Almost Equivalent Paradigms of Contextuality

Linde Wester

June 14, 2017

Abstract

Various frameworks that generalise the notion of contextuality in theories of physics have been proposed recently; one is a sheaf-theoretic approach by Abramsky and Brandenburger; a second is an equivalence-based approach by Spekkens. We combine the two approaches to derive a canonical method for detecting contextuality in models with preparations, transformations and sharp measurements, specifically in noise-free quantum circuits. In addition, we show that there is an isomorphism between respective categories of the two formalisms, which restricts to an isomorphism between the class of non-contextual theories in the sheaf sense and the class of factorizable non-contextual theories in the equivalence-based sense.

1 Introduction

Two Formalisms for Contextuality

Contextuality of quantum mechanics, which entails the impossibility of assigning predetermined outcomes to observables in a way that is independent of the method of observation, was first described by Kochen and Specker in [12]. Recently, two formalisms of different scope and nature have been proposed, which generalise the currently known examples of contextuality. The two approaches share the goal of seeking to express the notion of non-contextuality in a manner that is independent of the quantum formalism, and hence applicable to any operational theory. In this paper, we unify the sheaf-theoretic formalism by Abramsky and Brandenburger [3] with the equivalence-based notion of contextuality developed by Spekkens in [16]. We combine the advantages of each formalism to give a method for detecting contextuality in noise-free quantum circuits. This is helpful for an exploration of the connection between contextuality and quantum computing.

The Sheaf Approach

In the sheaf approach to contextuality, one defines contextuality as the non-existence of a joint probability distribution over the outcomes of joint measurements. It is formulated within the mathematical framework of sheaf theory, as the non-existence of a global section for a presheaf of measurement outcomes. The mathematical framework provides algorithmic methods to detect contextuality based on sheaf cohomology [6], as well as means of quantifying contextuality [1]. The approach can be applied to data generated by experiments without further knowledge or assumptions about the underlying physical description.

The Equivalence-Based approach

Contextuality in the equivalence-based approach is defined as the non-existence of certain ontological models for operational theories. Such an ontological model must be determined by the statistical data of the experiment only, and cannot depend on any additional data, regarded as the 'context'. This gives a natural explanation for operational equivalence of measurements and preparations: we cannot distinguish them because they correspond to the same ontological values. The formalism distinguishes three different types of contextuality: contextuality of preparations, transformations and of (unsharp) measurements. A theory as a whole is non-contextual if it is non-contextual for preparations, transformations, and for measurements simultaneously. As a result, contextuality can be experimentally tested, in a way that is robust to noise [13].

Overview

In this paper, we draw a formal connection between the sheaf-theoretic approach to contextuality on the one hand and general contextuality in the equivalence-based approach on the other hand. General contextuality is defined as the existence of at least one of the different types of contextuality.

We show that scenarios that can be described in the sheaf formalism correspond to operational theories and vice versa. A priori, experimental scenarios that are non-contextual in the sheaf-theoretic sense, may not be non-contextual in the equivalence-based sense, as they may depend on data other that the outcome statistics of the experiment. We show that whenever an experimental setting is non-contextual in the sheaf sense, we can eliminate any of these redundant data, making it equivalent to an experimental setting with no statistical redundancies. We use this minimal model to construct a non-contextual ontological representation for each scenario that is non-contextual in the sheaf sense. We call this the 'canonical' ontological model. In addition, we generalise a result of Abramsky and Brandenburger, to derive that an operational theory can be realised by a factorizable non-contextual ontological model in the equivalence-based sense, if and only if it can be realised by a 'canonical' non-contextual ontological model. By a result derived in [16], this implies that an operational theory with sharp measurements is non-contextual if and only if its 'canonical' representation is non-contextual. Noise-free quantum circuits can be represented as operational theories with sharp measurements. Consequently, we have a method for detecting contextuality in the circuit model. Our results depend on the way we choose to represent a physical setting in the abstract formalisms. The level of abstraction allows a certain degree of freedom in this respect. While the two formalisms are seemingly very different, they can be used to represent reality in equivalent ways. The choices we make in this paper to establish this equivalence offer an insight into the meaning of the formalisms and exemplify their nuances.

Outline

In Sections 2 we recall the equivalence-based approach to contextuality. We demonstrate that non-locality is a special case of measurement contextuality in the equivalence-based approach, and that measurement non-contextuality implies parameter independence of ontological models. We discuss Mermin's All Versus Nothing argument and Bell's scenario in the equivalence-based model. In Section 3, we recall the sheaf-theoretic approach. We give an alternative contextuality proof for the scenario for preparations and unsharp measurements given by Spekkens [16]. In Section 4, we introduce a method to construct a non-contextual ontological representation for each non-contextual empirical theory. We demonstrate how these ontological representations provide a simple method to detect contextuality in noise-free quantum circuits in Section 5. In Section 6.1 we introduce the categories $\mathcal{E}mp$ of empirical theories, $\mathcal{O}T$ of operational theories, and $\mathcal{O}R$ of ontological representations. In Section 6.2, we show how the correspondence between empirical and operational theories gives rise to a categorical isomorphism, which maps the subcategory of non-contextual empirical theories to the subcategory of operational theories that can be realised by a factorizable non-contextual ontological representation. In Section 7, we discuss contextuality for unsharp measurements, and give an example for which the two notions of contexuality differ.

2 The Equivalence-Based Approach

In this section we recall the equivalence-based approach to contextuality introduced by Spekkens in [16] and [14]. We discuss how this approach relates to joint measurements, factorizability, and parameter independence. The Kocher-Specker contextuality argument can be derived from the equivalence-based approach, as shown in [13]. We will demonstrate how it gives rise to Bell's non-locality scenario and Mermin's all versus nothing argument.

Consider two sets, P and M, of preparation procedures and measurement procedures, respectively. For each measurement $m \in M$ there is a set O^m of possible measurement outcomes. For each pair $(p, m) \in P \times M$, there exists a probability distribution $d_{p,m} : O^m \to [0, 1]$ over the set of possible outcomes O^m. The value $d_{p,m}(k)$ should be understood as the probability of obtaining the outcome k when a preparation p is performed, followed by a measurement m. We write D for the set of probability distributions $\{d_{p,m}\}_{p\in P, m\in M}$. An **operational theory** is defined by a tuple (P, M, D, O), where $O = \cup_m O^m$. In [16], operational theories also contain a set of transformation procedures, but we will not consider these here. This is a minor restriction since we can account for any transformation followed by a measurement by considering it as a new measurement. However, we lose the significance of compositionality of transformations.

Preparations and measurements are **statistically equivalent** when they are not distinguishable based on the measurement statistics in the operational theory. Let $p, p' \in P$ be preparations and let $m, m' \in M$ be measurements, this is expressed below.

$$\begin{array}{cccc} p \sim p' & \Leftrightarrow & d_{p,m} = d_{p',m} & \forall m \in M \\ m \sim m' & \Leftrightarrow & d_{p,m} = d_{p,m'} & \forall p \in P \\ (m,k) \sim (m',k') & \Leftrightarrow & d_{p,m}(k) = d_{p,m'}(k') & \forall p \in P \end{array}$$

An example of an operational theory is quantum theory. Equivalence classes of preparation procedures correspond to density matrices. Equivalence classes of measurement procedures correspond to POVM's. An **ontological representation** of the operational theory $A = (P, M, D, O)$ consists of a measurable topological space of ontological values Ω, together with sets of distribution functions $\mu = \{\mu_p : \Omega \to [0,1]\}_{p \in P}$ and $\xi = \{\xi_m(\lambda) : O^m \to [0,1]\}_{\lambda \in \Omega, m \in M}$. The distribution functions are such that they realise the measurement statistics of A, which is expressed by the formula below.

$$\int_{\Omega} \xi_m(\lambda)(k)\mu_p(\lambda)d\lambda = d_{p,m}(k) \qquad \forall p \in P, m \in M \tag{1}$$

An ontological model should be thought of as representing a physical system as it really is, while the operational theory merely describes our knowledge of the system, which may not be accurate or complete.

Remark 1. *If we assume that preparations, measurements and outcomes form separable measure spaces. By a result of Brandenburger and Keisler [5], we can take the space of ontological values to be the unit interval with the Borel measure without loss of generality.*

Definition 1. *An* **ontological representation** *is called* **preparation non-contextual** *if $\mu_p = \mu_{p'}$ whenever $p \sim p'$; it is called* **measurement non-contextual** *if $\xi_{k,m} = \xi_{k',m'}$ whenever $(m,k) \sim (m',k')$; it is called* **non-contextual** *if it is preparation non-contextual as well as measurement non-contextual. An* **operational theory** *is called* **non-contextual** *whenever there exists a non-contextual ontological model that realises the theory.*

One can verify that it is possible to find a non-contextual ontological representation for any operational theory. To see this, consider the ontological representation where the ontological states are given by equivalence classes of preparations $\Omega = \{[p]\}_{p \in P}$; the distribution function $\mu_q([p])$ is 1 iff $q \in [p]$ and 0 otherwise; the distribution $\xi_m([p])(k) = d_{p,m}(k)$. As a consequence, additional structure on an operational theory is required to characterise contextuality.

Postulate 1 (Convexity of Operational Theories)**.** *Operational theories are closed under convex combinations of preparations and measurements.*

In quantum theory, convex combinations can be seen as sampling over preparation or measurement procedures with probabilistic weights specified by the coefficients.

Postulate 2 (Preservation of Convexity)**.** *Let $p = c_1 p_1 + c_2 p_2$ be a convex combination of preparations, let $m = c_1 m_1 + c_2 m_2$ be a convex combination of measurements. We have the following equalities of distribution functions:*

$$\mu_{c_1 p_1 + c_2 p_2} = c_1 \mu_{p_1} + c_2 \mu_{p_2} \tag{2}$$

$$\xi_{c_1 m_1 + c_2 m_2} = c_1 \xi_{m_1} + c_2 \xi_{m_2} \tag{3}$$

$$\tag{4}$$

The two postulates are necessary conditions for the contextuality proof of preparations and unsharp measurements of 2-dimensional quantum systems given in [16]. It is also a necessary requirement for recovering any examples of contextuality that assume outcome determinism, including the Kochen Specker scenario and Mermin's 'All versus nothing' argument. We will demonstrate this for the latter in Section 2.2, using Lemma 1 below. A distribution function for a measurement m in an ontological model is **outcome-deterministic** if for any ontological value λ, one outcome of m occurs with certainty. This is the case when $\xi_m(\lambda)(k) \in \{0, 1\}$ for all $k \in O$ and $\lambda \in \Omega$. A distribution function is **outcome-deterministic almost everywhere** if it is outcome-deterministic up to a subset of Ω of measure 0. For a certain class of operational theories, any non-contextual ontological representation must be outcome-deterministic. We give two sufficient conditions below.

Definition 2. *A measurement m with outcome set O^m is* **perfectly predictable** *if for all $k \in O^m$ there exists a preparation p_k, such that $d_{p_k,m}(k') = \delta_{k,k'}$.*

Definition 3. *A preparation p^{mix} is* **maximally mixed** *if the following two conditions hold.*

1. *For every preparation p', p^{mix} is statistically equivalent to some convex combination of preparations containing p'.*
2. *For every perfectly predictable measurement m, p^{mix} is statistically equivalent to some convex combination of p_k for $k \in O^m$.*

Lemma 1. *Let m be a perfectly predictable measurement in an operational theory that satisfies postulates 1 and 2 and contains a maximally mixed preparation. The distribution function ξ_m is outcome-deterministic almost everywhere.*

Proof. Let $\Omega_p = \{\mu_p(\lambda) > 0\}$ be the topological support of μ_p. The topological support of a measure space is defined to be the largest closed subset such that every open neighbourhood of every point of the space has positive measure.

Perfect predictability implies that $\int \xi_m(\lambda)(k)\mu_{p_{k'}}(\lambda)d\lambda = \delta_{k,k'}$. If $k = k'$, this means that $\xi_m(\lambda)(k) = 1$ for all $\lambda \in \Omega_k$, since μ_{p_k} is a probability distribution. For $k \neq k'$, this implies that $\xi_m(\lambda)(k) = 0$ for $\lambda \in \Omega_{k'}$. This gives us the following description of ξ_m:

$$\xi_m(\lambda)(k) = \begin{cases} 1 & \text{if } \lambda \in \Omega_k \\ 0 & \text{if } \lambda \in \cup_{k' \neq k}\Omega_{k'} \end{cases}$$

It is left to prove that $\Omega \backslash \cup \Omega_k$ has measure zero. By preparation non-contextuality, there is one distribution $\mu_{p^{mix}}$ for all preparations that are statistically equivalent to the maximally mixed preparation p^{mix}. It follows that $\Omega \backslash \cup_p \Omega_p$ has measure zero. By the first condition of the maximally mixed preparation and postulate 2, it follows that $\Omega \backslash \Omega_{p^{mix}}$ must have measure zero. By the second condition, we know that $\Omega_{p^{mix}} \backslash \cup_{k \in O^m} \Omega_{p_k}$ has measure zero. It follows that $\Omega \backslash \cup_{k \in O^m} \Omega_{p_k}$ has measure zero. Hence ξ_m is outcome-deterministic almost everywhere. □

In quantum theory, PVM's are perfectly predictable. Any preparation of the maximally mixed state is a maximally mixed preparation. It follows that each preparation non-contextual ontological representation that satsfies postulate 2 of a PVM in quantum theory is outcome-deterministic. This result was given in [16].

Definition 4. *A set of N measurements $\{m_1, m_2, ..., m_N\}$ is* **jointly measurable** *if there exists a measurement m with the following features:*
(i) The outcome set of m is the Cartesian product of the outcome sets of $\{m_1, ..., m_N\}$
(ii) Let S be a subset of the index set $\{1, ..., n\}$. The outcome distributions for every joint measurement of any subset $\{m_s | s \in S\} \subset \{m_1, ..., m_N\}$ is recovered as the marginal of the outcome distribution of m for all preparations $p \in P$. Denoting a joint measurement of the subset S by m_S with a corresponding section $k_S \in O^{m_s}$, the condition can be expressed as

$$\forall S, \forall p : d_{p,m_S}(k) = \sum_{k \in O^m : \pi_S(k) = k_S} d_{p,m}(k). \tag{5}$$

Here, π_S is the projection function on the subset $\mathcal{E}(m_S) \subset \mathcal{E}(m)$.
(iii) The functions $\pi_S \circ m$ define measurements in the operational theory.

Conditions (i) and (ii) correspond to the definition of a joint measurement in [14].

An ontological representation is called **parameter independent** if for each measurement the effect on the ontological states is independent of any other measurement performed simultaneously. We can restrict the distribution function of a joint measurement m to a subset m' by the restriction function $\xi_m|_{m'}$, which is defined as $\xi_m|_{m'}(k')(\lambda) := \sum_{k:\pi_{m'}(k)=k'} \xi_m(k)(\lambda)$, where $\pi_{m'} : O^m \to O^{m'}$ projects the outcomes of m to the set of outcomes of m'. Parameter independence means that for two joint measurements m, n, the equality $\xi_m|_{m \cap n} = \xi_n|_{m \cap n}$ holds.

Lemma 2. *Any measurement non-contextual representation of an operational theory (P, M, D, O) is parameter independent. That is, for all joint measurements $m, n \in \mathcal{M}$ in a measurement cover $\mathcal{M}$, we have the following equalities:*

$$\xi^k_{m|m\cap n}(\lambda) = \xi^k_{m\cap n}(\lambda) = \xi^k_{n|m\cap n}(\lambda) \tag{6}$$

Proof. Let $m = \{m_1, ..., m_N\}$ be a jointly measurable set of measurement procedures of $\mathcal{M}$, let p be a preparation procedure, let K_s be the set $\{k \in O^m | \pi_s(k) = k_s\}$, for some $k_s \in O^{m_s}$.

By joint measurability and basic probability theory, we have the following sequence of equalities on the operational level:

$$\begin{aligned}\mathbb{P}(k_S|m_S;p) &= \sum_{k\in K_S} \mathbb{P}(k|m,p) \\ &= \mathbb{P}(K_S|m,p) \\ &= \mathbb{P}(k_S|\pi_S \circ m, p)\end{aligned}$$

By condition (iii) of joint measurements, $\pi_S \circ m$ is a well-defined measurement. It then follows from measurement non-contextuality that $\xi^{k_s}_{m_s} = \xi^{k_S}_{\pi_S\circ m}$, which implies

$$\xi^{k_S}_{m_S} = \sum_{k|\pi_S(k)=k_S} \xi_m(k) =: \xi_{m|m_S}(k_S) \tag{7}$$

□

We call an ontological model **factorizable** when for joint measurements $m = (m_1, ..., m_n)$, we can write $\xi_m(\lambda)(o) = \prod_{i=1,...,n} \xi_{m_i}(\lambda)(\pi_i(o))$. It was shown in Theorem 6 of [14] that any ontological model which is outcome-deterministic and measurement non-contextual, is factorizable.

2.1 Bell's Scenario

Consider an experiment where two parties can each choose from two different measurements with outcome set $\{0, 1\}$: a and a' for the first party; b and b' for the second party. The outcome statistics of each possible combination of measurements is organised in the table below. Each entry $a_{i,j}$ of the table represents the probability of obtaining outcome i for measurement j.

	(0,0)	(1,0)	(0,1)	(1,1)
(a, b)	1/2	0	0	1/2
(a', b)	3/8	1/8	1/8	3/8
(a, b')	3/8	1/8	1/8	3/8
(a', b')	1/8	3/8	3/8	1/8

(8)

This gives rise to an operational theory with one preparation, $P = \{p\}$. The set of measurements is defined as $M = \{a, a', b, b', (a, b), (a', b), (a, b'), (a', b')\}$. The outcome set for each elementary measurement is $O^m = \{0, 1\}$; for each tuple, the outcome set is $\{0, 1\}^2$. The set D is given by the distributions defined in the table, together with their marginals.

We assume perfect predictability and the existence of a maximally mixed state. Since any non-contextual ontological representation is in particular preparation non-contextual, it follows from perfect predictability and the existence of the maximally mixed state that any such representation must be outcome-deterministic, hence factorizable. As a result, every ontological state can be associated with a function $\{a, a', b, b'\} \to \{0, 1\}$ from the set of elementary measurements to the outcome set. This morphism maps each measurement to the outcome that occurs with probability 1 when the system is in ontological state λ. In other words, we can identify each ontological value λ with the outcome $(\alpha, \beta, \alpha', \beta') \in \{0, 1\}^4$. We obtain the probability that $a = \alpha, b = \beta, a' = \alpha', b' = \beta'$ by taking the weithed sum over all ontological values that correspond to this outcome according to equation 1. The weights in this sum are given by the distribution function of the preparation procedure. We denote this probability by $p_{\alpha\beta\alpha'\beta'}$. These probabilities should sum up to the values given in the table. The entries $a_{1,1}, a_{2,2}, a_{3,3}$, and $a_{1,4}$, give us the 4 equations below.

$$a_{1,1} : p_{0000} + p_{0010} + p_{0001} + p_{0011} = 1/2 \qquad a_{2,2} : p_{0010} + p_{1010} + p_{0011} + p_{1011} = 1/8$$
$$a_{3,3} : p_{0001} + p_{0101} + p_{0011} + p_{0111} = 1/8 \qquad a_{1,4} : p_{0000} + p_{0100} + p_{1000} + p_{1100} = 1/8$$

The left-hand-side of the sum of $a_{2,2}, a_{3,3}$ and $a_{1,4}$ should be greater than 1/2, since it contains all summands of $a_{1.1}$. However, the right-hand side of these equations sums to 3/8. As a result, the equations cannot be satisfied. As a consequence of this contradiction, a non-contextual ontological representation cannot exist.

2.2 Mermin's All Versus Nothing Argument

Suppose that we are given a preparation of the GHZ state $(|\uparrow\uparrow\uparrow\rangle + |\downarrow\downarrow\downarrow\rangle)/\sqrt{2}$ and a choice of Pauli X and Y measurements $\{X_i, Y_i\}_{i\in\{1,2,3\}}$, where the index i corresponds to the three different components of the GHZ state. The outcome set for each individual measurement is given by its eigenvalues, $\{-1, 1\}$.

This gives rise to an operational theory where $P = \{p^{GHZ}\}$, and $M = \{X_1, X_2, X_3, Y_1, Y_2, Y_3, X_1Y_2Y_3, Y_1Y_2X_3, Y_1X_2Y_3, X_1X_2X_3\}$. The triples represent joint measurements. The outcome sets O^m are $\{-1, 1\}$ for the elementary measurements and $\{-1, 1\}^3$ for the joint measurements. The elements of D are given by the Born rule. One can verify that for the triples of the measurements below, we get the following outcomes with certainty, where the right-hand-side is the product of the outcomes of the three individual measurements:

$$X_1Y_2Y_3 \quad = -1 \qquad Y_1Y_2X_3 \quad = -1 \qquad Y_1X_2Y_3 \quad = -1 \qquad X_1X_2X_3 \quad = 1$$

We will show that no ontic state λ in a non-contextual representation allows for probability distributions $\mu_m(\lambda)$ that are consistent with the support of this scenario. Suppose that there exists a non-contextual ontological representation for this operational theory, then in particular, this representation is preparation non-contextual. Assuming perfect predictability and the existence of a maximally mixed state, preparation non-contextuality of the ontological representation implies outcome determinism and factorizability. Given any ontological state λ of such representation, we can identify each of the measurements X_i, Y_i with the outcome that occurs with certainty. All four triples are joint measurements of their components, so by factorizability, their outcomes correspond to the product of the outcomes of the three components. In other words, μ assigns -1 or 1 to each X_i, Y_i, in a way that the equalities above are satisfied. It is easy to see that this is impossible: The product of the expressions on the left-hand-side must equal 1, since every measurement occurs twice, while the product of the right-hand-sides equals -1.

3 The Sheaf Approach

We recall the sheaf-theoretic approach to contextuality and non-locality, which was introduced by Abramsky and Brandenburger in [3] and [2]. Sheaves are a mathematical tool for describing how local data can be combined to obtain global information about a system. In this setting a **system type** consists of a discrete set X of measurement labels, together with a **measurement cover** $\mathcal{M} = \{C_i\}_{i \in I}$. This is an antichain of subsets of X, such that $\cup_{i \in I} C_i = X$. This means that for $C, C' \in \mathcal{M}$, we have the implication $C \subset C' \implies C = C'$. The measurement cover $\mathcal{M}$ represents the maximal sets of measurements that can be performed jointly. We write $\downarrow \mathcal{M}_A$ for the simplicial complex generated by $\mathcal{M}$.

We shall fix a set O of outcomes, which is the union of the sets of possible outcomes for each of the measurements in X. For each set of measurements $U \subset X$, a **section over** U is a function $U \to O$. We write O^U for the set of sections over U. The assignment $U \mapsto O^U$ defines a sheaf over the discrete topological space $\mathcal{E} : \mathcal{P}(X) \to Set$, which we call the **sheaf of events**. The restriction function, which is the remaining part of the data defining this sheaf, is given below.

$$\rho^U_{U'} := \mathcal{E}(U \subset U') : O^{U'} \mapsto O^U :: s \to s|_U$$

We call elements of $\mathcal{E}(X)$ **global sections** of measurement outcomes. Each global section consists of an assignment of an outcome to each of the measurements.

For any commutative semiring R and set X, an R-distribution d on X is a map $d : X \to R$ of finite support, such that

$$\sum_{x \in X} d(x) = 1$$

We write $\mathcal{D}_R(X)$ for the set of R-distributions on X. For a function of sets $f : X \to Y$, we define

$$\mathcal{D}_R(f) : \mathcal{D}_R(X) \to \mathcal{D}_R(Y) :: d \mapsto [y \mapsto \sum_{f(x)=y} d(x)]$$

It is easy to see that $\mathcal{D}_R$ is functorial. Hence, we can compose $\mathcal{E}$ with $\mathcal{D}_R$ to obtain a presheaf $\mathcal{D}_R\mathcal{E} : \mathcal{P}(X)^{op} \to Set$, which maps each set of measurements to the set of R-distributions over their sections. The ring of non-negative reals $\mathbb{R}_+$ corresponds to probability distributions over the outcomes; the ring of booleans $\mathbb{B}$ represents the possibility of outcomes.

The approach can be generalised to a presheaf over a small, thin category $D_R\mathcal{E} : \mathbf{C} \to Set$, as in [8]. A category is called thin when for each two objects A, B and each two morphisms $f, g : A \to B$, we have the equality $f = g$. This is another way of characterising a preorder. As in the equivalence-based approach, the order relation is given by a notion of joint measurement. Depending on the interpretation of an empirical theory, one could adopt different notions of joint measurement. In this paper we will use the notion of joint measurability given in Section 1. We recover the set of measurement labels X as the set given by all objects A, such that there is an arrow $A \to B$ to each object B, for which there exists an arrow $B \to A$. The measurement cover $\mathcal{M}$ corresponds to the set of those objects A, such that there is an arrow $B \to A$ from every object B, for which there exists an arrow $A \to B$. Note that when the thin category is a poset, we obtain the usual notions of measurement labels and a measurement cover.

A **state** for a system type $\mathbf{C}$ corresponds to a family σ for the cover $\mathcal{M}$, with respect to the presheaf $\mathcal{D}_R\mathcal{E}$. This is given by a distribution $\sigma_C \in \mathcal{D}_R\mathcal{E}(C)$ for each measurement context $C \in \mathcal{M}$. A state is called **no-signalling** when for all $C, C' \in \mathcal{M}$

$$\sigma_C|_{C \cap C'} = \sigma_{C'}|_{C \cap C'}$$

When a state σ is no-signalling, the restriction $\sigma|_m$ for $m \in Ob(\mathbf{C})$ corresponds to the probability distribution over the outcomes of m. We will denote this distribution by σ_m. There may be several states corresponding to the same distribution. These are **statistically equivalent states**. We call the tuple $(\mathbf{C}, S, O)$, where S is a collection of states for a system type $\mathbf{C}$, an **empirical theory**. We define **contextuality of a state** $\sigma \in S$ as the non-existence of a global section for the presheaf $\mathcal{D}_R\mathcal{E}$ given σ. An **empirical theory is contextual** if at least one of its states is contextual.

Quantum theory gives rise to an empirical model. The measurement labels correspond to POVM's, elements of a measurement cover are POVM's that form a joint measurement as in Definition 4, and states are given by density matrices.

3.1 Convexity in Empirical Theories

Many of the classic contextuality results about quantum mechanics, including Kochen-Specker scenarios, Hardy's paradox, Bell's scenario and Mermin's all versus nothing argument can be derived from the sheaf approach, as shown in [3] and [7]. In this section we will do the same for the contextuality argument of preparations and unsharp measurements described in [16]. We will demonstrate that the arguments are a direct consequence of Postulates 1 and 2, independent of the chosen notion of contextuality.

As Postulates 1 and 2 are formulated in terms of operational theories and ontological representations, we introduce their analogues for empirical models. In this setting, global sections can be seen as the counterpart of the non-contextual ontological representations. We explain this in Section 4.1. Note that states are by definition closed under convex combinations, which means that Postulate 1 is equivalent to Postulate 3 below.

Postulate 3 (Convexity of Empirical Theories)**.** *The set of measurement labels of an empirical theory is closed under convex combinations.*

Postulate 4 (Preservation of convexity)**.** *Let d be a global section for an empirical theory. Let p_1, p_2 be preparations that give rise to the states σ_{p_1} and σ_{p_2}, such that the convex combination $c_1 \cdot p_1 + c_2 \cdot p_2$ gives rise to the state $\sigma_{c_1 \cdot p_1 + c_2 \cdot p_2}$. Let m_1, m_2 be measurement labels. We have the following equalities*

$$d^{\sigma_{c_1 \cdot p_1 + c_2 \cdot p_2}} = c_1 \cdot d^{\sigma_{p_1}} + c_2 \cdot d^{\sigma_{p_2}} \tag{9}$$

$$d_{c_1 \cdot m_1 + c_2 \cdot m_2} = c_1 \cdot d_{m_1} + c_2 \cdot d_{m_2} \tag{10}$$

Lemma 3. *Let $(\mathbf{C}, S)$ be an empirical model with a global section d. Postulate 4 implies that convexity is preserved by the empirical model*

$$\sigma^{c_1 \cdot p_1 + c_2 \cdot p_2} = c_1 \cdot \sigma^{p_1} + c_2 \cdot \sigma^{p_2}$$

$$\sigma_{c_1 \cdot m_1 + c_2 \cdot m_2} = c_1 \cdot \sigma_{m_1} + c_2 \cdot \sigma_{m_2}$$

Proof. This follows immediately from the definition of a global section. □

We now give a proof for the contextuality scenarios of preparations and unsharp measurements in the sheaf formalism.The proofs depend on the fact that the empirical model of quantum theory does not preserve convexity.

3.1.1 Contextuality for Preparations

Consider the following set of states in quantum theory:

$$\begin{aligned} \psi_a &= (1,0) & \psi_b &= (1/2, \sqrt{3}/2) & \psi_c &= (1/2, -\sqrt{3}/2) \\ \psi_A &= (0,1) & \psi_B &= (\sqrt{3}/2, -1/2) & \psi_C &= (\sqrt{3}/2, 1/2) \end{aligned}$$

We define the empirical theory below, where each P_x is the measurement label that corresponds to the projection onto the quantum state ϕ_x. Furthermore, σ^x is the state in the empirical model that corresponds to the quantum state ϕ_x. The state σ^{xyz} is defined as $\frac{1}{3}\sigma^x + \frac{1}{3}\sigma^y + \frac{1}{3}\sigma^z$.

$X = \{P_a, P_A, P_b, P_B, P_c, P_C\}$ $\qquad$ $\mathcal{M} = \{\{P_a, P_A\}, \{P_b, P_B\}, \{P_c, P_C\}\}$

$S = \{\sigma^a, \sigma^A, \sigma^b, \sigma^B, \sigma^c, \sigma^C, \sigma^{abc}, \sigma^{ABC}\}$ $\qquad$ $O = \{0, 1\}$

The outcome set O indicates if the outcome corresponding to the projector of the POVM element occurs (1), or if it does not (0). For instance, for the section $s : P_a \mapsto 0$, $\sigma^a_{P_a}(s) = 0$, $\sigma^b_{P_a}(s) = \frac{1}{4}$, and $\sigma^c_{P_a}(s) = \frac{1}{4}$.

The convex combination $\frac{1}{3}\sigma^a + \frac{1}{3}\sigma^b + \frac{1}{3}\sigma^c$ gives rise to the maximally mixed state with constant distribution $\sigma^{mix}_{p_x} = \frac{1}{2}$ for any observable p_x. Suppose that there exists a global section d, by Postulate 4 and Lemma 3, this gives us the following equality.

$$\frac{1}{2} = \frac{1}{3}\sigma^a_{P_a} + \frac{1}{3}\sigma^b_{P_a} + \frac{1}{3}\sigma^c_{P_a}$$

It is easy to see that this cannot hold for any section. Working out the outcome probabilities for $P_a \mapsto 0$ gives us the contradiction below.

$$\frac{1}{2} = \frac{1}{3} \cdot 0 + \frac{1}{3} \cdot \frac{1}{4} + \frac{1}{3} \cdot \frac{1}{4}$$

3.1.2 Contextuality for Unsharp Measurements

Consider the following empirical theory

$$\begin{aligned}
&X = \{P_a, P_A, P_b, P_B, P_c, P_C, P_{abc}, P_{ABC}\} \\
&\mathcal{M} = \{\{P_a, P_A\}, \{P_b, P_B\}, \{P_c, P_C\}, \{P_{abc}, P_{ABC}\}\} \\
&S = \{\sigma^a, \sigma^A, \sigma^b, \sigma^B, \sigma^c, \sigma^C\} \\
&O = \{0, 1\}
\end{aligned}$$

The measurement label P_{abc} is the convex combination $\frac{1}{3}P_a + \frac{1}{3}P_b + \frac{1}{3}P_c$, P_{ABC} is defined similarly, and the other elements are as defined in section 3.1.1. In quantum theory, this gives us the measurement context $\{P_{abc}, P_{ABC}\} = \{\frac{1}{2}, \frac{1}{2}\}$. Suppose that this scenario has a global section. By Postulate 4 and Lemma 3, we have the equalities below for any state σ.

$$\frac{1}{2} = \sigma_{P_{abc}} = 1/3\sigma_{P_a} + 1/3\sigma_{P_b} + 1/3\sigma_{P_c} \tag{11}$$

$$\frac{1}{2} = \sigma_{P_{ABC}} = 1/3\sigma_{P_A} + 1/3\sigma P_B + 1/3\sigma P_C \tag{12}$$

It is easy to see that this does not hold for the given states. For example, if we take the state σ^a, the convexity condition together with the Born rule give us $\sigma^a_d(1) = \sigma^a_D(0) = \frac{3+\sqrt{3}}{6}$ and $\sigma_D(1) = \sigma_d(0) = \frac{1+\sqrt{3}}{6}$. This contradicts the outcome statistics of $\{p_{abc}, p_{ABC}\}$, which assign equal probability to each outcome for any measurement.

4 Unifying Approaches

We will explore the relation between empirical theories, operational theories, and ontological representations. Any no-signalling empirical theory $A = (\mathbf{C}_A, S_A, O_A)$ corresponds to an operational theory $Op(A) = (P_{Op(A)}, M_{Op(A)}, D_{Op(A)}, O_A)$ in the sense that the two theories describe the same experimental setting. The elements of the operational theory are defined below.

$$
\begin{aligned}
P_{Op(A)} &:= S_A \\
M_{Op(A)} &:= Ob(\mathbf{C}_A) \\
d_{m,\sigma}(k) &:= \sigma_m(s) \qquad\qquad \text{for } d_{m,\sigma} \in D_{Op(A)} \text{ and } s(m) = k
\end{aligned}
$$

Conversely, the preorder $\mathbf{C}_A$ corresponds to the set of measurements M of the operational theory, with the order relation given by the notion of joint measurement.

Remark 2. *Signalling empirical theories cannot be described as an operational theory. The reason is that while the 'same' measurement can have different outcome statistics in the empirical theory, depending on the context, this is not possible in an operational theory. A way to get around this is by treating restrictions of a context to a measurement as elementary measurements.*

We will show that every non-contextual empirical theory A gives rise to an non-contextual ontological representation for $Op(A)$. We will call this a canonical ontological model for the empirical theory. Finally, we prove the following theorem, which generalises the result in [3].

Theorem 4. *The following statements are equivalent for any no-signalling empirical theory A and its corresponding operational theory $Op(A)$*

1. *The empirical theory A admits a global section*
2. *The operational theory $Op(A)$ admits a canonical non-contextual ontological representation*
3. *The operational theory $Op(A)$ admits a factorizable non-contextual ontological representation*

4.1 A Canonical Ontological Representation

As a warm-up, we recall the canonical ontological representation for empirical theories with a global section d for each state σ, which was introduced in [3].

The ontological states are given by the global sections of outcomes, the distributions μ correspond to the global section of distribution functions and $\xi_m(s)(k)$ indicates whether s assigns the outcome k to the measurement m.

$$\Omega = \mathcal{E}(X) \qquad \mu_\sigma(s) = d(s) \qquad \xi_m(s)(k) = \delta_{s|_m(m),k}$$

It is easy to see that this ontological representation is generally not non-contextual, since sections may assign different outcomes to statistically equivalent measurements. Suppose that s is a section of measurement outcomes such that $s|_m \neq s|_n$ for $m \sim n$, then $\xi_m(s)(k) \neq \xi_n(s)(k)$. To get around this, we will prove that whenever a global section exists, we can find another global section that depends on equivalence classes of measurements only. It is not hard to see that the same holds for states.

4.2 Statistical Equivalence in Empirical Theories

We call two states $\sigma, \sigma' \in S$ and two measurement labels $m, m' \in Ob(\mathbf{C})$ **statistically equivalent** when $\sigma_m = \sigma'_m$ for all $m \in Ob(\mathbf{C})$ and $\sigma_m = \sigma_{m'}$ for all $\sigma \in S_A$, respectively. In that case we write $\sigma \sim \sigma'$ and $m \sim m'$.

Let $A = (\mathbf{C}_A, S_A)$ be an empirical theory. We construct a new empirical theory $\tilde{A} := (\mathbf{C}_A/{\sim}, \tilde{S}_A)$ by quotienting the objects of $\mathbf{C}$ by the equivalence relation. The new category $\mathbf{C}_A/\sim$ contains an arrow between two equivalence classes if there exists an arrow between two representatives of the classes. It is instructive to unfold the structure of this new empirical theory. For each object $[C]$ of $\mathbf{C}_A/\sim$ the new set of sections $\mathcal{E}([C])$ contains a (not necessarily unique) section $\tilde{s}$ for each $s \in \mathcal{E}(C)$. This section is defined as $\tilde{s}([C]) := s(C)$. The states in $\tilde{S}_A := \{\tilde{\sigma}\}_{\sigma \in S_A}$, are defined as $\tilde{\sigma}_{[C]}(\tilde{s}) := \sigma_C(s)$. The set $\tilde{S}_A$ is well-defined, because $[C] = [D]$ if and only if $\sigma_C = \sigma_D$ for each $\sigma \in S_A$.

Lemma 5. *Any empirical theory A admits a global section iff it admits a global section that only depends on equivalence classes of measurements of A.*

We will prove this Lemma formally in Section 6. Intuitively, it can be understood as follows: Any global section d of S_A can be restricted to a global section over a subcategory of $\mathbf{C}_A$ of representatives of $\mathbf{C}_A/\sim$. This restriction defines a global section for $\tilde{S}_A$. Conversely, any global section $\tilde{d}$ of $\tilde{A}$ defines a global section d for A, defined as $d(s) := \tilde{d}(\tilde{s})$ when s assigns the same value to all elements of an equivalence class, and $d(s) := 0$ otherwise.

We have shown how to deal with equivalence on the level of measurements. However, individual outcomes of measurements can be statistically equivalent, even when the measurements as a whole are not. This means that for some $s \in O^m$ and $s' \in O^{m'}$, $\sigma_m(s) = \sigma_{m'}(s')$ for all $\sigma \in S_A$. To eliminate this last form of statistical redundancy, we rewrite any such system type A as a system type A' with outcome set $\{0, 1\}$. The measurement labels of A' are given by the individual observables in each measurement. We denote each observable by a tuple (m, k) of a measurement and an outcome, so $X_{A'} := \{(m, k)\}_{m \in X_A, k \in O}$.

The measurement cover is given by the sets of observables that form a measurement in the original cover: $\mathcal{M}_{A'} := \{\{(m,k)|k \in O, m \in C\}_{C\in\mathcal{M}_A}\}$. The outcomes 0 and 1 indicate whether the outcome corresponding to the observable is observed, hence $S_{A'} := \{\sigma'|\sigma'_{(m,k)}(1) = \sigma_m(k)\}$. The support $\mathcal{E}(m)$ of each measurement m consists of those sections where exactly one observable in each measurement is assigned a 1, and all others are assigned a 0.

Note that the model A has a global section iff A' has a global section under the given restrictions. As a consequence of Lemma 5, A has a global section iff $\tilde{A}'$ has a global section. Hence, A contains a global section induced by a global section d for $\tilde{A}'$, which is only defined on equivalence classes.

4.3 Non-contextual Canonical Ontological Representations

We can now define a canonical ontological representation that preserves non-contextuality. Let A be an empirical theory with a global section d_σ for each state $\sigma \in S_A$, which only depends on the equivalence classes of the preparations. We make use of the minimal empirical theory $\tilde{A}'$ and its induced global sections $\tilde{d}_{\tilde{\sigma}}$ to define the canonical ontological representation $R(A) = (\Omega_A^{NC}, \{\mu_\sigma^{NC}\}_{\sigma\in S_A}, \{\xi_m^{NC}\}_{m\in\downarrow\mathcal{M}})$:

$$\Omega_{R(A)} := \mathcal{E}(X(\mathbf{C}/\sim)), \qquad \mu_\sigma^{NC}(s) := \tilde{d}_{\tilde{\sigma}}(\tilde{s}), \qquad \xi_m^{NC}(s)(k) := \delta_{\tilde{s}([m]),[k]}$$

Note that $\xi_m^{NC}(s)(k)$ is only defined when s is a section over n, so when this is not the case, we will take $\xi_m^{NC}(s)(k)$ to be 0. This representation generates the required outcome statistics, as shown below.

$$\begin{aligned}
\int_{\Omega_A^{NC}} \mu_\sigma^{NC}(s)\xi_m^{NC}(s)(k)ds &= \sum_{s\in\mathcal{E}(X(\mathbf{C}/\sim))} \tilde{d}_{\tilde{\sigma}}(\tilde{s})\delta_{\tilde{s}|_{[m]}([m]),[k]} \\
&= \sum_{s\in\mathcal{E}(X(\mathbf{C}))} d_\sigma(s)\delta_{s|_m(m),k} \\
&= d^A_{\sigma,m}(k)
\end{aligned}$$

The first equality holds by unfolding definitions of the canonical representation. The second equality holds because $d_\sigma(s)$ is only nonzero on those sections s that assign the same outcome to all equivalent measurements; therefore, we can extend the sum over $\mathcal{E}(X/\sim)$ to the sum over $\mathcal{E}(X)$. The last equality holds as both expressions are equal to $\sigma_m(s)(k)$.

This canonical ontological representation is by definition preparation non-contextual. On measurements, it is defined such that $m \sim m'$ implies $\xi_m^{NC} = \xi_{m'}^{NC}$; hence it is measurement non-contextual.

It is left to determine under which conditions an operational theory can be realised by a non-contextual empirical theory. To this end, we generalise Theorem 8.1 of [3].

Lemma 6. *For every factorizable, non-contextual ontological representation B, there exists an empirical theory A with a global section, such that $R(A)$ and B realise the same operational theory.*

Proof. Let B be a factorizable, measurement non-contextual ontological representation. The operational theory realised by B induces an empirical theory A where X_A is given by the minimal elements of the preorder of joint measurements. By Lemma 2, B is parameter independent. Every preparation $p \in P_B$ realises a state σ_p with a global section d_p for the sheaf of distributions induced by A. These are defined below for $r \in \mathcal{E}(m)$ and $s \in \mathcal{E}(X)$:

$$\sigma_p(r) := \int_{\Omega_B} \xi_m(\lambda)(r(m))\mu_P(\lambda)d\lambda \quad d_p(s) := \int_{\Omega_B} \prod_{m \in X_A} \xi_m(\lambda)(s|_m(m))\mu_P(\lambda)d\lambda \tag{13}$$

We need to verify that $R(A)$ and B realise the same measurement statistics. This follows from the equalities below, where we denote the canonical ontological representation by Ω'_A, μ', and ξ'.

$$\begin{aligned}
\int_{\tilde{s} \in \Omega'_A} \xi'_m(\tilde{s})(k)\mu'_p d\tilde{s} &= \sum_{\tilde{s} \in \mathcal{E}(X_A/\sim)} \delta_{\tilde{s}|_{[m]}([m]),[k]} \left[\tilde{d}_{\tilde{\sigma}_P}(\tilde{s}) \right] \\
&= \sum_{s \in \mathcal{E}(X_A)} \delta_{s|_m(m),k} \left[\int_{\lambda \in \Omega_B} \prod_{n \in X_A} \mu_p(\lambda)\xi_n(\lambda)(s|_n(n))d\lambda \right] \\
&= \int_{\lambda \in \Omega_B} \xi_m(\lambda)(k) \left[\sum_{s \in \mathcal{E}(X_A \backslash m)} \prod_{n \in X_A \backslash m} \xi_n(\lambda)(s|_n(n)) \right] \mu_p(\lambda)d\lambda \\
&= \int_{\lambda \in \Omega_B} \xi_m(\lambda)(k)\mu_p(\lambda)d\lambda
\end{aligned}$$

The first two equalities result from expanding definitions. For the third, we apply Fubini's theorem, split the sum and product, and rewrite the expression. The last equality holds because probability distributions sum to one over all the inputs. □

Proof of Theorem 4. For any empirical theory A, the canonical non-contextual ontological representation for $Op(A)$ is given by $R(A)$, which means that 1) $\Rightarrow$ 2). The canonical ontological representation $R(A)$ is factorizable; therefore, 2) $\Rightarrow$ 3). Finally, 3) $\Rightarrow$ 1) holds by Lemma 6. □

Corollary 7. *For the class of perfectly predictable operational theories with a maximally mixed preparation, an operational theory is non-contextual iff its canonical ontological representation is non-contextual.*

Proof. By Lemma 1, all preparation non-contextual ontological representations of operational theories in this class are outcome-deterministic. It follows that all non-contextual ontological representations are factorizable. □

5 Contextuality in the Circuit Model

Theorem 4 gives rise to a method for detecting some forms of contextuality in the circuit model, under the assumption that the quantum circuits are noise-free. This method is a first step to a general formalism for contextuality in the circuit model. This approach does not take statistical equivalences into account that arise from composition of gates. To accomodate for this type of equivalence, contextuality of transformations needs to be considered. However, as quantum circuits can be simulated classically, this will most likely not affect the contextuality of the circuit. We leave a further analysis for future work.

Contextuality plays a crucial role in measurement-based quantum computing (MBQC). As shown in [15] and [11], a MBQC scenario evaluates a mod 2 nonlinear function iff this scenario is strongly contextual. Since MBQC is a universal model for quantum computing, one may ask whether this holds for quantum computation in general. The most likely answer is 'no' since we can compute mod-2 non-linear gates with classical circuits. Indeed, it is easy to find a counterexample that shows that a strongly contextual MBQC scenario translates to a non-contextual quantum circuit, which implements the same computation.

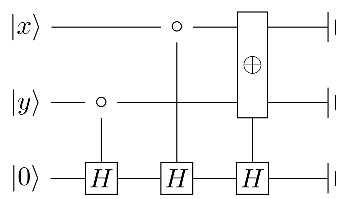
Any quantum circuit, or collection of quantum circuits, corresponds to an operational theory consisting of preparation procedures of qubits, transformation procedures (quantum gates), and measurement procedures (projective measurement in the computational basis). For simplicity, we consider operational theories with only preparations and measurements; therefore, we regard any combination of gates applied to a preparation as a separate preparation. We also allow preparations of mixed states and applications of mixed gates.

The measurements in a noise-free quantum circuit are perfectly predictable. Since the maximally mixed state is part of the operational theory, it follows from Corollary 7 that a quantum circuit is non-contextual whenever its canonical ontological representation is non-contextual. For a circuit with n qubits, the canonical set of ontological values is given by the sections over the measurements, so $\Omega = \{0,1\}^n$. The indicator function ξ_{m_i} for the measurement on the ith qubit is the deterministic function such that $\xi_{m_i}(s)$ for $s \in \Omega$ corresponds to the ith component of s. For each preparation p, the function μ_p is defined as $\mu_p(s) = P(s|p,m)$, where m is the measurement in the computational basis.

The standard example of contextuality in MBQC is given by the following scenario. We are given a GHZ state and perform the measurements M_a, M_b, M_c on the three components of the state, depending on an input of two bits x, y in the following way: Let P_0 be a Pauli X measurement and let P_1 be a Pauli Y measurement. Then $M_a := P_x$, $M_b := P_y$ and $M_c := P_{x\oplus y}$. This scenario implements the OR-gate on the two input bits x, y. It also corresponds to Mermin's all versus nothing scenario, as the inputs correspond to the following measurements:

$$00: \quad XXX \qquad 01: \quad XYY \qquad 10: \quad YXY \qquad 11: \quad YYX$$

The same OR-gate is implemented by the quantum circuit below, where we first have a controlled Hadamard gate on two qubits x and y and finally a controlled-H gate, controlled by the sum of the first two qubits. The output is given by the outcome of the measurement on the 3rd qubit.



We have the following deterministic distribution functions for the different inputs $\mu_{000}(000) = 1, \mu_{100}(101) = 1, \mu_{010}(011) = 1, \mu_{110}(111) = 1$, determined by the output probabilities of the circuit. Finally, the three measurements have the following indicator functions $\xi_1(xyz) = x, \xi_2(xyz) = y, \xi_3(xyz) = z$. None of the measurements or preparations is statistically equivalent to anothor; hence, this ontological representation is non-contextual. Note that under the preservation of convexity assumption, this circuit is non-contextual, as the circuit only contains pure states.

6 A Categorical Isomorphism

In this section, we show that the two formalisms can be used to represent physical reality in equivalent ways. To establish this, we use the mathematical framework of category theory. We show that the correspondence between empirical theories, operational theories and ontological representations discussed in the previous section gives rise to functors between suitable categories. In particular, we establish an isomorphism between the categories $\mathcal{E}mp$ of empirical theories and $\mathcal{O}T$ operational theories. This isomosrphism maps non-contextual empirical theories to operational theories that admit a factorizable non-contextual ontological representation.

6.1 A Categorical Setup

6.1.1 The Category of Empirical Theories

We will define **the category** $\mathcal{E}mp$ of **empirical theories** and **transformations that preserve contextuality and statistical equivalence**. The category $\mathcal{E}mp$ is an extension of the category of empirical models introduced in [9].

Definition 5. *A transformation between empirical theories is given by a triple $f = (f^S, f^{\mathcal{M}}, f^O)$ of maps between the set of states, the measurement cover and the set of outcomes, respectively. In addition, each assignment $C \mapsto f^{\mathcal{M}}(C)$ consists of a functor $f^C : C \to f^{\mathcal{M}}(C)$ of the subcategories of objects with an arrow to C and $f^{\mathcal{M}}(C)$, respectively.*

Note that if $\mathbf{C}_A$ and $\mathbf{C}_B$ are posets, f is a simplicial map $\downarrow \mathcal{M}_A \to \downarrow \mathcal{M}_B$. We write f for either component when it is clear from the context which one we mean. If a transformation satisfies the following equation, we can recover the statistical data of the domain from the statistical data of the image.

$$\sigma_C(s) = \sum_{s' \circ f^{\mathbf{C}} = f^O \circ s} f^S(\sigma)_{f^{\mathbf{C}}(C)}(s') \qquad\qquad \forall C \in \mathcal{M}_A \tag{14}$$

We will call such transformations **contextuality preserving** due to the following Lemma.

Lemma 8. *Let $f : A \to B$ be a transformation of empirical theories that satisfies equation (14) and let σ be a state of A. If σ does not admit a global section, then $f(\sigma)$ does not admit a global section.*

Proof. Suppose that $f(\sigma) \in B$ has a global section $\nu \in D_R\mathcal{E}(f(X_A))$. This means that $\nu|_{C'} = f(\sigma)_{C'}$ for all $C' \in X$. This induces a global section for σ, given by $\mu(s) = \sum_{s' \circ f^{\mathcal{M}} = f^O \circ s} \nu|_{f(X_A)}(s')$ in $D_R(\mathcal{E}(X_A))$.

□

When A and B are no-signalling models, $f^{\mathcal{M}}$ is simply a functor of categories $\mathbf{C}_A \to \mathbf{C}_B$. Equation 14 then simplifies to the equation below.

$$\sigma_m(s) = f^S(\sigma)_{f^C(m)}(s') \qquad\qquad \forall f^O \circ s(m) = s' \circ f^C(m)\ \forall m \in X_A \tag{15}$$

In addition to equation 14, we require morphisms to preserve statistical equivalence:

$$m \sim m' \quad \Rightarrow \quad f(m) \sim f(m') \qquad\quad \sigma \sim \sigma' \quad \Rightarrow \quad f(\sigma) \sim f(\sigma') \tag{16}$$

We can now prove the statement in Lemma 5, that any global section gives rise to a global sections defined on equivalence classes.

Proof of Lemma 5. Consider the quotient map $A \xrightarrow{q} \tilde{A}$ and any inclusion map $\tilde{A} \xrightarrow{i} A$, which is defined as follows: $[C]$ is mapped to some representative C such that $i^{\mathbf{C}}$ is a functor, and $\tilde{\sigma}$ is mapped to σ. It is easy to see that q and i are morphisms in $\mathcal{E}mp$. consequently, the proof follows from Lemma 8. □

Remark 3. *Another way to define transformations between empirical theories is given in [4]. Here, empirical models are defined in terms of Chu spaces and the function on states goes in the opposite direction. By that definition, contextuality of states would only be preserved by transformations that are surjective on states.*

6.1.2 The category of operational theories

Operational theories form a category $\mathcal{O}t$. Morphisms are tuples $f = (f^M, f^P, f^O) : A \to B$, such that $f^M : M_A \to M_B$, $f^P : P_A \to P_B$ and $f^O : O_A \to O_B$ preserve outcome statistics and statistical equivalence:

$$d_{f^P(p),f^M(m)}(f^O(k)) = d_{p,m}(k) \tag{17}$$

$$m \sim m' \Rightarrow f(m) \sim f(m') \qquad p \sim p' \Rightarrow f(p) \sim f(p') \tag{18}$$

This category is similar to the category of operational theories defined in [4].

6.1.3 The category of ontological representations

Objects in the category $\mathcal{O}R$ of ontological representations correspond to a pair of an ontological representation and its induced operational theory. Morphisms consist of triples of maps (f, f^μ, f^ξ), where $f : A \to B$ is a morphism of operational theories, and $f^\xi : \xi \mapsto \xi'$ and $f^\mu : \mu \mapsto \mu'$ are functions of sets. We require that the image of (f, f^μ, f^ξ) realises the operational theory in the image of f. This means that the images of the elements of μ and ξ coincide with the elements corresponding to the images of f^P and f^M. We express this as $f^\mu(\mu_p) = f^\mu(\mu)_{f^P(p)}$ and $f^\xi(\xi_m) = f^\xi(\xi)_{f^M(m)}$. In addition, one can deduce from equations 1 and 17 that the equality below holds.

$$\int_{\Omega_B} f^\xi(\xi_m)(\lambda)(f^O(k))f^\mu(\mu_p)(\lambda)d\lambda = \int_{\Omega_A} \xi_m(\lambda)(k)\mu_p(\lambda)d\lambda \tag{19}$$

Remark 4. *Note that the morphisms do not contain a component that maps between the sets of ontological values. This is because our goal is not to understand individual ontological representations, but to explore the existence of certain classes of ontological representations for operational theories.*

There is a forgetful functor $G : \mathcal{O}R \to \mathcal{O}T$ that maps each ontological representation to its corresponding operational theory. More precisely, it maps (Ω, ξ, μ) to $(\{\mu_p\}_{p\in P}, \{\xi_m\}_{m\in M}, D, O)$, where $d_{p,m} := \int_\Omega \mu_p(\lambda)\xi_m(\lambda))d\lambda$. The elements μ_P and ξ_M no longer represent distribution functions, but merely label the preparations and measurements.

Lemma 9. *Contextuality of operational theories is preserved by morphisms in* $\mathcal{O}T$

Proof. Let $f : A \to B$ be a morphism of operational theories. Let $(\Omega_B, \{\mu_p\}_{p\in P_B}, \{\xi_m\}_{m\in M_B})$ be a non-contextual ontological representation of B. This induces an ontological representation $(\Omega_B, \{\mu'_p\}_{p\in P_A}, \{\xi'_m\}_{m\in M_A})$, which is defined as $\mu'_p := \mu_{fp}$, $\xi'_m := \xi_{fm}$. Non-contextuality of this ontological representation is guaranteed by the equivalence preservation condition on f. It follows by contradiction that when A is contextual, B must be contextual. $\square$

6.2 Empirical and Operational theories

The assignment $A \mapsto Op(A)$ of an operational theory to each no-signalling empirical theory described in Section 4.1 gives rise to the functor below.

$$\begin{array}{ccc} \mathcal{E}mp & \xrightarrow{\qquad\qquad Op \qquad\qquad} & \mathcal{O}T \\ A & \longmapsto & Op(A) \\ (f^S, f^C, f^{O_A}) & \longmapsto & (f^M, f^P, f^{O_{Op(A)}}) \end{array}$$

Here, f^M is defined as the assignment on objects of $f^{\mathcal{M}}$, $f^P := f^S$ and $f^{O_A} = f^{O_{Op(A)}}$, since $O_A = O_{Op(A)}$. To verify that this is well-defined on morphisms, one needs to check that equations 17 and 18 hold. Since we only consider no-signalling empirical theories, this follows directly from equations 15 and 16. Functoriality is straightforward. We will show that it is in fact an isomorphism of categories. As discussed in section 4, the assignment is bijective on objects. To see that it is injective on morphisms, note that the functor $f^{\mathcal{M}}$ is completely determined by its assignments on objects, since $\mathbf{C}_A$ and $\mathbf{C}_B$ are thin categories. Surjectivity follows from the fact that equations 15 and 16 imply equations 17 and 18.

We will show that the isomorphism $\mathcal{E}mp \xrightarrow{Op} \mathcal{O}T$ maps non-contextual empirical theories to operational theories that admit a factorizable non-contextual ontological representation. In order to do so, we first examine how the canonical ontological representation described in Section 4.1 gives rise to a functor $\mathcal{E}mp \xrightarrow{R} \mathcal{O}R$. This functor maps each non-contextual empirical model to its canonical non-contextual ontological representation. It maps each morphism of empirical models to a morphism of ontological representations in an obvious way, such that the effect on the outcome statistics is the same in either model. Later we will show that the composition of this functor with the forgetful functor $\mathcal{O}R \to \mathcal{O}T$ equals $\mathcal{E}mp \xrightarrow{Op} \mathcal{O}T$ on the class of non-contextual empirical models.

Proposition 10. *For any choice of global sections, the assignment $A \mapsto R(A)$ defines an equivalence between the subcategory of non-contextual empirical theories and the subcategory of non-contextual, factorizable ontological representations. The image $Rf = (Rf, (Rf)^\mu, (Rf)^\xi)$ of each morphism $f = (f^S, f^{\mathcal{M}}, f^O)$ has the following components*

$$Rf := Op(f) \qquad (Rf)^\mu(\sigma_m) := f\sigma_{f(m)}, \qquad (Rf)^\xi(\xi_m)(s)(k) := \delta_{s(f(m)),k}$$

Proof. We need to verify that for each $f : A \to B$ in $\mathcal{E}mp$, $Rf : R(A) \to R(B)$ is a well-defined morphism in the category of ontological representations. It is easy to see that since f preserves statistical equivalence, Rf does too. By the following equations, Rf also satisfies equation 19.

$$\int_{\Omega_{R(A)}} \xi_m(\lambda)(k)\mu_\sigma(\lambda)d\lambda = \sum_{\lambda\in\mathcal{E}(m)} \delta_{\lambda(m),k}\sigma_m(\lambda) \tag{20}$$

$$= \sum_{\lambda'\in\mathcal{E}(f(m))} \delta_{\lambda'(f(m)),f^O(k)} f\sigma_{f(m)}(\lambda') \tag{21}$$

$$= \int_{\Omega_{R(B)}} Ff(\xi_m)(\lambda')(f^O(k))Rf(\mu_\sigma)(\lambda')d\lambda' \tag{22}$$

The equalities are obtained by unfolding definitions, application of equation 15 and rewriting the summation.

By Lemma 6, R is essentially surjective on the subcategory of factorizable non-contextual ontological representations. We will show that the functor R is injective on hom-sets. First of all, $\delta_{s(f(m)),k} = \delta_{s(g(m)),k)}$ for all $k \in O$ implies that $f^S = g^S$. Similarly, $R(f)^\mu = R(g)^\mu$ implies $f^C = g^C$. We will prove that F is surjective on hom-sets. Let $(g, g^\mu, g^\xi) : RA \to RB$ be a morphism in $\mathcal{O}R$. This corresponds to the morphism $g' : A \to B$ in $\mathcal{E}mp$ with components $g'^{\mathcal{M}}(m) = g^M(m)$ and $g'^S(\sigma)_{g'^{\mathcal{M}}(m(s))} = g^P(\mu_\sigma)(s)$ for $s \in \mathcal{E}(m)$. Since each ξ in the image of g is a delta function, it must be equal to $\delta_{\lambda(g(m)),k}$. Finally, to show that equation 18 holds, we take equation 19 and unfold the definitions of Ω_A, Ω_B, $g(\xi_M)$, and ξ_M. This gives us the equality below, which reduces to the second condition for transformations of empirical theories.

$$\sum_{s\in\mathcal{E}(m)} \delta_{s(m),k}\mu_\sigma(s) = \sum_{s\in\mathcal{E}(f(m))} \delta_{s(f(m)),k} g_\mu(\mu_\sigma)(s) \tag{23}$$

□

Theorem 11. *The isomorphism Op restricts to an isomorphism between the subcategory of non-contextual empirical theories and the subcategory of operational theories that do not admit a factorizable non-contextual ontological representation.*

Proof. Note that the following diagram commutes, where we write $\mathcal{E}mp^{NC}$ and $\mathcal{O}R^{FNC}$ for the subcategories of non-contextual empirical theories and factorizable non-contextual ontological representations, respectively.

$$\begin{tikzcd}
\mathcal{E}mp \arrow[rr, "Op"] & & \mathcal{O}T \\
\mathcal{E}mp^{NC} \arrow[u, hook] \arrow[dr, "R"] & & \mathcal{O}R \arrow[u, "G"'] \\
& \mathcal{O}R^{FNC} \arrow[ur, hook] &
\end{tikzcd}$$

□

Corollary 12. *For models with perfect predictability, R restricts to an isomorphism between non-contextual empirical theories and non-contextual operational theories.*

7 Non-factorizable representations and POVM's

In general, equivalence-based measurement contextuality implies sheaf-theoretic contextuality, but not necessarily the other way around. The two formalisms coincide for models with perfect predictability, as any non-contextual representation for sharp measurements is necessarily factorizable. Furthermore, any non-local scenario rules out non-factorizable ontological representations, as these would violate local causality.

One may wonder if there are interesting examples of scenarios where the two formalisms differ. Below we give a sufficient postulate under which the two formalisms are equal.

The following example shows that the two formalisms are not equal for all scenarios.

Example 1. *There are three parties, A, B, C, that each conduct a measurement with two outcomes, $\{0, 1\}$. It is possible for two parties to apply the measurement at the same time, but it is not possible to apply all three measurements simultaneously. The measurement statistics is such that for any joint measurement, the obtained outcome is $(0, 1)$ half of the time, and $(1, 0)$, half of the time.*

This scenario cannot be realised by sharp measurements in quantum mechanics. However, one can find a POVM for each joint measurement that margnalises to the required outcomes: $(0 \cdot P_{0,0}, \frac{1}{2} \cdot P_{0,1}, \frac{1}{2} \cdot P_{1,0}, 0 \cdot P_{1,1})$, where $P_{i,j}$ is the projector onto outcome (i, j).

Lemma 13. *The scenario in Example 1 is contextual in the sheaf sense, but non-contextual in the equivalence-based sense*

Proof. The marginal probabilities for each of the individual measurements are $\frac{1}{2}$ for either of the outcomes. It follows that all measurements are statistically equivalent, and hence, should not be distinguishable on the ontological level. This means that we can define the set $\Omega := \{*\}$ to be a singleton set. We set $\mu_p(*) = 1$ for any preparation of this scenario, $\xi_m(*)(0) = \xi_m(*)(1) = \frac{1}{2}$ for each of the elementary measurements, and $\xi_m(*)(0, 1) = \xi_m(*)(1, 0) = \frac{1}{2}$, for each of the joint measurements. On the other hand, it is not possible to define a factorizable non-contextual ontological representation. It is easy to see this, since without loss of generality, any global section of measurement outcomes to the presheaf describing this scenario must assign the same outcome to measurement A and B. But that means that it does not marginalise to an admissible outcome for the joint measurement of A and B. $\square$

For a complete comparison of the two notions, a better understanding of unsharp measurements is required. Another point of consideration is the extent to which the functors respect additional postulates. We have shown that for all known examples of contextuality conditional to postulates in the equivalence-based framework, the two notions coincide. However, this may not be the case in general. Ideally, one would like to have a specification of the class of scenarios and possibly postulates for which the formalisms are different. We leave this for future work.

References

[1] S. Abramsky, R. Soares Barbosa, and S. Mansfield. Quantifying contextuality via linear programming. 2016.

[2] S. Abramsky and A. Brandenburger. Axiomatizing physical theories in a sheaf-theoretic framework. *Unpublished note*, 2011.

[3] S. Abramsky and A. Brandenburger. The sheaf-theoretic structure of non-locality and contextuality. *New Journal of Physics 13*, 2011.

[4] S. Abramsky and C. Heunen. Operational theories and categorical quantum mechanics. 2013.

[5] S. Abramsky and J. Keisler. A canonical hidden-variable space. 2014.

[6] S. Abramsky, S. Mansfield, and R. Soares Barbosa. The cohomology of non-locality and contextuality. *QPL 2011*, 2011.

[7] S. Abramsky, S. Mansfield, R. Soares Barbosa, K. Kishida, and R. Lal. Contextuality, cohomology and paradox. 2015.

[8] S. Abramsky and M. Sadrzadeh. Semantic unification - a sheaf theoretic approach to natural language. *Lecture Notes in Computer Science*, 8222, 2014.

[9] Samson Abramsky and Adam Brandenburger. Axiomatizing physical theories in a sheaf-theoretic framework. 2011.

[10] J. Bell. On the einstein podolsky rosen paradox. *Physics 1*, 1964. 195.

[11] M. Hoban, E. T. Campbell, K. Loukopoulos, and D. E. Browne. Non-adaptive measurement-based quantum computation and multi-party bell inequalities. *New Journal of Physics*, 13, 2011.

[12] S. Kochen and E.P. Specker. The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics*, 17:59–87, 1967.

[13] R. Kunjwal and R. Spekkens. From the kochen specker theorem to non-contextuality inequalities without assuming determinism. 2015.

[14] Y. Liang, R. W. Spekkens, and H. M. Wiseman. Specker's parable of the over-protective seer: A road to contextuality, nonlocality and complementarity. *Physics Reports*, 2011. vol 506, issues 1-2, pp 1-39.

[15] Robert Raussendorf. Contextuality in measurement-based quantum computing. 2009.

[16] R. W. Spekkens. Contextuality for preparations, transformations, and unsharp measurements. *A Physical Review*, 71, 2005.

A
Title of Appendix A

Text of Appendix A is Here

B
Title of Appendix B