Imperial College London
Department of Physics

Contextuality in Foundations and Quantum Computation

Markus Frembs

Submitted in part fulfilment of the requirements
for the degree of Doctor of Philosophy at
Imperial College London, October 2019

To Jono, Paavo, and Tony.

Acknowledgements

This thesis would not have been written if not for the support of many people.

I am grateful first to my supervisor Terry Rudolph for his support for this project despite running a company in Silicon Valley. His ability to ask provoking questions have spurred a great deal of my passion for science.

My gratitude extends to Stephen Bartlett, who not only offered me financial support for visiting Sydney, Australia on several occasions, but who contributed with his deep physical intuition a large indefinite amount to our ongoing collaboration.

In a very special way I am grateful to Andreas Döring. He played a major role in moulding the central ideas of this project, and I have learnt innumerable things from him, including a refined appreciation for conceptual thought and mathematical beauty. His entirely voluntary commitment to this project in sharing his distinguished insight and providing me his maximum possible guidance have carried me throughout my PhD. I count our friendship among the most valuable outcomes of this project.

I would like to thank EPSRC, the Controlled Quantum Dynamics CDT, and Quantum Group Sydney for providing funding throughout my PhD. I am greatly indebted to the administrative staff, particularly Miranda Smith, Veena Dhulipala, and Naho Ollason.

I am also greatly indebted to my colleagues at Imperial College for many stimulating discussions, and even more for our companionship, which I deeply cherish. I owe a warm welcome as an 'orphan' to the Quantum Group in Sydney, and am especially grateful to Sam Roberts, our (Skype) discussions are always insightful and most enjoyable.

A great part of this journey I have walked by the immeasurable emotional support of many close friends. You likely underestimate your significance in this work, and more so my personal growth over the years. I cannot express how much I value your friendship, care, and love Daniel, Maxi, Karina, Cecile, Lauren, Chris, Charles, and Alex, as well as Veri, Anna, and Sally.

I would also like to mention, that when not with friends or colleagues, I have endured difficult times with the support of the motivational, uplifting, and always life-affirming tunes by Above & Beyond: Treasure really is measured in units of love!

Needless to say, I would never even have gotten to a place to undergo such a project without my family. I am particularly grateful to my parents, who have always encouraged me to pursue my own goals and have shown me their support in every way possible—latest by allowing me the genial atmosphere of being home with family during the write-up of this thesis. Very special thanks goes to my brothers who grounded me whenever I was going in circles during that time.

Abstract

Contextuality is a key concept in quantum theory. We reveal just how important it is by demonstrating that quantum theory builds on contextuality in a fundamental way: a number of key theorems in quantum foundations can be given a unified presentation in the topos approach to quantum theory, which is based on contextuality as the common underlying principle. We review existing results and complement them by providing contextual reformulations for Stinespring's and Bell's theorem.

Both have a number of consequences that go far beyond the evident confirmation of the unifying character of contextuality in quantum theory. Complete positivity of quantum channels is already encoded in contexts, nonlocality arises from a notion of composition of contexts, and quantum states can be singled out—among more general non-signalling correlations over the composite context structure—by a notion of time orientation in subsystems, thus solving a much discussed open problem in quantum information theory. We also discuss nonlocal correlations under the generalisation to orthomodular lattices and provide generalised Bell inequalities in this setting.

The dominant role of contextuality in quantum foundations further supports a recent hypothesis in quantum computation, which identifies contextuality as the resource for the supposed quantum advantage over classical computers. In particular, within the architecture of measurement-based quantum computation, the resource character of nonlocality and contextuality exhibits rather clearly.

We study contextuality in this framework and generalise the strong link between contextuality and computation observed in the qubit case to qudit systems. More precisely, we provide new proofs of contextuality as well as a universal implementation of computation in this setting, while emphasising the crucial role played by phase relations between measurement eigenstates. Finally, we suggest a fine-grained measure for contextuality in the form of the number of qubits required for implementation in the non-adaptive, deterministic case.

Contents

List of Figures

Declaration

I declare that the content of this thesis is entirely my own original work, except where appropriately referenced.

Copyright Disclaimer

Much of the content of this thesis is adapted and in rare cases directly taken from the following papers:

- A. Döring, M. Frembs, arXiv: `1910.09591 [math-ph]` (2019) [47]
- M. Frembs, A. Döring, arXiv: `1910.09596 [quant-ph]` (2019) [62]
- M. Frembs, S. Roberts, S. D. Bartlett, *New J. Phys.* **20**, 103011 (2018) [64]

Ch. 2 contains the essence of the first two preprints, in particular, Sec. 2.3 and Sec. 2.4 are largely based on [47], while the proof of Thm. 43 in Sec. 2.4.4 is a generalisation of the proof presented in [62]. The latter further contains a discussion on frame functions in composite systems, which closely resembles the analysis in Sec. 2.4.2. Finally, Sec. 3.1.4 in Ch. 3 reproduces in more detail the main result in [64].

Information derived from these works as well as published and unpublished work of others has been acknowledged in the text and references are given in the list of sources.

"I should have more faith," he said; "I ought to know by this time that when a fact appears opposed to a long train of deductions it invariably proves to be capable of bearing some other interpretation."

- Sir Arthur Conan Doyle

Chapter 1

Introduction

Recent success of machine learning algorithms, which mimic the brain in the form of neuronal networks, suggests that our cognitive abilities arise at least to some degree in a similar manner. Maybe the greatest challenge to science is to decipher whether our conscious experience is 'simply' that—a complex network of physical neurons trained over years of sensual input—or whether our perception of self rests on (fundamentally) different physics altogether [72]. A conclusive answer to this and the possibly related 'hard problem' of consciousness [31] seems far afield, yet certain aspects of the biological information processing in our brains undoubtedly affect our thinking. Clearly, this extends to our ability 'to do science'. One particularly limiting factor resulting from this is that our perception bias makes original thought, which defies the pattern recognition of physical processes relevant to our everyday experience, a true rarity. Few instances of people overcoming the boundary of this 'natural intuition' have triggered profound philosophical, scientific, and technological revolutions, and are among the greatest achievements of mankind.

Surely, the invention of quantum theory deserves its place in this list. What is more, it is also a special representative in that even a hundred years after its formulation, the scope of its revolutionary content has arguably not been fully grasped yet. It might even require a further substantial shift in perspective as debates on quantum foundations are still ongoing and with them the discussion about how to change our image of reality.

One feature of quantum theory with far-reaching consequences for our understanding of the world is beautifully captured in the following adaptation of a parable by the Swiss mathematician Ernst Specker (English translation largely taken from [107]).

At the Assyrian School of Prophets in Arba'ilu in the time of King Asarhaddon, there taught a seer from Nineva. He was a distinguished representative of his faculty (eclipses of the sun and moon) and aside from the heavenly bodies, his interest was almost exclusively in his daughter. His teaching success was limited; the subject proved to be dry and required a previous knowledge of mathematics which was scarcely available. If he did not find the student interest which he desired in class, he did find it elsewhere in overwhelming measure. His daughter had hardly reached a marriageable age when he was flooded with requests for her hand from students and young graduates. And though he did not believe that he would always have her by his side, she was in any case still too young and her suitors in no way worthy. In order that the suitors might convince themselves of their unworthiness, he promised them that she would be wed to the one who could solve a prediction task that was posed to them.

Each suitor was taken before a table on which stood four little boxes arranged in a square, [each of which might or might not contain a gem], and was asked to predict which of the boxes contained a gem and which did not. But no matter how many times they tried, it seemed impossible to succeed in this task. After each suitor had made his prediction, he was ordered by the father to open any two boxes within either a row or a column of the square arrangement. It always turned out that two opened boxes disproved the suitor's prediction. The daughter would have remained unmarried until the father's death, if not for the fact that, after the prediction of the son of a prophet [whom she fancied], she quickly opened two boxes herself, and the suitor's prediction [for these two boxes] was found, in this case, to be correct. Following the weak protest of her father that he had wanted two other boxes opened, she tried to open the remaining two. But this proved impossible where upon the father grudgingly admitted that the prediction, being unfalsified, was valid. [The daughter and the suitor were married and lived happily ever after.]

Assuming the seer was not merely a skilful trickster, Specker's parable puts into question a certain logical assumption that pervades physics well into the twentieth century. In order to motivate exactly what this assumption is, note that we constantly interpret data by means of inferring information about the world that we do not have direct access to.

A helpful illustration for this fact is the work of a famous detective trying to reconstruct a crime scene in Sir Arthur Conan Doyle's 'Silver Blaze' [51]:

"Is there any point to which you would wish to draw my attention?"

"To the curious incident of the dog in the night-time."

"The dog did nothing in the night-time."

"That was the curious incident."

From the rather obvious inference 'Silver Blaze was stolen from the stable in the night-time, therefore there must have been a perpetrator' and the masterly inference 'the dog did not bark in the night-time, therefore no-one unknown to the household approached the stable' the observant detective immediately arrives at the cunning conclusion 'Silver Blaze was stolen by a perpetrator, who was known to the household' as a logical necessity.

More abstractly, and for the propositional logic underlying a physical system, the same line of reasoning has that with implications $A \to B$ and $B \to C$, the implication $A \to C$ should also hold. Since ordering statements by deduction (abduction in the novel) is so deeply ingrained into our everyday lives, it is easy to overlook that this reasoning relies on the assumption that statements can always be ordered in this way. This might seem obvious, however, this assumption is not necessary. Specker points this out in the above parable, where implications exist between neighbouring boxes only, yet not for all boxes together; for this reason, the daughter cannot open the remaining boxes after the first two are opened. In other words, the events relating to the information content of the boxes containing a gem or not are not all simultaneously measurable. The inconspicuous assumption underlying classical deductive reasoning therefore is that all statements about a system are simultaneously verifiable, and implications between them can thus be related transitively as in the above example. Yet, the situation in Specker's parable is different—only certain subsets of statements about the system can be simultaneously verified, and implications between them thus be composed transitively.

But clearly, we can open four boxes and observe their content simultaneously, so why bother with this mathematical curiosity? In fact, it turns out that quantum theory behaves very similar to the boxes in Specker's parable—both the original version with three boxes and our adaptation. Similar only, since the exact scenarios do not exist, yet very similar ones do; for the version above, increasing the number of boxes arranged in a square from four to nine constitutes a similar example, which turns out to be quantum-realisable in the Mermin-Peres square (cf. Fig. 3.1 (a)). This raises the question why nature does not behave by our intuition, yet at the same time does not depart arbitrarily far from it. This question is an interesting one and has been addressed e.g. in [107], where the original version of Specker's parable is discussed.

On the other hand, since contextuality is likely not the only physical principle underlying nature, another question is at least as pressing: how important is the idea behind Specker's parable on the structure of logical implications—subject to the equivalence relation defined by simultaneous measurability—in nature, specifically, in quantum theory? Is it merely a curious feature of the latter or does it underlie physics on a fundamental level? This latter question has been brought to the forefront by Chris Isham and collaborators [23, 24, 75, 87], who suggested to construct theories of physics from the collection of classical perspectives or 'contexts'.

In Ch. 2 we review this idea in detail (cf. Sec. 2.1.2), collect several known results about the structure of simultaneous measurability in quantum theory (cf. Sec. 2.3), and extend them to nonlocality in composite systems (cf. Sec. 2.4). In particular, we provide inherently contextual reformulations for Stinespring's and Bell's theorem, based on a notion of contextual composition in place of the tensor product construction in the standard formalism. We also give a definition of general non-signalling theories over orthomodular lattices and derive Bell inequalities in this setting (cf. Sec. 2.5).

In accordance with previous results, our findings emphasise the crucial role played by contextuality. The emerging, alternative formulation of quantum theory heavily rests on the deep insights by Specker and Isham, and thus suggests a potential shift in perspective on quantum physics as a whole.

This has far reaching consequences not only for quantum foundations, but also its applications such as quantum computation. Contextuality has recently been suggested as the resource responsible for the supposed advantage of quantum over classical computers. Clearly, the fact that contextuality underlies quantum theory on a fundamental level, strongly supports this hypothesis. Yet, more than that is needed in order to convert this resource into useful computational power. We address this issue by studying contextuality in the particular computing architecture known as measurement-based (quantum) computation in Ch. 3. More precisely, we improve existing results on the resource character of contextuality (cf. Sec. 3.1), construct many inherently contextual examples (cf. Sec. 3.2), and suggest a classification of contextuality by means of the number of local subsystems required for the implementation of certain tasks in this framework (cf. Sec. 3.3).

We end by discussing potential avenues for future research along various directions in Ch. 4.

Chapter 2

Contextuality in Foundations of Quantum Theory

Quantum theory was developed in the beginning of the twentieth century by Planck, Einstein, Bohr, Sommerfeld, de Broglie, and many others in an attempt to understand the emission spectra of atoms. The theory was given a rigorous mathematical formulation only a few decades later by Heisenberg, Schrödinger, Born, Dirac, Wigner, von Neumann, Jordan, Weyl, and many more and has remained largely unchanged since then. We review important aspects of this algebraic structure in Sec. 2.2 and discuss key results in foundations in this algebraic setting. In particular, we will be concerned with the landmark theorems by Wigner, Gleason, Stinespring, Bell, and Kochen & Specker. Most obviously, contextuality is the subject of the Kochen-Specker theorem. Therefore, in Sec. 2.1.1 we give a detailed exposition of this theorem first, and discuss its topos-based reformulation by Isham, Butterfield, and Hamilton [23, 24, 75, 87] as a prototype for the reformulation of other theorems in foundations. For more background on the topos approach to quantum theory and other closely related ideas we refer to [1–4, 33, 53, 56, 60, 61, 78, 79, 115, 116]. A great benefit of this reformulation is that it gives a geometrical interpretation of the former, based on a generalised state space in the form of the spectral presheaf, which fundamentally builds on the concept of physical contextuality, which we define in Sec. 2.1.2. In fact, physical contextuality and its mathematical embodiment in the form of presheaves over the partial order of contexts prove more universal and lead to reformulations of other key theorems, namely Gleason's and Wigner's theorem. We review those in Sec. 2.3 and highlight their intimate relationship with contextuality. Measured by the significance of these theorems for quantum foundations, it is surprising that their contextual nature has not been elaborated on before, and should be seen as an important contribution to foundations in and of itself.

The connection between Bell's theorem, locality, and contextuality has been recognised before [6, 25, 57, 59, 92]. However, no contextual reformulation of the above type had previously been known. In order to bridge this gap, first, we give a derivation of the crucial assumption of factorisability in Bell's theorem from the perspective of classical state spaces and argue that it is naturally related to trivial physical contextuality in Sec. 2.4. In particular, we show how contextuality fundamentally relates to composition of systems. To this end, we provide a notion of composition based only on the context structure of a theory. Our key result is a reformulation of Bell's theorem in contextual form based on our notion of context composition and a choice of time direction in subsystems. This reformulation can also be understood as a generalisation of Gleason's theorem to composite systems, strengthening a previous result in [145].

As a consequence of Bell's theorem in contextual form we obtain a solution to a key problem in quantum foundations and quantum information theory concerning the restrictiveness of the no-signalling principle [15, 121, 125]. We show in detail that no-signalling constraints also arise as marginalisation constraints between contexts in the Bell presheaf. As such we find that no-signalling corresponds to our definition of composition of contexts and almost singles out quantum theory over the context structure corresponding to local von Neumann algebras: it only lacks a choice of time direction in subsystems.

Along the way we provide a reformulation of the more technical but nonetheless crucial dilation theorem by Stinespring (cf. Sec. 2.3.4). In particular, we prove that completely positive maps are naturally encoded on the level of contexts already. Both, Stinespring's and Bell's theorem thus also prove to be very closely connected with contextuality in ways not recognised previously. Finally, in Sec. 2.5 we discuss generalisations of our reformulation of Bell's theorem in contextual form by considering orthomodular lattices instead of projection lattices of von Neumann algebras, and we consider correlations in such theories. Sec. 2.6 summarises.

2.1 The Kochen-Specker theorem and contextuality

This section serves two purposes, it is meant as an introduction to the Kochen-Specker theorem and as a conceptual motivation for the notion of physical contextuality. First, in Sec. 2.1.1 we review the key idea behind the Kochen-Specker theorem and highlight the role contextuality plays in it. Second, in Sec. 2.1.2 we extract the physical principle inherent to contextuality and argue how it reveals a principal difference between classical and quantum physics. We fill it with mathematical content in Sec. 2.2, which will allow us to relate not only the Kochen-Specker theorem to contextuality, but many more key components of quantum theory, too.

2.1.1 The Kochen-Specker argument

Following Bell's seminal work [18], Kochen and Specker further refined the constraint on hidden variable models [101]. At least since the famous paper by Einstein, Podolski, and Rosen [54] there had been an ongoing debate about the possibility to understand quantum theory as a high-level description of a more fundamental theory, similar to thermodynamics, which Boltzmann gave a statistical underpinning based on classical physics in the form of Newtonian mechanics. Bell's theorem (cf. Sec. 2.4.4) puts strong constraints on such an interpretation. However, while Bell derives his conclusion in conjunction with a notion of locality, Kochen and Specker remove this additional assumption and simply ask whether it is at all possible to ascribe a classical state space to quantum theory. In particular, Kochen and Specker are concerned with the concept of a 'microstate', which can be understood as a deterministic assignment of measurement outcomes to all observables simultaneously. Importantly, microstates exist in classical systems as a consequence of the fact that classical observables are all simultaneously measurable. Famously, by the uncertainty principle observables are not simultaneously measurable in quantum theory, however, this does not necessarily imply that no microstates of the aforementioned type exist. Nevertheless, Kochen and Specker show that quantum theory does not admit microstates. We present the key steps of their argument before changing perspective and pinpointing the underlying notion of physical contextuality in Sec. 2.1.2.

Kochen and Specker ask whether quantum theory allows for an underlying 'classical' description. To this end, they first set out to clarify what conditions should be met by a classical interpretation of a theory. Note that two important ingredients to a physical theory are (i) a set of observables $\mathcal{O}$ and (ii) a set of states Σ. In a classical theory, observables are modeled as functions on some measurable space Σ: for every observable $a \in \mathcal{O}$, there exists a measurable function $f_a : \Sigma \to \mathbb{R}$.[1] If moreover every such function is promoted to the status of an observable, $\mathcal{O}$ has the structure of an algebra.

Forcing this model onto quantum theory, the probability of measuring a value in the interval $\Delta \subset \mathbb{R}$ with the observable $a \in \mathcal{O}$ given a quantum state ψ is given by

$$P_{a,\psi}(\Delta) = \mu_\psi(f_a^{-1}(\Delta)) \ .$$

Here, μ_ψ is the measure on the state space Σ corresponding to ψ. Clearly, any interpretation of quantum theory should also reproduce quantum mechanical expectation values,

$$\mathbb{E}_\psi(a) = \int_\Sigma d\mu_\psi(s) f_a(s) = \langle \psi, a\psi \rangle \ . \tag{2.1}$$

Yet, Eq. (2.1) by itself is not very restrictive, it allows for artificial state spaces such as the one constructed in [101]. Let $\Sigma = \mathbb{R}^\mathcal{O} = \{s : \mathcal{O} \to \mathbb{R}\}$ and set $f_a(s) = s(a)$ for all $a \in \mathcal{O}$. Then the product measure $\mu_\psi = \prod_{a\in\mathcal{O}} P_{a,\psi}$ trivially reproduces quantum mechanical probabilities: $\mu_\psi(f_a^{-1}(\Delta)) = \mu_\psi(\{s \mid s(a) \in \Delta\}) = P_{a,\psi}(\Delta)$.

This example works since no functional relations between observables are taken into account. However, whenever such relations exist on the level of the observables, it is natural to require these to be at least partially reflected in their (functional) representation. Taken to the extreme, one might thus require the full algebraic structure of quantum theory to be reflected in such a representation:

$$f_{a+b} = f_a + f_b \qquad\qquad f_{ab} = f_a \cdot f_b \qquad\qquad f_{\lambda a} = \lambda f_a \tag{2.2}$$

[1]Here, 'measurable' is used in the mathematical sense.

To find such a representation means to give an algebra homomorphism $\mathcal{O} \to \mathbb{R}^\Sigma$ from the algebra of quantum mechanical observables $\mathcal{O}$ to the algebra of real-valued functions on some measure space Σ. This is easily seen to be impossible by an example due to Bell. Consider the spin observables σ_x and σ_z with eigenvalues ± 1, as well as the spin observable $\sigma_{xz} := \frac{1}{\sqrt{2}}(\sigma_x + \sigma_z)$ corresponding to measuring the spin along the axis bisecting the measurement axes of σ_x and σ_z. Requiring this algebraic relation to be reflected in a functional representation $f_\sigma : \Sigma \to \mathbb{R}$ immediately yields a contradiction when evaluated on a microstate $s \in \Sigma$: $f_{\sigma_{xz}}(s) = \frac{1}{\sqrt{2}}(f_{\sigma_x}(s) + f_{\sigma_z}(s)) = \frac{1}{\sqrt{2}}(\pm 1 + \pm 1) \neq \pm 1 = f_{\sigma_{xz}}(s)$.

It is here that Bell, Kochen, and Specker have a deep structural insight: there is no reason to require algebraic constraints to be reflected between *all* quantum mechanical observables, only when the measured values can be inferred from one another such relations should hold. Since measured values can be inferred between simultaneously measurable observables, it is natural to *at least* require algebraic constraints to be reflected in the spectra of such observables. We call the set of measured values of an observable $a \in \mathcal{O}$ its *spectrum* and denote it by $\mathrm{sp}(a)$.

In order to evaluate the relevant constraints in Eq. (2.2), Kochen and Specker therefore introduce the notion of a *partial algebra*, i.e., an algebra with an equivalence relation called *'simultaneous measurability'* defined on it. Accordingly, a *partial algebra homomorphism* relaxes the conditions of an algebra homomorphism in Eq. (2.2) to hold between simultaneously measurable observables only [101]. Of special interest are partial algebra homomorphisms into $\mathbb{R}$.

Definition 1. *Let $\mathcal{O}$ be a partial algebra representing the observables of a physical theory. A* valuation function *('prediction function' in [101]) $v : \mathcal{O} \to \mathbb{R}$ is a map such that:*

(i) $v(a) \in \mathrm{sp}(a)$ (spectrum rule)

(ii) $v : \mathcal{O} \to \mathbb{R}^\Sigma$ is a partial algebra homomorphism

Clearly, if there exists a state space Σ underlying a theory, then every state $s \in \Sigma$ defines a valuation function $v_s : \mathcal{O} \to \mathbb{R}$ by evaluation, $v_s(a) = f_a(s)$. By ruling out the existence of valuation functions in quantum theory, Kochen and Specker conclude that no classical interpretation can be given for the latter, in particular, that no classical state space Σ exists.

In order to state the theorem, one only needs a minimum of projective geometry in three dimensions—strictly less than the full mathematical apparatus of quantum theory. However, the technical proof somewhat distracts from the important physical aspects and we therefore defer the discussion of the theorem until after a thorough treatment of the mathematical background in Sec. 2.2, which will also allow us to introduce substantial generalisations of the original result in [101]. Instead, in the next section we focus on the conceptual idea underlying the argument outlined in this section—the principle of physical contextuality.

2.1.2 Physical contextuality

The key ingredient to the argument in the last section is the restriction of the algebraic constraints between observables in Eq. (2.2). Building on [75, 101], in [47] we conceptualise this as follows.

Definition 2. *Let the observables of a physical system be given by a partial algebra $\mathcal{O}$ with equivalence relation called* simultaneous measurability*. An equivalence class of observables that are pairwise simultaneously measurable is called a* (maximal) context*. Moreover, $\mathcal{O}$ is called* physically contextual *if not all its observables are simultaneously measurable, i.e., if $\mathcal{O}$ is not itself a (maximal) context.*

Classical theories are not physically contextual by this definition, they contain a single maximal context and we will sometimes call them *single-context* theories for this reason (cf. [53]). Quantum theories, on the other hand, are a very special type of physically contextual theories. Importantly, any theory with physical contextuality still has contexts, i.e., equivalence classes of observables that are pairwise simultaneously measurable. In the extreme case, contexts consist of a single observable. What is more, there is a natural notion of *coarse-graining* arising from inclusion relations between subsets of equivalence classes called (non-maximal) *contexts*. From an information-theoretic perspective, coarse-graining captures the loss of information when going to smaller contexts. Contexts and their order relations therefore encode physical contextuality [1–4, 23, 24, 47, 75, 87, 101].

Definition 3. *Let the observables of a physical system be given by a partial algebra $\mathcal{O}$. The* context category $\mathcal{C}(\mathcal{O})$ *is the partial order of (non-maximal) contexts ordered by inclusion.*

Soon, we will assume contexts to carry additional structure, e.g., in quantum theory it is natural to take contexts to correspond to unital, commutative C^*-algebras or von Neumann algebras. Yet, we emphasise that the importance of physical contextuality lies in the order structure, which applies already at the level of mere sets. A related approach to contextuality setting off at this level of generality is the sheaf-theoretic framework in [6].

We remark that contextuality is used with different meanings in the literature [26, 81, 85, 101, 126, 134]. In order to clearly distinguish those from the one given in Def. 2 and Def. 3, we call the latter 'physical contextuality'. Notably, physical contextuality is inherently operational, since it is based on the notion of simultaneous measurability, yet there is no need to introduce 'contextual value assignments', 'counterfactual definiteness' or other often convoluted concepts.[2] Instead, we use our minimal version of (physical) contextuality mostly as a mathematical (bookkeeping) tool in order to study its restrictiveness in quantum theory.

Physical contextuality is a conceptual principle and (mostly) independent of a mathematical formalism. In order to study its role in quantum theory quantitatively, in the next section we give the necessary mathematical background on (algebraic) quantum theory and make the structure of physical contextuality explicit in this case. We will then see that physical contextuality is not only at the heart of the Kochen-Specker theorem, but also of other key theorems in quantum foundations, which obtain a natural reformulation in terms of the order structure between contexts (cf. Sec. 2.3).

[2] Instead, in order to make sense of probabilistic assignments and correlations in Bell's theorem later on, we will tacitly assume a notion of 'statistical regularity' [103], which in some form necessarily underlies any kind of scientific study. Arguably, this is a much less problematic principle than 'counterfactual definiteness'.

2.2 Mathematical background

Throughout this chapter we will take the view of algebraic quantum theory. This framework underlies not only quantum mechanics, but also quantum information theory and in large part (algebraic) quantum field theory. The key ingredient is the algebraic structure on observables, which are modelled mathematically by self-adjoint operators on a Hilbert space. The self-adjoint operators play a twofold role, first, they underlie the measurement process, second, they correspond to infinitesimal generators of time evolution. The latter aspect connects with the theory of Lie algebras while the former is captured by the probability calculus inherent to Jordan algebras. Both aspects are intricately interwoven into one multiplicative product yielding the structure of an associative von Neumann or C^*-algebra. We review this structure in detail over the next sections. In particular, we make explicit the idea that quantum theory arises via 'local-to-global' extensions from classical physics, where 'local' means 'within a single context' or 'classical' and 'global' refers to the collection of all contexts by means of physical contextuality. As will become clear soon, the latter corresponds to non-trivial Jordan structure and only indirectly to noncommutativity, which is often considered to be the essence of quantum.

We start with a general overview of algebraic quantum theory in Sec. 2.2.1 (for references, see [96, 97, 138, 139, 144]). In Sec. 2.2.2 we give a rigorous definition of physical contextuality within algebraic quantum theory. More precisely, we introduce some basic notions from category theory (for more details, we refer to e.g. [104, 106]) in order to define the context category and presheaves over the context category as the tools to study *local-to-global* problems, which provide the basis for the reformulations of many key theorems in contextual form in Sec. 2.3.

In Sec. 2.2.3 we connect physical contextuality with the Jordan algebra aspect in von Neumann algebras. The corresponding split of the associative product into a symmetric and an antisymmetric part lies at the heart of the dichotomy of observables in quantum theory, whose role will become particularly important in the study of Bell's theorem in Sec. 2.4.4. Since the relevant structures—Jordan algebras, order derivations, orientations etc.—are not well known outside of small communities, we provide some necessary background on (only) those notions needed for later theorems. Standard references with many more details include [8, 39].

2.2.1 Part I - Algebra of observables

Bounded operators on Hilbert space. A *Hilbert space* $\mathcal{H}$ is a vector space with inner product $(\cdot,\cdot)$ completed in the norm $||\cdot|| = |(\cdot,\cdot)|^2$. Recall that a sesquilinear form $(\cdot,\cdot) : V \times V \to \mathbb{C}$ on a vector space V is called *positive-definite* if $(x,x) \geq 0$ for all $x \in V$ and $(x,x) \geq 0 \Leftrightarrow x = 0$, and is called *conjugate-symmetric* or *Hermitian* if $(v,w) = \overline{(w,v)}$ for all $v,w \in V$. An *inner product* is a positive-definite Hermitian form.[3]

Every Hilbert space is in particular a *Banach space*, i.e., a normed vector space complete with respect to the induced norm topology. In finite dimension $n \in \mathbb{N}$, the prototypical Hilbert space is $\mathcal{H} = \mathbb{C}^n$ with *inner product* $(v,w) = \sum_{i=1}^n v_i \overline{w_i}$ for all $v,w \in \mathbb{C}^n$. A Hilbert space is called *separable* if it has a countable *orthonormal basis* $(v_i)_{i\in\mathbb{N}}$, i.e., a countable family of vectors $v_i \in \mathcal{H}$ such that $(v_i,v_j) = \delta_{ij}$, and for every $v \in \mathcal{H}$ there exist unique complex numbers $(c_i)_{i\in\mathbb{N}}$ such that $v = \sum_{i=1}^\infty c_i v_i$. In what follows, Hilbert spaces will be assumed separable.

A *linear operator* $a : V \to W$ is a linear map between vector spaces V, W. A linear operator between normed vector spaces V, W is called *bounded* if there exists $0 \leq K \in \mathbb{R}$ such that $||av||_W < K||v||_V$ for all $v \in V$. The smallest such K is the *operator norm of* a, denoted $||a|| := \inf\{K \in \mathbb{R} \mid ||av|| \leq K||v|| \ \forall v \in V\}$,[4] and a is *bounded* if $||a|| < \infty$.

Theorem. *A linear operator between normed spaces is bounded if and only if it is continuous.*

The set of bounded operators on Hilbert space $\mathcal{H}$ is denoted $\mathcal{B}(\mathcal{H})$ and forms a *Banach algebra*, i.e., it is a vector space, which is complete in the topology induced by the operator norm, and such that multiplication is continuous: $||ab|| \leq ||a|| \cdot ||b||$ for all $a,b \in \mathcal{B}(\mathcal{H})$.

For the purposes of this thesis, it will be enough to consider bounded operators, more precisely, observables will be mathematically represented by bounded self-adjoint operators on some Hilbert space.[5] The latter property refers to a further symmetry of the Banach algebra $\mathcal{B}(\mathcal{H})$: for $a \in \mathcal{B}(\mathcal{H})$, define a^* to be the operator such that $(av,w) = (v,a^*w)$ for all $v,w \in \mathcal{H}$. a^* is called the *(Hilbert) adjoint of* a and is provably unique. In finite dimensions the adjoint is

[3] We will follow the mathematical convention that sesquilinear forms are linear in the first (and conjugate linear in the second) argument: $\forall \lambda_1, \lambda_2 \in \mathbb{C}$, $\forall v_1, v_2, w \in \mathcal{H}$, $(\lambda_1 v + \lambda_2 v_2, w) = \lambda_1(v_1,w) + \lambda_2(v_2,w)$.

[4] Note that the operator norm depends on the respective norms on V and W.

[5] We remark that unbounded operators such as position and momentum in quantum mechanics can be treated by affiliating them with the von Neumann algebra of observables (cf. [96, 97]).

given by transposition and complex conjugation, $a^* = \overline{a^t}$. The following properties of operators $a \in \mathcal{B}(\mathcal{H})$ are of special interest. An operator is called *normal* if $aa^* = a^*a$, a normal operator is called *unitary* if $aa^* = a^*a = 1$ and *self-adjoint* if $a = a^*$. A self-adjoint operator is called *positive* if $(av, v) \geq 0$ for all $v \in \mathcal{H}$. We denote the set of positive operators by $\mathcal{B}(\mathcal{H})_+$. The set of self-adjoint operators forms a real vector space denoted by $\mathcal{B}(\mathcal{H})_{\mathrm{sa}} \subset \mathcal{B}(\mathcal{H})$, but not a subalgebra since $ab \notin \mathcal{B}(\mathcal{H})_{\mathrm{sa}}$ in general. Every operator $a \in \mathcal{B}(\mathcal{H})$ has a unique decomposition, $a = a_1 + ia_2$, where $a_1, a_2 \in \mathcal{B}(\mathcal{H})_{\mathrm{sa}}$, namely $a_1 = \frac{1}{2}(a + a^*)$, $a_2 = \frac{-i}{2}(a - a^*)$. It follows that a is normal if and only if $[a_1, a_2] = 0$.

C^*-algebras, von Neumann algebras. The map $^* : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ makes $\mathcal{B}(\mathcal{H})$ into a $*$-algebra: it defines an *antilinear involution*, i.e., for all $a, b \in \mathcal{B}(\mathcal{H})$, $\lambda \in \mathbb{C}$ it holds $(\lambda a)^* = \overline{\lambda} a^*$, $(a+b)^* = a^* + b^*$, $(ab)^* = b^*a^*$, and $(a^*)^* = a$. Since $\mathcal{B}(\mathcal{H})$ is also a Banach space, it is natural to require compatibility with the norm as well. In fact, the *C^*-identity*, $||a^*a|| = ||a^*|| \cdot ||a|| = ||a||^2$ for all $a \in \mathcal{B}(\mathcal{H})$, holds and makes $\mathcal{B}(\mathcal{H})$ into a *C^*-algebra*.

Definition 4. *A C^*-algebra $\mathcal{A}$ is a Banach algebra over the field of complex numbers, together with an antilinear involution $^* : \mathcal{A} \to \mathcal{A}$, which satisfies the C^*-identity,*

$$\forall a \in \mathcal{A}: \ ||a||^2 = ||a^*a|| \ .$$

The C^*-algebra $\mathcal{B}(\mathcal{H})$ thus carries a natural topology induced by the operator norm $|| \cdot ||$ known as *uniform operator topology*. However, this topology is in general not the only topology on $\mathcal{B}(\mathcal{H})$, several weaker topologies exist. For instance, a net $(a_i)_{i \in I}$ converges to a in the *strong operator topology* or *strongly* if and only if $||a_i v - av|| \to 0$ for all $v \in \mathcal{H}$.[6] This is the topology of pointwise convergence and generally weaker than the uniform operator topology. Of special interest to us is the *weak operator topology*. A net $(a_i)_{i \in I}$ converges to a in the weak operator topology or *weakly* if and only if $(v, a_i w) \to (v, aw)$ for all $v, w \in \mathcal{H}$. Requiring closure with respect to the weak operator topology leads to the definition of a *von Neumann algebra*.

[6] A *net* $f : I \to X$ is a map from a directed set I to a topological space X. A *directed set* I is a non-empty set with a preorder, i.e., a reflexive, transitive relation $\leq$ such that every pair of elements has an upper bound (cf. Sec. 2.2.2).

Definition 5. *A von Neumann algebra $\mathcal{N}$ is a unital subalgebra of $\mathcal{B}(\mathcal{H})$, i.e., a subalgebra including the multiplicative identity in $\mathcal{B}(\mathcal{H})$, closed in the weak operator topology.*

Every von Neumann algebra is in particular uniformly closed and thus a unital C^*-algebra. Note also that $\mathcal{B}(\mathcal{H})$ is a von Neumann algebra. Von Neumann algebras provide a natural mathematical representation of physical quantities.

The algebra of physical quantities is modeled by
a (noncommutative) von Neumann algebra.

We remark that weakly closed operator algebras can also be characterized abstractly, without reference to a Hilbert space, as follows: let X be a Banach space and denote by X^* its continous dual, i.e., the Banach space of bounded linear functionals $\phi : X \to \mathbb{C}$ (see below).

Definition 6. *A W^*-algebra $\mathcal{N}$ is a C^*-algebra that is the dual of some Banach space X. The latter is called the* predual of $\mathcal{N}$.

Nevertheless, by the Gelfand-Naimark representation theorem, Thm. 5 below, every von Neumann algebra $\mathcal{N}$ arises as a subalgebra $\mathcal{N} \subseteq \mathcal{B}(\mathcal{H})$ for some Hilbert space $\mathcal{H}$.

Linear functionals, states, and weights. A linear functional ϕ on a C^*-algebra $\mathcal{A}$ is a linear map $\phi : \mathcal{A} \to \mathbb{C}$. A linear functional ϕ is *positive* if $\phi(a) \geq 0$ for all $a \in \mathcal{A}_+$.

Theorem 1. *A linear functional ϕ is positive if and only if it is bounded and $||\phi|| = \phi(1)$.*

A *state* σ is a positive (bounded) linear functional that is normalised, i.e., $\sigma(1) = 1$. The set of states on $\mathcal{A}$ is denoted $\mathcal{S}(\mathcal{A})$ and forms a convex set since for every two states $\sigma_1, \sigma_2 \in \mathcal{S}(\mathcal{A})$ and $\lambda \in [0,1]$: $\lambda\sigma_1 + (1-\lambda)\sigma_2 \in \mathcal{S}(\mathcal{A})$. By the Krein-Milman theorem $\mathcal{S}(\mathcal{A})$ is the weakly closed convex hull of its extreme points, which are called *pure states* and denoted $\mathcal{E}(\mathcal{A})$. A non-zero linear functional on a commutative C^*-algebra $\mathcal{A}$ is a pure state if and only if it is multiplicative. The set of multiplicative linear functionals is called the *Gelfand spectrum of* $\mathcal{A}$,

$$\Sigma_{\mathcal{A}} := \{0 \neq \lambda : \mathcal{A} \to \mathbb{C} \text{ linear} \mid \forall a, b \in \mathcal{A} : \ \lambda(ab) = \lambda(a)\lambda(b)\} = \mathcal{E}(\mathcal{A}) \ . \tag{2.3}$$

The Gelfand spectrum is a compact Hausdorff space relative to the weak* topology.

Finally, weights generalise linear functionals in (infinite-dimensional) von Neumann algebras.

Definition 7. *Let $\mathcal{N}$ be a von Neumann algebra and $\mathcal{N}_+$ the set of positive elements of $\mathcal{N}$. An additive, homogeneous map $\omega : \mathcal{N}_+ \to [0, \infty]$, i.e., $\omega(a + b) = \omega(a) + \omega(b)$ and $\omega(\lambda a) = \lambda\omega(a)$ for all $a, b \in \mathcal{N}_+$, $\lambda \in \mathbb{R}$, is called a* weight on $\mathcal{N}$.

A weight is called *faithful* if $\omega(a) = 0$ implies $a = 0$. A weight is called *finite* if $\omega(1) < \infty$. A *positive linear functional* on $\mathcal{N}$ is a finite weight. If moreover $\omega(a^*a) = \omega(aa^*)$ for all $a \in \mathcal{N}$, then ω is called a *trace on $\mathcal{N}$*. Finally, a trace is called *semi-finite* if for all $a \in \mathcal{N}_+$ non-zero, there exists $b \in \mathcal{N}_+$ non-zero with $\omega(b) \leq \infty$ and $b \leq a$.

Trace-class operators and normal states. There are some important subspaces of $\mathcal{B}(\mathcal{H})$ arising as closed two-sided ideals. First, an operator $a : E \to F$ between Banach spaces is called *compact* if for all $D \subseteq E$ bounded, the closure $\overline{a(D)}$ is compact in F.

Theorem 2. *Let $\mathcal{H}_1, \mathcal{H}_2$ be Hilbert spaces. Every compact operator $a : \mathcal{H}_1 \to \mathcal{H}_2$ is the norm limit of finite linear sums $\sum_{i=1}^{n} c_i(\cdot, v_i)w_i$ with $c_i \in \mathbb{C}$, $v_i \in \mathcal{H}_1$, $w_i \in \mathcal{H}_2$.*

The space of compact operators on Hilbert space $\mathcal{H}$ is denoted $\mathcal{K}(\mathcal{H})$ and can be further refined as follows. Let $(v_i)_{i\in\mathbb{N}}$ be an orthonormal basis of a (separable) Hilbert space $\mathcal{H}$, and set

$$\forall a \in \mathcal{B}(\mathcal{H})_+ : \ \operatorname{tr}(a) := \sum_{i=1}^{\infty}(v_i, av_i) \ . \tag{2.4}$$

Definition 8. *The set $\mathcal{L}^1(\mathcal{H})$ of trace-class operators on $\mathcal{H}$ is the linear span of*

$$\mathcal{L}^1(\mathcal{H})_+ := \{a \in \mathcal{B}(\mathcal{H})_+ \mid \operatorname{tr}(a) < \infty\} \ .$$

The set of Hilbert-Schmidt operators on $\mathcal{H}$ is given by

$$\mathcal{L}^2(\mathcal{H}) := \{b \in \mathcal{B}(\mathcal{H}) \mid b^*b \in \mathcal{L}^1(\mathcal{H})\} \ .$$

Note that in finite dimensions every operator is trace-class.

Lemma. $\mathcal{L}^1(\mathcal{H}) \subset \mathcal{L}^2(\mathcal{H}) \subset \mathcal{K}(\mathcal{H})$ *define closed two-sided ideals in* $\mathcal{B}(\mathcal{H})$.

We will encounter trace-class operators via normal states in Gleason's theorem, Thm. 22.

Definition 9. *A state σ on a von Neumann algebra $\mathcal{N}$ is called* normal *or* ultraweakly continuous *if $\sigma(a_i) \to \sigma(a)$ for every monotone increasing net $(a_i)_{i\in I}$ of operators in $\mathcal{N}$ with least upper bound a. Equivalently, $\sigma(\sum_i p_i) = \sum_i \sigma(p_i)$ for all families of orthogonal projections $(p_i)_{i\in I}$.*

For every normal state $\sigma \in \mathcal{S}(\mathcal{N})$, $\mathcal{N} \subseteq \mathcal{B}(\mathcal{H})$ there exists a trace-class operator $\rho \in \mathcal{L}^1(\mathcal{H})$ such that $\sigma(a) = \mathrm{tr}(\rho a)$ for all $a \in \mathcal{N}$ and vice versa.

Projections and spectral theorem. The set of *projections* in a C^*-algebra $\mathcal{A}$ is the set of self-adjoint, idempotent operators $\mathcal{P}(\mathcal{A}) := \{p \in \mathcal{A}_{\mathrm{sa}} \mid p^2 = p\}$. Projections play a particularly important role for von Neumann algebras. Given a faithful representation in the bounded operators on some Hilbert space $\mathcal{H}$ (cf. Thm. 5 below), the projections in a von Neumann algebra $\mathcal{P}(\mathcal{N})$ are in bijective correspondence with closed subspaces of $\mathcal{H}$. Since the latter are naturally ordered by inclusion, this equips $\mathcal{P}(\mathcal{N})$ with a partial order, which is closely related to contextuality (see Sec. 2.2.2 for more details). Algebraically, this order reads as follows: let $p, q \in \mathcal{P}(\mathcal{N})$, then $p \leq q$ if and only if $pq = qp = p$.

Moreover, the following two definitions will become important in the classification of von Neumann algebras below. Two projections $p, q \in \mathcal{P}(\mathcal{N})$ are called *equivalent* if there exists a partial isometry u such that $p = u^*u$ and $q = uu^*$.[7] A projection $p \in \mathcal{P}(\mathcal{N})$ is called *finite* if there exists no other projection $q < p$, $q \in \mathcal{P}(\mathcal{N})$, which is equivalent to p.

In contrast to general C^*-algebras, von Neumann algebras contain many projections. More precisely, let $F \subseteq \mathcal{B}(\mathcal{H})$ and define the *commutant of* F,

$$F' := \{b \in \mathcal{B}(\mathcal{H}) \mid \forall a \in F : [a, b] = ab - ba = 0\} \ .$$

Besides the topological condition in Def. 5 and Def. 6, von Neumann algebras can further be defined purely algebraically by means of von Neumann's double commutant theorem.

[7] Recall that $u \in \mathcal{N}$ is a partial isometry if u^*u and uu^* are both projections. As a map between Hilbert spaces, $u \in \mathcal{N} \subseteq \mathcal{B}(\mathcal{H})$ is a partial isometry if it is an isometry on the complement of its kernel.

Theorem 3. (von Neumann, double commutant theorem) *Let $\mathcal{H}$ be a Hilbert space and $\mathcal{B}(\mathcal{H})$ the algebra of bounded operators. Let $\mathcal{N} \subseteq \mathcal{B}(\mathcal{H})$ be a $*$-subalgebra of $\mathcal{B}(\mathcal{H})$ containing the identity. Then the following are equivalent:*

(i) $\mathcal{N}$ is closed in the weak (strong) operator topology, i.e., $\mathcal{N}$ is a von Neumann algebra

(ii) $\mathcal{N} = \mathcal{N}'' = (\mathcal{N}')'$

Moreover, every von Neumann algebra $\mathcal{N}$ is generated by its projections $\mathcal{N} = \mathcal{P}(\mathcal{N})''$.[8] The latter follows from the *spectral theorem*—another important decomposition of self-adjoint operators. In finite dimensions it reads

$$\forall a \in M_n(\mathbb{C}) \ (:= \mathcal{B}(\mathbb{C}^n)) : \ a = \sum_{i=1}^{n} a_i p_i \ . \tag{2.5}$$

Here, $a_i \in \mathbb{R}$ is an *eigenvalue of* a, i.e., $av = a_i v$ for some $v \in \mathcal{H} = \mathbb{C}^n$,[9] and $p_i \in \mathcal{P}(\mathcal{H})$ denotes the projection onto the corresponding eigenspace. A straightforward generalisation of Eq. (2.5) to infinite dimensions exists only for *compact operators* (cf. Thm. 2). General bounded self-adjoint operators do not necessarily have eigenvalues in infinite dimensions. Instead, one defines a *spectral value* $\lambda \in \mathbb{C}$ of an operator $a \in \mathcal{A}$ in a C^*-algebra $\mathcal{A}$ to be such that $a - \lambda 1$ does not have a (two-sided) inverse in $\mathcal{A}$. The collection of spectral values is called the *spectrum of* a and denoted $\mathrm{sp}(a)$.

We also need the notion of *operator-valued measure.*

Definition 10. *Let X be a compact Hausdorff space and $\sigma(X)$ its Borel σ-algebra. Then $\varrho : \sigma(X) \to \mathcal{B}(\mathcal{H})$ is called a operator-valued measure if it is weakly finitely additive, i.e., for any finite collection of disjoint Borel sets $(B_i)_{i \in I}$ in the σ-algebra $\sigma(X)$,*

$$\forall v, w \in \mathcal{H} : \ \langle \varrho(\cup_{i \in I} B_i) v, w \rangle = \sum_{i \in I} \langle \varrho(B_i) v, w \rangle \ .$$

A special case are projection-valued (or spectral) measures $\varrho : \sigma(X) \to \mathcal{P}(\mathcal{H})$.

[8]Since projections correspond with measurement outcomes in experiments, this justifies the use of von Neumann algebras from an operational perspective.

[9]Note that eigenvalues of self-adjoint operators are necessarily real.

As a generalisation of Eq. (2.5) based on *projection-valued measures*, one then obtains.

Theorem 4. *Let $a \in \mathcal{B}(\mathcal{H})_{\mathrm{sa}}$. There exists a projection-valued measure $\mu_a : \sigma(\mathbb{R}) \to \mathcal{P}(\mathcal{H})$ such that*

$$a = \int_{\mathbb{R}} \lambda \; d\mu_a(\lambda) \; .$$

The spectral theorem underlies the Borel functional calculus, i.e., we can apply functions to operators as follows: let $f : X \to \mathbb{R}$ by a bounded measurable function, then there exists a unique bounded linear operator $f(a) : \mathcal{H} \to \mathcal{H}$ defined by $f(a) = \int_{\mathbb{R}} f(\lambda) d\mu_a(\lambda)$.

Since every operator $a \in \mathcal{N}$ is the unique sum of self-adjoint operators, $a = a_1 + ia_2$, $a_1, a_2 \in \mathcal{N}_{\mathrm{sa}}$, and every self-adjoint operator has a spectral resolution by the spectral theorem, it follows that $\mathcal{P}(\mathcal{N})$ generates $\mathcal{N}$. In this sense, the study of von Neumann algebras can be reduced to the study of its projections. Contextuality builds on this idea together with the essential properties inherent to $\mathcal{P}(\mathcal{N})$ to be discussed in Sec. 2.2.2.

Type classification for von Neumann algebras. Since von Neumann algebras are generated by their projections, their classification is in terms of projections also.

Definition 11. *The* center $\mathcal{Z}(\mathcal{N})$ *of a von Neumann algebra is the von Neumann subalgebra of those operators in $\mathcal{N}$ that commute with all other operators,*

$$\mathcal{Z}(\mathcal{N}) := \{a \in \mathcal{N} \mid \forall b \in \mathcal{N} : \; [a, b] = 0\} \; .$$

A von Neumann algebra $\mathcal{N}$ is called a factor *if it has trivial center.*

The building blocks of von Neumann algebras are factors, i.e., every von Neumann algebra has a unique decomposition into factors $\mathcal{N} = \int_X^{\oplus} \mathcal{N}_x d\mu(x)$.[10] In particular, the factor decomposition is encoded in the *central projections* $\mathcal{ZP} := \{p \in \mathcal{Z}(\mathcal{N}) \mid p^2 = p\}$. Furthermore, the classification of factors is intuitively given in terms of the 'size' of its projections. Since, up to rescaling for every factor there exists a unique trace 'measuring the size' of its projections (cf. Eq. (2.4)), this classification can equivalently be given in terms of the image under the trace.

[10] If $\mathcal{N}$ has finitely many factors, this simplifies to the direct sum $\mathcal{N} = \bigoplus_{i \in I} \mathcal{N}_i$.

There are three types of factors, each with sub-types. A factor $\mathcal{N}$ is of type *I* if it contains a minimal projection, i.e., $\exists p \in \mathcal{P}(\mathcal{N})$ such that $q < p$ for $q \in \mathcal{P}(\mathcal{N})$ implies $q = 0$. Every such factor is isomorphic to the algebra of bounded operators on some Hilbert space, and since there exists a Hilbert space for every cardinal number, factors of this type are further distinguished by the Hilbert space dimension: I_n, $n \in (\mathbb{N} \cup \infty)$. These correspond with the matrix algebras $M_n(\mathbb{C})$. The size of the projections of a factor of type I_n is given by the unique (standard) trace $\mathrm{tr}(a) = \sum_{i=1}^n a_{ii}$ for $a_{ij} \in M_n(\mathbb{C})$, which for projections takes values $\{1, \cdots, n\}$. A factor with no minimal projection is of type *II*, if it still contains a non-zero finite projection. One further distinguishes between factors of type II_1 and II_∞, the former having a finite identity operator. Accordingly, the trace takes values in $[0, 1]$ or $[0, \infty]$, respectively. Finally, a factor is of type *III* if it contains no non-zero finite projection. Non-zero projections in such factors are therefore necessarily infinite and, in fact, all non-zero projections have the same (infinite) 'size', i.e., the trace takes values $\{0, \infty\}$ only. Factors of this type are often indexed by a real number, III_λ, $\lambda \in [0, 1]$, which relates to their Connes spectra [39].

Representation theory. We will mostly be concerned with von Neumann algebras. Yet, many key structural theorems hold on the level of C^*-algebras. Importantly, the *Gelfand-Naimark representation theorem* proves that any C^*-algebra has a faithful representation in the bounded operators on some Hilbert space. Even in the most general case we may thus consider observables as bounded self-adjoint operators on some Hilbert space. At the heart of this theorem is the *Gelfand-Naimark-Segal (GNS) construction.*

Theorem 5. *Let $\mathcal{A}$ be a C^*-algebra and σ a state on $\mathcal{A}$. Then there is a $*$-representation π_σ of $\mathcal{A}$ on a Hilbert space $\mathcal{H}_\sigma$ and a unit vector $v_\sigma \in \mathcal{H}_\sigma$ such that $\sigma(a) = (v_\sigma, \pi_\sigma(a)v_\sigma)$ for all $a \in \mathcal{A}$.*

The Hilbert space constructed in the proof of this theorem arises from the pre-inner product $(a, b) := \sigma(b^*a)$ for all $a, b \in \mathcal{A}$. In particular, the direct sum over all pure states yields a faithful representation π acting on the direct sum of Hilbert spaces $\mathcal{H}_\sigma$ by $\pi(a)(\bigoplus_\sigma v_\sigma) = \bigoplus_\sigma \pi_\sigma(a)v_\sigma$.[11]

On the other hand, note that we directly obtain a faithful representation from Thm. 5 if we are given a faithful, normal state, i.e., a faithful, normal, finite and normalised weight.

[11] This is the representation constructed in the Gelfand-Naimark representation theorem [86].

Yet, some infinite-dimensional von Neumann algebras do not admit such states, in particular, the finiteness condition fails. Nevertheless, one can show that every von Neumann algebra possesses a faithful, normal, semi-finite weight. In order to deal with the infinite-dimensional case, one therefore generalises the construction in Thm. 5 to weights by defining an inner product from $(x, y) := \omega(y^* x)$ only for $x, y \in n^\omega := \{x \in \mathcal{N} \mid \omega(x^* x) < \infty\}$. The representation obtained from this construction is called semi-cyclic and is sometimes denoted $L^2(\mathcal{N})$. One can show that it does not depend on the choice of faithful, normal, semi-finite weight.

We will use representations with respect to such weights in combination with the Riesz-Fréchet theorem in order to identify states and linear functionals.

Theorem 6. *Let $\mathcal{H}$ be a Hilbert space and $\phi \in \mathcal{H}^*$. Then there exists $w \in \mathcal{H}$ such that $\phi(v) = (v, w)$ for all $v \in \mathcal{H}$ and $||w||_{\mathcal{H}} = ||\phi||_{\mathcal{H}^*}$.*

Importantly, by semi-finiteness of ω, for every $x \in \mathcal{N}_+$ there exists a monotone increasing net $(x_i)_{i \in I} \in \mathcal{N}_+$ with limit x (in the strong operator topology), yet with $\omega(x_i) < \infty$ for all $i \in I$. It is thus sufficient to keep the correspondence under the Riesz-Fréchet theorem for finite elements (and appropriate monotone increasing nets).

2.2.2 Part II - Context category and presheaves

Partial orders and lattices. We saw that the inclusion relations between closed subspaces of a Hilbert space $\mathcal{H}$ make $\mathcal{P}(\mathcal{N})$ into a partial order for every von Neumann algebra $\mathcal{N} \subseteq \mathcal{B}(\mathcal{H})$. Since partial orders will be crucial also for the mathematical representation of the context category from Def. 3 in Sec. 2.1.2, we provide some more background on this structure here.

Definition 12. *A* partial order *is a set P with a binary relation $\preceq$ that satisfies:*

(i) $\forall p \in P: \ p \preceq p$ (reflexivity)

(ii) $\forall p, q, r \in P: \ p \preceq q$ *and* $q \preceq r$*, then* $p \preceq r$ (transitivity)

(iii) $\forall p, q \in P: \ p \preceq q$ *and* $q \preceq p$*, then* $q = p$ (antisymmetry)

A partial order is also an antisymmetric preorder, the latter being a set with a reflexive and transitive relation. For two elements $p, q \in P$ define the *least upper bound* or *join*, $p \vee q := \inf\{r \in$

$P \mid p \preceq r,\ q \preceq r\}$, and *greatest lower bound* or *meet*, $p \wedge q := \sup\{r \in P \mid r \preceq p,\ r \preceq q\}$. Note that joins and meets are unique if they exist. A partial order is a *join-semilattice/meet-semilattice* if for any two elements $p, q \in P$, $p \vee q / p \wedge q$ exists. Moreover, if P contains $\bigvee_{i \in I} p_i / \bigwedge_{i \in I} p_i$ for any family of elements $(p_i)_{i \in I}$ then P is called a *complete join-semilattice/meet-semilattice*. A partial order is called a *(complete) lattice* if it is both a (complete) meet-semilattice and a (complete) join-semilattice. It is bounded, if it contains a *least element* 0 and *greatest element* 1. A lattice L is called *distributive* if

$$\forall p, q, r \in L: \ p \vee (q \wedge r) = (p \vee q) \wedge (p \vee r) \quad \text{(distributivity)} \ .$$

Furthermore, a notion of *complement* is given as follows: let P be a bounded lattice and $p \in P$, then any $q \in P$ such that $q \vee p = 1$ and $q \wedge p = 0$ is called a *complement of* p. A *complemented lattice* is a lattice L with an *orthocomplementation* $^\perp : L \to L$, mapping every element $p \in P$ to a complement $p^\perp$. Complements need not be unique, however, they are unique for bounded distributive lattices.

Theorem. *A* Boolean algebra *(or* lattice*) is a bounded, orthocomplemented, distributive lattice.*

Classical physics is built on Boolean logic and thus on distributive lattices. More general are *modular* lattices, which satisfy:

$$\forall p, q, r \in L, p \preceq r: \ p \vee (q \wedge r) = (p \vee q) \wedge r \quad \text{(modularity)}$$

It is easy to see that the projections in a von Neumann algebra $\mathcal{P}(\mathcal{N})$ form a complemented lattice: the least element $0 \in \mathcal{P}(\mathcal{N})$ corresponds to the zero projection, the greatest element $1 \in \mathcal{P}(\mathcal{N})$ corresponds to the identity projection, and the map $^\perp : \mathcal{P}(\mathcal{N}) \to \mathcal{P}(\mathcal{N})$ defined via $p^\perp = 1 - p$ defines an orthocomplementation. Yet, $\mathcal{P}(\mathcal{N})$ is distributive only if $\mathcal{N}$ is commutative. Nevertheless, $L := \mathcal{P}(\mathcal{N})$ satisfies a weakened notion of modularity.

$$\forall p, q \in L, p \preceq q: \ p \vee (p^\perp \wedge q) = q \quad \text{(orthomodularity)} \tag{2.6}$$

Definition 13. *A lattice is called* orthomodular *if Eq. (2.6) holds.*

Theorem 7. *Let $\mathcal{N}$ be a von Neumann algebra. Then $\mathcal{P}(\mathcal{N})$ is a complete orthomodular lattice.*

Despite not being a Boolean algebra, an orthomodular lattice L is a lattice built from Boolean algebras, i.e., it can be obtained by pasting together Boolean lattices in a suitable way [68, 117]. An important consequence is that whenever $p, q \in L$ are *orthogonal*, $p \perp q$, i.e., $q \preceq p^{\perp}$, then there exists a Boolean sublattice $B \subseteq L$ containing both elements $p, q \in B$. This foreshadows a general theme in this chapter: quantum theory can be understood as a collection of classical perspectives ('contexts') with corresponding interrelations. To make this idea precise, after introducing some basic notions from category theory, in the subsequent paragraphs we define the notion of context in more detail and relate it to state spaces in classical physics. Quantum theory will then emerge from gluing together multiple contexts in an appropriate sense.

Categories and functors. We review some basic categorical definitions (cf. [104, 106, 124]).

Definition 14. *A* category $\mathcal{C}$ *is a collection of objects* $\mathrm{Ob}(\mathcal{C})$ *and a collection of arrows* $\mathrm{Arr}(\mathcal{C})$ *with domains and codomains in* $\mathrm{Ob}(\mathcal{C})$ *such that:*

1. *for any arrows $f, g \in \mathrm{Arr}(\mathcal{C})$ with $f : A \to B, g : B \to C$, there is an arrow $g \circ f \in \mathrm{Arr}(\mathcal{C})$ such that $g \circ f : A \to C$,*

2. *composition of arrows is associative, i.e., for any arrows $f, g, h \in \mathrm{Arr}(\mathcal{C})$ with $f : A \to B$, $g : B \to C$, and $h : C \to D$, it holds that $h \circ (g \circ f) = (h \circ g) \circ f$,*

3. *there is an identity arrow $id_A : A \to A$ for every object $A \in \mathrm{Ob}(\mathcal{C})$ such that for every arrow $f : A \to B$ it holds that $f \circ id_A = id_B \circ f$*

The collection of all arrows from A to B is denoted $\mathrm{Hom}_{\mathcal{C}}(A, B)$ *and is called the* hom-set between A and B.[12] *To each category $\mathcal{C}$ there exists an* opposite category $\mathcal{C}^{\mathrm{op}}$ *with the same objects,* $\mathrm{Ob}(\mathcal{C}^{\mathrm{op}}) = \mathrm{Ob}(\mathcal{C})$*, but all arrows reversed, i.e., whenever $A, B \in \mathrm{Ob}(\mathcal{C})$ and $(f : A \to B) \in \mathrm{Arr}(\mathcal{C})$ then there exists an arrow $(f : B \to A) \in \mathrm{Arr}(\mathcal{C}^{\mathrm{op}})$.*

[12] Note that hom-sets are not necessarily sets. If $\mathrm{Hom}_{\mathcal{C}}(A, B)$ is a set for all objects $A, B \in \mathrm{Ob}(\mathcal{C})$, $\mathcal{C}$ is called *locally small.*

Most if not all mathematical structures form categories in the appropriate sense. The one most familiar is the category **Set**, whose objects are sets and whose arrows are functions. Adding more structure to the objects leads to different categories, for instance, the category **Top** has objects topological spaces (sets with a topology) and arrows continuous maps, whereas the category **Grp** has as objects groups (sets with a group multiplication) and arrows group homomorphisms. Another example for a category is a partial order.

Example 1. *A partial order* $(P, \preceq)$ *forms a category with objects the elements in the set* P, $\mathrm{Ob}((P, \preceq)) = P$, *and arrows* $(p \to q) \in \mathrm{Arr}((P, \preceq))$ *between elements* $p, q \in P$ *whenever* $p \preceq q$. *As* $\preceq$ *is reflexive we have* $p \preceq p$, *which constitutes the identity arrow* $id_p : p \to p$ *for all* $p \in P$. *Moreover, composition of arrows is associative by transitivity of the partial order,* $p \preceq (q \preceq r) = (p \preceq q) \preceq r$ *for all* $p, q, r \in P$.

As with any structure defined in mathematics, of particular interest are the corresponding maps. For categories maps are functors.

Definition 15. *A* (covariant) functor $F : \mathcal{C} \to \mathcal{D}$ *between categories* $\mathcal{C}, \mathcal{D}$ *is defined*

1. *on objects: for all* $C \in \mathrm{Ob}(\mathcal{C})$, *there exists* $D \in \mathrm{Ob}(\mathcal{D})$ *such that* $F(C) = D$,
2. *on arrows: for all* $(f : C \to C') \in \mathrm{Arr}(\mathcal{C})$, *there exists* $(F(f) : F(C) \to F(C'))$ *such that* $\forall f, g, g \circ f \in \mathrm{Arr}(\mathcal{C}) : F(g \circ f) = F(g) \circ F(f)$ *and* $\forall C \in \mathrm{Ob}(\mathcal{C}) : F \circ id_C = id_{F(C)}$.

A contravariant functor $F' : \mathcal{C} \to \mathcal{D}$ *is a covariant functor* $F : \mathcal{C}^{\mathrm{op}} \to \mathcal{D}$, *i.e., a functor, which reverses the composition of arrows. Equivalently, a contravariant functor is a covariant functor* $F : \mathcal{C} \to \mathcal{D}^{\mathrm{op}}$. *A* presheaf *is a contravariant functor,* $P : \mathcal{C} \to \mathbf{Set}$.

Example 2. *Consider the covariant powerset functor* $\mathcal{P} : \mathbf{Set} \to \mathbf{Set}$. *To each set* X, *we assign its power set* $X \mapsto \mathcal{P}(X)$, *where the powerset* $\mathcal{P}(X)$ *is the set of all subsets of* X. *On arrows it maps functions to functions,*

$$(f : X \to Y) \longmapsto \mathcal{P}(f) : S \to \{f(x) \mid x \in S\} \quad \forall S \in \mathcal{P}(X)\ . \tag{2.7}$$

This construction preserves identity arrows and composition, it thus defines a (covariant) functor.

A more relevant example for the study of presheaves over the partial order of contexts is the following.

Example 3. *Take the three element set $\{A, B, C\}$, which we think of as containing the three edges of a triangle. Inclusion of subsets yields a partial order $(\mathcal{P}(\{A, B, C\}), \subseteq)$, which also forms a category. We can thus construct a presheaf $\Delta : (\mathcal{P}(\{A, B, C\}), \subseteq) \to$ **Set** by assigning all strict total orders over subsets $S \subseteq \mathcal{P}(\{A, B, C\})$. For instance, the subset $\{A, B\} \subseteq \{A, B, C\}$ can be assigned two strict total orders $\Delta(\{A, B\}) = \{A < B, B < A\}$. Δ becomes a presheaf if we restrict the sets of possible total orders accordingly, e.g.*

$$\Delta(\{A, B\} \subseteq \{A, B, C\}) : \Delta(\{A, B, C\}) \to \Delta(\{A, B\}),\ (\{A, B, C\}, <) \mapsto (\{A, B\}, < |_{\{A,B\}})\ .$$

Functors can have a number of properties, a non-exhaustive list is the following. A functor $F : \mathcal{C} \to \mathcal{D}$ is called *faithful* if for any two objects $C, C' \in \mathrm{Ob}(\mathcal{C})$ the map $\mathrm{Hom}_{\mathcal{C}}(C, C') \to \mathrm{Hom}_{\mathcal{D}}(F(C), F(C'))$ is injective, *full* if the map $\mathrm{Hom}_{\mathcal{C}}(C, C') \to \mathrm{Hom}_{\mathcal{D}}(F(C), F(C'))$ is surjective, and *essentially surjective* if every object $D \in \mathrm{Ob}(\mathcal{D})$ is isomorphic to $F(C)$ for some $C \in \mathrm{Ob}(\mathcal{C})$. Importantly, a faithful and full ('fully faithful'), and essentially surjective functor induces an *equivalence of categories*. The latter concept is more readily stated in terms of natural isomorphisms involving a functor $F : \mathcal{C} \to \mathcal{D}$ and its 'inverse'.

Definition 16. *Let $F, G : \mathcal{C} \to \mathcal{D}$ be functors between categories $\mathcal{C}, \mathcal{D}$. A* natural transformation from F to G, *$\eta : F \to G$, is a family of morphisms such that*

1. *for every $C \in \mathrm{Ob}(\mathcal{C})$, there exists a morphism $\eta_C : F(C) \to G(C)$ between objects of $\mathcal{D}$,*

2. *for every morphism $(f : C \to C') \in \mathrm{Arr}(\mathcal{C})$, it holds that $\eta_{C'} \circ F(f) = G(f) \circ \eta_C$.*

Moreover, η is called a natural isomorphism *if η_C is an isomorphism in $\mathcal{D}$ for every object $C \in \mathrm{Ob}(\mathcal{C})$.*

Two categories $\mathcal{C}, \mathcal{D}$ are called equivalent *if there exist functors $F : \mathcal{C} \to \mathcal{D}$ and $G : \mathcal{D} \to \mathcal{C}$ with natural isomorphisms $\eta : GF \to I_{\mathcal{C}}$ and $\vartheta : FG \to I_{\mathcal{D}}$ to the identity functors $I_{\mathcal{C}}$ on $\mathcal{C}$ and $I_{\mathcal{D}}$ on $\mathcal{D}$, respectively. One speaks of a duality of categories if F, G are contravariant functors.*

Contexts and classical (Stone/Gelfand) dualities. Let $\mathcal{N}$ be a von Neumann algebra. By Thm. 7, $\mathcal{P}(\mathcal{N})$ is a complete orthomodular lattice—it is glued together from Boolean lattices. The latter can be interpreted as classical state spaces as follows. Given a Boolean algebra B, a *state* λ of B is a homomorphism $\lambda : B \to \{0, 1\}$. The set of all states of B is called its *spectrum*, and denoted $\Omega(B)$. Equipped with the *Stone topology* generated by the sets $U_b := \{\lambda \in \Omega(B) \mid \lambda(b) = 1\}$ for $b \in B$, Ω is a compact, totally disconnected Hausdorff space, called the *Stone space* of B.[13] A compact, extremely disconnected Hausdorff space is called *Stonean*. *Stone duality* lifts the correspondence between Boolean logic and topological spaces to categorical equivalences.

Theorem 8. *The category of totally (extremely) disconnected Hausdorff spaces and (open) continuous maps,* **Stone** (**Stonean**)*, is dually equivalent to the category of (complete) Boolean algebras and lattice homomorphisms preserving* $0, 1$*,* **BA** (**cBA**).[14]

On the other hand, recall the definition of the *Gelfand spectrum* for an abelian C^*-algebra $\mathcal{A}$ in Eq. (2.3), $\Sigma_\mathcal{A} := \{0 \neq \lambda : \mathcal{A} \to \mathbb{C} \text{ linear} \mid \forall a, b \in \mathcal{A} : \ \lambda(ab) = \lambda(a)\lambda(b)\}$, equipped with the weak* topology. The *Gelfand transformation*, $\mathcal{G} : \mathcal{A} \to C(\Sigma_\mathcal{A})$, defined by $\mathcal{G}(a) := (\overline{a} : \Sigma_\mathcal{A} \to \mathbb{C}, \overline{a}(\lambda) := \lambda(a))$ is an isometric $*$-isomorphism between C^*-algebras and gives rise to the categorical correspondence known as *Gelfand duality*.

Theorem 9. *The category of unital, commutative C^*-algebras and $*$-homomorphisms,* $\mathbf{ucC^*}$*, is dually equivalent to the category of compact Hausdorff spaces and continuous maps,* **KHaus**.

Recall that the elements in the Gelfand spectrum of a commutative C^*-algebra are multiplicative states, which in turn correspond with pure states. It follows that given a commutative C^*-algebra, the set of its pure states corresponds with the points in the Hausdorff space dual. In particular, every commutative von Neumann algebra thus constitutes such a classical state space, which additionally comes with a Boolean logic inherited from its projections, since for $\mathcal{N}$ commutative, $\mathcal{P}(\mathcal{N})$ is a complete Boolean algebra. Conversely, every classical state space is a commutative von Neumann algebra if a minimal notion of measurability is added. More precisely,

[13] Recall that a topological space is *totally disconnected* if its only connected components are singletons, and is *extremely disconnected* if the closure of every open subset is open.

[14] Here and throughout we mark categories in boldface.

a Stonean space Ω is called *Hyperstonean* if it admits sufficiently many normal measures: for any non-zero positive continuous function $f : \Omega \to \mathbb{R}$, there exists a positive normal measure μ with $\mu(f) > 0$. (See also [120] and Thm. 35 for more details on the relation with measurability.)

Theorem 10. *The category of commutative von Neumann algebras and normal, unital $*$-homomorphisms,* **cvNA***, is dually equivalent to the category of Hyperstonean spaces and open maps,* **HStone***.*

Combining classical logic of Boolean algebras with measurability thus inevitably points to the study of commutative von Neumann algebras. Succinctly,

Classical state spaces correspond with commutative von Neumann algebras.

In the quantum case, von Neumann algebras are generally noncommutative, nevertheless $\mathcal{N}$ contains many commutative von Neumann subalgebras.

Definition 17. *Let $\mathcal{N}$ be a von Neumann algebra representing the physical quantities of a theory. A* context *is a commutative von Neumann subalgebra $V \subseteq \mathcal{N}$.*

Physically, a context is a set of simultaneously measurable observables (cf. Def. 2). Mathematically, contexts are modeled by commutative von Neumann subalgebras of the noncommutative von Neumann algebra $\mathcal{N}$ describing the observables of a physical system. Every context provides a singular, classical perspective onto the physical system represented by $\mathcal{N}$.

Context category and coarse-graining. Let the algebra of physical quantities be modeled by a von Neumann algebra $\mathcal{N}$. The collection of all commutative von Neumann subalgebras ('contexts') carries a natural notion of *coarse-graining* in the form of ordering contexts by inclusion. This yields a partial order.

Definition 18. *Let $\mathcal{N}$ be a von Neumann algebra. The* context category *$\mathcal{V}(\mathcal{N})$ is the partial order of commutative von Neumann subalgebras $V \subseteq \mathcal{N}$ ordered by inclusion,*

$$\mathcal{V}(\mathcal{N}) := (\{\text{commutative von Neumann subalgebras of } \mathcal{N}\}, \subset) \,.$$

For $\mathcal{N} = \mathcal{B}(\mathcal{H})$ we use the shorthand $\mathcal{V}(\mathcal{H}) := \mathcal{V}(\mathcal{B}(\mathcal{H}))$.

The term 'context category' stems from the fact that every partial order defines a category of its own. $\mathcal{V}(\mathcal{N})$ is a meet-semilattice with least element the *trivial context*, $V^0 := \{1\}'' = \mathbb{C}1$, and maximal contexts the maximal abelian subalgebras of $\mathcal{N}$. A useful example to have in mind is the following.

Example 4. *Let $\mathcal{H} = \mathbb{C}^3$ and $\mathcal{N} = \mathcal{B}(\mathcal{H}) = M_3(\mathbb{C})$. Its context category $\mathcal{V}(\mathcal{H}) := \mathcal{V}(\mathcal{B}(\mathcal{H}))$ has three 'layers'. The lowest layer contains only the trivial context $V^0 := \mathbb{C}1$. It is contained in any context generated by a single rank-1 projection $p_1 \in \mathcal{P}(\mathcal{H})$ such as $V := \{p_1, 1\}'' = \mathbb{C}p_1 + \mathbb{C}(1-p_1)$. Finally, the 'top layer' contains the maximal contexts generated by three mutually orthogonal rank-1 projections, $V := \{p_1, p_2, p_3\}'' = \mathbb{C}p_1 + \mathbb{C}p_2 + \mathbb{C}p_3$.*

Inclusion relations arise by coarse-graining: given a maximal context $V := \{p_1, p_2, p_3\}''$, there are three subcontexts $\{p_1, (p_2 + p_3)\}'' = \{p_1, 1\}'' \subset V$, $\{p_2, (p_1 + p_3)\}'' = \{p_2, 1\}'' \subset V$, and $\{p_3, (p_1 + p_2)\}'' = \{p_3, 1\}'' \subset V$. Similarly, the trivial context arises by coarse-graining from $V = \{p, 1\}''$ via $V^0 = \{1\}'' = \{p + (1-p)\}'' \subseteq \{p, (1-p)\}'' = \{p, 1\}'' = V$.

Gluing of local data and global sections. A key ingredient to our programme are presheaves over the context category, they collect classical information of a particular kind *locally*, i.e., *within contexts*, together with the constraints describing how contexts are nested. We would like to study how restrictive these constraints are to information accessible *globally*, i.e., consistent *across contexts*. This naturally leads to *(global) sections*. We briefly review the basic concept behind this definition and embed it into the general study of sheaves.

Note that there is a natural topology on partial orders called the *Alexandrov topology*.

Definition 19. *Let P be a preorder. The upper (lower) Alexandrov topology $\tau_\uparrow$ ($\tau_\downarrow$) contains the upper (lower) sets,*

$$\tau_\uparrow = \{U \subseteq P \mid \forall x \in U, y \in P : x \preceq y \Rightarrow y \in U\} \ .$$

The closed sets are the lower (upper) sets,

$$\tau_\downarrow = \{U \subseteq P \mid \forall x \in U, y \in P : x \succeq y \Rightarrow y \in U\} \ .$$

$\mathcal{P}(\mathcal{N})$ with the upper (lower) Alexandrov topology $\tau_\uparrow$ ($\tau_\downarrow$) thus becomes a topological space. More generally, a presheaf $\underline{P} : P \to \mathcal{C}$ captures the idea of associating data to the open sets of a topological space (elements in a partial order) in such a way that reflects the inclusion relations between them.[15] In particular, given an open set $U \subseteq P$ and an open cover $(U_i)_{i\in I}$, i.e., U_i open and $\bigcup_{i\in I} U_i = U$, a *(local) section of* $\underline{P}$ *over* U corresponds to a collection of elements $(\gamma_{U_i})_{i\in I}$, $\gamma_{U_i} \in \underline{P}(U_i)$ that 'fit together' under restriction maps by means of functoriality, i.e., whenever $U_i \subseteq U_j$ then $\gamma_{U_i} = \underline{P}(U_i \subseteq U_j)(\gamma_{U_j})$. A *global section* is a section with $U = P$. We will mostly be concerned with the existence of (global) sections of presheaves over $\mathcal{V}(\mathcal{N})$, $\underline{P} : \mathcal{V}(\mathcal{N}) \to \mathbf{Set}$, equipped with e.g. the lower Alexandrov topology $\tau_\downarrow$.

Definition 20. *Let* $\underline{P} : \mathcal{V}(\mathcal{N}) \to \mathbf{Set}$ *be a presheaf over the context category* $\mathcal{V}(\mathcal{N})$ *corresponding to some von Neumann algebra* $\mathcal{N}$*. A* global section of $\underline{P}$ *is a collection of elements* $(\underline{P}(V))_{V\in\mathcal{V}(\mathcal{N})}$ *such that* $\underline{P}(\tilde{V} \subseteq V)(\underline{P}(V)) = \underline{P}(\tilde{V})$*. The collection of global sections of* $\underline{P}$ *is denoted by* $\Gamma(\underline{P})$*.*

Sections always exist locally, yet the existence of global sections depends on the type of constraints imposed by the order structure in $\mathcal{V}(\mathcal{N})$. To see how such constraints arise, we return to our toy example, Ex. 3.

Example 5. *Notice that* Δ *in Ex. 3 does have global sections, in fact, every strict total order on* $\{A, B, C\}$ *restricts to total orders on subsets. Yet, given strict total orders on a (covering) collection of subsets, there does not always exist a global section that restricts to them. For instance, the Penrose tribar (cf. Fig. 2.1) gives rise to local sections* $(A < B) \in \Delta(\{A, B\})$, $(B < C) \in \Delta(\{B, C\})$*, and* $(C < A) \in \Delta(\{A, C\})$*, which do not arise from a global section.*

Recall that classical states correspond with elements in the Gelfand spectrum corresponding to the abelian von Neumann algebra of physical quantities. A natural generalisation to noncommutative von Neumann algebras suggests that quantum states should correspond with global sections of some presheaf over the partial order of contexts. In Sec. 2.3 we will verify this hunch. In fact, the language of presheaves over the context category will prove capable of capturing many more aspects of quantum theory.

[15] Here and throughout we mark presheaves with an underscore.

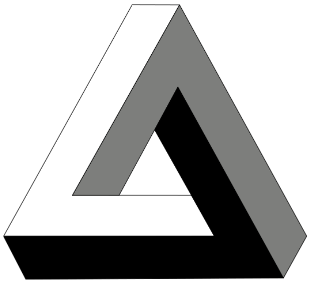


Figure 2.1: The Penrose tribar [122] can be viewed as a visualisation of three local sections that do not combine into a global section. It captures the idea of structures that satisfy some property locally but not necessarily globally.

2.2.3 Part III - Jordan algebras and orientations

We introduce an important classification of von Neumann algebras arising from the decomposition of its associative product into symmetric and antisymmetric component. Physically, this relates to the dichotomic function of self-adjoint operators as measurements and generators of time evolution in quantum theory. This section is mainly based on work in [46] and more generally in [8, 39], and serves as an introduction to Sec. 2.3.4 and Sec. 2.4.4.

Jordan algebras. The product on a von Neumann algebra $\mathcal{N}$ has the decomposition,

$$\forall a, b \in \mathcal{N}: \ ab = \frac{1}{2}(ab + ba) + \frac{1}{2}(ab - ba) =: \frac{1}{2} a \circ b + \frac{1}{2}[a, b] \ .$$

The latter antisymmetric bracket $[\cdot, \cdot]$ is known as the *commutator*, it makes $(\mathcal{N}, [\cdot, \cdot])$ into a Lie algebra. Lie algebras also arise in classical physics from Poisson brackets between measurable functions on a Poisson manifold. The former symmetric product is known as the *anticommutator* or *Jordan product* and makes $(\mathcal{N}, \circ)$ into a (special) Jordan algebra (cf. Def. 22). Jordan algebras arose out of an attempt to equip the observables in quantum theory with an algebraic relation [93, 94]. Importantly, the associative product in a von Neumann algebra does not close over its self-adjoint elements, i.e., $ab \notin \mathcal{N}_{\mathrm{sa}}$ for $a, b \in \mathcal{N}_{\mathrm{sa}}$ in general, but the Jordan product does. The Jordan product is commutative by construction, however, it is generally non-associative, $a \circ (b \circ c) \neq (a \circ b) \circ c$. (It is associative if and only if the associative product is commutative.) Jordan algebras are less studied, partly because their classical counterpart is trivial, i.e., associative. Note that the symmetric, associative product in the algebra of measurable functions on a Poisson manifold is given by pointwise multiplication. From this point of view, the non-classical aspects of quantum theory are more appropriately described by its non-trivial Jordan structure rather than noncommutativity, as is commonly stated. However, despite some efforts [55] Jordan algebras have not found resonance with a wider audience in physics. As we will see below, they are deeply intertwined with contextuality.

Definition 21. *A Jordan algebra $\mathcal{J}$ is an algebra with a product $\circ$, which satisfies $a \circ b = b \circ a$ and $a \circ (b \circ a^2) = (a \circ b) \circ a^2$ for all $a, b \in \mathcal{J}$.*

Most Jordan algebras arise by symmetrisation of an associative algebra [111].

Definition 22. *Let $\mathcal{A}$ be an associative algebra over a field $\mathbb{F}$ (not of characteristic 2). The vector space $\mathcal{A}$ equipped with the bilinear operation $\circ$ defined by*

$$\forall a, b \in \mathcal{A}: \ a \circ b := \frac{1}{2}(ab + ba)$$

is called the special Jordan algebra $\mathcal{J}(\mathcal{A})$ associated with $\mathcal{A}$.

Every Jordan algebra that does not arise in this way is called *exceptional*. The latter have been related to spin factors [8]. We will only be interested in special Jordan algebras $\mathcal{J}(\mathcal{N})$ associated with von Neumann algebras $\mathcal{N}$, we will therefore drop the classifier *special* in what follows.

A linear map between Jordan algebras $\Phi : \mathcal{J}(\mathcal{N}_1) \to \mathcal{J}(\mathcal{N}_2)$ is called a *Jordan homomorphism* if for all $a, b \in \mathcal{J}(\mathcal{N}_1)$: $\Phi(a \circ b) = \Phi(a) \circ \Phi(b)$. Recall that $\mathcal{N} = \mathcal{N}_{\mathrm{sa}} + i\mathcal{N}_{\mathrm{sa}}$ is the complexification of $\mathcal{J}(\mathcal{N})$. A Jordan $*$-homomorphism is a linear map $\Phi : \mathcal{N}_1 \to \mathcal{N}_2$ such that for all $a \in \mathcal{N}_1$: $\Phi(a^*) = \Phi(a)^*$. The corresponding definitions for isomorphisms read accordingly.

Contextuality and Jordan algebras. In a deep result by Dye, Jordan $*$-isomorphisms between Jordan algebras associated with von Neumann algebras have been related with automorphisms of their projection lattices.

Theorem 11. (Dye [52]) *Let $\mathcal{N}_1$, $\mathcal{N}_2$ be von Neumann algebras with no direct summands of type I_2. Every orthoisomorphism $\varphi : \mathcal{P}(\mathcal{N}_1) \to \mathcal{P}(\mathcal{N}_2)$ can be uniquely extended to a Jordan $*$-isomorphism $\Phi : \mathcal{N}_1 \to \mathcal{N}_2$.*

A reformulation of this result in terms of contexts arises by the close connection of the projection lattice and the context category as observed in [76].

Theorem 12. (Harding-Navara [76]) *Let $\mathcal{N}$ be a von Neumann algebra not isomorphic to $\mathbb{C} \oplus \mathbb{C}$ or to $M_2(\mathbb{C})$. Then the context category $\mathcal{V}(\mathcal{N})$ of $\mathcal{N}$ determines the projection lattice $\mathcal{P}(\mathcal{N})$ as an orthomodular lattice up to isomorphism. Conversely, the projection lattice $\mathcal{P}(\mathcal{N})$ determines the poset $\mathcal{V}(\mathcal{N})$ up to isomorphism.*

In fact, Harding and Navara's proof holds for all orthomodular lattices, which contain no maximal Boolean sublattices with only 4 elements (this is why the trivial cases $\mathcal{N} = \mathbb{C} \oplus \mathbb{C}$ or $\mathcal{N} = M_2(\mathbb{C})$ are excluded). The result shows that the context category, i.e., the collection of contexts together with their nesting relations, encodes the same amount of information as the projection lattice. This allows for a reformulation of Dye's theorem in contextual form.

Theorem 13. (Döring-Harding [50]) *Let $\mathcal{N}$ be a von Neumann algebra not isomorphic to $\mathbb{C} \oplus \mathbb{C}$ or $M_2(\mathbb{C})$. For every order automorphism $\varphi : \mathcal{V}(\mathcal{N}) \to \mathcal{V}(\mathcal{N})$, there is a unique Jordan $*$-automorphism $\Phi : (\mathcal{N}, \cdot) \to (\mathcal{N}, \cdot)$ such that $\varphi(V) = \Phi[V]$ for all $V \in \mathcal{V}(\mathcal{N})$.*

The essence of this theorem is that the mere order structure between contexts is rich enough to encode the algebra up to Jordan $*$-isomorphism. Recall that this order structure is trivial for classical systems, since there one only has a single (maximal) context. This is consistent with the fact that the Jordan product is trivial for classical systems, i.e., it reduces to the commutative (pointwise) multiplication of functions. Physical contextuality thus reflects the additional algebraic structure in quantum theory inherent to the Jordan product. Succinctly,

Contextuality is Jordan structure.

In other words, quantum theory is different from classical theory by its non-trivial context structure. This result sits somewhat opposite to the commonly recited doctrine that the essence of quantum theory is the non-vanishing of the commutator.

Contextuality and von Neumann algebras. It is an obvious question to ask whether two algebras that are Jordan $*$-isomorphic are already isomorphic as von Neumann algebras. If this was the case the same would necessarily hold for factors, yet this was shown not to be the case in the seminal work by Connes.

Theorem 14. (Connes [38]) *Two von Neumann algebras $\mathcal{N}_1$, $\mathcal{N}_2$ that are Jordan $*$-isomorphic need not be isomorphic as von Neumann algebras.*

Since $\mathcal{V}(\mathcal{N})$ encodes the same information as the Jordan algebra $\mathcal{J}(\mathcal{N})$ [50], contextuality also contains strictly less information than the associative product on $\mathcal{N}$. The difference between the former and the latter boils down to a choice of time direction on factors.

Order derivations, dynamical correspondences, and time orientations. The product in a C^*-algebra $\mathcal{A}$ contains more information than the associated Jordan product in $\mathcal{J}(\mathcal{A})$. For C^*-algebras it is natural to consider unital *Jordan-Banach algebras (JB-algebras)*, i.e., Banach algebras with a Jordan product. Moreover, for von Neumann algebras the additional topological condition on weak-closure is captured by (unital) *weakly-closed Jordan-Banach algebras (JBW-algebras)*. Clearly, maps between such algebras $\Phi : \mathcal{J}(\mathcal{A}_1) \to \mathcal{J}(\mathcal{A}_2)$ lift to maps between C^*/von Neumann algebras if and only if they preserve commutators. This extra information can be encoded more algebraically as shown by Connes, Alfsen, and Shultz [7, 40].

Definition 23. *An* order derivation *δ on a JB(W)-algebra $\mathcal{J}$ is a bounded linear operator such that $e^{t\delta}(\mathcal{J}_+) \subseteq \mathcal{J}_+$ for all $t \in \mathbb{R}$, i.e., $t \mapsto e^{t\delta}$ is a one-parameter group of order automorphisms.*

An order derivation δ is called *self-adjoint* if $\delta = \delta_a$ for some $a \in \mathcal{J}$, where $\delta_a : \mathcal{A} \to \mathcal{A}$, $\delta_a(b) = a \circ b$. An order derivation δ is *skew-adjoint* if $\delta(1) = 0$. The set of skew order derivations is denoted $\mathrm{OD}_s(\mathcal{J})$. Every order derivation can be decomposed uniquely as the sum of a self-adjoint and a skew-adjoint order derivation. Moreover, one has the following (cf. [7]).

Proposition 1. *If $\mathcal{J}(\mathcal{N}) = (\mathcal{N}_{\mathrm{sa}}, \circ)$ is the JBW-algebra associated with the self-adjoint part of a von Neumann algebra $\mathcal{N}$, then every order derivation on $\mathcal{J}(\mathcal{N})$ is of the form*

$$\delta_a : \mathcal{J}(\mathcal{N}) \to \mathcal{J}(\mathcal{N}), \quad \delta_a(b) := \frac{1}{2}(ab + ba^*)$$

for some $a \in \mathcal{N}$. An order derivation is self-adjoint if and only if $\delta = \delta_a$ for some self-adjoint $a \in \mathcal{N}_{\mathrm{sa}}$ and is skew-adjoint if and only if $\delta = \delta_{ia} = \frac{i}{2}[a, _]$for some $a \in \mathcal{N}_{\mathrm{sa}}$.

The importance of order derivations is that maps between Jordan algebras $\Phi : \mathcal{J}(\mathcal{N}_1) \to \mathcal{J}(\mathcal{N}_2)$ lift to maps between von Neumann algebras if and only if they preserve skew order derivations [46].

Proposition 2. *A (normal) unital Jordan homomorphism $\Phi : \mathcal{J}(\mathcal{N}_1) \to \mathcal{J}(\mathcal{N}_2)$ corresponding to von Neumann algebras $\mathcal{N}_1$, $\mathcal{N}_2$ extends to a (normal) unital homomorphism of von Neumann algebras if and only if*

$$\forall a \in (\mathcal{N}_1)_{\mathrm{sa}}, \forall t \in \mathbb{R} : \; \Phi \circ e^{t\delta_{ia}} = e^{t\delta_{i\Phi(a)}} \circ \Phi \; . \tag{2.8}$$

Proof. The proof is straightforward. First, by the exponential series expansion of $e^{t\delta_{ia}}$, Eq. (2.8) is equivalent to $\Phi \circ \delta_{ia} = \delta_{i\Phi(a)} \circ \Phi$ for all $a \in (\mathcal{N}_1)_{\mathrm{sa}}$. Hence, if Φ is a (normal) unital Jordan homomorphism such that Eq. (2.8) holds, we have

$$\forall a, b \in (\mathcal{N}_1)_{\mathrm{sa}} : \; \Phi(\frac{i}{2}[a, b]) = (\Phi \circ \delta_{ia})(b) = (\delta_{i\Phi(a)} \circ \Phi)(b) = \frac{i}{2}[\Phi(a), \Phi(b)] \; ,$$

so Φ preserves all commutators between self-adjoint operators. Since any operator $a \in \mathcal{N}_1$ can be decomposed uniquely as $a = a_1 + ia_2$ for $a_1, a_2 \in (\mathcal{N}_1)_{\mathrm{sa}}$, it easily follows that Φ preserves all commutators. Hence, it is a (normal) unital homomorphism $\Phi : \mathcal{N}_1 \to \mathcal{N}_2$ of von Neumann algebras. (The statement for normal morphisms holds as $e^{t\delta}$ is normal.)

Conversely, if $\Phi : \mathcal{N}_1 \to \mathcal{N}_2$ is a homomorphism of von Neumann algebras, then its restriction to $(\mathcal{N}_1)_{\mathrm{sa}}$ is a Jordan homomorphism onto $(\mathcal{N}_2)_{\mathrm{sa}}$ such that condition Eq. (2.8) holds. This completes the proof. □

In order to encode the additional structure inherent to the antisymmetric part of the associative product, i.e., the commutator, one thus needs skew order derivations. Prop. 1 defines the latter for JBW-algebras in relation to a given von Neumann algebra $\mathcal{N}$. Independent of $\mathcal{N}$ and more generally for JB-algebras, one defines a *dynamical correspondence*[16] ψ as a map into

[16]The name indicates the dual role of self-adjoint operators as observables and generators of dynamics (cf. [14]).

skew order derivations on $\mathcal{J}$: $\psi : \mathcal{J} \longrightarrow \mathrm{OD}_s(\mathcal{J})$ with $x \longmapsto \psi_x$ such that (i) $[\psi_x, \psi_y] = -[\delta_x, \delta_y]$ and (ii) $\psi_x(x) = 0$ for all $x, y \in \mathcal{J}$ [7]. Similarly to skew order derivations in Prop. 2, dynamical correspondences then classify associative products on unital JB(W)-algebras.

Theorem 15. (Alfsen-Shultz [7]) *A unital JB-algebra $\mathcal{A}$ is (isomorphic to) the self-adjoint part of a C^*-algebra if and only if there exists a dynamical correspondence on $\mathcal{A}$. In this case each dynamical correspondence ψ on $\mathcal{A}$ determines a unique Jordan compatible C^*-product such that $\psi_a b = \frac{i}{2}(ab - ba)$ for each pair $a, b \in \mathcal{A}$, and each Jordan compatible C^*-product arises in this way from a unique dynamical correspondence ψ on $\mathcal{A}$. The same conclusions hold with JBW in place of JB and von Neumann in place of C^*.*

Recall that by Thm. 14, two different von Neumann algebras can induce the same underlying Jordan algebra. A Jordan compatible C^*-product therefore means any associative product on $\mathcal{A}$, which reduces to the given Jordan product (and similarly for von Neumann algebras). By Thm. 15 the different associative products on C^*-algebras (von Neumann algebras) corresponding to the same unital JB(W)-algebra are classified by dynamical correspondences. For von Neumann algebras this can be further refined to factors by the following result in [8].

Theorem 16. *Let $\star_1$, $\star_2$ be two associative products on $\mathcal{J}(\mathcal{N}_1) \simeq \mathcal{J}(\mathcal{N}_2)$ corresponding to von Neumann algebras $\mathcal{N}_1$, $\mathcal{N}_2$, respectively. Then $\star_1$, $\star_2$ differ by a central projection $p \in \mathcal{Z}(\mathcal{N}_1) \simeq \mathcal{Z}(\mathcal{N}_2)$, which is 1 on the abelian part of $\mathcal{N}_1, \mathcal{N}_2$ in the following sense:*

$$\forall a, b \in \mathcal{J}(\mathcal{N}_1) \simeq \mathcal{J}(\mathcal{N}_2): \;\; a \star_2 b = p \star_1 a \star_1 b + (1 - p) \star_1 b \star_1 a$$

For single factors this boils down to a choice of sign in the commutator $a \star b = \pm[a, b]$. In particular, a dynamical correspondence thus corresponds to a unique sign choice for commutators in every factor. Moreover, by interpreting the parameter t in Prop. 2 as time, this sign choice corresponds with a choice of time direction in every factor.

Finally, these concepts can be lifted to the context category [46]. By Thm. 15 a von Neumann algebra $\mathcal{N} \leftrightarrow (\mathcal{J}(\mathcal{N}), \psi_{\mathcal{N}})$ is a JBW-algebra $\mathcal{J}(\mathcal{N})$ together with the dynamical correspondence encoding the time direction on $\mathcal{N}$ in every factor $\psi_{\mathcal{N}}$. By Thm. 12 in [76], the JBW-algebra $\mathcal{J}(\mathcal{N})$ is equivalently encoded in the context category $\mathcal{V}(\mathcal{N})$, hence, $\mathcal{N} \leftrightarrow (\mathcal{V}(\mathcal{N}), \psi_{\mathcal{N}})$ also.

To make this precise, note that every one-parameter group of order automorphisms, $t \mapsto e^{t\delta_{ia}}$ for $a \in \mathcal{N}_{\mathrm{sa}}$, also defines a one-parameter group of order automorphisms of $\mathcal{V}(\mathcal{N})$,

$$\begin{aligned} \widetilde{e^{t\delta_{ia}}} : \mathcal{V}(\mathcal{N}) &\longrightarrow \mathcal{V}(\mathcal{N}) \\ V &\longmapsto e^{\frac{i}{2}ta} V e^{\frac{-i}{2}ta} \,. \end{aligned}$$

We denote the group of order automorphisms of $\mathcal{V}(\mathcal{N})$, i.e., maps $\varphi : \mathcal{V}(\mathcal{N}) \to \mathcal{V}(\mathcal{N})$ such that $\varphi(\tilde{V}) \leq \varphi(V) \Leftrightarrow \tilde{V} \leq V$, by $\mathrm{Aut}(\mathcal{V}(\mathcal{N}))$. The following definition appears in [46].

Definition 24. *Let $\mathcal{N}$ be a von Neumann algebra and $\mathcal{V}(\mathcal{N})$ its context category. The map*

$$\begin{aligned} \widetilde{\psi} : \mathcal{N}_{\mathrm{sa}} \times \mathbb{R} &\longrightarrow \mathrm{Aut}(\mathcal{V}(\mathcal{N})) \\ (a, t) &\longmapsto \widetilde{e^{t\delta_{ia}}} \end{aligned}$$

is called the time orientation on order-automorphisms of $\mathcal{V}(\mathcal{N})$ *induced by $\mathcal{N}$. When $\mathcal{V}(\mathcal{N})$ is equipped with this time orientation, it is called the* oriented context category of $\mathcal{N}$, *denoted $\widetilde{\mathcal{V}(\mathcal{N})}$.*

Time orientations encode the forward time direction in a quantum system.

The equivalence $\mathcal{N} \leftrightarrow (\mathcal{V}(\mathcal{N}), \psi_{\mathcal{N}})$ for the context order $\mathcal{V}(\mathcal{N})$ according to Thm. 15 then reads.

Theorem 17. (Döring [46]) *Let $\mathcal{N}_1, \mathcal{N}_2$ be von Neumann algebras not isomorphic to $\mathbb{C} \oplus \mathbb{C}$ and with no type I_2 summands. There is a bijective correspondence between isomorphisms $\Phi : \mathcal{N}_1 \to \mathcal{N}_2$ of von Neumann algebras and order isomorphisms $\varphi : \mathcal{V}(\mathcal{N}_1) \to \mathcal{V}(\mathcal{N}_2)$ that preserve the orientations on $\mathcal{V}(\mathcal{N}_1)$ and $\mathcal{V}(\mathcal{N}_1)$ induced by $\mathcal{N}_1$ and $\mathcal{N}_2$, respectively.*

We have seen earlier that contextuality is Jordan structure associated with von Neumann algebras. Thm. 17 completes this picture by adding time orientations: as one might expect, as a dynamical concept time orientations relate to commutators (encoded by skew order derivations as described above) and one-parameter families resulting from exponentiation.

2.3 Contextuality and the fundamental theorems in quantum theory

In this section we use the mathematical representation of physical contextuality and the corresponding context category in quantum theory to provide reformulations of key theorems in foundations. The value of these reformulations is twofold: (i) contextuality emerges as an underlying physical principle, unifying many seemingly unrelated aspects in quantum theory, (ii) the key structural components of quantum theory arise from local-to-global obstructions over suitable presheaves over the context category. This section gives a more detailed account of the results presented in the recent preprint [47].

As motivated in Sec. 2.2.1, the algebra of physical observables is naturally modeled by some noncommutative von Neumann algebra. We therefore strive for maximal generality within this framework. A natural generalisation to orthomodular lattices is discussed in Sec. 2.5 for Bell's theorem, for which we define a suitable presheaf in Sec. 2.4.

2.3.1 The Kochen-Specker theorem

Continuing our discussion in Sec. 2.1.1, we first give the original statement of the Kochen-Specker theorem. We have already discussed the essence of the argument, it remains to fill in the technical details. Recall that observables in quantum theory are represented by bounded self-adjoint operators $\mathcal{O} = \mathcal{B}(\mathcal{H})_{\mathrm{sa}}$ on some Hilbert space $\mathcal{H}$ (cf. Sec. 2.2.1), and two observables $a, b \in \mathcal{B}(\mathcal{H})_{\mathrm{sa}}$ are simultaneously measurable if they commute $[a, b] = 0$. In finite dimensions, a self-adjoint operator a is a Hermitian matrix, i.e., $a \in M_n(\mathbb{C})$ and $a^* = a$. By the spectral theorem, Thm. 4, every Hermitian matrix has an eigenvalue decomposition $a = \sum_{i=1}^n a_i p_i$, where the eigenvalues make up the spectrum of the observable, $\mathrm{sp}(a) = \{a_1, \cdots, a_n\}$, and the p_i's are the projections onto the corresponding eigenspaces. By Def. 1 every valuation function is a partial algebra homomorphism, which restricts to a partial Boolean algebra homomorphism $v : \mathcal{P}(\mathbb{C}^n) \to \{0, 1\}$ for $\mathcal{P}(\mathbb{C}^n)$. In an impressive combinatorial effort involving a total of 117 vectors in $\mathbb{R}^3$, Kochen and Specker go on to prove that no such partial Boolean algebra homomorphism exists whenever the dimensions of the Hilbert space is at least three.

Theorem 18. (Kochen-Specker [101]) *Let $\mathcal{H}$ be a Hilbert space,* $\dim(H) \geq 3$*, and* $\mathcal{B}(\mathcal{H})$ *represent the algebra of physical quantities of some quantum system. Then there exists no valuation function as in Def. 1. In particular, there exists no classical state space for the quantum system.*

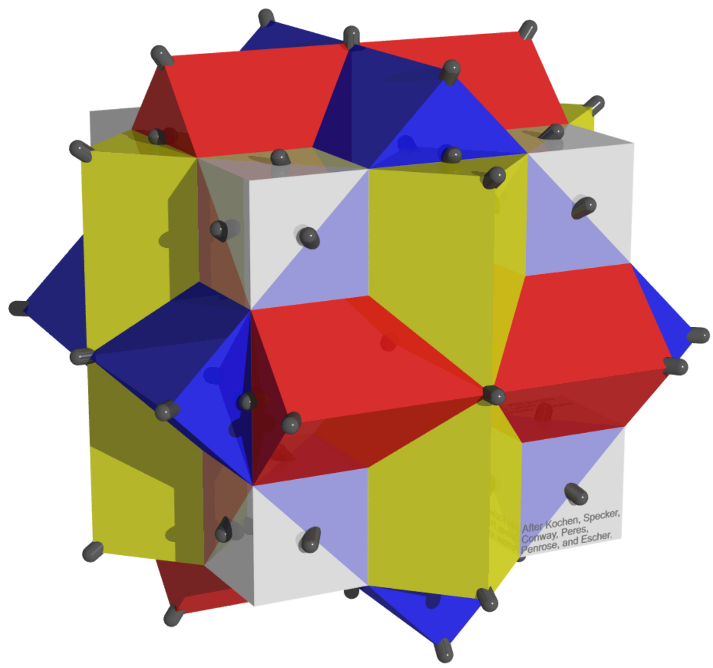


Figure 2.2: Geometrical visualisation taken from [136] of the 33 state vectors used in the proof of the Kochen-Specker theorem due to Peres [123].

Following this cornerstone result from 1967, several improvements on the proof of the theorem have been made, requiring less than the initially constructed 117 vectors. A geometrical proof due to Peres requires 33 vectors arranged along vertices and edges of the cubes in Fig. 2.2. It is not too hard to see that it is impossible to assign either 0 or 1 to every vector such that their sum over any set of orthogonal vectors equals 1. This immediately implies that no partial Boolean algebra homomorphism exists. The complexity of the argument can be further reduced by relaxing the minimal Hilbert space dimension to four. For instance, it is straightforward to see that the product constraints between spin-$\frac{1}{2}$ observables in the famous Mermin-Peres square (see Fig. 3.1 (a) in Ch. 3) lead to a similar obstruction.

As mentioned in the introduction, the algebraic content of quantum theory has largely survived unchanged. This holds even in quantum field theory, where infinite-dimensional algebras and superselection rules can be imposed on the algebraic level by von Neumann algebras (of type II and III) with non-trivial centre (cf. Sec. 2.2.1). It is thus interesting to consider generalisations of Thm. 18, which hold not only for von Neumann algebras of type I_n, $n \geq 3$, but for von Neumann algebras of arbitrary type. This has been done in [45].

Theorem 19. (Döring [45]) *Let $\mathcal{N}$ be a von Neumann algebra not only consisting of summands of type I_1, I_2, which represents the algebra of physical quantities of some quantum system. Then there exists no valuation function as in Def. 1. In particular, there exists no classical state space for the quantum system.*

The proof is based on Gleason's theorem and in this sense generalises a similar proof to Thm. 18 given by Bell in 1966 [18, 59].

In other words, there is no realist state space model for quantum theory, which assigns spectral values to all observables at once and preserves the functional relations between them. The power of Thm. 19 is that it relies on functional relations between simultaneously measurable observables only. It is the achievement of Bell, Kochen, and Specker to realise that the necessary algebraic constraints to be reflected in a classical, i.e., functional representation of quantum mechanical observables on some state space need only hold between commuting observables in order to show that such state spaces cannot exist.

Nonetheless, Isham, Butterfield, and Hamilton showed that this result can be given a geometric interpretation by introducing a generalised state space in the form of a certain presheaf over the partial order of contexts [23, 24, 75, 87]. While originally dealing with observables and order relations between them directly, later work takes a more algebraic viewpoint, which was further solidified in [1–4], and which is the one we will follow here.

Let $\mathcal{N}$ be a von Neumann algebra. As a consequence of the Borel functional calculus, two operators $a, b \in \mathcal{N}$ commute if and only if there exists another operator $c \in \mathcal{N}$ and Borel functions f, g such that $a = f(c)$, $b = g(c)$. Functional constraints are thus entirely encoded between commuting operators, i.e., in abelian subalgebras. Yet, they also relate noncommuting operators: given an observable $a \in \mathcal{N}$ there might be noncommuting observables $b, c \in \mathcal{N}$ and Borel functions h, k such that $a = h(b) = k(c)$.

The physical interpretation of the latter condition is the following. Note that Borel functions h, k are in general not injective and thus effectively 'wash out' some information. Under this notion of *coarse-graining*, operators can be related even if they are not simultaneously measurable. The algebraic constraints thus arise between commuting observables only, but are sensitive to the ways a given operator can arise as the coarse-graining of different, potentially non-commuting operators. This suggests the following structure: given a noncommutative algebra, consider all abelian subalgebras and order them by coarse-graining. Mathematically, this coarse-graining can be implemented by inclusion of abelian subalgebras resulting in the partial order of contexts $\mathcal{V}(\mathcal{N})$ as defined in Sec. 2.2.2.

Moreover, note that the partial algebra homomorphism property in Def. 1 imposes constraints on valuation functions for every inclusion relation in $\mathcal{V}(\mathcal{N})$. Since for any real Borel function f it holds that $\mathrm{sp}(f(a)) = \overline{f(\mathrm{sp}(a))}$ for all $a \in \mathcal{N}_{\mathrm{sa}}$, and $\mathrm{sp}(f(a)) = f(\mathrm{sp}(a))$ for all $a \in \mathcal{N}_{\mathrm{sa}}$ if f is also continuous, for von Neumann algebras this reads as follows: $v : \mathcal{N}_{\mathrm{sa}} \to \mathbb{R}$ is a valuation function if it satisfies the spectrum rule and the *functional composition principle*,

$$\forall f : V \to V \text{ continuous}, a \in \mathcal{N}_{\mathrm{sa}} : \quad v(f(a)) = f(v(a)) \ . \tag{2.9}$$

The partial algebra homomorphism condition in Def. 1 in general, and the functional composition principle in particular, thus suggest to map contexts to value assignments in a functorial way [23, 24, 75, 87]. First, note that in every context $V \in \mathcal{V}(\mathcal{N})$, a valuation function $\lambda : V \to \mathbb{C}$ is an algebra homomorphism, equivalently, a *character* or multiplicative/pure state, i.e., an element in the Gelfand spectrum Σ_V. We thus map every abelian subalgebra $V \in \mathcal{V}(\mathcal{N})$ to its Gelfand spectrum Σ_V. Second, the coarse-graining constraints between contexts $i_{\tilde{V}V} : \tilde{V} \hookrightarrow V$ correspond to non-injective, continuous functions f and can be imposed on elements in Gelfand spectra by restriction: for $\lambda \in \Sigma_V$, denote by $\lambda|_{\tilde{V}}$ the restriction of λ to the algebra $\tilde{V} \subset V$.

Definition 25. *Let $\mathcal{N}$ be a von Neumann algebra with context category $\mathcal{V}(\mathcal{N})$. The* spectral presheaf $\underline{\Sigma}(\mathcal{V}(\mathcal{N}))$ of $\mathcal{N}$ over $\mathcal{V}(\mathcal{N})$ *is the presheaf given* *(i) on objects: for all $V \in \mathcal{V}(\mathcal{N})$, let $\underline{\Sigma}_V := \Sigma_V$, the Gelfand spectrum of V,*

(ii) on arrows: for all $V, \tilde{V} \in \mathcal{V}(\mathcal{N})$, if $\tilde{V} \subseteq V$, let $\underline{\Sigma}(i_{\tilde{V}V}) : \underline{\Sigma}_V \longrightarrow \underline{\Sigma}_{\tilde{V}}$ with $\lambda \longmapsto \lambda|_{\tilde{V}}$.

If $\mathcal{N} = \mathcal{B}(\mathcal{H})$, we also write $\underline{\Sigma}(\mathcal{H})$ for $\underline{\Sigma}(\mathcal{V}(\mathcal{B}(\mathcal{H})))$.

The spectral presheaf was introduced in [87], it captures all constraints inherent to Def. 1, in particular, those inherent to Eq. (2.9). $\underline{\Sigma}(\mathcal{V}(\mathcal{N}))$ may thus be understood as a bookkeeping device: it encodes all algebraic constraints to potential value assignments in $\mathcal{V}(\mathcal{N})$. Importantly, note that a valuation function corresponds to a collection of characters $(\lambda_V)_{V \in \mathcal{V}(\mathcal{N})}$, i.e., a global section $\gamma \in \Gamma(\underline{\Sigma}(\mathcal{N}))$. The Kochen-Specker theorem is therefore equivalent to the following contextual reformulation [23, 24, 75, 87].

Theorem 20. (Kochen-Specker in contextual form) *Let $\mathcal{N}$ be a von Neumann algebra not only consisting of summands of type I_1, I_2. The spectral presheaf $\underline{\Sigma}(\mathcal{N})$ has no global sections.*

Note that Thm. 20 holds for all von Neumann algebras (not only consisting of summands of type I_1, I_2) and thus incorporates the generalisation of the original Kochen-Specker theorem in Thm. 19.

Thm. 20 is at first only a reformulation of Thm. 19, yet it adds a previously hidden geometrical perspective. While value assignments do exist *locally*, i.e., in every context, the algebraic constraints between simultaneously measurable observables, encoded in the coarse-graining maps, obstruct the existence of such an assignment *globally*, i.e., over all contexts of $\mathcal{V}(\mathcal{N})$. While the elements in $\underline{\Sigma}_V$ are points in a compact Hausdorff space and thus naturally give rise to the structure of a classical state space (cf. Thm. 9), there are no generalised points of this type in the quantum case, formally, $\Gamma(\underline{\Sigma}(\mathcal{N})) = \emptyset$.

The Kochen-Specker theorem is thus an example of a *local-to-global*-type obstruction to the existence of global sections of a corresponding presheaf (noncommutative space), in this case the spectral presheaf. We will see that many more theorems in the foundations of quantum theory attain similar reformulations for suitable notions of presheaves over the context order $\mathcal{V}(\mathcal{N})$.

2.3.2 Gleason's theorem

Gleason's theorem is another landmark result in quantum theory. It justifies the Born rule, which originally had the status of an axiom, from purely mathematical considerations.

Theorem 21. (Gleason [66]) *Let $\mathcal{H}$ be a Hilbert space and* $\dim(\mathcal{H}) \geq 3$ *finite. Then every probability measure* $\mu : \mathcal{P}(\mathcal{H}) \to [0,1]$ *over the projections on* $\mathcal{H}$ *corresponds to a density matrix* $\rho_\mu : \mathcal{H} \to \mathcal{H}$, $\rho_\mu \geq 0$, $\mathrm{tr}(\rho_\mu) = 1$ *such that* $\mu(p) = \mathrm{tr}(\rho_\mu p)$ *for all* $p \in \mathcal{P}(\mathcal{H})$.

Here, a *(finitely additive) probability measure* is a map $\mu : \mathcal{P}(\mathcal{H}) \to [0,1]$ such that $\mu(p+q) = \mu(p) + \mu(q)$ whenever $p, q \in \mathcal{P}(\mathcal{H})$, $pq = 0$ and $\mu(1) = 1$. While the original argument was for type I_n factors only, the validity of Thm. 21 was extended to arbitrary von Neumann algebras in [34, 151] (cf. [109]). In this setting, μ is further called completely (countably) additive if $\mu(\sum_{i\in I} p_i) = \sum_{i\in I} \mu(p_i)$ for every (countable) family of orthogonal projections $(p_i)_{i\in I}$.

Theorem 22. (Gleason-Christensen-Yaedon [34, 151]) *Let $\mathcal{N}$ be a von Neumann algebra with no summand of type I_2 and let* $\mu : \mathcal{P}(\mathcal{N}) \to \mathbb{R}$ *be a finitely additive probability measure on the projections of $\mathcal{N}$. There exists a unique state* $\sigma_\mu \in \mathcal{S}(\mathcal{N})$ *such that* $\mu(p) = \sigma_\mu(p)$ *for all* $p \in \mathcal{P}(\mathcal{N})$.

If μ is also completely additive then σ_μ is normal and of the form $\mu(p) = \sigma_\mu(p) = \mathrm{tr}(\rho_\mu p)$ *for all* $p \in \mathcal{P}(\mathcal{N})$ *and ρ_μ a positive trace-class operator with* $\mathrm{tr}(\rho_\mu) = 1$.

Recall that trace-class operators generalise density matrices to infinite dimensions. In finite dimensions every state is of this form, however, in infinite dimensions only normal states correspond to trace-class operators (cf. Sec. 2.2.1). Normal states satisfy $\sigma(\bigvee_{i\in I} p_i) = \sup_{i\in I}\sigma(p_i)$ for all families of pairwise orthogonal projections $(p_i)_{i\in I}$ and are thus easily seen to correspond to completely additive probability measures, i.e., $\mu(\bigvee_{i\in I} p_i) = \sum_{i\in I} \mu(p_i)$. Succinctly, every finitely additive measure bijectively corresponds to a state on $\mathcal{N}$ and every completely additive measure bijectively corresponds to a normal state on $\mathcal{N}$.[17]

For later reference we mention a further generalisation of Thm. 22, which is concerned with the codomain of probability measures on $\mathcal{P}(\mathcal{N})$. Rather than restricting measures to be real-valued, a Gleason-type theorem holds even for Banach space-valued measures [89].

[17] Clearly, for $\mathcal{N} \subseteq \mathcal{B}(\mathcal{H})$ with $\mathcal{H}$ separable, countable additivity is sufficient.

Let $\mathcal{P}(\mathcal{N})$ be the projection lattice of a von Neumann algebra $\mathcal{N}$, X a Banach space, and $\mu : \mathcal{P}(\mathcal{N}) \to X$ a map such that (i) $\mu(p+q) = \mu(p) + \mu(q)$ whenever $p, q \in \mathcal{P}(\mathcal{N})$, $pq = 0$ and (ii) $\sup\{||\mu(p)|| : p \in \mathcal{P}(\mathcal{N})\} < \infty$. Then μ is said to be a *finitely additive, X-valued measure on $\mathcal{P}(\mathcal{N})$*. Clearly, each bounded linear operator from $\mathcal{N}$ to X restricts to a finitely additive X-valued measure, conversely:

Theorem 23. (Mackey-Gleason-Bunce-Wright [89]) *Let $\mathcal{N}$ be a von Neumann algebra with no direct summand of type I_2. Then for each Banach space X, each X-valued measure $\mu : \mathcal{P}(\mathcal{N}) \to X$ has a unique extension to a bounded linear operator $\phi : \mathcal{N} \to X$.*

Note that when $\mathcal{N} = M_2(\mathbb{C})$ and X is one-dimensional, there exist examples of measures that fail to extend to linear functionals.

Thm. 21 and its generalisations, Thm. 22 and Thm. 23, are remarkable for the following reason: finite (complete) additivity imposes constraints between commuting projections only. On the other hand, by the non-contextual assignment of probabilities to projections $p \in \mathcal{P}(\mathcal{N})$, i.e., independent of which context p lies in, these constraints extend beyond commuting projections. Similar to the Kochen-Specker theorem, the constraints thus arise solely within contexts and via the inclusion relations between them. Yet, in contrast every context is no longer assigned its Gelfand spectrum but instead the set of finitely (completely) additive probability measures over it. Moreover, coarse-graining between probability distributions in different contexts is naturally encoded by marginalisation. In close analogy with [49], we make the following definition.

Definition 26. *Let $\mathcal{N}$ be a von Neumann algebra with context category $\mathcal{V}(\mathcal{N})$. The* (normal) probabilistic presheaf $\underline{\Pi}$ of $\mathcal{N}$ over $\mathcal{V}(\mathcal{N})$ *is the presheaf given*

(i) on objects: for all $V \in \mathcal{V}(\mathcal{N})$, let

$$\underline{\Pi}_V := \{\mu_V : \mathcal{P}(V) \to [0,1] \mid \mu_V \textit{ is a finitely (completely) additive probability measure}\} \ ,$$

(ii) on arrows: for all $V, \tilde{V} \in \mathcal{V}(\mathcal{N})$, if $\tilde{V} \subseteq V$, let $\underline{\Pi}(i_{\tilde{V}V}) : \underline{\Pi}_V \longrightarrow \underline{\Pi}_{\tilde{V}}$ with $\mu_V \longmapsto \mu_V|_{\tilde{V}}$.

Here, $\mu_V|_{\tilde{V}}$ denotes the marginalisation map, which sends $\mu_V : \mathcal{P}(V) \to [0,1]$ to $\mu_{\tilde{V}} : \mathcal{P}(\tilde{V}) \to [0,1]$ for $\tilde{V} \subseteq V$. $\underline{\Pi}$ can be seen as a generalisation of $\underline{\Sigma}$, since $\underline{\Pi}_V$ contains all convex linear combinations of elements in $\underline{\Sigma}_V$ in every context $V \in \mathcal{V}(\mathcal{N})$, and the marginalisation maps coincide with restriction in $\underline{\Sigma}$ between pure states.

In this reading, a finitely (completely) additive probability measure over the projections of $\mathcal{N}$ is a collection of finitely (completely) additive probability measures over contexts $(\mu_V)_{V\in\mathcal{V}(\mathcal{N})}$, i.e., a global section of the (normal) probabilistic presheaf $\underline{\Pi}$. Since every (normal) quantum state, i.e., every positive linear functional of norm one, defines a finitely (completely) additive probability measure over the projections of $\mathcal{N}$, it corresponds with a global section of $\underline{\Pi}$. However, it is not obvious that all global sections are of this form. Yet, by Gleason's theorem the obstructions in $\mathcal{V}(\mathcal{N})$ are such that the linear functionals in contexts uniquely extend to linear functionals on all of $\mathcal{N}$. This yields a contextual reformulation of Thm. 22 as was first noted in [45], and presented in similar form to here in [43, 48].

Theorem 24. (Gleason in contextual form (I)) *Let $\mathcal{N}$ be a von Neumann algebra with no summand of type I_2. There is a bijective correspondence between (normal) quantum states, i.e., states on $\mathcal{N}$, and global sections of the (normal) probabilistic presheaf $\underline{\Pi}$ of $\mathcal{N}$ over $\mathcal{V}(\mathcal{N})$.*

There is an obvious generalisation of Def. 26 to Banach space-valued measures in Thm. 23. What is more, the latter allows for a refinement of Gleason's theorem in contextual form. In fact, Gleason's theorem proves something slightly stronger than what is captured by Thm. 24: not only does every global section of the probabilistic presheaf (cf. Def. 26) assign a measure to every context, but it relates measures across contexts in a particular way. To see this, consider that we assign a 'probability' to every element $p \in \mathcal{P}(\mathcal{N})$ with respect to *every* projection $q \in \mathcal{P}(V_0)$ for (at least) one (maximal) reference context V_0. More precisely, we require the measures μ_V to decompose into measures over the elements in $\mathcal{P}(V_0)$ themselves. Mathematically, this means that we assign a set of measures $\mu_{V,q}$ in every context $V \in \mathcal{V}(\mathcal{N})$, labelled by the projections in the reference context $q \in \mathcal{P}(V_0)$, and additive for orthogonal projections $q, q' \in \mathcal{P}(V_0)$, $qq' = 0$. Comparing with Def. 26, probability measures in contexts then decompose as follows:

$$\mu_V = \sum_{\substack{q_i \in \mathcal{P}(V_0) \\ q_i q_j = 0, \bigvee_i q_i = 1}} \mu_{V,q_i}, \quad \forall V \in \mathcal{V}(\mathcal{N})$$

We can impose this constraint in a somewhat suggestive way by writing $\mu_{V,q_i} = v_i^* \varphi v_i$ for all $q_i \in \mathcal{P}(V_0)$ and thus $\mu_V = \sum_i v_i^* \varphi v_i = v^* \varphi v$, where $v \in \mathcal{K}$ for some appropriate Hilbert space $\mathcal{K}$ and $\varphi : \mathcal{P}(V) \hookrightarrow \mathcal{P}(\mathcal{K})$ an embedding, in particular, $\varphi(0) = 0$, $\varphi(1-p) = 1 - \varphi(p)$, and $\varphi(p + p') = \varphi(p) + \varphi(p')$ for all $p, p' \in \mathcal{P}(V)$, $pp' = 0$. Crucially, the condition that φ preserves orthogonality can be encoded locally by requiring additivity for all $\mu_{V,q}$, $q \in \mathcal{P}(V_0)$. We thus define the dilated probabilistic presheaf as follows.

Definition 27. *Let $\mathcal{N}$ be a von Neumann algebra with context category $\mathcal{V}(\mathcal{N})$. The* (normal) dilated probabilistic presheaf $\underline{\Pi}$ *of $\mathcal{N}$ over $\mathcal{V}(\mathcal{N})$*[18] *is the presheaf given*

(i) on objects: for all $V \in \mathcal{V}(\mathcal{N})$, let

$$\underline{\Pi}_V := \{\mu_V = v^* \varphi_V v \mid v \in \mathcal{K},^{19} \ \varphi_V : \mathcal{P}(V) \hookrightarrow \mathcal{P}(\mathcal{K}) \ \textit{(normal), and } \mu_V(1) = 1\} \ ,$$

(ii) on arrows: for all $V, \tilde{V} \in \mathcal{V}(\mathcal{N})$, if $\tilde{V} \subseteq V$, let

$$\underline{\Pi}(i_{\tilde{V}V}) : \underline{\Pi}_V \longrightarrow \underline{\Pi}_{\tilde{V}} \ \textit{with} \ \varphi_V \longmapsto \varphi_V|_{\tilde{V}} \ .$$

Theorem 25. (Gleason in contextual form (II)) *Let $\mathcal{N}$ be a von Neumann algebra with no summand of type I_2. There is a bijective correspondence between (normal) quantum states, i.e., states on $\mathcal{N}$, and global sections of the (normal) dilated probabilistic presheaf $\underline{\Pi}$ of $\mathcal{N}$ over $\mathcal{V}(\mathcal{N})$.*

Proof. It is easy to see that every (normal) state $\sigma \in \mathcal{S}(\mathcal{N})$ defines a finitely (completely) additive global section $\gamma_\sigma \in \Gamma(\underline{\Pi}(\mathcal{V}(\mathcal{N})))$ via its purification. Conversely, note that $\gamma = (\mu_V = v^* \varphi_V v)_{V \in \mathcal{V}(\mathcal{N})} \in \Gamma(\underline{\Pi}(\mathcal{V}(\mathcal{N})))$ defines a finitely (completely) additive map $\varphi : \mathcal{P}(\mathcal{N}) \to \mathcal{B}(\mathcal{K})$. By Thm. 23 φ uniquely extends to a bounded linear map $\Phi : \mathcal{N} \to \mathcal{B}(\mathcal{K})$ such that $\Phi|_{\mathcal{P}(\mathcal{N})} = \varphi$. In particular, we thus obtain a (normal) state $\sigma_\gamma = v^* \Phi v \in \mathcal{S}(\mathcal{N})$ for some $v \in \mathcal{K}$. □

[18] In a slight abuse of notation we will use the same notation $\underline{\Pi}(\mathcal{V}(\mathcal{N}))$ for the probabilistic and the dilated probabilistic presheaf of $\mathcal{N}$ over $\mathcal{V}(\mathcal{N})$.

[19] Note that we can choose $\mathcal{K}$ independently of contexts, since every measure admits a dilation for $\dim(\mathcal{K}) \geq \dim(\mathcal{N})$ (cf. Thm. 31). Importantly, this does not imply linearity, the latter requires Thm. 23 (see also [88]).

Note that, in general, the dilated probabilistic presheaf encodes more constraints than the probabilistic presheaf, since additivity not only holds with respect to the *overall* probability measure μ_V in Def. 26, but also with respect to the *individual* measures $\mu_{V,q_i} = v_i^* \varphi v_i$ (cf. Def. 27). While this is not immediately obvious for local systems, for which, by Thm. 24 and Thm. 25, probabilistic and dilated probabilistic presheaf have isomorphic global sections,[20] this correspondence breaks down for composite systems (cf. Sec. 2.3.4 and Sec. 2.4.4).

Thm. 24 (Thm. 25) is a reformulation of Gleason's theorem similar to the reformulation of the Kochen-Specker theorem by Isham and Butterfield in [87]. Its significance lies in the way it orders the components in the proof of Thm. 21. $\underline{\Pi}$ is a functor over the context category $\mathcal{V}(\mathcal{N})$ similar to $\underline{\Sigma}$, both theorems thus relate to (physical) contextuality in the same way, they classify global sections of their respective presheaves over the partial order of abelian subalgebras in $\mathcal{V}(\mathcal{N})$. Accordingly, both theorems answer a local-to-global problem. Thm. 20 asserts that despite the existence of value assignments locally, no such assignments are possible globally. The obstructions of $\mathcal{V}(\mathcal{N})$ on $\underline{\Sigma}$ are too restrictive. In contrast, for the probabilistic presheaf such global assignments do exist and correspond with quantum states exactly.

2.3.3 Wigner's theorem

One of the earliest cornerstone theorems in quantum theory is Wigner's theorem, it classifies the symmetries of a quantum system.

Theorem 26. (Wigner [147]) *Let $\mathcal{H}$ be a Hilbert space, $\dim(\mathcal{H}) \geq 2$, and let $\mathcal{P}_1(\mathcal{H})$ be the set of rank-1 projections on $\mathcal{H}$ (equivalently, $\mathcal{P}_1(\mathcal{H})$ is the projective Hilbert space). Every bijective map*

$$\varphi : \mathcal{P}_1(\mathcal{H}) \longrightarrow \mathcal{P}_1(\mathcal{H}), \quad p \longmapsto \varphi(p)$$

such that $\mathrm{tr}(\varphi(p), \varphi(q)) = \mathrm{tr}(p, q)$ for all $p, q \in \mathcal{P}_1(\mathcal{H})$ (i.e., transition probabilities are preserved) is implemented by conjugation with a unitary or anti-unitary operator u,

$$\forall p \in \mathcal{P}_1(\mathcal{H}) : \varphi(p) = upu^* \ .$$

[20] Note that we identify global sections $\gamma, \gamma' \in \Gamma(\underline{\Pi}(\mathcal{V}(\mathcal{N})))$ whenever $\mu_V^\gamma = \mu_V^{\gamma'}$ for all $V \in \mathcal{V}(\mathcal{N})$.

The first step in reformulating Thm. 26 in terms of contexts is to realise that the automorphism group $\mathrm{Aut}(\mathcal{P}_1(\mathcal{H}))$ is closely related to the automorphism group of $\mathcal{P}(\mathcal{H})$. The former encodes symmetries of transition probabilities as in Thm. 26, the latter is defined as follows.

Definition 28. *Let $\mathcal{N}$ be a von Neumann algebra. The automorphism group $\mathrm{Aut}(\mathcal{P}(\mathcal{N}))$ consists of all bijective maps $\varphi : \mathcal{P}(\mathcal{N}) \to \mathcal{P}(\mathcal{N})$ satisfying*

(i) $\forall p, q \in \mathcal{P}(\mathcal{N}) : \ (p \leq q) \Leftrightarrow (\varphi(p) \leq \varphi(q))$

(ii) $\forall p \in \mathcal{P}(\mathcal{N}) : \ \varphi(1-p) = 1 - \varphi(p)$

Different to most key theorems discussed in this dissertation, Wigner's theorem holds already in two dimensions instead of three. Nevertheless, the following equivalence once again requires three dimensions [29].

Theorem 27. (Cassinelli [29]) *Let $\mathcal{N}$ be a von Neumann algebra with no summand of type I_2. Then $\mathrm{Aut}(\mathcal{P}_1(\mathcal{N})) \simeq \mathrm{Aut}(\mathcal{P}(\mathcal{N}))$.*

Furthermore, by Thm. 12 automorphisms on the projection lattice $\mathcal{P}(\mathcal{N})$ correspond with order automorphisms on the context category $\mathcal{V}(\mathcal{N})$. The latter are classified in terms of Jordan $*$-automorphisms on $\mathcal{N}$ by Dye's theorem in contextual form, Thm. 13. For convenience, we restate it here.

Theorem 13. (Döring-Harding [50]) *Let $\mathcal{N}$ be a von Neumann algebra not isomorphic to $\mathbb{C} \oplus \mathbb{C}$ or $M_2(\mathbb{C})$. For every order automorphism $\varphi : \mathcal{V}(\mathcal{N}) \to \mathcal{V}(\mathcal{N})$, there is a unique Jordan $*$-automorphism $\Phi : (\mathcal{N}, \cdot) \to (\mathcal{N}, \cdot)$ such that $\varphi(V) = \Phi[V]$ for all $V \in \mathcal{V}(\mathcal{N})$.*

In order to relate this back to Wigner's original theorem, we need two more theorems on Jordan $*$-homomorphisms. The first is the following standard result in [95].

Theorem 28. *Let $\mathcal{N}_1, \mathcal{N}_2$ be von Neumann algebras and let $\Phi : \mathcal{N}_1 \to \mathcal{N}_2$ be a Jordan $*$-isomorphism. Then there exists a central projection $p \in \mathcal{Z}(\mathcal{N})$ such that Φ acts as a $*$-isomorphism on $p\mathcal{N}_1 p$ and as a $*$-anti-isomorphism on $(1-p)\mathcal{N}_1(1-p)$.*

In words, a Jordan $*$-isomorphism acts on every factor either as a $*$-isomorphism or $*$-anti-isomorphism. Clearly, the same applies to Jordan $*$-automorphisms. Finally, for factors $\mathcal{N} = \mathcal{B}(\mathcal{H})$ these correspond with unitaries and anti-unitaries by the following result (cf. [8]).

Proposition 3. *Every $*$-automorphism $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ is implemented by conjugation with a unitary operator, and every $*$-anti-automorphism is implemented by conjugation with an anti-unitary operator.*

Putting the pieces together, we obtain the following reformulation of Wigner's theorem [47].

Theorem 29. (Wigner in contextual form.) *Let $\mathcal{H}$ be a Hilbert space and $\dim(\mathcal{H}) \geq 3$. Every order automorphism $\varphi : \mathcal{V}(\mathcal{H}) \to \mathcal{V}(\mathcal{H})$ is implemented by conjugation with either a unitary or anti-unitary operator u,*

$$\forall V \in \mathcal{V}(\mathcal{H}) : \ \varphi(V) = uVu^* \ .$$

Proof. By Thm. 13 φ uniquely extends to a Jordan $*$-automorphism and with Thm. 28 decomposes into a sum of $*$-automorphisms and $*$-anti-automorphism over factors, equivalently, central projections. Since $\mathcal{N} = \mathcal{B}(\mathcal{H})$ is a factor, the only central projections are $0, 1$. Hence, $\mathcal{N} = 1\mathcal{N}1 \oplus 0\mathcal{N}0$ and φ acts non-trivially only on the first summand and either as a $*$-automorphism or $*$-anti-automorphism. By Prop. 3 the former corresponds to conjugation by a unitary, the latter to conjugation by an anti-unitary operator. □

Wigner's theorem, thus arises as a special case of Dye's theorem, namely for factors $\mathcal{N} = \mathcal{B}(\mathcal{H})$ of type I_n with $3 \leq n = \dim(\mathcal{H})$. Similarly, Wigner's theorem in contextual form, is a special case of Dye's theorem in contextual form, which lifts symmetries on the partial order of contexts to Jordan $*$-isomorphisms for arbitrary von Neumann algebras.

The contextual reformulation of Wigner's theorem is given in terms of symmetries of the partial order of contexts. In particular, we have not defined a corresponding presheaf as for the Kochen-Specker theorem and Gleason's theorem. Nevertheless, it is possible to define a presheaf, which encodes the partial of contexts and nothing more. In terms of this presheaf, Wigner's theorem relates automorphisms on this presheaf with conjugation by unitary or anti-unitary operators (for details, see [47]). Importantly, Wigner's theorem in contextual (presheaf) form places both unitary and anti-unitary operators on the same footing. In fact, we will find that both need to be considered when we study states on composite systems in Sec. 2.4.4.

2.3.4 Stinespring's theorem

Stinespring's theorem constitutes another cornerstone in mathematical quantum theory, it classifies completely positive maps. Recall that a linear map $\phi : \mathcal{A} \rightarrow \mathcal{B}(\mathcal{H})$ with $\mathcal{A}$ a C^*-algebra is called *n-positive* if $\mathrm{id} \otimes \phi : \mathcal{A}^{(n)} \rightarrow \mathcal{B}(\mathcal{H})^{(n)}$ is positive as an operator from the C^*-algebra of $n \times n$-matrices with entries in $\mathcal{A}$, $\mathcal{A}^{(n)} := M_n(\mathcal{A})$, into $\mathcal{B}(\mathcal{H})^{(n)} := M_n(\mathcal{B}(\mathcal{H}))$. ϕ is called *completely positive* if it is n-positive for all $n \in \mathbb{N}$. Clearly, every completely positive map is positive, however, a positive map is generally not completely positive. Completely positive maps play an important role in the study of quantum channels. The latter are defined as trace-preserving, completely positive maps.

Theorem 30. (Stinespring [135]) *Let $\mathcal{A}$ be a unital C^*-algebra, $\mathcal{H}$ a Hilbert space, and $\phi : \mathcal{A} \rightarrow \mathcal{B}(\mathcal{H})$ a linear function. Then a necessary and sufficient condition that ϕ have the form*

$$\forall a \in \mathcal{A} : \;\; \phi(a) = v^* \Phi(a) v \; ,$$

where $v : \mathcal{H} \rightarrow \mathcal{K}$ is a bounded linear transformation from $\mathcal{H}$ to a Hilbert space $\mathcal{K}$ and Φ is a $*$-representation of $\mathcal{A}$ into operators on $\mathcal{K}$, is that ϕ be completely positive.

Stinespring proved his theorem as a noncommutative generalisation of Naimark's dilation theorem. The latter classifies positive operator-valued measures (cf. Def. 10).

Theorem 31. (Naimark [114]) *Let ϱ be a positive $\mathcal{B}(\mathcal{H})$-valued measure on a compact Hausdorff space X. There exists a Hilbert space $\mathcal{K}$, a bounded operator $v : \mathcal{H} \rightarrow \mathcal{K}$, and a self-adjoint, spectral $\mathcal{P}(\mathcal{K})$-valued measure on X, φ, such that*

$$\forall B \in \sigma(X) : \;\; \varrho(B) = v^* \varphi(B) v \; .$$

Being a noncommutative generalisation of Thm. 31, the contextual character of Stinespring's theorem is somewhat implicit. In order to extract this contextual aspect explicitly, note that similar to Def. 18 of $\mathcal{V}(\mathcal{N})$, any noncommutative, unital C^*-algebra $\mathcal{A}$ can be decomposed into its abelian subalgebras ordered by inclusion, denoted analogously by $\mathcal{V}(\mathcal{A})$. (Complete) Positivity of $\phi : \mathcal{A} \rightarrow \mathcal{B}(\mathcal{H})$ implies, in particular, that the maps $\phi|_V : V \rightarrow \mathcal{B}(\mathcal{H})$ are positive in every context $V \in \mathcal{V}(\mathcal{A})$. Remarkably, for von Neumann algebras the other direction holds true by Gleason's theorem as we show in Thm. 34 below.

The proof consists of two parts. For the first part, note that $\varrho_V := \phi|_{\mathcal{P}(V)}$ defines a positive operator-valued measure in every context $V \in \mathcal{V}(\mathcal{N})$. In particular, by Stone duality, Thm. 8, every context $V \in \mathcal{V}(\mathcal{N})$ corresponds to an extremely disconnected compact Hausdorff space. More precisely, $\mathcal{P}(V)$ is a complete Boolean algebra and there is a lattice isomorphism $\alpha_V : \mathcal{P}(V) \to \mathrm{Cl}(\Sigma_V)$ into the (lattice of) clopen subsets of the Gelfand spectrum Σ_V defined by $\alpha_V(p) = \{\lambda \in \Sigma_V \mid \lambda(p) = 1\}$. Since the Borel σ-algebra on Σ_V has elements the clopen subsets $\mathrm{Cl}(\Sigma_V)$ and for any finite, disjoint family of Borel (clopen) subsets of Σ_V, B_i, it follows that $\alpha_V^{-1}(B_i) \in \mathcal{P}(\mathcal{H})$ is a family of orthogonal projections, it is easily seen that $\varrho_V \circ \alpha_V^{-1}(\cup_i B_i) = \phi(\sum_i p_i) = \sum_i \phi(p_i) = \sum_i \varrho_V \circ \alpha_V^{-1}(B_i)$.[21] Hence, the map $\varrho_V \circ \alpha_V^{-1}$ defines a positive operator-valued measure on Σ_V, equivalently, since α_V is an isomorphism, ϱ_V defines a positive operator-valued measure on $\mathcal{P}(V)$. We can thus apply Naimark's dilation theorem, Thm. 31, in every abelian subalgebra $V \in \mathcal{V}(\mathcal{N})$ to obtain Hilbert spaces $\mathcal{K}_V$, projection-valued measures $\varphi_V : \mathcal{P}(\mathcal{N}) \to \mathcal{P}(\mathcal{K}_V)$, and bounded operators $v_V : \mathcal{H} \to \mathcal{K}_V$ such that $\varrho_V = v_V^* \varphi_V v_V$.

Now note that every commutative von Neumann algebra $V \in \mathcal{V}(\mathcal{N})$ is isomorphic to the algebra $L^\infty(X, \mu)$ of bounded measurable functions on some (standard) measure space (X, μ) acting on the Hilbert space of square-integrable functions $L^2(X, \mu)$ by multiplication (cf. Thm. 35 below).[22] The positive operator-valued measures $\varrho_V : \mathcal{P}(V) \to \mathcal{B}(\mathcal{H})$ therefore give rise to the linear map ϕ_V by setting $\phi_V(f) = \int_{\Sigma_V} f(\lambda) d\varrho_V(\lambda)$ for all $f \in L^\infty(\Sigma_V)$ (cf. Thm. 4). Naimark's theorem, which lifts the positive operator-valued measures $\varrho_V = v_V^* \varphi_V v_V$ to spectral measures φ_V on $\mathcal{K}_V$ in every context, thus also lifts the positive linear maps $\phi_V = v_V^* \Phi_V v_V$ to C^*-algebra homomorphisms $\Phi_V(f) = \int_{\Sigma_V} f(\lambda) d\varphi_V(\lambda)$. It follows that Naimark's theorem is a special case of Stinespring's theorem for $\mathcal{A}$ abelian. Surprisingly, Stinespring's theorem also arises from Naimark's theorem when applied to the entire context category, i.e., by choosing the dilations φ_V in Thm. 31 consistently *across contexts*. To see this, we need another presheaf.

[21] The Borel σ-algebra on Σ_V arises from any standard measure space (X, μ) with $L^\infty(X, \mu) \simeq V$ by removing all sets of measure zero.

[22] Note that (X, μ) is not unique, yet modulo sets of measure zero we recover the duality in Thm. 10 (see also [120]). We sometimes use the abstract notation $L^\infty(\Sigma_V)$ to indicate any such measure space.

Definition 29. *Let $\mathcal{N}$ be a von Neumann algebra with no summand of type I_2 and $\mathcal{H}$ a Hilbert space. The* POVM presheaf $\underline{\Pi}^{\mathcal{H}}$ of $\mathcal{N}$ over $\mathcal{V}(\mathcal{N})$ *is the presheaf given*

(i) on objects: for all $V \in \mathcal{V}(\mathcal{N})$, let

$$\underline{\Pi}^{\mathcal{H}}_V := \{\varrho_V = v^*\varphi_V v \mid v : \mathcal{H} \to \mathcal{K},\text{[23]}\ \varphi_V : \mathcal{P}(V) \hookrightarrow \mathcal{P}(\mathcal{K})\}\ ,\text{[24]}$$

(ii) on arrows: for all $V, \tilde{V} \in \mathcal{V}(\mathcal{N})$, if $\tilde{V} \subseteq V$, let

$$\underline{\Pi}(i_{\tilde{V}V}) : \underline{\Pi}_V \longrightarrow \underline{\Pi}_{\tilde{V}} \text{ with } \varphi_V \longmapsto \varphi_V|_{\tilde{V}}\ .$$

The *POVM presheaf* is a natural generalisation of the dilated probabilistic presheaf. In fact, it is itself a special case of an even more general 'measure presheaf' for arbitrary Banach space-valued measures corresponding to the generalised version of Gleason's theorem, Thm. 23.

We would like to show that global sections of the POVM presheaf correspond with completely positive maps. This is almost the case. Instead, we obtain *decomposable* maps, i.e., maps of the form $\phi = v^*\Phi v$ with $\Phi : \mathcal{N} \to \mathcal{B}(\mathcal{K})$ a Jordan $*$-homomorphism. Similar to completely positive maps, which are characterised by Stinespring's theorem, Thm. 30, decomposable maps are characterised by a symmetrised positivity condition [137].

Theorem 32. (Størmer [137]) *Let $\mathcal{A}$ be a C^*-algebra and $\Phi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ linear. Then Φ is decomposable if and only if for all $n \in \mathbb{N}$, $(a_{ij}), (a_{ji}) \in M_n(\mathcal{A})_+$ implies $\Phi(a_{ij}) \in M_n(\mathcal{B}(\mathcal{H}))_+$.*

The next theorem proves that collections of spectral measures in contexts uniquely extend to decomposable maps.

Theorem 33. *Let $\mathcal{N}$ be a von Neumann algebra with no summand of type I_2, $\mathcal{H}$ a Hilbert space, and $\underline{\Pi}^{\mathcal{H}}(\mathcal{V}(\mathcal{N}))$ the corresponding POVM presheaf of $\mathcal{N}$ over $\mathcal{V}(\mathcal{N})$. There is a bijective correspondence between global sections $\Gamma(\underline{\Pi}^{\mathcal{H}}(\mathcal{V}(\mathcal{N})))$ and decomposable maps $\phi : \mathcal{N} \to \mathcal{B}(\mathcal{H})$.*

[23] Similarly to Def. 27, we may choose $\mathcal{K}$ independently of contexts, since by Thm. 31 a dilation exists e.g. for $\dim(\mathcal{K}) \geq \dim(\mathcal{N})\dim(\mathcal{H})$.

[24] Alternatively, by the discussion preceding Def. 29 one may think of φ_V as a spectral measure on Σ_V.

Proof. Clearly, every decomposable map $\phi : \mathcal{N} \to \mathcal{B}(\mathcal{H})$, $\phi = v^*\Phi v$ for $v : \mathcal{H} \to \mathcal{K}$ linear and $\Phi : \mathcal{N} \to \mathcal{B}(\mathcal{K})$ a Jordan $*$-homomorphism defines a global section $\gamma_\phi \in \Gamma(\underline{\Pi}^{\mathcal{H}}(\mathcal{V}(\mathcal{N})))$ since $\Phi|_{\mathcal{P}(\mathcal{N})} : \mathcal{P}(\mathcal{N}) \to \mathcal{P}(\mathcal{K})$ is an orthomorphism (a spectral measure in every context $V \in \mathcal{V}(\mathcal{N})$).

Conversely, a global section $\gamma = (\varrho_V = v^*\varphi_V v)_{V \in VN} \in \Gamma(\underline{\Pi}^{\mathcal{H}}(\mathcal{V}(\mathcal{N})))$ defines an orthomorphism $\varphi : \mathcal{P}(\mathcal{N}) \to \mathcal{P}(\mathcal{K})$, i.e.,

$$\forall p, q \in \mathcal{P}(\mathcal{N}), pq = 0 : \quad \varphi(p)\varphi(q) = 0 \quad \text{and} \quad \varphi(p+q) = \varphi(p) + \varphi(q) \; .$$

This follows since $\varphi = (\varphi_V)_{V \in \mathcal{V}(\mathcal{N})}$ defines a family of embeddings (spectral measures), in particular, there is an embedding φ_V with $p, q \in \mathcal{P}(V)$. By Thm. 23 φ uniquely extends to a bounded linear operator $\Phi : \mathcal{N} \to \mathcal{B}(\mathcal{K})$.

We therefore not only find that ϱ extends to a bounded linear operator $\phi_\gamma : \mathcal{N} \to \mathcal{B}(\mathcal{H})$ by Thm. 23, but we also obtain a *globally* defined dilation $(\phi_\gamma)|_{\mathcal{P}(\mathcal{N})} = v^*\varphi v$ with $\varphi : \mathcal{P}(\mathcal{N}) \to \mathcal{P}(\mathcal{K})$ an orthomorphism. The additional property that $\Phi|_{\mathcal{P}(\mathcal{N})} = \varphi$ is an orthomorphism further implies that Φ defines a Jordan $*$-homomorphism (see also [88]). To see this, it is enough to show that Φ also preserves squares, i.e., for every $a \in \mathcal{N}$ with spectral decomposition $a = \sum_i a_i p_i$,

$$\Phi(a^2) = \Phi(\sum_i a_i^2 p_i) = \sum_i a_i^2 \Phi(p_i) = \Phi(a)^2 \; .$$

Since $\{a, b\} = \frac{1}{2}(ab + ba) = \frac{1}{4}[(a+b)^2 - a^2 - b^2]$, this implies

$$\Phi(\{a, b\}) = \Phi(\frac{1}{4}[(a+b)^2 - a^2 - b^2]) = \frac{1}{4}[(\Phi(a) + \Phi(b))^2 - \Phi(a)^2 - \Phi(b)^2] = \{\Phi(a), \Phi(b)\} \; .$$

Finally, $\Phi(a^*) = \Phi(\sum_i \overline{a_i} p_i) = \sum_i \overline{a_i}\Phi(p_i) = \Phi(a)^*$. Hence, Φ in $\rho_\gamma = v^*\Phi v$ is a Jordan $*$-homomorphism. This completes the proof. □

In order to meet the assumptions of Thm. 23 we restricted to von Neumann algebras. Nevertheless most of the proof goes through for general C^*-algebras, in particular, Naimark's theorem still holds for unital C^*-algebras, which correspond to compact Hausdorff spaces by Thm. 9. One might thus hope to obtain a similar correspondence between order homomorphisms $\varphi : \mathcal{V}(\mathcal{A}_1) \to \mathcal{V}(\mathcal{A}_2)$ and Jordan $*$-homomorphisms $\Phi : \mathcal{J}(\mathcal{A}_1) \to \mathcal{J}(\mathcal{A}_2)$. This problem has been addressed in [74], where it is shown that every order isomorphism between the context categories lifts to a unique quasilinear Jordan $*$-isomorphism. The problem of extending the refomulation of Thm. 30 to C^*-algebras thus reduces to the following problem: When does a quasilinear map between C^*-algebras extend to a linear map? For general C^*-algebras, quasilinear maps are strictly weaker than linear maps, yet under certain additional constraints quasilinearity and linearity coincide [73]. Succinctly, a Gleason-type theorem is thus necessary for the reformulation of Stinespring's theorem over contexts.

Importantly, positive maps are more general than decomposable maps. In particular, without extending additivity of the local $\mathcal{B}(\mathcal{H})$-valued measures in contexts to their spectral dilations in Thm. 31, global sections of the probabilistic presheaf do not correspond with decomposable maps only, but with general positive maps. In contrast, global sections of the dilated probabilistic presheaf do correspond with decomposable maps, which allows us to rediscover Stinespring's theorem as follows. Recall that a Jordan $*$-homomorphism $\Phi : \mathcal{J}(\mathcal{N}) \to \mathcal{J}(\mathcal{B}(\mathcal{H}))$ lifts to a homomorphism of von Neumann algebras $\Phi : \mathcal{N} \to \mathcal{B}(\mathcal{H})$ if and only if it preserves orientations by Thm. 17 in Sec. 2.2. In fact, there is a canonical choice of orientation for every global section $\gamma \in \Gamma(\underline{\Pi}^{\mathcal{H}}(\mathcal{V}(\mathcal{N})))$ such that the corresponding map ϕ_γ in Thm. 33 is completely positive. This is the content of the following reformulation of Stinespring's theorem.

Theorem 34. (Stinespring in contextual form) *Let $\mathcal{N}$ be a von Neumann algebra with no summand of type I_2, $\mathcal{H}$ a Hilbert space, and $\underline{\Pi}^{\mathcal{H}}(\mathcal{V}(\mathcal{N}))$ the corresponding POVM presheaf of $\mathcal{N}$ over $\mathcal{V}(\mathcal{N})$. For every global section $\gamma \in \Gamma(\underline{\Pi}^{\mathcal{H}}(\mathcal{V}(\mathcal{N})))$ there exists a unique von Neumann algebra $\widetilde{\mathcal{N}}$ with $\mathcal{V}(\widetilde{\mathcal{N}}) \simeq \mathcal{V}(\mathcal{N})$, for which $\phi_\gamma : \widetilde{\mathcal{N}} \to \mathcal{B}(\mathcal{H})$ in Thm. 33 is completely positive.*

Proof. Let $\gamma \in \Gamma(\underline{\Pi}^{\mathcal{H}}(\mathcal{V}(\mathcal{N})))$. By Thm. 33 there exists a unique decomposable map $\phi_\gamma = v^*\Phi v$ with $v : \mathcal{H} \to \mathcal{K}$ a bounded linear operator and $\Phi : \mathcal{J}(\mathcal{N}) \to \mathcal{J}(\mathcal{B}(\mathcal{K}))$ a Jordan $*$-homomorphism. By Thm. 17 Φ becomes a homomorphism of von Neumann algebras if it also preserves orientations. This is the case if we 'pull back' the orientation on $\mathcal{B}(\mathcal{K})$ (respectively, $\mathcal{B}(\mathcal{H}) \supseteq v^*\Phi(\mathcal{B}(\mathcal{H}))v$) to $\mathcal{J}(\mathcal{N})$ by Kadison's theorem, Thm. 28, yielding a von Neumann algebra $\widetilde{\mathcal{N}}$ with $\mathcal{J}(\widetilde{\mathcal{N}}) \simeq \mathcal{J}(\mathcal{N})$ and thus $\mathcal{V}(\widetilde{\mathcal{N}}) \simeq \mathcal{V}(\mathcal{N})$ by Thm. 13. Hence, by Thm. 16 and Thm. 17 we find:

$$\Phi(ab) = \Phi(\{a, b\} + [a, b]) = \{\Phi(a), \Phi(b)\} + [\Phi(a), \Phi(b)] = \Phi(a)\Phi(b) \quad \forall a, b \in \widetilde{\mathcal{N}}$$

Clearly, Φ in $\phi_\gamma = v^*\Phi v$ becomes a $*$-homomorphism for this choice of orientation, which implies that $\phi_\gamma : \widetilde{\mathcal{N}} \to \mathcal{B}(\mathcal{H})$ is completely positive (cf. [135]). □

Stinespring's theorem, Thm. 30, thus also obtains a natural interpretation over contexts: we may view completely positive maps as global sections of the POVM presheaf over the partial order of contexts in Def. 29, with real-valued measures in Def. 27 replaced by positive operator-valued measures, locally dilated according to Naimark's theorem.

In particular, note that Thm. 34 offers a new, alternative proof for Thm. 30, at least for von Neumann algebras. In its original form, Thm. 30 has the following reading: it classifies completely positive maps $\phi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ from any C^*-algebra $\mathcal{A}$ into the bounded operators on some Hilbert space. There, the condition of complete positivity is a type of *global* condition on ϕ. Conversely, Thm. 34 allows for a different reading: viewing $\mathcal{A}$ as a partial order of commutative subalgebras, ϕ is clearly positive (in every commutative subalgebra) and can be lifted, first, to a spectral measure on some larger Hilbert space $\mathcal{K}$ *locally* by Naimark's theorem and, second, to a $*$-homomorphism under the canonical choice of commutators in $\Phi(\mathcal{N})$ by Thm. 34. This implies complete positivity *globally*. Succinctly, for von Neumann algebras (complete) positivity is a type of *local-to-global* property similar to linearity in Gleason's theorem and, in fact, a consequence of the latter in the form of Thm. 23.

It appears to be an advantage of the contextual perspective in the topos formalism, that positivity, dilations in contexts, and local time orientations are clearly differentiated, whereas these concepts are somewhat convoluted in the standard formalism, in the form of complete positivity. Note also that Thm. 34 is consistent with, and once again confirms the idea that quantum phenomena arise from phenomena in contexts together with their order relations.

The next section is concerned with Bell's theorem. The situation there is similar: let $\mathcal{H}$ in Def. 29 be such that we can find a representation $\mathcal{N}_2 \subseteq \mathcal{B}(\mathcal{H})$ (cf. Thm. 5). As a consequence, global sections of the (Bell) probabilistic presheaf over product contexts (cf. Eq. (2.24) below) correspond with quantum states on the composite system, in the form of completely positive maps, only if they also preserve commutators between von Neumann algebras $\mathcal{N}_1, \mathcal{N}_2$ corresponding to subsystems. As we will see, from a physical perspective, this extra structure imposes a consistency condition on time directions in subsystems.

2.4 Contextuality and Bell's theorem

Bell's seminal paper [17] responds to a longstanding conjecture by Einstein, Podolsky, and Rosen (EPR) [54], who claim quantum theory is only a statistical version of a more fundamental theory, similar to the relation between thermodynamics and statistical mechanics. Besides the probabilistic nature of quantum theory, this idea is motivated by certain nonlocal features present in the quantum formalism, believed to be resolved within the more fundamental theory.

As a response to EPR's thought experiment, Bell formalises EPR's assumption of an underlying space of hidden variables and derives a constraint for the maximal amount of correlations possible in such theories under the additional assumption of locality.[25] However, some quantum mechanically predicted and experimentally verified correlations [12, 65, 132] do not obey these constraints and thus cannot be reproduced by any local hidden variable model.

To give an example, we sketch a standard version of Bell's theorem. The CHSH inequality, named after Clauser, Horne, Shimony, and Holt, puts a bound on the outcome statistics of the nonlocal, bipartite quantity $c := a \times b + a \times b' + a' \times b - a' \times b'$ with local observables a, a' and b, b' and corresponding outcomes $A, A', B, B' \in \{-1, 1\}$ [36].[26] Assuming the existence of valuation functions in a classical (hidden variable) theory, the expectation value after repeated measurements is constrained by $\mathbb{E}_{\mathrm{cl}}(c) := \mathbb{E}_{\Sigma}(c) \leq 2$ (cf. Eq. (2.18)). Quantum correlations can be strictly stronger than classical (single-context) correlations. In particular, local spin-$1/2$ measurements $a(\alpha), a'(\alpha')$ and $b(\beta), b'(\beta')$, rotated by angles $\alpha = 0$, $\alpha' = \frac{\pi}{2}$ and $\beta = \frac{\pi}{4}$, $\beta' = -\frac{\pi}{4}$ evaluated on the Bell state $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, yield the expectation value $\mathbb{E}_{\mathrm{qm}}(c) := \langle \phi^+ | c | \phi^+ \rangle = 2\sqrt{2}$, and by Tsirelson's theorem exceed the classical bound maximally [141],

$$\mathbb{E}^{\max}_{\mathrm{cl}}(c) < \mathbb{E}^{\max}_{\mathrm{qm}}(c) < \mathbb{E}^{\max}_{\mathrm{ns}}(c) \ . \tag{2.10}$$

For later reference we have extended this inequality to the right by adding correlations in general non-signalling theories such as those arising from PR-boxes [125] (cf. Fig. 2.4). For the two-dimensional, bipartite case this leads to $\mathbb{E}^{\max}_{\mathrm{ns}}(c) = 4$ (see also Sec. 2.5.4).[27]

[25] For a discussion of the locality principles involved in Bell's theorem see also [25, 148]
[26] Note that c is not an observable, but derives from the statistical average over repeated measurements [103].
[27] We give a precise definition for the term 'general non-signalling theories' in Sec. 2.5.

As with the other theorems discussed in this article, we show that the essence in Bell's theorem is naturally encoded in the partial order of contexts, and we discuss the relation between contextuality and locality in this setting. The connection between these concepts has been highlighted before [6], here, we extend these results in several ways, in particular, we stress the importance of composition in this unified framework. This section gives a more detailed account of the results presented in the recent preprint [62].

We first recall the derivation of Bell's theorem, which also underlies the CHSH inequality above, emphasising the assumption of an underlying single-context state space, i.e., trivial physical contextuality, in Sec. 2.4.1. In Sec. 2.4.3 we argue how this assumption generalises to a multiple-context perspective after incorporating existing results on multipartite frame functions in Sec. 2.4.2. Finally, we show how the latter allow to upper bound correlations in theories that exhibit physical contextuality, i.e., for which not all observables are simultaneously measurable. In particular, we show how quantum correlations arise from global sections over the corresponding Bell presheaf in Sec. 2.4.4. Not surprisingly, the Bell presheaf is closely related to the (dilated) probabilistic presheaf, but is adapted to bipartite (or multipartite) systems. We finish with a discussion on correlations in general non-signalling theories, for which we give a definition in the language of presheaves over contexts in Sec. 2.5.

2.4.1 Correlations in classical theories

Classical state spaces

In Sec. 2.1.2 we have introduced the notion of physical contextuality as the mere order structure of contexts, i.e., collections of *simultaneously measurable* observables and their inclusion relations. In classical theories all observables are *simultaneously measurable*, hence, from the perspective of physical contextuality, they correspond to the trivial case of a single (maximal) context. In this section we give a derivation of Bell's theorem in the light of this assumption, in particular, we discuss the crucial notion of *composition* from the viewpoint of trivial physical contextuality.

As before it will be enough to consider the kinematics of the theory and we therefore start with a set of observables $\mathcal{O}$. Observables $a \in \mathcal{O}$ in classical theories are mathematically represented by measurable functions $f_a : \Sigma \to \mathbb{R}$ from some measure space (Σ, μ) to the real numbers. Σ is called the *(single-context) state space* of the theory. The elements $s \in \Sigma$ are called *microstates* and every $s \in \Sigma$ allows to assign truth values to propositions of the form '$a \in \Delta$' (read *'the observable a has a value within the Borel subset $\Delta \subset \mathbb{R}$'*):

$$\Theta(a \in \Delta, s) := \begin{cases} 1 \text{ if } s \in f_a^{-1}(\Delta) \\ 0 \text{ otherwise} \end{cases} \tag{2.11}$$

We can therefore speak of the *value of an observable $v_s(a)$ given the state $s \in \Sigma$* in the intuitive sense, i.e., through evaluation of the corresponding measurable function

$$v_s(a) := f_a(s) \; . \tag{2.12}$$

Valuation functions $v_s : \mathcal{O} \to \mathbb{R}$ in Eq. (2.12) are defined for all observables, in other words, every observable has an intrinsic (sharp) value in every state.[28] The observation that all observables *simultaneously* take deterministic values justifies to model *physical states* by points in some space Σ and observables by (measurable) functions $f_a : \mathcal{O} \to \mathbb{R}$ in the first place. This reasoning has to be revisited for non-classical theories, i.e., theories with non-trivial physical contextuality, and we will do so in the sections to follow. Note also that by this argument observables play the fundamental role, whereas states appear as a derived concept. This perspective will become important later, when we go over from single to multiple-context state spaces.

It is natural to equip $\mathcal{O}$ with the structure of an algebra. In fact, by modeling observables as functions we are automatically given a vector space structure as well as a product by pointwise multiplication of functions.[29] In fact, let (X, μ) be a σ-finite measure space, i.e., a measure space with a σ-finite measure μ, then $L^\infty(X, \mu)$ acts on the Hilbert space $L^2(X, \mu)$ by multiplication $\psi \to f\psi$ for all $\psi \in L^2(X, \mu)$ and $f \in L^\infty(X, \mu)$. If moreover X/N is a standard Borel space for $N := \{A \subset X \mid \mu(A) = 0\}$,[30] we obtain the following representation theorem (cf. [138]).

[28] The spectrum rule, $v_s(a) \in \mathrm{sp}(a) = \mathrm{Im}(f_a)$, is trivially satisfied. Note also that this does not imply determinism (as not all properties of s need to be directly observable).

[29] In doing so, we include the 'trivial' observable $e \in \mathcal{O}$ represented by the constant function $f_e = 1$. This observable simply asks the question 'Is the system there?', and the answer is always 'yes'.

[30] We will sometimes call a measure (X, μ) standard, if μ is a σ-finite measure and X modulo sets of measure

Theorem 35. *Every commutative von Neumann algebra on a separable Hilbert space is $*$-isomorphic to $l^\infty(\Omega)$, $L^\infty([0,1])$, or $L^\infty([0,1]) \oplus l^\infty(\Omega)$ for a countable set Ω, where μ is the counting or Lebesgue measure, respectively. Conversely, for every standard measure space (X, μ), $L^\infty(X, \mu)$ is a von Neumann algebra on a separable Hilbert space.*

With this analogy we will assume $\mathcal{O}$ to be a commutative von Neumann algebra, namely the algebra of measurable functions on a (standard) measure space. It is straightforward to extend the definition of valuation functions in Eq. (2.12) to this algebraic structure, namely for all $a, b \in \mathcal{O}$, $\lambda \in \mathbb{R}$, and $s \in \Sigma$ we set:

$$v_s(a \cdot b) := f_a(s) \cdot f_b(s), \quad v_s(a + b) := f_a(s) + f_b(s), \quad v_s(\lambda a) := \lambda f_a(s) \tag{2.13}$$

In other words, classical states $s \in \Sigma$ correspond to algebra homomorphisms $v_s : \mathcal{O} \to \mathbb{R}$. More generally, in the presence of physical contextuality, we may define generalised classical states to be valuation functions, i.e., partial algebra homomorphisms for which Eq. (2.13) only holds between algebras of simultaneously measurable observables, i.e., in contexts (cf. Def. 1). Clearly, this motivates the generalisation to partial algebra homomorphisms encountered in Sec. 2.1.1. Recall also that valuation functions play a central role in Thm. 19, which proves their non-existence within the setting of general von Neumann algebras, thus ruling out a classical state space picture. Bell's theorem attains a similar reformulation as a no-go-result for such classical states based on the additional assumption of composition. To see this in detail, in the remainder of this section we give a derivation of factorisability for classical (single-context) theories from composition defined by the canonical Cartesian product of state spaces.

Given two subsystems with (standard) measure spaces (Σ_1, μ_1), (Σ_2, μ_2), the product state space is defined as the Cartesian product $\Sigma_{1\&2} := \Sigma_1 \times \Sigma_2$ with product σ-algebra $\sigma_{1\&2}$ generated by elements $B_1 \times B_2$ for $B_1 \in \sigma_1$, $B_2 \in \sigma_2$, and the product measure $\mu_{1\&2} := \mu_1 \times \mu_2$ satisfies the condition,

$$\mu_{1\&2}(B_1 \times B_2) = \mu_1(B_1) \cdot \mu_2(B_2)\ .[31] \tag{2.14}$$

zero a standard Borel space.

[31] A product measure always exists, it is unique if the individual measures are also σ-finite (cf. Thm. 35). Note

In a similar way we obtain composite state spaces with multiple subsystems. Correspondingly, composite observables $a \in \mathcal{O}$ are represented by measurable functions $f_a : \Sigma \to \mathbb{R}$ on the composite state space $\Sigma = \times_{i=1}^n \Sigma_i$. Clearly, evaluation on elements $s \in \Sigma$ still yields algebra homomorphisms similarly to Eq. (2.13), hence, we obtain composite valuation functions $v_s : \mathcal{O} \to \mathbb{R}$ from the obvious generalisation of Eq. (2.12) to composite observables.

Importantly, the algebra of composite observables is generated by the algebras of its subsystems.[32] In order to obtain a generalisation of the truth values in Eq. (2.11), it is thus enough to consider tuples $\mathbf{a} = (a_1, \cdots, a_n) \in \mathcal{O}$ with $a_i \in \mathcal{O}_i$ for $i \in \{1, \cdots, n\}$ as well as measurable functions $\mathbf{f_a} : \Sigma \to \mathbb{R}^n$, $\mathbf{f_a}(s) := \times_{i=1}^n f_{a_i}(s_i)$ with $s \in \Sigma = \times_{i=1}^n \Sigma_i$. Namely, we define the truth value of the proposition '$\mathbf{a} \in \Delta$' with Borel set $\Delta := \times_{i=1}^n \Delta_i$ as follows:

$$\Theta(\mathbf{a} \in \Delta, s) := \begin{cases} 1 \text{ if } s \in \mathbf{f}_{\mathbf{a}}^{-1}(\Delta) \\ 0 \text{ otherwise} \end{cases} = \begin{cases} 1 \text{ if } s_i \in f_{a_i}^{-1}(\Delta_i) \ \forall i \\ 0 \text{ otherwise} \end{cases} = \prod_{i=1}^n \Theta(a_i \in \Delta_i, s_i) \qquad (2.15)$$

Statistical mixtures and joint probability distributions

The spectrum rule, $v(a) \in \mathrm{sp}(a)$, and functional composition principle in Eq. (2.9) are specific to pure states, (classical) mixed states on the other hand are modeled as statistical averages by means of normalised measures on state space $\mu : \Sigma \to \mathbb{R}$,

$$\int_\Sigma d\mu(s) = 1, \quad \forall s \in \Sigma : \ \mu(s) \geq 0 \,. \qquad (2.16)$$

The probability for the event corresponding to the Borel set $\Delta \subset \mathbb{R}$, when measuring the observable $a \in \mathcal{O}$ of a system in the mixed state μ, is then given by

$$\mu(\Delta \mid a) = \int_{\{s \in \Sigma | v_s(a) \in \Delta\}} d\mu(s) = \int_{f_a^{-1}(\Delta)} d\mu(s) = \int_\Sigma d\mu(s) \ \Theta(a \in \Delta, s) \,. \qquad (2.17)$$

In the last step we have used the indicator function $\Theta(a \in \Delta, s)$ in Eq. (2.11). For instance, the probability for obtaining a particular outcome A corresponds to the Borel set $\Delta := \{A\}$.

also that the Cartesian product extends to spaces with more structure, such as symplectic, Poisson manifolds etc.

[32]Here, 'generated' means under taking linear combinations, products, and pointwise limits.

Analogously, for product measures on a bipartite system we have by Eq. (2.14) and Eq. (2.15):

$$\begin{aligned}\mu(A,B \mid a,b) &= \int_{\Sigma} d\mu(s)\ \Theta((a,b)\in(A,B),s)\\ &= \left(\int_{\Sigma_1} d\mu_1(s_1)\ \Theta(a\in A,s_1)\right)\cdot\left(\int_{\Sigma_2} d\mu_2(s_2)\ \Theta(b\in B,s_2)\right)\\ &= \mu_1(A\mid a)\cdot\mu_2(B\mid b)\end{aligned}$$

A general mixed state on the composite system is then a statistical mixture of product measures,

$$\mu(A,B \mid a,b) = \int_{\Lambda} d\lambda\ p(\lambda)\ \mu_1(A\mid a,\lambda)\cdot\mu_2(B\mid b,\lambda)\ . \tag{2.18}$$

Note that $\mu(A,B \mid a,b)$ is a special type of *joint probability distribution*, i.e., a normalised measure on the composite system $\mu:\Sigma\rightarrow\mathbb{R}$, $\mu(s)\geq 0$ for all $s\in\Sigma$, and $\int_{\Sigma} d\mu(s)=1$. A joint probability distribution is called *factorisable* if it is of the form in Eq. (2.18).

The locality principle in factorisability is simply the condition that local measures depend on local outcomes and observables only. Clearly, by modeling the set of composite observables via the Cartesian product this is almost automatic—neither outcome nor observable affect the other factor in the product in Eq. (2.18).

The splitting of classical joint probability distributions according to factorisability thus fundamentally stems from the existence of local (single-context) state spaces with composition defined by the Cartesian product (cf. [13]). In other words, we have derived Eq. (2.18) from two assumptions:

(i) trivial physical contextuality, i.e., a single (maximal) context (in each subsystem) and

(ii) the Cartesian product of state spaces as the state space of the composite system.

Since condition (ii) is entirely natural for single-context state spaces, Eq. (2.18) can also be read as a consequence of *just* trivial physical contextuality.

Factorisability thus corresponds to composition given by the Cartesian product and by the above argument also to (trivial) physical contextuality. This suggests an intimate relationship between the underlying concepts:

contextuality $\longrightarrow$ composition $\longrightarrow$ locality

Admittedly, for now especially the first relation is a rather bold conjecture based on the case of trivial physical contextuality in classical systems only. However, in Sec. 2.4.3 we will see how these concepts are in fact closely related also in the multiple-context setting.

In a first attempt to define a notion of composition, which is suitable for the context structure in quantum systems and compatible with the more general (than factorisability) locality principle known as 'no-signalling', in the next section we study the composition behaviour of frame functions underlying the original proof of Gleason's theorem.

2.4.2 Composition of frame functions and Gleason's theorem

The content of this and the following sections is taken from [62].

Recall that measures on the projections of some Hilbert spaces $\mathcal{H}$ with $\dim(\mathcal{H}) \geq 3$ finite are classified by Gleason's theorem, Thm. 21. A closely related concept is that of *frame functions of weight* $W \in \mathbb{R}$ *on the unit sphere* $S(\mathcal{H})$: $f : S(\mathcal{H}) \rightarrow \mathbb{R}$ where $\sum_{j=1}^{d} f(v_j) = W$ for all orthonormal bases $(v_j)_{j=1}^{d} \in \mathrm{ONB}(\mathcal{H})$ with $d := \dim(\mathcal{H})$. In fact, Thm. 21 is a direct consequence of the following theorem about frame functions [66].

Theorem 36. *Let* $\dim(\mathcal{H}) \geq 3$ *finite. If* f *is a non-negative frame function of weight* $W \in \mathbb{R}^+$*, then there exists a density matrix* $\rho : \mathcal{H} \rightarrow \mathcal{H}$ *such that* $f(v) = W \mathrm{tr}(\rho p_v) = W \langle v|\rho|v\rangle$ *for all* $v \in S(\mathcal{H})$.[33]

Of course, we can apply Thm. 21 to composite quantum systems and consider frame functions $f : S(\mathcal{H}) \rightarrow \mathbb{R}$, where $\mathcal{H} = \otimes_{i=1}^{n} \mathcal{H}_i$ is the tensor product Hilbert space. However, in doing so we no longer restrict to outcomes of local measurements only. From an operational perspective the only outcomes accessible to local observers correspond to elements in $\sigma(\mathcal{H}) := \{v_1 \otimes \cdots \otimes v_n \in S(\mathcal{H}) \mid v_i \in S(\mathcal{H}_i)\}$. It is thus natural to consider *unentangled frame functions* with domain $\sigma(\mathcal{H}) \subsetneq S(\mathcal{H})$ and constraints restricted to $\mathrm{ONB}(\sigma(\mathcal{H})) \subsetneq \mathrm{ONB}(\mathcal{H})$ instead. This was studied in [145].

[33]Here and below, we identify rank-1 projections and vectors via $p_v = |v\rangle\langle v|$, and occasionally switch between the mathematical and the physically motivated Dirac notation of vectors $v \leftrightarrow |v\rangle$ (and dual vectors $v^* \leftrightarrow \langle v|$).

Theorem 37. (Wallach [145]) *Let $\mathcal{H} = \bigotimes_{i=1}^{n} \mathcal{H}_i$, $\dim(\mathcal{H}_i) \geq 3$ finite for all $i \in \{1, \cdots, n\}$, $n \in \mathbb{N}$. If $f : \sigma(\mathcal{H}) \to \mathbb{R}$ is a non-negative, unentangled frame function, then there exists a self-adjoint operator $t : \mathcal{H} \to \mathcal{H}$ such that $f(v) = \mathrm{tr}(tp_v) = \langle v|t|v\rangle$ for all $v \in \sigma(\mathcal{H})$.*

Note that in contrast to Thm. 21, Thm. 37 does not imply positivity of t.

A further restriction compared to $\mathrm{ONB}(\sigma(\mathcal{H}))$ are *frame functions over product bases*: $f : \sigma(\mathcal{H}) \to \mathbb{R}$ with $\sum_{j_1,\cdots,j_n=1}^{d_1,\cdots,d_n} f(v_{j_1,1} \otimes \cdots \otimes v_{j_n,n}) = W$, $d_i := \dim(\mathcal{H}_i)$ only on product bases, $\beta(\mathcal{H}) := \{(v_{j_1,1} \otimes \cdots \otimes v_{j_n,n})_{j_1,\cdots,j_n=1}^{d_1,\cdots,d_n} \mid (v_{j_i,i})_{j_i=1}^{d_i} \in \mathrm{ONB}(\mathcal{H}_i)\}$. Clearly, $S(\mathcal{H})$ contains many non-local states. But even unentangled bases cannot always be implemented with local operations and classical communication only [19], suggesting product bases as the most natural choice of constraints. Yet, it was shown that a similar result to Thm. 37 no longer holds for frame functions over product bases (cf. Prop. 5 in [145]).

To gain some insight into what is 'missing', it is helpful to consider examples of frame functions over product bases. Wallach gives a whole family of examples in [145], which are easily seen to correspond to signalling distributions. We thus add more constraints in the form of *no-signalling*: for $i \in \{1, \cdots, n\}$ with $(v_{j_i,i})_{j_i=1}^{d_i}, (w_{k_i,i})_{k_i=1}^{d_i} \in \mathrm{ONB}(\mathcal{H}_i)$ and $x_{l_r,r} \in S(\mathcal{H}_r)$ for all $l_r \in \{1, \cdots, d_r\}$, $r \neq i$,

$$\sum_{j_i=1}^{d_i} f(x_{l_r,r} \otimes v_{j_i,i}) = \sum_{k_i=1}^{d_i} f(x_{l_r,r} \otimes w_{k_i,i}) \,, \tag{2.19}$$

where we use the shorthand $x_{l_r,r} \otimes v_{j_i,i} := (x_{l_1,1} \otimes \cdots \otimes x_{l_{i-1},i-1} \otimes v_{j_i,i} \otimes x_{l_{i+1},i+1} \otimes \cdots \otimes x_{l_n,n})$. In light of PR-boxes one might still expect such *non-signalling frame functions* to be more general than quantum states. However, this turns out not to be the case. To show this we introduce yet another choice of basis: let $B \in \beta(\mathcal{H})$, $B' \in \mathrm{ONB}(\mathcal{H})$ and set $B' \sim B$ if there exists a sequence of unitaries $(U^m)_{m=1}^{N}$ such that $B^0 = B$, $B^m = U^m B^{m-1}$, $B^N = B'$ and where every unitary U^m acts non-trivially only on local subspaces of the form $x^m_{l_r,r} \otimes (v^m_{j_i,i} + v^m_{j'_i,i})$ with $x^m_{l_r,r} \otimes v^m_{j_i,i}, x^m_{l_r,r} \otimes v^m_{j'_i,i} \in B^k$. Importantly, the equivalence relation $\sim$ is independent of the choice of product basis $B \in \beta(\mathcal{H})$, it only depends on the split of Hilbert space $\mathcal{H} = \otimes_{i=1}^{n} \mathcal{H}_i$ (cf. Fig. 2.3). This follows since any $d_i \times d_i$-unitary matrix can be written as a product of two-level unitaries. In particular, any two bases related by local unitary transformations, i.e., unitaries acting on subsystems $\mathcal{H}_i$, are therefore related by $\sim$. We call the elements in $T(\beta(\mathcal{H})) := \{B' \in \mathrm{ONB}(H) \mid B' \sim B \in \beta(\mathcal{H})\}$ *twisted product bases*.[34] Note that $\beta(\mathcal{H}) \subsetneq T(\beta(\mathcal{H})) \subseteq \mathrm{ONB}(\sigma(\mathcal{H}))$. Clearly, the first inclusion is strict already for local dimensions $\dim(\mathcal{H}_i) \geq 2$. In fact, the latter inclusion is strict as well, there are unentangled bases in dimension at most ten that do not correspond to twisted product bases (cf. [102]).

Proposition 4. $T(\beta(\mathcal{H})) \subsetneq \mathrm{ONB}(\sigma(\mathcal{H}))$

Proof. Clearly, every twisted product basis is also an unentangled basis. The fact that the other direction fails is non-trivial, but can be concluded from a counterexample to Keller's tiling conjecture [102]: for $n \geq 10$ construct the following tiling of $\mathbb{R}^n$ by cubes of length 2 such that

(a) the centers of all cubes are in $\mathbb{Z}^n$,

(b) the tiling is $4\mathbb{Z}^n$-periodic,

(c) no two cubes have a complete facet in common.

More precisely, let $C := \{(x_1, \cdots, x_n) \mid -1 \leq x_i \leq 1 \quad \forall i \in \{1, \cdots, n\}\}$ denote a cube (of length 2). Then a tiling corresponds to 2^n equivalence classes of translates of C of the form $\mathbf{m} + C + 4\mathbb{Z}^n$ for

$$\mathbf{m} = (m_1, \cdots, m_n) \in \mathbb{Z}^n, \quad 0 \leq m_i \leq 3\,. \tag{2.20}$$

Next, consider the conditions: (i) $\mathbf{m}$ and $\mathbf{m}'$ have some $|m_i - m_i'| = 2$ and (ii) $\mathbf{m}$ and $\mathbf{m}'$ differ in two coordinate directions. Finally, denote by G_n and G_n^* two graphs, each of which has 4^n vertices labeled by the 4^n vectors in Eq. (2.20), where G_n has an edge between vertices $\mathbf{m}$ and $\mathbf{m}'$ if (i) holds, while G_n^* has an edge between vertices $\mathbf{m}$ and $\mathbf{m}'$ if (i) and (ii) hold.

Then a set $\mathcal{S}$ of 2^n vectors of the form in Eq. (2.20) yields a $4\mathbb{Z}^n$-periodic cube tiling if and only if $\mathcal{S}$ forms a clique in G_n; it yields a $4\mathbb{Z}^n$-periodic cube tiling with no two cubes having a complete facet in common if and only if $\mathcal{S}$ forms a clique in G_n^*.

We now translate this into a basis of $\mathcal{H} = (\mathbb{C}^2)^{\otimes 10}$. Consider the qubit states $|0\rangle$, $|1\rangle$,

[34] Twisted product bases arise in a similar (but more complex) way to the set of rotations of a Rubik's cube.

$|+\rangle := \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|-\rangle := \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ and define the correspondence $m_i \mapsto |\psi(m_i)\rangle$ by:

$$|\psi(m_i = 0)\rangle = |0\rangle_i \qquad\qquad |\psi(m_i = 1)\rangle = |+\rangle_i$$
$$|\psi(m_i = 2)\rangle = |1\rangle_i \qquad\qquad |\psi(m_i = 3)\rangle = |-\rangle_i$$

First, note that $|\psi(\mathcal{S})\rangle := \{|\psi(m_1)\rangle \otimes \cdots \otimes |\psi(m_n)\rangle \mid \mathbf{m} \in \mathcal{S}\}$ forms a basis of $(\mathbb{C}^2)^{\otimes 10}$: there are 2^{10} vectors and it is easily seen that $\langle\psi(\mathbf{m})|\psi(\mathbf{m}')\rangle = 0$ for $\mathbf{m}, \mathbf{m}' \in \mathcal{S}$, $\mathbf{m} \neq \mathbf{m}'$ by condition (i) above. Moreover, $|\psi(\mathcal{S})\rangle \in \sigma((\mathbb{C}^2)^{\otimes 10})$ is an unentangled basis by construction. However, any two vectors $|\psi(\mathbf{m})\rangle$, $|\psi(\mathbf{m}')\rangle$ for $\mathbf{m}, \mathbf{m}' \in \mathcal{S}$, $\mathbf{m} \neq \mathbf{m}'$ differ on at least two sites by condition (ii). It follows that no two-dimensional subspace of the form $x^m_{l_r,r} \otimes (v^m_{j_i,i} + v^m_{j_i',i})$ exists in $|\psi(\mathcal{S})\rangle$. Yet, any twisted product basis has at least one two-dimensional subspace of this form, hence, $|\psi(\mathcal{S})\rangle$ cannot be a twisted product basis. □

Thm. 37 fails for product bases, yet it *already* holds for frame functions over twisted product bases. Since the latter contain strictly fewer conditions than unentangled frame functions, this generalises Thm. 37.

Proposition 5. *Let $\mathcal{H} = \bigotimes_{i=1}^n \mathcal{H}_i$, $\dim(\mathcal{H}_i) \geq 3$ finite for all $i \in \{1, \cdots, n\}$, $n \in \mathbb{N}$. If $f : \sigma(\mathcal{H}) \to \mathbb{R}$ is a non-negative frame function over twisted product bases, then there exists a self-adjoint operator $t : \mathcal{H} \to \mathcal{H}$ such that $f(v) = \mathrm{tr}(tp_v) = \langle v|t|v\rangle$ for all $v \in \sigma(\mathcal{H})$.*

Proof. In the proof of Thm. 2 in [145] replace unentangled bases with twisted product bases in the inductive hypothesis. The case $n = 1$ still holds by Thm. 21. Consider $\mathcal{H} = \mathcal{H}_1 \otimes V$, $V = \otimes_{i=2}^n \mathcal{H}_i$ with $\dim(\mathcal{H}_i) \geq 3$ for all $1 \leq i \leq n$. If $(v_j)_{j=1}^{d_1} \in \mathrm{ONB}(\mathcal{H}_1)$ is an orthonormal basis of $\mathcal{H}_1$ and $(u^j_k)_{k=1}^{d_V} \in T(\beta(V))$ is a twisted product basis for V, then $(v_j \otimes u^j_k)_{j,k=1}^{d_1,d_V} \in T(\beta(\mathcal{H}))$ is a twisted product basis for $\mathcal{H}$. This follows since we can transform u^j_k for every j into a product basis on V by the assumption that $u^j_k \in T(\beta(V))$, and the fact that applying local unitaries on subspaces $\sum_{j_i=1}^{d_i} v_{j_r,r} \otimes v_{j_i,i}$ for all i we can transform between product bases in $\beta(V)$.

Since f is a twisted product frame function (of weight $W \in \mathbb{R}^+$), the function $f_v(u) = f(v \otimes u)$ is a non-negative twisted product frame function on V (of weight $W_v = W_{v_1} = W - \sum_{j=2,k=1}^{d_1,d_V} f(v_j \otimes u^j_k) \in \mathbb{R}^+$) for each $v \in \mathcal{H}_1$. By the inductive hypothesis we thus find $f_v(u) = \langle u|t_V(v)|u\rangle$ for all $u \in \sigma(V)$ with $t_V(v) : V \to V$ self-adjoint.

$$
\begin{array}{ccc}
|0\rangle|0+1\rangle & |0\rangle|0-1\rangle & |0+1\rangle|2\rangle \\
|1+2\rangle|0\rangle & |1\rangle|1\rangle & |0-1\rangle|2\rangle \\
|1-2\rangle|0\rangle & |2\rangle|1+2\rangle & |2\rangle|1-2\rangle
\end{array}
\quad \rightarrow \quad
\begin{array}{ccc}
|0\rangle|0+1\rangle & |0\rangle|0-1\rangle & |0+1\rangle|2\rangle \\
|1+2\rangle|0\rangle & |1\rangle|1\rangle & |0-1\rangle|2\rangle \\
|1-2\rangle|0\rangle & |2\rangle|1\rangle & |2\rangle|2\rangle
\end{array}
$$

$$
\downarrow
$$

$$
\begin{array}{ccc}
|0\rangle|0\rangle & |0\rangle|1\rangle & |0\rangle|2\rangle \\
|1+2\rangle|0\rangle & |1\rangle|1\rangle & |1\rangle|2\rangle \\
|1-2\rangle|0\rangle & |2\rangle|1\rangle & |2\rangle|2\rangle
\end{array}
\quad \leftarrow \quad
\begin{array}{ccc}
|0\rangle|0\rangle & |0\rangle|1\rangle & |0+1\rangle|2\rangle \\
|1+2\rangle|0\rangle & |1\rangle|1\rangle & |0-1\rangle|2\rangle \\
|1-2\rangle|0\rangle & |2\rangle|1\rangle & |2\rangle|2\rangle
\end{array}
$$

$$
\downarrow
$$

$$
\begin{array}{ccc}
|0\rangle|0\rangle & |0\rangle|1\rangle & |0\rangle|2\rangle \\
|1\rangle|0\rangle & |1\rangle|1\rangle & |1\rangle|2\rangle \\
|2\rangle|0\rangle & |2\rangle|1\rangle & |2\rangle|2\rangle
\end{array}
\quad = \quad \{|0\rangle, |1\rangle, |2\rangle\}^T \otimes \{|0\rangle, |1\rangle, |2\rangle\}
$$

Figure 2.3: The unentangled basis in the top left corner (cf. [19]) is transformed into a product basis (bottom left corner) by successively applying local unitaries, e.g., in the first step $(|2\rangle|1+2\rangle, |2\rangle|1-2\rangle) \to (|2\rangle|1\rangle, |2\rangle|2\rangle)$ where $|x\rangle|y\rangle := |x\rangle \otimes |y\rangle$ as well as $|x \pm y\rangle := \frac{1}{\sqrt{2}}(|x\rangle \pm |y\rangle)$.

Conversely, let $(u_k)_{k=1}^{d_V} \in T(\beta(V))$ be a twisted product basis for V and $(v_j^k)_{j=1}^{d_1} \in \mathrm{ONB}(\mathcal{H}_1)$ for every k, then $(v_j^k \otimes u_k)_{j,k=1}^{d_1,d_V} \in T(\beta(\mathcal{H}))$ is a twisted product basis for $\mathcal{H}$ (by a similar argument as before) and by the inductive hypothesis we conclude $f_u(v) := f(v \otimes u) = \langle v|t_{\mathcal{H}_1}(u)|v\rangle$ for all $v \in S(\mathcal{H}_1)$ with $t_{\mathcal{H}_1}(u) : \mathcal{H}_1 \to \mathcal{H}_1$ self-adjoint. The remainder of the proof is identical to the one of Thm. 37. $\square$

For instance, the unentangled basis in [19] is easily transformed into a product basis (cf. Fig. 2.3) and is thus in particular a twisted product basis. For frame functions over product bases consistency with such twisting operations is equivalent to no-signalling.

Lemma 1. *Non-negative, non-signalling frame functions $f : \sigma(\mathcal{H}) \to \mathbb{R}$ over product bases with finite local dimension $\dim(\mathcal{H}_i) \geq 3$ bijectively correspond to non-negative frame functions over twisted product bases.*

Proof. Let $x_{l_r,r} \in S(\mathcal{H}_r)$ for all $l_r \in \{1, \cdots, d_r\}$, $r \neq i$ and $(v_{j_i,i})_{j_i=1}^{d_i}, (w_{k_i,i})_{j_i=1}^{d_i} \in \mathrm{ONB}(\mathcal{H}_i)$ such that w.l.o.g. $|v_{1,i}\rangle\langle v_{1,i}| + |v_{2,i}\rangle\langle v_{2,i}| = |w_{1,i}\rangle\langle w_{1,i}| + |w_{2,i}\rangle\langle w_{2,i}|$ and $v_{j_i,i} = w_{k_i,i}$ for $3 \leq j_i = k_i \leq n$.

Then by no-signalling in Eq. (2.19),

$$\begin{aligned} f(x_{l_r,r}\otimes v_{1,i})+f(x_{l_r,r}\otimes v_{2,i}) &= \sum_{j_i=1}^{d_i} f(x_{l_r,r}\otimes v_{j_i,i}) - \sum_{j_i=3}^{d_i} f(x_{l_r,r}\otimes v_{j_i,i}) \\ &= \sum_{k_i=1}^{d_i} f(x_{l_r,r}\otimes w_{k_i,i}) - \sum_{k_i=3}^{d_i} f(x_{l_r,r}\otimes w_{k_i,i}) \\ &= f(x_{l_r,r}\otimes w_{1,i}) + f(x_{l_r,r}\otimes w_{2,i}). \end{aligned}$$

As twisted product bases are generated from local unitaries acting on two-dimensional subspaces of the form $x_{l_r,r}\otimes(v_{j_i,i}+v_{j'_i,i})$, f is also a frame function over twisted product bases. Conversely, for the latter Eq. (2.19) holds since it holds already for two-dimensional subspaces. $\square$

Theorem 38. *Let $\mathcal{H}=\bigotimes_{i=1}^{n}\mathcal{H}_i$, $\dim(\mathcal{H}_i)\geq 3$ finite for all $i\in\{1,\cdots,n\}$, $n\in\mathbb{N}$. If $f:\sigma(\mathcal{H})\rightarrow\mathbb{R}$ is a non-negative, non-signalling frame function over product bases, then there exists a self-adjoint operator $t:\mathcal{H}\rightarrow\mathcal{H}$ such that $f(v)=\mathrm{tr}(tp_v)=\langle v|t|v\rangle$ for all $v\in\sigma(\mathcal{H})$.*

Proof. This follows immediately from Lm. 1 and Prop. 5. $\square$

Note that our results only apply to finite local dimensions $\dim(\mathcal{H}_i)\geq 3$, this restriction in Thm. 38 is due to Thm. 21, equivalently Thm. 36. Nevertheless, generalisations of Thm. 21 to two dimensions exist based on (subsets of) positive operator-valued measures (POVMs) [22, 30, 150]. More precisely, non-negative frame functions $f:\mathcal{E}(\mathcal{H})\rightarrow\mathbb{R}^+$ of weight $W\in\mathbb{R}^+$ with domain $\mathcal{E}(\mathcal{H})$ the set of all *effects*, i.e., convex combinations of projections, and such that $\sum_{i\in I} f(E_i)=W$ whenever $\sum_{i\in I}E_i=1$, correspond to density matrices: $f(E)=W\mathrm{tr}(\rho E)$ for all $E\in\mathcal{E}(\mathcal{H})$ and $\dim(\mathcal{H})\geq 2$ finite.

Similarly, replacing $\sigma(\mathcal{H})$ by $\sigma(\mathcal{E}(\mathcal{H}))$ (equivalently, projection-valued measures (PVMs) by POVMs) in the otherwise analogous definitions of (twisted) product frame functions and no-signalling in Eq. (2.19), one obtains a generalisation to systems with dimension $\dim(\mathcal{H}_i)=2$.

Theorem 39. *Let $\mathcal{H}=\bigotimes_{i=1}^{n}\mathcal{H}_i$, $\dim(\mathcal{H}_i)\geq 2$ finite for all $i\in\{1,\cdots,n\}$, $n\in\mathbb{N}$. If $f:\sigma(\mathcal{E}(\mathcal{H}))\rightarrow\mathbb{R}$ is a non-negative, non-signalling frame function over product POVMs, then there exists a self-adjoint operator $t:\mathcal{H}\rightarrow\mathcal{H}$ such that $f(E)=\mathrm{tr}(tE)$ for all $E\in\sigma(\mathcal{E}(\mathcal{H}))$.*

Proof. By [22] frame functions over $\sigma(\mathcal{E}(\mathcal{H}))$ correspond to quantum states for every $\mathcal{H}_i$ in $\mathcal{H} = \bigotimes_{i=1}^n \mathcal{H}_i$. With this the inductive proof of Prop. 5 goes through also for $\dim(\mathcal{H}) = 2$. The same holds for the correspondence of the no-signalling condition in Eq. (2.19) and constraints on f arising from local transformations leaving convex combinations $\sum_i E_i \leq 1$ invariant in Lm. 1. □

We thus find that no-signalling is *almost* enough to restrict frame functions (of weight 1) over product bases to quantum states. To be precise, by Prop. 5 the correspondence is with self-adjoint operators of unit trace, which are positive on product states.

We combined this result with earlier work in [145] and added the important distinction between unentangled and twisted product basis frame functions, as no-signalling becomes redundant in the latter case by Lm. 1 and Prop. 5. A more direct way to study non-signalling probability distributions is by means of contextuality. In the next section we thus reformulate our results in contextual form and show how no-signalling arises as a subset of the marginalisation constraints over product contexts.

2.4.3 Contextuality, composition, and locality

Note that the derivation in Sec. 2.4.1 crucially depended on the assumption of an underlying classical state space with composition defined in terms of the Cartesian product. In this section we discuss alternative ways for composition, in particular, we motivate composition of systems based on observables and their context order instead of state spaces. In the subsequent sections we study the implications of this context structure for the spectral and the probabilistic presheaf. In doing so we again consider general von Neumann algebras, in particular, we extend and generalise the results on frame functions in the last section.

Composition via Cartesian products of state spaces. Recall that we defined composition of classical systems in terms of their state spaces, namely, via the product of the corresponding measure spaces. On the other hand, observables in classical theories are represented by measurable functions and every measurable function on the composite state space can be approximated by suitable limits of linear combinations of indicator functions (cf. Eq. (2.15)). In this sense, it does not matter whether we define composition in terms of states or observables for classical systems.[35]

Taking classical systems to be represented by commutative von Neumann algebras $\mathcal{N}_i$, $i = 1, 2$ with corresponding state spaces corresponding to the Gelfand spectra $\Sigma_i = \Sigma(\mathcal{N}_i) \simeq \Gamma(\underline{\Sigma}(\mathcal{V}(\mathcal{N}_i)))$,[36] this equivalence reads,

$$\Sigma_{1\&2} = \Sigma_1 \times \Sigma_2 = \Gamma(\underline{\Sigma}(\mathcal{V}(\mathcal{N}_1))) \times \Gamma(\underline{\Sigma}(\mathcal{V}(\mathcal{N}_2))) = \Gamma(\underline{\Sigma}(\mathcal{V}(\mathcal{N}_1) \times \mathcal{V}(\mathcal{N}_2))) \ . \tag{2.21}$$

Here, the final equality is with respect to the context product in Eq. (2.24) below.

By the Kochen-Specker theorem in contextual form, Thm. 20, $\Gamma(\underline{\Sigma}(\mathcal{V}(\mathcal{N})))$ is empty whenever $\mathcal{N}$ is a noncommutative von Neumann algebra (not only consisting of summands of type I_1, I_2). The equivalence in Eq. (2.21) thus breaks down for such algebras. Nevertheless, composition in terms of state spaces can be carried over to quantum systems if we define the state space of the composite system in terms of convex combinations of elements in the Cartesian product of global sections of the probabilistic presheaves of subsystems instead:

$$\Gamma_{1\&2} := \mathrm{Conv}(\Sigma_{1\&2}), \quad \Sigma_{1\&2} := \Gamma(\underline{\Pi}(\mathcal{N}_1)) \times \Gamma(\underline{\Pi}(\mathcal{N}_2))^{37} \tag{2.22}$$

Compare this with mixed states in classical theories, which are given by convex combinations of elements in the Cartesian product of global sections of the spectral presheaf by factorisability, Eq. (2.18). In fact, factorisability holds for *any* system—with or without local physical contextuality—as long as composition is defined by means of Eq. (2.22).

Proposition 6. *Let $\mathcal{N}_1$, $\mathcal{N}_2$ be possibly noncommutative von Neumann algebras and let the set of states on the composite system $\Gamma_{1\&2}$ be defined according to Eq. (2.22). Then all states $\gamma \in \Gamma_{1\&2}$ are factorisable and satisfy the Bell inequalities.*

[35] A similar duality holds even in the presence of physical contextuality for state spaces given by the non-empty sets of global sections of spectral presheaves (cf. Prop. 8 in Sec. 2.5.2).

[36] Recall that by Thm. 10, the category of commutative von Neumann algebras is dually equivalent to the category of Hyperstonean spaces, equivalently, standard measure spaces modulo set of measure zero by Thm. 35.

[37] In the context of generalised probabilistic theories, this construction is sometimes called the minimal tensor product [91].

Proof. Let $\gamma \in \Gamma_{1\&2}$. By definition, γ is a convex combination of product states $(\gamma_1, \gamma_2) \in \Sigma_{1\&2}$ with $\gamma_1 \in \Gamma(\underline{\Pi}(\mathcal{N}_1))$ and $\gamma_2 \in \Gamma(\underline{\Pi}(\mathcal{N}_2))$. Hence, there exists a measure space (Λ, μ_γ) with measure $\mu_\gamma : \Lambda \to \mathbb{R}_0^+$ such that for all $p \in \mathcal{P}(\mathcal{N}_1)$, $q \in \mathcal{P}(\mathcal{N}_2)$,

$$\gamma(p, q) = \int_\Lambda d\mu_\gamma(\lambda)\ \gamma_1(p \mid \lambda) \cdot \gamma_2(q \mid \lambda) = \int_\Lambda d\mu_\gamma(\lambda)\ \mathrm{tr}(\rho^\lambda_{\gamma_1} p) \cdot \mathrm{tr}(\rho^\lambda_{\gamma_2} q)\ . \tag{2.23}$$

The last equality follows by Gleason's theorem in contextual form, Thm. 24. Eq. (2.23) is just factorisability and the Bell inequalities thus necessarily hold. □

This argument is not restricted to states on von Neumann algebras, e.g. density matrices, but holds for arbitrary locally stochastic models with composition defined by the Cartesian product in Eq. (2.22). Every stochastic, factorisable model thus satisfies the Bell inequalities (cf. [35]). Moreover, it is interesting to note that by [57, 58] the existence of the latter is equivalent to the existence of a deterministic local hidden variable model of the composite system. In this sense, even stochastic, factorisable models with local physical contextuality, yet composition defined via Eq. (2.22), still correspond to single-context state spaces.

Succinctly, by Prop. 6 factorisability is a direct consequence of composition defined in terms of the Cartesian product on state spaces. Yet, while this construction is natural for classical theories (and as we will see in Sec. 2.5, also for generalised classical theories), this is no longer the case for more general states arising as global sections of the Bell presheaf defined on the composite system (cf. Eq. (2.24) below). Note that such theories are of great interest, since Thm. 19 rules out valuation functions, equivalently, global sections of the spectral presheaf, already in subsystems. We thus seek a unified notion of composition that relates to physical contextuality and incorporates both the classical and the quantum case.

Composition via contexts. It is clear from Prop. 6 and the above argument on stochastic, factorisable models (cf. [57, 58]) that Eq. (2.22) does not define composition for theories with physical contextuality. Instead, we have seen that on the level of frame functions, twisted product bases are a natural choice relating to no-signalling. In fact, there is a more direct way to encode no-signalling using contextuality. Recall that at the core of contextuality lies the notion of *simultaneous measurability*: we say that a physical system is contextual if not all its observables $\mathcal{O}$ can be measured simultaneously in every state. Clearly, classical systems are non-contextual and compose by the Cartesian product. Yet, also contextual systems can contain sets of simultaneously measurable observables called *contexts*. Shifting focus from states to observables and their context order, we define a notion of composition of contexts by the canonical product on partial orders, denoted $\mathcal{V}_{1\&2} := (\mathcal{V}_1 \times \mathcal{V}_2, \subseteq_{1\&2})$, and given by the Cartesian product of elements $\mathcal{V}_1 \times \mathcal{V}_2$ with order relations such that for all $\tilde{V}_1, V_1 \in \mathcal{V}_1$, $\tilde{V}_2, V_2 \in \mathcal{V}_2$:

$$(\tilde{V}_1, \tilde{V}_2) \leq (V_1, V_2) :\Longleftrightarrow \tilde{V}_1 \leq_1 V_1 \text{ and } \tilde{V}_2 \leq_2 V_2 \tag{2.24}$$

Accordingly, we define the spectral presheaf $\underline{\Sigma}_{1\&2} := \underline{\Sigma}(\mathcal{V}_{1\&2})$ and the probabilistic presheaf $\underline{\Pi}_{1\&2} := \underline{\Pi}(\mathcal{V}_{1\&2})$ over this product context order.

Recall that in this setup a measure on the projection lattice $\mu : \mathcal{P}(\mathcal{H}) \to [0,1]$ becomes a collection of probability distributions $(\mu_V)_{V \in \mathcal{V}(\mathcal{H})}$, one for every context. Moreover, *non-contextuality* further constrains these *across* different contexts: let $\mu_{\tilde{V}}$, μ_V be measures over contexts $\tilde{V}, V$, $\tilde{V} \leq V$, then $\mu_{\tilde{V}}$ is obtained from μ_V by marginalisation, denoted $\mu_{\tilde{V}} = \mu_V|_{\tilde{V}}$. Given the results on frame functions over product bases in [145], one might be sceptical whether global sections over product contexts always correspond to quantum states. However, the following lemma together with Lm. 1 shows that no-signalling is already contained in the contextual constraints between product contexts.

Lemma 2. *Global sections of the probabilistic presheaf (cf. Def. 26) over product contexts, $\gamma \in \Gamma(\underline{\Pi}(\mathcal{V}(\mathcal{H}_1) \times \mathcal{V}(\mathcal{H}_2)))$ with $\dim(\mathcal{H}_i)$ finite, bijectively correspond to non-negative frame functions of weight 1 over twisted product bases, $f : \sigma(\mathcal{H}) \to \mathbb{R}$.*

Proof. Clearly, a frame function over twisted product bases defines a global section on product contexts by $\gamma_f(p_{v_1} \otimes p_{v_2}) = f(v_1 \otimes v_2)$, $p_v = |v\rangle\langle v|$ for all $v_i \in \sigma(\mathcal{H}_i)$. Marginalisation between contexts follows from the constraints on f between twisted product bases.

Conversely, a global section over product contexts $V \in \mathcal{V}(\mathcal{H}_1) \times \mathcal{V}(\mathcal{H}_2)$ defines a map $f_\gamma : \sigma(\mathcal{H}) \to \mathbb{R}_0^+$ by $f_\gamma(v_{p_1} \otimes v_{p_2}) := \gamma(p_1 \otimes p_2)$ for all $p_i \in \mathcal{P}(\mathcal{H}_i)$. Moreover, it satisfies the constraints encoded in twisted product bases, which for global sections arise from marginalisation between product contexts of the form (and by symmetry for $i = 1 \leftrightarrow i = 2$):

$$\begin{aligned} V &:= V_1 \times \{p_{1,2}, p_{2,2}, (p_{1,2} + p_{2,2})^\perp\} \\ W &:= V_1 \times \{q_{1,2}, q_{2,2}, (p_{1,2} + p_{2,2})^\perp\} \\ \tilde{V} &:= V_1 \times \{(p_{1,2} + p_{2,2}), (p_{1,2} + p_{2,2})^\perp\} \end{aligned} \qquad \begin{tikzcd} V & & W \\ & \tilde{V} \arrow[ul, hook] \arrow[ur, hook] & \end{tikzcd}$$

Here, we defined contexts via their projections $p_{j_i,i} := |v_{j_i,i}\rangle\langle v_{j_i,i}|$, $q_{k_i,i} := |w_{k_i,i}\rangle\langle w_{k_i,i}|$ corresponding to product bases $(v_{j_i,i})_{j_i=1}^{d_i}, (w_{k_i,i})_{k_i=1}^{d_i} \in \mathrm{ONB}(\mathcal{H}_i)$ such that $V_1 = \{p_{1,1}, \cdots, p_{d_1,1}\}$ and $p_{1,2} + p_{2,2} = q_{1,2} + q_{2,2}$ (cf. proof of Lm. 1). $\square$

The generalisation to the multipartite setting is analogous. Note that in going over to contexts (equivalently from frame functions to measures) we achieve a type of trade-off: while we do not consider contexts corresponding to twisted product bases directly, there are more constraints between contexts that effectively contain the same information as frame functions over twisted product bases. In particular, no-signalling is contained in the marginalisation maps between product contexts. More precisely, by a similar argument to the one in Lm. 2, it is easy to see that the following conditions, which correspond with the no-signalling condition on frame functions in Eq. (2.19), are identical to those in Eq. (2.24) if one also demands transitivity:

$$(\tilde{V}_1, \tilde{V}_2) \subseteq_{\mathrm{ns}} (V_1, V_2) \;:\Longleftrightarrow\; (\tilde{V}_1 = V_1, \tilde{V}_2 \subseteq V_2) \text{ or } (\tilde{V}_1 \subseteq V_1, \tilde{V}_2 = V_2) \tag{2.25}$$

We thus obtain the following reformulation of Thm. 38.

Theorem 40. *For every global section of the probabilistic presheaf over product contexts* $\gamma \in \Gamma(\underline{\Pi}(\mathcal{V}(\mathcal{H}_1) \times \mathcal{V}(\mathcal{H}_2)))$ *with* $3 \leq \dim(\mathcal{H}_i)$ *finite, there exists a self-adjoint operator* $t : \mathcal{B}(\mathcal{H}_1) \otimes \mathcal{B}(\mathcal{H}_2) \to \mathbb{R}$ *such that* $\mathrm{tr}(t) = 1$ *and* $t(p_1 \otimes p_2) \geq 0$ *for all* $p_1 \in \mathcal{P}(\mathcal{H}_1)$, $p_2 \in \mathcal{P}(\mathcal{H}_2)$.

Proof. This follows directly from Lm. 2 and Thm. 38. $\square$

Thm. 40 (and Thm. 38) are very close to a bijective correspondence: for every global section of the probabilistic presheaf over product contexts there *exists* a corresponding self-adjoint operator t of unit trace. Moreover, note that if t is positive and appropriately normalised, i.e., it has unit partial traces, it defines a *unique* quantum state since local measurement statistics are sufficient to distinguish between arbitrary quantum states [149].[38] However, the probabilistic presheaf over product contexts is not quite enough to single out quantum states, since not all operators in Thm. 40 correspond to quantum states (cf. [100]). In order to relate non-signalling joint probability distributions over product contexts with quantum states, it will be crucial to consider the *dilated* probabilistic presheaf over product contexts.

Definition 30. *Let $\mathcal{N}_1, \mathcal{N}_2$ be von Neumann algebras with context categories $\mathcal{V}(\mathcal{N}_1), \mathcal{V}(\mathcal{N}_2)$. The (normal)* Bell presheaf *is the (normal) dilated probabilistic presheaf (cf. Def. 27) over product contexts, $\underline{\Pi}(\mathcal{V}(\mathcal{N}_1) \times \mathcal{V}(\mathcal{N}_2))$, restricted to locally normalised measures, i.e., $\mu_V|_{V_2} \in \underline{\Pi}_{V_1}$ and $\mu_V|_{V_1} \in \underline{\Pi}_{V_2}$ for all $V = V_1 \times V_2 \in \mathcal{V}(\mathcal{N}_1) \times \mathcal{V}(\mathcal{N}_2)$.*[39]

Composition via tensor products. Before we explore the consequences of context composition for the Bell presheaf in more detail, we end this section by mentioning a third way of defining composition, which in fact is the standard composition in quantum theory. There, the pure state space is the projective space $\mathbb{P}(\mathcal{H})$ corresponding to the Hilbert space $\mathcal{H}$. Given component systems with Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2$, the Hilbert space of the composite system has Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$ and pure state space

$$\mathbb{P}_{1\&2} := \mathbb{P}(\mathcal{H}_1 \otimes \mathcal{H}_2) \ .$$

Note that there are many more contexts for this kind of composition than for composition via contexts described above: the poset $\mathcal{V}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ contains many contexts that are not of the form $V_1 \otimes V_2$, which are the only contexts available in the poset $\mathcal{V}_{1\&2}$. In fact, the functor

$$\mathcal{V}_{1\&2} \longrightarrow \mathcal{V}(\mathcal{H}_1 \otimes \mathcal{H}_2), \qquad (V_1, V_2) \longmapsto V_1 \otimes V_2$$

is fully faithful, but not essentially surjective (surjective on objects). We say that $\mathcal{V}_{1\&2}$ contains only *product* (or *twisted product*) contexts (cf. [62]).

[38] This property of quantum theory is sometimes called local tomography [77].
[39] Note that the Bell presheaf assigns to every product context $V = V_1 \times V_2 \in \mathcal{V}(\mathcal{N}_1) \times \mathcal{V}(\mathcal{N}_2)$ the set of all dilations of joint probability distributions over V, in other words, all factorisable joint probability distributions.

2.4.4 Bell's theorem in contextual form

We combine the results obtained in the previous sections into a reformulation of Bell's theorem in terms of its restriction on state spaces. Recall that we defined composition in terms of the order structure of observables, which encodes physical contextuality. While this shift of perspective does not change the way systems compose in classical theories, it does change the way systems compose in the quantum case, where composite systems are usually defined in terms of the tensor product. In fact, composition of contexts yields a self-adjoint operator of norm one and thus almost a quantum state in the finite-dimensional case by Thm. 40. Our main theorem in this section provides the missing link to establish positivity of this operator and thus the desired bijective correspondence between global sections of the Bell presheaf and quantum states. As a consequence, we obtain a unified framework for composition and locality by means of physical contextuality, valid in both classical and quantum physics.

In a nutshell, global sections of the Bell presheaf correspond with (quantum) states on algebras with specific time orientations. Here, unlike in the local case, it is crucial to consider the dilated probabilistic presheaf in Def. 27, since only global sections of the latter correspond with Jordan $*$-homomorphisms, for which the consistency condition between local time orientations can be expressed in terms of dynamical correspondences, which lift the Jordan $*$-homomorphisms to $*$-homomorphisms (cf. Prop. 2). To make this explicit, we need the oriented context category from Def. 24. (For more details, see Sec. 2.2.3 as well as [7, 46].) Global sections of the Bell presheaf over the oriented context category need to be consistent with its inherent orientation.

Definition 31. *Let $\mathcal{N}_1$, $\mathcal{N}_2$ be von Neumann algebras with no summand of type I_2 and $\mathcal{V}(\mathcal{N}_1)$, $\mathcal{V}(\mathcal{N}_2)$ the corresponding context categories with respective time orientations $\tilde{\psi}_1$, $\tilde{\psi}_2$. A global section of the Bell presheaf $\gamma \in \Gamma(\underline{\Pi}(\mathcal{V}_{1\&2}))$ with $\mathcal{V}_{1\&2} = \mathcal{V}(\mathcal{N}_1) \times \mathcal{V}(\mathcal{N}_2)$ is called* orientation-preserving with respect to $\tilde{\psi} = (\tilde{\psi}_1, \tilde{\psi}_2)$ *if*

$$\forall a \in \mathcal{B}(\mathcal{N}_1), t \in \mathbb{R}: \quad \Phi_\gamma \circ e^{t\psi_1(a)} = e^{t\psi_2(\Phi_\gamma(a))} \circ \Phi_\gamma \, , \tag{2.26}$$

where Φ_γ is the Jordan $*$-homomorphism in Thm. 33.

The set of orientation-preserving global sections with respect to $\psi = (\psi_1, \psi_2)$ is denoted

$$\Gamma(\underline{\Pi}(\widetilde{\mathcal{V}_{1\&2}})) := \{\gamma \in \Gamma(\underline{\Pi}(\mathcal{V}_{1\&2})) \mid \gamma \text{ is orientation} - \text{preserving with respect to } \psi\} \ ,$$

where $\widetilde{\mathcal{V}_{1\&2}} := (\mathcal{V}_{1\&2}, \tilde{\psi}) = (\mathcal{V}(\mathcal{N}_1), \tilde{\psi}_1) \times (\mathcal{V}(\mathcal{N}_2), \tilde{\psi}_2) = \widetilde{\mathcal{V}(\mathcal{N}_1)} \times \widetilde{\mathcal{V}(\mathcal{N}_2)}$.

Being orientation-preserving is a property, which explicitly refers to composite systems. Nevertheless, since $\mathbb{C}$ can also be interpreted as a Jordan and von Neumann algebra, every global section $\gamma \in \Gamma(\underline{\Pi}(\mathcal{V}(\mathcal{N})))$ trivially defines a Jordan and von Neumann algebra homomorphism $\gamma : \mathcal{N} \to \mathbb{C}$ also for single systems. The condition in Eq. (2.26) is trivially satisfied in this case and γ therefore orientation-preserving. Note that Gleason's theorem in contextual form, Thm. 24 (Thm. 25), can therefore equivalently be phrased in terms of orientation-preserving global sections.

We will also need to relate completely positive maps with self-adjoint operators. In finite dimensions (and for single factors) this correspondence is established by the Choi-Jamiołkowski isomorphism between completely positive maps $\phi : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ and positive operators on the tensor product $M_{mn}(\mathbb{C}) \simeq M_m(\mathbb{C}) \otimes M_n(\mathbb{C})$ [32, 90].

Theorem 41. (Choi [32]) *Let $\phi : M_n(\mathbb{C}) \to M_m(\mathbb{C})$ be a linear map. Then ϕ is completely positive if and only if the (operator) matrix $\rho_\phi : M_{mn}(\mathbb{C}) \to M_{mn}(\mathbb{C})$, $\rho_\phi = \sum_{ij} E_{ij} \otimes \phi(E_{ij})$ is positive, where E_{ij} denotes the matrix with entry 1 in position (i, j) and 0 otherwise.*

In infinite dimensions and for general C^*-algebras the correspondence between positive and completely positive maps is slightly more complex. In particular, note that in finite dimensions the trace defines a special type of completely positive map: $\phi_0 : \mathcal{N}_1 \to \mathcal{N}_2$, $\phi_0 := 1_{\mathcal{N}_2} \mathrm{tr}_{\mathcal{N}_1}$. In infinite dimensions such a tracial state is neither guaranteed to exist nor is it unique if it does. Instead, we need an alternative reference map ϕ_0 as well as certain continuity conditions. We refer to [16] for the definitions. Once such a reference map is given, a similar correspondence to Thm. 41 holds.

Theorem 42. (Belavkin [16]) *Let $\mathcal{A}$ be a C^*-algebra and denote the space of bounded linear operators on Hilbert space $\mathcal{H}$ by $\mathcal{B}(\mathcal{H})$. Let $\phi, \phi_0 : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ be bounded completely positive* *maps and let $\mathcal{K}$ be a Hilbert space for a representation $\pi : A \to \mathcal{B}(\mathcal{K})$ in which ϕ_0 is spatial, i.e.,*

$$\phi_0(a) = v^* \pi(a) v, \quad \forall a \in \mathcal{A} \,,$$

where $v : \mathcal{H} \to \mathcal{K}$ is a bounded operator.

1. *ϕ is completely absolutely continuous with respect to ϕ_0 if and only if it has a spatial representation $\phi(a) = v'^* \pi(a) v'$ with $\pi(a)v' = \theta\pi(a)v$, where θ is a densely defined operator in the minimal $\mathcal{H}$, commuting with $\pi(\mathcal{A}) = \{\pi(a) \mid a \in \mathcal{A}\}$ on the lineal $\mathcal{D} = \{\sum_j \pi(a_j) F \eta_j\}$.*
2. *ϕ is strongly completely absolutely continuous with respect to ϕ_0 if and only if ϕ is spatial in $(\pi, \mathcal{H})$ and there exists a positive self-adjoint operator ρ, uniquely defined on $\mathcal{D}$, affiliated with the commutant $\pi(\mathcal{A})'$ and such that $\forall a \in \mathcal{A}$:*

$$\phi(a) = v^* \rho \pi(a) v = (\rho^{\frac{1}{2}} v)^* \pi(a) (\rho^{\frac{1}{2}} v) \,, \tag{2.27}$$

3. *ϕ is completely dominated by ϕ_0 if and only if Eq. (2.27) holds and ρ is bounded.*

We are now in the position to prove the following.

Theorem 43. *Let $\mathcal{N}_1$, $\mathcal{N}_2$ be von Neumann algebras with no summand of type I_2. There is a bijective correspondence between the set of finitely (completely) additive, orientation-preserving global sections of the Bell presheaf $\Gamma(\underline{\Pi}(\widetilde{\mathcal{V}_{1\&2}}))$, $\widetilde{\mathcal{V}_{1\&2}} := \widetilde{\mathcal{V}(\mathcal{N}_1)} \times \widetilde{\mathcal{V}(\mathcal{N}_2)}$ and the set of (normal) states $\mathcal{S}_{1\&2} := \mathcal{S}(\mathcal{N}_1 \bar{\otimes} \mathcal{N}_2)$ on the spatial tensor product algebra $\mathcal{N}_1 \bar{\otimes} \mathcal{N}_2$.*

Proof. It is not too to hard to see that every state $\sigma \in \mathcal{S}(\mathcal{N}_1 \bar{\otimes} \mathcal{N}_2)$ gives rise to an orientation-preserving global section of the Bell presheaf $\gamma_\sigma \in \Gamma(\underline{\Pi}(\widetilde{\mathcal{V}_{1\&2}}))$.

For the other direction, we proceed in several steps. We first show that associated with every global section there exists a linear operator on the tensor product algebra. This generalises Thm. 37, which deals with finite dimensions and single factors only. We then prove positivity of this operator from complete positivity of an associative map under the consistency condition between local time orientations. This step crucially hinges on the fact that we consider the dilated probabilistic presheaf. Finally, we establish the correspondence between positive self-adjoint operators of norm one and states on the tensor product algebra.

Linearity. Fix a context $V_1 \in \mathcal{V}(\mathcal{N}_1)$ and consider the corresponding partial order of contexts under inclusion inherited from $\mathcal{V}_{1\&2} = \mathcal{V}(\mathcal{N}_1) \times \mathcal{V}(\mathcal{N}_2)$ by restriction,

$$\mathcal{V}_{1\&2}(V_1) := \{V_1 \times V_2 \mid V_2 \in \mathcal{V}(\mathcal{N}_2)\} .$$

In every context $V = V_1 \times V_2 \in \mathcal{V}_{1\&2}$, the probability measure $\mu_V^\gamma \in \underline{\Pi}(\mathcal{V}_{1\&2})(V)$ corresponding to the global section γ can be written in terms of conditional probabilities as follows:

$$\forall p \in \mathcal{P}(V_1), q \in \mathcal{P}(V_2):\ \mu_V^\gamma(p,q) = \mu_{V_1}^\gamma(p)\mu_{V_2}^\gamma(q \mid p) = \mu_{V_1}^\gamma(p)\gamma_2^p(q) = \mu_{V_1}^\gamma(p)\sigma_2^p(q) \tag{2.28}$$

Here, $(\mu_{V_2}^\gamma(\_ \mid p))_{V_2 \in \mathcal{V}(\mathcal{N}_2)} =: \gamma_2^p \in \Gamma(\underline{\Pi}(\mathcal{V}_{1\&2}(V_1)))$ is a global section of the probabilistic presheaf $\underline{\Pi}(\mathcal{V}_{1\&2}(V_1))$, which also depends on $p \in \mathcal{P}(V_1)$. Since $\mathcal{V}_{1\&2}(V_1) \cong \mathcal{V}(\mathcal{N}_2)$, by the generalised version of Gleason's theorem, Thm. 22, global sections correspond with quantum states $\Gamma(\underline{\Pi}(\mathcal{V}(\mathcal{N}_2))) \cong \mathcal{S}(\mathcal{N}_2)$, in particular, γ_2^p corresponds with a state $\sigma_2^p \in \mathcal{N}_2$ (dependent on p). As $V_1 \in \mathcal{V}(\mathcal{N}_1)$ was arbitrary, Eq. (2.28) holds for all $V \in \mathcal{V}_{1\&2}$ and, hence, for all $p \in \mathcal{P}(\mathcal{N}_1)$.

Let $p = p_1 \vee p_2 = p_1 + p_2$ with $p_1, p_2 \in \mathcal{P}(\mathcal{N}_1)$ orthogonal, i.e., $p_1 p_2 = 0$. As γ is additive,

$$\mu_{V_1}^\gamma(p)\sigma_2^p = \mu_{V_1}^\gamma(p_1)\sigma_2^{p_1} + \mu_{V_1}^\gamma(p_2)\sigma_2^{p_2} .$$

It follows that the map $\varrho^\gamma(p) := \mu_{V_1}^\gamma(p)\sigma_2^p$ satisfies,

$$\varrho^\gamma(p) = \varrho^\gamma(p_1) + \varrho^\gamma(p_2)$$

for $p = p_1 \vee p_2 = p_1 + p_2$ with $p_1, p_2 \in \mathcal{P}(\mathcal{N}_1)$ orthogonal. Note that while $\mathcal{S}_2 := \mathcal{S}(\mathcal{N}_2)$ is a Banach space (as a closed subspace of the continuous dual of $\mathcal{N}_2$) it is not a priori clear that the image $\mathcal{S}_2'$ under $\varrho^\gamma : \mathcal{P}(\mathcal{N}_1) \to \mathcal{S}_2'$ is a Banach space. Consider therefore $\mu_{V_1}^\gamma(p_k)\gamma_2^{p_k}$ in $\mathcal{S}_2'$ where $p_k \to p$ in the weak operator topology on $\mathcal{N}_1$. Note that by symmetry of Eq. (2.28) and for $q \in \mathcal{P}(\mathcal{N}_2)$ arbitrary, we have the identity

$$\mu^\gamma_{V_1}(p_k)\sigma_2^{p_k}(q) = \sigma_1^q(p_k)\mu^\gamma_{V_2}(q) \ ,$$

where the right hand side converges to $\sigma_1^q(p)\mu^\gamma_{V_2}(q)$ with $\sigma_1^q \in \mathcal{S}_1 := \mathcal{S}(\mathcal{N}_1) \simeq \Gamma(\underline{\Pi}(\mathcal{V}(\mathcal{N}_1)))$ as $\mathcal{S}_1$ is a Banach space. It follows that $\mathcal{S}_2'$ is also a Banach space and, by definition, $\mathcal{S}_2'$ only contains bounded operators such that ϱ^γ is a finitely additive, $\mathcal{S}_2'$-valued measure on $\mathcal{P}(\mathcal{N}_1)$. By Thm. 23 ϱ^γ thus uniquely extends to a bounded linear operator $\phi_\gamma : \mathcal{N}_1 \to \mathcal{S}_2'$. Equivalently, we can understand this as a bounded linear operator $\sigma^\gamma : \mathcal{N}_1 \odot \mathcal{N}_2 \to \mathbb{C}$ with $\odot$ the algebraic tensor product by setting

$$\sigma^\gamma(\sum_{i=1}^n \lambda_i a_i \odot b_i) = \sum_{i=1}^n \lambda_i \phi_\gamma(a_i)(b_i), \quad \lambda_i \in \mathbb{C}, n \in \mathbb{N} \ . \tag{2.29}$$

We collect some important properties of σ^γ. First, σ^γ is a bounded linear functional on $\mathcal{N}_1 \odot \mathcal{N}_2$. Second, summing over mutually orthogonal projections $p_k \in \mathcal{P}(\mathcal{N}_1)$ with $\bigvee_k p_k = 1$ on the first subsystem, by additivity of ϕ_γ we obtain

$$\sum_k \sigma^\gamma(p_k \odot b) = \sum_k \phi_\gamma(p_k)(b) = \sum_k \mu^\gamma_{V_1}(p_k)\sigma_2^{p_k}(b) = \mu^\gamma_{V_1}(1)\sigma_2^1(b) = \sigma_2^1(b) =: \sigma_2(b) \ .$$

Hence, for $\sigma_2(b) \neq 0$ finite (complete) additivity implies $\sigma_1^\gamma(\_ \mid b) := \frac{\phi_\gamma(\_)(b)}{\sigma_2^1(b)} \in \mathcal{S}_1 \simeq \Gamma(\underline{\Pi}(\mathcal{V}(\mathcal{N}_1)))$, and thus formally, $\sigma^\gamma(a,b) = \sigma_1^\gamma(a \mid b) \cdot \sigma_2^\gamma(b)$. It follows immediately that σ^γ is normalised since σ_1^γ and σ_2^γ are; alternatively, $\sigma^\gamma(1) = \phi_\gamma(1)(1) = \mu^\gamma_{V_1}(1)\sigma_2^1(1) = 1$ since $\mu^\gamma_{V_1}$ and σ_2^1 are normalised.

Topology and tensor product. We need to show that σ^γ extends to a bounded linear functional on the spatial tensor product $\mathcal{N}_1 \bar{\otimes} \mathcal{N}_2$. Clearly, σ^γ can be extended to a linear functional on $\mathcal{N}_1 \otimes \mathcal{N}_2$ where $\otimes$ denotes any topological tensor product between Banach spaces (cf. [71]). However, in general this extension is not unique, i.e., given any cross norm on $\mathcal{B}(\mathcal{H})$ and a (faithful) $*$-representation $\pi : \mathcal{N}_1 \odot \mathcal{N}_2 \to \mathcal{B}(\mathcal{H})$ on some Hilbert space, there can be more than one linear functional on $\mathcal{N}_1 \otimes \mathcal{N}_2$, which restricts to σ^γ. In this case the bijective correspondence between states and global sections is lost.

However, this is not the case for the spatial tensor product $\mathcal{N}_1 \bar{\otimes} \mathcal{N}_2 \subseteq \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2)$, where $\mathcal{H}_1$, $\mathcal{H}_2$ denote Hilbert spaces for which there exist faithful $*$-representations $\pi_i : \mathcal{N}_i \to \mathcal{B}(\mathcal{H}_i)$. To see this, note that the operator σ^γ can be written as a linear combination of states $\sigma_1 \in \mathcal{S}_1$, $\sigma_2 \in \mathcal{S}_2$, e.g. by expanding $a = \sum_j a_j p_j \in \mathcal{N}_1$, $b = \sum_k b_k q_k$ in $\sigma^\gamma(a,b) = \sum_{jk} a_j b_k \sigma_1^\gamma(p_j \mid q_k) \cdot \sigma_2^\gamma(q_k)$. Furthermore, the spatial tensor product has the property that every product of states has a unique extension to a state on the tensor product $\sigma_{1\&2} = \sigma_1 \cdot \sigma_2$. (This is simply a consequence of the fact that the tensor product on Hilbert spaces is unique up to isomorphism.) ϕ_γ thus uniquely extends to a linear functional on the spatial tensor product.

Now note that a linear functional ϕ on a C^*-algebra is positive if and only if it is bounded and $||\phi|| = \phi(1)$, Thm. 1. Hence, if σ^γ is positive it follows that σ^γ is a state on the spatial tensor product $\mathcal{N}_1 \bar{\otimes} \mathcal{N}_2$. In the remainder of the proof we will thus show positivity of $\sigma^\gamma : \mathcal{N}_1 \bar{\otimes} \mathcal{N}_2 \to \mathbb{C}$ by proving complete positivity of the related map $\phi_\gamma : \mathcal{N}_1 \to \widetilde{S}_2$ and the correspondence between these maps through Thm. 41 and Thm. 42.

GNS Representation. To this end it will be useful to work with a $*$-representation of $\mathcal{N}_2$. Note that every C^*-algebra has a faithful representation in the bounded operators of some Hilbert space $\mathcal{B}(\mathcal{H}_2)$ by Thm. 5. Moreover, every von Neumann algebra possesses a faithful normal, semi-finite weight w_2 and thus allows to construct a $*$-representation $\pi^{w_2} : \mathcal{N}_2 \to \mathcal{B}(\mathcal{H}_2)$ using the Gelfand-Naimark-Segal construction. The latter defines a Hilbert space by completion of the inner product $(a,b) = w_2(b^* a)$ for all $a, b \in \mathcal{N}_2^\omega := \{x \in \mathcal{N}_2 \mid \omega_2(x^* x) < \infty\}$. We will use any such faithful $*$-representation to translate between states and operators by means of the Riesz-Fréchet theorem, Thm. 6: let $\phi \in \mathcal{H}_2^*$, then there exists $y \in \mathcal{H}_2$ such that $\phi(x) = (x,y)$ for all $x \in \mathcal{H}_2$. Moreover, the map sending $y \in \mathcal{H}_2$ to ϕ_y is an isometric isomorphism. Based on this identification we will use states and operators interchangeably in what follows.[40] In particular, note that there exists a unique bounded self-adjoint operator $\rho^\gamma = \sum_k c_k(\rho_{1,k}^\gamma \otimes \rho_{2,k}^\gamma) \in \mathcal{N}_1 \bar{\otimes} \mathcal{N}_2$, $c_k \in \mathbb{C}$ such that $\sigma^\gamma(a,b) = (a \bar{\otimes} b, \rho^\gamma) = \omega_{1\&2}((\rho^\gamma)^* \cdot (a \bar{\otimes} b)) = \omega_{1\&2}(\rho^\gamma \cdot (a \bar{\otimes} b)) = \sum_k c_k \, \omega_1(\rho_{1,k}^\gamma \cdot a) \cdot \omega_2(\rho_{2,k}^\gamma \cdot b)$ for all $a \in \mathcal{N}_1$, $b \in \mathcal{N}_2$. In finite dimensions, ρ^γ is the operator t in Thm. 37.

[40] In a slight abuse of notation we will also use the same symbol to denote maps $g : D \to \mathcal{B}(\mathcal{H})$ resulting from a corresponding map $g : D \to \mathcal{N}$ and a faithful representation $\pi : \mathcal{N} \to \mathcal{B}(\mathcal{H})$ (i.e. '$g = \pi \circ g$').

Complete positivity of ϕ_γ. With these identifications, it is easily seen that $\varrho^\gamma : \mathcal{P}(\mathcal{N}_1) \to \mathcal{S}_2'$ corresponds to a map $\varrho^\gamma : \mathcal{P}(\mathcal{N}_1) \to \mathcal{B}(\mathcal{H}_2)_+$. Moreover, by Def. 27 $\varrho^\gamma|_{\mathcal{P}(V)} = v^*\varphi_V v$ for $v : \mathcal{H}_2 \to \mathcal{K}$ and φ_V an embedding (spectral measure) in every context $V \in \mathcal{V}(\mathcal{N})$, in particular, it defines a global section of $\underline{\Pi}^{\mathcal{H}_2}(\mathcal{V}(\mathcal{N}_1))$. By Thm. 33 ϕ_γ is thus not only positive but also decomposable, i.e., $\phi_\gamma = v^*\Phi_\gamma v$ for $\Phi_\gamma : \mathcal{N}_1 \to \mathcal{B}(\mathcal{K})$ a Jordan $*$-homomorphism. Finally, since γ is orientation-preserving with respect to the dynamical correspondences on $\mathcal{N}_1$ and $\mathcal{N}_2$ (respectively, $\Phi_\gamma(\mathcal{N}_1) \subseteq \mathcal{B}(\mathcal{K})$), Φ_γ extends to a $*$-homomorphism by Thm. 17.

We are left to show that the orientation on $\mathcal{N}_2$ fixes the orientation on the von Neumann algebra $\mathcal{N}_2^v$, for which $\Phi_\gamma(\mathcal{N}_1) \subseteq \mathcal{N}_2^v \subseteq \mathcal{B}(\mathcal{K})$ and which restricts to $\mathcal{N}_2$ under v, i.e., $\mathcal{N}_2 = v^*\mathcal{N}_2^v v$. To see this, note first that the argument reduces to factors since we can apply Thm. 33 to each factor $p\mathcal{N}_2 p$ independently, where $p \in \mathcal{ZP}(\mathcal{N}_2)$ is a central projection in $\mathcal{N}_2$. v therefore preserves the factor decomposition on $\mathcal{N}_2$ by construction. Moreover, $\mathcal{B}(\mathcal{H}_2) \supseteq \mathcal{N}_2 \hookrightarrow \mathcal{N}_2^v \subseteq \mathcal{B}(\mathcal{K})$ is an embedding in every factor $p\mathcal{N}_2 p$, $p \in \mathcal{ZP}(\mathcal{N}_2)$, since the linear operator v from Thm. 33 is the projection onto the subspace corresponding to $\mathcal{H}_2$ in $\mathcal{K}$. From this it follows that the dynamical correspondence on $\mathcal{N}_2$ determines a unique dynamical correspondence on $\mathcal{N}_2^v \subseteq \mathcal{B}(\mathcal{K})$.

Positivity of σ^γ. We need to relate complete positivity of ϕ_γ to positivity of σ^γ. In finite dimensions this correspondence is established by Thm. 41. Explicitly, let $\mathcal{N}_1 = \mathcal{B}(\mathbb{C}^n) = M_n(\mathbb{C})$, $\mathcal{N}_2 = \mathcal{B}(\mathbb{C}^m) = M_m(\mathbb{C})$. Since ϕ_γ is completely positive, ρ_{ϕ_γ} is positive by Thm. 41, where $(\rho_{\phi_\gamma})_{ij} = \phi_\gamma(E_{ij})$ is the image of ϕ_γ under the Choi-Jamiołkowski isomorphism (cf. Thm. 41). Finally, by the correspondence between states and density matrices in finite dimensions, we have $(\rho \geq 0, \mathrm{tr}(\rho) = 1\colon \mathcal{N}_1 \otimes \mathcal{N}_2 \ni \rho) \leftrightarrow (\mathrm{tr}(\rho \cdot \_) \in \mathcal{S}(\mathcal{N}_1 \otimes \mathcal{N}_2))$, in particular, $\rho_{\phi_\gamma} \leftrightarrow \sigma^\gamma := \mathrm{tr}(\rho_{\phi_\gamma} \cdot \_)$. In summary, positivity of σ^γ thus follows from complete positivity of ϕ_γ. This proves Thm. 43 for finite-dimensional von Neumann algebras (with no summand of type I_2).

In infinite dimensions, the trace can be replaced by the canonical reference map given by the faithful weight $\omega_1 : \mathcal{N}_1 \to \mathbb{C}$ in the representation $\pi^{w_1} : \mathcal{N}_1 \to \mathcal{B}(\mathcal{H}_1)$. This yields a completely positive map $\phi_0 = 1_{\mathcal{N}_2}\omega_1 = v_0^*\pi_0 v_0$. Complete positivity of ϕ_0 means $\sum_{i,j=1}^n (\phi_0(a_{ij})\eta_i, \eta_j) \geq 0$ for all $\eta_i \in \mathcal{H}_2$, $n \in \mathbb{N}$ whenever $a_{ij} \in M_n(\mathcal{N}_1)_+$. Moreover, since ω_1 is faithful, $\omega_1(x) = 0$ implies $x = 0$. Hence, any sequence $(a_{ij})_m \in \mathcal{N}_1$ for which $\lim_{m\to\infty} \sum_{i,j=1}^n (\phi_0((a_{ij})_m)\eta_i, \eta_j) = 0$

necessarily converges as well, i.e., $\lim_{m\to\infty}(a_{ij})_m = 0$. (More precisely, let $a_{ij} \in M_n(\mathcal{N}_1)_+$ be a non-negative matrix. Then $\sum_{i,j=1}^n(\phi_0(a_{ij})\eta_i, \eta_j) = \sum_{i,j=1}^n(\pi_0(a_{ij})v_0\eta_i, v_0\eta_j) = 0$ for all $\eta_i \in \mathcal{H}_2$ implies $\pi_0(a_{ij}) = 0$, since π_0 is faithful, which implies $a_{ij} = 0$, since ω_1 is faithful.) But then the same also holds for any other completely positive map, i.e., $\lim_{m\to\infty}\sum_{i,j=1}^n(\phi_\gamma((a_{ij})_m)\eta_i, \eta_j) = 0$ for all $\eta_i \in \mathcal{H}_2$.

By definition ϕ_γ is thus strongly completely absolutely continuous with respect to ϕ_0 (cf. [16]), hence, Thm. 42 applies and there exists a positive operator $\rho_{\phi_\gamma} = \sum_k c_k\left(\rho^k_{1,\phi_\gamma}\bar{\otimes}\rho^k_{2,\phi_\gamma}\right)$ such that $\phi_\gamma = v_0^*\rho_{\phi_\gamma}\pi_0 v_0$. Note that this operator corresponds with σ^γ in Eq. (2.29),

$$\begin{aligned}
\sigma^\gamma(a,b) &= \phi_\gamma(a)(b) \\
&= \omega_2(\phi_\gamma(a)\cdot b) \\
&= \omega_2((v_0^*\rho_{\phi_\gamma}\pi_0(a)v_0)\cdot b) \\
&= \omega_2\left(\left(v_0^*\left(\sum_k c_k\ \rho^k_{1,\phi_\gamma}\bar{\otimes}\rho^k_{2,\phi_\gamma}\right)\pi_0(a)v_0\right)\cdot b\right) \\
&= \sum_k c_k\ \omega_1(\rho^k_{1,\phi_\gamma}\cdot a)\cdot\omega_2(\rho^k_{2,\phi_\gamma}\cdot b) = (a\bar{\otimes}b, \rho_{\phi_\gamma})\ .
\end{aligned}$$

In fact, since the latter is bounded, ϕ_γ is also completely dominated by ϕ_0 (cf. [16]). In finite dimensions ρ_{ϕ_γ} in Thm. 41 is the (noncommutative) Radon-Nikodym derivative with respect to the standard trace. Succinctly, σ^γ is positive and thus the unique state associated with γ.

Normality. Finally, we highlight that the arguments in the proof work for finitely additive as well as completely additive global sections. In particular, Thm.. 33 extends to complete additivity and normal Jordan $*$-homomorphisms. Hence, σ^γ is normal whenever γ is completely additive and vice versa. This proves the theorem. □

A few remarks are in order. First, without specifying time orientations explicitly, the only information accessible on the level of contexts is the Jordan $*$-homomorphism aspect in Φ_γ, in particular, the mere context structure supports *different* time orientations. Conversely, every global section of the Bell presheaf already corresponds with a quantum state for *some* choice of time orientation.

To see this, consider the partial order of composite contexts $\mathcal{V}_{1\&2} = \mathcal{V}_1 \times \mathcal{V}_2$, for which there exist $\mathcal{N}_1$, $\mathcal{N}_2$ von Neumann algebras with no summands of type I_2 such that $\mathcal{V}_1 = \mathcal{V}(\mathcal{N}_1)$, $\mathcal{V}_2 = \mathcal{V}(\mathcal{N}_2)$. Combining Thm. 43 and Thm. 34, it is not too hard to see that every finitely (completely) additive global section of the Bell presheaf $\gamma \in \Gamma(\underline{\Pi}(\mathcal{V}_{1\&2}))$ corresponds with a (normal) state $\sigma^\gamma \in \mathcal{S}_{1\&2}$ on the spatial tensor product algebra $\widetilde{\mathcal{N}_1}\bar{\otimes}\widetilde{\mathcal{N}_2}$ for some von Neumann algebras $\mathcal{J}(\widetilde{\mathcal{N}_1}) \subseteq \mathcal{J}(\mathcal{N}_1)$ and $\mathcal{J}(\widetilde{\mathcal{N}_2}) \simeq \mathcal{J}(\mathcal{N}_2)$. Namely, we may choose an orientation on $\mathcal{J}(\mathcal{N}_2)$ according to the orientation on $\mathcal{B}(\mathcal{K})$. By a similar argument as in Thm. 34, we can then 'pull back' this orientation to $\mathcal{J}(\mathcal{N}_1)$ by Kadison's theorem, Thm. 28, which lifts the Jordan $*$-homomorphism $\Phi : \mathcal{J}(\widetilde{\mathcal{N}_1}) \to \mathcal{J}(\widetilde{\mathcal{N}_2})$ to a $*$-homomorphism $\Phi : \widetilde{\mathcal{N}_1} \to \widetilde{\mathcal{N}_2}$. Hence, for every global section there *exists* (at least) one corresponding completely positive map $\phi_\gamma = v^*\Phi v$, which it turn corresponds with a state on the respective von Neumann algebras $\widetilde{\mathcal{N}_1}, \widetilde{\mathcal{N}_2}$ by the same arguments as in Thm. 43. (Yet, unlike in Thm. 34, there is no longer a canonical choice of orientation on $\mathcal{B}(\mathcal{K})$.)

Second, we note that ϕ_γ is completely positive *only* with respect to certain dynamical correspondences and thus von Neumann algebras, namely those for which Φ preserves commutators. In particular, one should not expect ϕ_γ to be completely positive with respect to *any* choice of dynamical correspondences $\psi_{\mathcal{N}_1}$ and $\psi_{\mathcal{N}_2}$ for the following reason: let $\phi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ be a decomposable map, i.e., there exists a Hilbert space $\mathcal{K}$, a bounded linear operator $v : \mathcal{H} \to \mathcal{K}$, and a Jordan $*$-homomorphism Φ such that $\phi = v^*\Phi v$. By Thm. 32 this is equivalent to the condition that for every matrix with $x_{ij}, x_{ji} \in M_n(\mathcal{A})_+$ also $\phi(x_{ij}) \in M_n(\mathcal{B}(\mathcal{H}))_+$. On the other hand, every Jordan $*$-isomorphism is the sum of a $*$-isomorphism and a $*$-anti-isomorphism $\Phi = \overrightarrow{\Phi} + \overleftarrow{\Phi}$ by Thm. 28. In particular, we have the following:

$$\sum_{ij}^{n}(\phi(x_{ij})x_i, x_j) = \sum_{ij}^{n}(\Phi(x_{ij})vx_i, vx_j) = \sum_{ik}^{n}(\overrightarrow{\Phi}(x_{ij})vx_i, vx_j) + \sum_{ik}^{n}(\overleftarrow{\Phi}(x_{ij})vx_i, vx_j) \geq 0$$

Since $\overrightarrow{\Phi}$, $\overleftarrow{\Phi}$ are (anti-)$*$-isomorphisms, by Stinespring's theorem the maps $\overrightarrow{\phi} = v^*\overrightarrow{\Phi}v$ and $\overleftarrow{\phi} = v^*\overleftarrow{\Phi}v$ satisfy the following conditions: $x_{ij} \in M_n(\mathcal{A})_+$ implies $\overrightarrow{\phi}(x_{ij}) \in M_n(\mathcal{B}(\mathcal{H}))_+$ and $x_{ji} \in M_n(\mathcal{A})_+$ implies $\overleftarrow{\phi}(x_{ji}) \in M_n(\mathcal{B}(\mathcal{H}))_+$. Now let t be the partial transpose on $M_n(\mathcal{A})$, then $t(x_{ij}) = x_{ji}$. Generally, $t(x_{ij}) \notin M_n(\mathcal{A})_+$ for $x_{ij} \in M_n(\mathcal{A})_+$ and thus also

$\sum_{ij}(\overleftarrow{\phi}(x_{ij})x_i, x_j) = \sum_{ij}(\overleftarrow{\phi}(t(x_{ji}))x_i, x_j) \ngeq 0$. We therefore cannot conclude that $x_{ij} \in M_n(\mathcal{A})_+$ implies $\sum_{ij}^n(\phi(x_{ij})x_i, x_j) \geq 0$. Noting that the partial transpose effectively changes the local time orientation on $M_n(\mathcal{A})$, we find that $\phi = \overrightarrow{\phi} + \overleftarrow{\phi}$ will generally not be completely positive with respect to any dynamical correspondence. Note that this argument applies in particular to the matrix $\rho_\phi = \sum_{ij} E_{ij} \otimes \phi(E_{ij}) \in M_{mn}(\mathbb{C})$ (cf. Thm. 41), and thus provides a link with entanglement in the famous Peres-Horodecki criterion [84].

We also point out the interesting fact that while individual contexts arise from the 'timeless' principle of simultaneous measurability, the order relations between contexts encode time directions, which are distinguished by certain entangled states. A closer look at the interplay between entanglement and time orientations will be postponed for later study.

Finally, we combine Thm. 43 and our earlier results to arrive at the following reformulation of Bell's theorem.

Theorem 44. (Bell's theorem in contextual form) *Let the observables of a physical system be represented by von Neumann algebras $\mathcal{N}_1$, $\mathcal{N}_2$ with no summand of type I_2. Then the local state spaces $\mathcal{S}_i = \Gamma(\underline{\Pi}(\mathcal{V}(\mathcal{N}_i)))$, $i = 1, 2$ with $\underline{\Pi}$ in Def. 27 compose as*

$$\mathcal{S}(\mathcal{N}_1 \bar{\otimes} \mathcal{N}_2) \simeq \Gamma(\underline{\Pi}(\widetilde{\mathcal{V}(\mathcal{N}_1)} \times \widetilde{\mathcal{V}(\mathcal{N}_2)})) \; .$$

Moreover, for $\mathcal{N}_i$ abelian the pure state spaces $\Sigma_i = \Sigma(\mathcal{N}_i) \simeq \Gamma(\underline{\Sigma}(\mathcal{V}(\mathcal{N}_i)))$, $i = 1, 2$ compose as

$$\Sigma_1 \times \Sigma_2 \simeq \Gamma(\underline{\Sigma}(\mathcal{V}(\mathcal{N}_1) \times \mathcal{V}(\mathcal{N}_2))) \; .$$

Proof. The first assertion is Thm. 43. The second assertion follows from the derivation of Bell's inequality in Sec. 2.4.1 (see also Prop. 8 below). □

Thm. 44 subsumes the different types of composition in classical and quantum theory into a single type of contextual composition. Furthermore, it incorporates the bounds on correlations—such as those in Eq. (2.10) arising from the corresponding locality conditions in the form of factorisability and no-signalling—in terms of the allowed state spaces. As a consequence, we obtain an immediate explanation for the bound on the right-hand side of Eq. (2.10), which can be interpreted as a bound on correlations from generalised Bell inequalities based on the assumption of no-signalling and context composition.

Bell's original theorem can be understood as a consequence of a classical state space picture with composition defined by the Cartesian product on the level of its pure state spaces. This famously breaks down in quantum theory, which violates the inequalities bounding correlations in such models. The same classical state spaces also arise from composing contexts (cf. Eq. (2.21)), yet the context order is trivial and easily remains unnoticed in classical theories.

Defining composition via contexts rather than states provides a unified notion for both classical and quantum theories. This is surprising since the product of contexts $\mathcal{V}_{1\&2} = \mathcal{V}_1 \times \mathcal{V}_2$ means a substantial reduction in the number of contexts compared to $\mathcal{V}(\mathcal{H}_1 \otimes \mathcal{H}_2)$. The latter constrains global sections of the corresponding dilated probabilistic presheaf $\underline{\Pi}(\mathcal{V}(\mathcal{H}_1 \otimes \mathcal{H}_2))$ to quantum states by Thm. 24, yet the Bell presheaf $\underline{\Pi}(\mathcal{V}_{1\&2})$ is only a sub-presheaf of the former and it is a priori not clear whether general non-signalling distributions in the form of global sections $\gamma \in \Gamma(\underline{\Pi}(\mathcal{V}_{1\&2}))$ contain super-quantum correlations. Remarkably, Thm. 43 still constrains global sections of the Bell presheaf to quantum states under the additional consistency condition of being orientation-preserving. In fact, all global sections correspond with quantum states for a suitable von Neumann algebra.

Note also that this sheds further light on a previous result in [15], where it was observed that 'locally quantum', non-signalling correlations do not exceed quantum correlations. This immediately follows for global sections of the Bell presheaf, which always correspond with quantum states for suitably chosen time orientations in subsystems.

A similar result was also reported by Colbeck and Renner in [37]. There it is shown that quantum theory is complete in the sense that no theory can contain more information if it agrees with quantum mechanical predictions and obeys a notion of free choice. The latter is closely related to our notion of composition, in particular, it also implies the no-signalling constraints in Eq. (2.25). Consistent probability assignments thus arise as global sections of the Bell presheaf, which are classified by Thm. 43.

The correspondence between the notions of locality inherent to factorisability and no-signalling, composition, and contextuality carries even further. Indeed, one might argue that considering single-context theories is too restrictive to acknowledge the power of Bell's theorem. In fact, we will find that factorisability not only corresponds to single-context state spaces but to all theories with 'classical' microstates in the form of global sections of a generalised spectral presheaf. We will address this in more detail in Sec 2.5.

2.5 General non-signalling theories

The setting of von Neumann algebras in the definition of the context category and presheaves defined over it provides a framework general enough to encompass both classical and quantum theories. One might nevertheless be interested in studying the principle of physical contextuality in more general scenarios. For instance, one could argue to give up the algebraic structure in von Neumann algebras globally, yet model every context by an abelian von Neumann algebra, which suggests to study Hyperstonean spaces (cf. Thm. 10). This approach seems especially interesting from the perspective that local structure is the only one directly accessible and, together with the nesting relations between contexts, proves to be behind many results of the theory as shown in Sec. 2.3 and Sec. 2.4. To accommodate for this, in the next section we generalise the framework of presheaves over the context category considerably by allowing for (Hyperstonean) orthomodular lattices (cf. Sec. 2.5.1).

In particular, we are interested in correlations in composite systems, for which we keep the notion of context composition in Eq. (2.24). Recall that for von Neumann algebras this leads to a genuinely different context structure for composite systems than given by the tensor product. There are by far more contexts and more restriction maps in $\underline{\Pi}(\mathcal{V}(\mathcal{H}_1 \otimes \mathcal{H}_2))$ than in $\underline{\Pi}(\mathcal{V}(\mathcal{H}_1) \times \mathcal{V}(\mathcal{H}_2))$, and global sections of the latter probabilistic presheaf are thus a priori less constrained than those of the former, which correspond with quantum states by Gleason's theorem in contextual form, Thm. 24. Remarkably, Thm. 44 shows that global sections of the Bell presheaf essentially correspond with quantum states for von Neumann algebras (with no summand of type I_2), which suggests context composition as an alternative for the tensor product. Clearly, for general orthomodular lattices the tensor product is no longer defined. Yet, as we will see below, the definition of the probabilistic presheaf over product contexts can be generalised and some aspects of Thm. 44 still apply in this case.

Note that alternative context structures have also been studied elsewhere [107, 133]. In [81] it was shown that a type of nonlocality persists under relaxation of the (non-contextual) identification of non-maximal operators, the consequences for the underlying context structure were explored in [63]. Other types of contextual structures have also been discussed in [26, 134].

2.5.1 Presheaves over orthomodular lattices

In Def. 25 and Def. 26 we introduced the spectral and the probabilistic presheaf for von Neumann algebras. [28] extends the former definition to orthomodular lattices and proves a generalised Stone duality for complete orthomodular lattices. We review the main definitions and results and extend them to the probabilistic presheaf.

In analogy to $\mathcal{V}(\mathcal{N})$ one defines a context category for orthomodular lattices (cf. Def. 13).

Definition 32. *For an orthomodular lattice L, let $\mathcal{B}(L)$ denote the* partial order of Boolean sublattices of L, *where the partial order on $\mathcal{B}(L)$ is given by inclusion. $\mathcal{B}(L)$ is also called the* context category *of L.*

It is straightforward to see that this lifts to a functor [28].

Proposition 7. *There is a functor $\mathcal{B}$:* **OML** $\rightarrow$ **Pos** *sending each orthomodular lattice L to its context category $\mathcal{B}(L)$ and each homomorphism $\varphi : L \rightarrow M$ of orthomodular lattices to the corresponding morphism $\phi : \mathcal{B}(L) \rightarrow \mathcal{B}(M)$.*

Recall that the spectral presheaf $\underline{\Sigma}$ of $\mathcal{N}$ over $\mathcal{V}(\mathcal{N})$ maps contexts into their Gelfand spectra Σ_V locally. The relevant duality underlying this functorial mapping is Thm. 10, which applies to the subcategory **HStonean** of **Stonean** in Thm. 8.

Given a general orthomodular lattice $L \in$ **OML** with context category $\mathcal{B}(L)$ consisting of Boolean algebras, one cannot apply Gelfand duality. Nevertheless, one may still assign a context its Stone space. Accordingly, [28] generalise the spectral presheaf to orthomodular lattices.

Definition 33. *Let L be an orthomodular lattice with context category $\mathcal{B}(L)$. The* spectral presheaf $\underline{\Sigma}$ of L over $\mathcal{B}(L)$ *is the presheaf given*

(i) on objects: for all $B \in \mathrm{Ob}(\mathcal{B}(L))$, let $\underline{\Sigma}_B := \Omega(B)$, the Stone space of B.

Here, $\underline{\Sigma}_B$ denotes the component of $\underline{\Sigma}(L)$ at B.

(ii) on arrows: for all $B, \tilde{B} \in \mathcal{B}(L)$, if $\tilde{B} \subseteq B$, let $\underline{\Sigma}(i_{\tilde{B}B}) : \underline{\Sigma}_B \longrightarrow \underline{\Sigma}_{\tilde{B}}$ with $\lambda \longmapsto \lambda|_{\tilde{B}}$.

Here, $\lambda|_{\tilde{B}}$ denotes the restriction of λ to the subalgebra $\tilde{B}$.

Similarly, we would like to extend the notion of probabilistic presheaf to orthomodular lattices. Note that for the spectral presheaf $\underline{\Sigma}(L)$ over the orthomodular lattice $L \in \mathbf{OML}$ we only give up the structure *between contexts*, whereas the structure *within contexts* remains the same by means of Stone duality. However, the same is no longer true for the probabilistic presheaf over general orthomodular lattices, since Boolean sublattices do not necessarily correspond with commutative von Neumann algebras.

In order to enforce the structure of commutative von Neumann algebras *within contexts* we need to lift Stone duality to Thm. 10, i.e., restrict to Hyperstonean spaces. The latter are in particular Stonean spaces and thus correspond with complete Boolean algebras by Thm. 8. It is therefore natural to restrict to complete orthomodular lattices. Still, not every Stonean space corresponds with a von Neumann algebra, one needs to further restrict to Hyperstonean spaces. Correspondingly, we will call an orthomodular lattice $L \in \mathbf{OML}$ *Hyperstonean* if its Boolean subalgebras correspond not only to Stonean but to Hyperstonean spaces, and write **HOML** for the category of Hyperstonean orthomodular lattices (with suitable morphisms, preserving this extra structure). With this definition, the following generalisation of $\underline{\Pi}$ from $\mathcal{V}(\mathcal{N})$ to Hyperstonean orthomodular lattices is again only concerned with the constraints *between contexts*.

Definition 34. *Let $L \in \mathbf{HOML}$ be a Hyperstonean orthomodular lattice with context category $\mathcal{B}(L)$. The* (normal) probabilistic presheaf $\underline{\Pi}$ of L over $\mathcal{B}(L)$ *is the presheaf given*

(i) on objects: for all $V \in \mathcal{B}(L)$,[41] *let*

$$\underline{\Pi}_V := \{\mu_V : \mathcal{P}(V) \to [0,1] \mid \mu_V \textit{ is a finitely (completely) additive probability measure}\} ,$$

(ii) on arrows: for all $V, \tilde{V} \in \mathcal{B}(L)$, if $\tilde{V} \subseteq V$, let $\underline{\Pi}(i_{\tilde{V}V}) : \underline{\Pi}_V \longrightarrow \underline{\Pi}_{\tilde{V}}$ with $\lambda \longmapsto \lambda|_{\tilde{V}}$.

Analogously, we can also generalise the definition of the dilated probabilistic presheaf to $L \in \mathbf{HOML}$. Note that while global sections of the probabilistic presheaf $\underline{\Pi}(\mathcal{V}(\mathcal{N}))$ correspond to quantum states by Thm. 24, for $L \in \mathbf{HOML}$, $\Gamma(\underline{\Pi}(\mathcal{B}(L)))$ might be the empty set (cf. [67]).

[41] We write $V \in \mathcal{B}(L)$ for $L \in \mathbf{HOML}$ to indicate that, as with $\mathcal{V}(\mathcal{N})$ for $\mathcal{N}$ a von Neumann algebra, contexts correspond with commutative von Neumann algebras since L is Hyperstonean.

Importantly, [28] shows that the spectral presheaf is a complete invariant of the orthomodular lattice, i.e., two orthomodular lattices L_1, L_2 are isomorphic if and only if their respective spectral presheaves are: $L_1 \sim L_2 \Leftrightarrow \underline{\Sigma}(\mathcal{B}(L_1)) \sim \underline{\Sigma}(\mathcal{B}(L_2))$. Since the probabilistic presheaf $\underline{\Pi}(\mathcal{B}(L))$ over $L \in \mathbf{HOML}$ contains more information than the spectral presheaf $\underline{\Sigma}(\mathcal{B}(L))$ and $\mathbf{HOML}$ is a subcategory of $\mathbf{OML}$, the same also holds for the former: $L_1 \sim L_2 \Leftrightarrow \underline{\Pi}(\mathcal{B}(L_1)) \sim \underline{\Pi}(\mathcal{B}(L_2))$.

2.5.2 Generalised classical state spaces

In this section we explore the consequences of context composition for the (generalised) spectral presheaf in Def. 33, as well as for the corresponding (classical) state spaces consisting of its global sections. Recall that Bell's theorem rules out factorisable hidden variable models, in particular, those defined in Eq. (2.22). In fact, the Cartesian product construction in Eq. (2.22) and factorisability are natural also for the generalised spectral presheaf. This follows as global sections of the spectral presheaf over the composite context category $\mathcal{B}_{1\&2} := \mathcal{B}(L_1) \times \mathcal{B}(L_2)$ for $L_1, L_2 \in \mathbf{OML}$ define valuation functions, which are easily seen to compose classically.

Proposition 8. *Let $L_1, L_2 \in \mathbf{OML}$ be orthomodular lattices with respective context categories $\mathcal{B}(L_1)$, $\mathcal{B}(L_2)$ and spectral presheaves $\underline{\Sigma}(\mathcal{B}(L_1))$, $\underline{\Sigma}(\mathcal{B}(L_2))$. For global sections of the spectral presheaf over the composite context category the following correspondence holds,*

$$\Gamma(\underline{\Sigma}(\mathcal{B}_{1\&2})) \simeq \Gamma(\underline{\Sigma}(\mathcal{B}(L_1))) \times \Gamma(\underline{\Sigma}(\mathcal{B}(L_2))) \; . \tag{2.30}$$

Proof. Clearly, $\gamma_1 \cdot \gamma_2 \in \Gamma(\underline{\Sigma}(\mathcal{B}_{1\&2}))$ for all $\gamma_1 \in \Gamma(\underline{\Sigma}(\mathcal{B}(L_1))$, $\gamma_2 \in \Gamma(\underline{\Sigma}(\mathcal{B}(L_2))$. Conversely, let $\gamma \in \Gamma(\underline{\Sigma}(\mathcal{B}_{1\&2}))$, then $\gamma_V = \gamma_{V_1} \cdot \gamma_{V_2}$ with $\gamma_{V_1} \in \underline{\Sigma}(\mathcal{B}(L_1))_{V_1}$, $\gamma_{V_2} \in \underline{\Sigma}(\mathcal{B}(L_2))_{V_2}$ for all contexts, $V = (V_1, V_2) \in \mathcal{B}_{1\&2}$. From this it easily follows that $\gamma_1 := (\gamma_{V_1})_{V_1 \in \mathcal{V}_1} \in \Gamma(\underline{\Sigma}(\mathcal{B}(L_1)))$ and $\gamma_2 := (\gamma_{V_2})_{V_2 \in \mathcal{V}_2} \in \Gamma(\underline{\Sigma}(\mathcal{B}(L_2)))$ (cf. proofs to Thm. 45 and Thm. 43). □

With Prop. 6 and Prop. 8, Bell's theorem provides a no-go-result for all classical state spaces of the form in Eq. (2.22), with pure states given by valuation functions, equivalently, global sections of the spectral presheaf for composite systems: the correlations in the outcome statistics are constrained by factorisability and therefore cannot account for (all) those arising in quantum mechanics.

More generally, assume that the state space of the system corresponds with the global sections of the probabilistic presheaf $\Gamma(\underline{\Pi}(\mathcal{B}(L)))$ with $L \in \mathbf{HOML}$. Note that $\Gamma(\underline{\Pi}(\mathcal{B}(L)))$ is in particular a convex set, the pure state space $\Gamma_{\mathrm{pure}}(\underline{\Pi}(\mathcal{B}(L)))$ therefore consists of all elements that cannot be written in terms of proper convex linear combinations of other elements. Clearly, the pure state space always contains global sections of the corresponding spectral presheaf $\Gamma(\underline{\Sigma}(\mathcal{B}(L))) \subseteq \Gamma_{\mathrm{pure}}(\underline{\Pi}(\mathcal{B}(L)))$. Moreover, for classical, i.e., single-context theories equality holds. By Prop. 8 the subset of the pure state space consisting of such microstates composes via the Cartesian product. The latter thus applies not only to classical (single-context), but to all theories for which $\Gamma(\underline{\Sigma}(\mathcal{B}(L))) = \Gamma_{\mathrm{pure}}(\underline{\Pi}(\mathcal{B}(L)))$. Since the latter compose via the Cartesian product, they satisfy factorisability and thus the Bell inequalities.

Clearly, this analysis breaks down for more general pure states $\gamma \in \Gamma_{\mathrm{pure}}(\underline{\Pi}(\mathcal{B}(L)))$, for which factorisability is replaced by the no-signalling marginalisation constraints. Even more drastically, this assumption breaks down if $\Gamma(\underline{\Sigma}(L_1))$, $\Gamma(\underline{\Sigma}(L_2))$ are empty. The latter is the case in quantum theory, where $L_1 = \mathcal{V}(\mathcal{N}_1)$, $L_2 = \mathcal{V}(\mathcal{N}_2)$ correspond to von Neumann algebras (with no summand of type I_2). Then we already know that valuation functions do not exist by the Kochen-Specker theorem, Thm. 19, and it is natural to interpret Bell's theorem as a consequence of the impossibility of a pure state space consisting of microstates $\gamma \in \Gamma(\underline{\Sigma}(\mathcal{B}(L)))$. In this reading Bell's theorem becomes a special case of the Kochen-Specker theorem, locality only plays a secondary role. The only exception are two-dimensional quantum systems, and by Prop. 8 composite systems with subsystems of local dimension two. In this special case, Bell's theorem may be interpreted as a stronger no-go-result than the Kochen-Specker theorem ruling out theories with state spaces arising from global sections of spectral presheaves.

The above analysis is concerned with the classical part of Bell's theorem in Thm. 44, which corresponds with the left-hand side of Eq. (2.10). In the next section we consider the right-hand side of Eq. (2.10), yet in the setting of (Hyperstonean) orthomodular lattices. As shown before, global sections of the probabilistic presheaf over product contexts are non-signalling, yet they no longer restrict to quantum states since Gleason's theorem does not apply to general orthomodular lattices. Nevertheless, a generalised statistical version survives even in this case and allows to compare correlations beyond those realised in classical and quantum theories.

2.5.3 Bayes' theorem and the Bell presheaf

In this section we comment on the implications of context composition on correlations arising from global sections of the generalised Bell presheaf[42] $\underline{\Pi}(\mathcal{B}_{1\&2})$ with $\mathcal{B}_{1\&2} := \mathcal{B}(L_1) \times \mathcal{B}(L_2)$ over (Hyperstonean) orthomodular lattices $L_1, L_2 \in \mathbf{HOML}$. We will refer to theories with state spaces given by global sections of the generalised Bell presheaf as *general non-signalling theories*.[43] The naming convention is justified since the marginalisation constraints between product contexts are equivalently encoded by no-signalling (cf. Eq. (2.25)).[44]

Recall that any joint distribution $\mu(A, B)$ over events A, B corresponding to random variables a, b satisfies a symmetric decomposition based on conditional probabilities $\mu(A \mid B)$ underlying Bayes' theorem,

$$\mu(A, B) = \mu(B \mid A) \cdot \mu(A) = \mu(A \mid B) \cdot \mu(B) \ . \tag{2.31}$$

In particular, Eq. (2.31) holds for probability distributions in every product context $V = (V_1, V_2) \in \mathcal{B}_{1\&2}$, yet there is a priori no reason why a similar decomposition should hold *simultaneously* over all contexts. For quantum theory this is guaranteed by Thm. 43, i.e., context composition and Gleason's theorem. More precisely, let ρ be a composite density matrix over Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. Then for all $p_1 \in \mathcal{P}(\mathcal{H}_1)$, $p_2 \in \mathcal{P}(\mathcal{H}_2)$ it holds

$$\rho(p_1, p_2) = \mathrm{tr}(\rho(p_1 \otimes p_2)) = \mathrm{tr}_1(\rho_1(p_2)p_1) \cdot \mathrm{tr}_2(\rho_2 p_2) =: \rho_1(p_1 \mid p_2) \cdot \rho_2(p_2) = \rho_2(p_2 \mid p_1) \cdot \rho_1(p_1) \ , \tag{2.32}$$

where $\rho_1(p_2) := \mathrm{tr}_2(\rho(1 \otimes p_2))$, $\rho_2 := \mathrm{tr}_1(\rho)$ and similarly $\rho_2(p_1) := \mathrm{tr}_1(\rho(p_1 \otimes 1))$, $\rho_1 := \mathrm{tr}_2(\rho)$.[45] Note that Eq. (2.32) immediately rules out PR-box distributions. This is essentially a different version of the argument in [15], which asserts that no-signalling and systems being 'locally quantum' restrict correlations to be quantum. What is more, Bayes' theorem becomes a type of *local-to-global* property similar to but independent of linearity in Gleason's theorem, Thm. 24.

[42] While of general interest, here we will not discuss the subtleties arising from restricting to the dilated probabilistic presheaf in the generalisation of the Bell presheaf in Def. 30 to Hyperstonean orthomodular lattices.

[43] Note that this definition is different from the definition of general probabilistic theories given, e.g. in [134].

[44] For instance, the PR-box distribution arises as a global section of the Bell presheaf over the restricted context order consisting of just four contexts corresponding to the four types of measurements Alice and Bob can perform in a CHSH experiment (see also Ex. 6 in Sec. 2.5.4).

[45] Note that $\frac{\rho_1(p_2)}{\mathrm{tr}_2(\rho_2 p_2)}$, $\frac{\rho_2(p_1)}{\mathrm{tr}_1(\rho_1 p_1)}$ are the post-measurement states after a local measurement on either subsystem.

Theorem 45. *Let $L_1, L_2 \in$ **HOML** be Hyperstonean orthomodular lattices and $\mathcal{B}(L_1)$, $\mathcal{B}(L_2)$ their respective context categories. For all $\gamma \in \Gamma_{1\&2} := \Gamma(\underline{\Pi}(\mathcal{B}_{1\&2}))$ with $\mathcal{B}_{1\&2} := \mathcal{B}(L_1) \times \mathcal{B}(L_2)$ there exist (unique) $\gamma_1 \in \Gamma(\underline{\Pi}(\mathcal{B}(L_1)))$, $\gamma_2 \in \Gamma(\underline{\Pi}(\mathcal{B}(L_2)))$ such that for all $p_1 \in L_1$, $p_2 \in L_2$:*

$$\gamma(p_1, p_2) = \gamma_1(p_1 \mid p_2) \cdot \gamma_2(p_2) = \gamma_1(p_2 \mid p_1) \cdot \gamma_2(p_1) \tag{2.33}$$

Proof. The proof is a shortened version of Thm. 43. First, fix a context $V_1 \in \mathcal{B}(L_1)$ and consider the corresponding partial order of contexts under inclusion inherited from $\mathcal{B}_{1\&2}$,

$$\mathcal{B}_{1\&2}(V_1) := \{V_1 \times V_2 \mid V_2 \in \mathcal{B}(L_2)\} \ .$$

The probability measure $\mu_V^\gamma := \gamma_V$ in the context $V = (V_1, V_2) \in \mathcal{B}_{1\&2}(V_1)$ corresponding to the global section $\gamma \in \Gamma_{1\&2}$ for $p_1 \in \mathcal{P}(V_1), p_2 \in \mathcal{P}(V_2)$ takes the form

$$\mu_V^\gamma(p_1, p_2) = \mu_{V_1}^\gamma(p_1) \cdot \mu_{V_2}^\gamma(p_2 \mid p_1) \ . \tag{2.34}$$

It follows from context composition that $\gamma_2^{p_1} := \mu_{V_2}^\gamma(\_ \mid p_1) \in \mu_{V_1}^\gamma(p_1) \cdot \Gamma(\underline{\Pi}(\mathcal{B}(L_2)))$ for all $p_1 \in \mathcal{P}(V_1)$. As $V_1 \in \mathcal{B}(L_1)$ was arbitrary, Eq. (2.34) holds for all $V \in \mathcal{B}_{1\&2}$, in particular, for all $p_1 \in L_1$. Furthermore, let $p_1 = q_1 \vee q_1'$ with $q_1, q_1' \in L_1$ orthogonal, i.e., there exists $V_1 \in \mathcal{B}(L_1)$ such that $q_1, q_1' \in \mathcal{P}(V_1)$ and $q_1 \leq q_1'^{\perp}$. As γ is finitely additive we also have

$$\mu_{V_1}^\gamma(p_1) \cdot \gamma_2^{p_1} = \mu_{V_1}^\gamma(q_1) \cdot \gamma_2^{q_1} + \mu_{V_1}^\gamma(q_1') \cdot \gamma_2^{q_1'} \ .$$

It follows that the map $\phi_\gamma : L_1 \longrightarrow \mathbb{R}_0^+ \cdot \Gamma(\underline{\Pi}(\mathcal{B}(L_2)))$, $p_1 \longmapsto \mu_{V_1}^\gamma(p_1) \cdot \gamma_2^{p_1}$ satisfies $\phi_\gamma(p_1) = \phi_\gamma(q_1) + \phi_\gamma(q_1')$ for $p_1 = q_1 \vee q_1'$ with $q_1, q_1' \in L_1$ orthogonal. Furthermore, let $\phi_{p_2}^\gamma(p_1) := \mu_{V_1}^\gamma(p_1) \cdot \gamma_2^{p_1}(p_2)$ and note that for every set of mutually orthogonal $q_1^i \in \mathcal{P}(V_1)$, $V_1 \in \mathcal{B}(L_1)$ with $\bigvee_i q_1^i = 1 \in L_1$:

$$\bigvee_i \phi_{p_2}^\gamma(q_1^i) = \mu_{V_1}^\gamma(1) \cdot \gamma_2^1(p_2) = \gamma_2^1(p_2) =: \gamma_2(p_2)$$

Hence, for $\gamma_2(p_2) \neq 0$ we have $\gamma_1(\_ \mid p_2) := \frac{\phi_{p_2}^\gamma}{\gamma_2(p_2)} \in \Gamma(\underline{\Pi}(\mathcal{B}(L_1)))$ and thus $\gamma(p_1, p_2) = \phi_\gamma(p_1)(p_2) = \phi_{p_2}^\gamma(p_1) = \gamma_1(p_1 \mid p_2) \cdot \gamma_2(p_2)$. The other direction follows by symmetry. □

Importantly, while linearity is a *local-to-global* property special to von Neumann algebras, the notion of Bayes' theorem only requires context composition in Eq. (2.24) and therefore holds in general non-signalling theories. Thm. 45 is thus a type of generalisation of the non-classical aspect of Bell's theorem in contextual form, Thm. 44, to Hyperstonean orthomodular lattices $L_1, L_2 \in \mathbf{HOML}$. It classifies the possible states on the composite system in terms of the states on subsystems. For von Neumann algebras, where $L_i = \mathcal{V}(\mathcal{N}_i)$, the latter bijectively correspond with states in $\mathcal{S}(\mathcal{N}_i)$ by Gleason's theorem, Thm. 22. Yet, for more general orthomodular lattices we do not have this additional structure, which allowed to derive the particular Bell inequalities for quantum theory such as Eq. (2.10). Instead, in the next section we consider correlations in generalised non-signalling theories in terms of global sections of the Bell presheaf directly and discuss a way to quantify correlations in such theories.

2.5.4 Correlations in general non-signalling theories

While the setting of von Neumann algebras is very rich and applies to existing classical and quantum frameworks, one might yet want to consider more general scenarios. In particular, from a foundations and information-theoretic perspective it is interesting to study correlations in general non-signalling theories, which by the definition in Sec. 2.5.1 arise by considering global sections of the Bell presheaf over Hyperstonean orthomodular lattices with composition of contexts as in Eq. (2.24). We discuss a familiar example in this setting.

Example 6. *The PR-box distribution arises as a global section of the Bell presheaf* $\gamma_{\mathrm{PR}} \in \Gamma(\underline{\Pi}(\mathcal{B}^{\mathrm{PR}}_{1\&2}))$ *over the partial order of contexts* $\mathcal{B}^{\mathrm{PR}}_{1\&2} = \mathcal{B}(L_1) \times \mathcal{B}(L_2)$*, where* $L_1 \simeq L_2$*,* $L_i = \{0, p_i, (1-p_i), q_i, (1-q_i), 1\}$ *as a set, and* $\mathcal{B}(L_i)$ *consists of contexts* $V_{p_i} = \{p_i, 1\}''$*,* $V_{q_i} = \{q_i, 1\}''$*, and* $V^0_i = \{1\}''$ *with only (non-trivial) order relations* $V^0_i \subset V_{p_i}, V_{q_i}$*.*

Clearly, the context structure in Ex. 6 is different to the context structure in $\mathcal{V}(\mathcal{N}_1 \otimes \mathcal{N}_2)$, which for noncommutative von Neumann algebras is also different to the (trivial) context structure in classical theories. We therefore seek a way to compare correlations in theories with different context categories. To this end we will define a 'distance' between global sections in the respective theories and some given reference set of probability measures.

Note first that by Thm. 35, in every context $V \in \mathcal{B}(L)$ a distance on probability measures $\mu_1, \mu_2 \in \underline{\Pi}_V$ can be defined as $\Delta(\mu_1, \mu_2) := ||\mu_1 - \mu_2||_1$, where $||\cdot||_1$ denotes the L_1-norm on the space of measurable functions on some standard measure space isomorphic to V.[46]

We want to define a similar distance also over multiple contexts and thereby compare correlations between theories based on different context structures. More precisely, we measure the distance to some given collection of reference probability measures defined over measurement outcomes of certain observables, denoted $\widetilde{\gamma}_0 = (\mu_{\tilde{V}})_{\tilde{V} \in \widetilde{\mathcal{V}}}$. Note that instead of writing the dependence on observables explicitly, we suggestively write a dependence on contexts $\tilde{V} \in \widetilde{\mathcal{V}}$. More precisely, we take every probability measure $\mu_{\tilde{V}}$ to correspond to the (smallest) context $\tilde{V}$ that contains all observables it is defined over. In particular, this implies that observables within such contexts are simultaneously measurable. However, in general this does not fix the context structure of a theory T (necessarily containing all relevant simultaneously measurable observables in the probability measures $\widetilde{\gamma}_0 = (\mu_{\tilde{V}})_{\tilde{V} \in \widetilde{\mathcal{V}}}$) uniquely. Correspondingly, there is a map $\xi_T : \widetilde{\mathcal{V}} \to \mathcal{B}(L)_T$, which embeds the abstract contexts $\tilde{V} \in \widetilde{\mathcal{V}}$ into the theory-specific context order $\mathcal{B}(L)_T$. In order to make this clear, we consider two extreme cases.

On the one hand, assume that $\widetilde{\gamma}_0$ defines a local section $\gamma_0 = (\mu_V)_{V \in \mathcal{V}}$ of $\underline{\Pi}(\mathcal{B}(L)_T)$, i.e., the map $\xi_T : \widetilde{\mathcal{V}} \to \mathcal{B}(L)_T$, which embeds observables into the theory-specific context structure $\mathcal{B}(L)_T$, is one-to-one. Then γ_0 is attained in that theory if it extends to a global section $\gamma \in \underline{\Pi}(\mathcal{B}(L)_T)$. However, not every local section arises from a global section (cf. Ex. 5), hence, in general we can only approximate γ_0. To to so we minimise the sum of distances in contexts given by the L_1-norm:

$$\Delta_T(\widetilde{\gamma}_0) := \inf_{\gamma \in \Gamma(\underline{\Pi}(\mathcal{B}(L)_T))} \sum_{V \in \mathcal{V}} \Delta(\gamma_V, \mu_V)$$

On the other hand, the theory T might assign every probability measure in $\widetilde{\gamma}_0 = (\mu_{\tilde{V}})_{\tilde{V} \in \widetilde{\mathcal{V}}}$ to the same context $V \in \mathcal{B}(L)_T$, i.e., the map $\xi_T : \widetilde{\mathcal{V}} \to \mathcal{B}(L)_T$ is many-to-one. In this case, we set

$$\Delta_T(\widetilde{\gamma}_0) = \inf_{\gamma \in \underline{\Pi}_V} \sum_{\tilde{V} \in \widetilde{\mathcal{V}}} \Delta(\gamma_V, \mu_{\xi(\tilde{V})}) \, . \tag{2.35}$$

[46]For X countable this equals the *total variation/statistical distance*, $\Delta(\mu_1, \mu_2) := \sup_{A \in \sigma(X)} |\mu_1(A) - \mu_2(A)|$.

The general case is a combination of these two, in fact, both arise from the same formula:

$$\Delta_T(\widetilde{\gamma}_0) := \inf_{\gamma\in\Gamma(\underline{\Pi}(\mathcal{B}(L)_T))} \sum_{\tilde{V}\in\widetilde{\mathcal{V}}} \Delta(\gamma_{\xi(\tilde{V})}, \mu_{\xi(\tilde{V})})$$

Note that $\Delta_T(\widetilde{\gamma}_0)$ depends on the context structure of the underlying theory T via the map $\xi_T : \widetilde{\mathcal{V}} \to \mathcal{B}(L)_T$, which embeds the implicit context structure in $\widetilde{\mathcal{V}}$ into $\mathcal{B}(L)_T$. For a given set of reference probability measures $\widetilde{\gamma}_0 = (\mu_{\tilde{V}})_{\tilde{V}\in\widetilde{\mathcal{V}}}$ we thus obtain inequalities of the form $\Delta_T(\widetilde{\gamma}_0) \leq \Delta_{T'}(\widetilde{\gamma}_0)$ by comparing theories with different context structures with respect to the values of their contextual distances to $\widetilde{\gamma}_0$. Note that we have not yet assumed observables to be composite, the argument therefore applies also to single systems and thus leads to genuine contextuality inequalities.

2.5.5 Generalised Bell inequalities

In the following we are interested in the special case of Bell inequalities, i.e., we consider composite context categories $\mathcal{B}_{1\&2}$. Moreover, Bell inequalities compare quantum with classical correlations. Classical theories contain a single (maximal) context, hence, ξ_{cl} collapses the implicit context structure $\widetilde{\mathcal{V}}$ in $\widetilde{\gamma}_0$ to a single context as in the second extreme case discussed in the last section. In order to obtain a non-trivial bound from Eq. (2.35) in this case, we necessarily need to consider probability measures $\widetilde{\gamma}_0 = (\mu_{\tilde{V}})_{\tilde{V}\in\widetilde{\mathcal{V}}}$, which are not simultaneously satisfiable by a factorisable probability distribution. The simplest such case arises by comparing probability measures over the outcomes of operators of the form $\{a\times b, a\times b', a'\times b, a'\times b'\}$:

$$\begin{aligned} \mu_{\tilde{V}_1}(A,B\mid a,b) &= \mu_{\tilde{V}_1}(A\mid a)\cdot\mu_{\tilde{V}_1}(B\mid b), & \mu_{\tilde{V}_2}(A,B'\mid a,b') &= \mu_{\tilde{V}_2}(A\mid a)\cdot\mu_{\tilde{V}_2}(B'\mid b'), \\ \mu_{\tilde{V}_3}(A',B\mid a',b) &= \mu_{\tilde{V}_3}(A'\mid a')\cdot\mu_{\tilde{V}_3}(B\mid b), & \mu_{\tilde{V}_4}(A',B'\mid a',b') &\neq \mu_{\tilde{V}_4}(A'\mid a')\cdot\mu_{\tilde{V}_4}(B'\mid b') \end{aligned} \tag{2.36}$$

Recall that every factorisable joint probability distribution is a convex mixture of product measures $\mu(A,B\mid a,b) = \mu(A\mid a)\cdot\mu(B\mid b)$, where $\mu(A\mid a)$ is the marginal distribution conditioned on local measurement settings a. Clearly, such distributions can at most satisfy three out of the four conditions above. This is precisely what the CHSH inequality measures.

Before we continue to apply our approach to the CHSH inequality explicitly, we slightly generalise this scenario in terms of the choice of $\widetilde{\gamma}_0$ as follows. Consider composite observables of the form $a_i \times b_j$, $i \in \{1, \cdots, n_1\}$, $j \in \{1, \cdots, n_2\}$, where a_i, b_j are local observables on the first, second subsystem, respectively. If we assume sufficiently many local measurement outcomes, we can choose a collection of general non-signalling probability measures (similar to the PR-box distribution) $\widetilde{\gamma}_0 = \gamma_{\mathrm{ns}}^{(n_1,n_2)} = (\mu_{\tilde{V}})_{\tilde{V} \in \widetilde{\mathcal{V}}}$, which violates a maximal number of constraints in factorisable probability distributions. More precisely, there are $(n_1 - 1)(n_2 - 1)$ product constraints of the form in Eq. (2.36), and violation of a product constraint as in Eq. (2.36) contributes 2 to the overall distance (cf. Fig. 2.4). With respect to such a collection γ_{ns}, every factorisable probability distribution is subject to the following bound:

$$\Delta_{\mathrm{cl}}(\gamma_{\mathrm{ns}}^{(n_1,n_2)}) = 2(n_1 - 1)(n_2 - 1) \tag{2.37}$$

Note that this is a non-trivial bound since $\Delta_{\mathrm{ns}}(\gamma_{\mathrm{ns}}^{(n_1,n_2)}) = 0$ and thus represents a Bell inequality. More precisely, by their very definition, the distributions in $\gamma_{\mathrm{ns}}^{(n_1,n_2)}$ maximise the distance to factorisable probability distributions, i.e., convex combinations of product measures by Eq. (2.18). In contrast, global sections in general non-signalling theories are constrained only by the generalised Bayes' theorem in Thm. 45. In principle, they are more general than factorisable distributions and may thus violate the constraint in Eq. (2.37).

We finish by considering the CHSH scenario from this viewpoint, i.e., we explicitly compute the correlation terms $\Delta_T(\gamma_{\mathrm{PR}})$ in the classical, quantum, and PR-box case for the simplest scenario, where $n_1, n_2 = 2$ and therefore $\gamma_{\mathrm{PR}} := \gamma_{\mathrm{ns}}^{(2,2)}$. For the latter the context structure is very simple and given in Ex. 6, it consists of the four maximal contexts $(\mathcal{B}_{1\&2}^{\mathrm{PR}})_{\max} = \{V_{p_1} \times V_{p_2}, V_{p_1} \times V_{q_2}, V_{q_1} \times V_{p_2}, V_{q_1} \times V_{q_2}\}$. The PR-box distribution defines a global section of the corresponding probabilistic presheaf $\gamma_{\mathrm{PR}} \in \underline{\Pi}(\mathcal{B}_{1\&2}^{\mathrm{PR}})$. Hence, $\Delta_{\mathrm{PR}}(\gamma_{\mathrm{PR}}) = 0$.

For the quantum case we have $\gamma_{\mathrm{qm}} \leftrightarrow \rho = |\psi^+\rangle\langle\psi^+|$, $\gamma_1 \leftrightarrow \mathrm{tr}_2(\rho)$, $\gamma_1(\_ \mid p_B) \leftrightarrow \mathrm{tr}_2(\rho p_B)$ and similarly, $\gamma_2 \leftrightarrow \mathrm{tr}_1(\rho)$ and $\gamma_2(\_ \mid p_A) \leftrightarrow \mathrm{tr}_1(\rho p_A)$. With the choice of parameters given at the beginning of Sec. 2.4 one computes $\Delta_{\mathrm{qm}}(\gamma_{\mathrm{PR}}) = 4 - 2\sqrt{2}$ (cf. Fig. 2.4).

$\mu_T(A,B\mid a,b)$		$A=-1$						$A=1$					
		$B=-1$			$B=1$			$B=-1$			$B=1$		
		cl	qm	ns	cl	qm	ns	cl	qm	ns	cl	qm	ns
a	b	$1/2$	$2+\sqrt{2}/8$	$1/2$	0	$2-\sqrt{2}/8$	0	0	$2-\sqrt{2}/8$	0	$1/2$	$2+\sqrt{2}/8$	$1/2$
	b'	$1/2$	$2+\sqrt{2}/8$	$1/2$	0	$2-\sqrt{2}/8$	0	0	$2-\sqrt{2}/8$	0	$1/2$	$2+\sqrt{2}/8$	$1/2$
a'	b	$1/2$	$2+\sqrt{2}/8$	$1/2$	0	$2-\sqrt{2}/8$	0	0	$2-\sqrt{2}/8$	0	$1/2$	$2+\sqrt{2}/8$	$1/2$
	b'	$1/2$	$2-\sqrt{2}/8$	0	0	$2+\sqrt{2}/8$	$1/2$	0	$2+\sqrt{2}/8$	$1/2$	$1/2$	$2-\sqrt{2}/8$	0

Figure 2.4: Correlations in the CHSH experiment [36], which has as reference distribution the non-signalling PR-box distribution $\widetilde{\gamma}_0 = \gamma_{\mathrm{PR}} = \gamma_{\mathrm{ns}}^{(2,2)}$ [125]. The latter is attained over the context structure $\mathcal{B}_{1\&2}^{\mathrm{PR}}$, and approximated by γ_{cl} over the trivial context structure in classical theories as well as by γ_{qm} over the context category $\mathcal{V}_{1\&2} = \mathcal{V}(\mathcal{N}_1) \times \mathcal{V}(\mathcal{N}_2)$ in quantum theory.

This in turn violates the corresponding (single-context) state space correlation term: there, the closest distribution to γ_{PR} is given by the distribution $\gamma_{\mathrm{cl}}(+,+) = \gamma_{\mathrm{cl}}(-,-) = \frac{1}{2}$, $\gamma_{\mathrm{cl}}(+,-) = \gamma_{\mathrm{cl}}(-,+) = 0$ (cf. Fig. 2.4), which has distance $\Delta(\gamma_{\mathrm{cl}}, \gamma_{\mathrm{PR}}) = 2$ in accordance with Eq. (2.37).

Note that $\mathbb{E}_T^{\max}(c_{\mathrm{ns}}^{2,2}) = 4 - \Delta_T(\gamma_{\mathrm{ns}}^{2,2})$, where $c_{\mathrm{ns}}^{2,2} = ab + ab' + a'b - a'b'$ is the (statistical) quantity in the CHSH experiment in Sec. 2.4. We thus rediscover the CHSH inequality in Eq. (2.10) as multiple-context distances to the non-signalling distributions in $\widetilde{\gamma}_0 = \gamma_{\mathrm{ns}}^{(2,2)} = \gamma_{\mathrm{PR}}$, which are purposefully chosen to reveal any departure from the product constraints imposed by factorisability.

2.6 Summary

In this chapter we studied contextuality in foundations of quantum theory. We gave a precise conceptual definition of this physical principle in the form of the partial order of contexts, colloquially, the collection of nested classical perspectives onto a physical system, and poured it into a rigorous mathematical formalism for the case of quantum theory. We highlighted that many integral properties of quantum theory arise as *local-to-global*-type constraints in the form of global sections of suitable presheaves over the context category, in particular, we provided new reformulations for Stinespring's and Bell's theorem. Not only do these reformulations solidify the unifying status of physical contextuality in quantum theory, but they come with significant improvements over existing results. First, Stinespring's theorem in contextual form shows that complete positivity—a crucial property of quantum channels—arises from positivity, dilations in contexts, and a canonical choice of time orientation. Second, Bell's theorem in contextual form unifies the two substantially different ways of composing systems in classical and quantum theory. Our reformulation thus relates to both faces of Bell's theorem: bounds on classical as well as quantum correlations, as in the CHSH inequality, Eq. (2.10). In particular, by defining composition via contexts rather than state spaces, we proved that global sections of the Bell presheaf correspond with quantum states unambiguously. This means a generalisation of Gleason's theorem to composite systems over the oriented context category. Moreover, by identifying no-signalling with the marginalisation constraints over the product context category, we showed that previous attempts at deriving the state space of quantum theory from no-signalling need to be complemented with a consistency condition on time orientations in subsystems. Consequently, no-signalling singles out quantum states over the context structure in quantum theory with appropriately chosen local time orientations.

Our results apply to general von Neumann algebras and thus to general algebraic quantum theory. As a possible generalisation to quantum theory, we defined general non-signalling theories by relaxing the context structure arising from von Neumann algebras to (Hyperstonean) orthomodular lattices. We embedded Bell's theorem into this setting and provided a method to compare correlations in theories with different context structure.

Concretely, we showed that the CHSH inequality arises from a multiple-context distance measure between global sections in such theories.

This research is naturally associated with the topos approach to quantum theory initiated by Isham, Butterfield, and Hamilton [23, 24, 75, 87] and further developed by Isham and Döring [1–4]. It should therefore also be understood as a successful test of the underlying deep insight to view a quantum system as the collection of its classical perspectives. Given its continued success, it is natural to ask for avenues of future research in this field.

One that arises out of this work, especially out of the reformulation of Bell's theorem over the composite context category, is a rigorous definition of composition of systems in the topos formalism in the form of a universal property for appropriately defined categorical objects. On a broader level, it would be interesting to elaborate on the inherent geometrical character of the formalism. Already suggested by Isham, the obstructions arising from contextuality via order relations might have a geometric origin and as such should be classified by means of geometrical invariants, e.g. by the study of Čech cohomology. Moreover, the clean separation of the dichotomic function of self-adjoint operators as observables and generators of time evolution suggests a further geometrisation of the generalised state spaces in the form of the spectral and probabilistic presheaf. This could help close the gap between classical and quantum theory, suggest ways for quantisation, and ultimately lead the way towards unification of quantum theory with other inherently geometric theories in physics, most desirably, gravity.

Further avenues for future research outside the immediate scope of the topos approach arise e.g. from the close resemblance of the proof techniques used in the reformulations of Stinespring's and Bell's theorem with existing criteria for entanglement in quantum information theory [121], in particular, separable states might correspond to global sections independent of local time orientations in subsystems. Also, it remains for future study to relate the 'contextual distance measure' between correlations in theories with different context structures to other approaches, in particular, the setting of Hyperstonean orthomodular lattices seems a promising starting point to cross-identify ideas in the topos and the graph-theoretic approach to contextuality.

Chapter 3

Contextuality in Quantum Computation

Quantum computation rests on the idea to process information by the laws of quantum rather than classical theory. There are at least two aspects to this. First, a unit of information in classical physics, known as a bit, has two possible configurations, and measurement ('read-out') simply reveals this information. On the other hand, a quantum mechanical unit of information, called a qubit, possesses a plethora of possible configurations. Measurement still reveals only one of two possible outcomes, yet in general only with some probability. Second, in modern computers classical information is processed in electrical circuits by performing logic gates in a well-ordered manner. Quantum information can be processed in a similar computing architecture, known as the circuit model of quantum computation. Yet, while likely the most broadly used one to date, several alternative models of computation exist.

One that is special to quantum theory is called measurement-based quantum computation (MBQC). This model builds on the quantum phenomenon called nonlocality, which exhibits in correlations between space-like separated parties that cannot be reproduced by any classical model according to Bell's theorem. The responsible quantum states are called entangled and are arguably at the heart of the mystery behind quantum mechanics. MBQC exploits such nonlocal correlations by performing a series of local measurements on a (highly) entangled resource state, and post-processing the resulting local measurement outcomes.

This idea goes back to the groundbreaking work of Raussendorf, Briegel, and Browne [21, 128, 130], who also showed that for suitable resource states MBQC presents a universal model for quantum computation. What is more, the framework is especially interesting for the study of one of the key open questions in the field of quantum computation. While quantum computers are believed to outperform classical computers, it is not clear how this advantage occurs and what the underlying resource is. From a theoretical point of view, as well as for the technical realisation of future quantum computers, a key challenge therefore is to identify structures in quantum theory, which provide a provable quantum advantage, and to classify and quantify such resources. In recent years, contextuality has been suggested to play this role [20, 42, 44, 85, 98, 110, 126], and since measurement-based quantum computation explicitly exploits contextuality in the form of nonlocality in composite systems, this has led to a number of strong results on the resource character of contextuality in this setting [64, 119, 127, 129, 143].

The goal of this chapter is to further study contextuality within this architecture. In doing so, we prove a computational criterion for contextuality in general measurement-based computation in Sec. 3.1, we construct new examples of contextual MBQC in Sec. 3.2, and identify a possible resource measure for contextuality in the form of the number of qudits required for implementation in Sec. 3.3.

3.1 Contextuality in measurement-based computation

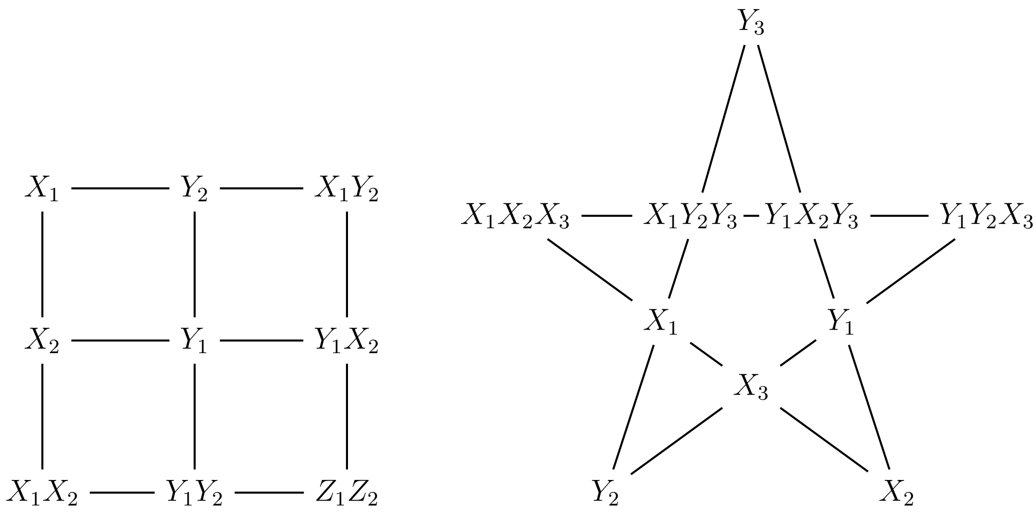
Following the broad objective to classify the resource for quantum advantage, we strive to classify contextuality in the form of nonlocality within the hybrid quantum computing architecture known as measurement-based computation (MBC). This framework is tailor-made for the study of nonlocality and largely independent of a particular physical implementation. Even on such abstract level, we prove a strong link between function computation and nonlocality, which generalises a previous result in [126]. This result has also been published in [64].

In Sec. 3.2 we will specialise this to the quantum case and further refine the classification of contextuality for measurement-based quantum computation (MBQC). Nevertheless, already in this section we will at times consider guiding examples arising within the latter. In fact, an illustrative example, which contains much of the general structure, is the following.

A prototypical example: the Anders-Browne NAND-gate

We discuss an example, which illustrates the close relation between contextuality and computation. Recall that contextuality in the form of nonlocality is at the heart of Bell's theorem. Nonlocal correlations arise, for instance, from the outcome statistics of local measurements performed on certain entangled resource states (cf. Sec. 2.4). Interestingly, such correlations can be expressed in a computational way, as was first realised in [10], where Mermin's famous proof of contextuality based on local measurements on a three qubit GHZ-state is turned into the computation of a NAND-gate. We briefly review both, the original contextuality argument by Mermin [112, 113] and the related computation due to Anders and Browne [10].

Consider the qubit operators in Fig. 3.1 (b) $\mathcal{O}_M$, which all have eigenvalues ± 1. Note also that the operators along any of the lines commute with each other, and thus define a context. Now assume there was a value assignment $v : \mathcal{O}_M \to \{-1, 1\}$ for the operators in Fig. 3.1 (b), then their eigenvalues are required to reflect the algebraic constraints on the level of operators according to Def. 1. In particular, multiplying operators in contexts results in the identity for all edges apart from the horizontal one, for which the product of operators yields negative identity.



(a) Mermin-Peres square. (b) Mermin's star.

Figure 3.1: Qubit operators (1_k's omitted) used in the contextuality proofs in (a) the Mermin-Peres square, and (b) Mermin's star [112, 113]. In both cases, product constraints between operators cannot be consistently reflected in local value assignments: multiplication of pre-assigned values ± 1 to measurements in such assignments (nodes in the diagrams) across contexts yields $+1$—every measurement appears in two contexts—whereas multiplication of operators in contexts (edges in the diagrams) yields -1 in one and $+1$ in all other contexts (cf. Eq. (3.1)).

This immediately implies a contradiction:

$$\prod_{O \in \mathcal{O}_M} v(O)^2 = 1 \neq -1 = \prod_{C \in \mathcal{C}(\mathcal{O}_M)} \left(\prod_{O \in C} v(O) \right) \tag{3.1}$$

Note that the operators XXX, XYY, YXY, and YYX in the horizontal context share the GHZ-eigenstate $|\psi_{\text{GHZ}}\rangle = \frac{1}{\sqrt{2}}(|001\rangle - |110\rangle)$.[1] We can use this to build a computation as follows.

Let $\mathbf{i} = (i_1, i_2)^\intercal \in \mathbb{Z}_2^2$ denote the input of the computation and consider a classical control computer, which selects measurements on individual qubits according to $M_k(c_k = 0) = X_k$ and $M_k(c_k = 1) = Y_k$ for $k \in \{1, 2, 3\}$ and $c_k = l_k(\mathbf{i})$ with linear functions $l_1(\mathbf{i}) = i_1$, $l_2(\mathbf{i}) = i_2$, and $l_3(\mathbf{i}) = i_1 \oplus i_2$. It is easy to see that these measurement settings define the global observables $M(\mathbf{i})$ in the context corresponding to the horizontal line in Fig. 3.1 (b):

$$M((0,0)^\intercal) = X \otimes X \otimes X \qquad M((1,0)^\intercal) = Y \otimes X \otimes Y$$
$$M((0,1)^\intercal) = X \otimes Y \otimes Y \qquad M((1,1)^\intercal) = Y \otimes Y \otimes X$$

[1] Here and throughout, we often omit the tensor product between product operators and product states.

The resource state $|\psi_{\mathrm{GHZ}}\rangle$ is an eigenvector of each global observable $M(\mathbf{i})$ with corresponding eigenvalues given by the sum of local measurement outcomes m_k, where m_k denotes the measured eigenvalue on qubit k in $|\psi_{\mathrm{GHZ}}\rangle$. While the local measurement outcomes are individually random, the global eigenvalues are deterministic and can be expressed in terms of the input as a Boolean function $o : \mathbb{Z}_2^2 \rightarrow \mathbb{Z}_2$, $o(\mathbf{i}) = \sum_{k=1}^{3} m_k$. The post-processed measurement outcomes of the local X and Y measurements thus yield the NAND-gate as output function of the computation,

$$(-1)^{o(\mathbf{i})} = (-1)^{\sum_{k=1}^{3} m_k} = (-1)^{\mathrm{NAND}(\mathbf{i})} . \tag{3.2}$$

Why is this interesting? Note that the classical computer evaluates linear functions only, we may therefore assume it to be capable of doing just that. On the other hand, the NAND-gate is clearly a nonlinear function. Hence, if we exclude any type of communication between the measurement sites until all measurements are performed (and assume a notion of 'statistical regularity' [103]), we must conclude that what boosts the power of the classical computer is the quantum resource. In essence, contextuality (nonlocality) acts as a resource that lifts the restricted complexity class of the classical control computer to universal classical computation. As it turns out, the relation between contextuality and computational advantage is not just a peculiarity of this example. To see this, in the next section we introduce the hybrid computing architecture, which underlies the above example yet generalises to arbitrarily correlated resources.

3.1.1 Definition of ld-MBC

In this section we will primarily be concerned with generalising the results in [126], which are valid beyond the quantum case but are restricted to very simple systems only. For this reason, we define measurement-based computation in a way that is independent of a physical implementation, and specialise the framework to the quantum case only in Sec. 3.2, where we discuss contextuality in measurement-based quantum computation in more detail. Our setup fits within the computational framework first introduced in [10] and further refined in [126], in order to study the computational power of correlated resources in general, which includes measurement-based quantum computation as a special case.

The Setup

A general measurement-based computation (MBC) consists of two components: a correlated resource, and a control computer with restricted computational power. The correlated resource consists of N local parties, each of which is allowed to exchange classical information with the control computer once. No communication between parties is allowed during the computation, and the correlations in their output are entirely due to interactions prior to the computation. During the exchange with the control computer, each party receives an input from the control computer (called the measurement setting), and returns an output (called the measurement outcome). The control computer combines the local measurement outcomes to produce the computational output.

We restrict the complexity of the classical control to $\mathbb{F}_d$-linear side-processing, where $\mathbb{F}_d$ denotes the finite field with $d = p^r$ elements for p prime and $r \in \mathbb{N}$. This greatly simplifies the analysis of contextuality as a resource in MBC, but will have to be lifted in future study in order to quantify any advantage of MBQC over universal classical computers.

Definition 35. *A ld-*MBC with classical input $\mathbf{i} \in \mathbb{F}_d^n$ and classical output $o(\mathbf{i}) \in \mathbb{F}_d$ *consists of N parties, each of which receives an input $c_k \in \mathbb{F}_d$ from the control computer, returns an outcome $m_k \in \mathbb{F}_d$, for $k = 1, \dots N$, and is restricted to linear side-processing as follows:*

1. *the choice of measurement bases $\mathbf{c} = (c_1, \cdots, c_N)^\intercal$ is related to the measurement outcomes $\mathbf{m}$ and the $\mathbb{F}_d$-valued classical input $\mathbf{i} = (i_1, \cdots, i_n)^\intercal \in \mathbb{F}_d^n$ via*

$$\mathbf{c} = T\mathbf{m} + C\mathbf{i} \mod d\,, \tag{3.3}$$

   for some $T \in Mat_N(\mathbb{F}_d)$ and $C \in Mat(N \times n, \mathbb{F}_d)$;

2. *for a suitable ordering of parties $1, \cdots, N$ the matrix T in Eq. (3.3) is lower triangular with vanishing diagonal. If $T = 0$ the ld-MBQ is called* non-adaptive *or* (temporally) flat*;*

3. *the computational output $o(\mathbf{i}) \in \mathbb{F}_d$ is a linear function of the local measurement outcomes*

$\mathbf{m} = (m_1, \cdots, m_N)^\intercal$ *such that for* $m_0 \in \mathbb{F}_d$ *and* $Z \in Mat(1 \times N, \mathbb{F}_d)$,

$$o(\mathbf{i}) = Z\mathbf{m} + m_0 \mod d \,. \tag{3.4}$$

In the above, $\mathrm{Mat}_N(\mathbb{F}_d)$ denotes the space of $N \times N$ matrices with entries in the finite field $\mathbb{F}_d$. Note that the setup does not specify the nature of the correlated resource or the measurements performed on them. In later sections, we will specify the framework to the special case of ld-MBQC, where the correlated resource is given by some entangled quantum state, and measurements correspond with quantum operators. However, the contextuality thresholds derived in this section hold on the level of general ld-MBC, and we thus defer a thorough introduction of the particular quantum implementation to Sec. 3.2.

Non-contextuality in ld-MBC

We define the notion of contextuality and nonlocality considered in the framework of ld-MBC. Similar to Ch. 2, we denote the set of observables by $\mathcal{O}$, and consider it equipped with an equivalence relation called *simultaneous measurability*. A context $C \subseteq \mathcal{O}$ consists of a set of simultaneously measurable observables, and we denote the set of all contexts by $\mathcal{C}$. However, contrary to the discussion in Ch. 2, a system is called *contextual*, if no non-contextual value assignment, i.e., no valuation function as in Def. 1 exists.[2] In the special case of product observables $a = a_1 \times a_2$, $b = b_1 \times b_2$, and $ab = a_1b_1 \times a_2b_2$, the locally measurable observables a_1 and b_1 (a_2 and b_2) compose individually, and by similar reasoning as in Sec. 2.4.3 we require local value assignments to do, too. Accordingly, a value assignment is a map $v : \mathcal{O} \to \mathbb{R}$ with the following properties (cf. Def. 1):

(*i*) $\forall a \in \mathcal{O}: \; v(a) \in \mathrm{sp}(a)$, where $\mathrm{sp}(a)$ denotes the set of measurement outcomes of a

(*ii*) $\forall C \in \mathcal{C}, \forall a, b, ab \in C : v(a)v(b) = v(ab)$

When no such local assignment exists, we say that the system is *nonlocal*.

[2] A system is sometimes called contextual, but not strongly contextual, if non-contextual value assignments exist, yet are not compatible with quantum theory (cf. [6, 126]). Here, we will not make this distinction.

The connection with ld-MBC is as follows. Each input $\mathbf{i} \in \mathbb{F}_d^n$ can be regarded as selecting a context $C(\mathbf{i})$ (that is, a set of simultaneously measurable observables) through

$$C(\mathbf{i}) = \{M_1(c_1(\mathbf{i})), \ldots, M_N(c_N(\mathbf{i})), M(\mathbf{i})\} \ . \tag{3.5}$$

We have included the global observable $M(\mathbf{i}) = M_1(c_1(\mathbf{i})) \times M_2(c_2(\mathbf{i})) \times \cdots \times M_N(c_N(\mathbf{i}))$ in each context as its measurement outcome is fixed in the deterministic case, and corresponds to the computational output $o(\mathbf{i})$ (that is inferred from outcomes of the local measurements). The task of finding a non-contextual hidden variable model is to find (perhaps many) value assignments to local observables that are consistent with the global value assignment. Since global value assignments correspond with the computational output in MBC, certain computations may not be compatible with non-contextual hidden variable models and thus constitute a proof of contextuality and nonlocality.

3.1.2 Contextuality in $l2$-MBC

As with the general case of quantum computation, it is natural to study what resource lifts the restricted classical control computer in ld-MBC to universal computation. In the simplest case $d = 2$ (with $\mathbb{F}_2 = \mathbb{Z}_2$), the conditions under which an $l2$-MBC allows for the computation of nonlinear Boolean functions—functions that would otherwise be beyond the capabilities of the control computer—have been well characterised.

Theorem 46. (Raussendorf [126]) *Let M be a $l2$-MBC, which deterministically evaluates a Boolean function $o : \mathbb{Z}_2^n \longrightarrow \mathbb{Z}_2$. If $o(\mathbf{i})$ is nonlinear in $\mathbf{i} \in \mathbb{Z}_2^n$, then M is contextual.*

In other words, if a $l2$-MBC can be described by a non-contextual hidden variable model, where measurement outcomes are pre-determined by value assignments, it is restricted to computing linear functions. Thm. 46 thus establishes a strong connection between function computation in the general computing architecture of $l2$-MBC and contextuality. Note also that Thm. 46 holds even in the adaptive case (cf. Def. 35).

Contextuality thus acts as a resource in the setting of $l2$-MBC. This is true, in particular, in any quantum implementation: if local measurements on a multi-qubit state can be used to evaluate nonlinear Boolean functions with only linear side-processing, then such computation constitutes a proof of contextuality—the possible local measurement outcomes cannot all be pre-assigned. Clearly, this generalises the computation of the NAND-gate on three qubits in [10], which arises from the explicit proof of contextuality in Mermin's star in Fig. 3.1 (b).

However, Thm. 46 restricts to the simplest type of local systems only, e.g. *qubits* in the quantum case. The latter have unusual properties from the perspective of contextuality. Single qubits are non-contextual by the Kochen-Specker theorem, Thm. 18, while entangled qubits exhibit state-independent contextuality using only Pauli observables in contrast to its qudit counterparts. It is therefore natural to ask whether Thm. 46 generalises to systems with size $d \geq 3$, in particular, whether the interplay between contextuality and nonlinearity holds more generally, or whether it crucially depends on some of the pathologies associated with qubit contextuality. In fact, the general case is not so straightforward [83], as we will see by considering some explicit examples within the qudit stabiliser formalism in the next section. For instance, certain nonlinear functions can be computed already within non-contextual ld-MBQC.

3.1.3 Examples from the qudit stabiliser formalism

In this section, we illustrate some of the subtleties involved in the case $d \geq 3$. We focus on a particularly interesting case of measurement-based quantum computation, where local measurements arise as gates in the qudit stabilizer formalism. Unlike the qubit case, the latter is non-contextual (in the sense defined in Sec. 3.1.1).[3]

Contrary to what one might naively expect, we will see that local measurements arising in the qudit stabilizer formalism possess a computational power that exceeds $\mathbb{F}_d$-linear processing. That is, nonlinear functions can be evaluated using a ld-MBQC that is entirely non-contextual, in stark contrast to the qubit case. This demonstrates that the relationship between contextuality and nonlinearity in the qubit case, more generally $d = 2$, is not the end of the story, and for qudits, more generally $d \geq 3$, we need a finer functional constraint.

[3]An explicit non-contextual hidden variable model for the qudit stabilizer theory is given by the discrete Wigner function defined in [70, 142].

Sympletic structure of qudit stabilizer formalism

In this section we specify the setting of ld-MBQC in the qudit stabiliser formalism with d prime, where measurements belong to the qudit Pauli group and are given by conjugation $M_k = U_k^{c_k} M_k(0) U_k^{-c_k}$, $c_k \in \mathbb{Z}_d$ with unitaries U_k in the Clifford group. For details, see e.g. [41].

Recall that the Pauli group $\mathcal{P}_d^{\otimes N}$ over $\mathbb{Z}_d$ is the group generated by N-fold tensor products of individual elements from $\langle X_k, Z_k, \omega 1_k \rangle$, $k \in \{1, \cdots, N\}$ with $X|q\rangle = |q+1\rangle$, $Z|q\rangle = \omega^q |q\rangle$, and $\omega = e^{\frac{2\pi i}{d}}$ a d-th root of unity.

These qudit Pauli operators can be conveniently represented (up to phase) by Weyl operators. A Weyl operator $W_{\mathbf{v}}$ for $\mathbf{v} = (\mathbf{a}, \mathbf{b})^\intercal \in \mathbb{Z}_d^{2N}$ is defined as the N-fold tensor product of local operators $W_{a,b} = \tau^{-ab} Z^a X^b$, where $\tau^2 = \omega$ and $a, b \in \mathbb{Z}_d$. Weyl operators are generalised Pauli operators with a particular choice of phase. Importantly, Weyl operators obey the defining commutation relation,

$$W_{\mathbf{v}} W_{\mathbf{w}} = \omega^{[\mathbf{v},\mathbf{w}]} W_{\mathbf{w}} W_{\mathbf{v}} \ , \tag{3.6}$$

where $[\mathbf{v}, \mathbf{w}] = \mathbf{v}^\intercal \sigma_{2N} \mathbf{w}$ for $\mathbf{v}, \mathbf{w} \in \mathbb{Z}_d^{2N}$ and symplectic matrix $\sigma_{2N} = \begin{bmatrix} 0_N & 1_N \\ -1_N & 0_N \end{bmatrix}$.

The Clifford group $\mathcal{C}_N(d) \subset \mathcal{U}(\mathbb{C}_d^{\otimes N})$ of $\mathcal{P}_d^{\otimes N}$ is the group of unitary operators such that $VPV^\dagger \in \mathcal{P}_d^{\otimes N}$ for all $P \in \mathcal{P}_d^{\otimes N}$, $V \in \mathcal{C}_N(d)$. All (N-qudit) Clifford operators $V \in \mathcal{C}_N(d)$ factorise,

$$V = U W_{\mathbf{x}}, \quad \mathbf{x} \in \mathbb{Z}_d^{2N} \ ,$$

into a Weyl operator $W_{\mathbf{x}}$ and an element of the group of symplectic Clifford operators $U \in \sigma\mathcal{C}_N(d)$. The latter are defined as automorphisms on the set of Weyl operators, i.e., for all $\mathbf{v} \in \mathbb{Z}_d^{2N}$ it holds that $U W_{\mathbf{v}} U^\dagger = W_{\mathbf{w}}$ for some $\mathbf{w} \in \mathbb{Z}_d^{2N}$, in fact, they preserve the underlying symplectic structure,

$$U W_{\mathbf{v}} U^{-1} = W_{C_U \mathbf{v}}, \text{ for some } C_U \in \mathrm{Sp}_{2N}(\mathbb{Z}_d) \ . \tag{3.7}$$

Here, the group $\mathrm{Sp}_{2N}(\mathbb{Z}_d)$ denotes the group of symplectic transformations, i.e., linear transformations $C : \mathbb{Z}_d^{2N} \longrightarrow \mathbb{Z}_d^{2N}$ such that $C^\intercal \sigma_{2N} C = \sigma_{2N}$.

With these preliminaries on the symplectic structure of the qudit stabiliser formalism, we study the transformation properties of Pauli observables under Clifford operations for the computational output of the corresponding ld-MBQC.

To this end, note that by the Weyl commutation relations in Eq. (3.6) and the fact that symplectic operators preserve the symplectic inner product in Eq. (3.7), we obtain the following relation for any Clifford unitary $V \in \mathcal{C}_N(d)$ acting on an individual qudit:[4]

$$\begin{aligned} V^c\, W_{\mathbf{v}}\, V^{-c} &= (UW_{\mathbf{x}})^c\, W_{\mathbf{v}}\, (UW_{\mathbf{x}})^{-c} \\ &= (UW_{\mathbf{x}})^{c-1}\, W_{C_U\mathbf{x}}\, W_{C_U\mathbf{v}}\, W_{-C_U\mathbf{x}}\, (W_{-\mathbf{x}}U^{-1})^{c-1} \\ &= (W_{C_U\mathbf{x}}\, \cdots\, W_{C_U^c\mathbf{x}})\, W_{C_U^c\mathbf{v}}\, (W_{-C_U^c\mathbf{x}} \cdots\, W_{-C_U\mathbf{x}}) \\ &= \omega^{[C_U\mathbf{x},C_U^c\mathbf{v}]+[C_U^2\mathbf{x},C_U^c\mathbf{v}]+\,\cdots\,+[C_U^c\mathbf{x},C_U^c\mathbf{v}]} W_{C_U^c\mathbf{v}} \\ &= \omega^{\sum_{j=0}^{c-1}[\mathbf{x},C_U^j\mathbf{v}]} W_{C_U^c\mathbf{v}} \end{aligned} \tag{3.8}$$

The phase in Eq. (3.8) is state-independent, it only depends on the Weyl commutation relations and the symplectic structure of the Clifford group. Yet, choosing local measurements $M_k(c_k) = V_k^{c_k} W_{\mathbf{v}_k} V_k^{-c_k}$ we can already construct ld-MBQCs with linear and nonlinear output.

Example 1: Linear output

As a first example, we examine the very restrictive case, where the controlled unitary operators in Eq. (3.8) are Pauli operators, i.e., $V = W_{\mathbf{x}}$, $U = 1 \in \sigma\mathcal{C}_1(d)$. Then the phase depends linearly on the linear input function $c = l(\mathbf{i})$, in fact, we simply obtain a variant of Eq. (3.6),

$$W_{\mathbf{x}}^{l(\mathbf{i})} W_{\mathbf{v}} W_{\mathbf{x}}^{-l(\mathbf{i})} = \omega^{l(\mathbf{i})[\mathbf{x},\mathbf{v}]} W_{\mathbf{v}}\ . \tag{3.9}$$

From Eq. (3.9) we infer that conjugation of a Pauli operator by Pauli operators results at most in multiplication of a phase, yet does not change the context. That is $M(\mathbf{i}) \propto M(\mathbf{0})$ and $C(\mathbf{i}) \propto C(\mathbf{0})$ for all inputs, meaning the output $o(\mathbf{i})$ is linearly related to $o(\mathbf{0})$. As a result, we are trivially restricted to non-contextuality.

[4]For clarity, we omit the subscript k labeling different qudit sites.

Example 2: Quadratic output

The situation changes if we apply a non-trivial symplectic Clifford operator in Eq. (3.8). In particular, we show how using control unitaries from the symplectic Clifford group $\sigma\mathcal{C}_N(d)$ allows us to compute quadratic functions due to the underlying symplectic structure.

Note that the generalised phase gate S for $d \geq 3$ is an element of the symplectic Clifford group,

$$S = \sum_{q=0}^{d-1} \tau^{q^2} |q\rangle\langle q| \ \in \ \sigma\mathcal{C}_1(d) \ ,$$

which up to phase acts on the generalised Pauli $X = W_{\mathbf{v}}, \mathbf{v} = (0,1)^\intercal$ by multiplication with Pauli Z, $S^k X S^{-k} = \tau^{-k} Z^k X$. Furthermore, consider the following state of $N = 2d$ qudits,

$$|\psi\rangle = \frac{1}{\sqrt{d}} \sum_{q=0}^{d-1} |q\rangle^{\otimes 2} \otimes |q+1\rangle^{\otimes 2} \otimes \cdots \otimes |q+d-1\rangle^{\otimes 2} \ .$$

We fix all linear functions $c_k = l_k(\mathbf{i}) = l(\mathbf{i})$ to be the same, and note that we have the following stabilizer relations,

$$\bigotimes_{k=1}^{2d} \left(S^{l(\mathbf{i})} W_{\mathbf{v}} S^{-l(\mathbf{i})} \right)_k |\psi\rangle = |\psi\rangle, \quad \mathbf{v} = (0,1)^\intercal \ , \tag{3.10}$$

where the parentheses $(\cdot)_k$ denote the subsystem on which the operator acts.

We can use Eq. (3.10) together with the symplectic structure of Weyl operators to implement quadratic output functions through accumulated symplectic products: without loss of generality, choose the first qudit and take $V_1 = (SW_{\mathbf{x}})_1$ for $\mathbf{x} = (0,-1)^\intercal$ in Eq. (3.10) such that $[\mathbf{x}, C_S^k \mathbf{v}] = k$, while leaving $V_k = S_k$ for $k \geq 2$, hence,

$$o(\mathbf{i}) = \sum_{j=0}^{l(\mathbf{i})-1} [\mathbf{x}, C_S^j \mathbf{v}] = \frac{l(\mathbf{i})(l(\mathbf{i})-1)}{2} \ . \tag{3.11}$$

Despite l being a linear function, the output function o is quadratic due to the symplectic structure of the Weyl group. In a similar vein, one obtains other nonlinear functions as well. This raises at least two questions: (i) What functions can be computed in ld-MBC?, and (ii) Is there a generalised contextuality threshold as in Thm. 46 for $d \geq 3$? We will address the former in Sec. 3.2, and give an answer to the latter in the next section.

3.1.4 Contextuality in ld-MBC

In this section we prove a generalisation of Thm. 46 to the case where d is a prime power. Despite the conceptual differences between the qubit and qudit case, which we highlighted in the previous section, we will provide a criterion for contextuality in ld-MBC, which only involves the degree of the output function. In particular, it does not depend on the particular implementation in quantum theory, but holds in full generality of non-adaptive, deterministic ld-MBC. Before we state the theorem, we need to provide some background first.

Let $\mathbb{F}_d$ be the finite field with $d = p^r$ (p prime and $r \in \mathbb{N}$) elements and denote by $\Omega_n^{\mathbb{F}_d} := \mathbb{F}_d[x_1, \ldots, x_n]$ the polynomial ring in n variables $x_1, \ldots, x_n \in \mathbb{F}_d$. For a monomial $\prod_{k=1}^n x_k^{e_k}$, $e_k \in \mathbb{N}$ is called the *partial degree corresponding with* x_k, $\sum_{k=1}^n e_k$ is called the *combined degree of* $\prod_{k=1}^n x_k^{e_k}$, and the *degree of* $f \in \Omega_n^{\mathbb{F}_d}$ is the greatest combined degree of all its monomials, denoted deg(f). We need the following standard result (cf. [108]).

Theorem 47. *Let $\mathbb{F}_d$ be the finite field with $d = p^r$ (p prime and $r \in \mathbb{N}$) elements, and $n \in \mathbb{N}$. Then every function $f : \mathbb{F}_d^n \longrightarrow \mathbb{F}_d$ is given by a polynomial $f \in \Omega_n^{\mathbb{F}_d}$ of partial degree less or equal to $d-1$ in each variable.*

Proof. Consider the Dirac delta function $\delta : \mathbb{F}_d^n \to \mathbb{F}_d$ defined as

$$\delta(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} = \mathbf{0} \\ 0 & \text{otherwise} \end{cases} . \tag{3.12}$$

We can represent δ as $\delta(\mathbf{x}) = \prod_{j=1}^n (1 - x_j^{d-1})$, which follows from Fermat's little theorem for d prime, and for general finite fields since every element in the multiplicative group $\mathbb{F}_d^\times$ has order a divisor of $d-1$. We can therefore express any function $f : \mathbb{F}_d^n \to \mathbb{F}_d$ as a linear combination of Dirac delta functions,

$$f(\mathbf{x}) = \sum_{\mathbf{y} \in \mathbb{F}_d^n} C(\mathbf{y})\delta(\mathbf{x} - \mathbf{y}) , \quad C(\mathbf{y}) \in \mathbb{F}_d . \qquad \square$$

Note that Thm. 47 is not true over infinite fields, where the corresponding ring of functions contains many non-polynomial in addition to polynomial functions.

Next, we characterise subspaces of $\Omega_n^{\mathbb{F}_d}$, which are invariant under pre- and post-composition with linear functions. We consider two obvious cases: the space of all functions $\Omega_n^{\mathbb{F}_d}$ and the space of linear functions

$$L_n^{\mathbb{F}_d} := \{l \in \Omega_n^{\mathbb{F}_d} \mid \forall \mathbf{x} \in \mathbb{F}_d^n : \; l(\mathbf{x}) = a_0 + \sum_{j=1}^n a_j x_j, \; a_j \in \mathbb{F}_d\} \,.$$

Clearly, the former is a subspace invariant under arbitrary function composition by Thm. 47. For the latter, note that any composition of linear functions results again in a linear function.

Aside from these two cases, there also exist other non-trivial subspaces stable under linear pre- and post-composition. Define the following subspaces for $1 \leq D \leq n(d-1)$,

$$\Omega_n^{\mathbb{F}_d}(D) := \left\langle \prod_{k=1}^n x_k^{e_k} \;\middle|\; e_k \in \mathbb{F}_d, \; \sum_{k=1}^n e_k \leq D \right\rangle_l \,, \tag{3.13}$$

where $\langle \cdot \rangle_l$ denotes the linear span. The function spaces $\Omega_n^{\mathbb{F}_d}(D)$ depend on the field $\mathbb{F}_d$, the number of inputs n, and the maximal *combined degree* D. In other words, $\Omega_n^{\mathbb{F}_d}(D)$ contains all polynomials $f : \mathbb{F}_d^n \to \mathbb{F}_d$ with $\deg(f) \leq D$. We prove a lemma, detailing the behaviour of these subspaces under linear pre- and post-composition.

Lemma 3. *$\Omega_n^{\mathbb{F}_d}(D)$ is invariant under linear pre- and post-composition for all $1 \leq D \leq n(d-1)$:*

$$L_1^{\mathbb{F}_d} \circ \Omega_n^{\mathbb{F}_d}(D) \circ L_n^{\mathbb{F}_d} = \Omega_n^{\mathbb{F}_d}(D)$$

Proof. Let $f \in \Omega_n^{\mathbb{F}_d}$ be a polynomial of degree $1 \leq \deg(f) = D \leq n(d-1)$. Clearly, $f \circ l \in \Omega_n^{\mathbb{F}_d}(D)$ for all $l \in L_n^{\mathbb{F}_d}$ since evaluating a polynonomial of some degree on linear functions results in a polynomial of at most that degree. Moreover, the same holds under post-composition with linear functions and we thus find

$$L_1^{\mathbb{F}_d} \circ \Omega_n^{\mathbb{F}_d}(D) \circ L_n^{\mathbb{F}_d} \subseteq_l \Omega_n^{\mathbb{F}_d}(D) \,.$$

On the other hand, $\Omega_n^{\mathbb{F}_d}(D)$ is generated by $L_1^{\mathbb{F}_d} \circ \Omega_n^{\mathbb{F}_d}(D) \circ L_n^{\mathbb{F}_d}$ since the identity is a linear function. This proves the lemma. $\square$

We thus conclude that the subspaces closed under linear pre- and post-composition are exactly $\Omega_n^{\mathbb{F}_d}(D)$ for $1 \leq D \leq n(d-1)$, in particular, $\Omega_n^{\mathbb{F}_d}(1) = L_n^{\mathbb{F}_d}$ and $\Omega_n^{\mathbb{F}_d}(n(d-1)) = \Omega_n^{\mathbb{F}_d}$.

With these preliminaries we prove the following generalisation of the contextuality threshold in Thm. 46 [64]. A similar result was discussed in [83] from the perspective of Bell inequalities.

Theorem 48. *Let M be a flat ld-MBC with $d = p^r$ (p prime and $r \in \mathbb{N}$), which deterministically evaluates a function $o : \mathbb{F}_d^n \rightarrow \mathbb{F}_d$. If $\deg(o) \geq d$, then M is contextual.*

Proof. Since M is non-contextual by assumption, measurement outcomes at local sites $k \in \{1, \cdots, N\}$ arise from functions $m_k : \mathbb{F}_d \rightarrow \mathbb{F}_d$. More precisely, there are maps $\phi_k : \mathbb{F}_d \rightarrow \mathcal{O}$, which assign every control input $c_k \in \mathbb{F}_d$ a local measurement function. Measurement corresponds with function evaluation (cf. Sec. 2.4.1) and yields definite outcomes since M is deterministic. Including linear pre-processing $l_k \in L_n^{\mathbb{F}_d}$ we thus have the following functional relations,

$$m_k = \phi_k \circ l_k \quad \forall k \in \{1, \cdots, N\} \, . \tag{3.14}$$

Moreover, the output function o in ld-MBC is determined from local measurement outcomes by linear post-processing. Hence, the entire non-contextual computation has functional signature as depicted in Fig. 3.2. By Lm. 3 it follows that the degree of o is constrained by the maximal degree of the local functions ϕ_k, which by Thm. 47 is at most $d-1$. Hence, $o \in \Omega_1^{\mathbb{F}_d}(d-1)$ and thus $\deg(o) \leq d-1$. This proves the theorem. □

Note that in the special case $d = 2$, the maps $\phi_k : \mathbb{Z}_2 \rightarrow \mathbb{Z}_2$ are necessarily linear. Hence, the corresponding subspace for the output functions in $l2$-MBC is simply the space of linear functions $\Omega_n^{\mathbb{F}_d}(1) = L_n^{\mathbb{F}_d}$. Linearity and non-contextuality thus coincide and we recover Thm. 46 in the non-adaptive case. In particular, the nonlinear NAND-gate in the Anders-Browne example in Sec. 3.1 constitutes a proof of contextuality, which is Mermin's star (cf. Fig. 3.1 (b)).

In the general case $d = p^r$ with p prime and $r \in \mathbb{N}$, local measurement outcomes still arise by evaluation of (measurement) functions $\phi_k : \mathbb{F}_d \longrightarrow \mathbb{F}_d$, yet such functions are not all linear, but correspond with polynomials of degree less than d by Thm. 47. It follows that for qudits with $d \geq 3$ certain nonlinear functions can be implemented already locally, such as the quadratic

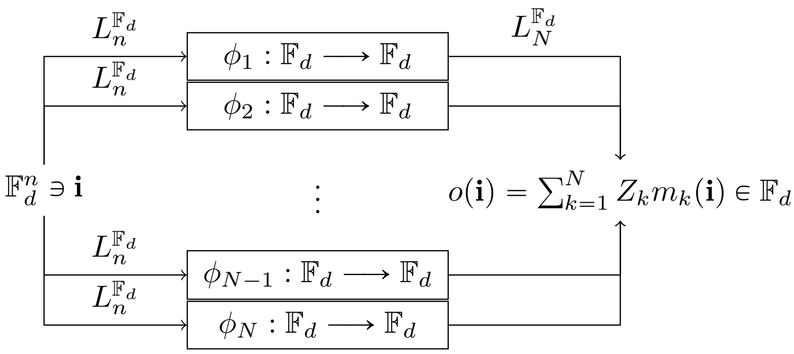


Figure 3.2: Schematic of functional signatures in non-contextual ld-MBC [64]. In a non-contextual setting, local value assignments $m_k : \mathbb{F}_d^n \to \mathbb{F}_d$ split into (classical) linear pre- and post-processing and local (quantum) measurements ϕ_k. The same holds for the output function $o : \mathbb{F}_d^n \to \mathbb{F}_d$, $o(\mathbf{i}) = \sum_{k=1}^N Z_k m_k(\mathbf{i})$ for some $Z_k \in \mathbb{F}_d$. Any additional complexity arises from the (quantum) measurements $\phi_k : \mathbb{F}_d \to \mathbb{F}_d$.

functions in the qudit stabiliser formalism in Sec. 3.1.3. Thm. 48 is thus in perfect agreement with non-contextuality of the qudit stabiliser formalism.

Note that similarly to Thm. 46, also Thm. 48 is independent of the particular physical implementation and thus holds in full generality of non-adaptive, deterministic ld-MBC. Apart from the adaptive case, Thm. 46 therefore arises as a special case of Thm. 48 for $d = 2$, where every function $\mathbb{Z}_2 \to \mathbb{Z}_2$ is linear. The fact that Thm. 46 holds in the adaptive case also turns out to be somewhat pathological. We spell this out in more detail in the next section.

3.1.5 Nonlocality, composition, and adaptivity

Unlike Thm. 46, Thm. 48 is restricted to the non-adaptive setting. That is, the measurement settings for the k-th qudit depend only on the input $\mathbf{i} \in \mathbb{F}_d^n$ and not previous measurement outcomes at sites $k' < k$.

From a computational perspective, the reason behind this restriction is that if we allowed for temporal ordering, we would effectively also allow for composition of functions. Clearly, the classification of function spaces $\Omega_n^{\mathbb{F}_d}(D)$ breaks down in this case. Nevertheless, we do have stability under composition for linear functions $L_n^{\mathbb{F}_d}$, which allows for temporal ordering in the qubit case: composition of linear functions yields linear functions. On the other hand, nonlinear functions $\phi : \mathbb{F}_d \longrightarrow \mathbb{F}_d$ will generally lift the control computer to universal computation in $\Omega_n^{\mathbb{F}_d}$ under composition.

Nevertheless, the computational restriction to non-adaptivity in the derivation of the threshold for contextuality in Thm. 48 should come as little surprise. Note that at the core of the framework of ld-MBC lies the identification of locally-measurable systems, and the power of correlations between these systems. Yet, nonlocal correlations are naturally 'size-dependent'. More precisely, if we coarse-grain the ld-MBC by grouping local systems together and allow for nonlocal measurements on those, then we will also change the threshold on contextuality in Thm. 48. A similar argument also applies under adaptivity, since the exchange of information between local parties will generally allow for the implementation of nonlocal measurements between those systems. While this is not the case for $l2$-MBC, where adaptivity still restricts to $L_n^{\mathbb{Z}_2}$ by Thm. 46, $l2$-MBC is nevertheless unstable under grouping systems together, e.g. nonlocal measurements on two qubits do implement functions $o:\mathbb{Z}_2^2 \rightarrow \mathbb{Z}_2$.

Succinctly, contextuality and nonlocality in ld-MBC are therefore not 'scale invariant' for arbitrary d. This in turn means that deriving a threshold for contextuality in the adaptive setting is harder and will usually be possible only if adaptivity is further restricted such as in $l2$-MBC.

3.1.6 The probabilistic case

Note that Thm. 46 and Thm. 48 apply to the deterministic case only. In this section we relax this restriction. For $l2$-MBC a probabilistic threshold for contextuality was given in [126]. More precisely, a $l2$-MBC is said to evaluate a Boolean function $f:\mathbb{Z}_2^n \rightarrow \mathbb{Z}_2$ with average success probability P, if $P = \frac{1}{2^n}\sum_{\mathbf{i}\in\mathbb{Z}_2^n} \mathrm{Prob}(o(\mathbf{i}) = f(\mathbf{i}))$. Moreover, define the average distance of a Boolean function $f:\mathbb{Z}_2^n \rightarrow \mathbb{Z}_2$ to the set linear functions $L_n^{\mathbb{Z}_2}$ by

$$\nu(f) := \frac{1}{2^n} \min_{l\in L_n^{\mathbb{Z}_2}} |\{\mathbf{i}\in\mathbb{Z}_2^n \mid f(\mathbf{i}) \neq l(\mathbf{i})\}| \, . \tag{3.15}$$

Theorem 49. *Let M be a $l2$-MBC, which probabilistically evaluates a nonlinear Boolean function $o:\mathbb{Z}_2^n \rightarrow \mathbb{Z}_2$ with success probability P. If $P > 1-\nu(o)$, then M is contextual.*

The optimal bound is attained for bent functions, which have maximal distance to the set of linear functions, and for which M is contextual if $P > (\frac{1}{2}) + (\frac{1}{2})^{\frac{n}{2}+1}$ [126].

We also remark that the bound in Thm. 49 can be refined by means of the *non-contextual fraction*. The latter arises in the sheaf-theoretic framework, which studies general empirical models, i.e., sets of probability distributions for measurements grouped into contexts (for details, see [6]). Any empirical model e has a convex decomposition into a contextual and a non-contextual part

$$e = \mathrm{NCF}(e)e + (1 - \mathrm{NCF}(e))e, \qquad \mathrm{NCF}(e) \in [0, 1] \ .$$

Here, a *non-contextual* empirical model has $\mathrm{NCF}(e) = 1$ and corresponds to a probability distribution over pure states, i.e., valuation functions or global sections of the corresponding event sheaf (cf. [6]). Correspondingly, an empirical model is called *contextual* if $\mathrm{NCF}(e) < 1$ and *strongly contextual* if $\mathrm{NCF}(e) = 0$.

With these definitions, [5] prove the following bound on the average success probability,

$$P \leq 1 - \mathrm{NFC}(e)\nu(o) \ .$$

Note that the non-contextual bound on the success probability in Thm. 49 crucially depends on the distance of the output function to the closest linear function $\nu(o)$. By Thm. 46 linear functions are those realisable in the non-contextual case. We can relax Thm. 48 to the probabilistic case in a similar way. First, we adjust the definition of the distance ν to inputs in finite fields,

$$\nu(f) := \frac{1}{d^n} \min_{g \in \Omega_n^{\mathbb{F}_d}(d-1)} |\{\mathbf{i} \in \mathbb{F}_d^n \mid f(\mathbf{i}) \neq g(\mathbf{i})\}| \ , \tag{3.16}$$

where the minimum is now taken over all polynomial functions of degree at most $d-1$ for d a prime power. Generalising the qubit case [126], we observe that non-contextuality bounds the average success probability of any ld-MBC, which evaluates the output function $o : \mathbb{F}_d^n \to \mathbb{F}_d$, again by

$$P \leq 1 - \nu(o) \ . \tag{3.17}$$

A violation of this inequality thus yields a proof of nonlocality and generalises the results in [5, 126], where only Boolean functions were considered [64].

Theorem 50. *Let M be a ld-MBC with $d = p^r$ (p prime and $r \in \mathbb{N}$), which evaluates a function $o : \mathbb{F}_d^n \to \mathbb{F}_d$, $\deg(o) \geq d$ with success probability P. If $P > 1 - \nu(o)$, then M is contextual.*

Proof. This follows immediately from the above discussion and the arguments in Thm. 49 under the generalisation of Eq. (3.15) to Eq. (3.16). □

Note that similar to the case $d = 2$, we achieve the optimal bound for functions that are farthest from the set of non-contextual output functions $\Omega_n^{\mathbb{F}_d}(d-1)$.

3.1.7 Summary

In this section we found a computational threshold for contextuality in non-adaptive ld-MBC with d a prime power. This generalises earlier results for $d = 2$ in [5, 126]. We have restricted to prime powers since these correspond with the number of elements in finite fields, over which the ring of functions in n variables coincides with the corresponding ring of polynomials by Thm. 47. Still, many of our arguments do not require this simplification and apply more generally.

As with general measurement-based computation, ld-MBC harnesses contextuality in the form of nonlocality: nonlocal correlations between spatially separated subsystems boost the computational power of the classical control computer capable of linear side-processing only. Thm. 48 can thus be understood as a computational version of Bell's theorem [83]: assuming local (single qudit) measurements reveal outcomes as in a local hidden variable model, computation is constrained to $\Omega_n^{\mathbb{F}_d}(d-1)$ (for some success probability). However, note that in contrast to the discussion in Sec. 2.4, MBC neglects non-contextual constraints at local sites.

In the next section we refine the results obtained in this section to the quantum case. Since quantum states correspond with global sections of the probabilistic presheaf, they are also subject to the coarse-graining constraints locally. The study of measurement-based quantum computation is therefore not only concerned with nonlocality, but with contextuality more generally. In order to study the role of contextuality as a resource in quantum computation, we therefore seek a computational classification of contextuality in ld-MBQC. As we will see, this classification crucially depends on certain local phase relations between eigenstates of local measurement operators, and thus on the explicit structure of states in quantum theory.

3.2 Contextuality in measurement-based quantum computation

Def. 35 in the last section defines ld-MBC as a framework based on correlated resources, independent of their physical implementation. In this section we are concerned with the quantum version of this setup known as ld-MBQC, where the correlated resource is given by some entangled quantum state.

3.2.1 Definition of ld-MBQC

We assume that the eigenvalues of local measurements $M_k(c_k)$ are of the form ω^{m_k} for $\omega = e^{\frac{2\pi i}{d}}$ and $m_k \in \mathbb{Z}_d$. Importantly, we will restrict to d prime such that $\mathbb{Z}_d \simeq \{\omega^m \mid m \in \mathbb{Z}\}$ is again a field. Note that operators M_k are not Hermitian, but we use the terminology 'measurement of M_k' to denote a projective measurement in the eigenbasis of M_k, where we associate the measurement outcome $m_k \in \mathbb{Z}_d$ with the eigenvalue ω^{m_k} (cf. [64, 127]). For a given input $\mathbf{i} \in \mathbb{Z}_d^n$, a 'global measurement' $M(\mathbf{i})$ is the tensor product of local measurements, and encodes the computational output. For simplicity, we will mostly restrict to non-adaptive, deterministic ld-MBQC. In this case, the product of all local measurements stabilizes the resource state, and we assume that the control computer evaluates the output function $o(\mathbf{i})$ by adding local measurement outcomes,

$$o(\mathbf{i}) = \sum_{k=1}^{N} m_k \mod d\ .$$

As before, we investigate whether the quantum resource state increases the computational power of the control computer, i.e., if the output function is of degree greater or equal to d by Thm 48.

In summary, we have the following definition for *non-adaptive* ld-MBQC (cf. Fig. 3.3).

Definition 36. *A* non-adaptive ld-MBQC with d prime, input string $\mathbf{i} \in \mathbb{Z}_d^n$, and output $o(\mathbf{i}) \in \mathbb{Z}_d$, *consists of the following components:*

1. *an N-qudit system each of local dimension d, where the overall resource state is represented by $|\psi\rangle \in (\mathbb{C}^d)^{\otimes N}$;*

2. *a set of measurement settings* $c_k = l_k(\mathbf{i})$ *for some* $\mathbb{Z}_d$*-linear functions* $l_k : \mathbb{Z}_d^n \rightarrow \mathbb{Z}_d$*, independent of previous measurement outcomes;*

3. *a set of measurements* M_k *on each qudit, each with* d *possible eigenvalues* ω^{m_k}*, where* $m_k \in \mathbb{Z}_d$ *is the measurement outcome;*[5]

4. *the computational output is the linear sum of local measurement outcomes* $\mathbf{m} = \{m_1, \cdots, m_N\} \in \mathbb{Z}_d^N$*,*

$$o(\mathbf{i}) = \sum_{k=1}^{N} m_k \mod d\,.^{6} \tag{3.18}$$

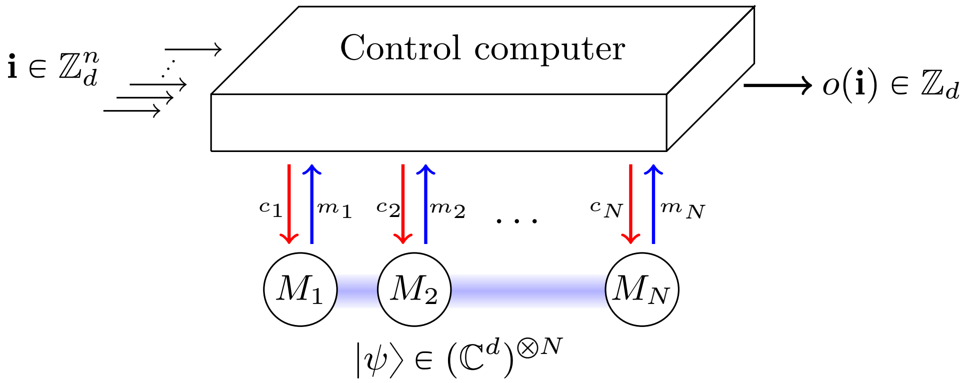


Figure 3.3: Schematic of ld-MBQC [64].

We remark that with suitably chosen resource states, such as (qudit) cluster states, *adaptive* ld-MBQC is universal for quantum computation [130, 152].

Phase relations in deterministic $l2$-MBQC

In this section we stress the importance of local phase relations between eigenstates of measurement operators in the qubit case, which will serve as a guideline for the constructive proofs in Sec. 3.2.2. We will further refine these phase relations in Sec. 3.3.3.

[5]Note that the measurement M_k is only constrained on outcomes, in particular, we do not restrict to conjugation of some reference measurement by unitaries arising as projective representations of $\mathbb{Z}_d$ such as in the qudit stabiliser formalism (cf. Eq. (3.24)).

[6]Note that this no restriction to the corresponding setting in Def. 35 since any linear post-processing can be encoded locally.

Note that every local measurement $M_k(c_k)$ defines a basis of eigenstates. Expressed in terms of the computational basis $|q\rangle$ with $q \in \mathbb{Z}_2$, for qubits any basis is of the form:

$$
\begin{aligned}
|\varphi,\vartheta\rangle &= \sin(\varphi)|0\rangle + e^{\pi i\vartheta}\cos(\varphi)|1\rangle & |0\rangle &= \sin(\varphi)|\varphi,\vartheta\rangle + \cos(\varphi)|\overline{\varphi,\vartheta}\rangle \\
|\overline{\varphi,\vartheta}\rangle &= \cos(\varphi)|0\rangle - e^{\pi i\vartheta}\sin(\varphi)|1\rangle & |1\rangle &= e^{-\pi i\vartheta}(\cos(\varphi)|\varphi,\vartheta\rangle - \sin(\varphi)|\overline{\varphi,\vartheta}\rangle)
\end{aligned}
$$

Recall that in ld-MBQC the output function $o(\mathbf{i}) = \oplus_{k=1}^{N} m_k$ arises as the parity of the individual measurement outcomes on local qudits. For qubits there are two different parity states and o encodes which of these two parities is obtained when measuring a given context. Henceforth, we choose as resource state the N-qubit GHZ-state (cf. [146]),

$$
|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle^N + |1\rangle^N) = \frac{1}{\sqrt{2}}\sum_{q=0}^{1} \otimes_{k=1}^{N} |q_k = q\rangle \ , \tag{3.19}
$$

and set the default measurement operator for the input $\mathbf{i} = 0$ to be the tensor product of Pauli-X measurements with eigenvectors $|+\rangle = |\frac{\pi}{4}, 0\rangle$, $|-\rangle = |\overline{\frac{\pi}{4}, 0}\rangle$. Clearly, $|\psi\rangle$ is a parity $+1$-eigenstate of this operator, the contributions of opposite parity cancel. In a similar way, one obtains the parity -1-eigenstate if the cancellations are such that contributions with positive parity vanish. More precisely, for other inputs the local measurement operators correspond to different bases. In particular, for $\vartheta_k \neq 0$ this results in the additional phase factor $e^{-\pi i\vartheta_k}$ for $|1\rangle_k$ in $|\psi\rangle$. In the prototypical Anders-Browne example on three qubits this reads as follows,

$$
\begin{aligned}
|\psi\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle) &\overset{X\underline{X}X}{=} \frac{1}{2}(|+++\rangle + |+--\rangle + |-+-\rangle + |--+\rangle) \\
&\overset{X\underline{Y}Y}{=} \frac{1}{2}(|+i\bar{i}\rangle + |+\bar{i}i\rangle + |-ii\rangle + |-\bar{i}\bar{i}\rangle) \\
&\overset{Y\underline{X}Y}{=} \frac{1}{2}(|i+\bar{i}\rangle + |i-i\rangle + |\bar{i}+i\rangle + |\bar{i}-\bar{i}\rangle) \\
&\overset{Y\underline{Y}X}{=} \frac{1}{2}(|ii-\rangle + |i\bar{i}+\rangle + |\bar{i}i+\rangle + |\bar{i}\bar{i}-\rangle) \ .
\end{aligned}
$$

Here, $|i\rangle$ and $|\bar{i}\rangle$ correspond to the basis with $\varphi = \frac{\pi}{4}$ and $\vartheta = \frac{1}{2}$. Note that this choice of local bases solves the following set of linear equations $\sum_{k=1}^{3} l_k(i_1, i_2) \cdot \vartheta_k = o(i_1, i_2)$, where $l_1(i_1, i_2) = i_1$, $l_2(i_1, i_2) = i_2$, $l_3(i_1, i_2) = i_1 \oplus i_2$ and $o(i_1, i_2) = 1 + (i_1 + 1) \cdot (i_2 + 1)$.

This example is illustrative in a number of ways as we show in Lm. 4 and Lm. 5 below. First, note that for deterministic l2-MBQC it is enough to consider $\varphi = \frac{\pi}{4}$.

Lemma 4. *The eigenstates of local measurements in deterministic l2-MBQC on a GHZ-state with $N \geq 3$ qubits are mutually unbiased with respect to the local basis of the resource state.*

Proof. Note that the measurement operators $M(\mathbf{i}) = \otimes_{k=1}^{N} M_k(c_k(\mathbf{i}))$ are such that $|\psi\rangle$ is a parity eigenstate of $M(\mathbf{i})$ for all $\mathbf{i} \in \mathbb{Z}_2^n$. For every input $\mathbf{i} \in \mathbb{Z}_2^n$, rewrite $|\psi\rangle$ in the local eigenbases corresponding to the $M_k(c_k(\mathbf{i}))$. This yields a superposition of product states $|\mathbf{m}\rangle = \otimes_{k=1}^{N}|m_k\rangle$, where we denote every product state by the Boolean vector $\mathbf{m} \in \mathbb{Z}_2^N$ with $m_k = 0$ for $|\varphi, \vartheta\rangle_k$ and $m_k = 1$ for $|\overline{\varphi, \vartheta}\rangle_k$ at site k. Clearly, the product state $|\mathbf{m}\rangle$ has parity $m = \oplus_{k=1}^{N} m_k$. Moreover, the coefficient to the product state $|\mathbf{m}\rangle$ reads

$$\prod_{k=1}^{N} \Phi^{m_k}(\varphi_k) + (-1)^m \prod_{k=1}^{N} \Phi^{m_k \oplus 1}(\varphi_k)\ , \tag{3.20}$$

where $\Phi^0(\varphi_k) = \sin(\varphi_k)$ and $\Phi^1(\varphi_k) = \cos(\varphi_k)$, since both $|0\rangle_k$ and $|1\rangle_k$ can have non-zero overlap with $|m_k\rangle$.[7] A parity eigenstate is on hand if all product states of some parity cancel. We thus have $\frac{2^N}{2}$ constraints from Eq. (3.20), both on absolute values and phases. Clearly, the constraints on absolute values are satisfied for $\varphi = \frac{\pi}{4}$. Moreover, for $N \geq 3$ all solutions are of this form. First, for $N \geq 3$ odd, consider pairs of constraints in Eq. (3.20) of the same parity:

$$\Phi^{m_k}(\varphi_k) \prod_{k' \neq k} \Phi^{m_{k'}}(\varphi_{k'}) + (-1)^m \Phi^{m_k \oplus 1}(\varphi_k) \prod_{k' \neq k} \Phi^{m_{k'} \oplus 1}(\varphi_{k'}) = 0$$
$$\Phi^{m_k}(\varphi_k) \prod_{k' \neq k} \Phi^{m_{k'} \oplus 1}(\varphi_{k'}) + (-1)^m \Phi^{m_k \oplus 1}(\varphi_k) \prod_{k' \neq k} \Phi^{m_{k'}}(\varphi_{k'}) = 0$$

These imply $\frac{\prod_{k' \neq k} \Phi^{m_{k'} \oplus 1}(\varphi_{k'})}{\prod_{k' \neq k} \Phi^{m_{k'}}(\varphi_{k'})} = -(-1)^m \frac{\Phi^{m_k}(\varphi_k)}{\Phi^{m_k \oplus 1}(\varphi_k)} = \frac{\prod_{k' \neq k} \Phi^{m_{k'}}(\varphi_{k'})}{\prod_{k' \neq k} \Phi^{m_{k'} \oplus 1}(\varphi_{k'})}$ and thus $|\sin(\varphi_k)| = |\cos(\varphi_k)|$, hence, $\varphi_k = \frac{\pi}{4}$. For N even, similar constraints yield $|\Phi^{m_k}(\varphi_k)\Phi^{m_{k'}}(\varphi_{k'})| = |\Phi^{m_k \oplus 1}(\varphi_k)\Phi^{m_{k'} \oplus 1}(\varphi_{k'})|$. For $N \neq 2$ we thus again find $\varphi_k = \frac{\pi}{4}$, since for another pair of constraints in Eq. (3.20) also $|\Phi^{m_k}(\varphi_k)\Phi^{m_{k'} \oplus 1}(\varphi_{k'})| = |\Phi^{m_k \oplus 1}(\varphi_k)\Phi^{m_{k'}}(\varphi_{k'})|$, hence, $\frac{|\Phi^{m_{k'} \oplus 1}(\varphi_{k'})|}{|\Phi^{m_{k'}}(\varphi_{k'})|} = \frac{|\Phi^{m_k}(\varphi_k)|}{|\Phi^{m_k \oplus 1}(\varphi_k)|} = \frac{|\Phi^{m_{k'}}(\varphi_{k'})|}{|\Phi^{m_{k'} \oplus 1}(\varphi_{k'})|}$. □

[7] Local measurements in the computational basis only change the resource state and can thus be neglected.

We are thus left with the phase constraints arising from Eq. (3.20). Generalising the Anders-Browne example above, let $\vartheta \in \mathbb{R}^N$, $o : \mathbb{Z}_2^n \to \mathbb{Z}_2$ and $\mathbf{l}(\mathbf{i}) = L\mathbf{i}$, where $\mathbf{i} \in \mathbb{Z}_2^n$ and $L \in \mathrm{Mat}(N \times n, \mathbb{Z}_2)$ arbitrary. Then any solution to the set of equations

$$\forall \mathbf{i} \in \mathbb{Z}_2^n : \; \mathbf{l}(\mathbf{i}) \cdot \vartheta = \sum_{k=1}^{N} l_k(\mathbf{i}) \vartheta_k = o(\mathbf{i}) \tag{3.21}$$

allows to construct a $l2$-MBQC. Namely, consider local measurement operators $X(\vartheta)$ with eigenstates $|\vartheta\rangle = |\frac{\pi}{4}, \vartheta\rangle$ and $|\overline{\vartheta}\rangle = |\overline{\frac{\pi}{4}, \vartheta}\rangle$. The resource GHZ-state $|\psi\rangle$ in Eq. (3.19) is then a parity $+1$-eigenstate of the operator $\otimes_{k=1}^{N} X_k(0)$. On the other hand, $|\psi\rangle$ is a parity -1-eigenstate of all operators of the form $\otimes_{k=1}^{N} X(\vartheta_k)$ for local operators such that $\sum_{k=1}^{m} \vartheta_k = 1$, e.g. $\vartheta_k = \frac{1}{m}$ for $k \in \{1, \cdots, m\}$ and $\vartheta_k = 0$ for $k \in \{m+1, \cdots, N\}$. Finding an $l2$-MBQC, which computes the output function o, thus reduces to finding a set of (linear) functions l_k, which satisfy the phase constraints in Eq. (3.21). In Sec. 3.2.2 we will use this technique in order to construct contextual examples—according to the threshold derived in Thm. 48—as well as a general classification of computation in flat, deterministic ld-MBQC.

Measurement in deterministic ld-MBQC

Recall that each party $k \in \{1, \cdots, N\}$ performs one of d (qudit) measurements $M_k(c_k)$ determined by a single input $c_k \in \mathbb{Z}_d$. We require the M_k to have (non-degenerate) eigenvalues of the form ω^{m_k} for $m_k \in \mathbb{Z}_d$, from which it follows that $M_k^d = 1$. Moreover, by the argument in Lm. 4, for deterministic ld-MBQC it is enough to consider local measurements with eigenstates mutually unbiased with respect to a reference basis, e.g. the computational basis in the qudit resource state $|\psi\rangle = \frac{1}{\sqrt{d}} \sum_{q=0}^{d-1} \otimes_{k=1}^{N} |q_k = q\rangle$. Note that we find such operators in

$$X(f)|q\rangle = f(q)|q+1\rangle \quad \text{with} \quad f : \mathbb{Z}_d \to U(1), \; \prod_{q=0}^{d-1} f(q) = 1 \; . \tag{3.22}$$

In fact, every non-degenerate local measurement operator with $M_k^d = 1$ is of this form.

Lemma 5. *Every local measurement operator M in deterministic ld-MBQC with d prime is of the form $M = X(f)$ in Eq. (3.22) for some function $f : \mathbb{Z}_d \to U(1)$ with $\prod_{q=0}^{d-1} f(q) = 1$.*

Proof. Let $M \in U(\mathbb{C}^d)$, $M^d = 1$ be non-degenerate, and denote its eigenvectors by $|m\rangle$, $m \in \mathbb{Z}_d$. By assumption, the corresponding basis is mutually unbiased with respect to $|q\rangle$, i.e., $|\langle q|m\rangle|^2 = \frac{1}{d}$ for all $q, m \in \mathbb{Z}_d$. Expressing the eigenvectors $|m\rangle$ of M in terms of the $|q\rangle$-basis thus yields

$$
\begin{aligned}
|m\rangle &= \frac{1}{\sqrt{d}}(|0\rangle + \omega^{-m} f(0)|1\rangle + \omega^{-2m} f(0)f(1)|2\rangle + \cdots + \omega^{-(d-1)m}\left(\prod_{q'=0}^{d-2} f(q')\right)|d-1\rangle) \\
&= \frac{1}{\sqrt{d}}\left(\sum_{q=0}^{d-1} \omega^{-qm} \prod_{q'=0}^{q-1} f(q')|q\rangle\right)
\end{aligned}
$$

for some $f : \mathbb{Z}_d \to U(1)$ with $\prod_{q=0}^{d-1} f(q) = 1$. $\square$

In the following, we will often consider operators with a reference phase θ,

$$
X(\theta, f)|q\rangle = \theta\chi^{f(q)}|q+1\rangle, \quad f : \mathbb{Z}_d \to \mathbb{Z}_d \ . \tag{3.23}
$$

The constraint in Eq. (3.23) then reads $\theta^d = \chi^{-\sum_{q=0}^{d-1} f(q)}$. Note that the rotation operators $X(\phi) = e^{i\phi}\sum_{q=0}^{d-2}|q+1\rangle\langle q| + e^{i(1-d)\phi}|0\rangle\langle d-1|$ defined in [105] for the construction of general proofs of contextuality are of this type, namely for $\theta = e^{i\phi}$, $\chi = e^{-id\phi}$ and $f(q) = \delta(q-(d-1))$ (see also, [140]). The inputs $c_k \in \mathbb{Z}_d$ to the measurement devices thus specify $M_k(c_k) = X(\theta(c_k), f(c_k))$ and are themselves determined in a linear way from the computational input $\mathbf{i} \in \mathbb{Z}_d^n$ according to the setup in Def. 36.

Finally, we comment on the structure of measurement operators considered here compared to those considered elsewhere [64, 126]. In particular, within the qudit stabiliser formalism (cf. Sec. 3.1.3) the classical control is often modeled by means of unitary conjugation

$$
M_k(c_k) = U_k^{c_k} M_k(0) U_k^{-c_k} \ , \tag{3.24}
$$

where $U_k^{c_k}$ is a unitary projective representation of $\mathbb{Z}_d$ for d prime, and $M_k(0)$ is some reference Pauli operator. In this case the local phases θ in Eq. (3.23) arise from the special Weyl commutation relations in Eq. (3.6) and Eq. (3.8).

3.2.2 Implementation of flat, deterministic ld-MBQC

In this section we employ the phase relations in Eq. (3.22) for computation in ld-MBQC by means of Eq. (3.21). We begin with an explicit example on three qutrits, which is contextual by the threshold in Thm. 48. Subsequently, we prove that any output function $o : \mathbb{Z}_d^n \to \mathbb{Z}_d$ for d prime can be constructed by explicitly implementing the δ-function.

A contextual qutrit example

Consider first the following three operators of the form in Eq. (3.23) on a three-dimensional system by their action on computational basis states $|q\rangle$ for $0 \leq q < 3$ and $\omega = e^{\frac{2\pi i}{3}}$:

$$M(0)|q\rangle = X|q\rangle := |q+1\rangle \quad M(1)|q\rangle := \theta_1 \omega^{q^2}|q+1\rangle \quad M(2)|q\rangle := \theta_2 \omega^{-q^2}|q+1\rangle \tag{3.25}$$

Obviously, $M(0)$ has order three. A quick computation shows that by choosing $\theta_1^3 = \omega$ and $\theta_2^3 = \omega^2$ we also have $M(1)^3 = 1$, $M(2)^3 = 1$. Moreover, we set $\theta_1\theta_2 = \omega$, e.g. $\theta_1 = e^{\frac{2\pi i}{9}}$ and $\theta_2 = e^{\frac{4\pi i}{9}}$. The eigenstates $|\theta_k^{m_k}\rangle$ of $M(k)$ with $|\theta_0^{m_0}\rangle = |x^{m_0}\rangle$ for $\theta_0 = 1$ are given as follows.

$$
\begin{aligned}
|\theta_1^0\rangle &= \frac{1}{\sqrt{3}}(|0\rangle + \theta_1|1\rangle + \theta_1^2\omega|2\rangle) & |0\rangle &= \frac{1}{\sqrt{3}}(|\theta_1^0\rangle + |\theta_1^1\rangle + |\theta_1^2\rangle)\\
|\theta_1^1\rangle &= \frac{1}{\sqrt{3}}(|0\rangle + \theta_1\omega^2|1\rangle + \theta_1^2\omega^2|2\rangle) & |1\rangle &= \frac{1}{\sqrt{3}\theta_1}(|\theta_1^0\rangle + \omega|\theta_1^1\rangle + \omega^2|\theta_1^2\rangle)\\
|\theta_1^2\rangle &= \frac{1}{\sqrt{3}}(|0\rangle + \theta_1\omega|1\rangle + \theta_1^2|2\rangle) & |2\rangle &= \frac{1}{\sqrt{3}\theta_1^2\omega}(|\theta_1^0\rangle + \omega^2|\theta_1^1\rangle + \omega|\theta_1^2\rangle)
\end{aligned}
$$

$$
\begin{aligned}
|\theta_2^0\rangle &= \frac{1}{\sqrt{3}}(|0\rangle + \theta_2|1\rangle + \theta_2^2\omega^2|2\rangle) & |0\rangle &= \frac{1}{\sqrt{3}}(|\theta_2^0\rangle + |\theta_2^1\rangle + |\theta_2^2\rangle)\\
|\theta_2^1\rangle &= \frac{1}{\sqrt{3}}(|0\rangle + \theta_2\omega^2|1\rangle + \theta_2^2|2\rangle) & |1\rangle &= \frac{1}{\sqrt{3}\theta_2}(|\theta_2^0\rangle + \omega|\theta_2^1\rangle + \omega^2|\theta_2^2\rangle)\\
|\theta_2^2\rangle &= \frac{1}{\sqrt{3}}(|0\rangle + \theta_2\omega|1\rangle + \theta_2^2\omega|2\rangle) & |2\rangle &= \frac{1}{\sqrt{3}\theta_2^2\omega^2}(|\theta_2^0\rangle + \omega^2|\theta_2^1\rangle + \omega|\theta_2^2\rangle)
\end{aligned}
$$

$$
\begin{aligned}
|x^0\rangle &= \frac{1}{\sqrt{3}}(|0\rangle + |1\rangle + |2\rangle) & |0\rangle &= \frac{1}{\sqrt{3}}(|x^0\rangle + |x^1\rangle + |x^2\rangle)\\
|x^1\rangle &= \frac{1}{\sqrt{3}}(|0\rangle + \omega^2|1\rangle + \omega|2\rangle) & |1\rangle &= \frac{1}{\sqrt{3}}(|x^0\rangle + \omega|x^1\rangle + \omega^2|x^2\rangle)\\
|x^2\rangle &= \frac{1}{\sqrt{3}}(|0\rangle + \omega|1\rangle + \omega^2|2\rangle) & |2\rangle &= \frac{1}{\sqrt{3}}(|x^0\rangle + \omega^2|x^1\rangle + \omega|x^2\rangle)
\end{aligned}
$$

Combining local measurement operators in Eq. (3.25) to global measurements on the three-qutrit resource GHZ-state

$$|\psi\rangle = \frac{1}{\sqrt{3}}(|000\rangle + |111\rangle + |222\rangle) \ , \tag{3.26}$$

allows to construct a contextual computation by Thm. 48.

Theorem 51. *The l3-MBQC with local measurements $M_k(c_k = l_k(\mathbf{i}))$ as in Eq. (3.25), where the input $\mathbf{i} = (i_1, i_2)^\intercal \in \mathbb{Z}_3^2$ sets the measurements via $\mathbb{Z}_3$-linear functions $l_1(\mathbf{i}) := i_1$, $l_2(\mathbf{i}) := i_2$, and $l_3(\mathbf{i}) := -i_1 - i_2$, is contextual when evaluated on the resource state in Eq. (3.26).*

Proof. Note that we have the following identities,

$$\begin{aligned}
M(0) \otimes M(0) \otimes M(0)|q\rangle^{\otimes 3} &= |q+1\rangle^{\otimes 3}, \\
M(1) \otimes M(1) \otimes M(1)|q\rangle^{\otimes 3} &= \omega|q+1\rangle^{\otimes 3}, \\
M(2) \otimes M(2) \otimes M(2)|q\rangle^{\otimes 3} &= \omega^2|q+1\rangle^{\otimes 3}, \\
\sigma(M(0) \otimes M(1) \otimes M(2))|q\rangle^{\otimes 3} &= \theta_1\theta_2|q+1\rangle^{\otimes 3} = \omega|q+1\rangle^{\otimes 3} \quad \forall \sigma \in S_3 \ ,
\end{aligned}$$

where we understand the permutation operator to act on the set of control inputs $\{c_1, c_2, c_3\}$. From this one readily computes the output function.

$$\begin{array}{lll}
o((0,0)^\intercal) = 0 & o((0,1)^\intercal) = 1 & o((0,2)^\intercal) = 1 \\
o((1,0)^\intercal) = 1 & o((1,1)^\intercal) = 1 & o((1,2)^\intercal) = 1 \\
o((2,0)^\intercal) = 1 & o((2,1)^\intercal) = 1 & o((2,2)^\intercal) = 2
\end{array}$$

In order to prove that this computation is contextual, we need to show that o is at least cubic according to Thm. 48. We assume to the contrary and make the ansatz,

$$g(\mathbf{i}) = \alpha_1 i_1^2 + \alpha_2 i_2^2 + \beta i_1 i_2 + \gamma_1 i_1 + \gamma_2 i_2 + \delta \ .$$

From $g((0,0)^\intercal) = 0$ we deduce $\delta = 0$. From $g((0,1)^\intercal) = 1$, $g((0,2)^\intercal) = 1$ we get $\alpha_2 + \gamma_2 = 1$,

$\alpha_2 + 2\gamma_2 = 1$, respectively, and thus $\alpha_2 = 1$, $\gamma_2 = 0$. By symmetry, $\alpha_1 = 1$, $\gamma_1 = 0$. Furthermore, from $g((1,1)^\intercal) = 1$, we get $\beta = 2$, hence, $g = i_1^2 + i_2^2 + 2i_1i_2 = (i_1+i_2)^2$. However, evaluation on the remaining inputs, $g((1,2)^\intercal) = g((2,1)^\intercal) = 0 \neq 1 = o((1,2)^\intercal) = o((2,1)^\intercal)$ and $g((2,2)^\intercal) = 1 \neq 2 = o((2,2)^\intercal)$, yields a contradiction, hence, o is not a quadratic function and must therefore be contextual by Thm. 48. In fact, one easily verifies $o(\mathbf{i}) = 2i_1^2i_2 + 2i_1i_2^2 + i_1^2 + i_2^2 + i_1i_2$. □

Clearly, this example is reminiscent of the Anders-Browne example for qubits. However, note that while the measurement operators in the qubit case can be implemented within the stabiliser formalism, $M(0)$, $M(1)$, and $M(2)$ in Eq. (3.25) cannot be part of the qutrit stabilizer formalism, since the latter is non-contextual for $d \geq 3$.

Note also that while the NAND-gate has maximal degree and thus generates arbitrary Boolean functions $o : \mathbb{Z}_2^2 \to \mathbb{Z}_2$ under linear pre- and post-composition, this is not the case for the output function in Thm. 51. In fact, we will prove shortly that arbitrary functions can be implemented in flat, deterministic ld-MBQC, yet this requires certain number of qudits. In particular, three qutrits are not sufficient to implement arbitrary functions $o : \mathbb{Z}_3^2 \to \mathbb{Z}_3$. We study this relation in more detail in Sec. 3.3.

The previous example is easily generalised to d prime, thus yielding an Anders-Browne-type example for qudits of arbitrary prime dimension.

Theorem 52. *There exists a contextual flat, deterministic ld-MBQC, d prime, on three qudits.*

Proof. Define the following d operators of the form in Eq. (3.23) by their action on computational basis states $|q\rangle$ for $0 \leq q \leq d-1$ and $\omega = e^{\frac{2\pi i}{d}}$ as follows:

$$M(0)|q\rangle := X|q\rangle = |q+1\rangle, \qquad M(c)|q\rangle := \theta(c)\omega^{cq^{d-1}}|q+1\rangle, \quad 1 \leq c \leq d-1 \tag{3.27}$$

Note first that $M(c)^d = 1$ if we set $\theta(c)^d = \omega^c$ and thus $\theta(c_1)^d\theta(c_2)^d\theta(d-c_1-c_2)^d = 1$. There is more freedom in choosing $\theta(c)$ and we set $\theta(c) = e^{\frac{c2\pi i}{d^2}}$. Similarly to the qutrit case in Thm. 51, we also specify $\mathbb{Z}_d$-linear functions $l_1(\mathbf{i}) := i_1$, $l_2(\mathbf{i}) := i_2$, and $l_3(\mathbf{i}) := -i_1 - i_2$ for $\mathbf{i} = (i_1, i_2)^\intercal \in \mathbb{Z}_d^2$, and take the resource state to be

$$|\psi\rangle = \frac{1}{\sqrt{d}} \sum_{q=0}^{d-1} |q\rangle^{\otimes 3} .$$

The output function of this computation then has the following form:

$$o(\mathbf{i}) = \begin{cases} 0 & \text{if } i_1, i_2 = 0 \\ 1 & \text{if } i_1 + i_2 \leq d \\ 2 & \text{if } i_1 + i_2 > d \end{cases} \tag{3.28}$$

Again, by Thm. 48 this ld-MBQC is contextual if $o(\mathbf{i})$ is at least of degree d. To prove this, note that the number of monomials in i_1, i_2 of (combined) degree at most $d-1$ is the same as the number of constraints in Eq. (3.28) for $i_1 + i_2 \leq d-1$. The latter thus fix the former uniquely, resulting in $g = (i_1 + i_2)^{d-1}$ if the computation is non-contextual. However, g does not satisfy the constraints for $i_1 + i_2 \geq d$. Hence, the output function in $o(\mathbf{i})$ must contain at least one term of degree d or greater and is therefore contextual. □

Similarly to the output function in Thm. 51, it is not hard to see that $o(\mathbf{i})$ in Eq. (3.28) does not contain a term of maximal degree $2(d-1)$ either. In order to prove that indeed every output function can be implemented in flat, deterministic ld-MBQC, in the next two sections we explicitly compute the n-dimensional δ-function for qubits and qudits of prime dimension.

Implementation of the n-dimensional δ-function on $2^n - 1$ qubits

In this section we give an explicit implementation for the computation of the n-dimensional Dirac δ-function

$$\delta(\mathbf{x}) := \begin{cases} 1 & \text{if } \mathbf{x} = \mathbf{0} \\ 0 & \text{elsewhere} \end{cases} \qquad \mathbf{x} \in \mathbb{Z}_d^n \tag{3.29}$$

for $d = 2$ as a flat, deterministic $l2$-MBQC. Clearly, this implies that any function $o : \mathbb{Z}_2^n \to \mathbb{Z}_2$ can be implemented in flat, deterministic $l2$-MBQC. We make some preliminary remarks first.

Note that for qubits the only non-trivial dependency on c in $M(c) = X(f(c))$ from Eq. (3.22) is linear and therefore yields operators of the form in Eq. (3.23):

$$M(\theta)|q\rangle := \theta\chi^q|q+1\rangle = e^{\pi i\vartheta}\chi^q|q+1\rangle, \qquad \theta^2 = \chi^{-1} \tag{3.30}$$

In particular, we have $M(\theta)^2 = 1$ and the eigenstates of $M(\theta)$ read:

$$\begin{aligned} |\vartheta^+\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + e^{\pi i\vartheta}|1\rangle) & \qquad |0\rangle &= \frac{1}{\sqrt{2}}(|\vartheta^+\rangle + |\vartheta^-\rangle) \\ |\vartheta^-\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - e^{\pi i\vartheta}|1\rangle) & \qquad |1\rangle &= \frac{1}{\sqrt{2}}e^{-\pi i\vartheta}(|\vartheta^+\rangle - |\vartheta^-\rangle) \end{aligned}$$

Similarly to the Pauli-Y operator in the Anders and Browne example, rewriting the resource state in the basis $M(\theta)$, we pick up the phase $e^{-\pi i\vartheta}$ for every local $|1\rangle$ state in $|\psi\rangle$.

Theorem 53. *The n-dimensional δ-function in Eq. (3.29) for $d = 2$ can be implemented on $N = 2^n - 1$ qubits within flat, deterministic $l2$-MBQC.*

Proof. Consider the resource state in Eq. (3.19) for $N = 2^n - 1$ as well as linear functions

$$l_{\mathbf{a}}(\mathbf{i}) := \oplus_{j=1}^n a_j i_j \ , \qquad \mathbf{0} \neq \mathbf{a} \in \mathbb{Z}_2^n \ .$$

We prove that this indeed computes the desired function for a suitable $\theta_k = \theta = e^{\pi i\vartheta}$ under the measurement procedure $0 \to M(0)$ and $1 \to M(\theta)$ according to Eq. (3.30).

First, consider the case of the input string containing exactly one non-zero entry, e.g. $\mathbf{i} = (1, 0, \cdots, 0)^\intercal$, and count the number of phases ϑ that we collect. As ϑ is independent of the site, this is simply the number of functions that i_1 appears in. There is one function in which it appears by itself, then $n-1$ functions where it appears together with another input, $\binom{n-1}{2}$ in which it appears together with two more inputs and so on. Overall the number of functions is

$$\sum_{k=0}^{n-1} \binom{n-1}{k} = 2^{n-1} \ .$$

For inputs containing two non-zero entries, e.g. $\mathbf{i} = (1, 1, 0, \cdots, 0)^\intercal$, we again count the number of appearances of, in this case, i_1 and i_2. Note that only those functions will contribute that contain exactly one, but not both of those entries, i.e.,

$$\sum_{k=0}^{n-2} \binom{2}{1} \binom{n-2}{k} = 2 \cdot 2^{n-2} = 2^{n-1} \, .$$

The general case with m non-zero entries reads as follows:

$$\sum_{k=1}^{\lceil \frac{m}{2} \rceil} \binom{m}{2k-1} \sum_{l=0}^{n-m} \binom{n-m}{l} = 2^{m-1} \cdot 2^{n-m} = 2^{n-1}$$

Hence, for all but the zero input we flip the overall parity in Eq. (3.19) if we set

$$(e^{\pi i \vartheta})^{2^{n-1}} = -1 \iff \vartheta = \frac{1}{2^{n-1}} + K, \quad 2^{n-1}K = 0 \mod 2 \, . \tag{3.31}$$

Finally, note that this setup computes the function $o(\mathbf{i}) = \delta(\mathbf{i}) + 1$, hence, we obtain the n-dimensional Dirac δ-function by simple post-processing. $\square$

Thm. 53 immediately implies the following.

Corollary 1. *Any function $o : \mathbb{Z}_2^n \to \mathbb{Z}_2$ can be computed within flat, deterministic l2-MBQC.*

Proof. This follows directly from Thm. 53, and the fact that every function $o : \mathbb{Z}_2^n \to \mathbb{Z}_2$ can be written as a sum of n-dimensional δ-functions, i.e., $o(\mathbf{i}) = \sum_{\mathbf{j} \in \mathbb{Z}_2^n} C(\mathbf{j})\delta(\mathbf{i} - \mathbf{j})$ for $C(\mathbf{j}) \in \mathbb{Z}_2$. $\square$

Thm. 53 and Cor. 1 are a consequence—and, as we will see in Sec. 3.3.3, an alternative proof—of an earlier result in [82].

Theorem 54. (Hoban-Campbell [82]) *In order to implement the n-dimensional δ-function in flat, deterministic l2-MBQC one requires $2^n - 1$ qubits.*

Note that this is in stark contrast to the scaling behaviour under adaptive $l2$-MBQC (cf. Sec. 3.1.5). For instance, the naive protocol using iterative NAND-gates requires a linear number of adaptive steps in the degree $\deg(\delta) = n$, and so does the number of qubits necessary.

Implementation of the n-dimensional δ-function on $d^n - 1$ qudits

The construction in the last section can be generalised to systems of arbitrary prime dimension. First, it is useful to review different parity states for qudits. To this end, consider the qutrit example in Sec. 3.2.2 with resource state $|\psi\rangle = \frac{1}{\sqrt{3}}(|000\rangle + |111\rangle + |222\rangle)$ again. We rewrite this in the eigenbasis of the generalised X measurement with action $X|q\rangle = |q+1\rangle$:

$$\begin{aligned}
|\psi\rangle &= \frac{1}{\sqrt{3}}\left(|000\rangle + |111\rangle + |222\rangle\right) \\
&= \frac{1}{9}\left((|x^0\rangle + |x^1\rangle + |x^2\rangle)^{\otimes 3} + (|x^0\rangle + \omega|x^1\rangle + \omega^2|x^2\rangle)^{\otimes 3} + (|x^0\rangle + \omega^2|x^1\rangle + \omega|x^2\rangle)^{\otimes 3}\right) \\
&= \frac{1}{9}\left(\sum_{\mathbf{m}\in\mathbb{Z}_3^3} |x^{m_1}x^{m_2}x^{m_3}\rangle + \sum_{\mathbf{m}\in\mathbb{Z}_3^3} \omega^{\sum_{k=1}^3 m_k}|x^{m_1}x^{m_2}x^{m_3}\rangle + \sum_{\mathbf{m}\in\mathbb{Z}_3^3} \omega^{2\sum_{k=1}^3 m_k}|x^{m_1}x^{m_2}x^{m_3}\rangle\right) \\
&= \frac{1}{9}\left(\sum_{\mathbf{m}\in\mathbb{Z}_3^3} (1 + \omega^{\sum_{k=1}^3 m_k} + \omega^{2\sum_{k=1}^3 m_k})|x^{m_1}x^{m_2}x^{m_3}\rangle\right) \\
&= \frac{1}{3}\left(\sum_{\substack{\mathbf{m}\in\mathbb{Z}_3^3,\\ \oplus_{k=1}^3 m_k = 0}} |x^{m_1}x^{m_2}x^{m_3}\rangle\right)
\end{aligned}$$

Hence, $|\psi\rangle$ has parity $\oplus_{k=1}^3 m_k = 0$ in this representation. For deterministic ld-MBQC we need to choose local measurements such that the resource state is a state of certain parity for every input when rewritten in the corresponding local bases. For instance, note that in the qutrit example in Thm. 51, any other choice of (local) measurements results in a similar decomposition with the difference that we pick up some additional phase factors,

$$|\psi\rangle = \frac{1}{9}\left(\sum_{\mathbf{m}\in\mathbb{Z}_3^3} (1 + \theta_1^{-n_1}\theta_2^{-n_2}\omega^{\sum_{k=1}^3 m_k} + \theta_1^{-2n_1}\theta_2^{-2n_2}\omega^{-n_1-2n_2+2\sum_{k=1}^3 m_k})|u^{m_1}u^{m_2}u^{m_3}\rangle\right) .$$

Here, n_1 and n_2 denote the number of measurements $M(1)$ and $M(2)$, respectively, and $u^{m_k} \in \{x^{m_k}(=\theta_0^{m_k}), \theta_1^{m_k}, \theta_2^{m_k}\}$. Given our choice of functions, we have $n_1 + n_2 = 0$, $n_1 = 3$, or $n_2 = 3$, which reproduces the output function in Sec. 3.2.2.

By allowing for more general phase relations than in the local measurement operators in

Eq. (3.27) acting on the N-qudit GHZ-state

$$|\psi\rangle = \frac{1}{\sqrt{d}}\sum_{q=0}^{d-1}|q\rangle^N = \frac{1}{\sqrt{d}}\sum_{q=0}^{d-1}\otimes_{k=1}^N|q_k = q\rangle \ , \tag{3.32}$$

a similar line of reasoning proves the following.

Theorem 55. *The n-dimensional δ-function in Eq. (3.29) for d prime can be implemented on $N = d^n - 1$ qudits within flat, deterministic ld-MBQC.*

Proof. The proof strategy is similar to the qubit case in Thm. 53, see App. 3.A for details. □

Note that the phase relations needed for this construction read

$$\chi = \omega^{-\frac{2}{d^{n-1}(d-1)}}, \qquad \theta(c) = \chi^{-\frac{(d-1)}{d}c} = \omega^{\frac{2c}{d^n}} \ . \tag{3.33}$$

In particular, for qubits we have $\omega = e^{\frac{2\pi i}{2}} = -1$. We thus recover the Anders-Browne example for $n = 2$ with $\chi = -1$, $\theta = \theta(1) = \chi^{-\frac{1}{2}} = i$, and $X(\theta) = Y$, as well as the phase relation for the n-dimensional qubit δ-function in Thm. 53, where $\theta = e^{\pi i\vartheta} = (-1)^{\frac{1}{2^{n-1}}}$ as in Eq. (3.31).

Importantly, the factor 2 in Eq. (3.33) implies a crucial difference between the qubit and qudit case: in general, the phase $\theta(c)$ is a d^{n+1}-th root of unity, however, for $d = 2$, θ is only a d^n-th root of unity by this additional factor. The contextual signature of the qubit stabiliser formalism somehow resides in this phase, in particular, it implies that any quadratic function can be computed within the qubit stabilizer formalism already. This is in contrast to the qudit stabiliser formalism, which is non-contextual [64, 70].

Finally, recall that the n-dimensional δ-function has polynomial representation $\delta(\mathbf{x}) = \prod_{k=1}^n(1 - x_k^{p-1})$ (cf. Eq. (3.12)), which contains a term of maximal degree. Hence, we have constructed a (maximally) contextual example for qudits in arbitrary prime dimension according to the threshold in Thm. 48. In particular, we have established the following.

Corollary 2. *Any function $o : \mathbb{Z}_d^n \to \mathbb{Z}_d$ for d prime can be computed within flat, deterministic ld-MBQC.*

Proof. This follows directly from Thm. 55, and the fact that every function $o : \mathbb{Z}_d^n \to \mathbb{Z}_d$ can be written as a sum of n-dimensional δ-functions, i.e., $o(\mathbf{i}) = \sum_{\mathbf{j}\in\mathbb{Z}_d^n} C(\mathbf{j})\delta(\mathbf{i} - \mathbf{j})$ for $C(\mathbf{j}) \in \mathbb{Z}_d$. □

Note that the number of qudits in the implementation of the δ-function in Thm. 55 counts $d^n - 1$, which generalises the qubit case, where $2^n - 1$ qubits are optimal by Thm. 54. As a means to further distinguish contextuality beyond the threshold in Thm. 48, we explore the question of optimality in more detail in the next section, i.e., we ask for the minimal number of qudits required to implement a given function in flat, deterministic ld-MBQC.

3.3 Towards a classification of contextuality in ld-MBQC

In this section we study computation in flat, deterministic ld-MBQC from a resource-theoretic perspective. In the last section we proved that any function $o : \mathbb{Z}_d^n \rightarrow \mathbb{Z}_d$ for d prime can be computed within flat, deterministic ld-MBQC. However, depending on the type of function computed, the number of qudits necessary for its implementation varies drastically. We therefore further study the structure of contextuality by asking for the optimal number of qudits needed to implement a function $o : \mathbb{Z}_d^n \rightarrow \mathbb{Z}_d$ in flat, deterministic ld-MBQC.

Definition 37. *Let $o : \mathbb{Z}_d^n \rightarrow \mathbb{Z}_d$, d prime. We call a flat ld-MBQC, which deterministically implements o* optimal*, if no other flat ld-MBQC exists, which deterministically implements o on fewer qudits. The optimal number of qudits is denoted $R(o)$.*

Note first that we have the freedom to manipulate any output function o by invertible linear transformations on the inputs via pre-processing. The resource cost $R(o)$ should thus be an invariant under affine transformations. For this reason we define an equivalence relation on all functions with signature $o : \mathbb{Z}_d^n \rightarrow \mathbb{Z}_d$ under affine transformations as follows:

$$o \sim o' :\Longleftrightarrow \exists M \in \mathrm{Mat}(n \times n, \mathbb{Z}_d),\ \mathrm{rk}(M) = n :\ o'(\mathbf{i}) = o(M\mathbf{i}) \tag{3.34}$$

Furthermore, in Sec. 3.2.2 we have seen how the n-dimensional δ-function can be implemented as a flat, deterministic ld-MBQC on $N = d^n - 1$ qudits. Hence, given an arbitrary function $o : \mathbb{Z}_d^n \rightarrow \mathbb{Z}_d$, one way to implement it is by naively adding all terms in the sum $o(\mathbf{i}) = \sum_{\mathbf{j} \in \mathbb{Z}_d^n} C(\mathbf{j})\delta(\mathbf{i} - \mathbf{j})$, $C(\mathbf{j}) \in \mathbb{Z}_d$. However, it is easy to see that the optimal number of qudits is only subadditive in this as well as its polynomial representation. It follows that in contrast to the computational bound in Thm. 48, which emerged from considering subspaces under linear pre- and post-composition, there is more to contextuality beyond that threshold.

In order to gain some intuition for the actual behaviour of $R(o)$, we first consider two specific quadratic Boolean functions in Sec. 3.3.1. For the general case, we study the map between different representations of functions in Sec. 3.3.2. In doing so we prove that our construction method for $l2$-MBQC in terms of the phase relations in Eq. (3.21) is generic (in the qubit case), we reproduce some known bounds for R obtained in previous sections, and connect its optimisation problem to known problems in circuit synthesis.

3.3.1 Examples

The computation of the NAND-gate in [10] requires three qubits and is optimal according to Thm. 54. For the output function $o : \mathbb{Z}_2^4 \to \mathbb{Z}_2$, $o(\mathbf{i}) = i_1 i_2 + i_3 i_4$ one can thus easily construct an $l2$-MBQC that requires six qubits. However, this function has an optimal cost of $R = 5$. To see this consider local phases $\theta_1 = \theta_2 = -\theta_3 = -\theta_4 = \theta_5 = i$ and linear functions:

$$
\begin{aligned}
&l_1(\mathbf{i}) := i_1 && l_2(\mathbf{i}) := i_2 && \\
&l_3(\mathbf{i}) := i_1 \oplus i_2 \oplus i_3 && l_4(\mathbf{i}) := i_1 \oplus i_2 \oplus i_4 && l_5(\mathbf{i}) := i_1 \oplus i_2 \oplus i_3 \oplus i_4
\end{aligned}
$$

As before we may implement this computation as a flat $l2$-MBQC using the generalised GHZ-state $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle^5 + |1\rangle^5)$ as resource, and local observables given by $c_k = 0 : M_k(\theta_k = 1)$, $c_k = 1 : M_k(\theta_k = \pm i)$ (cf. Eq. (3.30)). A straightforward computation then yields the output:

$$
\begin{array}{llll}
(0,0,0,0)^\intercal \longrightarrow 0 & (1,0,0,0)^\intercal \longrightarrow 0 & (1,0,0,1)^\intercal \longrightarrow 0 & (1,1,1,0)^\intercal \longrightarrow 1 \\
(1,1,0,0)^\intercal \longrightarrow 1 & (0,1,0,0)^\intercal \longrightarrow 0 & (0,1,1,0)^\intercal \longrightarrow 0 & (1,1,0,1)^\intercal \longrightarrow 1 \\
(0,0,1,1)^\intercal \longrightarrow 1 & (0,0,1,0)^\intercal \longrightarrow 0 & (1,0,1,0)^\intercal \longrightarrow 0 & (1,0,1,1)^\intercal \longrightarrow 1 \\
(1,1,1,1)^\intercal \longrightarrow 0 & (0,0,0,1)^\intercal \longrightarrow 0 & (0,1,0,1)^\intercal \longrightarrow 0 & (0,1,1,1)^\intercal \longrightarrow 1
\end{array}
$$

This is easily seen to reproduce the function $o(\mathbf{i}) = i_1 i_2 + i_3 i_4$, since $\prod_{k=1}^5 \theta^{l_k(\mathbf{i})} = (-1)^{o(\mathbf{i})}$.

As a second example, consider quadratic functions $\Sigma_2^n(x) = \sum_{i_1 < i_2, i_j \in \{1,\cdots,n\}} x_{i_1} x_{i_2}$. These have been discussed in [82], where it was proven that $R(\Sigma_2^n) = n + 1$. As with the first example, this shows that R is subadditive (in its polynomial representation), which rules out a straightforward computation, e.g. in terms of the δ-functions in Sec. 3.2.2, which are provably optimal. This additional complexity arises from an ambiguity in the representation of functions in terms of $\mathbb{Z}_d$-linear functions over the reals instead of polynomials over finite fields.

3.3.2 Polynomial vs $\mathbb{Z}_d$-linear function representation

In the discussion so far, we have encountered two representations for functions $o : \mathbb{Z}_d^n \rightarrow \mathbb{Z}_d$. In the case $d = 2$, i.e., for Boolean functions $f : \mathbb{Z}_2^n \rightarrow \mathbb{Z}_2$, they read

$$f(x) = \sum_{\mathbf{a}\in\mathbb{Z}_2^n} C_{\mathbf{a}} \left(\oplus_{j=1}^n a_j x_j\right) = \sum_{\mathbf{b}\in\mathbb{Z}_2^n} C_{\mathbf{b}} \left(\prod_{j=1}^n x_j^{b_j}\right) .$$

The latter is the polynomial representation of Thm. 47 with coefficients $C_{\mathbf{b}} \in \mathbb{Z}_2$, whereas the former representation is based on $\mathbb{Z}_2$-linear functions with real coefficients $C_{\mathbf{a}} \in \mathbb{R}$. Both sets of functions are linearly independent and generate the space of Boolean functions on bitstrings $\mathbf{x} \in \mathbb{Z}_2^n$. Hence, there exists a corresponding transformation between the coefficients $C_{\mathbf{a}}$, $C_{\mathbf{b}}$.

Note that every $\mathbb{Z}_2$-linear function can be written in terms of monomials since

$$\oplus_{j=1}^n x_j = \sum_{\mathbf{0}\neq\mathbf{b}\in\mathbb{Z}_2^n} (-2)^{W(\mathbf{b})-1} \prod_{j=1}^n x_j^{b_j} , \tag{3.35}$$

where $W(\mathbf{b}) := \sum_{j=1}^n b_j$ denotes the Hamming weight of $\mathbf{b} \in \mathbb{Z}_2^n$ (cf. [82]). More generally, we define $l_{\mathbf{a}} := \oplus_{j=1}^n a_j x_j$, $f_{\mathbf{b}} := 2^{W(\mathbf{b})-1} \prod_{j=1}^n x_j^{b_j}$ for $\mathbf{0} \neq \mathbf{a}, \mathbf{b} \in \mathbb{Z}_2^n$ and $l_{\mathbf{0}} = f_{\mathbf{0}} := 1$, as well as

$$\langle l_{\mathbf{a}}, f_{\mathbf{b}} \rangle := (-1)^{\sum_{j=1}^n a_j b_j - 1} , \qquad \langle l_{\mathbf{0}}, f_{\mathbf{0}} \rangle := 1 . \tag{3.36}$$

We then have the following map $\mathcal{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ between representations $f = \sum_{\mathbf{a}\in\mathbb{Z}_2^n} C_{\mathbf{a}} l_{\mathbf{a}} = \sum_{\mathbf{b}\in\mathbb{Z}_2^n} C_{\mathbf{b}} (\frac{1}{2^{W(\mathbf{b})-1}} f_{\mathbf{b}})$:

$$\forall \mathbf{a} \in \mathbb{Z}_2^n : \ \mathcal{F}(l_{\mathbf{a}}) := \sum_{\mathbf{b}\in\mathbb{Z}_2^n} \langle l_{\mathbf{a}}, f_{\mathbf{b}} \rangle f_{\mathbf{b}} \tag{3.37}$$

$\mathcal{F}$ is a real-linear map, it can be represented as a matrix with entries $\mathcal{F}_{l_{\mathbf{a}} f_{\mathbf{b}}} = \langle l_{\mathbf{a}}, f_{\mathbf{b}} \rangle = \pm 1$ (cf. Eq. (3.36)). In fact, for fixed n and with appropriate normalisation factor $N = 2^{-\frac{n}{2}}$, $\mathcal{F}$ is a Hadamard transform and thus in particular unitary, from which it follows that $(N\mathcal{F}_{l_{\mathbf{a}} f_{\mathbf{b}}})^{-1} = N\mathcal{F}_{f_{\mathbf{b}} l_{\mathbf{a}}} = N\mathcal{F}_{l_{\mathbf{a}} f_{\mathbf{b}}}$. Clearly, this generalises Eq. (3.35) and provides an explicit transformation between the two representations of Boolean functions underlying the $l2$-MBQC in Eq. (3.21).[8]

[8]Note the close relationship between $\mathcal{F}$ and the (multi-dimensional) discrete Fourier transform, as well as the latter's importance in existing quantum algorithms, e.g. Grover's and Shor's algorithm.

Note that a similar construction can be given in the qudit case with d prime, too. It is not hard to see that next to the set of monomials (by Thm. 47), also the set of d^n functions $l_{\mathbf{a}} := \oplus_{j=1}^n a_j x_j$, $\mathbf{a} \in \mathbb{Z}_d^n$ generates the space of functions $f : \mathbb{Z}_d^n \to \mathbb{Z}_d$. In particular, note that by the proof of Thm. 55, the Dirac δ-function has the following representation:

$$\delta(\mathbf{x}) = \prod_{j=1}^n \delta(x_j) = \sum_{\mathbf{a} \in \mathbb{Z}_d^n} \frac{2}{d^n} \left(\bigoplus_{j=1}^n a_j x_j \right) + 1 \tag{3.38}$$

We can thus construct a linear transformation between the coefficients $C_{\mathbf{b}} \in \mathbb{Z}_d$ over elements $\prod_{j=1}^n x_j^{b_j}$, $\mathbf{b} \in \mathbb{Z}_2^n$ in the polynomial representation, and the coefficients $C_{\mathbf{a}} \in \mathbb{R}$ over elements $l_{\mathbf{a}}$ in the $\mathbb{Z}_d$-linear representation, similar to the case $d = 2$.

The transformation $\mathcal{F}$ has a number of consequences for the implementation of temporally-flat, deterministic ld-MBQC, which we will exploit in the following sections.

3.3.3 Phase relations in flat, deterministic ld-MBQC

Given the N-qubit GHZ-state in Eq. (3.19) and operators in Eq. (3.30), different choices of local measurements simply translate into phase relations according to Eq. (3.21). The latter is a representation of the output function $o(\mathbf{i})$ in terms of $\mathbb{Z}_2$-linear functions $l_k(\mathbf{i})$. The local phases $\theta_k := e^{\pi i \vartheta_k}$ thus implement the coefficients $\vartheta_k = C_{\mathbf{a}} \in \mathbb{R}$ under the mapping $\mathcal{F}^{-1}(o(\mathbf{i}))$.[9] More generally for d prime, we have the following phase constraints.

Proposition 9. *In flat, deterministic ld-MBQC for d prime, and with local measurement operators in Eq. (3.22) acting on a GHZ resource state, the output function arises from the phase relations between the eigenbases of the local measurement operators as follows*

$$\prod_{1 \leq k \leq N} \left(\prod_{q'=0}^{q-1} f_k(c_k)(q') \right) = \omega^{q o(\mathbf{i})} \,. \tag{3.39}$$

Proof. We give the details in App. 3.B. □

[9] The coefficients $\vartheta_k \in \mathbb{R}$ in the $\mathbb{Z}_2$-linear representation of the Boolean function $o : \mathbb{Z}_2^n \to \mathbb{Z}_2$ are also known as the Walsh spectrum of o.

Note that Eq. (3.39) encodes constraints for every value of $q \in \mathbb{Z}_d$, where the case $q = 0$ is trivially satisfied. This leaves a single constraint in the qubit case, which reads $\prod_{1\leq k\leq N} \theta_k(l_k(\mathbf{i})) = \prod_{1\leq k\leq N} \theta_k^{l_k(\mathbf{i})} = e^{\pi i (\sum_{k=1}^{N} l_k(\mathbf{i})\vartheta_k)} = \omega^{o(\mathbf{i})}$ by Eq. (3.30),[10] and reproduces Eq. (3.21) for qubits. We thus find that our initial strategy of constructing explicit $l2$-MBQCs is generic. In particular, the transformation formula in Eq. (3.37), which translates the polynomial representation of the output function $o: \mathbb{Z}_2^n \rightarrow \mathbb{Z}_2$ into an equivalent representation in terms of fractions ϑ_k, applies to $l2$-MBQC via phases $\theta_k = e^{\pi i \vartheta_k} = \frac{1}{f_k(1)(0)}$ between the eigenbases of local measurement operators in Eq. (3.30). On the other hand, for qudits phase relations in Eq. (3.39) are more general than the constraints in Eq. (3.21). For instance, note the quadratic dependency on linear functions in the phase relations arising in the qudit stabiliser formalism (cf. Eq. (3.8) and Eq. (3.11)).

Eq. (3.37) has a number of consequences. For instance, the (polynomial) degree of the output function $o: \mathbb{Z}_2^n \rightarrow \mathbb{Z}_2$ places a lower bound on the coefficients $\min_k \vartheta_k \leq \frac{1}{2^{\deg(o)-1}}$ required for the implementation as $l2$-MBQC, from which it follows that only quadratic functions can be computed within the qubit stabilizer formalism, where phases arise as powers of $i = \theta = e^{\pi i \frac{1}{2}}$, hence, $\vartheta_k \geq \frac{1}{2^{2-1}}$. In the next sections we show how Eq. (3.37) also allows us to bound the number of qubits required for flat, deterministic $l2$-MBQC.

Optimal representation of functions in flat $l2$-MBQC

By means of the Hadamard transform in Eq. (3.37), given a polynomial output function, we always find an implementation as $l2$-MBQC. As we will see in the next sections, for monomials and other highly symmetric functions, this representation is optimal in the number of non-zero coefficients of its $\mathbb{Z}_2$-linear representation (and thus in the number of qubits in the implementation as $l2$-MBQC), whereas for more general functions this is not the case.

The ambiguity underlying such representations stems from the fact that given a Boolean function f, we may always add even multiples of other Boolean functions. Clearly, this does not change f in its polynomial representation in Thm. 47, however, it potentially changes the representation in terms of $\mathbb{Z}_2$-linear functions and thus the implementation as $l2$-MBQC. For instance, note that the representation of the output function $o: \mathbb{Z}_2^4 \rightarrow \mathbb{Z}_2$, $o(\mathbf{i}) = i_1 i_2 + i_3 i_4$ in

[10] Note that we set local reference measurements to correspond with the generalised Pauli-X operator, for which $\theta = 1$, and whose eigenstates define the reference basis of Lm. 4.

terms of $\mathbb{Z}_2$-linear functions in Sec. 3.3.1 arises by subtracting the 'zero term' $z = 4i_1i_2i_3i_4 - 2i_1i_2(i_3 + i_4 - 1) \in Z(o)$, i.e., it corresponds to the non-zero coefficients under $\mathcal{F}^{-1}(o - z)$.

More generally, let $f : \mathbb{Z}_2^n \to \mathbb{Z}_2$ be a Boolean function and define the linear space of *zero terms* as

$$Z(f) = \langle 2^m g \mid g : \mathbb{Z}_2^n \to \mathbb{Z}_2, n \geq m \geq 1 \rangle \ . \tag{3.40}$$

In addition to the linear equivalence relation in Eq. (3.34), we thus face the following problem.

Proposition 10. *The minimal number of qubits $R(o)$ required to implement a given polynomial output function $o : \mathbb{Z}_2^n \to \mathbb{Z}_2$ in flat, deterministic l2-MBQC is the minimal number of non-zero coefficients $C_{\mathbf{a}}$ of $\mathcal{F}^{-1}(o)$ in Eq. (3.37) under the equivalence relation $o \sim o' \iff o' = o + z$, $z \in Z(o)$ of Eq. (3.40).*

Proof. This follows immediately by the above discussion and the transformation in Eq. (3.37). $\square$

Importantly, Prop. 10 is responsible for the complexity of computing the optimal number of qubits in flat, deterministic $l2$-MBQC. In particular, we find a rich structure of contextuality beyond the mere distinction between contextual and non-contextual computation in Thm. 48.

Moreover, a similar relation also exists for $d > 2$, however, since $\mathcal{F}^{-1}$ only applies to operators with linear phase relations in Eq. (3.39), ld-MBQCs with fewer qudits than $R(o)$ might exist in that case.

It is interesting to note that similar minimisation problems to Prop. 10 also arise in circuit synthesis, e.g. the minimal number of T-gates can be related to the minimal number of $\mathbb{Z}_2$-linear functions with odd coefficients [27, 80]. Solving the latter further relates to minimum distance decoding in punctured Reed-Muller codes, which seems hard in general [9]. While our problem is slightly different—we are interested in all, not just odd terms in the $\mathbb{Z}_2$-linear representation—numerical calculation suggests that a straightforward extrapolation from known cases remains difficult.

Nevertheless, for certain functions the complexity of R under the equivalence relation in Eq. (3.40) simplifies. In the final two sections, we identify two such cases, which allows us to find the optimal implementation of monomials in MBQC, and provide an upper bound to elementary symmetric functions.

Optimal implementation of monomials as $l2$-MBQC

Given a general output function in its polynomial representation $o(\mathbf{i})$, we may use the transformation $\mathcal{F}^{-1}$ in Eq. (3.37) to obtain a representation in terms of $\mathbb{Z}_2$-linear functions and thus study the minimal number of qubits required for implementation as flat, deterministic $l2$-MBQC. In fact, for monomials the decomposition under $\mathcal{F}^{-1}$ is already optimal since there are no zero terms in Eq. (3.40), which could affect the minimisation in Prop. 10.

Theorem 56. *In order to implement the monomial $f : \mathbb{Z}_2^n \rightarrow \mathbb{Z}_2$, $f(x) = \prod_{j=1}^n x_j$ in flat, deterministic $l2$-MBQC one requires $N = 2^n - 1$ qubits.*

Proof. Note that f has degree $\deg(f) = n = W(\mathbf{b})$ for $\mathbf{b} = (1)^n := (1, \cdots, 1) \in \mathbb{Z}_2^n$, hence, by Eq. (3.37) it has coefficient $\frac{1}{2^{W(\mathbf{b})-1}} = \frac{1}{2^{n-1}}$. Explicitly, the coefficients in the $\mathbb{Z}_2$-linear representation under the transformation $\mathcal{F}^{-1}$ read:

$$\begin{aligned}
\mathcal{F}^{-1}(\prod_{j=1}^n x_j) &\overset{\mathbf{b}=(1)^n}{=} \mathcal{F}(\prod_{j=1}^n x_j^{b_j}) = \mathcal{F}(\frac{1}{2^{W(\mathbf{b})-1}} f_\mathbf{b}) \\
&= \frac{1}{2^{W(\mathbf{b})-1}} \sum_{\mathbf{a}\in\mathbb{Z}_2^n} \langle f_\mathbf{b}, l_\mathbf{a} \rangle l_\mathbf{a} = \frac{1}{2^{W(\mathbf{b})-1}} \sum_{\mathbf{a}\in\mathbb{Z}_2^n} (-1)^{1-W(\mathbf{a})} \oplus_{j=1}^n a_j x_j
\end{aligned}$$

Since these terms are all odd multiples of $\frac{1}{2^{W(\mathbf{b})-1}}$, they can only be reduced by a zero term of degree at least n, however, there are no such terms in $Z(f)$, hence, the representation under the transformation $\mathcal{F}^{-1}$ is already optimal. Finally, note that the overlap with $l_\mathbf{a}$, $\mathbf{a} = 0$ can be implemented by post-processing, leaving $N = 2^n - 1$ non-zero terms. $\square$

Note also that the n-dimensional δ-function arises from monomials by linear pre-composition in Eq. (3.34), hence, $R(\delta) = 2^n - 1$, which reproduces the bound in Thm. 54.

A similar computation for qudits yields the representation in Eq. (3.38). Yet again, there the map $\mathcal{F}^{-1}$ only applies to measurement operators for which the phase constraints in Eq. (3.39) depend linearly on the inputs, i.e., if there exist phases $\Xi_k(q)$ such that $f_k(c_k)(q) = \Xi_k(q)^{c_k}$.[11]

[11] Note that the operators in Eq. (3.41) in the proof of Thm. 55 are of this form, namely $f_k(c_k)(q) = \chi_k^{c_k(q^{d-1} - \frac{d-1}{d})}$ for $1 \leq q \leq d-1$.

Elementary symmetric functions

While for monomials the transformation in Eq. (3.37) is already optimal in the number of non-zero coefficients (and thus in the number of qubits in the implementation as $l2$-MBQC), this is no longer the case for more general polynomials. Nevertheless, for certain symmetric functions the minimisation problem for R in Prop. 10 under the equivalence relation in Eq. (3.40) simplifies.

As an example of this case, we consider elementary symmetric functions,

$$\Sigma_k^n(\mathbf{x}) \;=\; \sum_{\substack{i_1<\cdots<i_k \\ i_j\in\{1,\cdots,n\}}} x_{i_1}\cdots x_{i_k}, \qquad k\leq n\;.$$

Note that the multiple AND-function in [82] corresponds to the case $k=2$. Plugging Σ_k^n into the transformation in Eq. (3.37) results in a total number of terms equal to $\sum_{l=1}^{k}\binom{n}{l}$. However, we can minimise this number by (at least) $\binom{n}{k}-1$ as follows. We add the zero term $z\in Z(\Sigma_k^n)$,

$$\begin{aligned} z &= (-2)^{n-k}x_1\cdots x_n + (-2)^{n-k-1}\sum_{\substack{i_1<\cdots<i_{n-1} \\ i_j\in\{1,\cdots,n\}}} x_{i_1}\cdots x_{i_{n-1}} + \cdots + (-2)\sum_{\substack{i_1<\cdots<i_{k+1} \\ i_j\in\{1,\cdots,n\}}} x_{i_1}\cdots x_{i_{k+1}} \\ &= \sum_{l=0}^{n-k-1}(-2)^{n-k-l}\sum_{\substack{i_1<\cdots<i_{n-l} \\ i_j\in\{1,\cdots,n\}}} x_{i_1}\cdots x_{i_{n-l}}\;. \end{aligned}$$

By construction, Σ_k^n and z have the same (smallest) coefficient $\frac{1}{2^{k-1}}$, and we can thus compare the coefficients in their representation based on $\mathbb{Z}_2$-linear functions $l_{\mathbf{a}}$, $\mathbf{a}\in\mathbb{Z}_2^n$. Clearly, Σ_k^n+z contains the term $x_1\oplus\cdots\oplus x_n$ and thus $C_{W(\mathbf{a})=n}^{\Sigma_k^n+z}=\frac{(-1)^{k-1}}{2^{k-1}}$. For the terms of length $k\leq m<n$, the coefficients $C_{W(\mathbf{a})=m}^{\Sigma_k^n+z}$ contain contributions from all higher degree terms in the polynomial representation of Σ_k^n+z:

$$\begin{aligned} C_{W(\mathbf{a})=m}^{\Sigma_k^n+z} &= \frac{1}{2^{k-1}}(-1)^{(n-k)+(m-1)}\left(1-\binom{n-m}{n-m-1}+\binom{n-m}{n-m-2}-\cdots+(-1)^{n-m}\right) \\ &= \frac{1}{2^{k-1}}(-1)^{(n-k)+(m-1)}\left(\sum_{l=0}^{n-m}(-1)^l\binom{n-m}{n-m-l}\right)=0 \end{aligned}$$

Hence, with respect to monomials of degree $k \leq m$ in $\Sigma_k^n + z$, we have reduced the overall number of non-zero coefficients by $\binom{n}{k} - 1$. Note also that the coefficients of the remaining monomials of degree $1 \leq m < k$ are non-zero since there, the above sum is truncated and reads

$$\begin{aligned} C_{W(\mathbf{a})=m}^{\Sigma_k^n+z} &= \frac{1}{2^{k-1}}(-1)^{(n-k)+(m-1)}\left(1 - \binom{n-m}{n-m-1} + \binom{n-m}{n-m-2} - \dots + (-1)^{n-k}\binom{n-m}{k-m}\right) \\ &= \frac{1}{2^{k-1}}(-1)^{(n-k)+(m-1)}\left(\sum_{l=0}^{n-k}(-1)^l\binom{n-m}{n-m-l}\right) , \end{aligned}$$

thus leaving a total of $\sum_{l=1}^{k-1}\binom{n}{l} + 1$ terms in the $\mathbb{Z}_2$-linear representation in general. Still, numerical tests show that this number is only suboptimal, i.e., $R(\Sigma_k^n) \leq \sum_{l=1}^{k-1}\binom{n}{l} + 1$. Nevertheless, note the following: (i) $R(\Sigma_n^n) = \sum_{l=1}^{n-1}\binom{n}{l} + 1 = 2^n - 1$ reproduces the optimal number of qubits within $l2$-MBQC for monomials in Thm. 56, (ii) $R(\Sigma_2^n) = \sum_{l=1}^{2-1}\binom{n}{l} + 1 = n + 1$ reproduces the bound on qubits in Prop. 2 in [82], and (iii) we find cases such as $R(\Sigma_2^7) > R(\Sigma_3^3)$ despite $\deg(\Sigma_3^3) = 3 > 2 = \deg(\Sigma_2^7)$, which implies that—unlike the contextuality threshold in Thm. 48—the (polynomial) degree alone is not sufficient to compare functions with respect to their optimal representation in flat, deterministic ld-MBQC, thus confirming that the computational classification of contextuality has a richer substructure beyond the non-contextual case.

3.4 Summary

In this chapter we have studied contextuality in measurement-based computation. Based on the general framework in Def. 35 (cf. [10]), in Sec. 3.1 we first considered correlations in general theories and proved a strict bound on the space of functions computable in non-contextual measurement-based computation with $\mathbb{F}_d$-linear side-processing (ld-MBC), which generalises a previous result for $d = 2$ in [126].

In Sec. 3.2 we refined this bound by considering measurement-based quantum computation [21, 128] with $\mathbb{Z}_d$-linear side-processing (ld-MBQC) explicitly. In particular, building on and generalising results in [82, 126, 146], we proved that any output function $o : \mathbb{Z}_d^n \to \mathbb{Z}_d$ can be deterministically computed already in the non-adaptive case. We highlighted that a crucial resource in this setting is the number of qudits required for implementation and linked it to known hard problems in circuit synthesis. The key message is that (at least) for qubits, any $l2$-MBQC implements a given function in terms of $\mathbb{Z}_2$-linear functions with real-valued coefficients, which arise as phase relations between eigenvectors of local measurement operators. For qudits, the same construction applies, yet represents a special case of ld-MBQC only, thus potentially allowing for implementations requiring fewer than the number of non-zero terms under the discrete Fourier transform in Eq. (3.37). Nevertheless, even for qudits contextuality in ld-MBQC is closely related with the phase relations between local sites according to the more general relation in Eq. (3.39). In this sense, the phase relations between measurement operators in Eq. (3.22) fully classify flat, deterministic ld-MBQC.

We have considered measurement-based quantum computation since the resource character of nonlocality and contextuality exhibits rather clearly in this quantum computing architecture. In particular, the restriction on $\mathbb{Z}_d$-linear side-processing allows to state the quantum advantage in terms of the clear-cut complexity-theoretic difference of Thm. 48. While interesting applications might be found within this framework, the most interesting scenario will likely arise from considering universal classical side-processing.

3.A Proof of Theorem 55

We choose measurement operators of the form in Eq. (3.23) for prime dimension d as follows:

$$M(0)|q\rangle := X|q\rangle = |q+1\rangle, \qquad M(c)|q\rangle := \theta(c)\chi^{cq^{d-1}}|q+1\rangle, \quad 1 \leq c \leq d-1 \tag{3.41}$$

Then $M(c)^d = 1$ if we set $\theta(c)^d = \chi^{-(d-1)c}$. We find the following eigenstates

$$|\theta(c)^m\rangle = \frac{1}{\sqrt{d}}\sum_{q=0}^{d-1}(\omega^{-m}\theta(c))^q(\chi^c)^{(q-1)q^{d-1}}|q\rangle, \quad \forall c \in \mathbb{Z}_d$$

with corresponding expressions in terms of computational basis states,

$$|q\rangle = \frac{1}{\sqrt{d}}\frac{1}{\theta(c)^q\chi^{c(q-1)q^{d-1}}}\sum_{m=0}^{d-1}\omega^{qm}|\theta(c)^m\rangle, \quad \forall c \in \mathbb{Z}_d\ . \tag{3.42}$$

Consider the (d^n-1)-qudit resource state $|\psi\rangle = \frac{1}{\sqrt{d}}\sum_{q=0}^{d-1}|q\rangle^{\otimes d^n-1} = \frac{1}{\sqrt{d}}\sum_{q=0}^{d-1}\otimes_{k=1}^{d^n-1}|q_k = q\rangle$ and assume that the output function $o : \mathbb{Z}_d^n \to \mathbb{Z}_d$ is encoded in the phase relations as follows

$$\prod_{1\leq k\leq d^n-1}\theta(c_k(\mathbf{i}))^q\chi^{c_k(\mathbf{i})(q-1)q^{d-1}} = \omega^{qo(\mathbf{i})}\ . \tag{3.43}$$

Rewriting $|\psi\rangle$ in terms of the local measurement bases via Eq. (3.42) yields

$$\begin{aligned}
|\psi\rangle &= \frac{1}{\sqrt{d}}\sum_{q=0}^{d-1}\otimes_{k=1}^{d^n-1}\left(\frac{1}{\sqrt{d}}\frac{1}{\theta(c_k(\mathbf{i}))^q\chi^{c_k(\mathbf{i})(q-1)q^{d-1}}}\sum_{m_k=0}^{d-1}\omega^{qm_k}|\theta(c_k(\mathbf{i}))^{m_k}\rangle\right)\\
&= d^{\frac{-d^n}{2}}\sum_{q=0}^{d-1}\omega^{-qo(\mathbf{i})}\left(\sum_{\mathbf{m}\in\mathbb{Z}_d^{d^n-1}}\otimes_{k=1}^{d^n-1}\omega^{qm_k}|\theta(c_k(\mathbf{i}))^{m_k}\rangle\right)\\
&= d^{\frac{-d^n}{2}}\sum_{q=0}^{d-1}\left(\sum_{\mathbf{m}\in\mathbb{Z}_d^{d^n-1}}\omega^{q(\sum_{k=1}^{d^n-1}m_k-o(\mathbf{i}))}\otimes_{k=1}^{d^n-1}|\theta(c_k(\mathbf{i}))^{m_k}\rangle\right)\\
&= d^{\frac{-d^n+2}{2}}\left(\sum_{\substack{\mathbf{m}\in\mathbb{Z}_d^{d^n-1},\\ \oplus_{k=1}^{d^n-1}m_k=o(\mathbf{i})}}\otimes_{k=1}^{d^n-1}|\theta(c_k(\mathbf{i}))^{m_k}\rangle\right)\ .
\end{aligned}$$

Hence, we need to show that we can satisfy the phase relations in Eq. (3.43) for $o(\mathbf{i}) = \delta(\mathbf{i})$ by choosing suitable linear functions $c_k(\mathbf{i}) = l_k(\mathbf{i}) \in L_n^{\mathbb{Z}_d}$.

Note that in contrast to the qubit case, the space of (local) functions with signature $f : \mathbb{Z}_d \to \mathbb{Z}_d$ contains d linearly independent elements. Out of those we only consider non-zero, linear and homogeneous functions, i.e., $l_{\mathbf{a}} = \oplus_{j=1}^n a_j i_j$ for $\mathbf{0} \neq \mathbf{a} \in \mathbb{Z}_d^n$, for which we again count appearances of entries in our choice of functions.

First, consider a single non-zero entry, e.g. $\mathbf{i} = (i_1, 0, \cdots, 0)^\intercal \in \mathbb{Z}_d^n$. For every appearance of the entry i_1, there are functions with coefficients a_1 ranging over all of $\mathbb{Z}_d$, which results in the local phase factor,

$$\phi(q) := \prod_{a_1=0}^{d-1} \theta(a_1 i_1)^q \chi^{a_1 i_1 (q-1) q^{d-1}} = \theta^q \chi^{\sum_{c=0}^{d-1} c (q-1) q^{d-1}} = \theta^q \chi^{\frac{d(d-1)}{2}(q-1)q^{d-1}} \,,^{12} \tag{3.44}$$

where $\theta := \prod_{c=0}^{d-1} \theta(c)$. Furthermore, the number of functions with $a_1 \neq 0$ counts $\sum_{l=0}^{n-1} \binom{n-1}{l} (d-1)^l = d^{n-1}$, hence, the overall phase factor in Eq. (3.43) reads $\phi(q)^{d^{n-1}}$. Next, we consider an input with two non-zero entries, e.g., $\mathbf{i} = (i_1, i_2, 0, \cdots, 0)^\intercal \in \mathbb{Z}_d^n$. We need to be more careful about the counting in this case as in contrast to the qubit case, where $i_1 + i_2 = 0$ for $i_1, i_2 \neq 0$, this does not hold for qudits. For functions with non-zero coefficients $a_1, a_2 \neq 0$ we obtain the contribution $\theta(a_1 i_1 + a_2 i_2)$, in particular, we need to count how many of these linear combinations equate to 0 and thus do not add a phase. It is not hard to see that there are $(d-1)^2 - (d-1) = (d-1)(d-2)$ non-trivial combinations, hence, we end up with the following overall phase factor in Eq. (3.43),

$$(\phi(q))^{d^{n-2}} \cdot (\phi(q))^{d^{n-2}} \cdot (\phi(q)^{d-2})^{d^{n-2}} = (\phi(q)^d)^{d^{n-2}} = \phi(q)^{d^{n-1}} \,.$$

The first two contributions are due to functions, where either $a_1 = 0$ or $a_2 = 0$, the third arises from the functions with both $a_1, a_2 \neq 0$, out of which there are $((d-1)(d-2))^{d^{n-2}}$ (and where, by symmetry, we can always group $(d-1)$ together to obtain the phase $\phi(q)$ in Eq. (3.44)). This argument now generalises to input strings $\mathbf{i} \in \mathbb{Z}_d^n$ with m non-zero entries as follows.

[12] Note that we are abusing notation slightly by using modulo-d arithmetic over phases with different periods. However, as the functions are computed classically the input is always an element in $\mathbb{Z}_d$.

Denote the number of non-zero linear combinations of the form $\oplus_{j=I} a_j i_j$ with $I \subseteq \{1, \cdots, n\}$, $|I| = m$ by $g(m)$. Clearly, $g(1) = (d-1)$, more generally

$$g(m) = (d-1)^m - g(m-1) = (d-1)^m - (d-1)^{m-1} + g(m-2) \ \cdots = \sum_{l=0}^{m-1} (-1)^l (d-1)^{m-l} \, .$$

Now, there are d^{n-m} linear functions for every linear function with m non-zero entries i_k, and for each of those we have the following contribution:

$$\begin{aligned}
\sum_{k=1}^{m} \binom{m}{k} g(k) &= \sum_{k=1}^{m} \binom{m}{k} \left(\sum_{l=0}^{k-1} (-1)^l (d-1)^{k-l} \right) \\
&= \binom{m}{m} \left((d-1)^m - (d-1)^{m-1} + (d-1)^{m-2} - (d-1)^{m-3} + \cdots \right) \\
&\quad + \binom{m}{m-1} \left((d-1)^{m-1} - (d-1)^{m-2} + (d-1)^{m-3} - \cdots \right) \\
&\quad\quad + \binom{m}{m-2} \left((d-1)^{m-2} - (d-1)^{m-3} + \cdots \right) \\
&\quad\quad\quad \vdots \\
&= \sum_{k=0}^{m-1} (-1)^k (d-1)^{m-k} \left(\sum_{l=0}^{k} (-1)^l \binom{m}{m-l} \right) \\
&= \sum_{k=0}^{m-1} (-1)^k (d-1)^{m-k} \left((-1)^k \binom{m-1}{k} \right) \\
&= (d-1) \left(\sum_{k=0}^{m-1} (d-1)^k \binom{m-1}{k} \right) \\
&= (d-1) d^{m-1}
\end{aligned}$$

In here, the first factor $(d-1)$ again results in the phase $\phi(q)$ from Eq. (3.44), and we thus obtain the overall phase factor

$$\left(\phi(q)^{d^{m-1}}\right)^{d^{n-m}} = \phi(q)^{d^{n-1}} \, . \tag{3.45}$$

Finally, we relate this phase factor to the local phases $\theta(c)$ and χ. Since,

$$\theta^d = \left(\prod_{c=0}^{d-1} \theta(c) \right)^d = \prod_{c=0}^{d-1} \chi^{-(d-1)c} = \chi^{-(d-1)\sum_{c=0}^{d-1} c} = \chi^{-\frac{d(d-1)^2}{2}} \, , \tag{3.46}$$

we need to choose $\theta(c)$ for $1 \leq c \leq d-1$ such that $\theta = \chi^{-\frac{(d-1)^2}{2}}$, e.g. $\theta(c) := \chi^{-\frac{c(d-1)}{d}}$. Next, we insert Eq. (3.46) into Eq. (3.44) and compute the overall phase factors:

$$\phi(q)^{d^{n-1}} = (\chi^{-\frac{(d-1)^2}{2}q} \cdot \chi^{(q-1)\frac{d(d-1)}{2}q^{d-1}})^{d^{n-1}} = \begin{cases} 1 & \text{if } q = 0 \\ \chi^{-\frac{d^{n-1}d(d-1)}{2}} (\chi^{-d^{n-1}\frac{(d-1)}{2}})^q & \text{if } 1 \leq q \leq d-1 \end{cases}$$

We may thus set $\chi^{-\frac{d^{n-1}(d-1)}{2}} = \omega$, from which it follows that $(\chi^{-\frac{d^{n-1}(d-1)}{2}})^d = 1$, hence, $\phi(q)^{d^{n-1}} = \omega^q$ as required. In summary, we obtain the output function

$$o(\mathbf{i}) = \begin{cases} 0 & \text{if } \mathbf{i} = 0 \\ 1 & \text{if } \mathbf{i} \neq 0 \end{cases}, \tag{3.47}$$

from which we compute $\delta(\mathbf{i}) = (d-1)o(\mathbf{i}) + 1$ by linear post-processing.

3.B Proof of Proposition 9

We choose measurement operators for prime dimension d as in Eq. (3.22), i.e.,

$$X(f)|q\rangle = f(q)|q+1\rangle \quad \text{with} \quad f : \mathbb{Z}_d \to U(1) \text{ s.t. } \prod_{q=0}^{d-1} f(q) = 1 \,,$$

where $|q\rangle$ denotes the computational basis. Expressed in terms of this basis, the eigenstates read:

$$\begin{aligned} |m\rangle &= \frac{1}{\sqrt{d}} \left(|0\rangle + \omega^{-m} f(0)|1\rangle + \omega^{-2m} f(0)f(1)|2\rangle + \cdots + \omega^{-(d-1)m} \left(\prod_{q'=0}^{d-2} f(q') \right) |d-1\rangle \right) \\ &= \frac{1}{\sqrt{d}} \left(\sum_{q=0}^{d-1} \omega^{-qm} \prod_{q'=0}^{q-1} f(q')|q\rangle \right) \end{aligned}$$

Conversely, the computational basis expressed in terms of eigenstates of $X(f)$ reads

$$|q\rangle = \frac{1}{\sqrt{d}} \left(\frac{1}{\prod_{q'=0}^{q-1} f(q')} \sum_{m=0}^{d-1} \omega^{qm} |m\rangle \right) \,. \tag{3.48}$$

We also use the following N-qudit resource state,

$$|\psi\rangle = \frac{1}{\sqrt{d}} \sum_{q=0}^{d-1} |q\rangle^{\otimes N} = \frac{1}{\sqrt{d}} \sum_{q=0}^{d-1} \otimes_{k=1}^{N} |q_k = q\rangle \ .$$

We would like to obtain the same parity state $\oplus_{k=1}^{N} m_k = 1 \pmod{d}$ for all but the $\mathbf{i} = 0$ input. Rewriting $|\psi\rangle$ in terms of the local measurement bases via Eq. (3.48) yields

$$\begin{aligned}
|\psi\rangle &= \frac{1}{\sqrt{d}} \sum_{q=0}^{d-1} \otimes_{k=1}^{N} \left(\frac{1}{\sqrt{d}} \frac{1}{\prod_{q'=0}^{q-1} f_k(c_k)(q')} \sum_{m_k=0}^{d-1} \omega^{q m_k} |m_k\rangle \right) \\
&= d^{\frac{-d^n}{2}} \sum_{q=0}^{d-1} \omega^{-q o(\mathbf{i})} \left(\sum_{\mathbf{m} \in \mathbb{Z}_d^N} \otimes_{k=1}^{N} \omega^{q m_k} |m_k\rangle \right) \\
&= d^{\frac{-d^n}{2}} \sum_{q=0}^{d-1} \left(\sum_{\mathbf{m} \in \mathbb{Z}_d^N} \omega^{q(\sum_{k=1}^{N} m_k - o(\mathbf{i}))} \otimes_{k=1}^{N} |m_k\rangle \right) \\
&= d^{\frac{-d^n+2}{2}} \left(\sum_{\substack{\mathbf{m} \in \mathbb{Z}_d^N, \\ \oplus_{k=1}^{N} m_k = o(\mathbf{i})}} \otimes_{k=1}^{N} |m_k\rangle \right) ,
\end{aligned}$$

if the output function $o : \mathbb{Z}_d^n \to \mathbb{Z}_d$ is encoded in the phase relations as follows

$$\prod_{1 \le k \le N} \left(\prod_{q'=0}^{q-1} f_k(c_k)(q') \right) = \omega^{q o(\mathbf{i})} \ .$$

Chapter 4

Conclusion and Outlook

Quantum theory is maybe our most successful theory of nature. Even though it has been invented with rather specific problems in mind, its applicability now reaches far beyond its initially intended purpose; it underlies not just atoms but all of quantum chemistry and it survived the marriage with field theory and relativity in the formulation of the standard model. The repeated accuracy of its predictions is staggering, to date we still only have few clues to where possible amendments might lurk. One way of testing the scope of quantum theory further might be by means of future quantum computers. The exponential scaling in simulations of quantum systems is the main roadblock to extensive pharmaceutical progress, groundbreaking innovation in material sciences, and possibly probing new physics. Together with a handful of exciting quantum algorithms, such as Shor's factorisation of large numbers, they have since inspired a generation of researchers.

Shor's algorithm suggests that certain computational tasks run more efficiently on a quantum computer than on any classical computer. Assuming this is in fact the case, it is natural to ask what powers a quantum computer. Motivated by the clear indications that (physical) contextuality is a fundamental ingredient to quantum theory, in the latter part of this thesis we studied contextuality in measurement-based quantum computation. We strengthened a previous connection between contextuality and computation from qubit to qudit systems, and constructed explicit contextual examples complementing well-known qubit ones. We gave a universal implementation method within this setup and found a natural resource measure in the number of qubit systems required in the non-adaptive, deterministic case. As a common thread, we emphasised that contextuality in quantum theory is closely intertwined with phase relations between eigenstates of measurement operators.

A natural avenue for future research in this direction is to perform a thorough investigation of the phase relations discussed here, compared to those arising in the cohomological classification of contextuality such as in [127]. That contextuality might be given a geometrical explanation—as an obstruction to global sections of corresponding presheaves—was first conjectured by Isham within the topos formalism [1]. It has received increased attention in recent years, which led to a number of different formalisms [5, 119, 127, 131]. Yet, none has so far been able to derive complete invariants. We expect that a suitable abelian presheaf, whose group structure arises naturally from the quantum formalism—in contrast to the very general sheaf-theoretic study of cohomology in [5]—will allow to combine and refine existing results.

In Ch. 2 we demonstrated that quantum theory builds on contextuality in a fundamental way. In fact, the mere order structure of contexts suffices to derive the quantum symmetries described by Wigner's theorem, the spectral presheaf encodes the Kochen-Specker theorem in the form of a quantum state space without global sections, while global sections of the probabilistic presheaf bijectively correspond with quantum states by Gleason's theorem. To this list of illustrious results in foundations we added Stinespring's and Bell's theorem: with respect to the former, we showed that complete positivity, a key property of quantum channels, originates from dilations in and time orientations on contexts, and with respect to the latter, that nonlocality in Bell's theorem emerges naturally from a notion of composition of contexts. As a corollary to this result, we solved an outstanding problem in quantum information theory by providing a classification of quantum correlations among all non-signalling correlations under the additional notion of time orientations in subsystems. We remark that the close resemblance with entanglement criteria in the reformulation of Bell's theorem (in contextual form) should be read as a hint towards treating this time-directional symmetry as a local, intrinsically relational symmetry principle within the active research on emerging space-time structures from entanglement.

In summary, (up to a choice of time direction) the full richness of quantum theory resides in the mere order relations between observables imposed by the equivalence relation defined by simultaneous measurability. This is a glaring confirmation of just how significant the idea behind physical contextuality is, which was maybe first acknowledged by Niels Bohr, mathematically captured by Ernst Specker and Simon Kochen, and raised to the status of a physical principle in the topos approach to quantum theory by Chris Isham and collaborators. Probably the most powerful workhorse in this programme is Gleason's theorem, which proves that the linear structure in quantum theory emerges from its context order in a very non-trivial way. It is remarkable that physical contextuality is sufficient to recover this central aspect of quantum theory, yet it begs the question why the context structure is of the particular form inherent to (certain) Jordan algebras, thus excluding many more general orthomodular lattices and examples such as Specker's parable.

Possible attempts at resolving this problem might be the following. In the geometric formalism [11, 99], the linear structure arises by complementing the (classical) symplectic form with a (quantum) Riemannian metric. Their resemblance with Jordan and Lie algebra aspects in the topos formalism is likely not a coincidence and deserves a more detailed study in the future. Another result that recovers the quantum algebra from few algebraic principles including a notion of composition is [69]. We expect that our notion of composition over contexts will allow to strengthen this result and yield further insight into the linear structure underlying the algebra of quantum mechanical observables.

Finally, the geometric nature of the topos formalism, the appearance of classical physics within contexts, and the generality of von Neumann algebras provide a solid basis for incorporating relativistic quantum mechanics, quantum field theories, and ultimately theories of gravity into this programme.

This last endeavour is maybe the single most ambitious and most rewarding in all of physics. It thus reminds of the introductory remarks and Specker's parable [133]. It is certainly not clear whether nature has set us an impossible task. But even if not, her contextual character seems to urge us to acknowledge our own limitations in order to see beyond our singular perspectives on reality. If we accept this, maybe the parable brings out a whole new interpretation.

The Unknown,

The Grand Show,

The Choir Of The Stars

Interstellar

Theatre Play,

The Nebulae Curtain Falls

Imagination,

Evolution,

A Species From The Vale

Walks In Wonder

In Search Of

The Source Of The Tale

[118]

Bibliography

1. A. Döring, and C. J. Isham, *J. Math. Phys.* **49**, 053515 (2008).

2. A. Döring, and C. J. Isham, *J. Math. Phys.* **49**, 053516 (2008).

3. A. Döring, and C. J. Isham, *J. Math. Phys.* **49**, 053517 (2008).

4. A. Döring, and C. J. Isham, *J. Math. Phys.* **49**, 053518 (2008).

5. S. Abramsky, R. S. Barbosa, S. Mansfield, *Phys. Rev. Lett.* **119**, 050504 (2017).

6. S. Abramsky, A. Brandenburger, *New J. Phys.* **13**, 113036 (2011).

7. E. M. Alfsen, F. W. Shultz, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6596–6601 (1998).

8. E. M. Alfsen, F. W. Shultz, *State spaces of operator algebras: Basic theory, orientations, and C^*-products* (Birkhäuser, Boston, 2001).

9. M. Amy, M. Mosca, *IEEE Transactions on Information Theory* **65**, 4771–4784 (2019).

10. J. Anders, D. E. Browne, *Phys. Rev. Lett.* **102**, 050502 (2009).

11. A. Ashtekar, T. A. Schilling, in *On Einstein's Path: Essays in Honor of Engelbert Schucking*, ed. by A. Harvey (Springer New York, New York, NY, 1999), pp. 23–65, ISBN: 978-1-4612-1422-9.

12. A. Aspect, P. Grangier, G. Roger, *Phys. Rev. Lett.* **47**, 460–463 (1981).

13. J. Baez, *Lett. Math. Phys.* **13** (1987).

14. J. C. Baez, arXiv: `2006.14741 [math-ph]` (2020).

15. H. Barnum, S. Beigi, S. Boixo, M. B. Elliott, S. Wehner, *Phys. Rev. Lett.* **104**, 140401 (2010).

16. V. P. Belavkin, P. Staszewski, *Rep. Math. Phys.* **24**, 49–55 (1986).

17. J. S. Bell, *Physics* **1**, 195 (1964).

18. J. Bell, *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy* (Cambridge University Press, 2004), ISBN: 9780521523387.

19. C. H. Bennett *et al.*, *Phys. Rev. A* **59**, 1070–1091 (1999).

20. J. Bermejo-Vega, N. Delfosse, D. E. Browne, C. Okay, R. Raussendorf, *Phys. Rev. Lett.* **119**, 120505 (2017).

21. H. J. Briegel, R. Raussendorf, *Phys. Rev. Lett.* **86**, 910–913 (2001).

22. P. Busch, *Phys. Rev. Lett.* **91**, 120403 (2003).

23. J. Butterfield, C. J. Isham, *Int. J. Theor. Phys.* **38**, 827–859 (1999).

24. J. Butterfield, C. J. Isham, *Int. J. Theor. Phys.* **41**, 613–639 (2002).

25. J. Butterfield, *Br. J. Philos. Sci.* **43**, 41–83 (1992).

26. A. Cabello, S. Severini, A. Winter, *Phys. Rev. Lett.* **112**, 040401 (2014).

27. E. T. Campbell, M. Howard, *Phys. Rev. Lett.* **118**, 060501 (2017).

28. S. Cannon, A. Döring, presented at the Reality and Measurement in Algebraic Quantum Theory, ed. by M. Ozawa *et al.*, pp. 3–65.

29. G. Cassinelli, E. de Vito, P. Lahti, A. Levrero, *The Theory of Symmetry Actions in Quantum Mechanics* (Springer, 2004).

30. C. M. Caves, C. A. Fuchs, K. K. Manne, J. M. Renes, *Found. Phys.* **34**, 193–209 (2004).

31. D. J. Chalmers, *J. Conscious. Stud.* **2**, 200–219 (1995).

32. M.-D. Choi, *Linear Algebra Its Appl.* **10**, 285–290 (1975).

33. N. P. L. Chris Heunen, B. Spitters, *Deep Beauty*, ed. by H. Halvorson (2009).

34. E. Christensen, *Commun. Math. Phys.* **86**, 529–538 (1982).

35. J. F. Clauser, M. A. Horne, *Phys. Rev. Lett.* **10**, 526–535 (1974).

36. J. F. Clauser, M. A. Horne, A. Shimony, R. A. Holt, *Phys. Rev. Lett.* **23**, 880–884 (1969).

37. R. Colbeck, R. Renner, *Phys. Rev. Lett.* **108**, 150402 (2012).

38. A. Connes, *Ann. Math.* **101**, 536–554 (1975).

39. A. Connes, *Noncommutative Geometry* (Elsevier Science, 1995), ISBN: 9780080571751.

40. A. Connes, fr, *Ann. Inst. Fourier (Grenoble)* **24**, 121–155 (1974).

41. N. De Beaudrap, *Quantum Info. Comput.* **13**, 73–115 (2013).

42. N. de Silva, *arXiv e-prints*, arXiv: `1709.00013 [quant-ph]` (2017).

43. H. F. de Groote, *Observables IV: The Presheaf Perspective*, 2007, arXiv: `0708.0677 [math-ph]`.

44. N. Delfosse, P. Allard Guerin, J. Bian, R. Raussendorf, *Phys. Rev. X* **5**, 021003 (2015).

45. A. Döring, *Int. J. Theor. Phys.* **44**, 139–160 (2005).

46. A. Döring, *ArXiv e-prints*, arXiv: `1411.5558 [math.OA]` (2014).

47. A. Döring, M. Frembs, arXiv: `1910.09591 [math-ph]` (2019).

48. A. Döring, *Flows on generalised Gelfand spectra of nonabelian unital C^*-algebras and time evolution of quantum systems*, 2012, arXiv: `1212.4882 [math.OA]`.

49. A. Döring, *Adv. Sci. Lett.* **2** (Oct. 2008).

50. A. Döring, J. Harding, *Houston J. Math., Electronic Edition* **42**, 559–568 (2016).

51. A. C. Doyle, *The Memoirs of Sherlock Holmes* (George Newnes, 1894).

52. H. A. Dye, *Ann. Math.* **61**, 73–89 (1955).

53. D. A. Edwards, *Synthese* **42**, 1–70 (1979).

54. A. Einstein, B. Podolsky, N. Rosen, *Phys. Rev.* **47**, 777–780 (1935).

55. G. Emch, *Algebraic Methods in Statistical Mechanics and Quantum Field Theory* (Dover Publications, 2009), ISBN: 9780486472096.

56. B. Fauser, G. Raynaud, S. Vickers, *Electronic Proceedings in Theoretical Computer Science* **95**, 81–90 (2012).

57. A. Fine, *Phys. Rev. Lett.* **48**, 291–295 (1982).

58. A. Fine, *J. Math. Phys.* **23**, 1306–1310 (1982).

59. A. Fine, P. Teller, *Found. Phys.* **8**, 629–636 (1978).

60. C. Flori, *J. Math. Phys.* **51**, 053527 (2010).

61. C. Flori, *J. Math. Phys.* **54**, 032106 (2013).

62. M. Frembs, A. Döring, arXiv: `1910.09596 [quant-ph]` (2019).

63. M. Frembs, *No-signalling and contextuality*, Master Thesis, 2016.

64. M. Frembs, S. Roberts, S. D. Bartlett, *New J. Phys.* **20**, 103011 (2018).

65. M. Giustina *et al.*, *Phys. Rev. Lett.* **115**, 250401 (2015).

66. A. M. Gleason, in *The Logico-Algebraic Approach to Quantum Mechanics: Volume I: Historical Evolution*, ed. by C. A. Hooker (Springer Netherlands, Dordrecht, 1975), pp. 123–133.

67. R. Greechie, *J. Comb. Theory, Series A* **10**, 119–132 (1971).

68. R. Greechie, *J. Comb. Theory* **4**, 210–218 (1968).

69. E. Grgin, A. Petersen, *Commun. Math. Phys.* **50**, 177–188 (1976).

70. D. Gross, *J. Math. Phys.* **47**, 122107–122107 (2006).

71. A. Grothendieck, *Produits tensoriels topologiques et espaces nucléaires.* French (Providence, RI: American Mathematical Society (AMS), Reprint, 1955), vol. 16, p. 191, ISBN: 978-0-8218-1216-7; 978-0-8218-9957-1/ebook.

72. S. Hameroff, R. Penrose, *Phys. Life Rev.* **11**, 39–78 (2014).

73. J. Hamhalter, *J. Math. Anal. Appl.* **422**, 1103–1115 (2015).

74. J. Hamhalter, *J. Math. Anal. Appl.* **383**, 391–399 (2011).

75. J. Hamilton, C. J. Isham, J. Butterfield, *Int. J. Theor. Phys.* **39**, 1413–1436 (2000).

76. J. Harding, M. Navara, *Order* **28**, 549–563 (2011).

77. L. Hardy, W. K. Wootters, *Found. Phys.* **42**, 454–473 (2011).

78. C. Heunen, N. P. Landsman, B. Spitters, *Commun. Math. Phys.* **291**, 63–110 (2009).

79. C. Heunen, N. P. Landsman, B. Spitters, *Synthese* **186**, 719–752 (2012).

80. L. E. Heyfron, E. T. Campbell, *Quantum Science and Technology* **4**, 015004 (2018).

81. P. Heywood, M. L. G. Redhead, *Found. Phys.* **13**, 481–499 (1983).

82. M. J. Hoban, E. T. Campbell, K. Loukopoulos, D. E. Browne, *New J. Phys.* **13**, 023014 (2011).

83. M. J. Hoban, J. J. Wallman, D. E. Browne, *Phys. Rev. A* **84**, 062107 (6 2011).

84. M. Horodecki, P. Horodecki, R. Horodecki, *Phys. Lett. A* **223**, 1–8 (1996).

85. M. Howard, J. Wallman, V. Veitch, J. Emerson, *Nature* **510**, 351–355 (2014).

86. M. N. I. Gelfand, *Rec. Math. [Mat. Sbornik] N.S.* **12**, 197–217, ISSN: 2 (1943).

87. C. J. Isham, J. Butterfield, *Int. J. Theor. Phys.* **37**, 2669–2733 (1998).

88. L. J. Bunce, J. D. M. Wright, *Exposition. Math.* **11**, 91–95 (1993).

89. L. J. Bunce, J. D. M. Wright, *Bull. Amer. Math. Soc.* **26** (1992).

90. A. Jamiołkowski, *Rep. Math. Phys.* **3**, 275–278 (1972).

91. P. Janotta, H. Hinrichsen, *J. Phys. A* **47**, 323001 (2014).

92. J. P. Jarrett, *Nous* **18**, 569–589 (1984).

93. P. Jordan, *Nachr. Ges. Wiss. Göttingen* **41**, 209–214 (1933).

94. P. Jordan, J. v. Neumann, E. Wigner, *Ann. Math.* **35**, 29–64 (1934).

95. R. V. Kadison, *Ann. Math.* **54**, 325–338 (1951).

96. R. V. Kadison, J. R. Ringrose, *Fundamentals of the theory of operator algebras: Volume I, Elementary theory* (Academic Press, 1983).

97. R. V. Kadison, J. R. Ringrose, *Fundamentals of the theory of operator algebras: Volume II, Advanced theory* (Academic Press, 1983).

98. A. Karanjai, J. J. Wallman, S. D. Bartlett, *arXiv e-prints*, eprint: `1802.07744` (quant-ph) (2018).

99. T. W. B. Kibble, *Commun. Math. Phys.* **65**, 189–201 (1979).

100. M. Kläy, C. Randall, D. Foulis, *Int. J. Theor. Phys.* **26**, 199–219 (1987).

101. S. Kochen, E. P. Specker, *Journal of Mathematics and Mechanics* **17**, 59–87 (1967).

102. J. Lagarias, P. W. Shor, *Bull. Amer. Math. Soc.* **27** (1996).

103. J. P. Lambare, arXiv: `1911.00343 [quant-ph]` (2019).

104. S. Lane, *Categories for the Working Mathematician* (Springer New York, 1998), ISBN: 9780387984032.

105. J. Lawrence, *Phys. Rev. A* **89**, 012105 (2014).

106. F. W. Lawvere, S. H. Schanuel, *Conceptual Mathematics: A First Introduction to Categories* (Cambridge University Press, ed. 2, 2009).

107. Y.-C. Liang, R. W. Spekkens, H. M. Wiseman, *Phys. Rep.* **506**, 1–39 (2011).

108. R. Lidl, H. Niederreiter, *Finite Fields* (Cambridge University Press, ed. 2, 1996).

109. S. Maeda, *Rev. Math. Phys.* **1**, 235–290 (1989).

110. S. Mansfield, E. Kashefi, *Phys. Rev. Lett.* **121**, 230401 (2018).

111. K. Mccrimmon, *A Taste of Jordan Algebras* (Springer, 2004).

112. N. D. Mermin, *Rev. Mod. Phys.* **65**, 803–815 (1993).

113. N. D. Mermin, *Phys. Rev. Lett.* **65**, 3373–3376 (1990).

114. M. A. Naimark, *C. R. (Dokl.) Acad. Sci. URSS, n. Ser.* **41**, 359–361 (1943).

115. K. Nakayama, *J. Math. Phys.* **54**, 072102 (2013).

116. K. Nakayama, *J. Math. Phys.* **55**, 102103 (2014).

117. M. Navara, V. Rogalewicz, *Mathematische Nachrichten* **154**, 157–168 (1991).

118. Nightwish, *Shudder Before The Beautiful*, Nuclear Blast Records, 2015.

119. C. Okay, S. Roberts, S. Bartlett, R. Raussendorf, *ArXiv e-prints*, arXiv: `1701.01888` (2016).

120. D. Pavlov, arXiv: `2005.05284 [math.OA]` (2020).

121. M. Pawlowski *et al.*, *Nature* **461**, Publisher: Nature Publishing Group, 1101–1104 (2009).

122. *Penrose triangle*, https://en.wikipedia.org/wiki/Penrose_triangle, Accessed: 2020-01-31.

123. A. Peres, *J. Phys. A* **24**, L175–L178 (1991).

124. B. Pierce, B. C, B. Pierce, M. Garey, A. Meyer, *Basic Category Theory for Computer Scientists* (MIT Press, 1991), ISBN: 9780262660716.

125. S. Popescu, D. Rohrlich, *Found. Phys.* **24**, 379–385 (1994).

126. R. Raussendorf, *Phys. Rev. A* **88**, 022322 (2013).

127. R. Raussendorf, arXiv: 1602.04155 [quant-ph] (2016).

128. R. Raussendorf, H. J. Briegel, *Phys. Rev. Lett.* **86**, 5188–5191 (2001).

129. R. Raussendorf, D. E. Browne, N. Delfosse, C. Okay, J. Bermejo-Vega, *Phys. Rev. A* **95**, 052334 (2017).

130. R. Raussendorf, D. E. Browne, H. J. Briegel, *Phys. Rev. A* **68**, 022312 (2003).

131. F. Roumen, *Electronic Proceedings in Theoretical Computer Science* **236**, 174–201 (2017).

132. L. K. Shalm *et al.*, *Phys. Rev. Lett.* **115**, 250402 (2015).

133. E. Specker, *Dialectica* **14**, 239–246 (1960).

134. R. W. Spekkens, *Phys. Rev. A* **71**, 052108 (2005).

135. W. F. Stinespring, *Proc. Am. Math. Soc.* **6**, 211–216 (1955).

136. C. Stoica, *Global and local aspects of causality*, The Time Machine Factory [speakable and unspeakable] on Time Travel - TM 2019, (https://www.slideshare.net/CristiStoica2/cristi-stoica-global-and-local-aspects-of-causality-the-time-machine-factory-speakable-unspeakable-on-time-travel-in-turin-october-2012).

137. E. Størmer, *Proc. Am. Math. Soc.* **86**, 402–402 (1982).

138. M. Takesaki, *Theory of Operator Algebras I* (Springer Berlin Heidelberg, 2001), ISBN: 9783540422488.

139. M. Takesaki, *Theory of Operator Algebras II* (Springer Berlin Heidelberg, 2002), ISBN: 9783540429142.

140. W. Tang, S. Yu, C. H. Oh, *Phys. Rev. A* **95**, 012131 (1 2017).

141. B. S. Tsirelson, *Lett. Math. Phys.* **4**, 93–100 (1980).

142. V. Veitch, C. Ferrie, D. Gross, J. Emerson, *New J. Phys.* **14**, 113011 (2012).

143. V. Veitch, S. A. H. Mousavian, D. Gottesman, J. Emerson, *New J. Phys.* **16**, 013009 (2014).

144. J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Springer, 1996), ISBN: 9783540592075.

145. N. Wallach, *Contemp. Math.* **305** (2000).

146. R. F. Werner, M. M. Wolf, *Phys. Rev. A* **64**, 032112 (2001).

147. E. P. Wigner, *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren* (Friedrich Vieweg und Sohn, Braunschweig, 1931), pp. 251–254.

148. H. M. Wiseman, E. G. Cavalcanti, in *Quantum [Un]Speakables II: Half a Century of Bell's Theorem*, ed. by R. Bertlmann, A. Zeilinger (Springer International Publishing, Cham, 2017), pp. 119–142.

149. W. Wootters, in *Complexity, Entropy, and the Physics of Information*, ed. by W. Zurek, pp. 39–46.

150. V. Wright, S. Weigert, *J. Phys. A* **52** (2018).

151. F. J. Yeadon, *Bull. London Math. Soc.* **16**, 145–150 (1984).

152. D. L. Zhou, B. Zeng, Z. Xu, C. P. Sun, *Phys. Rev. A* **68**, 062303 (2003).