One world is (probably) just as good as many

Jeremy Steeger
jsteeger@uw.edu

Department of Philosophy
University of Washington

November 8, 2021

Abstract

One of our most sophisticated accounts of objective chance in quantum mechanics involves the Deutsch-Wallace theorem, which uses state-space symmetries to justify agents' use of the Born rule when the quantum state is known. But Wallace (2003, 2012) argues that this theorem requires an Everettian approach to measurement. I find that this argument is unsound. I demonstrate a counter-example by applying the Deutsch-Wallace theorem to the de Broglie-Bohm pilot wave theory.

Contents

1 Introduction

Both defenders and critics of Everettian quantum mechanics (EQM) know the so-called "problem of probability" well. Roughly put: Everettians claim that every quantum measurement outcome occurs, so what could it mean for events to be more or less probable? Surely, every event occurs with certainty, and so with probability one! Brown puts it strikingly, imagining that you awake from a dream in which Prof. X managed to toss a quantum coin:

> The next day, relieved in the knowledge that there is at most one Prof. X, you recall the moment in the dream when he claimed that the probability of heads for the biased coin was around 0.7; it was before you were aware of the bizarre consequences of tossing the coin. You now find yourself idly wondering *what Prof. X could have meant.* [...] From the God's-eye perspective, everything that could happen was happening, and there was no uncertainty about the outcome of the tosses. Was Prof. X not talking then about genuine probabilities at all? (2011, p. 9)

On the one hand, this story raises a genuine philosophical puzzle. But on the other, it can easily be used as a sort of incredulous-stare response to EQM. In light of physicist-philosophers making truly wild claims, it is natural to cast about for surface-level signs of contradiction. Probability-talk is an obvious place to start.

Which, unfortunately, puts defenders of EQM on the defensive. On some level, the situation is a bit odd. Everyone encounters similar skeptical hypotheses in childhood: "the sun rises and sets," but the earth actually does the turning; "the planets wander through the stars," but the stars are actually quite a bit further away.[1] We all have gone through the (not unchallenging!) process of resolving surface-level conflicts between theory and the immediate grammar of our experience. It seems distinctly uncharitable not to lend EQMers the same effort. But many do not. And so before EQMers can even get around to extolling the virtues of the view, they have to shore it up against charges of incoherence. No one ought to envy the dreams of EQMers, plagued as they are with bad-faith skeptics.

Upon reflection, the EQMer finds that this situation allows for a tempting rhetorical gestalt: perhaps they could show that the many-worlds interpretation makes sense of probability and makes *more* sense of it than non-branching theories. So perhaps a many-worlds-exclusive derivation of chance will be enough for the EQMer to pull the rug out from under the skeptic. Once put off-balance, they think, the skeptic will come around.

Enter the Deutsch-Wallace theorem, a derivation of chance values (i.e., the Born rule) from a thin, operational definition of objective probability and basic facts about the quantum state space within the EQM framework. This theorem has been much-celebrated within the philosophical literature—and rightly so! It is a lovely result that deepens our understanding of quantum probability. But couched in the above context, one can see the temptation to argue that this theorem *only* holds within the EQM framework.

Perhaps in this spirit, Wallace puts the following words into the mouth of his anti-skeptic:

> Anti-sceptic: We're totally used to probability in Everettian contexts. Okay, it might be *philosophically* a bit puzzling, but those puzzles don't really matter from the point of view of physics: practically speaking, we've got a sufficiently solid grip on probability to do science. In the single-world interpretation, though, we're worried that the whole idea of probability makes no sense at all. Failure to understand probability in a satisfactory

[1] I borrow the latter analogy from Wallace (2012, pp. 427–428).

way in the single-world interpretation isn't just problematic or puzzling: it's fatal. (2012, p. 246)

Wallace, in his own voice, validates the anti-skeptic's claim: "as long as probability talk is understood operationally," he writes, "the Everett interpretation is actually *better off* than non-branching theories in making sense of that talk" (2012, p. 275). In sum, Wallace seems to claim that the Deutsch-Wallace justification of the Born rule holds in a many-worlds approach and not in single-world approaches, and the failure of the latter to achieve anything similar is fatal.

But what if a single-worlder could make *just as good* use of the Deutsch-Wallace theorem? Then, by Wallace's lights, this single-worlder and the EQMer would be on even footing—at least regarding probability.[2]

At this point, it is probably worth stressing that the anti-skeptic's claim is about as dramatic as it gets. Other proponents of EQM defend a weakened version of it. Saunders asserts that "nothing comparable has been achieved for any other physical theory of chance" (2010, p. 184). Read affirms that Deutsch and Wallace's derivation of the Born rule "renders the notion of objective probability less mysterious in EQM than in antecedent physical theories" (2018, p. 5). But I would wager that the weakened form of the claim makes the underlying philosophical question all the more pressing: is it *exclusively* many-worlds that can avail itself of the derivation, or can a single-world approach do so, too?

This paper demonstrates how a proponent of the de Broglie-Bohm pilot wave theory can make excellent use of Wallace's proof of the Born rule. I then argue that the Bohmian's defense of the proof's premises is just as good as the Everettian's. Thus, by Wallace's lights, one world is just as good as many as far as explaining probability goes. So if we wish to find unique advantages for the Everettian, then we ought to look for them elsewhere.

To make this point, it helps to focus on Wallace's discussion of a simple, non-decision-theoretic version of his Born-rule proof, which I will call the symmetry theorem (2012, pp. 148–156).[3] This version uses an operational notion of chance, and Wallace directly asserts that a Bohmian cannot use it. A key premise in this theorem is *state supervenience*, the assumption that chances supervene on the wave function Ψ. Roughly, Wallace suggests that EQMers ought to justify state supervenience by noting that Ψ captures all that exists. He then asserts that particles' configurations in pilot wave theories violate the premise by breaking symmetries in Ψ. But the precise values of configurations q in Bohm's theory are, indeed, "hidden" from agents—at least as a practical limitation on their ability to measure a system (Bohm, 1952a,b). Bohm's theory does not take an agent to observe precise particle configurations q directly, but rather indirectly and approximately via the particle's selection of a particular branch of the overall wavefunction Ψ of the system and the measuring device (Barrett, 2019). Thus, an agent only ever has *approximate* knowledge of q before measurement—an epistemic fact that I call *q-ignorance*. Given that chance is "understood operationally" (as detailed in §2.1), q-ignorance secures state supervenience. The Bohmian is free to adopt Wallace's operational approach to chance, and so they can use the symmetry theorem to derive Born-rule chances. We can, indeed, justify probability in a single world—in fact, Wallace's theorem provides an example of how to do it.

But is this justification *just as good* as the Everettian's? I think so, at least on a charitable approach to many-worlds. On the "Hydra" reading, an agent's talk of some physical system

[2] I am not making any claims about how well these theories deal with locality. In fact, I hope this paper emphasizes locality as a deciding factor between interpretations by demonstrating the relative flexibility of the probability problem.

[3] Something like the symmetry theorem lies at the heart of several other recent derivations of the Born rule, including the self-locating uncertainty approach of Sebens and Carroll (2018) and the "envarience" approach of Zurek (2005, 2009). Both seem to require an Everettian approach. I focus exclusively on Wallace's derivation to provide one clear account of how to disentangle many-worlds assumptions from the symmetry theorem.

corresponds directly to some Ψ. Hydra semantics arguably justify state supervenience on their own. But the Hydra view requires users of quantum theory to accept speaking falsehoods the majority of the time. Thus, Wallace encourages us to adopt the more charitable "Lewisian" view, where system-talk corresponds, instead, to a particular branch of Ψ (2012, Ch. 7). To secure state supervenience for the Lewisian, Wallace appeals to what I call *self-ignorance*—the assumption that an agent cannot reliably know the branch-identity fixing their future path in a branching event (2012, p. 150). But, I argue, this assertion is *only* plausible due to environmental decoherence, the process responsible for generating Ψ's branching structure. In short, one can reasonably claim that decoherence places a limit on an agent's reliable knowledge via a principle that I will call *decoherence exclusivity*. From this principle, self-ignorance follows. But this principle also implies state supervenience *directly*, rendering appeals to further physical or metaphysical details superfluous. Indeed, the Bohmian can use the principle, too, and it provides a good explanation of q-ignorance as a bonus. This argument yields a precise sense in which one world is (probably) just as good as many.

The paper is structured as follows. Section 2 sets the stage by describing the minimal "operational" definition of probability needed for the subsequent results (§2.1). It then motivates the symmetry theorem by noting that it addresses a question that Gleason's theorem fails to answer (§2.2). Section 3 contains my argument, which proceeds as follows. §3.1 introduces the decoherent histories formalism as a framework to house the Hydra, Lewisian, and Bohmian views. §3.2 gives a precise statement of each view. §3.3 describes how each view justifies state supervenience with decoherence exclusivity and argues that self-ignorance and q-ignorance are epistemically on par; §3.4 assesses what goes wrong in Wallace's argument. §3.5 gives a precise statement of the symmetry theorem that applies to all three views. §4 concludes with a preliminary discussion of where the symmetry theorem sits with regards to existing Bohmian approaches to the Born rule—particularly those of Dürr, Goldstein, and Zanghì (1992, henceforth DGZ) and Valentini (2020).

2 Gleason does not give the last word on quantum probability

2.1 Setting the stage: what is an explanation of chance?

It may be useful to start by reviewing just what Deutsch and Wallace's symmetry theorem is doing and why it is so compelling. The symmetry theorem affords one sort of explanation of chance values, i.e., the objective probability values associated with some physical system. This explanation comes in three components: a theory of physical states, a theory of chances, and some link between the two. We intend this link to provide a telling answer to the following sort of why-question: given that the physical state of a system is χ, why should we assign measurement outcomes the chances ch, rather than others? We take a relevant answer to show how ch depends on what χ represents about the system in question.[4]

Our physical theory of interest is non-relativistic quantum mechanics (NRQM). NRQM specifies kinematics, dynamics, and observables for microscopic systems. We begin in the usual way, by associating a unit vector in a Hilbert space $\mathcal{H}_S$ with our system of interest, S. S could include an electron's spin, location, and any other observables that we intend to measure. We represent the preparation and measurement of such systems with projections onto linear subspaces of $\mathcal{H}_S$; let $\mathcal{P}(\mathcal{H}_S)$ denote the set of these projections. Recall that the spectra of the familiar self-adjoint observables are associated with projection-valued measures (PVMs)—and thereby with σ-algebras of projections $\mathcal{E} \subseteq \mathcal{P}(\mathcal{H}_S)$—via the spectral theorem.[5] We allow the system to be open, i.e., possibly

[4] I use van Fraassen's (1980, pp. 134–157) pragmatic model as a framework, but I do not require this model to give the final word on explanation.

[5] Recall that a *projection-valued measure*, for some measurable space (X, Σ) and some subset $\mathcal{E}$ of $\mathcal{P}(\mathcal{H}_S)$, is a map

coupled with its environment, E. E might include air, dust, photons, measurement devices, and so on; associate all these with a Hilbert space $\mathcal{H}_E$. We suppose that the S and E together describe a closed system, i.e., a system that is not subject to any other influences from *its* environment. This total system is described by a unit vector $\Psi \in \mathcal{H}_{SE} = \mathcal{H}_S \otimes \mathcal{H}_E$. Possible trajectories are given by the action of unitary maps U_t on unit vectors of $\mathcal{H}_{SE}$, while the Schrödinger equation,

$$i\hbar \frac{d}{dt}\Psi(t) = H\Psi(t), \tag{1}$$

for some Hamiltonian H acting on the system, picks out which of these trajectories are dynamical (once appropriate initial and boundary conditions are specified).[6] At any time t, the state ρ of the open system S is given by tracing out the environmental degrees of freedom; that is,

$$\rho(t) = \mathrm{Tr}_E |\Psi(t)\rangle\langle\Psi(t)|, \tag{2}$$

where $\rho(t)$ is a density operator on $\mathcal{H}_S$. In the special case that S is not coupled with E, then $\rho(t)$ will equal $|\psi\rangle\langle\psi|$ for some wave function ψ in $\mathcal{H}_S$.

Note that the partial trace specifies the relationship between various physical degrees of freedom at a given time. NRQM also specifies a temporal system-subsystem relationship that I will call the *quantum updating rule*. (This rule sometimes goes by the more familiar name of "the projection postulate"—but as we will soon see, we need not postulate it.) According to this rule, when a total system Ψ yields an outcome P_i at time t (via a "projective measurement"), the system at t is effectively described by $P_i\Psi$ (appropriately normalized). This rule is crucial for ensuring that we can reliably prepare states within a certain range. However, it does not entail anything so strong as "the collapse of the wave function." We will add to this skeleton to recover many-worlds and pilot wave theories that all deny this collapse (see §3.1). As such, we say for now that the system's dynamical state includes *at least* Ψ, which evolves according to equation (1).

My approach to the de Broglie-Bohm completion of NRQM will differ on one key point from the well-known approaches of DGZ (1992) and Valentini (2020). Unlike these authors, but in line with Wallace (2020), I suppose that NRQM is non-cosmological. In other words, I will not demand that some wave function Ψ describes the entire universe, regardless of whether I adopt a Bohmian or an Everettian attitude towards measurement. This choice raises the question of the precise relationship between my approach to the Born rule and those of DGZ and Valentini. In §4, I will briefly sketch how the symmetry theorem might still interest these cosmological Bohmians (although I leave a thorough investigation for future work). Until then, however, the reader should take Ψ to denote the wavefunction of a closed system that need not be the universe.

What about the chances of measurement outcomes for that system? It is widely agreed that states in a formal theory of chance ought to be functions that satisfy some (usually set-theoretic, logical, or algebraic) formalization of Kolmogorov's three axioms—namely, non-negativity, normality, and countable additivity. Various formalizations add surprising complications. The usual set-theoretic approach, for instance, turns out to be too strict for our purposes. So I will stick to a simple and general algebraic approach. Let Σ be a (possibly partial) σ-algebra with top and bottom elements $\top$ and $\bot$.[7] Elements of Σ represent events or utterances that given events occur. A chance function

$E : \Sigma \to \mathcal{E}$ that is non-negative, normalized, and countably additive. For ease of exposition, I will not be treating positive-operator-valued measurements (POVMs) in this paper. But note that we can recover all POVMs as PVMs on closed systems via Neumark's theorem. See Busch et al. (1995) for a comprehensive discussion of these concepts.

[6] Nothing hinges on the choice of the Schrödinger picture, here; the aim is merely to get one well-defined notion of kinematics and dynamics on the table.

[7] Recall that a partial σ-algebra is a partial complemented lattice, i.e., a lattice with partial operations $\bigvee$ and $\bigwedge$ which denote the least upper bound and greatest lower bound of a set of elements, respectively, and the operation $\neg$, which denotes the complement of an element.

$ch : \Sigma \to \mathbb{R}$ from the algebra of events to the real numbers must satisfy

$$ch(e) \geq 0, \tag{3}$$
$$ch(\top) = 1, \text{ and} \tag{4}$$
$$ch\left(\bigvee_i e_i\right) = \sum_i ch(e_i) \text{ when } e_i \leq \neg e_j \text{ for } i \neq j, \tag{5}$$

where the later condition holds only when the argument of the function is defined. Probability theory, like quantum theory, also has a system-subsystem relation—namely, conditional probability, the definition of which allows for modifications to the algebra Σ of events. For instance, suppose we wish to restrict our attention to only those elements of Σ smaller than or equal to e (informally, events that occur given that e occurs); these elements form a subalgebra, $\downarrow e$. Via the usual definition of conditional probability, the state ch on Σ yields the following state on the subsystem $\downarrow e$:

$$ch(f \,|\, e) := \frac{ch(f \wedge e)}{ch(e)}. \tag{6}$$

This way of thinking about conditional probability will prove instructive: to answer our why-question, we will strive to link the system-subsystem relations in both NRQM and probability theory.

Now we can begin to fill in the variables in our why-question: given that the physical state of a closed system is χ, where Ψ is at least a component of χ, why should we assign measurement outcomes $P_i \in \mathcal{P}(\mathcal{H}_{SE})$ the chances given by the Born rule, i.e.,

$$ch_\Psi(P_i) = \langle \Psi, P_i \Psi \rangle, \tag{7}$$

rather than others? As flagged above, a relevant answer shows how ch_Ψ depends on what χ represents. To motivate this dependence—and assess the goodness of an answer—it helps to say a bit more about what ch represents.

Many agree that whatever ch represents, it must account for the functional roles that chance-talk plays in our day-to-day and scientific reasoning. EQMers, in particular, identify two roles that chances must recover:

1. **the inferential link**, i.e., the chance of an event is measured (roughly) by (actual) relative frequencies of that event; and

2. **the credential link**, i.e., all else being equal, one's subjective degree of belief or credence in an event ought to equal the chance of that event,

where we suppose that repeatable processes yield chancy events (making good sense of relative frequencies). Note well that the "ought" in the credential link roughly implies "can." We assume that we can roughly measure chance values and use the results of such measurements to make predictions. One can concoct "chances" that are less accessible to agents—but these would fail to capture the function of "chances" in *scientific* reasoning.

Papineau (1996) introduces these roles under slightly different names, and Saunders (2010), Brown (2011), Wallace (2012), and Read (2018) all endorse them. Wallace assumes that the credential link benefits an agent's pragmatic aims, e.g., their desire to avoid losing money in bets. Thus, he calls it the "decision-theoretic link." However, as Brown (2011) notes, the link with credence need not be spelled out in terms of pragmatic decision-making—it could, instead, be a matter of epistemic (truth-seeking) aims. Additionally, as Saunders (2010) notes, there is another notable role that chance-talk plays in our discourse—namely,

3. **the link with uncertainty**, i.e., chance events, prior to their occurrence, are uncertain.

However, it is not clear that this last link is *essential* to chance-talk. So, assuming that the inferential and credential links capture the essential bits of this talk, we may try to define chance as the thing satisfying them.

Wallace follows this strategy. He formalizes an agent's credences with a function cr satisfying the probability axioms (3)–(6). Then he defines chance in terms of credence using Lewis's (1980) principal principle (PP).

> **Principal principle (PP).** ch is a chance function iff for any event e, if the theory t together with admissible background information b entails that $ch(e) = x$, then an agent ought to set their credence as $cr(e \mid b \wedge t) = x$.[8]

On the operational approach, it is crucial that b only includes information that agents can reliably access. As long as it does, agents can increase their credence in the right theory t by updating. Suppose our agent notes the number of times that e occurs in a large number of trials N all satisfying b. Let e_m represent them seeing e a total of m times. Assuming each e is independent, PP implies that the agent's prior credence $cr(e_m \mid b \wedge t)$ is given by the binomial distribution for $ch(e)$—which is well-approximated by a sharply-peaked Gaussian centered on the chance value. After seeing M occurrences of e, our agent updates using Bayes's theorem

$$cr(t \mid e_M \wedge b) = \frac{cr(e_M \mid b \wedge t)\, cr(t \mid b)}{cr(e_M \mid b)}, \tag{8}$$

which, recall, is a consequence of (6). Thus, their new credence in t will be very low unless t and b imply that $ch(e) \approx M/N$. Relative frequency thereby (roughly) measures chance. And in the idealized case where an agent has full credence in t, they will update to set their credence in e equal to the appropriate chance value given b. So if something plays the role of chance in PP (so defined!), it satisfies both the inferential and credential links. PP (once combined with Bayesian updating) thereby elucidates the standard scientific use of "chance."

This much—equations (3)–(6) and PP—suffices to establish a thin, operational definition of chance. Note well that the non-decision-theoretic version of Deutsch-Wallace that I aim to generalize assumes only this operational definition. I will argue that single-world theorists can use this version of Deutsch-Wallace, too, provided they assume at least this definition (although they may assume more).

The bare operational definition has a clear explanatory deficit. It posits Kolmogorov's rules (3)–(6) by fiat, and it does little to elucidate why these rules have anything to do with chance. Here, decision theory or relative frequencies may enter the story. But nothing in a single-world approach requires a commitment to any sort of frequentism. It just happens to be the most popular single-world theory of chance on hand. So while Wallace (2012, Ch. 4) seeks to motivate the anti-skeptic's view by offering a lengthy critique of frequentism, it is not clear that the single-world theorist needs anything so contentious to motivate Kolmogorov's rules.

I will not address the question of how the Bohmian *should* justify Kolmogorov's rules. One might rightly assert, then, that Wallace's full decision-theoretic derivation for the Everettian offers a deeper explanation than the Bohmian derivation that I present here. Unfortunately, I do not have the

[8] I borrow this statement of Lewis's PP from Wallace (2012, p. 141). PP is a specific formalization of the intuitions described by the credential and inferential links. The links themselves are ambiguous between Lewis's formalization and, e.g., those of Hall (2004) and Ismael (2008). The differences among various formalizations are relevant for chance functions that are self-undermining, i.e., that do not assign $ch(t) = 1$ (Pettigrew, 2012). But on the operational approach, we assume that chances are not self-undermining.

space to discuss an analog of this version for the Bohmian. However, note that Wallace (2012) takes the symmetry theorem to suffice to illustrate his claim that Bohmians cannot use his derivations (in Chapter 4). Prima facie, then, once we remove this barrier to Bohmians' use of the symmetry theorem, they ought to be able to make use of the full decision-theoretic derivation. They ought to be able to use the symmetry theorem with other justifications of Kolmogorov's rules, too—including frequentist ones. Subtleties abound, but I will leave the development of these stronger Bohmian derivations for future work.

So, let us make peace, for the moment, with simply positing Kolmogorov's rules. Many physicists, after all, are happy to do so. But these same physicists often raise a different concern about the Deutsch-Wallace derivation—namely, that Gleason's theorem provides all the explanation that quantum probabilities need. So with this much scaffolding in place, let us assess whether Gleason's theorem answers our why-question.

2.2 Gleason's theorem, measurement neutrality, and the coordination problem

Does Gleason's theorem tell us why we ought to assign chances ch when the physical state of a system includes Ψ? I argue that there are two reasons why the theorem does not give a satisfying answer. First, it *assumes* that chances are independent of the context of measurement—an empirical fact that merits explanation. Second, it faces a problem of coordination: it identifies the right set of chance states, but it does not favor any particular link between Ψ and ch over any other. Thus, the casual frequentist would be mistaken to think that Gleason's theorem renders the Deutsch-Wallace argument otiose.

The thought that it does originates in Barnum, Caves, Finkelstein, Fuchs, and Schack's (2000) infamous response to Deutsch's (1999) original decision-theoretic argument. "*By assuming that measurements are described by probabilities that are consistent with the Hilbert-space structure of the observables, Gleason's theorem derives in one shot the state-space structure of quantum mechanics and the probability rule*," Barnum et al. forcefully claim (2000, p. 1182; emphasis theirs). They are right about the state-space structure but wrong about the probability rule.

Specifically, they are right that Gleason's theorem pins down chance states' structure for open systems, i.e., for systems $\rho = \mathrm{Tr}_E|\Psi\rangle\langle\Psi|$. It is easy to see that the Born rule yields functions that must follow Kolmogorov's rules. In more detail: for a set $\mathcal{E}$ of projections that pairwise commute, define $\bigvee_i P_i$ as the projection onto the closed linear subspace spanned by the ranges of the P_i and define $\bigwedge_i P_i$ as the projection onto the intersection of those ranges, and for each P_i, define $\neg P_i := 1 - P_i$. Then $\mathcal{P}(\mathcal{H}_S)$ is a (partial) σ-algebra, and it is straightforward to check that the Born rule maps states in NRQM to functions that satisfy (3)–(5).[9] But it is a far subtler matter to verify that *only* Born-rule functions satisfy these rules. Gleason (1957) shows us one way to do it.[10]

Gleason's theorem. For $\dim(\mathcal{H}_S) \geq 3$, A function $ch : \mathcal{P}(\mathcal{H}_S) \to \mathbb{R}$ satisfies Kolmogorov's rules if and only if there exists some density operator ρ such that

$$ch(P_i) = \mathrm{Tr}(\rho P_i). \tag{9}$$

[9]Note that while $\mathcal{P}(\mathcal{H}_S)$ is a partial complemented lattice, it is *not* a Hilbert lattice; these latter lattices are complete, but non-intuitive from the standpoint of probability. For more technical details on Hilbert lattices, see Rédei (1998); for a critical perspective, see Kochen (2015).

[10]Note that by taking the domain of the chance function to be the partial σ-algebra $\mathcal{P}(\mathcal{H}_S)$, Gleason's theorem assigns probabilities to all projections on the Hilbert space—and so to all possible measurements, not just to those that can be done simultaneously. In particular, this choice of domain assumes that the probability values are *measurement-neutral*, or independent of a choice of additional, compatible measurements. For more on measurement-neutrality (which is also sometimes called "noncontextuality"), see Section §3.5 (and especially footnote 19).

Equation (9) gives one statement of the Born rule for open quantum systems. Thus, Gleason's theorem proves that there is a bijection between chance-states and QM-states for (nearly all) open quantum systems.[11]

But note two strange things about Gleason's approach. First, by taking the domain of the chance function to be the partial σ-algebra $\mathcal{P}(\mathcal{H}_S)$, Gleason assigns probabilities to all projections on the Hilbert space at once—and so to all possible measurements at once, including those that cannot be made simultaneously! As a consequence, he tacitly assumes that the probability values are *measurement-neutral*, or independent of which additional, compatible measurements we choose to perform. To illustrate, note that we can only empirically verify Born's rule for one choice of compatible measurements—i.e., one (non-partial) σ-subalgebra $\mathcal{E} \subset \mathcal{P}(\mathcal{H}_S)$—at a time. So suppose that $\mathcal{E}_1 \neq \mathcal{E}_2$ and $P_i \in \mathcal{E}_1 \cap \mathcal{E}_2$. It is not a logical necessity that the probability of P_i measured with $\mathcal{E}_1$ matches the probability of P_i measured with $\mathcal{E}_2$. It is an empirical fact that is worthy of explanation. But this point is obscured by picking the domain $\mathcal{P}(\mathcal{H}_S)$, which contains just the one copy of P_i. In short, if we aim to derive the Born rule from theory, it seems odd to assume that probabilities agree in different measurement contexts from the outset, as Gleason tacitly does.[12]

Second, Gleason's theorem does not favor any particular bijection between QM states and chance states over any other. The necessary and sufficient condition that it identifies for *ch* to be Kolmogorovian is that there is *some* density operator that yields it via the Born rule. Therefore, Gleason's theorem does not decide between the Born rule and, say, the Shmorn Rule—a rule which instructs the user of QM to first rotate a unit vector ninety degrees about some one-dimensional subspace before applying the Born rule. In other words, Gleason's theorem straightforwardly yields the following:

> **Shmleason's theorem.** For $\dim(\mathcal{H}_S) \geq 3$, a function $ch : \mathcal{P}(\mathcal{H}_S) \to \mathbb{R}$ satisfies Kolmogorov's rules if and only if there exists a density operator ρ' such that
>
> $$ch(P_i) := \mathrm{Tr}(U^\dagger \rho' U P_i). \tag{10}$$
>
> for some fixed unitary operator $U \neq I$.

For the proof, let ρ be the state from Gleason's theorem and define $\rho' := U\rho U^\dagger$. Now note that Shmleason's theorem "endorses" a different probability rule. Shmleason's theorem is just as sound as Gleason's theorem; as of yet, we have no reason to privilege (9) over (10).[13]

Thankfully, Gleason's theorem invokes very little of the structure of NRQM. In particular, it does not invoke the quantum updating rule. This rule is one key ingredient of the symmetry theorem. However, its physical significance has long been a matter of controversy—so we had better justify it if we wish for the symmetry theorem to answer our why-question properly! Following Wallace, I begin my justifications of the quantum updating rule by positing environmental decoherence.

[11] As Busch (2003) has shown, we can remove the parenthetical by attending to POVMs, which generalize PVMs. I will be sticking with PVMs for ease of exposition, but the proceeding arguments all generalize naturally to allow for POVMs.

[12] Wallace (2003) also emphasizes this point, and he notes that the symmetry theorem offers an explanation of measurement neutrality. I recover this explanation for the Bohmian in Section §3.5.

[13] In (2003, pp. 433–434), Wallace points to this underlying problem when he asserts that were we to try to use Gleason's theorem to derive the Born rule, we would still need to appeal to the "games" developed in Deutsch's original proof of the symmetry theorem. As Gill (2005) notes, it is possible to add one of Deutsch's assumptions to Gleason's theorem in order to derive the right bijection (viz., what appears below as the "normalization link"). But it is less clear that a similar strategy can offer an explanation of measurement neutrality.

3 A symmetry theorem for pilot wave and many-worlds theories

This section shows how both the Everettian and the Bohmian can use decoherence to secure the physical meaning of the quantum updating rule. Then, it argues that each theorist can use decoherence to give an equally good justification of state supervenience. To review the structure of the argument: §3.1 describes how to use decoherence to recover quasi-classical histories in an interpretation-neutral way, closely following the accounts of Wallace (2012) and Rosaler (2016). §3.2 uses this account as a framework to house two Everettian approaches to measurement, the Hydra and Lewisian views, in addition to the Bohmian view. §3.3 argues that the Lewisian and the Bohmian motivate state supervenience equally well, and §3.4 assesses what goes wrong with Wallace's argument that the Bohmian cannot do so. Finally, §3.5 states my interpretation-neutral version of the symmetry theorem.

3.1 An interpretation-neutral approach to decoherence

Recall the decoherence program's core idea: when a subsystem of interest couples with its environment, coherence among its pointer states leaks into the total system, leaving the subsystem in a mixture of these states. As Schlosshauer (2007) cogently argues, positing just this much on top of the bare quantum theory described above does not solve the measurement problem. Instead, it addresses two closely related issues: why some pointer bases (like Gaussian wave packets or spin eigenstates) seem to be preferred by given observations and why it is so difficult to observe the effects of quantum coherence at macroscopic levels. It resolves the former by specifying a physical mechanism—namely, the system-environment interaction—that picks out the preferred basis. It addresses the latter by positing that macroscopic systems undergo much quicker decoherence than their microscopic cousins. But the program does not try to explain the appearance of *specific* outcomes. This final question is one that the traditional interpretations of quantum theory (many-worlds and pilot wave theories among them) are poised to answer. All three of these elements combined provide one or another physical justification of the quantum updating rule as a description of repeatable preparations and measurements.

If the preceding is right, then decoherence does not favor any given traditional interpretation. Likewise, any tool that tracks it—including the decoherent histories formalism—cannot carry interpretive commitments. Nonetheless, it is worth reviewing the formalism, as its relationship to interpretations is often unclear. Gell-Mann and Hartle's (1990) original proposal does not spell out the role of measurement, and their subsequent works take interestingly divergent approaches. For example, while one may read Hartle (2010) as recommending a many-worlds approach, Gell-Mann and Hartle (2012) explicitly endorse just one history as "real." So, closely following Wallace (2012) and Rosaler (2016), I will briefly detail how I take their histories formalism to accommodate both many-worlds and pilot wave theories.

We start wth our total, closed system described by some state $\Psi \in \mathcal{H}_{SE}$ at time t_0. Now suppose that on a very brief timescale τ_D, the interaction between the system and the environment H_{SE} dominates the Hamiltonian in equation (1). Pick an $\mathcal{H}_S$-spanning set of states $\{\psi_j\}$, the pointer states, that are robust under the action of H_{SE}. More precisely, for some $\Delta t \ll \tau_D$, a system prepared in a pointer state (with some environmental "ready state" E) couples with its environment and evolves as

$$|\psi_j\rangle \otimes |E\rangle \xrightarrow{\Delta t} |\psi_j\rangle \otimes |E_j\rangle, \tag{11}$$

where E_j is some final environmental state. If, instead, the system begins in a superposition of pointer states, then due to the linearity of (1), the total system evolves as

$$(a|\psi_1\rangle + b|\psi_2\rangle) \otimes |E\rangle \xrightarrow{\Delta t} a|\psi_1\rangle \otimes |E_1\rangle + b|\psi_2\rangle \otimes |E_2\rangle. \tag{12}$$

This completes our first step of decoherence, which we suppose ends at t_1. Applying the partial trace of equation (2), we see that the final state of the system is given by

$$\rho_S(t_1) = |a|^2|\psi_1\rangle\langle\psi_1| + |b|^2|\psi_2\rangle\langle\psi_2| + a^*b\langle E_1|E_2\rangle\,|\psi_2\rangle\langle\psi_1| + b^*a\langle E_2|E_1\rangle\,|\psi_1\rangle\langle\psi_2|. \tag{13}$$

If the environment interacts strongly enough with the system, then we can suppose that $\langle E_1|E_2\rangle \simeq 0$. For example, take ψ_1 and ψ_2 to be well-separated, localized states of a heavy dust particle S, and take E to be a short-wavelength air particle scattering strongly off the dust. Then E_1 and E_2 would be roughly orthogonal states of the air particle. So we have

$$\rho_S(t_1) \simeq |a|^2|\psi_1\rangle\langle\psi_1| + |b|^2|\psi_2\rangle\langle\psi_2|, \tag{14}$$

an approximate mixture of pointer states. In this way, decoherence has (approximately) moved the initial superposition of the system into its environment.

Now let the system continue to evolve to a later time. Once the system is in a mixture of pointer states, the environment has very little effect on its further evolution—so, to a good approximation and for some intermediate time $\Delta t' \gg \tau_D$, the system and the environment evolve independently under their own self-Hamiltonians. Since this dynamical evolution is linear, each of the terms in equation (12) evolves into a new superposition of pointer states tensored with a new ready state of the environment. For example, the evolution might look like

$$a|\psi_1\rangle \otimes |E_1\rangle + b|\psi_2\rangle \otimes |E_2\rangle \xrightarrow{\Delta t'} a\,(a_1|\psi_1\rangle + a_2|\psi_2\rangle) \otimes |E_1'\rangle + b(b_1|\psi_1\rangle + b_2|\psi_2\rangle) \otimes |E_2'\rangle. \tag{15}$$

Then decoherence will occur again, but for *each* of the terms in the superposition on the right-hand side of equation (15). This process yields the total evolution

$$(a|\psi_1\rangle + b|\psi_2\rangle) \otimes |E\rangle \xrightarrow{\Delta t + \Delta t' + \Delta t} |\psi_1\rangle \otimes (aa_1|E_{1,1}\rangle + bb_1|E_{2,1}\rangle) + |\psi_2\rangle \otimes (aa_2|E_{1,2}\rangle + bb_2|E_{2,2}\rangle), \tag{16}$$

completing our second instance of decoherence (which we suppose ends at t_2). Note that decoherence ensures that either of $E_{1,1}$ or $E_{2,1}$ is approximately orthogonal with either of $E_{1,2}$ or $E_{2,2}$. But we also expect the information recorded by the first decoherence event to be distributed widely throughout the environmental degrees of freedom (so new bits of the environment, e.g., other particles in the air scattering off of our system, are doing the second bit of decoherence). Thus we expect the pairs $E_{1,1}, E_{2,1}$ and $E_{1,2}, E_{2,2}$ to be approximately orthogonal, as well. In this way, decoherence yields a natural branching structure wherein the environment records four different sequences of pointer states (ψ_1 then ψ_1, ψ_2 then ψ_1, ψ_1 then ψ_2, and ψ_2 then ψ_2), and none of these sequences interfere with each other. Naturally, this process may be iterated an arbitrary number of times.

Note well that the set of pointer states might be uncountable and overcomplete. Both Schlosshauer (2007, §2.8, §5.2) and Wallace (2012, §3) argue that a particularly natural choice of pointer states (for a single spinless particle) is the set of coherent states

$$\psi_{(q,p)}(x) = \langle x|q,p\rangle = ae^{-\lambda(x-q)^2}e^{ipx} \tag{17}$$

(for q and p reflecting "position" and "momentum" values ranging over the reals, and where λ denotes the width of the Gaussian wave-packet). For these pointers, an arbitrary initial state $\Psi \in \mathcal{H}_{SE}$ may be written as

$$\Psi(t_0) = \int dq_0 dp_0\, c(q_0,p_0)\,|q_0,p_0\rangle \otimes |E(q_0,p_0)\rangle, \tag{18}$$

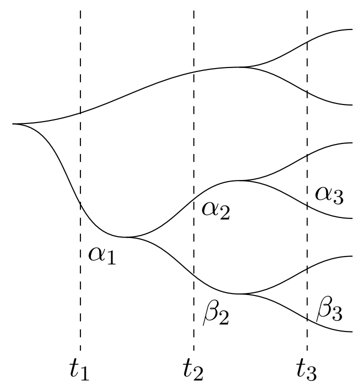


Figure 1: Schematic illustration of a branching history algebra with $n = 3$ steps of decoherence. Events have been labelled for the two histories α and β (note that $\alpha_1 = \beta_1$).

where $c(q_0, p_0)$ is the coefficient for the appropriate pointer state and $|E(q_0, p_0)\rangle \in \mathcal{H}_E$ is its associated environment state. To clean up our notation a bit, let us introduce the sequence $A = (A_1, \ldots, A_n)$ to represent a sequence of pointer states like $|q_0, p_0\rangle, \ldots, |q_n, p_n\rangle$ for n steps of decoherence. Let $c(A)$ be the product of transition amplitudes for the appropriate pointer states—so, for example, in equation (16), we have $c(\psi_1, \psi_2) = aa_2$ and $c(\psi_1, \psi_1) = aa_1$. At t_n, the state in equation (18) has evolved to

$$\Psi(t_n) = \int dq_0 \ldots dq_n dp_0 \ldots dp_n \, c(A) \, |q_n, p_n\rangle \otimes |E(A)\rangle, \tag{19}$$

where $|q_n, p_n\rangle$ is the final state in A and

$$\langle E(A')|E(A)\rangle \simeq 0 \text{ if } A' \neq A. \tag{20}$$

Equation (20) points to the idea that n steps of decoherence result in a number of incompatible environment states, each recording a different history of n pointer states.

Capitalizing on this idea, let us idealize equation (20) such that at each of the n steps, the environmental states are *exactly* orthogonal. Given this idealization, consider all possible pointer states A_i at time t_i and note that their associated projections

$$\alpha_i = |A_i\rangle\langle A_i| \otimes |E(A_1, \ldots, A_i)\rangle\langle E(A_1, \ldots, A_i)| \tag{21}$$

form a mutually orthogonal set.

To derive the Born rule, it would help to organize these projections within a familiar algebraic structure. Luckily, our idealization allows us to form a natural σ-algebra of events. First, note that each time-indexed projection α_i is associated with a PVM, which in turn specifies a σ-algebra $\mathcal{S}^i$ of projections on $\mathcal{H}_{SE}$. So, following Wallace (2012, pp. 95–96), define the *history algebra* $\{\mathcal{S}^i\}$ as the n-fold direct product of such σ-algebras of projections,

$$\left\{\mathcal{S}^i\right\} := \mathcal{S}^1 \times \ldots \times \mathcal{S}^n, \tag{22}$$

the elements of which are given by complete specifications of histories $\alpha = (\alpha_1, \ldots, \alpha_n)$ (for the appropriate initial state and dynamics, and where the algebraic operations are performed pointwise). Our idea now is to characterize when a history algebra witnesses the branching structure described above.

One natural way to do so is to consider the transition weights for projections in atomic coarse-grainings of the algebra. We say that a *coarse-graining* of $\left\{\mathcal{S}^i\right\}$ is a history algebra $\{\mathcal{C}^i\}$ where every projection in $\mathcal{C}^i$ is a sum of projections in $\mathcal{S}^i$ (for every i). A coarse-graining is *atomic* if each $\mathcal{C}^i$ is the free σ-algebra generated by some countable set of projections. Now define the transition weight between any two projections, for $t_i < t_j$ as follows:

$$\mathcal{T}(\alpha_i, \beta_j) := \frac{|\beta_j U(t_i, t_j) \alpha_i U(t_0, t_i)\Psi|^2}{|\alpha_i U(t_0, t_i)\Psi|^2}, \tag{23}$$

where $U(t_0, t_i)\Psi = \Psi(t_i)$ and $U(t_i, t_j)\Psi(t_i) = \Psi(t_j)$. We say that a history algebra is *branching* when each projection "receives weight" from just one of its predecessors, which amounts to satisfying the following condition.

> **Branching criterion.** A history algebra $\{\mathcal{S}^i\}$ is *branching* for Ψ (or Ψ-*branching*) when it admits an atomic coarse-graining $\{\mathcal{C}^i\}$ such that, for any $\alpha, \alpha', \beta \in \mathcal{C}$ and $t_i < t_j$,
>
> $$\text{if } \mathcal{T}(\alpha_i, \alpha'_j) \neq 0 \text{ and } \mathcal{T}(\beta_i, \alpha'_j) \neq 0, \text{ then } \alpha_i = \beta_i. \tag{24}$$
>
> If the above is satisfied for $\{\mathcal{C}^i\} = \{\mathcal{S}^i\}$, then we say that $\{\mathcal{S}^i\}$ is *strictly branching* (or *strictly* Ψ-*branching*).

Figure 1 gives a heuristic illustration of the branching that results. We need to invoke one final idealization before proceeding: we assume that every decohering kinematic trajectory is dynamically possible. As such, we can avail ourselves of environmental decoherence along any pointer basis for any suitable dynamics:

> **Decoherence availability**. Let $\{\mathcal{S}^i\}$ be an atomic history algebra with n steps of decoherence. For any $\Psi(t_0)$, any set of unitary maps $\{U_1, \ldots, U_n\}$ for which for which $\{\mathcal{S}^i\}$ is strictly Ψ-branching is available (where $U_i\Psi = \Psi(t_i)$ for each i).

The coarse-grained dynamics $\{U_i\}$ idealize, in part, the evolution generated by the system's self-Hamiltonian and the environment's. Thus, decoherence availability captures the idea that the system can evolve in any unitary manner during the time intervals $\Delta t'$ between decoherence events along an arbitrary pointer basis (as described above). In other words, the principle just ensures that we can perform all of what Wallace calls the "physically performable operations" (2012, p. 153).

We can now see how decoherence conditionally justifies the quantum updating rule. In short, (24) ensures that every branch is dynamically isolated from all the others. So as long as we can associate projective measurements with branches, the quantum updating rule holds. Thus, each branching event α_i creates a new closed system with the wave function $\Psi' = \alpha_i\Psi(t_i)/|\alpha_i\Psi(t_i)|$. We might loosely call this recursive specification of closed systems "subsystem-recursivity in time"; Figure 2 illustrates this recursivity. We also have what we might roughly call "subsystem-recursivity in space." At every time step t_i, the projected wave function $\Psi'(t_i) = |A_i\rangle \otimes |E(A_1, \ldots, A_i)\rangle$ is separable. Thus, the system's wave function $|A_i\rangle \in \mathcal{H}_S$ is decoupled from the environment, and it may be promoted to a closed system.[14] Assuming that agents reliably know pointer states, they can reliably know which new closed systems a branching event yields. Decoherence then justifies our repeated "discovery" of closed systems with specific, explicit wave functions Ψ—so long as we can explain why we see only one of the options in the post-branching superposition (or otherwise show that this question is ill-posed). Thus, to complete the justification of the quantum updating rule, we turn to how the many-worlds and pilot wave theorists respond to the question of specific outcomes.

[14] Wallace (2019) provides a rigorous, formal theory of subsystem-recursivity that informs my brief and qualitative discussion here.

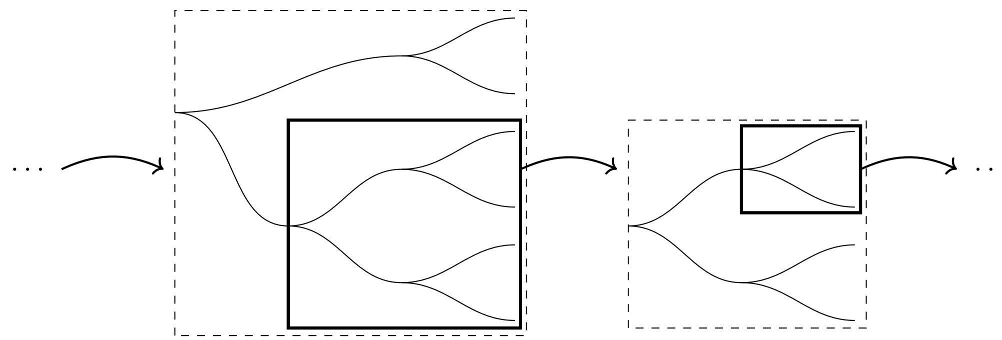


Figure 2: Subsystem-recursivity in time shared by both many-worlds and pilot wave theory.

3.2 The Hydra, Lewisian, and Bohmian views

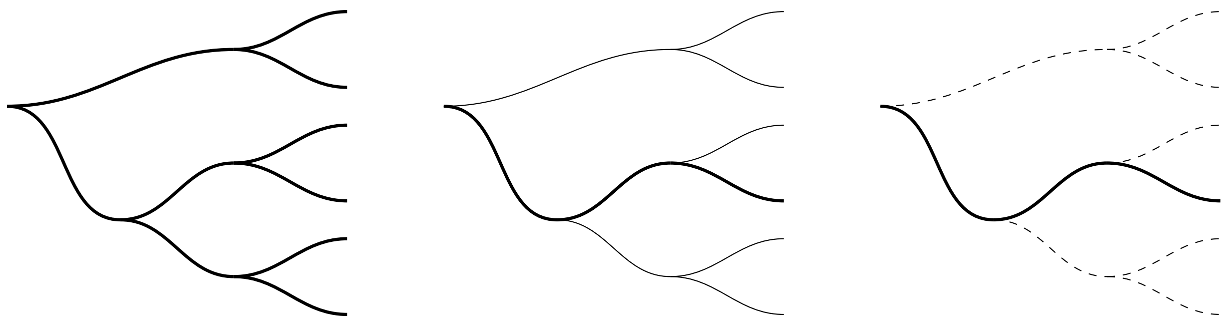
How do we square our talk of specific outcomes with the claim that every history (with non-zero weight) does, in fact, occur? Everettians have developed several distinct ways of answering this question. I will focus on just two of these: the Hydra view and the Lewisian view.[15]

On the Hydra view, we dig in our heels and claim that talk of specific outcomes is confused. If agents wish to talk about the physics of the universe, the thought goes, then they ought to talk about everything in it—so the claim that a spin-superposed electron is *only* measured in a spin-up state is just false. This view, illustrated in Figure 3a, instructs us to ignore the question of specific outcomes: strictly speaking, such outcomes are illusory. But this response is distinctly uncharitable to users of the theory. Suppose you and a friend toss a fair quantum coin to decide who leaves a tip for your waiter: if you claim that the coin did not land heads, clearing you of responsibility, you are, strictly speaking, wrong. By the lights of the Hydra view, users of quantum theory speak falsehoods the vast majority of the time. So the view is not terribly kind to its adherents (although it might, incidentally, turn out to be quite a bit kinder to waiters).

Notably, the Hydra view eliminates uncertainty about future events. Every proposition in $\{\mathcal{S}^i\}$ with non-zero weight is true. As flagged by the story of Prof. X, this view must sever the link between chance and uncertainty flagged in §2.1. On the one hand, the insight that we can sever this link is one of the modern Everettian program's crowning achievements. On the other, it shows another way in which the Hydra view is uncharitable: agents are, strictly speaking, incorrect to claim that chancy events are uncertain. Given these difficulties, it would be nice to have a more charitable alternative on hand, one that vindicates agents' talk of uncertainty.

Wallace presents the Lewisian view as one such alternative (a view inspired by, but importantly distinct from, David Lewis's views on personal identity; cf. Saunders and Wallace 2008, Saunders 2010, Wallace 2012). On this view, any individual referenced by an agent (including any person or any thing) ought to refer to an appropriate *part* of the universal state. Explicitly, suppose again that our agent claims to only see spin-up. Let α_t denote this measurement event, occurring in some appropriate history α. The agent's "spin-up" corresponds to some ordered pair $\chi = (\Psi, \alpha)$ of a universal state and total branch history, where α contains α_t. The state χ makes α true, and it makes all the other atomic propositions in $\left\{\mathcal{S}^i\right\}$ false. Saunders (2010, p. 192) provides a metaphysics

[15] I borrow this terminology from Wallace (2012, p. 281). These views have natural analogs that cut branches into Siderian stages rather than Lewisian worms—Wallace refers to these as the Disconnected view and the Stage view, respectively (2012, p. 282). Tappenden (2011) has extensively developed the Stage view, and he saliently notes that the view yields post-measurement, pre-observation uncertainty that is quite similar to single-world uncertainty about chancy events. But since the Hydra and Lewisian views suffice to make my point about the symmetry theorem, I will not cover the Disconnected and Stage views here.



(a) Hydra view (b) Lewisian view (c) pilot wave view

Figure 3: Schematic illustrations of the many-worlds completions of a branching history algebra on the Hydra and Lewisian views, as well as a sketch of the Bohmian completion.

compatible with these semantics by associating every complete history α with a particular spacetime worm, an individual object to which any observer can refer. Figure 3b gives a schematic illustration, singling out the history α (a complete history with a non-zero weight).

The pilot wave theory that I consider at length adapts Barrett's (2019) presentation of Bohm's (1952a, 1952b) theory. This theory posits (point-like) particles whose motion is governed by the wave function. We might view the wave function as a real, physical field or merely as a governing-law. In either case, its physical significance is chiefly dynamical. In turn, only one history is "real" in the sense that only one can (approximately!) describe the particles' actual trajectories. So the Bohmian state makes some $\alpha \in \left\{\mathcal{S}^i\right\}$ true, and it makes all the other atomic propositions false. The wave function still contains information regarding all the other possible paths (represented by the dotted lines in Figure 3c).[16]

We can complete our description of Bohmian particles with a bit more formalism. Specifically, we add states and dynamics directly describing the particles themselves. Describe a total system of N particles evolving in $\mathbb{R}^3$ with a point $q = (q_1, \ldots, q_N)$ in the configuration space $\mathbb{R}^{3N}$. For a given time t, Ψ is a spinor-valued function on this space. (Recall that "spinors" are vectors of complex numbers that allow for the description of systems that "have spin"; see Norsen (2014) for a more detailed discussion.) While equation (1) still governs Ψ, it leaves the evolution of the point-particles under-determined. We adopt one natural option, namely, the guiding equation

$$\frac{dq_i}{dt} = \frac{\hbar}{m_i}\mathrm{Im}\frac{\Psi^*\nabla_i\Psi}{\Psi^*\Psi}(q), \tag{25}$$

where $q_i = (x_i, y_i, z_i)$ is the position of the ith particle and $\nabla_i = (\partial/\partial x_i, \partial/\partial y_i, \partial/\partial z_i)$ is the gradient with respect to that position (and the products of spinors in the numerator and the denominator are scalar products). Just as it does for the Schrödinger dynamics, decoherence guarantees that these particle dynamics are always well-behaved: the motion of the particles is the same regardless of whether or not we apply the quantum updating rule to Ψ after decoherence. If we do apply the rule, we get one instance of what DGZ call an *effective wave function*—very roughly, a wave packet

[16]Note well that I am not making any specific ontological claims here about Ψ! The dotted lines in Figure 3c are meant to represent possible particle paths in (four-dimensional) spacetime, much like the solid lines in Figure 3b are meant to represent worms in (four-dimensional) spacetime. This view of Bohmian particles is compatible with taking Ψ to be an object that lives in an ontic $3N$-dimensional configuration space. It is compatible with taking Ψ to specify a multi-valued field in spacetime in the style of Romano (2020). It is also compatible with taking Ψ to be purely nomic, nothing more than a law governing the motions of particles in spacetime. The reader should feel free to choose whichever view of Ψ seems most natural! Nothing in the subsequent argument will hinge on this choice, so long as Ψ's main role remains the specification of point-particle dynamics.

in Ψ whose support contains the actual particle configuration and is macroscopically distinct from the supports of the other wave packets that comprise Ψ.[17] The complete state of a closed system, then, is $\chi = (\Psi, q)$, where and (1) and (25) specify the dynamics (and where Ψ may be an effective wave function). With our three views of measurement on the table, let us now turn to the crucial question of how each might justify state supervenience.

3.3 How each view justifies state supervenience

Recall that state supervenience, the claim that chances supervene on Ψ, is a crucial ingredient for the symmetry theorem—and it is the one ingredient that Wallace argues the Bohmians can't have. The Hydra view might seem to have an easy enough time justifying this principle: Ψ, after all, is the only thing that agents can talk about.

Plausibly, however, this fact about Hydra semantics does not give the right sort of explanation of why Ψ is so important to chance. It seems more relevant to this question that the dynamics of decoherence limit the ways that agents can prepare or measure quantum systems. Let us reify as a principle the claim that decoherence limits an agent's reliable knowledge to Ψ.

> **Decoherence exclusivity.** Agents gain reliable knowledge about quantum systems *exclusively* through decoherence along known pointer states, and so Ψ is the most that they can reliably know.

This claim is well-motivated by NRQM, which does not seem to contain (on its own) any other dynamical resources that could influence measurement—pathological cases such as the recoherence of branches notwithstanding. Elevating the claim to the level of an axiom primarily serves to rule out those possible pathological cases by fiat. But it also rules out any esoteric metaphysical possibilities that might arise on views less spartan than the Hydra one.

Note that this decoherence principle, on its own, *directly implies* state supervenience. We have stipulated that admissible background information in PP must be reliable. So if Ψ is the most that agents can reliably know, then chances must supervene on it; end of story. Moreover: decoherence exclusivity is not exclusive to any one view. The Hydra, Lewisian, and Bohmian views are all free to endorse it.

But an Everettian might think that the semantics of the Hydra view does explain state supervenience well, contra the plausible claim above. Such an Everettian might wonder if an analogous strategy is available to the Lewisian. Can they use semantics alone to justify state supervenience? If so, then they might have a leg up on the Bohmian, after all.

To answer this question, let us take a closer look at Saunders's (2010) version of Lewisian semantics. Saunders supposes that agents can refer to time-slices of quantum states along branches. The natural metaphysical correlate, here, is an instantaneous worm-slice, or *stage*. (Henceforth, I will simply talk about stages and worms, but bear in mind that one is always free to take a deflationary approach to these objects!) Let us denote stages with tuples $\chi = (\Psi, \alpha, \alpha_t)$, or α_t for short (when this usage is clear from context). There now arises a crucial question as to whether worms share their temporal parts before branching. In the case of two worms α and β, if their parts do so "overlap," then there is numerically one worm-slice $\alpha_t = \beta_t$ at a time t before branching, and numerically two slices after. Nevertheless, it might be the case that there are numerically two worm-slices at all times. In this case, one history of worm-slices might qualitatively agree with another up to the time of branching, at which time the former "diverges" from the latter.

[17] See §5 and fn. 20 of DGZ (1992) for precise definitions and a brief discussion of how decoherence yields effective wave functions. For a longer discussion, see Romano (2016).

It is clear that the divergence view, at least, provides one good way of vindicating uncertainty-talk. On divergence, a stage on a worm is bit like a driver on a forking road—but we suppose (somewhat counter-intuitively) that before the road forks, there are two drivers and two roads that just happen to coincide. One driver and one road follow the left fork; the others follow the right. A bit more precisely: when an agent thinks about their future at time t, this thought is possessed by a slice α_t, and this slice is distinct from any other slice β_t of a worm β whereon the agent sees a different outcome. Then we can say that, before branching, our agent is uncertain about whether their fully-extended worm is α or β. P. J. Lewis (2007) argues that things do not go quite as smoothly on the overlap view. Here, there is one driver and one road before the fork; $\alpha_t = \beta_t$ *numerically*. But if it is right that slices possess thoughts (rather than entire worms), then there is just one stage possessing one thought at t. And so the uncertainty seems to disappear. At the very least, care is needed to flesh out Lewisian semantics that can deliver on Wallace's promise of charity towards uncertain agents.

But regardless of whether the Lewisian chooses divergence or overlap, they need to do a bit more work to explain why agents cannot use worms or stages as background information in PP. Wallace provides a principle that does this work. He asserts that, on any Lewisian view, "each agent does have a unique future, but it is in principle impossible for him to possess reliable knowledge of that future" (2012, p. 150, fn. 25). In other words, an agent cannot know the identity of their history α in $\chi = (\Psi, \alpha)$ with a grain fine enough to fix future facts (regardless of how we divvy up their worm's temporal parts). I will call this assertion *self-ignorance*.

> **Self-ignorance.** An agent cannot reliably know the identity of the history α that determines which branch they will take before the branching.

Self-ignorance implies that chances cannot depend on α. The question, now, is how the Lewisian can justify this principle.

Does self-ignorance follow from the semantics alone? It might seem so, at first glance. After all, on both the divergence and overlap views, worms are spatially and temporally coincident—and so at least dynamically identical—before branching. But this observation alone does not imply self-ignorance on either view. For divergence, suppose that, say, my chair and the corresponding chair-wise arrangement of atoms are numerically two objects. These objects are spatially and temporally coincident, and they evolve the same way in time. But I see both, and I am clear on which is which. For overlap, suppose one version of me decides to remove my chair's armrests. Another decides otherwise. There is numerically one chair before the time my second self removes the armrests and two afterward. At the moment I make my decision, I am quite clear on which of the two future chairs is mine—even though there is numerically one chair at that moment! So in either case: why should a lack of dynamical difference before branching limit agents' knowledge?

The best way that I can see for the Lewisian to respond is to invoke decoherence exclusivity. With this principle, they can stop fretting about my armchair and simply appeal to the dynamics of decoherence to limit knowledge directly. The only price they have to pay is the admission that Bohmians can make this appeal, too.

For the sake of comparison, it is worth spelling out how decoherence exclusivity justifies the Bohmian's lack of reliable knowledge of particle configurations. Barrett (2019) nicely illustrates a pilot waver's ignorance of a configuration q with a toy model of a Stern-Gerlach experiment. In this toy model, an electron e in a superposition of spins travels a path B if its configuration lies in the support of the spin-down wave packet (and it otherwise travels a path A). Barrett then introduces a particle p that acts as an idealized measuring device. The particle's wavefunction entangles with the electron's such that p moves to a region b if and only if e took path B. Of this model, Barrett writes:

> Suppose that the recording particle p in the two-path experiment moves to region b. This does not tell us precisely where the electron is. It might be anywhere in the wave packet that traveled path B. Rather, it tells us *which wave packet* the configuration (and hence the electron) is associated with. That is, the empirical content of the record, what one can deduce from the value of the record, is given by *the effective wave function selected by the current particle configuration.* This is what an observer has epistemic access to given her measurement record. In this precise sense, this is what she sees. (2019, pp. 213–214, emphasis Barrett's)

Thus, the Bohmian agent only ever knows q approximately, at least in domains where this model of measurement applies. And conveniently, we have in our pocket a separate dynamical condition that will *guarantee* that this model applies: namely, decoherence exclusivity. The environmental record states of the decoherent histories formalism simply generalize Barrett's recording particle p. So decoherence exclusivity explains why agents cannot reliably know the precise value of q before measurement. This claim is the one that I call *q-ignorance* in the introduction.

> **q-ignorance.** An agent cannot reliably know a system's configuration q before measurement.

Bohm (1952a, 1952b) motivates q-ignorance as a "practical limitation" rather than as a consequence of decoherence exclusivity, as I do here, and I grant that not all Bohmians would be willing to accept my axiom-style approach. For one thing, decoherence exclusivity does not reduce all facts about agents to dynamics: it invokes "reliable knowledge" as a primitive. But I would wager that this decoherence principle is dynamical enough to be of interest to at least some pilot wavers. Indeed, as Callender (2007) notes, many Bohmians are willing to go further and simply posit the Born rule as an axiom (as there are plenty of non-dynamical laws in physics already).

It seems, then, that q-ignorance and self-ignorance enjoy equally good dynamical justifications via decoherence exclusivity. Moreover, the strategy of using Lewisian semantics alone to justify self-ignorance—and thereby state supervenience—does not look promising. So, by my tally, the Lewisian and the Bohmian motivate state supervenience equally well.

But if that's right, then Wallace's argument that no single-world theory can justify this supervenience principle must be unsound. In the next section, I diagnose where Wallace's argument goes wrong.

3.4 Wallace's argument that Bohmians cannot justify state supervenience

Here is Wallace's argument that only a many-worlds theory can justify state supervenience, in brief. Wallace's key premise is that if part of a measurement event breaks the symmetries of Ψ's deterministic evolution, then the chance of that event cannot supervene on it. With this premise in tow, Wallace supposes, contra the many-worlds hypothesis, that measurement is a single-world event. He claims that this measurement process is either stochastic or deterministic. If it is stochastic, he writes, we could either specify the process or leave it unstated. He argues that the former breaks Ψ's symmetries and that the latter is question-begging. If the process is deterministic, he claims, then we ought to posit a distribution over microstates—and then the actual microstate (e.g., for the pilot wave theory, the configuration q) breaks Ψ's symmetries. In either case, the only live options break Ψ's symmetries. So by the key premise, chances cannot supervene on Ψ in a single-world theory.

Wallace summarizes the moral:

> Whether we are considering a stochastic or a deterministic process, the problem is ultimately the same. We are attempting to use a dynamical symmetry between two possible

outcomes to argue that the outcomes are equally likely. But since only one outcome actually occurs, something must break the symmetry—be it the actual microconditions of the system, or the actual process that occurs in a stochastic situation. Either way, we have to build probabilistic assumptions into that symmetry-breaking process, and in doing so we effectively abandon the goal of explicating probability. (Wallace, 2012, pp. 147–148)

This objection originates in (Wallace, 2003, §7), where it arises as part of a response to Barnum et al. (2000). In addition to their claim that Gleason's theorem renders Deutsch's argument otiose, Barnum et al. level two other criticisms. First, they note that while Deutsch claims to derive probability using the non-probabilistic part of NRQM and decision theory, one can derive the probability axioms from decision-theoretic principles on their own. Second, they claim that Deutsch's proof contains a technical non-sequitur. Wallace (2003) handily dispenses with the first of these criticisms. In so many words: Deutsch is deriving the NRQM-chance link, not probability *tout court*. The supposed non-sequitur to which the second criticism points turns out to be a consequence of measurement neutrality, which in turn follows from the assumption of state supervenience (Wallace 2003, p. 432; Wallace 2012, p. 197). If we can defend state supervenience, then the criticisms of Barnum et al. (2000) dissolve. With the above argument, Wallace (2003, 2012) claims that EQM is unique in its ability to give such a defense.

As demonstrated in the previous section, this argument is not sound: the key premise that symmetry-breaking scuttles state supervenience is false. Decoherence exclusivity directly implies state supervenience for any of the three views of measurement. To be sure, interesting differences remain. But these differences cannot have any impact on an agent's *reliable knowledge*—and so they cannot have any impact on *chance* (as long as we are defining chance with PP). The Hydra, Lewisian, and Bohmian views each have equal claim to decoherence exclusivity. And the semantics-first strategy exclusive to the Lewisian does not look promising.

There is no denying that the Hydra view's proponents have a particularly elegant strategy for defending state supervenience: they simply note that Ψ is all they care to talk about. But my strategy for Lewisians and pilot wavers, while not quite as simple, seems no less elegant: they simply note that, due to decoherence, Ψ is all we can reliably know. Supervenience then follows from Wallace's operational approach to chance.[18] So, for the Lewisian and the Bohmian, the game is zero-sum. Nothing probabilistic is gained or lost when switching from many worlds to one.

We have secured state supervenience for each of the Hydra, Lewisian, and Bohmian views. To derive the Born rule, all that remains is to draw structural links between our chance states and our NRQM states.

3.5 The symmetry theorem

The symmetry theorem flows from a simple observation: if a system starts in a superposition of two pointer states with the same weights, we can construct symmetric measurements that yield the same final state. Explicitly, imagine a Stern-Gerlach apparatus measuring an electron in a superposition of spin-up and spin-down. The system evolves with the unitary map U on Ψ as

$$(a|\uparrow\rangle + b|\downarrow\rangle) \otimes |E\rangle \xrightarrow{U} a|\text{up}\rangle + b|\text{down}\rangle. \tag{26}$$

[18] In an early footnote, Wallace might concede that this argumentative strategy is open to the Bohmian. He notes that the relevant symmetries may only appear at the level of distributions on phase space (2003, p. 435, fn. 8). But using the symmetries in *distributions* raises the specter of circularity. My approach assigns only *dynamical* significance to the pilot wave Ψ (in the manner discussed in §3.2).

Now suppose we rotate the apparatus upside-down run the experiment again. The rotation is implemented by another unitary map U', codifying a symmetry between this measurement and the first. This upside-down experiment thus evolves with the map $U'U$ on Ψ as

$$(a|\uparrow\rangle + b|\downarrow\rangle) \otimes |E\rangle \xrightarrow{U'U} a|\text{down}\rangle + b|\text{up}\rangle. \tag{27}$$

Now if $a = b$, $U\Psi = U'U\Psi$. Thus, given state supervenience (and the rules of probability), the chances assigned to up and down both ought to equal one-half. In brief, the symmetry theorem uses decoherence availability and minimal chance-NRQM links to extend the above argument to arbitrary Ψ and arbitrary observables. Specifically, we need two premises that connect the structure of chance states to that of NRQM states. But if these premises hold for any one of the Hydra, Lewisian, or Bohmian views, then they hold for all three.

First, we assign chance functions to our branching history algebras that satisfy the rules of probability:

> **Probability.** For every Ψ-branching history algebra $\{\mathcal{S}^i\}$, there is a chance function $\mathbf{ch}_\Psi : \{\mathcal{S}^i\} \to \mathbb{R}$ that satisfies Kolmogorov's rules (3)–(6). This assignment defines a chance function $ch_\Psi : \mathcal{S}^i \to \mathbb{R}$ on each $\mathcal{S}^i$ as follows:
>
> $$ch_\Psi(\alpha_i) := \mathbf{ch}_\Psi\Big(\bigvee\{\gamma \,|\, \gamma_i = \alpha_i\}\Big)\,. \tag{28}$$

In other words, the chance that α_i is the ith event is given by the chance that any one of the histories γ with that event occurs. It is easy to check that ch_Ψ also follows all of Kolmogorov's rules (a property it inherits from $\mathbf{ch}_\Psi$). The Hydra, Lewisian, and Bohmian views may all adopt the operational definition of chance with PP given in §2.1. As flagged there, this definition stipulates Kolmogorov's rules, and so the reader might (rightly) demand more. But to maintain a tight focus on how the Bohmian can use the core theorem, I leave the question of motivating these rules for future work.

Second, we assume that normalization and conditionalization of chance and NRQM states agree with each another. Explicitly, we assume:

> **Structural links.** Two of them:
>
> 1. *Normalization link.* $ch_\Psi(\alpha_i) = 1$ if and only if $\Psi(t_i)$ lies in the range of α_i.
> 2. *Temporal link.* The conditional quantum state $(\Psi, \alpha_i) := \alpha_i\Psi(t_i)/|\alpha_i\Psi(t_i)|^2$ agrees with the definition of conditional probability, i.e.,
>
> $$\mathbf{ch}_{(\Psi,\alpha_i)}(\,\cdot\,) = \mathbf{ch}_\Psi\Big(\,\cdot\,\Big|\bigvee\{\gamma \,|\, \gamma_i = \alpha_i\}\Big)\,. \tag{29}$$

The normalization link connects NRQM states that lie in the range of projections with probability-one chance states. It is a weaker version of the eigenstate-eigenvalue link: a state lies in the range of a projection only if it possesses the relevant observable property (as opposed to "if and only if"). The temporal link simply coordinates the conditional chance and quantum states that are defined when some event α_i occurs. Each of the three views of measurement adds strictly more structure to NRQM-plus-decoherence. So if the above coordination of structure is justified for one view, it is justified for all three.

Before stating the symmetry theorem, we formalize state supervenience with the chance functions defined by *probability*.

State supervenience. If $\Psi(t_i) = \Phi(t_i)$, then $ch_\Psi(\alpha_i) = ch_\Phi(\alpha_i)$.

As flagged at the jump, state supervenience is the core assumption of the symmetry theorem. Accordingly, the theorem lives or dies on our ability to justify it. But the premise follows from decoherence exclusivity, and (as we have seen) all three views can adopt this axiom. The Hydra theorist might wish to forgo this axiom, as their semantics suffice to justify state supervenience. But decoherence exclusivity seems to be the best that either the Lewisian or the pilot wave theorist can do.

The preceding allows for a simple statement of the symmetry theorem for all three views:

The symmetry theorem. For a quantum system satisfying *decoherence availability*, *decoherence exclusivity*, *probability*, and *structural links*, the chance functions are given by the Born rule,

$$ch_\Psi(\alpha_i) = \langle \Psi(t_i), \alpha_i \Psi(t_i) \rangle. \tag{30}$$

And so the symmetry theorem explains why systems with the physical state χ including Ψ ought to be assigned the chance state *ch* on multiple accounts of what χ represents. I provide a proof in Appendix A.

While Saunders (2004) provides an explicitly operational derivation of the symmetry theorem, the role of decoherence in his proof is not immediately apparent. I borrow Wallace's (2012) more transparent strategy for my proof. Saunders also suggests that his operational derivation is only applicable to one pointer basis and serves as an alternative to Gleason's theorem (2004, pp. 1786–1787). But the above shows that the symmetry theorem applies to all pointer bases by attending to all the dynamical possibilities.

Moreover, while Wallace asserts that Gleason's theorem does not add anything to the symmetry argument (2003, p. 434), the former complements the latter. The symmetry theorem proves that the NRQM-chance link ought to be given by the Born rule, while Gleason's theorem shows that this rule does not miss any of the possible measurement-neutral chance states for open quantum systems. Explicitly, so long as the projections α have the form (21), the Born rule for open quantum systems, equation (9), follows from the Born rule for closed quantum systems, (30), for every density operator ρ. (This fact is an immediate consequence of Stinespring's (1955) dilation theorem, which guarantees that every density operator ρ can be expressed as $\mathrm{Tr}_E|\Psi\rangle\langle\Psi|$ for some $\mathcal{H}_E$ and some $\Psi \in \mathcal{H}_S \otimes \mathcal{H}_E$.)

Note, too, that the symmetry theorem derives, rather than assumes, measurement neutrality. As flagged in §2.2, Gleason tacitly assumes measurement neutrality by assigning probability functions directly to the partial σ-algebra, $\mathcal{P}(\mathcal{H}_S)$. But in a branching history algebra, we associate each full σ-subalgebra of $\mathcal{P}(\mathcal{H}_S)$ with a given PVM for one $\mathcal{S}^i$—so the same event occurring in different measurement contexts need not have the same chance by default. Measures defined on different history algebras need not agree where the algebras overlap. Rather, this agreement follows from the fact that our symmetry-derived measures match the Born rule. Moreover, the context-dependence of Bohmian properties is consistent with each of our four premises.[19] Indeed, the theorem gives one explanation for why the chances in the pilot wave theory do not depend on the context of measurement even while the specific outcomes of measurements do.

I do think that there is a sense in which the symmetry theorem might be easier to visualize when we assume the Hydra view rather than the Lewisian or the Bohmian view. On the Hydra view, we can put all branches on equal ontic footing, and so we are invited to view the dynamical symmetries

[19]Wallace (2012, p. 197) refers to measurement neutrality as "noncontextuality." But since this usage is at odds with the more prevalent notion of Kochen-Specker contextuality (1975), I will stick with the former term. The pilot wave theory is both measurement-neutral and (Kochen-Specker) contextual (because the values of self-adjoint observables typically depend on the context of measurement).

of Ψ in much the same way that we envision the spatial symmetries of a die or a coin. But as shown above, Ψ's symmetries ultimately have the same import for chance regardless of whether Ψ is the totality of ontology or merely a law governing particle motions—whether it is an "is" or a "tends to."

4 Discussion

On the one hand, this paper has been a work of criticism. I have argued that Wallace's (2012) claim that many-worlds theories are better off than single-world theories in making sense of probability fails by his own lights. While Wallace argues that only a many-worlds theorist can justify state supervenience, I have argued that the Lewisian and the Bohmian do so equally well by appealing to decoherence.

On the other hand, the main takeaway of this paper ought to be positive. I hope to have shown that the Deutsch-Wallace theorem should not be an item of niche interest, a feather in the cap of the most philosophically adventurous physicists—on the contrary, a whole host of approaches to quantum theory might appeal to it to explain the ascription of particular chance values. Note that this explanation presupposes not just Hilbert space structure but also decoherence-governed preparations and measurements. But I do not wish to claim that the theorem gives the final word on quantum probability. I merely wish to stress that it is one important step forward in our understanding, one that might hold interest for theorists of many different philosophical stripes.

In this spirit, I would like to conclude with a brief sketch of how the symmetry theorem interacts with other Bohmian approaches to the Born rule. For example, many Bohmians are happy to simply postulate this rule as an additional axiom of the theory (as flagged in §3.3). There are various reasons they might do so. On the one hand, they might think that "chance" should be a primitive, in which case the symmetry theorem holds no interest. On the other, they might simply be fine with granting axiom-status to various principles with epistemic primitives (like "knowledge" or "belief"). Then the symmetry argument shows how they are free to adopt a *strictly weaker* epistemic principle—decoherence exclusivity—from which the Born rule follows (in conjunction with the plausible structural constraints sketched above). Insofar as it is generally good to search for weaker axioms, the derivation is a boon for this sort of Bohmian. For them, the fact that decoherence exclusivity provides a (partially) dynamical explanation of chance values is just a bonus.

More speculatively, the symmetry theorem might be compatible with the approaches of both Dürr, Goldstein, and Zanghì (1992) and Valentini (2020). As flagged in §2.1, these authors' approaches to pilot wave theory differ from mine in at least one crucial way: they take a cosmological approach and assign a state $(\boldsymbol{\Psi}, \mathbf{q})$ to the entire universe. Very roughly, DGZ seek to *derive* q-ignorance (which they call "absolute uncertainty") as a consequence of one natural typicality measure over possible Bohmian universes, thereby ensuring that all empirical subsystems obey Born-rule statistics. Contrariwise, Valentini allows for the possibility of non-standard initial distributions over configurations of the universe. That means that some subsystems might fail to yield Born-rule statistics. But his non-standard theory can still save the phenomena, so long as it recovers the Born rule in suitable domains—which recovery is precisely the aim of Valentini's (2019) quantum H-theorem. Valentini criticizes DGZ's derivation on various grounds. But following Norsen (2018), one can read him as seeking similar derivations, albeit for *various* typicality measures as opposed to just one.

On Norsen's conciliatory approach, the symmetry theorem might be a valuable tool for cosmological Bohmians. To wit: decoherence exclusivity might serve as a principle for identifying suitable subsystems for the target of a Valentini-style derivation. That is, we might stipulate that NRQM is valid in precisely the domains where decoherence exclusivity applies. Then the symmetry theorem imposes the Born rule as a sort of consistency constraint. It restricts the possibility space of admissible typicality measures over universal configurations—where a measure is "admissible" just as long as it recovers the Born rule in the domains that satisfy decoherence exclusivity. One can view the results of DGZ and Valentini as probing this space of admissible typicality measures.

One clear virtue of this approach is that it does not stipulate any measure of typicality or chance: it derives all measures for suitable subsystems from state-space symmetries, and admissible measures on the universe follow from a consistency constraint. But this approach assuredly requires more work. For one thing, it invokes commitments that many Bohmians might find objectionable. DGZ might have a frequentist view of probability on which all facts about agents' ignorance must reduce to facts about (actual) frequencies. Their typicality argument might be compatible with this view, but the symmetry theorem might not be. The operational definition of chance central to the latter invokes agents' credences directly, so (at the very least) this definition would require a suitable frequentist justification. I do not take this brief discussion to present a definitive resolution of the debate between DGZ and Valentini. I only want to highlight how the symmetry approach is distinct from both DGZ's typicality argument and Valentini's H-theorem—as well as to point the way towards future work that might reconcile all three approaches.

The aim of the present work has been to highlight the flexibility and fecundity of the symmetry theorem as a conceptual resource, not to rock any particular Bohmian's boat. I have sketched an alternative to the approaches of DGZ and Valentini, but their proponents are free to ignore my sketch if they wish. My point is just that Wallace's attempts to reserve a symmetry-based strategy for Everettians do not succeed. Of course, applications of the symmetry theorem to various *other* single-world interpretations of NRQM are bound to involve idiosyncrasies that I have not treated here. To that, I can only say that I hope the above discussion provides a few important heuristics for future attempts to use the theorem more broadly.

Acknowledgements

The author would like to thank Harvey Brown, Benjamin H. Feintzeig, James Read, and David Wallace for their invaluable discussions and suggestions. Many thanks as well to the thorough and rigorous criticisms of several anonymous referees, without which this paper would be in a much sorrier state (though any remaining errors are, of course, the author's sole responsibility!). Finally, much thanks to the participants of the 2020 Michaelmas Oxford Philosophy of Physics Seminar for a fantastic discussion of this work. The author was supported during the completion of this work by the National Science Foundation under grants #1846560 and #2043089.

A Proof of the symmetry theorem

Before proving the symmetry theorem, we prove as a lemma an important condition that follows from the *structural links* assumption.

Branching link. When $\{\mathcal{S}^i\}$ is strictly Ψ-branching and $\mathcal{T}(\alpha_i, \beta_j) \neq 0$ for $t_i < t_j$,

$$ch_{(\Psi,\alpha_i)}(\beta_j) = \frac{ch_\Psi(\beta_j)}{ch_\Psi(\alpha_i)}. \tag{31}$$

Proof of the branching link. First, we show that chances respect branching, i.e., $\mathbf{ch}_\Psi(\gamma) = 0$ whenever

γ contains α_i, β_j such that $\mathcal{T}(\alpha_i, \beta_j) = 0$. Explicitly,

$$\mathbf{ch}_\Psi(\gamma) \leq \mathbf{ch}_\Psi\Big(\bigvee\{\gamma \,|\, \gamma_j = \beta_j \wedge \gamma_i = \alpha_i\}\Big) = \mathbf{ch}_\Psi\Big(\bigvee\{\gamma \,|\, \gamma_j = \beta_j\} \Big| \bigvee\{\gamma \,|\, \gamma_i = \alpha_i\}\Big) \tag{32}$$

$$= \mathbf{ch}_{(\Psi,\alpha_i)}\Big(\bigvee\{\gamma \,|\, \gamma_j = \beta_j\}\Big) \tag{33}$$

$$= ch_{(\Psi,\alpha_i)}(\beta_j) \tag{34}$$

$$= 0, \tag{35}$$

where (32) follows from additivity and the definition of conditional probability, (33) follows from *temporal link*, i.e., (29), (34) follows from the definition (28) in the *probability* assumption, and (35) follows from *normalization link*, the rules of probability, and the fact that $\mathcal{T}(\alpha_i, \beta_j) = 0$. The rules of probability then imply that $\mathbf{ch}_\Psi(\gamma) = 0$.

Now suppose that $\{\mathcal{S}^i\}$ is strictly Ψ-branching and $\mathcal{T}(\alpha_i, \beta_j) \neq 0$ for $t_i < t_j$. Note that

$$ch_{(\Psi,\alpha_i)}(\beta_j) = \mathbf{ch}_{(\Psi,\alpha_i)}\Big(\bigvee\{\gamma \,|\, \gamma_j = \beta_j\}\Big) \tag{36}$$

$$= \mathbf{ch}_\Psi\Big(\bigvee\{\gamma \,|\, \gamma_j = \beta_j\} \Big| \bigvee\{\gamma \,|\, \gamma_i = \alpha_i\}\Big) \tag{37}$$

$$= \frac{\mathbf{ch}_\Psi(\bigvee\{\gamma \,|\, \gamma_j = \beta_j \wedge \gamma_i = \alpha_i\})}{\mathbf{ch}_\Psi(\bigvee\{\gamma \,|\, \gamma_i = \alpha_i\})} \tag{38}$$

$$= \frac{\mathbf{ch}_\Psi(\bigvee\{\gamma \,|\, \gamma_j = \beta_j\})}{\mathbf{ch}_\Psi(\bigvee\{\gamma \,|\, \gamma_i = \alpha_i\})} \tag{39}$$

$$= \frac{ch_\Psi(\beta_j)}{ch_\Psi(\alpha_i)}, \tag{40}$$

where (36) follows from *probability*, (37) follows from *temporal link*, (38) follows from the definition of conditional probability (and a bit of algebra), (39) follows from $\mathbf{ch}_\Psi$ respecting branching and additivity (and our supposition), and (40) follows from (28). □

With *branching link* in hand, we can quickly prove the symmetry theorem.

Proof of the symmetry theorem. Following (Wallace, 2012, Ch. 4), we aim to prove that $ch_\Psi(\alpha_i) = \langle\Psi(t_i), \alpha_i\Psi(t_i)\rangle$ in four steps. First (i), we generalize the intuitive re-labeling argument to show that projections with equal Born weights must have equal chances—i.e., we show that ch_Ψ must be a function of Born weights. Second (ii), we invoke some of our environmental degrees of freedom to show that this function is increasing. Third (iii), we invoke N environmental degrees of freedom to show that function must equal the Born weight when it is rational. Fourth (iv), we use a simple limiting argument (and the second and third steps) to obtain agreement for arbitrary Born weights.

(i) Suppose $\langle\Psi(t_i), \alpha_i\Psi(t_i)\rangle = \langle\Psi(t_i), \beta_i\Psi(t_i)\rangle$.

First, suppose that both sides equal zero. Then $\Psi(t_i)$ lies in the range of both $\neg\alpha_i$ and $\neg\beta_i$, and so by *normalization link*, $ch_\Psi(\neg\alpha_i) = ch_\Psi(\neg\beta_i) = 1$. Thus, by *probability*, $ch_\Psi(\alpha_i) = ch_\Psi(\beta_i) = 0$.

Now suppose otherwise. By the above reasoning, we get that $ch_\Psi(\alpha_i) \neq 0$ and $ch_\Psi(\beta_i) \neq 0$. Next, we run the analog of the intuitive re-labelling argument. By *decoherence availability*, we can consider two different strictly branching history algebras for the next step of decoherence: one that gives a projection α_{i+1} weight from α_i and one that gives it weight from β_i. To distinguish these, let Ψ evolve with the first dynamics, and let Φ evolve with the second dynamics. Now define $\Phi(t_j) = \Psi(t_j)$ for $j \leq i$ and let $\Psi(t_{i+1}) = X\Psi(t_i)$ and $\Phi(t_{i+1}) = Y\Psi(t_i)$ for the unitary operators

$$X := V^\alpha \alpha_i + W^\alpha(1-\alpha_i) \qquad Y := V^\beta \beta_i + W^\beta(1-\beta_i) \tag{41}$$

where V^α and V^β share the range of α_{i+1} and W^α and W^β share the range of some mutually orthogonal projection β_{i+1}. By construction, $\Psi(t_{i+1}) = \Phi(t_{i+1})$.

Note, too, that $X\alpha_i\Psi(t_i)$ and $Y\beta_i\Psi(t_i)$ both lie in the range of α_{i+1}. So by *branching link*, we have

$$ch_{(\Psi,\alpha_i)}(\alpha_{i+1}) = \frac{ch_\Psi(\alpha_{i+1})}{ch_\Psi(\alpha_i)} = 1, \tag{42}$$

and similarly

$$ch_{(\Phi,\beta_i)}(\alpha_{i+1}) = \frac{ch_\Phi(\alpha_{i+1})}{ch_\Phi(\beta_i)} = 1. \tag{43}$$

By *state supervenience*, $ch_\Phi(\alpha_{i+1}) = ch_\Psi(\alpha_{i+1})$, and so $ch_\Psi(\alpha_i\ ) = ch_\Phi(\beta_i)$. Applying *state supervenience* once more, we get

$$ch_\Psi(\alpha_i) = ch_\Psi(\beta_i). \tag{44}$$

(Note that, again by *state supervenience*, this last equality holds even if our original history algebra was not strictly branching.)

(ii) Suppose $\langle\Psi(t_i), \alpha_i\Psi(t_i)\rangle > \langle\Psi(t_i), \beta_i\Psi(t_i)\rangle$.

By *decoherence availability*, we may assume that the $i+1$ step of decoherence is given by a unitary Y such that, for γ_{i+1} and ω_{i+1} two mutually orthogonal projections, we have

$$\begin{aligned}\langle Y^\dagger\Psi(t_i), (\gamma_{i+1}\vee\omega_{i+1})Y\Psi(t_i)\rangle &= \langle\Psi(t_i), \alpha_i\Psi(t_i)\rangle \\ \langle Y^\dagger\Psi(t_i), \gamma_{i+1}Y\Psi(t_i)\rangle &= \langle\Psi(t_i), \beta_i\Psi(t_i)\rangle\end{aligned} \tag{45}$$

where Y maps vectors in the range of α_i to the range of $\gamma_{i+1}\vee\omega_{i+1}$. Then, by step (i),

$$ch_\Psi(\gamma_{i+1}\vee\omega_{i+1}) = ch_\Psi(\alpha_i), \qquad ch_\Psi(\gamma_{i+1}) = ch_\Psi(\beta_i). \tag{46}$$

By *probability*,

$$ch_\Psi(\gamma_{i+1}\vee\omega_{i+1}) \geq ch_\Psi(\gamma_{i+1}) \tag{47}$$

and so

$$ch_\Psi(\alpha_i) \geq ch_\Psi(\beta_i). \tag{48}$$

(iii) Suppose $\langle\Psi(t_i), \alpha_i\Psi(t_i)\rangle$ is rational, i.e. equal to $\frac{M}{N}$ for some positive integers M, N.

By *decoherence availability*, we may pick some $\mathcal{H}_{SE}$-spanning sequence of orthogonal projections $\gamma^1, \ldots, \gamma^m, \ldots, \gamma^N$ that generate a σ-algebra containing α_i such that, for some Φ such that $\Phi(t_i) = \Psi(t_i)$,

$$\langle\Phi(t_i), \gamma^m\Phi(t_i)\rangle = \frac{1}{N} \tag{49}$$

for all m. By (i), $ch_\Phi(\gamma^m)$ must be independent of m. Thus, by *probability*, $ch_\Phi(\gamma^m) = \frac{1}{N}$.

Now let $\omega := \bigvee_{i\leq M}\gamma^i$. We have that $\langle\Psi(t_i), \alpha_i\Psi(t_i)\rangle = \langle\Phi(t_i), \omega\Phi(t_i)\rangle = \langle\Psi(t_i), \omega\Psi(t_i)\rangle$, so by step (i), $ch_\Psi(\alpha_i) = ch_\Psi(\omega)$. By *probability*, $ch_\Psi(\omega) = \sum_{m=1}^M ch_\Psi(\gamma^m) = \frac{M}{N}$, and so we get that

$$ch_\Psi(\alpha_i) = \frac{M}{N}. \tag{50}$$

(iv) Suppose $\langle\Psi(t_i), \alpha_i\Psi(t_i)\rangle = r \in [0,1]$, where r may not be rational.

By (i), ch is a function of Born-rule weights, i.e.

$$ch_\Psi(\alpha_i) = f(\langle\Psi(t_i), \alpha_i\Psi(t_i)\rangle) \tag{51}$$

for some $f : [0,1] \to [0,1]$. By step (ii), f is increasing; by step (iii), $f(M/N) = M/N$.

So let $\{a_i\}$ and $\{b_i\}$ be, respectively, increasing and decreasing sequences of rational numbers in $[0,1]$ converging to r. Since $f(b_i) = b_i$ for all i, $f(r) \leq r$. And since $f(a_i) = a_i$ for all i, $f(r) \geq r$—thus, $f(r) = r$, and so

$$ch_\Psi(\alpha_i) = \langle\Psi(t_i), \alpha_i\Psi(t_i)\rangle. \tag{52}$$

□

References

Barnum, H., Caves, C. M., Finkelstein, J., Fuchs, C. A., and Schack, R. (2000). Quantum probability from decision theory? *Proceedings of the Royal Society of London*, A456:1175–1182.

Barrett, J. A. (2019). *The Conceptual Foundations of Quantum Mechanics*. Oxford University Press, Oxford.

Bohm, D. (1952a). A suggested interpretation of the quantum theory in terms of "hidden" variables. I. *Physical Review*, 85(2):166–179.

Bohm, D. (1952b). A suggested interpretation of the quantum theory in terms of "hidden" variables. II. *Physical Review*, 85(2):180.

Brown, H. R. (2011). Curious and sublime: the connection between uncertainty and probability in physics. *Philosophical Transactions of the Royal Society of London A*, 369(1956):4690–4704.

Busch, P. (2003). Quantum states and generalized observables: a simple proof of Gleason's theorem. *Physical Review Letters*, 91(12):120403.

Busch, P., Grabowski, M., and Lahti, P. J. (1995). *Operational Quantum Physics*. Springer-Verlag, Berlin.

Callender, C. (2007). The emergence and interpretation of probability in Bohmian mechanics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 38(2):351–370.

Deutsch, D. (1999). Quantum theory of probability and decisions. *Proceedings: Mathematical, Physical and Engineering Sciences*, 455(1988):3129–3137.

Dürr, D., Goldstein, S., and Zanghì, N. (1992). Quantum equilibrium and the origin of absolute uncertainty. *Journal of Statistical Physics*, 67(5-6):843–907.

Gell-Mann, M. and Hartle, J. B. (1990). Quantum mechanics in the light of quantum cosmology. In Zurek, W. H., editor, *Complexity, Entropy and the Physics of Information*, page 425–459. Addison-Wesley, Redwood City.

Gell-Mann, M. and Hartle, J. B. (2012). Decoherent histories quantum mechanics with one real fine-grained history. *Physical Review A*, 85(6):062120.

Gill, R. D. (2005). On an argument of david deutsch. In Schürmann, M. and Franz, U., editors, *Quantum Probability and Infinite Dimensional Analysis: From Foundations to Applications*, QP—PQ: Quantum Probability and White Noise Analysis, pages 277–292. World Scientific.

Gleason, A. M. (1957). Measures on the closed subspaces of a Hilbert space. *Journal of Mathematics and Mechanics*, 6(6):885–893.

Hall, N. (2004). Two mistakes about credence and chance. *Australasian Journal of Philosophy*, 82(1):93–111.

Hartle, J. B. (2010). Quasiclassical realms. In Saunders, S., Barrett, J., Kent, A., and Wallace, D., editors, *Many Worlds? Everett, Quantum Theory, & Reality*, page 73–98. Oxford University Press, Oxford.

Ismael, J. (2008). Raid! Dissolving the big, bad bug. *Noûs*, 42(2):292–307.

Kochen, S. (2015). A reconstruction of quantum mechanics. *Foundations of Physics*, 45(5):557–590.

Kochen, S. and Specker, E. P. (1975). The problem of hidden variables in quantum mechanics. In *The Logico-Algebraic Approach to Quantum Mechanics: Volume I: Historical Evolution*, pages 293–328. Springer.

Lewis, D. (1980). A subjectivist's guide to objective chance. In Jeffrey, R. C., editor, *Studies in Inductive Logic and Probability*, volume 2, pages 263–294. University of California Press, Berkeley.

Lewis, P. J. (2007). Uncertainty and probability for branching selves. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 38(1):1–14.

Norsen, T. (2014). The pilot-wave perspective on spin. *American Journal of Physics*, 82(4):337–348.

Norsen, T. (2018). On the explanation of born-rule statistics in the de broglie-bohm pilot-wave theory. *Entropy*, 20(6):422.

Papineau, D. (1996). Many minds are no worse than one. *The British Journal for the Philosophy of Science*, 47(2):233–241.

Pettigrew, R. (2012). Accuracy, chance, and the principal principle. *Philosophical Review*, 121(2):241–275.

Read, J. (2018). In defence of Everettian decision theory. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*.

Rédei, M. (1998). *Quantum Logic in Algebraic Approach*. Kluwer Academic Publishers, Dordrecht.

Romano, D. (2016). Bohmian classical limit in bounded regions. In Felline, L., Ledda, A., Paoli, F., and Rossanese, E., editors, *New Directions in Logic and Philosophy of Science*. Lightning Source, Milton Keynes.

Romano, D. (2020). Multi-field and Bohm's theory. *Synthese*. `https://doi.org/10.1007/s11229-020-02737-6`.

Rosaler, J. (2016). Interpretation neutrality in the classical domain of quantum theory. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 53:54–72.

Saunders, S. (2004). Derivation of the Born rule from operational assumptions. *Proceedings: Mathematical, Physical and Engineering Sciences*, 460(2046):1771–1788.

Saunders, S. (2010). Chance in the Everett interpretation. In Saunders, S., Barrett, J., Kent, A., and Wallace, D., editors, *Many Worlds? Everett, Quantum Theory, & Reality*, book section 6, page 181–205. Oxford University Press, Oxford.

Saunders, S. and Wallace, D. (2008). Branching and uncertainty. *The British Journal for the Philosophy of Science*, 59(3):293–305.

Schlosshauer, M. (2007). *Decoherence and the Quantum-to-Classical Transition.* Springer Science & Business Media.

Sebens, C. T. and Carroll, S. M. (2018). Self-locating uncertainty and the origin of probability in Everettian quantum mechanics. *The British Journal for the Philosophy of Science*, 69(1):25–74.

Stinespring, W. F. (1955). Positive functions on C*-algebras. *Proceedings of the American Mathematical Society*, 6(2):211–216.

Tappenden, P. (2011). Evidence and uncertainty in Everett's multiverse. *British Journal for the Philosophy of Science*, 62(1):99–123.

Valentini, A. (2019). Foundations of statistical mechanics and the status of the born rule in de Broglie-Bohm pilot-wave theory.

Valentini, A. (2020). Foundations of statistical mechanics and the status of the born rule in de broglie-bohm pilot-wave theory. In Allori, V., editor, *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature*, page 423–478. World Scientific, Singapore.

van Fraassen, B. C. (1980). *The Scientific Image.* Clarendon Press.

Wallace, D. (2003). Everettian rationality: defending Deutsch's approach to probability in the Everett interpretation. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 34(3):415–439.

Wallace, D. (2012). *The Emergent Multiverse: Quantum Theory According to the Everett Interpretation.* Oxford University Press.

Wallace, D. (2019). Isolated systems and their symmetries, part II: local and global symmetries of field theories. `http://philsci-archive.pitt.edu/16624/`.

Wallace, D. (2020). On the plurality of quantum theories: Quantum theory as a framework, and its implications for the quantum measurement problem. In French, S. and Saatsi, J., editors, *Scientific Realism and the Quantum*, book section 5, page 78–102. Oxford University Press, Oxford.

Zurek, W. H. (2005). Probabilities from entanglement, Born's rule $p_k = |\psi_k|^2$ from envariance. *Physical Review A*, 71(5):052105.

Zurek, W. H. (2009). Quantum Darwinism. *Nature Physics*, 5(3):181.