Is the quantum state real? A review of ψ-ontology theorems

M. S. Leifer*†
Perimeter Institute for Theoretical Physics
31 Caroline St. N., Waterloo, ON, Canada, N2L 2Y5.

March 10, 2014

Abstract

Towards the end of 2011, Pusey, Barrett and Rudolph (PBR) derived a theorem that aimed to show that the quantum state must be ontic (a state of reality) in a broad class of realist approaches to quantum theory. This result attracted a lot of attention and controversy. The aim of this review article is to to review the background to the PBR Theorem, to provide a clear presentation of the theorem itself, and to review related work that has appeared since the publication of the PBR paper. In particular, this review:

- Explains what it means for the quantum state to be ontic or epistemic (a state of knowledge).
- Reviews arguments for and against an ontic interpretation of the quantum state as they existed prior to the PBR Theorem.
- Explains why proving the reality of the quantum state is a very strong constraint on realist theories in that it would imply many of the known no-go theorems, such as Bell's Theorem and the need for an exponentially large ontic state space.
- Provides a comprehensive presentation of the PBR Theorem itself, along with subsequent improvements and criticisms of its assumptions.
- Reviews two other arguments for the reality of the quantum state: the first due to Hardy and the second due to Colbeck and Renner, and explains why their assumptions are less compelling than the PBR Theorem.
- Reviews subsequent work aimed at ruling out stronger notions of what it means for the quantum state to be epistemic and points out open questions in this area.

The overall aim is not only to provide the background needed for the novice in this area to understand the current status, but also to discuss often overlooked subtleties that should be of interest to the experts.

Contents

*Email: matt@mattleifer.info
†Website: http://mattleifer.info

1 Introduction

In 1964, John Bell fundamentally changed the way that we think about quantum theory [1]. Shimony famously referred to tests of Bell's Theorem as "experimental metaphysics" [2], but I disagree with this characterization. What Bell's Theorem really shows us is that the foundations of quantum theory is a bona fide field physics, in which questions are to be resolved by rigorous argument and experiment, rather than remaining the subject of open-ended debate. In other words, it is a mistake to crudely divide quantum theory into its practical part and its interpretation, and to think of the latter as metaphysics, experimental or otherwise.

In the wake of Bell's Theorem, the study of entanglement and nonlocality has become a mainstream field of physics, particularly in light of its practical applications in quantum information science, but Bell's broader lesson—that the interpretation quantum theory should be approached as a rigorous science—has rather been missed. This is nowhere more evident than in debates about the status of the quantum state. The question of just what type of thing the quantum state, or wavefunction, represents, has been with us since the beginnings of quantum theory. The likes of de Broglie and Schrödinger initially wanted to view the wavefunction as a real physical wave, just like the waves of classical field theory, with perhaps some additional structure to account for particle-like or "quantum" properties [3]. In contrast, following Born's introduction of the probability rule [4], the Copenhagen interpretation advocated by Bohr, Heisenberg, Pauli et. al. came to view the wavefunction as a "probability wave" and denied the need for a more fundamental reality to underlie it [5]. In modern terms, most realist interpretations of quantum theory; such as many-worlds [6–8], de Broglie-Bohm theory [9–12], spontaneous collapse theories [13, 14], and modal interpretations [15]; view the wavefunction as part of reality, whereas those that follow more Copenhagenish lines [16–26] tend to view it as a representation of knowledge, information, or belief. The big advantage of the latter view is that the notorious collapse of the wavefunction can be explained as the effect of acquiring new information, no more serious than the updating of classical probabilities in the light of new data, rather than as an anomalous physical process that needs to be eliminated or explained away.

The question then is whether this is a necessary dichotomy. Is the only way to avoid having this weird multidimensional object as part of reality to give up on reality altogether, or can we reach a compromise in which there is a well-founded reality, but one in which the wavefunction only represents knowledge? This seems like a question that is ripe for attacking with the kind of rigour that Bell brought to nonlocality, and indeed Pusey, Barrett and Rudolph (PBR) have recently proven a theorem to the effect that the wavefunction

must be ontic (i.e. a state of reality), as opposed to epistemic (i.e. a state of knowledge) in a broad class of realist approaches to quantum theory [27].

Since then, there has been much discussion and criticism of the PBR Theorem in both formal [28–39] and informal venues [40–49], as well as a couple of attempts to derive the same conclusion as PBR from different assumptions [50, 51]. The PBR Theorem and its successors all employ auxiliary assumptions of varying degrees of reasonableness. Without these assumptions, it has been shown that the wavefunction may be epistemic [52]. Therefore, there has also been subsequent work aimed at ruling out stronger notions of what it means for the wavefunction to be epistemic, without using such auxiliary assumptions [53–58]. The aim of this review article is to provide the background necessary for understanding these results, to provide a comprehensive presentation and criticism of them, and to explain their implications.

One of the most intriguing things about proving that the wavefunction must be ontic is that it would imply a large number of existing no-go results, including Bell's Theorem [1] and excess baggage theorems [59–61](i.e. showing that the size of the ontic state space must be infinite and must scale exponentially with the number of systems). Therefore, even apart from its foundational significance, proving the reality of the wavefunction could potentially provide a powerful unification of the known no-go theorems, and may have applications in quantum information theory.

My aim is that this review should be accessible to as wide an audience as possible, but I have made three decisions about how to present the material that make my treatment somewhat more involved than those found elsewhere in the literature. Firstly, I adopt rigorous measure theoretic probability theory. It is common in the literature to specialize to finite sample spaces or to adopt physicist's rigour with respect to continuous spaces. The latter basically involves proving all results as if you were dealing with smooth and continuous probability densities and then hoping everything still works when you throw in a bunch of Dirac delta functions. Although a measure theoretic approach may reduce accessibility, there are important reasons for adopting it. It would be odd to attempt to prove the reality of the wavefunction within a framework that does not admit a model in which the wavefunction is real in the first place. Since the wavefunction involves continuous parameters, this means that there is no option of specializing to finite sample spaces. Furthermore, there are several special cases of interest for which the physicist's optimistic approach simply does not work, including the case where the wavefunction, and only the wavefunction, is the state of reality. Therefore, in order to cover all the cases of interest, there is really no option other than taking a measure theoretic approach. As an aid to accessibility, I outline how the main definitions and arguments specialize to the case of a finite sample space, which should be sufficient for those who do not wish to get embroiled in the technical details.

Secondly, it is common in the literature to assume that we are interested in modelling all pure states and all projective measurements on some finite dimensional Hilbert space, and to specialize results to that context. However, some results apply equally to subsets of states and measurements, which I call *fragments* of quantum theory. In addition, it is known that some fragments of quantum theory, have natural models in which the wavefunction is epistemic [62–64]. Therefore, I think it is important to emphasize the minimal structures in which the various results can be proved, rather than just assuming that we are trying to model all states and measurements on some Hilbert space.

The third presentation decision concerns my treatment of *preparation contextuality* (see §5.4 for the formal definition). The main issue we are interested in is whether pure quantum states must be ontic, since it is uncontroversial that mixed states can at least sometimes represent lack of knowledge about which of a set pure state was prepared. It is common in the literature to assume that each pure quantum state is represented by a unique probability measure over the possible states of reality, but I do not make this assumption. In a preparation contextual model, different methods of preparing a quantum state may lead to different probability measures. In fact, this must occur for mixed states [65], so it seems sensible to allow for the possibility that it might occur for pure states as well. In addition, some of the intermediate results to be discussed hold equally well for mixed states, but this can only be established by adopting definitions that are broad enough to encompass mixed states, which are necessarily preparation contextual.

These three presentation decisions mean that the definitions, statements of results, and proofs that appear in this review often differ from those in the existing literature. Generally, this is just a matter of making a

few obvious generalizations, without substantively changing the ideas. For this reason, I do not explicitly point out where such generalizations have been made.

The review is divided into three parts. Part I is a general review of the distinction between ontic and epistemic interpretations of the quantum state. It discusses the arguments that had been given for ontic and epistemic interpretations of the wavefunction prior to the discovery of the PBR Theorem. My aim in this part is to convince you that there is some merit to the epistemic interpretation and that previous arguments for the reality of the quantum state are unconvincing. In this part, I also give a formal definition of the class of models to which the PBR Theorem and related results apply, and define what it means for the quantum state to be ontic or epistemic within this class of models. Following this, I give a detailed discussion of the other no-go theorems that would follow as corollaries of proving the reality of the wavefunction.

Part II reviews the three theorems that attempt to prove the reality of the wavefunction: the PBR Theorem, Hardy's Theorem, and the Colbeck-Renner Theorem. The treatment of the PBR Theorem is the most detailed of the three, since it has attracted the largest literature and has been subject to the largest amount of confusion and criticism. In my view, it makes the strongest case of the three theorems for the reality of the wavefunction, although it is still not bulletproof, so I go to some lengths to sort the silly criticisms from the substantive ones. The assumptions behind the Hardy and Colbeck-Renner Theorems receive a more critical treatment, but the theorems are still presented in detail because they are interestingly related to other arguments about realist interpretations of quantum theory.

Part III deals with attempts to go beyond the rigid distinction between epistemic and ontic interpretations of the wavefunction by positing stronger constraints on epistemic interpretations. One of the aims of doing this is to remove the problematic auxiliary assumptions needed to prove the three main theorems, whilst still arriving at a conclusion that is morally similar. This part is shorter than the other two and mostly just summarizes the known results without proof. The reason for this is that many of the results are only preliminary and will likely be superceded by the time this review is published. My main aim in this part is to point out the most promising directions for future research.

Part I
The ψ-ontic/epistemic distinction

The results reviewed in this paper aim to show that the quantum state must be *ontic* (a state of reality) rather than *epistemic* (a state of knowledge). What does this mean and why is it important? The word "ontology" derives from the Greek word for "being" and refers to the branch of metaphysics that concerns the character of things that exist. In the present context, an *ontic state* refers to something that objectively exists in the world, independently of any observer or agent. In other words, ontic states are the things that would still exist if all intelligent beings were suddenly wiped out from the universe. On the other hand, "epistemology" is the branch of philosophy that studies of the nature and scope of knowledge. An *epistemic state* is therefore a description of what an observer currently knows about a physical system. It is something that exists in the mind of the observer rather than in the external physical world.

In classical mechanics, the distinction between ontic and epistemic states is fairly clear. A single Newtonian particle in one dimension has a position x and a momentum p and these are objective properties of the particle that exist independently of us. All other objective properties of the particle are functions of x and p. The ontic state of the particle is therefore the phase space point (x, p). This evolves according to Hamilton's equations

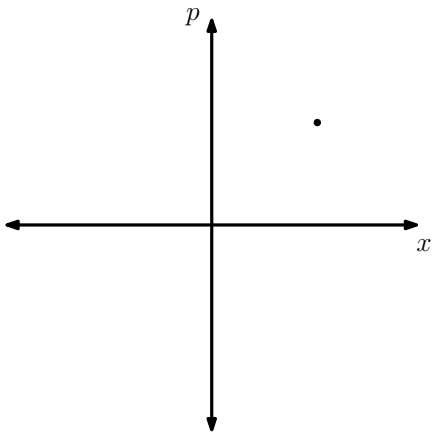
$$\frac{dp}{dt} = -\frac{\partial H}{\partial x} \qquad\qquad \frac{dx}{dt} = \frac{\partial H}{\partial q}, \tag{1}$$

where H is the Hamiltonian. On the other hand, if we do not know the exact position and momentum of the particle then our knowledge about its ontic state is represented by a probability density $f(x, p)$ over phase space. By applying Hamilton's equations to the individual phase space points on which $f(x, p)$ is supported,

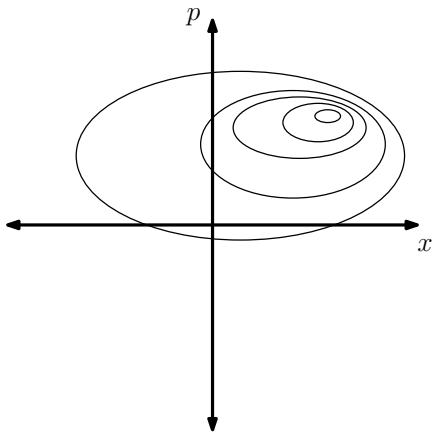
it can be shown that $f(x,p)$ evolves according to Liouville's equation

$$\frac{\partial f}{\partial t} = \frac{\partial H}{\partial x}\frac{\partial f}{\partial p} - \frac{\partial H}{\partial p}\frac{\partial f}{\partial x}. \tag{2}$$

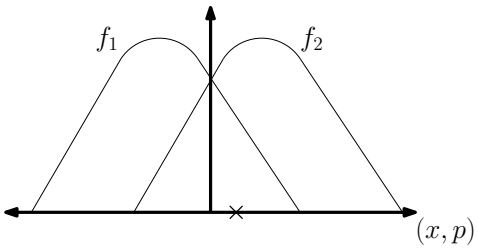
The probability density $f(x,p)$ is our epistemic state. See Fig. 1 for an illustration of the distinction between classical ontic and epistemic states. For other types of classical system the situation is analogous, the only difference being the dimension of the phase space, e.g. $6N$ dimensions for N particles in 3 dimensional space or a continuum for field systems. The phase space point is still the ontic state and a probability density over phase space is the epistemic state.



(a) An ontic state is a point in phase space.



(b) An epistemic state is a probability density on phase space. Contours indicate lines of equal probability density.



(c) An ontic state (cross) is deemed possible in more than one epistemic state (f_1 and f_2). Phase space has been schematically collapsed down to one dimension for illustrative purposes

Figure 1: The distinction between ontic and epistemic states in single particle classical mechanics.

Note that calling a probability density "epistemic" is controversial in some circles. It presupposes a broadly Bayesian interpretation of probability theory in which probabilities represent an agent's knowledge, information, or beliefs. Fortunately, the issue at stake does not really depend on this as it also appears in other interpretations of probability under a different name. What is important is that the states dubbed "epistemic" only have probabilistic import so they cannot be regarded as intrinsic properties of individual physical systems. The key property that this implies is that a given ontic state is deemed possible in more than one epistemic state.

On the Bayesian reading, this is due to the fact that different agents may have different knowledge about one and the same physical system. For example, perhaps Alice knows the position of a classical particle exactly but nothing about its momentum, whilst Bob knows the momentum precisely but nothing about its

position. Alice and Bob would then assign different probability distributions to the system, with the crucial property that they would overlap on the ontic state actually occupied by the system.

Other interpretations of probability exhibit the same property in a different way. For example, on a frequentist account of probability, probabilities represent the relative frequencies of occurrence of some property in an ensemble of independently and identically prepared systems. In this context, we would talk about a state being "statistical" rather than "epistemic". The statistical state of a system depends upon the choice of ensemble that the individual system is regarded as being a part of. For example, suppose a classical particle occupies the phase space point $(1m, 1kgms^{-1})$. If we regard it as part of an ensemble of particles that all have positive position, but some have negative momentum, then it will be assigned a different probability distribution than if we regard it as part of an ensemble of particles all of which have positive momentum, but some have negative position. In the former case, the probability distribution will have support on negative momentum phase space points and in the latter case it will have support on negative position phase space points. The point is that ensembles consist of more than one individual system and the same ontic state may occur as a part of many different ensembles. A frequentist will not be lead astray by substituting the word "statistical" for every occurrence of the word "epistemic" in this article, but the latter terminology is used here because it has become standard.

Interpretations of probability that involve single-case objective chances present more of a challenge for the ontic/epistemic distinction, since they imply that probabilities can at least sometimes be ontic. Nevertheless, I believe that an appropriate distinction can still be made in most of these theories. This discussion is deferred to Appendix §A since it is mainly of interest to those concerned with the philosophy of probability. However, it is worth mentioning that many of those who have felt the need to introduce objective chances have been motivated in part by the role that probability plays in physics, and in quantum theory in particular. Since quantum probabilities are functions of the wavefunction, they only present a novel issue for the interpretation of probability if the wavefunction itself is ontic because only then would quantum probabilities need to have a more objective status than they do in classical physics. Since the status of the wavefunction is precisely the question at issue, it is perhaps wise to defer judgement on the necessity of objective chances until the reality of the wavefunction is decided.

What is at stake then is the following question: When a quantum state $|\psi\rangle$ is assigned to a physical system, does this mean that there is some independently existing property of the individual system that is in one-to-one correspondence with $|\psi\rangle$ (up to a global phase), or is $|\psi\rangle$ simply a mathematical tool for determining probabilities, existing only in the minds and calculations of quantum theorists? This is perhaps the most hotly debated issue in all of quantum foundations. I refer to it as *the ψ-ontic/epistemic distinction* and use the terms ψ-ontic/ψ-epistemic to describe interpretations that adopt an ontic/epistemic view of the quantum state. Holders of the ψ-ontic view have been dubbed ψ-ontologists by Christopher Granade (a masters student in Rob Spekkens' quantum foundations course at Perimeter Institute in 2010) and, continuing in this vein, I refer to the reality of the quantum state as ψ-ontology and to theorems that attempt to establish this view as ψ-ontology theorems.

To avoid misunderstanding, note that the ontic/epistemic question should not be interpreted as asking whether quantum states are ontic independently of whether quantum theory is exactly true. It is not about whether the ultimate final theory of physics, if indeed such a thing exists, will feature quantum states as part of its ontology. We have little idea of what such a final theory might look like and consequently we have little idea of what reality is actually made of at the most fundamental level. Nevertheless, we can still ask what quantum theory itself *says* about reality. In other words, we are imagining a hypothetical world in which quantum theory is in fact a completely correct theory of physics, and asking whether quantum states would have to be ontological in that world. That world is very unlikely to be our actual world, so the question is really about the internal structure of quantum theory. More specifically, it is about what kinds of explanation are compatible with quantum theory. For example, a ψ-ontic view implies that we should draw analogies between quantum states and phase space points when comparing quantum and classical physics, and between the Schrödinger equation and Hamilton's equations, whereas a ψ-epistemic view says that the appropriate analogies are between quantum states and probability distributions, and between the Schrödinger equation and Liouville's equation. If nothing else, this strongly impacts how we are to understand the classical limit of

quantum theory. So, whilst the ontic/epistemic question may at first sight seem abstract and philosophical, it does in fact have concrete implications for physics.

The remainder of this part is structured as follows. §2 discusses arguments in favour of the ψ-epistemic view, with the aim of convincing you that ψ-ontology theorems are telling us something deep and surprising. For those that remain unconvinced, §3 reviews the main arguments for the reality of quantum states that were given prior to the discovery of ψ-ontology theorems. In my view, none of these are particularly compelling, so even someone who is already convinced of the reality of quantum states needs something like a ψ-ontology theorem if they aspire to defend their position with the same sort of conceptual force with which Bell derived nonlocality. Following this, §4 introduces the framework of ontological models, in which ψ-ontology theorems are proven, and gives the rigorous definition of the ψ-ontic/ψ-epistemic distinction. Finally, §5 discusses the implications of proving ψ-ontology, by showing that several existing no-go theorems can be derived from it.

2 Arguments for a ψ-epistemic interpretation

Before getting into the details of ψ-epistemic explanations, it is important to distinguish two kinds of ψ-epistemic interpretation. The most popular type are those variously described as anti-realist, instrumentalist, or positivist. Since these labels are often intended as terms of abuse, I prefer to call these approaches *neo-Copenhagen* in order to avoid implications for the philosophy of science that go way beyond how we choose to understand quantum theory. All such interpretations bear a family resemblance to the Copenhagen interpretation in that they are both ψ-epistemic and they deny the need for any deeper description of reality beyond quantum theory. Here, by "Copenhagen" I mean the views of Bohr, Heisenberg, Pauli et. al. (see e.g. [5]), which are clearly ψ-epistemic, rather than the view often found under this name in textbooks, which is actually due to Dirac [66] and von Neumann [67] and is more ambiguous about whether the wavefunction is real. If asked what quantum states represent knowledge about, neo-Copenhagenists are likely to answer that they represent knowledge about the outcomes of future measurements, rather than knowledge of some underlying observer-independent reality. Modern neo-Copenhagen views include the Quantum Bayesianism of Caves, Fuchs and Schack [16–18], the views of of Bub and Pitowsky [19], the quantum pragmatism of Healy [21], the relational quantum mechanics of Rovelli [22], the empiricist interpretation of W. M. de Muynck [23], as well as the views of David Mermin [24], Asher Peres [25], and Brukner and Zeilinger [26]. Some may quibble about whether all these interpretations resemble Copenhagen enough to be called neo-Copenhagen, but for present purposes all that matters is that these authors do not view the quantum state as an intrinsic property of an individual system and they do not believe that a deeper reality is required to make sense of quantum theory.

The second type of ψ-epistemic interpretation are those that are realist, in the sense that they do posit some underlying ontology. They just deny that the wavefunction is part of that ontology. Instead, the wavefunction is to be understood as representing our knowledge of the underlying reality, in the same way that a probability distribution on phase space represents our knowledge of the true phase space point occupied by a classical particle. There is evidence that Einstein's view was of this type [68]. Ballentine's statistical interpretation [69] is also compatible with this view in that he leaves open the possibility that hidden variables exist and only insists that, if they do exist, the wavefunction remains statistical (as a frequentist, Ballentine uses the term "statistical" rather than "epistemic"). More recently, Spekkens has been a strong advocate of this point of view [62].

Neo-Copenhagen and realist ψ-epistemic interpretations share much of the same explanatory structure, since they both view probability measures as the correct classical analogy for the wavefunction. Many of the arguments for adopting a ψ-epistemic interpretation apply equally to both of them. On the other hand, ψ-ontology theorems only apply to realist interpretations. This is to be expected as it would be difficult to prove that the wavefunction must be ontic in a framework that does not admit the existence of ontic states in the first place. Because of this, ψ-epistemicists always have the option of becoming neo-Copenhagen in the face of ψ-ontology theorems.

Realist ψ-epistemic interpretations are already strongly constrained by existing no-go theorems, such as Bell's Theorem [1] and the Kochen-Specker Theorem [70], which go some way to explaining why not

many concrete ψ-epistemic models have been proposed. However, there is no reason to view these results as decisive against realist ψ-epistemic interpretations any more than they are decisive against realist ψ-ontic interpretations. For example, Bohmian mechanics and spontaneous collapse theories still attract considerable support despite displaying nonlocality and contextuality, as the existing no-go theorems imply they must. Thus, we would be guilty of a double-standard if we ruled out realist ψ-epistemic interpretations on the basis of these results but still admitted the possibility of ψ-ontic ones. What is needed is a theorem that explicitly addresses the ψ-ontic/epistemic distinction, and this is the gap that ψ-ontology theorems are intended to fill.

In the remainder of this section, the main arguments in favour of ψ-epistemic interpretations are reviewed. Because we do not have a fully worked out realist ψ-epistemic model that covers the whole of quantum theory, it is helpful to introduce toy models that are similar to quantum theory in some respects, but in which the analogous notion to the quantum state is clearly epistemic. These are intended to demonstrate the kinds of explanation that are possible in ψ-epistemic theories. Spekkens' toy theory [62], which reinvigorated interest in realist ψ-epistemic models in recent years, is reviewed in §2.1. There are also ψ-epistemic models that cover fragments of quantum theory, e.g. just pure state preparations and projective measurements of a single qubit or just continuous variable systems when restricted to Gaussian states and operations. These are reviewed in §2.2. Finally, I review three further arguments for the ψ-epistemic view based on the fact that quantum theory can be viewed as a generalization of classical probability theory in §2.3, on the collapse of the wavefunction in §2.4, and on the size of the quantum state space in §2.5.

2.1 Spekkens toy bit

Spekkens introduced a toy theory [62] that qualitatively reproduces the physics of spin-1/2 particles (or any other instantiation of qubits) when they are prepared and measured in the x, y and z bases. The full version of Spekkens theory incorporates dynamics and composite systems, including reproducing some of the phenomena associated with entangled states but, for illustrative purposes, we restrict attention to the simplest case of a single *toy bit*. The toy theory is meant to demonstrate the explanatory power of ψ-epistemic interpretations by providing natural explanations of many quantum phenomena that are puzzling if the quantum state is ontic.

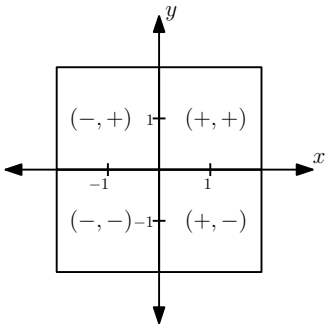


Figure 2: The ontic state space of Spekkens toy-bit. The ontic states are labelled (x, y) where x and y take values ± 1, which are abbreviated to $\pm$ for compactness.

A toy bit consists of a system that can be in one of four states, labelled $(-,-), (-,+), (+,-)$ and $(+,+)$. In Fig. 2, these are depicted laid out as a grid in the $x-y$ plane, with the origin lying at the center of the grid. The ontic states can then be thought of as representing the coordinates of the centers of the grid cells, given by (x, y), with $\pm$ short for ± 1. For concreteness, one can imagine that the cells of the grid represent four boxes and that the system is a ball that can be in one of them. The ontic state (x, y) then represents the state of affairs in which the ball is in the box centered on the coordinates (x, y).

The most fine grained description of the toy bit is always its ontic state, but we might not know exactly

which of the ontic states is occupied. In general, our knowledge of the system is described by a probability distribution over the four ontic states, and this probability distribution is our epistemic state. Spekkens imagines that there is a restriction on the set of epistemic states that may be assigned to the system, called the *knowledge-balance principle*, which is in some ways analogous to the uncertainty principle. Roughly speaking, the knowledge balance principle states that at most half of the information needed to specify the ontic state can be known at any given time. This means, for example, that if we know the x-coordinate with certainty then we cannot know anything about the y-coordinate. Given this restriction, there are six possible states of maximal knowledge, termed *pure states*, as shown in the left hand side of Fig. 3. The pure states are denoted $|x\pm)\,, |y\pm)\,, |z\pm)$ in analogy to the quantum states $|x\pm\rangle\,, |y\pm\rangle\,, |z\pm\rangle$ of a spin-1/2 particle.

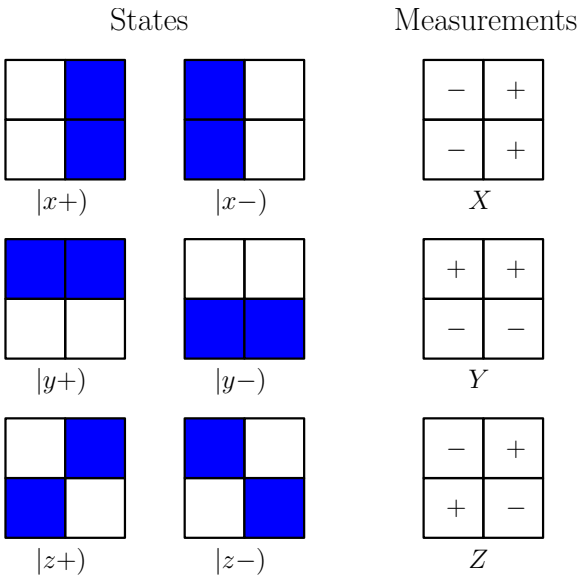


Figure 3: The allowed states and measurements of Spekkens' toy bit. For the states, a blue square indicates that the corresponding ontic state has probability 1/2 and a white square indicates probability 0. For the measurements, a square labelled + gives the +1 outcome with certainty and a square labelled − gives the −1 outcome with certainty.

Note that the epistemic states $|z\pm)$ are not states with a definite value of the z-coordinate. The system is two dimensional so it does not have a third coordinate. Instead, $|z+)$ is the state in which we know only that the x and y coordinates are equal and $|z-)$ is the state in which we know only that they are different. Defining $z = xy$, this is equivalent to saying that $|z\pm)$ is the state in which we know only that $z = \pm 1$.

Although the knowledge balance principle has been imposed by hand, it is easy to imagine that it could arise from a lack of fine-grained control over the system. For example, imagine a preparation device that pushes the ball to the left along the x-axis, but that same device also causes a random disturbance to the y-coordinate, such that the best we can do after operating the device is to assign the state $|x-)$.

Having described the epistemic states of the theory, the next task is to describe the measurements. Spekkens requires that measurements be *repeatable*, which means that if a measurement is repeated twice in succession then it should yield the same outcome both times. Also, the measurement should respect the knowledge balance principle, so that our epistemic state after the measurement contains at most half of the information required to specify the ontic state. In order to satisfy this second requirement, the measurement must necessarily cause a disturbance to the ontic state, since otherwise we could end up in a situation in which we know the ontic state exactly. For example, if a measurement of the x-coordinate could be implemented without disturbance then measuring the x-coordinate followed by measuring the y-coordinate would tell us the exact ontic state of the system.

There are three nontrivial measurements that can be implemented in such a way that they satisfy the two requirements: the X measurement reveals the x coordinate, the Y measurement reveals the y coordinate, and the Z measurement reveals the value of $z = xy$. These are illustrated on the right hand side of Fig. 3. Each of these measurements causes a random exchange between the pairs of ontic states that give the same outcome in the measurement. For example, if we perform an X measurement and get the $+1$ outcome, then with probability $1/2$ nothing happens and with probability $1/2$ the states $(+,-)$ and $(+,+)$ are exchanged. This ensures that we always end up in an epistemic state that satisfies the knowledge-balance principle at the end of the measurement, in this case $|x+)$. It is easy to see that this is the only type of disturbance that is compatible with both repeatability and the knowledge-balance principle. For example, for an X-measurement the random disturbance cannot exchange ontic states that have different values of the x-coordinate, e.g. $(+,+)$ and $(-,+)$, since this would violate repeatability.

The theory described so far makes exactly the same predictions as quantum theory for sequences of measurements in the x, y and z directions of spin-1/2 particles prepared in one of the states $|x\pm\rangle$, $|y\pm\rangle$ and $|z\pm\rangle$ if we identify these six states with $|x\pm)$, $|y\pm)$ and $|z\pm)$ and the Pauli observables σ_x, σ_y and σ_z with X, Y and Z. It can thus be regarded as a hidden variable theory for this kind of experiment. Further, the quantum states are epistemic in this representation, as they are each represented by probability distributions that have support on two ontic states and nonorthogonal states overlap, e.g. $|x+)$ and $|y+)$ both assign probability $1/2$ to the ontic state $(+,+)$.

Several features of quantum theory that are puzzling on the ψ-ontic view are present in this theory and have very natural explanations. Firstly, consider the fact that nonorthogonal pure states cannot be perfectly distinguished by a measurement, e.g. if either the state $|x+\rangle$ or the state $|y+\rangle$ is prepared, and you do not know which, then there is no measurement that will enable you to deduce this information with certainty. If quantum states are ontic then the two preparations correspond to distinct states of reality and it is puzzling that we cannot detect this difference. On the other hand, the toy theory states $|x+)$ and $|y+)$ overlap on the ontic state $(+,+)$ and this will be occupied by the system 50% of the time whenever $|x+)$ or $|y+)$ is prepared. When this does happen, there is nothing about the ontic state of the system that could possibly tell you whether $|x+)$ or $|y+)$ was prepared. Therefore, we must fail to distinguish the two preparations at least 50% of the time. The overlap of the two epistemic states accounts for their indistinguishability.

Another feature of quantum theory that is easily accounted for in Spekkens' model is the no-cloning theorem. In quantum theory, there is no transformation that copies both of two nonorthogonal states. For example, there is no device that operates with certainty and outputs both $|x+\rangle \otimes |x+\rangle$ when $|x+\rangle$ is input and $|y+\rangle \otimes |y+\rangle$ when $|y+\rangle$ is input. On the ψ-ontic view this is puzzling because the two states represent distinct states of reality, so one might expect that this distinctness could be detected and then copied over to another system. Again, this is easily explained in Spekkens' model in terms of the overlap between the epistemic states $|x+)$ and $|y+)$. Fig. 4 shows the inputs and outputs of the hypothetical toy-theory cloning machine. The two input states overlap on the ontic state $(+,+)$ and this occurs 50% of the time regardless of which input state is prepared. Since the cloning machine only has access to the ontic state, it must do the same thing to the state $(+,+)$, regardless of whether it occurs because $|x+)$ was prepared or because $|y+)$ was prepared. Therefore, 50% of the time, the input must get mapped to the same set of ontic states, with the same probabilities, regardless of which state was prepared, so there must be at least a 50% overlap of the output states of any physically possible device for these two input states. In contrast, the output states of the hypothetical cloning machine only overlap on the ontic state $((+,+),(+,+))$ and this must only occur 25% of the time at the output for either input state. Therefore, the cloning machine is impossible.

In Spekkens' toy theory, both indistinguishability and no-cloning follow from the more general fact that a stochastic map cannot decrease the overlap of two probability distributions. In quantum theory, there is a similar result that no transformation that can be implemented with certainty can decrease the inner product between two pure states [71]. This suggests that the inner product of two quantum states is analogous to the overlap between two probability distributions, and this analogy would be most easily explained if quantum states with nonzero inner product were literally represented by overlapping probability distributions on some ontic state space, i.e. by a ψ-epistemic interpretation.

Finally, consider the fact that mixed states in quantum theory have more than one decomposition into a

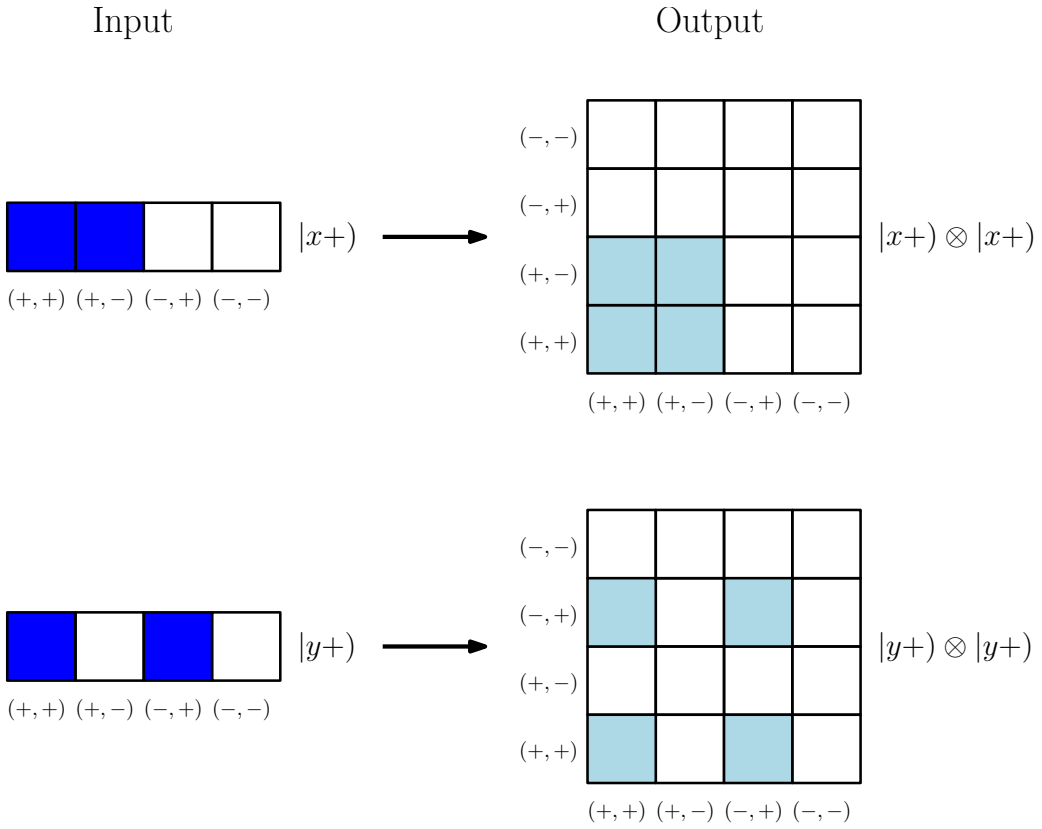


Figure 4: Inputs and outputs of a hypothetical toy bit cloning machine. In order to represent a two toy bit state, the ontic state space of a single toy bit is represented along one dimension. For the outputs, the horizontal axis represents the first toy bit and the vertical axis represents the second toy bit. Dark blue represents ontic states occupied with 50% probability and light blue represents those occupied with 25% probability. The inputs overlap on an ontic state that they both assign 50% probability, but the outputs only overlap on an ontic state that they both assign 25% probability.

convex sum of pure states. For example, the maximally mixed state of a spin-1/2 particle is $I/2$, where I is the identity operator, and this can be written alternatively as

$$\frac{I}{2} = \frac{1}{2}\left|x+\right\rangle\left\langle x+\right| + \frac{1}{2}\left|x-\right\rangle\left\langle x-\right| \tag{3}$$

$$= \frac{1}{2}\left|y+\right\rangle\left\langle y+\right| + \frac{1}{2}\left|y-\right\rangle\left\langle y-\right| \tag{4}$$

$$= \frac{1}{2}\left|z+\right\rangle\left\langle z+\right| + \frac{1}{2}\left|z-\right\rangle\left\langle z-\right|. \tag{5}$$

Physically speaking, this means that if we prepare a spin-1/2 particle in the $|x+\rangle$ state with probability 1/2 and in the $|x+\rangle$ state with probability 1/2 then no experiment can tell the difference between this ensemble and that formed by preparing it in the $|y+\rangle$ state with probability 1/2 and in the $|y-\rangle$ state with probability 1/2 (and similarly for $|z\pm\rangle$). On a ψ-ontic view this is puzzling because the $|x\pm\rangle$ states are ontologically distinct from the $|y\pm\rangle$ (and $|z\pm\rangle$) states so this difference should be detectable. However, in Spekkens' theory this nonuniqueness of decomposition is easily explained because preparing $|x+)$ with probability 1/2 and $|x-)$ with probability 1/2 leads to exactly the same distribution over ontic states as preparing $|y+)$ with probability 1/2 and $|y-)$ with probability 1/2 (and similarly for $|z\pm)$). This is illustrated in Fig. 5. Note that this is only possible because the distributions corresponding to nonorthogonal quantum states overlap.

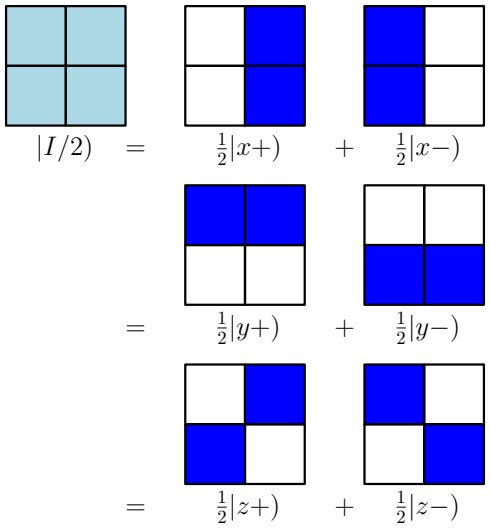


Figure 5: Multiple pure-state decompositions of a mixed state in Spekkens toy theory. The maximally mixed state $|I/2)$ can be written as a 50/50 mixture in three different ways. Light blue indicates a probability of 1/4.

What has been presented in this section is just a small fraction of the quantum phenomena that are accounted for in Spekkens' toy model. Many more can be found in [62], but I hope the present discussion has conveyed a flavour of the type of explanation that is possible in a realist ψ-epistemic theory.

2.2 Models for fragments of quantum theory

Spekkens' toy model is qualitatively similar to the stabilizer fragment of quantum theory, which consists of the set of states that are joint eigenstates of maximal commutative subgroups of the Pauli group (i.e. the group generated by tensor products of the identity and the three Pauli operators) and has dynamics given by unitaries that map the Pauli group to itself (see [72] for details of the stabilizer formalism and [73] for a presentation of Spekkens' toy theory that closely resembles it). The stabilizer fragment is important in quantum information theory as it contains all the states and operations needed for quantum error correction, as well as a number of other quantum protocols. Spekkens' model does not exactly reproduce the stabilizer fragment when dynamics and entanglement are taken into account, but other models have been proposed that do reproduce fragments of quantum theory exactly in a ψ-epistemic manner.

First of all, Spekkens' toy theory has been generalized to larger dimensions [63] and to continuous variable systems [64]. It turns out that for odd dimensional Hilbert spaces, Spekkens' model reproduces the stabilizer fragment of quantum theory exactly. For continuous variable systems, Spekkens' model reproduces the Gaussian fragment of quantum theory, in which all states are Gaussian and the transformations and measurements preserve the Gaussian nature of the states. A Gaussian state is one that has a Gaussian Wigner function. For a single particle, the Wigner function is defined in terms of the density operator ρ as $W(x,p) = \int_{-\infty}^{+\infty} ds e^{ips/\hbar} \left\langle x - \frac{s}{2} \right| \rho \left| x + \frac{s}{2} \right\rangle$ and is a pseudo-probability distribution, i.e. it is normalized to 1 but it does not have to be positive. Gaussian functions are in fact positive so in this case $W(x,p)$ can be regarded as a probability distribution and unsurprisingly these are the epistemic states in Spekkens' continuous variable theory, with the ontic states being the phase space points (x,p).

Kochen and Specker gave a model for a single qubit that is ψ-epistemic [70]. They were not actually trying to generate a ψ-epistemic model, but rather to provide a counterexample to their eponymous theorem in 2-dimensions in order to show that the theorem requires a Hilbert space of ≥ 3 dimensions for its proof. Nevertheless, their model is a paradigmatic example of a ψ-epistemic theory. The details of this model are

presented in §4.3 after we have introduced the formalism for realist ψ-epistemic models more rigorously. Along similar lines, Lewis et. al. [52] and Aaronson et. al. [53] have constructed ψ-epistemic models that work for all finite dimensional systems. These models were developed as technical counterexamples to certain conjectures about ψ-ontology theorems and as such they are not very elegant or plausible. They are discussed in context in §7.5.

2.3 Generalized probability theory

Apart from specific models, there are also qualitative arguments in favour of the ψ-epistemic view. The first of these is that quantum theory can be viewed as a noncommutative generalization of classical probability theory. Classically, consider the algebra $\mathcal{A}$ of random variables on a sample space under pointwise addition and multiplication. A probability distribution can then be regarded as a positive functional $\mu : \mathcal{A} \to \mathbb{R}$ that assigns to each random variable its expectation value. The quantum generalization of this is to replace the commutative algebra $\mathcal{A}$ by the noncommutative algebra $\mathfrak{B}(\mathcal{H})$ of bounded operators on a Hilbert space $\mathcal{H}$. A quantum state ρ is isomorphic to a positive functional f_ρ on $\mathfrak{B}(\mathcal{H})$ given by $f_\rho(M) = \mathrm{Tr}(M\rho)$. In fact, by a theorem of von Neumann [67], all positive linear functionals on $\mathfrak{B}(\mathcal{H})$ that are normalized such that $f(I) = 1$, where I is the identity operator, are of this form provided the dimension of the Hilbert space is ≥ 3.

Both $\mathcal{A}$ and $\mathfrak{B}(\mathcal{H})$ are examples of von Neumann algebras, and a generalization of classical measure theoretic probability can be developed by defining generalized probability distributions to be positive normalized functionals on such algebras [74, 75]. This generalized theory has both classical probability theory and quantum theory as special cases. In this theory, quantum states are playing the same role in the quantum case that probability measures play in the classical case, and so it is natural to interpret quantum states and classical probabilities as the same kind of entity. Since classical probabilities are usually interpreted epistemically, it is natural to interpret quantum states in the same way.

This line of argument would not be too convincing if noncommutative probability theory were just a formal mathematical generalization with no practical applications. However, the theory has a rich array of applications in quantum statistical mechanics, and especially in quantum information theory. The full machinery of von Neumann algebras is not often needed in quantum information, as we are usually dealing with finite dimensional systems. Nevertheless, whenever the analogy is made between classical probability distributions and density operators, and between stochastic maps and quantum operations, generalized probability is at play in the background. For example, in quantum compression theory [76], a density operator on a finite Hilbert space is viewed as the correct generalization of a classical information source with finite alphabet, which would be described by a classical probability distribution. Similarly, a quantum channel is described by a quantum operation, and this is viewed as generalizing a classical channel, which would be modelled as a stochastic map.

In fact, it is difficult to find any area of quantum information and computing in which probabilities are not viewed as the correct classical analogs of quantum states, and this includes areas that concern themselves exclusively with pure states and unitary transformations. For example, the standard circuit model of quantum computing [77] only employs pure states and unitaries, but quantum computational complexity classes are most often defined as generalizations of classical probabilistic complexity classes (see [78] for definitions of the complexity classes mentioned in this section.). The class BQP, usually thought of as the set of problems that can be solved efficiently on a quantum computer, is sometimes loosely described as the quantum version of P, the class of problems that can be solved in polynomial time on a deterministic classical computer, but in fact it is the generalization of BPP, the set of problems that can be solved in polynomial time on a probabilistic classical computer with probability $> 2/3$. All over quantum computing theory we find the analogy made to classical probabilistic computing, and not to classical deterministic computing.

It seems then, that if we take quantum information and computing seriously, then we must take generalized probability theory seriously as well. On these and other grounds, I have argued elsewhere [79, 80] that quantum theory is indeed best viewed as a generalization of probability theory. The details of this would take us too far afield, but suffice to say there are good reasons for viewing quantum states as analogous to probability distributions and, if we do that, we should try to interpret them both in the same sort of way.

2.4 The collapse of the wavefunction

A straightforward resolution of the collapse of the wavefunction, the measurement problem, Schrödinger's cat and friends is one of the main advantages of ψ-epistemic interpretations. Recall that the measurement problem stems from the fact that there are two different ways of propagating a quantum state forward in time. When the system is isolated and not being observed, the quantum state is evolved smoothly and continuously according to the Schrödinger equation. On the other hand, when a measurement is made on the system, the quantum state must be updated according to the projection postulate, leading to the instantaneous and discontinuous collapse of the wavefunction. Since a measurement is presumably just some type of physical interaction between system and apparatus, this poses the problem of why it is not also modelled by Schrödinger evolution. However, doing so leads to seemingly absurd situations, such as Schrödinger's eponymous cat ending up in a superposition of being alive and dead at the same time.

The measurement problem is not so much resolved by ψ-epistemic interpretations as it is dissolved by them. It is revealed as a pseudo-problem that we were wrong to have placed so much emphasis on in the first place. This is because the measurement problem is only well-posed if we have already established that the quantum state is ontic, i.e. that it is a direct representation of reality. Only then does a superposition of dead and alive cats necessarily represent a distinct physical state of affairs from a definitely alive or definitely dead cat. On the other hand, if the quantum state only represents what we know about reality then the cat may perfectly well be definitely dead or alive before we look, and the fact that we describe it by a superposition may simply reflect the fact that we do not know which possibility has occurred yet.

2.5 Excess baggage

According to the ψ-ontologist, a single qubit contains an infinite amount of information because a pure state of a qubit is specified by two continuous complex parameters (ignoring normalization). For example, Alice could encode an arbitrarily long bit string as the decimal expansion of the amplitude of the $|0\rangle$ state. However, according to the Holevo bound [81], only a single bit of classical information can be encoded in a qubit in such a way that it can be reliably retrieved. If the quantum state truly exists in reality, it is puzzling that we cannot detect all of this extra information. Hardy has coined the term "ontological excess baggage" to refer to this phenomenon [59]. It seems that ψ-ontologists are attributing a lot more information to the state of reality than required to explain our observations.

The ψ-epistemic response to this is to note that a classical probability distribution is also specified by continuous parameters. A probability distribution over a single classical bit requires two real parameters (again ignoring normalization). If probabilities were intrinsic properties of individual systems then this would present a similar puzzle as there would be an infinite amount of information in a single bit. However, classical bits are in fact always either in the state zero or one and the probabilities simply represent our knowledge about that value. In reality, there is just as much information in a classical bit as we can extract from it, namely one bit. If the quantum state is epistemic, then the same resolution is available to the problem of excess baggage. The continuous parameters required to specify the state of a qubit simply represent our knowledge about it, and the actual ontic state of the qubit, whatever that may be, might only contain a finite amount of information.

The excess baggage problem is exacerbated by considering how the state space scales with the number of qubits. A pure state of n qubits is specified by 2^n complex parameters, but only n bits can be reliably encoded according to the Holevo bound. However, the number of parameters required to specify a probability distribution over n bits also scales exponentially, so the ψ-epistemic resolution of the problem is still available.

In response to this, ψ-ontologists might be inclined to point out that the number of bits that can be reliably encoded in n qubits depends on how exactly the communication task is defined. If Alice and Bob have pre-shared entanglement then Alice can send $2n$ bits to Bob in n qubits via superdense coding [82]. Similarly, qubits perform better than classical bits in random access coding [83], wherein Bob is not required to reliably retrieve all of the bits that Alice sends, but only a limited number of them of his choice. However, the amount of information that Alice can send to Bob does not scale exponentially with the number of qubits in any of these protocols, so there is still an excess baggage problem.

3 Arguments for a ψ-ontic interpretation

Having reviewed the arguments in favour of ψ-epistemic interpretations, we now look at those that had been put forward in favor of the reality of quantum states prior to the discovery of ψ-ontology theorems. Despite receiving a good deal of support, I hope to convince you that they are far from compelling. Thus, even those who are already convinced of the reality of the quantum state should be interested in establishing their claim rigorously via ψ-ontology theorems.

A big difficulty in extracting arguments for ψ-ontology from the literature is that the majority of authors neglect the possibility of realist ψ-epistemic theories. Instead, they seem to think that either the wavefunction must be real, or else we must adopt some kind of neo-Copenhagen approach. Thus, many purported arguments for the reality of the wavefunction are really just arguments for the reality of something, regardless of whether that thing is the wavefunction. Since realist ψ-epistemic interpretations already accept the need for an objective reality, such arguments can be dismissed in the present context. From amongst these arguments, I have attempted to sift out those that say something more substantive about the wavefunction specifically. I have found four broad classes of argument, each of which is discussed in turn in this section. §3.1 discusses the argument from interference, §3.2 discusses the argument from the eigenvalue-eigenstate link, §3.3 discusses the argument from existing realist interpretations of quantum theory, and finally §3.4 discusses the argument from quantum computation.

3.1 Interference

> We choose to examine a phenomenon which is impossible, *absolutely* impossible, to explain in any classical way, and which has in it the heart of quantum mechanics. In reality, it contains the *only* mystery. — R. P. Feynman [84] [Emphasis in original]

Following Feynman, single particle interference phenomena, such as the double slit experiment, are often viewed as containing the essential mystery of quantum theory. The problem of explaining the double slit experiment is usually presented as a dichotomy between explaining it in terms of a classical wave that spreads out and travels through both slits or in terms of a classical particle that travels along a definite trajectory that goes through only one slit. Neither of these explanations can account for both the interference pattern and the fact that it is built out of discrete localized detection events. A wave would not produce discrete detection events and a classical particle would not be affected by whether or not the other slit is open. This is taken as evidence that no classical description can work, and that something more Copenhagen-like must be at work.

Of course, the dichotomy between either classical waves or particles is a false one. If we allow the state of reality to be something more general, i.e. some sort of quantum stuff that we do not necessarily understand yet, then many additional explanations of the experiment become available. For example, there is the Bohmian picture in which both a wave and a particle exist, and the motion of the particle is guided by the wave. The wave then explains the interference fringes, whilst the particle explains the discrete detection events. This is by no means the only possibility, but it does highlight the gap in the usual argument. Nevertheless, in a realist picture, it seems that something wavelike needs to exist in order to explain the interference fringes, and the obvious candidate is the wavefunction.

However, in order to arrive at the conclusion that the wavefunction must be real, greater leeway has been given in determining what the ontic state might be like compared to the original argument, which intended to rule out both particles and waves. Given this, we should be careful to rule out other possibilities rigorously, rather than jumping to the conclusion that the wavefunction must be real. In this broader context, the only thing that the double slit experiment definitively establishes is that there must be some sort influence that travels through both slits in order to generate the interference pattern. It does not establish that this influence must be a wavefunction.

In fact, interference phenomena occur in some of the previously discussed ψ-epistemic models, so the inference from interference to the reality of the wavefunction is incorrect. In Spekkens' toy theory, a notion of coherent superposition can be introduced such that, for example, $|y+\rangle$ is a coherent superposition of $|x+\rangle$

and $|x+\rangle$). Such superpositions are preserved under dynamical evolution, so there is a superposition principle in the theory (see [62] for details). Further, since all two-dimensional Hilbert spaces are created equal, there is nothing special about the interpretation of the toy bit in terms of a spin-1/2 particle. It could equally well be a model of any other two-dimensional system. For example, consider the two dimensional subspace of an optical mode spanned by the vacuum state $|0\rangle$ and the state $|1\rangle$ where it contains one photon. The toy bit state $|x+\rangle$ can be reinterpreted as $|0\rangle$ and $|x-\rangle$ as $|1\rangle$, and by doing so a whole host of Mach-Zehnder interferometry experiments can be qualitatively reproduced by the theory [85]. This includes not only basic interferometry, but also such seemingly paradoxical effects as the delayed choice experiment [86] and the Elitzur-Vaidman bomb test [87]. In this theory, there is always a fact of the matter about which arm of the interferometer the photon travels along, but it does not fall afoul of the standard waves vs. particles argument because the vacuum state has structure. The epistemic state $|x+\rangle$ is compatible with two possible ontic states $(+,+)$ and $(+,-)$ so when a photon is in the left arm of an interferometer and no photon is in the right arm then there is still a bit of information travelling along the right arm of the interferometer that can be used to convey information about whether or not its path was blocked. There is an influence that travels through both arms, but that influence is not a wavefunction.

Interference phenomena also occur in all of the models discussed in §2.2 simply because they reproduce fragments of quantum theory exactly and those fragments contain coherent superpositions. It is arguable whether the mechanisms explaining interference in all these models are plausible, but the main point is that the direct inference from interference to the reality of the wavefunction is blocked by them. If there is an argument from interference to be made then it will need to employ further assumptions. Hardy's ψ-ontology theorem, discussed in §9, can be viewed as an attempt at doing this, but, in light of the way that interference is modelled in Spekkens' toy theory, its assumptions do not seem all that plausible.

Ultimately, the intuition behind the argument from interference stems from an analogy with classical fields. Because wavefunctions can be superposed, they exhibit interference. Prior to the discovery of quantum theory, the only entities in physics that obeyed a superposition principle and exhibited interference were classical fields, and these were definitely intended to be taken as real. For example, the value of the electromagnetic field at some point in spacetime is an objective property that can be measured by observing the motion of test charges. The interference of wavefunctions is then taken as evidence that they should be interpreted as something similar to classical fields.

However, the analogy between wavefunctions and fields is only exact for a single spinless particle, for which the wavefunction is essentially just a field on ordinary three dimensional space. This breaks down for more than one particle, due to the possibility of entanglement. The size of the quantum state space scales exponentially with the number of systems, leading to the previously discussed excess baggage problem. The wavefunction can no longer be viewed as field on ordinary three-dimensional space, so the analogy with a classical field should be viewed with skepticism. In combination with the fact that interference phenomena can be modelled ψ-epistemicly, the argument from interference is far from compelling.

3.2 The eigenvalue-eigenstate link

The eigenvalue-eigenstate link refers to the tenet of orthodox quantum theory that when a system is in an eigenstate $|m\rangle$ of an observable M with eigenvalue m then M is a property of the system that has value m. Conversely, when the state is not an eigenstate of M then M is not a property of the system. In other words, the properties of a system consist of all the observables of which the quantum state is an eigenstate and nothing else. These properties are taken to be objectively real, independently of the observer.

This leads to an argument for the reality of the wavefunction because the quantum state of a system is determined uniquely by the set of observables of which it is an eigenstate. Indeed, it is determined uniquely by just a single observable, since $|\psi\rangle$ is an eigenstate of the projector $|\psi\rangle\langle\psi|$ with eigenvalue 1 and (up to a global phase) it is the only state in the $+1$ eigenstate of $|\psi\rangle\langle\psi|$. The argument is then that, if a system has a set of definite properties, and those properties uniquely determine the wavefunction, then the wavefunction itself must be real.

Roger Penrose is perhaps the most prominent advocate of this argument, so here it is in his own words.

> One of the most powerful reasons for rejecting such a subjective viewpoint concerning the reality of $|\psi\rangle$ comes from the fact that whatever $|\psi\rangle$ might be, there is always—in principle, at least–a primitive measurement whose **YES** space consists of the Hilbert space ray determined by $|\psi\rangle$. The point is that the physical state $|\psi\rangle$ (determined by the ray of complex multiples of $|\psi\rangle$) is uniquely determined by the fact that the outcome **YES**, for this state, is certain. No other physical state has this property. For any other state, there would merely be some probability, short of certainty, that the outcome will be **YES**, and an outcome of **NO** might occur. Thus, although there is no measurement which will tell us what $|\psi\rangle$ actually is, the physical state $|\psi\rangle$ is uniquely determined by what it asserts must be the result of a measurement that might be performed on it...
>
> To put the point a little more forcefully, imagine that a quantum system has been set up in a known state, say $|\phi\rangle$, and it is computed that after a time t the state will have evolved, under the action of U, into another state $|\psi\rangle$. For example, $|\phi\rangle$ might represent the state 'spin up' ($|\phi\rangle = |\uparrow\rangle$) of an atom of spin $\frac{1}{2}$, and we can suppose that it has been put in that state by the action of some previous measurement. Let us assume that our atom has a magnetic moment aligned with its spin (i.e. it is a little magnet pointing to the spin direction). When the atom is placed in a magnetic field, the spin direction will precess in a well-defined way, that can be accurately computed as the action of U, to give some new state, say $|\psi\rangle = |\rightarrow\rangle$, after a time t. Is this computed state to be taken seriously as part of physical reality? It is hard to see how this can be denied. For $|\psi\rangle$ has to be prepared for the possibility that we might choose to measure it with the primitive measurement referred to above, namely that whose **YES** space consists precisely of the multiples of $|\psi\rangle$. Here, this is the spin measurement in the direction $\rightarrow$. The system has to know to give the answer **YES**, with certainty for that measurement, whereas no spin state of the atom other than $|\psi\rangle = |\rightarrow\rangle$ could guarantee this. — Roger Penrose, quoted in [88].

There are two possible responses to this argument. The first response is to simply deny the eigenvalue-eigenstate link. That is, even if Alice assigns the state $|\psi\rangle$ to the system, it does not follow that the system has objectively real properties corresponding to the observables of which $|\psi\rangle$ is an eigenstate. It simply means that Alice thinks she can predict the outcomes of measurements of those observables with certainty. To put the point a little more forcefully, suppose that Roger, who believes in the eigenvalue-eigenstate link argument, assigns the state $|\psi\rangle$ to the system and hence believes that $|\psi\rangle$ is an objective property of the system. Alice, however, persists in assigning a different state $|\phi\rangle$, which is non-orthogonal to $|\psi\rangle$. Is there anything that Roger can do to convince Alice that she is wrong and that $|\psi\rangle$ is the objectively correct state to assign to the system? Well, one thing he could do is to measure the observable $|\psi\rangle\langle\psi|$, so suppose he does this and obtains the $+1$ outcome (or the **YES** outcome in Penrose's terminology). However, this is not enough to convince Alice because her state, being non-orthogonal to $|\psi\rangle$, also assigns a nonzero probability to the $+1$ outcome. Roger might think that he can get around this by preparing a large number of systems using the same preparation device and then performing the same measurement on all of them. If all the outcomes turn out to be $+1$ then surely this should be enough evidence to convince Alice that the state $|\psi\rangle$ is an objective property of the system? Unfortunately, this is not the case either as all that has been established is a property of the ensemble of systems, not an intrinsic property of an individual system. In particular, Alice may disagree that the preparations are independent and identically distributed. For example, she might think that the appropriate state to assign to the first system, or to some small subset of all the systems, is $|\phi\rangle$, and that the appropriate state for the rest of the systems is $|\psi\rangle$. In this case, Alice agrees with Roger that the observed outcome has high probability, but they still completely disagree about the state of the first system. It seems that there is nothing that Roger can do to convince Alice that $|\psi\rangle$ must be an objective property of the system.

This counterargument is the most appropriate response for a neo-Copenhagenist, but the realist ψ-epistemicist has an even more compelling counterargument available. Consider the epistemic state $|x+)$ in Spekken's toy theory. This state assigns the definite value $+1$ to the X measurement and indeed it is the only allowed epistemic state in the theory that does this. In fact, all of the pure states of a toy bit $|x\pm)\,, |y\pm)$ and $|z\pm)$ are uniquely determined by the definite value that they assign to one of the measurements. Fol-

lowing Penrose's reasoning, we would then conclude that these states are objective properties of the system. However, this is not the case since the objective properties of the system are those that are determined by the ontic state, and each ontic state is compatible with more than one epistemic state. For example, the ontic state $(+,+)$ is compatible with $|x+\rangle$, $|y+\rangle$ and $|z+\rangle$. Given the complete specification of reality, the epistemic state is underdetermined.

One way of exposing the error in the eigenvalue-eigenstate argument is to note that, in the toy theory, the fact that observables uniquely determine epistemic states is a consequence of the knowledge-balance principle and not a fundamental fact about reality. For example, without the knowledge-balance principle, it would be permissible to have an epistemic state that assigns probability $2/3$ to $(+,+)$ and $1/3$ to $(+,-)$. Just like $|x+\rangle$, this state assigns probability 1 to the $+1$ outcome of the X measurement and this is the only measurement that is assigned a definite value. If this state were allowed then it would no longer be possible to mistake $|x+\rangle$ for an objective property of the system. Penrose has mistaken the set of states that it is possible to prepare with current experiments for the set of all logically possible states.

Another way of exposing the error is to look at the restrictions on measurements in the toy theory. The measurements only reveal coarse-grained information about the ontic state. Without the knowledge-balance principle it would be permissible to conceive of a more fine-grained measurement that reveals the ontic state exactly. This measurement is a definite property of the system because it is determined uniquely by the ontic state. However, specifying this observable no longer uniquely determines the epistemic state. For example, if we learn that the ontic state is $(+,+)$ then this is compatible with $|x+\rangle$, $|y+\rangle$ and $|z+\rangle$.

In conclusion, the mistake in the eigenvalue-eigenstate argument is to assume that the observables that we can actually measure in experiments form the sum total of all the properties of the system and to assume that the set of states that we can actually prepare are the sum total of all logically conceivable states. Without these assumptions, the argument is simply false.

3.3 Existing realist interpretations

There are a handful of fully worked out realist interpretations of quantum theory, including many-worlds [6–8], de Broglie-Bohm theory [9–12], spontaneous collapse theories [13, 14] and modal interpretations [15]. In each of these interpretations the wavefunction is part of the ontic state, so there is an argument from lack of imagination to be made: since all the interpretations of quantum theory that we have managed to come up with that are uncontroversially realist have a real wavefunction, then the wavefunction must be real.

I admit that it behoves the realist ψ-epistemicist to try to construct a fully worked out interpretation. However, absence of evidence is not the same thing as evidence of absence. Nevertheless, I have frequently heard this argument made in private conversations. Some people seem to think that since we have a bunch of well worked out interpretations, we ought to simply pick one of them and not bother thinking about other possibilities. Ever since the inception of quantum theory we have been beset by the problem of quantum jumps, by which I mean that quantum theorists are liable to jump to conclusions.

Despite the obvious weakness of this argument, there is a more subtle point to be made. John Bell was motivated to work on his eponymous theorem by noting that de Broglie-Bohm theory exhibited nonlocality. He wanted to know if this was just a quirk of de Broglie-Bohm theory or an inescapable property of any realist interpretation of quantum theory. With this motivation, he ended up proving the latter. The lesson of this is that if we find that all realist interpretations of quantum theory share a property that some find objectionable then we ought to determine whether or not this is a necessary property. However, this is a motivation for developing ψ-ontology theorems, rather than regarding the matter as settled a priori.

3.4 Quantum computation

The final argument I want to consider is due to David Deutsch, who put it forward as an argument in favor of the many-worlds interpretation. However, I think the argument can be adapted, more generally, into an argument for the reality of the wavefunction. Here is the argument in Deutsch's own words.

> To predict that future quantum computers, made to a given specification, will work in the ways I

> have described, one need only solve a few uncontroversial equations. But to explain exactly *how* they will work, some form of multiple-universe language is unavoidable. Thus quantum computers provide irresistible evidence that the multiverse is real. One especially convincing argument is provided by quantum algorithms [...] which calculate more intermediate results in the course of a single computation than there are atoms in the visible universe. When a quantum computer delivers the output of such a computation, we shall know that those intermediate results must have been computed somewhere, because they were needed to produce the right answer. So I issue this challenge to those who still cling to a single-universe world view: *if the universe we see around us is all there is, where are quantum computations performed?* I have yet to receive a plausible reply. — David Deutsch [89] [Emphasis in original].

Quantum algorithms that offer exponential improvement over existing classical algorithms, such as Shor's factoring algorithm [90], start by putting a quantum system in a superposition of all possible input strings. Then, some computation is done on each of the strings before using interference effects between them to elicit the answer to the computation. If each of the branches of the wavefunction were not individually real, whether or not they are interpreted in a many-worlds sense, then where does the computation get done?

This is not exactly an argument for the reality of the wavefunction, but it is at least an argument that the size of ontic state space should scale exponentially with the number of qubits, and that the ontic state should contain pieces that look like the branches of a wavefunction. However, whilst I agree with Deutsch that an interpretation of quantum theory should offer an explanation of how quantum computations work, it is not at all obvious that the explanation must be a direct translation of what happens to the wavefunction. The argument would perhaps be more compelling if there were known exponential speedups for problems where we think that the best we can do classically is to just search through an exponentially large set of solutions, since we could then argue that a quantum computer must be doing just that. This would be the case if we had such a speedup for the travelling salesman problem, or any other NP complete problem. The sort of problems for which we do have exponential speedup, such as factoring, are more subtle than this. They lie in NP, but are not NP complete. If we were to find an efficient classical algorithm for these problems then it would not cause the whole structure of computational complexity theory to come crashing down. If such an algorithm exists, then whatever deeper theory underlies quantum theory may be exploiting this same structure to perform the quantum computation.

Even if such a scenario does not play out, Deutsch's argument is not decisive against realist ψ-epistemic interpretations. Since we have not yet constructed a viable interpretation of this sort that covers the whole of quantum theory, who knows what explanations such a theory might provide? Therefore, as Deutsch says, explaining quantum computation ought to be viewed as a challenge for the ψ-epistemic program rather than an argument against it.

4 Formalizing the ψ-ontic/epistemic distinction

Hopefully, by this point I have convinced you that it is worth trying to settle the question of the reality of the wavefunction rigorously. The aim of this section is to provide a formal definition of what it means for the quantum state to be ontic or epistemic within a realist model of quantum theory. This is usually done within the framework of *ontological models*. This is really no different from the framework that Bell used to prove his eponymous theorem, and an ontological model is sometimes alternatively known as a hidden variable theory. However, I prefer the term "ontological model" because there is a lot of confusion about the meaning of the term "hidden variables". Following the example of Bohmian mechanics, a hidden variable theory is often thought to be a theory in which some additional variables are posited alongside the wavefunction, which is itself conceived of as ontic from the start. Since the reality of the wavefunction is precisely the point at issue, we definitely want to include models in which it is not assumed to be real within our framework. In addition, in order to cover orthodox quantum theory, we want our framework to include models in which the wavefunction is the *only* thing that is real, i.e. there are no additional hidden variables. A further confusion is the commonly held view that a hidden variable theory must restore determinism, whereas we want to

allow for the possibility that nature might be genuinely stochastic. For these reasons, I prefer to use the term "ontological model". It is either the same thing as a hidden variable theory or more general, depending on how general you thought hidden variable theories were in the first place.

Whilst the Hardy and Colbeck-Renner Theorems involve assumptions about how dynamics are represented in an ontological model, the PBR Theorem only involves prepare-and-measure experiments, i.e. a system is prepared in some quantum state and is then immediately measured and discarded. Therefore, we deal with prepare-and-measure experiments first and defer discussion of dynamics until it is needed. A ψ-ontology theorem aims at proving that any ontological model that reproduces quantum theory must have ontic quantum states. This does not apply to arbitrary fragments of quantum theory, since we have seen in §2.2 that there are fragments that can be modelled with epistemic quantum states. In order to understand both cases, we need to define ontological models for fragments of quantum theory rather than just for quantum theory as a whole. The formal definition of a prepare-and-measure fragment of quantum theory is given in §4.1 and then §4.2 explains how these are represented in ontological models, with examples given in §4.3. Based on this, §4.4 gives the formal definition of what it means for the quantum state to be ontic or epistemic.

4.1 Prepare and measure experiments

In a prepare-and-measure experiment, the experimenter prepares a system in some quantum state and then immediately measures an observable. After the observable is measured, the system is discarded, but the experimenter can repeat the whole process of preparing and measuring as many times as she likes in order to build up frequency statistics. Each run of the experiment is assumed to be statistically independent of the others and we assume that the experimenter can choose which measurement to perform independently of the choice of preparation.

For completeness, we consider the most general type of quantum state—a density operator—and the most general type of observable —a Positive Operator Valued Measure (POVM). Readers unfamiliar with these concepts should consult a standard textbook, such as [77] or [91].

Definition 4.1. A *prepare-and-measure (PM) fragment* $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ of quantum theory consists of a Hilbert space $\mathcal{H}$, a set $\mathcal{P}$ of density operators on $\mathcal{H}$, and a set $\mathcal{M}$ of POVMs on $\mathcal{H}$. The probability of obtaining the outcome corresponding to a POVM element $E \in M$ when performing a measurement $M \in \mathcal{M}$ on a system prepared in the state $\rho \in \mathcal{P}$ is given by the Born rule

$$\mathrm{Prob}(E|\rho, M) = \mathrm{Tr}\left(E\rho\right). \tag{6}$$

For many of the results reviewed here, the PM fragment under consideration is the one in which $\mathcal{P}$ is the set of all pure states on $\mathcal{H}$ and $\mathcal{M}$ consists of measurements in a set of complete orthonormal bases. The formalism of PM fragments allows the sets of states and measurements required to make ψ-ontology theorems go through to be made explicit. Additionally, many of the intermediate results used in proving ψ-ontology theorems apply to PM fragments in general, including those that feature mixed states and POVMs, so it is worth introducing fragments at this level of generality.

4.2 Ontological models

The idea of an ontological model of a PM fragment $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ is that there is some set Λ of *ontic states* that give a complete specification of the properties of the system as they exist in reality. When a quantum system is prepared in a state $\rho \in \mathcal{P}$, what really happens is that the system occupies one of the ontic states $\lambda \in \Lambda$. However, the preparation procedure may not completely control the ontic state, so our knowledge of the ontic state is described by a probability measure μ over Λ. This means that Λ needs to be a measurable space, with a σ-algebra Σ, and that $\mu : \Sigma \to [0, 1]$ is a σ-additive function satisfying $\mu(\Lambda) = 1$. For those unfamiliar with measure-theoretic probability, for a finite space Σ would just be the set of all subsets of Λ and σ-additivity reduces to $\mu(\Omega_1 \cup \Omega_2) = \mu(\Omega_1) + \mu(\Omega_2)$ where Ω_1 and Ω_2 are disjoint subsets. In general, different methods of preparing the same quantum state may result in different probability measures. This

is especially true of mixed states, since they do not have unique decompositions into a convex mixture of pure states. If one prepares a mixed state by choosing randomly from one of the pure states in such a decomposition, then one can prove that the probability distribution must in general depend on the choice of decomposition. This is known as *preparation contextuality*, and is discussed in more detail in §5.4. For this reason, a quantum state ρ is associated with a set Δ_ρ of probability measures rather than just a single unique measure. Note that it is possible to find models in which pure states correspond to unique measures, and much of the literature implicitly assumes this type of model. However, it turns out that this assumption is not necessary, so we allow for the possibility that even a pure state is represented by a set of probability measures for the sake of generality.

Turning now to measurements, a measurement $M \in \mathcal{M}$ might not reveal λ exactly but depend on it only probabilistically. This could be because nature is fundamentally stochastic, but it could also arise in a deterministic theory if the response of the measuring device depends not only on λ but also on degrees of freedom within the measuring device that are not under the experimenter's control (see [92] for a discussion of this). To account for this, each POVM element $E \in M$ is represented by a positive measurable response function $\xi_E^M : \Lambda \to [0,1]$, where $\xi_E^M(\lambda)$ is the conditional probability of obtaining the outcome E given that the ontic state is λ and the measurement M is performed. Since each λ must give rise to a well defined probability distribution for the outcomes of M, the response functions must satisfy $\sum_{E\in M} \xi_E^M(\lambda) = 1$ for all $\lambda \in \Lambda$. See Fig. 6 for an illustration of the response functions corresponding to a measurement. Note that the response functions generally depend not only on the POVM element E, but also the measurement M in which it appears. If E appears in two different measurements, M and N, then it need not be the case that $\xi_E^M = \xi_E^N$. This is to allow for *measurement contextuality*, which says that $\xi_E^M \neq \xi_E^N$ must sometimes occur in certain types of model. The most familiar form of this is the Kochen-Specker Theorem [70], which applies to deterministic models, but it has been generalized to POVMs and nondeterministic models by Spekkens [65]. Measurement contextuality is discussed further in 5.2.

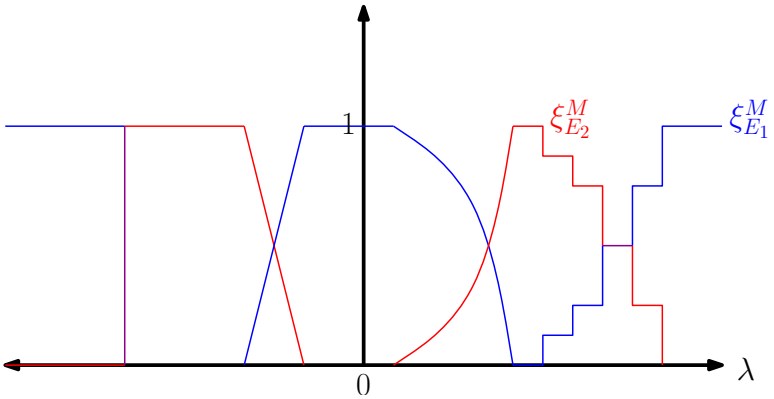


Figure 6: Possible response functions $\xi_{E_1}^M$ and $\xi_{E_2}^M$ for a two outcome measurement $M = \{E_1, E_2\}$ in an ontological model. The values of the response functions are the conditional probabilities of obtaining the outcome E_j when M is measured and the ontic state is λ. As a result, they must sum to 1 everywhere. For illustrative purposes, the ontic state space is represented as a 1-dimensional line, but it may actually be an arbitrary measurable space.

Finally, if the ontological model is to reproduce the quantum predictions of $\mathfrak{F}$ then, for each $\rho \in \mathcal{P}$, every $\mu \in \Delta_\rho$ must satisfy

$$\int_\Lambda \xi_E^M(\lambda) d\mu(lambda) = \mathrm{Tr}\,(E\rho)\,, \tag{7}$$

for all $M \in \mathcal{M}, E \in M$.

Summarizing, we have

Definition 4.2. An *ontological model* $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ of a prepare-and-measure fragment $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ consists of a measurable space (Λ, Σ), where Λ is called the *ontic state space*, a function Δ that maps each $\rho \in \mathcal{P}$ to a set of probability measures $\Delta(\rho) = \Delta_\rho$, and a function ξ that maps each pair $M \in \mathcal{M}, E \in M$ to a measurable response function $\xi(M, E) = \xi_E^M : \Lambda \to [0, 1]$ such that

$$\forall M \in \mathcal{M}, \lambda \in \Lambda, \qquad \sum_{E \in M} \xi_E^M(\lambda) = 1. \tag{8}$$

The ontological model *reproduces the quantum predictions* if, for all $\rho \in \mathcal{P}$, each $\mu \in \Delta_\rho$ satisfies

$$\forall M \in \mathcal{M}, E \in M, \qquad \int_\Lambda \xi_E^M(\lambda) d\mu(\lambda) = \mathrm{Tr}\left(E\rho\right). \tag{9}$$

In what follows, it is convenient to use projectors $|\psi\rangle\langle\psi|$ to represent pure states instead of vectors $|\psi\rangle$ to avoid the global phase ambiguity. The set of projectors onto the pure states of $\mathcal{H}$ is known as the *projective Hilbert space* of $\mathcal{H}$. The notation $[\psi] = |\psi\rangle\langle\psi|$ and $\Delta_\psi = \Delta_{[\psi]}$ is used to reduce clutter. Similarly, if a measurement M consists of projectors onto pure states $[\phi]$ then we use the notation ξ_ϕ^M as shorthand for $\xi_{[\phi]}^M$.

4.3 Examples of Ontological Models

Example 4.3 (Spekkens' toy bit)**.** Spekkens' toy bit can be recast as an ontological model of the prepare-and-measure fragment $\langle \mathbb{C}^2, \mathcal{P}, \mathcal{M} \rangle$ where

$$\mathcal{P} = \{[x+], [x-], [y+], [y-], [z+], [z-], I/2\}, \tag{10}$$

and

$$\mathcal{M} = \{\{[x+], [x-]\}, \{[y+], [y-]\}, \{[z+], [z-]\}\}. \tag{11}$$

The ontic state space is $\Lambda = \{(+,+), (+,-), (-,+), (-,-)\}$. Since Λ is a finite set, the integral in Eq. (9) is just a sum

$$\sum_{\lambda \in \Lambda} \xi_E^M(\lambda)\mu(\lambda) = \mathrm{Tr}\left(E\rho\right), \tag{12}$$

where, abusing notation slightly, we write $\mu(\lambda)$ for $\mu(\{\lambda\})$.

In this model, each quantum state is represented by a unique probability measure. The six probability functions $|x\pm), |y\pm), |z\pm)$ illustrated in Fig. 3 represent the states $[x\pm], [y\pm], [z\pm]$, and $|I/2)$ from Fig. 5 represents the maximally mixed state $I/2$. The response functions are as described in Fig. 3. For example, for the $X = \{[x+], [x-]\}$ measurement we have

$$\begin{array}{rclcl} \xi_{x+}^X(+,+) & = & \xi_{x+}^X(+,-) & = & 1 \\ \xi_{x+}^X(-,+) & = & \xi_{x+}^X(-,-) & = & 0 \\ \xi_{x-}^X(+,+) & = & \xi_{x-}^X(+,-) & = & 0 \\ \xi_{x-}^X(-,+) & = & \xi_{x-}^X(-,-) & = & 1. \end{array} \tag{13}$$

It is easy to check that this reproduces the quantum predictions.

Example 4.4 (The Beltrametti-Bugajski model [93])**.** This model is essentially a translation of the orthodox interpretation of quantum theory into the language of ontological models. The prepare-and-measure fragment is $\langle \mathbb{C}^d, \mathcal{P}, \mathcal{M} \rangle$ where $\mathcal{P}$ contains all the pure states on $\mathbb{C}^d$ and $\mathcal{M}$ consists of all POVMs on $\mathbb{C}^d$. The idea of the model is that the quantum state, and only the quantum state, represents reality. Therefore, the ontic state space Λ is just the set of pure states $[\lambda]$ for $|\lambda\rangle \in \mathbb{C}^d$, i.e. states differing by a global phase are identified so Λ is the projective Hilbert space of $\mathbb{C}^d$. This space carries a natural topology induced by the inner product, and we take Σ to be the Borel σ-algebra of this topology (see [94] for details).

A pure quantum state $[\psi]$ is then represented by the point measure

$$\mu(\Omega) = \delta_{[\psi]}(\Omega) = \begin{cases} 1 & \text{if } [\psi] \in \Omega \\ 0 & \text{if } [\psi] \notin \Omega. \end{cases} \tag{14}$$

Since the quantum state is the ontic state in this model, it should come as no surprise that the response functions simply specify the quantum probabilities, i.e. if a POVM $M \in \mathcal{M}$ contains the operator E then

$$\xi_E^M([\lambda]) = \mathrm{Tr}\left(E\,[\lambda]\right). \tag{15}$$

We then trivially have

$$\int_\Lambda \xi_E^M([\lambda]) d\mu([\lambda]) = \int_\Lambda \mathrm{Tr}\left(E\,[\lambda]\right) d\delta_{[\psi]}([\lambda]) = \mathrm{Tr}\left(E\,[\psi]\right), \tag{16}$$

so the model reproduces the quantum predictions.

The model can be extended to mixed states by writing them as convex combinations of pure states. For example, the maximally mixed state of a spin-1/2 particle can be written as

$$\frac{I}{2} = \frac{1}{2}\,[x+] + \frac{1}{2}\,[x-] \tag{17}$$

so its predictions can be reproduced by the measure

$$\mu = \frac{1}{2}\delta_{[x+]} + \frac{1}{2}\delta_{[x-]} \tag{18}$$

Note however that this is a preparation contextual model because preparing the maximally mixed state by mixing a different set of states, e.g.

$$\frac{I}{2} = \frac{1}{2}\,[y+] + \frac{1}{2}\,[y-] \tag{19}$$

yields a different distribution

$$\mu = \frac{1}{2}\delta_{[y+]} + \frac{1}{2}\delta_{[y-]} \tag{20}$$

which reproduces the quantum predictions just as well. In general, Δ_ρ consists of one convex combination of point measures for each of the different ways of writing ρ as a mixture of pure states.

Example 4.5 (The Bell Model [95])**.** In his review of no-go theorems for hidden variable theories [95], Bell introduced an ontological model for measurements in orthonormal bases on two-dimensional systems. The obvious generalization to arbitrary finite dimensional systems is presented here, as it is needed in §7.5. Bell intended his model as a pedagogical device to point out the flaws in previous no-go theorems, and he was primarily interested in whether a deterministic theory could reproduce the quantum predictions. In modern terms, the Bell model can be thought of as a minimal modification of the Beltrametti-Bugajski model, intended to make it deterministic.

The PM fragment of interest is $\mathfrak{F} = \langle \mathbb{C}^d, \mathcal{P}, \mathcal{M} \rangle$, where $\mathcal{P}$ consists of all pure states and $\mathcal{M}$ consists of all measurements of the form $\mathcal{M} = \{[\phi_j]\}_{j=1}^d$, where $\{|\phi_j\rangle\}_{j=1}^d$ is an orthonormal basis for $\mathbb{C}^d$.

The generalized Bell model employs an ontic state space $\Lambda = \Lambda_1 \times \Lambda_2$ that is the cartesian product of two state spaces. As in Beltrametti-Bugajski, Λ_1 is the projective Hilbert space of $\mathbb{C}^d$, with Borel σ-algebra Σ_1. Λ_2 is the unit interval $[0, 1]$, with Borel σ-algebra Σ_2, representing an additional hidden variable. The measurable space is then $(\Lambda_1 \times \Lambda_2, \Sigma_1 \otimes \Sigma_2)$, where the tensor product σ-algebra $\Sigma_1 \otimes \Sigma_2$ is the σ-algebra generated by sets of the form $\Omega_1 \times \Omega_2$ with $\Omega_1 \in \Sigma_1$, $\Omega_2 \in \Sigma_2$.

The quantum state $[\psi]$ is represented by a product measure

$$\mu(\Omega) = \int_{\Lambda_2} \mu_1(\Omega_{\lambda_2}) d\mu_2(\lambda_2), \tag{21}$$

where μ_1 and μ_2 are probability measures on Λ_1 and Λ_2 respectively and $\Omega_{\lambda_2} = \{[\lambda_1] \in \Lambda_1 | ([\lambda_1], \lambda_2) \in \Omega\}$. As in the Beltrametti-Bugajski model, $[\lambda_1]$ represents the quantum state so $\mu_1 = \delta_{[\psi]}$. The other variable λ_2 is uniformly distributed so μ_2 is just the uniform measure on $[0, 1]$.

The outcome of a measurement $M = \{[\phi_j]\}_{j=1}^d$ is determined as follows. For each $[\lambda_1]$, the unit interval is divided into d subsets of length $\mathrm{Tr}([\phi_j][\lambda_1])$. If λ_2 is in the jth such subset then the $[\phi_j]$ outcome occurs with certainty. It does not matter how we choose the subsets so long as they are disjoint and of the right length. One way of doing it is to pick the jth set to be the interval

$$\sum_{k=1}^{j-1} \mathrm{Tr}([\phi_k][\lambda_1]) \leq \lambda_2 < \sum_{k=1}^{j} \mathrm{Tr}([\phi_k][\lambda_1]), \tag{22}$$

as illustrated in Fig. 7.

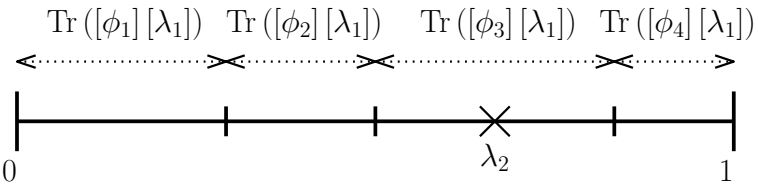


Figure 7: Example of the response functions of the Bell model for a measurement in $\mathbb{C}^4$. The unit interval is divided into 4 subintervals of length $\mathrm{Tr}([\phi_j][\lambda_1])$. If λ_2 is in the jth interval then the outcome will be $[\phi_j]$. In the case depicted, the outcome is $[\phi_3]$.

This corresponds to the response functions

$$\xi_{\phi_j}^M([\lambda_1], \lambda_2) = \begin{cases} 1 & \text{if } \sum_{k=1}^{j-1} \mathrm{Tr}([\phi_k][\lambda_1]) \leq \lambda_2 < \sum_{k=1}^{j} \mathrm{Tr}([\phi_k][\lambda_1]) \\ 0 & \text{otherwise.} \end{cases} \tag{23}$$

It is straightforward to see that this model reproduces the quantum predictions. If the state prepared is $[\psi]$ then the point measure μ_1 implies that $[\lambda_1] = [\psi]$. Therefore, the length of the jth subset of the unit interval will be $\mathrm{Tr}([\phi_j][\psi])$. Since μ_2 is the uniform measure on $[0, 1]$, the probability of λ_2 being in a subset of $[0, 1]$ is just the total length of the subset, so the observed probability of obtaining outcome $[\phi_j]$ is $\mathrm{Tr}([\phi_j][\psi])$, as required.

Example 4.6 (The Kochen-Specker model [70])**.** Kochen and Specker introduced an ontological model for measurements in orthonormal bases on a two-dimensional system. The prepare-and-measure fragment is $\langle \mathbb{C}^2, \mathcal{P}, \mathcal{M} \rangle$ where $\mathcal{P}$ consists of all pure states on $\mathbb{C}^2$ and $\mathcal{M}$ consists of all measurements of the form $M = \{[\phi], [\phi^\perp]\}$ with $\{|\phi\rangle, |\phi^\perp\rangle\}$ an orthonormal basis.

A pure state $[\psi]$ in a two dimensional Hilbert space can be represented as a point $\vec{\psi} = (\sin\vartheta\cos\varphi, \sin\vartheta\sin\varphi, \cos\vartheta)$ on the surface of a unit 2-sphere S_2 by choosing a representative vector $|\psi\rangle$ via

$$|\psi\rangle = \cos(\frac{\vartheta}{2})|z+\rangle + e^{\imath\varphi}\sin(\frac{\vartheta}{2})|z-\rangle, \tag{24}$$

where $0 \leq \vartheta < \pi$ and $-\pi < \varphi \leq \pi$. This is known as the Bloch sphere representation (see Fig. 8).

In the Kochen-Specker model, the ontic state space is the unit sphere $\Lambda = S_2$, which can be thought of as the Bloch sphere. The quantum state $[\psi]$ is corresponds to a measure μ that can be written as

$$\mu(\Omega) = \int_\Omega p(\vec{\lambda}) \sin\vartheta d\vartheta d\varphi, \tag{25}$$

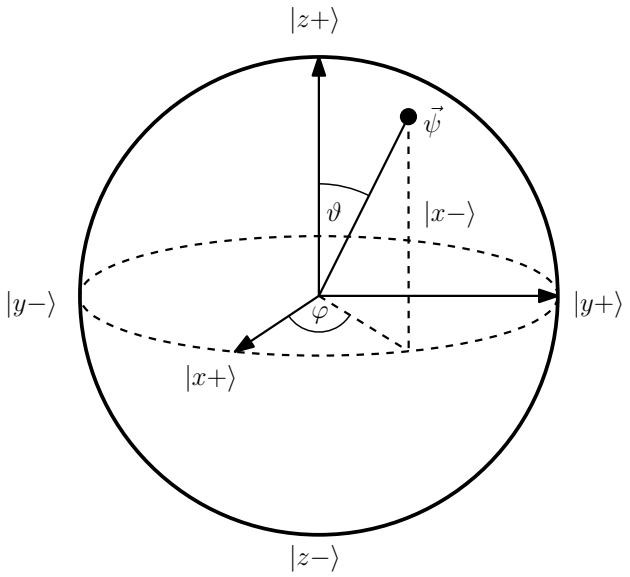


Figure 8: The Bloch sphere representation of a qubit.

where $\vec{\lambda} = (\sin\vartheta\cos\varphi, \sin\vartheta\sin\varphi, \cos\vartheta)$ and the density p is given by

$$p(\vec{\lambda}) = \frac{1}{\pi}\Theta(\vec{\psi}\cdot\vec{\lambda})\vec{\psi}\cdot\vec{\lambda}, \tag{26}$$

where Θ is the Heaviside step function

$$\Theta(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x \leq 0. \end{cases} \tag{27}$$

This density is nonzero on the hemisphere defined by all vectors that subtend an angle of less than $\pi/2$ with $\vec{\psi}$ and it takes values proportional to the cosine of the angle between $\vec{\lambda}$ and $\vec{\psi}$.

For a measurement in the orthonormal basis $\{[\phi], [\phi^\perp]\}$, the outcome $[\phi]$ is obtained if the angle between $\vec{\lambda}$ and $\vec{\phi}$ is smaller than the angle between $\vec{\lambda}$ and $\vec{\phi^\perp}$. Otherwise the outcome $[\phi^\perp]$ is obtained. This corresponds to the response functions

$$\xi_\phi^M(\lambda) = \Theta(\vec{\phi}\cdot\vec{\lambda}) \tag{28}$$

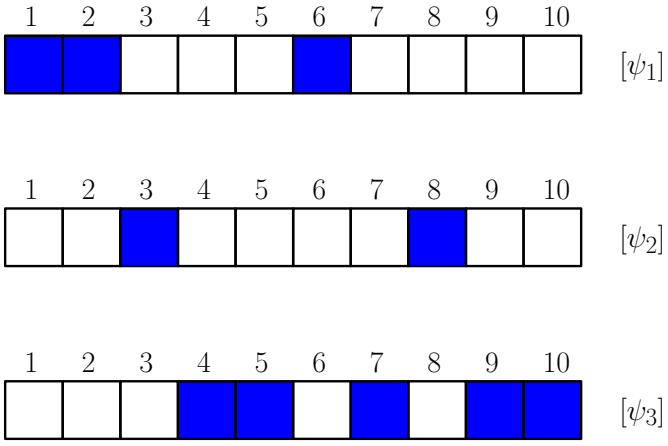
$$\xi_{\phi^\perp}^M(\lambda) = 1 - \xi_\phi^M(\lambda). \tag{29}$$

A proof that this reproduces the quantum predictions is given in Appendix B.

4.4 Defining ψ-ontic/epistemic models

We are now in a position to formally define what it means for the quantum state to be real within an ontological model. The definition presented here is the one used by PBR [27], which is a slightly more rigorous version of a definition originally introduced by Harrigan and Spekkens [68]. It is uncontroversial that mixed states at least sometimes represent knowledge about which of a set of pure states was prepared, so they are at least partially epistemic. For this reason, ψ-ontology theorems are only concerned with proving the reality of pure quantum states. A ψ-ontic model is then one in which, if the pure state $[\psi]$ is prepared, then $[\psi]$ is part of the ontic state of the system. In other words, the ontic state space can be thought of as being composed of the set of pure quantum states along with possibly some extra hidden variables. This

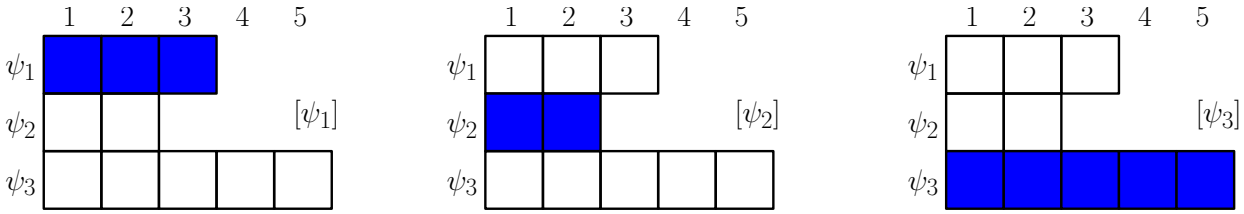
will be the case if the measures corresponding to distinct pure state preparations do not overlap with one another (see Fig. 9). Conversely, the ψ-epistemic explanations of quantum phenomena discussed in §2.1 depend crucially on having overlap between the measures representing different quantum states, so whether or not there is overlap the key issue.



(a) Ontic model with nonoverlapping quantum states

Old label	New label
1	$(\psi_1, 1)$
2	$(\psi_1, 2)$
3	$(\psi_2, 1)$
4	$(\psi_3, 1)$
5	$(\psi_3, 2)$
6	$(\psi_1, 3)$
7	$(\psi_3, 3)$
8	$(\psi_2, 2)$
9	$(\psi_3, 4)$
10	$(\psi_3, 5)$

(b) Relabelling the ontic states



(c) Equivalent model in which the quantum state is explicitly part of the ontology.

Figure 9: In an ontological model, if the probability measures corresponding to distinct quantum states do not overlap, then the ontic states can be relabelled such that the quantum state is explicitly part of the ontology. Fig. (a) depicts the measures corresponding to three quantum states on a discrete ontic state space consisting of the integers from 1 to 10. Blue boxes indicate the ontic states on which the measures have support. Fig. (b) shows a one-to-one map to a new ontic state space in which the quantum states are an explicit part of the ontology. The result is shown in Fig. (c).

The notion of when two probability measures overlap can be formalized using the variational distance.

Definition 4.7. The *variational distance* between two probability measures μ and ν on a measurable space (Λ, Σ) is

$$D(\mu, \nu) = \sup_{\Omega \in \Sigma} |\mu(\Omega) - \nu(\Omega)| \,. \tag{30}$$

Note that taking the absolute value in Eq. (30) is optional, since if $\nu(\Omega) > \mu(\Omega)$ for some $\Omega \in \Sigma$ then

$$\mu(\Lambda \backslash \Omega) - \nu(\Lambda \backslash \Omega) = 1 - \mu(\Omega) - 1 + \nu(\Omega) = \nu(\Omega) - \mu(\Omega), \tag{31}$$

so there is always another measurable set for which the difference of the measures is the same, but μ is larger than ν. Thus, the variational distance can equivalently be defined as

$$D(\mu, \nu) = \sup_{\Omega \in \Sigma} (\mu(\Omega) - \nu(\Omega)) \,, \tag{32}$$

or indeed

$$D(\mu, \nu) = \sup_{\Omega \in \Sigma} (\nu(\Omega) - \mu(\Omega)) \,. \tag{33}$$

The variational distance is a metric on the set of probability measures and it has the following operational interpretation. Suppose that a system is prepared according to one of two preparation procedures, P_1 or P_2, where P_1 corresponds to the measure μ and the P_2 to the measure ν, each case having an equal a priori probability. You are then told the actual value of λ and you wish to make the best possible guess as to which of the two preparation procedures was used. Your best strategy will be to choose a set $\Omega \subseteq \Lambda$ and guess P_1 if $\lambda \in \Omega$ and P_2 if $\lambda \notin \Omega$. More generally, you could use a probabilistic procedure, but convexity implies that this cannot increase your probability of success. Your probability of success is then

$$\text{Prob}(P_1)\text{Prob}(\Omega|P_1) + \text{Prob}(P_2)\text{Prob}(\Lambda\backslash\Omega|P_2) = \frac{1}{2}\mu(\Omega) + \frac{1}{2}\nu(\Lambda\backslash\Omega) \tag{34}$$

$$= \frac{1}{2}\left(1 + \mu(\Omega) - \nu(\Omega)\right), \tag{35}$$

so the maximum probability of success over all such strategies is $\frac{1}{2}\left(1 + D(\mu,\nu)\right)$. Success can occur with unit probability only when $D(\mu,\nu) = 1$, so in this case λ effectively determines which probability density was prepared uniquely.

More rigorously, $D(\mu,\nu) = 1$ is equivalent to the existence of a measurable set $\Omega \in \Sigma$ such that $\mu(\Omega) = 1$ and $\nu(\Omega) = 0$. An optimal guessing strategy then consists of guessing μ if $\lambda \in \Omega$ and ν if $\lambda \notin \Omega$.

If μ and ν are dominated by a third measure m, i.e.

$$\mu(\Omega) = \int_\Omega p(\lambda) dm(\lambda) \tag{36}$$

$$\nu(\Omega) = \int_\Omega q(\lambda) dm(\lambda), \tag{37}$$

for some m-measurable densities p and q, then

$$D(\mu,\nu) = \frac{1}{2}\int_\Lambda |p(\lambda) - q(\lambda)|\, dm(\lambda), \tag{38}$$

which is often the most convenient form for computation.

For those unfamiliar with measure-theoretic probability, a measure m dominates a measure μ if, whenever $m(\Omega) = 0$ then $\mu(\Omega) = 0$. If this holds then μ can be written as a density with respect to m. The measure $m = \frac{1}{2}(\mu + \nu)$ dominates both μ and ν, and a similarly for any finite set of probability measures. For an uncountable set of measures there need not exist a measure that dominates all of them. For example, there is no measure that dominates all of the point measures $\delta_{[\psi]}$ occurring in the Beltrametti-Bugajski model, which is one of the reasons why we work with probability measures instead of probability densities in the present treatment.

Based on the variational distance, we can define what it means for two quantum states to have no overlap in an ontic model.

Definition 4.8. Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M}\rangle$ be a PM-fragment and let $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ be an ontological model of it. A pair of quantum states $\rho, \sigma \in \mathcal{P}$ are *ontologically distinct* in Ξ if, for all $\mu \in \Delta_\rho, \nu \in \Delta_\sigma$,

$$D(\mu,\nu) = 1, \tag{39}$$

otherwise they are *ontologically indistinct*.

The idea here is that, if ρ and σ are ontologically distinct then, regardless of how they are prepared, they can be perfectly distinguished given knowledge of the exact ontic state. In order to take preparation contextuality into account, all possible pairs of distributions must have no overlap.

If a set of states is operationally distinguishable, i.e. there exists a quantum measurement that perfectly distinguishes them, then they must also be ontologically distinct. The intuition behind this is straightforward. Consider the case of a finite ontic state space. If the measures corresponding to two operationally distinguishable quantum states did have nontrivial overlap then there would be a finite probability of an

ontic state occurring that is assigned a finite probability according to both of them, and if this happens then the ontic state cannot be used to deduce which quantum state was prepared with certainty. Hence, such a model could not reproduce the statistics of the distinguishing measurement, which does allow this deduction. More formally,

Definition 4.9. Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment. A set of quantum states $\{\rho_j\} \subseteq \mathcal{P}$ is *operationally distinguishable* (often abbreviated to *distinguishable* when the context is clear) in $\mathfrak{F}$ if there exists a POVM $\{E_j\} \in \mathcal{M}$ such that

$$\mathrm{Tr}\left(E_j \rho_k\right) = \delta_{jk}. \tag{40}$$

Theorem 4.10. *Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment and let $\mathcal{D} \subseteq \mathcal{P}$ be a set of states that is operationally distinguishable in $\mathfrak{F}$. Then, every pair $\rho, \sigma \in \mathcal{D}$, $\rho \neq \sigma$, is ontologically distinct in any ontological model of $\mathfrak{F}$ that reproduces the quantum predictions.*

Proof. Let $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ be an ontological model of $\mathfrak{F}$ that reproduces the quantum predictions and let $M \in \mathcal{M}$ be a measurement that distinguishes the states in $\mathcal{D}$. Then, for $\rho, \sigma \in \mathcal{D}$, $\rho \neq \sigma$, there exists a POVM element $E \in M$ such that

$$\mathrm{Tr}\left(E\rho\right) = 1 \qquad \text{and} \qquad \mathrm{Tr}\left(E\sigma\right) = 0. \tag{41}$$

In the ontological model, this translates to

$$\int_\Lambda \xi_E^M(\lambda) d\mu(\lambda) = 1 \tag{42}$$

$$\int_\Lambda \xi_E^M(\lambda) d\nu(\lambda) = 0, \tag{43}$$

for all $\mu \in \Delta_\rho$, $\nu \in \Delta_\sigma$. In order to satisfy Eq. (42), ξ_E^M must be equal to 1 on a set Ω that is of measure one according to μ.

Assume that ρ and σ are not ontologically distinct. Then, there exist $\mu \in \Delta_\rho$, $\nu \in \Delta_\sigma$ such that $D(\mu, \nu) < 1$. This is equivalent to saying that all sets that are of measure 1 according to μ are of nonzero measure according to to ν. In particular, this applies to Ω, so $\nu(\Omega) > 0$. However, ξ_E^M is equal to 1 on Ω and $\Omega \subseteq \Lambda$, so from Eq. (43) we have

$$0 = \int_\Lambda \xi_E^M(\lambda) d\nu(\lambda) \tag{44}$$

$$\geq \int_\Omega \xi_E^M(\lambda) d\nu(\lambda) \tag{45}$$

$$= \int_\Omega d\nu(\lambda) = \nu(\Omega) > 0, \tag{46}$$

which is a contradiction. □

Although operationally distinguishable quantum states are always ontologically distinct, arbitrary pairs of states need not be ontologically distinct. If all pairs of pure states are nonetheless ontologically distinct then the ontological model is ψ-ontic.

Definition 4.11. An ontological model of $\langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ is *ψ-ontic* if all pairs of pure quantum states $[\psi], [\phi] \in \mathcal{P}$, $[\psi] \neq [\phi]$, are ontologically distinct. Otherwise the model is *ψ-epistemic*.

This definition captures the idea that quantum states are real in an ontological model if the distributions corresponding to distinct pure states do not overlap. Defining a ψ-epistemic model to be the negation of this is extremely permissive. For example, a model in which only a single pair of pure states have any overlap, all other pairs being ontologically distinct, would be ψ-epistemic according to this definition. The ψ-epistemic explanations of quantum phenomena would not apply to such a model. The ψ-ontology theorems discussed

in part II aim to rule out even this permissive notion of a ψ-epistemic model, so they obviously rule out any less permissive definition as well. Nevertheless, because ψ-ontology theorems require questionable auxiliary assumptions, it is still interesting to consider stronger notions of what it means for the quantum state to be epistemic. This is done in in §5.3 and part III.

Note that a ψ-ontic model is not necessarily ψ-*complete*, where the latter means that the wavefunction, and only the wavefunction, is the ontic state. ψ-complete models are obviously ψ-ontic, but in general ψ-ontic models may involve other variables in addition to the wavefunction. A lot of confusion may be avoided by clearly separating the notions of ψ-ontic and ψ-complete models, particularly since much of the literature on hidden variable theories focuses on the question of whether quantum theory is complete, i.e. it is assumed that the wavefunction is real a priori and the only question is whether anything else needs to be added to it. Obviously, such a framework is not suited to discussing the question of whether the wavefunction must be real in the first place.

Of the examples considered in §4.3, Spekkens' toy theory and the Kochen-Specker model are ψ-epistemic, and the Beltrametti-Bugajski and Bell models are ψ-ontic. In Spekkens' theory, the ontic state space is finite, so the integral in Eq. (38) can be performed with respect to the counting measure, which yields the sum

$$D(\mu,\nu) = \frac{1}{2}\sum_{\lambda} |\mu(\lambda) - \nu(\lambda)|. \tag{47}$$

For toy bit, λ takes the values $(+,+),(+,-),(-,+)$ and $(-,-)$. If μ is the $|x+)$ measure, corresponding to the quantum state $[x+]$, and ν is $|y+)$, corresponding to $[y+]$ then we have

$$D(\mu,\nu) = \frac{1}{2}\left(\left|\frac{1}{2}-\frac{1}{2}\right| + \left|\frac{1}{2}-0\right| + \left|0-\frac{1}{2}\right| + |0-0|\right) = \frac{1}{2}. \tag{48}$$

Thus, the model is ψ-epistemic because $D(\mu,\nu) < 1$, as we expect for overlapping measures. Appendix B provides a proof that the Kochen-Specker model is ψ-epistemic by showing that it satisfies the stronger notion of being maximally ψ-epistemic to be discussed in §5.3.

The Beltrametti-Bugajski model is in fact ψ-complete, because its ontic state space consists of the set of pure states and preparing a given pure state causes the corresponding ontic state to be occupied with certainty. It follows that the model is ψ-ontic because each pure state corresponds to a distinct point measure. Thus, for $[\psi] \neq [\phi]$ the corresponding measures are $\delta_{[\psi]}$ and $\delta_{[\phi]}$. The set $\{[\psi]\}$ is measure one according to $\delta_{[\psi]}$ and measure zero according to $\delta_{[\phi]}$. Similarly, the Bell model is ψ-ontic, where now the set $\{[\psi]\} \times [0,1]$ is measure one for the measure corresponding to $[\psi]$ and measure zero for the measure corresponding to $[\phi]$. It is, however, not ψ-complete because of the additional component $[0,1]$ of the ontic state space.

5 Implications of ψ-ontology

Before discussing ψ-ontology theorems, it is worth pausing to consider some of their implications. One of the most interesting things about establishing the reality of the quantum state within the ontological models framework is that it would imply a lot of existing no-go results as simple consequences. Thus, even if you are a neo-Copenhagenist who rejects the ontological models framework outright, you should still be interested in ψ-ontology theorems as potentially the most powerful class of no-go results that we currently have. Additionally, the ontological models framework can be thought of as an attempt to simulate quantum theory using classical resources. Just as Bell's Theorem has implications for the difference between quantum communication complexity and for device independent quantum cryptography [96], ψ-ontology theorems might become a useful tool in quantum information theory.

The main implications of ψ-ontology are illustrated in Fig. 10, and each of them is discussed in this section. There are two main strands of implications, one based on the size of the ontic state space and one based on results related to contextuality. The results of the contextuality strand can alternatively be derived as consequences of the Kochen-Specker Theorem in place of ψ-ontology, so Kochen-Specker contextuality is also discussed below.

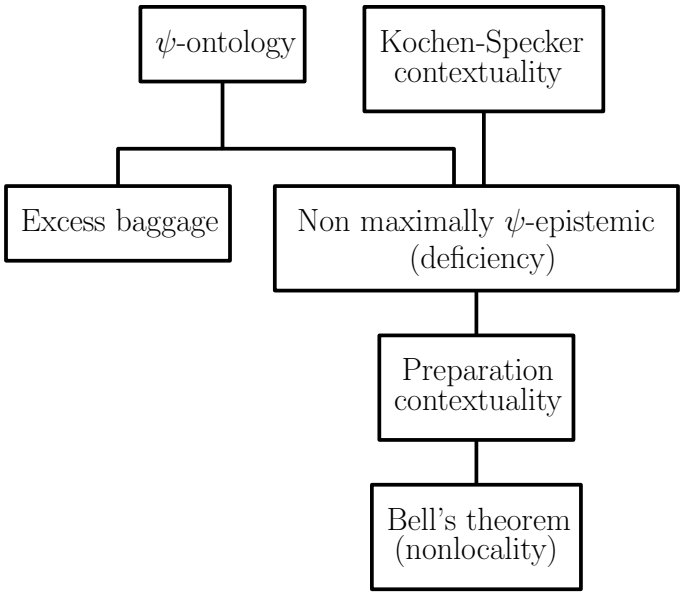


Figure 10: Hierarchy of constraints on ontological models of quantum theory. Properties that appear higher in the diagram imply those that they are connected to lower down. Non maximal ψ-epistemicity can be derived from either ψ-ontology or Kochen-Specker contextuality.

5.1 Excess baggage

The idea of excess baggage has already been discussed in §2.5 as the tension between the infinite amount of information required to specify the quantum state of a qubit and the fact that it can only be used to reliably transmit a single bit. Similarly, the number of parameters required to specify a quantum state scales exponentially with the number of systems, but the amount of information that can be transmitted scales only linearly.

Excess baggage theorems show that this tension cannot be avoided in an ontological model by providing lower bounds on the size of the ontic state space required to reproduce quantum theory. The first such result, from which the "excess baggage" terminology originates, was due to Hardy, who proved that an infinite number of ontic states are required to reproduce the predictions of any quantum system, even just a qubit [59]. Subsequently, Montina showed that the number of parameters required to specify an ontic state must scale exponentially with the number of systems [60, 61].

If it can be proved that an ontological model of quantum theory must be ψ-ontic then it follows immediately that the ontic state space must be uncountably infinite and that the number of parameters required to specify an ontic state must scale exponentially with the number of systems. This is because, in a ψ-ontic model, there must be at least as many ontic states as there are quantum states.

5.2 Kochen-Specker contextuality

Kochen-Specker contextuality is not directly related to ψ-ontology, but many of the consequences of ψ-ontology can alternatively be derived from it.

The notion of contextuality first arose in Kochen and Specker's attempt to prove a no-go theorem for hidden variable theories [70]. Kochen and Specker's definition of contextuality only deals with projective measurements, but this has since been generalized and given a more operational spin by Spekkens [65], and we follow his approach here. The basic idea is that, if two things are operationally equivalent in quantum theory, i.e. if they always give rise to the exact same observable probabilities, then they should be represented the same way in an ontological model. Applied to measurements, this is formally defined as follows.

Definition 5.1. Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment and let $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ be an ontological model of it. Ξ is *measurement noncontextual* if, whenever a pair of measurements $M, N \in \mathcal{M}$ share a common POVM element $E \in M$, $E \in N$, then $\xi_E^M = \xi_E^N$. Otherwise, it is *measurement contextual*.

If a measurement M contains the POVM element E, then, for a system prepared in the state ρ, the outcome E occurs with probability $\mathrm{Tr}(E\rho)$. If another measurement N also contains E then the outcome E still has the same probability $\mathrm{Tr}(E\rho)$ in the measurement N as it did in M. Since there is nothing in the quantum predictions that distinguishes E occurring in M from E occurring in N, a noncontextual model should represent them in the same way. The particular measurement M in which E occurs is called the *measurement context* of E.

The classic example of the same measurement operator occurring in more than one context is to take two orthonormal bases, $\{|\phi_j\rangle\}_{j=1}^d$ and $\{|\phi'_j\rangle\}_{j=1}^d$, such that $|\phi_1\rangle = |\phi'_1\rangle$. Such a pair can be constructed from a unitary U that leaves $|\phi_1\rangle$ invariant via $|\phi'_j\rangle = U|\phi_j\rangle$. Then, the two measurements $M = \{[\phi_j]\}_{j=1}^d$ and $N = \{[\phi'_j]\}_{j=1}^d$ share the common projector $[\phi_1]$. This can only happen nontrivially if $d \geq 3$, since if $d = 2$ and $U|\phi_1\rangle = |\phi_1\rangle$ then $U|\phi_2\rangle$ can only differ from $|\phi_2\rangle$ by a global phase. For non projective POVMs, nontrivial examples can be constructed for $d = 2$ as well (see [65] for details), but we are only concerned with the traditional Kochen-Specker notion of contextuality here, which only applies to projective measurements.

Definition 5.2. Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment where $\mathcal{M}$ consists of projective measurements. An ontological model $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ of $\mathfrak{F}$ is *Kochen-Specker (KS) noncontextual* if it is both:

- Outcome deterministic: $\forall M \in \mathcal{M}, E \in M, \lambda \in \Lambda,\ \xi_E^M(\lambda) \in \{0, 1\}$.
- Measurement noncontextual.

Otherwise the model is *KS contextual*.

In other words, KS noncontextuality only applies to projective measurements and is the combination of Spekkens' notion of measurement noncontextuality with outcome determinism, i.e. the idea that the ontic state should determine the outcome of a projective measurement with certainty. The Kochen-Specker Theorem [70], and other proofs of KS contextuality [97–103], show that it is impossible to construct a KS noncontextual model for all projective measurements in Hilbert spaces of dimension ≥ 3.

Note that a model may be KS contextual either by being measurement contextual or by being being nondeterministic. It can of course be both, but either one on its own is sufficient to reproduce the quantum predictions. The Beltrametti-Bugajski model is an example of a model that is nondeterministic, but measurement noncontextual because the response functions $\xi_E^M([\lambda]) = \mathrm{Tr}(E[\lambda])$ just mimic the quantum probabilities, which do not depend on M. On the other hand, the Bell model is deterministic, but measurement contextual because the way in which the unit interval is divided up depends on the ordering of the measurement operators.

Unlike ψ-ontology, KS contextuality does not obviously imply excess baggage, but many of the other implications of ψ-ontology are implied by KS contextuality. Therefore, one might wonder whether ψ-ontology implies KS contextuality. This is not the case, as the Bell model in two-dimensions is ψ-ontic but can be converted into a KS noncontextual model by reordering the way in which the unit interval is divided, e.g. by making the first interval always correspond to the measurement outcome in the northern hemisphere of the Bloch sphere. Similarly, there are ψ-epistemic models that are KS contextual, one of which is described in §7.5, so ψ-ontology and KS contextuality are inequivalent constraints on ontological models.

The following characterization of KS noncontextual models will be useful in what follows.

Definition 5.3. Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment and let $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ be an ontological model of it. For every $M \in \mathcal{M}$, $E \in M$, let

$$\Lambda^{(M,E)} = \left\{ \lambda \in \Lambda \middle| \xi_E^M(\lambda) = 1 \right\}. \tag{49}$$

The *cosupport* Λ^E of a POVM element E is then

$$\Lambda^E = \cap_{\{M \in \mathcal{M} | E \in M\}} \Lambda^{(M,E)}. \tag{50}$$

In other words, the cosupport of E is the set of ontic states that always return the outcome E with certainty in any measurement context that contains it.

Theorem 5.4. *Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment, where $\mathcal{M}$ consists of projective measurements. If an ontological model $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ of $\mathfrak{F}$ is KS noncontextual then, for all $M \in \mathcal{M}$, $E \in M$, $\rho \in \mathcal{P}$, every $\mu \in \Delta_\rho$ satisfies*

$$\int_\Lambda \xi_E^M(\lambda) d\mu(\lambda) = \mu(\Lambda^E). \tag{51}$$

Proof. If Ξ is KS noncontextual then ξ_E^M is independent of its measurement context M, so, for any M containing E, $\Lambda^E = \Lambda^{(E,M)} = \{\lambda \in \Lambda | \xi_E^M(\lambda) = 1\}$. Furthermore, $\xi_E^M(\lambda)$ is either equal to 0 or 1 everywhere, so we only need integrate over those λ for which it is equal to 1. This gives

$$\int_\Lambda \xi_E^M(\lambda) d\mu(\lambda) = \int_{\Lambda^E} \xi_E^M(\lambda) d\mu(\lambda) \tag{52}$$

$$= \int_{\Lambda^E} d\mu(\lambda) \tag{53}$$

$$= \mu(\Lambda^E), \tag{54}$$

as required. □

The converse to this theorem does not hold, i.e. if Eq. (51) always holds then the model may still be KS contextual. This is because there may be subsets of the ontic state space that are of measure zero according to all the probability measures that occur in the model. The values of the response functions on these sets are not constrained by Eq. (51). However, a partial converse can be obtained by excising some of the measure zero sets from the ontic state space. Since they are of measure zero, their removal does not substantively affect the structure of the model.

Definition 5.5. Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment. An ontological model $\tilde{\Xi} = (\tilde{\Lambda}, \tilde{\Sigma}, \tilde{\Delta}, \tilde{\xi})$ of $\mathfrak{F}$ is a *measure zero revision* of another ontological model $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ of $\mathfrak{F}$ if

- The ontic state space $\tilde{\Lambda}$ differs from Λ only by the removal of a measure zero set. Formally, $\tilde{\Lambda} \in \Sigma$ and for all $\rho \in \mathcal{P}$ and $\mu \in \Delta_\rho$, $\mu(\tilde{\Lambda}) = 1$. Further, $\tilde{\Sigma}$ is the sub-σ-algebra of Σ containing all $\Omega \in \Sigma$ such that $\Omega \subseteq \tilde{\Lambda}$.

- The probability measures of $\tilde{\Xi}$ are the restrictions of the probability measures of Ξ to $(\tilde{\Lambda}, \tilde{\Sigma})$. Formally, for all $\rho \in \mathcal{P}$, there is a surjective mapping $f : \Delta_\rho \to \tilde{\Delta}_\rho$ defined by

  $$f(\mu)(\Omega) = \mu(\Omega), \tag{55}$$

  for all $\Omega \in \tilde{\Sigma}$.

- The response functions of $\tilde{\Xi}$ are the restrictions of the response functions of Ξ to $\tilde{\Lambda}$. Formally, for all $M \in \mathcal{M}$, $E \in M$ and $\lambda \in \tilde{\Lambda}$, $\tilde{\xi}_E^M(\lambda) = \xi_E^M(\lambda)$.

Preforming a measure zero revision does not represent a substantive change in the following sense. Imagine you are passively observing an experimenter who is performing a sequence of prepare-and-measure experiments in the fragment $\mathfrak{F}$. Suppose that, in addition to observing the choices of preparation and measurement that the experimenter makes and the measurement outcomes she obtains, you also get to see the exact ontic state λ in every run of the experiment. You write down a long list of data consisting of the preparation performed, the ontic state, the choice of measurement, and the measurement outcome in each run of the experiment. This data would allow you to distinguish between an ontological model Ξ and a measure zero revision of it $\tilde{\Xi}$ only if an ontic state in $\Lambda \backslash \tilde{\Lambda}$ happens to be occupied in at least one run of the experiment, since this can only happen in Ξ and the probabilistic predictions made by the two models are otherwise exactly the same. However, the probability of this happening is zero because the set of ontic states removed

has zero probability according to all the probability measures in the theory. To all intents and purposes then, the two models tell the same story about reality.

A partial converse to Theorem 5.4 can be obtained as follows.

Theorem 5.6. *Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment, where $\mathcal{M}$ consists of projective measurements and is at most countable. Let $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ be an ontological model of $\mathfrak{F}$. If, for all $M \in \mathcal{M}$, $E \in M$, $\rho \in \mathcal{P}$, every $\mu \in \Delta_\rho$ satisfies*

$$\int_\Lambda \xi_E^M(\lambda) d\mu(\lambda) = \mu(\Lambda^E), \tag{56}$$

then there exists a measure zero revision of Ξ that is KS noncontextual.

Proof. The measure zero revision $\tilde{\Xi} = (\tilde{\Lambda}, \tilde{\Sigma}, \tilde{\Delta}, \tilde{\xi})$ is constructed as follows. First, let $M = \{E_j\} \in \mathcal{M}$ and note that, for all $\rho \in \mathcal{P}$, every $\mu \in \Delta_\rho$ satisfies

$$\begin{aligned} 1 = \mu(\Lambda) &= \int_\Lambda d\mu(\lambda) && (57)\\ &= \int_\Lambda \sum_j \xi_{E_j}^M(\lambda) d\mu(\lambda) && (58)\\ &= \sum_j \int_\Lambda \xi_{E_j}^M(\lambda) d\mu(\lambda) && (59)\\ &= \sum_j \mu(\Lambda^{E_j}), && (60) \end{aligned}$$

where the second line follows from $\sum_j \xi_{E_j}^M(\lambda) = 1$ and the fourth from Eq. (56).

The sets Λ^{E_j} are disjoint because if $\xi_{E_j}^M(\lambda) = 1$ then it must be the case that $\xi_{E_k}^M(\lambda) = 0$ for all $k \neq j$ in order to satisfy $\sum_j \xi_{E_j}^M(\lambda) = 1$. Therefore, $\Lambda^M = \cup_j \Lambda^{E_j}$ is a set of measure one according to μ. Now set $\tilde{\Lambda} = \cap_{M \in \mathcal{M}} \Lambda^M$. This is also measure one according to μ, since it is the intersection of an at most countable number of measure one sets. The probability measures in the revision are then obtained by restriction to $\tilde{\Lambda}$.

For any measurement $M = \{E_j\} \in \mathcal{M}$, the response functions $\tilde{\xi}_{E_j}^M$ obtained by restriction of $\xi_{E_j}^M$ to $\tilde{\Lambda}$ are equal to 1 on Λ^{E_j} and are zero elsewhere on $\tilde{\Lambda}$. The latter follows because $\tilde{\Lambda} \subseteq \cup_j \Lambda^{E_j}$ and we already saw that $\xi_{E_j}^M$ is equal to zero on Λ^{E_k} for $k \neq j$. Thus, the model is outcome deterministic. Further, it is measurement noncontextual because Λ^{E_j} does not depend on the measurement context. Thus, the model is KS noncontextual. □

5.3 Maximally ψ-epistemic models

The concept of a maximally ψ-epistemic model was introduced by Maroney as a stronger notion of what it means for an ontological model to be ψ-epistemic [56]. An equivalent concept was introduced earlier by Harrigan and Rudolph [92], under the terminology "non deficient model". As you might imagine from the name, a ψ-ontic model cannot be maximally ψ-epistemic, but, up to measure one revision, KS contextual models cannot be maximally ψ-epistemic either [55, 57, 92].

The real interest in the concept of a maximally ψ-epistemic model is that it enables one to devise measures of the extent to which a model is ψ-epistemic, and hence to go beyond the sharp dichotomy between ψ-epistemic and ψ-ontic implied by the definitions adopted so far [54, 56, 104]. This aspect is discussed in §12.2, but ruling out maximally ψ-epistemic models is also useful as a stepping stone between ψ-ontology/KS contextuality and the next step in the contextuality strand, which is preparation contextuality.

The basic idea is that, in order to justify the ψ-epistemic explanation of distinguishability, we need more than just that the probability measures corresponding to two different quantum states, $[\psi]$ and $[\phi]$, should have nonzero overlap. Ideally, when measuring a system prepared in the state $[\psi]$, all of the probability of obtaining the outcome corresponding to the projector $[\phi]$ should be accounted for by the overlap region

between the probability measures corresponding to the two states. Since we are comparing the overlap of $[\psi]$ and $[\phi]$ considered as a states with the probability of obtaining $[\phi]$ as a measurement outcome, this only makes sense if we are dealing with a PM fragment $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ for which, for every pure state $[\phi] \in \mathcal{P}$, there exists a measurement $M \in \mathcal{M}$ such that $[\phi] \in M$. Therefore, this is a standing assumption for the remainder of this section.

Definition 5.7. Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment and $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ an ontological model of it. The model Ξ is *maximally ψ-epistemic* if, for all pure states $[\psi], [\phi] \in \mathcal{P}$, for all $\mu \in \Delta_\psi$, $\nu \in \Delta_\phi$ and $M \in \mathcal{M}$ with $[\phi] \in M$,

$$\int_\Omega \xi_\phi^M(\lambda) d\mu(\lambda) = \int_\Lambda \xi_\phi^M(\lambda) d\mu(\lambda), \tag{61}$$

for all sets $\Omega \in \Sigma$ such that $\nu(\Omega) = 1$.

To unpack this definition a little, let $\mu \in \Delta_\psi$, let $\nu \in \Delta_\phi$, let M be a measurement that contains $[\phi]$ and let ξ_ϕ^M be the corresponding response function. Suppose Ω is a set of measure one according to ν and Ω' is a set of measure one according to μ. Then, Eq. (61) is equivalent to

$$\int_{\Omega \cap \Omega'} \xi_\psi^M(\lambda) d\mu(\lambda) = \int_\Lambda \xi_\psi^M(\lambda) d\mu(\lambda). \tag{62}$$

Whatever one might mean by the overlap region between μ and ν, it should be the intersection of a set that is measure one according to μ with a set that is measure one according to ν, and since Eq. (62) must hold for all such sets, it guarantees that the probability of obtaining outcome $[\phi]$ when measuring a system prepared in the state $[\psi]$ is entirely accounted for by any such region.

Note that, instead of Eq. (61), previous works [55–57, 92] imposed the requirement that

$$\int_{\Lambda_q} \xi_\phi^M(\lambda) p(\lambda) d\lambda = \int_\Lambda \xi_\phi^M(\lambda) p(\lambda) d\lambda, \tag{63}$$

where $d\lambda$ is a measure that dominates μ and ν, p and q are densities that represent them, and $\Lambda_q = \{\lambda \in \Lambda | q(\lambda) > 0\}$ is the support of the probability density q. However, the probability measure ν does not correspond to a unique density q, since densities that differ on a set of measure zero according to $d\lambda$ represent the same probability measure. This means that Λ_q is not uniquely specified by ν, since setting the values of q to zero on a measure zero set would change Λ_q but not ν. Therefore, to use this definition, one has to imagine that the ontological model specifies a particular density representation for each measure, rather than just the measure itself. On its own, this older definition is also not strong enough to entail that a ψ-ontic model cannot be maximally ψ-epistemic or that Kochen-Specker contextual models cannot be maximally ψ-epistemic. To remedy this, one can adopt a standing assumption that all probability measures appearing in the model are dominated by, i.e. absolutely continuous with respect to, a canonical measure $d\lambda$, as was done explicitly in [55], and this must be regarded as an implicit assumption in previous works [56, 57, 92]. However, we do not really want to make this assumption, because some models of interest do not satisfy it. For example, an uncountable set of point measures on distinct points is not dominated by any measure, and these occur in the Beltrametti-Bugajski and Bell models.

It is straightforward to verify that Spekkens' toy theory is maximally ψ-epistemic and a proof that the Kochen-Specker model is maximally ψ-epistemic is given in Appendix B. On the other hand, a ψ-ontic model that reproduces the quantum predictions cannot be maximally ψ-epistemic, so the Beltrametti-Bugajski and Bell models are not maximally ψ-epistemic.

Theorem 5.8. *Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment such that $\mathcal{P}$ contains at least one pair of nonorthogonal pure states and let $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ be an ontological model of $\mathfrak{F}$ that reproduces the quantum predictions. If Ξ is maximally ψ-epistemic then it is ψ-epistemic.*

Proof. Assume that Ξ is ψ-ontic. Then, for every pair of nonorthogonal pure states $[\psi], [\phi] \in \mathcal{P}$, every pair $\mu \in \Delta_\psi$, $\nu \in \Delta_\phi$ satisfies $D(\mu, \nu) = 0$, which is equivalent to saying that there exists a set Ω that is measure

1 according to ν and measure 0 according to μ. However, if Ξ is maximally ψ-epistemic then, for all $\mu \in \Delta_\psi$, $\nu \in \Delta_\phi$, and $M \in \mathcal{M}$ containing $[\phi]$,

$$\int_\Omega \xi_\phi^M(\lambda) d\mu(\lambda) = \int_\Lambda \xi_\phi^M(\lambda) d\mu(\lambda). \tag{64}$$

This implies that,

$$\begin{aligned} \mu(\Omega) &= \int_\Omega d\mu(\lambda) && (65)\\ &\geq \int_\Omega \xi_\phi^M(\lambda) d\mu(\lambda) && (66)\\ &= \int_\Lambda \xi_\phi^M(\lambda) d\mu(\lambda) && (67)\\ &= \mathrm{Tr}\left([\phi]\,[\psi]\right) > 0, && (68) \end{aligned}$$

where the second line follows from the fact that $\xi_\phi^M(\lambda) \in [0,1]$ and the fourth from the fact that Ξ reproduces the quantum predictions. Thus, Ω cannot be of measure zero according to μ, which contradicts the assertion that the model is ψ-ontic. □

Further, if we restrict attention to models of measurements in a countable set of complete orthonormal bases, a maximally ψ-epistemic model has a measure one revision that is KS noncontextual.

Theorem 5.9. *Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment where, for each $[\phi] \in \mathcal{M}$ the set of $M \in \mathcal{M}$ that contain $[\phi]$ is at most countable. Let $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ be an ontological model of $\mathfrak{F}$ that reproduces the quantum predictions. If Ξ is maximally ψ-epistemic then, for all pure states $[\psi], [\phi] \in \mathcal{P}$, every $\mu \in \Delta_\psi$ satisfies*

$$\int_\Lambda \xi_\phi^M(\lambda) d\mu(\lambda) = \mu(\Lambda^\phi), \tag{69}$$

for every $M \in \mathcal{M}$ that contains $[\phi]$, and where Λ^ϕ is the cosupport of $[\phi]$.

Proof. Let $\nu \in \Delta_\phi$ and consider an $M \in \mathcal{M}$ that contains $[\phi]$. Then, in order to reproduce the quantum predictions we must have

$$\int_\Lambda \xi_\phi^M(\lambda) d\nu(\lambda) = \mathrm{Tr}\left([\phi]\,[\phi]\right) = 1. \tag{70}$$

In order for this to be true, there must be a set $\Omega \in \Sigma$ such that $\nu(\Omega) = 1$ and ξ_ϕ^M is equal to one on Ω. Any such set satisfies $\Omega \subseteq \Lambda^{(M,\phi)}$ and so $\Lambda^{(M,\phi)}$ is also a set of measure one according to ν. Hence $\Lambda^\phi = \cap_{\{M \in \mathcal{M} | [\phi] \in M\}} \Lambda^{(M,\phi)}$ is also of measure one according to ν, since it is the intersection of an at most countable collection of measure one sets. Since ξ_ϕ^M is equal to 1 on Λ^ϕ, we have

$$\int_{\Lambda^\phi} \xi_\phi^M(\lambda) d\mu(\lambda) = \int_{\Lambda^\phi} d\mu(\lambda) = \mu(\Lambda^\phi). \tag{71}$$

Thus, Eq. (69) is a special case of Eq. (61), so it must hold for a maximally ψ-epistemic model. □

Corollary 5.10. *Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment in which $\mathcal{M}$ is at most countable and consists of complete orthonormal projective measurements. Suppose that, for all $M \in \mathcal{M}$, $[\phi] \in M$, it is also the case that $[\phi] \in \mathcal{P}$. Then, any maximally ψ-epistemic ontological model of $\mathfrak{F}$ has a measure one revision that is KS noncontextual.*

The proof just consists of combining Theorem 5.9 with Theorem 5.6.

5.4 Preparation contextuality

The notion of preparation contextuality is due to Spekkens [65]. Recall that the idea of noncontextuality is that things that are operationally equivalent, i.e. always lead to the same observable probabilities, should be represented in the same way in an ontological model. Preparation noncontextuality is just this idea applied to states rather than measurements. This section shows that, when extended to mixed states, non maximally ψ-epistemic models must be preparation contextual, a result that was first pointed out in [55].

Definition 5.11. In an ontological model $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ of a PM fragment $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$, a state $\rho \in \mathcal{P}$ is *preparation noncontextual* if Δ_ρ only contains a single measure. Otherwise ρ is *preparation contextual*. Similarly, Ξ itself is preparation noncontextual if every $\rho \in \mathcal{P}$ is preparation noncontextual, and otherwise Ξ is preparation contextual.

Whenever two preparation procedures result in the same quantum state ρ, there is no measurement that can distinguish between them because, according to quantum theory, all of the outcome probabilities for every measurement are exactly the same. Thus, the general principle of noncontextuality implies that all methods of preparing ρ should result in the same probability measure.

Note that much of the literature on ψ-ontology implicitly assumes preparation noncontextuality for pure states by using a unique measure μ_ψ for each pure state $[\psi]$ under consideration. This assumption is fairly harmless for a couple of reasons. Firstly, the necessity of preparation contextuality has so far only been proved for mixed states, so assuming a unique measure for pure states is not ruled out by any existing no-go theorem. Secondly, adapting results to preparation contextual models is usually just a matter of modifying the definitions in a fairly obvious way that does not require proofs to be modified substantively. Nevertheless, I prefer to take the possibility of preparation contextuality into account explicitly because it is oddly asymmetric to allow for contextual measurements but not contextual states, and because it allows results to be proved under weaker assumptions.

In order to prove preparation contextuality for mixed states, a further assumption is required about how convex combinations of quantum states should be represented in ontological models. To understand this, let $\rho = \sum_j p_j \sigma_j$ be a convex decomposition of a mixed state ρ into other (pure or mixed) states σ_j, i.e. $0 \leq p_j \leq 1$ and $\sum_j p_j = 1$. One method of preparing ρ is to generate a classical random variable that takes value j with probability p_j (e.g. by flipping coins, throwing dice, or any other suitable method), prepare the state σ_j if the value of the classical variable is j, and then discard and forget the value of the classical variable.

In an ontological model, it is reasonable to assume that, in the above mixing procedure, if μ_j is the probability measure corresponding to the method used to prepare σ_j, then $\sum_j p_j \mu_j$ is the measure corresponding to preparing ρ by this mixing procedure. This is because the randomness used to generate the classical variable could come from a source that is completely independent of the system under investigation, in which case we would expect that the ontic state of the system only depends on it via its effect on which of the σ_j is prepared. More formally,

Definition 5.12. Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment, and let $\rho \in \mathcal{P}$. An ontological model $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ of $\mathfrak{F}$ *respects convex decompositions* of ρ if, for every set of states $\{\sigma_j\} \subseteq \mathcal{P}$ such that $\rho = \sum_j p_j \sigma_j \in \mathcal{P}$ for some coefficients p_j satisfying $0 \leq p_j \leq 1$ and $\sum_j p_j = 1$, for all possible choices of $\mu_j \in \Delta_{\sigma_j}$, the measure $\sum_j p_j \mu_j$ is in Δ_ρ. The ontological model *respects convexity* if it respects convex decompositions of every $\rho \in \mathcal{P}$.

For the purpose of connecting non maximally ψ-epistemic models with preparation contextuality, recall that an ontological model $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ of a PM fragment $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ is non maximally ψ-epistemic iff there are two nonorthogonal states $[\psi], [\phi] \in \mathcal{P}$ and a measurement $M \in \mathcal{M}$ that contains $[\phi]$, such that, for some $\mu \in \Delta_\psi, \nu \in \Delta_\phi$, there exists an $\Omega \in \Sigma$ such that $\nu(\Omega) = 1$ but

$$\int_\Omega \xi_\phi^M(\lambda) d\mu(\lambda) < \int_\Lambda \xi_\phi^M(\lambda) d\mu(\lambda). \tag{72}$$

Theorem 5.13. *Let $\mathfrak{F} = \langle \mathbb{C}^d, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment and let $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ be an ontological model of it that reproduces the quantum predictions, and is not maximally ψ-epistemic.*

For some $[\psi], [\phi] \in \mathcal{P}$ and $M \in \mathcal{M}$ containing $[\phi]$ that satisfy Eq. (72), assume that $\sigma^{\psi^\perp} = \frac{1}{d-1}(I - [\psi])$, $\sigma^{\phi^\perp} = \frac{1}{d-1}(I - [\phi])$ and the maximally mixed state I/d are also in $\mathcal{P}$.

Then, if Ξ respects convex decompositions of I/d, I/d is preparation contextual in Ξ.

Proof. The maximally mixed state has convex decompositions

$$\begin{aligned} \frac{I}{d} &= \frac{1}{d}[\psi] + \frac{d-1}{d}\sigma^{\psi^\perp} \\ &= \frac{1}{d}[\phi] + \frac{d-1}{d}\sigma^{\phi^\perp}. \end{aligned}$$

Since Ξ respects convex decompositions of I/d, for every $\mu \in \Delta_\psi$ and $\mu^\perp \in \Delta_{\sigma^{\psi^\perp}}$, the measure $\frac{1}{d}\mu + \frac{d-1}{d}\mu^\perp$ is in $\Delta_{I/d}$. Similarly, for every $\nu \in \Delta_\phi$ and $\nu^\perp \in \Delta_{\sigma^{\phi^\perp}}$, the measure $\frac{1}{d}\nu + \frac{d-1}{d}\nu^\perp$ is in $\Delta_{I/d}$. Assume that I/d is preparation noncontextual in Ξ, so that $\Delta_{I/d}$ only contains one probability measure ζ. Then, we have

$$\begin{aligned} \zeta &= \frac{1}{d}\mu + \frac{d-1}{d}\mu^\perp \\ &= \frac{1}{d}\nu + \frac{d-1}{d}\nu^\perp. \end{aligned}$$

Now, since $[\psi]$, $[\phi]$ and M satisfy Eq. (72), and the model reproduces the quantum predictions, there must exist a $\mu \in \Delta_\psi$, a $\nu \in \Delta_\phi$, and a set $\Omega \in \Sigma$ such that $\nu(\Omega) = 1$, but

$$\int_\Omega \xi_\phi^M(\lambda) d\mu(\lambda) < \mathrm{Tr}\left([\phi][\psi]\right). \tag{73}$$

This means that, in order to reproduce the quantum predictions, there must be a set $\Omega' \subseteq \Lambda \backslash \Omega$ such that ξ_ϕ^M is nonzero everywhere on Ω' and $\mu(\Omega') > 0$. It is also the case that $\nu(\Omega') = 0$ because Ω' is a subset of $\Lambda \backslash \Omega$.

Again, since Ξ reproduces the quantum predictions, we must have

$$\int_\Omega \xi_\phi^M(\lambda) d\nu^\perp(\lambda) = \mathrm{Tr}\left([\phi]\sigma^{\phi^\perp}\right) = 0. \tag{74}$$

This implies that $\nu^\perp(\Omega') = 0$ because ξ_ϕ^M is nonzero everywhere on Ω'. Therefore we have

$$\begin{aligned} \zeta(\Omega') &= \frac{1}{d}\nu(\Omega') + \frac{d-1}{d}\nu^\perp(\Omega') \\ &= 0, \end{aligned}$$

but also

$$\begin{aligned} \zeta(\Omega') &= \frac{1}{d}\mu(\Omega') + \frac{d-1}{d}\mu^\perp(\Omega') \\ &\geq \frac{1}{d}\mu(\Omega') \\ &> 0, \end{aligned}$$

which is a contradiction. Therefore, I/d must be preparation contextual in Ξ. □

Similar arguments can be used to establish the preparation contextuality of other mixed states. If σ and τ are states such that $\mathrm{Tr}([\phi]\sigma) = 0$ and

$$p[\psi] + (1-p)\tau = q[\phi] + (1-q)\sigma, \tag{75}$$

for some $0 \leq p, q \leq 1$, then the state $\rho = p\,[\psi] + (1-p)\tau$ must be preparation contextual by the same argument given above. If the PM fragment under consideration consists of all states and measurements on some Hilbert space, and if Eq. (61) fails for every pair of pure states, then this is enough to show that every mixed state must be preparation contextual. However, it has not yet been shown that Eq. (61) must fail for all pairs of pure states.

5.5 Bell's Theorem

In the previous section, we established that, if a model is non maximally ψ-epistemic, then the maximally mixed state must be preparation contextual, provided the fragment contains the states $\sigma^{\psi^\perp} = \frac{1}{d-1}\,(I - [\psi])$ and $\sigma^{\phi^\perp} = \frac{1}{d-1}\,(I - [\phi])$ orthogonal to two states $[\psi]$ and $[\phi]$ that fail to satisfy the condition of maximal ψ-epistemicity. This section shows that this type of preparation contextuality can be used to prove Bell's Theorem. The argument is based on an idea due to Barrett [105] that has appeared in the literature in an informal version [55]. Due to the chain of implications we have established so far, this shows that either ψ-ontology or KS contextuality is sufficient to prove Bell's Theorem.

The Einstein-Podolsky-Rosen argument [106] establishes that ψ-complete models are nonlocal, but a ψ-ontic model may involve other variables in addition to the wavefunction. That nonlocality follows from the weaker assumption of ψ-ontology is implicit in an argument that Einstein made in correspondence with Schrödinger. This was first pointed out by Harrigan and Spekkens in [68], where they also provide a more formal version of the argument. The proof provided here follows a similar line of argument, but employs the weaker assumption of preparation contextuality for the maximally mixed state described above.

The basic idea runs as follows. Suppose Alice and Bob share a composite system with Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^d$, prepared in the entangled state $[\Phi^+]_{AB}$, where

$$\left|\Phi^+\right\rangle_{AB} = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} |j\rangle_A \otimes |j\rangle_B\,, \tag{76}$$

and $\{|j\rangle\}_{j=0}^{d-1}$ is an orthonormal basis for $\mathbb{C}^d$. If Alice performs the measurement $\left\{[\psi]_A^T, I_A - [\psi]_A^T\right\}$, where ^{T} denotes transpose in the $|j\rangle$ basis and I_A is the identity operator on $\mathcal{H}_A$, then, with probability $1/d$ she obtains the $[\psi]_A^T$ outcome and the state of Bob's system gets updated to $[\psi]_B$, and with probability $(d-1)/d$ she gets the $I_A - [\psi]_A^T$ outcome and the state of Bob's system gets updated to $\sigma_B^{\psi^\perp}$. If we allow Alice to postselect on her measurement outcome, then we can regard this as a method of preparing Bob's system in the states $[\psi]_B$ and $\sigma_B^{\psi^\perp}$. If she subsequently forgets and discards her measurement outcome then we can also regard this as a method of preparing Bob's system in the maximally mixed state by mixing $[\psi]_B$ with $\sigma_B^{\psi^\perp}$. Similar remarks apply if Alice measures $\{[\phi]_A^T, I_A - [\phi]_A^T\}$, which, by postselection, prepares Bob's system in the state $[\phi]_B$ or $\sigma_B^{\phi^\perp}$ and, upon forgetting the measurement outcome, prepares it in the maximally mixed state by mixing $[\phi]_B$ with $\sigma_B^{\phi^\perp}$.

Now, given an ontological model for the composite system that satisfies Bell's locality condition, we can use it to derive an ontological model for these preparation procedures on Bob's system alone. This model must be preparation noncontextual for the state I/d because locality implies that the probability of getting a particular ontic state for the derived model on Bob's system cannot depend on whether Alice chooses to measure $\left\{[\psi]^T, I - [\psi]^T\right\}$ or $\left\{[\phi]^T, I - [\phi]^T\right\}$. However, in order to construct a model for all possible preparations and measurements of this form that reproduces the quantum predictions, preparation noncontextuality must fail for some such pair because a maximally ψ-epistemic model is impossible (either because we are assuming ψ-ontology or because of the Kochen-Specker Theorem). Therefore, Bell's locality condition must fail. The remainder of this section formalizes this argument.

The type of fragment relevant to Bell's Theorem involves states on a bipartite system and local measurements. This is formalized in the concept of a product measurement fragment.

Definition 5.14. A *product measurement fragment* $\mathfrak{F}_{AB} = \langle \mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{P}_{AB}, \mathcal{M}_A \times \mathcal{M}_B \rangle$ is a PM fragment on a composite system with Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, where $\mathcal{M}_A$ is a set of POVMs on $\mathcal{H}_A$, $\mathcal{M}_B$ is a set of POVMs on $\mathcal{H}_B$ and $\mathcal{M}_A \times \mathcal{M}_B$ denotes the set of all POVMs of the form $(M_A, M_B) = \{E_j \otimes F_k\}$, where $M_A = \{E_j\} \in \mathcal{M}_A$ and $M_B = \{F_k\} \in \mathcal{M}_B$.

To prove Bell's Theorem, we only need consider a single bipartite state ρ_{AB}, so for the remainder of this section we assume that $\mathcal{P}_{AB}$ just contains a single state ρ_{AB}. Most of the concepts can be generalized beyond this in an obvious way.

The first step is to reinterpret experiments on the composite system as prepare-and-measure experiments on Bob's system alone. At the operational level, when the system is prepared in the state ρ_{AB}, Alice makes a measurement M_A on system A and Bob makes a measurement M_B on system B, quantum theory predicts the outcome probabilities

$$\text{Prob}(E, F|M_A, M_B) = \text{Tr}_{AB}\left(E \otimes F \rho_{AB}\right), \tag{77}$$

where $E \in M_A$ and $F \in M_B$.

These probabilities can be rewritten in terms of states and measurements on $\mathcal{H}_B$ alone by imagining that Alice performs her measurement before Bob performs his, and considering how the state of Bob's system gets updated in light of Alice's measurement result. Upon learning that the outcome of M_A was E, the state of Bob's system gets updated to

$$\rho_{B|E} = \frac{\text{Tr}_A\left(E \otimes I_B \rho_{AB}\right)}{\text{Prob}(E|M_A)}, \tag{78}$$

where $\text{Prob}(E|M_A) = \text{Tr}_{AB}\left(E \otimes I_B \rho_{AB}\right)$ is the probability of Alice's measurement outcome. Eq. (77) can then be rewritten as

$$\text{Prob}(E, F|M_A, M_B) = \text{Prob}(E|M_A)\text{Tr}_B\left(F\rho_{B|E}\right). \tag{79}$$

By postselecting on obtaining the outcome E, Alice's measurement can be viewed as a method of preparing Bob's system in the state $\rho_{B|E}$. In doing so, E gets reinterpreted as specifying a preparation procedure for system B, instead of a measurement on A, so we now want to condition the probabilities on E. This gives

$$\text{Prob}(F|M_A, E, M_B) = \frac{\text{Prob}(E, F|M_A, M_B)}{\text{Prob}(E|M_A)} = \text{Tr}_B\left(F\rho_{B|E}\right). \tag{80}$$

We also want to consider the case where Alice forgets her measurement outcome outcome. This prepares Bob's system in the state $\rho_B = \text{Tr}_A\left(\rho_{AB}\right)$ and we have, for every $M_A \in \mathcal{M}_A$,

$$\rho_B = \sum_{E \in M_A} \text{Prob}(E|M_A)\rho_{B|E}. \tag{81}$$

We can now define the conditional fragment on Bob's system consisting of all the states that Alice can prepare on his system in this way.

Definition 5.15. Let $\mathfrak{F}_{AB} = \langle \mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{P}_{AB}, \mathcal{M}_A \times \mathcal{M}_B \rangle$ be a product measurement fragment, where $\mathcal{P}_{AB}$ contains a single state ρ_{AB}. The *conditional* fragment on B given A is $\mathfrak{F}_{B|A} = \langle \mathcal{H}_B, \mathcal{P}_{B|A}, \mathcal{M}_B \rangle$, where $\mathcal{P}_{B|A}$ consists of all states of the form $\rho_{B|E}$ as given by Eq. (78) where $E \in M_A$ for some $M_A \in \mathcal{M}_A$, as well as the state $\rho_B = \text{Tr}_A\left(\rho_{AB}\right)$.

Given an ontological model of $\mathfrak{F}_{AB}$, it is natural to think that an ontological model for the conditional PM fragment $\mathfrak{F}_{B|A}$ might be derivable from it. However, in general, an ontological model for $\mathfrak{F}_{AB}$ cannot be cleanly separated into a part that depends on system A and a part that depends on system B unless it satisfies Bell's condition of local causality. To see this, suppose we have an ontological model $\Xi_{AB} = (\Lambda_{AB}, \Sigma_{AB}, \Delta_{AB}, \xi_{AB})$ of $\mathfrak{F}_{AB}$. If Alice chooses to make the measurement M_A and Bob chooses to make the measurement M_B then in general, not assuming locality, this is represented by a set of response functions $\xi_{AB}(M_A, M_B, E \otimes F)(\lambda) = \xi_{E\otimes F}^{M_A, M_B}(\lambda)$, where $E \in M_A$ and $F \in M_B$. The model then predicts that the observed probability of obtaining the outcomes E and F is given by

$$P(E, F|M_A, M_B) = \int_\Lambda \xi_{E\otimes F}^{M_A, M_B}(\lambda) d\mu(\lambda), \tag{82}$$

where $\mu \in \Delta_{\rho_{AB}}$.

At the operational level, the conditional probability $\text{Prob}(F|M_A, E, M_B)$ could be rewritten in terms of a state on $\mathcal{H}_B$ that depends on E alone and the measurement operator F, which is also defined on $\mathcal{H}_B$. The analogue at the ontological level would be to absorb the dependence on M_A and E into the measure μ and be left with a response function that only depends on M_B and F. Unfortunately, the response function $\xi_{E\otimes F}^{M_A,M_B}(\lambda)$ need not factor cleanly into a term that only depends on M_A and E and a term that only depends on M_B and F. In order to obtain such a decomposition, a locality assumption is required.

Definition 5.16. An ontological model $\Xi_{AB} = (\Lambda_{AB}, \Sigma_{AB}, \Delta_{AB}, \xi_{AB})$ of a product measurement fragment $\mathfrak{F}_{AB} = \langle \mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{P}_{AB}, \mathcal{M}_A \times \mathcal{M}_B \rangle$ is *Bell local* if, for all $M_A \in \mathcal{M}_A$, $E \in M_A$, $M_B \in \mathcal{M}_B$, $F \in M_B$, and $\lambda \in \Lambda_{AB}$,

$$\xi_{AB}(M_A, M_B, E \otimes F)(\lambda) = \xi_A(M_A, E)(\lambda)\xi_B(M_B, F)(\lambda), \tag{83}$$

where ξ_A is a map from pairs $M_A \in \mathcal{M}_A$, $E \in M_A$ to response functions $\xi_A(M_A, E) = \xi_E^{M_A}$ on Λ_{AB} satisfying $\sum_{E\in M_A} \xi_E^{M_A}(\lambda) = 1$ for all $M_A \in \mathcal{M}_A$, $\lambda \in \Lambda_{AB}$, and similarly for ξ_B.

The map ξ_{AB} specifying response functions in a Bell local model is denoted $\xi_{AB} = \xi_A \times \xi_B$.

Due to differences in notation, it may not be immediately obvious that this is the locality condition that is usually used to prove Bell's Theorem. Recall that, in general, the interpretation of a response function $\xi(M, E)(\lambda) = \xi_E^M(\lambda)$ is that it is the conditional probability $P(E|M, \lambda)$ of obtaining the outcome E given that measurement M is performed and the ontic state is λ. In these terms, Eq. (83) can be rewritten as

$$P(E, F|M_A, M_B, \lambda) = P(E|M_A, \lambda)P(F|M_B, \lambda), \tag{84}$$

which is a more familiar way of writing Bell locality.

To derive Bell's Theorem, there is one further issue to deal with. Usually, it is assumed that there is an unique probability measure corresponding to the state ρ_{AB}, so that the observed probabilities are given by

$$P(E, F|M_A, M_B) = \int_\Lambda \xi_E^{M_A}(\lambda)\xi_F^{M_B}(\lambda)d\mu(\lambda), \tag{85}$$

for some fixed measure μ. In contrast, here we are allowing for preparation contextuality, so there may be more than one measure in $\Delta_{\rho_{AB}}$, any of which could potentially appear in this equation. To facilitate an easier proof, we restrict attention to a particular method of preparing ρ_{AB} that always results in the same measure μ and build an ontological model for the conditional fragment $\mathfrak{F}_{B|A}$ using just μ rather than all of the measures in $\Delta_{\rho_{AB}}$. This does not weaken the conclusion because the result holds regardless of which $\mu \in \Delta_{\rho_{AB}}$ is chosen.

Given a measure $\mu \in \Delta_{\rho_{AB}}$ in a Bell local model of $\mathfrak{F}_{AB}$, we can define measures conditional on M_A and $E \in M_A$ as follows,

$$\mu_{M_A,E}(\Omega) = \frac{1}{P(E|M_A)} \int_\Omega \xi_E^{M_A}(\lambda)d\mu(\lambda), \tag{86}$$

where $P(E|M_A) = \int_\Lambda \xi_E^{M_A}(\lambda)d\mu(\lambda)$ is the observed probability of obtaining the outcome E for the measurement M_A. Eq. (85) can then be rewritten as

$$P(E, F|M_A, M_B,) = P(E|M_A) \int_\Lambda \xi_F^{M_B}(\lambda)d\mu_{M_A,E}(\lambda), \tag{87}$$

and conditioning on E gives

$$P(F|M_A, E, M_B) = \int_\Lambda \xi_F^{M_B}(\lambda)d\mu_{M_A,E}(\lambda), \tag{88}$$

which is the analogue of Eq. (80) at the ontological level.

Finally, if Alice does not make a measurement, or forgets her result, then we have

$$P(F|M_A, M_B) = \int_\Lambda \xi_F^{M_B}(\lambda)d\mu(\lambda). \tag{89}$$

Thus, an ontological model for the conditional fragment can be defined as follows.

Definition 5.17. Let $\mathfrak{F}_{AB} = \langle \mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{P}_{AB}, \mathcal{M}_A \times \mathcal{M}_B \rangle$ be a product measurement fragment where $\mathcal{P}_{AB}$ contains a single state ρ_{AB} and let $\Xi_{AB} = \langle \Lambda_{AB}, \Sigma_{AB}, \Delta_{AB}, \xi_A \times \xi_B \rangle$ be a Bell local ontological model of it. The *conditional ontological model* of the conditional fragment $\mathfrak{F}_{B|A}$ induced by $\mu \in \Delta_{AB}(\rho_{AB})$ is $\Xi_{B|A} = (\Lambda_{AB}, \Sigma_{AB}, \Delta_{B|A}, \xi_B)$, where, for each $M_A \in \mathcal{M}_A$, $E \in M_A$, $\Delta_{\rho_{B|E}}$ contains the measure $\mu_{M_A,E}$ defined in Eq. (86), and $\Delta_{\rho_B} = \{\mu\}$.

It is straightforward to check that if Ξ_{AB} reproduces the quantum predictions of $\mathfrak{F}_{AB}$ then $\Xi_{B|A}$ reproduces the quantum predictions of $\mathfrak{F}_{B|A}$

We now specialize to the fragment of interest for proving Bell's Theorem. This is $\mathfrak{F}_{AB} = \langle \mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{P}_{AB}, \mathcal{M}_A \times \mathcal{M}_B \rangle$, where $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^d$, $\mathcal{P}_{AB} = \{[\Phi^+]_{AB}\}$, $\mathcal{M}_A$ consists of all projective measurements of the form $M_A = \{E_0, E_1\}$ where $E_0 = [\psi]$, $E_1 = I_A - [\psi]$ for some $|\psi\rangle \in \mathbb{C}^d$ and $\mathcal{M}_B$ consists of all measurements on $\mathcal{H}_B$. Given the available measurements in $\mathcal{M}_A$, $\rho_{B|E_0}$ is always a pure state and in fact Alice can prepare Bob's system in any pure state by choosing an appropriate E_0, i.e. $[\psi]_A^T$ if she wants to prepare $[\psi]_B$.

Theorem 5.18. *Let $\Xi_{AB} = (\Lambda_{AB}, \Sigma_{AB}, \Delta_{AB}, \xi_A \times \xi_B)$ be a Bell local model for the fragment defined above that reproduces the quantum predictions for Alice's measurements, i.e. $\int_\Lambda \xi_E^{M_A}(\lambda) d\mu(\lambda) = Tr\left(E \otimes I_B [\Phi^+]_{AB}\right)$. Let $\mathfrak{F}_{B|A}$ be the corresponding conditional fragment and let $\Xi_{B|A}$ be the conditional ontological model induced by some $\mu \in \Delta_{AB}([\Phi^+]_{AB})$. Then, $\Xi_{B|A}$ is preparation noncontextual and respects convex decompositions of the maximally mixed state $I_B/d = Tr_A\left([\Phi^+]_{AB}\right)$*

Proof. Given the choice of $[\Phi^+]_{AB}$, there is only one POVM $M_A = \{E_0, E_1\}$ that can lead to a given state $\rho_{B|E_j}$. The possible conditional states on Bob's side are $\rho_{B|E_0} = [\psi]_B$ and $\rho_{B|E_1} = \sigma_B^{\psi^\perp} = \frac{1}{d-1}(I_B - [\psi]_B)$, which occur when Alice makes the measurement $\{[\psi]_A^T, I_A - [\psi]_A^T\}$, and so each of Alice's measurements determine a unique pair of conditional states for Bob. Therefore, each $\Delta_{\rho_{B|E_j}}$ contains only a single measure, so the model is preparation noncontextual for these states. Similarly, by construction, only one measure was placed in $\Delta_{I/d}$.

The only convex combinations that appear in $\mathfrak{F}_{B|A}$ are those of the form

$$\frac{I_B}{d} = \sum_{j=0}^{1} \text{Prob}(E_j|M_A)\rho_{B|E_j}, \tag{90}$$

where $\text{Prob}(E_j|M_A) = \text{Tr}_{AB}\left(E_j \otimes I_B [\Phi^+]_{AB}\right)$. Therefore, the result follows if

$$\mu = \sum_{j=0}^{1} \text{Prob}(E_j|M_A)\mu_{M_A,E_j}. \tag{91}$$

To see this, note that $\sum_j \xi_{E_j}^{M_A}(\lambda) = 1$ for all $\lambda \in \Lambda_{AB}$, so, for $\Omega \in \Sigma_{AB}$,

$$\begin{aligned}
\mu(\Omega) &= \int_\Omega d\mu(\lambda) && (92)\\
&= \int_\Omega \left(\sum_j \xi_{E_j}^{M_A}(\lambda) \right) d\mu(\lambda) && (93)\\
&= \sum_j \int_\Omega \xi_{E_j}^{M_A}(\lambda) d\mu(\lambda) && (94)\\
&= \sum_j P(E_j|M_A) \frac{1}{P(E_j|M_A)} \int_\Omega \xi_{E_j}^{M_A}(\lambda) d\mu(\lambda) && (95)\\
&= \sum_j P(E_j|M_A) \mu_{M_A,E_j}(\Omega), && (96)
\end{aligned}$$

where $P(E_j|M_A) = \int_\Lambda \xi_{E_j}^{M_A}(\lambda) d\mu(\lambda)$ is the probability of Alice obtaining outcome E_j when measuring M_A according to the ontological model Ξ. By assumption, the model reproduces the quantum predictions for Alice's measurements, so $P(E_j|M_A) = \text{Prob}(E_j|M_A)$, from which the result follows. □

Corollary 5.19. *For $d \geq 3$, no Bell local ontological model of the fragment $\mathfrak{F}_{AB}$ can reproduce the quantum predictions.*

Proof. By Definition 5.17, an ontological model $\Xi_{B|A}$ of the conditional fragment $\mathfrak{F}_{B|A}$ that reproduces the quantum predictions can be derived from a Bell local model Ξ_{AB} that reproduces the quantum predictions of $\mathfrak{F}_{AB}$. By Theorem 5.18, $\Xi_{B|A}$ respects convex decompositions of I_B/d and is preparation noncontextual. However, the fragment $\mathfrak{F}_{B|A}$ contains all pure states and all measurements on $\mathbb{C}^d$. Thus, by the Kochen-Specker Theorem, the model must be KS contextual for $d \geq 3$, which, by Corollary 5.10, implies that it cannot be maximally ψ-epistemic, which, by Theorem 5.13, implies that I_B/d must be preparation contextual in $\Xi_{B|A}$. Thus, by contradiction, there can be no model Ξ_{AB} that is both Bell local and reproduces the quantum predictions. □

Obviously, this is a rather convoluted way of proving Bell's Theorem. However, note that if it can be shown, under suitable assumptions, that all ontological models must be ψ-ontic, then we could replace the implication from the Kochen-Specker Theorem with the implication from ψ-ontology. Thus, establishing ψ-ontology would imply Bell's Theorem as a corollary.

Part II
ψ-ontology theorems

Having understood the implications of ψ-ontology, it is finally time to look at the extent to which it can be established. The first ψ-ontology theorem that was discovered is due to Pusey, Barrett and Rudolph (PBR). Of the available results, it is the most widely discussed in the literature and I think it makes the strongest case for ψ-ontology, so it receives the longest treatment in §7, including a discussion of generalizations and criticisms. Following this, I discuss two further ψ-ontology theorems: Hardy's Theorem in §9, and the Colbeck-Renner Theorem in §10.

Before getting to these theorems, I discuss the concept of *antidistinguishability* in §6 because it plays a key role in both the PBR Theorem and Hardy's Theorem, as well as some of the results discussed in Part III. Both Hardy's Theorem and the Colbeck-Renner Theorem involve assumptions about how dynamics are represented in ontological models, so this is discussed in §8 after the discussion of PBR is completed.

6 Antidistinguishability

According to Theorem 4.10, sets of distinguishable states are pairwise ontologically distinct in an ontological model. However, ψ-ontology requires that all pairs of pure states must be ontologically distinct, and most pairs of pure states are not orthogonal. Therefore, to prove ψ-ontology, it is useful to consider a weaker concept.

Definition 6.1. Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment. A set of quantum states $\{\rho_j\} \subseteq \mathcal{P}$ is *antidistinguishable* in $\mathfrak{F}$ if there exists a POVM $\{E_j\} \in \mathcal{M}$ such that

$$\text{Tr}\left(E_j \rho_j\right) = 0. \tag{97}$$

Recall that the definition of distinguishability states that there should exist a POVM $\{E_j\}$ such that $\text{Tr}\left(E_j \rho_k\right) = \delta_{jk}$ for all j. This is equivalent to requiring $\text{Tr}\left(E_j \rho_j\right) = 1$ for all j by the following argument. Firstly, $\text{Tr}\left(E_j \rho_j\right) = 1$ is obviously a special case of $\text{Tr}\left(E_j \rho_k\right) = \delta_{jk}$. For the converse direction note that if there were a $k \neq j$ such that $\text{Tr}\left(E_j \rho_k\right) > 0$ then $\text{Tr}\left(\left(E_j + E_k\right) \rho_k\right) > 1$, but this cannot be the case if

ρ_k is a density operator and $\{E_m\}$ is a POVM because then we must have $1 = \mathrm{Tr}\,(\rho_k) = \mathrm{Tr}\,(\sum_m E_m \rho_k) \geq \mathrm{Tr}\,((E_j + E_k)\,\rho_k)$. Therefore, the difference between distinguishability and antidistinguishability is simply the replacement of $\mathrm{Tr}\,(E_j \rho_j) = 1$ with $\mathrm{Tr}\,(E_j \rho_j) = 0$. To understand what this means, suppose that one of the states in a set $\mathcal{D}$ is prepared and you do not know which. If $\mathcal{D}$ is distinguishable then there is a measurement for which each outcome identifies that a specific member of $\mathcal{D}$ was definitely the one prepared. On the other hand, if $\mathcal{D}$ is antidistinguishable then there is a measurement for which each outcome identifies that a specific member of $\mathcal{D}$ was definitely not the one that was prepared. In this sense, the two concepts are opposites.

Antidistinguishability is a weaker property than distinguishability because, if the measurement outcome j tells us that ρ_j was definitely prepared, then it also tells us that ρ_k was definitely not prepared for all $k \neq j$. Therefore, a distinguishing POVM can be converted to an antidistinguishing one just by permuting the outcome indices in such a way that none of them are left invariant, e.g. if there are n POVM elements then this can be done by mapping $j \to j + 1 \mod n$. For sets of three or more states, antidistinguishability is strictly weaker than distinguishability, as there exist sets of nonorthogonal states that can be antidistinguished (see Fig. 11 for an example). For a set of two states, antidistinguishability is equivalent to distinguishability because if you know that one of the states was not the one prepared then you know that it must have been the other one.

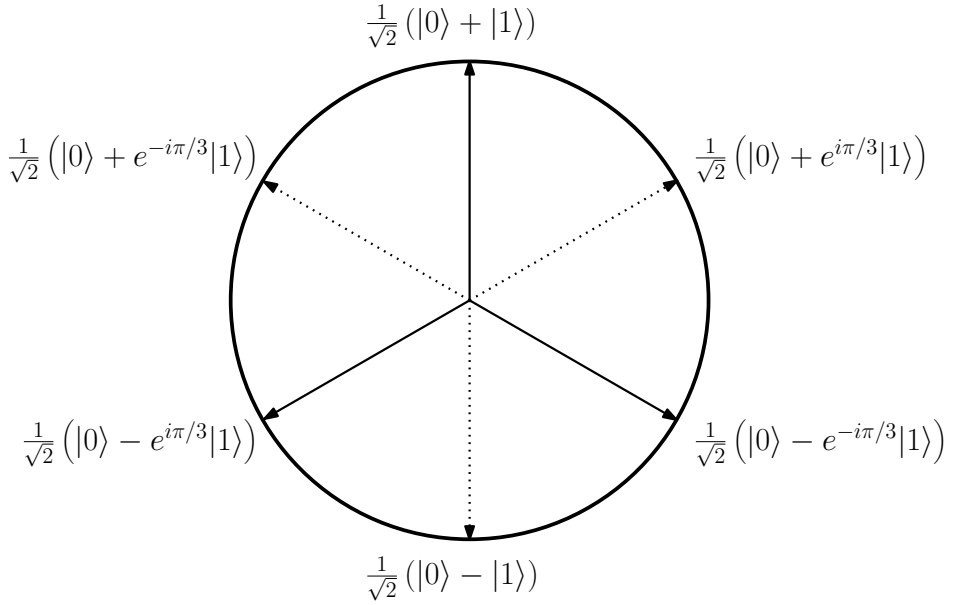


Figure 11: A set of nonorthogonal states that can be antidistinguished, and the POVM that antidistinguishes them. The diagram represents the equator of the Bloch sphere and the three antidistinguishable states are shown by the solid arrows. The dotted arrows show three states that are each orthogonal to one of the states in the antidistinguishable set. A POVM can be formed by scaling the projectors onto the three dotted states by a factor of 2/3. By virtue of the orthogonality relations, this is an antidistinguishing POVM.

Like distinguishability, antidistinguishability has consequences for the overlaps of probability measures in operational models. However, whilst distinguishability constrains the pairwise overlaps, antidistinguishability only constrains the n-way overlaps, where n is the number of states in the antidistinguishable set. Since the variational distance only applies to two measures, a more general notion of overlap is needed to capture this.

Definition 6.2. A *countable partition* $\{\Omega_k\}$ of a measurable space (Λ, Σ) is an at most countable set of disjoint subsets $\Omega_k \in \Sigma$, $\Omega_k \cap \Omega_r = \emptyset$ for $k \neq r$, such that $\cup_k \Omega_k = \Lambda$. $T(\Lambda)$ denotes the set of countable partitions of (Λ, Σ).

Definition 6.3. The *overlap* of an at most countable set $\{\mu_j\}$ of probability measures on a sample space (Λ, Σ) is given by

$$L\left(\{\mu_j\}\right) = \inf_{\{\Omega_k\}\in T(\Lambda)} \left(\sum_k \min_j \left(\mu_j(\Omega_k)\right) \right). \tag{98}$$

It can be shown that if m is a measure that dominates $\{\mu_j\}$ then the overlap is given by

$$L\left(\{\mu_j\}\right) = \int_\Lambda \min_j \left(p_j(\lambda)\right) dm(\lambda), \tag{99}$$

where p_j is a density that represents μ_j with respect to m, i.e. $\mu_j(\Omega) = \int_\Omega p_j(\lambda) dm(\lambda)$. For a finite set of measures, a dominating measure always exists. This form of the overlap is often most useful for calculations.

Note that, if $\{\mu_j\}$ contains n measures then the optimization in Eq. (98) can be restricted to partitions that contain n subsets. This is because, if a partition contains $m > n$ subsets, then the same μ_j is bound to be the minimum for more than one Ω_k and these sets can then be replaced by their union without affecting the result. Therefore, the overlap is given by

$$L(\{\mu_j\}) = \min_{\{\Omega_k\}\in T_n(\Lambda)} \left(\sum_{k=1}^{n} \min_j \left(\mu_j(\Omega_k)\right) \right), \tag{100}$$

where $T_n(\Lambda)$ is the set of n-fold partitions of Λ. In fact, this can be simplified further to

$$L(\{\mu_j\}) = \min_{\{\Omega_j\}\in T_n(\Lambda)} \left(\sum_j \mu_j(\Omega_j) \right), \tag{101}$$

because the minimizing partition can always be relabelled such that k gets mapped to the j that minimizes $\mu_j(\Omega_k)$.

The overlap has the following operational interpretation. Suppose that a system is prepared according to one of n preparation procedures P_j, where P_j corresponds to the measure μ_j, each case having equal a priori probability $1/n$. You are then told the actual value of λ and your task is to announce a j between 1 and n such that P_j was *not* the procedure used to prepare the system, i.e. you want to antidistinguish the probability measures. If you adopt a deterministic strategy then the best you can do is to partition the space into n subsets Ω_j and announce that P_j was not the preparation procedure used when $\lambda \in \Omega_j$. The probability of failure when using this strategy is $\frac{1}{n}\sum_{j=1}^{n} \mu_j(\Omega_j)$ so the minimal probability of failure is $\frac{1}{n} L(\{\mu_j\})$. By convexity, a probabilistic strategy cannot do any better than this. Therefore, $L(\{\mu_j\}) = 0$ corresponds to the case where the probability measures can be antidistinguished perfectly.

The overlap generalizes the variational distance in the following sense.

Theorem 6.4. *For two probability measures μ, ν on (Λ, Σ), $L(\mu, \nu) = 1 - D(\mu, \nu)$, where D is the variational distance.*

Proof. From Eq. (101),

$$\begin{aligned} L(\mu,\nu) &= \inf_{\Omega\in\Sigma} \left(\mu(\Omega) + \nu(\Lambda\backslash\Omega)\right) && (102)\\ &= \inf_{\Omega\in\Sigma} \left(\mu(\Omega) + 1 - \nu(\Omega)\right) && (103)\\ &= 1 - \sup_{\Omega\in\Sigma} \left(\nu(\Omega) - \mu(\Omega)\right) && (104)\\ &= 1 - \sup_{\Omega\in\Sigma} \left|\mu(\Omega) - \nu(\Omega)\right| && (105)\\ &= 1 - D(\mu,\nu). && (106) \end{aligned}$$

□

Because of this, the condition for two quantum states ρ and σ to be ontologically distinct in an ontological model $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ can be restated as: for all $\mu \in \Delta_\rho$ and $\nu \in \Delta_\sigma$, $L(\mu, \nu) = 0$.

The reason for introducing the more general n-way overlaps is that, being a weaker concept than distinguishability, antidistinguishability only places constraints on the n-way overlaps rather than on pairwise overlaps. To understand the intuition behind these constraints, consider the case of a finite ontic state space and suppose that there is an ontic state λ that is assigned nonzero probability by every measure μ_j in a set of probability measures $\{\mu_j\}_{j=1}^n$ corresponding to an antidistinguishable set $\{\rho_j\}_{j=1}^n$. Then, λ must assign zero probability to each of the outcomes of the antidistinguishing measurement. However, the probabilities that λ assigns to the measurement outcomes must form a probability distribution, so they cannot all be zero. Hence, no such λ can exist.

In order to formulate this argument rigorously, we have to deal with measure zero sets. For this reason, it is better to formulate the argument in terms of the overlap.

Theorem 6.5. *Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a prepare-and-measure fragment and suppose that $\{\rho_j\} \subseteq \mathcal{P}$ is antidistinguishable in $\mathfrak{F}$. Then, in any ontological model $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ of $\mathfrak{F}$ that reproduces the quantum predictions,*

$$L(\{\mu_j\}) = 0, \tag{107}$$

for every possible choice of probability measures $\mu_j \in \Delta_{\rho_j}$.

Proof. Let $M = \{E_j\}$ be the antidistinguishing POVM, let $\mu_j \in \Delta_{\rho_j}$, and let m be a measure that dominates $\{\mu_j\}$ such that, for $\Omega \in \Sigma$, $\mu_j(\Omega) = \int_\Omega p_j(\lambda) dm(\lambda)$ for some densities p_j. In order to reproduce the predictions of M, the ontological model must satisfy

$$\int_\Lambda \xi_{E_j}^M(\lambda) d\mu_j(\lambda) = \int_\Lambda \xi_{E_j}^M(\lambda) p_j(\lambda) dm(\lambda) = \mathrm{Tr}\left(E_j \rho_j\right) = 0. \tag{108}$$

However,

$$\int_\Lambda \xi_{E_j}^M(\lambda) p_j(\lambda) dm(\lambda) \geq \int_\Lambda \xi_{E_j}^M(\lambda) \min_k \left(p_k(\lambda)\right) dm(\lambda), \tag{109}$$

so

$$\int_\Lambda \xi_{E_j}^M(\lambda) \min_k \left(p_k(\lambda)\right) dm(\lambda) = 0. \tag{110}$$

Now, sum both sides over j. Because, $\sum_j \xi_{E_j}^M(\lambda) = 1$ for every $\lambda \in \Lambda$, this gives

$$\int_\Lambda \min_k \left(p_k(\lambda)\right) dm(\lambda) = 0, \tag{111}$$

or in other words $L\left(\{\mu_k\}\right) = 0$. □

Finally, a few remarks are in order about terminology. Antidistinguishability was first introduced by Caves, Fuchs and Shack in their study of the compatibility of quantum state assignments, where it was called "post-Peierls (PP) incompatibility" [107]. The concept did not attract much attention until its use in the PBR Theorem, but since then a more systematic study of antidistinguishability has appeared [108], where it goes under the name "conclusive exclusion of quantum states". Therefore, if you see the terms "PP incompatibility" and "conclusive exclusion" in the literature then you can rest assured that they mean the same thing as antidistinguishability. I think "antidistinguishability" is a more suggestive name, but time will tell which terminology becomes standard.

7 The Pusey-Barrett-Rudolph Theorem

The PBR Theorem is the most prominent ψ-ontology theorem. It is based on an assumption known as the Preparation Independence Postulate (PIP), which essentially says that composite systems prepared in

a product state should be independent of one another. It is this postulate that allows bounds on n-way overlaps coming from antidistinguishability to be leveraged into the bounds on pairwise overlaps needed to establish ψ-ontology. The PIP is presented in its simplest form in §7.1 in order to facilitate the simplest possible proof of the main result in §7.2. However, the PIP is the most controversial assumption of the PBR Theorem, so its motivation is discussed in §7.3, weakenings of it that still allow the PBR Theorem to go through are discussed in §7.4, its necessity for proving the theorem is established in §7.5, and criticism of it is discussed in §7.6. Finally, other criticisms of the PBR Theorem, not directed at the PIP, are discussed in §7.7.

7.1 The Preparation Independence Postulate

The Preparation Independence Postulate (PIP) is an assumption about how composite systems should be modelled when the subsystems are prepared independently of one another. Suppose Alice prepares her system in a quantum state ρ_A and Bob prepares his system in a quantum state ρ_B, such that their joint state is a product $\rho_A \otimes \rho_B$. Suppose further that Alice has an ontological model with ontic state space (Λ_A, Σ_A) that reproduces the quantum predictions for her system in isolation and that μ_A is the probability measure that it assigns to her preparation procedure. Similarly, Bob has an ontological model for his system with ontic state space (Λ_B, Σ_B) that assigns the measure μ_B to his preparation procedure. Then, the PIP says that there ought to be an ontological model for the joint system with ontic state space $(\Lambda_A \times \Lambda_B, \Sigma_A \otimes \Sigma_B)$ in which the product measure

$$\mu_A \times \mu_B(\Omega) = \int_{\Lambda_B} \mu_A(\Omega_{\lambda_B}) d\mu_B(\lambda_B), \tag{112}$$

reproduces the quantum predictions for $\rho_A \otimes \rho_B$, where $\Omega_{\lambda_B} = \{\lambda_A \in \Lambda_A | (\lambda_A, \lambda_B) \in \Omega\}$.

This assumption looks superficially plausible, and indeed it is not much commented upon in the PBR paper. However, since it is the most controversial assumption of the PBR Theorem, it is worth pausing to define it a bit more rigorously. Discussion of the motivation for the PIP and criticism of it is deferred until §7.3–§7.6.

To start with, we need to understand how prepare-and-measure fragments on subsystems can be combined to form composites.

Definition 7.1. Let $\mathfrak{F}_A = \langle \mathcal{H}_A, \mathcal{P}_A, \mathcal{M}_B \rangle$ and $\mathfrak{F}_B = \langle \mathcal{H}_B, \mathcal{P}_B, \mathcal{M}_B \rangle$ be PM fragments. The *direct product* of $\mathfrak{F}_A$ and $\mathfrak{F}_B$ is

$$\mathfrak{F}_A \times \mathfrak{F}_B = \langle \mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{P}_A \times \mathcal{P}_B, \mathcal{M}_A \times \mathcal{M}_B \rangle, \tag{113}$$

where $\mathcal{P}_A \times \mathcal{P}_B$ consists of all product states $\rho_A \otimes \rho_B$ with $\rho_A \in \mathcal{P}_A$ and $\rho_B \in \mathcal{P}_B$, and $\mathcal{M}_A \times \mathcal{M}_B$ consists of all POVMs of the form $\{E_j^A \otimes E_k^B\}$ with $\{E_j^A\} \in \mathcal{M}_A$ and $\{E_k^B\} \in \mathcal{M}_B$.

A *general product* (or just product) of $\mathfrak{F}_A$ and $\mathfrak{F}_B$ is any PM fragment of the form

$$\mathfrak{F}_{AB} = \langle \mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{P}_{AB}, \mathcal{M}_{AB} \rangle, \tag{114}$$

where $\mathcal{P}_A \times \mathcal{P}_B \subseteq \mathcal{P}_{AB}$ and $\mathcal{M}_A \times \mathcal{M}_B \subseteq \mathcal{M}_{AB}$.

$\mathfrak{F}_A$ and $\mathfrak{F}_B$ are called the *factors* of $\mathfrak{F}_{AB}$.

The direct product fragment represents a situation in which there are two separated systems and we can only do separate prepare-and-measure experiments on them individually. It only allows product states and local measurements, so there is no possibility of entanglement or even classical correlation between the two systems. A general product fragment contains all the states and measurements in the direct product, but it may, in addition, include some extra correlated or entangled states and measurements on the joint system.

Any two ontological models $\Xi_A = (\Lambda_A, \Sigma_A, \Delta_A, \xi_A)$ and $\Xi_B = (\Lambda_B, \Sigma_B, \Delta_B, \xi_B)$ for $\mathfrak{F}_A$ and $\mathfrak{F}_B$ can be combined to form an ontological model $\Xi_{AB} = (\Lambda_{AB}, \Sigma_{AB}, \Delta_{AB}, \xi_{AB})$ for the direct product $\mathfrak{F}_A \times \mathfrak{F}_B$. To do so, set $\Lambda_{AB} = \Lambda_A \times \Lambda_B$ and $\Sigma_{AB} = \Sigma_A \otimes \Sigma_B$, let $\Delta_{AB}(\rho_A \otimes \rho_B)$ be the set of measures of the form

$\mu_{AB} = \mu_A \times \mu_B$, where $\mu_A \in \Delta_A(\rho_A)$ and $\mu_B \in \Delta_B(\rho_B)$, and set $\xi^{M_A,M_B}_{E_j \otimes E_k}(\lambda_A, \lambda_B) = \xi^{M_A}_{E_j}(\lambda_A)\xi^{M_B}_{E_k}(\lambda_B)$. Then, for $\mu_{AB} \in \Delta_{AB}(\rho_A \otimes \rho_B)$, we have

$$\int_{\Lambda_{AB}} \xi^{M_A,M_B}_{E_j \otimes E_k}(\lambda_A, \lambda_B) d\mu_{AB}(\lambda_A, \lambda_B) = \int_{\Lambda_A} \xi^{M_A}_{E_j}(\lambda_A) d\mu_A(\lambda_A) \times \int_{\Lambda_B} \xi^{M_B}_{E_k}(\lambda_B) d\mu_B(\lambda_B), \tag{115}$$

where $\mu_A \in \Delta_A(\rho_A)$ and $\mu_B \in \Delta_B(\rho_B)$. From this, it is easy to see that if Ξ_A reproduces the quantum predictions for $\mathfrak{F}_A$ and Ξ_B reproduces the quantum predictions for $\mathfrak{F}_B$, then Ξ_{AB} reproduces the quantum predictions for $\mathfrak{F}_A \times \mathfrak{F}_B$.

The key point is that, in order to model a direct product fragment, we only need to take the cartesian product of the subsystem ontic state spaces and we only need to use product measures. The PIP states that this should remain true even if the composite fragment is expanded to include entangled measurements (but still no entangled states). A little more terminology is required to state this formally.

Definition 7.2. Let $\mathfrak{F}_A = \langle \mathcal{H}_A, \mathcal{P}_A, \mathcal{M}_B \rangle$ and $\mathfrak{F}_B = \langle \mathcal{H}_B, \mathcal{P}_B, \mathcal{M}_B \rangle$ be prepare-and-measure fragments. A *product state fragment* is a product fragment $\mathfrak{F}_{AB} = \langle \mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{P}_{AB}, \mathcal{M}_{AB} \rangle$ where $\mathcal{P}_{AB} = \mathcal{P}_A \times \mathcal{P}_B$, i.e. it only includes product states, but it may have entangled measurements.

Definition 7.3. A pair of ontological models $\Xi_A = (\Lambda_A, \Sigma_A, \Delta_A, \xi_A)$, $\Xi_B = (\Lambda_B, \Sigma_B, \Delta_B, \xi_B)$ for fragments $\mathfrak{F}_A = \langle \mathcal{H}_A, \mathcal{P}_A, \mathcal{M}_A \rangle$ and $\mathfrak{F}_B = \langle \mathcal{H}_B, \mathcal{P}_B, \mathcal{M}_B \rangle$ are *compatible* with an ontological model $\Xi_{AB} = (\Lambda_{AB}, \Sigma_{AB}, \Delta_{AB}, \xi_{AB})$ for a product state fragment $\mathfrak{F}_{AB} = \langle \mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{P}_A \times \mathcal{P}_B, \mathcal{M}_{AB} \rangle$ if Ξ_{AB} satisfies

- The Cartesian Product Assumption (CPA): $\Lambda_{AB} = \Lambda_A \times \Lambda_B$, $\Sigma_{AB} = \Sigma_A \otimes \Sigma_B$.
- The No-Correlation Assumption (NCA): $\forall \rho_A \in \mathcal{P}_A, \rho_B \in \mathcal{P}_B$, $\Delta_{AB}(\rho_A \otimes \rho_B) = \Delta_A(\rho_A) \times \Delta_B(\rho_B)$ is the set of measures of the form $\mu_A \times \mu_B$, where $\mu_A \in \Delta_A(\rho_A)$ and $\mu_B \in \Delta_B(\rho_B)$. This means that $\mu_{AB} \in \Delta_{AB}(\rho_A \otimes \rho_B)$ is of the form

$$\mu_{AB}(\Omega) = \int_{\Lambda_B} \mu_A(\Omega_{\lambda_B}) d\mu_B(\lambda_B), \tag{116}$$

  where $\Omega_{\lambda_B} = \{\lambda_A \in \Lambda_A | (\lambda_A, \Lambda_B) \in \Omega\}$, $\mu_A \in \Delta_A(\rho_A)$ and $\mu_B \in \Delta_B(\rho_B)$.

The CPA says that, when modelling product state preparations, there are no additional properties of the joint system that are not derived from the properties of the individual system. In other words, there are no "genuinely nonlocal properties" that are irrelevant for local measurements but come into play when making joint measurements of the two systems. The NCA says that, further, product states can be modelled as product measures.

Definition 7.4. A pair of ontological models Ξ_A and Ξ_B for fragments $\mathfrak{F}_A$ and $\mathfrak{F}_B$ that reproduce the quantum predictions satisfy the *Preparation Independence Postulate (PIP)* if, for all product state fragments $\mathfrak{F}_{AB}$ with $\mathfrak{F}_A$ and $\mathfrak{F}_B$ as factors, there exists an ontological model Ξ_{AB} that is compatible with Ξ_A and Ξ_B and reproduces the quantum predictions.

Note that, it is very important that the PIP only applies to product state fragments and not to general product fragments. In other words, it does not have any implications for how entangled states should be represented in an ontological model. Suppose we have an ontological model $(\Lambda_{AB}, \Sigma_{AB}, \Delta_{AB}, \xi_{AB})$ for a product state fragment $\mathfrak{F}_{AB}$ that satisfies the PIP and we want to extend it to a more general product fragment $\mathfrak{F}'_{AB}$ that differs from $\mathfrak{F}_{AB}$ by the addition of some entangled states. Then, we have to allow for the fact that the extended ontological model $(\Lambda'_{AB}, \Sigma'_{AB}, \Delta'_{AB}, \xi'_{AB})$ might have extra ontic states that have zero measure whenever we prepare a product state, but nonzero measure for some entangled state. Therefore, all we can conclude about the ontic state space of the extended model from the PIP is that $\Lambda_A \times \Lambda_B \subseteq \Lambda'_{AB}$, that $\Sigma_A \otimes \Sigma_B$ is a σ-subalgebra of Σ'_{AB}, and that product state preparations have measure zero outside $\Lambda_A \times \Lambda_B$ subset, but there may be other ontic states that come into play when we prepare entangled states. Similarly, probability measures that induce correlations between Λ_A and Λ_B may come into play when we prepare entangled states.

The reason this is important is that ψ-ontic models do not satisfy $\Lambda'_{AB} = \Lambda_A \times \Lambda_B$ and it would be poor form to prove a ψ-ontology theorem only to find that ψ-ontic models are also ruled out by your assumptions. To understand this, consider the Beltrametti-Bugajski model for two Hilbert spaces $\mathcal{H}_A$, $\mathcal{H}_B$ and for the composite system $\mathcal{H}_A \otimes \mathcal{H}_B$. For the two systems individually, we have the two projective Hilbert spaces consisting of all pure states $[\psi]_A$ and $[\phi]_B$ as the ontic state spaces. If we only want to model a product state fragment then we only need states of the form $[\psi]_A \otimes [\phi]_B$, and the set of such states is isomorphic to the cartesian product of the two projective Hilbert spaces. Similarly, the point measure on $[\psi]_A \otimes [\phi]_B$ can be written as the product of a point measure at $[\psi]_A$ with a point measure at $[\psi]_B$. Therefore, the model satisfies the PIP. On the other hand, if we want to model entangled states then we need additional ontic states $[\psi]_{AB}$ that cannot be written as products, and the point measure at $[\psi]_{AB}$ cannot be written as a product of measures on the ontic state spaces of the factors. Therefore, modelling entangled states requires additional ontic states and probability measures, but this is not a violation of the PIP because the PIP only applies to product state fragments.

For ease of exposition, we have so far confined attention to composites of two subsystems. The PBR Theorem actually requires the generalization to n systems, but only in the case of n identical factors. The required generalization should be obvious, but for completeness here are the definitions.

Definition 7.5. Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment. An *n fold product state fragment* with factors $\mathfrak{F}$ is any fragment $\mathfrak{F}^n$ of the form

$$\mathfrak{F}^n = \langle \mathcal{H}^{\otimes n}, \mathcal{P}^{\times n}, \mathcal{M}' \rangle, \tag{117}$$

where $\mathcal{P}^{\times n}$ consists of states of the form $\otimes_{j=1}^n \rho_j$ for $\rho_j \in \mathcal{P}$, and $\mathcal{M}^{\times n} \subseteq \mathcal{M}'$ where $\mathcal{M}^{\times n}$ consists of all POVMs of the form $\{\otimes_{j=1}^n E_{k_j}^{(j)}\}$ for $\{E_{k_j}^{(j)}\} \in \mathcal{M}$.

Definition 7.6. An ontological model $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ of a fragment $\mathfrak{F}$ is *compatible* with an ontological model $\Xi^n = (\Lambda^n, \Sigma^n, \Delta^n, \xi^n)$ for an n-fold product state fragment $\mathfrak{F}^n$ with factors $\mathfrak{F}$ if it satisfies:

- The Cartesian Product Assumption (CPA): $\Lambda^n = \Lambda^{\times n}$, $\Sigma^n = \Sigma^{\otimes n}$.
- The No-Correlation Assumption (NCA): $\Delta^n(\otimes_{j=1}^n \rho_j)$ consists of product measures of the form $\times_{j=1}^n \mu_j$, where $\mu_j \in \Delta(\rho_j)$.

Definition 7.7. An ontological model for a fragment $\mathfrak{F}$ that reproduces the quantum predictions satisfies the PIP if it is compatible with an ontological model for any n-fold product state fragment $\mathfrak{F}^n$ with factors $\mathfrak{F}$ that reproduces the quantum predictions.

7.2 The main result

Recalling definition 4.11 of a ψ-ontic model, we are now in a position to prove the main result of the PBR paper.

Theorem 7.8 (The PBR Theorem)**.** *Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a prepare-and-measure fragment where $\mathcal{P}$ contains all pure states on $\mathcal{H}$. Then, any ontological model of $\mathfrak{F}$ that reproduces the quantum predictions and satisfies the PIP is ψ-ontic.*

The proof of this theorem is based on two lemmas. The first establishes a connection between antidistinguishability of product states and ontological distinctness, and the second establishes that the required antidistinguishability holds. Because they are quite technical, it is worth pausing to sketch a special case.

Example 7.9. Consider the two states $[z+]$ and $[x+]$, where $|x+\rangle = \frac{1}{\sqrt{2}}(|z+\rangle + |z-\rangle)$. These two states are not orthogonal, so we cannot use distinguishability to establish their ontological distinctness. Since there are only two of them, they cannot be antidistinguishable either. However, now consider two copies of the system and the four states $[z+] \otimes [z+]$, $[z+] \otimes [x+]$, $[x+] \otimes [z+]$ and $[x+] \otimes [x+]$. It turns out that these four

states are antidistinguishable. The measurement required to antidistinguish them is entangled and consists of the projectors onto the following four orthonormal vectors.

$$|\phi_1\rangle = \frac{1}{\sqrt{2}}\left(|z+\rangle\,|z-\rangle + |z-\rangle\,|z+\rangle\right) \tag{118}$$

$$|\phi_2\rangle = \frac{1}{\sqrt{2}}\left(|z+\rangle\,|x-\rangle + |z-\rangle\,|x+\rangle\right) \tag{119}$$

$$|\phi_3\rangle = \frac{1}{\sqrt{2}}\left(|x+\rangle\,|z-\rangle + |x-\rangle\,|z+\rangle\right) \tag{120}$$

$$|\phi_4\rangle = \frac{1}{\sqrt{2}}\left(|x+\rangle\,|x-\rangle + |x-\rangle\,|x+\rangle\right), \tag{121}$$

where $|x-\rangle = \frac{1}{\sqrt{2}}\left(|z+\rangle - |z-\rangle\right)$. It is easy to check that $|\phi_1\rangle$ is orthogonal to $|z+\rangle \otimes |z+\rangle$, $|\phi_2\rangle$ is orthogonal to $|z+\rangle \otimes |x+\rangle$, $|\phi_3\rangle$ is orthogonal to $|x+\rangle \otimes |z+\rangle$, and $|\phi_4\rangle$ is orthogonal to $|x+\rangle \otimes |x+\rangle$, as required.

The second lemma shows that a similar thing happens for any pair of pure states $[\psi_0], [\psi_1]$. Specifically, the set of states $\{[\psi_{k_1}] \otimes [\psi_{k_2}] \otimes \ldots \otimes [\psi_{k_n}]\}$, where $k_j \in \{0,1\}$, is always antidistinguishable for sufficiently large n.

Returning to the example, by Theorem 6.5, antidistinguishability implies that, in an ontological model, the probability measures corresponding to $[z+] \otimes [z+]$, $[z+] \otimes [x+]$, $[x+] \otimes [z+]$ and $[x+] \otimes [x+]$ can have no four way overlap. However, if the PIP is assumed, this implies that $[z+]$ and $[x+]$ must be ontologically distinct. As a rough sketch of this argument, consider the case of a finite ontic state space, let μ_{z+} and μ_{x+} be a pair of measures associated with $[z+]$ and $[x+]$ in the ontological model, and assume that there is some λ to which they both assign nonzero probability. According to the PIP, the four product states must be represented by measures of the form $\mu_{z+} \times \mu_{z+}$, $\mu_{z+} \times \mu_{x+}$, $\mu_{x+} \times \mu_{z+}$, and $\mu_{x+} \times \mu_{x+}$. Each of these assigns a nonzero probability to (λ, λ), since both μ_{z+} and μ_{x+} assign nonzero probability to λ, so there is a common ontic state to which all four measures assign nonzero probability. However, antidistinguishability implies that there can be no 4 way overlap between the measures so there is a contradiction, and hence $[z+]$ and $[x+]$ must be ontologically distinct.

The first lemma is a more formal version of this argument that applies to the states $\{[\psi_{k_1}] \otimes [\psi_{k_2}] \otimes \ldots \otimes [\psi_{k_n}]\}$ and deals with the general measure theoretic case.

Lemma 7.10. *Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment, let $\rho_0, \rho_1 \in \mathcal{P}$, and let $\mathfrak{F}^n = \langle \mathcal{H}^{\otimes n}, \mathcal{P}^{\times n}, \mathcal{M}' \rangle$ be an n-fold product state fragment. Let $\mathcal{D}$ be the set of all states of the form $\otimes_{j=1}^n \rho_{k_j}$ where $k_j \in \{0,1\}$. If $\mathcal{D}$ is antidistinguishable in $\mathfrak{F}^n$ then, in any ontological model of $\mathfrak{F}$ that reproduces the quantum predictions and satisfies the PIP, ρ_0 and ρ_1 are ontologically distinct.*

Proof. Let $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ be an ontological model for $\mathfrak{F}$ and let $\Xi^n = (\Lambda^n, \Sigma^n, \Delta^n, \xi^n)$ be a compatible ontological model for $\mathfrak{F}^n$ that reproduces the quantum predictions, which has to exist according to the PIP. Let $\boldsymbol{k} = (k_1, k_2, \ldots, k_n)$, $\rho_{\boldsymbol{k}} = \otimes_{j=1}^n \rho_{k_j}$, and $\mu_{\boldsymbol{k}} \in \Delta^n(\rho_{\boldsymbol{k}})$. Since $\mathcal{D}$ is antidistinguishable, Theorem 6.5 implies that $L\left(\{\mu_{\boldsymbol{k}}\}\right) = 0$, but, due to the PIP, $\mu_{\boldsymbol{k}} = \times_{j=1}^n \mu_{k_j}$ for some $\mu_{k_j} \in \Delta(\rho_{k_j})$. Let m be a measure that dominates μ_0 and μ_1 and let p_0 and p_1 be densities that represent them. Then, we have

$$\int_{\Lambda_1} \ldots \int_{\Lambda_n} \min_{\boldsymbol{k} \in \{0,1\}^n} \left(\prod_{j=1}^n p_{k_j}(\lambda_j) \right) dm(\lambda_1) \ldots dm(\lambda_n) = 0. \tag{122}$$

However, since each p_{k_j} is a positive function, the minimum is attained by minimizing each factor individually. Therefore, we have

$$\left[\int_\Lambda \min_{k \in \{0,1\}} \left(p_k(\lambda)\right) dm(\lambda) \right]^n = 0, \tag{123}$$

and this implies that $L(\mu_0, \mu_1) = \int_\Lambda \min_k \left(p_k(\lambda)\right) dm(\lambda) = 0$, which is the condition for ontological distinctness of ρ_0 and ρ_1. □

Lemma 7.11. *Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment where $\mathcal{P}$ contains all pure states on $\mathcal{H}$. Then, for any pair of pure states $[\psi_0], [\psi_1] \in \mathcal{P}$, $[\psi_0] \neq [\psi_1]$, there exists an n and a product state fragment $\mathfrak{F}^n$ such that the set of states $\left\{\otimes_{j=1}^n \left[\psi_{k_j}\right]\right\}$ for $k_j \in \{0,1\}$ is antidistinguishable in $\mathfrak{F}^n$.*

Proof. Let $\boldsymbol{k} = (k_1, k_2, \ldots, k_n)$ and $[\psi_{\boldsymbol{k}}] = \otimes_{j=1}^n \left[\psi_{k_j}\right]$. We have to prove that, for some n, there exists a POVM $\{E_{\boldsymbol{k}}\}$ on $\mathcal{H}^{\otimes n}$ that antidistinguishes $\{[\psi_{\boldsymbol{k}}]\}$, since, if such a measurement exists, it will be contained in some product state fragment. The global phase ambiguity can be exploited to choose representative vectors $|\psi_0\rangle$ and $|\psi_1\rangle$, such that $\langle\psi_0|\psi_1\rangle$ is a positive real number. Since $|\psi_0\rangle, |\psi_1\rangle$ span a two-dimensional subspace, we can always find an orthonormal basis $\{|0\rangle, |1\rangle, \ldots\}$ for $\mathcal{H}$ such that

$$|\psi_0\rangle = \cos\frac{\theta}{2}|0\rangle + e^{\imath\beta}\sin\frac{\theta}{2}|1\rangle \tag{124}$$

$$|\psi_1\rangle = \cos\frac{\theta}{2}|0\rangle - e^{\imath\beta}\sin\frac{\theta}{2}|1\rangle, \tag{125}$$

where $\cos\theta = \langle\psi_0|\psi_1\rangle$ and $0 \leq \theta \leq \pi/2$. Here, β is a phase that may be chosen arbitrarily by setting the global phase of $|1\rangle$. It is introduced for later convenience.

In defining the measurement $\{E_{\boldsymbol{k}}\}$, we can restrict attention to the subspace spanned by the vectors $|\boldsymbol{k}\rangle = |k_1\rangle \otimes |k_2\rangle \otimes \ldots \otimes |k_n\rangle$ because the states $[\psi_{\boldsymbol{k}}]$ all lie in this subspace. Hence, the projector onto the orthogonal complement of this subspace can be added to one of the measurement operators without affecting the outcome statistics.

To construct the antidistinguishing measurement, set $E_{\boldsymbol{k}} = [\phi_{\boldsymbol{k}}]$ for a set of orthonormal vectors $\{|\phi_{\boldsymbol{k}}\rangle\}$. The orthonormal vectors are given by

$$|\phi_{\boldsymbol{k}}\rangle = \frac{1}{\sqrt{2^n}}\left(e^{-\imath\alpha}|00\ldots00\rangle + \sum_{\{\boldsymbol{m}\in\{0,1\}^n | \boldsymbol{m}\neq(0,0,\ldots,0)\}} (-1)^{\boldsymbol{k}\cdot\boldsymbol{m}}|\boldsymbol{m}\rangle\right), \tag{126}$$

where $\boldsymbol{k}\cdot\boldsymbol{m} = \sum_j k_j m_j$. To see that these vectors are orthonormal, note that

$$\begin{aligned}\langle\phi_{\boldsymbol{k}}|\phi_{\boldsymbol{l}}\rangle &= \frac{1}{2^n}\left(1 + \sum_{\boldsymbol{m}\neq(0,0,\ldots,0)} (-1)^{(\boldsymbol{k}+\boldsymbol{l})\cdot\boldsymbol{m}}\right) \\ &= \frac{1}{2^n}\sum_{\boldsymbol{m}\in\{0,1\}^n} (-1)^{(\boldsymbol{k}+\boldsymbol{l})\cdot\boldsymbol{m}},\end{aligned} \tag{127}$$

where $\boldsymbol{k}+\boldsymbol{l}$ denotes componentwise addition modulo 2, and the second line follows because $(\boldsymbol{k}+\boldsymbol{l})\cdot(0,0,\ldots,0) = 0$. Now, if $\boldsymbol{k} = \boldsymbol{l}$ then $(\boldsymbol{k}+\boldsymbol{l}) = (0,0,\ldots,0)$, so all the terms are 1 and there are 2^n of them, so $\langle\phi_{\boldsymbol{k}}|\phi_{\boldsymbol{k}}\rangle = 1$. If $\boldsymbol{k} \neq \boldsymbol{l}$ then $\boldsymbol{k}+\boldsymbol{l}$ has some positions where the component is 1. However, there are as many vectors $\boldsymbol{m}$ with an odd number of 1's on these positions as there are with an even number of 1's on these positions, so overall there are as many -1 terms as there are $+1$ terms in the sum, and hence $\langle\phi_{\boldsymbol{k}}|\phi_{\boldsymbol{l}}\rangle = 0$.

Now, the relevant measurement probabilities are $\mathrm{Tr}\left(E_{\boldsymbol{k}}\left[\psi_{\boldsymbol{k}}\right]\right) = |\langle\phi_{\boldsymbol{k}}|\psi_{\boldsymbol{k}}\rangle|^2$ and these should be zero if

$\{E_{\boldsymbol{k}}\}$ antidistinguishes $\{[\psi_{\boldsymbol{k}}]\}$. We have,

$$\langle\phi_{\boldsymbol{k}}|\psi_{\boldsymbol{k}}\rangle = \frac{1}{\sqrt{2^n}}\left(e^{\imath\alpha}\langle 00\ldots0| + \sum_{\boldsymbol{m}\neq(0,0,\ldots,0)}(-1)^{\boldsymbol{k}\cdot\boldsymbol{m}}\langle\boldsymbol{m}|\right)\bigotimes_{j=1}^{n}\left(\cos\frac{\theta}{2}|0\rangle + (-1)^{k_j}e^{\imath\beta}\sin\frac{\theta}{2}|1\rangle\right) \tag{128}$$

$$= \frac{1}{\sqrt{2^n}}\left(\left(\cos\frac{\theta}{2}\right)^n e^{\imath\alpha} + \sum_{\boldsymbol{m}\neq(0,0,\ldots,0)}(-1)^{\boldsymbol{k}\cdot\boldsymbol{m}}\left(\cos\frac{\theta}{2}\right)^{n-|\boldsymbol{m}|}\left(\sin\frac{\theta}{2}\right)^{|\boldsymbol{m}|}e^{\imath|\boldsymbol{m}|\beta}(-1)^{\boldsymbol{k}\cdot\boldsymbol{m}}\right) \tag{129}$$

$$= \frac{1}{\sqrt{2^n}}\left(\left(\cos\frac{\theta}{2}\right)^n e^{\imath\alpha} + \sum_{r=1}^{n}\binom{n}{r}\left(\cos\frac{\theta}{2}\right)^{n-r}\left(\sin\frac{\theta}{2}\right)^r e^{\imath r\beta}\right) \tag{130}$$

$$= \frac{1}{\sqrt{2^n}}\left(\cos\frac{\theta}{2}\right)^n\left(e^{\imath\alpha} + \left(1+e^{\imath\beta}\tan\frac{\theta}{2}\right)^n - 1\right), \tag{131}$$

where $|\boldsymbol{m}|$ denotes the Hamming norm $|\boldsymbol{m}| = \sum_j m_j$.

For a given θ, this can be made to equal zero if it is possible to choose n, α and β such that

$$e^{\imath\alpha} + \left(1+e^{\imath\beta}\tan\frac{\theta}{2}\right)^n - 1 = 0. \tag{132}$$

The required α will exist provided there exist choices of n and β such that

$$\left|1 - \left(1+e^{\imath\beta}\tan\frac{\theta}{2}\right)^n\right| = 1. \tag{133}$$

Such values exist if the curve of $f_n(\beta) = 1 - \left(1+e^{\imath\beta}\tan\frac{\theta}{2}\right)^n$ intersects the unit circle in the complex plane for some value of n. Since f_n is continuous it suffices to find, for some fixed n, one point outside the unit circle and one point inside it.

Choose n such that

$$n \geq \frac{1}{\log_2\left(\tan\frac{\theta}{2}+1\right)}, \tag{134}$$

so that $\tan\frac{\theta}{2} \geq 2^{\frac{1}{n}} - 1$. Then,

$$f_n(0) = 1 - \left(1+\tan\frac{\theta}{2}\right)^n \tag{135}$$

$$\leq 1 - \left(1+2^{\frac{1}{n}}-1\right)^n \tag{136}$$

$$= -1, \tag{137}$$

so $f_n(0)$ is either on or outside the unit circle. On the other hand

$$f_n(\pi) = 1 - \left(1-\tan\frac{\theta}{2}\right)^n, \tag{138}$$

and since $0 \leq \theta \leq \pi/2$ we have $0 \leq \tan\frac{\theta}{2} \leq 1$ so $0 \leq f_n(\pi) \leq 1$. Hence, $f_n(\pi)$ is either on or inside the unit circle. □

Proof of theorem 7.15. Consider any pair of pure states $[\psi_0], [\psi_1] \in \mathcal{P}$. By Lemma 7.11, there exists a product state fragment $\mathfrak{F}^n$ in which the states $\{[\psi_{k_1}] \otimes [\psi_{k_2}] \otimes \ldots \otimes [\psi_{k_n}]\}$ are antidistinguishable. Then, by the PIP and Lemma 7.10, $[\psi_0]$ and $[\psi_1]$ must be ontologically distinct in any ontological model of $\mathfrak{F}$. Since this applies to all pairs of pure states, any such ontological model must be ψ-ontic. □

7.3 Motivation for the PIP

The PIP says that ontological models for subsystems should be compatible with models for product states on the composite system. Compatibility breaks down into two assumptions: the CPA, which says that the ontic state space of the composite system should be a cartesian product of the ontic states spaces of the subsystems, and the NCA, which says that product states should be modelled by product measures. We consider the motivations for each of these requirements in turn.

The CPA is a weakening of a kind of locality assumption known as *Einstein separability*. The terminology is due to Howard [109], who defines it as the idea that "two spatially separated systems possess their own separate real states". Einstein formulated this idea in the context of his arguments for the incompleteness of quantum theory, and expressed it as follows.

> If one asks what is characteristic of the realm of physical ideas independently of the quantum-theory, then above all the following attracts our attention: the concepts of physics refer to a real external world, i.e. ideas are posited of things that claim a 'real existence' independent of the perceiving subject (bodies, fields, etc.), and these ideas are, on the other hand, brought into as secure a relationship as possible with sense impressions. Moreover, it is characteristic of these physical things that they are conceived of as being arranged in a space-time continuum. Further, it appears to be essential for this arrangement of the things introduced in physics that, at a specific time, these things claim an existence independent of one another, insofar as these things 'lie in different parts of space'. Without such an assumption of the mutually independent existence (the 'being-thus') of spatially distant things, an assumption which originates in everyday thought, physical thought in the sense familiar to us would not be possible. Nor does one see how physical laws could be formulated and tested without such a clean separation. Field theory has carried out this principle to the extreme, in that it localizes within infinitely small (four-dimensional) space-elements the elementary things existing independently of one another that it takes as basic, as well as the elementary laws it postulates for them. — Albert Einstein [110].

The PBR argument only makes use of product state preparations, so we can always imagine that the individual systems are prepared very far away from each other and are only later brought together to perform the entangled measurement. If implemented this way, the individual systems would occupy spatially separated regions at their point of preparation and so, according to separability, they ought to possess their own separate real states. Implicit in this is the idea that there are no inherently global joint properties of the composite system that are not determined by the properties of the individual systems. In the language of ontological models, this means that ontic state spaces should compose according to the cartesian product.

Einstein thought that separability should always hold, regardless of whether we are preparing product states or entangled states. In light of Bell's Theorem, the case for such a general separability assumption is significantly weakened. Separability is not actually required to prove Bell's Theorem [111], but, if we are contemplating a world in which the effects of measurement can be transmitted instantaneously across space, then it makes sense to also contemplate a world in which there are inherently global properties as well. Additionally, since we are in the business of proving ψ-ontology theorems and the quantum state of an entangled system would, if ontic, be such an inherently global property, we had better not introduce any assumptions that rule them out. For this reason, the CPA is restricted to product state preparations. It says that separability should hold, not in general, but only when systems are prepared in product states.

The motivation for assuming separability for product states is that we generally think that experiments on separated systems are independent of one another. It should not be necessary to gather our system here on Earth together with one on Mars in order to determine all of the parameters relevant to our Earth-bound experiment. Of course, when performing experiments involving systems on Earth that are correlated with those on Mars, what happens on Mars is very relevant, but the CPA only applies to product states. If we allow that genuinely global properties may be relevant even to an isolated system then we open up a Pandora's box. It could well be necessary to gather together every system in the universe in order to determine all the parameters that are relevant for our system here on Earth.

At this point, it is worth noting that the assumptions behind no-go theorems are often designed to mirror operational features of quantum theory at the ontological level. This is perhaps clearest in the case of contextuality. Preparation noncontextuality says that if there is no difference between two preparation procedures in terms of the observable statistics they predict, i.e. they are represented by the same quantum state, then there should be no difference between them at the ontological level either, i.e. they should be represented by the same probability measure over the ontic states. If this assumption does not hold then an explanatory gap is opened because, if two preparations are represented by different probability measures, then one would generally expect to be able to pick up this difference in the observed statistics. To prevent this from happening, the way that measurements are represented by response functions has to be fine-tuned so that the difference between the two probability densities is washed out by averaging. This fine-tuning has no natural explanation within the ontological model, other than that it is needed in order to reproduce the quantum predictions. Similarly, on an operational level, local measurements on entangled states cannot be used to send signals, so it makes sense to demand locality when modelling them ontologically. Otherwise, we would have to explain why the purported nonlocal influences cannot be used to signal. Proofs of preparation contextuality and Bell's Theorem therefore expose explanatory gaps in the ontological models framework.

In this vein, separability can be motivated by the operational principle known as *local tomography*. This says that the state of a joint system, and hence the probabilities it predicts for any measurement, are completely determined by the statistics of local measurements made on the subsystems. For example, the state of a two-qubit system is completely determined by the statistics of measurements of the Pauli observables σ_x, σ_y, and σ_z made on each qubit individually and their correlations. There are no parameters of the joint state that require bringing the two subsystems together and making a joint measurement in order to determine them. Therefore, it makes sense to posit an underlying ontology that does not involve genuinely global properties either. Without this assumption, we would have to explain why the genuinely global properties do not lead to observable parameters that can only be determined by bringing the systems together.

Of course, the whole point of no-go theorems, and the reason they are surprising, is that they show that such operationally motivated assumptions are not actually viable at the ontological level, at least within the usual framework. They each imply fine-tunings for which we currently have no good explanation. My point here is just that separability bears a family resemblance to the assumptions behind other no-go theorems, so if you think that preparation noncontextuality and Bell's local causality have intuitive appeal then the same should hold for separability as well. If we find, as we have, that some of these assumptions are not actually viable then it makes sense to explore the consequences of weakened versions of them, which maintain some of the appeal but are not yet ruled out. From this perspective the CPA, as a weakening of separability, is a reasonable thing to posit.

The NCA can similarly be motivated as mirroring an operational feature of quantum theory at the ontological level. When two systems are prepared in a product state they are completely uncorrelated from each other. No quantum measurement will ever reveal any pre-existing correlation. Therefore, it makes sense to think that the systems are uncorrelated at the ontological level as well. If not, then it is puzzling that this correlation does not show up in measurement statistics.

Perhaps the CPA and NCA are not so self-evidently true that they must never be questioned, but they are at least as solid as the assumptions that go into other no-go theorems. Therefore, the PBR Theorem presents us with a dilemma. Either we must give up the CPA, the NPA, or ψ-epistemicism, and each choice opens up an explanatory gap. I think this is at least as interesting as the dilemmas posed by Bell's Theorem and contextuality.

7.4 Weakening the assumptions

The assumptions of the PBR Theorem have so far been presented in their most intuitively accessible form. However, it is possible to weaken them somewhat without affecting the conclusion. This has been pointed out by Hall [33] and by Schlosshauer and Fine [38]. The cost of doing this is that the weakened assumptions are less clearly motivated by operational properties of quantum theory. Nevertheless, it is interesting to identify the weakest set of assumptions under which the theorem can be proved.

First of all, note that the only quantum predictions used in the PBR argument involve measurement outcomes that are assigned probability 0 by some quantum state via antidistinguishability. It would make no difference to the argument if ontological models were only required to reproduce the probability 0 predictions of quantum theory, instead of requiring that they reproduce all of the quantum predictions exactly. Thus, the requirement of reproducing the quantum predictions may be replaced by the following.

Definition 7.12. An ontological model $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ of a prepare-and-measure fragment $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ *reproduces the quantum preclusions* if, for all $\rho \in \mathcal{P}$, each $\mu \in \Delta_\rho$ satisfies

$$\int_\Lambda \xi_E^M(\lambda) d\mu(\lambda) = 0, \tag{139}$$

for all $M \in \mathcal{M}$, $E \in M$, such that $\mathrm{Tr}\,(E\rho) = 0$.

Secondly, the only place in the PBR argument where the PIP is used is in Lemma 7.10, where it licenses the implication from the antidistinguishability of the states $\rho_{\boldsymbol{k}} = \otimes_{j=1}^n \rho_{k_j}$ to the ontological distinctness of ρ_0 and ρ_1 via the inference from $L(\{\mu_{\boldsymbol{k}}\}) = 0$ to $L(\mu_0, \mu_1) = 0$. If we want to avoid using the PIP, then we could just assume this directly as follows.

Definition 7.13. An ontological model $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ for a fragment $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ is *compact* with respect to an ontological model $\Xi^n = (\Lambda^n, \Sigma^n, \Delta^n, \xi^n)$ for the n-fold product state fragment $\mathfrak{F}^n = \langle \mathcal{H}^{\otimes n}, \mathcal{P}^{\times n}, \mathcal{M}' \rangle$ with factors $\mathfrak{F}$ if, for all pairs $\rho_0, \rho_1 \in \mathcal{P}$, whenever $L(\{\mu_{\boldsymbol{k}}\}) = 0$ for all choices of $\mu_{\boldsymbol{k}} \in \Delta^n(\rho_{\boldsymbol{k}})$ (where $\rho_{\boldsymbol{k}} = \otimes_{j=1}^n \rho_{k_j}$), then $L(\mu_0, \mu_1) = 0$ for all choices of $\mu_0 \in \Delta(\rho_0), \mu_1 \in \Delta(\rho_1)$.

Definition 7.14. An ontological model for a fragment $\mathfrak{F}$ that reproduces the quantum preclusions satisfies *compactness* if it is compact with respect to an ontological model that reproduces the quantum preclusions for any n-fold product state fragment $\mathfrak{F}^n$ with factors $\mathfrak{F}$.

The condition of compactness is originally due to Schlosshauer and Fine [38], and is a slight generalization of a condition that Hall called "compatibility" [33]. Here, I have presented it in a slightly more rigorous form in order to take care of measure zero issues and to allow for preparation contextuality. In this form, the meaning of compactness may be somewhat obscure, so it is helpful to consider the special case of a finite ontic state space. In this case, when $n = 2$, compactness says is that, if a measure μ_0 corresponding to ρ_0 assigns a nonzero probability to the ontic state λ and a measure μ_1 corresponding to ρ_1 likewise assigns a nonzero probability to λ, then there ought to be probability densities μ_{00}, μ_{01}, μ_{10} and μ_{11} representing the states $\rho_0 \otimes \rho_0$, $\rho_0 \otimes \rho_1$, $\rho_1 \otimes \rho_0$ and $\rho_1 \otimes \rho_1$ of the composite system that all assign nonzero probability to some ontic state λ'. The advantage of this formulation is that it does not assume that the ontic state space of the composite system has a cartesian product structure because it does not specify how λ', which is an ontic state of the composite, is related to λ, which is an ontic state of the subsystem. It also does not rule out the possibility of correlations between the two systems at the ontological level.

However, the motivation for assuming compactness is somewhat obscure. Unlike the PIP, it is not a weakened form of separability, and has no obvious operational motivation. It is doubtful that one would come up with compactness without having the idea of cartesian products and product measures in mind in the first place. It amounts to simply assuming that Lemma 7.10 is true by fiat. Without any examples of natural models that satisfy compactness but not the PIP, it is not clear why one would make such an assumption. Life would be very simple if we always just raised the lemmas needed to prove theorems into assumptions instead of proving them. Nevertheless, the PBR Theorem can be stated in the more general form.

Theorem 7.15 (The "Generalized" PBR Theorem)**.** *Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a prepare-and-measure fragment where $\mathcal{P}$ contains all pure states on $\mathcal{H}$. Then, any ontological model of $\mathfrak{F}$ that reproduces the quantum preclusions and satisfies compactness is ψ-ontic.*

The proof is just the same as the original PBR Theorem, except for the removal of Lemma 7.10.

7.5 Necessity of the PIP

The PIP is the most substantive assumption that goes into the PBR Theorem. All of the rest of its assumptions are standard for the ontological models framework and are common to the vast majority of no-go theorems for realist models of quantum theory, such as Bell's Theorem and the Kochen-Specker Theorem. Therefore, any criticism of PBR that is not directed against the PIP could equally well be directed against these other results. Whilst I do not wish to belittle such criticisms, and will deal with them in §7.7, it is more important to deal with objections that are specific to PBR, and hence to the PIP. Before doing so, we should check that the PIP is a necessary assumption, since if ψ-ontology can be established without the PIP then the question of whether to accept it is moot.

The Kochen-Specker model shows that the PIP is necessary for the case of a qubit. However, the case of a qubit is rather special and the existence of this model does not rule out the possibility that ontological models for systems of larger dimension might have to be ψ-ontic. Lewis, Barrett, Jennings and Rudolph (LBJR) showed that it is in fact possible to construct a ψ-epistemic model for measurements in orthonormal bases in all dimensions [52]. Subsequently, Aaronson, Bouland, Chua and Lowther (ABCL) provided an alternative construction based on the same basic idea [53]. Of course, none of these models satisfy the PIP. Whilst the LBJR model is ψ-epistemic, the ABCL construction additionally shows that, for any pair $[\psi]$, $[\phi]$ of nonorthogonal states, a ψ-epistemic model can be constructed in which measures corresponding to $[\psi]$ and $[\phi]$ have nonzero overlap. Since the main idea is the same, we focus on the ABCL result here.

Both LBJR and ABCL start from the Bell model described in Example 4.5. This is a ψ-ontic model, but by a simple modification it can be made ψ-epistemic. Recall that, in the Bell model, the ontic state space is $\Lambda = \Lambda_1 \times \Lambda_2$, where Λ_1 is the set of pure states on $\mathbb{C}^d$ and Λ_2 is the unit interval. A state $[\psi]$ is represented by the product measure $\mu = \mu_1 \times \mu_2$, where $\mu_1 = \delta_{[\psi]}$ is the point measure at $[\psi]$ on Λ_1 and μ_2 is the uniform measure on $\Lambda_2 = [0, 1]$. For each $[\lambda_1] \in \Lambda_1$, the response functions are set by dividing up the interval $[0, 1]$ into subsets of length $\mathrm{Tr}\left([\phi_j]\,[\lambda_1]\right)$ that give the $[\phi_j]$ outcome with certainty. In Example 4.5 this was done simply by dividing $[0, 1]$ up into consecutive intervals of the form

$$\sum_{k=1}^{j-1} \mathrm{Tr}\left([\phi_k]\,[\lambda_1]\right) \leq \lambda_2 < \sum_{k=1}^{j} \mathrm{Tr}\left([\phi_k]\,[\lambda_1]\right), \tag{140}$$

as illustrated in Fig. 7. However, this division could be done in a different way, providing the total length of the subset that gives $[\phi_j]$ with certainty remains equal to $\mathrm{Tr}\left([\phi_j]\,[\lambda_1]\right)$. The first step in converting the Bell model into a ψ-epistemic model is to change this division.

Following ABCL, let $[a]$ and $[b]$ be two nonorthogonal states that we would like to make ontologically indistinct. Given a measurement $M = \{[\phi_j]\}_{j=1}^{d}$, let σ be a permutation of $(1, 2, \ldots, d)$ such that

$$\begin{aligned} \min\left(\mathrm{Tr}\left(\left[\phi_{\sigma(1)}\right][a]\right), \mathrm{Tr}\left(\left[\phi_{\sigma(1)}\right][b]\right)\right) &\geq \min\left(\mathrm{Tr}\left(\left[\phi_{\sigma(2)}\right][a]\right), \mathrm{Tr}\left(\left[\phi_{\sigma(2)}\right][b]\right)\right) &(141)\\ &\geq \ldots &(142)\\ &\geq \min\left(\mathrm{Tr}\left(\left[\phi_{\sigma(d)}\right][a]\right), \mathrm{Tr}\left(\left[\phi_{\sigma(d)}\right][b]\right)\right). &(143)\end{aligned}$$

Now, the unit interval is divided up in exactly the same way as before, except with respect to the permuted indices instead of the original ordering of outcomes i.e. the outcome $\left[\phi_{\sigma(j)}\right]$ is obtained with certainty if

$$\sum_{k=1}^{j-1} \mathrm{Tr}\left(\left[\phi_{\sigma(k)}\right][\lambda_1]\right) \leq \lambda_2 < \sum_{k=1}^{j} \mathrm{Tr}\left(\left[\phi_{\sigma(k)}\right][\lambda_1]\right). \tag{144}$$

So far, the model is still ψ-ontic because we have only modified the response functions, so the probability measures still have their point measure terms. The next step is to show that there is an $\epsilon > 0$ such that all ontic states in the set

$$\Omega_{a,b} = \{[a], [b]\} \times [0, \epsilon], \tag{145}$$

always return the $\left[\phi_{\sigma(1)}\right]$ outcome in any measurement. Thus, any weight that a probability measure assigns to this region of the ontic state space can be redistributed in an arbitrary way without affecting the observed probabilities, and by doing so the probability measures associated with $[a]$ and $[b]$ can be made to overlap.

To prove this, let $|a\rangle$, $|b\rangle$ and $|\phi_j\rangle$ be vector representatives of $[a]$, $[b]$ and $[\phi_j]$. It is sufficient to show that there exists an $\epsilon > 0$ such that, for all measurements $M = \{[\phi_j]\}_{j=1}^d$, there exists a j such that $|\langle\phi_j|a\rangle| > \epsilon$ and $|\langle\phi_j|b\rangle| > \epsilon$. Then, whatever $\left[\phi_{\sigma(1)}\right]$ is, it must satisfy $\left|\left\langle\phi_{\sigma(1)}\middle|a\right\rangle\right| > \epsilon$ and $\left|\left\langle\phi_{\sigma(1)}\middle|b\right\rangle\right| > \epsilon$ because σ orders the outcomes in decreasing order of $\min\left(|\langle\phi_j|a\rangle|, |\langle\phi_j|b\rangle|\right)$. To show this, note that

$$|\langle a|b\rangle| = \left|\langle a| \left(\sum_{j=1}^{d} |\phi_j\rangle\langle\phi_j|\right) |b\rangle\right| \tag{146}$$

$$\leq \sum_{j=1}^{d} |\langle a|\phi_j\rangle \langle\phi_j|b\rangle| \tag{147}$$

$$\leq \sum_{j=1}^{d} |\langle a|\phi_j\rangle| \, |\langle\phi_j|b\rangle|, \tag{148}$$

where we have used the fact that $\sum_{j=1}^d |\phi_j\rangle\langle\phi_j|$ is a resolution of the identity and the triangle inequality. The largest term in the sum must be larger than the average, so we have

$$\max_j \left(|\langle a|\phi_j\rangle| \times |\langle\phi_j|b\rangle|\right) \geq \frac{|\langle a|b\rangle|}{d}, \tag{149}$$

but since $|\langle a|\phi_j\rangle|$ and $|\langle\phi_j|b\rangle|$ are between 0 and 1, this means

$$\max_j \left(\min\left(|\langle a|\phi_j\rangle|, |\langle\phi_j|b\rangle|\right)\right) \geq \frac{|\langle a|b\rangle|}{d}, \tag{150}$$

and since $[a]$ and $[b]$ are nonorthogonal, $|\langle a|b\rangle|/d > 0$, so setting $\epsilon = |\langle a|b\rangle|/d$ gives the desired result.

Now, we can replace the measure associated with $[a]$ with

$$\mu(\Omega) = \int_{\Lambda_2} \mu_{\lambda_2}(\Omega_{\lambda_2}) d\mu_2(\lambda_2), \tag{151}$$

where as before μ_2 is the uniform measure on Λ_2 and $\Omega_{\lambda_2} = \{[\lambda_1] \in \Lambda_1 | ([\lambda_1], \lambda_2) \in \Omega\}$, but now μ_{λ_2} is a measure on Λ_1 that depends on λ_2 via

$$\mu_{\lambda_2} = \begin{cases} \delta_{[a]} & \text{if } \lambda_2 > \epsilon \\ \epsilon\mu_{\Omega_{a,b}} & \text{if } \lambda_2 \leq \epsilon, \end{cases} \tag{152}$$

where $\mu_{\Omega_{a,b}}$ is an arbitrary measure on $\Omega_{a,b}$ which, for example, could be the uniform measure. Similarly for $[b]$, with $\delta_{[a]}$ replaced by $\delta_{[b]}$. Since both measures now agree on $\Omega_{a,b}$ and assign nonzero probability to it, the model is ψ-epistemic.

This model is only ψ-epistemic in a fairly weak sense, since only a single pair of pure states has any overlap. However, the same idea can be used as the basis for constructing a model in which every pair of nonorthogonal pure states has overlap (see §11 for an outline of the construction).

The PBR Theorem implies that the ABCL model must necessarily fail to satisfy the PIP. To see this, we have to show that, given an ABCL model for $\mathbb{C}^d$, there is some n for which there is no model for product states on the composite system $\left(\mathbb{C}^d\right)^{\otimes n}$ with which it is compatible. Given that the ABCL construction can be used to make any single pair of pure states ontologically indistinct, we can apply it to the states $[z+]$ and $[x+]$ in $\mathbb{C}^2$ from Example 7.9. Then, compatibiity implies that the model for product states on $\mathbb{C}^2 \otimes \mathbb{C}^2$ would have to have nonzero four-way overlap for the states $[z+] \otimes [z+]$, $[z+] \otimes [x+]$, $[x+] \otimes [z+]$

and $[x+] \otimes [x+]$, since these are represented by products of the measures representing $[z+]$ and $[x+]$ in the ABCL model, which have nonzero pairwise overlap. However, nonzero overlap for these four states is ruled out by their antidistinguishability, so no such model can exist. Similarly, whichever pair of quantum states we choose to make ontologically indistinct, the PBR Theorem rules out compatibility with models for the composite system for some n.

7.6 Criticism of the PIP

Since the PIP is necessary for proving the PBR Theorem, it is important to settle the question of whether it is reasonable. In this section, two criticisms of the PIP are discussed, directed at the CPA and NCA respectively. In §7.6.1, a weakening of the CPA due to Emerson, Serbin, Sutherland and Veitch (ESSV) [32] is discussed along with a ψ-epistemic model that they proposed for the experiment of Example 7.9 based on an extension of Spekkens' toy theory. In my view, this is a reasonable weakening of the PIP, but the ESSV model reduces the amount of overlap between nonorthogonal states compared with Spekkens' theory, so it is not clear whether it can be extended to model all of quantum theory in a ψ-epistemic way. §7.6.2 discusses a criticism due to Hall [33] that the NCA does not follow from causality considerations of the type considered in Bell's Theorem. Whilst this argument is formally correct, I think it misses the point as the NCA is not motivated by causality in the first place.

7.6.1 Criticism of the CPA

In my view, the weakest part of the PIP is the CPA, i.e. the idea that there should be no global properties of a system that are not reducible to properties of its subsystems when it is prepared in a product state. Recall that, in order to model direct product fragments, we can always get away with taking the cartesian product of ontic state spaces and using product measures. Therefore, the only time global properties would necessarily have to play a role is when a joint measurement is made, e.g. a measurement in an entangled basis. Such measurements already involve an interaction between the two systems at the operational level, so invoking global properties to explain such experiments does not seem to open up a huge gap between the structure of the ontological model and the operational structure of the experiment. The argument that it might be necessary to bring a system on Earth together with a system on Mars in order to characterize its properties does not hold water if the genuinely global properties do not have any effect on the outcomes of local measurements. It would still be possible to work with separate systems completely independently of each another, in blissful ignorance of the global properties, until we decide to do an experiment that necessarily involves bringing the systems together.

Recently, Emerson, Serbin, Sutherland and Veitch (ESSV) [32] have shown how a ψ-epistemic model of this type can reproduce the predictions of Example 7.9. They first outline a weakening of the PIP that allows for global properties, but still has an effective cartesian product structure and lack of correlation when we restrict attention to local experiments. They then show how the statistics of the antidistinguishable states $[z+] \otimes [z+], [z+] \otimes [x+], [x+] \otimes [z+], [x+] \otimes [x+]$ can be modelled in a ψ-epistemic way under the entangled measurement considered in Example 7.9.

Definition 7.16. A pair of ontological models $\Xi_A = (\Lambda_A, \Sigma_A, \Delta_A, \xi_A)$, $\Xi_B = (\Lambda_B, \Sigma_B, \Delta_B, \xi_B)$ for fragments $\mathfrak{F}_A = \langle \mathcal{H}_A, \mathcal{P}_A, \mathcal{M}_A \rangle$ and $\mathfrak{F}_B = \langle \mathcal{H}_B, \mathcal{P}_B, \mathcal{M}_B \rangle$ are *weakly compatible* with an ontological model $\Xi_{AB} = (\Lambda_{AB}, \Sigma_{AB}, \Delta_{AB}, \xi_{AB})$ for a product state fragment $\mathfrak{F}_{AB} = \langle \mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{P}_A \times \mathcal{P}_B, \mathcal{M}_{AB} \rangle$ if Ξ_{AB} satisfies

- The Weak Cartesian Product Assumption (WCPA): $\Lambda_{AB} = \Lambda_A \times \Lambda_B \times \Lambda_{\mathrm{NL}}$, where Λ_{NL} represents some global degrees of freedom not reducible to properties of system A and system B alone, and $\Sigma_{AB} = \Sigma_A \otimes \Sigma_B \otimes \Sigma_{\mathrm{NL}}$, where Σ_{NL} is a σ-algebra over Λ_{NL}.
- The Weak No-Correlation Assumption (WNCA): $\forall \rho_A \in \mathcal{P}_A, \rho_B \in \mathcal{P}_B$, all $\mu_{AB} \in \Delta_{AB}(\rho_A \otimes \rho_B)$ satisfy

$$\mu_{AB}(\Omega_{AB} \times \Lambda_{\mathrm{NL}}) = \mu_A \times \mu_B(\Omega_{AB}), \tag{153}$$

for all $\Omega_{AB} \in \Sigma_A \otimes \Sigma_B$ for some $\mu_A \in \Delta_A(\rho_A)$ and $\mu_B \in \Delta_B(\rho_B)$, i.e. the marginal measure on $\Lambda_A \times \Lambda_B$ satisfies the NCA.

In general, the ontic state space for a composite system need not break down neatly into a product of local ontic state spaces and global properties. The WCPA is therefore a relatively mild generalization of the CPA. The WNCA then says that, if we only have access to λ_A and λ_B, then the probability measures corresponding to product states look just like product measures, even though they may in fact be correlated via the third variable λ_{NL}. Further, not only do they look like product measures, they look just like the product of measures that the ontological models Ξ_A and Ξ_B would assign to the two systems individually. If we assume that Ξ_A and Ξ_B are adequate for modelling $\mathfrak{F}_A$ and $\mathfrak{F}_B$ individually, it follows that local measurements can be modelled by products of the response functions ξ_A and ξ_B assigned by these models, so the outcomes of local measurements need only depend on λ_A and λ_B. Therefore, under local measurements, this type of model is indistinguishable from one that satisfies the PIP. It is only under joint measurements that there is a difference, since the outcomes of these may depend on λ_{NL}.

Definition 7.17. A pair of ontological models Ξ_A and Ξ_B for fragments $\mathfrak{F}_A$ and $\mathfrak{F}_B$ satisfy the *Weak Preparation Independence Postulate (WPIP)* if, for all product state fragments $\mathfrak{F}_{AB}$ with $\mathfrak{F}_A$ and $\mathfrak{F}_B$ as factors, there exists an ontological model Ξ_{AB} that is weakly compatible with Ξ_A and Ξ_B.

For the ESSV model, take $\mathcal{P}_A = \mathcal{P}_B = \{[z+], [x+]\}$ and let $\mathcal{M}_A$ and $\mathcal{M}_B$ be the set of measurements considered in Spekkens' toy theory, as described in Example 4.3. For the product state fragment $\mathfrak{F}_{AB}$ we include, in addition to the local measurements, the measurement $M = \{[\phi_1], [\phi_2], [\phi_3], [\phi_4]\}$, where

$$|\phi_1\rangle = \frac{1}{\sqrt{2}} \left(|z+\rangle |z-\rangle + |z-\rangle |z+\rangle \right) \tag{154}$$

$$|\phi_2\rangle = \frac{1}{\sqrt{2}} \left(|z+\rangle |x-\rangle + |z-\rangle |x+\rangle \right) \tag{155}$$

$$|\phi_3\rangle = \frac{1}{\sqrt{2}} \left(|x+\rangle |z-\rangle + |x-\rangle |z+\rangle \right) \tag{156}$$

$$|\phi_4\rangle = \frac{1}{\sqrt{2}} \left(|x+\rangle |x-\rangle + |x-\rangle |x+\rangle \right), \tag{157}$$

which is the antidistinguishing measurement from Example 7.9.

Since $\mathcal{P}_A$ and $\mathcal{P}_B$ are a subset of the states modelled by the toy theory, the ontological models for the individual systems can just be taken to be toy-bits. The ontic state spaces Λ_A and Λ_B are both $((+,+),(+,-),(-,+),(-,-))$ and $[z+]$ and $[x+]$ are modelled by the distributions $|z+)$ and $|x+)$ shown in Fig. 3. However, to model the composite system, an extra variable with ontic state space $\{0,1\}$ is introduced, called the *nonlocal bit*. The total ontic state space is just the cartesian product of this with the two toy-bit state spaces, so the WCPA is satisfied. The measures μ_{z+z+}, μ_{z+x+}, μ_{x+z+} and μ_{x+x+}, corresponding to the four product states $[z+] \otimes [z+]$, $[z+] \otimes [x+]$, $[x+] \otimes [z+]$ and $[x+] \otimes [x+]$, are just as they would be in the toy theory, except that we also have to specify the distribution of the nonlocal bit. If the two quantum states prepared are different (i.e. $[z+] \otimes [x+]$ or $[x+] \otimes [z+]$), and the local ontic state is $((+,+),(+,+))$, then the nonlocal bit takes the value 1. In all other cases it takes the value 0. The resulting distributions are illustrated in Fig. 12. By construction, marginalizing these distributions over the nonlocal bit just gives the product distributions that one would assign in the toy theory, so the WNCA is satisfied.

The response functions for local measurements are just products of those from the toy-theory, so it remains to specify the response function for the antidistinguishing measurement $M = \{[\phi_1], [\phi_2], [\phi_3], [\phi_4]\}$. These are illustrated in Fig 13. It is straightforward to check that they reproduce the quantum predictions when applied to the distributions given in Fig. 12.

Inspection of Fig. 12 shows that this model is ψ-epistemic. Each pair of distributions overlaps on exactly one cell where the nonlocal bit 0 in both cases, and both distributions assign 25% probability to this cell. The variational distance between each pair is therefore 3/4. Note, however, that in the usual toy theory without the nonlocal bit, the variational distances would be 1/2, so we have gained the ability to model the

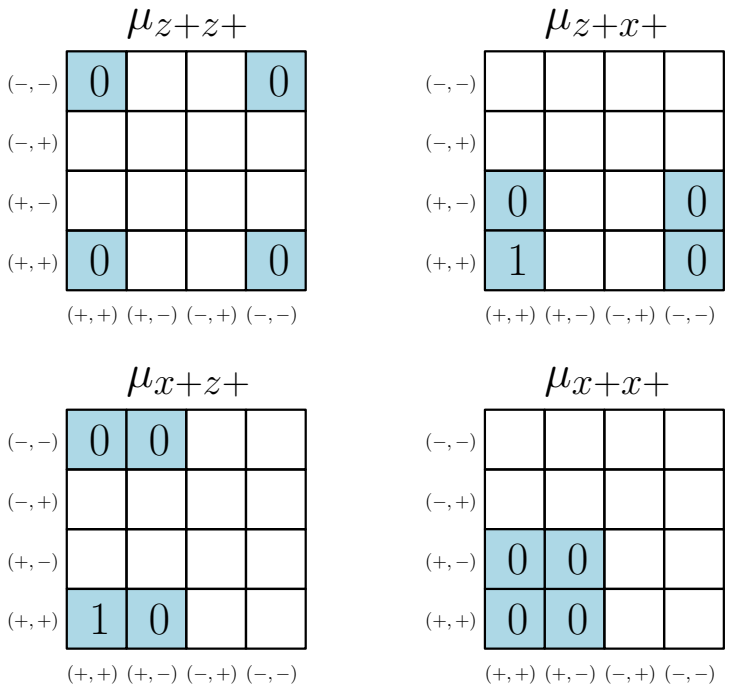


Figure 12: The probability distributions for composite systems in the ESSV model. The grid represents the local ontic state spaces, with the ontic state of the first toy-bit varying along the horizontal and the second along the vertical. Light blue cells have 25% probability and white cells have 0 probability. The values of the nonlocal bit are shown for each cell that has nonzero probability.

antidistinguishing measurement by reducing the overlaps between states. Thus, although the ESSV model shows that adopting the WPIP can evade the conclusion of the PBR Theorem, it does not show that a viable ψ-epistemic model for the whole of quantum theory can be constructed that satisfies the WPIP. It could be that modelling additional entangled measurements would require the overlaps to be reduced further, and perhaps they might have to go down to zero in order to model all possible measurements.

In contrast, the ABCL model shows that ψ-epistemic models are possible in general, but it does not satisfy the WPIP because the ontic state space it assigns to $\mathbb{C}^d \otimes \mathbb{C}^d$ does not factorize neatly into local state spaces for each system and nonlocal degrees of freedom. It may be possible to rewrite the model in such a way that the ontic state space does factorize, but then it is not clear that the WNCA would be satisfied. Therefore, it is an open question whether a ψ-epistemic ontological model exists that satisfies the WPIP and reproduces the whole of quantum theory.

7.6.2 Criticism of the NCA

The PIP has also been criticized on other grounds by Hall [33] (this criticism is also mentioned by Schlosshauer and Fine [38]). Hall objects to the NCA on the grounds that, even if they are spacelike separated, the events corresponding to preparing two systems have a common past, so their lack of correlation cannot be derived from causality by the same sort of reasoning that Bell used to motivate his locality condition. More explicitly, suppose we have a model that does satisfy the CPA, so that we can associate separate ontic states λ_A and λ_B to Alice and Bob's systems. Suppose that the intersection of the past lightcones of Alice and Bob's preparation events also has some physical properties described its own ontic state λ_{past}. For concreteness, suppose that all the ontic state spaces are finite, i.e. the ontic state spaces of Alice's system, Bob's system, and their common past are all finite. The argument does not depend on this, but it makes things conceptually simpler.

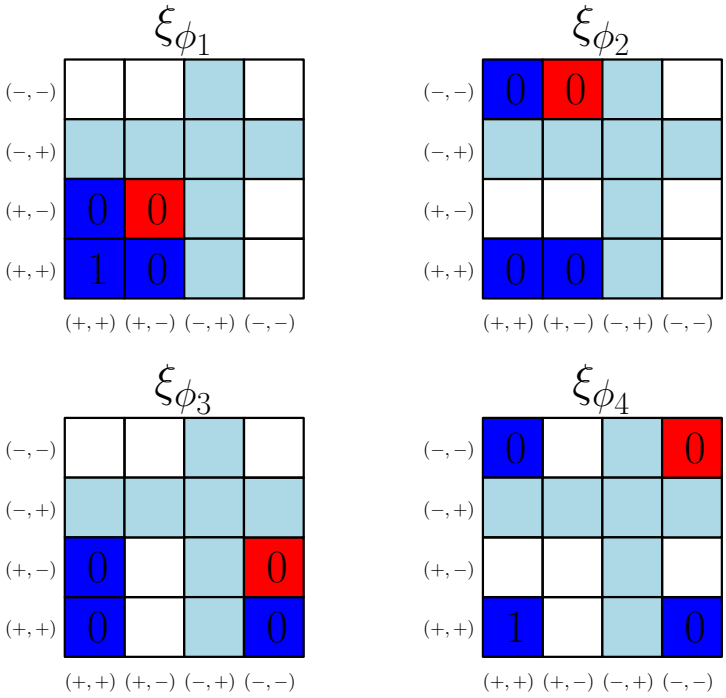


Figure 13: Response functions for the antidistinguishing measurement in the ESSV model. Cells with no specified value for the nonlocal bit give the same outcome probability regardless of the value of the bit. White cells indicate probability 0 and light blue cells indicate 25% probability. Blue cells indicate 50% probability, but only if the nonlocal bit takes the value specified, otherwise they give probability 0. Similarly, red cells indicate 100% probability, but only if the nonlocal bit is as specified, otherwise they give probability 0.

Suppose that Alice decides to prepare the quantum state $[\psi]_A$ and Bob decides to prepare the quantum state $[\phi]_B$. In general, the resulting λ_A and λ_B might depend on both Alice and Bob's choices of preparation and the ontic state λ_{past} of their common past, so the preparation procedure would be specified by conditional probabilities $P(\lambda_A, \lambda_B | [\psi]_A, [\phi]_B, \lambda_{\text{past}})$. Given that the preparation procedures might occur at spacelike separation, it is not unreasonable to impose a factorization of probabilities akin to Bell locality, which would read

$$P(\lambda_A, \lambda_B | [\psi]_A, [\phi]_B, \lambda_{\text{past}}) = P(\lambda_A | [\psi]_A, \lambda_{\text{past}}) P(\lambda_B | [\phi]_B, \lambda_{\text{past}}), \tag{158}$$

However, this is not enough to entail the NCA because the measure μ corresponding to the preparation of $[\psi]_A \otimes [\phi]_B$ would then be,

$$\mu(\lambda_A, \lambda_B) = \sum_{\lambda_{\text{past}}} P(\lambda_A | [\psi]_A, \lambda_{\text{past}}) P(\lambda_B | [\phi]_B, \lambda_{\text{past}}) P(\lambda_{\text{past}}), \tag{159}$$

where $P(\lambda_{\text{past}})$ is the prior distribution over the variables in the common past. This can induce correlations between λ_A and λ_B due to their common dependence on λ_{past}.

This argument is correct, but all it shows is that the NCA cannot be regarded as a causality assumption akin to Bell locality. It does not imply that the NCA is necessarily unreasonable. To infer that, one would have to believe that the *only* reasonable type of assumption to make in a no-go theorem is one that follows from Bell-type locality. In contrast, §7.3 discussed a more general strategy for positing reasonable assumptions, which is to look at the operational structure of quantum theory and try to impose similar structure at the ontological level. In the case of the NCA, we noted that a product state displays no correlations in any quantum measurement, so it makes sense to posit that no such correlations exist at the ontological level either. It is this, and not Bell locality, that is the motivation for the NCA. Dropping the

NCA in light of the PBR Theorem is not an obviously crazy thing to do, but it is the PBR Theorem itself rather than causality arguments that should motivate this move. The same reasoning applies to all other no-go theorems that are not motivated by locality. For example, I think it is fairly clear that proofs of preparation contextuality are a good reason for dropping the assumption of preparation noncontextuality, but the fact that the latter cannot be derived from Bell locality is a much less compelling reason for doing so.

7.7 Other Criticisms of the PBR Theorem

Apart from criticism of the PIP, several other criticisms of the PBR Theorem have been raised. §7.7.1 discusses a criticism due to Drezet [30, 31] and Schlosshauer and Fine [38] based on the idea that response functions should depend on the quantum state in addition to the ontic state. I argue that this criticism is simply a misunderstanding of what is meant by the term "ontic state" in the ontological models framework. I then discuss two further criticisms due to Schlosshauer and Fine [38]. The first criticism, discussed in §7.7.2, is a claim that the ψ-ontic/ψ-epistemic distinction is merely conventional because one kind of model can be converted into a structurally equivalent model of the other kind. The second criticism, discussed in §7.7.3, is based on the idea that modelling detector inefficiencies offers a way out of the dilemma posed by PBR. Whilst this is strictly speaking true, it only concerns the practical implementation of PBR-type experiments and has no impact on the PBR Theorem as a structural result about quantum theory itself, assuming ideal experiments. If one accepts the PIP, error analysis can be used to constrain this escape route in more practical experiments. Finally, in §7.7.4, I discuss criticisms that reject some aspect of the ontological models framework outright. These criticisms are mostly based on a neo-Copenhagen point of view, and thus are easy to deal with as the PBR Theorem was never intended to rule out such approaches.

Since the criticisms in this section are directed against the ontological models framework in general, they are not specific to PBR but could also be directed against Hardy's Theorem, the Colbeck-Renner Theorem and the results discussed in Part III. I discuss them here because they were made as responses to the PBR Theorem and not these other results, but this is simply because the PBR Theorem was the first and is still the most prominent ψ-ontology theorem.

7.7.1 Response functions should depend on the quantum state

Several authors have pointed out that, in the ontological models framework, the response functions ξ_E^M are assumed to be independent of the quantum state ρ that is prepared [30, 31, 38]. If a dependence $\xi_E^{M,\rho}$ is allowed then the theorem can be trivially evaded. The response functions can simply ignore the ontic state and return the quantum probabilities by setting $\xi_E^{M,\rho}(\lambda) = \mathrm{Tr}\,(E\rho)$, since we then have, for any probability measure μ,

$$\int_\Lambda \xi_E^{M,\rho}(\lambda) d\mu(\lambda) = \mathrm{Tr}\,(E\rho) \int_\Lambda d\mu(\lambda) \tag{160}$$

$$= \mathrm{Tr}\,(E\rho)\,. \tag{161}$$

The ontic state space and probability densities can then be anything at all, so the model can trivially be made to satisfy the PIP. For example, $\mathbb{C}^d$ could be associated with the ontic state space $\Lambda_d = \{1, 2, \ldots, d\}$ and $\mathbb{C}^d \otimes \mathbb{C}^{d'}$ with the Cartesian product $\Lambda_d \times \Lambda_{d'}$, which is isomorphic to $\Lambda_{dd'}$. Then, the uniform density can be used to represent all quantum states, which makes the model trivially ψ-epistemic.

Of course, proponents of this view do not have this sort of model in mind as a realistic candidate for describing quantum theory. They think the response functions should depend on the ontic states in some way in addition to the quantum state. The above model is just intended to show how trivial the ψ-ontic/ψ-epistemic distinction becomes when response functions are allowed to depend on the quantum state.

I think there are two intuitions behind this sort of objection. The first is based on elementary probability theory and the second on a misunderstanding of how conventional hidden variable theories, such as de Broglie-Bohm theory, are meant to fit into the ontological models framework. First of all, putting aside

everything we know about the ontological models framework for the moment, suppose we are interested in some prepare-and-measure fragment $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ and we have a theory for reproducing its predictions that involves some sort of additional variable λ that takes values in a set Λ. For the purposes of this argument, assume that Λ, $\mathcal{P}$ and $\mathcal{M}$ are finite, so that elementary probability theory can be used. This restriction can easily be lifted, but dealing with the measure theoretic complications would obscure the argument.

A quite general way that such a theory could be formulated is in terms of a joint probability distribution $P(\lambda, E, M, \rho)$ over all the variables involved, where $P(\lambda, E, M, \rho)$ is specified for every $\rho \in \mathcal{P}$, $M \in \mathcal{M}$, $E \in M$, $\lambda \in \Lambda$. In order to determine whether such a model reproduces the quantum predictions, we need to determine the conditional probabilities $P(E|\rho, M)$, i.e. the probability for an outcome given the choice of state and measurement, because quantum theory tells us this should equal $\mathrm{Tr}\,(E\rho)$. This can be computed as follows.

$$P(E|\rho, M) = \sum_{\lambda} P(E, \lambda|\rho, M) = \sum_{\lambda} P(E|\lambda, \rho, M) P(\lambda|\rho, M), \tag{162}$$

where the conditional probabilities are defined in terms of the joint probability in the usual way and the second equality follows from the law of total probability.

On the other hand, in the ontological models framework, the same quantity would be computed as

$$P(E|\rho, M) = \sum_{\lambda} \xi_E^M(\lambda) \mu(\lambda), \tag{163}$$

where $\mu \in \Delta_\rho$ and the usual integral has been replaced by a sum because Λ is finite. Comparing these two expressions, in an ontological model we have

$$P(E|\lambda, \rho, M) = \xi_E^M(\lambda) \tag{164}$$

$$P(\lambda|\rho, M) = \mu(\lambda). \tag{165}$$

The left hand side of Eq. (164) depends on ρ, but the right hand side does not. Similarly, the left hand side of Eq. (165) depends on M, but the right hand side does not. Despite appearances, both sides of Eq. (165) depend on ρ because μ itself is ρ dependent given that it must me a member of Δ_ρ. Therefore, it seems that the ontological models framework implicitly assumes that the following conditional independences hold.

$$P(E|\lambda, \rho, M) = P(E|\lambda, M) \tag{166}$$

$$P(\lambda|\rho, M) = P(\lambda|\rho). \tag{167}$$

The second conditional independence, that λ should not depend on the choice of measurement, was noted in the context of PBR by Hall [33], but it is not all that controversial. It is usually justified by the idea that the measurement setting is a "free choice" that may be chosen by the experimenter long after the preparation is completed. Things can be set up such that the measurement choice is determined by something that should be independent of the rest of the experiment, such as a random number generator. Theories in which dependence of λ on M nevertheless still holds in the underlying ontology are often called *superdeterministic*. There is a minority that seriously advocates superdeterminism, but this loophole exists in almost all no-go theorems for ontological models of quantum theory, e.g. it applies to Bell's Theorem as well. Therefore, if it is taken seriously as a response to PBR then it must be taken seriously for these other results as well. One way in which dependence on M can happen is if there are retrocausal influences that travel backwards in time. Personally, I am not opposed to developing retrocausal theories as a response to all the quantum no-go theorems, although I am not convinced it is the right direction either. Nonetheless, this criticism is not specific to PBR.

The first conditional independence, that the measurement outcome should be independent of the quantum state given the ontic state and the choice of measurement, is the one that is objected to more frequently. However, this is not really a substantive assumption, but rather it is part of the very meaning of the term "ontic state". The ontic state is supposed to comprise all the properties of the system that exist in reality.

In addition to its own setting, the response of the measurement device is only supposed to depend on those properties of the system that exist in reality, so the ontic state is the only information about the preparation procedure that it receives. If there is an additional dependence on the quantum state then that simply means that we have made an incorrect assertion about what the ontic state actually is. It must include all the information that is required to determine the response of the measurement device. Therefore, saying that the response function should depend on the quantum state is tantamount to saying that the quantum state is part of the ontic state, and it is very easy to prove that the quantum state is ontic if you assume that it is ontic from the outset. In conclusion, I agree with the critics that this conditional independence is an assumption of the PBR Theorem, but it is an assumption that is part of the definition of the term "ontic state", rather than something that can be eliminated in order to arrive at a more general notion of what it means for a model to be ψ-epistemic that still conveys the same meaning.

I think that proponents of this objection have been misled by the way in which hidden variable theories, and de Broglie-Bohm theory in particular, have traditionally been presented. It is often thought that the aim of a hidden variable theory should be to restore determinism, and so the problem of developing such a theory is often phrased in terms of whether quantum theory is "complete". The terminology "complete" suggests taking the idea that elements of the existing quantum formalism represent reality for granted, and only asking whether anything else needs to be added to it. It is obviously critical to not take this point of view if the reality of the wavefunction is the very thing under investigation. The criticism of Drezet [30, 31] exemplifies this mistake. He suggests that de Broglie-Bohm theory is a counterexample to the PBR argument. In de Broglie-Bohm theory, particles have well-defined positions which evolve deterministically, and the probability distribution assigned to the particles is given by $|\psi(x)|^2$, where $\psi(x)$ is the wavefunction. Drezet's argument is that, if we view the particle positions as the ontic states of the system, then their distributions overlap for any pair of nonorthogonal states because in this case the $|\psi(x)|^2$ distributions overlap. He claims that this makes the theory ψ-epistemic.

Now, in the conventional understanding of de Broglie-Bohm theory, the wavefunction is understood to be part of the ontic state in addition to the particle positions. It is true that the particle positions are in some sense more fundamental than the wavefunction, and they are often called the "primitive ontology" [112–114] or the "local beables" of the theory [115]. The particle positions are supposed to be the things in the theory that provide a direct picture of what reality looks like to us, e.g. when we observe the pointer of a measurement device pointing to a specific value then it is the positions of the particles that make up the pointer that determine this. Nevertheless, the wavefunction is still needed as part of the ontology because it determines how the particles move via the guidance equation. The response of a measurement device to an interaction with a system it is measuring depends on the wavefunction of the system as well as the particle positions, so the wavefunction is still part of the ontic state, even if it is in some sense less primitive than the particle positions. Of course, Drezet can get away with having only the particle positions comprise the ontic state if he allows the response functions to depend on the wavefunction separately. Indeed, as pointed out by Schlosshauer and Fine [38], this is often how the response functions are written in de Broglie-Bohm theory and other hidden variable theories, but this is because that framework was addressing the issue of completeness, which assumes that the wavefunction is real, rather than the question of whether the quantum state is real in the first place. As I have argued, it is part of the definition of an ontic state that it suffices to completely determine how a measurement device will react to the system, so if your response functions depend on the wavefunction then the wavefunction is ontic and there is nothing left to prove.

7.7.2 The ψ-ontic/ψ-epistemic distinction is conventional

Fine and Schlosshauer [38] claim that the distinction between ψ-ontic and ψ-epistemic models is merely conventional because a ψ-ontic model can be converted into a "structurally equivalent" model that is ψ-epistemic and vice versa. They do not actually define the term "structurally equivalent", so it is perhaps best to look at the procedures they propose for converting models.

Firstly, starting from a ψ-epistemic model it is trivial to construct a ψ-ontic one. Simply take the new ontic state space to be the cartesian product of the existing ontic state space with the set of pure quantum states. Then, for each pure state preparation, take the existing probability measures and form the product

with a point measure on the same pure state in the new component of the ontic state space. These are the probability measures of the new model. Finally, extend the response functions to the new ontic state space in the most trivial way, by having them not depend on the new factor at all. Because of the point measures, the new model is ψ-ontic, but because the response functions completely ignore this component of the ontic state space, the model makes the exact same predictions as the original one.

The construction that Fine and Schlosshauer intend for converting a ψ-ontic model into a ψ-epistemic model is less obvious. They simply refer to the LBJR paper for this [52], which uses the same sort of construction as the ABCL model, discussed in §7.5. Recall that the idea is to find regions of the ontic state space associated with two different quantum states that make the exact same predictions for all measurements. One can then redistribute any weight that probability measures assign to this region such that measures associated with distinct quantum states now match on this region, and hence they now overlap. Fine and Schlosshauer seem to think that such regions can always be found, but this is not the case. For example, in the Beltrametti-Bugajski model, the ontic state space is just the set of pure quantum states and each quantum state is represented by a different point measure on this space. The response functions simply return the quantum probabilities, so each pair of ontic states makes different predictions for some quantum measurement. Therefore, there are no regions of the ontic state space corresponding to distinct quantum states that make identical predictions. Thus, this construction cannot be used to generate a ψ-epistemic model in this case.

To be fair, Fine and Schlosshauer confine their attention to deterministic models, but they do not prove that appropriate regions of the ontic state space can always be found even in this case. It might be interesting to investigate this, but nevertheless determinism is deliberately not an assumption of the ontic models framework, so the fact that there are non-deterministic models to which their construction does not apply is enough to defeat the criticism.

Further, since Schlosshauer and Fine do not define what they mean by "structurally equivalent", it is not clear what their objection is in the first place. The only requirement I can infer from their paper is that "structurally equivalent" models should make the same predictions for all quantum measurements. If this is really all that the term means then any two models that reproduce the quantum predictions would be structurally equivalent. For example, the Beltrametti-Bugajski model and de Broglie-Bohm theory would be structurally equivalent by this criterion. It is clear, however, that they are not *explanatorially* equivalent. Beltrametti-Bugajski is simply a more precise version of the orthodox interpretation of quantum theory in which the quantum state, and only the quantum state, is the state of reality. This brings with it all the attendant problems of measurement and the collapse of the wavefunction. On the other hand de Broglie-Bohm solves these problems by introducing additional variables. Whether or not you think it is plausible as a fundamental theory, it does not have a measurement problem. Therefore, reproducing the same set of predictions does not mean that two theories are equally viable.

Of course, to some extent, Fine and Schlosshauer are just pointing out that, without the PIP, both ψ-epistemic and ψ-ontic models are possible. However, this means that criticism of the PBR Theorem should be directed specifically at the PIP, and not at whether the ψ-ontic/ψ-epistemic distinction makes sense in the first place.

7.7.3 Detector inefficiencies

Fine and Schlosshauer also make another objection based on what would happen if detector inefficiencies were included in our models. The basic idea is that, if each measurement includes an additional outcome corresponding to the system failing to be detected, then overlap between states can be restored. A simple way of doing this is to imagine that the ontic state space has a special ontic state that always causes the detection to fail. If there is a nonzero probability of detection failure in our experiments for every pure state, then the probability measures corresponding to them can always be made to overlap by assigning nonzero probability to the special ontic state.

In response to this, firstly note that one could regard the PBR Theorem as a theorem about the structure of quantum theory itself, independently of any ideas about how to test it experimentally. In this regard, one can imagine idealized state preparations and measurements that conform to the quantum predictions

exactly and have zero probability of failure. On this view, considerations of detection failure are irrelevant.

However, any real experiment does have some failure probability and one never reproduces the predictions of quantum theory exactly. I have not discussed error analysis in this review, but PBR do provide such an analysis in their paper [27]. The aim here is to show that if the experimental results differ by ϵ from the quantum predictions then the overlap of probability measures must be much smaller than the inner products of the corresponding quantum states. This poses just as much of a problem for the ψ-epistemic view as the exact version of PBR, since the overlaps should be comparable to the inner products in order for the ψ-epistemic explanations of indistinguishability, no cloning, etc. to work. I believe that the PBR error analysis is sound in this regard. However, it does require the PIP, since the fact that independently prepared systems are represented by product measures is needed to relate the probabilities for an experiment involving n systems to the overlaps of probability measures for a single system. Fine and Schlosshauer instead advocate the compactness postulate, discussed in §7.4, which does not imply any quantitative relationship between the probability measures for a single system and for n systems. This is perhaps why they think that detection efficiency is a more serious problem, but my response to that is to just advocate adopting the PIP instead, as compactness does not really have a compelling motivation of its own.

One other way of reading Fine and Schlosshauer's objection is as a prediction that detection efficiencies will be close to zero if preparations and measurements involved in the PBR Theorem are actually performed in the lab. In this way, the antidistinguishing measurement can almost always just fail to return an outcome, so that the probabilities of each outcome for each quantum state really are close to zero. This is quite implausible for special cases that only involve a small number of systems, such as in Example 7.9, since we can already do fairly reliable entangled measurements on small numbers of qubits [116]. It is possible, however, that quantum theory will break down for some suitably large number of systems, and large n is needed to test the PBR Theorem for pairs of quantum states that have large inner product. It would of course be very interesting if quantum theory did break down in this way, but I do not think that the idea that quantum theory would break down just to preserve the ψ-epistemic interpretation of quantum states is very plausible. Indeed, a successor theory may very well involve a new notion of the state of a system, so the question of whether the quantum state is real in such a theory would not necessarily be meaningful. It is better to think of the PBR Theorem as a result about the type of explanations that are possible within quantum theory itself.

7.7.4 Rejecting the ontological models framework

Finally, several authors have objected to the adoption of the ontological models framework wholesale [34, 41, 45]. Usually, these objections come from those who adopt neo-Copenhagen approaches, so I would say that this is just a misunderstanding of the intended scope of the PBR Theorem, which was never intended to rule out such interpretations.

For example, Griffiths [41] wonders why anyone would still be interested in the ontological models framework at all, given that existing results like Bell's Theorem and the Kochen-Specker Theorem already make it look quite implausible. To this I would respond that, to the extent that explicitly nonlocal and contextual theories like de Broglie-Bohm theory, spontaneous collapse theories and modal interpretations are currently taken seriously, the framework is interesting because it encapsulates them and allows us to study what other constraints must be satisfied by theories in this category. Beyond that, the ontological models framework is interesting as a model of how to simulate quantum systems with classical resources, so even if the framework is without foundational significance, it is still relevant to quantum information theory. Of course, Griffiths thinks he has a superior approach in the form of decoherent/consistent histories, but to my mind the best way of understanding this approach is either as a neo-Copenhagen interpretation or as a way of formulating the branching structure in a many-worlds interpretation. Both of these are beyond the intended scope of the ontological models framework and the PBR Theorem.

Similarly, Hofmann [34] thinks that the conclusion of PBR can be avoided by allowing exotic probability theories, such as those involving negative probabilities. Such theories have a long history in quantum theory, with the most famous example being the Wigner function [117]. However, exotic probability theories are usually couched in neo-Copenhagen or operationalist terms, i.e. it is fine to use whatever mathematical

object you like to represent unobservable things, so long as you always derive an ordinary positive probability distribution for observable measurement outcomes. In the ontological models framework, the ontic state λ is supposed to be something that objectively exists independently of the experimenter. Although we may not know the exact value of λ and may only be able to detect coarse-grained properties of it in our experiments, it is supposed to be knowable in principle, even if only by a hypothetical super-quantum agent who is not subject to the same limitations as us. For such an agent, the probability assigned to λ has to have one of its conventional meanings, e.g. in terms of frequencies, betting odds, etc., since it is not conceptually different from any other probability. Therefore, only requiring probabilities to be positive for things that we can observe is not good enough to constitute a realist interpretation. One would need to specify what it means to assign a negative probability to λ for someone who can know λ exactly and, as far as I am aware, no such interpretation of negative probabilities exists.

8 Dynamics in ontological models

The remaining two ψ-ontology theorems—Hardy's Theorem and the Colbeck-Renner Theorem—make use of assumptions about how dynamics are represented in ontological models. Two distinct scenarios are relevant. Firstly, we extend the notion of a PM fragment to include the possibility of performing discrete unitary transformations between preparation and measurement. As with measurements, we do not want to assume that this dynamics is deterministic at the ontological level, so a unitary is represented by a stochastic transformation on the ontic state space. This type of dynamics is discussed in §8.1. An important property of stochastic transformations is that they cannot increase the variational distance between probability measures. Therefore, if two quantum states are ontologically indistinct then they remain ontologically indistinct after applying a unitary transformation.

Using just unitary dynamics, it is possible to prove versions of both the Hardy and Colbeck-Renner Theorems, but they fail to show that all nonorthogonal pure states must be ontologically distinct. Instead, they show that, if the inner product of a pair of pure states is less than some quantity that depends on the Hilbert space dimension, then they must be ontologically distinct. This quantity tends to 1 in the limit of infinite dimension, so we do obtain full blown ψ-ontology theorems in this limit.

The trick to extending these results into full blown ψ-ontology theorems for arbitrary dimensions is to consider a different type of dynamical transformation. This involves appending an ancillary system in a fixed state to the system of interest and is discussed in §8.2. This does not change the inner product of the original pure states, but it does increase the dimension of the Hilbert space. The dimension of the ancillary system can then be chosen so that the required inequality is satisfied in the larger Hilbert space. However, our definition of a fragment assumes that the Hilbert space is fixed and adding an ancillary system changes the Hilbert space. Therefore, in the present framework, appending a system should be viewed as a dynamical mapping between two *different* fragments and, at the ontological level, as a mapping between two *different* ontological models. If appending an ancilla is modelled by a stochastic transformation in the same sort of way as unitary dynamics within a fixed fragment, then it also shares the property that ontologically indistinct quantum states remain so after such a transformation.

The property that dynamics preserve ontological indistinctness, be they unitary or the appending of ancillas, is the only assumption about dynamics needed to prove the Colbeck-Renner Theorem. Therefore, we state this as an explicit assumption so that the Colbeck-Renner argument can be formulated in terms of PM fragments with just this additional assumption. On the other hand, Hardy's Theorem involves an additional assumption about dynamics, to be discussed in §9.1, so it requires the extended notion of a fragment that includes unitary transformations.

8.1 Unitary dynamics

Definition 8.1. A *prepare-measure-transform (PMT) fragment* of quantum theory $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M}, \mathcal{T} \rangle$ consists of a Hilbert space $\mathcal{H}$, a set $\mathcal{P}$ of density operators on $\mathcal{H}$, a set $\mathcal{M}$ of POVMs on $\mathcal{H}$, and a set $\mathcal{T}$ of

unitary operators on $\mathcal{H}$ that contains the identity. Additionally, $\mathcal{P}$ is closed under the action of $\mathcal{T}$, i.e. if $\rho \in \mathcal{P}$ and $U \in \mathcal{T}$ then $U\rho U^\dagger \in \mathcal{P}$.

The quantum probability of obtaining the outcome $E \in M$ of a measurement $M \in \mathcal{M}$ when the state $\rho \in \mathcal{P}$ is prepared and the transformation $U \in \mathcal{T}$ is applied between preparation and measurement is

$$\text{Prob}\left(E|\rho, M, U\right) = \text{Tr}\left(EU\rho U^\dagger\right). \tag{168}$$

The constraint that $\mathcal{P}$ should be closed under the action of $\mathcal{T}$ is imposed because preparing the state ρ followed by implementing U provides a method of preparing the state $U\rho U^\dagger$. The identity is assumed to be in $\mathcal{T}$ because we want a PMT fragment to be an extension of a PM fragment, so it should be possible to do nothing between preparation and measurement, i.e. Eq. (168) reduces to the probability rule for a PM fragment when U is the identity. In addition, it is natural to impose further consistency constraints. Firstly, it is common to assume that if $U, V \in \mathcal{T}$ then $VU \in \mathcal{T}$, because if we can implement U and V separately then we could apply them one after the other. This makes $\mathcal{T}$ a semigroup. Similarly, it is usual to assume that $\mathcal{M}$ is closed under the action of $\mathcal{T}$, i.e. if $\{E_j\} \in \mathcal{M}$ and $U \in \mathcal{T}$ then $\{U^\dagger E_j U\} \in \mathcal{M}$. This is because applying U before measuring $\{E_j\}$ is a method of measuring $\{U^\dagger E_j U\}$. These additional constraints are not imposed here because they are not required for the Hardy or Colbeck-Renner Theorems.

In order to construct an ontological model for a PMT fragment, states and measurements are represented on an ontic state space (Λ, Σ) as before. The only novel issue is how to represent transformations. As with measurements, we do not want to assume that $U \in \mathcal{T}$ acts deterministically on the ontic states, so in general U is represented by a Markov kernel γ with source and target both equal to (Λ, Σ). This means that γ is a measurable function that associates, to each $\lambda \in \Lambda$, a probability measure γ_λ on (Λ, Σ). For $\Omega \in \Sigma$, $\gamma_\lambda(\Omega)$ is the conditional probability that the ontic state will end up in Ω after the transformation, given that it started in the ontic state λ. For a finite ontic state space $\gamma_\lambda(\{\lambda'\})$ is the probability that the dynamics causes λ to make a transition to λ', and thus a Markov kernel is just the measure theoretic generalization of a transition matrix, which is used to model stochastic dynamics for a system with finite state space. If the system is assigned probability measure μ before the transformation then afterwards it is assigned the measure ν, where

$$\nu(\Omega) = \int_\Lambda \gamma_\lambda(\Omega) d\mu(\lambda). \tag{169}$$

As with states, there is the possibility that different methods of implementing U might lead to different Markov kernels, in which case the model is *transformation contextual* [65]. To account for this, each $U \in \mathcal{T}$ is associated with a set Γ_U of Markov Kernels, rather than just one.

The consistency constraint that $\mathcal{P}$ is closed under $\mathcal{T}$ also needs to be reflected at the ontological level. Suppose $\rho, \sigma \in \mathcal{P}$ and $\sigma = U\rho U^\dagger$ for some $U \in \mathcal{T}$. For every $\mu \in \Delta_\rho$ and $\gamma \in \Gamma_U$, it should be the case that $\nu \in \Delta_\sigma$, where ν is given by Eq. (169). This is because any method of preparing ρ and applying U is a method of preparing σ.

Definition 8.2. An *ontological model* $\Xi = (\Lambda, \Sigma, \Delta, \xi, \Gamma)$ of a PTM fragment $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M}, \mathcal{T} \rangle$ consists of an ontic state space (Λ, Σ), a function Δ that maps each $\rho \in \mathcal{P}$ to a set of probability measures $\Delta(\rho) = \Delta_\rho$, a function ξ that maps each pair $M \in \mathcal{M}, E \in M$ to a response function $\xi(M, E) = \xi_E^M : \Lambda \to [0, 1]$ such that

$$\forall M \in \mathcal{M}, \lambda \in \Lambda, \qquad \sum_{E \in M} \xi_E^M(\lambda) = 1, \tag{170}$$

and a function Γ that maps each $U \in \mathcal{T}$ to a set of Markov kernels $\Gamma(U) = \Gamma_U$, i.e. $\gamma \in \Gamma_U$ is a measurable function $\gamma : \lambda \to \gamma_\lambda$ where γ_λ is a probability measure on (Λ, Σ).

In addition, for every $\rho, \sigma \in \mathcal{P}$, $U \in \mathcal{T}$ such that $\sigma = U\rho U^\dagger$, for every $\mu \in \Delta_\rho$ and $\gamma \in \Gamma_U$, it must be the case that $\nu \in \Delta_\sigma$, where

$$\nu(\Omega) = \int_\Lambda \gamma_\lambda(\Omega) d\mu(\lambda). \tag{171}$$

The ontological model *reproduces the quantum predictions* if, for all $\rho \in \mathcal{P}$, each $\mu \in \Delta_\rho$ satisfies

$$\forall M \in \mathcal{M}, E \in M, \qquad \int_\Lambda \xi_E^M(\lambda) d\mu(\lambda) = \text{Tr}\left(E\rho\right). \tag{172}$$

Note that we do not have to explicitly impose that, for all $\rho \in \mathcal{P}$, $U \in \mathcal{T}$, each pair $\mu \in \Delta_\rho$, $\gamma \in \Gamma_U$ satisfies

$$\forall M \in \mathcal{M}, E \in M, \qquad \int_\Lambda \int_\Lambda \xi_E^M(\lambda') d\gamma_\lambda(\lambda') d\mu(\lambda) = \mathrm{Tr}\left(EU\rho U^\dagger\right), \tag{173}$$

since this is implied by the consistency constraints on states and the probability densities that represent them.

In fact, as in the PBR Theorem, Hardy's Theorem only depends on the weaker requirement that the model reproduces the quantum preclusions, which means that

$$\int_\Lambda \xi_E^M(\lambda) d\mu(\lambda) = 0, \tag{174}$$

whenever $\mathrm{Tr}\,(E\rho) = 0$.

The following standard result is the basis of the most common assumption about dynamics used to prove ψ-ontology theorems.

Theorem 8.3. *Let μ and ν be probability measures on a measurable space (Λ, Σ), let γ be a Markov kernel with source (Λ, Σ) and target (Λ', Σ'), and let μ' and ν' be the measures on (Λ', Σ') resulting from the action of γ on μ and ν, i.e. for $\Omega' \in \Sigma'$,*

$$\mu'(\Omega') = \int_\Lambda \gamma_\lambda(\Omega') d\mu(\lambda) \tag{175}$$

$$\nu'(\Omega') = \int_\Lambda \gamma_\lambda(\Omega) d\nu(\lambda). \tag{176}$$

Then, $D(\mu', \nu') \leq D(\mu, \nu)$, where D is the variational distance.

Proof. Let m be a measure that dominates μ and ν and let p and q be corresponding densities. Then, for all $\Omega' \in \Sigma'$

$$|\mu'(\Omega') - \nu'(\Omega')| = \left| \int_\Lambda \gamma_\lambda(\Omega') \left(p(\lambda) - q(\lambda)\right) dm(\lambda) \right| \tag{177}$$

$$\leq \int_\Lambda \left|\gamma_\lambda(\Omega') \left(p(\lambda) - q(\lambda)\right)\right| dm(\lambda), \tag{178}$$

where the second line follows from the triangle inequality. However, since $0 \leq \gamma_\lambda(\Omega') \leq 1$, we have

$$\int_\Lambda \left|\gamma_\lambda(\Omega') \left(p(\lambda) - q(\lambda)\right)\right| dm(\lambda) \leq \int_\Lambda \left|p(\lambda) - q(\lambda)\right| dm(\lambda) = D(\mu, \nu). \tag{179}$$

Hence, for all $\Omega' \in \Sigma'$,

$$|\mu'(\Omega') - \nu'(\Omega')| \leq D(\mu, \nu). \tag{180}$$

Therefore, this also applies for the supremum of the left hand side, so $D(\mu', \nu') \leq D(\mu, \nu)$. □

Corollary 8.4. *Let $\langle \mathcal{H}, \mathcal{P}, \mathcal{M}, \mathcal{T} \rangle$ be a PMT fragment and let $\Xi = (\Lambda, \Sigma, \Delta, \xi, \Gamma)$ be an ontological model of it. If $\rho, \sigma \in \mathcal{P}$ are ontologically indistinct then $U\rho U^\dagger$ and $U\sigma U^\dagger$ are also ontologically indistinct for all $U \in \mathcal{T}$.*

Proof. If ρ and σ are ontologically indistinct then there exist $\mu \in \Delta_\rho$ and $\nu \in \Delta_\sigma$ such that $D(\mu, \nu) < 1$. However, for any $\gamma \in \Gamma_U$, $\mu' \in \Delta_{U\rho U^\dagger}$ and $\nu' \in \Delta_{U\sigma U^\dagger}$, where, for all $\Omega \in \Sigma$,

$$\mu'(\Omega) = \int_\Lambda \gamma_\lambda(\Omega) d\mu(\lambda) \tag{181}$$

$$\nu'(\Omega) = \int_\Lambda \gamma_\lambda(\Omega) d\nu(\lambda). \tag{182}$$

Hence, by Theorem 8.3, $D(\mu', \nu') \leq D(\mu, \nu) < 1$, so $U\rho U^\dagger$ and $U\sigma U^\dagger$ are also ontologically indistinct. □

If the PMT fragment of interest contains all unitary transformations on $\mathcal{H}$, then corollary 8.4 implies that whether or not two pure states are ontologically indistinct depends only on their inner product. This is because, if $[\psi'] = U[\psi]U^\dagger$ and $[\phi'] = U[\phi]U^\dagger$, then $[\psi] = U^\dagger[\psi']U$ and $[\phi] = U^\dagger[\phi']$, so applying Corollary 8.4 to both U and $U^\dagger$ implies that $[\psi]$ and $[\phi]$ are ontologically distinct iff $[\psi']$ and $[\phi']$ are.

In fact, Corollary 8.4 is the only property of the way that unitary dynamics is represented that is used in proving the Colbeck-Renner Theorem. Therefore, for the purposes of proving that theorem, all of the detailed considerations about how unitary dynamics are represented can be replaced by the following assumption.

Definition 8.5. Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment and let $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ be an ontological model of it. Ξ *preserves ontological indistinctness* with respect to a unitary operator U if, for every $\rho, \sigma \in \mathcal{P}$ that are ontologically indistinct in Ξ, $U\rho U^\dagger$ and $U\sigma U^\dagger$ are also ontologically distinct in Ξ whenever $U\rho U^\dagger, U\sigma U^\dagger \in \mathcal{P}$.

The benefit of this assumption is that, if we can show that $[\psi]$ and $[\phi]$ are ontologically distinct then it follows that all pairs $[\psi']$ and $[\phi']$ where $\mathrm{Tr}([\phi'][\psi']) = \mathrm{Tr}([\phi][\psi])$ must also be ontologically distinct, so to prove ψ-ontology we only need to prove that there exists a pair of ontologically distinct states for every value of the inner product.

The assumption that unitary dynamics preserves ontological indistinctness stands independently of how U is represented in an ontological model, but it is worth bearing in mind that the representation of U as a stochastic transformation is really what motivates it.

8.2 Appending ancillas

So far, we have considered unitary dynamics on a fixed Hilbert space. More generally, the Hilbert space of the system may change during the course of the experiment. For example, the dimension may be reduced if part of the system is absorbed into the environment. For present purposes, we only need to consider a very particular kind of change, in which the experimenter appends an additional system in a fixed quantum state τ to the system under investigation. Specifically, if the system is originally described by a state ρ_A on a Hilbert space $\mathcal{H}_A$, then after appending the ancilla it is described by the state $\rho_A \otimes \tau_B$ on the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, where $\mathcal{H}_B$ is the Hilbert space of the ancillary system. In order to prove full-blown ψ-ontology theorems for arbitrary dimensional Hilbert spaces, the following assumption is used.

Definition 8.6. Let $\mathfrak{F}_A = \langle \mathcal{H}_A, \mathcal{P}_A, \mathcal{M}_A \rangle$ and $\mathfrak{F}_B = \langle \mathcal{H}_B, \mathcal{P}_B, \mathcal{M}_B \rangle$ be PM fragments and let $\mathfrak{F}_{AB} = \langle \mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{P}_{AB}, \mathcal{M}_{AB} \rangle$ be a product fragment with factors $\mathfrak{F}_A$ and $\mathfrak{F}_B$. Let $\Xi_A = (\Lambda_A, \Sigma_A, \Delta_A, \xi_A)$ be an ontological model of $\mathfrak{F}_A$ and let $\Xi_{AB} = (\Lambda_{AB}, \Sigma_{AB}, \Delta_{AB}, \xi_{AB})$ be an ontological model of $\mathfrak{F}_{AB}$. Ξ_{AB} *preserves ontological indistinctness* with respect to Ξ_A if, whenever $\rho_A, \sigma_A \in \mathcal{P}_A$ are ontologically indistinct in Ξ_A, $\rho_A \otimes \tau_B$ and $\sigma_A \otimes \tau_B$ are ontologically indistinct in Ξ_{AB} for all $\tau_B \in \mathcal{P}_B$.

The motivation for this is similar to that for the assumption of preserving ontological distinctness with respect to unitary transformations. The action of appending an ancilla is a type of dynamics, and so it should be represented by a stochastic transformation at the ontological level. The main difference from the unitary case is that we are dealing with a transformation that changes the underlying Hilbert space, and our definition of a PM fragment assumes a fixed Hilbert space. Therefore, appending an ancilla is a map between two distinct fragments and so it is represented by a stochastic map between two different ontological models. Specifically, if (Λ_A, Σ_A) is the ontic state space of Ξ_A and $(\Lambda_{AB}, \Sigma_{AB})$ is the ontic state space of Ξ_{AB}, then appending an ancilla would be represented by a Markov kernel γ with source (Λ_A, Σ_A) and target $(\Lambda_{AB}, \Sigma_{AB})$. By Theorem 8.3, this cannot increase the total variation distance of the measures representing quantum states and hence, by the same argument used to prove Corollary 8.4, it preserves ontological indistinctness.

At first, it may seem surprising that appending ancillas should be modelled at the ontological level in the same sort of way as unitary dynamics, since they are very different types of operations. As additional motivation, note that the most general type of dynamics of a quantum system, is described by a Completely-Positive Trace-Preserving (CPT) map and the more general claim is that CPT maps should be represented by stochastic transformations (see e.g. [65] for a discussion of this). Both unitary dynamics and appending

an ancilla are examples of CPT maps, so their representation in terms of stochastic maps are special cases of this general idea.

9 Hardy's Theorem

Hardy has proven a ψ-ontology theorem [51] based on an assumption about how dynamics should be represented in an ontological model known as *ontic indifference*. This assumption is rather unnatural for a ψ-epistemic theory, but nonetheless Hardy's Theorem is of interest due to its close connection with the argument for the reality of the wavefunction based on interference, which was discussed in §3.1. Hardy's Theorem can be regarded as the missing step in the inference from "something must go through both slits" to "that thing must be the wavefunction". The reason for the failure of the argument from interference can therefore be pinpointed more precisely as the failure of ontic indifference.

§9.1 describes the ontic indifference assumption and the way it can be violated in a ψ-epistemic theory, as well as explaining Hardy's motivation for introducing it. §9.2 presents a sketch of a special case of Hardy's Theorem in terms of a simple Mach-Zehnder interferometry experiment. This helps to clarify the relation to the argument from interference. The full theorem and its proof are given in §9.3.

9.1 Ontic Indifference

To understand Hardy's assumption, first consider an ontological model with a finite ontic state space. Ontic indifference then says that, if a pure state $[\psi]$ is invariant under the action of a unitary U, i.e. $U[\psi]U^\dagger = [\psi]$, then there should be a method of implementing U such that all the ontic states that get assigned a nonzero probability by $[\psi]$ are left invariant. As usual, this needs to be modified to accommodate the general measure-theoretic case.

Definition 9.1. Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M}, \mathcal{T} \rangle$ be a PMT fragment and let $\Xi = (\Lambda, \Sigma, \Delta, \xi, \Gamma)$ be an ontological model of it. A state $\rho \in \mathcal{P}$ satisfies *ontic indifference* in Ξ if, for every $U \in \mathcal{T}$ such that $U\rho U^\dagger = \rho$, for every $\mu \in \Delta_\rho$ there exists a $\gamma \in \Gamma_U$ and a set $\Omega \in \Sigma$ such that $\mu(\Omega) = 1$ and $\gamma_\lambda(\Omega') = \delta_\lambda(\Omega')$ for all $\Omega' \in \Sigma, \Omega' \subseteq \Omega$, where δ_λ is the point measure at λ.

The model Ξ satisfies *ontic indifference* if every pure state $[\psi] \in \mathcal{P}$ satisfies ontic indifference in Ξ. It satisfies *restricted ontic indifference* if there exists a pure state $[\psi] \in \mathcal{P}$ that satisfies ontic indifference in Ξ.

To see why ontic indifference is suspect from a ψ-epistemic point of view, it suffices to consider a model with a finite ontic state space Λ. It is then unclear why ontic indifference should hold for a quantum state that is represented by a probability measure with support on more than one ontic state, and there obviously must be such states in a ψ-epistemic model. For example, suppose that $[\psi]$ is represented by the uniform distribution over Λ. Then, any permutation of the ontic states leaves this distribution invariant and hence could potentially represent the dynamics of a unitary that leaves $[\psi]$ invariant without contradicting the quantum predictions. More generally, the measures corresponding to $[\psi]$ need only be fixed points of the stochastic transformations representing the unitaries that leave them invariant, and even this is stronger than required, since the stochastic transformations could also map between different members of Δ_ψ.

In fact, this is precisely what happens in Spekkens' toy theory. Consider the state $[z+]$, which is invariant under the action of the unitary operator σ_z. In Spekkens' theory, $[z+]$ is represented by the distribution $|z+)$ that has equal support on $(+,+)$ and $(-,-)$ and is zero elsewhere on the ontic state space. The transformation σ_z can be represented by the permutation

$$(+,+) \to (-,-) \tag{183}$$
$$(+,-) \to (-,+) \tag{184}$$
$$(-,+) \to (+,-) \tag{185}$$
$$(-,-) \to (+,+). \tag{186}$$

This leaves $|z+\rangle$ invariant but does not satisfy ontic indifference because it swaps the two states $(+,+)$ and $(-,-)$ in the support of $|z+\rangle$. It is straightforward to check that this permutation acts appropriately on all the other distributions as well. For example, $[x+]$ gets mapped to $[x-]$ under σ_z, and, in the toy theory, $|x+\rangle$ has equal support on $(+,+)$ and $(+,-)$ and is zero elsewhere, whereas $|x-\rangle$ has equal support on $(-,+)$ and $(-,-)$ and is zero elsewhere. Since the permutation maps $(+,+)$ to $(-,-)$ and $(+,-)$ to $(-,+)$, it maps $|x+\rangle$ to $|x-\rangle$ as required.

Given that Spekkens' toy theory is an archetypal example of a ψ-epistemic theory, the fact that it does not satisfy ontic indifference is evidence that ontic indifference is not a reasonable assumption for a ψ-ontology theorem. Nevertheless, Hardy does provide a motivation for it based on locality, which reveals an interesting connection to the argument from interference.

Hardy's motivation runs as follows. Consider a single photon which can be in one of d different spatial modes, labelled $x_0, x_1, \ldots, x_{d-1}$. For example, the modes might represent the arms of an interferometer, as depicted in Fig. 14. The state in which the photon is in mode x_j is written as $[x_j]$. Suppose that, at the ontological level, $[x_j]$ is to be thought of as a state in which the photon is literally *in* mode x_j so it corresponds to a situation in which there is literally nothing relevant to the behaviour of the photon located in any of the other modes.

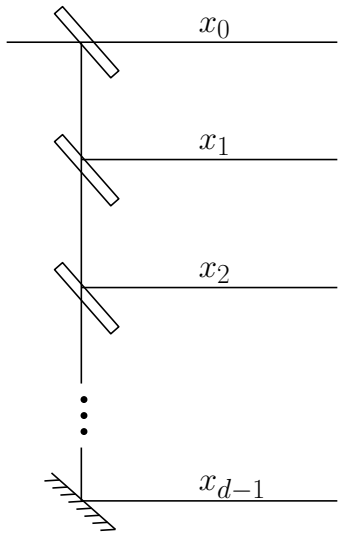


Figure 14: One way of instantiating d spatial modes in an interferometer. A single photon is passed through the first beamsplitter from the left and is repeatedly split at $d-1$ beamsplitters, with the dth mode reflected by the mirror at the bottom.

To model this situation, the ontic state space is assumed to decompose into sets of ontic states $\Lambda^{(j)}$, which are localized at each mode, along with possibly some additional degrees of freedom $\Lambda_{\rm NL}$ which are not so localized. The total ontic state space is then assumed to be $\Lambda = \left(\oplus_{j=0}^{d-1}\Lambda^{(j)}\right) \oplus \Lambda_{\rm NL}$. As a concrete example, each mode might have exactly one ontic state, with the ontic state corresponding to mode x_j labelled by the integer j. The ontic state space would then be $\{0,1,\ldots,d-1,\ldots\} = \{0\} \oplus \{1\} \oplus \ldots \oplus \{d-1\} \oplus \ldots$, where the second $\ldots$ is to leave room for possible additional ontic states not localized in a mode. More generally, the structure would be similar, but there could be multiple ontic states corresponding to each mode. Then, the state $[x_j]$ represents a situation in which only ontic states in $\Lambda^{(j)}$ can be occupied, so $\Lambda^{(j)}$ would be a measure one set according to any probability measure in Δ_{x_j}.

Now, any unitary that leaves $[x_0]$ invariant can be implemented in such a way that it only involves manipulating modes x_1–x_{d-1}, say by adding phase shifters to them or combining them at beamsplitters. If implemented in this way, it makes sense, by locality, to think that this would have no effect on ontic states localized at x_0 and hence ontic indifference would be satisfied for $[x_0]$.

This argument for ontic indifference is based on locality, so it only works for states like $[x_j]$, which are spatially localized. Fortunately, Hardy's Theorem can be proved under restricted ontic indifference, which only requires that ontic indifference should hold for a single pure state, so, in the case of our photon modes setup, this can always be chosen to be $[x_0]$. Hence the assumption need never be applied to a state that is not localized.

Even if the argument from locality were sound, we should be skeptical of imposing locality requirements on ontological models because of Bell's Theorem. However, the argument can be evaded without even giving up on locality, since it is really the assumption that $[x_j]$ corresponds to a set of ontic states localized at x_j that is at fault. Moving to a Fock space description, in which $[n]_{x_j}$ is the state in which there are n photons in mode x_j, makes this objection clearer. In Fock space, the state $[x_0]$ is written as $[1]_{x_0} \otimes [0]_{x_1} \otimes \ldots \otimes [0]_{x_{d-1}}$, which explicitly shows that modes x_1–x_{d-1} are in their vacuum state. From quantum field theory, we know that the vacuum is not a featureless void, but has some sort of structure. Therefore, it makes sense that, at the ontological level, there might be more than one ontic state associated with the vacuum, and a transformation that does not affect things localized at x_0 might still act nontrivially on these vacuum ontic states.

As an explicit example of this, we can construct a "second quantized" version of Spekkens' toy theory, in which we allow each mode to have at most one photon. The states $[0]_{x_j}$, $[1]_{x_j}$, $[+]_{x_j}$, $[-]_{x_j}$, $[+i]_{x_j}$ and $[-i]_{x_j}$, where

$$|\pm\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle \pm |1\rangle\right) \tag{187}$$

$$|\pm i\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle \pm i\,|1\rangle\right), \tag{188}$$

are isomorphic to the states $[z+]$, $[z-]$, $[x+]$, $[x-]$, $[y+]$ and $[y-]$ of a spin-1/2 particle. Therefore, since Spekkens' toy theory provides a model for these spin-1/2 states under measurements in the corresponding bases, it also provides a model for the corresponding photon states and measurements. In this model, the ontic state space for a mode is $\Lambda^{(j)} = \{(+,+)_{x_j}, (+,-)_{x_j}, (-,+)_{x_j}, (-,-)_{x_j}\}$ and the ontic state spaces for modes compose via the cartesian product rather than the direct sum. So, for example, in the case of two modes x_0 and x_1, the total ontic state space would be $\Lambda = \Lambda^{(0)} \times \Lambda^{(1)}$, which is

$$\{((+,+)_{x_0},(+,+)_{x_1}),((+,+)_{x_0},(+,-)_{x_1}),((+,+)_{x_0},(-,+)_{x_1}),((+,+)_{x_0},(-,-)_{x_1}), \tag{189}$$

$$((+,-)_{x_0},(+,+)_{x_1}),((+,-)_{x_0},(+,-)_{x_1}),((+,-)_{x_0},(-,+)_{x_1}),((+,-)_{x_0},(-,-)_{x_1}), \tag{190}$$

$$((-,+)_{x_0},(+,+)_{x_1}),((-,+)_{x_0},(+,-)_{x_1}),((-,+)_{x_0},(-,+)_{x_1}),((-,+)_{x_0},(-,-)_{x_1}), \tag{191}$$

$$((-,-)_{x_0},(+,+)_{x_1}),((-,-)_{x_0},(+,-)_{x_1}),((-,-)_{x_0},(-,+)_{x_1}),((-,-)_{x_0},(-,-)_{x_1})\}. \tag{192}$$

The vacuum state $[0]_{x_j}$ is represented by a equal mixture of $(+,+)_{x_j}$ and $(-,-)_{x_j}$, just like the $[z+]$ state in the spin-1/2 case, and the state $[1]_{x_j}$ with one photon is represented by an equal mixture of $(+,-)_{x_j}$ and $(-,+)_{x_j}$, just like the $[z-]$ state in the spin-1/2 case. The state $[x_0] = [1]_{x_0} \otimes [0]_{x_1}$ is then represented by the product of these two distributions, which is an equal mixture of $((+,-)_{x_0},(+,+)_{x_1})$, $((+,-)_{x_0},(-,-)_{x_1})$, $((-,+)_{x_0},(+,+)_{x_1})$ and $((-,+)_{x_0},(-,-)_{x_1})$. A transformation acting locally on x_1 can then switch the states $(+,+)_{x_1}$ and $(-,-)_{x_1}$, in violation of ontic indifference, whilst leaving the distribution invariant. Thus, the assumption that local ontic state spaces compose according to the direct sum, rather than locality per se., is the main problem with Hardy's argument for ontic indifference. Since we expect vacuum states to have structure, this is not a good assumption on which to build a ψ-ontology theorem.

Nevertheless, the direct sum construction that motivates ontic indifference is closely related to the argument from interference discussed in §3.1. To see this connection, it is helpful to outline a simple special case of Hardy's Theorem.

9.2 An example

Consider a Mach-Zehnder interferometer as depicted in Fig. 15. We will outline the argument that ontic indifference implies that the state $[x_0]$, representing a photon in the upper arm of the interferometer, must be ontologically distinct from $[\psi]$, where $|\psi\rangle = \frac{1}{\sqrt{2}}(|x_0\rangle + |x_1\rangle)$, representing an equal superposition of both paths. The argument in this section assumes a finite ontic state space. A measure-theoretic argument is given for the general case in the next section.

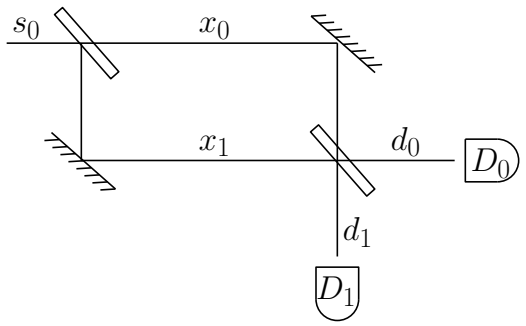


Figure 15: The Mach-Zehnder interferometer used in the example of Hardy's Theorem.

On passing through the second beamsplitter, $|x_0\rangle$ gets mapped to $\frac{1}{\sqrt{2}}(|d_0\rangle + |d_1\rangle)$ and $|x_1\rangle$ gets mapped to $\frac{1}{\sqrt{2}}(|d_0\rangle - |d_1\rangle)$, from which we infer that $|\psi\rangle$ gets mapped to $|d_0\rangle$. Hence, if the state $[x_0]$ is prepared then the detectors D_0 and D_1 will fire with 50/50 probability and if the state $[\psi]$ is prepared then D_0 will fire with certainty. For the practically inclined, note that the state $[\psi]$ can be prepared by passing a photon from the source s_0 through the first beamsplitter as in Fig. 15 and the state $[x_0]$ can be prepared by removing the first beamsplitter and passing a photon from s_0 directly into the upper arm.

The outcome of this experiment can be altered by placing a π phase shifter in the lower arm of the interferometer, e.g. by altering its path length by half a wavelength relative to the upper arm. This leaves $|x_0\rangle$ invariant, but maps $|\psi\rangle$ to $|\phi\rangle = \frac{1}{\sqrt{2}}(|x_0\rangle - |x_1\rangle)$, which will now cause the detector D_1 to fire with certainty. This is a typical example of interference, replacing constructive interference (D_0 fires without the phase shifter) with destructive interference (D_1 fires with the phase shifter).

In an ontological model, assume that there is some ontic state λ that is assigned nonzero probability by both $[x_0]$ and $[\psi]$. We will show that, under the assumption of ontic indifference, this leads to a contradiction.

Since $[\psi]$ and $[\phi]$ are orthogonal, by Theorem 4.10 they must be ontologically distinct. This can also be seen more directly because every ontic state assigned nonzero probability by $[\psi]$ must cause D_0 to fire with certainty, whereas every ontic state assigned nonzero probability by $[\phi]$ must cause D_1 to fire with certainty. Therefore, $[\phi]$ must assign zero probability to the ontic state λ.

Now consider the action of the unitary U that adds a π phase shift to the mode x_1. Since this leaves $[x_0]$ invariant, the assumption of ontic indifference implies that λ must be left invariant. However, since U maps $[\psi]$ to $[\phi]$, it must also cause λ to transition to an ontic state that causes D_1 to fire with certainty, and no such ontic state has nonzero probability according to $[\psi]$. Therefore λ cannot be left invariant, which is a contradiction. We conclude that there can be no λ assigned nonzero probability by both $[x_0]$ and $[\psi]$.

By considering Hardy's motivation for ontic indifference, this can be connected to the argument from interference as follows. Assume that the ontic state space decomposes into a direct sum $\Lambda = \Lambda^{(0)} \oplus \Lambda^{(1)} \oplus \Lambda_{\mathrm{NL}}$. The states in $\Lambda^{(0)}$ are localized in the upper arm of the interferometer, representing the photon definitely taking the upper path, and the states in $\Lambda^{(1)}$ are localized in the lower arm of the interferometer, representing the photon definitely taking the lower path. If we assume that the placement of a π phase shifter in the lower arm has no effect on the ontic states in $\Lambda^{(0)}$ then we can deduce that $[\psi]$ cannot assign nonzero probability to states in $\Lambda^{(0)}$ by the same argument as above. Switching the roles of x_0 and x_1, so that the phase shifter is now placed or not placed in x_0, we can deduce the same for $\Lambda^{(0)}$. Therefore, $[\psi]$ must assign all its weight to Λ_{NL}, which is not localized to either arm. You could say that this represents a situation in which the

photon "travels along both paths", but it would be more accurate to say that it only has global degrees of freedom, not localized on either path. Further, if we assume that $[x_0]$ and $[x_1]$ are entirely supported on $\Lambda^{(0)}$ and $\Lambda^{(1)}$ respectively, then the three quantum states $[\psi]$, $[x_0]$ and $[x_1]$ must be ontologically distinct. This is a bit further along the road to proving ψ-ontology than the more naive version of the argument from interference. Hardy's result then shows that the same sort of assumption can be used to prove a full blown ψ-ontology theorem, which seems to vindicate the intuition behind the argument from interference.

Of course, if we give up the idea that the ontic state space must decompose as a direct sum of local state spaces plus some global degrees of freedom, then there is no motivation for ontic indifference, so the above argument does not go through. In fact, the second quantized version of Spekkens' toy theory is able to account for almost all the qualitative features of Mach-Zehnder interferometry in a local and ψ-epistemic way [85]. Thus, any claim that these experiments entail the reality of the wavefunction or violate locality must involve implicitly assuming something like ontic indifference. The most important aspect of Hardy's analysis is to bring this assumption to the fore more clearly.

9.3 The main result

We are now in a position to state and prove Hardy's main result. The proof strategy used here is due to Patra, Pironio and Massar [118] and is conceptually simpler than Hardy's original proof.

Theorem 9.2. *Let $\mathcal{P}$ be the set of pure states on $\mathbb{C}^d$ for $d \geq 3$. Then, there exists a set of measurements $\mathcal{M}$ and a set of unitaries $\mathcal{T}$ such that any ontological model of the fragment $\mathfrak{F} = \langle \mathbb{C}^d, \mathcal{P}, \mathcal{M}, \mathcal{T} \rangle$ that reproduces the quantum preclusions and satisfies restricted ontic indifference has the property that any pair of states $[\psi], [\phi] \in \mathcal{P}$ satisfying $Tr([\phi][\psi]) \leq (d-1)/d$ must be ontologically distinct.*

Before proving this theorem, note that, if we form a product of $\mathfrak{F}$ with another fragment $\mathfrak{F}'$ and assume that there exits an ontological model of the larger fragment that reproduces the quantum preclusions and preserves ontological indistinctness with respect to our ontological model of $\mathfrak{F}$, then, for any pair $[\psi], [\phi] \in \mathcal{P}$, we can always choose the dimensionality of the Hilbert space of $\mathfrak{F}'$ such that $\mathrm{Tr}([\phi] \otimes [0][\psi] \otimes [0]) \leq (d-1)/d$ is satisfied, where $[0]$ is any fixed pure state of $\mathfrak{F}'$. Hardy's Theorem implies that $[\psi] \otimes [0]$ and $[\phi] \otimes [0]$ must be ontologically distinct, and then the assumption that appending ancillas preserves ontological indistinctness implies that $[\psi]$ and $[\phi]$ must be ontologically distinct in the original model. Since this applies to any pair of pure states, the original ontological model must be ψ-ontic. The case $d = 2$ is dealt with in a similar way.

The proof relies on the following three lemmas.

Lemma 9.3. *Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M}, \mathcal{T} \rangle$ be a PTM fragment, let $\Xi = (\Lambda, \Sigma, \Delta, \xi, \Gamma)$ be an ontological model of it, let $[\psi]$ be a state that satisfies ontic indifference in Ξ, let $\{U_j\} \subseteq \mathcal{T}$ be an at most countable set of unitaries that leave $[\psi]$ invariant and suppose that $[\phi]$ is ontologically indistinct from $[\psi]$ in Ξ. Then there exist measures $\mu_j \in \Delta_{U_j[\phi]U_j^\dagger}$, such that $L(\{\mu_j\}) > 0$, where L is the overlap.*

Proof. Let $\nu \in \Delta_\psi$ and suppose that Ω_j is a measure one set left invariant by some $\gamma^j \in \Gamma_{U_j}$, i.e. $\gamma^j_\lambda(\Omega'_j) = \delta_\lambda(\Omega'_j)$ for $\lambda \in \Omega_j$ and $\Omega'_j \subseteq \Omega_j$. Then, $\Omega = \cap_j \Omega_j$ is also measure one according to ν because it is the intersection of an at most countable set of measure one sets, and Ω is left invariant by each γ^j. Since $[\psi]$ and $[\phi]$ are ontologically indistinct, there exists a $\mu \in \Delta_\phi$ such that $\mu(\Omega) > 0$. Now, the measure μ_j, defined by $\mu_j(\Omega') = \int_\Lambda \gamma^j_\lambda(\Omega') d\mu(\lambda)$ for $\Omega' \in \Sigma$, is in $\Delta_{U_j[\phi]U_j^\dagger}$, and $\mu_j(\Omega') = \mu(\Omega')$ for $\Omega' \subseteq \Omega$.

Now, consider a partition of Λ into sets Ω^k. Then,

$$\sum_k \min_j \left[\mu_j(\Omega^k)\right] = \sum_k \min_j \left[\mu_j\left(\Omega^k \cap \Omega\right) + \mu_j\left(\Omega^k \cap (\Lambda\backslash\Omega)\right)\right] \tag{193}$$

$$\geq \sum_k \left(\min_j \left[\mu_j\left(\Omega^k \cap \Omega\right)\right] + \min_j \left[\mu_j\left(\Omega^k \cap (\Lambda\backslash\Omega)\right)\right]\right) \tag{194}$$

$$\geq \sum_k \min_j \left[\mu_j\left(\Omega^k \cap \Omega\right)\right] \tag{195}$$

$$= \sum_k \mu(\Omega^k \cap \Omega) = \mu(\Omega) > 0. \tag{196}$$

Since this holds for all partitions it must also hold for the infimum, and hence $L(\{\mu_j\}) > 0$. □

The next lemma is also used in the proof of the Colbeck-Renner Theorem.

Lemma 9.4. *Let $\{|j\rangle\}_{j=0}^{d-1}$ be a set of $d \geq 3$ orthonormal vectors in $\mathbb{C}^{d'}$ for $d' \geq d$, i.e. not necessarily a complete basis, and consider the vectors*

$$|\psi\rangle = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} |j\rangle\,, \tag{197}$$

and

$$|\phi\rangle = \frac{1}{\sqrt{d-1}} \sum_{j=1}^{d-1} e^{\imath\varphi_j}\,|j\rangle\,. \tag{198}$$

Then, the phases φ_j may be chosen such that $\langle\phi|\psi\rangle$ is real and takes any value between 0 and $\sqrt{\frac{d-1}{d}}$.

Proof. First consider the case where d is odd. In this case, set $\varphi_j = \varphi$ for $1 \leq j \leq \frac{d-1}{2}$ and $\varphi_j = -\varphi$ for $\frac{d+1}{2} \leq j \leq d-1$. Then, $\langle\phi|\psi\rangle = \sqrt{\frac{d-1}{d}}\cos\varphi$, which may take any value between 0 and $\sqrt{\frac{d-1}{d}}$ by varying φ from 0 to $\pi/2$.

When d is even, set $\varphi_1 = 0$, $\varphi_j = \phi$ for $2 \leq j \leq \frac{d}{2}$, and $\varphi_j = -\phi$ for $\frac{d}{2}+1 \leq j \leq d-1$. Then, $\langle\phi|\psi\rangle = \sqrt{\frac{1}{d(d-1)}}(1 + (d-2)\cos\varphi)$. This is equal to $\sqrt{\frac{d-1}{d}}$ for $\varphi = 0$ and is equal to 0 for $\varphi = \pi - \cos^{-1}\left(\frac{1}{d-2}\right)$ and varies continuously in between, so again any positive real number between 0 and $\sqrt{\frac{d-1}{d}}$ can be obtained. □

Lemma 9.5. *Let $[\psi]$ and $[\phi]$ be two nonorthogonal states on $\mathbb{C}^d$ for $d \geq 3$ such that $Tr\left([\phi]\,[\psi]\right) \leq (d-1)/d$. Then, there exist unitaries $\{U_j\}_{j=0}^{d-1}$ such that $U_j\,[\psi]\,U_j^\dagger = [\psi]$ and such that $\{[\phi_j]\}$ is antidistinguishable, where $[\phi_j] = U_j\,[\phi]\,U_j^\dagger$.*

Proof. Let $|\psi\rangle$ and $|\phi\rangle$ be vector representatives of $[\psi]$ and $[\phi]$. Then,

$$|\phi\rangle = \alpha\,|\psi\rangle + \beta\,\left|\psi^\perp\right\rangle\,, \tag{199}$$

where $\alpha = \langle\phi|\psi\rangle$ and $\left\langle\psi^\perp\middle|\psi\right\rangle = 0$. By appropriate choice of the global phase of $|\psi\rangle$, it is possible to make α real and positive.

Let $\{|j\rangle\}_{j=0}^{d-1}$ be a basis in which $|\psi\rangle$ is represented as $|\psi\rangle = \frac{1}{\sqrt{d}}\sum_{j=0}^{d-1}|j\rangle$ and consider the vector

$$|\phi_0\rangle = \sum_{j=0}^{d-1} e^{\imath\varphi_j}\,|j\rangle\,. \tag{200}$$

By Lemma 9.4, it is possible to choose the phases φ_j such that $\langle\phi_1|\psi\rangle = \alpha$, since α is between 0 and $\sqrt{\frac{d-1}{d}}$.

Making this choice of phases, it follows that

$$|\phi_0\rangle = \alpha\,|\psi\rangle + \beta\,\left|\psi_0^{\perp}\right\rangle, \tag{201}$$

where $\left\langle\psi_0^{\perp}\middle|\psi\right\rangle = 0$. Let U be a unitary that acts as $U\,|\psi\rangle = |\psi\rangle$ and $U\left|\psi^{\perp}\right\rangle = \left|\psi_0^{\perp}\right\rangle$. It follows that $|\phi_0\rangle = U\,|\phi\rangle$.

Now, consider the unitary V, where $V\,|j\rangle = |j+1 \mod d\rangle$, which leaves $|\psi\rangle$ invariant. Then, $U_j = V^j U$ also leaves $|\psi\rangle$ invariant for $j = 0, 1, \ldots, d-1$ and the vectors $|\phi_j\rangle = U_j\,|\psi\rangle$ satisfy $\langle j|\phi_j\rangle = 0$, since, by construction, $\langle 0|\phi_0\rangle = 0$. Thus, the states $\{[\phi_j]\}_{j=0}^{d-1}$ are antidistinguished by the measurement $\{[j]\}_{j=0}^{d-1}$ as required. □

Proof of Theorem 9.2. Let $\Xi = (\Lambda, \Sigma, \Delta, \xi, \Gamma)$ be an ontological model of the PMT fragment $\mathfrak{F}$ to be constructed. For the purposes of contradiction, assume that two states $[\psi], [\phi] \in \mathcal{P}$ satisfying $\mathrm{Tr}\left([\phi]\,[\psi]\right) \leq (d-1)/d$ are ontologically indistinct in Ξ and let $[\psi']$ be a state that satisfies ontic indifference in Ξ. Without loss of generality, we may assume that $[\psi] = [\psi']$. If not, let U be a unitary such that $[\psi'] = U\,[\psi]\,U^{\dagger}$ and assume $U \in \mathcal{T}$. Then the states $[\psi']$ and $[\phi'] = U\,[\phi]\,U^{\dagger}$ satisfy $\mathrm{Tr}\left([\phi']\,[\psi']\right) \leq (d-1)/d$ because the inner product is preserved under unitary transformations. Also, by Corollary 8.4, if $[\psi]$ and $[\phi]$ are ontologically indistinct then so are $[\psi']$ and $[\phi']$ so it is sufficient to derive a contradiction for these two states.

By Lemma 9.5, there exist unitaries $\{U_j\}_{j=0}^{d-1}$ that leave $[\psi]$ invariant such that $\left\{U_j\,[\phi]\,U_j^{\dagger}\right\}_{j=0}^{d-1}$ is antidistinguishable. Assume that each U_j is in $\mathcal{T}$ and that the antidistinguishing measurement is in $\mathcal{M}$. Then, by Theorem 6.5, for every set of probability measures $\{\mu_j\}_{j=0}^{d-1}$ with $\mu_j \in \Delta_{U_j[\phi]U_j^{\dagger}}$, $L(\{\mu_j\}) = 0$, but by Lemma 9.3 there must exist a set such that $L(\{\mu_j\}) > 0$. Thus, we have a contradiction so $[\psi]$ and $[\phi]$ must be ontologically distinct. □

10 The Colbeck-Renner Theorem

Our final ψ-ontology theorem is due to Colbeck and Renner [119]. They originally proved a ψ-ontology theorem as a byproduct of a broader no-go theorem that aimed to rule out any ontological model of quantum theory in which more precise predictions can be made given knowledge of the ontic state than if you only know the quantum state [50,120,121]. This broader theorem is quite tricky to understand, both conceptually and technically, and its assumptions have been criticised [122–124]. Fortunately, Colbeck and Renner have recently put forward a simplified argument for ψ-ontology based on similar ideas that bypasses much of the complexity [119]. At the time of writing, this argument has not yet appeared in print, so I present my own reconstruction of it here.

The Colbeck-Renner Theorem is based on a locality assumption known as *parameter independence*. The condition of Bell locality, discussed in §5.5, can be decomposed as the conjunction of parameter independence and another assumption called *outcome independence*. Roughly speaking, parameter independence says that the measurement outcome at Bob's side should not depend on Alice's choice of measurement and outcome independence says that the measurement outcome at Bob's side should not depend on Alice's measurement outcome. For a typical Bell inequality experiment, it is in principle possible to reproduce the quantum predictions by violating outcome independence and leaving parameter independence intact, so the Colbeck-Renner Theorem rules out the class of ψ-epistemic models that opt for this resolution to the conundrum presented by Bell's Theorem. Nevertheless, Bell's Theorem should lead us to be skeptical of imposing locality assumptions on ontological models. Indeed, many notable realist approaches to quantum theory, such as de Broglie-Bohm theory, violate parameter independence. As well as ruling out ψ-epistemic models, parameter independence rules out a wide class of viable ψ-ontic models, so it is perhaps an unreasonably strong assumption for a ψ-ontology theorem.

It should be noted that Colbeck and Renner claim that their theorem is not directly based on parameter independence, but rather on a broader assumption that they call "free choice", aimed at capturing the idea

that certain variables, such as the settings of preparation and measurement devices, are chosen freely by the experimenter and not constrained by other parameters of the system. Parameter independence is just one consequence of their free choice assumption. However, whether their assumption truly captures the meaning of making a free choice is one of the most controversial aspects of the Colbeck-Renner argument. My own opinion is that, whilst some aspects of free choice are captured by the Colbeck-Renner assumption, it is at least partly a locality assumption as well because it implies parameter independence. Therefore, I prefer to avoid the free choice controversy and just assume parameter independence directly.

Understanding the Colbeck-Renner Theorem requires quite a bit of background material, so the discussion is broken down into four sections. §10.1 gives a formal definition of parameter independence and shows how, given an ontological model for a bipartite system, it can be used to derive an ontological model for the reduced states on a subsystem. §10.2 introduces chained Bell measurements, which are the main technical tool used in proving the Colbeck-Renner Theorem. §10.3 discusses the equiprobability theorem, which shows that, in a model that satisfies parameter independence, all ontic states associated with a maximally entangled state must assign equal probability to all the outcomes of some local measurement in an orthonormal basis. Finally, §10.4 leverages the equiprobability theorem to prove the main result.

10.1 Parameter Independence

As in Bell's Theorem, the PM fragment of interest for the Colbeck-Renner Theorem is a product measurement fragment, consisting of states and local measurements on a bipartite system (see Definition 5.14). §5.5 discussed the concept of a conditional fragment on Bob's system that results from viewing Alice's measurements as preparation procedures for states on Bob's system. Given an ontological model for a product measurement fragment that satisfies Bell locality, it is possible to derive a model for the conditional fragment from it. Here, we are concerned with a smaller fragment that results from tracing out one of the systems, but not post-selecting on measurement outcomes.

Definition 10.1. Let $\mathfrak{F}_{AB} = \langle \mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{P}_{AB}, \mathcal{M}_A \times \mathcal{M}_B \rangle$ be a product measurement fragment. The *marginal fragment on* A is $\mathfrak{F}_A = \langle \mathcal{H}_A, \mathcal{P}_A, \mathcal{M}_A \rangle$, where $\mathcal{P}_A$ consists of all states of the form $\rho_A = \mathrm{Tr}_B(\rho_{AB})$ for $\rho_{AB} \in \mathcal{P}_{AB}$. The marginal fragment on B is defined similarly, with the roles of A and B interchanged.

In general, an ontological model $\Xi_{AB} = (\Lambda_{AB}, \Sigma_{AB}, \Delta_{AB}, \xi_{AB})$ for a product measurement fragment $\mathfrak{F}_{AB} = \langle \mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{P}_{AB}, \mathcal{M}_A \times \mathcal{M}_B \rangle$ specifies response functions $\xi_{AB}(M_A, M_B, E \otimes F) = \xi_{E\otimes F}^{M_A,M_B}$ for every $M_A \in \mathcal{M}_A$, $M_B \in \mathcal{M}_B$, $E \in M_A$ and $F \in M_B$. If the model reproduces the quantum predictions, then for all $\rho_{AB} \in \mathcal{P}_{AB}$,

$$\int_{\Lambda_{AB}} \xi_{E\otimes F}^{M_A,M_B}(\lambda) d\mu(\lambda) = \mathrm{Tr}_{AB}(E \otimes F \rho_{AB}), \tag{202}$$

for every $\mu \in \Delta_{AB}(\rho)$. Summing both sides of this equation over $F \in M_B$ gives

$$\int_{\Lambda_{AB}} \xi_{E}^{M_A,M_B}(\lambda) d\mu(\lambda) = \mathrm{Tr}_{A}(E \rho_{A}), \tag{203}$$

where $\rho_A = \mathrm{Tr}_B(\rho_{AB})$ and $\xi_E^{M_A,M_B} = \sum_{F\in M_B} \xi_{E\otimes F}^{M_A,M_B}$. For a fixed M_B, $\{\xi_E^{M_A,M_B}\}_{E\in M_A}$ is a set of response functions for the measurement M_A of the marginal fragment. Further, if μ is used to represent ρ_A, this choice of response functions reproduces the quantum predictions. The only difficulty is that $\xi_E^{M_A,M_B}$ depends on M_B as well as M_A, so there is an ambiguity about which choice of M_B should be used to model the marginal fragment. The assumption of parameter independence removes this ambiguity.

Definition 10.2. An ontological model $\Xi_{AB} = (\Lambda_{AB}, \Sigma_{AB}, \Delta_{AB}, \xi_{AB})$ of a product measurement fragment $\mathfrak{F}_{AB} = \langle \mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{P}_{AB}, \mathcal{M}_A \times \mathcal{M}_B \rangle$ satisfies *parameter independence* if, for all $M_A \in \mathcal{M}_A$ and $M_B \in \mathcal{M}_B$, $\sum_{E\in M_A} \xi_{E\otimes F}^{M_A,M_B}$ does not depend on M_A and $\sum_{F\in M_B} \xi_{E\otimes F}^{M_A,M_B}$ does not depend on M_B.

This assumption can be recast in a more familiar form by recalling the interpretation of a response function $\xi_E^M(\lambda) = P(E|M,\lambda)$ as the conditional probability that the outcome E is obtained in measurement

M when the ontic state is λ. Then, parameter independence reads

$$P(E|M_A, M_B, \lambda) = P(E|M_A, \lambda) \tag{204}$$
$$P(F|M_A, M_B, \lambda) = P(F|M_B, \lambda). \tag{205}$$

Parameter independence is a locality assumption that says that Alice's measurement outcome should not depend on Bob's choice of measurement and vice versa. It is a weaker requirement than Bell locality, as it can be shown that the conjunction of parameter independence and *outcome independence*, which reads

$$P(E|M_A, M_B, F, \lambda) = P(E|M_A, M_B, \lambda) \tag{206}$$
$$P(F|M_A, E, M_B, \lambda) = P(F|M_A, M_B, \lambda), \tag{207}$$

is equivalent to Bell locality [125, 126]. For present purposes, the point of this is that, whilst the full Bell locality condition is needed to derive a well-defined ontological model for conditional fragments, parameter independence is sufficient to have unambiguous models for marginal fragments.

Definition 10.3. Let $\mathfrak{F}_{AB} = \langle \mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{P}_{AB}, \mathcal{M}_A \times \mathcal{M}_B \rangle$ be a product measurement fragment and let $\Xi_{AB} = (\Lambda_{AB}, \Sigma_{AB}, \Delta_{AB}, \xi_{AB})$ be an ontological model of it that satisfies parameter independence. The *marginal model on A*, $\Xi_A = (\Lambda_{AB}, \Sigma_{AB}, \Delta_A, \xi_A)$ is an ontological model for the marginal fragment $\mathfrak{F}_A = \langle \mathcal{H}_A, \mathcal{P}_A, \mathcal{M}_A \rangle$, where, for $\rho_A \in \mathcal{P}_A$,

$$\Delta_A(\rho_A) = \cup_{\{\rho_{AB} \in \mathcal{P}_{AB} | \mathrm{Tr}_B(\rho_{AB}) = \rho_A\}} \Delta_{AB}(\rho_{AB}), \tag{208}$$

and, for $M_A \in \mathcal{M}_A$, $E \in M_A$

$$\xi_A(M_A, E) = \sum_{F \in M_B} \xi_{AB}(M_A, M_B, E \otimes F), \tag{209}$$

for all $M_B \in \mathcal{M}_B$.

Eq. (203) implies that if Ξ_{AB} reproduces the quantum predictions for $\mathfrak{F}_{AB}$ then the marginal model Ξ_A reproduces the quantum predictions for the marginal fragment $\mathfrak{F}_A$.

10.2 Chained Bell measurements

The main engine of the proof of the Colbeck-Renner Theorem is the *chained Bell measurements*. These were originally developed in the context of a proof of Bell's Theorem that involved chaining together several tests of the Clauser-Horne-Shimony-Holt inequality [127], resulting in the chained Bell inequalities [128,129], which explains the terminology.

We are interested in the product measurement fragment $\mathfrak{F} = \langle \mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{P}_{AB}, \mathcal{M}_A \times \mathcal{M}_B \rangle$, where $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^2$ and both $\mathcal{M}_A$ and $\mathcal{M}_B$ contain N measurements in orthonormal bases. For later convenience, Alice's measurements are labelled by even integers and Bob's by the odd integers , so we have $M_A^a \in \mathcal{M}_A$ for $a \in E_N = \{0, 2, \ldots, 2N-2\}$ and $M_A^b \in \mathcal{M}_B$ for $b \in O_N = \{1, 3, \ldots, 2N-1\}$. The measurements are described by pairs of orthogonal rank-1 projectors $M_A^a = \{[\phi_0^a]_A, [\phi_1^a]_A\}$, $M_B^b = \{[\phi_0^b]_B, [\phi_1^b]_B\}$ onto the orthonormal bases $\{|\phi_0^a\rangle_A, |\phi_1^a\rangle_A\}$ and $\{|\phi_0^b\rangle_B, |\phi_1^b\rangle_B\}$. If the state ρ_{AB} is prepared, Alice measures M_A^a, and Bob measures M_B^b, then the outcome probabilities are

$$\mathrm{Prob}\left([\phi_j^a], [\phi_k^b] | M_A^a, M_B^b, \rho_{AB}\right) = \mathrm{Tr}_{AB}\left([\phi_j^a] \otimes [\phi_k^b] \rho_{AB}\right). \tag{210}$$

For later convenience, we introduce two random variables X and Y that take values 0 and 1 to represent Alice and Bob's outcomes and write

$$\mathrm{Prob}(X = j, Y = k | a, b, \rho_{AB}) = \mathrm{Prob}\left([\phi_j^a], [\phi_k^b] | M_A^a, M_B^b, \rho_{AB}\right). \tag{211}$$

Now consider the following correlation measure for these outcome probabilities.

$$I_N(\rho_{AB}) = \text{Prob}(X = Y|0, 2N-1, \rho_{AB}) + \sum_{\{(a,b)|a\in E_N, b\in O_N, |a-b|=1\}} \text{Prob}(X \neq Y|a, b, \rho_{AB}). \tag{212}$$

The structure of this measure is illustrated in Fig. 16. Note that smaller values of this measure indicate that the measurement outcomes are more highly correlated.

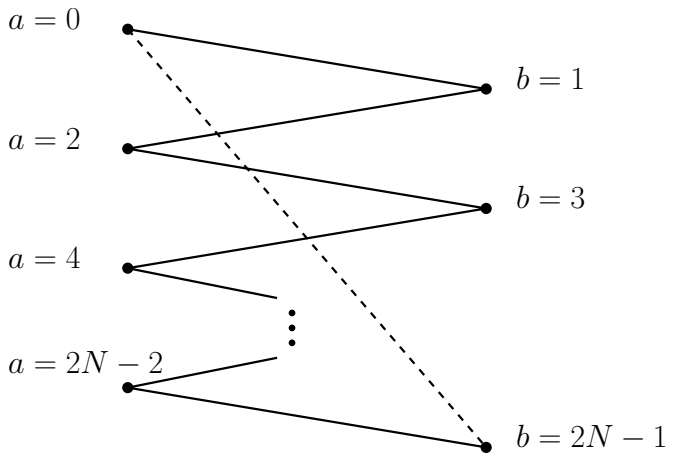


Figure 16: The structure of the correlation measure given in Eq. (212). A solid line between a and b represents the probability that the outcomes of M_a^A and M_b^A are different and a dotted line represents the probability that they are the same. The correlation measure is then the sum of the terms represented by each line.

Now, let $\{|0\rangle, |1\rangle\}$ be an orthonormal basis for $\mathbb{C}^2$, and consider the quantum state $[\Phi^+]_{AB}$, where

$$\left|\Phi^+\right\rangle_{AB} = \frac{1}{\sqrt{2}}\left(|00\rangle_{AB} + |11\rangle_{AB}\right). \tag{213}$$

Let $\vartheta_j^a = \left(\frac{a}{2N} + j\right)\pi$, and suppose that the bases for Alice and Bob's measurements are given by

$$\left|\phi_j^a\right\rangle_A = \cos\left(\frac{\vartheta_j^a}{2}\right)|0\rangle_A + \sin\left(\frac{\vartheta_j^a}{2}\right)|1\rangle_A \tag{214}$$

$$\left|\phi_k^b\right\rangle_B = \cos\left(\frac{\vartheta_k^b}{2}\right)|0\rangle_A + \sin\left(\frac{\vartheta_k^b}{2}\right)|1\rangle_B. \tag{215}$$

These measurements are illustrated on the Bloch sphere in Fig. 17.

With these particular choices, the outcome probabilities are

$$\text{Prob}\left(X = j, Y = k|a, b, \left[\Phi^+\right]_{AB}\right) = \frac{1}{2}\cos^2\left(\left(\frac{a-b}{2N} + j - k\right)\frac{\pi}{2}\right), \tag{216}$$

from which it is straightforward to compute that

$$I_N(\left[\Phi^+_{AB}\right]) = 2N\sin^2\frac{\pi}{4N} \leq \frac{\pi^2}{8N}, \tag{217}$$

so that the correlation measure tends to zero as $N \to \infty$.

Without going through the whole calculation, it is fairly easy to see why this should be the case. For the state $[\Phi^+]_{AB}$, when Alice performs the measurement M_A^a and obtains the outcome j, Bob's state gets

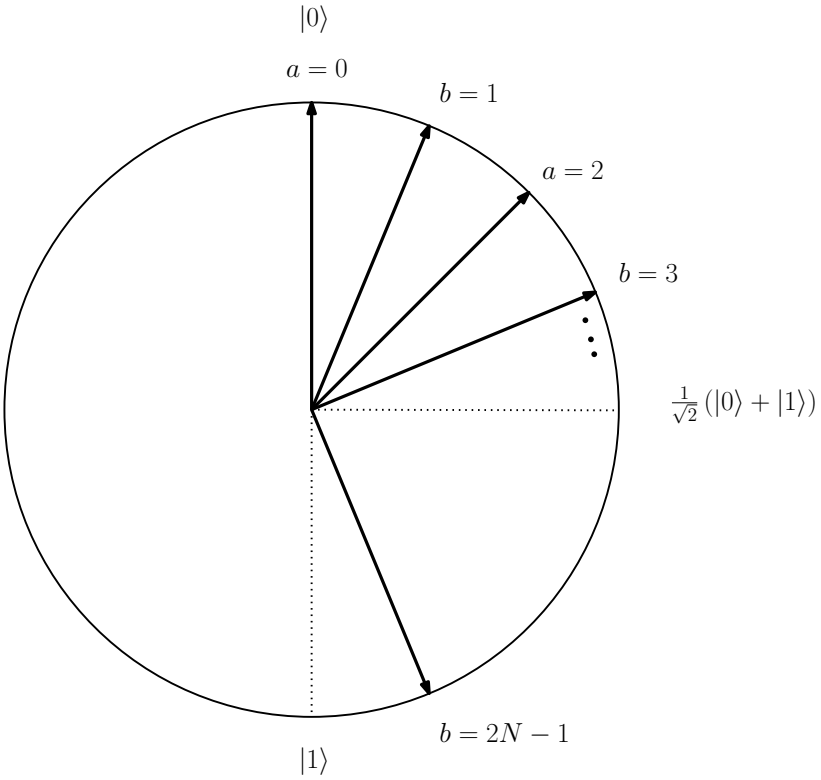


Figure 17: Alice and Bob's measurement bases represented on a great circle of the Bloch sphere.

updated to $\left[\phi_j^a\right]_B$. The condition $|a-b| = 1$ in the sum in Eq. (212) means that, in the limit of large N, the terms in this sum involve basis states for Bob that are very close to Alice's basis states. Thus, since Bob's system collapses to the basis state that Alice obtained in her measurement, the probability of getting the same outcome is very high, and hence $P(j \neq k|a, b, [\Phi^+]_{AB})$ is close to zero. Similarly, the state $\left|\phi_j^{2N-1}\right\rangle_B$ is almost orthogonal to $\left|\phi_j^0\right\rangle_B$, so the probability of getting equal outcomes in the first term is also close to zero.

10.3 The equiprobability theorem

The aim of this section is to prove the main technical result that the Colbeck-Renner Theorem relies upon. We first deal with the case where Alice and Bob's Hilbert spaces are two dimensional, and then extend this to higher dimensions.

Theorem 10.4. *Let $\mathfrak{F}_{AB} = \langle \mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{P}_{AB}, \mathcal{M}_A \times \mathcal{M}_B \rangle$ be a product measurement fragment where $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^2$, $\mathcal{P}_{AB}$ contains the maximally entangled state $[\Phi^+]_{AB}$, where*

$$\left|\Phi^+\right\rangle_{AB} = \frac{1}{\sqrt{2}}\left(|00\rangle_{AB} + |11\rangle_{AB}\right), \tag{218}$$

for some orthonormal basis $\{|0\rangle, |1\rangle\}$, $\mathcal{M}_A$ contains measurements in every orthonormal basis on $\mathcal{H}_A$ and $\mathcal{M}_B$ contains measurements in every orthonormal basis on $\mathcal{H}_B$.

Let $\Xi_{AB} = (\Lambda_{AB}, \Sigma_{AB}, \Delta_{AB}, \xi_{AB})$ be an ontological model of $\mathfrak{F}_{AB}$ that reproduces the quantum predictions and satisfies parameter independence.

Then, for any $\mu \in \Delta_{AB}([\Phi^+]_{AB})$, there exists a set $\Omega \in \Sigma_{AB}$ such that $\mu(\Omega) = 1$ and, for the measurement $M_A = \{[0]_A, [1]_A\}$, the marginal model on A satisfies $\xi_0^{M_A}(\lambda) = \xi_1^{M_A}(\lambda) = \frac{1}{2}$ for $\lambda \in \Omega$.

Before proving this theorem, it is instructive to see how it can be used to establish ontological distinctness for some pairs of states. Consider the state $[0]_A \otimes [0]_B$. To reproduce the quantum predictions, the marginal model on A must predict that the $[1]_A$ outcome has zero probability. This means that any $\nu \in \Delta_{AB}([0]_A \otimes [0]_B)$ must assign zero measure to any set $\Omega \in \Sigma_{AB}$ for which $\xi_1^{M_A}$ is nonzero. However, according to Theorem 10.4, for any $\mu \in \Delta_{AB}([\Phi^+]_{AB})$ there is such a set that is measure one according to μ. Thus, for any $\mu \in \Delta_{AB}([\Phi^+]_{AB})$ and any $\nu \in \Delta_{AB}([0]_A \otimes [0]_B)$, there is a set that is measure 1 according to μ and measure zero according to ν, so $[\Phi^+]_{AB}$ and $[0]_A \otimes [0]_B$ are ontologically distinct. The Colbeck-Renner Theorem generalizes this argument to arbitrary pairs of pure states.

The proof of Theorem 10.4 relies on the following two lemmas.

Lemma 10.5. *Let X and Y be random variables that take values in the set $\{0, 1, \ldots, d-1\}$ and let $P(X), P'(X), Q(Y), Q'(Y)$ be probability distributions over them. For $0 \leq p \leq 1$, let*

$$P''(X) = pP(X) + (1-p)P'(X) \tag{219}$$
$$Q''(Y) = pQ(Y) + (1-p)Q'(Y). \tag{220}$$

Then,

$$D\left(P''(X), Q''(Y)\right) \leq pD\left(P(X), Q(Y)\right) + (1-p)D\left(P'(X), Q'(Y)\right), \tag{221}$$

where

$$D\left(P(X), Q(Y)\right) = \frac{1}{2}\sum_{j=0}^{d-1} \left|P(X=j) - Q(Y=j)\right|, \tag{222}$$

is the variational distance.

Proof.

$$D\left(P''(X), Q''(Y)\right) = \frac{1}{2}\sum_{j=0}^{d-1} \left|p\left(P(X=j) - Q(Y=j)\right) + (1-p)\left(P'(X=j) - Q'(X=j)\right)\right| \tag{223}$$
$$\leq \frac{1}{2}\sum_{j=0}^{d-1} \left(\left|p\left(P(X=j) - Q(Y=j)\right)\right| + \left|(1-p)\left(P'(X=j) - Q'(X=j)\right)\right|\right) \tag{224}$$
$$= p\frac{1}{2}\sum_{j=0}^{d-1} \left|P(X=j) - Q(Y=j)\right| + (1-p)\frac{1}{2}\sum_{j=0}^{d-1} \left|P'(X=j) - Q'(Y=j)\right| \tag{225}$$
$$= pD\left(P(X), Q(Y)\right) + (1-p)\left(P'(X), Q'(Y)\right), \tag{226}$$

where the second step follows from the triangle inequality. □

Lemma 10.6. *Let X and Y be random variables that take values in the set $\{0, 1, \ldots, d-1\}$ and let $P(X)$ and $P(Y)$ be probability distributions over them. Let $P(X, Y)$ be a joint distribution such that $P(X) = \sum_k P(X, Y=k)$ and $P(Y) = \sum_j P(X=j, Y)$. Then,*

$$D\left(P(X), P(Y)\right) \leq P(X \neq Y), \tag{227}$$

where D is the variational distance.

Proof. Let $P_{\neq}(X, Y)$ be the conditional probability distribution of X and Y given that $X \neq Y$, let $P_{=}(X, Y)$ be the conditional probability distribution of X and Y given that $X = Y$, let $p_{=} = P(X = Y)$ and let $p_{\neq} = P(X \neq Y)$. Then, by the law of total probability

$$P(X, Y) = p_{\neq}P_{\neq}(X, Y) + p_{=}P_{=}(X, Y). \tag{228}$$

Taking the marginals of this equation gives

$$P(X) = p_{\neq} P_{\neq}(X) + p_{=} P_{=}(X) \tag{229}$$
$$P(Y) = p_{\neq} P_{\neq}(Y) + p_{=} P_{=}(Y). \tag{230}$$

Then, by Lemma 10.5,

$$D(P(X), P(Y)) \leq p_{\neq} D(P_{\neq}(X), P_{\neq}(Y)) + p_{=} D(P_{=}(X), P_{=}(Y)) \tag{231}$$
$$\leq p_{\neq} = P(X \neq Y). \tag{232}$$

□

Proof of Theorem 10.4. Consider the chained Bell measurements $M_A^a = \{[\phi_0^a]_A, [\phi_1^a]_A\}$ and $M_B^b = \left\{[\phi_0^b]_B, [\phi_1^b]_B\right\}$ defined in §10.2 and note that $M_A^0 = \{[0]_A, [1]_A\}$. From Eq. (217), the correlation measure I_N defined in Eq. (212) satisfies $I_N([\Phi^+]_{AB}) \leq \frac{\pi^2}{8N}$. We can also define a similar correlation measure within the ontological model, given by

$$I_N(\lambda) = P(X = Y|0, 2N-1, \lambda) + \sum_{\{(a,b)|a\in E_N, b\in O_N, |a-b|=1\}} P(X \neq Y|a, b, \lambda), \tag{233}$$

where

$$P(X = j, Y = k|a, b, \lambda) = \xi^{M_A^a, M_B^b}_{[\phi_j^a]_A \otimes [\phi^b]_B}(\lambda). \tag{234}$$

Now, by parameter independence

$$P(X = j|a, b, \lambda) = P(X = j|a, \lambda) = \xi^{M_A^a}_{[\phi_j^a]}(\lambda), \tag{235}$$

and

$$P(Y = k|a, b, \lambda) = P(Y = k|b, \lambda) = \xi^{M_B^b}_{[\phi_k^b]}(\lambda), \tag{236}$$

where $\xi^{M_A^a}_{[\phi_j^a]}$ and $\xi^{M_B^b}_{[\phi_k^b]}$ are the response functions of the marginal models on A and B respectively.

Let $\tilde{X} = (X + 1) \mod 2$, so that if $X = 0$ then $\tilde{X} = 1$ and vice versa. Then, $P(X = Y|0, 2N-1, \lambda) = P(\tilde{X} \neq Y|0, 2N-1, \lambda)$. Now, by parameter independence and Lemma 10.6,

$$I_N(\lambda) \geq D(P(\tilde{X}|0, \lambda), P(Y|2N-1, \lambda)) + \sum_{\{(a,b)|a\in E_N, b\in O_N, |a-b|=1\}} D(P(X|a, \lambda), P(Y|b, \lambda)) \tag{237}$$
$$\geq D(P(\tilde{X})|0, \lambda), P(X|0, \lambda)), \tag{238}$$

where the second step follows from the triangle inequality for the variational distance and the structure of the correlation measure shown in Fig. 16.

Since the ontological model reproduces the quantum predictions we have, for any $\mu \in \Delta_{AB}([\Phi^+]_{AB})$, $\int_\Lambda I_N(\lambda) d\mu(\lambda) = I_N([\Phi^+]_{AB})$ and in the limit $N \to \infty$, this is equal to 0. Therefore, using Eq. (238) gives

$$\int_\Lambda D(P(\tilde{X}|0, \lambda), P(X|0, \lambda)) d\mu(\lambda) = 0. \tag{239}$$

Using Eq. (235), this is equivalent to

$$\int_\Lambda \left| \xi_0^{M_A^0}(\lambda) - \xi_1^{M_A^0}(\lambda) \right| d\mu(\lambda) = 0. \tag{240}$$

This can only happen if there exists a set $\Omega \in \Sigma_{AB}$ such that $\mu(\Omega) = 1$ and $\xi_0^{M_A^0}(\lambda) = \xi_1^{M_A^0}(\lambda)$ for $\lambda \in \Omega$. But since $\xi_0^{M_A^0}(\lambda) + \xi_1^{M_A^0}(\lambda) = 1$, this means that $\xi_0^{M_A^0}(\lambda) = \xi_1^{M_0}(\lambda) = \frac{1}{2}$ for $\lambda \in \Omega$. □

The next step is to generalize this result to arbitrary dimensions.

Theorem 10.7. *Let $\mathfrak{F}_{AB} = \langle \mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{P}_{AB}, \mathcal{M}_A \times \mathcal{M}_B \rangle$ be a product measurement fragment where $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^d$, $\mathcal{P}_{AB}$ contains the maximally entangled state $[\Phi^+]_{AB}$, where*

$$\left|\Phi^+\right\rangle_{AB} = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} |jj\rangle_{AB}\,, \tag{241}$$

for some orthonormal basis $\{|j\rangle\}_{j=0}^{d-1}$, $\mathcal{M}_A$ contains measurements in every orthonormal basis on $\mathcal{H}_A$, and $\mathcal{M}_B$ contains measurements in every orthonormal basis on $\mathcal{H}_B$.

Let $\Xi_{AB} = (\Lambda_{AB}, \Sigma_{AB}, \Delta_{AB}, \xi_{AB})$ be an ontological model of $\mathfrak{F}_{AB}$ that reproduces the quantum predictions and satisfies parameter independence.

Then, for any $\mu \in \Delta_{AB}([\Phi^+]_{AB})$, there exists a set $\Omega \in \Sigma_{AB}$ such that $\mu(\Omega) = 1$ and, for the measurement $M_A = \{[j]_A\}_{j=0}^{d-1}$, the marginal model on A satisfies $\xi_j^{M_A}(\lambda) = \frac{1}{d}$ for $\lambda \in \Omega$.

Proof. The chained Bell measurements can be extended to higher dimensions by defining $\left[\phi_j^a\right]_A \left[\phi_k^b\right]_B$ as before for $j,k \in \{0,1\}$ and setting $\left[\phi_j^a\right]_A = [j]_A$, $\left[\phi_k^b\right] = [k]_B$ for $2 \leq j,k \leq d-1$. Now, define two random variables X and Y taking values $0, 1, \ldots d-1$ and, as before, set

$$\text{Prob}\left(X = j, Y = k | a, b, \left[\Phi^+\right]_{AB}\right) = \text{Tr}_{AB}\left(\left[\phi_j^a\right]_A \otimes \left[\phi_k^b\right]_B \left[\Phi^+\right]_{AB}\right). \tag{242}$$

Now, condition these probabilities on the event $C = \{X \in \{0,1\} \cap Y \in \{0,1\}\}$, i.e. define

$$\text{Prob}\left(X = j, Y = k | a, b, \left[\Phi^+\right]_{AB}, C\right) = \frac{\text{Prob}\left(X = j, Y = k | a, b, [\Phi^+]_{AB}\right)}{\sum_{j',k'=0}^{1} \text{Prob}\left(X = j, Y = k | a, b, [\Phi^+]_{AB}\right)}, \tag{243}$$

for $j,k \in \{0,1\}$. Straightforward calculation shows that $\text{Prob}\left(X = j, Y = k | a, b, [\Phi^+]_{AB}, C\right)$ is equal to the $d = 2$ case of the unconditioned probability distribution.

Similarly, at the ontological level, for $j,k \in \{0,1\}$, we can define the conditional probability distribution

$$P(X = j, Y = k | a, b, \lambda, C) = \frac{1}{N(\lambda)} \xi^{M_A^a, M_B^b}_{\left[\phi_j^a\right] \otimes \left[\phi_k^b\right]}(\lambda), \tag{244}$$

where

$$N(\lambda) = \sum_{j,k=0}^{1} \xi^{M_A^a, M_B^b}_{\left[\phi_j^a\right] \otimes \left[\phi_k^b\right]}(\lambda). \tag{245}$$

Given that the model reproduces the quantum predictions, we must have

$$\text{Prob}\left(X = j, Y = k | a, b, \left[\Phi^+\right]_{AB}, C\right) = \int_{\Lambda_{AB}} \frac{1}{N(\lambda)} \xi^{M_A^a, M_B^b}_{\left[\phi_j^a\right] \otimes \left[\phi_k^b\right]}(\lambda) d\mu(\lambda), \tag{246}$$

for $j,k \in \{0,1\}$ and $\mu \in \Delta_{AB}([\Phi^+]_{AB})$. Then, Theorem 10.4 implies that $\frac{1}{N(\lambda)} \xi_0^{M_A^0}(\lambda) = \frac{1}{N(\lambda)} \xi_1^{M_A^0}(\lambda)$ for a measure one set $\Omega_{0,1}$ according to μ and hence $\xi_0^{M_A^0}(\lambda) = \xi_1^{M_A^0}(\lambda)$ for $\lambda \in \Omega_{0,1}$.

Now, the same argument can be repeated by placing the nontrivial part of the chained Bell measurements on a different subspace, e.g. on the subspace spanned by $|1\rangle$ and $|2\rangle$ rather than the subspace spanned by $|0\rangle$ and $|1\rangle$ and conditioning on $\{X \in \{1,2\} \cap Y \in \{1,2\}\}$ rather than $\{X \in \{0,1\} \cap Y \in \{0,1\}\}$. In total, this shows that, for all $j,k \in \{0,1,\ldots,d-1\}$ there exists sets $\Omega_{j,k}$ that are measure one according to μ and such that $\xi_j^{M_A^0}(\lambda) = \xi_k^{M_A^0}(\lambda)$ for $\lambda \in \Omega_{j,k}$. Now, $\Omega = \cap_{j,k=0}^{d-1} \Omega_{j,k}$ is also measure one according to μ because it is the intersection of a finite collection of measure one sets. This implies that, for every j and k, we have $\xi_j^{M_A^0}(\lambda) = \xi_k^{M_A^0}(\lambda)$ for $\lambda \in \Omega$. Finally, $\sum_{j=0}^{d-1} \xi_j^{M_A^0}(\lambda) = 1$ implies that $\xi_j^{M_A^0}(\lambda) = \frac{1}{d}$ for $\lambda \in \Omega$. □

10.4 The main result

Theorem 10.8. *Let $\mathfrak{F}_{AB} = \langle \mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{P}_{AB}, \mathcal{M}_A \times \mathcal{M}_B \rangle$ be a product measurement fragment where $\mathcal{H}_A = \mathcal{H}_B = \mathbb{C}^d$ for $d \geq 3$, $\mathcal{P}_{AB}$ contains all pure states on $\mathcal{H}_A \otimes \mathcal{H}_B$, $\mathcal{M}_A$ contains all measurements in orthonormal bases on $\mathcal{H}_A$, and $\mathcal{M}_B$ contains all measurements in orthonormal bases on $\mathcal{H}_B$.*

Let $\Xi_{AB} = (\Lambda_{AB}, \Sigma_{AB}, \Delta_{AB}, \xi_{AB})$ be an ontological model of $\mathfrak{F}_{AB}$ that satisfies parameter independence, preserves ontological indistinctness with respect to all unitaries on $\mathcal{H}_A \otimes \mathcal{H}_B$, and reproduces the quantum predictions. Then, any pair of pure states $[\psi]_{AB}, [\phi]_{AB} \in \mathcal{P}_{AB}$ that satisfies $Tr_{AB}([\psi]_{AB}[\phi]_{AB}) \leq \frac{d-1}{d}$ is ontologically distinct in Ξ_{AB}.

Before proving this theorem, note that, as with Hardy's Theorem, assuming that appending ancillas preserves ontological indistinctness allows this to be converted into a full blown ψ-ontology theorem. In this case, there are two issues to deal with. The first is that the the condition on the inner product is not satisfied for all states and the second is that not all fragments have the tensor product structure of a product measurement fragment. To deal with the first issue, consider a fragment $\mathfrak{F}_A = \langle \mathcal{H}_A, \mathcal{P}_A, \mathcal{M}_A \rangle$ where $\mathcal{H}_A = \mathbb{C}^d$, $\mathcal{P}_A$ contains all pure states on $\mathcal{H}_A$, and $\mathcal{M}_A$ contains all measurements in orthonormal bases on $\mathcal{H}_A$. A given pair of pure states, $[\psi]_A$ and $[\phi]_A$, either satisfy the condition $\mathrm{Tr}_A([\phi]_A[\psi]_A) \leq \frac{d-1}{d}$ or they do not. If they do not then we can introduce a product fragment $\mathfrak{F}_{AA'} = \langle \mathcal{H}_A \otimes \mathcal{H}_{A'}, \mathcal{P}_{AA'}, \mathcal{M}_{AA'} \rangle$, where we again assume that $\mathcal{P}_{AA'}$ contains all pure states and $\mathcal{M}_{AA'}$ contains all measurements in orthonormal bases. The dimension of $\mathcal{H}_{A'}$ is chosen such that $[\psi]_A \otimes [0]_{A'}$ and $[\phi]_A \otimes [0]_{A'}$ satisfy the inner product condition, where $[0]_{A'}$ is a fixed pure state. If we assume that an ontological model of $\mathfrak{F}_{AA'}$ must preserve ontological indistinctness with respect to an ontological model of $\mathfrak{F}_A$, then proving the ontological distinctness of $[\psi]_A \otimes [0]_{A'}$ and $[\phi]_A \otimes [0]_{A'}$ is enough to establish the ontological indistinctness of $[\psi]_A$ and $[\phi]_A$ in the original model.

To deal with the issue of tensor product structure, first of all, to avoid cluttered notation, relabel AA' as A. Then, we can take $\mathfrak{F}_B = \langle \mathcal{H}_B, \mathcal{P}_B, \mathcal{M}_B \rangle$ to be a second copy of $\mathfrak{F}_A$, i.e. it has the same Hilbert space, states and measurements, but just a different system label. Then, form a product measurement fragment $\mathfrak{F}_{AB} = \langle \mathcal{H}_A \otimes \mathcal{H}_B, \mathcal{P}_{AB}, \mathcal{M}_A \times \mathcal{M}_B \rangle$ with factors $\mathfrak{F}_A$ and $\mathfrak{F}_B$, where $\mathcal{P}_{AB}$ contains all pure states on $\mathcal{H}_A \otimes \mathcal{H}_B$. If $\mathrm{Tr}_A([\phi]_A[\phi]_A) \leq \frac{d-1}{d}$ then $\mathrm{Tr}_{AB}([\phi]_A \otimes [0]_B[\psi]_A \otimes [0]_B) = \mathrm{Tr}_A([\phi]_A[\phi]_A) \leq \frac{d-1}{d}$ for any fixed pure state $[0]_B \in \mathcal{P}_B$. Assuming that an ontological model of $\mathfrak{F}_{AB}$ must preserve ontological indistinctness with respect to an ontological model of $\mathfrak{F}_A$ means that proving ontological distinctness of $[\psi]_A \otimes [0]_B$ and $[\phi]_A \otimes [0]_B$ is enough to prove the ontological distinctness of $[\psi]_A$ and $[\phi]_B$. If we assume further that an ontological model of $\mathfrak{F}_{AB}$ must satisfy the conditions of the theorem, i.e. parameter independence, preserving ontological indistinctness with respect to unitaries, and reproducing the quantum predictions, then any ontological model of $\mathfrak{F}_A$ must be ψ-ontic.

Proof of Theorem 10.8. By choosing the global phases appropriately, we can find vector representatives, $|\psi\rangle_{AB}$ and $|\phi\rangle_{AB}$, of $[\psi]_{AB}$ and $[\phi]_{AB}$ such that $\langle\psi|\phi\rangle_{AB}$ is real and positive. Since $\mathrm{Tr}([\phi]_{AB}[\psi]_{AB}) \leq \frac{d-1}{d}$, we have $0 \leq \langle\phi|\psi\rangle_{AB} \leq \sqrt{\frac{d-1}{d}}$.

Let $\{|j\rangle\}_{j=0}^{d-1}$ be an orthonormal basis for $\mathbb{C}^d$ and consider the vectors

$$\left|\Phi^+\right\rangle_{AB} = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} |jj\rangle_{AB}\,, \tag{247}$$

and

$$|\eta\rangle_{AB} = \frac{1}{\sqrt{d-1}} \sum_{j=1}^{d-1} e^{i\varphi_j} |jj\rangle_{AB}\,, \tag{248}$$

By Lemma 9.4, the phases φ_j can be chosen in such a way that $\langle\eta|\Phi^+\rangle_{AB} = \langle\phi|\psi\rangle_{AB}$. Thus, there exists a unitary operator U such that $U|\psi\rangle_{AB} = |\Phi^+\rangle$ and $U|\phi\rangle_{AB} = |\eta\rangle_{AB}$. Given that Ξ_{AB} preserves ontological indistinctness with respect to unitaries, if $[\Phi^+]_{AB}$ and $[\eta]_{AB}$ are ontologically distinct in Ξ_{AB} then $[\psi]_{AB}$ and $[\phi]_{AB}$ must also be ontologically distinct.

Let $\mu \in \Delta_{AB}([\Phi^+]_{AB})$ and $\nu \in \Delta_{AB}([\eta]_{AB})$. Consider the measurement $M_A = \{[j]_A\}_{j=0}^{d-1}$. By Theorem 10.7, the response function $\xi_0^{M_A}$ of the marginal model on A must be equal to $1/d$ on a set Ω that is measure one according to μ. However, in order to reproduce the quantum predictions, we must have

$$\int_{\Lambda_{AB}} \xi_0^{M_A}(\lambda) d\nu(\lambda) = 0, \tag{249}$$

because $|\eta\rangle_{AB}$ is orthogonal to $|0\rangle_A$. Thus, $\xi_0^{M_A}$ can only be nonzero on a set that is of measure zero according to ν. Hence, $\mu(\Omega) = 1$ and $\nu(\Omega) = 0$, so $[\Phi^+]_{AB}$ and $[\eta]_{AB}$ are ontologically distinct. □

Part III
Beyond the ψ-ontic/epistemic distinction

Part II presented three theorems that aimed to rule out ψ-epistemic ontological models. In each case, auxiliary assumptions were required to prove the theorem: the preparation independence postulate for the PBR Theorem, ontic indifference for Hardy's Theorem, and parameter independence for the Colbeck-Renner Theorem. Each of these auxiliary assumptions are question begging. We also saw, in §7.5, that ψ-epistemic theories are possible if we do not make any auxiliary assumptions.

However, as noted in §4.4, the definition of a ψ-epistemic model is highly permissive in that it only requires probability measures corresponding to two states to have overlap in an ontological model. Most pairs of nonorthogonal states may still be ontologically distinct and furthermore the overlaps that do exist may be arbitrarily small. Clearly, more than this is needed if the ψ-epistemic explanations of quantum phenomena like the indistinguishability of nonorthogonal states are to be viable. Therefore, one might hope that viable ψ-epistemic interpretations might be ruled out without making auxiliary assumptions by imposing stronger requirements on the overlaps in ontological models. This part discusses four proposals for such requirements.

§11 discusses the requirement that all pairs of nonorthogonal states should correspond to measures that have nonzero overlap. Aaronson et. al. [53] have shown that such models exist in all dimensions, but they have also shown that imposing additional assumptions can rule them out. One of these is that the model should satisfy a fairly reasonable symmetry requirement, but the other severely restricts the class of ontic state spaces. §12 discusses the requirement that ontological models should be continuous, in the sense that quantum states with large inner product should correspond to measures with large overlap. Two different notions of continuity have been discussed in the literature. Patra, Pironio and Massar have proven a theorem ruling out models that satisfy a very strong notion of continuity [118]. However, their notion of continuity is so strong that it rules out many quite reasonable ψ-epistemic models. This is discussed§12.1. §12.2 discusses a version of Lipschitz continuity, which says that the ratio of overlaps in an ontological model to the quantum probabilities should be bounded. This can be used as the basis of measuring the degree to which a model is ψ-epistemic by comparing the overlaps of measures in the ontological model with the inner products of the corresponding quantum states. The idea here is that if the former is small for a pair of states with large inner product then the ψ-epistemic explanation of the indistinguishability of quantum states is not really viable. Several results bounding these overlaps have recently been obtained [54, 56, 104]. Finally, §13 discusses the concept of a *sometimes* ψ-ontic model and its converse, a *never* ψ-ontic model. In a *sometimes* ψ-ontic model, for each pure state there is a region of the ontic state space to which its probability measures assign nonzero measure, but this region is assigned zero measure by the probability measures corresponding to all other pure states. Thus, if the system happens to occupy an ontic state in one of these regions then the quantum state that was prepared can be deduced uniquely. The system need not always occupy such an ontic state though; hence the terminology "sometimes" ψ-ontic. In the context of reproducing the predictions of all projective measurements for all pure states on $\mathbb{C}^d$, all ψ-epistemic models that have been proposed to date are sometimes ψ-ontic for $d \geq 3$, but the question of whether this must necessarily be the case is open.

Even if you think the additional assumptions of the ψ-ontology theorems discussed in Part II are reasonable, there are still important reasons to look for stronger notions of ψ-epistemicity that might be ruled out without them. In §5, we discussed the implications of assuming that the quantum state is ψ-ontic, ranging from excess baggage to preparation contextuality and nonlocality. However, if your claim of ψ-ontology is based on the existing theorems then the implications from ψ-ontology to these other features of ontological models inherit the auxiliary assumptions made in the existing ψ-ontology theorems. Since we can already prove these things without such additional assumptions, e.g. the standard proof of Bell's Theorem does not require assuming the preparation independence postulate or making assumptions about how dynamics is represented, the claim that ψ-ontology theorems are the strongest known no-go theorems for ontological models is somewhat weakened. From this point of view, proving that models must be sometimes ψ-ontic would be particularly interesting, since all of the implications discussed in §5 would follow from it. This is also discussed in §13.

The discussion in this part is less detailed than in Parts I and II. This is because going beyond the ψ-ontic/ψ-epistemic distinction is a relatively new development and many of the known results are only provisional. I expect that some of the material in this section will be obsolete by the time this review is published, so it is more important to outline the main ideas than to give detailed proofs.

11 Pairwise ψ-epistemic models

One of the deficiencies in the definition of a ψ-epistemic model is that it only requires that a single pair of pure states is ontologically indistinct. Clearly, in order to explain the indistinguishability of every pair of nonorthogonal states in terms of the overlap of the corresponding probability measures, all such pairs should to be ontologically indistinct. This is what is meant by a pairwise ψ-epistemic model.

Definition 11.1. Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment and let $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ be an ontological model of it. Ξ is *pairwise ψ-epistemic* if all pairs of nonorthogonal pure states in $\mathcal{P}$ are ontologically indistinct in Ξ.

Aaronson, Bouland, Chua and Lowther (ABCL) have shown that pairwise ψ-epistemic models exist for the fragment consisting of the set of all pure states and all measurements in orthonormal bases in any finite dimensional Hilbert space [53]. They also prove a theorem ruling out such models by imposing some additional assumptions. One of their assumptions is a quite reasonable symmetry requirement, but the other is that the ontic state space describing a system with Hilbert space $\mathbb{C}^d$ is either the projective Hilbert space, i.e. the set of pure states on $\mathbb{C}^d$, or the set of unitary operators on $\mathbb{C}^d$. This is obviously a very restricted setting, so their theorem should be regarded as a provisional step towards ruling out pairwise ψ-epistemic models by symmetry requirements. Here, I give a rough outline of their construction of a pairwise ψ-epistemic theory in finite dimensional Hilbert spaces and give a brief account of the symmetry assumption behind their theorem.

§7.5 described ABCL's construction of an ontological model wherein any fixed pair of nonorthogonal pure states could be made ontologically indistinct. The key to constructing a pairwise ψ-epistemic model is to find a way of mixing ontological models together such that any states that are ontologically indistinct in the original models remain so in the mixture. Then, if one mixes sufficiently many ABCL models the result will be pairwise ψ-epistemic. A mixture of two ontological models is defined as follows.

Definition 11.2. Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment and let $\Xi^1 = (\Lambda^1, \Sigma^1, \Delta^1, \xi^1)$ and $\Xi^2 = (\Lambda^2, \Sigma^2, \Delta^2, \xi^2)$ be ontological models of it. The model $\Xi^3 = (\Lambda^3, \Sigma^3, \Delta^3, \xi^3) = p\Xi^1 + (1-p)\Xi^2$, where $0 \leq p \leq 1$, is defined as follows.

- $\Lambda^3 = \Lambda^1 \oplus \Lambda^2$ and Σ^3 consists of all sets of the form $\Omega^1 \oplus \Omega^2$ for $\Omega^1 \in \Sigma^1$, $\Omega^2 \in \Sigma^2$.
- For all $\rho \in \mathcal{P}$, Δ^3_ρ consists of all measures of the form $p\mu^1 + (1-p)\mu^2$ for $\mu^1 \in \Delta^1_\rho$ and $\mu^2 \in \Delta^2_\rho$.
- For all $M \in \mathcal{M}$, $E \in M$, $(\xi^3)^M_E = (\xi^1)^M_E + (\xi^2)^M_E$.

A few words of clarification are in order here. First of all, the direct sum of two spaces is essentially just the two spaces placed side-by-side. More formally, $\Lambda_1 \oplus \Lambda_2$ is the union of the set of objects of the form $(1, \lambda_1)$ with the set of objects of the form $(2, \lambda_2)$. This is different from the union of Λ_1 and Λ_2, since Λ_1 and Λ_2 might contain some common elements. The addition of the integer label is to ensure that we have two copies of such common elements, with the label indicating whether it is the copy that belongs to Λ_1 or to Λ_2. However, to simplify notation, I generally omit the integer label. When μ^1 is a probability measure on (Λ^1, Σ^1), I also denote by μ^1 the probability measure on (Λ^3, Σ^3) that satisfies $\mu^1(\Lambda^2) = 0$ and agrees on Σ^1. Similarly, when $(\xi^1)^M_E$ is a response function of Λ^1 the same notation denotes the function on Λ^3 that agrees on Λ^1 and is zero on Λ^2. Similar remarks apply under the exchange of 1 and 2. It is then straightforward to check that if Ξ^1 and Ξ^2 reproduce the quantum predictions then so does Ξ^3.

Theorem 11.3. *Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment and and let $\Xi^1 = (\Lambda^1, \Sigma^1, \Delta^1, \xi^1)$ and $\Xi^2 = (\Lambda^2, \Sigma^2, \Delta^2, \xi^2)$ be ontological models of it. Let $\Xi^3 = p\Xi^1 + (1-p)\Xi^2$ for some $0 < p < 1$. If $[\psi]$ and $[\phi]$ are ontologically indistinct in either Ξ^1 or Ξ^2 then they are ontologically in distinct in Ξ^3.*

Proof. Without loss of generality, assume that $[\psi]$ and $[\phi]$ are ontologically indistinct in Ξ^1, since the proof for Ξ^2 follows the same logic. Then, there exist $\mu^1 \in \Delta^1_\psi$ and $\nu^1 \in \Delta^1_\phi$ such that $c = D(\mu^1, \nu^1) < 1$, where D is the variational distance. Let $\mu^2 \in \Delta^2_\psi$ and $\nu^2 \in \Delta^2_\phi$. Then, $\mu^3 = p\mu^1 + (1-p)\mu^2 \in \Delta^3_\psi$ and $\nu^3 = p\nu^1 + (1-p)\nu^2 \in \Delta^3_\phi$. Let $\Omega \in \Sigma^3$ and define $\Omega^1 = \Omega \cap \Lambda^1$ and $\Omega^2 = \Omega \cap \Lambda^2$. Then, for any measure μ on (Λ^3, Σ^3), $\mu(\Omega) = \mu(\Omega^1) + \mu(\Omega^2)$, but we also have $\mu^1(\Omega^2) = \nu^1(\Omega^2) = \mu^2(\Omega^1) = \nu^2(\Omega^1) = 0$. Hence,

$$\begin{aligned} \left|\mu^3(\Omega) - \nu^3(\Omega)\right| &= \left|p\mu^1(\Omega^1) + p\nu^1(\Omega^1) + (1-p)\mu^2(\Omega^2) + (1-p)\nu^2(\Omega^2)\right| && (250)\\ &\leq p\left|\mu^1(\Omega^1) - \nu^1(\Omega^1)\right| + (1-p)\left|\mu^2(\Omega^2) - \nu^2(\Omega^2)\right| && (251)\\ &\leq pc + (1-p) && (252)\\ &= 1 - (1-c)p, && (253) \end{aligned}$$

and this is < 1 because $0 \leq c < 1$ and $p > 0$. Since this holds for any $\Omega \in \Sigma^3$ it also holds for the supremum. Hence, $D(\mu^3, \nu^3) < 1$ and so $[\psi]$ and $[\phi]$ are ontologically indistinct in Ξ^3. □

The rough idea of the ABCL construction is now quite simple to describe. Let $\mathfrak{F} = \langle \mathbb{C}^d, \mathcal{P}, \mathcal{M} \rangle$ be the PM fragment where $\mathcal{P}$ consists of all pure states on $\mathbb{C}^d$ and $\mathcal{M}$ consists of all measurements in orthonormal bases. In §7.5, we showed that, for any pair of nonorthogonal states $[\psi], [\phi] \in \mathcal{P}$, there exists an ontological model $\Xi_{\psi,\phi}$ in which $[\psi]$ and $[\phi]$ are ontologically indistinct. Therefore, by Theorem 11.3, the model $p\Xi_{\psi,\phi} + (1-p)\Xi_{\psi',\phi'}$ has both $([\psi], [\phi])$ and $([\psi'], [\phi'])$ as ontologically indistinct pairs. This construction can then be iterated in order to obtain a model in which any finite set of pairs of nonorthogonal pure states are ontologically distinct.

The remaining difficulty is to extend this construction to an infinite number of pairs, i.e. all pairs of pure states on $\mathbb{C}^d$. The details of this are quite technical but straightforward once one understands the basic idea. The interested reader should consult [53].

ABCL go on to prove a theorem ruling out pairwise ψ-epistemic models under an additional symmetry requirement. They also impose that the ontic state space for a quantum system with Hilbert space $\mathbb{C}^d$ is either the projective Hilbert space of $\mathbb{C}^d$ or the set of unitary operators on $\mathbb{C}^d$. In the more general context of arbitrary models, their symmetry assumption runs as follows.

Definition 11.4. Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M}, \mathcal{T} \rangle$ be a PMT fragment and let $\Xi = (\Lambda, \Sigma, \Delta, \xi, \Gamma)$ be an ontological model of it. Ξ satisfies *ABCL symmetry* if, for all pairs $[\psi] \in \mathcal{P}$ and $U \in \mathcal{T}$ such that $U[\psi]U^\dagger = [\psi]$, there exists a $\mu \in \Delta_\psi$ and a $\gamma \in \Gamma_U$ such that

$$\mu(\Omega) = \int_\Lambda \gamma_\lambda(\Omega) d\mu(\lambda). \tag{254}$$

This says that, if a unitary leaves a pure state invariant, then there should be a probability measure representing that state and a Markov kernel representing the unitary such that the measure is invariant

under the stochastic map representing the unitary. ABCL then show that, for the fragment consisting of all pure states, all measurements in orthonormal bases, and all unitaries on $\mathbb{C}^d$, if the ontic state space is either the projective Hilbert space of $\mathbb{C}^d$ or the set of unitary operators on $\mathbb{C}^d$ and the dynamics is represented by the usual action of the unitary group on these spaces, then no ABCL symmetric and pairwise ψ-epistemic model exists.

Imposing additional symmetry requirements of the type suggested by ABCL is a promising direction for ψ-ontology theorems, but their setup is so restricted that one cannot really draw any firm conclusions about ontic models in general from it. The proof itself is quite technical and not obviously generalizable to other ontic state spaces, so the details are omitted here.

12 Continuity

The idea of continuity is that if two pure quantum states are close to one another, in the sense of having large inner product, then there should be probability measures representing them that have large overlap. Imposing continuity is not a bad idea, since presumably models in which a small change in experimental conditions leads to a large change at the ontological level would be quite contrived. The most basic notion of continuity runs as follows.

Definition 12.1. Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment where $\mathcal{P}$ consists of pure states and let $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ be an ontological model of it. Ξ is *continuous* if, for all $\epsilon > 0$, there exists a $\delta > 0$ such that if $\mathrm{Tr}\left([\phi]\,[\psi]\right) > 1 - \delta$ for some pure states $[\psi], [\phi] \in \mathcal{P}$, then there exists $\mu \in \Delta_\psi$ and $\nu \in \Delta_\phi$ such that $L(\mu, \nu) > 1 - \epsilon$.

This notion of continuity is quite permissive because it imposes no constraints on how δ should be related to ϵ. It is easy to see that any pairwise ψ-epistemic model is continuous in this sense, since all it requires is that the overlap of measures corresponding to nonorthogonal states should be bounded away from zero. Since ABCL have shown that pairwise ψ-epistemic models exist, this notion of continuity cannot be used to derive a no-go theorem.

Therefore, it is interesting to look at stronger notions of continuity that might be reasonable for a ψ-epistemic model. In the remainder of this section, two such notions are discussed. §12.1 discusses a notion of continuity due to Patra, Pironio and Massar [118] and discusses a theorem that they proved to rule such models out. I argue that their definition of continuity is unreasonable for a ψ-epistemic model. §12.2 discusses a different notion based on Lipschitz continuity, which has been used in a number of recent works to derive measures of the extent to which a model is ψ-epistemic. The best known bound shows that the ratio of the overlap to the inner product must decrease exponentially in Hilbert space dimension for some families of states.

12.1 The Patra-Pironio-Massar Theorem

PPM's continuity assumption runs as follows.

Definition 12.2. Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment, where $\mathcal{P}$ consists of pure states, let $\delta > 0$ and, for $[\psi] \in \mathcal{P}$, let $B^\delta_\psi = \left\{ [\phi] \in \mathcal{P} \middle| \sqrt{\mathrm{Tr}\left([\phi]\,[\psi]\right)} \geq 1 - \delta \right\}$ be the ball of radius δ centered at $[\psi]$. An ontological model $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ of $\mathfrak{F}$ is *PPM-δ-continuous* if, for every pure state $[\psi] \in \mathcal{P}$ and every set $\{[\phi_j]\} \subseteq \mathcal{P}$ where $[\phi_j] \in B^\delta_\psi$, there exist probability measures $\mu_j \in \Delta_{\phi_j}$ such that $L(\{\mu_j\}) > 0$.

For the case of a finite ontic state space, the PPM-δ-continuity assumption says that, for every set of quantum states that are sufficiently close some common state $[\psi]$, there should be at least one ontic state to which they all assign nonzero probability. This is stronger than the basic notion of continuity given in Definition 12.1, which only constrains pairwise overlaps.

Based on their assumption, PPM prove the following theorem.

Theorem 12.3. *Let $\mathfrak{F} = \langle \mathbb{C}^d, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment with $\{[k]\}_{k=0}^{d-1} \in \mathcal{M}$ for some orthonormal basis $\{|k\rangle\}_{k=0}^{d-1}$. If $[\psi] \in \mathcal{P}$ and $\{[\phi_j]\}_{j=0}^{d-1} \subseteq \mathcal{P}$, where*

$$|\psi\rangle = \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} |k\rangle \tag{255}$$

$$|\phi_j\rangle = \frac{1}{\sqrt{d-1}} \sum_{k \neq j} |k\rangle \,, \tag{256}$$

then there is no PPM-δ-continuous ontological model of $\mathfrak{F}$ that reproduces the quantum preclusions for $\delta \geq 1 - \sqrt{(d-1)/d}$.

For an infinite dimensional Hilbert space, this implies that there is no PPM-δ-continuous model for any $\delta > 0$. As with the Hardy and Colbeck-Renner Theorems, the constraint on δ can be removed by assuming that adding ancillas preserves ontological indistinctness.

Proof of Theorem 12.3. Let $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ be an ontological model of $\mathfrak{F}$ and let $\mu_j \in \Delta_{\phi_j}$. The states $[\phi_j]$ all lie within B_ψ^δ for $\delta = 1 - \sqrt{(d-1)/d}$. However, the states $[\phi_j]$ are antidistinguished by the measurement $\{[k]\}_{k=0}^{d-1}$ and hence, by Theorem 6.5, $L(\{\mu_j\}) = 0$. □

This proof is remarkably simple, but it does not mean much as PPM-δ-continuity is unreasonably strong. In order to see the problem, it is helpful to consider a simple example. Therefore, consider how the fragment $\mathfrak{F} = \langle \mathbb{C}^d, \mathcal{P}, \mathcal{M} \rangle$ might be modelled, where $\mathcal{P}$ just contains the states $[\psi]$, $[\phi_j]$ and $\mathcal{M}$ only contains the measurement $M = \{[k]\}_{k=1}^{d}$. A natural way of doing this is to use the ontic state space $\Lambda = \{0, 1, \ldots, d-1\}$, model $[\psi]$ by the uniform distribution, and model $[\phi_j]$ by the distribution that is zero on j and uniformly distributed over the rest of the ontic state space. Then, it is straightforward to see that setting the response functions $\xi_k^M(j) = \delta_{j,k}$ yields a model that reproduces the quantum predictions. Furthermore, this model is pairwise ψ-epistemic, and hence satisfies the more basic notion of continuity, because $D(\mu_j, \mu_k) = 1/(d-1)$ for $j \neq k$.

From this, it is easy to see that the problem with PPM-δ-continuity comes from ontological states that are assigned a small weight in the distribution. If a distribution corresponding to some quantum state assigns a small weight to all of the ontic states in its support then one way of making a small change to that distribution is to set the weight assigned to one of the ontic states to zero and redistribute it over the other ontic states. If those distributions all happen to represent quantum states that are close to the original one, and there is no reason why they should not, then these quantum states will have no ontic state that is common to all of them, and hence PPM-δ-continuity will be violated.

In fact, the same criticism applies to classical models. Suppose we have a system consisting of a ball that can be in one of d boxes, labelled $0, 1, \ldots, d-1$. The system is prepared in one of the following ways. Either a box is chosen uniformly at random and the ball is placed in that box, or first an integer j from 0 to $d-1$ is specified and box j is removed, then the ball is placed uniformly at random in one of the remaining boxes, before finally returning box j to its place. Suppose you are given a table which specifies the probabilities of finding the ball in each of the boxes for each of these $d+1$ preparation procedures. You are, however, not told anything about where these probabilities come from, i.e. you are not told about the ball and boxes but rather just given a table that tells you that if preparation j is made then the probability of getting outcome k is 0 if $k = j$ and $1/(d-1)$ otherwise. Your task is to try and come up with a model of what is going on in reality in order to generate these statistics; a task perfectly analogous to the project of constructing an ontological model for some fragment of quantum theory. If you imposed upon yourself the analog of PPM-δ-continuity, that sets of preparations having close operational predictions must have an ontic state in common, you would never be able to come up with the ball and box model that is actually generating the probabilities and, in fact, no model satisfying this condition exists. That, in a nutshell, is what is wrong with PPM-δ-continuity. A condition that rules out a classical probabilistic description of a classical probabilistic model is clearly way to strong to prove anything meaningful about quantum theory.

12.2 Lipschitz continuity

In order to for the ψ-epistemic explanation of the indistinguishability of quantum states to be viable, quantum states that have large inner product should correspond to probability measures with large overlap, but quantum states that are almost orthogonal need have very little overlap at the ontic level. In between these extremes, an amount of overlap that scales with the inner product in some way is needed. This motivates the definition of Lipschitz continuity.

Definition 12.4. Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment where $\mathcal{P}$ consists of pure states and where, for every $[\psi] \in \mathcal{P}$, there exists an $M \in \mathcal{M}$ such that $[\psi] \in M$. Let $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ be an ontological model of $\mathfrak{F}$. Ξ is Lipschitz continuous if there exists a $K > 0$ such that

$$L(\mu, \nu) \geq K \mathrm{Tr}\left([\phi]\,[\psi]\right), \tag{257}$$

for some $\mu \in \Delta_\psi$, $\nu \in \Delta_\phi$.

Bounds on K have been obtained by Maroney [56] and later by Barrett et. al. [54]. The measures of overlap used in these works differ, but this only affects the bounds by a constant factor. The best published bound on K shows that there exist fragments with Hilbert space $\mathbb{C}^d$ such that $K < 4/(d-1)$ for any ontological model that reproduces the quantum preclusions and the same construction shows that $K < 2/d$ in prime power dimensions [54]. In forthcoming work, I show that there exists a positive constant c such that $K \leq de^{-cd}$ [104], so the classical overlaps must be exponentially smaller than the quantum probabilities.

These constructions typically work by showing that, in any ontological model reproducing the quantum predictions, there exist families of states for which the ratio $L(\mu, \nu)/\mathrm{Tr}\left([\phi]\,[\psi]\right)$ is bounded. This still applies even without assuming Lipschitz continuity, i.e. there exist pairs of states in $\mathbb{C}^d$ such that the ratio of overlap to inner product is exponentially small. Lipschitz continuity is only required to infer that this must be the case for all pairs of states. Unfortunately, the pairs of states used in the known constructions have $\mathrm{Tr}\left([\phi]\,[\psi]\right) \to 0$ for $d \to \infty$, so all we can really say without Lipschitz continuity is that almost orthogonal states have overlap that is exponentially smaller than their inner product.

What is really needed is a bound on the overlap for families of quantum states in increasing dimension where the inner product remains fixed as the dimension increases. If the overlap still tends to zero in large dimension then this would be very strong evidence that the ψ-epistemic explanation of indistinguishability is not viable in any ontological model.

13 Never ψ-ontic models

To conclude, I would like to discuss one further strengthening of the notion of a ψ-epistemic model, about which very little is currently known. This is the notion of a never ψ-ontic model.

Definition 13.1. Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment and let $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ be an ontological model of it. Ξ is *sometimes ψ-ontic* if, for all pure states $[\psi] \in \mathcal{P}$, there exists a $\Omega \in \Sigma$ such that $\mu(\Omega) > 0$ for some $\mu \in \Delta_\psi$, but for all other pure states $[\phi] \in \mathcal{P}$, $[\phi] \neq [\psi]$, every $\nu \in \Delta_\phi$ has $\nu(\Omega) = 0$. Otherwise the model is *never ψ-ontic*.

Roughly speaking, in a sometimes ψ-ontic model, each pure state has a special region of the ontic state space all to itself. If you know that the ontic state occupies this region then you can identify the quantum state that was prepared with probability one. However, the ontic state need not always occupy such a region; hence the terminology "sometimes" ψ-ontic. In contrast, in a never ψ-ontic model, every region of the ontic state space is shared nontrivially by more than one quantum state. Considerations about the degree of overlap of pairs of states do not really bear on the question of whether a never ψ-ontic model is possible, since arbitrary overlaps may occur outside the special regions in a sometimes ψ-ontic model.

The reason why the notion of a sometimes ψ-ontic model is interesting is that all of the implications of ψ-ontology discussed in §5 can be derived from sometimes ψ-ontology instead. Whilst the fact that a maximally

ψ-epistemic model is impossible is enough to derive preparation contextuality and Bell's Theorem, it does not imply excess baggage. However, excess baggage follows directly from sometimes ψ-ontology, since there must be at least as many ontic states as there are pure quantum states if each quantum state is to be assigned its own region of ontic state space. Further, a sometimes ψ-ontic model cannot be maximally ψ-epistemic, so we obtain all the implications of that as well.

Theorem 13.2. *Let $\mathfrak{F} = \langle \mathcal{H}, \mathcal{P}, \mathcal{M} \rangle$ be a PM fragment and let $\Xi = (\Lambda, \Sigma, \Delta, \xi)$ be an ontological model of it. If Ξ is sometimes ψ-ontic then it is not maximally ψ-epistemic.*

Proof. Assume that Ξ is maximally ψ-epistemic. Then, for all $[\psi], [\phi] \in \mathcal{P}$, all $\mu \in \Delta_\psi$, all $\nu \in \Delta_\phi$, and all $\Omega \in \Sigma$ such that $\nu(\Omega) = 1$,

$$\int_\Omega \xi_E^M(\lambda) d\mu(\lambda) = \int_\Lambda \xi_E^M(\lambda) d\mu(\lambda), \tag{258}$$

for all $M \in \mathcal{M}$, $E \in M$.

Let Ω' be the region uniquely assigned to $[\psi]$ so that $\nu(\Omega') = 0$ and $\mu(\Omega') > 0$. Then Ω can be chosen such that $\Omega \cap \Omega' = \emptyset$. Summing Eq. (258) over E then gives $\mu(\Omega) = \mu(\Lambda) = 1$. However, $\mu(\Omega)$ cannot be equal to one because $\mu(\Omega') > 0$ and Ω and Ω' are disjoint, so this would make $\mu(\Lambda) > 1$. Hence, there is a contradiction, so Ξ cannot be maximally ψ-epistemic. □

The Kochen-Specker model is never ψ-ontic, but all the existing ψ-epistemic models for Hilbert spaces of dimension ≥ 3 are sometimes ψ-ontic, so whether never ψ-ontic models exist in all dimensions is an open question. Since it is impossible to prove ψ-ontology without auxiliary assumptions, sometimes ψ-ontology is perhaps the strongest result about the reality of the quantum state that we could hope to prove without such assumptions. For this reason, I consider this to be the most important open question in this area.

14 Discussion and conclusions

14.1 Summary

This review article was divided into three parts. In the first part, I explained the distinction between ontic and epistemic interpretations of the quantum state, outlined the pre-existing arguments for epistemic and ontic interpretations, and explained how many existing no-go theorems would follow from proving that the quantum state is ontic. The aim of this part was to convince you that the ψ-ontic/epistemic distinction is interesting, that the question is unresolved by qualitative arguments, and that it is the sort of issue that should be addressed with the rigour that Bell brought to nonlocality.

The second part discussed the three existing ψ-ontology theorems: the PBR Theorem, Hardy's Theorem, and the Colbeck-Renner Theorem. Each of these results involve auxiliary assumptions, of varying degrees of reasonableness. Hardy's ontic indifference assumption is not really appropriate for a ψ-epistemic theory, but it does allow the flaw in the argument for the reality of the wavefunction from interference to be exposed more clearly. Colbeck and Renner's assumption of parameter independence is doubtful in light of Bell's Theorem. Whilst violations of Bell inequalities can arise from the failure of outcome independence alone, many viable realist interpretations of quantum theory, such as de Broglie-Bohm theory, violate parameter independence instead, and one really wants the scope of a no-go theorem to include as many viable realist interpretations as possible. In my view, the preparation independence postulate of the PBR Theorem is the best of the bunch. It is not completely unassailable, since weakening it slightly to allow for genuinely nonlocal degrees of freedom does potentially allow for a viable ψ-epistemic interpretation. Nevertheless, the PIP is satisfied in theories like de Broglie-Bohm, so, unlike the Colbeck-Renner Theorem, the PBR Theorem does explain why the wavefunction must be real in theories of that type.

Without auxiliary assumptions, ψ-epistemic models do exist, so the third part of this review discussed strengthenings of the notion of a ψ-epistemic model may be needed in order to make the ψ-epistemic explanations of phenomena like the indistinguishability of quantum states and the no cloning theorem go through. Whilst the case is not yet watertight, bounds on the ratio of the overlap of probability measures to the

quantum probabilities of the states they represent are beginning to make ψ-epistemic explanations look implausible within the ontological models framework.

14.2 Open questions

Part of the aim of this review was to provide the necessary background necessary for researchers who would like to begin working in this area. In this regard, several open questions were raised in the main text, and solving some of them would help to put the non viability of ψ-epistemic interpretations on a more secure footing. For easy reference, the following list collects all the open problems together in one place. The section numbers indicate where in the review you can find a more detailed description.

1. §5.3: Does the condition of maximal ψ-epistemicity have to fail for every pair of nonorthogonal pure states? If this is true then the ψ-epistemic explanation of the distinguishability of quantum states would not hold perfectly for any pair of pure states. It would also provide a proof that every mixed state must be preparation contextual.

2. §7.7.2: Can a deterministic ψ-ontic model always be converted into a ψ-epistemic one by finding regions of the ontic state space associated with different pure states that make the exact same predictions? One of Schlosshauer and Fine's criticisms of the PBR Theorem was based on the idea that this can always be done. It is not true for nondeterministic theories, but it may be true for deterministic theories.

3. §11: Can reasonable symmetry requirements be used to rule out pairwise ψ-epistemic models without restrictive assumptions on the nature of the ontic state space? The theorem of Aaronson et. al. only applies if the ontic state space is projective Hilbert space or the set of unitary operators, which is a very restrictive class of models.

4. §12.2: Does the overlap of probability measures representing nonorthogonal states have to tend to zero in increasing Hilbert space dimension for pairs of states with a fixed inner product? The existing results only imply this for families of states that become orthogonal in the infinite dimensional limit, so this only implies that the ψ-epistemic explanation of indistinguishability is implausible for states that are almost distinguishable.

5. §13: Does a never ψ-ontic model exist for the fragment consisting of all pure states and projective measurements in $\mathcal{C}^d$ for $d \geq 3$? In my view, this is the most important open question, as proving that all models must be sometimes ψ-ontic has the same implications as ψ-ontology.

6. Do ψ-ontology results have applications in quantum information theory, beyond those of the implications discussed in §5? The ontological models framework is already regarded as implausible by some physicists, but it is worth investigating anyway because it provides a way of simulating quantum experiments using classical resources. Results derived in this framework often go on to have applications in quantum information. In this regard, Montina proved an upper bound on the classical communication complexity of simulating the identity channel for a qubit using the Kochen-Specker model, which is maximally ψ-epistemic [130]. It is possible that the nonexistence of such models in higher dimensions could be used to prove lower bounds on the communitcation complexity of the identity channel in higher dimensions.

14.3 Experiments

Probably the most important issue not discussed so far in this review is the question of whether the reality of the quantum state can be established experimentally. When it comes to experiments in the foundations of quantum theory, tests of Bell's Theorem are somewhat of a gold standard. Violation of a Bell inequality rules out a wide class of local realistic theories independently of the details of quantum theory. Consequently, the experimenter does not have to assume much about how their experiment is represented within quantum theory. From their observed statistics they simply get a yes/no answer as to whether local hidden variable

theories are viable, modulo the known loopholes. In modern parlance, tests of Bell's Theorem are *device independent* [96].

In contrast, a test of the reality of the quantum state would not be device independent simply because the "quantum state" is the thing we are testing the reality of, and that is a theory dependent notion. Consequently, one has to assume that our quantum theoretical description of the way that our preparation devices work is more or less accurate, in the sense that they are approximately preparing the quantum states the theory says they are, in order to test the existing ψ-ontology results. Therefore, it is desirable to have a more theory independent notion of whether a given set of observed statistics imply that the "probabilistic state", i.e. some theory-independent generalization of the quantum state, must be real. It is not obvious whether this can be done, but if it can then experimental tests of ψ-ontology results would become much more interesting.

Of course, one can still perform non device independent experimental tests. This amounts to trying to prepare the states, perform the transformations, and make the measurements involved in a ψ-ontology result and checking that the quantum predictions are approximately upheld. Due to experimental error, the agreement will never be exact, but one can bound the overlap between probability measures representing quantum states instead of showing that it must be exactly zero. For the special case of the PBR Theorem given in Example 7.9 [131], this has been done using two ions in an ion trap. However, the experimental result only shows that the overlap in probability measures must be smaller than the quantum probability, and not that it must be close to zero. This is quite far from establishing the reality of the quantum state, since for that one would want to test many pairs of quantum states with a variety of different inner products, and the PBR measurement for states with large inner product requires an entangled measurement on a large number of quantum systems. This is not likely to be feasible until we have a general purpose quantum computer. Similarly, the original version of the Colbeck-Renner Theorem, i.e. the one that aimed to rule out ontological models with greater predictive power than quantum theory, has been tested using a quantum optical system [132] and there has also been an optical test of the Patra-Pironio-Massar Theorem on PPM-continuous models [58].

14.4 Future directions

Assuming that future results drive the final nails into ψ-epistemic explanations within the ontological models framework, the final question I want to address is where to go next. One option is to embrace the reality of the wavefunction by adopting one of the existing realist interpretations that fits into the ontological models framework, e.g. de Broglie-Bohm theory, spontaneous collapse theories or modal interpretations. Another is to adopt a neo-Copenhagen interpretation.

The first option is unappealing because adopting one of these realist interpretations opens up an explanatory gap. Namely, given that a ψ-epistemic interpretation would provides compelling explanations of a whole variety of quantum phenomena, it is puzzling that the quantum state should nevertheless be ontic. Of course, the existing no-go results, such as Bell's Theorem, also imply explanatory gaps, e.g. if a theory is explicitly nonlocal then why can this not be used to send a superluminal signal? For this reason, those inclined to the ψ-epistemic view are likely to have rejected these interpretations already on the basis of these other gaps.

Being neo-Copenhagen is always an option, but the merits of such an move depend on the degree to which one believes that realism is desirable. This is not the place to get into the debate between realism and antirealism and whether neo-Copenhagen views are compatible with some weakened notion of realism. Suffice to say that the viability of these interpretations turns on issues that are far deeper than the reality of the wavefunction. For my part, I think that if one denies the existence of an observer-independent reality then it becomes very difficult to maintain a clear notion of explanation at all. Closing explanatory gaps by denying the need for any explanation at all does not seem that appealing to me.

The only remaining option then is to adopt a realist interpretation that does not fit in to the ontological models framework. There are several possibilities, most of them highly speculative.

- Many-worlds: The many-worlds interpretation [6–8] is not contained within the ontological models

framework because the latter assumes that measurements yield a single unique outcome. Many-worlds is a commonplace retreat for realists who want to avoid introducing nonlocality in the face of Bell's Theorem. However, the conventional many-worlds interpretation is not a retreat for ψ-epistemicits because it is based on the idea that the quantum state is a literal description of reality. Nevertheless, since ψ-ontology Theorems do not apply to theories that involve many-worlds, it is possible that there is a viable ψ-epistemic interpretation that involves them. This is doubtless an unappealing option to both existing many-worlds advocates and ψ-epistemicists. For many-worlds advocates, the reason to take multiple worlds seriously is because they are encoded in the branching structure of the wavefunction. If you take away the reality of the wavefunction then you take away their reason for believing in them. Similarly, this is the only real reason for taking many worlds seriously, so someone who does not believe in the reality of the wavefunction is unlikely to have found the existence of many-worlds plausible in the first place. Nevertheless, it is logically possible that the universe could be described by some structure that does not imply a unique wavefunction, but does support the existence of many-worlds.

- Histories approaches: Histories interpretations, such as consistent histories [133] and Sorkin's co-event formalism [134, 135] are based on taking the spacetime picture provided by the path integral seriously. Consistent histories does not fit into the ontological models framework as it decrees that not all sets of histories can be assigned a probability and there may not be a unique description of the universe in terms of a single history. Griffiths argues that the consistent histories interpretation is appropriately ψ-epistemic [136]. In contrast, the co-event formalism is based on a modified logic and probability theory. I explained my doubts about realist interpretations of exotic probability theories in §7.7.4 and, in any case, Wallden [39] provides evidence that a result analogous to the PBR Theorem may hold in the coevent formalism.

- Retrocausality: In the ontological models framework, it is assumed that the probability measure representing a quantum state is independent of the choice of future measurement setting. If this were not the case then the ψ-epistemic interpretation of quantum phenomena could be maintained by having the measures corresponding to antidistinguishable states have no overlap when the antidistinguishing measurement is made, but nonzero overlap when other measurements are made. One way that dependence on the measurement setting may occur is if there is a direct causal influence, travelling backwards in time, from the measurement to the preparation. Several authors have argued that there are independent reasons for adopting a retrocausal approach to quantum theory [137–139], not least because it might allow an appropriately local resolution to the dilemma posed by Bell's Theorem. The transactional interpretation [140, 141] is explicitly retrocausal and the two-state vector formalism [142] can be read in a retrocausal way. However, both of these theories posit an ontic wavefunction. If we are to maintain ψ-epistemic explanations then we instead need to look for retrocausal ontological models that posit a deeper reality underlying quantum theory that does not include the quantum state.

- Relationalism: The ontological models framework assumes that systems have their own intrinsic properties, encompassed by the ontological state λ. Relationalism posits that systems do not have intrinsic properties, but only properties relative to other systems. The usual analogy is with the concept of position. One cannot talk about the position of a particle without setting down some coordinate system, and this implicitly means that we are measuring position with respect to some other physical system that provides a reference frame. There is an obvious commonality with Everett's relative state approach [6], except that we want a theory of this type that is ψ-epistemic. Rovelli's relational quantum mechanics [22] is a theory of this type in which the wavefunction is supposed to be epistemic, but he defines relational properties in terms of the global wavefunction and it is not clear how they supposed to be determined if the wavefunction is not real.

In conclusion, I think we should try to find a way of understanding quantum theory that closes as many of the explanatory gaps opened by no-go theorems as possible. This is because an interpretation that merely accommodates the known facts about quantum theory, rather than explaining them, is unlikely to yield principles that can reliably guide us towards future physical theories. Since ψ-ontology implies many of

the existing no-go theorems, the gap opened by ψ-ontology results should be taken at least as seriously as the others. This means that we should investigate the speculative roads less travelled described above, in addition to others that we have not thought of yet. The chances than any one of them will bear fruit may be slim, but the rewards if they do will be worth the effort.

Acknowledgements

I would like to thank Jonathan Barrett, Sarah Croke, Joseph Emerson, Lucien Hardy, Teiko Heinosaari, Ravi Kunjwal, Klaas Landsman, Peter Lewis, Owen Maroney, Matt Pusey, Terry Rudolph, Max Schlosshauer and Rob Spekkens for useful discussions. This research was supported in part by the FQXi Large Grant "Time and the Structure of Quantum Theory" and by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Research and Innovation.

A The ontic/epistemic distinction and objective chance

In this appendix, the question of whether a distinction equivalent to the Bayesian distinction between ontic and epistemic states can be made in theories of objective chance. As in the frequency theory, we are interested in whether objective chances can be viewed as intrinsic properties of individual systems, or whether they must be defined with reference to other systems, e.g. an ensemble of similar systems or the conditions prevailing in the environment of the system. An intrinsic property of the system is something like the charge of an electron or the hidden variable state of a quantum system, if the latter are presumed to exist. In a realist approach to physics, these things correspond to ontological features of the system and they cannot be changed just by viewing the environment of the system differently. Changing them requires an intervention in the system itself.

Deciding whether or not objective chances are intrinsic properties is difficult because there is no universally agreed upon theory of objective chance. Fortunately, the question only depends on a couple of broad features of the theory.

Firstly, some authors posit that objective chance is compatible with determinism so that, for example, the probabilities involved in classical statistical mechanics can be viewed as objective chances. Others think that objective chances only make sense if there is a genuine stochasticity in nature, with quantum theory providing the prime example of a theory that involves such genuine chance. It should be clear that objective chance cannot be intrinsic in any theory that is compatible with determinism. This is because, in a deterministic theory, the intrinsic properties of an isolated system determine its future uniquely, so they could only ever give rise to objective chances of 0 or 1. Therefore, to make such a theory work, one has to refer either to an ensemble of systems, as in the frequency theory, or to the conditions surrounding the system. As an example of the latter, one can imagine a specification of the way in which a coin should be tossed such as "a strong flip between thumb and forefinger" that is specific enough to license the assignment of a fixed objective chance, but vague enough that it does not determine the outcome of the toss uniquely. Of course, it is questionable whether such a notion makes sense, but the point is that, if one does take this view, then probabilities are not intrinsic properties of systems. Either they refer to ensembles or they refer to the conditions of interaction between a system and its environment, and specifying these in different levels of detail would lead to different probability assignments, just as in the frequency case.

However, many philosophers of objective chance, including Popper [143] and Lewis [144], take the view that objective chances only make sense in a genuinely stochastic universe. In this case, one can require that the objective chances of an experimental outcome are only the same if the intrinsic properties of the system before the experiment are identical in all relevant details. The prime example of this would be to say that the objective chances of obtaining a given outcome in a quantum measurement on two different systems are the same iff the systems are described by the same quantum state $|\psi\rangle$ prior to the experiment. This presents more of a problem for the distinction we are trying to make, since ostensibly the state $|\psi\rangle$ only refers to the

system itself.

However, it should be noted that the quantum state example rests on questionable assumptions about the interpretation of quantum theory. It assumes that $|\psi\rangle$ can be regarded as an intrinsic property of the system and, further, that $|\psi\rangle$ is a complete description of the system. In an operationalist approach, the assertion that $|\psi\rangle$ is an intrinsic property of a system is denied. Instead, it is a description of those facts about the device that prepared the system that are relevant for predicting future measurement outcomes. In other words, it is a condensed description of a set of knob settings, meter readings, etc. that refer to a piece of experimental apparatus external to the system. These are not intrinsic properties of the system, so the relevant distinction is maintained.

Of course, both Popper and Lewis intend a more realist interpretation of quantum theory. However, even if quantum states are intrinsic properties of quantum systems, they need not be complete. For example, in de Broglie-Bohm theory one also needs to specify the positions of particles. The theory is deterministic when both the quantum state and particle positions are specified, so we are back in the position of having to define objective chances in a deterministic theory, in which case they are not intrinsic. In de Broglie-Bohm theory, the statement that a system is described by a state $|\psi\rangle$ really means that it is part of an ensemble of systems described by the quantum state $|\psi\rangle \otimes |\psi\rangle \otimes \ldots$ and in which the particle positions are distributed according to $|\psi(x)|^2$. Since there is freedom to look at subensembles where the positions are not distributed according to $|\psi(x)|^2$, and these would allow prediction with greater accuracy than quantum theory, systems described by the same quantum state do not have the same intrinsic properties in all relevant detail.

Nevertheless, although quantum theory is a prime motivation for objective chances, theories of objective chance are usually independent of the details of physics. Thus, we can ignore the quantum motivation and just look at the theories of chance actually proposed to determine whether chances are intrinsic. In this regard, an important distinction is whether or not a theory of chance is *Humean*. Roughly speaking, a Humean theory is one in which the chances are defined in terms of the facts on the ground, i.e. facts about the universe that could form part of our experience of it (see [145] for a more precise definition). This means that objective chances cannot be defined in terms of things like $|\psi\rangle$, which do not form part of our experience. Lewis [144] was a proponent of Humean theories of chance and his favoured *best system* theory is really just a modification of frequentism. More specifically, he thought that chances were specified via a tradeoff between accurately capturing the relative frequencies and simplicity. Thus, if a large number of coin flips is performed several different times, and on each occasion the relative frequency of heads obtained was close to, but not exactly, 1/2, varying in a seemingly random way, Lewis would say that this licenses assigning an objective chance of 1/2 to the coin flips. This differs from the frequency theory in that it does not demand that probabilities are exactly the relative frequencies in some real or hypothetical, finite or infinite, sequence of experiments. Of course, defining an objective tradeoff between predictability and simplicity is difficult, but for present purposes all that matters is that again objective chances are not intrinsic properties of systems, but are instead defined with respect to an ensemble. Likewise, I think that any Humean theory of chance makes chance a non-intrinsic property because chances would have to be defined in terms of observable facts, and I do not see how this could be done without referring to ensembles or to the surroundings of the system.

Therefore, the only chance theories that really pose a problem for the distinction we are trying to make are those that are non-Humean, i.e. they posit that chances do not supervene on facts that could form part of our experience. Of these, the most prominent is Popper's propensity theory [143]. Propensities are dispositional properties, i.e. a system has a disposition to produce a certain outcome in an experiment. Propensity theories are broadly classified as either long-run propensity theories or single-case propensity theories. In a long-run theory, a propensity is read as a disposition to produce a certain relative frequency of outcomes in the long run. Since this refers to an ensemble, again there is no problem distinguishing this type of property from intrinsic properties of individual systems. On the other hand, single case propensities are read as a disposition to produce a certain outcome in a single experiment. These are much more problematic for the distinction we are trying to make, as they do not refer to entities other than the individual system. Thus, in a single case propensity theory, it may not be possible to make a clean distinction that is analogous to the Bayesian distinction between ontic and epistemic states.

Before concluding, note that many philosophers take a more laissez faire approach to objective chances.

Whilst they believe that objective chances exist, they do not commit to a specific theory and are instead content to specify some rules that they must obey, such as Lewis' principal principle [144]. This is a perfectly reasonable attitude to take, but it is unreasonable to expect that the question of whether chances are intrinsic properties can be settled at this level of generality. Some further features of objective chance would need to be specified, such as whether or not they ought to be Humean.

In summary, most theories of objective chance seem to admit a distinction equivalent to the Bayesian distinction between ontic and epistemic states in that chances refer to non intrinsic facts about a system. The problematic theories are single-case, such as the single-case propensity theory. However, it seems a bit of a stretch to adopt this theory in order to avoid investigating the question of whether quantum states are ontic or epistemic, particularly since the interpretation of quantum theory is a prime motivation for introducing objective chances in the first place.

B The Kochen-Specker model

This appendix proves that the Kochen-Specker model of Example 4.6 reproduces the quantum predictions and is maximally ψ-epistemic. Recall that, in the Kochen-Specker model, the quantum state $[\psi]$ is represented by a unique probability measure

$$\mu(\Omega) = \int_\Omega p(\vec{\lambda}) \sin\vartheta d\vartheta d\varphi, \tag{259}$$

over the Bloch sphere, where the density p is given by

$$p(\vec{\lambda}) = \frac{1}{\pi}\Theta\left(\vec{\psi}\cdot\vec{\lambda}\right)\vec{\psi}\cdot\vec{\lambda}, \tag{260}$$

and Θ is the Heaviside step function. The response functions for a measurement $\left\{[\phi], \left[\phi^\perp\right]\right\}$ are given by

$$\xi^M_\phi(\lambda) = \Theta(\vec{\phi}\cdot\vec{\lambda}) \tag{261}$$

$$\xi^M_{\phi^\perp}(\lambda) = 1 - \xi^M_\phi(\lambda). \tag{262}$$

To prove that this reproduces the quantum predictions, we need to show that, for any pair of states $[\psi]$ and $[\phi]$,

$$\int_\Lambda \xi^M_\phi(\vec{\lambda}) d\mu(\vec{\lambda}) = \mathrm{Tr}\left([\phi][\psi]\right). \tag{263}$$

The corresponding equation for $\left[\phi^\perp\right]$ will then be automatically satisfied because

$$\int_\Lambda \xi^M_{\phi^\perp}(\vec{\lambda}) d\mu(\vec{\lambda}) = 1 - \int_\Lambda \xi^M_\phi(\vec{\lambda}) d\mu(\vec{\lambda}). \tag{264}$$

To prove Eq. (263), it is convenient to choose a parameterization of the Bloch sphere such that both $\vec{\psi}$ and $\vec{\phi}$ lie on the equator. We can further choose $\vec{\psi}$ to point along the x axis so that $\vec{\psi} = (1,0,0)$ and $\vec{\phi} = (\cos\varphi_\phi, \sin\varphi_\phi, 0)$ for some angle $-\pi < \varphi_\phi \leq \pi$. Using Eq. (24), this means that the right hand side of Eq. (263) is

$$\mathrm{Tr}\left([\phi][\psi]\right) = |\langle\phi|\psi\rangle|^2 = \frac{1}{2}\left(1 + \cos(\varphi_\phi)\right). \tag{265}$$

Now, expanding the left hand side of Eq. (263) gives

$$\int_\Lambda \xi^M_\phi(\vec{\lambda}) d\mu(\vec{\lambda}) = \frac{1}{\pi}\int_\Lambda \Theta(\vec{\phi}\cdot\vec{\lambda})\Theta\left(\vec{\psi}\cdot\vec{\lambda}\right)\vec{\psi}\cdot\vec{\lambda}\sin\vartheta d\vartheta d\varphi. \tag{266}$$

Since $\vec{\lambda} = (\sin\vartheta\cos\varphi, \sin\vartheta\sin\varphi, \cos\vartheta)$, we have

$$\vec{\psi}\cdot\vec{\lambda} = \sin\vartheta\cos\varphi \tag{267}$$

$$\vec{\phi}\cdot\vec{\lambda} = \sin\vartheta\cos\varphi\cos\varphi_\phi + \sin\vartheta\sin\varphi\sin\varphi_\phi \tag{268}$$

$$= \sin\vartheta\cos\left(\varphi - \varphi_\phi\right). \tag{269}$$

Due to he Heaviside step functions, we only need to integrate over the region where both of these are positive. This is the intersection of $-\frac{\pi}{2} < \varphi < \frac{\pi}{2}$ and $-\frac{\pi}{2} + \varphi_\phi < \varphi < \frac{\pi}{2} + \varphi_\phi$. When φ_ϕ is positive this is the interval $-\frac{\pi}{2} + \varphi_\phi < \varphi < \frac{\pi}{2}$ and when φ_ϕ is negative this is the interval $-\frac{\pi}{2} < \vartheta < \frac{\pi}{2} + \varphi_\phi$. Consider first the case where φ_ϕ is positive. Then Eq. (266) reduces to

$$\int_\Lambda \xi_\phi^M(\vec{\lambda}) d\mu(\vec{\lambda}) = \int_0^\pi \sin^2\vartheta d\vartheta \int_{-\frac{\pi}{2}+\varphi_\phi}^{\frac{\pi}{2}} \cos\varphi d\varphi \tag{270}$$

$$= \frac{1}{2}\left[\sin\varphi\right]_{-\frac{\pi}{2}+\varphi_\phi}^{\frac{\pi}{2}} \tag{271}$$

$$= \frac{1}{2}\left(1 + \sin\left(\frac{\pi}{2} - \varphi_\phi\right)\right) \tag{272}$$

$$= \frac{1}{2}\left(1 + \cos\varphi_\phi\right), \tag{273}$$

as required. The case where φ_ϕ is negative gives the same result because sin is an odd function.

To prove that the model is maximally ψ-epistemic, let ν be the probability measure associated with $[\phi]$, i.e.

$$\nu(\Omega) = \int_\Omega q(\vec{\lambda}) \sin\vartheta d\vartheta d\varphi, \tag{274}$$

where

$$q(\vec{\lambda}) = \frac{1}{\pi}\Theta\left(\vec{\phi}\cdot\vec{\lambda}\right)\vec{\phi}\cdot\vec{\lambda}. \tag{275}$$

We then need to show that

$$\int_\Omega \xi_\phi^M(\vec{\lambda}) d\mu(\vec{\lambda}) = \int_\Lambda \xi_\phi^M(\vec{\lambda}) d\mu(\vec{\lambda}), \tag{276}$$

for any Ω such that $\nu(\Omega) = 1$. Assume that φ_ϕ is positive (the negative case follows the same logic). Let,

$$\Omega_\psi = \left\{\vec{\lambda} \in \Lambda \middle| 0 < \vartheta < \pi, -\frac{\pi}{2} < \varphi < \frac{\pi}{2}\right\} \tag{277}$$

$$\Omega_\phi = \left\{\vec{\lambda} \in \Lambda \middle| 0 < \vartheta < \pi, -\frac{\pi}{2} + \varphi_\phi < \varphi < \frac{\pi}{2} + \varphi_\phi\right\}. \tag{278}$$

Note that, for any measurable set Ω

$$\int_\Omega \xi_\phi^M(\vec{\lambda}) d\mu(\vec{\lambda}) = \int_{\Omega\cap\Omega_\psi} \xi_\phi^M(\vec{\lambda}) d\mu(\vec{\lambda}), \tag{279}$$

because $p(\vec{\lambda})$ is zero outside Ω_ψ. Note also that Ω_ϕ is a measure one set according to ν because $q(\vec{\lambda})$ is zero outside this set. However, in proving that the model reproduces the quantum predictions, we showed that

$$\int_\Lambda \xi_\phi^M(\vec{\lambda}) d\mu(\vec{\lambda}) = \int_{\Omega_\psi\cap\Omega_\phi} \xi_\phi^M(\vec{\lambda}) d\mu(\vec{\lambda}) = \mathrm{Tr}\left([\phi][\psi]\right), \tag{280}$$

and hence

$$\int_{\Omega_\phi} \xi_\phi^M(\vec{\lambda}) d\mu(\vec{\lambda}) = \int_{\Omega_\psi\cap\Omega_\phi} \xi_\phi^M(\vec{\lambda}) d\mu(\vec{\lambda}) = \int_\Lambda \xi_\phi^M(\vec{\lambda}) d\mu(\vec{\lambda}), \tag{281}$$

so we have the required property for the special case of the set Ω_ϕ.

Now let Ω be any other set that is of measure one according to ν. We can write Ω as the union of two disjoint sets via

$$\Omega = (\Omega\cap\Omega_\phi) \cup (\Omega\backslash\Omega_\phi). \tag{282}$$

The set $\Omega\backslash\Omega_\phi$ is of measure zero according to ν because $q(\vec{\lambda})$ is zero outside $\Omega\phi$. This means that $\Omega\cap\Omega_\phi$ is of measure one. Further, ξ_ϕ^M is also zero outside Ω_ϕ so

$$\int_{\Omega\backslash\Omega_\phi} \xi_\phi^M(\vec{\lambda}) d\mu(\vec{\lambda}) = 0. \tag{283}$$

Therefore, we only need to show that

$$\int_{\Omega\cap\Omega_\phi} \xi_\phi^M(\vec{\lambda})d\mu(\vec{\lambda}) = \int_{\Omega_\phi} \xi_\phi^M(\vec{\lambda})d\mu(\vec{\lambda}). \tag{284}$$

To see this, note that μ and ν are absolutely continuous with respect to one another on $\Omega_\psi \cap \Omega_\phi$. Since, $\Omega \cap \Omega_\phi$ is of measure one according to ν, $\Omega_\phi \setminus (\Omega \cap \Omega_\phi)$ is of measure zero according to ν and hence, by absolute continuity, $\Omega_\psi \cap (\Omega_\phi \setminus (\Omega \cap \Omega_\phi))$ is of measure zero according to both ν and μ. Thus,

$$\int_{\Omega_\phi} \xi_\phi^M(\vec{\lambda})d\mu(\vec{\lambda}) = \int_{\Omega_\psi\cap\Omega_\phi} \xi_\phi^M(\vec{\lambda})d\mu(\vec{\lambda}) \tag{285}$$

$$= \int_{\Omega_\psi\cap\Omega_\phi\cap\Omega} \xi_\phi^M(\vec{\lambda})d\mu(\vec{\lambda}) + \int_{\Omega_\psi\cap(\Omega_\phi\setminus(\Omega\cap\Omega_\phi))} \xi_\phi^M(\vec{\lambda})d\mu(\vec{\lambda}) \tag{286}$$

$$= \int_{\Omega_\psi\cap\Omega_\phi\cap\Omega} \xi_\phi^M(\vec{\lambda})d\mu(\vec{\lambda}) + 0 \tag{287}$$

$$= \int_{\Omega\cap\Omega_\phi} \xi_\phi^M(\vec{\lambda})d\mu(\vec{\lambda}), \tag{288}$$

as required.

References

[1] J. S. Bell. On the Einstein Podolsky Rosen paradox. *Physics*, 1(3):195–200, 1964. URL: `http://philoscience.unibe.ch/documents/TexteHS10/bell1964epr.pdf`.

[2] A. Shimony. Contextual hidden variables theories and Bell's inequalities. *British Journal for Philosophy of Science*, 35:25–45, 1984. `doi:10.1093/bjps/35.1.25`.

[3] G. Bacciagaluppi and A. Valentini. *Quantum Theory at the Crossroads: Reconsidering the 1927 Solvay Conference*. Cambridge University Press, 2009. `arXiv:quant-ph/0609184`.

[4] M. Born. On the quantum mechanics of collisions. In J. A. Wheeler and W. H. Zurek, editors, *Quantum Theory and Measurement*, chapter I.2. Princeton University Press, 1983.

[5] J. Faye. Copenhagen interpretation of quantum mechanics. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Fall 2008 edition, 2008. URL: `http://plato.stanford.edu/archives/fall2008/entries/qm-copenhagen`.

[6] H. Everett. "Relative state" formulation of quantum mechanics. *Reviews of Modern Physics*, 29:454–462, 1957. `doi:10.1103/RevModPhys.29.454`.

[7] B. S. DeWitt and R. N. Graham, editors. *The Many-Worlds Interpretation of Quantum Mechanics*. Princeton University Press, 1973.

[8] D. Wallace. *The Emergent Multiverse: Quantum Theory according to the Everett Interpretation*. Oxford University Press, 2012.

[9] L. de Broglie. The new dynamics of quanta. In G. Bacciagaluppi and A. Valentini, editors, *Quantum Theory at the Crossroads: Reconsidering the 1927 Solvay Conference*, pages 373–406. Cambridge University Press, 2009. (page nos. refer to arXiv version). `arXiv:quant-ph/0609184`.

[10] D. Bohm. A suggested interpretation of the quantum theory in terms of "hidden" variables. i. *Physical Review*, 85:166–179, 1952. `doi:10.1103/PhysRev.85.166`.

[11] D. Bohm. A suggested interpretation of the quantum theory in terms of "hidden" variables. ii. *Physical Review*, 85(2):180–193, 1952. doi:10.1103/PhysRev.85.180.

[12] D. Dürr and S. Teufel. *Bohmian Mechanics*. Springer, 2009. URL: http://www.springer.com/physics/quantum+physics/book/978-3-540-89343-1.

[13] G. C. Ghirardi, A. Rimini, and T. Weber. Unified dynamics for microscopic and macroscopic systems. *Physical Review D*, 34(2):470–491, 1986. doi:10.1103/PhysRevD.34.470.

[14] A. Bassi, K. Lochan, S. Satin, T. P. Singh, and H. Ulbricht. Models of wave-function collapse, underlying theories, and experimental tests. *Reviews of Modern Physics*, 85(2):471–527, 2013. arXiv:1204.4325, doi:10.1103/RevModPhys.85.471.

[15] O. Lombardi and D. Dieks. Modal interpretations of quantum mechanics. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Fall 2013 edition, 2013. URL: http://plato.stanford.edu/archives/fall2013/entries/qm-modal/.

[16] C. A. Fuchs. QBism, the perimeter of quantum bayesianism. 2010. arXiv:1003.5209.

[17] C. A. Fuchs. Quantum bayesianism at the perimeter. *Physics in Canada*, 66(2):77–81, 2010. URL: http://www.cap.ca/en/article/quantum-bayesianism-perimeter, arXiv:1003.5182.

[18] C. A. Fuchs, N. D. Mermin, and R. Schack. An introduction to QBism with an application to the locality of quantum mechanics. 2013. arXiv:1311.5253.

[19] I. Pitowsky. Quantum mechanics as a theory of probability. In W. Demopoulos and I. Pitowsky, editors, *Physical Theory and its Interpretation: Essays in Honour of Jeffrey Bub*, volume 72 of *The Western Ontario Series in Philosophy of Science*, chapter 10, pages 213–240. Springer, 2006. arXiv:quant-ph/0510095.

[20] J. Bub and I. Pitowsky. Two dogmas about quantum mechanics. In S. Saunders, J. Barrett, A. Kent, and D. Wallace, editors, *Many Worlds? Everett, Quantum Theory and Reality*, chapter 14, pages 433–459. Oxford University Press, 2010. arXiv:0712.4258, doi:10.1093/acprof:oso/9780199560561.003.0016.

[21] R. Healey. Quantum theory: A pragmatist approach. *British Journal for The Philosophy of Science*, 63(4):729–771, 2012. PhilSci:http://philsci-archive.pitt.edu/8620/. arXiv:1008.3896, doi:10.1093/bjps/axr054.

[22] C. Rovelli. Relational quantum mechanics. *International Journal of Theoretical Physics*, 35(8):1637–1678, 1996. arXiv:quant-ph/9609002, doi:10.1007/BF02302261.

[23] W. M. de Muynck. *Foundations of Quantum Mechanics, an Empricist Approach*, volume 127 of *Fundamental Theories of Physics*. Kluwer, 2002.

[24] N. D. Mermin. Annotated interview with a QBist in the making. Based on Mermin's contribution to M. Schlosshauer, editor, *Elegance and Enigma: The Quantum Interviews*. Springer, 2011., 2013. arXiv:1301.6551.

[25] A. Peres. *Quantum Theory: Concepts and Methods*, volume 72 of *Funamental Theories of Physics*. Kluwer, 1995.

[26] C. Brukner and A. Zeilinger. Information and fundamental elements of the structure of quantum theory. In L. Castell and O. Ischebeck, editors, *Time, Quantum and Information*. Springer, 2003. arXiv:quant-ph/0212084.

[27] M. F. Pusey, J. Barrett, and T. Rudolph. On the reality of the quantum state. *Nature Physics*, 8:475–478, 2012. arXiv:1111.3328, doi:10.1038/nphys2309.

[28] J. Bub. Quantum correlations and the measurement problem. 2012. arXiv:arXiv:1210.6371.

[29] J. Bub. Bananaworld: Quantum mechanics for primates. 2012. arXiv:1211.3062.

[30] A. Drezet. Can quantum mechanics be considered as statistical? an analysis of the PBR theorem. 2012. arXiv:1203.2475.

[31] A. Drezet. Should the wave-function be a part of the quantum ontological state? *Progress in Physics*, 4, 2012. URL: http://www.ptep-online.com/index_files/2012/PP-31-04.PDF, arXiv:1209.2565.

[32] J. Emerson, D. Serbin, C. Sutherland, and V. Veitch. The whole is greater than the sum of the parts: on the possibility of purely statistical interpretations of quantum theory. 2013. arXiv:1312.1345.

[33] Michael J. W. Hall. Generalisations of the recent Pusey-Barrett-Rudolph theorem for statistical models of quantum phenomena. 2011. arXiv:1111.6304.

[34] H. F. Hofmann. The quantum state should be interpreted statistically. 2011. arXiv:1112.2446.

[35] S. Mansfield. On the reality of observable properties. 2013. arXiv:1306.3216.

[36] D. J. Miller. Alternative experimental protocol to demonstrate the Pusey-Barrett-Rudolph theorem. *Physical Review A*, 87(1):014103, 2013. arXiv:1202.6465, doi:10.1103/PhysRevA.87.014103.

[37] M. Schlosshauer and A. Fine. A no-go theorem for the composition of quantum systems. 2013. arXiv:1306.5805.

[38] M. Schlosshauer and A. Fine. Implications of the Pusey-Barrett-Rudolph quantum no-go theorem. *Phys. Rev. Lett.*, 108(26):260404, 2012. arXiv:1203.4779, doi:10.1103/PhysRevLett.108.260404.

[39] Petros Wallden. Distinguishing initial state-vectors from each other in histories formulations and the pbr argument. *Foundations of Physics*, 2013. Published online first. arXiv:1211.2084, doi:10.1007/s10701-013-9759-8.

[40] S. Aaronson. The quantum state cannot be interpreted as something other than a quantum state, Nov 2011. URL: http://www.scottaaronson.com/blog/?p=822.

[41] R. B. Griffiths. Who needs hidden variables? *The Quantum Times (Newsletter of the APS Topical Group on Quantum Information)*, 6(4):4–5, 2012. URL: http://www.aps.org/units/gqi/newsletters/upload/vol6num4.pdf.

[42] M. S. Leifer. PBR, EPR and all that jazz. *The Quantum Times (Newsletter of the APS Topical Group on Quantum Information)*, 6(3):1–3, 2011. URL: http://www.aps.org/units/gqi/newsletters/upload/vol6num3.pdf.

[43] M. S. Leifer. Response to Griffiths. *The Quantum Times (Newsletter of the APS Topical Group on Quantum Information)*, 6(4):6, 2012. URL: http://www.aps.org/units/gqi/newsletters/upload/vol6num4.pdf.

[44] M. S. Leifer. Can the quantum state be interpreted statistically?, Nov 2011. URL: http://mattleifer.info/2011/11/20/can-the-quantum-state-be-interpreted-statistically/.

[45] L. Motl. Nature hypes anti-QM crackpot paper by Pusey et. al., Nov 2011. URL: http://motls.blogspot.co.uk/2011/11/nature-hypes-anti-qm-crackpot-paper-by.html.

[46] E. S. Reich. A boost for quantum reality. *Nature*, 485(7397):157–158, May 2012. doi:10.1038/485157a.

[47] E. S. Reich. Quantum theorem shakes foundations. *Nature*, 2011. doi:10.1038/nature.2011.9392.

[48] R. W. Spekkens. Why I am not a psi-ontologist. video of talk given at Perimeter Institute 08/05/2012, 2012. URL: http://pirsa.org/12050021/.

[49] D. Wallace. Guest post: David Wallace on the physicality of the quantum state, Nov 2011. URL: http://blogs.discovermagazine.com/cosmicvariance/2011/11/18/guest-post-david-wallace-on-the-physicality-of-the-quantum-state/.

[50] R. Colbeck and R. Renner. Is a system's wave function in one-to-one correspondence with its elements of reality? *Phys. Rev. Lett.*, 108(15):150402, 2012. arXiv:1111.6597, doi:10.1103/PhysRevLett.108.150402.

[51] L. Hardy. Are quantum states real? *International Journal of Modern Physics B*, 27:1345012, 2013. arXiv:1205.1439, doi:10.1142/S0217979213450124.

[52] P. G. Lewis, D. Jennings, J. Barrett, and T. Rudolph. Distinct quantum states can be compatible with a single state of reality. *Phys. Rev. Lett.*, 108(15):150404, 2012. arXiv:1201.6554, doi:10.1103/PhysRevLett.109.150404.

[53] S. Aaronson, A. Bouland, L. Chua, and G. Lowther. Psi-epistemic theories: The role of symmetry. 2013. arXiv:1303.2834.

[54] J. Barrett, E. G. Cavalcanti, R. Lal, and O. J. E. Maroney. No ψ-epistemic model can fully explain the indistinguishability of quantum states. 2013. arXiv:1310.8302.

[55] M. S. Leifer and O. J. E. Maroney. Maximally epistemic interpretations of the quantum state and contextuality. *Physical Review Letters*, 110(12):120401, 2013. arXiv:1208.5132, doi:10.1103/PhysRevLett.110.120401.

[56] O. J. E. Maroney. How statistical are quantum states? 2012. arXiv:1207.6906.

[57] O. J. E. Maroney. A brief note on epistemic interpretations and the Kochen-Speker theorem. 1207.7192, 2012.

[58] M. K. Patra, L. Olisager, F. Duport, J. Safioui, S. Pironio, and S. Massar. Experimental refutation of a class of ψ-epistemic models. *Physical Review A*, 88(3):032112, 2013. arXiv:1306.0414, doi:10.1103/PhysRevA.88.032112.

[59] L. Hardy. Quantum ontological excess baggage. *Stud. Hist. Phil. Mod. Phys.*, 35(2):267–276, 2004. doi:10.1016/j.shpsb.2003.12.001.

[60] A. Montina. Exponential complexity and ontological theories of quantum mechanics. *Phys. Rev. A*, 77(2):022104, 2008. arXiv:0711.4770, doi:10.1103/PhysRevA.77.022104.

[61] A. Montina. State-space dimensionality in short-memory hidden-variable theories. *Phys. Rev. A*, 83(3):032107, 2011. arXiv:1008.4415, doi:10.1103/PhysRevA.83.032107.

[62] R. W. Spekkens. Evidence for the epistemic view of quantum states: A toy theory. *Phys. Rev. A*, 75(3):032110, 2007. arXiv:quant-ph/0401052, doi:10.1103/PhysRevA.75.032110.

[63] O. Schreiber and R. W. Spekkens. The power of epistemic restrictions in axiomatizing quantum theory: from trits to qutrits. Unpublished work. Videos of talks discussing this material are available: R.W. Spekkens, Talk, July 17, University of Oxford, http://categorieslogicphysics.wikidot.com/people#robspekkens. R. W. Spekkens, Talk, Aug. 10, 2008, Perimeter Institute, http://pirsa.org/09080009/.

[64] S. D. Bartlett, T. Rudolph, and R. W. Spekkens. Reconstruction of gaussian quantum mechanics from liouville mechanics with an epistemic restriction. *Phys. Rev. A*, 86(1):012103, 2012. arXiv:1111.5057, doi:10.1103/PhysRevA.86.012103.

[65] R. W. Spekkens. Contextuality for preparations, transformations, and unsharp measurements. *Physical Review A*, 71(5):052108, 2005. arXiv:quant-ph/0406166, doi:10.1103/PhysRevA.71.052108.

[66] P. A. M. Dirac. *The Principles of Quantum Mechanics*. Oxford University Press, fourth edition, 1958.

[67] J. von Neumann. *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, 1955.

[68] N. Harrigan and R. W. Spekkens. Einstein, incompleteness, and the epistemic view of quantum states. *Found. Phys.*, 40(2):125–157, 2010. arXiv:0706.2661, doi:10.1007/s10701-009-9347-0.

[69] L. E. Ballentine. The statistical interpretation of quantum mechanics. *Rev. Mod. Phys.*, 42(4):358–381, 1970. doi:10.1103/RevModPhys.42.358.

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

[71] A. Chefles and S. M. Barnett. Quantum state separation, unambiguous discrimination and exact cloning. *Journal of Physics A: Mathematical and General*, 31(50):10097, 1998. arXiv:quant-ph/9808018, doi:10.1088/0305-4470/31/50/007.

[72] D. Gottesman. The Heisenberg representation of quantum computers. In S. P. Corney, R. Delbourgo, and P. D. Jarvis, editors, *Group 22: Proceedings of the 12th International Colloquium on Group Theoretical Methods in Physics*, pages 32–43. International Press, 1999. arXiv:quant-ph/9807006.

[73] M. F. Pusey. Stabilizer notation for Spekkens' toy theory. *Foundations of Physics*, 42(5):688–708, 2012. arXiv:1103.5037, doi:10.1007/s10701-012-9639-7.

[74] M. Rédei and S. J. Summers. Quantum probability theory. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 38(2):390–417, 2007. arXiv:quant-ph/0601158, doi:10.1016/j.shpsb.2006.05.006.

[75] D. Petz. *Quantum Information Theory and Quantum Statistics*. Theoretical and Mathematical Physics. Springer, 2008.

[76] B. Schumacher. Quantum coding. *Physical Review A*, 51(4):2738, 1995. doi:10.1103/PhysRevA.51.2738.

[77] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[78] Complexity zoo wiki. Accessed 30th Nov. 2013. URL: https://complexityzoo.uwaterloo.ca/Complexity_Zoo.

[79] M. S. Leifer and R. W. Spekkens. Towards a formulation of quantum theory as a causally neutral theory of bayesian inference. *Physical Review A*, 88(5):052130, 2013. arXiv:1107.5849, doi:10.1103/PhysRevA.88.052130.

[80] M. S. Leifer. "It from bit" and the quantum probability rule. Winning entry in 2013 FQXi essay contest: It From Bit or Bit From It? eprint arXiv:1311.0857, 2013. URL: http://fqxi.org/community/forum/topic/1821, arXiv:1311.0857.

[81] A. S. Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problems of Information Transmission*, 9:177–183, 1973.

[82] C. H. Bennett and S. J. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Physical Review Letters*, 69(20):2881–2884, 1992. doi:10.1103/PhysRevLett.69.2881.

[83] A. Ambainis, A. Nayak, A. Ta-Shma, and U. Vazirani. Dense quantum coding and a lower bound for 1-way quantum automata. In *STOC '99 Proceedings of the thirty-first annual ACM symposium on Theory of computing*, pages 376–383. ACM, 1999. arXiv:quant-ph/9804043, doi:10.1145/301250.301347.

[84] R. P. Feynman, R. B. Leighton, and M. Sands. *The Feynman Lectures on Physics: Volume III*, chapter 1-1, pages 1–1. Basic Books, millenium edition, 2010. URL: http://www.feynmanlectures.caltech.edu/III_01.html#Ch1-S1.

[85] E. Martin, M. S. Leifer, and R. W. Spekkens. Interference phenomena in episepistemic restricted theories. In preparation. The theory is briefly described in R. W. Spekkens, Talk, Feb. 27, 2008, Perimeter Institute, http://pirsa.org/08020051/, 2014.

[86] J. A. Wheeler. The 'past' and the 'delayed-choice' double-slit experiment. In A. R. Marlow, editor, *Mathematical Foundations of Quantum Theory*, pages 9–48. Elsevier, 1978. doi:10.1016/B978-0-12-473250-6.50006-6.

[87] A. C. Elitzur and L. Vaidman. Quantum mechanical interaction-free measurements. *Foundations of Physics*, 23(7):987–997, 1993. arXiv:hep-th/9305002, doi:10.1007/BF00736012.

[88] C. A. Fuchs. Penrose tiles. In *Coming of Age with Quantum Information: Notes on a Paulian Idea*, pages 402–408. Cambridge University Press, 2011. (page nos. refer to arXiv version). arXiv:quant-ph/0105039.

[89] D. Deutsch. David deutschs many worlds. *Frontiers magazine*, December 1998. URL: http://daviddeutsch.physics.ox.ac.uk/structure/Documents/Non-Technical/Frontiers.html.

[90] P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, 1997. arXiv:quant-ph/9508027, doi:10.1137/S0097539795293172.

[91] T. Heinosaari and M. Ziman. *The Mathematical Language of Quantum Theory: From Uncertainty to Entanglement*. Cambridge University Press, 2011.

[92] N. Harrigan and T. Rudolph. Ontological models and the interpretation of contextuality. 2007. arXiv:0709.4266.

[93] E. G. Beltrametti and S. Bugajski. A classical extension of quantum mechanics. *Journal of Physics A: Mathematical and General*, 28(12):3329, 1995. doi:10.1088/0305-4470/28/12/007.

[94] W. Stulpe. Topologies and measurable structure on the projective Hilbert space. 2007. arXiv:0708.1208.

[95] J. S. Bell. On the problem of hidden variables in quantum mechanics. *Reviews of Modern Physics*, 38(3):447–452, 1966. doi:10.1103/RevModPhys.38.447.

[96] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner. Bell nonlocality. 2013. arXiv:1303.2849.

[97] N. D. Mermin. Simple unified form for the major no-hidden-variables theorems. *Physical Review Letters*, 65(27):3373–3376, 1990. doi:10.1103/PhysRevLett.65.3373.

[98] A. Peres. Two simple proofs of the Kochen-Specker theorem. *Journal of Physics A: Mathematical and General*, 24(4):L175–L178, 1991. doi:10.1088/0305-4470/24/4/003.

[99] R. Clifton. Getting contextual and nonlocal elements-of-reality the easy way. *American Journal of Physics*, 61:443, 1993. doi:10.1119/1.17239.

[100] A. Cabello, J. M. Estebaranz, and G. García-Alcaine. Bell-Kochen-Specker theorem: A proof with 18 vectors. *Physics Letters A*, 212(4):183–187, 1996. arXiv:quant-ph/9706009, doi:10.1016/0375-9601(96)00134-X.

[101] A. A. Klyachko, M. A. Can, S. Binicioğlu, and A. S. Shumovsky. Simple test for hidden variables in spin-1 systems. *Physical Review Letters*, 101(2):020403, 2008. arXiv:0706.0126, doi:10.1103/PhysRevLett.101.020403.

[102] A. Cabello, S. Severini, and A. Winter. (Non-)contextuality of physical theories as an axiom. 2010. arXiv:1010.2163.

[103] Y. C. Liang, R. W. Spekkens, and H. Wiseman. Specker's parable of the overprotective seer: A road to contextuality, nonlocality and complementarity. *Physics Reports*, 506(1–2):1–39, 2011. arXiv:1010.1273, doi:10.1016/j.physrep.2011.05.001.

[104] M. S. Leifer. ψ-epistemic models are exponentially bad at explaining the distinguishability of quantum states. in preparation, 2013.

[105] J. Barrett. Any bipartite Bell-type proof of non-locality can be converted into a proof of preparation contextuality. Private communication, 2004.

[106] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 47(10):777–780, 1935. doi:10.1103/PhysRev.47.777.

[107] C. M. Caves, C. A. Fuchs, and R. Schack. Conditions for compatibility of quantum state assignments. *Phys. Rev. A*, 66(6):062111, 2002. arXiv:quant-ph/0206110, doi:10.1103/PhysRevA.66.062111.

[108] S. Bandyopadhyay, R. Jain, J. Oppenheim, and Perry C. Conclusive exclusion of quantum states. 2013. arXiv:1306.4683.

[109] D. Howard. Einstein on locality and separability. *Studies in History and Philosophy of Science*, 16(3):171–201, 1985. doi:10.1016/0039-3681(85)90001-9.

[110] A. Einstein. Quanten-mechanik und wirklichkeit. *Dialectica*, 2(3–4):320–324, 1948. English translation by Howard in [109]. doi:10.1111/j.1746-8361.1948.tb00704.x.

[111] J. Henson. Non-separability does not relieve the problem of Bell's theorem. *Foundations of Physics*, 43(8):1008–1038, 2013. arXiv:1302.7188, doi:10.1007/s10701-013-9730-8.

[112] D. Dürr, S. Goldstein, and N. Zanghi. Quantum equilibrium and the origin of absolute uncertainty. *Journal of Statistical Physics*, 67(5–6):843–907, 1992. arXiv:quant-ph/0308039, doi:10.1007/BF01049004.

[113] S. Goldstein. Quantum theory witwith observers—part one. *Physics Today*, 51(3):42–46, March 1998. doi:10.1063/1.882184.

[114] S. Goldstein. Quantum theory witwith observers—part two. *Physics Today*, 51(4):38–42, April 1998. doi:10.1063/1.882241.

[115] J. S. Bell. The theory of local beables. In *Speakable and unspeakable in quantum mechanics*, pages 52–62. Cambridge University Press, 2004. URL: http://cdsweb.cern.ch/record/980036/files/197508125.pdf.

[116] A. J. Leggett. Testing the limits of quantum mechanics: motivation, state of play, prospects. *Journal of Physics: Condensed Matter*, 14(15):R415–R451, 2002. URL: http://physics.illinois.edu/people/Leggett/PhysicaC-2002.pdf, doi:10.1088/0953-8984/14/15/201.

[117] E. Wigner. On the quantum correction for thermodynamic equilibrium. *Physical Review*, 40(5):749–759, 1932. doi:10.1103/PhysRev.40.749.

[118] M. K. Patra, S. Pironio, and S. Massar. No-go theorem for ψ-epistemic models based on a continuity assumption. *Physical Review Letters*, 111(9):090402, 2013. arXiv:1211.1179, doi:10.1103/PhysRevLett.111.090402.

[119] R. Renner. Talk: Does freedom of choice imply that the wavefunction is real? Online Q+ hangout, 29th October 2013. URL: http://www.youtube.com/watch?v=wWJz2gJ13xI.

[120] R. Colbeck and R. Renner. No extension of quantum theory can have improved predictive power. *Nature Communications*, 2:411, 2011. arXiv:1005.5173, doi:10.1038/ncomms1416.

[121] R. Colbeck and R. Renner. The completeness of quantum theory for predicting measurement outcomes. 2012. arXiv:1208.4123.

[122] G. Ghirardi and R. Romano. On the completeness of quantum mechanics and the interpretation of the state vector. *Journal of Physics: Conference Series*, 442(1):012002, 2013. 1302.6278. doi:10.1088/1742-6596/442/1/012002.

[123] G. Ghirardi and R. Romano. About possible extensions of quantum theory. *Found. Phys.*, 43(7):881–894, 2013. arXiv:1301.5040, doi:10.1007/s10701-013-9724-6.

[124] G. Ghirardi and R. Romano. Comment on "is a system's wave function in one-to-one correspondence with its elements of reality?". 2013. arXiv:1302.1635.

[125] J. Jarrett. On the physical significance of the locality conditions in the Bell arguments. *Noûs*, 18(4):569–589, 1984. URL: http://www.jstor.org/stable/2214878.

[126] A. Shimony. Controllable and uncontrollable non-locality. In *Search for a Naturalistic World View*, volume 2, chapter 10, pages 130–139. Cambridge University Press, 1993. doi:10.1017/CBO9781139172196.010.

[127] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt. Proposed experiment to test local hidden-variable theories. *Physical Review Letters*, 23(15):880–884, 1969. doi:10.1103/PhysRevLett.23.880.

[128] S. L. Braunstein and C. M. Caves. Chained Bell inequalities. In M. Kafatos, editor, *Bell's Theorem, Quantum Theory and Conceptions of the Universe*, pages 27–36. Springer, 1989. doi:10.1007/978-94-017-0849-4_5.

[129] S. L. Braunstein and C. M. Caves. Wringing out better Bell inequalities. *Nuclear Physics B – Proceedings Supplements*, 6:211–221, 1989. URL: http://www-users.cs.york.ac.uk/schmuel/papers/90/bcB90.pdf, doi:10.1016/0920-5632(89)90441-6.

[130] A. Montina. Epistemic view of quantum states and communication complexity of quantum channels. *Phys. Rev. Lett.*, 109(11):110501, 2012. arXiv:1206.2961, doi:10.1103/PhysRevLett.109.110501.

[131] D. Nigg, T. Monz, P. Schindler, E. Martinez, M. Chwalla, M. Hennrich, R. Blatt, M. F. Pusey, T. Rudolph, and J. Barrett. Can different quantum state vectors correspond to the same physical state? an experimental test. 2012. arXiv:eprintarXiv:1211.0942.

[132] T. E. Stuart, J. A. Slater, R. Colbeck, R. Renner, and W. Tittel. An experimental test of all theories with predictive power beyond quantum theory. *Phys. Rev. Lett.*, 109(2):020402, 2012. arXiv:1105.0133, doi:10.1103/PhysRevLett.109.020402.

[133] R. B. Griffiths. *Consistent Quantum Theory*. Cambridge University Press, 2002.

[134] R. D. Sorkin. Quantum dynamics without the wavefunction. *Journal of Physics A: Mathematical and Theoretical*, 40(12):3207, 2007. arXiv:quant-ph/0610204, doi:10.1088/1751-8113/40/12/S20.

[135] R. D. Sorkin. An exercise in "anhomomorphic logic". *Journal of Physics: Conference Series*, 67:012018, 2007. arXiv:quant-ph/0703276, doi:10.1088/1742-6596/67/1/012018.

[136] R. B. Griffiths. Epistemic restrictions in Hilbert space quantum mechanics. *Physical Review A*, 88(4):042122, 2013. arXiv:1308.4176, doi:10.1103/PhysRevA.88.042122.

[137] O. Costa de Beauregard. Time symmetry and interpretation of quantum mechanics. *Foundations of Physics*, 6(5):539–559, 1976. doi:10.1007/BF00715107.

[138] H. Price. Does time-symmetry imply retrocausality? how the quantum world says "maybe"? *Studies in History and Philosophy of Modern Physics*, 43(2):75–83, 2012. arXiv:1002.0906, doi:10.1016/j.shpsb.2011.12.003.

[139] H. Price and K. Wharton. Dispelling the quantum spooks – a clue that Einstein missed? 2013. arXiv:1307.7744.

[140] J. G. Cramer. The transactional interpretation of quantum mechanics. *Reviews of Modern Physics*, 58(3):647–687, 1986. doi:10.1103/RevModPhys.58.647.

[141] R. E. Kastner. *The Transactional Interpretation of Quantum Mechanics: The Reality of Possibility*. Cambridge University Press, 2013.

[142] Y. Aharonov and L. Vaidman. The two-state vector formalism of quantum mechanics. In J. G. Muga, R. Sala Mayato, and I. L. Egusquiza, editors, *Time in Quantum Mechanics*, volume 72 of *Lecture Notes in Physics*, chapter 13, pages 369–412. Springer, 2002. arXiv:quant-ph/0105101, doi:10.1007/3-540-45846-8_13.

[143] K. Popper. The propensity interpretation of probability. In A. Eagle, editor, *Philosophy of Probability: Contemporary Readings*, chapter 28, pages 488–497. Routledge, 2011.

[144] D. Lewis. A subjectivist's guide to objective chance. In A. Eagle, editor, *Philosophy of Probability: Contemporary Readings*, chapter 27, pages 458–487. Routledge, 2011.

[145] A. Eagle. Physical probability: Introduction. In A. Eagle, editor, *Philosophy of Probability: Contemporary Readings*, chapter 21, pages 339–354. Rou, 2011.