : \mathbf{Set} \longrightarrow \mathbf{Set}$$

assigns to a set S the set $\mathcal{D}_R(S)$ of R-distributions on S, and to any function $f : S \to T$, the function[4]

$$\mathcal{D}_R(f) : \mathcal{D}_R(S) \longrightarrow \mathcal{D}_R(T) :: d \longmapsto \lambda t. \sum_{\substack{s \in S: \\ f(s) = t}} d(s).$$

DEFINITION II.10. The **presheaf of event R-distributions** $\mathcal{D}_R\mathcal{E}$ is defined as the composition $\mathcal{D}_R \circ \mathcal{E}$. That is,

- For all $U \subseteq X$, $\mathcal{D}_R\mathcal{E}(U)$ is the set of R-distributions over $\mathcal{E}(U)$.
- For all $U \subseteq U' \subseteq X$, the corresponding restriction map of $\mathcal{D}_R\mathcal{E}$ is defined as

$$\mathcal{D}_R\mathcal{E}(U \subseteq U') : \mathcal{D}_R\mathcal{E}(U') \longrightarrow \mathcal{D}_R\mathcal{E}(U) :: d \longmapsto d|_U,$$

  where, for all $s \in \mathcal{E}(U)$,

$$d|_U(s) := \sum_{\substack{t \in \mathcal{E}(U'): \\ t|_U = s}} d(t).$$

Contrary to the sheaf of events, the presheaf $\mathcal{D}_R\mathcal{E}$ fails to satisfy the sheaf condition. In fact, as we shall see in Section 5, the impossibility of merging local probability distributions to obtain a global one is a central aspect of the sheaf–theoretic definition of non-locality and contextuality.

[4]The notation we use for the definition of $\mathcal{D}_R(f)$ is borrowed from lambda calculus. The map $\lambda x.f(x)$ is the map that takes x as an input and outputs $f(x)$.

4.4. Empirical models. As anticipated in the introductory sections of this chapter, an empirical model is an assignment of a probability distribution for each context of a measurement scenario. We require such probabilities to satisfy an additional compatibility property, which is a generlisation of **no-signalling**, or *no-disturbance* [**GRW80**]. In its traditional formulation, the no-signalling principle states that, given a Bell-type scenario, the choice of measurement by one (or more) parties should not affect the probability distributions of the other parties. In the familiar setting of a $(2,2,2)$ scenario, this corresponds to the following statement: given Alice's choice of measurement $a \in \{a_1, a_2\}$, her probability of observing outcome $o_A \in O_a$ should be independent of Bob's choice of measurement:

$$\sum_{o_B \in O_{b_1}} \mathsf{Prob}(o_A, o_B \mid a, b_1) = \sum_{o_B \in O_{b_2}} \mathsf{Prob}(o_A, o_B \mid a, b_2)$$

This means that the probability distributions over the contexts $\{a, b_1\}$ and $\{a, b_2\}$ marginalise to the same distribution on the intersection $\{a\} = \{a, b_1\} \cap \{a, b_2\}$:

$$\mathsf{Prob}(o_A, o_B \mid a, b_1)|_{\{a\}} = \mathsf{Prob}(o_A, o_B \mid a, b_2)|_{\{a\}}, \ \forall o_A, o_B.$$

The requirement of no-signalling is due to the fact that we assume Alice and Bob (and any other party involved in any Bell-type scenario) to be space-like separated in the relativistic sense. In this case, a violation of no-signalling corresponds to faster-than-light exchange of information between the parties, which contradicts the laws of special relativity. It has been proved several times and with different techniques that the probability distributions predicted by quantum mechanics in Bell-type quantum scenarios do not violate no-signalling [**GRW80, Bus82, Jor83, Shi84, Red87, SB93, Ken95**], and it has been recently shown that the result can be naturally extended to arbitrary quantum scenarios [**AB11a**], although the relativistic interpretation provided above fails in the absence of a multipartite setting.

This discussion motivates the following definition:

Definition II.11. A (no-signalling) **empirical model** over a measurement scenario $\langle X, \mathcal{M}, (O_m) \rangle$ is a compatible family

$$\{e_C \in \mathcal{D}_R\mathcal{E}(C)\}_{C \in \mathcal{M}}.$$

Explicitly, it is a family of probability distributions e_C over each context $C \in \mathcal{M}$ such that

$$e_C|_{C \cap C'} = e_{C'}|_{C \cap C'},$$

for any $C, C' \in \mathcal{M}$. If $R = \mathbb{R}_{\geq 0}$, the non–negative reals, we say that the empirical model is **probabilistic**, if $R = \mathbb{B}$, the Booleans, it is called **possibilistic**.

A probabilistic empirical model $\{e_C\}_{C \in \mathcal{M}}$ on a scenario $\Sigma = \langle X, \mathcal{M}, (O_m) \rangle$ is said to be **quantum realisable** if Σ is quantum realisable and there exists a quantum state $|\psi\rangle$ such that each probability distribution e_C is given by the Born rule applied to $|\psi\rangle$.

By choosing different semirings R, one can define other various kinds of empirical models. Each class of models may present different contextuality properties. For instance, the case where $R = \mathbb{R}$ allows for *negative probabilities*, which have been thoroughly studied in relation with quantum mechanics and proved to give rise to a fully non-contextual theory [**Wig32, Dir42, Moy49, Fey87, SR93, AB11a, AB14**].

4.5. Possibilistic models. Possibilistic models represent the main subject of study of this thesis. Intuitively, they can be thought of as *possibilistic collapses* of probabilistic models, where the values of the individual proabilities are neglected, and only the information regarding which events are possible (i.e. with probability > 0) is taken into account. More formally, a probabilistic model $e = \{e_C \in \mathcal{D}_{\mathbb{R}_{\geq 0}}\mathcal{E}(C)\}_{C\in\mathcal{M}}$ gives rise to a possibilistic model $\tilde{e} = \{\tilde{e}_C \in \mathcal{D}_{\mathbb{B}}\mathcal{E}(C)\}_{C\in\mathcal{M}}$, where $\tilde{e}_C := \chi_{\mathsf{supp}(e_C)} : \mathcal{E}(C) \to \mathbb{B}$ is the indicator function of $\mathsf{supp}(e_C)$. For instance, the PR-box model of Table II.3 collapses to the **possibility table** II.7.

TABLE II.7. The possibilistic collapse of the PR-box model. Each event labelled with a '1' is possible; those labelled with a '0' are impossible.

A	B	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
a_1	b_1	1	0	0	1
a_1	b_2	1	0	0	1
a_2	b_1	1	0	0	1
a_2	b_2	0	1	1	0

However, possibilistic models do not solely arise in this form: there exist possibilistic models that are not the possibilistic collapse of any probabilistic model, as shown in [**Abr13b, ABK$^+$16**]. This brings us to the following equivalent definition of possibilistic models in sheaf theory.

Given a probabilistic model $e = \{e_C\}_{C\in\mathcal{M}}$ over a scenario $\langle X, \mathcal{M}, (O_m)\rangle$, its possibilistic collapse $\{\tilde{e}_C\}_{C\in\mathcal{M}}$ can be described as a subpresheaf $\mathcal{S}_e$ of $\mathcal{E}$, where, for each subset $U \subseteq X$, $S_e(U)$ identifies the subset of possible events at U:

$$\mathcal{S}_e(U) := \{s \in \mathcal{E}(U) \mid s|_{U\cap C} \in \mathsf{supp}\,(e_C|_{U\cap C}) \;\; \forall C \in \mathcal{M}\} \subseteq \mathcal{E}(U). \tag{II.4}$$

By abstracting from this situation, one can reformulate the definition of possibilistic empirical model (Definition II.11) as follows:

DEFINITION II.12. A **possibilistic empirical model** over a scenario $\langle X, \mathcal{M}, (O_m)\rangle$ is a subpresheaf $\mathcal{S}$ of $\mathcal{E}$ such that

(1) $\mathcal{S}(C) \neq \emptyset$ for all $C \in \mathcal{M}$.
(2) $\mathcal{S}$ is **flasque beneath the cover**, i.e. the restriction map $\rho^{U'}_U = \mathcal{S}(U \subseteq U')$ is surjective whenever $U \subseteq U' \subseteq C$ for some context $C \in \mathcal{M}$.
(3) Every compatible family on the cover $\{s_C \in \mathcal{S}(C)\}_{C\in\mathcal{M}}$ induces a global section $g \in \mathcal{S}(X)$ such that $g|_C = s_C$ for all $C \in \mathcal{M}$. Note that this section is unique as $\mathcal{S}$ is a subpresheaf of the sheaf $\mathcal{E}$.

Condition (1) ensures that there is at least one possible event at each context. Condition (3) says that a family of possible events that agree on their common variables gives rise to a possible global assignments of outcomes to each measurement, thus establishing a correspondence between global sections and compatible families. Condition (2) is the least trivial, and can be interpreted as a possibilistic version of no-signalling. Indeed, on the usual bipartite scenario, if Alice chooses measurement $a \in \{a_1, a_2\}$, and we denote $U = \{a\}$, $U' = \{a, b_1\} \in \mathcal{M}$ and $U'' = \{a, b_2\} \in \mathcal{M}$, we know that both $\rho^{U'}_U$ and $\rho^{U''}_U$ are surjective, as $\mathcal{M} \ni U' \supseteq U \subseteq U'' \in \mathcal{M}$. Therefore, a possible event $\{a \mapsto o_A\} \in \mathcal{S}(U)$ arises as a restriction of both a possible event $\{a \mapsto o_A, b_1 \mapsto o_B\} \in \mathcal{S}(U')$ and a possible event $\{a \mapsto o_A, b_2 \mapsto o_B\} \in \mathcal{S}(U'')$. This means that the event $\{a \mapsto o_A\}$ is possible *regardless* of Bob's choice of measurement, which corresponds to the statement of no-signalling.

Another important consequence of the three conditions is that a possibilistic model $\mathcal{S}$ is uniquely determined by its value on the contexts. Indeed, values $\mathcal{S}(U)$ for $U \subseteq C$ for some $C \in \mathcal{M}$ are fixed by Condition 2, and Condition 3 determines the values for U above the cover.

Let us list some examples of possibilistic empirical models. Both the PR-box (Table II.7) and the GHZ model (Table II.6) are usually considered in their possibilistic forms. Another important example of a possibilistic model is the **Hardy model** [**Har92, Har93**], which is defined on a $(2,2,2)$ scenario and presented in Table II.8. This model is realisable in quantum mechanics and has been used to give a proof of non-locality *without inequalities.*

TABLE II.8. The possibilistic Hardy model.

A	B	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
a_1	b_1	1	1	1	1
a_1	b_2	0	1	1	1
a_2	b_1	0	1	1	1
a_2	b_2	1	1	1	0

On more general non-Bell scenarios, a representative class of examples is that of **Kochen–Specker models** [**AB11a, MB13**], of which the models used in the proof of the Kochen–Specker theorem [**KS67**] are special cases.

DEFINITION II.13. Let $\Sigma := \langle X, \mathcal{M}, O = \{0,1\}\rangle$ be a measurement scenario such that its context all have the same cardinality. The **Kochen–Specker (KS) model** on Σ is the possibilistic model $\mathcal{S} : \mathbf{Open}(X)^{op} \to \mathbf{Set}$ defined as follows: for all $C \in \mathcal{M}$,

$$\mathcal{S}(C) := \{s \in \mathcal{E}(C) \mid o(s) = 1\}$$

where, given an event $s \in \mathcal{E}(C) = O^C$,

$$o(s) := |\{m \in C \mid s(m) = 1\}|.$$

That is, given any context C, the possible sections of $\mathcal{S}(C)$ are those that assign the outcome 1 to exactly one measurement.

The simplest non-trivial example of a KS model is the **Specker's triangle** [**Spe60, LSW11**], defined on the non-Bell type scenario identified by $X = \{a,b,c\}$ and $\mathcal{M} = \{\{a,b\},\{b,c\},\{a,c\}\}$, and presented in Table II.9.

4.6. Bundle diagrams. Possibilistic models can be effectively visualised by taking advantage of the simplicial complex description of measurement scenarios introduced in Section 3.2 to construct **bundle diagrams** [**ABK$^+$15, Car17, BO18**]. A bundle diagram is comprised of two elements: a base space, constituted by the simplicial complex representing the scenario, and a fibre which reproduces the possible events of the

TABLE II.9. The Specker's triangle model.

Contexts	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
$\{a,b\}$	0	1	1	0
$\{b,c\}$	0	1	1	0
$\{a,c\}$	0	1	1	0

empirical model. More specifically, given a scenario $\langle X, \mathcal{M}, (O_m)\rangle$ and its simplicial representation $\Sigma = \downarrow_{\subseteq}\mathcal{M}$, we let Σ lie as the 'base' of the diagram. Above each vertex $m \in X$ is a discrete fibre representing the set of its possible outcomes O_m (see e.g. Figure II.2). The sections of an empirical model $\mathcal{S}$ on $\langle X, \mathcal{M}, (O_m)\rangle$ can be visualised as edges (or, more generally, faces) in the fibre. In Figure II.5, the bundle diagrams of the Hardy model (Table II.8), the PR-box (Table II.3) and Specker's triangle (Table II.9) are depicted.

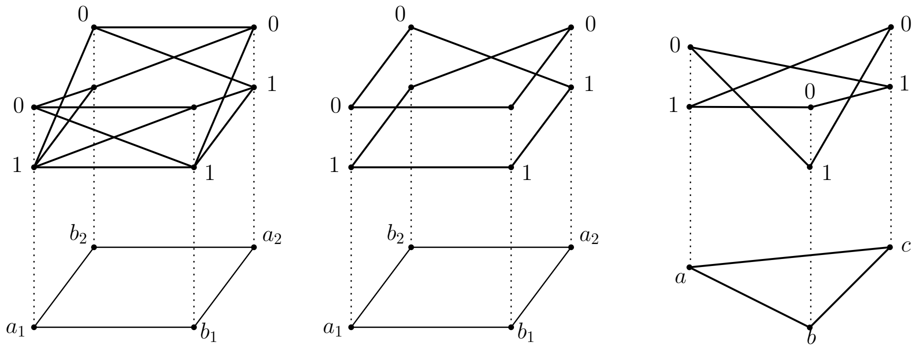


FIGURE II.5. Bundle diagrams of the Hardy, PR-box and Specker's triangle models.

Of course, this graphical representation can be used only when the scenario is sufficiently simple. For instance, the GHZ model of Table II.6 is hard to visualise due to the 3-dimensionality of the base space, although a promising attempt has been made in [**BO18**]. In spite of this, many of the models we will deal with are representable, and bundle diagrams will prove to be a valuable tool to heuristically investigate their contextual properties.

4.7. Modelisation of experimental data. The requirement of no-signalling is of paramount importance for the development of the techniques presented in this thesis. However, real experimental data are often noisy and may contain small traces of signalling in their distributions, due to measurement errors, influence of unwanted external factors, or simply by the finiteness of the sample.[5] It is thus important to clarify how to model noisy experimental data in the strict theoretical framework we adopt in our work.

[5]Note that, while the theory deals with *probability distributions*, actual experiments only provide *relative frequencies*.

Probabilistic empirical models can be alternatively represented as real vectors. Let $d := \sum_{C \in \mathcal{M}} |\mathcal{E}(C)|$. Given an empirical model $e = \{e_C \in \mathcal{D}_{\mathbb{R}_{\geq 0}} \mathcal{E}(C)\}_{C \in \mathcal{M}}$ on a scenario $\Sigma = \langle X, \mathcal{M}, (O_m) \rangle$, one can rewrite it as a d-dimensional real vector $\mathbf{V}_e$ defined as follows: for all $C \in \mathcal{M}$ and $s_C \in \mathcal{E}(C)$,

$$\mathbf{V}_e[\langle C, s_C \rangle] := e_C(s_C).$$

Let $\mathsf{NS}(\Sigma) \subseteq \mathbb{R}^d$ denote the set of no-signalling models over Σ, seen as vectors in $\mathbb{R}^d$. The points of $\mathsf{NS}(\Sigma)$ are those vectors $\mathbf{V}$ of non-negative real numbers satisfying the normalisation equations, that is, for all $C \in \mathcal{M}$,

$$\sum_{s_c \in \mathcal{E}(C)} \mathbf{V}[\langle C, s_C \rangle] = 1,$$

and the compatibility (or no-signalling) conditions: for all $C, C' \in \mathcal{M}$ and $t \in \mathcal{E}(C \cap C')$,

$$\sum_{\substack{s_C \in \mathcal{E}(C): \\ s_C|_{C \cap C'} = t}} \mathbf{V}[\langle C, s_C \rangle] = \sum_{\substack{s_{C'} \in \mathcal{E}(C'): \\ s_{C'}|_{C \cap C'} = t}} \mathbf{V}[\langle C, s_{C'} \rangle],$$

which are all linear. Therefore, $\mathsf{NS}(\Sigma)$, being specified by a set of linear constraints, is a *polytope*, called the **no-signalling polytope** of the scenario Σ [**PBS11, Pop14**].

With this premise, given a signalling empirical model e, seen as a vector $\mathbf{V}_e \subseteq \mathbb{R}^d$ obtained from noisy experimental data, one can choose its nearest element $\mathbf{V}_{\hat{e}}$ on the no-signalling polytope as a suitable modelisation of the experimental results. Then, all the techniques discussed in this thesis can be applied to the no-signalling empirical model $\hat{e}$.

5. Non-locality and contextuality

The phenomenon of non-locality has been presented by Einstein, Podolsky and Rosen [**EPR35**] as a paradoxical aspect of quantum mechanics which, in their view, proved that a quantum state does not constitute a complete description of the state of a system. Specifically, they showed that the postulates of quantum physics allow the measurement of the position and momentum of a pair of entangled particles to violate Heisenberg's uncertainty principle [**Hei27**], unless the very act of measuring one particle instantaneously affects the other. This 'spooky action at a distance' violates the laws of relativity, and was consequently deemed a contradiction. Their proposed solution was to reinterpret the probabilistic aspect of quantum predictions as a fundamentally incomplete knowledge of reality, rather than a faithful ontological representation of the state of a system. According to this viewpoint, measurements have deterministic well-defined outcomes, or **hidden variables**, regardless of whether they are performed or not. Although these values may not be accessible – hence the term *hidden* – they constitute the genuine *elements of reality* of the system. Then, the probability distribution observed for each measurement context in a quantum experiment simply arises as the marginal of a global probability distribution over the hidden variables.

In sheaf theoretic terms, given a measurement scenario $\langle X, \mathcal{M}, (O_m) \rangle$, the canonical set of hidden variables is the one of global sections of the sheaf $\mathcal{E}$, i.e. global assignments of outcomes to the measurements. This choice does not cause any loss of generality, as justified in [**AB11a**] via a generalisation of a famous result by Fine [**Fin82**].

A global section $d \in \mathcal{D}_R\mathcal{E}(X)$ of the presheaf $\mathcal{D}_R\mathcal{E}$ specifies a distribution over the set $\mathcal{E}(X)$ of hidden variables. Thus, the requirement of **non-locality** can be formulated as follows: *given an empirical model $\{e_C \in \mathcal{D}_{\mathbb{R}_{\geq 0}}\mathcal{E}(C)\}_{C\in\mathcal{M}}$ over a Bell-type scenario $\langle X, \mathcal{M}, (O_m)\rangle$, there exists a global distribution $d \in \mathcal{D}_{\mathbb{R}_{\geq 0}}\mathcal{E}(X)$ such that $d|_C = e_C$ for all $C \in \mathcal{M}$.*

Bell's no go theorem [**Bel64**] shows that no physical theory of local hidden variables can reproduce the predictions of quantum mechanics, effectively establishing non-locality as a fundamental feature of reality, rather than an undesired property of quantum mechanics. Kochen–Specker's theorem [**KS67**] extended Bell's result to measurements scenarios involving non-local measurements, introducing the more general concept of *contextuality*. By abstracting from these two key results, we are finally able to introduce the definition of non-locality and contextuality in sheaf theoretic terms:

Definition II.14. Let $e = \{e_C \in \mathcal{D}_R\mathcal{E}(C)\}_{C\in\mathcal{M}}$ be an empirical model over a scenario $\langle X, \mathcal{M}, (O_m)\rangle$. We say that e is **contextual** if there is no global section $d \in \mathcal{D}_R\mathcal{E}(S)$ such that $d|_C = e_C$ for all $C \in \mathcal{M}$. We say that it is **non-local** if, in addition, the scenario $\langle X, \mathcal{M}, (O_m)\rangle$ is Bell-type.

If $R = \mathbb{B}$, we say that $\{e_C\}_{C\in\mathcal{M}}$ is **possibilistically** or **logically contextual**.[6,7]

Notice how non-locality arises as a special case of contextuality. This unified view of the two phenomena is one of the key advantages of the sheaf theoretic approach.

5.1. A hierarchy of contextuality. So far, the only concrete example of a contextuality argument we have encountered is the one based on logical Bell inequalities of Section 5, used to prove the non-locality of Bell's model. Bell's model is probabilistic, and this aspect plays a crucial role in the argument, as it allows to associate probabilities to the propositional formulae involved in the proof. However, there are some cases where contextuality can be witnessed even at the level of possibilities, with purely logical arguments. This type of contextuality arguments was developed by Heywood & Redhead [**HR83**], Greenberger, Horne, Shimony, & Zeilinger [**GHZ89, GHSZ90**], whose proof was subsequently simplified by Mermin [**Mer90a, Mer90b**], and Hardy [**Har92, Har93**].

Since these proofs rely solely on the possibilistic structure of empirical models, the kind of contextual behaviour observed appears to be somewhat stronger than the one featured by Bell's model. The high-level description provided by sheaf theory allows to make this claim rigorous, and establish a clear hierarchy of different strengths of contextuality.

In Section 4.5 we showed how probabilistic models give rise to possibilistic ones, and how such models can be defined as subpresheaves of $\mathcal{E}$. In Definition II.14, we introduced a notion of contextuality both for probabilistic and possibilistic models. A natural question to ask is *what is the relation between the probabilistic contextuality of a model $\{e_C\}_{C\in\mathcal{M}}$, and the possibilistic contextuality of its collapse $\{\tilde{e}_C\}_{C\in\mathcal{M}}$?*

[6] The notion of *possibilistic* (or *logical*) *contextuality* will be equivalently reformulated in different terms in Definition II.15, which is the one we will adopt for the rest of the thesis.

[7] When $R = \mathbb{R}_{\geq 0}$, we will sometimes call the model *probabilistically contextual*, to emphasise the difference with possibilistic contextuality.

Via the isomorphism $\mathbb{B}^{\mathcal{E}(X)} \cong \mathcal{P}(\mathcal{E}(X))$, a possibilistic global distribution $d \in \mathcal{D}_{\mathbb{B}}\mathcal{E}(X)$ consistent with the collapse $\{\tilde{e}_C\}_{C\in\mathcal{M}}$ of a probabilistic model $\{e_C\}_{C\in\mathcal{M}}$ can be identified with a set of global assignments that exactly restricts to the set of possible local assignments at each context.

The existence of such a global section is clearly a weaker requirement than the existence of a *probabilistic* global section for $\{e_C\}_{C\in\mathcal{M}}$, which in addition must marginalise to the individual local probabilities. Thus, one can clearly see that possibilistic contextuality implies probabilistic contextuality.

Possibilistic contextuality can also be rephrased in terms of subpresheaves of $\mathcal{E}$ by formalising the discussion above on the role of a possibilistic global distribution. Let $S_e : \mathcal{P}(X)^{op} \to \mathbf{Set}$ be the possibilistic model obtained by collapsing a probabilistic model $e = \{e_C\}_{C\in\mathcal{M}}$ as in (II.4). Then, the set $S_e(X)$ of global sections contains all the global assignments in $\mathcal{E}(C)$ that are consistent with e, i.e. such that their restriction to every context C is in the support of e_C. Possibilistic contextuality arises when $S_e(X)$ is not large enough to account for *all* of the local events that e deems possible. In other words, there exists (at least one) local section $s \in S_e(C)$ which does not extend to any global section in $S_e(X)$, that is, for all $g \in S_e(X)$, $g|_C \neq s$. This can be interpreted as the fact that the locally observed event s cannot be explained by a classical hidden variable.

In extreme cases, $S_e(X)$ could be empty. This means that *none* of the locally observed events can be explained classically. We refer to this phenomenon as *strong contextuality.*

By abstracting the discussion above, we introduce the general definition of logical and strong contextuality.

Definition II.15. Let $\mathcal{S} : \mathcal{P}(X)^{op} \to \mathbf{Set}$ be a possibilistic empirical model over a scenario $\langle X, \mathcal{M}, (O_m)\rangle$. We say that $\mathcal{S}$ is

- **Logically** or **possibilistically contextual**, denoted $\mathsf{LC}(\mathcal{S})$, if there exists a local section $s \in \mathcal{S}(C)$, with $C \in \mathcal{M}$, such that s is not the restriction of any global section in $\mathcal{S}(X)$. In this case, we say that $\mathcal{S}$ is *logically* or *possibilistically contextual at* s, and write $\mathsf{LC}(\mathcal{S}, s)$.
- **Strongly contextual**, denoted $\mathsf{SC}(\mathcal{S})$, if $\mathcal{S}(X) = \emptyset$, or, equivalently, if $\mathsf{LC}(\mathcal{S}, s)$ for all local sections $s \in \mathcal{S}(C)$ for all contexts $C \in \mathcal{M}$.

From these definitions and the discussion above, it is clear that we have a *hierarchy* of levels of contextuality:

Strong contextuality $\Rightarrow$ Logical contextuality $\Rightarrow$ (probabilistic) Contextuality

In the following list of examples is contained the proof of the strictness of these implications:

Example II.16.

- Bell's model (Table II.2) is probabilistically contextual, as shown in Section 2.3, but it is *not* logically contextual. Indeed, its possibilistic collapse arises as a restriction of the global distribution $d \in \mathcal{D}_{\mathbb{B}}\mathcal{E}(X)$ defined as follows: for all $g \in \mathcal{E}(X)$,

$$d(g) = 1 \iff g(a_1) + g(b_1) = 0 \mod 2.$$

- The Hardy model is logically contextual. Indeed, the local section
$$s = \{a_1 \mapsto 0, b_1 \mapsto 0\} \in \mathcal{E}(\{a_1, b_1\})$$
cannot be extended to a global one. Indeed, a hypothetical global section $g \in \mathcal{E}(X)$ consistent with the model and such that $g|_C = s_C$ would satisfy $g(a_1) = 0$, and $g(b_1) = 0$. By simply looking at the possibility table of the model (Table II.8), we have
$$g(b_1) = 0 \;\Rightarrow\; g(a_2) = 1 \;\Rightarrow\; g(b_2) = 0 \;\Rightarrow\; g(a_1) = 1, \tag{II.5}$$
which is a contradiction. However, the model is *not* strongly contextual, as it contains the global section
$$\{a_1 \mapsto 1, b_1 \mapsto 1, a_2 \mapsto 0, b_2 \mapsto 0\} \in \mathcal{E}(X).$$
- The GHZ model is strongly contextual. To see this, one can observe that the

TABLE II.10. Four rows of the GHZ model.

A	B	C	000	001	010	011	100	101	110	111
a_1	b_1	c_1	1	0	0	1	0	1	1	0
a_1	b_2	c_2	0	1	1	0	1	0	0	1
a_2	b_1	c_2	0	1	1	0	1	0	0	1
a_2	b_2	c_1	0	1	1	0	1	0	0	1

support of the model, displayed in Table II.10, is characterised by the following equations in $\mathbb{Z}_2$:
$$a_1 \oplus b_1 \oplus c_1 = 0, \qquad a_2 \oplus b_1 \oplus c_2 = 1,$$
$$a_1 \oplus b_2 \oplus c_2 = 1, \qquad a_2 \oplus b_2 \oplus c_1 = 1.$$
It is sufficient to sum all these equations to obtain $0 = 1$. This contradiction shows that the model does not admit any global section, and is therefore strongly contextual.

The very last proof of the strong contextuality of the GHZ state belongs to the general class of *All-vs-Nothing arguments* to which the next section is devoted.

5.2. All-vs-Nothing arguments. Among the various proofs of contextual behaviour in quantum mechanics, a class of arguments dubbed *All-vs-Nothing* stands out as being particularly abundant in the literature. All-vs-Nothing (AvN) arguments are proofs of strong contextuality which rest on the observation that the possible local assignments of an empirical model satisfy a system of parity equations that admit no global solution.

The first instance of an AvN argument is due to Mermin, who coined the term to describe his proof of the strong contextuality of the GHZ state [**Mer90a, Mer90b**], which was presented in Example II.16. Since then, AvN arguments have been extensively used to produce other examples of strongly contextual models in quantum physics [**Wae14**], especially in stabiliser quantum mechanics [**Got97**].

In [**ABK$^+$15**] Abramsky et al. proposed a generalisation of this class of proofs, which takes into account systems of linear equations for any ring R, greatly enhancing the scope of their applicability both within and beyond quantum physics. Let us briefly review this approach.

Let R be a ring, and consider a measurement scenario $\langle X, \mathcal{M}, R\rangle$, where each measurement produces an outcome in R

Definition II.17. An R**-linear equation** is a triple $\varphi = \langle C, a, b\rangle$, where $C \in \mathcal{M}$, $a : C \to R$, and $b \in R$. We denote $C = V_\varphi$. An event $s \in \mathcal{E}(C)$ **satisfies** φ if

$$\sum_{m\in C} a(m)s(m) = b.$$

This lifts to the level of systems of R-linear equations and sets of assignments: given a system of equations Γ, let

$$\mathbb{M}(\Gamma) := \{s \in \mathcal{E}(C) \mid s \models \varphi,\ \forall\varphi \in \Gamma\}$$

denote the set of events in $\mathcal{E}(C)$ that satisfy every equation φ in Γ. Similarly, given a set of events $S \subseteq \mathcal{E}(C)$, let

$$\mathbb{T}_R(S) := \{\varphi \mid s \models \varphi,\ \forall s \in S\}$$

be the set of equations satisfied by all events in S. With this premise, given an empirical model $\mathcal{S}$ on $\langle X, \mathcal{M}, R\rangle$, we may associate its R-linear theory to it, which readily leads to the definition of an AvN argument:

Definition II.18. The R**-linear theory** of a model $\mathcal{S}$ is

$$\mathbb{T}_R(\mathcal{S}) := \bigcup_{C\in\mathcal{M}} \mathbb{T}_R(\mathcal{S}(C)) = \{\varphi \mid s \models \varphi,\ \forall s \in \mathcal{S}(V_\varphi)\}.$$

We say that $\mathcal{S}$ is **AvN**, written $\mathsf{AvN}_R(\mathcal{S})$, if $\mathbb{T}_R(\mathcal{S})$ is inconsistent. That is, if there is no global assignment $g : X \to R$ such that $g|_{V_\varphi} \models \varphi$ for all $\varphi \in \mathbb{T}_R(\mathcal{S})$.

The inconsistency of the associated system of equations is in fact a proof of strong contextuality for the model in question.

Proposition II.19 (Proposition 7 of [**ABK⁺15**]). *An* AvN_R *model is strongly contextual.*

Proof. Suppose $\mathcal{S}$ is not strongly contextual. Then there exists $g \in \mathcal{E}(R)$ such that $g|_C \in \mathcal{S}(C)$ for all $C \in \mathcal{M}$. It follows that, given any $\varphi \in \mathbb{T}_R(\mathcal{S})$, $g|_{V_\varphi} \in \mathcal{S}(V_\varphi)$, which implies $g|_{V_\varphi} \models \varphi$. Thus $\mathbb{T}_R$ is consistent. □

This result gives rise to the notion of **AvN contextuality**, which is strictly stronger than strong contextuality [**ABK⁺15**].

5.3. Contextuality in bundle diagrams. Possibilistic forms of contextuality, i.e. logical and strong, can often be graphically visualised in the bundle representation of empirical models. For instance, the simple argument used in Example II.16 to show that the Hardy model is contextual can be reproduced geometrically, as shown in Figure II.6. Section $s = (a_1, b_1) \mapsto (0, 0)$ is marked in red. A global section corresponds to a closed loop around the bundle. In the central diagram of Figure II.6, we display an attempt to extend the s to a closed loop, which follows the list of implications used

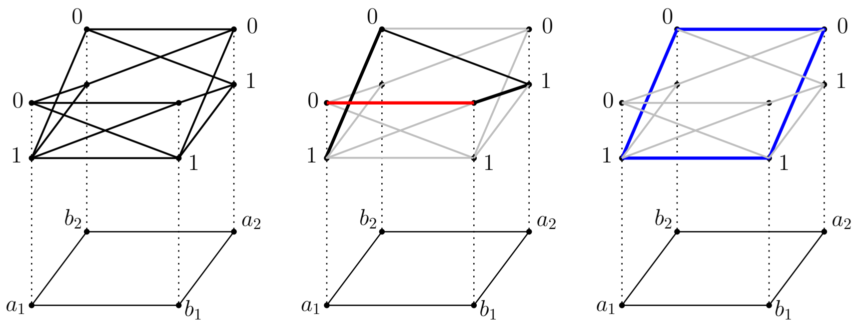


FIGURE II.6. A topological visualisation of the contextual properties of the Hardy model

in (II.5). One can clearly see that s cannot be extended to such a closed loop, thus we conclude that the model is logically contextual at s.

The diagram on the right of Figure II.6 highlights in blue a global section consistent with the model. This shows that the Hardy model is *not* strongly contextual. It is sufficient to glance at the bundle diagram of the PR-box and the Specker's triangle displayed in Figure II.5 to see that these model do not contain any global section, and therefore are strongly contextual.

Although the insight on contextuality provided by these bundle-like representations of empirical models might simply look like a nice visualisation of the phenomenon with little potential for general results, it does tell us something of pivotal importance for this thesis, that is, it shows rather neatly that contextuality can be perceived as a purely topological property. The fact that contextuality finds such a compelling description in sheaf theory, which is ultimately a topological theory, can also be considered as evidence for this. The simple idea of *extending a local section to a closed loop* will play a key role in developing a complex topological apparatus suited to model and study contextuality.

5.4. Vorob'ev's theorem. A natural question to ask concerning contextuality is whether it is possible to characterise those measurement scenarios $\langle X, \mathcal{M}, (O_m)\rangle$ on which one can define contextual empirical models. A classical result due to Vorob'ev [**Vor62**], and rewritten more generally in [**Bar15a**] to fit the sheaf–theoretic framework, deals precisely with this problem. Vorob'ev's theorem states that it is impossible to witness contextuality on a scenario $\langle X, \mathcal{M}, (O_m)\rangle$ if and only if its simplicial complex description Σ is *acyclic* in the database–theoretic sense of [**BFM$^+$81, BFMY83, Fag83, FMU82**]:[8]

DEFINITION II.20. Let $\langle X, \mathcal{M}, (O_m)\rangle$ be a measurement scenario, and let $\Sigma = \downarrow_{\subseteq}\mathcal{M}$ be its simplicial representation. For each context $C \in \mathcal{M}$, i.e. each facet σ_C of Σ, we denote by π_C the set of vertices of Σ which belong to σ_C and not to any other facet.

$$\pi_C := \{x \in V(\Sigma) \mid (x \in \tau \Rightarrow \tau \subseteq \sigma_C), \forall \tau \in \Sigma\}$$

[8]Of course, this notion did not exist at the time of the formulation of Vorob'ev's theorem, who characterised the property rather imprecisely. The later introduction of the concept of acyclic databases allowed to rephrase the theorem in the form we present here.

If $\pi_C \neq \emptyset$ for some $C \in \mathcal{M}$, we say that there is a **Graham-reduction** step from Σ to the subcomplex

$$\Sigma' := \Sigma|_{V(\Sigma)\setminus\pi_C} = \{\sigma \in \Sigma \mid \sigma \cap \pi_C = \emptyset\} = \{\sigma \setminus \pi_C \mid \sigma \in \Sigma\}$$

comprised of all the vertices except the ones in π_C. In this case, the Graham-reduction from Σ to Σ' is denoted by $\Sigma \rightsquigarrow \Sigma'$. The scenario is said to be **acyclic** if there is a sequence of Graham-reduction steps

$$\Sigma =: \Sigma_0 \rightsquigarrow \Sigma_1 \rightsquigarrow \cdots \rightsquigarrow \Sigma_n = \{\emptyset\}.$$

In Figure II.7 we illustrate an example of Grahm reduction in the case of both a cyclic and acyclic cover. In red it is highlighted the vertex remove at each step.

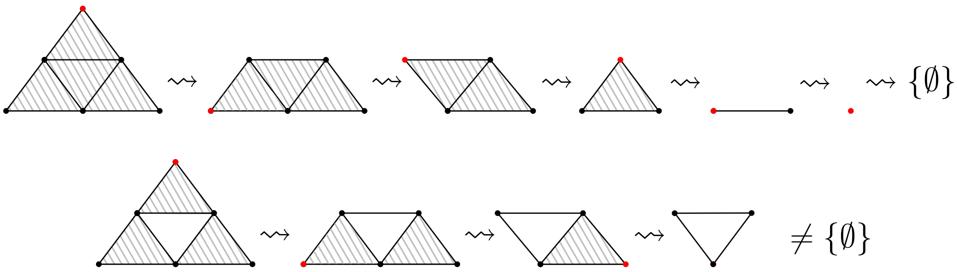


FIGURE II.7. Example of a cyclic and acyclic scenario.

In the formulation of [**Bar15a**], Vorob'ev's theorem is stated as follows:

THEOREM II.21 (Vorob'ev's theorem). *Let $\langle X, \mathcal{M}, (O_m)\rangle$ be a scenario. Any probabilistic or possibilistic empirical model defined on $\langle X, \mathcal{M}, (O_m)\rangle$ is non-contextual if and only if $\Sigma = \downarrow_{\subseteq}\mathcal{M}$ is acyclic.*

This key result will be further explored and refined in Chapter V.

6. The cohomology of non-locality and contextuality

The problem of extending local sections to global ones is well-studied in sheaf theory. In fact, it is safe to say that this very question is the main motivation underpinning the development of the theory and, more specifically, of sheaf cohomology.

Sheaf cohomology was originally introduced by Leray [**Ler45**] and later clarified by the work of Koszul [**Kos47a, Kos47b, Kos51**], Cartan [**Car49**], Borel [**Bor51**] and Serre [**Ser55, Ser56, Ser57**]. It found striking applications to classic problems in algebraic topology, such as Weil's proof of De Rham's equations [**Wei52**], theorems A and B for Stein manifolds [**Ste40**], proved by Cartan & Serre [**CS48**] and utilised to solve the two Cousin problems [**Cou95**], and theorems A and B for coherent sheaves [**Ser55**]. All of these longstanding problems share a similar trait: they concern the extendability of local features to global ones. This is exactly the kind of problem with which we are dealing when studying contextuality, and it is therefore natural to use sheaf cohomology to study this phenomenon.

A sheaf-cohomological framework suited to study logical forms of contextuality was developed in [**AMB12**] and extended in [**ABK$^+$15, Car15, Car17, Car18**]. In this thesis, we will adopt and refine this viewpoint. Although we will not go into details, we shall mention that the cohomological approach to contextuality has been studied using other kinds of cohomology, such as cyclic and order cohomology of effect algebras [**Rou17**], simplicial cohomology [**ORBR17**], and group cohomology [**Rau16, ORBR17, OTR18**], and now represents a well–established line of research.

6.1. Sheaf cohomology. General sheaf cohomology deals with presheaves of abelian groups, and thus it is not immediately obvious how to apply it to study possibilistic empirical models, which are merely presheaves of sets. This difficulty is overcome by considering an **AbGrp**-valued presheaf which *represents* the model in a suitable way:

Definition II.22. Let $\mathcal{S} : \mathcal{P}(X)^{op} \to \mathbf{Set}$ be a possibilistic empirical model over a scenario $\langle X, \mathcal{M}, (O_m) \rangle$. An **AbGrp**-valued presheaf $\mathcal{F} : \mathcal{P}(X)^{op} \to \mathbf{AbGrp}$ is said to be an **abelian representation** of $\mathcal{S}$ if it is such that

(1) It satisfies conditions (1), (2) and (3) of Definition II.12.
(2) There exists an injection $i : \mathcal{S} \hookrightarrow \mathcal{F}$ with $i_C(s) \neq 0 \in \mathcal{F}(C)$ for all $C \in \mathcal{M}$ and for each $s \in \mathcal{S}(C)$.

An abelian representation of $\mathcal{S}$ simply embeds the local sections into an abelian group. Condition (2) of the definition above ensures that none of the local sections of $\mathcal{S}$ is mapped to the null element of the group, which is nothing but an artificial addition to each set of local sections with no physical interpretation.

In practice, given a presheaf $\mathcal{S}$, we use $\mathcal{F} := F_R\mathcal{S}$ as its representation, where R is a ring,[9] and $F_R : \mathbf{Set} \to \mathbf{AbGrp}$ is the free functor on R, which maps a set X to the group of *formal R-linear combinations* of its elements:

$$F_R(X) := \left\{ \sum_{i \in I} \lambda_i \cdot x_i \;\middle|\; |I| < \infty, \lambda_i \in R, x_i \in X \ \forall i \in I \right\}.$$

The restriction maps of $\mathcal{F} = F_R\mathcal{S}$ are obtained by linearly extending the ones of $\mathcal{S}$. In a slight abuse of notation, we will denote by $\rho^{U'}_U$ both the restriction maps of $\mathcal{S}$ and those of $\mathcal{F}$. The injection $\mathcal{S} \hookrightarrow F_R\mathcal{S}$ is given by the following trivial collection of maps: for all $U \subseteq X$,

$$i_U : \mathcal{S}(U) \hookrightarrow F_R(U) :: s \mapsto 1 \cdot s.$$

Although this might seem a minor alteration, this *abelian approximation* of empirical models will play a crucial role in the existence of *false negatives*, an issue that will be thoroughly analysed in Chapters III and IV.

The apparatus of sheaf cohomology is abstract and complex, but for our purposes it is sufficient to consider its simplest form, namely **Čech cohomology**. This is due to the fact that sheaf cohomology coincides with Čech cohomology whenever the presheaf in question is defined on a paracompact space [**God58**]. The presheaves describing empirical models are defined on a discrete space, which, in particular, is paracompact. So we can limit ourselves to Čech cohomology without any loss of generality.

[9] Usually, $R = \mathbb{Z}$ or $\mathbb{Z}_2$.

6.1.1. Čech cohomology.

Definition II.23. Let $\mathcal{M} \subseteq \mathbf{Open}(X)$ be a collection of open subsets of a space X. A **q-simplex of the nerve** of $\mathcal{M}$ is a tuple $\sigma = \langle C_0, \ldots, C_q \rangle$ of elements of $\mathcal{M}$ with non-empty intersection. The set of q-simplices of $\mathcal{M}$ is denoted

$$\mathcal{N}(\mathcal{M})^q := \left\{ \sigma = \langle C_0, \ldots, C_q \rangle \in \mathcal{M}^{q+1} \mid |\sigma| \neq \emptyset \right\},$$

where

$$|\sigma| := \bigcap_{i=0}^{q} C_i.$$

The collection of all simplices $\mathcal{N}(\mathcal{M}) := \{\emptyset\} \cup \bigcup_{q \geq 0} \mathcal{N}(\mathcal{M})^q$ is essentially an abstract simplicial complex with the added structure of an order for its faces.

Although this definition is completely general, in this thesis $\mathcal{M}$ will always be the measurement cover of a scenario with measurement set X.

For all $q \geq 0$ and each $0 \leq j \leq q+1$, we define the boundary maps $\partial_j : \mathcal{N}(\mathcal{M})^{q+1} \to \mathcal{N}(\mathcal{M})^q$ by

$$\partial_j(C_0, \ldots, C_{q+1}) := (C_0, \ldots, C_{j-1}, \hat{C}_j, C_{j+1}, \ldots, C_{q+1}),$$

where $\hat{C}_j$ is to denote that element C_j has been removed from the list.

Definition II.24. Let $\mathcal{F} : \mathbf{Open}(X)^{op} \to \mathbf{AbGrp}$ be a presheaf.[10] The **(augmented) Čech cochain complex** of $\mathcal{F}$ is defined as the sequence

$$0 \xrightarrow{\delta^{-1} := 0} C^0(\mathcal{M}, \mathcal{F}) \xrightarrow{\delta^0} C^1(\mathcal{M}, \mathcal{F}) \xrightarrow{\delta^1} \ldots,$$

where

- For each $q \geq 0$,
$$C^q(\mathcal{M}, \mathcal{F}) := \bigoplus_{\sigma \in \mathcal{N}(\mathcal{M})^q} \mathcal{F}(|\sigma|)$$
is the abelian group of **q-cochains**.
- For each $q \geq 0$, the **q-th coboundary map** $\delta^q : C^q(\mathcal{M}, \mathcal{F}) \to C^{q+1}(\mathcal{M}, \mathcal{F})$ is defined by
$$\delta^q(\omega)(\sigma) := \sum_{j=0}^{q+1} (-1)^j \rho^{|\partial_j(\sigma)|}_{|\sigma|}(\omega(\partial_j \sigma)),$$
where we have used the fact that $|\sigma| \subseteq |\partial_j(\sigma)|$. By convention, $\delta^{-1} := 0$.

A straightforward calculation yields the following proposition.

Proposition II.25 ([**AMB12**]). *For all $q \geq -1$, $\delta^{q+1} \circ \delta^q = 0$.*

Thus the object of the definition is indeed a cochain complex. **Čech cohomology** $\check{H}^*(\mathcal{M}, \mathcal{F})$ is defined as the cohomology of this augmented cochain complex:

Definition II.26. Given a presheaf $\mathcal{F} : \mathbf{Open}(X)^{op} \to \mathbf{AbGrp}$, we define, for all $q \geq 0$,

[10] In our case, $\mathcal{F}$ will always be an abelian representation of a possibilistic empirical model $\mathcal{S}$.

- The group of q–**cocycles** as

$$Z^q(\mathcal{M}, \mathcal{F}) := \ker(\delta^q).$$

- The group of q–**coboundaries** as

$$B^q(\mathcal{M}, \mathcal{F}) := \operatorname{im}(\delta^{q-1}).$$

- The q–**th Čech cohomology group** of $\mathcal{F}$ as the quotient

$$\check{H}^q(\mathcal{M}, \mathcal{F}) := Z^q(\mathcal{M}, \mathcal{F})/B^q(\mathcal{M}, \mathcal{F}).$$

In our study, we will always assume that $\mathcal{M}$ is **connected**, which means that given any $C, C' \in \mathcal{M}$, there exists a sequence $C = C_0, \ldots, C_n = C'$ such that $C_i \cap C_{i+1} \neq \emptyset$ for all $0 \leq i \leq n-1$. Note that this assumption does not cause any loss of generality as it is always possible to study contextuality in the individual connected components.

Proposition II.27 ([**AMB12**]). *There is a one–to–one correspondence between cocycles in $Z^0(\mathcal{M}, \mathcal{F}) \cong \check{H}^0(\mathcal{M}, \mathcal{F})$ and compatible families $\{r_C \in \mathcal{F}(C)\}_{C \in \mathcal{M}}$*

6.1.2. Relative cohomology. We shall be concerned with extendability of local sections at a fixed context $C_0 \in \mathcal{M}$. For this reason, we define the **relative cohomology** of $\mathcal{F}$. To do so, we introduce two auxiliary presheaves. Firstly

$$\mathcal{F}|_{C_0} : \mathbf{Open}(X)^{op} \to \mathbf{AbGrp} :: U \mapsto \mathcal{F}(U \cap C_0).$$

The restriction to C_0 yields a morphism of sheaves $p^{C_0} : \mathcal{F} \Rightarrow \mathcal{F}|_{C_0}$ given by

$$p^{C_0}_U : \mathcal{F}(U) \to \mathcal{F}|_{C_0}(U) :: r \mapsto r|_{C_0 \cap U}.$$

Each $p^{C_0}_U$ is surjective as $\mathcal{F}$ is flasque beneath the cover and $U \cap C_0 \subseteq C_0 \in \mathcal{M}$. The second presheaf is defined by $\mathcal{F}_{\tilde{C}_0}(U) := \ker(p^{C_0}_U)$. To summarise, we have the following exact sequence of presheaves

$$\mathbf{0} \Longrightarrow \mathcal{F}_{\tilde{C}_0} \Longrightarrow \mathcal{F} \overset{p^{C_0}}{\Longrightarrow} \mathcal{F}|_{C_0}, \tag{II.6}$$

which can be lifted to cochains

$$0 \longrightarrow C^0(\mathcal{M}, \mathcal{F}_{\tilde{C}_0}) \hookrightarrow C^0(\mathcal{M}, \mathcal{F}) \xrightarrow{\bigoplus_C p^{C_0}_C} C^0(\mathcal{M}, \mathcal{F}|_{C_0}) \longrightarrow 0,$$

where exactness on the right follows by surjectivity of all the $p^{C_0}_C$.

6.2. Cohomology obstructions. The map δ^0 can be correstricted to a map $\tilde{\delta^0} := \delta^0 \mid^{Z^1(\mathcal{M}, \mathcal{F})}$ whose kernel is $Z^0(\mathcal{M}, \mathcal{F}) \cong \check{H}^0(\mathcal{M}, \mathcal{F})$ and whose cokernel is isomorphic to $\check{H}^1(\mathcal{M}, \mathcal{F})$, and the same procedure can be applied to $\mathcal{F}|_{C_0}$ and $\mathcal{F}_{\tilde{C}_0}$. Therefore, by applying the snake lemma, we obtain the following diagram:

$$\begin{tikzcd}
 & \check{H}^0(\mathcal{M},\mathcal{F}_{\tilde{C}_0}) \arrow[r] \arrow[d] & \check{H}^0(\mathcal{M},\mathcal{F}) \arrow[r] \arrow[d] & \check{H}^0(\mathcal{M},\mathcal{F}|_{C_0}) \arrow[d] \arrow[dddll, rounded corners, "\gamma_{C_0}"] & \\
0 \arrow[r] & C^0(\mathcal{M},\mathcal{F}_{\tilde{C}_0}) \arrow[r] \arrow[d] & C^0(\mathcal{M},\mathcal{F}) \arrow[r] \arrow[d] & C^0(\mathcal{M},\mathcal{F}|_{C_0}) \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & Z^1(\mathcal{M},\mathcal{F}_{\tilde{C}_0}) \arrow[r] \arrow[d] & Z^1(\mathcal{M},\mathcal{F}) \arrow[r] \arrow[d] & Z^1(\mathcal{M},\mathcal{F}|_{C_0}) \arrow[d] & \\
 & \check{H}^1(\mathcal{M},\mathcal{F}_{\tilde{C}_0}) \arrow[r] & \check{H}^1(\mathcal{M},\mathcal{F}) \arrow[r] & \check{H}^1(\mathcal{M},\mathcal{F}|_{C_0}) &
\end{tikzcd}$$

The homomorphism γ_{C_0} is called the **connecting homomorphism** relative to the context C_0. The following elementary result is introduced without proof in [**AMB12**]. We will give here a short proof for the purpose of introducing the isomorphism ϕ^0, which will be generalised in Chapter III.

Lemma II.28. *Given a context $C_0 \in \mathcal{M}$, we have*

$$\check{H}^0(\mathcal{M},\mathcal{F}|_{C_0}) \cong \mathcal{F}(C_0).$$

Proof. By Proposition II.27, elements of $Z^0(\mathcal{M},\mathcal{F}|_{C_0}) \cong \check{H}^0(\mathcal{M},\mathcal{F}|_{C_0})$ are compatible families of $\mathcal{F}|_{C_0}$. Thus by condition (3) of Definition II.12, for each $s = \langle s_C \rangle_{C\in\mathcal{M}} \in Z^0(\mathcal{M},\mathcal{F}|_{C_0})$ there exists a unique global section, which we denote by $\phi^0(s) \in \mathcal{F}|_{C_0}(X) = \mathcal{F}(C_0)$, that restricts to each element s_C. This defines an assignment

$$\phi^0 : \check{H}^0(\mathcal{M},\mathcal{F}|_{C_0}) \longrightarrow \mathcal{F}(C_0),$$

whose inverse is

$$\psi^0 : \mathcal{F}(C_0) \longrightarrow \check{H}^0(\mathcal{M},\mathcal{F}|_{C_0}) :: s_0 \mapsto \langle s_0|_{C\cap C_0} \rangle_{C\in\mathcal{M}} \,. \tag{II.7}$$

One can easily verify that these assignments are group homomorphisms. $\square$

Thanks to this lemma, we can introduce the following definition:

Definition II.29. Let C_0 be a context of a scenario $\langle X, \mathcal{M}, (O_m) \rangle$, and let $\mathcal{F}$ be an abelian representation of a model $\mathcal{S}$ on the scenario. For any local section $r_0 \in \mathcal{F}(C_0)$, the element

$$\gamma_{C_0}(r_0) \in \check{H}^1(\mathcal{M},\mathcal{F}_{\tilde{C}_0})$$

is called the **cohomology obstruction** of r_0. With a slight abuse of terminology, given a local section $s_0 \in \mathcal{S}(C_0)$, we will call its **cohomology obstruction** the element

$$\gamma_{C_0}(i_{C_0}(s_0)) \in \check{H}^1(\mathcal{M},\mathcal{F}_{\tilde{C}_0}),$$

where $i : \mathcal{S} \hookrightarrow \mathcal{F}$ is the injection of Definition II.22.

The reason why it is called an *obstruction* is clarified by the following proposition:

Proposition II.30 ([**AMB12**]). *Let $\mathcal{F}$ be a an abelian representation of a model $\mathcal{S}$ on a scenario $\langle X, \mathcal{M}, (O_m)\rangle$, let $C_0 \in \mathcal{M}$ be a context and $r_0 \in \mathcal{F}(C_0)$. Then, there exists a compatible family $\{r_C \in \mathcal{F}(C)\}_{C\in\mathcal{M}}$ such that $r_{C_0} = r_0$ if and only if the obstruction of r_0* vanishes*, i.e. $\gamma_{C_0}(r_0) = 0$.*

This results motivates the following definition, which is the cohomological counterpart of Definition II.15.

Definition II.31. Let $\mathcal{S}$ be an empirical model on a scenario $\langle X, \mathcal{M}, (O_m)\rangle$, and consider a local section $s_0 \in \mathcal{S}(C_0)$.

- $\mathcal{S}$ is **cohomologically logically contextual at** s_0, denoted $\mathsf{CLC}\,(\mathcal{S}, s_0)$, if $\gamma_{C_0}(s_0) \neq 0$. We say that $\mathcal{S}$ is **cohomologically logically contextual**, denoted $\mathsf{CLC}(\mathcal{S})$, if $\mathsf{CLC}\,(\mathcal{S}, s)$ for some section s.
- $\mathcal{S}$ is **cohomologically strongly contextual** , denoted $\mathsf{CSC}\,(\mathcal{S})$, if $\mathsf{CLC}\,(\mathcal{S}, s)$ for all sections s.

The main result of [**AMB12**] provides a sufficient condition for an empirical model to be contextual:

Theorem II.32. *Let $\mathcal{S}$ be an empirical model on a scenario $\langle X, \mathcal{M}, (O_m)\rangle$. Given a section s_0 of $\mathcal{S}$, we have*

$$\mathsf{CLC}(\mathcal{S}, s_0) \Rightarrow \mathsf{LC}(\mathcal{S}, s_0),$$
$$\mathsf{CSC}\,(\mathcal{S}) \Rightarrow \mathsf{SC}\,(\mathcal{S})\,.$$

Note that cohomology only provides a sufficient condition for contextuality, which, as we shall see in detail in Chapter III, is not necessary in general. Although Proposition II.30 says that cohomology gives rise to a *complete invariant* for the extendability of local sections to global sections of an abelian presheaf $\mathcal{F}$, this does not generalise to sections of a presheaf of sets $\mathcal{S}$. This is due to the fact that global sections of $\mathcal{F}$ are not necessarily global sections of $\mathcal{S}$. This aspect gives rise to *false negatives*, an issue which will be extensively studied in Chapters III and IV.

Despite this, cohomology has been proved to correctly detect contextuality in a variety of models, including the GHZ model, the PR Boxes, the Peres–Mermin magic square [**Per90, Mer90b, Mer93**], and all $\neg\mathsf{GCD}$ models [**AB11a**]. All these instances can be shown to be part of the vast class of models admitting All-vs-Nothing arguments of contextuality. In fact, in [**ABK$^+$15**], it has been shown that cohomology does correctly captures this very general kind of contextual behaviour. This constitutes the main motivation underpinning our study, in the context of this thesis, of All-vs-Nothing arguments in quantum mechanics, presented in Chapter VI.

CHAPTER III

The cohomology of contextuality: extensions and limitations

Summary

This chapter illustrates new insights into different aspects of the application of sheaf cohomology to the study of contextuality. Many of the results presented here are limitative in character, and highlight important shortcomings of the theory. In particular, we analyse the issue of false negatives, and show that, in its present formulation, sheaf cohomology does not constitute a complete invariant for strong contextuality, not even under symmetry and connectedness restrictions on the measurement cover, disproving a previous conjecture. We extend the theory by generalising cohomology obstructions to higher cohomology groups. Such higher obstructions give rise to a refinement of the notion of cohomological contextuality: different 'levels' of contextuality are organised in a hierarchy of logical implications. Finally, we present an alternative description of the first cohomology group in terms of $\mathcal{F}$-torsors, resulting in a new interpretation of the obstructions.

1. Overview

The pioneering work on the application of sheaf cohomology to contextuality by Abramsky, Barbosa & Mansfield [**AMB12**] – reviewed in Chapter II – presented a sufficient condition for contextuality based on the concept of *cohomology obstruction*. Although this method has been proved to correctly detect non-classical behaviour in a variety of scenarios, the authors pointed out that it does not constitute a complete invariant for contextuality, as witnessed by the existence of *false negatives*. These findings motivate further research on the actual power of detection of cohomological obstructions.

In this chapter, we illustrate new insights into the properties of sheaf cohomology with the ultimate goal of understanding how false negatives arise. In particular, we aim to give an answer to some of the open questions left by [**AMB12, ABK$^+$15**]:

- *Where does cohomological contextuality sit in the hierarchy of contextuality?* Theorem II.32 says that cohomological contextuality is a stronger property than regular contextuality, but it does not fully characterise its position in the hierarchy of contextuality reviewed in Chapter II. It would be especially desirable to understand the relation between cohomological contextuality and strong contextuality. In this respect, it was conjectured in [**AMB12**, Conjecture 8.1] that, despite the existence of false negatives, cohomology is a complete invariant for

strong contextuality under suitable – though unspecified – assumptions on the measurement scenario.
- *Can higher cohomology groups be used to study contextuality?*
  The obstruction defined in [**AMB12**] is an element of the first Čech cohomology group. In classical problems involving sheaf cohomology, as well as in the related domain of *obstruction theory* [**Whi40, Ste51, EM54**], the cohomology obstructions also arise in *higher* cohomology groups. Therefore, it is natural to ask whether higher cohomology groups play any role in the detection of contextual false negatives, as suggested in [**ABK$^+$15**].
- *Can obstructions be characterised independently of cohomology?*
  Cohomology obstructions are highly abstract concepts, and have little interpretation in the physical setting of experimental scenarios. Can the same power of detection be achieved through a more concrete description?

We outline our results:

- We disprove Conjecture 8.1 of [**AMB12**] by providing an explicit example of an empirical model, defined on a simple $(2, 2, 4)$ Bell-type scenario verifying any reasonable form of connectedness and symmetry, which is strongly contextual yet cohomologically non-contextual. This counterexample clarifies the hierarchical structure of contextuality, as shown in Figure III.3.

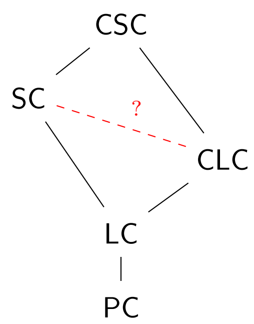


FIGURE III.1. Hierarchy proposed in [**AMB12**].

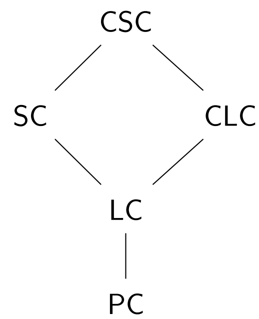


FIGURE III.2. Actual hierarchy.

FIGURE III.3. Hasse diagrams of the hierarchical structure of contextuality as proposed in [**AMB12**], and proven in the present chapter.

- We generalise cohomology obstructions to higher cohomology groups. These *higher obstructions* yield a refinement of the notion of cohomological contextuality: for each $q \geq 0$, we say that a model is q-cohomologically contextual if the q-th obstruction does not vanish.
- We show that these new levels of contextuality are organised in a precise hierarchy, described by the Hasse diagram of Figure III.4.
- We highlight a crucial limitation of higher obstructions, namely that they cannot be applied to study contextuality in no-signalling empirical models.

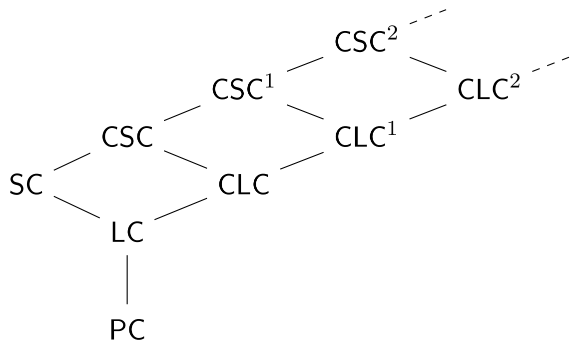


FIGURE III.4. A Hasse diagram of the hierarchy of higher cohomology obstructions introduced in this chapter

- We give a new description of the first cohomology group and, in particular, of the cohomology obstructions, using $\mathcal{F}$-torsors for an abelian representation $\mathcal{F}$.

The content of this chapter has been published in [**Car17**].

Outline of the chapter. We start, in Section 2, by presenting a concrete example of a cohomological false negative and studying the typical form of a false global section. In Section 3 we introduce the counterexample to Conjecture 8.1 of [**AMB12**]. Section 4 presents the generalisation of cohomology obstructions to higher cohomology groups and investigates its consequences. Finally, in Section 5, we present the alternative description of cohomology obstructions as $\mathcal{F}$-torsors.

2. False negatives in cohomology

We begin by introducing the issue of false negatives [**AMB12**], i.e. empirical models whose contextuality is not properly detected by cohomology. Consider the Hardy model $\mathcal{S}$ of Table II.8, and let us enumerate its possible sections as in Table III.1.

TABLE III.1. An enumeration of the possible sections of the Hardy model.

A	B	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
a_1	b_1	s_1	s_2	s_3	s_4
a_1	b_2		s_5	s_6	s_7
a_2	b_1		s_8	s_9	s_{10}
a_2	b_2	s_{11}	s_{12}	s_{13}	

In Example II.16, we proved that the model is logically contextual at s_1, as this section cannot be extended to a global section. A topological visualisation of this proof was provided in Figure II.6.

Let $\mathcal{F} := F_R\mathcal{S}$, where R is an arbitrary ring, and consider s_1 as an element of $\mathcal{F}(\{a_1, b_1\})$.[1] Because the local sections of $\mathcal{F}$ are formal linear combinations of sections of $\mathcal{S}$, section $s_1 \in \mathcal{F}(\{a_1, b_1\})$ *can* be extended to a global section of $\mathcal{F}$, namely the one corresponding to the following compatible family:

$$\{s_1,\ s_6 - s_7 + s_5,\ s_8,\ s_{12}\}.$$

Its compatibility can be easily verified: here, we only explicitly show the non-trivial step:

$$\begin{aligned}(s_6 - s_7 + s_5)|_{\{a_1\}} &= s_6|_{\{a_1\}} - s_7|_{\{a_1\}} + s_5|_{\{a_1\}} = \{a_1 \mapsto 0\} - \{a_1 \mapsto 1\} + \{a_1 \mapsto 1\} \\ &= \{a_1 \mapsto 0\} = s_1|_{\{a_1\}}; \\ (s_6 - s_7 + s_5)|_{\{b_2\}} &= s_6|_{\{b_2\}} - s_7|_{\{b_2\}} + s_5|_{\{b_2\}} = \{b_2 \mapsto 1\} - \{b_2 \mapsto 1\} + \{b_2 \mapsto 0\} \\ &= \{b_2 \mapsto 0\} = s_{12}|_{\{b_2\}}.\end{aligned}$$

By Proposition II.30, we conclude that $\gamma(s_1) = 0$, which means that cohomology is unable to detect the logical contextuality of the Hardy model, resulting in a **false negative**.[2]

A topological interpretation of this proof can be found in Figure III.5. Here, the bundle diagram of the model is presented both in its 3-dimensional and planar form. The planar representation is particularly handy to visualise false negatives, and will be used extensively in Chapter IV. In order to recover the original diagram, it is sufficient to 'glue' back together the two ends of the planar version.

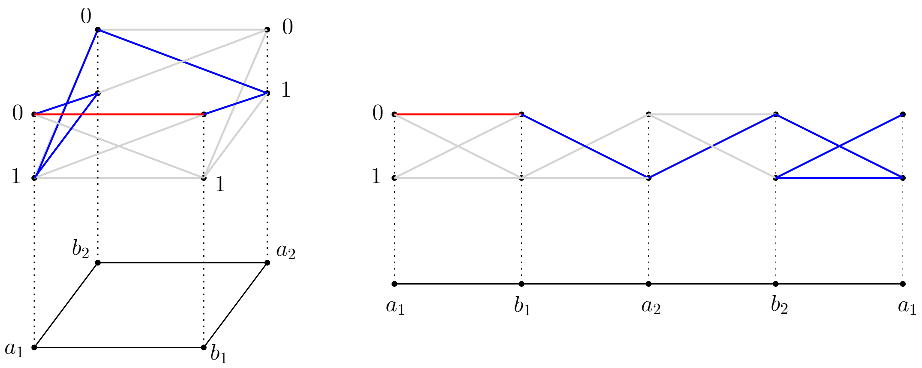


FIGURE III.5. A cohomological false negative for the Hardy model.

In the picture, section s_1 is highlighted in red, whereas the global section containing it is marked in blue. Notice that this is indeed a closed loop, but it features a 'twist' over context $\{a_1, b_2\}$. Such a twisted loop is obviously not a valid global section of $\mathcal{S}$.

The existence of such false negatives shows that sheaf cohomology does not constitute a complete invariant for contextuality.

[1]Here, we used the injection $i : \mathcal{S} \hookrightarrow \mathcal{F}$ which, in this case, is given by the trivial embedding $s_1 \mapsto 1 \cdot s_1$. From now on, we will pass from any $s \in \mathcal{S}(C)$ to $s := i_C(s) \in \mathcal{F}(C)$ without any comments. With a slight abuse of notation, we will also denote $s \in \mathcal{F}(C)$ for $i_C(s)$.

[2]In [**AMB12**], the situation we have just described is usually referred to as a *false positive*. However, we believe the term used here is more fitting as, after all, we are testing for contextuality.

3. A false negative for strong contextuality

In [**AMB12**], it is brought to attention that, although cohomology can fail to detect logical contextuality as in the case of the Hardy model, it is rather difficult to construct a strongly contextual false negative. Indeed, cohomology is able to detect the strong contextuality of a variety of well-known models, including GHZ states, PR Boxes, the Peres–Mermin magic square, all $\neg\mathsf{GCD}$ models [**AMB12**], and the whole class of models admitting All-vs-Nothing arguments [**ABK**$^+$**15**].

The only cohomological false negative for strong contextuality that has appeared in the literature [**AMB12**] is the Kochen-Specker model $\mathcal{S}$ for the cover

$$\{A,B,C\},\{B,D,E\},\{C,D,E\},\{A,D,F\},\{A,E,G\}. \tag{III.1}$$

The false negative has been observed using $F_{\mathbb{Z}_2}\mathcal{S}$ as an abelian representation for $\mathcal{S}$. Let us show how it arises.

In Table III.2, we introduce a list of variables in $\mathbb{Z}_2$ for each of the 15 possible sections of the model to determine whether it is possible to construct a global section for the abelian representation $F_{\mathbb{Z}_2}\mathcal{S}$, or, equivalently, a compatible family.

TABLE III.2. Variables for the possible sections of the Kochen-Specker model on the cover (IV.18).

Contexts	$(1,0,0)$	$(0,1,0)$	$(0,0,1)$
$\{A,B,C\}$	a	b	c
$\{B,D,E\}$	d	e	f
$\{C,D,E\}$	g	h	i
$\{A,D,F\}$	j	k	l
$\{A,E,G\}$	m	n	o

The compatibility conditions of a presumed compatible family for $\mathcal{F}$ translate into equations modulo 2. First of all, we have

$$\begin{aligned} a &= j = m & e &= h = k \\ b &= d = g = c & f &= i = n \end{aligned}$$

Moreover,

$$\begin{aligned} a\oplus c &= d\oplus f & b\oplus c &= n\oplus o & g\oplus i &= j\oplus l \\ a\oplus b &= h\oplus i & d\oplus f &= j\oplus l & g\oplus h &= m\oplus o \\ b\oplus c &= k\oplus l & d\oplus e &= m\oplus o & k\oplus l &= n\oplus o \end{aligned}$$

From these equations it follows that

$$\begin{aligned} a &= i = j = m = n = o \\ b &= c = d = e = g = h = k = l \end{aligned}$$

Thus, we can rewrite Table III.2 using the only two free variables a and b, as shown in Table III.3. Thanks to this table, one can immediately verify that the model is strongly contextual. Indeed, in order to construct a compatible family, we are only allowed to choose one section per context to which we assign 1, while the others must be zero. By simply looking at Table III.3 we can see that this is clearly impossible.

TABLE III.3. Table III.2 rewritten given compatibility equations.

Contexts	$(1,0,0)$	$(0,1,0)$	$(1,0,0)$
$\{A,B,C\}$	a	b	b
$\{B,D,E\}$	b	b	a
$\{C,D,E\}$	b	b	a
$\{A,D,F\}$	a	b	b
$\{A,E,G\}$	a	a	a

However, if we let $a = 1$ and $b = 0$ we obtain the following cohomological compatible family

$$\left\{s_{\{A,B,C\},A},\ s_{\{B,D,E\},E},\ s_{\{C,D,E\},E},\ s_{\{A,D,F\},A},\ s_{\{A,E,G\},A} \oplus s_{\{A,E,G\},E} \oplus s_{\{A,E,G\},G}\right\}$$

where we have used the following notation for sections of a Kochen-Specker model: given a context C and a measurement $m \in C$, the section $s_{C,m}$ is the section that maps m to 1 and every other $x \in C$ to 0 [**MB13**].

The false negative we have just presented features a rather asymmetrical structure. In particular "the existence of measurements belonging to a single context, namely F and G, seems to be crucial" [**AMB12**] for the manifestation of the false negative. Due to these limitations, the following conjecture was made:

CONJECTURE III.1 (Conjecture 8.1 of [**AMB12**]). *Under suitable assumptions of symmetry and connectedness of the cover, the cohomology obstruction is a complete invariant for strong contextuality.*

We introduce a counterexample to this conjecture. Consider the model $\mathcal{S}$ described in Table III.4. This model is defined on a $(2,2,4)$ Bell-type scenario, which is extremely

TABLE III.4. A possibilistic model on a $(2,2,4)$ scenario. This model is a counterexample to Conjecture III.1.

A	B	00	01	10	02	20	03	30	11	12	21	13	31	22	23	32	33
a_1	b_1	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	1
a_1	b_2	1	0	1	0	0	0	0	1	0	1	0	0	1	0	1	1
a_2	b_1	1	0	1	0	0	0	0	1	0	1	0	0	1	0	1	1
a_2	b_2	0	1	0	0	0	0	1	0	1	0	0	0	0	1	0	0

simple and verifies any reasonable form of symmetry and connectedness. The bundle diagram of the model is presented on the left hand side of Figure III.6.

By carefully analysing the picture, one verifies that none of the sections can be extended to a compatible family of $\mathcal{S}$ (i.e. a closed path containing one and only one section per context), but each one of them is contained in a compatible family of $\mathcal{F} := F_{\mathbb{Z}}\mathcal{S}$, namely a closed path similar to the one generating the false negative for the Hardy model (cf Figure III.5). For instance, we show this feature by considering the section $s_0 := (a_1, b_1) \mapsto (0,0)$: from the central diagram of Figure III.6 it appears clear that this section is non-extendable to a compatible family of $\mathcal{S}$, while the diagram on the

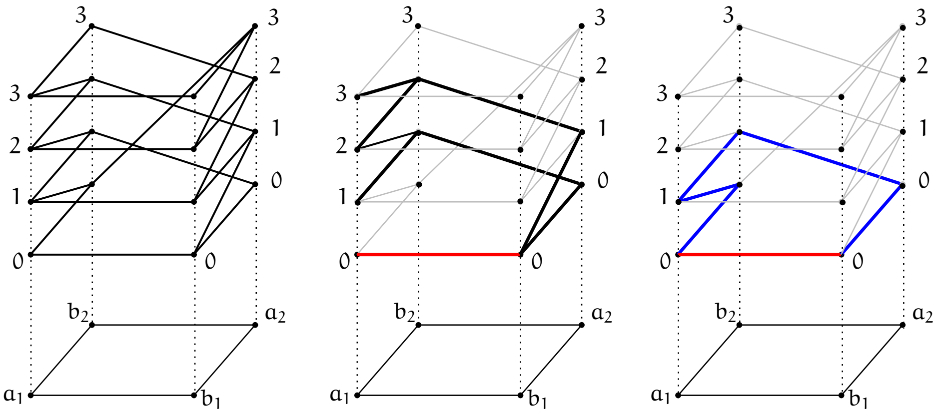


FIGURE III.6. (Left) – The bundle diagram of $\mathcal{S}$; (Center) – a visual proof of strong contextuality: none of the local sections can be extended to a closed loop; (Right) – a visual proof of false negativity: every section is part of a twisted cohomology loop.

right-hand side shows that s_0 is part of a compatible family for $\mathcal{F}$, explicitly defined by

$$\{s_0,\ (a_2, b_1) \mapsto (0,0),\ (a_2, b_2) \mapsto (0,1),$$
$$[(a_1, b_1) \mapsto (1,1)] - [(a_1, b_1) \mapsto (1,0)] + [(a_1, b_1) \mapsto (0,0)]\}.$$

The reader can verify that a false negative exists for all of the local sections in the model. This task is made significantly simpler by considering the planar representation of the bundle diagram of the model, presented in Figure III.7. Here, the reader can visually verify both that every section cannot be extended to a closed loop (i.e. that $\mathcal{S}$ is strongly contextual), and that every section is part of a compatible family for $\mathcal{F}$. For instance, we show another false negative for section $(a_1, b_2) \mapsto (0,0)$.

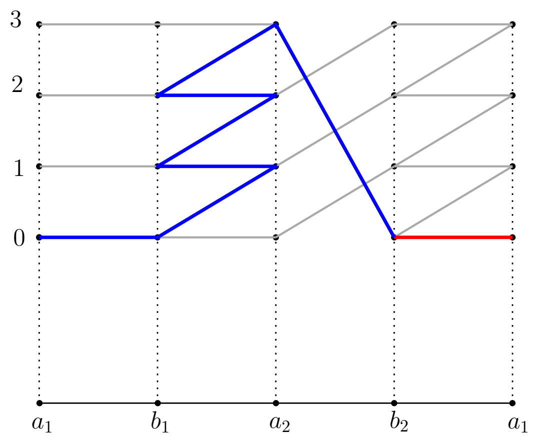


FIGURE III.7. The planar representation of the bundle diagram of $\mathcal{S}$. In blue, a cohomological false negative for the section $(a_1, b_2) \mapsto (0,0)$, marked in red.

We conclude that this model is strongly contextual but not cohomologically contextual (not even cohomologically logically contextual), essentially disproving Conjecture 8.1 of [**AMB12**].[3]

4. Extension to higher cohomology groups

The sheaf cohomological method developed in [**AMB12**] involves only the first Čech cohomology group, which contains the obstructions. The existence of extreme false negatives such as the one presented in the previous section motivates a deeper inspection of the higher cohomology groups in search of information on how such extreme cases arise. We will introduce here a generalisation of cohomology obstructions to higher-dimensional cohomology groups.

Let $\mathcal{F}$ be an abelian representation of an empirical model $\mathcal{S}$ on a scenario $\langle X, \mathcal{M}, O\rangle$. Let $q \geq 0$ be an integer and fix a context $C_0 \in \mathcal{M}$. To each section $s_0 \in \mathcal{F}(C_0)$ we associate a q-relative cochain $c^q_{s_0} \in C^q(\mathcal{M}, \mathcal{F}|_{C_0})$ defined by

$$c^q_{s_0}(\omega) := s_0|_{C_0 \cap |\omega|}, \ \forall \omega \in \mathcal{N}(\mathcal{M})^q.$$

This assignment determines a homomorphism $\psi^q : \mathcal{F}(C_0) \to C^q(\mathcal{M}, \mathcal{F}|_{C_0})$ which generalises the isomorphism ψ^0 of (II.7). Although ψ^q is not an isomorphism in general, it is always injective, which means that different sections in $\mathcal{F}(C_0)$ are mapped to distinct elements of $C^q(\mathcal{M}, \mathcal{F}|_{C_0})$.

Lemma III.2. *For each $q \geq 0$, the homomorphism ψ^q is injective.*

Proof. Let $s_0 \in \ker(\psi^q)$. Then $c^q_{s_0} = 0$, thus in particular $0 = c^q_{s_0}(\underbrace{C_0, \ldots, C_0}_{q+1 \text{ many}}) = s_0$. Therefore, $\ker(\psi^q) = 0$ and the homomorphism is injective. □

An important aspect of ψ^0 is that its image lies in $Z^0(\mathcal{M}, \mathcal{F}|_{C_0}) \cong \check{H}^0(\mathcal{M}, \mathcal{F}|_{C_0})$. However, this is not necessarily the case in higher dimensions. The following lemma shows that this feature is generalisable only in even dimensions.

Lemma III.3. *Let $q \geq 0$. The image of ψ^q is contained in $Z^q(\mathcal{M}, \mathcal{F}|_{C_0})$ if and only if q is even.*

Proof. Let $s_0 \in \mathcal{F}(C_0)$. For any $\omega \in \mathcal{N}(\mathcal{M})^{q+1}$, we have

$$\begin{aligned}\delta^q\left(c^q_{s_0}\right)(\omega) &= \sum_{k=0}^{q+1}(-1)^k \rho^{|\partial_k\omega|}_{|\omega|}\left(c^q_{s_0}(\partial_k\omega)\right) = \sum_{k=0}^{q+1}(-1)^k \rho^{|\partial_k\omega|}_{|\omega|}\left(s_0|_{C_0\cap|\partial_k\omega|}\right) \\ &= \sum_{k=0}^{q+1}(-1)^k s_0|_{C_0\cap|\omega|} = s_0|_{C_0\cap|\omega|} \cdot \sum_{k=0}^{q+1}(-1)^k\end{aligned}$$

The last sum is an alternating sum. Therefore, $\delta^q\left(c^q_{s_0}\right)(\omega) = 0$ if and only if q is even. □

[3]The open-endedness of the statement of the conjecture leaves room for a small minority of special cases where cohomology is indeed a full invariant of strong contextuality. An example is given in [**Man13**], where it is shown that the conjecture is true for the extremely limited class of symmetric Kochen-Specker models satisfying a condition due to Daykin and Häggkvist [**DH81**].

Given a $q \geq 0$, we can generalise the construction of the connecting homomorphism γ to the order $2q$. For each $\sigma \in \mathcal{N}(\mathcal{M})^{2q}$, the exact sequence (II.6) yields an exact sequence

$$0 \xrightarrow{0} \mathcal{F}_{\tilde{C}_0}(|\sigma|) := ker(p^{C_0}_{|\sigma|}) \to \mathcal{F}(|\sigma|) \xrightarrow{p^{C_0}_{|\sigma|}} \mathcal{F}|_{C_0}(|\sigma|) \longrightarrow 0.$$

Surjectivity on the right is due to the fact that $|\sigma| \cap C_0 \subseteq C_0$ for all σ, hence the map $p^{C_0}_{|\sigma|}$ is surjective since $\mathcal{S}$ is flasque beneath the cover.

We can sum these morphisms for every $\sigma \in \mathcal{N}(\mathcal{M})^{2q}$ and lift exactness to the chain level:

$$0 \xrightarrow{0} C^{2q}(\mathcal{M}, \mathcal{F}_{\tilde{C}_0}) \to C^{2q}(\mathcal{M}, \mathcal{F}) \xrightarrow{\bigoplus_\sigma p^{C_0}_{|\sigma|}} C^{2q}(\mathcal{M}, \mathcal{F}|_{C_0}) \longrightarrow 0. \tag{III.2}$$

Then, we take the correstriction $\tilde{\delta}^{2q}$ of the $2q$-th coboundary maps to Z^{2q+1} and obtain

$$\begin{tikzcd}
0 \arrow[r] & C^{2q}(\mathcal{M}, \mathcal{F}_{\tilde{C}_0}) \arrow[r] \arrow[d, "\tilde{\delta}^{2q}"] & C^{2q}(\mathcal{M}, \mathcal{F}) \arrow[r] \arrow[d, "\tilde{\delta}^{2q}"] & C^{2q}(\mathcal{M}, \mathcal{F}|_{C_0}) \arrow[r] \arrow[d, "\tilde{\delta}^{2q}"] & 0 \\
0 \arrow[r] & Z^{2q+1}(\mathcal{M}, \mathcal{F}_{\tilde{C}_0}) \arrow[r] & Z^{2q+1}(\mathcal{M}, \mathcal{F}) \arrow[r] & Z^{2q+1}(\mathcal{M}, \mathcal{F}|_{C_0}) &
\end{tikzcd}$$

Finally, we apply the snake lemma to this diagram and obtain the q-th connecting homomorphism $\tilde{\gamma}^q_{C_0}$.

(III.3)

$$\begin{tikzcd}
 & & & \mathcal{F}(C_0) \arrow[d, "\psi^{2q}"] & \\
 & Z^{2q}(\mathcal{M}, \mathcal{F}_{\tilde{C}_0}) \arrow[r] \arrow[d] & Z^{2q}(\mathcal{M}, \mathcal{F}) \arrow[r] \arrow[d] & Z^{2q}(\mathcal{M}, \mathcal{F}|_{C_0}) \arrow[d] \arrow[dddll, bend left, "\tilde{\gamma}^q_{C_0}"] & \\
0 \arrow[r] & C^{2q}(\mathcal{M}, \mathcal{F}_{\tilde{C}_0}) \arrow[r] \arrow[d] & C^{2q}(\mathcal{M}, \mathcal{F}) \arrow[r] \arrow[d] & C^{2q}(\mathcal{M}, \mathcal{F}|_{C_0}) \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & Z^{2q+1}(\mathcal{M}, \mathcal{F}_{\tilde{C}_0}) \arrow[r] \arrow[d] & Z^{2q+1}(\mathcal{M}, \mathcal{F}) \arrow[r] \arrow[d] & Z^{2q+1}(\mathcal{M}, \mathcal{F}|_{C_0}) \arrow[d] & \\
 & \check{H}^{2q+1}(\mathcal{M}, \mathcal{F}_{\tilde{C}_0}) \arrow[r] & \check{H}^{2q+1}(\mathcal{M}, \mathcal{F}) \arrow[r] & \check{H}^{2q+1}(\mathcal{M}, \mathcal{F}|_{C_0}) &
\end{tikzcd}$$

This construction naturally culminates in the following definition:

Definition III.4. Let $s_0 \in \mathcal{F}(C_0)$. We define the **q-th cohomological obstructionof** s_0 as the element

$$\gamma^q_{C_0}(s_0) := \tilde{\gamma}^q_{C_0}(\psi^{2q}(s_0)) \in \check{H}^{2q+1}(\mathcal{M}, \mathcal{F}).$$ [4]

The empirical model $\mathcal{S}$ underlying $\mathcal{F}$ is defined to be

- **cohomologically logically q-contextual at a section** s_0, or $\mathsf{CLC}^q(\mathcal{S}, s_0)$, if $\gamma^q_{C_0}(s_0) \neq 0$. We say that $\mathcal{S}$ is **cohomologically logically q-contextual** if $\mathsf{CLC}^q(\mathcal{S}, s)$ for some section s.
- **cohomologically strongly q-contextual**, or $\mathsf{CSC}^q(\mathcal{S})$, if $\mathsf{CLC}^q(\mathcal{S}, s)$ for all s.

For $q = 0$ one recovers the original notion of cohomological contextuality, introduced in Definition II.31.

Note that, due to parity arguments needed to achieve this definition, the cohomological obstruction is generalisable only to odd-dimensional cohomology groups.

In the case $q = 0$, Proposition II.30 tells us that the vanishing of the cohomological obstruction is equivalent to the existence of a compatible family in $\mathcal{F}$ containing s_0. The analogous result for higher obstructions is the following:

Proposition III.5. *Given a $q \geq 0$, a context $C_0 \in \mathcal{M}$ and a local section $s_0 \in \mathcal{F}(C_0)$, $\gamma^q_{C_0}(s_0) = 0$ if and only if there exists a family $s \in Z^{2q}(\mathcal{M}, \mathcal{F})$ such that*

$$p^{C_0}_{|\sigma|}(s(\sigma)) = c^{2q}_{s_0}(\sigma) = s_0|_{C_0 \cap |\sigma|} \ \forall \sigma \in \mathcal{N}(\mathcal{M})^{2q}. \tag{III.4}$$

Proof. $\gamma^q(s_0) = 0 \Leftrightarrow \tilde{\gamma}^q(c^{2q}_{s_0}) = 0 \Leftrightarrow c^{2q}_{s_0} \in \ker(\tilde{\gamma}^q)$. Since $\tilde{\gamma}^q$ is defined using the snake lemma, it is part of an exact sequence. Therefore, $c^{2q}_{s_0} \in \ker(\tilde{\gamma}^q)$ if and only if there exists a family $s \in Z^{2q}(\mathcal{M}, \mathcal{F})$ such that (III.4) is verified. □

4.1. A hierarchy of cohomological obstructions. Remarkably, higher cohomology obstructions are organised in a precise hierarchy of implications. In the following proposition we show that, if an obstruction vanishes at order $q \geq 0$, it must vanish at any higher order $q' \geq q$.

Theorem III.6. *Let $\mathcal{F}$ be an abelian presheaf representing an empirical model $\mathcal{S}$ on a scenario $\langle X, \mathcal{M}, O \rangle$. Let $s_0 \in \mathcal{F}(C_0)$. Then $\mathsf{CLC}^{q+1}(\mathcal{S}, s_0) \Rightarrow \mathsf{CLC}^q(\mathcal{S}, s_0)$ for all $q \geq 0$.*

Proof. We will show $\neg\mathsf{CLC}^q(\mathcal{S}, s_0) \Rightarrow \neg\mathsf{CLC}^{q+1}(\mathcal{S}, s_0)$. Suppose $\neg\mathsf{CLC}^q(\mathcal{S}, s_0)$, that is, $\gamma^q_{C_0}(s_0) = 0$. By Proposition III.5 there exists a family $s \in Z^{2q}(\mathcal{M}, \mathcal{F})$ such that $p^{C_0}_{|\sigma|}(s(\sigma)) = c^{2q}_{s_0}(\sigma)$ for all $\sigma \in \mathcal{N}(\mathcal{M})^{2q}$. For all $\tau \in \mathcal{N}(\mathcal{M})^{2q+2}$, we define

$$f(s)(\tau) := \rho^{|\partial_{2q+1}\partial_{2q+2}\tau|}_{|\tau|}(s(\partial_{2q+1}\partial_{2q+2}\tau)) = s(\partial_{2q+1}\partial_{2q+2}\tau)|_{|\tau|}.$$

Notice that $f(s)(\tau) \in \mathcal{F}(|\tau|)$, thus $f(s) \in C^{2q+2}(\mathcal{M}, \mathcal{F})$. We can actually show that $f(s)$ is in $Z^{2q+2}(\mathcal{M}, \mathcal{F})$ as follows. Given an arbitrary $\nu \in \mathcal{N}(\mathcal{M})^{2q+3}$, we have

[4]Note that if $q = 0$ this definition coincides with the one of cohomological obstruction given before, hence $\gamma^0_{C_0} = \gamma_{C_0}$.

$$
\begin{aligned}
\delta^{2q+2}(f(s))(\nu) &= \sum_{k=0}^{2q+3}(-1)^k \rho_{|\nu|}^{|\partial_k\nu|}(f(s)(\partial_k\nu)) \\
&= \sum_{k=0}^{2q+3}(-1)^k \rho_{|\nu|}^{|\partial_k\nu|}\rho_{|\partial_k\nu|}^{|\partial_{2q+1}\partial_{2q+2}\partial_k\nu|}(s(\partial_{2q+1}\partial_{2q+2}\partial_k\nu)) \\
&= \sum_{k=0}^{2q+3}(-1)^k \rho_{|\nu|}^{|\partial_{2q+1}\partial_{2q+2}\partial_k\nu|}(s(\partial_{2q+1}\partial_{2q+2}\partial_k\nu)) \\
&= \sum_{k=0}^{2q+1}(-1)^k \rho_{|\nu|}^{|\partial_{2q+1}\partial_{2q+2}\partial_k\nu|}(s(\partial_{2q+1}\partial_{2q+2}\partial_k\nu)) \\
&\qquad + \rho_{|\nu|}^{|\partial_{2q+1}\partial_{2q+2}\partial_{2q+2}\nu|}(s(\partial_{2q+1}\partial_{2q+2}\partial_{2q+2}\nu)) \\
&\qquad - \rho_{|\nu|}^{|\partial_{2q+1}\partial_{2q+2}\partial_{2q+3}\nu|}(s(\partial_{2q+1}\partial_{2q+2}\partial_{2q+3}\nu))
\end{aligned}
\tag{III.5}
$$

Notice that the last two terms of the sum cancel out since, trivially,

$$\partial_{2q+2}\partial_{2q+2}\nu = \partial_{2q+2}\partial_{2q+3}\nu.$$

Hence,

$$
\begin{aligned}
\delta^{2q+2}(f(s))(\nu) &\overset{(\text{III.5})}{=} \sum_{k=0}^{2q+1}(-1)^k \rho_{|\nu|}^{|\partial_{2q+1}\partial_{2q+2}\partial_k\nu|}(s(\partial_{2q+1}\partial_{2q+2}\partial_k\nu)) \\
&= \sum_{k=0}^{2q+1}(-1)^k \rho_{|\nu|}^{|\partial_k\partial_{2q+1}\partial_{2q+2}\nu|}(s(\partial_k\partial_{2q+1}\partial_{2q+2}\nu)),
\end{aligned}
\tag{III.6}
$$

where the last equality is valid since now $0 \le k \le 2q+1$ and thus it is unimportant whether we cancel the k-th term before or after having canceled the $(2q+2)$-th and the $(2q+1)$-th. We can now relabel $\partial_{2q+1}\partial_{2q+2}\nu =: \tilde{\nu} \in \mathcal{N}(\mathcal{M})^{2q+1}$ and obtain

$$
\begin{aligned}
\delta^{2q+2}(f(s))(\nu) &\overset{(\text{III.6})}{=} \sum_{k=0}^{2q+1}(-1)^k \rho_{|\nu|}^{|\partial_k\tilde{\nu}|}(s(\partial_k\tilde{\nu})) = \sum_{k=0}^{2q+1}(-1)^k s(\partial_k\tilde{\nu})|_{|\nu|} \\
&= \sum_{k=0}^{2q+1}(-1)^k \left(s(\partial_k\tilde{\nu})|_{|\tilde{\nu}|}\right)|_{|\nu|} = \left(\sum_{k=0}^{2q+1}(-1)^k \left(s(\partial_k\tilde{\nu})|_{|\tilde{\nu}|}\right)\right)|_{|\nu|} \\
&= \left(\sum_{k=0}^{2q+1}(-1)^k \rho_{|\tilde{\nu}|}^{|\partial_k\tilde{\nu}|}(s(\partial_k\tilde{\nu}))\right)|_{|\nu|} = \delta^{2q}(s)(\tilde{\nu})|_{|\nu|} = 0|_{|\nu|} = 0,
\end{aligned}
$$

where the second-to-last equality is due to the fact that $s \in Z^{2q}(\mathcal{M}, \mathcal{F})$.

Let $\sigma \in \mathcal{N}(\mathcal{M})^{2q+2}$, and let $\tilde{\sigma} := \partial_{2q+1}\partial_{2q+2}\sigma$. We have

$$\begin{aligned} p^{C_0}_{|\sigma|}(f(s)(\sigma)) = f(s)(\sigma)|_{|\sigma|\cap C_0} &= s(\partial_{2q+1}\partial_{2q+2}\sigma)|_{|\sigma|\cap C_0} = s(\tilde{\sigma})|_{|\sigma|\cap C_0} = s(\tilde{\sigma})|_{|\tilde{\sigma}|\cap|\sigma|\cap C_0} \\ &= \left(s(\tilde{\sigma})|_{|\tilde{\sigma}|\cap C_0}\right)|_{|\sigma|} = \left(p^{C_0}_{|\tilde{\sigma}|}(s(\tilde{\sigma}))\right)|_{|\sigma|} = \left(c^{2q}_{s_0}(\tilde{\sigma})\right)|_{|\sigma|} = \left(s_0|_{|\tilde{\sigma}|\cap C_0}\right)|_{|\sigma|} \\ &= s_0|_{|\tilde{\sigma}|\cap|\sigma|\cap C_0} = s_0|_{|\sigma|\cap C_0} = c^{2q+2}_{s_0}(\sigma). \end{aligned}$$

By Proposition III.5 this implies $\gamma^{q+1}_{C_0}(s_0) = 0$. □

This result reveals the existence of an infinite number of *levels* of contextuality organised in the following hierarchy:

$$\begin{tikzcd} \mathsf{CSC}(\mathcal{S}) \arrow[d, Rightarrow] & \mathsf{CSC}^1(\mathcal{S}) \arrow[l, Rightarrow, "\text{III.6}"'] \arrow[d, Rightarrow] & \dots \arrow[l, Rightarrow] & \mathsf{CSC}^q(\mathcal{S}) \arrow[l, Rightarrow] \arrow[d, Rightarrow] & \mathsf{CSC}^{q+1}(\mathcal{S}) \arrow[l, Rightarrow, "\text{III.6}"'] \arrow[d, Rightarrow] & \dots \arrow[l, Rightarrow] \\ \mathsf{CLC}(\mathcal{S}) & \mathsf{CLC}^1(\mathcal{S}) \arrow[l, Rightarrow, "\text{III.6}"] & \dots \arrow[l, Rightarrow] & \mathsf{CLC}^q(\mathcal{S}) \arrow[l, Rightarrow] & \mathsf{CLC}^{q+1}(\mathcal{S}) \arrow[l, Rightarrow, "\text{III.6}"] & \dots \arrow[l, Rightarrow] \end{tikzcd}$$

4.2. Higher cohomology groups cannot be used to study contextuality. Despite the successful refinement of the notion of cohomological contextuality, it turns out that higher obstructions cannot be applied to the study of contextuality in no-signalling empirical models:

THEOREM III.7. *No-signalling empirical models are cohomologically q-non-contextual for any $q > 0$.*

PROOF. Consider an abelian presheaf $\mathcal{F}$ representing an empirical model $\mathcal{S}$ on a scenario $\langle X, \mathcal{M}, O\rangle$, where $\mathcal{M} := \{C_i\}_{i\in I}$. Let $C_0 \in \mathcal{M}$ be an arbitrary context, and $s_{C_0} \in \mathcal{F}(C_0)$ an arbitrary section. By no-signalling, there exists a family $\{s_{C_i} \in \mathcal{F}(C_i)\}_{i\in I}$ such that $s_{C_i}|_{C_i\cap C_0} = s_{C_0}|_{C_i\cap C_0}$ for all i. We define $z \in C^2(\mathcal{M},\mathcal{F})$ by the expression

$$z(\omega) := s_{\partial_0\partial_2\omega}|_{|\omega|} \ \in \mathcal{F}(|\omega|) \ \forall \omega \in \mathcal{N}(\mathcal{M})^2.$$

More explicitly, given an $\omega := (C_i, C_j, C_k) \in \mathcal{N}(\mathcal{M})^2$, we define

$$z(C_i, C_j, C_k) := s_{C_j}|_{C_i\cap C_j\cap C_k} \ \in \mathcal{F}(C_i \cap C_j \cap C_k).$$

Given a general $\sigma := (C_i, C_j, C_k, C_l) \in \mathcal{N}(\mathcal{M})^3$, we have

$$\begin{aligned} \delta^2(z)(\sigma) &= z(C_j, C_k, C_l)|_{|\sigma|} - z(C_i, C_k, C_l)|_{|\sigma|} + z(C_i, C_j, C_l)|_{|\sigma|} - z(C_i, C_j, C_k)|_{|\sigma|} \\ &= s_{C_k}|_{|\sigma|} - s_{C_k}|_{|\sigma|} + s_{C_j}|_{|\sigma|} - s_{C_j}|_{|\sigma|} = 0, \end{aligned}$$

thus $z \in Z^2(\mathcal{M},\mathcal{F})$. Moreover, for any general $\omega = (C_i, C_j, C_k) \in \mathcal{N}(\mathcal{M})^2$ we have

$$\begin{aligned} p^{C_0}_{|\omega|}(z(\omega)) = z(\omega)|_{|\omega|\cap C_0} &= s_{C_j}|_{C_i\cap C_j\cap C_k\cap C_0} = \left(s_{C_j}|_{C_j\cap C_0}\right)|_{C_i\cap C_j\cap C_k\cap C_0} \\ &= \left(s_{C_0}|_{C_j\cap C_0}\right)|_{|\omega|\cap C_0} = s_{C_0}|_{|\omega|\cap C_0} = c^2_{s_{C_0}}(\omega). \end{aligned}$$

By Proposition III.5, this result implies $\gamma^1_{C_0}(s_{C_0}) = 0$, and by Theorem III.6, we conclude $\neg\mathsf{CLC}^q(\mathcal{S})$ for all $q > 0$. □

This negative result puts an end to the discussion on the role of higher cohomology in contextuality. However, it has been suggested that the implications of Theorem III.6 can potentially be used to study the signalling structure of empirical models [**Kis16a**]. We aim to investigate this aspect in future work.

5. An alternative description of the first cohomology group

Since higher cohomology groups cannot be used to infer information on how false negatives arise, we devote the last section of this chapter to a detailed study of the first cohomology group $\check{H}^1(\mathcal{M}, \mathcal{F}_{\tilde{C}_0})$. As explained in [**AMB12**], this group is of crucial importance for the cohomological study of contextuality, as it contains all of the obstructions to the existence of global sections. Its relevance has been also previously highlighted by Penrose in his *On the cohomology of impossible figures* [**Pen92**] (see also [**PP58**]), which presents "intriguing resemblances" with our study [**ABK$^+$15**], as we also pointed out in the introduction. Yet a full grasp of the nature of its elements is still to be achieved. We propose here a description of $\check{H}^1$ based on the notion of $\mathcal{F}$-torsors, as well as some considerations on the connecting homomorphism γ.

5.1. The connecting homomorphisms. The first step in understanding cohomological obstructions is studying the connecting homomorphisms. We present here some insights on how the properties of γ can give us information on the type of contextuality of an empirical model.

PROPOSITION III.8. *Let $\mathcal{F}$ be an abelian presheaf representing an empirical model $\mathcal{S}$ on a scenario $\langle X, \mathcal{M}, O\rangle$. The model is cohomologically strongly contextual if and only if γ_C is injective for all $C \in \mathcal{M}$.*

PROOF. By Proposition II.30, $\mathcal{S}$ is cohomologically strongly contextual if and only if $\gamma_C(s) \neq 0$ for all contexts $C \in \mathcal{M}$ and all sections $s \in \mathcal{F}(C)$. This is equivalent to say that $\ker(\gamma_C) = 0$ for all $C \in \mathcal{M}$. □

Thanks to this result, we can give a lower bound for the cardinality of $\check{H}^1(\mathcal{M}, \mathcal{F}_{\tilde{C}_0})$ in the case of cohomologically strongly contextual models:

$$\mathsf{CSC}(\mathcal{S}) \Rightarrow |\check{H}^1(\mathcal{M}, \mathcal{F}_{\tilde{C}_0})| \geq |\mathcal{F}(C_0)|.$$

On the other hand, given a $\mathsf{CLC} \wedge \neg\mathsf{CSC}$ model, Proposition III.8 implies that there exist two distinct sections that give rise to the same non-zero cohomological obstruction.

The injectivity of a single connecting homomoprhism is a sufficient condition for the strong contextuality of an empirical model.

PROPOSITION III.9. *Let $\mathcal{F}$ be an abelian presheaf representing an empirical model $\mathcal{S}$ on a scenario $\langle X, \mathcal{M}, O\rangle$. If there exists a $C_0 \in \mathcal{M}$ such that γ_{C_0} is injective, then $\mathcal{S}$ is strongly contextual.*

PROOF. If $\mathcal{S}$ is not strongly contextual, there must exist a context $\bar{C} \in \mathcal{M}$ and a section $s \in \mathcal{S}(\bar{C})$ that is extendable to a compatible family $\sigma := \{s_C \in \mathcal{S}(C)\}_{C\in\mathcal{M}}$. Consider the section s_{C_0} of this family. It is trivially an extendable local section since it is part of the compatible family σ, thus $\neg\mathsf{LC}(\mathcal{S}, s_{C_0})$. By Theorem II.32, this implies $\neg\mathsf{CLC}(\mathcal{S}, s_{C_0})$ or, equivalently, $\gamma_{C_0}(s_{C_0}) = 0$. $\mathcal{F}$ represents $\mathcal{S}$, thus s_{C_0} is non-zero in $\mathcal{F}(C_0)$, hence we conclude that $\ker(\gamma_{C_0}) \neq 0$, which means that γ_{C_0} is not injective. □

Notice that these two propositions clarify how CSC is a stronger condition than SC: we need all the connecting homomorphisms $\{\gamma_C\}_{C\in\mathcal{M}}$ to be injective in order to conclude that a model is CSC, but it is sufficient to have a single injective γ_C to conclude that it is SC.

5.2. $\mathcal{F}$-torsors. In this section, we review the concept of *torsor relative to a presheaf.* We will then proceed to show how to re-interpret cohomology obstructions as torsors.

Definition III.10. Let $\mathcal{F} : \mathbf{Open}(X)^{op} \to \mathbf{AbGrp}$ be a presheaf of abelian groups over a topological space X. An **$\mathcal{F}$-presheaf** is a presheaf of sets T over X equipped with a morphism of presheaves $\phi : \mathcal{F} \times T \Rightarrow T$ such that, for each open $U \subseteq X$, the map

$$\phi_U : \mathcal{F}(U) \times T(U) \to T(U) :: (g,t) \mapsto g \bullet t$$

is a left action of $\mathcal{F}(U)$ on $T(U)$.

Given two $\mathcal{F}$-presheaves T and T', a **morphism of $\mathcal{F}$-presheaves** from T to T' is a natural transformation $\psi : T \Rightarrow T'$ such that ψ_U is equivariant for all open $U \subseteq X$. That is, for all $t \in T(U)$, for all $g \in \mathcal{F}(U)$,

$$\psi_U(g \bullet t) = g \bullet \psi_U(t).$$

An $\mathcal{F}$-presheaf is a generalisation of the notion of *G-presheaf* (where G is a group), also referred to as *principal G-bundle* or *principal homogeneous space* [**Châ44, Châ46**], which has been proved to give rise to valuable alternative descriptions of elements of the first cohomology group [**LT58**]. $\mathcal{F}$-presheaves present a similar connection in the case of sheaf cohomology [**Dus75, Mil16**], which we shall now adapt to fit the study of contextuality. We start by introducing the concept of $\mathcal{F}$-torsor.

Definition III.11. An **$\mathcal{F}$-torsor** T is an $\mathcal{F}$-presheaf such that

(1) There exists an open cover $\mathcal{V}$ of X that **trivialises** T, i.e. such that $T(V) \neq \emptyset$ for all $V \in \mathcal{V}$.
(2) The actions $\phi_U : \mathcal{F}(U) \times T(U) \to T(U)$ are simply transitive. That is, for all $s,t \in T(U)$, there exists a unique $g \in \mathcal{F}(U)$ such that $g \bullet s = t$.

The simplest example of $\mathcal{F}$-torsor is the **trivial $\mathcal{F}$-torsor** $\mathcal{UF}$,[5] where the action is simply given by $g.\mathcal{U}(h) := \mathcal{U}(g+h)$. We denote by $\mathsf{Trs}_{\mathcal{F}}$ the set of isomorphism classes of $\mathcal{F}$-torsors. One can show that an $\mathcal{F}$-torsor T is isomorphic to the trivial $\mathcal{F}$-torsor if and only if $T(X) \neq \emptyset$ [**Mil16**].

5.3. $\mathcal{F}$-torsors and contextuality. In this final section, we adapt the discussion carried out so far to fit the contextuality framework. Let $\mathcal{F}$ be an abelian presheaf representing an empirical model $\mathcal{S}$ on a scenario $\langle X, \mathcal{M}, O\rangle$, with $\mathcal{M} := \{C_i\}_{i\in I}$. Let

$$\mathsf{Trs}(\mathcal{M},\mathcal{F}) := \{T \in \mathsf{Trs}_{\mathcal{F}} \mid T \text{ is trivialised by } \mathcal{M}\},$$

seen as a pointed set with the isomorphism class of the trivial $\mathcal{F}$-torsor as distinguished element. The following proposition clarifies the connection between $\mathcal{F}$-torsors and the

[5]Here, $\mathcal{U} : \mathbf{AbGrp} \to \mathbf{Set}$ denotes the forgetful functor. To avoid confusion, we will not explicitly show its presence: the trivial $\mathcal{F}$-torsor will be simply denoted by $\mathcal{F}$.

first Čech cohomology group containing the obstructions to the extension of local sections. This result is an adaptation of the well-known correspondence between torsors and cohomology [**Mil16**].

PROPOSITION III.12. *There is a bijection of pointed sets*[6] $\mathsf{Trs}(\mathcal{M},\mathcal{F}) \cong \check{H}^1(\mathcal{M},\mathcal{F})$.

PROOF. Let $T \in \mathsf{Trs}(\mathcal{M},\mathcal{F})$. Because $\mathcal{M}$ trivialises T, we can arbitrarily choose a collection $\{t_i \in T(C_i)\}_{i\in I}$. By simple transitivity, for all $i,j \in I$, there exists a unique $g_{ij} \in \mathcal{F}(C_i \cap C_j)$ such that $g_{ij} \bullet t_i|_{C_i\cap C_j} = t_j|_{C_i\cap C_j}$. We also have

$$\begin{aligned}(g_{jk}|_{C_i\cap C_j\cap C_k} + g_{ij}|_{C_i\cap C_j\cap C_k}) \bullet t_i|_{C_i\cap C_j\cap C_k} &= g_{jk}|_{C_i\cap C_j\cap C_k} \bullet \left(g_{ij} \bullet t_i|_{C_i\cap C_j}\right)|_{C_i\cap C_j\cap C_k} \\ &= g_{jk}|_{C_i\cap C_j\cap C_k} \bullet \left(t_j|_{C_i\cap C_j}\right)|_{C_i\cap C_j\cap C_k} \\ &= \left(g_{jk} \bullet t_j|_{C_j\cap C_k}\right)|_{C_i\cap C_j\cap C_k} \\ &= t_k|_{C_i\cap C_j\cap C_k} \\ &= g_{ik} \bullet t_i|_{C_i\cap C_j\cap C_k},\end{aligned}$$

which implies $g_{jk}|_{C_i\cap C_j\cap C_k} + g_{ij}|_{C_i\cap C_j\cap C_k} = g_{ki}|_{C_i\cap C_j\cap C_k}$ for all $i,j,k \in I$ by the uniqueness part of simple transitivity. This equation says that $\check{T}$, defined by $\check{T}(C_i, C_j) := g_{ij}$ for all $i,j \in I$, is a 1-cocycle, i.e. an element of $Z^1(\mathcal{M},\mathcal{F})$. Let

$$F : \mathsf{Trs}(\mathcal{M},\mathcal{F}) \to \check{H}^1(\mathcal{M},\mathcal{F}) :: T \mapsto [\check{T}].$$

In order to show that this map is well-defined, we need to prove that $\check{T}$ is independent of the choice of the family $\{t_i\}_{i\in I}$. Suppose we choose $\{t'_i \in T(C_i)\}_{i\in I}$ instead, then we obtain a family $\{g'_{ij} \in \mathcal{F}(C_i \cap C_j)\}_{i,j\in I}$ as before. By simple transitivity, for each $i \in I$ there exists an element $g_i \in \mathcal{F}(C_i)$ such that $g_i \bullet t'_i = t_i$. Thus, we obtain a family $H := \{g_i \in \mathcal{F}(C_i)\}_{i\in I}$. We have

$$\begin{aligned}\left(g'_{ij} + g_i|_{C_i\cap C_j}\right) \bullet t'_i|_{C_i\cap C_j} &= g'_{ij} \bullet \left(g_i|_{C_i\cap C_j} \bullet t'_i|_{C_i\cap C_j}\right) = g'_{ij} \bullet t_i|_{C_i\cap C_j} \\ &= t_j|_{C_i\cap C_j}, \ \forall i,j \in I.\end{aligned}$$

On the other hand,

$$\begin{aligned}\left(g_j|_{C_i\cap C_j} + g_{ij}\right) \bullet t'_i|_{C_i\cap C_j} &= g_j|_{C_i\cap C_j} \bullet \left(g_{ij} \bullet t'_i|_{C_i\cap C_j}\right) = g_j|_{C_i\cap C_j} \bullet t'_j|_{C_i\cap C_j} \\ &= t_j|_{C_i\cap C_j}, \ \forall i,j \in I.\end{aligned}$$

Again, by the uniqueness part of simple transitivity, this implies $g'_{ij} + g_i|_{C_i\cap C_j} = g_j|_{C_i\cap C_j} + g_{ij}$ for all $i,j \in I$, which is equivalent to say $\delta^0(H)(C_i, C_j) = g'_{ij} - g_{ij}$ for all $i,j \in I$. Consequently, it does not matter whether we define $\check{T}(C_i, C_j) := g_{ij}$ or $\check{T}(C_i, C_j) := g'_{ij}$ since these two 1-cocycles are cohomologous.

Notice that F maps the trivial $\mathcal{F}$-torsor to $0 \in \check{H}^1(\mathcal{M},\mathcal{F})$, thus it is a morphism of pointed sets. To prove that F is a bijection, we introduce an inverse $G : \check{H}^1(\mathcal{M},\mathcal{F}) \to \mathsf{Trs}(\mathcal{M},\mathcal{F})$. Given $[z] \in \check{H}^1(\mathcal{M},\mathcal{F})$, we define the presheaf $G([z]) : \mathbf{Open}(X)^{op} \to$ **AbGrp** as follows: for all $U \subseteq X$,

[6] The distinguished point of $\mathsf{Trs}(\mathcal{M},\mathcal{F})$ is the trivial $\mathcal{F}$-torsor, whereas the one of $\check{H}^1(\mathcal{M},\mathcal{F})$ is 0. The bijection maps one to the other.

$$G([z])(U) := \left\{ (t_i)_i \in \bigoplus_{i \in I} \mathcal{F}(C_i \cap U) \;\middle|\; t_i|_{C_{i,j} \cap U} - t_j|_{C_{i,j} \cap U} = z(C_i, C_j)|_{C_{i,j} \cap U}, \forall i, j \in I \right\},$$

where we have used the special notation $C_{i,j} := C_i \cap C_j$ for the sake of simplicity.

The restriction maps are given by

$$G([z])(U \subseteq U') :: (t'_i)_{i \in I} \mapsto (t'_i|_{C_i \cap U})_{i \in I}.$$

We define an $\mathcal{F}$-action on $G([z])$ by the expression

$$g \bullet (t_i)_{i \in I} := (t_i - g|_{C_i \cap U})_i,$$

for any $g \in \mathcal{F}(U)$.

We need to show that $G([z]) \in \mathsf{Trs}(\mathcal{M}, \mathcal{F})$. To do so, we prove that for any context $C_j \in \mathcal{M}$, there exists an isomorphism of $\mathcal{F}|_{C_j}$-presheaves $\mathcal{F}|_{C_j} \Rightarrow G([z])|_{C_j}$ (recall that $\mathcal{F}$ denotes the trivial $\mathcal{F}$-torsor). Consider a $U \subseteq C_j$. The map

$$h^j_U : \mathcal{F}|_{C_j}(U) \to G([z])|_{C_j}(U) :: g \mapsto \left(z(C_i, C_j)|_{C_i \cap C_j \cap U} - g|_{C_i \cap C_j \cap U} \right)_{i \in I}$$

is an isomorphism with inverse

$$k^j_U : G([z])|_{C_j}(U) \to \mathcal{F}|_{C_j}(U) :: (t_i)_{i \in I} \mapsto -t_j.$$

In fact, h^j_U is equivariant since

$$\begin{aligned} g \bullet h^j_U(h) &= g \bullet \left(z(C_i, C_j)|_{C_i \cap C_j \cap U} - h|_{C_i \cap C_j \cap U} \right)_{i \in I} \\ &= \left(z(C_i, C_j)|_{C_i \cap C_j \cap U} - h|_{C_i \cap C_j \cap U} - g|_{C_i \cap C_j \cap U} \right) \\ &= h^j_U(\mathcal{U}(g + h)) = h^j_U(g \bullet h), \end{aligned}$$

where the last action is the one of the trivial $\mathcal{F}$-torsor. Moreover, k^j_U is indeed the inverse of h^j_U:

$$h^j_U \left(k^j_U \left((t_i)_{i \in I} \right) \right) = h_U(-t_j) = \left(z(C_i, C_j)|_{C_i \cap C_j \cap U} + t_j \right)_{i \in I} = (t_i - t_j + t_j)_{i \in I} = (t_i)_{i \in I},$$

and

$$k^j_U(h^j_U(g)) = k_U \left(\left(z(C_i, C_j)|_{C_i \cap C_j \cap U} - g|_{C_i \cap C_j \cap U} \right)_{i \in I} \right) = -z(C_j, C_j)|_{C_j \cap U} + g = g,$$

where the last equality is due to the fact that z is a 1-cocycle. Since $\mathcal{F}|_{C_j} \cong G([z])|_{C_j}$ for all contexts C_j, we know that $G([z])$ is an $\mathcal{F}$-torsor trivialised by the measurement cover $\mathcal{M}$.

We also need to show that the definition of G is independent of the choice of the representative z of the 1-cocycle $[z]$. Suppose we take a cohomologous 1-cocycle z'. Then there exists a family $h := \{h_i \in \mathcal{F}(C_i)\}_{i \in I}$ such that $z'(C_i, C_j) - z(C_i, C_j) = \delta^0(h)$. Then we can define an isomorphism of $\mathcal{F}$-torsors $g([z]) \cong g([z'])$ induced by the maps

$$\psi_U : G([z])(U) \to G([z'])(U) :: (t_i)_{i \in I} \mapsto (h_i|_{C_i \cap U} + t_i)_{i \in I}.$$

In fact, this map is equivariant since

$$\begin{aligned} g \bullet \psi_U \left((t_i)_{i \in I} \right) &= g \bullet \left((h_i|_{C_i \cap U} + t_i)_{i \in I} \right) = (h_i|_{C_i \cap U} + t_i - g|_{C_i \cap U})_{i \in I} \\ &= \psi_U \left((t_i - g|_{C_i \cap U})_{i \in I} \right) = \psi_U \left(g \bullet (t_i)_{i \in I} \right), \end{aligned}$$

and its inverse is clearly

$$G([z'])(U) \to G([z])(U) :: (t'_i)_{i\in I} \mapsto (t'_i - h_i|_{C_i\cap U})_{i\in I}.$$

We can finally show that G is the inverse of F.

- Let $T \in \mathsf{Trs}(\mathcal{M}, \mathcal{F})$. We want to show that $T \cong G([\check{T}])$. Let $U \subseteq X$, and suppose that $\check{T}$ is defined with respect to the family $\{t_i \in T(C_i)\}_{i\in I}$. Consider an element $s \in T(U)$ and the induced family $\{s_i \in T(U \cap C_i)\}_{i\in I} := \{s|_{C_i\cap U}\}_{i\in I}$.[7] By simple transitivity, for each $i \in I$ there is a unique $g_i \in \mathcal{F}(C_i \cap U)$ such that $g_i \bullet s_i = t_i|_{C_i\cap U}$. This allows us to define the isomorphism

  $$\phi_U : T(U) \to G([\check{T}])(U) :: s \to (g_i)_{i\in I}$$

  We leave to the reader the rather simple verification of the fact it is actually an isomorphism, but we explicitly show that it is equivariant. To see this, let $h \in \mathcal{F}(U)$. We have $\phi_U(h \bullet s) = (k_i)_{i\in I}$, where, for all $i \in I$, k_i is the unique element in $\mathcal{F}(C_i \cap U)$ such that $k_i \bullet (h \bullet s)|_{C_i\cap U} = t_i|_{C_i\cap U}$. More explicitly, k_i is the unique element such that

  $$k_i \bullet (h|_{C_i\cap U} \bullet s_i) = t_i|_{C_i\cap U},$$

  which is equivalent to

  $$(k_i + h|_{C_i\cap U}) \bullet s_i = t_i|_{C_i\cap U}.$$

  On the other hand, $h \bullet \phi_U(s) = h \bullet (g_i)_{i\in I} = (g_i - h|_{C_i\cap U})_{i\in I}$. Since

  $$(g_i - h|_{U\cap C_i}) \bullet (h|_{C_i\cap U} \bullet s_i) = (g_i - h|_{C_i\cap U} + h|_{C_i\cap U}) \bullet s_i = g_i \bullet s_i = t_i|_{C_i\cap U},$$

  we conclude by simple transitivity that $k_i = g_i - h|_{U\cap C_i}$ for all $i \in I$, which leads to $h \bullet \phi_U(s) = \phi_U(h \bullet s)$.
- Let $[z] \in \check{H}^1(\mathcal{M}, \mathcal{F})$. We want to show that $F(G([z])) = [z]$. We construct the family $\{t_k \in G([z])(C_k)\}$ given by $t_k := (z(C_i, C_k))_{i\in I}$ and we use it to define $F(G([z]))$ by setting, for all $i, j \in I$, $F(G([z]))(C_i, C_j)$ to be the unique element $g_{ij} \in \mathcal{F}(C_i \cap C_j)$ such that $g_{ij} \bullet t_j|_{C_i\cap C_j} = t_i|_{C_i\cap C_j}$. Notice that

  $$\begin{aligned} z(C_l, C_k) \bullet t_k|_{C_l\cap C_k} &= z(C_l, C_k) \bullet (z(C_i, C_k)|_{C_i\cap C_l\cap C_k})_{i\in I} \\ &= (z(C_i, C_k)|_{C_i\cap C_l\cap C_k} - z(C_l, C_k)|_{C_i\cap C_l\cap C_k})_{i\in I} \\ &= (z(C_i, C_l))_{i\in I} = t_l|_{C_l\cap C_k}. \end{aligned}$$

  Therefore, by simple transitivity, $g_{ij} = z(C_i, C_j)$ for all $i, j \in I$, proving $F(G([z])) = [z]$.

□

This bijection equips $\mathsf{Trs}(\mathcal{M}, \mathcal{F})$ with a group structure. The addition of two $\mathcal{F}$-torsors is defined componentwise at each subset $U \subseteq X$ by

$$G([z])(U) + G([w])(U) := G([z] + [w])(U)$$

[7] Note the similarities with the construction of cohomological obstruction in [**AMB12**], where we take a no-signalling family for the initial section.

for all $[z], [w] \in \check{H}^1(\mathcal{M}, \mathcal{F})$. Clearly, the above bijection becomes an isomorphism of abelian groups with respect to this addition.

This results implies that the elements of the first cohomology group $\check{H}^1(\mathcal{M}, \mathcal{F}_{\tilde{C}_0})$ relative to a context $C_0 \in \mathcal{M}$ (and, in particular, cohomological obstructions) can be seen as isomorphism classes of $\mathcal{F}_{\tilde{C}_0}$-torsors trivialised by the measurement cover $\mathcal{M}$.

Until now, elements of $\check{H}^1$ could only be identified via the abstract equations imposed by the rigid definition of cohomology. The reason why we believe the new description might be more satisfactory, is that despite their seemingly sophisticated definition, torsors are rather simple objects, as explained by Baez in [**Bae09**]. In the simplest terms, an $\mathcal{F}_{\tilde{C}_0}$-torsor is the presheaf $\mathcal{F}_{\tilde{C}_0}$ having lost its identity in each group $\mathcal{F}_{\tilde{C}_0}(U)$, for $U \subseteq X$. Rather than describing the local sections at each $\mathcal{F}_{\tilde{C}_0}(U)$, it measures their difference. This aspect is particularly important for the purpose of studying contextuality as it allows to effectively capture the non-compatibility of families of local sections by evaluating their difference at the intersection of contexts.

Discussion

Sheaf cohomology is a powerful method for the detection of contextuality. However, our work has highlighted some decisive limitations. In particular, we showed that it cannot provide a full invariant for contextuality (neither logical nor strong), not even under reasonably strong symmetry and connectedness assumptions. We proved that cohomology obstructions can be naturally generalised to higher cohomology groups, resulting in a new hierarchy of contextuality levels. However, nothing can be gained from this refinement in the study of no-signalling empirical models.

It shall be mentioned that another approach to higher cohomology, based on the work of Steenrod [**Ste51**] and Eilenberg [**EM54**] on *obstruction theory*, has been attempted by the author. This method involves envisioning empirical models as *simplicial fibrations* over the simplicial complexes describing the scenarios, as suggested in [**Bar15b**], and use obstruction theory to study contextuality in the guise of a cross-section construction problem. This perspective was developed by taking advantage of alternative description of obstruction theory based on *Potsnikov towers* [**Pos51**], and their simplicial versions [**DK84, May92**]. However, this method proved to be much more difficult to implement as it requires models to be defined as Kan fibrations over Kan simplicial sets (rather than simplicial complexes), due to the necessity of defining homotopy classes of the spaces in question. This condition dramatically restricts the number of empirical models to which the theory can be applied, and it does not appear to offer any advantage in terms of false negatives.

The results of this chapter could potentially undermine our topological viewpoint on contextuality, and demand further study in search for a complete cohomology invariant for contextuality. Chapter IV will address this issue in detail with new methods and constructions.

In the last section of this chapter, we have provided an alternative description of the first cohomology group using $\mathcal{F}$-torsors. Although this approach is still at a developing stage, it allows to understand cohomological obstructions as a more concrete mathematical entity. The torsor viewpoint also presents intriguing connections with *gauge theory* [**Wey29, YM54b, YM54a**], which, in turn, appears to be the right framework to naturally formalise the concept of bundle diagram by describing empirical models as principal bundles (see also [**Cun19**]). A discrete version of gauge theory, developed in [**Man87**], also opens up the possibility of developing a different algebraic topological method to detect contextuality, based on *holonomy* [**Car26**] rather than cohomology [**AK17**]. This new perspective, which presents striking similarities with the work of Simon [**Sim83**] on the application of holonomy theory to study classic topics in quantum physics such as the Berry phase [**Ber84**] and the quantum adiabatic theorem [**BF28**], is currently under development by the author in collaboration with Samson Abramsky.

CHAPTER IV

Towards a complete cohomological invariant for non-locality and contextuality

Summary

This chapter introduces a cohomological invariant for non-locality and contextuality which is applicable to the large majority of empirical models, and conjectured to be complete. The issue of cohomological false negatives introduced in Chapter III is solved by presenting a novel construction, which derives the so-called line *versions of empirical models, exposing their deeper topological structure. The power of the invariant is demonstrated in a large number of examples.*

1. Overview

The limitations of the Čech cohomological approach to contextuality highlighted in Chapter III motivate further research to achieve a complete topological invariant. Apart from sheaf cohomology, a number of other cohomology theories have been proposed to investigate contextual features, yet none of these frameworks have solved the issue of false negatives satisfactorily. In [**Rau16**], Raussendorf introduced an approach to contextuality based on group cohomology, which was later expanded and integrated with elements of simplicial cohomology in [**ORBR17, OTR18**]. This viewpoint is limited to the class of arguments used in measurement based quantum computation [**RB01b**], which is a particular subclass of all-vs-nothing arguments [**Aas18**], for which Čech cohomology has already been proved to be a complete invariant [**ABK$^+$15**]. The work of Roumen [**Rou17**] made some important steps forward in the direction of a full invariant. His approach, based on cyclic cohomology [**Con81, Con83, Con85**] and order cohomology [**FW98, Pul06**] of effect algebras [**FB94**], gives rise to a false negative–free obstruction. However, the result is obtained at the expense of the practical computability of the cohomological invariant, which appears so complex that no application to concrete empirical models has been presented yet.

In this chapter, we introduce a different viewpoint on models and scenarios, which is particularly conductive for the application of the usual Čech cohomology framework. In contrast with Chapter III, where the emphasis was put on sharpening cohomology tools, here we focus on finding the best topological representation of empirical models to prevent the occurrence of false negatives. This approach leads to the construction of so-called *line* versions of models and scenarios, i.e. series of modified structures that capture the deeper topological properties of a presheaf by explicitly encapsulating the extendability properties of its local sections.

The result is a complete sheaf cohomological invariant applicable to a very large class of empirical models, whose power of detection can be easily tested in a variety of situations, and whose practical computability is not compromised. Although it has not been possible to prove that the invariant applies to *all* models, we conjecture it works universally. Indeed, the class of models that could potentially escape the net is so restricted that it has not been possible to produce an example of such a model, much less a false negative for it.

We provide an exhaustive list of concrete examples where the line model construction solves the issue of false negatives. This includes all of the problematic models presented in Chapter III and many more.

The content of this chapter has been presented at the 15th International Conference on Quantum Physics and Logic. A pre-print is available at [**Car18**].

Outline of the chapter. We start in Section 2 by introducing *line* models and scenarios and present a number of examples. Section 3 investigates the contextual properties of line models. In Section 4 we focus on the special class of cyclic models and study their structure. Section 5 presents a complete cohomological invariant for contextuality over cyclic models, and demonstrates its power on a large number of examples, including all the cohomological false negatives introduced in previous chapters. The invariant is then extended to an extremely large class of models in Section 6, and demonstrated on the remaining known false negatives. Finally, in Section 7, we show how the line construction can be naturally extended to probabilistic empirical models.

2. Line models and scenarios

In this section, we introduce the construction of line models and scenarios. The rationale behind their conception can be explained more easily through the following concrete example.

2.1. A motivating example. We introduce another example of cohomological false negative that will serve as a guide for the development of our strategy. Consider the model described by Table IV.1, whose bundle diagram is depicted in Figure IV.1.

TABLE IV.1. A logically contextual empirical model.

A	B	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
a_1	b_1	1	0	0	1
a_1	b_2	1	0	1	1
a_2	b_1	1	0	0	1
a_2	b_2	0	1	1	0

By simply looking at the diagram, one can see that section $s := (a_1, b_1) \mapsto (1,1)$, highlighted in red, is not part of any compatible family, which means that the model is logically contextual at s. However, given the abelian representation $\mathcal{F} := F_{\mathbb{Z}_2}\mathcal{S}$ of $\mathcal{S}$, we

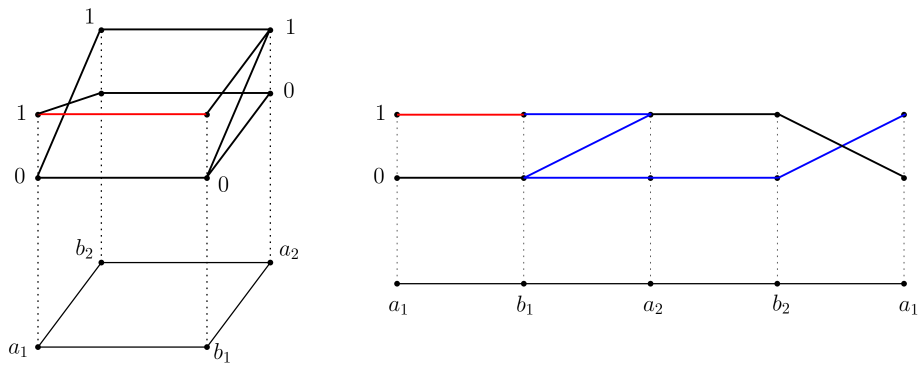


FIGURE IV.1. A cohomology false negative for section $(a_1, b_1) \mapsto (1, 1)$, highlighted in red

can see that s is part of the compatible family

$$\{s,(a_2, b_1) \mapsto (1,1) \ + \ (a_2, b_1) \mapsto (1,0) \ + \ (a_2, b_1) \mapsto (0,0), \\ (a_2, b_2) \mapsto (0,0), (a_1, b_2) \mapsto (0,1)\},$$

which means that the model is not cohomologically logically contextual at s. The 'false' global section is highlighted in blue in Figure IV.1, featuring the typical 'Z'-shaped path of cohomological false negatives. For future reference, loops of this kind will be called *cohomology loops*, or *non-standard loops*.

Now, suppose that, in the process of trying to extend s to form a closed loop, we could 'force' the selection of section $(a_2, b_1) \mapsto (1, 1)$ for the context $\{a_2, b_1\}$. This would disallow the 'Z' path higlighted in blue in Figure IV.1, which is ultimately responsible for the existence of a false negative. It would then be possible to conclude that s cannot be extended to a closed loop, even the ones allowed by linear combinations typical of cohomology.

Our strategy will closely follow this idea. Instead of focusing on a single section, we aim at capturing all the possible ways to extend a section to its immediately adjacent contexts. To do this, we introduce the concept of *line version* of an empirical model, a presheaf whose sections correspond to 'forced' extensions of local sections to adjacent contexts. By repeating this construction a sufficient amount of times, it should be possible to determine whether a section can be extended globally.

2.2. Line scenarios. In order to introduce line models, we will need to modify the structure of the scenario on which the empirical model in question is defined. We recall that the cover $\mathcal{M}$ of any scenario will always be assumed to be connected.

DEFINITION IV.1. Let $\langle X, \mathcal{M}, (O_m) \rangle$ be a measurement scenario. We define the **first line scenario** of $\langle X, \mathcal{M}, (O_m) \rangle$ as the scenario

$$\langle X, \mathcal{M}, (O_m) \rangle^{(1)} := \left\langle X^{(1)}, \mathcal{M}^{(1)}, (O_m^{(1)})_{m \in X^{(1)}} \right\rangle .$$

where

- $X^{(1)} := \mathcal{M}$, i.e. the measurements of the line scenario are the contexts of the original scenario.

- If $\mathcal{M}$ contains a single context C, we let $\mathcal{M}^{(1)} := \{\{C\}\}$.[1] Otherwise, we have $|\mathcal{M}| \geq 2$, and define
$$\mathcal{M}^{(1)} := \left\{\{C, C'\} \subseteq \mathcal{M} \mid C \neq C' \text{ and } C \cap C' \neq \emptyset\right\}.$$
That is, a context of the line scenario is a pair of intersecting contexts of the original one.
- For all $C \in \mathcal{M}$, $O^{(1)}_C := \mathcal{E}(C)$, where $\mathcal{E} : \mathcal{P}(X)^{op} \to \mathbf{Set}$ is the sheaf of events of $\langle X, \mathcal{M}, (O_m)\rangle$.

Proposition IV.2. *Let $\langle X, \mathcal{M}, (O_m)\rangle$ be a measurement scenario. Then its first line scenario is well-defined.*

Proof. First of all, note that $X^{(1)}$ is finite because X is finite. We clearly have $\mathcal{M}^{(1)} \subseteq \mathcal{P}(X^{(1)})$. Hence, we only need to show that $\mathcal{M}^{(1)}$ is a cover and an antichain. If $\mathcal{M}$ contains a single context C, this is trivially verified. Indeed, $\mathcal{M}^{(1)} = \{\{C\}\}$, thus the antichain condition is trivial. Moreover, we have
$$\bigcup \mathcal{M}^{(1)} = \{C\} = \mathcal{M} = X^{(1)}.$$
Now, suppose $|\mathcal{M}| \geq 2$. We have
$$\bigcup \mathcal{M}^{(1)} = \bigcup_{\substack{C, C' \in \mathcal{M} \\ C \cap C' \neq \emptyset \\ C \neq C'}} \{C, C'\} = \mathcal{M} = X^{(1)}.$$
Indeed,

- The inclusion $\bigcup \mathcal{M}^{(1)} \subseteq \mathcal{M} = X^{(1)}$ is trivial, given that each $M \in \mathcal{M}^{(1)}$ is included in $\mathcal{M}$ by definition.
- Let $C \in \mathcal{M}$. Since $|\mathcal{M}| \geq 2$, there exists a distinct $C' \in \mathcal{M}$. Since $\mathcal{M}$ is connected, there exists a sequence $C = C_0, \dots C_n = C'$ such that $C_i \cap C_{i+1} \neq \emptyset$, hence $C \in \{C_0, C_1\} \subseteq \bigcup \mathcal{M}^{(1)}$.

The antichain condition is also easily verifiable since any $\{C, C'\} \in \mathcal{M}^{(1)}$ has cardinality 2 by definition, thus inclusion implies equality. $\square$

It is worth spelling out the definition of the sheaf of events of the first line scenario, which we will denote by $\mathcal{E}^{(1)}$.We have $\mathcal{E}^{(1)} : \mathcal{P}(X^{(1)})^{op} \to \mathbf{Set}$, where, given a $\mathcal{U} \subseteq X^{(1)}$, we have
$$\mathcal{E}^{(1)}(\mathcal{U}) := \prod_{C \in \mathcal{U}} O^{(1)}_C = \prod_{C \in \mathcal{U}} \mathcal{E}(C),$$
with restriction maps given by the obvious projections.

To have a better understanding of how the first line scenario is defined, we give an example in Figure IV.2. On the left-hand side we show a simplicial complex representation of the measurement cover
$$\mathcal{M} = \{\{a, b, c\}, \{b, c, d\}, \{a, c, d\}, \{a, b, d\}, \{b, e, f\}, \{e, g\}\}$$
over the set $X = \{a, b, c, d, e, f, g\}$, already seen in Example II.5. On the right-hand side, we have the simplicial representation of the cover $\mathcal{M}^{(1)}$ of the first line scenario.

[1]This special case will never be used in practice, as clarified later in Remark IV.4.

Notice that, despite the simplicial complex of the original scenario having dimension 2,

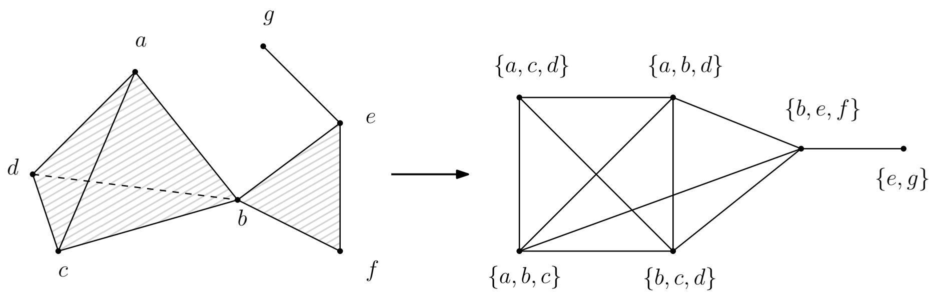


FIGURE IV.2. A measurement scenario (left) and its first line version (right). Note that the tetrahedron $\{a, b, c, d\}$ is hollow.

the complex of the first line scenario is a graph. This is due to the fact that the first line scenario describes the 1-simplices of the nerve of the original scenario, and thus has dimension 1. For this reason, all line scenarios are represented by a graph.

We will often need to repeat the procedure of modifying the original scenario into its line version. This leads to the following recursive definition:

DEFINITION IV.3. Let $\langle X, \mathcal{M}, (O_m)\rangle$ be a measurement scenario. For any $k \in \mathbb{N}$, we define the **k-th line scenario** of $\langle X, \mathcal{M}, (O_m)\rangle$, denoted by $\langle X, \mathcal{M}, (O_m)\rangle^{(k)}$, as

$$\langle X, \mathcal{M}, (O_m)\rangle^{(0)} := \langle X, \mathcal{M}, (O_m)\rangle ,$$

if $k = 0$, and

$$\langle X, \mathcal{M}, (O_m)\rangle^{(k)} := \left(\langle X, \mathcal{M}, (O_m)\rangle^{(k-1)}\right)^{(1)} ,$$

if $k \geq 1$. Proposition IV.2 ensures that all the higher-level line scenarios are well-defined.

The origin of the name *line* we use here is due to the fact that if we consider our construction from a purely combinatorial standpoint, it appears to be a generalisation of the *line graph* construction of graph theory [**Whi32, Kra43, HN60**] to the level of simplicial complexes, as pointed out by Roberson [**Rob18**].[2] The generalisation only involves the first line scenario, where we transition from an abstract simplicial complex describing $\mathcal{M}$ to a graph describing $\mathcal{M}^{(1)}$. For any $k \geq 2$, the two operations coincide: the graph representing $\mathcal{M}^{(k)}$ is exactly the line graph of the graph representing $\mathcal{M}^{(k-1)}$. Although this connection will not be explored much further in this thesis, it certainly deserves additional investigation. In particular, it would be interesting to know whether the existing computational methods for line graphs, e.g. the linear time algorithms for line graph recognition of Roussopoulos [**Rou73**] and Lehot [**Leh74**], can be effectively used to better understand contextuality.

We conclude this section with the following remark, which justifies slightly stricter assumptions on the measurement cover of first line scenarios.

[2] In fact, earlier versions of our work used the term *joint* instead. We decided to change the name in light of the connection with graph theory.

Remark IV.4. Consider a scenario $\langle X, \mathcal{M}, (O_m)\rangle$. If there exists a k such that $\mathcal{M}^{(k)}$ contains a single context, then one can show that this necessarily implies that $\langle X, \mathcal{M}, (O_m)\rangle$ is acyclic in the database-theoretic sense of Section 5.4 of Chapter II. By Vorob'ev's theorem (Theorem II.21), we know that we cannot witness contexutal behaviour in such a scenario. Therefore, from now on, we will always assume $|\mathcal{M}^{(k)}| \geq 2$ for all $k \geq 0$.

2.3. Line models. In this section we introduce the concept of *line model* and we analyse some of its basic properties.

Definition IV.5. Let $\mathcal{S}$ be an empirical model on a scenario $\langle X, \mathcal{M}, (O_m)\rangle$. The **first line model** $\mathcal{S}^{(1)}$ is an empirical model on the first line scenario $\langle X, \mathcal{M}, (O_m)\rangle^{(1)}$, defined as follows: for all $\mathcal{U} \subseteq X^{(1)}$,

$$\mathcal{S}^{(1)}(\mathcal{U}) := \left\{ (s_C)_{C\in\mathcal{U}} \in \prod_{C\in\mathcal{U}} \mathcal{S}(C) \;\middle|\; s_C|_{C\cap C'} = s_{C'}|_{C\cap C'} \;\forall C, C' \in \mathcal{U} \right\}.$$

The restriction maps are inherited from $\mathcal{E}^{(1)}$.

Note that, in particular, for elements $\mathscr{C} = \{C, C'\} \subseteq \mathcal{M}$ of the cover $\mathcal{M}^{(1)}$, $\mathcal{S}^{(1)}(\mathscr{C})$ coincides with the following pullback:

$$\begin{tikzcd}
\mathcal{S}^{(1)}(\mathscr{C}) \arrow[r] \arrow[d] \arrow[dr, phantom, "\lrcorner"] & \mathcal{S}(C) \arrow[d, "\rho^{C}_{C\cap C'}"] \\
\mathcal{S}(C') \arrow[r, "\rho^{C'}_{C\cap C'}"'] & \mathcal{S}(C\cap C')
\end{tikzcd}$$

Before we embark on a detailed analysis of line models, we point out an important feature of compatible sections of $\mathcal{S}^{(1)}$ which will be used throughout:

Lemma IV.6. *Consider two distinct contexts $\mathscr{C} = \{C, C'\}$, $\mathscr{D} = \{D, D'\}$ in $\mathcal{M}^{(1)}$ with non-empty intersection (w.l.o.g. we assume $C = D$). Then, two sections $s = (s_C, s_{C'}) \in \mathcal{S}^{(1)}(\mathscr{C})$ and $t = (t_D, t_{D'}) \in \mathcal{S}^{(1)}(\mathscr{D})$ are compatible if and only if $s_C = t_D$. In particular, a compatible family for $\mathcal{S}^{(1)}$ cannot contain two different local sections of $\mathcal{S}$ over the same context of $\mathcal{M}$.*

Proof. We have

$$s|_{\mathscr{C}\cap\mathscr{D}} = (s_C, s_{C'})|_{\mathscr{C}\cap\mathscr{D}} = (s_C, s_{C'})|_{\{C\}} = s_C,$$
$$t|_{\mathscr{C}\cap\mathscr{D}} = (t_D, t_{D'})|_{\mathscr{C}\cap\mathscr{D}} = (t_D, t_{D'})|_{\{D\}} = t_D.$$

Hence $s|_{\mathscr{C}\cap\mathscr{D}} = t|_{\mathscr{C}\cap\mathscr{D}}$ if and only if $s_C = t_D$. $\square$

The following proposition shows that the first line model is well defined.

Proposition IV.7. *Let $\mathcal{S}$ be an empirical model. Then, $\mathcal{S}^{(1)}$ is a well-defined empirical model.*

Proof. First of all, note that $\mathcal{S}^{(1)}$ is a subpresheaf of $\mathcal{E}^{(1)}$. Indeed,

$$\mathcal{S}^{(1)}(\mathcal{U}) \subseteq \prod_{C\in\mathcal{U}} \mathcal{S}(C) \subseteq \prod_{C\in\mathcal{U}} \mathcal{E}(C) = \mathcal{E}^{(1)}(\mathcal{U}).$$

Now, we need to verify conditions 1, 2 and 3 of Definition II.12.

(1) Let $\mathscr{C} = \{C, C'\} \in \mathcal{M}^{(1)}$. Because $\mathcal{S}$ is an empirical model, we know that $\mathcal{S}(C) \neq \emptyset$, given that $C \in \mathcal{M}$. Let $s_C \in \mathcal{S}(C)$. Since $\mathcal{S}$ is flasque beneath the cover, and because $C \cap C' \subseteq C' \in \mathcal{M}$, the restriction map $\rho^{C'}_{C\cap C'} : \mathcal{S}(C') \to \mathcal{S}(C \cap C')$ is surjective. Therefore, there exists $s_{C'} \in \mathcal{S}(C')$ such that

$$\rho^{C'}_{C\cap C'}(s_{C'}) = s_{C'}|_{C\cap C'} = s_C|_{C\cap C'}.$$

Hence, $(s_C, s_{C'}) \in \mathcal{S}^{(1)}(\{C, C'\})$.

(2) Let $\mathcal{U} \subseteq \mathcal{U}' \subseteq \mathscr{C}$ for some context $\mathscr{C} = \{C, C'\} \in \mathcal{M}^{(1)}$. If $\mathcal{U}$ or $\mathcal{U}'$ are empty, then the condition is trivially verified as $\rho^{\mathcal{U}}_{\emptyset}$ is obviously surjective. If $\mathcal{U} = \mathcal{U}'$ then $\rho^{\mathcal{U}'}_{\mathcal{U}}$ is the identity, which is also surjective. The only non–trivial case arises when $\mathcal{U} \subsetneq \mathcal{U}' = \mathscr{C}$. W.l.o.g., suppose $\mathcal{U} = \{C\}$ (the other case being $\mathcal{U} = \{C'\}$), and let $s_C \in \mathcal{S}(C)$. Because $\mathcal{S}$ is flasque beneath the cover, the restriction map $\rho^{C'}_{C\cap C'} : \mathcal{S}(C') \to \mathcal{S}(C \cap C')$ is surjective. Hence, there exists an $s_{C'} \in \mathcal{S}(C')$ such that $s_{C'}|_{C\cap C'} = s_C|_{C\cap C'}$. Thus, $(s_C, s_{C'}) \in \mathcal{S}^{(1)}(\mathscr{C})$, and

$$(s_C, s_{C'})|_{\mathcal{U}} = (s_C, s_{C'})|_{\{C\}} = s_C,$$

which shows that $\rho^{\mathcal{U}'}_{\mathcal{U}}$ is surjective.

(3) Let $F := \{(s_C, s_{C'})_{\mathscr{C}}\}_{\mathscr{C}\in\mathcal{M}^{(1)}}$ be a compatible family[3] for $\mathcal{S}^{(1)}$, which means that $(s_C, s_{C'})_{\mathscr{C}} \in \mathcal{S}^{(1)}(\mathscr{C})$ for all $\mathscr{C} \in \mathcal{M}^{(1)}$, and

$$(s_C, s_{C'})|_{\mathscr{C}\cap\mathscr{D}} = (s_D, s_{D'})|_{\mathscr{C}\cap\mathscr{D}}$$

for all $\mathscr{C} = \{C, C'\}$ and $\mathscr{D} = \{D, D'\}$ in $\mathcal{M}^{(1)}$. The family F induces the global section

$$g := (s_C)_{C\in\mathcal{M}} \in \prod_{C\in\mathcal{M}} \mathcal{S}(C) \subseteq \mathcal{S}^{(1)}\left(X^{(1)}\right),$$

which is well-defined by Lemma IV.6. The fact that $g_C|_{C\cap C'} = g_{C'}|_{C\cap C'}$ for all $C, C' \in \mathcal{M}$ is trivially verified given that $g_C = s_C$ and $(s_C, s_{C'}) \in \mathcal{S}^{(1)}(\{C, C'\})$. □

We will often need to repeat the procedure of taking the first line model, which leads to the following definition, reminiscent of Definition IV.3.

Definition IV.8. Let $\mathcal{S}$ be an empirical model on a scenario $\langle X, \mathcal{M}, (O_m)\rangle$. For any $k \in \mathbb{N}$, we define the **k-th line model of $\mathcal{S}$**, denoted by $\mathcal{S}^{(k)}$, as follows: if $k = 0$, then $\mathcal{S}^{(0)} := \mathcal{S}$. If $k \geq 1$, then

$$\mathcal{S}^{(k)} := \left(\mathcal{S}^{(k-1)}\right)^{(1)}.$$

[3]A more precise notation would be $\left\{(s_C, s_{C'})_{\{C,C'\}}\right\}_{\{C,C'\}\in\mathcal{M}^{(1)}}$, but we will often use the one we adopt here for the sake of simplicity.

Proposition IV.7 guarantees that all the higher-level line models are well-defined empirical models.

We end this section with two important remarks.

REMARK IV.9. Consider an empirical model $\mathcal{S}$ on a scenario $\langle X, \mathcal{M}, (O_m)\rangle$. There is a key subtlety in the definition of the line models of $\mathcal{S}$ which we will exploit in some of the proofs in Section 4. Let $C \in \mathcal{M}^{(k-1)}$. The subtlety consists in the following equality, which simply follows by definition:

$$\mathcal{S}^{(k)}(\{C\}) = \mathcal{S}^{(k-1)}(C). \tag{IV.1}$$

Consider two distinct contexts $\mathscr{C} = \{C, C'\}, \mathscr{D} = \{D, D'\} \in \mathcal{M}^{(k)}$, such that $\mathscr{C} \cap \mathscr{D} \neq \emptyset$. W.l.o.g. we suppose $C = D$, so that $\mathscr{C} \cap \mathscr{D} = \{C\}$. Suppose we have a section $s_{\mathscr{C}} \in \mathcal{S}^{(k)}(\mathscr{C})$. Then, because of (IV.1), the restricted section

$$s_{\mathscr{C}}|_{\mathscr{C}\cap\mathscr{D}} = s_{\mathscr{C}}|_{\{C\}}$$

can be seen both as an element of $\mathcal{S}^{(k)}(\{C\})$, or, equivalently, as an element of $\mathcal{S}^{(k-1)}(C)$. In the latter case, we will denote the restricted section as

$$s_{\mathscr{C}_1}|_C \in \mathcal{S}^{(k-1)}(C).$$

So, to summarise, $s_{\mathscr{C}}|_{\{C\}} \in \mathcal{S}^{(k)}(\{C\})$, while $s_{\mathscr{C}_1}|_C \in \mathcal{S}^{(k-1)}(C)$.

REMARK IV.10. Let $\mathcal{S}$ be an empirical model on a scenario $\langle X, \mathcal{M}, (O_m)\rangle$. By definition, the possible sections of $\mathcal{S}^{(1)}$ are pairs of sections of $\mathcal{S}$. Similarly, sections of $\mathcal{S}^{(2)}$ are *pairs of pairs* of sections of $\mathcal{S}$. In general, sections of $\mathcal{S}^{(k)}$ are *pairs of pairs ... of pairs* (k times) of sections of $\mathcal{S}$. For our purposes, given a section s of $\mathcal{S}$, we will need to list those sections of $\mathcal{S}^{(k)}$ that *contain* s. To do this, we will use the flatten function, whose name is borrowed from popular programming languages. This function takes a section $t_{\mathscr{C}} \in \mathcal{S}^{(k)}(\mathscr{C})$ (which is a pair of pairs ... of pairs (k times) of sections of $\mathcal{S}^{(k)}$) as argument and returns a single set containing all the sections of $\mathcal{S}^{(k)}$ that appear in $t_{\mathscr{C}}$. For instance, for $k = 3$, we have

$$\mathsf{flatten}\,[(((s_1, s_2), (s_3, s_4)), ((s_5, s_6), (s_7, s_8)))] = \{s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8\}$$

2.4. Interpretation and examples. In Section 2.1, we briefly sketched our strategy to avoid false negatives, which consists of considering multiple compatible local sections at the same time, instead of focusing on a single one. The notion of line model embodies precisely this idea. As discussed above, local sections of the first line model of an empirical model $\mathcal{S}$ are pairs of sections of $\mathcal{S}$ above adjacent contexts. This allows one to 'force' the selection of the sections on adjacent contexts in the original model, thus reducing the chances of the existence of a false negative. Higher-level line models further refine this approach and allow to consider three, four, k compatible sections at the same time. These statements will be made precise in Section 4, but it is worth giving some examples that will guide us through the technical results.

Let us start by illustrating the first line model of the Hardy model, displayed in Table II.8. Recall that $X = \{a_1, a_2, b_1, b_2\}$, $\mathcal{M} = \{\{a_1, b_1\}, \{a_1, b_2\}, \{a_2, b_1\}, \{a_2, b_2\}\}$ and $O_m = \{0, 1\}$ for all $m \in X$. Let $C_1 := \{a_1, b_1\}$, $C_2 := \{a_1, b_2\}$, $C_3 := \{a_2, b_1\}$ and $C_4 := \{a_2, b_2\}$. Then we have $X^{(1)} = \mathcal{M}$ and

$$\mathcal{M}^{(1)} = \{\{C_1, C_2\}, \{C_2, C_4\}, \{C_3, C_4\}, \{C_1, C_3\}\}.$$

In Table IV.2, we recall the enumeration of the sections of the Hardy model already presented in Table III.1.

TABLE IV.2. An enumeration of the possible sections of the Hardy model.

A	B	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
a_1	b_1	s_1	s_2	s_3	s_4
a_1	b_2		s_5	s_6	s_7
a_2	b_1		s_8	s_9	s_{10}
a_2	b_2	s_{11}	s_{12}	s_{13}	

Thanks to this notation, we can represent the planar bundle diagram for the first line model of the Hardy model in Figure IV.3.

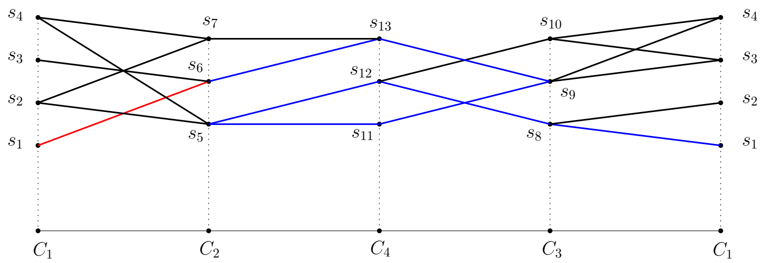


FIGURE IV.3. The first line model of the Hardy model. In red, the only section containing s_1 in the context $\{C_1, C_2\}$. In blue, a cohomology loop containing s_1.

Compare this to Figure III.5, where we highlighted in red section s_1, which is not part of any compatible family. Consider the context $\mathscr{C} = \{C_1, C_2\}$: the only section at $\mathscr{C}$ of the first line model containing s_1 is (s_1, s_6), marked in red in Figure IV.3. Notice that this section is not part of any compatible family in the line model either. In Figure III.5, we provided a cohomology loop containing s_1, which is responsible for the existence of a false negative. Note that, in the case of the first line model, it is no longer possible to create 'Z' shaped paths above a single context (this fact is not a coincidence, as we will se in Lemma IV.29 and more generally in Theorem IV.31), however, it is still possible to find a more complex cohomology loop containing (s_1, s_6), namely

$$\{(s_1, s_6), (s_6, s_{13}) - (s_5, s_{11}) + (s_5, s_{12}), (s_{13}, s_9) - (s_{11}, s_9) + (s_{12}, s_8), (s_8, s_1)\},$$

which is highlighted in blue in Figure IV.3.

Let us now consider the model of Table IV.1. In Table IV.3, we give an enumeration of its possible sections. With this enumeration, we illustrate the first line model as a planar bundle diagram in Figure IV.4. We have already shown that s_2 is not part of any compatible family, but it is part of a cohomology loop, which gives rise to a false negative. In the line model, the only section over $\{C_1, C_2\}$ containing s_2 is (s_2, s_5). Note that not only (s_2, s_5) is not part of any compatible family, but it appears not to be part of any cohomology loop either. We have successfully removed the 'Z' path responsible for the false negative. This fact accurately reflects the discussion on this model carried

TABLE IV.3. An enumeration of the possible sections of the model IV.1.

A	B	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
a_1	b_1	s_1			s_2
a_1	b_2	s_3		s_4	s_5
a_2	b_1	s_6			s_7
a_2	b_2		s_8	s_9	

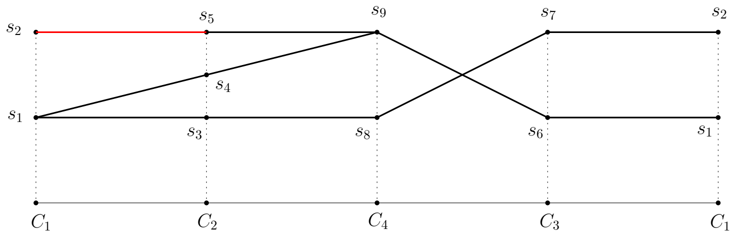


FIGURE IV.4. The first line model of the model given by Table IV.1. In red, the section S

out in Section 2.1. By imposing the joint selection of s_2 *and* s_5, we have successfully removed the false negative. A formal proof of this fact will be given in Section 4.

It is fairly easy to see that, because the first line model is logically contextual at (s_2, s_5), which is the only section containing s_2, the underlying model is logically contextual at s_2. However, the relation between the contextual properties of line models and the original ones may not be immediately clear. We will give all the details about this question in the following section.

3. The contextuality of line models

By looking at the examples of the previous sections, a natural question to ask is: what conclusions can we draw on an empirical model by looking at the contextuality properties of its line models? The answer is given by the following results.

PROPOSITION IV.11. *Consider an empirical model $\mathcal{S}$ on a measurement scenario $\langle X, \mathcal{M}, (O_m)\rangle$. Let $C \in \mathcal{M}$ and $s \in \mathcal{S}(C)$. The following are equivalent.*

(1) The model $\mathcal{S}$ is logically contextual at s, i.e. $\mathsf{LC}(\mathcal{S}, s)$.
(2) There exists a $C' \in \mathcal{M}$ with $C \neq C'$ and $C \cap C' \neq \emptyset$ such that, for all $t \in \mathcal{S}(C')$ verifying $t|_{C\cap C'} = s|_{C\cap C'}$, we have $\mathsf{LC}\left(\mathcal{S}^{(1)}, (s,t)\right)$
(3) For all $C' \in \mathcal{M}$ with $C \neq C'$ and $C \cap C' \neq \emptyset$, for all $t \in \mathcal{S}(C')$ verifying $t|_{C\cap C'} = s|_{C\cap C'}$, we have $\mathsf{LC}\left(\mathcal{S}^{(1)}, (s,t)\right)$.

PROOF. The fact that (3) implies (2) is trivial.

- (2) $\Rightarrow$(1): Suppose $\neg\mathsf{LC}(\mathcal{S}, s)$. Then, there exists a family $F := \{s_C \in \mathcal{S}(C)\}_{C\in\mathcal{M}}$, compatible for $\mathcal{S}$, such that $s_C = s$. We want to show that, for all $C' \in \mathcal{M}$ with $C \cap C' \neq \emptyset$, there exists $t \in \mathcal{S}(C')$ verifying $t|_{C\cap C'} = s|_{C\cap C'}$ such that $\neg\mathsf{LC}\left(\mathcal{S}^{(1)}, (s,t)\right)$.

Consider the family

$$F' := \Big\{(s_K, s_{K'}) \in \mathcal{S}^{(1)}(\{K, K'\})\Big\}_{\{K,K'\}\in\mathcal{M}^{(1)}}.$$

This family is well-defined: $(s_K, s_{K'})$ is indeed in $\mathcal{S}^{(1)}(\{K, K'\})$ by compatibility of F. Moreover, it is compatible for $\mathcal{S}^{(1)}$ by its own definition and Lemma IV.6, since we chose exactly one section for each context of $\mathcal{M}$. Let $C' \in \mathcal{M}$ with $C \cap C' \neq \emptyset$, and consider $t := s_{C'} \in \mathcal{S}(C')$. Then $(s, t) = (s_C, s_{C'}) \in F'$, which proves that $t|_{C\cap C'} = s|_{C\cap C'}$ (as $t = s_{C'}$ and $s = s_C$). Moreover, $\neg\mathsf{LC}\left(\mathcal{S}^{(1)}, (s, t_{C'})\right)$, since $t = s_{C'}$ is part of the compatible family F.

- (1) $\Rightarrow$(3): Suppose there exists a C' with $C \cap C' \neq \emptyset$, such that there exists a $t \in \mathcal{S}(C')$, with $t|_{C\cap C'} = s|_{C\cap C'}$, verifying $\neg\mathsf{LC}\left(\mathcal{S}^{(1)}, (s, t)\right)$. This means that there exists a family

$$F := \Big\{(v_K, v_{K'}) \in \mathcal{S}^{(1)}(\{K, K'\})\Big\}_{\{K,K'\}\in\mathcal{M}^{(1)}},$$

compatible for $\mathcal{S}^{(1)}$, such that $(v_C, v_{C'}) = (s, t)$. Consider the family $F' := \{v_K \in \mathcal{S}(K)\}_{K\in\mathcal{M}}$. This family contains precisely one local section for each context of $\mathcal{M}$ by connectedness of the cover. Moreover, each such global section is well-defined by compatibility of F (see Lemma IV.6). Now, note that F' is a compatible family for $\mathcal{S}$. Indeed, given $K, K' \in \mathcal{M}$, because $(v_K, v_{K'}) \in \mathcal{S}^{(1)}(\{K, K'\})$, we must have $v_K|_{K\cap K'} = v_{K'}|_{K\cap K'}$. Therefore, because $v_C = s$, the section s is contained in the compatible family F', proving that $\neg\mathsf{LC}\,(\mathcal{S}, s)$.

□

This proposition provides a characterisation of how logical contextuality propagates through the line scenarios. This result has a much more elegant form when we turn our attention to strong contextuality. Indeed, it turns out that strong contextuality is completely preserved through the construction of line models:

Corollary IV.12. *Let $\mathcal{S}$ be an empirical model on a scenario $\langle X, \mathcal{M}, (O_m)\rangle$. Then $\mathcal{S}$ is strongly contextual if and only if $\mathcal{S}^{(1)}$ is strongly contextual.*

Proof. Recall that strong contextuality is equivalent to logical contextuality for all the sections of the model. Suppose $\mathsf{SC}\,(\mathcal{S})$. Let $C, C' \in \mathcal{M}$ be any two distinct contexts such that $C \cap C' \neq \emptyset$. Let $(s_C, t_{C'}) \in \mathcal{S}^{(1)}(\{C, C'\})$ be an arbitrary section. We want to show that $\mathcal{S}^{(1)}$ is logically contextual at $(s_C, t_{C'})$. Since $\mathcal{S}$ is strongly contextual, it is in particular logically contextual at s_C. By Proposition IV.11, this implies that, for all $C' \in \mathcal{M}$ with $C \cap C' \neq \emptyset$, for all $t \in \mathcal{S}(C')$ verifying $t|_{C\cap C'} = s_C|_{C\cap C'}$, we have $\mathsf{LC}\left(\mathcal{S}^{(1)}, (s_C, t)\right)$. In particular, if we take $t := t_{C'}$, we have $\mathsf{LC}\left(\mathcal{S}^{(1)}, (s_C, t_{C'})\right)$.

For the converse, suppose $\mathsf{SC}\left(\mathcal{S}^{(1)}\right)$. Let $C \in \mathcal{M}$ and take an arbitrary section $s \in \mathcal{S}(C)$. Let $C' \in \mathcal{M}$ such that $C \cap C' \neq \emptyset$. Because $\mathsf{SC}(\mathcal{S}^{(1)})$, we know that for all $t \in \mathcal{S}(C')$ verifying $t|_{C\cap C'} = s|_{C\cap C'}$, we have $\mathsf{LC}\left(\mathcal{S}^{(1)}, (s, t)\right)$. By Proposition IV.11, we conclude that $\mathsf{LC}\,(\mathcal{S}, s)$. □

Proposition IV.11 motivates the following definition.

Definition IV.13. Let $\mathcal{S}$ be an empirical model on a scenario $\langle X, \mathcal{M}, (O_m)\rangle$. Let $s \in \mathcal{S}(C)$ be a local section at some context $C \in \mathcal{M}$, and $k \geq 1$. We say that $\mathcal{S}$ is $\mathsf{LC}^{(k)}$ at s, and write $\mathsf{LC}^{(k)}(\mathcal{S}, s)$, if we have $\mathsf{LC}\left(\mathcal{S}^{(k)}, t\right)$ for all local sections t of $\mathcal{S}^{(k)}$ such that $s \in \mathsf{flatten}(t)$ (cf. Remark IV.10).

By applying simple inductive arguments, we immediately have the following additional corollaries of Proposition IV.11:

Corollary IV.14. *Let $\mathcal{S}$ be an empirical model on a scenario $\langle X, \mathcal{M}, (O_m)\rangle$. Let $C \in \mathcal{M}$ and $s \in \mathcal{S}(C)$. Then the following are equivalent:*

(1) $\mathcal{S}$ is logically contextual at s, i.e. $\mathsf{LC}(\mathcal{S}, s)$.
(2) There exists a $k \geq 1$ such that $\mathsf{LC}^{(k)}(\mathcal{S}, s)$
(3) $\mathsf{LC}^{(k)}(\mathcal{S}, s)$ for all $k \geq 1$.

Corollary IV.15. *Let $\mathcal{S}$ be an empirical model on a scenario $\langle X, \mathcal{M}, (O_m)\rangle$. The following are equivalent:*

(1) $\mathcal{S}$ is strongly contextual.
(2) There exists a $k \geq 1$ such that $\mathcal{S}^{(k)}$ is strongly contextual
(3) $\mathcal{S}^{(k)}$ is strongly contextual for all $k \geq 0$

Note that there is no need to extend Definition IV.13 to strong contextuality as this would be equivalent to regular strong contextuality by Corollary IV.15.

4. Cyclic models and their properties

Before we prove the main results of the chapter, we will need to introduce the notions of *path*, *cycle*, and *cyclic model*, and thoroughly inspect their properties. We start with an important remark:

Remark IV.16. Let $\langle X, \mathcal{M}, (O_m)\rangle$ be a measurement scenario. By definition, for each $k \geq 1$, the contexts of $\mathcal{M}^{(k)}$ are sets of contexts of $\mathcal{M}^{(k-1)}$. In order to avoid confusion between the contexts of $\mathcal{M}^{(k)}$ and those of $\mathcal{M}^{(k-1)}$ we will denote them using different calligraphic styles. The typical hierarchy we will use is the following:

$$c \in \mathcal{M}^{(k-2)} \to C \in \mathcal{M}^{(k-1)} \to \mathscr{C} \in \mathcal{M}^{(k)} \to \mathfrak{C} \in \mathcal{M}^{(k+1)}.$$

Note that the hierarchy will always be the same, but we will *not* fix a calligraphic style for a specific k, as we will have to deal with many different cases.

4.1. Paths and cycles. Let us inspect some of the properties of line scenarios. We start by introducing the notions of *path* and *cycle*.

Definition IV.17. Let $\langle X, \mathcal{M}, (O_m)\rangle$ be a measurement scenario and $n, k \geq 1$. An **n-path for $\mathcal{M}^{(k)}$** is a set $\mathscr{D} := \{C_1, \ldots, C_{n+1}\} \subseteq \mathcal{M}^{(k-1)}$ of $n+1$ distinct contexts of $\mathcal{M}^{(k-1)}$ such that $C_i \cap C_{i+1} \neq \emptyset$ for all $1 \leq i \leq n$. It is called an **$(n+1)$-cycle** if, in addition, $C_{n+1} \cap C_1 \neq \emptyset$. An n-path $\mathscr{D}$ is called *chordal* if there exist two non-consecutive indices i, j, with $\{i, j\} \neq \{1, n+1\}$, such that $C_i \cap C_j \neq \emptyset$.

We can think of an n-path for $\mathcal{M}^{(k)}$ as a sequence of distinct vertices in the graph generated by $\mathcal{M}^{(k)}$. This corresponds to the graph-theoretic notion of *simple path*.

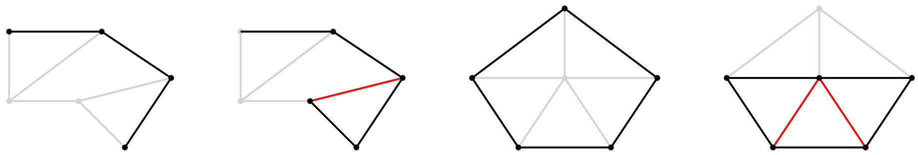


FIGURE IV.5. Different types of paths for $\mathcal{M}^{(k)}$. The grey graph represents $\mathcal{M}^{(k)}$. From left to right: a chordless 4-path, a chordal 4-path (chord highlighted in red), a chordless 5-cycle and a chordal 5-cycle (chords highlighted in red).

Similarly, (chordal) cycles for $\mathcal{M}^{(k)}$ correspond to *(chordal) simple cycles* in graph theory. In Figure IV.5, we give some graphical examples.

REMARK IV.18. In graph theory, a simple path can be equivalently described by the sequence of edges connecting the vertices. Similarly, an n-path $\mathscr{D}_\bullet = \{C_1, \dots, C_{n+1}\} \subseteq \mathcal{M}^{(k-1)}$ for $\mathcal{M}^{(k)}$ can be specified by the set

$$\underline{\mathfrak{D}} = \{\{C_1, C_2\}, \{C_2, C_3\}, \dots, \{C_n, C_{n+1}\}\} \subseteq \mathcal{M}^{(k)},$$

containing contexts of $\mathcal{M}^{(k)}$, i.e. edges of the graph generated by $\mathcal{M}^{(k)}$. The set $\underline{\mathfrak{D}}$ will be referred to as the **edge representation** of the path $\mathscr{D}_\bullet$. To avoid confusion, from now on, we will denote $\mathscr{D}_\bullet$ for the vertex representation and $\underline{\mathfrak{D}}$ for the edge representation. Notice that, while an n-path in vertex representation contains $n+1$ elements, an n-path in edge representation only contains n elements. If $\mathscr{D}_\bullet$ is an $(n+1)-$*cycle*, then its edge representation is

$$\underline{\mathfrak{D}} = \{\{C_1, C_2\}, \{C_2, C_3\}, \dots, \{C_n, C_{n+1}\}, \{C_1, C_{n+1}\}\} \subseteq \mathcal{M}^{(k)}.$$

In this case, both the vertex and the edge representation contain $(n+1)$ elements.

4.1.1. Special properties of 3-cycles. Cycles for $\mathcal{M}^{(k)}$ of size 3 present some peculiarities that deserve to be discussed in detail in order to avoid confusion. The reason is that, although they are technically chordless, one of their edges could be seen as a chord connecting the remaining two. A key aspect of chordless n-cycles for $\mathcal{M}^{(k)}$ of size $n \geq 4$, which will be proved in Proposition IV.23, is that they must be generated by n-cycles for $\mathcal{M}^{(k-1)}$. This is not generally true for 3-cycles. Indeed, we could potentially have a 3-cycle $\mathscr{D}_\bullet = \{C_1, C_2, C_3\}$ for $\mathcal{M}^{(k)}$ which is generated by a star-shaped configuration of the C_i's, seen as edges of $\mathcal{M}^{(k-1)}$, as shown in Figure IV.6.

We will refer to this kind of 3-cycles as **non-proper 3 cycles for $\mathcal{M}^{(k)}$**. Any cycle which is not a non-proper 3 cycle will be referred to as **proper**. A proper 3-cycle for $\mathcal{M}^{(k)}$ is pictured in Figure IV.7.

At the end of Section 2.2, we argued that line scenarios can be seen as a generalisation of line graphs. In light of this connection, the peculiarities of 3-cycles we have just described are not surprising. Indeed, Whitney's isomorphism theorem [**Whi32**] shows that any two connected graphs with isomorphic line graphs are isomorphic, *with the exception* of the 3-cycle and the star graph on 4 vertices. In general, some of the results of the following subsection can also be recovered as special cases of Whitney's isomorphism theorem. However, we shall not forget that the graphs describing line scenarios carry a

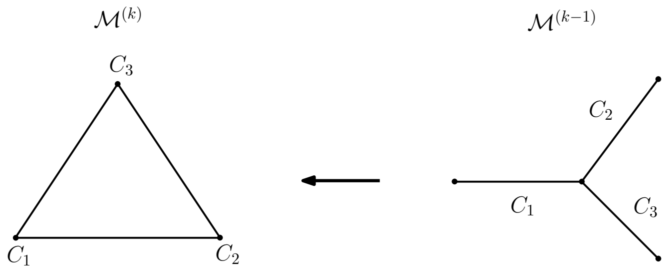


FIGURE IV.6. A non-proper 3-cycle $\mathscr{D}_\bullet = \{C_1, C_2, C_3\}$ for $\mathcal{M}^{(k)}$.

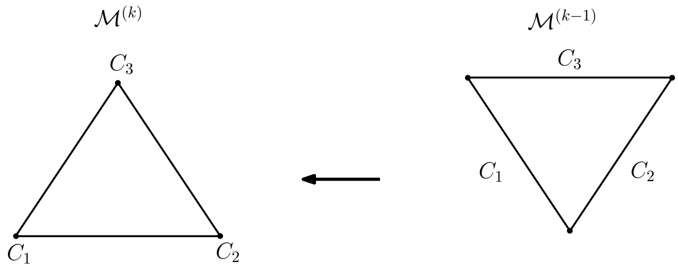


FIGURE IV.7. A proper 3-cycle $\mathscr{D}_\bullet = \{C_1, C_2, C_3\}$ for $\mathcal{M}^{(k)}$.

richer structure given by measurements and contexts, which is important to retain for our purposes. This motivates proving these results independently from graph theory.

4.2. Fundamental properties of paths and cycles. In this section, we will present some key properties of paths and cycles that will play a crucial role in the proofs of the main results of the chapter. The first proposition, for instance, shows that paths and cycles are preserved when taking the line version of a scenario, in the sense that they naturally give rise to paths and cycles in the new scenario.

PROPOSITION IV.19. *Let $\langle X, \mathcal{M}, (O_m)\rangle$ be a measurement scenario. Let $k \geq 1$, and let $\underline{\mathscr{D}} := \{C_1, \ldots, C_n\} \subseteq \mathcal{M}^{(k)}$ be an n-path for $\mathcal{M}^{(k)}$, where $n \geq 2$. Then, the set*

$$\underline{\mathfrak{D}} = \{\mathcal{K}_1, \ldots, \mathcal{K}_{n-1}\} := \{\{C_1, C_2\}, \{C_2, C_3\}, \ldots, \{C_{n-1}, C_n\}\}$$

is a chordless $(n-1)$-path for $\mathcal{M}^{(k+1)}$. Moreover, if $\underline{\mathscr{D}}$ is a cycle, then

$$\underline{\mathfrak{D}} = \{\mathcal{K}_1, \ldots, \mathcal{K}_n\} := \{\{C_1, C_2\}, \{C_2, C_3\}, \ldots, \{C_{n-1}, C_n\}, \{C_n, C_1\}\}$$

is a chordless n-cycle for $\mathcal{M}^{(k+1)}$.

PROOF. The elements of $\underline{\mathfrak{D}}$ are all distinct because the elements of $\underline{\mathscr{D}}$ are all distinct. Moreover, $\underline{\mathfrak{D}} \subseteq \mathcal{M}^{(k)}$ because $C_i \cap C_{i+1} \neq \emptyset$ for all $1 \leq i \leq n-1$. If $n = 2$, then clearly $\underline{\mathfrak{D}} = \{\mathcal{K}_1\}$ is a 1-path for $\mathcal{M}^{(k+1)}$. Now, assume $n > 2$. We have $\mathcal{K}_i \cap \mathcal{K}_{i+1} = \{C_i, C_{i+1}\} \cap \{C_{i+1}, C_{i+2}\} = \{C_{i+1}\} \neq \emptyset$ for all $1 \leq i \leq n-2$, which shows that $\underline{\mathfrak{D}}$ is an $(n-1)$-path. If $\underline{\mathscr{D}}$ is a cycle, we also have $\mathcal{K}_n \cap \mathcal{K}_1 = \{C_n, C_1\} \cap \{C_1, C_2\} = \{C_1\} \neq \emptyset$, which proves that $\underline{\mathfrak{D}}$ is an n-cycle. To prove that $\underline{\mathfrak{D}}$ is chordless, suppose by contradiction that there exist two non-consecutive indices $1 \leq i, j \leq n-1$ (if $\underline{\mathscr{D}}$ is a cycle we suppose

$1 \leq i, j \leq n$ and $\{i, j\} \neq \{1, n\}$), such that $\mathcal{K}_i \cap \mathcal{K}_j \neq \emptyset$. Then $\{C_i, C_{i+1}\} \cap \{C_j, C_{j+1}\} \neq \emptyset$, which contradicts the fact that the C_i's are all distinct. □

We define the notion of *cyclic scenario.*

DEFINITION IV.20. A measurement scenario $\langle X, \mathcal{M}, (O_m) \rangle$ is called **cyclic** if $\mathcal{M}^{(1)}$ is a chordless cycle (in edge representation).

REMARK IV.21. In Section 5.4 of Chapter II, we introduced the notion of *acyclicity* of a measurement scenario. This concept should not be interpreted as the negation of the notion of cyclicity introduced here, as the two do not coincide. To clarify the distinction, consider the scenario of Figure IV.2. This scenario is cyclic in the Vorob'ev sense: after having performed Graham reductions by removing vertices g, e and f, we are left with an irreducible hollow tetrahedron spanned by vertices a, b, c, d. However, its first line version, also depicted in Figure IV.2, is *not* a cycle in the graph-theoretic sense. Hence the scenario is cyclic in the Vorob'ev sense, but non-cyclic in the sense introduced here.

Thanks to Proposition IV.19, we immediately have the following:

PROPOSITION IV.22. *Let $\langle X, \mathcal{M}, (O_m) \rangle$ be a cyclic scenario, and let $n := |\mathcal{M}|$. Then $\mathcal{M}^{(k)}$ is a chordless n-cycle in edge representation for all $k \geq 1$. In particular, $\langle X, \mathcal{M}, (O_m) \rangle^{(l)}$ is cyclic for all $l \geq 0$.*

Consider a measurement scenario $\langle X, \mathcal{M}, (O_m) \rangle$ and let $k \geq 2$, $n \geq 2$. Let $\mathscr{D}_\bullet := \{C_1, \ldots, C_{n+1}\} \subseteq \mathcal{M}^{(k-1)}$ be an n-path for $\mathcal{M}^{(k)}$. By definition of a path, we know that there exist $k_1, \ldots k_n \in \mathcal{M}^{(k-2)}$ such that $\{k_i\} = C_i \cap C_{i+1}$ for all $1 \leq i \leq n$. Let $k_{n+1} \in \mathcal{M}^{(k-2)}$ be such that $C_{n+1} \setminus C_n = \{k_{n+1}\}$, and let $k_0 \in \mathcal{M}^{(k-2)}$ be such that $C_1 \setminus C_2 = \{k_0\}$. Notice that, if $\mathscr{D}_\bullet$ is a proper $(n+1)$-cycle, then $k_{n+1} = k_0$. With this notation, we can prove the following proposition, which is almost a converse of Proposition IV.19: it shows that any chordless n-path (resp. proper n-cycle) in the k-th line scenario always comes from an $(n+1)$-path (resp. n-cycle) on the $(k-1)$-th line scenario.

PROPOSITION IV.23. *Let $\langle X, \mathcal{M}, (O_m) \rangle$ be a measurement scenario, let $k \geq 2$ and $2 \leq n < |\mathcal{M}^{(k-1)}|$. Let $\mathscr{D}_\bullet := \{C_1, \ldots, C_{n+1}\} \subseteq \mathcal{M}^{(k-1)}$ be a chordless n-path for $\mathcal{M}^{(k)}$. Then, the set*

$$D_\bullet = \{k_0, k_1, \ldots k_{n+1}\} \subseteq \mathcal{M}^{(k-2)}$$

is an $(n+1)$-path for $\mathcal{M}^{(k-1)}$.

From this result, it immediately follows that, if $\mathscr{D}_\bullet$ is a proper $(n+1)$-cycle, since $k_0 = k_{n+1}$, then $D_\bullet = \{k_1, \ldots, k_{n+1}\} \subseteq \mathcal{M}^{(k-2)}$ is an $(n+1)$-cycle for $\mathcal{M}^{(k-1)}$.

PROOF. Let us start by verifying that the k_i's are all distinct. Firstly, notice that $k_0 \neq k_j$ for all $1 \leq j \leq n$, since otherwise we would have $C_1 \cap C_{j+1} = \{k_0\} \neq \emptyset$, which either contradicts the fact that $\{k_0\} = C_1 \setminus C_2$ (when $j = 1$), or it contradicts the fact $\mathscr{D}_\bullet$ is chordless (when $j > 1$). One can prove that $k_{n+1} \neq k_j$ for all $1 \leq j \leq n$ in the same way. Now, suppose there are two distinct indices $1 \leq i, j \leq n$ such that $k_i = k_j$. Then, i and j are consecutive because otherwise we would have $C_i \cap C_j = \{k_i\} \neq \emptyset$, which contradicts the fact that $\mathscr{D}_\bullet$ is chordless. Thus, we only need to prove that $k_i \neq k_{i+1}$ for all $1 \leq i \leq n-1$. Suppose by contradiction there exists an i such that $k_i = k_{i+1}$. Then

$$C_i \cap C_{i+1} = \{k_i\} = \{k_{i+1}\} = C_{i+1} \cap C_{i+2},$$

which implies $C_i \cap C_{i+2} = \{k_i\} \neq \emptyset$, which contradicts the fact that $\mathscr{D}_\bullet$ is chordless. We are only left to prove that consecutive k_i's intersect. Because the k_i's are all distinct, we know that $C_i = \{k_{i-1}, k_i\}$ for all $1 \leq i \leq n+1$. Since $C_1, \ldots, C_{n+1} \in \mathcal{M}^{(k-1)}$, this implies $k_{i-1} \cap k_i \neq \emptyset$ for all $1 \leq i \leq n+1$. ☐

Remark IV.24. Proposition IV.23 allows us to establish a conventional notation. Suppose we have a chordless n-path $\mathscr{D}_\bullet := \{C_1, \ldots, C_{n+1}\}$ for $\mathcal{M}^{(k)}$, with $n \geq 1$ and $k \geq 2$. If $\mathscr{D}_\bullet$ is a cycle, we assume it is proper. Proposition IV.23 shows that we can relabel the components c_i^1, c_i^2 of each $C_i = \{c_i^1, c_i^2\}$ in such a way that $c_i^2 = c_{i+1}^1$ for all $1 \leq i \leq n$. If $\mathscr{D}_\bullet$ is a cycle, we also have $c_{n+1}^2 = c_1^1$. This notation will be used extensively in many of the proofs of this chapter. To clarify how it is constructed, we provide a graphical representation in Figure IV.8.[4]

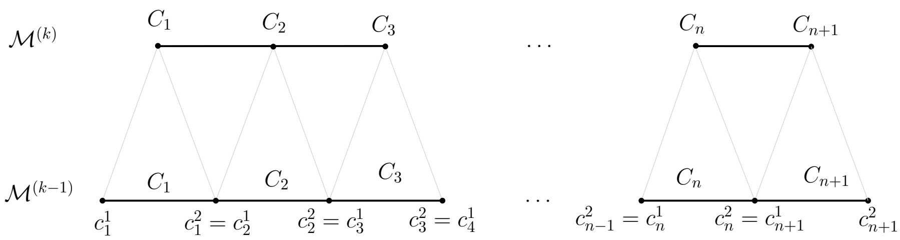


Figure IV.8. The standard notation for n-paths.

Suppose we have a cyclic scenario $\langle X, \mathcal{M}, (O_m)\rangle$, and a n-path $\underline{\mathfrak{D}} := \{\mathscr{C}_1, \ldots \mathscr{C}_n\} \subseteq \mathcal{M}^{(k)}$ for $\mathcal{M}^{(k)}$, with $n < |\mathcal{M}|$. Because $\mathcal{M}^{(1)}$ is chordless, we know by Proposition IV.22 that $\mathcal{M}^{(k)}$ is a chordless $|\mathcal{M}|$-cycle. Thus, the path $\underline{\mathfrak{D}}$ must be chordless as well, as the existence of a chord for $\underline{\mathfrak{D}}$ would imply the existence of a chord for $\mathcal{M}^{(k)}$. Therefore, we can use the notation of Remark IV.24 to formulate the following proposition, which states that, for cyclic scenarios, an n-path for $\mathcal{M}^{(k)}$ always comes from an $(n+1)$-path in $\mathcal{M}^{(k-1)}$.

Proposition IV.25. *Let $\langle X, \mathcal{M}, (O_m)\rangle$ be a cyclic scenario, and let $k \geq 2$, $2 \leq n < |\mathcal{M}|$. Let $\underline{\mathfrak{D}} := \{\mathscr{C}_1, \ldots \mathscr{C}_n\} \subseteq \mathcal{M}^{(k)}$ be an n-path for $\mathcal{M}^{(k)}$. Then the set*

$$\underline{\mathscr{D}} := \{C_1^1, C_1^2, C_2^2, \ldots, C_n^2\} \tag{IV.2}$$

is an $(n+1)$-path for $\mathcal{M}^{(k-1)}$.

Proof. Rewrite the n-path $\underline{\mathfrak{D}}$ in vertex representation: $\mathfrak{D}_\bullet = \{C_1^1, C_1^2, C_2^2, \ldots, C_n^2\}$. Since $\langle X, \mathcal{M}, (O_m)\rangle$ is cyclic, we know by Proposition IV.19 that $|\mathcal{M}^{(k-1)}| = |\mathcal{M}|$ for all $k \geq 1$. Therefore, we have $2 \leq n < |\mathcal{M}^{(k-1)}|$ and we can apply Proposition IV.23 to $\mathfrak{D}_\bullet = \{C_1^1, C_1^2, C_2^2, \ldots, C_n^2\}$ to conclude that $\mathscr{D}_\bullet := \{k_0, k_1, \ldots, k_{n+1}\} \subseteq \mathcal{M}^{(k-2)}$ is an

[4]In other words, we relabel the components of the C_i's in such a way that $k_0 = c_1^1$ and $k_i = c_i^2$ for all $1 \leq i \leq n+1$, where $k_0, \ldots k_{n+1}$ are defined as in Proposition IV.23

$(n+1)$-path for $\mathcal{M}^{(k-1)}$, where the element k_0 is such that $\{k_0\} = C_1^1 \setminus C_1^2$, the element k_{n+1} is such that $\{k_{n+1}\} = C_n^2 \setminus C_{n-1}^2$, and each k_i is such that $\{k_i\} = C_i^2 \cap C_{i+1}^2$ for all $1 \leq i \leq n-1$. Now, notice that $C_1^1 = \{k_0, k_1\}$, and $C_i^2 = \{k_{i-1}, k_i\}$ for all $2 \leq i \leq n-1$. Therefore, the edge representation of the $(n+1)$-path $\mathscr{D}_\bullet$ is in fact $\underline{\mathscr{D}}$ defined in (IV.2). $\square$

5. The cohomology of cyclic models: a complete invariant

In this section we will formalise the intuitive idea discussed at the end of Section 2.4, and generalise it to prove that we can always find a cohomological witness for contextuality in the line models of a cyclic empirical model.

5.1. Preliminaries. First of all, we need to introduce some preliminary definitions. Let $\mathcal{S}$ be an empirical model on a measurement scenario $\langle X, \mathcal{M}, (O_m)\rangle$. We will choose, as a representative for each line model $\mathcal{S}^{(k)}$, the presheaf of abelian groups

$$\mathcal{F}^{(k)} := F_{\mathbb{Z}_2}\mathcal{S}^{(k)} : \mathcal{P}(X^{(k)})^{op} \longrightarrow \mathbf{AbGrp}. \tag{IV.3}$$

We can now formulate the following definition, which is a natural extension of Definition IV.13 to account for cohomology.

Definition IV.26. Let $\mathcal{S}$ be an empirical model on a scenario $\langle X, \mathcal{M}, (O_m)\rangle$, with representative $\mathcal{F}$ as in (IV.3). Let s be a local section of $\mathcal{S}$. In view of the results of Section 3, we say that $\mathcal{S}$ **is** $\mathsf{CLC}^{(k)}$ **at** s, and write $\mathsf{CLC}^{(k)}(\mathcal{S}, s)$, if we have $\mathsf{CLC}(\mathcal{S}^{(k)}, t)$ for every local section t of $\mathcal{S}^{(k)}$ such that $s \in \mathsf{flatten}(t)$.

We can use this definition to extend Theorem II.32 to line models:

Theorem IV.27. *Let $\mathcal{S}$ be an empirical model. Given a section s of $\mathcal{S}$, if there exists a $k \geq 0$ such that $\mathsf{CLC}^{(k)}(\mathcal{S}, s)$, then $\mathsf{LC}(\mathcal{S}, s)$. Moreover, $\mathsf{CSC}(\mathcal{S}^{(k)}) \Rightarrow \mathsf{SC}(\mathcal{S})$.*

Proof. Suppose $\mathsf{CLC}^{(k)}(\mathcal{S}, s)$, i.e. $\mathsf{CLC}(\mathcal{S}^{(k)}, t)$ for every local section t of $\mathcal{S}^{(k)}$ such that $s \in \mathsf{flatten}(t)$. By Theorem II.32, it follows that $\mathsf{LC}(\mathcal{S}^{(k)}, t)$ for all t such that $s \in \mathsf{flatten}(t)$. In other words, we have $\mathsf{LC}^{(k)}(\mathcal{S}, s)$ (cf. Definition IV.13). By Corollary IV.14, this implies that $\mathcal{S}$ is logically contextual at s.

Now, suppose $\mathsf{CSC}(\mathcal{S}^{(k)})$, then, by Theorem II.32, we have $\mathsf{SC}(\mathcal{S}^{(k)})$. By Corollary IV.15 we conclude that $\mathsf{SC}(\mathcal{S})$. $\square$

We now introduce the notion of *partial family.*

Definition IV.28. Let $\mathcal{S}$ be an empirical model over a scenario $\langle X, \mathcal{M}, (O_m)\rangle$, and $n, k \geq 1$. An n**-partial family for** $\mathcal{F}^{(k)}$ is a family

$$\left\{ f_{\mathscr{C}} \in \mathcal{F}^{(k)}(\mathscr{C}) \right\}_{\mathscr{C} \in \underline{\mathfrak{D}}}$$

over an n-path $\underline{\mathfrak{D}} = \{\mathscr{C}_1, \ldots, \mathscr{C}_n\} \subseteq \mathcal{M}^{(k)}$, which is compatible for $\mathcal{F}^{(k)}$, and satisfies the following conditions:

$$f_{\mathscr{C}_1}|_{C_1^1} \in \mathcal{S}^{(k-1)}(C_1^1), \tag{IV.4}$$

$$f_{\mathscr{C}_n}|_{C_n^2} \in \mathcal{S}^{(k-1)}(C_n^2), \tag{IV.5}$$

(cf. Remark IV.24 for notation). A partial family is called *standard* if there exists a family $\{s_{\mathscr{C}} \in \mathcal{S}^{(k)}(\mathscr{C})\}_{\mathscr{C}\in\underline{\mathfrak{D}}}$, compatible for $\mathcal{S}^{(k)}$ such that

$$s_{\mathscr{C}_1}|_{C_1^1} = f_{\mathscr{C}_1}|_{C_1^1}, \tag{IV.6}$$

$$s_{\mathscr{C}_n}|_{C_n^2} = f_{\mathscr{C}_n}|_{C_n^2}. \tag{IV.7}$$

In this case, $\{s_{\mathscr{C}} \in \mathcal{S}^{(k)}(\mathscr{C})\}_{\mathscr{C}\in\underline{\mathfrak{D}}}$ is called a *standard form of* $\left\{f_{\mathscr{C}} \in \mathcal{F}^{(k)}(\mathscr{C})\right\}_{\mathscr{C}\in\underline{\mathfrak{D}}}$.

Note that a 1-partial family for $\mathcal{F}^{(k)}$ is simply a single section $f \in \mathcal{F}^{(k)}(\mathscr{C})$ over a context $\mathscr{C} \in \mathcal{M}^{(k)}$, which verifies conditions (IV.4) and (IV.5).

We have introduced partial families in order to model the typical cohomology false negative. Indeed, non-standard partial families are nothing but partial families of $\mathcal{F}^{(k)}$ (i.e. families of linear combinations of sections of $\mathcal{S}^{(k)}$) that cannot be replaced by simple families of $\mathcal{S}^{(k)}$, just like a cohomology false negative is a compatible family for $\mathcal{F}$ which cannot be replaced by a compatible family of $\mathcal{S}$. Conditions (IV.4) and (IV.5) simply insure that the family always starts and ends at a single point, i.e. a local section of $\mathcal{S}^{(k_1)}$. We give some graphical intuition on partial families in Figure IV.9, to clarify this concept. Throughout the rest of the section, we will show how non-standard families can be suppressed by applying the line model construction a sufficient amount of times.

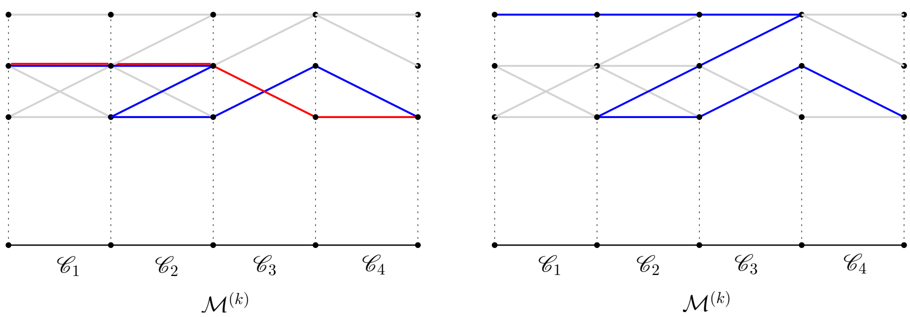


FIGURE IV.9. Two examples of 4-partial families for $\mathcal{F}^{(k)}$ (in blue) over the 4-path $\{\mathscr{C}_1, \mathscr{C}_2, \mathscr{C}_3, \mathscr{C}_4\} \subseteq \mathcal{M}^{(k)}$. On the left, a standard family, with a standard form highlighted in red. On the right, a non-standard partial family.

5.2. A complete cohomological invariant for contextuality in cyclic models. We will now show how to get rid of non-standard partial families. This procedure will require a number of intermediate steps.

The following lemma is called the *no-Z lemma* because it formalises the idea, introduced in Section 2.1, that first line models do not contain 'Z' shaped paths which typically give rise to false negatives in cohomology.

LEMMA IV.29 (No-Z lemma). *Let $\mathcal{S}$ be an empirical model over a measurement scenario $\langle X, \mathcal{M}, (O_m)\rangle$. Let $k \geq 1$, and $\mathscr{C} = \{C^1, C^2\} \in \mathcal{M}^{(k)}$. Every 1-partial family for $\mathcal{F}^{(k)}$ over $\mathscr{C}$ of the form*

$$f_{\mathscr{C}} = (s_1, t_1) + (s_2, t_1) + (s_2, t_2), \tag{IV.8}$$

(where $s_i \in \mathcal{S}^{(k-1)}(C^1)$*,* $t_i \in \mathcal{S}^{(k-1)}(C^2)$ *for all* $i = 1, 2$*) is standard.*

PROOF. Let $f_{\mathscr{C}}$ be a 1-partial family defined by (IV.8). Because (s_1, t_1), (s_2, t_1) and (s_2, t_2) are all in $\mathcal{S}^{(k)}(\mathscr{C})$, we know that

$$s_1|_{C^1 \cap C^2} = t_1|_{C^1 \cap C^2} = s_2|_{C^1 \cap C^2} = t_2|_{C^1 \cap C^2} \tag{IV.9}$$

Therefore, $(s_1, t_2) \in \mathcal{S}^{(k)}(\mathscr{C})$, and we have

$$(s_1, t_2)|_{C^1} = s_1 = s_1 + \underbrace{2 \cdot s_2}_{=0} = f_{\mathscr{C}}|_{C^1}$$

$$(s_1, t_2)|_{C^2} = t_2 = t_2 + \underbrace{2 \cdot t_1}_{=0} = f_{\mathscr{C}}|_{C^2},$$

which correspond to conditions (IV.6) and (IV.7) (we have used the fact that the coefficients are in $\mathbb{Z}_2$, hence $2 = 0$). This proves that (s_1, t_2) is the standard form of $f_{\mathscr{C}}$. ☐

We will now generalise the no-Z lemma to all the 1-partial families for $\mathcal{F}^{(k)}$. The proof essentially consists of a recursive algorithm which takes a 1-partial family as input, and outputs a standard form by repeatedly applying the no-Z lemma to the first three segments of the partial family, which – we show – are always in a 'Z' shape.

LEMMA IV.30. *Let* $\mathcal{S}$ *be an empirical model on a scenario* $\langle X, \mathcal{M}, (O_m) \rangle$*, and let* $k \geq 1$*. All the* 1*-partial families for* $\mathcal{F}^{(k)}$ *are standard.*

PROOF. A 1-partial family is a single section $f_{\mathscr{C}} \in \mathcal{F}^{(k)}(\mathscr{C})$ over a single context $\mathscr{C} = \{C^1, C^2\} \in \mathcal{M}^{(k)}$, which verifies conditions (IV.4) and (IV.5). We provide an algorithm that constructs a standard form $s_{\mathscr{C}} \in \mathcal{S}^{(k)}(\mathscr{C})$. Let us start by enumerating the possible sections at C^1 and C^2 by denoting $\mathcal{S}^{(k-1)}(C^1) = \{s_1, s_2, \ldots, s_n\}$, and $\mathcal{S}^{(k-1)}(C^2) = \{t_1, t_2, \ldots, t_m\}$. Let

$$I := \{(i, j) \in [n] \times [m] \mid (s_i, t_j) \in \mathcal{S}^{(k)}(\mathscr{C})\},$$

where $[l] := \{1, 2, \ldots, l\}$. By definition of $\mathcal{F}^{(k)}$, the section $f_{\mathscr{C}}$ can be written as a formal linear combination of sections in $\mathcal{S}^{(k)}(\mathscr{C})$:

$$f_{\mathscr{C}} = \sum_{(i,j) \in I} \alpha_{ij} \cdot (s_i, t_j),$$

where $\alpha_{ij} \in \mathbb{Z}_2$.

If $f_{\mathscr{C}} \in \mathcal{S}^{(k)}(\mathscr{C})$, then we are done, as $f_{\mathscr{C}}$ is already in standard form. Otherwise, we know by (IV.4) that $f_{\mathscr{C}}|_{C^1} \in \mathcal{S}^{(k-1)}(C^1)$. Therefore, we can assume w.l.o.g. that $f_{\mathscr{C}}|_{C^1} = s_1$. Because of this, there exists a $j_1 \in [m]$ such that $\alpha_{1j_1} = 1$, and we can assume w.l.o.g. that $j_1 = 1$, which means that $f_{\mathscr{C}}$ contains the section $(s_1, t_1) \in \mathcal{S}^{(1)}(\mathscr{C})$ in its summands, i.e.

$$f_{\mathscr{C}} = (s_1, t_1) + \sum_{(i,j) \neq (1,1)} \alpha_{ij}(s_i, t_j) \tag{IV.10}$$

By equation (IV.5), we know that $f_{\mathscr{C}}|_{C^2} \in \mathcal{S}^{(k-1)}(C^2)$ and we can denote $f_{\mathscr{C}}|_{C^2} = t_l$, for some $l \in [m]$. If $l = 1$, then we can immediately return (s_1, t_1) as the standard form of $f_{\mathscr{C}}$. Otherwise we assume $l \neq 1$.

Claim 1. There exists an index $i_1 \neq 1$, $i_1 \in [n]$, such that $(i_1, 1) \in I$ and $\alpha_{i_1 1} = 1$. W.l.o.g. we let $i_1 = 2$.

Proof Suppose *ab absurdo* $\alpha_{i1} = 0$ for all $1 \neq i \in [n]$ such that $(i, 1) \in I$. Then, given (IV.10), we have

$$f_{\mathscr{C}} = (s_1, t_1) + \sum_{i,j:\ j \neq 1} \alpha_{ij}(s_i, t_j).$$

This implies

$$f_{\mathscr{C}}|_{C^2} = t_1 + \sum_{i,j:\ j \neq 1} \alpha_{ij} t_j,$$

which always contains the summand $t_1 \neq t_l$, and thus can never equal t_l, which is a contradiction. ■

Claim 1 shows that $f_{\mathscr{C}}$ always contains the summand (s_2, t_1), i.e.

$$f_{\mathscr{C}} = (s_1, t_1) + (s_2, t_1) + \sum_{\substack{(i,j) \neq (1,1) \\ (i,j) \neq (2,1)}} \alpha_{ij}(s_i, t_j). \tag{IV.11}$$

Claim 2. There exists an index $j_2 \in [m]$, $j_2 \neq 1$, such that $(2, j_2) \in I$ and $\alpha_{2j_2} = 1$. W.l.o.g. we let $j_2 = 2$.

Proof Suppose by contradiction that $\alpha_{2j} = 0$ for all $1 \neq j \in [m]$ such that $(2, j) \in I$. Then, given (IV.11), we have

$$f_{\mathscr{C}} = (s_1, t_1) + (s_2, t_1) + \sum_{\substack{(i,j) \neq (1,1) \\ i \neq 2}} \alpha_{ij}(s_i, t_j).$$

This implies

$$f_{\mathscr{C}}|_{\mathscr{C}^1} = s_1 + s_2 + \sum_{\substack{(i,j) \neq (1,1) \\ i \neq 2}} \alpha_{ij} s_i,$$

which always contains the summand s_2, and thus can never equal s_1, which is a contradiction. ■

Claim 2 shows that $f_{\mathscr{C}}$ always contains the summand (s_2, t_2), i.e.

$$f_{\mathscr{C}} = (s_1, t_1) + (s_2, t_1) + (s_2, t_2) + \sum_{\substack{(i,j) \neq (1,1) \\ (i,j) \neq (2,1) \\ (i,j) \neq (2,2)}} \alpha_{ij}(s_i, t_j).$$

Notice how the first three summands are exactly the same as in (IV.8). This means that these first three 'steps' of the partial family $f_{\mathscr{C}}$ are in a 'Z' shape. This allows us to apply the no-Z lemma and substitute the Z with a section in $\mathcal{S}^{(k)}(\mathscr{C})$, as shown in Figure IV.10.

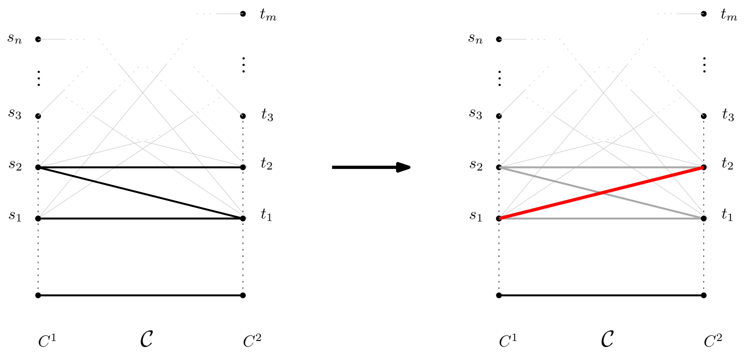


FIGURE IV.10. A visualisation of the proof. On the left-hand side the 'Z' shape at the beginning of the partial family. On the right-hand side, the substitution of the 'Z' with a section of $\mathcal{S}^{(k)}(\mathscr{C})$.

In other words, by the no-Z lemma (Lemma IV.29), $(1, 2)$ must be in I, and the section $(s_1, t_2) \in \mathcal{S}^{(k)}(\mathscr{C})$ is the standard form of the partial family $(s_1, t_1) + (s_2, t_1) + (s_2, t_2)$.

If $l = 2$, then (s_1, t_2) is the standard form of $f_{\mathscr{C}}$ and we are done. Otherwise we can re-input the partial family

$$f'_{\mathscr{C}} := (s_1, t_2) + \sum_{\substack{(i,j)\neq(1,1)\\(i,j)\neq(2,1)\\(i,j)\neq(2,2)}} \alpha_{ij}(s_i, t_j)$$

into the algorithm. Notice that this family has two non-zero coefficients less than the original one, thus it strictly decreases in size. Therefore, the algorithm obviously terminates as there is only a finite amount of sections. □

The following theorem is the key result of the chapter. It shows that, on cyclic scenarios, all the n-partial families for $\mathcal{F}^{(k)}$, where $n \leq k$, can be replaced by a standard form of the same size. In other words, any potential cohomological false negative of size $n \leq k$ can be erased. This fact will lead us to a fundamental result, namely that, on a cyclic scenario, it is sufficient to take the $(|\mathcal{M}| - 1)$-th line model to remove every cohomology false negative with certainty (Theorem IV.32).

THEOREM IV.31. *Let $\mathcal{S}$ be an empirical model on a cyclic scenario $\langle X, \mathcal{M}, (O_m)\rangle$. Let $k \geq 1$ and let n be such that $n \leq k$ and $n < |\mathcal{M}|$. All the n-partial families for $\mathcal{F}^{(k)}$ are standard.*

PROOF. We will proceed by induction on k. The base case $k = 1$ is a special case of Lemma IV.30. We will now suppose $k \geq 2$.

Let $P := \left\{f_{\mathscr{C}} \in \mathcal{F}^{(k)}(\mathscr{C})\right\}_{\mathscr{C}\in\mathfrak{D}}$ be an n-partial family for $\mathcal{F}^{(k)}$ over the n-path $\mathfrak{D} = \{\mathscr{C}_1, \dots, \mathscr{C}_n\} \subseteq \mathcal{M}^{(k)}$. Notice that, for $n = 1$, the result follows directly from Lemma IV.30. Suppose $n \geq 2$. Because $\langle X, \mathcal{M}, (O_m)\rangle$ is cyclic, by Proposition IV.22 we know that $\mathcal{M}^{(k)}$ is a chordless $|\mathcal{M}|$-cycle. Moreover, $\mathfrak{D}$ is not a cycle, since $n < |\mathcal{M}|$. Thus we can use the notation of Remark IV.24. Let

$$\underline{\mathscr{D}}' := \{K_1, K_2, \dots K_{n-1}\} = \{C_1^2, C_2^2, \dots C_{n-1}^2\} \subseteq \mathcal{M}^{(k-1)}.$$

Then $\underline{\mathscr{D}}'$ is an $(n-1)$-path for $\mathcal{M}^{(k-1)}$. Indeed, by Proposition IV.25 we know that $\underline{\mathscr{D}}' \cup \{C_1^1, C_n^2\}$ is an $(n+1)$-path for $\mathcal{M}^{(k-1)}$.

Define a family $P' := \left\{t_{K_i} \in \mathcal{F}^{(k-1)}(K_i)\right\}_{i=1}^{n-1}$ by

$$t_{K_i} := f_{\mathscr{C}_i}|_{K_i}, \tag{IV.12}$$

(cf. Remark IV.9).

Claim 1. The family P' is an $(n-1)$-partial family for $\mathcal{F}^{(k-1)}$.

Proof Let us start by proving that P' is compatible for $\mathcal{F}^{(k-1)}$. Let $1 \leq i \leq n-2$. We have

$$\begin{aligned} t_{K_i}|_{K_i \cap K_{i+1}} &= (f_{\mathscr{C}_i}|_{K_i})\,|_{K_i \cap K_{i+1}} \overset{(*)}{=} \left(f_{\mathscr{C}_i}|_{\mathscr{C}_i \cap \mathscr{C}_{i+1}}\right)|_{K_i \cap K_{i+1}} \\ &\overset{(\dagger)}{=} \left(f_{\mathscr{C}_{i+1}}|_{\mathscr{C}_i \cap \mathscr{C}_{i+1}}\right)|_{K_i \cap K_{i+1}} \overset{(*)}{=} \left(f_{\mathscr{C}_{i+1}}|_{K_i}\right)|_{K_i \cap K_{i+1}} \\ &= f_{\mathscr{C}_{i+1}}|_{K_i \cap K_{i+1}} = \left(f_{\mathscr{C}_{i+1}}|_{K_{i+1}}\right)|_{K_i \cap K_{i+1}} \\ &= t_{K_{i+1}}|_{K_i \cap K_{i+1}}, \end{aligned}$$

where we have used the fact that $\mathscr{C}_i \cap \mathscr{C}_{i+1} = \{C_i^2\} = \{K_i\}$ in the equalities $(*)$ (cf. Remark IV.9), and the fact that P is compatible for $\mathcal{F}^{(k)}$ in equality $(\dagger)$. With the usual notation $K_i := \{k_i^1, k_i^2\}$, since $\{k_1^1, k_1^2, k_2^2, \dots, k_{n-1}^2\}$ is an n-path for $\mathcal{M}^{(k-2)}$ by Proposition IV.25, we know that k_1^1 is the first vertex of the path, which means that $\{k_1^1\} = C_1^1 \cap C_1^2$. In view of Remark IV.9, we have

$$\begin{aligned} t_{K_1}|_{k_1^1} &= (f_{\mathscr{C}_1}|_{K_1})\,|_{k_1^1} = \left(f_{\mathscr{C}_1}|_{C_1^2}\right)|_{k_1^1} = \left(f_{\mathscr{C}_1}|_{C_1^2}\right)|_{C_1^1 \cap C_1^2} = f_{\mathscr{C}_1}|_{C_1^1 \cap C_1^2} \\ &= \left(f_{\mathscr{C}_1}|_{C_1^1}\right)|_{C_1^1 \cap C_1^2} = \left(f_{\mathscr{C}_1}|_{C_1^1}\right)|_{k_1^1} \end{aligned}$$

Because $f_{\mathscr{C}_1}|_{C_1^1} \in \mathcal{S}^{(k-1)}(C_1^1)$ by condition (IV.4), we must have

$$t_{K_1}|_{k_1^1} = \left(f_{\mathscr{C}_1}|_{C_1^1}\right)|_{C_1^1 \cap C_1^2} \in \mathcal{S}^{(k-1)}(k_1^1), \tag{IV.13}$$

hence P' satisfies (IV.4).

We prove (IV.5) essentially in the same way: we start by a simple observation, namely that, because $K_{n-1} = C_{n-1}^2$, we have $\{K_{n-1}\} = \mathscr{C}_{n-1} \cap \mathscr{C}_n$. Therefore,

$$t_{K_{n-1}} = f_{\mathscr{C}_{n-1}}|_{K_{n-1}} = f_{\mathscr{C}_{n-1}}|_{C_{n-1} \cap C_n} = f_{\mathscr{C}_n}|_{C_{n-1} \cap C_n} = f_{\mathscr{C}_n}|_{K_{n-1}}, \tag{IV.14}$$

where we have used the fact that P is compatible in the third equality. Now, with a similar argument as before, given that $\{k_{n-1}^2\} = C_{n-1}^2 \cap C_n^2$, we have

$$\begin{aligned} t_{K_{n-1}}|_{k_{n-1}^2} &\overset{\text{(IV.14)}}{=} (f_{\mathscr{C}_n}|_{K_{n-1}})\,|_{k_{n-1}^2} = \left(f_{\mathscr{C}_n}|_{C_{n-1}^2}\right)|_{k_{n-1}^2} = \left(f_{\mathscr{C}_n}|_{C_{n-1}^2}\right)|_{C_{n-1}^2 \cap C_n^2} \\ &= \quad f_{\mathscr{C}_n}|_{C_{n-1}^2 \cap C_n^2} = \left(f_{\mathscr{C}_n}|_{C_n^2}\right)|_{C_{n-1}^2 \cap C_n^2} = \left(f_{\mathscr{C}_n}|_{C_n^2}\right)|_{k_{n-1}^2}. \end{aligned}$$

Because $f_{\mathscr{C}_n}|_{C_n^2} \in \mathcal{S}^{(k-1)}(C_1^1)$ by condition (IV.5), we must have

$$t_{K_{n-1}}|_{k_{n-1}^2} = \left(f_{\mathscr{C}_n}|_{C_n^2}\right)|_{C_{n-1}^2 \cap C_n^2} \in \mathcal{S}^{(k-1)}(k_{n-1}^2), \tag{IV.15}$$

which means that P' satisfies (IV.5). ■

Because P' is an $(n-1)$-partial family for $\mathcal{F}^{(k-1)}$, by inductive hypothesis, we know that P' is standard. Let $S := \{s_{K_i} \in \mathcal{S}^{(k-1)}(K_i)\}_{i=1}^{n-1}$ be a standard form of P', i.e. S is compatible for $\mathcal{S}^{(k-1)}$ and it is such that

$$s_{K_1}|_{k_1^1} = t_{K_1}|_{k_1^1}, \tag{IV.16}$$

$$s_{K_{n-1}}|_{k_{n-1}^2} = t_{K_{n-1}}|_{k_{n-1}^2}. \tag{IV.17}$$

Consider the family $G := \{g_{\mathscr{C}_i} \in \mathcal{S}^{(k)}(\mathscr{C}_i)\}_{i=1}^n$, defined as follows:

$$g_{\mathscr{C}_i} := \begin{cases} \left(f_{\mathscr{C}_1|_{C_1^1}}, s_{K_1}\right) & \text{if } i = 1, \\ \left(s_{K_{n-1}}, f_{\mathscr{C}_n}|_{C_n^2}\right) & \text{if } i = n, \\ \left(s_{K_{i-1}}, s_{K_i}\right) & \text{for all } 2 \leq i \leq n-1. \end{cases}$$

Claim 2. The family G is a standard form for P.

Proof First of all, we need to check that $g_{\mathscr{C}_i}$ is indeed an element of $\mathcal{S}^{(k)}(\mathscr{C}_i)$ for all $1 \leq i \leq n$. We have

$$\left(f_{\mathscr{C}_1}|_{C_1^1}\right)|_{C_1^1 \cap K_1} = \left(f_{\mathscr{C}_1}|_{C_1^1}\right)|_{C_1^1 \cap C_1^2} \overset{\text{(IV.13)}}{=} t_{K_1}|_{k_1^1} \overset{\text{(IV.16)}}{=} s_{K_1}|_{k_1^1} = s_{K_1}|_{C_1^1 \cap K_1}.$$

Similarly,

$$\begin{aligned} \left(f_{\mathscr{C}_n}|_{C_n^2}\right)|_{K_{n-1} \cap C_n^2} &= \left(f_{\mathscr{C}_n}|_{C_n^2}\right)|_{C_{n-1}^2 \cap C_n^2} \overset{\text{(IV.15)}}{=} t_{K_{n-1}}|_{k_{n-1}^2} \overset{\text{(IV.17)}}{=} s_{K_{n-1}}|_{k_{n-1}^2} \\ &= s_{K_{n-1}}|_{K_{n-1} \cap C_n^2} \end{aligned}$$

Finally, let $2 \leq i \leq n-1$. We readily have

$$s_{K_{i-1}}|_{K_{i-1} \cap K_i} = s_{K_i}|_{K_{i-1} \cap K_i}$$

by the simple fact that S is compatible for $\mathcal{S}^{(k-1)}$.

The fact that G satisfies equations (IV.6) and (IV.7) for P trivially follows from the very definition of G, indeed

$$\begin{aligned} g_{\mathscr{C}_1}|_{\mathscr{C}_1^1} &= \left(f_{\mathscr{C}_1|_{C_1^1}}, s_{K_1}\right)|_{C_1^1} = f_{\mathscr{C}_1}|_{C_1^1}; \\ g_{\mathscr{C}_n}|_{\mathscr{C}_n^2} &= \left(s_{K_{n-1}}, f_{\mathscr{C}_n}|_{C_n^2}\right)|_{C_n^2} = f_{\mathscr{C}_n}|_{C_n^2}. \end{aligned}$$

■

Thanks to this claim, we have successfully proved that P is standard. □

We can now introduce a complete cohomology characterisation of logical and strong contextuality for cyclic scenarios:

THEOREM IV.32. *Let $\mathcal{S}$ be an empirical model on a cyclic scenario $\langle X, \mathcal{M}, (O_m)\rangle$. Let $C \in \mathcal{M}$ and $s \in \mathcal{S}(C)$. Then,*

$$\mathsf{LC}(\mathcal{S}, s) \Leftrightarrow \mathsf{CLC}^{(n)}(\mathcal{S}, s),$$

where $n := |\mathcal{M}| - 1$. Moreover,

$$\mathsf{SC}(\mathcal{S}) \Leftrightarrow \mathsf{CSC}(\mathcal{S}^{(n)})$$

PROOF. The implications $\mathsf{CLC}^{(n)}(\mathcal{S}, s) \Rightarrow \mathsf{LC}(\mathcal{S}, s)$ and $\mathsf{CSC}(\mathcal{S}^{(n)}) \Rightarrow \mathsf{SC}(\mathcal{S})$ have already been proven in Theorem IV.27.

To prove the converse, we will show that $\neg\mathsf{CLC}^{(n)}(\mathcal{S}, s) \Rightarrow \neg\mathsf{LC}(\mathcal{S}, s)$. Suppose $\neg\mathsf{CLC}^{(n)}(\mathcal{S}, s)$. By Definition IV.26, this implies that there exists a context $\mathscr{C}_0 \in \mathcal{M}^{(n)}$ and a section $t \in \mathcal{S}^{(n)}(\mathscr{C}_0)$ such that $s \in \mathsf{flatten}(t)$ and $\neg\mathsf{CLC}(\mathcal{S}^{(n)}, t)$. Thus, there exists a compatible family

$$F := \left\{ f_{\mathscr{C}} \in \mathcal{F}^{(n)}(\mathscr{C}) \right\}_{\mathscr{C} \in \mathcal{M}^{(n)}}$$

such that $f_{\mathscr{C}_0} = t$. Because $\langle X, \mathcal{M}, (O_m)\rangle$ is cyclic, we know by Proposition IV.22 that $\mathcal{M}^{(n-1)}$ is also cyclic. Theferefore,

$$\mathcal{M}^{(n)} = \{\mathscr{C}_0, \mathscr{C}_1, \ldots, \mathscr{C}_n\}$$

is a chordless $|\mathcal{M}|$-cycle, which implies that $\{\mathscr{C}_1, \ldots, \mathscr{C}_n\}$ is a chordless n-path (in edge representation) for $\mathcal{M}^{(n)}$. Let

$$P := \left\{ f_{\mathscr{C}_i} \in \mathcal{F}^{(n)}(\mathscr{C}_i) \right\}_{i=1}^{n}.$$

Then P is a n-partial family for $\mathcal{F}^{(n)}$, indeed it is compatible because F is compatible, and we have, with the usual notation

$$f_{\mathscr{C}_1}|_{C_1^1} = f_{\mathscr{C}_1}|_{\mathscr{C}_0 \cap \mathscr{C}_1} \overset{(*)}{=} f_{\mathscr{C}_0}|_{\mathscr{C}_0 \cap \mathscr{C}_1} = t|_{\mathscr{C}_0 \cap \mathscr{C}_1} = t|_{\{C_1^1\}} = t|_{C_1^1} \in \mathcal{S}^{(n-1)}(C_1^1),$$

and

$$f_{\mathscr{C}_n}|_{C_n^2} = f_{\mathscr{C}_n}|_{\mathscr{C}_n \cap \mathscr{C}_0} \overset{(*)}{=} f_{\mathscr{C}_0}|_{\mathscr{C}_n \cap \mathscr{C}_0} = t|_{\mathscr{C}_n \cap \mathscr{C}_0} = t|_{\{C_n^2\}} = t|_{C_n^2} \in \mathcal{S}^{(n-1)}(C_n^2),$$

where we have used the fact that F is compatible in equalities $(*)$. By Theorem IV.31, we know that P is standard. Thus there exists a compatible family

$$P' := \{ s_{\mathscr{C}_i} \in \mathcal{S}^{(n)}(\mathscr{C}_i) \}_{i=1}^{n}$$

such that

$$s_{\mathscr{C}_1}|_{C_1^1} = f_{\mathscr{C}_1}|_{C_1^1} = t|_{\mathscr{C}_0 \cap \mathscr{C}_1},$$
$$s_{\mathscr{C}_n}|_{C_n^2} = f_{\mathscr{C}_n}|_{C_n^2} = t|_{\mathscr{C}_n \cap \mathscr{C}_0}.$$

Therefore, the family

$$P' \cup \{t\} = \{ s_{\mathscr{C}_i} \in \mathcal{S}^{(n)}(\mathscr{C}_i) \}_{i=1}^{n} \cup \{t\}$$

is a compatible family for $\mathcal{S}^{(n)}$ that contains t. Thus we have $\neg\mathsf{LC}(\mathcal{S}^{(n)}, t)$, which means that $\neg\mathsf{LC}^{(n)}(\mathcal{S}, s)$. It follows from Corollary IV.14 that $\mathcal{S}$ is *not* logically contextual at s.

Suppose now $\neg\mathsf{CSC}(\mathcal{S}^{(n)})$. Then there exists a section t of $\mathcal{S}^{(n)}$ such that $\neg\mathsf{LC}(\mathcal{S}^{(n)}, t)$. Consider an arbitrary section s of $\mathcal{S}$ such that $s \in \mathsf{flatten}(t)$. Then we have $\neg\mathsf{LC}^{(n)}(\mathcal{S}, s)$, and we can apply the same argument used before to show that this implies $\neg\mathsf{LC}(\mathcal{S}, s)$, which in turn implies $\neg\mathsf{SC}(\mathcal{S})$. $\square$

This theorem tells us that if we want to study the contextuality of a cyclic scenario $\langle X, \mathcal{M}, (O_m)\rangle$, it is sufficient to analyse the cohomology of its $(|\mathcal{M}|-1)$-th line model to assert with certainty which sections give rise to contextual behaviour. This is a major step forward, as this method allows us to get rid of all the false negatives presented in this thesis and many others, as we shall see in the rest of the chapter.

5.3. Examples. In this section we will show how this method applies to the false negatives appeared in the literature, presented in the previous sections and chapters.

The model of Table IV.1. We have already shown that the model $\mathcal{S}$ presented in Table IV.1 displays a cohomology false negative for the section s_2 (cf. Table IV.3), and suggested that it vanishes as soon as we consider its first line model $\mathcal{S}^{(1)}$. We can give a formal proof that this is true. In Figure IV.11 we present once again the bundle diagram of the first line model, where we introduce variables $a, b, c, d, e, f, g, h, i, j \in \mathbb{Z}_2$ that represent the coefficients to give to every section of $\mathcal{S}^{(1)}$ in order to construct a compatible family for $\mathcal{F}^{(1)}$.

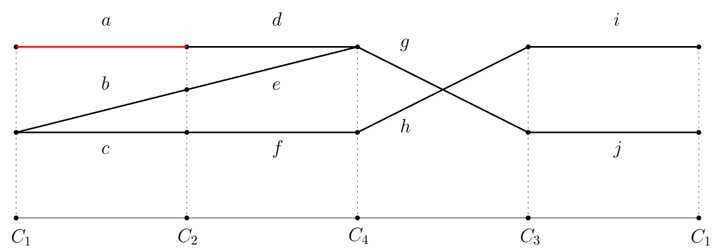


FIGURE IV.11. The bundle diagram of $\mathcal{S}^{(1)}$ with the variables in $\mathbb{Z}_2$ corresponding to each section. The seciton s_2, responsible for logical contextuality (cf. Section 2.1) is highlighted in red.

The compatibility conditions of a presumed compatible family for $\mathcal{F}^{(1)}$ can be summarised in the following equations:

$$\begin{array}{llll} a = d, & & & \\ b = e, & d \oplus e = g, & h = i, & i = a, \\ & f = h, & g = j, & j = b \oplus c. \\ c = f, & & & \end{array}$$

Because the family must contain s_2, which is marked in red in Figure IV.11, we must have $a = 1$ and $b = c = 0$. It follows directly that $e = f = h = i = 0$ and that $d = g = j = 1$. However, since $j = b \oplus c$, this leads to $1 = 0 \oplus 0 = 0$, which is obviously a contradiction. We have just proved that the cohomology of $\mathcal{S}^{(1)}$ does detect the logical contextuality of $\mathcal{S}$ at s_2.

Note that in this case, although $|\mathcal{M}| = 4$, it was not necessary to take the third line model of $\mathcal{S}$ to remove the false negative, as suggested by Theorem IV.32. Indeed, the bound $|\mathcal{M}| - 1$ is the one that gives us absolute certainty about the non-existence of a false negative. However, as we have just shown, it might be sufficient to take a lower level line model to remove any false negative from the model.

The Hardy model. The Hardy model (cf. Table II.8) is perhaps the most well-studied example of cohomological false negative for contextuality [**AMB12, ABK$^+$15, Car17**]. We have shown in Figure IV.3 that its first line model still results in a cohomological false negative for the section s_1, at which the Hardy model $\mathcal{S}$ is logically contextual. In Figure IV.12, we present the bundle diagram of the second line model.

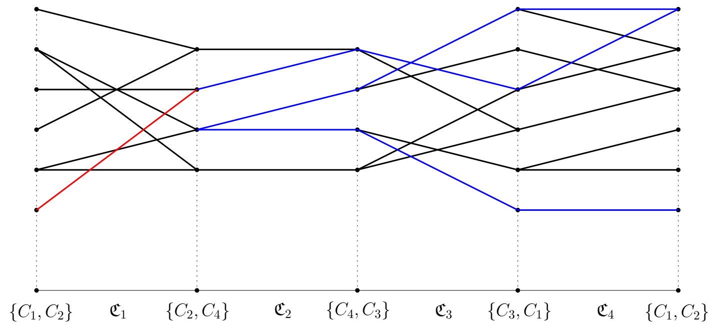


FIGURE IV.12. The bundle diagram of the second line model of the Hardy model. The red section is the only section of $\mathcal{S}^{(2)}$ that contains the original section s_1. In blue, a false negative for the red section

Notice that, even in this case, we still have a compatible family for $\mathcal{F}^{(2)}$ containing the only section of $\mathcal{S}^{(2)}$ over $\mathfrak{C}_1$ that contains s_1. Thus, we must consider the third line model to get rid of the false negative. The third line model of the Hardy model is presented in Figure IV.13, where we have highlighted in red the only section over $\{\mathfrak{C}_1, \mathfrak{C}_2\}$ containing s_1.

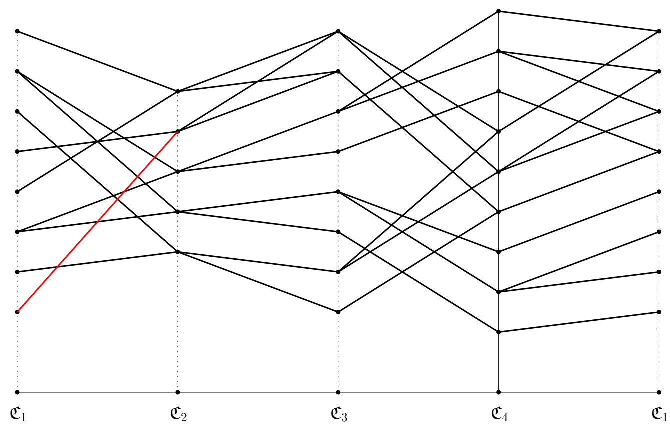


FIGURE IV.13. The bundle diagram of the third line model of the Hardy model. The section marked in red is the only section containing s_1.

Since $|\mathcal{M}| = 4$, Theorem IV.32 assures that cohomology does detect contextuality at the red section. This can be graphically checked by highlighting all the possible attempts to extend the red section to a compatible family for $\mathcal{F}^{(4)}$, as shown in Figure IV.14. In particular, we show all the possibilities to extend the section starting from left to right.

Note that the choice of direction is irrelevant, and the reader can verify that the same is true if we try to extend right to left instead.

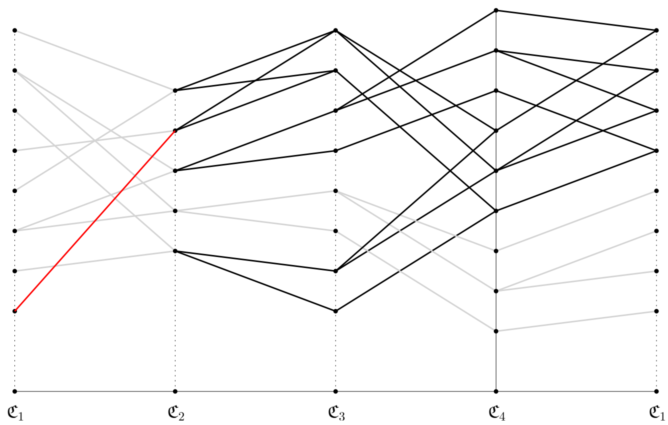


FIGURE IV.14. It is impossible to extend the red section to a compatible family for $\mathcal{F}^{(4)}$.

This graphical proof can easily be converted into a formal proof following the same idea as in the previous paragraph. Note that the fact that we had to consider the third line model of the Hardy model in order to get rid of the false negative shows that the bound $|\mathcal{M}| - 1$ of Theorem IV.32 is tight.

The false negative for strong contextuality of Table III.4. The false negative presented in Chapter III (Table III.4) is particularly interesting because it concerns all the sections of the model. Indeed, the model is cohomologically non-contextual despite being strongly contextual. In other words, there is a cohomology false negative for every single section of the model. In Figure IV.15, we depict the bundle diagram of the first line model.

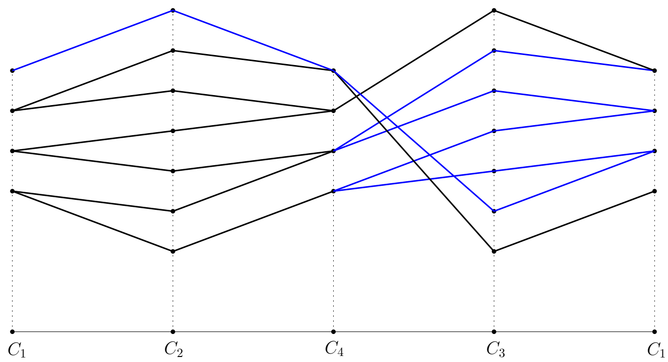


FIGURE IV.15. The first line model $\mathcal{S}^{(1)}$ of the false negative from Chapter III. In blue, a compatible family for $\mathcal{F}^{(1)}$.

Notice that, for each section, it is still possible to find a compatible family for $\mathcal{F}^{(1)}$ that contains it, giving rise to a false negative. For example, we have highlighted one such compatible family in blue, which constitutes a false negative for the contextuality of the top section for the context $\{C_1, C_2\}$. Therefore, we need to consider the second line model, whose bundle diagram is depicted in Figure IV.16 Even in this case, it is still

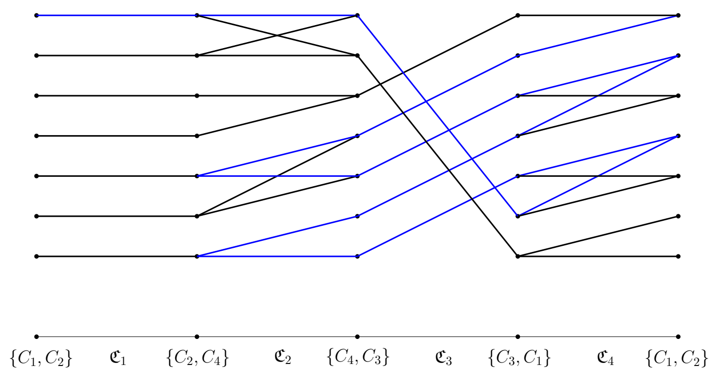


FIGURE IV.16. The second line model of the false negative from Chapter III. In blue, a compatible family for $\mathcal{F}^{(2)}$.

possible to find a cohomology false negative for each section of the model (see e.g. the blue loop highlighted in Figure IV.16).

In Figure IV.17 the bundle diagram of the third line model is shown. Once again,

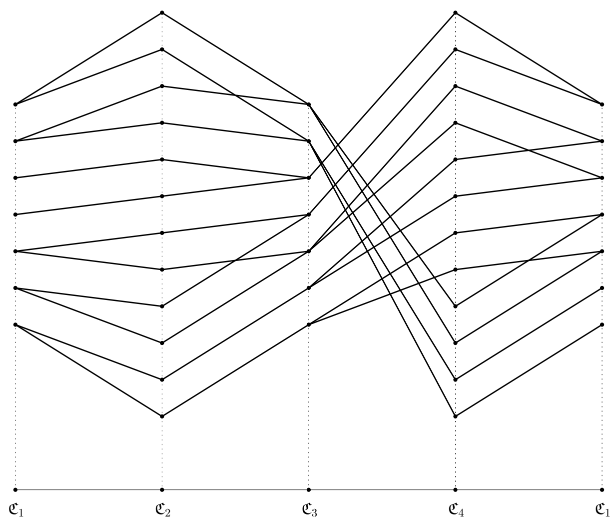


FIGURE IV.17. The third line model of the false negative from [**Car17**].

because $|\mathcal{M}| = 4$, we know by Theorem IV.32, that cohomology detects contextuality at every section of the model. This can be checked graphically. For example, in Figure IV.18 we show that it is never possible to extend the section marked in red to a compatible family for $\mathcal{F}^{(3)}$. The reader can verify that this is true for any section of the model.

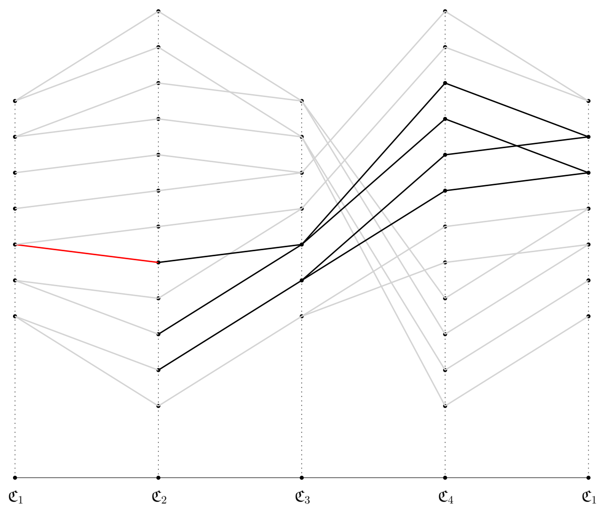


FIGURE IV.18. A graphical proof of the fact that cohomology of $\mathcal{F}^{(3)}$ does detect contextuality at the section marked in red. This is true for any section of the model.

6. Extending the invariant to general models

In the previous section, we have successfully defined a full cohomological invariant for contextuality for all cyclic models. The goal of this section is to extend this result to arbitrary models. In particular, will show that the invariant can be extended to a very large class of scenarios. This result will lead us to conjecture that the invariant works universally.

Although cyclic models constitute only a fraction of all the possible empirical models, they play a crucial role in the study of contextuality. In Chapter II, we presented Vorob'ev's theorem (Theorem II.21), which states that the existence of 'cycles' in the database theoretic sense of Definition II.20 is a necessary condition for contextuality. As mentioned earlier, in Remark IV.21, the database-theoretic notion of cyclicity is not strictly equivalent to the one we introduced in this chapter. However, it is easy to prove that any non-acyclic cover in the sense of Definition II.20 must contain at least one cyclic subcover in the sense defined here. The fact that the existence of cycles in the cover is necessary for contextuality suggests that contextual features can be observed by focusing uniquely on the cycles.

To convey this idea, we introduce here the notion of *cyclic contextuality property (CCP)*. The contextual features of models satisfying the CCP can always be recovered by looking at cycles in the cover $\mathcal{M}^{(1)}$.

DEFINITION IV.33. Let $\mathcal{S}$ be an empirical model on a scenario $\langle X, \mathcal{M}, (O_m) \rangle$. We say that $\mathcal{S}$ has the **cyclic contextuality property (CCP)** if, for each local section s of $\mathcal{S}$ such that $\mathsf{LC}(\mathcal{S}, s)$, there exists a cycle $\mathscr{D}_\bullet \subseteq \mathcal{M}$ for $\mathcal{M}^{(1)}$ (called a **contextual cycle of** s) such that $\mathsf{LC}(\mathcal{S}|_{\mathscr{D}_\bullet}, s)$, where $\mathcal{S}|_{\mathscr{D}_\bullet}$ is the model obtained by restricting $\mathcal{S}$ to the subcover $\mathscr{D}_\bullet$.

Most empirical models satisfy the CCP. To give an idea of how common this property is, it is sufficient to say that all the models that have appeared in the literature on the sheaf description of contextuality share this property. It shall also be mentioned that the author has not been able to find any example of an empirical model which does not satisfy this property.

Remarkably, the cohomological invariant introduced in the previous sections can be immediately extended to all models satisfying with the CCP.

PROPOSITION IV.34. *Let $\mathcal{S}$ be a model on a general scenario $\langle X, \mathcal{M}, (O_m)\rangle$, and suppose $\mathcal{S}$ has the CCP. For all sections s of $\mathcal{S}$, we have*

$$\mathsf{LC}(\mathcal{S}, s) \Leftrightarrow \mathsf{CLC}^{(n-1)}(\mathcal{S}, s),$$

where n denotes the size of any contextual cycle of s.

PROOF. The implication $\mathsf{CLC}^{(n-1)}(\mathcal{S}, s) \Rightarrow \mathsf{LC}(\mathcal{S}, s)$ follows from Theorem IV.27. Now, suppose $\mathsf{LC}(\mathcal{S}, s)$. Let $\mathscr{D}_\bullet \subseteq \mathcal{M}$ be a contextual cycle of s. By definition, we have $\mathsf{LC}(\mathcal{S}|_{\mathscr{D}_\bullet}, s)$. The model $\mathcal{S}|_{\mathscr{D}_\bullet}$ is defined on the cyclic scenario $\mathscr{D}_\bullet$, thus we can apply Theorem IV.32 to conclude that $\mathsf{CLC}^{(n-1)}(\mathcal{S}|_{\mathscr{D}_\bullet}, s)$, which readily implies $\mathsf{CLC}^{(n-1)}(\mathcal{S}, s)$. □

Thanks to this simple proposition, we can extend Theorem IV.32 to models satisfying the CCP over general scenarios. To prove this, we will need the following result:

PROPOSITION IV.35. *Let $\mathcal{S}$ be an empirical model on a scenario $\langle X, \mathcal{M}, (O_m)\rangle$. For all $k \geq 0$ and every section s of $\mathcal{S}$, we have*

$$\mathsf{CLC}^{(k)}(\mathcal{S}, s) \Rightarrow \mathsf{CLC}^{(l)}(\mathcal{S}, s) \ \forall l \geq k.$$

Similarly,

$$\mathsf{CSC}(\mathcal{S}^{(k)}) \Rightarrow \mathsf{CSC}(\mathcal{S}^{(l)}) \ \forall l \geq k.$$

PROOF. We are going to prove that $\neg\mathsf{CLC}^{(k+1)}(\mathcal{S}, s) \Rightarrow \neg\mathsf{CLC}^{(k)}(\mathcal{S}, s)$, and the result will follow by induction. Suppose $\neg\mathsf{CLC}^{(k+1)}(\mathcal{S}, s)$. Then there exists a context $\mathscr{C} = \{C_1, C_2\} \in \mathcal{M}^{(k+1)}$ and a section $t = (t_1, t_2) \in \mathcal{S}^{(k+1)}(\mathscr{C})$ (where $t_1 \in \mathcal{S}^{(k)}(C_1)$ and $t_2 \in \mathcal{S}^{(k)}(C_2)$) such that $s \in \mathsf{flatten}(t)$ and $\neg\mathsf{CLC}(\mathcal{S}^{(k+1)}, t)$. In particular, this means that there exists a compatible family

$$F = \{t_{\mathcal{K}} \in \mathcal{F}^{(k+1)}(\mathcal{K})\}_{\mathcal{K} \in \mathcal{M}^{(k+1)}}$$

such that $t_{\mathscr{C}} = t$. Given a context $C \in \mathcal{M}^{(k)}$ we know that there exists a $C' \in \mathcal{M}^{(k)}$ such that $\{C, C'\} \in \mathcal{M}^{(k+1)}$. Let $u_C := t_{\{C,C'\}}|_C$. This is well-defined because, given a different $C'' \in \mathcal{M}^{(k)}$ such that $\{C, C''\} \in \mathcal{M}^{(k+1)}$, we have

$$\begin{aligned} t_{\{C,C''\}}|_C &\overset{(*)}{=} t_{\{C,C''\}}|_{\{C\}} = t_{\{C,C''\}}|_{\{C,C'\}\cap\{C,C''\}} \overset{(\dagger)}{=} t_{\{C,C'\}}|_{\{C,C'\}\cap\{C,C''\}} \\ &= t_{\{C,C'\}}|_{\{C\}} \overset{(*)}{=} t_{\{C,C'\}}|_C, \end{aligned}$$

where we have used compatibility of F in (†), and applied what discussed in Remark IV.9 in (∗). Thus we can define the family

$$F' := \{u_C \in \mathcal{F}^{(k)}(C)\}_{C \in \mathcal{M}^{(k)}}.$$

We can show that F' is a compatible family for $\mathcal{F}^{(k)}$ as follows: suppose $C, C' \in \mathcal{M}^{(k)}$, $C \neq C'$ and $C \cap C' \neq \emptyset$, then

$$u_C|_{C\cap C'} = \left(t_{\{C,C'\}}|_C\right)|_{C\cap C'} = t_{\{C,C'\}}|_{C\cap C'} = \left(t_{\{C,C'\}}|_{C'}\right)|_{C\cap C'} = u_{C'}|_{C\cap C'}.$$

Now, because $s \in \mathsf{flatten}(t)$, we can suppose w.l.o.g. that $s \in \mathsf{flatten}(t_1)$. Moreover, because $t_{\mathscr{C}} = t = (t_1, t_2)$, we have $u_{C_1} = t_1$. Thus F' is a compatible family which contains t_1. We conclude that $\neg\mathsf{CLC}(\mathcal{S}^{(k)}, t_1)$, which implies $\neg\mathsf{CLC}^{(k)}(\mathcal{S}, s)$, as $s \in \mathsf{flatten}(t_1)$. The same argument can be used to prove that $\neg\mathsf{CSC}(\mathcal{S}^{(k+1)}) \Rightarrow \neg\mathsf{CSC}(\mathcal{S}^{(k)})$. □

From these two propositions, we immediately have the following theorem:

Theorem IV.36. *Let $\mathcal{S}$ be a model on a scenario $\langle X, \mathcal{M}, (O_m)\rangle$, and suppose $\mathcal{S}$ has the CCP. Let N denote the size of the largest cycle in $\mathcal{M}^{(1)}$. We have*

$$\mathsf{LC}(\mathcal{S}, s) \Leftrightarrow \mathsf{CLC}^{(N-1)}(\mathcal{S}, s),$$

for all section s of $\mathcal{S}$. Moreover, we have

$$\mathsf{SC}(\mathcal{S}) \Leftrightarrow \mathsf{CSC}(\mathcal{S}^{(N-1)}).$$

This result shows that, if a model has the CCP, studying its contextuality is equivalent to study the cohomological contextuality of its $(N-1)$-th line model. In other words, cohomological contextuality on $\mathcal{S}^{(N-1)}$ is a full invariant for contextuality on the original model. As in the case of cyclic scenarios, note that it might be possible to erase cohomological false negatives for a particular model even at a lower level.

Obviously, we usually do not know a priori whether a model satisfies the CCP, however, as we mentioned earlier, this property is extremely common among empirical models, which means that this method is widely applicable. In the following section we will give some examples to support this claim.

6.1. Examples.

A simple scenario. Let us start with the model displayed in Table IV.4. A

Table IV.4. The empirical model $\mathcal{S}$.

Contexts	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
$\{a,b\}$	0	1	1	0
$\{a,d\}$	1	0	1	1
$\{b,c\}$	1	1	0	1
$\{b,d\}$	1	0	0	1
$\{c,d\}$	0	1	1	0

bundle diagram representation of the model can be found in Figure IV.19. By simply looking at the diagram, it is easy to see that the section $(b,d) \mapsto (1,1)$, marked in red, cannot be extended to any compatible family for $\mathcal{S}$. However, it can be extended to a compatible family for $\mathcal{F}$, as shown in blue.

Using the enumeration specified in Table IV.5, we can represent the first line model $\mathcal{S}^{(1)}$ as a bundle diagram in Figure IV.20.

Notice that section s_{10}, marked in red, cannot be extended to a compatible family for $\mathcal{F}^{(1)}$ for the cycle $\{\{b,d\},\{b,c\},\{c,d\}\}$ (all the possibilities to extend section (s_8, s_10),

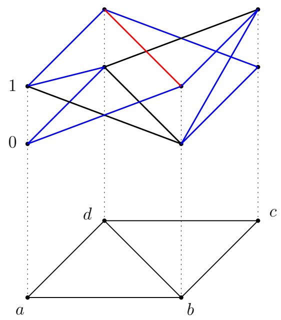


FIGURE IV.19. The bundle diagram of the model displayed in Table IV.4. In blue is highlighted the cohomological false negative for the section $(b, d) \mapsto (1, 1)$, marked in red.

TABLE IV.5. An enumeration of the sections of $\mathcal{S}$. The model is logically contextual at s_{10}

Contexts	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
$\{a, b\}$		s_1	s_2	
$\{a, d\}$	s_3		s_4	s_5
$\{b, c\}$	s_6	s_7		s_8
$\{b, d\}$	s_9			s_{10}
$\{c, d\}$		s_{11}	s_{12}	

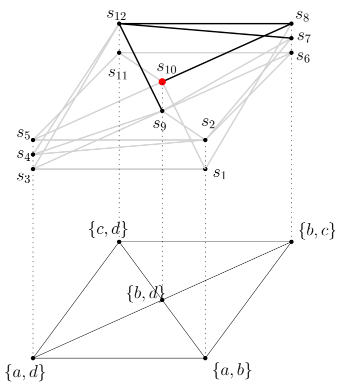


FIGURE IV.20. The bundle diagram of $\mathcal{S}^{(1)}$. In black, all the possibilities to extend the only section on $\{\{b, d\}, \{b, c\}$ containing s_{10}, i.e. (s_8, s_{10}), to a cohomology loop on the cycle $\{\{b, d\}, \{b, c\}, \{c, d\}\}$ by proceeding clockwise. They all fail to be compatible.

by proceeding clockwise are highlighted in black). In particular, this means that the cohomological false negative has been deleted. Note that in this case it was sufficient to derive the first line model to avoid a false negative. The size of the largest cycle in this scenario is 4, thus, in general, we would have to consider the third line model to remove any false negative with absolute certainty.

The Kochen-Specker model of [AMB12]. The only cohomological false negative on a non-cyclic model that has appeared in the literature is the Kochen-Specker model for the cover

$$\{A,B,C\},\{B,D,E\},\{C,D,E\},\{A,D,F\},\{A,E,G\}, \tag{IV.18}$$

introduced in [**AMB12**]. We have already shown how the false negative arises in Section 3 of Chapter III. In particular, we assigned variables $a, b, \ldots, o \in \mathbb{Z}_2$ for each of the local sections of $\mathcal{S}$ and showed that compatibility constraints force these variables to be organised as in Table III.3, which we report again here for clarity (Table IV.6)

Table IV.6. Table III.3 reported from Chapter III.

Contexts	$(1,0,0)$	$(0,1,0)$	$(1,0,0)$
$\{A,B,C\}$	a	b	b
$\{B,D,E\}$	b	b	a
$\{C,D,E\}$	b	b	a
$\{A,D,F\}$	a	b	b
$\{A,E,G\}$	a	a	a

By looking at this table, we were able to prove that $\mathcal{S}$ is strongly contextual. However, by letting $a = 1$ and $b = 0$ one obtains the following 'false' compatible family

$$\begin{aligned}\{&s_{\{A,B,C\},A},\ s_{\{B,D,E\},E},\ s_{\{C,D,E\},E},\ s_{\{A,D,F\},A},\\ &\qquad s_{\{A,E,G\},A} \oplus s_{\{A,E,G\},E} \oplus s_{\{A,E,G\},G}\},\end{aligned} \tag{IV.19}$$

which is a false negative for logical contextuality at sections $s_{\{A,B,C\},A}$, $s_{\{B,D,E\},E}$, $s_{\{C,D,E\},E}$ and $s_{\{A,D,F\},A}$. An important aspect of this family, which we have not highlighted in Chapter III, is that the only other compatible families we have for $\mathcal{F}$, namely the ones obtained by setting $a = 0, b = 1$ or $a = b = 1$ in Table IV.6, do not give rise to false negatives for any section of $\mathcal{S}$. Indeed, each one of these families contains exclusively sections of $\mathcal{F}$ that are not in $\mathcal{S}$. Therefore, we only need to show that the line model construction erases the unique false compatible family (IV.19) to conclude.

We will now show that it is sufficient to derive the first line model $\mathcal{S}^{(1)}$ to remove this false global secction. First of all, we represent the first line model using bundle diagrams. For each context $C = \{c_1, c_2, c_3\}$ of $\mathcal{M}$, there are exactly three possible sections, namely s_{C,c_1}, s_{C,c_2} and s_{C,c_3}. Therefore, for each vertex of $\mathcal{M}^{(1)}$, there are three distinct vertices in its fiber, which we will label with s_{C,c_1}, s_{C,c_2} and s_{C,c_3} from bottom to top (this labelling is not shown in the pictures for the sake of readability of the diagrams). Using this convention, we have depicted the bundle diagram of the first line model in Figure IV.21 (the colored vertices are only used as a visual reference).

In Figure IV.22, we give a different representation of the model by decomposing it into two planar diagrams. The top diagram corresponds to the cycle that constitutes the perimeter of the pentagon, while the bottom one corresponds to the star-shaped cycle in the centre. The colored circles in the fibers represent the four sections for which we have a cohomological false negative in the original model, as explained at the bottom of the picture. In the same picture, we have also introduced variables

$$a, b, \ldots, y, z, \tilde{a}, \tilde{b}, \ldots, \tilde{u}, \tilde{v}$$

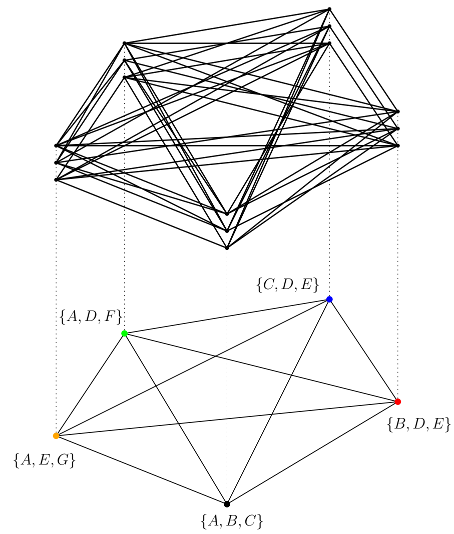


FIGURE IV.21. The bundle diagram of the first line model of the Kochen-Specker model on the cover (IV.18).

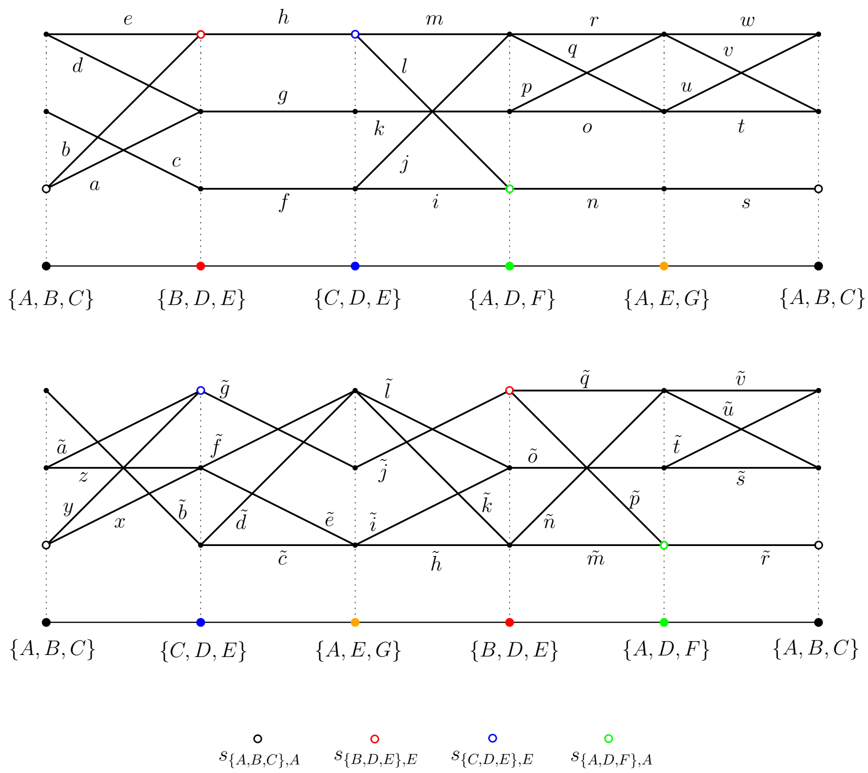


FIGURE IV.22. The Kochen-Specker model of [**AMB12**] decomposed in two cyclic planar diagrams. The top diagram corresponds to the perimeter of the pentagon; the bottom diagram refers to the central 'star'. We introduce one variable in $\mathbb{Z}_2$ for each section of the model.

in $\mathbb{Z}_2$ for each of the possible sections of the first line model. We will now show that the cohomological false negative no longer exists. To do so, we list all the equations imposed by compatibility conditions, which can be obtained from the diagram of Figure IV.22:

$$\begin{aligned}
a \oplus d &= g \\
b \oplus e &= h \\
c &= f \\
a \oplus d &= \tilde{o} \\
b \oplus e &= \tilde{p} \oplus \tilde{q} \\
c &= \tilde{m} \oplus \tilde{n} \\
a \oplus d &= \tilde{i} \oplus \tilde{l} \\
b \oplus e &= \tilde{j} \\
c &= \tilde{h} \oplus \tilde{k} \\
a \oplus b &= \tilde{r} \\
c &= \tilde{s} \oplus \tilde{u} \\
d \oplus e &= \tilde{t} \oplus \tilde{v} \\
a \oplus b &= s \\
c &= t \oplus v \\
d \oplus e &= u \oplus w \\
a \oplus b &= x \oplus y \\
c &= z \oplus \tilde{a} \\
d \oplus e &= \tilde{b} \\
f &= i \oplus j \\
g &= k \\
h &= l \oplus m \\
f &= b_1 \\
g &= x \oplus z \\
h &= y \oplus \tilde{a} \\
f &= \tilde{h} \oplus \tilde{i} \\
g &= \tilde{j} \\
h &= \tilde{k} \oplus \tilde{l} \\
f &= \tilde{h} \oplus \tilde{k} \\
g &= \tilde{i} \oplus \tilde{l} \\
h &= \tilde{j}
\end{aligned}$$

$$\begin{aligned}
f &= \tilde{m} \oplus \tilde{n} \\
g &= \tilde{o} \\
h &= \tilde{p} \oplus \tilde{q} \\
i \oplus l &= n \\
k &= o \oplus p \\
j \oplus m &= q \oplus r \\
i \oplus l &= \tilde{r} \\
k &= \tilde{s} \oplus \tilde{t} \\
j \oplus m &= \tilde{u} \oplus \tilde{v} \\
i \oplus l &= \tilde{m} \oplus \tilde{p} \\
k &= \tilde{o} \\
j \oplus m &= \tilde{n} \oplus \tilde{q} \\
i \oplus j &= \tilde{b} \\
k &= x \oplus z \\
l \oplus m &= y \oplus \tilde{a} \\
i \oplus j &= \tilde{c} \oplus \tilde{d} \\
k &= \tilde{e} \oplus \tilde{f} \\
l \oplus m &= \tilde{g} \\
n &= s \\
o \oplus q &= t \oplus u \\
p \oplus r &= v \oplus w \\
n &= \tilde{h} \oplus \tilde{i} \\
o \oplus q &= \tilde{j} \\
p \oplus r &= \tilde{k} \oplus \tilde{l} \\
n &= \tilde{c} \oplus \tilde{l} \\
o \oplus q &= \tilde{g} \\
p \oplus r &= \tilde{d} \oplus \tilde{f} \\
n &= \tilde{m} \oplus \tilde{p} \\
o \oplus p &= \tilde{o} \\
q \oplus r &= \tilde{n} \oplus \tilde{q}
\end{aligned}$$

$$\begin{aligned}
n &= \tilde{r} \\
o \oplus r &= \tilde{s} \oplus \tilde{t} \\
q \oplus r &= \tilde{u} \oplus \tilde{v} \\
s &= x \oplus y \\
t \oplus v &= z \oplus \tilde{a} \\
u \oplus w &= \tilde{b} \\
s &= \tilde{r} \\
t \oplus v &= \tilde{s} \oplus \tilde{u} \\
u \oplus w &= \tilde{t} \oplus \tilde{v} \\
s &= \tilde{c} \oplus \tilde{e} \\
t \oplus w &= \tilde{g} \\
v \oplus w &= \tilde{d} \oplus \tilde{f} \\
s &= \tilde{h} \oplus \tilde{i} \\
t \oplus u &= \tilde{j} \\
v \oplus w &= \tilde{k} \oplus \tilde{l} \\
\tilde{b} &= \tilde{c} \oplus \tilde{d} \\
x \oplus z &= \tilde{e} \oplus \tilde{f} \\
y \oplus \tilde{a} &= \tilde{g} \\
x \oplus y &= \tilde{r} \\
z \oplus \tilde{a} &= \tilde{s} \oplus \tilde{u} \\
\tilde{b} &= \tilde{t} \oplus \tilde{v} \\
\tilde{c} \oplus \tilde{e} &= \tilde{h} \oplus \tilde{i} \\
\tilde{g} &= \tilde{j} \\
\tilde{d} \oplus \tilde{f} &= \tilde{k} \oplus \tilde{l} \\
\tilde{h} \oplus \tilde{k} &= \tilde{m} \oplus \tilde{n} \\
\tilde{i} \oplus \tilde{l} &= \tilde{o} \\
\tilde{j} &= \tilde{p} \oplus \tilde{q} \\
\tilde{m} \oplus \tilde{p} &= \tilde{r} \\
\tilde{o} &= \tilde{s} \oplus \tilde{t} \\
\tilde{n} \oplus \tilde{q} &= \tilde{u} \oplus \tilde{v}
\end{aligned}$$

With the aid of a computer, we can easily find the solutions to this system of equations. The free variables are a, b, i, o, $\tilde{a}$, $\tilde{c}$, $\tilde{m}$, $\tilde{s}$ and $\tilde{t}$, and we must have

$$\begin{aligned}
&c = f = g = h = k = n = s = \tilde{b} = \tilde{g} = \tilde{j} = \tilde{o} = \tilde{r} = a \oplus b \\
&j = l = a \oplus b \oplus i \\
&p = q = a \oplus b \oplus o \\
&u = v = a \oplus b \oplus t \\
&y = z = a \oplus b \oplus \tilde{a} \\
&\tilde{d} = \tilde{e} = \tilde{h} = \tilde{l} = a \oplus b \oplus \tilde{c} \\
&\tilde{n} = \tilde{p} = a \oplus b \oplus \tilde{m} \\
&\tilde{t} = \tilde{u} = a \oplus b \oplus \tilde{s}
\end{aligned} \tag{IV.20}$$

Consider section $s_{\{A,B,C\},A}$ of the original model (marked with a red circle in Figure IV.22). The only section of $\mathcal{S}^{(1)}$ at the context $\{\{A,B,C\},\{A,E,G\}\} \in \mathcal{M}^{(1)}$ that contains $s_{\{A,B,C\},A}$ is $s := (s_{\{A,B,C\},A}, s_{\{A,E,G\},A})$, whose corresponding variable is s. If we impose $s = 1$ and $t = u = v = w = 0$, we can see that these values are not consistent with the constraints (IV.20) imposed by compatibility of a presumed compatible family for cohomology. Indeed, we have

$$0 = u = a \oplus b \oplus t = s \oplus t = 1 \oplus 0 = 1.$$

This means that section $s = (s_{\{A,B,C\},A}, s_{\{A,E,G\},A})$ cannot be extended to a compatible family for $\mathcal{F}^{(1)}$. In other words, the cohomological false negative for $s_{\{A,B,C\},A}$ has vanished.

In the same way, $(s_{\{B,D,E\},E}, s_{\{C,D,E\},E})$ is the only section of $\mathcal{S}^{(1)}$ at the context $\{\{B,D,E\},\{C,D,E\}\} \in \mathcal{M}^{(1)}$ that contains both $s_{\{B,D,E\},E}$ and $s_{\{C,D,E\},E}$. The corresponding variable is h, and if we impose $h = 1$ and $g = f = 0$, we have an immediate contradiction since $h = f = g$ by (IV.20). Thus we conclude, using the same argument as before, that the cohomological false negative for the contextuality of $\mathcal{S}$ at sections $s_{\{B,D,E\},E}$ and $s_{\{C,D,E\},E}$ has vanished.

Finally, to show that we have removed the false negative for $s_{\{A,D,F\},A}$, it is sufficient to argue that $(s_{\{A,D,F\},A}, s_{\{A,E,G\},A})$ is the only section of $\mathcal{S}^{(1)}$ at the context $\{\{A,D,F\},\{A,E,G\}\} \in \mathcal{M}^{(1)}$ that contains it. The corresponding variable is n, and if we impose $n = 1$, $o = p = q = r = 0$, we have

$$0 = p = a \oplus b \oplus o = n \oplus o = 1 \oplus 0 = 1,$$

which is again a contradiction.

7. Probabilistic line models

We conclude this chapter by showing that the line model construction can be naturally generalised to probabilistic empirical models. Although the cohomological techniques at our disposal cannot be used to study probabilistic contextuality, we still believe the construction deserves to be further investigated as, just like in the possibilistic case, it exposes the hidden structure of local extendability properties of probabilistic models.

Suppose we have a probabilistic empirical model $e = \{e_C \in \mathcal{D}_R\mathcal{E}(C)\}_{C\in\mathcal{M}}$ on a measurement scenario $\langle X, \mathcal{M}, (O_m)\rangle$. We define the *first line model* of e as follows:

DEFINITION IV.37. The **first line model** of e is the probabilistic model $e^{(1)} = \{e^{(1)}_{\mathscr{C}}\}_{\mathscr{C}\in\mathcal{M}^{(1)}}$ on the scenario $\langle X, \mathcal{M}, (O_m)\rangle^{(1)}$, where, for each $\mathscr{C} = \{C^1, C^2\}$, the distribution $e^{(1)}_{\mathscr{C}} : \mathcal{E}^{(1)}(\mathscr{C}) = \mathcal{E}(C^1) \times \mathcal{E}(C^2) \to R$ is defined as follows:

$$e^{(1)}_{\mathscr{C}}(s,t) := \begin{cases} \frac{e_{C^1}(s)e_{C^2}(t)}{I} & \text{if } s|_{C^1\cap C^2} = t|_{C^1\cap C^2} \text{ and } I \neq 0 \\ 0 & \text{otherwise,} \end{cases} \tag{IV.21}$$

where $I := e_{C^1}|_{C^1\cap C^2}(s|_{C^1\cap C^2}) = e_{C^2}|_{C^1\cap C^2}(t|_{C^1\cap C^2})$.

The following proposition shows that the first line model is a well-defined empirical model over $\langle X, \mathcal{M}, (O_m)\rangle^{(1)}$.

PROPOSITION IV.38. *Let e be a probabilistic empirical model over a measurement scenario $\langle X, \mathcal{M}, (O_m)\rangle$. Then, $e^{(1)}$ is a well-defined model on $\langle X, \mathcal{M}, (O_m)\rangle^{(1)}$. Moreover, the following is true for all $\mathscr{C} = \{C^1, C^2\} \in \mathcal{M}^{(1)}$:*

$$\begin{aligned} e^{(1)}_{\mathscr{C}}|_{C^1} &= e_{C^1} \\ e^{(1)}_{\mathscr{C}}|_{C^2} &= e_{C^2}, \end{aligned} \tag{IV.22}$$

in other words, we recover the original model by marginalising the first line model to the original contexts.[5]

PROOF. We will start by proving equations (IV.22), and the result will follow immediately after. Let $\mathscr{C} = \{C^1, C^2\}$. Let

$$U := C^1 \setminus C^2, \qquad V := C^1 \cap C^2, \qquad W := C^2 \setminus C^1.$$

Then, $C^1 = U \sqcup V$ and $C^2 = W \sqcup V$. Thus,

$$\begin{aligned} e_{C^1} &\in \mathcal{D}_R\mathcal{E}(U \sqcup V) \cong \mathcal{D}_R(\mathcal{E}(U) \times \mathcal{E}(V)), \\ e_{C^2} &\in \mathcal{D}_R\mathcal{E}(W \sqcup V) \cong \mathcal{D}_R(\mathcal{E}(W) \times \mathcal{E}(V)). \end{aligned}$$

Moreover,

$$\mathcal{E}^{(1)}(\mathscr{C}) = \mathcal{E}(C^1) \times \mathcal{E}(C^2) = \mathcal{E}(U \sqcup V) \times \mathcal{E}(V \sqcup W) \cong \mathcal{E}(U) \times \mathcal{E}(V) \times \mathcal{E}(V) \times \mathcal{E}(W).$$

With this premise, one can rewrite the definition (IV.21) of $e^{(1)}_{\mathscr{C}}$ as a map

$$e^{(1)}_{\mathscr{C}} : \mathcal{E}(U) \times \mathcal{E}(V) \times \mathcal{E}(V) \times \mathcal{E}(W) \longrightarrow R$$

given by

$$e^{(1)}_{\mathscr{C}}(u, v, v', w) := \begin{cases} \frac{e_{C^1}(u,v)e_{C^2}(v,w)}{e_{C^1}|_V(v)} & \text{if } v = v' \text{ and } e_{C^1}|_V(v) \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

[5]The proper notation for equations (IV.22) should be $e^{(1)}_{\mathscr{C}}|_{\{C^i\}}$, for $i = 1, 2$. However, since $\mathcal{E}^{(1)}(\{C^1\}) = \mathcal{E}(C^1)$ by definition, we can see $e^{(1)}_{\mathscr{C}}|_{\{C^i\}}$ both as an element of $\mathcal{D}_R\mathcal{E}^{(1)}(\{C^i\})$ and as an element of $\mathcal{D}_R\mathcal{E}(C^i)$ (cf. Remark IV.9). The notation we use here is to emphasise the latter interpretation.

For any $(u,v) \in \mathcal{E}(U) \times \mathcal{E}(V)$, with $e_{C^1}|_V(v) \neq 0$, we have

$$\begin{aligned} e^{(1)}_{\mathscr{C}}|_{C^1}(u,v) &= \sum_{(v',w)\in\mathcal{E}^{(1)}(C^2)} e^{(1)}_{\mathscr{C}}(u,v,v',w) = \sum_{w\in\mathcal{E}(W)} e^{(1)}_{\mathscr{C}}(u,v,v,w) \\ &= \sum_{w\in\mathcal{E}(W)} \frac{e_{C^1}(u,v)e_{C^2}(v,w)}{e_{C^1}|_V(v)} = \frac{e_{C^1}(u,v)}{e_{C^1}|_V(v)} \sum_{w\in\mathcal{E}(W)} e_{C^2}(v,w) \\ &= \frac{e_{C^1}(u,v)}{e_{C^1}|_V(v)} e_{C^2}|_V(v) = \frac{e_{C^1}(u,v)}{e_{C^1}|_V(v)} e_{C^1}|_V(v) = e_{C^1}(u,v). \end{aligned}$$

and for $e_{C^1}|_V(v) = 0$, both sides are zero. Similarly, we prove that $e^{(1)}|_{C^2} = e_{C^2}$, to complete the proof of equations (IV.22). From this, it immediately follows that $e^{(1)}_{\mathscr{C}}$ is a distribution, indeed

$$\begin{aligned} \sum_{s\in\mathcal{E}^{(1)}(\mathscr{C})} e^{(1)}_{\mathscr{C}}(s) &= \sum_{\substack{s\in\mathcal{E}^{(1)}(\mathscr{C}) \\ s|_\emptyset = *}} e^{(1)}_{\mathscr{C}}(s) = e^{(1)}_{\mathscr{C}}|_\emptyset(*) = \left(e^{(1)}_{\mathscr{C}}|_{C^1}\right)|_\emptyset(*) = e_{C^1}|_\emptyset(*) \\ &= \sum_{\substack{s\in\mathcal{E}(C^1) \\ s|_\emptyset = *}} e_{C^1}(s) = \sum_{s\in\mathcal{E}(C^1)} e_{C^1}(s) = 1. \end{aligned}$$

Compatibility of $e^{(1)}$ is also a consequence of equations (IV.22), indeed, given two intersecting contexts $\mathscr{C}_1 = \{C^1_1, C^2_1\}$ and $\mathscr{C}_2 = \{C^1_2, C^2_2\}$,

$$e^{(1)}_{\mathscr{C}_1}|_{\mathscr{C}_1\cap\mathscr{C}_2} = e^{(1)}_{\mathscr{C}_1}|_{\{C^2_1\}} = e^{(1)}_{\mathscr{C}_1}|_{C^2_1} = e_{C^2_1} = e^{(1)}_{\mathscr{C}_2}|_{\{C^2_1\}} = e^{(1)}_{\mathscr{C}_2}|_{\mathscr{C}_1\cap\mathscr{C}_2} \tag{IV.23}$$

□

7.1. The contextuality of probabilistic line models. We investigate the contextual properties of line models.

Lemma IV.39. *Let*

$$D : \mathcal{E}^{(1)}(X^{(1)}) = \mathcal{E}^{(1)}(\mathcal{M}) = \prod_{C\in\mathcal{M}} \mathcal{E}(C) \longrightarrow R$$

be a global distribution for $e^{(1)} = \{e^{(1)}_{\mathscr{C}}\}_{\mathscr{C}\in\mathcal{M}^{(1)}}$. *Then,* $D\left(\langle s_C\rangle_{C\in\mathcal{M}}\right) = 0$ *for all* $\langle s_C\rangle_{C\in\mathcal{M}}$ *such that* $\{s_C\}_{C\in\mathcal{M}}$ *is non-compatible for* $\mathcal{E}$ *with respect to* $\mathcal{M}$.

Proof. Suppose $\{s_C\}_{C\in\mathcal{M}}$ is non compatible for $\mathcal{E}$ with respect to $\mathcal{M}$, which means that there exist $C, C' \in \mathcal{M}$ such that $s_C|_{C\cap C'} \neq s_{C'}|_{C\cap C'}$. Because of this, we must have $C \cap C' \neq \emptyset$. Therefore, $\{C, C'\} \in \mathcal{M}^{(1)}$, and we have

$$D|_{\mathscr{C}}(s_C, s_{C'}) = e^{(1)}_{\mathscr{C}}(s_C, s_{C'}) = 0,$$

where the last equality follows directly from (IV.21). Thus,

$$0 = D|_{\mathscr{C}}(s_C, s_{C'}) = \sum_{\substack{\langle p_C\rangle\in\mathcal{E}^{(1)}(\mathcal{M}) \\ p_C = s_C \\ p_{C'} = s_{C'}}} \underbrace{D\left(\langle p_C\rangle_{C\in\mathcal{M}}\right)}_{\geq 0}.$$

From this equation, it follows that

$$D\left(\langle p_C\rangle_{C\in\mathcal{M}}\right)=0\ \forall\,\langle p_C\rangle\in\mathcal{E}^{(1)}(\mathcal{M})\text{ s.t. } p_C=s_C \text{ and } p_{C'}=s_{C'}.$$

In particular, $D(\langle s_C\rangle_{C\in\mathcal{M}})=0$. $\square$

The following result shows that contextuality is preserved under the line model construction:

THEOREM IV.40. *Let $e=\{e_C\}_{C\in\mathcal{M}}$ be a probabilistic empirical model. If e is contextual, then $\mathrm{e}^{(1)}=\{e^{(1)}_{\mathscr{C}}\}_{\mathscr{C}\in\mathcal{M}^{(1)}}$ is contextual.*

PROOF. Suppose $e^{(1)}$ is non-contextual. Then, there exists a global distribution

$$D:\prod_{C\in\mathcal{M}}\mathcal{E}(C)\longrightarrow R$$

such that $D|_{\mathscr{C}}=e^{(1)}_{\mathscr{C}}$ for all $\mathscr{C}\in\mathcal{M}^{(1)}$. We define the following map:

$$d:\mathcal{E}(X)\longrightarrow R::s\longmapsto D(\langle s|_C\rangle_{C\in\mathcal{M}})$$

We will now prove that $d|_C=e_C$ for all $C\in\mathcal{M}$.

Thanks to Lemma IV.39, we can show that $d|_C=D|_{\{C\}}$ for all $C\in\mathcal{M}$. Indeed, for a given $t\in\mathcal{E}^{(1)}(\{C\})=\mathcal{E}(C)$, we have

$$D|_{\{C\}}(t)=\sum_{\substack{\langle t_C\rangle\in\mathcal{E}^{(1)}(\mathcal{M})\\ t_C=t}}D(\langle t_C\rangle)\overset{IV.39}{=}\sum_{\substack{\langle t_C\rangle\in\mathcal{E}^{(1)}(\mathcal{M})\\ t_C=t\\ \{t_C\}\text{comp.}}}D(\langle t_C\rangle)\tag{IV.24}$$

Because $\mathcal{E}$ is a sheaf, for each compatible family $\{t_C\}_{C\in\mathcal{M}}$ there exists a unique $p\in\mathcal{E}(X)$ such that $p|_C=t_C$ for all $C\in\mathcal{M}$. Thus, we have

$$\begin{aligned}D|_{\{C\}}(t)&\overset{(IV.24)}{=}\sum_{\substack{\langle t_C\rangle\in\mathcal{E}^{(1)}(\mathcal{M})\\ t_C=t\\ \{t_C\}\text{comp.}}}D(\langle t_C\rangle)=\sum_{\substack{p\in\mathcal{E}(X)\\ p|_C=t}}D(\langle p|_C\rangle)\\&=\sum_{\substack{p\in\mathcal{E}(X)\\ p|_C=t}}d(p)=d|_C(t)\end{aligned}\tag{IV.25}$$

Now, let $C\in\mathcal{M}$. By the usual assumptions, there exists a $C'\in\mathcal{M}$ such that $C\cap C'\neq\emptyset$, which means that $\mathscr{C}:=\{C,C'\}\in\mathcal{M}^{(1)}$, and we have

$$d|_C\overset{(IV.25)}{=}D|_{\{C\}}=(D|_{\mathscr{C}})\,|_{\{C\}}=e^{(1)}_{\mathscr{C}}|_{\{C\}}\overset{(IV.22)}{=}e_C.$$

Finally, we can show that d is a distribution using the same argument as in (IV.23):

$$d|_\emptyset=(d|_C)\,|_\emptyset=e_C|_\emptyset=1,$$

where C is any context in $\mathcal{M}$. We conclude that d is a global section for e, which means that e is non-contextual. $\square$

Note that the converse of this theorem is not true. In fact, it is easy to see from the results of the previous sections that it is not even true in the case of possibilistic empirical models.

Discussion

The line model construction has successfully solved the issue of false negatives for a very large class of empirical models, namely those that satisfy the CCP. This result represents a major step forward in the direction of a complete topological characterisation of possibilistic and strong contextuality.

There are strong indications suggesting that the cohomological invariant introduced in this chapter is not limited to models satisfying the CCP, but can be extended to all empirical models. In order to give rise to a false negative, a model $\mathcal{S}$ on a scenario $\langle X, \mathcal{M}, (O_m) \rangle$ would have to satisfy all of the following:

(1) $\mathcal{S}$ is contextual.
(2) $\langle X, \mathcal{M}, (O_m) \rangle$ is non-cyclic.
(3) $\mathcal{S}$ does not satisfy the CCP.
(4) $\mathcal{S}$ gives rise to a cohomology false negative in *all* its line versions at *all* its sections.

In particular, conditions (3) and (4) seem particularly difficult to be simultaneously satisfied. We already argued that it is extremely difficult, if not impossible, to identify a contextual model that does not have the CCP. On top of this, the model would need to feature a false negative at each of its sections which is preserved through all of its line models. For this reason, we put forward the following conjecture:

CONJECTURE IV.41. *Given a general model $\mathcal{S}$ on a scenario $\langle X, \mathcal{M}, (O_m) \rangle$, there exists a $k \geq 0$ such that*

$$\mathsf{LC}(\mathcal{S}, s) \Leftrightarrow \mathsf{CLC}^{(k)}(\mathcal{S}, s),$$

for all sections s of $\mathcal{S}$. In other words, the cohomology of line models represents a full invariant for contextuality.

The proof of this conjecture remains an open question.

CHAPTER V

Contextuality and generic inference: a theory of disagreement

Summary

The goal of this chapter is to establish a strong link between two apparently unrelated topics: the study of conflicting information in the formal framework of valuation algebras, and the phenomena of non-locality and contextuality. In particular, we show that these highly non-classical phenomena are mathematically equivalent to a general notion of disagreement. *This result generalises previously observed connections between contextuality, relational databases and constraint satisfaction problems, and further proves that contextuality is not a phenomenon limited to quantum mechanics, but pervades various domains of mathematics and computer science. The connection allows to translate theorems, methods and algorithms between different fields. We take advantage of this strong interaction to develop new algorithms of generic inference for the detection of non-locality and contextuality, which outperform the current methods.*

1. Overview

The high level of generality provided by the sheaf theoretic description of contextuality has sparked the establishment of a number of significant and unexpected connections with other apparently unrelated fields. Indeed, the elemental idea of contextuality as a discordance between local consistency and global inconsistency is so powerful and flexible to be abundantly observable even outside of the scope of quantum physics. The sheaf-theoretic definition of contextuality has been linked to relational database theory [**Abr13a, Bar15a**], robust constraint satisfaction [**AGK13, ABdSZ17**], natural language semantics [**AS14**], logical paradoxes [**ABK$^+$15, Kis16b, ABCP17, dS17**], and other fields [**Abr14b**]. This profusion of examples motivates the search for a general theory of *contextual semantics*, an all-comprehensive approach able to capture the essence and structure of contextual behaviour.

In this chapter, we propose such a general framework, based on the idea of *disagreement* between information sources. A natural theory to model concepts such as *information* and *knowledge* is the one of *valuation algebras*, introduced by Shenoy [**She89, SSS$^+$90**]. These abstract structures mirror the fundamental properties one naturally attributes to the notion of knowledge and capture an extremely wide range of instances, including relational databases [**KS96**], constraint satisfaction problems [**KS00**] and propositional logic [**She94, KHM99**]. Therefore, not only does the theory developed in this chapter naturally specialises to the aforementioned examples of contextuality outside quantum mechanics, but it allows to witness this phenomenon in a much larger class of fields, ranging from belief potentials to predicate logic, from linear systems to probability distributions.

The purpose of this chapter is to introduce a general vocabulary for contextual behaviour, based on the notion of *disagreement*, which can then be used to translate theorems, methods and algorithms from one field to the other. The generality of the valuation algebraic framework makes the scope for potential new results extremely wide.

The problem of detecting disagreement can be effectively modelled as an *inference problem*. Inference problems have been studied extensively in the literature, and numerous algorithms have been specifically developed to solve them. In particular, the paradigm of *local computation* – introduced by Spiegelhalter and Lauritzen [**SL88**] to solve inference problems in Bayesian networks – has proven particularly effective, and was consequently generalised to the valuation framework by the work of Shafer and Shenoy [**She89, SS91**], Kohlas [**Koh03**], and Pouly [**Pou08, Pou10**], who collectively developed the theory of *generic inference*. In this chapter, we will take advantage of generic inference to develop faster algorithms to detect contextuality.

The content of this chapter has been developed in collaboration with Samson Abramsky, and has been partially published in [**AC19**].

Outline of the chapter. In Section 2, we introduce valuation algebras and inference problems. Sections 3 and 4 are devoted to the presentation of a large amount of examples of valuation algebras. In Section 5, we present the connection between valuation prealgebras and sheaf theory. A general definition of *disagreement* is presented in Section 6, along with some key examples. Section 7 established the connection between contextuality and the general notions of local and global disagreement. In Section 8, we show that, in many relevant valuation algebras, detecting disagreement is in fact an inference problem, and introduce the concept of *complete disagreement*. Section 9 deals with the connection between disagreement and logical forms of contextuality. Finally, in Section 10, we take advantage of the general theory developed so far to develop new algorithms of generic inference to detect contextuality, and study their performance.

2. Valuation algebras and generic inference

We will adopt the language of *valuation algebras*, a general framework suited to model concepts such as *knowledge* and *information*, as well as their fundamental properties.

2.1. Valuation prealgebras.

Definition V.1. Let V be a set of variables. A **valuation prealgebra** over V is a set Φ equipped with two operations:

(1) Labelling: $\Phi \to \mathcal{P}(V) :: \phi \mapsto d(\phi)$.
(2) Projection: $\Phi \times \mathcal{P}(V) \to \Phi :: (\phi, S) \mapsto \phi^{\downarrow S}$, for all $S \subseteq d(\phi)$,

such that the following axioms are satisfied:

(A1) *Projection*: Given $\phi \in \Phi$ and $S \subseteq d(\phi)$,

$$d\left(\phi^{\downarrow S}\right) = S.$$

(A2) *Transitivity*: Given $\phi \in \Phi$ and $S \subseteq T \subseteq d(\phi)$,

$$\left(\phi^{\downarrow T}\right)^{\downarrow S} = \phi^{\downarrow S}.$$

(A3) *Domain*: Given $\phi \in \Phi$,

$$\phi^{\downarrow d(\phi)} = \phi.$$

The elements of a valuation prealgebra are called **valuations**. A set of valuations is called a **knowledgebase**. A set of variables $D \subseteq V$ is called a **domain**. The **domain of a valuation** ϕ is the set $d(\phi)$.

Intuitively, a valuation $\phi \in \Phi$ represents information about the possible values of a finite set of variables $d(\phi) = \{x_1, \dots x_n\} \subseteq V$, which constitutes the domain of ϕ. For any finite set of variables $S \subseteq V$, we denote by

$$\Phi_S := \{\phi \in \Phi \mid d(\phi) = S\}$$

the set of valuations with domain S. Thus,

$$\Phi = \bigcup_{S \subseteq V} \Phi_S.$$

The projection operation can be interpreted as the natural process of *focusing* information over a set of variables to the subset relevant for a given problem. In other words, projecting a valuation ϕ to a subset $S \subseteq d(\phi)$ corresponds to disregarding the information carried by ϕ on variables in $d(\phi) \setminus S$, which are assumed to be irrelevant.

2.2. Valuation algebras. Another important aspect of information is that local pieces of knowledge can be combined to achieve a better collective understanding of all the variables involved. This additional operation gives rise to the notion of *valuation algebra*.

DEFINITION V.2. Let V be a countable set of variables. A **valuation algebra** over V is a valuation prealgebra Φ equipped with the additional **combination** operation:

$$- \otimes - : \Phi \times \Phi \longrightarrow \Phi :: (\phi, \psi) \mapsto \phi \otimes \psi.$$

We require combination to satisfy the following axioms:

(A4) *Commutative Semigroup*: $(\Phi, \otimes)$ is associative and commutative.

(A5) *Labelling*: For all $\phi, \psi \in \Phi$,

$$d(\phi \otimes \psi) = d(\phi) \cup d(\psi)$$

(A6) *Combination*: For $\phi, \psi \in \Phi$, with $d(\phi) = S$, $d(\psi) = T$ and $U \subseteq V$ such that $S \subseteq U \subseteq S \cup T$,

$$(\phi \otimes \psi)^{\downarrow U} = \phi \otimes \psi^{\downarrow U \cap T}$$

The motivation behind axioms (A4) and (A5) can be easily explained intuitively: (A4) states that if information comes in pieces, the order of their combination should not affect the overall result. On the other hand, (A5) says that the combination of two pieces of information yields knowledge on the variables contained in the union of their domains. The combination axiom appears to be more subtle. Assume we have some knowledge about a domain in order to answer a given question. Then, (A6) states how the answer is affected if a new information piece is added. The new information can either be fully aggregated with the existing knowledge, and then projected to the specified domain, or first strapped of its irrelevant parts and then combined with the rest. These two are equivalent.

2.3. Information algebras. Besides axioms (A1)–(A6), it is often desirable to add some additional postulates, which collectively give rise to the notion of *information algebra*.

Definition V.3. Let Φ be a valuation algebra over a set of variables V.

- We say that Φ has **neutral elements** if it satisfies
  (A7) *Commutative monoid*: For each $S \subseteq V$, there exists a *neutral element* $e_S \in \Phi_S$ such that
  $$\phi \otimes e_S = e_S \otimes \phi = \phi$$
  for all $\phi \in \Phi_S$. Such neutral elements must satisfy the following identity:
  $$e_S \otimes e_T = e_{S \cup T}$$
  for all subsets $S, T \subseteq V$.
- We say that Φ has **null elements** if it satisfies
  (A8) *Nullity*: For each $S \subseteq V$ there exists a *null element* $z_S \in \Phi_S$ such that
  $$\phi \otimes z_S = z_S \otimes \phi = z_S.$$
  Moreover, for all $S, T \subseteq V$ such that $S \subseteq T$, we have, for each $\phi \in \Phi_T$,
  $$\phi^{\downarrow S} = z_S \Longleftrightarrow \phi = z_T.$$
- We say that Φ is **idempotent** if it satisfies
  (A9) *Idempotency*: For all $\phi \in \Phi$ and $S \subseteq d(\phi)$, it holds that
  $$\phi \otimes \phi^{\downarrow S} = \phi$$
- If Φ satisfies axioms (A7)–(A9), it is called an **information algebra**

The rationale behind these additional axioms is intuitively clear. Neutral elements correspond to 'irrelevant information', in the sense that they do not improve any other information we combine them with. Null elements, on the other hand, can be interpreted as destructive information, i.e. knowledge that corrupts any other valuation to the point of making it useless. Idempotency is the signature axiom of qualitative or logical, rather than quantitative, e.g. probabilistic, information. It says that counting *how many times* we have a piece of information is irrelevant.

2.4. Inference problems. If we consider valuations as pieces of information, we are naturally drawn to formulate the classic problem of extracting relevant knowledge about a given query out of a certain knowledgebase. In the valuation algebra theory, such a task is called an *inference problem*, and is formally defined as follows:

Definition V.4. Given a valuation algebra Φ, a knowledgebase $\{\phi_1, \dots, \phi_n\} \subseteq \Phi$, and a set of domains $\mathcal{D} = \{D_1, \dots, D_k\}$, with $D_i \subseteq d(\phi_1 \otimes \cdots \otimes \phi_n)$, we call an **inference problem** the task of computing

$$\phi^{\downarrow D_i} = (\phi_1 \otimes \cdots \otimes \phi_n)^{\downarrow D_i}, \ \forall 1 \leq i \leq k.$$

The valuation $\phi = (\phi_1 \otimes \cdots \otimes \phi_n)$ is called **joint valuation** or **objective function**, while each domain D_i is called a **query**. If $|\mathcal{D}| = 1$ the problem is called a **single-query inference problem**, otherwise it is called **multi-query**.

Before listing some examples, we need to introduce the concepts of *frame* and *tuple.* Consider a set of variables $V = \{x_1, x_2, \dots\}$. For each variable $x \in V$, we denote by Ω_x its **frame**, which can be thought of as a set of possible values for the variable x. A **tuple** with finite domain $S \subseteq V$ is a function

$$\mathbf{x} : S \longrightarrow \coprod_{x \in S} \Omega_x$$

that assigns to each $v \in S$ a value $\mathbf{x}(v) \in \Omega_v$. Note that, a tuple $\mathbf{x} : S \to \coprod_{x \in S} \Omega_x$ can be seen as an element of

$$\Omega_S := \prod_{x \in S} \Omega_x$$

For this reason, we will often write $\mathbf{x} \in \Omega_S$ and refer to such a tuple using the term S-tuple. Just like valuations, tuples can also be *projected.* Given an S-tuple $\mathbf{x}$ and a subset $T \subseteq S$, we denote by $\mathbf{x}_{\downarrow T}$ the tuple $\pi_T(\mathbf{x}) \in \Omega_T$, where π_T denotes the cartesian projection.[1]

3. Basic examples

Both the present section and the next are devoted to the introduction of an exhaustive list of examples of valuation and information algebras. The reason for such an extensive catalog of examples is to demonstrate the versatility of the valuation algebraic framework, which will consequently allow the reader to grasp the generality of the results presented in this chapter. This first section introduces the most elementary examples constituting the building blocks of our general discussion. We leave to the reader the verifications of all the axioms of a valuation algebra for each example (some of the proofs can be found in [**PK12**] and [**Koh12**]).

Indicator functions. An indicator function i with domain $S \subseteq V$ identifies a subset $U \subseteq \Omega_S$. More specifically, i is a map $i : \Omega_S \to \{0, 1\}$ defined by

$$i(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in U. \\ 0 & \text{if } \mathbf{x} \notin U. \end{cases}$$

Combination is defined as follows: given two indicator functions i_1, i_2 with domains S and T respectively, for all $\mathbf{x} \in \Omega_{S \cup T}$,

$$i_1 \otimes i_2(\mathbf{x}) := i_1(\mathbf{x}_{\downarrow S}) \cdot i_2(\mathbf{x}_{\downarrow T}) = \min\left\{i_1(\mathbf{x}_{\downarrow S}), i_2(\mathbf{x}_{\downarrow T})\right\}.$$

[1] We have used the different notation $(\cdot)_{\downarrow(-)}$ for the projection of tuples to distinguish it from the projection $(\cdot)^{\downarrow(-)}$ of valuation algebras.

Given an indicator i with domain S, a subset $T \subseteq S$, and a T-tuple $x \in \Omega_T$, we define the projection of i to T by maximisation:

$$i^{\downarrow T}(\mathbf{x}) := \max_{\mathbf{y} \in \Omega_{S \setminus T}} i(\mathbf{x}, \mathbf{y}). \tag{V.1}$$

For each $S \subseteq V$, the neutral element $e_S \equiv 1$ is the function that assigns 1 to each $\mathbf{x} \in \Omega_S$. The null element of Φ_S is the constant function $z_S \equiv 0$, which assigns 0 to each $\mathbf{x} \in \Omega_S$. It can be easily shown that this algebra is idempotent, which implies that it constitutes an information algebra.

We can use tables as a handy representation of indicators. Indeed, an indicator function with domain S can be seen as an $|S|$-dimensional table with $|\Omega_S|$ entries in $\{0, 1\}$. This is shown in the following example, which also introduces the first instance of an inference problem.

Example V.5. Suppose we have a set of variables $V = \{a_1, b_1, a_2, b_2\}$ with frames $\Omega_x := \{0, 1\}$ for all $x \in V$. Suppose we have the following indicator functions:

$$i_1 := \begin{array}{cc|c} a_1 & b_1 & \\ \hline 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{array} \qquad i_2 := \begin{array}{cc|c} a_1 & b_2 & \\ \hline 0 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{array}$$

$$i_3 := \begin{array}{cc|c} a_2 & b_1 & \\ \hline 0 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{array} \qquad i_4 := \begin{array}{cc|c} a_2 & b_2 & \\ \hline 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{array}$$

We are interested in solving the following inference problem:

$$(i_1 \otimes i_2 \otimes i_3 \otimes i_4)^{\downarrow \{a_2, b_2\}}.$$

Notice that we already have information about the query $\{a_2, b_2\}$ given by the indicator i_4. However, we want to take into account all the available knowledge to give a precise answer. We have

$$i_1 \otimes i_2 \otimes i_3 \otimes i_4 = \begin{array}{cccc|c} a_1 & b_1 & a_2 & b_2 & \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{array}$$

From here, we can compute

$$(i_1 \otimes i_2 \otimes i_3 \otimes i_4)^{\downarrow\{a_2,b_2\}} = \begin{array}{cc|c} a_2 & b_2 & \\ \hline 0 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{array}$$

Notice that the final result is different from i_4. This can be interpreted as an inconsistency in the information shared by i_1, i_2, i_3, i_4 on the variables a_2, b_2. This kind of inconsistency corresponds to the more general notion of *disagreement*, which will be defined later, in Section 6. The careful reader will have noticed that the valuations taken into account in this example correspond to the possibilistic distributions over the four contexts of the Hardy model. This is a hint of what we will present in the next sections: a formal description of contextuality as a particular case of general notion of disagreement.

Relational databases. Consider the data table V.1, taken from [**Abr13a**].

TABLE V.1. A simple data table from [**Abr13a**]

branch-name	account-no	customer-name	balance
Cambridge	10991-06284	Newton	2,567.53
Hanover	10992-35671	Leibniz	11,245.75
...	...	...	...

We identify the list of **attributes**

$$\{\text{branch-name, account-no, customer-name, balance}\}$$

labelling the columns of the table. In the language of valuation algebras, attributes correspond to variables. Each entry of the table is a tuple specifying a value for each of the attributes. Thus, the full table is nothing but a set of tuples, which corresponds to the notion of *relation* in database theory. A database normally consists of a set of such relations.

Abstracting from this example, given a set of attributes V, we define a relation over $S \subseteq V$ as a set $R \subseteq \Omega_S$. The domain of R, often called *schema* of R in database theoretic terms, is then $d(R) = S$. Combination is given by the *natural join*: Let R_1 and R_2 be two relations with domain S and T respectively, then

$$R_1 \otimes R_2 := R_1 \bowtie R_2 = \{\mathbf{x} \in \Omega_{S \cup T} \mid \mathbf{x}_{\downarrow S} \in R_1 \ \wedge \ \mathbf{x}_{\downarrow T} \in R_2\}.$$

Given a relation R with domain $d(R) = S$, and a subset $T \subseteq S$, we define projection as follows:

$$R^{\downarrow T} := \{\mathbf{x}_{\downarrow T} \mid \mathbf{x} \in R\}. \tag{V.2}$$

For each $S \subseteq V$, we define the neutral element as $e_S := \Omega_S$. On the other hand, the null element is defined as $z_S := \emptyset$. One can verify that this algebra is idempotent, and thus constitutes an information algebra.

EXAMPLE V.6. Let us give an example of a simple inference problem. Consider a database comprised of three relations $R_{1,2,3}$ which collect information about the geographical locations of cities and countries:

$R_1 =$

continent	country
Europe	UK
America	USA
Europe	FRA
Asia	CHI
...	...

$R_2 =$

country	city
FRA	Paris
GER	Frankfurt
ESP	Valencia
ITA	Milan
...	...

$R_3 =$

continent	city
Europe	Paris
Africa	Algeri
Asia	Beijing
America	Miami
...	...

Now, suppose that whoever collected the data did not address the issue of different cities having the same name, so that $R_{1,2,3}$ may contain the following sub-relations:

$\tilde{R}_1 =$

continent	country
America	CAN
Europe	UK

$\tilde{R}_2 =$

country	city
CAN	London
UK	Boston

$\tilde{R}_3 =$

continent	city
Europe	London
America	Boston

where $\tilde{R}_3$ contains information about the cities of London (UK) and Boston (USA), while $\tilde{R}_2$ refers to the smaller cities of London, Ontario (CA), and Boston, Lincolnshire (UK). Then, if we combine $\tilde{R}_1$ with $\tilde{R}_2$, we obtain.

$\tilde{R}_1 \otimes \tilde{R}_2 =$

continent	country	city
America	Canada	London
Europe	UK	Boston

By combining this with $\tilde{R}_3$ we obtain an empty table, that is $\tilde{R}_1 \otimes \tilde{R}_2 \otimes \tilde{R}_3 = \emptyset$. Just like in the example of the previous paragraph on indicator functions, this fact can be interpreted as *disagreement* among the different sources of information, which in this case is clearly generated by the issue of different cities with the same name. This kind of inconsistency, i.e. the non-existence of a *global relation* is precisely the notion that was proven to be equivalent to contextuality in [**Abr13a, Bar15a**]. In fact, notice how the database comprised of $\tilde{R}_{1,2,3}$ can be visually represented as a bundle diagram which is equivalent to the one of the Specker's triangle (Figure II.5), as shown in Figure V.1. One of the goals of this chapter is to extend this connection to the more general level of

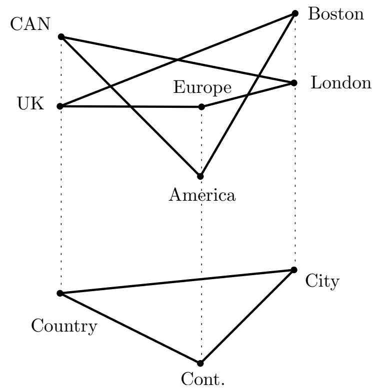


FIGURE V.1. A bundle representation of the database $\{\tilde{R}_{1,2,3}\}$. Equivalent to the Specker's triangle (Figure II.5).

valuation algebras.

REMARK V.7. The valuation algebra of relational databases is in fact equivalent to the one of indicator functions. Indeed, a relation $R \subseteq \Omega_S$ in Φ_S can be interpreted as an indicator function via the following map

$$f_S : \mathcal{P}(\Omega_S) \longrightarrow 2^{\Omega_S} :: R \longmapsto i_R, \tag{V.3}$$

where $i_R(\mathbf{x}) = 1$ if and only if $\mathbf{x} \in R$. Moreover, the maps f_S preserve both composition and projection, i.e. for all $R_1 \in \Phi_S$ and $R_2 \in \Phi_T$,

$$f_{S\cup T}(R_1 \otimes R_2) = f_S(R_1) \otimes f_T(R_2),$$

and, for all $R \in \Phi_S$ and $T \subseteq S$,

$$f_T(R^{\downarrow T}) = (f_S(R))^{\downarrow T}.$$

Therefore, all the operations coincide, and the two valuation algebras are equivalent.

Semiring valuation algebras. Let $\langle R, +, \cdot, 0, 1\rangle$ be a commutative semiring and V a set of variables. A semiring valuation with domain $S \subseteq V$ is a function

$$\phi : \Omega_S \longrightarrow R.$$

We define the three operation as follows:

(1) *Labelling*: $d : \Phi \to \mathcal{P}(V)$, with $d(\phi) = S$ if $\phi \in \Phi_S$.
(2) *Combination*: $\otimes : \Phi \times \Phi \to \Phi$, where, for all $\phi, \psi \in \Phi$ and $\mathbf{x} \in \Omega_{d(\phi)\cup d(\psi)}$, we have

$$(\phi \otimes \psi)(\mathbf{x}) := \phi\left(\mathbf{x}_{\downarrow d(\phi)}\right) \cdot \psi\left(\mathbf{x}_{\downarrow d(\psi)}\right).$$

(3) *Projection*: $\downarrow$: $\Phi \times \mathcal{P}(V) \to \Phi$, where, for all $\phi \in \Phi$, $T \subseteq d(\phi)$ and $\mathbf{x} \in \Omega_T$, we have

$$\phi^{\downarrow T}(\mathbf{x}) := \sum_{\mathbf{y} \in \Omega_{S \setminus T}} \phi(\mathbf{x}, \mathbf{y}).$$

The resulting valuation algebra is idempotent only when R is idempotent. For instance, when $R = \mathbb{B}$, in which case it coincides with the one of indicator functions. The neutral element $e_S \equiv 1$ is the function that assigns 1 to each $\mathbf{x} \in \Omega_S$. The null element $z_S \equiv 0$ is the 0-function.

Semiring valuations are often referred to as *arithmetic potentials*, or R-**potentials**. In the case where $R = \mathbb{R}_{\geq 0}$, the corresponding valuation algebra is the one of **probability potentials**.

Set potentials - belief functions. Set potentials are a general framework useful to model information in the presence of uncertainty. Similarly to the case of semiring valuation algebras, one can define a valuation algebra of set potentials for every choice of a semiring. Let R be a semiring. Given a set of variables V, a *set potential* with domain $d(m) = S \subseteq V$ is a map

$$m : \mathcal{P}(\Omega_S) \longrightarrow R$$

We can define the combination of two set potentials m_1 and m_2 with domains S and T respectively as follows: for all $U \subseteq \Omega_{S \cup T}$,

$$m_1 \otimes m_2(U) := \sum_{\substack{W \subseteq \Omega_S, X \subseteq \Omega_T \\ W \bowtie X = U}} m_1(W) \cdot m_2(X).$$

Given a set potential m and a subset $T \subseteq d(m)$, we define, for all $U \subseteq \Omega_T$,

$$m^{\downarrow T}(U) := \sum_{\substack{W \subseteq \Omega_S \\ W_{\downarrow T} = U}} m(W),$$

where $W_{\downarrow T} = \{\mathbf{x}_{\downarrow T} \mid \mathbf{x} \in W\}$. Similarly to arithmetic potentials, this valuation algebra is idempotent only when R is idempotent. The neutral element e_S for a subset $S \subseteq V$ is defined as follows: for all $U \subseteq \Omega_S$,

$$e_S(U) := \delta_{U, \Omega_S} = \begin{cases} 1 & \text{if } U = \Omega_S \\ 0 & \text{otherwise.} \end{cases}$$

The null element is the constant function $z_S \equiv 0$, which assigns 0 to each subset $U \subseteq \Omega_S$.

In the special case where $R = \mathbb{R}_{\geq 0}$, set potentials are called **belief functions** (inference over belief functions is also known as Dempster–Shafer theory [**Dem08, Sha76**]).

4. Advanced examples: language and models

A special class of examples, which deserves particular attention, is the one arising from logic. In this section, we will start by presenting the valuation algebras associated to propositional logic. Then, we will generalise this example to a vast class of algebras arising from general notions of *language* and *models*.

4.1. Propositional logic. Often, information concerns the truth value of logical propositions. This is the main focus of propositional logic, which can be elegantly captured by the framework of valuation algebras. Let us start by briefly reviewing some of the basic definitions.

4.1.1. Syntax and semantics. Let $P := \{p_1, p_2, \dots\}$ be a countable set of *propositional symbols*. The language of propositional logic $\mathcal{L}_P$ over P consists of *well-formed formulae* (wff), defined as follows:

(1) Each element $p_i \in P$ and the symbols $\top$, $\bot$ are wffs and are called *atomic*.
(2) If φ is a wff, then $\neg\varphi$ is a wff.
(3) If φ and ψ are wffs, then $\varphi \wedge \psi$ is a wff
(4) If φ is a wff and $p_i \in P$, then $(\exists p_i)\varphi$ is a wff.

Every wff is generated from atomic formulae via a finite number of applications of rules 2,3 and 4. For convenience, the following operations are customarily added to the language:

- $\varphi \vee \psi := \neg(\neg\varphi \wedge \neg\psi)$
- $\varphi \rightarrow \psi := \neg\varphi \vee \psi$
- $\varphi \leftrightarrow \psi := (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$

For each subset $Q \subseteq P$ of the symbols, one can define the language $\mathcal{L}_Q \subseteq \mathcal{L}_P$ by simply replacing P by Q in rule (1).[2]

The semantics of the language is represented by *valuations*, i.e. maps $v : P \to \{0, 1\}$. Valuations can be extended to propositional formulae $\hat{v} : \mathcal{L}_P \to \{0, 1\}$ with the following rules:

(1) $\hat{v}(\top) = 1$ and $\hat{v}(\bot) = 0$
(2) $\hat{v}(p_i) = v(p_i)$ for all $p_i \in P$
(3) $\hat{v}(\neg\varphi) = \begin{cases} 1 & \text{if } \hat{v}(\varphi) = 0 \\ 0 & \text{if } \hat{v}(\varphi) = 1 \end{cases}$
(4) $\hat{v}(\varphi \wedge \psi) = \begin{cases} 1 & \text{if } \hat{v}(\varphi) = 1 \text{ and } \hat{v}(\psi) = 1 \\ 0 & \text{otherwise} \end{cases}$
(5) $\hat{v}((\exists p)\varphi) = \hat{v}(\varphi[p/\top] \vee \varphi[p/\bot])$

where $\varphi[p/q]$ denotes the formula obtained by replacing all occurrencies of the symbol p in φ by $\top$. We say that a valuation v *satisfies* a propositional formula φ if $\hat{v}(\varphi) = 1$. Then, v is called a *model* of φ, and we write $v \models \varphi$.

We define, for every set Γ of formulas, the set of models for the formulae contained in Γ:

$$\mathscr{M}(\Gamma) := \{v : v \models \varphi, \ \forall\varphi \in \Gamma\}.$$

In a similar way, for each set S of valuations, let

$$\mathscr{T}(S) := \{\varphi : v \models \varphi, \ \forall v \in S\}$$

denote the set of formulae satisfied by all the valuations in S. Both $\mathscr{M}$ and $\mathscr{T}$ can be seen as order reversing maps between the powerset of $\mathcal{L}_P$ and the one of valuations:

[2]Note that we explicitly do not require the same for rule (4), so that, if $\phi \in \mathcal{L}_Q$ only contains variables in Q and $p_{i_1}, \dots, p_{i_n} \in P \setminus Q$, then $\exists p_{i_1} \dots p_{i_n}\varphi$ is still a formula of $\mathcal{L}_Q$.

$$\begin{tikzcd} \mathcal{P}(\mathcal{L}_P) \arrow[r, bend left, "\mathscr{M}"] & \mathcal{P}\left(\mathcal{P}(\mathcal{L}_P)\right) \arrow[l, bend left, "\mathscr{T}"] \end{tikzcd}.$$

Together, $\mathscr{M}$ and $\mathscr{T}$ constitute a *Galois connection*, which describes the relation between syntax and semantics [**EKMS92**]:

Definition V.8. Given two partially ordered sets A and B, an (antitone) **Galois connection** between A and B is a pair $\langle f, g\rangle$ of order-reversing maps $f : A \to B, g : B \to A$ such that $a \leq g \circ f(a)$ for all $a \in A$, and $b \leq f \circ g(b)$ for all $b \in B$.

4.1.2. The information algebra of propositional information. We shall now introduce an information algebra suited to describe propositional logic.

Consider the lattice $\mathscr{L}$ of finite subsets of P. We interpret each element $Q \subseteq P$ of $\mathscr{L}$ as a question regarding the truth values of the propositions $p_i \in Q$.

Given a valuation $v : P \to \{0,1\}$ and a subset $Q \in \mathscr{L}$, the restriction $v|_Q : Q \to \{0,1\}$ is called an *interpretation* of the language $\mathcal{L}_Q \subseteq \mathcal{L}_P$. We denote by $\mathcal{I}_Q$ the set of possible interpretations of $\mathcal{L}_Q$. For $w \in \mathcal{I}_Q$ and a wff φ of $\mathcal{L}_Q$, we write $w \models_Q \varphi$ if $w = v|_Q$ and $v \models \varphi$ for some $v : P \to \{0,1\}$. Then, we extend the definitions of $\mathscr{M}$ and $\mathscr{T}$ to interpretations: for all subsets $Q \subseteq P$ and $\Gamma \subseteq \mathcal{L}_Q$, let

$$\mathscr{M}_Q(\Gamma) := \{w \in \mathcal{I}_Q : w \models_Q \varphi,\ \forall \varphi \in \Gamma\}.$$

Similarly, for all $S \subseteq \mathcal{I}_Q$, let

$$\mathscr{T}_Q(S) := \{\varphi \in \mathcal{L}_Q : w \models_Q \varphi,\ \forall w \in S\}.$$

Once again, the operators $\mathscr{M}_Q$ and $\mathscr{T}_Q$ constitute a Galois connection.

The subsets of $\mathcal{I}_Q$ are called **information sets**, and they constitute a valuation algebra with the following operations:

- *Labelling*: Given an information set $M \subseteq \mathcal{I}_Q$, its label is defined to be
$$d(M) = Q.$$
- *Combination*: For all $M_1 \subseteq \mathcal{I}_Q$ and $M_2 \subseteq \mathcal{I}_U$, let
$$M_1 \otimes M_2 := M_1 \bowtie M_2 = \{w \in \mathcal{I}_{Q\cup U} : w|_Q \in M_1 \wedge w|_U \in M_2\}.$$
- *Projection*: Given an information set M and a domain $Q \subseteq d(M)$, we define
$$M^{\downarrow Q} := \{w|_Q : w \in M\}.$$

The neutral elements of the algebra are the sets $e_Q := \mathcal{I}_Q$, while the null elements are $z_Q := \emptyset$. The algebra is clearly idempotent, hence it is an information algebra.

4.1.3. Algebra of sentences. In addition to the algebra of propositional information sets, one can define an information algebra of propositional formulae.

Suppose we have two information sets M_1, M_2 of the form $M_1 = \mathscr{M}_Q(\Gamma_1)$ and $M_2 = \mathscr{M}_U(\Gamma_2)$ for $\Gamma_1 \subseteq \mathcal{L}_Q$ and $\Gamma_2 \subseteq \mathcal{L}_U$. Then, the formulae of Γ_1 and Γ_2 can be seen as sentences of the language $\mathcal{L}_{Q\cup U}$, enabling us to define the *vacuous extension* of $M \subseteq \mathcal{I}_Q$ to $Q \cup U$:

$$M^{\uparrow Q\cup U} := \{w \in \mathcal{I}_{Q\cup U} : w|_Q \in M\}.$$

It is easy to see that, for $M = \mathscr{M}_Q(\Gamma)$, we have $M^{\uparrow Q\cup U} = \mathscr{M}_{Q\cup U}(\Gamma)$. Therefore,

$$M_1 \otimes M_2 = M_1^{\uparrow Q\cup U} \cap M_2^{\uparrow Q\cup U} = \mathscr{M}_{Q\cup U}(\Gamma_1) \cap \mathscr{M}_{Q\cup U}(\Gamma_2) = \mathscr{M}_{Q\cup U}(\Gamma_1 \cup \Gamma_2).$$

Moreover, if $M = \mathscr{M}_U(\Gamma)$ for some set of formulae $\Gamma \in \mathcal{L}_U$, we have, for all $Q \subseteq U$,

$$M^{\downarrow Q} = \{w|_Q : w \in \mathscr{M}_U(\Gamma)\} = \mathscr{M}_Q((\exists p_{i_1}, \dots, p_{i_n})\Gamma)$$

where $U \setminus Q = \{p_{i_1}, \dots p_{i_n}\}$, and $(\exists p_{i_1}, \dots, p_{i_n})\Gamma := \{(\exists p_{i_1}, \dots, p_{i_n})\,\varphi : \varphi \in \Gamma\}$. In other words, projection corresponds to existential quantification of the original information.

This observation is the key to the definition of an algebra of sentences associated to the algebra of information sets. Indeed, to each $M = \mathscr{M}_Q(\Gamma)$ of $\mathcal{I}_Q$ one can associate its theory $\mathscr{T}_Q(M) = \mathscr{T}_Q(\mathscr{M}_Q(\Gamma))$. Because $\mathscr{T}_Q$ and $\mathscr{M}_Q$ constitute a Galois connection, one can show that the operator $\mathcal{C}_Q := \mathscr{T}_Q \circ \mathscr{M}_Q$ is a *closure operator*, i.e. it satisfies the following:

$$\begin{aligned}
&\Gamma \subseteq \mathcal{C}_Q(\Gamma),\\
&\Gamma_1 \subseteq \Gamma_2 \Longrightarrow \mathcal{C}_Q(\Gamma_1) \subseteq \mathcal{C}_Q(\Gamma_2),\\
&\mathcal{C}_Q(\mathcal{C}_Q(\Gamma)) = \mathcal{C}_Q(\Gamma).
\end{aligned}$$

For this reason, subsets $\Gamma \subseteq \mathcal{L}_Q$ of the form $\Gamma = \mathcal{C}_Q(\Gamma)$ are called *closed.* The algebra of formulae is thus in fact an algebra of closed sets of sentences, with the following operations.

- *Labelling*: Given a closed set of formulae $\Gamma \subseteq \mathcal{L}_Q$, we define
$$d(\Gamma) := Q.$$
- *Combination*: Given two closed sets $\Gamma_1 \subseteq \mathcal{L}_Q$ and $\Gamma_2 \subseteq \mathcal{L}_U$, we define
$$\Gamma_1 \otimes \Gamma_2 := \mathcal{C}_{Q\cup U}(\Gamma_1 \cup \Gamma_2).$$
- *Projection*: For $\Gamma \subseteq \mathcal{L}_U$ and $Q \subseteq U$, let
$$\Gamma^{\downarrow Q} := \mathcal{C}_Q((\exists p_{i_1} \dots p_{i_n})\Gamma),$$
where $U \setminus Q = \{p_{i_1} \dots p_{i_n}\}$.

The neutral elements of this algebra are the tautologies $e_Q := \mathcal{C}_Q(\emptyset)$, while the null elements are $z_Q := \mathcal{L}_Q$. The combination operation is idempotent, thus the algebra of sentences is an information algebra.

4.2. General languages. The concepts of *language* and *model* encountered in propositional logic can be seen in more general terms. A language $\mathcal{L}$ could be simply defined as a set of possible sentences, without considering their syntactic structure. Similarly, information can be seen as the set of possible answers $\mathbb{M}$ (or models) to a set of questions in $\mathcal{L}$, regardless of the way such answers are formulated. In order to capture the idea of a model $m \in M$ being an answer to $s \in \mathcal{L}$, we assume a binary relation $\models \subseteq \mathcal{L} \times \mathbb{M}$. Much of the structure of propositional information can be recovered in this general setting. This idea is expressed, in the valuation algebra literature, through the concept of *context.*[3] A context is a triple $\langle \mathcal{L}, \mathbb{M}, \models \rangle$ and its function is to express

[3]This term shall not be confused with the notion of context in contextuality.

information carried by sentences. Given a set of sentences $S \subseteq \mathcal{L}$ and a set of models $M \subseteq \mathbb{M}$, we define:

$$\mathscr{M}(S) := \{m \in \mathbb{M} : m \models s, \forall s \in S\},$$
$$\mathscr{T}(M) := \{s \in \mathcal{L} : m \models s, \forall m \in M\}.$$

As before, $\mathscr{M}$ and $\mathscr{T}$ constitute a Galois connection, and thus their composition $\mathcal{C} := \mathscr{T} \circ \mathscr{M}$ is a closure operator.

In propositional logic, we have a family of contexts $\{\langle \mathcal{L}_Q, \mathbb{M}_Q, \models_Q \rangle\}_{Q \subseteq P}\}$ which are linked together by restriction of interpretations and embeddings of formulae. We will now introduce a generic construction which abstracts this situation and extends it to general languages. Let us start by introducing the concept of *tuple system*, which subsumes the properties of interpretations in propositional logic.

Definition V.9. A **tuple system** over $\mathcal{P}(V)$, where V is a set of variables, is a set T equipped with two operations $d : T \to \mathcal{P}(V)$ and $\downarrow : T \times \mathcal{P}(V) \to T$ satisfying the following axioms:

(T1) If $Q \subseteq d(\mathbf{t})$, then $d(\mathbf{t}_{\downarrow Q}) = Q$.
(T2) If $Q \subseteq U \subseteq d(\mathbf{t})$, then $\left(\mathbf{t}_{\downarrow U}\right)_{\downarrow Q} = \mathbf{t}_{\downarrow Q}$.
(T3) If $d(\mathbf{t}) = Q$, then $\mathbf{t}_{\downarrow Q} = \mathbf{t}$.
(T4) For $d(\mathbf{t}) = Q$, $d(\mathbf{u}) = U$ such that $\mathbf{t}_{\downarrow Q \cap U} = \mathbf{u}_{\downarrow Q \cap U}$, there exists $\mathbf{g} \in T$ such that $d(\mathbf{g}) = Q \cup U$, $\mathbf{g}_{\downarrow Q} = t$ and $\mathbf{g}_{\downarrow U} = \mathbf{u}$.
(T5) For $d(\mathbf{t}) = Q$ and $Q \subseteq U$, there exists $\mathbf{g} \in T$ such that $d(\mathbf{g}) = U$ and $\mathbf{g}_{\downarrow Q} = \mathbf{t}$.

Apart from cartesian tuples, which constitute the most basic instance of a tuple system, many other frameworks can be captured by this notion. Here are two examples which are relevant for our study:

Example V.10. As mentioned above, propositional interpretations constitute a tuple system. Let T be the set of all interpretation of language of propositional logic $\mathcal{L}_P$. Given an interpretation $w \in \mathcal{I}_Q \subseteq T$, we define $d(w) := Q$, while the operation $\downarrow$ coincides with function restriction. We will leave to the reader the verification of axioms (T1)-(T5).

Example V.11. A more complex example of a tuple system is the one of *distributions over a semiring*. Let R be a semiring, and suppose we have a set of variables V with frames Ω_v for each $v \in V$. Let $T := \bigcup_{U \subseteq V} \mathcal{D}_R(\Omega_U)$ be the set of all R-distributions over tuples on subsets of V. We shall now prove that T is a tuple system.

The domain $d(D)$ of a distribution $D \in \mathcal{D}_R(\Omega_U)$ is defined to be U, whereas projection $\downarrow$ is defined as follows: for all $Q \subseteq U$ and $D \in \mathcal{D}_R(\Omega_U)$, for all $\mathbf{x} \in \Omega_Q$,

$$D_{\downarrow Q}(\mathbf{x}) := \sum_{\substack{\mathbf{y} \in \Omega_U : \\ \mathbf{y}_{\downarrow Q} = \mathbf{x}}} D(\mathbf{y}).$$

Axioms (T1)-(T5) can be verified as follows:

(T1): Let $D \in \mathcal{D}_R(\Omega_U)$, and $Q \subseteq U$. We have $|\mathsf{supp}(D_{\downarrow Q})| < \infty$ by definition of $D_{\downarrow Q}$ and the fact that $|\mathsf{supp}(D)| < \infty$. Moreover,

$$\sum_{\mathbf{x} \in \Omega_Q} D_{\downarrow Q}(\mathbf{x}) = \sum_{\mathbf{x} \in \Omega_Q} \sum_{\substack{\mathbf{y} \in \Omega_U \\ \mathbf{y}_{\downarrow Q} = \mathbf{x}}} D(\mathbf{y}) = \sum_{\mathbf{y} \in \Omega_U} D(\mathbf{y}) = 1.$$

Therefore, $D_{\downarrow Q}(\mathbf{x}) \in \mathcal{D}_R(\Omega_Q)$, which means that $d(D_{\downarrow Q}(\mathbf{x})) = Q$.

(T2): Let $D \in \mathcal{D}_R(\Omega_S)$, and $Q \subseteq U \subseteq S$. For all $\mathbf{x} \in \Omega_Q$,

$$\begin{aligned} \left(D_{\downarrow U}\right)_{\downarrow Q}(\mathbf{x}) &= \sum_{\substack{\mathbf{y} \in \Omega_U \\ \mathbf{y}_{\downarrow Q} = \mathbf{x}}} D_{\downarrow U}(\mathbf{y}) = \sum_{\substack{\mathbf{y} \in \Omega_U \\ \mathbf{y}_{\downarrow Q} = \mathbf{x}}} \sum_{\substack{\mathbf{z} \in \Omega_S \\ \mathbf{z}_{\downarrow U} = \mathbf{y}}} D(\mathbf{z}) \\ &= \sum_{\substack{\mathbf{z} \in \Omega_S \\ \mathbf{z}_{\downarrow Q} = \mathbf{x}}} D(\mathbf{z}) = D_{\downarrow Q}(\mathbf{x}) \end{aligned}$$

(T3): Let $D \in \mathcal{D}_R(\Omega_Q)$. For all $\mathbf{x} \in \Omega_Q$,

$$D_{\downarrow Q}(\mathbf{x}) = \sum_{\substack{\mathbf{y} \in \Omega_Q \\ \mathbf{y}_{\downarrow Q} = \mathbf{x}}} D(\mathbf{y}) = D(\mathbf{x}).$$

(T4): The proof of this axiom coincides with the proof of equations (IV.22) in Proposition IV.38, up to minor alterations.

(T5): Let $D \in \mathcal{D}_R(\Omega_Q)$, and $Q \subseteq U$. Let $\mathbf{x} \in \Omega_Q$. By axiom (T5) of the tuple system of cartesian tuples, we know that there exists a tuple $\mathbf{g}^{\mathbf{x}} \in \Omega_U$ such that $\mathbf{g}^{\mathbf{x}}_{\downarrow Q} = \mathbf{x}$. This gives us a subset $S := \{\mathbf{g}^{\mathbf{x}}\}_{\mathbf{x} \in \Omega_Q} \subseteq \Omega_U$. We define a distribution $G : \Omega_U \to R$ as follows: for all $\mathbf{y} \in \Omega_U$,

$$G(\mathbf{y}) := \begin{cases} D(\mathbf{x}) & \text{if } \mathbf{y} = \mathbf{g}^{\mathbf{x}} \in S \\ 0 & \text{otherwise.} \end{cases}$$

Then, G is a distribution, indeed

$$\sum_{\mathbf{y} \in \Omega_U} G(\mathbf{y}) = \sum_{\mathbf{g}^{\mathbf{x}} \in \Omega_U} G(\mathbf{g}^{\mathbf{x}}) = \sum_{\mathbf{x} \in \Omega_Q} D(\mathbf{x}) = 1.$$

Moreover, for all $\mathbf{z} \in \Omega_Q$,

$$G_{\downarrow Q}(\mathbf{z}) = \sum_{\substack{\mathbf{y} \in \Omega_U : \\ \mathbf{y}_{\downarrow Q} = \mathbf{z}}} G(\mathbf{y}) = \sum_{\substack{\mathbf{g}^{\mathbf{x}} \in S : \\ \mathbf{g}^{\mathbf{x}}_{\downarrow Q} = \mathbf{z}}} G(\mathbf{g}^{\mathbf{x}}) = G(\mathbf{g}^{\mathbf{z}}) = D(\mathbf{z}).$$

The general concept of tuple system is the key to the generalisation of the algebra of propositional information sets and propositional formulae.

4.2.1. Information algebra of general information sets. Suppose $\mathbb{M}$ is a tuple system over a set of variables V. For each subset $S \subseteq V$ we define

$$\mathbb{M}_S := \{\mathbf{m} \in \mathbb{M} : d_{\mathbb{M}}(m) = S\}.$$

The valuation algebra structure of propositional information sets introduced in Section 4.1 can be generalised. We call an *information set* a subset $M \subseteq \mathbb{M}_Q$ for some $Q \subseteq V$. Given two information sets $M_1 \subseteq \mathbb{M}_Q$ and $M_2 \subseteq \mathbb{M}_U$, we define their combination

$$M_1 \otimes M_2 := \{\mathbf{m} \in \mathbb{M}_{Q\cup U} : \mathbf{m}_{\downarrow Q} \in M_1 \wedge \mathbf{m}_{\downarrow U} \in M_2\}.$$

Projection of an information set $M \subseteq \mathbb{M}_U$ to a subset $Q \subseteq U$ is given by

$$M^{\downarrow Q} := \{\mathbf{m}_{\downarrow Q} : \mathbf{m} \in M\}.$$

The labelling function d is defined in the obvious way (i.e. $d(M) = Q \Leftrightarrow Q \in \mathbb{M}_Q$). The algebra is idempotent and has both neutral elements $e_Q = \mathbb{M}_Q$, and null elements $z_Q = \emptyset$ for all $Q \subseteq V$. It is thus an information algebra.

Example V.12.

- The valuation algebra of relational databases can be seen as a special case of the algebra we have just defined. The tuple system in question is the one of cartesian tuples: $\mathbb{M}_S := \Omega_S$.
- The propositional information algebra is based on the tuple system of interpretations, introduced in Example V.10. In this case, $\mathbb{M}_S = \mathcal{I}_S$.
- By taking the tuple set of semiring distributions defined in Example V.11, we obtain the information algebra of information sets constituted by

$$\mathbb{M}_S = \mathcal{D}_R(\Omega_S),$$

where $S \subseteq V$.

4.2.2. Information algebra of general sentences. Given a tuple-system, and assuming the existence of a language expressing information in it, it is often possible to construct an algebra of formulae associated to the algebra of general information sets. In this paragraph, we list the necessary conditions for its definition.

Let $\mathbb{M}$ be a tuple-system over a set of variables V. For all $Q \subseteq V$, we assume the existence of a language $\mathcal{L}_Q$ to express information in $\mathbb{M}_Q$. More specifically, we suppose a sequence of logical contexts

$$\{\langle \mathcal{L}_Q, \mathbb{M}_Q, \models_Q\rangle\}_{Q\subseteq V}.$$

Moreover, we assume that these nested logical contexts are related by embeddings which preserve information in the following sense: for all $Q \subseteq U \subseteq V$ we require the existence of an embedding $f_{Q,U} : \mathcal{L}_Q \longrightarrow \mathcal{L}_U$ such that the pair $\langle f_{Q,U}, g_{U,Q}\rangle$, where

$$g_{U,Q} : \mathbb{M}_U \longrightarrow \mathbb{M}_Q :: \mathbf{m} \longmapsto \mathbf{m}_{\downarrow Q},$$

constitutes an *infomorphism*, i.e. it satisfies

$$\mathbf{m} \models_U f_{Q,U}(\mathbf{s}) \Longleftrightarrow g_{U,Q}(\mathbf{m}) \models_Q s,\ \forall \mathbf{m} \in \mathcal{M}_U, \forall \mathbf{s} \in \mathcal{L}_Q. \tag{V.4}$$

Notice that this condition is satisfied in particular for propositional logic, where $f_{Q,U} : \mathcal{L}_Q \hookrightarrow \mathcal{L}_U$ is simply the inclusion. Indeed, trivially, if $\varphi \in \mathcal{L}_Q$,

$$w \models_U \varphi \Longleftrightarrow w|_Q \models_Q \varphi.$$

The definition of the algebra of propositional sentences of Section 4.1.3 rests on existential quantification. In order to replicate this structure, we will require the general language $\mathcal{L}_V$ to be equipped with a generalised existential quantifier, which maintains part of the basic semantic interpretation of $\exists$ in propositional logic.

The quantifier $\exists$ of propositional logic behaves as follows: given subsets $Q \subseteq U \subseteq V$ such that $U \setminus Q = \{p_{i_1}, \dots, p_{i_n}\}$, it turns a formula $\varphi \in \mathcal{L}_U$ into a formula of $\mathcal{L}_Q$, namely $(\exists p_{i_1}, \dots, p_{i_n})\,\varphi$. Morevoer, the following property holds: for all interpretations $w \in \mathcal{I}_U$ and formulae $\varphi \in \mathcal{L}_U$, we have

$$w \models_U \varphi \iff w|_Q \models_Q (\exists p_{i_1}, \dots, p_{i_n})\,\varphi. \tag{V.5}$$

By abstracting from this, we will require a general language $\mathcal{L}_V$ to be equipped with a quantifier §, which, given $Q \subseteq U$, turns a sentence $s \in \mathcal{L}_U$ into $\S_Q\; s \in \mathcal{L}_Q$. In addition, § must satisfy the following properties:

(1) For all $Q \subseteq U \subseteq V$, for all models $\mathbf{m} \in \mathbb{M}_U$ and sentences $s \in \mathcal{L}_U$,
$$\mathbf{m} \models_U s \iff \mathbf{m}_{\downarrow Q} \models_Q \S_Q\; s$$
(2) For all $Q \subseteq U \subseteq V$, for all models $\mathbf{m} \in \mathbb{M}_Q$ and sentences $s \in \mathcal{L}_Q$,
$$\mathbf{m} \models_Q \S_Q\; f_{Q,U}(s) \iff \mathbf{m} \models_Q s$$

Property (1) is simply the generalisation of (V.5). Property (2) translates to a complete tautology in the case of propositional logic,[4] but needs to be stated for general models and languages, where the embeddings f may have a more complex structure.

We are finally ready to introduce an algebra of sentences for the family

$$\{\langle \mathcal{L}_Q, \mathbb{M}_Q, \models_Q \rangle, f, g, \S\} \tag{V.6}$$

by generalising the discussion carried out in Section 4.1.3. The valuations of the algebra are closed sets of sentences $S \subseteq \mathcal{L}_Q$. This means that $S = \mathcal{C}_Q(S)$, where $\mathcal{C}_Q = \mathscr{T}_Q \circ \mathscr{M}_Q$ denotes the closure operator associated to the context $\langle \mathcal{L}_Q, \mathbb{M}_Q, \models_Q \rangle$. Combination is defined as follows: given $S_1 \subseteq \mathcal{L}_Q$ and $S_2 \subseteq \mathcal{L}_U$, we have

$$S_1 \otimes S_2 := \mathcal{C}_{Q \cup U}\left(f_{Q,Q\cup U}(S_1) \cup f_{U,Q\cup U}(S_2)\right).$$

Given a set of sentences $S \subseteq \mathcal{L}_U$, projection to a subdomain $Q \subseteq U$ is given by

$$S^{\downarrow Q} := \mathcal{C}_Q(\S_Q\; S).$$

This algebra is clearly idempotent, and it has both neutral elements $e_Q := \mathcal{C}_Q(\emptyset)$, and null elements $z_Q := \mathcal{L}_Q$. It is thus an information algebra.

4.3. Examples. Apart from propositional logic, there are many frameworks whose information structure can be described in this general sense. Here are some examples.

Predicate logic. Many logical frameworks outside of propositional logic share the same basic structure. We show here how predicate logic can be treated in this abstract sense, and refer the reader to [**WM99**] for further examples in the domain of logic.

The vocabulary of predicate logic comprises a countable set of variables $V := \{x_1, x_2 \dots\}$, a countable set of predicate symbols $P := \{p_1, p_2, \dots\}$, the logical constants $\top$ and $\bot$, the connectors $\wedge, \neg$, and the existential quantifier $\exists$. To each p_i we associate its *rank* $\mathsf{rank}(p_i) \in \mathbb{N}$. Formulae of predicate logic are obtained by applying a finite amount of times the following rules:

[4] In this case, since $f_{Q,U}(\varphi) = \varphi$, and φ only contains variables in Q, there are no variables over which to quantify, thus the statement reduces to $\mathbf{m} \models_Q \varphi \Leftrightarrow \mathbf{m} \models_Q \varphi$.

(1) $\top$, $\bot$ and all strings of the form $p_i x_{i_1} \dots x_{i_\rho}$, where $\rho = \mathsf{rank}(p_i)$, are formulae.
(2) If φ is a formula, then $\neg\varphi$ and $(\exists x_i)\varphi$ are formulae.
(3) If φ and ψ are formulae, then $\varphi \wedge \psi$ is a formula.

The predicate language is the set $\mathcal{L}$ of all formulae. For each subset $Q \subseteq V$, the language $\mathcal{L}_Q \subseteq \mathcal{L}$ is the language where only variables from Q are allowed.

Let $\mathcal{R} = \langle U, R_1, R_2, \dots \rangle$ be a *relational structure*. This means that U is a nonempty set called *universe*, and each $R_i \subseteq U^{\mathsf{rank}(p_i)}$ is a relation of arity $\mathsf{rank}(p_i)$. A *valuation* is a sequence $v = (v_i)_{i=1}^{\infty}$ of elements of the universe (i.e. $v_i \in U$ for all i). We denote by U^ω the set of all possible valuations, end denote

$$v^{\Rightarrow i} := \{u \in U^\omega : u_j = v_j, \forall j \neq i\}.$$

Just like in the case of propositional logic, valuations are used to assign truth values $\hat{v}(\varphi) \in \{0,1\}$ to each formula in $\varphi \in \mathcal{L}$. More specifically, we have

(1) $\hat{v}(\top) = 1$ and $\hat{v}(\bot) = 0$.
(2) $\hat{v}(p_i x_{i_1} \dots x_{i\rho}) = \begin{cases} 1 & \text{if } (v_{i_j})_{j=1}^{\rho} \in R_i, \\ 0 & \text{otherwise.} \end{cases}$
(3) $\hat{v}(\neg\varphi) = \begin{cases} 1 & \text{if } \hat{v}(\varphi) = 0, \\ 0 & \text{if } \hat{v}(\varphi) = 1. \end{cases}$
(4) $\hat{v}((\exists x_i)\varphi) = \begin{cases} 1 & \text{if there exists } u \in v^{\Rightarrow i} \text{ s.t. } \hat{u}(\varphi) = 1, \\ 0 & \text{otherwise.} \end{cases}$
(5) $\hat{v}(\varphi \wedge \psi) = \begin{cases} 1 & \text{if } \hat{v}(\varphi) = 1 \text{ and } \hat{v}(\psi) = 1, \\ 0 & \text{otherwise.} \end{cases}$

We say that a valuation $v \in U^\omega$ is a *model* of a formula $\varphi \in \mathcal{L}$ in the structure $\mathcal{R}$ if $\hat{v}(\varphi) = 1$, and write $v \models \varphi$. This defines a context $\langle \mathcal{L}, U^\omega, \models \rangle$. Moreover, for each subset $S \subseteq V$ we have a context $\left\langle \mathcal{L}_S, U^S, \models_S \right\rangle$, given by projecting valuations onto S. Clearly, U^S constitutes a tuple system with cartesian projection. Furthermore, for all $Q \subseteq U \subseteq V$, we have $\mathcal{L}_Q \subseteq \mathcal{L}_U$. Thus we can define the embedding $f_{Q,U} : \mathcal{L}_Q \hookrightarrow \mathcal{L}_U$ simply as the inclusion map, and it is easy to check that the infomorphism condition (V.4) is verified. Therefore, we can define the valuation algebra of *predicate information sets*.

Since $\mathcal{L}$ is equipped with the quantifier $\exists$, which readily verifies conditions (1) and (2), one can define the algebra of predicate sentences as explained above.

Linear systems. Let K be a field and $V = \{x_1, x_2, \dots, x_n\}$ a set of variables. For each subset $W = \{x_{i_1}, \dots, x_{i_m}\} \subseteq V$, the language $\mathcal{L}_W$ is constituted by all linear equations e over variables contained in W:

$$e := \left(\sum_{j=1}^{m} \lambda_j x_j = b_j \right),$$

where $\lambda_j, b_j \in K$ for all $1 \leq j \leq m$. Models $\mathbf{s} = \langle s_1, \dots, s_m \rangle \in K^m$ are m-tuples valued in K. The relation $\mathbf{s} \models_W e$ means that $\mathbf{s}$ is a solution to the equation e. These notions define a context $\langle \mathcal{L}_W, \mathbb{M}_W, \models_W \rangle$ for each subset $W \subseteq V$. Clearly, $\mathbb{M}_W$ is a tuple system, with projection given by simple cartesian projection. Moreover, since $\mathcal{L}_Q \subseteq \mathcal{L}_U$

for all $Q \subseteq U$, we can define a sequence of embeddings $f_{Q,U} : \mathcal{L}_Q \hookrightarrow \mathcal{L}_U$ that can be easily proved to verify condition (V.4). Therefore, we can define the valuation algebra of information sets as explained above.

Although uncommon, it is possible to equip the language $\mathcal{L}$ with an existential quantifier verifying conditions (1) and (2). This allows the definition of an associated algebra of sentences for linear equations.

Linear inequalities. The same discussion holds for linear inequalities. Given a field K and a finite set of variables $V = \{x_1, \dots, x_n\}$, we define, for each $W = \{x_{i_1}, \dots, x_{i_m}\} \subseteq V$ the language $\mathcal{L}_W$, constituted by linear inequalities over variables contained in W:

$$i := \left(\sum_{j=1}^{m} \lambda_j x_j \leq b_j \right)$$

where $\lambda_j, b_j \in K$ for all $q \leq j \leq m$. The same conclusions about the valuation algebras associated to linear equations hold for linear inequalities.

Constraint satisfaction problems. Constraint satisfaction problems (CSP) are particularly useful to model various kinds of problems in mathematics and computer science thanks to their general and versatile formulation. A CSP is a triple $\langle X, D, C \rangle$, where $X = \{x_1, \dots x_n\}$ is a set of variables, $D = \{D_1, \dots D_n\}$ is the set of the respective domains[5] of values, and $C = \{c_1, \dots c_m\}$ is a set of constraints. Each variable x_i can take values in its domain D_i. A constraint $c_i \in C$ is a pair $\langle T_i, R_i \rangle$, where $T_i := \{x_{i_1}, \dots, x_{i_k}\} \subseteq X$ and R_i is a k-ary relation $R \subseteq \prod_{j=1}^{k} D_{i_j}$. The set T_i is also called the *scheme* of c_i, and it is denoted by $\mathsf{scheme}(c_i)$, whereas $R_i := \mathsf{rel}(c_i)$.
An *evaluation* of the variables is a map from a subset of the variables $S \subseteq X$

$$\mathbf{v} : S \longrightarrow \prod_{x_i \in S} D_i.$$

that assigns to each variable $x_j \in S$ a value $v(x_j)$ in its domain D_j. Such a map can be seen as an element of

$$D_S := \prod_{x_i \in S} D_i.$$

For each $S \subseteq X$, we define $\mathbb{M}_S$ to be the set of all evaluations on S, i.e. $\mathbb{M}_S := D_S$. The associated language $\mathcal{L}_S$ is defined by

$$\mathcal{L}_S := \{c_i \in C : \mathsf{scheme}(c_i) \subseteq S\}.$$

We say that an evaluation $\mathbf{v}$ of the variables in S satisfies a constraint $c = \langle T, R \rangle$, if $\mathbf{v}_{\downarrow S \cap T} \in R$. An evaluation is called *consistent* when it does not violate any constraint. It is called *complete* if it includes all variables in X. It is called a *solution* if it is consistent and complete.

Satisfiability of a constraint defines a relation $\models_S$ for each subset $S \subseteq X$: given $\mathbf{v} \in \mathbb{M}_S$ and $c \in \mathcal{L}_S$,

$$\mathbf{v} \models_S c \iff S \cap \mathsf{scheme}(c) = \emptyset \text{ or } \mathbf{v}_{\downarrow S \cap \mathsf{scheme}(c)} \in \mathsf{rel}(c).$$

[5] This shall not be confused with the term *domain* of the valuation algebra formalism. Rather, the domain of a variable in this setting is analogue to the concept of frame in valuation algebras.

Thus we obtain a sequence of contexts $\langle \mathcal{L}_S, \mathbb{M}_S, \models_S \rangle$. Moreover, because $\mathbb{M}_S = D_S$ is a tuple system, and $\mathcal{L}_Q \subseteq \mathcal{L}_U$ for all $Q \subseteq U$, we can trivially define embeddings $f_{Q,U} : \mathcal{L}_Q \hookrightarrow \mathcal{L}_U$, which satisfy the informorphism condition (V.4). Thus, we can define the algebra of information sets as described above.

5. Valuation algebras and sheaf theory

Remarkably, many of the properties of valuation algebras can be effectively captured by sheaf theory. Just like presheaves deal with the restriction and localisation of topological structures and their extendability through a 'gluing' process, valuation algebras model the focus of knowledge and information, and represent the natural framework to study how local information can be extended through a 'combination' process.

A valuation prealgebra Φ on a set of variables V is nothing but a presheaf over the discrete space V:

$$\Phi : \mathcal{P}(V)^{op} \longrightarrow \mathbf{Set}, \tag{V.7}$$

where $\Phi(S) := \Phi_S$, and

$$\Phi(S \subseteq T) := \rho_S^T : \Phi_T \longrightarrow \Phi_S :: \phi \longmapsto \phi^{\downarrow T}.$$

Indeed, by (A3), we have, for all $S \subseteq V$ and for all $\phi \in \Phi_S$,

$$\rho_S^S(\phi) = \phi^{\downarrow S} = \phi^{\downarrow d(\phi)} \overset{(A3)}{=} \phi,$$

and, by (A2), for all $S \subseteq T \subseteq U \subseteq V$ and $\phi \in \Phi_U$,

$$\rho_S^T \circ \rho_T^U(\phi) = \left(\phi^{\downarrow T}\right)^{\downarrow S} \overset{(A2)}{=} \phi^{\downarrow S} = \rho_S^U(\phi).$$

This sheaf-theoretic perspective allows us to capture the restriction or localisation of the information carried by a valuation algebra. We conclude that a valuation algebra is simply a presheaf (V.7) equipped with a combination operation which satisfies axioms (A4)-(A6). Unfortunately, combination cannot generally be characterised in categorical terms. However, there exist some general constructions which will play a central role in our study, as we shall see in Section 8.2.

5.1. Examples. In many concrete cases, the presheaf Φ can be decomposed into a sequential composition of functors. Before presenting a few examples, let us introduce some definitions. Let V be a set of variables, the **frame functor** $\Omega : \mathcal{P}(V)^{op} \to \mathbf{Set}$ is defined as follows: for all $U \subseteq U' \subseteq V$, we have

$$\Omega(U) := \Omega_U,$$
$$\Omega(U \subseteq U') := \left(\mathsf{proj}_U^{U'} : \Omega_{U'} \longrightarrow \Omega_U :: \mathbf{x} \longmapsto \mathbf{x}_{\downarrow U}\right).$$

Secondly, for all semirings $\langle R, +, \cdot, 0, 1\rangle$, we define the functor $F_R : \mathbf{Set} \to \mathbf{Set}$ as follows: for all sets X, and every function $f : X \to Y$, we have

$$F_R(X) := \{\phi : X \to R\},$$
$$F_R(f) : F_R(X) \longrightarrow F_R(Y) :: \phi \longmapsto \lambda y. \sum_{\substack{x \in X \\ f(x)=y}} \phi(x).$$

With this premise, we can analyse the functorial structure of the examples of Sections 3 and 4:

- *Indicator functions*: The valuation prealgebra of indicator functions on a set of variables V, with frames Ω_x for each $x \in V$, can be described by the presheaf I, which is defined by the following composition of functors:

$$\mathsf{I} : \mathcal{P}(V)^{op} \xrightarrow{\Omega} \mathbf{Set} \xrightarrow{F_{\mathbb{B}}} \mathbf{Set},$$

- *Relational databases*: Given a set of attributes V, we can describe the valuation prealgebra of relational databases as the presheaf $\mathcal{R}$, defined as follows:

$$\mathcal{R} : \mathcal{P}(V)^{op} \xrightarrow{\Omega} \mathbf{Set} \xrightarrow{\mathcal{P}} \mathbf{Set}.$$

  Remark V.13. The functor $\mathcal{P}$ and $F_{\mathbb{B}}$ are naturally isomorphic. Indeed, the natural transformation $\eta : \mathcal{P} \Rightarrow F_{\mathbb{B}}$ defined, for all sets U and all subsets $Q \subseteq U$, by

$$\eta_U(Q) := \lambda u. \begin{cases} 1 & \text{if } u \in Q \\ 0 & \text{if } u \notin Q. \end{cases}$$

  can be easily proved to be an isomorphism. This equivalence translates into a natural isomorphism $\mathsf{I} \cong \mathcal{R}$, which encodes in category theoretic terms the equivalence between the valuation prealgebra of indicator functions and the one of relational databases outlined in Remark V.7.

- *Semiring valuation algebras*: Given a commutative semiring R and a set of variables V, the R-semiring valuation prealgebra can be described by the following composition:

$$\mathsf{SR}_R : \mathcal{P}(V)^{op} \xrightarrow{\Omega} \mathbf{Set} \xrightarrow{F_R} \mathbf{Set}.$$

  In particular, we see that in the case where $R = \mathbb{B}$, this valuation prealgebra coincides with the one of indicator functions.
- *Belief functions*: Given a commutative semiring R and a set of variables V, the valuation prealgebra of set potentials is represented by the following composition:

$$\mathsf{BF}_R : \mathcal{P}(V)^{op} \xrightarrow{\Omega} \mathbf{Set} \xrightarrow{\mathcal{P}} \mathbf{Set} \xrightarrow{F_R} \mathbf{Set}$$

- *Language and models*: Let V be a set of variables, and suppose we have a context system $\langle \mathcal{L}, \mathbb{M}, \models \rangle$. We define the *model functor* $\mathbb{M} : \mathcal{P}(V)^{op} \to \mathbf{Set}$ as follows: for all $Q \subseteq U \subseteq V$,

$$\mathbb{M}(Q) := \mathbb{M}_Q,$$
$$\mathbb{M}(Q \subseteq U) := \left(r^U_Q : \mathbb{M}_U \to \mathbb{M}_Q :: \mathbf{x} \mapsto \mathbf{x}_{\downarrow Q} \right)$$

  where $\mathbf{x}_{\downarrow Q}$ denotes the projection of the tuple system $\mathbb{M}$.[6]
  - *Information sets*: The valuation prealgebra of information sets for the context system $\langle \mathcal{L}, \mathbb{M}, \models \rangle$ is described by the following presheaf:

$$\mathsf{IS} : \mathcal{P}(V)^{op} \xrightarrow{\mathbb{M}} \mathbf{Set} \xrightarrow{\mathcal{P}} \mathbf{Set}.$$

[6]Notice that the functor $\mathbb{M}$ is just a general version of the functor Ω. $\mathbb{M}$ coincides with Ω when the tuple system is constituted by cartesian tuples.

– *Algebra of sentences*: Suppose $\langle \mathcal{L}, \mathbb{M}, \models, f, g, \S \rangle$ is a general family of nested logical contexts as in (V.6). We define the *category of languages* **Lng** as the poset category whose objects are languages $\mathcal{L}_Q$ for $Q \subseteq P$, ordered by inclusion. Then, the prealgebra of sentences associated to $\langle \mathcal{L}, \mathbb{M}, \models \rangle$ is defined as follows:

$$\mathsf{AS} : \mathcal{P}(V)^{op} \xrightarrow{\mathcal{L}} \mathbf{Lng}^{op} \xrightarrow{\mathcal{CP}} \mathbf{Set},$$

where $\mathcal{L}(Q) := \mathcal{L}_Q$ for all $Q \subseteq P$, $\mathcal{L}(Q \subseteq U) := (\mathcal{L}_Q \subseteq \mathcal{L}_U)$ for all $Q \subseteq U \subseteq P$, and $\mathcal{CP}$ denotes the *closed subsets functor*, which is defined as follows:

$$\mathcal{CP}(\mathcal{L}_Q) := \{\mathcal{C}_Q(S) : S \subseteq \mathcal{L}_Q\}$$
$$\mathcal{CP}(\mathcal{L}_Q \subseteq \mathcal{L}_U) : \mathcal{CP}(\mathcal{L}_U) \to \mathcal{CP}(\mathcal{L}_Q) :: \mathcal{C}_U(S) \mapsto \mathcal{C}_Q(\S_Q\, S).$$

where $\mathcal{C}_Q$ denotes the closure operator of $\langle \mathcal{L}_Q, \mathbb{M}_Q, \models_Q \rangle$, and $\S$ is the quantifier of $\mathcal{L}$.

6. A theory of disagreement

Inference problems are the main subject of study in the valuation algebra literature, as they capture the very essence of information as a carrier of knowledge used to answer specific questions. However, such problems do not take into account another fundamental concept related to knowledge: disagreement. A solution to an inference problem is nothing but the most informative answer one can get on a specific question given the available information, regardless of the quality of the sources. This means that if some of the sources provide incorrect or incomplete information, the solution of the problem will be corrupted. In other words, disagreement between the different sources will lead to a less informative answer. In this chapter we will show how to model disagreement in valuation algebra knowledgebases using the sheaf-theoretic description introduced in the previous section. Later on, we will link this discussion with the study of non-locality and contextuality.

6.1. Defining disagreement. Consider a valuation algebra Φ, on a set of variables V, and let $K := \{\phi_1, \dots, \phi_n\}$ be a knowledgebase. We let $C_i := d(\phi_i)$, and $\mathcal{M} := \{C_i\}_{i=1}^n$.

A natural way to say that $\phi_1, \dots, \phi_n$ **agree** is to say that there exists a valuation $\gamma \in \Phi_X$ such that

$$\gamma^{\downarrow C_i} = \phi_i, \ \forall 1 \leq i \leq n. \tag{V.8}$$

Concretely, this means that the information carried by each individual valuation ϕ_i comes as a restriction of a 'truth' valuation which is implicitly agreed upon by all the sources. Given this premise, we say that $\phi_1, \dots, \phi_n$ **disagree** if such a global valuation γ does not exist.

In sheaf theoretic terms, for any subset $U \subseteq X$, each valuation $\phi \in \Phi_U$ is a local section at U, while a valuation $\gamma \in \Phi_X$ is a global section. Therefore, saying that $\phi_1, \dots, \phi_n$ agree is equivalent to say that there exists a global section for the family $\{\phi_1, \dots, \phi_n\}$.

6.1.1. Local disagreement. Notice that a necessary condition for (V.8) to hold is that

$$\phi_i^{\downarrow C_i \cap C_j} = \phi_j^{\downarrow C_i \cap C_j}, \ \forall 1 \leq i, j \leq n, \tag{V.9}$$

indeed,

$$\phi_i^{\downarrow C_i \cap C_j} \overset{(\text{V.8})}{=} \left(\gamma^{\downarrow C_i}\right)^{\downarrow C_i \cap C_j} \overset{(\text{A4})}{=} \gamma^{\downarrow C_i \cap C_j} \overset{(\text{A4})}{=} \left(\gamma^{\downarrow C_j}\right)^{\downarrow C_i \cap C_j} \overset{(\text{V.8})}{=} \phi_j^{\downarrow C_i \cap C_j}$$

This is to say that, if there is agreement between the sources, then, in particular, each pair of sources agree on their common variables. Condition (V.9) defines the notion of *local agreement*. Local agreement of valuations $\phi_1, \ldots \phi_n$ corresponds to the sheaf theoretic notion of *compatibility*. Therefore, it makes sense to introduce the following definition:

Definition V.14. Let Φ be a valuation algebra on a set of variables V. A **locally agreeing knowledgebase** is a compatible family $\{\phi_{C_i}\}_{i=1}^n$ for the presheaf $\Phi : \mathcal{P}(V)^{op} \to \mathbf{Set}$. We say that the ϕ_C's **agree globally** if there is a global section $\gamma \in \Phi(X)$ such that $\gamma|_C = \phi_C$ for all $C \in \mathcal{M}$. Such a global section is called a **truth valuation**.

6.2. Examples. Local agreement between sources is clearly easier to verify in real world scenarios than global disagreement. If two sources are directly in contradiction with each other, such an inconsistency should be immediately apparent. In [**ZG18**], a sheaf-theoretic framework for the study of such direct contradictions is introduced. For instance, the following example – taken from real sources – is given:

Example V.15. Let us consider breast cancer screening guidlines from three different accredited sources:

(1) *Screening with mammography and clinical breast exam annually*
(2) *Biennal screening mammography is recommended*
(3) *Women aged 50 to 54 years should get mammograms every year. Women aged 55 years and older should switch to mammograms every 2 years* [...]

We can represent the information provided by these medical bodies as a knowledgebase of the information algebra of relational databases. We have the variables $\{a, e, f\}$ representing age intervals, exam type and exam frequency respectively. The frames for each variable are $\Omega_a = \{54^-, 54^+\}$, $\Omega_e = \{\text{MG}, \text{CBE}\}$, $\Omega_f = \{\text{Y}, 2\text{Y}\}$. Each source $i = 1, 2, 3$ from the example can be described by 3 relations R_i such that $d(R_{1,2}) = \{e, f\}$ and $d(R_3) = \{a, e, f\}$, defined by

$$R_1 = \{\langle \text{M}, \text{Y} \rangle, \langle \text{CBE}, \text{Y} \rangle\},$$
$$R_2 = \{\langle \text{M}, 2\text{Y} \rangle\},$$
$$R_3 = \{\left\langle 54^-, \text{M}, \text{Y} \right\rangle, \left\langle 54^+, \text{M}, \text{Y} \right\rangle\}.$$

It is easy to see that, for instance, the three sources do not agree locally on which exam to undergo. Indeed,

$$R_1^{\downarrow \{e\}} = \{M, CBE\} \neq \{M\} = R_{2,3}^{\downarrow \{e\}}$$

However, it is not always the case that locally-agreeing information leads to global agreement as defined in (V.8). Such inconsistencies are much more subtle and difficult to detect, and they will constitute the main subject of our discussion. An example of locally agreeing sources which do not agree globally can be obtained by slightly tweaking the previous example.

Example V.16. Suppose the guidelines are now as follows:

(1) *Screening with mammography annually, clinical breast exam annually or biannually*
(2) *Women aged 50 to 54 years should get mammograms. Women aged 55 years and older should switch to clinical breast exams*
(3) *Women aged 50 to 54 years should undergo an exam every year. Women aged 55 years and older should be examined every 2 years*

In this case we have

$$R_1 = \{\langle \mathrm{M}, \mathrm{Y}\rangle, \langle \mathrm{CBE}, \mathrm{Y}\rangle, \langle \mathrm{CBE}, 2\mathrm{Y}\rangle\},$$
$$R_2 = \{\langle 54^-, \mathrm{M}\rangle, \langle 54^+, \mathrm{CBE}\rangle\},$$
$$R_3 = \{\langle 54^-, \mathrm{Y}\rangle, \langle 54^+, 2\mathrm{Y}\rangle\},$$

with $d(R_1) = \{e, f\}$, $d(R_2) = \{a, e\}$ and $d(R_3) = \{a, f\}$. It is easy to see that all these sources agree locally, i.e. they do not directly contradict each other. However, we can show that they globally disagree. Indeed, the only global section of the presheaf induced by $K = \{R_1, R_2, R_3\}$ is

$$G := \{\langle \mathrm{M}, \mathrm{Y}, 54^-\rangle, \langle \mathrm{CBE}, 2\mathrm{Y}, 54^+\rangle\},$$

and we have

$$G^{\downarrow d(R_1)} = \{\langle \mathrm{M}, \mathrm{Y}\rangle, \langle \mathrm{CBE}, 2\mathrm{Y}\rangle\} \neq R_1.$$

We can easily represent visualise this knowledgebase as a bundle diagram, pictured in Figure V.2. In this case, the facets of the base complex are the domains of $R_{1,2,3}$, and the fibers represent the frames of each variable.

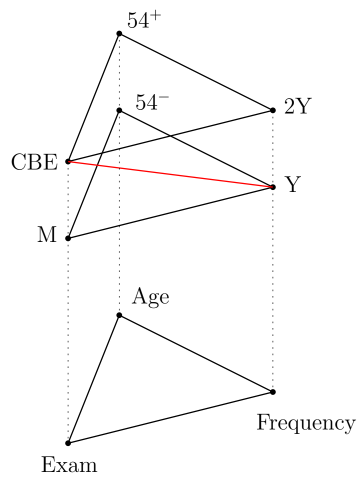


Figure V.2. The knowledgebase $\{R_1, R_2, R_3\}$ represented as a bundle diagram. The red edge, which represents tuple $\langle \mathrm{CBE}, \mathrm{Y}\rangle$, is not part of any global valuation

The parallelism with contextuality is evident. This is another instance of a database which does not admit a universal relation, a property that has been proved to be equivalent to contextuality in [**Abr13a**]. We have already encountered a similar instance of global disagreement in Example V.6.

The main focus of this chapter will be to study locally agreeing sources that disagree globally and show that such a discrepancy is mathematically equivalent to the notion of contextuality. Then, we will use this connection to translate theorems, methods and algorithms from one framework to the other.

Let us show other examples of locally agreeing valuations that are globally disagreeing, to prove how widespread this phenomenon is in various domains of mathematics.

Example V.17. Consider the following system of linear equations in $\mathbb{Z}_2$:

$$\begin{aligned} e_1 &:= (x_1 \oplus x_2 \oplus x_3 = 1) \\ e_2 &:= (x_1 \oplus y_2 \oplus y_3 = 0) \\ e_3 &:= (y_1 \oplus x_2 \oplus y_3 = 0) \\ e_4 &:= (y_1 \oplus y_2 \oplus x_3 = 0) \end{aligned} \tag{V.10}$$

Let $V := \{x_{1,2,3}, y_{1,2,3}\}$ denote the set of variables involved and consider the valuation algebra of information sets arising from linear equations discussed in Section 4. For all $1 \leq i \leq 4$, let ϕ_i denote the set of solutions to the equation e_i. In other words,

$$\phi_i := \mathcal{M}_{d(e_i)}(e_i), \ \forall 1 \leq i \leq 4.$$

Then, we can see that $|d(\phi)_i \cap d(\phi_j)| = 1$ for all $1 \leq i, j \leq 4$ (i.e. each equation has exactly one variable appearing in common with any other equation), and

$$\phi_i^{\downarrow d(\phi)_i \cap d(\phi_j)} = \{0, 1\},$$

where we have identified 0 and 1 with the tuples $d(\phi)_i \cap d(\phi_j) \mapsto 0$ and $d(\phi)_i \cap d(\phi_j) \mapsto 1$ respectively. Therefore, the knowledgebase $K = \{\phi_1, \ldots, \phi_4\}$ agrees locally. However, it is easy to see that system (V.10) is inconsistent. Indeed, since each variable appears twice on the right hand side, the sum of all the equations yields $0 = 1$. Therefore, there is no global section for the knowledgebase K. We can see that, in the case of linear euqations, global agreement corresponds to the existence of a solution.

Example V.18. Consider the following problem. We want to colour a political map of the geographical region surrounding Malawi using 3 colours – say red, green and yellow – with the condition that adjacent countries should be coloured differently. A blank map is pictured in Figure V.3.

We can model this problem as a CSP. The set of variables X is constituted by a variable for each of the 5 countries in the map, i.e.

$$X = \{\mathsf{MOZ}, \mathsf{MWI}, \mathsf{TZA}, \mathsf{ZMB}, \mathsf{ZWE}\}$$

The domain for each variable is the set of colours we can attribute to the country, i.e. $D = \{g, r, y\}$, and it is the same for each variable. Let

$$S := \{\langle g, r \rangle, \langle g, y \rangle, \langle r, y \rangle\}.$$

FIGURE V.3. A blank map of the geographical region surrounding Malawi.

There are 8 constraints c_i, $1 \leq i \leq 8$ defined by $\mathsf{rel}(c_i) = S$ for all $1 \leq i \leq 8$ and

$$T_1 = \{\mathsf{MOZ}, \mathsf{MWI}\}, T_2 = \{\mathsf{MOZ}, \mathsf{TZA}\}, T_3 = \{\mathsf{MOZ}, \mathsf{ZMB}\}, T_4 = \{\mathsf{MOZ}, \mathsf{ZWE}\},$$
$$T_5 = \{\mathsf{MWI}, \mathsf{TZA}\}, T_6 = \{\mathsf{MWI}, \mathsf{ZWE}\}, T_7 = \{\mathsf{TZA}, \mathsf{ZMB}\}, T_8 = \{\mathsf{ZMB}, \mathsf{ZWE}\},$$

where $T_i = \mathsf{scheme}(c_i)$ for all $1 \leq i \leq 8$. Consider the valuation algebra of information sets for CSPs defined in Section 4. The knowledgebase for this problem is constituted by 8 valuations $\{\phi_i\}_{i=1}^{8}$ defined by

$$\phi_i := \mathcal{M}_{T_i}(C) = \{\mathbf{v} \in D_{T_i} : \mathbf{v} \models_{T_i} c_j,\ \forall 1 \leq j \leq 8\}.$$

For instance

$$\phi_1 = \{\langle \mathsf{MOZ}, \mathsf{MWI} \rangle \mapsto \langle g, r \rangle\ ,\ \text{or}\ \ \langle g, y \rangle\ ,\ \text{or}\ \ \langle r, y \rangle\}.$$

It is easy to see that all the valuations agree locally, indeed, for all $1 \leq i, j \leq 8$, we either have $T_i \cap T_j = \emptyset$, in which case local agreement is trivially satisfied, or $T_i \cap T_j$ has exactly one element t, and we have

$$\phi_i^{\downarrow T_i \cap T_j} = \{t \mapsto g \text{ or } r \text{ or } y\} = \phi_j^{\downarrow T_i \cap T_j}.$$

However, the ϕ_i's disagree globally. Indeed, if there was a valuation γ which projects onto each ϕ_i, it would imply that there is a solution to the CSP, and thus a colouring of the map using only three colours. It is easy to see that this is not the case by simply looking at the *constraint graph* (cf. Figure V.4) of the problem, and conclude that its chromatic number is 4.

EXAMPLE V.19. To cite an example from logic, consider the the famous liar's paradox, which consists of the sentence

$$S : S \text{ is false.}$$

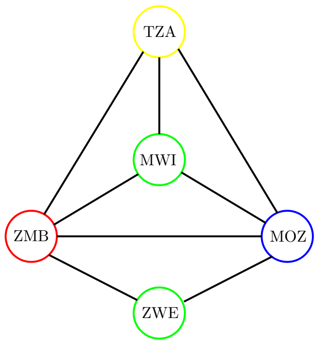


FIGURE V.4. The constraint graph of the CSP. A 4-colouring of the graph is shown. The chromatic number of the graph is 4.

We can generalise this to a *liar cycle* of length n, i.e. a sequence of statements:

$$\begin{aligned} S_1 &: S_2 \text{ is true,} \\ S_2 &: S_3 \text{ is true,} \\ &\vdots \\ S_{n-1} &: S_n \text{ is true,} \\ S_n &: S_1 \text{ is false.} \end{aligned}$$

These statements can be modelled as a series of formulae in propositional logic. Let $V := \{s_1, \ldots, s_n\}$ be a set of variables, each representing one of the statements S_i above. The n liar cycle can be rewritten as follows:

$$\begin{aligned} s_1 &\leftrightarrow s_2 \\ s_2 &\leftrightarrow s_3 \\ &\vdots \\ s_{n-1} &\leftrightarrow s_n \\ s_n &\leftrightarrow \neg s_1 \end{aligned} \tag{V.11}$$

We define the following valuations:

$$\phi_n := \mathcal{M}_{\{s_1,s_n\}}(s_n \leftrightarrow \neg s_1) = \{\mathbf{v} :: \langle s_1, s_n\rangle \mapsto \langle 1,0\rangle \text{ or } \langle 0,1\rangle\},$$

and for all $1 \leq i \leq n-1$,

$$\phi_i := \mathcal{M}_{\{s_i,s_{i+1}\}}(s_i \leftrightarrow s_{i+1}) = \{\mathbf{v} :: \langle s_i, s_{i+1}\rangle \mapsto \langle 0,0\rangle \text{ or } \langle 1,1\rangle\}.$$

It is easy to see that the ϕ_i's agree locally. Indeed, both 0 and 1 are valid assignments for a single variable, which is the most two valuations can have in common. On the other hand, the fact that the liar cycle gives rise to a paradox means that the valuations do not agree globally. Indeed, a global assignment of truth values to each $s_1, \ldots, s_n$ consistent with equations (V.11) is impossible, as they collectively yield $s_1 \leftrightarrow \neg s_1$. A more formal proof of this will be given in Section 9, when we will have the tools to characterise global truth functions for the algebra in question.

The same situation can be described using the information algebra of propositional formulae. For all $1 \leq i \leq n-1$, let

$$\phi_i := \mathcal{C}_{\{s_i,s_{i+1}\}}(\{s_i \leftrightarrow s_{i+1}\}),$$

and

$$\phi_n := \mathcal{C}_{\{s_1,s_n\}}(\{s_n \leftrightarrow s_1\}),$$

where, we recall, $\mathcal{C}_Q = \mathscr{T}_Q \circ \mathscr{M}_Q$ is the closure operator. Then, by definition of $\mathcal{C}$, each ϕ_i contains all the formulae satisfied by the valuations $\mathbf{v} :: \langle s_i, s_{i+1}\rangle \mapsto \langle 0,0\rangle$ or $\langle 1,1\rangle$, i.e.

$$\phi_i = \{s_i \leftrightarrow s_{i+1},\ s_i \rightarrow s_{i+1},\ s_{i+1} \leftrightarrow s_i,\ (s_i \rightarrow s_{i+1}) \wedge (s_{i+1} \leftrightarrow s_i),\ \dots\},$$

and similarly for ϕ_n. Now, the knowledgebase $\{\phi_1, \dots, \phi_n\}$ agrees locally, indeed, for all $1 \leq i \leq n-1$,

$$\begin{aligned}
\phi_i^{\downarrow\{s_i,s_{i+1}\}\cap\{s_{i+1},s_{i+2}\}} &= \phi_i^{\downarrow\{s_{i+1}\}} = \mathcal{C}_{\{s_{i+1}\}}\left(\exists s_i\ \{s_i \leftrightarrow s_{i+1},\ s_i \rightarrow s_{i+1},\ s_{i+1} \leftrightarrow s_i,\ \dots\}\right) \\
&= \mathscr{T}_{\{s_{i+1}\}}(\{\mathbf{v} : s_{i+1} \mapsto 0 \text{ or } 1\}) \\
&= \{s_{i+1}, \neg s_{i+1}\} \\
&= \mathcal{C}_{\{s_{i+1}\}}(\exists\ s_{i+2}\{s_i \leftrightarrow s_{i+1},\ s_i \rightarrow s_{i+1},\ s_{i+1} \leftrightarrow s_i,\ \dots\}) \\
&= \phi_{i+1}^{\downarrow\{s_i,s_{i+1}\}\cap\{s_{i+1},s_{i+2}\}}.
\end{aligned}$$

One can prove in the same way that $\phi_n^{\downarrow\{s_1,s_n\}\cap\{s_1,s_2\}} = \phi_1^{\downarrow\{s_1,s_n\}\cap\{s_1,s_2\}}$. However, once again, the paradoxical nature of the liar cycle corresponds to the fact that one cannot find a global truth valuation for this knowledgebase. This will be proved formally in Section 8.4.

7. Disagreement and contextuality

After having introduced a general concept of disagreement and presented many examples of locally agreeing knowledgebases that disagree globally ranging from relational databases to propositional logic, it is now time to reveal the link between contextuality and disagreement underpinning the theory presented thus far. We have purposely kept the connection implicit until now, so as to show how the theory of disagreement can be developed completely independently of the concept of contextuality.

No-signalling and local agreement. Let $e = \{e_C \in \mathcal{D}_R\mathcal{E}(C)\}_{C\in\mathcal{M}}$ be an empirical model over a measurement scenario $\langle X, \mathcal{M}, (O_m)\rangle$. Let us take X as a set of variables on which to build a suitable valuation algebra. As a first note, it is straightforward to show that the sheaf of event $\mathcal{E}$ coincides with the frame functor $\Omega : \mathcal{P}(X)^{op} \rightarrow \mathbf{Set}$, where, for each measurement $m \in X$, its frame is given by $\Omega_m := O_m$.[7]

We can interpret each e_C as a valuation of the algebra Φ of R-potentials. From this viewpoint, the empirical model e is nothing but a knowledgebase of Φ. Then, the property of no-signalling corresponds precisely to local agreement. Thus, to summarise, one can think of an empirical model as a locally agreeing knowledgebase of Φ.

[7] From now on, we will use $\mathcal{E}$ and Ω interchangeably. In particular, a local section in $\mathcal{E}(U) = \Omega_U$ can be denoted either as a function $s \in \mathcal{E}(U)$, or as a tuple $\mathbf{x} \in \Omega_U$, in which case the restriction is denoted $\rho_U^{U'}(\mathbf{x}) = \mathbf{x}_{\downarrow U}$.

Contextuality and global disagreement. By considering no-signalling empirical models as locally agreeing knowledgebases of the algebra of R-potentials, a striking connection with the theory of disagreement arises: non-locality and contextuality are just a special instance of a locally agreeing knowledgebase which disagrees globally:

THEOREM V.20. *Let $e = \{e_C\}_{C\in\mathcal{M}}$ be an empirical model. Then e is contextual if and only if the locally-agreeing knowledgebase $K = \{e_C\}_{C\in\mathcal{M}} \subseteq \Phi$ disagrees globally.*

PROOF. If e is non-contextual, then there exists a global distribution $d : \mathcal{E}(X) \to R$ such that $d|_C = e_C$ for all $C \in \mathcal{M}$. Hence d is a global truth valuation for the knowledgebase K. Conversely, suppose K agrees globally, which means that there exists a global R-potential $d : \Omega_X = \mathcal{E}(X) \to R$ such that $d^{\downarrow C} = e_C$ for all $C \in \mathcal{M}$. The only thing we need to prove is that d is an R-distribution, i.e. that it is normalised. Let $*$ denote the unique $\emptyset$-tuple and let $C \in \mathcal{M}$ be an arbitrary context. We have:

$$\sum_{g\in\mathcal{E}(X)} d(g) = \sum_{\substack{g\in\mathcal{E}(X):\\ g_{\downarrow\emptyset}=*}} d(g) = d^{\downarrow\emptyset}(*) = \left(d^{\downarrow C}\right)^{\downarrow\emptyset}(*) = (e_C)^{\downarrow\emptyset}(*)$$
$$= \sum_{\substack{s\in\mathcal{E}(C):\\ s_{\downarrow\emptyset}=*}} e_C(s) = \sum_{s\in\mathcal{E}(C)} e_C(s) = 1.$$

□

Therefore, from a structural perspective, these counterintuitive phenomena of quantum physics are equivalent to all of the examples introduced in Section 6, and any instance of local agreement vs global disagreement arising from the valuation algebra framework.

This result is a major generalisation of the connections observed in [**Abr13a, Bar15a**] and [**AGK13, ABdSZ17**], which are limited to relational databases and CSPs respectively. It further proves that contextuality is not a phenomenon limited to quantum physics, but it is a general concept which pervades various domains, most of which are completely unrelated to quantum mechanics.

This connection can be further explored and expanded to take into account different kinds of contextuality, such as possibilistic and strong.[8] In order to do this, we need to further refine the notion of valuation algebra, and identify particular structures that will turn the problem of detecting disagreement in an inference problem, as shown for instance by the basic examples V.5 and V.6 in Section 3. This will be the main focus of the next section.

8. Disagreement, complete disagreement and inference problems

Studying global disagreement amounts to looking for a global truth, which is shared by all the sources of information. It is thus natural to ask whether it is possible to recover it from the collective information of the sources and the structure of the valuation

[8]Possibilistic contextuality is actually already captured by the connection we have just presented as, we recall, possibilistic models are compatible families $\{e_C \in \mathcal{D}_{\mathbb{B}}\mathcal{E}(C)\}_{C\in\mathcal{M}}$. However, we would like to rephrase this result using the terminology of Definition II.12, where possibilistic empirical models are defined as presheaves $\mathcal{S} : \mathcal{P}(X)^{op} \to \mathbf{Set}$. We will do this formally in Section 9.

algebra. It turns out that, in a variety of situations, the global truth valuation can appear only in one form, which makes the problem of finding it significantly easier and, crucially, equivalent to an inference problem. In order to prove this, we will need to introduce the concept of an *ordered valuation algebra*.

8.1. Ordered valuation algebras. Given a valuation algebra Φ on a set of variables V, and two valuations $\phi, \psi \in \Phi_S$ for some $S \subseteq V$, one could raise the following question: *how does the information carried by ϕ compare to the one carried by ψ?* In other words, *is there a way of quantifying the amount of information represented by a valuation?* The answer to this question is given by extending the present framework to the one of *ordered valuation algebras* [**Hae04**]. An ordered valuation algebra is a valuation algebra equipped with a completeness relation $\preceq$, which aims to capture how informative a valuation is with respect to others.

Definition V.21. Let Φ be a valuation algebra with null elements on a set of variables V. Then, Φ is an **ordered valuation algebra** if there exists a partial order $\preceq$ on Φ such that the following additional axioms are verified:

(A10) *Partial order*: For all $\phi, \psi \in \Phi$, $\phi \preceq \psi$ implies $d(\phi) = d(\psi)$. Moreover, for every $S \subseteq V$ and $\Psi \subseteq \Phi_S$, the infimum $\inf(\Psi)$ exists.

(A11) *Null element*: For all $S \subseteq V$, we have
$$\inf(\Phi_S) = z_S.$$

(A12) *Monotonicity of combination*: For all $\phi_1, \phi_2, \psi_1, \psi_2 \in \Phi$ such that $\phi_1 \preceq \phi_2$ and $\psi_1 \preceq \psi_2$ we have
$$\phi_1 \otimes \psi_1 \preceq \phi_2 \otimes \psi_2.$$

(A13) *Monotonicity of projection*: For all $\phi, \psi \in \Phi$, if $\phi \preceq \psi$ then
$$\phi^{\downarrow S} \preceq \psi^{\downarrow S}$$
for all $S \subseteq d(\phi) = d(\psi)$.

It can be shown that all the instances of valuation algebras presented in the previous sections can be ordered. We list the definitions of the partial orders in each case (some of the proofs can be found in [**Koh12**]).

- *Indicator functions*: Given $i_1, i_2 \in \Phi_S$, we have
$$i_1 \preceq i_2 \iff \mathsf{supp}(i_1) \subseteq \mathsf{supp}(i_2).$$
- *Semiring valuation algebras* and *set potentials*: Given two arithmetic potentials $\phi, \psi : \Omega_S \to R$, we define
$$\phi \preceq \psi \iff \phi(\mathbf{x}) \leq \psi(\mathbf{x}),\ \forall \mathbf{x} \in \Omega_S.$$
Similarly, given two set potentials $m_1, m_2 \in \Phi_S$,
$$m_1 \preceq m_2 \iff m_1(U) \leq m_2(U),\ \forall U \subseteq \Omega_S.$$
- *General information sets: relational databases, propositional logic, etc.*: In all these cases, the order is simply given by inclusion. That is, for information sets $M_1, M_2 \in \Phi_Q$, we have
$$M_1 \preceq M_2 \Leftrightarrow M_1 \subseteq M_2.$$

In particular, for relations $R_1, R_2 \in \Phi_S$,

$$R_1 \preceq R_2 \Leftrightarrow R_1 \subseteq R_2.$$

- *Algebra of sentences*: The order structure in all algebras of sentences given by any context is also induced by inclusion. However, the order is reversed, i.e. , given two sets of sentences $\Gamma_1, \Gamma_2 \in \Phi_Q$, we have

$$\Gamma_1 \preceq \Gamma_2 \Leftrightarrow \Gamma_2 \supseteq \Gamma_1.$$

Indeed, we have already shown that the least informative element $\inf(\Phi_Q) = z_Q$ is the whole language $\mathcal{L}_Q$.

Once again, we can incorporate some of the axioms in the structure of the presheaf (V.7) by simply rewriting it as

$$\Phi : \mathcal{P}(V)^{op} \longrightarrow \mathbf{Pos},$$

where **Pos** denotes the category of posets and monotone maps. To be more precise, one could write $\Phi : \mathcal{P}(V)^{op} \to \mathcal{C}$, where $\mathcal{C}$ is the full subcategory of **Pos** whose objects are monoids with a complete meet-semilattice structure which is compatible with multiplication (this would capture all the axioms except for (A11)). We will not pursue this idea for the sake of simplicity.

8.2. A general construction for composition. An interesting aspect brought to light by the order structure of valuation algebras is that the composition laws of many algebras are uniquely characterised by the same categorical construction.

Let Φ be an ordered valuation prealgebra, viewed as a presheaf. Thanks to the universal property of products of the category **Set**, we have, for all $S, T \subseteq V$, the following diagram:

$$\begin{tikzcd}
\Phi(S) & \Phi(S) \times \Phi(T) \arrow[l, "\pi_1"'] \arrow[r, "\pi_2"] & \Phi(T) \\
& \Phi(S \cup T) \arrow[ul, "\rho_S^{S\cup T}"] \arrow[u, dashed, "\langle \rho_S^{S\cup T}, \rho_T^{S\cup T} \rangle"] \arrow[ur, "\rho_T^{S\cup T}"'] &
\end{tikzcd}$$

This leads to the following definition:

Definition V.22. An **adjoint valuation algebra** is an ordered valuation algebra Φ such that its combination operation $\otimes$ is the right adjoint of the map $\langle \rho_S^{S\cup T}, \rho_T^{S\cup T} \rangle$, defined in the diagram above. In this case, $\otimes$ is the unique map satisfying the following conditions:

$$\mathsf{id}_{\Phi(S\cup T)} \leq \otimes \circ \left\langle \rho_S^{S\cup T}, \rho_T^{S\cup T} \right\rangle, \tag{V.12}$$

$$\left\langle \rho_S^{S\cup T}, \rho_T^{S\cup T} \right\rangle \circ \otimes \leq \mathsf{id}_{\Phi(S)\times\Phi(T)}, \tag{V.13}$$

where $\leq$ is the pointwise order inherited by the partial order $\preceq$ of the algebra.

Let Φ be an adjoint valuation algebra. Condition (V.12) means the following: for all $S, T \subseteq V$, and any $\phi \in \Phi_{S\cup T}$,

$$\phi \preceq \phi^{\downarrow S} \otimes \phi^{\downarrow T}. \tag{V.14}$$

On the other hand, condition (V.13) means that for all $\phi \in \Phi_S$ and $\psi \in \Phi_T$, the following two equations hold:

$$\begin{aligned} (\phi \otimes \psi)^{\downarrow S} &\preceq \phi, \\ (\phi \otimes \psi)^{\downarrow T} &\preceq \psi. \end{aligned} \tag{V.15}$$

A large portion of the valuation algebras encountered in the first sections of the chapter are in fact adjoint valuation algebras, as we shall now prove:

Proposition V.23. *The information algebra of general information sets and its associated algebra of sentences (cf. Sections 4.2.1 and 4.2.2) are adjoint information algebras.*

Before proceding with the proof, let us recall that this result is extremely widely applicable: it includes relational databases, propositional information, propositional sentences, linear equations, linear inequalities, constraint satisfaction problems and many more.

Proof. Suppose $\langle \mathcal{L}, \mathbb{M}, \models, f, g, \S \rangle$ is a general family of nested logical contexts as in (V.6). We start by proving the proposition for the algebra of general information sets:

- Let $M \subseteq \mathbb{M}_{Q\cup U}$, where $Q, U \subseteq V$.

$$\begin{aligned} M^{\downarrow Q} \otimes M^{\downarrow U} &= \{\mathbf{x}_{\downarrow Q} : \mathbf{x} \in M\} \otimes \{\mathbf{x}_{\downarrow U} : \mathbf{x} \in M\} \\ &= \left\{\mathbf{v} \in \mathbb{M}_{Q\cup U} : \text{ there exist } \mathbf{x}, \mathbf{y} \in M \text{ s.t. } (\mathbf{v}_{\downarrow Q} = \mathbf{x}_{\downarrow Q}) \wedge (\mathbf{v}_{\downarrow U} = \mathbf{y}_{\downarrow U})\right\}. \end{aligned}$$

  Then, clearly, $M \subseteq M^{\downarrow Q} \otimes M^{\downarrow U}$.
- Now, let $M_1 \subseteq \mathbb{M}_Q$ and $M_2 \subseteq \mathbb{M}_U$. We have

$$\begin{aligned} (M_1 \otimes M_2)^{\downarrow Q} &= \{\mathbf{v} \in \mathbb{M}_{Q\cup U} : (\mathbf{v}_{\downarrow Q} \in M_1) \wedge (\mathbf{v}_{\downarrow U} \in M_2)\}^{\downarrow Q} \\ &= \left\{\mathbf{v}_{\downarrow Q} : \mathbf{v} \in \mathbb{M}_{Q\cup U} \wedge (\mathbf{v}_{\downarrow Q} \in M_1) \wedge (\mathbf{v}_{\downarrow U} \in M_2)\right\} \subseteq M_1. \end{aligned}$$

  One proves that $(M_1 \otimes M_2)^{\downarrow U} \subseteq M_2$ in the same way.

This concludes the proof for the algebra of information sets. Let us now prove the result for the associated algebra of sentences:

- Let $S \subseteq \mathcal{L}_{Q\cup U}$ be a closed set of sentences, i.e. $\mathcal{C}_{Q\cup U}(S) = S$. We want to show that $S \preceq S^{\downarrow Q} \otimes S^{\downarrow U}$. This is equivalent to

$$S \supseteq S^{\downarrow Q} \otimes S^{\downarrow U}.$$

  By definition,

$$S = \mathcal{C}_{Q\cup U}(S) = \{\varphi \in \mathcal{L}_{Q\cup U} : \mathbf{m} \models_{Q\cup U} \varphi,\ \forall \mathbf{m} \text{ s.t. } \mathbf{m} \models_{Q\cup U} S\}.$$

Moreover,

$$
\begin{aligned}
S^{\downarrow Q} \otimes S^{\downarrow U} &= \mathcal{C}_Q(\S_Q\ S) \otimes \mathcal{C}_U(\S_U\ S) \\
&= \mathcal{C}_{Q\cup U}\left(f_{Q,Q\cup U}(\mathcal{C}_Q(\S_Q\ S)) \cup f_{U,Q\cup U}(\mathcal{C}_U(\S_U\ S))\right) \\
&= \{\varphi \in \mathcal{L}_{Q\cup U} : \mathbf{m} \models_{Q\cup U} \varphi,\ \forall \mathbf{m} \in \mathbb{M}_{Q\cup U} \\
&\quad \text{s.t. } \mathbf{m} \models_{Q\cup U} f_{Q,Q\cup U}(\mathcal{C}_Q(\S_Q\ S)) \cup f_{U,Q\cup U}(\mathcal{C}_U(\S_U\ S))\}
\end{aligned} \tag{V.16}
$$

Now, let $\varphi \in S^{\downarrow Q} \otimes S^{\downarrow U}$. Let $\mathbf{m} \in \mathbb{M}_{Q\cup U}$ be such that $\mathbf{m} \models_{Q\cup U} S$. We want to show that $\mathbf{m} \models_{Q\cup U} \varphi$. By (V.16), it is sufficient to prove that

$$
\mathbf{m} \models_{Q\cup U} f_{Q,Q\cup U}(\mathcal{C}_Q(\S_Q\ S)) \cup f_{U,Q\cup U}(\mathcal{C}_U(\S_U\ S)). \tag{V.17}
$$

We have

$$
\begin{aligned}
\text{(V.17)} \iff &\begin{cases} \mathbf{m} \models_{Q\cup U} f_{Q,Q\cup U}(\mathcal{C}_Q(\S_Q\ S)), \\ \mathbf{m} \models_{Q\cup U} f_{U,Q\cup U}(\mathcal{C}_U(\S_U\ S)). \end{cases} \\
\overset{\text{(V.4)}}{\iff} &\begin{cases} \mathbf{m}_{\downarrow Q} \models_Q \mathcal{C}_Q(\S_Q\ S), \\ \mathbf{m}_{\downarrow U} \models_U \mathcal{C}_U(\S_U\ S). \end{cases} \\
\iff &\begin{cases} \mathbf{m}_{\downarrow Q} \models_Q \S_Q\ S, \\ \mathbf{m}_{\downarrow U} \models_U \S_U\ S, \end{cases}
\end{aligned}
$$

and this is true by property (1) of the definition of §, since $\mathbf{m} \models_{Q\cup U} S$ by hypothesis.

- Let $S_1 \subseteq \mathcal{L}_Q$ and $S_2 \subseteq \mathcal{L}_U$ be closed sets. We want to show that $(S_1 \otimes S_2)^{\downarrow Q} \preceq S_1$ and $(S_1 \otimes S_2)^{\downarrow U} \preceq S_2$, which are equivalent to

$$
\begin{aligned}
(S_1 \otimes S_2)^{\downarrow Q} &\supseteq S_1, \\
(S_1 \otimes S_2)^{\downarrow U} &\supseteq S_2.
\end{aligned}
$$

By definition,

$$
\begin{aligned}
(S_1 \otimes S_2)^{\downarrow Q} &= \mathcal{C}_Q\left(\S_Q\ \ (\mathcal{C}_{Q\cup U}(f_{Q,Q\cup U}(S_1) \cup f_{U,Q\cup U}(S_2))\right) \\
&= \{\varphi \in \mathcal{L}_Q : \mathbf{m} \models_Q \varphi,\ \forall \mathbf{m} \in \mathcal{M}_Q \text{ s.t.} \\
&\quad \mathbf{m} \models_Q \S_Q\ \ (\mathcal{C}_{Q\cup U}(f_{Q,Q\cup U}(S_1) \cup f_{U,Q\cup U}(S_2))\}
\end{aligned}
$$

Let $\varphi \in S_1$. Let $\mathbf{m} \in \mathbb{M}_Q$ be such that

$$
\mathbf{m} \models_Q \S_Q\ \ (\mathcal{C}_{Q\cup U}(f_{Q,Q\cup U}(S_1) \cup f_{U,Q\cup U}(S_2)))\,. \tag{V.18}
$$

We want to prove that $\mathbf{m} \models_Q \varphi$. Let $\mathbf{v} \in \mathbb{M}_{Q\cup U}$ be such that

$$
\mathbf{v} \models_{Q\cup U} f_{Q,Q\cup U}(S_1) \cup f_{U,Q\cup U}(S_2) \tag{V.19}
$$

Then, in particular, $\mathbf{v} \models_{Q\cup U} f_{Q,Q\cup U}(\varphi)$. This implies

$$
f_{Q,Q\cup U}(\varphi) \in \mathcal{C}_{Q\cup U}(f_{Q,Q\cup U}(S_1) \cup f_{U,Q\cup U}(S_2)),
$$

and thus, by (V.18),

$$
\mathbf{m} \models_Q \S_Q\ f_{Q,Q\cup U}(\varphi).
$$

From here, we can conclude that $\mathbf{m} \models_Q \varphi$ by property (2) of the definition of §. This concludes the proof of $S_1 \subseteq (S_1 \otimes S_2)^{\downarrow Q}$. The other inequality can be shown in the same way.

□

8.2.1. Constructing a truth valuation. The most important aspect of adjoint valuation algebras is that, given a globally agreeing knowledgebase, it is possible to construct a truth function by combining the information in the knowledgebase. This is proved in the following proposition, which generalises Proposition 2.3 in [**Abr13a**].

PROPOSITION V.24. *Let Φ be an adjoint valuation algebra on a set of variables V. Let $K = \{\phi_1, \ldots, \phi_n\} \subseteq \Phi$ be a knowledgebase. Let*

$$\gamma = \bigotimes_{i=1}^{n} \phi_i. \tag{V.20}$$

Then $\phi_1, \ldots, \phi_n$ agree globally if and only if $\gamma^{\downarrow d(\phi_i)} = \phi_i$. In this case, γ is the most informative of all the possible truth valuations.

PROOF. Suppose $\delta \in \Phi_V$ is a truth valuation for K, i.e. $\delta^{\downarrow d(\phi_i)} = \phi_i$ for all $1 \leq i \leq n$. Since Φ is adjoint, we have

$$\delta \overset{(V.14)}{\preceq} \bigotimes_{i=1}^{n} \delta^{\downarrow d(\phi_i)} = \bigotimes_{i=1}^{n} \phi_i = \gamma$$

Moreover, because projection is monotone by axiom (A13), we have

$$\phi_i \preceq \delta^{\downarrow d(\phi_i)} \overset{(A13)}{\preceq} \gamma^{\downarrow d(\phi_i)} \overset{(V.15)}{\preceq} \phi_i.$$

Thus γ is a truth valuation for K. □

8.3. Detecting disagreement is an inference problem. Thanks to Proposition V.24, the quest for a global truth valuation becomes a much easier task. This, in turn, makes the problem of detecting disagreement significantly simpler. In fact, we can reformulate it using the now familiar concept of inference problem. Given a knowledgebase $\{\phi_1, \ldots, \phi_n\}$, it is sufficient to solve the problem

$$(\phi_1 \otimes \cdots \otimes \phi_n)^{\downarrow d(\phi_i)} \tag{V.21}$$

for all $1 \leq i \leq n$. Then, the knowledgebase agrees globally if and only if the solution to each problem is ϕ_i.

EXAMPLE V.25. Consider the inference problem for indicator functions introduced in Example V.5. In our computations, we showed that

$$(i_1 \otimes i_2 \otimes i_3 \otimes i_4)^{\downarrow \{a_1, b_1\}} \neq i_1.$$

Thus we can immediately conclude that the valuations $i_1, \ldots, i_4$ disagree globally.

8.4. Complete disagreement. Let $\phi_1, \phi_2 \in \Phi$ be two valuations of an adjoint information algebra Φ, and let $d(\phi_1) = S$, $d(\phi_2) = T$, with $S \cap T \neq \emptyset$. To say that ϕ_1 and ϕ_2 disagree amounts to say that not *all* the information carried by ϕ_1 and ϕ_2 can be preserved by combining them. However, *some* of this information is preserved, namely the quantities $\psi_1 := (\phi_1 \otimes \phi_2)^{\downarrow S}$ and $\psi_2 := (\phi_1 \otimes \phi_2)^{\downarrow T}$. Indeed, $\psi_1 \preceq \phi_1$ and $\psi_2 \preceq \phi_2$ represent exactly the portion of information on which the original valuations *do* agree. This can be easily shown by arguing that ψ_1 and ψ_2 agree on their common variables:

$$\psi_1^{\downarrow S\cap T} = \left((\phi_1 \otimes \phi_2)^{\downarrow S}\right)^{\downarrow S\cap T} = (\phi_1 \otimes \phi_2)^{\downarrow S\cap T} = \left((\phi_1 \otimes \phi_2)^{\downarrow T}\right)^{\downarrow S\cap T} = \psi_2^{\downarrow S\cap T}.$$

However, there may be situations where ψ_1 and ψ_2 are null elements of the algebra. This corresponds to a situation where ϕ_1 and ϕ_2 disagree *completely*. In this case, we have

$$\phi_1 \otimes \phi_2 = z_{S\cup T}.$$

The liar cycle of Example V.19 gives a compelling example of complete disagreement. Let us compute the global valuation $\gamma := \bigotimes_{i=1}^n \phi_n$. We have

$$\bigotimes_{i=1}^{n-1} \phi_i = \{\mathbf{v} \in \mathcal{I}_V : \mathbf{v}(\langle s_1, \ldots, s_n\rangle) = \langle 0, 0, \ldots, 0\rangle \text{ or } \langle 1, 1, \ldots, 1\rangle\}$$

Hence,

$$\begin{aligned}\gamma &= \{\mathbf{v} \in \mathcal{I}_V : \mathbf{v}(\langle s_1, \ldots, s_n\rangle) = \langle 0, 0, \ldots, 0\rangle \text{ or } \langle 1, 1, \ldots, 1\rangle\} \\ &\quad \otimes \left\{\mathbf{v} \in \mathcal{I}_{\{s_1, s_n\}} : \mathbf{v}(\langle s_1, s_n\rangle) = \langle 1, 0\rangle \text{ or } \langle 0, 1\rangle\right\} = \emptyset = z_V.\end{aligned}$$

Hence, we conclude that the knowledgebase in question disagrees completely, despite agreeing locally.

In light of this discussion, we introduce the following definition:

Definition V.26. Let Φ be an information algebra, and consider a knowledgebase $\{\phi_1, \ldots, \phi_n\}$ over a set of variables V. We say that $\phi_1, \ldots, \phi_n$ **disagree completely** if $\gamma := \bigotimes_{i=1}^n \phi_i = z_V$, or, equivalently by axiom (A8), if there exists a $1 \leq i \leq n$ such that

$$(\phi_1 \otimes \cdots \otimes \phi_n)^{\downarrow d(\phi_i)} = z_{d(\phi_i)},$$

One can easily show that the knowledgebases of Examples V.6, V.17 and V.18 all disagree completely.

8.4.1. Probabilistic contextuality and complete disagreement. Back in Section 7, we showed that probabilistic contextuality is an instance of disagreement. Here, we show that it can also be seen as an instance of complete disagreement by considering a different valuation algebra.

Suppose we have an empirical model $\{e_C \in \mathcal{D}_R\mathcal{E}(C)\}_{C\in\mathcal{M}}$ on a measurement scenario $\langle X, \mathcal{M}, (O_m)\rangle$. We encode the locally observed probability distributions e_C into a knowledgebase for the adjoint valuation algebra of information sets for the tuple system of probability distributions introduced in Example V.12. The knowledgebase is constituted by single element sets containing the probabilities e_C:

$$K := \{\{e_C\}\}_{C\in\mathcal{M}}.$$

Note that this knowledgebase agrees locally, indeed, for all $C, C' \in \mathcal{M}$,

$$\{e_C\}^{\downarrow C\cap C'} = \{e_C^{\downarrow C\cap C'}\} = \{e_C|_{C\cap C'}\} \overset{(*)}{=} \{e_{C'}|_{C\cap C'}\} = \{e_C^{\downarrow C\cap C'}\} = \{e_C\}^{\downarrow C\cap C'},$$

where we have used no-signalling in equality $(*)$.

By Proposition V.29, a global truth valuation γ for K exists if and only

$$\gamma = \bigotimes_{C\in\mathcal{M}} \{e_C\} = \{d \in \mathcal{D}_R\mathcal{E}(X) : d_{\downarrow C} = e_C,\ \forall C \in \mathcal{M}\}.$$

Therefore, a hypothetical global truth is a set containing all the possible global probability distributions that marginalise to the empirically observed ones. By definition, the model $\{e_C\}_{C\in\mathcal{M}}$ is contextual if and only if $\gamma = \emptyset = z_X$, where, we recall, z_X denotes the null element of the valuation algebra. Thus, we have just proved the following:

Proposition V.27. *Let $\{e_C \in \mathcal{D}_R\mathcal{E}(C)\}_{C\in\mathcal{M}}$ be an empirical model on a scenario $\langle X, \mathcal{M}, (O_m)\rangle$. The model is contextual if and only if K disagrees completely.*

9. Disagreement and possibilistic forms of contextuality

The notions of disagreement and complete disagreement for adjoint information algebras allow to extend the connection with contextuality observed in Section 6 to the level of logical and strong contextuality. Just like in the case of probabilistic contextuality, a logically contextual empirical model can be seen as a locally agreeing knowledgebase which disagrees globally. In this case, the valuation algebra in question is the one of indicator functions.

Let $\mathcal{S} : \mathcal{P}(X)^{op} \to \mathbf{Set}$ be an empirical model on a scenario $\langle X, \mathcal{M}, (O_m)\rangle$. We associate to it the following knowledgebase

$$K := \{i_{\mathcal{S}(C)} \in \mathsf{I}_C\}_{C\in\mathcal{M}},$$

where $i_{\mathcal{S}}(C) : \mathcal{E}(C) = \Omega_C \to \{0,1\}$ is the indicator function of the set $\mathcal{S}(C) \subseteq \mathcal{E}(C) = \Omega_C$.

Lemma V.28. *The knowledgebase K agrees locally.*

Proof. Let $\mathbf{x} \in \mathcal{E}(C\cap C') = \Omega_{C\cap C'}$. We have

$$i_{\mathcal{S}(C)}^{\downarrow C\cap C'}(\mathbf{x}) = \max_{\mathbf{y}\in\mathcal{E}(C\setminus C')} i_{\mathcal{S}(C)}(\mathbf{x},\mathbf{y}) = \begin{cases} 1 & \text{if } \exists \mathbf{y} \in \mathcal{E}(C\setminus C') : \langle \mathbf{x},\mathbf{y}\rangle \in \mathcal{S}(C), \\ 0 & \text{otherwise.} \end{cases}$$

$$= \begin{cases} 1 & \text{if } \mathbf{x} \in \mathcal{S}(C\cap C') \\ 0 & \text{otherwise} \end{cases} = i_{\mathcal{S}(C\cap C')}$$

where the penultimate equality follows from the fact that $\mathcal{S}$ is flasque beneath the cover, and $C\cap C' \subseteq C \in \mathcal{M}$. With the same argument we show that $i_{\mathcal{S}(C')}^{\downarrow C\cap C'} = i_{\mathcal{S}(C\cap C')}$, and we conclude

$$i_{\mathcal{S}(C)}^{\downarrow C\cap C'} = i_{\mathcal{S}(C\cap C')} = i_{\mathcal{S}(C')}^{\downarrow C\cap C'},$$

which means that K agrees locally. □

This proposition reiterates the idea, introduced in Section 7, that empirical models correspond to locally agreeing knowledgebases.

PROPOSITION V.29. *The knowledgebase K disagrees globally if and only if $\mathcal{S}$ is logically contextual. It disagrees completely if and only if $\mathcal{S}$ is strongly contextual.*

PROOF.

- Suppose K disagrees globally. Let $\gamma := \bigotimes_{C\in\mathcal{M}} i_{\mathcal{S}(C)}$. By Proposition V.24, there exists a context $C_0 \in \mathcal{M}$ such that $\gamma^{\downarrow C_0} \neq i_{\mathcal{S}(C_0)}$. By Proposition V.23, this implies $\gamma^{\downarrow C_0} \prec i_{\mathcal{S}(C_0)}$, which means that there exists a local section $\mathbf{x} \in \mathcal{S}(C_0)$ such that $\gamma^{\downarrow C_0}(\mathbf{x}) \lneq i_{\mathcal{S}(C_0)}(\mathbf{x})$. Hence, $\gamma^{\downarrow C_0}(\mathbf{x}) = 0$ and $i_{\mathcal{S}(C_0)}(\mathbf{x}) = 1$. We will now show that $\mathcal{S}$ is logically contextual at $\mathbf{x}$. Suppose $\neg\mathsf{LC}(\mathcal{S}, \mathbf{x})$ by contradiction. Then there exists a global section $\mathbf{g} \in \mathcal{S}(X)$ such that $\mathbf{g}_{\downarrow C_0} = \mathbf{x}$. Because $\mathbf{g} \in \mathcal{S}(X)$, $\mathbf{g}_{\downarrow C} \in \mathcal{S}(C)$ for all $C \in \mathcal{M}$, which means that $i_{\mathcal{S}(C)}(\mathbf{g}^{\downarrow C}) = 1$ for all $C \in \mathcal{M}$. Therefore,
$$\gamma(\mathbf{g}) = \prod_{C\in\mathcal{M}} i_{\mathcal{S}(C)}(\mathbf{g}_{\downarrow C}) = 1.$$
This implies
$$\gamma^{\downarrow C_0}(\mathbf{x}) = \gamma^{\downarrow C_0}(\mathbf{g}_{\downarrow C_0}) = \max_{\substack{\mathbf{y}\in\Omega_X \\ \mathbf{y}_{\downarrow C_0} = \mathbf{g}_{\downarrow C_0}}} \gamma(\mathbf{y}) = 1$$
which is a contradiction.

  Now, Suppose $\mathcal{S}$ is logically contextual at a section $\mathbf{x} \in \mathcal{S}(C_0)$. We have
$$\gamma^{\downarrow C_0}(\mathbf{x}) = \max_{\substack{\mathbf{y}\in\Omega_X \\ \mathbf{y}_{\downarrow C_0} = \mathbf{x}}} \gamma(\mathbf{y}) = \max_{\substack{\mathbf{y}\in\Omega_X \\ \mathbf{y}_{\downarrow C_0} = \mathbf{x}}} \prod_{C\in\mathcal{M}} i_{\mathcal{S}(C)}(\mathbf{y}_{\downarrow C}). \tag{V.22}$$
  Suppose by contradiction $\gamma^{\downarrow C_0}(\mathbf{x}) = 1$. By (V.22), there exists a $\mathbf{y} \in \Omega_X = \mathcal{E}(X)$ such that $\mathbf{y}_{\downarrow C} \in \mathcal{S}(C)$ and $\mathbf{y}_{\downarrow C_0} = \mathbf{x}$. By condition 3 of the definition of a possibilistic empirical model, this implies that $\mathbf{y} \in \mathcal{S}(X)$, which means that it is a global section containing $\mathbf{x}$. Thus, we reach a contradiction as we have just proved $\neg\mathsf{LC}(\mathcal{S}, \mathbf{x})$.
- We will now prove that K disagrees completely if and only if $\mathcal{S}$ is strongly contextual. Recall that the null element of the algebra is the zero function $z_X \equiv 0$. We have
$$\begin{aligned}\neg\mathsf{SC}(\mathcal{S}) &\Leftrightarrow \exists\mathbf{g} \in \mathcal{E}(X) : i_{\mathcal{S}(C)}(\mathbf{g}_{\downarrow C}) = 1 \ \forall C \in \mathcal{M} \\ &\Leftrightarrow \exists\mathbf{g} \in \mathcal{E}(X) : \gamma(\mathbf{g}) = \prod_{C\in\mathcal{M}} i_{\mathcal{S}(C)}(\mathbf{g}_{\downarrow C}) = 1 \Leftrightarrow \gamma \not\equiv 0.\end{aligned}$$

□

9.1. Detecting logical and strong contextuality is an inference problem. Thanks to Proposition V.29 and the results of the previous sections, we can easily translate the problem of detecting logical and strong contextuality into inference problems. This aspect is particularly important as it allows to develop new algorithms for the detection of contextuality based on efficient generic inference methods, as we shall see in Section 10.

The following proposition follows immediately from Proposition V.29 and the results of Sections 8.3 and 8.4:

PROPOSITION V.30. *Let $\mathcal{S} : \mathcal{P}(X) \to$ **Set** be an empirical model over a scenario $\langle X, \mathcal{M}, (O_m) \rangle$. Then*

- *The model $\mathcal{S}$ is logically contextual if and only if there exists a $\overline{C} \in \mathcal{M}$ such that*

$$\left(\bigotimes_{C \in \mathcal{M}} i_{\mathcal{S}(C)} \right)^{\downarrow \overline{C}} \neq i_{\mathcal{S}(\overline{C})} \tag{V.23}$$

- *The model $\mathcal{S}$ is strongly contextual if and only if, for all $\overline{C} \in \mathcal{M}$,*

$$\left(\bigotimes_{C \in \mathcal{M}} i_{\mathcal{S}(C)} \right)^{\downarrow \overline{C}} = z_{\overline{C}}. \tag{V.24}$$

It follows that, in order to determine whether a model $\mathcal{S}$ is strongly contextual, one has to solve a single inference problem. On the other hand, to determine whether $\mathcal{S}$ is logically contextual, one has to solve anywhere from 1 to $|\mathcal{M}|$ distinct problems.

10. Inference algorithms for contextuality

From a computational perspective, the problem of detecting non-locality and contextuality is highly complex. This aspect has been highlighted in a recent paper by Abramsky, Gottlob & Kolaitis, which shows that recognising non-locality in n-partite Bell-type scenarios is NP-complete [**AGK13**]. Despite this, computational explorations of contextual behaviour [**Man13, MB12**] have proven useful for various tasks, such as the classification of logically non-local quantum states [**AC14**], and the partial characterisation of strongly non-local 3-qubit quantum states [**ABC$^+$17**], which will be presented in Chapter VII. For this reason, it is highly desirable to improve the complexity of current algorithms for the identification of contextuality.

The translation of the question of searching for non-locality and contextuality in the form of an inference problem, expressed in Proposition V.30, paves the way for the application of efficient inference algorithms to test for these phenomena.

In this section, we introduce new algorithms for logical and strong contextuality based on mainstream methods of generic inference. Although it is impossible to precisely characterise their complexity for general contextuality – as it highly depends on the structure of the measurement scenario – we show that they significantly outperform current methods for non-locality, most notably in its logical form, in (n, k, l) scenarios.

10.1. Current algorithms and their complexity. Current algorithms for the detection of non-locality and contextuality rely mostly on solving linear systems of equations. Let us briefly review how such systems are constructed.

Given a scenario $\Sigma = \langle X, \mathcal{M}, (O_m) \rangle$, one defines its *incidence matrix* $\mathbf{M}$, which describes the way global events restrict to local events. The rows of $\mathbf{M}$ are indexed by $\bigsqcup_{C \in \mathcal{M}} \mathcal{E}(C)$, i.e. the local events at each context, while its columns are indexed by $\mathcal{E}(X)$, the global events. For all $C \in \mathcal{M}$, $s_C \in \mathcal{E}(C)$ and $g \in \mathcal{E}(X)$, let

$$\mathbf{M}[\langle C, s_C \rangle, g] := \begin{cases} 1 & \text{if } g|_C = s_C, \\ 0 & \text{otherwise.} \end{cases}$$

Let R be a semiring and consider an empirical model $e = \{e_C \in \mathcal{D}_R\mathcal{E}(C)\}_{C\in\mathcal{M}}$ on Σ. We represent it as an $\sum_{C\in\mathcal{M}} |\mathcal{E}(C)|$-dimensional vector with coefficients in R, defined as follows: for all $C \in \mathcal{M}$ and $s_C \in \mathcal{E}(C)$,

$$\mathbf{V}_e[\langle C, s_C\rangle] := e_C(s_C).$$

We augment both $\mathbf{M}$ and $\mathbf{V}_e$ into $\mathbf{M}'$ and $\mathbf{V}'_e$ by adding an extra row, every entry of which is $1 \in R$.[9] Finally, we introduce an $|\mathcal{E}(X)|$-dimensional vector of unknowns $\mathbf{X}$. A result of [**AB11a**] shows that the solutions to the system

$$\mathbf{M}'\mathbf{X} = \mathbf{V}'_e \tag{V.25}$$

correspond bijectively to global sections for e. Therefore, determining whether e is contextual amounts to establish whether (V.25) has any solution.

10.1.1. Probabilistic and strong contextuality for probabilistic models. For probabilistic empirical models, i.e. when $R = \mathbb{R}_{\geq 0}$, we add to (V.25) the condition $\mathbf{X} \geq 0$ and solve over $\mathbb{R}$. The computational complexity of establishing whether a linear system in $\mathbb{R}$ with n variables has any solution has a complexity of $\mathcal{O}(n^\epsilon)$, with $\epsilon \geq 2$. The best known bound is $\epsilon = 2.373$, achieved by Williams' improved version [**Wil12**] of the Coppersmith–Winograd algorithm [**CW90**]. By applying this method to solve (V.25), one obtains an algorithm for the detection of probabilistic contextuality with complexity[10]

$$\mathcal{O}\left(|\mathcal{E}(X)|^\epsilon\right). \tag{V.26}$$

In the case of an (n, k, l) Bell-type scenario, we have $|\mathcal{E}(X)| = l^{nk}$. Thus, the complexity of detecting probabilistic non-locality is

$$\mathcal{O}\left(l^{\epsilon nk}\right). \tag{V.27}$$

This method can be improved to enable the computation of the *contextual fraction* of an empirical model [**ABM17**], which not only allows to determine whether a model exhibits probabilistic contextuality, but also quantifies the amount of non-classicality present in its distributions. In order to determine the contextual fraction, one relaxes system (V.25) into the following linear programming problem:

$$\begin{array}{ll} \texttt{Find} & \mathbf{c} \in \mathbb{R}^{|\mathcal{E}(X)|} \\ \texttt{maximising} & \mathbf{1}\cdot\mathbf{c} \\ \texttt{subject to} & \mathbf{Mc} \leq \mathbf{V}_e \\ \texttt{and} & \mathbf{c} \geq 0 \end{array} \tag{V.28}$$

While solutions to (V.25) correspond to global sections for e, this problem merely looks for *subdistributions*[11] c compatible with e, i.e. such that $c|_C(s_C) \leq e_C$ for all $C \in \mathcal{M}$. These subdistributions can be interpreted as non-contextual 'approximations' of e. Hence, a solution c to (V.28) is nothing but the best possible classical approximation

[9] This procedure is needed to ensure that the solutions to the system (V.25), presented later, are normalised.

[10] The extra condition $\mathbf{X} \geq 0$ can be proven not to influence the overall complexity of establishing whether the system has any solution.

[11] A subdistribution is a the same as a distribution, except for the requirement of normalisation.

to the model e. Its weight $w(c)$ constitutes the *non-contextual* (or *non-local*) *fraction* $\mathsf{NCF}(e)$ of the model:

$$\mathsf{NCF}(e) = w(c) := \sum_{\substack{C \in \mathcal{M} \\ s_C \in \mathcal{E}(C)}} c(s_C) \in [0,1]$$

The *contextual* (or *non-local*) *fraction* is then defined to be $\mathsf{CF}(e) := 1 - \mathsf{NCF}(e)$. A key result of [**ABM17**] shows that a model e is probabilistically contextual if and only if $\mathsf{CF}(e) > 0$, and it is strongly contextual if and only if $\mathsf{CF}(e) = 1$. Hence this method can be used to detect both probabilistic and strong contextuality.

There are many methods to solve linear programming problems. In 2018, Cohen, Lee & Song presented a new improved *path-following interior point algorithm* for linear programming which has the same complexity as matrix multiplication [**CLS18**]. To the author's knowledge, this is the best known algorithm in terms of worst-case running time. Therefore, solving (V.28) has the same complexity as solving (V.25). Hence the complexities introduced in (V.26) and (V.27) apply both to probabilistic and strong contextuality. However, it is important to remark that, for strong contextuality, this statement is only valid for probabilistic empirical models. If we are required to recognise strongly contextual behaviour in a possibilistic model whose probabilistic structure is unknown (or does not exist), solving system (V.25) for $R = \mathbb{B}$ is the only viable option. This brings us to the next section:

10.1.2. Logical and strong contextuality for possibilistic models. In the presence of a possibilistic empirical model, that is when $R = \mathbb{B}$, the system of equations (V.25) takes the form of a *boolean satisfiability (SAT) problem*, which is notoriously difficult to solve, as it constitutes the first known instance of NP-complete problem [**Coo71, Lev73**]. The complexity of solving a SAT problem with n clauses is $\mathcal{O}\left(2^{\delta n}\right)$, where $\delta > 0$ (the best known runtime has $\delta = 0.386$ [**Sch99, PPSZ05**]). The SAT problem generated by (V.25) has $\sum_{C \in \mathcal{M}} |\mathcal{E}(C)|$ clauses, thus the worst-case complexity of detecting contextuality on a general scenario $\langle X, \mathcal{M}, (O_m) \rangle$ is

$$\mathcal{O}\left(2^{\delta \cdot \left(\sum_{C \in \mathcal{M}} |\mathcal{E}(C)|\right)}\right).$$

Consequently, the complexity of detecting logical and strong non-locality for a possibilistic model on an (n, k, l) scenario is

$$\mathcal{O}\left(2^{\delta \cdot (kl)^n}\right), \tag{V.29}$$

which is extremely inefficient. It is therefore natural to ask whether this method can be improved. On this subject, Mansfield & Fritz proved that polynomial time algorithms can be implemented for all $(2, k, 2)$ and $(2, 2, l)$ scenarios, but also conjectured the general problem for (n, k, l) scenarios to be NP-hard [**MF12, AGK13, Sim18**].

10.2. Inference algorithms. In the following sections, we propose a different approach to the question of detecting contextuality. It is based on methods of generic inference which rest on the idea of *local computation*, a scheme that involves combining information at the local level of the relevant variables for the query.

Before focusing on the solutions of problems (V.23) and (V.24), which are central to the identification of non-classical behaviour, we are going to present these methods in complete generality. For this purpose, we introduce a general single-query inference problem which will serve as a model for the discussion carried out in the following pages.

Let Φ be a valuation algebra, and $K = \{\phi_1, \ldots, \phi_m\} \subseteq \Phi$ a knowledgebase. We will consider the following problem:

$$(\phi_1 \otimes \cdots \otimes \phi_m)^{\downarrow D}, \tag{V.30}$$

where $D \subseteq d\,(\phi_1 \otimes \cdots \otimes \phi_m)$.

10.3. The fusion algorithm. The fusion algorithm [**She92**] is the simplest method for the solution of inference problems. It rests on the idea of sequentially eliminating those variables that do not appear in the query of the problem, a procedure inspired by traditional dynamic programming [**BB72**]. The process of variable elimination can be formally defined as follows:

DEFINITION V.31. Let Φ be a valuation algebra on a set of variables V. Given a valuation $\phi \in \Phi$ and a variable $x \in d(\phi)$ we define the **elimination** of variable x from ϕ to be

$$\phi^{-x} := \phi^{\downarrow d(\phi) \setminus x}.$$

We will take advantage of the following property of variable elimination, proved in Theorem 3.1 of [**PK12**]: given a knowledgebase K, eliminating a variable x only affects those valuations in K whose domain contains x. That is,

$$\left(\bigotimes K\right)^{-x} = \left(\bigotimes_{\phi \in K:\ x \in d(\phi)} \phi\right)^{-x} \otimes \left(\bigotimes_{\phi \in K:\ x \notin d(\phi)} \phi\right).$$

Consequently, once an elimination sequence for the variables not in the query is established, these variables can be eliminated *locally* where they appear. This is significantly more efficient than combining all the information in K before eliminating the irrelevant variables. Following this intuition, one obtains a simple procedure to solve (V.30) which consists of the following steps:

(1) Define an enumeration of the variables to be eliminated: $\{x_1, \ldots, x_n\}$.
(2) For each x_i,
    (a) Combine all the valuations that contain variable x_i into a new valuation ψ.
    (b) Eliminate x_i from ψ.
    (c) Update K by erasing all the valuations involved in the process and replacing them with ψ.
(3) Once all the variables have been eliminated, return the combination of what is left.

The fusion method is described more formally in Algorithm 1.

10.4. The complexity of the fusion algorithm. Unfortunately, it is impossible to characterise the complexity of the fusion algorithm in general as it highly depends on the definition of combination and projection of the algebra in question. For our purposes, we will restrict ourselves to the algebra of indicator functions over a set of variables V. In this case, combining valuations $\phi_1, \ldots, \phi_k$ involves going through each element of Ω_S,

Algorithm 1 General fusion algorithm

Input: $\{\phi_1, \dots, \phi_m\}$, D
Output: $(\phi_1, \dots, \phi_m)^{\downarrow D}$

procedure
  $\Psi \leftarrow \{\phi_1, \dots, \phi_m\}$
  for $x \in \bigcup_i d(\phi_i) \setminus D$ **do**
    $\Gamma \leftarrow \{\phi_i \in \Psi \mid x \in d(\phi_i)\}$
    $\psi \leftarrow \bigotimes \Gamma$
    $\Psi \leftarrow (\Psi \setminus \Gamma) \cup \{\psi^{-x}\}$
  end for
  return $\bigotimes \Psi$
end procedure

where $S = \bigcup_{i=1}^{k} d(\phi_i)$, and performing elementary operations. Thus, the complexity of combining $\phi_1, \dots, \phi_k$ is bounded by $\mathcal{O}(d^{|S|})$, where $d := \max_{x \in V} |\Omega_x|$ denotes the size of the largest frame. Indeed, $|\Omega_S| = \prod_{v \in S} |\Omega_v| \leq d^{|S|}$. Similarly, one proves that the complexity of projection of a valuation with domain S is also bounded by $\mathcal{O}(d^{|S|})$.

Now, suppose we have an elimination sequence $\{x_1, \dots, x_n\}$ for the solution of the general inference problem (V.30) using the fusion algorithm. When variable x_i is eliminated, the algorithm combines all valuations whose domain contain x_i. Thus, if m_i valuations contain x_i before iteration i, its elimination requires $m_i - 1$ combinations and one variable elimination.[12] Hence, by denoting $T := d(\phi_1 \otimes \cdots \otimes \phi_m)$, the total complexity of the fusion process is given by

$$\mathcal{O}\left(\sum_{i \in T \setminus D} m_i \cdot d^{|S_i|} \right), \tag{V.31}$$

where

$$S_i := \bigcup_{\substack{\phi \in K: \\ x_i \in d(\phi)}} d(\phi)$$

is the union of all domains containing x_i. In order to keep track of the size of S_i for each iteration i, it is convenient to graphically represent its functioning in the form of a graph, as proposed by Shenoy in [**She96**].

10.4.1. Graphical representation of the fusion algorithm. Suppose we apply the fusion algorithm to the inference problem (V.30). Let $L_0 := \{d(\phi_i)\}_{i=1}^m$ denote the list of domains of the valuations in K. Given a sequence $\{x_1, \dots, x_n\}$ of variables to be eliminated, we progressively build a labelled undirected graph $(\mathsf{V}, \mathsf{E}, \lambda)$, where V is the set of nodes, E is the set of edges, and $\lambda : \mathsf{V} \to \mathcal{P}\left(\bigcup_{i=1}^m d(\phi_i)\right)$ is a labelling function. The purpose of this graph is to describe the amount of operations needed at each step of the fusion algorithm according to the specified elimination sequence.

We assume the graph to be empty at the start, i.e. $\mathsf{V} = \mathsf{E} = \emptyset$. At the i-th iteration of the algorithm, when variable x_i is eliminated in the fusion process, we compute the

[12]The complexity of variable elimination can be easily proved to equal the complexity of projection.

union S_i of all domains that contain x_i, that is

$$S_i := \bigcup_{\substack{S \in L_{i-1}: \\ x_i \in S}} S \tag{V.32}$$

Then, we let

$$L_i := (L_{i-1} \setminus \{S \in L_{i-1} : x_i \in S\}) \cup \{S_i \setminus \{x_i\}\},$$

and define a new node i of the graph with label $\lambda(i) := S_i$. This node is tagged with a colour and added to the graph. Then, we go through every other coloured node $v \in \mathsf{V}$ of the graph: if v contains variable x_i, then we remove its colour and we add an edge $(i, v) \in E$. This process is repeated until all the variables have been eliminated, and corresponds to the **for** cycle of Algorithm 1. The final combination $\bigotimes \Psi$ is represented by adding one last coloured node labelled by D and connecting it to all remaining coloured nodes. Finally, all colours are removed.

It can be shown that, for any elimination sequence, the resulting graph is in fact a tree. More specifically, it is a *join tree* [**PK12**]:

DEFINITION V.32. A labelled tree $(\mathsf{V}, \mathsf{E}, \lambda)$ is a **join tree** if for any two nodes $i, j \in \mathsf{V}$ and $x \in \lambda(i) \cap \lambda(j)$, $x \in \lambda(k)$ for all nodes k on the path between i and j.

In the interest of time, we will not list here any general examples of the construction of join trees associated with fusion processes (a large number of instances can be found in [**She96, PK12**]). However, we will present examples specific to the application of the fusion algorithm to detecting contextuality in Sections 10.5.4 and 10.8.3.

10.4.2. Complexity considerations. The join tree associated to a run of the fusion algorithm based on a certain elimination sequence describes the functioning of the algorithm and keeps track of the complexity of the combinations performed at each iteration by labelling the corresponding node with S_i, whose cardinality plays a central role in (V.31). The size of these labels is bounded by $\max_{i \in \mathsf{V}} |\lambda(i)|$, a quantity which is usually expressed through the concept of *treewidth*:

DEFINITION V.33. The **treewidth** of a join tree $(\mathsf{V}, \mathsf{E}, \lambda)$ is given by

$$\omega := \max_{i \in \mathsf{V}} |\lambda(i)| - 1.$$

The **treewidth of an inference problem**, denoted by ω^*, is the minimum treewidth over all join trees created from all possible elimination sequences.[13]

At any iteration of the fusion algorithm, the locally computed valuation ψ is reinserted in the current knowledgebase of the algorithm (Ψ in Algorithm 1). Thus, we have

$$\sum_{i \in T \setminus D} m_i \leq m + |T \setminus D| \leq m + |T|.$$

Given (V.31), these observations collectively determine an overall complexity of

$$\mathcal{O}\left((m + |T|) \cdot d^{\omega+1}\right). \tag{V.33}$$

[13]The decrement of 1 in this definition is due to technical reasons which are not relevant for the purpose of this chapter (see [**PK12**] for details).

Of course, this complexity highly depends on the elimination sequence chosen. An optimal elimination sequence minimises the treewidth of the corresponding join tree, and yields an optimal complexity of

$$\mathcal{O}\left((m+|T|)\cdot d^{\omega^*+1}\right). \tag{V.34}$$

10.5. A fusion algorithm for contextuality. Thanks to the connection between contextuality and inference problems implemented in the previous sections, the fusion algorithm can be naturally adapted to detect logical and strong contextuality in empirical models. By proposition V.30, the detection of both logical and strong contextuality rests on the solution of the following inference problem for the valuation algebra of indicator functions:

$$\left(\bigotimes_{C\in\mathcal{M}} i_{\mathcal{S}(C)}\right)^{\downarrow\overline{C}}. \tag{V.35}$$

The following algorithms are based on solving (multiple copies of) (V.35) using the fusion algorithm.

10.5.1. Strong contextuality. In order to determine whether a model $\mathcal{S}$ on a scenario $\langle X,\mathcal{M},(O_m)\rangle$ is strongly contextual, it is sufficient to solve inference problem (V.35) for an arbitrary context $\overline{C}\in\mathcal{M}$, and check whether the answer is $z_{\overline{C}}$. In the interest of reducing complexity, it is convenient to take the largest context in $\mathcal{M}$, as it minimises the amount of variables to be eliminated. Given this premise, we present the fusion method for strong contextuality in Algorithm 2.

Algorithm 2 Fusion algorithm for strong contextuality

Input: $\mathcal{S}$, $\langle X,\mathcal{M},(O_m)\rangle$
Output: $\mathsf{SC}(\mathcal{S})$.

procedure
    $\Psi \leftarrow \{i_{\mathcal{S}(C)}\}_{C\in\mathcal{M}}$
    $\overline{C} \leftarrow$ largest context of $\mathcal{M}$
    for $m \in X\setminus\overline{C}$ **do**
        $\Gamma \leftarrow \{i_{\mathcal{S}(C)} \mid m\in C\}$
        $\psi \leftarrow \bigotimes\Gamma$
        $\Psi \leftarrow (\Psi\setminus\Gamma)\cup\{\psi^{-m}\}$
    end for
    if $\bigotimes\Psi = z_{\overline{C}}$ **then return** `true`
    else return `false`
    end if
end procedure

10.5.2. Logical contextuality. The fusion method for logical contextuality, presented in Algorithm 3, is slightly more complicated. In this case, one needs to solve the inference problem (V.35) for multiple contexts $C\in\mathcal{M}$ until a solution different from $i_{\mathcal{S}(C)}$ is found. If such a solution is not found after having solved the problem for all $C\in\mathcal{M}$, we can conclude that the model is not logically contextual.

Algorithm 3 Fusion algorithm for logical contextuality

Input: $\mathcal{S}$, $\langle X, \mathcal{M}, (O_m)\rangle$
Output: $\mathsf{LC}(\mathcal{S})$.

```
1: procedure
2:     Ψ ← {i_S(C)}_C∈M
3:     for C ∈ M do
4:         for m ∈ X \ C do
5:             Γ ← {i_S(C) | m ∈ C}
6:             ψ ← ⨂ Γ
7:             Ψ ← (Ψ \ Γ) ∪ {ψ^−m}
8:         end for
9:         Ψ_C ← ⨂ Ψ
10:        if Ψ_C ≠ i_S(C) then return true
11:        end if
12:    end for
13:    return false
14: end procedure
```

10.5.3. The complexity of the fusion algorithm for contextuality. Consider a general measurement scenario $\Sigma := \langle X, \mathcal{M}, (O_m)\rangle$, and let $O \in (O_m)_{m\in X}$ denote its largest outcome set. Using (V.33), the complexity of detecting strong contextuality of an empirical model on Σ using the fusion algorithm is

$$\mathcal{O}\left((|\mathcal{M}| + |X|) \cdot |O|^{\omega+1}\right), \tag{V.36}$$

For logical contextuality, a worst-case factor of $|\mathcal{M}|$ must be added to account for the **for** cycle of line 3 in Algorithm 3:

$$\mathcal{O}\left(|\mathcal{M}| \cdot (|\mathcal{M}| + |X|) \cdot |O|^{\omega+1}\right). \tag{V.37}$$

Unfortunately, it is impossible to bound ω in general as it greatly depends on the structure of the measurement scenario and the elimination sequence chosen. However, it turns out that we *can* compute the optimal treewidth ω^* in the case of (n, k, l) scenarios, as we shall see in the following section.

10.5.4. The complexity of the fusion algorithm for non-locality in (n, k, l) scenarios. Let $\Sigma = \langle X, \mathcal{M}, O\rangle$ be an (n, k, l) Bell-type scenario. By definition, X is partitioned into subsets $\{X_i\}_{i=1}^n$. We will adopt the following notation: for all $1 \le i \le n$,

$$X_i := \{1_i, 2_i, \ldots, k_i\},$$

so that i_j denotes the i-th measurement for the j-th party.

Let $\mathcal{S}$ be an empirical model on Σ. Both Algorithms 2 and 3 are based on the solution of (V.35). W.l.o.g. we assume $\overline{C} = \{1_1, 1_2, \ldots, 1_n\}$, and propose the following elimination sequence:

$$2_1, 3_1, \ldots, k_1, 2_2, 3_2, \ldots, k_2, \ldots, 2_n, 3_n, \ldots, k_n. \tag{V.38}$$

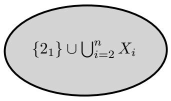
Let us build the join tree corresponding to this sequence. At the beginning of the algorithm, we have $L_0 = \mathcal{M}$.

- *Elimination of* 2_1:

$$S_1 = \{2_1\} \cup \bigcup_{i=2}^{n} X_i$$

$$L_1 = \{\{i_1, \ldots, i_n\} \in \mathcal{M} \mid i_1 \neq 2_1\} \cup \left\{ \bigcup_{i=2}^{n} X_i \right\}$$

We add the first node of the join tree, labelled with S_1, and tagged with a colour:

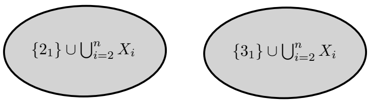


- *Elimination of* 3_1:

$$S_2 = \{3_1\} \cup \bigcup_{i=2}^{n} X_i$$

$$L_2 = \{\{i_1, \ldots, i_n\} \mid i_1 \neq 2_1, 3_1\} \cup \left\{ \bigcup_{i=2}^{n} X_i \right\}$$

Another node is added to the tree. The first node does not contain variable 3_1, thus no edge is added, and no colour is removed:

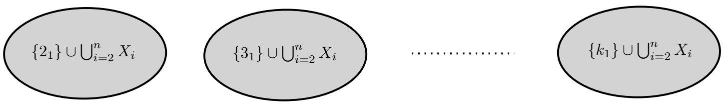


- *Elimination of* $4_1, \ldots, k_1$: The same procedure is repeated until iteration $k-1$, where variable k_1 is eliminated:

$$S_{k-1} = \{k_1\} \cup \bigcup_{i=2}^{n} X_i$$

$$L_{k-1} = \{\{1_1, i_2 \ldots, i_n\}\} \cup \left\{ \bigcup_{i=2}^{n} X_i \right\}$$

At this point, the tree is just a collection of $k-1$ disjoint coloured nodes:

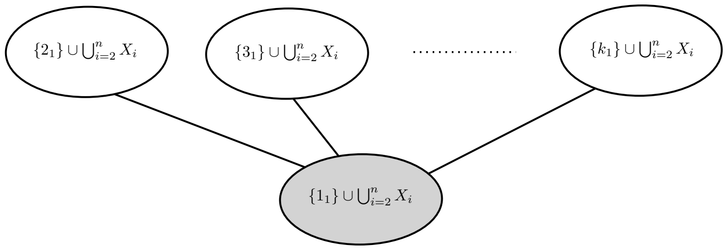


- *Elimination of* 2_2:

$$S_k = \{1_1\} \cup \bigcup_{i=2}^{n} X_i$$

$$L_k = \{\{1_1, i_2, \ldots, i_n\} \mid i_2 \neq 2_2\} \cup \left\{ \bigcup_{i=2}^{n} X_i \setminus \{2_2\} \right\}$$

The node associated with this iteration is connected to all the preceding nodes, since they all contain variable 2_2. Consequently, all their colours are removed:

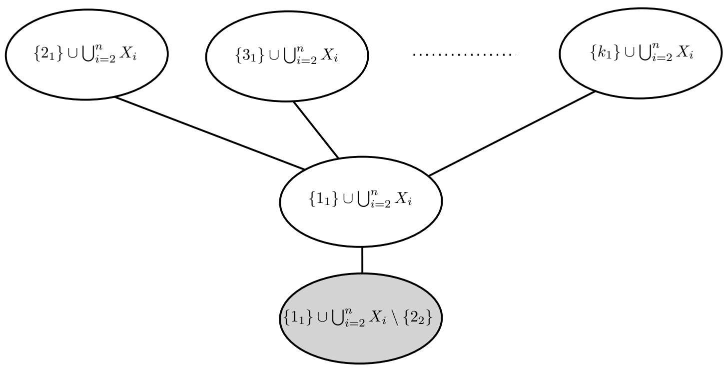


- *Elimination of* 3_2:

$$S_{k+1} = \{1_1\} \cup \bigcup_{i=2}^{n} X_i \setminus \{2_2\}$$

$$L_{k+1} = \{\{1_1, i_2 \ldots, i_n\} \mid i_2 \neq 2_2, 3_2\} \cup \left\{ \bigcup_{i=2}^{n} X_i \setminus \{2_2, 3_2\} \right\}$$

The node corresponding to this iteration is connected to the previous since its label contains variable 3_2:



- *Elimination of* $4_2, \ldots, k_n$: The construction of the join tree for the remaining variables is straightforward: each node is connected to the one corresponding to the previous iteration, and the labels are strictly decreasing in size. The complete join tree is pictured in Figure V.5.

The treewidth of this join tree is $\omega = k(n-1)$. It is crucial to remark that *any* elimination sequence gives rise to a join tree with a treewidth of a at least $k(n-1)$. Indeed, the elimination of any variable $i_j \in X_j$ at the beginning of the algorithm gives rise to a node labelled by

$$\{i_j\} \cup \bigcup_{l \neq j} X_l.$$

We conclude that $\omega^* = k(n-1)$, which means that elimination sequence (V.38) is *optimal* for the solution of problem (V.35).

Thanks to this discussion we obtain the following optimised complexities for Algorithms 2 and 3:

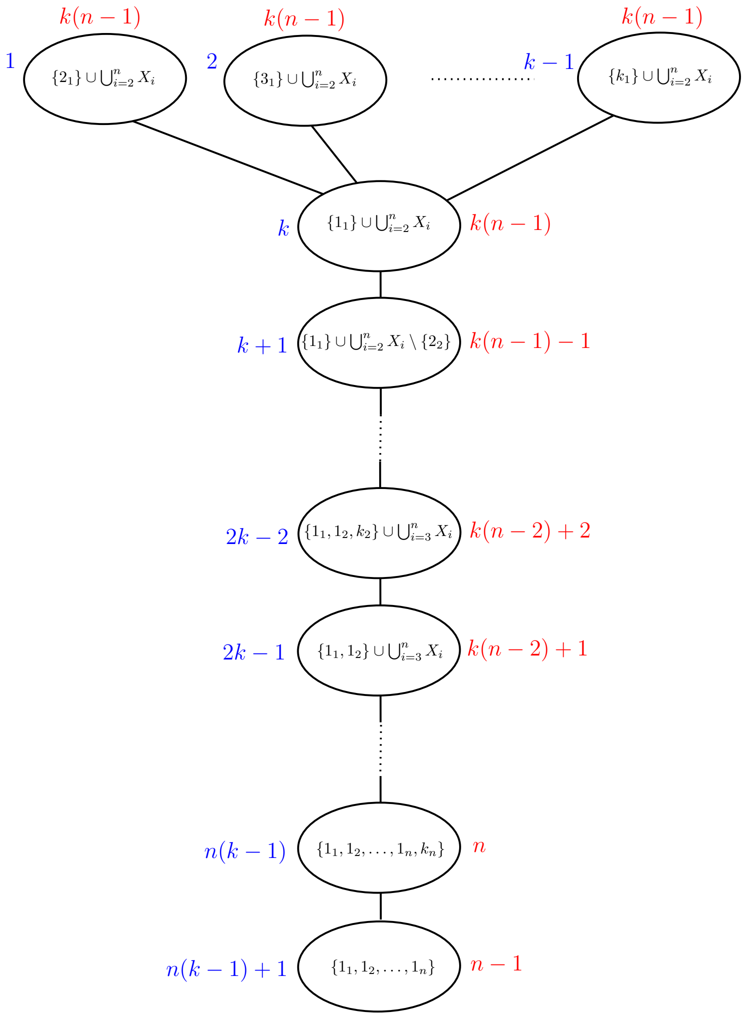


FIGURE V.5. The join tree corresponding to a solution of (V.35) using elimination sequence (V.38). The blue numbers indicate the iteration of each node, while the red numbers display their width.

- By (V.36), the optimal complexity of determining whether a model $\mathcal{S}$ on an (n, k, l) scenario is strongly non-local using Algorithm 2 is:

$$\mathcal{O}\left((k^n + kn) \cdot l^{k(n-1)+1}\right) = \mathcal{O}\left(k^n \cdot l^{k(n-1)+1}\right). \tag{V.39}$$

- By (V.37), the optimal complexity of determining whether $\mathcal{S}$ is logically non-local using 3 is:

$$\mathcal{O}\left(k^n(k^n + kn) \cdot l^{k(n-1)+1}\right) = \mathcal{O}\left(k^{2n} \cdot l^{k(n-1)+1}\right). \tag{V.40}$$

By comparing these results to (V.27) and (V.29), we conclude that both these algorithms are significantly faster than the current state of the art, reviewed in Section 10.1.

The improvement is particularly substantial for logical contextuality, as well as strong contextuality for possibilistic models whose probabilistic structure is unknown.

10.6. Message passing schemes and the collect algorithm. In this section, we introduce a different viewpoint on local computation which eventually leads to a second algorithm for the detection of contextuality based on the *collect algorithm* for single-query inference problems.

The fundamental difference between the collect algorithm and the fusion algorithm is that the join tree construction is separated from the inference process. In the case of fusion, the join tree is simply a description of the functioning of the algorithm. Instead, the collect algorithm requires the construction of a join tree first, and then functions as a message-passing process, guided by the structure of the tree. The key advantage is that performance may potentially be boosted by choosing a suitable join tree. Let us present some details on what 'suitable' means in this case.

10.6.1. Covering join trees. The join tree associated with a run of the fusion algorithm satisfies the following property, which simply follows from (V.32): for each element ϕ_i of the knowledgebase of the inference problem, there exists a node j of the tree such that $d(\phi_i) \subseteq \lambda(j)$. Thanks to this feature, one can assign each knowledgebase factor to the first node of the join tree containing its domain. Of course, some nodes may contain more than one valuation, while others may remain empty. Assuming the valuation algebra in question has neutral elements, we shall assign valuation $e_{\lambda(j)}$ to each of the empty nodes j. Furthermore, if a node contains multiple knowledgebase factors, they shall be combined into a unique valuation. As a result, we obtain a join tree such that each node $i \in \mathsf{V}$ contains exactly one valuation ψ_i, where $d(\psi_i) \subseteq \lambda(i)$. Each ψ_i has to be intepreted as the *initial content* of node i.

With this premise, the fusion process can be interpreted as a *message-passing scheme* [**SS90**], where nodes act as virtual processors that communicate by exchanging messages. At each iteration of the algorithm, the corresponding node receives a valuation by its predecessors, combines it with its initial content, eliminates its designated variable, and sends the result to its unique neighbor corresponding to a later step of the process. More precisely, at step i of the algorithm, when variable x_i is eliminated, node i contains valuation η_i (which consists of its initial information ψ_i combined with all messages sent to i by its predecessors), then it computes $\eta_i^{-x_i}$ and sends the result to its immediate successor, or *child node* $\mathsf{ch}(i)$. The child node, in turn, will compute $\eta_{\mathsf{ch}(i)} \otimes \eta_i^{-x_i}$ and set it as its current value. This process is reiterated until the last node, which will eventually contain the solution to the inference problem.

The idea of a message-passing scheme for the solution of an inference problem can be implemented independently from the join tree structure imposed by the fusion process. For this purpose, we introduce the general notion of *covering join tree*, which embodies the same characteristics of the join trees associated to the fusion algorithm, yet only depends on the inference problem.

DEFINITION V.34. Let $\mathcal{T} := (\mathsf{V}, \mathsf{E}, \lambda)$ be a join tree.

- $\mathcal{T}$ is said to **cover** domains $S_1, \dots, S_m$ if for all $1 \leq i \leq m$ there exists a node $j \in \mathsf{V}$ such that $S_i \subseteq \lambda(j)$.

- $\mathcal{T}$ is a **covering join tree** for the inference problem $(\phi_1 \otimes \cdots \otimes \phi_m)^{\downarrow D}$ if it covers the query[14] D and the knowledgebase domains $d(\phi_1), \ldots, d(\phi_m)$.

In the covering join tree associated to the fusion algorithm, the node numbering is determined by the elimination sequence. This fact plays a key role in establishing the sending and receiving nodes of a message. If a join tree is constructed independently from the fusion process, such a numbering has to be imposed artificially. To achieve this, one identifies a node covering the query as the *root node* and assigns the number $|\mathsf{V}|$ to it. Then, each edge is 'directed' towards the root node, and it is possible to establish an order of the nodes such that, if j is a node on the path from i to $|V|$, then $i < j$. This enumeration, in turn, determines the notions of *parent node*, *child node*, *leaf* and *separator*:

Definition V.35. Let $\mathcal{T} = (\mathsf{V}, \mathsf{E}, \lambda)$ be a covering join tree for an inference problem, and assume its nodes are ordered as described above.

- The **parents** of a node $i \in \mathsf{V}$ are defined by the set
$$\mathsf{pa}(i) := \{j \in \mathsf{V} \mid j < i \text{ and } (i, j) \in \mathsf{E}\}.$$
- Nodes without parents are called **leaves**
- The **child** $\mathsf{ch}(i)$ of a node $i < |\mathsf{V}|$ is the unique node $j \in \mathsf{V}$ such that $(i, j) \in \mathsf{E}$ and $i < j$.
- The **separator** $\mathsf{sep}(i)$ of a node $i < |\mathsf{V}|$ is the set
$$\mathsf{sep}(i) := \lambda(i) \cap \lambda(\mathsf{ch}(i)).$$

We now have all the necessary tools for the definition of the collect algorithm.

10.6.2. The collect algorithm. Broadly speaking, the collect algorithm is the most general and efficient method for the solution of single-query inference problems. It presupposes a covering join tree for the inference problem, numbered according to the procedure described above.

Given a such a tree for the problem $(\phi_1 \otimes \cdots \otimes \phi_m)^{\downarrow D}$, we determine a *covering assignment* $a : \{1, \ldots, m\} \to \mathsf{V}$ such that $d(\phi_i) \subseteq \lambda(a(i))$ for all $1 \leq i \leq m$. Then, to each node $i \in \mathsf{V}$, we associate the valuation

$$\psi_i^{(1)} := e_{\lambda(i)} \otimes \bigotimes_{j : a(j) = i} \phi_j, \tag{V.41}$$

called the **initial content** of node i. This is simply a translation of what discussed in the message-passing interpretation of the fusion algorithm: at the beginning of the procedure, each node contains the combination of all the valuations whose domain is covered by its label. The addition of $e_{\lambda(i)}$ simply ensures that $d\left(\psi_i^{(1)}\right) = \lambda(i)$.

With this premise, the collect algorithm can be simply described by the following three rules:

- Each node sends a message to its child after it has received all messages from its parents. Hence leaves send their messages right away.

[14]Note that D can contain more than one query. This definition thus applies to multi-query inference problems as well.

- After having received a message, the node updates its current content by combining it with the incoming message.
- When a node is ready to send, it computes its message by projecting its current content on its separator. Then, it sends the message to its child.

In order to keep track of the content of the nodes at each step of the algorithm, we will denote by $\psi_i^{(j)}$ the content of node i before step j. The algorithm functions as follows:

(1) For each vertex $i = 1, \dots, |\mathsf{V}|$,
   (a) Node i computes the message
$$\mu_{i \to \mathsf{ch}(i)} := \psi_i^{(i)\downarrow \mathsf{sep}(i)}.$$
   (b) This message is sent to the child node $\mathsf{ch}(i)$
   (c) The child node $\mathsf{ch}(i)$ combines the message with its current content:
$$\psi_{\mathsf{ch}(i)}^{(i+1)} = \psi_{\mathsf{ch}(i)}^{(i)} \otimes \mu_{i \to \mathsf{ch}(i)}.$$
   All other node contents remain unchanged:
$$\psi_j^{(i+1)} = \psi_j^{(i)}, \ \forall j \neq i.$$

(2) Return the content $\psi_{|\mathsf{V}|}^{(|\mathsf{V}|)}$ of the root node restricted to the query.

In [**PK12**] (Theorem 3.6), it is formally proven that $\left(\psi_{|\mathsf{V}|}^{(|\mathsf{V}|)}\right)^{\downarrow D}$ is indeed the solution of the inference problem. The collect procedure is described more formally in Algorithm 4.

Algorithm 4 General collect algorithm

Input: $\{\phi_1, \dots, \phi_m\}$, D, $\left(\mathsf{V}, \mathsf{E}, \lambda, \left\{\psi_i^{(1)}\right\}_{i \in \mathsf{V}}\right)$
Output: $(\phi_1, \dots, \phi_m)^{\downarrow D}$

procedure
  for $i = 1, \dots, |\mathsf{V}|$ **do**
    $\mu_{i \to \mathsf{ch}(i)} \leftarrow \left(\psi_i^{(i)}\right)^{\downarrow \mathsf{sep}(i)}$.
    $\psi_{\mathsf{ch}(i)}^{(i+1)} \leftarrow \psi_{\mathsf{ch}(i)}^{(i)} \otimes \mu_{i \to \mathsf{ch}(i)}$.
  end for
  return $\left(\psi_{|\mathsf{V}|}^{|\mathsf{V}|}\right)^{\downarrow D}$
end procedure

10.7. The complexity of the collect algorithm. The separation of the join tree construction procedure from the local computation process is the key to the potentially reduced complexity of the collect algorithm. Just like in the case of the fusion algorithm, the treewidth of the join tree determines the size of the largest combination to be performed, which is the dominant complexity factor $d^{\omega+1}$. However, the number of join tree nodes is no longer strictly related to m and $|T|$ as in (V.33). In Algorithm 4, each node combines all incoming messages from its parents. Because $|\mathsf{E}| = |\mathsf{V}| - 1$, this implies that a total of $|\mathsf{V}| - 1$ combinations are performed, to which we add a total of $|\mathsf{V}|-1$ projections to the separators needed to define each message, plus the final projection onto the query. As a result, we obtain a factor of $2(|\mathsf{V}|-1)+1$, which yields a total complexity of

$$\mathcal{O}\left(|\mathsf{V}|\cdot d^{\omega+1}\right). \tag{V.42}$$

Note that this complexity does not take into account the computation of the initial content of each node according to (V.41), which may significantly increase the running time of the algorithm.

Since the fusion algorithm gives rise to a covering join tree for the inference problem, the collect algorithm is at least as efficient as the fusion process. However, depending on the inference problem, it might be possible to construct a more efficient covering join tree, thus reducing the complexity of the method. For this reason, the construction of covering join trees for inference problem becomes of fundamental importance, and constitutes a broad research field by itself [**Ros70, BB72, AK91, CM94, HL99**]. It shall be noted, though, that the dominant factor $d^{\omega+1}$ in (V.42) for the worst case *cannot* be improved from (V.33). Indeed, it has been proven that for any covering join tree of a given treewidth there always exists an elimination sequence for the fusion algorithm that produces a join tree with the same treewidth [**Arn85, DP89**].

10.8. A collect algorithm for contextuality. Just like the fusion algorithm, the collect method can be applied to detect both logical and strong contextuality.

10.8.1. Strong contextuality. Let $\mathcal{S}$ be a model on a scenario $\langle X, \mathcal{M}, (O_m)\rangle$. Once again, the strategy is to use the collect algorithm to solve (V.35), and check whether the result is $z_{\overline{C}}$. Of course, one has to construct a covering join tree for (V.35) first, and compute the initial content $\psi_i^{(1)}$ of each node. This task depends on the measurement scenario, and it is thus impossible to provide any specific guidelines other than the aforementioned general heuristics on the construction of covering trees.

Once a covering tree $(\mathsf{V},\mathsf{E},\lambda)$ for (V.35) is provided and a covering assignment $a : \mathcal{M}\to\mathsf{V}$ is determined, one computes the initial content of each node using combination of indicator functions, following (V.41):

$$\psi_i^{(i)} = e_{\lambda(i)} \otimes \bigotimes_{\substack{C\in\mathcal{M}:\\ a(C)=i}} i_{\mathcal{S}(C)}.$$

The collect procedure for the detection of strong contextuality can now be presented in Algorithm 5. By (V.42), the algorithm has a complexity of

$$\mathcal{O}\left(|\mathsf{V}|\cdot |O|^{\omega+1}\right), \tag{V.43}$$

where O denotes the largest outcome set.

10.8.2. Logical contextuality. The detection of logical contextuality is slightly more complex. In this case, the problem requires solving multiple copies of (V.35), and thus takes the form of a multi-query inference problem, where the queries are contained in $\mathcal{M}$. Therefore, one can use a covering join tree for the multi-query problem

$$\left(\bigotimes_{C\in\mathcal{M}} i_{\mathcal{S}(C)}\right)^{\downarrow\mathcal{M}}$$

Algorithm 5 Collect algorithm for strong contextuality

Input: $\mathcal{S}$, $\langle X, \mathcal{M}, (O_m)\rangle$, $\left(\mathsf{V}, \mathsf{E}, \lambda, \left\{\psi_i^{(1)}\right\}_{i\in\mathsf{V}}\right)$ covering join tree for (V.35)
Output: $\mathsf{SC}(\mathcal{S})$

procedure
  for $i = 1, \ldots, |\mathsf{V}|$ **do**
    $\mu_{i\to\mathsf{ch}(i)} \leftarrow \left(\psi_i^{(i)}\right)^{\downarrow\mathsf{sep}(i)}$.
    $\psi_{\mathsf{ch}(i)}^{(i+1)} \leftarrow \psi_{\mathsf{ch}(i)}^{(i)} \otimes \mu_{i\to\mathsf{ch}(i)}$.
  end for
  if $\left(\psi_{|\mathsf{V}|}^{(|\mathsf{V}|)}\right)^{\downarrow C} = z_{\overline{C}}$ **then return** `true`
  else return `false`
  end if
end procedure

to solve each copy of the problem, with one caveat: the ordering of the tree has to be changed at each iteration to guarantee that the root note contains the current query. This is simply achieved by redirecting the edges of the tree towards the new root node, and defining a new order following the same procedure as before. With this premise, the collect algorithm for logical contextuality can be presented in Algorithm 6. By (V.42),

Algorithm 6 Collect algorithm for logical contextuality

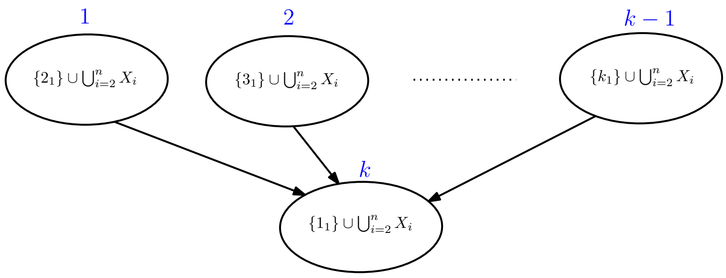
Input: $\mathcal{S}$, $\langle X, \mathcal{M}, (O_m)\rangle$, $\left(\mathsf{V}, \mathsf{E}, \lambda, \left\{\psi_i^{(1)}\right\}_{i\in\mathsf{V}}\right)$ covering join tree for (V.35)
Output: $\mathsf{LC}(\mathcal{S})$

procedure
  for $i = |\mathsf{V}|, |\mathsf{V}| - 1, \ldots, 1$ **do**
    Re-order the tree, imposing $i = |\mathsf{V}|$.
    for $j = 1, \ldots, |\mathsf{V}|$ **do**
      $\mu_{j\to\mathsf{ch}(j)} \leftarrow \left(\psi_j^{(j)}\right)^{\downarrow\mathsf{sep}(j)}$.
      $\psi_{\mathsf{ch}(j)}^{(j+1)} \leftarrow \psi_{\mathsf{ch}(j)}^{(j)} \otimes \mu_{j\to\mathsf{ch}(j)}$.
    end for
    for $C \in \mathcal{M}$ covered by j **do**
      if $\left(\psi_{|\mathsf{V}|}^{|\mathsf{V}|}\right)^{\downarrow C} \neq i_{\mathcal{S}(C)}$ **then return** `true`
      end if
    end for
    Cancel re-ordering.
  end for
  return `false`
end procedure

this procedure has a complexity of

$$O\left(|\mathcal{M}| \cdot |\mathsf{V}| \cdot |O|^{\omega+1}\right). \tag{V.44}$$

10.8.3. The complexity of the collect algorithm for non-locality in (n, k, l)-scenarios. In order to give a more concrete example for Algorithms 5 and 6 and their complexity, we consider once again the problem of detecting non-locality in (n, k, l) scenarios. Let $\Sigma = \langle X, \mathcal{M}, O\rangle$ be an (n, k, l) Bell-type scenario. Consider the following covering join tree for problem (V.35), where $\overline{C} = \{1_1, \dots, 1_n\}$. The order of the nodes is displayed in blue.



Note that this tree is optimal both in terms of its width and number of nodes. Indeed, it has width $\omega = k(n-1)$ – which we proved to be optimal in Section 10.5.4 – and only k nodes, which is clearly the minimal amount needed to cover all contexts without increasing ω.

Each node covers k^{n-1} contexts. Thus, in order to compute their initial content, one has to perform $(k^{n-1} - 1)$ combinations for each of the k nodes, which costs

$$\mathcal{O}\left(k\left(\left(k^{n-1}-1\right)l^{\omega+1}\right)\right) = \mathcal{O}\left(k^n \cdot l^{k(n-1)+1}\right). \tag{V.45}$$

We can now evaluate the complexity of Algorithms 5 and 6:

Strong contextuality. By (V.43), the collect algorithm requires $\mathcal{O}\left(k \cdot l^{k(n-1)+1}\right)$ operations. By adding these to (V.45), we obtain an overall complexity of

$$\mathcal{O}\left(k^n \cdot l^{k(n-1)+1}\right),$$

which coincides with the complexity of the fusion algorithm for strong contextuality (V.39). We conclude that, in this case, nothing is gained by using the collect algorithm over fusion.

Logical contextuality. Algorithm 6 essentially runs the collect procedure k times (in the worst case scenario), each with a different root node. At the end of every iteration, there is an additional **for** cycle (lines 8-11), which performs one projection for each of the contexts covered by the current root node. Each node covers k^{n-1} contexts, hence for each of the k runs we have the usual complexity (V.42) of performing the collect procedure *plus* an additional $k^{n-1} \cdot l^{\omega+1}$ operations, leading to a total of

$$\mathcal{O}\left(k\left(k \cdot l^{\omega+1} + k^{n-1} \cdot l^{\omega+1}\right)\right) = \mathcal{O}\left(k^n \cdot l^{k(n-1)+1}\right).$$

The addition of the complexity of computing the initial contents (V.45) does not change the result, hence this is the overall complexity of detecting logical non-locality using the collect method. Therefore, the collect algorithm for logical non-locality is strictly more efficient than the fusion algorithm, and further improves the current complexity of (V.29).

Discussion

We have presented a general definition of different forms of disagreement between information sources in the abstract framework of valuation algebras. In particular, we identified three kinds of disagreement: local, global and complete, and presented many examples of each of them using different valuation algebras. A particular attention has been given to instances of knowledgebases which agree locally but disagree globally. By recovering part of the valuation algebraic formalism in sheaf-theoretic terms, we showed that contextuality is simply a special case of such a knowledgebase, where the valuation algebra in question is the one of R-potentials, while strong contextuality is a special case of complete disagreement for the algebra of indicator functions. This result is a vast generalisation of the previously observed connections between contextuality and relational databases, constraint satisfaction problems, and logical paradoxes, and constitutes a promising attempt to establish a general theory of contextual semantics. The main advantage of such an abstract and flexible treatment is that it significantly widens the scope for the observation of contextual behaviour, and could potentially lead to the transfer of results and methods for disagreement across the many different fields captured by the valuation algebraic framework.

We have only started to explore this potential by applying methods of generic inference to develop algorithms for contextuality, and we firmly believe that much more can be done in this direction. For instance, one could hope to extend the cohomological invariant for contextuality developed in Chapter IV to a general invariant for disagreement in valuation algebras, which could then be used to study e.g. the solvability of constraint satisfaction problems such as graph colourability, or to detect inconsistencies in large databases. Another possible research path would be to generalise Vorob'ev's theorem to the level of valuation algebras, so as to recover as special cases both its original formulation concerning probability distributions, and its translation in terms of contextuality (Theorem II.21). Then, one could translate the result to other valuation algebras, and establish new interesting connections.

In the last part of the chapter, we listed two methods for the detection of logical forms of contextuality based on popular algorithms of generic inference: the fusion and collect algorithms. We showed that these techniques outperform current algorithms over (n, k, l) scenarios, especially for logical contextuality. Although the collect algorithm is arguably the most efficient known method for general inference problem, there are countless other techniques which may offer better runtime on specific scenarios. We shall investigate this possibility in future work.

CHAPTER VI

A complete characterisation of All-vs-Nothing arguments for stabiliser states

Summary

All-vs-Nothing arguments constitute an important class of proofs of contextuality in quantum mechanics. Since their first appearance in the works of Mermin, other examples have subsequently been presented, most notably in stabiliser quantum mechanics, where they are routinely used to produce instances of strongly contextual quantum states. However, no general method of identifying the states giving rise to this kind of proofs was known until now. In this chapter, we take advantage of the general formulation of AvN arguments by Abramsky et al. to give a complete characterisation of the stabiliser states giving rise to AvN arguments, which can be used to produce an exhaustive list of strongly contextual multi-qubit states. This is achieved through a combinatorial characterisation of AvN arguments, the AvN triple Theorem, whose proof makes use of the theory of graph states. This leads to other surprising structural results, such as that every AvN argument can essentially be reduced to Mermin's original proof.

1. Overview

Quantum physics provides many examples of strong contextuality. Among the first to observe this phenomenon was Mermin, who proved the GHZ state to be strongly contextual using the argument reviewed in Example II.16, which he dubbed 'all-vs-nothing' due to the sharp contrast between the 0% classical probability of a certain event to occur versus its 100% quantum probability [**Mer90a, Mer93**].

Mermin's original formulation has been recently abstracted and generalised by Abramsky and coworkers into the formal definition of *AvN arguments*, reviewed in Section 5.2 of Chapter II [**ABK$^+$15**]. The class of models admitting this kind of contextuality proofs has been subsequently shown to present suitable topological properties, both in the setting of sheaf cohomology described in this thesis [**ABK$^+$15, Aas18**], and the one proposed by Okay et al. [**ORBR17**]. For this reason, we are interested in understanding how such models arise in Nature.

One of the key results of [**ABK$^+$15**] is that this high-level description can be conveniently adapted to prove strong contextuality for a large class of states in quantum theory, notably in stabiliser quantum mechanics [**Got97**]. AvN contextual stabiliser states have recently gained considerable interest as a fundamental resource for quantum computing, particularly in the setting of measurement based quantum computation (MBQC)

[**GC99, RB01b, RB01a, KLM01, RBB03, Nie03, Leu04, RBB03, BBD**$^+$**09, Rau13**]. Contextual behaviour of this kind is also known to be necessary for increasing computational power in certain models of MBQC with restricted classical co-processing [**Rau13**]. We have already encountered an example of this in Section 2.3 of Chapter II, where we reviewed a result by Anders & Browne on how the GHZ state empowers a linear classical co-processor to implement the non-linear OR function [**AB09**]. This finding was subsequently extended by Dunjko, Kapourniotis & Kashefi [**DKK16**], who showed that strong non-locality also enables the function to be implemented in a secure delegated way.

These celebrated results motivate a thorough investigation of the stabiliser world in search for states that give rise to AvN arguments for strong contextuality. Although some general constructions to produce instances of AvN contextuality have been introduced before [**GTHB05, Wae14**], a method to identify *all* such states is yet to be presented. In this chapter, we solve this problem by providing a complete characterisation of AvN arguments for stabiliser states. This leads to a computational method to generate myriad examples of strongly contextual models in quantum mechanics, as well as the corresponding proofs of their contextuality. Let us briefly outline our results:

- We show that generalised AvN arguments on stabiliser states can be completely characterised by *AvN triples*, proving part of what was previously known as the *AvN triple conjecture* [**Abr14a**]. AvN triples are triples of elements of the Pauli group on n qubits satisfying various combinatorial properties. The presence of such a triple in a stabiliser group was proven in [**ABK**$^+$**15**] to be a sufficient condition for AvN contextuality. We show that the converse of this statement is also true for the case of *maximal* stabiliser subgroups, or, equivalently, for stabiliser states. This is achieved by adopting the language of *graph states* [**SW01b, HDE**$^+$**06**], which is particularly suited to model stabiliser states. This new description of AvN arguments leads to a surprising structural result concerning the way such arguments arise:
- We prove that any AvN argument on an n-qubit stabiliser state can be reduced to an AvN argument only involving three qubits which are local Clifford-equivalent to the tripartite GHZ state. This conclusion seems to suggest that the GHZ state is somewhat 'universal' among stabiliser states admitting AvN proofs of contextuality.
- We present a computational method to generate all AvN arguments for n-qubit stabiliser states, based on the characterisation introduced above. Notice that we do not limit ourselves to identifying the AvN contextual states, but we provide an exhaustive list of actual AvN proofs of contextuality for each of them.

The content of this chapter has been developed in collaboration with Samson Abramsky, Rui Soares Barbosa and Simon Perdrix, and has been published in [**ABCP17**].

Outline of the chapter. In Section 2, we review Mermin's original All-vs-Nothing argument. Section 3 introduces the stabiliser formalism and shows how stabiliser subgroups induce XOR theories, which can be used to derive AvN arguments for the corresponding stabiliser states. The characterisation of AvN arguments is introduced in

Section 4, where the AvN triple theorem is proved using graph states. Finally, in Section 5, we present the applications of the theorem and illustrate the computational method to generate AvN triples.

2. Mermin's proof

Before embarking in a detailed analysis of All-vs-Nothing arguments in stabiliser quantum mechanics, it is convenient to review Mermin's original proof of the strong contextuality of the GHZ state in this specific setting.

We start by definiting the **Pauli operators**, dichotomic one-qubit local observables corresponding to measuring spin in the x, y, and z axes, with eigenvalues ± 1:

$$X \coloneqq \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ Y \coloneqq \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ Z \coloneqq \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \in \mathbb{M}_2(\mathbb{C})$$

These matrices are self-adjoint, have eigenvalues ± 1, and satisfy the following relations: for all $P, Q \in \{X, Y, Z\}$,

$$\begin{aligned} P^2 &= I, \\ PQ &= -QP. \end{aligned} \tag{VI.1}$$

Consider a tripartite experimental setting where each party $i = 1, 2, 3$ can choose to perform a Pauli measurement in $\{X_i, Y_i\}$ on the i-th qubit of the GHZ state[1]

$$|\mathsf{GHZ}\rangle \coloneqq \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle).$$

Once measurement $P_i \in \{X_i, Y_i\}$ is executed, experimenter i obtains the eigenvalue $\bar{P}_i \in \{\pm 1\}$ as outcome. By direct calculation we can show that the following holds:

$$X_1 \otimes X_2 \otimes X_3 \, |\mathsf{GHZ}\rangle = |\mathsf{GHZ}\rangle \, .$$

In other words, one could say that the operator $X_1 \otimes X_2 \otimes X_3$ *stabilises* the GHZ state. This implies that the expected value of measuring X at each site is

$$\langle \mathsf{GHZ}| \ X_1 \otimes X_2 \otimes X_3 \ |\mathsf{GHZ}\rangle = 1.$$

Consequently, since the eigenvalues of any joint measurement are the products of eigenvalues at each site, we conclude that

$$\bar{X}_1 \cdot \bar{X}_2 \cdot \bar{X}_3 = 1,$$

where $\bar{X}_i = \pm 1$ denotes the eigenvalue produced after measurement X_i is performed. Using the isomorphism $\{+1, -1, \times\} \cong \{0, 1, \oplus\}$, we rewrite this equation in the following form:

$$\bar{X}_1 \oplus \bar{X}_2 \oplus \bar{X}_3 = 0.$$

This means that the possible joint events of the corresponding context of the underlying empirical model are the ones with an *even* number of 1's.

A similar argument can be applied to the joint measurement $X_1 \otimes Y_2 \otimes Y_3$. In this case the GHZ state is *anti-stabilised*:

$$X_1 \otimes Y_2 \otimes Y_3 \, |\mathsf{GHZ}\rangle = (-1) \cdot |\mathsf{GHZ}\rangle \, .$$

[1]The subscript i for a Pauli measurement P_i has the only purpose of distinguishing the same measurement performed by different parties

It follows that the outcomes obtained by each party when measuring $X_1 \otimes Y_2 \otimes Y_3$ must satisfy

$$\bar{X}_1 \oplus \bar{Y}_2 \oplus \bar{Y}_3 = 1.$$

By repeating the same discussion for $Y_1 \otimes X_2 \otimes Y_3$ and $Y_1 \otimes Y_2 \otimes X_3$ we obtain the following system of equations:

$$\begin{aligned} \bar{X}_1 \oplus \bar{X}_2 \oplus \bar{X}_3 &= 0 & \qquad \bar{Y}_1 \oplus \bar{X}_2 \oplus \bar{Y}_3 &= 1 \\ \bar{X}_1 \oplus \bar{Y}_2 \oplus \bar{Y}_3 &= 1 & \qquad \bar{Y}_1 \oplus \bar{Y}_2 \oplus \bar{X}_3 &= 1, \end{aligned}$$

This system characterises the entries of the corresponding 4 contexts of the possibilistic GHZ model, displayed in Table VI.1 (cf. Table II.6).

1	2	3	000	001	010	011	100	101	110	111
…	…	…	…	…	…	…	…	…	…	…
X_1	X_2	X_3	1	0	0	1	0	1	1	0
X_1	Y_2	Y_3	0	1	1	0	1	0	0	1
Y_1	X_2	Y_3	0	1	1	0	1	0	0	1
Y_1	Y_2	X_3	0	1	1	0	1	0	0	1
…	…	…	…	…	…	…	…	…	…	…

TABLE VI.1. A partial table for the support of the GHZ model.

It is straightforward to see that this system is inconsistent. Indeed, if we sum all the equations, we obtain $0 = 1$, as each variable appears twice on the left-hand side. This means that it is impossible to find a global assignment $\{X_1, Y_1, X_2, Y_2, X_3, Y_3\} \to \{0, 1\}$ consistent with the model, showing that the GHZ state is strongly contextual.

3. The stabiliser world

We shall now show how the general definition of AvN arguments presented in Section VI of Chapter II naturally specialises to the setting of stabiliser quantum mechanics [**Got97, NC00, Cav14**]. Let us start by introducing the *Pauli n-group*, which embodies the algebraic properties of multi-partite local measurements:

DEFINITION VI.1. Let $n \geq 1$ be an integer. The **Pauli n-group** $\mathscr{P}_n$ is comprised of elements of the form

$$\alpha(P_1, \ldots, P_n),$$

where $(P_i)_{i=1}^n$ is an n-tuple of Pauli operators, $P_i \in \{X, Y, Z, I\}$, with global phase $\alpha \in \{\pm 1, \pm i\}$. Elements of $\mathscr{P}_n$ will be denoted by bold capital letters $\mathbf{P} = \alpha(P_i)_{i=1}^n$. Multiplication is defined by

$$\alpha(P_1, \ldots, P_n)\beta(Q_1, \ldots, Q_n) = \gamma(R_1, \ldots, R_n),$$

where $P_i Q_i = \gamma_i R_i$, $\gamma = \alpha\beta(\prod_i \gamma_i)$. The unit is $\mathbf{I} := (I_i)_{i=1}^n$.

The group $\mathscr{P}_n$ acts on the Hilbert space of n-qubits $\mathbb{H}_n := (\mathbb{C}^2)^{\otimes n}$ via the action

$$\alpha(P_i)_{i=1}^n \bullet |\psi\rangle := \alpha P_1 \otimes \cdots \otimes P_n |\psi\rangle . \tag{VI.2}$$

Elements of the Pauli n-group shall be interpreted as n-partite local measurements. In other words, a Pauli $P \in \mathscr{P}_n$ determines, up to phase, a context of an n-partite measurement scenario where each party can perform X, Y, Z or I on their respective qubit of shared a state.

In this chapter, we shall be concerned with scenarios whose set of contexts is determined by a subgroup of $\mathscr{P}_n$. More specifically, we are interested in studying the empirical model arising from the application of joint measurements in a subgroup $S \leq \mathscr{P}_n$ to an n-qubit state which is *stabilised* under (VI.2) by all elements of S. This leads to the definition of *stabiliser subspaces* and *subgroups*:

DEFINITION VI.2. Given a subgroup $S \leq \mathscr{P}_n$, the **stabiliser** of S is the linear subspace

$$V_S := \{|\psi\rangle \in \mathbb{H}_n \mid \mathbf{P} \bullet |\psi\rangle = |\psi\rangle\,,\ \forall \mathbf{P} \in S\} \subseteq \mathbb{H}_n.$$

Similarly, given a subspace $V \subseteq \mathbb{H}_n$, the **stabiliser** of V is the subgroup

$$S_V := \left\{\mathbf{P} \in \mathscr{P}_n \mid \mathbf{P} \bullet |\psi\rangle = |\psi\rangle\,,\ \forall\,|\psi\rangle \in V\right\}.$$

A subgroup $S \leq \mathscr{P}_n$ that stabilises a non trivial subspace of $\mathbb{H}_n$ is called a **stabiliser subgroup**. That is to say, $S \leq \mathscr{P}_n$ is a stabiliser subgroup if and only if V_S is non-trivial.

This definition establishes a Galois connection

$$\begin{tikzcd} \mathsf{SG}(\mathscr{P}_n) \arrow[r, bend left] & \mathsf{SS}(\mathbb{H}_n), \arrow[l, bend left] \end{tikzcd}$$

between subgroups of $\mathscr{P}_n$ and subspaces of $\mathbb{H}_n$, both ordered by inclsion, which underpins the following well-known relation between the rank of a subgroup $S \leq \mathscr{P}_n$ and the dimension of V_S [**NC00**]:

$$\text{(VI.3)} \qquad \mathsf{rank}(S) = k \quad \Leftrightarrow \quad \dim(V_S) = 2^{n-k}.$$

In particular, when S is a maximal stabiliser subgroup, i.e. $k = n$, we have $\dim V_S = 1$. Hence, S stabilises a unique state, up to global phase.

DEFINITION VI.3. We call a state stabilised by a maximal subgroup $S \leq \mathscr{P}_n$ a **stabiliser state**.

Let us conclude this section by presenting the following elementary result of stabiliser theory, which will be used throughout:

PROPOSITION VI.4. *Let $S \leq \mathscr{P}_n$ be a stabiliser subgroup.*

(1) The elements of S have global phase ± 1.
(2) S is abelian.

PROOF. Let us start by observing that $-\mathbf{I} \notin S$. Indeed, $-\mathbf{I} \in S$ implies $|\psi\rangle = -|\psi\rangle$ for all $|\psi\rangle \in V_S$, which contradicts the non-triviality of V_S. This, in turn, immediately leads to a proof of (1): let $\mathbf{P} = \alpha(P_i)_{i=1}^n \in S$, if $\alpha = \pm i$, we have $\mathbf{P}^2 = \alpha^2(I_i)_{i=1}^n = -(I_i)_{i=1}^n = -\mathbf{I}$. In order to prove (2) note that, by (VI.1) and (1), two elements $\mathbf{P}, \mathbf{Q} \in S$ either commute or anti-commute. Suppose by contradiction that they anticommute, and let $|\psi\rangle \in V_S$. We have

$$|\psi\rangle = \mathbf{Q}\,|\psi\rangle = \mathbf{QP}\,|\psi\rangle = -\mathbf{PQ}\,|\psi\rangle = -\,|\psi\rangle\,.$$

Thus $|\psi\rangle = 0$, which contradicts the non-triviality of V_S. $\square$

3.1. Stabiliser subgroups induce XOR theories. Consider a quantum measurement scenario where n parties share an n-qubit state and can each choose to perform a local Pauli operator X, Y, or Z on their respective qubit. The set of measurements is thus $\mathscr{X} = \bigsqcup_{i=1}^{n}\{X_i, Y_i, Z_i\}$, and the contexts are subsets of $\mathscr{X}$ that contain at most one element for each index i, establishing an $(n, 3, 2)$ Bell-type scenario.[2] In analogy with Mermin's proof, we will now show how elements of the Pauli group that stabilise a state generate parity conditions for the possible joint outcomes of the corresponding empirical model.

Let $\mathbf{P} = \alpha(P_i)_{i=1}^{n} \in \mathscr{P}_n$ be an element of the Pauli n-group and $|\psi\rangle \in \mathbb{H}_n$ a state stabilised by $\mathbf{P}$. By Proposition VI.4, we have $\alpha = \pm 1$. Then,

$$\alpha(P_1 \otimes \cdots \otimes P_n)\,|\psi\rangle = |\psi\rangle$$

and so $|\psi\rangle$ is an α-eigenvector of $P_1 \otimes \cdots \otimes P_n$. Consequently, the expected value satisfies

$$\langle\psi|P_1 \otimes \cdots \otimes P_n|\psi\rangle = \alpha.$$

This means that, given n qubits prepared on the state $|\psi\rangle$, any joint outcome $s \in \mathbb{Z}_2^n$ resulting from measuring P_i at each qubit i must have even (resp. odd) parity when $\alpha = 1$ (resp. $\alpha = -1$). Therefore, to any stabiliser group $S \leq \mathscr{P}_n$ we associate an XOR theory

$$\mathbb{T}_{\mathbb{Z}_2}(S) \coloneqq \{\varphi_{\mathbf{P}} \mid \mathbf{P} \in S\},$$

where, if $\mathbf{P} = (-1)^a (P_i)_{i=1}^{n}$ with $a \in \mathbb{Z}_2$,

$$\varphi_{\mathbf{P}} \coloneqq \left(\bigoplus_{\substack{i \in \{1,\dots,n\} \\ P_i \neq I}} \bar{P}_i \;=\; a \right).$$

Given a stabilser subgroup $S \leq \mathscr{P}_n$, we are interested in the empirical model arising from applying measurements in S to a state in V_S. In this case, the possible local sections at each contexts $\mathbf{P} \in S$ are completely determined by solutions to $\varphi_{\mathbf{P}}$, hence the model is independent of the chosen state of V_S.

The definition of the theory $\mathbb{T}_{\mathbb{Z}_2}(S)$ allows us to translate the general notion of an AvN empirical model (Definition II.18) to the specific case of stabiliser subgroups:

Definition VI.5. We say that S is **AvN** if $\mathbb{T}_{\mathbb{Z}_2}(S)$ is inconsistent.

These scenarios give rise to the same empirical model, since the possible events at a context $\mathbf{P}$ are completely characterised as solutions to $\varphi_{\mathbf{P}}$. The following proposition is a direct consequence of Proposition II.19:

Proposition VI.6. *Given an AvN subgroup $S \leq \mathscr{P}_n$ and any state $|\psi\rangle \in V_S$, the n-partite empirical model realised by $|\psi\rangle$ under the Pauli measurements in S is strongly contextual.*

Indeed, the inconsistency of $\mathbb{T}_{\mathbb{Z}_2}(S)$ implies the impossibility of finding a global assignment compatible with the support of the empirical model.

[2]In this chapter, we will temporarily denote the set of measurements by $\mathscr{X}$ in order to avoid confusion with the Pauli measurement X.

4. Characterising AvN arguments

Using the general theory of AvN arguments for stabiliser states reviewed in the last section, we present a characterisation of AvN arguments based on the combinatorial concept of an AvN triple [**ABK+15**].

4.1. AvN triples. Since AvN subgroups give rise to strongly contextual empirical models, we are naturally interested in characterising this property. In [**ABK+15**], this problem is addressed by introducing the notion of AvN triple. Here, we rephrase the definition in slightly more general terms:

Definition VI.7 (cf. [**ABK+15**, Definition 3]). An **AvN triple** in $\mathscr{P}_n$ is a triple $\langle \mathbf{E}, \mathbf{F}, \mathbf{G} \rangle$ of elements of $\mathscr{P}_n$ with global phases ± 1 that pairwise commute and that satisfy the following conditions:

(1) For each $i = 1, \ldots, n$, at least two of E_i, F_i, G_i are equal.
(2) The number of i such that $E_i = G_i \neq F_i$, all distinct from I, is odd.

Note that the only difference with respect to the original definition from [**ABK+15**] is that we allow elements of an AvN triple to have global phase -1.

A key result from [**ABK+15**] is that AvN triples provide a sufficient condition for All-vs-Nothing proofs of strong contextuality. A similar argument shows that this is still true for the slightly more general notion of AvN triple.

Theorem VI.8 (cf. [**ABK+15**, Theorem 4]). *Any subgroup S of $\mathscr{P}_n$ containing an AvN triple is AvN.*

Proof. Let $\langle \mathbf{E}, \mathbf{F}, \mathbf{G} \rangle$ be an AvN triple in $\mathscr{P}_n$, where $\mathbf{E} = (-1)^a (E_i)_{i=1}^n$, $\mathbf{F} = (-1)^b (F_i)_{i=1}^n$ and $\mathbf{G} = (-1)^c (G_i)_{i=1}^n$, with $a, b, c \in \mathbb{Z}_2$. We denote by N_F the number of i's such that $E_i = G_i \neq F_i$ and $E_i, F_i, G_i \neq I$, which is odd by (2). The global phase of $\mathbf{H} \coloneqq \mathbf{EFG}$ is given by

$$(-1)^{a+b+c+\mathsf{N}_F} = (-1)^{a \oplus b \oplus c \oplus 1}.$$

Indeed, we get a local phase -1 for each of the N_F indices where $E_i = G_i \neq F_i$, since $E_i F_i G_i = -E_i G_i F_i = -F_i$. Thus, if we consider the four equations corresponding to these elements in the XOR theory of the subgroup, summing their right-hand sides yields $a \oplus b \oplus c \oplus (a \oplus b \oplus c \oplus 1) = 1$. On the other hand, by 1, we have $\{E_i, F_i, G_i\} = \{P_i, Q_i\}$ with at least two elements equal to P_i. Thus, by (VI.1), the product $E_i F_i G_i$ will be Q_i up to a phase, and so, as this is absorbed into the global phase, $H_i = Q_i$. This means that each column of the four equations contains 2 P_i's and 2 Q_i's. Therefore, summing all the four equations we obtain $0 = 1$. $\square$

It was conjectured in [**Abr14a**] that the presence of an AvN triple in a stabiliser subgroup is also necessary for the existence of an All-vs-Nothing proof of strong contextuality. This is the *AvN triple conjecture*. The intuition is that short AvN proofs suffice, and it is based on the observation that any AvN argument that has appeared in the literature can be seen to come down to exhibit an AvN triple. We will now prove the conjecture for the case of maximal stabiliser subgroups or, equivalently, for stabiliser states by taking advantage of the graph state formalism, which is briefly reviewed in the following subsection.

4.2. Graph states. Graph states are special types of multi-qubit states that can be represented by a graph.

DEFINITION VI.9. Let $G = (V, E)$ be an undirected graph. For each $u \in V$, consider the element $\mathbf{G}^u = (G_v^{(u)})_{v \in V} \in \mathscr{P}_{|V|}$ with global phase $+1$ and components

$$G_v^{(u)} = \begin{cases} X & \text{if } v = u \\ Z & \text{if } v \in \mathcal{N}(u) \\ I & \text{otherwise,} \end{cases} \tag{VI.4}$$

where $\mathcal{N}(u)$ denotes the neighborhood of u, i.e. the set of all the vertices adjacent to u. The **graph state** $|G\rangle$ associated to G is the unique[3] state stabilised by the subgroup generated by these elements,

$$S_G = \langle \{\mathbf{G}^u \mid u \in V\} \rangle \,. \tag{VI.5}$$

One of the key properties of graph states is their generality with respect to stabiliser states, as stated in the following theorem due to Schlingemann [**Sch02**]. Recall the definition of the local Clifford (LC) group on n qubits

$$\mathsf{C}_1^n := \{U \in \mathbb{U}(2^n) \mid U \mathscr{P}_n U^\dagger = \mathscr{P}_n\},$$

where U acts by conjugation on elements of $\mathscr{P}_n$ via the representation of $\alpha(P_i)_{i=1}^n$ as the operator $\alpha P_1 \otimes \cdots \otimes P_n$. Two states $|\psi\rangle, |\psi'\rangle \in \mathbb{H}_n$ are said to be LC-equivalent whenever there is a $U \in \mathsf{C}_1^n$ such that $|\psi'\rangle = U|\psi\rangle$.

THEOREM VI.10 ([**Sch02, GKR02**]). *Any stabiliser state $|S\rangle$ is LC-equivalent to some graph state $|G\rangle$, i.e. $|S\rangle = U|G\rangle$ for some LC unitary $U \in \mathsf{C}_1^{|V|}$.*

An instance of this result will be particularly important for our discussion. Consider the n-partite GHZ state

$$|\mathsf{GHZ}(n)\rangle = \frac{1}{\sqrt{2}}(|0\rangle^{\otimes n} + |1\rangle^{\otimes n}).$$

We apply a local Clifford transformation consisting of a Hadamard unitary

$$H := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

at every qubit of GHZ except the k-th one (where $1 \leq k \leq n$ can be chosen arbitrarily) to obtain

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left(|+ \cdots + 0_k + \cdots +\rangle \quad + \quad |- \cdots - 1_k - \cdots -\rangle \right).$$

The stabiliser group of $|\psi\rangle$ is generated by the elements $\mathbf{E}, \mathbf{F}^1, \mathbf{F}^2, \ldots, \mathbf{F}^{k-1}, \mathbf{F}^{k+1}, \ldots, \mathbf{F}^n$, where

$$E_i = \begin{cases} X & \text{if } i = k \\ Z & \text{if } i \neq k \end{cases} \quad \text{and} \quad F_i^j = \begin{cases} X & \text{if } i = j \\ Z & \text{if } i = k \\ I & \text{otherwise} \end{cases}$$

[3]Uniqueness follows from (VI.3) since there is an independent generator for each vertex.

By definition of a graph state, this list of stabilisers coincides with the *star graph* centred at the vertex corresponding to the k-th qubit (Figure VI.1). Since k was chosen arbitrarily, all the graph states corresponding to star graphs on n vertices and different centres are LC-equivalent.

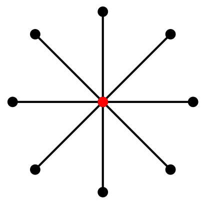


FIGURE VI.1. Example of a star graph $G = (V, E)$ with $|V| = 9$. The graph state $|G\rangle$ is LU-equivalent to the 9-partite GHZ state.

Another important property of graph states is that they allow to characterise LC-equivalence between them by a simple operation on the underlying graphs. This is the notion of local complementation, introduced by Bouchet in [**Bou93**]

DEFINITION VI.11. Given a graph $G = (V, E)$ and a vertex $v \in V$, the **local complement of G at v**, denoted by $G \star v$, is obtained by complementing the subgraph of G induced by the neighborhood $\mathcal{N}(v)$ of v and leaving the rest of the graph unchanged.

The following theorem is due to Van den Nest, Dehaene, and De Moor [**VdNDDM04**] (see also [**HDE**$^+$**06**]).

THEOREM VI.12 ([**VdNDDM04**, Theorem 3]). *By local complementation of a graph $G = (V, E)$ at some vertex $v \in V$ one obtains an LC-equivalent graph state. Moreover, two graph states $|G\rangle$ and $|G'\rangle$ are LC-equivalent if and only if the corresponding graphs are related by a sequence of local complementations, i.e.*

$$G' = G \star v_1 \cdots \star v_n$$

for some $v_1, \ldots v_n \in V$.

Thanks to this theorem, we can show that the n-partite GHZ state is LC-equivalent both to the state corresponding to the star graph (cf. Figure VI.1), and the one corresponding to the complete graph on n vertices. Indeed, it is sufficient to choose a vertex v of the complete graph and apply a local complementation to it to obtain a star graph centred at v, as illustrated in Figure VI.2.

4.3. The AvN triple theorem and its consequences. In this section, we prove the theorem characterising AvN arguments on stabiliser states.

Firstly, we need to make some observations. Note that the Born rule is invariant under any unitary action acting simultaneously on the measurement by conjugation and on the state. Therefore, if we have a quantum realisable empirical model specified by a state $|\psi\rangle$ and a set of measurements $\mathcal{X}$, then given any unitary U, the empirical model specified by the state $U|\psi\rangle$ and the set of measurements $U\mathcal{X}U^\dagger = \{UAU^\dagger \mid A \in \mathcal{X}\}$ is equivalent to the original one, in the sense that it assigns the same probabilities, which of course implies that it has the same contextuality properties.

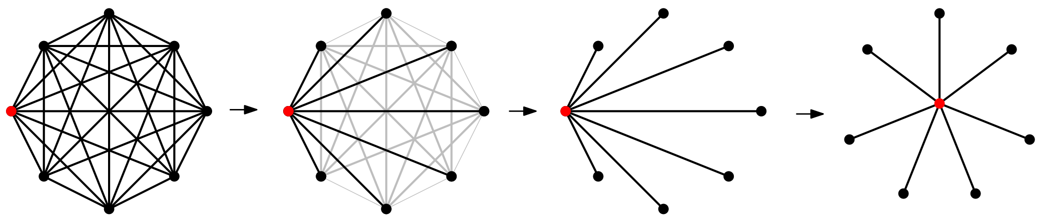


FIGURE VI.2. The star graph centred at a vertex v, marked in red, can be obtained from the complete graph by local complementation at the vertex v.

In the particular case when $|\psi\rangle$ is a stabiliser state for the subgroup $S \leq \mathscr{P}_n$, and U is a LC-operation, then the state $U|\psi\rangle$ is a stabiliser state for the subgroup $USU^\dagger = \{U\mathbf{P}U^\dagger \mid \mathbf{P} \in S\}$.

An important fact we shall need is that AvN triples are sent to AvN triples by such LC operations. The reason is that LC operations are composed of local unitaries that act as permutations on the set $\{\pm X, \pm Y, \pm Z\}$, and therefore preserve all the conditions of Definition VI.7.

We now show that AvN triples fully characterise All-vs-Nothing arguments for stabiliser states, and that a tripartite GHZ state is always responsible for the existence of such an AvN proof of strong contextuality.

THEOREM VI.13 (AvN Triple Theorem). *A maximal stabiliser subgroup S of $\mathscr{P}_n$ is AvN if and only if it contains an AvN triple. The AvN argument can be reduced to one concerning only three qubits. The state induced by the subgraph for these three qubits is LC-equivalent to a tripartite GHZ state.*

PROOF. Sufficiency follows from Proposition VI.8. So, suppose that the maximal stabiliser subgroup S is AvN. Let $|\psi\rangle$ be the stabiliser state corresponding to S. Since any stabiliser state is LC-equivalent to a graph state by Theorem VI.10, and since *LC* transformations preserve AvN triples, we can suppose without loss of generality that $|\psi\rangle$ is a graph state $|G\rangle$ induced by a graph $G = (V, E)$, and consequently that $S = S_G$ as in (VI.5).

Given that the empirical model obtained from the state $|G\rangle$ and local Pauli operators is strongly contextual, there must exist at least one vertex u with degree at least 2, i.e. $|\mathcal{N}(u)| \geq 2$. Indeed, if G has no such vertex, G is a union of disconnected edges and vertices, which implies that $|G\rangle$ is a tensor product of 1-qubit and 2-qubit states, which do not present strongly contextual behaviour for any choice of local measurements [**BMT05, ABC$^+$17**].

Let $u \in V$ have degree ≥ 2 and let v, w be two distinct vertices in $\mathcal{N}(u)$. We have two possible cases:

(1) There is an edge between v and w. Then, in accordance with (VI.4), the elements $\mathbf{G}^u, \mathbf{G}^v, \mathbf{G}^w$ of S_G have the form:[4]

$$\begin{array}{rcccl} \mathbf{G}^u: & X_u & Z_v & Z_w & [I \text{ or } Z \text{ on all other qubits}] \\ \mathbf{G}^v: & Z_u & X_v & Z_w & [I \text{ or } Z \text{ on all other qubits}] \\ \mathbf{G}^w: & Z_u & Z_v & X_w & [I \text{ or } Z \text{ on all other qubits}], \end{array} \tag{VI.6}$$

which are easily seen to constitute an AvN triple.

(2) There is no edge between v and w. Then, we have

$$\begin{array}{rcccl} \mathbf{G}^u: & X_u & Z_v & Z_w & [I \text{ or } Z \text{ on all other qubits}] \\ \mathbf{G}^v: & Z_u & X_v & I_w & [I \text{ or } Z \text{ on all other qubits}] \\ \mathbf{G}^w: & Z_u & I_v & X_w & [I \text{ or } Z \text{ on all other qubits}] \end{array}$$

and the elements $\langle \mathbf{G}^u, \mathbf{G}^u\mathbf{G}^v, \mathbf{G}^u\mathbf{G}^w \rangle$ form an AvN triple:

$$\begin{array}{rcccl} \mathbf{G}^u: & X_u & Z_v & Z_w & [I \text{ or } Z \text{ on all other qubits}] \\ \mathbf{G}^u\mathbf{G}^v: & Y_u & Y_v & Z_w & [I \text{ or } Z \text{ on all other qubits}] \\ \mathbf{G}^u\mathbf{G}^w: & Y_u & Z_v & Y_w & [I \text{ or } Z \text{ on all other qubits}]. \end{array}$$

Notice that in both cases the AvN argument is reduced to just three qubits. Moreover, by the discussion at the end of the previous subsection, we know that the state corresponding to the subgraph induced by u, v, w in either of these two cases is LC-equivalent to a tripartite GHZ state: in Case 1 we have a complete graph on three vertices, while in Case 2 we have a star graph centred at u. □

The second part of the above result means that the essence of the contradiction is witnessed by looking at only three qubits. In fact, in the contexts being considered, the experimenters at the remaining $n-3$ parties either perform no measurement or a Z measurement. We could imagine that, in trying to build a consistent global assignment of outcomes in $\mathbb{Z}_2$ to all the measurements, each of these $n-3$ parties i is allowed to freely choose a value 0 or 1 for the variable $\bar{Z}_i$. Then, the equations for the variables representing the measurements of the remaining three parties would be those of the usual GHZ argument, up to flipping an even number of the values on the right-hand side. In terms of the state, we can use the *partial inner product* operation described e.g. in [**SW10**][5] to apply the eigenvectors corresponding to the chosen values for the other $n-3$ parties to $|G\rangle$, resulting in a three-qubit pure state which is LC-equivalent to the GHZ state.

From this theorem, we immediately obtain the following corollaries:

Corollary VI.14. *A graph state $|G\rangle$ is strongly contextual if and only if G has a vertex of degree at least* 2.

Corollary VI.15. *Every strongly contextual 3-qubit stabiliser state is LC-equivalent to the GHZ state.*

The graph theoretic arguments used in the proof of the AvN triple theorem present some similarities to the techniques developed in [**GTHB05**] to derive Bell inequalities

[4]The notation in (VI.6) indicates that $G_u^{(u)} = X$, $G_v^{(u)} = G_w^{(u)} = Z$, and $G_z^{(u)}$ is either Z or I for every other vertex $z \in V \setminus \{u, v, w\}$, and analogously for the other lines.

[5]This is actually the application of a linear map to a vector under Map-State duality [**AC08**].

that are maximally violated by some graph states. In fact, Corollary VI.14 also follows from Theorem 1 in [**GTHB05**]. However, it is important to note that the AvN triple theorem goes beyond the graph state formalism to expose the underlying algebraic structure of AvN arguments, which ultimately rests on AvN triples.

Let us present some examples to clarify the statement of Theorem VI.13.

Example VI.16. Cluster states are a fundamental resource in measurement-based quantum computation [**AB09, RB01b, Rau03, RBB02**]. The 4-qubit 2-dimensional cluster state is described by the graph in Figure VI.3.

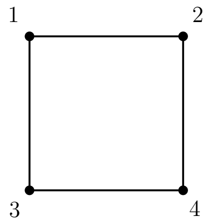


Figure VI.3. The 4-qubit 2-dimensional cluster-state.

Its stabiliser group S is generated by the following elements of $\mathscr{P}_4$:

$$\begin{array}{rcccc}
\mathbf{G}^1: & X_1 & Z_2 & Z_3 & I_4 \\
\mathbf{G}^2: & Z_1 & X_2 & I_3 & Z_4 \\
\mathbf{G}^3: & Z_1 & I_2 & X_3 & Z_4 \\
\mathbf{G}^4: & I_1 & Z_2 & Z_3 & X_4
\end{array}$$

The stabiliser group S contains the following 4 AvN triples, corresponding to the triples of qubits highlighted in Figure VI.4:

$$\begin{array}{rcccc c rcccc}
\mathbf{G}^1: & X_1 & Z_2 & Z_3 & I_4 & \quad & \mathbf{G}^2: & Z_1 & X_2 & I_3 & Z_4 \\
\mathbf{G}^1\mathbf{G}^2: & Y_1 & Y_2 & Z_3 & Z_4 & & \mathbf{G}^2\mathbf{G}^1: & Y_1 & Y_2 & Z_3 & Z_4 \\
\mathbf{G}^1\mathbf{G}^3: & Y_1 & Z_2 & Y_3 & Z_4 & & \mathbf{G}^2\mathbf{G}^4: & Z_1 & Y_2 & Z_3 & Y_4 \\
\\
\mathbf{G}^3: & Z_1 & I_2 & X_3 & Z_4 & & \mathbf{G}^4: & I_1 & Z_2 & Z_3 & X_4 \\
\mathbf{G}^3\mathbf{G}^1: & Y_1 & Z_2 & Y_3 & Z_4 & & \mathbf{G}^4\mathbf{G}^2: & Z_1 & Y_2 & Z_3 & Y_4 \\
\mathbf{G}^3\mathbf{G}^4: & Z_1 & Z_2 & Y_3 & Y_4 & & \mathbf{G}^4\mathbf{G}^3: & Z_1 & Z_2 & Y_3 & Y_4
\end{array} \tag{VI.7}$$

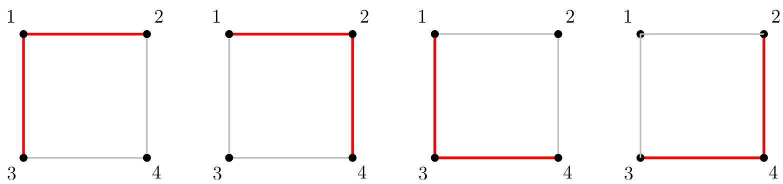


Figure VI.4. Qubits generating the AvN triples of (VI.7). Each triple of qubits is LC-equivalent to GHZ.

Example VI.17. Consider the graph state $|G\rangle$ represented in Figure VI.5.

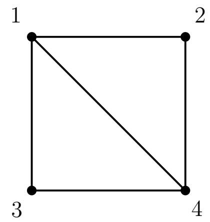


FIGURE VI.5. A general graph G.

Its stabiliser S is generated by the following elements of $\mathscr{P}_4$:

$$\begin{array}{rcccc} \mathbf{G}^1: & X_1 & Z_2 & Z_3 & Z_4 \\ \mathbf{G}^2: & Z_1 & X_2 & I_3 & Z_4 \\ \mathbf{G}^3: & Z_1 & I_2 & X_3 & Z_4 \\ \mathbf{G}^4: & Z_1 & Z_2 & Z_3 & X_4 \end{array}$$

and contains the following AvN triples:

$$\begin{array}{rcccc} \mathbf{G}^1: & X_1 & Z_2 & Z_3 & Z_4 \\ \mathbf{G}^3: & Z_1 & I_2 & X_3 & Z_4 \\ \mathbf{G}^4: & Z_1 & Z_2 & Z_3 & X_4 \end{array} \qquad \begin{array}{rcccc} \mathbf{G}^1: & X_1 & Z_2 & Z_3 & Z_4 \\ \mathbf{G}^2: & Z_1 & X_2 & I_3 & Z_4 \\ \mathbf{G}^4: & Z_1 & Z_2 & Z_3 & X_4 \end{array}$$
$$\begin{array}{rcccc} \mathbf{G}^1: & X_1 & Z_2 & Z_3 & Z_4 \\ \mathbf{G}^1\mathbf{G}^2: & Y_1 & Y_2 & Z_3 & I_4 \\ \mathbf{G}^1\mathbf{G}^3: & Y_1 & Z_2 & Y_3 & I_4 \end{array} \qquad \begin{array}{rcccc} \mathbf{G}^4: & Z_1 & Z_2 & Z_3 & X_4 \\ \mathbf{G}^4\mathbf{G}^2: & I_1 & Y_2 & Z_3 & Y_4 \\ \mathbf{G}^4\mathbf{G}^3: & I_1 & Z_2 & Y_3 & Y_4 \end{array} \tag{VI.8}$$

which correspond to the triples of qubits illustrated in Figure VI.6.

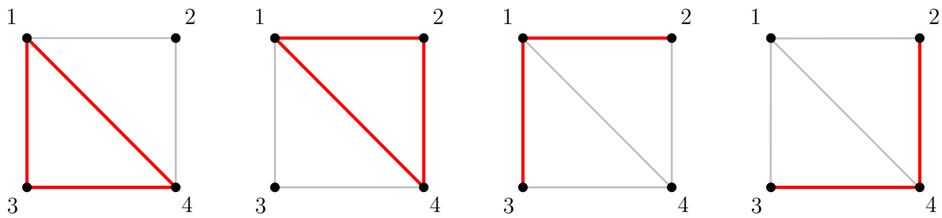


FIGURE VI.6. Qubits generating the AvN triples of (VI.8). Each triple of qubits is LC-equivalent to GHZ.

5. Applications

In this section we take advantage of the characterisation introduced above to develop a computational method to identify all the possible AvN arguments.

5.1. Counting AvN triples. We start by introducing an alternative definition of AvN triple.

DEFINITION VI.18 (Alternative Definition of AvN triple). An *AvN triple* in the Pauli n-group $\mathscr{P}_n$ is a triple $\langle \mathbf{E}, \mathbf{F}, \mathbf{G} \rangle$ with global phases ± 1, such that

1. For each $i = 1, \ldots, n$, at least two of E_i, F_i, G_i are equal.
2. The number N_G of i's such that $E_i = F_i \neq G_i$, all distinct from I, is odd.

(3) The number N_E of i's such that $E_i \neq F_i = G_i$, all distinct from I, is odd.
(4) The number N_F of i's such that $E_i = G_i \neq F_i$, all distinct from I, is odd.

The equivalence of the two definitions follows directly from the following lemma.

Lemma VI.19. *Let $n \geq 3$. Suppose $\mathbf{E}, \mathbf{F}, \mathbf{G} \in \mathscr{P}_n$ have global phase ± 1 and are such that for each $i = 1, \ldots, n$, at least two of E_i, F_i, G_i are equal. Then $\mathbf{E}, \mathbf{F}, \mathbf{G}$ commute pairwise if and only if $\mathsf{N}_E, \mathsf{N}_F$ and N_G have the same parity.*

Proof. Given two arbitrary elements $\mathbf{H}, \mathbf{K} \in \mathscr{P}_n$ we have

$$\mathbf{HK} = (-1)^{\mathsf{N}(\mathbf{H},\mathbf{K})}\mathbf{KH},$$

where

$$\mathsf{N}\,(\mathbf{H}, \mathbf{K}) := |\{i \mid H_i \neq K_i \ \wedge \ H_i, K_i \neq I\}|.$$

Thus, $\mathbf{E}$ and $\mathbf{F}$ commute if and only if $\mathsf{N}\,(\mathbf{E}, \mathbf{F})$ is even. By hypothesis, for each i, at least two of E_i, F_i, G_i are equal, hence

$$\begin{aligned}\mathsf{N}\,(\mathbf{E}, \mathbf{F}) &= |\{i \mid G_i = E_i \neq F_i \ \wedge \ E_i, F_i, G_i \neq I\}| + |\{i \mid E_i \neq F_i = G_i \ \wedge \ E_i, F_i, G_i \neq I\}| \\ &= \mathsf{N}_F + \mathsf{N}_E.\end{aligned}$$

Therefore,

$$\mathbf{E}, \mathbf{F} \text{ commute} \quad \Leftrightarrow \quad \mathsf{N}_E \text{ and } \mathsf{N}_F \text{ have the same parity.}$$

Similarly,

$$\begin{aligned}\mathbf{F}, \mathbf{G} \text{ commute} \quad &\Leftrightarrow \quad \mathsf{N}_F \text{ and } \mathsf{N}_G \text{ have the same parity} \\ \mathbf{E}, \mathbf{G} \text{ commute} \quad &\Leftrightarrow \quad \mathsf{N}_E \text{ and } \mathsf{N}_G \text{ have the same parity,}\end{aligned}$$

and the result follows. □

Note that this new definition can be used to derive an alternative proof of the fact that any AvN triple for n-partite states can be reduced to an AvN triple that only involves 3 qubits, in accordance with Theorem VI.13. Indeed, given an AvN triple $\langle \mathbf{E}, \mathbf{F}, \mathbf{G} \rangle$ in $\mathscr{P}_n$, since $\mathsf{N}_G, \mathsf{N}_E, \mathsf{N}_F$ are odd, we can always choose 3 indices $1 \leq i_1, i_2, i_3 \leq n$ such that

$$E_{i_1} = F_{i_1} \neq G_{i_1}, \qquad E_{i_2} \neq F_{i_2} = G_{i_2}, \qquad E_{i_3} = G_{i_3} \neq F_{i_3}$$

Clearly, the elements of the triple restricted to these indices constitute an AvN triple in $\mathscr{P}_3$ and therefore an AvN argument.

The rationale for introducing Definition VI.18 is that it allows to better understand AvN triples from a computational perspective. We show a first example by providing a closed formula for the number of AvN triples in $\mathscr{P}_n$.

Proposition VI.20. *Let $n \geq 3$. The number of AvN triples in $\mathscr{P}_n$ is given by*

$$8 \cdot \left[\sum_{k=1}^{\frac{1}{2}(n+[n])-1} \binom{n}{2k+1} \binom{k+1}{k-1} \cdot 6^{2k+1} \cdot 22^{n-2k-1} \right],$$

where $[n] \in \mathbb{Z}_2$ denotes the parity of n.

PROOF. The factor of $2^3 = 8$ corresponds to the possible choices of global phase ± 1 for each element in the triple.

By Definition VI.18, an AvN triple $\langle \mathbf{E}, \mathbf{F}, \mathbf{G} \rangle$ is essentially determined by three odd numbers $\mathsf{N}_E, \mathsf{N}_F, \mathsf{N}_G$. Their sum $\mathsf{S} \coloneqq \mathsf{N}_E + \mathsf{N}_F + \mathsf{N}_G \leq n$ can be seen as the number of columns of the triple that play an active part in the AvN argument. Let us compute the amount of AvN triples having S "relevant" columns. We start by counting the number of solutions to the equation

$$\mathsf{N}_E + \mathsf{N}_F + \mathsf{N}_G = \mathsf{S},$$

where $\mathsf{N}_E, \mathsf{N}_F, \mathsf{N}_G, \mathsf{S}$ are all odd numbers. Let $k, e, f, g \geq 0$ be integers such that $\mathsf{S} = 2k+1$ and $\mathsf{N}_i = 2i + 1$ for $i = E, F, G$. We have

$$\mathsf{N}_E + \mathsf{N}_F + \mathsf{N}_G = \mathsf{S} \quad \Leftrightarrow \quad 2e+1+2f+1+2g+1 = 2k+1 \quad \Leftrightarrow \quad e+f+g = k-1$$

One can show[6] that the number of solutions to this equation is $\binom{k+1}{k-1}$. By condition (1) of Definition VI.18 we must choose two observables in $\{X, Y, Z\}$ (the order counts) in each of the S relevant columns, for a total of 6^{2k+1} possibilities. Finally, there are 8 possible configurations of each of the remaining $n - S$ non-relevant columns, namely

$$\begin{array}{ccccccccc} I & P & P & I & I & P & P & I \\ I & P & I & P & I & P & I & P \\ I & P & I & I & P & I & P & P \end{array}$$

where P has to be chosen in $\{X, Y, Z\}$ for a total of $(3\cdot 7+1)^{n-\mathsf{S}} = 22^{n-2k-1}$ possibilities. Hence, the number of AvN triples in $\mathscr{P}_n$ having $\mathsf{S} = 2k+1$ relevant columns is

$$\mathsf{N}_\mathsf{S} \coloneqq \binom{k+1}{k-1} \cdot 6^{2k+1} \cdot 22^{n-2k-1}.$$

Now, the amount of odd numbers of relevant columns $\mathsf{S} \leq n$ that we can select is given by

$$\sum_{k=1}^{\frac{1}{2}(n+[n])-1} \binom{n}{2k+1},$$

and the result follows. $\square$

5.2. Generating AvN triples. We devote this last section to the presentation of a computational method to generate all the AvN triples contained in $\mathscr{P}_n$. Until now, we only had a rather limited number of examples of quantum-realisable models featuring All-vs-Nothing proofs of strong contextuality. Thanks to the AvN triple theorem VI.13, the technique we introduce allows us to find all such models for a sufficiently small n.

Check vectors [**NC00**] are a useful way to represent elements of $\mathscr{P}_n$ in a computation-friendly way:

DEFINITION VI.21. Given an element $\mathbf{P} \coloneqq \alpha(P_i)_{i=1}^n \in \mathscr{P}_n$, its **check vector** $r(\mathbf{P})$ is a $2n$-vector

$$r(\mathbf{P}) = (x_1, x_2, \ldots, x_n, z_1, z_2, \ldots, z_n) \in \mathbb{Z}_2^{2n}$$

[6]This can be easily done e.g. by stars and bars [**Fel68**]

whose entries are defined as follows

$$(x_i, z_i) = \begin{cases} (0,0) & \text{if } P_i = I \\ (1,0) & \text{if } P_i = X \\ (1,1) & \text{if } P_i = Y \\ (0,1) & \text{if } P_i = Z. \end{cases}$$

Every check vector $r(\mathbf{P})$ completely determines $\mathbf{P}$ up to phase (i.e. $r(\mathbf{P}) = r(\alpha\mathbf{P})$ for all $\alpha \in \{\pm 1, \pm i\}$). We can use this representation to express the conditions for an AvN triple. More specifically, we represent an AvN triple, up to the global phases of each of its elements, as a matrix $M \in M_{3\times 2n}(\mathbb{Z}_2)$ whose rows are the check vectors of each element of the triple. Condition 1 of Definition VI.18 can be rewritten as

$$\forall 1 \leq j \leq n.\ \exists i, k \in \{1,2,3\}.\ \begin{cases} M_{i,j} = M_{k,j} \\ M_{i,n+j} = M_{k,n+j} \end{cases} \tag{VI.9}$$

The numbers $\mathsf{N}_E, \mathsf{N}_F, \mathsf{N}_G$ can also be easily computed. For instance, N_G equals the cardinality of the set

$$\{ i \in \{1,2,3\} \mid M_{1,j} = M_{3,j} \wedge M_{1,n+j} = M_{3,n+j} \ \wedge\ (M_{1,j} \neq M_{2,j} \vee M_{1,n+j} \neq M_{2,n+j}) \\ \wedge\ (M_{1,j} \neq 0 \vee M_{1,n+j} \neq 0)\ \wedge\ (M_{2,j} \neq 0 \vee M_{2,n+j} \neq 0) \ \}$$

Hence, in order to find all the AvN triples in $\mathscr{P}_n$ we need to solve the following problem:

$$\begin{array}{ll} \texttt{Find all} & M \in M_{3\times 2n}(\mathbb{Z}_2) \\ \texttt{such that} & M \text{ verifies (VI.9)}, \\ & \mathsf{N}_G,\ \mathsf{N}_E, \text{ and } \mathsf{N}_F \text{ are odd}, \end{array}$$

which is easily programmable.

An implementation of this method using Mathematica [**WR14**] can be found in [**Car16**], where we present the algorithm and the resulting list of all 216 AvN triples in $\mathscr{P}_3$ and all 19008 AvN triples in $\mathscr{P}_4$, disregarding the choice of global phases ± 1 for each element – in order to get the total number of AvN triples from Proposition VI.20, note that these numbers need to be multiplied by a factor of 8 to account for the choice of these global phases. By Theorem VI.13, this list generates all the possible AvN arguments for 3-qubit and 4-qubit stabiliser states.

Discussion

The recent formalisation and generalisation of All-vs-Nothing arguments in stabiliser quantum mechanics [**ABK+15**] allowed us to study their properties from a purely mathematical standpoint.

Thanks to this framework, we have introduced an important characterisation of AvN arguments for stabiliser states based on the combinatorial concept of AvN triple [**ABK+15**], leading to a computational technique to identify all such arguments. It remains an open question whether the AvN triple conjecture still holds for non-maximal stabiliser subgroups. We aim to investigate this question in future work.

The graph state formalism, which played a crucial rôle in the proof of the AvN triple theorem, also allowed us to infer an important structural feature of AvN arguments, namely that any such argument can be reduced to an AvN proof on three qubits, which is essentially a standard GHZ argument. This result shows in particular that the GHZ state is the only 3-qubit stabiliser state, up to LC-equivalence, admitting an AvN argument for strong contextuality.

Our computations provide a very large number of quantum-realisable strongly contextual empirical models admitting AvN arguments. These new models could potentially find applications in quantum information and computation, as well as contributing to the ongoing theoretical study of strong contextuality as a key feature of quantum mechanics.

The results of this chapter shall also be considered as part of the wider line of research focusing on characterising the quantum states that give rise to possibilistic forms of non-locality and contextuality. For instance, a major step forward in the classification of multi-partite pure states generating logically contextual empirical models has been made by Abramsky & Constantin in [**AC14**]. A natural question to ask in light of the conclusions of this chapter is whether it is possible to achieve a similar result for strong contextuality, i.e. a complete characterisation that goes beyond stabiliser quantum mechanics to include general n-qubit states. As it turns out, this question is considerably harder than the one addressed in [**AC14**], even for $n = 3$. Some decisive progress in this direction will be presented in the next chapter, where we analyse the minimum quantum resources needed to achieve maximal degrees of strong non-locality.

CHAPTER VII

Minimum quantum resources for strong non-locality

Summary

We analyse the minimum quantum resources needed to realise strong non-locality, as exemplified e.g. by the classical GHZ construction. We show that no two-qubit system, with any finite number of local measurements, can realise strong non-locality. For three-qubit systems, we show that strong non-locality can only be realised in the GHZ SLOCC class, and with equatorial measurements. However, we show that in this class there is an infinite family of states not LU-equivalent to the GHZ state that realise strong non-locality with finitely many measurements. These states have decreasing entanglement between one qubit and the other two, necessitating an increasing number of local measurements on the latter.

1. Overview

In this chapter, our aim is to analyse the minimum quantum resources needed to realise strong non-locality. The results of Chapter VI show that this phenomenon can be observed in experimental settings involving one-qubit local Pauli measurements on n-qubit stabiliser states, for $n \geq 3$. Here, we extend the scope of our search for strongly non-local behaviour beyond stabiliser quantum mechanics, and study general n-qubit states subject to single-qubit local measurements. Our focus on qubits is motivated by a result of Heywood and Redhead [**HR83**], which proves that strong non-locality can be achieved with a bipartite system, but with a qutrit on each site.

Let us summarise our conclusions:

- Our first result is limitative in character. We show that strong non-locality *cannot* be realised by a two-qubit system with any finite number of local measurements. Thus no bipartite Bell-type experiment with finitely many measurement settings can realise strong non-locality using a two-qubit state as a resource.

  There is a subtle counterpoint to this in a result from [**BKP06**], which shows that using a maximally entangled bipartite state, and an infinite family of local measurements, strong non-locality is achieved "in the limit" in a suitable sense. More precisely, as more and more measurements from the family are used, the local fraction – the part of the behaviour which can be accounted for by a local model (see Section 10.1.1 of Chapter V) – tends to 0, or equivalently the non-local fraction tends to 1. There is an interesting connection to this in our results for the tripartite case.

However, there is a practical advantage in being able to witness strong non-locality with a fixed finite number of measurements. If one wishes to design an experimental test for maximal non-locality, it is desirable that one can increase precision, i.e. increase the lower bound on the non-local fraction, without needing to expand the experimental setup – in particular, the number of measurement settings required to be performed – but rather by simply performing more runs of the same experiment.

- Having shown that strong non-locality cannot be realised in the two-qubit case, we turn to the analysis of three-qubit systems. Of course, we know by the results of Chapter VI that strong non-locality *can* be achieved in this case. Our aim is to analyse for which states, and with respect to which measurements, can strong non-locality be achieved. We use the classification into *Stochastic local operations and classical communication (SLOCC) classes* for tripartite qubit systems from [**DVC00**]. According to this analysis, there are two maximal SLOCC classes, the GHZ and W classes. Below these, there are the degenerate cases of products of an entangled bipartite state with a one-qubit state, e.g. $AB{-}C$. By the previous result, these degenerate cases cannot realise strong non-locality. We furthermore show that no state in the W class can realise strong non-locality, for any choice of finitely-many local measurements.
- This leaves us with the GHZ SLOCC class. We use the detailed description of this class as a parameterised family of states from [**DVC00**]. We first show that any state in this class witnessing strong non-locality with finitely many local measurements must satisfy a number of constraints on the parameters. In particular, the state must be balanced in the sense that the coefficients in its unique linear decomposition into a pair of product states have the same complex modulus. We furthermore show that *only equatorial measurements need be considered* (the equators being uniquely determined by the state) – no other measurements can contribute to a strong non-locality argument.
- Having thus narrowed the possibilities for realising strong non-locality considerably, we find a new infinite family of models displaying strong non-locality using states within the GHZ SLOCC class that are not LU-equivalent to the GHZ state. The states in this family start from GHZ and tend in the limit to the state $|\Phi^+\rangle \otimes |+\rangle$ in the AB–C class with maximal entanglement on the first two qubits, and in product with the third. This family is actually closely related to the construction from [**BKP06**] in which an increasing number of measurements on a bipartite maximally entangled state eventually squeezes the local fraction to zero in the limit. Our family is obtained by adding a third qubit to this setup, with two available local measurements, and some entanglement between the first two qubits and the third one, thus allowing strong non-locality to be witnessed with a finite number of measurements. There is a trade-off between the number of measurement settings available on the first two qubits – and, consequently, the lower bound for the non-local fraction these measurements can witness – and the amount of entanglement necessary between the third qubit and the original two.

The content of this chapter has been developed in collaboration with Samson Arbamsky, Rui Soares Barbosa, Nadish de Silva, Kohei Kishida and Shane Mansfield. The results have been published in [**ABC$^+$17**].

Outline of the chapter. Section 2 summarises some background material on non-locality and entanglement classification of three-qubit states, Section 3 shows that strong non-locality cannot be witnessed by two-qubit states and a finite number of local measurements; Section 4 does the same for three-qubit states in the SLOCC class of W; Section 5 deals with states in the SLOCC class of GHZ, deriving conditions on these necessary for strong non-locality; Section 6 presents the family of strong non-locality arguments using states in the GHZ-SLOCC class.

2. Background

2.1. Quantum models and scenarios. In this chapter, we shall be concerned only with $(n, k, 2)$ Bell-type scenarios. More specifically, our ideal n-partite scenario will be of the form $\langle X, \mathcal{M}, O\rangle$, where $X = X_1 \sqcup \cdots \sqcup X_n$, and $O = \{\pm 1\}$ to model the outcome of a single-qubit observable. Joint measurement in X will be denoted as tuples $\mathbf{m} = \langle m_1, \ldots, m_n\rangle \in X_1 \times \cdots \times X_n$, since they are determined by a choice of measurement for each of the n-parties (so $\mathcal{M} \cong \prod_{i=1}^n X_i$).

Given any joint measurement $\mathbf{m} \in \mathcal{M}$, the corresponding set of events $\mathcal{E}(\mathbf{m})$ is isomorphic to O^n, thus the empirical models considered in this chapter will be of the form $\left\{e_{\mathbf{m}} \in \mathcal{D}_{\mathbb{R}_{\geq 0}}(O^n)\right\}_{\mathbf{m}\in\mathcal{M}}$. Given a vector of outcomes $\mathbf{o} = \langle o_1, \ldots, o_n\rangle \in O^n$, the probability $e_{\mathbf{m}}(\mathbf{o})$ of obtaining the joint outcomes $\mathbf{o}$ upon performing the measurements $\mathbf{m}$ at each site will often be denoted as follows:

$$e_{\mathbf{m}}(\mathbf{o}) = \mathsf{Prob}(\mathbf{o}|\mathbf{m}) = \mathsf{Prob}(o_1, \ldots, o_n | m_1, \ldots, m_n).$$

Since we are interested in studying strong non-locality, we shall be particularly concerned with finding global assignments $g = \bigsqcup_{i=1}^n g_i \colon \bigsqcup_{i=1}^n X_i \longrightarrow O$, where each g_i is a map $g_i : X_i \to O$. We recall that such a global assignment is compatible with a probabilistic model $\{e_{\mathbf{m}}\}_{\mathbf{m}\in\mathcal{M}}$ if and only if, for any choice of measurements $\mathbf{m} = \langle m_1, \ldots, m_n\rangle \in \mathcal{M}$, we have

$$e_{\mathbf{m}}(g(\mathbf{m})) = \mathsf{Prob}(g(\mathbf{m})|\mathbf{m}) = \mathsf{Prob}(g_1(m_1), \ldots, g_n(m_n)|m_1, \ldots, m_n) > 0,$$

where $g(\mathbf{m}) = \langle g_1(m_1), \ldots, g_n(m_n)\rangle$.

The Bloch sphere representation of one-qubit pure states will be useful: assuming a preferred orthonormal basis $\{|0\rangle, |1\rangle\}$ of $\mathbb{C}^2$, we shall use the notation

$$|\theta, \varphi\rangle := \cos\frac{\theta}{2}|0\rangle + e^{i\varphi}\sin\frac{\theta}{2}|1\rangle$$

for any $\theta \in [0, \pi]$ and $\varphi \in [0, 2\pi)$.

Any single-qubit projective measurement is fully determined by specifying such a normalised vector in $\mathbb{C}^2$, namely the pure state corresponding to the $+1$ eigenvalue or outcome. Hence, the set of local measurements for a single qubit is labelled by

$$\mathsf{LM} = [0, \pi] \times [0, 2\pi)$$

The quantum measurement determined by $(\theta, \varphi) \in \mathsf{LM}$ has eigenvalues $O = \{+1, -1\}$ with the eigenvector corresponding to outcome $o \in O$ given by:

$$|\theta, \varphi \mapsto o\rangle := \begin{cases} |\theta, \varphi\rangle & \text{if } o = +1 \\ |\pi - \theta, \varphi + \pi\rangle & \text{if } o = -1 \end{cases}$$

Throughout this chapter, we shall be considering the n-partite measurement scenario with $X_i = \mathsf{LM}$ for every site. Measurement contexts correspond to a choice of single qubit measurements for each of the n sites, represented by a tuple $(\boldsymbol{\theta}, \boldsymbol{\varphi}) = \langle(\theta_1, \varphi_1), \ldots, (\theta_n, \varphi_n)\rangle$. Performing all the measurements of a context in parallel yields an outcome $\mathbf{o} = \langle o_1, \ldots, o_n\rangle \in O^n$. The vector corresponding to this outcome is denoted

$$|\boldsymbol{\theta}, \boldsymbol{\varphi} \mapsto \mathbf{o}\rangle := |\theta_1, \varphi_1 \mapsto o_1\rangle \otimes \cdots \otimes |\theta_n, \varphi_n \mapsto o_n\rangle .$$

We shall also find it useful to write

$$|\boldsymbol{\theta}, \boldsymbol{\varphi}\rangle := |\theta_1, \varphi_1\rangle \otimes \cdots \otimes |\theta_n, \varphi_n\rangle = |\boldsymbol{\theta}, \boldsymbol{\varphi} \mapsto \langle +1, \ldots, +1\rangle\rangle$$

for the vector corresponding to the joint outcome assigning $+1$ at every site.

An n-qubit state $|\psi\rangle$ determines an empirical model $e^{|\psi\rangle}$ for this measurement scenario:

$$e^{|\psi\rangle}_{(\boldsymbol{\theta}, \boldsymbol{\varphi})}(\mathbf{o}) = \mathsf{Prob}^{|\psi\rangle}(o_1, \ldots, o_n | (\theta_1, \varphi_1), \ldots, (\theta_n, \varphi_n)) := |\langle \boldsymbol{\theta}, \boldsymbol{\varphi} \mapsto \mathbf{o} | \psi\rangle|^2.$$

We are concerned with checking for strongly non-local behaviour on such a model. This amounts to checking for the existence of maps $g_i \colon \mathsf{LM} \longrightarrow O$ for each site such that for any choice of measurements $(\boldsymbol{\theta}, \boldsymbol{\varphi})$, the corresponding outcome has positive probability:

$$\begin{aligned} e_{(\boldsymbol{\theta}, \boldsymbol{\varphi})}(g(\boldsymbol{\theta}, \boldsymbol{\varphi})) &= \mathsf{Prob}^{|\psi\rangle}(g_1(\theta_1, \varphi_1), \ldots, g_n(\theta_n, \varphi_n) | (\theta_1, \varphi_1), \ldots, (\theta_n, \varphi_n)) \\ &= |\langle \boldsymbol{\theta}, \boldsymbol{\varphi} \mapsto g(\boldsymbol{\theta}, \boldsymbol{\varphi}) | \psi\rangle|^2 > 0. \end{aligned}$$

Given that these are quantum probabilities, we can rephrase this condition in terms of non-vanishing amplitudes: $\langle \boldsymbol{\theta}, \boldsymbol{\varphi} \mapsto g(\boldsymbol{\theta}, \boldsymbol{\varphi}) | \psi\rangle \neq 0$.

The following fact will be used throughout. Suppose we want to check the consistency with the empirical model of a given global assignment $g = \bigsqcup_{i=1}^n g_i$. If this assignment satisfies

$$\forall i \in \{1, \ldots, n\}.\ g_i(\theta, \varphi) = -g_i(\pi - \theta, \varphi + \pi), \tag{VII.1}$$

that is, measurements with $+1$ eigenstates diametrically opposed in the Bloch spehere (i.e. measurements that are the negation of each other) are assigned opposite outcomes, then

$$|\theta, \varphi \mapsto g_i(\theta, \varphi)\rangle = \begin{cases} |\theta, \varphi\rangle & \text{if } g_i(\theta, \varphi) = +1 \\ |\pi - \theta, \varphi + \pi\rangle & \text{if } g_i(\theta, \varphi) = -1 \quad (\Leftrightarrow g_i(\pi - \theta, \varphi + \theta) = +1) \end{cases}$$

meaning that $|\boldsymbol{\theta}, \boldsymbol{\varphi} \mapsto g(\boldsymbol{\theta}, \boldsymbol{\varphi})\rangle = |\boldsymbol{\theta}', \boldsymbol{\varphi}'\rangle$ with $g_i(\theta'_i, \varphi'_i) = +1$ for all i. In other words, should we wish to calculate the amplitude for a joint outcome $\mathbf{o}$ on a given context $(\boldsymbol{\theta}, \boldsymbol{\varphi})$, we may equivalently calculate the amplitude for the joint outcome $\langle +1, \ldots, +1\rangle$ on a new context $(\boldsymbol{\theta}', \boldsymbol{\varphi}')$ obtained by substituting $\theta_i \mapsto \pi - \theta_i$ and $\varphi_i \mapsto \pi + \varphi_i$ for all i such that $o_i = -1$. Therefore, it suffices to verify the equation $\langle \boldsymbol{\theta}, \boldsymbol{\varphi} \mapsto g(\boldsymbol{\theta}, \boldsymbol{\varphi}) | \psi\rangle \neq 0$ for all contexts whose measurements are all assigned $+1$. Indeed, the same is true if (VII.1) is relaxed to simply say that $g_i(\pi - \theta, \varphi + \pi) = -1 \Rightarrow g_i(\theta, \varphi) = +1$. Incidentally, even though we shall not need this fact, note that if there is any global assignment consistent with the model, there will be one that satisfies (VII.1), for this would only require a subset of the conditions.

We conclude this subsection with two observations regarding these particular quantum empirical models. First, note that local unitaries (LU) on the state don't affect non-locality, or indeed strong non-locality, of the resulting empirical model. This follows from the fact that by moving from the Schrödinger to the Heisenberg picture, we may equivalently leave the state fixed and apply the corresponding unitaries to the sets of available local measurements. Since the available local measurements are all the projective one-qubit measurements, a local unitary, which can be seen as a rotation of the Bloch sphere, merely maps this set to itself. Secondly, if we are dealing with a product state of n-qubits, $|\psi\rangle = |\psi_1\rangle \otimes \cdots \otimes |\psi_n\rangle$, then the resulting empirical model is necessarily local. This is because the probabilities factorise:

$$\mathsf{Prob}^{|\psi\rangle}(\mathbf{o}|(\boldsymbol{\theta},\boldsymbol{\varphi})) = |\langle \boldsymbol{\theta},\boldsymbol{\varphi} \mapsto \mathbf{o}|\psi\rangle|^2 = \left|\prod_{i=1}^{n} \langle \theta_i, \varphi_i \mapsto o_i|\psi_i\rangle\right|^2 = \prod_{i=1}^{n} |\langle \theta_i, \varphi_i \mapsto o_i|\psi_i\rangle|^2 .$$

2.2. SLOCC classes of three-qubit states. A classification of multipartite quantum states by their degree of entanglement is given by the notion of *LOCC* (local operations and classical communication) equivalence [**BBPS96, Nie99, KLM99**]. A protocol is said to be LOCC if it is of the following form: each party may perform local measurements and transformations on their system, and may communicate measurement outcomes to the other parties, so that local operations may be conditioned on measurement outcomes anywhere in the system. A state $|\psi_1\rangle$ is LOCC-convertible to a state $|\psi_2\rangle$ if there exists a LOCC protocol that *deterministically* produces $|\psi_2\rangle$ when starting with $|\psi_1\rangle$. Intuitively, such a protocol cannot increase the degree of entanglement and so we think of $|\psi_1\rangle$ as being at least as entangled as $|\psi_2\rangle$. The notion of LOCC-convertibility defines a preorder[1] on multipartite states that in turn yields a notion of LOCC-equivalence of states: the states $|\psi\rangle$ and $|\phi\rangle$ are LOCC-equivalent when $|\psi\rangle$ is LOCC-convertible to $|\phi\rangle$ and *vice versa*. The LOCC-convertibility preorder then naturally defines a partial order on the collection of LOCC equivalence classes of states.

A coarser classification of multipartite quantum states is given by relaxing the requirement that our conversion protocols succeed deterministically to the requirement that they succeed with non-zero probability [**BPR$^+$00**]. The previous paragraph holds true for SLOCC (stochastic LOCC) *mutatis mutandis*. Note that equivalence of two states under LU transformations implies their SLOCC-equivalence. More generally, two states are SLOCC-equivalent if and only if they are related by an invertible local operator (ILO) [**DVC00**].

Dür, Vidal, and Cirac [**DVC00**] classified the SLOCC classes of three-qubit systems and found there to be exactly six classes (see Figure VII.1). The GHZ and W states are representatives of the two maximal, non-comparable classes. Three intermediate classes are characterised by bipartite entanglement between two of the qubits, which are in a product with the third. Finally, the minimal class is given by product states.

[1] A preorder is a reflexive and transitive relation; i.e. it is like a partial order except that it can deem two distinct elements equivalent.

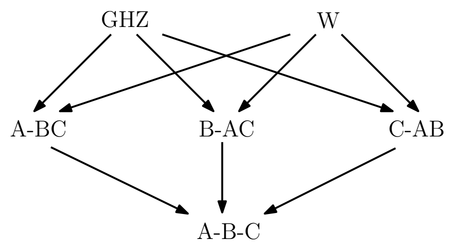


FIGURE VII.1. Hasse diagram of the partial order of three-qubit SLOCC classes.

By the last observation in the previous section, it is obvious that a state in the A–B–C class cannot realise non-locality, and that the case of a state in one of the intermediate classes can be reduced to that of the two qubits that are entangled. Hence, we shall first discuss strong non-locality for two-qubit states and then proceed in turn to each of the maximal SLOCC classes of three-qubit states, W and GHZ.

3. Two-qubit states are not strongly non-local

Every two-qubit state can be written, up to LU, uniquely as

$$|\psi\rangle = \cos\delta\,|00\rangle + \sin\delta\,|11\rangle\,, \tag{VII.2}$$

where $\delta \in [0, \frac{\pi}{4}]$. The state (VII.2) is either: the product state $|00\rangle$, which is obviously non-contextual since it is separable, when $\delta = 0$; or an entangled state in the SLOCC class of the Bell state $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, when $\delta > 0$.

THEOREM VII.1. *Two-qubit states do not admit strongly non-local behaviour.*

PROOF. This proof rests on defining an explicit global assignment $g : \mathsf{LM} \sqcup \mathsf{LM} \to O$ consistent with the possible events of the empirical model. More specifically, the map g is obtained by assigning outcome $+1$ to one hemisphere of the Bloch sphere, and -1 to the other, with special conditions on the poles and a slight asymmetry between the two parties.

We start by computing the amplitude $\langle \boldsymbol{\theta}, \boldsymbol{\varphi} | \psi\rangle$ of measuring $(\boldsymbol{\theta}, \boldsymbol{\varphi}) = \langle(\theta_1, \varphi_1), (\theta_2, \varphi_2)\rangle$ on the general state (VII.2) and obtaining joint outcome $\langle +1, +1\rangle$:

$$\langle \boldsymbol{\theta}, \boldsymbol{\varphi} | \psi\rangle = \cos\delta \cos\frac{\theta_1}{2}\cos\frac{\theta_2}{2} + \sin\delta\sin\frac{\theta_1}{2}\sin\frac{\theta_2}{2}e^{-i(\varphi_1+\varphi_2)}$$

Since $\delta = 0$ gives rise to a product state, we will assume $\delta \neq 0$.

We define the following maps:

$$g_1 \colon \mathsf{LM} \longrightarrow O :: (\theta, \varphi) \longmapsto \begin{cases} +1 & \text{if } \theta = \pi \text{ or } \left(\theta \neq 0 \text{ and } \varphi \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right)\right) \\ -1 & \text{if } \theta = 0 \text{ or } \left(\theta \neq \pi \text{ and } \varphi \in \left[\frac{\pi}{2}, \frac{3\pi}{2}\right)\right) \end{cases}$$

$$g_2 \colon \mathsf{LM} \longrightarrow O :: (\theta, \varphi) \longmapsto \begin{cases} +1 & \text{if } \theta = \pi \text{ or } \left(\theta \neq 0 \text{ and } \varphi \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right]\right) \\ -1 & \text{if } \theta = 0 \text{ or } \left(\theta \neq \pi \text{ and } \varphi \in \left(\frac{\pi}{2}, \frac{3\pi}{2}\right]\right) \end{cases}$$

and let $g := g_1 \sqcup g_2 \colon \mathsf{LM} \sqcup \mathsf{LM} \longrightarrow O$ be a global assignment. A graphical representation of the map g can be found in Figure VII.2.

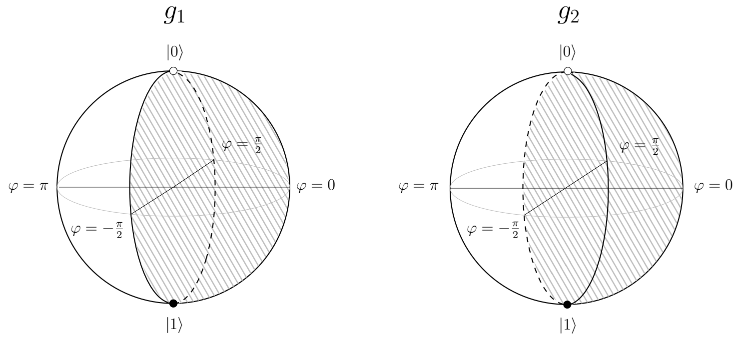


FIGURE VII.2. Graphical representation of the global assignment g. The shaded region corresponds to the measurements mapped to $+1$ by g.

Let $(\boldsymbol{\theta}, \boldsymbol{\varphi})$ be a context whose individual measurements are mapped to $+1$ by g (see Section 2.1 for why this is sufficient). In particular, it holds that $\theta_1, \theta_2 \neq 0$. Since $\delta \neq 0$, we have

$$s := \sin\delta \sin\frac{\theta_1}{2} \sin\frac{\theta_2}{2} > 0 \qquad \text{and} \qquad c := \cos\delta \cos\frac{\theta_1}{2} \cos\frac{\theta_2}{2} \geq 0.$$

If $\theta_1 = \pi$ or $\theta_2 = \pi$, then $c = 0$, which implies $\langle \boldsymbol{\theta}, \boldsymbol{\varphi} | \psi \rangle = s e^{-i(\varphi_1+\varphi_2)} \neq 0$. Otherwise, $\varphi_1 \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right)$, $\varphi_2 \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right]$ and $\langle \boldsymbol{\theta}, \boldsymbol{\varphi} | \psi \rangle = c + s e^{-i(\varphi_1+\varphi_2)}$ is the sum of a positive real number and a non-zero complex number. For it to be zero, the latter must be real and negative, hence

$$\varphi_1 + \varphi_2 = \pi \mod 2\pi,$$

which cannot be satisfied in the domain of φ_1, φ_2. $\square$

4. W-SLOCC states are not strongly non-local

A general state in the SLOCC class of the W state $|\mathsf{W}\rangle = \frac{1}{\sqrt{3}}(|001\rangle + |010\rangle + |100\rangle)$ can be written, up to LU, as

$$|\psi_{\mathsf{W}}\rangle = \sqrt{a}|001\rangle + \sqrt{b}|010\rangle + \sqrt{c}|100\rangle + \sqrt{d}|000\rangle, \tag{VII.3}$$

where $a, b, c \in \mathbb{R}_{>0}$ and $d := 1 - (a+b+c) \in \mathbb{R}_{\geq 0}$. Indeed, we can obtain $|\psi_{\mathsf{W}}\rangle$ from $|\mathsf{W}\rangle$ by applying the following ILO to $|\mathsf{W}\rangle$:

$$\begin{pmatrix} \sqrt{a} & \sqrt{b} \\ 0 & \sqrt{c} \end{pmatrix} \otimes \begin{pmatrix} \sqrt{3} & 0 \\ 0 & \frac{\sqrt{3b}}{\sqrt{a}} \end{pmatrix} \otimes I.$$

In order to prove that W-SLOCC states are not strongly non-local, we will need the following lemma, which generalises the argument used in the proof of Theorem VII.1 to show that the amplitude could not be zero.

LEMMA VII.2. *Let $z_1, \ldots, z_m \in \mathbb{C}$, and $r \in \mathbb{R}_{\geq 0}$. If*

$$\sum_{i=1}^{m} z_i + r = 0, \tag{VII.4}$$

then one of the following holds: (i) $z_1 = \cdots = z_m = r = 0$*; (ii) there exists a* $z_k \in \mathbb{R}_{<0}$*; (iii) there exists* $1 \leq k, l \leq m$ *such that* $\mathrm{Arg}(z_k) \in (0, \pi)$ *and* $\mathrm{Arg}(z_l) \in (-\pi, 0)$*.*

PROOF. If all the z_i are real, then, since r is non-negative, we must have either (i) or (ii). Now, suppose there is a $1 \leq k \leq m$ such that $\mathrm{Im}(z_k) \neq 0$. By (VII.4), we have $\sum_{i=1}^{n} \mathrm{Im}(z_i) = 0$. Thus,

$$\sum_{i \neq k} \mathrm{Im}(z_i) = -\mathrm{Im}(z_k) \quad \Leftrightarrow \quad \sum_{i \neq k} |z_i| \sin(\mathrm{Arg}(z_i)) = -|z_k| \sin(\mathrm{Arg}(z_k)).$$

Hence, there exists at least one $l \neq k$ for which the sign of $\mathrm{Im}(z_l)$ is opposite to that of $\mathrm{Im}(z_k)$, which implies that z_l and z_k are in different sides of the real axis, implying the condition about $\mathrm{Arg}(z_l)$ and $\mathrm{Arg}(z_k)$. □

THEOREM VII.3. *States in the SLOCC class of W do not admit strongly non-local behaviour.*

PROOF. Similarly to the bipartite case of Theorem VII.1, the key idea of the proof is the definition of a global assignment $g : \mathsf{LM} \sqcup \mathsf{LM} \sqcup \mathsf{LM} \to O$ whose restriction to each context is contained in the support of the model. Once again, g is obtained by partitioning the Bloch sphere into two hemispheres to which are assigned different outcomes, with asymmetric polar conditions across the parties.

We start by computing the amplitude $\langle \boldsymbol{\theta}, \boldsymbol{\varphi} | \psi_{\mathsf{W}} \rangle$ of measuring $(\boldsymbol{\theta}, \boldsymbol{\varphi})$ on the general state (VII.3) and obtaining joint outcome $\langle +1, +1, +1 \rangle$:

$$\begin{aligned} \langle \boldsymbol{\theta}, \boldsymbol{\varphi} | \psi_{\mathsf{W}} \rangle = {} & \underbrace{\sqrt{a} \left(\cos\frac{\theta_1}{2} \cos\frac{\theta_2}{2} \sin\frac{\theta_3}{2} e^{-i\varphi_3} \right)}_{=: z_3 \in \mathbb{C}} + \underbrace{\sqrt{b} \left(\cos\frac{\theta_1}{2} \cos\frac{\theta_3}{2} \sin\frac{\theta_2}{2} e^{-i\varphi_2} \right)}_{=: z_2 \in \mathbb{C}} \\ & + \underbrace{\sqrt{c} \left(\cos\frac{\theta_2}{2} \cos\frac{\theta_3}{2} \sin\frac{\theta_1}{2} e^{-i\varphi_1} \right)}_{=: z_1 \in \mathbb{C}} + \underbrace{\sqrt{d} \left(\cos\frac{\theta_1}{2} \cos\frac{\theta_2}{2} \cos\frac{\theta_3}{2} \right)}_{=: r \in \mathbb{R}_{\geq 0}}. \end{aligned} \tag{VII.5}$$

Define the following functions:

$$h = g_1 = g_2 \colon \mathsf{LM} \longrightarrow O :: (\theta, \varphi) \longmapsto \begin{cases} +1 & \text{if } \theta = 0 \text{ or } (\theta \neq \pi \text{ and } \varphi \in (-\pi, 0]) \\ -1 & \text{if } \theta = \pi \text{ or } (\theta \neq 0 \text{ and } \varphi \in (0, \pi]) \end{cases}$$

$$g_3 \colon \mathsf{LM} \longrightarrow O :: (\theta, \varphi) \longmapsto \begin{cases} +1 & \text{if } \theta = \pi \text{ or } (\theta \neq 0 \text{ and } \varphi \in (-\pi, 0]) \\ -1 & \text{if } \theta = 0 \text{ or } (\theta \neq \pi \text{ and } \varphi \in (0, \pi]) \end{cases}$$

and let $g := h \sqcup h \sqcup g_3 \colon \mathsf{LM} \sqcup \mathsf{LM} \sqcup \mathsf{LM} \longrightarrow O$ be a global assignment. The map g is graphically represented in Figure VII.3.

Let $(\boldsymbol{\theta}, \boldsymbol{\varphi})$ be a context whose individual measurements are mapped to $+1$ by g. In particular, $\theta_1, \theta_2 \neq \pi$ and $\theta_3 \neq 0$. Since $a > 0$, we have

$$|z_3| = \sqrt{a} \cos\frac{\theta_1}{2} \cos\frac{\theta_2}{2} \sin\frac{\theta_3}{2} > 0,$$

which implies $z_3 \neq 0$. Now, if $\theta_3 = \pi$, then $z_1 = z_2 = r = 0$ and $\langle \boldsymbol{\theta}, \boldsymbol{\varphi} | \psi_{\mathsf{W}} \rangle = z_3 \neq 0$.

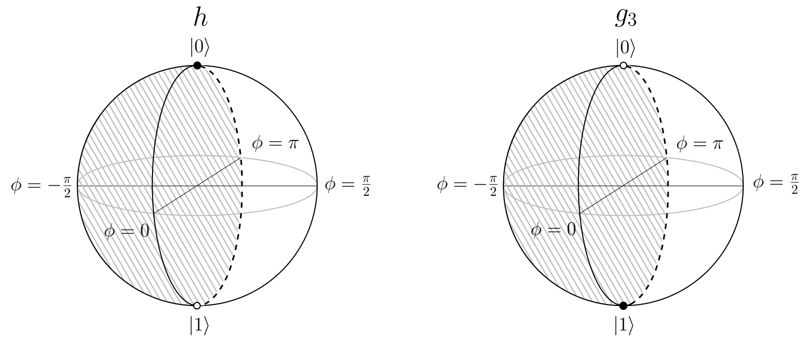


FIGURE VII.3. Graphical representation of the global assignment g. The shaded region corresponds to the measurements mapped to $+1$ by g.

Otherwise, $\theta_3 \neq \pi$ and $\varphi_3 \in (-\pi, 0]$, implying that $\mathrm{Arg}(z_3) = -\varphi_3 \in [0, \pi)$. For $i = 1, 2$, we either have $\theta_i = 0$ or $\varphi_i \in (-\pi, 0]$, implying that $z_i = 0$ or $\mathrm{Arg}(z_i) = -\varphi_i \in [0, \pi)$. Using Lemma VII.2, we conclude that $\langle \boldsymbol{\theta}, \boldsymbol{\varphi} | \psi_{\mathsf{W}} \rangle \neq 0$: (i) fails because $z_3 \neq 0$, while (ii) and (iii) fail because $\mathrm{Arg}(z_i) \in [0, \pi)$ whenever $z_i \neq 0$. □

5. Strong non-locality in the SLOCC class of GHZ

5.1. The n-partite GHZ state and local equatorial measurements. Before we tackle the general case of GHZ-SLOCC states, we consider the GHZ state itself. We show that equatorial measurements are the only relevant ones in the study of strong non-locality for this state. In fact, this holds for the general n-partite GHZ state,

$$|\mathsf{GHZ}(n)\rangle := \frac{1}{\sqrt{2}} \left(|0\rangle^{\otimes n} + |1\rangle^{\otimes n} \right),$$

and consequentely, in light of the remark towards the end of Section 2.1, for any state in its LU class. In the next section, we generalise this result to arbitrary states in the SLOCC class of the tripartite GHZ state, and study conditions for strong non-locality within this class.

THEOREM VII.4. *Any strongly non-local behaviour of $|\mathsf{GHZ}(n)\rangle$ can be witnessed using only equatorial measurements. That is, there is a global assignment g consistent with the model $e^{|\mathsf{GHZ}(n)\rangle}$ in all contexts that are not exclusively composed of equatorial measurements.*

PROOF. The proof is achieved using a construction of a global assignment similar to the ones previously discussed. First, we derive the formula for the amplitude $\langle \boldsymbol{\theta}, \boldsymbol{\varphi} | \mathsf{GHZ}(n) \rangle$ of measuring $(\boldsymbol{\theta}, \boldsymbol{\varphi})$ and obtaining joint outcome $\langle +1, \ldots, +1 \rangle$:

$$\langle \boldsymbol{\theta}, \boldsymbol{\varphi} | \mathsf{GHZ}(n) \rangle = \frac{1}{\sqrt{2}} \left(\prod_{i=1}^{n} \cos \frac{\theta_i}{2} + e^{-i \sum_{i=1}^{n} \varphi_i} \prod_{i=1}^{n} \sin \frac{\theta_i}{2} \right).$$

Consider the function

$$h \colon \mathsf{LM} \longrightarrow O :: (\theta, \varphi) \longmapsto \begin{cases} +1 & \text{if } \theta \in \left[0, \frac{\pi}{2}\right] \\ -1 & \text{if } \theta \in \left(\frac{\pi}{2}, \pi\right] \end{cases}$$

i.e. h assigns $+1$ to the equator and the northern hemisphere, and -1 to the southern hemisphere. Let $g := \bigsqcup_{i=1}^{n} h \colon \bigsqcup_{i=1}^{n} \mathsf{LM} \longrightarrow O$. We show that this global assignment is consistent with the probabilities at all contexts that include at least a non-equatorial measurement.

Let $(\boldsymbol{\theta}, \boldsymbol{\varphi})$ be a context whose measurements are mapped to $+1$ by g. In particular, $\theta_i \leq \frac{\pi}{2}$ for all i. If $\langle \boldsymbol{\theta}, \boldsymbol{\varphi} | \mathsf{GHZ}(n) \rangle = 0$, then

$$\prod_{i=1}^{n} \cos \frac{\theta_i}{2} = -e^{-i(\sum_{i=1}^{n} \varphi_i)} \prod_{i=1}^{n} \sin \frac{\theta_i}{2}$$

Taking the modulus of both sides and dividing the right-hand by the left-hand side yields:

$$\prod_{i=1}^{n} \tan \frac{\theta_i}{2} = 1$$

which is verified if and only if $\theta_i = \frac{\pi}{2}$ for all $1 \leq i \leq n$. $\square$

5.2. Balanced GHZ-SLOCC states and local equatorial measurements. A general state in the SLOCC class of the GHZ state can be written, up to LU, as

$$|\psi_{\mathsf{GHZ}}\rangle = \sqrt{K}(\cos\delta\,|000\rangle + \sin\delta e^{i\Phi}\,|\varphi_1\rangle\,|\varphi_2\rangle\,|\varphi_3\rangle), \tag{VII.6}$$

where $K = (1 + 2\cos\delta\sin\delta\cos\alpha\cos\beta\cos\gamma\cos\Phi)^{-1}$, and

$$|\varphi_1\rangle = \cos\alpha\,|0\rangle + \sin\alpha\,|1\rangle\,, \quad |\varphi_2\rangle = \cos\beta\,|0\rangle + \sin\beta\,|1\rangle\,, \quad |\varphi_3\rangle = \cos\gamma\,|0\rangle + \sin\gamma\,|1\rangle\,,$$

for some $\delta \in (0, \pi/4]$, $\alpha, \beta, \gamma \in (0, \pi/2]$, and $\Phi \in [0, 2\pi)$. Indeed, $|\psi_{\mathsf{GHZ}}\rangle$ is obtained from $|\mathsf{GHZ}\rangle$ via the ILO

$$\sqrt{2K}\begin{pmatrix} \cos\delta & \sin\delta\cos\alpha e^{i\Phi} \\ 0 & \sin\delta\sin\alpha e^{i\Phi} \end{pmatrix} \otimes \begin{pmatrix} 1 & \cos\beta \\ 0 & \sin\beta \end{pmatrix} \otimes \begin{pmatrix} 1 & \cos\gamma \\ 0 & \sin\gamma \end{pmatrix}.$$

In order to prove the results of this section, it is convenient to describe $|\psi_{\mathsf{GHZ}}\rangle$ in a slightly different form. By applying local unitaries, we can rewrite it as

$$|\psi_{\mathsf{GHZ}}\rangle = \sqrt{K}(\cos\delta\,|v_{\lambda_1}\rangle\,|v_{\lambda_2}\rangle\,|v_{\lambda_3}\rangle + \sin\delta e^{i\Phi}\,|w_{\lambda_1}\rangle\,|w_{\lambda_2}\rangle\,|w_{\lambda_3}\rangle), \tag{VII.7}$$

where

$$|v_\lambda\rangle = |\lambda, 0\rangle = \cos\frac{\lambda}{2}\,|0\rangle + \sin\frac{\lambda}{2}\,|1\rangle\,, \quad |w_\lambda\rangle = |\pi - \lambda, 0\rangle = \sin\frac{\lambda}{2}\,|0\rangle + \cos\frac{\lambda}{2}\,|1\rangle \tag{VII.8}$$

for some $\lambda_i \in [0, \frac{\pi}{2})$, $i = 1, 2, 3$. The action of this LU can be thought of as choosing a new orthonormal basis for each qubit: a graphical illustration of this process can be found in Figure VII.4. A key advantage of this LU-equivalent description of a general state in the GHZ SLOCC class is that the equator of the i-th qubit's Bloch sphere coincides with the great circle that bisects the i-th components of the two unique product states that form a linear decomposition of the state. Note that any state in the GHZ SLOCC class thus uniquely defines an equator in each Bloch sphere. It is to the measurements lying on these that we refer as being **equatorial**.

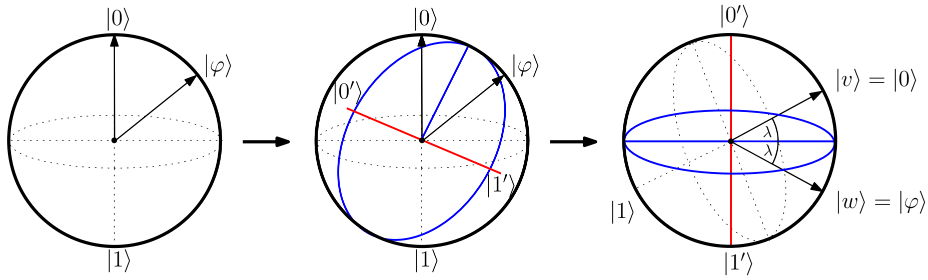


FIGURE VII.4. Choice of a new basis $\{|0'\rangle, |1'\rangle\}$ for each qubit that allows the state to be described in the form (VII.7).

DEFINITION VII.5. We say that a state in the GHZ SLOCC class is **balanced** if the coefficients in its unique linear decomposition into a pair of product states have the same complex modulus – when the state is written in the form (VII.7), this corresponds to having $\delta = \frac{\pi}{4}$, hence $\cos\delta = \sin\delta = \frac{1}{\sqrt{2}}$.

LEMMA VII.6. *Let $|v_\lambda\rangle$ and $|w_\lambda\rangle$ be given as in* (VII.8)*, with $\lambda \in [0, \pi/2)$, and consider a measurement (θ, φ) with $\theta \in [0, \pi/2)$, i.e. with $+1$ eigenstate in the 'northern hemisphere'. Then $|\langle\theta, \varphi|v_\lambda\rangle| > |\langle\theta, \varphi|w_\lambda\rangle|$.*

PROOF. We have

$$|\langle\theta,\varphi|v_\lambda\rangle| > |\langle\theta,\varphi|w_\lambda\rangle| \Leftrightarrow \left|\cos\frac{\theta}{2}\cos\frac{\lambda}{2} + \sin\frac{\theta}{2}\sin\frac{\lambda}{2}e^{-i\varphi}\right| > \left|\cos\frac{\theta}{2}\sin\frac{\lambda}{2} + \sin\frac{\theta}{2}\cos\frac{\lambda}{2}e^{-i\varphi}\right|$$
$$\Leftrightarrow \left|1 + \tan\frac{\lambda}{2}\tan\frac{\theta}{2}e^{-i\varphi}\right| > \left|\tan\frac{\lambda}{2} + \tan\frac{\theta}{2}e^{-i\varphi}\right|,$$

where, for the last step, we divide both sides by $\cos\frac{\lambda}{2}\cos\frac{\theta}{2}$, which is never 0 since $\lambda, \theta \in [0, \pi/2)$. Let $x := \tan\frac{\lambda}{2}$ and $y := \tan\frac{\theta}{2}$, then

$$\begin{aligned}|1 + xye^{-i\varphi}| > |x + ye^{-i\varphi}| &\Leftrightarrow |1 + xy(\cos\varphi - i\sin\varphi)| > |x + y(\cos\varphi - i\sin\varphi)|\\ &\Leftrightarrow 1 + 2xy\cos\varphi + x^2y^2 > x^2 + 2xy\cos\varphi + y^2\\ &\Leftrightarrow 1 + x^2y^2 - x^2 - y^2 > 0 \Leftrightarrow (1-x^2)(1-y^2) > 0\end{aligned}$$

and this is always verified since $x, y \in [0, 1)$ by the definition of the domains of θ and λ. $\square$

We use this lemma to generalise Theorem VII.4 to arbitrary states in the SLOCC class of the tripartite GHZ state.

THEOREM VII.7. *A state in the SLOCC class of GHZ that displays strong non-locality must be balanced. Moreover, any such strongly non-local behaviour can be witnessed using only equatorial measurements.*

PROOF. The proof of this theorem can be derived by taking advantage of the special properties of balanced states and combining them with the argument used for Theorem

VII.4. As before, we compute the amplitude $\langle \boldsymbol{\theta}, \boldsymbol{\varphi} | \psi_{\mathsf{GHZ}} \rangle$:

$$\langle \boldsymbol{\theta}, \boldsymbol{\varphi} | \psi_{\mathsf{GHZ}} \rangle = \sqrt{K} \left(\cos\delta \prod_{i=1}^{3} \langle \boldsymbol{\theta}, \boldsymbol{\varphi} | v_{\lambda_i} \rangle + \sin\delta e^{i\Phi} \prod_{i=1}^{3} \langle \boldsymbol{\theta}, \boldsymbol{\varphi} | w_{\lambda_i} \rangle \right)$$

Take $h \colon \mathsf{LM} \longrightarrow O$ as defined in the proof of Theorem VII.4 and let $g := h \sqcup h \sqcup h$. We claim that g is consistent with the empirical probabilities at all contexts that include at least a non-equatorial measurement.

Let $(\boldsymbol{\theta}, \boldsymbol{\varphi})$ be a context whose measurements are all mapped to $+1$ by g. In particular, $\theta_i \leq \frac{\pi}{2}$ for $i = 1, 2, 3$. If $\langle \boldsymbol{\theta}, \boldsymbol{\varphi} | \psi_{\mathsf{GHZ}} \rangle = 0$, then

$$\cos\delta \prod_{i=1}^{3} \langle \boldsymbol{\theta}, \boldsymbol{\varphi} | v_{\lambda_i} \rangle = -\sin\delta e^{i\Phi} \prod_{i=1}^{3} \langle \boldsymbol{\theta}, \boldsymbol{\varphi} | w_{\lambda_i} \rangle ,$$

and taking the complex modulus of both sides,

$$\cos\delta \prod_{i=1}^{3} | \langle \boldsymbol{\theta}, \boldsymbol{\varphi} | v_{\lambda_i} \rangle | = \sin\delta \prod_{i=1}^{3} | \langle \boldsymbol{\theta}, \boldsymbol{\varphi} | w_{\lambda_i} \rangle |$$

Since $\delta \in (0, \pi/4]$ we have $\cos\delta \geq \sin\delta$, with equality iff $\delta = \frac{\pi}{4}$. By Lemma VII.6, we conclude that this equation can only be satisfied if $\delta = \frac{\pi}{4}$ (i.e. the state is balanced) and $\theta_i = \frac{\pi}{2}$ for $i = 1, 2, 3$ (i.e. all the measurements are equatorial). $\square$

5.3. Further restrictions. The theorem above allows us to reduce the scope of our search for strongly non-local behaviour in the SLOCC class of GHZ to: (i) balanced states, i.e. those of the form

$$|\mathsf{B}_{\boldsymbol{\lambda},\Phi}\rangle := \sqrt{\frac{K}{2}} (|v_{\lambda_1}\rangle |v_{\lambda_2}\rangle |v_{\lambda_3}\rangle + e^{i\Phi} |w_{\lambda_1}\rangle |w_{\lambda_2}\rangle |w_{\lambda_3}\rangle),$$

determined by a tuple $\boldsymbol{\lambda} = \langle \lambda_1, \lambda_2, \lambda_3 \rangle \in \left[0, \frac{\pi}{2}\right)^3$ and a phase Φ, where $|v_\lambda\rangle$ and $|w_\lambda\rangle$ are given as in (VII.8); (ii) local equatorial measurements in the sense defined above, i.e. those with $+1$ eigenstate

$$|\varphi\rangle := \left| \frac{\pi}{2}, \varphi \right\rangle = \frac{1}{\sqrt{2}} (|0\rangle + e^{i\varphi} |1\rangle)$$

for $\varphi \in [0, 2\pi)$. Given this premise, we are interested in understanding when the amplitude function $\langle \boldsymbol{\varphi} | \mathsf{B}_{\boldsymbol{\lambda},\Phi} \rangle$ is 0. We have:

$$\begin{aligned} \langle \boldsymbol{\varphi} | \mathsf{B}_{\boldsymbol{\lambda},\Phi} \rangle = 0 &\Leftrightarrow \prod_{i=1}^{3} \langle \varphi_i | v_{\lambda_i} \rangle + e^{i\Phi} \prod_{i=1}^{3} \langle \varphi_i | w_{\lambda_i} \rangle = 0 \\ &\Leftrightarrow \prod_{i=1}^{3} \langle \varphi_i | w_{\lambda_i} \rangle = -e^{-i\Phi} \prod_{i=1}^{3} \langle \varphi_i | v_{\lambda_i} \rangle \\ &\Leftrightarrow \prod_{i=1}^{3} \langle \varphi_i | w_{\lambda_i} \rangle = -e^{-i\Phi} \prod_{i=1}^{3} e^{-i\varphi_i} \overline{\langle \varphi_i | w_{\lambda_i} \rangle} \qquad \text{(VII.9)} \end{aligned}$$

$$\Leftrightarrow \prod_{i=1}^{3} e^{i\varphi_i} \langle \varphi_i | w_{\lambda_i} \rangle \, \overline{\langle \varphi_i | w_{\lambda_i} \rangle}^{-1} = -e^{-i\Phi}$$

$$\Leftrightarrow \prod_{i=1}^{3} e^{i\varphi_i} \left(\frac{\langle \varphi_i | w_{\lambda_i} \rangle}{| \langle \varphi_i | w_{\lambda_i} \rangle |} \right)^2 = -e^{-i\Phi}$$

$$\Leftrightarrow \sum_{i=1}^{3} (\varphi_i + 2\mathrm{Arg} \langle \varphi_i | w_{\lambda_i} \rangle) = \pi - \Phi \mod 2\pi$$

where to get (VII.9) we use

$$\langle \varphi | v_\lambda \rangle = \frac{1}{\sqrt{2}} \left(\cos\frac{\lambda}{2} + \sin\frac{\lambda}{2} e^{-i\varphi} \right) = \frac{e^{-i\varphi}}{\sqrt{2}} \left(\cos\frac{\lambda}{2} e^{i\varphi} + \sin\frac{\lambda}{2} \right) = e^{-i\varphi} \overline{\langle \varphi | w_\lambda \rangle}.$$

and for the last step we take the argument of two complex numbers of norm 1. Defining

$$\beta(\lambda, \varphi) := \varphi + 2\mathrm{Arg} \langle \varphi | w_\lambda \rangle = \varphi - 2\arctan\left(\frac{\sin\frac{\lambda}{2} \sin\varphi}{\cos\frac{\lambda}{2} + \sin\frac{\lambda}{2}\cos\varphi} \right),$$

we can rewrite the condition above as

$$\text{(VII.10)} \qquad \langle \boldsymbol{\varphi} | \mathsf{B}_{\boldsymbol{\lambda},\Phi} \rangle = 0 \quad \Leftrightarrow \quad \sum_{i=1}^{3} \beta(\lambda_i, \varphi_i) = \pi - \Phi \mod 2\pi$$

Proposition VII.8. *If $\lambda_1 + \lambda_2 + \lambda_3 > \frac{\pi}{2}$, the state $|\mathsf{B}_{\boldsymbol{\lambda},0}\rangle$ does not admit strongly non-local behaviour.*

Proof. We start by showing that the map $\beta(\lambda, \varphi)$, seen as a function of φ, is strictly increasing for all $\lambda \in \left[0, \frac{\pi}{2}\right)$. To see this, it is sufficient to compute the derivative:

$$\forall \lambda \in \left[0, \frac{\pi}{2}\right), \varphi \in [0, 2\pi). \quad \frac{\partial}{\partial\varphi}\beta(\lambda, \varphi) = \frac{\cos\lambda}{1 + \cos\varphi\sin\lambda}.$$

This is strictly positive since $\cos\lambda > 0$ and $\cos\varphi \sin\lambda > -1$ since $0 \le \sin\lambda < 1$.

Now, define a function $h\colon [0, 2\pi) \longrightarrow O$ by

$$h(\varphi) := \begin{cases} +1 & \text{if } \varphi \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right] \\ -1 & \text{if } \varphi \in \left(\frac{\pi}{2}, \frac{3\pi}{2}\right] \end{cases}$$

and let $g := h \sqcup h \sqcup h$. Take a context $\boldsymbol{\varphi}$ whose measurements are assigned $+1$ by g, i.e. $\varphi_i \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right]$. Using the fact that $\beta(\lambda, -)$ is increasing, we have

$$\left| \sum_{i=1}^{3} \beta(\lambda_i, \varphi_i) \right| \le \sum_{i=1}^{3} |\beta(\lambda_i, \varphi_i)| \le \sum_{i=1}^{3} \beta\left(\lambda_i, \frac{\pi}{2}\right) = \sum_{i=1}^{3} \left(\frac{\pi}{2} - \lambda_i \right) = \frac{3\pi}{2} - \sum_{i=1}^{3} \lambda_i$$
$$< \frac{3\pi}{2} - \frac{\pi}{2} = \pi.$$

Consequently, $\sum_{i=1}^{3} \beta(\lambda_i, \varphi_i) \neq \pi \mod 2\pi$, hence by (VII.10), $\langle \boldsymbol{\varphi} | \mathsf{B}_{\boldsymbol{\lambda},0} \rangle \neq 0$ as required. $\square$

6. A family of strongly non-local three-qubit models

THEOREM VII.9. *Let $m \in \mathbb{N}_{>0}$ and $N := 2m$ an even number. Consider the tripartite measurement scenario with $X_1 = X_2 = \{0, \ldots, N-1\}$ and $X_3 = \left\{0, \frac{N}{2}\right\}$. The empirical model determined by the state $\left|\mathsf{B}_{\langle 0,0,\lambda_N\rangle,0}\right\rangle$, where $\lambda_N := \frac{\pi}{2} - \frac{\pi}{N}$, with the measurement label i at each site interpreted as the local equatorial measurement $\cos\frac{i\pi}{N}\sigma_X + \sin\frac{i\pi}{N}\sigma_Y$ (i.e. the measurement with $+1$ eigenstate $\left|\frac{\pi}{2}, i\frac{\pi}{N}\right\rangle$), is strongly non-local.*

PROOF. This proof rests on deriving, using the algebraic structure of $\mathbb{Z}_{2N}$, a (conditional) system of linear equations over $\mathbb{Z}_2$ that must be satisfied by any global assignment consistent with the possible events of the empirical model, yet does not admit any solution. This seems to be closely related to an AvN argument, but does not quite fit this setting. The reason is that the system of linear equations that a global assignment g must satisfy depends on the value that g assigns to a particular measurement. In that sense, this could be seen as a conditional version of an AvN argument.

Consider a context $\langle i, j, k\rangle \in X_1 \times X_2 \times X_3$, with $i, j \in \{0, \ldots, N-1\}$, $k \in \{0, m\}$, and a triple of outcomes $\langle a_i, b_j, c_k\rangle \in \mathbb{Z}_2^3$ for the measurements in the context.[2] From equation (VII.10), we know that measuring $\langle i, j, k\rangle$ and obtaining outcomes $\langle a_i, b_j, c_k\rangle$ has probability zero if and only if

$$\beta\left(0, i\frac{\pi}{N} + a_i\pi\right) + \beta\left(0, j\frac{\pi}{N} + b_j\pi\right) + \beta\left(\frac{\pi}{2} - \frac{\pi}{N}, k\frac{\pi}{N} + c_k\pi\right) = \pi \mod 2\pi \tag{VII.11}$$

With simple computations, we can show that $\beta(0, \varphi) = \varphi$ for all $\varphi \in [0, 2\pi)$, and that

$$\beta\left(\frac{\pi}{2} - \frac{\pi}{N}, c_0\pi\right) = c_0\pi \quad \text{and} \quad \beta\left(\frac{\pi}{2} - \frac{\pi}{N}, \frac{\pi}{2} + c_m\pi\right) = (-1)^{c_m}\frac{\pi}{N}. \tag{VII.12}$$

An arbitrary global assignment is defined by choosing outcomes for all the measurements in $X_1 \sqcup X_2 \sqcup X_3$:

$$a_0, \ldots, a_{N-1}, b_0, \ldots, b_{N-1}, c_0, c_m \in \mathbb{Z}_2.$$

By (VII.11) and (VII.12), such an assignment is consistent with the probabilities of the empirical model at every context if and only if

$$\begin{cases} i\frac{\pi}{N} + a_i\pi + j\frac{\pi}{N} + b_j\pi + c_0\pi \neq \pi & \mod 2\pi \quad \forall i, j \in \{0, \ldots, N-1\} \\ i\frac{\pi}{N} + a_i\pi + j\frac{\pi}{N} + b_j\pi + (-1)^{c_m}\frac{\pi}{N} \neq \pi & \mod 2\pi \quad \forall i, j \in \{0, \ldots, N-1\} \end{cases}$$

We will proceed to show that this system admits no solution, which implies strong non-locality. By identifying the group $\left\{k\frac{\pi}{N} \mid k \in \mathbb{Z}_{2N}\right\}$ with $\mathbb{Z}_{2N}$, we can equivalently rewrite

$$\begin{cases} i + a_iN + j + b_jN + c_0N \neq N & \mod 2N \quad \forall i, j \\ i + a_iN + j + b_jN + (-1)^{c_m} \neq N & \mod 2N \quad \forall i, j \end{cases}$$

$$\Leftrightarrow$$

$$\begin{cases} i + j + N(a_i \oplus b_j \oplus c_0) \neq N & \mod 2N \quad \forall i, j \\ i + j + (-1)^{c_m} + N(a_i \oplus b_j) \neq N & \mod 2N \quad \forall i, j \end{cases}$$

[2] For this proof, we have relabelled $+1, -1, \times$ as $0, 1, \oplus$, just like we did for the AvN arguments of Chapter VI

$$\Leftrightarrow$$

$$\begin{cases} a_i \oplus b_j \oplus c_0 = 0 & \forall i,j \text{ s.t. } i+j=0 \\ a_i \oplus b_j \oplus c_0 = 1 & \forall i,j \text{ s.t. } i+j=N \\ \\ a_i \oplus b_j = 0 & \forall i,j \text{ s.t. } i+j+(-1)^{c_m}=0 \\ a_i \oplus b_j = 1 & \forall i,j \text{ s.t. } i+j+(-1)^{c_m}=N. \end{cases}$$

$$\Leftrightarrow$$

$$\begin{cases} a_0 \oplus b_0 \oplus c_0 = 0 \\ a_i \oplus b_{N-i} \oplus c_0 = 1 & \forall i \text{ s.t. } 1 \le i \le N-1 \\ \\ a_i \oplus b_{N-i-1} = 1 & \forall i \text{ s.t. } 0 \le i \le N-1 \qquad \text{if } c_m = 0 \\ \\ a_0 \oplus b_1 = 0 \\ a_1 \oplus b_0 = 0 & \qquad\qquad\qquad\qquad\qquad\quad \text{if } c_m = 1 \\ a_i \oplus b_{N+1-i} = 1 & \forall i \text{ s.t. } 2 \le i \le N-1 \end{cases}$$

Since $N = 2m$ is even, if we sum all the N equations from the first two lines we obtain

$$\bigoplus_{i=0}^{N-1} a_i \oplus \bigoplus_{j=0}^{N-1} b_j = 1.$$

On the other hand, if we sum any of the other two groups of N equations we get

$$\bigoplus_{i=0}^{N-1} a_i \oplus \bigoplus_{j=0}^{N-1} b_j = 0,$$

showing that the system is unsatisfiable regardless of whether $c_m = 0$ or $c_m = 1$. □

This new family of strongly non-local three-qubit systems is tightly connected to a construction on two-qubit states due to Barrett, Kent, and Pironio [**BKP06**]. In particular, our empirical models restricted to the first two parties coincide, up to a rotation of the equatorial measurements, to those used in [**BKP06**]. The local fraction of these bipartite empirical models tends to zero as the number of measurements increases, but obviously none of them are strongly non-local. Despite the lack of strong non-locality in the bipartite systems constructed in [**BKP06**], we show that it is possible to witness strongly non-local behaviour with a finite amount of measurements by adding a third qubit with some entanglement, and only two local measurements – Pauli X and Y – available on it. An interesting aspect is that there is a trade-off between the number of measuring settings available on the first two qubits and the amount of entanglement between the third qubit and the system comprised of the other two.

We illustrate this by computing the bipartite von Neumann entanglement entropy between the first two qubits and the third, i.e. the von Neumann entropy of the reduced state of $\left|\mathsf{B}_{\langle 0,0,\lambda\rangle,0}\right\rangle$ corresponding to the third qubit, as a function of λ. Let ρ_{ABC} denote the density matrix of $|\mathsf{B}_{\langle 0,0,\lambda\rangle,0}\rangle$. The reduced density matrix corresponding to the third qubit is

$$\begin{aligned}\rho_C(\lambda) &= \mathrm{Tr}_{AB}[\rho_{ABC}] = \langle 00|_{AB}\rho_{ABC}|00\rangle_{AB} + \langle 11|_{AB}\rho_{ABC}|11\rangle_{AB} \\ &= \frac{1}{2}\begin{pmatrix} 1 & 2\cos\frac{\lambda}{2}\sin\frac{\lambda}{2} \\ 2\cos\frac{\lambda}{2}\sin\frac{\lambda}{2} & 1\end{pmatrix}.\end{aligned}$$

The eigenvalues of $\rho_C(\lambda)$ are $\epsilon_\pm(\lambda) := \frac{1}{2}(1 \pm \sin\lambda)$. Hence, by rewriting $\rho_C(\lambda)$ in its eigenbasis, we can easily compute the von Neumann entropy S_C as a function of λ:

$$S_C(\lambda) := -\,\mathrm{Tr}\left[\rho_C(\lambda)\log_2\rho_C(\lambda)\right] = -\epsilon_+(\lambda)\log_2\epsilon_+(\lambda) - \epsilon_-(\lambda)\log_2\epsilon_-(\lambda)$$

The plot of the function $S_C(\lambda)$ is shown in Figure VII.5. Notice that the entangle-

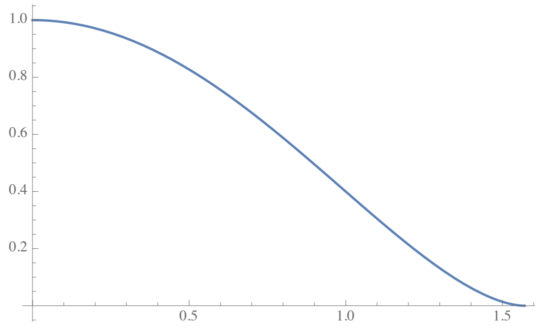


FIGURE VII.5. Von Neumann entanglement entropy between the third qubit of $|\mathsf{B}_{\langle 0,0,\lambda\rangle,0}\rangle$ and the other two as a function of λ.

ment entropy is maximal, i.e. equal to 1, when $N = 2$, in which case $\lambda_2 = 0$ and so $|\mathsf{B}_{\langle 0,0,\lambda_2\rangle,0}\rangle = |\mathsf{GHZ}\rangle$. This corresponds to the usual GHSZ argument with Pauli measurements X,Y for each qubit. On the other hand, $S(\lambda)$ becomes arbitrarily small as $N \to \infty$, when $\lambda_N \to \frac{\pi}{2}$ and $|\mathsf{B}_{\langle 0,0,\lambda_N\rangle,0}\rangle$ approaches the state $|\Phi^+\rangle \otimes |+\rangle$, which has no entanglement between the first two qubits and the third.

Discussion

Our analysis of strong non-locality for three-qubit systems has been quite extensive. We shall discuss a number of directions for further research.

(1) First, it remains to complete our classification of all instances of three-qubit strong non-locality.
(2) The family of strongly non-local models introduced in Section 6 does not fit the framework of AvN arguments exactly. Nevertheless, our proof of strong non-locality does make essential use of the algebraic structure of $\mathbb{Z}_{2N}$ (or the circle group), in what amounts to a conditional version of an AvN argument. One may wonder whether a similar property will hold for all instances of three-qubit strong non-locality.
(3) This family also highlights an inter-relationship between non-locality, entanglement and the number of measurements available, and raises the question of whether this is an instance of a more general relationship.

(4) Finally, while the present results provide necessary conditions for strong non-locality in three-qubit states, the more general question of characterising strong non-locality of n-qubit states, where little is known about SLOCC classes, remains open.

CHAPTER VIII

Conclusion

In this thesis, we have aimed to develop a solid understanding of the mathematical structure of logical forms of contextuality, both in quantum physics and in more general terms. This analysis has been conducted following four different research paths: the study of non-locality and contextuality as a topological property (Chaptets III and IV), the search for instances of contextual behaviour outside of quantum mechanics (Chapter V), the development of new techniques to detect contextuality (Chapters III-V), and the study of logical forms of contextuality exhibited by multi-qubit states (Chapters VI and VII). Our work in each of these areas has led to the establishment of interesting new results and novel perspectives.

The outcomes of Chapter III, despite their restrictive trait, represent an important step forward in understanding the cohomology theory developed in [**AMB12, ABK$^+$15**]. The investigation of the structure of false negatives carried out here is the key element leading to the refinement introduced later, in Chapter IV. This study has brought in its wake new developments of the sheaf cohomology framework, such as the hierarchy of higher cohomology obstructions and the reinterpretation of the usual obstructions as $\mathcal{F}$-torsors. Although these results are not immediately useful for the purposes of this thesis, they offer new theoretical insights on the general question of extending local sections of a presheaf to global ones. In particular, we identify the following possible research directions that may be worth exploring:

- Although higher obstructions cannot be used to study contextuality, they do carry information concerning the *local* extendability of sections. It would be interesting to determine whether these can be used to study *signalling* instead. In particular, one might hope to infer new structural results on models that violate the no-signalling condition. Examples of such models are considered, for instance, in the "contextuality by default" framework [**DKC16**], which is particularly suited to model contextual behaviour in psychology experiments.
- More generally, it would be interesting to determine whether higher obstructions may be used to characterise the extendability of local sections of arbitrary presheaves. Since higher cohomology groups play a substantial role in general obstruction theory [**Ste51, EM54**], which concerns the extendability of continuous maps in topology, it seems reasonable to envision applications of the hierarchy presented in this thesis to study the extendability properties of any given presheaf.
- The new perspective on cohomology obstructions provided by $\mathcal{F}$-torsors is still to be explored. This formalism presents intriguing connections with gauge theory, which seems to be the natural framework to formally capture the bundle

diagram description of empirical models. More specifically, one might hope to reach a compelling definition of empirical models as principal bundles, and apply alternative tools of topology, such as holonomy theory, to characterise contextuality.

In Chapter IV, we have presented a complete cohomology invariant for non-locality and contextuality, showing that, in most cases, these phenomena can be fully characterised using topological tools. We proved that the invariant works for a very large class of empirical models, which includes all the known instances of false negatives. However, there is still considerable margin for improvement. We believe future research efforts should be directed towards completing the following tasks:

- Proving Conjecture IV.41. This is certainly the main open question left by Chapter IV. A crucial step to achieve this result is understanding the cyclic contextuality property (CCP). A thorough analysis of this property would either rule out the existence of any model violating it, or lead either to the key to the proof of the conjecture, or to a counterexample.
- Developing computational methods based on sheaf cohomology to detect contextuality. Computational algebraic topology has been increasingly used in various fields of science, where sheaf cohomology and particularly persistent (co)homology [**ELZ02, CZCG05**] have found remarkable applications to areas as diverse as image analysis [**BEK10, CIdSZ08**], signal analysis [**PH15**], fractal geometry [**MS12**], viral evolution [**CCR13**] and bacteria classification [**OD16**]. It is thus natural to seek to develop new topological algorithms, based on the cohomology invariant introduced in this thesis, to search for contextual behaviour in large datasets.

In Chapter V, we have introduced a general definition of contextuality using the flexible language of valuation algebras, which allows to conveniently capture its nature as a fundamental gap between local agreement and global disagreement in information sources, in analogy with the topological interpretation of local consistency vs global inconsistency. This new formulation has provided a common theory for the examples of contextual behaviour outside quantum mechanics observed in previous work [**Abr13a, AGK13, AS14**], and thus constitutes a promising attempt to develop a general theory of contextual semantics. Moreover, it provides inspiration for the potential establishment of numerous other examples of contextuality beyond quantum theory, which may arise in any of the countless domains whose essence can be captured by valuation algebras. This potential remains largely unexplored, and shall be investigated in future research. In addition, we plan to take advantage of the formalism to translate theorems and results about disagreement across different valuation algebras. Our goal is both to develop new techniques for the study of contextual behaviour, as exemplified e.g. by the algorithms presented in this thesis, and to find new applications of the current contextuality-detection methods. For instance, consider Example V.18, where we showed that the unsolvability of a graph colouring problem is mathematically equivalent to contextuality. This connection could be exploited to translate general results on graph colourability to contextuality: to mention an example, although hardly relevant, the four-colour theorem can be rephrased as a statement about the contextuality of certain models defined over scenarios whose simplicial complex representation is a non-planar graph. On the other hand, the same connection could lead to the development of a cohomology theory to study graph colourability, following the methods introduced in Chapter IV.

The new algorithms for the detection of logical forms of contextuality developed at the end of Chapter V are a prime example of the usefulness of the valuation algebraic framework. By translating the problem of recognising contextual behaviour into an inference problem, we have been able to take advantage of the theory of generic inference to construct new methods to detect logical and strong contextuality. We also proved that these algorithms have better worst-case complexity than the current state of the art for the detection of non-locality in (n, k, l) scenarios. Unfortunately, due to time constraints, we have not been able to implement them in a programming language. This will be one of the prime goals of our future research on the subject.

After having explored logical forms of contextuality from a highly general and abstract viewpoint, the two last chapters of the thesis focus on contextuality as a purely quantum phenomenon. The results presented here offer new perspectives and potential resources for quantum computation, with a large number of new examples of multi-qubit states exhibiting strong contextuality. In Chapter VI, we introduced a complete characterisation of All-vs-Nothing arguments for stabiliser states, which is expressed through the combinatorial concept of AvN triple. This has allowed to describe a general method to identify all instances of All-vs-Nothing contextuality for stabiliser states on any number of qubits. This result may be of particular interest for the development of measurement based quantum computation, which makes use of strongly contextual stabiliser states as fundamental resources to achieve faster computation. Moreover, it has led to the establishment of a new structural result, namely that any AvN argument can essentially be reduced to Mermin's original proof of the strong contextuality of the GHZ state. We identify two possible improvements of the results presented here:

- Extending the characterisation from maximal stabiliser subgroups – which correspond to stabiliser states – to arbitrary stabiliser subgroups. We believe this could be achieved by taking advantage of graph codes [**SW01a**], a natural generalisation of the graph state formalism used to prove our main result.
- Implementing efficient algorithms to solve the problem of generating AvN triples. The Mathematica code we use in [**Car16**] is powerful enough to generate examples up to 5 qubits in a reasonable time. However, if we plan to find examples for a higher number of qubits, faster algorithms need to be introduced.

Finally, in Chapter VII, we determine the minimum quantum resources needed to achieve strong non-locality, and provide a partial characterisation of the three-qubit states that exhibit this peculiar kind of contextuality. We show that strong non-locality cannot be realised by any two-qubit system with any finite number of local measurements, and prove that for 3-qubit systems, it can only arise in models involving states in the GHZ SLOCC class and equatorial measurements, along with additional minor restrictions. Within this class of models, we identify a new infinite family of models displaying strong non-locality using states that are not LU-equivalent to the GHZ state. An interesting aspect of this class of states is that there is a trade-off between the number of measurement settings available on the first two qubits and the amount of entanglement between the third qubit and the first two. Let us list some possible directions for future research:

- First and foremost, the classification of all three-qubit states exhibiting strong non-locality remains to be completed.
- The similarities between the family of strongly non-local states introduced in this chapter and the scenarios considered in [**BKP06**] suggests a possible application for experimental tests of contextual behaviour. In [**BKP06**], it is shown that the contextual fraction of the empirical model obtained by applying certain single-qubit measurements to the Bell state can be pushed arbitrarly close to 1 by increasing the number of measurements allowed for each party. This means that experimental bounds for non-locality can be improved by allowing more measurements to be performed. Our family of strongly non-local states suggests that a similar experimental programme can be implemented by introducing a third qubit with some entanglement with the Bell state instead.
- More generally, it would be interesting to understand the precise relation between non-locality, entanglement and the number of measurements allowed at each site.
- Finally, it remains to extend the search for strongly non-local behaviour beyond three-qubit states. This task seems particularly ambitious, since none of the techniques implemented in this theory can be naturally generalised to any higher number of qubits due to a fundamental lack of knowledge about their infinitely many SLOCC classes.

Bibliography

[Aas18] Sivert Aasnaess, *Comparing two cohomological obstructions for contextuality*, Master's thesis, University of Oxford, 2018.

[AB09] Janet Anders and Dan E Browne, *Computational power of correlations*, Physical Review Letters **102** (2009), no. 5, 050502.

[AB11a] Samson Abramsky and Adam Brandenburger, *The sheaf-theoretic structure of non-locality and contextuality*, New Journal of Physics **13** (2011), no. 11, 113036, Eprint available at arXiv:1102.0264 [quant-ph].

[AB11b] ———, *A unified sheaf-theoretic account of non-locality and contextuality*, arXiv preprint arXiv:1701.00242 [quant-ph] (2011).

[AB14] ———, *An operational interpretation of negative probabilities and no-signalling models*, Horizons of the mind: A tribute to Prakash Panangaden (Franck van Breugel, Elham Kashefi, Catuscia Palamidessi, and Jan Rutten, eds.), Lecture Notes in Computer Science, vol. 8464, Springer International Publishing, 2014, pp. 59–75.

[ABC⁺17] Samson Abramsky, Rui Soares Barbosa, Giovanni Carù, Nadish de Silva, Kohei Kishida, and Shane Mansfield, *Minimum quantum resources for strong non-locality*, 12th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2017) (Dagstuhl, Germany), Leibniz International Proceedings in Informatics (LIPIcs), Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2017.

[ABCP17] 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 of London A: Mathematical, Physical and Engineering Sciences **375** (2017), no. 2106.

[ABdSZ17] Samson Abramsky, Rui Soares Barbosa, Nadish de Silva, and Octavio Zapata, *The quantum monad on relational structures*, arXiv:1705.07310 [quant-ph], 2017.

[ABK⁺15] Samson Abramsky, Rui Soares Barbosa, Kohei Kishida, Raymond Lal, and Shane Mansfield, *Contextuality, cohomology and paradox*, 24th EACSL Annual Conference on Computer Science Logic (CSL 2015) (Dagstuhl, Germany) (Stephan Kreutzer, ed.), Leibniz International Proceedings in Informatics (LIPIcs), vol. 41, Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2015, pp. 211–228.

[ABK⁺16] 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** (2016), 58 – 65, Foundations of Probability Theory in Psychology and Beyond.

[ABM17] Samson Abramsky, Rui Soares Barbosa, and Shane Mansfield, *Contextual fraction as a measure of contextuality*, Physical Review Letters **119** (2017), 050504.

[Abr13a] Samson Abramsky, *Relational databases and Bell's theorem*, In search of elegance in the theory and practice of computation: Essays dedicated to Peter Buneman (Val Tannen, Limsoon Wong, Leonid Libkin, Wenfei Fan, Wang-Chiew Tan, and Michael Fourman, eds.), Lecture Notes in Computer Science, vol. 8000, Springer Berlin Heidelberg, 2013, Eprint available at arXiv:1208.6416 [cs.LO], pp. 13–35.

[Abr13b] ———, *Relational hidden variables and non-locality*, Studia Logica **101** (2013), no. 2, 411–452, Eprint available at arXiv:1007.2754 [quant-ph].

[Abr14a] ———, *All-versus-Nothing arguments, stabiliser subgroups, and the "AvN triple" conjecture*, Talk at 10 years of categorical quantum mechanics workshop, University of Oxford, October 2014.

[Abr14b] ———, *Contextual semantics: From quantum mechanics to logic, databases, constraints, and complexity*, Bulletin of the EATCS **113** (2014), 137–163, Eprint available at arXiv:1406.7386 [quant-ph].

[AC08] Samson Abramsky and Bob Coecke, *Categorical quantum mechanics*, Handbook of Quantum Logic and Quantum Structures: Quantum Logic (Kurt Engesser, Dov M Gabbay, and Daniel Lehmann, eds.), Elsevier, 2008, pp. 261–324.

[AC14] Samson Abramsky and Carmen Constantin, *A classification of multipartite states by degree of non-locality*, Proceedings 10th International Workshop on Quantum Physics and Logic, Castelldefels (Barcelona), Spain, 17th to 19th July 2013 (Bob Coecke and Matty Hoban, eds.), Electronic Proceedings in Theoretical Computer Science, vol. 171, Open Publishing Association, 2014, Eprint available at arXiv:1412.5213 [quant-ph], pp. 10–25.

[AC16] Samson Abramsky and Giovanni Carú, *Identifying all-vs-nothing arguments in stabiliser theory*, Informal pre-proceedings of the 13th International Conference on Quantum Physics and Logic (QPL 2016), June 2016.

[AC19] Samson Abramsky and Giovanni Carù, *Non-locality, contextuality and valuation algebras: a general theory of disagreement*, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences (2019).

[ADR82] Alain Aspect, Jean Dalibard, and Gérard Roger, *Experimental test of Bell's inequalities using time-varying analyzers*, Physical Review Letters **49** (1982), no. 25, 1804–1807.

[AFLS15] Antonio Acín, Tobias Fritz, Anthony Leverrier, and Ana Belén Sainz, *A combinatorial approach to nonlocality and contextuality*, Communications in Mathematical Physics **334** (2015), no. 2, 533–628, Eprint available at arXiv:1212.4084 [quant-ph].

[AGK13] Samson Abramsky, Georg Gottlob, and Phokion G. Kolaitis, *Robust constraint satisfaction and local hidden variables in quantum mechanics*, Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence (Francesca Rossi, ed.), AAAI Press, 2013, pp. 440–446.

[AH12] Samson Abramsky and Lucien Hardy, *Logical Bell inequalities*, Physical Review A **85** (2012), no. 6, 062114, Eprint available at arXiv:1203.1352 [quant-ph].

[AK91] Russell Almond and Augustine Kong, *Optimality issues in constructing a markov tree from graphical models*, Res. Rep. A-3, Harvard Univ., Dept. of Stat (1991).

[AK17] Samson Abramsky and Simon Kochen, Private communication during the conference *Contextuality: Conceptual Issues, Operational Signatures, and Applications*, Perimeter institute, July 2017.

[AMB12] Samson Abramsky, Shane Mansfield, and Rui Soares Barbosa, *The cohomology of non-locality and contextuality*, Proceedings 8th International Workshop on Quantum Physics and Logic, Nijmegen, Netherlands, October 27–29, 2011 (Bart Jacobs, Peter Selinger, and Bas Spitters, eds.), Electronic Proceedings in Theoretical Computer Science, vol. 95, Open Publishing Association, 2012, Eprint available at arXiv:1111.3620 [quant-ph], pp. 1–14.

[Arn85] Stefan Arnborg, *Efficient algorithms for combinatorial problems on graphs with bounded decomposability — a survey*, BIT Numerical Mathematics **25** (1985), no. 1, 1–23.

[AS14] Samson Abramsky and Mehrnoosh Sadrzadeh, *Semantic unification: A sheaf theoretic approach to natural language*, Categories and types in logic, language, and physics: Essays Dedicated to Jim Lambek on the occasion of his 90th birthday (Claudia Casadio, Bob Coecke, Michael Moortgat, and Philip Scott, eds.), Lecture Notes in Computer Science, vol. 8222, Springer, 2014, Eprint available at arXiv:1403.3351 [cs.CL], pp. 1–13.

[Bae09] J. Baez, *Torsors made easy*, 2009, available at `http://math.ucr.edu/home/baez/torsors.html`.

[Bar14] Rui Soares Barbosa, *On monogamy of non-locality and macroscopic averages: examples and preliminary results*, Proceedings 11th workshop on Quantum Physics and Logic, Kyoto, Japan, 4-6th June 2014 (Bob Coecke, Ichiro Hasuo, and Prakash Panangaden, eds.), Electronic Proceedings in Theoretical Computer Science, vol. 172, Open Publishing Association, 2014, Eprint available at arXiv:1412.8541 [quant-ph], pp. 36–55.

[Bar15a] ———, *Contextuality in quantum mechanics and beyond*, Ph.D. thesis, University of Oxford, 2015.

[Bar15b] ———, *Note on the bundle perspective*, Unpublished notes, University of Oxford, 2015.

[BB72] Umberto Bertele and Francesco Brioschi, *Nonserial dynamic programming*, Academic Press, 1972.

[BB14] Charles H Bennett and Gilles Brassard, *Quantum cryptography: public key distribution and coin tossing.*, Theor. Comput. Sci. **560** (2014), no. 12, 7–11.

[BBD+09] H. J. Briegel, D. E. Browne, W. Dür, R. Raussendorf, and M. Van den Nest, *Measurement-based quantum computation*, Nature Physics **5** (2009), 19 EP –.

[BBPS96] Charles H. Bennett, Herbert J. Bernstein, Sandu Popescu, and Benjamin Schumacher, *Concentrating partial entanglement by local operations*, Physical Review A **53** (1996), 2046–2052.

[BEK10] P. Bendich, H. Edelsbrunner, and M. Kerber, *Computing robustness and persistence for images*, IEEE Transactions on Visualization and Computer Graphics **16** (2010), no. 6, 1251–1260.

[Bel64] John S. Bell, *On the Einstein-Podolsky-Rosen paradox*, Physics **1** (1964), no. 3, 195–200.

[Bel66] ———, *On the problem of hidden variables in quantum mechanics*, Reviews of Modern Physics **38** (1966), no. 3, 447–452.

[Bel87] ———, *Speakable and unspeakable in quantum mechanics: Collected papers on quantum philosophy*, Cambridge University Press, 1987.

[Ber84] Michael Victor Berry, *Quantal phase factors accompanying adiabatic changes*, Proc. R. Soc. Lond. A **392** (1984), no. 1802, 45–57.

[BF28] Max Born and Vladimir Fock, *Beweis des adiabatensatzes*, Zeitschrift für Physik **51** (1928), no. 3-4, 165–180.

[BFM+81] Catriel Beeri, Ronald Fagin, David Maier, Alberto Mendelzon, Jeffrey Ullman, and Mihalis Yannakakis, *Properties of acyclic database schemes*, STOC '81: Proceedings of the thirteenth annual ACM Symposium on Theory of Computing, STOC'81, ACM, 1981, pp. 355–362.

[BFMY83] Catriel Beeri, Ronald Fagin, David Maier, and Mihalis Yannakakis, *On the desirability of acyclic database schemes*, Journal of the ACM **30** (1983), no. 3, 479–513.

[BK05] Sergey Bravyi and Alexei Kitaev, *Universal quantum computation with ideal clifford gates and noisy ancillas*, Phys. Rev. A **71** (2005), 022316.

[BKP06] Jonathan Barrett, Adrian Kent, and Stefano Pironio, *Maximally nonlocal and monogamous quantum correlations*, Physical Review Letters **97** (2006), no. 17, 170409, Eprint available at arXiv:quant-ph/0605182.

[BMT05] Gilles Brassard, André Allan Méthot, and Alain Tapp, *Minimum entangled state dimension required for pseudo-telepathy*, Quantum Information & Computation **5** (2005), no. 4, 275–284.

[BO18] Kerstin Beer and Tobias J Osborne, *Contextuality and bundle diagrams*, 2018, Eprint available at arXiv:1802.08424 [quant-ph].

[Bor51] Armand Borel, *Cohomologie des espaces homogènes*, Séminaire Bourbaki **1** (1948-1951).

[Bou93] André Bouchet, *Recognizing locally equivalent graphs*, Discrete Mathematics **114** (1993), no. 1, 75 – 86.

[BPR+00] Charles H. Bennett, Sandu Popescu, Daniel Rohrlich, John A. Smolin, and Ashish V. Thapliyal, *Exact and asymptotic measures of multipartite pure-state entanglement*, Physical Review A **63** (2000), 012307.

[Bus82] PJ Bussey, *"super-luminal communication" in einstein-podolsky-rosen experiments*, Physics Letters A **90** (1982), no. 1-2, 9–12.

[CAB12] Earl T. Campbell, Hussain Anwar, and Dan E. Browne, *Magic-state distillation in all prime dimensions using quantum reed-muller codes*, Phys. Rev. X **2** (2012), 041021.

[Car26] Élie Cartan, *Les groupes d'holonomie des espaces généralisés*, Acta mathematica **48** (1926), no. 1-2, 1–42.

[Car49] Henri Cartan, *Algebraic topology*, Harvard University, 1949.

[Car15] Giovanni Carù, *Detecting contextuality: sheaf cohomology and all-vs-nothing arguments*, Master's thesis, University of Oxford, 2015.

[Car16] Giovanni Carù, *Generating AvN triples*, Mathematica notebook, 2016, Available at Github.

[Car17] Giovanni Carù, *On the cohomology of contextuality*, Proceedings 13th International Conference on Quantum Physics and Logic, Glasgow, Scotland, 6-10 June 2016 (Ross Duncan and Chris Heunen, eds.), Electronic Proceedings in Theoretical Computer Science, vol. 236, Open Publishing Association, 2017, Eprint available at arXiv:1701.00242 [quant-ph], pp. 21–39.

[Car18] Giovanni Carú, *Towards a complete cohomology invariant for non-locality and contextuality*, 2018, Eprint available at arXiv:1807.04203 [quant-ph].

[Cav14] Carlton Caves, *Stabilizer formalism for qubits*, Internal report, September 2014.

[CCR13] Joseph Minhow Chan, Gunnar Carlsson, and Raul Rabadan, *Topology of viral evolution*, Proceedings of the National Academy of Sciences **110** (2013), no. 46, 18566–18571.

[Châ44] François Châtelet, *Variations sur un theme de H. Poincaré*, Annales scientifiques de l'École Normale Supérieure, vol. 61, Elsevier, 1944, pp. 249–300.

[Châ46] ———, *Méthode galoisienne et courbes de genre 1*, Annales de l'Université de Lyon, Section A, vol. 89, 1946, pp. 40–49.

[CHSH69] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt, *Proposed experiment to test local hidden-variable theories*, Physical Review Letters **23** (1969), no. 15, 880–884.

[CIdSZ08] Gunnar Carlsson, Tigran Ishkhanov, Vin de Silva, and Afra Zomorodian, *On the local behavior of spaces of natural images*, International Journal of Computer Vision **76** (2008), no. 1, 1–12.

[CLS18] Michael B Cohen, Yin Tat Lee, and Zhao Song, *Solving linear programs in the current matrix multiplication time*, 2018, Eprint available at arXiv:1810.07896 [quant-ph].

[CM94] Andrés Cano and Serafín Moral, *Heuristic algorithms for the triangulation of graphs*, International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Springer, 1994, pp. 98–107.

[Con81] A Connes, *Spectral sequence and homology of currents for operator algebras*, Math. Forschungsinst. Oberwolfach Tagungsber **41** (1981), no. 81, 27–9.

[Con83] Alain Connes, *Cohomologie cyclique et foncteurs Extn*, CR Acad. Sci. Paris Sér. I Math **296** (1983), no. 23, 953–958.

[Con85] ———, *Non-commutative differential geometry*, Publications Mathématiques de l'Institut des Hautes Études Scientifiques **62** (1985), no. 1, 41–144.

[Coo71] Stephen A. Cook, *The complexity of theorem-proving procedures*, Proceedings of the Third Annual ACM Symposium on Theory of Computing (New York, NY, USA), STOC '71, ACM, 1971, pp. 151–158.

[Cou95] Pierre Cousin, *Sur les fonctions de n variables complexes*, Acta Math. **19** (1895), 1–61.

[CS48] Henri Cartan and Jean-Pierre Serre, *Applications de la théorie générale à divers problèmes globaux*, Séminaire Henri Cartan **4** (1948), no. 1951-1952, 1951–1952.

[CSW14] Adán Cabello, Simone Severini, and Andreas Winter, *Graph-theoretic approach to quantum correlations*, Phys. Rev. Lett. **112** (2014), 040401.

[Cun19] Marcelo Terra Cunha, *On measures and measurements: a fibre bundle approach to contextuality*, arXiv:1903.08819 [quant-ph], 2019.

[CW90] Don Coppersmith and Shmuel Winograd, *Matrix multiplication via arithmetic progressions*, Journal of symbolic computation **9** (1990), no. 3, 251–280.

[CZCG05] Gunnar Carlsson, Afra Zomorodian, Anne Collins, and Leonidas J. Guibas, *Persistent barcodes for shapes*, International Journal of Shape Modeling **11** (2005), no. 02, 149–187.

[Dem08] Arthur P Dempster, *Upper and lower probabilities induced by a multivalued mapping*, Classic Works of the Dempster-Shafer Theory of Belief Functions, Springer, 2008, pp. 57–72.

[DH81] David E Daykin and Roland Häggkvist, *Degrees giving independent edges in a hypergraph*, Bulletin of the Australian Mathematical Society **23** (1981), no. 1, 103–109.

[Dir42] Paul A. M. Dirac, *Bakerian lecture: The physical interpretation of quantum mechanics*, Proceedings of the Royal Society of London, Series A: Mathematical and Physical Sciences **180** (1942), no. 980, 1–40.

[DK84] WG Dwyer and DM Kan, *An obstruction theory for diagrams of simplicial sets*, Nederl. Akad. Wetensch. Indag. Math **46** (1984), no. 2, 139–146.

[DKC16] Ehtibar N. Dzhafarov, Janne V. Kujala, and Victor H. Cervantes, *Contextuality-by-default: A brief overview of ideas, concepts, and terminology*, Quantum Interaction (Cham) (Harald Atmanspacher, Thomas Filk, and Emmanuel Pothos, eds.), Springer International Publishing, 2016, pp. 12–23.

[DKK16] Vedran Dunjko, Theodoros Kapourniotis, and Elham Kashefi, *Quantum-enhanced secure delegated classical computing*, Quantum Information & Computation **61** (2016), no. 1, 61–86.

[DP89] Rina Dechter and Judea Pearl, *Tree clustering for constraint networks*, Artificial Intelligence **38** (1989), no. 3, 353–366.

[dS17] Nadish de Silva, *Logical paradoxes in quantum computation*, 2017, Eprint available at arXiv:1709.00013 [quant-ph].

[Dus75] John Williford Duskin, *Simplicial methods and the interpretation of" triple" cohomology*, vol. 163, American Mathematical Soc., 1975.

[DVC00] W. Dür, G. Vidal, and J. I. Cirac, *Three qubits can be entangled in two inequivalent ways*, Phys. Rev. A **62** (2000), 062314.

[Eke91] Artur K. Ekert, *Quantum cryptography based on Bell's theorem*, Phys. Rev. Lett. **67** (1991), 661–663.

[EKMS92] M. Erne, J. Koslowski, A. Melton, and G.E. Strecker, *A primer on Galois connections*, York Academy of Science, 1992.

[ELZ02] Edelsbrunner, Letscher, and Zomorodian, *Topological persistence and simplification*, Discrete & Computational Geometry **28** (2002), no. 4, 511–533.

[EM54] Samuel Eilenberg and Saunders MacLane, *On the groups $h(\pi, n)$, iii: Operations and obstructions*, Annals of Mathematics (1954), 513–557.

[EPR35] Albert Einstein, Boris Podolsky, and Nathan Rosen, *Can quantum-mechanical description of physical reality be considered complete?*, Physical Review **47** (1935), no. 10, 777–780.

[Fag83] Ronald Fagin, *Acyclic database schemes (of various degrees): A painless introduction*, Colloquium on Trees in Algebra and Programming, Springer, 1983, pp. 65–89.

[FB94] D. J. Foulis and M. K. Bennett, *Effect algebras and unsharp quantum logics*, Foundations of Physics **24** (1994), no. 10, 1331–1352.

[Fel68] William Feller, *An introduction to probability theory and its applications: volume I*, John Wiley & Sons New York, 1968.

[Fey87] Richard P. Feynman, *Negative probability*, Quantum implications: Essays in honour of David Bohm (Basil J. Hiley and F. David Peat, eds.), Methuen, 1987, pp. 235–248.

[Fin82] Arthur Fine, *Hidden variables, joint probability, and the Bell inequalities*, Physical Review Letters **48** (1982), no. 5, 291–295.

[FMU82] Ronald Fagin, Alberto O. Mendelzon, and Jeffrey D. Ullman, *A simplified universal relation assumption and its properties*, ACM Transactions on Database Systems **7** (1982), 343–360.

[FW98] David V Feldman and Alexander Wilce, *Abelian extensions of quantum logics*, International journal of theoretical physics **37** (1998), no. 1, 39–43.

[GC99] Daniel Gottesman and Isaac L. Chuang, *Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations*, Nature **402** (1999), no. 6760, 390–393.

[GE07] D. Gross and J. Eisert, *Novel schemes for measurement-based quantum computation*, Phys. Rev. Lett. **98** (2007), 220503.

[GHSZ90] Daniel M. Greenberger, Michael A. Horne, Abner Shimony, and Anton Zeilinger, *Bell's theorem without inequalities*, American Journal of Physics **58** (1990), no. 12, 1131–1143.

[GHZ89] Daniel M. Greenberger, Michael A. Horne, and Anton Zeilinger, *Going beyond Bell's theorem*, Bell's theorem, quantum theory, and conceptions of the universe (M. Kafatos, ed.), Fundamental Theories of Physics, vol. 37, Kluwer, 1989, pp. 69–72.

[GKR02] Markus Grassl, Andreas Klappenecker, and Martin Rötteler, *Graphs, quadratic forms, and quantum codes*, Proceedings IEEE International Symposium on Information Theory (ISIT 2002), 2002, p. 45.

[God58] Roger Godement, *Topologie algébrique et théorie des faisceaux*, Hermann, 1958.

[Got97] Daniel Gottesman, *Stabilizer codes and quantum error correction*, 1997.

[Gro96] Lov K. Grover, *A fast quantum mechanical algorithm for database search*, Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing (New York, NY, USA), STOC '96, ACM, 1996, pp. 212–219.

[GRW80] Giancarlo Ghirardi, Alberto Rimini, and Tullio Weber, *A general argument against superluminal transmission through the quantum mechanical measurement process*, Lettere al Nuovo Cimento Series 2 **27** (1980), no. 10, 293–298.

[GTHB05] Otfried Gühne, Géza Tóth, Philipp Hyllus, and Hans J. Briegel, *Bell inequalities for graph states*, Physical Review Letters **95** (2005), no. 12, 120405.

[Hae04] Rolf Haenni, *Ordered valuation algebras: a generic framework for approximating inference*, International Journal of Approximate Reasoning **37** (2004), no. 1, 1–41.

[Har92] Lucien Hardy, *Quantum mechanics, local realistic theories, and Lorentz-invariant realistic theories*, Physical Review Letters **68** (1992), no. 20, 2981–2984.

[Har93] ———, *Nonlocality for two particles without inequalities for almost all entangled states*, Physical Review Letters **71** (1993), no. 11, 1665–1668.

[HDE+06] Marc Hein, Wolfgang Dür, Jens Eisert, Robert Raussendorf, Maarten Van den Nest, and Hans J Briegel, *Entanglement in graph states and its applications*, preprint arXiv:quant-ph/0602096, 2006.

[Hei27] W Heisenberg, *W. heisenberg, z. phys. 43, 172 (1927).*, Z. Phys. **43** (1927), 172.

[HL99] Rolf Haenni and Norbert Lehmann, *Efficient hypertree construction*, Institute of Informatics University of Fribourg, 1999.

[HN60] Frank Harary and Robert Z Norman, *Some properties of line digraphs*, Rendiconti del Circolo Matematico di Palermo **9** (1960), no. 2, 161–168.

[HR83] Peter Heywood and Michael L. G. Redhead, *Nonlocality and the Kochen-Specker paradox*, Foundations of Physics **13** (1983), no. 5, 481–499.

[HWVE14] Mark Howard, Joel Wallman, Victor Veitch, and Joseph Emerson, *Contextuality supplies the 'magic' for quantum computation*, Nature **510** (2014), no. 7505, 351–355.

[Jor83] Thomas F Jordan, *Quantum correlations do not transmit signals*, Physics Letters A **94** (1983), no. 6-7, 264.

[Ken95] John Bernard Kennedy, *On the empirical foundations of the quantum no-signalling proofs*, Philosophy of Science **62** (1995), no. 4, 543–560.

[KHM99] J Kohlas, R Haenni, and S Moral, *Propositional information systems*, Journal of Logic and Computation **9** (1999), no. 5, 651–681.

[Kis16a] Kohei Kishida, Private communication during the 13th International Conference on Quantum Physics and Logic, Strathclyde University, Glasgow, June 2016.

[Kis16b] ———, *Logic of local inference for contextuality in quantum physics and beyond*, arXiv:1605.08949 [quant-ph], 2016.

[KLM99] Adrian Kent, Noah Linden, and Serge Massar, *Optimal entanglement enhancement for mixed states*, Physical Review Letters **83** (1999), 2656–2659.

[KLM01] E. Knill, R. Laflamme, and G. J. Milburn, *A scheme for efficient quantum computation with linear optics*, Nature **409** (2001), no. 6816, 46–52.

[Kni05] E. Knill, *Quantum computing with realistically noisy devices*, Nature **434** (2005), 39 EP –.

[Koh03] Jürg Kohlas, *Information algebras*, Springer, 2003.

[Koh12] ______, *Information algebras: Generic structures for inference*, Springer Science & Business Media, 2012.

[Kos47a] Jean-Louis Koszul, *Algebre-sur les opérateurs de dérivation dans un anneau*, Comptes rendus hebdomadaires des séances de l'académie des sciences **225** (1947), no. 4, 217–219.

[Kos47b] Jean-Louis Koszul, *Sur l'homologie des espaces homogènes*, C. R. Acad. Sci. Paris **225** (1947), 477–479.

[Kos51] Jean-Louis Koszul, *Cohomologie des espaces fibrés différentiables et connexions*, Séminaire Bourbaki **1** (1948-1951), 313–319.

[Kra43] J Krausz, *Démonstration nouvelle d'une théoreme de whitney sur les réseaux*, Mat. Fiz. Lapok **50** (1943), no. 1, 75–85.

[KS67] Simon Kochen and Ernst P. Specker, *The problem of hidden variables in quantum mechanics*, Journal of Mathematics and Mechanics **17** (1967), no. 1, 59–87.

[KS96] Jürg Kohlas and Robert F Stärk, *Information algebras and information systems*, Citeseer, 1996.

[KS00] Jürg Kohlas and Prakash P Shenoy, *Computation in valuation algebras*, Handbook of defeasible reasoning and uncertainty management systems, Springer, 2000, pp. 5–39.

[KT85] Leonid A. Khalfin and Boris S. Tsirelson, *Quantum and quasi-classical analogs of bell inequalities*, Symposium on the Foundations of Modern Physics 1985, 1985, pp. 441–460.

[Leh74] Philippe GH Lehot, *An optimal algorithm to detect a line graph and output its root graph*, Journal of the ACM (JACM) **21** (1974), no. 4, 569–575.

[Ler45] Jean Leray, *Sur la forme des espaces topologiques et sur les points fixes des representations*, J. Math. Pures Appl. **9** (1945), 95–248.

[Leu04] Debbie W Leung, *Quantum computation by measurements*, International Journal of Quantum Information **2** (2004), no. 01, 33–43.

[Lev73] L. A. Levin, *Universal sequential search problems*, Probl. Peredachi Inf. **9** (1973), no. 3, 115–116.

[LSW11] Yeong-Cherng Liang, Robert W. Spekkens, and Howard M. Wiseman, *Specker's parable of the overprotective seer: a road to contextuality, nonlocality and complementarity*, Physics Reports **506** (2011), no. 1–2, 1–39, Eprint available at arXiv:1010.1273 [quant-ph].

[LT58] Serge Lang and John Tate, *Principal homogeneous spaces over abelian varieties*, American Journal of Mathematics **80** (1958), no. 3, 659–684.

[Man87] NS Manton, *Connections on discrete fibre bundles*, Communications in Mathematical Physics **113** (1987), no. 2, 341–351.

[Man13] Shane Mansfield, *The mathematical structure of non-locality and contextuality*, DPhil thesis, University of Oxford, 2013.

[May92] J Peter May, *Simplicial objects in algebraic topology*, vol. 11, University of Chicago Press, 1992.

[MB12] Shane Mansfield and Rui Soares Barbosa, *Quantifying contextuality in quantum probability tables*, *Mathematica* package, 2012.

[MB13] ______, *Extendability in the sheaf-theoretic approach: Construction of Bell models from Kochen-Specker models*, (Pre)Proceedings of 10th Wokshop on Quantum Physics and Logic (QPL X), ICFo Barcelona, 2013, Eprint available at arXiv:1402.4827 [quant-ph].

[Mer90a] N. David Mermin, *Quantum mysteries revisited*, American Journal of Physics **58** (1990), no. 8, 731–734.

[Mer90b] ______, *Simple unified form for the major no-hidden-variables theorems*, Physical Review Letters **65** (1990), no. 27, 3373–3376.

[Mer93] ______, *Hidden variables and the two theorems of John Bell*, Reviews of Modern Physics **65** (1993), 803–815.

[MF12] Shane Mansfield and Tobias Fritz, *Hardy's non-locality paradox and possibilistic conditions for non-locality*, Foundations of physics **42** (2012), no. 5, 709–719.

[Mil16] James S Milne, *Etale cohomology (pms-33)*, vol. 33, Princeton university press, 2016.

[Moy49] José Enrique Moyal, *Quantum mechanics as a statistical theory*, Mathematical Proceedings of the Cambridge Philosophical Society **45** (1949), no. 01, 99–124.

[MS12] Robert MacPherson and Benjamin Schweinhart, *Measuring shape with topology*, Journal of Mathematical Physics **53** (2012), no. 7, 073516.

[MS16] Carl A. Miller and Yaoyun Shi, *Robust protocols for securely expanding randomness and distributing keys using untrusted quantum devices*, J. ACM **63** (2016), no. 4, 33:1–33:63.

[MY98] D. Mayers and A. Yao, *Quantum cryptography with imperfect apparatus*, Proceedings 39th Annual Symposium on Foundations of Computer Science (Cat. No.98CB36280), Nov 1998, pp. 503–509.

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

[Nie99] M. A. Nielsen, *Conditions for a class of entanglement transformations*, Physical Review Letters **83** (1999), 436–439.

[Nie03] Michael A. Nielsen, *Quantum computation by measurement and quantum memory*, Physics Letters A **308** (2003), no. 2, 96 – 100.

[OD16] Marc Offroy and Ludovic Duponchel, *Topological data analysis: A promising big data exploration tool in biology, analytical chemistry and physical chemistry*, Analytica Chimica Acta **910** (2016), 1 – 11.

[ORBR17] Cihan Okay, Sam Roberts, Stephen D. Bartlett, and Robert Raussendorf, *Topological proofs of contextuality in qunatum mechanics*, Quantum Info. Comput. **17** (2017), no. 13-14, 1135–1166.

[OTR18] Cihan Okay, Emily Tyhurst, and Robert Raussendorf, *The cohomological and the resource-theoretic perspective on quantum contextuality: common ground through the contextual fraction*, 2018, Eprint available at arXiv:1806.04657 [quant-ph].

[PBS11] Stefano Pironio, Jean-Daniel Bancal, and Valerio Scarani, *Extremal correlations of the tripartite no-signaling polytope*, Journal of Physics A: Mathematical and Theoretical **44** (2011), no. 6, 065303, Eprint available at arXiv:1101.2477 [quant-ph].

[Pen92] Roger Penrose, *On the cohomology of impossible figures*, Leonardo (1992), 245–247.

[Per90] Asher Peres, *Incompatible results of quantum measurements*, Physics Letters A **151** (1990), no. 3-4, 107–108.

[PH15] Jose A. Perea and John Harer, *Sliding windows and persistence: An application of topological methods to signal analysis*, Foundations of Computational Mathematics **15** (2015), no. 3, 799–838.

[PK12] Marc Pouly and Jürg Kohlas, *Generic inference: A unifying theory for automated reasoning*, John Wiley & Sons, 2012.

[Pop14] Sandu Popescu, *Nonlocality beyond quantum mechanics*, Nature Physics **10** (2014), 264 EP –.

[Pos51] Mikhail M Postnikov, *Determination of the homology groups of a space by means of the homotopy invariants*, Doklady Akademii Nauk SSSR (N.S.), no. 76, 1951, pp. 359–362.

[Pou08] Marc Pouly, *A generic framework for local computation*, Ph.D. thesis, Université de Fribourg, 2008.

[Pou10] ———, *Nenok — a software architecture for generic inference*, International Journal on Artificial Intelligence Tools **19** (2010), no. 01, 65–99.

[PP58] Lionel S. Penrose and Roger Penrose, *Impossible objects: A special type of visual illusion*, British Journal of Psychology **43** (1958), 31–33.

[PPSZ05] Ramamohan Paturi, Pavel Pudlák, Michael E. Saks, and Francis Zane, *An improved exponential-time algorithm for k-sat*, J. ACM **52** (2005), no. 3, 337–364.

[PR94] Sandu Popescu and Daniel Rohrlich, *Quantum nonlocality as an axiom*, Foundations of Physics **24** (1994), no. 3, 379–385.

[Pul06] Sylvia Pulmannová, *Extensions of partially ordered partial abelian monoids*, Czechoslovak Mathematical Journal **56** (2006), no. 1, 155–178.

[Ras85] Peter Rastall, *Locality, Bell's theorem, and quantum mechanics*, Foundations of Physics **15** (1985), no. 9, 963–972.

[Rau03] Robert Raussendorf, *Measurement-based quantum computation with cluster states*, Ph.D. thesis, Ludwig-Maximilians-Universität München, October 2003.

[Rau13] ______, *Contextuality in measurement-based quantum computation*, Physical Review A **88** (2013), no. 2, 022322.

[Rau16] ______, *Cohomological framework for contextual quantum computations*, 2016, Eprint available at arXiv:1602.04155v3 [quant-ph].

[RB01a] Robert Raussendorf and Hans Briegel, *Computational model underlying the one-way quantum computer*, arXiv:0108067 [quant-ph], 2001.

[RB01b] Robert Raussendorf and Hans J. Briegel, *A one-way quantum computer*, Physical Review Letters **86** (2001), no. 22, 5188–5191.

[RBB02] Robert Raussendorf, Daniel E Browne, and Hans J Briegel, *The one-way quantum computer – a non-network model of quantum computation*, Journal of Modern Optics **49** (2002), no. 8, 1299–1306.

[RBB03] Robert Raussendorf, Daniel E. Browne, and Hans J. Briegel, *Measurement-based quantum computation on cluster states*, Phys. Rev. A **68** (2003), 022312.

[Red87] M Redhead, *Non-locality, incompleteness and realism*, Oxford University Press, Oxford. Rietdijk, CW (1966),"A Rigorous Proof of Determinism Derived from the Special Theory of Relativity", Philosophy of Science **43** (1987), 598–609.

[Rob18] David E. Roberson, Private communication during the Workshop on Logical Aspects of Quantum Information, Lorentz Institute, Leiden, July 2018.

[Ros70] Donald J Rose, *Triangulated graphs and the elimination process*, Journal of Mathematical Analysis and Applications **32** (1970), no. 3, 597 – 609.

[Rou73] Nicholas D Roussopoulos, *A max $\{m, n\}$ algorithm for determining the graph h from its line graph g*, Information Processing Letters **2** (1973), no. 4, 108–112.

[Rou17] Frank Roumen, *Cohomology of effect algebras*, Proceedings 13th International Conference on Quantum Physics and Logic, Glasgow, Scotland, 6-10 June 2016 (Ross Duncan and Chris Heunen, eds.), vol. 236, Open Publishing Association, 2017, pp. 174–201.

[SB93] Harald Scherer and Paul Busch, *Problem of signal transmission via quantum correlations and einstein incompleteness in quantum mechanics*, Physical Review A **47** (1993), no. 3, 1647.

[Sch99] T Schoning, *A probabilistic algorithm for k-sat and constraint satisfaction problems*, Foundations of Computer Science, 1999. 40th Annual Symposium on, IEEE, 1999, pp. 410–414.

[Sch02] D Schlingemann, *Stabilizer codes can be realized as graph codes*, Quantum Information & Computation **2** (2002), no. 4, 307–323.

[Ser55] Jean-Pierre Serre, *Faisceaux algébriques cohérents*, Annals of Mathematics (1955), 197–278.

[Ser56] ______, *Géométrie algébrique et géométrie analytique*, Annales de l'institut Fourier, vol. 6, Association des Annales de l'Institut Fourier, 1956, pp. 1–42.

[Ser57] ______, *Sur la cohomologie des variétés algébriques*, Gauthier-Villars, 1957.

[Sha76] Glenn Shafer, *A mathematical theory of evidence*, vol. 42, Princeton university press, 1976.

[She89] Prakash P. Shenoy, *A valuation-based language for expert systems*, International Journal of Approximate Reasoning **3** (1989), no. 5, 383–411.

[She92] Prakash P Shenoy, *Valuation-based systems: A framework for managing uncertainty in expert systems*, Fuzzy logic for the management of uncertainty, John Wiley & Sons, Inc., 1992, pp. 83–104.

[She94] ______, *Consistency in valuation-based systems*, ORSA Journal on Computing **6** (1994), no. 3, 281–291.

[She96] ______, *Axioms for dynamic programming*, A. Gammerman (ed.), Computational Learning and Probabilitic Reasoning, John Wiley & Sons Ltd, 1996, pp. 259–275.

[Shi84] Abner Shimony, *Controllable and uncontrollable non-locality*, Foundations of quantum mechanics in the light of new technology (1984), 225–230.

[Sho99] Peter W. Shor, *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer*, SIAM Review **41** (1999), no. 2, 303–332.

[Sho03] Peter W Shor, *Why haven't more quantum algorithms been found?*, Journal of the ACM (JACM) **50** (2003), no. 1, 87–90.

[Sim83] Barry Simon, *Holonomy, the quantum adiabatic theorem, and berry's phase*, Physical Review Letters **51** (1983), no. 24, 2167.

[Sim18] Andrew W Simmons, *On the computational complexity of detecting possibilistic locality*, Journal of Logic and Computation **28** (2018), no. 1, 203–217.

[SL88] David J. Spiegelhalter and Steffen L. Lauritzen, *Sequential updating of conditional probabilities on directed graphical structures*, Networks **20** (1988), no. 5, 579–605.

[Spe60] Ernst Specker, *Die Logik nicht gleichzeitig entscheidbarer Aussagen*, Dialectica **14** (1960), 239–246, Translation by M. P. Seevinck, The logic of non-simultaneously decidable propositions, eprint available at arXiv:1103.4537 [physics.hist-ph].

[Spe05] R. W. Spekkens, *Contextuality for preparations, transformations, and unsharp measurements*, Phys. Rev. A **71** (2005), 052108.

[SR93] E. C. George Sudarshan and Tony Rothman, *A new interpretation of Bell's inequalities*, International Journal of Theoretical Physics **32** (1993), no. 7, 1077–1086.

[SS90] Prakash P. Shenoy and Glenn Shafer, *Axioms for probability and belief-function propagation*, Uncertainty in Artificial Intelligence (Ross D. Shachter, Tod S. Levitt, Laveen N. Kanal, and John F. Lemmer, eds.), Machine Intelligence and Pattern Recognition, vol. 9, North-Holland, 1990, pp. 169 – 198.

[SS91] Glenn R Shafer and Prakash P Shenoy, *Local computation in hypertrees*, Tech. Rept. School of Business, University of Kansas, 1991.

[SSS$^+$90] Prakash P Shenoy, Glenn Shafer, Ross D. Shachter, Tod S. Levitt, Laveen N. Kanal, and John F. Lemmer, *Axioms for probability and belief-function propagation*, vol. 9, pp. 169–198, North-Holland, 1990.

[Ste40] Karl Stein, *Topologische bedingungen für die existenz analytischer funktionen komplexer veränderlichen zu vorgegebenen nullstellenflächen*, Mathematische Annalen **117** (1940), no. 1, 727–757.

[Ste51] Norman Steenrod, *The topology of fiber bundles, volume 14 of*, Princeton Mathematical Series (1951), 16.

[SW01a] D. Schlingemann and R. F. Werner, *Quantum error-correcting codes associated with graphs*, Phys. Rev. A **65** (2001), 012308.

[SW01b] Dirk Schlingemann and Reinhard F Werner, *Quantum error-correcting codes associated with graphs*, Physical Review A **65** (2001), no. 1, 012308.

[SW10] Benjamin Schumacher and Michael Westmoreland, *Quantum processes, systems and information*, Cambridge University Press, 2010.

[Tsi80] Boris S. Tsirel'son, *Quantum generalizations of Bell's inequality*, Letters in Mathematical Physics **4** (1980), 93–100.

[VdNDDM04] Maarten Van den Nest, Jeroen Dehaene, and Bart De Moor, *Graphical description of the action of local Clifford transformations on graph states*, Physical Review A **69** (2004), 022316.

[VdNMDB06] Maarten Van den Nest, Akimasa Miyake, Wolfgang Dür, and Hans J. Briegel, *Universal resources for measurement-based quantum computation*, Phys. Rev. Lett. **97** (2006), 150504.

[Vor62] Nikolai N. Vorob'ev, *Consistent families of measures and their extensions*, Theory of Probability and its Applications (Teoriya Veroyatnostei i ee Primeneniya) **7** (1962), no. 2, 147–163 (English: N. Greenleaf, trans.), 153–159 (Russian).

[VV14] Umesh Vazirani and Thomas Vidick, *Fully device-independent quantum key distribution*, Phys. Rev. Lett. **113** (2014), 140501.

[Wae14] Mordecai Waegell, *Primitive nonclassical structures of the n-qubit pauli group*, Phys. Rev. A **89** (2014), 012321.

[Wat09] John Watrous, *Quantum computational complexity*, Encyclopedia of complexity and systems science (2009), 7174–7201.

[Wei52] André Weil, *Sur les théorèmes de rham*, Commentarii Mathematici Helvetici **26** (1952), 119–145.

[Wey29] Hermann Weyl, *Gravitation and the electron*, Proceedings of the National Academy of Sciences **15** (1929), no. 4, 323–334.

[Whi32] Hassler Whitney, *Congruent graphs and the connectivity of graphs*, American Journal of Mathematics **54** (1932), no. 1, 150–168.

[Whi40] Hassler Whitney, *On the theory of sphere-bundles*, Proceedings of the National Academy of Sciences **26** (1940), no. 2, 148–153.

[Wig32] Eugene Wigner, *On the quantum correction for thermodynamic equilibrium*, Physical Review **40** (1932), no. 5, 749–759.

[Wil12] Virginia Vassilevska Williams, *Multiplying matrices faster than coppersmith-winograd*, Proceedings of the Forty-fourth Annual ACM Symposium on Theory of Computing (New York, NY, USA), STOC '12, ACM, 2012, pp. 887–898.

[WM99] Nic Wilson and Jérôme Mengin, *Logical deduction using the local computation framework*, European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty, Springer, 1999, pp. 386–396.

[WR14] Inc. Wolfram Research, *Mathematica*, version 10.0 ed., Wolfram Research, Inc., 2014.

[YM54a] Chen-Ning Yang and Robert L Mills, *Conservation of isotopic spin and isotopic gauge invariance*, Physical Review **96** (1954), no. 1, 191.

[YM54b] ———, *Isotopic spin conservation and a generalized gauge invariance*, Physical Review **95** (1954), no. 1, 631.

[ZG18] Wlodek Zadrozny and Luciana Garbayo, *A sheaf model of contradictions and disagreements. preliminary report and discussion*, 2018, Eprint available at arXiv:1801.09036 [cs.CL].

List of Symbols

$(-)^{\downarrow}$ Projection operation of a valuation algebra
$(-)^{-x}$ Variable elimination operation of a valuation algebra
$(-)_{\downarrow}$ Projection operation of a tuple system
$\mathbb{B}$ Boolean semiring
$\mathcal{C}_Q$ Closure operator associated to the Galois connection $\langle \mathscr{T}_Q, \mathscr{M}_Q \rangle$ over a domain Q
$\mathsf{ch}(i)$ Child node of a node i of a join tree
$\mathsf{CLC}(\mathcal{S})$ Cohomological logical contextuality of $\mathcal{S}$
$\mathsf{CLC}(\mathcal{S}, s)$ Cohomological logical contextuality of $\mathcal{S}$ at section s.
C_1^n Clifford group on n qubits
$\mathcal{M}^{(q)}$ Measurement cover for the q-th line scenario
$\mathsf{CSC}(\mathcal{S})$ Cohomological strong contextuality of $\mathcal{S}$.
$\mathcal{D}_R$ R-distribution functor
$\mathcal{D}_R\mathcal{E}$ Presheaf of event R-distributions
$\downarrow_{\subseteq}$ Downward closure
$\mathcal{E}$ Sheaf of events
$\mathcal{E}^{(q)}$ Sheaf of events of the q-th line scenario
$\mathcal{F}$ Abelian representation of a possibilistic model.
$\mathsf{flatten}$ Flatten function
$\mathcal{F}^{(q)}$ Abelian representation of $\mathcal{S}^{(q)}$
$\mathsf{CLC}^q(\mathcal{S})$ Cohomological logical q-contextuality of $\mathcal{S}$
$\mathsf{CLC}^q(\mathcal{S}, s)$ Cohomological logical q-contextuality of $\mathcal{S}$ at section s
$\mathsf{CSC}^q(\mathcal{S})$ Cohomological strong q-contextuality of $\mathcal{S}$
$\mathbb{H}_n$ Hilbert space of n qubits
$\bowtie$ Natural join
$\mathsf{LC}(\mathcal{S})$ Logical contextuality of $\mathcal{S}$.
$\mathsf{LC}(\mathcal{S}, s)$ Logical contextuality of $\mathcal{S}$ at section s.
$\rightsquigarrow$ Graham reduction step
$\mathcal{M}$ Measurement cover
$\mathfrak{D}_\bullet$ Cycle in vertex representation
$\mathsf{NS}(\Sigma)$ No-signalling polytope of a scenario $\Sigma = \langle X, \mathcal{M}, (O_m) \rangle$
$X^{(q)}$ Set of measurements for the q-th line scenario
$\mathcal{S}^{(q)}$ q-th line model
$\mathcal{N}(\mathcal{M})$ Nerve of a cover $\mathcal{M}$.
Ω Frame functor
Ω_x Frame of a variable x
$\oplus$ Addition modulo 2.

$\otimes$	In valuation algebras: combination operation
$\mathcal{P}$	Powerset
$\mathcal{P}_{\mathsf{fin}}$	Finite powerset
$\mathsf{pa}(i)$	Set of parent nodes of a node i of a join tree
π_A	Cartesian projection onto A
$\mathscr{P}_n$	Pauli group on n qubits
$\preceq$	Partial order of an ordered valuation algebra
$e^{(1)}$	First probabilistic line model of an empirical model e
$\S$	General quantifier over for an algebra of propositional sentences
$\rho_U^{U'}$	(Pre)sheaf restriction map from U' to U
$\mathbb{R}_{\geq 0}$	Non-negative reals
$\mathcal{S}$	Possibilistic empirical model.
$\mathcal{S}_e$	Possibilistic empirical model presheaf generated by the support of a probabilistic model $e = \{e_C\}_{C \in \mathcal{M}}$
$\mathsf{SC}(\mathcal{S})$	Strong contextuality of $\mathcal{S}$
$\langle X, \mathcal{M}, (O_m) \rangle^{(q)}$	q-th line scenario of $\langle X, \mathcal{M}, (O_m) \rangle$
$\mathsf{sep}(i)$	Separator set of a node i of a join tree
$\star$	Local complementation operation
$\mathscr{M}_Q$	Model operator over a domain Q
supp	Support of a function
$\mathscr{T}_Q$	Theory operator over a domain Q
$\mathsf{Trs}(\mathcal{M}, \mathcal{F})$	Set of elements in $\mathsf{Trs}_{\mathcal{F}}$ which are trivialised by $\mathcal{M}$
$\mathsf{Trs}_{\mathcal{F}}$	Set of isomorphism classes of $\mathcal{F}$-torsors
$\underline{\mathscr{D}}$	Cycle in edge representation
$d(\phi)$	Domain of a valuation ϕ
e_S	Neutral element of a valuation algebra for domain S
F_R	Free functor on a ring R.
O	Unique set of outcomes
$O^{(q)}$	q-th line scenario outcome set
O_m	Set of outcomes for measurement m.
X	Set of measurements
z_S	Null element of a valuation algebra for domain S

Index

abelian representation, see empirical model, possibilistic, abelian representation of33
acyclicity (Vorob'ev), see scenario, acyclic32
All-vs-Nothing argument, **30**, 160
antichain, **17**
AvN contextuality, **30**, 160
AvN triple conjecture, **161**

balanced state, **183**
belief function, **108**
Bell state, **12**
Bell's model, **12**
bundle diagram, **24**, 30
  planar representation, 42

(augmented) Čech cochain complex, 34
Čech cohomology, **35**
check vector, **169**
closure operator, 111
coboundary, **35**
  coboundary map, **34**
cochain, **34**
cocycle, **35**
cohomological q-th contextuality, see contextuality, cohomological contextuality, higher cohomological logical and strong contextuality48
cohomological contextuality, see contextuality, cohomological logical and strong contextuality37
cohomology obstruction, **36**
  higher cohomology obstruction, **48**
connecting homomorphism, **36**, 47, 51
constraint satisfaction problem (CSP), **117**
context, 10, **16**
context (valuation algebras), 111
contextual cycle, see cycle over a line scenario, contextual cycle87
contextual fraction, 16, 138
contextuality, **27**
  cohomological logical contextuality, **37**
    higher cohomological logical contextuality, **48**
  cohomological strong contextuality, **37**
    higher cohomological strong contextuality, **48**
  hierarchy, **28**, 40
  $\mathsf{LC}^{(k)}$-type contextuality, **70**
  possibilistic contextuality, **27, 28**
  strong contextuality, **28**
cover, 16
  connected, **35**
cycle over a line scenario, **70**
  chordal cycle, **70**
  contextual cycle, **87**
  proper cycle, **71**
cyclic contextuality property (CCP), **87**

disagreement (valuation algebras), 120
  complete disagreement, **133**
  global disagreement, **120**
  local disagreement, **121**
domain (valuation algebras), see valuation, domain of101

empirical model, 11, **22**
  possibilistic, **23**
    abelian representation of, **33**
  probabilistic, **22**
  quantum realisable, **22**
event, *see also* sheaf of events, 11, **20**

$\mathcal{F}$-presheaf, **52**
$\mathcal{F}$-torsor, **52**
false negative, **42**, 44, 60
frame functor, **118**
fusion algorithm, 139
  fusion algorithm for contextuality, **142**

Galois connection, **110**
Graham-reduction, **32**
graph state, **162**

Greenberger-Horne-Zeilinger (GHZ) model, 15, 27, **29**, 43
Greenberger-Horne-Zeilinger (GHZ) state, **15**

Hardy model, **24**, 42, 66, 105
hidden variable, 26
hierarchy of contextuality, see contextuality, hierarchy28
higher cohomological contextuality, see contextuality, cohomological contextuality, higher cohomological logical and strong contextuality48

indicator function, **103**
inference problem, **103**
  joint valuation of, **103**
  multi-query, **103**
  single-query, **103**
  treewidth of, **141**
infomorphism, 114
information algebra, **102**
information set, 114
  information set of CSPs, **118**
  information set of linear inequalities, **117**
  information set of linear systems, **117**
  predicate information set, **116**
  propositional information set, **110**
invertible local operator (ILO), 177

join tree, **141**
  covering assignment, 148
  covering join tree, **148**
  node of, **141**
  treewidth of, **141**

knowledgebase, **101**
Kochen–Specker (KS) model, **24**, 43

line graph, 63, 71
line model, 64, **65**
  k-th line model, **65**
  first line model, **64**
  probabilistic line model, **95**
line scenario, 61, **63**
  k-th line scenario, **63**
  first line scenario, **61**
local Clifford (LC) group, **162**
local complementation, **163**
Local operations and classical communication (LOCC) class, **177**
logical Bell inequality, 14
logical contextuality, see possibilistic contextuality28

measurement, 10, **16**
  compatible measurements, **16**
  equatorial, **182**
measurement context, see context10
measurement cover, see cover16
measurement scenario, see scenario16

natural join, **106**
nerve, **34**
no-signalling, **22**
  no-signalling polytope, **26**
no-Z lemma, **76**
node of a join tree, **141**
  child node, **148**
  leaf, **148**
  parent node, **148**
  separator of, **148**
non-locality, **27**, 27

partial family, **75**
  standard, **76**
path over a line scenario, **70**
  chordal path, **70**
  edge representation of, **71**
  vertex representation of, **71**
Pauli group, **158**
Peres–Mermin magic square model, 43
planar representation, see bundle diagram, planar representation42
Popescu–Rohrlich (PR)-box, 12, **23**, 43
possibility table, **23**
presheaf, **20**
  compatible family of, **20**
  compatible sections of, **20**
  flasque beneath the cover, **23**, 47
  of a valuation algebra, **118**
  of event R-distributions, **21**
  restriction map of, **20**
  section of, **20**
    global, **20**
  subpresheaf, **20**
probability table, **12**

q-th cohomology obstruction, see cohomology obstruction, higher cohomology obstruction48

R-distribution, **21**
  R-distribution functor, **21**
R-linear equation, **30**
R-linear theory, **30**
R-potential, **108**
relational database, **105**
  attribute of, **105**
relative cohomology, **35**

scenario, **16**
  (n, k, l)-type, **17**
  acyclic (Vorob'ev), **32**
  Bell-type, **17**
  cyclic (line), **73**
  quantum realisable, **17**
set of measurements, **16**
set of outcomes, **16**
sheaf, **20**
  of events, **20**
sheaf cohomology, 32
simplicial complex, **18**
  dimension of, **18**
  face of, **18**
    dimension of, **18**
  facet of, **18**
  vertex of, **18**
Specker's triangle model, **24**, 107
stabiliser state, **159**
stabiliser subgroup, **159**
stabiliser subspace, **159**
Stochastic local operations and classical communication (SLOCC) class, 177
subpresheaf, see presheaf, subpresheaf20

treewidth, **141**
truth valuation, **121**
tuple system, **112**
  of cartesian tuples, 112
  of probability distributions, 112
  of propositional interpretations, **112**

Vorob'ev's theorem, **32**
valuation
  domain of, **101**
valuation algebra, **101**
  adjoint valuation algebra, **129**
  combination operation of, **101**
  idempotent valutation algebra, **102**
  labelling operation of, **100**
  neutral elements of, **102**
  null elements of, **102**
  ordered valuation algebra, **128**
  semiring valuation algebra, **107**
  valuation algebra of formulae, 110
  valuation algebra of information sets, 110, 114
  valuation algebra of probability potentials, **108**
  valuation algebra of set potentials, **108**
  variable elimination, **139**
  variable of, 101
    frame of, 103
    tuple of, 103
valuation prealgebra, **100**

XOR-theory, 160