# Understanding sequential measurements in $\psi$-epistemic ontological models

by

Joshua B. Ruebeck

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Masters of Science
in
Physics

Waterloo, Ontario, Canada, 2019

## Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

## Abstract

Since the famous debates of Einstein & Bohr, physicists have argued about the nature of the quantum state. Is it best thought of as describing the-way-things-are out there in the world, or merely as a description of our knowledge of such things? In more recent years, this has been termed the $\psi$-epistemic/$\psi$-ontic debate, and that distinction has been given a mathematical definition within the ontological models formalism. This formalism is a framework for describing a large class of interpretations of quantum mechanics. Here we show that consideration of sequential measurements, and the fact that the quantum state changes during measurements, has been a neglected topic in this area as it places nontrivial restrictions on the structure of $\psi$-epistemic ontological models.

We do this by finding a general restriction on the structure of $\psi$-epistemic models, although not one that is strong enough to rule them out categorically. We then apply this restriction to all of the known examples of $\psi$-epistemic ontological models and show that they can't represent sequential measurements. We also present a new version of the ontological models formalism, based on hidden Markov models, which we develop briefly and describe how it may be useful for either proving a no-go theorem (i.e. '$\psi$-epistemic models can't exist') or for learning more about the structure of $\psi$-epistemic models and how to construct them.

## Acknowledgements

This thesis could not have been written without the mentorship of Piers Lillystone, who took the time during his own PhD studies to introduce me to quantum foundations and to help me develop the ideas presented in this thesis (sorry for all of the Venn diagrams). I would also like to thank my supervisor Joseph Emerson for giving me the opportunity to study these topics for the last two years.

I almost certainly would not have made it through my graduate studies without the moral support of all of the friends I've made in my time at Waterloo and my long-distance friends who provided their support via text and video chats. My family, despite deciding that I am now less employable than my musical-theater-major little brother, has been so supportive of my education throughout my life and continues to be an anchor as I figure out what comes next.

More professionally speaking, I would like to thank Rob Spekkens for detailed feedback on this thesis, as well as Matt Leifer, David Schmid, Jim Crutchfield, and Samuel Loomis for taking the time to discuss this project with me and share their expertise at various points during its development. I am also thankful for financial support from an Ontario Graduate Scholarship, the University of Waterloo, the Canada First Research Excellence Fund, and the NSERC Discovery program.

# Table of Contents

# List of Symbols

| | |
|---|---|
| $\psi, \phi, \alpha, \beta$ | Quantum states |
| $[\psi]$ | Projector onto $\|\psi\rangle$ (shorthand for $\|\psi\rangle\langle\psi\|$) |
| $\rho$ | Density matrix |
| $M_k$ | The $k$th operator of measurement $M$ |
| $\Pi$ | A projective measurement operator |
| $U$ | Unitary transformation |
| | |
| $P$ | Preparation |
| $T$ | Transformation |
| $M$ | Measurement |
| $U_T$ | The unitary transformation corresponding to an operational transformation $T$ |
| $\rho_P$ | The density matrix corresponding to an operational preparation $P$ |
| $P_\psi$ | A (nonunique) preparation corresponding to the quantum state $\|\psi\rangle$ |
| $T_U$ | A (nonunique) transformation corresponding to the unitary $U$ |
| $M_{\Pi_i}$ | A (nonunique) measurement corresponding to the measurement operators $\{\Pi_i\}$ |
| $k$ | An outcome of an experiment |
| $a$ | Action (either a preparation, transformation, or measurement) |
| $\mathrm{Pr}_Q$ | The probability distribution associated with operational quantum theory |
| $\boldsymbol{P}$ | The set of all preparations; sim. for $\boldsymbol{T}, \boldsymbol{M}, \boldsymbol{K}, \boldsymbol{A}$ |
| | |
| $\overleftrightarrow{k}$ | String of outcomes (sim. for actions $\overleftrightarrow{a}$) |
| $\overleftarrow{k}$ | String of past outcomes |
| $\overset{\bullet\!\!\rightarrow}{k}$ | String of future outcomes, including present |
| $\overset{\circ\!\!\rightarrow}{k}$ | String of future outcomes, excluding present |

| | |
|---|---|
| $\lambda$ | Ontic state |
| $\mathbf{\Lambda}$ | Ontic state space |
| $\mu$ | Preparation distribution |
| $\Gamma$ | Transition matrix |
| $\xi$ | Response function |
| $\eta$ | State update rule |
| Supp | Support of a probability distribution |
| $\Delta_\psi$ | Support of a quantum state $\psi$ |
| | |
| $\Theta$ | Heaviside theta function |
| $\delta$ | Dirac delta function or Kronecker delta, depending on context |
| $\vec{\psi}$ | Bloch vector corresponding to the quantum state $\lvert\psi\rangle$ |
| $R_U$ | Rotation of the Bloch sphere corresponding to the unitary $U$ |
| $\mathcal{PH}^{d-1}$ | Projective Hilbert space of dimension $d-1$ |
| | |
| $Z, X$ | Generalized $Z, X$ operators |
| $\omega$ | root of unity |
| $[\cdot,\cdot]$ | Symplectic inner product |
| $T_{(\mathbf{x},\mathbf{z})} = T_\lambda$ | Generalized Pauli operators |
| $A_{(\mathbf{x},\mathbf{z})} = A_\lambda$ | Phase point operators |
| | |
| $X, \boldsymbol{X}, x$ | A random variable, its alphabet, and an element of its alphabet |
| $H$ | Classical Shannon entropy |
| $I$ | Classical mutual information (Shannon) |
| $E, E_{\overleftrightarrow{A}}$ | Excess entropy |
| $C, C_{\overleftarrow{A}}$ | Statistical complexity |
| $\mathcal{E}, \mathcal{E}_{\overleftarrow{A}}$ | Epistemicity |

# List of Figures

# Chapter 1

# Introduction

> ***Phaedrus:*** *What is the other principle, Socrates?*
> ***Socrates:*** *The second principle is that of division into species according to the natural formation, where the joint is, not breaking any part as a bad carver might.*
>
> – Plato, *Phaedrus*; trans. Benjamin Jowett

Plato gave us many things; among them, some passages ripe for overquoting out of context. The paraphrase "cut nature at the joints" has become a favorite of both scientists and philosophers over the millennia since it was written, and not for bad reason. We physicists especially like to think of ourselves as artist-meets-craftsperson, like a (good) carver. We like to think that nature has joints where it ought to be divided, and we like to think that we can find and divide those joints.

While philosophers have always argued about interpretations of scientific theories and whether or not they've cut any joints at all, many physicists were generally pretty happy with their progress until quantum physics came along at the turn of the 20th century. Suddenly, the joints moved out of focus, shifted, and, to some, disappeared completely. Back in the good old days of classical physics, billiard balls with mass but no spacial extent were strange, but considered idealizations at worst; the electric field may travel through no medium but space, but we got used to that. Perhaps one exception is the action-at-a-distance of Newtonian gravity, which Newton himself didn't like [44], but which was eventually resolved by general relativity.

But quantum physics! A particle that's 'here and there but also neither,' or a cat that's both dead and alive—clear nonsense. One of the first popular interpretations, and one that still throws its weight around today, simply said that 'our human, mesoscopic, classical language can't describe the goings-on of microscopic, quantum entities, and that's that.' This was Heisenberg's interpretation [90], one of a number of interpretations grouped under the heading of the Copenhagen Interpretation(s). While many folks agree with most of his assessment—that statements like 'here and there but also neither' are indeed nonsense[1]—it would be awfully disappointing if 'and that's that' really is the full extent of what we can say about microscopic physics.

So, fuelled by a mixture of faith, hope, and optimism, physicists and philosophers have been toiling away for the last hundred years to make some sense of that nonsense, birthing the field that is the subject of this thesis, quantum foundations. We continue to look for the natural joints, whether or not they actually exist. Many proposals have been put forward, and thoroughly developed by their proponents: Bohr's complementarity [14], de Broglie and Bohm's respective pilot wave theories [13, 85], Everett's universal wavefunction [29], the many-worlds interpretation [23, 24], various modal interpretations [25, 59, 86], collapse models [10, 67], QBism [36], relational quantum mechanics [71], the $\psi$-complete interpretation [12], pragmatist interpretations [48], non-interpretations [37], and more. This is not to mention aspirational interpretations that are based on toy theories and desiderata that have not yet been satisfied [5, 6, 26, 27, 47, 50, 53, 72, 81]. Indeed, anyone who asks 100 physicists for their personal preferred interpretation of quantum theory had better be prepared to hear at least 101 answers [77]. A review of this interpretation zoo is far beyond the scope of this thesis, and to my knowledge has not been carried out. In its absence, the Stanford Encyclopedia of Philosophy and Wikipedia provide useful starting points for the interested reader.

As if the chaos and diversity of this zoo weren't evidence enough that our joint-cutting is going badly, we can also look outside quantum theory to its archnemesis and chief rival, general relativity. A common catch-phrase of physicists is that 'quantum mechanics and general relativity are incompatible.' What they generally mean by this is that if one takes the chunk of math that is general relativity, picks up the lump of math that is quantum mechanics, and then attempts to smush them together, then untamable infinities, ghost particles, and other nonsensical entities emerge. This is generally seen as a pragmatic issue, where we would like to model as much of the universe as we can via math and physics, and we expect that a unified theory of quantum gravity will model more of the universe than either quantum theory or general relativity does on its own.

However, the issue goes much deeper: quantum theory and general relativity seem to be *metaphysically incompatible*. In other words, they divide nature at different joints. For

[1] To emphasize, I think that the constant use of phrases like "here and there at the same time" in both the popular press and formal scientific writing does nobody any favors and communicates very little except a demonstration that our language really is ill-equipped to describe whatever these goings-on really are.

example, the word *time* means wildly different things in each of these theories. As such, they are not just incompatible in the mathematical sense but perhaps also incommensurable in the sense of Kuhn, Feyerabend, and others [33, 54]. This assertion is, of course, dependent on one's interpretations of both general relativity and quantum theory; general relativity has essentially a single mainstream realist interpretation, and none of the popular realist interpretations for quantum theory are compatible with it. The only existing interpretations that are metaphysically compatible are the instrumentalist readings of each, which essentially amount to Heisenberg's 'and that's that.' To emphasize, this metaphysical incompatibility is not solely motivated by the desire for a unified theory of quantum gravity; whether or not one believes such a thing exists or is achievable, it would at the very least be desirable for our theories to give us a coherent picture of the universe, even if they are mathematically incompatible.

That said, we can turn this line of reasoning around. Just as we pursue the natural joints of the world despite any guarantee of their existence, we can do the same with a unified theory of physics. In this pursuit, the perspective of the previous paragraph has another reading: it should be no surprise that math-smushing was a failed exercise, when the real-world referents of that math were incompatible to begin with. Thus quantum foundations can serve as a resource in our search for a unified theory, if we approach it intentionally as such. Hardy suggests something along these lines [44], but moves too quickly (for me) to his own interpretation and a mathematization of these ideas without exploring sufficiently the role that existing interpretations and frameworks may play in this project. In particular, I think that any sort of pragmatic, quantum-gravity-oriented foundational program will have to be pluralist in its outlook. By this I mean that, rather than arguing about which interpretation is 'true' (after all, as Hardy and others point out, we can be fairly certain that none of them are), we can instead ask "What does each of these interpretations teach us about quantum theory and the world?" This recognizes that, even if none of the interpretations are 'true,' probably each one has something informative to say about quantum theory or its relationship to the world.

Returning to the interpretation zoo, one way in which we can make practical use (and some modicum of sense) of this zoo is to search for commonalities that allow us to categorize the interpretations in a variety of ways. One such way of categorizing them is by their response to the question "Is the quantum state real?" Discussion of this question can be dated back to the famous debates of Bohr and Einstein, but the modern version of this debate was largely kicked off by [47]. As always, 'real' is not entirely well-defined, but in this context most folks have in mind some notion of objective, observer-independent, 'out-there-in-the-world' existence. Cabello [15] provides the most up-to-date taxonomy based on interpretations' answer to this question; while most of its content is not novel, it makes

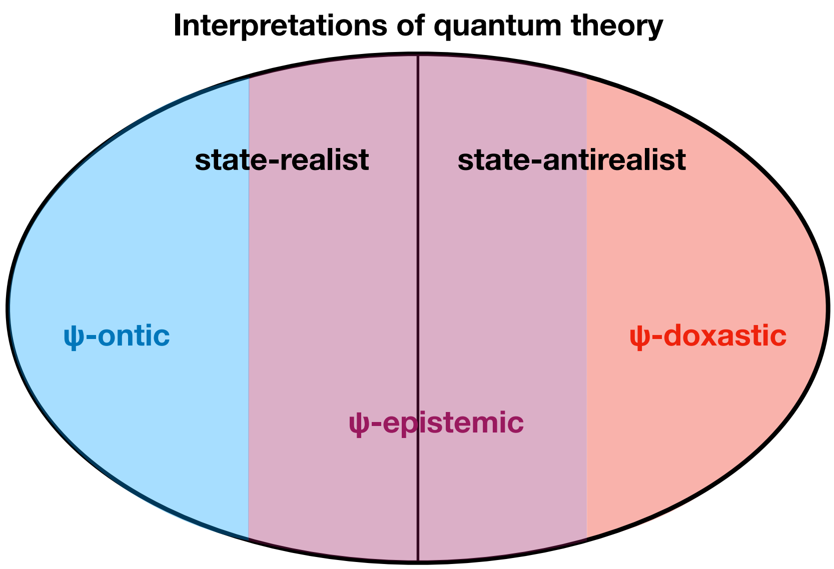



Figure 1.1: A Venn diagram categorizing interpretations of quantum theory, based on Cabello's classification scheme. I have renamed his "type-I" and "type-II" as state-realist and state-antirealist, respectively, which gains the advantage of specificity but of course loses the fuzziness that can be helpful in capturing a wider range of interpretations. These labels refer to whether or not an interpretation maintains that states (that is, *any* kind of states, not just quantum ones) have an objective existence; they are divided in the diagram by the black line. Additionally, Cabello reserves $\psi$-ontic (the quantum state is real) and $\psi$-epistemic (the quantum state is a state of knowledge) for state-realist interpretations, while I have extended $\psi$-epistemic into the state-antirealist region and used Schack's term, $\psi$-doxastic, for interpretations that claim that the quantum state is a subjective expression of belief.

a few important distinctions that are relevant to the present work. I have adapted this taxonomy with some of my own terminology and some of Schack's [76], and represented it as the Venn diagram in Fig. 1.1. In addition to their stance on the reality of the quantum state, interpretations are also distinguished by their stance on the reality of *any* potential state describing a system. I term these two categories *state-realist* and *state-antirealist*. Of course, although the categories in the Venn diagram are drawn as mutually exclusive and complete, any attempt at a categorization of a set as fuzzy as the interpretation zoo will not actually be so.

The question "Is the quantum state real?" then has three answers that have appeared in the literature, though more are probably possible. The first, "Yes, quantum states are objective elements of the world" corresponds to $\psi$-ontic[2] interpretations. This category is the most straightforward and requires little explication; examples include pilot wave theories, many-worlds, and the $\psi$-complete interpretation. This third interpretation, while not widely discussed, says that the (single-world) quantum state is in fact all that there is; other $\psi$-ontic interpretations generally specify more machinery.

The second answer that appears in the literature is "No, the quantum state represents *knowledge* about reality." Interpretations taking this view are variously termed epistemic, statistical, or $\psi$-epistemic, though we will use the latter here. The obvious next question is "Knowledge about what?" and different $\psi$-epistemic interpretations answer said question in a few different ways. Most state-realist $\psi$-epistemic interpretations say that the quantum state represents knowledge about that non-quantum state that they believe is real. State-antirealist $\psi$-epistemic interpretations must, ipso facto, be a little more creative about their answer to the latter question. One example is instrumentalist interpretations like Copenhagen interpretations and non-interpretations, which say that the quantum state just represents knowledge about future experimental results, and nothing more.

$\psi$-doxastic interpretations provide the third potential answer: "No, the quantum state represents the *beliefs* of agents." As mentioned above, the border between these and state-antirealist $\psi$-epistemic interpretations is a fuzzy one. In fact, the only interpretation that Cabello puts into this final category (and, as far as I know, the only interpretation that labels itself as such) is QBism. The advantage here is to avoid the question "Knowledge about what?" although many disadvantages are incurred in the process. In particular, QBism has been accused of being solipsistic and overly anti-realist in general. They deny both, instead arguing for "participatory realism" about structural elements of quantum theory like Born's rule [35, 83].

[2]Etymologies: $\psi$ is the standard symbol for a wavefunction or quantum state, and so is used here to refer to the quantum state. *-ontic* comes from the Ancient Greek *on*, meaning *being*, *existence*, *essence*; *-epistemic* from *episteme*, meaning *knowledge*; and *-doxastic* from *doxa*, meaning *belief.*

In order to clarify these positions, it might help to consider them in analogy to the everyday experience of flipping a coin and covering it as soon as it lands. Given that we don't know how it landed, we might assign it a 50% chance of being heads-up and a 50% chance of being tails-up. This probability distribution represents our *knowledge* of the coin. $\psi$-epistemic interpretations maintain that the quantum state is like this probability distribution. A state-realist will say that, before we look at it, the coin still has some definite, objective state; in this case, either heads or tails. $\psi$-ontic interpretations then say that the proper analogy for the quantum state is this objective state, rather than the probability distribution. State-antirealists, on the other hand, say that there is no fact of the matter about the coin, no real state[3]. $\psi$-doxastic interpretations would then say that the quantum state properly corresponds to the probability distribution, but viewed as a description of belief (usually expressed as a willingness to bet on the outcome of the coin-flip) rather than knowledge.

The present work is concerned with state-realist $\psi$-epistemic interpretations, not out of some conviction that states must be real, but because that class allows for a quantitative mathematical treatment via the ontological models formalism [45, 47, 72, 80]. The ontological models formalism, described in Chapter 2, captures almost all state-realist interpretations, with the exception of many-worlds interpretations and those that reject classical probability theory as a way to describe the behavior of these states. Generalized Probabilistic Theories (GPTs) [8], while not usually viewed as interpretations, are the only examples that I know of the latter case. Many-worlds interpretations are $\psi$-ontic; GPTs, not really being interpretations per say, are agnostic to the $\psi$-epistemic/$\psi$-ontic distinction. As such, all examples of state-realist $\psi$-epistemic interpretations that I am aware of are describable by the ontological models formalism, so for convenience I will use "state-realist $\psi$-epistemic interpretations" as synonymous with "$\psi$-epistemic interpretations that can be described in the ontological models formalism." The ontological models formalism, as a framework that describes a large number of interpretations, is a tool that is well-suited for pursing the pluralist/generalist project that I described earlier. In particular, it gives us a precise mathematical criterion distinguishing between $\psi$-ontic and $\psi$-epistemic interpretations, which allows us to really discuss them categorically, rather than as collections of examples.

State-realist $\psi$-epistemic interpretations are a very attractive class of interpretations, as sort of the best-of-all-worlds compromise. $\psi$-ontic interpretations, while often straight-

[3]The analogy breaks down a little bit here—there is debate about whether state-(anti)realism at the quantum level implies state-(anti)realism at the macroscopic level of coins. Most people, regardless of their interpretation of quantum theory, would agree that in the case of the coin there is a fact-of-the-matter about the coin before it is observed.

forward, suffer from a number of known issues including the measurement problem, unexplained physical collapse of wavefunctions, nonlocality, and others. Many of the desirable features of $\psi$-ontic interpretations can be retained by loosening to state-realism rather than $\psi$-onticity specifically, although we do know from Bell's theorem and subsequent experiments that any state-realist interpretation must be manifestly nonlocal [11]. On the other hand, there are many qualitative arguments for why quantum mechanics has the character of a theory describing knowledge [5, 27, 34, 81], which suggests that $\psi$-epistemic interpretations provide a very natural understanding of quantum theory.

There is one catch that I haven't mentioned: there is a severe dearth of actual state-realist $\psi$-epistemic interpretations. There are some nicely behaved and well-developed ones that only describe certain subtheories of quantum theory (i.e. they disallow certain experimental sequences), are toy theories that are qualitatively similar but nonidentical to quantum theory, or can only describe a qubit (a two-level quantum system). For a few years, there was doubt as to whether or not such interpretations exist at all for full quantum theory. Eventually a couple of examples were constructed that apply to the whole of quantum theory, but they are rather unnatural and have no basis in anything physical; they are based solely on mathematical abstractions. However, as proofs of concept, these examples aim to show that state-realist $\psi$-epistemic interpretations do exist in some form.

One of the main contributions of this work is to show that this last conclusion is no longer warranted. With the exception of a couple of specific interpretations [17, 58, 64, 81] and two other studies [16, 52], all work in the ontological models formalism to date has considered experiments that involve only a single measurement. Most relevantly, all examples of state-realist $\psi$-epistemic interpretations (in dimension $d \geq 3$) only consider a single measurement. We extend the formalism in order to include the possibility of sequential measurements, and in doing so show that the existing examples of state-realistic $\psi$-epistemic interpretations for full quantum theory are no longer valid.

Importantly, this is not a categorical result like "There are no state-realistic $\psi$-epistemic interpretations of quantum theory." Such a result is called a no-go theorem; the most famous example is Bell's theorem [11] which says that any state-realist interpretation must be nonlocal; i.e. for certain experimental setups, a state must extend beyond the lightcone of the event it is describing[4]. Such no-go theorems exist for $\psi$-epistemic interpretations, but all make one additional assumption beyond the basic structure of the ontological models formalism. This additional assumption can be interpreted as the negation of a property that $\psi$-epistemic models are required to have.

[4] More specifically, the causal influence corresponding to a state update under measurement must travel faster than the speed of light, but this is not the usual phrasing of this property.

For example, the first no-go theorem that was explicitly concerned with the $\psi$-epistemic/$\psi$-ontic distinction[5] was the PBR theorem due to Pusey, Barrett, and Rudolph [68]. The PBR theorem's extra assumption is called the preparation independence postulate (PIP), which roughly says that, for two systems whose quantum states are uncorrelated, the real states of the system must also be uncorrelated. It shows that, under the addition of the PIP, no $\psi$-epistemic interpretation can exist. Equivalently, any existing $\psi$-epistemic interpretation must violate the preparation independence postulate. Thus, although these types of results are usually presented as ruling out some type of $\psi$-epistemic interpretation, they are perhaps more properly understood as restrictions on the structure thereof. Because of the (previously accepted) existence of $\psi$-epistemic interpretations, it was understood that any no-go theorem needs an additional assumption. Other no-go theorems [1, 19, 43, 61] have been proven for other additional assumptions/properties and are reviewed thoroughly, along with the rest of the $\psi$-epistemic/$\psi$-ontic debate, in [55].

The results of this thesis differ in two ways from these no-go theorems. First, as already stated, we do not make a categorical statement, but evaluate a set of examples. Second, we do not make an additional assumption to do so[6]. Because we rule out *all* known examples in this fashion, we reopen the possibility of a no-go theorem for state-realist $\psi$-epistemic interpretations as long as that theorem takes into account sequential measurements. This result is interesting, because it was previously the accepted wisdom that "A straightforward resolution of the collapse of the wavefunction, the measurement problem, Schrödinger's cat and friends is one of the main advantages of $\psi$-epistemic interpretations" [55]. All of the 'paradoxes' listed here have to do with how the quantum state changes during measurement, which is only important when considering sequential measurements. While this accepted wisdom may still be true for state-antirealist interpretations, we will show that within the class of state-realist interpretations, those that are $\psi$-epistemic are actually at a *dis*advantage when it comes to describing state update under measurement relative to $\psi$-ontic interpretations. Thus, while these 'paradoxes' are obviously still an issue for $\psi$-ontic interpretations, state-realist $\psi$-epistemic interpretations do not necessarily resolve them in the straightforward way expressed in the quote above.

While reopening the possibility of a no-extra-assumptions no-go theorem is exciting, it would of course be even more exciting to prove such a theorem. In Chapter 5, we'll

[5] Bell's theorem is the first (valid) no-go theorem of any kind in quantum theory, and can be read as ruling out certain types of $\psi$-epistemic interpretations. The PBR theorem can also be read as having less to do with $\psi$-epistemic interpretations and more directly with the PIP [28].

[6] Although it seems fairly self-evident to myself and others that modeling sequential measurement is not an additional assumption, we further justify this position in the conclusion (Chapter 6) for the benefit of any naysayers.

draw from computational mechanics [21, 78] to construct a new set of tools for attempting this, and describe our progress so far along this track. It is, of course, still possible that state-realist $\psi$-epistemic models *do* exist—luckily, anything we learn using computational mechanics that doesn't support a no-go theorem will most likely help in the construction of a $\psi$-epistemic model, so this new framework is interesting regardless of the actual final answer to the question, "is the quantum state real?" This is especially so because other properties of interest to the quantum foundations community, like contextuality, also have nice expressions in this framework.

Of course, regardless of this hypothetical result, it was clear before the present work and even clearer after it that state-realist $\psi$-epistemic interpretations are not as natural an explanation of quantum theory as was previously thought—all of the previously known no-go theorems point towards unnatural constructions at best. As such, it is already worth thinking about what it might mean to move outside of the ontological models formalism. Computational mechanics also provides some resources for this line of thought; we'll discuss some potential avenues, along with their implications, in Chapter 6.

Chapters 2–4 appeared in [74] (authored with Piers Lillystone and Joseph Emerson), with the exception of Sections 2.2, 2.7, 3.2.2, and 4.3. The sections that did appear in [74] have been expanded, some extensively.

# Chapter 2

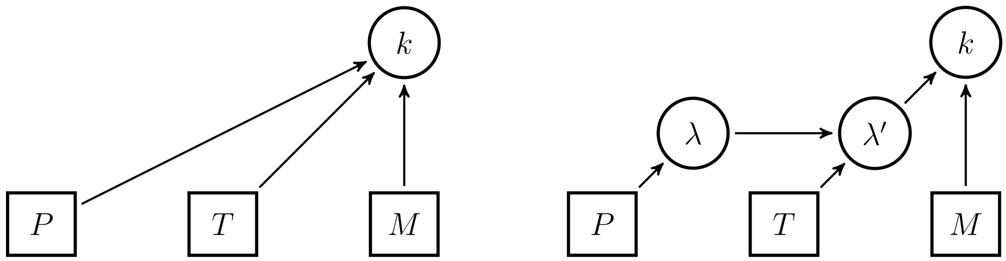
# Extending the ontological models formalism for sequential measurements

> *...and the judge wasn't gonna look at the twenty-seven 8 by 10 colored glossy pictures with the circles and arrows and a paragraph on the back of each one explainin' what each one was to be used as evidence against us...*
>
> – Arlo Guthrie, *Alice's Restaurant*

We begin by reviewing the standard treatment of the ontological models formalism for prepare-and-measure-once experiments (Sections 2.1–2.3). We then add in a description of state update under measurement and reformulate the entire framework from a stochastic processes/hidden Markov models point of view (Sections 2.4–2.6). This formalizes and motivates our definition of state update while simultaneously clarifying a number of the underlying assumptions of the ontological models formalism. We close by including some remarks pointing towards interesting consequences for current notions of contextuality when including sequential measurements (Section 2.7).

## 2.1 The ontological models formalism

Before we define ontological models, we need to define an *operational theory* [80]. In this framework, we describe an experiment via a preparation, a transformation, and a




(a) An operational theory

(b) An ontological model of the operational theory

Figure 2.1: Influence diagrams for (a) operational theories and (b) ontological models. The boxes represent choices made by the experimenter, circles represent uncontrolled random variables, and arrows represent possible causal influences. Time proceeds from left to right. In this particular case, $P, T$, and $M$ represent an experimenter's choice of preparation, transformation, and measurement, respectively. $k$ represents the outcome of the experiment, while $\lambda$ and $\lambda'$ are the ontic state of the system at two different times.

measurement. The operational theory must first specify a set $\boldsymbol{P}$ of allowable preparations, a set $\boldsymbol{T}$ of allowable transformations, and a set $\boldsymbol{M}$ of allowable measurements. Then, for any $P \in \boldsymbol{P}$, $T \in \boldsymbol{T}$, and $M \in \boldsymbol{M}$, it must define the probability

$$\Pr(k|M, T, P) \tag{2.1}$$

of some outcome $k \in \boldsymbol{K}$ for a particular experiment. $\boldsymbol{K}$ is the set of possible outcomes, which we usually take to be $\mathbb{Z}$ for convenience. An influence diagram [51] of this setup is shown in Fig. 2.1a; as $k$ is the only uncontrolled random variable (circle), and it has causal arrows coming from all three of $P, T, M$, we see that Eq. 2.1 is exactly what's needed to specify the behavior of this influence diagram.

Note that the use of this framework is not limited to quantum theory; for example, $P$ could be flipping a coin into a box and covering it, $T$ could be shaking the box, and $M$ could be opening the box to see which side of the coin is face-up. This generality is one of the factors that allows us to draw such strong connections to the theory of stochastic processes in Section 2.6.

When we include transformations, this setting is called the *prepare-transform-measure* operational framework, and when we omit transformations it is the *prepare-measure* framework. Since later we will be considering multiple measurements, we refer to it here as the *prepare-and-measure-once* framework. Often we take $\boldsymbol{P}$ to be the set of pure quantum state preparations, $\boldsymbol{T}$ to be the full set unitary maps on a Hilbert space, and $\boldsymbol{M}$ to be all projective measurements on this Hilbert space. In this case, we say we are describing the full quantum theory[1]; in contrast, a *subtheory* is described by taking subsets of $\boldsymbol{P}, \boldsymbol{T}, \boldsymbol{M}$ for the full quantum theory. For example, in quantum information settings we often consider only measurements in the standard basis.

When modeling full (pure-state) quantum theory, we have the operational probability distribution

$$\Pr_Q(k|M,T,P) = \operatorname{tr}\Big(M_k U_T \rho_P U_T^\dagger M_k^\dagger\Big). \tag{2.2}$$

Here $\rho_P$ is the density matrix corresponding to the preparation $P$; $U_T$ is the unitary corresponding to the transformation $T$, and $M_k$ is the $k$th measurement operator of the measurement $M$.

An *ontological model* [45, 72, 80] supplements this operational point of view by asserting that a system has a state $\lambda$, called an *ontic state*. To specify an ontological model, we first choose an *ontic state space* $\boldsymbol{\Lambda}$, which is the set of all possible states that the system can attain. Then, preparations are described by a *preparation distribution* $\mu(\lambda|P)$, which is the probability of obtaining some ontic state $\lambda \in \boldsymbol{\Lambda}$ given the preparation $P$. Transformations are described by a *transition matrix* $\Gamma(\lambda'|\lambda, T)$, which is the probability of preparing a new state $\lambda'$ given the previous state $\lambda$ and the choice $T$ of transformation. Finally, measurements are represented by a *response function* $\xi(k|\lambda, M)$ which describes the probability of an outcome $k$ given the ontic state $\lambda$ and the choice of measurement $M$. We say that an ontological model successfully reproduces quantum theory if

$$\int_{\boldsymbol{\Lambda}} \mathrm{d}\lambda' \int_{\boldsymbol{\Lambda}} \mathrm{d}\lambda\, \xi(k|\lambda', M)\Gamma(\lambda'|\lambda, T)\mu(\lambda|P) = \Pr_Q(k|M,T,P) \tag{2.3}$$
$$\forall P \in \boldsymbol{P}, T \in \boldsymbol{T}, M \in \boldsymbol{M},$$

or, in a prepare-and-measure-once experiment,

$$\int_{\boldsymbol{\Lambda}} \mathrm{d}\lambda\, \xi(k|\lambda, M)\mu(\lambda|P) = \Pr_Q(k|M,P) \tag{2.4}$$
$$\forall P \in \boldsymbol{P}, M \in \boldsymbol{M}.$$

In both cases we've written $\Pr_Q$ for the quantum probability Eq. 2.2, but in general it would be whatever probability distribution is specified by the operational theory. Note that this consistency condition amounts to marginalizing over the ontic states $\lambda, \lambda'$—more on this later.

[1]Larger sets can be considered (e.g. including mixed states, CPTP maps, and/or non-projective measurements), but all of the models studied in this paper fit the given definition. In addition, a $\psi$-onticity theorem proven for pure quantum theory will hold for mixed-state quantum theory.

## 2.2 Contextuality

For the most part, this thesis is not explicitly concerned with contextuality [53, 80], but we give its definition(s) here since it is important to ensure that we do not assume noncontextuality. We will also present some preliminary observations about contextuality's relationship to sequential measurements in Section 2.7, although this line of investigation has not been fully developed. We focus on generalized contextuality [80] rather than Kochen-Specker contextuality because (a) accounting for potential contextuality in the generalized sense is sufficient to account for potential contextuality in the Kochen-Specker sense and (b) this is the definition that is affected by the inclusion of sequential measurements.

In [80], Spekkens defines two preparations $P, P'$ as *operationally equivalent* if they give the same measurement statistics:

$$P \underset{\text{op}}{\simeq} P' \iff \Pr(k|M, P) = \Pr(k|M, P') \qquad \forall M \in \boldsymbol{M}, k \in \boldsymbol{K}. \tag{2.5}$$

The operational equivalence of measurements and transformations is defined similarly but separately:

$$M \underset{\text{op}}{\simeq} M' \iff \Pr(k|M, P) = \Pr(k|M', P) \qquad \forall P \in \boldsymbol{P}, k \in \boldsymbol{K}. \tag{2.6}$$

$$T \underset{\text{op}}{\simeq} T' \iff \Pr(k|M, T, P) = \Pr(k|M, T', P) \qquad \forall P \in \boldsymbol{P}, M \in \boldsymbol{M}, k \in \boldsymbol{K}. \tag{2.7}$$

Roughly, noncontextuality is the condition that operational equivalence implies equivalence at the ontological level as well. A model is then *preparation noncontextual* if

$$P \underset{\text{op}}{\simeq} P' \implies \mu(\lambda|P) = \mu(\lambda|P') \qquad \forall P, P' \in \boldsymbol{P}. \tag{2.8}$$

Similarly, it is *measurement noncontextual* or *transformation noncontextual* if, respectively,

$$M \underset{\text{op}}{\simeq} M' \implies \xi(k|\lambda, M) = \xi(k|\lambda, M') \qquad \forall M, M' \in \boldsymbol{M}, \tag{2.9}$$

$$T \underset{\text{op}}{\simeq} T' \implies \Gamma(\lambda'|\lambda, T) = \Gamma(\lambda'|\lambda, T') \qquad \forall T, T' \in \boldsymbol{T}. \tag{2.10}$$

A model is *preparation contextual* if it is not preparation noncontextual, etc. We must make sure that we don't assume any kind of noncontextuality, as noncontextual models of quantum theory do not exist [80]. The accidental assumption of noncontextuality was the error in von Neumann's 'proof' that no ontological models (then called hidden variable theories) of quantum theory exist [89][2].

[2] An important historical note: this error was pointed out by Hermann in 1935 [49], though ignored by the physics community until it was re-discovered by Bell in 1966 [11].

## 2.3 $\psi$-epistemic ontological models

Within the ontological models formalism, we can define a set of precise criteria to distinguish $\psi$-epistemic interpretations from $\psi$-ontic interpretations. Consider first the $\psi$-epistemic criterion, proposed in [47] as a test for whether an interpretation admits at least some quantum states that are not uniquely determined by the underlying state of reality. We begin by defining the *support* $\mathrm{Supp}(f(\cdot))$ of a distribution as the set of ontic states on which it is nonzero. i.e., for a preparation distribution,

$$\mathrm{Supp}(\mu(\cdot|P)) = \{\lambda \in \mathbf{\Lambda} : \mu(\lambda|P) > 0\}. \tag{2.11}$$

Then, following [52], we account for potential preparation contextuality by defining

$$\Delta_\psi = \bigcup_{P_\psi \in \boldsymbol{P}_\psi} \mathrm{Supp}(\mu(\cdot|P_\psi)) \tag{2.12}$$

where $\boldsymbol{P}_\psi \subseteq \boldsymbol{P}$ is the set consisting of every possible preparation corresponding to the quantum state $|\psi\rangle$. We refer to $\Delta_\psi$ as the support of a quantum state $|\psi\rangle$, to distinguish it from the support of a particular preparation $P_\psi$. Then a pair of states $|\psi\rangle, |\phi\rangle$ is *ontologically distinct* in a particular model if $\Delta_\phi \cap \Delta_\psi = \emptyset$, and *ontologically indistinct* otherwise[3]. This leads us to the standard definition of a $\psi$-epistemic ontological model [47, 55]:

**Definition 1** ($\psi$-epistemic)**.** *An ontological model is $\psi$-epistemic if there exists a pair of states $|\psi\rangle, |\phi\rangle$ that are ontologically indistinct; i.e. $\exists\, |\psi\rangle, |\phi\rangle : \Delta_\phi \cap \Delta_\psi \neq \emptyset$.*

The motivation for this definition can be seen most clearly via Fig. 2.2. On the left hand side is a depiction of the ontic space for a $\psi$-epistemic model, where some of the supports overlap. On the right is the ontic space for a $\psi$-ontic model; since none of the supports overlap, they must partition the state space. As such, there is exactly one quantum state corresponding to each point in the ontic state space—even if it wasn't included explicitly in the ontology, one 'might as well' have named the ontic state ($\psi$, extra information).

As noted in [55], this definition is highly permissive in the sense that, if an ontological model were to contain only a single pair of quantum states that are ontologically indistinct,

[3] For this purposes of this paper we assume there exists a measure that is absolutely continuous with respect to all other measures in the ontological model. Therefore, we can work with probability densities, rather than the full measure-theoretic treatment. While this assumption is not strictly true in all of our models, it does not affect our results. A strict treatment of many of the models/definitions presented here can be found in [55].

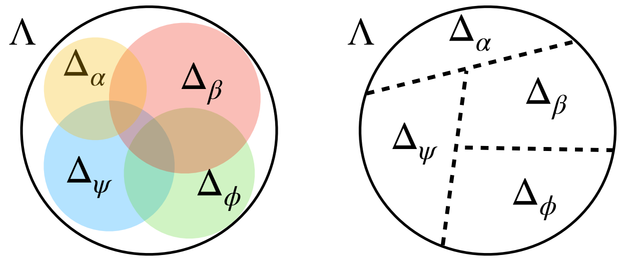



Figure 2.2: Venn diagrams for the ontic state space of a $\psi$-epistemic OM (left) and a $\psi$-ontic OM (right). The regions shown are the supports of four unique quantum states; although they are shown partitioning the state space in the $\psi$-ontic diagram, this is done for illustrative purposes and really only holds for the set of all quantum states.

then it would not achieve the full explanatory power expected of the $\psi$-epistemic viewpoint; this is exactly the case with the $\text{ABCL}_0$ model discussed in Section 3.2.5.

There are, however, proposals to strengthen the notion of $\psi$-epistemicity, two of which are relevant to our discussion [55, 63].

**Definition 2** (Pairwise $\psi$-epistemic)**.** *An ontological model is* pairwise $\psi$-epistemic *if, for all pairs* $|\psi\rangle, |\phi\rangle$ *of nonorthogonal quantum states,* $|\psi\rangle$ *and* $|\phi\rangle$ *are ontologically indistinct.*

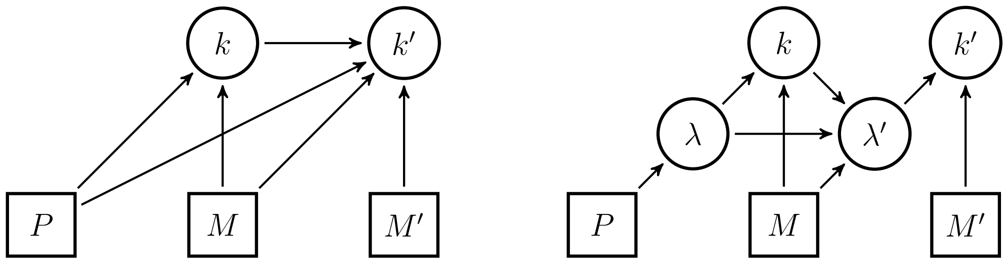
**Definition 3** (Never $\psi$-ontic)**.** *An ontological model is* never $\psi$-ontic *if every ontic state* $\lambda \in \mathbf{\Lambda}$ *is in the support of at least two quantum states:*

$$\forall \lambda \in \mathbf{\Lambda} : \exists \psi, \phi : \lambda \in \Delta_\psi \cap \Delta_\phi. \tag{2.13}$$

Note that both of these definitions imply the weaker notion of $\psi$-epistemicity, but are independent from one another.

## 2.4 Adding sequential measurements

The standard definitions of operational theories and ontological models involve a single measurement and a single measurement outcome despite the fact many important quantum experiments (e.g. Stern-Gerlach, double slit [75]) and quantum algorithms (e.g. measurement-based error correction [39]) involve sequential measurements. Thus we will refer to the usual definition of prepare-measure as prepare-and-measure-once experiments.




(a) An operational theory

(b) An ontological model of the operational theory

Figure 2.3: Influence diagrams for a prepare-and-measure-twice operational theory and its ontological models. We have essentially replaced the transformation in Fig. 2.1 with a measurement, and made the necessary adjustments of adding another outcome $k$ and more causal arrows as described in the text.

Here we introduce sequential measurements; an influence diagram for two measurements is shown in Fig. 2.3. It is described by the probability distribution

$$\Pr(k', k|M', M, P) = \Pr(k'|k, M', M, P)\Pr(k|M, P). \tag{2.14}$$

In the case of quantum mechanics, this operational probability becomes

$$\Pr(k', k|M', M, P) = \mathrm{tr}\Big(M'_{k'} M_k \rho_P M_k^\dagger {M'}_{k'}^\dagger\Big). \tag{2.15}$$

It is also typical in quantum foundations to describe measurements by positive operator valued measures (POVMs). However, they do not fully specify how a measurement updates a state. Although one can obtain a POVM $\{E_k\}$ from a set of generalized measurement operators $\{M_k\}$ by the relation $E_k = M_k^\dagger M_k$, the decomposition of $\{E_k\}$ into $\{M_k\}$ is not unique. Thus although $\boldsymbol{M}$ is often described by POVMs, consideration of state update requires that we specify generalized measurement operators instead.

As an example of when this is important, consider a coarse-graining of the measurement $\{M_k\} = \{[0], [1], [2]\}$, where we denote the projector onto a state

$$[\psi] = |\psi\rangle\langle\psi| \tag{2.16}$$

as in [55]. We can either coarse-grain coherently, i.e. measure $\{M'_k\} = \{[0] + [1], [2]\}$ or we can coarse-grain decoherently by measuring $\{M_k\}$ and then combining outcomes 0 and 1 into a single measurement result and 'forgetting' which one actually occurred. While these two processes are represented by the same POVM $\{E_k\} = \{[0]+[1],[2]\}$, their state update behavior is different: if the state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ is measured, it will stay the same in the coherent case or update to the mixed state $\frac{1}{2}([0]+[1])$ in the decoherent case.

We are now in a position to supplement the definition of an ontological model in order to model sequential measurement. In textbook quantum theory, the state updates during a measurement in a way that depends on the previous state, the measurement procedure, and the measurement outcome: Lüder's rule states that for a measurement outcome corresponding to a projector $\Pi_k$, a quantum state $|\psi\rangle$ will update during measurement to

$$|\psi'\rangle = \frac{\Pi_k |\psi\rangle}{\sqrt{\langle\psi|\Pi_k|\psi\rangle}}. \tag{2.17}$$

We allow for dependence on all of these things by representing this via a *state update rule* $\eta(\lambda'|k,\lambda,M)$. As with $\mu$, $\Gamma$, and $\xi$, we require that $\eta$ is normalized. Although this object looks very similar to the transition matrix for transformations, it is distinguished by two important features which we emphasize by choosing a new symbol to represent it.

The first distinction is simple, in that $\eta$ depends on a measurement outcome $k$, while $\Gamma$ does not; this is analogous to the fact that generally in quantum theory we can only implement measurement update maps probabilistically (i.e. by post-selecting on a not-necessarily-deterministic measurement outcome).

The second distinction is the fact that $\eta(\lambda'|k,\lambda,M)$ is not defined for all $\lambda \in \mathbf{\Lambda}$. Roughly speaking, it doesn't make sense to ask "What is the new state $\lambda'$ after measuring state $\lambda$ and obtaining outcome $k$?" if the outcome $k$ could not have occurred given the previous state $\lambda$. Thus $\eta(\lambda'|k,\lambda,M)$ is well-defined only for $\lambda \in \mathrm{Supp}(\xi(k|\cdot))$. This property of $\eta$ is analogous to the fact that Eq. (2.17) is only well-defined when its denominator $\langle\psi|\Pi_k|\psi\rangle \neq 0$, i.e. only when the measurement $\Pi_k$ could have responded to the pre-measurement quantum state $|\psi\rangle$.

This second distinction is central to the main result of this paper; by showing that $\eta$ is non-normalizable on some domain, we are able to conclude that that domain cannot be part of the support of $\xi$.

The consistency condition with quantum theory is given by equations similar to Eqs. 2.3 and 2.4 where we marginalize out $\lambda$; a formal statement will be postponed until Section 2.6, where we will also provide more formal treatments of the properties of $\eta$ and other extensions of the ontological models formalism discussed so far.

Although, as stated above, we do not assume noncontextuality of any kind, most of our results hold for a projector independent of its full measurement context. We are explicit about this when it is the case, and write $\xi(\Pi|\lambda)$ rather than $\xi(k=0|\lambda, M=\{\Pi,\ldots\})$ for notational convenience; $\eta(\lambda'|\lambda,\Pi)$ is defined similarly.

## 2.5 Some easy cases of state update rules

There are two cases in which, given a prepare-and-measure-once ontological model for quantum theory, we can always augment it with a state update rule for measurement. First, if we only include rank-1 projective measurements in the subtheory we're modeling, we can simply re-prepare in the measured (unique, pure) state:

$$\eta(\lambda'|k,\lambda,M_{\{\Pi_i\}}) = \mu(\lambda'|P_{\Pi_k}) \quad \text{for } \operatorname{tr}(\Pi_k)=1. \tag{2.18}$$

This is normalized for all $\lambda$ since $\mu$ is normalized, and faithfully reproduces quantum statistics since $\mu$ does. It is independent of the previous state $\lambda$, which, besides being unsatisfying in an explanatory sense, is also not possible in general (this follows from Section 4.1).

Second, it is quick to prove, again by construction, that $\psi$-ontic models can always be given a state update rule. Since there is a unique quantum state $|\psi_\lambda\rangle$ associated with every ontic state $\lambda$, we can define

$$\eta(\lambda'|k,\lambda,M) = \mu\left(\lambda' \middle| P = \frac{M_k[\psi_\lambda]M_k^\dagger}{\operatorname{tr}\left(M_k[\psi_\lambda]M_k^\dagger\right)}\right). \tag{2.19}$$

Again, normalization and faithfulness follow because $\mu$ has these properties. Note that this construction works for any kind of measurement, not just projective measurements.

These observations together suggest that in order to find anything interesting involving state update, we ought to examine higher-rank measurements in $\psi$-epistemic models. This suspicion will be confirmed by the main result of this paper, which applies to exactly these types of measurements and models.

## 2.6 A new perspective: ontological models as hidden Markov models of stochastic channels

In the context of sequential measurements, it is illuminating to motivate the definition of an ontological model from the point of view of the hidden Markov models (HMM) literature. We do this in order to (a) provide a rigorous treatment of sequential-measurement

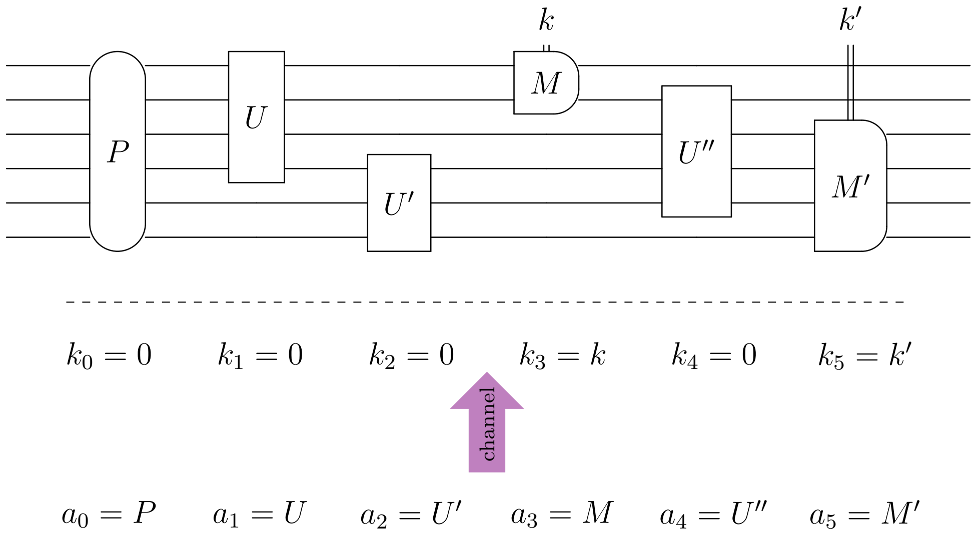


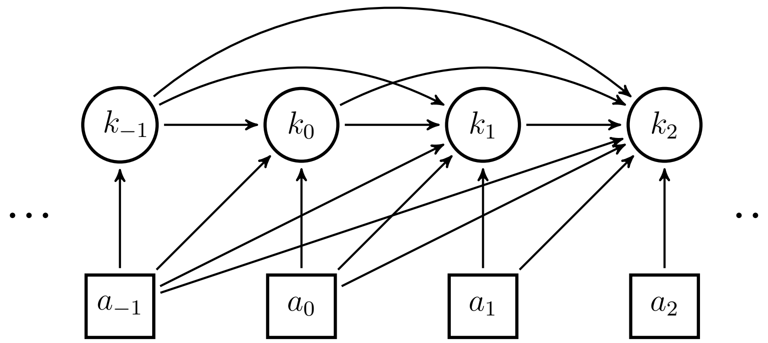

Figure 2.4: A quantum circuit can be pictured as a stochastic channel, as described in the text. The inputs to the channel $a_t$ are the choice of operation, and the outputs $k_t$ report the results of measurements (a trivial output $k_t = 0$ occurs if the input was a preparation or a transformation).

scenarios presented informally in Section 2.1 and (b) clarify the assumptions that define the ontological models framework. Additionally, this leads directly to the computational mechanics approach that we will describe in Chapter 5.

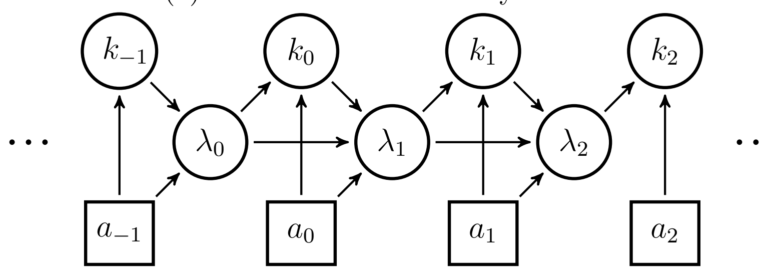
In an abstract sense, we picture a quantum circuit as a memoryful stochastic channel (Fig. 2.4). The channel that we often discuss with regards to a quantum circuit is the (quantum) channel that takes the input quantum state and maps it to the output quantum state. For present purposes, we will instead think of it as a channel, used repeatedly at each time step, from the experimenter to individual measurement outcomes. Pictorially, one might think of this as 'rotating the channel ninety degrees' in a circuit diagram.

Another, perhaps more concrete, presentation is shown in Fig. 2.5 and involves simply extending the influence diagrams in Fig. 2.3 to an arbitrary number of measurements. The only difference is that now we don't differentiate between preparations, transformations, and measurements, thinking of any of those as a choice of *action* $a$ on the system. Accordingly, there is an output for every input; if the input was a preparation or transformation, we specify a trivial, deterministic output $k = 0$ that carries no information.

For a moment, we will move to the usual setting of computational mechanics, which examines bi-infinite strings of input and output symbols (this is to anticipate further discus-




(a) A causal and stationary channel




(b) A hidden Markov model of the above channel

Figure 2.5: Influence diagrams for stochastic channels. We no longer distinguish between preparations, transformations, and measurements, referring to all of them as "actions" $a$. Additionally, we have extended to an arbitrary number of actions rather than jut two or three. Note that no arrows point backwards in time, which corresponds to causality (Def. 5). In the hidden Markov model the state $\lambda_t$ mediates all causal influences through time (Def. 6), which corresponds to the fact that no arrows move past the states $\lambda_t$ without going through them.

sion of computational mechanics in Chapter 5). As in the operational theories framework, our inputs and outputs could come from any alphabet; we denote the set of possible inputs/actions as $\boldsymbol{A}$, and the set of possible outputs as $\boldsymbol{K}$. For example, we could use $\boldsymbol{A} = \boldsymbol{K} = \{0, 1\}$ and thus describe binary input-output processes as in [7]. Later, we will restrict to semi-infinite sequences that begin with a preparation, as these are the input sequences that have well-defined behavior in quantum theory.

We denote the string of all inputs as $\overleftrightarrow{a_0} = \ldots a_{-2}a_{-1}a_0a_1a_2\ldots$ , and similarly the string of outputs $\overleftrightarrow{k_0} = \ldots k_{-2}k_{-1}k_0k_1k_2\ldots$. The subscript 0 indicates in both cases the time that we take as an origin/reference point. The channel is then described by the conditional probability distribution

$$\Pr(\overleftrightarrow{k_0}|\overleftrightarrow{a_0}). \tag{2.20}$$

This is the analog of Eq. (2.1) in the operational theories framework. Following [7], we denote substrings with $a_{t:t+L} = a_t a_{t+1} \ldots a_{t+L-1}$. We also denote the past $\overleftarrow{a_t} = a_{-\infty:t}$, future (including present) $\overset{\bullet\!\!\rightarrow}{a}_t = a_{t:\infty}$, and future (excluding present) $\overset{\circ\!\!\rightarrow}{a}_t = a_{t+1:\infty}$. We can now define two properties of stochastic channels [7]:

**Definition 4** (Stationary)**.** *A* stationary *channel is one that has time-translation symmetry, so statistics are not affected by our choice of time-origin:*

$$\begin{aligned} \Pr(k_{t:t+L}|\overleftrightarrow{a_t}) &= \Pr(k_{0:L}|\overleftrightarrow{a_0}) \qquad \textit{and} \\ \Pr(\overleftrightarrow{k_t}|\overleftrightarrow{a_t}) &= \Pr(\overleftrightarrow{k_0}|\overleftrightarrow{a_0}) \qquad \forall t, L, \overleftrightarrow{a}. \end{aligned} \tag{2.21}$$

**Definition 5** (Causal)**.** *A* causal *channel is one for which a finite output substring depends only on input symbols in its past:*

$$\Pr(k_{t:t+L}|\overleftrightarrow{a}) = \Pr(k_{t:t+L}|\overleftarrow{a}_{t+L}). \qquad \forall t, L, \overleftarrow{a} \tag{2.22}$$

It is shown in [7] that a channel satisfying these two properties can be specified entirely by the single-symbol recurrence relation

$$\Pr(k|a, \overleftarrow{a}, \overleftarrow{k}). \tag{2.23}$$

From here on, we drop the time-origin subscript because the condition of stationarity means that it is redundant. Note that this expression does not imply a Markov process, since it depends in general on the entire histories $\overleftarrow{a}$ and $\overleftarrow{k}$. All we mean by single-symbol is that we are not specifying the probabilities over the whole future, just a single output symbol. We now construct an HMM as follows:

**Definition 6** (Hidden Markov Model)**.** *A* hidden Markov model *(HMM) of a stationary, causal channel is specified by an additional random variable $\lambda$ taking values in a state space $\mathbf{\Lambda}$. It is given a joint probability distribution over $\overleftrightarrow{\lambda}, \overleftrightarrow{a}, \overleftrightarrow{k}$ so that the recurrence relation above (Eq. 2.23) becomes*

$$\Pr(k, \lambda'|a, \lambda, \overleftarrow{a}, \overleftarrow{k}, \overleftarrow{\lambda}) = \Pr(k, \lambda'|a, \lambda), \tag{2.24}$$

where we've used a prime to denote the next time step. In other words, the state $\lambda$ renders the future conditionally independent of the past and induces a Markov process over the state space $\mathbf{\Lambda}$ that mediates the channel statistics. An influence diagram [51] of a stationary, causal channel is shown in Figure 2.5 before and after the specification of an HMM.

To see that specification of an HMM as in Eq. 2.24 is equivalent to the definition of an ontological model given in Section 2.1, we factor the probability distribution from Eq. 2.24 and look separately at the cases where $a$ is a preparation, transformation, or measurement:

$$\begin{aligned} &\Pr(k, \lambda'|\lambda, a) \\ &\quad = \Pr(\lambda'|k, \lambda, a)\Pr(k|\lambda, a) \\ &\quad = \begin{cases} \mu(\lambda'|P)\delta_{k,0} & a = P \in \boldsymbol{P} \\ \Gamma(\lambda'|\lambda, T)\delta_{k,0} & a = T \in \boldsymbol{T} \\ \eta(\lambda'|k, \lambda, M)\xi(k|\lambda, M) & a = M \in \boldsymbol{M} \end{cases} . \end{aligned} \tag{2.25}$$

Here $\delta$ is the Kronecker delta, which we use to assign a trivial output for preparations and transformations. The state update map $\eta$ emerges naturally from this perspective of a quantum experiment as a stochastic process, and here we see another reason why it is only defined in the support of $\xi$. If we take the joint distribution $\Pr(k, \lambda'|\lambda, a)$ to be the more fundamental object, then it is clear we can obtain $\xi$ directly by marginalization

$$\xi(k|\lambda, M) = \int_{\mathbf{\Lambda}} \mathrm{d}\lambda' \Pr(k, \lambda'|\lambda, M) \tag{2.26}$$

which is always well defined, and then find $\eta$ by rearranging Eq. 2.25:

$$\eta(\lambda'|k, \lambda, M) = \frac{\Pr(k, \lambda'|\lambda, M)}{\xi(k|\lambda, M)} \tag{2.27}$$

Thus clearly $\eta(\lambda'|k, \lambda, M)$ is only well-defined when $\xi(k|\lambda, M) \neq 0$.

Definitions 4–6 constitute an equivalent formulation of the ontological models formalism. The assumptions of this construction can be broken down as follows: (a) quantum theory is described by a stochastic channel, (b) this channel is stationary, (c) it is causal, and (d) we assign the system a state which acts as an HMM of the channel. The authors of [56] identify (c) and (d), calling them non-retrocausality and $\lambda$-mediation, respectively. We will address these assumptions further in Chapter 6, and discuss what they might mean for the search for a $\psi$-epistemic interpretation of quantum theory outside of the ontological models formalism.

For now, note that we have said nothing in this analogy about the bogeyman that is "reality"—the state $\lambda$ defined here is not quite an operational object, but there is no stipulation that it must be "real." Of course, the state being real is often the *point* of this whole construction, but being able to discuss the ontological models formalism outside of a realist framework should lend some clarity as to when and where results derived in this formalism apply. In particular, it shifts the focus from the word "real" to the word "state." This is why we introduced the term "state-(anti)realist" in the introduction; those categories are sometimes referred to simply as "realist" and "antirealist," which are meaningless words when not used with a referent.

## 2.7 Aside: contextuality and sequential measurements

As already stated, this thesis is not overly concerned with contextuality. However, we take a moment to define the analogue of generalized contextuality in the HMM picture, since this immediately raises some interesting questions for future investigation. As we have unified preparations, measurements, and transformations in this picture, we can also unify the definition of operational equivalence for two actions $a, a'$:

$$a \underset{\text{op}}{\sim} a' \iff \Pr(\overleftrightarrow{k} | \overleftarrow{a}\, a\, \overset{\circ}{\overrightarrow{a}}) = \Pr(\overleftrightarrow{k} | \overleftarrow{a}\, a'\, \overset{\circ}{\overrightarrow{a}}) \qquad \forall \overleftrightarrow{k}, \overleftarrow{a}, \overset{\circ}{\overrightarrow{a}} \,. \tag{2.28}$$

We can also define a model to be *noncontextual* (overall) if

$$a \underset{\text{op}}{\sim} a' \implies \Pr(\overleftrightarrow{k}, \overleftrightarrow{\lambda} | \overleftarrow{a}\, a\, \overset{\circ}{\overrightarrow{a}}) = \Pr(\overleftrightarrow{k}, \overleftrightarrow{\lambda} | \overleftarrow{a}\, a'\, \overset{\circ}{\overrightarrow{a}}) \qquad \forall \overleftrightarrow{k}, \overleftrightarrow{\lambda}, \overleftarrow{a}, \overset{\circ}{\overrightarrow{a}} \,. \tag{2.29}$$

This definition clearly coincides with Spekkens' in the case of preparations and transformations, but differs slightly in the case of measurements as our definition requires operational equivalence to hold over all subsequent observations as well. For destructive measurements, these definitions coincide.

However, for nondestructive measurements, many fewer measurements qualify as operationally equivalent if we take into account state update. For example, we can take the measurement described in Section 2.4 where we coarse-grain the measurement $\{[0], [1], [2]\}$ either coherently or decoherently. According to Spekkens' definition, since these are represented by the same POVM, they are operationally equivalent for the purposes of discussing measurement contextuality. However, since they have different state update behavior, they are not operationally equivalent under the definition of Eq. (2.28). This is exactly the type of measurement used in the proof that any model of quantum theory must be measurement contextual [80]. It is thus an open question whether or not models of quantum theory must be contextual in the sense of Eq. (2.29) that takes into account state update behavior for nondestructive measurements.

This formulation also highlights a particularly pernicious form of contextuality that, as far as I know, has not been identified in the literature. This is the assumption that, when enacting a preparation, the new ontic state is independent of the previous ontic state. As far as I can tell, it is not necessarily the case that this is true. To formalize this, it makes sense to call an action $a$ an "operational preparation" if it removes all dependence on the past (this definition subsumes preparations in the operational theory sense along with rank-1 measurements):

$$\Pr(\overset{\circ\rightarrow}{k} \mid \overleftarrow{k}, \overleftarrow{a}, k, a, \overset{\circ\rightarrow}{a}) = \Pr(\overset{\circ\rightarrow}{k} \mid k, a, \overset{\circ\rightarrow}{a}). \tag{2.30}$$

However, it does not follow from this that $a$ must prepare an ontic state that is independent of the previous ontic state:

$$a \text{ is an operational preparation } \not\Longrightarrow \; Pr(\lambda'|\lambda, k, a) = Pr(\lambda'|, k, a). \tag{2.31}$$

That this implication holds is implicitly assumed by the definition of the preparation distribution $\mu$ for ontological models.

To show how this issue might arise in a particular example, consider preparing the state $|0\rangle$ on a qubit by measuring in the standard basis, and then applying a correctional unitary if the result $|1\rangle$ was obtained. This is, operationally, a preparation: it completely determines all future observations of the system. However, the system existed before the preparation (as is the case with physical qubits used in NMR, ion traps, and more; the main exception is photonic systems) and thus we might assume that it had some state before the preparation. While the *quantum* state is now independent of this previous state, it does not follow that the ontic state must be.

It is unclear exactly how much of an issue this is; it may be that it is perfectly legitimate to marginalize over the previous ontic state, although it is not clear that that will always be well-defined. Additionally, it is clearly related to the preparation independence assumption for Bell's theorem [11], and seems more general but may in fact not be.

# Chapter 3

# A dictionary of ontological models

*All God's critters got a place in the choir*
*Some sing low, some sing higher*
*Some sing out loud on the telephone wire*
*And some just clap their hands, paws, or anything they got now*

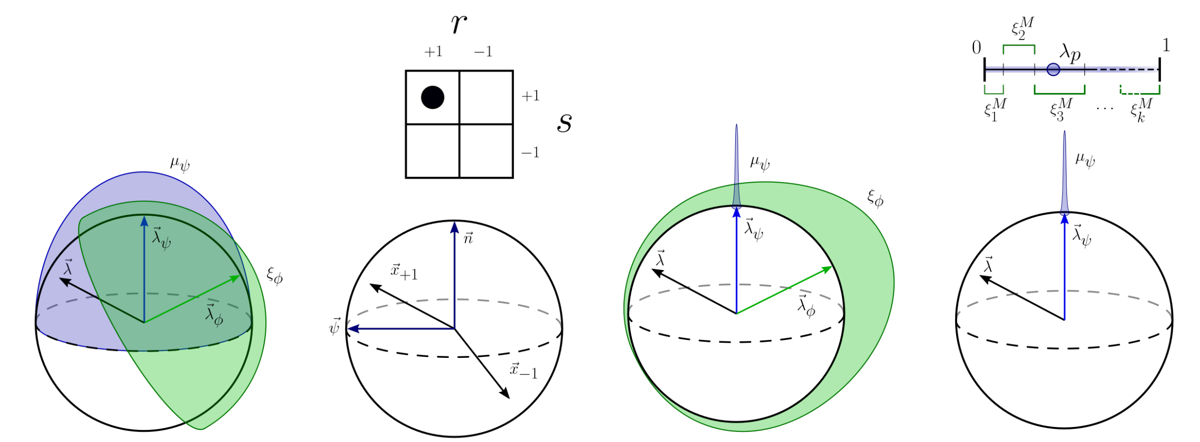
– Bill Staines & Janet Wheeler, *All God's Critters*

We provide a number of examples of ontological models from the literature, illustrating some of the properties described in the previous sections. For each model, we specify its state update rule if it exists; those that can't represent state update are noted but a proof is postponed until Chapter 4. All of these state-update rules have not appeared previously in the literature, with the exception of the one for Montina's model, though its presentation is new. Although many of these models can be easily defined for arbitrary types of measurements, we only consider projective measurements for simplicity and notational consistency.

## 3.1 $\psi$-epistemic models of a qubit

### 3.1.1 Kochen-Specker model

The Kochen-Specker (KS) model of a qubit [47, 53, 55] is an exemplar of what we look for in a $\psi$-epistemic theory, with the unfortunate feature that it only works in $d = 2$ dimensions. It is both pairwise $\psi$-epistemic and never $\psi$-ontic, and provides a very intuitively pleasing




(a) Kochen-Specker model (b) Montina model (c) BB model (d) Bell Model

Figure 3.1: Visualizations of the state space, preparation distributions, and response functions for (a) the Kochen-Specker model, (b) Montina's model, (c) the Beltrametti-Bugajski model for $d = 2$, and (d) Bell's model for $d = 2$. Blue represents the support of preparations, and green the support of the response functions, where possible. Black objects are generic elements of the state space. Figure created by Piers Lillystone, used with permission.

interpretation of the statistical nature of quantum theory. We take the ontic space to be the unit sphere $S^2$, and denote by $\vec{\psi}$ the Bloch vector corresponding to $|\psi\rangle$ under the usual mapping. Preparations and measurement outcomes are represented by distributions over hemispheres, with response functions uniform and preparation distributions peaked towards the center (Fig. 3.1a). Unitary transformations are represented by rotations of the sphere. Since the only nontrivial measurements on a qubit are rank-1 measurements, this is a case where we can use the state update rule described in Eq. 2.18 and just re-prepare the measured state for our update rule:

$$
\begin{aligned}
\mathbf{\Lambda} &= S^2 \\
\mu(\vec{\lambda}|P_\psi) &= \frac{1}{\pi}\Theta(\vec{\psi}\cdot\vec{\lambda})\vec{\psi}\cdot\vec{\lambda} \\
\Gamma(\vec{\lambda}'|\vec{\lambda},T_U) &= \delta(\vec{\lambda}'-R_U\vec{\lambda}) \\
\xi(k|\vec{\lambda},M_\phi) &= \Theta(k\vec{\phi}\cdot\vec{\lambda}) \\
\eta(\vec{\lambda}'|k,\vec{\lambda},M_\phi) &= \frac{1}{\pi}\Theta(k\vec{\phi}\cdot\vec{\lambda}')k\vec{\phi}\cdot\vec{\lambda}'
\end{aligned}
\tag{3.1}
$$

Here $\Theta$ is the Heaviside step function, $R_U$ is the rotation of the Bloch sphere corresponding to a unitary $U$, and $k \in \{+1,-1\}$. This particular state update rule is not particularly satisfactory in an explanatory sense, since it is independent of the previous ontic state.

An actual proof that this model reproduces quantum mechanics has not been presented in the literature, so for completeness we include it in Appendix A.

Rudolph notes in [72] that in the Kochen-Specker model, one should update the state to a cosine distribution peaked at the measurement vector, which is exactly what we have done with our generic update rule. This clearly reproduces quantum statistics given that the rest of the apparatus does, but does not give an account of the physical process (the measurement invasiveness) that would cause one's distribution to update in that fashion. In the same paper, Rudolph notes the apparent arbitrariness of the cosine probability distribution used in the KS model; he would prefer uniform distributions (or at least something related to the maximum entropy principle) such as those he utilizes in the rest of the paper. Montina notes in [62] that his model, further articulated in [64], satisfies this criterion by simply adding a second vector.

### 3.1.2 Montina's model

In [64], Montina introduces an ontological model based on the Kochen-Specker model that explicitly accounts for measurement invasiveness. The model was constructed to show that state update in a qubit can be successfully modeled by only updating a finite amount of information in the ontic state. The original presentation is not stated in terms of the ontological models formalism, and only explicitly models measurements in the standard basis. As such, we will take some space to formally express the model and extend it to model measurement in arbitrary bases.

We include this model here for two reasons. First, it is one of the few ontological models in the literature that has explicitly considered state update under measurement.

Second, it demonstrates that the generic rank-1 update (Eq. 2.18) that we used for the Kochen-Specker model is not the only possibility; even though all measurements in this model are rank-1, $\eta$ has nontrivial dependence on the previous ontic state $\lambda$. Thus just because we *can* construct a trivial update rule in some cases does not mean that there is then nothing interesting to investigate. It also includes these features while remaining pairwise $\psi$-epistemic and never $\psi$-ontic.

We first present Montina's model as he describes it, and then derive $\mu, \xi, \Gamma, \eta$ based on this description. The ontic space is

$$\mathbf{\Lambda} = S^2 \times S^2 \times \{-1, 1\} \times \{-1, 1\}. \tag{3.2}$$

We write an arbitrary ontic state as $\lambda = (\vec{x}_{+1}, \vec{x}_{-1}, r, s)$, where

$$\vec{x}_{+1}, \vec{x}_{-1} \in S^2 \tag{3.3}$$

$$r, s \in \{-1, 1\} \tag{3.4}$$

Montina refers to the vectors $\vec{x}_{+1}, \vec{x}_{-1}$ as noise functions, but we include them in the ontic state of the qubit since, if they weren't part of the ontic state, where would they exist? As desired by Rudolph [72], the marginal distribution over $\vec{x}_i$ is uniform at all times[1], and they evolve deterministically according to the rotation $R_U$ that corresponds to the unitary evolution of the qubit; i.e., if $\vec{x}_i$ is the vector before a transformation and $\vec{x}'_i$ is the one after, then

$$\vec{x}'_i = R_U \vec{x}_i. \tag{3.5}$$

$r$ is the bit that is updated by each measurement. It acts as an index on which vector ($x_{+1}$ or $x_{-1}$) to 'pay attention to.' Upon measurement in the standard basis, $r$ updates to

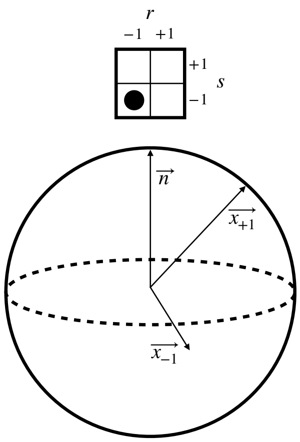
$$r' = \text{sgn}[(\vec{x}_{+1} \cdot \vec{n})^2 - (\vec{x}_{-1} \cdot \vec{n})] \tag{3.6}$$

where $\vec{n}$ is a unit vector along the $z$-axis. This means that we 'pay attention to' whichever vector was closest to the axis of measurement at the time of the most recent measurement.

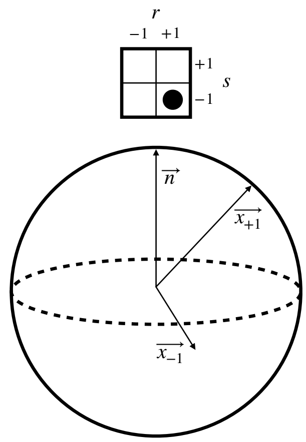
Finally, $s \in \{-1, 1\}$ determines the result of a measurement in the standard basis. $s$ flips sign whenever the vector $\vec{x}_r$ (that we are 'paying attention to') crosses the equator (the geodesic normal to $\vec{n}$) during a transformation. An example of an evolution of the ontic state is shown in Fig. 3.2.

Montina shows that the conditional probabilities for consecutive measurements obtained from this model align exactly with those of quantum mechanics.

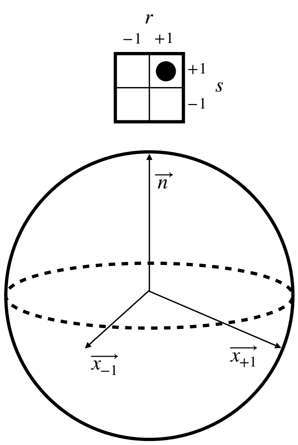
[1] We use $\vec{x}_i$ to refer generically to either/both of the vectors.



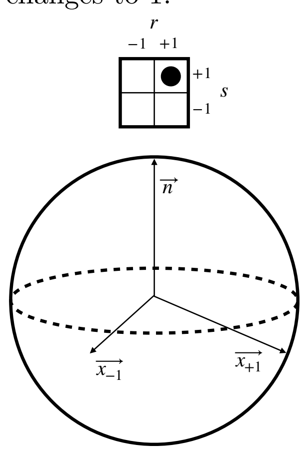

(a) Just before the first measurement, we have some initial values for the ontic state.




(b) Just after the first measurement, $r$ updates to the index of the vector that is more closely aligned with the axis along $n$. In this case, this is $\vec{x}_{+1}$ so $r$ changes to 1.




(c) Just after the unitary transformation and just before the second measurement, both of the vectors have been rotated with the same rotation matrix $R_U$ corresponding to the unitary. $r = 1$ tells us that we should be "paying attention to" $\vec{x}_{+1}$, which has changed hemispheres during the transformation, so $s$ changes sign from its previous value.




(d) The second measurement reports the value of $s$, in this case $+1$. Just after the second measurement, we again must update $r$. This time, $\vec{x}_{+1}$ is still more aligned with the $z$-axis so $r$ remains 1.

Figure 3.2: An example of the evolution for a particular ontic state of Montina's model, in an experiment consisting of two measurements in the standard basis separated by a unitary transformation.

Explicitly, Montina only models standard basis measurements with arbitrary unitary transformations; the implicit process for other measurements is "rotate, measure, rotate back." For preparations, we measure in the standard basis then rotate our result to the desired state using a transformation. We can make both of these implicit processes explicit, and in the process derive exact expressions for Montina's model as expressed in the ontological models formalism. The reader who is uninterested in the derivation may skip to Eq. (3.20).

We begin with $\Gamma$, the transition matrix for some transformation $T_U$ corresponding to a unitary $U$. This just amounts to wrapping (3.5) in a delta function and also writing down the update rule for $s$ using a Heaviside theta. $r$ remains unchanged.

$$\Gamma(\lambda'|\lambda, T_U) = \delta\left(\vec{x}'_{+1} - R_U\vec{x}_{+1}\right)\delta\left(\vec{x}'_{-1} - R_U\vec{x}_{-1}\right)\theta\left[ss'\left(\vec{x}'_r\cdot\vec{n}\right)\left(\vec{x}_r\cdot n\right)\right]\theta[rr'] \tag{3.7}$$

Next, we describe preparations of a pure quantum state $|\psi\rangle$, which we construct by the process:

1. Measure along $\vec{n}$ and get result $s$
2. Apply rotation $R_{s\vec{\psi}}$, where $R_{\vec{m}}$ is the rotation such that $\vec{m} = R_{\vec{m}}\vec{n}$.

We evaluate the probability distribution at each of these steps, denoting by $\lambda = (\vec{x}_{+1}, \vec{x}_{-1}, r, s)$ the state after the measurement (step 1) and by $\lambda' = (\vec{x}'_{+1}, \vec{x}'_{-1}, r', s')$ the state after the rotation (step 2). Our goal is to obtain the expression $\mu(\lambda'|\psi)$. Since $\vec{x}_{+1}, \vec{x}_{-1}, s$ are all independent, we can write

$$\Pr(\lambda) = \Pr(r|\vec{x}_{+1}, \vec{x}_{-1}, s)\Pr(x_{+1})\Pr(x_{-1})\Pr(s) \tag{3.8}$$

where all three of the latter distributions are uniform:

$$\Pr(x_{+1}) = \Pr(x_{-1}) = \frac{1}{4\pi}, \qquad \Pr(s) = \frac{1}{2}. \tag{3.9}$$

$r$ is determined by which $\vec{x}$ is closer to $\vec{n}$, so we have

$$\Pr(r|\vec{x}_{+1}, \vec{x}_{-1}, s) = \theta\left[r\mathrm{sgn}\left[\left(\vec{x}_{+1}\cdot\vec{n}\right)^2 - \left(\vec{x}_{+1}\cdot\vec{n}\right)^2\right]\right] \tag{3.10}$$

and can write the joint distribution as

$$\Pr(\lambda) = \frac{1}{32\pi^2}\theta\left[r\mathrm{sgn}\left[\left(\vec{x}_{+1}\cdot\vec{n}\right)^2 - \left(\vec{x}_{+1}\cdot\vec{n}\right)^2\right]\right]. \tag{3.11}$$

Now to find the distribution after the rotation, we multiply this expression by $\Gamma$ and integrate out $\lambda$:

$$\Pr(\lambda'|P_\psi) = \sum_{r,s} \int \mathrm{d}\vec{x}_{+1}\, \mathrm{d}\vec{x}_{-1}\ \Gamma(\lambda'|\lambda, T_{s\vec{\psi}})\Pr(\lambda). \tag{3.12}$$

We first note that

$$\delta\left[\vec{x}'_i - R_{s\vec{\psi}}\vec{x}_i\right] = \delta\left[R^{-1}_{s\vec{\psi}}\vec{x}'_i - \vec{x}_i\right] \tag{3.13}$$

and

$$R^{-1}_{s\vec{\psi}}\vec{x}'_r \cdot \vec{n} = \vec{x}'_r \cdot R_{s\vec{\psi}}\vec{n} = s\vec{x}'_r \cdot \vec{\psi}. \tag{3.14}$$

Using these two expressions and integrating out the $\delta$-functions right away gives

$$\Pr(\lambda'|\psi) = \sum_{r,s} \theta\left[s's\left(s\vec{x}'_r \cdot \vec{\psi}\right)\left(\vec{x}'_r \cdot \vec{n}\right)\right]\theta[r'r] \tag{3.15}$$

$$\cdot \frac{1}{32\pi^2}\theta\left[r\mathrm{sgn}\left[\left(\vec{x}'_{+1} \cdot \vec{\psi}\right)^2 - \left(\vec{x}'_{-1} \cdot \vec{\psi}\right)^2\right]\right] \tag{3.16}$$

$$= \frac{1}{(4\pi)^2}\theta\left[s'\left(\vec{x}'_{r'} \cdot \vec{\psi}\right)\left(\vec{x}'_{r'} \cdot \vec{n}\right)\right]\theta\left[r'\mathrm{sgn}\left[\left(\vec{x}'_{+1} \cdot \vec{\psi}\right)^2 - \left(\vec{x}'_{-1} \cdot \vec{\psi}\right)^2\right]\right] \tag{3.17}$$

so we have (removing primes)

$$\mu(\lambda|P_\psi) = \frac{1}{(4\pi)^2}\theta\left[s\left(\vec{x}_r \cdot \vec{\psi}\right)\left(\vec{x}_r \cdot \vec{n}\right)\right]\theta\left[r\,\mathrm{sgn}\left[\left(\vec{x}_{+1} \cdot \vec{\psi}\right)^2 - \left(\vec{x}_{-1} \cdot \vec{\psi}\right)^2\right]\right] \tag{3.18}$$

Using the same techniques, it is much easier to write down the response function

$$\xi(k|\lambda, M_\phi) = \theta\left[ks\left(\vec{x}_r \cdot \vec{n}\right)\left(\vec{x}_r \cdot \vec{\phi}\right)\right] \tag{3.19}$$

Summarizing these constructions, we can write

$$\begin{aligned}
\mathbf{\Lambda} &= S^2 \times S^2 \times \{-1,+1\} \times \{-1,+1\} \\
\lambda &= (\vec{x}_{+1}, \vec{x}_{-1}, r, s) \\
\mu(\lambda|P_\psi) &= \frac{1}{(4\pi)^2}\Theta\left[s\left(\vec{x}_r \cdot \vec{\psi}\right)\left(\vec{x}_r \cdot \vec{n}\right)\right]\Theta\left[r\left[\left(\vec{x}_{+1} \cdot \vec{\psi}\right)^2 - \left(\vec{x}_{-1} \cdot \vec{\psi}\right)^2\right]\right] \\
\Gamma(\lambda'|\lambda, T_U) &= \delta\left(\vec{x}'_{+1} - R_U\vec{x}_{+1}\right)\delta\left(\vec{x}'_{-1} - R_U\vec{x}_{-1}\right)\Theta\left[ss'\left(\vec{x}'_r \cdot \vec{n}\right)\left(\vec{x}_r \cdot \vec{n}\right)\right]\Theta[rr'] \\
\xi(k|\lambda, M_\phi) &= \Theta\left[ks\left(\vec{x}_r \cdot \vec{n}\right)\left(\vec{x}_r \cdot \vec{\phi}\right)\right] \\
\eta(\lambda'|k, \lambda, M_\phi) &= \delta\left(\vec{x}'_{+1} - \vec{x}_{+1}\right)\delta\left(\vec{x}'_{-1} - \vec{x}_{-1}\right)\Theta\left[ss'(\vec{x}_r \cdot \vec{n})\left(\vec{x}_r \cdot \vec{\phi}\right)(\vec{x}_{r'} \cdot \vec{n})\left(\vec{x}_{r'} \cdot \vec{\phi}\right)\right] \\
&\quad \cdot \Theta\left[r'\left[\left(\vec{x}_{+1} \cdot \vec{\phi}\right)^2 - \left(\vec{x}_{-1} \cdot \vec{\phi}\right)^2\right]\right]
\end{aligned} \tag{3.20}$$

We prove that this construction reproduces quantum mechanics in Appendix A.

Note that, for measurement in an arbitrary basis, both $r$ and $s$ are updated (although the $\vec{x}_i$ remain static). This is in contrast to the case for standard basis measurements, where only $r$ is updated. The standard basis case led to Montina's claim that state update under measurement is accounted for by updating a single bit, but it is clear from the form of $\eta$ above that by including all measurements in our subtheory we have caused both bits to be updated during measurement.

## 3.2 Models of full quantum theory for arbitrary dimension

### 3.2.1 Beltrametti-Bugajski model

The Beltrametti-Bugajski (BB) model [12, 47, 55] is perhaps the simplest ontological model that describes a system of arbitrary dimension. Although it is $\psi$-ontic, it is the starting point for the construction of the next three models in this section. For a $d$-dimensional quantum system, we take the ontic space to be the quantum state space, which we denote $\mathcal{PH}^{d-1}$ (the projective Hilbert space of dimension $d-1$). Preparations, transformations, measurements, and state update rules then follow directly from the usual quantum rules:

$$\begin{aligned}
\mathbf{\Lambda} &= \mathcal{PH}^{d-1} \\
\mu(\lambda|P_\psi) &= \delta(|\lambda\rangle - |\psi\rangle) \\
\Gamma(\lambda'|\lambda, T_U) &= \delta(|\lambda'\rangle - U|\lambda\rangle) \\
\xi(k|\lambda, M_{\{\Pi_i\}}) &= \langle\lambda|\,\Pi_k\,|\lambda\rangle \\
\eta(\lambda'|k, \lambda, M_{\{\Pi_i\}}) &= \delta\left(|\lambda'\rangle - \frac{\Pi_k\,|\lambda\rangle}{\sqrt{\langle\lambda|\,\Pi_k\,|\lambda\rangle}}\right)
\end{aligned} \tag{3.21}$$

This provides an example of the generic update-rule for $\psi$-ontic models (Eq. 2.19), and is depicted in Fig. 3.1c. The BB model is also termed $\psi$-complete, since it defines no additional ontology besides the quantum state.

Note that, although the BB model takes the pure quantum states as its ontic state space, it can still model mixed-state quantum theory, since any density matrix can be written (nonuniquely) as a convex combination of pure states. This is why we discuss $\psi$-onticity rather than $\rho$-onticity, since the BB model is "$\rho$-epistemic" and works for all dimensions, including state update.

### 3.2.2 $\rho$-complete model

That said, we can still define a model that takes the set of density matrices as its ontology rather than the set of pure states. This model has not (to my knowledge) been proposed in the literature, mostly because it is seen as an unnecessarily cumbersome version of the BB model.

$$\Lambda = \{\text{density matrices}\} \tag{3.22}$$

$$\mu(\lambda|P_\rho) = \delta(\lambda - \rho) \tag{3.23}$$

$$\Gamma(\lambda'|\lambda, T_{\mathcal{C}}) = \delta(\lambda' - \mathcal{C}(\lambda)) \tag{3.24}$$

$$\xi(k|\lambda, M_{\{M_i\}}) = \operatorname{tr}\left(M_k \lambda M_k^\dagger\right) \tag{3.25}$$

$$\eta(\lambda'|k, \lambda, M_{\{M_i\}}) = \delta(\lambda' - M_k \lambda M_k^\dagger) \tag{3.26}$$

Here $\mathcal{C}$ is a CPTP map (potentially noisy quantum channel), and $\{M_i\}$ is a set of generalized (not necessarily projective) measurement operators. If we close this model under convex combination, it's horrendously contextual, but interestingly we don't have to include convexity in order to reproduce all of quantum theory (while convexity is required for the BB model to reproduce mixed-state quantum theory). This model does not have many interesting properties from the usual perspective of quantum foundations work, but we will discuss it further when we arrive in Chapter 5.

### 3.2.3 Bell's model

Lewis et al. [57] extended a model of a qubit originally proposed by Bell [11] to arbitrary dimension, which can be seen as a modification of the Beltrametti-Bugajski model [55]. It is in some ways analogous to the de Broglie-Bohm interpretation of continuum quantum theory, in that it is $\psi$-ontic while adding just a little bit of extra ontology in order to render it *outcome deterministic*, which means that for any given ontic state, every measurement has a predetermined outcome: $\xi(k|\lambda, M) \in \{0,1\} \quad \forall k, \lambda, M$. Bell, a big fan of the de Broglie-Bohm interpretation, created this model with the intent of demonstrating that an outcome deterministic model of finite-dimensional quantum theory is possible.

The ontic space is the Cartesian product of the projective Hilbert space with the unit interval $[0,1]$. Now we write $\lambda$ as an ordered pair $\lambda = (|\lambda\rangle, p_\lambda)$ where $|\lambda\rangle \in \mathcal{PH}^{d-1}$, as in the Beltrametti-Bugajski model, and $p_\lambda \in [0,1]$. Preparations remain essentially the same, becoming a product distribution of a delta function on the quantum state space with a uniform distribution over the unit interval. The response functions divide up the unit interval into lengths corresponding to probabilities of measuring each outcome, and respond with outcome $k$ when $p_\lambda$ is in the corresponding interval (Fig. 3.1d). This has the effect of making the model outcome deterministic.

$$\begin{aligned}
\mathbf{\Lambda} &= \mathcal{PH}^{d-1} \times [0,1] \\
\mu(\lambda|P_\psi) &= \delta(|\lambda\rangle - |\psi\rangle) \\
\Gamma(\lambda'|\lambda, T_U) &= \delta(|\lambda'\rangle - U\,|\lambda\rangle) \\
\xi(k|\lambda, M_{\{\Pi_i\}}) &= \Theta\left[p_\lambda - \sum_{j=0}^{k-1} \mathrm{tr}(\Pi_j[\lambda])\right]\Theta\left[-p_\lambda + \sum_{j=0}^{k} \mathrm{tr}(\Pi_j[\lambda])\right] \\
\eta(\lambda'|k, \lambda, M_{\{\Pi_i\}}) &= \delta\left(|\lambda'\rangle - \frac{\Pi_k\,|\lambda\rangle}{\sqrt{\langle\lambda|\,\Pi_k\,|\lambda\rangle}}\right)
\end{aligned} \tag{3.27}$$

Since this model is still $\psi$-ontic, we once again use the generic state update rule for $\psi$-ontic models. In this case, we can also see that this works because of the structure of the preparations as product distributions. Since every state has a uniform distribution over $p_\lambda$, and this is uncorrelated with $|\lambda\rangle$, we don't need to update any information about $p_\lambda$ and so can just re-use the Beltrametti-Bugajski model update rule, attached to a uniform distribution over $p_\lambda$. This uniform distribution over $p_\lambda$ is not particularly visible in the definitions of $\mu$ and $\eta$ because it is equal to 1. That said, it is present and amounts to a randomization step in the state update, which is necessary for an outcome-deterministic model.

### 3.2.4 LJBR model

We now proceed to our first example of a $\psi$-epistemic model in $d \geq 3$. In [57], Lewis et al. construct a $\psi$-epistemic model based on their generalization of Bell's model that we described in the previous section. Referred to here as the LJBR model, it is motivated by the observation that the order of segments in the response function of Bell's model does not matter: roughly, a re-ordering of these segments allows arbitrary modification of preparation distributions within a subset of the ontic space, so they can be made to overlap. We present here a description of the version of this model that has a preferred basis rather than a single preferred state; Lewis et al. describe it briefly but don't specify it explicitly, spending more time on motivation and construction of the simpler version with a single preferred state [57].

The LJBR model has the same ontic space as the Bell model, so we again write ontic states as $\lambda = (|\lambda\rangle\,, p_\lambda)$. It is constructed in a preferred basis $\{|j\rangle\}$, which we use in defining two helper functions. First,

$$z_j(|\lambda\rangle) = \inf_{|\phi\rangle:\mathrm{tr}([j][\phi])\geq 1/d} \mathrm{tr}([\lambda][\phi]). \tag{3.28}$$

Note that $z_j(|\lambda\rangle) > 0$ if and only if $\mathrm{tr}([j][\lambda]) > \frac{d-1}{d}$, so $z_j(|\lambda\rangle)$ is nonzero for at most a single element of the preferred basis; we denote this unique vector (if it exists) as $|j_\lambda\rangle$. Second, we implicitly define a permutation $\pi_{M,\lambda}$ for each measurement $M$ and ontic state $\lambda$:

$$\mathrm{tr}\big(M_{\pi_{M,\lambda}(0)}[j_\lambda]\big) \geq \mathrm{tr}\big(M_{\pi_{M,\lambda}(1)}[j_\lambda]\big) \geq \cdots \geq \mathrm{tr}\big(M_{\pi_{M,\lambda}(|M|-1)}[j_\lambda]\big). \tag{3.29}$$

If there is no $|j_\lambda\rangle$, i.e. $z_j(|\lambda\rangle) = 0$ for all $j$, then we take $\pi_{M,\lambda}$ to be the identity permutation. The final piece we need before defining the model itself is a set

$$\mathcal{S}_j = \{\lambda \mid z_j(|\lambda\rangle) > 0\} \tag{3.30}$$

defined for each basis vector. Without further ado, the full specification of the model:

$$\begin{aligned}
\mathbf{\Lambda} &= \mathcal{PH}^{d-1} \times [0,1] \\
\lambda &= (|\lambda\rangle, p_\lambda) \\
\mu(\lambda|P_\psi) &= \delta(|\lambda\rangle - |\psi\rangle) \prod_j \Theta[p_\lambda - z_j(|\psi\rangle)] + \sum_j z_j(|\psi\rangle)\mu_{\mathcal{S}_j}(\lambda) \\
\xi(k|\lambda, M_{\{\Pi_i\}}) &= \Theta\left[p_\lambda - \sum_{l=0}^{k-1} \mathrm{tr}\big(\Pi_{\pi_{M,\lambda}(l)}[\lambda]\big)\right] \Theta\left[-p_\lambda + \sum_{l=0}^{k} \mathrm{tr}\big(\Pi_{\pi_{M,\lambda}(l)}[\lambda]\big)\right]
\end{aligned} \tag{3.31}$$

where $\mu_{\mathcal{S}_j}(\lambda)$ is the uniform distribution over $\mathcal{S}_j$ (although it could be any distribution whose support is exactly equal to $\mathcal{S}_j$).

Roughly, all quantum states $|\psi\rangle$ with $\mathrm{tr}([j][\psi]) > \frac{d-1}{d}$ will have support on $\mathcal{S}_j$, and so will all be ontologically indistinct from one another. The permutation included in the definition of the measurements is constructed so that this shared support does not affect the prepare-and-measure-once statistics: the measurement ordered first by the permutation has a support which entirely contains $\mathcal{S}_j$. This model is not pairwise $\psi$-epistemic, nor is it never $\psi$-ontic. Note additionally that it was originally only defined for rank-1 projective measurements, but it works just as well for higher-rank projective measurements (or, in fact, any kind of measurement) without modification.

As advertised, this model cannot model sequential measurement/state update. The proof will be presented in Section 4.2.1. For the moment, note that the new preparation distribution is no longer a product distribution over the two parts of the state space, so the state update rule from the Bell model shouldn't work. Indeed, it is straightforward to combine that state update rule with $\mu$ and $\xi$ from (3.31) and check that this is the case. We also haven't specified the transition matrix $\Gamma$, as it turns out that this model can't represent transformations either (Section 4.3.4).

### 3.2.5 ABCL models

In [1], Aaronson et al. construct two $\psi$-epistemic models also based on Bell's model. The first, which we will call $\text{ABCL}_0$ , is very closely related to the LJBR model but is not identical; rather than continuous regions of quantum states which overlap, this model has exactly one pair of quantum states which are ontologically indistinct. However, it gains the feature that *any* two nonorthogonal quantum states can be chosen as the single pair that overlaps. As such, $\text{ABCL}_0$ is really a class of models parametrized by the choice of two states $|\alpha\rangle, |\beta\rangle$ which will be the two states that are ontologically indistinct.

The second, $\text{ABCL}_1$ , is a convex mixture (to be defined) of the $\text{ABCL}_0$ model constructed for all pairs $|\alpha\rangle, |\beta\rangle$ and is intended to demonstrate the possibility of a pairwise $\psi$-epistemic model. This is the only known example of a pairwise $\psi$-epistemic model in $d \geq 3$, but it still is not never $\psi$-ontic [55, 63]. These models have come under criticism for their "unnaturalness," but we show here that their problems go deeper due to an inability to represent state-update.

We begin with $\text{ABCL}_0$ , defining a couple of helper functions like in the LJBR model. Rather than ordering measurements with respect to traces with a preferred basis, we use the defining states $|\alpha\rangle, |\beta\rangle$ and a function

$$g_{\alpha\beta}(\Pi) = \min\{\text{tr}(\Pi[\alpha]), \text{tr}(\Pi[\beta])\}. \tag{3.32}$$

We now define a new permutation $\sigma_M$[2] for each measurement $M$ [55]:

$$g_{\alpha\beta}(M_{\sigma_M(0)}) \geq g_{\alpha\beta}(M_{\sigma_M(1)}) \geq \cdots \geq g_{\alpha\beta}(M_{\sigma_M(|M|-1)}). \tag{3.33}$$

[2] To be precise, we should label this with $\alpha, \beta$ as well to emphasize that it belongs to the model defined by that particular pair of states.

With this, we can specify the ABCL$_0$ model.

$$
\begin{aligned}
&\mathbf{\Lambda} = \mathcal{PH}^{d-1} \times [0,1] \\
&\lambda = (|\lambda\rangle, p_\lambda) \\
&\mu(\lambda|P_\psi) = \begin{cases} \Big[\Theta(p_\lambda - \varepsilon)\delta(|\lambda\rangle - |\psi\rangle) \\ \quad + \frac{1}{2}\Theta(\varepsilon - p_\lambda)[\delta(|\lambda\rangle - |\alpha\rangle) + \delta(|\lambda\rangle - |\beta\rangle)]\Big] & \text{if } |\psi\rangle = |\alpha\rangle, |\beta\rangle \\ \delta(|\lambda\rangle - |\psi\rangle) & \text{otherwise} \end{cases} \\
&\xi(k|\lambda, M_{\{\Pi_i\}}) = \Theta\left[p_\lambda - \sum_{j=0}^{k-1} \operatorname{tr}\big(\Pi_{\sigma_M(j)}[\lambda]\big)\right]\Theta\left[-p_\lambda + \sum_{j=0}^{k} \operatorname{tr}\big(\Pi_{\sigma_M(j)}[\lambda]\big)\right]
\end{aligned} \tag{3.34}
$$

for $\varepsilon \leq \frac{|\langle\alpha|\beta\rangle|}{d}$. Now the preparation distributions for $|\alpha\rangle$ and $|\beta\rangle$ overlap on $\{|\alpha\rangle, |\beta\rangle\} \times [0, \varepsilon]$; as in the LJBR model, the permutation in $\xi$ ensures that the preparation change doesn't affect the prepare-and-measure-once statistics by making sure measurements whose support must contain this overlap region are ordered first.

We outline the ABCL$_1$ model schematically and refer the reader to [1, 55] for details. Given two ontological models specified by $\mathbf{\Lambda}_1, \mu_1, \xi_1$ and $\mathbf{\Lambda}_2, \mu_2, \xi_2$ respectively, the authors define a convex combination of these models as $\mathbf{\Lambda}_3, \mu_3, \xi_3$ such that

$$
\begin{aligned}
\mathbf{\Lambda}_3 &= \mathbf{\Lambda}_1 \oplus \mathbf{\Lambda}_2 \\
\mu_3 &= p\mu_1 + (1-p)\mu_2 \\
\xi_3 &= \xi_1 + \xi_2
\end{aligned} \tag{3.35}
$$

Here $p \in (0, 1)$ is some mixing parameter. If there's overlap between two states in either of models 1 or 2, then model 3 has overlap on these states. The ABCL$_1$ model is then defined essentially as a convex mixture of the ABCL$_0$ models for *all* pairs $|\alpha\rangle, |\beta\rangle$, taking care with respect to the uncountable size of this set.

In order to include state update in a convex combination of ontological models, the most obvious (and perhaps only) option is to specify

$$\eta_3 = \eta_1 + \eta_2. \tag{3.36}$$

Once again, these models cannot represent state update consistently, which we prove in Section 4.2.2. ABCL$_0$ also can't represent transformations, but ABCL$_1$ representing transformations is not strictly ruled out by any of our results. That said, it seems unlikely that it can, as any transformation would have to map *between* the submodels that have been combined (Section 4.3.4).

## 3.3 Models of subtheories

Although we have dealt so far with models that include the full set of preparations, transformations, and measurements for pure-state quantum theory, there is the possibility that state update will be better accounted for in $\psi$-epistemic models of subtheories. It turns out that although the stabilizer subtheory can be represented by a $\psi$-epistemic model, the more general Kitchen Sink model which models any finite subtheory cannot in general represent state update under measurement.

### 3.3.1 Kitchen Sink model

The Kitchen Sink model is a $\psi$-epistemic ontological model for any finite subtheory of quantum theory [46, Section IIIC]. Given a finite set of projective measurements $\mathcal{M} = \{M^{(i)}\}$, we choose our ontic states to be a list of measurement outcomes $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_{|\mathcal{M}|})$. That is, $\lambda_i = k$ means that if $M^{(i)} = \{\Pi_j^{(i)}\}$ is measured when the system is in the state $\lambda$, the outcome $\Pi_k$ will occur with certainty. For a system of dimension $d$, the maximum number of projectors in any given measurement is $d$, so we pad all of our measurements with 0s until they have $d$ elements. The Kitchen Sink model is then defined by

$$
\begin{aligned}
\boldsymbol{\Lambda} &= \mathbb{Z}_d^{|\mathcal{M}|} \\
\mu(\lambda|\psi) &= \prod_{i=1}^{|\mathcal{M}|} \operatorname{tr}\Big(\Pi_{\lambda_i}^{(i)}[\psi]\Big) \\
\xi(k|\lambda, M^{(i)}) &= \delta(k, \lambda_i)
\end{aligned} \tag{3.37}
$$

The Kitchen Sink is pairwise $\psi$-epistemic for almost all subtheories, and also never $\psi$-ontic if we include all pure states in our subtheory. Transformations can additionally be modeled under the assumption of a closed subtheory, which means that any transformation included in the subtheory must map states in the subtheory to other states in the subtheory, and similarly for measurements. In other words, for all transformations $T$ and measurements $M^{(i)}$, there exists a $j$ such that $\{M_k^{(j)}\} = \{T(M_k^{(i)})\}$. We denote this as $j = T(i)$, which allows for a nice expression of transformations:

$$\Gamma(\lambda'|\lambda, T) = \prod_{j=1}^{m} \delta(\lambda'_{T(j)}, \lambda_j) \tag{3.38}$$

We can only rule out the Kitchen Sink model for certain subtheories, as it is easy to construct subtheories with trivial update rules (i.e. by only including rank-1 measurements). That said, our requirements are few and are satisfied by the multi-qu$p$it stabilizer subtheory, arguably the most important subtheory of quantum theory. Specifically, we only need to include two states $|\alpha\rangle, |\beta\rangle$ and two measurements $M^{(1)} = \{\Pi, \mathbb{I} - \Pi\}, M^{(2)}$ satisfying

$$\langle\alpha|\beta\rangle \neq 0 \tag{3.39}$$

$$\langle\alpha|\Pi|\alpha\rangle \neq 0 \tag{3.40}$$

$$\langle\beta|\Pi|\beta\rangle \neq 0 \tag{3.41}$$

$$M^{(2)} \text{ distinguishes } \Pi|\alpha\rangle \text{ and } \Pi|\beta\rangle \tag{3.42}$$

The first is required because we don't expect orthogonal states to be ontologically indistinct. The next two stipulate that there is a nonzero chance of obtaining an outcome $\Pi$ when measuring $|\alpha\rangle$ and $|\beta\rangle$, so that its support overlaps with their support. The last condition implies that the post-measurement states are ontologically distinct [55]; note that the existence of such a measurement in any case requires that $\Pi|\alpha\rangle$ and $\Pi|\beta\rangle$ are orthogonal.

The proof that such a subtheory cannot represent measurement update (in Section 4.2.3) demonstrates that even subtheories have trouble with state update. In particular, the stabilizer subtheory satisfies the requirements in Eqs. 3.39–3.42, so it cannot be modeled by the Kitchen Sink. That said, we can show that the stabilizer subtheory still supports a $\psi$-epistemic interpretation using other models.

### 3.3.2 Qu$p$it stabilizer subtheory

We focus on the stabilizer subtheory for prime dimensional systems. For the case $p = 2$, Lillystone and Emerson construct a $\psi$-epistemic model of the $n$-qubit stabilizer formalism that successfully represents state update under measurement [58]. This model starts from the Kitchen Sink model and augments $\mathbf{\Lambda}$ so that the problematic overlaps of the Kitchen sink are removed. This model is not pairwise $\psi$-epistemic, but a modified version (see appendix of [58]) is never-$\psi$-ontic. We don't present the construction here because it is significantly more convoluted than the odd prime case. This reflects the often-observed ill-behaved nature of the qubit stabilizer subtheory.

For $p \geq 3$, the stabilizer subtheory admits a particularly nice $\psi$-epistemic ontological model constructed from the discrete Wigner function [40, 88, 91], and so we restrict to odd prime $p$ for the rest of this section. We refer to qu*d*its of o̲dd p̲rime dimension as quo̲*p̲*its[3], with dimension $p$.

Before we specify our ontological model, we must describe the operational theory; that is, the stabilizer subtheory of quantum mechanics. The construction begins with the generalization of the operators $X$ and $Z$ to a single quo*p*it, which are defined by their action on the standard basis $\{|j\rangle\}$ for $j = 0, 1, \ldots, p-1$:

$$X|j\rangle = |j+1\rangle \tag{3.43}$$

$$Z|j\rangle = \omega^j |j\rangle \tag{3.44}$$

$$\omega = e^{2\pi i/p} \tag{3.45}$$

All integer arithmetic is done $\bmod p$. The full set of generalized Pauli operators on $n$ quo*p*its is then given by

$$T_{(\mathbf{x},\mathbf{z})} = \bigotimes_{j=0}^{n-1} \omega^{x_j z_j/2} X^{x_j} Z^{z_j} \tag{3.46}$$

$$\text{for } \mathbf{x} = (x_0, x_{+1}, \ldots, x_{n-1}) \in \mathbb{Z}_p^n$$

$$\text{and } \mathbf{z} = (z_0, z_1, \ldots, z_{n-1}) \in \mathbb{Z}_p^n$$

Note that, while the Pauli operators for qubits are both Hermetian and unitary, the generalized Pauli operators are only unitary.

We also define the symplectic inner product as

$$[(\mathbf{x},\mathbf{z}),(\mathbf{x}',\mathbf{z}')] = \mathbf{z}\cdot\mathbf{x}' - \mathbf{x}\cdot\mathbf{z}', \tag{3.47}$$

where $\cdot$ is the usual dot product on $\mathbb{Z}_p^n$. We will sometimes abbreviate $\lambda \in \mathbb{Z}_p^n \times \mathbb{Z}_p^n$ rather than writing $(\mathbf{x},\mathbf{z})$, as this will later be an element of our ontic space. We can now define the *phase-point operators* $A_\lambda$, which are a symplectic Fourier transform of the Pauli operators:

$$A_\lambda = \frac{1}{p^n} \sum_{\lambda' \in \mathbb{Z}_p^n \times \mathbb{Z}_p^n} \omega^{[\lambda,\lambda']} T_{\lambda'} \tag{3.48}$$

[3] pronounced [ˈkwɑ pɪts]

Stabilizer measurements are the easiest to define. The are just measurements of the Pauli operators $T_{(\mathbf{x},\mathbf{z})}$; in other words, projections onto their eigenspaces. Each Pauli operator has eigenvalues $\{\omega^k\}$ for $k \in \mathbb{Z}_p$. Each of the $p$ eigenvalues have equally sized eigenspaces, and thus each eigenspace has dimension $p^{n-1}$. We can reformulate this in the measurement operator picture rather than an observable picture, in order to match the rest of this thesis and allow for some explicit calculations down the line. Let $\Pi_k^{(\mathbf{x},\mathbf{z})}$ be the projector onto the $k$th eigenspace of $T_{(\mathbf{x},\mathbf{z})}$, i.e. the eigenspace with eigenvalue $\omega^k$. The reader can check that

$$\Pi_k^{(\mathbf{x},\mathbf{z})} = \frac{1}{p}\sum_{m=0}^{p-1}\left(\omega^{-k}T_{(\mathbf{x},\mathbf{z})}\right)^m. \tag{3.49}$$

Then, the allowable stabilizer measurements are

$$M^{(\mathbf{x},\mathbf{z})} = \{\Pi_k^{(\mathbf{x},\mathbf{z})}\} \qquad \text{for } \mathbf{x},\mathbf{z} \in \mathbb{Z}_p^n. \tag{3.50}$$

We emphasize once more that each of these projectors is rank-$p^{n-1}$; as observed in Section 2.5, any projector with rank greater than one raises a potential difficulty in representing state update.

Next, we define the stabilizer states. A *stabilizer group* $G$ is defined as a set of $p^n$ mutually commuting Pauli operators multiplied by a phase $\omega^{-k}$ (where $k \in \mathbb{Z}_p$) [88]. This group can be specified by a set of $n$ generators, but care must be taken to make sure that they actually generate the whole group $G$ rather than a subgroup. There is a unique state which is an eigenvector of all of these operators with eigenvalue +1; we say that $G$ *stabilizes* this state. A stabilizer state, then, is a state which is stabilized by a stabilizer group $G$. To make this more concrete, we describe $G$ by a set $S$ of $n$ triples $(\mathbf{x},\mathbf{z},k)$; these triples specify the generators of $G$. In symbols, we define $S$ implicitly by

$$G = \langle\{\omega^{-k}T_{(\mathbf{x},\mathbf{z})} | (\mathbf{x},\mathbf{z},k) \in S\}\rangle, \tag{3.51}$$

where $\langle\cdot\rangle$ indicates the group generated by the set $\cdot$. We must be careful, however, to make sure that these generators commute and generate the full group. For the first condition, two Pauli operators $T_\lambda, T_{\lambda'}$ commute if and only if the symplectic inner product $[\lambda,\lambda'] = 0$ [87]. For the second condition, $S$ will generate the full group if and only if all of its elements $(\mathbf{x},\mathbf{z})$ are linearly independent (ignoring $k$) [17]. The explicit expression for the density operator of a stabilizer state corresponding to $S$ is then

$$\rho_S = \prod_{(\mathbf{x},\mathbf{z},k)\in S}\Pi_k^{(\mathbf{x},\mathbf{z})}. \tag{3.52}$$

For an alternative expression as a sum of Paulis that is sometimes more computationally convenient, see Appendix A.4.

Finally, the transformations of the stabilizer subtheory are called Clifford transformations. These are the transformations that map the set of Pauli operators to itself. In other words, it is the normalizer of the Pauli group. We do not discuss transformations further, as that would unnecessarily complicate our exposition.

With our operational theory defined, we can now describe the ontological model given by the discrete Wigner function. In general, the Wigner function can actually describe the full quantum theory, but in that case some of the probabilities become negative and it is a quasiprobability representation (QPR) [30–32] rather than an ontological model. This QPR is defined as

$$\begin{aligned}
\mathbf{\Lambda} &= \mathbb{Z}_p^n \times \mathbb{Z}_p^n \\
\mu(\lambda|P_\psi) &= \frac{1}{p^n}\,\mathrm{tr}(A_\lambda[\psi]) \\
\xi(k|\lambda, M_{\{\Pi_i\}}) &= \mathrm{tr}(\Pi_k A_\lambda). \\
\eta(\lambda'|k,\lambda, M_{\{\Pi_i\}}) &= \frac{1}{p^n}\frac{\mathrm{tr}(A_\lambda \Pi_k A_{\lambda'}\Pi_k)}{\mathrm{tr}(\Pi_k A_\lambda)}.
\end{aligned} \tag{3.53}$$

State-update rules are not usually discussed explicitly in QPRs, but state update is a completely positive (CP) map, and CP maps are described by an expression like $\eta$ for QPRs [30]. To double-check, we have also derived it explicitly in Appendix A.

If, however, we restrict to modeling preparations, transformations, and measurements in the qu$p$it stabilizer subtheory, then the representation is positive and it forms a well-defined ontological model [40, 41]. Using our notation for stabilizer states and measurements from above, we have

$$\begin{aligned}
\mu(\lambda|P_S) &= \frac{1}{p^n}\prod_{(\mathbf{x},\mathbf{z},k)\in S} \delta_{k,[(\mathbf{x},\mathbf{z}),\lambda]} \\
\xi(k|\lambda, M^{(\mathbf{x},\mathbf{z})}) &= \delta_{k,[(\mathbf{x},\mathbf{z}),\lambda]} \\
\eta(\lambda'|\,k,\lambda, M^{(\mathbf{x},\mathbf{z})}) &= \frac{1}{p}\frac{\delta_{k,[(\mathbf{x},\mathbf{z}),\lambda']}}{\delta_{k,[(\mathbf{x},\mathbf{z}),\lambda]}}\sum_{c\in\mathbb{Z}_p}\delta_{\lambda-\lambda',c(\mathbf{x},\mathbf{z})}
\end{aligned} \tag{3.54}$$

where $\delta$ is the Kronecker delta. We present a derivation of these expressions from the QPR version (Eq. 3.53) in Appendix A. Note that the delta function in the denominator of $\eta$ that may be zero; this is fine, since that delta function is equivalent to $\xi$ and $\eta$ only needs to be defined in the support of $\xi$. This model is both pairwise $\psi$-epistemic and never $\psi$-ontic.

Although it is known that the stabilizer states (measurements) are exactly the pure states (projective measurements) for which $\mu$ ($\xi$) is positive [40, 41], the positivity of $\eta$ has not been investigated. Our expression above amounts to a proof in one direction, that it is positive for any stabilizer measurement. It is an open question whether or not these are the *only* projective measurements for which $\eta$ is positive, though it does seem likely.

Tangentially, if we extend the Wigner function to the full quantum theory as a QPR, we get negatively represented state update, as expected. One consequence of this is that some state updates can't be normalized, so the Wigner function state update must include a renormalization step not allowed in ontological models or quasiprobability representations. For example, in a single qutrit where $\lambda$ takes the form $(x, z) \in \mathbb{Z}_3 \times \mathbb{Z}_3$, the projector $\Pi = [0] + [1]$ has

$$\operatorname{tr}\big(A_{(0,2)}\Pi A_{\lambda'}\Pi\big) \begin{cases} 2/3 & \lambda' = (0,2) \\ -1/3 & \lambda' = (1,2) \\ -1/3 & \lambda' = (2,2) \\ 0 & \text{otherwise} \end{cases}. \tag{3.55}$$

Since these add to 0, $\eta(\lambda'|\lambda, \Pi)$ is not normalizable for $\lambda = (0, 2)$. Alternatively, the joint distribution $\Pr(k, \lambda'|\lambda, M) = \xi(k|\lambda, M)\eta(\lambda'|k, \lambda, M)$ is always well-defined and normalizable, so if we restrict inquiry to that object then renormalization is not required. Although further discussion of state update under measurement in quasi-probability representations is beyond the scope of this paper, we note that Theorem 1 does not hold for quasi-probability representations so this could be one potential direction for related future work.

# Chapter 4

# Restrictions on ontological models from sequential measurements

*If I had a hammer,*
*I'd hammer in the morning*
*I'd hammer in the evening*
*All over this land*

– Pete Seeger and Lee Hays, *The Hammer Song*

In this chapter, we finally get around to discussing the restrictions that state update places on ontological models. This consists of one main theorem (Section 4.1) which we then apply to all three currently known examples of $\psi$-epistemic models in dimension $d \geq 3$ in order to show that they cannot represent state update (Section 4.2). Although we don't have a general $\psi$-epistemicity no-go theorem, we have taken our main theorem and applied it to the two most obvious contexts that might lead to such a theorem (Section 4.3). This leads to derivations of already-known restrictions on ontological models, which we include anyway as a demonstration of state update's relationship to these other properties.

## 4.1 The hammer: a general restriction on $\psi$-epistemic models via state update

We now prove our central claim that consideration of a rule for state update under measurement has consequences for the response function of an ontological model, so that consistent state update puts restrictions on how one may represent even a prepare-and-measure-once experiment. We begin with a lemma that articulates a general property of the update rule $\eta$, and then examine its consequences for response functions $\xi$.

**Lemma 1.** *Suppose we have an ontological model with ontic space* $\mathbf{\Lambda}$*, preparation distributions* $\mu(\lambda|P)$*, indicator functions* $\xi(k|\lambda, M)$*, and state update maps* $\eta(\lambda'|k, \lambda, M)$*. For a particular ontic state* $\lambda$ *and measurement projector* $\Pi$*, we define the set* $S_{\lambda,\Pi}$ *of quantum states that one could obtain after measurement of any quantum state consistent with* $\lambda$*:*

$$S_{\lambda,\Pi} = \left\{ \frac{\Pi\,|\phi\rangle}{\sqrt{\langle\phi|\Pi|\phi\rangle}} \,\middle|\, \forall\,|\phi\rangle : \lambda \in \Delta_\phi \right\}. \tag{4.1}$$

*It is then true that, independently of the measurement context of* $\Pi$*,*

$$\mathrm{Supp}(\eta(\cdot|\lambda, \Pi)) \subseteq \bigcap_{|\psi\rangle \in S_{\lambda,\Pi}} \Delta_\psi. \tag{4.2}$$

*Proof.* Suppose that measuring a state $|\phi\rangle$ with a measurement $M$ results in the updated state $|\psi\rangle$ when we get outcome $k$, where $\Pi$ is the $k$th projector in $M$. Then let $P_{M,k,P_\phi} \in \boldsymbol{P}_\psi$ be the preparation procedure associated with post-selection of this measurement outcome after a particular preparation $P_\phi$. It must be normalized on $\Delta_\psi$:

$$\begin{aligned}
1 &= \int_{\Delta_\psi} \mathrm{d}\lambda'\, \mu(\lambda'|P_{M,k,P_\phi}) \\
&= \int_{\Delta_\psi} \mathrm{d}\lambda' \int_{\Delta_\phi} \mathrm{d}\lambda\, \eta(\lambda'|k, \lambda, M)\mu(\lambda|P_\phi) \\
&= \int_{\Delta_\phi} \mathrm{d}\lambda\, \mu(\lambda|P_\phi) \int_{\Delta_\psi} \mathrm{d}\lambda'\, \eta(\lambda'|k, \lambda, M)
\end{aligned}$$

Since $\eta$ is always positive, normalization of $\mu(\lambda|P_\phi)$ then implies that

$$\int_{\Delta_\psi} \mathrm{d}\lambda'\, \eta(\lambda'|k, \lambda, M) = 1 \qquad \forall \lambda \in \Delta_\phi.$$

If $\eta$ is normalized on a region, its support must be contained in that region. Thus for all $\lambda$ that are consistent with some preparation $|\phi\rangle$ that could result in the post-measurement state $|\psi\rangle$,

$$\mathrm{Supp}(\eta(\cdot|k, \lambda, M)) \subseteq \Delta_\psi.$$

The fact that this is true for all $|\psi\rangle$ that could result from the measurement leads to Eq. 4.2. □

We note that Lemma 1 is easy to account for in $\psi$-ontic theories and for rank-1 measurements, since in both cases $S_{\lambda,\Pi}$ has a single element. This is why we were able to write down update rules for these situations in Section 2.5. Outside of these trivial cases, Eq. (4.2) is a very restrictive condition; depending on the structure of $S_{\lambda,\Pi}$, the intersection may be a very small set. In particular, if any two of the post-selected quantum states are orthogonal, then $S_{\lambda,\Pi}$ is empty.

**Theorem 1** (Main theorem)**.** *Suppose that a projector* $\Pi$ *maps two states* $|\alpha\rangle, |\beta\rangle$ *to ontologically distinct states* $\Pi|\alpha\rangle, \Pi|\beta\rangle$*. Then the response function for* $\Pi$ *cannot have support on the ontic overlap of* $|\alpha\rangle, |\beta\rangle$ *for any measurement context of* $\Pi$*; i.e.*

$$\Delta_{\Pi|\alpha\rangle} \cap \Delta_{\Pi|\beta\rangle} = \emptyset \implies \xi(\Pi|\lambda) = 0 \quad \forall \lambda \in \Delta_\alpha \cap \Delta_\beta. \tag{4.3}$$

*Proof.* Pick some $\lambda \in \Delta_\alpha \cap \Delta_\beta$. By Lemma 1, $\mathrm{Supp}(\eta(\cdot|\lambda, \Pi)) = \emptyset$ so $\eta(\lambda'|\lambda, \Pi)$ is not normalizable for this $\lambda$. As discussed in Section 2.1, this is only allowable if $\xi(\Pi|\lambda) = 0$. □

Both the lemma and the theorem hold for non-projective measurements as well. We emphasize that this result does not say anything *directly* about the overlap of the supports of quantum states, just their overlap within the support of a particular response function. In the following section, we deploy this theorem by showing that, in every known $\psi$-epistemic model for $d \geq 3$, measurements of this type exist *and* have support on the relevant overlaps, leading to contradiction and demonstrating that these models cannot reproduce state update under measurement.

## 4.2 The nails: known $\psi$-epistemic models cannot represent sequential measurement

### 4.2.1 LJBR

This model is the first to fall to Theorem 1:

**Theorem 2.** *The LJBR model cannot represent state update under measurement in dimension* $d \geq 3$*.*

*Proof.* The general idea of the proof is to find a measurement which maps any two nonidentical states $|\alpha\rangle, |\beta\rangle$ to two again nonidentical states $\Pi|\alpha\rangle, \Pi|\beta\rangle$ which both have no support on any of the $\mathcal{S}_j$, and so are ontologically distinct.

Consider the preferred basis $|j\rangle$ of the LJBR model and the generalized $x$-basis defined by

$$|X_k\rangle = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} \omega^{jk} |j\rangle, \qquad \omega = e^{2\pi i/d}. \tag{4.4}$$

These $x$-basis states have the property $\mathrm{tr}([X_k][j]) = \frac{1}{d}$ for all $j, k$. There must exist two elements $|X_{k_1}\rangle, |X_{k_2}\rangle$ of the $x$-basis such that $|\alpha\rangle, |\beta\rangle$ differ on that two-dimensional subspace or else $|\alpha\rangle, |\beta\rangle$ would be identical. Pick two such elements, and consider the projector $\Pi = [X_{k_1}] + [X_{k_2}]$. The quantum overlap of the post-measurement state $\Pi|\alpha\rangle$ with any basis vector $|j\rangle$ is, using the submultiplicativity of the trace,

$$\frac{\mathrm{tr}([j]\Pi[\alpha]\Pi)}{\mathrm{tr}(\Pi[\alpha])} \leq \mathrm{tr}([j]\Pi) = \frac{2}{d} \leq \frac{d-1}{d} \tag{4.5}$$

and the same is true for $\Pi|\beta\rangle$. As described above, only states with $\mathrm{tr}([j][\psi]) > \frac{d-1}{d}$ have overlap with any other states in the LJBR model, so the post-measurement states are ontologically distinct; by Theorem 1, $\xi(\Pi|\lambda) = 0$ for all $\lambda \in \mathcal{S}_j$ for all $j$ since $|\alpha\rangle$ and $|\beta\rangle$ were arbitrary.

However, when measured in the context of the rest of the rank-1 $x$-basis projectors, $\Pi$ will be ordered first by $\pi_{M,\lambda}$ for all $\lambda$ since

$$\mathrm{tr}(\Pi[j]) = \frac{2}{d} > \frac{1}{d} \tag{4.6}$$

for all $j$. Thus, by the construction of the LJBR model, $\xi(\Pi|\lambda) = 1$ for all $\lambda \in \mathcal{S}_j$, resulting in a contradiction. □

### 4.2.2 ABCL

**Theorem 3.** *The ABCL$_0$ model cannot represent state update under measurement in dimension $d \geq 3$.*

*Proof.* Call the two states defining the model $|\alpha\rangle, |\beta\rangle$. Let $\Pi = [\alpha] + [\gamma]$, where $|\gamma\rangle$ is some state such that

$$\langle\alpha|\gamma\rangle = 0 \quad \text{and} \quad 0 < |\langle\gamma|\beta\rangle|^2 < 1 - |\langle\alpha|\beta\rangle|^2. \tag{4.7}$$

Such a state exists by the results of Section 4.3.1. Under this measurement, $|\alpha\rangle$ maps to $|\alpha\rangle$ and $|\beta\rangle$ does not get mapped to either $|\alpha\rangle$ or $|\beta\rangle$. Thus the post-measurement states are ontologically distinct, so by Theorem 1, $\xi(\Pi|\lambda) = 0$ for all $\lambda \in \Delta_\alpha \cap \Delta_\beta$.

For the other half of the contradiction, note that $\mathrm{tr}(\Pi[\alpha]) = 1$ means $g_{\alpha\beta}(\Pi) = \mathrm{tr}(\Pi[\beta]) > 0$ (since $|\alpha\rangle, |\beta\rangle$ are nonorthogonal) and $g_{\alpha\beta}(\mathbb{I} - \Pi) = 1 - \mathrm{tr}(\Pi[\alpha]) = 0$, so $\Pi$ is ordered first by $\sigma_M$ when measured in the context $M = \{\Pi, \mathbb{I} - \Pi\}$. Thus $\xi(\Pi|\lambda) = 1$ for all $\lambda \in \Delta_\alpha \cap \Delta_\beta$, resulting in a contradiction. □

The failure of the $\mathrm{ABCL}_1$ model to reproduce state update follows directly from the failure of the $\mathrm{ABCL}_0$ model.

**Theorem 4.** *The ABCL₁ model cannot represent state update under measurement in dimension $d \geq 3$.*

*Proof.* When we take a convex combination of models, we see that

$$\begin{aligned} \mathrm{Supp}(\xi_3(k|\cdot, M)) &= \mathrm{Supp}(\xi_1(k|\cdot, M)) \cup \mathrm{Supp}(\xi_2(k|\cdot, M)) \\ \mathrm{Supp}(\eta_3(\cdot|k, \lambda, M)) &= \mathrm{Supp}(\eta_1(\cdot|k, \lambda, M)) \cup \mathrm{Supp}(\eta_2(\cdot|k, \lambda, M)) \end{aligned}$$

Thus if either of models 1 or 2 violates Theorem 1, model 3 must violate it as well. Since all of the $\mathrm{ABCL}_0$ models being mixed violate Theorem 1, so must $\mathrm{ABCL}_1$ . □

### 4.2.3 Kitchen Sink

**Theorem 5.** *For any finite subtheory containing states and measurements satisfying the conditions given in Eqs. 3.39–3.42, the Kitchen sink model cannot model state update under measurement.*

*Proof.* Theorem 1 implies that, if $\Pi$ maps $|\alpha\rangle, |\beta\rangle$ to ontologically distinct states, then

$$\int_{\mathbf{\Lambda}} \mathrm{d}\lambda\, \xi(\Pi|\lambda)\mu(\lambda|\alpha)\mu(\lambda|\beta) = 0. \tag{4.8}$$

We evaluate this quantity for the Kitchen Sink's response functions and preparation distributions, using the states and measurement $M^{(1)}$ satisfying Eqs. 3.39–3.42:

$$\begin{aligned}
&\int_{\mathbf{\Lambda}} \mathrm{d}\lambda\, \xi(k=0|\lambda, M^{(1)})\mu(\lambda|\alpha)\mu(\lambda|\beta) \\
&\quad = \sum_{\lambda \in \mathbb{Z}_r^{|\mathcal{M}|}} \delta(0,\lambda_1) \prod_{j=1}^{|\mathcal{M}|} \mathrm{tr}\Big(M^{(j)}_{\lambda_j}[\alpha]\Big)\, \mathrm{tr}\Big(M^{(j)}_{\lambda_j}[\beta]\Big) \\
&\quad = \sum_{\lambda_1 \in \mathbb{Z}_r} \delta(0,\lambda_1)\, \mathrm{tr}\Big(M^{(1)}_{\lambda_1}[\alpha]\Big)\, \mathrm{tr}\Big(M^{(1)}_{\lambda_1}[\beta]\Big) \\
&\qquad\qquad \cdot \prod_{j=2}^{|\mathcal{M}|} \sum_{l \in \mathbb{Z}_r} \mathrm{tr}\Big(M^{(j)}_{l}[\alpha]\Big)\, \mathrm{tr}\Big(M^{(j)}_{l}[\beta]\Big) \\
&\quad = \mathrm{tr}\Big(M^{(1)}_{0}[\alpha]\Big)\, \mathrm{tr}\Big(M^{(1)}_{0}[\beta]\Big) \\
&\qquad\qquad \cdot \prod_{j=2}^{|\mathcal{M}|} \sum_{l \in \mathbb{Z}_r} \mathrm{tr}\Big(M^{(j)}_{l}[\alpha]\Big)\, \mathrm{tr}\Big(M^{(j)}_{l}[\beta]\Big)
\end{aligned} \tag{4.9}$$

This final expression will be zero if and only if at least one of its factors is 0. The first two factors are nonzero by Eqs. 3.40 and 3.41. The rest of the factors are nonzero due to Eq. (3.39) and the completeness condition on the measurements. Thus Eq. 4.9 is nonzero and we have a contradiction, so state update cannot be represented faithfully. □

## 4.3 Can our result lead directly to a general $\psi$-epistemicity no-go theorem?

Given the fact that Theorem 1 has ruled out all known examples of $\psi$-epistemic ontological models, the natural question to ask next is whether or not it can rule out $\psi$-epistemic models categorically. However, this task is complicated by two issues that emerge when we are not discussing a specific ontological model. First, how do we ensure that the post-measurement quantum states are ontologically distinct? We address this by introducing what we call orthogonalizing measurements. Second, all that we can generally conclude is that a response function $\xi = 0$ on a certain domain, without concluding anything directly about whether or not two quantum states can overlap. To account for this, we manufacture situations in which, by nature of some well-known features of ontological models, the response function $\xi$ is *also* required to be 1 on a certain domain, reaching a contradiction in a similar way to the proofs in the previous section. Unfortunately, the two obvious ways to set this up both result in restrictions that were already known from considerations other than state update. We go through their proofs anyway to demonstrate their relation to state update.

### 4.3.1 Orthogonalizing measurements

The only obvious way to apply Theorem 1 outside the context of a specific ontological model is to use a measurement of the following kind:

**Definition 7** (Orthogonalizing projector)**.** *A projector* $\Pi$ orthogonalizes *two states* $|\alpha\rangle$ *and* $|\beta\rangle$ *if* $\langle\beta|\,\Pi\,|\alpha\rangle = 0$*,* $\langle\alpha|\Pi|\alpha\rangle \neq 0$*, and* $\langle\beta|\Pi|\beta\rangle \neq 0$*.*

This means that if we measure $|\alpha\rangle$ and $|\beta\rangle$ with a measurement containing $\Pi$, then upon obtaining the outcome $\Pi$ (which has nonzero probability) we know that they have been mapped to orthogonal states. Orthogonal states are always ontologically distinct [80], so we can apply Theorem 1.

Before we apply orthogonalizing measurements to ontological models, we will discuss some of their interesting features. As long as we are considering systems in $d \geq 3$ dimensions, it is easy to construct two states that are orthogonalized by any given rank-2 measurement and easy to construct a (rank-2) measurement that orthogonalizes any two given (non-identical) states. The constructions are similar, so we focus on the latter. Given any two states $|\alpha\rangle, |\beta\rangle$, we can define a basis such that

$$\begin{aligned}|\alpha\rangle &= |0\rangle \\ |\beta\rangle &= \langle\alpha|\beta\rangle\,|0\rangle + \sqrt{1-|\langle\alpha|\beta\rangle|^2}\,|1\rangle\end{aligned} \tag{4.10}$$

and two more states $|\gamma\rangle, |\delta\rangle$ with a single free parameter $\langle\beta|\gamma\rangle$:

$$\begin{aligned}|\gamma\rangle &= \frac{\langle\beta|\gamma\rangle\,|1\rangle + \sqrt{1-|\langle\alpha|\beta\rangle|^2-|\langle\beta|\gamma\rangle|^2}\,|2\rangle}{\sqrt{1-|\langle\alpha|\beta\rangle|^2}} \\ |\delta\rangle &= \frac{\beta_1^*\gamma_2^*\,|0\rangle - \beta_0^*\gamma_2^*\,|1\rangle + \beta_0^*\gamma_1^*\,|2\rangle}{\sqrt{1-|\langle\beta|\gamma\rangle|^2}}\end{aligned} \tag{4.11}$$

where $\beta_1 = \langle 1|\beta\rangle, \gamma_2 = \langle 2|\gamma\rangle$, etc. Since, as can be checked from their definitions, $\langle\delta|\gamma\rangle = 0$, we can use these to define a projector $\Pi = [\delta] + [\gamma]$. Then we see that since $\langle\alpha|\gamma\rangle = \langle\beta|\delta\rangle = 0$, the relations

$$\begin{aligned}\langle\beta|\,\Pi\,|\alpha\rangle &= \langle\beta|\gamma\rangle\langle\gamma|\alpha\rangle + \langle\beta|\delta\rangle\langle\delta|\alpha\rangle = 0\\ \langle\beta|\Pi|\beta\rangle &= |\langle\beta|\gamma\rangle|^2 \neq 0\\ \langle\alpha|\Pi|\alpha\rangle &= |\langle\alpha|\delta\rangle|^2 = \frac{1-|\langle\alpha|\beta\rangle|^2 - |\langle\beta|\gamma\rangle|^2}{1-|\langle\beta|\gamma\rangle|^2} \neq 0\end{aligned} \tag{4.12}$$

are satisfied as long as $0 < |\langle\beta|\gamma\rangle| < \sqrt{1-|\langle\alpha|\beta\rangle|^2}$. Thus $\Pi = [\gamma] + [\delta]$ orthogonalizes $|\alpha\rangle, |\beta\rangle$ for any value of the free parameter $\langle\beta|\gamma\rangle$ in the range specified above.

Orthogonalizing measurements clearly have applications to distinguishability and can be used for unambiguous state discrimination [79]: if we obtain the outcome $\Pi$ from a measurement, then we can make a subsequent measurement that distinguishes with certainty whether $|\alpha\rangle$ or $|\beta\rangle$ was initially prepared. Defining

$$g_{\alpha\beta}(\Pi) = \min\{\mathrm{tr}(\Pi[\alpha]), \mathrm{tr}(\Pi[\beta])\}, \tag{4.13}$$

this protocol guarantees a success probability of at least $g_{\alpha\beta}(\Pi)$ regardless of the relative frequencies with which $|\alpha\rangle$ or $|\beta\rangle$ are prepared. As such, we might expect a bound on $g_{\alpha\beta}(\Pi)$ that is a function of the quantum overlap $\langle\alpha|\beta\rangle$:

**Theorem 6.** *If a projector $\Pi$ orthogonalizes $|\alpha\rangle$ and $|\beta\rangle$, then $g_{\alpha\beta}(\Pi) \leq 1 - |\langle\alpha|\beta\rangle|$. Furthermore, this is a tight bound.*

*Proof.* We use the submultiplicativity of trace along with $\mathrm{tr}([\alpha]\Pi[\beta]) = 0$:

$$\begin{aligned}\mathrm{tr}([\alpha](\mathbb{I}-\Pi)[\beta](\mathbb{I}-\Pi)) &\leq \mathrm{tr}([\alpha](\mathbb{I}-\Pi))\,\mathrm{tr}([\beta](\mathbb{I}-\Pi))\\ \mathrm{tr}([\alpha][\beta]) &\leq (1-\mathrm{tr}([\alpha]\Pi))(1-\mathrm{tr}([\beta]\Pi))\\ \mathrm{tr}([\alpha]\Pi) &\leq 1 - \frac{|\langle\alpha|\beta\rangle|^2}{1-\mathrm{tr}([\beta]\Pi)}\\ \implies g_{\alpha\beta}(\Pi) &\leq \min\{1-\frac{|\langle\alpha|\beta\rangle|^2}{1-\mathrm{tr}([\beta]\Pi)}, \mathrm{tr}([\beta]\Pi)\}\end{aligned}$$

Because the first argument of min is a continuous and monotonically decreasing function of the second argument, the right-hand-side will be maximized when the two arguments are equal. This sets $\mathrm{tr}([\beta]\Pi) = 1 - |\langle\alpha|\beta\rangle|$, so

$$g_{\alpha\beta}(\Pi) \leq 1 - |\langle\alpha|\beta\rangle|\,. \tag{4.14}$$

Finally, to see that this bound is tight, we note that setting $\langle\beta|\gamma\rangle = \sqrt{1 - |\langle\alpha|\beta\rangle|}$ in Eq. 4.11 gives an example that saturates the bound. $\square$

The quantity $g_{\alpha\beta}(\Pi)$ was used to define the LJBR model discussed in Section 3.2.4 which, through the lens of orthogonalizing measurements, underscores the relationship between distinguishability and $\psi$-epistemic models.

Although we have shown above that orthogonalizing measurements are generic in the full quantum theory, one might wonder whether they exist in standard subtheories like the stabilizer subtheory described in Section 3.3.2. Although for a single qu$p$it ($p$ a prime), the stabilizer subtheory contains only rank-1 measurements, the $n$-qu$p$it stabilizer subtheory contains rank-$p^{n-1}$ measurements that describe measuring a single qu$p$it. Some of these higher-rank measurements can be orthogonalizing measurements: picking a two-qu$p$it subspace, we define the stabilizer states

$$\begin{aligned} |\alpha\rangle &= |0\rangle \left( |0\rangle + |1\rangle + \cdots + |p-1\rangle \right) \\ |\beta\rangle &= |00\rangle + |11\rangle + \cdots + |p-1\rangle\, |p-1\rangle \end{aligned} \tag{4.15}$$

which are stabilized by $S_\alpha = \langle Z_1, X_2\rangle$ and $S_\beta = \langle Z_1 Z_2^{p-1}, X_1 X_2\rangle$, respectively[1]. By examination, we see that if we measure $Z_2$ (i.e. measure in the standard basis on the second qubit), any outcome besides $|0\rangle$ will orthogonalize these two states.

A related idea to orthogonalizing measurements, called partitioning measurements, was defined in [52]. In our language, a partitioning measurement for a set $\{|\psi_i\rangle\}$ of states is one where each measurement outcome orthogonalizes at least one pair of states in $\{|\psi_i\rangle\}$. We can also fine-grain orthogonalizing measurements (or partitioning measurements) in order to construct antidistinguishing measurements such as those used in the proof of the PBR theorem [68]. This relationship will be explored further in the next two sections.

### 4.3.2 Application 1: three-way overlap

As suggested above, we want to find some way to show that $\xi$ is not zero on some subset of the ontic space in order to derive a contradiction. The easiest way to do this is to find some state $|\alpha\rangle$ contained in the image of $\Pi$, so that the support of $\xi(\Pi|\lambda)$ must contain the support $\Delta_\alpha$ of the quantum state [45]. In other words,

$$\langle\alpha|\Pi|\alpha\rangle = 1 \implies \forall\lambda \in \Delta_\alpha : \xi(\Pi|\lambda) = 1. \tag{4.16}$$

[1] Here $Z_1$ indicates a $Z$ gate on the first qubit, $X_2$ indicates an $X$ gate on the second qubit, etc.

We then consider two other states $|\beta\rangle, |\gamma\rangle$ that are orthogonalized by $\Pi$, so $\xi(\Pi|\lambda) = 0 \quad \forall\lambda \in \Delta_\beta \cap \Delta_\gamma$. Clearly then $\Delta_\alpha \cap \Delta_\beta \cap \Delta_\gamma = \emptyset$.

For which triples of states $|\alpha\rangle, |\beta\rangle, |\gamma\rangle$ does such a projector $\Pi$ exist? We first note that, if $\Pi$ orthogonalizes $|\beta\rangle, |\gamma\rangle$, then we can write

$$\Pi = \frac{\Pi[\beta]\Pi}{\mathrm{tr}(\Pi[\beta])} + \frac{\Pi[\gamma]\Pi}{\mathrm{tr}(\Pi[\gamma])}, \tag{4.17}$$

where each term is a rank-1 projector. Since $\Pi$ orthogonalizes $|\beta\rangle$ and $|\gamma\rangle$, it follows that $\mathrm{tr}\left(\frac{\Pi[\beta]\Pi}{\mathrm{tr}(\Pi[\beta])}[\gamma]\right) = 0$, and vice versa. Also note that, since $\mathrm{tr}(\Pi[\alpha]) = 1$, we can infer that $I - \Pi$ will have 0 probability of responding to $\alpha$.

Now consider the measurement

$$M = \{\frac{\Pi[\beta]\Pi}{\mathrm{tr}(\Pi[\beta])}, \frac{\Pi[\gamma]\Pi}{\mathrm{tr}(\Pi[\gamma])}, I - \Pi\}. \tag{4.18}$$

The first outcome has 0 probability of responding to $|\gamma\rangle$, the second outcome has 0 probability of responding to $|\beta\rangle$, and the final outcome has 0 probability of responding to $|\alpha\rangle$. This type of measurement is called an antidistinguishing measurement. When this type of measurement exists for a set of states, they are called antidistinguishable [68] or post-Peierls incompatible [18]. These arguments work in reverse, so we see that our desired construction of measurement exists if and only if $|\alpha\rangle, |\beta\rangle, |\gamma\rangle$ are antidistinguishable. Unfortunately, it is already known that an antidistinguishable set of states must have no mutual overlap [9], as this is the main tool leveraged in the PBR theorem [68]. It is interesting to note that in [18], the authors show that three pure states are antidistinguishable if and only if

$$\begin{gathered}\left(1 - |\langle\alpha|\beta\rangle|^2 - |\langle\beta|\gamma\rangle|^2 - |\langle\gamma|\alpha\rangle|^2\right)^2 \geq 4\,|\langle\alpha|\beta\rangle|^2\,|\langle\beta|\gamma\rangle|^2\,|\langle\gamma|\alpha\rangle|^2 \\ \text{and} \\ |\langle\alpha|\beta\rangle|^2 + |\langle\beta|\gamma\rangle|^2 + |\langle\gamma|\alpha\rangle|^2 \leq 1,\end{gathered} \tag{4.19}$$

which gives an easily checkable condition for three states being able to have mutual overlap.

### 4.3.3 Application 2: state update + unitaries ≈ CPTP maps

Three-way overlap is fairly well understood in the literature, and doesn't actually bear directly on the question of the possibility of a $\psi$-epistemic model. We now try to find a restriction on pairwise overlaps. The obvious place to start would be to use the setup from the last section, and simply set $|\gamma\rangle = |\alpha\rangle$; unfortunately, this immediately fails the test in Eq. 4.19. As such, we set our sights on a slightly weaker condition:

**Lemma 2.** *Given nonorthogonal* $|\alpha\rangle, |\beta\rangle$, *and a target overlap* $0 < t \leq 1$, *then there exists* $\Pi$ *s.t.*

$$\operatorname{tr}(\Pi[\alpha]) = 1 \tag{4.20}$$

$$\frac{\operatorname{tr}(\Pi[\alpha]\Pi[\beta])}{\operatorname{tr}(\Pi[\alpha])\operatorname{tr}(\Pi[\beta])} = t \tag{4.21}$$

*if and only if* $t \geq \operatorname{tr}([\alpha][\beta])$.

*Proof.* Let $\Pi = [\alpha] + [\gamma]$ and choose a basis such that

$$|\alpha\rangle = |0\rangle \tag{4.22}$$

$$|\beta\rangle = \beta_0 |0\rangle + \sqrt{1 - |\beta_0|^2}\, |1\rangle \tag{4.23}$$

$$|\gamma\rangle = \gamma_1 |1\rangle + \sqrt{1 - |\gamma_1|^2}\, |2\rangle . \tag{4.24}$$

Eq. 4.20 is satisfied by the definition of $\Pi$, and Eq. 4.21 requires

$$|\gamma_1|^2 = \frac{|\beta_0|^2(1-t)}{t(1-|\beta_0|^2)}. \tag{4.25}$$

Then

$$|\gamma_1|^2 \geq 0 \iff |\beta_0|^2(1-t) \geq 0 = \text{True} \tag{4.26}$$

$$|\gamma_1|^2 \leq 1 \iff t \geq |\beta_0|^2 \tag{4.27}$$

where $|\beta_0|^2 = \operatorname{tr}([\alpha][\beta])$. □

Why is this useful? In [55, Section 8.1], Leifer shows[2] that, if unitary transformations are included in the operational theory, then ontological distinctness can depend only on the inner product of the states:

$$\operatorname{tr}([\alpha][\beta]) = \operatorname{tr}([\alpha][\beta']) \implies \left(\Delta_\alpha \cap \Delta_\beta = \emptyset \iff \Delta_{\alpha'} \cap \Delta_{\beta'} = \emptyset\right) \tag{4.28}$$

[2]Leifer doesn't claim that this proof is original, but also gives no attribution in the section, and I can't find it anywhere else in the literature, so I will attribute it to him.

Combining this with Lemma 2 and Theorem 1, we see that if a particular pair of quantum states $|\alpha\rangle, |\beta\rangle$ is ontologically distinct, and $|\alpha\rangle', |\beta\rangle'$ have a larger inner product than them, then we can choose a projector that maps $|\alpha\rangle', |\beta\rangle'$ to a pair of states with the same inner product as $|\alpha\rangle, |\beta\rangle$, and thus conclude that they are ontologically distinct. We have arrived at the much stronger condition

**Theorem 7.**

$$\mathrm{tr}([\alpha][\beta]) \leq \mathrm{tr}\big([\alpha]'[\beta]'\big) \wedge \Delta_{\alpha'} \cap \Delta_{\beta'} = \emptyset \implies \Delta_\alpha \cap \Delta_\beta = \emptyset. \tag{4.29}$$

Again, this is not exactly a new restriction. In [55, Section 8.1], Leifer notes that this result can be obtained by including CPTP maps in the operational theory. However, we have come to the same conclusion while remaining in pure-state quantum theory; this suggests that, for a $\psi$-onticity theorem, we may not need to look at mixed state quantum theory, but rather pure state quantum theory is sufficient. This is the tack we follow in Chapter 5.

### 4.3.4 A note on the relationship between state update and transformations

As noted in the previous section, transformations also play a role in restricting the structure of $\psi$-epistemic ontological models [55, Section 8.1]. If an ontological model successfully represents all unitary transformations, then $|\langle\psi|\phi\rangle| = |\langle\psi'|\phi'\rangle|$ implies that $|\psi\rangle, |\phi\rangle$ are ontologically distinct if and only if $|\psi'\rangle, |\phi'\rangle$ are ontologically distinct. If the model also includes all CPTP maps, then $|\langle\psi|\phi\rangle| \geq |\langle\psi'|\phi'\rangle|$ implies that $|\psi'\rangle, |\phi'\rangle$ are ontologically distinct if $|\psi\rangle, |\phi\rangle$ are.

It immediately follows that the LJBR and $\mathrm{ABCL}_0$ models cannot faithfully represent unitary transformations. In each model there exist quantum states which are ontologically distinct from every other state; pick one of these states, and it is easy to find examples of pairs of ontologically distinct states with any inner product.

However, transformations cannot necessarily rule out the $\mathrm{ABCL}_1$ model: since it is pairwise $\psi$-epistemic, the unitary condition could *in principle* be satisfied. That said, the transformation rule would be complicated because it would have to map *between* models that are mixed together, so it is certainly an open question whether this is actually possible.

The Kitchen Sink in fact *can* represent measurements, as described in Section 3.3.1, so it seems that (at least for this set of examples), state update puts a more severe restriction on ontological models than transformations do.

# Chapter 5

# Towards a $\psi$-epistemicity no-go theorem via computational mechanics

*Well I'm on my way*
*I don't know where I'm going*
*I'm on my way*
*Taking my time*
*But I don't know where*

– Paul Simon, *Me and Julio Down by the Schoolyard*

Now that we've established that a $\psi$-epistemicity no-go theorem without any additional assumptions is once again a possibility, the natural course of action is to start trying to prove one. Unfortunately, as described in Section 4.3, the results described so far do not lead us directly to one. In this section, we describe a powerful framework that may be able to prove such a no-go theorem, and describe the way in which we think it may do so. The framework is based on computational mechanics, so we begin by describing that research program (Section 5.1); next, we discuss how it relates/applies to quantum mechanics and the problem under consideration (Section 5.2). Finally, we present the progress so far and ideas for how to wring a $\psi$-onticity theorem out of this framework (Section 5.4).

Due to the in-progress nature of this work, this chapter is significantly less well organized than the previous ones and may contain small or large errors.

## 5.1 Introducing computational mechanics

Computational mechanics [21, 78] is a research program stretching back 40 years that aims to understand the informational and computational properties of stochastic processes and complex systems. In many ways, it is similar to quantum foundations in its conception and ethos, and might be described as "foundations of complex systems," "foundations of information processing," or something the like. Its main tool, the $\epsilon$-machine, is the (unique) hidden Markov model that is simultaneously maximally predictive and minimally entropic; the uniqueness of an $\epsilon$-machine allows one to interpret its properties as intrinsic properties of the stochastic process that it represents. The whole of computational mechanics is built on top of Shannon information theory, for which we include a brief introduction in Appendix B.

We will introduce the basic concepts of computational mechanics via much hand-waving and under-rug-sweeping, as a full review of the subject is far outside the scope of this thesis. For a recent summary of the history and wide applicability of this research program, we recommend [21]. For a rigorous mathematical treatment of the basic tools, [78] provides a comprehensive and clear development[1].

We pick up the thread of thought, and more importantly notation, from Section 2.6 where we showed that ontological models are equivalent to hidden Markov models. We have had to compromise between the notations for ontological models and computational mechanics. We follow the (computational mechanics) convention that probability theory uses capitalized letters in a particular way: $X$ is a random variable, $\boldsymbol{X}$ is its alphabet (it's usually $\mathcal{X}$ but we want this available for Greek letters as well) and $x \in \boldsymbol{X}$ is a particular element of its alphabet. We follow usual simplifying conventions to write $\Pr_X(x) = \Pr(X = x) = \Pr(x)$. The case where this differs from standard ontological models is in $\Lambda$, which is now the random variable representing the ontic state rather than the state space (which is $\boldsymbol{\Lambda}$, as we have used so far in this work).

For a moment, rather than describing an input-output process $\Pr(\overleftrightarrow{K} | \overleftrightarrow{A})$, we will look at just a bi-infinite process with no input; these are described by a distribution $\Pr(\overleftrightarrow{K})$. In this case, a hidden Markov model of this process is some random variable $\Lambda$ that renders the future conditionally independent of the past; in symbols, $\Pr(\overset{\bullet\!\!\rightarrow}{K} | \overleftarrow{K}, \Lambda) = \Pr(\overset{\bullet\!\!\rightarrow}{K} | \Lambda)$. This is equivalent to saying that the past and future share no mutual information when conditioned on the state $\Lambda$:

$$I[\overleftarrow{K}; \overset{\bullet\!\!\rightarrow}{K} | \Lambda] = 0. \tag{5.1}$$

[1] Incidentally, this paper is the first place that I encountered the quote from Plato that opens this thesis.

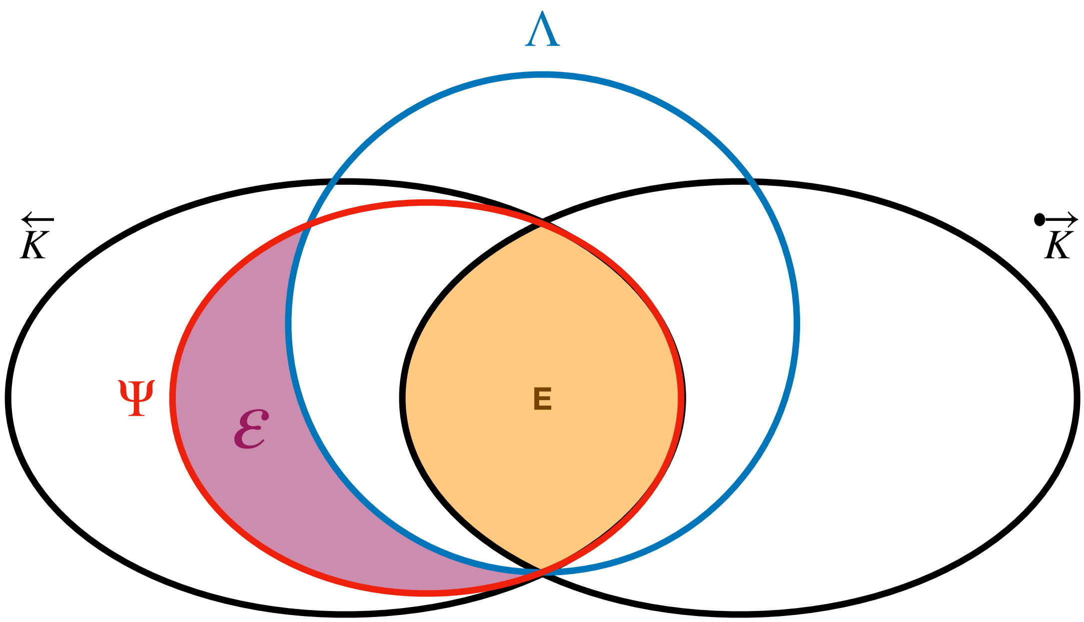



Figure 5.1: An I-diagram for a process with no input. The past and future outputs are represented by the $K$s with arrows, the $\epsilon$-machine by $\Psi$, and a generic HMM by $\Lambda$. The epistemicity $\mathcal{E}$ and excess entropy $E$ are defined in the text.

We will only use the information-theoretic statement of conditional independence from here forward, as it is more succinct. It is also very easy to visualize via I-diagrams [92] like Fig. 5.1. These look very much like Venn diagrams, and are not just a heuristic analogy for informational quantities but a strict formal representation. We introduce them more thoroughly in Appendix B, but roughly speaking, the region assigned to a random variable represents its entropy $H$; the overlaps between random variables are mutual informations $I$; and regions that are in one circle but excluded from another correspond to conditional entropies and conditional mutual informations. While the actual size of a region is arbitrary, the point of the diagram is that regions are additive. In this particular case, Eq. 5.1 is represented in Fig. 5.1 by the fact that $\Lambda$ completely contains the overlap between $\overleftarrow{K}$ and $\overset{\bullet\!\!\rightarrow}{K}$. This region, labeled $E = I[\overleftarrow{K}; \overset{\bullet\!\!\rightarrow}{K}]$, is called the *excess entropy* of the process and is a (usually loose) lower bound on the entropy of any HMM $\Lambda$.

The $\epsilon$-machine of this process can be defined in two equivalent ways [84]. Chronologically the first to be defined, a *history* $\epsilon$-machine has states which are equivalence classes of pasts generated by the equivalence relation $\underset{\epsilon}{\sim}$:

$$\overleftarrow{k} \underset{\epsilon}{\sim} \overleftarrow{k}' \iff \Pr(\overset{\bullet\!\!\rightarrow}{k} | \overleftarrow{k}) = \Pr(\overset{\bullet\!\!\rightarrow}{k} | \overleftarrow{k}') \qquad \forall \overset{\bullet\!\!\rightarrow}{k} . \tag{5.2}$$

This relation says that two pasts are equivalent if they make equivalent predictions for the future; in that case, we group those pasts into a single state, and only keep track of the state (the equivalence class) rather than the particular pasts. The fact that these states for form an HMM of the process should be fairly intuitive, but is also explicated further and proven in [78].

Equivalently, A *generator* $\epsilon$-machine is an unifilar irreducible HMM with probabilistically distinct states [7, Definition 3]. Irreducible means that every state is reachable from every other state with a finite probability. Unifilar means that $H[\Lambda'|K, \Lambda] = 0$, where $\Lambda'$ is the state one time step after $\Lambda$; in other words, state update maps (and transformations and preparations) are deterministic. Probabilistically distinct states means that $\forall \lambda, \lambda' \in \mathbf{\Lambda} : \exists \overset{\bullet\!\!\rightarrow}{k} : \Pr(\overset{\bullet\!\!\rightarrow}{k} |\lambda) \neq \Pr(\overset{\bullet\!\!\rightarrow}{k} |\lambda')$. In this case $\lambda, \lambda'$ are any two states that might obtain in the present.

The $\epsilon$-machine is shown on the information diagram in Fig. 5.1 labeled by $\Psi$ (for reasons that will be discussed shortly). Notice that it is contained entirely within the past outputs, as from the history $\epsilon$-machine definition it is easy to see that $H[\Psi|\overleftarrow{K}] = 0$.

As mentioned above, due to its uniqueness, some properties of the $\epsilon$-machine can be considered properties of the process. In particular, the *statistical complexity* $C = H[\Psi]$ is an important quantity that we will return to later.

That wraps up the most important elements in the usual setting for computational mechanics, which concerns the $\epsilon$-machines of no-input stochastic processes. Recently, the computational mechanics group has also discussed input-output processes [7]. These are processes of the type $\Pr(\overleftrightarrow{K}|\overleftrightarrow{A})$ that we discussed in Section 2.6. Although only history $\epsilon$-machines have been discussed in this setting, it is likely that they are still equivalent to generator $\epsilon$-machines. In this setting, we need just a few modifications to the definitions that we will rattle off quickly; in each case, it amounts to finding the properly motivated place to put $\overleftarrow{A}$ and $\overset{\bullet\!\!\rightarrow}{A}$.

An HMM for an input-output process is (equivalently to Def. 6) a random variable $\Lambda$ that satisfies

$$I[\overset{\bullet\!\!\rightarrow}{K}; \overleftarrow{K}, \overleftarrow{A}|\Lambda, \overset{\bullet\!\!\rightarrow}{A}] = 0. \tag{5.3}$$

This suggests that we should define the excess entropy as

$$E_{\overleftrightarrow{A}} = I[\overset{\bullet\!\!\rightarrow}{K}; \overleftarrow{K}, \overleftarrow{A}| \overset{\bullet\!\!\rightarrow}{A}] \tag{5.4}$$

so that it still serves as a lower bound on $H[\Lambda]$. We have subscripted with $\overleftrightarrow{A}$ to emphasize that, while the channel $\Pr(\overleftrightarrow{K}|\overleftrightarrow{A})$ is fixed, this quantity will in general depend on the choice of input process $\Pr(\overleftrightarrow{A})$. Similarly, we now denote the statistical complexity as [7]

$$C_{\overleftarrow{A}} = H[\Psi], \tag{5.5}$$

although the definition does not change. The statistical complexity only depends on the past inputs, hence its subscript $\overleftarrow{A}$. The history $\epsilon$-machine of a process is now defined by the equivalence relation [7]

$$(\overleftarrow{k}, \overleftarrow{a}) \sim_{\epsilon} (\overleftarrow{k}', \overleftarrow{a}') \iff \Pr(\overrightarrow{k} | \overleftarrow{k}, \overleftarrow{a}, \overrightarrow{a}) = \Pr(\overrightarrow{k} | \overleftarrow{k}', \overleftarrow{a}', \overrightarrow{a}) \qquad \forall \overrightarrow{k}, \overrightarrow{a} \,. \tag{5.6}$$

The generator $\epsilon$-machine for input-output processes has not been investigated in the computational mechanics literature, but it is likely that a straightforward generalization of the no-input case will still be equivalent to the history $\epsilon$-machine [7]. For the three properties above, we generalize by saying that the definition of irreducible is the same, but now 'unifilar' means

$$H[\Lambda'|K, \Lambda, A] = 0 \tag{5.7}$$

and 'probabilistically distinct states' is given by the condition [7]

$$\forall \lambda_1, \lambda_2 : \exists \overrightarrow{a}, \overrightarrow{k} : \Pr(\overrightarrow{k} | \lambda_1, \overrightarrow{a}) \neq \Pr(\overrightarrow{k} | \lambda_2, \overrightarrow{a}). \tag{5.8}$$

Whether constructed from the history or generator point of view, the $\epsilon$-machine is a unique HMM satisfying the given conditions[2]. That said, it is not the only HMM of the process. So, a natural question to ask is how the $\epsilon$-machine compares to other HMMs of the process, which is exactly the question discussed in [73] for binary Markov processes (with no input). In that work, I showed that the space of HMMs for those processes is large, and that in general those HMMs can have smaller entropy than the $\epsilon$-machine (As discussed later, this is a sufficient condition for those models to be considered something like $\psi$-epistemic). Unfortunately the $\epsilon$-machine is the only HMM that has been studied for input-output processes; as we will discuss in the next section, the ontological models formalism is expressly concerned with comparing alternate HMMs to the $\epsilon$-machine for input output processes. This means that if we are to prove a $\psi$-epistemicity no-go theorem, a challenging prerequisite will be to merge these two previously studied branches of computational mechanics. We take some baby steps in this work, and although this

[2]For the input-output case, equivalence of the history and generator constructions has not yet been rigorously proven, but probably holds.

process of merging will certainly be interesting in its own right, it is unlikely that we will have to develop it quite as far as the individual branches if our only intention is to prove a $\psi$-onticity theorem for quantum theory.

## 5.2 What does this have to do with quantum theory?

In order to convince the reader that the previous section was worthwhile, we will proceed with a few key observations:

**1: $\psi$ is the $\epsilon$-machine of pure-state quantum theory, and $\rho$ is the $\epsilon$-machine of mixed-state quantum theory.** To be precise, by $\psi$ we mean the Beltrametti-Bugajski model from Section 3.2.1 and by $\rho$ we mean the $\rho$-ontic model of Section 3.2.2. In both cases, we have already proven that the models are HMMs of their respective processes, so we just need to check the three conditions that define a generator $\epsilon$-machine.

**Theorem 8.** *The Beltrametti-Bugajski model is the $\epsilon$-machine of pure-state quantum theory.*

*Proof.*

1. The model is irreducible because we can always apply a unitary transformation to get from any one state to another.[3]

2. It is unifilar because all three of $\mu, \Gamma$, and $\eta$ are $\delta$ functions and so have zero entropy.

3. The states are probabilistically distinct, because if we measure a particular quantum state $|\psi\rangle$ with a measurement including $[\psi]$, $\psi$ is the only quantum state that gives the outcome $[\psi]$ with certainty.

□

[3] In this and the following case, it's actually sufficient that we can re-prepare any quantum state/density matrix that we like. For some reason I find this unsatisfactory, hence the tack taken above.

**Theorem 9.** *The $\rho$-ontic model is the $\epsilon$-machine of mixed-state quantum theory.*

*Proof.*

1. Irreducibility is a little more complicated in this case (if we don't take the easy-out of preparations), as we don't have a set of reversible transformations between all density matrices. If we're trying to get from $\rho \to \rho'$, then if $\mathrm{tr}\big(\rho'^2\big) \leq \mathrm{tr}(\rho^2)$, there exists a $CPTP$ map that maps $\rho \to \rho'$. Otherwise, we can use some projective measurement to map $\rho \to \rho''$ where $\rho''$ is some pure state density matrix, and then find a CPTP map to map $\rho'' \to \rho'$.

2. Again, $\mu, \Gamma$, and $\eta$ are all delta functions, so the model is unifilar.

3. Finally, the states are probabilistically distinct by Gleason's theorem [38].

□

Recall that the $\epsilon$-machine is unique for any particular process; as a sanity check, we can make sure that $\psi$ does *not* satisfy the conditions of an $\epsilon$-machine for mixed-state quantum theory even though it is a valid HMM. In particular, although it is still irreducible and still has probabilistically distinct states, it is no longer unifilar. This is because, for a non-unitary CPTP map, a pure quantum state will be mapped to a mixed quantum state, which in this case would be represented as a convex combination of pure quantum states. Thus it can be mapped to more than one state, and so is not unifilar.

**2: Computational mechanics gives a succinct, precise expression of $\psi$-epistemicity and other quantities of interest in quantum foundations.** Besides noticing the $\epsilon$-machines above, my other main motivation for applying computational mechanics to ontological models is the Refinement Lemma [78, Lemma 7]. This lemma says that any model in a certain class of HMMs (not all HMMs) must be a fine graining (i.e. refinement) "almost everywhere" of the $\epsilon$-machine states. "Almost everywhere" refers to the fact that this may not hold on measure-0 subsets, which are irrelevant to the functioning of the model. This sounds a lot like the partitioning of the state space that happens in $\psi$-ontic models shown in Fig. 2.2, and is in fact the exact same condition. A later paper [22] expresses this concept that a model $\mathcal{R}$ is a 'refinement almost everywhere' of the $\epsilon$-machine $\mathcal{S}$ as the equivalent condition $H[\mathcal{S}|\mathcal{R}] = 0$. Re-expressing in terms of our relevant variables, with $\Psi$ as our $\epsilon$-machine and $\Lambda$ our rival model, this reads $H[\Psi|\Lambda] = 0$ and is exactly equivalent to the definition of $\psi$-onticity in the ontological models formalism.

Thus, we define the *epistemicity* $\mathcal{E}_{\overleftarrow{A}}$ of a model as

$$\mathcal{E}_{\overleftarrow{A}} = H[\Psi|\Lambda]. \tag{5.9}$$

With this, we can say that a model $\Lambda$ is $\psi$-ontic if $\forall \overleftarrow{A} : \mathcal{E}_{\overleftarrow{A}} = 0$, and $\psi$-epistemic if $\exists \overleftarrow{A} : \mathcal{E}_{\overleftarrow{A}} > 0$. One advantage of this formulation is that we now have a single real number to describe this distinction, rather than a property of probability distributions. Additionally, all of the "measure-0" caveats that accompany the usual definition are automatically accounted for by this one.

A side-effect is that we have sort of implicitly defined a notion of $\psi$-epistemicity that *applies to any stochastic process*, not just quantum theory. We will call this notion $\epsilon$-epistemicity, where we simply replace $\Psi$ in Eq. (5.9) by the relevant $\epsilon$-machine for the process. This allows us to say that, although mixed state quantum theory may not allow $\psi$-epistemic interpretations, it *does* allow an $\epsilon$-epistemic interpretation, which in this case we might call $\rho$-epistemic. This is because, as described above, the Beltrametti-Bugajski model is such a model for mixed state quantum theory. We will return to this concept in the next section.

Montina briefly touches on a related criterion in [65], but refers to $I[\Psi; \Lambda]$ rather than $\mathcal{E} = H[\Psi|\Lambda]$[4]. These quantities are related by $I[\Psi; \Lambda] = H[\Psi] - \mathcal{E}$, but because of (a) the quantity $H[\Psi]$ in the equation and (b) some pesky infinities, they are not sufficient to describe one another. It is tempting to say that $I[\Psi; \Lambda] = H[\Psi]$ is an equivalent condition to $\mathcal{E} = 0$, but this fails due to the fact that any of the three quantities may be infinite; thus even if $\mathcal{E} > 0$, it is possible that $I[\Psi; \Lambda] = H[\Psi] = \infty$ (this is the case for the LJBR model, as Montina notes). The mutual information is still an interesting quantity, in particular for the communication complexity problem addressed by Montina, but it does not fully capture the notion of $\psi$-epistemicity.

There are some other properties of ontological models that have nice information-theoretic descriptions. For example, outcome determinism is expressible as the (single-symbol) condition

$$H[K|\Lambda, A] = 0. \tag{5.10}$$

This says that, if you know the ontic state and the choice of action/measurement $a$, then you can predict the next outcome $k$ with certainty.

There is also a form of preparation noncontextuality that is expressible as

$$I[\Lambda; \overleftarrow{K}, \overleftarrow{A}|\Psi] = 0. \tag{5.11}$$

This says that the experimental procedure by which one arrives at a state $\lambda$ is irrelevant if you know the quantum state $\psi$ that results from that experimental procedure. In the

[4]We leave out the subscript on $\mathcal{E}$ for this case because Montina considers only the uniform distribution over preparations as his input.

context of pure state quantum theory that we are mostly working in, this condition is equivalent to preparation noncontextuality for pure states (PNCfPS), which is a very mild condition that every model in Chapter 3 satisfies. In the context of mixed state quantum theory, we get the usual definition of generalized preparation noncontextuality, which is known to be an unsatisfiable condition for ontological models [80]. It would be interesting to see whether it is possible to prove this result using the present information-theoretic framework, as a warm-up for a $\psi$-epistemicity no-go theorem.

As described in Section 2.7, there is a new form of preparation contextuality that becomes apparent in the HMM picture. In our new information theory notation, Eq. 2.31 can be expressed as

$$I[\overset{\bullet\!\rightarrow}{K}; \overleftarrow{K}, \overleftarrow{A} | K, A] = 0 \not\Rightarrow I[\Lambda'; \Lambda | K, A] = 0. \tag{5.12}$$

These last couple of paragraphs raise the question of whether or not transformation contextuality and measurement contextuality are expressible in information-theoretic terms. So far I have not found a way of doing this, but it may exist. Similarly, Kochen-Specker contextuality, deficiency, pairwise $\psi$-epistemicity, never-$\psi$-onticity, and other properties do not have immediately obvious expressions.

**3: Computational mechanics provides a powerful framework for bounding these quantities.** Because of the positivity of most Shannon information quantities, along with the many additive relationships between them, it is very common to be able to put bounds on these quantities. For a toy example, take the simple I-diagram for a no-input process in Fig. 5.1. In that case, all of the atoms of the I-diagram are positive, so it is easy to read off from the diagram that

$$\mathcal{E} \leq C - E, \tag{5.13}$$

where $C = H[\Psi]$ is defined above. Recall that both $C$ and $E$ are properties of the process, and not the model $\Lambda$, so we have *bounded a property of the model by properties of the process*. Thus, if it were the case that $C - E = 0$ for some process, we could conclude that it admits no $\epsilon$-epistemic model. The input-output process version of this gets significantly more complicated, but also gives us more powerful tools to work with, as we will describe in the next section.

## 5.3 Review: computational mechanics and quantum theory

This is not the first work to look at the relationship between computational mechanics and quantum theory. Most of the work that combines the two does so in a way that contrasts with our own: rather than using classical HMMs to model quantum processes, they use a quantum version of HMMs to model classical processes [2–4, 42, 60, 66, 70, 82]. As such, their results are not directly relevant to the subject under consideration.

The exception is a recent work by Cabello et al. [16] which actually applies computational mechanics directly to quantum foundations in much the same way that we do here. The authors use it in combination with Landauer's principle in order to evaluate the thermodynamic cost of state-realist interpretations of quantum theory. In this way, they claim to show that any ontological model with finite memory must dissipate an unbounded amount of heat during every measurement, which is surely an undesirable consequence. Unfortunately, the work makes a subtle but severe technical error that renders their result a tautology.

This occurs when the authors claim that an $\epsilon$-machine is the HMM producing minimum heat. In their argument, they show that it is minimal among *prescient rivals*, which are actually a severely restricted set of HMMs. The fact that there can exist non-prescient HMMs with smaller state-entropy than the $\epsilon$-machine (and that they are in some sense generic) was shown in [73]. In particular, prescient rivals are a proper subset of $\psi$-ontic models, so the authors have accidentally assumed $\psi$-onticity. This contradicts their other assumption about ontological models having finite memory, since quantum states are continuous. It is then no surprise that an infinite amount of information must be updated during measurement, since we must update the entirety of the quantum state. In order to show that their results actually holds for all ontological models, they would need to essentially prove a $\psi$-epistemicity no-go theorem, which would then render their result trivial.

## 5.4 Preliminary results & future directions

To begin with, we have drawn the general I-diagram for a general input-output process, its $\epsilon$-machine $\Psi$, and an arbitrary HMM $\Lambda$ in Fig. 5.2. Immediately, we see that the epistemicity is no longer a single atom but the sum of nine $\mathcal{E}_i$, which have been defined in the diagram: $\mathcal{E}_{\overleftarrow{A}} = \sum_{i=1}^{9} \mathcal{E}_i$. Additionally, whereas the no-input case has all-nonnegative

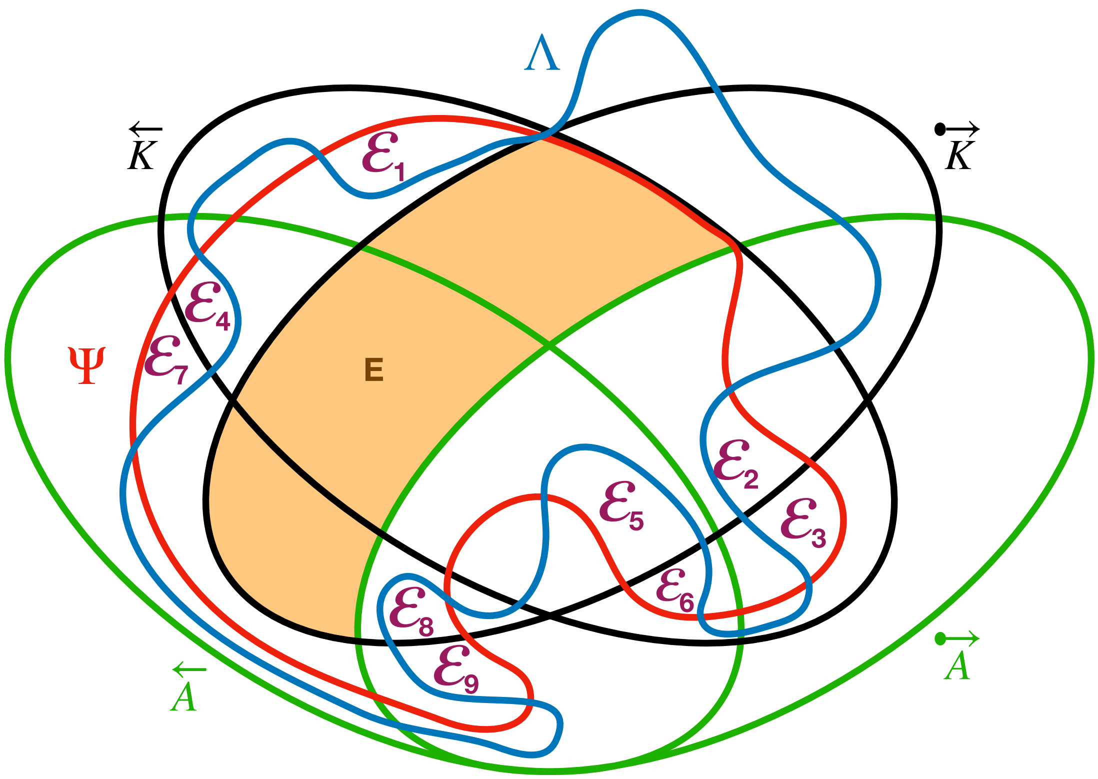



Figure 5.2: An I-diagram of an input-output process, its $\epsilon$-machine $\Psi$, and an arbitrary HMM $\Lambda$. The $A$s are inputs, and $K$s are outputs. The epistemicity $\mathcal{E}$ and excess entropy $E$ are defined in the text; this diagram defines the $\mathcal{E}_i$ which sum to $\mathcal{E}$. This figure is affectionately named the 'squiggle monster.'

atoms, many of the $\mathcal{E}_i$ (and other atoms) can now be negative, which significantly complicates the bounding process. In particular, using $\ngeq 0$ to mean "not *necessarily* nonnegative" we can read off from the diagram that

$$
\begin{array}{l|l}
\mathcal{E}_1 = H[\Psi|\overleftrightarrow{A},\Lambda] \geq 0 & \mathcal{E}_6 = I[\Psi;\overleftarrow{A};\overset{\bullet\!\!\rightarrow}{A};\overleftarrow{K}|\overset{\bullet\!\!\leftrightarrow}{K},\Lambda] \ngeq 0 \\
\mathcal{E}_2 = I[\Psi;\overset{\bullet\!\!\leftrightarrow}{K}|\overleftarrow{A},\Lambda] \geq 0 & \mathcal{E}_7 = H[\Psi|\overleftarrow{K},\overset{\bullet\!\!\rightarrow}{A},\Lambda] \geq 0 \\
\mathcal{E}_3 = I[\Psi;\overset{\bullet\!\!\rightarrow}{A}|\overset{\bullet\!\!\leftrightarrow}{K},\overleftarrow{A},\Lambda] \geq 0 & \mathcal{E}_8 = I[\Psi;\overset{\bullet\!\!\leftrightarrow}{K}|\overleftarrow{K},\Lambda] \geq 0 \\
\mathcal{E}_4 = I[\Psi;\overleftarrow{K};\overleftarrow{A}|\overset{\bullet\!\!\rightarrow}{A},\Lambda] \ngeq 0 & \mathcal{E}_9 = I[\Psi;\overset{\bullet\!\!\rightarrow}{A}|\overleftrightarrow{K},\Lambda] \geq 0 \\
\mathcal{E}_5 = I[\Psi;\overleftrightarrow{K};\overleftrightarrow{A}|\Lambda] \ngeq 0 &
\end{array}
\tag{5.14}
$$

However, we can group the negative quantities with positive ones in such a way that their sum is positive; for example,

$$
\begin{aligned}
\mathcal{E}_1 + \mathcal{E}_4 &= I[\Psi;\overleftarrow{K}|\overset{\bullet\!\!\rightarrow}{A},\Lambda] \geq 0 \\
\mathcal{E}_4 + \mathcal{E}_7 &= I[\Psi;\overleftarrow{A}|\overset{\bullet\!\!\rightarrow}{A},\Lambda] \geq 0 \\
\mathcal{E}_3 + \mathcal{E}_6 + \mathcal{E}_9 &= I[\Psi;\overset{\bullet\!\!\rightarrow}{A}|\overset{\bullet\!\!\leftrightarrow}{K},\Lambda] \geq 0.
\end{aligned}
$$

These combinations are non-unique, as can be seen from the fact that $\mathcal{E}_4$ can be combined with either $\mathcal{E}_1$ or $\mathcal{E}_7$ to form a positive quantity. Actually, it is less about the positivity of $\mathcal{E}_i$ and more about the positivity of the surrounding atoms, but these are usually the same condition. Deciding which ones are relevant/boundable will be part of the required future work.

Above, we claimed that, although more complicated, input-output processes also give us stronger tools. What are these? Recall that $\mathcal{E}_{\overleftarrow{A}}$ depends only on the past inputs $\overleftarrow{A}$ due to causality. This means that, if we are able to bound it by some quantity, say $F_{\overleftrightarrow{A}}$ that depends on the whole input, then we can factor $\Pr(\overleftrightarrow{A}) = \Pr(\overset{\bullet\!\!\rightarrow}{A}|\overleftarrow{A})\Pr(\overleftarrow{A})$ and we have the much stronger bound

$$\mathcal{E}_{\overleftarrow{A}} \leq \inf_{\overset{\bullet\!\!\rightarrow}{A}|\overleftarrow{A}} F_{\overleftrightarrow{A}}. \tag{5.15}$$

In particular, if there exists some experiment $\overset{\bullet\!\!\rightarrow}{A}|\overleftarrow{A}$ such that $F_{\overleftrightarrow{A}} = 0$, then we can conclude $\mathcal{E}_{\overleftarrow{A}} = 0$. This kind of existence criterion is much easier to satisfy than the bare $\forall$ quantifier in our definition of $\psi$-epistemicity. This does not hold for every $\mathcal{E}_i$ that makes up $\mathcal{E}_{\overleftarrow{A}}$, but there is at least one decomposition of $\mathcal{E}_{\overleftarrow{a}}$ into quantities which only depend on

| | Stabilizer | Qubit | Pure state | Mixed state |
|---|---|---|---|---|
| Infinite input alphabet | × | ✓ | ✓ | ✓ |
| Continuous $\epsilon$-machine states | × | ✓ | ✓ | ✓ |
| Non-Markovian | ✓ | × | ✓ | ✓ |
| Repeatable measurements | ✓ | ✓ | ✓ | × |
| State-measurement complementarity | ✓ | ✓ | ✓ | × |
| Reversible transformations | ✓ | ✓ | ✓ | × |
| Tomography | ✓ | ✓ | ✓ | ✓ |

Table 5.1: Properties (rows) of various operational theories (columns). Stabilizer refers to the pure state stabilizer subtheory for power of odd prime dimensional systems (i.e. $n$-quo$p$its). Qubit refers to pure state qubit dynamics. Pure state refers to pure state quantum theory for $d \geq 3$, and Mixed state refers to mixed state quantum theory for $d \geq 3$. The properties are described in the text.

the past inputs:

$$\mathcal{E}_{\overleftarrow{A}} = (\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3) + (\mathcal{E}_4 + \mathcal{E}_5 + \mathcal{E}_6) + (\mathcal{E}_7 + \mathcal{E}_8 + \mathcal{E}_9) \tag{5.16}$$

$$H[\Psi|\,\Lambda] = H[\Psi|\,\overleftarrow{A}, \Lambda] + I[\Psi; \overleftarrow{K}; \overleftarrow{A}\,|\,\Lambda] + H[\Psi|\,\overleftarrow{K}, \Lambda] \tag{5.17}$$

Unfortunately the middle term can be negative as it is a three-way mutual information.

The tough part, of course, will be figuring out which sums of $\mathcal{E}_i$ to bound, and how to bound them. However, we are not flying completely blind. We have three examples of theories that are closely related to pure state quantum theory, but which each admit a $\epsilon$-epistemic interpretation: the stabilizer subtheory for $n$ quo$p$its, full quantum theory for a qubit ($d = 2$), and mixed state quantum theory (i.e. the operational theory that includes mixed states, CPTP maps, and non-projective measurements). Note that $\psi$ is still the $\epsilon$-machine for the stabilizer subtheory, by essentially the same proof that it is such for pure state quantum theory. Our goal, then, is to find some set of properties of pure state quantum mechanics that are not all simultaneously satisfied by the other three theories. One strange thing to note is that both stabilizers and qubits are subtheories of pure state quantum theory, while pure state quantum theory is a subtheory of mixed state quantum theory. This means that, if it is true that any model of pure state quantum theory must be $\epsilon$-ontic, then $\epsilon$-epistemicity is not a property that extends from theory to subtheory or vice versa.

We have listed some properties that seem like they may be relevant in Table 5.1. The first two rows single out the stabilizer subtheory. They are expressible as $\exists A : H[A] = \infty$, and $\exists \overleftarrow{A} : H[\Psi] = \infty$, respectively. It's unclear how to apply these to get bounds on anything. The third row singles out the single-qubit case; all single qubit (pure-state) dynamics are Markovian, since the only nontrivial measurements have rank 1. The condition of non-Markovianity can be expressed as $I[\overset{\circ\!\!\rightarrow}{K}; \overleftarrow{K}, \overleftarrow{A} | K, \overset{\bullet\!\!\rightarrow}{A}] > 0$, i.e. the present symbol/input pair does not form an HMM of the process. It is especially unclear how this might be applied, since we're really looking for quantities that are restricted to be 0. However, we note that it has been shown that $\epsilon$-epistemic models are generic for binary Markov processes [73], so this is one promising route (unfortunately the converse, which is what we're really after, has not been shown).

The next three rows single out mixed-state quantum mechanics. By repeatable measurements, we mean that *every* measurement in the theory can be repeated; in informational terms, $\forall a : H[K|K', A = a, A' = a] = 0$ (this notation is a little weird, but it still expresses a meaningful quantity). By state-measurement complementarity, we mean that for every $\epsilon$-machine state, there exists a measurement that will respond deterministically to that state, but not to any other: $\forall \psi : \exists a : H[k|\Psi = \psi, A = a] = 0 \wedge H[k|\Psi \neq \psi, A = a] > 0$. Again, the notation is weird and the application unclear. Reversible transformations also probably has an information-theoretic expression, but I think the point has been made.

The final row of the table, tomography, does not distinguish between these theories but it does give us a quantity exactly of the form in Eq. (5.15). In particular, in any theory that allows for tomography, we have

$$\forall \overleftarrow{A} : \exists (\overset{\bullet\!\!\rightarrow}{A} | \overleftarrow{A}, \overleftarrow{K}) : H[\Psi | \overset{\bullet\!\!\rightarrow}{A}, \overset{\bullet\!\!\rightarrow}{K}] = 0. \tag{5.18}$$

The one caveat is that we have allowed the future choice of experiment to also depend on the past outcomes; although this creates some issues with our formal construction of HMMs in Section 2.6, it's probably fine. We can demonstrate the above property by the following (very informal) construction: the future experiment $\overset{\bullet\!\!\rightarrow}{A}$ has a block structure, where each block consists of first repeating $\overleftarrow{A}$. If the sequence of outcomes is not the same as $\overleftarrow{k}$, then begin the next block. For each sequence that did reproduce the same outcomes, we are guaranteed to have the same $\psi$ at the end, conditioned on the actual values of $\overleftarrow{a}, \overleftarrow{k}$, as we had after the actual initial experiment. Thus, we have a way of repeatedly preparing $\psi$, and can perform a tomography experiment to determine $\psi$ for each possible $\overleftarrow{a}, \overleftarrow{k}$, and thus reconstructing $\Psi$ (I did say there would be hand-waving).

Presuming that this method actually works, and we don't have to worry about the fact that the entropy in Eq. 5.18 is infinite for any finite future experiment and only becomes 0 in the infinite-time limit, and the infinite-time limit is well defined in general, then we still have to actually use this result to bound some of the $\mathcal{E}_i$. All three of $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3$ are contained within the region $H[\Psi| \overset{\bullet\rightarrow}{A}, \overset{\bullet\rightarrow}{K}]$, but unfortunately they appear to depend on the future input so the trick from Eq. (5.15) won't work. Additionally, the rest of the atoms in that region are mostly negative, so it's not necessarily clear that $\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3$ is actually upper bounded by the value of $H[\Psi| \overset{\bullet\rightarrow}{A}, \overset{\bullet\rightarrow}{K}]$. Additionally, it seems likely from the computational mechanics literature that tomography is a generic property of $\epsilon$-machines, as it seems related to the notion of synchronization that is central to their development [22].

There is a special case that makes everything easier to analyze, which is models that are PNCfPS. In that case, almost all of $\Lambda$ becomes contained in $\Psi$, which means that more regions are positive and bounds become easier to come by. In particular, the bound in the previous paragraph holds in this case. As such, it may be that PNCfPS $\implies$ $\psi$-ontic, which would be a very interesting result and would be consistent with the set of examples that we ruled out, which are all PNCfPS.

Whether or not this technique actually pans out to prove a $\psi$-onticity theorem is anyone's guess. While the initial assessments of this chapter are optimistic, we've certainly ended up finding that the situation is daunting at best. Regardless, pursuing this line of reasoning should almost certainly provide some new insights both into the structure of ontological models and into quantum theory's relationship to other probabilistic theories. For example, if we prove that some $\mathcal{E}_i$ are 0 but not others, then we have at least gained information about 'where the epistemicity can come from.'

# Chapter 6

# Conclusion

> *There is a theory which states that if ever anyone discovers exactly what the Universe is for and why it is here, it will instantly disappear and be replaced by something even more bizarre and inexplicable. There is another theory which states that this has already happened.*
>
> – Douglas Adams, *The Restaurant at the End of the Universe*

Throughout this thesis we've been promising a couple of discussions in the conclusion, which we can now address.

First of all, we have repeatedly referred to our results as utilizing no additional assumptions, when we have clearly added a consideration of state update. While state update is certainly an extra piece of mathematical structure, it enters the picture in a significantly different way than the additional assumptions of existing no-go theorems [1, 19, 43, 61, 68]. In particular, it is introduced at the operational level—we say that we want to model sequential measurements, and the requirement of state update follows directly[1]. This introduction at the operational level is in contrast to the additional assumptions in the literature, which are all introduced at the ontological level and directly impose restrictions on the structure of the ontological model. In other words, state update is a thing that we want to model, whereas the 'additional assumptions' are about *how* we model it.

[1] One could perhaps argue that sequential measurements should be described by combining all of the measurements into a single joint distribution and describing that as a single measurement step; however, this is not a description of how things actually happen in experiments and is in contradiction with our HMM analogy. It also suggests a rather unattractive 'measurement at the end of the universe' picture, where we somehow need to wait until every measurement has been done before asserting that anything at all has happened.

Second, we may provide some (perhaps premature) speculation about where to head with state-realist $\psi$-epistemic interpretations if we must leave the ontological models formalism behind. As described in the introduction, this seems likely, independent of the actual existence or not of a $\psi$-epistemic ontological model. This is because all of the restrictions posed by previous no-go theorems still hold, and, perhaps most importantly, Bell's theorem tells us that any such model will be nonlocal anyway. To further this discussion, we recall the set of assumptions that characterize the ontological models formalism set out in Section 2.6:

1. We can represent nature via stochastic channels
2. This process is causal (time-ordered)
3. This process is stationary (time-translation invariant)
4. A system has a state
5. The states of a system render the past and future observations of the system conditionally independent of one another.

The first item on this list isn't so much an assumption as a statement of fact; even Generalized Probabilistic Theories—which postulate non-classical probability theories for the dynamics of states—reproduce the operational classical probabilities of quantum theory.

The fourth item on this list is the point of departure for state-antirealist $\psi$-epistemic interpretations and its negation has already been explored in many examples of such interpretations—while it is clearly an attractive option due to its many adherents, the benefits of a state-realist version are perhaps worth enough that we should not be so hasty to abandon the search for one.

If we want to explore state-realist $\psi$-epistemic interpretations outside of the ontological models formalism, this leaves the possibility of negating items 2,3, and 5. In my few discussions with senior members of the quantum foundations community, number 5 seems to be the obvious answer to many people. Dropping it would essentially amount to adding more causal arrows to the influence diagram in Fig. 2.5b which can move past a state without being mediated by it. While this is probably the least radical option, it has its own issues. First of all, this conditional independence is the *definition* of a state for a hidden Markov model and it is very unclear whether an object that does not have this property can be interpreted as any reasonable kind of state. Second, the nice mathematical structure that accompanies hidden Markov becomes not so nice with the inclusion of these extra causal influences. One major reason to use hidden Markov models (and, in fact, the justification for their name) is that they model an arbitrarily non-Markovian statistical process by inducing a Markov process over the "hidden" states (this Markovianity follows from assumption 5). Finally, it is hard (for me) to see a physical basis for this type of interpretation. What would it mean for a measurement taken today to depend on a state that the system held last week?

At this point we're down to numbers two and three, which, you might guess, is where I want to start trying to poke holes in the ontological models formalism. First, the downsides to picking these two: they are very basic assumptions about how we construct theories and how we understand time via special relativity. Indeed, a direct negation of either of these two leads to many of the same issues as the negation of number 5. But I'm suspecting that, if this is the joint we need to cut, then we may need something more radical than simple negation.

In particular, I will just point out that time, both as a general concept and as an element of physical theories, is not well understood [69]. Theorizing about time has been a thread of thought dating back to Plato, and we're still struggling to make sense of it, whether it's the subjective experience, the cognitive structure, the thing that is correlated with increases in (statistical mechanics) entropy, a reversible parameter, or an element of a dynamic spacetime distinguished from our familiar and friendly spacial dimensions by a single minus sign in the metric.

Given this, and the many other ways in which a number of quantum foundations research programs point towards time and causality as two of the central aspects of the strangeness of quantum theory, it seems natural to conclude that these are some of the big problems that we need to solve in order to make sense of quantum theory. Moreover, they provide a direct link to start trying to think about quantum foundations as existing in spacetime, rather than in quantum circuits and agents' minds as it has tended to for the last twenty years or so. In the spirit of foundations-as-path-towards-quantum-gravity articulated in the introduction, this seems like a very appealing way to go.

Speculation aside, we have shown that consideration of sequential measurements, and the state update rule needed to represent them, places nontrivial restrictions on the structure of $\psi$-epistemic ontological models. These restrictions were used to demonstrate that none of the known examples of $\psi$-epistemic ontological models, defined for a single-measurement experiment, can be supplemented with a state update rule. As such, the door is wide open for a no-go theorem ruling out $\psi$-epistemic ontological models without any additional assumptions. This flips the script on the usual view that $\psi$-epistemic interpretations have an advantage over $\psi$-ontic interpretations in describing this behavior.

The importance of sequential measurements suggested an analogy between ontological models and hidden Markov models, which further invited an application of computational mechanics. These connections provide a new look at the structure of ontological models, and promise to inform either a no-go theorem or the construction of a $\psi$-epistemic model.

# References


[1] S. Aaronson, A. Bouland, L. Chua, and G. Lowther. $\psi$-epistemic theories: The role of symmetry. *Physical Review A*, 88(3), Sept. 2013. doi:10.1103/PhysRevA.88.032111.

[2] C. Aghamohammadi, S. P. Loomis, J. R. Mahoney, and J. P. Crutchfield. Extreme Quantum Advantage for Rare-Event Sampling. *Physical Review X*, 8(1):011025, Feb. 2018, arXiv:1707.09553. doi:10.1103/PhysRevX.8.011025.

[3] C. Aghamohammadi, J. R. Mahoney, and J. P. Crutchfield. The Ambiguity of Simplicity. Feb. 2016, arXiv:1602.08646.

[4] C. Aghamohammadi, J. R. Mahoney, and J. P. Crutchfield. Extreme Quantum Advantage when Simulating Classical Systems with Long-Range Interaction. *Scientific Reports*, 7(1):6735, July 2017. doi:10.1038/s41598-017-04928-7.

[5] L. E. Ballentine. The Statistical Interpretation of Quantum Mechanics. *Reviews of Modern Physics*, 42(4):358–381, Oct. 1970. doi:10.1103/RevModPhys.42.358.

[6] L. E. Ballentine, Y. Yang, and J. P. Zibin. Inadequacy of Ehrenfest's theorem to characterize the classical regime. *Physical Review A*, 50(4):2854–2859, Oct. 1994. doi:10.1103/PhysRevA.50.2854.

[7] N. Barnett and J. P. Crutchfield. Computational Mechanics of Input-Output Processes: Structured transformations and the $\epsilon$-transducer. *Journal of Statistical Physics*, 161(2):404–451, Oct. 2015, arXiv:1412.2690. doi:10.1007/s10955-015-1327-5.

[8] H. Barnum and A. Wilce. Post-Classical Probability Theory. May 2012, arXiv:1205.3833.

[9] J. Barrett, E. G. Cavalcanti, R. Lal, and O. J. E. Maroney. No $\psi$-Epistemic Model Can Fully Explain the Indistinguishability of Quantum States. *Physical Review Letters*, 112(25), June 2014. doi:10.1103/PhysRevLett.112.250403.

[10] A. Bassi, K. Lochan, S. Satin, T. P. Singh, and H. Ulbricht. Models of wave-function collapse, underlying theories, and experimental tests. *Reviews of Modern Physics*, 85(2):471–527, Apr. 2013. doi:10.1103/RevModPhys.85.471.

[11] J. S. Bell. On the Problem of Hidden Variables in Quantum Mechanics. *Reviews of Modern Physics*, 38(3):447–452, July 1966. doi:10.1103/RevModPhys.38.447.

[12] E. G. Beltrametti and S. Bugajski. A classical extension of quantum mechanics. *Journal of Physics A: Mathematical and General*, 28(12):3329, 1995. doi:10.1088/0305-4470/28/12/007.

[13] D. Bohm. A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. I. *Physical Review*, 85(2):166–179, Jan. 1952. doi:10.1103/PhysRev.85.166.

[14] J. Bub. Why Bohr was (Mostly) Right. Nov. 2017, arXiv:1711.01604.

[15] A. Cabello. Interpretations of quantum theory: A map of madness. pages 138–144, 2017, arXiv:1509.04711. doi:10.1017/9781316494233.009.

[16] A. Cabello, M. Gu, O. Gühne, J.-Å. Larsson, and K. Wiesner. Thermodynamical cost of some interpretations of quantum theory. *Physical Review A*, 94(5), Nov. 2016. doi:10.1103/PhysRevA.94.052127.

[17] L. Catani and D. E. Browne. Spekkens' toy model in all dimensions and its relationship with stabilizer quantum mechanics. *New Journal of Physics*, 19(7):073035, July 2017, arXiv:1701.07801. doi:10.1088/1367-2630/aa781c.

[18] C. M. Caves, C. A. Fuchs, and R. Schack. Conditions for compatibility of quantum-state assignments. *Physical Review A*, 66(6):062111, Dec. 2002. doi:10.1103/PhysRevA.66.062111.

[19] R. Colbeck and R. Renner. A system's wave function is uniquely determined by its underlying physical state. *New Journal of Physics*, 19(1):013016, Jan. 2017. doi:10.1088/1367-2630/aa515c.

[20] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley-Interscience, Hoboken, N.J, 2nd ed edition, 2006.

[21] J. P. Crutchfield. The Origins of Computational Mechanics: A Brief Intellectual History and Several Clarifications. Oct. 2017, arXiv:1710.06832.

[22] J. P. Crutchfield, C. J. Ellison, R. G. James, and J. R. Mahoney. Synchronization and control in intrinsic and designed computation: An information-theoretic analysis of competing models of stochastic computation. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 20(3):037105, Sept. 2010. doi:10.1063/1.3489888.

[23] D. Deutsch. *The Fabric of Reality: The Science of Parallel Universes– and Its Implications*. Allen Lane, New York, 1997.

[24] B. S. Dewitt and N. Graham. *The Many Worlds Interpretation of Quantum Mechanics*. Princeton Legacy Library, 2016.

[25] D. Dieks and P. E. Vermaas. *The Modal Interpretation of Quantum Mechanics*. Springer Netherlands, Dordrecht, 1998.

[26] A. Einstein, B. Podolsky, and N. Rosen. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Physical Review*, 47(10):777–780, May 1935. doi:10.1103/PhysRev.47.777.

[27] J. Emerson. Quantum Chaos and Quantum-Classical Correspondence. Nov. 2002, arXiv:quant-ph/0211035.

[28] J. Emerson, D. Serbin, C. Sutherland, and V. Veitch. The whole is greater than the sum of the parts: On the possibility of purely statistical interpretations of quantum theory. Dec. 2013, arXiv:1312.1345.

[29] H. Everett. "Relative State" Formulation of Quantum Mechanics. *Reviews of Modern Physics*, 29(3):454–462, July 1957. doi:10.1103/RevModPhys.29.454.

[30] C. Ferrie. Quasi-probability representations of quantum theory with applications to quantum information science. *Reports on Progress in Physics*, 74(11):116001, Nov. 2011, arXiv:1010.2701. doi:10.1088/0034-4885/74/11/116001.

[31] C. Ferrie and J. Emerson. Frame representations of quantum mechanics and the necessity of negativity in quasi-probability representations. *Journal of Physics A: Mathematical and Theoretical*, 41(35):352001, Sept. 2008, arXiv:0711.2658. doi:10.1088/1751-8113/41/35/352001.

[32] C. Ferrie and J. Emerson. Framed Hilbert space: Hanging the quasi-probability pictures of quantum theory. *New Journal of Physics*, 11(6):063040, June 2009, arXiv:0903.4843. doi:10.1088/1367-2630/11/6/063040.

[33] P. Feyerabend. *Against Method*. Verso, London ; New York, 4th ed edition, 2010.

[34] C. A. Fuchs. Quantum Mechanics as Quantum Information (and only a little more). May 2002, arXiv:quant-ph/0205039.

[35] C. A. Fuchs. On Participatory Realism. Jan. 2016, arXiv:1601.04360.

[36] C. A. Fuchs, N. D. Mermin, and R. Schack. An introduction to QBism with an application to the locality of quantum mechanics. *American Journal of Physics*, 82(8):749–754, Aug. 2014. doi:10.1119/1.4874855.

[37] C. A. Fuchs and A. Peres. Quantum Theory Needs No 'Interpretation'. *Physics Today*, 53(3):70, Jan. 2007. doi:10.1063/1.883004.

[38] A. Gleason. Measures on the Closed Subspaces of a Hilbert Space. *Indiana University Mathematics Journal*, 6(4):885–893, 1957. doi:10.1512/iumj.1957.6.56050.

[39] D. Gottesman. An Introduction to Quantum Error Correction and Fault-Tolerant Quantum Computation. Apr. 2009, arXiv:0904.2557.

[40] D. Gross. Hudson's Theorem for finite-dimensional quantum systems. *Journal of Mathematical Physics*, 47(12):122107, Dec. 2006, arXiv:quant-ph/0602001. doi:10.1063/1.2393152.

[41] D. Gross. Non-negative Wigner functions in prime dimensions. *Applied Physics B*, 86(3):367–370, Feb. 2007, arXiv:quant-ph/0702004. doi:10.1007/s00340-006-2510-9.

[42] M. Gu, K. Wiesner, E. Rieper, and V. Vedral. Quantum mechanics can reduce the complexity of classical models. *Nature Communications*, 3:762, Mar. 2012. doi:10.1038/ncomms1761.

[43] L. Hardy. Are quantum states real? *International Journal of Modern Physics B*, 27(01n03):1345012, Jan. 2013. doi:10.1142/S0217979213450124.

[44] L. Hardy. The Construction Interpretation: A Conceptual Road to Quantum Gravity. July 2018, arXiv:1807.10980.

[45] N. Harrigan and T. Rudolph. Ontological models and the interpretation of contextuality. Sept. 2007, arXiv:0709.4266.

[46] N. Harrigan, T. Rudolph, and S. Aaronson. Representing probabilistic data via ontological models. Sept. 2007, arXiv:0709.1149.

[47] N. Harrigan and R. W. Spekkens. Einstein, incompleteness, and the epistemic view of quantum states. *Foundations of Physics*, 40(2):125–157, Feb. 2010, arXiv:0706.2661. doi:10.1007/s10701-009-9347-0.

[48] R. Healey. Quantum Theory: A Pragmatist Approach. Aug. 2010, arXiv:1008.3896.

[49] G. Hermann. Die naturphilosophischen Grundlagen der Quantenmechanik. *Naturwissenschaften*, 23(42):718–721, Oct. 1935. doi:10.1007/BF01491142.

[50] D. Howard. Einstein on locality and separability. *Studies in History and Philosophy of Science Part A*, 16(3):171–201, Sept. 1985. doi:10.1016/0039-3681(85)90001-9.

[51] R. A. Howard and J. E. Matheson. Influence Diagrams. *Decision Analysis*, 2(3):127–143, Sept. 2005. doi:10.1287/deca.1050.0020.

[52] A. Karanjai, J. J. Wallman, and S. D. Bartlett. Contextuality bounds the efficiency of classical simulation of quantum processes. Feb. 2018, arXiv:1802.07744.

[53] S. Kochen and E. Specker. The Problem of Hidden Variables in Quantum Mechanics. *Indiana University Mathematics Journal*, 17(1):59–87, 1967. doi:10.1512/iumj.1968.17.17004.

[54] T. S. Kuhn and I. Hacking. *The Structure of Scientific Revolutions*. The University of Chicago Press, Chicago ; London, fourth edition edition, 2012.

[55] M. S. Leifer. Is the quantum state real? An extended review of $\psi$-ontology theorems. *Quanta*, 3(1):67, Nov. 2014, arXiv:1409.1570. doi:10.12743/quanta.v3i1.22.

[56] M. S. Leifer and M. F. Pusey. Is a time symmetric interpretation of quantum theory possible without retrocausality? *Proc. R. Soc. A*, 473(2202):20160607, June 2017. doi:10.1098/rspa.2016.0607.

[57] P. G. Lewis, D. Jennings, J. Barrett, and T. Rudolph. Distinct Quantum States Can Be Compatible with a Single State of Reality. *Physical Review Letters*, 109(15), Oct. 2012. doi:10.1103/PhysRevLett.109.150404.

[58] P. Lillystone and J. Emerson. A Contextual $\psi$-Epistemic Model of the n-Qubit Stabilizer Formalism. Apr. 2019, arXiv:1904.04268.

[59] O. Lombardi and D. Dieks. Modal Interpretations of Quantum Mechanics. *The Stanford Encyclopedia of Philosophy*.

[60] J. R. Mahoney, C. Aghamohammadi, and J. P. Crutchfield. Occam's Quantum Strop: Synchronizing and Compressing Classical Cryptic Processes via a Quantum Channel. *Scientific Reports*, 6:20495, Feb. 2016. doi:10.1038/srep20495.

[61] S. Mansfield. Reality of the quantum state: Towards a stronger $\psi$ -ontology theorem. *Physical Review A*, 94(4), Oct. 2016. doi:10.1103/PhysRevA.94.042124.

[62] A. Montina. Approximate simulation of entanglement with a linear cost of communication. *Physical Review A*, 84(4):042307, Oct. 2011. doi:10.1103/PhysRevA.84.042307.

[63] A. Montina. Comment on math overflow question: 'psi-epistemic theories' in 3 or more dimensions, Dec. 2012.

[64] A. Montina. Dynamics of a qubit as a classical stochastic process with time-correlated noise: Minimal measurement invasiveness. *Physical Review Letters*, 108(16), Apr. 2012, arXiv:1108.5138. doi:10.1103/PhysRevLett.108.160501.

[65] A. Montina. Epistemic view of quantum states and communication complexity of quantum channels. *Physical Review Letters*, 109(11), Sept. 2012, arXiv:1206.2961. doi:10.1103/PhysRevLett.109.110501.

[66] M. S. Palsson, M. Gu, J. Ho, H. M. Wiseman, and G. J. Pryde. Experimentally modeling stochastic processes with less memory by the use of a quantum processor. *Science Advances*, 3(2):e1601302, Feb. 2017. doi:10.1126/sciadv.1601302.

[67] P. Pearle. Collapse models. In H.-P. Breuer and F. Petruccione, editors, *Open Systems and Measurement in Relativistic Quantum Theory*, volume 526, pages 195–234. Springer Berlin Heidelberg, 1999. doi:10.1007/BFb0104404.

[68] M. F. Pusey, J. Barrett, and T. Rudolph. On the reality of the quantum state. *Nature Physics*, 8(6):476–479, May 2012, arXiv:1111.3328. doi:10.1038/nphys2309.

[69] H. Reichenbach. *The Direction of Time*. Number 65 in Campus. Univ. of California P, Berkeley, 1971.

[70] P. M. Riechers, J. R. Mahoney, C. Aghamohammadi, and J. P. Crutchfield. Minimized state complexity of quantum-encoded cryptic processes. *Physical Review A*, 93(5):052317, May 2016. doi:10.1103/PhysRevA.93.052317.

[71] C. Rovelli. Relational quantum mechanics. *International Journal of Theoretical Physics*, 35(8):1637–1678, Aug. 1996. doi:10.1007/BF02302261.

[72] T. Rudolph. Ontological Models for Quantum Mechanics and the Kochen-Specker theorem. Aug. 2006, arXiv:quant-ph/0608120.

[73] J. B. Ruebeck, R. G. James, J. R. Mahoney, and J. P. Crutchfield. Prediction and generation of binary Markov processes: Can a finite-state fox catch a Markov mouse? *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 28(1):013109, Jan. 2018. doi:10.1063/1.5003041.

[74] J. B. Ruebeck, P. Lillystone, and J. Emerson. $\psi$-epistemic interpretations of quantum theory have a measurement problem. Dec. 2018, arXiv:1812.08218.

[75] J. J. Sakurai and J. Napolitano. *Modern Quantum Mechanics*. Addison-Wesley, Boston, 2nd edition, 2011.

[76] R. Schack. Participatory realism, June 2016.

[77] M. Schlosshauer, J. Kofler, and A. Zeilinger. A Snapshot of Foundational Attitudes Toward Quantum Mechanics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 44(3):222–230, Aug. 2013, arXiv:1301.1069. doi:10.1016/j.shpsb.2013.04.004.

[78] C. R. Shalizi and J. P. Crutchfield. Computational Mechanics: Pattern and Prediction, Structure and Simplicity. July 1999, arXiv:cond-mat/9907176.

[79] D. Spehner. Quantum correlations and distinguishability of quantum states. *Journal of Mathematical Physics*, 55(7):075211, July 2014, arXiv:1407.3739. doi:10.1063/1.4885832.

[80] R. W. Spekkens. Contextuality for preparations, transformations, and unsharp measurements. *Physical Review A*, 71(5), May 2005, arXiv:quant-ph/0406166. doi:10.1103/PhysRevA.71.052108.

[81] R. W. Spekkens. In defense of the epistemic view of quantum states: A toy theory. *Physical Review A*, 75(3), Mar. 2007, arXiv:quant-ph/0401052. doi:10.1103/PhysRevA.75.032110.

[82] J. Thompson, A. J. P. Garner, V. Vedral, and M. Gu. Using quantum theory to simplify input–output processes. *npj Quantum Information*, 3(1):6, Feb. 2017. doi:10.1038/s41534-016-0001-3.

[83] C. G. Timpson. Quantum Bayesianism: A Study. Apr. 2008, arXiv:0804.2047.

[84] N. F. Travers and J. P. Crutchfield. Equivalence of History and Generator Epsilon-Machines. Nov. 2011, arXiv:1111.4500.

[85] A. Valentini. *On the Pilot-Wave Theory of Classical, Quantum, and Subquantum Physics*. Ph.D. Thesis, SISSA, Trieste, 1991.

[86] B. C. Van Fraassen. *Quantum Mechanics: An Empiricist View*. Clarendon Press ; Oxford University Press, Oxford [England] : New York, 1991.

[87] V. Veitch. *Negative Quasi-Probability in the Context of Quantum Computation*. PhD thesis, University of Waterloo, 2013.

[88] V. Veitch, S. A. H. Mousavian, D. Gottesman, and J. Emerson. The Resource Theory of Stabilizer Computation. *New Journal of Physics*, 16(1):013009, Jan. 2014, arXiv:1307.7171. doi:10.1088/1367-2630/16/1/013009.

[89] J. Von Neumann. *Mathematical Foundations of Quantum Mechanics*. Princeton Landmarks in Mathematics and Physics. Princeton University Press, Princeton Chichester, 1996.

[90] Werner Heisenberg. *Physics & Philosophy*. Unwin University Books, 1958.

[91] W. K. Wootters. A Wigner-function formulation of finite-state quantum mechanics. *Annals of Physics*, 176(1):1–21, May 1987. doi:10.1016/0003-4916(87)90176-X.

[92] R. W. Yeung. A new outlook of Shannon's information measures. *IEEE Trans. Information Theory*, 37:466–474, 1991. doi:10.1109/18.79902.

# APPENDICES

# Appendix A

# Proofs that models reproduce quantum theory

## A.1 Kochen-Specker model

*Proof: The KS model reproduces quantum statistics.* We first evaluate

$$\begin{aligned}\Pr(\phi|\psi) &= \int_\Lambda d\lambda \xi(\phi|\lambda)\mu(\lambda|\phi) \\ &= \int_\Lambda d\lambda \frac{1}{\pi}\Theta(\vec{\psi}\cdot\vec{\lambda})\,\vec{\psi}\cdot\vec{\lambda}\,\Theta(\vec{\phi}\cdot\vec{\lambda}).\end{aligned} \tag{A.1}$$

It is convenient to pick a basis for $\mathbb{R}^3$ such that $\vec{\psi} = \hat{x}$ and $\vec{\phi}$ lies in the half-plane $z = 0, y > 0$ an angle $\gamma$ away from the $x$-axis (and thus from $\vec{\psi}$). We parametrize $\vec{\lambda}$ by the usual polar angle $\alpha$ and azimuthal angle $\beta$. Summarizing,

$$\begin{aligned}\vec{\psi} &= \hat{x} \\ \vec{\phi} &= \cos(\gamma)\,\hat{x} + \sin(\gamma)\,\hat{y} \\ \vec{\lambda} &= \sin(\alpha)\cos(\beta)\,\hat{x} + \sin(\alpha)\sin(\beta)\,\hat{y} + \cos(\alpha)\,\hat{z}\end{aligned}$$

With the usual differential element $d\lambda = d\alpha\, d\beta\, \sin\alpha$, we can then calculate dot products and substitute into Eq. A.1:

$$\Pr(\phi|\psi) = \frac{1}{\pi}\int_0^\pi d\alpha \int_{-\pi}^{\pi} d\beta\, \sin\alpha\, \Theta[\sin\alpha\cos\beta]\, \Theta[\sin\alpha(\cos\beta\cos\gamma + \sin\beta\sin\gamma)]\sin\alpha\cos\beta$$

Since $\alpha$ goes from 0 to $\pi$, $\sin\alpha > 0$ so it doesn't affect the Heaviside function in either case. Thus the first gives us b$\Theta(\cos\beta)$, which restricts $\beta$ to range from $-\pi/2$ to $\pi/2$. The second can be written $\Theta(\cos(\beta-\gamma))$; since we chose $\vec{\phi}$ to lie in the $+y$ half-plane, $\gamma \in [0,\pi]$ and so $\beta$ is further restricted to be less than $\pi/2-\gamma$. Absorbing the Heaviside functions into our integration limits then gives

$$\begin{aligned}
\Pr(\phi|\psi) &= \frac{1}{\pi}\int_{-\pi/2}^{\pi/2-\gamma} d\beta\,\cos\beta \int_0^{\pi} d\alpha\,\sin^2\alpha \\
&= \frac{1}{2}\int_{-\pi/2}^{\pi/2-\gamma} d\beta\,\cos\beta \\
&= \frac{1}{2}\left(\sin(\pi/2-\gamma) - \sin(-\pi/2)\right) \\
&= \frac{1}{2}\left(\cos\gamma + 1\right) \\
&= \frac{1}{2}\left(\vec{\psi}\cdot\vec{\phi} + 1\right) \\
&= |\langle\phi|\psi\rangle|^2
\end{aligned}$$

□

## A.2 Montina's model

Montina proved in [64] that his model reproduces quantum mechanics. However, we ought to check that our explicit ontological models presentation of his model still does. Since the transformations were built into everything, let's just evaluate the prepare-measure case:

$$\begin{aligned}
\Pr(k|M_\phi, P_\psi) &= \int_\Lambda d\lambda\,\xi(k|\lambda, M_\phi)\mu(\lambda|P_\psi) \\
&= \sum_{r,s}\int \mathrm{d}\vec{x}_{+1}\,\mathrm{d}\vec{x}_2\,\Theta\Big[ks(\vec{x}_r\cdot\vec{n})(\vec{x}_r\cdot\vec{\phi})\Big] \\
&\quad \cdot\frac{1}{4\pi^2}\Theta\Big[r\mathrm{sgn}[(\vec{x}_{+1}\cdot\vec{\psi})^2 - (\vec{x}_{-1}\cdot\vec{\psi})^2]\Big]\Theta\Big[s(\vec{x}_r\cdot\vec{\psi})(\vec{x}_r\cdot\vec{n})\Big]
\end{aligned}$$

After substituting above, we first evaluate the sum over $s$ (examining only the terms that depend on $s$):

$$\sum_s \Theta\Big[s(\vec{x}_r\cdot\vec{\psi})(\vec{x}_r\cdot\vec{n})\Big]\Theta\Big[ks(\vec{x}_r\cdot\vec{n})(\vec{x}_r\cdot\vec{\phi})\Big] = \Theta\Big[k(\vec{x}_r\cdot\vec{\psi})(\vec{x}_r\cdot\vec{\phi})\Big]$$

This can be seen from the fact that, of the two terms in the sum, the first theta function will be 1 for exactly one of them. This is the term where $\mathrm{sgn}(\vec{x}_r \cdot \vec{n}) = \mathrm{sgn}(s\vec{x}_r \cdot \vec{\psi})$. So, we can substitute this expression into the second theta function and eliminate the sum. Then we have an $s^2$ term, which is equal to 1 (since $s = \pm 1$), and we get the result on the right hand side. Substituting this back into the original integral, we get

$$\Pr(k|M_\phi, P_\psi) = \sum_r \int \mathrm{d}\vec{x}_{+1}\,\mathrm{d}\vec{x}_2\,\frac{1}{4\pi^2}\Theta\Big[k(\vec{x}_r \cdot \vec{\psi})(\vec{x}_r \cdot \vec{\phi})\Big]\Theta\Big[r\mathrm{sgn}[(\vec{x}_{+1} \cdot \vec{\psi})^2 - (\vec{x}_{-1} \cdot \vec{\psi})^2]\Big]. \tag{A.2}$$

This is the integral that Montina shows is consistent with quantum theory, so we're done.

## A.3 Kitchen Sink

### Normalization

$\xi$ is clearly normalized. $\Gamma$ is normalized due to the assumption of a closed subtheory; $U$ acts as a permutation on the $M^j$s so the product of delta functions is satisfied exactly once.

For $\mu$,

$$\sum_\lambda \mu(\lambda|\rho) = \sum_{\lambda \in \mathbb{Z}_r^m} \prod_{i=1}^m \mathrm{tr}\big(E^i_{\lambda_i}\rho\big) \tag{A.3}$$

$$= \prod_{i=1}^m \sum_{l \in \mathbb{Z}_r} \mathrm{tr}\big(E^i_l \rho\big) \tag{A.4}$$

$$= 1 \tag{A.5}$$

Where in the last step we've used the normalization condition on the POVM. Since $\eta$ amounts to re-preparing in the measured (rank-1) state, it is normalized because $\mu$ is.

### Prepare-measure

We proceed similarly to how we showed the normalization of $\mu$:

$$
\begin{aligned}
\sum_{\lambda} \xi(k|\lambda, M^i)\mu(\lambda|\rho) &= \sum_{\lambda} \delta(k, \lambda(M^i)) \prod_{j=1}^{m} \operatorname{tr}\Big(E^j_{\lambda(M^j)}\rho\Big) \\
&= \sum_{\lambda} \delta(k, \lambda(M^i)) \operatorname{tr}\Big(E^1_{\lambda(M^1)}\rho\Big) \operatorname{tr}\Big(E^2_{\lambda(M^2)}\rho\Big) \cdots \operatorname{tr}\big(E^m_{\lambda(M^m)}\rho\big) \\
&= \sum_{\lambda_1 \in \mathbb{Z}_r} \sum_{\lambda_2 \in \mathbb{Z}_r} \cdots \sum_{\lambda_m \in \mathbb{Z}_r} \delta(k, \lambda_i) \operatorname{tr}\big(E^1_{\lambda_1}\rho\big) \operatorname{tr}\big(E^2_{\lambda_2}\rho\big) \cdots \operatorname{tr}\big(E^m_{\lambda_m}\rho\big) \\
&= \sum_{\lambda_1 \in \mathbb{Z}_r} \operatorname{tr}\big(E^1_{\lambda_1}\rho\big) \cdots \sum_{\lambda_i \in \mathbb{Z}_r} \delta(k, \lambda_i) \operatorname{tr}\big(E^i_{\lambda_i}\rho\big) \cdots \sum_{\lambda_m \in \mathbb{Z}_r} \operatorname{tr}\big(E^m_{\lambda_m}\rho\big) \\
&= \sum_{\lambda_i \in \mathbb{Z}_r} \delta(k, \lambda_i) \operatorname{tr}\big(E^i_{\lambda_i}\rho\big) \\
&= \operatorname{tr}\big(E^i_k\rho\big)
\end{aligned}
$$

which is the desired quantum probability.

**Transformations (unitary)**

$$\begin{aligned}
\sum_{\lambda}\Gamma(\lambda'|\lambda,U)\mu(\lambda|\rho) &= \sum_{\lambda\in\mathbb{Z}_r^m}\left[\prod_{j=1}^{m}\delta(\lambda'(UM^jU^\dagger),\lambda(M^j))\right]\left[\prod_{i=1}^{m}\mathrm{tr}\Big(E^i_{\lambda(M^i)}\rho\Big)\right] \\
&= \sum_{\lambda\in\mathbb{Z}_r^m}\prod_{j=1}^{m}\delta(\lambda'(UM^jU^\dagger),\lambda(M^j))\,\mathrm{tr}\Big(E^j_{\lambda(M^j)}\rho\Big) \\
&= \prod_{j=1}^{m}\sum_{l\in\mathbb{Z}_r}\delta(\lambda'(UM^jU^\dagger),l)\,\mathrm{tr}\big(E^j_l\rho\big) \\
&= \prod_{j=1}^{m}\mathrm{tr}\Big(E^j_{\lambda'(UM^jU^\dagger)}\rho\Big) \\
&= \prod_{j=1}^{m}\mathrm{tr}\Big(U^\dagger E^j_{\lambda'(M^j)}U\rho\Big) \\
&= \prod_{j=1}^{m}\mathrm{tr}\Big(E^j_{\lambda'(M^j)}U\rho U^\dagger\Big) \\
&= \mu(\lambda'|U\rho U^\dagger)
\end{aligned}$$

We get from the fourth line to the fifth line by reindexing, since assuming a closed subtheory means that $U$ induces a permutation on the set of $M^j$s.

## A.4 The Wigner Function

When working with the Wigner function and stabilizer subtheory, we make free use of the following identities, all derivable by plugging in the relevant definitions:

$$\begin{aligned}
T_\lambda T_{\lambda'} &= \omega^{2^{-1}[\lambda,\lambda']}T_{\lambda+\lambda'} \\
\mathrm{tr}(T_\lambda) &= p^n\delta_{\lambda,\mathbf{0}} \\
\mathrm{tr}(T_\lambda T_{\lambda'}) &= p^n\delta_{\lambda,-\lambda'} \\
T_\lambda^\dagger &= T_\lambda^{-1} = T_{-\lambda} \\
(T_\lambda)^m &= T_{m\lambda} \\
A_\lambda A_{\lambda'} &= \omega^{2[\lambda',\lambda]}T_{2(\lambda-\lambda')} \\
\mathrm{tr}(A_\lambda T_{\lambda'}) &= \omega^{[\lambda',\lambda]}
\end{aligned}$$

**Derivation of OM presentation from QPR presentation**

First, we evaluate $\xi$ for a stabilizer measurement. We use $\sigma, \sigma' \ldots \in \mathbb{Z}_p^n \times \mathbb{Z}_p^n$ as dummy variables for summing over the ontic space and $l, m$ as dummy variables for summing over $\mathbb{Z}_p$.

$$\begin{aligned}
\xi(k|\lambda, M^{(\mathbf{x},\mathbf{z})}) &= \operatorname{tr}\Big(\Pi_k^{(\mathbf{x},\mathbf{z})} A_\lambda\Big) \\
&= \frac{1}{p^{n+1}} \sum_m \omega^{-km} \sum_\sigma \omega^{[\lambda,\sigma]} \operatorname{tr}\big(T^m_{(\mathbf{x},\mathbf{z})} T_\sigma\big) \\
&= \frac{1}{p^{n+1}} \sum_m \sum_\sigma \omega^{[\lambda,\sigma]-km+2^{-1}[m(\mathbf{x},\mathbf{z}),\sigma]} \operatorname{tr}\big(T_{m(\mathbf{x},\mathbf{z})+\sigma}\big) \\
&= \frac{1}{p} \sum_m \sum_\sigma \omega^{[\lambda,\sigma]-km+2^{-1}[m(\mathbf{x},\mathbf{z}),\sigma]} \delta_{m(\mathbf{x},\mathbf{z}),-\sigma} \\
&= \frac{1}{p} \sum_m \omega^{[\lambda,-m(\mathbf{x},\mathbf{z})]-km+2^{-1}[m(\mathbf{x},\mathbf{z}),-m(\mathbf{x},\mathbf{z})]} \\
&= \frac{1}{p} \sum_m \omega^{-m(k-[(\mathbf{x},\mathbf{z}),\lambda])} \\
&= \delta_{k,[(\mathbf{x},\mathbf{z}),\lambda]}
\end{aligned}$$

In the last step, we use the fact that the sum over powers of the $p$th root of unity $\omega$ on a vector space over $\mathbb{Z}_p$ is equal to the size of the vector space if the power is 0 for all elements of the vector space, and 0 otherwise.

We now carry out a similar, if nastier, calculation for $\eta$:

$$
\begin{aligned}
&\eta(\lambda'|k,\lambda,M^{(\mathbf{x},\mathbf{z})})\\
&= \frac{1}{p^n \operatorname{tr}\left(\Pi_k^{(\mathbf{x},\mathbf{z})} A_\lambda\right)} \operatorname{tr}\left(\Pi_k^{(\mathbf{x},\mathbf{z})} A_\lambda \Pi_k^{(\mathbf{x},\mathbf{z})} A_{\lambda'}\right)\\
&= \frac{1}{p^{3n+2}\delta_{k,[(\mathbf{x},\mathbf{z}),\lambda]}} \sum_{l,m}\sum_{\sigma,\sigma'} \omega^{[\lambda,\sigma]+[\lambda',\sigma']-km-kl} \operatorname{tr}\left(T_{l(\mathbf{x},\mathbf{z})} T_\sigma T_{m(\mathbf{x},\mathbf{z})} T_{\sigma'}\right)\\
&= \frac{1}{p^{3n+2}\delta_{k,[(\mathbf{x},\mathbf{z}),\lambda]}} \sum_{l,m}\sum_{\sigma,\sigma'} \omega^{[\lambda,\sigma]+[\lambda',\sigma']-km-kl+2^{-1}[l(\mathbf{x},\mathbf{z}),\sigma]+,^{-1}[m(\mathbf{x},\mathbf{z}),\sigma']} \operatorname{tr}\left(T_{l(\mathbf{x},\mathbf{z})+\sigma} T_{m(\mathbf{x},\mathbf{z})+\sigma'}\right)\\
&= \frac{1}{p^{3n+2}\delta_{k,[(\mathbf{x},\mathbf{z}),\lambda]}} \sum_{l,m}\sum_{\sigma,\sigma'} \omega^{[\lambda,\sigma]+[\lambda',\sigma']-km-kl+2^{-1}[l(\mathbf{x},\mathbf{z}),\sigma]+,^{-1}[m(\mathbf{x},\mathbf{z}),\sigma']} p^n \delta_{\sigma',-\sigma-l(\mathbf{x},\mathbf{z})-m(\mathbf{x},\mathbf{z})}\\
&= \frac{1}{p^{2n+2}\delta_{k,[(\mathbf{x},\mathbf{z}),\lambda]}} \sum_{l,m}\sum_{\sigma,\sigma'} \omega^{[\lambda,\sigma]-[\lambda',\sigma]-l[\lambda',(\mathbf{x},\mathbf{z})]-m[\lambda',(\mathbf{x},\mathbf{z})]-km-kl+l2^{-1}[(\mathbf{x},\mathbf{z}),\sigma]-m2^{-1}[(\mathbf{x},\mathbf{z}),\sigma]}\\
&\qquad\qquad\qquad\qquad \cdot \omega^{-lm2^{-1}[(\mathbf{x},\mathbf{z}),(\mathbf{x},\mathbf{z})]-m^2 2^{-1}[(\mathbf{x},\mathbf{z}),(\mathbf{x},\mathbf{z})]}\\
&= \frac{1}{p^{2n+2}\delta_{k,[(\mathbf{x},\mathbf{z}),\lambda]}} \sum_{\sigma} \omega^{[\lambda,\sigma]-[\lambda',\sigma]} \sum_{l} \omega^{l(2^{-1}[(\mathbf{x},\mathbf{z}),\sigma]-[\lambda',(\mathbf{x},\mathbf{z})]-k)} \sum_{m} \omega^{-m(2^{-1}[(\mathbf{x},\mathbf{z}),\sigma]+[\lambda',(\mathbf{x},\mathbf{z})]+k)}\\
&= \frac{1}{p^{2n+2}\delta_{k,[(\mathbf{x},\mathbf{z}),\lambda]}} \sum_{\sigma} \omega^{[\lambda-\lambda',\sigma]} p\delta_{k,[(\mathbf{x},\mathbf{z}),\lambda'+2^{-1}\sigma]} p\delta_{k,[(\mathbf{x},\mathbf{z}),\lambda'-2^{-1}\sigma]}\\
&= \frac{1}{p^{2n}\delta_{k,[(\mathbf{x},\mathbf{z}),\lambda]}} \sum_{\sigma} \omega^{[\lambda-\lambda',\sigma]} \delta_{k,[(\mathbf{x},\mathbf{z}),\lambda']} \delta_{0,[(\mathbf{x},\mathbf{z}),\sigma]}\\
&= \frac{1}{p^{2n}\delta_{k,[(\mathbf{x},\mathbf{z}),\lambda]}} \delta_{k,[(\mathbf{x},\mathbf{z}),\lambda']} \sum_{\sigma} \omega^{[\lambda-\lambda',\sigma]} \delta_{0,[(\mathbf{x},\mathbf{z}),\sigma]}\\
&= \frac{1}{p^{2n}\delta_{k,[(\mathbf{x},\mathbf{z}),\lambda]}} \delta_{k,[(\mathbf{x},\mathbf{z}),\lambda']} \sum_{c\in\mathbb{Z}_p} p^{2n-1} \delta_{\lambda-\lambda',c(\mathbf{x},\mathbf{z})}\\
&= \frac{1}{p\delta_{k,[(\mathbf{x},\mathbf{z}),\lambda]}} \delta_{k,[(\mathbf{x},\mathbf{z}),\lambda']} \sum_{c\in\mathbb{Z}_p} \delta_{\lambda-\lambda',c(\mathbf{x},\mathbf{z})}
\end{aligned}
$$

The final step follows from the fact that the delta function $\delta_{0,[(\mathbf{x},\mathbf{z}),\sigma]}$ identifies a subspace of $\mathbb{Z}_p^n \times \mathbb{Z}_p^n$; which is a vector space over the field $\mathbb{Z}_p$. On this subspace, the exponent $[\lambda-\lambda',\sigma]$ is 0 for all $\sigma$ if and only if $\lambda-\lambda' = c(\mathbf{x},\mathbf{z})$ for some $c \in \mathbb{Z}_p$. The constraint $0 = [(\mathbf{x},\mathbf{z}),\sigma]$ is one-dimensional, so the subspace it identifies has size $p^{2n-1}$.

Finally, we evaluate $\mu$. We do this in two steps, the first of which is to re-express $\rho_S$

as a sum of Paulis rather than a product of projectors. This proof is much easier if we use the notation $S = \{(\lambda_i, k_i)\}$, where $|S| = n$ (the number of quo$p$its):

$$\rho_S = \prod_{i=1}^{n} \Pi_{k_i}^{\lambda_i} \tag{A.6}$$

$$= \prod_{i=1}^{n} \frac{1}{p} \sum_{m_i \in \mathbb{Z}_p} \omega^{-m_i k_i} T_{m_i \lambda_i} \tag{A.7}$$

$$= \frac{1}{p^n} \sum_{m \in \mathbb{Z}_p^n} \prod_{i=1}^{n} \omega^{-m_i k_i} T_{m_i \lambda_i} \tag{A.8}$$

$$= \frac{1}{p^n} \sum_{m \in \mathbb{Z}_p^n} \omega^{-\sum_i m_i k_i} T_{\sum_i m_i \lambda_i} \tag{A.9}$$

The first step was to substitute the definition of $\rho_S$, the second to substitute the definition of the projector. We then interchange the sum and the product, which results in a sum over $\mathbb{Z}_p^n$ rather than $\mathbb{Z}_p$; we've implicitly defined $m = (m_1, \ldots, m_n)$. The last step combines the phases and combines the Paulis—the latter is a special case of the composition rule that we can use because all of the elements of $S$ commute with one another, so the phase one gets when combining Paulis disappears.

It is now easy to directly evaluate $\mu$ using this expression:

$$\mu(\lambda|\rho_S) = \frac{1}{p^n} \operatorname{tr}(A_\lambda \rho_S) \tag{A.10}$$

$$= \frac{1}{p^n} \frac{1}{p^n} \sum_{m \in \mathbb{Z}_p^n} \omega^{-\sum_i k_i m_i} \operatorname{tr}\big(A_\lambda T_{\sum_i m_i \lambda_i}\big) \tag{A.11}$$

$$= \frac{1}{p^{2n}} \sum_{m \in \mathbb{Z}_p^n} \omega^{-\sum_i m_i k_i} \omega^{[\sum_i m_i \lambda_i, \lambda]} \tag{A.12}$$

$$= \frac{1}{p^{2n}} \sum_{m \in \mathbb{Z}_p^n} \omega^{\sum_i m_i ([\lambda_i, \lambda] - k_i)} \tag{A.13}$$

$$= \frac{1}{p^{2n}} \prod_{i=1}^{n} \sum_{m_i \in \mathbb{Z}_p} \omega^{m_i ([\lambda_i, \lambda] - k_i)} \tag{A.14}$$

$$= \frac{1}{p^{2n}} \prod_{i=1}^{n} \delta_{k_i, [\lambda_i, \lambda])} \tag{A.15}$$

This is equivalent to the expression given in the main text when translated back to the usual notation.

### Derivation of state-update rule

From the Wigner function for states, we can derive the state-update rule. We are looking for $\eta_{M_k}(\lambda'|\lambda)$ such that

$$\mu(\lambda'|P_{\mathcal{E}_{M_k}(\rho)}) = \sum_{\lambda} \eta(\lambda'|k, \lambda, M)\mu(\lambda|P_\rho) \tag{A.16}$$

where $\mathcal{E}_{M_k}(\rho) = \frac{M_k \rho M_k}{\mathrm{tr}\left(M_k \rho M_k^\dagger\right)}$ is the channel described by Luder's rule. To find it, we expand the left-hand side of this equation:

$$\mu(\lambda'|P_{\mathcal{E}_{M_k}(\rho)}) = \frac{1}{p^n} \mathrm{tr}\left( A_{\lambda'} \frac{M_k \rho M_k^\dagger}{\mathrm{tr}\left(M_k \rho M_k^\dagger\right)} \right) \tag{A.17}$$

$$= \frac{1}{p^n} \frac{\mathrm{tr}\left(A_{\lambda'} M_k \rho M_k^\dagger\right)}{\mathrm{tr}\left(M_k \rho M_k^\dagger\right)} \tag{A.18}$$

$$= \frac{1}{p^n} \frac{\mathrm{tr}\left(A_{\lambda'} M_k (\sum_\lambda \mu_\rho(\lambda) A_\lambda) M_k^\dagger\right)}{\mathrm{tr}\left(M_k (\sum_\lambda \mu_\rho(\lambda) A_\lambda) M_k^\dagger\right)} \tag{A.19}$$

$$= \frac{1}{p^n} \frac{\sum_\lambda \mathrm{tr}\left(A_{\lambda'} M_k A_\lambda M_k^\dagger\right) \mu_\rho(\lambda)}{\sum_\lambda \mathrm{tr}\left(M_k A_\lambda M_k^\dagger\right) \mu_\rho(\lambda)} \tag{A.20}$$

where we have used the orthonormality of the $A_{\lambda'}$ to write $\rho = \sum_\lambda \mu_\rho(\lambda) A_\lambda$. Thus we can identify

$$\eta_{M_k}(\lambda'|\lambda) = \frac{1}{N(\lambda)} \frac{1}{p^n} \mathrm{tr}\left(A_{\lambda'} M_k A_\lambda M_k^\dagger\right) \tag{A.21}$$

where $N(\lambda)$ is a normalizing constant that may depend on $\lambda$. Summing over $\lambda'$ and setting the result equal to one gives $N(\lambda) = \mathrm{tr}\left(M_k A_\lambda M_k^\dagger\right)$

# Appendix B

# A (very) brief introduction to Shannon information theory

We introduce the very basics of classical information theory along with more specific points that are relevant to Chapter 5. We refer the interested reader to [20] as the canonical introductory/reference text for Shannon information theory.

Shannon information is most straightforward and well-behaved when applied to discrete random variables. The fundamental quantity is the discrete *mutual information* $I_{\mathrm{disc}}$ between two random variables, defined as

$$I_{\mathrm{disc}}[X;Y] = \sum_{x\in\boldsymbol{X},y\in\boldsymbol{Y}} \Pr(x,y)\log\frac{\Pr(x,y)}{\Pr(x)\Pr(y)}. \tag{B.1}$$

The log is usually taken base 2, in which case the mutual information has units of bits; other choices of base just change the units. We take $0\log 0 = 0$, $0\log\frac{0}{0} = 0$, etc. as these functions are well-behaved in the limits towards 0. The mutual information is bounded by

$$0 \leq I_{\mathrm{disc}}[X;Y] \leq \min\{\log|\boldsymbol{X}|, \log|\boldsymbol{Y}|\}. \tag{B.2}$$

It is 0 if and only if $X$ and $Y$ are independent; in this case, $\Pr(x,y) = \Pr(x)\Pr(y)$ so every element of the sum includes a term $\log(1) = 0$. Note that it is symmetric in its inputs, i.e. $I_{\mathrm{disc}}[X;Y] = I_{\mathrm{disc}}[Y;X]$.

Before we move on to other informational quantities, we first give the more general definition of mutual information which applies to any random variable, including continuous and singular random variables. Let $\Delta = \{\Delta_i\}$ be some finite partitioning of an alphabet for a random variable, i.e. $|\Delta| < \infty$ and

$$\bigcup_i \Delta_i = \boldsymbol{X} \qquad \text{and} \qquad \Delta_i \cap \Delta_j = \emptyset \text{ if } i \neq j. \tag{B.3}$$

Then we can define a coarse-grained version of that random variable as

$$\Pr(X^\Delta = i) = \sum_{x \in \Delta_i} \Pr(x)$$
$$\text{or}$$
$$\Pr(X^\Delta = i) = \int_{\Delta_i} \mathrm{d}x \Pr(x), \tag{B.4}$$

depending on whether $X$ is a discrete or continuous random variable (If it's mixed, a mix of sums and integrals essentially works)[1]. The bounds and symmetry of $I_{\text{disc}}$ carry over to $I$, with the caveat that the upper bond $\min\{\log|\boldsymbol{X}|, \log|\boldsymbol{Y}|\}$ may be infinite.

Finally, this allows us to define the general form for mutual information as a supremum over partitions $\Delta^{(1)}$ and $\Delta^{(2)}$:

$$I[X;Y] = \sup_{\Delta^{(1)},\Delta^{(2)}} I_{\text{disc}}[X^{\Delta^{(1)}}; Y^{\Delta^{(2)}}] \tag{B.5}$$

If $X$ and $Y$ are both discrete random variables, then $I_{\text{disc}}[X;Y] = I[X;Y]$, so we only need to work with $I$ from here on.

The mutual information allows for definitions of all of the rest of the quantities of interest in this paper. The *Shannon entropy*, or just *entropy*, of a random variable $X$ is

$$H[X] = I[X;X], \tag{B.6}$$

so it is sometimes also called the self-information. It describes the amount of uncertainty present in the random variable, which is equivalent to the average number of bits needed to communicate that random variable across a noiseless channel.

The *joint entropy* $H[X,Y]$ of two variables is simply the entropy of their joint random variable, and doesn't need further definition. The *conditional entropy* is defined as

$$H[X|Y] = H[X,Y] - H[Y], \tag{B.7}$$

[1]There are some issues with Riemann integrability here that are resolved by a proper measure theoretic treatment, but in this appendix we are interested in communicating the general idea rather than precise mathematics

and describes the uncertainty present in $X$ if one knows the value of $Y$. An equivalent definition is

$$H[X|Y] = H[X] - I[X;Y]. \tag{B.8}$$

We can then use this to define the conditional mutual information:

$$\begin{aligned} I[X;Y|Z] &= H[X|Z] - H[X|Y,Z] && \text{(B.9)} \\ &= H[Y|Z] - H[Y|X,Z] && \text{(B.10)} \end{aligned}$$

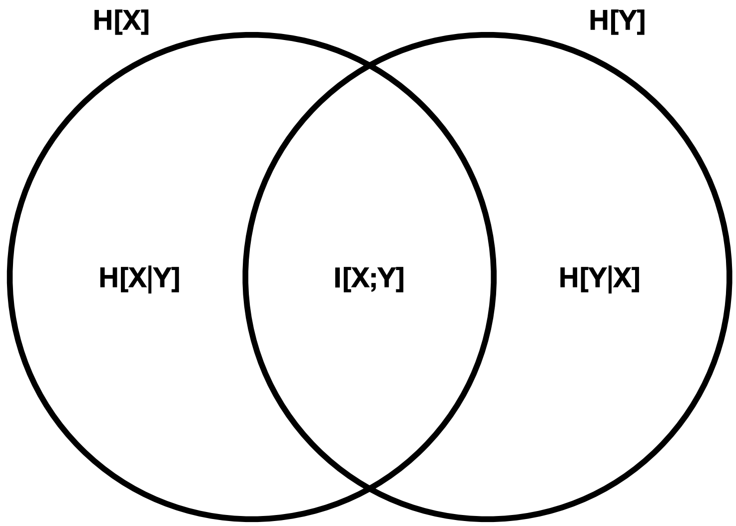
In general, while it is possible to do all of this symbolically, it is usually easier to use I-diagrams [20, 92]. This representation makes use of the fact that the information quantities above are measures on sets; thus they are in strict analogy to set operations and can be represented rigorously by Venn diagrams as in Fig. B.1. More particularly, the additive relationships between these quantities behaves like set unions, set division, etc. Looking at the first I-diagram which has two random variables, we see that the regions labelled $I[X;Y]$ and $H[X|Y]$ together make up the circle labelled $H[X]$; we can then conclude that $I[X;Y] + H[X|Y] = H[X]$, which is indeed the same as Eq. (B.8). The whole area of the diagram is the joint entropy $H[X,Y]$, and we can read off that this is equivalent to $I[X;Y] + H[X|Y] + H[Y|X]$, and thus use the diagram to generate new relationships. The main feature of the three-variable diagram that makes it different from the two-variable diagram is the existence of the three-way overlap that is labelled by the tripartite mutual information $I[X;Y;Z]$, so far undefined. While operational notions of tripartite mutual information are notoriously hard to pin down, the I-diagram provides solid enough motivation for this new quantity, and we can read off its definition as

$$I[X;Y;Z] = I[X;Y] - I[X;Y|Z] \tag{B.11}$$

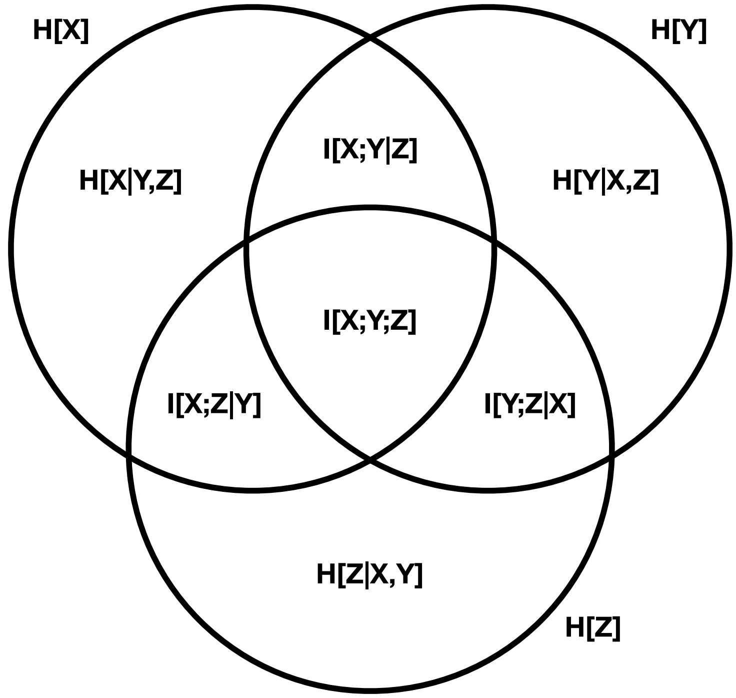
or equivalently by swapping around $X$, $Y$, and $Z$ on the right hand side. One has to be careful as soon as I-diagrams have three-way overlaps, because the tripartite mutual information is no longer positive and thus our information measures technically form a *signed* measure on sets. Also, just because we see that $I[X;Y|Z]$ is 'contained' within $I[X;Y]$, the potential negativity of $I[X;Y;Z]$ does not allow us to conclude that $I[X;Y|Z] \leq I[X;Y]$ (which is indeed false in general).

To summarize the use of I-diagrams in reasoning about information measures, we write down the formal substitutions: joint random variables correspond to unions $(, \to \cup)$, mutual informations to intersections $(; \to \cap)$, and conditionals to set division $(| \to -)$ [92].

Generically, continuous random variables have infinite entropy $H$. However, a Dirac delta function will have 0 entropy, and a weighted sum of Dirac deltas will have finite




(a) Two random variables




(b) Three random variables

Figure B.1: Examples of generic I-diagrams, described in the text.

entropy as calculated by treating the resulting distribution as a discrete distribution. This can be seen by revisiting the supremum over coarse-grainings from the definition of $I$, this time evaluated for $H$. For a random variable $X \sim \delta(x_0)$, we take

$$H[X] = \sup_{\Delta} H_{\text{disc}}[X^{\Delta}]. \tag{B.12}$$

In the coarse-graining, the delta function will exist in only the subset $\Delta_i$ such that $x_0 \in \Delta_i$. Thus the coarse-grained distribution will be 1 or 0 everywhere, and the supremum over entropies will be 0. We ignored some issues having to do with boundaries and Riemann integration, but, as above, these are resolved by a proper measure-theoretic treatment. The result for a weighted sum of Dirac deltas follows similarly. The results above may be summed up by saying that the finitude of a random variable depends not on whether its *alphabet* is continuous or discrete, but rather whether or not its *support* is continuous or discrete.