http://www.swansea.ac.uk/library/researchsupport/ris-support/

The Emergent Copenhagen Interpretation of Quantum Mechanics

Timothy J. Hollowood[*]
Department of Physics, Swansea University,
Swansea, SA2 8PP, UK
(Dated: March 3, 2015)

We introduce a new and conceptually simple interpretation of quantum mechanics based on reduced density matrices of sub-systems from which the standard Copenhagen interpretation emerges as an effective description of macroscopically large systems. This interpretation describes a world in which definite measurement results are obtained with probabilities that reproduce the Born rule. Wave function collapse is seen to be a useful but fundamentally unnecessary piece of prudent book keeping which is only valid for macro-systems. The new interpretation lies in a class of modal interpretations in that it applies to quantum systems that interact with a much larger environment. However, we show that it does not suffer from the problems that have plagued similar modal interpretations like macroscopic superpositions and rapid flipping between macroscopically distinct states. We describe how the interpretation fits neatly together with fully quantum formulations of statistical mechanics and that a measurement process can be viewed as a process of ergodicity breaking analogous to a phase transition. The key feature of the new interpretation is that joint probabilities for the ergodic subsets of states of disjoint macro-systems only arise as emergent quantities. Finally we give an account of the EPR-Bohm thought experiment and show that the interpretation implies the violation of the Bell inequality characteristic of quantum mechanics but in a way that is rather novel. The final conclusion is that the Copenhagen interpretation gives a completely satisfactory phenomenology of macro-systems interacting with micro-systems.

I. INTRODUCTION

The central mystery of quantum mechanics is present even in the simplest measurement on a qubit. Solving the Schrödinger equation for a suitable Hamiltonian gives an evolution of the form

$$
\begin{aligned}
\big(c_+|z^+\rangle + c_-|z^-\rangle\big) &\otimes |A_0\rangle \\
&\longrightarrow c_+|z^+\rangle \otimes |A_+\rangle + c_-|z^-\rangle \otimes |A_-\rangle ,
\end{aligned}
\tag{1}
$$

where the two states $|A_\pm\rangle$ are macroscopically distinct states of the measuring device. But how can this be consistent with the fact that when any experiment of this type is performed a definite outcome occurs either $|A_+\rangle$ or $|A_-\rangle$? The Copenhagen interpretation[1] solves the problem by "collapsing the wave function," i.e. choosing only one of the distinct terms on the right-hand side with probabilities $|c_\pm|^2$, respectively, on the grounds that the measuring device is macroscopic. The universal success of applying the rule disguises the fact that it is really only a rule of thumb: when is a system sufficiently macroscopic that it qualifies for collapse? This arbitrary separation of systems into microscopic and macroscopic is the famous Heisenberg cut.

We can measure how macroscopically distinct the two states of the measuring device $|A_\pm\rangle$ are by estimating their inner product. Let us suppose the measuring device has a macroscopically large number of microscopic

degrees-of-freedom N. States are macroscopically distinct if all the microscopic degrees-of-freedom are separated by a macroscopic scale L. If ℓ is a characteristic microscopic length scale in the system and assuming, say, Gaussian wave functions for the microscopical degrees-of-freedom spread over the scale ℓ, the matrix elements between macroscopically distinct states is roughly

$$
\Delta \equiv \big|\langle A_+|A_-\rangle\big| \sim \exp\big[-NL^2/\ell^2\big] .
\tag{2}
$$

Just to get a feel for the numbers, suppose $N \sim 10^{20}$, $\ell \sim 10^{-10}\,\mathrm{m}$ (atomic size) and $L \sim 10^{-4}\,\mathrm{m}$, giving

$$
\boxed{\Delta \sim e^{-10^{32}}}
$$

In the following we will use Δ to denote a generic scale characterising the inner products of macroscopically distinct states. The estimate above is intended as a guide and the fact that this generic scale is so small will play an important role in this work.

There is an interesting analogue of the measurement problem in classical statistical mechanics. Consider an Ising ferromagnet above its Curie temperature. In a typical configuration, the spins point randomly up or down and there is no net magnetization. In the standard interpretation of classical statistical mechanics, the ensemble average captures a time average of the dynamics of the underlying microscopic state of the system. This is a statement of ergodicity: over time, interactions ensure that the microscopic state explores all the available states with a probability given by the usual Boltzmann factor. Now, if the temperature is reduced below the Curie temperature, there is a phase transition and the magnet finds itself in an ordered state where the spins tend to line up in the same direction, either up or down, and the system

---

[1] The Copenhagen interpretation is not really a completely settled set of ideas. We are using the term to stand for the way that most working physicists successfully use quantum mechanics in practice without having to even think about foundational issues.

develops a net magnetization. At a microscopic level, ergodicity is broken and for a given initial micro-state the system effectively only explores half of the overall ensemble over time and it is each of these sub-ensembles that has a net magnetization.

This picture of ergodicity breaking provides a possible template for solving the quantum measurement problem. The key feature that allows for its breaking in classical statistical mechanics is that the state of the system has two dual aspects: the ensemble and the micro-state. Of course, the fact that we deal with ensembles is because we, as macroscopic observers, are ignorant as to the exact micro-state at any particular time. When the system is in equilibrium, microscopic interactions in the system ensure that the micro-state explores the set of available states ergodically and a suitably coarse-grained time average is equal to the ensemble average. However, when the magnet undergoes the phase transition to the low temperature phase, the system becomes highly sensitive to which of the sub-ensembles the micro-state lies in. After the transition, ergodicity is broken and the micro-state only explores a sub-ensemble over time and so the time average is captured by the sub-ensemble. What is crucial for this mechanism is that underlying the ensemble is the existence of a real micro-state, even though this is hidden from the macroscopic point-of-view.

In order to carry this over into the quantum mechanical measurement problem, we need the analogues of the ensemble and the micro-state or what is called an "epistemic state" and an "ontic state".[2]

quantum mech.		statistical mech.
epistemic state	⟷	ensemble
ontic state	⟷	micro-state

If we had such a dual aspect, or different *modalities*, of the quantum state then we would have the possibility of applying an argument that involves the breaking of ergodicity to the measurement problem.

We do not expect that the macroscopic measuring device A will be in a pure quantum state, rather, it will be described by a reduced density matrix $\hat{\rho}_A$ as a consequence of its interaction with its environment. This density matrix is the epistemic state, the analogue of the ensemble in classical statistical mechanics. However, quantum mechanics in the standard formulation offers no analogue of an ontic state. In fact any attempt to provide an additional specification of the quantum state is tantamount to the introduction of hidden variables and, as shown by Bell [42], generally leads to predictions that are not consistent with standard quantum mechanics and experiments. So the conclusion is that classical statistic mechanics cannot provide a perfect analogue of what happens in a quantum measurement unless there is some

way to introduce addition information into the quantum state, the ontic information, but in a way that would be consistent with the usual predictions of quantum mechanics including the violation of Bell's theorem. A clue to how this could be done, is provided by a class of *modal interpretations* of quantum mechanics. In some of these interpretations, a system, being a sub-system of a larger system, is described by a reduced density matrix, the epistemic state. But as well as having a density matrix, the system also carries an ontic state in the form of one of the eigenvector of the reduced density matrix. These are the two modalities of the quantum state. Once we allow
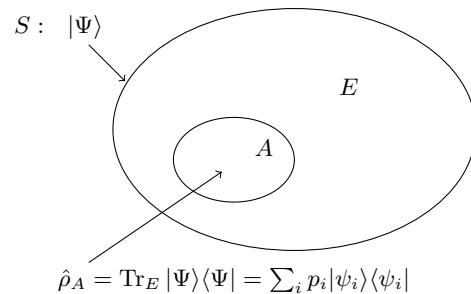

FIG. 1. A sub-system A of a large quantum system $S = A+E$. The quantum state of A is described by the reduced density matrix $\hat{\rho}_A$. In the new interpretation the ontic state of A at a given time corresponds to one of the eigenvectors of $\hat{\rho}_A$.

the system to have its own ontic state then this opens up the possibility to solve the measurement problem by using an ergodicity argument along the lines of classical statistical mechanics.[3] The question that we address in this paper is whether a complete and consistent interpretation of quantum mechanics can be constructed along these lines. Our claim—building on [1]—is that there exists a conceptually simple interpretation that can fill the ugly gap in the Copenhagen interpretation between unitary time evolution of states and the collapse of the wave function by having these features apply to the two different modes of the quantum state. At the same time it appear to evade the consequences of Bell's theorem and its generalizations to be fully consistent with the usual predictions of quantum mechanics. We call this new interpretation the *emergent Copenhagen interpretation* for reasons that will themselves emerge. It lies firmly in the class of modal interpretations [2–10][4] but is rather different in some fundamental ways that ensure it does not suffer from the problems that plague other modal interpretations.

---

[2] The latter were called "internal states" in [1] after [32, 33].

---

[3] The fact that ergodicity breaking should play some role in the measurement problem crystalized in discussions with Jacob Barandes.

[4] The book by Vermaas [11] gives an eloquent account of the class of modal interpretations, both their successes and problems, as well as having a full list of references.

The new interpretation, like other modal interpretations, works hand-in-hand with the phenomenon of decoherence in quantum systems [16–22] in order to describe a satisfactory phenomenology of macroscopic systems and is closely allied to the modern understanding of statistical mechanics as it arises from the quantum mechanics of systems interacting with large environments or baths [23–30]. Classical behaviour, the Born rule and wave function collapse are not put in by hand but are seen to emerge as an approximate phenomenology of systems with a macroscopic number of degrees-of-freedom. Heisenberg's cut is replaced by a continuous spectrum of "classicality". Finally, the new interpretation leads to a solution of the measurement problem via the mechanism of ergodicity breaking.

The paper is organised as follows. In section II we introduce the new interpretation by defining ontic states in section II A. In section II B, we discuss disjoint sub-systems and how their ontic states can relate to the ontic states of their union. In section II C, we discuss sub-systems in general and the extent to which quantum descriptions of a system are to be thought of as effective theories with an in-build ultra-violet cut off. Section II D then considers the dynamics of ontic states. This takes the form of a stochastic process which must satisfy a number of conditions. Most importantly, as we discuss in section II E, it must be coarse-grained at the scale of the ultra-violet cut off in the temporal domain to avoid problems of other modal interpretations. Section III discusses how a recognisable classical ontology can emerge for macro-systems. This involves a discussion of how a classical ontology involves a patching together of ontic states of a number of macro-systems embedded in a larger environment. We show that it is meaningful to define joint probabilities for disjoint systems but only in an emergent sense. Section III B explains how the new interpretation is related to modern formulations of statistical mechanics built on quantum mechanics. Section IV is devoted to a discussion of measurement. We first show in section IV A how a simple model without an environment can account for some features of measurement but also has a number of problems. In section IV B, these problems are all resolved in a more realistic model that includes the environment as well as a measuring device that is not 100% efficient. In section V we discuss the classical experiment of Bohm based on the original thought experiment of Einstein, Podolsky and Rosen and show the new interpretation gives a description that reproduces that of Copenhagen quantum mechanics but without the need to collapse the wave function. Finally, in section VI, we draw some conclusions.

II. THE EMERGENT COPENHAGEN INTERPRETATION

A. Ontic States

A key feature of the new interpretation, as in some other modal interpretations, is the focus on sub-systems of larger systems $A \subset S$: see figure 1. For simplicity, we assume that S is large enough so that its quantum dynamics is to an excellent approximation unitary. In most cases, we can assume that the state of S is pure $|\Psi(t)\rangle$; indeed, much of the what we say will be independent of the exact state of S whether mixed or pure.[5] For each sub-system, $A \subset S$ for which the Hilbert space of S factorizes as $\mathcal{H}_S = \mathcal{H}_A \otimes \mathcal{H}_E$, we can define a reduced density matrix by tracing over the Hilbert space of the complement E:

$$\hat{\rho}_A(t) = \text{Tr}_E |\Psi(t)\rangle\langle\Psi(t)| . \qquad (3)$$

It is a theorem that the reduced density $\hat{\rho}_A(t)$ has a discrete spectrum:

$$\hat{\rho}_A(t)|\psi_i(t)\rangle = p_i(t)|\psi_i(t)\rangle , \qquad (4)$$

where the $\{p_i(t)\}$ are a set of real numbers with $0 \leq p_i(t) \leq 1$ and $\sum_i p_i(t) = 1$. The reduced density matrix $\hat{\rho}_A$ is the *epistemic state* of A.

> **The ontic state:** at a particular instant of time, the new interpretation asserts that A is actually in one of the eigenstates $|\psi_i(t)\rangle$. The state that is actually realised is called the *ontic state*.

This property assignment is done at an instant of time t and therefore a degeneracy in the $\{p_i(t)\}$ is not realistic. However, dealing with problems that arise from degeneracies, or more realistically near-degeneracies, as time evolves is key to building a successful modal interpretation. Note, also that the time-dependence of the ontic state $|\psi_i(t)\rangle$ refers to the time at which the decomposition (4) is made and it is important that these states do not generally solve the Schrödinger equation.[6]

Unlike other modal interpretations, we do not directly interpret $p_i(t)$ as the probability that A is in the ontic state $|\psi_i(t)\rangle$, although this will emerge in certain situations. In fact, the more fundamental probabilities in the new interpretation are conditional probabilities $p_{i|j}(t, t_0)$ that, given the system was in the ontic state $|\psi_j(t_0)\rangle$ at an earlier time $t_0 < t$, the system is in the ontic state $|\psi_i(t)\rangle$ at time t. It is a hypothesis that these conditional probability are related to the single-time probabilities $p_i(t)$

---

[5] Note that the mixed states that we have are all "improper mixtures" and we do not need the concept of a "proper mixture".

[6] In fact, since A is generally interacting with E there is no concept of a Schrödinger equation applying within the sub-system A.

via

$$p_i(t) = \sum_j p_{i|j}(t, t_0) p_j(t_0) \qquad (5)$$

Given this, there are two ways that the $p_i(t)$ can be interpreted as single-time probabilities:

Initial condition: the key equation (5) shows that is an unambiguous way to define the single-time probabilities $p_i(t)$ via an initial condition. If at t_0 the state of S is a tensor product state $|\psi_0(t_0)\rangle \otimes |\phi_0(t_0)\rangle$, then the ontic state at t_0 is uniquely $|\psi_0(t_0)\rangle$. In that case, $p_i(t) = p_{i|0}(t, t_0)$ is the probability that the ontic state is $|\psi_i(t)\rangle$ at time $t > t_0$.

Equilibrium: we will see, in section III B, that there is another definition that is valid when A is a macro-system in equilibrium with its environment E so that $\hat{\rho}_A(t)$ is only slowly varying. In this case, for a characteristic time scale τ, $p_{i|j}(t, t_0)$, with $t = t_0 + \tau$, becomes independent of j, the initial state. In that case, $p_i(t) = p_{i|j}(t, t_0)$ (approximately time-independent) is the probability of finding the system in the ontic state $|\psi_i(t)\rangle$, independent of the initial state $|\psi_j(t_0)\rangle$.

One might imagine that the assignment of a particular ontic state is tantamount to specifying a kind of hidden variable. As we will see this is potentially misleading because the behaviour of ontic states is not at all like standard hidden variables. In particular, the ontic states of A are not global property assignments and the recognition that they are only intrinsic properties from the perspective of A in relation to the rest of the total system is known in the literature on modal interpretations as *relationalism* or *perspectivalism* [12–15]. We want to emphasize that this is not really a philosophical stance but is simply acknowledging what it means to perform a trace that involves summing over states in a disjoint sub-system to A.

There is an important additional detail to mention here. Since the total system $A + E$ is in a pure state $|\Psi\rangle$, then assuming $d_A \leq d_E$, where $d_A = \dim \mathcal{H}_A$, etc., because the environment E is typically much bigger than the sub-system of interest A, each ontic states of A, say $|\psi_i\rangle$, is precisely correlated with a particular ontic state of E, which we call the *mirror ontic state* and denote $|\tilde{\psi}_i\rangle$. This follows because $\hat{\rho}_A$ and $\hat{\rho}_E$ have the same non-vanishing spectrum and one way to exhibit the correlation is via the Schmidt decomposition of $|\Psi\rangle$:

$$|\Psi\rangle = \sum_i \sqrt{p_i}|\psi_i\rangle \otimes |\tilde{\psi}_i\rangle . \qquad (6)$$

The states $|\psi_i\rangle \otimes |\tilde{\psi}_i\rangle$ are a set of d_A orthogonal vectors in the $d_A d_E$ dimensional Hilbert space of $A + E$. So the property assignment of $|\psi_i\rangle$ is always precisely correlated with the mirror assignment of $|\tilde{\psi}_i\rangle$ to its complement. We

will see in section II D that the mirror ontic states play an important role in the dynamics of ontic states.

Finally, we should point out that ontic states are irreducible in the sense that there is no further notion of probability on top of their inherent probability. In this regard, we do not assume a priori the Born rule which we will have to ultimately derive from the behaviour of ontic states in realistic situations.

B. Disjoint Sub-systems

It is important that once we trace down to the Hilbert space $\mathcal{H}_A$ factor, we potentially forgo any knowledge of the ontic states of other disjoint sub-systems. Generally working with the sub-space A is only good for asking inclusive questions regarding the dynamics of ontic states in a disjoint sub-system. In particular, as mentioned in the last section, this means that ontic assignments for A cannot generally be taken as global property assignments.

The only information we can have on disjoint sub-systems A and B are the epistemic and ontic states of the combined system $A + B$. The ontic states of the latter will not generally be related to products of ontic states of A and B: see figure 2. This means that it is not
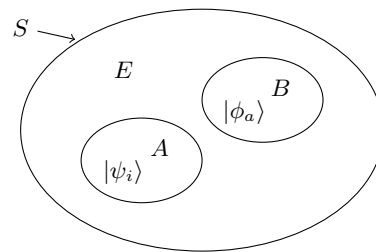

FIG. 2. Two disjoint systems A and B interacting with a large environment E. Ontic states of A and B are $|\psi_i\rangle$ and $|\phi_a\rangle$, respectively, However, ontic states of $A + B$ are not generally equal to tensor products $|\psi_i\rangle \otimes |\phi_a\rangle$ and the ontology is "quantum".

generally consistent to make joint ontic assignments and define joint probabilities to sub-systems A and B. However, in section III A, we will see that under suitable circumstances such joint assignment and probabilities can emerge when A and B are macro-systems embedded in a much larger environment.

In addition, as we saw in the last section, the ontic state $|\psi_i\rangle$ of a sub-system A is precisely correlated with the mirror ontic state $|\tilde{\psi}_i\rangle$ of the E, the complement of A in the sense that $A + E$ is the total system. However, in this case the pair of ontic states have no precise relation to that of $A + E$ which is the pure state $|\Psi\rangle$ in (6).

C. Sub-systems and Effective Theories

Since each sub-system $A \subset S$ enjoys its own set of ontic states and, if the latter are to represent a property assignment, how are we to understand deformations in the definition of the sub-system? This involves discrete changes when we decide to move degrees-of-freedom from the environment E to A, and vice-versa. The new interpretation must explain how a recognisably stable classical ontology can emerge out of the myriad of different possible sub-systems and their associated ontic states. Essentially it does this by recognising that classical states are not directly identified with a particular ontic state of a particular sub-system but rather to coarse grained histories of ontic states and these are not sensitive to the precise microscopic definition of A. We may change the definition of A by moving microscopic degrees-of-freedom into and out of A without affecting the collective behaviour of A that arises from coarse-grained time averages of a time sequence of ontic states. In fact, we will see that for the sub-system A that is in equilibrium with the environment E, its "classical state" is identified with the ensemble or epistemic state $\hat{\rho}_A(t)$. The collective dynamics defined by the ensemble is expected to be stable with respect to microscopic re-definitions of A.

Before we describe time dependence within the new interpretation, it is important that we establish the limitations of a particular quantum mechanical description of a physical system. Many of the problems with existing modal interpretations result from making unrealistic assumptions about the range of validity of the quantum description. Analysing a quantum system involves identifying an appropriate Hilbert space and Hamiltonian such that the resulting dynamics is unitary. However, such descriptions can only be approximately valid above some particular length or time scale.[7] Equivalently, using the uncertainty principle, below a particular momentum or energy scale. For instance, consider a scalar particle. At low enough momenta, non-relativistic quantum mechanics is a good approximation and particle number is effectively a conserved quantity so it make sense to write down effective theories by taking a sector of the full Hilbert space with one particle $\mathcal{H}_1$. Such a description, however, will break down when the momentum increases and relativistic effects become important. This is governed by a momentum scale mc, a length scale $\hbar/mc$ (the Compton wavelength), an energy scale mc^2 and a temporal scale $\hbar/mc^2$. In fact, non-relativistic quantum mechanics based on the one-particle sector leads to violations of causality for measurements based on finite spatial regions on the scale $\hbar/mc$. In that case, to recover a consistent and causal description one must allow particles to be created and destroyed and the effective theory in the one-particle sector is no longer valid; one must instead work with a much larger Hilbert space containing all-particle sectors $\mathcal{H}_0 \oplus \mathcal{H}_1 \oplus \cdots$ in order to have a unitary description. Of course, this is where quantum field theory becomes the more appropriate formulation. At even higher momenta, this description could break down, for instance, if the particle were a composite. At some high momentum scale the constituents would become important and a different effective theory would be needed.

The message here is this that, when we analyse a typical quantum system, we are inevitably doing a low momentum (or low energy, large distance/time scale) approximation. In this effective description there is no sense in which the effective Hilbert space $\mathcal{H}_A$ is a factor of the total Hilbert space of the universe—assuming that the latter even makes sense. In other words, it is not even clear that A has a well-defined parent system in the sense that $\mathcal{H}_S = \mathcal{H}_A \otimes \mathcal{H}_E$. However, in the spirit of effective theory one can imagine that we can identify E with all the degrees-of-freedom at the scale of the effective theory that directly interact, or are entangled, with A. The only requirement is that S is chosen to be large enough in order to achieve an approximately unitary description of the overall dynamics. For overall consistency, it is important that the detailed nature of E, the so-called "environment" or "bath", is largely irrelevant for the behaviour of the sub-system A. This turns out to be the case as long as $d_E \gg d_A$. One could say that it is crucial that the environment is present but that its details are largely irrelevant, including the overall state of S, except in very special situations where the initial state is very non-generic. The latter occurs when a carefully designed measuring device is interacting with a microscopic quantum system.

Given that the analyses of quantum systems are only effective descriptions valid above some particular length/time scale, the so-called ultra-violet cut off, it is important that that the new interpretation yields a formalism that that is not sensitive to phenomena on the cut off scale. This does not means that the new interpretation is not applicable to shorter distances or times, but in order to be valid at more refined scales one would need to apply it to the more fundamental effective theory that takes over at this scale. The importance of this observation is that many of the problems suffered by other modal interpretations result from issues that involve arbitrarily short distance and time scales well beyond the validity of the effective theory. The new interpretation, on the contrary, is immune from these difficulties because it acknowledges the inherent limitations of an effective description. The fact that analyses of quantum systems are only effective, means that there are intrinsic errors to any calculation which involve powers of the characteristic length/time scale to the length/time ultra-violet cut off.

---

[7] This point was made in [1] and elaborated in discussions with Jacob Barandes [46].

D. Ontic Dynamics

If the issue of how to define ontic states at a given instant of time is simple to state, the issue of how ontic states change in time is a subject fraught with problems for all modal interpretations. Our conclusion is that previous approaches are fundamentally flawed because they make unrealistic expectations as to the limits of the validity of the analysis. They do not recognise the key fact, discussed in the last section, that an analysis of a quantum system is generally only an effective one and so is valid only on distance and time scales that are greater than some specific ultra-violet cut off. It is then perhaps no surprise that problems arise when one tries to construct a theory of ontic dynamics that is continuous in time.

A key principle, therefore, is that it only makes sense to define ontic dynamics on a temporal coarse graining scale that is the ultra-violet cut off scale η. So for a particle, for instance, this would be $\eta \sim \hbar/mc^2$. As long as this scale is smaller than any characteristic decoherence time scale of the system under study then the effective description is valid. We denote the

$$\boxed{\text{characteristic decoherence time scale} = \tau}$$

Then the validity of the effective theory requires that

$$\eta \ll \tau \ . \tag{7}$$

The scale τ will be defined more precisely later. The condition (7) ensures that (i) the effective theory is a valid description and (ii) the coarse graining appears smooth at the scale of time-dependent phenomena of the system. To summarise, our coarse graining scale lies at ultra-violet cut off scale of the effective theory and for consistency of the effective description, as long as (7) is satisfied and a Markov condition respected, the whole formalism is then insensitive to the exact value of η since the discretization errors are of order η/τ. However, we cannot attempt to take the cut off $\eta \to 0$ and remain within the validity of the effective theory. Note that even the dynamics of the epistemic state governed by the Schrödinger equation cannot be considered more fundamental because it is also only an equation valid within the effective theory.

Since it is the interaction between A and E that is responsible for the time dependence of the probabilities $p_i(t)$, on the time scale τ, and the reason that the ontic states $|\psi_i(t)\rangle$ do not satisfy the Schrödinger equation, the fact that the coarse graining scale $\eta \ll \tau$ means that over a time step η the set of probabilities and ontic states do not change much. The implication is that there exists a unique one-to-one mapping between ontic states at time t with those at $t+\eta$. We can use the freedom to permute the labels of the ontic states at each time step to ensure that the mapping associates $|\psi_i(t)\rangle$ with $|\psi_i(t+\eta)\rangle$ in the sense that

$$\langle \psi_j(t+\eta)|\psi_i(t)\rangle \sim \delta_{ij} + \mathcal{O}(\eta/\tau) \ , \tag{8}$$

in which case, the associated individual probabilities only change by a small amount

$$p_i(t+\eta) \sim p_i(t)\big(1 + \mathcal{O}(\eta/\tau)\big) \ . \tag{9}$$

It is important to emphasise that this continuity condition (8) is defined at the temporal scale η and we are not at liberty to take the limit $\eta \to 0$ since this would go beyond the domain of applicability of the effective theory. This is fortunate because it saves the new interpretation from the scourge of macro-flips, a disease that infects other modal interpretations. A macro-flip occurs when the eigenvalues $p_i(t)$ and $p_j(t)$, for two macroscopically distinct ontic states, try to cross. But generically the eigenvalues "repel" each other and this leads to an extremely rapid flip, over a time scale of order $\tau\Delta$, of the ontic state from $|\psi_i\rangle$ to the macroscopically distinct state $|\psi_j\rangle$, or vice-versa. Given the importance of this issue, we describe it more fully in section II E. Having said that, apart from solving the problem of macro-flips, having a small but finite cut off η will not affect the dynamics over the physically relevant time scale τ since the discretisation errors involve powers of $\eta/\tau \ll 1$.

Given that the total system $S = A+E$ is assumed to be a pure state $|\Psi(t)\rangle$, the time dependence of the epistemic state $\hat{\rho}_A(t)$ is determined by solving the Schrödinger equation for S giving

$$\hat{\rho}_A(t) = \text{Tr}_E\left[\hat{U}(t,t_0)|\Psi(t_0)\rangle\langle\Psi(t_0)|\hat{U}(t_0,t)\right] \ . \tag{10}$$

where $\hat{U}(t,t_0)$ is the unitary time evolution operator in S. We want to emphasize that it is really only meaningful to describe this epistemic dynamics on the coarse graining scale η.

The problem before us is to write down a similar dynamical equation for the probability that the system A is in the ontic state $|\psi_i(t+\eta)\rangle$ given that it was in the ontic state $|\psi_j(t)\rangle$:

$$p_{i|j}(t) \equiv p_{i|j}(t+\eta, t) \ . \tag{11}$$

These conditional probabilities, along with the Markov property discussed below, define a stochastic process. In order to be consistent with the probability constraint (5), we must have

$$p_i(t+\eta) = \sum_j p_{i|j}(t)p_j(t) \ , \tag{12}$$

but this does not determine the stochastic process uniquely so its definition is a hypothesis. However, we will argue there are some additional natural constraints which lead us to a unique process:

Markov: ontic states carry no memory of their previous history and so the conditional probability to be in the ontic state $|\psi_i(t+\eta)\rangle$ should only depends on the ontic state $|\psi_j(t)\rangle$ and not on ontic states at earlier times. This condition is fundamental to our

whole approach since it ensures that we can build up the more coarse grained dynamics in terms of the microscopic transitions (11) at the ultra-violet scale η.

Locality: the stochastic process should be driven by the local interaction between A and E and so in order that it leads to a local form of dynamics we require that it depends only on the initial and final ontic states $|\psi_j(t)\rangle$ and $|\psi_i(t+\eta)\rangle$ of A and their ontic mirrors $|\tilde{\psi}_j(t)\rangle$ and $|\tilde{\psi}_i(t+\eta)\rangle$ of E as well as $\hat{H}_{\text{int}}$ the part of the total Hamiltonian

$$\hat{H} = \hat{H}_A \otimes I_E + I_A \otimes \hat{H}_E + \hat{H}_{\text{int}} , \qquad (13)$$

that describes the local coupling between A and E.[8]

Ergodicity: generically we require that any state can reach any other state in a finite number of steps. However, this must break down for two macroscopically distinct states $|\psi_i\rangle$ and $|\psi_j\rangle$, for which we require that there should only be a minute probability of order Δ (a typical inner product of two macroscopically distinct states described in section I) of a transition between them over a time scale τ.

As long as the stochastic process has these properties, then its actual microscopic details are largely irrelevant to the behaviour of macro-systems. The Markov property is very natural given that in the new interpretation the ontic state of a system is just a property assignment at a particular time and has no memory of previous ontic states. Moreover imposing this condition is fundamental because it means that the ultra-violet dynamics $p_{i|j}(t)$ determines the whole stochastic process since over a series of time steps $t_n = t + (n-1)\eta$, according to the Markov property,

$$p_{j_N|j_1}(t_N, t_1) = \sum_{j_2, \ldots, j_{N-1}} \left[\prod_{n=1}^{N-1} p_{j_{n+1}|j_n}(t_n) \right] . \quad (14)$$

The resulting stochastic process is therefore a conceptually simple discrete-time Markov chain.

We now turn to the definition of the ultra-violet dynamics. A useful observation involves the matrix elements

$$V_{ij}(t) = \left(p_i(t+\eta)p_j(t) \right)^{1/2}$$
$$\times \operatorname{Re} \langle \psi_i(t+\eta)| \otimes \langle \tilde{\psi}_i(t+\eta)|\hat{U}(t+\eta, t)|\psi_j(t)\rangle \otimes |\tilde{\psi}_j(t)\rangle , \quad (15)$$

where $|\tilde{\psi}_i\rangle$ are mirror ontic states of E defined in (6) and $\hat{U}(t+\eta, t)$ is the unitary time evolution operator in $A+E$.

Note that these matrix elements are completely symmetrical between A and E and the resulting dynamics will consequently respect the exact correlation between their ontic states. One finds from this definition and (6), that

$$p_i(t+\eta) = \sum_j V_{ij}(t) , \qquad p_j(t) = \sum_i V_{ij}(t) , \quad (16)$$

so it it tempting to relate

$$V_{ij}(t) \longleftrightarrow p_{i|j}(t)p_j(t) . \quad (17)$$

But we cannot have equality here, because the matrix elements $V_{ij}(t)$ are not necessarily valued between 0 and 1. However, we can proceed as follows. If we have labelled the states to be consist with the continuity condition (8), it follows that since

$$p_i(t+\eta) - p_i(t) = \sum_j \left[V_{ij}(t) - V_{ji}(t) \right] , \quad (18)$$

the right-hand side must be small of order η/τ. We can define the ultra-violet process by taking, for $i \neq j$,[9]

$$p_{i|j}(t) = \frac{1}{p_j(t)} \max \left[V_{ij}(t) - V_{ji}(t), 0 \right] \quad (19)$$

and

$$p_{i|i}(t) = 1 - \sum_{j \neq i} p_{j|i}(t) . \quad (20)$$

This process satisfies the probability constant (12).

It is important to realize that there is no guarantee that the process defined above is consistent in the sense that the conditional probabilities $p_{i|j}(t)$ are valued in the interval $[0, 1]$. In fact, the consistency conditions are

$$\sum_{j \neq i} p_{j|i}(t) \leq 1 , \qquad \forall i . \quad (21)$$

As $\eta \to 0$, the elements $p_{i|j}(t)$, $i \neq j$, can be made arbitrarily small and so the process can always be made consistent in this limit. Hence, there is an upper bound on how big the cut off η can be taken. In order to investigate the this, we can interpret δp_i for a single time step as being due to a mismatch between the flows into and out of the ith state, that is

$$\delta p_i = \sum_{j \neq i} p_{i|j} p_j - \sum_{j \neq i} p_{j|i} p_i . \quad (22)$$

---

[8] Issues involving locality should properly be formulated in terms of relativistic quantum field theory, so our notion here is more primitive.

[9] Using these expressions it is easy to see that the process defined in (19) and (20) above agrees with the one defined in [1]: to compare formulae the quantities $p_{ij}^{(n)}$ in [1] are the conditional probabilities $p_{i|j}(t_n)$ here.

The net flow out of the ith state involves the sum on the left-hand side of (21). If we define the decoherence time scale τ by

$$\tau = \eta \Big[\sup_i \sum_{j \neq i} p_{j|i} \Big]^{-1} . \qquad (23)$$

So consistency of the process requires $\eta \ll \tau$. Note that if $\eta \not\ll \tau$ then this does not imply a breakdown of the formalism but rather a breakdown of the validity of the effective theory: one should go to a more fundamental effective theory valid at smaller distance/time scales.

Generically, both terms on the right-hand side of (22) will be of the same order, so that

$$\delta p_i \sim \mathcal{O}\big(p_i \eta / \tau \big) . \qquad (24)$$

However, later, we will describe the situation when A is in equilibrium with the environment, in that case the p_i will be approximately constant on account of a balance between the flows into and out of the states in (22). It is important to notice that ontic states with very small probabilities $p_j(t)$ do not give anomalously large values of $p_{i|j}(t)$ as might be inferred from (19) because the factor of $p_j(t)$ in the denominator is generally balanced by a factor of a similar order in the numerator.

In the limit $\eta \ll \tau$, we can evaluate the matrix elements (15) in perturbation theory:

$$V_{ij} = \frac{\eta \sqrt{p_i p_j}}{\hbar} \, \text{Im} \, \langle \psi_i | \otimes \langle \tilde{\psi}_i | \hat{H}_{\text{int}} | \psi_j \rangle \otimes | \tilde{\psi}_j \rangle + \cdots . \quad (25)$$

The fact that the matrix elements in (25) only depends on the coupling $\hat{H}_{\text{int}}$ and the ontic states $| \psi_i \rangle$ and $| \psi_j \rangle$ and their mirrors $| \tilde{\psi}_i \rangle$ and $| \tilde{\psi}_j \rangle$ encapsulates the locality requirement. Note that at this leading order $V_{ij} = -V_{ji}$.

In addition, the fact that the transition probabilities depend on a matrix elements (25) involving the states $| \psi_i \rangle \otimes | \tilde{\psi}_i \rangle$ means that the stochastic process will be seen to satisfy the ergodicity requirement. Essentially, if the two states are macroscopically distinct states, then we can expect $p_{i|j}(t)$ will be suppressed by a factor of order Δ relative to the generic situation. The probability that the system will make a transition between the ontic states will therefore be vanishing small. In fact, over a time T the chance that the system will make a transition from one macroscopically distinct state to another would be order $T \Delta / \tau$. It is clear, given the crude estimate of Δ in section I, we would have to wait of the order of $e^{10^{32}}$ times the age of the universe to see such a transition.

It is worth emphasising that the ultra-violet stochastic dynamics we have defined is not unique. The ambiguity corresponds to changing

$$\delta p_{i|j} = \frac{\Theta_{ij}}{p_j} , \quad \sum_i \Theta_{ij} = \sum_j \Theta_{ij} = 0 , \qquad (26)$$

subject to the constraint (21). However, there are no obvious quantities Θ_{ij} that could be defined that are at the same time consistent with the ergodicity and locality requirements, i.e. defined by the 2 states $| \psi_i \rangle \otimes | \tilde{\psi}_i \rangle$ and $| \psi_j \rangle \otimes | \tilde{\psi}_j \rangle$ and the interaction Hamiltonian H_{int} between A and E. Moreover, our process is very closely related to that written down by Bell in his theory of beables [31].[10] We take the process that we have defined as being a hypothesis on the same level as the Schrödinger equation that determines the dynamics of epistemic state. However, it is worth emphasising that:

> As long as the ergodicity condition is satisfied, along with the key probability relation (5), the microscopic details of the stochastic process are actually irrelevant for reproducing standard Copenhagen interpretation phenomenology of macro-systems which are in equilibrium with their environment (as described in section III B).

Finally, it is worth making clear the point that, although we have introduced an auxiliary stochastic process to define the dynamics of ontic states, this is completely different from dynamical collapse models discussed in the literature; for example in the review [43]. The latter involve stochastic modifications of Schrödinger's equation itself, in other words they involve introducing stochastic dynamics for the epistemic state which is a completely different philosophy from the one we are setting out here.

E. The Continuum Process and Macro-Flips

It is tempting, even though it runs counter to the methodology of effective theory, to take the stochastic process that we defined in the last section and take the cut off $\eta \to 0$, in order to define a continuum process. However, typically one should expects pathologies to arise when effective theories are pushed beyond their range of validity. Indeed, in the present case, a pathology manifests as the existence of micro-flips that represent a severe problem for existing modal interpretations that insist on following ontic states continuously in time. In this section, we will describe how they arise and how the new interpretation avoids them.

The ontic states of a sub-system $| \psi_i(t) \rangle$ are defined continuously in time and so it is tempting to define ontic dynamics that is also continuous in time. In fact, the continuum limit of (18) takes the form

$$\frac{dp_i}{dt} = \sum_j J_{ij} , \qquad J_{ij} = -J_{ji} , \qquad (27)$$

with

$$J_{ij} = \frac{2 \sqrt{p_i p_j}}{\hbar} \, \text{Im} \, \langle \psi_i | \otimes \langle \tilde{\psi}_i | \hat{H}_{\text{int}} | \psi_j \rangle \otimes | \tilde{\psi}_j \rangle , \qquad (28)$$

---

[10] In fact, in the continuum limit $\eta \to 0$ it corresponds to taking the reduced density matrix ρ_A as the beable.

which is equal to the perturbative form of $(V_{ij} - V_{ji})/\eta$ using (25). Written in this form, manifests the fact that if A does not interact with E then J_{ij} vanishes.[11] In that case, a continuum Markov process can be defined whose master equation takes the form

$$\frac{dp_i}{dt} = \sum_{j \neq i} \left(T_{ij}p_j - T_{ji}p_i \right) , \qquad (29)$$

corresponding to transitions into and out of $|\psi_i\rangle$. For the stochastic process satisfying the ergodicity condition defined in the last section

$$T_{ij} = \frac{1}{p_j} \max \left(J_{ij}, 0 \right) . \qquad (30)$$

These are the transition probabilities originally suggested by Bell [31] and further analysed in the context of modal interpretations by Bacciagaluppi and Dickson [10]. This is the continuum version of (19) obtained in the limit $\eta \to 0$.

Now we turn to the issue of macro-flips. The problem of the continuous stochastic process above, occurs when two eigenvalues of $\hat{\rho}_A(t)$, say $p_+(t)$ and $p_-(t)$, associated to macroscopically distinct ontic states, try to cross as illustrated in 3.

If the eigenvalues actually *do* cross then there is no problem because then one can define the continuity of the ontic states in time by imposing analyticity across the point of degeneracy. This ensures that there are no abrupt transitions between the two macroscopically distinct states. However, generically the eigenvalues will not cross and, on the contrary, there will be a crossover. In order to analyse what happens we can isolate the important part of the reduced density matrix in the 2-dimensional subspace spanned by the exactly orthogonal macroscopically distinct states $|\phi_\pm\rangle$ roughly constant in the neighbourhood of the degenerate point at $t = 0$:

$$\rho_{\text{subspace}} \sim \begin{pmatrix} p_0 + a_1 t & p_0 \Delta \\ p_0 \Delta & p_0 + a_2 t \end{pmatrix} . \qquad (31)$$

Here, a_i and p_0 are real constants and we define $a = (a_1 - a_2)/2$. Note that the a_i will be of the order τ^{-1}, where τ is, as previously, a characteristic decoherence scale. Here, Δ is a measure of the typical inner-product of two macroscopically distinct states. As described in section I, for a typical macroscopic system we might have $\Delta \sim e^{-10^{32}}$. In the absence of the tiny inner product between the macroscopically distinct states, i.e. when $\Delta = 0$, the ontic states (eigenvectors) are $|\psi_i\rangle = |\phi_i\rangle$

and the eigenvalues actually cross at $t = 0$. However, when $\Delta \neq 0$ the eigenvalues are

$$p_\pm(t) = p_0 + \frac{a_1 + a_2}{2} t \pm \sqrt{(at)^2 + (p_0 \Delta)^2} \qquad (32)$$

and the level crossing is avoided since, although the two eigenvalues become close, they never actually cross. Near the crossover, the eigenvectors are approximately

$$|\psi_\pm(t)\rangle = \cos\theta(t)|\phi_\pm\rangle \pm \sin\theta(t)|\phi_\mp\rangle , \qquad (33)$$

where

$$\tan\theta(t) = \frac{at + \sqrt{(at)^2 + (p_0 \Delta)^2}}{p_0 \Delta} , \qquad (34)$$

which goes from 0 to $\frac{\pi}{2}$ as t increases through the crossover. The time for the crossover to occur is order $\tau\Delta$. This reveals the problem: by avoiding the degeneracy the system exhibits an instability in the sense that it flips between the macroscopically distinct states $|\phi_+\rangle$ and $|\phi_-\rangle$ in a very short time of order $\tau\Delta$. This time scale, for any system, will be much smaller than even the Planck scale and therefore certainly much smaller than the cut off scale η. So an interpretation based on the continuous time stochastic process suffers from unacceptably rapid switching between macroscopically distinct states—a macro-flip—whenever probabilities try to cross. Of course such macro-flips are completely unphysical and one can only conclude that the continuous-time stochastic process, pushing as it does beyond the domain of applicability of the effective theory, is fundamentally flawed.
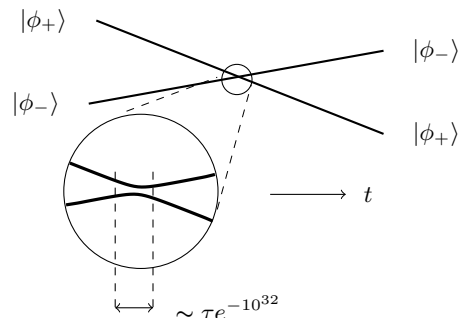

FIG. 3. A crossover where there is an approximate degeneracy as $p_+(t)$ and $p_-(t)$ try to cross but ultimately "repel" on a time scale of order $\tau\Delta$. Any continuous-time stochastic process will involve a rapid switching between macroscopically distinct states $|\phi_\pm\rangle$—a macro-flip. On the contrary, our coarse-grained stochastic process cannot "resolve" the crossover and a macro-flip does not occur even though the cut-off scale η is much smaller than any other characteristic time scales in the problem: $\tau \gg \eta \gg \tau\Delta$.

On the contrary, our coarse-grained stochastic process defined in section II D is immune from macro-flips. The reason is that the crossover scale $\tau\Delta \ll \eta$, even though

---

[11] Note that the expression for J_{ij} seems to be missing a term involving a time derivative compared with [1, 10]. However, it is easy to see that, since $\partial_t(|\psi_j\rangle \otimes |\tilde{\psi}_j\rangle) = \partial_t|\psi_j\rangle \otimes |\tilde{\psi}_j\rangle + |\psi_j\rangle \otimes \partial_t|\tilde{\psi}_j\rangle$ and the sets of states $\{|\psi_j\rangle\}$ and $\{|\tilde{\psi}_j\rangle\}$ are orthonormal, this term actually vanishes. In addition, the total Hamiltonian may be replaced by $\hat{H}_{\text{int}}$ for the same reason.

$\eta \ll \tau$, and so a crossover simply cannot be resolved by the coarse-grained stochastic process. The smoothness condition (8) then ensures that across the time step that includes the crossover the ontic states are preserved $|\psi_i(t+\eta)\rangle \approx |\psi_i(t)\rangle \approx |\phi_i\rangle$, to order η/τ, and no dramatic macro-flip of the ontic state occurs.

III. THE CLASSICAL LIMIT

A. Emergent Classicality

It is key feature of the new interpretation that what we ordinarily understand as classical behaviour only emerges in appropriate situations and is not put in by hand. The fact that a successful phenomenology of macro-systems arises rests largely on the fact that the stochastic process defined in the last section satisfies the ergodicity condition.

More fundamentally, the new interpretation makes statements about ontology that are precise but they are not necessarily of a classical kind. For instance, as discussed in section II B, an interpretation based on reduced density matrices cannot generally make joint property assignments to two disjoint systems A and B let alone associate probabilities to them. The only statements that can be made involving both A and B are via the combined system $A + B$ and in this case its ontic states will generally not be a tensor product of ontic states of the sub-systems, i.e. $|\psi_i\rangle \otimes |\phi_a\rangle$. The relation between ontic states of the three systems A, B and $A + B$ will be more fuzzy and potentially contradictory. One could say that the fragments of reality cannot be drawn together to form a consistent whole. We call this a "quantum" ontology.

However, a familiar "classical" ontology can be an emergent phenomenon in the following sense. Suppose A and B are two disjoint weakly-interacting or causally separated macro-systems in, or close to, equilibrium and hence strongly entangled with the environment E with $d_A, d_B \ll d_E$. We expect in these circumstances that the ontic states of $A + B$ will indeed approximately factorize into a tensor product of the ontic states of A and B. The mismatch will involve typically minute order Δ effects:

$$|\Phi_{m(i,a)}\rangle = |\psi_i\rangle \otimes |\phi_a\rangle + \mathcal{O}(\Delta) \ , \qquad (35)$$

where $m = m(i,a)$ is a 1-to-1 map. What this means is the descriptions provided by A, B and $A + B$ can be integrated into a consistent whole, at least to high degree of accuracy.

In this context, it is meaningful to make joint property assignments for A and B and we can interpret, in an emergent sense, the probability $p_{m(i,a)}$ as the joint probability for a pair of ontic states $|\psi_i\rangle$ and $|\phi_a\rangle$ of A and B:

$$p(i,a) \overset{\text{emergent}}{=} p_{m(i,a)} \ . \qquad (36)$$

It is the view from $A + B$ that is needed to follow any potential correlations between the ontic states of A and B. Note also that the ontic states of A or B still cannot generally be taken as global property assignments outside of the triplet of sub-systems A, B and $A + B$.

The emergent joint probabilities satisfy the usual probability relations, but only to order Δ,

$$\sum_a p(i,a) = p_i + \mathcal{O}(\Delta) \ ,$$
$$\sum_i p(i,a) = p_a + \mathcal{O}(\Delta) \ , \qquad (37)$$

although $\sum_{ia} p(i,a) = 1$.

In general, the ontic states of A and B can be correlated in a classical (non-entangled) sense when $\hat{\rho}_{A+B} \neq \hat{\rho}_A \otimes \hat{\rho}_B$ meaning that

$$p(i,a) \neq p_i \, p_a \ . \qquad (38)$$

The case when A and B are not correlated corresponds to when $\hat{\rho}_{A+B} = \hat{\rho}_A \otimes \hat{\rho}_B$ and $p(i,a) = p_i p_a$.

Note that we will meet an example in section V an example where the ontic states of $A + B$ are tensor product states of A and B but the factors of one of them are not the ontic states of the corresponding sub-system. So simply being a tensor product state is not sufficient to have a "classical" ontology.

The picture above of emergent joint ontic assignments and joint probabilities can be generalized to many weakly interacting or causally disconnected macro-systems $A_1 + \cdots + A_n$. An ontic state of the parent system will be, to order Δ, a tensor product state $|\Phi_{m(i^{(1)},...,i^{(n)})}\rangle = |\psi_{i^{(1)}}^{(1)}\rangle \otimes \cdots \otimes |\psi_{i^{(n)}}^{(n)}\rangle$, and so emergent joint probabilities can be defined of the form

$$p\big(i^{(1)}, \ldots, i^{(n)}\big) \overset{\text{emergent}}{=} p_{m(i^{(1)},...,i^{(n)})} \ . \qquad (39)$$

The picture we have here is that the emergent classical world involves patching together very slightly different descriptions—differing at $\mathcal{O}(\Delta)$—of the same systems from the point-of-view of A_i, $A_i + A_j$, $A_i + A_j + A_k$, etc. For more microscopic systems this integration of ontic states becomes more ambiguous and a classical description evaporates to be replaced by an ontology that is truly quantum. We can quantify the degree of classicality in terms of the generic scale Δ. So we can expect systems to exhibit quantum fuzziness when Δ is not so small so that the relation between the ontic states becomes ambiguous and joint ontic assignments and joint probabilities cannot be consistently defined.

B. Link with Statistical Mechanics

A key requirement of an interpretation of quantum mechanics is to explain how the classical behaviour of macroscopic systems emerges. Macro-systems with many

degrees-of-freedom are complicated systems whose collective behaviour is captured by the techniques of statistical mechanics. It seems natural that any quantum origin of classical behaviour must, at the very least, be able to give a consistent foundation to classical statistical mechanics. In fact, we might hope that such an understanding would put classical statistical mechanics on a firmer conceptual footing given that it is still, somewhat surprisingly, a controversial subject. In particular, there is no consensus on the role of probability, the meaning of entropy and the relation of ensemble averages to time averages.

In the last few year a rather different and intrinsically quantum approach to the subject has been developed [23–30]. In this approach, ensembles arise at the quantum level when a system A is entangled with a large thermal bath, or environment, E. So the total system $S = A + E$ can be in a pure state but, nevertheless, the sub-system A has a reduced density matrix that defines an ensemble $\hat{\rho}_A$. In this point-of-view, the thermodynamic entropy of A is precisely the entanglement entropy of the sub-system $S = -\text{Tr}(\hat{\rho}_A \log \hat{\rho}_A)$ which is non-vanishing when A is non-trivially entangled with the bath.

When the bath is much bigger than the system $d_E \gg d_A$ there are some very powerful principles that emerge.[12] For almost any pure state of the total system in some subspace $\mathcal{H}_R \subset \mathcal{H}_S$ described by some global constraint R (preserved under time evolution) the reduced density matrix of the system A is approximately equal to

$$\hat{\rho}_A \approx \frac{\text{Tr}_E \hat{\mathscr{P}}_R}{d_R} \, , \qquad (40)$$

where $\hat{\mathscr{P}}_R$ is the projection operator on the subspace $\mathcal{H}_R$. If the global constraint is on the energy and the interaction between the system A and the environment E is sufficiently weak then it is straightforward to show that $\hat{\rho}_A$ is approximately the canonical ensemble,

$$\hat{\rho}_A \approx \frac{e^{-\beta \hat{H}_A}}{Z} \, , \qquad Z = \text{Tr}_A \, e^{-\beta \hat{H}_A} \, . \qquad (41)$$

However, it is important that the principle applies to other possible constraints R on the total system, including the absence of a constraint, and also to cases where the interaction between the system and bath is not small.

The implications for statistical mechanics are evident. If the total system starts out in almost any state in the subspace $\mathcal{H}_R$, then the state of the sub-system A is approximately the state (40) independent of time. This describes a situation where the system is in equilibrium with its environment. In this analysis, the entropy becomes an objective property of the state of A caused by entanglement with E.

More general questions involve situations when the system starts off in a state which is not in equilibrium [28, 29]. The results are summarised below:[13]

Equilibration: subject to some reasonable conditions, every pure state of S is such that a small sub-system $A \subset S$ will equilibrate meaning that $\hat{\rho}_A(t)$ approaches a limit which fluctuates about a constant. Note that the initial state does not need to be "typical", in other words even though the overwhelming number of states of S are such that A is in equilibrium already, it is also true of initial states where A is far from equilibrium. This includes tensor product states $|\Psi\rangle = |\psi\rangle \otimes |\phi\rangle$.

Bath independence: in the case that the initial state is a tensor product $|\Psi\rangle = |\psi\rangle \otimes |\phi\rangle$, the equilibrium state of A is independent of the state of the bath $|\phi\rangle$.

Sub-system independence: there are general conditions under which the equilibrium state of A is independent of $|\psi\rangle$ the initial state of A. However, there are also non-generic situations for which the equilibrium state depends sensitively on $|\psi\rangle$.

It is clear that the focus on sub-systems means that this re-formulation of statistical mechanics is closely related with modal quantum mechanics. Moreover, the new interpretation adds a new and important detail to the story through the existence of the ontic state of the sub-system A. This is analogous to the micro-state of classical statistical mechanics. However, it is important to point out that its dynamics, described by the stochastic process described in section II D, is conceptually simpler than the dynamics of micro-states in classical statistical mechanics because in the quantum case the number of ontic states is always finite and the stochastic process is a simple discrete-time Markov chain.

When the system A equilibrates, $\hat{\rho}_A(t)$ and its eigenvalues will fluctuate around a slowly varying quasi-equilibrium and the underlying Markov chain becomes approximately *homogeneous*: that is the transition matrix $p_{i|j}(t)$ becomes time independent over time scales of order τ. Of course, there may be much slower time dependence for on scales $\gg \tau$. The rate of flow into and out of each ontic state in (22) approximately balances. Under generic conditions, although the microscopic transitions between a pair of states only go one way, over a finite number n of time steps $p_{i|j}(t + n\eta, t)$ is a matrix whose entries are all > 0 and hence the equilibrium process is *regular*. The meaning of this is that any ontic state is only a finite number of time steps away from any other state. The fact that the process is a *regular* homogeneous Markov chain implies that it is also *ergodic* and then it

―――――

―――――

is a standard result that

$$\lim_{n \to \infty} p_{i|j}(t + n\eta, t) = p_i(t + n\eta) \;, \qquad (42)$$

independent of j, and so, whatever the initial state, after a large number of steps, the probability distribution is equal to $p_i(t)$. More precisely, one can show that the number of time steps must be at least order τ/η; in other words, one must wait for a time of the order of τ, the decoherence time defined by (23), for the memory of the initial state to be lost:

$$\boxed{\text{equilibrium:} \qquad p_{i|j}(t + \tau, t) \approx p_i(t + \tau)} \qquad (43)$$

This means that, in equilibrium, the actual ultra-violet details of the stochastic process are hidden over time scales of order τ, the decoherence time.[14]

In addition, ergodicity means that, in equilibrium, the time average of a temporal sequence of ontic states over a time scale of order τ is well approximated by the ensemble average described by $\hat{\rho}_A(t)$. This gives us another way to interpret the single-time probabilities $p_i(t)$: when A has reached equilibrium with the bath, the single-time probabilities $p_i(t)$ are approximately constant and are equal to the probability that, in a suitably coarse-grained time average, the system is in the ontic state $|\psi_i(t)\rangle$, independent of the initial state in the past. From the point-of-view of the emergent classical view, the ontic state of a macro-system in equilibrium is hidden and the probabilities $p_i(t)$ can be given the ignorance interpretation:

> When a system is in equilibrium with its environment, the ensemble associated to the—approximately constant—reduced density matrix $\hat{\rho}_A(t)$ captures the time average over the dynamics of the ontic state. So we can associate the classical description of a macro-system with the collective behaviour of the ensemble $\hat{\rho}_A(t)$.

It is noteworthy that the equilibrium state $\hat{\rho}_A(t)$ does not depend on the details of the initial state of the bath, however, the same cannot be said in all circumstances for the initial state of A. In certain situations, discussed in [28], the final equilibrium epistemic state of A can depend very sensitively on the initial state. This is connected to a breakdown of ergodicity of the stochastic process. We

can expect this to happen when the equilibrium state of A has sets of ontic states which are macroscopically distinct. In these conditions, the matrix elements $V_{ij}(t)$ between states in different sets will be minute of order $\eta\Delta/\tau$. In this case, ergodicity of the stochastic process is broken and once the system has equilibrated the resulting time average of a temporal sequence of ontic states is then only captured by a sub-ensemble of $\hat{\rho}_A(t)$. We could summarise the situation by saying that:

> In the new interpretation, both the correlata—the ergodic subsets of states described by the sub-ensembles—and the correlations, in the form of the joint probabilities (36), are emergent quantities.

In this situation, the dynamics of the underlying ontic state as described by the stochastic process is very sensitive to the initial state of A. This is precisely what happens in a phase transition in a statistical system like the Ising magnetic described in the introduction as the temperature is lowered. But it also describes what happens in a quantum measurement where the measuring device is designed to be very sensitive to the quantum state of the system being measured. We will argue in §IV, that when the measuring device equilibrates with the environment there is a breaking of ergodicity of the underling Markov chain such that each ergodic component corresponds to a distinct measurement outcome. Collapse of the wave function corresponds to a tidying up exercise in which one removes the ergodic components that are not reachable from the component that corresponds to the particular measurement outcome that is realized, as illustrated in figure 4.
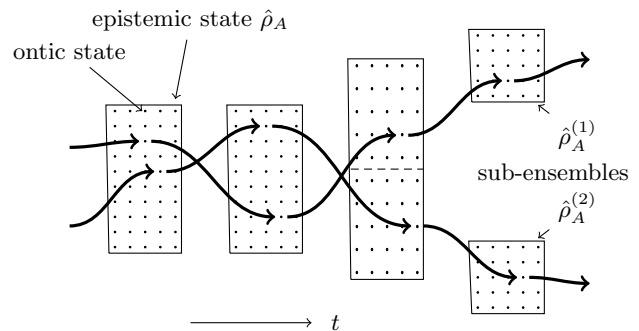

FIG. 4. An illustration of the process of ergodicity breaking during a measurement. Here, the set of of ontic states splits into two distinct sub-ensembles between which the probability of transition is vanishingly small. Two ontic histories are shown that end up in different sub-ensembles after the measurement.

---

[14] As a simple example of an equilibrium process, suppose that at equilibrium A is maximally entangled with E, so $p_i \approx 1/d_A$. If we define the process by taking one of each pair $\{p_{i|j}, p_{j|i}\}$ randomly and giving it the value p, while the other vanishes. In this case, if one takes $p_{i|j}(t + T, t) - p_{i|k}(t + T, t)$ for some randomly chosen distinct i, j and k, then this approaches zero as $\exp(-T/\tau)$, with the time to approach equilibrium $\tau = 2\eta/(pd_A)$, valid when $\tau \gg \eta$. On the other hand, the decoherence time is defined in (23) as $\eta\big[\sum_{j \neq i} p_{j|i}\big]^{-1} \approx 2\eta/(pd_A) = \tau$ so this confirms that the equilibrium time is equal to the decoherence time.

IV. MEASUREMENTS AND THE COLLAPSE OF THE WAVE FUNCTION

In this section we show how the new interpretation can give a convincing account of the measurement process. There are some benefits in starting with a simple model, showing how it successfully describes certain aspects but also how it has certain limitations. Then we introduce a more sophisticated model which preserves the good features and solves the problems. This shows that models need to be realistic otherwise one can be led to misleading conclusions. Along the way, we will need to use intuition from the theory of decoherence, e.g. [16–22].

A. A Naïve Model

The simplest set up involves an idealized von Neumann measurement on a microscopic system. Let us suppose the system P is some finite dimensional quantum system with some basis states $|\xi_i\rangle$ that are eigenstates of the operator associated to the observable we want to measure. The system P is then coupled to a measuring device A through a Hamiltonian which acts as

$$\hat{H}|\xi_i\rangle \otimes |\psi\rangle = |\xi_i\rangle \otimes \hat{H}^{(i)}|\psi\rangle \ , \qquad (44)$$

If the initial state of the combined system is taken to be the non-entangled state

$$|\Psi(0)\rangle = \Big[\sum_i c_i|\xi_i\rangle\Big] \otimes |\psi(0)\rangle \ , \qquad (45)$$

then using the linearity of the Schrödinger Equation we have

$$|\Psi(t)\rangle = \sum_i c_i|\xi_i\rangle \otimes |\psi_i(t)\rangle \ , \qquad (46)$$

where $|\psi_i(t)\rangle$ is the state that evolves from $|\psi(0)\rangle$ via the effective Hamiltonian $\hat{H}^{(i)}$. In order for A to be efficacious, it must be that the states $|\psi_i(t)\rangle$ become macroscopically distinct after a macroscopic time T. This is the process of decoherence [16–22]. We expect the inner product of the states $|\psi_i(t)\rangle$ to exhibit a behaviour roughly of the form

$$\big|\langle\psi_i(t)|\psi_j(t)\rangle\big| \sim \exp\big[-NX(t)^2/\ell^2\big] \ , \qquad (47)$$

for $i \neq j$, where $X(t)$ describes how the distance between the microscopic constituents of the measuring device behaves between the two measurement outcomes. We expect that this goes from 0 to the macroscopic scale L at the end of the measurement at $t = T$, and so

$$\big|\langle\psi_i(T)|\psi_j(T)\rangle\big| \sim \Delta \ , \qquad i \neq j \ . \qquad (48)$$

We can estimate how fast decoherence occurs compared with the macroscopic measuring time T by assuming that

$X(t)$ is linear in t. As an example, by taking the same values for N, ℓ and L as in section I, we find

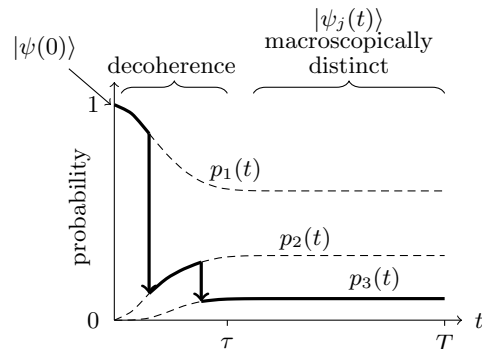$$\frac{\tau}{T} \sim \frac{\ell}{L\sqrt{N}} \leq e^{-16} \ . \qquad (49)$$

FIG. 5. An example of how the probabilities $p_i(t)$ might behave for the case $d_P = 3$. During the decoherence period $0 \leq t \leq \tau$, the states $|\psi_i(t)\rangle$ become macroscopically distinct while the probabilities change in time. At the end of the measurement at $T \gg \tau$, the eigenstates of the reduced density matrix are approximately the states $|\psi_i(T)\rangle$ that occur with probability $p_i(T)$. Also show in bold is a hypothetic solution of the stochastic process in this case with two transitions.

Now we apply the new interpretation to the model. The reduced density matrix of the measuring device is

$$\hat{\rho}_A(t) = \sum_i |c_i|^2 |\psi_i(t)\rangle\langle\psi_i(t)| \ . \qquad (50)$$

We can now follow the ontic states in time as decoherence occurs [1, 34]: see figure 5. In this case, one can verify that only the transition probabilities $p_{2|1}(t)$, $p_{3|1}(t)$ and $p_{3|2}(t)$ are non-vanishing and so the histories of ontic states are rather simple as illustrated in figure 5. After the measurement is complete $t = T$, the ontic states of A are approximately equal to the $|\psi_i(T)\rangle$ up to terms which are order Δ. But, as emphasised earlier, Δ is so much smaller than any other dimensionless scale in the problem and in particular much smaller than the errors intrinsic in the effective theory and so can safely be ignored. Given that initially the measuring device has a unique ontic state $|\psi(0)\rangle$, the corresponding probability to be in the ontic state $|\psi_i(T)\rangle$ after the measurement is

$$p_i(T) = |c_i|^2 + \mathcal{O}(\Delta) \ . \qquad (51)$$

So the new interpretation yields a satisfactory phenomenology in this simplest measurement model and, furthermore, yields the Born rule for the probabilities. There are, however, problems with the model:

(i) When the initial microscopic system has a degeneracy $|c_i| = |c_j|$ for some $i \neq j$. In the case of exact degeneracy one can easily show that the associated eigenvectors of $\hat{\rho}_A(T)$ after the measurement are no longer

$|\psi_i(T)\rangle$ and $|\psi_j(T)\rangle$, but rather involves the macroscopic superpositions [34]

$$\frac{1}{\sqrt{2}}\big(|\psi_i(T)\rangle \pm |\psi_j(T)\rangle\big) \ . \tag{52}$$

There is a superficial answer to this problem which says that a degeneracy of the state of P requires infinite fine tuning that is unrealistic in practice [46]. Whilst this is true, one can show that even when there is no exact degeneracy the onset of decoherence can be delayed in an artificial way by having nearly degenerate states. If $|c_1|^2 = \frac{1}{2} + e^{-s}$ and $|c_2|^2 = \frac{1}{2} - e^{-s}$ then the decoherence time is proportional to $\sqrt{s}$ and so can be increased by tuning close to the degenerate point.

(ii) The model does not satisfy our requirement that the ontic states of macroscopic systems are robust against microscopic re-definitions of the sub-system. In this case, if we re-define A to include the microscopic system P itself then the ontic state of $A + P$ now becomes the pure state of the total system (46) with probability 1. So the slightly re-defined system, with the addition of one microscopic degree-of-freedom, has led to a completely different ontology involving unacceptable superpositions of macroscopically distinct states.

(iii) As it stands the model describes the case of a perfect measuring device that makes no errors. In realistic situations one could imagine that A makes errors which manifests by less than perfect correlation between the states of P and A in (46). A more realistic situation has

$$|\Psi(t)\rangle = \sum_i f_{ij}(t)|\xi_i\rangle \otimes |\psi_j(t)\rangle \ . \tag{53}$$

with $f_{ij}(t) \rightarrow f_{ij}^{(0)}$, a constant, after the measurement time $t = T$. Now the reduced density matrix of the measuring device is

$$\hat{\rho}_A(T) = \sum_{ijk} f_{ij}^{(0)} f_{ik}^{(0)*} |\psi_j(T)\rangle\langle\psi_k(T)| \tag{54}$$

and since the matrix $f_{ij}^{(0)}$ need not be diagonal the ontic states will involve superpositions of the macroscopically distinct states $|\psi_i(T)\rangle$ [35].

B. A More Realistic Model

We now turn to a more sophisticated but realistic model which solves all three problems above. The new feature is that the measuring device is also interacting with an environment E and so the complete system is $P + A + E$. The initial state of the system is

$$|\Psi(0)\rangle = \Big[\sum_i c_i|\xi_i\rangle \Big] \otimes |\Phi_0(0)\rangle \ . \tag{55}$$

Here, $|\Phi_0\rangle$ is the initial state of the measuring device plus the environment from which one deduces the reduced

density matrix of A by tracing out the environment:

$$\hat{\rho}_A(0) = \text{Tr}_E |\Phi_0(0)\rangle\langle\Phi_0(0)| \ , \tag{56}$$

with eigenstates and eigenvalues

$$\hat{\rho}_A(0)|\psi_a(0)\rangle = p_a(0)|\psi_a(0)\rangle \ . \tag{57}$$

We assume that the measuring device and environment are in equilibrium and so in a highly entangled state. In that case, the emergent classical view is ignorant of the exact ontic state of A. At equilibrium, the stochastic process described in section II D is ergodic and it is the macroscopic time average of this process that corresponds to what we think of as the initial classical state of the measuring device and this is captured by performing the ensemble average over $\hat{\rho}_A(0)$.

Evolving forward in time, and including measuring inefficiencies, we expect to have a state of the form

$$|\Psi(t)\rangle = \sum_{ij} f_{ij}(t)|\xi_i\rangle \otimes |\Phi_j(t)\rangle \ . \tag{58}$$

After the decoherence time $t > \tau$ and so certainly at the end of the measurement $t = T$, the states of $A + B$ are macroscopically distinct:

$$\big|\langle\Phi_i(T)|\Phi_j(T)\rangle\big| \sim \Delta \ , \qquad \text{for } i \neq j \tag{59}$$

and $f_{ij}(t) \rightarrow f_{ij}^{(0)}$, a constant matrix. Note that unitarity requires that the matrix $f_{ij}^{(0)}$ satisfies

$$\sum_{ij} \big|f_{ij}^{(0)}\big|^2 = 1 + \mathcal{O}(\Delta) \ . \tag{60}$$

The reduced density matrix of the measuring device plus environment $A + E$ takes the form

$$\hat{\rho}_{A+E}(T) = \sum_{ijk} f_{ij}^{(0)} f_{ik}^{(0)*} |\Phi_j\rangle\langle\Phi_k| \ , \tag{61}$$

and so the ontic states of $A + E$ are

$$\sum_j Z_j^{(i)}|\Phi_j(T)\rangle + \mathcal{O}(\Delta) \ , \tag{62}$$

where the $Z_j^{(i)}$ are eigenvectors of the matrix with elements:

$$M_{jk} = \sum_i f_{ij}^{(0)} f_{ik}^{(0)*} \ ,$$
$$\sum_j M_{jk} Z_k^{(i)} = \lambda^{(i)} Z_j^{(i)} \ . \tag{63}$$

So, unfortunately, the ontic states involve macroscopic superpositions of states of $A + E$ of the kind that we saw in the simple measurement model in the last section.

However, if we examine the ontic states of A, which is the physically relevant sub-system, a more satisfactory

picture emerges. Since the states of the environment are to a high degree orthogonal for $i \neq j$, the reduced density matrix for A takes the form

$$\hat{\rho}_A(T) = \sum_{ij} \left| f_{ij}^{(0)} \right|^2 \hat{\rho}_A^{(j)}(T) + \mathcal{O}(\Delta) , \quad (64)$$

where we have defined the component reduced density matrices, one for each of the outcomes,

$$\hat{\rho}_A^{(j)}(T) = \text{Tr}_E \, |\Phi_j(T)\rangle\langle\Phi_j(T)| . \quad (65)$$

Since the states $|\Phi_i(T)\rangle$ have such a small inner product (59), the component density matrices commute to high degree of accuracy

$$[\hat{\rho}_A^{(i)}(T), \hat{\rho}_A^{(j)}(T)] = \mathcal{O}(\Delta) . \quad (66)$$

The ontic states $|\psi_a(T)\rangle$, the eigenvectors of $\hat{\rho}_A(T)$, will, therefore, split up into d_P mutually ergodically inaccessible sets $\mathcal{E}^{(i)}$, as well as a set $\mathcal{E}_0$ of approximately null eigenvectors that play no important role. The component density matrices take the form

$$\hat{\rho}_A^{(i)}(T) = \sum_{a \in \mathcal{E}^{(i)}} p_a^{(i)} |\psi_a(T)\rangle\langle\psi_a(T)| + \mathcal{O}(\Delta) . \quad (67)$$

Therefore the ontic state $|\psi_a(T)\rangle$, $a \in \mathcal{E}^{(j)}$, has $\hat{\rho}_A(T)$ eigenvalue

$$p_a(T) = \sum_i \left| f_{ij}^{(0)} \right|^2 p_a^{(j)} + \mathcal{O}(\Delta) . \quad (68)$$

It does not involve a macroscopic superposition—at least to order Δ—and so compared to (54), one can see that the environment has the effect of yielding an acceptable phenomenology.

Under the stochastic process, even for macroscopic time scales, there is only a minute probability for $|\psi_a(T)\rangle$ to make a transition into another state $|\psi_b(T)\rangle$ in a different ergodic set $a \in \mathcal{E}^{(i)}$ and $b \in \mathcal{E}^{(j)}$, for $i \neq j$. For fixed j, the group of ontic states $|\psi_a(T)\rangle$, $a \in \mathcal{E}^{(j)}$, are not macroscopically distinct, and so ontic states can effectively only make transitions within this set. We can, in principle, use the stochastic process to work out the probability $p_{a|b}(T, 0)$ that an initial ontic state $|\psi_b(0)\rangle$ evolves to $|\psi_a(T)\rangle$ using (14). However, according to our discussion below equation (57), since the measuring device is assumed to be in equilibrium with the environment its ontic states are undergoing constant transitions. The time average of this process is captured by taking the ensemble average over $\hat{\rho}_A(0)$. With this initial averaging, we can use (14) to compute the probabilities that the system is finally in the ontic state $|\psi_a(T)\rangle$:

$$\sum_b p_{a|b}(T, 0) p_b(0) = p_a(T) . \quad (69)$$

The result is simply the single-time probability in (68). Turning to the final state, as far as the emergent classical ontology is concerned, the ontic states in an ergodic set $\mathcal{E}^{(j)}$ cannot be distinguished and are also undergoing constant transitions between themselves. Therefore it is only meaningful to compute the inclusive probability that the system ends up in the jth set:

$$p^{(j)}(T) = \sum_{a \in \mathcal{E}^{(j)}} p_a(T) = \sum_i \left| f_{ij}^{(0)} \right|^2 + \mathcal{O}(\Delta) , \quad (70)$$

where we used the fact that $\text{Tr}\,\hat{\rho}_A^{(j)} = 1$ so that $\sum_{a \in \mathcal{E}^{(j)}} p_a^{(j)} = 1 + \mathcal{O}(\Delta)$. It is particularly noteworthy that the final result here is actually independent of the detailed form of the microscopic stochastic process that we introduced in section II D because we averaged over the initial ontic state and the calculated an inclusive probability for a sub-ensemble in the final state. The non-trivial role of the stochastic process is that it was defined in such a way so as to lead to ergodicity breaking. Note that for a perfect measuring device $f_{ij}(t) = c_i \delta_{ij}$ and so $p^{(j)}(T) = |c_j|^2$ which is just the Born rule.

Notice that, although the individual probabilities $p_a(t)$ can depend on time as a result of the interaction of A with the environment, the sum over the group is independent of time to order Δ, assuming, as we have, that the measuring device settles down and ceases to make errors for $t > T$. It is clear that the more realistic model solves problems (i) and (iii). Firstly, (iii) is solved because the ontic states $|\psi_a(T)\rangle$ do not involve macroscopic superpositions. In addition, (i) is solved because the near degeneracy $|c_i| \approx |c_j|$ no longer leads to macroscopic superpositions or artificially extended de-coherence times.

Now we turn to problem (ii). The reduced density matrix of $A + P$ is

$$\hat{\rho}_{A+P}(T) = \sum_{ij} \left| f_{ij}^{(0)} \right|^2 |\Xi^{(j)}\rangle\langle\Xi^{(j)}| \otimes \hat{\rho}_A^{(j)}(T) + \mathcal{O}(\Delta) , \quad (71)$$

where

$$|\Xi^{(j)}\rangle = \left[\sum_i \left| f_{ij}^{(0)} \right|^2 \right]^{-1/2} \sum_i f_{ij}^{(0)} |\xi_i\rangle \quad (72)$$

and so the ontic states are

$$|\Xi^{(j)}\rangle \otimes |\psi_a(T)\rangle + \mathcal{O}(\Delta) , \quad a \in \mathcal{E}^{(j)} . \quad (73)$$

The probabilities are identical to (68). Note that the ontic states (73) are those of A not entangled with states of P. Hence, focussing on $A + P$ rather than A does not lead to the disastrous change in the ontology seen in the simple model. However, unless the measuring device is perfect, the states $|\Xi^{(j)}\rangle$ are not the ontic states of P, they are not even orthogonal.

C. Collapse of the Wave Function

After the measurement has been completed, the stochastic process ensures that there are effectively no

transitions between the states in different ergodic sets. As far as A is concerned, if the state $|\psi_a(T)\rangle$ has $a \in \mathcal{E}^{(j)}$ then for all practical purposes for calculating future dynamics, it would be prudent, but not necessary, to remove the other terms $i \neq j$ from A's reduced density matrix (64) leading to the replacement

$$\hat{\rho}_A(T) \rightsquigarrow \hat{\rho}_A^{(j)}(T) . \qquad (74)$$

So the collapse of the wave function in the new interpretation is just the innocuous process of removing terms which could only have an effect of order Δ in the future. As we have repeatedly emphasised such effects are many orders of magnitude smaller than other kinds of systematic errors that are inherent in the model. So if we condition the future of a system on its present state—the ergodic sub-ensemble of ontic states at a given time—then collapse of the wave function is a harmless procedure of removing terms that are ergodically inaccessible.

D. Measurement of Continuous Quantum Systems

Measurements on microscopic systems with continuous eigenvectors, like the position of a particle, have caused problems for modal interpretations [37–39]. In this case one might expect a generalization of (46) to be

$$|\Psi(t)\rangle = \int dx \, \xi(x)|x\rangle \otimes |\psi_x(t)\rangle . \qquad (75)$$

Here, $|\psi_x(t)\rangle$ is a macroscopic state of A indicating that the particle is at x. The reduced density matrix of the measuring device is then

$$\hat{\rho}_A(t) = \int dx \, |\xi(x)|^2 |\psi_x(t)\rangle\langle\psi_x(t)| . \qquad (76)$$

The potential problem is that even though the states $|\psi_x(t)\rangle$ are expected to become approximately orthogonal, for instance, one expects

$$\big|\langle\psi_x(T)|\psi_y(T)\rangle\big| \sim \exp\big[-\lambda N(x-y)^2\big] , \qquad (77)$$

the eigenvectors of $\hat{\rho}_A(T)$, contrary to naïve expectations, are not localized in x. In fact, if the original wave function of the particle $\xi(x)$ is spread out over a range δx then the ontic states of the measuring device involve a spread of the states $|\psi_x(t)\rangle$ over the same range. In other words, the ontic states of A would involve macroscopic superpositions. This is obviously a potential disaster for any modal interpretation.

The loop-hole in this thinking was identified in [1]. In fact we should be on alert for any argument, like that above, that relies on the smoothness of wave functions down to be arbitrarily small scales as this is not realistic and in the spirit of effective theory. In fact non-relativistic quantum mechanics based on the one-particle truncation of the multi-particle Hilbert space

breaks down on length scales $\hbar/mc$. So any attempt to measure the position of the particle down to these scale will inevitably involve particle creation and annihilation. But in a more practical sense, any realistic description would acknowledge the fact that A would have some inherent finite resolution scale. A simple way to model this is to imagine that the states of A respond to the particle's position in finite bins $[x_j, x_{j+1}]$, where $x_j = x_0 + j\epsilon$, where ϵ is the resolution scale. As along as $\epsilon \gg \hbar/mc$ the description of the measurement process within non-relativistic quantum mechanics will be valid. In that case (75) should be replaced by

$$|\Psi(t)\rangle = \sum_j \int_{x_j}^{x_{j+1}} dx \, \xi(x)|x\rangle \otimes |\psi_j(t)\rangle . \qquad (78)$$

After the measurement, the ontic states of A are approximately one of the discrete macroscopically distinct states $|\psi_j(t)\rangle$ and so a satisfactory phenomenology without macroscopic superpositions ensues.

In [1] a similar issue was shown to arise in an experiment involving the monitoring of a decaying system. In this case if the measuring device is taken to have infinite temporal resolution then superpositions of macroscopically distinct states arise. However, once the finite temporal resolution scale of a realistic measuring device is taken into account the problem with macroscopic superpositions evaporates.

V. EPR-BOHM AND BELL

In this section, we discuss the new interpretation in the context of Bohm's classic thought experiment [16, 40], based originally on the classic paper by Einstein, Podolsky, and Rosen [41], and the implications for Bell's theorem [42].

It is useful for simplicity to avoid introducing a separate environment. Our experience from section IV A suggests that we may do this as long as (i) we avoid degeneracies in the state of microscopic system (ii) have perfect measuring devices. The point is that the microsystem acts as a surrogate environment for the measuring devices and the ontic states of the measuring devices act as proxies for the ergodic sets of ontic states of the more realistic situation with a large environment.

A. The Thought Experiment

The EPR-Bohm set-up begins with a pair of qubits initially prepared in the entangled state that we take to be

$$|\Phi\rangle = c_+|z^+z^-\rangle + c_-|z^-z^+\rangle . \qquad (79)$$

Note that we assume that $c_\pm$ are generic to avoid degeneracies. The ontic states of 1 or 2 are each one of the pair $|z^\pm\rangle$ with probabilities $|c_\pm|^2$, or vice-versa, respectively.
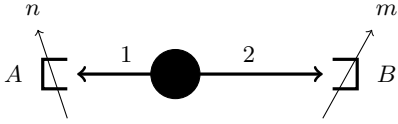
FIG. 6. The EPR-Bohm thought experiment. Two qubits in the entangled state $|\Phi\rangle$ are produced at the source and then recoil back-to-back towards 2 qubit detectors A and B designed to measure the component of the spin along directions n and m, respectively. In our set up, we choose an inertial frame for which the interaction between A and 1 happens before B and 2.

The ontic state of $1 + 2$ is uniquely $|\Phi\rangle$. Given our discussion in section III A, we can say that the systems 1, 2 and $1 + 2$ are related in a "quantum" way and there is no notion of a joint property assignment or associated probability for the ontic states of 1 and 2.

In order to perform measurements, we add two spatially-separated spin detectors A and B, where A detects the spin of particle 1 and B detects the spin of particle 2 as in figure 6. The complete system consists of qubits 1 and 2, measuring devices A and B and the environment, although as mentioned above in the present context we will ignore the environment in this simplified analysis. It is important that even though there is no genuine environment, the ontic state of $A + B$ is always related to A and B in a classical way, that is as a tensor product of the ontic states of A and B as in (35) but in this simplified model exactly. So the description from the point-of-view of A, B and $A + B$ can be integrated into a consistent global description throughout the experiment. Another important point is that the mirror ontic states to those of $A + B$ are the ontic states of $1 + 2$ themselves so that the ontic dynamics of $A + B$ depends implicitly on $1 + 2$ as is clear from the matrix elements in (25).

Let us analyse what happens if A is set to measure the spin component in the n direction, at an angle θ to the z-axis in the (x, z) plane, and B is set to measure the spin in a direction m at an angle ϕ to the z-axis in the (x, z) plane.[15] This set-up is generic enough for our purposes.

The initial state of the overall system is

$$|\Psi(t_1)\rangle = |A_0 B_0\rangle \otimes |\Phi\rangle . \qquad (80)$$

Suppose, in a certain inertial frame, the interaction between A and 1 happens first at time $t_A > t_1$. After a short decoherence time, the state of 1 becomes entangled with A and the state of the total system becomes for $t_2 > t_A$

$$|\Psi(t_2)\rangle = |A_+ B_0\rangle \otimes |n^+ \psi^+\rangle + |A_- B_0\rangle \otimes |n^- \psi^-\rangle , \ (81)$$

---

[15] The eigenstates are $|n^\pm\rangle = \cos(\theta/2)|z^\pm\rangle \pm \sin(\theta/2)|z^\mp\rangle$ and $|m^\pm\rangle = \cos(\phi/2)|z^\pm\rangle \pm \sin(\phi/2)|z^\mp\rangle$.

where

$$|\psi^\pm\rangle = c_\pm \cos(\theta/2)|z^\mp\rangle \pm c_\mp \sin(\theta/2)|z^\pm\rangle . \qquad (82)$$

Note that these states are neither normalised or orthogonal.

Assuming in this simple model that the states $|A_\pm\rangle$ are exactly orthogonal, the ontic state of $1 + 2$ has changed from $|\Phi\rangle$ to either of the non-entangled (and non-normalized states) $|n^\pm \psi^\pm\rangle$ with probabilities $\langle \psi^\pm | \psi^\pm \rangle$. The ontic state of 1 is now one of the pair $|n^\pm\rangle$ and these are now perfectly matched with the ontic states of $1 + 2$.

The ontic states of $1 + 2$, after the local interaction between A and 1, are, of course, the analogues of the collapsed wave functions of the Copenhagen interpretation. So the puzzle seems to be that a local interaction between A and 1 has changed the ontic state of $1 + 2$ and this seems to have led to a non-local change of the ontic state of qubit 2. But it turns out this is not correct. The key point point is that the interaction between A and 1 changes the ontic states of 1 and $1 + 2$ but not 2. To see this, the reduced density matrix of 2 before the interaction between A and 1 is

$$\hat{\rho}_2(t_1) = |c_+|^2 |z^-\rangle\langle z^-| + |c_-|^2 |z^+\rangle\langle z^+| , \qquad (83)$$

while after the interaction,

$$\hat{\rho}_2(t_2) = |\psi^+\rangle\langle \psi^+| + |\psi^-\rangle\langle \psi^-| . \qquad (84)$$

But the states $|\psi^\pm\rangle$ are neither normalized or orthogonal and using (82) one can show that actually

$$\hat{\rho}_2(t_1) \equiv \hat{\rho}_2(t_2) , \qquad (85)$$

reflecting the fact that ontic assignments are local and an interaction between A and 1 cannot affect the ontic state of the causally separated qubit 2. The implication is that the ontic state of qubit 2 remains one of the pair $|z^\pm\rangle$ throughout the interaction between A and qubit 1. To be clear, the interaction between A and 1 does not change the possible ontic states $|z^\pm\rangle$ of 2 but also the actual ontic state of 2. The latter fact follows from the stochastic process describing the ontic states of 2 given that the interaction Hamiltonian $\hat{H}_{\text{int}}$ is itself local, that is of the form of a sum of two local interactions:
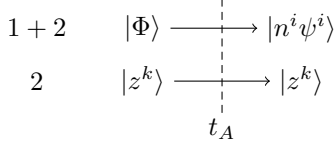
$$\hat{H}_{\text{int}} = \hat{H}_{\text{int}}^{A,1} + \hat{H}_{\text{int}}^{B,2} , \qquad (86)$$

where the first (second) term is non-vanishing for t in the neighbourhood of t_A (t_B). This implies

$$p_{z^\pm | z^\pm}(t_2, t_1) = 1 . \qquad (87)$$

So although the interaction between A and 1 changes the ontic state of $1 + 2$ from $|\Phi\rangle$ to one of the pair $|n^\pm \psi^\pm\rangle$ and this seems to have fundamentally changed the state of 2, however, the ontic states of $1 + 2$ cannot be broken down in terms of the ontic states of its sub-systems 1

and 2—at least for $t < t_B$. Consequently, there is an apparent duality for qubit 2 from the point-of-view of $1 + 2$ compared with 2:

$$
\begin{array}{ccc}
1+2 & |\Phi\rangle \longrightarrow & |n^i \psi^i\rangle \\
2 & |z^k\rangle \longrightarrow & |z^k\rangle
\end{array}
$$
$$
t_A
$$

It is tempting to say that the behaviour of the state of $1 + 2$ here is non-local. However, this is potentially misleading because in order to talk about locality within the system $1 + 2$ requires us to break $1 + 2$ down in terms of its sub-systems 1 and 2. But, to reiterate, we cannot break down the states $|n^i \psi^i\rangle$ ontically in terms of 1 and 2 because the descriptions via $1 + 2$ and 2 cannot be integrated into a consistent whole. We will have more to say about this and its implications in section V B.

To emphasize, for $t < t_B$, the ontic state of $1 + 2$ is not a tensor product of ontic states of 1 and 2 so there is no sense in which the ontology is "classical" in the sense described in section III A. So in figure 7, which summarizes the ontic dynamics of 1, 2 and $1 + 2$, we label it as "quantum".
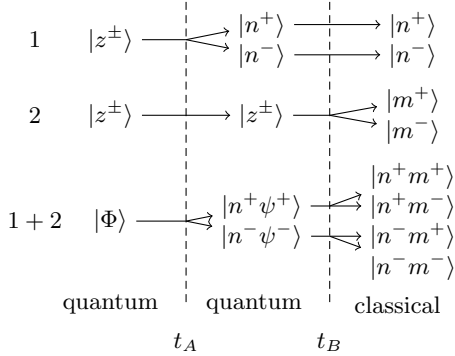

FIG. 7. Snapshots of the ontic states of 1, 2 and $1+2$ at times $t_1 < t_A < t_2 < t_B < t_3$ in the given inertial frame. For $t = t_1$ and t_2, the ontic states of $1+2$ are not tensor products of the ontic states of 1 and 2 and so the ontology is "quantum". In the last time step $t = t_3$ the relation between states becomes "classical" in the sense that the ontic states of $1+2$ are tensor products of those of 1 and 2.

On the other hand the behaviour of $A + B$, A and B is always classical as illustrated in figure 8. The initial ontic state of $A + B$ is $|A_0 B_0\rangle$ while after the measurement at A the ontic state is then one of the pair $|A_\pm B_0\rangle$ with probabilities

$$
p_{A_\pm B_0 | A_0 B_0}(t_2, t_1) = \langle \psi^\pm | \psi^\pm \rangle , \tag{88}
$$

for $t_1 < t_A < t_2$. These follow from applying (5) along with the uniqueness of the initial ontic state.

Finally, at time t_B in this frame, B interacts with qubit 2 and after a further short decoherence time, the state becomes

$$
|\Psi(t_3)\rangle = \sum_{ij=\pm} |A_i B_j\rangle \otimes |n^i m^j\rangle \langle m^j | \psi^i \rangle . \tag{89}
$$

After this interaction the triplet of systems 1, 2 and $1+2$ now has "classical" ontology in the sense that the ontic states of $1 + 2$ are one of the quartet $|n^i m^j\rangle$ and these are tensor products of the ontic states of 1 and 2. For the measuring devices, the final ontic states are one of the quartet $|A_i B_j\rangle$. It is important is that (i) the first interaction between A and 1 does not change the ontic state of B and (ii) the second interaction between B and 2 does not change the ontic state of A. This is guaranteed by the locality of the interaction Hamiltonian (86):

$$
p_{B_0 | B_0}(t_2, t_1) = 1 , \qquad p_{A_i | A_j}(t_3, t_2) = \delta_{ij} . \tag{90}
$$

The implication is that the ontic dynamics has the tree-like structure as shown in figure 8. Therefore, once again using (5), the only non-vanishing probabilities are

$$
p_{A_i B_j | A_i B_0}(t_3, t_2) = \frac{\left| \langle m^j | \psi^i \rangle \right|^2}{\langle \psi^i | \psi^i \rangle} , \tag{91}
$$

for $t_2 < t_B < t_3$. The final probabilities follow from (5), or by composing (88) and (91),

$$
p_{A_i B_j | A_0 B_0}(t_3, t_1) = \left| \langle m^j | \psi^i \rangle \right|^2 . \tag{92}
$$

As we have emphasised, the ontic states of A, B and $A + B$ are just tensor products of those of A and B for all t. In the more realistic model with an environment, this relation will be emergent as in (35). Given the classical ontology, it is meaningful to define joint probabilities as in (36):

$$
p(A_i B_j) \stackrel{\text{emergent}}{=} p_{A_i B_j | A_0 B_0}(t_3, t_1) . \tag{93}
$$

These emergent probabilities are exactly what would have been predicted on the basis of Born's rule. So during the experiment the entanglement between 1 and 2 is converted into an emergent classical correlation between A and B. But what is crucial is that this correlation arises after the *local* interaction between 1 and A and separately between 2 and B and not through any non-local interaction between A and B.

B. Bell's Theorem

Now we turn to a discussion of Bell's theorem [31, 42]. The first point is that, since the predictions agree with the conventional analysis of the Copenhagen interpretation and the Born rule, it must be that the new interpretation implies a violation of the Bell inequality.

In fact it is easy to see that the new interpretation violates *outcome independence*:
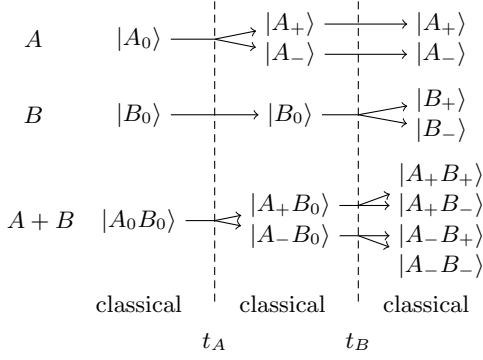
$$
p(B_j | A_i, n, m) = p(B_j | n, m) . \tag{94}
$$

FIG. 8. Snapshots of the ontic states of the measuring devices A, B and $A+B$ at t_1, t_2 and t_3. The relation between the three systems A, B and $A+B$ is always classical, as the ontic states of $A+B$ are just tensor products of the ontic states of A with those of B. However, for $t = t_3$, the measuring devices are, nevertheless, correlated in a classical sense because $\hat{\rho}_{A+B} \neq \hat{\rho}_A \otimes \hat{\rho}_B$.

To see this note that the conditional probability above can only be defined in the context of the ontic dynamics of sub-system $A + B$; indeed,

$$p(B_j|A_i, n, m) \equiv p_{A_iB_j|A_iB_0}(t_3, t_2)$$
$$= \frac{\left|\langle m^j|\psi^i\rangle\right|^2}{\langle\psi^i|\psi^i\rangle} \ . \tag{95}$$

We can now pinpoint where the dependence on the state of A arises in the interaction between B and 2. Although the sub-system $A + B$ can be broken down in terms of its component parts A and B because their ontology is aways "classical", the ontic dynamics of $A + B$ depends on the mirror ontic state of $1+2$, and $1+2$ does not have a "classical" ontology for $t < t_B$. The relation between the ontic states of $A+B$ and the associated mirror ontic states of $1 + 2$ are as follows:
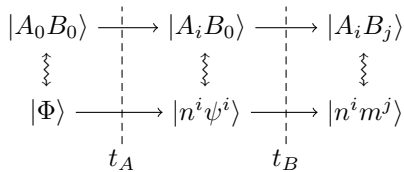
It is the entanglement of the mirror ontic state $|\Phi\rangle$ that leads, after the first interaction at $t = t_A$, to the mirror ontic states $|n^i\psi^i\rangle$. Including only the important tensor product factors, we have

It is important to remember that $|\psi^i\rangle$ are not the ontic states of qubit 2 even though they are tensor product

factors of the ontic states of $1 + 2$ for $t = t_2$. This in turn implies that when B interacts with qubit 2 at t_B, the ontic dynamics becomes implicitly dependent on the ontic state of A through the mirror ontic state $|n^i\psi^i\rangle$. In this interaction, locality ensures that only the tensor product factors of B and qubit 2 are relevant but the state of A influences the dynamics via this initial tensor product factor of qubit 2, that is $|\psi^i\rangle$, and this leads to the A_i dependent probabilities (95):

Note that during the interaction between B and qubit 2 the state of causally separated $|A_i\rangle$ is inert but is included in the above to see the correlation with the states $|\psi^i\rangle$.

On the other hand, the probabilities satisfy *parameter independence*:

$$p(A_\pm|n, m) = p(A_\pm|n) \ , \tag{96}$$

since

$$p(A_\pm|m, n) \equiv \sum_{j=\pm} p_{A_\pm B_j|A_0 B_0}(t_3, t_1)$$
$$= |c_\pm|^2 \cos^2(\theta/2) + |c_\mp|^2 \sin^2(\theta/2) \ , \tag{97}$$

independent of m, i.e. ϕ. This latter condition expresses the fact that what happens at A does not depend on what is measured at B. The violation of outcome independence but observance of parameter independence is just as in the Copenhagen interpretation and a violation of either implies a violation of the Bell inequality. So what we have shown is that the new interpretation violates Bell's inequality for the same reason the Copenhagen interpretation does because the initial state of the qubits $|\Phi\rangle$ is entangled. The novelty is that it does this without invoking the collapse of the wave function. Instead what happens is a curious and apparently non-local change in the ontic state of $1+2$ from $|\Phi\rangle$ to $|n^i\psi^i\rangle$ when A and 1 interact. On closer inspection, though, it is hard to say this is a non-local process because it does not make sense to analyse the states $|n^i\psi^i\rangle$ in terms of the ontic states of their constituent qubits. So this is the "spooky action at a distance" of quantum mechanics laid bare. In any event there is no violation of causality.

Another point to make is that the assignment of ontic states clearly depends on the inertial frame chosen to view the experiment. For instance in our chosen frame, where the measurement is made at A before B, the ontic states of $A + B$ after the first measurement are $|A_\pm B_0\rangle$. But in another frame, where the measurement is made at B before A, the ontic states of $A + B$ after the first measurement are $|A_0 B_\pm\rangle$. This, of course, is no surprise, but modal interpretations have been claimed to be inconsistent with special relativity [44]. Fortunately these

claims have shown to rest on the false assumption that the ontic states of a sub-system are global property assignments and so are unwarranted [12].

VI. DISCUSSION

The emergent Copenhagen interpretation is a completely self-contained interpretation of the quantum mechanics that builds on and modifies earlier proposals that are collectively known as modal interpretations. The key ingredient is the fact that one cannot isolate a quantum system from its environment. The interaction between the two allows one to define the dual notion of the epistemic and ontic state of the sub-system. The former evolves according to the Schrödinger equation (for the total system), whilst the latter evolves according to a stochastic process. This dualism is analogous to the ensemble and micro-state of classical statistical mechanics and allows for the solution of the measurement problem by invoking an ergodicity argument familiar from the discussion of a phase transition. Although we have presented the quantum interpretation as an analogue of classical statistical mechanics, in fact the former should be taken as a proper definition of the latter. In this regard, we are taking the viewpoint of [23–30] but reintroducing the notion of a micro-state in the form of the ontic state. However, unlike the situation of micro-states and ergodicity in classical statistical mechanics, the dynamics of ontic states follows from a very simple Markov process.

The most striking feature of the interpretation is that it reproduces the phenomenology of the Copenhagen interpretation for macro-systems: definite measurement outcomes occur with a probability determined by the Born rule and the collapse of the wave function is just an innocuous process of removing ergodically inaccessible parts of the epistemic state. However, there is no sharp Heisenberg cut between microscopic systems and macroscopic systems. One can quantify the degree of classicality as the time required to see a transition between two hypothetically macroscopically distinct states; as order τ/Δ, where Δ is the magnitude of the inner product of the two states. A shortened description of new interpretation will appear in [45].

Finally, we can put Schrödinger's cat out of it misery. According to the new interpretation, it is either alive or dead but these states are not pure states rather they are associated to ergodic subsets of ontic states that are, for all physically relevant time scales, mutually inaccessable. So the cat may be either alive or dead but in either case in equilibrium with—hence strongly entangled with—the environment.

[1] T. J. Hollowood, "The Copenhagen Interpretation as an Emergent Phenomenon," J. Phys. A **46**, 325302 (2013) [arXiv:1302.4228 [quant-ph]].

[2] H. P. Krips, "Two Paradoxes in Quantum Mechanics," Phil. Sci. **36**, 145 (1969).

[3] H. P. Krips. "Statistical interpretation of quantum theory," Am. J. Phys. **43**, 420 (1975).

[4] H. P. Krips. "The Metaphysics of Quantum Theory," Clarendon Press Oxford, 1987.

[5] B. C. Van Fraassen, "A Formal Approach to the Philosophy of Science," In: R. Colodny, editor, *Paradigms and Paradoxes: The Philosophical Challenge of the Quantum Domain*, pages 303–366. University of Pittsburgh Press, 1972.

[6] N. D. Cartwright, "Van Fraassen's Modal Model of Quantum Mechanics," Phil. Sci. **41**, 199 (1974).

[7] B. C. Van Fraassen, "Quantum Mechanics: An Empiristicist View," Oxford University Press, 1991.

[8] J. Bub, "Quantum mechanics without the projection postulate," Found. Phys. **22**, 737 (1992).

[9] P. E. Vermaas and D. Dieks, "The Modal Interpretation of Quantum Mechanics and Its Generalization to Density Operators," Found. Phys. **25**, 145 (1995).

[10] G. Bacciagaluppi and W. M. Dickson, "Dynamics for Modal Interpretations," Found. Phys. **29**, 1165 (1999).

[11] P. E. Vermaas, "A Philosopher's Understanding of Quantum Mechanics: Possibilities and Impossibilities of a Modal Interpretation," Cambridge University Press, 1999.

[12] J. Berkovitz and M. Hemmo, "A New Modal Interpretation of Quantum Mechanics in Terms of Relational Properties," In: W. Demopoulos and I. Pitowsky, editors, *Physical Theory and its Interpretation Essays in Honor of Jeffrey Bub*, pages 1–28. The Western Ontario Series in Philosophy of Science, Springer, 2006.

[13] C. Rovelli, "Relational Quantum Mechanics," International Journal of Theoretical Physics **35**, 1637 (1996).

[14] D. Dieks, "Objectivity in Perspective: Relationism in the Interpretation of Quantum Mechanics," Foundations of Physics **39**, 760 (2009).

[15] G. Bene and D. Dieks, "A Perspectival Version of the Modal Interpretation of Quantum Mechanics and the Origin of Macroscopic Behavior," Foundations of Physics **32**, 645 (2002).

[16] D. Bohm, "Quantum Theory," Courier Dover Publications, 1951.

[17] E. Joos and H. D. Zeh, "The emergence of classical properties through interaction with the environment,"

Zeitschrift für Physik B Condensed Matter **59**, 223 (1985).

[18] E. Joos, editor. "Decoherence and the appearance of a classical world in quantum theory," Springer, 2003.

[19] W. H. Zurek, "Decoherence and the transition from quantum to classical—revisted," quant-ph/0306072.

[20] M. Schlosshauer, "Decoherence, the measurement problem, and interpretations of quantum mechanics," Rev. Mod. Phys. **76**, 1267 (2005).

[21] M. Schlosshauer and K. Camilleri, "The quantum-to-classical transition: Bohr's doctrine of classical concepts, emergent classicality, and decoherence," arXiv:0804.1609.

[22] H. P. Breuer and F. Petruccione, "The Theory of Open Quantum Systems," Oxford University Press, 2002.

[23] S. Popescu, A. J. Short, and A. Winter, "The foundations of statistical mechanics from entanglement: Individual states vs. averages," arXiv quant-ph/0511225.

[24] S. Popescu, A. J. Short, and A. Winter, "Entanglement and the foundations of statistical mechanics," Nature Physics **2**, 754 (2006).

[25] P. Bocchieri and A. Loinger, "Ergodic foundation of quantum statistical mechanics," Phys. Rev. 114, 948951 (1959).

[26] S. Lloyd, "Black Holes, Demons and the Loss of Coherence," Thesis Rockefeller Univ. (1988).

[27] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, "Canonical Typicality," Phys. Rev. Lett. **96**, 050403 (2006).

[28] N. Linden, S. Popescu, A. J. Short, and A. Winter, "Quantum mechanical evolution towards thermal equilibrium," Phys. Rev. E **79**, 061103 (2009) [arXiv:0812.2385 [quant-ph]].

[29] A. J. Short, "Equilibration of quantum systems and subsystems," New Journal of Physics **13**, 053009 (2011).

[30] J. Gemmer, M. Michel, and G. Mahler, "Quantum Thermodynamics: Emergence of Thermodynamic Behavior within Composite Quantum Systems," Berlin: Springer, 2004.

[31] J. S. Bell, "Speakable and Unspeakable in Quantum Mechanics: Collected papers on quantum philosophy," Cambridge University Press, 2004.

[32] A. Sudbery, "Diese Verdammte Quantenspringerei," Stud. Hist. Phil. Mod. Phys. **33**, 387 (2002).

[33] G. Bene, "Quantum reference systems: A New framework for quantum mechanics," Physica A **242**, 529 (1997) [arXiv:quant-ph/9703021].

[34] A. Albrecht, "Following a "collapsing" wave function," Phys. Rev. D **48**, 3768 (1993).

[35] D. Albert and B. Loewer, "Wanted dead or alive: two attempts to solve Schrödinger's paradox," In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, pages 277–285, JSTOR, 1990.

[36] G. Bacciagaluppi and M. Hemmo, "Modal interpretations, decoherence and measurements," Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics **27**, 239 (1996).

[37] G. Bacciagaluppi, M. J. Donald, and P. E. Vermaas, "Continuity and discontinuity of definite properties in the modal interpretation," Helvetica Physica Acta **68**, 679 (1995).

[38] M. J. Donald, "Discontinuity and continuity of definite properties in the modal interpretation," In *"The Modal Interpretation of Quantum Mechanics"*, pages 213–222. Springer, 1998.

[39] D. N. Page, "Quantum Uncertainties in the Schmidt Basis Given by Decoherence," arXiv:1108.2709 [hep-th].

[40] D. Bohm and Y. Aharonov, "Discussion of Experimental Proof for the Paradox of Einstein, Rosen, and Podolsky," Phys. Rev. **108**, 1070 (1957).

[41] A. Einstein, B. Podolsky, and N. Rosen, "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" Phys. Rev. **47**, 777 (1935).

[42] J. S. Bell, "On the Einstein-Podolsky-Rosen Paradox," Physics **1**, 195 (1964).

[43] A. Bassi and G. Ghirardi, "Dynamical reduction models," Phys. Rept. 379, 257 (2003) [quant-ph/0302164].

[44] W. Myrvold, "Modal interpretations and relativity," Foundations of Physics **32**, 1773 (2202).

[45] T. J. Hollowood, "New Modal Quantum Mechanics." *to appear*.

[46] J. Barandes, *private communication*.