Contextuality under weak assumptions

Contextuality under weak assumptions

Andrew W. Simmons
Department of Physics, Imperial College London, London SW7 2AZ, United Kingdom

Joel J. Wallman
Institute for Quantum Computing and Department of Applied Mathematics, University of Waterloo, Waterloo, Canada

Hakop Pashayan
Centre for Engineered Quantum Systems, School of Physics, The University of Sydney, Sydney, NSW 2006, Australia

Stephen D. Bartlett
Centre for Engineered Quantum Systems, School of Physics, The University of Sydney, Sydney, NSW 2006, Australia

Terry Rudolph
Department of Physics, Imperial College London, London SW7 2AZ, United Kingdom

Abstract. The presence of contextuality in quantum theory was first highlighted by Bell, Kochen and Specker, who discovered that for quantum systems of three or more dimensions, measurements could not be viewed as deterministically revealing pre-existing properties of the system. More precisely, no model can assign deterministic outcomes to the projectors of a quantum measurement in a way that depends only on the projector and not the context (the full set of projectors) in which it appeared, despite the fact that the Born rule probabilities associated with projectors are independent of the context. A more general, operational definition of contextuality introduced by Spekkens, which we will term "probabilistic contextuality", drops the assumption of determinism and allows for operations other than measurements to be considered contextual. Even two-dimensional quantum mechanics can be shown to be contextual under this generalised notion. Probabilistic noncontextuality represents the postulate that elements of an operational theory that cannot be distinguished from each other based on the statistics of arbitrarily many repeated experiments (they give rise to the same operational probabilities) are ontologically identical. In this paper, we introduce a framework that enables us to distinguish between different noncontextuality assumptions in terms of the relationships between the ontological representations of objects in the theory given a certain relation between their operational representations. This framework can be used to motivate and define a "possibilistic" analogue, encapsulating the idea that elements of an operational theory that cannot be unambiguously distinguished operationally can also not be unambiguously distinguished ontologically. We then prove that possibilistic noncontextuality is equivalent to an alternative notion of noncontextuality proposed by Hardy. Finally, we demonstrate that these weaker noncontextuality assumptions are sufficient to prove alternative versions of known "no-go" theorems that constrain ψ-epistemic models for quantum mechanics.

1. Introduction

One way in which quantum mechanics differs strongly from classical mechanics is the existence of incompatible observables; if we want to think that measurements reveal properties of a system, then we must reconcile this with the fact that there exist pairs of properties that cannot be simultaneously measured. By considering the same observables appearing in different *contexts*, that is, measured alongside different sets of other observables, Bell [1], as well as Kochen and Specker [2] showed that measurements could not be thought of as simply revealing underlying properties of the system in a way that was independent of the context in which the observable was measured. This property of quantum mechanics is now referred to as *contextuality*.

Discussions of contextuality often focus on scenarios in which an element of a operational theory such as quantum mechanics manifests itself in two different contexts, such as two different decompositions of a density matrix; or an observable being measured in two different ways, alongside different sets of co-measurable observables. These manifestations are treated identically by the operational theory, always leading to the same probabilities. In fact, this is why the same notation is used for the objects in the first place, as a context-independent symbol is all that is needed to calculate probabilities. However there is no formal argument to be made that these elements which are operationally context-independent should also be ontologically context-independent: this must be taken axiomatically. Motivation for such an axiom cannot be logically deduced or extrapolated from experimental data, but rather must be based on aesthetic principles (e.g., Leibniz's principle [3]). We call such axioms "noncontextuality assumptions". This idea, that operationally indistinguishable objects should also be considered ontologically identical is a form of noncontextuality, henceforth referred to as "probabilistic contextuality". We see that it somewhat encapsulates the hypothesis, reminiscent of Occam's razor, that these objects give rise to identical properties because they are ontologically identical. We will introduce a framework which allows us to compare and contrast different noncontextuality axioms by viewing them as assumptions about ontological properties motivated by operational data.

This framework can be used to motivate and define a "possibilistic" analogue to probabilistic noncontextuality. Possibilistic noncontextuality reflects the assumption that two elements of a physical theory share the same operational possibilities also share the same ontological possibilities. The notion of grouping states by the sets of events that they assign nonzero probability is not entirely novel: this viewpoint emerges naturally in the setting of classical and quantum probability theory from considerations of what it means for different parties to hold different but compatible beliefs about a system such as those of Brun, Finkelstein and Mermin [4] or Caves, Fuchs and Schack [5]. This notion of possibilistic noncontextuality is strictly weaker than that of probabilistic contextuality. We will demonstrate that analogues of some results originally proved using the stronger notion of probabilistic noncontextuality in fact hold with this weaker assumption: for example, it was shown by Spekkens [6] that the assumption of noncontextuality for preparations is incompatible with the operational predictions of quantum mechanics, and it was shown by Morris [7] and by Chen and Montina [8] that any ontological model obeying noncontextuality for measurements must be ψ-ontic. We demonstrate that similar results can be proven using only the weaker, possibilistic, notion of noncontextuality.

2. Notions of Noncontextuality

Noncontextuality is a property of an ontological model (also known as a hidden variable model). In order to formulate a generalized notion of contextuality, we begin by briefly reviewing the framework of ontological models, which allow for realistic descriptions of experimental procedures within an operational theory [6, 9].

2.1. Ontological models and operational theories

An operational theory assigns probabilities $\Pr(k|\mathcal{P},\mathcal{T},\mathcal{M})$ to outcomes k occurring when procedures for a preparation $\mathcal{P}$, a transformation $\mathcal{T}$, and finally a measurement $\mathcal{M}$ are implemented. Any particular $\mathcal{P}$, $\mathcal{M}$ or $\mathcal{T}$ will be denoted an *element* of the operational theory. Quantum mechanics can be regarded as an operational theory by identifying preparation procedures with density matrices ρ, transformations with

unitary operators U acting on density matrices via conjugation, and K-outcome measurements with positive-operator-valued measures (POVM) $\mathcal{M} = \{E_1, \dots E_K\}$. The probability for an outcome k given a preparation described by ρ and a transformation by U is then given by the Born rule,

$$\Pr(k|\rho, U, \mathcal{M}) = \mathrm{tr}[E_k U \rho U^\dagger]. \tag{1}$$

An ontological model for such an operational theory consists of a set Λ, equipped with a σ-algebra Σ, of possible ontic (or "real") states alongside ontological representations of preparations, transformations and measurements, dependent on each other only via states $\lambda \in \Lambda$. A preparation procedure $\mathcal{P}$ is represented within the ontological model by the preparation of a system in an ontic state λ sampled according to some measure $\mu_\mathcal{P}$ on Λ. A transformation $\mathcal{T}$ in the operational theory is represented in the ontological model by a map that, for a system in the ontic state λ, samples a new $\lambda' \in \Lambda$ according to the conditional probability measure $\Gamma_\mathcal{T}(\lambda'|\lambda)$ on Λ. Finally, a K-outcome measurement $\mathcal{M}$ in the operational theory is represented in the ontological model by a response function ξ such that, for a system in the ontic state λ, the measurement outcome $k \in \{1, \dots, K\} =: \mathbb{N}_K$ is sampled from the conditional probability distribution $\{\xi_\mathcal{M}(k|\lambda)\}$ on $\mathbb{N}_K$.

The ontological model reproduces the predictions of the operational theory if

$$\Pr(k|\mathcal{P}, \mathcal{T}, \mathcal{M}) = \int \mathrm{d}\lambda\, \mathrm{d}\lambda'\, \mu_\mathcal{P}(\lambda) \Gamma_\mathcal{T}(\lambda'|\lambda) \xi_\mathcal{M}(k|\lambda') \tag{2}$$

for all $k, \mathcal{P}, \mathcal{T}, \mathcal{M}$. While all constructions in this paper hold for general operational theories, we will only explicitly consider ontological models of quantum mechanics. For brevity, we will omit transformations and focus solely on preparations and measurements, though all definitions can be directly extended to account for transformations.

The Beltrametti-Bugajski model [10] is perhaps the simplest example of an ontological model of pure-state quantum mechanics. In it, pure quantum states are treated as being physical states, so we have $\Lambda = \mathcal{H}$; $\mu_{|\psi\rangle}(|\lambda\rangle) = \delta_{|\psi\rangle}(|\lambda\rangle)$; and $\xi_M(E||\lambda\rangle) = \mathrm{tr}(E|\lambda\rangle\langle\lambda|)$.

We note that for the rest of this paper we will be assuming that our ontological models have a convex structure and therefore preparations (or other operational procedures) can be statistically mixed with each other. Without this assumption many of these noncontextuality relations become trivial. An example of a model without such an assumption is quantum mechanics restricted to pure states, unitary transformations, and projective measurements.

2.2. Operational and ontological relations

In this section, we define a generalized *noncontextuality assumption* to be an assumption that a particular relation at the level of the operational theory implies a (usually similar) relation at the level of the ontological theory. It is a scheme for making conclusions about ontology based on a theory's operational predictions. A noncontextuality assumption allows reasoning about ontological properties based on observable (operational) phenomena.

More formally, an *operational relation* $\sim$ will be a symmetric, reflexive relation over elements in the operational theory (i.e., preparations, transformations and/or measurements); an *ontological relation* $\approx$ is a symmetric, reflexive relation over elements of the ontological model. For clarity, $\sim$ and $\approx$ will always be used to denote operational and ontological relations respectively.

Definition 1 (Noncontextuality Assumption) *A noncontextuality assumption is a statement that when two objects in the operational theory are related by a specified* operational relation $\sim$*, their ontological representations must be related by an* ontological relation $\approx$*. For example, applied to preparations these are statements of the form*

$$\mathcal{P}_1 \sim \mathcal{P}_2 \Rightarrow \mu_{\mathcal{P}_1} \approx \mu_{\mathcal{P}_2} \tag{3}$$

Notably, we do not require that an operational or ontological relation be transitive by definition, and therefore do not restrict ourselves to only consider equivalence relations for this purpose. In practice, however, many interesting noncontextuality assumptions are defined using equivalence relations and elements related by such will be referred to as *equivalent*.

In general, the operational relation is a condition that specifies which elements of the operational theory are related to each other. The noncontextuality assumption and ontological relation then state how the ontic representation of two equivalent elements of the operational theory are related to each other. Given a noncontextuality assumption **X**, we say that an operational theory exhibits **X** contextuality if in any ontological model that correctly reproduces the operational theory there exists some pair of operational elements related by $\overset{\mathbf{x}}{\sim}$ whose ontological representations are not related by $\overset{\mathbf{x}}{\approx}$.

While noncontextuality assumptions can apply to preparations, transformations and measurements, in this section we will use preparations as an example.

2.3. Probabilistic Noncontextuality

As discussed, in quantum mechanics there are many elements such as density matrices, observables, and POVM elements, whose operational behaviour is context-independent. The notion of contextuality due to Spekkens [6] captures this tension by making the assumption that objects that are operationally identical are also ontologically identical. We will present this form of noncontextuality assumption using our notation here.

Definition 2 (ProbOp) *Two preparation procedures $\mathcal{P}_1$ and $\mathcal{P}_2$ in an operational model are probabilistically equivalent (denoted $\mathcal{P}_1 \overset{Prob}{\sim} \mathcal{P}_2$) if for all outcomes k of all measurement procedures $\mathcal{M}$,*

$$\Pr(k|\mathcal{P}_1, \mathcal{M}) = \Pr(k|\mathcal{P}_2, \mathcal{M}). \tag{4}$$

In the case of quantum mechanics, all operational predictions for a preparation procedure $\mathcal{P}$ are completely encoded in the associated density matrix $\rho_{\mathcal{P}}$. Therefore we have

$$\mathcal{P}_1 \overset{\mathbf{Prob}}{\sim} \mathcal{P}_2 \Leftrightarrow \rho_{\mathcal{P}_1} = \rho_{\mathcal{P}_1}. \tag{5}$$

As a particular example, the maximally mixed state can be prepared in many ways, such as by tracing over half of a maximally-entangled pair, randomly applying a unitary operation or by preparing a system in a basis and forgetting which element was prepared. All of these methods are probabilistically equivalent for a single qubit.

Having specified an operational relation, the next step is to specify an ontological relation.

Definition 3 (ProbOn) *Two preparation distributions $\mu_{\mathcal{P}_1}$, $\mu_{\mathcal{P}_2}$ in an ontological model are* probabilistically equivalent*, denoted $\mu_{\mathcal{P}_1} \overset{Prob}{\approx} \mu_{\mathcal{P}_2}$, if and only if $\mu_{\mathcal{P}_1}(\lambda) = \mu_{\mathcal{P}_2}(\lambda)\ \forall\lambda$.*

Definition 4 (Probabilistic Noncontextuality) *For any two preparations $\mathcal{P}_1, \mathcal{P}_2$ related by **ProbOp**, their ontological representations $\mu_{\mathcal{P}_1}, \mu_{\mathcal{P}_2}$ are related by **ProbOn**. That is, the probabilistic noncontextuality assumption states that*

$$\mathcal{P}_1 \overset{Prob}{\sim} \mathcal{P}_2 \Rightarrow \mu_{\mathcal{P}_1} \overset{Prob}{\approx} \mu_{\mathcal{P}_2}. \tag{6}$$

In the quantum case, this translates to assuming

$$\mu_{\mathcal{P}} = \mu_{\rho_{\mathcal{P}}}. \tag{7}$$

So, in a probabilistically noncontextual model, if there are two preparation procedures that cannot be statistically distinguished by any measurements then those preparations lead to identical physical situations. In a probabilistically contextual model, then there are different physical situations that nonetheless lead to identical operational predictions for any measurement: the inability to distinguish them is caused by lack of sufficiently fine-grained measurement.

We can look to electromagnetism in order to find a classical example of an application of this principle. Naïvely, one might expect an operational preparation of an electromagnetic system to correspond to a specific magnetic vector potential as an ontic state. However, different vector potentials correspond to the same experimental predictions if they differ by a curl-free vector field. The application of the probabilistic noncontextuality assumption, then, would be to say that all such preparations are really ontologically identical; they result in the same distribution over ontic states. The vector potential is often said to be

unphysical because of the property that different potentials lead to identical experimental predictions; this line of argument would lead one to conclude that not only are such distributions over ontic states identical, but that they are a delta function. That is, we quotient the state space to form a new ontic state $\tilde{\Lambda} = \Lambda/\!\sim$, where $A \sim A'$ if and only if $\nabla(A - A') = 0$. This will ensure that any such pair of ontic states have the same representative in $\tilde{\Lambda}$. The assumptions leading to such an ontological refinement are thus strictly stronger than those leading only to probabilistic noncontextuality.

We see that both the operational relation **ProbOp** and the ontological relation **ProbOn** are *probabilistic* in that they depend on the full set of (exact) probabilities associated with the object. These probabilistic relations are very restrictive and as a result, $\overset{\mathbf{Prob}}{\sim}$ is a very selective relation, leading to a restriction on ontological models that is more easily satisfied than that of a less selective relation. However, those elements that are related have a very strong ontological condition applied by **ProbOn**.

Contextuality is often considered to be a generalisation of nonlocality; by Fine's theorem [11], the factorisability condition that characterises local distributions in a Bell scenario is interchangeable with the assumption of outcome-determinism for ontological models. With this assumption, nonlocality manifests itself as noncontextuality with respect to contexts chosen jointly by a spacelike separated Alice and Bob. Some notions of contextuality, such as that of Kochen and Specker[2]; Klyachko, Can, Cetiner, Bincioglu and Shumovsky [12]; Abramsky and Brandenburger [13]; and Acín, Fritz, Leverrier and Belén Sainz [14] assume these deterministic models for their notions of contextuality. Probabilistic noncontextuality is a step more general, not necessarily assuming determinism, but recovering these notions of noncontextuality when we restrict our gaze to outcome-deterministic ontological models. It can be considered a generalised form of the type of nonlocality identified by Bell; a scenario demonstrates probabilistic contextuality if its precise operational probabilities cannot be fully explained by a noncontextual (rather than nonlocal) model. As such, under certain assumptions, probabilistic noncontextuality can be detected by inequalities that are robust to noise, as an analogue of Bell inequalities [12, 15].

2.4. Possibilistic Noncontextuality

Another natural choice for both the operational and ontological relations is a *possibilistic* one, in which we consider only the possibilities of operational and ontological events, rather than their probabilities. Such possibilistic considerations also appear naturally in the setting of consistency conditions for agents beliefs about the state of a system, which have been previously studied in the literature. In the work of Brun, Finkelstein and Mermin (BFM) [4], a consistent set of state assignments is a set of density matrices which have some overlap in their operational possibilities. The different parties can assign different density matrices to a system, but as long as there exist some states in the mutual support of all of them, then there is an event that can occur that all of them agree is possible, meaning that their initial state assignments are consistent. In that of Caves, Fuchs and Schack (CFS) [5], the strongest compatibility criterion that can be applied is to ask whether or not the density matrices assigned to the state by each of the parties are in the same equivalence class, defined in this possibilistic sense. If they are, then all of them agree on which events are possible and impossible, even if they disagree on the precise probabilities to assign to each event. Both of these conditions are naturally possibilistic rather than probabilistic.

Such considerations motivate us to define possibilistic operational and ontological relations.

Definition 5 (PossOp) *Two preparation procedures $\mathcal{P}_1$ and $\mathcal{P}_2$ in an operational theory are possibilistically equivalent, denoted $\mathcal{P}_1 \overset{\mathbf{Poss}}{\sim} \mathcal{P}_2$, if for all outcomes k of all measurement procedures $\mathcal{M}$,*

$$\Pr(k|\mathcal{P}_1, \mathcal{M}) = 0 \Leftrightarrow \Pr(k|\mathcal{P}_2, \mathcal{M}) = 0. \tag{8}$$

In the quantum case, two preparation procedures $\mathcal{P}_1$ and $\mathcal{P}_2$ are possibilistically equivalent if and only if the density matrices they give rise to have the same kernel (or equivalently, the same support). The kernel, which is spanned by the states with which ρ has no overlap, completely defines the possibilistic structure of measurements.

$$\mathcal{P}_1 \overset{\mathbf{Poss}}{\sim} \mathcal{P}_2 \Leftrightarrow \ker \rho_{\mathcal{P}_1} = \ker \rho_{\mathcal{P}_2} \tag{9}$$

where $\ker M = \{v : Mv = 0\}$ is the kernel of M.

Definition 6 (Support) *The support of a measure μ is the largest set $S(\mu)$ such that every open set which has non-empty intersection with $S(\mu)$ has nonzero measure.*

More succinctly, but a little less accurately, the support $S(\mu) = \{\lambda : \mu(\lambda) > 0\}$.

Definition 7 (PossOn) *Two preparation distributions $\mu_{\mathcal{P}_1}$, $\mu_{\mathcal{P}_2}$ in an ontological model are* possibilistically equivalent*, denoted $\mu_{\mathcal{P}_1} \overset{\boldsymbol{Poss}}{\approx} \mu_{\mathcal{P}_2}$, if $\mathcal{S}(\mu_{\mathcal{P}_1}) = \mathcal{S}(\mu_{\mathcal{P}_2})$.*

Definition 8 (Possibilistic Noncontextuality) *For any two preparations $\mathcal{P}_1, \mathcal{P}_2$ related by **PossOp**, their ontological representations $\mu_{\mathcal{P}_1}, \mu_{\mathcal{P}_2}$ are related by **PossOn**. That is,*

$$\mathcal{P}_2 \overset{\boldsymbol{Poss}}{\sim} \mathcal{P}_2 \Rightarrow \mu_{\mathcal{P}_1} \overset{\boldsymbol{Poss}}{\approx} \mu_{\mathcal{P}_2}. \tag{10}$$

In the quantum case, this translates to assuming

$$\ker \rho_{\mathcal{P}_1} = \ker \rho_{\mathcal{P}_2} \Rightarrow \mu_{\mathcal{P}_1} \overset{\mathbf{Poss}}{\approx} \mu_{\mathcal{P}_2}. \tag{11}$$

A fully possibilistic noncontextuality assumption, then, would be that if two preparation procedures are possibilistically equivalent operationally, then their ontological representations are also possibilistically equivalent.

A possibilistically noncontextual model for quantum mechanics would provide a natural explanation for why there exist different preparation procedures that cannot be unambiguously discriminated [16]; such preparations lead to ensembles of ontic states that themselves cannot be unambiguously distinguished. Therefore, the lack of the ability to perform unambiguous discrimination in a theory could be interpreted as property of the distributions over ontic states themselves, rather than emerging from a lack of sufficiently fine-grained measurements.

Additionally, we find that the assumption of possibilistic preparation noncontextuality causes the quantum consistency conditions, mentioned above, for density matrix assignment to coincide with classical notions of consistency of beliefs. If we have a collection of agents, each can describe their subjective beliefs about the state of a quantum system via a density matrix ρ_i. According to the BFM consistency criterion, these assignments are consistent if the intersection of the supports of the ρ_i is empty; that is, that there is some subspace of the Hilbert space which each party agrees is in the support of their density matrix. Since the assumption of possibilistic noncontextuality for preparations uniquely defines a set of ontic states associated with this agreed support, this quantum compatibility condition reduces exactly to the classical notion of "strong consistency", which is that there exists some physical state that each agent agrees could represent the true state of the system. The strongest compatibility condition considered by Caves, Fuchs, and Schack, which they denote *equal supports*, is that the supports of each of the ρ_i are identical; in the same way as above, the assumption of possibilistic preparation noncontextuality means this condition reduces to the classical concept of *concordance*: that the agents agree on which states are possible and impossible, although their actual probability assignments may differ.

2.5. Hardy Noncontextuality

We can now bring other previously-studied notions of noncontextuality within this framework. One such notion is known as "Hardy" or "logical" noncontextuality [13, 17, 18], and when applied to preparations corresponds to the assumption that two preparations leading to identical operational predictions must be compatible with the same set of ontic states. As with probabilistic noncontextuality, Hardy noncontextuality has been considered previously under the assumption of outcome determinism, such as in the treatment of Abramsky *et al.* Again, we study a more general notion without this restriction. While probabilistic noncontextuality can be considered a generalisation of Bell nonlocality, Hardy noncontextuality can be considered a generalisation of Hardy's proof of nonlocality. That is, a scenario exhibits Hardy noncontextuality if there are possibilities (rather than detailed probabilities) that cannot be explained within a noncontextual ontological model.

We can formulate Hardy noncontextuality using the operational and ontological relations defined above.

Definition 9 (HardyNC) *An ontological model is Hardy noncontextual iff $\mathcal{P}_1 \overset{\boldsymbol{Prob}}{\sim} \mathcal{P}_2 \Rightarrow \mu_{\mathcal{P}_1} \overset{\boldsymbol{Poss}}{\approx} \mu_{\mathcal{P}_2}$.*

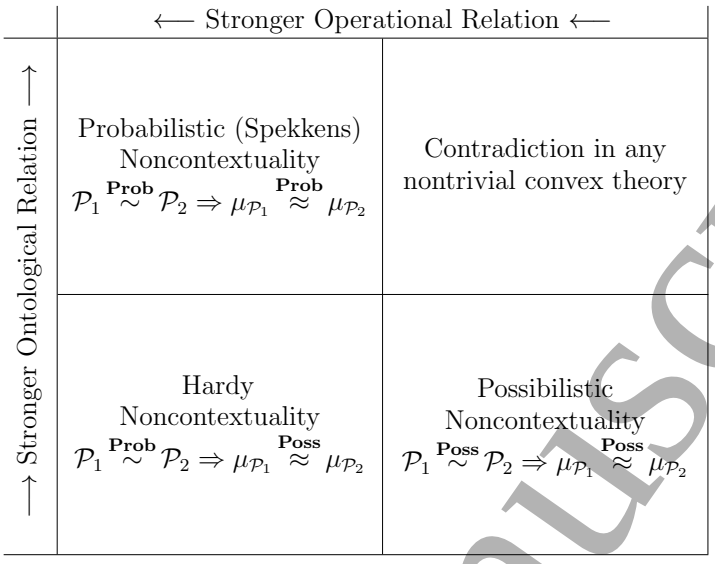

Figure 1. A square showing the different types of contextuality assumptions, under the assumption of convexity. If convexity is not demanded, there are multiple theories that fulfil the stronger form of noncontextuality indicated as a contradiction above, such as quantum mechanics restricted to pure states and unitary operations.

The assumption of Hardy noncontextuality is strictly weaker than probabilistic noncontextuality as it uses the same operational relation but imposes a weaker ontological relation on the identified elements. Furthermore, Hardy noncontextuality is directly implied by possibilistic noncontextuality because

$$\mathcal{P}_1 \overset{\textbf{Prob}}{\sim} \mathcal{P}_2 \Rightarrow \mathcal{P}_1 \overset{\textbf{Poss}}{\sim} \mathcal{P}_2. \tag{12}$$

Intuitively, the assumption of Hardy noncontextuality appears to also be strictly weaker than of possibilistic noncontextuality because the same ontological relation is being enforced in both cases, while possibilistic noncontextuality has a operational relation that is easier to satisfy and so relates more elements of the theory. For example, all density matrices of the form $p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1|$ for $p \in (0,1)$ are *possibilistically* equivalent and yet no two such density matrices are *probabilistically* equivalent.

We now prove, however, that Hardy noncontextuality and possibilistic noncontextuality are equivalent for ontological models of quantum mechanics. Before doing this, we note that any two preparation procedures that result in the same density matrix being prepared must have the same ontic supports under an assumption of possibilistic, probabilistic, or Hardy noncontextuality. Hence, we will use the notation $\mathcal{S}(\mu_\rho) \equiv \mathcal{S}(\mu_{\mathcal{P}})$ for any $\mathcal{P}$ that results in a preparation of ρ.

Theorem 1 *An ontological model of finite-dimensional quantum mechanics is possibilistically noncontextual if and only if it is Hardy noncontextual.*

Proof. We will present the proof for preparation procedures here and defer the proofs for transformations and measurements to the appendix. From eq. (12), possibilistic preparation noncontextuality implies Hardy preparation noncontextuality: any elements related under the probabilistic operational relation are also related under the possibilistic operational relation. So, we only need to prove the converse, that is, that Hardy noncontextuality implies that density matrices with the same kernel can be represented by distributions with the same support over Λ.

Let ρ_0 and ρ_1 be density matrices with the same kernel and with smallest nonzero and largest eigenvalues $\alpha_{j,\min}$ and $\alpha_{j,\max}$ for $j = 0, 1$ respectively. We define

$$\sigma_0 = \frac{1}{1 - \frac{\alpha_{0,\min}}{\alpha_{1,\max}}} \left(\rho_0 - \frac{\alpha_{0,\min}}{\alpha_{1,\max}} \rho_1 \right). \tag{13}$$

We see that σ_0 is positive semi-definite, as the largest eigenvalue of $\frac{\alpha_{0,\min}}{\alpha_{1,\max}}\rho_1$ is $\alpha_{0,\min}$ and ρ_0 and ρ_1 have the same kernel. We note that we must have that $\alpha_{0,\min} \leq \alpha_{1,\max}$, since they have support on subspaces of equal dimensions, say, d, and so both ρ_1 and ρ_0 have d nonzero eigenvalues that sum to 1. Hence, we have $\alpha_{0,\min} = \alpha_{1,\max}$ only in the case that each is the completely mixed state over their support, and $\alpha_{0,\min} < \alpha_{1,\max}$ otherwise. It is easily verified that $\mathrm{tr}(\sigma_0) = 1$, so σ_0 is a density matrix.

Hence ρ_0 can be prepared as a convex mixture of ρ_1 and σ_0 as

$$\rho_0 = \left(1 - \frac{\alpha_{0,\min}}{\alpha_{1,\max}}\right)\sigma_0 + \frac{\alpha_{0,\min}}{\alpha_{1,\max}}\rho_1. \tag{14}$$

Therefore there exists a preparation procedure that prepares ρ_0 such that

$$\mathcal{S}(\mu_{\rho_0}) = \mathcal{S}(\mu_{\sigma_0)} \cup \mathcal{S}(\mu_{\rho_1}) \supseteq \mathcal{S}(\mu_{\rho_1}) \tag{15}$$

We can then repeat the argument for ρ_1. Therefore $\mathcal{S}(\mu_{\rho_0}) = \mathcal{S}(\mu_{\rho_1})$ for any ρ_0, ρ_1 with the same kernel, completing the proof. □

One immediate consequence of this is that it permits the more natural motivation of possibilistic noncontextuality to be applied to the better-known concept of Hardy noncontextuality, providing a motivation for its assumption independent of the fact that it is a weaker noncontextuality assumption than is than probabilistic noncontextuality.

3. Restrictions on ontological models due to possibilistic noncontextuality

The assumption of probabilistic preparation noncontextuality is incompatible with the operational predictions of quantum theory [6]. We now show that even the seemingly weaker notion of possibilistic noncontextuality is already incompatible with quantum theory.

Theorem 2 *Any ontological model of quantum mechanics is possibilistically contextual for preparations.*

Proof. Let

$$|\phi\rangle = \cos\left(\frac{\phi}{2}\right)|0\rangle + \sin\left(\frac{\phi}{2}\right)|1\rangle \tag{16}$$

for some $0 < \phi < \pi/2$. We note that all mixed states are related under the possibilistic operational relation. In particular then, a mixed state made up of $|\phi\rangle\langle\phi|$ and $|-\pi/2\rangle\langle-\pi/2|$, and one made up of $|-\phi\rangle\langle-\phi|$ and $|\pi/2\rangle\langle\pi/2|$ are related. Therefore, such matrices must have the same support.

$$\mathcal{S}(|-\phi\rangle\langle-\phi|) \cup \mathcal{S}\left(\left|\frac{\pi}{2}\right\rangle\left\langle\frac{\pi}{2}\right|\right) = \mathcal{S}(|\phi\rangle\langle\phi|) \cup \mathcal{S}\left(\left|-\frac{\pi}{2}\right\rangle\left\langle-\frac{\pi}{2}\right|\right). \tag{17}$$

Taking the intersection of both sides with $\mathcal{S}(\left|\frac{\pi}{2}\right\rangle\left\langle\frac{\pi}{2}\right|)$ gives

$$\mathcal{S}\left(\left|\frac{\pi}{2}\right\rangle\left\langle\frac{\pi}{2}\right|\right) = \left[\mathcal{S}(|\phi\rangle\langle\phi|) \cap \mathcal{S}\left(\left|\frac{\pi}{2}\right\rangle\left\langle\frac{\pi}{2}\right|\right)\right] \cup \left[\mathcal{S}\left(\left|-\frac{\pi}{2}\right\rangle\left\langle-\frac{\pi}{2}\right|\right) \cap \mathcal{S}\left(\left|\frac{\pi}{2}\right\rangle\left\langle\frac{\pi}{2}\right|\right)\right] \tag{18}$$

$$= \left[\mathcal{S}(|\phi\rangle\langle\phi|) \cap \mathcal{S}\left(\left|\frac{\pi}{2}\right\rangle\left\langle\frac{\pi}{2}\right|\right)\right] \tag{19}$$

$$\subseteq \mathcal{S}(|\phi\rangle\langle\phi|) \tag{20}$$

where the second line follows since $\left|\frac{\pi}{2}\right\rangle$ and $\left|-\frac{\pi}{2}\right\rangle$ are completely operationally distinguishable, and hence must be fully ontologically distinct. This shows that the set of states consistent with $|\pi/2\rangle\langle\pi/2|$ is a subset of those consistent with $|\phi\rangle\langle\phi|$, and so any operational predictions compatible with a preparation of $|\pi/2\rangle\langle\pi/2|$must also be consistent with a preparation of $|\phi\rangle\langle\phi|$. However, in order to be consistent with quantum mechanics, preparing $\left|\frac{\pi}{2}\right\rangle$ and measuring in the basis $\{|\phi\rangle, |\pi + \phi\rangle\}$ must give the outcome $|\pi + \phi\rangle$ with nonzero probability. □

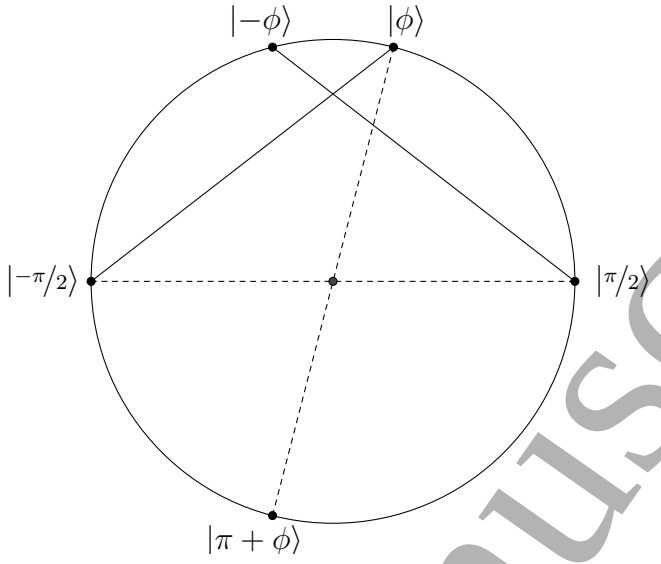

Figure 2. Diagrams showing states that can be used in the proof of theorem 2.

The above proof is similar to Hardy's proof of Bell's theorem based only upon possibilistic arguments [19]. In some sense, considering these possibilistic restrictions on ontological models is forcing them to obey the same logical structure as the operational theory, in this case quantum mechanics. This is because we can think of working with possibilities as a generalised probability model, in which we work with the Boolean (logical) semiring in which an event is only prescribed the symbol 0 (impossible) or 1 (possible).

Another known restriction on ontological models under the assumption of probabilistic measurement noncontextuality is that such models cannot be ψ-epistemic, a result first shown by Morris [7] and by Chen and Montina [8]. Here, we will show an analogous result, in which we use a strictly weaker assumption than probabilistic measurement noncontextuality.

We define two sets associated with a response function for outcome k of a measurement $\mathcal{M}$ in an ontological model:

$$\mathcal{R}[\xi_M(k|\lambda)] := \{\lambda : \xi_{\mathcal{M}}(k|\lambda) = 1\} \tag{21}$$

$$\mathcal{T}[\xi_M(k|\lambda)] := \{\lambda : \xi_{\mathcal{M}}(k|\lambda) > 0\}. \tag{22}$$

We will show that any ontological model in which these sets depend only upon the POVM element E_k associated with outcome k, and so can be denoted by $\mathcal{R}(E_k)$ and $\mathcal{T}(E_k)$ respectively, imposes restrictions on which states can have ontic overlap. We will call such a model *faithful*, as it is a stronger form of the faithfulness condition introduced in [20]. We note that this condition is implied by probabilistic noncontextuality.

Theorem 3 *Let $|\psi_1\rangle$ and $|\psi_2\rangle$ be two states in a Hilbert space of dimension $d \geq 2^N$, with $|\langle\psi_1|\psi_2\rangle|^2 \leq \cos^{2N}(2\tan^{-1}(2^{1/2N}-1))$. Then in any faithful ontological model they have disjoint supports:*

$$\mathcal{S}(\mu_{|\psi_1\rangle}) \cap \mathcal{S}(\mu_{|\psi_2\rangle}) = \emptyset. \tag{23}$$

Proof. Let $\mathcal{P}_1$ and $\mathcal{P}_2$ be fixed preparation procedures for $|\psi_1\rangle$ and $|\psi_2\rangle$ and let $|\psi_1\rangle = |a\rangle^{\otimes N}$ and $|\psi_2\rangle = |b\rangle^{\otimes N}$ where $|a\rangle = |\phi\rangle$ and $|b\rangle = |-\phi\rangle$ from eq. (16). This form for our states can be made without loss of generality, since for any pair of states $|\varphi_1\rangle$, $|\varphi_2\rangle$, there exists a unitary rotation mapping them into a pair of states of our chosen form, and unitary actions must preserve overlaps of support. If $|\langle\varphi_1|\varphi_2\rangle|^2 = |\langle\psi_1|\psi_2\rangle|^2$, then there exists a unitary U such that $U|\varphi_1\rangle = |\psi_1\rangle$ and $U|\varphi_2\rangle = |\psi_2\rangle$, and so any state that is in the support of $|\varphi_1\rangle$ and $|\varphi_2\rangle$ is mapped by such a transformation into a state that is in the support of $|\psi_1\rangle$ and $|\psi_2\rangle$ since this is a legitimate preparation procedure for those states.

By considering each of the measurements formed by independently measuring $\{|a\rangle, |\bar{a}\rangle\}$ or $\{|b\rangle, |\bar{b}\rangle\}$ on each of the N qubits (where $|\bar{a}\rangle$ denotes the orthogonal state to $|a\rangle$, and similarly for $|b\rangle$), we have

$$\mathcal{S}(\mu_{|\psi_1\rangle}) \cap \mathcal{S}(\mu_{|\psi_2\rangle}) \subseteq \cap_{\vec{\nu}\in\{a,b\}^N} \mathcal{R}(|\vec{\nu}\rangle\langle\vec{\nu}|), \tag{24}$$

except perhaps on an unphysical set of measure 0, as either $|a\rangle^{\otimes N}$ or $|b\rangle^{\otimes N}$ is orthogonal to all but one outcome in each measurement. To see this, consider the case in which we measure with respect to $\{|a\rangle, |\bar{a}\rangle\}$ on the first qubit, and $\{|b\rangle, |\bar{b}\rangle\}$ on the second. Here, $|a\rangle|a\rangle$ is incompatible with $|\bar{a}\rangle|b\rangle$ and $|\bar{a}\rangle|\bar{b}\rangle$, with a similar condition holding for $|b\rangle|b\rangle$. The only outcome that is compatible with both $|a\rangle|a\rangle$ and $|b\rangle|b\rangle$ is $|a\rangle|b\rangle$. We note that by assumption, we have $|\langle a|b\rangle|^2 \leq \cos^2(2\tan^{-1}(2^{1/2N} - 1))$ and so we can apply the construction in [21], so there exists a rank-1 projector-valued measure (PVM) $M = \{\Pi_{\vec{\nu}} : \vec{\nu} \in \{a, b\}^N\}$ such that $\Pi_{\vec{\nu}}|\vec{\nu}\rangle\langle\vec{\nu}| = 0$ for all $\vec{\nu}$ [21]. For all $\vec{\nu}$, $\Pi_{\vec{\nu}}|\vec{\nu}\rangle\langle\vec{\nu}| = 0$ and so there exist PVMs containing $\Pi_{\vec{\nu}}$ and $|\vec{\nu}\rangle\langle\vec{\nu}|$, for example $\{|\vec{\nu}\rangle\langle\vec{\nu}|, \Pi_{\vec{\nu}}, \mathbb{1} - |\vec{\nu}\rangle\langle\vec{\nu}| - \Pi_{\vec{\nu}}\}$. Hence

$$\mathcal{R}(|\vec{\nu}\rangle\langle\vec{\nu}|) \cap \mathcal{T}(\Pi_{\vec{\nu}}) = \emptyset \tag{25}$$

for all $\vec{\nu}$, since a lambda in the intersection of $\mathcal{R}(|\vec{\nu}\rangle\langle\vec{\nu}|$ and $\mathcal{T}(\Pi_{\vec{\nu}})$ would, in a measurement containing both such as the one given above, have to yield the $|\vec{\nu}\rangle\langle\vec{\nu}|$ outcome with probability 1, and the $\Pi_{\vec{\nu}}$ outcome with nonzero probability, a clear contradiction. Therefore

$$\begin{aligned}
\mathcal{S}(\mu_{|\psi_1\rangle}) \cap \mathcal{S}(\mu_{|\psi_2\rangle}) &\subseteq \left[\cap_{\vec{\nu}\in\{a,b\}^n} \mathcal{R}(|\vec{\nu}\rangle\langle\vec{\nu}|)\right] \cap \Lambda & (26)\\
&\subseteq \left[\cap_{\vec{\nu}\in\{a,b\}^n} \mathcal{R}(|\vec{\nu}\rangle\langle\vec{\nu}|)\right] \cap \left[\cup_{\vec{\nu}\in\{a,b\}^n} \mathcal{T}(\Pi_{\vec{\nu}})\right] & (27)\\
&\subseteq \emptyset, & (28)
\end{aligned}$$

giving the desired result, where the second line follows as some outcome must occur when the measurement M is performed. □

Since $\lim_{N\to\infty} \cos^{2N}(2\tan^{-1}(2^{1/2N} - 1)) = 1$, which can be seen by observing that the expression has a Laurent series expansion at infinity of $1 - 1/x + O(1/x^2)$, any two states can be shown to be ontologically disjoint if we are equipped with a large enough Hilbert space. The result does however differ from those of Morris [7] and of Chen and Montina [8] insofar as this dimensional constraint is present. As in the PBR theorem [21], we need a large quantum dimension in order to be able to have a measurement that can distinguish between states with large overlap. We note that in particular, any faithful ontological models for an infinite dimensional space will be ψ-ontic, since no two such states can share any ontological support. In general, any quantum system can be thought of as part of a larger quantum system, so results such as this demonstrate an inherent tension between possibilistic measurement noncontextuality and ψ-epistemic ontological models. Theorem 3 can also be thought of as demonstrating that the assumption of preparation independence used by PBR can be replaced by that of faithfulness, as applied here.

We are left to consider which noncontextuality assumptions mandate the property of faithfulness for ontological models. In particular, is it a consequence of any of the noncontextuality assumptions we have explored? We shall see that this is dependent on the requirements one sets on an operational theory. We shall need to introduce the formulation of probabilistic noncontextuality as it is applied to measurements.

Definition 10 (ProbOp-M) *Two measurement outcomes $k^{(1)}_{\mathcal{M}_1}$, belonging to a measurement procedure $\mathcal{M}_1$, and $k^{(2)}_{\mathcal{M}_2}$, belonging to a measurement procedure $\mathcal{M}_2$, in an operational model are probabilistically equivalent (denoted $k^{(1)}_{\mathcal{M}_1} \overset{Prob}{\sim} k^{(2)}_{\mathcal{P}_2}$) if for all preparation procedures $\mathcal{P}$,*

$$\Pr(k^{(1)}|\mathcal{P}, \mathcal{M}_1) = \Pr(k^{(2)}|\mathcal{P}, \mathcal{M}_2). \tag{29}$$

Definition 11 (ProbOn-M) *Two measurement effects $\xi_{M_1}(k_1|\lambda)$, $\xi_{M_2}(k_2|\lambda)$ in an ontological model are* probabilistically equivalent*, denoted $\xi_{M_1}(k_1|\lambda) \overset{Prob}{\approx} \xi_{M_2}(k_2|\lambda)$, if and only if $\xi_{M_1}(k_1|\lambda) = \xi_{M_2}(k_2|\lambda)$ $\forall\lambda$.*

Definition 12 (Probabilistic Noncontextuality for Measurements) *For any two preparations $k^{(1)}_{\mathcal{M}_1} k^{(2)}_{\mathcal{M}_2}$ related by **ProbOp-M**, their ontological representations $\xi_{M_1}(k_1|\lambda), \xi_{M_2}(k_2|\lambda)$ are related by **ProbOn-M**. That is, the probabilistic noncontextuality assumption states that*

$$\mathcal{P}_1 \overset{\textbf{Prob}}{\sim} \mathcal{P}_2 \Rightarrow \xi_{M_1}(k_1|\lambda) \overset{\textbf{Prob}}{\approx} \xi_{M_2}(k_2|\lambda). \tag{30}$$

We see that the notion of measurement noncontextuality captures two related ideas: that any way of performing a measurement (for example, a specific Naimark dilation for a POVM) yields to ontologically identical predictions; and that the probabilities of specific outcomes are independent of the context in which they appear.

Under this assumption of possibilistic measurement noncontextuality, first note that both the conditions $\xi_M(k|\lambda) = 1$ and $\xi_M(k|\lambda) > 0$ can be recognised by considering only the possibilities of alternative outcomes. By the definition of possibilistic noncontextuality, we then demand that these possibilities, and equally the impossibilities, are independent of the specific context in which an observable appears. However, for a system in a given ontic state one cannot necessarily conclude that an outcome certain in one context will be certain in all other contexts in which it appears. The noncontextuality assumption taken at face value will only tell us that it cannot be impossible in any other context. However, we can note that in any context in which this outcome, E, appears, we can coarse-grain the other measurement results by "forgetting" which of the alternative outcomes happened. These coarse-grainings yield identical probabilistic predictions, and in a context in which E is certain, the coarse-graining is impossible. Therefore, these coarse-grainings are impossible in all contexts, and E is certain in all contexts in which it appears.

The weakness of this argument lies in whether or not we can really consider the coarse-grainings to yield identical probabilistic predictions with each other. If our scenario is restricted to preparing a quantum state and then measuring it, then this assumption is justified; this is the position taken for example by Spekkens [6]. However, if we consider that a second measurement procedure could be implemented afterwards, necessitating the existence of a state update rule, then the assumption does not hold and we must use a different notion of noncontextuality. One such candidate would be a *trichotomistic* notion of noncontextuality, in which our ontological relation, as applied to two related measurements, is that each outcome must be considered impossible, possible-but-not-certain, or certain, independent of the context in which it appears. We can assign these classes the symbols 0, 1/2 and 1‡. In any case, this trichotomistic notion of contextuality is also strictly weaker than probabilistic noncontextuality and so a weaker notion is used for this proof regardless of one's position of the equivalence of these coarse-grainings over measurement outcomes.

4. Tests of possibilistic noncontextuality cannot be robust to experimental error

Recently, probabilistic noncontextuality inequalities have been demonstrated and subsequently shown to be experimentally violated [22, 23]. In order to avoid finite precision loopholes, the probabilistic noncontextuality inequalities are derived assuming probabilistic noncontextuality for both preparations and measurements. However, as we now prove, no such inequalities can be demonstrated for possibilistic noncontextuality.

Theorem 4 *Any operational prediction of quantum mechanics can be approximated arbitrarily well by an ontological model obeying universal possibilistic noncontextuality; that is, the conjunction of possibilistic noncontextuality for preparations and for measurements.*

Proof. Let $\rho_\epsilon = (1-\epsilon)\rho + \epsilon I_d/d$ for any fixed $\epsilon > 0$ and any state ρ. For any two density matrices ρ and σ,

$$\rho_\epsilon \overset{\textbf{Poss}}{\sim} \sigma_\epsilon. \tag{31}$$

Therefore we can approximate all quantum mechanical predictions by operators that are all possibilistically equivalent. The Beltrametti-Bugajski model then gives a possibilistic noncontextual model for all states as follows. Let $\Lambda = \mathcal{H}$ and $\mu_{|\psi_\epsilon\rangle}(|\lambda\rangle) = (1-\epsilon)\delta(\lambda - \psi) + \epsilon\mu_H$ where μ_H is the uniform Haar measure. We can

‡ This concept of noncontextuality is impossible to replicate within the Abramsky-Brandenburger sheaf-theoretic framework because these values do not form a semiring; coarse-graining two elements that are possible-but-not-certain can be either possible-but-not-certain or certain. Any framework for contextuality based on matching probabilistic predictions in semirings cannot express this sort of notion.

then extend the model to general mixed states by taking convex combinations of $|\psi_\epsilon\rangle\langle\psi_\epsilon|$, so that $\mathcal{S}(\rho_\epsilon) = \mathcal{H}$ for all states ρ_ϵ. The same argument also applies to transformations and measurements. □

Now, any proposed noncontextuality inequality for probabilistic noncontextuality will necessarily be a function only of observed test statistics compared to some constant value, so we can see that for any inequality, there exist universally possibilistically noncontextual models that differ from the quantum statistics by an arbitrarily small value. Further, any such inequality of some use would have to be saturated by quantum mechanics. Thus, finite experimental uncertainty makes impossible the existence of any conclusive proof that reality is not universally possibilistically noncontextual. However, were reality probabilistically noncontextual, there would be some amount by which the operational predictions would have to be altered and this would be experimentally verifiable. Hence, while any particular possibilistically noncontextual theory can be experimentally refuted, it is not possible to, for example using an analogue of a Bell inequality, to refute the entire set of possibilistically noncontextual theories. It is worth noting that while there are claims [24] to have an experimental verification of Hardy's theorem, these experiments actually verify a *probabilistic* version of Hardy's paradox, enabling the existence of inequalities that are not ϵ-closely approximated by quantum-mechanical predictions.

5. Further generalisation

In the previous sections we have developed a framework for expressing arbitrary examples of noncontextuality relations, and explored the relationships between different forms. In this section, we will demonstrate the power of this framework to consider novel forms of contextuality assumptions, and explore which ontological and operational relations are sensible under certain subjective criteria. Up to this point, all of our operational and ontological relations have been equivalence relations, although our definition of such relations does not require transitivity. In fact, there are natural motivations for notions of noncontextuality that do not have this property.

We could choose an operational relation identifying elements which have an operational distinguishability of no more than ϵ: a nontransitive assumption. Likewise, we can have an ontological relation that says that related preparation distributions should have a classical distinguishability bounded above by $f(\epsilon)$ for some f; enforcing a "similar elements of the theory are represented similarly in the ontological model" idea. An idea of this kind has been explored by Winter [25].

Definition 13 ($\overset{\epsilon}{\sim}$) *For two preparation procedures $\mathcal{P}_1$ and $\mathcal{P}_2$, we have $\mathcal{P}_1 \sim_\epsilon \mathcal{P}_2 \Leftrightarrow D(\mathcal{P}_1, \mathcal{P}_2) \leq \epsilon$, where $D(\mathcal{P}_1, \mathcal{P}_2)$ is the operational distinguishability of the two preparations. Assuming quantum theory, this is given by $D(\mathcal{P}_1, \mathcal{P}_2) = |\rho_1 - \rho_2|_1$, where ρ_1 and ρ_2 are the density operators associated with preparations $\mathcal{P}_1$ and $\mathcal{P}_2$ respectively, and $|\cdot|_1$ is the trace norm.*

Definition 14 ($\overset{\epsilon}{\approx}_f$) *Two preparation distributions μ_1, μ_2 are related $\mu_1 \approx_\epsilon \mu_2$ if and only if $\int_\Lambda |\mu_{\mathcal{P}_1} - \mu_{\mathcal{P}_2}| \, \mathrm{d}\lambda \leq f(\epsilon)$*

Up until now, a notion of noncontextuality has consisted of an operational relation, an ontological relation, and a simple noncontextuality assumption saying that the former implies the latter. Here we see that we have sets of operational and ontological relations, and our noncontextuality assumption can be thought of as a kind of axiom schema: for any preparation procedures $\mathcal{P}_1 \overset{\epsilon}{\sim} \mathcal{P}_2$, we have $\mu_{\mathcal{P}_1} \overset{\epsilon}{\approx}_f \mu_{\mathcal{P}_2}$.

6. Concluding Remarks

We have introduced a general framework for the postulation and interpretation of noncontextuality assumptions; we have seen that this view of a noncontextuality assumption as a statement that allows inference of ontological properties from operational ones is both powerful in its descriptive capacity and its ability to highlight novel assumptions.

Using this framework, we have explored weaker notions of contextuality than that of Spekkens[6], and we have seen that one such noncontextuality assumption of this kind, possibilistic contextuality, encapsulates the kind of contextuality present in Hardy's proof of Bell's theorem, and that probabilistic contextuality

encapsulates the kind present in that of Bell. An interesting open question, then, is to ask what the corresponding contextuality assumptions are that encapsulate Kochen-Specker contextuality, the analogue of the kind present in the GHZ proof of Bell's theorem [26].

For any scientific realist, the ultimate aim of scientific inquiry is to be able to make statements about the true state of the world, as far as is possible. To be able to make any statements about the ontological nature of the world, we need some sort of noncontextuality assumption in order to allow our operational knowledge to transfer into this domain. Knowing, then, which noncontextuality assumptions are tenable within a given operational scenario, and the relative strengths of these assumptions, is essential in making any claim, however tentative, about the real nature of things.

Acknowledgements

AWS would like to thank Angela Xu for her helpful discussions, and acknowledges support from Cambridge Quantum Computing Limited, and from EPSRC, *via* the Centre for Doctoral Training in Controlled Quantum Dynamics. SDB acknowledges support from the ARC *via* the Centre of Excellence in Engineered Quantum Systems (EQuS), project number CE110001013.

[1] Bell J S 2002 *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press)
[2] Kochen S and Specker E 1967 *Journal of Mathematics and Mechanics* **17**
[3] Spekkens R W 2016 Leibniz's principle of the identity of indiscernables as a foundational principle for quantum theory information-Theoretic interpretations of quantum mechanics
[4] Brun T A, Finkelstein J and Mermin N D 2002 *Phys. Rev. A* **65**
[5] Caves C, Fuchs C and Schack R 2002 *Phys. Rev. A* **66**
[6] Spekkens R W 2005 *Phys. Rev. A* **71**
[7] Morris R 2009 Topics in quantum foundations: Ontological models, and distinguishability as a resource
[8] Chen Z and Montina A 2011 *Phys. Rev. A* **83**
[9] Harrigan N and Rudolph T 2007 URL `http://arxiv.org/abs/0709.4266v1`
[10] Beltrametti E G and Bugajski S 1995 *J. Phys. A: Math Gen* **28** 3329–3343
[11] Fine A 1982 *Phys. Rev. Lett.* **48**
[12] Klyachko A A, Can M A, Cetiner S B and Shumovsky A S 2008 *Phys. Rev. Lett.* **101**
[13] Abramsky S and Brandenburger A 2011 *New Journal of Physics* **13**
[14] Acin A, Fritz T, Leverrier A and Belén Sainz A 2015 *Comm. Math. Phys* **334**
[15] experimental test of noncontextuality without unphysical idealizations A 2016 *Nature Communications*
[16] Ivanovic I D 1987 *Phys. Lett. A* **123** 257–259
[17] Lal R 2014 Personal Communication
[18] Liang Y C, Spekkens R W and Wiseman H M 2011 *Physics Reports* **506** 1–39
[19] Hardy L 1993 *Phys. Rev. Lett.* **71**(11) 1665–1668 URL `http://link.aps.org/doi/10.1103/PhysRevLett.71.1665`
[20] Rudolph T 2006 Ontological models for quantum mechanics and the Kochen-Specker theorem arXiv:quant-ph/0608120
[21] Pusey M F, Barrett J and Rudolph T 2012 *Nature Physics* **8** 475–478
[22] Kunjwal R and Spekkens R W 2015 *Phys. Rev. Lett.* **115**
[23] Mazurek M D, Pusey M F, Kunjwal R, Resch K J and Spekkens R W An experimental test of noncontextuality without unwarranted idealisations
[24] Lundeen J S and Steinberg A M 2009 *Phys. Rev. Lett.* **102**
[25] Winter A 2014 *Journal of Physics A: Mathematical and Theoretical* **47**
[26] Greenberger D, Horne M, Shimony A and Zeilinger A 1990 *American Journal of Physics* **58** 1131–43

Appendix A.

Just as we defined the support of a preparation procedure above, we will denote the *support of a measurement effect* with response function $\xi(\lambda)$ as

$$\mathcal{R}(\xi) := \{\lambda \in \Lambda | \xi(\lambda) > 0\}. \tag{A.1}$$

We can also define a notion of the *support of a transformation procedure* as

$$\mathcal{U}(\Gamma) := \{\{\lambda_1, \lambda_2\} \in \Lambda \otimes \Lambda | \Gamma(\lambda_2|\lambda_1) > 0\}. \tag{A.2}$$

Theorem 5 *Possibilistic measurement noncontextuality is equivalent to Hardy measurement noncontextuality.*

Proof. Again, it is clear that possibilistic measurement noncontextuality implies Hardy measurement noncontextuality, so we only need to prove the converse.

Quantum mechanically, POVM elements are positive semidefinite matrices, and the assumption of possibilistic measurement noncontextuality is the identification of the ontic supports of any two such preparation procedures whose associated density matrices share the same kernel over $\mathbb{R}^2$. The proof here echoes the proof above strongly: if we have two matrices E_1 and E_2, that share a kernel, and correspond to two measurement effects, we can write $E_1 = aE_2 + bE_3$, for E_3 some matrix with a support strictly contained within that of the other E_i. In a similar fashion to the preparations case, we can construct a POVM element with the same matrix as E_1 as a convex sum of E_2 and E_3, demonstrating that $\mathcal{R}(E_1) = \mathcal{R}(E_2) \cup \mathcal{R}(E_3)$. We also have, by symmetry $\mathcal{R}(E_2) = \mathcal{R}(E_1) \cup \mathcal{R}(E_3)$, and therefore $\mathcal{R}(E_1) = \mathcal{R}(E_2) = \mathcal{R}(\not\Vdash_{\mathrm{supp}(E_i)})$. This completes the proof. □

The above proofs admit a slight generalisation, in that they can show that the possibilistic and probabilistic operational relations lead to identical restrictions for any operational relation based on a function $f(x)$ with $f(x) = f(0)$ iff $x = 0$, and $f(x + y) \geq f(x)$, as defined in section 5.

We note that in this proof, we are explicitly using the same assumption that we called out in the proof of theorem 3, namely that we are considering a coarse-graining of two measurement outcomes via classical post-processing to be possibilistically identical to an ontological coarse graining of the relevant POVM elements. In that section, dropping this implicit assumption weakened our axioms and required we move to an additional noncontextuality assumption. However, for this result, dropping this assumption makes the proof trivial because we are proving a much weaker statement. It can be easily checked that this proof, therefore, holds in both the case in which this assumption is made or not.

Theorem 6 *Possibilistic transformation noncontextuality is equivalent to Hardy transformation noncontextuality.*

Proof. Once again, it is clear that possibilistic transformation noncontextuality implies Hardy transformation noncontextuality, so we only need to prove the converse.

Consider two transformation procedures Γ_1 and Γ_2 with the property that $\forall \xi, \mu$,

$$\int_\Lambda \mathrm{d}\lambda \mathrm{d}\lambda' \mu(\lambda) \Gamma_1(\lambda'|\lambda) \xi(\lambda') > 0 \Leftrightarrow \int_\Lambda \mathrm{d}\lambda \mathrm{d}\lambda' \mu(\lambda) \Gamma_2(\lambda'|\lambda) \xi(\lambda') > 0. \tag{A.3}$$

In general, these are associated with some representations of completely positive, trace preserving maps T_1 and T_2. Consider their action on the state $|\psi\rangle\langle\psi|$: they map it to $\sigma_1 = T_1|\psi\rangle\langle\psi|T_1^\dagger$, and $\sigma_2 = T_2|\psi\rangle\langle\psi|T_2^\dagger$ respectively. These two matrices must share a kernel, or else we would be in contradiction with our assumed property. An example of two maps meeting these criteria might be two dephasing channels with different dephasing strengths.

Note that the most general kind of measurement we can perform to enact tomography on such a transformation procedure is to prepare some entangled state, send part of the entangled state through the transformation procedure, and then follow this up by a joint measurement. By the possibilistic notion of noncontextuality, outcomes of such experiments must yield either a zero probability for both transformations, or a nonzero probability for both transformations. Consider now the Choi-Jamiolkowski isomorphism as applied to our two CPTP maps T_i, leading to two channel-states $\tau^{(i)}$. We can see that in general, we require for an entangled initial state σ_{AB} and an entangled measurement E_{AC}, that

$$\mathrm{tr}\left(\sigma_{AB}\tau^{(1)}_{BC}E_{AC}\right) > 0 \quad \Longleftrightarrow \quad \mathrm{tr}\left(\sigma_{AB}\tau^{(2)}_{BC}E_{AC}\right) > 0. \tag{A.4}$$

Taking a trace over C for appropriately chosen E_{AC}, *viz* a Bell state $|\Phi_+\rangle\langle\Phi_+|$ reduces this to the already-proved case of preparation noncontextuality. □