

Correlation and combinatorics: causal contextuality and spin systems



Amy Jade Searle

St Hugh's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2024

Acknowledgements

I came to work with my supervisors through somewhat unconventional means, it mostly being the result of both Samson and Vlatko having a very optimistic approach to research, as well as a (perhaps somewhat) lenient approach to paperwork. This has allowed me a good amount of freedom in research, which I am hugely grateful to both Samson and Vlatko for. It moreover would not have been possible without the generous combined Smith-Westlake scholarship and EPSRC funding, or the support of the FoQaCiA network.

I would also like to thank my collaborators, including Rui, who also supervised many parts of my project, and Joey. Sharing research interests is a huge blessing, and useful for being spoken out of dead ends and into more promising directions.

I am—like many—indebted to the teachers and mentors who have encouraged me over the years, and without whom I am sure to have made very little progress. This includes the schoolteachers Mrs Osmond, Ms de Villiers and Mrs Jacobs, who are each in some way responsible for my love of mathematics.

It also includes Sinéad Griffin, who supervised my during my exchange year at Berkeley, and who has in many ways shaped the passion I have for science and the scientific process.

Oxford, the quirky place, has the unique problem that many leave as quickly as they come. Nevertheless through this process I have met many interesting, thoughtful and determined people who have all in some way left their mark. I am thankful also to those who, like me, have lingered around in Oxford. A special thanks to Pedr and Juani for some memorable moments on the bike, Toby and Faiz for many a sunday night dinner, and Olly and Pat for a productive week writing up next to the fire and admiring Skeggy, and the ALP Quantum group: Lucia, Lodovico, Simone, Maria, Giuseppe, Aditya, Nicetu, and Tristan. Thank you for making the office such a comfy spot (albeit one with some serious design flaws). Also, and especially, thank you to Eva, for being an unstoppable source of brilliant reading recommendations, a feverish dancer, my writing confidante. You taught me a lot.

The research group at INL in Braga welcomed me wholeheartedly, and it was a special part of my DPhil. Thank you especially to Anita and Davide for hosting me at their house, and the rest of the group for taking me in with open arms.

I would also like to thank the cycling gang up north—Izzy, George, Sam, Lucy, Addi—who started as cycling friends but quickly developed into my UK based family. And Ella Halliday, who I’m still afraid to wear my cycling gear around (and that’s the sign of a good friendship).

Thank you to Agnes, who has already given me so many wonderful memories. You are the best. Also Lazlo and the cats, for keeping us company.

My family have supported me in ways too numerous to enumerate, and this is something I am learning not to take for granted. Mom, dad, Mike, grandma Louie, you really did make this possible. Mom, your successes instilled in young me a belief that anything is possible, and dad, you taught me to wonder about the order of things. Mike, you are my best friend and thank you for always believing that what I pursue is interesting. Grandma, it has been a blessing to live so close to you these past four years and hear your stories (of which you have many).

I would like to dedicate this thesis to my late friend Freya, who had such a profound impact on me.

Abstract

Contextuality is a nonclassical feature of quantum systems—exhibited by data that is produced empirically or theoretically—which in the realm of sheaf theory is characterised by local consistency but global inconsistency. A large part of this thesis is concerned with studying how this signature of nonclassicality is apparent also when measurements are embedded in some causal structure, and so motivating the study of *causal contextuality*.

We begin with temporal correlations, which occur when measurements are performed sequentially on the system, and in which the definition of nonclassicality becomes sensitive to memory resources with which the classical system is equipped. For certain types of such memory, we show that there exists a map from the temporal setup to a (appropriately defined) contextuality setup, such that every nonclassical temporal empirical model satisfying no-signalling constraints consistent with the memory function corresponds to a contextual empirical model on this constructed scenario—one can view this also as a simulation of a subset of the temporal correlations by the contextuality setup. The existence of such a map allows us to apply a result from Vorob'ev in order to say, for any temporal setup and choice memory function, whether nonclassical correlations can arise. We then study causal setups by employing the notion of strategy from game semantics. We in particular show how ‘playing off’ Nature strategies, corresponding to adaptive hidden variables, against Experimenter strategies, which may also be adaptive, realises the classical correlations of certain causal setups from the literature. We show that adaptivity on the side of the Experimenter, by reducing the sets of measurements empirical data is obtained over, can remove the inconsistencies that are imperative for the observation of contextuality.

In the second part of the thesis, we study spin Hamiltonians on random graphs, focusing on exact descriptions in the thermodynamic limit. By utilising the graphon, which is the limit object of a dense random graphs sequence, we are able to derive analytical results for certain graphons and certain choice of Hamiltonian. Our overarching result is that the equilibrium physics in the thermodynamic limit is described by a set of coupled equations containing the graphon, and which describes product, i.e. unentangled, states.

Contents

1	Introduction	1
1.1	Combinatorics and correlation in contextuality	2
1.2	Local consistency and global inconsistency	7
1.3	Beyond spatial correlations	8
1.4	Spins Hamiltonians on Graphs	9
1.5	Motivation and contributions	10
I	Background material	12
2	Sheaf theory and simplicial complexes	13
2.1	Consistent data via sheaves and presheaves	14
2.2	Simplicial complexes	17
2.3	The distribution monad and conditional distributions	21
2.4	Bundle diagrams	22
3	The sheaf approach to contextuality	24
3.1	The setup	26
3.2	Deterministic systems and classicality	31
3.3	Probabilistic systems and empirical models	32
3.4	Contextuality	34
3.5	Kochen-Specker scenarios and the Kochen-Specker Theorem	39
3.6	Simulation and quantum resources	41
3.6.1	Subscenarios from restriction	44
3.7	Measurement based quantum computing	45
3.8	Vorob'ev's Theorem and measurement covers	49

4	Causality in a quantum information setting	51
4.1	Classical probabilities on causal DAGs	53
4.2	Bayesian networks: classicality as Markovianity	54
4.3	Causal Bell scenarios	56
4.4	The Leggett-Garg setup and macrorealism	62
II	Research contributions	66
5	Flattening temporal correlations to contextuality correlations	67
5.1	The setup	71
5.2	Empirical models for temporal scenarios	76
5.3	Mapping temporal setups to contextuality setups	81
5.4	Mapping empirical models	84
5.5	No-signalling constraints beyond the AoT constraints	85
5.6	Combinatorial conditions for classicality	89
5.7	Θ -memory	94
5.8	Finite state machines	98
5.9	Allowing outputs to be stored in memory	100
5.9.1	Parity machines and related examples	101
5.9.2	Nonclassicality and memory with outputs	102
5.10	Discussion	104
6	Causality and contextuality: a game semantics approach	106
6.1	Causal contextuality	107
6.2	The strategy presheaves	111
6.3	Covers and empirical models	112
6.4	The sheaf axiom and $\mathcal{S}^N$	113
6.5	Causal measurement scenarios: examples	117
6.5.1	Causal Bell scenarios	117
6.5.2	Contextuality scenarios	121
6.5.3	Temporal scenarios	124
6.6	Adaptive Experimenter strategies	127
6.6.1	Anders-Browne example	133
6.6.2	HLP scenarios	134
6.7	Discussion	138

7 Spin Hamiltonians on graphs	140
7.1 Spins on graphs	143
7.1.1 Operators	145
7.2 Phase transitions	146
7.3 Spin Hamiltonians	148
7.3.1 The classical Ising model and Max-Cut	149
7.3.2 The XXZ model	151
7.3.3 The transverse field Ising model	151
7.3.4 Results concerning non-lattice geometries	152
7.4 Thermodynamic limit of spin systems on dense random graphs	154
7.4.1 Sequences of Erdős-Rényi graphs	154
7.4.2 Dense graph sequences	159
7.4.3 Thermodynamic physics on dense graph limits	167
7.4.4 The Ising model on $W(x, y) = xy$ and $W(x, y) = p$	170
7.4.5 The transverse field Ising model on $W(x, y) = \sqrt{xy}$	173
7.5 Discussion	175
8 Conclusion	178
8.1 Summary	178
8.2 Future research directions	179
8.2.1 Vorob'ev-type results for causal measurement scenarios	179
8.2.2 Using Θ -consistent strategies to realise outcome memory . . .	180
8.2.3 Complexity classifications of spin systems on graph in the thermodynamic limit	180
Bibliography	182

Chapter 1

Introduction

— Summary —

This introductory chapter motivates the topics of study of this thesis. Specifically, the use of combinatorics in the fields of contextuality and many body physics is discussed, and based on such observations, a motivation given for the research directions taken. We conclude with a summary of the contributions contained in the thesis.

The era of quantum information will impact many areas of the technology landscape [111]. Crucial to using quantum technologies, however, is understanding which features that these systems exhibit which cannot be produced by any classical system. Broadly, in prepare and measure setups, which includes Bell [21] and contextuality [91, 129] setups as a special case, we would like to understand if the system being measured is exhibiting quantum features—which we for now simply define to be features which systems obeying the laws of classical physics cannot exhibit—and this is done by checking the frequencies of measurement outcomes, and determining whether the resulting probability distribution could have been classically produced [16, 88, 73, 130]. If not, the system being measured is considered a ‘resource’ in some sense; the correlations that it produces can be used, for example, to perform tasks with a higher probability of success than if the correlations had originated from a classical system [100, 70, 138]. On the other hand, quantum annealers work by embedding some mathematical problem in the Hamiltonian of the quantum system, so that by allowing the system to relax to its ground state we are able to find

the solution to the problem. Here, the interest is in how hard it is to classically compute the ground state [83, 82]. If it is extremely difficult, then the quantum system gives us a way of solving it by engineering the correct Hamiltonian, and then allowing for it to relax. So in this context, we are interested in which Hamiltonians have ground states which are difficult to compute classically, or, alternatively the negation of this, which is which have ground states which can be efficiently solved classically—either exactly, or approximately with bounded error.

It turns out that in both cases combinatorial objects, in the form of simplicial complexes and graph limits, are important. This is perhaps not a surprise: both the commutation relation of quantum mechanics and interaction patterns of spin Hamiltonians describe the relation between objects, these objects being measurements and spins respectively, and it is combinatorics which allows us to express these binary or n -ary relations.

This thesis explores prepare and measure scenarios in which there is some causal dependency between measurements, and spin Hamiltonians on random graphs, emphasising structural aspects along the way. We begin by giving an overview of how this topic crops up in the field of contextuality.

1.1 Combinatorics and correlation in contextuality

There are a number of frameworks which describe contextuality, but central to all of them is the idea that a physical theory, which might be a hypothesis about how the universe behaves or might be purely of theoretical interest, assigns **probabilities** to **events**. To simplify matters, we consider an experimental setup of a particular kind: a system $\mathcal{S}$ may be measured with **observables** from the set X . Each observable x has outcome set O_x , and we say that two observables x and y are **compatible** if they can be performed simultaneously, which in quantum physics corresponds to commutativity. We write Σ for the set of compatible subsets of X , and refer to this as the **compatibility complex**. An element $\sigma \in \Sigma$ denotes a compatible set of observables.

A **measurement event** is a tuple $\langle x, o \rangle$, or an assignment of an outcome o to the measurement x . Since in any round of experiment only compatible observables can be simultaneously measured, experiments yield collections of measurement events for each compatible subset of the set of observables X . More succinctly, the data obtained in any round of experiment is captured using the shorthand

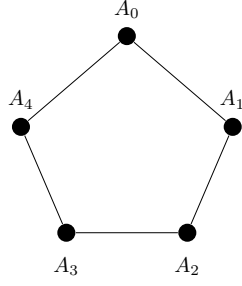


Figure 1.1: One choice of compatibility complex for 5 measurements, labelled A_0 through to A_5 .

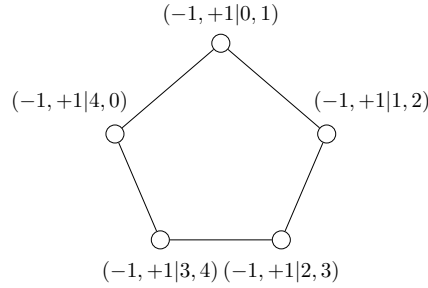


Figure 1.2: An induced subgraph of the exclusivity graph corresponding to the compatibility complex shown in Figure 1.1, in the case that each measurement is dichotomic, or has two outcomes.

$(o_1 o_2 \dots o_n | x_1 x_2 \dots x_n)$ where $x_1, x_2, \dots$ are compatible observables and o_i the outcome upon measuring x_i . This will be referred to as an **experimental event**, or simply an **event**.

In the graph theoretic approach of Cabello, Serveni and Winter, [36] the exclusivity graph E corresponding to any experimental setup $\mathcal{M} = \langle X, \Sigma, O \rangle$ is built as follows. Vertices consist of events, and vertices share an edge if these events are mutually exclusive, which is if both contain an observable x but send it to different outcomes. Suppose the setup consists of five observables $A_1 - A_5$ and each can have outcome -1 or 1 . The compatibility complex is displayed in Figure 1.1.

Moreover, a choice of induced subgraph of the exclusivity graph resulting from the above compatibility constraints is the pentagon displayed in Figure 1.2.

Cabello, Serveni and Winter showed that certain induced subgraphs of the exclusivity graph, such as the one above, correspond to inequalities which bound the correlations of the experimental setup. This can also be seen by viewing the exclusivity graph as a weighted graph, as in [138]. In particular, a theory associates with each event $(o_i, o_j | A_i, A_j)$, where $i, j \in \{0, 1, 2, 3, 4\}$, a probability $P(o_i, o_j | A_i, A_j)$. It can be shown that if one considers a theory which has a **noncontextual hidden**

variable, or NCHV, explanation, then these probabilities must satisfy

$$\sum_{i=0}^4 P(-1, +1 | A_i, A_{i+1}) \leq 2. \quad (1.1)$$

The NCHV bound for a sum of probabilities, which in this case is 2, can be shown to be equal to the independence number $\alpha(G)$ of the associated exclusivity graph with the events in the sum occurring as vertices. The independence number of a graph is the maximum number of vertices which can be selected which are non-adjacent, and the problem of determining $\alpha(G)$ for a given graph G is called the maximum stable set problem (MSSP) [101]—the solution can be constructed as the solution to a linear program. Equivalently, the independence number is the maximum number of 1's that can be obtained by assigning either 0 or 1 to each vertex, with the constraint that each clique should contain only a single vertex assigned to 1. Classical theory, a NCHV theory, therefore produces correlations which are bounded by the independence number of G .

It is known that quantum theory can exceed the bound above. Nevertheless, it can be shown that the Lovász number of G , written $\Theta(G)$, places a bound on the quantum realisable correlations, which for the pentagon gives the upper bound

$$\sum_{i=0}^4 P_Q(-1, +1 | A_i, A_{i+1}) \leq \sqrt{5}. \quad (1.2)$$

The Lovász number, or the Theta function, as introduced by Lovász, can be seen as a semidefinite program relaxation of the MSSP, and in particular in [101] Lovász is able to show that $\alpha(G) \geq \Theta(G)$. And, finally, we consider theories which obey the exclusivity principle, which requires that if two events are exclusive—so that they cannot occur together—then the sum of their probabilities cannot exceed one. It can be shown that for such theories we must have

$$\sum_{i=0}^4 P_E(-1, +1 | A_i, A_{i+1}) \leq \frac{5}{2}, \quad (1.3)$$

for this example, where $\frac{5}{2}$ is the fractional packing number of the pentagon graph. More generally, this bound is given by $\alpha^*(G)$, the fractional packing number of the associated graph G , which can be seen as a relaxation of the problem of finding the independence number by allowing for assignments to vertices in $[0, 1]$ (rather

than just $\{0,1\}$). As such, $\alpha^*(G) \geq \alpha(G)$ [29]. Moreover, it always holds that $\alpha^*(G) \geq \Theta(G)$ [90].

Surprisingly, from just the observation that probabilities must obey the exclusivity principle, an observation due to Specker, has allowed for the derivation of inequalities which bound the sum of probabilities associated with any exclusivity graph of events G . Moreover, these logical conditions arise naturally as structural constraints—as certain parameters on a graph.

Note that if for the choice of exclusivity graph

$$\alpha(G) = \Theta(G) = \alpha^*(G) \tag{1.4}$$

then the associated inequality cannot be used to differentiate between correlations arising from classical theory, quantum theory, and theories obeying the exclusivity principle, as all will be bounded in the same way. An interesting question is for which experimental setup $\mathcal{M}$ does there exist at least one exclusivity graph G with

$$\alpha(G) \neq \alpha^*(G), \tag{1.5}$$

so that not all correlations in the setup are classically realisable. One can instead understand this in terms of the compatibility complex of the setup, which comes about naturally not in the graph approach so far discussed, but in the sheaf approach of Abramsky and Brandenburger [6]. This result is a repurposing of a Theorem due to Vorob'ev, originally cast in the context of game theory [137]; in this context the theorem allows the following question to be answered:

For which compatibility complexes Σ do there exist no-signalling probability distributions which are contextual?

Here, no-signalling correlations encompass the correlations which arise from theories satisfying the exclusivity principle [58], and for our purposes will be the most general ‘allowed’ types of correlation, since it singles out those that act in accordance with the theory of special relativity. Vorob'ev showed that the complex **not** being acyclic is a necessary and sufficient condition for contextuality to arise [1]. This result is quite remarkable—it guarantees that, if Σ is not acyclic, one can assign probabilities to events such that (i) these probabilities are no-signalling and (ii) they are contextual. Moreover, if Σ is acyclic, the theorem guarantees that no such

¹See section [2.2] for the definition of acyclicity

assignment of probabilities exists. Vorob'ev's Theorem has been used in the context of quantum contextuality by Barbosa in [15, 14], Xu and Cabello in [144], and Budroni and Morchio in [32]. In some literature, a graph is instead used to represent compatibility, so that an edge connects two observables which are compatible. The corresponding simplicial complex is obtained as the clique complex of the graph, since cliques correspond to sets of observables which are pairwise compatible. It is known that the simplicial complex is acyclic if the graph is chordal (see [20] for this result), and so graph chordality instead becomes the property of interest.

The general programme of the above research is the following: given a logical relationship which must hold amongst a set of events, derive inequalities which must be satisfied by the associated probabilities. These inequalities are referred to as 'Boole's conditions of possible experience', and it was Pitowsky who first understood Boole's conditions of possible experience in the context of correlation polytopes, and their implication for Quantum theory [117].

How does one formally express these logical conditions? As done in Abramsky and Hardy [8], and Abramsky [2], one can associate with each event $e_i = (o_0, o_1, \dots, o_n | x_0, x_1, \dots, x_n)$ a propositional formula ϕ_i such that ϕ_i is **true** when event e_i has occurred and **false** otherwise. In other words, we introduce $(x_0 = o_0)$ which is **true** when x_0 obtains the value o_0 . The proposition ϕ_i will therefore be

$$\phi_i = (x_0 = o_0) \wedge (x_1 = o_1) \wedge \dots \wedge (x_n = o_n) \quad (1.6)$$

and will be true when event e_i occurs, where event e_0 corresponds to the top vertex of the pentagon and we label moving clockwise.

Consider the five events associated with the five vertices of the pentagon graph of Figure 1.2, and associate each $i \in \{0, 1, 2, 3, 4\}$ with a propositional formula ϕ_i . Note that $\phi_3 \wedge \phi_4$ is contradictory, since it contains $(A_4 = -1) \wedge (A_4 = +1)$ which is always **false**. More generally, the collection $\{\phi_i\}$ of propositions is said to be K -consistent if the size of the largest subfamily of these propositions that is satisfiable is K . As is shown in by Abramsky and Hardy in [8], classical probability theory then implies that

$$\sum_i p_i \leq K, \quad (1.7)$$

where p_i is the probability of the event e_i associated with ϕ_i occurring. For the example, this gives

$$\sum_i p_i \leq 2, \quad (1.8)$$

since the maximum consistent subfamily is 2. This gives the same bound as the independence number of G : since an independent set of the graph considers vertices which do not share an edge, and since edges correspond to exclusive or contradicting events, a consistent statement must be an independent set. Inequalities of the type in [1.7] are referred to as logical Bell inequalities, and a Theorem in [8] is able to relate violation of these inequalities to contextuality of the system. In particular, the theorem states

A probabilistic model over any cover is contextual if and only if it violates one of finitely many logical Bell inequalities.

This can be combined with Vorob'ev's theorem to relate the structure of the compatibility complex, in particular whether it possesses acyclicity, to the violation of logical Bell inequalities.

1.2 Local consistency and global inconsistency

The sheaf approach allows one to visualise apparent paradoxes, and especially, for our purposes, those arising in contextuality, quite clearly. One of the most well-known of these paradoxes occurs in the Bell setup, and can be stated as: violation of a logical Bell inequality is consistent with the model assigning non-zero probability to contradicting logical statements. For Bell, this was the observation that both (i) realism and (ii) locality could not be simultaneously satisfied in the Bell setup [21]. Pitowsky observed, however, that these paradoxes are never observed in a *single* system, but rather inferred from the empirical data collected from a collection of systems which have been identically prepared. To quote him (this is really a quote stolen from the exposition of Pitowsky's work by Abramsky in [2]):

A violation of Boole's conditions of possible experience cannot be encountered when all the frequencies concerned have been measured on a single sample. Such a violation simply entails a logical contradiction; 'observing' it would be like 'observing' a round square. We expect Boole's conditions to hold even when the frequencies are measured on distinct, large random samples. But they are systematically violated, and there is no easy way out (see below). We thus live 'on the edge of a logical contradiction'.

In the sheaf approach, contextuality is visualised via **local consistency** but **global inconsistency**, akin to that in Escher's 'Ascending and Descending' shown below.

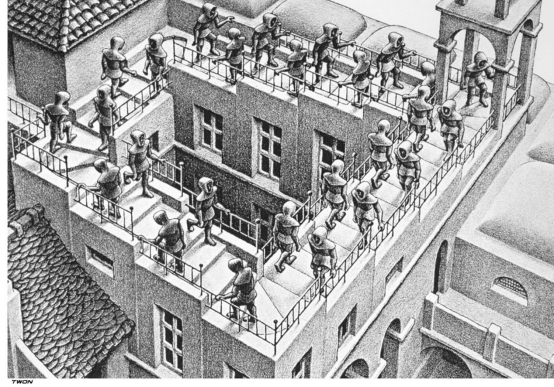


Figure 1.3: 'Ascending and Descending' by M.C. Escher.

If one looks at any of the four sides of the square staircase, it looks as if they are ascending; and yet, it is impossible for a staircase which always ascends to close on itself, as in the image. The observation of contextuality in empirical data is exactly analogous: while for the partial views of the system we obtain on compatible measurements there is no contradiction, it seems that when one tries to extend this description to all observables of the system simultaneously, there is an apparent paradox. Pitowsky's observation is that one never observes this paradox in a single system, but rather in the data arising from a number of identically prepared systems. Vorob'ev's Theorem on the other hand is a statement about for which compatibility relations such inconsistencies can arise—for example, it is imperative that the staircase be cyclic for the paradox to be observed. We aim to use the sheaf approach to show that this exact phenomena is responsible also for the nonclassicality present in temporal and more general causal correlations.

1.3 Beyond spatial correlations

So far, the logical constraints which have been placed on the assignment of probabilities to events arise entirely due to the compatibility structure Σ of the observables. However, nonclassicality can also be demonstrated for sequences of incompatible measurements on a single system, so that measurements (i) rather than being space-like separated, are separated in time and (ii) need not be compatible in the sense so far defined. The lack of compatibility simply means that the occurrence of an event

can depend, in principle in an arbitrary way, on events which occurred before. We therefore want to study how the results such as those above generalise to more involved logical relationships between events, such as the event e_i having some logical dependency on the event e_j . This occurs naturally if one includes causal constraints on the set of observables, so that observable A_i must always occur after event A_j , or more concretely the observables occur at timelike separated spacetime points $\mathbf{x}_i$ and $\mathbf{x}_j$.

The questions of interest are

- i How can we represent such conditions combinatorially?
- ii How do these conditions affect the resulting correlations which can form, and how does this differ for different choices of physical theory?

1.4 Spins Hamiltonians on Graphs

Given a graph G , we attach to each vertex of G a particle, such as a d dimensional spin or a fermionic mode, and build a Hamiltonian $H(G)$ such that for each edge $(i, j) \in G$ the Hamiltonian contains a 2-local term describing some type of interaction between only two of the particles.

Interestingly, for supersymmetric Hamiltonians describing interacting fermions it can be shown that supersymmetric ground states which have $l+1$ nodes occupied by a fermion are in one to one correspondence with the number of l -dimensional holes in the independence complex of G [37, 75]. This provides a direct link between solving for the ground state and finding holes in the independence complex of G , the latter of which is a well known problem in the field of algebraic topology.

Results of these kind, which directly attach the difficulty of solving the ground state of a Hamiltonian $H(G)$ to a structural property of the graph G give insight into those Hamiltonians which might be hard to simulate classically.

It is this problem that we are interested in, but as opposed to the hard-core boson Hamiltonian of [56], we study 2-local Hamiltonians describing the interaction of d -dimensional spins according to some graph G . We moreover focus on the case in which (i) G is a random graph and (ii) the thermodynamic limit, which is the limit in which the number of interacting particles tends to infinity.

1.5 Motivation and contributions

Overall, there are two motivations for the research in this thesis:

1. A large motivation is the extension of results of the type above to setups in which there is some dependency between observables.
2. Another motivation is the study of how structure—expressed in terms of graphs or more general combinatorial objects—influences the resulting probability distributions in other contexts. Namely, the study of n spins, each of dimension d , such that the interaction terms in the Hamiltonian H are specified by a graph G with spins as vertices

We summarise the contributions:

- i An application of Vorob'ev's Theorem in the temporal setting, where the memory of the classical system is constrained in some way, is given in Chapter [5](#). This allows us to say, given a particular experimental setup and choice of memory function, whether nonclassicality can arise in such a way that the no-signalling of the non-classical model matches that of a classical system with memory described by the memory function. This is reminiscent of the case in Bell scenarios—any non-local correlation for the Bell setup must still obey the same no-signalling relations as one which is classical. We moreover relate classicality of a set of correlations to its simulability by a (probabilistic) classical machine, with finite state machines being a specific example.

Publications. (preprint) Searle, A., Barbosa, R.S., Abramsky, S. Mapping Temporal Correlations to Contextuality Correlations. *To be made publicly accessible in the near future, and will appear as a talk at Quantum Physics and Logic in July this year.*

- ii The sheaf approach to contextuality [\[6\]](#), and to causal Bell scenarios [\[66\]](#) is extended to general causal setups in Chapter [6](#). This extension models dependency between measurement events, arising due to causal constraints, via an enabling relation. This terminology we use, and in particular the central objects of study which are causal games, are reminiscent of that used in research on classical event structures by Winskel and Nielsen [\[141\]](#) and Kahn and Plotkin [\[80\]](#). We discuss also how adaptive strategies for the Experimenter allow us to capture a subset of examples considered by Henson, Lal and Pusey in [\[73\]](#), where classicality is captured by Markovianity with respect to a directed acyclic graph on the random variables in the setup.

Publications. Abramsky, S*, Barbosa, RS*, and Searle, A.* 2024. Combining contextuality and causality: a game semantics approach. Phil. Trans. R. Soc. A.38220230002. <http://doi.org/10.1098/rsta.2023.0002>.

iii For certain dense random and weighted graph sequences (G_n) , where n indicates the number of vertices of the graph, the equilibrium state of the 2-local Hamiltonian $H(G_n)$ is derived in the thermodynamic limit. This is the contents of Chapter 7. The result follows from utilising graphons, a graphon being the limit of a dense sequence of graphs. The thermodynamic limit corresponds to the limit in which $n \rightarrow \infty$. The two main results are

- (a) Define the sequence (G_n) (respectively $(\bar{G}_n)$) of random graphs (weighted graphs) from stochastic sampling (weighted sampling) of a graphon W . In the limit $n \rightarrow \infty$, we have $f(H(G_n)) = f(H(\bar{G}_n))$, where f is the free energy. This guarantees that observables converge in this limit.
- (b) In the limit $n \rightarrow \infty$, the equilibrium state of $H(\bar{G}_n)$ is governed by three coupled nonlinear integrals. Such integrals are known as Fredholm equations in the literature, and the graphon forms what is called the kernel of the integral. Such equations, if solvable, give an analytic solution for the system.

These results are contained in [128] which came about as a generalisation of the result in [134], which contains the special case in which the graphs are Erdős-Rényi graphs.

Publications.

- (a) Searle, A. and Tindall, J. “Thermodynamic limit of spin systems on random graphs”. In: Phys. Rev. Res. 6 (1 2024), p. 013011. doi: 10.1103/PhysRevResearch.6.013011. url: <https://link.aps.org/doi/10.1103/PhysRevResearch.6.013011>.
- (b) Tindall, J., Searle, A., Alhajri, A. et al. Quantum physics in connected worlds. Nat Commun 13, 7445 (2022). <https://doi.org/10.1038/s41467-022-35090-y>.

Part I

Background material

Chapter 2

Sheaf theory and simplicial complexes

— Summary —

This chapter covers the relevant background material on sheaf theory and abstract simplicial complexes. Sheaves are first defined using topological spaces, after which a generalised version of the sheaf axiom is given for any category formed from a poset, rather than that specifically corresponding to the topological space. Terminology pertaining to simplicial complexes, such as acyclicity, are covered, and will be necessary in later chapters, when introducing Vorob'ev's Theorem [137].

We begin with an introduction to some basic concepts concerning sheaves and simplicial complexes, which will be assumed knowledge for the remainder of the thesis. This is by no means an extensive background on the material, and its aim will be to give the required definitions, but also give the reader some intuition for the objects which in later chapters will become central to our study of nonclassical correlations.

After reading the introduction, in which I spoke about a graph theoretic approach to contextuality, it might not at first be obvious why we now instead turn to what is often seen as the heavy mathematical machinery of sheaves. Mike Robinson, in his university lecture series on sheaf theory, answers this succinctly by stating that ‘if one wants to derive a framework which tracks data over a space, such that data must satisfy consistency constraints imposed by the space, one always derives the theory

of sheaves’. That is, presheaves and sheaves are in some sense more basic than other approaches, because the notion of local consistency and global inconsistency, which is central to the emergence of nonclassicality or contextuality, is already well-defined via the obstructions to global sections for the presheaf. Another motivation for working with sheaves is that many categorical constructions which have come about in other contexts—such as in database theory or game semantics—can be readily reused for our own purposes.

Sheaves have a long history of development (see for example [55]) and there are a number of different definitions, depending on the level of generality one wants and their intended use. The form in which we view them, as objects which allows us to attach data to topological spaces, or more generally to posets, was conceived when Jean Leray in 1946 introduced a way of assigning modules to closed sets [96]. After a number of subsequent developments, it was Alexander Grothendieck in 1955 who introduced presheaves [69], and in particular understood how the categorical setting allowed for sheaves to be generalised: rather than assigning data to topological spaces one needed only a Grothendieck topology, which abstracts the notion of the cover of a topological space to a cover in categorical terms. It was Godement in 1958 who published the first complete volume on sheaf theory [65]. Despite the flourishing field which followed Leray’s research, some of which was done as a prisoner of war, it took some time for his ideas to be appreciated and accepted; “Les Mathematical Reviews demand‘erent ‘à quoi ça peut servir?’”, or, “The Mathematical Reviews asked ‘What good can this do?’” [125]. The tension between the perceived success of his work and the realised importance of it might offer some consolation to the modern mathematician and scientist.

While the definition of presheaf and sheaf we consider is in some sense a basic one, their introduction will still to some extent rely on simple notions from category theory. The definitions of category, functor, and commuting diagram will be assumed but can be read about on any one of the numerous websites which cover this material, such as [ncatlab](#) or [Math3ma](#). The reader interested in a more detailed or mathematical introduction should consult a standard textbook on the matter, or the thesis of Justin Curry titled ‘Sheaves, Cosheaves, and Applications’ [46] which gives a rigorous account.

2.1 Consistent data via sheaves and presheaves

Sheaves and presheaves are essentially useful for attaching **data** to **topological spaces**. Let us illustrate this with an example. In everyday life, we are tasked

with processing the visual data which we receive from the objects around us and processing this data (in the visual cortex) in order to make sense of the world around us. Here, the topological space is the manifold of space time we observe, and the data is visual data which is converted from photons to pictures by our eyes and brain. However, we cannot simultaneously take in all of our surroundings— we are unfortunately only capable of looking at a window of the world or a piece of this topological space. The visual cortex, in processing this data, glues together the pictures from two different windows of space time. This can be done in a consistent fashion because, provided the time difference between the collection of data for the region of space A and the region of space B is diminishing, it is expected that the images agree on their overlap. We therefore say that visual data is **glueable**.

The first part, in which data is attached to a topological space, is formally captured by an object called a presheaf.

Definition 2.1.1. *A presheaf $\mathcal{F}$ of sets on a topological space X is given by two pieces of information:*

1. *For each open set U of X , a set $\mathcal{F}(U)$ (called the set of sections).*
2. *For each pair of open sets $V \subseteq U$ of X , a restriction map $\rho_V^U : \mathcal{F}(U) \rightarrow \mathcal{F}(V)$ such that*
 - (a) *for all U , $\rho_U^U = id_U$.*
 - (b) *Whenever $W \subseteq V \subseteq U$, $\rho_W^U = \rho_W^V \circ \rho_V^U$.*

A sheaf is a presheaf which satisfies additional axioms which ensures that this data is well-behaved.

Definition 2.1.2. *A sheaf $\mathcal{F}$ of sets is a presheaf which satisfies the following additional condition. Let $\{U_i\}_{i \in I}$ be a collection of sets in X and $U = \bigcup_{i \in I} U_i$, where U is also an open set in X . Take a family of sections $\{s_i\}_{i \in I}$ with each $s_i \in \mathcal{F}(U_i)$ which is compatible, meaning for every $i, j \in I$, $s_i|_{U_i \cap U_j} = s_j|_{U_i \cap U_j}$. Then it holds that*

1. *(Existence of gluing) There exists a global section $s \in \mathcal{F}(U)$ with $s|_{U_i} = s_i$.*
2. *(Uniqueness of gluing) Let $s, t \in \mathcal{F}(U)$ be such that $s|_{U_i} = t|_{U_i}$ for all $i \in I$. Then $s = t$.*

Notice that in order to check the above sheaf condition, we require data only on an indexed cover $\{U_i\}_{i \in I}$ of the presheaf and not on the entirety of the topological space X .

It will be convenient to work with the category-theoretic definition of presheaves and sheaves, which reduces to the definition above as a special case. Here, a presheaf is simply a contravariant functor $\mathcal{F} : \mathbf{C}^{\text{op}} \rightarrow \mathbf{Set}$ which sends

1. Any object $U \in \text{Ob}(\mathbf{C})$ to an object $\mathcal{F}(U) \in \text{Ob}(\mathbf{Set})$
2. Any morphism $f : V \rightarrow U$ in $\mathbf{C}$ to a morphism $\mathcal{F}(f) : \mathcal{F}(U) \rightarrow \mathcal{F}(V)$ in $\mathbf{Set}$,

subject to the conditions of functoriality which are (i) $\mathcal{F}(\text{id}_U) = \text{id}_{\mathcal{F}(U)}$ and (ii) if $f : U \rightarrow V$ and $g : V \rightarrow W$ in $\mathbf{C}$, then $\mathcal{F}(g \circ f) = \mathcal{F}(f) \circ \mathcal{F}(g)$.

Conventionally, an element $s \in \mathcal{F}(U)$ is called a **section** and the morphism $\mathcal{F}(f) : \mathcal{F}(U) \rightarrow \mathcal{F}(V)$ which is the image of $f : V \rightarrow U$ in $\mathbf{C}$ is referred to as the **restriction map**. We write $\rho_V^U : \mathcal{F}(U) \rightarrow \mathcal{F}(V) :: s \mapsto s|_V$. Thus $s|_V$ is the restriction of the section s to U .

Definition 2.1.3 (Presheaf (Categorical definition)). *A presheaf is a contravariant functor $F : \mathbf{C}^{\text{op}} \rightarrow \mathbf{Set}$.*

The presheaf over a topological space X arises as an example of this by defining the category $\text{Op}(X)$ of open subsets of X . It has as objects the open sets U of the topological space and as morphisms inclusions $U \hookrightarrow V$ of open sets.

We will be most interested in the case in which the category $\mathbf{C}$ arises from a poset $(P, \leq)$, meaning that the category has as objects elements $p \in P$ and a morphism $p \rightarrow q$ iff $p \leq q$ in P . We recall that given two elements x and y of a poset P , define the **meet**, $x \wedge y$, and **join**, $x \vee y$, to be the greatest lower bound and least upper bound respectively,

$$x \wedge y = \max\{z \mid z \leq x, z \leq y\} \quad x \vee y = \min\{z \mid z \geq x, z \geq y\}, \quad (2.1)$$

and this definition may be extended to a subset $\{p_i\}_{i \in I}$ of P as $\bigwedge_i p_i$ and $\bigvee_i p_i$ respectively. The join and meet is never guaranteed to exist, and for our purposes we require that the poset is **bounded complete**, which is when all bounded sets $\{p_i\}_{i \in I}$ have a join $\bigvee_i p_i$.

In order to define a sheaf on the category P , a notion of cover is needed, since the gluing condition relies on such a cover being defined. However, our interest will not be in proving that the gluing condition for *all* covers, as is required by

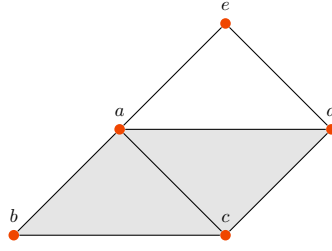


Figure 2.1: An abstract simplicial complex, which can be seen as the gluing together of n -dimensional simplexes along certain faces. Here, a shaded grey region indicates a ‘filled in’ face; for example, the shaded triangle on the vertices a , b and c indicates that the three-dimensional face $\{a, b, c\}$ is included in the complex.

the definition of a sheaf, but rather with respect to a single cover which will be physically motivated by the problem.

Therefore, rather than require the full definition with the machinery of Grothendieck topologies, we can simply check when the gluing axiom, which can be defined for just a single family $\{f_i : p_i \rightarrow p\}$ of morphisms, holds.

Definition 2.1.4 (Sheaf axiom for family of morphisms). *Given a presheaf $F : P \rightarrow \mathbf{Set}$, it is said to satisfy the sheaf axiom for the family or morphisms $\{f_i : p_i \rightarrow p\}$ if, whenever we have a compatible family of elements $\{s_i \in F(p_i)\}_{i \in I}$ which are compatible— in the sense that for any $i, j \in I$, if $p_i \leq p_j$ then $s_j|_{p_i} = s_i$ — then there exists a unique $s \in F(p)$ such that $s|_{p_i} = s_i$ for all $i \in I$.*

In the case that P has as its elements the open sets of a topological space X , so that $P = \mathcal{P}(X)$, we will be interested in covers $\{p_i\}_{i \in I}$ which can be represented by a combinatorial object called a **simplicial complex**, and in which case the join is simply the union operator on sets.

2.2 Simplicial complexes

A simplicial complex is an object which allows for relations beyond pairwise relations amongst some choice of objects to be captured. We are in particular concerned with **abstract simplicial complexes** (ASC), where the position of the points (e.g. in a n -dimensional real space) are not specified, although oftentimes this will be done implicitly in order to visualise them, such as that example in Figure 2.1. We will represent higher dimensional faces of the simplicial complex, beyond 2-dimensional ones which are lines, as filled in polyhedra.

Definition 2.2.1 ((Finite) Abstract Simplicial Complex). *A simplicial complex on a finite set X is a collection Σ of non-empty subsets of X subject to two requirements:*

1. *for each vertex x in X , the singleton $\{x\}$ is in Σ , and*
2. *if τ is in Σ and $\sigma \subset \tau$ then σ must also be in Σ .*

We call an element $\sigma \in \Sigma$ the **face** of the simplicial complex, and $\dim(\sigma) := |\sigma| - 1$ is the **dimension** of the face σ . In particular, a face with $n + 1$ vertices is n -dimensional. Given a simplicial complex Σ , the set of faces of Σ which are not included in any other faces, and are therefore maximal, will be denoted $\max(\Sigma)$:

$$\max(\Sigma) := \{\tau \in \Sigma \mid \nexists \sigma \in \Sigma. \tau \subset \sigma\}. \quad (2.2)$$

Now, instead of assigning data to a cover $\{U_i\}_{i \in I}$ of a topological space X , we directly assign data to the faces of a simplicial complex Σ . As with graphs, there are classes of simplicial complex which are of special interest because they occur frequently in application.

Consider the simplicial complex $\Delta_A := \mathcal{P}(A)$, which consists of all subsets of the finite set A . This is known as a **simplex**. On the other end of the spectrum, the simplicial complex in which every face either contains a single element or is empty is called the **discrete complex**. In particular, given a set A we define the discrete complex

$$\mathcal{D}_A := \{\emptyset\} \cup \{\{v\} \mid v \in A\}. \quad (2.3)$$

In such a complex, every context is disjoint and contains a single element. We will write $\mathcal{D}_k$ with $k \in \mathbb{N}$ for the discrete complex on $[k]$ where we use the conventional notation $[k] = \{1, 2, 3, \dots, k\}$. Similarly, we will write Δ_k for the simplex on the set $[k]$.

Given two simplicial complexes Σ and Θ , a simplicial map $\pi : \Sigma \rightarrow \Theta$ is a map which preserves the compatibility structure of the complex.

Definition 2.2.2 (Simplicial Map). *A simplicial map $\pi : \Sigma \rightarrow \Theta$ between two simplicial complexes Σ and Θ with vertex sets X and Y respectively is a function $f : X \rightarrow Y$ such that*

$$\forall \sigma \in \Sigma. \{f(x) \mid x \in \sigma\} \in \Theta. \quad (2.4)$$

Given a subset $Y \subseteq X$ of the vertex set X of Σ , the **induced subcomplex** is defined

$$\Sigma|_Y := \{\sigma \cap Y \mid \sigma \in \Sigma\}. \quad (2.5)$$

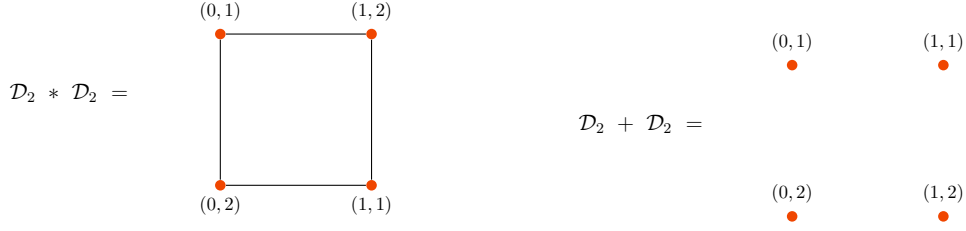


Figure 2.2: Two different product operations that can be performed on abstract simplicial complexes (ASCs). *Left* . The simplicial join of two ASCs. *Right* . The coproduct of two ASCs.

Note that there is always a simplicial map $\Sigma|_Y \rightarrow \Sigma$ from an induced subcomplex to the original simplicial complex, which is simply an inclusion.

Let the set of all abstract simplicial complexes be **ASC**. An operation on simplicial complexes is a map $\mathbf{ASC} \times \mathbf{ASC} \rightarrow \mathbf{ASC}$, and is of special interest because it tells us how two (or more) simplicial complexes can be combined to produce a new, more involved, simplicial complex. Two operations of interest are the simplicial join $*$ and the coproduct $+$ which are defined

1. $\Sigma * \Theta := \{\sigma \sqcup \theta | \sigma \in \Sigma \wedge \theta \in \Theta\},$
2. $\Sigma + \Theta := \{\{0\} \times \tau | \tau \in \Sigma\} \cup \{\{1\} \times \tau | \tau \in \Theta\}.$

Here, $\sigma \sqcup \theta$ denotes the disjoint union of two sets, or symbolically $\sigma \sqcup \theta := (\{0\} \times \sigma) \cup (\{1\} \times \tau)$. Taking for example $\Sigma = \Theta = \mathcal{D}_2$, where recall $\mathcal{D}_2 = \{\emptyset, \{1\}, \{2\}\}$, the result of $\mathcal{D}_2 * \mathcal{D}_2$ is shown on the left of Figure 2.2 and the result of $\mathcal{D}_2 + \mathcal{D}_2$ on the right. Categorically, **ASC** can be thought of as a category with objects being simplicial complexes and morphisms simplicial maps.

When considering simplicial complexes as the choice of cover of a particular presheaf, it will become important to know if it possesses a certain property known as **acyclicity**, which can be thought of as the generalisation of the corresponding property of graphs, where it refers to the absence of cycles. We will give the algorithmic definition of acyclicity, in which a simplicial complex is acyclic if and only if we can remove vertices from it in a certain way so as to produce the zero complex $\Delta_0 := \{\emptyset\}$, which contains no vertices.

If vertices can be removed from the simplicial complex it is said to be **Graham reducible**.

Definition 2.2.3 (Graham reducible). *We say that a simplicial complex Σ is Graham reducible if there exists a face $\sigma \in \max(\Sigma)$ which has vertex $x \in \Sigma$ such that the vertex x is not in any other maximal faces.*

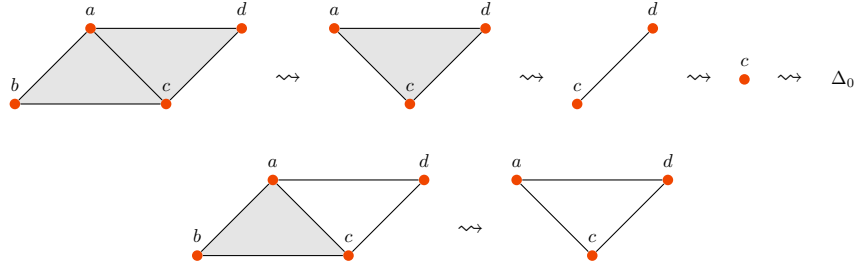


Figure 2.3: *Top.* Successful Graham reduction of an acyclic simplicial complex. *Bottom.* A simplicial complex which cannot be Graham reduced to Δ_0 and is therefore not acyclic.

The removal of the vertex x to form the simplicial complex $\Sigma' := \{\sigma \setminus \{x\} \mid \sigma \in \Sigma\}$ is called a Graham reduction step. We write $\Sigma \rightsquigarrow \Sigma'$.

Definition 2.2.4 (Acyclic Simplicial Complex). *A simplicial complex Σ is said to be acyclic if there is a series of Graham reduction steps $\Sigma \rightsquigarrow \Sigma' \rightsquigarrow \dots \rightsquigarrow \Delta_0$, ending in the zero complex Δ_0 .*

Some simple examples are useful to illustrate what is meant by acyclicity. First consider the acyclic simplicial complex shown in the top half of Figure 2.3 together with its Graham reduction to the complex Δ_0 . Looking first at the leftmost complex we see that vertices a and d both occur in a single maximal face of the simplicial complex. We begin by removing a , which leaves us with the next simplicial complex in that row, where $\rightsquigarrow$ is used to indicate the process of Graham reduction. After this, every vertex occurs in one maximal face which is simply the face $\{b, c, d\}$, and so any one of them can be removed, which is chosen to be b resulting in the next simplicial complex. In this manner, we are left with the simplicial complex $\{d\}$ containing only the single vertex d . This can clearly be reduced to Δ_0 , and thus this simplicial complex can be seen to be acyclic.

On the other hand, the simplicial complex shown in the bottom half of Figure 2.3 is not acyclic. This is because although we can initially reduce it as before by removing the vertex a , after this each vertex occurs in *two* maximal faces, and therefore none qualify to be removed in a Graham reduction step.

We will also require the following proposition, which relates the acyclicity of a simplicial complex to the acyclicity of its induced subcomplexes.

Proposition 2.2.5. *Given a simplicial complex Σ , if there exists an induced subcomplex K which is Graham irreducible, then this implies that Σ is Graham irreducible.*

Proof. It is enough to show that for an acyclic simplicial complex every induced subcomplex is Graham reducible to a disjoint union of points (since we would like the negation of this). However, [103] defines acyclicity as the property of having no induced subcomplex which is biconnected and non-trivial (i.e. not a simplex; having more than one maximal face), and this definition of acyclicity and the version which relies on a series of Graham reduction steps were shown to be identical in [14]. This alternative definition makes it clear that for an acyclic simplicial complex, every induced subcomplex is reducible to a disjoint union of points. $\square$

2.3 The distribution monad and conditional distributions

In this thesis we will be working with sheaves $\mathcal{F}$ in which elements $s \in \mathcal{F}(U)$ will be physically interpreted as behaviours of a system on the space U . In the case that U is a set of measurements, for example, a behaviour is an assignment of an outcome $s(o)$ to every measurement $m \in U$. However, we would like also to describe probabilistic systems. Thus, given some set U of measurements, we would like to attach some probability $w(s)$ to each deterministic $s \in \mathcal{F}(U)$.

This is done formally as follows. Given a set X and a commutative semiring $R = (R, +, 0, \cdot, 1)$, let $\mathcal{D}_R(X)$ be the set which has as its elements functions $w : X \rightarrow R$ such that

$$\sum_{s \in X} w(s) = 1, \quad (2.6)$$

i.e. it is normalised. Moreover, w is required to have finite support, where the support is defined to be the subset of X for which $w(s) \neq 0$.

Semirings of particular interest will be

1. *The reals.* $R = (\mathbb{R}, +, 0, \times, 1)$
2. *The booleans.* $R = (\mathbb{B}, \vee, 0, \wedge, 1)$.

In fact, $\mathcal{D}_R : \mathbf{Set} \rightarrow \mathbf{Set}$ is a monad, or a presheaf $\mathbf{Set} \rightarrow \mathbf{Set}$. Given a set $X' \subseteq X$, we restrict an element $w \in \mathcal{D}_R(X)$ to an element $w|_{X'} \in \mathcal{D}_R(X')$ where

$$w|_{X'}(s') = \sum_{s \in X; s|_{X'} = s'} w(s). \quad (2.7)$$

This allows us to speak of, for $R = \mathbb{R}$, probability distributions over the set X .

Central to the theory of probability is the notion of **conditional probability**. This categorical account of conditional probability will follow Jacobs [79] and Furber and Jacobs [62].

Given a distribution $w \in \mathcal{D}_R(X)$, we can form the conditional distribution $w|_p$ where $p : X \rightarrow \{0, 1\}$ is a predicate on X assigning a truth value (0 or 1) to elements of X . While in [62] they consider fuzzy predicates which are more general maps $p : X \rightarrow [0, 1]$, we will only be interested in the sharp case sending elements to 0 or 1. A predicate p singles out certain elements of X which are known to be true—thus the conditional distribution $w|_p$, read “ w given p ”, when acting on an element $s \in X$ gives the probability of s given that the set $\{r | p(r) = 1\}$ is true. This conditional distribution is defined

$$w|_p(s) := \frac{w(s)p(s)}{\sum_{r \in X} w(r) \cdot p(r)}. \quad (2.8)$$

2.4 Bundle diagrams

We have seen that a presheaf $\mathcal{F} : \mathcal{P}(X) \rightarrow \mathbf{Set}$ is a sheaf with respect to the cover $\Sigma \subseteq \mathcal{P}(X)$ of X if the compatible family $\{s_\sigma\}_{\sigma \in \Sigma}$ of sections has a unique global section $h \in \mathcal{F}(X)$. However, even if $\mathcal{F}$ is a sheaf, its precomposition with the distribution monad $\mathcal{D}_R \circ \mathcal{F}$ may not be. The sheaf property, or lack thereof, for the compatible family $\{p_\sigma\}_{\sigma \in \Sigma}$ can be conveniently captured via a **bundle diagram**.

The bundle diagram has as its base the simplicial complex, and above each vertex x in the simplicial complex is the fibre of values that x can take on under $\mathcal{F}$. For any n -dimensional face $\sigma \in \Sigma$ of the simplicial complex, a section $s \in \mathcal{F}(\sigma)$ which is contained in the support of p_σ is represented via an n -simplex connecting the assignment values contained in s at each vertex $x \in \sigma$. For example, if the simplicial complex and fibres are as in the diagram below, then taking $\sigma = \{a_1, b_1\}$, and $s \in \text{supp}(p_\sigma)$ to be a function $\sigma \rightarrow \{0, 1\}$ sending $a_1 \mapsto 0$ and $b_1 \mapsto 0$, we represent this by the line connecting 0 and 0 above the face $\{a_1, b_1\}$, as shown in green.

A global section $h \in \mathcal{F}(X)$ is a closed path through the fibres, as highlighted in red above, since a global section defines a single outcome for each variable. Conveniently, then, the failure of $\mathcal{D}_R \circ \mathcal{F}$ to be a sheaf can be observed by finding a line, such as that highlighted in green, which does not extend to a closed path. This is because it implies that there exists a local section s in the support of p_σ which does not correspond to any global section with support on each $\sigma \in \Sigma$. It therefore

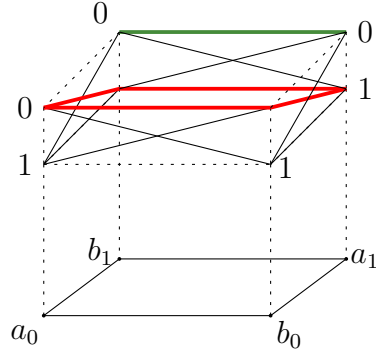


Figure 2.4: An example bundle diagram.

cannot arise from the restriction of a global section $h \in \mathcal{F}(X)$ which is in the support $\text{supp}(X)$, but if $\mathcal{D}_R \circ \mathcal{F}$ is a sheaf then this would always be the case. We will analyse this in more detail in the following Chapter.

Chapter 3

The sheaf approach to contextuality

— Summary —

This chapter covers Abramsky and Brandenburger’s sheaf approach to contextuality [7], and relevant developments which came out of this approach. This includes the resource theory of contextuality [3], quantum advantage due to contextuality in a measurement based quantum computing setting [119, 10], the contextual fraction which gives a measure of contextuality [4], and Vorob’ev’s Theorem [137], which allows one to single out the contextuality setups which can host nonclassical correlations.

Famously, in 1935 Einstein, Podolsky and Rosen wrote a paper titled “Can Quantum-Mechanical Description of Physical Reality be Considered Complete?” which presented what became known as the EPR paradox [50]. It questioned whether quantum mechanics could indeed be considered a complete theory of Nature, which John S. Bell showed to be equivalent to quantum theory being in tension with the assumptions of (1) reality and (2) locality [21]. In this single stroke, the foundations of physics which had up to that point seemed unshakeable, were put into question.

Another way to view the type of non-classicality that occurs in general Bell setups is to see it as it being impossible to assign an ontic state to the system (i.e. a

state in which all observables of the system are well-defined) due to the observation that in quantum mechanics such observables can be context dependent; in particular this can happen even when no-signalling is satisfied. The context refers to the set of measurements which a measurement m occurs together with, and for the value of m to be context-dependent is in stark opposition with the classical viewpoint. Systems which behave like this are said to be **contextual**, much like Escher’s stairs which perpetually ascend and end at the same position (see Introduction).

Quantum contextuality however encompasses a much broader class of examples besides general Bell setups. The demonstration of contextuality relies only on a measurement set together with a compatibility relation, which for quantum physics is the binary commutation relation, so that Bell setups arise as examples in which commutation arises because observers, or parties, have access to nonoverlapping subsets of the Hilbert space. Many approaches to contextuality exist, including Spekkens’ ontological approach [131], the hypergraph approach of Cabello, Severini and Winter [36], the sheaf theoretic approach of Abramsky and Brandenburger [7], and contextuality by default of Dzhafarov [49]. We will work within Abramsky Brandenburger’s sheaf theoretic approach. The power of this approach is that it allows for the import of tools from sheaf theory to the context of quantum contextuality. To name a few examples, this has included development of

- **Logical Bell Inequalities** and proof of their completeness by Abramsky and Hardy [8];
- **Cohomological tools** for contextuality by Abramsky, Mansfield and Barbosa [9], in particular using Čech cohomology to characterise contextuality;
- The **contextual fraction** as a generalisation of the non-local fraction by Mansfield and Barbosa [4];
- A definition of **simulation** of a system $\mathcal{S}$ by another system $\mathcal{S}'$, including a **resource theoretic formulation** of contextuality by Abramsky, Barbosa, Karvonen and Mansfield, and Karvonen, respectively [3, 81].

While it is interesting to demonstrate that a quantum system may exhibit contextuality for the same reason that it is interesting that systems exhibit Bell non-classicality—because of the separation in correlations from an identical classical system—the broader motivation in working with contextuality is to demonstrate some form of computational advantage. In this direction, two rather important bodies of work have been [10] and [119]. In [10] Anders and Brown demonstrated

that certain input states can raise the computational power of the classical control computer in measurement based quantum computing (MBQC). In particular, it was shown that deterministic computation of a NAND gate is possible given access to a GHZ resource state and a classical control computer which is only able to perform CNOT and XOR operations. This is an explicit demonstration of promotion of the classical computational power from subuniversal (or linear) to universal, since the NAND gate is nonlinear. On the other hand, Raussendorf showed that contextuality has a role to play by demonstrating that MBQCs which compute a non-linear Boolean function with high probability are contextual.

3.1 The setup

We will be studying a single restricted type of setting in which we have access to a black-box, into which we feed inputs and it returns outputs. The input specifies the measurement to be made, and the output is the outcome of the performed measurement. Inside the black box lies a system $\mathcal{S}$ which is what is measured. The setup is indicated in diagrammatic form in Figure 3.1 for clarity.

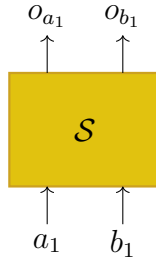


Figure 3.1: A system $\mathcal{S}$, represented by a ‘black’ box. The measurements a_1 and b_1 are fed into the box, indicating that they are performed on the system, and corresponding outcomes o_{a_1} and o_{b_1} are produced by the box.

Here, commuting measurements a_1 and b_1 are fed into the box, which can be thought of measurements a_1 and b_1 being performed on the system. The results of measurement are then fed out as o_{a_1} and o_{b_1} respectively ^[1]. A single run of the experiment is therefore both a choice of measurements and the receiving of an outcome for each of these choices. The question is then whether, over many runs, the resulting probability distribution on outcomes which $\mathcal{S}$ produces has classical explanation.

¹An alternative viewpoint, which will be useful later on, is that we ‘ask’ the system $\mathcal{S}$ questions a_1 and b_1 , and the outputs are the system’s answers to these questions.

In each round of experiment a commuting set of measurements must be chosen. To this end we introduce a simplicial complex Σ on the set X of possible measurements, such that a face $C \in \Sigma$ indicates a commuting set of measurements— C will be referred to as a **measurement context**. For example, suppose that the set of measurements is $X = \{1, 2, 3, 4\}$ and the simplicial complex $\Sigma := \mathcal{P}(X)$. Thus $\{1, 2, 3, 4\} \in \Sigma$ and all measurements commute. We will however be interested in setups in which some of the measurements do not commute, since it is the non-commutativity, and in particular the measurement of a measurement m in more than one measurement context, which allows for the possibility of contextuality.

The last piece of information which must be specified is the outcome set O_m belonging to each measurement m . For clarity of notation and without loss of generality, we will take O_m to be identical for every measurement m and given by O . Oftentimes we will take $O := \{0, 1\}$ in which case the measurement is said to be dichotomic. It follows that a **measurement scenario** can be summarised by the tuple $\mathcal{M} = \langle X, \Sigma, O \rangle$.

Definition 3.1.1 (Measurement Scenario). *A measurement scenario is a triple $\mathcal{M} = \langle X, \Sigma, O \rangle$ where*

1. X is the set of measurements,
2. $\Sigma \subseteq \mathcal{P}(X)$ is the commutation relation on the set of measurements,
3. O is the set of outcomes for each measurement.

We consider as an example the tripartite setup, which is a Bell-type example arising when three agents, A , B and C , each have two measurements which they can perform.

Example 3.1.2 (Tripartite Setup). *The tripartite setup has measurement set $X = \{a_0, a_1, b_0, b_1, c_0, c_1\}$ and simplicial complex displayed in Figure [3.2](#).*

Maximal faces of Σ consist of a single choice of measurement for each agent. Note, moreover, that every measurement occurs in more than one maximal face of Σ . Oftentimes measurement are dichotomic so that $O = \{0, 1\}$.

A measurement scenario does not make reference to the theory of quantum mechanics, but it encapsulates this as a specific example. In this case, specifying the measurements amounts to for each $m_i \in X$ specifying a **projection-valued measure** (PVM) [\[47\]](#) on the Hilbert space $\mathcal{H}$. Recall that a PVM consists of a set of

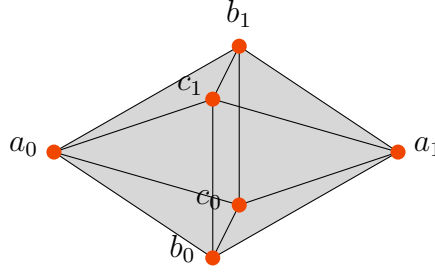


Figure 3.2: The compatibility complex for a three party Bell setup, in which each party has access to two measurements.

projectors $M := \{P_j\}_{j \in O}$ with j denoting the outcome associated with the projector, such that (i) projectors are orthogonal, $[P_i, P_j] = 0$ and (ii) they sum to give a resolution of the identity, $\sum_{j \in O} P_j = I$.

Let $\mathbf{M}$ be a list of PVMs, then a triple $\langle \mathbf{M}, \Sigma, O \rangle$ is a **quantum measurement scenario**, where $\sigma \in \Sigma \iff \forall M_i, M_j \in \Sigma. [M_i, M_j] = 0$. We show how the well-studied Bell-CHSH setup [42] can be captured in this framework.

Example 3.1.3 (The Bell-CHSH Setup). *The CHSH Setup has measurement scenario $\mathcal{M} = \langle X, \mathcal{M}, O \rangle$ where*

1. $X = \{a_0, a_1, b_0, b_1\}$
2. $\Sigma = \{\{a_0, b_0\}, \{a_0, b_1\}, \{a_1, b_0\}, \{a_1, b_1\}, \{a_0\}, \{a_1\}, \{b_0\}, \{b_1\}\}$
3. $O = \{0, 1\}$

The simplicial complex Σ is shown below.

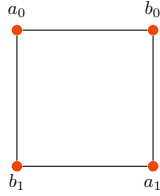


Figure 3.3: Compatibility complex of the Bell-CHSH setup.

Here, a_i indicates a measurement performed by a party Alice and b_i a measurement by a party Bob. We may be more specific and give the corresponding quantum

realisations. Let M_{a_0} be measurement in the computational Z basis, so that it consists of projectors

$$\begin{cases} P_1^{a_0} = |0\rangle \langle 0| \\ P_{-1}^{a_0} = |1\rangle \langle 1| \end{cases} \quad (3.1)$$

and M_{a_1} be in the X or Hadamard basis so that the projectors are instead

$$\begin{cases} P_1^{a_1} = |+\rangle \langle +| \\ P_{-1}^{a_1} = |-\rangle \langle -| \end{cases} \quad (3.2)$$

We define projectors for $b_{0,1}$ such that b_0 lies at an angle θ

$$\begin{cases} P_1^{b_0} = [\cos \theta |0\rangle + \sin \theta |1\rangle] [\cos \theta \langle 0| + \sin \theta \langle 1|] \\ P_{-1}^{b_0} = [-\sin \theta |0\rangle + \cos \theta |1\rangle] [-\sin \theta \langle 0| + \cos \theta \langle 1|] \end{cases} \quad (3.3)$$

and

$$\begin{cases} P_1^{b_1} = [\cos (\pi/2 - \theta) |0\rangle + \sin (\pi/2 - \theta) |1\rangle] [\cos (\pi/2 - \theta) \langle 0| + \sin (\pi/2 - \theta) \langle 1|] \\ P_{-1}^{b_1} = [-\sin (\pi/2 - \theta) |0\rangle + \cos (\pi/2 - \theta) |1\rangle] [-\sin (\pi/2 - \theta) \langle 0| + \cos (\pi/2 - \theta) \langle 1|] \end{cases} \quad (3.4)$$

Note that the PVMs M_{a_i} and M_{b_j} act on different subsets of the Hilbert space, so that the required commutation relations are automatically satisfied.

The above example is, as with the tripartite setup, a Bell setup, or a setup in which measurements are performed by parties who each have access to some factor $\mathcal{H}_i$ of the total Hilbert space $\mathcal{H}$, such that $\mathcal{H} = \bigotimes \mathcal{H}_i$. In examples such as this, experimenters independently choose a single measurement to perform, so that the compatibility structure as specified by Σ will naturally contain, in its maximal faces, one measurement for each experimenter. This is clear when studying the simplicial complexes of the tripartite and bipartite (i.e. Bell-CHSH) setups.

This specific structure of Bell setups can alternatively be seen as a consequence of the way in which measurements scenarios compose in parallel. Given a setup $\mathcal{M}_1$ with compatibility structure specified by Σ_1 and a setup $\mathcal{M}_2$ with compatibility structure defined by Σ_2 , the parallel composition defines how to produce a larger

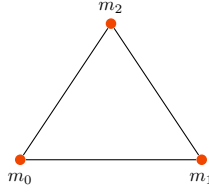


Figure 3.4: A compatibility complex for three measurements known as the Kochen-Specker triangle. It describes the case in which any two measurements are compatible, but not the set $\{m_0, m_1, m_2\}$.

measurement scenario $\mathcal{M}$ by placing $\mathcal{M}_1$ and $\mathcal{M}_2$ in parallel.

Definition 3.1.4 (Parallel Composition of Measurement Scenarios). *Given a measurement scenario $\mathcal{M}_1 = \langle X_1, \Sigma_1, O_1 \rangle$ and a measurement scenario $\mathcal{M}_2 = \langle X_2, \Sigma_2, O_2 \rangle$ define their parallel composition $\mathcal{M}_1 * \mathcal{M}_2$ to be the measurement scenario $\mathcal{M} = \langle X, \Sigma, O \rangle$ where*

1. $X := X_1 \cup X_2$ (assume X_1 and X_2 are disjoint w.l.o.g),
2. $\Sigma := \Sigma_1 * \Sigma_2$ where $\Sigma_1 * \Sigma_2$ denotes the simplicial join of the complexes,
3. $O_m = \begin{cases} O_1 & \text{when } m \in X_1 \\ O_2 & \text{when } m \in X_2. \end{cases}$

All Bell scenarios can be produced from the parallel composition of measurement scenarios for the individual ‘parties’, which have compatibility relations described by discrete simplicial complexes. Contextuality, on the other hand, encompasses all measurement scenarios. The simplicial complex in Figure 3.4, for example, cannot be realised by any Bell setup.

This is known as the Kochen-Specker triangle, and the associated measurement scenario is the simplest—in terms of number of measurements—to exhibit contextuality, although such contextuality is not quantum realisable^[2]. This is because if the PVMs in a set σ commute pairwise then the set is a compatible set of measurements, $\sigma \in \Sigma$. As a result, since $\{M_0, M_1, M_2\}$ commute pairwise, we must have the face $\{M_0, M_1, M_2\} \in \Sigma$ which prohibits the possibility of the ‘empty triangle’ displayed above.

Defining composition is useful for developing resource theories, which in this case would be the resource theory of contextuality. Later we will see that if it can be

²Note that if one allows for more general quantum measurements, the **positive operator-valued measures** (POVMs) then such correlations are achievable [\[97\]](#)

shown that contextuality does not increase under a choice of composition operator, then that operation is free in the associated resource theory.

3.2 Deterministic systems and classicality

We have introduced measurement scenarios, which capture a particular experimental setup. Given access to a black box $\mathcal{S}$ and a measurement scenario $\mathcal{M}$, we can **run** an experiment in which a set of commuting measurements is chosen, fed into the system, and a set of outcomes obtained. We indicate this for the CHSH setup below. We can model this behaviour of the system $\mathcal{S}$ as the function $s_\sigma : \sigma = \{a_1, b_1\} \rightarrow O$

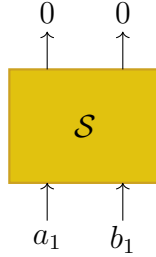


Figure 3.5: Diagrammatic representation of the behaviour that might be exhibited by a black box in a single round of experiment.

defined as

$$s_\sigma :: \{a_1 \mapsto 0; b_1 \mapsto 0\}. \quad (3.5)$$

We call the assignment of an outcome o to a measurement m a **measurement event**, and write $\langle m, o \rangle$ to denote it. Given a subset of measurements $U \subseteq X$, we can represent the simultaneous allocation of an outcome to each measurement in U by the function

$$s : U \rightarrow O \quad (3.6)$$

so that $s(m)$ is the outcome upon measurement of $m \in X$. There will be $|O|^{|U|}$ such functions, and we let $\mathcal{E}(U)$ denote this set of functions. We will refer to functions $s \in \mathcal{E}(U)$ as **sections**, and the set $\mathcal{E}(U)$ as **the set of sections over U** .

Note that the given $\mathcal{E}$ is in fact a presheaf. In particular, recalling definition [2.1.1](#), for each inclusion of measurement sets $V \subseteq U$, we define a restriction map $\rho_V^U : \mathcal{E}(U) \rightarrow \mathcal{E}(V)$ which sends $s \in \mathcal{E}(U)$ to the function $s|_V \in \mathcal{E}(V)$ with $s|_V(m) := s(m)$. It can be checked that this satisfies both conditions in definition [2.1.1](#).

Recall that a sheaf is a presheaf which satisfies additional gluing conditions. It can be checked that $\mathcal{E}$ does satisfy these additional conditions— as a result given $U \subseteq$

X and a cover $\{U_i\}_{i \in I}$ with $s_i \in \mathcal{E}(U_i)$ for each $i \in I$ such that $s_i|_{U_i \cap U_j} = s_j|_{U_i \cap U_j}$, we can always glue these together to form a unique $f \in \mathcal{E}(U)$ with $f|_{U_i} = s_i$ for all $i \in I$.

Physically, taking Σ , the sets of observables which can be simultaneously measured, to be the cover and $U = X$, and letting s_σ correspond to a deterministic mapping of $\sigma \in \Sigma$, $\mathcal{E}$ being a sheaf says that we can always build a function $f : X \rightarrow O$ which restricts down to give the observed behaviour on each $\sigma \rightarrow O$. Note that this is independent of Σ , and $\mathcal{E}$ is a sheaf regardless of the chosen cover.

We will see, when considering a probabilistic system $\mathcal{S}$, this is not the case, and that this is precisely the distinction between classical and contextual systems.

3.3 Probabilistic systems and empirical models

Sections describe deterministic behaviours of a system. However, systems, including classical systems, do not need to be deterministic. To deal with this, we precompose the event sheaf $\mathcal{E}$ with the distribution functor $\mathcal{D}_R$, which was defined in Section 2.3. We therefore obtain, for each subset of measurements $U \subseteq X$, the set of all normalised probability distributions on the set of sections on U .

Recall that given a function $f : X \rightarrow Y$ between sets, the functor $\mathcal{D}_R$ sends this to a map which marginalises the probability distribution associated with X to that associated with Y :

$$\mathcal{D}_R(f) : \mathcal{D}_R(X) \rightarrow \mathcal{D}_R(Y) :: p \mapsto \left[y \mapsto \sum_{f(x)=y} p(x) \right]. \quad (3.7)$$

Therefore given $U, V \subseteq X$ where $U \subseteq V$, and given a distribution $p \in \mathcal{D}_R \circ \mathcal{E}(V)$, we can restrict this to get a probability distribution $p|_U \in \mathcal{D}_R \circ \mathcal{E}(U)$, where we define, for each $f \in \mathcal{E}(U)$

$$p|_U(f) := \sum_{\substack{g \in \mathcal{E}(V) \\ g|_U = f}} p(g). \quad (3.8)$$

Empirically, we only obtain certain probability distributions: one for each commuting set of measurements $\sigma \in \Sigma$. This is in practice implemented by preparing an unlimited (or very large) number of identically prepared systems $\mathcal{S}_1, \mathcal{S}_2, \dots$ and measuring each in turn, for varying choices of measurement set. The systems $\mathcal{S}_i$ may be classical, quantum, or more generally any system which satisfies the axioms of a **general probabilistic theory** (GPT) [17].

In this manner a table of probabilities can be built up, such as the one below (which is for the CHSH setup).

		(0, 0)	(0, 1)	(1, 0)	(1, 1)
a_0	b_0	1/2	0	0	1/2
a_1	b_0	1/2	0	0	1/2
a_0	b_1	1/2	0	0	1/2
a_1	b_1	0	1/2	1/2	0

Table 3.1: Example empirical model for the CHSH setup.

Note, however, that the information contained in a table such as Table 3.1 is identical to taking a compatible family $\{e_\sigma\}_{\sigma \in \Sigma}$ for the presheaf $\mathcal{D}_R \circ \mathcal{E}$ over the cover Σ . Each e_σ in this family is contained in the row labelled by σ in the table. For example, the $\frac{1}{2}$ in the row ' $a_0 \ b_0$ ' and column (0,0) indicates that the local section $s : \{a_0 \mapsto 0, b_0 \mapsto 0\}$ occurs with probability $\frac{1}{2}$, or $e_{\{a_0, b_0\}}(s) = 1/2$.

Tables such as the one above are referred to as **empirical models**.

Definition 3.3.1 (Empirical Model). *Let $\mathcal{M} = \langle X, \Sigma, O \rangle$ be a contextuality measurement scenario. An empirical model e consists of a compatible family of probability distributions $\{e_\sigma\}_{\sigma \in \Sigma}$ on the cover Σ for the presheaf $\mathcal{D}_R \circ \mathcal{E}$. Compatibility means that given e_σ and $e_{\sigma'}$ then we should have*

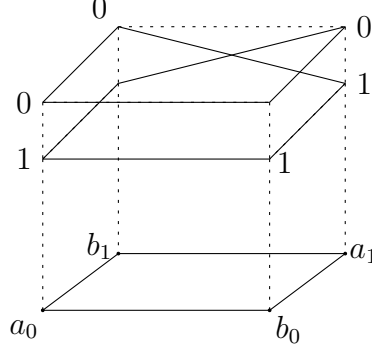
$$e_\sigma|_{\sigma \cap \sigma'} = e_{\sigma'}|_{\sigma \cap \sigma'}. \quad (3.9)$$

We write $\text{EM}(\mathcal{M})$ to represent the set of all empirical models on the scenario $\mathcal{M}$.

The condition that the family of probability distributions $\{e_\sigma\}_{\sigma \in \Sigma}$ be compatible is the **no-disturbance condition** [118], which, for Bell scenarios, is the **no-signalling** condition. It says that the probability of obtaining an outcome for a measurement m should not depend on which other measurements it is performed together with. For Bell setups this can be enforced by ensuring that measurements are space-like separated; for contextuality setups since measurements could have support on overlapping subsets of the Hilbert space, no-disturbance cannot be physically guaranteed but we take it as an axiom. The definition of empirical model therefore guarantees that no-signalling, and more generally no-disturbance, is satisfied.

The data of an empirical model may also be represented via a **bundle diagram**. The bundle diagram for the empirical model given in Table 3.1 is shown below. The base of the bundle diagram is the simplicial complex Σ corresponding to the scenario $\mathcal{M}$. The fibres above each $\sigma \in \Sigma$ are the set of sections $\mathcal{E}(\sigma)$. The empirical model

is realised by representing a dark simplex for a section $s \in \mathcal{E}(\sigma)$ which is observed. In particular, for the CHSH setup, since each maximal context contains just two measurements, the empirical model given in table 3.1 simply has a line connecting $a_0 = i$ and $b_0 = j$ along the edge $\{a_0, b_0\}$ if the column (i, j) in the row $\{a_0, b_0\}$ occurs with non-zero probability.



As we had the quantum realisation of a particular measurement scenario, we may have the quantum realisation of an empirical model. Given a quantum state $|\phi\rangle$, the data for an empirical model is specified for each $\sigma \in \Sigma$ by calculating

$$e_\sigma(s) = \langle \phi | \left(\prod_{M_i; i \in \sigma} P_{s(i)}^i \right) | \phi \rangle. \quad (3.10)$$

Recall that P_o^i is the projection corresponding to the outcome o for the measurement M_i .

3.4 Contextuality

We define an empirical model as a compatible family of probability distributions for the presheaf $\mathcal{D}_R \circ \mathcal{E}$. For certain covers, and therefore for certain empirical models, the sheaf condition will be met so that there exists a global section projecting down to give the local sections of the empirical model. This is precisely what is meant by contextuality.

Definition 3.4.1 (Contextuality). *A model $e \in \mathbf{EM}(\mathcal{M})$ is contextual if there exists no global probability distribution $d \in \mathcal{D}_R \circ \mathcal{E}(\Sigma)$ such that*

$$e_\sigma := d|_\sigma \quad (3.11)$$

for all measurement contexts $\sigma \in \Sigma$.

Contextuality can therefore be interpreted as a property of empirical models which contain data that cannot be glued together into a consistent global picture. We will oftentimes say that the empirical model is **extendable** if a global section does exist. In [6], it is moreover shown that since global sections can be viewed as deterministic local hidden variables (LHVs), contextuality of an empirical model implies that no such hidden variable description for it exists.

Besides deciding whether an empirical model is contextual or not, empirical models which are not extendable can also be classified according to the strength of contextuality. To emphasise this, and before characterising these different types of contextuality, we consider two examples below. The difference in proofs of contextuality in these examples highlights these different ‘levels’ of contextuality.

Example 3.4.2 (Peres-Mermin’s ‘magic’ square). *The Peres-Mermin magic square [115, 107, 106] refers to the ‘square’ below, where each row and each column is a measurement context and the σ_α s are the Pauli matrices. Also shown is the result of multiplication of each row and column; note that multiplication of every row and the first two columns gives the identity operator 1 while multiplication of the last column gives -1 .*

$1 \otimes \sigma_x$	$\sigma_x \otimes 1$	$\sigma_x \otimes \sigma_x$	1
$\sigma_y \otimes 1$	$1 \otimes \sigma_y$	$\sigma_y \otimes \sigma_y$	1
$\sigma_y \otimes \sigma_x$	$\sigma_x \otimes \sigma_y$	$\sigma_z \otimes \sigma_z$	1
1	1	-1	

The reason this is contextual is the following: there exists no assignment $s : X \rightarrow \{-1, 1\}$ where X is the set of 9 measurements in the table, such that the product of these assignments gives the value expected from multiplication of the appropriate row or column. Taking row 1 as an example, we require that

$$s(1 \otimes \sigma_x)s(\sigma_x \otimes 1)s(\sigma_x \otimes \sigma_x) = 1. \quad (3.12)$$

This can be satisfied by, for example, taking

$$s(1 \otimes \sigma_x) = s(\sigma_x \otimes 1) = s(\sigma_x \otimes \sigma_x) = 1. \quad (3.13)$$

However all 6 of the equations corresponding to the 3 rows and 3 columns cannot be simultaneously satisfied in this manner.

To see this, label row i R_i and column j C_j . Then note that

$$R_1 \cdot R_2 \cdot R_3 \cdot C_1 \cdot C_2 \cdot C_3 = -1, \quad (3.14)$$

however, the left hand side of this equation is equal to

$$[s(1 \otimes \sigma_x)]^2 [s(\sigma_x \otimes 1)]^2 [s(1 \otimes \sigma_y)]^2 \dots [s(\sigma_z \otimes \sigma_z)]^2. \quad (3.15)$$

This quantity must be equal to 1, since each $[s(-, -)]^2$ term must equal 1. This together with equation [3.14](#) gives a contradiction, because they imply $1 = -1$.

		(0, 0)	(0, 1)	(1, 0)	(1, 1)
a_0	b_0	1/2	0	0	1/2
a_1	b_0	3/8	1/8	1/8	3/8
a_0	b_1	3/8	1/8	1/8	3/8
a_1	b_1	1/4	1/4	1/4	1/4

Table 3.2: The Bell-CHSH empirical model.

Example 3.4.3 (Contextuality of the Bell-CHSH model). *We consider the Bell-CHSH empirical model displayed in the table [3.2](#). It arises as follows. Two qubits are prepared in the entangled state*

$$|\phi^+\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle) \quad (3.16)$$

and one distributed to Alice and one to Bob. Each agent has access to the same two measurements (albeit on their factor of the Hilbert space), which are the following two,

$$a_0 = b_0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad a_1 = b_1 = \begin{pmatrix} 0 & e^{-i\pi/3} \\ e^{i\pi/3} & 0 \end{pmatrix}. \quad (3.17)$$

Physically this corresponds to two measurements separated by an angle of $\frac{\pi}{3}$ on the X-Y plane of the Bloch sphere. Noncontextuality of this empirical model requires that there be a probability distribution over global sections marginalising to give the data in the table. An example global section is

$$f :: \{a_0 \mapsto 0, a_1 \mapsto 1, b_0 \mapsto 0, b_1 \mapsto 1\}. \quad (3.18)$$

There are $2^4 = 16$ such global sections. An example of a probability distribution over such sections is

$$1/8 \vec{f}_1 + 1/8 \vec{f}_2 + 3/8 \vec{f}_3 + 3/8 \vec{f}_4 \quad (3.19)$$

where

$$\left\{ \begin{array}{l} f_1 :: \{a_0 \mapsto 0, a_1 \mapsto 0, b_0 \mapsto 0, b_1 \mapsto 1\} \\ f_2 :: \{a_0 \mapsto 0, a_1 \mapsto 1, b_0 \mapsto 0, b_1 \mapsto 0\} \\ f_3 :: \{a_0 \mapsto 1, a_1 \mapsto 0, b_0 \mapsto 1, b_1 \mapsto 1\} \\ f_4 :: \{a_0 \mapsto 1, a_1 \mapsto 1, b_0 \mapsto 1, b_1 \mapsto 0\}. \end{array} \right. \quad (3.20)$$

Nevertheless this choice does not marginalise to give the desired empirical data. It can be shown, via contradiction, that no choice of global section produces the data in Table 3.2, so that this is an example of a contextual model.

In the first example, contextuality arose by showing that there exists *no* global section consistent with the observed data. That is, given a probabilistic empirical model $e \in \mathbf{EM}(\mathcal{M})$, we associate with it a possibilistic model $\text{poss}(e)$ by taking for each $\sigma \in \Sigma$, the possibilistic model in the contexts to be $\mathbf{supp}(e_\sigma)$. **Possibilistic extendability**, then, refers to whether there is a Boolean distribution d on $\mathcal{E}(X)$ which restricts to $\mathbf{supp}(e_\sigma)$ for every context $\sigma \in \Sigma$. This is equivalent to specifying a non-empty subset S_e of $\mathcal{E}(X)$:

$$S_e := \{s \in \mathcal{E}(X) \mid \forall \sigma \in \Sigma. s|_\sigma \in \mathbf{supp}(e_\sigma)\}. \quad (3.21)$$

By deriving a contradiction when supposing that such a global section does exist, we showed for the Peres-Mermin square that $S_e = \emptyset$ and therefore the setup is strongly contextual.

The second example instead demonstrated contextuality by showing that there exists no global probability distribution which marginalises to give the observed local distributions, which is a probabilistic proof of contextuality. The difference between these two proofs is a result of the following **hierarchy of contextuality**.

Definition 3.4.4 (Hierarchy of Contextuality). *Let e be a probabilistic empirical model on $\mathcal{M}$. We say that e is*

1. **Contextual** if there exists no global section $h \in \mathcal{D}_R \circ \mathcal{E}(X)$ restricting to give the local sections e_σ for all $\sigma \in \Sigma$.

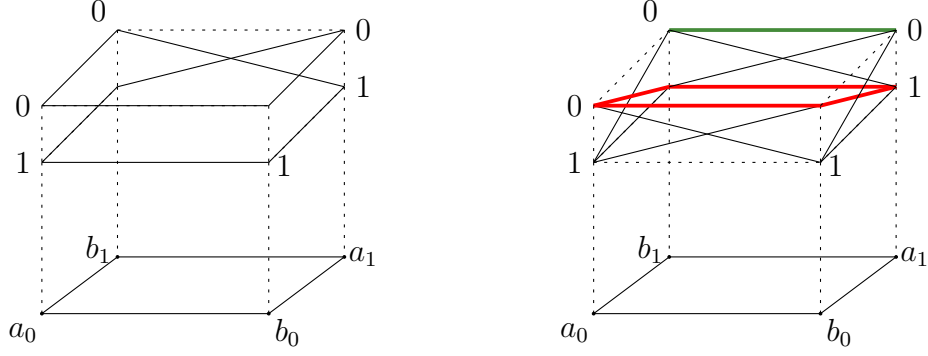


Figure 3.6: Example bundle diagrams. *Left.* The PR Box. *Right.* The Hardy model.

2. **Logically contextual** if there is an $s \in \text{supp}(e_\sigma)$ which does not arise as the restriction of a consistent global assignment $g \in S_e$.
3. **Strongly contextual** if $S_e = \emptyset$.

We express the hierarchy as follows

$$\text{Contextuality} < \text{Logical contextuality} < \text{Strong contextuality} \quad (3.22)$$

indicating that, for example, logical contextuality $\implies$ contextuality, and strong contextuality $\implies$ contextuality.

Bundle diagrams are useful for visualising logical and strong contextuality. Since a global section in the bundle diagram is a closed path traversing all the fibres once, logical contextuality is present when there exists a local section, in this case a line, which does not extend to a closed path and strong contextuality is present when no line extends to a path. We display the bundle diagrams of the PR box (left) and Hardy model (right) below. Both of the associated empirical models are contextual: for the PR box, there exists no closed path in the bundle diagram, while for the Hardy model there is closed path (highlighted in red), but there is also a line which does not extend into a closed path (highlighted in green). The PR box is strongly contextual and the Hardy model logically contextual.

3.5 Kochen-Specker scenarios and the Kochen-Specker Theorem

Contextuality is a property of an empirical model $e \in \text{EM}(\mathcal{M})$ defined on a measurement scenario $\mathcal{M}$. It has been shown that there are various ways to prove that a given empirical model is contextual, and such methods vary depending on the type of contextuality.

In certain cases, given just the measurement scenario $\mathcal{M}$ it is possible to guarantee that every state of the system (i.e. for quantum mechanics every wavefunction $|\psi\rangle$), when measured, will result in a contextual empirical model. This is known as **state-independent contextuality** [34]. The Peres-Mermin magic square, for example, gives a contradiction regardless of the state $|\psi\rangle$ that is measured.

Kochen-Specker scenarios are a class of scenarios for which contextuality is always state-independent strong contextuality. We describe such scenarios first in the quantum mechanical setting, where they naturally arise, and then explain how these are captured in the sheaf approach developed above.

In quantum mechanics, an idealised measurement is described by a rank-1 projector on a Hilbert space. For a Hilbert space of n -dimensions, any set $\{P_1, \dots, P_n\}$ of orthogonal rank-1 projectors satisfies

1. $P_i P_j = 0$ for $i \neq j$;
2. $\sum_{i=1}^n P_i = 1$.

Each projector can be viewed as a ray spanning a vector $\mathbf{v}_i$ in $\mathbb{R}^n$ [3] and a projector will be associated with either a ‘yes’ answer (1 outcome) or ‘no’ answer (0 outcome). In this terminology, given a set $\mathcal{B} \subseteq \mathbb{R}^n$ of n orthogonal projectors items 1 and 2 imply that given a state (i) orthogonal projectors cannot both give 1 (yes) upon measurement and (ii) at least one of the projectors must give 1 (yes) upon measurement. Thus there exists a map $f : \mathcal{B} \rightarrow \{0, 1\}$ with $\sum_{\mathbf{v} \in \mathcal{B}} f(\mathbf{v}) = 1$.

The question that Kochen and Specker asked, and ultimately showed not to be true in their 1967 paper [91], is whether one can given a set $S \subseteq \mathbb{R}^n$ of vectors always find an assignment $f : S \rightarrow \{0, 1\}$ such that for every set $\mathcal{B}$ of orthogonal vectors contained in S , $\sum_{\mathbf{v} \in \mathcal{B}} f(\mathbf{v}) = 1$.

The above result, known as the Kochen-Specker Theorem, can be viewed, as with empirical models, as an obstruction in extending truth assignments of these projectors in a set of orthogonal vectors $\mathcal{B}$ to truth assignments on the whole set S .

³In general the Hilbert space is a complex vector space, but we specialise to the real case.

If no such extension exists, the set S is called a Kochen-Specker set, or simply a KS set. The original proof of the theorem relied on a set of 117 vectors. Since then, the smallest KS set is the 18 vector set in 4 dimensions introduced in [35], which was shown to be minimal in [145].

KS setups and the KS models can be seen as specific instances of the machinery developed by Abramsky and Brandenburger in [7]. A KS setup contains an equal number of measurements per context. Thus, for a KS scenario $\langle X, \Sigma, O \rangle$ we have

$$\forall \sigma \in \max(\Sigma). |\sigma| = n, \quad (3.23)$$

for some fixed integer n . A KS empirical model is one which has support only on the sections which send a single measurement in a context to the outcome 1, and everything else to 0. Thus,

$$s \in \text{supp}(e_\sigma) \implies \exists m \in \sigma. s(m) = 1 \wedge \forall m' \in \sigma \setminus \{m\}. s(m') = 0. \quad (3.24)$$

If an empirical model $e \in \text{EM}(\mathcal{M})$ satisfies equation [3.24] we write e^{KC} ; we denote as $\text{EM}^{\text{KC}}(\mathcal{M})$ the set of all Kochen-Specker models on $\mathcal{M}$.

Example 3.5.1 (The Kochen-Specker Triangle). *The simplest contextual KS model is ‘the contextual triangle’, which is displayed in the table below, alongside the bundle diagram which makes clear strong contextuality of the model.*

Note that the only possible sections in each maximal context are those which map a single measurement to the outcome 1.

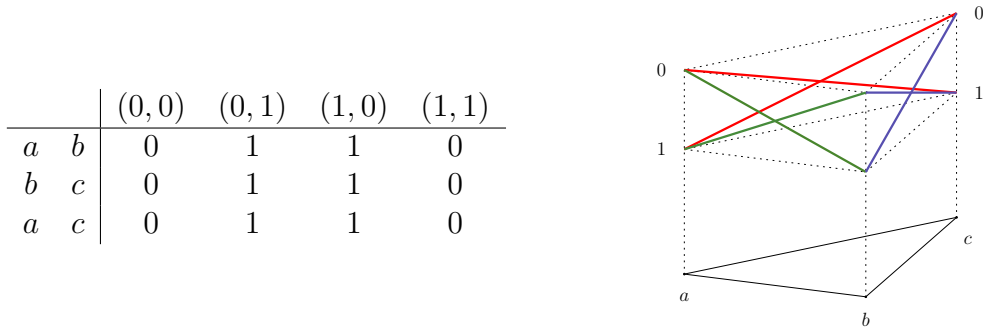


Figure 3.7: The empirical model and bundle diagram for Example [3.5.1].

3.6 Simulation and quantum resources

Resource theories have proved to be fundamental in quantum information [39]. A resource theory consists of a set of states which are considered free together with a set of allowed or free operations, these both being dependent upon the particular resource theory at hand. Importantly, the set of free states is invariant upon application of the free resource—the free resource always converts a free state into another free state—and the resource states are then those which cannot be generated starting from a free state and using free operations. In order to quantify this resourcefulness, one introduces a monotone map from the set of states to the real numbers, such that this number cannot increase when a state is acted upon by the free operations of the theory. For example, in the resource theory of entanglement, free operations correspond local operations and classical communication (LOCC), and the resource states are entangled states which cannot be generated by LOCC starting from a separable (i.e. unentangled) state.

Here, we consider the resource theory of contextuality and non-locality and how this relates to simulation of one empirical model by another as studied in [3] and [81]. This includes introducing the contextual fraction as a measure of contextuality, which will be the resource monotone of the theory, as well as defining the allowed operations on empirical models [4].

We begin by considering operations on empirical models. We write $e : \langle X, \Sigma, O \rangle$ when e is an empirical model on $\langle X, \Sigma, O \rangle$. The below operations are considered ‘free’ because they can be classically implemented, i.e. they require no additional contextual resources.

Probabilistic mixing. A new empirical model is produced from a probabilistic mix of $e : \langle X, \Sigma, O \rangle$ and $e' : \langle X, \Sigma, O \rangle$ with probabilistic weight $\lambda \in [0, 1]$. This corresponds to tossing a weighted coin to choose whether to run e or e' . The new empirical model $e +_\lambda e' : \langle X, \Sigma, O \rangle$ is given by

$$(e +_\lambda e')_\sigma(s) := \lambda e_\sigma(s) + (1 - \lambda) e'_\sigma(s), \quad (3.25)$$

for $\sigma \in \Sigma$ and $s \in \mathcal{E}(\sigma)$.

Controlled choice Given two empirical models on distinct measurement scenarios $e : \langle X, \Sigma, O \rangle$ and $e' : \langle X', \Sigma', O' \rangle$ a new measurement scenario $e \& e' : \langle X \sqcup X', \Sigma + \Sigma', [O, O'] \rangle$ is defined. Here, $X \sqcup X'$ is the disjoint union of measurement sets, and $\Sigma + \Sigma'$ is

the coproduct of measurement scenarios. Moreover, $[O, O']$ indicates that

$$\begin{aligned} [O, O']_x &= O_x \quad ; \quad x \in X \\ [O, O']_x &= O'_x \quad ; \quad x \in X'. \end{aligned} \tag{3.26}$$

We define the empirical model on this new scenario in terms of e and e' as

$$\begin{aligned} (e \& e')_{\{0\} \times \sigma} &:= e_\sigma \quad ; \quad \sigma \in \Sigma \\ (e \& e')_{\{1\} \times \sigma} &:= e'_\sigma \quad ; \quad \sigma \in \Sigma' \end{aligned} \tag{3.27}$$

so that the compatibility of the new scenario ensures that only one of the empirical models is run.

Tensor Product Given two empirical models $e : \langle X, \Sigma, O \rangle$ and $e' : \langle X', \Sigma', O' \rangle$ a new one is formed by running these two simultaneously. The measurement scenario is $e \otimes e' : \langle X \sqcup X', \Sigma * \Sigma', [O, O'] \rangle$ where now the simplicial join is instead performed. A face $\sigma \in \Sigma * \Sigma'$ may consist of both measurements from X and X' . Taking $\sigma \sqcup \sigma'$, $s \in \mathcal{E}(\sigma)$ and $s' \in \mathcal{E}(\sigma')$, define the tensor product empirical model

$$(e \otimes e')_{\sigma \sqcup \sigma'}((s, s')) := e_\sigma(s) e'_{\sigma'}(s'). \tag{3.28}$$

We have mentioned that a resource theory requires a measure over the states which is monotone for the allowed operations. The measure is taken to be the **contextual fraction** CF. This is defined as follows. Any empirical model $e : \langle X, \Sigma, O \rangle$ can be decomposed as

$$e = e^{\text{NC}} +_\lambda e' \tag{3.29}$$

(see equation [3.25](#)) with e^{NC} noncontextual and for some $\lambda \in [0, 1]$. Such decompositions are not unique, and so there will be many possible values of λ — the non-contextual fraction $\text{NCF}(e)$ is the largest value of λ for which equation [3.29](#) holds, $\text{NCF}(e) = \max\{\lambda | e^{\text{NC}} +_\lambda e'\}$.

The contextual fraction is

$$\text{CF}(e) = 1 - \text{NCF}(e), \tag{3.30}$$

which is monotone under the operations defined before. In particular, [4](#)

1. $\text{CF}(e +_\lambda e') \leq \lambda \text{CF}(e) + (1 - \lambda) \text{CF}(e')$
2. $\text{CF}(e \& e') = \max\{\text{CF}(e), \text{CF}(e')\}$
3. $\text{CF}(e \otimes e') = \text{CF}(e) + \text{CF}(e') - \text{CF}(e)\text{CF}(e')$

The resource theory framework setup above is one way to speak about ‘resourcefulness’ of an empirical model, in terms of the contextual fraction.

One can instead grade empirical models in terms of simulations, where one empirical model is considered more ‘powerful’, so to speak, than a second one if it can simulate this latter one. Consider two measurement scenarios $\mathcal{M}_X = \langle X, \Sigma, O \rangle$ and $\mathcal{M}_Y = \langle Y, \Theta, P \rangle$.

A morphism $\langle \pi, h \rangle : \mathcal{M}_Y \rightarrow \mathcal{M}_X$ consists of

- i The simplicial map $\pi : \Sigma \rightarrow \Theta$ which is a function $\pi : X \rightarrow Y$ which preserves compatibility (see Section [2.2](#)).
- ii The family of maps $h_U : \mathcal{E}_Y(\pi(U)) \rightarrow \mathcal{E}_X(U)$ for each $U \in \mathcal{P}(X)$. Equivalently, a family of maps $h_x : P_{\pi(x)} \rightarrow O_x$ on output sets.

Note that in [81](#), [16](#) a simplicial relation is used instead of a simplicial map, the former being more general. For our purposes, however, a simplicial map will be sufficient. A morphism $\langle \pi, h \rangle : \mathcal{M}_Y \rightarrow \mathcal{M}_X$ describes (i) how to map measurements $m \in X$ to measurements $\pi(x) \in Y$ and (ii) how to map a section $s \in \mathcal{E}_Y(\pi(U))$ for any $U \in \mathcal{P}(X)$ —equivalently a choice of outcomes for the measurements in $\pi(U)$ —to a section $s' \in \mathcal{E}_X(U)$ —equivalently a choice of outcome for the measurements in U .

An experimenter with access to an empirical model $d : \langle Y, \Theta, P \rangle$ can **simulate** an empirical model on $\langle X, \Sigma, O \rangle$: for any choice of measurements $\sigma \in \Sigma$ the experimenter instead performs the measurements $\pi(\sigma)$, obtaining the distribution $d_{\pi(\sigma)}$ on sections. Each of these sections is then mapped back, via h_σ , to sections on σ , resulting in an empirical model $e : \langle X, \Sigma, O \rangle$. This procedure is shown in Figure [3.8](#). We give the formal definition below.

Definition 3.6.1 (Simulation of Empirical Models). *Given measurement scenarios $\mathcal{M}_X : \langle X, \Sigma, O \rangle$ and $\mathcal{M}_Y : \langle Y, \Theta, P \rangle$ and the empirical models $e : \mathcal{M}_X$ and $d : \mathcal{M}_Y$, we say that d deterministically simulates e if there is a morphism $\langle \pi, h \rangle : \mathcal{M}_Y \rightarrow \mathcal{M}_X$ such that*

$$e_\sigma(s) = d_{\pi(\sigma)}(h_\sigma(s)). \quad (3.31)$$

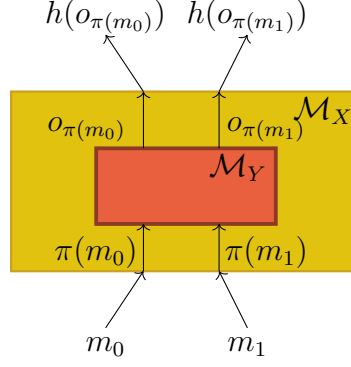


Figure 3.8: A diagrammatic representation of a simulation of an empirical model e on $\mathcal{M}_X$ by an empirical model d on $\mathcal{M}_Y$.

In other words, e is the pushforward along $h_\sigma : \mathcal{E}_Y(\pi(\sigma)) \rightarrow \mathcal{E}_X(\sigma)$ of d , or $e_\sigma = h_\sigma^* d_\sigma$.

One would expect that the contextuality of the two models is important when determining when one model can simulate another. This idea of comparing empirical models based on simulability is an idea which holds in general, but here we will need only the case concerning deterministic simulations and strong contextuality. As such, we give a special case of the relevant theorem from [81] below, which says that an empirical model can only simulate a strongly contextual empirical model if it itself is strongly contextual.

Theorem 3.6.2 (Karvonen [81]). *Let $\langle \pi, h \rangle : d \mapsto e$ be a deterministic simulation. If e is strongly contextual, d is as well.*

Therefore, rather than grading empirical models based on the contextual fraction, one can compare empirical models pairwise based on whether there exists a simulation between them.

We consider a type of simulation which arises when the scenario $\mathcal{M}_X$ is a restriction of the scenario $\mathcal{M}_Y$.

3.6.1 Subscenarios from restriction

Naturally, one can restrict a measurement scenario $\mathcal{M} = \langle X, \Sigma, O \rangle$ to a scenario $\mathcal{M}|_Y = \langle Y, \Sigma|_Y, O \rangle$ on the subset $Y \subseteq X$ of measurements. Here, $\Sigma|_Y := \{\sigma \cap Y \mid \sigma \in \Sigma\}$. Then we have a morphism $\langle i, \text{id} \rangle : \mathcal{M} \rightarrow \mathcal{M}'$ where $i : Y \rightarrow X$ is an inclusion and the map on outcomes is simply the identity, since the outcome set for any $y \in Y$ for the two scenarios is identical.

Therefore, there is a simulation $e \mapsto e|_Y$ which allows one to simulate the restricted empirical model $(e|_Y)_\sigma := e_\sigma$ from the empirical model e . The restricted model is therefore only contextual if the original empirical model is.

3.7 Measurement based quantum computing

Contextuality has so far been considered from quite an abstract perspective, with certain examples arising by considering spatially separated parties who can perform measurements on a shared resource, quantum or otherwise. Here we see how contextuality arises in computation, namely in measurement based quantum computing, a paradigm of quantum computing which utilises quantum measurement to implement an algorithm. The idea is that n parties, each with access to a qubit in a multipartite state $|\Phi\rangle$, perform a measurement in the basis specified by an input q_k , thereby obtaining an outcome s_k . The state $|\Phi\rangle$ is referred to as a **resource state**. For resource states $|\Phi\rangle$ which fall into the category of cluster states, MBQC is a universal model of quantum computation [120]. Moreover, as shown in [11] and [119], contextuality allows one to explicitly increase the computational power of the classical control computer. Anders and Browne in [11] demonstrate that non-locality promotes a subuniversal classical power of Pauli measurements and mod2-linear classical processing to universal. Raussendorf takes this further in [119] by showing that contextuality is necessary in order for an MBQC to compute a non-linear Boolean function with high probability.

We recall an l2-MBQC protocol as defined by Raussendorf in [119], where l2 refers to the mod2-linear constraint imposed on the side of classical processing.

Definition 3.7.1 (l2-MBQC [119]). *An l2-MBQC is a procedure employing local quantum measurements via measurement parties to produce a classical output $\mathbf{o}$ from a classical input $\mathbf{i}$ as follows.*

1. *For each party k , $k = 1..n$, there is a binary choice for the measurement basis, $q_k \in \{0, 1\}$.*
2. *For each party k and each $q_k \in \{0, 1\}$, the measurement outcome is binary-valued, $s_k \in \{0, 1\}$.*
3. *The computational output $\mathbf{o}$ is bitwise a parity of measurement outcomes $\mathbf{s} = (s_1, s_2, \dots, s_n)^T$*

$$\mathbf{o} = Z\mathbf{s} \pmod{2}. \quad (3.32)$$

4. The choice of measurement bases $\mathbf{q} = (q_1, q_2, \dots, q_n)^T$ is related to the measurement outcomes $\mathbf{s}$ and the binary-valued classical input $\mathbf{i} = (i_1, \dots, i_l)^T$ via

$$\mathbf{q} = T\mathbf{s} + Q\mathbf{i} \pmod{2}. \quad (3.33)$$

5. For a suitable ordering of the parties $1..n$, the matrix T is lower triangular with vanishing diagonal.

An l 2-MBQC M therefore takes a vector of inputs $\mathbf{i}$ and produced the computational output $\mathbf{o}$ from this. It can be useful to instead view this as a function

$$o : (\mathbb{Z}_2)^l \rightarrow (\mathbb{Z}_2)^n \quad (3.34)$$

where recall l is the size of $\mathbf{i}$, and n the size of $\mathbf{o}$, so that the function o simply maps a binary input string to a binary output string.

One of the main results of paper [119] is the following theorem.

Theorem 3.7.2 (Raussendorf [119]). *Let M be an l 2-MBQC deterministically evaluating a function $o : (\mathbb{Z}_2)^l \rightarrow \mathbb{Z}_2$. If o is nonlinear, then M is strongly contextual.*

This generalises the result of Anders and Browne, which we will consider as an example below, to any l 2-MBQC which *deterministically* evaluates a function. Also in [119], a more general theorem is given which applies also to non-deterministic functions and does not require strong contextuality.

Theorem 3.7.3 (Raussendorf [119]). *Let M be an l 2-MBQC evaluating a function $o : (\mathbb{Z}_2)^n \rightarrow \mathbb{Z}_2$. Then M is contextual if it succeeds with an average probability $P_S > 1 - \frac{d_H(o)}{2^n}$, where $d_H(o)$ is the Hamming distance of o from the closest linear function.*

Therefore if M is noncontextual, the probability of it being able to evaluate the function o is limited by the Hamming distance of o from some linear function. In particular, if $d_H(o) = 0$ so that o is linear, M is contextual if $P_S > 1$ which, it being a probability, can never occur. An l 2-MBQC which computes linear functions therefore is always noncontextual. We now consider the Anders-Browne construction as a specific example of Theorem [3.7.2].

	+++	++-	+ - +	+ - -	- + +	- + -	- - +	- - -
$\sigma_x \sigma_x \sigma_x$	1	0	0	1	0	1	1	0
$\sigma_x \sigma_y \sigma_y$	0	1	1	0	1	0	0	1
$\sigma_y \sigma_x \sigma_y$	0	1	1	0	1	0	0	1
$\sigma_y \sigma_y \sigma_x$	0	1	1	0	1	0	0	1

Table 3.3: The empirical model used by Anders and Browne in [10].

Example 3.7.4 (Anders Browne). *Anders and Browne in [11] use as a resource the GHZ state*

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|001\rangle - |110\rangle) \quad (3.35)$$

together with a classical computer restricted to NOT and CNOT operations. They are then able to show that measurement of this state can, supplemented with feed-forward, deterministically compute a OR gate.

The measurement contexts for the setup together with their possible joint outputs are shown in Table 3.3

The l2-MBQC which implements this has as $q_i \in \{0, 1\}$ as the binary input for the three parties so that $i \in \{1, 2, 3\}$. We therefore take q_1 to correspond to the input for agent A, q_2 for B, and q_3 for C.

The choice of measurement that agent i performs when receiving input q_i is $O_i(q_i)$, and here we take $O_i(0) = \sigma_x$ and $O_i(1) = \sigma_y$ for all i . Each party then produces an outcome s_i which we can group into the tuple $\mathbf{s} = (s_1, s_2, s_3)$.

In order to compute the (nonlinear) OR gate, the l2-MBQC must take two input bits $\mathbf{i} = (i_1, i_2)$ and produce a single output bit o such that o is the OR of i_1 and i_2 . The classical control computer may use linear processing of the form in equation 3.33 to compute the choice of measurement basis from the classical inputs and outputs (with causal flow being respected).

We see from the table that the input for C (party 3) is the XOR of the inputs for A and B. In particular

$$\begin{aligned} 0, 0 &\mapsto 0 \\ 0, 1 &\mapsto 1 \\ 1, 0 &\mapsto 1 \\ 1, 1 &\mapsto 0 \end{aligned} \quad (3.36)$$

To this end we define

$$q_1 = i_1, \quad q_2 = i_2, \quad q_3 = i_1 + i_2 \pmod{2}. \quad (3.37)$$

The outcome o is defined as the XOR of the measurement outcomes

$$o = s_1 + s_2 + s_3 \pmod{2} \quad (= s_1 \oplus s_2 \oplus s_3). \quad (3.38)$$

We now check that this computes the OR function. Analysing the table, it holds that

$$\sigma_x \otimes \sigma_x \otimes \sigma_x |\Psi\rangle = (\lambda_1^x \oplus \lambda_2^x \oplus \lambda_3^x) |\Psi\rangle = +1 |\Psi\rangle \quad (3.39)$$

$$\sigma_x \otimes \sigma_y \otimes \sigma_y |\Psi\rangle = (\lambda_1^x \oplus \lambda_2^y \oplus \lambda_3^y) |\Psi\rangle = -1 |\Psi\rangle \quad (3.40)$$

$$\sigma_y \otimes \sigma_x \otimes \sigma_y |\Psi\rangle = (\lambda_1^y \oplus \lambda_2^x \oplus \lambda_3^y) |\Psi\rangle = -1 |\Psi\rangle \quad (3.41)$$

$$\sigma_y \otimes \sigma_y \otimes \sigma_y |\Psi\rangle = (\lambda_1^y \oplus \lambda_2^y \oplus \lambda_3^y) |\Psi\rangle = -1 |\Psi\rangle \quad (3.42)$$

where λ_i^p is the eigenvalue of the Pauli observable σ_p and it can be shown that it is related to the outcome via $\lambda_i^p = (-1)^{s_i}$ i.e. we have a group isomorphism $\langle \{0, 1\}, \oplus \rangle \cong \langle \{+1, -1\}, \cdot \rangle$ relating eigenvalues to outcomes.

The computation is

$$\begin{aligned} i_1, i_2 &\mapsto o \\ 0, 0 &\mapsto 0 \\ 0, 1 &\mapsto 1 \\ 1, 0 &\mapsto 1 \\ 1, 1 &\mapsto 0 \end{aligned} \quad (3.43)$$

which is indeed the OR function.

The empirical model in Table 3.3, which is a state dependent version of Mermin's star with the GHZ state, is known to exhibit contextuality, and it is this contextuality which allows for a nonlinear computation (the OR gate) to be done by a linear classical computer.

There are a number of ways to show that the model is contextual. For example, adding equations 3.39 to 3.42 (mod 2) results in the contradiction $0 = 1$. The contextual behaviour also shows up in the associated bundle diagrams, which was analysed

for this example in [19]. The bundle diagram corresponding to the empirical model given in Table 3.3 is shown below. The model is logically contextual: sections fail to extend to global sections. It is moreover strongly contextual, as the set of global sections which project down to give the correct support on e is empty.

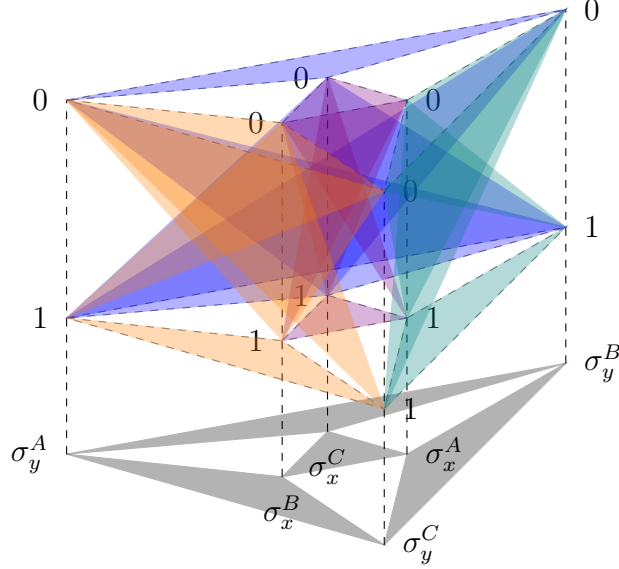


Figure 3.9: Bundle diagram for the Anders Browne setup. Figure reproduced with permission of the authors in [19].

3.8 Vorob'ev's Theorem and measurement covers

The question which we have been trying to answer so far is whether a given empirical model exhibits correlations which can be reproduced by a system obeying the laws of classical physics, which one may want to think of as a type of hidden variable theory. This corresponds, in the sheaf theoretic framework, to whether there exists a global section which restricts down to give the local probability distributions of the empirical model.

A related question is, given a measurement scenario $\mathcal{M}$, whether there exists any empirical models in the set $\mathbf{EM}(\mathcal{M})$ which are contextual. The answer to this will come from Vorob'ev's Theorem, originally stated in the realm of game theory. We give it in the form expressed in [15].

Theorem 3.8.1 (Vorob'ev's Theorem [137, 15]). *Let Σ be a simplicial complex. Then any empirical model defined on Σ is extendable if and only if Σ is acyclic.*

Recall that we say that an empirical model is extendable if there exists a global section extending the local sections of the empirical model in a consistent fashion, so that the empirical model is classical. In Section 2.2 the definition of acyclicity of a simplicial complex was given. It was defined via Graham reduction steps: an acyclic complex is one which cannot be reduced to the zero complex Δ_0 via a series of Graham reduction steps, with a single step being the removal of a vertex from a maximal context, which belongs to no other maximal context.

Vorob'ev's theorem tells us not only that for certain—acyclic—measurement scenarios there can be no contextual empirical models, it also in the case of cyclicity, guarantees that there is at least one contextual empirical model in $\text{EM}(\mathcal{M})$.

Vorob'ev's theorem can also be viewed in terms of simulations: given a scenario $\mathcal{M} = \langle X, \Sigma, O \rangle$ suppose that there exists a Graham reduction step $\Sigma \rightsquigarrow \Sigma'$ where $\Sigma' = \{\sigma \setminus \{x\} \mid \sigma \in \Sigma\}$, so that the vertex x is removed. Then it can be shown that there exists a morphism $\langle \pi, h \rangle : \langle X \setminus \{x\}, \Sigma', O \rangle \rightarrow \langle X, \Sigma, O \rangle$ [81]. As a result we obtain a simulation $e|_{X \setminus \{x\}} \rightarrow e$, and so contextuality of e implies contextuality of $e|_{X \setminus \{x\}}$. Now, if Σ is acyclic then this can be done for Graham reduction steps right down to the zero complex Δ_0 , which cannot host contextuality. Thus an empirical model $e : \langle X, \Sigma, O \rangle$ cannot be contextual if Σ is acyclic.

Chapter 4

Causality in a quantum information setting

— Summary —

This chapter summarises three approaches to modelling causal correlations in a quantum information setting. These are (i) Bayesian networks generalised from Pearl’s classical version, told through the lens of the Henson, Lal, and Pusey criterion [73], which gives a way of classifying which networks exhibit correlations that cannot be classically realised, (ii) causal Bell scenarios as captured by Gogioso and Pinzani using sheaves [66], and (iii) Leggett-Garg correlations [95], and specifically Mansfield’s sheaf approach in [104] to capturing such correlations.

While contextuality, a generalisation of Bell non-locality, has given key insight into how nonclassicality can increase the computational power of a system, it requires (or is equivalent to) performing all measurements simultaneously, so that there can be no causal dependency between measurements. For this reason there has been an increasing trend to study nonclassical correlations in more general causal setups; usually, these are generalisations of Bell or non-locality scenarios rather than contextuality scenarios.

The field of causality naturally has enormous application also outside of quantum information, it being concerned with the broad programme of uncovering causal re-

relationships between random variables. Graphical representations involving directed acyclic graphs (DAGs), which gained prominence due to Pearl [114], have been utilised for decades by mathematicians and computer scientists alike to uncover causal dependencies in various and diverse contexts, ranging from ecology to economics, and everything in between. Pearl’s programme involved using graph theory to understand independence requirements which are imposed on classical random variables due to the causal structure in which they are embedded. This has far reaching consequences because it allows for the disentanglement of correlation from causation, which is imperative in many applied sciences when definitive statements are to be made about causation.

In the context of quantum information, we are interested in what bounds a causal structure places on the correlations which can form if one uses classical, quantum, or a more general information theory. This link was first established by Wood and Spekkens [143]. Bell’s scenario, for example, is captured by asserting that classical correlations are Markov with respect to the Bell DAG. The no-signalling conditions, which must be obeyed by any theory and not just classical probability theory, are also captured by the causal independence constraints imposed by the DAG.

A number of frameworks have been developed with the aim of investigating non-classical correlations in arbitrary causal structures. Many of these have been an effort towards the problem of causal inference, which is to try to determine which causal relations are compatible with the probability distributions over a collection of observed variables. One can also move in the opposite direction, and, fixing a causal relation among variables, determine the set of all possible observed probability distributions. Here we take the second point of view.

In the setup studied by Henson, Lal and Pusey in [73], a DAG has as its nodes random variables and an edge from one node to another indicates direct causal dependency. On the other hand, Gogioso and Pinzani [66] consider a DAG of agents with access to an input and output set, such that all information (in the form of inputs and outputs) in the backward light cone of an agent is signalled. Fritz in [61], in contrast, worked under the assumptions of macrorealism and noninvasivity, motivated by the occurrence of these assumptions in the Leggett-Garg setup, and, on the other hand, in [59] considered nodes to act as update ‘gates’ for hidden variables, which amounts to allowing for, at each measurement, an arbitrary probabilistic update of the hidden variable. As such, the noninvasivity assumption of macrorealism does not hold. Mansfield in [104] showed how macrorealism can be captured in a sheaf-theoretic fashion. Finally, but not least of all, the inflation technique was introduced by Wolfe, Spekkens and Fritz in [142], where, as with

Henson Lal and Pusey, nodes are taken to be random variables—both hidden and visible—and edges represent direct causal influences between variables.

For brevity, we cover the cases studied by Henson Lal and Pusey, Gogioso and Pinzani, and Mansfield (covering that of Leggett-Garg too), in turn. Each of these indicates a different interpretation of causal DAGs, with the others adopting similar frameworks to one of the two, so that a lot can be covered by just these two cases.

A natural starting point is the Bayesian networks of Pearl as realised in the form of graphs, which gives a setting in which to consider random variables (measurement setting inputs and measurement outputs). While this will allow us to study the generalisation of Bell scenarios to arbitrary causal scenarios, it does not entail a full generalisation of contextuality scenarios, as one does not usually speak of agents in contextuality (as doing so implies factorisation of the Hilbert space). For this reason, and as we will later see, we must move beyond such graph theoretic treatments.

4.1 Classical probabilities on causal DAGs

A directed graph G consists of a set of vertices $V(G)$ and an ordered edge set $E(G) \subseteq V(G) \times V(G)$, so that we write $v \rightarrow w$ if $(v, w) \in E(G)$. We will often write V and E for the vertex and edge set, but occasionally it will be useful to differentiate between edge and vertex sets arising from different graphs. A cycle in G is a sequence of edges which ends at the same vertex at which the sequence started. For example, the sequence of edges $a \rightarrow b \rightarrow c \rightarrow a$ is a cycle in G . A graph which contains no cycles is called acyclic.

Given a vertex $v \in V$, we call a vertex c for which there exists an edge $v \rightarrow c$ a child of v , and we denote the set of all such children vertices $\text{ch}_G(v)$, or $\text{ch}(v)$ if the graph of v is clear from the context. Similarly, a parent of v is a vertex $p \in V$ for which there exists an edge $p \rightarrow v$. We denote the set of parents $\text{pa}_G(v)$ (or $\text{pa}(v)$). Finally, a path in G is a sequence of edges $v_1 \rightarrow v_2 \rightarrow v_3 \rightarrow \dots$ in G , and we denote the set of all paths $P(G)$. If there exists a path in G from a vertex v to a vertex w then we write $v \rightsquigarrow w$.

The way in which one interprets the vertices of G depends on the framework one chooses to use.

4.2 Bayesian networks: classicality as Markovianity

Henson Lal and Pusey (HLP) attach to each vertex $i \in V$ a random variable $X^{(i)}$, so that the set of random variables is $\{X^{(i)}\}_{i \in V}$. For convenience we will label vertices from 1 to n , n being the total number of vertices in V . These random variables are (i) inputs to agents, determining the measurement setting (ii) outputs obtained upon measurement (iii) hidden variables. An edge $i \rightarrow j$ in G indicates that the random variable $X^{(i)}$ has a direct causal influence on $X^{(j)}$. The value of a random variable $X^{(i)}$ is denoted $x^{(i)}$, and the set of values of the parents of $X^{(i)}$ in G will be, as with the random variable itself, denoted $\text{pa}(x^{(i)})$, and similarly $\text{ch}(x^{(i)})$ for the set of values of the children nodes.

Pearl introduced the notion of a probability distribution on the random variables $X^{(i)}$ being Markovian with respect to a particular causal structure.

Definition 4.2.1 (Markov Condition [114]). *Let G be a DAG. A probability distribution p is Markov relative to G if p satisfies*

$$p(x^{(1)}, \dots, x^{(n)}) = \prod_i p(x^{(i)} | \text{pa}(x^{(i)})). \quad (4.1)$$

The Markov condition requires that the probability of the random variables $X^{(i)}$ obtaining value $x^{(i)}$ be conditional only on the values obtained by its parent nodes $\text{pa}(x^{(i)})$ in G . HLP then define classicality with respect to this condition.

Definition 4.2.2 (Classical Bayesian Network [114]). *A (classical) Bayesian network is a pair (p, G) where G is a DAG, and p is a probability distribution that is Markov relative to G .*

This contrasts the sheaf approach, where classicality arises purely as an obstruction to the presheaf of distributions on events being a sheaf. The Bell-CHSH setup, for example, is described by the DAG pictured below.

Thus the set of nodes is $\{X, Y, A, B, \Lambda\}$. The Markov condition of equation [4.1] implies

$$p(abxy\lambda) = p(a|x\lambda)p(b|y\lambda)p(x)p(y)p(\lambda). \quad (4.2)$$

This, in turn, implies the following types of conditional independence

1. $p(a|xy\lambda) = p(a|x\lambda)$

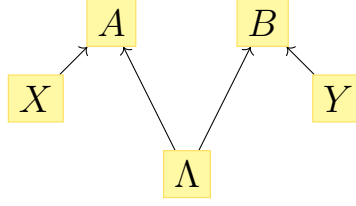


Figure 4.1: The DAG of the Bell-CHSH setup.

2. $p(a|xy) = p(a|x)$ where $p(a|xy) = \sum_b p(ab|xy)$.

Conditional independences such as that in the former follow directly from the Markov condition which requires that random variables be dependent only on parent nodes in G . In the latter equation the variable λ has been marginalised over and the result is the no-signalling condition. The set of probability distributions which obey these conditional independences are often referred to as no-signalling, to reflect that these independences are satisfied, and the set is written **NS**.

While conditional independences of the random variables in a DAG G are derivable from the Markov condition and probability calculus, it is handy to have a systematic way of uncovering them. To this end one introduces d -separation, which essentially describes two variables as independent if they are disconnected in a particular way in G . The notion of d -separation was developed by Geiger, Verma and Pearl [63]. We will not require the definition of d -separation, but mention it to emphasise that such relations are of high importance and there are methods dedicated to systematically uncovering them.

In the DAGs so far considered, such as the Bell DAG in Figure 4.1, every node is considered the same. However, it will be necessary that certain nodes be hidden, meaning that the value of the corresponding random variable is never observed. Such hidden nodes—such as the node Λ in the Bell DAG—correspond to points at which a state is prepared, or source nodes, and the classical description presupposes that such source nodes admit a deterministic LHV description. To this end HLP introduce **generalised DAGs** or **GDAGs** and corresponding **general Bayesian networks** as, respectively, graphs in which nodes are labelled as ‘visible’ or ‘hidden’ and networks which can host any generalised correlations. The idea is that these generalised probabilities should originate from an operational theory and be ‘generalised Markovian’ with respect to G . The three cases considered are

1. General probabilistic theory, or the set of probabilities **GPT**;
2. Quantum theory, or the set of probabilities **Q**;

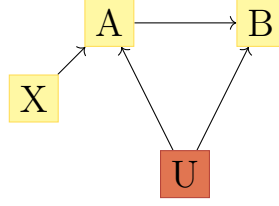


Figure 4.2: The instrumental setup, in which the outcome of the measurement by party A is used as the input by party B . Such behaviour is, in our terminology, adaptive.

3. Classical probability theory, or the set of probabilities $\mathbf{C}$.

We will not give the full presentation of generalised Bayesian networks, but introduce an example which demonstrates how causal correlations are captured in these DAGs.

Example 4.2.3 (Instrumental Setup). *The instrumental scenario DAG is shown in Figure 4.2*

The setup consists of three visible nodes, $\{X, A, B\}$, and the hidden node $\{U\}$. Concretely, this can be thought of as the variable random X being supplied to A as a choice of measurement setting, who in turn sends the result of measurement to B , who again uses this as a choice of measurement setting. U can be seen as a source of shared randomness shared between A and B .

It is shown in [135] that in the case $|X| = |A| = 2$ and $|B|$ arbitrary, there can be no separation between $\mathbf{C}$, $\mathbf{Q}$ and $\mathbf{GPT}$. In other words, we have $\mathbf{C} = \mathbf{Q} = \mathbf{GPT}$. On the other hand, when $|X| = 3$ and $|A| = |B| = 2$, they show that $\mathbf{C} \subsetneq \mathbf{Q} \subsetneq \mathbf{GPT}$, so that there exist quantum correlations which are not classically realisable and $\mathbf{GPT}$ correlations which are not quantum realisable.

HLP develop an algorithm which **classifies the interesting DAG's**. What they mean by this is that the algorithm is able to discern whether, for a given DAG, it holds that $\mathbf{C} \subsetneq \mathbf{NS}$, since if these are equal then all probability distributions on G which satisfy the required no-signalling constraints are in fact classically realisable. We note, in closing, that an important feature of their framework is that it is able to capture examples in which there are multiple source nodes, including the triangle network of [121].

4.3 Causal Bell scenarios

Gogioso and Pinzani in [66] generalised the non-local scenarios of the sheaf theoretic formulation to scenarios in which there is some causal structure present. They do

this by introducing a poset $(\Omega, \leq)$ —a set together with a partial order on that set—of agents, and for each agent $\omega \in \Omega$, an input set I_ω and an output set O_ω . Altogether, then, a measurement scenario is captured by a triple $\langle \Omega, \{I_\omega\}_{\omega \in \Omega}, \{O_\omega\}_{\omega \in \Omega} \rangle$. We will refer to such setups as **causal Bell scenarios** moving forward, or, occasionally, as Gogioso-Pinzani or GP scenarios.

Begin by letting $\Lambda(\Omega)$ denote the set of all down-closed subsets of Ω with respect to $\leq$. A down-closed subset, or equivalently lower set, is a subset $S \subseteq \Omega$ where

$$\forall \omega \in \Omega, \forall \omega' \in S. \omega \leq \omega' \implies \omega \in S. \quad (4.3)$$

Choose for each $\omega \in S$ a choice of subset of inputs $U_\omega \subseteq I_\omega$. A section is then defined to be a function which is causal with respect to $\leq$. This is a function

$$f : \prod_{\omega \in S} U_\omega \rightarrow \prod_{\omega \in S} O_\omega \quad (4.4)$$

which factorises as a collection of functions $\{f_\omega\}_{\omega \in \Omega}$ where

$$f_\omega : \prod_{\nu \in \downarrow \omega} U_\nu \rightarrow O_\omega. \quad (4.5)$$

Here, $\downarrow \omega$ is the subset of Ω which contains all $\nu \in \Omega$ for which $\nu \leq \omega$. We say it is the **lower set generated by ω** .

Given such a section f we can restrict it to a section $f|_L$ on a lower set $L \subseteq S$,

$$f|_L : \prod_{\omega \in L} U_\omega \rightarrow \prod_{\omega \in L} O_\omega \quad (4.6)$$

as follows. Introduce the restriction maps

$$\rho_{S,L}^U : U_S \rightarrow U_L$$

$$\rho_{S,L}^O : O_S \rightarrow O_L$$

where for brevity U_S is now used to denote $\prod_{\omega \in S} U_\omega$. Thus for any $\mathbf{i} \in U_S$, $\rho_{S,L}^U(\mathbf{i}) = \mathbf{i}|_L$ where $\mathbf{i}|_L$ is a tuple with entries only at elements ω which lie in L . The restriction

map $\rho_{S,L}^O$ is defined analogously. We then define

$$f|_L(\mathbf{i}|_L) = f(\mathbf{i}|_L). \quad (4.7)$$

The requirement that $f : \prod_{\omega \in \Omega} I_\omega \rightarrow \prod_{\omega \in \Omega} O_\omega$ be causal is equivalent to requiring that, for any $S \subseteq \Omega$ (it need not necessarily be down-closed) the following diagram commute.

$$\begin{array}{ccc} I_\Omega & \xrightarrow{f} & O_\Omega \\ \rho_{\Omega, \downarrow S}^U \downarrow & & \downarrow \rho_{\Omega, S}^O \\ U_{\downarrow S} & \dashrightarrow & O_S. \end{array}$$

Let $\mathcal{E}_C((S, \{U_\omega\}_{\omega \in S}))$ be the set of causal Bell functions on $(S, \{U_\omega\}_{\omega \in S})$. Moreover, define the set

$$\mathcal{L} := \{(S, \{U_\omega\}_{\omega \in S}) \mid S \in \Lambda(\Omega) \wedge \forall \omega \in S. U_\omega \subseteq I_\omega\}. \quad (4.8)$$

Precomposing the event sheaf $\mathcal{E}_C$ with the distribution monad gives probability distributions over deterministic causal functions, and specifically a global section is an element

$$p \in \mathcal{D}_R \left(\prod_{\omega \in \Omega} (I_{\downarrow \omega} \rightarrow O_\omega) \right). \quad (4.9)$$

We will write $p[f_{\omega_1}, f_{\omega_2}, \dots]$ to denote the evaluation of the probability distribution at a choice of causal function factorising according to $f_{\omega_1}, f_{\omega_2}, \dots$, or simply $p[f]$.

Given $L \subseteq S$ we may marginalise a distribution p on causal functions over S to one over L via

$$(p|_L)[f'] := \sum_{\substack{f \in \mathcal{E}_C(S) \\ f|_L = f'}} p[f]. \quad (4.10)$$

An empirical model is a compatible family for the presheaf $\mathcal{D}_R \circ \mathcal{E}_C$ over the cover $\mathcal{C}$ which contains elements of the form $(L, \{i(\omega)\}_{\omega \in L})$ for $L \in \Lambda(\Omega)$. An element $(L, \{i(\omega)\}_{\omega \in L})$ of the cover consists of a single input for each agent in L , so we will from now on label it instead as $\mathbf{i}$.

Definition 4.3.1 (Causal Bell empirical model). *Let $\mathcal{M} = \langle \Omega, \underline{I}, \underline{O} \rangle$ be a causal measurement scenario. An empirical model e is a compatible family of distributions $e_{\mathbf{i}} \in \mathcal{D}_R \circ \mathcal{E}_C(\mathbf{i})$ for $\mathbf{i} \in \bigcup_{S \in \Lambda(\Omega)} (\prod_{\omega \in S} I_\omega)$.*

Contextuality then arises as an obstruction to a global section for the compatible family defined by e .

There is an alternative way to define empirical models, and often this definition will be preferred as it is more intuitive. Operationally, an empirical model is for each input that the set of agents can choose a distribution on observed outcomes which obeys the desired no-signalling constraints imposed by the causal structure. Mathematically, this is just a conditional distribution

$$e : I_\Omega \rightarrow \mathcal{D}_R(O_\Omega) \quad (4.11)$$

so that $e(\mathbf{i}) \in \mathcal{D}_R(O_\Omega)$ is a distribution on outputs given the choice $\mathbf{i}$ of inputs. We write $e(\mathbf{i})[\mathbf{o}]$ to denote the evaluation at an outcome $\mathbf{o}$, which is more commonly written $p(\mathbf{o}|\mathbf{i})$. The no-signalling conditions are expressed by the following commuting diagram

$$\begin{array}{ccc} I_\Omega & \xrightarrow{e} & \mathcal{D}_R(O_\Omega) \\ \rho_{\Omega, \downarrow S}^U \downarrow & & \downarrow \rho_{\Omega, S}^O \\ U_{\downarrow S} & \dashrightarrow & \mathcal{D}_R(O_S). \end{array}$$

Since a global section is a distribution $d \in \mathcal{D}_R(\prod_{\omega \in \Omega} (I_{\downarrow \omega} \rightarrow O_\omega))$, an empirical model, if arising from a global section d , will have probabilities

$$e_d(\mathbf{i})[\mathbf{o}] = \sum_{\substack{f_{\omega_1}, f_{\omega_2}, \dots, f_{\omega_n} \\ \mathbf{o}(\omega) = f_\omega(\mathbf{i}|_{\downarrow \omega})}} d[f_{\omega_1}, f_{\omega_2}, \dots, f_{\omega_n}]. \quad (4.12)$$

Example 4.3.2 (Temporal Bell-CHSH). *Consider the setup with two parties $\Omega = \{A, B\}$ with $A \leq B$. A causal global section is a function $f : I_A \times I_B \rightarrow O_A \times O_B$ which factorises as $f(i_A, i_B) = (f_A(i_A), f_B(i_A, i_B))$, since $A \leq B$ in Ω the function f_B depends on both inputs. A probabilistic global section is a distribution $d \in \mathcal{D}_R((I_A \rightarrow O_A) \times (I_A \times I_B \rightarrow O_B))$ such that for any $f : I_A \rightarrow O_A$ and $g : I_A \times I_B \rightarrow O_B$, $d[f, g]$ is the probability of the system responding with deterministic functions f and g .*

On the other hand, empirical data is obtained from observation as follows (this can be done whether a global section exists or not). Actually observing the system

gives two distributions:

$$p : I_A \rightarrow \mathcal{D}(O_A) \quad (4.13)$$

$$q : I_A \times I_B \times O_A \rightarrow \mathcal{D}(O_B). \quad (4.14)$$

Then for $R \in O_A$ and $S \in O_B$ the empirical model is defined $e(i, j)[R, S] = p(i)[R]q(i, j, R)[S]$. Moreover if it arises from a global section then

$$e(i, j)[R, S] = \sum_{f, g, R=f(i), S=g(i, j)} d[f, g]. \quad (4.15)$$

The equation above is more familiarly expressed

$$p(a_i = R, b_j = S) = \sum_{\lambda \in \Lambda} p(\lambda) p(a_i = R | \lambda) p(b_j = S | i, \lambda), \quad (4.16)$$

where the set Λ of hidden variables is in one-to-one correspondence with deterministic global sections. Such factorisation can be shown explicitly from equation [4.15](#), and is a straightforward generalisation of the proof in [\[6\]](#).

Compare the expression in equation [4.16](#) to that of the classical Markovian condition in equation [4.1](#). Suppose we try to find a DAG with Markov conditional identical to the factorisation of equation [4.16](#). There are four possible GDAGs shown below.

for which the Markovianity condition reduces to, respectively

$$p(a, b | x, y) = \sum_{\lambda} p(\lambda) p(a | \lambda, x) p(b | a, y) \quad (4.17)$$

or

$$p(a, b | x, y) = \sum_{\lambda} p(\lambda) p(a | \lambda, x) p(b | \lambda, a, y). \quad (4.18)$$

or

$$p(a, b | x, y) = \sum_{\lambda} p(\lambda) p(a | \lambda, x) p(b | \lambda, x, y). \quad (4.19)$$

or

$$p(a, b | x, y) = \sum_{\lambda} p(\lambda) p(a | \lambda x) p(b | \lambda, a, y, x). \quad (4.20)$$

In the equations above, a corresponds to the measured value of the random variable

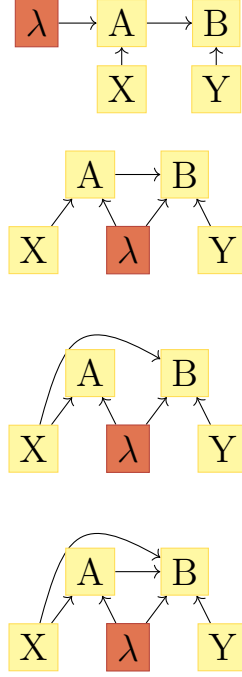


Figure 4.3: Four GDAGs which might be associated with the temporal Bell-CHSH setup.

A , and $p(a|x)$ translates into the probability of a being obtained given that the measurement setting was x . Thus, expressing these equations instead in the language of equation [4.16](#) gives

$$p(a_i = R, b_j = S) = \sum_{\lambda \in \Lambda} p(\lambda) p(a_i = R|\lambda) p(b_j = S|R) \quad (4.21)$$

and

$$p(a_i = R, b_j = S) = \sum_{\lambda \in \Lambda} p(\lambda) p(a_i = R|\lambda) p(b_j = S|R, \lambda). \quad (4.22)$$

and

$$p(a_i = R, b_j = S) = \sum_{\lambda \in \Lambda} p(\lambda) p(a_i = R|\lambda) p(b_j = S|i, \lambda). \quad (4.23)$$

and

$$p(a_i = R, b_j = S) = \sum_{\lambda \in \Lambda} p(\lambda) p(a_i = R|\lambda) p(b_j = S|i, R, \lambda). \quad (4.24)$$

In the third equation we obtain the same factorisation condition as in equation [4.1](#). In fact, $p(b_j = S|i, R, \lambda) = p(b_j = S|i, \lambda)$ since given the NCHV λ and the input at A i one can reconstruct S . Thus equations [4.23](#) and [4.24](#) are the same for correlations which admit a NCHV description.

A GP scenario therefore emerges on the assumption that *all* information from the past is signalled. The interpretation, however, is not quite right: in an HLP setup, which hinges on definitions from Pearl’s use of Bayesian networks to study correlation, the arrow $A \rightarrow B$ signifies that the output at A is forwarded to party B ; instead, for a GP scenario, it signifies that a hidden variable which may contain all the information about the past is passed from party A to party B . Therefore a better and perhaps more elegant diagrammatic representation GP scenarios is through DAG’s in the following form [†]

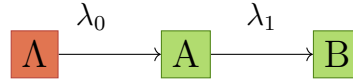


Figure 4.4: Diagrammatic representation of the temporal Bell-CHSH setup.

Here, a hidden variable λ_0 which we identify with a map $f : I_A \rightarrow O_A$ is measured by A , thereby revealing the output $f(i)$ where i is the input at a . Then, a hidden variable λ_1 is passed from A to B which is really a section $f : I_A \times I_B \rightarrow O_A \times O_B$, so that the output at B that is obtained is $f(i, j)$, where j is the choice of input at B . In every round of experiment, we only observe that the hidden variable (i.e. global section), if the system is behaving in accord with one, maps $i \mapsto f(i)$ and $j \mapsto f(i, j)$. Here we will simply draw the DAG on the agents to represent such a setup:



Figure 4.5: Alternative diagrammatic representation of the temporal Bell-CHSH setup, in which the source node is omitted.

4.4 The Leggett-Garg setup and macrorealism

The third and final example we consider are correlations which are **macrorealistic** [95]. Famously, these correlations were studied by Leggett and Garg in a simple setup consisting of a single measurement a which is performed at three distinct times t_1 , t_2 and t_3 [95]. Assuming the system being measured evolves in between these measurements, which would be unitary Hamiltonian evolution for a quantum system, we can work within the Heisenberg picture and consider time-evolved measurements

[†]In this representation, d -separation cannot be used to derive conditional independences.

a_1 , a_2 and a_3 instead, where $a_i = a(t_i)$. In each round of experiment, one of the three sequences a_1a_2 , a_1a_3 or a_2a_3 may be performed.

Leggett and Garg showed that if we take classicality to include the assumptions of macroscopic realism and non-invasiveness, then a quantum system in this setup is able to produce correlations which cannot be produced by any classical system [95]. We state these assumptions below for clarity:

1. *Macroscopic Realism.* The system is always in one of several distinct macroscopic states at any given time.
2. *Non-invasiveness.* It is possible to measure the system to determine its state without changing the state of the system or affecting its subsequent dynamics.

Based on these assumptions, a macro-realistic hidden variable theory can be derived. In particular, it can be shown that for a macro-realistic system undergoing a sequence of measurements by N parties each with a given measurement setting, the probability of obtaining the string of outcomes $(o_0, \dots, o_n)$ given a string of input settings $(i_0, \dots, i_n)$ factorises as follows:

$$p(o_0, \dots, o_n | i_0, \dots, i_n) = \sum_{\lambda \in \Lambda} p(\lambda) p(o_0 | i_0) p(o_1 | i_1) \dots p(o_n | i_n). \quad (4.25)$$

Here, λ is a macrorealistic HV, and Λ a set of macrorealistic HVs.

The validity of these assumptions placed on a classical system have been extensively debated [43]. Showing that macrorealism holds in a loophole free way is difficult because of the temporal nature of the correlations. While in a Bell setup one can ensure that there is no unaccounted for signalling in the classical system by separating the parties sufficiently and appealing to special relativity, here it is not possible to guarantee that a measurement performed at a later time is not impacted by the value obtained by a measurement at an earlier time.

Mansfield in [104] shows how such correlations can be captured in a sheaf theoretic way. Given a poset $(X, \leq)$ of measurements, with $m_1 < m_2$ indicating that m_1 occurred before m_2 define an order $\preceq$ on $\mathcal{P}(X)$ as for any $A, B \in \mathcal{P}(X)$,

$$A \preceq B \implies \downarrow(A \cap B) = A. \quad (4.26)$$

A measurement scenario is then defined $\mathcal{M} = \langle (X, \leq), \mathcal{C}, O \rangle$ where

i $(X, \leq)$ is a poset of measurements;

- ii $\mathcal{C}$ is a choice of subset of $\mathcal{P}(X)$ which is down closed, $\downarrow_{\preceq} \mathcal{C} = \mathcal{C}$;
- iii O is a set of outputs for each measurement.

An element $\sigma \in \mathcal{C}$ is called a **context**, in analogy with contextual setups. Note that while for a contextuality setup we require down closure of the cover with respect to the subset relation $\subseteq$, here it is instead with respect to the order $\preceq$. In particular, $(\mathcal{P}(X), \preceq)$ is now the poset that is used.

A section on an element $U \in \mathcal{P}(X)$ is a function

$$s : U \rightarrow O :: m \mapsto o. \quad (4.27)$$

This choice of section encodes the assumptions of macrorealism. In particular, a **global section** is an assignment $f : X \rightarrow O$ sending each measurement $m \in X$ to a predefined outcome $f(m)$ regardless of any previous measurements made. The set of sections over U is denoted $\mathcal{E}(U)$, as this sheaf coincides with the one used in standard contextuality scenarios.

The difference arises when defining empirical models, since no-signalling constraints should now be relaxed to account for the imposed causal structure.

Definition 4.4.1 (Causal Empirical Model [105]). *A causal empirical model is a choice $e_C \in \mathcal{D}_R \circ \mathcal{E}(C)$ for each $C \in \mathcal{C}$ such that for any $C, C' \in \mathcal{C}$ with $C \leq C'$*

$$e_{C'}|_C = e_C. \quad (4.28)$$

Thus an empirical model is simply a compatible family for the presheaf $\mathcal{D}_R \circ \mathcal{E}$ on the cover $\mathcal{C}$.

We consider the Leggett-Garg setup as an example.

Example 4.4.2 (Leggett-Garg setup [95]). *Define $X = \{a_1, a_2, a_3\}$ with order $a_i < a_j$ if $i < j$. The set of contexts is defined*

$$\mathcal{C} = \{\{a_1, a_2\}, \{a_2, a_3\}, \{a_1, a_3\}, \{a_1\}, \{a_2\}, \{a_3\}\}, \quad (4.29)$$

which is down-closed with respect to $\preceq$. The outcome set is taken to be $\{0, 1\}$. The empirical model displayed in Table 4.6 (not quantum realisable) obeys the required no-signalling relations:

$$e_{\{a_i, a_j\}}|_{\{a_j\}} = e_{\{a_j\}} \quad (4.30)$$

Nevertheless, it is not classical—as can be seen from the bundle diagram, there are no candidates for a global section and the diagram contains no closed paths.

One can also define a quantum realisable model which is not classical by this definition. Define the single qubit observables a_i for $i \in \{1, 2, 3\}$ as

$$a_i = \mathbf{a}_i \cdot \sigma \quad (4.31)$$

where $\sigma = \begin{pmatrix} \sigma_x \\ \sigma_y \\ \sigma_z \end{pmatrix}$ is a vector of Pauli matrices and $\mathbf{a}_i$ indicates the vector on the Bloch sphere along which the qubit is to be measured. Such observables are PVMs with projectors

$$P_1^{a_i} = \frac{\mathbb{1} + a_i}{2} \quad P_{-1}^{a_i} = \frac{\mathbb{1} - a_i}{2}. \quad (4.32)$$

It is known that choosing the vectors $\mathbf{a}_i$ such that the angle between $\mathbf{a}_1$ and $\mathbf{a}_2$, and $\mathbf{a}_2$ and $\mathbf{a}_3$ is $\pi/3$, maximises violation of the Leggett-Garg inequality [33] (for PVMs).

		(0, 0)	(0, 1)	(1, 0)	(1, 1)
a_1	a_2	1/2	0	0	1/2
a_2	a_3	1/2	0	0	1/2
a_1	a_3	0	1/2	1/2	0

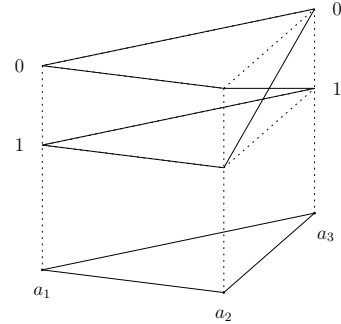


Figure 4.6

Figure 4.7: An example empirical model and associated bundle diagram for the Leggett-Garg setup.

Note that if an empirical model is classical, it satisfies additional no-signalling requirements. In particular, it is actually no-signalling with respect to the cover $\downarrow_{\subseteq} \mathcal{C}$ which is the cover for a contextuality setup. This will be considered in more detail in the following chapter.

Part II

Research contributions

Chapter 5

Flattening temporal correlations to contextuality correlations

— Summary —

This is the first chapter containing new research contributions. We study correlations which arise when measurements are sequentially performed on a system, and with the caveat that the systems memory is constrained in some way. We show that, for certain choices of memory, such correlations can be simulated by an appropriately defined contextuality scenario. This result is ‘the Mapping Theorem’ (Theorem 5.7.2). This construction then allows us to lift Vorob'ev's Theorem [137] to such temporal scenarios, allowing us to say, for which choices of setup and memory, there can be correlations which are not classically realisable (with the given constraint on memory), giving us the second main Theorem of the chapter (Theorem 5.6.1).

We have seen that contextuality is an important feature of quantum systems which distinguishes them from their classical counterparts; in quantum systems measurement contexts are important and distributions on measurement outcomes become context specific, albeit in a way that the no-signalling principle is still satisfied. Such an observation has created a fundamental shift in the way quantum systems are to be viewed, but has also led to important discoveries concerning the role of contextuality in information processing tasks—not least of all, and as we

have discussed in earlier Chapters, contextuality has been shown to be necessary for increasing the computational power in certain models of computation [11, 119].

Nevertheless, the correlations which arise in contextuality scenarios are by assumption no-disturbance, and therefore equivalent to performing all measurement simultaneously in a single round. Temporal correlations describe correlations which are time-like separated, such as those which arise with the repeated measurement of a single system at different time steps. The consideration of temporal correlations is not new, and, in analogy with the spatial case in which contextuality inequalities are ubiquitous [89, 41], inequalities can be derived which separate classical and quantum correlations in such setups, with violations of such inequalities corresponding to a quantum advantage in certain information processing tasks [27, 146]. However, although there are many similarities with results on spatial correlations, the assumptions of classicality are different, and of a more subtle kind. Most notably, Leggett and Garg supposed that macrorealism was a valid axiom of classicality in the temporal regime, however loophole free violation of inequalities derived under these assumptions have so far not been possible [53]. The subtlety arises due to the possibility of signalling in the temporal case, making it important to carefully consider different models of classicality and the classical resources they require. In short, in the temporal case quantum advantage in certain information processing tasks is always relative to some bounds on the classical system, such as memory bounds [31, 30, 136, 28].

This use of memory resource theoretic arguments in the context of quantum contextuality, or more generally nonclassicality, has been done in two different ways. On the one hand, Kleinmann et al. [88] and Fagundes and Kleinmann [54] consider how many bits of (a classical) memory are required in order to simulate contextual correlations. Since in a contextuality setup commutativity of measurements allows one to perform measurements sequentially without any effect on the resulting empirical data, one is allowing the classical machine to utilise this temporal order in order to simulate the resulting correlations. In these works it is shown that simulating the Peres-Mermin scenario requires a classical system—a finite state machine in particular—with at least three internal states. On the other hand, Hoffman et al. [74] and Budroni et al. [31] consider memory as a resource for simulating the temporal correlations arising when measuring a single system sequentially, whether such measurements are compatible or not. In [74] the polytope of temporal quantum correlations is characterised, while in [31] inequalities are established for the simplest scenario which bound classical, quantum, and GPT correlations.

Here, we cast into the sheaf theoretic framework a setup in which sequential measurements are performed on a single system. The production of a sequence of outcomes from a sequence of measurements will be called a **strategy**, and we subject such strategies to certain memory requirements: only a subset of all past measurements and outcomes can be retained. As is the case with contextuality, only certain empirical models can be considered classical, in the sense that the probability distributions which constitute the empirical model are consistent with a global section of the relevant presheaf. Alternatively, one can ask whether there exists a classical machine $\mathbf{E}$, which produces global strategies with some probability, which could have produced the empirical model.

In this case when this memory stores only measurements (outcomes are prohibited), we can map the empirical models containing temporal correlations to contextuality scenarios. This allows us to lift results from contextuality scenarios to the temporal scenario, and, in particular, we show that we can construct a simplicial complex which, if acyclicity, implies that there exists no nonclassical empirical models on the associated temporal scenario. This result is a direct consequence of use Vorob'ev's theorem on the image scenario.

While this result is restricted to cases in which $\mathbf{E}$ does not store outcomes, we can nevertheless show that

- i An empirical model generated by a classical machine that can store past measurements and their outcomes can be realised by a classical machine that stores only measurements if one increases the allowed number of stored measurements. We analyse this trade-off for certain examples of empirical models.
- ii We can for any measurement scenario which admits correlations which cannot be produced by a classical machine storing only the k preceding measurements, find at least one set of correlations which also cannot be produced by classical machine storing the k preceding measurements and outcomes.

Our perspective is somewhat different to that of [31] and [74] in that we do not consider directly the memories in the quantum and GPT cases, which they define to be the maximal number of states perfectly distinguishable with a single measurement. Instead, we define a classical memory via a classical memory function, which keeps track of some subset of the previous measurement events, and require that the quantum and more general 'allowed' correlations obey the same no-signalling constraints that the resulting classical correlations do. Macrorealistic correlations, for example, correspond to the classical system having no memory (which has been observed already in [31]), and the resulting no-signalling correlations are much more

prohibitive than the general case of unbounded memory—in fact, and as discussed in Section 5.5, macrorealism implies that there is no signalling in the setup whatsoever.

5.1 The setup

We study scenarios in which a sequence of measurements is performed on a single system, thereby producing a sequence of corresponding measurement outcomes. Measurements that can be performed on the system are drawn from a finite set X , and the output of each measurement is assumed to lie in a finite set O . The choice of sequence may vary but may be constrained in some way.

We specify the allowed sequences by collecting them in a prefix-closed subset $\Sigma \subseteq X^*$, so that each element $\sigma \in \Sigma$ is a sequence of measurements from X . Here, X^* denotes the set of all finite sequences which can be constructed from X , and we order elements in X^* via the prefix relation; that is, if τ is a prefix sequence of σ , then $\tau \leq \sigma$ in X^* . We indicate the concatenation of two sequences $\sigma_1, \sigma_2 \in X^*$ by $\sigma_1\sigma_2$ and let $\sigma[i]$ and $\sigma[-i]$ be the i th measurement in σ indexing from the beginning and end of the sequence respectively. Finally, each measurement is taken to have output lying in the set O , which for now we take to be the same for every measurement.

Definition 5.1.1 (Temporal measurement scenario). *A temporal measurement scenario is a tuple $\mathcal{M} = \langle X, \Sigma, O \rangle$, where*

- i* X is a finite set of measurements;
- ii* $\Sigma \subseteq X^*$ is a choice of measurement sequences to be performed;
- iii* O is a set of outcomes.

Note that Σ is not a simplicial complex—nevertheless, it still keeps track of which measurements are observed in a single run of the experiment, albeit such data is now inherently ordered. It will be convenient to require that Σ is down-closed with respect to $\leq$, so that $\downarrow_{\leq} \Sigma = \Sigma$. A **measurement event** is a tuple $\langle x, o \rangle$, where $x \in X$ is a choice of measurement and $o \in O$ the outcome obtained upon measurement. We will also write $\langle x, \perp \rangle$ to signify a measurement event with null entry for the measurement outcome. Therefore, in every round of measurement on the system $\mathcal{S}$, a sequence of measurement events is obtained. Concretely, to a temporal measurement scenario $\mathcal{M}$ we associate a **set of histories** $\mathcal{H}(\mathcal{M})$, defined as follows

$$\mathcal{H}(\mathcal{M}) := \{ \langle m_0, o_0 \rangle \langle m_1, o_1 \rangle \dots \langle m_n, o_n \rangle \mid m_0 m_1 \dots m_n \in \Sigma \wedge \forall i \in [0, n]. o_i \in O \} \quad (5.1)$$

An element $h \in \mathcal{H}(\mathcal{M})$ consists of a sequence of measurement events which could be observed in a round of measurement.

A subset of measurement scenarios are those which arise from agents sequentially acting on the system, with each agent only being allowed to perform a measurement from a specified set. This is described by a linearly ordered set of agents $\mathcal{A}$ together with, for each agent $A \in \mathcal{A}$, an input set I_A and an output set O_A . We call such scenarios **temporal Bell scenarios**, to match the terminology in the spatial case. Naturally, a temporal Bell scenario is a causal Bell scenario $\mathcal{M}_\Omega = \langle \Omega, \underline{I}, \underline{O} \rangle$ where Ω is a linear order.

Example 5.1.2 (Temporal Bell Scenarios). *Consider a linearly ordered set of agents $\mathcal{A}$. In each round of experiment, agent $A \in \mathcal{A}$ performs the measurement indicated by their input $i \in I_A$. We therefore define the set of measurements*

$$X := \{(a, i) | A \in \mathcal{A} \wedge i \in I_A\}. \quad (5.2)$$

Numbering the agents from 1 to N according to the linear order, the set of maximal sequences which can be performed is

$$\max(\Sigma) := \{(1, i_1)(2, i_2)\dots, (N, i_N) | \forall j \in [1, N]. i_j \in I_j\}. \quad (5.3)$$

The scenarios we have so far considered are theory independent. Usually, however, we are interested in quantum scenarios as a special case; and as in the spatial case these measurements are taken to be projectors on a Hilbert space.

Example 5.1.3 (Quantum Temporal Measurement Scenario). *Let $\mathbf{M}$ be a list of Projection-Valued Measures, or PVMs for short ^[1]. Then a triple $\langle \mathbf{M}, \Sigma, O \rangle$ is a quantum temporal measurement scenario, with Σ denoting which sequences of PVMs can be performed and O the possible outcomes of a PVM in $\mathbf{M}$.*

We can for example consider the temporal CHSH setup (see ^[60] and Section ^[4.3]). The set of PVMs consists of four measurements,

$$\mathbf{M} = \{M_{(a,0)}, M_{(a,1)}, M_{(b,0)}, M_{(b,1)}\}. \quad (5.4)$$

Sequences which can be performed are given by

$$\Sigma = \{M_{(a,0)}, M_{(a,0)}M_{(b,0)}, M_{(a,0)}M_{(b,1)}, M_{(a,1)}, M_{(a,1)}M_{(b,0)}, M_{(a,1)}M_{(b,1)}\}, \quad (5.5)$$

¹Note that the choice of PVMs is not the most general choice for a quantum system. More generally, one can consider generalised quantum measurements, or Lüder's measurements, in which case any correlation of the current example can be achieved with a quantum system—see ^[61]

which note is a prefix-closed subset of $\mathbf{M}^*$.

We now consider what behaviours of the system are allowed classically. In a single run of the experiment, a sequence of measurements $\sigma = m_0 m_1 \dots m_i \dots m_n$ is mapped to a sequence of outcomes. For a deterministic classical system, earlier measurements (and their outcomes) can be signalled to later measurements—in practice the system would store information from the past in a memory of some kind. To this end define a function

$$f : \Sigma \rightarrow O :: \sigma \mapsto o \quad (5.6)$$

such that if one inputs the sequence $\sigma = m_0 \dots m_n$ to the system $\mathcal{S}$ then the system returns the sequence of outputs $f(m_0)f(m_0 m_1) \dots f(m_0 \dots m_n)$. Note that here the function is defined so that $f(\sigma)$ returns an outcome only for the final measurement in σ . Therefore, measurements are mapped to outcomes conditional on the previous measurements which have been made. Note also that we have not considered outputs—why do we not take the output at m_n to be a function of this measurement as well as the n previous measurement events, or symbolically $f(\langle m_0, o_0 \rangle \langle m_1, o_1 \rangle \dots \langle m_n, \perp \rangle)$?

We consider the temporal CHSH setup to see why in this case outcomes need not be considered.

Example 5.1.4 (Temporal CHSH Setup). *The setup is described by a measurement scenario $\mathcal{M} = \langle X, \Sigma, O \rangle$ where*

1. $X = \{(a, 0), (a, 1), (b, 0), (b, 1)\}$
2. $\Sigma = \{(a, 0), (a, 1), (a, 0)(b, 0), (a, 0)(b, 1), (a, 1)(b, 0), (a, 1)(b, 1)\}$
3. $O = \{0, 1\}$.

A section is a function $f : U \rightarrow O$ where $U \subseteq \Sigma$. For example for $U = \{(a, 0), (a, 0)(b, 0)\}$ explicitly it is a mapping

$$f :: \{(a, 0) \mapsto i, (a, 0)(b, 0) \mapsto j\}, \quad (5.7)$$

with $i, j \in \{0, 1\}$. Including outputs, a section over U is a mapping

$$f :: \{\langle (a, 0), \perp \rangle \mapsto i, \langle (a, 0), o_{(a, 0)} \rangle \langle (b, 0), \perp \rangle \mapsto j'\}, \quad (5.8)$$

where j' can take on different values depending on the value of $o_{(a,0)}$. However since the only possible value for $o_{(a,0)}$ to take on is that that was actually obtained, i , this becomes

$$f :: \{ \langle (a, 0), \perp \rangle \mapsto i, \langle (a, 0), i \rangle \langle (b, 0), \perp \rangle \mapsto j \}, \quad (5.9)$$

which is identical to that without outputs.

Another way to see this is that since the system is responding deterministically, knowing the full history of measurements is enough to reconstruct all past measurement outcomes.

Given f , we define a function f^* on $\sigma = m_0 \dots m_n$ in Σ to be

$$f^*(m_0 \dots m_n) := f(m_0) f(m_0 m_1) \dots f(m_0 \dots m_n), \quad (5.10)$$

so that f^* instead returns for a choice of sequence of measurements a corresponding sequence of outcomes. In certain cases it will be convenient to use this representation.

Let $\mathcal{S}(\Sigma)$ denote the set of all possible functions as in equation [5.6](#); we will refer to this as the set of **global strategies** over U . Then for any $U \in \mathcal{P}(\Sigma)$ we can let $\mathcal{S}(U)$ be the set of strategies on U i.e. any $f \in \mathcal{S}(U)$ is a function

$$f : U \rightarrow O :: (\sigma \in U) \mapsto o. \quad (5.11)$$

Naturally, strategies over a set $U \in \mathcal{P}(\Sigma)$ can be restricted to a subset $U' \subset U$ with $U' \in \mathcal{P}(\Sigma)$, because if one knows how sequences are mapped in U then one knows how they are mapped in only a subset of those sequences. We will write this as $\mathcal{S}(U)|_{U'}$, where each $f \in \mathcal{S}(U)$ is restricted as $f|_{U'} :: (\sigma \in U') \mapsto (o \in O)$ where $f|_{U'}(\sigma) := f(\sigma)$ —note that $\sigma \in U'$ implies that $\sigma \in U$, so this is a well-defined restriction map. It can be checked that $\mathcal{S}$ together with this choice of restriction map satisfies the axioms of a presheaf. It can moreover be checked that $\mathcal{S}$ is a valid sheaf.

So far we have allowed for the system to have a lot of flexibility in the way it can respond to a measurement. By allowing a measurement to have different outcomes depending on **any** previous measurement made, we are allowing the system to store an arbitrary number of previous measurements— in other words, we are allowing the system to have an unbounded memory of a particular kind. However, as would usually be the case in real systems, we can suppose that this is restricted—for example, the system might only be capable of storing the previous k measurements

in memory at any given time. Limiting memory to only the previous k measurements (and not outcomes) is rather restrictive, but it turns out to be important for understanding more general types of memory.

We introduce a function, called the k -lookback, $lk_k : \Sigma \rightarrow X^*$ such that given a sequence $\sigma = m_0 \dots m_n$ it is mapped under this function as $lk_k(m_0 m_1, \dots, m_n) = m_{\max(0, n-k)} \dots m_n$. The k -lookback therefore either returns the $k+1$ -length tail of σ , or the entirety of σ , depending on which is longer. Note that the string that is returned is of length $k+1$ and not k . This is because we would like the lookback to refer to the k measurements *before* the measurement being performed in that time step.

We introduce a property which we refer to as lk_k -consistency which a section $f \in \mathcal{S}(U)$ may satisfy. A lk_k -consistent section is one which maps measurements which have had the k same measurement events before them to the same outcome.

Definition 5.1.5. *A strategy $f \in \mathcal{S}(U)$ is said to be lk_k -consistent if for every pair of sequences $\sigma, \tau \in \Sigma$ it holds that*

$$lk_k(\sigma) = lk_k(\tau) \implies f(\sigma) = f(\tau). \quad (5.12)$$

Requiring a system have support only on the lk_k -consistent strategies can be seen as placing a specific type of memory restriction on the system. We may think of the memory as having k ‘space holders’, which correspond to the k previous measurement made. Once a new measurement is made, a measurement is kicked off the 0 placeholder and each measurement moves down one space to make space for the latest measurement. The other justification will come later, where we will see that empirical models which have global sections that contain support only on the lk_k -lookback strategies may have, at any measurement, signalling only from a window of depth k ; in other words, there is no signalling from measurements prior to the k measurements before.

Now for each set $U \in \mathcal{P}(\Sigma)$, we denote as $\mathcal{S}_{lk_k}(U)$ the subset of $\mathcal{S}(U)$ containing only strategies which are lk_k -consistent. We obtain the following order of inclusions

$$\mathcal{S}_{lk_0} \subseteq \mathcal{S}_{lk_1} \subseteq \mathcal{S}_{lk_2} \subseteq \dots \subseteq \mathcal{S}_{lk_\infty} = \mathcal{S}. \quad (5.13)$$

This reflects that the smaller the value of k the less strategies there are which satisfy the lk_k -consistency property, and in particular that if a strategy σ is lk_k -consistent then it is automatically lk_{k+1} -consistent. When we study distributions

on these strategies, the extreme ends of k capture notions of classicality already studied in the literature: when k is maximal we recover the notion of classicality of Gogioso and Pinzani [67] (instantiated to a special case of linear causal order). On the other hand, when k is minimal, a deterministic classical system is a map from each measurement to an outcome, without any consideration of which measurements had been performed before. In this case, we capture the correlations produced by macro-realistic models, first introduced by Leggett and Garg [95] and subsequently studied in other contexts [60, 105]. We give a more precise statement of these observations in a later section.

5.2 Empirical models for temporal scenarios

So far we have considered deterministic behaviours for a system, as captured by lk_k -consistent strategies, over a particular temporal setup described by the scenario $\mathcal{M} = \langle X, \Sigma, O \rangle$. As with standard contextuality, an empirical model will be a collection of distributions on such deterministic behaviours for an appropriate cover. While for contextuality scenarios this cover arises naturally as the commuting subsets of the measurement set, which form a simplicial complex, here instead we must define an appropriate cover of $\mathcal{P}(\Sigma)$.

Performing a particular sequence of measurements $\sigma \in \Sigma$ produces a probability distribution e_σ on corresponding output sequences, i.e. on assignments of outcomes to the set $\downarrow\sigma$ of prefix sequences of σ . The collection $e = (e_\sigma)_{\sigma \in \Sigma}$ is called an **empirical model**.

Definition 5.2.1 (Temporal empirical model). *Given a temporal measurement scenario $\mathcal{M} = \langle X, \Sigma, O \rangle$, an empirical model is a choice of R -distribution $e_\sigma \in \mathcal{D}_R \circ \mathcal{S}(\downarrow\sigma)$ for each sequence $\sigma \in \Sigma$ which are compatible on their intersection i.e. such that for any $\sigma', \sigma \in \Sigma$,*

$$e_\sigma|_{\downarrow\sigma' \cap \downarrow\sigma} = e_{\sigma'}|_{\downarrow\sigma' \cap \downarrow\sigma}. \quad (5.14)$$

The above compatibility constraints imply that there can be no signalling from the future to the past, which in the literature are referred to as the **Arrow of Time** (AoT) constraints [43]. Denote the set of all temporal empirical models associated with measurement scenario $\mathcal{M}$, $\text{EM}(\mathcal{M})$. For convenience, we will denote the cover $\mathcal{C} := \{\downarrow\sigma | \sigma \in \Sigma\}$.

As always, classicality of an empirical model e corresponds to the existence of a global section $h \in \mathcal{D}_R \circ \mathcal{S}(\Sigma)$. However, we can require also that this global section have support only on the lk_k -consistent strategies. This amounts to requiring that

the classical system may only use the past k measurements to determine the outcome of any measurement.

Definition 5.2.2 (lk_k -classicality). *An empirical model $e \in \text{EM}(\mathcal{M})$ is said to be lk_k -classical if there exists a global section $h \in \mathcal{D}_R \circ \mathcal{S}_{\text{lk}_k}(\Sigma)$ such that $h|_{\downarrow\sigma} = e_\sigma$ for all $\sigma \in \Sigma$.*

We would also like to define the empirical models which are lk_k -consistent, in the sense that they obey the same no-signalling relations as the lk_k -classical ones. Analogous to Bell scenarios with no temporal order, if *deterministically* measurements may only depend on the k preceding measurement choices, then *probabilistically* there can only be signalling from a window of depth k in the past. This motivates the introduction of the following additional no-signalling constraints². For any $\downarrow\sigma, \downarrow\sigma' \in \mathcal{C}$, consider the sets $C \subseteq \downarrow\sigma$ and $C' \subseteq \downarrow\sigma'$ with $\text{lk}_k(C) = \text{lk}_k(C')$. We require additionally that

$$e_\sigma|_C = e_{\sigma'}|_{C'}. \quad (5.15)$$

This for $k = \infty$ reduces to the usual AoT constraints, and for $k = 0$ imposes that the probability of any measurement to obtain a certain outcome does not depend on any of the previous measurements made. We consider a simple example to show how such additional no-signalling constraints are expressed in practice.

Example 5.2.3. *Consider the scenario with $X = \{a, a', b\}$ and $\Sigma = \{a, a', ab, a'b\}$. An empirical model e on this scenario consists of the distributions*

$$\begin{cases} e_a \in \mathcal{D}_R \circ \mathcal{S}(\downarrow\{a\}) \\ e_{a'} \in \mathcal{D}_R \circ \mathcal{S}(\downarrow\{a'\}) \\ e_{ab} \in \mathcal{D}_R \circ \mathcal{S}(\downarrow\{ab\}) \\ e_{a'b} \in \mathcal{D}_R \circ \mathcal{S}(\downarrow\{a'b\}) \end{cases} \quad (5.16)$$

subject to the AoT constraints, for example

$$e_{ab}|_{\{a\}} = e_a. \quad (5.17)$$

²Note that for contextuality setups the no-signalling empirical models correspond to affine combinations of deterministic (classical) empirical models. While we suppose that here the same is true and that the lk_k -consistent empirical models should be affine combinations of the lk_k -classical ones, we do not prove this and leave it for future work.

Equation [5.15](#) on the other hand for $k = 0$ requires also that for $\sigma = ab$ and $\sigma' = a'b$,

$$e_{ab}|_{\{ab\}} = e_{a'b}|_{\{a'b\}}. \quad (5.18)$$

This expresses that there is no signalling of the measurement choice (a or a') to b .

Probability distributions of the form $e_{ab}|_{\{ab\}}$ can be thought of as a type of conditional probability distribution—it expressed the probability of b to obtain a particular outcome given that the measurement before was a .

Definition 5.2.4 (lk_k -Consistent Empirical Model). *Given a measurement scenario $\mathcal{M}$ and an empirical model $e \in \text{EM}(\mathcal{M})$, we say that e is lk_k -consistent when for any $\downarrow\sigma, \downarrow\sigma'$ with $\sigma, \sigma' \in \Sigma$, and sets $C \subseteq \downarrow\sigma$ and $C' \subseteq \downarrow\sigma'$ with $\text{lk}_k(C) = \text{lk}_k(C')$, the empirical model satisfies*

$$e_\sigma|_C = e_{\sigma'}|_{C'}. \quad (5.19)$$

We will write $\text{EM}^{(\text{lk}_k)}(\mathcal{M})$ to denote the set of lk_k -consistent empirical models. Moreover, we write $\text{AoT}^{(k)}$ to denote the set of empirical models which obey the AoT constraints in addition to the additional constraints from equation [5.19](#), $\text{C}^{(k)}$ to denote the subset which are lk_k -classical, and $\text{Q}^{(k)}$ to denote the subset which are quantum realisable.

Recall that noncontextuality is synonymous with the existence of a hidden variable model $\langle \Lambda, h, i \rangle$ where Λ is a set of deterministic noncontextual hidden variables, $h \in \mathcal{D}_R(\Lambda)$ a probability distribution over them, and $i : \Lambda \rightarrow \mathcal{E}(X)$ associates a global section with each hidden variable. Here, we do not wish to make any hypotheses about the way in which physical systems—especially classical or quantum systems—produce outcomes when sequentially measured. The assumptions of macrorealism, for example, hypothesise about how a classical system should behave under measurement; namely that it should exist in a definite state at all times and measurements should be noninvasive. We instead define classicality with respect to memory bounds on a classical machine, which circumvents a debatable philosophical issue by appealing to resource-theoretic notions, where in this case the resource of interest is memory.

Definition 5.2.5 (E-machine). *Let $\mathcal{M} = \langle X, \Sigma, O \rangle$ be a measurement scenario. We define a classical machine E to be a quadruple $\langle X, O, F, h \rangle$, where*

1. X is the input set
2. O is the output set

C	$(1, 1)$	$(1, -1)$	$(-1, 1)$	$(-1, -1)$
$(a, 0) (b, 0)$	$1/8(3 + 2\sqrt{2})$	$1/8(3 - 2\sqrt{2})$	$1/8$	$1/8$
$(a, 0) (b, 1)$	$1/8(3 + 2\sqrt{2})$	$1/8$	$1/8(3 - 2\sqrt{2})$	$1/8$
$(a, 1) (b, 0)$	$1/8$	$1/8(3 - 2\sqrt{2})$	$1/8$	$1/8(3 + 2\sqrt{2})$
$(a, 1) (b, 1)$	$1/8(3 - 2\sqrt{2})$	$1/8(3 + 2\sqrt{2})$	$1/8$	$1/8$

Table 5.1

3. $F \subseteq \{f :: (\sigma \in \Sigma) \mapsto (o \in \prod_{m \in \sigma} O)\}$, is the set of possible input-output mappings over Σ inputs strings

4. h is a probability distribution on the set F .

The set F may be unconstrained, but usually we do wish to impose some constraints. For finite state machines, for example, in which the output is dependent only on the current input and current state, the size of the state set has an influence on which strings of outputs are producible.

A function $f :: (\sigma \in \Sigma) \mapsto (o \in \prod_{m \in \sigma} O)$, corresponds to the starred version g^* of a strategy $g \in \mathcal{S}(\Sigma)$. Thus, we can define F to be the set of functions satisfying lk_k -consistency. In particular, if $f \in F$ maps $i_0 i_1 \dots i_n \mapsto o_0 o_1 \dots o_n$ then the lk_k -constraint corresponds to requiring that

$$o_k = q(i_k, i_{k-1}, \dots, i_{\max n-(k+1), 0}), \quad (5.20)$$

so that any output in the sequence is only a function of the previous k inputs as well as the current input. We will write $\mathbf{E}_{\text{lk}_k}$ to represent an E-machine with F constrained in this way.

Given an empirical model e , we now instead ask whether it is classical with respect to an E-machine.

Definition 5.2.6 ($\mathbf{E}_{\text{lk}_k}$ -classical). *Given a measurement scenario $\mathcal{M}$, an empirical model $e \in \mathbf{EM}(\mathcal{M})$ is said to be $\mathbf{E}_{\text{lk}_k}$ -classical if there exists an E-machine $\mathbf{E}_{\text{lk}_k} = \langle X, O, F, h \rangle$ such that $h|_{\downarrow \sigma} = e_\sigma$ for all $\sigma \in \Sigma$.*

Note that all strategies in the support of e_σ for each $\sigma \in \Sigma$ can be produced in a single run of $\mathbf{E}_{\text{lk}_k}$; however, this does not imply that e is $\mathbf{E}_{\text{lk}_k}$ -classical.

Example 5.2.7 (Temporal CHSH). *We consider again the temporal CHSH setup studied in [61].*

An empirical model is a compatible family for the presheaf $\mathcal{D}_R \circ \mathcal{S}$ on the cover indexed over Σ . Consider a maximal element of this cover which is of the form

$\{(a, i), (a, i)(b, j)\}$ for $i, j \in \{0, 1\}$ i.e. $\downarrow\sigma$ for $\sigma = \{(a, i)(b, j)\}$. The AoT constraints imply

$$e_{(a,i)} = e_{(a,i)(b,j)}|_{(a,i)} \quad (5.21)$$

which is usually written

$$p(a_i = o_1) = \sum_{o_2} p(a_i = o_1, b_j = o_2) \quad (5.22)$$

and implies that the outcome at agent A does not depend on the choice of measurement or outcome at B . On the other hand, no such constraint is implied for B .

This data is equivalent to that in tabulated form, and we give an example in Table 5.1; this is the empirical model generated by two sequential dichotomic measurements on a qubit described by Bloch sphere vectors $\mathbf{a}_i$ and $\mathbf{b}_j$ for $i, j \in \{0, 1\}$, with an initial state given by Bloch sphere vector $\mathbf{v}$. It is known that the temporal CHSH inequality is maximally violated for any choice of $\mathbf{v}$ [27] provided

$$\begin{cases} \mathbf{a}_0 = \frac{1}{\sqrt{2}} (\mathbf{b}_0 + \mathbf{b}_1) \\ \mathbf{a}_1 = \frac{1}{\sqrt{2}} (\mathbf{b}_0 - \mathbf{b}_1) \end{cases} \quad (5.23)$$

We take

$$\mathbf{b}_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{b}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad (5.24)$$

and the initial state $|0\rangle$. The resulting empirical model is not extendable under the assumption of macrorealism for classical correlations. This can be checked using the incidence matrix approach of [6], or observed from trial and error. The empirical model in Table 5.1 moreover satisfies the AoT constraints which can easily be checked.

As we will see, macrorealism corresponds to E_{lk_0} -classicality. On the other hand, since the empirical model here is not lk_0 -consistent, it instead makes sense only to ask whether it is E_{lk_1} -classical, which it is. This highlights that our approach requires that nonclassical models obey the same no-signalling relations as the classical ones.

5.3 Mapping temporal setups to contextuality setups

For contextuality scenarios, we saw in section [3.8](#) that the problem of finding empirical models which have correlations in the set $\mathbf{NS}$ but not in the set $\mathbf{C}$ can be reduced to asking whether the simplicial complex describing the cover of the corresponding measurement scenario $\mathcal{M}$ is acyclic. We wish to derive a similar result for temporal measurement scenarios, except that in this case we cannot apply Vorob'ev's Theorem directly. We expect that separation between $\mathbf{C}$ and $\mathbf{AoT}$ depends on the bound k on the classical systems memory capacity, while application of the Theorem to the cover Σ directly depends in no way on k .

There is at least one way this can be solved, which is by mapping a temporal measurement scenario $\mathcal{M}$ to a set of contextuality measurement scenarios $\{\mathcal{M}'_k\}_{k \in I}$ indexed by an integer k , such that the empirical models on these latter measurements scenarios are in one-to-one correspondence with $\mathbf{lk}_k$ -consistent empirical models on $\mathcal{M}$. The upshot of doing so is that Vorob'ev's Theorem applied to these new scenarios describes, for each choice of k , whether there are $\mathbf{lk}_k$ -nonclassical empirical models on $\mathcal{M}$.

Theorem 5.3.1 (Mapping Theorem). *For all $k \in \mathbb{N} \cup \{\infty\}$ there is a map $\mathcal{C}_{\mathbf{lk}_k}$ from temporal measurement scenarios to contextuality measurement scenarios, such that empirical models $w \in \mathbf{EM}(\mathcal{C}_{\mathbf{lk}_k}(\mathcal{M}))$ can be pulled back via $\mathcal{C}_{\mathbf{lk}_k}^*$ to $\mathbf{lk}_k$ -consistent empirical models $e \in \mathbf{EM}^{(\mathbf{lk}_k)}(\mathcal{M})$ on the temporal measurement scenario $\mathcal{M}$. This map preserves and reflects nonclassicality, meaning that an empirical model w on $\mathcal{M}'_k$ which is in the domain of $\mathcal{C}_{\mathbf{lk}_k}$ is contextual if and only if e is a $\mathbf{lk}_k$ -nonclassical model.*

The proof of this theorem will be given explicitly in the following subsection. We begin by defining the map $\mathcal{C}_{\mathbf{lk}_k}$ which ‘flattens’ the temporal measurement scenario to a contextuality measurement scenario. Given any subset $U \subseteq \Sigma$ of measurement sequences, define

$$\mathcal{C}_{\mathbf{lk}_k}(U) := \{\mathbf{lk}_k(\sigma) \mid \sigma \in U\}. \quad (5.25)$$

Given a temporal measurement scenario $\mathcal{M} = \langle X, \Sigma, O \rangle$, define a contextuality scenario $\mathcal{M}'_k = \langle X'_k, \Sigma'_k, O'_k \rangle$ as

1. $X'_k := \mathcal{C}_{\mathbf{lk}_k}(\Sigma)$,
2. $\Sigma'_k := \{\mathcal{C}_{\mathbf{lk}_k}(C) \mid C \in \downarrow_{\subseteq} \mathcal{C}\}$,
3. $(O'_k)_{\mathbf{lk}_k \sigma} := O_m$ where $\sigma[-1] = m$.

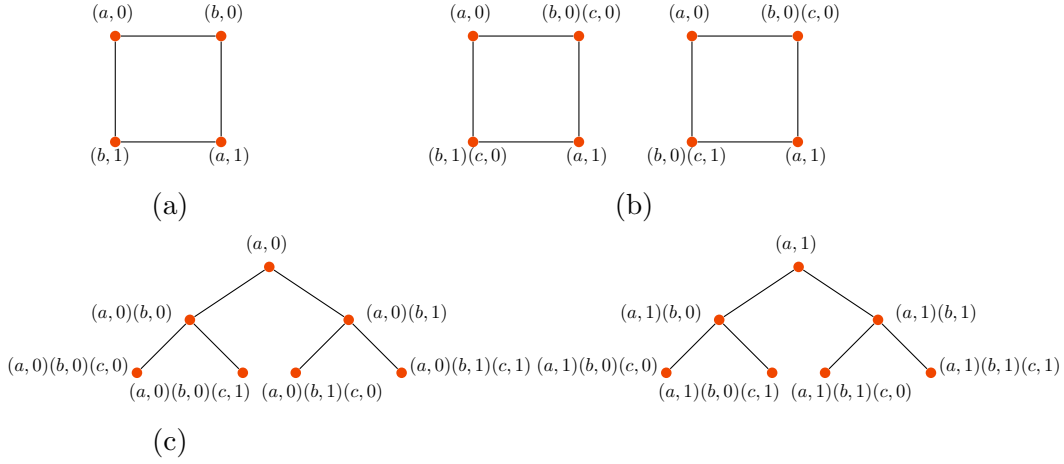


Figure 5.1: The image simplicial complexes, or subcomplexes thereof, under the map $\mathcal{C}_{\text{lk}_k}$ for the three party temporal Bell scenario, with binary inputs for agents. (a) image complex for $\text{lk}_k = 0$, (b) subcomplex of the image complex for $\text{lk}_k = 1$ (c) image complex for $\text{lk}_k = 2$.

The measurement set of the k th image scenario $\mathcal{M}'_k$ is the set of k -lookbacks of the sequences in Σ . Thus, for $k = 0$ we obtain a set of single measurements (analogous to the usual contextuality case) and in the case $k = \infty$ we obtain the set of sequences Σ . Next, a set of k -lookbacks forms a context $K \in \Sigma'_k$ only when their preimage under the k -lookback function is a valid context in $\downarrow_{\subseteq} \mathcal{C}$. Finally, the output set of the string of measurements $\text{lk}_k(\sigma)$ describing the k -lookback of the sequence σ is just given by the output set O_m of the terminal measurement m in that sequence.

Example 5.3.2 (3 agent temporal Bell scenario). *Consider the mapping $\mathcal{C}_{\text{lk}_k}$ of temporal Bell scenario with three agents, $\mathcal{A} = \{A, B, C\}$ where $A < B < C$, and in which each agent has two inputs, $I_A = I_B = I_C = \{0, 1\}$. The measurement scenario is therefore*

1. $X = \{(a, 0), (a, 1), (b, 0), (b, 1), (c, 0), (c, 1)\}$
2. $\sigma \in \Sigma \iff \exists i, j, k \in \{0, 1\}. (\sigma = (a, i)(b, j)(c, k) \vee \sigma = (a, i)(b, j) \vee \sigma = (a, i))$
3. $\forall m \in X. O_m := \{0, 1\}$.

The cover $\downarrow_{\subseteq} \mathcal{C} = \{C \mid \exists \sigma \in \Sigma. C \subseteq \downarrow \sigma\}$. There are three possible mappings in this case: $\mathcal{C}_{\text{lk}_0}, \mathcal{C}_{\text{lk}_1}$ and $\mathcal{C}_{\text{lk}_2}$, since all mappings $\mathcal{C}_{\text{lk}_k}$ for $k > 2$ are identical to the $k = 2$ case.

1. The $\mathcal{C}_{\text{lk}_0}$ mapping

- $X'_0 = \{m | m \in X\}$
- Σ'_0 is shown in Figure 5.1(a). Two important points are that (i) the simplicial complex is not Graham reducible and (ii) It is identical to the simplicial complex Σ for the spatial 2-party Bell scenario.
- $\forall m \in X'. O_m \in \{0, 1\}$.

2. The $\mathcal{C}_{lk_1}$ mapping

•

$$X'_1 = \{(a, 0), (a, 1), (a, 0)(b, 0), (a, 0)(b, 1), (a, 1)(b, 0), (a, 1)(b, 1), \\ (b, 0)(c, 0), (b, 0)(c, 1), (b, 1)(c, 0), (b, 1)(c, 1)\}$$

- The simplicial complex is too large to be informative when drawn, but it contains the induced subcomplexes shown in Figure 5.1(b). Note that the left subcomplex encodes the incompatibility of lookbacks: the outcome of $(c, 0)$ can only be obtained either after $(b, 0)$ or after $(b, 1)$ — not both. On the other hand, the right simplicial complex encodes the commutation relation of measurements: $(c, 1)$ and $(c, 0)$ cannot be simultaneously measured.
- $O_{\sigma=\dots m} := O_m$

The $\mathcal{C}_{lk_2}$ mapping

- $X'_3 := \Sigma$
- $\Sigma'_3 := \mathcal{C}$, which is shown in Figure 5.1(c).
- $O_{\sigma=\dots m} := O_m$

As is clear from the above example, for different values of k the mapping produces a different cover for the sheaf (as encoded by Σ'_k). For $k \geq 2$, the image scenario is actually identical to $\mathcal{M}$.

The next part of the theorem describes how the pullback of this map, denoted $\mathcal{C}_k^*$, can produce from an empirical model $w^{(k)} \in \text{EM}(\mathcal{M}'_k)$ a valid lk_k -consistent empirical model $e^{(k)} \in \text{EM}(\mathcal{M})$.

5.4 Mapping empirical models

We will begin by showing that lk_k -consistent strategies on any subset $U \subseteq \Sigma$ are in bijection with sections on the subset $U' := \mathcal{C}_{\text{lk}_k}(U)$.

Consider the set of lk_k -consistent strategies on U , recalling that we denote this set as $\mathcal{S}_{\text{lk}_k}(U)$, as well as the set of sections $\mathcal{E}(U')$, where $\mathcal{E}$ is the event sheaf as usually defined. We define a function

$$\alpha_U : \mathcal{S}_{\text{lk}_k}(U) \rightarrow \mathcal{E}(U') :: f \mapsto s \quad (5.26)$$

which sends a lk_k -consistent strategy $f \in \mathcal{S}_{\text{lk}_k}(U)$ to a section $s \in \mathcal{E}(U')$. Define s to be the function

$$s(l) := f(\sigma) \quad (5.27)$$

where l is the k -lookback of the sequence σ . Note that although we could have σ, σ' with $\text{lk}_k(\sigma) = \text{lk}_k(\sigma')$ so that the above seems ambiguous, both give the same definition of s due to f being lk_k -consistent.

Proposition 5.4.1. *The function α_U is a bijection for all $U \subseteq \Sigma$.*

Proof. It has an inverse α_U^{-1} which sends $s \in \mathcal{E}(U')$ to $f \in \mathcal{S}_{\text{lk}_k}(U)$ where $f(\sigma) := s(\text{lk}_k(\sigma))$. The result is a map which lk_k -consistent. $\square$

For any $U \subseteq \Sigma$, using the pullback $\mathcal{C}_{\text{lk}_k}^* \mathcal{E}(U) := \mathcal{E}(\mathcal{C}_{\text{lk}_k}(U))$ we obtain sections on the k -lookback of sequences in U directly, and these are in bijection with those on $\mathcal{S}_{\text{lk}_k}(U)$. This bijection extends also to the probability distributions on these sections, so that we obtain

$$\mathcal{D}_R \circ \mathcal{C}_{\text{lk}_k}^* \mathcal{E}(U) \cong^{\alpha_U} \mathcal{D}_R \circ \mathcal{S}_{\text{lk}_k}(U). \quad (5.28)$$

Take an empirical model $w \in \mathbf{EM}(\mathcal{M}'_k)$. Recall that this consists of a probability distribution $w_L \in \mathcal{D}_R \circ \mathcal{E}(L)$ for each $L \in \Sigma'_k$. Now define an empirical model e to be

$$e_C(f) = w_{\mathcal{C}_{\text{lk}_k}(C)}(\alpha_C(f)). \quad (5.29)$$

Note that this is identical to the expression we obtained when considering simulations, so that alternatively e is the pushforward along the morphism of measurement scenarios $\langle C_{\text{lk}_k}, \alpha \rangle : \mathcal{M}'_k \rightarrow \mathcal{M}$. The final part of the Theorem is to show that this map reflects and preserves classicality.

Lemma 5.4.2. *The induced map between empirical models $w \mapsto e$ given in equation 5.29 preserves and reflects classicality.*

Proof. We start by showing that classicality is preserved. Assume e is classical, so that there exists a global section $h \in \mathcal{D}_R \circ \mathcal{S}_{\text{lk}_k}(\Sigma)$ which marginalises to e on the cover $\mathcal{C}$. Let $\delta_s(s')$ be equal to 1 iff $s = s'$, and 0 otherwise. Here, $s, s' \in \mathcal{F}(U)$ where $\mathcal{F} \in \{\mathcal{E}, \mathcal{S}_{\text{lk}_k}\}$ and $U \subseteq X'_k$ for $\mathcal{F} = \mathcal{E}$, and $U \subseteq \Sigma$ for $\mathcal{F} = \mathcal{S}_{\text{lk}_k}$. Then we have that (leaving the subscript off α to simplify presentation)

$$\begin{aligned} w &= e_{\mathcal{C}_{\text{lk}_k}(C)}(\alpha(f)) \\ &= \sum_{s \in \mathcal{S}_{\text{lk}_k}(\Sigma)} \delta_{s|\mathcal{C}_{\text{lk}_k}(C)}(\alpha(f)) h(s) \\ &= \sum_{\alpha^{-1}(s) \in \mathcal{E}(X'_k)} \delta_{\alpha^{-1}(s)|C}(f) \alpha^{-1}(h)(\alpha^{-1}(s)) \end{aligned}$$

where the first line follows by definition (equation 5.29), the second from e being a classical empirical model and therefore the existence of a global section $s \in \mathcal{S}_{\text{lk}_k}(\Sigma)$, and the last can be seen to be true by checking equality of each term (the bijection ensures this holds). Thus $\alpha^{-1}(h)$ is a valid global section for w .

Showing that classicality is reflected follows similarly, by first assuming that w is classical. □

We have now completed the proof of Theorem 5.7.2.

5.5 No-signalling constraints beyond the AoT constraints

The AoT constraints contained in equation 5.14 express that the probability of a measurement obtaining a particular outcome should be independent of the choice of measurements at a later time. This can be implemented in a loophole free way by ensuring measurements are time-like separated.

However, as we have briefly discussed, the k -lookback constraints impose additional no-signalling constraints. Such additional constraints on an empirical model $e \in \text{EM}(\mathcal{M})$ become clear when considering the image scenario $\mathcal{M}'_k := \mathcal{C}_{\text{lk}_k}(\mathcal{M})$ and its corresponding empirical models $\text{EM}(\mathcal{M}'_k)$.

Take $w \in \mathbf{EM}(\mathcal{M}'_k)$. Note that we can have two sets in the cover $C_1, C_2 \in \downarrow_{\subseteq} \mathcal{C}$ which have intersection $C_1 \cap C_2$ that is not in $\mathcal{C}$, but which do have intersection in $\downarrow_{\subseteq} \mathcal{C}$. Therefore, we obtain the no-signalling constraint on w :

$$w_{\mathcal{C}_{\text{lk}_k}(C_1)}|_{\mathcal{C}_{\text{lk}_k}(C_1) \cap \mathcal{C}_{\text{lk}_k}(C_2)} = w_{\mathcal{C}_{\text{lk}_k}(C_2)}|_{\mathcal{C}_{\text{lk}_k}(C_1) \cap \mathcal{C}_{\text{lk}_k}(C_2)}. \quad (5.30)$$

Example 5.5.1. Consider again the measurement scenario of the three party temporal Bell setup, and consider the set of empirical models $w \in \mathbf{EM}(\mathcal{M}'_1)$ on the image scenario $\mathcal{M}'_1 := \mathcal{C}_1(\mathcal{M})$. Consider the sets $C_1 = \{(a, 0), (a, 0)(b, 0), (a, 0)(b, 0)(c, 0)\}$ and $C_2 = \{(a, 1), (a, 1)(b, 0), (a, 1)(b, 0)(c, 0)\}$. Under $\mathcal{C}_1$ these become

$$C'_1 := \mathcal{C}_1(C_1) = \{(a, 0), (a, 0)(b, 0), (b, 0)(c, 0)\} \quad (5.31)$$

$$C'_2 := \mathcal{C}_1(C_2) = \{(a, 1), (a, 1)(b, 0), (b, 0)(c, 0)\} \quad (5.32)$$

which have nonempty intersection $C'_1 \cap C'_2 = \{(b, 0)(c, 0)\}$. Therefore, we obtain the no-signalling constraint

$$w_{C'_1}|_{C'_1 \cap C'_2} = w_{C'_2}|_{C'_1 \cap C'_2} \quad (5.33)$$

or

$$w_{\{(a,0),(a,0)(b,0),(b,0)(c,0)\}}|_{\{(b,0)(c,0)\}} = w_{\{(a,1),(a,1)(b,0),(b,0)(c,0)\}}|_{\{(b,0)(c,0)\}}. \quad (5.34)$$

In Figure 5.2 we give an abstract representation of the convex set of correlations which can be achieved under different assumptions. The largest set, the grey set labelled $\mathbf{AoT}^{(k)}$, represents those correlations which obey the AoT no-signalling constraints as well as the additional ones imposed by lk_k -consistency. The green set $\mathbf{Q}^{(k)}$ represents the subset of these which are quantum realisable. Finally, the yellow subset $\mathbf{C}^{(k)}$ is the set of $\mathbf{E}_{\text{lk}_k}$ -classical correlations.

What do these new no-signalling equations express? They express that there can be no signalling of the measurement choice outside of the k -lookback window. For k maximal, such no-signalling constraints coincide with the AoT constraints imposed by the arrow of time. For k minimal and equal to zero, the no-signalling constraints are identical to those of a contextuality setup, which are referred to as the ‘no-disturbance’ conditions. These additional constraints for the case $k = 0$ are, in the literature, also known as **no-signalling in time** (NSIT) constraints [71].

The Mapping Theorem allows us to prove the two following propositions.

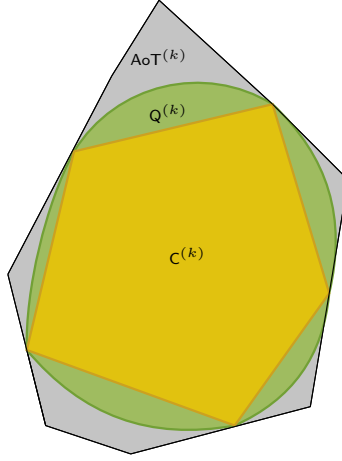


Figure 5.2: The convex sets of empirical models which are lk_k -consistent and E_{lk_k} -classical (denoted $\mathbf{C}^{(k)}$), or E_{lk_k} -nonclassical (denoted $\mathbf{Q}^{(k)}$ for the quantum realisable, while $\text{AoT}^{(k)}$ denotes those correlations obeying AoT in addition to the k -lookback window constraints)

Proposition 5.5.2. *Take a temporal measurement scenario $\mathcal{M}$. If a empirical model $e \in \text{EM}(\mathcal{M})$ is lk_0 -classical, or equivalently extendable, then this implies that the system obeys macroscopic realism.*

Proposition 5.5.3. *Take a temporal Bell scenario $\mathcal{M}$. The set of empirical models $e \in \text{EM}(\mathcal{M})$ is in bijection with the set of empirical models on $\mathcal{M}$ as defined by Gogioso and Pinzani in [66].*

We prove each of these in turn. We start with Proposition 5.5.2.

Proof. If e is lk_k -classical then there exists a global section $h \in \mathcal{D}_{\mathbb{R}} \circ \mathcal{S}_0(\Sigma)$. From Theorem 5.7.2, we can equivalently consider whether the image empirical model $w \in \mathcal{C}_0(e)$ is classical, since classicality is preserved and reflected. We have seen that sections on $\mathcal{M}'$ are of the form $s : V \rightarrow \mathcal{O}$ with V being some set of measurements. The proof of factorisation then follows directly from a proof contained in [6]. Given a section $s \in \mathcal{O}^{X'}$ let $\delta_s(s')$ be the distribution which is 1 when $s = s'$ and zero otherwise. Moreover, given a subset $V \subseteq X'$ we can restrict the distribution $\delta_s|_V$. It can be shown that $\delta_s|_V = \delta_{s|_V}$, where in the former case it is the delta distribution which is restricted, while in the latter case it is the section s which is restricted; therefore, we can interchange between these two restrictions. It then follows that

$$w_\sigma(s) = \sum_{s' \in \mathcal{O}', s'|_\sigma = s} h(s') = \sum_{s' \in \mathcal{O}^{X'}} \delta_{s'|_\sigma}(s) \cdot h(s') = \sum_{s' \in \mathcal{O}^{X'}} \delta_{s'}|_\sigma(s) \cdot h(s'). \quad (5.35)$$

Moreover, it can be checked that the following factorisation holds

$$\delta_{s'}|_{\sigma}(s) = \prod_{m \in \sigma} \delta_{s'}|_{\{m\}}(s|\{m\}). \quad (5.36)$$

Altogether this says that

$$e_{\sigma}(s) = \sum_{s' \in O^{X'}} \prod_{m \in \sigma} \delta_{s'}|_{\{m\}}(s|\{m\}) \cdot h(s'). \quad (5.37)$$

If we take $\sigma = m_{i_0}^0, \dots, m_{i_n}^n$ (so that the choice of measurement at time step j is $\{m_0^j, m_1^j\}$) and $s(m_i^k) = o_k$, then this is exactly equation [4.25](#) where each $s' \in O^{X'}$ is identified with a hidden variable $\lambda \in \Lambda$ and $h(s')$ gives the probability of obtaining the hidden variable s' . $\square$

We now prove Proposition [5.5.3](#).

Proof. We describe the bijective correspondence in the specific case $N = 2$, so with $\Omega = \{A, B\}$ and $A < B$. It is easy to see how it generalises to $N > 2$, although notationally quite a bit more involved. Therefore we have a temporal Bell scenario $\mathcal{M} = \langle \Omega, \underline{I}, \underline{O} \rangle$ and we take $I = O = \{0, 1\}$ for both A and B . In order to show a bijection of the set of empirical models $\mathbf{EM}^{\text{GP}}(\mathcal{M})$ and $\mathbf{EM}(\mathcal{M})$, it is enough to show that sections on an arbitrary element $(S, \{U_{\omega}\}_{\omega \in S})$ are in bijection with strategies on a corresponding element of $\mathcal{P}(\Sigma)$. Take for example $S = \{A, B\}$ and $U_A = \{0, 1\}$ and $U_B = \{0\}$. Then a section over $(S, \{U_{\omega}\}_{\omega \in S})$ is a pair of functions

$$\begin{aligned} f_A & : U_A \rightarrow O_A \\ f_B & : U_A \times U_B \rightarrow O_B. \end{aligned} \quad (5.38)$$

Consider the subset $V \subseteq \Sigma$ defined $V = \{(a, 0), (a, 1), (a, 0)(b, 0), (a, 1)(b, 0)\}$. A strategy over V is a function

$$s : V \rightarrow O. \quad (5.39)$$

Clearly, there is a bijective correspondence between sections on U and strategies on V by defining $s((a, i)) = f_A(i)$ and $s((a, i)(b, j)) = f_B(i, j)$. This extends to the empirical models, which are simply distributions on these sections or strategies. $\square$

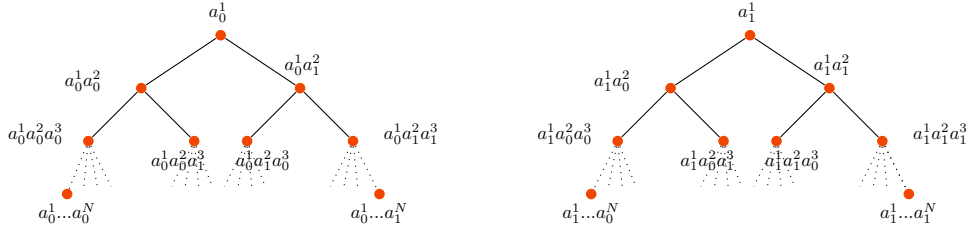


Figure 5.3: The compatibility complex for the image scenario of the N party temporal Bell scenario in the case that $k \geq N - 1$. Here, any branch of the tree is a context of the simplicial complex, so that the displayed complex is acyclic.

5.6 Combinatorial conditions for classicality

Combining Vorob'ev's Theorem with the Mapping Theorem above results in the following Theorem.

Theorem 5.6.1. *Let $\mathcal{M}$ be a temporal measurement scenario. Every empirical model $e \in \text{EM}^{(l_k)}(\mathcal{M})$ is E_{l_k} -classical iff the corresponding contextuality measurement scenario $\mathcal{C}_{l_k}(\mathcal{M})$ has acyclic simplicial complex Σ'_k .*

Proof. Suppose first that the simplicial complex Σ'_k associated with the measurement scenario $\mathcal{C}_{l_k}(\mathcal{M})$ is acyclic. From Theorem 5.7.2, every empirical model on top of Σ'_k is noncontextual. Each of these corresponds to an empirical model on $\mathcal{M}$, which, together with the property that $\mathcal{C}_{l_k}$ preserves and reflects classicality of e , is enough to say that every empirical model on $\mathcal{M}$ must, too, be k -lookback classical. On the other hand, suppose that every l_k -consistent empirical model on $\mathcal{M}$ is l_k -classical. Since every empirical model w on $\mathcal{C}_{l_k}(\mathcal{M})$ pullsback to give one of these empirical model on $\mathcal{M}$, and moreover this map preserves and reflects classicality, it must be that every empirical model $w \in \text{EM}(\mathcal{C}_{l_k}(\mathcal{M}))$ is noncontextual. From 3.8.1, we therefore know that the simplicial complex Σ'_k must be acyclic. $\square$

One immediately sees that all simplicial complexes of a measurement scenario $\mathcal{M}$ arising under the mapping $\mathcal{C}_{l_k}$ for $k \geq \max(|\sigma|) - 1$ are tree-like and, most importantly, acyclic. We display the general form of Σ'_k for these parameter regimes below (to save space, we switch notation letting a_i^m denote the m -th measurement of agent i , rather than (i, m)).

We therefore arrive at the following proposition.

Proposition 5.6.2. *Let $\mathcal{M} = \langle X, \Sigma, O \rangle$ be a temporal measurement scenario, and let N be the length of the longest sequence in Σ , $N := \max_{\sigma \in \Sigma} |\sigma|$. Then for all $k \geq N - 1$, every empirical model $e \in \text{EM}^{(l_k)}(\mathcal{M})$ is classical.*

Proof. Immediate. □

For the special case of temporal Bell scenarios, N is equal to the number of agents in the setup. We might wonder, then, what combinatorial conditions on Σ ensure that Σ'_k is not acyclic, thereby guaranteeing that an empirical model that is not classically realisable—but satisfies the k -lookback no-signalling requirement—exists.

We give a condition on the agents and their inputs which is both sufficient and necessary for Σ'_k to not be acyclic.

Proposition 5.6.3. *Let $\mathcal{M} = \langle X, \Sigma, O \rangle$ correspond to a temporal Bell scenario with N linearly ordered agents $\mathcal{A}$ with I_j denoting the input set of an agent $j \in \mathcal{A}$. The simplicial complex Σ'_k is Graham irreducible if and only if there exist at least two agents i and j , w.l.o.g we take $i < j$, and with $|I_i| \geq 2$ and $|I_j| \geq 2$ such that either*

1. $j - i > k$ or
2. There exists an $i < j < l$ with $l - i > k$ and $l - j \leq k$.

Proof. 1. **sufficiency.** For item 1, note that if $j - i > k$ then measurements performed by agent i are never contained in the k -lookback of agent j (and vice versa). Thus, we obtain a square as an induced subcomplex. We recall from proposition [2.2.5](#) that if there exists an induced subcomplex K which is Graham irreducible, then this implies that Σ is Graham irreducible. Take $\sigma_1, \sigma_2, \sigma_3, \sigma_4 \in \Sigma$ where

$$\begin{aligned} \sigma_1[-1] &= (i, 0) \\ \sigma_2[-1] &= (i, 1) \\ \sigma_3[-1] &= (j, 0) \\ \sigma_4[-1] &= (j, 1) \end{aligned} \tag{5.40}$$

and where for every pair of sequences $\sigma_m[r] = \sigma_n[r]$ for all other agents r (not including i and j). Then the image scenario $\mathcal{C}_{\text{lk}_k}(\mathcal{M})$ will have as a subcomplex the simplicial complex with vertices

$$\text{lk}(\sigma_1), \text{lk}_k(\sigma_2), \text{lk}_k(\sigma_3), \text{lk}_k(\sigma_4) \tag{5.41}$$

and maximal faces

$$\left\{ \begin{array}{l} \{\text{lk}_k(\sigma_1), \text{lk}_k(\sigma_3)\} \\ \{\text{lk}_k(\sigma_1), \text{lk}_k(\sigma_4)\} \\ \{\text{lk}_k(\sigma_2), \text{lk}_k(\sigma_3)\} \\ \{\text{lk}_k(\sigma_2), \text{lk}_k(\sigma_4)\} \end{array} \right. \quad (5.42)$$

which is a square and therefore irreducible via a sequence of Graham reduction steps. For item 2, if $l-i > k$ for some party l then the measurements performed by agent i are never in the k -lookback of agent l . On the contrary, if $l-j \leq k$ then l will contain measurements performed by j in its k -lookback. We are again able to find an induced subcomplex of the image scenario $\mathcal{C}_{\text{lk}_k}(\mathcal{M})$ which is not Graham reducible. This time, take $\sigma_1, \sigma_2, \sigma_3, \sigma_4 \in \Sigma$ where

$$\begin{aligned} \sigma_1[-1] &= (i, 0) \\ \sigma_2[-1] &= (i, 1) \\ \sigma_3[-1] &= (l, 0) \quad \& \quad \sigma_3[j] = (j, 0) \\ \sigma_4[-1] &= (l, 0) \quad \& \quad \sigma_4[j] = (j, 1) \end{aligned} \quad (5.43)$$

and where for every pair of sequences $\sigma_m[r] = \sigma_n[r]$ for all other agents r (not including i, j and l). Then the image scenario $\mathcal{C}_{\text{lk}_k}(\mathcal{M})$ will again have a square as the induced subcomplex with vertices

$$\text{lk}(\sigma_1), \text{lk}_k(\sigma_2), \text{lk}_k(\sigma_3), \text{lk}_k(\sigma_4)\} \quad (5.44)$$

and maximal faces

$$\left\{ \begin{array}{l} \{\text{lk}_k(\sigma_1), \text{lk}_k(\sigma_3)\} \\ \{\text{lk}_k(\sigma_1), \text{lk}_k(\sigma_4)\} \\ \{\text{lk}_k(\sigma_2), \text{lk}_k(\sigma_3)\} \\ \{\text{lk}_k(\sigma_2), \text{lk}_k(\sigma_4)\} \end{array} \right. \quad (5.45)$$

2. **necessity.** In order to prove necessity we prove the converse statement—that is, if the condition on the input set sizes and the separation between agents i

and j does not hold then this implies that Σ'_k is acyclic, and therefore reducible. Therefore, suppose first that we do not have at least two agents i and j with both $|I_i| \geq 2$ and $|I_j| \geq 2$. We denote measurements for i as (i, s) with $s \in I_i$ and same for j . We therefore have either only one agent in $\mathcal{A}$ with $I_i \geq 2$. Then the number of maximal sequences in Σ is given by $|I_i|$, and the image scenario has $|I_i|$ vertices: $\mathcal{C}_{|k_k}(\sigma)$ for each $\sigma \in \Sigma$. This image is always acyclic. This is because vertices $v \in \Sigma'_k$, which we recall are elements of X^* and therefore a string of measurements, will either (a) contain a single measurement of the form (j, s) or (b) not contain any (j, s) 's. All vertices for which (b) holds will fall into a single context $C \in \Sigma'_k$. Vertices for which (a) holds will always lie in a single maximal context $C_{b_i} \in \Sigma'_k$, and therefore all vertices for which (a) holds can be removed in a Graham reduction step. This leaves us only with vertices of type (b). But since these lie in the single context C , and after removal of vertices (a) this is the only context removing, the complex is acyclic. Now suppose we do have $|I_i| \geq 2$ and $|I_j| \geq 2$ for some agents i and j , but that both conditions 1 and 2 do not hold. If conditions 1 and 2 do not hold, then the vertices containing (j, s) always occur in a single maximal context, and can be removed. Likewise, vertices containing neither (j, s) nor (i, s') vertices occur in a single maximal context and can be removed in a Graham reduction step. This now leaves vertices containing (i, s) measurements. These cannot be in the same context, since they never occur in the same sequence. The result is an acyclic complex. $\square$

Key to the proof of sufficiency is the fact that if either (i) or (ii) holds then the image complex $\mathcal{C}_{|k_k}(\Sigma)$ will be Graham irreducible due to the presence of a square. This then allows us to invoke Vorob'ev's Theorem. Necessity on the other hand follows by considering all cases in which we do not have at least two agents with at least two inputs, and showing that this always results in an acyclic simplicial complex.

It is informative to study the subcomplexes displayed in Figure 5.1(b), which correspond to a 3 agent temporal Bell scenario with $k = 1$. The left-hand subcomplex corresponds to the case in which $i = 0$, which corresponds to agent A , and $j = 1$ which corresponds to agent B . Then since for $l = 2$, we have $|l - i| = |2 - 0| = 2 > 1$ and $|l - j| = |2 - 1| = 1 \leq 1$ we get a square subcomplex which is acyclic, so that the entire simplicial complex must be acyclic.

On the other hand, the right-hand subcomplex corresponds to the case in which $i = 0$ and $j = 2$ corresponding to agents A and C respectively. Then since $|j - i| =$

$2 > 1$ there exists a subcomplex of $\mathcal{C}_1(\Sigma)$ which is a square. As such, we expect to be able to construct an empirical model $e \in \text{EM}^{(\text{lk}_k)}(\mathcal{M})$ which is nonclassical; we do this below simplifying to one input for B (so that we consider acyclicity which arises due to (i)).

Example 5.6.4 (3-party temporal Bell). *We now construct a non-classical empirical model for the 3-party temporal Bell scenario, with the input at agent B taken to be the singleton $\{0\}$ and the inputs at A and C to be $\{0, 1\}$.*

	000	001	010	011	100	101	110	111
$(a, 0) (b, 0) (c, 0)$	1/4	0	1/4	0	0	1/4	1/4	0
$(a, 0) (b, 0) (c, 1)$	1/4	0	1/4	0	0	1/4	1/4	0
$(a, 1) (b, 0) (c, 0)$	1/4	0	1/4	0	0	1/4	1/4	0
$(a, 1) (b, 0) (c, 1)$	0	1/4	0	1/4	1/4	0	1/4	0

Table 5.2

It can be checked that it satisfies the required no-signalling constraints:

$$e_{\{(a,i),(a,i)(b,j),(a,i)(b,j)(c,k)\}}(s)|_{\{(a,i)\}} = e_{\{(a,i),(a,i)(b,j),(a,i)(b,j)(c,k)\}}(s)|_{\{(a,i)\}}$$

$$e_{\{(a,i),(a,i)(b,j),(a,i)(b,j)(c,k)\}}(s)|_{\{(a,i)(b,j)(c,k)\}} = e_{\{(a,i'),(a,i')(b,j),(a,i')(b,j)(c,k)\}}(s)|_{\{(a,i)(b,j)(c,k)\}}.$$

It is therefore a valid lk_1 -consistent empirical model, so we have an empirical model $e \in \text{EM}^{(\text{lk}_1)}(\mathcal{M})$. We will show that the e is non-classical by showing that there is no 1-lookback global strategy which is consistent with the support of this empirical model.

Proposition 5.6.5. *The empirical model defined in table [5.2](#) is lk_1 -consistent but nevertheless E_{lk_1} -nonclassical.*

Proof. Consider the image scenario $\mathcal{C}_1(\mathcal{M})$ and the empirical model $w \in \text{EM}(\mathcal{C}_1(\mathcal{M}))$ of which e is the pullback. We will prove that w must be non-classical by contradiction, and so begin by assuming that there exists a probability distribution h on global strategies $\mathcal{E}(X'_1)$ consistent with w . As a weaker condition, we would like to find a single section $s \in \mathcal{E}(X'_1)$ which is consistent with the support of w . This section will be of the following form

$$f :: \{(a, 0) \mapsto o_1, (a, 0)(b, 0) \mapsto o_2, (b, 0)(c, 0) \mapsto o_3, (b, 0)(c, 1) \mapsto o_4, (a, 1) \mapsto o_5, (a, 1)(b, 0) \mapsto o_6\}.$$

For example, looking at table [5.2](#) we might take $o_1 = 0$ and $o_2 = 0$, then we must take $o_3 = 0$ (it is the only way that this section will have support on entries which occur with non-zero probability) and $o_4 = 0$. No matter the choice of o_5 and o_6 , there will be a probability in the table which is zero but which f contains in its support. This is true regardless of how $o_1 - o_6$ are defined, as can be checked. Thus, w is non-classical and therefore e is non-classical (from Theorem [5.7.2](#)). $\square$

5.7 Θ -memory

We now consider the case in which a measurement can depend on up to k previous measurements, although these need not be the k preceding measurements. This can be envisioned by, at time step T , the memory register storing k measurements which occurred previously ³. At time step $T + 1$ a measurement m is performed and the system can either (i) kick off one of the k measurements and replace it with m or (ii) keep it as is.

We first briefly introduce some additional terminology concerning sequences: we write $\tau \preceq \sigma$ when τ is a subsequence of σ , meaning it can be obtained from σ by deleting certain measurements but not changing the order of the remaining measurements. Moreover if $\tau \leq \sigma$ we write $\sigma - \tau$ to denote the tail of the sequence of σ (where the tail refers to measurements at the end of the sequence) not contained in τ . Recall also that we write $\tau \leq \sigma$ if τ is a prefix sequence of σ .

We then define a Θ -function to be a function $\Theta : \Sigma \rightarrow X^*$ which satisfies the following three properties

1. $\Theta(\sigma) \preceq \sigma$ for all $\sigma \in \Sigma$,
2. $\Theta(\sigma)[-1] = \sigma[-1]$ for all $\sigma \in \Sigma$, and
3. for $\sigma_1 \leq \sigma_2$, $\Theta(\sigma_2) = s_1 s_2$ where $s_1 \preceq \Theta_k(\sigma_1)$ and $s_2 \preceq \sigma_2 - \sigma_1$.

The first requirement says that Θ should map any sequence σ to a subsequence of σ . The second requirement is that the final value in the measurement sequence $\Theta(\sigma)$ be the terminal measurement of σ . Finally, the third requirement is that given two sequences σ_1, σ_2 with σ_1 a prefix sequence of σ_2 , the value of Θ at σ_2 should be a concatenation of a subsequence of Θ at σ_1 and a subsequence of the tail of σ_2 not in σ_1 .

³If less than k measurements have occurred previously, all previous measurements are stored

We then say that a strategy $f \in \mathcal{S}(U)$ is Θ -consistent if

$$\Theta(\sigma) = \Theta(\tau) \implies f(\sigma) = f(\tau), \quad (5.46)$$

and given $U \subseteq \Sigma$ and a choice of Θ , write $\mathcal{S}_\Theta(U)$ for the set of Θ -consistent strategies on U .

The lk_k function arises as a special case of this by defining $\Theta(\sigma) = \text{lk}_k(\sigma)$. It is straightforward then to define what is meant by $\mathbf{E}_\Theta$ -classicality; this is a straightforward generalisation of Definition [5.2.6](#).

Definition 5.7.1 ($\mathbf{E}_\Theta$ -classicality). *An empirical model $e \in \mathbf{EM}(\mathcal{M})$ is said to be $\mathbf{E}_\Theta$ -classical if there exists an $\mathbf{E}$ -machine $E_\Theta = \langle X, O, F_\Theta, h \rangle$ such that $h|_{\downarrow\sigma} = e_\sigma$ for all $\sigma \in \Sigma$.*

It is of particular interest whether an empirical model has support on Θ -functions which are restricted to having in their codomain subsequences of length at most $k+1$. This amounts to considering functions which store arbitrary l measurements from the past, with $l \leq k$, as well as the terminal measurement m on σ . For example, the lookback_l function is for each $l \leq k$ an example of such a function.

Consider the Θ -functions which satisfy for fixed k and all $\sigma \in \Sigma$

$$|\Theta(\sigma)| \leq k + 1, \quad (5.47)$$

and write L_k for the set of Θ -functions which satisfy this property, bearing in mind that the actual memory occupies just the first k space holders. We can at this point either consider empirical models which are constrained to use a single $\Theta \in L_k$ in each round of measurement, or one which can probabilistically choose any $\Theta \in L_k$ at the start of measurement. The latter is more powerful as it encompasses the former. For $U \subseteq \Sigma$ we define the set of strategies

$$\mathcal{S}_{L_k}(U) := \bigcup_{\Theta \in L_k} \mathcal{S}_\Theta(U). \quad (5.48)$$

A strategy $s \in \mathcal{S}_{L_k}(U)$ is Θ -consistent for some $\Theta \in L_k$. We can then take lk_k -classicality to be as before, where we replace $\mathcal{S}_{\text{lk}_k}$ with $\mathcal{S}_{L_k}$. We refer to these as L_k -classical empirical models, where $\mathbf{E}_{L_k}$ refers to the probabilistic $\mathbf{E}_{L_k}$ -machine $\langle X, O, F_{L_k}, h \rangle$.

It turns out that these empirical models can still be associated with those on an image scenario, but the image scenario is formed from the controlled choice operation

(see Section 3.6)— these latter scenarios will be the images of $\mathcal{M}$ under a map $\mathcal{C}_\Theta$ for fixed $\Theta \in L_k$.

We recall the **controlled choice** operation. Given two measurement scenarios $\mathcal{M} = \langle X, \Sigma, O \rangle$ and $\mathcal{M}' = \langle X', \Sigma', O' \rangle$ and empirical models $e \in \mathbf{EM}(\mathcal{M})$ and $e' \in \mathbf{EM}(\mathcal{M}')$, we can form a measurement scenario $\langle X \sqcup X', \Sigma + \Sigma', [O, O'] \rangle$ and empirical model $e \& e'$ which is the controlled choice of e and e' . The choice of compatibility $\Sigma + \Sigma' := (\{0\} \times \Sigma) \cup (\{1\} \times \Sigma')$ ensures that in any round a choice of consistent subset of measurements comes from either $\mathcal{M}$ or $\mathcal{M}'$. The empirical model $e \& e'$ is defined

$$\begin{aligned} (e \& e')_{(\{0\} \times \sigma)} &:= e_\sigma \quad \text{when } \sigma \in \Sigma \\ (e \& e')_{(\{1\} \times \sigma)} &:= e'_\sigma \quad \text{when } \sigma \in \Sigma'. \end{aligned} \tag{5.49}$$

Theorem 5.7.2 (Generalised Mapping Theorem). *For all $k \in \mathbb{N} \cup \{\infty\}$ there is a map $\mathcal{C}_{L_k}$ from temporal measurement scenarios to contextuality measurement scenarios, such that empirical models $w \in \mathbf{EM}(\mathcal{C}_{L_k}(\mathcal{M}))$ can be pulled back via $\mathcal{C}_{L_k}^*$ to Θ -consistent empirical models $e \in \mathbf{EM}^{(L_k)}(\mathcal{M})$ on the temporal measurement scenario $\mathcal{M}$. This map preserves and reflects classicality, meaning that an empirical model w on $\mathcal{M}'_k$ which is in the image of $\mathcal{C}_{L_k}$ is contextual if and only if $e^{(L_k)}$ is an L_k -consistent nonclassical model.*

The first thing that must be shown is that for fixed $\Theta \in L_k$ the map $\mathcal{C}_\Theta$ is a straightforward generalisation of the $\mathcal{C}_{l_k}$ functions. Define, for any $\sigma \in \Sigma$ and a choice of Θ , $\mathcal{C}_\Theta(\sigma) = \Theta(\sigma)$. Furthermore we write $\mathcal{C}_\Theta(U) = \{\Theta(\sigma) | \sigma \in U\}$, so that given $\mathcal{M}$, $\mathcal{C}_\Theta(\mathcal{M})$ is the contextuality measurement scenario with

1. $X' = \Theta(\Sigma)$
2. $\Sigma' = \{\Theta(C) | C \in \mathcal{C}\}$
3. $O'_{\Theta(\sigma)} = O_{\sigma[-1]}$.

We can then similarly show that there is an isomorphism of sets $\mathcal{C}_\Theta^* \mathcal{E}(U) \cong^\alpha \mathcal{S}_\Theta(U)$, which allows us to pull back an empirical model $w \in \mathbf{EM}(\mathcal{C}_\Theta(\mathcal{M}))$ to define one $e^\Theta \in \mathbf{EM}^{(\Theta)}(\mathcal{M})$ by defining

$$e_C^{(\Theta)}(f) = w_{\mathcal{C}_\Theta(C)}(\alpha(f)), \tag{5.50}$$

where $f \in \mathcal{S}_\Theta(C)$ is a Θ -consistent strategy on C .

Now define $\mathcal{C}_{L_k}(\mathcal{M})$ for any $\mathcal{M}$ to be the coproduct

$$\sum_{\Theta \in L_k} \mathcal{C}_{\Theta}(\mathcal{M}), \quad (5.51)$$

and likewise we can define an empirical model $e^{(L_k)}$ to be the pullback of the empirical model

$$w_{\Theta_1} \& w_{\Theta_2} \& \dots \& w_{\Theta_{|L_k|}} \quad (5.52)$$

where $w_{\Theta_i} \in \text{EM}(\mathcal{C}_{\Theta_i}(\mathcal{M}))$.

Example 5.7.3 (3 party temporal Bell). *Take the three party Bell scenario, in which $\max \Sigma = \{(a, 0)(b, 0)(c, 0), (a, 0)(b, 0)(c, 1), (a, 0)(b, 1)(c, 0), (a, 1)(b, 0)(c, 1)\}$, so that agent B only has one choice of measurement.*

Consider the set L_1 of Θ -functions. This set contains three functions: $\Theta_0 := lk_0$, $\Theta_1 := lk_1$ and Θ_2 which we define below.

$$\Theta_2((a, i)) = (a, i)$$

$$\Theta_2((a, i)(b, j)) = (a, i)(b, j)$$

$$\Theta_2((a, i)(b, j)(c, k)) = (a, i)(c, k).$$

Let $\langle X'_i, \Sigma'_i, O'_i \rangle$ be the image scenario $\mathcal{C}_{\Theta_i}(\mathcal{M})$. Then $\mathcal{C}_{L_1}(\mathcal{M})$ is the scenario

$$\langle X'_1 \sqcup X'_2 \sqcup X'_3, \Sigma'_1 + \Sigma'_2 + \Sigma'_3, [O'_1, O'_2, O'_3] \rangle \quad (5.53)$$

. Valid L_1 -consistent empirical models on the temporal scenario $\mathcal{M}$ are pullbacks of empirical models $w_1 \& w_2 \& w_3$ where $w_i \in \text{EM}(\mathcal{C}_{\Theta_i}(\mathcal{M}))$.

We recall that the controlled choice of empirical models is a free operation in the resource theory of contextuality, covered in Section [3.6](#). The contextual fraction of the controlled choice empirical model is related to the contextual fractions of the individual empirical models as

$$\text{CF}(w) = \max_{\Theta_i \in L_k} \text{CF}(w_i). \quad (5.54)$$

Thus the empirical model w is contextual provided at least one of w_i is. Conversely, the empirical model w can be noncontextual only if all of w_i are. A similar ob-

servation is contained in the following proposition, which gives the corresponding Vorob'ev type results.

Proposition 5.7.4. *Fix a temporal measurement scenario $\mathcal{M} = \langle X, \Sigma, O \rangle$. Every empirical model $e \in \mathbf{EM}^{(L_k)}(\mathcal{M})$ is L_k -classical if and only if for every $\Theta \in L_k$, the image complex $\mathcal{C}_\Theta(\Sigma)$ is acyclic.*

Proof. This follows straightforwardly from the observation that coproduct of simplicial complexes $\sum_i \Sigma_i$ is acyclic if and only if every Σ_i is acyclic. $\square$

5.8 Finite state machines

We have defined a machine to be a tuple $\mathbf{E} = \langle X, O, F, h \rangle$. Oftentimes, rather than define a subset F of realisable input-output functions, one gives the machine a finite number of states and refers to it as a **finite state machine** [124]. A finite state machine (FSM), given its internal state and a choice of input produces an output. This process, of deterministically producing outputs given the machines current state and the choice of input, is described jointly via a transition function and output function.

Definition 5.8.1 (Finite State Machine [124]). *A finite state machine $\mathbf{E}$ is a 6-tuple $\langle S, S_0, X, O, T, Q \rangle$ consisting of*

- *A finite set of states S*
- *An initial state $S_0 \in S$*
- *A finite input alphabet X*
- *A finite output alphabet O*
- *A transition function $T : S \times X \rightarrow S$*
- *An output function $Q : S \times X \rightarrow O$*

A strategy will be defined in terms of the output function Q of the FSM via

$$f(x_0 x_1 \dots x_n) := Q(s_0, x_0) Q(s_1, x_1) \dots Q(s_n, x_n). \quad (5.55)$$

We can then generate the set F of realisable input-output functions over some choice of $\Sigma \subseteq X^*$.

We first show that Q and T can be defined such that the strategies realised by $\mathbf{E}$ are $\mathbf{lk}_k$ -consistent.

Proposition 5.8.2. *For appropriately defined functions Q and T and for a sufficient number of internal states S , the FSM E generates lk_k -consistent input-output functions.*

Proof. Take an arbitrary $U \subseteq \Sigma$ and consider the set of lk_k -consistent strategies $\mathcal{S}_{lk_k}(U)$. Recall that such strategies are in bijection with the set of sections $\mathcal{E}(\mathcal{C}_{lk_k}(U))$. Take the internal states of the system to be $S = \{1, 2, \dots, d\}$. We require that, for a fixed measurement choice $x \in X$, the set of internal states be large enough to allow the k -lookback for all sequences ending in x to be stored. To this end define

$$L(x) := \{lk_k(\sigma) \mid \sigma \in \Sigma, \sigma[-1] = x\} \quad (5.56)$$

and let

$$d := \max_{x \in X} |L(x)|. \quad (5.57)$$

For non-local setups *without* repeating measurements we require $d = |I|^k$, where I is the set of inputs of an agent.

Now define a bijection for each $x \in X$ as $\gamma_m : L(x) \rightarrow S$ which assigns to each lookback for x a state in S . This lets us in turn define the transition and output functions T and Q in terms of a section $f \in \mathcal{C}_{lk_k}^* \mathcal{E}(U)$ as

$$Q(x_i, s_i) = f(\gamma_{x_i}^{-1}(s_i)) \quad (5.58)$$

$$T(m_i, s_i) = \gamma_{x_{i+1}}(\gamma_{x_i}^{-1}(s_i)x_{i+1}). \quad (5.59)$$

where in the definition of T , $\gamma_{x_i}^{-1}(s_i)x_{i+1}$ is the concatenation of the sequences $\gamma_{x_i}^{-1}(s_i)$ and x_{i+1} . The output function uses the lookback for measurement x_i corresponding to the state s_i to define the output in terms of the lk_k -consistent section f . The transition function simply updates the lookback to include the next measurement to be performed x_{i+1} and maps this back to the corresponding internal state. $\square$

Since an E_{lk_k} consistent empirical models can be realised by an FSM of sufficient internal dimension, which uses its internal states to store the lookback of measurements, we can automatically extend Proposition 5.6.2 to derive an upper bound on the number of states needed to realise any empirical model for a given setup.

Proposition 5.8.3. *Let $\mathcal{M} = \langle X, \Sigma, O \rangle$ be a temporal measurement scenario. Every temporal Bell scenario $e \in \text{EM}(\mathcal{M})$ can be produced by an FSM with internal dimension $d \geq \max_{x \in X} |L_k(x)|$ where $k = N - 1$ and $N = \max_{\sigma \in \Sigma} |\sigma|$. Here the set*

$L_k(x)$ for each measurement $x \in X$ is defined as the total number of inequivalent k -lookbacks for a measurement $x \in X$,

$$L_k(x) := \{lk_k(\sigma) | \sigma \in \Sigma \wedge \sigma[-1] = x\}. \quad (5.60)$$

Proof. Take the FSM to have output and transition functions as in equations [5.58](#)–[5.59](#). Since $k = N - 1$, from Proposition [5.6.2](#) it is classical. $\square$

We now turn to a question we have left lingering, which is whether one can say anything interesting in the current framework about the case in which strategies can be dependent on past measurements *and* outcomes.

5.9 Allowing outputs to be stored in memory

Most generally, a system with bounded memory can choose which information from the past to store. Without access to information outside of the system, such as a clock, the information which the system can store is (i) a measurement (or equivalently a choice of measurement setting) (ii) an outcome (iii) A function of these. Naturally, then, we wonder how allowing the system to store outputs in addition to, or instead of, measurements changes the set of producible correlations.

Our first observation is that allowing for the system to store previous outputs, rather than measurements, still allows for the system to keep track of previous measurements by encoding these in the outputs. The result, however, is that the strategies are constrained to use outputs solely as a type of memory. Nevertheless, in the extreme case, where measurements have countably infinite output sets, for every empirical model $e \in \mathbf{EM}(\mathcal{M})$ there exists an empirical model $e' \in \mathbf{EM}(\mathcal{M})$ whose output can be projected onto the first $m \in \mathbb{N}$ values to give an empirical model identical to e , but which is classical. The idea is that the second empirical model uses the output set to store, in addition to the output of the first empirical model, all measurements which have previously occurred. Much like Hilbert’s Hotel, in which a fully occupied hotel with infinitely many rooms may still accommodate additional guests, our output sets, which we take to be infinitely big, can always store additional information. As such, we can store the full lookbacks in the output register. This observation makes it clear that the size of the output set is important in determining to what extent including the outputs is important.

The technique we have so far been using to study temporal correlations—namely, mapping them to contextual ones—no longer works when including outputs. The

problem arises because labelling a measurement with the memory at the point at which the measurement is performed (which so far has been limited to strings of measurements, such as the k -lookback) can result, when outcomes are included in this memory, in the no-signalling relations no longer being satisfied.

Nevertheless, we can still make some concrete statements about this case.

5.9.1 Parity machines and related examples

A standard example for classical sequential machines, such as finite state machines, is the parity checker. The parity checker for a string of inputs $x_0...x_l$ will output $o_l = 0$ if $x_0...x_l$ is of even parity (i.e. it contains an even number of 1's), and $o_l = 1$ if it is of odd parity (or contains an odd number of 1's).

A system which is able to store the previous outcome can implement the parity checker deterministically. The table below defines given o_{l-1} and x_l what the output o_l will be, by simply performing an XOR operation on the two bits. This clearly

o_{l-1}	x_l	$o_{l-1} \oplus x_l$
0	0	0
0	1	1
1	0	1
1	1	0

Table 5.3: The parity checker.

implements the parity checker. Now we ask what the minimal, in the sense of the number of inputs Θ contains in its codomain, Θ -function is that can produce the parity checker. If the entire history of inputs is stored the parity can be checked by computing $x_0 \oplus x_1 \oplus ... \oplus x_n$. Note however that for any bitstring, we can find a substring of length at most 2 with parity identical to the original. In fact, any bitstring which has odd parity always contains 1 as a substring, so we define $\Theta(x_0x_1...x_l) = 1x_l$ if $x_0x_1...x_l$ has odd parity. On the other hand, a bit string can contain only 1's and have even parity. Thus we define $\Theta(x_0x_1...x_l) = 0x_l$ if $x_0x_1...x_l$ has even parity and contains at least one 0, or if the bitstring contains no 0's then $\Theta(x_0x_1...x_l) = 11x_l$. Thus $\Theta \in L_3$, so that while only one outcome and one measurement is needed, if only measurements are to be stored then at least three are needed.

This is not such a bad trade-off. Consider now a system with a larger output set $|O| = z$, and suppose we have the empirical model in which the output o_l is the number of 1's mod z in the bitstring $x_0...x_l$, and so implements the parity checker

when $z = 2$. In this case it can similarly be shown that the minimal Θ -function is in the set L_{z+1} . Therefore, the number of extra measurements which need to be stored scales as the size of the output set.

5.9.2 Nonclassicality and memory with outputs

We denote as $\text{lookback}_k^{(e)} : \mathcal{S}(\Sigma) \times \Sigma \rightarrow \Delta_{\mathcal{M}}^*$ the map which sends a tuple (s, σ) consisting of a strategy s and a sequence σ to the $k + 1$ measurement events $\langle m_{\max(n-k,0)}, o_{\max(n-k,0)} \rangle \dots \langle m_{n-i}, o_{n-i} \rangle \langle m_n, \perp \rangle$ where $o_i := s(m_0 \dots m_{i-1} m_i)$. Note that this is strategy dependent, unlike lk_k .

Nevertheless, it is straightforward to extend previous definitions of lk_k -consistent strategies and lk_k -classical empirical models to $\text{lk}_k^{(e)}$. In particular, $s \in \mathcal{S}(\Sigma)$ is $\text{lk}_k^{(e)}$ -consistent if for all $\sigma, \sigma' \in \Sigma$,

$$\text{lk}_k^{(e)}(s, \sigma) = \text{lk}_k^{(e)}(s, \sigma') \implies s(\sigma) = s(\sigma'). \quad (5.61)$$

In this section we show that a slightly different version of one direction of Theorem 5.6.1 extends also to empirical models which are $\text{lk}_k^{(e)}$ -consistent.

Proposition 5.9.1. *Let $\mathcal{M}$ be a temporal measurement scenario. If the measurement scenario $\mathcal{C}_{\text{lk}_k}(\mathcal{M})$ has simplicial complex Σ'_k which is not acyclic, then there is at least one empirical model on $\mathcal{M}$ which cannot be generated by a classical machine $E_{\text{lk}_k^{(e)}}$.*

The above is interesting because it says that we can always not only find an empirical model which is E_{lk_k} -nonclassical, but also one which is $E_{\text{lk}_k^{(e)}}$ -nonclassical. Since a machine which can store the k previous measurements and their outcomes is significantly more powerful (from a memory resource point of view) this is in some sense surprising.

The theorem essentially follows because we can define an empirical model d on $\mathcal{C}_{\text{lk}_k}(\mathcal{M})$ which simulates an empirical model on the restricted scenario $\mathcal{C}_{\text{lk}_k}(\mathcal{M})|_Z$, where Z corresponds to the vertices of any square subcomplex of Σ'_k —Proposition 5.6.3 guarantees that such a restriction always exists. Using the fact that one can always define a contextual empirical model on this restricted scenario, we show that the empirical model d which simulates e (i) must also be contextual and (ii) can be defined in such a way that it pulls back to give a $\text{lk}_k^{(e)}$ -nonclassical model on $\mathcal{M}$. We will show this step-by-step.

Given a measurement scenario $\mathcal{M} = \langle X, \Sigma, O \rangle$, where Σ has vertices in the set X , we can restrict it to a scenario $\langle X', \Sigma|_{X'}, O \rangle$ where $X' \subseteq X$. We write $\mathcal{M}|_{X'}$ for the restricted scenario.

Note in terms of resource theoretic definitions that we have a deterministic morphism $\mathcal{M} \rightarrow \mathcal{M}|_{X'}$ since we always have a simplicial map $\pi : \Sigma|_{X'} \rightarrow \Sigma$ and for each $x' \in X'$ we have a map $\text{id} : O_{\pi(x)} = O_x$ which is simply the identity map.

Given an empirical model $e \in \text{EM}(\mathcal{M})$, we define the restricted model $(e|_{X'})_\sigma := e_\sigma$ since $\sigma \in \Sigma$ for any $\sigma \in \Sigma|_{X'}$. One can always simulate $e|_{X'}$ with e by pushing e forward along $\langle \pi, \text{id} \rangle$:

$$e|_{X'} = \langle \pi, \text{id} \rangle_* e. \quad (5.62)$$

The following Theorem has been shown for more general simulations in [3] and [81]. Here, we only give the case for empirical models which are restrictions, since it is the case which we are concerned with.

Lemma 5.9.2 ([81], [3]). *Given a contextuality measurement scenario $\mathcal{M} = \langle X, \Sigma, O \rangle$ consider an empirical model $d \in \text{EM}(\mathcal{M}|_{X'})$ on the restricted scenario $\langle X', \Sigma|_{X'}, O \rangle$, where $X' \subseteq X$. If d is the restriction of an empirical model $e \in \text{EM}(\mathcal{M})$, then contextuality of $d = e|_{X'}$ implies contextuality of e .*

Now we are in a position to prove proposition [5.9.1].

Proof. Suppose that $\mathcal{C}_{\text{lk}_k}(\mathcal{M})$ is not acyclic. By Proposition [5.6.3] this implies that there are at least two agents q and p , $q \leq p$ with either (i) $j - i > k$ or (ii) there exists an $i < j < l$ with $l - i > k$ and $l - j \leq k$. Consider case (i), as case (ii) follows analogously. If (i) holds then we restrict $\mathcal{C}_{\text{lk}_k}(\mathcal{M})$ to the vertices

$$\begin{aligned} &\{\text{lk}_k((1, i_1)(2, i_2) \dots (p, i_p)), \text{lk}_k((1, i_1)(2, i_2) \dots (p, i'_p)), \\ &\text{lk}_k((1, i_1)(2, i_2) \dots (q, i_q)), \text{lk}_k((1, i_1)(2, i_2) \dots (q, i'_q))\} \end{aligned} \quad (5.63)$$

W.l.o.g. we take $i_p = i_q = 0$ and $i'_p = i'_q = 1$. The restriction of Σ'_k to this set of vertices, which we call V , will contain four faces of the form

$$\{\text{lk}_k((1, i_1)(2, i_2) \dots (p, l_p)), \text{lk}_k((1, i_1)(2, i_2) \dots (q, l_q))\} \quad (5.64)$$

for $l_p, l_q \in \{0, 1\}$. In other words, the restricted complex is a square. We define an empirical model e on the restricted scenario to be a PR box [4]. We now wish to define an empirical model d on $\mathcal{C}_{\text{lk}_k}(\mathcal{M})$ which simulates e and which pullsback

⁴Note that this could equally be something quantum realisable, such as the CHSH empirical model

to give a $\text{lk}_k^{(e)}$ -nonclassical empirical model on $\mathcal{M}$. Note that since $p - q > k$ we have that $\text{lk}_k((1, i_1)(2, i_2) \dots (p, l_p))$ and $\text{lk}_k((1, i_1)(2, i_2) \dots (q, l_q))$ do not overlap on any agents. Define the empirical model d such that in any context C the nodes not in V are deterministically sent to 0. We would like to show that the pullback $C_k^*(d)$ is $\text{E}_{\text{lk}_k^{(e)}}$ -nonclassical. Since it is guaranteed to be E_{lk_k} -nonclassical (from Theorem 5.6.1), the set of lk_k -consistent global sections consistent with the support of $C_k^*(d)$ is empty (as we chose a PR Box). Note also that for the example we have constructed a section which has support on $C_k^*(d)$ is lk_k -consistent iff it is $\text{lk}_k^{(e)}$ -consistent, which implies that the set of global $\text{lk}_k^{(e)}$ -consistent strategies is also empty, and thus $C_k^*(d)$ is $\text{E}_{\text{lk}_k^{(e)}}$ -nonclassical. $\square$

5.10 Discussion

We have considered the sequential analogue of a contextuality setup, in which no-signalling constraints are relaxed to the AoT, or Arrow of Time, constraints. By utilising a sheaf theoretic approach, we modelled a setup in which the amount of information a system can store from the past is bounded. Starting with storage of just the previous k measurements, we showed that, given a particular setup captured by a temporal measurement scenario $\mathcal{M}$, there exist nonclassical correlations (resulting from affine combinations of lk_k -consistent strategies) only if the simplicial complex Σ'_k is not acyclic. This simplicial complex is the one associated with a contextuality measurement scenario $\mathcal{C}_k(\mathcal{M})$ which is the image of $\mathcal{M}$ under the map $\mathcal{C}_k$. Moreover, we showed that for temporal Bell scenarios, there are certain conditions on the order of agents and their input sets which, when met, guarantee that nonclassicality can arise for some lk_k -consistent model on $\mathcal{M}$.

We next showed how this results generalises when replacing the lk_k functions with Θ -functions, which allow for storage or arbitrary measurements from the past. We considered the set of strategies which are L_k -consistent meaning that they are Θ -consistent for some Θ -function which stores at most k measurements. Again, nonclassicality of L_k -consistent empirical models on a temporal scenario $\mathcal{M}$ relies on an image complex $\mathcal{C}_{L_k}$ being acyclic.

Finally, we discussed the case in which measurements and outputs are stored, concluding with proposition 5.9.1 which gives a Vorob'ev-type result (albeit only one direction) for the case in which the k immediate measurements and outcomes are stored by the system.

There are a few possible directions for future work. Extending the results developed in this chapter to causal measurement scenarios, as defined in [5] and where we have both temporal and spatial correlation, would be informative. Doing so, even just for the case $\mathcal{C}_\infty$ and for Bell scenarios, would shed light on which arrangement of agents can give rise to nonclassical empirical models. Are the linearly ordered agents considered here, for example, the only case in which one can never achieve non-classical AoT correlations? While we have mostly considered examples in which the measurement strings in σ are finite, oftentimes in sequential computation one allows the system to run indefinitely, thereby generating strings of outcomes of arbitrary length. It might therefore be informative to specialise to such cases in order to understand separations between quantum and classical correlations in these setups. Finally, it would be useful to study further the classical and no-signalling polytopes which arise from lk_k -consistent strategies; while it is clear that the former arises as the convex hull of the deterministic empirical models, it is intuitive, but not proven, that the no-signalling polytope on the other hand should arise as affine combinations of these models.

Chapter 6

Causality and contextuality: a game semantics approach

— Summary —

We generalise from temporal setups to causal setups, albeit now with no restriction on memory. Our key insight is that there are two types of causality which arise in the literature—that which only the Experimenter (and not the system being measured) has access to, and that which both the Experimenter and the system, here called Nature, have access to. The notion of classicality is sensitive to which definition is used. To this end, we introduce Nature and Experimenter strategies, and associated presheaves, which must respect the imposed causality. Crucially, Experimenter strategies can be adaptive. We show how this framework reduces to those studied in previous sections—the usual contextuality scenarios [7], the causal Bell scenarios of Gogioso and Pinzani [66], temporal setups of chapter [5], and certain scenarios from the Bayesian network literature (what we term HLP scenarios) [73]. Part of the material contained in this chapter is from [5], although there a slightly different framework is used.

There has been an increasing trend to try to incorporate causal constraints into Bell-type, and more generally contextuality, setups, thereby launching the programme of *causal contextuality*. While the study of temporal scenarios and the

temporal correlations that they host gives some insight into nonclassical correlations beyond contextual correlations, it is not of course the most general treatment of causal correlation, which would encompass both spatial and temporal elements.

Developing a thorough understanding of causal contextuality is a difficult task, not least of all because incorporating causal constraints can be done in more than one way, and these are nonequivalent. There are (at least) two prevailing viewpoints in the literature: on the one hand, causal dependencies allow a classical control computer to feed-forward measurement outcomes to determine future measurement choices, which is the type of causality inherent in MBQC but also in the causal network models considered by Henson, Lal and Pusey [73] and, independently, Fritz [59]. On the other hand, as studied by Gogioso and Pinzani [66], causal constraints allow a NCHV to assign outcomes to a measurement based on measurement events—measurements together with their outcomes—which have occurred in the past.

We view the setup as a type of concurrent game played between Nature and an Experimenter [1], with a unique set of measurement events (i.e. a history) arising for each choice of Nature and Experimenter strategy. Experimenter strategies choose which measurements should be performed, allowing for feedforward to be enforced by an experimenter, and Nature strategies respond to Experimenter strategies by supplying an output for the performed measurement. We show that while Gogioso and Pinzani’s setups arise naturally when introducing causal structure which both Nature and the Experimenter can use in forming a strategy, Henson Lal and Pusey’s setups (or a class of such setups) on the other hand allow the Experimenter to use adaptivity which Nature should be agnostic to.

Here, we introduce a framework which is able to differentiate between the above types of causality, by explicitly separating that due to the environment, which Nature can use, and that due to the Experimenter. Both have an effect on the observed contextuality of the system; for example, while the Bell-CHSH scenario in Figure 6.1(a) is known to exhibit nonclassicality—as for example demonstrated by violation of the CHSH inequalities—neither the Instrumental setup in the case $|X| = 2$ in Figure 6.1(b) nor the causal Bell-CHSH setup in Figure 6.1(c) are able to exhibit nonclassicality, but for different reasons.

6.1 Causal contextuality

It will be useful to differentiate between variables which the Experimenter has control over (measurement settings) and those which Nature has control over (measurement outcomes). Here by Nature we mean the system at hand, and at present we

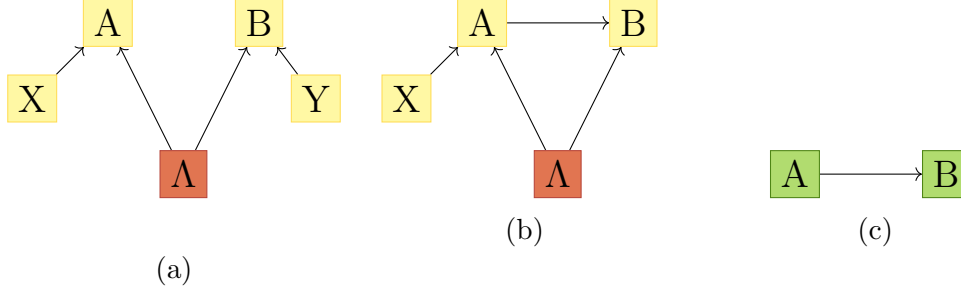


Figure 6.1: Types of two party setup. (a) The two party Bell setup, (b) The Instrumental setup and (c) Two party causal Bell scenario with $A \leq B$.

make no assumptions about how Nature behaves and only require that, when the experimenter poses a question or set of questions, Nature supplies an answer to each question. In practice, questions are asked by an experimenter performing a measurement, and answers given by Nature proving an outcome. Cells will correspond to points of a computation where questions are posed or answers are given, and to this end we introduce a set Γ_E of Experimenter cells and a set Γ_N of Nature cells.

To each $c \in \Gamma_E$ we associate a set of values V_Q^c which is a set of questions, and to each $c \in \Gamma_N$ a set of values V_A^c which is a set of answers. Thus, a single ‘run’ corresponds to filling certain cells in Γ with a value from their set of possible values. A tuple $\langle c, v \rangle$ where c is a cell and v is a value in that cell is called an event; when $c \in \Gamma_E$ we call this an Experimenter event, while when $c \in \Gamma_N$ we call this a Nature event. Nature events are usually referred to in the literature as *measurement events*, since they correspond to a measurement being assigned some outcome.

Now in order to model a causal unfolding of events, we write $s \vdash c$ where s is a set of events and $c \in \Gamma$, and $\vdash$ is called the enabling relation. The implication of this is that c can only be filled once certain events have occurred, and it indicates a type of dependency relation. Cells for which $\emptyset \vdash c$ can be filled automatically, and are called initially enabled. For flat scenarios, every experimenter cell is initially enabled, and any question $q \in V_Q^i$ at a cell $i \in \Gamma^E$ will enable a corresponding Nature cell $n(q) \in \Gamma^N$ which gives an answer to this question.

Altogether, we define a **causal measurement game** as follows.

Definition 6.1.1 (Causal Measurement Game). *A causal measurement game is a five-tuple $M = \langle \Gamma, Q, A, \{V_c\}_{c \in \Gamma}, \vdash \rangle$ where*

1. Γ is a countable set whose elements are called cells. We partition $\Gamma = \Gamma^E \sqcup \Gamma^N$ where Γ^E are Experimenter cells and Γ^N are Nature cells.
2. A set Q of questions.

3. A set A of answers.

4. By cases,

- If $c \in \Gamma^E$ a set $V_c \subseteq Q$, of questions
- If $c \in \Gamma^N$ a set $V_c \subseteq A$ of answers.

5. An enabling relation $\vdash \subseteq \mathcal{P}(\Delta_M) \times \Gamma$, where $\Delta_M := \{(c, v) | c \in \text{Gamma}, v \in V_c\}$ is the set of all possible events (both Nature and Experimenter).

Given a set of events $s \subseteq \Delta_M$, we define

$$\text{dom}(s) := \{c | \langle c, v \rangle \in s\} \quad (6.1)$$

so that $\text{dom}(s)$ is the set of cells associated with the events in s . Further, we say that the set of events s is **consistent** if for any events $\langle c, v \rangle, \langle c', v' \rangle \in s$ it holds that $c = c' \implies v = v'$.

In the following Chapter we will write $\langle \Gamma, V, \vdash \rangle$ under the assumption that Experimenter cells are subsets of a ‘question’ set, such as the set of observables on a Hilbert space, while Nature cells are subsets of an ‘answer set’, which is usually a finite set $[k] = \{1, 2, 3, \dots, k\}$. Our definition of a causal measurement game is almost identical to that of Kahn and Plotkin’s information matrices, which were introduced to model a causal unfolding of events [80]. While they consider cells of a single type, here cells are divided into two types. Likewise, the values that these cells take on tend to be thought of in different ways : while the filling of Experimenter cells corresponds to a question being asked, a Nature cell corresponds to the ‘answer’ to a question, or outcome of measurement. Otherwise, the interpretation is identical: cells which are enabled are filled with values, which for Kahn-Plotkin is called a decision rather than an event. Given an information matrix, Kahn and Plotkin introduce *configurations*, which are sets of decisions in which every decision is enabled. We introduce a corresponding notion of **histories of a measurement scenario**, which generalises the notion of history which we had in the temporal case.

Definition 6.1.2 (Set of Histories). *Given a measurement scenario $\mathcal{M}$, we define a history to be a subset of events $h \subseteq \Delta_M$ for which it holds that (i) it is consistent and (ii) every cell $c \in \text{dom}(h)$ is enabled, or, formally, $c \in \text{dom}(h)$ implies $s \vdash c$ for some subset $s \subseteq h$. We denote the set of all such histories $\mathcal{H}(\mathcal{M})$.*

A history is a set of events which can be observed in any run of the experiment. A **strategy** will be a subset of histories which denotes a recipe for playing the game,

and will be Nature or Experimenter specific. To introduce strategies, we need the accessibility relation $\triangleright$. While the enabling relation specifies which events are *needed* in order for a cell to be filled, the accessibility relation instead specifies which sets of events the cell *can* be filled after. We therefore write $s \triangleright c$ where s is a consistent set of events and $c \in \Gamma$ a cell if and only if

- $c \notin \text{dom}(s)$,
- $\exists t \subseteq s$ with $t \vdash x$.

We further write $s \triangleright^E c$ or $s \triangleright^N c$ when $c \in \Gamma^E$ and $c \in \Gamma^N$ respectively. We now define Nature strategies.

Definition 6.1.3 (Nature Strategy). *A nature strategy is a subset $\sigma \subseteq \mathcal{H}(\mathcal{M})$ such that*

- *Given $h, t \in \mathcal{H}(\mathcal{M})$, if $h \in \sigma$ and $t \subseteq h$ then $t \in \sigma$. We say that σ is downward closed.*
- *If $s \in \sigma$ and $s \triangleright^N c$ then there exists a unique $v \in V_c^A$ such that $s \cup \{\langle c, v \rangle\} \in \sigma$.*
- *If $s \in \sigma$ and $s \triangleright^E c$ then for all $v \in V_c^Q$ it holds that $s \cup \{\langle c, v \rangle\} \in \sigma$.*

A strategy for Nature therefore provides a unique answer to all Nature cells which are accessible, but allows for accessible experimenter cells to be filled with all possible sets of questions which the experimenter might ask. The strategy is therefore deterministic at Nature cells but nondeterministic at Experimenter cells. On the other hand, we define Experimenter strategies.

Definition 6.1.4 (Experimenter Strategy). *An experimenter strategy is a subset $\sigma \subseteq \mathcal{H}(M)$ such that*

- *Given $h, t \in \mathcal{H}(\mathcal{M})$, if $h \in \sigma$ and $t \subseteq \sigma$ then $t \in \sigma$.*
- *If $s \in \sigma$ and $s \triangleright^E c$ then there exists a unique $v \in V_c^Q$ such that $s \cup \{\langle c, v \rangle\} \in \sigma$.*
- *If $s \in \sigma$ and $s \triangleright^N c$ then for all $v \in V_c^A$ it holds that $s \cup \{\langle c, v \rangle\} \in \sigma$.*

Note that Nature strategies and Experimenter strategies are dual. Let $\mathcal{S}^N(M)$ and $\mathcal{S}^E(M)$ be the set of Nature and Experimenter strategies respectively associated with M .

6.2 The strategy presheaves

We define a **substructure relation** which is an order on measurement games, and will allow us to define strategy presheaves which generalise the event and strategy presheaves of the contextuality and temporal cases respectively.

Given two causal measurement games $G = \langle \Gamma_G, Q_G, A_G, \{V_{G,i}\}_{i \in \Gamma_G}, \vdash_G \rangle$ and $H = \langle \Gamma_H, Q_H, A_H, \{V_{H,i}\}_{i \in \Gamma_H}, \vdash_H \rangle$ write $G \leq H$ if

- i $\Gamma_G^E \subseteq \Gamma_H^E$ and $\Gamma_G^N \subseteq \Gamma_H^N$
- ii $Q_G \subseteq Q_H$
- iii $A_G \subseteq A_H$
- iv For all $c \in \Gamma_G$, $V_{G,c} \subseteq V_{H,c}$
- v For all $s \subseteq \Delta_G$, $c \in \Gamma_G$, $s \vdash_G c$ iff $s \vdash_H c$.

Items i – iv are intuitive and simply require that G have a subset of the cells that H has and moreover that the values which these cells can be filled with in G is a subset of the values that same cell can be filled with in H . Thus, any event, Experimenter or Nature, which occurs in G can occur in H . Item v says moreover that a cell is enabled in G in the same way it is enabled in H , provided the set of events which enable it can occur in G .

Given a nature strategy $n \in \mathcal{S}^N(H)$ we can restrict this to $n|_G := n \cap \mathcal{H}(G)$. The same holds for Experimenter strategies; a strategy $e \in \mathcal{S}^E(H)$ can be restricted $e|_G := e \cap \mathcal{H}(G)$. We show that this is a well-defined restriction map in the sense that these restrictions are valid Nature and Experimenter strategies on the restricted scenario.

Proposition 6.2.1. *Given two measurement scenarios G and H for which $G \leq H$ and a Nature strategy $n \in \mathcal{S}^N(H)$, then the restriction $n|_G$ is a Nature strategy of the scenario G , or $n|_G \in \mathcal{S}^N(G)$.*

Proof. First of all, since n is down closed it holds that $n \cap \mathcal{H}(G)$ will be, since $\mathcal{H}(G)$ is a down closed subset of $\mathcal{H}(H)$ and down-closure is preserved under intersection. Next, suppose $s \in n$ and $s \triangleright_G^N c$, then we have $s \triangleright_H^N c$ and so since n is a strategy in H it holds that there exists a unique $v \in V_c^A$ such that $s \cup \langle c, v \rangle \in n$. Moreover, if $s \in n$ and $s \triangleright_G^E c$ then it automatically holds that $s \triangleright_H^E c$ and therefore, again, since n is a strategy for H that for all $v \in V_c^Q$, $s \cup \{\langle c, v \rangle\} \in n$. Therefore, n is a valid strategy for G , too. $\square$

Proposition 6.2.2. *Given two measurement scenarios G and H for which $G \trianglelefteq H$ and an Experimenter strategy $e \in \mathcal{S}^E(H)$, then the restriction $e|_G$ is a Experimenter strategy of the scenario G , or $e|_G \in \mathcal{S}^E(G)$.*

The proof for experimenter strategies follows essentially in exactly the same way as that of Nature strategies in Proposition [6.2.1](#).

The fact that $n|_G \in \mathcal{S}^N(G)$ means that the restriction map which we have defined turns $\mathcal{S}^N$ into a presheaf—specifically, it is a presheaf from the category of substructures over M to the category **Set**. The category of substructures of M will be denoted **Sub**(M) where

1. **Objects** of **Sub**(M) are substructures of M
2. **Morphisms** of **Sub**(M) are arrows $G \rightarrow H$ for substructures for which $G \trianglelefteq H$.

As such we have the presheaves $\mathcal{S}^N : \mathbf{Sub}(M)^{\text{op}} \rightarrow \mathbf{Set}$ and $\mathcal{S}^E : \mathbf{Sub}(M)^{\text{op}} \rightarrow \mathbf{Set}$, which are presheaves over a poset with $(\mathbf{Sub}(M), \trianglelefteq)$ being the poset.

6.3 Covers and empirical models

In quantum mechanics, the commutation relation restricts which measurements can be performed simultaneously. On the other hand, for temporal setups, we saw that only certain sequences $\sigma \in X^*$ are observed, a restriction which is not fundamental as with the commutation relation, but which nevertheless allowed us to show that there is an obstruction in moving from the local observed distributions to a global distribution, and that this obstruction witnesses a kind of nonclassicality in the system.

In the current more general setting, in any run of the experiment we observe the statistics (measurement outcomes) associated with a substructure $G \trianglelefteq M$ of the measurement game M . Such substructures are collected into a cover $\mathcal{C} = \{G_i\}_{i \in I}$ of M . We recall that when defining a presheaf over a poset, the join of the elements in the cover should give the top element, $\bigvee_{i \in I} G_i = M$. We will require moreover that $\mathcal{C}$ be down closed with respect to the order on **Sub**(M), so that $\downarrow \mathcal{C} = \mathcal{C}$. The pair $(M, \mathcal{C})$ consisting of a measurement game together with a cover is called a **causal measurement scenario**. The associated empirical models are compatible families with respect to this cover for the Nature strategy presheaf $\mathcal{D}_R \circ \mathcal{S}^N$.

Definition 6.3.1 (Causal empirical model). *Given a causal scenario $\mathcal{M} = (M, \mathcal{C})$ a causal empirical model is a family of R -distributions $\{e_G\}_{G \in \mathcal{C}}$ with $e_G \in \mathcal{D}_R \circ \mathcal{S}^N(G)$, such that for any $G, H \in \mathcal{C}$ with $G \leq H$,*

$$e_H|_G = e_G. \quad (6.2)$$

We write $\text{EM}(\mathcal{M})$ for the set of empirical models on the causal scenario $\mathcal{M}$.

Nonclassicality is, as with contextuality and temporal scenarios, present when we cannot extend the local sections of the empirical model to a single unique global section. In this case we refer to nonclassicality as **causal contextuality**.

Definition 6.3.2 (Causal Contextuality). *An empirical model e is said to be causally contextual if there exists no global section $h \in \mathcal{D}_R \mathcal{S}^N(M)$ such that $h|_G = e_G$ for all $G \in \mathcal{C}$.*

While here the definition of empirical model and notion of contextuality make reference only to Nature strategies, later we will see how Experimenter strategies influence which Nature strategies are observed and thereby whether contextuality is observed—sometimes, an Experimenter strategy can reduce the observed contexts to an extent such that no local to global obstruction is observed. For now, we show how, even without Experimenter strategies, we can capture contextuality scenarios, temporal scenarios, and causal Bell scenarios.

6.4 The sheaf axiom and $\mathcal{S}^N$

Note that while in order for $\mathcal{S}^N$ to be considered a sheaf, we require that it satisfy the sheaf axiom—which required there be a unique gluing of a compatible family—in our case we are only interested in whether this property is satisfied for a single *specific* cover, which is physically motivated by the problem at hand. Since failure of the composition $\mathcal{D}_R \circ \mathcal{S}^N$ to satisfy the sheaf axiom for a choice of compatible family e_i is used in the definition of contextuality of an empirical model, it is imperative that for a deterministic classical empirical model defined instead over $\mathcal{S}^N$ there be no possibility of contextuality arising. In other words, if $\mathcal{S}^N$ is a sheaf, then contextuality of an empirical model cannot falsely arise due to pathologies in the way that $\mathcal{S}^N$ is defined.

To summarise,

1. We would like to show that the sheaf axiom holds with respect to the family of morphisms $F = \{f_i : G_i \rightarrow M\}$ defined by $\mathcal{C}$. In particular, we define $f_i : G_i \rightarrow M \in M \iff G_i \in \mathcal{C}$. Note that this does not imply that $\mathcal{S}^N$ is a sheaf, only that it obey the sheaf axiom with respect to a single cover.
2. The choice of cover $\mathcal{C}$ is physically motivated by the problem at hand. In other words, we do not give a method for ‘automatically’ defining $\mathcal{C}$ given a measurement game—this would be an interesting area for future direction, however.

We show that while for a compatible family $\{\sigma_i\}_{i \in I}$ for the presheaf $\mathcal{S}^N$ on the cover $\mathcal{C}$ the gluing property can always be satisfied, uniqueness does not necessarily hold. We will see nevertheless that for all cases we are interested in, uniqueness does hold.

Proposition 6.4.1. *The presheaf $\mathcal{S}^N$ satisfies the gluing axiom of the sheaf conditions.*

Proof. Take a compatible family $\{\sigma_i\}_{i \in I}$ for the presheaf $\mathcal{S}^N$. Therefore, for any $G_i \leq G_j$ with $G_i, G_j \in \mathcal{C}$ we have $\sigma_j|_{G_i} = \sigma_i$. Define $\sigma = \bigcup_{i \in I} \sigma_i$; since $\sigma|_{G_i} = \sigma_i$ for all $i \in I$ any global section must be a superset of σ and we take for now this to be the candidate gluing of the local sections. First, it is down-closed as the union of down-closed sets. Next suppose for some $s \in \sigma$, $s \triangleright^N c$. Then we must have $s \in \sigma_i$ for at least one $i \in I$. Suppose we have $s \in \sigma_1$ and $s \in \sigma_2$. Then since σ_i ($i \in \{1, 2\}$) are strategies if $c \in \Gamma_{G_i}$ then $s \triangleright_{G_i}^N c$ and there exists a unique $\langle c, v_i \rangle \in \sigma_i$ with $s \cup \{\langle c, v_i \rangle\} \in \sigma_i$. In order to show that σ is deterministic at Nature cells we must show that $v_1 = v_2$. On the other hand, if for no σ_i with $s \in \sigma_i$ we have $s \triangleright_{G_i}^N c$ (because $c \notin \Gamma_{G_i}$) then we can simply take any $v \in V_c$ and extend σ to $\sigma' := \sigma \cup (s \cup \{\langle c, v \rangle\})$, which still satisfies down-closure (although this extension, which is not unique, will cause problems with the uniqueness property of the sheaf condition).

(Showing $v_1 = v_2$ in first case) Since we require down-closure of $\mathcal{C}$, observe that there exists (at least one) substructure G_k in the cover with $G_k \leq G_1$ and $G_k \leq G_2$ defined as follows. It has as its set of cells $\Gamma_{G_k} \supseteq \text{dom}(s) \cup \{c\}$ and for any $c' \in \text{dom}(s) \cup \{c\}$, $(V_{G_k})_{c'} = (V_{G_1})_{c'}$ where w.l.o.g we assume $(V_{G_1})_{c'} \subseteq (V_{G_2})_{c'}$. Therefore $s \in \Delta_{G_k}$.

Now, $s \in \mathcal{H}(G_1)$ and $s \in \mathcal{H}(G_2)$ because by assumption $s \in \sigma_1$ and $s \in \sigma_2$ and by definition of Nature strategies σ_i is a down-closed subset of $\mathcal{H}(G_i)$. Since s is a history it must be that every $d \in \text{dom}(s)$ is enabled by some subset $r_d \subseteq s$, so

$r_d \vdash_{G_i} d$ for $i \in \{1, 2\}$. This implies, by definition of substructure, that $r_d \vdash_{G_k} d$. Furthermore, $s \triangleright_{G_i}^N c$ implies that $s \triangleright_{G_k}^N c$ so that $s \cup \langle c, v_i \rangle$ for any $v_i \in V_c$ is in $\mathcal{H}(G_k)$.

Since the family of σ_i 's is compatible, we must have $\sigma_1|_{G_k} = \sigma_k$ and $\sigma_2|_{G_k} = \sigma_k$, where recall restriction is defined $\sigma_i|_{G_k} = \sigma_i \cap \mathcal{H}(G_k)$. But we saw $s \cup \{\langle c, v_i \rangle\} \in \sigma_i$. All of this together gives,

$$\begin{aligned} s \cup \{\langle c, v_1 \rangle\} &\in \sigma_1|_{G_k} = \sigma_k \\ s \cup \{\langle c, v_2 \rangle\} &\in \sigma_2|_{G_k} = \sigma_k. \end{aligned} \tag{6.3}$$

But, since σ_k a Nature strategy so that it must be deterministic at Nature cells, there is a unique v such that $s \cup \{\langle c, v \rangle\} \in \sigma_k$ which implies that $v_1 = v_2 = v$.

Now suppose we have $s \triangleright^E c$ in σ . Then again we have $s \in \sigma_i$ for at least one $i \in I$. Consider the subset of the index set I defined

$$J := \{i \in I \mid s \in \sigma_i \wedge s \triangleright_{G_i}^E c\}. \tag{6.4}$$

Then if $V_c = \bigcup_{j \in J} (V_{G_j})_c$ then we have that $s \cup \langle c, v \rangle \in \sigma$ for all $v \in V_c$. This follows because since σ_i are strategies, for all $j \in J$, $s \cup \{\langle c, v \rangle\} \in \sigma_j$ for all $v \in (V_{G_j})_c$. Since every value in V_c is contained in one of the $(V_{G_j})_c$ this ensures that σ satisfies the required property.

If on the other hand we do not have $V_c = \bigcup_{j \in J} (V_{G_j})_c$, calling $W = V_c \setminus \bigcup_{j \in J} (V_{G_j})_c$, we can always extend σ to $\sigma' = \sigma \cup (\bigcup_{v \in W} \langle c, w \rangle)$. Note that such an extension is unique. $\square$

Proposition 6.4.2. *The gluing σ' described above is not necessarily unique.*

Proof. We give an example to demonstrate that uniqueness does not always hold. Consider the measurement game $M = \langle \Gamma, V, \vdash \rangle$ with two Experimenter cells and two Nature cells:

1. $\Gamma^E = \{A, B\}$, $\Gamma^N = \{a, b\}$
2. $c \in \{A, B, a\}$. $V_c = \{0\}$, $V_b = \{0, 1\}$
3. $\emptyset \vdash A$, $\{\langle A, 0 \rangle\} \vdash a$, $\{\langle a, 0 \rangle\} \vdash B$, $\{\langle B, 0 \rangle\} \vdash b$.

Then consider a cover with two top elements (i.e. they are not, in the poset, below any other elements of the cover) which we call $G_1 = \langle \Gamma_1, V_1, \vdash_1 \rangle$ and $G_2 = \langle \Gamma_2, V_2, \vdash_2 \rangle$ and define them respectively by

1. $\Gamma_1 = \{A, a\}$ and $\Gamma_2 = \{B, b\}$,
2. $\forall c \in \Gamma_1. (V_1)_c = \{0\}$ and $(V_2)_B = \{0\}$, $(V_2)_b = \{0, 1\}$
3. $\emptyset \vdash A$, $\{\langle A, 0 \rangle\} \vdash a$, $\{\langle B, 0 \rangle\} \vdash b$.

Note that we no longer have any enabling for the cell B in the game G_2 (and therefore the set of histories will be empty).

Now, taking $\sigma_1 = \{\{\langle A, 0 \rangle\}, \{\langle A, 0 \rangle, \langle a, 0 \rangle\}\}$ and $\sigma_2 = \emptyset$, their union must be extended in some way in order to be a valid Nature strategy on M . We could for example take the extension

$$\sigma = \{\{\langle A, 0 \rangle\}, \{\langle A, 0 \rangle, \langle a, 0 \rangle\}, \{\langle A, 0 \rangle, \langle a, 0 \rangle, \langle B, 0 \rangle\}, \{\langle A, 0 \rangle, \langle a, 0 \rangle, \langle B, 0 \rangle, \langle b, 1 \rangle\}\} \quad (6.5)$$

or, on the other hand, the extension

$$\sigma = \{\{\langle A, 0 \rangle\}, \{\langle A, 0 \rangle, \langle a, 0 \rangle\}, \{\langle A, 0 \rangle, \langle a, 0 \rangle, \langle B, 0 \rangle\}, \{\langle A, 0 \rangle, \langle a, 0 \rangle, \langle B, 0 \rangle, \langle b, 0 \rangle\}\}. \quad (6.6)$$

Both work, but send b to different outcomes, and therefore uniqueness does not hold. Note that taking $=_{\downarrow \leq} \{G_1, G_2\}$ does not affect this behaviour.

□

One way to ensure that it is unique is to ensure that a ‘good’ cover $\mathcal{C}$ is chosen. Note that if $\sigma = \bigcup_{i \in I} \sigma_i$ does not need to be extended to ensure it is a valid strategy for M , then uniqueness holds. This is because if $\sigma \in \mathcal{S}^N(M)$ and $\sigma' \in \mathcal{S}^N(M)$ we must have $\sigma \subseteq \sigma'$ since σ' must contain the union of the σ_i ’s to satisfy $\sigma'|_{G_i} = \sigma_i$ for all $i \in I$. But this implies that there exists some $s \in \sigma$ with $s \triangleright_i c$, which would invalidate σ as a strategy unless extended, but we have shown that it need not be extended.

Nevertheless, for the covers we use $\sigma := \bigcup_{i \in I} \sigma_i$ is a valid strategy and therefore both gluing and uniqueness hold, so that $\mathcal{S}^N$ is a valid sheaf. While we do not do so, it would be informative to give a characterisation of the covers which always satisfy uniqueness.

6.5 Causal measurement scenarios: examples

6.5.1 Causal Bell scenarios

Consider a causal Bell scenario $\mathcal{M} = \langle \mathcal{G}, \{I_v\}_{v \in V(\mathcal{G})}, \{O_v\}_{v \in V(\mathcal{G})} \rangle$ where $\mathcal{G}$ is a DAG specifying causal dependencies of agents, I_v specifies an input set for each agent, and O_v an output set.

Define an associated causal measurement game $M_{\mathcal{G}} = \langle \Gamma, V, \vdash \rangle$ where

i $\Gamma^E = V(G)$ and $\Gamma^N = \{a_i | A \in V(G) \wedge i \in I_A\}$;

ii By cases,

- If $A \in \Gamma^E$ then $V_A = I_A$
- If $a_i \in \Gamma^N$ then $V_{a_i} = O_A$

iii By cases,

- If $A \in \Gamma^E$ then $\bigcup_{A' \in \text{pa}(A)} \{\langle a'_i, o \rangle\} \vdash A$ for any $i \in I_{A'}$ and $o \in O_{A'}$,
- If $a_i \in \Gamma^N$ then $\{\langle A, i \rangle\} \vdash a_i$

A set of histories $h \in \mathcal{H}(M)$ is a set of events $h \subseteq \Delta_M$ which is consistent and secure. This is therefore, for any $S \in \Lambda(V(\mathcal{G}))$, the set

$$\bigcup_{A \in S} \{\langle A, i \rangle, \langle a_i, o \rangle\}. \quad (6.7)$$

It is possible also, for any $h \in \mathcal{H}(M)$, to define an order on events in h . Define for any two events $e, e' \in h$

$$e \leq_h e' \iff \forall y \in \mathcal{H}(\mathcal{M}). e' \in y \ \& \ y \subseteq h \implies e \in y. \quad (6.8)$$

We will write $e <_h e'$ for $e \leq_h e'$ and $e \neq e'$.

Example 6.5.1. For the DAG below, we have the relations $A < B$ and $A < C$.

A maximal element in $\mathcal{H}(M)$ will be of the form

$$\{\langle A, i \rangle, \langle a_i, o_a \rangle, \langle B, j \rangle, \langle b_j, o_b \rangle, \langle C, k \rangle, \langle c_k, o_c \rangle\}. \quad (6.9)$$

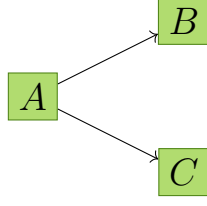


Figure 6.2: A three party causal Bell scenario. Party A causally precedes B and C , so that the outcomes at B and C may depend on both the input and outputs at A .

We find that this element has orderings

$$\langle A, i \rangle < \langle a_i, o_a \rangle \quad (6.10)$$

and similar for B and C , so that an agent picks a choice of measurement before Nature supplies the answer. Also, we have

$$\langle A, i \rangle < \langle B, j \rangle \quad \langle A, i \rangle < \langle C, k \rangle \quad (6.11)$$

and

$$\langle a_i, o_a \rangle < \langle b_j, o_b \rangle \quad \langle a_i, o_a \rangle < \langle c_k, o_c \rangle. \quad (6.12)$$

These express the causal ordering of the agents, as well as their measurements.

Then define the cover $\mathcal{C}$ to consist of substructures $G_{\mathbf{i}}$ where $\mathbf{i}$ is a tuple consisting of a single choice of input $\mathbf{i}(A) \in I_A$ for every $A \in S$ where S is down-closed in $V(G)$. The substructure $G_{\mathbf{i}}$ then has for each $A \in \Gamma^E$ a single value $V_A = \{\mathbf{i}(A)\}$, and restricts the set of Nature cells to $\Gamma_{\mathbf{i}}^N := \{a_l | \mathbf{i}(A) = l\}$. Thus there is one Nature cell associated with every agent in the graph. More generally, for a subset $S \subseteq V(\mathcal{G})$ which is down-closed w.r.t. the order on $\mathcal{G}$, and a choice $U_A \subseteq I_A$ for each $A \in S$ —which is summarised by a choice of tuple $(S, \{U_A\}_{A \in S})$ —we have the associated element $G_{(S, U)} \in \mathbf{Sub}(M)$ defined as follows,

$$\text{i } \Gamma_{(S, U)}^E = S \quad \Gamma_{(S, U)}^N = \{a_i | A \in S \wedge i \in U_A\};$$

ii By cases,

- If $A \in \Gamma_{(S, U)}^E$ then $V_A = U_A$
- If $a_i \in \Gamma_{(S, U)}^N$ then $V_{a_i} = O_A$

iii By cases,

- If $A \in \Gamma_{(S, \underline{U})}^E$ then $\bigcup_{A' \in \text{pa}(A)} \{\langle a'_i, o \rangle\} \vdash A$ for any $i \in U_{A'}$ and $o \in O_{A'}$,
- If $a_i \in \Gamma_{(S, \underline{U})}^N$ then $\{\langle A, i \rangle\} \vdash a_i$

Since it is these substructures that we are most interested in, rather than consider the full category $(\mathbf{Sub}(M), \leq)$ we consider the subcategory corresponding to the subposet $\{G_{(S, \underline{U})} | S \in \Lambda(V(\mathcal{G})) \wedge U_A \subseteq I_A\}$. We call this P , and it still has top element M .

It is then not difficult to show that for flat scenarios, or those scenarios in which $\mathcal{G}$ contains no edges, we have a bijection between Nature strategies on a substructure $G_{(S, \underline{U})}$ and sections $s \in \mathcal{E}(W)$ where $W = \Gamma_{(S, \underline{U})}^N$. In other words, we recover the usual notion of contextuality for flat scenarios. More generally, we recover the notion of causal contextuality studied by Gogioso and Pinzani in [66].

Proposition 6.5.2. *Consider the causal scenario $\mathcal{M} = (M_{\mathcal{G}}, \mathcal{C})$ where $M_{\mathcal{G}}$ is the causal game associated with the causal Bell scenario $\mathcal{M}' = \langle \mathcal{G}, \{I_v\}_{v \in V(\mathcal{G})}, \{O_v\}_{v \in V(\mathcal{G})} \rangle$. There exists a bijection α between the set of empirical models $\mathbf{EM}(\mathcal{M})$ and the set of GP (see Section 4.3) empirical models $\mathbf{EM}^{GP}(\mathcal{M}')$. Moreover, an empirical model $e \in \mathbf{EM}(\mathcal{M})$ is causally contextual if and only if $\alpha(e)$ is GP contextual.*

Proof. Note that if for every causal game $G_{\mathbf{i}}$ in the cover $\mathcal{C}$ we can define an associated $F(G_{\mathbf{i}}) \in \mathcal{C}^{GP}$ where recall the GP cover is the set $\bigcup_{S \in \Lambda(V(\mathcal{G}))} (\prod_{\omega \in S} I_{\omega})$ such that (i) there exists a bijection α between the sets $\mathcal{S}^N(G_{\mathbf{i}})$ and $\mathcal{E}_C(F(G_{\mathbf{i}}))$ and moreover (ii) if F itself is a bijection between covers, then we can automatically extend the bijection α to a bijection probability distributions on these sets, and this together with (ii) ensures a bijection of empirical models—we call this bijection α' .

We therefore begin by defining for any $G_{\mathbf{i}} \in \mathcal{C}$, $F(G_{\mathbf{i}}) = \mathbf{i}$. Since

$$\mathbf{i} \in \bigcup_{S \in \Gamma(V(\mathcal{G}))} \left(\prod_{\omega \in S} I_{\omega} \right), \quad (6.13)$$

this is an element of the cover $\mathcal{C}^{GP}$. Moreover, it is a bijection by taking $F^{-1}(\mathbf{i}) = G_{\mathbf{i}}$. For a Nature strategy $n \in \mathcal{S}^N(G_{\mathbf{i}})$ define a section $\alpha(n) \in \mathcal{E}(\mathbf{i})$ to be the section sending $\mathbf{i}(a)$ to o provided $\langle a_{\mathbf{i}(a)}, o \rangle \in n$. On the other hand given $f \in \mathcal{E}(\mathbf{i})$ define $\alpha^{-1}(f)$ to be the nature strategy which contains Nature decisions $\langle a_{\mathbf{i}(a)}, f(\mathbf{i}(a)) \rangle$. Experimenter decisions are, independent of f , fixed as $\langle a, \mathbf{i}(a) \rangle$.

Since we have shown (i) and (ii) hold, we have a bijection of empirical models $\mathbf{EM}^{GP}(\mathcal{M}')$ and $\mathbf{EM}(\mathcal{M})$. In order to show that contextuality is preserved and reflected in mapping a causal empirical model e to a GP empirical model e' , we now consider the top elements of the posets $\mathbf{Sub}(M)$ and $\mathcal{L}$ and associated global strategies.

The top element of $\mathbf{Sub}(M)$ is M and the top element of $\mathcal{L}$ is $(V(\mathcal{G}), \underline{I})$. As on the cover, given a global strategy $m \in \mathcal{S}^N(M)$ let the causal section $\alpha(m) \in \mathcal{E}((V(\mathcal{G}), \underline{I}))$ send $m(\mathbf{i})$ to the associated tuple of outcomes $\mathbf{o}$ provided that there exists a history $h \in m$ which contains for each agent A the decision $\langle A_{\mathbf{i}(a)}, \mathbf{o}(a) \rangle$ (naturally h will also contain Experimenter decisions $\langle a, \mathbf{i}(a) \rangle$). On the other hand given a causal section $g \in \mathcal{E}((V(\mathcal{G}), \underline{I}))$ define a strategy $\alpha^{-1}(g)$ to contain as its maximal elements the histories $h_{\mathbf{i}, \mathbf{o}}$ which contain decisions $\langle a_{\mathbf{i}(a)}, \mathbf{o}(a) \rangle$ and for all agents A . Now, suppose $e \in \mathbf{EM}(\mathcal{M})$ is causally contextual. Then there exists a probabilistic global strategy $p \in \mathcal{S}^N(M)$ where

$$\begin{aligned}
e_{G_{\mathbf{i}}}(n) &= \sum_{\substack{m \in \mathcal{S}^N(M); \\ m|_{G_{\mathbf{i}}=n}}} p(m) = \sum_{m \in \mathcal{S}^N(M)} \delta_{m|_{G_{\mathbf{i}}}(n)} p(m) \\
&= \sum_{\alpha(m) \in \mathcal{E}(\underline{I})} \delta_{\alpha(m)|_{\mathbf{i}}(\alpha(n))} \alpha'(p)(\alpha(m)) \\
&= e'_{\mathbf{i}}(\alpha(n)).
\end{aligned} \tag{6.14}$$

where e' is noncontextual with global section $\alpha'(p)$. The proof in the opposite direction, starting assuming contextuality of e' is exactly analogous since every step in the above deduction is reversible. $\square$

We consider the cover $\mathcal{C}$ of flat scenarios in more detail. In the two party case, $\mathbf{i}$ can either have support on both agents A and B in which case we write it as (i_a, i_b) , or on one or the other in which case we write it as $(i_a, -)$ or $(-, i_b)$ when it has support on only A and only B respectively. The set $\mathcal{C}$ has properties similar to a simplicial complex. We have

- i If $G \in \mathcal{C}$ and $G' \in P$ with $G' \trianglelefteq G$ then $G' \in \mathcal{C}$.
- ii $M = \bigvee_{G \in \mathcal{C}} G$.

which look identical to the axioms of a simplicial complex defined over a set X , except that the subset relation is replaced by $\trianglelefteq$ $\square$

In order to visualise it, we take the set $\{G_{(-,0)}, G_{(-,1)}, G_{(0,-)}, G_{(1,-)}\}$ to correspond to the one dimensional faces (i.e. the vertices), while the set $\{G_{(0,0)}, G_{(0,1)}, G_{(1,0)}, G_{(1,1)}\}$ corresponds to the two-dimensional faces. The cover $\mathcal{C}$ is therefore visualised by the simplicial complex

¹Note that a simplicial complex can, alternative to the definition given in Chapter 2, have the condition that requires singletons are contained in it with the condition that the union of all elements give the set X .

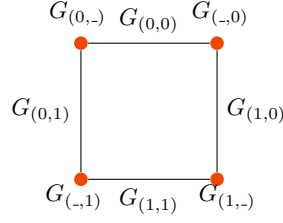


Figure 6.3: Visualisation of the cover for the Bell-CHSH setup when the base category is constructed from the poset of measurement games.

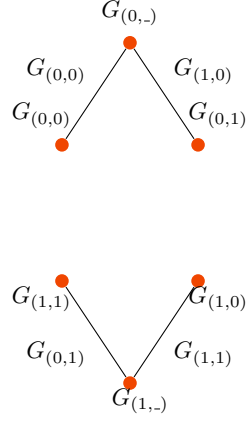


Figure 6.4: Visualisation of the cover for the temporal CHSH setup when the base category is constructed from the poset of measurement games. Note that, unlike the case in which there is no order on the parties A and B , measurement games of the form $G_{(-,i)}$ are not in the cover.

For non-flat scenarios the cover $\mathcal{C}$ still satisfies the generalised simplicial complex axioms (i) and (ii), however, the choice of visualisation is not so straightforward. Taking $A \rightarrow B$ in $\mathcal{G}$ the cover now contains elements of the form $G_{(i_A,-)}$, but not of the form $G_{(-,i_B)}$. The previous representation no longer makes sense; we instead consider all elements in $\mathcal{C}$ as vertices and a labelled edge H between $G_1, G_2 \in \mathcal{C}$ indicates that $H = G_1 \wedge G_2$. More generally, an n dimensional face is an element $H \in \mathcal{C}$ with $H = G_1 \wedge G_2 \wedge \dots \wedge G_n$. For $A \rightarrow B$ we obtain

6.5.2 Contextuality scenarios

Contextuality scenarios arise as a special case in which we have a single Experimenter cell E which is initially enabled, and where the cover $\mathcal{C}$ consists of substructures G_σ where $V_E = \sigma$ and σ is a commuting set of measurements.

Therefore in this case, the commuting structure of the measurement set is im-

posed by the commutation relation, and not enforced due to space-time separations between agents. Recall that a global section is a function $s : X \rightarrow O$, and that X is the top element in the category $\mathbf{Op}(X)$ over which we take our cover for the empirical model. Here, the corresponding top element M is a measurement game in which the agent has choice of any measurement in X . The set of Nature strategies $\mathcal{S}^N(M)$ is then in bijection with the set of sections O^X .

More explicitly, given a contextuality scenario $\mathcal{M}_C = \langle X, \Sigma, O \rangle$, define a set of measurement games $\{G_U\}_{U \in \mathcal{P}(X)}$ where $G_U = \langle \Gamma, V, \vdash \rangle$ is defined 2

1. $\Gamma_U^E = \{E\}$ and $\Gamma_U^N = U$
2. $V_E = U$
3. For any $x \in U$, $V_x = O_x$
4. We have the enabling

$$\emptyset \vdash E \quad \forall x \in U. \{\langle E, x \rangle\} \vdash x,$$

so that the single Experimenter is initially enabled, and any measurement is enabled once it has been chosen by the Experimenter.

The corresponding set of histories is given by

$$\mathcal{H}(G_U) = \bigcup_{x \in U, o \in O_x} \{\{\langle E, x \rangle\}, \{\langle E, x \rangle, \langle x, o \rangle\}\}. \quad (6.15)$$

An Experimenter strategy e is the subset

$$e = \bigcup_{o \in O_x} \{\{\langle E, x \rangle\}, \{\langle E, x \rangle, \langle x, o \rangle\}\} \quad (6.16)$$

for a fixed $x \in U$. The strategy chooses some $x \in U$ but allows for any outcome for that measurement. On the other hand, a Nature strategy n is the subset

$$n = \bigcup_{x \in U} \{\{\langle E, x \rangle\}, \{\langle E, x \rangle, \langle x, o \rangle\}\} \quad (6.17)$$

which specifies a (single) outcome for each $x \in U$. As a result, we have that

$$\mathcal{E}(U) \cong \mathcal{S}^N(G_U). \quad (6.18)$$

The set of sections on U is in bijection with the set of Nature strategies on G_U . Consider as a concrete example the Peres-Mermin square.

Example 6.5.3 (Peres-Mermin Square). *The Peres-Mermin square is shown in Table 6.1, where we have replaced the quantum measurements with letters (this simplifies the notation from the usual Pauli exposition without losing any of the interesting features). Recall that the contextuality scenario is $\mathcal{M} = \langle X, \Sigma, O \rangle$ where the*

a	b	c	1
A	B	C	1
α	β	γ	1
1	1	-1	

Table 6.1: Peres-Mermin Square

measurement set is $X = \{a, b, c, A, B, C, \alpha, \beta, \gamma\}$, the maximal contexts in Σ are the 3 rows and 3 columns of the table which we label from σ_1 to σ_6 , and the output for each measurement lies in the set $O = \{-1, 1\}$. We saw in the background material that contextuality of this setup can be explained by considering an empirical model $e \in \mathbf{EM}(\mathcal{M})$ which has support consistent with the table above. For example, considering the context $\{a, b, c\}$ (row 1), the empirical model should only have support on the sections which send either all three measurements to 1 or two of them to -1 and the third to 1—this is the only way to ensure that their product be equal to 1. We now consider how this can be captured by a causal measurement scenario, instead. Take $\mathcal{M} = \langle M, \mathcal{C} \rangle$ where the causal game $M = \langle \Gamma, V, \vdash \rangle$ is defined

1. $\Gamma^E = \{E\}$ and $\Gamma^N = X$;
2. $V_E = \{X\}$ and for any $x \in \Gamma^N$, $V_x = \{-1, 1\}$;
3. $\emptyset \vdash E$ and for all $x \in \Gamma^N$, $\langle E, x \rangle \vdash x$.

The cover for the empirical model consists of substructures G_i for $i \in [1, 6]$ where G_i has $\Gamma_i^N = \sigma_i$. A Nature strategy on for example G_1 with $\Gamma_1^N = \{a, b, c\}$ is

$$\{\{\langle E, a \rangle, \langle a, o_a \rangle\}, \{\langle E, b \rangle, \langle b, o_b \rangle\}, \{\langle E, c \rangle, \langle c, o_c \rangle\}\}. \quad (6.19)$$

6.5.3 Temporal scenarios

We begin with temporal Bell scenarios. Given the triple $\langle \mathcal{A}, \underline{I}, \underline{O} \rangle$, with $\mathcal{A}$ a total order so that we take $\mathcal{A} = \{1, 2, 3, \dots, N\}$, the set of experimenter cells is

$$\Gamma^E = \mathcal{A}. \quad (6.20)$$

Recall that the set of measurements is defined

$$X := \{k_i | k \in [1, N] \wedge i \in I_k\}. \quad (6.21)$$

We define the set of Nature cells

$$\Gamma^N = X, \quad (6.22)$$

so that we have one Nature cell for each observable $k_i \in X$ of the system. Each choice of measurement $i \in I_k$ for an agent k enables the corresponding Nature cell k_i , or symbolically

$$\{\langle k, i \rangle\} \vdash k_i. \quad (6.23)$$

Moreover, a decision by Nature on an outcome for that measurement enables the following agent to choose their measurement.

$$\{\langle k_i, o_k \rangle\} \vdash k + 1. \quad (6.24)$$

An element h of the set of histories $\mathcal{H}(\mathcal{M})$ generated in this way is of the form

$$h = \{\langle 0, i_0 \rangle, \langle 0_{i_0}, o_0 \rangle, \langle 1, i_1 \rangle, \langle 1_{i_1}, o_1 \rangle, \dots, \langle n, i_n \rangle, \langle n_{i_n}, o_n \rangle\}. \quad (6.25)$$

When the history h is realised, it expresses that the set of decisions above occurred. Thus note that the set $\mathcal{H}(\mathcal{M})$ can be viewed as a decision tree, which for temporal setups is particularly useful visually. The tree for just the first two agents A and B is shown in Figure [6.5](#).

Each branch is a history $h \in \mathcal{H}(\mathcal{M})$. We highlight a particular example of a history. Note that the order on h is

$$\langle 0, i_0 \rangle <_h \langle 0_{i_0}, o_0 \rangle <_h \langle 1, i_1 \rangle <_h \langle 1_{i_1}, o_1 \rangle <_h \dots <_h \langle n, i_n \rangle <_h \langle n_{i_n}, o_n \rangle, \quad (6.26)$$

so that there is a total order on events (as we would expect). First, Experimenter

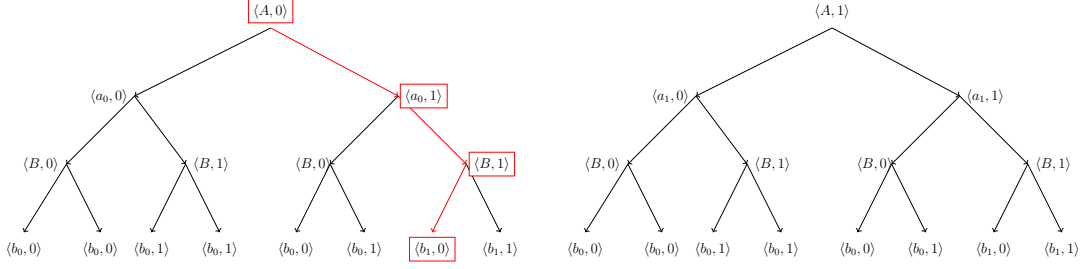


Figure 6.5

Figure 6.6: An example of a valid history for a temporal Bell scenario, as highlighted in red. The full tree represents the entire set of possible histories for the scenario.

0 chooses a measurement, then Nature provides an outcome to that measurements, then Experimenter 1 chooses a measurement, Nature provides a corresponding outcome, and so on.

We next consider Nature strategies $\mathcal{S}^N(\mathcal{M})$, which are deterministic at measurement outcomes but nondeterministic at measurement choices. Consider the temporal CHSH setup, with two agents A and B each with two inputs and two outputs. We display a Nature strategy as a subset of the set of histories graphically below.

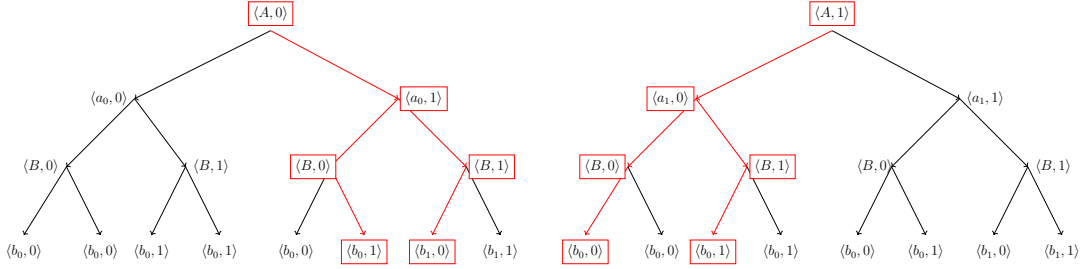


Figure 6.7: A valid Nature strategy, highlighted in red, as a subset of the set of histories. Note that the strategy is deterministic at Nature cells.

On the other hand, an Experimenter strategy $e \in \mathcal{S}^E(\mathcal{M})$ is deterministic on measurement choices but not on outcomes, as can be seen in the highlighted Experimenter strategy below,

so that agent B chooses their measurement based on the outcome at A . Note that together, a nature strategy n and an experimenter strategy e determine a unique history $h \in \mathcal{H}(\mathcal{M})$. We write this as $\langle n|e \rangle := n \cap e$, which takes the branch which lies at the overlap of the two trees.

Now consider a general temporal measurement scenario $\langle X, \Sigma, O \rangle$. Here we need at least N experimenter cells where $N := \max_{\sigma \in \Sigma} |\sigma|$, so we define

$$\Gamma^E = \{1, 2, \dots, N\}. \quad (6.27)$$

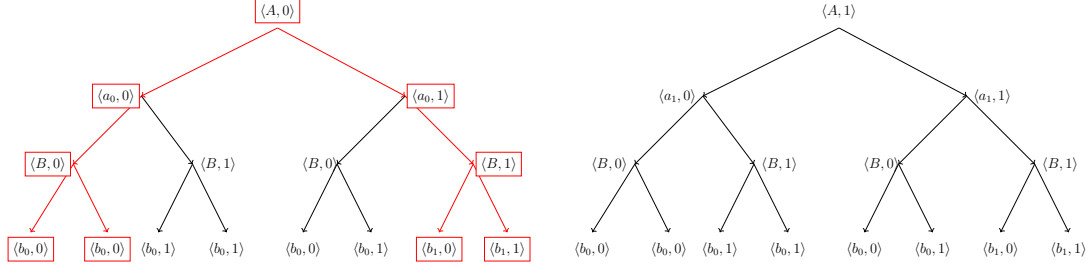


Figure 6.8: A valid Experimenter strategy, highlighted in red, as a subset of the set of histories. Note that the strategy is deterministic at Experimenter cells.

Since Σ dictates the allowed sequences of measurements, we take X_i to be the set of measurements which occur at position i in sequences $\sigma \in \Sigma$, $X_i := \{ m \mid \exists \sigma \in \Sigma. \sigma[i] = m \}$. Then any experimenter cell $i \in [1, N]$ can take on values in the set X_i ,

$$V_i = X_i \quad (6.28)$$

A measurement m in X_i is enabled when experimenter i fills their cell with the value m ; there is a slight subtlety however which arises with repeated measurements. If for $i \neq j$ we have a $\sigma \in \Sigma$ with $\sigma[i] = m$ and $\sigma[j] = m$, then the Nature cell m must be filled twice in a particular sequence, which is not allowed in the current framework (the resulting history does not obey the *consistency* axiom). One gets around this by labelling measurements with their time-step, which amounts to taking a Heisenberg picture stance, in which it is the operators that evolve and the wavefunction is stationary. We therefore define a Nature cell to be a measurement labelled with the time step at which it is performed,

$$\Gamma^N = \bigcup_{i \in [1, N]} \{i\} \times X_i \quad (6.29)$$

and a measurement (i, m) is enabled only when an experimenter at time step i fills their cell with m

$$\{\langle i, m \rangle\} \vdash_N (i, m). \quad (6.30)$$

The experimenter at time step i is enabled once the measurement cell at time step $i - 1$ has obtained an outcome,

$$\{\langle (i - 1, m), o \rangle\} \vdash i. \quad (6.31)$$

Note that the set of histories generated includes all combinations of measurement

events which arise when each Experimenter can choose from their set X_i , while we would like sequences to come from Σ . This can be fixed by taking as the cover the set $\{G_\sigma\}_{\sigma \in \Sigma}$ which only allows the experimenter cell i to be filled with the single value $\sigma[i]$.

We now consider how Experimenter strategies play a role, and in particular how Experimenter strategies allow us to consider examples which include adaptivity. This includes both HLP examples and a more general version of Anders and Browne’s OR gate implementation to a double OR gate implementation, with the second having as its input the output of the previous gate.

6.6 Adaptive Experimenter strategies

To understand why those examples studied by Henson Lal and Pusey fall outside of the current framework, consider the *Instrumental Setup*, introduced in Section 4.2. We will use the diagrammatic representation introduced there, too. The instrumental setup is described by the following DAG,

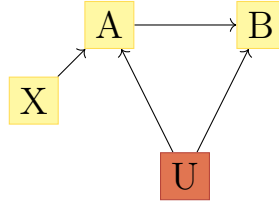


Figure 6.9: The Instrumental setup as a Bayesian network DAG

Recall that unlabelled edges between nodes indicate the sending of classical information (e.g. a bit) from the source node to the sink node. Oftentimes this will indicate sending the output of a measurement to be used as a measurement setting for a future agent. Labelled edges correspond to the sending of hidden variables.

Note that the setup in Figure 6.9 is different from the following setup, in which each agent has their own measurement setting, but the hidden variable is sent from A to B after measurement. This corresponds to a temporal CHSH setup.



Figure 6.10: The temporal CHSH setup as a causal Bell scenario DAG.

While we have seen that the framework is able to capture the type of causality that appears in this second digram, it cannot capture the type of adaptivity used in the Instrumental setup.

This is because while the enabling $\vdash$ in the definition of causal games is an enabling at the level of Nature i.e. any hidden variables are allowed to use this causal structure, here instead we would like to introduce an enabling at the level of the Experimenter only. What this means in practice is that the hidden variable λ is distributed prior to measurement (to all parties, if it is a Bell scenario), and after that, measurement with Experimenter imposed adaptivity can take place. This is exactly how the Instrumental setup is carried out. As in Figure 6.9, $\lambda \in U$ is distributed to A and B, and simultaneously an input $x \in X$ is distributed to A. The measurement a at A is performed and must be a function of λ and x , after which it is sent to B so that the measurement at B may be a function of λ and a . This is different from Figure 6.10, where A's input and output may be signalled. Alternatively, one can take the viewpoint that the hidden variable can update in between measurements (see Section 4.3).

To this end we introduce the enabling relation $\models \subseteq \mathcal{P}(\Delta_{\mathcal{M}}) \times \Gamma$ such that

$$\begin{aligned} X \models^E c &\implies \exists X' \subseteq X. X' \vdash^E c, \\ X \models^N c &\iff X \vdash^N c. \end{aligned} \tag{6.32}$$

This ensures that the experimenter imposed enabling (i) imposes causality constraints which respect those of $\vdash$ for Experimenter cells, and (ii) imposes no additional constraints for Nature cells.

The set of adaptive histories $\mathcal{H}^A(\mathcal{M})$ is defined analogously to $\mathcal{H}(\mathcal{M})$, except that we require cells in any $h \in \mathcal{H}^A(\mathcal{M})$ be secure with respect to $\models$. Note that

$$\mathcal{H}^A(\mathcal{M}) \subseteq \mathcal{H}(\mathcal{M}). \tag{6.33}$$

Moreover let $s \blacktriangleleft c$ be the accessibility relation with respect to $\models$, instead.

An adaptive strategy is then defined

Definition 6.6.1 (Adaptive Experimenter Strategy). *An adaptive Experimenter strategy is a subset $\tau \subseteq \mathcal{H}^A(\mathcal{M})$ such that*

- *Given $h, t \in \mathcal{H}^A(\mathcal{M})$, if $h \in \tau$ and $t \subseteq h$ then $t \in \tau$.*
- *If $s \in \tau$ and $s \blacktriangleleft^E c$ then there exists a unique $v \in V_c^Q$ such that $s \cup \{\langle c, v \rangle\} \in \tau$.*
- *If $s \in \tau$ and $s \blacktriangleleft^N c$ then for all $v \in V_c^A$ it holds that $s \cup \{\langle c, v \rangle\} \in \tau$.*

A special case we will be interested in is flat Bell scenarios with some adaptivity specified by a DAG $\mathcal{G}$ on the set of Experimenter cells Γ^N . The enabling is then defined

$$\bigcup_{A' \in \text{pa}_{\mathcal{G}}(A)} \{\langle a'_i, o \rangle\} \models A, \quad (6.34)$$

so that an agent A can perform their measurement once every agent in the set of parents of A' in $\mathcal{G}$, or $\text{pa}_{\mathcal{G}}(A)$, has performed their i th measurement a'_i **and** obtained an outcome. We will write $\models_{\mathcal{G}}$ for an adaptive enabling which arises in this way, and $\tau(\mathcal{G})$ for the corresponding choice of strategy to make its dependency on $\mathcal{G}$ explicit.

Let us analyse the effect on empirical models first through example, before giving the general definition. Recall that the PR Box empirical model can be captured with the following bundle diagram.

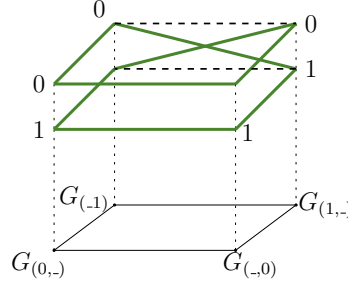


Figure 6.11: The PR box bundle diagram.

Such an empirical model is strongly contextual because there exists no closed paths in the bundle diagram.

The cover for the causal scenario consists of the substructures $\{G_{(i_a, i_b)}\}_{i_a, i_b \in \{0,1\}}$ where $G_{(i_a, i_b)}$ is restricted to having $V_A = \{i_a\}$ and $V_B = \{i_b\}$ —thus, each face in the base of the bundle diagram corresponds to a substructure $G_{(i_a, i_b)}$, and for example the line which represent the assignment $\{a_0 \mapsto 1, b_1 \mapsto 1\}$ is the Nature strategy with maximal element

$$\{\langle A, 0 \rangle, \langle B, 1 \rangle, \langle a_0, 1 \rangle, \langle b_1, 1 \rangle\}. \quad (6.35)$$

The substructure $G_{(i_a, -)}$ has only two valid Nature strategies, one containing the event $\langle a_{i_a}, 0 \rangle$ and the other the event $\langle a_{i_a}, 1 \rangle$. The stalk above $G_{(i_a, -)}$ in the bundle diagram therefore contains two possible assignments, which are these strategies. Now suppose that the exact same PR Box is measured, but the experimenter decides to choose the measurement for agent B based on the outcome of the agent A . Thus,

we have $A \rightarrow B$ in $\mathcal{G}$, and the adaptivity is defined

$$\begin{aligned} \emptyset \models_{\mathcal{G}} A & \quad \{\langle A, i \rangle\} \models_{\mathcal{G}} a_i \\ \{\langle a_i, o_1 \rangle\} \models_{\mathcal{G}} B & \quad \{\langle B, i \rangle\} \models_{\mathcal{G}} b_i \end{aligned} \tag{6.36}$$

One type of valid experimenter strategy τ is shown in Figure 6.12, where the choice of measurement at B is given by the outcome at A (note that there are other experimenter strategies consistent with the given enabling).

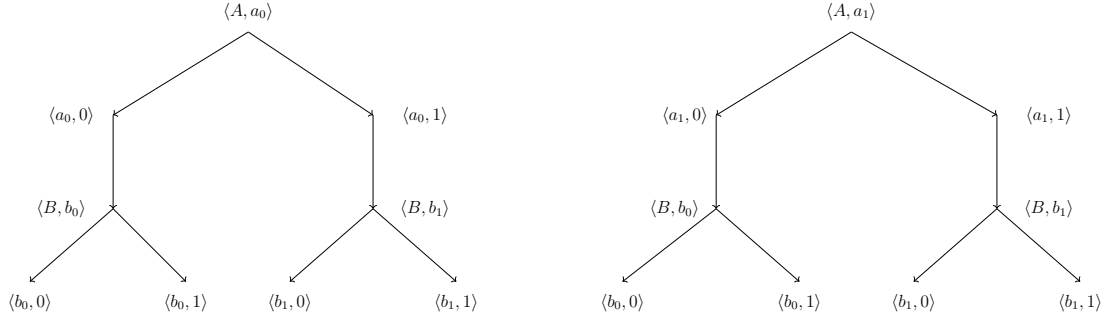


Figure 6.12: A choice of Experimenter strategy for the CHSH setup. The experimenter chooses their cell value based on the outcome of the measurement (either a_0 or a_1) which was performed before.

Therefore, in each round of experiment

1. Experimenter A chooses either measurement a_0 or a_1 to perform.
2. Nature responds with an outcome from the binary set $\{0, 1\}$.
3. Experimenter B chooses their measurement based on the outcomes of A 's measurement; they choose b_i ($i \in \{0, 1\}$) when outcome i is obtained at A .
4. Nature responds with an outcome from the binary set $\{0, 1\}$.

The type of strategies we actually observe on the cover $\{G_{(i,j)}\}$ of substructures will be $\langle \sigma | \tau \rangle$. Taking for example $V_A = \{0\}$ and $V_B = \{0\}$, only two strategies from $\Gamma^N(G_{(0,0)})$ are nonempty when taking the inner product with τ restricted to this substructure. These have maximal elements

$$\max(\sigma) = \{\langle A, 0 \rangle, \langle a_0, 0 \rangle, \langle B, 0 \rangle, \langle b_0, 0 \rangle\} \tag{6.37}$$

$$\max(\sigma') = \{\langle A, 0 \rangle, \langle a_0, 0 \rangle, \langle B, 0 \rangle, \langle b_0, 1 \rangle\} \tag{6.38}$$

and only σ is in the support of the PR box empirical model, since there is no line between $\langle a_0, 0 \rangle$ and $\langle b_0, 1 \rangle$. Doing this for each substructure, we obtain the diagram below.

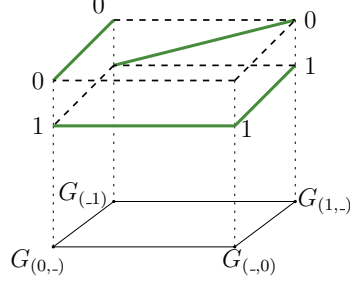


Figure 6.13: The bundle diagram when the Experimenter strategy is implemented on a PR box. The number of behaviours, or strategies, which are observed is reduced.

The resultant distribution that we observe is of course no longer no-signalling from A to B . Moreover, what does it mean for the resultant model to be non-contextual, since only a subset of the distribution of the PR Box is actually observed?

The problem at hand is that we are looking for a global section for the no-signalling model before experimenter strategies are imposed, and therefore a global section $h \in \mathcal{D}_R \circ \Gamma^N(\Gamma)$ of probabilistic Nature strategies which fits with the observed data. Before the experimenter strategy is imposed, the empirical model which a particular system **would have** produced is a family of distributions $\{e_C \in \mathcal{D}_R \circ \Gamma^N(G_{(i,j)})\}_{i,j \in \{0,1\}}$. What we observe nevertheless are the conditional distributions (see Section 2.3)

$$e_G|_{p_G}(\sigma) := \frac{e_G(\sigma)p_G(\sigma)}{\sum_{\sigma' \in \Gamma^N(G)} p_G(\sigma')e_G(\sigma')} \quad (6.39)$$

where $\sigma \in \Gamma^N(G)$ and the predicate for the conditional distribution is defined

$$p_G(s) = 1 \iff \langle \sigma | (\tau|_G) \rangle \neq \emptyset. \quad (6.40)$$

The question therefore becomes whether there exists an empirical model e which results in the observed distributions above **and** such that e has a global section. Note that the empirical model e which we construct need not coincide with the one which would actually be produced if measurement outcomes were not fed forward. In the case of the instrumental setup, we can in fact find such an empirical model which is noncontextual.

Consider the no-signalling empirical model displayed below, where we have extended the observed sections to create a valid no-signalling model.

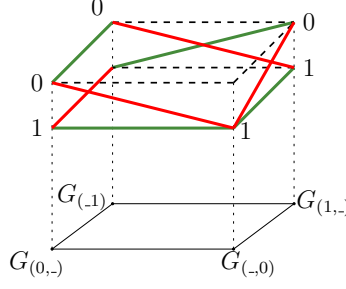


Figure 6.14: A reconstruction of *observed* bundle diagram in Figure 6.13 which is classical.

This is non-contextual, because every line extends into a closed path in the bundle diagram. In particular, there are two closed paths and every line is part of one of them. The global section displayed in the image can be written down in terms of nature strategies as

$$|p\rangle = 0.5\vec{s}_1 + 0.5\vec{s}_2 \quad (6.41)$$

where ³

$$\begin{aligned} & \{\{\langle a_0, 0 \rangle, \langle b_0, 0 \rangle\}, \{\langle a_1, 1 \rangle, \langle b_0, 0 \rangle\}, \{\langle a_0, 0 \rangle, \langle b_1, 0 \rangle\}, \{\langle a_1, 1 \rangle, \langle b_1, 0 \rangle\}\} \subset s_1 \\ & \{\{\langle a_0, 1 \rangle, \langle b_0, 0 \rangle\}, \{\langle a_1, 0 \rangle, \langle b_0, 0 \rangle\}, \{\langle a_0, 1 \rangle, \langle b_1, 1 \rangle\}, \{\langle a_1, 0 \rangle, \langle b_1, 1 \rangle\}\} \subset s_2 \end{aligned}$$

The example above demonstrates that the choice of Experimenter strategy has an effect on the contextuality of the observed data, essentially by restricting which data is observed. We define the τ –**conditional empirical model** to be the empirical model obtained if an Experimenter with access to an empirical model e follows the Experimenter strategy τ .

Definition 6.6.2 (τ –Conditional Empirical Model). *Given an adaptive experimenter strategy τ and an empirical model $e \in \mathbf{EM}(\mathcal{M})$, the τ –conditional empirical model is given by the conditional distributions*

$$e_G|_{p_G}(s) := \frac{e_G(s)p_G(s)}{\sum_{s' \in \Gamma^N(G)} p_G(s')e_G(s')} \quad (6.42)$$

³we disregard Experimenter events in such strategies for clarity of presentation—ambiguity is not lost since each Experimenter cell is associated with a unique value in each Nature strategy.

where $s \in \Gamma^N(G)$, $G \in \mathcal{C}$, and the predicate for the conditional distribution is defined

$$p_G^\tau(s) = 1 \iff \langle s | (\tau|_G) \rangle \neq \emptyset. \quad (6.43)$$

We can then define contextuality of a τ -conditional empirical model.

Definition 6.6.3 (Contextuality of a t -Conditional Empirical Model). *The τ -conditional empirical model $e_G|_{p_G^\tau}$ for $G \in \mathcal{C}$ is noncontextual if there exists some empirical model $w \in \mathbf{EM}(\mathcal{M})$ with for all $G \in \mathcal{C}$*

$$e_G|_{p_G^\tau}(s) = w_G|_{p_G^\tau}(s) \quad (6.44)$$

such that w is a noncontextual empirical model. If there exists no such empirical model $w \in \mathbf{EM}(\mathcal{M})$, then it is contextual.

Note that introduction of an adaptive strategy makes it *less likely* that the resulting empirical model is contextual, since the number of observed strategies is reduced. We now consider the Anders-Browne and Henson-Lal-Pusey examples.

6.6.1 Anders-Browne example

Anders and Browne's implementation of the OR gate (see Section 3.7) uses the entanglement of a GHZ state to promote a subuniversal classical computer to one of universal computing power.

While the Nature strategies correspond to those of a three party (Alice, Bob, Charlie) Bell scenario, the input for Charlie is chosen by the Experimenter as the XOR of the inputs for Alice and Bob—we require that Nature i.e. the hidden variable is oblivious to this. We introduce additional (Nature agnostic) causal structure

$$\emptyset \models A, \quad \emptyset \models B, \quad \{\langle A, i \rangle, \langle B, j \rangle\} \models C. \quad (6.45)$$

Thus Charlie must wait until the inputs of Alice and Bob have been chosen until they can choose theirs.

In particular, we define the Experimenter strategy τ (M is defined in the usual way for flat Bell scenarios) to be that with maximal elements (for fixed i and j)

$$\{\langle A, i \rangle, \langle B, j \rangle, \langle C, i \oplus j \rangle, \langle (a, i), o_1 \rangle, \langle (b, j), o_2 \rangle, \langle (c, i \oplus j), o_3 \rangle\} \quad (6.46)$$

with $o_1, o_2, o_3 \in \{0, 1\}$, so that there are 2^3 maximal elements in τ . The τ -conditional empirical model is identical to the Bell empirical model on contexts $\{(a, i), (b, j), (c, i \oplus j)\}$ and zero otherwise.

We can consider a second OR gate with Experimenter cells A', B' and C' , again with the input at C' being dependent on the inputs at A' and B' , but with the additional caveat that the input at A' is the XOR $o_1 \oplus o_2 \oplus o_3$ of the outcomes of the outcomes obtained in the previous OR gate.

In this case we define the enabling

$$\begin{aligned} \emptyset \models A, \quad \emptyset \models B, \quad \{\langle A, i \rangle, \langle B, j \rangle\} \models C \\ \emptyset \models B', \quad \{\langle (a, i), o_1 \rangle, \langle (b, j), o_2 \rangle, \langle (c, k), o_3 \rangle\} \models A' \quad \{\langle A', i \rangle, \langle B', j \rangle\} \models C'. \end{aligned} \tag{6.47}$$

Now, while B' need not wait for the first gate to be performed, the agent A' must wait for Nature cells (a, i) , (b, j) and (c, k) to be filled (i.e. outcomes obtained for the measurements), so that the outcomes $o_1 - o_3$ affect the choice of measurement at A' .

The appropriate adaptive Experimenter strategy τ has as its set of maximal elements

$$\begin{aligned} \{\langle A, i \rangle, \langle B, j \rangle, \langle C, i \oplus j \rangle, \langle (a, i), o_1 \rangle, \langle (b, j), o_2 \rangle, \langle (c, i \oplus j), o_3 \rangle, \langle A', o_1 \oplus o_2 \oplus o_3 \rangle, \\ \langle B', m \rangle, \langle C', (o_1 \oplus o_2 \oplus o_3) \oplus m \rangle, \langle (a', o_1 \oplus o_2 \oplus o_3), o'_1 \rangle, \langle (b', m), o'_2 \rangle, \\ \langle (c', (o_1 \oplus o_2 \oplus o_3) \oplus m), o'_3 \rangle\}, \end{aligned} \tag{6.48}$$

where every maximal element in τ has the same choice of i, j and m , but different combinations of $o_1 - o_3$ and $o'_1 - o'_3$.

6.6.2 HLP scenarios

Recall that the scenarios considered by Henson et al. in [73] are DAGs which consist of visible nodes and hidden nodes. We specialise to the graphs which

- i Have a single hidden node Λ
- ii All other nodes are either (a) ‘agent’ nodes A^i which connect to the hidden node or (b) ‘input’ nodes I^j which do not connect to the hidden node and connect directly to one or more agent nodes.

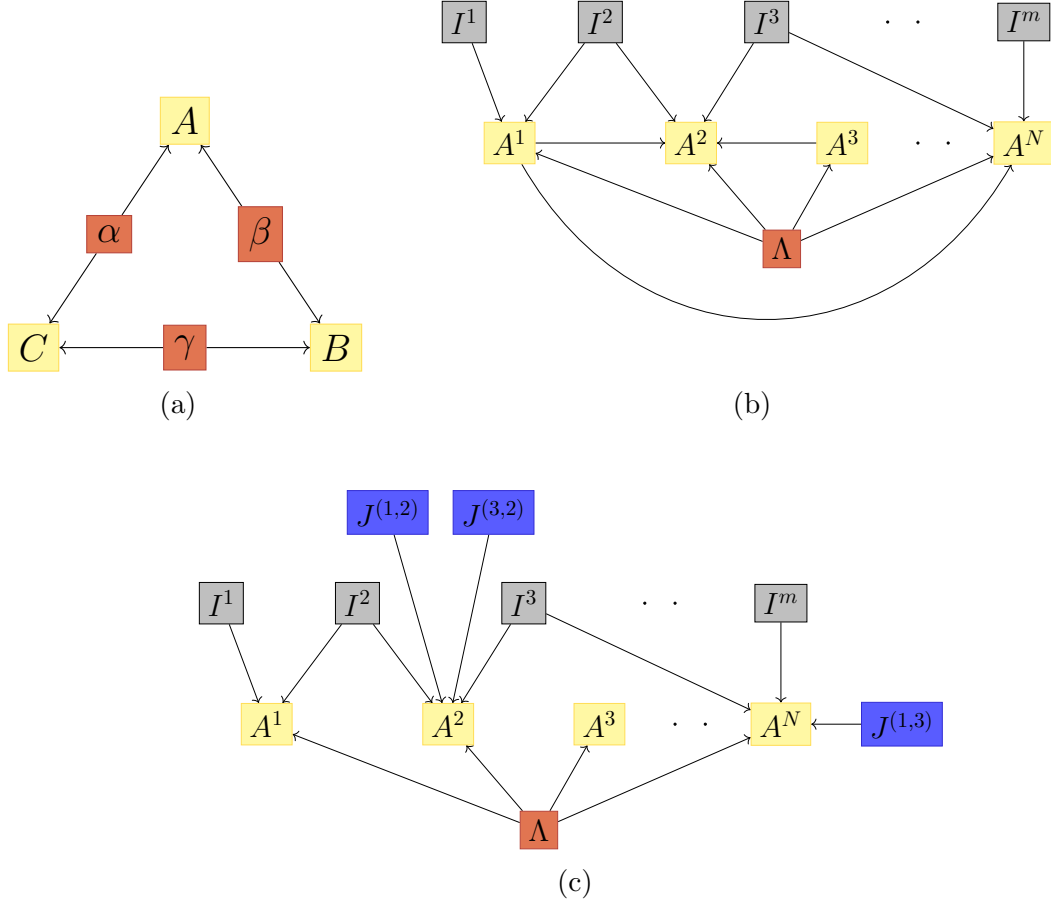


Figure 6.15: (a) The triangle network from [121], which contains three hidden nodes and therefore does not fall into the class of graphs we consider. It hosts quantum correlations when α , β and γ are taken to be quantum sources. (b) A general type of graph $\mathcal{K}$ from the class of graphs which we consider here, which are a subset of those considered in [73]. The restriction of $\mathcal{K}$ to agent (yellow) vertices is referred to as $\mathcal{G}$. (c) The graph $\mathcal{K}'$ obtained from $\mathcal{K}$ by for every edge $A^i \rightarrow A^k$ in $\mathcal{G}$ introducing an input node $J^{(i,k)}$ and edge $J^{(i,k)} \rightarrow A^k$.

The triangle network [121] for example (see Figure 6.15(a)) is not a valid DAG of this type because there is more than once source node. Figure 6.15(b) gives a depiction of the types of graphs we consider: they can be divided up into three layers, layer 1 consisting of the hidden node, layer 2 consisting of the agent nodes, and layer 3 consisting of the input nodes. We require that edges only pass from layer 1 to later 2, layer 3 to layer 2, and within layer 2. We will label the entire DAG $\mathcal{K}$ and the restriction of this DAG to the nodes in layer 2 (i.e. the agent nodes) $\mathcal{G}$.

For such examples, casting them into the causal game framework allows us to show that the Markov condition can be derived by assumption of classicality in terms of the existence of a global section.

For simplicity, we let all random variables A^i and I^i be binary, so that they may

only take on values in the set $\{0, 1\}$. The proof can easily be generalised to variables taking on values in any finite cardinality set. We proceed as follows: from $\mathcal{K}$ we obtain a new graph $\mathcal{K}'$ shown in Figure 6.15(c) by deleting each edge $A^i \rightarrow A^j$ in $\mathcal{K}$ between agents and introducing an input node $J^{(i,k)}$ which sends an input from A^i to A^k . The causal flow in $\mathcal{G}$ will then instead arise from an Experimenter strategy imposed on the agent nodes.

We will first introduce the setup in Figure 6.15(c) and afterwards consider how to implement the causality of $\mathcal{G}$ in subfigure (b), bearing in mind that in the Henson-Lal-Pusey framework an arrow between two nodes simply indicates that the value obtained by the random variable at the source node is sent to the sink node.

Figure 6.15(c) depicts a flat Bell scenario in which an agent A^i may have input from multiple nodes. We therefore take

$$V_{A^i} = \prod_{i \in \text{pa}_{\mathcal{K}'}(A^i)} \{0, 1\}. \quad (6.49)$$

By convention let the first values of tuples in V_{A^i} correspond to nodes $J^{(k,i)} \in \text{pa}_{\mathcal{K}'}(A^i)$. The setup therefore is a standard Bell setup except that an Experimenter cell A^i has as its possible values binary tuples of length $|\text{pa}(A^i)|$. A measurement which A^i performs is denoted $a_v^{(i)}$ where $v \in V_{A^i}$, since we would like a unique measurement to be associated with each value with which the experimenter can fill their cell. Finally, we take the standard Bell cover—games in the cover have a single input for each agent, or a subset of agents.

If noncontextual the setup has a global section $h \in \mathcal{D}_R \mathcal{S}^N(M)$ and obeys the factorisation⁴

$$e_{G_i}(\sigma) = \sum_{g \in \mathcal{S}^N(M)} h(g) \prod_{m \in \text{dom}(\Gamma_{G_i}^N)} \delta_{\sigma \cap \Delta_m}(g \cap \Delta_m) \quad (6.50)$$

for any $G_i \in \mathcal{C}$ (the standard Bell cover), and where $\sigma \in \mathcal{S}^N(G_i)$ and $\Delta_m := \{\langle m, o \rangle \mid o \in V_m\}$.

We now consider an Experimenter strategy which implements adaptivity according to $\mathcal{G}$ —the Experimenter strategy enforces that a measurement $a_v^{(i)}$ only occur if parents A^j of A^i in $\mathcal{G}$ obtain outcomes specified by v .

Take the strategy $\tau(\mathcal{G})$ which for any $A \in \Gamma^E$ with $\text{pa}_{\mathcal{G}}(A) = \{A^1, A^2, \dots, A^k\}$

⁴This is essentially a rewording of Proposition 3.1 (page 10) in [6]

satisfies the condition

$$(h \in \tau(\mathcal{G})) \wedge \left(\{ \langle A^1, i_1 \rangle, \langle a_{i_1}^{(1)}, o_1 \rangle, \langle A^2, i_2 \rangle, \langle a_{i_2}^{(2)}, o_2 \rangle, \dots, \langle A^k, i_k \rangle, \langle a_{i_k}^{(k)}, o_k \rangle \} \subset h \right) \implies \langle A, o \rangle \in \tau(\mathcal{G}) \wedge o[i] = o_i, \quad (6.51)$$

so that the i th element of the tuple of inputs for agent A corresponds to the output of its i th parent in $\mathcal{G}$.

Now consider the observed $\tau(\mathcal{G})$ -conditional empirical model when the Experimenters are arranged in the DAG $\mathcal{G}$. It is only non-zero on Nature strategies σ for which $p_{G_i}^{\tau(\mathcal{G})}(\sigma) = 1$, or

$$\langle \sigma | (\tau(\mathcal{G})|_{G_i}) \rangle \neq \emptyset. \quad (6.52)$$

Since a Nature strategy σ on G_i generally, for N agents, has maximal element

$$\left\{ \langle A^1, i_1 \rangle, \langle a_{i_1}^{(1)}, o_1 \rangle, \langle A^2, i_2 \rangle, \langle a_{i_2}^{(2)}, o_2 \rangle, \dots, \langle A^N, i_N \rangle, \langle a_{i_N}^{(N)}, o_N \rangle \right\} \quad (6.53)$$

those which satisfy equation [6.52](#) will have $i_k[1, l] = \prod_{A^j \in \text{pa}(A^k)} o_j$, where $l = |\text{pa}(A^k)|$, so that the first l entries of its inputs correspond to the outputs of its parents in $\mathcal{G}$ (other inputs may be ‘free’).

Proposition 6.6.4. *The τ -conditional empirical model e^τ where $e \in \text{EM}(\mathcal{M})$ is an empirical model on the flat Bell scenario $\mathcal{M} = \langle M, \mathcal{C} \rangle$ and $\tau(\mathcal{G})$ is an adaptive strategy according to $\mathcal{G}$, is contextual if and only if the HLP scenario is Markovian with respect to $\mathcal{K}$.*

Proof. Recall what it means for a probability distribution P on a set of random variables $X_i, i \in [1, n]$ to be Markovian with respect to a graph $\mathcal{G}$. It means that P factorises as

$$P(x^{(1)}, \dots, x^{(n)}) = \prod_i P(x^{(i)} | \text{pa}(x^{(i)})). \quad (6.54)$$

Now consider a DAG $\mathcal{G}$ of the type we consider. Thus nodes can be partitioned into sets A, I and $\{\Lambda\}$ consisting of Agent nodes A , input nodes I and the single hidden node Λ . We would like to instead consider the factorisation condition,

$$P(a^{(1)}, \dots, a^{(N)} | i^{(1)} \dots i^{(m)}) = \sum_{\lambda} P(\lambda) \prod_{j=1}^N P(a^{(j)} | \text{pa}(a^{(j)})). \quad (6.55)$$

Note that equation [6.54](#) holds if and only if equation [6.55](#) holds. This follows from the definition of conditional probability: rewriting equation [6.54](#) in terms of the

partition of variables gives

$$P(a^{(1)}, \dots, a^{(N)}, i^{(1)}, \dots, i^{(m)}, \lambda) = P(\lambda) \prod_{j=1}^N P(a^{(j)} | \text{pa}(a^{(j)})) \prod_{j=1}^m P(i^{(j)} | \text{pa}(i^{(j)})). \quad (6.56)$$

equation [6.55](#) then follows from the definition of conditional probability, where the term $\prod_{j=1}^m P(i^{(j)} | \text{pa}(i^{(j)}))$ cancels out. Moreover, the sum is taken over all λ since this is a hidden variable. So it remains to prove equation [6.55](#) holds if and only if the τ -conditional empirical model on $\mathcal{G}$ is noncontextual. In this expression $\text{pa}(a^{(j)})$ may refer to either other agent nodes or input nodes. Introduce the associated variable $a_{\text{pa}(a^{(j)})}^{(j)}$ —here, $a_{\text{pa}(a^{(j)})}^{(j)}$ is the value of the random variable when the value of parent nodes is known. Thus $P\left(a_{\text{pa}(a^{(j)})}^{(j)}\right) = P(a^{(j)} | \text{pa}(a^{(j)}))$. equation [6.55](#) becomes

$$P\left(a_{\text{pa}(a^{(1)})}^{(1)}, \dots, a_{\text{pa}(a^{(N)})}^{(N)}\right) = \sum_{\lambda} P(\lambda) \prod_{j=1}^N P\left(a_{\text{pa}(a^{(j)})}^{(j)}\right). \quad (6.57)$$

□

6.7 Discussion

This chapter considered the problem of capturing causal notions of nonclassicality with a sheaf theoretic approach. In doing so, and in separating the causality which Nature can use and that which it cannot and is imposed at the Experimenter level, we were able to differentiate between the setups considered by Henson Lal and Pusey, and those considered by Gogioso and Pinzani. Certain setups of the former arise as Experimenter strategies over flat Bell scenarios, while the latter allow Nature, or the hidden variable, to update measurement outcomes in between measurements.

Both have an effect on the observed contextuality. In HLP setups, as was the case for the instrumental setup, experimenter strategies reduce the observed local strategies, in a sense reducing the likelihood that contextuality is present in the observed empirical data. For GP scenarios, allowing hidden variables to store information from the past makes such a model significantly more powerful, and therefore increases the number of empirical models which can be described via a deterministic hidden variable theory. As we saw in the previous chapter, any temporal Bell scenario for which agents occur in a total order cannot be contextual (or nonclassical).

In future, it would be instructive to study whether the framework holds up for general HLP scenarios—the answer at first sight seems to be that it does not; in

the sheaf approach a hidden variable arises implicitly as a global section, and introducing multiple independent hidden variables, or source nodes, requires imposing some structure on global sections which respects the presence of these independent sources. This essentially arises as factorisation constraints on a global distribution, and becomes difficult to write down for the general case. However, it might be possible to incorporate the hidden nodes, or sources, explicitly as cells.

Further, an outstanding problem is to understand the structure of the cover $\mathcal{C}$, and its relations to abstract cell complexes which are defined in the literature. The possibility of extension of Vorob'ev's theorem to more general complexes, rather than just ASCs, would be another possible direction, and the literature on cellular sheaves [46] could prove a good starting point. This could extend structural results beyond those of Henson, Lal and Pusey, which can identify graphs for which $\mathbf{C} \neq \mathbf{NS}$.

Chapter 7

Spin Hamiltonians on graphs

— Summary

This chapter studies spin systems on random graphs in the limit that the number of spins tends to infinity, which in physics literature is known as ‘the thermodynamic limit’. This is done using the theory of graphons, or graph limits. We give a brief background on solving spins systems on different geometries, and build up some intuition for the main results by considering the classical Ising model on an Erdos Renyi graph. We then cover the two main results of the chapter. The first of these is Theorem 7.4.2, which states that, with high probability, the free energy of a spin Hamiltonian on a random graph converges in the thermodynamic limit to the free energy of the same Hamiltonian on a weighted, deterministic graph (which we define). The second result is that the thermodynamic limit can, as a result of Theorem 7.4.2, be found using mean-field theory. In this way it is shown that in this limit the observables of the system can be calculated by solving three coupled integral equations, which we give explicitly. The material contained in this Chapter appears in [133] and [128].

Many-body Hamiltonians, which describe the interaction of a large number of particles, are of high interest for at least two reasons. On the one hand, they offer a setting in which to study phenomena, both classical and quantum, which might occur in real systems, such phenomena being diverse and including magnetic effects [126], high levels of entanglement as displayed by the SYK model [86, 123],

which has facilitated the study of quantum gravity, and even early universe phase transitions [68]. In this case, usually, one studies how the ground state or equilibrium state of the system changes as some parameter in the Hamiltonian is changed. In ‘critical’ parameter regimes, which are parameter regimes in which the ground state undergoes abrupt changes, the system can exhibit interesting features which in the case of quantum systems usually entails high levels of entanglement, and as witnessed via diverging entanglement entropy near the critical point.

The other motivation is a consequence of the inherent complexity of these many-body Hamiltonians, due to the exponential increase in Hilbert space dimension which occurs with only a linear increase in particle number. They therefore offer a natural setting to benchmark classical and quantum techniques of solving problems—which are in a broad sense deemed difficult—whereby these problems are encoded as the ground state of the Hamiltonian. Demonstrations of ‘quantum advantage’ often employ such an approach, such as IBM’s kicked Ising experiment which simulated the kicked Ising quantum system on the heavy hexagon lattice using a quantum processor of 127 qubits [85], and which was hastily followed by an efficient simulation of the setup using classical methods [133].

A more—in some sense—direct, although not always straightforward approach, is a complexity class approach, which has lead to the field of Quantum Hamiltonian Complexity (QHC) [36]. The goal of QHC is to quantify the hardness of solving Hamiltonians (with certain locality constraints) by establishing connections to computational complexity theory. Notably, the problem of determining the ground state of a k –local Hamiltonian, or one in which any interaction term in the Hamiltonian occurs between at most k particles, is known to be **QMA-complete** [82], the quantum version of **NP-complete**. Here, **NP** refers to those problems which are verifiable in polynomial time, and the appended ‘complete’ refers to the fact that it is in **NP** and it is at least as hard as every problem in **NP** [44, 12]. We will demonstrate how complexity results arise naturally in the Ising model, but will not otherwise consider this approach in the work that follows. Nevertheless, it is interesting, and useful to bear in mind, that the choice of interaction pattern, which for 2–local Hamiltonians is simply a graph G —in which edges dictate pairwise interactions in the Hamiltonian—will affect the hardness of deriving analytical results to computational problems. We show that in the thermodynamic limit, which is when the number of particles approaches infinity, and in the case that G is a random graph, computing the ground state is equivalent to computing the ground state of the same Hamiltonian on a (non-random) weighted graph $\bar{G}$. Sometimes, this latter computation will be straightforward. In terms of complexity class considerations, such a

result puts the problem of finding the ground state of certain 2-local Hamiltonians in the thermodynamic limit in P, which is a result already known [26].

Our focus is on a Hamiltonian H of the form

$$H = \sum_{(i,j) \in E(G)} h_{ij} + \sum_{i \in V(G)} h_i \quad (7.1)$$

for a simple graph G with n vertices. We in particular study spin systems, so that at each vertex of G lies a d -dimensional spin system, and h_{ij} is an interaction term for the spins at vertex i and vertex j .

Methods for solving systems of this type, both exactly and approximately, are numerous and depend on the specifics of G and the choice of h_{ij} and h_i . Notably, for certain one dimensional chains, the Bethe ansatz [23], and Baxter’s generalisation thereof [18], can be used to find the exact eigenspectrum of certain Hamiltonians. The Jordan Wigner transformation, which maps spin operators to fermion creation and annihilation operators in such a way that the new Hamiltonian is easier to solve, has also been used to analytically solve certain cases [98, 116]. Spin wave theory and mean field theory are useful approximate methods, both of which consider fluctuations about some classical or mean field solution [13]. In terms of numerical approaches, the Ising model can be to a high degree of accuracy solved using Markov Chain Monte Carlo (MCMC) simulation, and in particular using the Metropolis-Hastings algorithm [109, 72]. For more complex systems, tensor network simulations have been hugely successful [147]; this class of techniques works by encoding the wavefunction as the tensor contraction of a network of tensors, this representation depending on the interaction structure of the Hamiltonian. A class of tensor networks states are the matrix product states (MPS) [127] which are able to describe states with area law entanglement (in contrast to the volume law entanglement that is in principle valid for states occupying the full Hilbert space) [51]. These states have been successful for a large class of problems, and in particular for one-dimensional spin chains, where local interactions ensure that entanglement between subparts is reduced. The density matrix renormalization group (DMRG) [139] then allows for optimisation of the MPS such that the resulting state is the ground state of a given Hamiltonian.

Here, we utilise mean field theory to study the thermodynamic limit of H with G a dense random graph. Our analysis is exact in this limit, and gives a good approximation for large n . This result, to our knowledge, is the first of its kind concerning the thermodynamic limit of spin systems on dense random graphs. We supplement this with tensor network calculations, and are able to demonstrate that

these calculations converge to the mean-field results in the limit $n \rightarrow \infty$, but are also able to capture the quantum fluctuations inherent to the finite n solution and go beyond mean-field in such parameter regimes. Central to our study will be the graphon, a continuous object which captures the statistics of large random graphs—or more precisely is the limit object of a sequence of random graphs—and introduced by Lovász in [102], to form a coupled set of integral equations whose solution yield expressions for global observables of the system.

We begin with an introductory section on the background of spin systems and known exact and numerical methods for particular and general geometries, followed by a section considering our results in the specific case of Erdős-Rényi random graphs, and finally a section on the limit of spin systems on random graphs.

7.1 Spins on graphs

Consider a graph $G = (V, E)$ with vertex set V and edge set $E \subseteq V \times V$ with corresponding cardinalities $|V| = n$ and $|E| = n_E$. We consider simple graphs meaning (i) edges in E cannot be of the form (v, v) , or we require no *self-edges* and (ii) If E contains an edge (v, w) then there can be no other edge between nodes v and w , or there can be no *multi-edges*. We associate with a graph G an n by n matrix $\mathbf{A}$ such that $A_{v,w} = 1$ iff $(v, w) \in E$ and is zero otherwise. The matrix $\mathbf{A}$ is known as the adjacency matrix of the graph.

To each vertex $v \in V$ we associate a $(2s + 1)$ –dimensional Hilbert space $\mathcal{H}_v = \mathbb{C}^{(2s+1)}$. Each of these will be the Hilbert space of a spin s particle, so that a basis for the Hilbert space is defined by taking the spin s representation of $\text{SU}(2)$, or the special unitary group of dimension 2. We take $\{S_v^x, S_v^y, S_v^z\}$ which obey the commutation relation

$$[S_v^a, S_v^b] = i\epsilon_{abc}S_v^c, \quad a, b, c \in \{x, y, z\}. \quad (7.2)$$

i.e. that of the Lie bracket for the Lie group $\mathfrak{su}(2)$. Here, ϵ_{abc} is the Levi-Civita symbol. An orthogonal basis for $\mathcal{H}_v$ can be taken to be the basis that diagonalises one of the S_v^i matrices (or indeed, any observable).

In the two-dimensional case (so that $s = \frac{1}{2}$) the spin observable is characterised by the matrices $\sigma_v^x, \sigma_v^y, \sigma_v^z$ where σ_v^α is the Pauli matrix σ^α on site v and we recall

that the Pauli matrices are defined

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (7.3)$$

The basis is taken to be $\{|\downarrow\rangle_v, |\uparrow\rangle_v\}$ where $|\downarrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|\uparrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, so that the basis diagonalises σ_v^z . We say $|\uparrow\rangle$ is *spin up* and $|\downarrow\rangle$ is *spin down*—although this only holds in the z basis. From now on we will instead write $|\uparrow\rangle_v = |0\rangle_v$ and $|\downarrow\rangle_v = |1\rangle_v$, according to convention, for the basis on the Hilbert space at site i . We write

$$\langle i|j\rangle_v = \delta_{ij} \quad i, j \in \{0, 1\} \quad (7.4)$$

to denote the inner product.

The Hilbert space of the total system on a graph G will be a tensor product of the individual ones,

$$\mathcal{H}_G = \bigotimes_{v \in V} \mathcal{H}_v = (\mathbb{C}^{(2s+1)})^{\otimes |V|}, \quad (7.5)$$

and this Hilbert space is spanned by orthonormal elements of the form $\bigotimes_{v \in V} |l_v\rangle_v$, with $l_v \in \{0, 1\}$. Their inner product is

$$\bigotimes_{v \in V} \langle l_v|_v \cdot \bigotimes_{v \in V} |r_v\rangle_v = \prod_{v \in V} \langle l_v|r_v\rangle_v. \quad (7.6)$$

We write an element of this basis as $|\sigma\rangle$ where σ is a string of length $n = |V(G)|$ of 0's and 1's. There are therefore 2^n such basis states, and we write this family as $\{|\sigma\rangle\}$.

Any element $|\psi\rangle$ of the Hilbert space $\mathcal{H}_G$ can be expanded as a linear combination of the family of orthonormal basis states $\{|\sigma\rangle\}$,

$$|\psi\rangle = \sum_{\sigma} \alpha_{\sigma} |\sigma\rangle. \quad (7.7)$$

A quantum state can equivalently, rather than its representation above as a vector in a Hilbert space, be expressed as a density operator, which is an operator on the Hilbert space and therefore a matrix. Given a vector state $|\psi\rangle$, the associated

density operator is defined

$$\rho = |\psi\rangle\langle\psi|, \quad (7.8)$$

and, importantly, certain states expressed in terms of density operators have no associated vector representation. These are the mixed states, which have decomposition as a probabilistic mix of vector states. A state which is not mixed is said to be a pure state.

An important property of a pure state is its **bipartite entanglement entropy** [112], which defines a measure of entanglement across a choice of partition of the system. This is done by measuring the uncertainty in the statistics of the reduced density operator corresponding to one side of the partition, but, importantly, the quantum state in this definition is assumed to be pure and not mixed. We will consider the Von Neumann entropy, defined as follows. Taking the first k particles to be in partition A and the remaining $n - k$ particles to be in partition B , the entanglement entropy is the Von Neumann entropy of the reduced state on either of the partitions,

$$S(\rho) = -\text{Tr}_A [\rho_A \ln(\rho_A)], \quad (7.9)$$

where $\rho_A = \text{Tr}_B [\rho]$ is the reduced density operator on the A partition. The interpretation is as follows: if there is a large amount of entanglement between the partitions A and B , then the reduced state ρ_A will be proportional to the identity matrix, which maximises equation 7.9.

7.1.1 Operators

A Hamiltonian H is a Hermitian operator $H : \mathcal{H}_G \rightarrow \mathcal{H}_G$ which send any $|\phi\rangle$ to $H|\phi\rangle \in \mathcal{H}_G$. For a system composed of n spins which are d dimensional, the Hamiltonian will be a square matrix of dimension $d^n \times d^n$. The eigenspectrum of H is of special interest because it gives the energy spectrum of the system. Therefore, labelling an eigenstate $|i\rangle$ and calling the corresponding energy E_i , it holds that $H|i\rangle = E_i|i\rangle$. The spectrum of H is the collection $\{E_i\}$ of energy eigenvalues. A system at zero temperature will favour the lowest energy state, which is called the ground state and written $|0\rangle$ (note that this is a different state from the $|0\rangle$ used to denote $|\uparrow\rangle$).

At a non-zero temperature T the system will not be in the ground state, but in a thermal equilibrium state in which each state $|i\rangle$ has probability p_i of being

occupied, this probability being temperature dependent. It can be shown that

$$p_i = \frac{\exp(-\beta E_i)}{Z} \quad (7.10)$$

where $\beta = 1/T$ (setting $k_B = 1$) and Z is a normalisation constant called the partition function and has expression $Z = \sum_i \exp(-\beta E_i)$.

Equivalently, the state of the system at thermal equilibrium is the ‘Gibbs state’ which is written

$$\rho = \frac{\exp(-\beta H)}{Z}. \quad (7.11)$$

The Gibbs state ρ is the unique density matrix that minimises the free energy functional, where the Helmholtz free energy density of an n spin Hamiltonian H is defined

$$f(H) := \frac{-1}{\beta n} \ln (\text{Tr}(\exp(-\beta H))), \quad (7.12)$$

and determines many important macroscopic observables of the system— therefore if $f(H_1) = f(H_2)$ for two systems described by Hamiltonians H_1 and H_2 , these observables on these systems, when in thermal equilibrium, agree.

The Hamiltonian is an example of a linear operator on $\mathcal{H}_G$ i.e. maps of the form $O : \mathcal{H}_G \rightarrow \mathcal{H}_G$, and we will denote the set of such operators $L(\mathcal{H}_G)$. Expectation values of such operators give the average value that is observed on many identically prepared systems; if the system is in state $|\psi\rangle$ the expectation is calculated as

$$\langle O \rangle := \langle \psi | O | \psi \rangle, \quad (7.13)$$

or, for a density operator ρ ,

$$\langle O \rangle = \text{Tr}(\rho O). \quad (7.14)$$

If the operator is an observable of the system, we require moreover that it is hermitian, so that expectation values are always real.

7.2 Phase transitions

A phase transition is the abrupt change in the state of the system for an infinitesimal change in a governing parameter. If continuous or otherwise called second order, it can be characterised by an order parameter, a quantity that is zero in one phase (disordered) but has a finite value in another phase (ordered). The choice of order

parameter depends on the system at hand, and oftentimes there is more than one order parameter of interest.

Quantum phase transitions are a special class which occur at zero temperature, $T = 0$, and by varying a non-thermal parameter such as an external field. The idea is that, as this parameter is varied, the spin configuration σ which is energetically favourable, or superposition of such configurations, changes, and these changes might be very abrupt and occur together with high levels of correlation in the system.

Here we are concerned with phase transitions in quantum and classical spin systems, which in certain phases display magnetic properties but can also display high levels of entanglement. The classical case is well understood: for large temperature T the system is disordered, and any given spin has equal probability of pointing up or down. This is referred to as paramagnetism. As this decreases, spins favour a lower energy state which is to align with their neighbours. At some critical temperature T_C , the material becomes ferromagnetic, and spins spontaneously choose one direction to align: either *up* or *down*. At $T = 0$ the system is in the ground state.

In terms of order parameter for spin systems, a natural candidate is the average magnetisation per site for a choice of axis $\alpha \in \{x, y, z\}$ which has expression

$$M^\alpha := \frac{1}{n} \sum_{i=1}^n \langle \sigma_i^\alpha \rangle. \quad (7.15)$$

For the classical spin system discussed above, M^z is 0 in the paramagnetic phase and 1 or -1 in the ferromagnetic phase.

Other types of magnetic behaviour can be exhibited, such as the anti-alignment of spins, or antiferromagnetism. Consider a Hamiltonian of the form

$$H = \sum_{(i,j) \in E(G)} J \sigma_i^z \sigma_j^z \quad (7.16)$$

with $J > 0$. Then the ground state must be one in which as many neighbouring spins in G as possible are anti-aligned, since $J \sigma_i^z \sigma_j^z |\uparrow\rangle_i |\uparrow\rangle_j > 0$ while $J \sigma_i^z \sigma_j^z |\uparrow\rangle_i |\downarrow\rangle_j < 0$. Again considering the change in phase as the temperature is varied, a good order parameter $\square$ is a measure of antiferromagnetism in the system,

$$C_{\text{AFM}} = \frac{1}{n_E} \sum_{(i,j) \in E(G)} \langle \sigma_i^z \sigma_j^z \rangle \quad (7.17)$$

¹note that the order parameter M^z is not useful as it is zero in the antiferromagnetic state

which is -1 in the antiferromagnetic state and 0 in the paramagnetic state.

Order parameters are examples of critical exponents, which are parameters describing the behaviour of systems near critical points (but need not measure order versus disorder). A common example are correlation functions, which measure to what extent the wavefunctions at two points i and j , which in our case are taken to be nodes of the graph G , are correlated. The two point spin correlator between site i and j is

$$C_{i,j}^\alpha = \langle \sigma_i^\alpha \sigma_j^\alpha \rangle. \quad (7.18)$$

One is usually interested in the scaling of these parameters, such as whether this falls off exponentially, logarithmically, or according to some power law as the distance between spins i and j is increased. Since we work with spins on a graph rather than lattice, distance is in terms of the minimum path length between i and j in G .

Having covered the relevant background material, we now move on to specific instances of the general spin Hamiltonian, and briefly survey some known results. This survey will be largely incomplete, and the interested reader should consult the much more complete bodies of work [\[48\]](#), [\[126\]](#).

7.3 Spin Hamiltonians

We consider a general spin Hamiltonian of the form

$$H(G) = \sum_{(v,w) \in E} h_{v,w} + \sum_{v=1}^n h_v, \quad (7.19)$$

with $h_{v,w}$ and h_v being, respectively, one and two-body operators acting on the subscripted vertices. We will focus on the choice of a spin $1/2$ system with

$$\begin{aligned} h_{v,w} &= J_x \sigma_v^x \sigma_w^x + J_y \sigma_v^y \sigma_w^y + J_z \sigma_v^z \sigma_w^z, & J_x, J_y, J_z &\in \mathbb{R} \\ h_v &= H_x \sigma_v^x + H_y \sigma_v^y + H_z \sigma_v^z, & H_x, H_y, H_z &\in \mathbb{R}. \end{aligned} \quad (7.20)$$

The Hamiltonian $H(G)$ describes spins on a graph G interacting according to a general interaction term $h_{v,w}$.

For the special case of a 1-dimensional (1D) chain, this Hamiltonian can be solved exactly for certain parameter regimes. More generally, this cannot be exactly solved, especially for the case of D -dimensional lattices for $D > 1$, although there

is an enormous and growing literature on approximate methods. We consider some specific examples.

7.3.1 The classical Ising model and Max-Cut

The Ising Hamiltonian is a classical Hamiltonian which can be shown to be universal. What this means is that any other classical spin Hamiltonian has spectrum which can be reproduced in the low-lying spectrum of the Ising model [45].

The Ising model is defined by the Hamiltonian

$$H_{\text{Ising}} = \sum_{i,j} J_{i,j} s_i s_j + \sum_i h_i s_i. \quad (7.21)$$

Here, we work with the n Boolean variables $s_i \in \{0, 1\}$ rather than with the Pauli matrices, since a two-dimensional classical spin system is a bit. In the case that the inter-spin coupling $J_{i,j} = J$ is independent of spins, the Hamiltonian of equation [7.21] is a special instance of the general Hamiltonian in equation [7.19] where $J_z = J$ and $H_z = h_i$ and all other parameters are zero. In this subspace of the Hilbert space, all eigenvalues are in the z basis, which allows us instead to work with the s_i variables.

Solving the Ising model means finding a spin configuration s [2] which minimises H_{Ising} . Ising himself solved the model on a one-dimensional lattice, showing that there is no phase transition [77]. Onsager produced the two-dimensional solution, and with a lengthy and complex calculation showed that there is a phase transition in two dimensions [113]. In higher dimensions, one typically uses numerical approaches such as Monte Carlo methods, while various approximations can also be used. One type of approximation is mean-field theory, which in parameter regimes where correlations are weak can work well, but breaks down near the critical point. Mean-field theory is also useful when the graph at hand is dense, so that each particle is bonded to of order n other particles—it is known that in this case product state descriptions are valid [25].

There is an interesting connection between the Ising model and the field of computational complexity; specifically, the ground state of the Ising model on a graph G without external fields is the solution to the weighted Max-Cut problem, which is NP-hard. The Max-Cut problem is stated

²We use s rather than σ to differentiate between the quantum and classical cases.

Given a weighted graph G , find a partition of the vertices into sets S and T such that the sum of the weights of the edges between S and T is at a maximum.

The Ising model with zero field strength on a graph G is written

$$H_{\text{Ising}} = \sum_{(i,j) \in E(G)} J_{i,j} s_i s_j. \quad (7.22)$$

Studying the Hamiltonian above, it becomes apparent that in order to minimise the action of the Hamiltonian on a state s , s should maximise the sum of the weights of neighbouring spins s_i and s_j with $(i,j) \in E(G)$ pointing in opposite directions. Thus solving the Ising model on the weighted graph G is equivalent to finding the **Max-Cut** of G .

Now consider the unweighed case in which

$$H_{\text{Ising}} = \sum_{(i,j) \in E(G)} s_i s_j, \quad (7.23)$$

and take G to be an Erdős-Rényi graph in which any two vertices $i, j \in V(G)$ with $i \neq j$ are connected with probability p . The Hamiltonian can therefore be expressed

$$H_{\text{Ising}} = \sum_{i,j \in V(G); i < j} P_{i,j} s_i s_j, \quad (7.24)$$

where $P_{i,j}$ is a random variable taking on the value 1 with probability p and 0 with probability $1 - p$. This is different from the variable $J_{i,j}$ in equation [7.22](#) which is not probabilistic and fixed at a single value for each choice of i and j . Nevertheless, it is tempting to replace $P_{i,j}$ with its expected value which is p , in which case we are studying the weighted case in which every weight is constant and equal to p , or $J_{i,j} = p$. The result we wish to show is that in the limit of diverging vertex number $n \rightarrow \infty$ for the graph G , the free energies of the ground state of the Hamiltonian in equation [7.22](#) with $J_{i,j} = p$ is identical to the Hamiltonian in equation [7.24](#). In fact the result will be somewhat more general than this, but it is useful to bear this example in mind.

7.3.2 The XXZ model

We now consider a version of equation [7.19](#) which contains noncommuting terms and can therefore host quantum phenomena, such as entanglement. It is known as the XXZ Heisenberg model, and we define $J_x = -J$, $J_y = -J$, $J_z = \Delta$, $H_z = H_x = H_y = 0$ so that the Hamiltonian reads

$$H(G) = \sum_{(v,w) \in E(G)} -J(\sigma_v^x \sigma_w^x + \sigma_v^y \sigma_w^y) + \Delta \sigma_v^z \sigma_w^z, \quad (7.25)$$

where we take $J > 0$, and from now on simply set it to $J = 1$.

We will first summarize results for the one-dimensional model, since this is the most well studied example, for different parameter regimes of Δ . In this case, G is defined by labelling vertices 1 to n and requiring $(i, i+1) \in E(G)$ where $i = j \bmod n$ for $j \in \mathbb{N}$. The result is a one-dimensional lattice with periodic boundary conditions.

Ferromagnetic Phase– For $\Delta < -1$, the XXZ chain is in a ferromagnetic phase, with spins aligning in either a positive or negative z direction. There is therefore a single ground state with no entanglement present, and the total magnetisation is given by

$$M^Z = \pm \frac{1}{2}n. \quad (7.26)$$

Néel Phase– For $\Delta > +1$ the XXZ chain is in the antiferromagnetic Ising or Néel phase, where spins are anti-aligned. Note that there are two (degenerate) ground states which satisfy this property.

XY Phase– The phase occurring in the parameter regime $-1 < \Delta < 1$ is more interesting: the excitation spectrum is gapless and correlation functions show power-law behaviour. The XY phase is critical. The transition point $\Delta = 1$ is a Berezinskii–Kosterlitz–Thouless or BKT transition. On the other hand the point $\Delta = -1$ corresponds to a (discontinuous) first order transition between the ferromagnetic and XY phases.

7.3.3 The transverse field Ising model

The transverse field Ising model is another example of a quantum model. In addition to the terms present in the usual classical Ising model, it has a ‘transverse field’ term, which does not commute with the interaction term. We define $J_x = J_y = 0$,

$J_z = -J$, $H_z = H_y = 0$ and $H_x = -h$. The Hamiltonian therefore reads

$$H_{\text{TFI}} = \sum_{(v,v') \in E} -J(\sigma_v^z \sigma_{v'}^z) - \sum_{v \in V} h \sigma_v^x. \quad (7.27)$$

When $|h| < 1$, for $J > 0$ the model is ferromagnetic and for $J < 0$ is antiferromagnetic. This is the ordered phase. On the other hand, when $|h| > 1$, spins instead align (or anti-align) with the field and, in the z direction, is disordered. At $|h| = 1$, the system undergoes a quantum phase transition and exists in a gapless phase. This is a special case of the XY Heisenberg model, another model which is often considered in the literature.

7.3.4 Results concerning non-lattice geometries

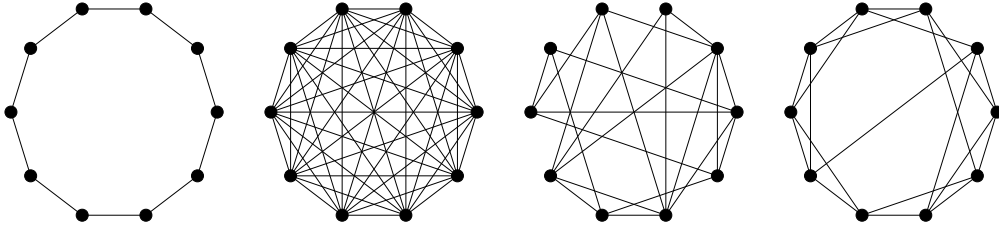


Figure 7.1: Examples of finite simple graphs. From left to right, *cycle graph*, C_n , *Complete graph* K_n , *Erdős-Rényi graph* $G(n,p)$, *Small world graph* $G(n,m,p)$.

So far, with the exception of the Ising model, we have discussed only results on lattices, and only on one-dimensional lattices at that. However, the physical properties of interacting systems are strongly dependent upon which components interact, so that it is important, and informative, to study these Hamiltonians on more general interaction patterns. Doing so has led to numerous interesting results: frustration causes the manifestation of exotic phases of matter such as spin liquids [132], and small-world effects alter the underlying universality class of ordered-disordered phase transitions [38].

In terms of physical realisations of such geometries, recent experimental advances have made it possible to go beyond lattice geometries for a range of platforms, including atoms trapped in cavities [40], photonic waveguides [76], Moiré Heterostructures [84], trapped ions [78] and superconducting circuits [140]. In particular, proposals now exist to use trapped ion arrays to engineer many body spin Hamiltonians defined over arbitrary graphs [93, 110].

From a theoretical point of view, however, the difficulty of solving the many-body problem (especially in the quantum regime), limits a general characterization of how network topology influences strongly correlated systems. Nevertheless, some special cases of more diverse geometries have been investigated, which we briefly discuss below.

Complete and Erdős-Rényi graphs A complete graph is a graph in which every pair of vertices shares an edge, and the complete graph on n vertices is denoted K_n . Both the Ising model and XY models have been studied in some detail on complete graphs, and are referred to as the Curie-Weiss [92, 52] and Lipkin-Meshkov-Glick (LMG) [99, 108, 64] models respectively. The high symmetry in such systems—namely, that they are permutationally invariant under the exchange of any spin—often makes calculations of ground states and studying phase transitions much simpler.

The Curie-Weiss model is exactly solvable and, like the one dimensional model, undergoes a phase transition from an ordered (AFM or FM) state to a disordered one at some critical temperature T_C [92, 52]. The model has also been studied with random-couplings in [24], which is equivalent to taking the graph G to be an Erdős-Rényi graph. They show that the free energy in the thermodynamic limit of this model approaches the that of the standard Curie Weiss model; such a result is a specific case of the general result we will prove in this Chapter.

The LMG model [99, 108, 64], like the XY Ising model in one dimension, for positive h undergoes a phase transition either from a ferromagnetic or antiferromagnetic (ordered) state, or from a critical XY state for $\gamma = 0$, to a paramagnetic state as the field is increased. It is known that the entanglement entropy is singular at the critical point, and diverges logarithmically near this point [94]. As noted by the authors in [94], “Results emerging from our analysis are surprisingly similar to those found for the one- dimensional XY chain”. The thermodynamic behaviour has been explicitly studied in [122].

Power Law Decaying Models Another type of geometry which has gained some interest is that in which spin interactions are beyond the nearest neighbours of the spin (as would be the case for lattices), but decay with some power law $1/r^\alpha$, where r measures the distance between spins and α controls the rate of decay of such interactions. In [57] the XXZ model is studied with power law decaying interactions on lattices of dimension d , and three regimes are found (i) For $\alpha > d + \sigma_c$ (σ_c is a phase dependent parameter) the system has properties similar to that of the $\alpha = 0$ short range limit (ii) For $\alpha < d$ the system reproduces the properties of the infinite-range limit $\alpha = 0$ in the thermodynamic limit. (iii) in between these values of α , the

system behaves differently from the $\alpha = 0$ and $\alpha = \infty$ limits in the thermodynamic limit (as opposed to the other regimes of α).

7.4 Thermodynamic limit of spin systems on dense random graphs

We are interested in to what extent properties of the graph G determine the spectrum of the Hamiltonian $H(G)$ given in equation [7.19](#). Random graphs are of particular interest because (i) they represent a large subset of graphs and (ii) One hopes that the thermodynamic limit, or the limit in which the number of spins $n \rightarrow \infty$, can be studied by utilising the large literature on random graphs limits [102](#).

Thus we study the thermodynamic behaviour of a system on a random graph G of node size n and with Hamiltonian $H(G)$ by studying the sequence of Hamiltonians $(H(G_n))$ and taking $n \rightarrow \infty$. It turns out that a sequence $(H(\bar{G}_n))$ can be defined, where $\bar{G}_n$ is a weighted graph on n nodes, such that the free energies of $H(G_n)$ and $H(\bar{G}_n)$ are identical in the limit $n \rightarrow \infty$.

We begin by presenting the main Theorem for the case in which the sequence (G_n) is a sequence of Erdős-Rényi graphs for fixed edge probability p . This is the material contained in the paper [134](#).

Subsequently, we show how this Theorem can be significantly generalised to any sequence of converging dense random graphs (G_i) , and give the proof. This is done by utilising the graphon, a continuous limit object of a sequence of dense graphs.

7.4.1 Sequences of Erdős-Rényi graphs

We consider an XXZ Hamiltonian on a graph G , now including a scaling factor of $\frac{1}{n}$ which will prevent energy divergences later on.

$$H_{\text{XXZ}}(G) = \frac{1}{n} \sum_{(v,v') \in E} -J(\sigma_v^x \sigma_{v'}^x + \sigma_v^y \sigma_{v'}^y) + \Delta \sigma_v^z \sigma_{v'}^z. \quad (7.28)$$

We restrict our numerics to $\Delta \geq 0$ regime.

An Erdős-Rényi graph on n vertices is constructed by connecting any two vertices v and $v \neq v'$ with probability p . We will denote such a graph as $\mathcal{G}(n, p)$ for notational convenience, but in reality $\mathcal{G}(n, p)$ specifies an ensemble of graphs and our random graph is a draw from this ensemble. An ensemble is a pair $\langle \Omega, P \rangle$, where Ω is the set

of all simple undirected graphs on n vertices, of which there are 2^N with $N = \binom{n}{2}$, and P assigns a probability to each graph in Ω

$$P(G) = p^m(1-p)^{N-m}, \quad (7.29)$$

where m is the number of edges in G . Note that when $p = \frac{1}{2}$ this reduces to

$$P(G) = \left(\frac{1}{2}\right)^N, \quad (7.30)$$

independent of m . Thus any graph is equally likely.

The Hamiltonian from equation [7.28](#) then becomes

$$H(G(n, p)) = \frac{1}{n} \sum_{\substack{v, v'=1 \\ v > v'}} -P_{v, v'} J(\sigma_v^x \sigma_{v'}^x + \sigma_v^y \sigma_{v'}^y) + \Delta \sigma_v^z \sigma_{v'}^z, \quad (7.31)$$

where $P_{v, v'}$ is a random variable which takes value 1 with probability p and 0 otherwise. The difficulty in solving equation [7.31](#) stems from the presence of this random variable $P_{v, v'}$ inside the summation—the Hamiltonian is a sum over two-body terms $h_{v, v'}$ whose presence is determined by $P_{v, v'}$ and is not known a priori. Now, consider instead the graph $\bar{G}(n, p)$ which is a weighted graph on n vertices where every vertex v and $v' \neq v$ are attached with an edge of weight p . It's associated Hamiltonian is

$$H(\bar{G}(n, p)) = \frac{1}{n} \sum_{\substack{v, v'=1 \\ v > v'}} -p J(\sigma_v^x \sigma_{v'}^x + \sigma_v^y \sigma_{v'}^y) + \Delta \sigma_v^z \sigma_{v'}^z, \quad (7.32)$$

where the vertex-pair dependent random variable $p_{v, v'}$ has been replaced with the constant term p . This may be removed from the summation, so that equation [7.32](#) is equivalent to

$$H(\bar{G}(n, p)) = \frac{p}{n} \sum_{\substack{v, v'=1 \\ v > v'}} -J(\sigma_v^x \sigma_{v'}^x + \sigma_v^y \sigma_{v'}^y) + \Delta \sigma_v^z \sigma_{v'}^z. \quad (7.33)$$

This new Hamiltonian is significantly easier to solve, because it commutes with the total spin operator $S^2 := (S^x)^2 + (S^y)^2 + (S^z)^2$ we can diagonalise H by working with the irreducible representation (irrep) of the total spin operator. Notice also

that this is equivalent to the XXZ Hamiltonian on the complete graph on n vertices K_n , with a scaling of p .

We begin by rewriting the Hamiltonian in equation [7.33](#) using the global spins operators $S^\alpha = \sum_i \sigma_i^\alpha$ for $\alpha \in x, y, z$:

$$H(\bar{G}(n, p)) = \frac{p}{n} \left(-J (S^2 - (S^z)^2) - 2 \right) + \Delta (S^z - 2) \Big) . \quad (7.34)$$

The basis states for the desired irrep will be written $|S, M\rangle$ where

$$\begin{aligned} S^Z |S, M\rangle &= M |S, M\rangle \\ S^2 |S, M\rangle &= 4S(S+1) |S, M\rangle . \end{aligned} \quad (7.35)$$

Since we are considering spin $\frac{1}{2}$ particles, the spin S can take on values $S \in \{0, 1, \dots, \frac{1}{2}n\}$ (assuming that n is even) and for fixed S we have $M \in \{-S, -S+1, \dots, 0, \dots, S-1, S\}$.

Thus we are transforming the tensor product representation which arises when taking n spin $\frac{1}{2}$ particles into the tensor sum representation, which is summarised by

$$|S, M\rangle = \sum_{\sigma_1 + \sigma_2 + \dots + \sigma_n = M} c(s_1, s_2, \dots, s_n; \sigma_1, \sigma_2, \dots, \sigma_n) |\sigma_1 \sigma_2, \dots, \sigma_n\rangle , \quad (7.36)$$

where $s_i = \frac{1}{2}$ since we are considering spin $\frac{1}{2}$ particles. The coefficients $c(s_1, s_2, \dots, s_n; \sigma_1, \sigma_2, \dots, \sigma_n)$ will be products of the Clebsch-Gordan coefficients for two particles, $c_{CG}(s_i, s_j; \sigma_i, \sigma_j)$.

The eigenvalues of the global Hamiltonian of equation [7.34](#) are found by applying it to the basis states $|S, M\rangle$. Their degeneracies $D_{S,M}$ follow from Clebsch-Gordan coefficients. It is found that

$$\lambda_{S,M} = \frac{-Jp}{n} (4S(S+1) - M^2) + 2Jp + \frac{p\Delta}{n} (M^2 - n)$$

$$D_{S,M} = \binom{n-1}{n/2 - S - \delta_{S,0}} - \binom{n-1}{n/2 - S - 1 - \delta_{S,0}}$$

In particular, any $|S, M\rangle$ with $S = \frac{n}{2}$ is a ground state.

The main Theorem of paper [134](#) is that equations [7.31](#) and [7.32](#) have identical free energies in the thermodynamic limit. As a consequence, all thermodynamic observables are equal in this limit.

Theorem 7.4.1. *Consider the XXZ Hamiltonians associated with (i) an Erdős-Rényi graph $G(n, p)$ on n vertices and (ii) the associated weighted graph $\bar{G}(n, p)$ on n vertices. For any $\delta > 0$, there exists an $\bar{n} > 0$ and $k > 0$ such that for any $n > \bar{n}$ it holds that*

$$P(|f(H(G(n, p))) - f(H(\bar{G}(n, p)))| < kn^{-1/2}) \geq 1 - \delta \quad (7.37)$$

so that in the thermodynamic limit, an n can be chosen such that there is a high probability that $|f(H(G(n, p))) - f(H(\bar{G}(n, p)))| = 0$.

We will not give a proof of this Theorem, since it will be encompassed later as a specific example of a more general Theorem. Note that we do not average the free energy over many realisations of graphs $G(n, p)$ drawn from the Erdős-Rényi ensemble—the result is that for a *single* graph drawn from such a distribution and in the limit $n \rightarrow \infty$, with high probability (that approaching 1) the difference between observables scales at $\mathcal{O}(n^{-\frac{1}{2}})$.

This result has an interesting implication. The Rado graph is the unique limit of the sequence $(G(n, 0.5))$ of Erdős-Rényi graphs with $p = \frac{1}{2}$ as $n \rightarrow \infty$. Since it is known to contain all simple graphs as induced subgraphs, Theorem 7.4.1 implies that for one randomly picks a simple graph in the limit $n \rightarrow \infty$ then asymptotically almost surely the physics of the XXZ graph will simply be, up to a scaling factor of p , given by $H(K_n)$, where K_n is the complete graph on n vertices.

This conclusion is remarkable because the physics of $H(K_n)$ is not many-body. As we have seen, it can be solved using the $|S, M\rangle$ basis states for the n dimensional representation of $SU(2)$, so that we are in effect solving a single large spin of dimension n . The original many-body problem, on the other hand, has 2^n basis states.

In terms of phase transitions, we are interested in that between the XY and AFM states as Δ changes, as already discussed for a 1D chain. Here, we focus on $\Delta > 0$. The relevant order parameters are

$$\begin{aligned} C_{\text{AFM}} &= \frac{1}{n_E} \sum_{(v, v') \in E} \langle \sigma_v^z \sigma_{v'}^z \rangle, \\ C_{\text{XY}} &= \frac{1}{n_E} \sum_{\substack{v, v'=1 \\ v > v'}}^L \langle \sigma_v^x \sigma_{v'}^x + \sigma_v^y \sigma_{v'}^y \rangle. \end{aligned} \quad (7.38)$$

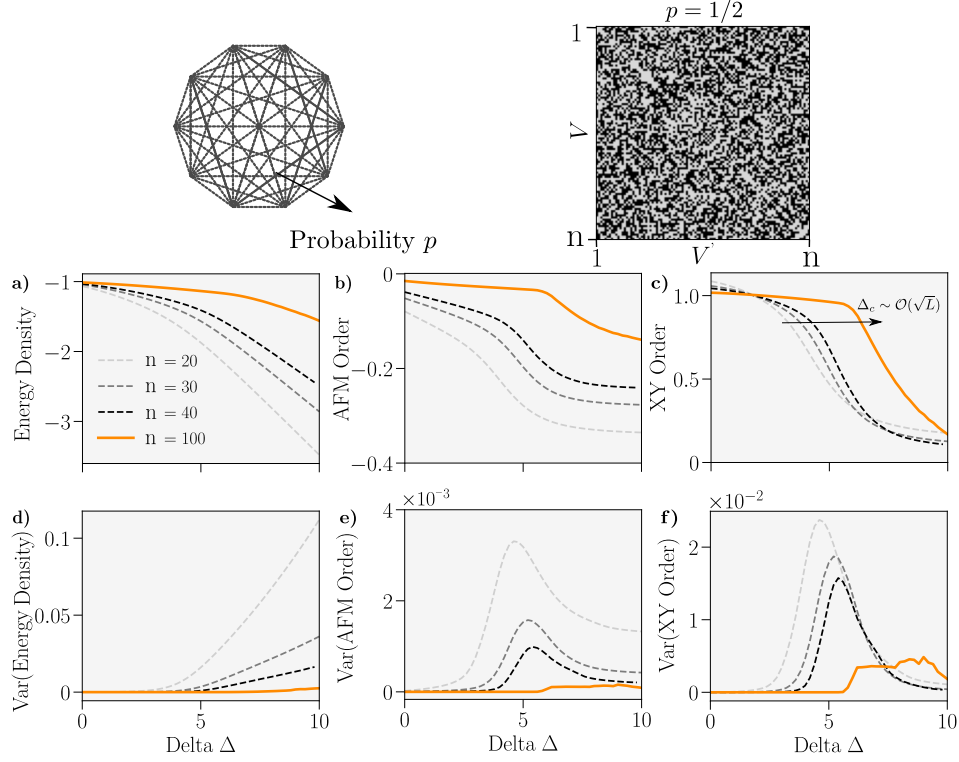


Figure 7.2: Properties of the ground-state of the spin 1/2 XXZ Hamiltonian on the Erdős-Rényi (ER) graph. A graph schematic is provided top left. We set $p = 1/2$ here and an example adjacency matrix for $L = 100$ is shown for this parameter in the top right. Data is obtained, for a given Δ , by calculating the ground state for n of draws of the ER graph from its ensemble and then averaging (plots a-c) or taking the variance (plots d-f). System sizes are coded by colour. **a-c)** Energy density $\frac{1}{n}\langle H \rangle$, anti-ferromagnetic order C_{AFM} and XY order C_{XY} - see equation (7.38) - versus Δ . **d-f)** Variance of the corresponding upper observables. We used $m = 100, 100, 100$ and 10 for $n = 20, 30, 40$ and 100 respectively. Image taken from [134].

We plot the results in Figure 7.2, which were obtained numerically by performing MPS calculations of the ground state. Whilst on a finite-size system we observe a ground-state transition between the XY and AFM phases, we find that the critical point drifts and is disappearing as $L \rightarrow \infty$ due to an apparent scaling of $\Delta_C \mathcal{O}(\sqrt{L})$.

Naturally, for small n , there is no reason to believe that their XY and AFM orders, as measured by C_{XY} and C_{AFM} respectively, should be similar. Nevertheless, as n increases, we expect such fluctuations to decrease. In order to ensure that fluctuations are suppressed as $n \rightarrow \infty$ so that the order parameters C_{AFM} and C_{XY} are well-defined for large system sizes, we calculate the variance of these parameters. This is defined

$$\text{Var}(C) = \frac{1}{m-1} \sum_{i=1}^m (C^i - \bar{C})^2. \quad (7.39)$$

As shown in Figure 7.2, this variance decreases with n , indicating that the ground-state properties are converging to a single fixed limit for $n \rightarrow \infty$.

This is in some sense a negative result: picking a graph from the Erdős-Rényi ensemble for $p = 0.5$ is equivalent to picking one from the space of all graphs on n vertices, with each being equally likely. As n becomes very large, no matter what graph you choose, the physics of the XXZ model on this graph will with high probability be trivial. We therefore are interested in types of graph which have small measure in this space, but which *do* give interesting physics in the limit $n \rightarrow \infty$.

When one considers irregular graphs instead—which are graphs where the variance of the degree of vertices in the graph is non vanishing—a phase transition does occur. For such examples, then the XXZ Hamiltonian does host complex many-body phases of matter, demonstrating that the result of Theorem 7.4.1 does not apply to any choice of dense random graph G_n . These examples are discussed and ground states found using matrix product state calculations in [134], and include for example defining G_n to have, for each vertex, a degree which is distributed according to a uniform distribution.

7.4.2 Dense graph sequences

By looking at the XXZ models as an example, we have seen that if one considers a sequence $(G(n, p))_n$ of Erdős-Rényi graphs, and on the other hand considers the sequence $(\bar{G}(n, p))_n$ where $\bar{G}(n, p)$ is a complete graph on n vertices with edge weights p , then the free energies of $H(G(n, p))$ and $H(\bar{G}(n, p))$ agree in the limit $n \rightarrow \infty$.

We show how this generalises for any convergent sequence (G_n) of dense random graphs; we define a procedure which defines a corresponding graph sequence $(\bar{G}_n)$

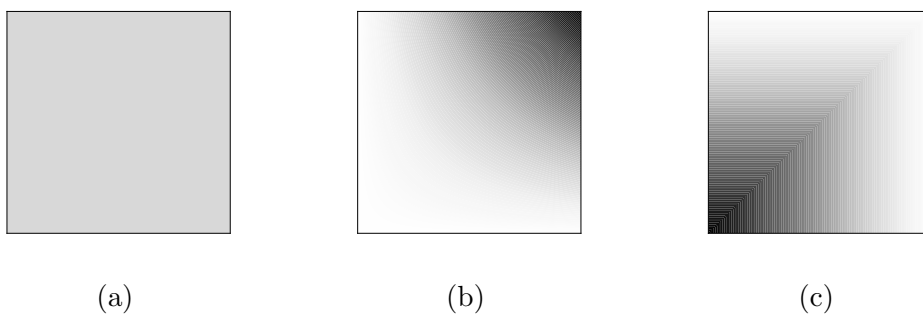


Figure 7.3: example graphons

such that $|f(H(G_n)) - f(H(\bar{G}_n))|$ vanishes in the limit $n \rightarrow \infty$.

Proving this statement utilises the graphon, which was shown by Lovász to be the limit of a sequence of dense graphs [102]. A graphon is a continuous function $W : [0, 1] \times [0, 1] \rightarrow [0, 1]$ so that for each choice $(x, y) \in [0, 1]^2$, $W(x, y)$ evaluates to a number in the interval $[0, 1]$. Loosely, this can be interpreted as the probability of vertices x and y , in a continuum of vertices, being connected. Some example graphons are shown in Figure 7.3. On the left is the constant graphon $W(x, y) = p$, central is the graphon $W(x, y) = xy$ and on the right we show the growing uniform attachment graphon, $W(x, y) = 1 - \max(x, y)$.

The graphon W is the limit object of a sequence of dense graphs (G_n) in the sense that the relative likelihood of finding any simple graph F inside of W and G_n converges for $n \rightarrow \infty$. Some intuition can be gained by observing the ‘visual convergence’ of the adjacency matrices of graphs G_n to the graphon W . Taking for example a sequence of Erdős-Rényi graphs, we see that they converge to the constant graphon $W(x, y) = p$ below.

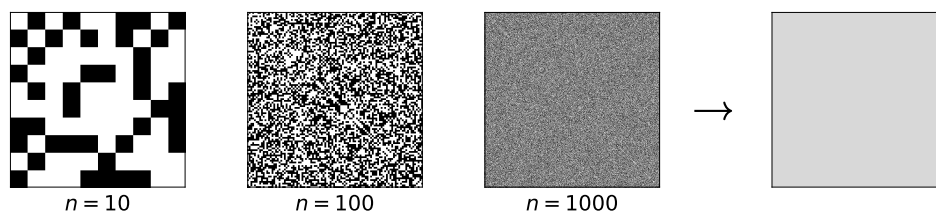


Figure 7.4: A sequence of adjacency matrices corresponding to a sequence of random graph realisations of size n , converging to the graphon $W(x, y) = p$.

As another example, the graphon $W(x, y) = 1 - \max(x, y)$ is the limit of the sequence of graphs (G_n) constructed as follows. G_1 is simply a single vertex, G_2 is constructed by adding a second vertex and with probability $1/2$ connecting it to the

first vertex. In general, at step n an n th vertex is added and each pair of vertices (i, j) which do not share an edge are connected with probability $1/n$.

Importantly, it is also possible to move in the opposite direction and given a graphon $W(x, y)$ construct sequences of finite graphs whose limit is $W(x, y)$. These finite graphs can be constructed via one of two methods: *stochastic* or *weighted* sampling of $W(x, y)$ and we use G_n^s and G_n^w to refer to their respective realisation over the vertices $v = 1 \dots n$. They can be constructed by defining the quantity

$$P_{v,v'} = n^2 \int_{I_v \times I_{v'}} W(x, y) dx dy, \quad I_v = [(v-1)/n, v/n]. \quad (7.40)$$

We then define the stochastic and weighed graph sequences (G_n^s) and (G_n^w) respectively as follows.

1. *Stochastic sampling.* The graph G_n^s for any n has adjacency matrix $A_{v,v'} = 1$ with probability $P_{v,v'}$ and $A_{v,v'} = 0$ otherwise.
2. *Weighted sampling.* The graph G_n^w for any n has adjacency matrix $A_{v,v'} = P_{v,v'}$.

These sequences are guaranteed to converge to the graphon $W(x, y)$ in the limit $n \rightarrow \infty$.

The theorem we wish to prove is the following.

Theorem 7.4.2. *Let $f(H) = -\frac{1}{n\beta} \ln(\text{Tr}(\exp(-\beta H)))$ be the free energy density of a $d^n \times d^n$ many-body Hamiltonian, with $\beta \in \mathbb{R}_{\geq 0}$ and d the dimension of the local Hilbert space. Let G_n^s and G_n^w be the stochastic and weighted realisations on n vertices of a graphon $W(x, y)$ respectively. For an arbitrary set of real, finite values for the parameters $\{J^x, J^y, J^z, h^x, h^y, h^z\}$ and for any $\delta > 0$, there exists a $\bar{n}' > 0$ and $k' > 0$ such that for any $n > \bar{n}'$ the following is true*

$$P(|f(H(G_n^s)) - f(H(G_n^w))| < k'n^{-1/2}) \geq 1 - \delta. \quad (7.41)$$

As such, for $n \rightarrow \infty$, asymptotically almost surely, the difference between the free energy densities vanishes.

The proof will rely on the following lemma, which we prove first.

Lemma 7.4.3. *Let $(A_n) = A_1, A_2, \dots$ and $(B_n) = B_1, B_2, \dots$ be sequences of square hermitian matrices, with the matrices A_n and B_n being of dimension $d^n \times d^n$ for some positive finite d . For any index n define $D_n := A_n - B_n$ and let $\lambda_{D,n}$ be*

the largest eigenvalue, in terms of absolute value, of D_n . If $|\lambda_{D,n}| = \mathcal{O}(n^\gamma)$ then $|f(A_n) - f(B_n)| = \mathcal{O}(n^{\gamma-1})$, where $f(M)$ denotes the free energy of the hermitian matrix M , which vanishes for $\gamma < 1$ as $n \rightarrow \infty$.

Here the free energy of a matrix M of dimension $d^n \times d^n$ is defined

$$f(A) = -\frac{1}{n\beta} \ln(\text{Tr}(e^{-\beta A})); \quad \beta \in \mathbb{R}_{\geq 0}. \quad (7.42)$$

Proof. Order the eigenvalues of $\hat{A}$, $\hat{B}$ and $\hat{D}$ as $\lambda_1^\eta, \lambda_2^\eta, \dots, \lambda_{d^n}^\eta$, with $\lambda_i^\eta \geq \lambda_{i-1}^\eta$ and $\eta = \hat{A}, \hat{B}$ or $\hat{D}$. By Weyl's inequality we have $\lambda_i^B + \lambda_1^D \leq \lambda_i^A \leq \lambda_i^B + \lambda_{d^n}^D$. Now, it follows from this inequality and $|\lambda_{\text{Max}}^D| = \mathcal{O}(n^\gamma)$ that $|\lambda_i^A - \lambda_i^B| = \mathcal{O}(n^\gamma) \forall i$. We can therefore write $\lambda_i^A = \lambda_i^B + c_i$ where c_i is a real number such that $|c_i| = \mathcal{O}(n^\gamma)$.

Now define the absolute free energy density difference of $\hat{A}$ and $\hat{B}$, substituting in $\lambda_i^A = \lambda_i^B + c_i$:

$$\Delta f = |f(\hat{A}) - f(\hat{B})| = \left| \frac{1}{n} \ln \left(\frac{\text{Tr}(\exp(-\beta \hat{A}))}{\text{Tr}(\exp(-\beta \hat{B}))} \right) \right| = \left| \frac{1}{n} \ln \left(\frac{\sum_{i=1}^{d^n} \exp(-\beta(c_i)) \exp(-\beta(\lambda_i^B))}{\sum_{i=1}^{d^n} \exp(-\beta(\lambda_i^B))} \right) \right|. \quad (7.43)$$

Define the following vectors of dimension d^n :

$$\begin{aligned} \mathbf{v} &= (\exp(-\beta \lambda_1^A/2), \exp(-\beta \lambda_2^A/2), \dots, \exp(-\beta \lambda_{d^n}^A/2))^T, \\ \mathbf{c} &= (\exp(-\beta c_1), \exp(-\beta c_2), \dots, \exp(-\beta c_{d^n}))^T, \end{aligned} \quad (7.44)$$

and the matrix $\hat{C} = \text{Diag}(\mathbf{c})$. Observe that

$$\Delta f = \left| \frac{1}{n} \ln \left(\frac{\langle \mathbf{v} | \hat{C} | \mathbf{v} \rangle}{\langle \mathbf{v} | \mathbf{v} \rangle} \right) \right| = \left| \frac{1}{n} \ln \left(\langle \tilde{\mathbf{v}} | \hat{C} | \tilde{\mathbf{v}} \rangle \right) \right|, \quad (7.45)$$

where we have defined $\tilde{\mathbf{v}} = \mathbf{v} / \sqrt{\langle \mathbf{v} | \mathbf{v} \rangle}$. As $\langle \tilde{\mathbf{v}} | \tilde{\mathbf{v}} \rangle = 1$, the argument of the logarithm is clearly bounded by the largest eigenvalue (by magnitude) λ_{Max}^C of $\hat{C}$. From $|c_i| = \mathcal{O}(n^\gamma) \forall i$ we have $|\lambda_{\text{Max}}^C| = \exp(\mathcal{O}(n^\gamma))$ and thus

$$\Delta f = \left| \frac{1}{n} \ln(\exp(\mathcal{O}(n^\gamma))) \right| = \mathcal{O}(n^{\gamma-1}), \quad (7.46)$$

completing our proof of Lemma 7.4.3 (the sign on the argument of the exponential $\exp(\mathcal{O}(n^\gamma))$ is not of concern as we are taking the absolute value of its logarithm). $\square$

We now prove Theorem [7.4.2](#).

Proof. We begin by defining the following operator:

$$\begin{aligned} D_n &:= H(G_n^S) - H(G_n^W) = \sum_{\alpha=x,y,z} D_n^\alpha \\ &= \sum_{\alpha=x,y,z} \frac{1}{n} \left(\sum_{\substack{i,j=1 \\ i < j}}^n A_{ij}^S \hat{\sigma}_i^\alpha \hat{\sigma}_j^\alpha - \sum_{\substack{i,j=1 \\ i < j}}^n A_{ij}^W \hat{\sigma}_i^\alpha \hat{\sigma}_j^\alpha \right), \end{aligned} \quad (7.47)$$

where A_{ij}^S are the matrix elements of G_n^S and A_{ij}^W are the matrix elements of G_n^W , finite stochastic and weighted realisations of some graphon $W(x, y)$.

We proceed to evaluate the eigenvalues of the operator D_n^α . As such, consider its eigenstates $|\sigma_1^\alpha, \dots, \sigma_n^\alpha\rangle$, where $\hat{\sigma}_i^\alpha |\sigma_1^\alpha, \dots, \sigma_n^\alpha\rangle = \mu_i |\sigma_1^\alpha, \dots, \sigma_n^\alpha\rangle$, with $\mu_i = \pm 1$, depending on whether the i th spin is pointing ‘up’ or ‘down’ in that basis. We will define $\mu_{ij} := \mu_i \mu_j$ and consider the eigenvalue

$$\begin{aligned} \langle \sigma_1^\alpha, \dots, \sigma_n^\alpha | D_n^\alpha | \sigma_1^\alpha, \dots, \sigma_n^\alpha \rangle &= \\ \frac{1}{n} \left(\sum_{\substack{i,j=1 \\ i < j}}^n A_{ij}^S \mu_{ij} - \sum_{\substack{i,j=1 \\ i < j}}^n A_{ij}^W \mu_{ij} \right), \end{aligned} \quad (7.48)$$

We will proceed to show that, independent of the eigenstate $|\sigma_1^\alpha, \dots, \sigma_n^\alpha\rangle$, this eigenvalue grows, at most, as $n^{1/2}$ in the large n limit. From there we can invoke Weyl’s inequality to show that the eigenvalues of D_n grow asymptotically as $n^{1/2}$ and subsequently invoke Lemma 1 (from above) to complete the proof of Theorem [7.4.2](#).

To prove the $n^{1/2}$ growth of equation [\(7.48\)](#), we begin with the Hoeffding inequality. This states that for independent random variables $Y_1, \dots, Y_n$ for which $a_i \leq Y_i \leq b_i$ then the sum $S_n := Y_1 + Y_2 + \dots + Y_n$ is bounded as

$$P(|S_n - \mathbb{E}(S_n)| \geq t) \leq 2 \exp\left(-\frac{2t^2}{\sum_{i=1}^n (a_i - b_i)^2}\right), \quad (7.49)$$

with the factor of 2 stemming from the fact we have incorporated both the upper and lower Hoeffding bounds together. We will apply this bound to equation [\(7.48\)](#).

Let us construct the set $X := \{X_{ij}\}$ which consists of the $\frac{n(n-1)}{2}$ random variables $X_{ij} := A_{ij}^S \mu_{ij}$ (we have $i, j = 1 \dots n$ and $i > j$). Observe that since $A_{ij}^S \in \{0, 1\}$ (the

graph G_n^S is simple) and μ_{ij} is either 1 or -1 , we have that $-1 \leq X_{ij} \leq 1$. Also, $\mathbb{E}(X_{ij}) = \mathbb{E}(A_{ij}^S)\mu_{ij} = A_{ij}^W\mu_{ij}$, where $\mathbb{E}(A_{ij}^S)$ denotes the expected value. In our case, however, all of the X_{ij} quantities are not independent—the sign of X_{ij} is determined from the sign of X_{ik} and X_{kj} . This can be dealt with by applying the Hoeffding bound to the X_{ij} 's with positive and negative sign separately.

A given eigenvector $|\sigma_1^\alpha, \dots, \sigma_n^\alpha\rangle$ will consist of a number of spins pointing up ($\sigma_i = +1$) and the remainder pointing down ($\sigma_i = -1$). Define $M = \sum_{k=1}^n \sigma_k$, then it can be checked that of the $\frac{n(n-1)}{2}$ parameters μ_{ij} , $\frac{n^2+M^2-2n}{4}$ are positive and $\frac{n^2-M^2}{4}$ are negative. We therefore partition the set $X := \{X_{ij}\}$ into two sets as follows

$$\begin{cases} x^A := \{X_{ij} | \mu_{ij} = 1\}, \\ x^B := \{X_{ij} | \mu_{ij} = -1\}. \end{cases}$$

We would like also to keep track of the values $A_{ij}^W\mu_{ij}$, so we partition the set $\{\mu_{ij}A_{ij}^W\}$ into

$$\begin{cases} \bar{x}^A := \{\mu_{ij}A_{ij}^W | \mu_{ij} = 1\}, \\ \bar{x}^B := \{\mu_{ij}A_{ij}^W | \mu_{ij} = -1\}. \end{cases}$$

We can now invoke the Hoeffding bound on each set separately, since each set now consists on independent random variables. We have two sums on which to invoke the bound: $S^A := \sum_{l=1}^{\frac{M^2+n^2-2n}{4}} x_l^A$ and $S^B := \sum_{l=1}^{\frac{n^2-M^2}{4}} x_l^B$, where we use x_l^A and x_l^B to refer to elements from x^A and x^B respectively. Likewise we will use $\bar{x}_i^A$ and $\bar{x}_i^B$ to refer to individual elements of $\bar{x}^A$ and $\bar{x}^B$ respectively. Then observe that, due to $\mathbb{E}(X_{ij}) = A_{ij}^W\mu_{ij}$ we have $\mathbb{E}(S^A) = \sum_{l=1}^{\frac{M^2+n^2-2n}{4}} \bar{x}_l^A$ and $\mathbb{E}(S^B) = \sum_{l=1}^{\frac{n^2-M^2}{4}} \bar{x}_l^B$ equation (7.49) then gives two bounds:

$$\begin{cases} P(|S^A - \mathbb{E}(S^A)| \geq t_1) \leq 2 \exp\left(\frac{-2t_1^2}{n^2+M^2-2n}\right), \\ P(|S^B - \mathbb{E}(S^B)| \geq t_2) \leq 2 \exp\left(\frac{-2t_2^2}{n^2-M^2}\right). \end{cases}$$

We combine these bounds to obtain:

$$\begin{aligned}
& P \left(\left| \sum_{\substack{i,j=1 \\ i < j}}^n \mu_{ij} A_{ij}^S - \sum_{\substack{i,j=1 \\ i < j}}^n \mu_{ij} A_{ij}^W \right| \geq t_1 + t_2 \right) \\
& \leq 4 \exp \left(\frac{-2t_1^2}{n^2 + M^2 - 2n} \right) \exp \left(\frac{-2t_2^2}{n^2 - M^2} \right). \tag{7.50}
\end{aligned}$$

Since we fixed the magnetisation M of the eigenstate in deriving the above bound, we should take the union bound over $\binom{n}{\frac{1}{2}(n+M)}$ eigenstates with magnetisation M . We will denote an eigenstate with magnetisation M as $|\sigma_M\rangle$, resulting in the new bound

$$\begin{aligned}
& \bigcup_{|\sigma_M\rangle} P \left(\left| \sum_{\substack{i,j=1 \\ i < j}}^n \mu_{ij} A_{ij}^S - \sum_{\substack{i,j=1 \\ i < j}}^n \mu_{ij} A_{ij}^W \right| \geq t_1 + t_2 \right) \\
& \leq 4 \binom{n}{\frac{1}{2}(n+M)} \exp \left(\frac{-2t_1^2}{n^2 + M^2 - 2n} \right) \exp \left(\frac{-2t_2^2}{n^2 - M^2} \right). \tag{7.51}
\end{aligned}$$

Now observe that the following is true

$$\begin{aligned}
& \binom{n}{\frac{1}{2}(n+M)} \exp \left(\frac{-2t_1^2}{n^2 + M^2 - 2n} \right) \exp \left(\frac{-2t_2^2}{n^2 - M^2} \right) \\
& \leq \binom{n}{\frac{n}{2}} \exp \left(-\frac{t_1^2 + t_2^2}{n^2} \right), \tag{7.52}
\end{aligned}$$

where $n \in \mathbb{N}$, $|M| \leq n$ and $t_1, t_2 \in \mathbb{R}^+$.

This leads us to the following

$$\begin{aligned}
& \bigcup_M \bigcup_{|\sigma_M\rangle} P \left(\left| \sum_{\substack{i,j=1 \\ i < j}}^n \mu_{ij} A_{ij}^S - \sum_{\substack{i,j=1 \\ i < j}}^n \mu_{ij} A_{ij}^W \right| \geq t_1 + t_2 \right) \\
& \leq 4(n+1) \binom{n}{\frac{n}{2}} \exp \left(-\frac{t_1^2 + t_2^2}{n^2} \right), \tag{7.53}
\end{aligned}$$

where the union bound has again been used.

Now let $\delta = 4(n+1) \binom{n}{\frac{n}{2}} \exp \left(-\frac{t_1^2 + t_2^2}{n^2} \right)$, which rearranging gives us an expression for $\sqrt{t_1^2 + t_2^2}$, and, moreover, note that $t_1 + t_2 = \sqrt{t_1^2 + t_2^2 + 2t_1 t_2} \geq \sqrt{t_1^2 + t_2^2}$. We

have

$$\delta = 4(n+1) \binom{n}{\frac{n}{2}} \exp \left(-\frac{t_1^2 + t_2^2}{n^2} \right) \quad (7.54)$$

$$\implies \frac{\delta}{4(n+1) \binom{n}{\frac{n}{2}}} = \exp \left(-\frac{t_1^2 + t_2^2}{n^2} \right) \quad (7.55)$$

$$\implies \frac{4(n+1) \binom{n}{\frac{n}{2}}}{\delta} = \exp \left(\frac{t_1^2 + t_2^2}{n^2} \right) \quad (7.56)$$

$$\implies t_1^2 + t_2^2 = n^2 \ln \left[\frac{4(n+1) \binom{n}{\frac{n}{2}}}{\delta} \right]. \quad (7.57)$$

We can therefore write

$$\bigcup_M \bigcup_{|\sigma_M\rangle} P \left(\left| \sum_{\substack{i,j=1 \\ i < j}}^n \mu_{ij} A_{ij}^S - \sum_{\substack{i,j=1 \\ i < j}}^n \mu_{ij} A_{ij}^W \right| \geq \sqrt{n^2 \ln \left[\frac{4(n+1) \binom{n}{\frac{n}{2}}}{\delta} \right]} \right) \leq \delta. \quad (7.58)$$

Thus, with high probability, we can bound our expression by $\sqrt{n^2 \ln \left[\frac{4(n+1) \binom{n}{\frac{n}{2}}}{\delta} \right]}$. We wish to bound this expression in terms of powers of n . Using the fact $\binom{n}{\frac{n}{2}} \sim \sqrt{\frac{2}{n\pi}} 2^n$ for large n , and also that using $2^n = e^{n \ln(2)}$, we arrive at

$$\sqrt{n^2 \ln \left[\frac{4(n+1) \binom{n}{\frac{n}{2}}}{\delta} \right]} \sim \sqrt{n^2 \ln \left[\frac{4(n+1) \sqrt{2n\pi} e^{n \ln(2)}}{\delta} \right]}, \quad (7.59)$$

for large n . The sub expression $\ln \left[\frac{4(n+1) \sqrt{2n\pi} e^{n \ln(2)}}{\delta} \right]$ can be written as a sum of logarithmic functions:

$$\ln \left[\frac{4(n+1) \sqrt{2n\pi} e^{n \ln(2)}}{\delta} \right] = \ln \left[\frac{4\sqrt{2/\pi}}{\delta} \right] + \ln \left[\frac{n+1}{\sqrt{n}} \right] + \ln [e^{n \ln(2)}], \quad (7.60)$$

so that we need to bound each of these terms individually. The first and third are clearly $\mathcal{O}(1)$ and $\mathcal{O}(n)$ respectively. Using that $\ln(\sqrt{n}) = \frac{1}{2} \ln(n)$ and that $\ln(n) \leq n-1$, the second term can be seen to be at most $\mathcal{O}(n)$. Recall that we defined delta in terms of n ; we call this value of the parameter $\bar{n}$, which is fixed once δ has been chosen.

Therefore, altogether, we can say that for a chosen $\delta > 0$ there exists a constant $k > 0$ and an $\bar{n} > 0$ such that

$$\bigcup_M \bigcup_{|\sigma_M\rangle} P \left(\left| \sum_{\substack{i,j=1 \\ i < j}}^n \mu_{ij} A_{ij}^S - \sum_{\substack{i,j=1 \\ i < j}}^n \mu_{ij} A_{ij}^W \right| \geq kn^{3/2} \right) \leq \delta \quad \forall n > \bar{n}. \quad (7.61)$$

Since δ can be as small as we please, as long as we choose n to be greater than $\bar{n}$, then with high probability the expression is bounded by $kn^{3/2}$ for k a constant greater than 0. Now, using this we can say that for any choice of $\delta > 0$, there exists a $\bar{n}' > 0$ and $k' > 0$ such that

$$P \left(|\langle \sigma_1^\alpha, \dots, \sigma_n^\alpha | D_n^\alpha | \sigma_1^\alpha, \dots, \sigma_n^\alpha \rangle| < k'n^{\frac{1}{2}} \right) \geq 1 - \delta \quad \forall M, |\sigma\rangle, n > \bar{n}'. \quad (7.62)$$

From Weyl's inequality we know that the eigenvalues of the operator $D_n = H(G_n^S) - H(G_n^W)$ are therefore, with high probability, bounded as $k'n^{1/2}$. From here we can invoke Lemma [7.4.3](#) with $\kappa = 1/2$ and the proof is complete. $\square$

We have therefore established our goal of generalising the result for a sequence of Erdős-Rényi graphs to that of any dense graph sequence. While we have showed convergence, we have not in any way *solved* the Hamiltonian on a random graph G_n in the limit $n \rightarrow \infty$ —we do not know anything of its physics. Theorem [7.4.2](#) now establishes that we can instead work with the Hamiltonian $H(G_n^w)$ and take the limit $n \rightarrow \infty$ of the solution. We therefore now turn to the problem of determining the equilibrium physics of $H(G_n^w)$.

7.4.3 Thermodynamic physics on dense graph limits

Solving a Hamiltonian H requires finding the spin configuration $|\sigma\rangle$ which has minimum energy under H . Since the size of the Hilbert space grows as $(2s+1)^n$ for spins s , for large n the problem becomes intractable, which is referred to as the **many-body problem**. Mean-field theory, on the other hand, for dense interactions is known to become exact in the limit $n \rightarrow \infty$.

Mean field theory is generally an approximate method, and gets around the intractable many-body problem by introducing, for each spin, a field that averages out the effect of all other spins that it interacts with. In general, this is of course at the detriment of the accuracy of the resulting solution. At critical points, in particular, the mean-field approximation often breaks down because of the long

range order (through entanglement) that develops in the system. Nevertheless, if interactions are dense then in the limit $n \rightarrow \infty$ the treatment becomes exact.

We would like to replace each spin operator σ_v^α with its thermodynamic average $\langle \sigma_i^\alpha \rangle = \text{Tr}(\rho \sigma_i^\alpha)$ and a fluctuation term around this average:

$$\sigma_i = \langle \sigma_v^\alpha \rangle + \delta \sigma_i^\alpha. \quad (7.63)$$

The two body terms $\sigma_v^\alpha \sigma_{v'}^\alpha$ become

$$\begin{aligned} \sigma_v^\alpha \sigma_{v'}^\alpha &= \langle \sigma_v^\alpha \rangle \langle \sigma_{v'}^\alpha \rangle + \langle \sigma_v^\alpha \rangle \delta \sigma_{v'}^\alpha + \delta \sigma_v^\alpha \langle \sigma_{v'}^\alpha \rangle + \delta \sigma_v^\alpha \delta \sigma_{v'}^\alpha \\ &= \langle \sigma_v^\alpha \rangle \langle \sigma_{v'}^\alpha \rangle + \langle \sigma_v^\alpha \rangle (\sigma_{v'}^\alpha - \langle \sigma_{v'}^\alpha \rangle) + (\sigma_v^\alpha - \langle \sigma_v^\alpha \rangle) \langle \sigma_{v'}^\alpha \rangle + \mathcal{O}(\delta^2) \end{aligned} \quad (7.64)$$

where in the second line we have replaced the $\delta \sigma_i^\alpha$ terms to get an expression in terms of σ_i^α and $\langle \sigma_i^\alpha \rangle$ only, and have also neglected the term of order δ^2 . Note that the first term is simply a constant (it is not a quantum operator), and will only shift the entire spectrum of the Hamiltonian by a constant amount. We therefore also neglect this constant, too.

The Hamiltonian becomes

$$H = \frac{1}{n} \sum_{v, v' \leq v \in V(G)} \sum_{\alpha=x,y,z} A_{v,v'} J^\alpha (\langle \sigma_{v'}^\alpha \rangle \sigma_v^\alpha + \langle \sigma_v^\alpha \rangle \sigma_{v'}^\alpha) + \sum_{v \in V} \sum_{\alpha=x,y,z} h^\alpha \sigma_v^\alpha \quad (7.65)$$

which, by re indexing the first summation, is equivalent to

$$H = \frac{1}{n} \sum_{v, v' \in V(G)} \sum_{\alpha=x,y,z} A_{v,v'} J^\alpha \langle \sigma_{v'}^\alpha \rangle \sigma_v^\alpha + \sum_{v \in V} \sum_{\alpha=x,y,z} h^\alpha \sigma_v^\alpha. \quad (7.66)$$

Note that this last equation can be written as a sum over local Hamiltonians, $H = \sum_{v \in V(G)} H_v$, where

$$H_v := \sum_{\alpha=x,y,z} \sigma_v^\alpha \left[\frac{1}{n} \left(\sum_{v' \in V(G)} A_{v,v'} J^\alpha \langle \sigma_{v'}^\alpha \rangle \right) + h^\alpha \right]. \quad (7.67)$$

Under this approximate Hamiltonian, the equilibrium state at an inverse temperature β is

$$\rho = \frac{\exp(-\beta H)}{\text{Tr}(\exp(-\beta H))} = \bigotimes_{v \in V(G)} \frac{\exp(-\beta H_v)}{\text{Tr}[\exp(-\beta H_v)]} = \bigotimes_{v \in V(G)} \rho_v \quad (7.68)$$

so that the equilibrium state is a product of the mixed states on individual vertices of the graph. The explicit expression is found to be, using the derived expression for H_v ,

$$\rho_v = \begin{pmatrix} \lambda_v - \lambda_v^z \tanh(\beta \lambda_v) & -(\lambda_v^x - i \lambda_v^y) \tanh(\beta \lambda_v) \\ -(\lambda_v^x + i \lambda_v^y) \tanh(\beta \lambda_v) & \lambda_v + \lambda_v^z \tanh(\beta \lambda_v) \end{pmatrix} \quad (7.69)$$

where

$$\lambda_v^\alpha = \frac{1}{n} \left(\sum_{v' \in V(G)} A_{v,v'} J^\alpha \langle \sigma_{v'}^\alpha \rangle \right) + h^\alpha \quad (7.70)$$

and

$$\lambda_v = \sqrt{(\lambda_v^x)^2 + (\lambda_v^y)^2 + (\lambda_v^z)^2}. \quad (7.71)$$

Note that there are $3n$ of these λ_v^α terms. The expectation values $\langle \sigma_v^\alpha \rangle$ can be calculated from ρ_v as

$$\langle \sigma_v^\alpha \rangle = \text{Tr}(\sigma_v^\alpha \rho_v). \quad (7.72)$$

Putting this into the expression for the λ_v^α terms in equation [7.70](#) gives a self consistency equation which is

$$\lambda_v^\alpha = \frac{-1}{n} \left(\sum_{v'} \frac{J^\alpha A_{v,v'} \lambda_{v'}^\alpha \tanh(\beta \lambda_{v'})}{\lambda_v} \right) + h^\alpha. \quad (7.73)$$

Solving these $3n$ coupled equations and putting them in equation [7.69](#) gives an expression for the local density operator ρ_v in terms of quantities which will be specified by the particular choice of Hamiltonian: J^α , h^α , $A_{v,v'}$, β and n . The global state is then completely specified as the product $\bigotimes_{v \in V(G)} \rho_v$.

Suppose now that we have two graph sequences (G_n) and $(\bar{G}_n)$ which arise, respectively, from stochastic and weighted sampling of the graphon W which is the limit of these sequences.

If one can solve equations [7.73](#) for the choice of graph $\bar{G}_n$, then, provided n is large, we expect the solution to be exact in the thermodynamic limit, and a good approximation for large n . Moreover, from Theorem [7.4.2](#), we expect the solution to also be exact for the stochastic sequence G_n .

The thermodynamic limit of equations [7.73](#) is taken as follows. Imagining the n vertices v placed along a line of length 1 with finite distance $\delta(x) = 1/n$ between two vertices, we take the limit $\delta(x) \rightarrow dx$. Therefore the self consistency equations

become

$$\lambda^\alpha(x) = -J^\alpha \int_0^1 \frac{W(x, y) \lambda^\alpha(y) \tanh [\beta \Lambda(y)]}{\Lambda(y)} dy + h^\alpha. \quad (7.74)$$

Alternatively, this can be derived by inserting the expression for $A_{v,v'}$ which arises from weighted sampling of a graphon W . equations of the form $\int_a^b K(t, s) f(s) ds$ are called Fredholm integral equations, and $K(t, s)$ is referred to as the kernel. We therefore, in the continuous limit, have three coupled Fredholm equations (for x , y and z) with the graphon as their Kernel.

We now consider two examples in which these equations are analytically solvable.

7.4.4 The Ising model on $W(x, y) = xy$ and $W(x, y) = p$

We consider the classical Ising model on the graphons $W(x, y) = p$ and $W(x, y) = xy$. For the former, stochastic sampling produces an Erdős-Rényi graph, while, for the latter, it produces finite graphs in which the probability of attaching two vertices $i, j \in [1, n]$ is the product ij/n^2 (which, as required for a probability distribution, lies between 0 and 1).

The coupled equations [7.74](#) allow us to analytically find the equilibrium state at temperature β for both of these graphons.

The classical Ising model is realised when we take Hamiltonian parameters $J^x = J^y = h^x = h^y = h^z = 0$ and $J^z = -1$. Substituting these values into equation [7.74](#) gives

$$\lambda^z(x) = \int_0^1 W(x, y) \tanh [\beta \lambda^z(y)] dy. \quad (7.75)$$

For $W(x, y) = p$, substitution into the above equation removes dependency of $\lambda^z(x)$ on the variable x . We therefore have

$$\lambda^z(x) = \int_0^1 p \tanh [\beta \lambda^z] dy = p \tanh [\beta \lambda^z]. \quad (7.76)$$

The magnetisation is given by

$$M^Z = \tanh(\beta p M^Z). \quad (7.77)$$

It is known that in the all to all case we have $M^Z = \tanh(\beta M^Z)$ [92](#), so that in the thermodynamic limit the magnetisation is simply a rescaling, via p , of that on the complete graph.

On the other hand, taking $W(x, y) = xy$, the $\lambda^z(x)$ expression is

$$\lambda^z(x) = \int_0^1 xy \tanh[\beta \lambda^z(y)] dy = y \left[\int_0^1 x \tanh[\beta \lambda^z(y)] dy \right]. \quad (7.78)$$

Since the integral is definite, it will be some function of β only, so that it must hold that $\lambda^z(x) = xf(\beta)$. Substituting this into equation [7.78](#) gives

$$\frac{c}{\beta} = \int_0^1 y \tanh(cy) dy \quad (7.79)$$

where $c = \beta f(\beta)$. Integrating this, either by hand or using an Integration solver gives

$$\int_0^1 y \tanh(cy) dy = - \left[\frac{-12c^2 + 12\text{Li}_2(-e^{-2c}) - 24c \log(e^{-2c} + 1) + \pi^2}{24c^2} \right], \quad (7.80)$$

where Li_2 is the dilogarithm function. Altogether then the equation we obtain for c is

$$\frac{24c^3}{\beta} = 12c^2 + 24c \ln(1 + e^{-2c}) - 12\text{Li}_2(-e^{-2c}) - \pi^2. \quad (7.81)$$

In Figure [7.5\(a\)](#) we plot the total magnetisation M^z for finite realisations of both graphons $W(x, y) = p$ and $W(x, y) = xy$ using Markov Chain Monte Carlo (MCMC) simulations, which is the standard numerical approach for the Ising model. We also plot the asymptotic behaviour of the solutions for the thermodynamic limit which have been derived above. As can be seen, away from the critical point there is good agreement between the derived thermodynamic equilibrium and that predicted by the numerical simulation.

We can also analyse the effect of introducing inhomogeneous coupling: as can be seen from the plot, the approach to the ferromagnetic state with $M^z = 1$ is slower. The reason becomes clear when studying Figure [7.5\(b\)](#), which plots the local magnetisation $\sigma^z(x)$ for $x \in [0, 1]$ and for a range of inverse temperatures β ; we see that for small x the local magnetisation remains close to zero, even well past the critical point for β . Note that the black rectangle corresponds to on-site magnetisations being everywhere zero for values of β before the critical point.

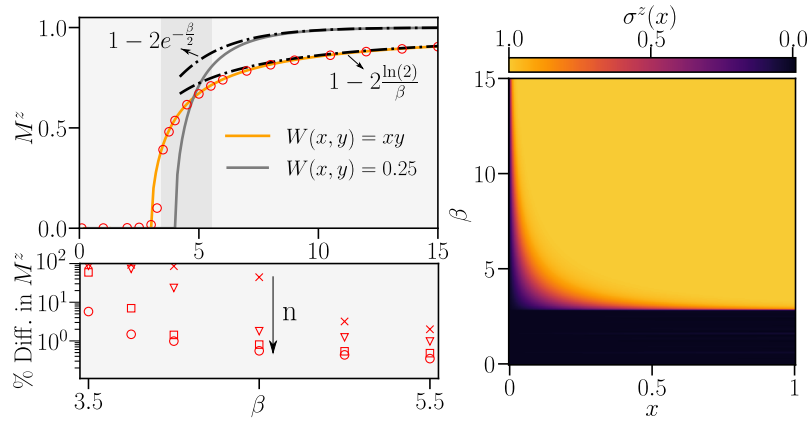


Figure 7.5: Magnetisation of the classical Ising model for the graphons: $W(x, y) = xy$ and $W(x, y) = \frac{1}{4}$. **a)** Total Magnetisation density M^z versus inverse temperature β for $n \rightarrow \infty$. Black dashed-dotted lines give the asymptotic, derived by taking the large β limit of the respective closed-form equations. Red circles correspond to Monte Carlo simulations of finite, $n = 800$, randomly sampled graphs G_n^S derived from $W(x, y) = xy$. Bottom) Percentage difference in M^z for the exact result in the thermodynamic limit versus finite-size Monte Carlo simulations at several n (crosses are $n = 100$, triangles $n = 200$, squares $n = 400$ and circles $n = 800$). For each β and n , 100 stochastic samples G_n^S are realised and the data (both top and bottom plots) is averaged over these. Further details are provided in the SM. **b)** On-site magnetisation $\sigma^z(x)$ versus β and x for the graphon $W(x, y) = xy$ in the thermodynamic limit. Image from [\[128\]](#).

7.4.5 The transverse field Ising model on $W(x, y) = \sqrt{xy}$

The Ising model is a classical spin system. We now turn to the transverse field Ising model, which, due to the presence of an external field which does not commute with interaction terms in the Hamiltonian, does host quantum effects. This will be witnessed by non-zero entanglement entropy close to the critical point.

In this case the parameters of the Hamiltonian are $J^x = J^y = h^y = h^z = 0$, $h^x = -1$ and $J^z = -1$ which results in Hamiltonian

$$H(G) = - \sum_{(v,v') \in E(G)} A_{v,v'} \sigma_v^x \sigma_{v'}^x - \sum_{v \in V(G)} \sigma_v^z. \quad (7.82)$$

As before, we are then interested in the ground or equilibrium state of $H(G_n)$ in the thermodynamic limit $n \rightarrow \infty$. We find that $\lambda^z(x)$ is given by the expression

$$\lambda^z(x) = - \int_0^1 \frac{W(x, y) \lambda^z(y) \tanh \left[\beta \sqrt{\lambda^z(y)^2 + h^2} \right]}{\sqrt{h^2 + \lambda^z(y)^2}} dy. \quad (7.83)$$

For $W(x, y) = p$ this becomes

$$\lambda^z(x) = -p \int_0^1 \frac{\lambda^z(y) \tanh \left[\beta \sqrt{\lambda^z(y)^2 + h^2} \right]}{\sqrt{h^2 + \lambda^z(y)^2}} dy. \quad (7.84)$$

Using, as before, the fact that the above implies $\lambda^z(x) = \lambda^z(y)$ i.e. there is no x dependence, the above becomes

$$\lambda^z = -p \frac{\lambda^z \tanh \left[\beta \sqrt{(\lambda^z)^2 + h^2} \right]}{\sqrt{(\lambda^z)^2 + h^2}} \quad (7.85)$$

which, simplifying, gives

$$(\lambda^z)^2 + h^2 = p^2 \tanh^2 \left[\beta \sqrt{(\lambda^z)^2 + h^2} \right]. \quad (7.86)$$

We now take the limit $\beta \rightarrow \infty$, corresponding to taking the temperature to 0, to give the ground state solution:

$$\lambda^z = \sqrt{p^2 - h^2}. \quad (7.87)$$

The magnetisation M^z is therefore

$$M^z = -\frac{\sqrt{p^2 - h^2}}{\sqrt{(p^2 - h^2) + h^2}} = \sqrt{1 - \frac{h^2}{p^2}}. \quad (7.88)$$

We consider a more interesting choice of graphon, taking $W = \sqrt{xy}$, and for which finite realisations have edge density which is inhomogeneous. equation [7.83](#) tells us that in the limit $\beta \rightarrow \infty$

$$\lambda^z(x) = \sqrt{x} \int_0^1 \frac{\sqrt{y} \lambda^z(y)}{\sqrt{h^2 + \lambda^z(y)^2}} dy \quad (7.89)$$

so that $\lambda^z(x) = \sqrt{x} f(h)$. Substituting this into the above expression yields

$$1 = \int_0^1 \frac{y}{\sqrt{h^2 + y f(h)^2}} dy. \quad (7.90)$$

This can be solved quite straightforwardly by parts, giving $f(h)$ as the solution to the implicit equation

$$\frac{2(2h^3 - 2h^2 \sqrt{h^2 + g^2(h)} + g^2(h) \sqrt{h^2 + g^2(h)})}{3g^4(h)} = 1, \quad (7.91)$$

which has the solution

$$g(h) = \begin{cases} \frac{\sqrt{2}}{3} \sqrt{1 + (1 - 3h) \sqrt{1 + 6h}} & h < \frac{1}{2}, \\ 0 & \text{otherwise.} \end{cases} \quad (7.92)$$

The transverse and longitudinal magnetisations are determined by the integrals $M^x = \int_0^1 \frac{h}{\sqrt{h^2 + g^2(h)x}} dx$ and $M^z = -\int_0^1 \frac{g(h) \sqrt{x}}{\sqrt{h^2 + g^2(h)x}} dx$ respectively. The first, by direct integration, gives

$$M^x = \begin{cases} \frac{6h}{3h + \sqrt{2 + 9h^2 + (2 - 6h) \sqrt{1 + 6h}}} & h < \frac{1}{2}, \\ 1 & \text{otherwise.} \end{cases} \quad (7.93)$$

The second can be done by an extensive series of trigonometric substitutions and

results in the following closed form expression

$$M^z(x) = \begin{cases} \frac{(1+s)(2\sqrt{(1-s)(4+18h^2-4s)}-9h^2(2\ln(3)+2\ln(h)-2\ln(-\sqrt{2-2s}+\sqrt{2+9h^2-2s})))}{108h^2(-1+2h)} & \\ 0 & \text{otherwise} \end{cases}, \quad (7.94)$$

where $s = (-1 + 3h)\sqrt{1 + 6h}$.

The results are plotted in Figure 7.6. We see that the approach to a ferromagnetic state with $M^z = 1$, as well as the approach away from the state with $M^x = 1$, as induced by increasing the transverse field, happens marginally faster for interactions governed by $W(x, y) = \sqrt{xy}$ versus those governed by $W(x, y) = p$. This arises because the magnetic susceptibility of the system at $x = 0$ is a Dirac-delta function for $W(x, y) = \sqrt{xy}$, $\lim_{h \rightarrow 0} \frac{d\sigma^x(x)}{dh} = \delta(x)$, which can be derived from equation 7.92. This is observed in Figure 7.6 at small x , where it is most susceptible it can be seen that the system is locally most susceptible the field. This, on the other hand, is no the case for $W(x, y) = p$.

We moreover calculate expectation values for a finite random realisation of the graphon $W(x, y) = \sqrt{xy}$ for system size $n \sim 100$. This is done using density matrix renormalisation group calculations on a matrix product state ansatz, i.e. a numerical calculation. These are in good agreement with the analytic calculation. Moreover, these finite size calculations allow us to calculate the Von Neumann entanglement entropy in the system, and we see that this diverges logarithmically near the transition point. Interestingly, for a fully connected graph, the model is known as the Lipkin-Meshkov-Glick model and it is known that in this case, too, the entanglement entropy at the critical point is diverging logarithmically [94]. Since we know that in this case, due to the commutation of the Hamiltonian with the total spin operator, the system is effectively a single large spin (see Section 7.4.1), it is interesting that for a different choice of graphon, and one which has highly irregular finite realisations, there is no effect on entanglement entropy scaling. This could suggest a universal mechanism underlying such scaling.

7.5 Discussion

We have considered the problem of finding the ground or equilibrium state of spin Hamiltonians on a choice of dense graph G in the thermodynamic limit. We have seen that, oftentimes in this limit the problem in fact becomes easier to solve; in the

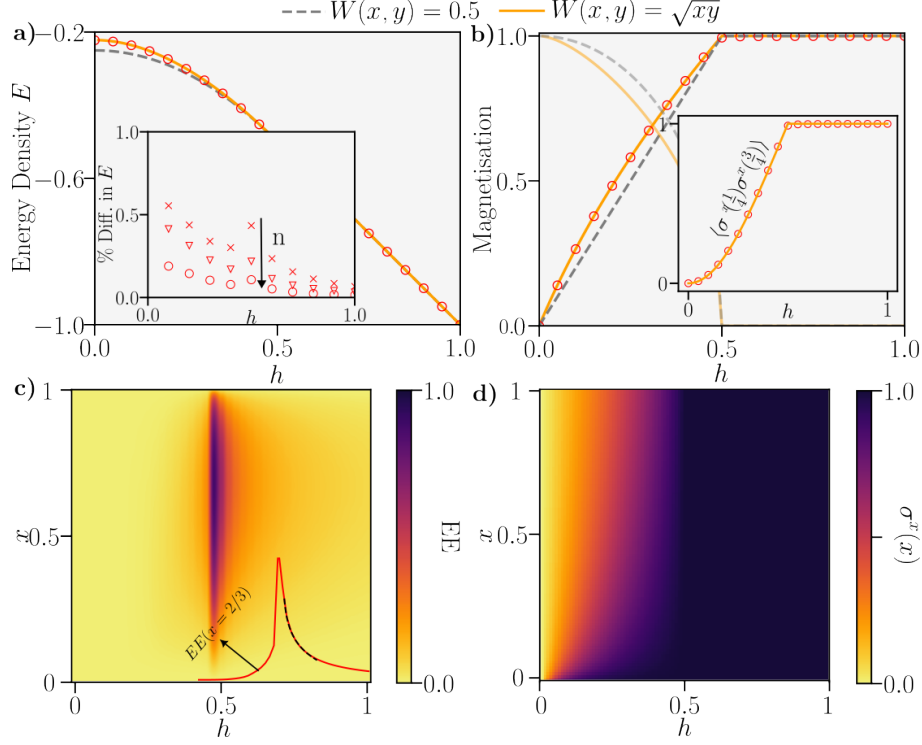


Figure 7.6: Properties of the ground state of the transverse field Ising model on the graphon $W(x, y) = \sqrt{xy}$. Results for the constant graphon are included for reference. **a)** Energy density versus transverse field strength h . The orange line represents the analytic solution in the thermodynamic limit. Markers represent numerical calculations averaged over 100 finite stochastic realisations of $W(x, y) = \sqrt{xy}$ on $n = 400$ sites. Inset) Percentage difference between the ground state energy calculated on $n = 100, 200$ and 400 (cross, triangle and circle marker respectively) site random-exchange realisations of $W(x, y)$ and the exact solution for $n \rightarrow \infty$. **b)** Total transverse (unfaded) and longitudinal (faded) Magnetisation densities of the ground state. Inset) Two-point correlator $\langle \sigma^x(\frac{1}{4}) \sigma^x(\frac{3}{4}) \rangle$. **c)** Von-Neumann Entanglement Entropy (EE) of $W(x, y) = \sqrt{xy}$ averaged over 100 stochastic realisations on $n = 400$ sites. The partition is between the first xn sites of the system and the remaining $(1-x)n$ sites. The red curve corresponds to the entanglement entropy at $x = 2/3$ for $h = 0 \rightarrow 1$. The dotted black line is the fit $EE(x = 2/3) = -0.136 \log_2(h - 0.5) - 0.089$. **d)** Analytical result for the on-site magnetisation $\sigma^x(x)$ versus transverse field strength h and position x for the graphon $W(x, y) = \sqrt{xy}$. Image from [128].

case that G is random we may instead study the behaviour of the same Hamiltonian instead on the ‘average’ weighted graph $\bar{G}$, since the free energies of the two coincide in this limit.

As the simplest example, an Erdős-Rényi graph $G(n, p)$ has corresponding average graph which is the complete graph with edge weights p , which is useful because spin Hamiltonians are in many cases, due to the high symmetry, easier to solve on these latter graphs. Remarkably, since for $p = 0.5$ the ensemble of graphs from which one samples has equal probability attached to every simple graph, this implies that in the thermodynamic limit, if one picks any simple graph then the probability of it having physics of the complete graph—which does not display complex physics—is asymptotically almost surely 1. Therefore, in order to achieve complex physics in this limit it is necessary to pick ‘exceptional’ graphs, or graphs with measure 0 in this limit. In terms of complexity class considerations, this implies that the free energy of $H(G)$ is computable in polynomial time for almost all graphs.

Nevertheless, we can consider other random graph ensembles and their associated graphons. Unexpectedly, even when the graphon is highly irregular—meaning its finite realisations are— we are able to analytically solve for the thermodynamic ground state. This leaves open an important question to address, which is for which choices of graph is the physics in the thermodynamic limit not well approximated by a product state. We have seen that dense graphs, even random dense graphs, are well approximated by MPS simulations due to product state or mean-field descriptions becoming exact in the thermodynamic limit. It is also known that these work well for sparse Hamiltonians with low entanglement, or with certain tree-like geometries which are particularly amenable to such descriptions [133]. Such results imply that non-dense interactions with non-tree-like geometries would be the first candidates to test. As suggested in [133], tensor networks can in fact be used to benchmark which many-body Hamiltonians become difficult for classical simulation techniques.

Chapter 8

Conclusion

8.1 Summary

The first part of this thesis considered the problem of broadening the definition of contextuality to include nonclassical correlations in more general setups—namely, setups in which there is some dependency between measurements. We work within the sheaf approach of Abramsky and Brandenburger, in which contextuality arises when local sections, which are obtained via measurement on a system, cannot be extended to a global section, such global sections corresponding to distributions over deterministic noncontextual hidden variables. In this more general setting, nonclassicality again arises as an obstruction to the extension of local sections to a global section.

In Chapter 5 we studied the special case of temporal measurement scenarios, which consist of a sequence of measurements being performed on a single system. A strategy supplies an outcome for each measurement in a sequence such that, in the most general case, the outcome of a measurement may depend on any of the preceding measurements and outcomes. We study restricted classes of strategies, and define what it means for an empirical model to be classical with respect to these restricted strategies. We then see that we can get a Vorob'ev-type result for each subclass of strategies, which identifies whether the setup can host any nonclassical correlations which obey the same no-signalling constraints as the classical models (for a choice of subclass of strategies), albeit this holds only when past measurements are remembered (and not outcomes). We analyse the case in which outcomes are stored, showing that one can use Θ -consistent strategies to mimic strategies that store outcomes, resulting in a trade-off between the cardinality of

the output set and the cardinality of the number of measurements the Θ -function can store. Finally, we show that one direction of a Vorob'evtype result can be found for lk_k^e -nonclassicality.

In Chapter 6 we studied general causal setups. This involved introducing Experimenter and Nature cells, so that the filling of an Experimenter cell corresponds to a choice of measurement being made and the choice of Nature cell to an outcome being assigned to the measurement. Causal flow is constrained by an enabling relation. We show that this framework is able to capture many special cases of interest, including those of Gogioso and Pinzani for causal Bell scenarios [66], contextuality scenarios, and the temporal scenarios of the previous chapter. We showed moreover that adaptive experimenter strategies allow us to capture also certain Bayesian networks, and discussed how such adaptivity, by reducing the observed measurement contexts, reduces the likelihood that an obstruction to the extension of an empirical model is observed.

Chapter 7 shifted the focus of study to many body Hamiltonians, and specifically those of n spins with interaction pattern described by a graph G . We in particular studied the thermodynamic limit of these systems, making use of the graphon or graph limit of the sequence (G_i) of graphs as $i \rightarrow \infty$. We show that, using a mean field ansatz, the Physics in this limit (i.e. the observables of the system) is captured by three coupled equations, and given by a product state. Such results are reminiscent of those in [25], which show that Hamiltonians dense graphs can often be solved with product state ansatzes.

8.2 Future research directions

8.2.1 Vorob'ev-type results for causal measurement scenarios

In Chapter 5 we made use of Vorob'ev's Theorem in the context of temporal correlations, these correlations are by assumption arising from repeated measurement of a single system. This was done by mapping such setups to contextuality setups, in a way such that classicality is preserved. Two further developments of this line of research could be

- i The development of a Graham reduction process which respect the temporal ordering of measurements, so that one could obtain a Vorob'ev-type result

on the temporal measurement scenario directly. For the related notion of collapsibility, there has been some work in this direction by Belton et al. [22].

- ii The extension of such observations to causal measurement scenarios. This could be a generalisation for presheaves over covers $\mathcal{C}$ which obey the ‘generalised simplicial complex’ axioms which were given in Chapter 6.

8.2.2 Using Θ -consistent strategies to realise outcome memory

At the end of Chapter 5 we analysed the trade-off between using (deterministic) Θ -consistent strategies and a strategy in which the outcome of a measurement can depend only on the previous outcome. In particular, we saw that the number of measurements the Θ -strategy must store is linear in the size of the output set. It would be interesting to derive such bounds more generally, and so to understand which Θ -consistent strategies can always realise a $\text{lk}_k^{(e)}$ -consistent strategy for some choice of k . Doing so would give an understanding of how ‘powerful’, so to speak, the Θ -functions are as models of memory.

Related to this is understanding better the resource which is at play here. While in [31] it is the ‘number of perfectly distinguishable states’ which is the relevant resource, here it is the amount of information, or number of bits, which are signalled. It would be informative to more explicitly consider this point of view.

8.2.3 Complexity classifications of spin systems on graph in the thermodynamic limit

Our treatment of the thermodynamic limit of spin Hamiltonians resulted in a set of coupled integral equations which contain the graphon, or graph limit, of the dense graphs of study. In certain cases, we see that these equations are analytically solvable, thereby giving an exact expression for the ground state in terms of a product of spin wavefunctions. Brandão and Harrow [25] have demonstrated that for a large class of Hamiltonians, a solution to the k -local Hamiltonian problem is well approximated by product (i.e. unentangled) states. In particular, this applies to 2-local Hamiltonians with high degree, which is the exact type of Hamiltonian we consider, which agrees with our mean-field derivation (a product state ansatz is implicitly taking a mean-field approach). Although the general local Hamiltonian problem is QMA-complete [87], the class of Hamiltonians considered in [25] is in NP. Our results imply that in the thermodynamic limit, only graphs with measure 0 in

the space of all simple graphs lie outside of $\mathcal{P}$. For this reason it would be interesting to better understand the implications of our work in the realm of complexity theory classifications, which might give interesting candidate graphs for which the mean-field approach breaks down.

Bibliography

- [1] S. Abramsky and P.-A. Mellies. “Concurrent games and full completeness”. In: **Proceedings. 14th Symposium on Logic in Computer Science (Cat. No. PR00158)**. 1999, pp. 431–442. DOI: [10 . 1109 / LICS . 1999 . 782638](https://doi.org/10.1109/LICS.1999.782638).
- [2] Samson Abramsky. “Classical Logic, Classical Probability, and Quantum Mechanics”. In: **Quantum, Probability, Logic: The Work and Influence of Itamar Pitowsky**. Ed. by Meir Hemmo and Orly Shenker. Cham: Springer International Publishing, 2020, pp. 1–17. ISBN: 978-3-030-34316-3. DOI: [10 . 1007 / 978-3-030-34316-3_1](https://doi.org/10.1007/978-3-030-34316-3_1). URL: https://doi.org/10.1007/978-3-030-34316-3_1.
- [3] Samson Abramsky, Rui Soares Barbosa, Martti Karvonen and Shane Mansfield. “A comonadic view of simulation and quantum resources”. In: **2019 34th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)**. IEEE. 2019, pp. 1–12.
- [4] Samson Abramsky, Rui Soares Barbosa and Shane Mansfield. “Contextual Fraction as a Measure of Contextuality”. In: **Physical Review Letters** 119.5 (Aug. 2017). ISSN: 1079-7114. DOI: [10 . 1103 / physrevlett . 119 . 050504](https://doi.org/10.1103/physrevlett.119.050504). URL: <http://dx.doi.org/10.1103/PhysRevLett.119.050504>.
- [5] Samson Abramsky, Rui Soares Barbosa and Amy Searle. “Combining contextuality and causality: a game semantics approach”. In: **Philosophical Transactions of the Royal Society A: Mathematical, Physical and**

- Engineering Sciences** 382.2268 (2024), p. 20230002. DOI: [10.1098/rsta.2023.0002](https://doi.org/10.1098/rsta.2023.0002). eprint: <https://royalsocietypublishing.org/doi/pdf/10.1098/rsta.2023.0002>. URL: <https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2023.0002>.
- [6] Samson Abramsky and Adam Brandenburger. “The sheaf-theoretic structure of non-locality and contextuality”. In: **New Journal of Physics** 13.11 (2011), p. 113036.
 - [7] Samson Abramsky and Adam Brandenburger. “The sheaf-theoretic structure of non-locality and contextuality”. In: **New Journal of Physics** 13.11 (Nov. 2011), p. 113036. ISSN: 1367-2630. DOI: [10.1088/1367-2630/13/11/113036](https://doi.org/10.1088/1367-2630/13/11/113036). URL: <http://dx.doi.org/10.1088/1367-2630/13/11/113036>.
 - [8] Samson Abramsky and Lucien Hardy. “Logical Bell inequalities”. In: **Physical Review A** 85.6 (June 2012). ISSN: 1094-1622. DOI: [10.1103/physreva.85.062114](https://doi.org/10.1103/physreva.85.062114). URL: <http://dx.doi.org/10.1103/PhysRevA.85.062114>.
 - [9] Samson Abramsky, Shane Mansfield and Rui Soares Barbosa. “The cohomology of non-locality and contextuality”. In: **8th International Workshop on Quantum Physics and Logic (QPL 2011)**. Ed. by Bart Jacobs, Peter Selinger and Bas Spitters. Vol. 95. Electronic Proceedings in Theoretical Computer Science. 2012, pp. 1–14. DOI: [10.4204/EPTCS.95.1](https://doi.org/10.4204/EPTCS.95.1).
 - [10] Janet Anders and Dan E. Browne. “Computational Power of Correlations”. In: **Physical Review Letters** 102.5 (Feb. 2009). ISSN: 1079-7114. DOI: [10.1103/physrevlett.102.050502](https://doi.org/10.1103/physrevlett.102.050502). URL: <http://dx.doi.org/10.1103/PhysRevLett.102.050502>.
 - [11] Janet Anders and Dan E Browne. “Computational power of correlations”. In: **Physical Review Letters** 102.5 (2009), p. 050502.

- [12] S. Arora and B. Barak. **Computational Complexity: A Modern Approach**. Cambridge University Press, 2006. ISBN: 978-0-521-42426-4. URL: <https://theory.cs.princeton.edu/complexity/book.pdf>.
- [13] Assa Auerbach. “Spin Wave Theory”. In: **Interacting Electrons and Quantum Magnetism**. New York, NY: Springer New York, 1994, pp. 113–127. ISBN: 978-1-4612-0869-3. DOI: [10.1007/978-1-4612-0869-3_11](https://doi.org/10.1007/978-1-4612-0869-3_11). URL: https://doi.org/10.1007/978-1-4612-0869-3_11.
- [14] Rui Barbosa. “Contextuality in quantum mechanics and beyond”. 2015.
- [15] Rui Soares Barbosa. “On monogamy of non-locality and macroscopic averages: examples and preliminary results”. In: **11th International Workshop on Quantum Physics and Logic (QPL 2014)**. Ed. by Bob Coecke, Ichiro Hasuo and Prakash Panangaden. Vol. 172. Electronic Proceedings in Theoretical Computer Science. Open Publishing Association, 2014, pp. 36–55. DOI: [10.4204/eptcs.172.4](https://doi.org/10.4204/eptcs.172.4).
- [16] Rui Soares Barbosa, Martti Karvonen and Shane Mansfield. “Closing Bell Boxing Black Box Simulations in the Resource Theory of Contextuality”. In: **Outstanding Contributions to Logic**. Springer International Publishing, 2023, 475–529. ISBN: 9783031241178. DOI: [10.1007/978-3-031-24117-8_13](https://doi.org/10.1007/978-3-031-24117-8_13). URL: http://dx.doi.org/10.1007/978-3-031-24117-8_13.
- [17] Jonathan Barrett. “Information processing in generalized probabilistic theories”. In: **Phys. Rev. A** 75 (3 2007), p. 032304. DOI: [10.1103/PhysRevA.75.032304](https://doi.org/10.1103/PhysRevA.75.032304). URL: <https://link.aps.org/doi/10.1103/PhysRevA.75.032304>.
- [18] Rodney J Baxter. “Partition function of the Eight-Vertex lattice model”. In: **Annals of Physics** 70.1 (1972), pp. 193–228. ISSN: 0003-4916. DOI: [https://doi.org/10.1016/0003-4916\(72\)90335-1](https://doi.org/10.1016/0003-4916(72)90335-1). URL: <https://www.sciencedirect.com/science/article/pii/0003491672903351>.

- [19] Kerstin Beer and Tobias J. Osborne. “Contextuality and bundle diagrams”. In: **Phys. Rev. A** 98 (5 2018), p. 052124. DOI: [10.1103/PhysRevA.98.052124](https://doi.org/10.1103/PhysRevA.98.052124). URL: <https://link.aps.org/doi/10.1103/PhysRevA.98.052124>.
- [20] Catriel Beeri, Ronald Fagin, David Maier and Mihalis Yannakakis. “On the Desirability of Acyclic Database Schemes”. In: **J. ACM** 30.3 (1983), 479–513. ISSN: 0004-5411. DOI: [10.1145/2402.322389](https://doi.org/10.1145/2402.322389). URL: <https://doi.org/10.1145/2402.322389>.
- [21] J. S. Bell. “On the Einstein Podolsky Rosen paradox”. In: **Physics Physique Fizika** 1 (3 1964), pp. 195–200. DOI: [10.1103/PhysicsPhysiqueFizika.1.195](https://link.aps.org/doi/10.1103/PhysicsPhysiqueFizika.1.195). URL: <https://link.aps.org/doi/10.1103/PhysicsPhysiqueFizika.1.195>.
- [22] Robin Belton, Robyn Brooks, Stefania Ebli, Lisbeth Fajstrup, Brittany Terese Fasy, Nicole Sanderson and Elizabeth Vidaurre. **Combinatorial Conditions for Directed Collapsing**. 2022. arXiv: [2106.01524 \[math.AT\]](https://arxiv.org/abs/2106.01524).
- [23] H. A. Bethe. “Zur Theorie der Metalle. I. Eigenwerte und Eigenfunktionen der linearen Atomkette”. In: **Zeit. für Physik** 71 (1931), p. 205. DOI: [10.1007/BF01341708](https://doi.org/10.1007/BF01341708).
- [24] Anton Bovier and Véronique Gayrard. “The thermodynamics of the Curie-Weiss model with random couplings”. In: **Journal of Statistical Physics** 72.3 (1993), pp. 643–664. DOI: [10.1007/BF01048027](https://doi.org/10.1007/BF01048027). URL: <https://doi.org/10.1007/BF01048027>.
- [25] Fernando G. S. L. Brandão and Aram W. Harrow. “Product-State Approximations to Quantum States”. In: **Communications in Mathematical Physics** 342.1 (2016), pp. 47–80. DOI: [10.1007/s00220-016-2575-1](https://doi.org/10.1007/s00220-016-2575-1). URL: <https://doi.org/10.1007/s00220-016-2575-1>.
- [26] Sergey Bravyi. **Efficient algorithm for a quantum analogue of 2-SAT**. 2006. arXiv: [quant-ph/0602108 \[quant-ph\]](https://arxiv.org/abs/quant-ph/0602108).

- [27] Caslav Brukner, Samuel Taylor, Sancho Cheung and Vlatko Vedral. **Quantum Entanglement in Time**. 2004. arXiv: [quant-ph/0402127](https://arxiv.org/abs/quant-ph/0402127) [quant-ph].
- [28] Costantino Budroni. “Contextuality, memory cost and non-classicality for sequential measurements”. In: **Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences** 377.2157 (2019), p. 20190141. DOI: [10.1098/rsta.2019.0141](https://doi.org/10.1098/rsta.2019.0141). eprint: <https://royalsocietypublishing.org/doi/pdf/10.1098/rsta.2019.0141>. URL: <https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2019.0141>.
- [29] Costantino Budroni, Adán Cabello, Otfried Gühne, Matthias Kleinmann and Jan-Åke Larsson. “Kochen-Specker contextuality”. In: **Reviews of Modern Physics** 94.4 (Dec. 2022). ISSN: 1539-0756. DOI: [10.1103/revmodphys.94.045007](https://doi.org/10.1103/revmodphys.94.045007). URL: <http://dx.doi.org/10.1103/RevModPhys.94.045007>.
- [30] Costantino Budroni and Clive Emary. “Temporal Quantum Correlations and Leggett-Garg Inequalities in Multilevel Systems”. In: **Physical Review Letters** 113.5 (July 2014). ISSN: 1079-7114. DOI: [10.1103/physrevlett.113.050401](https://doi.org/10.1103/physrevlett.113.050401). URL: <http://dx.doi.org/10.1103/PhysRevLett.113.050401>.
- [31] Costantino Budroni, Gabriel Fagundes and Matthias Kleinmann. “Memory cost of temporal correlations”. In: **New Journal of Physics** 21.9 (2019), p. 093018. DOI: [10.1088/1367-2630/ab3cb4](https://doi.org/10.1088/1367-2630/ab3cb4). URL: <https://doi.org/10.1088%2F1367-2630%2Fab3cb4>.
- [32] Costantino Budroni and Giovanni Morchio. “The extension problem for partial Boolean structures in quantum mechanics”. In: **Journal of Mathematical Physics** 51.12 (Dec. 2010), p. 122205. ISSN: 0022-2488. DOI: [10.1063/1.3523478](https://doi.org/10.1063/1.3523478). eprint: [https://pubs.aip.org/aip/jmp/article-pdf/doi/10.1063/1.3523478/13353322/122205_1_online.pdf](https://pubs.aip.org/aip/jmp/article-pdf/doi/10.1063/1.3523478/13353322/122205%5B1%5D_online.pdf). URL: <https://doi.org/10.1063/1.3523478>.

- [33] Costantino Budroni, Tobias Moroder, Matthias Kleinmann and Otfried Gühne. “Bounding Temporal Quantum Correlations”. In: **Phys. Rev. Lett.** 111 (2 2013), p. 020403. DOI: [10.1103/PhysRevLett.111.020403](https://doi.org/10.1103/PhysRevLett.111.020403). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.111.020403>.
- [34] Adán Cabello. “Experimentally Testable State-Independent Quantum Contextuality”. In: **Phys. Rev. Lett.** 101 (21 2008), p. 210401. DOI: [10.1103/PhysRevLett.101.210401](https://doi.org/10.1103/PhysRevLett.101.210401). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.101.210401>.
- [35] Adán Cabello, JoséM. Estebaranz and Guillermo García-Alcaine. “Bell-Kochen-Specker theorem: A proof with 18 vectors”. In: **Physics Letters A** 212.4 (1996), pp. 183–187. ISSN: 0375-9601. DOI: [https://doi.org/10.1016/0375-9601\(96\)00134-X](https://doi.org/10.1016/0375-9601(96)00134-X). URL: <https://www.sciencedirect.com/science/article/pii/037596019600134X>.
- [36] Adán Cabello, Simone Severini and Andreas Winter. “Graph-Theoretic Approach to Quantum Correlations”. In: **Physical Review Letters** 112.4 (Jan. 2014). ISSN: 1079-7114. DOI: [10.1103/physrevlett.112.040401](https://doi.org/10.1103/physrevlett.112.040401). URL: <http://dx.doi.org/10.1103/PhysRevLett.112.040401>.
- [37] Chris Cade and P. Marcos Crichigno. “Complexity of Supersymmetric Systems and the Cohomology Problem”. In: **Quantum** 8 (Apr. 2024), p. 1325. ISSN: 2521-327X. DOI: [10.22331/q-2024-04-30-1325](https://doi.org/10.22331/q-2024-04-30-1325). URL: <http://dx.doi.org/10.22331/q-2024-04-30-1325>.
- [38] Paulo R. A. Campos, Viviane M. de Oliveira and F. G. Brady Moreira. “Small-world effects in the majority-vote model”. In: **Phys. Rev. E** 67 (2 2003), p. 026104. DOI: [10.1103/PhysRevE.67.026104](https://doi.org/10.1103/PhysRevE.67.026104). URL: <https://link.aps.org/doi/10.1103/PhysRevE.67.026104>.
- [39] Eric Chitambar and Gilad Gour. “Quantum resource theories”. In: **Reviews of Modern Physics** 91.2 (Apr. 2019). ISSN: 1539-0756. DOI: [10.1103/RevModPhys.91.021001](https://doi.org/10.1103/RevModPhys.91.021001).

-
- [revmodphys.91.025001](#). URL: <http://dx.doi.org/10.1103/RevModPhys.91.025001>.
- [40] Joonhee Choi, Hengyun Zhou, Helena S. Knowles, Renate Landig, Soonwon Choi and Mikhail D. Lukin. “Robust Dynamic Hamiltonian Engineering of Many-Body Spin Systems”. In: **Phys. Rev. X** 10 (3 2020), p. 031002. DOI: [10.1103/PhysRevX.10.031002](#). URL: <https://link.aps.org/doi/10.1103/PhysRevX.10.031002>.
- [41] Raman Choudhary, Rui Soares Barbosa and Adán Cabello. **Lifting non-contextuality inequalities**. 2024. arXiv: [2401.12349 \[quant-ph\]](#).
- [42] John F. Clauser, Michael A. Horne, Abner Shimony and Richard A. Holt. “Proposed Experiment to Test Local Hidden-Variable Theories”. In: **Phys. Rev. Lett.** 23 (15 1969), pp. 880–884. DOI: [10.1103/PhysRevLett.23.880](#). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.23.880>.
- [43] Lucas Clemente and Johannes Kofler. “No Fine Theorem for Macrorealism: Limitations of the Leggett-Garg Inequality”. In: **Phys. Rev. Lett.** 116 (15 2016), p. 150401. DOI: [10.1103/PhysRevLett.116.150401](#). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.116.150401>.
- [44] Stephen A. Cook. “The complexity of theorem-proving procedures”. In: **Proceedings of the third annual ACM symposium on Theory of computing** (1971). URL: <https://api.semanticscholar.org/CorpusID:7573663>.
- [45] Gemma De las Cuevas and Toby S. Cubitt. “Simple universal models capture all classical spin physics”. In: **Science** 351.6278 (Mar. 2016), 1180–1183. ISSN: 1095-9203. DOI: [10.1126/science.aab3326](#). URL: <http://dx.doi.org/10.1126/science.aab3326>.
- [46] Justin Curry. **Sheaves, Cosheaves and Applications**. 2014. arXiv: [1303.3255 \[math.AT\]](#).

- [47] E. B. Davies and J. T. Lewis. “An operational approach to quantum probability”. In: **Communications in Mathematical Physics** 17.3 (1970), pp. 239–260. DOI: [10.1007/BF01647093](https://doi.org/10.1007/BF01647093). URL: <https://doi.org/10.1007/BF01647093>.
- [48] Amit Dutta, Gabriel Aeppli, Bikas K. Chakrabarti, Uma Divakaran, Thomas F. Rosenbaum and Diptiman Sen. **Quantum phase transitions in transverse field spin models: from statistical physics to quantum information**. 2015. arXiv: [1012.0653 \[cond-mat.stat-mech\]](https://arxiv.org/abs/1012.0653).
- [49] Ehtibar N Dzhafarov, Janne V Kujala and Victor H Cervantes. “Contextuality-by-default: a brief overview of ideas, concepts, and terminology”. In: **Quantum Interaction: 9th International Conference, QI 2015, Filzbach, Switzerland, July 15-17, 2015, Revised Selected Papers 9**. Springer. 2016, pp. 12–23.
- [50] A. Einstein, B. Podolsky and N. Rosen. “Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?” In: **Phys. Rev.** 47 (10 1935), pp. 777–780. DOI: [10.1103/PhysRev.47.777](https://link.aps.org/doi/10.1103/PhysRev.47.777). URL: <https://link.aps.org/doi/10.1103/PhysRev.47.777>.
- [51] J. Eisert, M. Cramer and M. B. Plenio. “Colloquium: Area laws for the entanglement entropy”. In: **Rev. Mod. Phys.** 82 (1 2010), pp. 277–306. DOI: [10.1103/RevModPhys.82.277](https://link.aps.org/doi/10.1103/RevModPhys.82.277). URL: <https://link.aps.org/doi/10.1103/RevModPhys.82.277>.
- [52] Richard S. Ellis and Charles M. Newman. “The statistics of Curie-Weiss models”. In: **Journal of Statistical Physics** 19.2 (1978), pp. 149–161. DOI: [10.1007/BF01012508](https://doi.org/10.1007/BF01012508). URL: <https://doi.org/10.1007/BF01012508>.
- [53] Clive Emary, Neill Lambert and Franco Nori. “Leggett–Garg inequalities”. In: **Reports on Progress in Physics** 77.1 (Dec. 2013), p. 016001. ISSN:

- 1361-6633. DOI: [10.1088/0034-4885/77/1/016001](https://doi.org/10.1088/0034-4885/77/1/016001). URL: <http://dx.doi.org/10.1088/0034-4885/77/1/016001>.
- [54] Gabriel Fagundes and Matthias Kleinmann. “Memory cost for simulating all quantum correlations from the Peres–Mermin scenario”. In: **Journal of Physics A: Mathematical and Theoretical** 50.32 (2017), p. 325302. DOI: [10.1088/1751-8121/aa7ab3](https://doi.org/10.1088/1751-8121/aa7ab3). URL: <https://dx.doi.org/10.1088/1751-8121/aa7ab3>.
- [55] Florence Dowdell Fasanelli. “THE CREATION OF SHEAF THEORY”. In: (Jan. 1981). DOI: [10.57912/23887509.v1](https://doi.org/10.57912/23887509.v1). URL: https://aura.american.edu/articles/thesis/THE_CREATION_OF_SHEAF_THEORY/23887509.
- [56] Paul Fendley, Kareljan Schoutens and Jan de Boer. “Lattice Models with Supersymmetry”. In: **Physical Review Letters** 90.12 (Mar. 2003). ISSN: 1079-7114. DOI: [10.1103/physrevlett.90.120402](https://doi.org/10.1103/physrevlett.90.120402). URL: <http://dx.doi.org/10.1103/PhysRevLett.90.120402>.
- [57] Irénée Frérot, Piero Naldesi and Tommaso Roscilde. “Entanglement and fluctuations in the XXZ model with power-law interactions”. In: **Phys. Rev. B** 95 (24 2017), p. 245111. DOI: [10.1103/PhysRevB.95.245111](https://doi.org/10.1103/PhysRevB.95.245111). URL: <https://link.aps.org/doi/10.1103/PhysRevB.95.245111>.
- [58] T. Fritz, A. B. Sainz, R. Augusiak, J Bohr Brask, R. Chaves, A. Leverrier and A. Acín. “Local orthogonality as a multipartite principle for quantum correlations”. In: **Nature Communications** 4.1 (2013), p. 2263. DOI: [10.1038/ncomms3263](https://doi.org/10.1038/ncomms3263). URL: <https://doi.org/10.1038/ncomms3263>.
- [59] Tobias Fritz. “Beyond Bell’s Theorem II: Scenarios with Arbitrary Causal Structure”. In: **Communications in Mathematical Physics** 341.2 (Nov. 2015), 391–434. ISSN: 1432-0916. DOI: [10.1007/s00220-015-2495-5](https://doi.org/10.1007/s00220-015-2495-5). URL: <http://dx.doi.org/10.1007/s00220-015-2495-5>.

- [60] Tobias Fritz. “Quantum correlations in the temporal Clauser-Horne-Shimony-Holt (CHSH) scenario”. In: **New Journal of Physics** 12.8, 083055 (Aug. 2010), p. 083055. DOI: [10.1088/1367-2630/12/8/083055](https://doi.org/10.1088/1367-2630/12/8/083055). arXiv: [1005.3421](https://arxiv.org/abs/1005.3421) [[quant-ph](#)].
- [61] Tobias Fritz. “Quantum correlations in the temporal Clauser-Horne-Shimony-Holt (CHSH) scenario”. In: **New Journal of Physics** 12.8 (Aug. 2010), p. 083055. ISSN: 1367-2630. DOI: [10.1088/1367-2630/12/8/083055](https://doi.org/10.1088/1367-2630/12/8/083055). URL: <http://dx.doi.org/10.1088/1367-2630/12/8/083055>.
- [62] Robert Furber and Bart Jacobs. “Towards a Categorical Account of Conditional Probability”. In: **Electronic Proceedings in Theoretical Computer Science** 195 (Nov. 2015), 179–195. ISSN: 2075-2180. DOI: [10.4204/eptcs.195.14](https://doi.org/10.4204/eptcs.195.14). URL: <http://dx.doi.org/10.4204/EPTCS.195.14>.
- [63] Dan Geiger, Tom S. Verma and Judea Pearl. **d-Separation: From Theorems to Algorithms**. 2013. arXiv: [1304.1505](https://arxiv.org/abs/1304.1505) [[cs.AI](#)].
- [64] A.J. Glick, H.J. Lipkin and N. Meshkov. “Validity of many-body approximation methods for a solvable model: (III). Diagram summations”. In: **Nuclear Physics** 62.2 (1965), pp. 211–224. ISSN: 0029-5582. DOI: [https://doi.org/10.1016/0029-5582\(65\)90864-3](https://doi.org/10.1016/0029-5582(65)90864-3). URL: <https://www.sciencedirect.com/science/article/pii/0029558265908643>.
- [65] R. Godement. **Topologie Algebrique Et Theorie Des Faisceaux**. Actualités scientifiques et industrielles. Hermann, 1964.
- [66] Stefano Gogioso and Nicola Pinzani. “The sheaf-theoretic structure of definite causality”. In: **18th International Conference on Quantum Physics and Logic (QPL 2021)**. Ed. by Chris Heunen and Miriam Backens. Vol. 343. Electronic Proceedings in Theoretical Computer Science. Open Publishing Association, 2021, pp. 301–324. DOI: [10.4204/EPTCS.343.13](https://doi.org/10.4204/EPTCS.343.13).

- [67] Stefano Gogioso and Nicola Pinzani. “The Sheaf-Theoretic Structure of Definite Causality”. In: **Electronic Proceedings in Theoretical Computer Science** 343 (2021), pp. 301–324. DOI: [10.4204/eptcs.343.13](https://doi.org/10.4204/eptcs.343.13). URL: <https://doi.org/10.4204/eptcs.343.13>.
- [68] S. M. Griffin, M. Lilienblum, K. T. Delaney, Y. Kumagai, M. Fiebig and N. A. Spaldin. “Scaling Behavior and Beyond Equilibrium in the Hexagonal Manganites”. In: **Phys. Rev. X** 2 (4 2012), p. 041022. DOI: [10.1103/PhysRevX.2.041022](https://link.aps.org/doi/10.1103/PhysRevX.2.041022). URL: <https://link.aps.org/doi/10.1103/PhysRevX.2.041022>.
- [69] A. Grothendieck. **A General Theory of Fibre Spaces: With Structure Sheaf**. National science foundation research project on geometry of function space. University of Kansas, Department of mathematics, 1955.
- [70] Shashank Gupta, Debashis Saha, Zhen-Peng Xu, Adán Cabello and A. S. Majumdar. “Quantum Contextuality Provides Communication Complexity Advantage”. In: **Phys. Rev. Lett.** 130 (8 2023), p. 080802. DOI: [10.1103/PhysRevLett.130.080802](https://link.aps.org/doi/10.1103/PhysRevLett.130.080802). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.130.080802>.
- [71] J. J. Halliwell. “Leggett-Garg inequalities and no-signaling in time: A quasiprobability approach”. In: **Physical Review A** 93.2 (Feb. 2016). ISSN: 2469-9934. DOI: [10.1103/PhysRevA.93.022123](https://doi.org/10.1103/PhysRevA.93.022123). URL: <http://dx.doi.org/10.1103/PhysRevA.93.022123>.
- [72] W. K. Hastings. “Monte Carlo Sampling Methods Using Markov Chains and Their Applications”. eng. In: **Biometrika** 57.1 (1970), pp. 97–109. ISSN: 00063444.
- [73] Joe Henson, Raymond Lal and Matthew F Pusey. “Theory-independent limits on correlations from generalized Bayesian networks”. In: **New Journal of Physics** 16.11 (2014), p. 113043.

- [74] Jannik Hoffmann, Cornelia Spee, Otfried Gühne and Costantino Budroni. “Structure of temporal correlations of a qubit”. In: **New Journal of Physics** 20.10 (2018), p. 102001. DOI: [10.1088/1367-2630/aae87f](https://doi.org/10.1088/1367-2630/aae87f). URL: <https://dx.doi.org/10.1088/1367-2630/aae87f>.
- [75] L. Huijse and K. Schoutens. **Supersymmetry, lattice fermions, independence complexes and cohomology theory**. 2009. arXiv: [0903.0784](https://arxiv.org/abs/0903.0784) [[cond-mat.str-el](https://arxiv.org/archive/math)].
- [76] C.-L. Hung, Alejandro González-Tudela, J. Ignacio Cirac and H. J. Kimble. “Quantum spin dynamics with pairwise-tunable, long-range interactions”. In: **Proceedings of the National Academy of Sciences** 113.34 (Aug. 2016). ISSN: 1091-6490. DOI: [10.1073/pnas.1603777113](https://doi.org/10.1073/pnas.1603777113). URL: <http://dx.doi.org/10.1073/pnas.1603777113>.
- [77] E. Ising. “Beitrag zur Theorie des Ferromagnetismus”. In: **Zeitschrift für Physik** 31 (Feb. 1925), pp. 253–258. DOI: [10.1007/BF02980577](https://doi.org/10.1007/BF02980577).
- [78] R. Islam, C. Senko, W. C. Campbell, S. Korenblit, J. Smith, A. Lee, E. E. Edwards, C.-C. J. Wang, J. K. Freericks and C. Monroe. “Emergence and Frustration of Magnetism with Variable-Range Interactions in a Quantum Simulator”. In: **Science** 340.6132 (2013), pp. 583–587. DOI: [10.1126/science.1232296](https://doi.org/10.1126/science.1232296). URL: <https://www.science.org/doi/abs/10.1126/science.1232296>.
- [79] Bart Jacobs. “A Categorical Basis for Conditional Probability”. In: (February 2016). URL: <http://dx.doi.org/10.4204/EPTCS.195.14>.
- [80] Gilles Kahn and Gordon D. Plotkin. “Concrete domains”. In: **Theoretical Computer Science** 121.1-2 (1993), pp. 187–277.
- [81] Martti Karvonen. “Categories of empirical models”. In: **15th International Conference on Quantum Physics and Logic (QPL 2018)**. Ed. by Peter

- Selinger and Giulio Chiribella. Vol. 287. *Electronic Proceedings in Theoretical Computer Science*. 2019, pp. 239–252. DOI: [10.4204/EPTCS.287.14](https://doi.org/10.4204/EPTCS.287.14).
- [82] Julia Kempe, Alexei Kitaev and Oded Regev. “The Complexity of the Local Hamiltonian Problem”. In: **SIAM Journal on Computing** 35.5 (2006), pp. 1070–1097. DOI: [10.1137/S0097539704445226](https://doi.org/10.1137/S0097539704445226). eprint: <https://doi.org/10.1137/S0097539704445226>. URL: <https://doi.org/10.1137/S0097539704445226>.
- [83] Julia Kempe and Oded Regev. **3-Local Hamiltonian is QMA-complete**. 2003. arXiv: [quant-ph/0302079](https://arxiv.org/abs/quant-ph/0302079) [quant-ph].
- [84] Dante M. Kennes, Martin Claassen, Lede Xian, Antoine Georges, Andrew J. Millis, James Hone, Cory R. Dean, D. N. Basov, Abhay N. Pasupathy and Angel Rubio. “Moiréheterostructures as a condensed-matter quantum simulator”. In: **Nature Physics** 17.2 (2021), pp. 155–163. DOI: [10.1038/s41567-020-01154-3](https://doi.org/10.1038/s41567-020-01154-3). URL: <https://doi.org/10.1038/s41567-020-01154-3>.
- [85] Youngseok Kim, Andrew Eddins, Sajant Anand, Ken Xuan Wei, Ewout van den Berg, Sami Rosenblatt, Hasan Nayfeh, Yantao Wu, Michael Zaletel, Kristan Temme and Abhinav Kandala. “Evidence for the utility of quantum computing before fault tolerance”. In: **Nature** 618.7965 (2023), pp. 500–505. DOI: [10.1038/s41586-023-06096-3](https://doi.org/10.1038/s41586-023-06096-3). URL: <https://doi.org/10.1038/s41586-023-06096-3>.
- [86] A. Kitaev. **A Simple Model Of Quantum Holography**. April 7, 2015 and May 27, 2015. DOI: [10.1103/PhysRevX.2.041022](https://doi.org/10.1103/PhysRevX.2.041022). URL: <http://online.kitp.ucsb.edu/online/entangled15/kitaev/>, <http://online.kitp.ucsb.edu/online/entangled15/kitaev2/>.

- [87] Alexei Y. Kitaev, A. H. Shen and Mikhail N. Vyalyi. “Classical and Quantum Computation”. In: **Graduate Studies in Mathematics**. 2002. URL: <https://api.semanticscholar.org/CorpusID:265878561>.
- [88] Matthias Kleinmann, Otfried Gühne, José R Portillo, Jan Åke Larsson and Adán Cabello. “Memory cost of quantum contextuality”. In: **New Journal of Physics** 13.11 (2011), p. 113011. DOI: [10.1088/1367-2630/13/11/113011](https://doi.org/10.1088/1367-2630/13/11/113011). URL: <https://dx.doi.org/10.1088/1367-2630/13/11/113011>.
- [89] Alexander A. Klyachko, M. Ali Canlu, familyi=u., given=Sinem, giveni=S., , and Alexander S. Shumovsky. “Simple Test for Hidden Variables in Spin-1 Systems”. In: **Phys. Rev. Lett.** 101 (2 2008), p. 020403. DOI: [10.1103/PhysRevLett.101.020403](https://doi.org/10.1103/PhysRevLett.101.020403). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.101.020403>.
- [90] Donald E. Knuth. **The sandwich theorem**. 1993. arXiv: [math/9312214](https://arxiv.org/abs/math/9312214) [[math.CO](https://arxiv.org/abs/math/9312214)]. URL: <https://arxiv.org/abs/math/9312214>.
- [91] Simon Kochen and E. P. Specker. “The Problem of Hidden Variables in Quantum Mechanics”. In: **The Logico-Algebraic Approach to Quantum Mechanics: Volume I: Historical Evolution**. Ed. by C. A. Hooker. Dordrecht: Springer Netherlands, 1975, pp. 293–328. ISBN: 978-94-010-1795-4. DOI: [10.1007/978-94-010-1795-4_17](https://doi.org/10.1007/978-94-010-1795-4_17). URL: https://doi.org/10.1007/978-94-010-1795-4_17.
- [92] M Kochmański, T Paszkiewicz and S Wolski. “Curie–Weiss magnet—a simple model of phase transition”. In: **European Journal of Physics** 34.6 (Oct. 2013), 1555–1573. ISSN: 1361-6404. DOI: [10.1088/0143-0807/34/6/1555](https://doi.org/10.1088/0143-0807/34/6/1555). URL: <http://dx.doi.org/10.1088/0143-0807/34/6/1555>.
- [93] S Korenblit, D Kafri, W C Campbell, R Islam, E E Edwards, Z-X Gong, G-D Lin, L-M Duan, J Kim, K Kim and C Monroe. “Quantum simulation of spin models on an arbitrary lattice with trapped ions”. In: **New Journal of**

- Physics** 14.9 (Sept. 2012), p. 095024. ISSN: 1367-2630. DOI: [10.1088/1367-2630/14/9/095024](https://doi.org/10.1088/1367-2630/14/9/095024). URL: <http://dx.doi.org/10.1088/1367-2630/14/9/095024>.
- [94] José I. Latorre, Román Orús, Enrique Rico and Julien Vidal. “Entanglement entropy in the Lipkin-Meshkov-Glick model”. In: **Physical Review A** 71.6 (June 2005). ISSN: 1094-1622. DOI: [10.1103/physreva.71.064101](https://doi.org/10.1103/physreva.71.064101). URL: <http://dx.doi.org/10.1103/PhysRevA.71.064101>.
- [95] A. J. Leggett and Anupam Garg. “Quantum mechanics versus macroscopic realism: Is the flux there when nobody looks?” In: **Phys. Rev. Lett.** 54 (9 1985), pp. 857–860. DOI: [10.1103/PhysRevLett.54.857](https://doi.org/10.1103/PhysRevLett.54.857). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.54.857>.
- [96] Jean Leray. “L’anneau d’homologie d’une représentation”. In: **Comptes rendus hebdomadaires des séances de l’Académie des Sciences** 222 (1946), 1366–1368.
- [97] Yeong-Cherng Liang, Robert W. Spekkens and Howard M. Wiseman. “Specker’s parable of the overprotective seer: A road to contextuality, nonlocality and complementarity”. In: **Physics Reports** 506.1 (2011), pp. 1–39. ISSN: 0370-1573. DOI: <https://doi.org/10.1016/j.physrep.2011.05.001>. URL: <https://www.sciencedirect.com/science/article/pii/S0370157311001517>.
- [98] Elliott Lieb, Theodore Schultz and Daniel Mattis. “Two soluble models of an antiferromagnetic chain”. In: **Annals of Physics** 16.3 (1961), pp. 407–466. ISSN: 0003-4916. DOI: [https://doi.org/10.1016/0003-4916\(61\)90115-4](https://doi.org/10.1016/0003-4916(61)90115-4). URL: <https://www.sciencedirect.com/science/article/pii/0003491661901154>.
- [99] H.J. Lipkin, N. Meshkov and A.J. Glick. “Validity of many-body approximation methods for a solvable model: (I). Exact solutions and perturbation theory”. In: **Nuclear Physics** 62.2 (1965), pp. 188–198. ISSN: 0029-5582.

- DOI: [https://doi.org/10.1016/0029-5582\(65\)90862-X](https://doi.org/10.1016/0029-5582(65)90862-X). URL: <https://www.sciencedirect.com/science/article/pii/002955826590862X>.
- [100] Matteo Lostaglio and Gabriel Senno. “Contextual advantage for state-dependent cloning”. In: **Quantum** 4 (Apr. 2020), p. 258. ISSN: 2521-327X. DOI: [10.22331/q-2020-04-27-258](https://doi.org/10.22331/q-2020-04-27-258). URL: <http://dx.doi.org/10.22331/q-2020-04-27-258>.
- [101] L. Lovasz. “On the Shannon capacity of a graph”. In: **IEEE Transactions on Information Theory** 25.1 (1979), pp. 1–7. DOI: [10.1109/TIT.1979.1055985](https://doi.org/10.1109/TIT.1979.1055985).
- [102] László Lovász and Balázs Szegedy. “Limits of dense graph sequences”. In: **Journal of Combinatorial Theory, Series B** 96.6 (2006), pp. 933–957. ISSN: 0095-8956. DOI: <https://doi.org/10.1016/j.jctb.2006.05.002>. URL: <https://www.sciencedirect.com/science/article/pii/S0095895606000517>.
- [103] 1953 Maier David. **The theory of relational databases**. Computer software engineering series. Rockville, Md.: Computer Science Press, 1983. Chap. xv, 637 pages : illustrations ; 24 cm. ISBN: 0914894420; 9780914894421.
- [104] Shane Mansfield. **A unified approach to contextuality and violations of macrorealism**. https://www1.psych.purdue.edu/~ehtibar/workshop/files/leacht_prague.pdf. 2017.
- [105] Shane Mansfield. **A unified approach to contextuality and violations of macrorealism**. lecture given at QCQMB workshop, Prague. 2017.
- [106] N. David Mermin. “Hidden variables and the two theorems of John Bell”. In: **Rev. Mod. Phys.** 65 (3 1993), pp. 803–815. DOI: [10.1103/RevModPhys.65.803](https://doi.org/10.1103/RevModPhys.65.803). URL: <https://link.aps.org/doi/10.1103/RevModPhys.65.803>.

- [107] N. David Mermin. “Simple unified form for the major no-hidden-variables theorems”. In: **Phys. Rev. Lett.** 65 (27 1990), pp. 3373–3376. DOI: [10.1103/PhysRevLett.65.3373](https://doi.org/10.1103/PhysRevLett.65.3373). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.65.3373>.
- [108] N. Meshkov, A.J. Glick and H.J. Lipkin. “Validity of many-body approximation methods for a solvable model: (II). Linearization procedures”. In: **Nuclear Physics** 62.2 (1965), pp. 199–210. ISSN: 0029-5582. DOI: [https://doi.org/10.1016/0029-5582\(65\)90863-1](https://doi.org/10.1016/0029-5582(65)90863-1). URL: <https://www.sciencedirect.com/science/article/pii/0029558265908631>.
- [109] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller and Edward Teller. “Equation of State Calculations by Fast Computing Machines”. eng. In: **The Journal of Chemical Physics** 21.6 (1953), pp. 1087–1092. ISSN: 0021-9606.
- [110] C. Monroe, W. C. Campbell, L.-M. Duan, Z.-X. Gong, A. V. Gorshkov, P. W. Hess, R. Islam, K. Kim, N. M. Linke, G. Pagano, P. Richerme, C. Senko and N. Y. Yao. “Programmable quantum simulations of spin systems with trapped ions”. In: **Rev. Mod. Phys.** 93 (2 2021), p. 025001. DOI: [10.1103/RevModPhys.93.025001](https://doi.org/10.1103/RevModPhys.93.025001). URL: <https://link.aps.org/doi/10.1103/RevModPhys.93.025001>.
- [111] Ashley Montanaro. “Quantum algorithms: an overview”. In: **npj Quantum Information** 2.1 (2016), p. 15023. DOI: [10.1038/npjqi.2015.23](https://doi.org/10.1038/npjqi.2015.23). URL: <https://doi.org/10.1038/npjqi.2015.23>.
- [112] Michael A. Nielsen and Isaac L. Chuang. **Quantum Computation and Quantum Information: 10th Anniversary Edition**. Cambridge University Press, 2010.
- [113] Lars Onsager. “Crystal Statistics. I. A Two-Dimensional Model with an Order-Disorder Transition”. In: **Phys. Rev.** 65 (3-4 1944), pp. 117–149. DOI:

- 10.1103/PhysRev.65.117. URL: <https://link.aps.org/doi/10.1103/PhysRev.65.117>.
- [114] Judea Pearl. **Causality. Models, Reasoning, and Inference**. 2nd ed. Cambridge, UK: Cambridge University Press, 2009. ISBN: 978-0-521-89560-6. DOI: [10.1017/CB09780511803161](https://doi.org/10.1017/CB09780511803161).
- [115] Asher Peres. “Incompatible results of quantum measurements”. In: **Physics Letters A** 151.3 (1990), pp. 107–108. ISSN: 0375-9601. DOI: [https://doi.org/10.1016/0375-9601\(90\)90172-K](https://doi.org/10.1016/0375-9601(90)90172-K). URL: <https://www.sciencedirect.com/science/article/pii/037596019090172K>.
- [116] Pierre Pfeuty. “The one-dimensional Ising model with a transverse field”. In: **Annals of Physics** 57.1 (1970), pp. 79–90. ISSN: 0003-4916. DOI: [https://doi.org/10.1016/0003-4916\(70\)90270-8](https://doi.org/10.1016/0003-4916(70)90270-8). URL: <https://www.sciencedirect.com/science/article/pii/0003491670902708>.
- [117] Itamar Pitowsky. “George Boole’s ‘Conditions of Possible Experience’ and the Quantum Puzzle”. In: **The British Journal for the Philosophy of Science** 45.1 (1994), pp. 95–125. ISSN: 00070882, 14643537. URL: <http://www.jstor.org/stable/687963> (visited on 07/05/2024).
- [118] Ravishankar Ramanathan, Akihito Soedaski, familyi=n., given=Paweł, giveni=P., ,, and Dagomir Kaszlikowski. “Generalized Monogamy of Contextual Inequalities from the No-Disturbance Principle”. In: **Phys. Rev. Lett.** 109 (5 2012), p. 050404. DOI: [10.1103/PhysRevLett.109.050404](https://link.aps.org/doi/10.1103/PhysRevLett.109.050404). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.109.050404>.
- [119] Robert Raussendorf. “Contextuality in measurement-based quantum computation”. In: **Physical Review A** 88.2 (Aug. 2013). ISSN: 1094-1622. DOI: [10.1103/PhysRevA.88.022322](https://doi.org/10.1103/PhysRevA.88.022322). URL: <http://dx.doi.org/10.1103/PhysRevA.88.022322>.

- [120] Robert Raussendorf, Daniel E. Browne and Hans J. Briegel. “Measurement-based quantum computation on cluster states”. In: **Physical Review A** 68.2 (Aug. 2003). ISSN: 1094-1622. DOI: [10.1103/physreva.68.022312](https://doi.org/10.1103/physreva.68.022312). URL: <http://dx.doi.org/10.1103/PhysRevA.68.022312>.
- [121] Marc-Olivier Renou, Elisa Bäumer, Sadra Boreiri, Nicolas Brunner, Nicolas Gisin and Salman Beigi. “Genuine Quantum Nonlocality in the Triangle Network”. In: **Phys. Rev. Lett.** 123 (14 2019), p. 140401. DOI: [10.1103/PhysRevLett.123.140401](https://doi.org/10.1103/PhysRevLett.123.140401). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.123.140401>.
- [122] Pedro Ribeiro, Julien Vidal and Rémy Mosseri. “Exact spectrum of the Lipkin-Meshkov-Glick model in the thermodynamic limit and finite-size corrections”. In: **Physical Review E** 78.2 (Aug. 2008). ISSN: 1550-2376. DOI: [10.1103/physreve.78.021106](https://doi.org/10.1103/physreve.78.021106). URL: <http://dx.doi.org/10.1103/PhysRevE.78.021106>.
- [123] Subir Sachdev and Jinwu Ye. “Gapless spin-fluid ground state in a random quantum Heisenberg magnet”. In: **Phys. Rev. Lett.** 70 (21 1993), pp. 3339–3342. DOI: [10.1103/PhysRevLett.70.3339](https://doi.org/10.1103/PhysRevLett.70.3339). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.70.3339>.
- [124] J. Sakarovitch and R. Thomas. **Elements of Automata Theory**. Cambridge University Press, 2009. ISBN: 9781139643795. URL: <https://books.google.com.ar/books?id=2RX3AQAQBAJ>.
- [125] M. Schmidt. **Hommes de science: 28 portraits**. Hermann, 1990. ISBN: 9782705661243.
- [126] U. Schollwöck, J. Richter, D.J.J. Farnell and R.F. Bishop. **Quantum Magnetism**. Lecture Notes in Physics. Springer Berlin Heidelberg, 2008. ISBN: 9783540400660.

- [127] Ulrich Schollwöck. “The density-matrix renormalization group in the age of matrix product states”. In: **Annals of Physics** 326.1 (2011). January 2011 Special Issue, pp. 96–192. ISSN: 0003-4916. DOI: <https://doi.org/10.1016/j.aop.2010.09.012>. URL: <https://www.sciencedirect.com/science/article/pii/S0003491610001752>.
- [128] Amy Searle and Joseph Tindall. “Thermodynamic limit of spin systems on random graphs”. In: **Phys. Rev. Res.** 6 (1 2024), p. 013011. DOI: [10.1103/PhysRevResearch.6.013011](https://doi.org/10.1103/PhysRevResearch.6.013011). URL: <https://link.aps.org/doi/10.1103/PhysRevResearch.6.013011>.
- [129] M. P. Seevinck. **E. Specker: ”The logic of non-simultaneously decidable propositions” (1960)**. 2011. arXiv: [1103.4537 \[physics.hist-ph\]](https://arxiv.org/abs/1103.4537).
- [130] Cornelia Spee, Costantino Budroni and Otfried Gühne. “Simulating extremal temporal correlations”. In: **New Journal of Physics** 22.10 (2020), p. 103037. DOI: [10.1088/1367-2630/abb899](https://doi.org/10.1088/1367-2630/abb899). URL: <https://dx.doi.org/10.1088/1367-2630/abb899>.
- [131] R. W. Spekkens. “Contextuality for preparations, transformations, and unsharp measurements”. In: **Physical Review A** 71.5 (May 2005). ISSN: 1094-1622. DOI: [10.1103/physreva.71.052108](https://doi.org/10.1103/physreva.71.052108). URL: <http://dx.doi.org/10.1103/PhysRevA.71.052108>.
- [132] K Takano, K Kubo and H Sakamoto. “Ground states with cluster structures in a frustrated Heisenberg chain”. In: **Journal of Physics: Condensed Matter** 8.35 (1996), p. 6405. DOI: [10.1088/0953-8984/8/35/009](https://doi.org/10.1088/0953-8984/8/35/009). URL: <https://dx.doi.org/10.1088/0953-8984/8/35/009>.
- [133] Joseph Tindall, Matthew Fishman, E. Miles Stoudenmire and Dries Sels. “Efficient Tensor Network Simulation of IBM’s Eagle Kicked Ising Experiment”. In: **PRX Quantum** 5.1 (Jan. 2024). ISSN: 2691-3399. DOI: [10.1103/PRXQuantum.5.011001](https://doi.org/10.1103/PRXQuantum.5.011001).

-
- [prxquantum.5.010308](https://doi.org/10.1103/PRXQuantum.5.010308). URL: <http://dx.doi.org/10.1103/PRXQuantum.5.010308>.
- [134] Joseph Tindall, Amy Searle, Abdulla Alhajri and Dieter Jaksch. “Quantum physics in connected worlds”. In: **Nature Communications** 13.1 (2022), p. 7445. DOI: [10.1038/s41467-022-35090-y](https://doi.org/10.1038/s41467-022-35090-y). URL: <https://doi.org/10.1038/s41467-022-35090-y>.
- [135] Thomas Van Himbeeck, Jonatan Bohr Brask, Stefano Pironio, Ravishankar Ramanathan, Ana Belén Sainz and Elie Wolfe. “Quantum violations in the Instrumental scenario and their relations to the Bell scenario”. In: **Quantum** 3 (Sept. 2019), p. 186. ISSN: 2521-327X. DOI: [10.22331/q-2019-09-16-186](https://doi.org/10.22331/q-2019-09-16-186). URL: <http://dx.doi.org/10.22331/q-2019-09-16-186>.
- [136] Lucas B. Vieira and Costantino Budroni. “Temporal correlations in the simplest measurement sequences”. In: **Quantum** 6 (2022), p. 623. DOI: [10.22331/q-2022-01-18-623](https://doi.org/10.22331/q-2022-01-18-623). URL: <https://doi.org/10.22331/q-2022-01-18-623>.
- [137] N. N. Vorob’ev. “Consistent Families of Measures and Their Extensions”. In: **Theory of Probability & Its Applications** 7.2 (1962), pp. 147–163. DOI: [10.1137/1107014](https://doi.org/10.1137/1107014).
- [138] Rafael Wagner, Anita Camillini and Ernesto F. Galvão. “Coherence and contextuality in a Mach-Zehnder interferometer”. In: **Quantum** 8 (Feb. 2024), p. 1240. ISSN: 2521-327X. DOI: [10.22331/q-2024-02-05-1240](https://doi.org/10.22331/q-2024-02-05-1240). URL: <http://dx.doi.org/10.22331/q-2024-02-05-1240>.
- [139] Steven R. White. “Density matrix formulation for quantum renormalization groups”. In: **Phys. Rev. Lett.** 69 (19 1992), pp. 2863–2866. DOI: [10.1103/PhysRevLett.69.2863](https://link.aps.org/doi/10.1103/PhysRevLett.69.2863). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.69.2863>.

- [140] Samuel A. Wilkinson and Michael J. Hartmann. “Superconducting quantum many-body circuits for quantum simulation and computing”. In: **Applied Physics Letters** 116.23 (June 2020). ISSN: 1077-3118. DOI: [10.1063/5.0008202](https://doi.org/10.1063/5.0008202). URL: <http://dx.doi.org/10.1063/5.0008202>.
- [141] Glynn Winskel and Mogens Nielsen. “Models for Concurrency”. In: **Handbook of Logic in Computer Science, Vol. IV**. Ed. by Samson Abramsky, Dov Gabbay and Tom Maibaum. Oxford University Press, 1995, pp. 1–148.
- [142] Elie Wolfe, Robert W. Spekkens and Tobias Fritz. “The Inflation Technique for Causal Inference with Latent Variables”. In: **Journal of Causal Inference** 7.2 (2019), p. 20170020. DOI: [doi:10.1515/jci-2017-0020](https://doi.org/10.1515/jci-2017-0020). URL: <https://doi.org/10.1515/jci-2017-0020>.
- [143] Christopher J Wood and Robert W Spekkens. “The lesson of causal discovery algorithms for quantum correlations: causal explanations of Bell-inequality violations require fine-tuning”. In: **New Journal of Physics** 17.3 (Mar. 2015), p. 033002. ISSN: 1367-2630. DOI: [10.1088/1367-2630/17/3/033002](https://doi.org/10.1088/1367-2630/17/3/033002). URL: <http://dx.doi.org/10.1088/1367-2630/17/3/033002>.
- [144] Zhen-Peng Xu and Adán Cabello. “Necessary and sufficient condition for contextuality from incompatibility”. In: **Phys. Rev. A** 99 (2 2019), p. 020103. DOI: [10.1103/PhysRevA.99.020103](https://doi.org/10.1103/PhysRevA.99.020103). URL: <https://link.aps.org/doi/10.1103/PhysRevA.99.020103>.
- [145] Zhen-Peng Xu, Jing-Ling Chen and Otfried Gühne. “Proof of the Peres Conjecture for Contextuality”. In: **Phys. Rev. Lett.** 124 (23 2020), p. 230401. DOI: [10.1103/PhysRevLett.124.230401](https://doi.org/10.1103/PhysRevLett.124.230401). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.124.230401>.
- [146] Tian Zhang, Oscar Dahlsten and Vlatko Vedral. **Quantum correlations in time**. 2020. arXiv: [2002.10448 \[quant-ph\]](https://arxiv.org/abs/2002.10448).

- [147] Yiqing Zhou, E. Miles Stoudenmire and Xavier Waintal. “What Limits the Simulation of Quantum Computers?” In: **Phys. Rev. X** 10 (4 2020), p. 041038. DOI: [10.1103/PhysRevX.10.041038](https://doi.org/10.1103/PhysRevX.10.041038). URL: <https://link.aps.org/doi/10.1103/PhysRevX.10.041038>.