Psi-Epistemic Theories: The Role of Symmetry

Scott Aaronson∗, Adam Bouland†, Lynn Chua‡, George Lowther§

Abstract

Formalizing an old desire of Einstein, “ψ-epistemic theories” try to reproduce the predictions of quantum mechanics, while viewing quantum states as ordinary probability distributions over underlying objects called “ontic states.” Regardless of one’s philosophical views about such theories, the question arises of whether one can cleanly rule them out, by proving no-go theorems analogous to the Bell Inequality. In the 1960s, Kochen and Speck (who first studied these theories) constructed an elegant ψ-epistemic theory for Hilbert space dimension d = 2, but also showed that any deterministic ψ-epistemic theory must be “measurement contextual” in dimensions 3 and higher. Last year, the topic attracted renewed attention, when Pusey, Barrett, and Rudolph (PBR) showed that any ψ-epistemic theory must “behave badly under tensor product.” In this paper, we prove that even without the Kochen-Specker or PBR assumptions, there are no ψ-epistemic theories in dimensions d ≥ 3 that satisfy two reasonable conditions: (1) symmetry under unitary transformations, and (2) “maximum nontriviality” (meaning that the probability distributions corresponding to any two non-orthogonal states overlap). The proof of this result, in the general case, uses some measure theory and differential geometry. On the other hand, we also show the surprising result that without the symmetry restriction, one can construct maximally-nontrivial ψ-epistemic theories in every finite dimension d.

1 Introduction

Debate has raged for almost a century about the interpretation of the quantum state. Although a quantum state evolves in a unitary and deterministic manner according to the Schrödinger equation, measurement is a probabilistic process in which the state is postulated to collapse to a single eigenstate. This is often viewed as an unnatural and poorly-understood process.

Epistemic theories have been proposed as alternatives to standard quantum mechanics. In these theories, a quantum state merely represents probabilistic information about a “real, underlying” physical state (called the ontic state). Perhaps not surprisingly, several no-go theorems have been proven that strongly constrain the ability of these epistemic theories to reproduce the predictions of standard quantum mechanics. Most famously, the Kochen-Specker theorem [1] showed that in Hilbert space dimensions d ≥ 3, no epistemic theory can be both deterministic and “noncontextual” (meaning that whether an eigenstate ψ gets returned as a measurement outcome is independent of which other states are also in the measurement basis). More recently, the Pusey-Barrett-Rudolph (PBR) theorem [2] showed that nontrivial ψ-epistemic theories are inconsistent, if the
ontic distribution for a product state $|\psi\rangle \otimes |\phi\rangle$ is simply the tensor product of the ontic distribution for $|\psi\rangle$ with the ontic distribution for $|\phi\rangle$. Even more recently, papers by Maroney [3] and Leifer and Maroney [4] prove the impossibility of a “maximally $\psi$-epistemic theory,” in which the overlap of the ontic distributions for all non-orthogonal states fully accounts for the uncertainty in distinguishing them via measurements.

In this paper, we study what happens if one drops the Kochen-Specker and PBR assumptions, and merely asks for a $\psi$-epistemic theory in which the ontic distributions overlap for all non-orthogonal states.

Formally, a $\psi$-epistemic theory in $d$ dimensions specifies:

1. A measurable space $\Lambda$, called the ontic space (the elements $\lambda \in \Lambda$ are then the ontic states).

2. A function mapping each quantum state $|\psi\rangle \in H_d$ to a probability measure $\mu_\psi$ over $\Lambda$, where $H_d$ is the Hilbert space in $d$ dimensions.

3. For each orthonormal measurement basis $M = \{\phi_1, \ldots, \phi_d\}$, a set of $d$ response functions $\{\xi_{k,M}(\lambda) \in [0, 1]\}$, which give the probability that an ontic state $\lambda$ would produce a measurement outcome $\phi_k$.

The theory is deterministic if the response functions take values only in $\{0, 1\}$. The Kochen-Specker theorem then states that, in dimensions $d \geq 3$, any deterministic theory must have response functions that depend nontrivially on $M$.

To reproduce the predictions of quantum mechanics, the response functions must satisfy the following two conditions:

\[
\int_{\Lambda} \xi_{k,M}(\lambda) \mu_\psi(\lambda) \, d\lambda = |\langle \phi_k | \psi \rangle|^2, \tag{1}
\]

\[
\sum_{i=1}^{d} \xi_{i,M}(\lambda) = 1 \quad \forall \lambda, M. \tag{2}
\]

Here Equation (1) corresponds to the Born rule of quantum mechanics, while Equation (2) says that the probabilities of the possible measurement outcomes must always sum to 1, even when ontic states are considered individually (rather than as elements of probability distributions). Note that Equations (1) and (2) are logically independent of each other.

The conditions above can easily be satisfied by setting $\Lambda = \mathbb{CP}^{d-1}$, the complex projective space consisting of unit vectors in $H_d$ up to an arbitrary phase, and $\mu_\psi(\lambda) = \delta(\lambda - \psi)$, where $\delta$ is the Dirac delta function, and $\xi_{k,M}(\lambda) = |\langle \phi_k | \lambda \rangle|^2$. But that simply gives an uninteresting restatement of quantum mechanics, since the $\mu_\psi$'s for different $\psi$'s have disjoint supports.\(^1\) Thus, let $\text{Supp}(\mu_\psi) \subseteq \Lambda$ be the support of $\mu_\psi$. Then we call a $\psi$-epistemic theory nontrivial if there exist $\psi \neq \phi$ such that $\mu_\psi$ and $\mu_\phi$ have total variation distance less than 1, i.e.

\[
\frac{1}{2} \int_{\Lambda} |\mu_\psi(\lambda) - \mu_\phi(\lambda)| \, d\lambda < 1. \tag{3}
\]

If $\mu_\psi$ and $\mu_\phi$ have total variation distance less than 1, then we say that $|\psi\rangle$ and $|\phi\rangle$ have “nontrivial overlap”. Otherwise we say they have “trivial overlap”. Note that it’s possible for $|\psi\rangle$ and $|\phi\rangle$ to have trivial overlap even if $\mu_\psi$ and $\mu_\phi$ have intersecting supports (this can happen if $\text{Supp}(\mu_\psi) \cap \text{Supp}(\mu_\phi)$ has measure 0).

\(^1\)Indeed, some authors would not even call this trivial theory “$\psi$-epistemic”; they would insist on calling it “$\psi$-ontic” instead. In this paper, we adopt a uniform definition of $\psi$-epistemic theories, but then distinguish between trivial and nontrivial such theories.
Note also that if $|\psi\rangle$ and $|\phi\rangle$ are orthogonal, then if we set $|\phi_1\rangle = |\psi\rangle$ and $|\phi_2\rangle = |\phi\rangle$, the conditions $|\langle \phi_1 | \psi \rangle| = |\langle \phi_2 | \phi \rangle| = 1$ and $|\langle \phi_2 | \psi \rangle| = |\langle \phi_1 | \phi \rangle| = 0$ imply that $\mu_\psi$ and $\mu_\phi$ have trivial overlap. Hence, we call a theory maximally nontrivial if the overlap is only trivial for orthogonal states: that is, if all nonorthogonal states $|\psi\rangle$, $|\phi\rangle$ have nontrivial overlap.

In a maximally nontrivial theory, some of the uncertainty of quantum measurement is explained by the overlap between the distributions corresponding to nonorthogonal states. Recently Maroney [3] and Leifer and Maroney [4] showed that it is impossible to have a “maximally $\psi$-epistemic theory” in which all of the uncertainty is explained by the overlap of distributions. Specifically, they require that, for all quantum states $|\psi\rangle$, $|\phi\rangle$,

$$\int_{\text{Supp}(|\phi\rangle)} \mu_\psi(\lambda) d\lambda = |\langle \phi | \psi \rangle|^2. \quad (4)$$

Here we are asking for a much weaker condition, in which only some of the uncertainty in measurement statistics is explained by the overlap of distributions, and we do not impose any conditions on the amount of overlap.

Another property that we might like a $\psi$-epistemic theory to satisfy is symmetry. Namely, we call a $\psi$-epistemic theory symmetric if $\Lambda = \mathbb{C}^d$ and the probability distribution $\mu_\psi(\lambda)$ is symmetric under unitary transformations that fix $|\psi\rangle$—or equivalently, if $\mu_\psi$ is a function $f_\psi$ only of $|\langle \psi | \lambda \rangle|$. We stress that this function is allowed to be different for different $\psi$: symmetry only applies to each $\mu_\psi$ individually. This makes our no-go theorem for symmetric theories stronger.

If additionally $\mu_\psi$ is a fixed function $f$ only of $|\langle \psi | \lambda \rangle|$, then we call the theory strongly symmetric. Note that, if a theory is strongly symmetric, then in order to apply a unitary $U$ to a state $|\psi\rangle$, one can simply apply $U$ to the ontic states. So strongly symmetric theories have a clear motivation: namely, they allow us to keep the Schrödinger equation as the time evolution of our system.

In dimension 2, there exists a strongly symmetric and maximally nontrivial theory found by Kochen and Specker [1]. In dimensions $d \geq 3$, Lewis et al. [5] found a nontrivial $\psi$-epistemic theory for all finite $d$. However, their theory is not symmetric and is far from being maximally nontrivial.

In this paper, we first give a construction of a maximally nontrivial $\psi$-epistemic theory for arbitrary $d$. Our theory builds on that of Lewis et al. [5], and was first constructed in a post on MathOverflow [6]. Unfortunately, this theory is rather unnatural and is not symmetric. We then prove that it is impossible to construct a maximally nontrivial theory that is symmetric, for Hilbert space dimensions $d \geq 3$. In other words, if we want maximally nontrivial theories in 3 or more dimensions, then we either need an ontic space $\Lambda$ other than $\Lambda = \mathbb{C}^{d-1}$, or else we need ontic distributions $\mu_\psi$ that “single out preferred directions in Hilbert space,” depending on more than just the inner product $|\langle \psi | \lambda \rangle|$. 

2 Nonsymmetric, Maximally Nontrivial Theory

By considering $\Lambda = \mathbb{C}^{d-1} \times [0, 1]$, Lewis et al. [5] found a deterministic, nontrivial $\psi$-epistemic theory for all finite $d$. They raised as an open problem whether a maximally nontrivial theory exists. In this section, we answer their question in the affirmative. Specifically, we first show that, for any two nonorthogonal states, we can construct a theory such that their probability distributions overlap. We then take a convex combination of such theories to obtain a maximally nontrivial theory.

Lemma 1. Given any two nonorthogonal quantum states $|a\rangle$, $|b\rangle$, there exists a $\psi$-epistemic theory $T(a, b) = (\Lambda, \mu, \xi)$ such that $\mu_a$ and $\mu_b$ have nontrivial overlap. Moreover, for $T(a, b)$, there exists $\varepsilon > 0$ such that $\mu_a$ and $\mu_b$ have nontrivial overlap for all $|a'\rangle$, $|b'\rangle$ such that

$$||a - a'||, ||b - b'|| < \varepsilon.$$
Proof. Our ontic state space will be \( \Lambda = \mathbb{C}^{d-1} \times [0, 1] \). Given an orthonormal basis \( M = \{ \phi_1, \ldots, \phi_d \} \), we first sort the \( \phi_i \)'s in decreasing order of \( \min(|\langle \phi_i | a \rangle|, |\langle \phi_i | b \rangle|) \). Then the outcome of measurement \( M \) on ontic state \((\lambda, p)\) will be the smallest positive integer \( i \) such that

\[
|\langle \phi_1 | \lambda \rangle|^2 + \cdots + |\langle \phi_i - 1 | \lambda \rangle|^2 \leq p \leq |\langle \phi_1 | \lambda \rangle|^2 + \cdots + |\langle \phi_i | \lambda \rangle|^2.
\]

In other words, \( \xi_{\iota, M}(|\lambda|, p) = 1 \) if \( i \) satisfies the above and no \( j < i \) does, and is 0 otherwise. If we assume that \( \mu_\psi(|\lambda|, p) = \delta(|\lambda| - |\psi|) \) for all \( p \in [0, 1] \), then it can be verified that \( T(a, b) \) is a valid \( \psi \)-epistemic theory, albeit so far a trivial one.

We now claim that there exists \( \epsilon > 0 \) such that, for all orthonormal bases \( M = \{ \phi_1, \ldots, \phi_d \} \), there exists \( i \) such that \( |\langle \phi_i | a \rangle| \geq \epsilon \) and \( |\langle \phi_i | b \rangle| \geq \epsilon \). Indeed, by the triangle inequality, we can let \( \epsilon = |\langle a | b \rangle|/d \), and \( \epsilon > 0 \) since \( |\langle a | b \rangle| > 0 \). This means that, for all measurements \( M \) and all \( p \in [0, \epsilon] \), the outcome is always \( i = 1 \) when \( M \) is applied to either of the ontic states \((|a|, p)\) or \((|b|, p)\).

Following Lewis et al. [5], we can “mix” the probability distributions \( \mu_a \) and \( \mu_b \), or have them intersect in the region \( p \in [0, \epsilon] \), without affecting the Born rule statistics for any measurement. Explicitly, we can let

\[
E_{a,b} = \{|a\}, |b\} \times [0, \epsilon],
\]

so that all \( \lambda \in E_{a,b} \) give the same measurement outcome \( \phi_1 \) for all measurements \( M \). Then any probability assigned by \( \mu_a \) or \( \mu_b \) to states within \( E_{a,b} \) can be redistributed over \( E_{a,b} \) without changing the measurement statistics. Thus, we can define \( \mu_a \) such that the weight it originally placed on \(|a| \times [0, \epsilon]\) is now placed uniformly on \( E_{a,b} \). More formally, we set

\[
\mu_a(|\lambda|, x) = \begin{cases} 
\delta(|\lambda| - |a|) & \text{if } x > \epsilon \\
\epsilon \mu_{E_{a,b}}(|\lambda|, x) & \text{if } x \leq \epsilon.
\end{cases}
\]

where \( \mu_{E_{a,b}} \) is the uniform distribution over \( E_{a,b} \). We similarly define \( \mu_b \). This then yields a theory with nontrivial overlap between \(|a|\) and \(|b|\).

Furthermore, suppose we have \(|a'|, |b'|\), such that \(||a - a'||, ||b - b'|| < \frac{\epsilon}{2}\). Then by continuity, we can similarly mix the distributions \( \mu_{a'} \) and \( \mu_{b'} \), or have them intersect each other in the region \( p \in [0, \frac{\epsilon}{2}] \), without affecting any measurement outcome. Note that the procedure of sorting the basis vectors of \( M \) might cause the measurement outcome to change discontinuously. However, this is not a problem since the procedure depends only on \(|a|\) and \(|b|\), which are fixed, and hence occurs uniformly for all \(|a'|\) and \(|b'|\) defined as above.

Lemma 1 implies that for any two nonorthogonal states \(|a|\) and \(|b|\), we can construct a theory where \( \mu_{a'} \) and \( \mu_{b'} \) have nontrivial overlap for all \(||a - a'||, ||b - b'|| < \epsilon\), for some \( \epsilon > 0 \). To obtain a maximally nontrivial theory, such that any two nonorthogonal vectors have probability distributions that overlap, we take a convex combination of such \( \psi \)-epistemic theories.

Given two \( \psi \)-epistemic theories \( T_1 = (\Lambda_1, \mu_1, \xi_1) \) and \( T_2 = (\Lambda_2, \mu_2, \xi_2) \) and a constant \( c \in (0, 1) \), we define the new theory \( cT_1 + (1-c)T_2 = (\Lambda_c, \mu_c, \xi_c) \) by setting \( \Lambda_c = (\Lambda_1 \times \{1\}) \cup (\Lambda_2 \times \{2\}) \) and \( \mu_c = c \mu_1 + (1-c) \mu_2 \). For any \((\lambda, i) \in \Lambda_c\), we then define \( \xi_c \) to equal \( \xi_i \) on \( \Lambda_i \).

The following is immediate from the definitions.

**Lemma 2.** \( cT_1 + (1-c)T_2 \) is a \( \psi \)-epistemic theory. Furthermore, if \( T_1 \) mixes the probability distributions \( \mu_{\psi}, \mu_{\phi} \) of two states \(|a|\) and \(|b|\), and \( T_2 \) mixes \( \mu_{a'} \) and \( \mu_{b'} \), then \( cT_1 + (1-c)T_2 \) mixes both pairs of distributions, assuming \( c \notin \{0,1\} \).

Note that the ontic state space of a convex combination of theories contains a copy of each of the original ontic spaces \( \Lambda_1 \) and \( \Lambda_2 \). If \( \Lambda_1 = \Lambda_2 \), it is natural to ask if we could get away with keeping
only one copy of the ontic state space. Unfortunately the answer in general is no. Suppose that we let \( \Lambda = \Lambda_1 = \Lambda_2 \), let \( \mu_c = c\mu_1 + (1-c)\mu_2 \), and let \( \xi_c = c\xi_1 + (1-c)\xi_2 \). Then the probability of measuring outcome \( i \) under measurement \( M \) and ontic distribution \( \mu_{c\psi} \) is

\[
\int_{\Lambda} (c\xi_{1,M}(\lambda) + (1-c)\xi_{2,M}(\lambda)) (c\mu_{1,\psi}(\lambda) + (1-c)\mu_{2,\psi}(\lambda)) \, d\lambda
\]

which will not in general reproduce the Born rule due to unwanted cross terms. This is why it is necessary to keep two copies of the ontic state space when taking a convex combination of theories.

Using Lemmas 1 and 2, we now construct a maximally nontrivial \( \psi \)-epistemic theory. Let \( T(a,b) \) be the theory returned by Lemma 1 given \( |a\rangle, |b\rangle \in H_d \). Also, for all positive integers \( n \), let \( A_n \) be a \( 1/n \)-net for \( H_d \), that is, a finite subset \( A_n \subseteq H_d \) such that for all \( |a\rangle \in H_d \), there exists \( |a'| \in A_n \) satisfying \( ||a-a'|| < 1/n \). By making small perturbations, we can ensure that \( \langle a|b \rangle \neq 0 \) for all \( |a\rangle, |b\rangle \in A_n \). Then our theory \( T \) is defined as follows:

\[
T = \frac{6}{\pi^2} \sum_{n=1}^{\infty} \frac{1}{n^2} \left( \frac{1}{|A_n|^2} \sum_{a,b \in A_n} T(a,b) \right). 
\]  

(Of course, in place of \( 6/(\pi^2 n^2) \), we could have chosen any infinite sequence summing to unity.) This yields a maximally nontrivial theory, since it can be verified that \( \mu_a \) and \( \mu_b \) have nontrivial overlap for all nonorthogonal states \( |a\rangle \) and \( |b\rangle \). Note that the ontic space is now \( \mathbb{C}P^{d-1} \times [0,1] \times \mathbb{N} \), which has the same cardinality as \( \mathbb{C}P^{d-1} \). It is thus possible to map this theory into a theory that uses \( \Lambda = \mathbb{C}P^{d-1} \) as its ontic space, using a bijection between the ontic spaces. However, it is clear that under such a bijection the theory becomes less symmetric: the quantum state \( |a\rangle \) no longer has any association with the state \( |a\rangle \) in the ontic space, and the measure is also very unnatural.

### 3 Nonexistence of Symmetric, Maximally Nontrivial Theories

We now turn to showing that it is impossible to construct a symmetric maximally nontrivial theory, in dimensions \( d \geq 3 \). Recall that a theory is called symmetric if

1. \( \Lambda = \mathbb{C}P^{d-1} \), and
2. for any quantum state \( |\psi\rangle \), the associated probability distribution \( \mu_{\psi} \) is invariant under unitary transformations that preserve \( |\psi\rangle \).

Specifically, if \( U \) is a unitary transformation such that \( U|\psi\rangle = |\psi\rangle \), then we require that \( \mu_{U\psi}(U\lambda) = \mu_{\psi}(\lambda) \). This implies that \( \mu_{\psi}(\lambda) \) is a function only of \( \langle \psi|\lambda \rangle^2 \): that is,

\[
\mu_{\psi}(\lambda) = f_{\psi} (|\langle \psi|\lambda \rangle|^2)
\]

for some nonnegative function \( f_{\psi} \). In other words, the probability measure \( \mu_{\psi} \), associated with state \( \psi \) must be a measure \( \nu_{\psi} \) on the unit interval which has been “stretched out” onto \( H_d \) over curves of constant \( |\langle \psi|\lambda \rangle|^2 \). If additionally we assume that for any \( U \), \( \mu_{U\psi}(U\lambda) = \mu_{\psi}(\lambda) \), or equivalently that \( \mu_{\psi}(\lambda) = f (|\langle \psi|\lambda \rangle|^2) \) for some fixed nonnegative function \( f \), the theory is called strongly symmetric.

In this section, we will first prove a number of facts about symmetric, maximally nontrivial theories in general. Using these facts, we will then show that no strongly symmetric, maximally nontrivial theory exists in dimension 3 or higher. Restricting to the strongly symmetric case will make the proof considerably easier. Later we will show how to generalize to the “merely” symmetric case.
Figure 1: Diagram of maximally nontrivial theory in $d = 2$ on the Bloch sphere. The shaded region corresponds to $\text{Supp}(\mu_\psi)$.

As mentioned earlier, Kochen and Specker proved that a strongly symmetric, maximally nontrivial $\psi$-epistemic theory exists in dimension $d = 2$ [1]. In their theory, which is illustrated in Figure 1, the ontic space is $\Lambda = \mathbb{C}P^1$, and the response functions for a given basis $M = \{\phi_1, \phi_2\}$ are

\[
\begin{align*}
\xi_{1,M}(\lambda) &= 1 \text{ if } |\langle \lambda|\phi_1 \rangle| \geq |\langle \lambda|\phi_2 \rangle| \text{ or } 0 \text{ otherwise,} \\
\xi_{2,M}(\lambda) &= 1 \text{ if } |\langle \lambda|\phi_2 \rangle| \geq |\langle \lambda|\phi_1 \rangle| \text{ or } 0 \text{ otherwise.}
\end{align*}
\]

Hence the response functions are deterministic and partition the ontic space. Intuitively, the result of a measurement on any ontic state is the state in the measurement basis to which it is closest. For any quantum state $|\psi\rangle \in H_2$, the probability distribution over $\Lambda$ is given by

\[
\mu_\psi(\lambda) = \begin{cases} 
\frac{2}{\pi} \left( |\langle \lambda|\psi \rangle|^2 - \frac{1}{2} \right) & \text{if } |\langle \lambda|\psi \rangle|^2 > \frac{1}{2} \\
0 & \text{otherwise.}
\end{cases}
\]

It can readily be verified that this theory satisfies the conditions in Equations 1 and 2 above, and has the properties of being strongly symmetric and maximally nontrivial. It is also maximally $\psi$-epistemic in the sense described by Maroney [3] and Maroney and Leifer [4].

Given a measurement outcome $\psi$ and a basis $M$ containing $\psi$, we define their nonzero set to be the set of ontic states $\lambda$ such that the response function $\xi_{\psi,M}(\lambda)$ gives a nonzero probability of returning $\psi$ when $M$ is applied:

\[
\text{Nonzero}(\xi_{\psi,M}) = \{\lambda : \xi_{\psi,M}(\lambda) \neq 0\}.
\]

Clearly in any $\psi$-epistemic theory, $\text{Supp}(\mu_\psi) \subseteq \text{Nonzero}(\xi_{\psi,M})$ for any measurement basis $M$ that contains $\psi$, because the state $|\psi\rangle$ must return measurement outcome $\psi$ with probability 1 for any such $M$. Harrigan and Rudolph [9] call a $\psi$-epistemic theory deficient if there exists a quantum state $|\psi\rangle$ and measurement basis $M$ containing $\psi$ such that

\[
\text{Supp}(\mu_\psi) \subsetneq \text{Nonzero}(\xi_{\psi,M}).
\]

In other words, a theory is deficient if there exists an ontic state $\lambda$ such that $\lambda$ has a nonzero probability of giving the measurement outcome corresponding to $|\psi\rangle$ for some $M$, even though $\lambda \notin \text{Supp}(\mu_\psi)$. Informally, this can be thought of as a “one-sided friendship” between $|\psi\rangle$ and $\lambda$.

It was first pointed out by Rudolph [8], and later shown by Harrigan and Rudolph [9], that theories in dimension $d \geq 3$ must be deficient. In this section, we prove that as a result of
deficiency, it is impossible to have a symmetric, maximally nontrivial theory with \( d \geq 3 \). We derive a contradiction by showing that if the theory is maximally nontrivial, then there exist orthogonal states \(|\psi\rangle,|\phi\rangle\), and a measurement basis \( M \) containing \(|\psi\rangle\), such that if \(|\phi\rangle\) is measured, then the outcome \(|\psi\rangle\) is returned with nonzero probability, contradicting the laws of quantum mechanics. We do this by choosing \(|\psi\rangle\) orthogonal to \(|\phi\rangle\) such that its associated response function \( \xi_{\psi,M} \) has nonzero measure in the deficient region, and proving that there exists \(|\phi\rangle\) such that \( \text{Supp}(\mu_\phi) \) overlaps with \( \text{Nonzero}(\xi_{\psi,M}) \) in the deficient region. Thus the deficiency imposed by the symmetry assumption makes it impossible for a maximally nontrivial theory to exist in \( d \geq 3 \).

We start with a few preliminary results on symmetric, maximally nontrivial theories. As stated previously, we know that \( \mu_\psi \) is generated by stretching a probability measure \( \nu_\psi \) on the unit interval over \( H_d \) along lines of constant \( |\langle \psi|\lambda \rangle|^2 \). By the Lebesgue decomposition theorem, \( \nu_\psi \) can be written uniquely as a sum of two measures \( \nu_{\psi,C} \) and \( \nu_{\psi,S} \), where \( \nu_{\psi,C} \) is absolutely continuous with respect to the Lebesgue measure over the unit interval, and \( \nu_{\psi,S} \) is singular with respect to that measure. Here we say \( \nu_{\psi,C} \) is “absolutely continuous” with respect to the Lebesgue measure, we mean that it assigns zero measure to any set of Lebesgue measure zero. When we say \( \nu_{\psi,S} \) is “singular”, we mean that its support is confined to a set of Lebesgue measure zero.

By the Radon-Nikodym theorem, due to its absolute continuity \( \nu_{\psi,C} \) has a probability density function \( g_\psi(x) \) that is a function, not a pseudo-function or delta function. To simplify our analysis, first we will show that the absolutely continuous part of the distribution is non-zero. We will later use only the absolutely continuous part of the distribution in our proof.

**Lemma 3.** For any state \(|\psi\rangle\) in a symmetric, maximally nontrivial theory, \( \nu_{\psi,C} \) has positive total measure, i.e. \( \int_0^1 g_\psi(x)dx > 0 \).

**Proof.** Suppose there exists a state \(|\psi\rangle\) such that \( \nu_{\psi,C} \) is 0 almost everywhere. Then only the singular part \( \nu_{\psi,S} \) of \( \nu_\psi \) has positive measure on \([0,1]\). Consider any state \(|\phi\rangle\) which is not orthogonal to \(|\psi\rangle\). By the Lebesgue decomposition theorem, \( \nu_\phi \) can also be written uniquely as a sum of its absolutely continuous and singular parts, \( \nu_{\phi,C} \) and \( \nu_{\phi,S} \), with respect to the Lebesgue measure over \([0,1]\). Since \( \nu_{\psi,S} \) has support on a set of measure zero, \( \mu_{\phi,C} \) and \( \nu_{\psi,S} \) have trivial overlap, where \( \mu_{\omega,A} \) denotes the distribution over \( \mathbb{C}^{d-1} \) obtained by stretching out \( \nu_{\omega,A} \) over lines of constant \( |\langle \omega|\lambda \rangle|^2 \). Also, \( \mu_{\psi,S} \) and \( \mu_{\phi,S} \) have trivial overlap because the intersection of surfaces of constant \( |\langle \lambda|\psi \rangle|^2 \) and \( |\langle \lambda|\phi \rangle|^2 \) has dimension \( 2d-4 \) (two less than the dimension of \( \mathbb{C}^{d-1} \)), so \( \text{Supp}(\mu_{\psi,S}) \cap \text{Supp}(\mu_{\phi,S}) \) must have measure zero with respect to either measure. Thus the variation distance between \( \mu_\psi \) and \( \mu_\phi \) is 1, contradicting maximum nontriviality.

From now on, we will assume \( \mu_\psi \) is generated only from the absolutely continuous part \( \nu_{\psi,C} \), so that \( \mu_\psi \) has as probability density function \( f_\psi(|\langle \psi|\lambda \rangle|^2) \) where \( f_\psi \) is a function, not a pseudo-function. We can do this without loss of generality, as our proof will not depend on the normalization of the probability distributions.

Next, let the distance between two states \( \psi \) and \( \phi \) be defined by their scaled radial distance (also called the Fubini-Study metric):

\[
||\psi - \phi|| = \frac{2}{\pi} \arccos(|\langle \psi|\phi \rangle|).
\]

For any state \(|\psi\rangle\in H_d\), with probability distribution \( \mu_\psi(\lambda) = f_\psi(|\langle \psi|\lambda \rangle|^2) \), we define the radius of \( \mu_\psi \) to be the distance between \(|\psi\rangle\) and the furthest away state at which \( \mu_\psi \) has substantial density:

\[
r_\psi = \sup \left\{ r : \forall \delta > 0 \int_{|\lambda - \delta| < ||\psi - \lambda|| < r} \mu_\psi(\lambda) d\lambda > 0 \right\}.
\]  

(14)
Lemma 4. For a symmetric theory, given any two states $|\psi\rangle, |\phi\rangle$, we have $||\psi - \phi|| \geq r_\psi + r_\phi$ if and only if $|\psi\rangle$ and $|\phi\rangle$ have trivial overlap.

Proof. Suppose that $||\psi - \phi|| \geq r_\psi + r_\phi$ but $|\psi\rangle$ and $|\phi\rangle$ have nontrivial overlap. Then $\text{Supp}(\mu_\psi) \cap \text{Supp}(\mu_\phi)$ has nonzero measure, and for any $\lambda$ in that set, the triangle inequality implies that $r_\psi + r_\phi \geq ||\psi - \lambda|| + ||\phi - \lambda|| \geq ||\psi - \phi||$. Thus $r_\psi$ and $r_\phi$ satisfy $r_\psi + r_\phi = ||\psi - \phi||$, which is a contradiction since $||\psi - \lambda|| + ||\phi - \lambda|| = ||\psi - \phi||$ only on a set of measure zero.

Now suppose that $||\psi - \phi|| < r_\psi + r_\phi$. Consider the region of $\lambda$ such that $||\psi - \phi|| \leq ||\psi - \lambda|| + ||\phi - \lambda|| < r_\psi + r_\phi$. From Equation (14), we can choose from this region a subregion of $\lambda$ such that both $||\psi - \lambda||$ and $||\phi - \lambda||$ are less than and arbitrarily close to $r_\psi$ and $r_\phi$ respectively, in order that this subregion has nonzero measure with respect to both $\mu_\psi$ and $\mu_\phi$. This would then imply that $|\psi\rangle$ and $|\phi\rangle$ have nontrivial overlap. $\Box$

Corollary 1. If $|\psi\rangle$ and $|\phi\rangle$ are orthogonal, then $r_\psi + r_\phi \leq 1$.

Proof. If $\langle \psi | \phi \rangle = 0$, then $||\psi - \phi|| = 1$. Since any orthogonal $|\psi\rangle$ and $|\phi\rangle$ have trivial overlap, Lemma 4 implies that $r_\psi + r_\phi \leq 1$. $\Box$

Lemma 5. Given any maximally nontrivial and symmetric theory in $d \geq 3$, for any state $|\psi\rangle \in H_d$, we have $r_\psi = \frac{1}{d}$.

Figure 2: From left to right: pictorial representations of the proof that $r_\psi \leq \frac{1}{d}$ in dimension 3, the proof that $r_\psi = \frac{1}{d}$, and the form of the $\mu_\psi$’s that we ultimately deduce (with $r_\psi = \frac{1}{d}$ for all $|\psi\rangle \in H_d$). The shaded regions are the supports of the respective probability distributions.

Proof. We first show that $r_\psi \leq \frac{1}{d}$ for all $|\psi\rangle \in H_d$, which we illustrate in the left side of Figure 2 for the case where $d = 3$. Suppose there exists $|\psi\rangle$ such that $r_\psi = \frac{1}{d} + \varepsilon$ for some $\varepsilon > 0$. From Corollary 1, for all $|\phi\rangle$ orthogonal to $|\psi\rangle$, we have $r_\phi \leq \frac{1}{d} - \varepsilon$. In dimension $d \geq 3$, there exist nonorthogonal states $|\phi\rangle, |\phi'\rangle$ such that $\langle \psi | \phi \rangle = \langle \psi | \phi' \rangle = 0$, and $|\phi\rangle \neq |\phi'\rangle$. Then $r_\phi + r_\phi' \leq 1 - 2\varepsilon$. If we choose $|\phi\rangle, |\phi'\rangle$ such that $1 - 2\varepsilon < ||\phi - \phi'|| < 1$, then from Lemma 4, we have that $\mu_\phi$ and $\mu_\phi'$ trivial overlap even though $\langle \phi | \phi' \rangle \neq 0$. This contradicts the theory being maximally nontrivial.

We now show that $r_\psi \geq \frac{1}{d}$ for all $|\psi\rangle \in H_d$, as illustrated in the center of Figure 2. Suppose there exists $|\psi\rangle$ such that $r_\psi = \frac{1}{d} - \varepsilon$ for some $\varepsilon > 0$. Since $r_\phi' \leq \frac{1}{d}$ for all $|\phi'\rangle \in H_d$, thus $r_\psi + r_\phi' \leq 1 - \varepsilon$. If we choose $|\psi\rangle, |\phi'\rangle$ such that $1 - \varepsilon < ||\psi - \phi'|| < 1$, then $\mu_\psi$ and $\mu_\phi'$ have trivial overlap from Lemma 4 even though $\langle \phi | \phi' \rangle \neq 0$. This again contradicts maximum nontriviality. $\Box$

This immediately implies the following:
Corollary 2. In dimensions \( d \geq 3 \), a symmetric \( \psi \)-epistemic theory is maximally nontrivial if and only if for any state \( |\psi⟩ \) and for all \( \delta > 0 \) the measure \( \mu_\psi \) integrated over the following region is non-zero:

\[
\left\{ \lambda : \frac{1}{2} \leq |⟨\psi|\lambda⟩|^2 \leq \frac{1}{2} + \delta \right\}
\]

(15)

Moreover, \( \text{Supp}(\mu_\psi) \) has measure zero on the set of \( \lambda \) such that \( |⟨\psi|\lambda⟩|^2 < \frac{1}{2} \).

Proof. By Lemma 5, for any state \( |\psi⟩ \) we have \( r_\psi = \frac{1}{2} \). By rewriting the distance between states in terms of their inner product, the Corollary follows from the definition of \( r_\psi \) in Equation 14. \( \square \)

In Lemma 5, we showed that the radius \( r_\psi \) of every state \( \psi \) in a maximally nontrivial symmetric theory is \( \frac{1}{2} \). We now use this to show that a certain set of ontic states is deficient. Recall that we say a set \( S \) is deficient for measurement \( M \) if \( S \) is not in \( \text{Supp}(\mu_\phi) \) for any \( \phi \in M \).

Corollary 3. Given any symmetric, maximally nontrivial \( \psi \)-epistemic theory in \( d \geq 3 \), for any measurement basis \( M = \{\phi_i\}_{i=1}^d \), the region

\[
R_M = \left\{ \lambda : |⟨\phi_i|\lambda⟩|^2 < \frac{1}{2}, i = 1, \ldots, d \right\}
\]

is deficient except on a set of measure zero. (Note that, by elementary geometry, \( R_M \) has positive measure if and only if \( d \geq 3 \).)

Proof. By Corollary 2, for all \( i = 1, \ldots, d \), the set \( \text{Supp}(\phi_i) \) must have measure zero over the region \( R_M \). However, Equation 2 implies that any \( \lambda \in R_M \) must be in \( \text{Nonzero}(\xi_{i,M}) \) for some \( i \) even if it is not in \( \text{Supp}(\phi_i) \). This means that \( R_M \) is deficient except possibly on a set of measure zero. \( \square \)

In general, deficiency occurs in any theory in \( d \geq 3 \) even without the symmetry assumption, as proved by Harrigan and Rudolph [9] using the Kochen-Specker theorem [1]. In Corollary 3, we showed that symmetry implies a specific type of deficiency.

We will now leverage the deficiency described in Corollary 3 to prove that symmetric, maximally nontrivial theories do not exist in dimension \( d \geq 3 \). For ease of presentation, we will first prove the theorem assuming that the theory is strongly symmetric. Recall that a theory is strongly symmetric if \( \mu_\psi \) is a fixed function of \( |⟨\psi|\lambda⟩| \). So not only is the distribution for each state \( \psi \) symmetric about \( \psi \), but furthermore the distribution for every state is identical. Then, in the next section, we will generalize the proof to the “merely” symmetric case.

To show that no strongly symmetric, maximally nontrivial theory exists, we first prove two simple results for \( \psi \)-epistemic theories in general. These results will help us to derive a contradiction for strongly symmetric, maximally nontrivial theories.

Lemma 6. Given any two orthogonal states \( |\phi⟩ \) and \( |\psi⟩ \), the set \( \text{Supp}(\mu_\phi) \cap \text{Nonzero}(\xi_{\psi,M}) \) has measure zero for all measurements \( M \) that contain \( \psi \).

Proof. Suppose to the contrary that \( \text{Supp}(\mu_\phi) \cap \text{Nonzero}(\xi_{\psi,M}) \) has positive measure for some measurement \( M \) containing \( \psi \). Then by definition, if the state \( |\phi⟩ \) is measured using \( M \), the outcome corresponding to \( |\psi⟩ \) is returned with nonzero probability. But since \( |⟨\psi|\phi⟩|^2 = 0 \), this contradicts the Born rule (Equation (1)). \( \square \)

Lemma 7. For any \( \alpha \in \Lambda \), let \( B_\varepsilon(\alpha) = \{\lambda : ||\lambda - \alpha|| < \varepsilon\} \) be an \( \varepsilon \)-ball around \( \alpha \), for some \( \varepsilon > 0 \). Given a measurement basis \( M = \{\phi_i\}_{i=1}^d \), there exists some \( j \) such that

\[
\int_{B_\varepsilon(\alpha)} \xi_{j,M}(\lambda) \, d\lambda > 0.
\]

(16)
Proof. For any such \( \alpha \in \Lambda \) the following holds,
\[
\int_{B_\varepsilon(\alpha)} \sum_{i=1}^{d} \xi_{i,M}(\lambda) \, d\lambda = \int_{B_\varepsilon(\alpha)} 1 \, d\lambda > 0.
\]

This then implies that there exists some \( j \) such that
\[
\int_{B_\varepsilon(\alpha)} \xi_{j,M}(\lambda) \, d\lambda > 0.
\]

Figure 3: Pictorial representation of the deficiency region for \( d = 3 \). The shaded regions are the supports of the respective probability measures, and the middle unshaded region \( R_M \) is deficient.

Using these two results, we can now prove that in dimension \( d \geq 3 \), there exists no strongly symmetric, maximally nontrivial \( \psi \)-epistemic theory.

**Theorem 1.** There exists no strongly symmetric, maximally nontrivial \( \psi \)-epistemic theory in dimension \( d \geq 3 \).

**Proof.** Suppose we have a symmetric, maximally nontrivial theory in dimension \( d \geq 3 \), and we fix a measurement basis \( M = \{ \phi_i \}_{i=1}^{d} \). From Corollary 3, there exists a deficiency region given by
\[
R_M = \left\{ \lambda : |\langle \phi_i | \lambda \rangle|^2 < \frac{1}{2}, i = 1, \ldots, d \right\},
\]

perhaps minus a set of measure zero. This is illustrated in Figure 3 for the case where \( d = 3 \).

Consider \( |\alpha\rangle = \frac{1}{\sqrt{d}} (|\phi_1\rangle + \cdots + |\phi_d\rangle) \), which is contained in the deficiency region. Given \( \varepsilon > 0 \), let \( B_\varepsilon(\alpha) = \{ \lambda : ||\lambda - \alpha|| < \varepsilon \} \) be the \( \varepsilon \)-ball around \( |\alpha\rangle \). We choose \( \varepsilon \) such that \( B_\varepsilon(\alpha) \) is contained in \( R_M \). From Lemma 7, there exists some \( j \) such that \( B := B_\varepsilon(\alpha) \cap \text{Nonzero}(\xi_{j,M}) \) has nonzero measure. Without loss of generality, we assume that \( j = 1 \).

Let \( \nu \) be the measure obtained by averaging \( \mu_\psi \) over all states \( |\psi\rangle \) orthogonal to \( |\phi_1\rangle \), and let \( A \) be the set of all \( \lambda \) such that \( |\langle \phi_1 | \lambda \rangle|^2 < \frac{1}{2} \). Since the theory is strongly symmetric, \( \nu \) must be a function only of \( |\langle \phi_1 | \lambda \rangle|^2 \). Moreover, each of the measures \( \mu_\psi \) assigns positive measure to the region of states \( \lambda \) such that \( |\langle \psi | \lambda \rangle|^2 \) is close to \( \frac{1}{2} \), hence the averaged measure \( \nu \) assigns positive measure to every open subset of \( A \), and therefore in particular to \( B \). This contradicts Lemma 6, which implies that each of the averaged measures \( \mu_\psi \) must assign zero measure to \( \text{Nonzero}(\xi_{1,M}) \) and hence \( B \).
3.1 Proof of Generalized No-Go Theorem

We now generalize our proof of Theorem 1 to the “merely” symmetric case, where the probability distributions $\mu_\psi$ can vary with $\psi$. First note that our previous proof does not immediately carry over. Since the probability distributions can vary as $\psi$ changes, it is possible that the distributions for states orthogonal to $\phi_1$ might be able to “evade” the set $B$ in the proof of Theorem 1 which returns answer $\phi_1$ under measurement $M$, while maintaining some density near their outer radii.

To see how this might occur, consider the following one dimensional example: Let $\Lambda = \mathbb{R}$ be the real line. Construct $B \subseteq [0,1]$ to be a “fat Cantor set” on $[0,1]$ as follows. Initially set $B = [0,1]$. In step 1, remove the middle $1/4$ of this interval, so that $B = [0,\frac{3}{8}] \cup \left[\frac{5}{8},1\right]$. At the $i$th step, remove the middle $\frac{1}{2^i}$ of each of the $2^i$ remaining intervals. Continue indefinitely. The resulting set $B$ is called a “fat Cantor set” because it is nowhere dense (so contains no intervals), yet has positive Lebesgue measure on $[0,1]$.

For each point $x \in \mathbb{R}$, let $\mu_x$ be the uniform distribution on $[x-1,x+1]$ with $B$ removed. Then $\mu_x$ is absolutely continuous with respect to the Lebesgue measure for all $x \in \mathbb{R}$, and furthermore has positive measure on $[1+x-\varepsilon,1+x]$ for all $\varepsilon > 0$. However, despite the fact that $B$ has positive measure, the distributions $\mu_x$ never intersect $B$. The worry is that our distributions in $\mathbb{CP}^{d-1}$ could likewise evade the set $B$ in our proof, foiling our contradiction. This worry is related to a variant of the Kakeya/Besicovitch problem, as we discuss in Section 4.

We can extend Theorem 1 without solving a Kakeya-like problem, but to do so we will need a result about the differential geometry of $\mathbb{CP}^{d-1}$. Interestingly, we will use the fact that we are working in a complex Hilbert space; we believe the proof could be adapted to a real Hilbert space, but it would be much less convenient.

Discussing the differential geometry of $\mathbb{CP}^{d-1}$ is easiest if we first to pick a gauge for $\mathbb{CP}^{d-1}$, that is, if we pick a representative from each equivalence class of vectors which differ only by a global phase. We use the following gauge: let $|\alpha\rangle = \frac{1}{\sqrt{d}}(|\phi_1\rangle + \cdots + |\phi_d\rangle)$. For each equivalence class, we pick a representative $u$ such that $\langle \alpha | u \rangle$ is real and positive. This uniquely identifies representatives for all equivalence classes of states, except those orthogonal to $\alpha$. Moreover, this way of choosing a gauge is continuous and smooth near $\alpha$; more precisely, equivalence classes which are close to one another have representatives which are also close to one another. This allows us to integrate over the manifold near $\alpha$ using these representatives. Using this gauge, we now prove the following.

Lemma 8. Let $M$ be a measurement basis $\{\phi_i\}$, let $|\alpha\rangle$ be defined as above, and let $d \geq 3$. Then there exist $d$ vectors $u_1 \ldots u_d$ in $\mathbb{CP}^{d-1}$ such that

- $\langle u_i | \phi_i \rangle = 0$ for all $i$.
- $\langle u_i | \alpha \rangle = \frac{1}{\sqrt{2}}$ for all $i$.
- The tangent vectors $t_i$ to the geodesics from $u_i$ to $\alpha$ are linearly independent at $\alpha$ when the tangent space is viewed as a real vector space.

Proof. Let $a = \sqrt{\frac{d}{2(d-1)^2}}$ and $b = \sqrt{\frac{d-2}{4(d-1)}}$. Then we define $u_1, \ldots u_{d-2}$ as follows:

\[ u_i = \left( \sum_{j \neq i} a |\phi_j\rangle \right) + ib |\phi_{i+1}\rangle - ib |\phi_{i+2}\rangle, \]
For the last two vectors, we set

$$u_{d-1} = \left( \sum_{j \neq d-1} a|\phi_j\rangle \right) + ib|\phi_d\rangle - ib|\phi_1\rangle,$$

$$u_d = \left( \sum_{j \neq d} a|\phi_j\rangle \right) + b|\phi_1\rangle - b|\phi_2\rangle.$$

Note that the coefficients in $u_d$ are all real, unlike for the other $d-1$ vectors. It is straightforward to verify that $\langle u_i|\phi_i\rangle = 0$ and $\langle u_i|\alpha\rangle = \frac{1}{\sqrt{2}}$ for all $i$. Furthermore, we can compute the tangent vectors $t_i$ as follows. The geodesics from $|u_i\rangle$ to $|\alpha\rangle$ in the Fubini-Study metric can be parameterized by

$$\gamma(t) = \cos(t)|u_i\rangle + \sin(t)|\alpha\rangle$$

These geodesics lie entirely within our choice of gauge. Therefore $t_i$ is the projection of $\gamma'(t)|_{t=\pi/2}$ onto the plane orthogonal to $\alpha$, which is

$$t_i = u_i - \langle \alpha|u_i\rangle \alpha$$

Since our gauge is fixed, there is no ambiguity that $u_i$ or $t_i$ could be multiplied by a global phase.

We now verify that the $t_i$’s are linearly independent. Suppose that $c_1t_1 + \ldots + c_dt_d = 0$, with $c_i$ real. Note that $\langle \alpha|u_i\rangle \alpha$ has all real coefficients, so a coefficient of $t_i$ is imaginary if and only if the corresponding coefficient of $u_i$ is imaginary. Since $c_1t_1 + \ldots + c_dt_d = 0$, in particular the imaginary terms in $|\phi_i\rangle$ must sum to zero for all $i$. For $i = 3\ldots d$, only the terms $c_{i-2}t_{i-2}$ and $c_{i-1}t_{i-1}$ contain imaginary multiples of $|\phi_i\rangle$. Hence this constraint implies $c_{i-2} = c_{i-1}$. Additionally, $c_1t_1$ is the only term containing an imaginary multiple of $|\phi_2\rangle$, so we must have $c_1 = 0$. Therefore $c_1 = c_2 = \ldots = c_{d-2} = 0$. Since $c_{d-1}t_{d-1}$ is the only term containing an imaginary multiple of $|\phi_1\rangle$, we must have $c_{d-1} = 0$, and hence $c_d = 0$ as well. Therefore the $t_i$’s are linearly independent.

Note that in a real Hilbert space, the analogous version of Lemma 8 is false because the dimension of the tangent space at $\alpha$ is only $d - 1$. In a complex Hilbert space the dimension of the tangent space is $2d - 2$, so the tangent space can contain $d$ linearly independent vectors assuming $d \geq 2$.

We now show that Lemma 8 implies the existence of a set $B$ of positive measure, on which every $\mu_{u_i}$ is “equivalent” to the Lebesgue measure, in the sense that if $S \subseteq B$ has positive Lebesgue measure, then $S$ has positive measure under each $\mu_{u_i}$.

**Lemma 9.** Let $u_i$ and $|\alpha\rangle$ be as defined in Lemma 8. Then there exists a set $B$ in the neighborhood of $|\alpha\rangle$, of positive Lebesgue measure, such that the $\mu_{u_i}$ are equivalent to the Lebesgue measure on $B$.

**Proof.** By our symmetry condition, each $\mu_{u_i}$ is a product measure between a measure $\nu_{u_i}$ on $[0,1]$ and a uniform measure on surfaces of constant $|\langle u_i|\lambda\rangle|$. Note that the normal vectors to surfaces of constant $|\langle u_i|\lambda\rangle|$ at $|\alpha\rangle$ are the tangent vectors $t_i$ defined in Lemma 8, and are hence linearly independent. Since $\mathbb{C}P^{d-1}$ is a smooth Riemannian manifold, these $t_i$ form a local coordinate system in the $\varepsilon$ neighborhood of $|\alpha\rangle$, which we denote $N_{\varepsilon}(|\alpha\rangle)$. If we associate a coordinate $x_i$ with each $t_i$, the integral over $N_{\varepsilon}(|\alpha\rangle)$ can be parameterized as

$$\int g(x_1, \ldots x_d, y_1 \ldots y_{d-2})dx_1 \ldots dx_d dy_1 \ldots dy_{d-2}$$

12
Here \( g(x_1...x_d) \) is the square root of the metric, which is strictly positive in the neighborhood of \(|\alpha\rangle\).

Also, \( dx_i \) is the Lebesgue integral over the coordinate \( x_i \), and the \( y_i \) are coordinates corresponding to the remaining \( d - 2 \) dimensions of the space.

Now consider the set

\[
B = N_\varepsilon(|\alpha\rangle) \cap \text{Supp}(\mu_{u_1}) \cap \text{Supp}(\mu_{u_2}) \cap \ldots \cap \text{Supp}(\mu_{u_d})
\]

Trivially each of the \( \mu_{u_i} \)'s is equivalent to the Lebesgue measure on \( B \). Note that each \( \text{Supp}(\mu_{u_i}) \) is a union of surfaces \( S_i \) of constant \( \langle \lambda | u_i \rangle \) which are perpendicular to \( t_i \) at \( \alpha \). If \( \varepsilon \) is sufficiently small these surfaces have negligible curvature, so they look like orthogonal hyperplanes in the \( x_i \) coordinate system. Let \( \varepsilon_i \) be the Lebesgue measure on \( \text{Supp}(\mu_{u_i}) \cap B \). If the surfaces \( S_i \) had zero curvature, the Lebesgue measure of \( B \) would simply be the product of the measures \( \prod_i \varepsilon_i \), since the \( x_i \) are orthogonal coordinates. Since the surfaces have slight curvature, and the coordinates \( x_i \) are not truly orthogonal, the above calculation has to be changed slightly. Specifically, for sufficiently small \( \varepsilon \) the Lebesgue measure of \( B \) can be approximated by \( g \prod_i \varepsilon_i \), where \( g \) is the value of the metric \( g(x_1...x_d) \) at \( \alpha \). This quantity is strictly positive since each \( \varepsilon_i > 0 \) by Corollary 2, the metric \( g \) is strictly positive, and each \( \mu_{u_i} \) is absolutely continuous with respect to the Lebesgue measure. Hence \( B \) has positive Lebesgue measure.

\[
\square
\]

From these two lemmas, the proof of our main theorem follows, because the orthogonality of each \( u_i \) to \( \phi_i \) and the Born rule implies the set \( B \) cannot give any outcome with positive probability under measurement, which contradicts the fact that each element in \( B \) must give some outcome under measurement.

**Theorem 2.** There exists no symmetric, maximally nontrivial \( \psi \)-epistemic theory in dimension \( d \geq 3 \).

**Proof.** By Lemmas 8 and 9 there is a measurement basis \( M = \{\phi_1, \ldots, \phi_d\} \) and vectors \( u_1, \ldots, u_d \) such that each \( u_i \) is orthogonal to \( \phi_i \).

Furthermore, there is a set \( B \) of positive measure such that each \( \mu_{u_i} \) is equivalent to the Lebesgue measure on \( B \). Therefore by the Born rule, for each \( i \) we must have

\[
\int_B \mu_{u_i}(\lambda) \xi_{i,M}(\lambda) d\lambda = 0
\]

Since each \( \mu_{u_i} \) is equivalent to the Lebesgue measure on \( B \), this implies

\[
\int_B \xi_{i,M}(\lambda) d\lambda = 0
\]

But also, since for each state \( \lambda \), \( \sigma_i \xi_{i,M}(\lambda) = 1 \), we have that

\[
\sum_i \int_B \xi_{i,M}(\lambda) d\lambda = \int_B d\lambda > 0
\]

which is a contradiction. \( \square \)

## 4 Conclusions and Open Problems

In this paper, we gave a construction of a maximally nontrivial theory in arbitrary finite dimensions. However, the theory we constructed is not symmetric and is rather unnatural. We then proved that
symmetric, maximally nontrivial $\psi$-epistemic theories do not exist in dimensions $d \geq 3$ (in contrast to the $d = 2$ case, where the Kochen-Specker theory provides an example). Our impossibility proof made heavy use of the symmetry assumption. As for the assumption $d \geq 3$, we used that in two places: firstly and most importantly, to get a nonempty deficiency region (in Corollary 3), and secondly, to prove that $r_\psi = \frac{1}{2}$ in Lemma 5.

It might be possible to relax our symmetry assumption and to consider different ontic spaces, since deficiency holds for any $\psi$-epistemic theory in $d \geq 3$ even without a symmetry assumption. Also, in our proof, we did not use the specific form of the Born rule, only the fact that projection of $|\psi\rangle$ onto $|\phi\rangle$ must occur with probability 0 if $\langle \psi | \phi \rangle = 0$. Additional properties of the Born rule might place further constraints on $\psi$-epistemic theories.

Interestingly, trying to generalize the proof of Theorem 1 directly to obtain a proof of Theorem 2 gives rise to a variant of the Kakeya/Besicovitch problem. Recall that to prove Theorem 1, we showed that ontic states in a set $B$ in the neighborhood of $\alpha$ returned value $j$ under measurement, and yet the average measure of states orthogonal to $j$ had nontrivial support on $B$. Now if the measures $\mu_\psi = f_\psi(|\langle \psi | \lambda \rangle|^2)$ vary with $\psi$, it remains open whether or not the measures of states orthogonal to $j$ must have support on $B$, or if instead it is possible for them to “evade” $B$ to avoid contradicting the Born rule.

Placing this problem in the plane rather than in $\mathbb{CP}^{d-1}$, we obtain a clean Kakeya-like problem as follows. Let $S$ be a subset of $\mathbb{R}^2$ with the following property. For all $x \in \mathbb{R}^2$ and $\varepsilon > 0$, $S$ contains a set of circles, centered at $x$, that has positive Lebesgue measure within the annulus $\{y : |y - x| \in [1 - \varepsilon, 1]\}$. Can the complement of $S$ have positive Lebesgue measure? This question has been discussed on MathOverflow [7] but remains open.

Here are some additional open problems.

- An obvious problem is whether symmetric and nontrivial (but not necessarily maximally nontrivial) theories exist in dimensions $d \geq 3$.
- How does the size of the deficiency region scale as the dimension $d$ increases?
- In the maximally nontrivial theory we constructed, the overlap between any two nonorthogonal states $|\psi\rangle, |\phi\rangle$ is vanishingly small: like $(\varepsilon/d)^{O(d)}$ as a function of the dimension $d$ and inner product $\varepsilon = |\langle \psi | \phi \rangle|$. Is it possible to construct a theory with substantially higher overlaps – say, $(\varepsilon/d)^{O(1)}$? (Note that if $d \geq 3$, then the result of Leifer and Maroney [4] says that the overlap cannot achieve its “maximum” value of $\varepsilon^2$.)
- Can we construct $\psi$-epistemic theories with the property that an ontic state $\lambda$, in the support of an ontic distribution $\mu_\psi$, can never be used to recover the quantum state $\psi$ uniquely? (This question was previously asked by Leifer and Maroney [4], as well as by A. Montina on MathOverflow [6].)
- What can be said about the case of infinite-dimensional Hilbert spaces?

5 Acknowledgments

We thank Andrew Drucker for discussions about the measure-theoretic aspects of this paper, and Matt Leifer and Terry Rudolph for discussions about $\psi$-epistemic theories in general. We also thank Terence Tao, Sean Eberhard, Ramiro de la Vega, Scott Carnahan, and Alberto Montina for discussions on MathOverflow.
References

[1] S. Kochen and E. P. Specker, “The problem of hidden variables in quantum mechanics,” J. Math. Mech. 17, 59 (1967).

[2] M. F. Pusey, J. Barrett and T. Rudolph, “On the reality of the quantum state,” Nature Physics 8, 476 (2012).

[3] O. J. E. Maroney, “How statistical are quantum states?” (2012), arXiv:1207.6906

[4] M. S. Leifer and O. J. E. Maroney, “Maximally epistemic interpretations of the quantum state and contextuality” (2012), arXiv:1208.5132v1

[5] P. G. Lewis, D. Jennings, J. Barrett and T. Rudolph, “The quantum state can be interpreted statistically,” (2012), arXiv: quant-ph/1201.6554v1

[6] http://mathoverflow.net/questions/95537/psi-epistemic-theories-in-3-or-more-dimensions

[7] http://mathoverflow.net/questions/108507/a-kakeya-like-problem-must-a-union-of-annuli-fill-the-plane

[8] T. Rudolph, “Ontological models for quantum mechanics and the Kochen-Specker theorem,” (2006), arXiv: quant-ph/0608120v1

[9] N. Harrigan and T. Rudolph, “Ontological models and the interpretation of contextuality,” (2007), arXiv: quant-ph/0709.4266v1