2013

Maximally Epistemic Interpretations of the Quantum State and Contextuality

Matthew S. Leifer
Chapman University, leifer@chapman.edu

O. J. E. Maroney
University of Oxford

Recommended Citation

Leifer, M.S., Maroney, O.J.E., 2013. Maximally Epistemic Interpretations of the Quantum State and Contextuality. *Phys. Rev. Lett.* 110, 120401. doi:10.1103/PhysRevLett.110.120401

Maximally Epistemic Interpretations of the Quantum State and Contextuality

M. S. Leifer* and O. J. E. Maroney[1,†]

[1]*Faculty of Philosophy, University of Oxford, 10 Merton Street, Oxford, OX1 4JJ, United Kingdom*
(Received 25 August 2012; published 20 March 2013)

We examine the relationship between quantum contextuality (in both the standard Kochen-Specker sense and in the generalized sense proposed by Spekkens) and models of quantum theory in which the quantum state is maximally epistemic. We find that preparation noncontextual models must be maximally epistemic, and these in turn must be Kochen-Specker noncontextual. This implies that the Kochen-Specker theorem is sufficient to establish both the impossibility of maximally epistemic models and the impossibility of preparation noncontextual models. The implication from preparation noncontextual to maximally epistemic then also yields a proof of Bell's theorem from an Einstein-Podolsky-Rosen-like argument.

DOI: 10.1103/PhysRevLett.110.120401 PACS numbers: 03.65.Ta, 03.65.Ud

The nature of the quantum state has been debated since the early days of quantum theory. Is it a state of knowledge or information (an epistemic state), or is it a state of physical reality (an ontic state)? One of the reasons for being interested in this question is that many of the phenomena of quantum theory are explained quite naturally in terms of the epistemic view of quantum states [1]. For example, the fact that nonorthogonal quantum states cannot be perfectly distinguished is puzzling if they correspond to distinct states of reality. However, on the epistemic view, a quantum state is represented by a probability distribution over ontic states, and nonorthogonal quantum states correspond to overlapping probability distributions. Indistinguishability is explained by the fact that preparations of the two quantum states would sometimes result in the same ontic state, and in those cases there would be nothing existing in reality that could distinguish the two.

Several theorems have recently been proved showing that the quantum state must be an ontic state [2,3]. Most of these have been proved within the ontological models framework [4], which generalizes the hidden variable approach used to prove earlier no-go results, such as Bell's theorem [5] and the Bell-Kochen-Specker theorem [6,7]. However, each of these new theorems rests on auxiliary assumptions, of varying degrees of reasonableness. For example, the Pusey-Barrett-Rudolph theorem [2] assumes that the ontic states of two systems prepared in a product state are statistically independent, and the Colbeck-Renner result [3] employs a strong "free choice" assumption that rules out deterministic theories *a priori*. An explicit counterexample shows that these proofs cannot be made to work without such auxiliary assumptions [8].

The requirement of ontic quantum states is perhaps the strongest constraint on hidden variable theories that has been proved to date. It immediately implies preparation contextuality (within the generalized approach to contextuality of Spekkens [9]), a version of Bell's theorem, and that the ontic state space must be infinite, with a number of parameters that increases exponentially with Hilbert space dimension. See Ref. [10] for a discussion of these implications. However, the auxiliary assumptions used in the proofs of the onticity of quantum states carry over into these corollaries whereas the original proofs of these results [5,9,11] did not require them. For this reason, it is interesting to look for results addressing the distinction between ontic and epistemic quantum states that are weaker than completely ontic, but can be proved without auxiliary assumptions, since such results may sit near the top of a hierarchy of no-go theorems.

Recently, one of us introduced a stronger notion of what it means for the quantum state to be epistemic and proved that it is incompatible with the predictions of quantum theory without any auxiliary assumptions [12]. An ontological model is maximally ψ-epistemic if the quantum probability of obtaining the outcome $|\phi\rangle$ when measuring a system prepared in the state $|\psi\rangle$ is entirely accounted for by the overlap between the corresponding probability distributions in the ontological model. This property is required if the epistemic explanation of the indistinguishability of nonorthogonal states is to be strictly true. It is satisfied by the ψ-epistemic model of two-dimensional Hilbert spaces proposed by Kochen and Specker [4,7], and its analog is satisfied by the epistemic toy model of Spekkens [1].

In this Letter, we explain how this stronger notion of quantum state epistemicity relates to other no-go theorems, particularly the traditional notion of noncontextuality used in proofs of the Kochen-Specker theorem and Spekkens' notion of preparation contextuality. Briefly, Kochen-Specker noncontextuality applies to deterministic models, and says that if an outcome corresponding to some projector is certain to occur in one measurement then outcomes corresponding to the same projector in other measurements must also be certain to occur. Preparation noncontextuality says that preparation procedures corresponding to the same density operator must be assigned the same probability distribution. Our results can be summarized as

$$\text{Preparation noncontextual} \Rightarrow \text{Maximally}\,\psi\text{-epistemic} \Rightarrow \text{Kochen-Specker noncontextual.} \tag{1}$$

Both implications are strict, which we demonstrate with specific examples of models that are Kochen-Specker noncontextual but not maximally ψ-epistemic, and maximally ψ-epistemic but not preparation noncontextual. Since the no-go theorem for maximally epistemic models does not require auxiliary assumptions, these implications provide a stronger proof of preparation contextuality and Bell's theorem than those obtained from other no-go theorems for ψ-epistemic models.

We are interested in ontological models that reproduce the quantum predictions for a set of prepare-and-measure experiments. The experimenter can perform measurements of a set of orthonormal bases $\mathcal{M} = \{M_1, M_2, \ldots\}$ on the system. Let $\mathcal{P} = \cup_{M\in\mathcal{M}} M$ denote the set of quantum states that occur in one or more of these bases. Prior to the measurement, the experimenter can prepare the system in any of the states in $\mathcal{P}$.

An ontological model for $\mathcal{M}$ specifies a measure space $(\Lambda, d\lambda)$ of ontic states. Each state $|\psi\rangle \in \mathcal{P}$ is associated with a probability distribution $\mu_\psi(\lambda)$ [13,15] over Λ and each measurement $M \in \mathcal{M}$ is associated with a set of positive response functions $\xi_M(\alpha|\lambda)$ that satisfy $\sum_{|\alpha\rangle\in M} \xi_M(\alpha|\lambda) = 1$ for all $\lambda \in \Lambda$. The ontological model is required to reproduce the Born rule, which means that $\forall\ |\psi\rangle \in \mathcal{P}, M \in \mathcal{M}, |\alpha\rangle \in M$,

$$\int_\Lambda \xi_M(\alpha|\lambda)\mu_\psi(\lambda)d\lambda = |\langle\alpha|\psi\rangle|^2. \tag{2}$$

For each state $|\psi\rangle \in \mathcal{P}$, define $\Lambda_\psi = \{\lambda|\mu_\psi(\lambda) > 0\}$. We assume that $\Lambda = \cup_{|\psi\rangle\in\mathcal{P}}\Lambda_\psi$, since otherwise there will be superfluous ontic states that are never prepared. Two important facts, of which we make repeated use, are that in order to reproduce $|\langle\psi|\psi\rangle|^2 = 1$ in Eq. (2), for every M that contains $|\psi\rangle$ it must be the case that $\xi_M(\psi|\lambda) = 1$ almost everywhere on Λ_ψ, and for all orthogonal $|\phi\rangle \in M$, such that $|\langle\phi|\psi\rangle|^2 = 0$, $\xi_M(\phi|\lambda) = 0$ almost everywhere on Λ_ψ. This implies that $\Lambda_\psi \cap \Lambda_\phi$ is of measure zero for orthogonal $|\psi\rangle$ and $|\phi\rangle$.

A ψ-ontic ontological model is one in which, for any pair of nonorthogonal quantum states $|\psi\rangle \neq |\phi\rangle$, $\Lambda_\psi \cap \Lambda_\phi$ is of measure zero. This means that, if one knows the ontic state then the prepared quantum state can be identified almost surely. Conversely, if $\Lambda_\psi \cap \Lambda_\phi$ has positive measure for some pair of states, then the model is ψ-epistemic.

An ontological model is maximally ψ-epistemic if $\int_{\Lambda_\phi} \mu_\psi(\lambda)d\lambda = |\langle\phi|\psi\rangle|^2$ for every $|\psi\rangle$, $|\phi\rangle \in \mathcal{P}$. Since $\xi_M(\phi|\lambda) = 1$ almost everywhere on Λ_ϕ, then $\int_{\Lambda_\phi} \xi_M(\phi|\lambda)\mu_\psi(\lambda)d\lambda = |\langle\phi|\psi\rangle|^2$. The probability of obtaining the outcome $|\phi\rangle$ when measuring a system prepared in the state $|\psi\rangle$ is entirely accounted for by the overlap between μ_ψ and μ_ϕ.

The traditional notion of noncontextuality used in proofs of the Kochen-Specker theorem is the combination of two conditions: 1. An ontological model is outcome deterministic if $\xi_M(\alpha|\lambda) \in \{0, 1\}$ almost everywhere on Λ, for all $M \in \mathcal{M}$, $|\alpha\rangle \in M$. 2. An ontological model is measurement noncontextual if, whenever $M, M' \in \mathcal{M}$ contain a common state $|\alpha\rangle$, $\xi_M(\alpha|\lambda) = \xi_{M'}(\alpha|\lambda)$ almost everywhere on Λ.

Theorem 1: If an ontological model of $\mathcal{M}$ is maximally ψ-epistemic then it is also outcome deterministic and measurement noncontextual.

Proof.—The proof closely parallels that of the "quantum deficit theorem" [16]. As mentioned above, for any $|\phi\rangle \in \mathcal{P}$, $\xi_M(\phi|\lambda) = 1$ almost everywhere on Λ_ϕ for every $M \in \mathcal{M}$ that contains $|\phi\rangle$. Hence, it is also equal to 1 almost everywhere on $\Lambda_\phi \cap \Lambda_\psi$, since Λ_ψ has positive measure. In order to reproduce the Born rule, the ontological model must satisfy

$$\int_\Lambda \xi_M(\phi|\lambda)\mu_\psi(\lambda)d\lambda = |\langle\phi|\psi\rangle|^2, \tag{3}$$

but a maximally ψ-epistemic theory must also satisfy

$$\int_{\Lambda_\phi} \xi_M(\phi|\lambda)\mu_\psi(\lambda)d\lambda = |\langle\phi|\psi\rangle|^2. \tag{4}$$

Given that these two equations must hold for all $|\psi\rangle \in \mathcal{P}$, comparing them yields $\xi_M(\phi|\lambda) = 0$ almost everywhere on $\Lambda\backslash\Lambda_\phi$. Thus, the model is outcome deterministic. Since the same argument holds for every M in which $|\phi\rangle$ appears, the model is measurement noncontextual. ■

The implication in this theorem is strict; i.e., there exist Kochen-Specker noncontextual models that are not maximally ψ-epistemic. An example is provided by the Bell-Mermin model [6,17], in which $\mathcal{M}$ consists of all orthonormal bases in a two-dimensional Hilbert space. The ontic state space of the model consists of the Cartesian product of two copies of the unit sphere $\Lambda = S_2 \times S_2$, and we denote the ontic states as $\lambda = (\vec{\lambda}_1, \vec{\lambda}_2)$, where $\vec{\lambda}_j \in S_2$. For a state $|\psi\rangle$, let $\vec{\psi}$ denote the corresponding Bloch vector. The distribution associated with $|\psi\rangle$ in the ontological model is a product $\mu_\psi(\lambda) = \mu_\psi(\vec{\lambda}_1)\mu_\psi(\vec{\lambda}_2)$, where $\mu_\psi(\vec{\lambda}_1) = \delta(\vec{\lambda}_1 - \vec{\psi})$ is a point measure on $\vec{\psi}$ [18] and $\mu_\psi(\vec{\lambda}_2) = \frac{1}{4\pi}$ is the uniform measure on S_2. It is easy to see that this model is not maximally ψ-epistemic because $\Lambda_\psi \cap \Lambda_\phi = \varnothing$ for distinct $|\psi\rangle$ and $|\phi\rangle$ due to the δ-function term. In fact the model is ψ ontic.

The response functions of the model are

$$\xi_M(\alpha|\lambda) = \Theta(\vec{\alpha}\cdot(\vec{\lambda}_1 + \vec{\lambda}_2)), \tag{5}$$

where Θ is the Heaviside step function

$$\Theta(x) = 1, \qquad x > 0 \tag{6}$$

$$= 0, \qquad x \leq 0. \tag{7}$$

This model is outcome deterministic because Θ only takes the values 0 and 1, and it is measurement noncontextual because the right-hand side of Eq. (5) does not depend on M. It is straightforward to check that the model reproduces the quantum predictions.

In order to understand the connection between maximally ψ-epistemic models and preparation contextuality, we need to describe how (proper) mixtures are represented in ontological models. Assume that, in addition to preparing the pure states in $\mathcal{P}$, the experimenter can also prepare mixtures of them by generating classical randomness (by flipping coins, rolling dice, etc.) with probability distribution p_j and then preparing a different state $|\psi_j\rangle \in \mathcal{P}$ depending on the outcome, resulting in the density operator $\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|$ [19]. The classical randomness is assumed to be independent of the ontic state of the quantum system, so that the distribution over ontic states associated with preparing the ensemble $\mathcal{E} = \{p_j, |\psi_j\rangle\}$ is $\mu_{\mathcal{E}} = \sum_j p_j \mu_{\psi_j}(\lambda)$. An ontological model is preparation noncontextual if $\mu_{\mathcal{E}}(\lambda)$ depends only on the density operator ρ, and not on the specific ensemble decomposition, $\mathcal{E} = \{p_j, |\psi_j\rangle\}$, used to prepare it. Otherwise the model is preparation contextual.

Theorem 2: Suppose an ontological model of $\mathcal{M}$ is not maximally ψ-epistemic, so that there exist states $|\psi\rangle$, $|\phi\rangle \in \mathcal{P}$ such that

$$\int_{\Lambda_\phi} \mu_\psi(\lambda) d\lambda < |\langle\phi|\psi\rangle|^2. \tag{8}$$

Then, if $\mathcal{P}$ includes the states $|\psi^\perp\rangle$, $|\phi^\perp\rangle$ that satisfy $\langle\psi^\perp|\psi\rangle = 0$ and $\langle\phi^\perp|\phi\rangle = 0$, and are in the subspace spanned by $|\psi\rangle$ and $|\phi\rangle$, then the model is also preparation contextual.

Proof.—By Eq. (2), for any M containing $|\phi\rangle$,

$$|\langle\phi|\psi\rangle|^2 = \int_\Lambda \xi_M(\phi|\lambda)\mu_\psi(\lambda)d\lambda \tag{9}$$

$$\geq \int_{\Lambda_\phi} \xi_M(\phi|\lambda)\mu_\psi(\lambda)d\lambda = \int_{\Lambda_\phi} \mu_\psi(\lambda)d\lambda, \tag{10}$$

where the last line follows because $\xi_M(\phi|\lambda) = 1$ almost everywhere on Λ_ϕ. By assumption, the inequality must be strict, so we have

$$\int_{\Lambda_\phi} \xi_M(\phi|\lambda)\mu_\psi(\lambda)d\lambda < \int_\Lambda \xi_M(\phi|\lambda)\mu_\psi(\lambda)d\lambda. \tag{11}$$

This means that there must be a set Ω of ontic states that is disjoint from Λ_ϕ, is assigned nonzero probability by μ_ψ, and is such that $\xi_M(\phi|\lambda) > 0$ for $\lambda \in \Omega$. Now, consider the two mixed preparations: 1. Prepare $|\psi\rangle$ with probability $1/2$ and $|\psi^\perp\rangle$ with probability $1/2$. 2. Prepare $|\phi\rangle$ with probability $1/2$ and $|\phi^\perp\rangle$ with probability $1/2$. The resulting density operators, ρ_1 and ρ_2, satisfy $\rho_1 = \rho_2 = \frac{1}{2}\Pi$, where Π is the projector onto the subspace spanned by $|\psi\rangle$ and $|\phi\rangle$. Let $\Lambda_1 = \Lambda_\psi \cup \Lambda_{\psi^\perp}$ and $\Lambda_2 = \Lambda_\phi \cup \Lambda_{\phi^\perp}$ be the supports of the corresponding distributions, $\mu_1 = \frac{1}{2}(\mu_\psi + \mu_{\psi^\perp})$ and $\mu_2 = \frac{1}{2}(\mu_\phi + \mu_{\phi^\perp})$, in the ontological model. Now, $\Lambda_1 \cap \Omega$ is assigned nonzero probability by μ_1, whereas μ_2 assigns probability zero to Ω. This is because Λ_ϕ is disjoint from Ω by definition and $\mu_{\phi^\perp}$ must assign zero probability to any set of ontic states that assign nonzero probability to $|\phi\rangle$ in a measurement of any orthonormal basis that contains it. Hence μ_1 and μ_2 must be distinct because their supports differ by a set of positive measure. ■

A simple corollary of this theorem is that, whenever the states $|\psi^\perp\rangle$ and $|\phi^\perp\rangle$ are in $\mathcal{P}$ for every $|\psi\rangle, |\phi\rangle \in \mathcal{P}$, then any preparation noncontextual ontological model is also maximally ψ-epistemic. As in the case of Kochen-Specker contextuality, this implication is strict; i.e., there are maximally ψ-epistemic models that are preparation contextual. An example is provided by the Kochen-Specker model [7], which again takes $\mathcal{M}$ to be all orthonormal bases in a two-dimensional Hilbert space. This time, the ontic state space is just a single copy of the unit sphere $\Lambda = S_2$ and the ontic states are unit vectors $\vec{\lambda} \in S_2$. The probability distribution associated with a quantum state $|\psi\rangle$ is

$$\mu_\psi(\vec{\lambda}) = \frac{1}{\pi}\Theta(\vec{\psi}\cdot\vec{\lambda})\vec{\psi}\cdot\vec{\lambda} \tag{12}$$

and the response function associated with a quantum state $|\phi\rangle$ is

$$\xi_M(\phi|\lambda) = \Theta(\vec{\phi}\cdot\vec{\lambda}). \tag{13}$$

It is straightforward to check that this model reproduces the quantum predictions.

The model is maximally ψ-epistemic because $\Lambda_\phi = \{\lambda|\Theta(\vec{\phi}\cdot\vec{\lambda}) = 1\}$ and thus

$$|\langle\phi|\psi\rangle|^2 = \int_\Lambda \xi_M(\phi|\lambda)\mu_\psi(\vec{\lambda})d\lambda \tag{14}$$

$$= \int_\Lambda \Theta(\vec{\phi}\cdot\vec{\lambda})\mu_\psi(\vec{\lambda})d\lambda \tag{15}$$

$$= \int_{\Lambda_\phi} \mu_\psi(\vec{\lambda})d\lambda. \tag{16}$$

On the other hand, the model is preparation contextual as can be seen by considering the two preparations 1. Prepare $|+z\rangle$ with probability $1/2$ and $|-z\rangle$ with probability $1/2$. 2. Prepare $|+x\rangle$ with probability $1/2$ and $|-x\rangle$ with probability $1/2$. Both preparations correspond to the maximally mixed state, but the distributions $\frac{1}{2}(\mu_{+z} + \mu_{-z})$ and $\frac{1}{2}(\mu_{+x} + \mu_{-x})$ are different. In particular, both μ_{+z}

and μ_{-z} are zero on the equator whereas μ_{+x} and μ_{-x} are both nonzero here.

Theorem 1 implies that any proof of the Kochen-Specker theorem is sufficient to establish that maximally ψ-epistemic models are impossible for Hilbert spaces of dimension greater than two. Unlike the proof in Ref. [12], however, this does not establish a bound on how close to maximally ψ-epistemic one can get. Further, the Kochen-Specker theorem allows a finite precision loophole [20] that can be exploited to allow noncontextual theories to get arbitrarily close to quantum statistics, so it seems unlikely that this proof could be made robust against experimental error.

Combining the two theorems also shows that the Kochen-Specker theorem is enough to establish preparation contextuality. While it was known that preparation noncontextuality implies outcome determinism for models of quantum theory [9], it is a novel implication that it also implies measurement noncontextuality. This demonstrates that the distinction between ontic and epistemic quantum states is useful for understanding the relationship between existing no-go theorems.

Finally, the type of preparation contextuality established by our results can be used to prove Bell's theorem. Briefly, if $|\psi\rangle$ and $|\phi\rangle$ are states such that

$$\int_{\Lambda_\phi} \mu_\psi(\lambda) d\lambda < |\langle\phi|\psi\rangle|^2 \tag{17}$$

then we can demonstrate nonlocality using a maximally entangled state $\frac{1}{\sqrt{2}}(|\psi\rangle_A|\psi\rangle_B + |\psi^\perp\rangle_A|\psi^\perp\rangle_B)$. Since the reduced density matrix on Bob's system is

$$\rho = \frac{1}{2}(|\psi\rangle\langle\psi| + |\psi^\perp\rangle\langle\psi^\perp|) \tag{18}$$

$$= \frac{1}{2}(|\phi\rangle\langle\phi| + |\phi^\perp\rangle\langle\phi^\perp|), \tag{19}$$

by the Schrödinger-HJW theorem [21] there are two measurements that Alice can perform, the first of which will collapse Bob's system to $|\psi\rangle$ or $|\psi^\perp\rangle$ with 50:50 probabilities, and the second of which will collapse it to $|\phi\rangle$ or $|\phi^\perp\rangle$ with 50:50 probabilities. However, Theorem 2 establishes that these two ensembles cannot correspond to the same probability distribution over ontic states. Thus, the distribution on Bob's side must depend on the choice of measurement that Alice makes, which implies Bell nonlocality. This argument generalizes the proof of Ref. [4], which showed that local theories would have to be ψ-epistemic. In fact, we see that they would have to be maximally ψ-epistemic. Filling in the formal details of this argument can be done in a similar way to Ref. [4].

While the impossibility of a maximally ψ-epistemic theory clarifies what can be proved about contextuality based on the distinction between ontic and epistemic interpretations of the quantum state without auxiliary assumptions, it is not sufficient to establish the constraints on the size of the ontic state space that follow from having fully ontic quantum states [11]. If one could prove, without auxiliary assumptions, that the support of every distribution in an ontological model must contain a set of states that are not shared by the distribution corresponding to any other quantum state, then these results would follow. Whether this can be proved is an important open question.

We would like to thank Chris Timpson for helpful discussions. O. J. E. M. is supported by the John Templeton Foundation.

*matt@mattleifer.info
http://mattleifer.info
†owen.maroney@philosophy.ox.ac.uk

[1] R. W. Spekkens, Phys. Rev. A **75**, 032110 (2007).
[2] M. F. Pusey, J. Barrett, and T. Rudolph, Nat. Phys. **8**, 476 (2012).
[3] R. Colbeck and R. Renner, Phys. Rev. Lett. **108**, 150402 (2012); M. J. W. Hall, arXiv:1111.6304; L. Hardy, arXiv:1205.1439; M. Schlosshauer and A. Fine, Phys. Rev. Lett. **108**, 260404 (2012).
[4] N. Harrigan and R. W. Spekkens, Found. Phys. **40**, 125 (2010).
[5] J. S. Bell, Physics (Long Island City, N.Y.) **1**, 195 (1964); Reprinted in J. S. Bell, *Speakable and Unspeakable In Quantum Mechanics*, (Cambridge University Press, England, Cambridge, 2004), 2nd ed., pp. 13–21.
[6] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966).
[7] S. Kochen and E. P. Specker, J. Math. Mech. **17**, 59 (1967).
[8] P. G. Lewis, D. Jennings, J. Barrett, and T. Rudolph, Phys. Rev. Lett. **109**, 150404 (2012).
[9] R. W. Spekkens, Phys. Rev. A **71**, 052108 (2005).
[10] M. Leifer, The Quantum Times **6**, 1 (2011), newsletter of the APS Topical Group on Quantum Information.
[11] L. Hardy, Stud. Hist. Phil. Mod. Phys. **35**, 267 (2004); A. Montina, Phys. Rev. A **77**, 022104 (2008).
[12] O. J. E. Maroney, arXiv:1207.6906.
[13] More generally, Λ is a measurable space and states are associated with probability measures ν_ψ. For ease of exposition, we have assumed that Λ is equipped with a canonical measure $d\lambda$ with respect to which all the measures ν_ψ are absolutely continuous, so that we have well-defined densities $\mu_\psi = d\nu_\psi/d\lambda$. This assumption is not strictly necessary. A more general measure theoretic treatment will appear in Ref. [14].
[14] M. Leifer, "Bounds on the Epistemic Interpretation of Quantum States from Contextuality Inequalities" (to be published).
[15] More generally, a distribution is associated with the procedure for preparing a pure state rather than the state itself to allow for preparation contextuality. Assuming preparation noncontextuality for pure states does no harm in the present context because our argument only establishes preparation contextuality for mixed states.
[16] N. Harrigan and T. Rudolph, arXiv:0709.4266.
[17] N. D. Mermin, Rev. Mod. Phys. **65**, 803 (1993).
[18] This does not satisfy the absolute continuity assumption, but a more technical version of Theorem 1 holds without it [14].

[19] For present purposes, it would be sufficient to allow only 50:50 mixtures.

[20] D. A. Meyer, Phys. Rev. Lett. **83**, 3751 (1999); A. Kent, *ibid.* **83**, 3755 (1999); R. K. Clifton and A. Kent, Proc. R. Soc. A **456**, 2101 (2000); J. Barrett and A. Kent, Stud. Hist. Phil. Mod. Phys. **35**, 151 (2004); R. Hermens, *ibid.* **42**, 214 (2011).

[21] E. Schrödinger, Proc. Cambridge Philos. Soc. **32**, 446 (1936); L. P. Hughston, R. Jozsa, and W. K. Wootters, Phys. Lett. A **183**, 14 (1993).