Evidence for the epistemic view of quantum states: A toy theory

Robert W. Spekkens
Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario, Canada N2L 2Y5
(Received 11 October 2005; revised manuscript received 2 November 2006; published 19 March 2007)

We present a toy theory that is based on a simple principle: the number of questions about the physical state of a system that are answered must always be equal to the number that are unanswered in a state of maximal knowledge. Many quantum phenomena are found to have analogues within this toy theory. These include the noncommutativity of measurements, interference, the multiplicity of convex decompositions of a mixed state, the impossibility of discriminating nonorthogonal states, the impossibility of a universal state inverter, the distinction between bipartite and tripartite entanglement, the monogamy of pure entanglement, no cloning, no broadcasting, remote steering, teleportation, entanglement swapping, dense coding, mutually unbiased bases, and many others. The diversity and quality of these analogies is taken as evidence for the view that quantum states are states of incomplete knowledge rather than states of reality. A consideration of the phenomena that the toy theory fails to reproduce, notably, violations of Bell inequalities and the existence of a Kochen-Specker theorem, provides clues for how to proceed with this research program.

DOI: 10.1103/PhysRevA.75.032110 PACS number(s): 03.65.Ta, 03.67.−a

I. INTRODUCTION

In this article, we introduce a simple toy theory based on a principle that restricts the amount of knowledge an observer can have about reality. Although not equivalent to quantum theory nor even competitive as an explanation of empirical phenomena, it reproduces in detail a large number of phenomena that are typically taken to be characteristically quantum. This, and the fact that the object analogous to the quantum state in the toy theory is a state of incomplete knowledge, are the grounds upon which we argue for our thesis: that quantum states are also states of incomplete knowledge.

We begin by clarifying the distinction between states of reality and states of knowledge. To be able to refer to it conveniently, we introduce the qualifiers *ontic* (from the Greek *ontos,* meaning "to be") and *epistemic* (from the Greek *epistēmē,* meaning "knowledge"). An *ontic state* is a state of reality, whereas an *epistemic state* is a state of knowledge. To understand the content of the distinction, it is useful to study how it arises in the uncontroversial context of classical physics.

The first notion of state that students typically encounter in their study of classical physics is the one associated with a point in phase space. This state provides a complete specification of all the properties of the system (in particle mechanics, such a state is sometimes called a "Newtonian state"). It is an ontic state. On the other hand, when a student learns classical *statistical* mechanics, a new kind of state is introduced, corresponding to a probability distribution over the phase space (sometimes called a "Liouville state"). This is an epistemic state. The critical difference between a point in phase space and a probability distribution over phase space is not that the latter is a function. An electromagnetic field configuration is a function over three-dimensional space, but is nonetheless an ontic state. What is critical about a probability distribution is that the relative height of the function at two different points is not a property of the system (unlike the relative height of an electromagnetic field at two points in space). Rather, this relative height represents the relative likelihood that some agent assigns to the two ontic states associated with those points of the phase space. The distribution describes only what this agent knows about the system.

There is one case wherein the distinction between an ontic state and an epistemic state breaks down, and that is for epistemic states describing complete knowledge, because the latter also contain a complete specification of a system's properties. For example, states of complete knowledge in a classical theory are represented by Dirac-δ functions on phase space, and these are associated one-to-one with the points of phase space. The epistemic states with which we shall be interested in this paper—the ones with which we hope to associate quantum states—are those describing *incomplete* knowledge.

A widespread view among physicists and philosophers of physics is that pure quantum states are ontic states. Only mixed quantum states are taken to be epistemic states, specifically, states of incomplete knowledge about which pure quantum state is really occurrent. In a variant of this view, even the mixed quantum states are interpreted as ontic (this approach is motivated by the fact that a mixed state may be expressed as a convex sum of pure states in many different ways). We shall describe proponents of both of these viewpoints as proponents of the *ontic view* of quantum states. In contrast, the thesis we wish to defend is that *all* quantum states, mixed and pure, are states of incomplete knowledge. This viewpoint will be referred to as the *epistemic view* of quantum states.

The ontic view of quantum states has a long history in the interpretation of quantum mechanics. Schrödinger initially interpreted the quantum state as a physical wave, and never wholly abandoned this view. In the classic textbooks of Dirac [1] and of von Neumann [2], the quantum state is taken to provide a complete specification of the properties of a system. This is also true of both collapse theories [3–5] and Everett-type interpretations [6,7]. Even within the popular hidden variable theories, such as the de Broglie-Bohm theory [8–10] and the modal interpretation [11–13], although the quantum state has an epistemic role to play in specifying the probability distribution over hidden variables, it is fundamentally an ontic state insofar as it acts as a guiding wave, causally influencing the dynamics of the hidden variables. The tension between the epistemic and ontic roles of the quantum state in these interpretations has understandably troubled many authors, and although efforts have been made to reduce the tension [10], these have tended to assign *less* rather than more epistemic significance to the quantum state.

The epistemic view, although less common than the ontic view, also has a long tradition. As we shall see in detail further on, Einstein's argument for the incompleteness of quantum mechanics (which is most clear in his correspondence with Schrödinger [14] but was made famous in the Einstein, Podolsky, and Rosen (EPR) paper [15]) is an argument for an epistemic view of quantum states. The work of Ballentine on the statistical interpretation [16,17] can be interpreted as a defense of the epistemic view, as can that of Emerson [18]. Peierls was also an early advocate of this interpretation of the quantum state [19]. It is only recently, with the advent of quantum information theory, that the epistemic view has become more widespread. The work of Caves, Fuchs, and Schack has been particularly notable in promoting this view [20–22], with Fuchs's manifesto on the subject [23] being one of the most eloquent and convincing to date. The present work owes much of its inspiration to this research program, in particular, the idea of deriving quantum phenomena from a principle that maximal information is incomplete and cannot be completed [20,24].

Despite the fact that the epistemic view appears to be on the rise, our impression is that many would-be supporters have failed to completely abandon their ontic preconceptions, perhaps due to the ubiquity of ontic language in the literature and perhaps due to a vague feeling that the epistemic path is one that has been shown to be inconsistent. We hope through this article to correct some of these misconceptions and to increase the respectability of this viewpoint.

We shall argue for the superiority of the epistemic view over the ontic view by demonstrating how a great number of quantum phenomena that are mysterious from the ontic viewpoint appear natural from the epistemic viewpoint. These phenomena include interference, noncommutativity, entanglement, no cloning, teleportation, and many others. Note that the distinction we are emphasizing is whether the phenomena can be understood *conceptually*, not whether they can be understood as mathematical consequences of the formalism, since the latter type of understanding is possible regardless of one's interpretation of the formalism. The greater the number of phenomena that appear mysterious from an ontic perspective but natural from an epistemic perspective, the more convincing the latter viewpoint becomes. For this reason, the article devotes much space to elaborating on such phenomena.

Of course, a proponent of the ontic view might argue that the phenomena in question are not mysterious *if* one abandons certain preconceived notions about physical reality. The challenge we offer to such a person is to present a few simple *physical* principles by the light of which all of these phenomena become conceptually intuitive (and not merely mathematical consequences of the formalism) within a framework wherein the quantum state is an ontic state. Our impression is that this challenge cannot be met. By contrast, a single information-theoretic principle, which imposes a constraint on the amount of knowledge one can have about any system, is sufficient to derive all of these phenomena in the context of a simple toy theory, as we shall demonstrate.

A few words are in order about the motivation for such a principle. In Liouville mechanics, states of incomplete knowledge exhibit phenomena analogous to those exhibited by pure quantum states. Among these are the existence of a no-cloning theorem for such states [20,25], the impossibility of discriminating such states with certainty [20,26], the lack of exponential divergence of such states (in the space of epistemic states) under chaotic evolution [27], and, for correlated states, many of the features of entanglement [28]. On the other hand, states of *complete* knowledge do not exhibit these phenomena. This suggests that one would obtain a better analogy with quantum theory if states of complete knowledge were somehow impossible to achieve, that is, if somehow *maximal* knowledge was always *incomplete* knowledge [20,24,29]. This idea is borne out by the results of this paper. In fact, the toy theory suggests that the restriction on knowledge should take a particular form, namely, that one's knowledge be quantitatively equal to one's ignorance in a state of maximal knowledge.

It is important to bear in mind that one cannot derive quantum theory from the toy theory, nor from any simple modification thereof. The problem is that the toy theory is a theory of incomplete knowledge about local and noncontextual hidden variables, and it is well known that quantum theory cannot be understood in this way [30,32,33]. This prompts the obvious question: if a quantum state is a state of knowledge, and it is not knowledge of local and noncontextual hidden variables, then what is it knowledge about? We do not at present have a good answer to this question. We shall therefore remain completely agnostic about the nature of the reality to which the knowledge represented by quantum states pertains. This is not to say that the question is not important. Rather, we see the epistemic approach as an unfinished project, and this question as the central obstacle to its completion. Nonetheless, we argue that even in the absence of an answer to this question, a case can be made for the epistemic view. The key is that one can hope to identify phenomena that are characteristic of states of incomplete knowledge regardless of what this knowledge is about.

The outline of the paper is as follows. In Sec. II, we introduce our foundational principle—that there is a balance between knowledge and ignorance in a state of maximal knowledge—and define our measures of knowledge and ignorance. From this starting point, and a few other assumptions, we derive the toy theory. We begin in Sec. III by considering the simplest possible system that can satisfy the principle. In Sec. IV we consider pairs of these systems, and in Sec. V, triplets. For each of these cases, we determine the epistemic states, measurements, and transformations that are allowed by the principle, as well as the manner in which epistemic states must be updated after a measurement. Along the way, we draw attention to various analogues of quantum phenomena. Some additional analogues are enumerated in Sec. VI, while in Sec. VII we identify some quantum phenomena that are *not* reproduced by the toy theory and consider what these teach us about how to proceed with the epistemic research program. In Sec. VIII, we discuss related work, specifically, Kirkpatrick's playing card model [34], Hardy's toy theory [35], Smolin's "lockboxes" [36], Zeilinger's foundational approach [37], and Wootters' discrete Wigner function [38]. We conclude in Sec. IX with some questions for future research. Some additional material is presented in the Appendices, namely, a discussion of why the toy theory for N elementary systems cannot be understood as a restriction upon quantum theory for N qubits, and of the significance of our results for information-theoretic derivations of quantum theory.

II. THE KNOWLEDGE BALANCE PRINCIPLE

The toy theory is built on the following foundational principle:

> If one has maximal knowledge, then for every system, at every time, the amount of knowledge one possesses about the ontic state of the system at that time must equal the amount of knowledge one lacks.

We call this the *knowledge balance principle*. As stated, it is not sufficiently explicit, because the manner of quantifying the amount of knowledge one possesses and the amount one lacks has yet to be specified. Although the measure of knowledge that we adopt is very simple, it is not a conventional one, and consequently we must define it carefully.

We begin by introducing the notion of a *canonical* set of yes-no questions. This is a set of yes-no questions that is sufficient to fully specify the ontic state, and that has a minimal number of elements. To clarify this notion, consider a situation wherein there are four possible ontic states. A set of questions that will determine which of the four applies is as follows: "Is it 1, or not?," "Is it 2, or not?." "Is it 3, or not?," and "Is it 4, or not?." This questioning scheme is inefficient however. A more efficient scheme divides the set of possibilities into two with every question. Indeed, one can fully specify the ontic state with just two questions, for instance: "Is it in the set $\{1,2\}$, or not?" and "Is it in the set $\{1,3\}$, or not?." As there are four answers to two yes-no questions, two is the minimal number of questions that can possibly specify which of four states applies. So the two questions just described form a canonical set. Note also that there can be many canonical sets of questions. For instance, a different pair of questions, namely, "Is it in the set $\{1,2\}$, or not?" and "Is it in the set $\{2,3\}$, or not?" also form a canonical set.

With the notion of a canonical set in hand, we can define our measure of knowledge. It is simply the maximum number of questions for which the answer is known, in a variation over all canonical sets of questions. Our measure of ignorance is simply the difference between this number and the total number of questions in the canonical set.

The knowledge balance principle, made specific with our measure of knowledge, is the starting point of the toy theory. There are, however, a few other assumptions that shall go into its derivation, to which we now turn.

We assume that all physical systems are such that there can be a balance of knowledge about them. This implies that the number of yes-no questions in a canonical set must be a multiple of two, because if this were not the case, one could not have an equality between the number of questions answered and the number of questions unanswered. The simplest possible case is to have just two questions in the canonical set. Because a canonical set is, by definition, the minimal sufficient set of questions required to specify the ontic state, it follows that for two questions there are four possible ontic states. Thus the simplest possible system in the toy theory has four ontic states. We call this an *elementary system.*

We shall also assume that every system is built of elementary systems. For a pair of elementary systems, there are four questions in the canonical set, and sixteen possible ontic states in all. For N systems, there are $2N$ questions in the canonical set and 2^{2N} possible ontic states. This "reductionist" assumption will have very significant consequences in the development of the toy theory, as the knowledge balance principle will yield more constraints for composite systems: not only must there be a balance of knowledge and ignorance for the whole, but for every part of the whole, right down to the smallest subsystems.

The motional degree of freedom for all systems is treated classically, a background of flat space-time is assumed and every elementary system is taken to exist at a point in space.

We also assume that the outcome of a reproducible measurement depends only on the ontic state of the system being measured. Moreover, we assume that a transformation applied to one system can only affect the ontic state of that system, and not the ontic state of others. If the systems are spatially separated, this amounts to an assumption of locality. Further, we shall assume that an observer's state of knowledge about a system does not dictate what can and cannot be done to the system, nor does it ever determine the change that occurs in the system's ontic state during a measurement. This assumption is motivated by the implausibility of there being a causal relation between the mental state of the agent and the ontic state of the apparatus or the system.

Finally, we assume that information gain about a system is always possible. This will allow us to infer the existence of a disturbance when a reproducible measurement is performed, rather than inferring the impossibility of reproducible measurements.

III. ELEMENTARY SYSTEMS

A. Epistemic states

An elementary system is one for which the number of questions in the canonical set is two, and consequently the number of ontic states is four. Although it takes two yes-no questions to specify the ontic state, the answer to only one of these can be known according to the knowledge balance principle. Thus, the epistemic states for which the balance occurs are those which identify the ontic state of the system to be one of two possibilities. Denoting the four ontic states by "1," "2," "3," and "4," and disjunction by the symbol "$\vee$" (read as "or"), we can specify the possible epistemic states as disjunctions of the ontic states. In all, there are six states of maximal knowledge, namely,

$$1 \vee 2,$$

$$3 \vee 4,$$

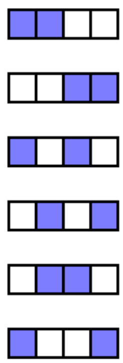
$$1 \vee 3,$$

$$2 \vee 4,$$

$$2 \vee 3,$$

$$1 \vee 4. \tag{1}$$

It is useful to represent these graphically as follows:



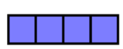
(2)

with the understanding that the four cells represent the four ontic states, and the filled cells denote the set in which the actual ontic state of the system is known to lie.

For a single elementary system, the only way to have less than maximal knowledge is for both questions in the canonical set to be unanswered. This corresponds to the epistemic state

$$1 \vee 2 \vee 3 \vee 4. \tag{3}$$

It is denoted pictorially by



(4)

A single elementary system in the toy theory is analogous to a system described by a two-dimensional Hilbert space in quantum theory, called a *qubit* in quantum information theory. In particular, the six epistemic states describing maximal knowledge of a single elementary system are analogous to the following six pure qubit states

$$1 \vee 2 \Leftrightarrow |0\rangle,$$

$$3 \vee 4 \Leftrightarrow |1\rangle,$$

$$1 \vee 3 \Leftrightarrow |+\rangle,$$

$$2 \vee 4 \Leftrightarrow |-\rangle,$$

$$2 \vee 3 \Leftrightarrow |+i\rangle,$$

$$1 \vee 4 \Leftrightarrow |-i\rangle, \tag{5}$$

where $|\pm\rangle = \frac{1}{\sqrt{2}}|0\rangle \pm |1\rangle$ and $|\pm i\rangle = \frac{1}{\sqrt{2}}|0\rangle \pm i|1\rangle$, while the single state of nonmaximal knowledge is analogous to the completely mixed state for a qubit, that is,

$$1 \vee 2 \vee 3 \vee 4 \Leftrightarrow I/2, \tag{6}$$

where I is the identity operator on the two-dimensional Hilbert space. The rest of this section will demonstrate the extent of this analogy. Note, however, that the choice of which three epistemic states to associate with $|0\rangle, |+\rangle$, and $|+i\rangle$ is simply a convention.

1. Disjointness

It is useful to define the *ontic support* of an epistemic state to be the set of ontic states which are consistent with it. For instance, the ontic support of $1 \vee 2$ is the set $\{1,2\}$. If the intersection of the ontic supports of a pair of epistemic states is empty, then those states are said to be *disjoint.* A set of epistemic states are said to be disjoint if they are pairwise disjoint. The relation of disjointness is analogous to the relation of orthogonality among quantum states. The fact that there are pairs of epistemic states which are nondisjoint demonstrates that there exists an analogue of nonorthogonality in the toy theory.

2. Fidelity

One can also introduce a measure of the degree of nondisjointness, equivalently, a measure of the distance between a pair of epistemic states in the space of such states. A standard measure of distance between two probability distributions, $\mathbf{p} = (p_k)_k$ and $\mathbf{q} = (q_k)_k$, is the *classical fidelity*, defined by $F(\mathbf{p},\mathbf{q}) = \Sigma_k \sqrt{p_k}\sqrt{q_k}$. If we treat the epistemic states of the toy theory as uniform probability distributions, for instance, associating the distribution $(1/2,1/2,0,0)$ with $1 \vee 2$, and $(1/4,1/4,1/4,1/4)$ with $1 \vee 2 \vee 3 \vee 4$, then we can use the classical fidelity as a measure of distance. For the epistemic states of a single elementary system, the fidelity between a pair takes one of four values: the value 0 if they are disjoint, such as $1 \vee 2$ and $3 \vee 4$; the value $1/2$ if they are nondisjoint states of maximal knowledge, such as $1 \vee 2$ and $1 \vee 3$; the value $1/\sqrt{2}$ if one is a state of maximal knowledge, the other not, such as $1 \vee 2$ and $1 \vee 2 \vee 3 \vee 4$; and the value 1 if the elements of the pair are identical.

The analogous measure of distance between quantum states is the quantum fidelity [39], which is defined for a pair of density operators, ρ and σ, as $\mathrm{Tr}|\sqrt{\rho}\sqrt{\sigma}|$. In the case of a pair of pure states, $|\psi\rangle$ and $|\chi\rangle$, the fidelity is simply the inner product squared, $|\langle\psi|\chi\rangle|^2$. It turns out that the classical fidelities between pairs of epistemic states are precisely equal to the quantum fidelities for the analogous pairs of quantum states under the mapping of Eq. (5). For instance, the quantum fidelity between $|0\rangle$ and $|1\rangle$ is 0, between $|0\rangle$ and $|+\rangle$ is $1/2$, between $|0\rangle$ and $I/2$ is $1/\sqrt{2}$, and between any state and itself is 1.

3. Compatibility

Another useful relation to introduce is that of compatibility. Two epistemic states are said to be *compatible* if the intersection of their ontic supports is the ontic support of a valid epistemic state. Thus, the epistemic states $1\vee 2$ and $1\vee 2\vee 3\vee 4$ have the ontic states 1 and 2 in common, and are therefore compatible, while $1\vee 2$ and $2\vee 3$ have only the ontic state 2 in common, and are therefore incompatible. Whenever two observers are describing the same system, their epistemic states must be compatible. This follows from the fact that if these individuals pool their information they will rule out any ontic state that either one of them rules out, which is equivalent to taking the intersection of the ontic supports of their epistemic states. If their epistemic states were incompatible, this would result in a final epistemic state that violated the knowledge balance principle. Note that this implies that if two observers both have maximal knowledge of a system, then their states of knowledge must be identical; there is always intersubjective agreement among maximally informed observers. This relation of compatibility is analogous to the Brun-Finkelstein-Mermin compatibility relation for quantum states, according to which two states are compatible whenever the intersection of their supports (in Hilbert space) is not null [40].

4. Convex combination

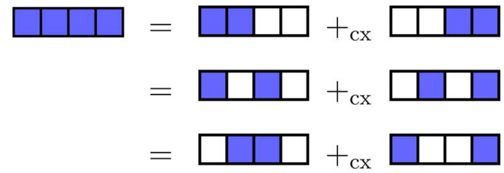
We now introduce a way of combining epistemic states that is analogous to the convex combination (or incoherent superposition) of quantum states. A pair of epistemic states in the toy theory must satisfy two conditions for the convex combination to be defined. The first condition is that they be disjoint. The second condition is that the union of their ontic supports must form the ontic support of a valid epistemic state. If both conditions are met, then the epistemic state that results by taking the union of the ontic supports of the pair is defined to be the convex combination of that pair. Thus, the convex combination of $1\vee 2$ and $3\vee 4$ is $1\vee 2\vee 3\vee 4$, while the convex combination of $1\vee 2$ and $1\vee 3$ is undefined, as is the convex combination of $1\vee 2$ and $1\vee 2\vee 3\vee 4$. The convex combination of a larger set of epistemic states is defined similarly.

Note that in addition to being sometimes undefined, the convex combination of a set of epistemic states in the toy theory also differs from the convex combination of a set of quantum states in there being nothing analogous to a convex sum with unequal weights.

It is useful to introduce the terms *mixed* and *pure* to specify whether or not an epistemic state can be obtained as a convex combination of distinct epistemic states or not. For a single elementary system, the epistemic states $1\vee 2$, $3\vee 4$, $1\vee 3$, $2\vee 4$, $1\vee 4$, and $2\vee 3$ are pure, while the epistemic state $1\vee 2\vee 3\vee 4$ is mixed. There are in fact many convex decompositions of $1\vee 2\vee 3\vee 4$. Denoting convex combination by the symbol "$+_{\text{cx}}$," we have

$$\begin{aligned}1\vee 2\vee 3\vee 4 &= (1\vee 2)+_{\text{cx}}(3\vee 4) = (1\vee 3)+_{\text{cx}}(2\vee 4)\\ &= (2\vee 3)+_{\text{cx}}(1\vee 4).\end{aligned} \tag{7}$$

Graphically,



(8)

This is analogous to the fact that in quantum theory, the completely mixed state of a qubit, $I/2$, has convex decompositions

$$\begin{aligned}I/2 &= \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -|\\ &= \frac{1}{2}|+i\rangle\langle +i| + \frac{1}{2}|-i\rangle\langle -i|.\end{aligned} \tag{9}$$

Thus, the toy theory mirrors quantum theory in admitting multiple convex decompositions of a mixed state into pure states. This multiplicity is a direct result of the fact that in the toy theory, pure epistemic states are states of incomplete knowledge.

5. A geometric representation of the space of epistemic states

In quantum theory, the Bloch sphere (or, more precisely, the Bloch ball) offers a useful geometric representation of the quantum states of a qubit and the relations of orthogonality and convex combination that hold among them [39]. Specifically, orthogonal quantum states are represented by antipodal points on the sphere, and every convex decomposition of a mixed state is associated with a convex polytope that contains in its interior the point representing the mixed state, with the vertices of the polytope representing the elements of the convex decomposition [41]. Similarly, the epistemic states for an elementary system in the toy theory can be represented by a subset of the points inside a unit ball. Disjoint epistemic states are represented by antipodal points, and convex decompositions of the mixed epistemic state are represented by line segments, the end points of which are the elements of the decomposition. The two pictures are presented for comparison in Fig. 1.

6. Coherent superposition

One can also introduce a way of combining epistemic states that is analogous to the coherent superposition of quantum states. What we seek is a binary operation that takes a pair of pure epistemic states to another pure epistemic state (unlike the operation of convex combination we have just introduced, which takes a pair of pure states to a mixed state). Suppose the two epistemic states we seek to combine are of the form $a\vee b$ and $c\vee d$ (here, of course, $a,b,c,d \in \{1,2,3,4\}$ and $a\neq b, c\neq d$). We assume that they are disjoint, so that $a,b\neq c,d$. Moreover, we adopt the convention that $a<b$ and $c<d$. One can define four new pure epistemic states from these two, namely, $a\vee c$, $a\vee d$, $b\vee c$, and $b\vee d$. We can think of these as the result of applying four distinct binary operations to the original pair of states. Denoting these four operations by $+_1$, $+_2$, $+_3$, and $+_4$, we have

$$(a\vee b)+_1(c\vee d) = a\vee c,$$

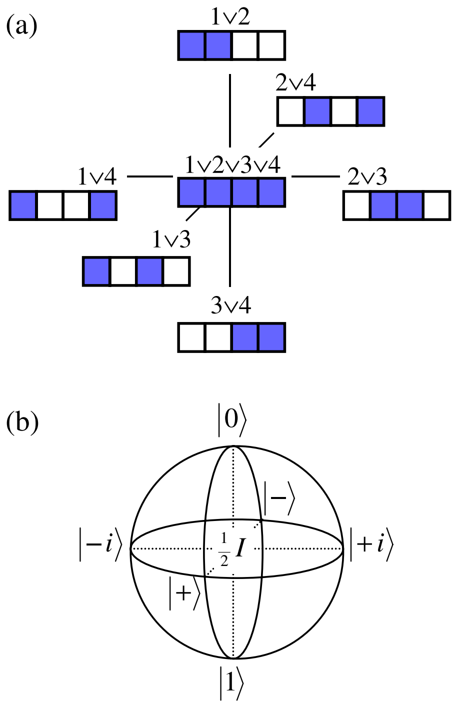


FIG. 1. (Color online) (a) A representation of the space of epistemic states in the toy theory, and (b) the Bloch ball representation of the states in quantum theory.

$$(a \vee b) +_2 (c \vee d) = b \vee d,$$

$$(a \vee b) +_3 (c \vee d) = b \vee c,$$

$$(a \vee b) +_4 (c \vee d) = a \vee d. \tag{10}$$

The first operation can be described as follows: take the ontic state of lowest index from the first epistemic state, and the ontic state of lowest index from the second epistemic, then define a new epistemic state in terms of these. The other three operations can be defined similarly. All that differs is whether one takes the ontic state with the lowest or highest index from each epistemic state. The convention we have chosen is that $+_1$, $+_2$, $+_3$, and $+_4$ are associated with low-low, high-high, high-low, and low-high. We call these *coherent binary operations.*

The four possible coherent binary operations acting on $1\vee 2$ and $3\vee 4$ yield

$$(1 \vee 2) +_1 (3 \vee 4) = 1 \vee 3, \tag{11}$$

$$(1 \vee 2) +_2 (3 \vee 4) = 2 \vee 4, \tag{12}$$

$$(1 \vee 2) +_3 (3 \vee 4) = 2 \vee 3, \tag{13}$$

$$(1 \vee 2) +_4 (3 \vee 4) = 1 \vee 4, \tag{14}$$

acting on $1\vee 3$ and $2\vee 4$, they yield

$$(1 \vee 3) +_1 (2 \vee 4) = 1 \vee 2, \tag{15}$$

$$(1 \vee 3) +_2 (2 \vee 4) = 3 \vee 4, \tag{16}$$

$$(1 \vee 3) +_3 (2 \vee 4) = 2 \vee 3, \tag{17}$$

$$(1 \vee 3) +_4 (2 \vee 4) = 1 \vee 4, \tag{18}$$

and acting on $2\vee 3$ and $1\vee 4$, they yield

$$(2 \vee 3) +_1 (1 \vee 4) = 1 \vee 2, \tag{19}$$

$$(2 \vee 3) +_2 (1 \vee 4) = 3 \vee 4, \tag{20}$$

$$(2 \vee 3) +_3 (1 \vee 4) = 1 \vee 3, \tag{21}$$

$$(2 \vee 3) +_4 (1 \vee 4) = 2 \vee 4. \tag{22}$$

These relations should be compared with the following relations among quantum states:

$$\sqrt{2}^{-1}(|0\rangle + |1\rangle) = |+\rangle, \tag{23}$$

$$\sqrt{2}^{-1}(|0\rangle - |1\rangle) = |-\rangle, \tag{24}$$

$$\sqrt{2}^{-1}(|0\rangle + i|1\rangle) = |+i\rangle, \tag{25}$$

$$\sqrt{2}^{-1}(|0\rangle - i|1\rangle) = |-i\rangle, \tag{26}$$

and

$$\sqrt{2}^{-1}(|+\rangle + |-\rangle) = |0\rangle, \tag{27}$$

$$\sqrt{2}^{-1}(|+\rangle - |-\rangle) = |1\rangle, \tag{28}$$

$$\sqrt{2}^{-1}(|+\rangle + i|-\rangle) = e^{i\pi/4}|-i\rangle, \tag{29}$$

$$\sqrt{2}^{-1}(|+\rangle - i|-\rangle) = e^{-i\pi/4}|+i\rangle, \tag{30}$$

and

$$\sqrt{2}^{-1}(|+i\rangle + |-i\rangle) = |0\rangle, \tag{31}$$

$$\sqrt{2}^{-1}(|+i\rangle - |-i\rangle) = i|1\rangle, \tag{32}$$

$$\sqrt{2}^{-1}(|+i\rangle + i|-i\rangle) = e^{i\pi/4}|+\rangle, \tag{33}$$

$$\sqrt{2}^{-1}(|+i\rangle - i|-i\rangle) = e^{-i\pi/4}|-\rangle. \tag{34}$$

Note that the combinations we have enumerated do not exhaust the possibilities, since for the operations $+_3$ and $+_4$, the order of the arguments in the operation is important. That is, $+_3$ and $+_4$ are not commutative operations. For instance, $(1\vee 2)+_3(3\vee 4)=2\vee 3$ while $(3\vee 4)+_3(1\vee 2)=1\vee 4$. The same sensitivity to ordering is found in quantum theory for coherent superpositions with relative phases $\pi/2$ and $3\pi/2$. For instance, $\sqrt{2}^{-1}(|0\rangle+i|1\rangle)=|+i\rangle$, while $\sqrt{2}^{-1}(|1\rangle+i|0\rangle)=i|-i\rangle$.

It is natural to associate the operations $+_1$, $+_2$, $+_3$, and $+_4$ with coherent superpositions of two quantum states where the relative weights are equal and the relative phases of the second term to the first are $0, \pi, \pi/2$, and $3\pi/2$, respectively,[1]

$$+_1 \Leftrightarrow 0,$$

$$+_2 \Leftrightarrow \pi,$$

$$+_3 \Leftrightarrow \pi/2,$$

$$+_4 \Leftrightarrow 3\pi/2. \tag{35}$$

Under this association of toy-theoretic operations with quantum operations and under the association of epistemic states with quantum states expressed in Eq. (5), the relations (11)–(22) parallel (modulo global phases) the relations (23)–(34), with two notable exceptions. Given the form of the relations (17) and (18), and the fact that $2 \vee 3$ maps to $|+i\rangle$ and $1 \vee 4$ maps to $|-i\rangle$ under Eq. (5), one would expect the right-hand side of Eq. (29) to be proportional to $|+i\rangle$ and the right-hand side of Eq. (30) to be proportional to $|-i\rangle$ rather than vice versa. Note that one cannot achieve a better analogy by modifying the associations adopted in Eqs. (5) and (35). For instance, by associating $2 \vee 3$ with $|-i\rangle$ and $1 \vee 4$ with $|+i\rangle$, the relations (17) and (18) can be made to parallel the relations (29) and (30), however, in this case the relations (13) and (14) fail to parallel (25) and (26). This curious failure of the analogy shows that an elementary system in the toy theory is not simply a constrained version of a qubit.

There are two other important respects in which our coherent binary operations for a single elementary system differ from those one finds in quantum theory for a qubit. First, whereas any pair of quantum states of a qubit can be coherently superposed, the binary operations in the toy theory are not defined for arbitrary pairs of epistemic states. Specifically, they are not defined for nondisjoint epistemic states. Second, whereas there are a continuum of different types of coherent superposition of a pair of quantum states of a qubit, corresponding to all possible relative weights and all possible relative phases, there are only four coherent binary operations in the toy theory.

B. Transformations

We now consider the sorts of transformations of the ontic states that are allowed by the knowledge balance principle. Imagine a transformation that takes two different ontic states, say 1 and 2, to a single ontic state, say 3. If the epistemic state prior to the transformation was $1 \vee 2$, then after the transformation, one would be certain that the ontic state was 3. But such an epistemic state violates the knowledge balance principle, therefore this transformation is not allowed. A similar example can be devised for any many-to-one map. Thus, all such maps are ruled out by the principle.

We are left with the one-to-one maps and the one-to-many maps. We focus on the former here, since these correspond to the reversible maps. Clearly, these are simply the set of permutations of the four ontic states.

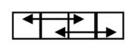
TABLE I. The class structure of the group S_4 of permutations of four elements.

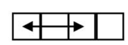
(1^4)	(31)	(21^2)	(2^2)	(4)
(1)(2)(3)(4)	(234)(1)	(12)(3)(4)	(12)(34)	(1234)
	(243)(1)			(1432)
		(13)(2)(4)	(13)(24)	
	(134)(2)			(1243)
	(143)(2)	(14)(2)(3)	(14)(23)	(1342)
	(124)(3)	(23)(1)(4)		(1324)
	(142)(3)			(1423)
		(24)(1)(3)		
	(123)(4)			
	(132)(4)	(34)(1)(2)		

One can describe permutations in terms of cycles. For instance, the permutation $a \to a$, $b \to c \to d \to b$ involves two cycles: a one-cycle, $a \to a$, and a three-cycle $b \to c \to d \to b$. In cycle notation, this permutation is written as $(a)(bcd)$. The set of permutations of 4 elements is the group S_4, containing 24 elements. Permutations with the same number of cycles form a class. We list the elements of S_4, and their class structure in Table I. If an element is written alone, it is its own inverse, whereas elements appearing in pairs are each other's inverses.

The valid transformations may be usefully represented graphically by arrows between the ontic states. For instance,

$(123)(4)$: 

$(13)(24)$: 

$(13)(2)(4)$: 

(1234) :  (36)

It is interesting to determine how the set of epistemic states are transformed under a permutation of the ontic states. For instance, the permutation $(123)(4)$ leads to the following map on the epistemic states

$$1 \vee 2 \to 2 \vee 3,$$

$$3 \vee 4 \to 1 \vee 4,$$

$$1 \vee 3 \to 1 \vee 2,$$

$$2 \vee 4 \to 3 \vee 4,$$

$$2 \vee 3 \to 1 \vee 3,$$

$$1 \vee 4 \to 2 \vee 4. \tag{37}$$

[1]Further justification for this analogy is provided in Sec. III B.

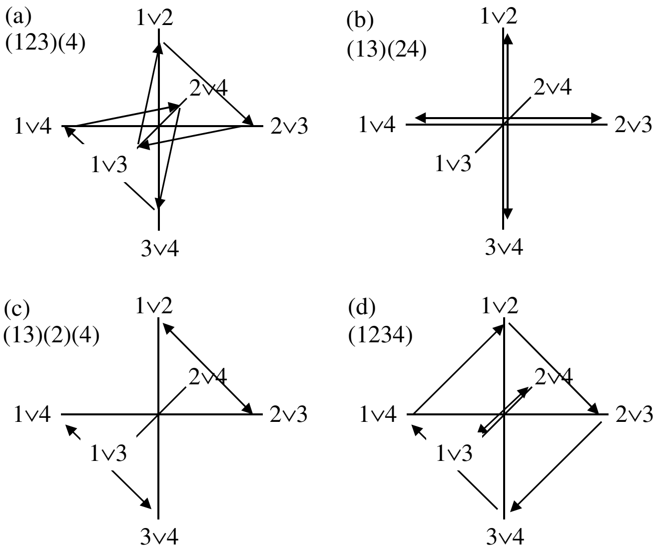


FIG. 2. How four permutations of the ontic states appear in the Bloch sphere representation of the space of epistemic states.

Representing the epistemic states in the "Bloch sphere" picture, we see that this permutation appears as a rotation by 120° about the axis that points in the $\hat{x}+\hat{y}+\hat{z}$ direction, as seen in Fig. 2(a). Similarly, the permutation (13)(24) appears as a rotation by 180° about the $\hat{x}$ axis, as seen in Fig. 2(b). These permutations are analogous to unitary maps in Hilbert space, which appear as rotations in the Bloch sphere. These two examples might lead one to think that *all* permutations appear as rotations in the Bloch sphere picture, but this is not the case. A permutation such as (13)(2)(4) is a reflection about the plane spanned by $\hat{x}$ and $(\hat{y}+\hat{z})$, as seen in Fig. 2(c), while (1234) involves a rotation of 90° about $\hat{x}$ and a reflection about the plane spanned by $\hat{y}$ and $\hat{z}$, as seen in Fig. 2(d). These are analogous to antiunitary maps in Hilbert space. Antiunitary maps do not represent possible evolutions of a system in quantum theory because evolution is assumed to be continuous in time and in the limit of short times only unitary maps go over to the identity map; this is the content of Wigner's theorem. Given that transformations in the toy theory are not continuous but discrete, there is no analogous constraint that would serve to rule out those transformations that are analogous to antiunitaries.

Note that the set of all reversible transformations for an elementary system corresponds to the symmetry group of a tetrahedron the four vertices of which are located along the $\hat{x}-\hat{y}+\hat{z}$ axis, the $-\hat{x}+\hat{y}+\hat{z}$ axis, the $\hat{x}+\hat{y}-\hat{z}$ axis, and the $-\hat{x}-\hat{y}-\hat{z}$ axis. These vertices are associated with the ontic states $1,2,3$, and 4, respectively.

Coherent superposition revisited

The analogy between these transformations and unitary maps lends support to the notion that the coherent binary operations introduced in the previous section are the correct analogues of coherent superpositions. In quantum theory, the notion of coherent superposition is typically defined in terms of the linear sum of vectors in Hilbert space. Given that it is not obvious whether or how the space of epistemic states considered here may be embedded in a vector space, this sort of definition is clearly not available. However, one can also define the notion of coherent superposition in quantum theory using transformations, and this sort of definition *does* have an analogue in the toy theory, as we presently demonstrate.

If a reversible transformation takes the quantum state $|0\rangle$ to $|1\rangle$, then a square root of that transformation (one that when acted twice yields the transformation) takes $|0\rangle$ to an equal superposition of $|0\rangle$ and $|1\rangle$. Recalling that unitaries are rotations on the Bloch sphere (Fig. 1), and denoting a rotation through an angle ϕ about the $\hat{n}$ axis by $R_{\hat{n}}(\phi)$, it is clear that $R_{\hat{y}}(\pi)$ and $R_{\hat{x}}(\pi)$ are both rotations that take $|0\rangle$ to $|1\rangle$. The rotations $R_{\hat{y}}(\pi/2)$and $R_{\hat{y}}(-\pi/2)$ are square roots of the first, and $R_{\hat{x}}(\pi/2)$ and $R_{\hat{x}}(-\pi/2)$ of the second. Applying these four different rotations to $|0\rangle$, we obtain the states $|+\rangle$, $|-\rangle$, $|+i\rangle$, and $|-i\rangle$, respectively. It follows that the square root transformations provide an alternative way of defining these coherent superpositions of $|0\rangle$ and $|1\rangle$.

This suggests that an equal superposition of the epistemic states $1\vee 2$ and $3\vee 4$ is obtained by acting on $1\vee 2$ with the square root of a permutation that maps $1\vee 2$ to $3\vee 4$. The permutations (14)(23) and (13)(24) both map $1\vee 2$ to $3\vee 4$ [these are the analogues of $R_{\hat{y}}(\pi)$, and $R_{\hat{x}}(\pi)$]. The permutation (14)(23) has two roots, namely, (1342) and (1243), as does the permutation (13)(24), namely, (1234) and (1432). [The graphical depictions of (13)(24) and (1234) in Fig. 2 illustrate this fact.] Acting these four permutations on $1\vee 2$ yields the four epistemic states $1\vee 3$, $2\vee 4$, $2\vee 3$, and $1\vee 4$ (which are precisely the analogues of $|+\rangle$, $|-\rangle$, $|+i\rangle$, and $|-i\rangle$).

Now consider how this works for an arbitrary pair of disjoint epistemic states, $a\vee b$ and $c\vee d$, where we conventionally assume that $a<b$ and $c<d$. The permutations that take $a\vee b$ to $c\vee d$ are $(ad)(bc)$ and $(ac)(bd)$. Each of these has a pair of square roots yielding in total the four permutations $(acdb)$, $(abdc)$, $(abcd)$, and $(adcb)$. These define four coherent binary operations on the epistemic states that are analogous, respectively, to equal superposition with relative phases 0, π, $\pi/2$, and $3\pi/2$. Acting them on $a\vee b$, we find

$$(acdb)(a\vee b)=a\vee c=(a\vee b)+_1(c\vee d),$$

$$(abdc)(a\vee b)=b\vee d=(a\vee b)+_2(c\vee d),$$

$$(abcd)(a\vee b)=b\vee c=(a\vee b)+_3(c\vee d),$$

$$(adcb)(a\vee b)=a\vee d=(a\vee b)+_4(c\vee d), \tag{38}$$

where the final equality follows from the definitions of the binary operations $+_1$, $+_2$, $+_3$, and $+_4$, presented in Eq. (10). This confirms our previous interpretation of these operations [see Eq. (35)] and reinforces our claim that there is a good analogue of coherent superposition in this toy theory.

C. No universal state inverter

Given the aspects of the toy theory developed so far, we can already demonstrate an analogy to a characteristically quantum phenomenon, namely, the impossibility of building a universal state inverter. For a single qubit, a universal state inverter is a device that deterministically maps every pure quantum state to the orthogonal quantum state, that is,

$$|\psi\rangle \rightarrow |\bar{\psi}\rangle \quad \text{for all } |\psi\rangle, \tag{39}$$

where $\langle\psi|\bar{\psi}\rangle = 0$. Such a map cannot be physically implemented because it is not unitary [42].

The analogous task in the toy theory is to deterministically map every pure epistemic state of an elementary system to the one that is disjoint with it. Thus, we require

$$1 \vee 2 \leftrightarrow 3 \vee 4,$$

$$1 \vee 3 \leftrightarrow 2 \vee 4,$$

$$2 \vee 3 \leftrightarrow 1 \vee 4. \tag{40}$$

But this transformation is impossible since it does not correspond to any permutation of the ontic states; the first two conditions together imply that $1 \leftrightarrow 4$ and $2 \leftrightarrow 3$, which is in contradiction with the third condition.

The impossibility of universal state inversion in both quantum theory and the toy theory can also be seen by noting that it would appear as an inversion about the origin in the Bloch ball representation, and such an inversion cannot be achieved by any rotation, nor by any combination of the rotations and reflections that are allowed in the toy theory.

D. Measurements

We now turn to the nature of measurements in the toy theory. We shall here consider only measurements that are *reproducible* in the sense that if repeated upon the same system, they yield the same outcome. For this to be possible, the epistemic state after the measurement must rule out all of the ontic states that are not consistent with the outcome (otherwise, the epistemic state would not reflect the fact that a different outcome cannot occur upon repetition).

The knowledge balance principle imposes restrictions on the sort of reproducible measurement that can be implemented. Again, we start by ruling out a certain kind of measurement, namely one which identifies whether or not the ontic state is in a singleton set. To be specific, consider the measurement which determines whether the ontic state is 1 or not. The "not 1" outcome identifies the ontic state as being either 2 or 3 or 4. Now, if in this measurement the outcome 1 occurs (and nothing prevents it from occurring when the initial epistemic state deems it to be possible), then by virtue of the assumed reproducibility of measurements, the epistemic state after the measurement must rule out the ontic states 2, 3, and 4. But this would mean that after the measurement one would be certain that the ontic state was 1, and such a state of knowledge violates the knowledge balance principle. Thus, the measurement considered is not allowed.

Clearly, the fewest ontic states that can be associated with a single outcome of a measurement is two. Thus, the only valid reproducible measurements are those which partition the four ontic states into two sets of two ontic states. There are only three such partitionings:

$$\{1 \vee 2, 3 \vee 4\},$$

$$\{1 \vee 3, 2 \vee 4\},$$

$$\{1 \vee 4, 2 \vee 3\}. \tag{41}$$

In our pictorial representation, we can represent these as

I	I	II	II

I	II	I	II

I	II	II	I

(42)

where in each case the two sets are distinguished by a roman numeral. These three partitionings are analogous to the following three bases in quantum theory:

$$\{1 \vee 2, 3 \vee 4\} \Leftrightarrow \{|0\rangle, |1\rangle\},$$

$$\{1 \vee 3, 2 \vee 4\} \Leftrightarrow \{|+\rangle, |-\rangle\},$$

$$\{1 \vee 4, 2 \vee 3\} \Leftrightarrow \{|+i\rangle, |-i\rangle\}. \tag{43}$$

We call the set of ontic states associated with a particular outcome the *ontic support* of that outcome. If the initial epistemic state has its ontic support inside the ontic support of a particular outcome, then that outcome is certain to occur, otherwise, the outcome is not determined by the initial epistemic state. For instance, suppose the epistemic state is $1 \vee 2$, so that graphically we have

(44)

If one performs the measurement that distinguishes $1 \vee 2$ from $3 \vee 4$, depicted

I	I	II	II

(45)

then the first outcome is certain to occur. On the other hand, if one performs the measurement that distinguishes $1 \vee 3$ from $2 \vee 4$, depicted

I	II	I	II

(46)

then the outcome is not determined. Nonetheless, one can say *something* in this case, namely, that in a large ensemble of such experiments, one expects the two outcomes to occur with equal frequency.[2] This is analogous to what occurs in quantum theory: if the initial quantum state is one of the elements of the orthogonal basis associated with the measurement, then the outcome associated with that element is certain to occur, while if it is not, only the expected relative frequencies of the outcomes are determined by the quantum state.

[2]This presumes that the relative frequency of different ontic states in the ensemble is equivalent to the probability distribution defined by the epistemic state. This assumption can be questioned. See, for instance, the work of Valentini [10].

E. Measurement update rule

Suppose the initial epistemic state is $1 \vee 2$, a reproducible measurement of $1 \vee 3$ versus $2 \vee 4$ is performed, and the outcome $1 \vee 3$ occurs. In this case, one can retrodict that the ontic state of the system must have been 1 prior to the measurement. This is not in conflict with the knowledge balance principle since the latter does not place restrictions on what one can know, at a given time, about the ontic state at an earlier time. The principle *does*, however, place restrictions on what one can know, at a given time, about the ontic state at that time. If it were the case that the system's ontic state was known to be unaltered in the process of measurement, then one's description of the system prior to the measurement would apply also after the measurement. But then, one would know the system to be in the ontic state 1 after the measurement, and this *is* in violation of the knowledge balance principle. Since we assume that information gain through measurements is always possible, we must conclude that measurement causes an unknown disturbance to the ontic state of the system.

In our particular example, the assumption that the measurement is reproducible implies that the epistemic state after the measurement must rule out the ontic states 2 and 4. Thus, the only final epistemic state that makes the measurement result reproducible and abides by the knowledge balance principle is $1 \vee 3$.

It follows that the nature of the unknown disturbance must be such that although one knows that the ontic state that applied prior to the measurement was 1, all one knows about the ontic state that applies after the measurement is that it is 1 or 3. Thus, the unknown disturbance must ensure that

$$1 \rightarrow 1 \vee 3. \tag{47}$$

Similarly, if the initial epistemic state was $3 \vee 4$ and a measurement of $1 \vee 3$ versus $2 \vee 4$ found the outcome $1 \vee 3$, one could infer that prior to the measurement, the ontic state must have been 3. However, in order to have reproducibility and to abide by the knowledge balance principle, it must be the case that after the measurement, the ontic state is only know to be 1 or 3. Thus, the unknown disturbance must ensure that

$$3 \rightarrow 1 \vee 3. \tag{48}$$

These two conditions can be satisfied by assuming that the measurement induces either the identity permutation (1)(3) or the two-cycle permutation (13) on the ontic states that are consistent with the outcome, but *which* of these two permutations occurs is not known. For instance, if the ontic state was 1, then either it remains 1 or it becomes 3, and all that can be said of the ontic state that applies after the measurement is that it is 1 or 3. (Note that one need not specify what occurs to the ontic states 2 and 4 since these have been ruled out by the outcome of the measurement.)

This is generalized as follows. In a measurement of $a \vee b$ versus $c \vee d$, if the outcome $a \vee b$ occurs, then either the identity permutation (a)(b) occurs (i.e., nothing happens to the system) or the two-cycle permutation (ab) occurs (if the ontic state is a, it becomes b and vice versa), but it is unknown which.

Note that the possible permutations resulting from a measurement depend only on the identity and outcome of the measurement and not on the initial epistemic state. This is appropriate, since the nature of someone's knowledge of a system should not influence how the ontic state of the system evolves during a measurement. By the same token, whether or not the system is initially correlated with other systems should not influence the nature of the evolution of the ontic state of the system during a measurement, because the presence or absence of such correlation is a feature of an observer's knowledge of the system, not a property of the system itself.[3] Thus, although we have derived the nature of the unknown disturbance by considering an example where the system being measured is not correlated with any other system, the results obtained must also be applicable when such correlation is present. We will therefore make use of the results derived above when we consider measurements on one member of a pair of systems in Secs. IV A, IV B, and IV G.

In the case considered here, where the system of interest is uncorrelated with all other systems, the nature of the transformation of the ontic states for reproducible maximally informative measurements implies a particularly simple rule for updating the epistemic state. The final epistemic state has ontic support equal to the ontic support of the outcome obtained in the measurement. This is analogous to the update rule for a reproducible maximally informative measurement in quantum theory, where the final quantum state is simply the eigenvector associated with the outcome obtained in the measurement.

We now consider a few more quantum phenomena for which we can provide an analogue in the toy theory.

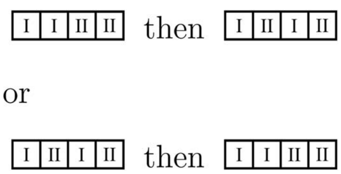
F. Noncommutativity of measurements

In quantum theory, the order in which measurements occur is important for the outcome that is obtained in these measurements. For instance, implementing a reproducible measurement of the basis $\{|0\rangle, |1\rangle\}$ followed by a reproducible measurement of the basis $\{|+\rangle, |-\rangle\}$ in general has different results from the case where they are implemented in the opposite order. Specifically, if the quantum state is $|0\rangle$ initially, then if the measurement of $\{|0\rangle, |1\rangle\}$ comes first, it will yield the outcome $|0\rangle$ with certainty. On the other hand, if it comes second, then the outcomes $|0\rangle$ and $|1\rangle$ will occur with equal probability. The reason is that the intervening measurement of $\{|+\rangle, |-\rangle\}$ collapses the quantum state to $|+\rangle$ or $|-\rangle$, and the latter states make the outcome of $\{|0\rangle, |1\rangle\}$ completely unpredictable.

Similarly, the order in which measurements occur in the toy theory also has a bearing on the outcomes obtained. In-

[3]The only way in which the initial epistemic state could influence the evolution of the ontic state is if there was a physical influence exerted by the mental state of the observer on the physical system. In our derivation of the toy theory, we are explicitly rejecting this sort of possibililty.

deed, the example just provided has a perfect analogue in the toy theory. We consider implementing a reproducible measurement associated with the partitioning $\{1\vee 2, 3\vee 4\}$ followed by the reproducible measurement associated with the partitioning $\{1\vee 3, 2\vee 4\}$, and the same measurements in reverse order:



(49)

Suppose that initially the epistemic state is $1\vee 2$,

(50)

If the measurement of $\{1\vee 2, 3\vee 4\}$ comes first, it will yield the outcome $1\vee 2$ with certainty. On the other hand, if it comes second, then the outcomes $1\vee 2$ and $3\vee 4$ will occur with equal frequency. The reason is that the measurement of $\{1\vee 3, 2\vee 4\}$ causes the epistemic state to be updated to $1\vee 3$ or $2\vee 4$, and each of these epistemic states makes the outcome of $\{1\vee 2, 3\vee 4\}$ completely unpredictable.

G. Interference

Another quantum phenomenon that the toy theory reproduces qualitatively is interference. We offer the following paradigmatic example of interference in quantum theory. Consider the following three experiments:

(a) Prepare $|0\rangle$, then measure $\{|+\rangle, |-\rangle\}$.

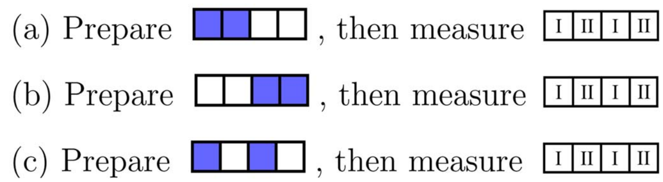
(b) Prepare $|1\rangle$, then measure $\{|+\rangle, |-\rangle\}$.

(c) Prepare $\sqrt{2}^{-1}(|0\rangle+|1\rangle)$, then measure $\{|+\rangle, |-\rangle\}$.

The probability distribution over the outcomes is $(1/2, 1/2)$ for (a), $(1/2, 1/2)$ for (b), and $(1, 0)$ for (c). The probability zero for the outcome $|-\rangle$ in case (c) is, of course, a result of the destructive interference between the amplitude for this outcome in states $|0\rangle$ and $|1\rangle$.

Interference is often cited as evidence against the epistemic view of quantum states. The argument runs as follows. If quantum states are associated with probability distributions over some hidden reality, then the only way one could possibly understand a coherent superposition of quantum states (so the argument goes) is as a convex combination of the associated probability distributions with weights given by the amplitudes squared. In particular, the distribution associated with the state $\sqrt{2}^{-1}(|0\rangle+|1\rangle)$ must be a convex sum, with equal weights, of the distributions associated with $|0\rangle$ and $|1\rangle$. But given that in a measurement of $\{|+\rangle, |-\rangle\}$ the $|-\rangle$ outcome occurs with probability $1/2$ for both $|0\rangle$ and $|1\rangle$, if $\sqrt{2}^{-1}(|0\rangle+|1\rangle)$ corresponded to a convex sum of these possibilities, one would still expect the $|-\rangle$ outcome to occur with probability $1/2$, not probability zero.

All this argument demonstrates, however, is a lack of imagination concerning the interpretation of coherent superposition within an epistemic view. We have already seen in Sec. III A how in the toy theory one can define some binary operations that are distinct from convex combination. The possibility of representing coherent superposition and convex combination differently within an epistemic view is what makes interference understandable. This is made clear through the toy theory version of the interference experiment discussed above. Recall from Eq. (35) of Sec. III A that the toy theory analogue of the coherent superposition $\sqrt{2}^{-1}(|0\rangle+|1\rangle)$ is $(1\vee 2)+_1(3\vee 4)$ which is simply $1\vee 3$. This is *not* the equally weighted probabilistic sum of the two epistemic states, which would be the epistemic state $1\vee 2\vee 3\vee 4$. Thus, the analogue of the three experiments are as follows:



It is straightforward to see that the probability distributions over the outcomes are $(1/2, 1/2)$ for (a), $(1/2, 1/2)$ for (b), and $(1, 0)$ for (c). Thus, the empirical signature of interference is reproduced.

Interference phenomena have led interpreters of quantum theory to conclude that whatever an equally weighted coherent superposition of two possibilities might be, it is not the "or" of those possibilities nor the "and" of those possibilities. This is certainly the case in the toy theory. The coherent combination of a pair of disjoint pure epistemic states is neither the "or" nor the "and" of those states, but rather a sampling of the ontic states from each.

IV. PAIRS OF ELEMENTARY SYSTEMS

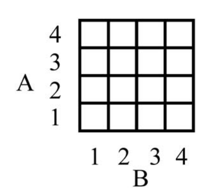
A. Epistemic states

The simplest composite system is a pair of elementary systems. Since each elementary system has four ontic states, the pair has sixteen ontic states. We can represent the ontic states of the pair by conjunctions of the possible ontic states of the constituents. Representing conjunction by "·" (read as "and"), the sixteen possibilities are

$$1\cdot 1,\ 1\cdot 2,\ 1\cdot 3,\ 1\cdot 4,\ 2\cdot 1,\ 2\cdot 2,\ 2\cdot 3,\ 2\cdot 4,$$

$$3\cdot 1,\ 3\cdot 2,\ 3\cdot 3,\ 3\cdot 4,\ 4\cdot 1,\ 4\cdot 2,\ 4\cdot 3,\ 4\cdot 4. \quad (51)$$

We can represent these graphically by a 4×4 array of boxes, where the rows represent the different ontic states of system A, and the columns represent the different ontic states of system B. Specifically, we take the box in the jth row from the bottom and kth column from the left to represent the ontic state $j\cdot k$.

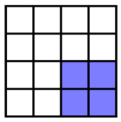


(52)

Since each system has two questions in a canonical set, the pair has four questions in a canonical set. The knowledge balance principle ensures that only two of these four questions may be answered in a state of maximal knowledge. This corresponds to knowing the ontic state to be among four of the sixteen possibilities. The pure epistemic states are therefore disjunctions of four ontic states, for instance,

$$(1 \cdot 3) \vee (1 \cdot 4) \vee (2 \cdot 3) \vee (2 \cdot 4), \tag{53}$$

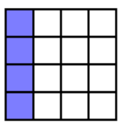
which can be represented graphically by



(54)

where we have dropped the labels on the rows and columns for convenience.

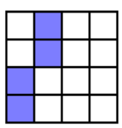
By applying the knowledge balance principle to each of the systems in the pair individually, we obtain a further constraint: at most a single question can be answered about the ontic state of each of the systems. Thus, an epistemic state for AB of the form



(55)

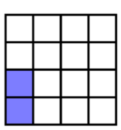
although satisfying the principle as it applies to the composite AB, violates the principle as it applies to the system B because the ontic state of B is known to be 1.

The epistemic state for AB of the form



(56)

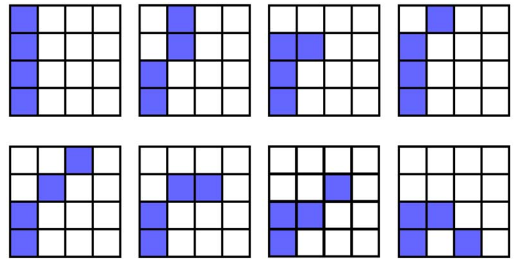
is also ruled out by application of the principle to the individual systems. Here, it is not the marginals that are the problem. Rather, the problem is that a reproducible measurement of $1 \vee 2$ versus $3 \vee 4$ on A, which has outcome $1 \vee 2$ for instance, allows one to rule out 3 and 4 as possibilities for the ontic states of A after the measurement, and, as established earlier, causes the ontic state of A to undergo an unknown permutation: either (1)(2) or (12). However, this leaves the final epistemic state of AB as



(57)

which corresponds to more knowledge about AB than is allowed by the principle. We have here made use of the assumption that the transformation that applies to A is the same whether A is correlated with B or not, since correlation is a feature of an observer's knowledge and therefore cannot determine the nature of the physical transformation.

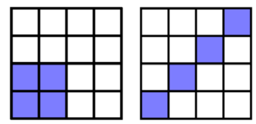
The full set of epistemic states that violate the knowledge balance principle in some way are



(58)

together with any epistemic state that can be obtained from one of these by a permutation of A and B, or by a permutation of the rows among themselves or by a permutation of the columns among themselves or by any combination of these operations.

It follows that the valid states of maximal knowledge for a pair of systems are of two types, represented as



(59)

together with those that can be obtained from the operations just described. These are analogous in quantum theory to product states and maximally entangled states, respectively.

Epistemic states for composite systems can be classified according to whether they describe correlations between the systems or not. An epistemic state is said to describe correlations between a pair of systems if some form of knowledge acquisition about one of the systems leads the bearer of this epistemic state to refine their description of the other system. It is clear that by these lights epistemic states of the first type are uncorrelated while those of the second type are correlated.

The general form of the first type of epistemic state is

$$(a \vee b) \cdot (c \vee d), \tag{60}$$

where $a,b,c,d \in \{1,2,3,4\}$ and $a \neq b$, $c \neq d$. These states are a conjunction of states of maximal knowledge for each of the systems, and thus satisfy the principle as it applies to the subsystems. Note that one can distribute the conjunction over the disjunction to rewrite the epistemic state as

$$(a \cdot c) \vee (a \cdot d) \vee (b \cdot c) \vee (b \cdot d), \tag{61}$$

verifying that it is a disjunction of four ontic states and thus satisfies the principle as it applies to the pair. Some examples of uncorrelated epistemic states are

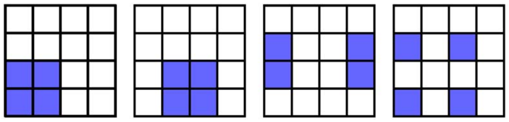
$$(1 \vee 2) \cdot (1 \vee 2),$$

$$(1 \vee 2) \cdot (2 \vee 3),$$

$$(2 \vee 3) \cdot (1 \vee 4),$$

$$(1 \vee 3) \cdot (1 \vee 3), \tag{62}$$

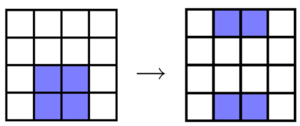
which are represented graphically as



(63)

By Eq. (5), these are analogous to the product states $|0\rangle|0\rangle$, $|0\rangle|+i\rangle$, $|+i\rangle|-i\rangle$, and $|+\rangle|+\rangle$, respectively.

Since such an epistemic state is simply a "product" of the marginals for A and B, when a measurement on A is implemented, only the marginal for A is updated, and this occurs in precisely the manner described in Sec. III E. For instance, if the epistemic state for the composite is $(1\vee 2)\cdot(2\vee 3)$, and a measurement of $1\vee 4$ versus $2\vee 3$ on system A finds the outcome $1\vee 4$, the final state is $(1\vee 4)\cdot(2\vee 3)$,



(64)

The general form of the second type of allowed epistemic state is

$$(a\cdot e)\vee(b\cdot f)\vee(c\cdot g)\vee(d\cdot h), \tag{65}$$

where $a,b,c,d,e,f,g,h\in\{1,2,3,4\}$ and $a\neq b\neq c\neq d$, $e\neq f\neq g\neq h$. Note that the marginal epistemic states for A and B are $1\vee 2\vee 3\vee 4$. Examples of such states are

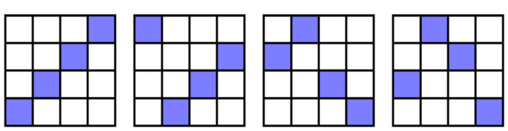
$$(1\cdot 1)\vee(2\cdot 2)\vee(3\cdot 3)\vee(4\cdot 4),$$

$$(1\cdot 2)\vee(2\cdot 3)\vee(3\cdot 4)\vee(4\cdot 1),$$

$$(1\cdot 4)\vee(2\cdot 3)\vee(3\cdot 1)\vee(4\cdot 2),$$

$$(1\cdot 4)\vee(2\cdot 1)\vee(3\cdot 3)\vee(4\cdot 2), \tag{66}$$

which are depicted as



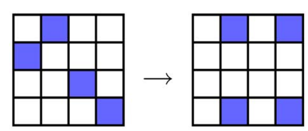
(67)

For such epistemic states, nothing is known about the ontic states of the individual systems, but everything is known about the relation between them. In the first example, for instance, the two systems are known to be in the same ontic state. In the second example, the ontic state of B has an index that is one greater (modulo 4) than the ontic state of A. In other words, the ontic state of B is related to the ontic state of A by the permutation (1234). In the third and fourth examples, the permutations are (1423) and (142)(3), respectively. There is an epistemic state of this sort for every permutation of the four ontic states, and thus 24 in all. These are represented graphically by the 24 ways of filling only one box in every row and column.

The following picture emerges. Unlike in classical theories, wherein one can know the relation between two systems completely *and* know their individual ontic states, in the toy theory we have a trade-off. In a state of maximal knowledge, *either* one can know as much as is possible to know about the individual ontic states of a pair of systems, in which case one has an answer to a single question about each, yielding an uncorrelated epistemic state, *or* one can know as much as is possible to know about the relation between the two systems, in which case one knows the answers to two questions about their relation, yielding a correlated epistemic state. It has been argued by Brukner, Zukowski and Zeilinger [43] (within the context of a different interpretational approach) that this sort of account captures the essence of entanglement.

It is worth noting that epistemic states of the second type are not only correlated, they are *perfectly* correlated, that is, for *any* form of knowledge acquisition about one of the systems, the description of the other is refined. Further on, we shall consider epistemic states describing imperfect correlations, for instance, in Eq. (80) of Sec. IV C.

It is useful to examine in detail how a perfectly correlated epistemic state is updated if a reproducible measurement is implemented on one of the subsystems. We describe this for a generic epistemic state of the form given in Eq. (65), and a generic measurement which distinguishes $a\vee b$ from $c\vee d$. Upon obtaining the outcome $a\vee b$, the ontic states $c\cdot g$ and $d\cdot h$ for the composite are ruled out. Thus one immediately sees that the marginal for B after the measurement will be $e\vee f$. Moreover, as discussed in Sec. III E, the measurement causes system A to undergo an unknown permutation, namely, (a)(b) or (ab). The first case yields $a\cdot e$ and $b\cdot f$ as possible final ontic states of the composite, while the second case yields $b\cdot e$ and $a\cdot f$. The final epistemic state is therefore the disjunction of these four possibilities, which is simply $(a\vee b)\cdot(e\vee f)$. As an example, if the epistemic state for AB is initially $(1\cdot 4)\vee(2\cdot 3)\vee(3\cdot 1)\vee(4\cdot 2)$ and a measurement of $2\vee 3$ versus $1\vee 4$ on A finds the outcome $1\vee 4$, the epistemic state is updated to $(1\vee 4)\cdot(2\vee 4)$,



(68)

The marginal for B is updated from $1\vee 2\vee 3\vee 4$ to $2\vee 4$, so there has been a refinement of the description of B as a result of the measurement on A.

B. Remote steering

"Steering" is the name given by Schrödinger to the phenomenon that lies at the heart of the Einstein-Podolsky-Rosen argument for the incompleteness of quantum theory [15]. We shall present the phenomenon in a manner that is closer to the account given by Einstein in his correspondence with Schrödinger [14] than to the account found in the EPR paper. Alice and Bob each hold a qubit, denoted A and B, respectively, and the pair AB is described by the quantum state $\frac{1}{\sqrt{2}}(|0\rangle|0\rangle+|1\rangle|1\rangle)$. Suppose Alice chooses to measure the $\{|0\rangle,|1\rangle\}$ basis (in a reproducible way) on system A. In this case, with probability 1/2 she obtains the outcome $|0\rangle$ and (following the standard collapse rule) she updates the quantum state of the pair to $|0\rangle|0\rangle$,

$$\frac{1}{\sqrt{2}}(|0\rangle|0\rangle + |1\rangle|1\rangle) \rightarrow |0\rangle|0\rangle, \tag{69}$$

while with probability 1/2, she obtains the outcome $|1\rangle$ and updates the quantum state of the pair to $|1\rangle|1\rangle$,

$$\frac{1}{\sqrt{2}}(|0\rangle|0\rangle + |1\rangle|1\rangle) \rightarrow |1\rangle|1\rangle. \tag{70}$$

On the other hand, if Alice chooses to measure the $\{|+\rangle,|-\rangle\}$ basis on system A, then with probability 1/2 she obtains the outcome $|+\rangle$ and updates the quantum state of the pair to $|+\rangle|+\rangle$,

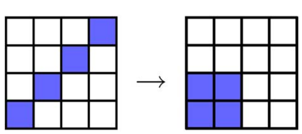
$$\frac{1}{\sqrt{2}}(|0\rangle|0\rangle + |1\rangle|1\rangle) \rightarrow |+\rangle|+\rangle, \tag{71}$$

and with probability 1/2 she obtains the outcome $|-\rangle$ and updates the quantum state of the pair to $|-\rangle|-\rangle$,

$$\frac{1}{\sqrt{2}}(|0\rangle|0\rangle + |1\rangle|1\rangle) \rightarrow |-\rangle|-\rangle. \tag{72}$$

Note that for one choice of Alice's measurement, the final quantum state for B is either $|0\rangle$ or $|1\rangle$ whereas for the other choice, it is either $|+\rangle$ or $|-\rangle$. In a 1935 paper discussing this phenomenon, Schrödinger remarks (Ref. [44], p. 555): "It is rather discomforting that the theory should allow a system to be steered or piloted into one or the other type of state at the experimenter's mercy in spite of his having no access to it." Indeed, if the quantum state is interpreted as a state of reality, so that $|0\rangle$, $|1\rangle$, $|+\rangle$, and $|-\rangle$ are mutually exclusive states of reality, then Alice's choice of measurement can directly influence the reality in Bob's laboratory. If the collapse occurs instantaneously, as is generally assumed, this would correspond to a nonlocal influence. To be precise, it would lead to a failure of local causality, in the sense defined by Bell [45].

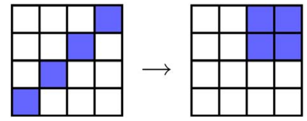
However, this example of the steering phenomenon does not imply a failure of local causality if one adopts an epistemic view of quantum states.[4] Indeed, we now show that the particular example of steering described above has a precise analogue in the toy theory despite the fact that the latter is explicitly local. Here is how it works. Alice and Bob each hold an elementary system, denoted A and B, respectively, and Alice's epistemic state for the pair AB is $(1\cdot 1)\vee(2\cdot 2)\vee(3\cdot 3)\vee(4\cdot 4)$. Suppose Alice implements the reproducible measurement on A that distinguishes $1\vee 2$ from $3\vee 4$. With probability 1/2 she obtains the outcome $1\vee 2$ and, given the results of the previous section, she must update her state of knowledge to $(1\vee 2)\cdot(1\vee 2)$. Graphically,



(73)

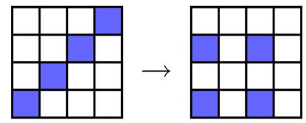
[4]Of course, a failure of locality *is* implied by correlations that violate Bell's inequalities [33], and consequently there is nothing analogous to such correlations in the toy theory. This will be discussed in Sec. VII.

On the other hand, if the outcome $3\vee 4$ occurs then Alice updates her epistemic state for the pair to $(3\vee 4)\cdot(3\vee 4)$,



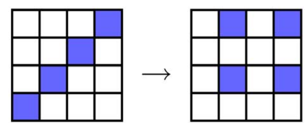
(74)

Alice could also choose to implement the measurement that distinguishes $1\vee 3$ from $2\vee 4$. She again obtains both outcomes with equal probability. If the outcome is $1\vee 3$, she updates her epistemic state for the pair to $(1\vee 3)\cdot(1\vee 3)$,



(75)

while if the outcome is $2\vee 4$, she updates her epistemic state to $(2\vee 4)\cdot(2\vee 4)$,



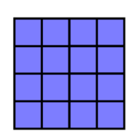
(76)

Note that the right-hand sides of Eqs. (73)–(76) are precisely analogous to those of Eqs. (69)–(72) under the mapping of Eq. (5).

The important point to note about the steering phenomenon in the toy theory is that the choice of measurement at A does not change the ontic state at B. The measurement *does* sometimes lead to a disturbance, but this is a disturbance to the ontic state of A. The only change associated with B is to Alice's *knowledge* of B. Suppose, for instance, that the ontic state of AB was initially $1\cdot 1$. Alice only knows that it is $(1\cdot 1)$ or $(2\cdot 2)$ or $(3\cdot 3)$ or $(4\cdot 4)$, and therefore initially assigns the marginal $1\vee 2\vee 3\vee 4$ to B. If she measures $1\vee 2$ versus $3\vee 4$ on A, she will obtain the outcome $1\vee 2$ (by virtue of A being in ontic state 1), and she will update her marginal for B to $1\vee 2$. If, on the other hand, she measures $1\vee 3$ versus $2\vee 4$, then she will obtain the outcome $1\vee 3$ (by virtue of A being in ontic state 1), and she will update her marginal for B to $1\vee 3$. In both cases, B remains in the ontic state 1 throughout. Alice has simply narrowed down the possibilities in two different ways.

C. Epistemic states of nonmaximal knowledge

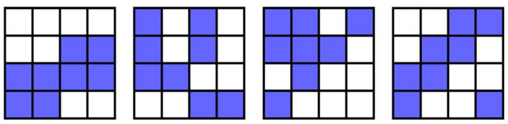
One way to have nonmaximal knowledge of a pair of systems is to know nothing about their ontic state. This corresponds to the epistemic state $(1\vee 2\vee 3\vee 4)\cdot(1\vee 2\vee 3\vee 4)$, depicted as



(77)

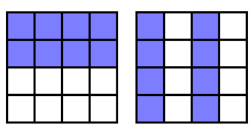
It is analogous to the completely mixed state for two qubits, $I/2\otimes I/2$.

In the case of a single elementary system, we found that knowing nothing was the *only* way to have nonmaximal knowledge. In the case of two elementary systems, however, there are other possibilities. Since there are four questions in the canonical set, one could know the answer to just one of these, rather than two or none. This corresponds to an ontic support with eight elements. These epistemic states are also highly constrained by the knowledge balance principle. Their marginals must be valid epistemic states for the individual subsystems, and they must be mapped to valid epistemic states under the update rule for measurements on one of the subsystems. Some examples of epistemic states of nonmaximal knowledge that contain eight ontic states but still violate the principle in some way are



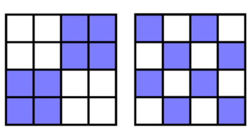
(78)

The epistemic states of nonmaximal knowledge that abide by the principle are again found to be of two types. The first type is essentially a conjunction of a pure epistemic state for one system and a mixed epistemic state for the other. Examples are $(3\vee4)\cdot(1\vee2\vee3\vee4)$ and $(1\vee2\vee3\vee4)\cdot(1\vee3)$, which are graphically depicted as



(79)

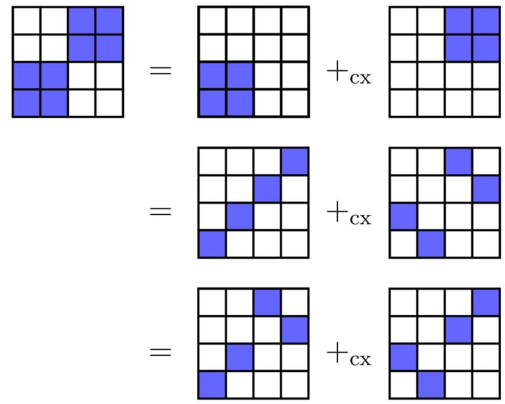
and which are analogous to the density operators $|1\rangle\langle1|\otimes I/2$ and $I/2\otimes|+\rangle\langle+|$, respectively. These are uncorrelated states. The second type of state is more interesting. Examples are $[(1\vee2)\cdot(1\vee2)]\vee[(3\vee4)\cdot(3\vee4)]$ and $[(1\vee3)\cdot(2\vee4)]\vee[(2\vee4)\cdot(1\vee3)]$, which are depicted as



(80)

which are analogous to the density operators $\frac{1}{2}|0\rangle\langle0|\otimes|0\rangle\langle0|+\frac{1}{2}|1\rangle\langle1|\otimes|1\rangle\langle1|$ and $\frac{1}{2}|+\rangle\langle+|\otimes|-\rangle\langle-|+\frac{1}{2}|-\rangle\langle-|\otimes|+\rangle\langle+|$. These are correlated states. Measurements upon one system require an update of the epistemic state of the other. For instance, if the initial epistemic state is $[(1\vee2)\cdot(1\vee2)]\vee[(3\vee4)\cdot(3\vee4)]$ and a measurement of $1\vee2$ versus $3\vee4$ is implemented on system A, then the final epistemic state of the pair is $(1\vee2)\cdot(1\vee2)$ if the outcome $1\vee2$ is obtained, and $(3\vee4)\cdot(3\vee4)$ if the outcome $3\vee4$ is obtained. Note however that other measurements, such as a measurement of $1\vee3$ versus $2\vee4$, do not lead to an update of the marginal of the nonmeasured system. Thus, the correlation is not perfect, in the sense defined in Sec. IV A. The same sort of thing occurs for the density operator $\frac{1}{2}|0\rangle\langle0|\otimes+\frac{1}{2}|1\rangle\langle1|\otimes|1\rangle\langle1|$. There is correlation for measurements in the $\{|0\rangle,|1\rangle\}$ basis, but none for measurements in the $\{|+\rangle,|-\rangle\}$ basis. The existence of a distinction between epistemic states exhibiting perfect correlations and those exhibiting imperfect correlations is analogous to the existence of a distinction, in quantum theory, between states that are said to be quantum correlated, or entangled, and those that are said to be merely *classically* correlated.[5]

Note that states of nonmaximal knowledge are mixed states. Indeed, they may be viewed as convex combinations of pure states in several different ways. For instance,

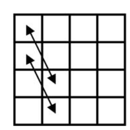


(81)

Coherent binary operations on the pure epistemic states of a pair of systems could also be defined, but we shall not do so here. Note that our definitions of disjointness and compatibility and of the fidelity between epistemic states, presented in the context of a single elementary system in Sec. III A, are applicable for composite systems as well.

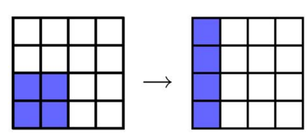
D. Transformations

The transformations that can be performed upon a pair of systems is a subset of the set of permutations of the sixteen ontic states. It is a subset because some permutations take valid epistemic states to invalid ones. For instance, the permutation



(82)

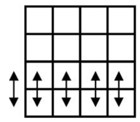
is invalid because it leads to the transition



(83)

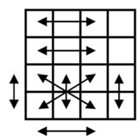
Independent permutations of each subsystem's ontic states are among the subset of allowed permutations of the composite's ontic states. For instance, the permutation (12)(3)(4) on system A yields

[5]Again, this is not to say that perfect correlations in the toy theory have *all* the features of quantum correlations. In particular, they do not violate any Bell inequality.



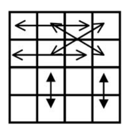
(84)

and the permutation (12)(3)(4) on A and (13)(2)(4) on B yields



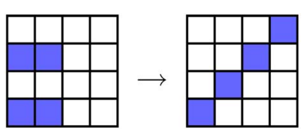
(85)

Clearly, such local permutations cannot change the degree of correlation between the systems: uncorrelated states are transformed into uncorrelated states and correlated states are transformed into correlated states. These permutations are analogous to local unitary operations in quantum theory, which do not change the degree of entanglement. Other permutations *can* alter the degree of correlation, and are thus analogous to entangling operations in quantum theory. One such permutation is



(86)

which yields the transition



(87)

It is analogous to the controlled-NOT operation for a pair qubits [39].

E. No cloning

Given the nature of transformations for a pair of elementary systems, it is possible to prove the existence of a no-cloning theorem. We begin by reviewing this theorem in the context of the ontic view of quantum states [46,47]. A cloning process for a set of states $\{|\psi_i\rangle\}$ is defined as a transformation satisfying

$$|\psi_i\rangle|\chi\rangle \rightarrow |\psi_i\rangle|\psi_i\rangle \tag{88}$$

for all $|\psi_i\rangle$, where $|\chi\rangle$ is an arbitrary fixed state. The idea is that the quantum state of A is unknown and the goal is to implement a transformation that leaves system B in this unknown state.

The simplest case to consider is when the set contains two states $\{|\psi_1\rangle,|\psi_2\rangle\}$. If $|\psi_1\rangle$ and $|\psi_2\rangle$ are nonorthogonal states, then the cloning process is impossible because it does not preserve inner products, and so cannot be a unitary map. For instance, a cloning process for the set $\{|1\rangle,|+\rangle\}$ satisfies

$$|1\rangle|0\rangle \rightarrow |1\rangle|1\rangle,$$

$$|+\rangle|0\rangle \rightarrow |+\rangle|+\rangle, \tag{89}$$

where we have taken the arbitrary initial state of system B to be $|0\rangle$. The inner product squared between the two possible initial states is $|\langle 1|+\rangle\langle 0|0\rangle|^2=1/2$, while the inner product squared between the two possible final states is $|\langle 1|+\rangle \times\langle 1|+\rangle|^2=1/4$. (Note that allowing irreversible operations does not help since these can only increase, not *decrease*, the fidelity between the input states. [39])

If one adopts an epistemic view of quantum states, then the question of whether cloning is possible or not is a question of whether the epistemic state that pertains to one system can be made to be also applicable to another without creating correlations between the two systems. It is *not* the question of whether the *ontic* state of a system can be duplicated in another (which would create correlations). Here is the manner in which it is defined in the toy theory. A cloning process for a set of epistemic states $\{(a_i \vee b_i)\}_i$ is defined as a transformation satisfying

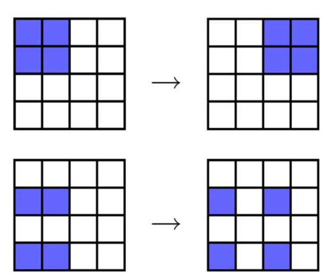
$$(a_i \vee b_i)\cdot(c \vee d) \rightarrow (a_i \vee b_i)\cdot(a_i \vee b_i) \tag{90}$$

for all epistemic states $a_i \vee b_i$ in the set, where $c \vee d$ is an arbitrary initial epistemic state for B. The cloning process cannot be implemented for nondisjoint epistemic states because it leads to a decrease in the classical fidelity (defined in Sec. III A) and because this fidelity is preserved under permutations (and is nondecreasing under any valid mixture of permutations). This is easily illustrated by an example. The analogue of the cloning process for the set $\{|1\rangle,|+\rangle\}$ is the cloning process for the set $\{3\vee 4, 1\vee 3\}$. We require

$$(3\vee 4)\cdot(1\vee 2) \rightarrow (3\vee 4)\cdot(3\vee 4),$$

$$(1\vee 3)\cdot(1\vee 2) \rightarrow (1\vee 3)\cdot(1\vee 3), \tag{91}$$

where we have taken the arbitrary initial epistemic state for B to be $1\vee 2$. Graphically, this is depicted as



(92)

Imagine that the upper and lower diagrams are superimposed on top of one another. It is then easy to see that there are two ontic states in the overlap of the two possible initial epistemic states, namely, the ontic states $3\cdot 1$ and $3\cdot 2$, whereas there is only one ontic state in the overlap of the two possible final epistemic states, namely, $3\cdot 3$. However, a permutation of the ontic states can only change the places wherein the two epistemic states overlap, not the *number* of places where they overlap. Thus, the cloning process is not a permutation and therefore is impossible in the toy theory.[6]

As it turns out, one does not actually need a restriction on knowledge to obtain a no-cloning theorem. By defining cloning in terms of epistemic states rather than ontic states, one obtains a no-cloning theorem for sets of nondisjoint epistemic states, even in classical theories [20,25]. A restriction on knowledge *is* necessary however in order to have *pure* states that are nondisjoint, which is necessary if there is to be a no-cloning theorem for *pure* states. In this sense, the toy theory is more analogous to quantum theory, vis-a-vis cloning possibilities, than any classical theory.

F. No broadcasting

Broadcasting is a process wherein one's state of knowledge about a system is duplicated in the marginals of a pair of systems while allowing that these systems might become correlated [48]. This differs from cloning insofar as the latter does not allow for such correlation. A broadcasting process for a set of density operators $\{\rho_i\}$ has the form

$$\rho_i \otimes \sigma \rightarrow W_i, \tag{93}$$

where W_i is a density operator for the composite AB that has marginals

$$\mathrm{Tr}_A(W_i) = \mathrm{Tr}_B(W_i) = \rho_i. \tag{94}$$

Broadcasting is only possible in quantum theory for a set of commuting density operators [48].

The simplest case to consider is when the set $\{\rho_i\}$ contains only pure states. Since nonorthogonal pure states do not commute, these cannot be broadcast. However, one does not need the result of Ref. [48] to see this. It follows immediately from the fact that any quantum state for AB with pure marginals is uncorrelated. That is, if ρ_i is a pure density operator, then the only way to satisfy Eq. (94) is to have $W_i = \rho_i \otimes \rho_i$. This implies that the only way to duplicate a pure state of a system in the marginals of a pair of systems is if the pair ends up in a product state. But this is simply cloning, and cloning of nonorthogonal pure states is impossible.

It may seem that the no-broadcasting theorem for pure quantum states tells us nothing that was not already contained in the no-cloning theorem. However, the former does capture something that the latter does not, namely, that pure states can never arise as the marginals of a correlated state. Although this is mathematically obvious given the formalism of quantum theory, it is a conceptually significant fact in the context of the epistemic view, since pure states are states of incomplete knowledge within the epistemic view, and it is natural to expect states of incomplete knowledge to arise as the marginals of a correlated state. Indeed, in a classical theory any state of incomplete knowledge can arise as the marginal of a correlated state, and consequently a broadcasting process exists for any set of epistemic states, even though a cloning process need not. Specifically, one can broadcast any set of epistemic states classically as follows. Measure the ontic state of system A (which can be done classically), prepare B in this ontic state, then forget the outcome of the measurement. The result is that the marginals for A and B will be whatever the initial epistemic state for A was, and the two systems will also be known to be perfectly correlated.

The fact that classically one can broadcast any set of epistemic states while quantum mechanically one can only broadcast commuting quantum states, may appear to challenge the view that quantum states are states of incomplete knowledge. However, the toy theory provides an enlightening example of how broadcasting of arbitrary sets of epistemic states may be ruled out. First note that the classical protocol for achieving broadcasting does not work in a toy theory universe since one cannot measure the ontic state of a system. The set of epistemic states in the toy theory for which it is easiest to see that there is *no* protocol that can achieve broadcasting is a set of nondisjoint pure epistemic states. The key to the proof is the fact that in the toy theory, as in quantum theory, pure epistemic states never arise as the marginals of correlated states. Consequently, when we demand that the marginals of the final epistemic state be pure we also demand that the two system be uncorrelated. This implies that a broadcasting process for pure states is simply a cloning process (as we defined it above), and as we saw in the previous section, such a process is impossible in the toy theory. This proof has the same structure as the one we provided for quantum theory. In this case, however, we can identify the conceptual underpinnings of the fact that pure epistemic states never arise as the marginals of correlated states.

Recall that the pure epistemic states in the toy theory are states of maximal knowledge. Thus, if every system is described by a pure epistemic state, one has maximal knowledge about each system. One cannot also have knowledge of the relations among the systems (that is, a correlated epistemic state), since this would exceed what is allowed by the knowledge balance principle. For example, if the marginal epistemic states for a pair of elementary systems are $a \vee b$ and $e \vee f$, respectively, then the only possible epistemic state for the pair is $(a \vee b)\cdot(e \vee f)$, which is an uncorrelated state.

Simply assuming that maximal information is incomplete is not sufficient to conclude that broadcasting of pure states will be impossible. For this, it needs to be the case that having maximal knowledge of A and maximal knowledge of B constitutes having maximal knowledge of the composite AB. The knowledge balance principle ensures that this is the case in the toy theory.

No broadcasting for *mixed* epistemic states also admits an analogue in the toy theory, but we do not consider it here.

G. Measurements

We now consider the measurements that may be performed upon a pair of systems. Every partitioning of the set of sixteen ontic states into four disjoint pure epistemic states yields a maximally informative measurement. If all of these correspond to uncorrelated epistemic states, we have a measurement such as

[6] In this case, the fidelity between the initial epistemic states is 1/2 whereas between the final epistemic states it is 1/4.

III	III	IV	IV
III	III	IV	IV
I	I	II	II
I	I	II	II

(95)

where the different roman numerals represent the different outcomes. This is simply a conjunction of a measurement upon the first system and a measurement upon the second, in this case a measurement of $\{1\vee 2, 3\vee 4\}$ on both. We can represent the measurement on the pair by the partitioning

$$\{S_I, S_{II}, S_{III}, S_{IV}\}, \tag{96}$$

where

$$S_I = (1 \vee 2) \cdot (1 \vee 2),$$

$$S_{II} = (1 \vee 2) \cdot (3 \vee 4),$$

$$S_{III} = (3 \vee 4) \cdot (1 \vee 2),$$

$$S_{IV} = (3 \vee 4) \cdot (3 \vee 4). \tag{97}$$

This is analogous to the product basis

$$\{|0\rangle|0\rangle, |0\rangle|1\rangle, |1\rangle|0\rangle, |1\rangle|1\rangle\} \tag{98}$$

in quantum theory. We call measurements of this type *product measurements.*

Other examples of product measurements can be obtained by permuting the rows and columns in the above example, for instance,

III	IV	III	IV
I	II	I	II
III	IV	III	IV
I	II	I	II

IV	III	III	IV
II	I	I	II
II	I	I	II
IV	III	III	IV

IV	III	III	IV
II	I	I	II
IV	III	III	IV
II	I	I	II

(99)

which are analogous to the bases

$$\{|+\rangle|+\rangle, |+\rangle|-\rangle, |-\rangle|+\rangle, |-\rangle|-\rangle\},$$

$$\{|+i\rangle|+i\rangle, |+i\rangle|-i\rangle, |-i\rangle|+i\rangle, |-i\rangle|-i\rangle\},$$

$$\{|+\rangle|+i\rangle, |+\rangle|-i\rangle, |-\rangle|+i\rangle, |-\rangle|-i\rangle\}, \tag{100}$$

respectively. Another form that a product measurement can take is

III	IV	III	IV
III	IV	III	IV
I	I	II	II
I	I	II	II

which is analogous to the product basis

$$\{|0\rangle|0\rangle, |0\rangle|1\rangle, |1\rangle|+\rangle, |1\rangle|-\rangle\} \tag{101}$$

in quantum theory.

If the disjoint epistemic states are all perfectly correlated, then we have a measurement such as

IV	III	II	I
III	IV	I	II
II	I	IV	III
I	II	III	IV

(102)

This example is analogous to the Bell basis

$$\{|\Phi^+\rangle, |\Phi^-\rangle, |\Psi^+\rangle, |\Psi^-\rangle\}, \tag{103}$$

where

$$|\Phi^\pm\rangle = \sqrt{2}^{-1}(|0\rangle|0\rangle \pm |1\rangle|1\rangle),$$

$$|\Psi^\pm\rangle = \sqrt{2}^{-1}(|0\rangle|1\rangle \pm |1\rangle|0\rangle). \tag{104}$$

Other examples of measurements composed entirely of correlated epistemic states include

III	IV	I	II
IV	III	II	I
II	I	IV	III
I	II	III	IV

II	IV	III	I
I	III	IV	II
IV	II	I	III
III	I	II	IV

III	II	IV	I
IV	I	III	II
I	IV	II	III
II	III	I	IV

(105)

There also exist measurements that are composed of some uncorrelated and some correlated epistemic states, for instance,

IV	IV	II	I
IV	IV	I	II
II	I	III	III
I	II	III	III

(106)

which is analogous to a measurement of the basis

$$\{|\Phi^+\rangle, |\Phi^-\rangle, |0\rangle|1\rangle, |1\rangle|0\rangle\}. \tag{107}$$

We call measurements that contain correlated epistemic states *joint* measurements since they cannot be implemented by separate measurements on the individual systems. Note that joint measurements can only be implemented directly if the systems are not spatially separated.

H. Mutually unbiased measurements

In quantum theory, two bases are said to be *mutually unbiased* if all the pairwise fidelities between elements from the two bases have the same value. Thus, bases $\{|\psi_i\rangle\}$ and $\{|\chi_j\rangle\}$ are mutually unbiased if $|\langle\psi_i|\chi_j\rangle|^2$ is independent of i and j. The number of mutually unbiased bases (MUBs) that can be constructed depends on the dimensionality d of the Hilbert space. For d a power of a prime, there are $d+1$ MUBs [49].

For a single qubit, the number of MUBs that can be constructed is three. An example of such a triplet of MUBs is

$$\{|0\rangle, |1\rangle\}, \{|+\rangle, |-\rangle\}, \{|+i\rangle, |-i\rangle\}. \tag{108}$$

For a pair of qubits, one can construct five MUBs, an example being

$$\{|0\rangle|0\rangle, |0\rangle|1\rangle, |1\rangle|0\rangle, |1\rangle|1\rangle\},$$

$$\{|+\rangle|+\rangle, |-\rangle|-\rangle, |+\rangle|-\rangle, |-\rangle|+\rangle\},$$

$$\{|-i\rangle|-i\rangle, |+i\rangle|+i\rangle, |-i\rangle|+i\rangle, |+i\rangle|-i\rangle\},$$

$$\{I \otimes U|\Phi^{+}\rangle, I \otimes U|\Phi^{-}\rangle, I \otimes U|\Psi^{+}\rangle, I \otimes U|\Psi^{-}\rangle\},$$

$$\{I \otimes V|\Phi^{+}\rangle, I \otimes V|\Phi^{-}\rangle, I \otimes V|\Psi^{+}\rangle, I \otimes V|\Psi^{-}\rangle\}, \quad (109)$$

where U is the unitary map that corresponds to a clockwise rotation by 120° about the $\hat{x}+\hat{y}+\hat{z}$ axis in the Bloch sphere, and $V=U^{-1}$ [38].

As discussed previously, the analogue in the toy theory of a basis of states is a set of disjoint epistemic states that yield a partitioning of the full set of ontic states. We call two such partitionings *mutually unbiased* if all pairwise classical fidelities (defined in Sec. III A) between elements from the two partitionings have the same value. For a pair of pure epistemic states, the classical fidelity is proportional to the number of ontic states they have in common. It follows that the number of mutually unbiased partitionings (MUPs) for a single elementary system is three,

I	I	II	II

I	II	I	II

I	II	II	I

(110)

There exist sets of five MUPs for a pair of elementary systems, an example being the set

III	III	IV	IV
III	III	IV	IV
I	I	II	II
I	I	II	II

III	IV	III	IV
I	II	I	II
III	IV	III	IV
I	II	I	II

IV	III	III	IV
II	I	I	II
II	I	I	II
IV	III	III	IV

II	IV	III	I
I	III	IV	II
IV	II	I	III
III	I	II	IV

III	II	IV	I
IV	I	III	II
I	IV	II	III
II	III	I	IV

(111)

We conjecture that the number of MUPs for any number of elementary systems is equal to the number of MUBs for the same number of qubits.

I. Dense coding

By transmitting a single qubit from Alice to Bob, the most classical information that can be communicated is one classical bit. This is a consequence of Holevo's theorem. However, if Alice and Bob initially share an entangled pair of qubits, then they can communicate two bits of classical information by transmitting a single qubit. This is known as dense coding [50].

The phenomenon is surprising because it is unclear how adding a resource of entanglement can possibly increase the capacity for communication given that the distribution of the resource may occur at a time prior to Alice even deciding which message she wishes to send, and need not involve any transmission from Alice to Bob; they may both simply receive their half of the entangled pair from a third party. The puzzle is sufficiently acute that some have suggested that the additional bit of information travels backwards in time through the channel that established the entanglement. A different sort of resolution of the puzzle is suggested by the analogue of dense coding in the toy theory. In order to see the extent of the analogy, we begin by presenting the quantum protocol.

A pair of qubits, A and B, described by the entangled state $|\Phi^{+}\rangle=\sqrt{2}^{-1}(|0\rangle|0\rangle+|1\rangle|1\rangle)$, are distributed to Alice and Bob (A to Alice and B to Bob). Depending on which of four messages, 00, 01, 10, or 11, Alice wishes to communicate to Bob, she implements one of four transformations on A corresponding to unitary operators I, σ_z, σ_x, and $i\sigma_y$ (where σ_x, σ_y, and σ_z are the Pauli operators [39]). These transformations map $|\Phi^{+}\rangle$ to the four Bell states, $|\Phi^{+}\rangle, |\Phi^{-}\rangle, |\Psi^{+}\rangle$, and $|\Psi^{-}\rangle$, respectively. Since these are orthogonal, they can be distinguished with certainty. Thus, if Alice sends qubit A to Bob, he holds the pair and can perform a measurement of the Bell basis to determine which of the four messages Alice wished to communicate. In this way, Alice has succeeded in communicating two bits of information to Bob.

In the toy theory, it is also true that the transmission of a single elementary system (without a shared resource of correlation) can only communicate a single classical bit. The reason is as follows. Although a single elementary system has four ontic states, allowing it to carry two bits of classical information, Alice cannot prepare the system to be in precisely one of these ontic states, nor can Bob measure which of the four ontic states describes the system. The best Alice can do is to choose which state of incomplete knowledge describes the system after her preparation procedure. Thus, she could encode one bit of classical information by choosing to perform one or the other of two preparations associated with the epistemic states

 and 

(112)

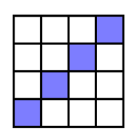
Bob can distinguish which preparation was implemented by subjecting the system to the measurement of the form

I	I	II	II

(113)

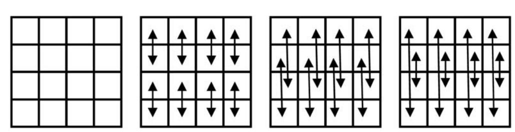
It should be clear that one classical bit is the most that Alice can communicate to Bob in this way.

On the other hand, if Alice and Bob initially each hold one half of a pair of elementary systems that are correlated, then Alice can communicate two bits to Bob. Here is a protocol that achieves this. Suppose that initially Alice holds an elementary system A and Bob holds an elementary system B, and these are known to be described by the epistemic state



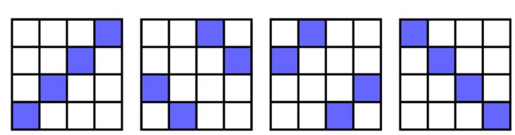
(114)

Alice can, depending on which of four messages she wishes to send, perform one of four permutations on A, namely, (1)(2)(3)(4), (12)(34), (13)(24), or (14)(23), graphically,



(115)

These map the initial epistemic state to the four epistemic states



(116)

There is a measurement that distinguishes these four epistemic states, namely,

IV	III	II	I
III	IV	I	II
II	I	IV	III
I	II	III	IV

(117)

Thus, if Alice sends the system A to Bob, he can implement this measurement and determine which of the four messages Alice wished to communicate.

One can summarize these facts about the toy theory as follows. Every elementary system has the inherent capability of encoding two bits of classical information. However, the knowledge balance principle imposes a restriction that prevents Alice and Bob from making use of this capacity unless they initially share correlated systems. In a toy theory universe, one *cannot* come to know which of four possible ontic states describe a single system, because one cannot learn two bits of information about a single system. However, one *can* come to know which of four possible relations hold between two systems, because one can learn two bits of information about a *pair* of systems. Moreover, one can fix which of these four relations holds by acting on just one of the systems.

Note that the toy theory yields an interesting new perspective on how to compare quantum and classical information theories: rather than comparing a single qubit to a single classical bit, as is conventionally done, the toy theory suggests that it is more appropriate to compare a single qubit to *two* classical bits.

J. Nonmaximally informative measurements

In addition to the measurements considered in Sec. IV G, there are measurements that are not maximally informative. These do not answer as many questions as are allowed by the knowledge balance principle. An example of a *product* measurement that is nonmaximally informative is one that is trivial for one of the systems. For instance, the measurement that is trivial on B and distinguishes $1\vee 2$ from $3\vee 4$ on A is depicted by

II	II	II	II
II	II	II	II
I	I	I	I
I	I	I	I

(118)

Joint measurements can also fail to be maximally informative. For instance, the measurement

II	II	I	I
II	II	I	I
I	I	II	II
I	I	II	II

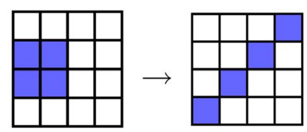
(119)

yields information about the relation between the two systems, but is not as informative as it could be. Indeed, it can be obtained by coarse graining of the outcomes of the measurement described by Eq. (95) or the one described by Eq. (102).

K. Measurement update rule

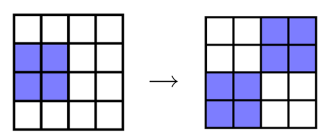
The unknown disturbance associated with measurements that act upon a *single* elementary system has been described in Sec. III E. The unknown disturbance associated with a product measurement is simply a conjunction of such disturbances on the individual subsystems. For joint measurements, we must see what the knowledge balance principle dictates. If the measurement is maximally informative, then in order for it to be repeatable, the updated epistemic state must assign zero probability to all the ontic states that are inconsistent with the outcome that occurred. But no more ontic states can receive probability zero without violating the principle. Thus, the ontic support of the final epistemic state must be the ontic support S of the measurement outcome. This must be true regardless of the initial epistemic state. This implies that an unknown permutation must occur as the result of the measurement, specifically, a permutation drawn uniformly from any set that has the property of randomizing the elements of S (the set of all permutations of the elements of S, for instance, has this property).

For instance, if the initial state is $(2\vee 3)\cdot(1\vee 2)$, and a reproducible measurement of the form of Eq. (102) (analogous to the Bell basis) finds the outcome I, then we have

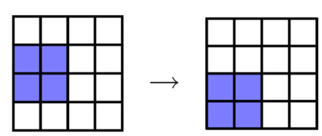


(120)

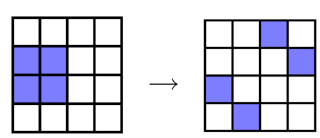
The situation is more complicated for joint measurements that are not maximally informative. Suppose, for instance, that the initial state is $(2\vee 3)\cdot(1\vee 2)$, and that a reproducible measurement of the form of Eq. (119) finds the outcome I. There are many update rules that are consistent with the reproducibility of the measurement. For instance,



(121)



(122)



(123)

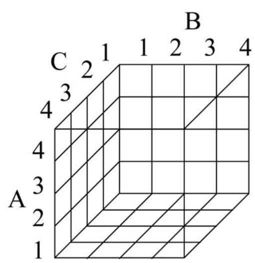
Indeed, any epistemic state appearing in Eq. (81) could be the final epistemic state while still yielding reproducibility.

It turns out that the update rule is not uniquely defined in this case. This is completely analogous to what occurs in quantum theory. There, reproducible measurements are represented by sets of projectors and these fail to be maximally informative if some of the projectors have rank greater than 1. Such a reproducible measurement can be associated with many different maps (differing for instance in the degree with which they maintain coherence in the subspaces defined by the higher rank projectors). Which map applies in a given instance depends on how the measurement is implemented. One update rule, however, is particularly common. This is the one wherein the final quantum state is the projection of the initial quantum state into the subspace associated with the outcome that occurs. We can define an analogous update rule in the toy theory: the final epistemic state is the one with the highest classical fidelity with the initial epistemic state. In this form, the analogy to the quantum update rule is apparent, since the quantum state that is the projection of the initial quantum state into the subspace associated with the outcome is the element of that subspace that has the maximal inner product with the initial quantum state. In the example provided above, this particular update rule corresponds to the second rule we depicted, that is, Eq. (122).

V. TRIPLETS OF ELEMENTARY SYSTEMS

A. Epistemic states

For three elementary systems, each of which has four ontic states, there are 64 ontic states in all. We can represent these by a $4\times4\times4$ grid of boxes, with the three systems labeled by A, B, and C,



(124)

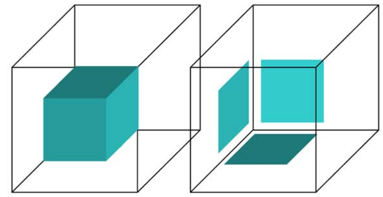
There are six yes-no questions in a canonical set for the three systems. In a state of maximal knowledge, three questions are answered and three are unanswered, which implies that a state of maximal knowledge contains eight ontic states. Any pair of systems in the triplet must also abide by the knowledge balance principle, so that the marginal distributions for the pairs must all be valid epistemic states for a pair of elementary systems.

The pure epistemic states that are allowed by the knowledge balance principle are of three types: (1) no correlations between any of the systems, (2) correlations between one pair of the systems, and (3) correlations between all three systems. We shall see that these are analogous, respectively, to product states, products of a Bell state and a pure state, and the so-called Greenberger-Horne-Zeilinger (GHZ) states [51].

The uncorrelated epistemic states are of the form

$$(1 \vee 2) \cdot (1 \vee 2) \cdot (1 \vee 2) \tag{125}$$

to within local permutations. The marginals over any pair of systems are pure uncorrelated epistemic states. For instance, the marginal on AB is simply $(1\vee2)\cdot(1\vee2)$. We can represent the epistemic states graphically as collections of solid colored $1\times1\times1$ blocks in our $4\times4\times4$ grid, and the marginals as the shadows of these blocks. For instance, the example of Eq. (125) is represented graphically as follows:



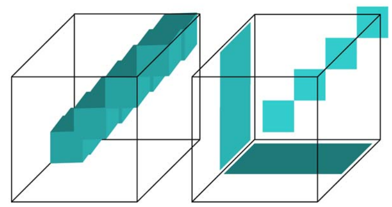
(126)

This is analogous to the quantum state $|0\rangle|0\rangle|0\rangle$.

The pair-correlated epistemic states are of the form

$$[(1\cdot1)\vee(2\cdot2)\vee(3\cdot3)\vee(4\cdot4)]\cdot(1\vee2), \tag{127}$$

to within local permutations. This is represented graphically as



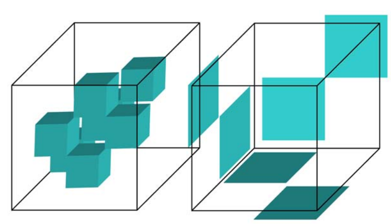
(128)

and is analogous to $\frac{1}{\sqrt{2}}(|0\rangle|0\rangle+|1\rangle|1\rangle)|0\rangle$.

The triplet-correlated epistemic states have the form

$$\begin{aligned}&(1\cdot1\cdot1)\vee(1\cdot2\cdot2)\vee(2\cdot1\cdot2)\vee(2\cdot2\cdot1)\\&\quad\vee(3\cdot3\cdot3)\vee(3\cdot4\cdot4)\vee(4\cdot3\cdot4)\vee(4\cdot4\cdot3)\end{aligned} \tag{129}$$

to within local permutations. The marginals over every pair of elementary systems are correlated mixed states. For the particular example we have provided, they are all of the form $[(1\vee2)\cdot(1\vee2)]\vee[(3\vee4)\cdot(3\vee4)]$. This is represented graphically as follows



(130)

This epistemic state is analogous to the GHZ state for three qubits of the form $\frac{1}{\sqrt{2}}(|0\rangle|0\rangle|0\rangle+|1\rangle|1\rangle|1\rangle)$ which has marginals $\frac{1}{2}|00\rangle\langle00|+\frac{1}{2}|11\rangle\langle11|$ over every pair of subsystems.

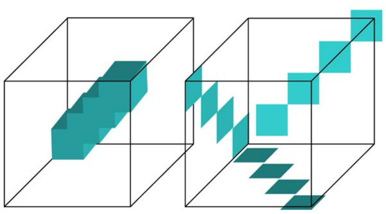
B. The monogamy of pure entanglement

In quantum theory, a system can be pure entangled with only *one* other system. The reason is that if A and B are pure entangled, then the reduced density operator over AB is a pure state. However, for the composite AB to be entangled with another system, the reduced density operator of AB must be mixed. Consequently, there is no entanglement between AB and any other system, and thus no entanglement between A and any other system besides B. This feature of pure state entanglement is sometimes referred to as the *monogamy* of entanglement [52].

From the epistemic perspective, the monogamy of pure entanglement is a monogamy of perfect correlations. In both classical theories and the toy theory, a pair of systems are perfectly correlated if one knows the precise relation between their ontic states. Perfect correlations are monogamous if a system can only be perfectly correlated with one other.

Classical statistical theories are polygamous when it comes to perfect correlations. For instance, it is possible to know that three systems, A, B, and C, are in precisely the same ontic state. In this case, A is perfectly correlated with B and perfectly correlated with C.

The toy theory, however, forbids such polygamy. We demonstrate this in the case of three elementary systems by supposing the contrary and deriving a contradiction with the knowledge balance principle. Suppose three elementary systems, A, B, and C, are all pairwise perfectly correlated. This would imply that for every ontic state of A there was associated a unique ontic state of B and a unique ontic state of C. For instance, one way for A, B, and C to be perfectly correlated would be if they were known to be in precisely the same ontic state, that is, if the epistemic state was $(1\cdot 1\cdot 1)\vee(2\cdot 2\cdot 2)\vee(3\cdot 3\cdot 3)\vee(4\cdot 4\cdot 4)$. This epistemic state and its marginals are represented graphically as follows:



(131)

But this is not one of the three valid forms of epistemic state for a triplet of elementary systems. The problem is that it contains only four ontic states rather than eight, which is the minimum number that is allowed by the knowledge balance principle.

C. Teleportation

A teleportation protocol in quantum theory makes use of a pair of qubits that are maximally entangled and a classical channel in order to transfer the applicability of an unknown quantum state from a qubit in Alice's possession to one in Bob's possession [53]. We shall begin by providing a standard account of how teleportation works within an ontic (rather than epistemic) view of quantum states.

A pair of qubits, denoted A and B, are prepared in the quantum state $|\Phi^+\rangle = \sqrt{2}^{-1}(|0\rangle|0\rangle + |1\rangle|1\rangle)$, after which A is given to Alice and B to Bob. A third party, Victor, prepares another system, denoted A', in the quantum state $|\psi\rangle$, and passes it to Alice. The identity of A''s quantum state is unknown to Alice and Bob. Their task is to implement a protocol that leaves B in the quantum state $|\psi\rangle$. The initial quantum state of $A'AB$ is

$$|\psi\rangle|\Phi^+\rangle. \tag{132}$$

It turns out that this can be rewritten as follows:

$$\frac{1}{2}|\Phi^+\rangle|\psi\rangle + \frac{1}{2}(I\otimes\sigma_z)|\Phi^+\rangle\sigma_z|\psi\rangle + \frac{1}{2}(I\otimes\sigma_x)|\Phi^+\rangle\sigma_x|\psi\rangle + \frac{1}{2}(I\otimes i\sigma_y)|\Phi^+\rangle(i\sigma_y)|\psi\rangle. \tag{133}$$

Note that $(I\otimes\sigma_z)|\Phi^+\rangle = |\Phi^-\rangle$, $(I\otimes\sigma_x)|\Phi^+\rangle = |\Psi^+\rangle$, and $(I\otimes i\sigma_y)|\Phi^+\rangle = |\Psi^-\rangle$, so that the states for $A'A$ in this decomposition are just the elements of the Bell basis. If Alice measures the Bell basis on $A'A$ and obtains the outcome associated with the unitary operator U, where $U\in\{I,\sigma_z,\sigma_x,i\sigma_y\}$, then the quantum state of $A'AB$ is updated to

$$(I\otimes U)|\Phi^+\rangle U|\psi\rangle. \tag{134}$$

If she classically communicates to Bob the identity of U—only two bits of information are required to do so—then Bob can apply the inverse of U to B to leave $A'AB$ in the state

$$(I\otimes U)|\Phi^+\rangle|\psi\rangle. \tag{135}$$

Thus, at the end of the protocol, the quantum state of B is $|\psi\rangle$, as required, and $A'A$ is left in one of the Bell states. The protocol succeeds regardless of the identity of $|\psi\rangle$, so Alice and Bob need not know its identity. Note that if system A' is entangled with a fourth system, C, then the quantum state of $A'C$ is transferred to BC, which is known as *entanglement swapping*.

Teleportation is often thought to be surprising because it takes an infinite amount of information to completely specify a quantum state, but somehow this state can be transferred from one system to another given the transmission of only two bits of classical information. This fact is only surprising, however, if one takes on ontic view of the quantum state. From the perspective that quantum states are states of incomplete knowledge, teleportation is a protocol wherein someone's *knowledge* about the system A' becomes applicable to the system B, and, as we shall see, a transfer of the applicability of a state of knowledge from A' to B requires much less communication from Alice to Bob. We demonstrate this first in the context of a classical theory, and then in the context of the toy theory, where there is a strong analogue to the quantum protocol.

In a classical theory, a transfer of the applicability of a state of knowledge is easily achieved. Suppose Victor describes system A' by some probability distribution $p(x)$ over its ontic states, and suppose that Alice and Bob do not know the nature of this distribution. Nonetheless, Alice can simply measure the ontic state of system A', then communicate this information to Bob, and Bob can prepare system B to be in this particular ontic state. Assuming that Victor knows that they have implemented this protocol, but does not know the outcome of Alice's measurement, he will assign the marginal distribution $p(x)$ to B. However, teleportation requires more than just getting the marginal distribution for B to reflect the initial marginal distribution for A'—the correlations of A' to other systems must also be reproduced. Since Victor initially describes A' as uncorrelated with all other systems, he should, in the end, describe B as uncorrelated with all other systems. The protocol we have just described does not quite achieve teleportation because Victor ends up describing B as perfectly correlated with A'. However, this problem is easily fixed: Alice can simply randomize the ontic state of A' at the end of the protocol.

Note that this protocol only requires Alice to communicate to Bob an amount of information that is sufficient to specify the ontic state of A', and this is in general much less than is required to specify Victor's epistemic state (for instance, there might be a finite number of ontic states, but an infinite number of epistemic states). Note also that this classical protocol succeeds without Alice and Bob requiring any resource of classical correlations.

In the toy theory, this protocol does not work because Alice cannot measure the precise ontic state of A'. Nonetheless, teleportation can be achieved if Alice and Bob initially share correlated systems. Here is how it works. Suppose Alice holds an elementary system A and Bob holds an elementary system B, and the pair is described by the epistemic state $(1\cdot1)\vee(2\cdot2)\vee(3\cdot3)\vee(4\cdot4)$ (analogous to $|\Phi^+\rangle$). It is known that A and B are in the same ontic state, but it is not known what this state is. A third party, Victor, sends to Alice a system A', which he describes by the epistemic state $a\vee b$, where the identity of a and b are unknown to Alice and Bob (analogous to the unknown state $|\psi\rangle$). Victor's initial epistemic state for $A'AB$ is

$$(a\vee b)\cdot((1\cdot1)\vee(2\cdot2)\vee(3\cdot3)\vee(4\cdot4)). \tag{136}$$

Although Alice cannot determine which of the four possible ontic states applies to A', she can determine which of four relations hold between A' and A. For instance, she can determine whether the permutation that relates A to A' is (1)(2)(3)(4), (12)(34), (13)(24), or (14)(23). This is simply the measurement of Eq. (102) (analogous to the Bell basis), applied to $A'A$.

Suppose the permutation relating A to A' is found to be P. Since the permutation that related B to A prior to the measurement was identity, one can conclude that the permutation that related B to A' prior to the measurement was P. Since Victor's state of knowledge about the initial ontic state of A' (where by "initial" we mean prior to the measurement) is $a\vee b$, it follows that, upon learning the outcome of Alice's measurement, his state of knowledge about the initial ontic state of B is $P[a]\vee P[b]$ where $P[a]$ is the image of a under the permutation P. Victor knows that Alice's measurement does not cause a physical disturbance to B, so his state of knowledge about its *final* ontic state (where by "final" we mean *after* the measurement) is also $P[a]\vee P[b]$. On the other hand, A' and A *do* suffer an unknown permutation due to Alice's measurement, which causes the epistemic state for the pair to be updated to $(1\cdot P[1])\vee(2\cdot P[2])\vee(3\cdot P[3])\vee(4\cdot P[4])$, the state appropriate for finding A to be related to A' by the permutation P. Thus, after Alice's measurement, Victor's epistemic state for $A'AB$ is

$$((1\cdot P[1])\vee(2\cdot P[2])\vee(3\cdot P[3])\vee(4\cdot P[4]))\cdot(P[a]\vee P[b]).$$

To complete the teleportation protocol, Alice communicates the outcome of her measurement, the permutation P, to Bob. Since there are four possible outcomes, this requires Alice to communicate two bits of information to Bob. Upon learning P, Bob applies its inverse to B. Thus, Victor's epistemic state at the end of the protocol is

$$((1\cdot P[1])\vee(2\cdot P[2])\vee(3\cdot P[3])\vee(4\cdot P[4]))\cdot(a\vee b).$$

The epistemic state $a\vee b$, which was applicable to A' at the start of the protocol, is now applicable to B. The epistemic state for the pair $A'A$ is left as one of the four correlated epistemic states that are analogous to the Bell states. This is the analogue of teleportation. Had Victor initially known A' to have a particular correlation with a fourth system, C, then at the end of the protocol, he would judge B to have this correlation with C. This is the analogue of entanglement swapping.

It should be noted that even if Victor does not learn the outcome of Alice's measurement, at the end of the protocol he still describes B by the epistemic state he initially assigned to A' (since he knows that Bob will implement the inverse of P, regardless of the identity of P). Note also that we could have chosen a different initial correlated epistemic state for AB, or a different basis of correlated epistemic states for Alice's measurement on $A'A$ [for instance, the basis associated with the four permutations (1)(2)(3)(4), (1234), (13)(24), and (1432)] and teleportation could still be achieved. These freedoms are analogous to freedoms that are present in the quantum protocol.

In the toy theory, even though it takes more than two bits of information to specify which of the six possible epistemic states applies (the analogue of the continuum of quantum states), the applicability of an unknown epistemic state can clearly be transferred from one system to another using only two bits of information. This is precisely what is achieved by the protocol we have described. In fact, a transfer of the applicability of a description from one place to another does not require *any* communication between those locations. Suppose that Alice refrains from sending Bob the two bits of information specifying the outcome of her measurement, but does send this information to Victor. It is still the case that in one quarter of the trials, namely those where Alice finds the ontic states of A' and A to be identical, the applicability of Victor's epistemic state is transferred from system A' to system B. The way in which the applicability of an epistemic state is transferred from one system to another is not by information transmission between the systems, but by information transmission to the individual who is describing the systems.

This toy version of a teleportation protocol is essentially the one provided by Hardy [35], modulo the choice of the set of permutations. Note, however, that his goal was to distinguish teleportation from nonlocality, not to provide an argument for the epistemic view of quantum states. His point, that not every phenomenon involving entanglement involves nonlocality, is reinforced by other examples we have considered here, such as remote steering, dense coding, and the monogamy of entanglement.

VI. FURTHER ANALOGUES

There are some more analogies between the toy theory and quantum theory which we have opted not to present in detail. Nonetheless, it is worth pointing out some of these, lest the phenomena in question be mistaken as uniquely quantum.

(a) The existence of unsharp measurements. In quantum theory, measurements on a system are typically associated with projective-valued measures (PVMs) or, equivalently, Hermitian operators on the system's Hilbert space. These are known as sharp measurements. There are other sorts of measurements on a system, called unsharp measurements, which are associated with positive-operator valued measures (POVMs) on the system's Hilbert space. They may arise by a convex combination of sharp measurements, or by coupling the system to an auxiliary system (called an *ancilla*) and performing a sharp measurement on the composite [39]. In the toy theory, one can also contemplate certain convex combinations of measurements, and one can implement effective measurements on a system by coupling to an ancilla and measuring the composite. These constitute unsharp measurements in the toy theory.

(b) The existence of irreversible transformations. A transformation on a system in quantum theory is reversible if it is associated with a unitary map. An arbitrary transformation, however, is associated with a completely positive trace-preserving linear map [39], which can be nonunitary. These can arise as a result of a convex combination of unitary maps, or by coupling the system to an ancilla and applying a reversible transformation to the pair. Again, operations of these sorts are allowed in the toy theory, and so irreversible transformations arise there as well.

The main features of state discrimination tasks in quantum theory [54] are also reproduced in the toy theory.

(c) No deterministic error-free discrimination of nonorthogonal states.

(d) The possibility of *in*deterministic error-free discrimination of nonorthogonal states (also known as unambiguous discrimination).

(e) No information gain without disturbance in discrimination of nonorthogonal states. This latter phenomenon accounts for the possibility of key distribution in quantum theory [55,56]. It follows that one expects key distribution to be possible in the toy theory as well.

The toy theory also contains analogues of a few recently discovered phenomena involving product bases, namely:

(f) The existence of locally indistinguishable product bases [57].

(g) The existence of unextendible product bases (that is, product bases for which no additional product state can be found that is orthogonal to every element of the basis) [58].

The phenomena (f) is sometimes referred to as "nonlocality without entanglement" (in fact, this was the title of Ref. [57]). This description is perhaps inappropriate given that the toy theory is explicitly local (in Bell's sense) and yet reproduces this phenomena.

(h) The fact that for every outcome of a maximally informative measurement, there is a unique quantum state that yields that outcome with certainty. To be specific, in a measurement associated with the basis $\{|\psi_i\rangle\}$, the only state that yields outcome i with certainty is $|\psi_i\rangle$. This feature cannot be captured within an epistemic approach if one allows for arbitrary probability distributions over the ontic states. The reason is that within such an approach, measurements correspond to a partitioning of the ontic states into disjoint sets, and a particular outcome of a measurement is only certain to occur if the epistemic state prior to the measurement has its ontic support within the ontic support of that outcome. But if arbitrary distributions are allowed, then there will be many epistemic states with the same ontic support. In the toy theory, on the other hand, there are no two distinct epistemic states with the same ontic support, since only uniform distributions are allowed. As a result, there is a one-to-one correspondence between the outcomes of maximally informative measurements and the pure epistemic states.

We have not exhausted the list of quantum phenomena that have analogues in the toy theory, however the point should be clear: the toy theory captures a good deal of quantum theory.

VII. PHENOMENA THAT ARE NOT REPRODUCED

There is a certain satisfaction in being able to reproduce quantum phenomena in a theory that admits a simple interpretation. Nonetheless, what is even more interesting is to identify the quantum phenomena that *cannot* be reproduced by the toy theory, since these now present the greatest challenge to the proponent of the epistemic view, and since these provide the best clues for determining what other conceptual ingredients, besides the idea that maximal information is incomplete, are at play in quantum theory.

Here are some features of quantum theory that are absent from the toy theory:

(i) Contextuality (i.e., the existence of a Kochen-Specker theorem [30,32]).

(ii) Nonlocality (i.e., the existence of a Bell theorem [33]).

(iii) The continuum of quantum states, measurements, and transformations.

(iv) The fact that convex combination and coherent superposition are full rather than partial binary operations on the space of quantum states.

(v) The fact that two levels of a fundamentally three-level system behave like a fundamentally two-level system.

(vi) The possibility of an exponential speed-up relative to classical computation, assuming certain computational problems are classically hard.

We shall consider each of these in turn.

A. Contextuality and nonlocality

The Kochen-Specker theorem [30,32] and Bell's theorem [33] state that any hidden variable theory that is local or noncontextual cannot reproduce all the predictions of quantum theory. The toy theory is, by construction, a local and noncontextual hidden variable theory. Thus, it cannot possibly capture all of quantum theory. In the face of these no-go theorems, a proponent of the epistemic view is forced to accept alternative possibilities for the nature of the ontic states to which our knowledge pertains in quantum theory. It is here that the novel conceptual ingredients are required. Note that since nonlocality is an instance of contextuality [59], the latter can be considered as the more fundamental of the two phenomena. Indeed, if quantum theory can be derived from a principle asserting that maximal information is incomplete and some other conceptual ingredient, then contextuality may be our best clue as to what this other conceptual ingredient must be.

B. Continuum of states, measurements, and transformations

The finite cardinality of epistemic states, reproducible measurements and reversible transformations in the toy theory is due to the fact that these are associated respectively with uniform distributions over, partitionings of, and permutations of a finite set of ontic states.

Of course, by allowing nonuniform probability distributions over the ontic states, measurements whose outcomes are determined only probabilistically by the ontic states, and probabilistic combinations of permutations, one could have a continuum of distinct epistemic states, measurements and transformations over a finite number of ontic states.

As it turns out, however, such a theory cannot reproduce the predictions of quantum theory. The proof is as follows. For every pair of pure quantum states, one can find a measurement and an outcome of this measurement such that the first quantum state assigns zero probability to this outcome while the second assigns to it a nonzero probability. This implies that the first quantum state does not contain in its ontic support any state that is in the ontic support of the measurement outcome, while the second quantum state does. It follows that every pure quantum state has an ontic support that is unique to it, equivalently, no two quantum states have the same ontic support. Since there are a continuum of pure quantum states, there must be a continuum of distinct subsets of the ontic states, which is only possible if the full set of ontic states is a continuum. This proof is due to Hardy [60]. (Of course, in practice one cannot verify that the number of distinct pure quantum states is really a continuum as opposed to being very large but finite, so all one can strictly conclude is that there must be a very large number of ontic states.)

Given these considerations, one is immediately led to the idea of modifying the toy theory to allow for a continuum of ontic states. In this case, there would be an infinite number of questions in the canonical set. However, if one were to keep the knowledge balance principle intact, this would imply that a single elementary system was capable of encoding an infinite number of classical bits, in contrast to the single classical bit that can be encoded in a qubit. Thus, if this variant of the toy theory is to be analogous to quantum theory, it must also involve some modification of the foundational principle; we must consider other ways to guarantee that knowledge is incomplete. An obvious choice is to assume that for N systems, only N of the infinite number of questions in a canonical set can be answered. (There would obviously be a great imbalance of knowledge in this case, since one's ignorance would always far exceed one's knowledge.) This choice, however, has significant problems. The most notable is the fact that there are an infinite number of mutually unbiased partitionings of an infinite set and therefore such a theory would have an infinite number of mutually unbiased measurements. By contrast, in quantum theory there are only three mutually unbiased bases for a qubit, and five for a pair of qubits. Other options for modifying the foundational principle are required here.

C. Full binary operations on epistemic states

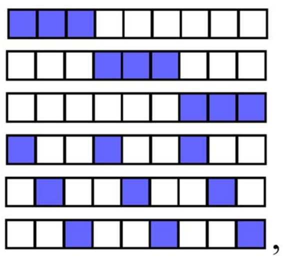
We have seen that there are two types of binary operations defined for epistemic states in the toy theory, analogous to convex combinations and coherent superpositions of quantum states. However, these operations are partial; they are not defined for every pair of epistemic states.

It might therefore seem desirable to close the set of epistemic states in the toy theory under convex combination with arbitrary probability distributions. In this case, the set of allowed epistemic states for a single elementary system would have the shape of an octahedron in the Bloch sphere picture. Hardy's toy theory, for instance, has this feature [35]. Such a variant of our toy theory has also been considered by Halvorson [61]. However, there is an important sense in which such a theory is *less* analogous to quantum theory than the one presented in this paper. The toy theory shares with quantum theory the feature that every mixed state has multiple convex decompositions into pure states, whereas in this modified version, there are many mixed states that have unique decompositions. Similarly, in the toy theory, as in quantum theory, every mixed state has a "purification"—a correlated state between the system of interest and another of equal size such that the marginal over the system of interest is equal to the mixed state in question—whereas in the modified version, there are many mixed states that have only a single purification.

The problem with the modified theory is that although convex combination has been extended to a full binary operation rather than a partial one, the coherent binary operations have not been so extended. Moreover, although one has allowed arbitrary weights in the convex combinations, one has not allowed the analogue of arbitrary amplitudes and phases for the coherent binary operations. It is likely that a better analogy with quantum theory can be obtained only if *both* operations are generalized. Unfortunately, it is unclear how to do so in a conceptually well-motivated way.

D. Embedding two-level systems in three-level systems

The toy theory we have described does not contain anything analogous to a three-level quantum system (called a "qutrit" in quantum information theory). Nonetheless, a variant of the toy theory does. One simply needs to change the measure of knowledge to one that refers to ternary questions (having three possible answers) rather than binary questions. We can then introduce canonical sets of ternary questions, and measure knowledge in terms of these. The knowledge balance principle then dictates that in a state of maximal knowledge, the maximum number of ternary questions for which the answer is known must equal the number for which the answer is unknown.[7] The simplest possible system one can consider is completely specified by the answers to a pair of ternary questions and thus has nine ontic states. In a state of maximal knowledge one has the answer to one of these questions, which corresponds to knowing that the system is in one of three ontic states. For instance, the epistemic states for an elementary system in such a theory are represented graphically as



etc. (137)

Although this variant of the toy theory does a good job of reproducing quantum phenomena involving qutrits, it cannot be combined, in any obvious way, with the original toy theory. For instance, two disjoint epistemic states of a toy qutrit are not isomorphic to two disjoint epistemic states of a toy qubit, since the former involve six ontic states, and the latter four. This is in contrast to quantum theory, where two levels of a fundamentally three-level system are isomorphic to a fundamentally two-level system.[8]

Similarly, a pair of qubits is described in the same way as a fundamentally four-level system in quantum theory. We could define a variant of the toy theory involving tertiary questions (having four answers), which would yield an analogue of a fundamentally four-level system, but this theory would be distinct from the original toy theory applied to a pair of elementary systems. For instance, in the original toy theory a pair of systems must satisfy the knowledge balance principle at the level of the pair and at the level of the individual systems, but nothing analogous to the latter constraint occurs in the variant involving tertiary questions.

E. Exponential speed-ups in computation

If it is indeed the case that quantum computers offer an exponential speed-up over classical computers for certain computational problems [39] (we currently do not have a proof that these problems are in fact difficult for a classical computer), then such a speed-up would be a feature of quantum theory that is not reproduced by the toy theory. This is clear since the toy theory can be efficiently simulated classically. N elementary systems of the toy theory can be modeled by $2N$ classical bits and every operation in the toy theory has a counterpart in the classical model since the toy theory involves a *restriction*, relative to the full classical model, on the permissible preparations, transformations and measurements. Thus, if quantum theory *does* offer an algorithmic speed-up, this is likely to be connected is some way to the other phenomena that the toy theory fails to reproduce, such as the contextuality and nonlocality of quantum theory. In this vein, note that some quantum information-processing tasks that offer an advantage over their classical counterparts have already been shown to have such a connection, specifically, random access codes [64] and communication complexity problems [65].

A distinction between those quantum phenomena that are due to maximal information being incomplete and those quantum phenomena that arise from some other conceptual ingredient is likely to be very useful in the field of quantum information theory, where there is currently a paucity of intuitions regarding what sorts of information-processing tasks can be implemented more successfully in a quantum universe than in a classical universe.

VIII. RELATED WORK

Kirkpatrick has considered a model of a system with two variables wherein it is assumed that the measurement of one variable causes a randomization in the value of the other [34]. This model exhibits noncommutativity of measurements as well as an analogue of interference. The manner in which these phenomena arise for a single elementary system in the toy theory is no different. Kirkpatrick does not, however, consider the possibility of transformations nor the case of multiple systems. Our conclusions are also quite different. While Kirkpatrick emphasizes the classicality of his model, we have tried to focus on the toy theory's innovation relative to a classical theory, namely, that maximal information is incomplete.

Hardy has introduced a toy theory very similar to the one described here [35]. The elementary systems within his theory also have four ontic states. Hardy postulates restrictions on the sorts of measurements that are possible, and a disturbance upon measurement that randomizes the ontic state among the possibilities consistent with the measurement outcome. This implies restrictions on the sorts of epistemic states that apply after a measurement. He also postulates that permutations of the ontic states of a single system are possible transformations.

In its treatment of a single system, Hardy's toy theory is essentially the same as the one presented here, although he has suggested that any convex combination should be allowed, in which case the set of epistemic states is the convex hull of the ones we consider (an octahedron on the Bloch sphere) [67]. Some of the disadvantages associated with adopting this set of epistemic states were discussed in Sec. VII. For multiple systems, the differences between Hardy's theory and the one presented here are more significant. Specifically, the set of measurements allowed in Hardy's theory is larger than the set picked out by the knowledge balance principle. For instance, many of the epistemic states that are forbidden by the knowledge balance principle, such as those displayed in Eq. (58) (except for the first of these) and those

[7]This theory is likely to be closely connected with the theory of qutrits confined to stabilizer states. The latter have been shown by Gross to have positive discrete Wigner representation [62] and thus admit a local noncontextual hidden variable model.

[8]The existence of such an isomorphism is one of the axioms in Hardy's axiomatization of quantum theory [63].

displayed in Eq. (78), are allowed in Hardy's toy theory. But excluding such states was critical to obtaining a good analogy with quantum theory, so Hardy's theory provides a weaker analogy to quantum theory than the one presented here. Note however that Hardy invented his theory for the purpose of demonstrating the possibility of a local theory that exhibits teleportation, and for this it is quite sufficient.

Smolin has constructed several toy models involving "lockboxes" [36]. The motivation for his work is to reproduce certain information-theoretic phenomena which have been suggested as postulates for quantum theory, specifically, no superluminal signaling, no broadcasting, no bit commitment and key distribution. One of Smolin's models, involving pairs of lockboxes, succeeds in this task. It assumes, however, that every pair of lockboxes bears a unique label and this assumption has recently been criticized as unphysical [68].

There is an interesting connection between Smolin's theory and our own. By abandoning the assumption of unique labels, and by formulating Smolin's model in a different manner, one obtains a variant of the toy theory. Suppose that every elementary system (a single lockbox in Smolin's terminology) has two possible ontic states, and thus only a single yes-no question that can be asked of it. Now assume that the answer to this question is always unknown. Denoting the two ontic states by 1 and 2, it follows that the only valid epistemic state for a single system is $1\vee 2$, and there are no nontrivial measurements. However, permuting 1 and 2 does not increase one's knowledge and is therefore an allowed transformation. For a pair of such systems, there are four possible ontic states, and thus two yes-no questions that can be asked of the pair. Assume that one can know the answer to *one* of these questions. Recalling that the marginals on the individual systems must be $1\vee 2$, it follows that the only valid epistemic states for the pair are $(1\cdot 1)\vee(2\cdot 2)$ and $(1\cdot 2)\vee(1\cdot 2)$. This corresponds to knowing that the ontic states of the two systems are the same, or knowing that they are different. The only possible measurement on the pair is the one that determines whether the ontic states of the two are the same or different. Note that a permutation on either system takes one epistemic state to the other. It is this last feature which is critical for establishing the impossibility of bit commitment.

Because of the assumption of unique labels for pairs, Smolin's model did not incorporate the possibility of correlation between more than two systems. By the lights of our reformulation however, it is natural to assume that for three systems (and three yes-no questions) one could still only have the answer to a single question, while for four systems, one could have the answer to two, and so forth. Although the resulting theory will not provide as good an analogy to quantum theory as does our toy theory, it would be interesting to explore the differences, since this is likely to shed light on how much work is being done by the assumption of a *balance* of knowledge and ignorance and how much is being done by the assumption of maximal knowledge being incomplete.

The above models all resemble the toy theory insofar as they are local noncontextual hidden variable theories. They do not, however, share the foundational principle from which the toy theory was derived. By contrast, Zeilinger has advocated an approach to quantum theory which is operational, denying any hidden ontic states, but which adopts a similar foundational principle [37]. Zeilinger's principle is that N elementary systems represent the truth values of N propositions. The propositions to which Zeilinger is referring are propositions stating the outcomes of measurements on the system, rather than propositions about the ontic state of the system. In particular, these propositions concern the outcomes of measurements associated with a set of mutually unbiased bases (Zeilinger calls these "mutually complementary measurements"). Note that the structure of the set of measurements in quantum theory is taken for granted in this approach; the existence of a particular number of mutually unbiased bases for an elementary system is assumed rather than derived. Had one assumed that there was only a single measurement for every elementary system, then Zeilinger's principle would be consistent with knowing the truth values for all propositions pertaining to a system and would therefore yield a classical theory. In the toy theory, the ratio of the number of known propositions to the total number of propositions which pertain to a system is fixed by the assumption of a balance between knowledge and ignorance.

Finally, Wootters has recently introduced a representation of the quantum states of N qubits as real functions on a discrete space of 4^N elements [38]. This is a generalized Wigner function representation of the quantum states. Since these functions can be negative, they cannot be interpreted as epistemic states. Nonetheless, this approach is likely to facilitate the comparison of quantum theory to the toy theory.

IX. CONCLUSIONS

We have considered the consequences of a principle of equality between knowledge and ignorance to the structure of the set of possible states of knowledge. We have examined the manner in which such states of knowledge may be decomposed into convex sums, decomposed into "coherent" sums, transformed, inverted, updated, remotely "steered," cloned, broadcast, teleported, and so forth. In all of these respects we have found that they resemble quantum states. This is strongly suggestive that quantum states should be interpreted as states of incomplete knowledge.

The toy theory contains almost no *physics.* The motional degree of freedom was assumed classical, and there were no masses or charges or forces or fields or Hamiltonians anywhere in the theory. Although this is a shortcoming from the perspective of obtaining an empirically adequate theory, it helps make the case for the epistemic view. Specifically, it supports the idea that a great number of quantum phenomena, and in particular all the phenomena that the toy theory reproduces, have nothing to do with physics, but rather concern only the manipulation of our information about the world.[9] Since the spectra of atoms are not reproduced in the toy theory, these might well be indicative of some real physics, but no cloning and quantum teleportation, for instance, are probably not.

[9]This idea has also been defended by Fuchs [24].

The following three questions for future research suggest themselves:

(i) Can we derive the knowledge balance principle from a physical principle governing the interactions between systems, treating observers as physical systems? Most scientific realists seek a theory that is *universal*, wherein apparatuses and observers are physical systems like any other rather than unanalyzed primitives that appear in the axioms of the theory. Thus, even if one could derive quantum theory from a set of axioms that included a principle of maximal information being incomplete, the question of whether and how this principle could be justified by some *physical* principle, governing all systems, including observers, would be left open. It may be useful to begin by attempting to answer this question in the context of the toy theory, rather than in the context of quantum theory.

(ii) What are the ontic states of which quantum states are states of knowledge? Within the context of the research program outlined here, this question captures the central mystery of quantum theory. Contextuality and nonlocality imply that there must be some modification, relative to classical theories, of our conception of reality if we are to interpret quantum states as states of incomplete knowledge about this reality. Specifically, there cannot be local systems with attributes that are measured in a noncontextual way. Many who adopt an epistemic interpretation of the quantum state abandon the notion that the knowledge represented by the quantum state is knowledge of a preexisting reality. Rather, it is assumed that the quantum state can only represent someone's knowledge about the outcomes of future measurements, or, more generally, the outcomes of future interventions into the world, for instance, whether or not there will be an audible click in a certain detector [23]. However, a proponent of the epistemic view is not forced to this conclusion. Noncontextual hidden variables and the outcomes of future interventions do not exhaust the possibilities for the sample space over which states of knowledge could be defined. We feel that the most promising avenue for the epistemic program is to investigate these other possibilities.

(iii) Is there a second principle that can capture the missing quantum phenomena? A principle stating that maximal knowledge is incomplete knowledge is likely to serve as a foundational principle in a simple axiomatization of quantum theory. This is the claim that we argue is made plausible by the strength of the analogy between the toy theory and quantum theory. Nonetheless, this principle is insufficient for deriving quantum theory. It is intriguing to speculate that we are lacking just one additional conceptual ingredient, just one extra principle about reality, from which all the phenomena of quantum theory, including contextuality and nonlocality, might be derived. To find a plausible candidate for a second such principle, it may be useful to adopt a similar strategy to the one used here to argue for the first principle: do not attempt to derive all of quantum theory, but rather focus on the more modest goal of reproducing a variety of quantum phenomena, even if only qualitatively and in the context of some incomplete and unphysical theory. In particular, attempt to reproduce those phenomena that the toy theory fails to reproduce. Armed with a conceptual innovation that captures the essence of the missing quantum phenomena, a path to quantum theory might suggest itself.

ACKNOWLEDGMENTS

I am indebted to Joseph Emerson and Chris Fuchs for opening my eyes to the epistemic view. I am also thankful to Stephen Bartlett, Chris Fuchs, Lucien Hardy, Terry Rudolph, Peter Turner, and Antony Valentini for many helpful comments.

APPENDIX A: WHY THE TOY THEORY IS NOT A RESTRICTION OF QUANTUM THEORY

The strong similarity of the toy theory to quantum theory might lead one to believe that the epistemic states, measurements, and transformations that apply to N elementary systems in the toy theory are simply subsets of the states, measurements, and transformations that apply to N qubits in quantum theory. This is not the case however. First, there is the fact that the coherent binary operations in the toy theory are not precisely analogous to coherent superpositions in quantum theory, as described in Sec. III A. Second, there is the fact that the set of transformations in the toy theory includes permutations analogous to antiunitary maps, which do not arise in a restricted version of quantum theory. A third fact is that the nature of the correlations for mutually unbiased measurements is different in the two theories, as we now demonstrate.

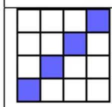
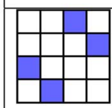
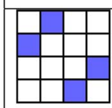
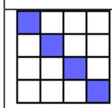
Suppose a pair of qubits is described by one of the four Bell states $|\Phi^+\rangle, |\Phi^-\rangle, |\Psi^+\rangle$, or $|\Psi^-\rangle$, and that one of three mutually unbiased measurements are implemented on each qubit: $\{|0\rangle, |1\rangle\}$ on each qubit, $\{|+\rangle, |-\rangle\}$ on each qubit, or $\{|+i\rangle, |-i\rangle\}$ on each qubit. For each state and each possible measurement, one obtains either correlation between the outcomes (the same outcome for each qubit), or anticorrelation (different outcomes for the two qubits). The results are summarized in Table II, where "C" denotes correlation and "A" denotes anticorrelation. One notes that in all cases there are an odd number of anticorrelations.

We can consider the analogous experiment in the toy theory. A pair of elementary systems are described by one of four pure correlated epistemic states (heading the rows of Table III), and one of three mutually unbiased measurements is implemented on each system (heading the columns of Table III). Again, one finds either correlated or anticorrelated outcomes, however, the number of anticorrelations is always even. Since one cannot achieve an even number of anticorrelations for any quantum state, it is clear that the toy theory for N elementary systems is not simply a restricted version of quantum theory for N qubits.

TABLE II. Correlations (C) and anticorrelations (A) for different measurements on each qubit of a pair prepared in one of the Bell states.

	$\{\|0\rangle, \|1\rangle\}$	$\{\|+\rangle, \|-\rangle\}$	$\{\|+i\rangle, \|-i\rangle\}$
$\|\Phi^+\rangle$	C	C	A
$\|\Phi^-\rangle$	C	A	C
$\|\Psi^+\rangle$	A	C	C
$\|\Psi^-\rangle$	A	A	A

TABLE III. Correlations (C) and anticorrelations (A) for mutually unbiased measurements given correlated epistemic states analogous to the Bell states.

	I I II II	I II I II	I II II I
	C	C	C
	C	A	A
	A	C	A
	A	A	C

APPENDIX B: RELEVANCE TO QUANTUM AXIOMATICS

There has recently been much interest in the possibility of deriving some or all of the quantum formalism from information-theoretic axioms. Fuchs has popularized the question [23], and it has been addressed in many recent articles [23,36,61,66,68]. The toy theory shows that many of the information-theoretic effects one finds in quantum theory are not unique to the latter, and this has important consequences for some proposed axiomatizations.

For instance, it is likely that in the toy theory key distribution [56] is possible, as discussed in Sec. VI. Moreover, it is likely that arbitrarily concealing and arbitrarily binding bit commitment [69,70] is not possible in the toy theory. For instance, the fact that there is an analogue of remote steering, as demonstrated in Sec. IV B, shows that an analogue of the Bennett-Brassard 1984 protocol for bit commitment [56] will not be secure against Alice. We have not *rigorously* established the possibility of key distribution and the impossibility of bit commitment, since to do so properly is a nontrivial task. Nonetheless, our results strongly suggest the falsity of an informal conjecture that the possibility of key distribution and the impossibility of bit commitment together imply quantum theory [71].

Recently, Clifton, Bub, and Halvorson (CBH) [66] have shown that within the context of a C^* algebraic framework, one can derive quantum theory from three information-theoretic postulates: the impossibility of superluminal information transfer through measurements, the impossibility of broadcasting, and the impossibility of bit commitment.

As we have shown, broadcasting is impossible in the toy theory, and since the theory is explicitly local, there is clearly no superluminal information transfer through measurement. Moreover, as discussed above, it is very likely that bit commitment is impossible in the toy theory. These facts do not, however, challenge the CBH characterization theorem since the toy theory does not fall within the C^* algebraic framework. For instance, convex combination is only a partial binary operation within the toy theory and is not defined for arbitrary probability distributions, features that are required within the C^* algebraic framework [61]. Two possibilities suggest themselves: either the assumption of a C^* algebraic framework rules out physically reasonable theories, or a closer examination of those features of the toy theory which cause it to fall outside this framework will show that it is not physically reasonable after all. Similar conclusions can be drawn form the work of Smolin [36].

Although the toy theory might ultimately be a setback for the CBH approach insofar as it leads one to question the innocence of the assumption of a C^* algebraic framework, the fact that it is derived from a simple information-theoretic principle, the knowledge balance principle, and the fact that it is so close in spirit to quantum theory suggests that the prospects for an axiomatization of quantum theory that is predominantly information theoretic are actually quite good.

[1] P. A. M. Dirac, *Principles of Quantum Mechanics*, 4th ed. (Oxford University Press, Oxford, 1958).
[2] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, NJ, 1955).
[3] G. C. Ghirardi, A. Rimini, and T. Weber, Phys. Rev. D **34**, 470 (1986).
[4] P. Pearle, Phys. Rev. A **39**, 2277 (1989).
[5] P. Pearle, in *Open Systems and Measurement in Relativistic Quantum Theory*, edited by F. Petruccione and H. P. Breuer (Springer-Verlag, Berlin, 1999).
[6] H. Everett III, Rev. Mod. Phys. **29**, 454 (1957).
[7] J. A. Barrett, *The Quantum Mechanics of Minds and Worlds* (Oxford University Press, Oxford, 1999).
[8] D. Bohm and B. J. Hiley, *The Undivided Universe* (Routledge, New York, 1993).
[9] P. R. Holland, *The Quantum Theory of Motion* (Cambridge University Press, New York, 1993).
[10] A. Valentini, Ph.D. thesis, The International School for Advanced Studies, Trieste, Italy, 1992.
[11] S. Kochen, in *Symposium on the Foundations of Modern Physics*, edited by P. Lahti and P. Mittelstaedt (World Scientific, Singapore, 1985), p. 151; D. Dieks, Ann. Phys. **7**, 174 (1988); R. Healey, *The Philosophy of Quantum Mechanics* (Cambridge University Press, Cambridge, 1989).
[12] *The Modal Interpretation of Quantum Mechanics*, edited by D. Dieks and P. Vermaas (Kluwer Academic Publishers, Boston, 1998).
[13] G. Bacciagaluppi, Ph.D. thesis, University of Cambridge, 1996.
[14] D. Howard, Stud. Hist. Philos. Sci **16**, 171 (1985); A. Fine, *The Shaky Game: Einstein, Realism, and the Quantum Theory* (University of Chicago Press, Chicago, 1986), p. 35.

[15] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
[16] L. E. Ballentine, Rev. Mod. Phys. **42**, 358 (1970).
[17] L. E. Ballentine, Y. Yang, and J. P. Zibin, Phys. Rev. A **50**, 2854 (1994).
[18] J. V. Emerson, Ph.D. thesis, Simon Fraser University, Vancouver, Canada, 2001, e-print quant-ph/0211035.
[19] R. Peierls, *Surprises in Theoretical Physics* (Princeton University Press, Princeton, 1979).
[20] C. M. Caves and C. A. Fuchs, e-print quant-ph/9601025.
[21] C. M. Caves, C. A. Fuchs, and R. Schack, J. Math. Phys. **43**, 4537 (2002).
[22] C. M. Caves, C. A. Fuchs, and R. Schack, Phys. Rev. A **65**, 022305 (2002).
[23] C. A. Fuchs, e-print quant-ph/0205039; C. A. Fuchs, J. Mod. Opt. **50**, 987 (2003).
[24] C. A. Fuchs, e-print quant-ph/0105039.
[25] A. Daffertshofer, A. R. Plastino, and A. Plastino, Phys. Rev. Lett. **88**, 210601 (2002).
[26] T. Rudolph and R. W. Spekkens, Phys. Rev. A **70**, 052306 (2004).
[27] L. E. Ballentine, in *Fundamental Problems in Quantum Physics*, edited by M. Ferrero and A. van der Merwe (Kluwer Academic, Dordrecht, 1995).
[28] D. Collins and S. Popescu, Phys. Rev. A **65**, 032321 (2002).
[29] C. M. Caves, C. A. Fuchs, and R. Schack, Phys. Rev. A **65**, 022305 (2002).
[30] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966). Reprinted in *Speakable and Unspeakable in Quantum Mechanics*, Ref. [31], Chap. 1.
[31] J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, New York, 1987).
[32] S. Kochen and E. P. Specker, J. Math. Mech. **17**, 59 (1967).
[33] J. S. Bell, Physics (Long Island City, N.Y.) **1**, 195 (1964). Reprinted in *Speakable and Unspeakable in Quantum Mechanics*, Ref. [31], Chap. 2.
[34] K. A. Kirkpatrick, Found. Phys. Lett. **16**, 199 (2003).
[35] L. Hardy, e-print quant-ph/9906123.
[36] J. A. Smolin, Quantum Inf. Comput. **5**, 161 (2005).
[37] A. Zeilinger, Found. Phys. **29**, 631 (1999).
[38] W. K. Wootters, IBM J. Res. Dev. **48**, 99 (2004); quant-ph/0306135.
[39] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, New York, 2000).
[40] T. A. Brun, J. Finkelstein, and N. D. Mermin, Phys. Rev. A **65**, 032315 (2002).
[41] R. W. Spekkens and T. Rudolph, Quantum Inf. Comput. **2**, 66 (2002).
[42] P. Rungta, V. Buzek, C. M. Caves, M. Hillery, and G. J. Milburn, Phys. Rev. A **64**, 042315 (2001).
[43] Č. Brukner, M. Zukowski, and A. Zeilinger, e-print quant-ph/0106119.
[44] E. Schrödinger, Proc. Cambridge Philos. Soc. **31**, 553 (1935).
[45] J. S. Bell, *Speakable and Unspeakable in Quantumn Mechanics* , Ref. [31], Chap. 7.
[46] W. K. Wootters and W. Zurek, Nature (London) **299**, 802 (1982).
[47] D. Dieks, Phys. Lett. **92A**, 271 (1982).
[48] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. Schumacher, Phys. Rev. Lett. **76**, 2818 (1996).
[49] W. K. Wootters and B. D. Fields, Ann. Phys., **191**, 363 (1989).
[50] C. H. Bennett and S. J. Wiesner, Phys. Rev. Lett. **69**, 2881 (1992).
[51] D. Greenberger, M. A. Horne, and A. Zeilinger, in *Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, edited by M. Kafatos (Kluwer Academic, Dordrecht, 1989), p. 69; N. D. Mermin, Phys. Today **43**, 9 (1990).
[52] V. Coffman, J. Kundu, and W. K. Wootters, Phys. Rev. A **61**, 052306 (2000).
[53] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. Wootters, Phys. Rev. Lett. **70**, 1895 (1993).
[54] A. Chefles, Contemp. Phys. **41**, 401 (2000).
[55] C. A. Fuchs, Fortschr. Phys. **46**, 535 (1998); C. A. Fuchs and K. Jacobs, Phys. Rev. A **63**, 062305 (2001).
[56] C. H. Bennett and G. Brassard, in *Proceedings of the IEEE International Conference on Computer, Systems, and Signal Processing, Bangalore, India* (IEEE, New York, 1984), pp. 175–179.
[57] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **59**, 1070 (1999).
[58] C. H. Bennett, D. P. DiVincenzo, T. Mor, P. W. Shor, J. A. Smolin, and B. M. Terhal, Phys. Rev. Lett. **82**, 5385 (1999).
[59] N. D. Mermin, Rev. Mod. Phys. **65**, 803 (1993).
[60] L. Hardy, Stud. Hist. Philos. Mod. Phys. **35B**, 267 (2004).
[61] H. Halvorson, Stud. Hist. Philos. Mod. Phys. **35B**, 277 (2004).
[62] D. Gross, e-print quant-ph/0602001.
[63] L. Hardy, e-print quant-ph/0101012.
[64] E. Galvão, Ph.D. thesis, Oxford University, Oxford, UK, 2002, e-print quant-ph/0212124.
[65] Č. Brukner, M. Zukowski, J.-W. Pan, A. Zeilinger, Phys. Rev. Lett. **92**, 127901 (2004).
[66] R. Clifton, J. Bub, and H. Halvorson, Found. Phys. **33**, 1561 (2003).
[67] L. Hardy (private communication).
[68] H. Halvorson and J. Bub, Quantum Inf. Comput. **5**, 170 (2005).
[69] D. Mayers, Phys. Rev. Lett. **78**, 3414 (1997); H.-K. Lo and H. F. Chau, *ibid.* **78**, 3410 (1997).
[70] R. W. Spekkens and T. Rudolph, Phys. Rev. A **65**, 012310 (2001).
[71] See Ref. [24], pp. 96–98; J. Bub, Found. Phys. **31**, 735 (2001).