

## Research



**Cite this article:** Hance JR, Hossenfelder S. 2022 The wave function as a true ensemble. *Proc. R. Soc. A* **478**: 20210705. <https://doi.org/10.1098/rspa.2021.0705>

Received: 4 September 2021

Accepted: 6 June 2022

**Subject Areas:**

quantum physics

**Keywords:**

ensembles, wave function, Bell's Theorem, statistical independence, quantum foundations

**Author for correspondence:**

Jonte R. Hance

e-mail: [jonte.hance@bristol.ac.uk](mailto:jonte.hance@bristol.ac.uk)

# The wave function as a true ensemble

Jonte R. Hance<sup>1</sup> and Sabine Hossenfelder<sup>2</sup>

<sup>1</sup>Quantum Engineering Technology Laboratories, Department of Electrical and Electronic Engineering, University of Bristol, Woodland Road, Bristol BS8 1US, UK

<sup>2</sup>Frankfurt Institute for Advanced Studies, Ruth-Moufang-Str. 1, D-60438 Frankfurt am Main, Germany

JRH, 0000-0001-8587-7618; SH, 0000-0002-2515-3842

In quantum mechanics, the wave function predicts probabilities of possible measurement outcomes, but not which individual outcome is realized in each run of an experiment. This suggests that it describes an ensemble of states with different values of a hidden variable. Here, we analyse this idea with reference to currently known theorems and experiments. We argue that the  $\psi$ -ontic/epistemic distinction fails to properly identify ensemble interpretations and propose a more useful definition. We then show that all local  $\psi$ -ensemble interpretations which reproduce quantum mechanics violate statistical independence. Theories with this property are commonly referred to as superdeterministic or retrocausal. Finally, we explain how this interpretation helps make sense of some otherwise puzzling phenomena in quantum mechanics, such as the delayed choice experiment, the Elitzur–Vaidman bomb detector and the extended Wigner's friends scenario.

## 1. Introduction

The wave function is weird. Its most salient feature—that it merely predicts probabilities for measurement outcomes, rather than the outcomes themselves—suggests that quantum mechanics is an emergent, average description of an underlying, more complicated dynamics. In this underlying theory, the time-evolution of the system would be determined and measurement outcomes could be predicted. It is because we lack information about the details of the initial state that we can only make a probabilistic prediction.

© 2022 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License <http://creativecommons.org/licenses/by/4.0/>, which permits unrestricted use, provided the original author and source are credited.

That the wave function might be an emergent description of yet-to-be-discovered underlying physics is often called a hidden variable interpretation. The hidden variables are the information that is missing in quantum mechanics. This straightforward explanation for the strange properties of the wave function, however, has seemingly been disfavoured by Bell's and related theorems [1,2], and, more recently, by the Pusey–Barrett–Rudolph theorem [3,4] (hereafter PBR-theorem). We here want to look closer at what these theorems actually say about the wave function as an ensemble of different values of the hidden variables.

The PBR-theorem in particular builds on a classification of models put forward by Harrigan & Spekkens [5]. Their framework attempts to mathematically formalize the distinction between a wave function that represents reality ( $\psi$ -ontic) and one that represents knowledge ( $\psi$ -epistemic). That this is a false dichotomy was previously pointed out in [6]. We here want to spell out another problem with the  $\psi$ -ontic/epistemic framework—it does not identify the hidden variables theories that Bell was trying to rule out.

We begin with defining in what sense we take the wave function to be an ensemble in §2, and explain the relation to the  $\psi$ -ontic/epistemic framework in §3. In §4, we will discuss the most important consequence, that is the different definition of what the hidden variables describe, and how to interpret it. In §5, we will revisit the PBR-theorem in this light. In §6, we discuss statistical independence (SI), locality, and the sometimes used classification of superdeterminism and retrocausality for violations of SI. Finally, in §7, we will explain how the  $\psi$ -ensemble interpretation helps make sense of quantum mechanics.

## 2. The $\psi$ -ensemble interpretation

We will begin by explaining what exactly we mean by the wave function being an ensemble. We do not mean the Ensemble, or Statistical, interpretation of quantum mechanics [7], according to which the wave function describes an ensemble of identically prepared states in identical experiments. It arguably does, but if the states are indeed identical, that's not much of an ensemble. Similarly, we do not mean Smolin's ensemble interpretation [8], where the ensemble considered is the ensemble of all the systems in the same quantum state in the Universe.

We mean instead that the supposedly identically prepared states in supposedly identical experiments are in fact different: they have different hidden variables and this is the reason why the measurement outcomes can be different for identical wave functions. We will refer to this interpretation of the wave function as the  $\psi$ -ensemble interpretation.

In our  $\psi$ -ensemble interpretation, we have an underlying theory with variables that we will collectively name  $\kappa$  and we will denote the state-space of all  $\kappa$  with  $K$ . The  $\kappa$  are generically time-dependent,  $\kappa(t)$ . These  $\kappa(t)$ s are not the same as Bell's hidden variables, as will become clear in a moment.

In the  $\psi$ -ensemble interpretation, the wave function in quantum mechanics emerges as an average description from the hidden variables theory, much like thermodynamics emerges from statistical mechanics. This interpretation suggests itself because of the apparent similarity of the von Neumann–Dirac equation with the Liouville equation, and also, as was pointed out in [9], because the naive  $\hbar \rightarrow 0$  limit of quantum mechanics results in a statistical theory.

We use the term 'hidden variables' because it has become common terminology. However, we want to stress that these variables are not *a priori* unmeasurable. They are 'hidden' merely in the sense that they do not appear in quantum mechanics. If the underlying theory was better understood, they could well become measurable one day. That is to say, the  $\psi$ -ensemble is an interpretation of the wave function in quantum mechanics, but it implies the existence of an underlying hidden variables theory. This underlying theory is better referred to as a completion or modification. This is as opposed to interpretations, such as the Transactional Interpretation [10,11] or Decoherent Histories [12], which have the same ontological basis as standard quantum mechanics, or Modal interpretations [13] that do not rely on an underlying completion (and do not violate SI—we will get to this point below).

We assume that the underlying theory is deterministic, that is, if one specifies an initial state, one can use the theory to uniquely calculate the outcome of a measurement.

$|\psi\rangle$  is, as usual, an element of a projected Hilbert space,  $\mathcal{H}$ . It fulfils the Schrödinger equation. We are considering a generic quantum mechanical measurement, in which  $|\psi\rangle$  is prepared at time  $t_p$  and measured at time  $t_m$ . The measurement is described by an orthonormal basis of pointer states that we will denote  $|I\rangle$ , where  $I \in \{0 \dots N - 1\}$  and  $N$  is the dimension of the Hilbert space.  $|I\rangle$  could be describing multiple different detectors, and might be a product-state or an entangled state when expressed in a basis of the individual detectors. The basis  $|I\rangle$  implicitly contains the measurement settings at the time of measurement.

For simplicity, we will in the following assume that the basis  $|I\rangle$  is time-independent. This does not mean that the detector setting cannot change. It merely means that if the detector setting before the measurement was different from the setting at the actual measurement, then it was not described by  $|I\rangle$ .

In quantum mechanics now, the Schrödinger evolution of  $|\psi\rangle$  will generically result in a state that is not one of the detector eigenstates at the time of measurement. In this case, we assume that  $|\psi(t_m)\rangle$  is an ensemble of different  $\kappa(t_m)$ , that is, a distribution  $\mu(\kappa(t))$  over  $K$ . This means in the  $\psi$ -ensemble theory we have a map

$$\mathcal{P}(K) \ni \mu(\kappa(t)) \rightarrow |\psi(t)\rangle \in \mathcal{H}, \quad (2.1)$$

where  $\mathcal{P}(K)$  is the space of all normalizable probability distributions on  $K$ . That is to say, we interpret  $|\psi(t_m)\rangle$  as an ensemble because we empirically know it corresponds to different measurement outcomes.

For each pointer state  $|I\rangle$ , there is a subset of hidden variables,  $\{\kappa\}_I$  that will lead to this outcome. We will refer to these subsets  $\{\kappa\}_I$  as ‘clusters’. The measurement reveals which cluster the hidden variable of the actual state belonged to. Hence, the measurement reveals some information about the hidden variable.

The reason we have for considering this interpretation is that the wave function update is non-local when interpreted as a physical process. This makes it difficult to combine quantum theory with general relativity. If we hence introduce a  $\psi$ -ensemble interpretation, our motivation is to develop an underlying theory which restores locality, causality and determinism. Of course, one can conceive of  $\psi$ -ensembles that are not local, but these (unsurprisingly) are not useful for restoring locality. Therefore, for the rest of this paper, we will only consider  $\psi$ -ensembles that are locally causal in Bell’s sense [14–16].

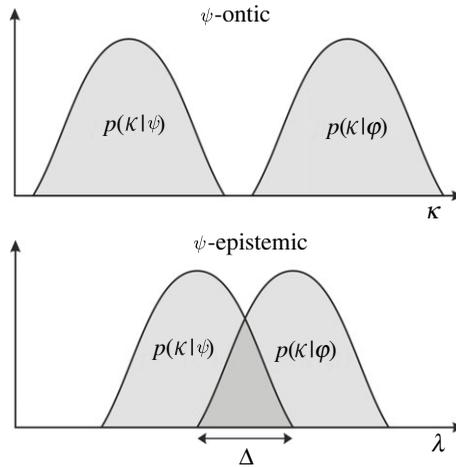
### 3. Difference between $\psi$ -ensemble and $\psi$ -epistemic

The introduction of the  $\psi$ -ensemble in the previous section sounds superficially very similar to the definition of a  $\psi$ -epistemic model introduced by Harrigan & Spekkens [5]. According to Harrigan and Spekkens (hereafter HS), a model is  $\psi$ -epistemic if one particular value of the hidden variable can correspond to more than one wave function.

As a simple example, consider you have a two-state system with detector eigenstates  $|0\rangle$  and  $|1\rangle$ . You could prepare the state so that under the Schrödinger-evolution it goes to  $|0\rangle$  with certainty. Or you could prepare it so that it goes to  $(|0\rangle + |1\rangle)/\sqrt{2}$ . If the wave function was  $\psi$ -epistemic, then the latter case must have some overlap with the former in the underlying hidden-variables theory because it can result in the same measurement outcome. At least that is the idea.

Formally, HS first define the probability  $p(\kappa|\psi)$  that the quantum wave function  $\psi$  is described by  $\kappa$ , where  $\kappa \in K$  [17,18]. Next, they define the probability of getting a particular measurement outcome given  $\kappa$ . While their framework can be used for any measurement, here we merely need the probability  $\mathcal{A}(\psi|\kappa)$  that the state  $\kappa$  is measured as  $\psi$ . From this, one obtains the support  $K_\psi \subset K$  for the wave function  $\psi$

$$\forall \kappa \in K_\psi, \mathcal{A}(\psi|\kappa) = 1. \quad (3.1)$$



**Figure 1.** Harrigan and Spekkens's  $\psi$ -epistemic and  $\psi$ -ontic models of reality [5]. Wave functions  $\psi$  and  $\varphi$  each have probability distributions over state-space. In a  $\psi$ -epistemic model, these can overlap over a subspace. A state  $\kappa$  within this overlap can then be represented both by  $\psi$  and  $\varphi$ . However, in  $\psi$ -ontic models, each state can be represented by at most one wave function. The amount of overlap is quantified by the parameter  $\Delta$ . For  $\psi$ -ontic models,  $\Delta = 0$ .

They then take a second wave function,  $\varphi$ , with its own probability distribution  $p(\kappa|\varphi)$ , and ask what is the probability that the  $\kappa$ s which contribute to  $\varphi$  also contribute to  $\psi$ . They quantify this overlap as

$$\Delta = \int_{K_\psi} p(\kappa|\varphi) d\kappa. \quad (3.2)$$

If  $\Delta > 0$ , HS interpret the wave functions in such a model as epistemic states, because the same underlying  $\kappa$  could contribute to both  $\psi$  and  $\varphi$ . Models with non-zero overlap  $\Delta$  are correspondingly called  $\psi$ -epistemic, whereas models without overlap are  $\psi$ -ontic (for illustration, see figure 1).

There are several problems with the HS  $\psi$ -ontic/epistemic distinction. One of them is purely linguistic. According to the HS definition, for  $\psi$ -ontic theories, the wave function in quantum mechanics itself is ontic. It seems odd to declare something ontic that can't be observed, especially given that Bohr already argued  $\psi$  is far better interpreted as knowledge. Settling a 100-year-old debate by definition is not insightful.

One might put this aside as historical nitpicking, but this linguistic issue reflects a deeper problem. If there's no observational requirement tied to calling something ontic, then every model can be made ontic just by declaring the wave function to be part of the hidden variables [4,19]. Such an easily malleable definition of 'ontic' is not what one wants to base theorems on.

It is also highly confusing that, according to the HS-definition, most hidden variables models which have been proposed so far are actually  $\psi$ -ontic. Think back to the example with the two state system that evolves into  $(|0\rangle + |1\rangle)/\sqrt{2}$  but is measured in either  $|0\rangle$  or  $|1\rangle$ . While the  $(|0\rangle + |1\rangle)/\sqrt{2}$  state must have been made up of different underlying states (because otherwise it could not give rise to different outcomes) there is no reason why any one of those underlying states must have been *the same* as those of a system prepared to evolve into  $|0\rangle$  or  $|1\rangle$ . Hence, according to the HS-definition, a  $\psi$ -ontic model can solve the measurement problem as well as a  $\psi$ -epistemic one.

We believe our definition of a  $\psi$ -ensemble to be more descriptive of hidden variables models, and less ambiguous than the  $\psi$ -epistemology proposed by HS. Whether or not the state prepared to evolve into  $(|0\rangle + |1\rangle)/\sqrt{2}$  had overlap with any other wave function, it will always be an ensemble if the underlying hidden variables theory solves the measurement problem. That is to say a  $\psi$ -ensemble could be either  $\psi$ -ontic or  $\psi$ -epistemic according to the HS-definition.

## 4. Hidden variables label trajectories

In our definition of a  $\psi$ -ensemble, we have been careful to point out that naturally the underlying hidden variables theory will be dynamic, hence the hidden variables will be functions of time. This is of utmost importance for locality considerations, as those usually pertain to arguments about the distribution of the hidden variables. (We will discuss this in §6.)

If we evaluate the distribution of hidden variables at the time of measurement, then we can infer the measurement outcome directly from that. However, for Bell's theorem, and also for the PBR-theorem, one normally considers the distribution of variables at the time of preparation instead. These two distributions will in general not be identical. The problem is, if one merely takes the distribution at the time of preparation, this will *not* constitute the hidden variables of Bell's theorem.

To see this, note that since we assume the underlying theory is deterministic, we can always take the distribution at the time of preparation,  $t_p$ , and evolve it forward with whatever is the transport function of the theory. Assuming time-translation invariance, let us denote the transport function as  $T(t_1 - t_0, \cdot)$  where the free slot is for the hidden variables. It has the property that

$$\kappa(t_1) = T(t_1 - t_0, \kappa(t_0)), T(0, \cdot) \equiv \text{Id}. \quad (4.1)$$

Now arguably the entire information that is necessary to determine the outcome at  $t_m$  is *both* the initial distribution at preparation  $\kappa(t_p)$  and the transport function  $T(t_m - t_p, \cdot)$ .

Bell, therefore, in the derivation of his inequality, correctly defines the hidden variables as the 'full specification' of the information necessary to predict the outcome [14]. In his own words, the 'values of these variables together with the state vector determine precisely the results of individual measurements' [20]. However, an initial state without an evolution law does not determine a final state.

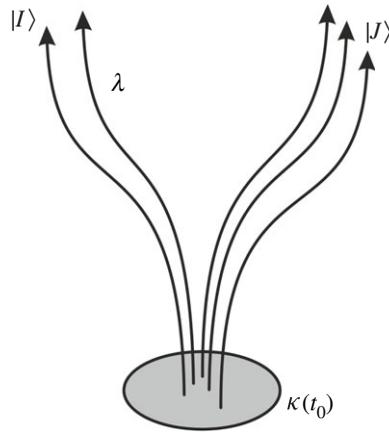
Therefore, if we call Bell's hidden variables (as usual)  $\lambda$ , then  $\lambda(t_0) = (\kappa(t_0), T(t_m - t_0, \kappa(t_0)))$  and one can now treat the hidden variables as time-independent. The additional information in the transport function becomes redundant in case one defines the distribution of the hidden variables at the time of measurement already ( $t_0 = t_m$ ) but normally one chooses  $t_p < t_0 < t_m$ .

Pictorially this means that Bell's hidden variables (the  $\lambda$ ) can be interpreted as labels for histories (in state space), whereas the HS hidden variables (that we called  $\kappa$ ) define a state (figure 2). For this reason, assuming that Bell's hidden variables belong to a particular moment in the history of the (entire) state makes no sense. We will comment in §6 on what this means for locality requirements that one normally uses for the distribution of the hidden variables.

A particularly illustrative example for hidden variables that define trajectories is the model proposed in [21,22]. In this case, the hidden variables simply are the detector eigenstates towards which the prepared state evolves. Given that we cannot measure the state before we measure the state, this is the most minimal assumption that one can make.

Following the terminology introduced in [15], we note that a deterministic model may not be predictable, in the sense that a model may contain variables from which one *could* calculate measurement outcomes when one had them, and yet one cannot find out the values of the hidden variables to actually do that. Bohmian mechanics is an example of such a deterministic, yet unpredictable mode. The  $\psi$ -ensembles we consider here are all deterministic. They may or may not also be predictable. In the model laid out in [21,22], the hidden variables are incalculable which also makes the model unpredictable.

A few words are in order here about the sense in which the  $\psi$ -ensemble is local. In the  $\psi$ -ensemble, the hidden variables are *in general* not localized in the sense that they do not belong to some compact region of space because they actually describe trajectories. The initial values  $\kappa(t_0)$  are in general not localized either because they are not uniquely defined. Suppose, for example, you have a distribution of variables in space and instead replace them with their Fourier-transformation. That would contain the same information but one may be localized, the other not. However, the initial states are localized if one considers experiments in which the prepared state is localized and is detected in localized detectors, which is for all practical purposes always the



**Figure 2.** Sketch to illustrate the relation between the hidden variables  $\kappa(t)$  which define the state, and the hidden variables  $\lambda$  which define the entire trajectory.  $|I\rangle$  and  $|J\rangle$  depict two different detector eigenstates. The trajectories that go to the eigenstate  $|I\rangle$  belong to cluster  $\{\kappa\}_I$ , and those which go to  $|J\rangle$  belong to cluster  $\{\kappa\}_J$ , respectively.

case, and certainly in Bell-type experiments. It is for this reason that the  $\psi$ -ensembles can fulfil the definition of local causality, as Bell intended.

We should add here that it is quite possibly the case that the type of model we consider, in which the hidden variables actually describe trajectories, is not what Bell meant when he conceived of his definition of local causality. But regardless of what Bell may have meant, such models can formally fulfil the condition of local causality.

## 5. The PBR theorem revisited

Let us then look at what happens when one confuses these two sets of hidden variables. The PBR-theorem [3], in a nutshell, shows that models which are  $\psi$ -epistemic according to the HS-definition are incompatible with observations. This superficially sounds similar to Bell's theorem which takes on hidden variables theories, but the PBR-theorem works entirely differently. For this reason, one cannot draw conclusions from it about  $\psi$ -ensemble theories. It is not our intention here to criticize the PBR theorem which has instilled much fruitful discussion and helped sharpen terminology. We merely want to clarify what conclusions can be drawn from it.

The above discussion on  $\psi$ -ensembles is relevant as the hidden variables in the HS-definition merely define the state. They do not also specify the time-evolution of the underlying hidden variable system. If one hence does not define the state at the time of measurement, then the hidden variables in the HS-definition do not determine the measurement outcome. How could they if they do not contain information about the evolution law? Now one may say that a definition is just that, a definition. However, for the PBR-theorem, one treats the hidden variables *as if* they did determine the measurement outcome, which implicitly assumes that the transport function contains no further information.

If however the transport function contains necessary information to determine the outcome, then one runs into the following problem. For the PBR-theorem, one considers the distribution of hidden variables at preparation in the HS-definition. However, those do not tell us what the outcome is, so no conclusions about whether the theory is viable or not can be derived. If one instead replaces the variables in the PBR-theorem with the variables which actually determine the outcome—i.e. those which contain information about the transport function—then one needs an additional assumption. This assumption is that the complete variables are not correlated with the detector settings. That is, one needs the assumption that Bell called SI. This is because if SI

is not fulfilled, then one of the assumptions for the PBR-theorem, the preparation independence postulate (PIP), is not fulfilled.

Bell defined SI as the absence of correlations between the hidden variables and the two detector settings which he considered in a particular experiment. However, of course, one can define SI in general as

$$\mu(\lambda|S) = \mu(\lambda), \quad (5.1)$$

where  $S$  are the detector settings (of whatever experiment), and  $\mu$  is the probability distribution of the hidden variables. As laid out in [23], violating this condition does not necessarily require correlations between the settings and the distribution of the hidden variables over the state space; the correlations can instead be induced by the structure of the state space.

SI sneaks into the PBR-theorem because the HS-definition for a  $\psi$ -epistemic model assumes that the hidden variables which define the state at preparation have nothing to do with the measurement. And that sounds reasonable, until one realizes that if one wants to say anything about measurement outcomes one needs to know the time-evolution of the hidden variables too.

The PIP now says that if we prepare two independent (spatially separate) experiments, then the distributions of the hidden variables should factorize. In the PBR argument, one first considers two experiments in isolation. Then one combines the two so-prepared states—using PIP—and measures them instead with a smartly chosen basis of distinguishable states. From this, one derives a contradiction: If the distribution of the hidden variables was actually the product of the distributions for the separate measurements, then one cannot reproduce the predictions of quantum mechanics.

Needless to say, the theorem is correct where the mathematics is concerned, but the relevant physical assumption was that the distributions of the hidden variables for the individual experiments already did not depend on the measurement settings. So of course if one multiplies them, they still do not depend on the measurement settings and one cannot reproduce the predictions of quantum mechanics. That is, in the  $\psi$ -epistemic framework, we have  $SI \Rightarrow PIP$ .<sup>1</sup>

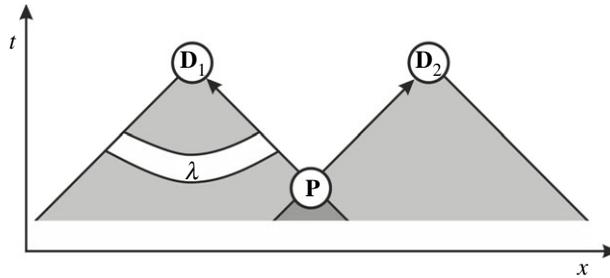
However, the relevant assumption was the independence of the complete hidden variables (the  $\lambda$ s which determine the outcome) from the measurement settings, not the factorization. In particular, the initial distribution of the variables,  $\kappa(t_p)$ , might well factorize at preparation.

An example for a model in which this happens is [24]. This model uses one (complex valued) hidden variable for each dimension of the Hilbert space. These hidden variables, which correspond to  $\kappa$ , are uniformly distributed in the unit disc. Hence, the model fulfils the requirements for the PBR-theorem and yet reproduces the predictions of quantum mechanics, seemingly defying the theorem. The reason is that the dynamical law of the model depends on the detector settings, and so do the full hidden variables  $\lambda$ . When one uses these variables, the model violates PIP, hence no contradiction with the theorem arises. However, since the variables then no longer describe the initial state at preparation, it is unclear why they should fulfil PIP in the first place.

We may note that this issue does not appear in Bell's theorem, simply because Bell considers a situation where the state is prepared at one place but measured in two, whereas PBR considers a situation where the state is prepared in two places but measured in one. In Bell's case, one has no reason to assume that the distribution of the hidden variables factorizes.

For completeness, let us mention that the approach presented in [25] to rule out (HS)  $\psi$ -epistemic models works by constructing sets of incompatible states and deriving properties about the overlap of probability distributions from measurements on those. This proof implicitly assumes that the probability distribution of the hidden variables does not depend on the measurement settings.

<sup>1</sup>Note that it is not in general the case that  $\neg PIP \Rightarrow \neg SI$  because it could be that other assumptions of the  $\psi$ -epistemic framework are violated rather than SI.



**Figure 3.** A diagram showing the two backwards light cones from detectors  $D_1$  and  $D_2$ , and their overlap at point of preparation  $P$ .

## 6. Properties of $\psi$ -ensemble theories

### (a) Statistical independence

We know from Bell's theorem that any  $\psi$ -ensemble theory which is locally causal (in Bell's sense) and reproduces the predictions of quantum mechanics must violate SI. Indeed, that one can reproduce quantum mechanical correlations while respecting local causality provided SI is violated is quite possibly the reason why Bell (and others after him) sought other ways to rule out SI, notably by arguing that it requires fine-tuning (discussed below).

However, one does not need Bell's theorem to see that local and causal  $\psi$ -ensemble models will violate SI. We just need to note that if we want states in the hidden variables theory to evolve locally and within the light-cone into detector eigenstates, then we need information about the detector before the prepared state reaches the detector—otherwise how do we know what state to evolve into?

For the purposes of this paper, we shall say that a hidden variables theory solves the measurement problem if it respects Bell's local causality and gives rise to detector eigenstates in each single run of an experiment with a probability distribution that agrees (to within measurement accuracy) with that of quantum mechanics. There are other ways to look at the measurement problem [26], but for the purposes of this present paper, this simplified notion that specifically applies to hidden variables theories will suffice.

To illustrate what solutions to the measurement problem have to do with SI, consider the generic setting of an experiment with two detectors (figure 3). The state is prepared at  $P$  and measured at detectors  $D_1$  and  $D_2$ . One may have in mind for example a single photon that goes through a beam splitter. The state does not have to be entangled. We want to know how it evolves in the underlying hidden variables theory.

When the prepared state arrives at  $D_1$  it needs to know the setting at  $D_2$ . Yet, it is clear that if that information wasn't available already at  $P$ , then it can't have gotten there within the light cone. Hence, both detector settings  $D_1$  and  $D_2$  must have been either in the initial distribution of hidden variables or the evolution law. In either case, SI is violated.

To be sure, we have drawn an extreme case where the prepared state actually moves with the speed of light. If one took a massive particle, the last moment in which information about  $D_2$  must be available for the underlying state going to  $D_1$  could be after preparation. But that does not change the fact that SI must have been violated because information about the detector setting must have been in the full specification of the hidden variables.

It is a corollary of this—and not the starting point—that any  $\psi$ -ensemble theory that solves the measurement problem and reproduces quantum mechanics will violate the assumptions of Bell's theorem and the PBR-theorem. It follows that experiments which show the violation of a Bell inequality or the CHSH inequality [27], or the agreement of the measurement proposed by PBR with the predictions of quantum mechanics [28], cannot rule out  $\psi$ -ensemble theories. A different type of experiment is necessary to test such theories.

Understanding that SI is violated *because* that is required to solve the measurement problem also explains why one need not expect such violations to occur in non-quantum experiments, like the often quoted tobacco trials to infer a correlation between smoking and cancer (see [29] and references therein). No preparation of such an experiment will ever result in a superposition of measurement outcomes. Even if one used a quantum experiment to, say, sample mice into two groups (one exposed to tobacco smoke, the other not), then the quantum superposition would be destroyed before one even wrote down the outcome of the quantum experiment. But SI is only violated between the preparation of an experiment that, under the Schrödinger-evolution, evolves into a superposition of detector eigenstates, and the measurement of that state.

It must be understood here that we use the term ‘measurement’ in the same sense as in quantum mechanics. That is, it does not necessarily require an observer or even a detector, it just requires a sufficient amount of interactions with the environment. We do not need to be specific about this because this aspect of the measurement problem was already solved by decoherence and einselection [30]. A detector is any system that decoheres the prepared state quickly enough, and the detector eigenstates are defined as those which are stable in that process. We merely need violations of SI to explain why we always measure a detector eigenstate and not, as decoherence predicts, a mixed state.

It is worth stressing that the statements made in this section are not suggested interpretations of mathematical expressions, they are conclusions drawn from empirical facts. We know, as a matter of fact, that measurements yield definite results. If the results come about locally, causally, and deterministically from the prepared state, then the underlying theory must violate SI. We also know empirically that we never observe superpositions of detector eigenstates, hence violations of SI are negligible for macroscopic objects.

Theories which violate statistical independence are often further broken down into those which are superdeterministic and those that are retrocausal. The superdeterministic ones were recently further broken down into those that are and are not supermeasured [23]. However, the terms superdeterministic and retrocausal have been used to mean different things by different authors which has made the situation very confusing. For example, some authors assume superdeterminism is deterministic [29], while others do not [31]. Some equate retrocausality with future input dependence [32], others do not [33,34]. Some claim that retrocausality is just mislabelled superdeterminism [35]. Yet other authors have proposed to use the terms all-at-once or atemporal [36]. None seem to assign much importance to the ensemble that the wave function should describe, which is at the centre of our definition.

We believe that the  $\psi$ -ensemble is a more useful classification because it captures those approaches which are candidates to solve the measurement problem(s).

## (b) Locality

It is quite common to get confused here about how such a requirement—that the evolution law depends on the measurement settings at the time of measurement—can possibly be local. However, keep in mind that the hidden variables  $\lambda$  label trajectories, not initial states, of the underlying theory. The trajectories necessarily connect the prepared state with the detectors locally and causally (if they did not, we could not detect the prepared state). Hence, the requirement that the hidden variables are correlated with the detector settings at measurement can come about entirely by local propagation and by local interactions. (Again, we do not claim that such a theory necessarily has to be local. It is just that this is the case we are interested in.)

Such  $\psi$ -ensemble theories that violate SI fulfil Bell’s criterion of local causality because if the information about both detector settings was available already at preparation, then it is unnecessary to later add it.

### (c) Fine-tuning

It is frequently argued that theories which violate SI would need a lot of detail to specify the initial state for the detector setting at the time of measurement, and that they are thence fine-tuned ‘conspiracies’ [37]. This initial state is often believed to be required long before the experiment even began, possibly as far back as the Big Bang.

However, the idea that we need to define an initial state for the detector setting is patently absurd. In quantum mechanics, we also never define the initial condition at the Big Bang that will give rise to a specific detector setting. Indeed, we never do this in any theory. It is in practice unnecessary to ever write down this initial state, regardless of whether it is feasible in the first place. We simply make predictions for any possible measurement setting. Hence, violating SI does not require a lot of detail; it merely requires the detector settings at the time of measurement—an input we also use in ordinary quantum mechanics.

A second fine-tuning argument [38] has it that it would require specifying a lot of details for the distribution of hidden variables in order to prevent violations of Born’s rule that would allow for superluminal signalling. However, leaving aside that the absolute impossibility of superluminal signalling (rather than its observed rarity) is a postulate, not an empirical fact, we can arguably always achieve that Born’s rule is fulfilled simply by postulating that it is fulfilled, which one also does in quantum mechanics.

That is, these two arguments use a double-standard: assumptions are not called fine-tuned when used in normal quantum mechanics (detector settings at measurement and Born’s rule are required to calculate probabilities) but are called fine-tuned when used in other frameworks. As already argued in [39], such a notion of fine-tuning is scientifically meaningless and can at best be considered metaphysical. (Fine-tuning in these contexts has been discussed in more detail by Adlam [40].)

## 7. Using the $\psi$ -ensemble to make sense of quantum experiments

Of course using a  $\psi$ -ensemble, even if deterministic, local, and causal does not remove the weirdness of quantum mechanics. It remains in the property that for a spatially distributed detector, a state in the underlying theory that goes to one particular eigenstate must have had information about that part of the detector to which it *did not* go. Again, this requirement can be formulated purely in local terms (if the state would have gone to the other part of the detector, it would have interacted with it), but this is where the famed quantum weirdness goes. If the underlying hidden variables theory is local and causal and deterministic, then the evolution which the actual state takes—and which we ultimately observe—takes into account what would have happened if the state had gone elsewhere. We will now explain how this nicely explains some otherwise very unintuitive quantum behaviour.

### (a) The double-slit experiment

To warm up, we consider the familiar double-slit experiment. The odd thing about this experiment is that the behaviour of the quantum particle changes depending on whether we know which slit the particle goes through. If we do not know, the particle seems to go through both slits and interferes with itself, like we expect for a wave. But if we measure which slit the particle goes through, the interference pattern vanishes.

This seemingly strange behaviour makes total sense from the  $\psi$ -ensemble perspective because the trajectory of the state in the underlying hidden variables theory depends on what is being measured.

### (b) Wheeler’s delayed-choice experiment

The delayed-choice experiment [41] comes in a large number of variations, but in the simplest version, one changes the variable that is being measured after the prepared state has begun its

propagation. In normal quantum mechanics, this is even more difficult to understand than the double slit experiment because it seems that the prepared state must somehow ‘change its mind’ as the setting is changed.

However, from the  $\psi$ -ensemble perspective this is clearly not what happened. This is because the trajectories depend on the detector setting *at the time of measurement*. What the settings were before that or how often they were changed is irrelevant.

### (c) The Elitzur–Vaidman bomb detector

The Elitzur–Vaidman bomb detector allows one to detect whether a bomb is live (would blow up when it detects a photon) without triggering the bomb [42]. In this experiment, the bomb acts as a detector. Whether the bomb is live or a dud therefore constitutes two different detector settings. From the  $\psi$ -ensemble perspective, it is hence rather unsurprising that the case in which the bomb does not blow up contains information about whether the bomb is live, because this is just a type of detector setting.

### (d) Counterfactual computation/communication

In counterfactual computation [43], one infers the result of a computer query without interacting with the computer. This experiment is a variant of the Elitzur–Vaidman experiment and can be explained the same way. The computer acts as a detector depending on what the outcome of the calculation is. In the  $\psi$ -ensemble interpretation, the particle which one measures contains information about the setting of the detectors which it was not detected in, which allows one to infer the result of the computation.

The same logic also applies to Salih *et al.*'s counterfactual communication device and related protocols [44–46], and so protocols for imaging [47] and for quantum information transfer derived from it [48,49].

### (e) The extended Wigner’s Friend scenario

The previous examples all work more or less the same way. The extended Wigner’s Friend experiment [50] is a more recent proposal that beautifully highlights the problems one runs into when rejecting hidden variables theories, but works entirely differently.

The extended Wigner’s Friend scenario is a Bell-type test with two entangled particles, one measured by Charlie and one by Debbie (the friends) in spatially separated laboratories. Each of the two observers is observed by a super-observer, Alice and Bob, respectively. Charlie and Debbie measure an entangled state and Alice and Bob then further measure correlations between Alice’s and Bob’s measurement outcomes. It turns out that, given suitable measurement settings, the observers cannot all agree on what the measurement results are.

The issue with this experiment is a fuzzy definition of what constitutes a measurement. As already pointed out in [51,52], one can interpret the friends’ measurements as either having resulted in a definite outcome, or as still being in an entangled quantum state. The resulting contradiction does not occur if one has a theory in which a measurement outcome is an unambiguous result of the time-evolution.

In a  $\psi$ -ensemble theory, the friends either make a measurement and violations of SI are for all practical purposes destroyed. Then the super-observers can no longer measure quantum correlations between the friends. Or the friends really just make a transformation on a quantum state, rather than a measurement, in which case the super-observers can well observe further quantum effects.

## 8. Conclusion

We have argued here that the  $\psi$ -ontic/epistemic distinction is not descriptive of hidden variables models that can restore locality and causality by giving a physical meaning to the wave function

update. Instead, they can better be described as  $\psi$ -ensemble theories. If they are local, causal, deterministic and reproduce the predictions of quantum mechanics, then such theories violate the assumptions of both Bell's theorem and the PBR-theorem. We have shown that taking this possibility seriously can help making sense of some otherwise strange quantum phenomena.

**Data accessibility.** This article has no additional data.

**Authors' contributions.** J.R.H.: conceptualization, formal analysis, funding acquisition, investigation, methodology, project administration, visualization, writing—original draft, writing—review and editing; S.H.: conceptualization, formal analysis, funding acquisition, investigation, methodology, project administration, visualization, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

**Conflict of interest declaration.** We declare we have no competing interest.

**Funding.** S.H. acknowledges support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under grant no. HO 2601/8-1. J.R.H. is supported by the University of York's EPSRC DTP grant no. EP/R513386/1, and the EPSRC Quantum Communications Hub (funded by the EPSRC grant no. EP/M013472/1).

**Acknowledgements.** We thank James Ladyman, Sophie Inman, Tim Palmer and John Rarity for useful discussions.

## References

- Bell JS. 1964 On the Einstein Podolsky Rosen paradox. *Phys. Phys. Fizika* **1**, 195–200. (doi:10.1103/PhysicsPhysiqueFizika.1.195)
- Clauser JF, Horne MA, Shimony A, Holt RA. 1969 Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.* **23**, 880–884. (doi:10.1103/PhysRevLett.23.880)
- Pusey MF, Barrett J, Rudolph T. 2012 On the reality of the quantum state. *Nat. Phys.* **8**, 475–478. (doi:10.1038/NPHYS2309)
- Leifer MS. 2014 Is the quantum state real? An extended review of  $\psi$ -ontology theorems. *Quanta* **3**, 67. (doi:10.12743/quanta.v3i1.22)
- Harrigan N, Spekkens RW. 2010 Einstein, incompleteness, and the epistemic view of quantum states. *Found. Phys.* **40**, 125–157. (doi:10.1007/s10701-009-9347-0)
- Hance JR, Rarity J, Ladyman J. 2022 Could wavefunctions simultaneously represent knowledge and reality? *Quantum Stud.: Math. Found.* (doi:10.1007/s40509-022-00271-3)
- Ballentine LE. 1970 The statistical interpretation of quantum mechanics. *Rev. Mod. Phys.* **42**, 358–381. (doi:10.1103/RevModPhys.42.358)
- Smolin L. 2012 A real ensemble interpretation of quantum mechanics. *Found. Phys.* **42**, 1239–1261. (doi:10.1007/s10701-012-9666-4)
- Klein U. 2012 What is the limit  $\hbar \rightarrow 0$  of quantum theory?. *Am. J. Phys.* **80**, 1009–1016. (doi:10.1119/1.4751274)
- Cramer JG. 1986 The transactional interpretation of quantum mechanics. *Rev. Mod. Phys.* **58**, 647–687. (doi:10.1103/RevModPhys.58.647)
- Kastner RE. 2013 *The transactional interpretation of quantum mechanics: the reality of possibility*. Cambridge, UK: Cambridge University Press.
- Gell-Mann M, Hartle JB. 1996 Quantum mechanics in the light of quantum cosmology. In *Foundations Of Quantum Mechanics In The Light Of New Technology: Selected Papers from the Proc. of the First through Fourth Int. Symp. on Foundations of Quantum Mechanics*, pp. 347–369. World Scientific. (doi:10.1142/9789812819895\_0036)
- Lombardi O, Dieks D. 2021 Modal interpretations of quantum mechanics. In *The Stanford Encyclopedia of Philosophy* (ed. EN Zalta). Metaphysics Research Lab, Stanford University, winter 2021 edition. URL: <https://plato.stanford.edu/archives/win2021/entries/qm-modal/>.
- Bell JS. 2004 *Speakable and unspeakable in quantum mechanics: Collected papers on quantum philosophy*, 2nd edn. Cambridge, UK: Cambridge University Press.
- Cavalcanti EG, Wiseman HM. 2012 Bell nonlocality, signal locality and unpredictability (or what Bohr could have told Einstein at Solvay had he known about Bell experiments). *Found. Phys.* **42**, 1329–1338. (doi:10.1007/s10701-012-9669-1)
- Norsen T. 2011 John S. Bell's concept of local causality. *Am. J. Phys.* **79**, 1261–1275. (doi:10.1119/1.3630940)

17. Harrigan N, Rudolph T, Aaronson S. 2007 Representing probabilistic data via ontological models. (<http://arxiv.org/abs/0709.1149>)
18. Maroney OJE. 2012 How statistical are quantum states? (<http://arxiv.org/abs/1207.6906>). URL: <https://arxiv.org/pdf/1207.6906.pdf>.
19. Schlosshauer M, Fine A. 2012 Implications of the Pusey-Barrett-Rudolph quantum no-go theorem. *Phys. Rev. Lett.* **108**, 260404. (doi:10.1103/PhysRevLett.108.260404)
20. Bell JS. 1966 On the problem of hidden variables in quantum mechanics. *Rev. Mod. Phys.* **38**, 447–452. (doi:10.1103/RevModPhys.38.447)
21. Palmer TN. 2009 The invariant set postulate: a new geometric framework for the foundations of quantum theory and the role played by gravity. *Proc. R. Soc. A* **465**, 3165–3185. (doi:10.1098/rspa.2009.0080)
22. Palmer TN. 2020 Discretization of the Bloch sphere, fractal invariant sets and Bell's theorem. *Proc. R. Soc. A* **476**, 20190350. (doi:10.1098/rspa.2019.0350)
23. Hance JR, Hossenfelder S, Palmer TN. 2021 Supermeasured: violating statistical independence without violating statistical independence. (<http://arxiv.org/abs/2108.07292>)
24. Donadi S, Hossenfelder S. 2020 A superdeterministic toy model. (<http://arxiv.org/abs/2010.01327>)
25. Barrett J, Cavalcanti EG, Lal R, Maroney OJE. 2014 No  $\psi$ -epistemic model can fully explain the indistinguishability of quantum states. *Phys. Rev. Lett.* **112**, 250403. (doi:10.1103/PhysRevLett.112.250403)
26. Maudlin T. 1995 Three measurement problems. *Topoi* **14**, 7–15. (doi:10.1007/BF00763473)
27. Aspect A, Grangier P, Roger G. 1981 Experimental tests of realistic local theories via Bell's theorem. *Phys. Rev. Lett.* **47**, 460–463. (doi:10.1103/PhysRevLett.47.460)
28. Ringbauer M, Duffus B, Branciard C, Cavalcanti EG, White AG, Fedrizzi A. 2015 Measurements on the reality of the wavefunction. *Nat. Phys.* **11**, 249–254. (doi:10.1038/nphys3233)
29. Hossenfelder S, Palmer T. 2020 Rethinking superdeterminism. *Front. Phys.* **8**, 139. (doi:10.3389/fphy.2020.00139)
30. Zurek WH. 2003 Decoherence, einselection, and the quantum origins of the classical. *Rev. Mod. Phys.* **75**, 715–775. (doi:10.1103/RevModPhys.75.715)
31. Sen I, Valentini A. 2020 Superdeterministic hidden-variables models I: non-equilibrium and signalling. *Proc. R. Soc. A* **476**, 20200212. (doi:10.1098/rspa.2020.0212)
32. Wharton KB, Argaman N. 2020 Colloquium: Bell's theorem and locally mediated reformulations of quantum mechanics. *Rev. Mod. Phys.* **92**, 021002. (doi:10.1103/RevModPhys.92.021002)
33. Price H. 1994 A neglected route to realism about quantum mechanics. *Mind* **103**, 303–336. (doi:10.1093/mind/103.411.303)
34. Price H. 2012 Does time-symmetry imply retrocausality? How the quantum world says 'maybe'? *Stud. Hist. Phil. Sci. B* **43**, 75–83. (doi:10.1016/j.shpsb.2011.12.003)
35. Nikolaev V, Vervoort L. 2022 Aspects of superdeterminism made intuitive. (<http://arxiv.org/abs/2205.10616>)
36. Adlam E. 2022 Two roads to retrocausality. (<http://arxiv.org/abs/2201.12934>)
37. Sen I, Valentini A. 2020 Superdeterministic hidden-variables models II: conspiracy. *Proc. R. Soc. A* **476**, 20200214. (doi:10.1098/rspa.2020.0214)
38. Wood CJ, Spekkens RW. 2015 The lesson of causal discovery algorithms for quantum correlations: causal explanations of bell-inequality violations require fine-tuning. *New J. Phys.* **17**, 033002. (doi:10.1088/1367-2630/17/3/033002)
39. Hossenfelder S. 2020 Superdeterminism: a guide for the perplexed. (<http://arxiv.org/abs/2010.01324>)
40. Adlam E. 2021 Contextuality, fine-tuning and teleological explanation. *Found. Phys.* **51**, 106. (doi:10.1007/s10701-021-00516-y)
41. Wheeler JA. 1978 The 'past' and the 'delayed-choice' double-slit experiment. In *Mathematical foundations of quantum theory*, pp. 9–48. Elsevier. (<http://dx.doi.org/doi:10.1016/B978-0-12-473250-6.50006-6>)
42. Elitzur AC, Vaidman L. 1993 Quantum mechanical interaction-free measurements. *Found. Phys.* **23**, 987–997. (doi:10.1007/BF00736012)
43. Hosten O, Rakher MT, Barreiro JT, Peters NA, Kwiat PG. 2006 Counterfactual quantum computation through quantum interrogation. *Nature* **439**, 949–952. (doi:10.1038/nature04523)

44. Salih H, Li Z-H, Al-Amri M, Zubairy MS. 2013 Protocol for direct counterfactual quantum communication. *Phys. Rev. Lett.* **110**, 170502. (doi:10.1103/PhysRevLett.110.170502)
45. Salih H, McCutcheon W, Hance JR, Rarity J. 2022 The laws of physics do not prohibit counterfactual communication. *npj Quantum Inf.* **8**, 60. (doi:10.1038/s41534-022-00564-w)
46. Hance JR, Ladyman J, Rarity J. 2021 How quantum is quantum counterfactual communication? *Found. Phys.* **51**, 12. (doi:10.1007/s10701-021-00412-5)
47. Hance JR, Rarity J. 2021 Counterfactual ghost imaging. *npj Quantum Inf.* **7**, 1–7. (doi:10.1038/s41534-021-00411-4)
48. Salih H, Hance JR, McCutcheon W, Rudolph T, Rarity J. 2020 Deterministic teleportation and universal computation without particle exchange. (<http://arxiv.org/abs/2009.05564>)
49. Salih H, Hance JR, McCutcheon W, Rudolph T, Rarity J. 2021 Exchange-free computation on an unknown qubit at a distance. *New J. Phys.* **23**, 013004. (doi:10.1088/1367-2630/abd3c4)
50. Frauchiger D, Renner R. 2018 Quantum theory cannot consistently describe the use of itself. *Nat. Commun.* **9**, 3711. (doi:10.1038/s41467-018-05739-8)
51. Relaño A. 2018 Decoherence allows quantum theory to describe the use of itself. (<http://arxiv.org/abs/1810.07065>)
52. Żukowski M, Markiewicz M. 2021 Physics and metaphysics of Wigner's friends: even performed premeasurements have no results. *Phys. Rev. Lett.* **126**, 130402. (doi:10.1103/PhysRevLett.126.130402)