

Summing to uncertainty: On the necessity of additivity in deriving the Born rule

Jiaxuan Zhang ^{*}*Department of Physics, University of Oxford, Oxford OX1 3PU, United Kingdom*

(Received 22 May 2026; revised 28 July 2026; accepted 4 August 2026; published 13 October 2026)

The emergence of intrinsic probability is one of the central problems in quantum mechanics, and investigating the origin of the Born rule provides the most direct approach to this problem. Many derivations of the Born rule have been proposed based on different postulates and frameworks. Identifying which postulates are genuinely necessary is therefore essential for understanding the origin of quantum probability. Here we prove that two postulates commonly employed in derivations of the Born rule, additivity and noncontextuality, are not equivalent and thus cannot be unified into a single postulate. We further show that the additivity postulate is indispensable in five major derivations of the Born rule: Gleason's theorem, Busch's extension of Gleason's theorem, the Deutsch-Wallace theorem, Zurek's envariance derivation, and the Finkelstein-Hartle theorem. We demonstrate that these derivations rely either on explicit additivity assumptions or on implicit forms of additivity embedded in the original arguments and that removing or weakening additivity gives rise to significant problems in the derivations. Our results support the view that additional postulates with intrinsic probabilistic properties should be introduced to derive the Born rule.

DOI: [10.1103/mqy7-s2db](https://doi.org/10.1103/mqy7-s2db)

I. INTRODUCTION

Ever since the early development of quantum mechanics, its intrinsic probabilistic nature has been one of the most puzzling problems for physicists. This problem is described by the famous *quantum measurement problem*. Briefly speaking, the measurement problem concerns the question of why a quantum system evolves continuously and deterministically according to the Schrödinger equation when it is not measured, while a measurement introduces a completely different, discontinuous, and random process [1,2]. The probabilistic result of a measurement is described by the *Born rule* [3,4]:

$$p(A) = \text{Tr}[\rho A],$$

where A is a projective operator or a *positive operator-valued measure* (POVM) effect. Also, if A is an observable, then the Born rule gives us the expectation value for the measurement [4].

In the standard formalism of quantum mechanics, the Born rule is introduced as the last axiom [4–7]. Unlike the two simple axioms of special relativity, which have clear physical intuitions, the quantum postulates are often criticized for being complicated and lacking transparent physical interpretations [8,9]. For example, why should quantum states be represented by vectors in Hilbert spaces? Why are quantum measurements represented by projections onto subspaces of

Hilbert spaces? Most importantly, why must we postulate the Born rule, a second dynamical law that appears to be unrelated to the Schrödinger equation? And where does its probabilistic nature originate? The goal of all existing interpretations of quantum mechanics is to address this measurement problem [1,2]. The Born rule can therefore be regarded as a core of the measurement problem. As a result, there have been many attempts to derive it rather than postulate it [7,10–15]. Some of these derivations also provide, to a certain extent, reasonable explanations for other quantum postulates, like the ones for quantum states [10] and measurement observables [7].

It has long been desired by many researchers that the Born rule, which is the only explicitly probabilistic postulate in quantum mechanics, could be derived from the other nonprobabilistic quantum postulates. This expectation is particularly strong among supporters of deterministic quantum interpretations, such as the *many-worlds interpretation* (MWI) [7,14–16]. Unfortunately, this attempt does not appear to be successful. All existing derivations of the Born rule involve the introduction of additional assumptions beyond standard quantum mechanics [17], even without considering the debates over the validity of each derivation. Vaidman has discussed several important derivations in Ref. [17] and concludes that extra postulates are unavoidable in deriving the Born rule. Although we agree with this argument, we believe that a more detailed analysis of existing derivations and their underlying postulates should be presented to support it.

In many original works on deriving the Born rule, the assumptions employed are not presented clearly or rigorously enough. As a result, it is difficult to determine which postulates are genuinely additional, which are indispensable for the derivation, and whether some of the extra assumptions can be derived from the nonprobabilistic quantum axioms or from other additional assumptions. There are a few authors

^{*}Contact author: jiaxuan.zhang@physics.ox.ac.uk

attempting to combine some additional assumptions—Logiurato and Smerzi have claimed that the *additivity* assumption can be replaced by the *noncontextuality* assumption [18]. Answering these questions is essential for locating the origin of probability in the Born rule and in quantum mechanics.

The focus of this work is the significance of the additivity assumption in derivations of the Born rule. First, we prove that additivity and noncontextuality are independent properties that cannot substitute for one another. In particular, we show that the proposed substitution in Ref. [18] is unsuccessful. Additivity is a fundamental property of *measures* and *probability measures* in mathematics [19] and therefore carries an intrinsic probabilistic property. In contrast, the noncontextuality of measurement functions does not contain any explicit probabilistic feature. Some authors have also argued that the probabilistic nature of the Born rule must ultimately originate from additional assumptions that themselves possess probabilistic property [17,20]. Our results provide clearer and more convincing evidence supporting this viewpoint.

Second, we prove that, among the five different derivations analyzed in this work (Gleason's theorem, Busch's extension of Gleason's theorem, the Deutsch-Wallace theorem, Zurek's envariance derivation, and the Finkelstein-Hartle theorem), none of the proposed "equivalent substitutes" genuinely plays the role of an additivity assumption. All five derivations ultimately require *countable additivity* itself. In some cases, additivity is explicitly assumed as a core postulate and plays a central role throughout the proof. In others, we identify the problems that arise in the absence of additivity and show how they can be resolved by introducing a countable additivity assumption. The five derivations we study represent three major streams of methods for deriving the Born rule, which we refer to as the Gleason-type, Deutsch-type, and Finkelstein-Hartle theorem-type approaches. The classification adopted here is based primarily on the mathematical structure of the derivations rather than on their physical motivations. A large number of later and more recent works also rely heavily on the mathematical structures of these three types (Gleason type [21–25], Deutsch type [18,26,27], and Finkelstein-Hartle theorem type [28]), although they are motivated by diverse physical ideas. This suggests that our analysis of the importance of additivity can also be extended to these newer derivations.

The structure of this paper is as follows. In Sec. II, we introduce the background material necessary for our discussion, including the nonprobabilistic postulates of quantum mechanics and three commonly used additional assumptions: additivity, noncontextuality, and normalization. In Sec. III, we analyze the relationships between the additivity assumption and the other two and prove that no equivalence relation exists among them. In Sec. IV, we examine the role of the additivity assumption in five prominent derivations of the Born rule. We also refer to them as two Gleason-type derivations [10,11] and three MWI derivations [12–14]. We call them MWI derivations because they were originally proposed by supporters of the many-worlds interpretation and exhibit certain MWI characteristics. However, these three derivations do not rely on MWI and still work within the framework of standard quantum mechanics. We will show that many problems encountered in these derivations arise from the absence of an

explicit additivity assumption. Although some of these issues have already been noted in the literature, our analysis provides a novel and illuminating perspective on their underlying cause.

II. PRELIMINARIES

In this section, we introduce the necessary background knowledge for the analyses in this paper. Some comes from quantum mechanics, and the rest comes from measure theory in mathematics. For our purpose, we will adjust the expressions of some definitions from measure theory. We also give rigorous definitions for three important properties or assumptions: additivity, noncontextuality, and normalization, distinguishing between the different meanings of these concepts that appear in the literature under the same name. This will help us clarify how the apparent equivalence of two concepts can emerge from carelessness in the definition.

A. Nonprobabilistic quantum postulates

The postulates of the standard quantum formalism first come from the works of Dirac [5] and von Neumann [6]. Combined with the definitions used in modern textbooks and literature [4,7], we list all the nonprobabilistic quantum postulates as follows:

(QM1) *State postulate*: A pure quantum state is represented by a unit vector $|\psi\rangle$ in a Hilbert space. A mixed state is represented by a convex combination of the density matrices of the pure states.

(QM2) *Evolution postulate*: Quantum states evolve under unitary transformations, or equivalently, according to the Schrödinger equation.

(QM3) *Observable postulate*: The measurement operator of any measurable physical property can be described by an observable

$$\Omega_A = \sum_i a_i A_i,$$

where $A_i = |a_i\rangle\langle a_i|$ is the projective operator of the eigenstate $|a_i\rangle$ and a_i is the corresponding eigenvalue.

We can also denote the measurement operators as a set $\{A_i, a_i\}_i$. This indicates that we can perform the measurement for a single projective operator without specifying the entire set. This raises the question of contextuality, which we will discuss later.

More general quantum measurements are not only described by projections but also by arbitrary positive semidefinite operators. Therefore, the Observable postulate has a variant:

(QM3') *POVM postulate*: A general quantum measurement is represented by a POVM:

Definition 1. A POVM $\{A_i\}$ is a set of operators, or POVM effects A_i , that satisfies the following properties:

- (i) For all i , $A_i \geq 0$.
- (ii) $\sum_i A_i = I$, where I is the identity operator.

Finally, we have the Eigenstate postulate, which describes a special case of quantum measurement that is deterministic and continuous:

(QM4) *Eigenstate postulate*: If a state is in an eigenstate $|a_i\rangle$ of the observable Ω_A , then the probability of obtaining a_i when measuring Ω_A is 1.

B. Measure and relevant additional assumptions

The observable postulate (QM3) describes the mathematical expression for the operator of a measurement operation. To obtain the measurement result, we need another function that takes measurement operators and states as input and outputs the results we can observe in experiments. In standard quantum mechanics, this function is given by the Born rule. However, since our purpose is to derive the Born rule, we only assume it to be a real-valued function here. This assumption is reasonable since we never observe imaginary values from experiments.

(M1) *Real function postulate*: Once we have the state ρ and the measurement (positive semidefinite) operator A , the quantum measurement process can be assumed to be a map or a function that maps the operator and the state to the extended real line:

$$\mu(\rho, A) = r \in \mathbb{R} \cup \{\pm\infty\}.$$

Usually, μ is assumed to be a measure in the literature on derivations of the Born rule [10,11], which directly introduces many important properties, such as additivity and non-negativity. However, since our goal is to study the necessity of each assumption, we only assume μ to be a real-valued function. We include $\{\pm\infty\}$ in the codomain, since infinite values are allowed by the definition of a measure [19]. This can cause problems related to continuity, which we will discuss in Sec. IV. Also, this output real number is not assumed to be a probability.

Note that the measurement function introduced in (M1) takes two arguments, ρ and A . However, throughout this paper, we usually write only $\mu(A)$. In most statements, this notation means that the corresponding statement is intended to hold for every admissible quantum state ρ , and ρ is omitted for simplicity. In Gleason-type derivations, by contrast, a quantum state is not postulated at the outset, so the measurement function is genuinely of the form $\mu(A)$, with no state argument. We use the more general notation $\mu(\rho, A)$ in (M1) to provide a unified framework for comparing the different derivations considered in this work.

We now discuss some special properties that this function may have. The definitions of additivity, noncontextuality, and normalization are given below, each with two different versions. This division may seem tedious at first, but it will be shown to be essential for clarifying the relationships between these properties or assumptions.

Definition 2 (Refs. [10,11]). The function μ is *countable additive* if for any countable set $\{A_i\}$ of mutually compatible operators A_i 's,

$$\mu\left(\sum_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

Here mutual *compatibility* means that the A_i are mutually disjoint if the A_i are projections and that $\sum_i A_i \leq I$ if the A_i are POVM effects.

This is the additivity assumed by Gleason and Busch in their derivations of the Born rule [10,11]. We will show below that this additivity usually (but not necessarily) involves an implicit noncontextuality assumption.

Definition 3 (Ref. [19]). The function μ is *finite additive* if for any finite set $\{A_i\}$, it satisfies

$$\mu\left(\sum_{i=1} A_i\right) = \sum_{i=1} \mu(A_i).$$

Definition 4. The function μ is *observable noncontextual* (ONC) if the measurement result $\mu(A)$ of a measurement operator A is independent of the operation of other compatible measurements.

This is the usual meaning of noncontextuality in the literature [18,21,29]. The compatible measurements are referred to as the *context* of the measurement. Here A and the compatible measurements can be projective operators, POVM effects, or observables.

Definition 5. The measurement function μ is *additivity noncontextual* (ANC) if it has an additivity property, and this additivity property is independent of the context of the measurement.

For example, let $\{A, B, C_1, C_2\}$ and $\{A, B, D_1, D_2\}$ be two different sets of measurements. The operators in both sets sum to I , and $\{C_1, C_2\} \neq \{D_1, D_2\}$. If ANC does not hold, we may have $\mu(A + B | \{C_1, C_2\}) = \mu(A | \{C_1, C_2\}) + \mu(B | \{C_1, C_2\}) \neq \mu(A + B | \{D_1, D_2\}) = \mu(A | \{D_1, D_2\}) + \mu(B | \{D_1, D_2\})$.

This property can be regarded as a subproperty of additivity.

Remark 1. We clarify here the conceptual distinction between ANC and ONC.

First, regarding the target object, ANC constrains some additivity property of μ and thus always appears together with additivity. That is, we ask whether $\mu(\sum_j A_{kj} | \{A_i\}_i) = \sum_j \mu(A_{kj} | \{A'_i\}_i)$ ($A_k = \sum_j A_{kj}$), for two different contexts $\{A_i\}_i$ and $\{A'_i\}_i$, both of which are compatible with the measurement operators; ONC, by contrast, concerns only a single fixed operator A_k in μ —it constrains how μ treats one fixed operator across different contexts, namely the measurement setups in which it is embedded as an outcome. That is, we ask whether $\mu(A_k | \{A_i\}_i) = \mu(A_k | \{A'_i\}_i)$. ONC is independent of additivity and therefore never compares an operator A_k with a different operator $\sum_j A_{kj}$ that happens to equal it after summation, because in that comparison the latter is no longer “ A_k embedded elsewhere.”

Second, as a result, regarding the flexibility of the measurement operator A_k in $\mu(A_k | \{A_i\}_i)$, ANC allows us to decompose a coarse operator A_k into finer compatible pieces A'_{kj} 's and constrains that summing their values is equivalent to treating A_k as directly given. This fine-graining is not supported by ONC, because its entire definition is stated in terms of a single fixed operator's context independence rather than operator identities that emerge through addition. Even if $A_k = \sum_j A_{kj}$ numerically, they are syntactically distinct objects. Because of this lack of flexibility with respect to addition, ONC is overall weaker than ANC and can be derived from the latter (see Proposition 2).

Definition 6. The function μ is *normalized* if $\mu(I) = 1$.

This is the usual definition of normalization for a probability measure [11]. In this work, whenever we refer to *normalization*, we mean this definition. However, some literature uses the term normalization to refer to another meaning below.

Definition 7. The function μ is *strongly normalized* if, for any countable set $\{A_i\}$ of mutually compatible operators (projections or POVMs) satisfying $\sum_i A_i = I$, we always have $\sum_i \mu(A_i) = 1$.

This property is not merely normalization, but a combination of normalization and additivity. It is misleading to call it normalization, but we define it in this way to help clarify the confusion present in some literature [18].

Additionally, the State postulate (QM1) also contains a normalization property and a non-negativity property. However, these properties do not imply normalization or non-negativity for the measurement function μ .

III. RELATION BETWEEN ADDITIVITY AND OTHER ASSUMPTIONS

In this section, we discuss the relation between additivity, noncontextuality, and normalization. First, we prove the nonequivalence of additivity and noncontextuality. We present counterexamples of functions μ that satisfy one property but not the other. We also rephrase the proof of their equivalence presented in Ref. [18], and identify the problem in it. At the end of this section, we prove that additivity can be derived from strong normalization but not normalization.

A. Additivity and noncontextuality

Noncontextuality and additivity are two additional assumptions commonly used in different existing derivations of the Born rule. Logiurato and Smerzi claim that these two assumptions are equivalent, namely, that each can be derived from the other [18]. We will show that they are not equivalent. The apparent equivalence arises from an ambiguity in the definitions of these assumptions.

Proposition 1. For a measurement function μ , ONC and (countable or finite) additivity are independent properties and cannot be derived from one another.

Proof. We prove the proposition by two counterexamples.

First, it is clear that additivity cannot be derived from ONC. A simple counterexample is

$$\mu(A) = \left(\frac{\text{Tr}[A]}{d} \right)^2,$$

where d is the dimension of the Hilbert space to which the operator A maps. It is clear that μ satisfies ONC. However, for rank-one projective operators A, B in a two-dimensional space,

$$\mu(A + B) = 1 \neq \mu(A) + \mu(B) = \frac{1}{2}.$$

Thus, μ does not satisfy additivity.

On the other hand, noncontextuality cannot be obtained from additivity alone. By additivity alone, we mean additivity without ANC. It is difficult to separate the two, because one may claim that, without ANC, additivity does not actually hold. However, additivity can still hold in an unusual form.

A simple counterexample is the *Equal rule*. We assign $\mu(A_i) = 1/N$ to every operator A_i in any set of operators, where N is the number of operators in that set. Let the two sets of projective operators be $\{A_1, A_2, A_3\}$ and $\{B_1, B_2, B_3, B_4\}$, where $A_1 + A_2 = B_1 + B_2$ but $A_i \neq B_j$ for any $i, j \in [2]$. Then

$$\mu(A_1 + A_2) = 2/3 = \mu(A_1) + \mu(A_2)$$

$$\mu(B_1 + B_2) = 1/2 = \mu(B_1) + \mu(B_2)$$

$$\mu(A_1 + A_2) \neq \mu(B_1 + B_2).$$

This μ satisfies additivity without ANC or ONC.

The nonequivalence of additivity and noncontextuality is therefore proved. ■

Now we explain the problem in Ref. [18] that leads to the incorrect conclusion that additivity and noncontextuality are equivalent. First, the problem in the derivation of additivity from noncontextuality is straightforward. Note that the first counterexample in the proof of Proposition 1 satisfies normalization but not strong normalization. The problematic proof in Ref. [18] arises from using strong normalization as the normalization assumption, which already implies additivity (see Lemma 1 below).

Second, the problem in the derivation of noncontextuality from additivity is more subtle. It arises from assuming ANC together with additivity, which already includes ONC (see Proposition 2 below). We first briefly explain the proof in Ref. [18]:

The logic of the proof in Ref. [18] can be illustrated by a simple example. Suppose we have two sets of projective operators $\{A_1, A_2, \dots\}$ and $\{B_1, B_2, \dots\}$. Let $A_1 = B_1$, and the other A_i and B_i are generally different. We assume additivity and normalization but not noncontextuality. This means that we may have

$$\mu(A_1) \neq \mu(B_1).$$

From normalization and additivity, we can obtain strong normalization (see Lemma 1). We then have

$$1 - \mu(A_1) = 1 - \mu(B_1)$$

$$\mu(A_2 + A_3 + \dots) \neq \mu(B_2 + B_3 + \dots).$$

The authors of Ref. [18] then claim that, since the operators $A_2 + A_3 + \dots = B_2 + B_3 + \dots$ and they also share the same context $A_1 = B_1$, even if ONC does not hold, one must still have $\mu(A_2 + A_3 + \dots) = \mu(B_2 + B_3 + \dots)$. This leads to a contradiction, and ONC is then derived from additivity.

We now show why this proof is incorrect.

Consider a different setup. Let $A_1 + A_2 = B_1 + B_2$, but $A_i \neq B_j$ for any $i, j \in [2]$. The remaining A_i and B_i are generally different. Again, we assume additivity and normalization but not noncontextuality. This means that we have

$$\mu(A_1 + A_2) \neq \mu(B_1 + B_2)$$

$$\mu(A_3 + \dots) \neq \mu(B_3 + \dots).$$

In this case, there is no contradiction, since the contexts of $(A_3 + \dots)$ and $(B_3 + \dots)$ are not the same.

Looking more closely at the setup discussed above, one may claim that additivity already implies

$$\mu(A_1 + A_2) = \mu(B_1) + \mu(B_2) = \mu(B_1 + B_2).$$

This is precisely where the confusion in the definitions arises. This equation holds only if ANC holds. ANC is a stronger noncontextuality assumption than ONC, and the latter can always be derived from the former.

Proposition 2. If a measurement function μ satisfies ANC, then it satisfies ONC.

Proof. We prove the result by contradiction. Suppose that we have two different contexts, or sets of operators, $\{A_i\}_i$ and $\{A'_j\}_j$, where $A_1 = A'_1$, but

$$\mu(A_1|\{A_i\}_i) \neq \mu(A'_1|\{A'_j\}_j).$$

Since μ satisfies ANC, the additivity property holds independently of the context. Therefore, we have

$$\mu(A_1 + A_2) = \mu(A_1) + \mu(A_2) = \mu(A'_1) + \mu(A_2).$$

This implies $\mu(A_1|\{A_i\}_i) = \mu(A'_1|\{A'_j\}_j)$, contradicting our assumption. Hence, ANC implies ONC. ■

Although ANC is often implicitly assumed in additivity assumptions, it is important to recognize that noncontextuality is then derived from a stronger version of noncontextuality rather than from additivity itself. Without ANC, ONC cannot be derived from additivity alone. The Equal rule counterexample in the proof of Proposition 1 clearly shows this.

Remark 2. The Equal rule counterexample in Proposition 1 also applies to the case considered in Ref. [18]. Suppose we have two sets of projective operators $\{A_1, A_2\}$ and $\{B_1, B_2, B_3\}$, where $A_1 = B_1$. Since both sets sum to I , we know that $A_2 = B_2 + B_3$. By the Equal rule, we obtain

$$\mu(A_2) = \frac{1}{2} \neq \mu(B_2 + B_3) = \frac{2}{3} = \mu(B_2) + \mu(B_3).$$

It may seem confusing that we obtain this inequality when the operator $A_2 = B_2 + B_3$ and the context $A_1 = B_1$ are both equivalent. The correct interpretation is that, for the measurement $\{B_1, B_2, B_3\}$, we can only calculate $\mu(B_2 + B_3)$ using additivity. Otherwise, we have no way to assign a value to $\mu(B_2 + B_3)$. If we measure $B_2 + B_3$ directly, then the measurement operator changes. Therefore, the measurement operators are not the same in the two measurements, although the context is.

Thus, we have shown that noncontextuality and additivity are independent properties. The ONC assumption cannot be derived from additivity without ANC, and the derivation in the opposite direction is even more clearly problematic. In the next subsection, we show why adopting strong normalization as the normalization assumption can falsely suggest that additivity can be derived from noncontextuality or indeed from any property one chooses.

B. Additivity and normalization

The aim of this subsection is to prove the independence of additivity and normalization: We have already presented a counterexample in the proof of Proposition 1 that satisfies normalization but not additivity or strong normalization. Counterexamples that satisfy additivity but not normalization are trivial to construct. The reason why these two independent assumptions are treated as equivalent by some [18] is that strong normalization is mistakenly used as the normalization assumption. However, the former is actually a combined property of the latter and additivity, as proved below.

Lemma 1. μ is strongly normalized if and only if μ is countably additive and normalized.

Notice that if the set of operators is finite rather than countable in the definition of strong normalization, we can only prove that μ is finitely additive.

Proof. If μ is strongly normalized, we know that $\mu(I) = 1$. And for mutually compatible A_i 's, we can express the strong normalization property using two partitions. First, considering the partition $\{\{\sum_i A_i\}, \{I - \sum_i A_i\}\}$, we obtain

$$\mu\left(\sum_i A_i\right) + \mu\left(I - \sum_i A_i\right) = 1.$$

Second, considering the finer partition $\{\sum_i A_i, \{I - \sum_i A_i\}\}$, we obtain

$$\sum_i \mu(A_i) + \mu\left(I - \sum_i A_i\right) = 1.$$

Subtracting the two equations gives

$$\mu\left(\sum_i A_i\right) = \sum_i \mu(A_i).$$

Conversely, if μ is countably additive, then for any countable set of mutually compatible operators $\{A_i\}$ such that $\sum_i A_i = I$,

$$\mu(I) = \mu\left(\sum_i A_i\right) = \sum_i \mu(A_i).$$

Since μ is normalized, we have $\sum_i \mu(A_i) = 1$. ■

Thus, if strong normalization is used as the normalization assumption, additivity is guaranteed. Since normalization is a very elementary and natural assumption, it is easy to mistakenly conclude that, with this seemingly harmless normalization assumption, additivity follows trivially from other assumptions such as noncontextuality. This conclusion is incorrect.

Together with the counterexamples discussed at the beginning of this subsection, we have shown that normalization and additivity are also independent properties and cannot be derived from each other.

IV. ADDITIVITY IN EXISTING DERIVATIONS OF THE BORN RULE

We have proved in the previous section that additivity cannot be derived from noncontextuality or normalization. In this section, we will show the crucial role played by the additivity assumption in five important existing derivations of the Born rule. Two of these are Gleason's theorem [10] and its variant [11]. The other three are derivations proposed by supporters of the MWI [12–14]. Although the latter three proofs exhibit clear features of MWI, their mathematical arguments do not rely on any special assumptions specific to MWI. Thus, all five derivations can be regarded as independent of quantum interpretations.

We will briefly introduce each derivation, including its proof structure and postulates. This allows us to identify the additivity assumptions that appear in these derivations and to clarify how they function in each case. A qualitative

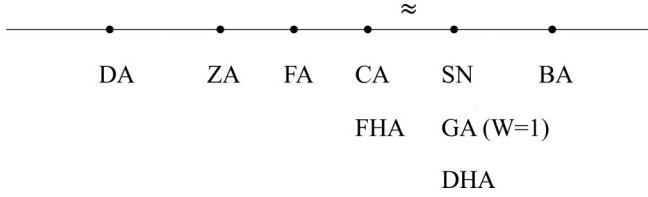


FIG. 1. Different additivity assumptions are shown along the axis, with their strength increasing from left to right. The abbreviations are as follows: Deutsch's additivity (DA), Zurek's weak additivity (ZA), finite additivity for projections (FA), countable additivity for projections (CA), the additivity assumption introduced for the Finkelstein-Hartle theorem (FHA), strong normalization (SN), Gleason's frame function (GA), Deutsch's hidden additivity (DHA), and Busch's additivity (BA). The $\approx$ indicates that, although the property on the right-hand side is stronger overall, this additional strength comes from extra properties, such as normalization. The strength of additivity itself is the same on both sides of the $\approx$ symbol.

comparison of the strengths of the different additivity assumptions used in these proofs is presented in Fig. 1. We will not discuss all debates surrounding each derivation in detail, since doing so would require substantial additional material and would distract from the focus of this paper. Instead, we focus on issues related to additivity.

A. Gleason's theorem

In this and the following subsections, we justify the following observation:

Observation 1. The countable additivity assumption is the central and indispensable assumption in the Gleason-type derivations of the Born rule: Gleason's derivation and Busch's derivation.

We first introduce Gleason's theorem. Proved in 1957, Gleason's theorem is one of the earliest results deriving the Born rule [10], and it is often regarded as one of the most important and reliable results in the field [18,20,29]. It replaces the commonly criticized strong additivity assumption in von Neumann's first derivation [6] with a more reasonable additivity assumption and derives the Born rule with a rigorous mathematical framework. The main concerns about Gleason's theorem are that it does not have a very clear physical intuition [7,18,30,31] and that it has an exception for two-dimensional Hilbert spaces [10]. Although Gleason's original proof is often regarded as intricate [32], its structure can be briefly reformulated as follows:

We begin by summarizing the postulates assumed by Gleason. The only nonprobabilistic quantum postulate used by Gleason is the Observable postulate (QM3). He also uses the Real Function postulate (M1). The function μ that he defines takes a single projection A_i as its input, so the measurement context can be left undetermined. The remaining postulates used by Gleason are listed below.

(G1) *Frame Function postulate (additivity and normalization).* Gleason assumes μ to be a frame function:

Definition 8. A frame function $\mu(|x\rangle)$ of weight W is a real-valued function defined on the set of unit vectors in a separable real or complex Hilbert space $\mathcal{H}$, such that if $\{|i\rangle\}_i$

forms an orthonormal basis of $\mathcal{H}$, then

$$\sum_i \mu(|i\rangle) = W.$$

Note that this definition is equivalent to our strong normalization when $W = 1$. Therefore, Gleason's frame function is a combined assumption of additivity and normalization. The normalization property is discussed separately below.

When $\dim \mathcal{H} < \aleph_0$, the additivity of the frame function is finite additivity; when $\dim \mathcal{H} = \aleph_0$, it is countable additivity.

(G2) *Non-negativity postulate:* For $\forall |x\rangle$, $\mu(|x\rangle) \geq 0$.

(G3) *Normalization postulate:* From Gleason's additivity, we can define $\mu(I) = W$.

At first glance, the normalization assumption may seem not very useful. One might think that, given Gleason's additivity with ANC, $\mu(I)$ is guaranteed to be a fixed value W and that W can always be rescaled. However, from the definition of a measure, it is possible that $W = +\infty$. This possibility may lead to a failure in the proof of the continuity of μ . Therefore, the normalization assumption and the non-negativity assumption are both introduced to exclude the infinite outputs.

Less importantly, if we want to interpret μ as a probability, then we should assume normalization and rescale W to 1 at the end to make the result physically meaningful and consistent with the Born rule.

(G4) *ONC postulate:* μ depends only on its input vector, namely the single projective operator measured, and not on the measurement context. That is, for any two different contexts $\{x_i\}_i$ and $\{x'_i\}_i$, as long as $x_k = x'_k$, we have

$$\mu(|x_k\rangle|\{x_i\}_i) = \mu(|x'_k\rangle|\{x'_i\}_i).$$

The noncontextuality assumption is not stated clearly in Gleason's original paper [10]. However, in modern discussions of Gleason's theorem, noncontextuality is often the first postulate to be mentioned and is regarded as very important [18,20,22,26,31]. This emphasis may be influenced by the early contextual counterexample given by Bell, which satisfies Gleason's additivity and other postulates except noncontextuality, as well as by the connection of this counterexample with hidden variable interpretations of quantum mechanics [29].

We also list two other definitions used in Gleason's proof to clarify some concepts in the discussion below. Note that these are not assumptions.

Definition 9 (Ref. [10]). Given a frame function μ on $\mathcal{H}$, we say it is regular if and only if there exists a Hermitian operator ρ on $\mathcal{H}$ such that

$$\mu(|x\rangle) = \langle x|\rho|x\rangle.$$

Definition 10 (Ref. [10]). A completely real subspace $\mathcal{K}$ of a Hilbert space $\mathcal{H}$ is a linear subspace of $\mathcal{H}$ such that for any $|x\rangle, |y\rangle \in \mathcal{K}$, $\langle x|y\rangle \in \mathbb{R}$.

It is easy to prove that a completely real subspace $\mathcal{K}$ is always a real Hilbert space. However, $\mathcal{H}$ itself does not need to be a real Hilbert space.

We now present the key ideas of Gleason's proof [10]. Note that the order here is not exactly the same as in Gleason's original paper but is rearranged for clarity.

First, it is easy to show that a frame function μ on $\mathcal{H}$ remains a frame function when its domain is restricted to

any subspace of $\mathcal{H}$. Moreover, if $\mu \geq 0$, then it remains non-negative on any subspace.

Second, it is straightforward that any real or complex Hilbert space $\mathcal{H}$ with $\dim \mathcal{H} \geq 3$ always contains three-dimensional completely real subspaces $\mathcal{H}' \cong \mathbb{R}^3$.

It is also straightforward that every completely real two-dimensional subspace of $\mathcal{H}$ is a linear subspace of some completely real three-dimensional subspace $\mathcal{H}' \cong \mathbb{R}^3$. Furthermore, if a non-negative frame function μ is regular on $\mathbb{R}^3$, then it is also regular on any completely real two-dimensional subspace of $\mathbb{R}^3$.

Third, Gleason proves that if a non-negative frame function μ is regular on every completely real two-dimensional subspace of $\mathcal{H}$, then μ is regular on the whole space $\mathcal{H}$ (Lemma 3.3 and Lemma 3.4 in Ref. [10]).

Therefore, the crucial step is to prove that any non-negative frame function μ is regular on $\mathbb{R}^3$. Gleason achieves this by first proving that any continuous frame function on $\mathbb{R}^3$ is regular (Theorem 2.3 in Ref. [10]) and then proving that any non-negative frame function μ on $\mathbb{R}^3$ is continuous (Theorem 2.8 in Ref. [10]).

The additivity of the frame function is essential in both steps. Clearly, a continuous function without additivity on $\mathbb{R}^3$ is not necessarily regular, and a non-negative function without additivity on $\mathbb{R}^3$ is not necessarily continuous.

Moreover, we cannot prove that an additive function on $\mathbb{R}^3$ is continuous without the non-negativity assumption. The additivity of the frame function only guarantees linearity for rational coefficients, that is, $\mu(qA) = q\mu(A)$ for all $q \in \mathbb{Q}$. However, over the real coefficients there exist many pathological additive functions that are discontinuous and unbounded. Gleason therefore introduces the non-negativity assumption to exclude these discontinuous frame functions [10] (see more in Appendix A 5).

Thus, both additivity and non-negativity are indispensable in Gleason's proof. Additivity plays the more important role as the central postulate. Without it, the entire proof would fail.

B. Busch's extension of Gleason's theorem to POVMs

Busch extends Gleason's result in a 2003 paper to more general quantum measurements, represented by POVMs [11]. The advantage of Busch's proof is that it is also valid for the two-dimensional Hilbert space, where Gleason's theorem does not single out the Born rule as the only solution. Busch's proof is also simpler, clearer, and more intuitive. The price, however, is an additivity assumption much stronger than the one used by Gleason. Note that Busch's additivity is stronger because it concerns POVM effects. The probabilistic result generated from assumptions about POVM effects can be suspicious, since they do not correspond to physical values of isolated systems as projective operators [33]. Nevertheless, the additivity applies only to compatible POVM effects, so the criticism [29,34] of von Neumann's strong additivity assumption (which applies to any measurements, including incompatible ones) does not apply to Busch's additivity assumption.

Both Gleason and Busch make explicit additivity assumptions in their proofs, so we focus on analyzing the role of additivity in their proofs in the previous and current

subsections. In the other three derivations, the additivity assumption is not always presented clearly in the original works. We will analyze problems that this omission causes.

We first briefly introduce Busch's postulates and the structure of his proof.

Busch assumes the POVM postulate (QM3') and the Real Function postulate (M1). For (M1), Busch actually assumes μ to be a probability measure, but we separate the properties of normalization and non-negativity. The remaining postulates by Busch are listed below:

(G-B1) *Countable additivity postulate for POVM*: μ satisfy the additivity property

$$\mu(A + B + \dots) = \mu(A) + \mu(B) + \dots$$

if $|\{A, B, \dots\}| \leq \aleph_0$ and $A + B + \dots \leq I$, i.e., the POVM effects are compatible.

(G-B2) *Non-negativity postulate*: $\mu \geq 0$.

In Gleason's proof, non-negativity is introduced to ensure continuity. It plays the same role in Busch's proof. Busch does not prove this continuity in his original paper. We provide a proof in Theorem A5 in Appendix A 5.

(G-B3) *Normalization postulate*: $\mu(I) = 1$.

As in Gleason's proof, this assumption is useful for ruling out infinite output values and for aligning the result with the Born rule.

(G-B4) *ONC postulate*: For any POVM effect A , $\mu(A)$ depends only on A and not on other compatible measurements performed simultaneously.

We present another definition of linear functional here to clarify Busch's proof. This is not an assumption.

Definition 11 (Ref. [35]). For a vector space V over the field K , $f : V \rightarrow K$ is called a *linear functional* if for any $\vec{u}, \vec{v} \in V$ and any scalar $c \in K$,

- (i) $f(\vec{u} + \vec{v}) = f(\vec{u}) + f(\vec{v})$
- (ii) $f(c\vec{u}) = cf(\vec{u})$

Busch's proof is structured as follows [11]. It is a proven result that any linear functional $\mu(A)$ can be written as $\mu(A) = \text{Tr}[\rho A]$ [6]. Since Busch assumes countable additivity for POVM effects at the outset, what remains is to derive homogeneity, i.e., $\mu(rA) = r\mu(A)$ for $r \in \mathbb{R}$. The proof for $r \in \mathbb{Q}$ is straightforward. Then continuity is needed to extend homogeneity to real r 's.

Although Busch's additivity assumption is much stronger than Gleason's, establishing the required continuity (which differs from the continuity discussed in Gleason's proof; see Remark A3 in Appendix A 5 for details) still relies on the normalization and non-negativity assumptions. Nevertheless, additivity for POVM effects substantially simplifies the proof and also plays a central role at each step of Busch's derivation. This completes our proof of Observation 1.

C. Deutsch's derivation with decision theory

In this and the following subsections, we justify the following observation:

Observation 2. The countable additivity assumption must be introduced, either explicitly or implicitly, in Deutsch-type derivations of the Born rule: Deutsch's and Zurek's derivations. Moreover, it cannot be derived from the weaker additivity assumption suggested by Zurek. Replacing

countable additivity with Zurek's weak additivity leads to the problem that the Born rule does not apply to all quantum states.

We first introduce Deutsch's derivation of the Born rule. Deutsch presented his proof in 1999 [12]. This is a significant result in the field. It aims to derive the probabilistic Born rule from nonprobabilistic postulates of quantum mechanics and decision theory, which Deutsch describes as "Deriving a 'tends to' from a 'does.'" [12] His method was later refined by Wallace and is therefore also called the *Deutsch-Wallace theorem* [36]. Discussions of whether this method can successfully achieve its goal have continued [13,37–40].

For our purposes, we mainly discuss two problems in Deutsch's derivation: the continuity problem pointed out by Barnum [41] and the dimensionality problem pointed out by Caves [42]. Note that both criticisms were originally raised in the context of Zurek's derivation (see Sec. IV D) but also apply to Deutsch's derivation. Both problems are closely related to the additivity assumption. In short, although Deutsch's additivity assumption appears at first sight to be very different from countable additivity, countable additivity can in fact be derived from Deutsch's postulates. Together with another hidden non-negativity assumption in Deutsch's derivation, this guarantees the continuity of the function μ , so continuity is not a genuine problem. However, since Deutsch's hidden additivity concerns only the complete set of projective operators, it does not imply the ANC property. Moreover, he uses a hidden ONC assumption not from the beginning of the proof but only at a later step for states with unequal amplitudes. This leaves room for contextual non-Born-rule counterexamples in low-dimensional Hilbert spaces. To analyze these problems more clearly, we first briefly review Deutsch's postulates and proof.

Deutsch starts his proof from a value function $\mathcal{V}[|\psi\rangle, X]$, where X is an observable, and proves that this value function yields the expectation value calculated by the Born rule, if we regard quantum measurement as a game played by a rational player following certain decision theory rules. In this game, the reward of each game is equal to the measurement outcome, i.e., a certain eigenvalue. In this work, for consistency of comparison, we can set the eigenvalue of each eigenstate $|x_i\rangle$ to be 1. Then the expectation value for any outcome is simply the probability, and we can write $\mu[|\psi\rangle, X] = \mathcal{V}[|\psi\rangle, X]$.

Although the rationality used by Deutsch has been repeatedly questioned in the literature [38,40], the rules that the player must follow in this quantum measurement game are given clearly. We summarize all of Deutsch's assumptions as follows.

The nonprobabilistic quantum postulates used by Deutsch are the State postulate (QM1), the Observable postulate (QM3), and the Eigenstate postulate (QM4). The Real Function postulate (M1) is also assumed in the value function setup. The remainder of his explicit and implicit postulates are listed below:

(D1) *Deutsch's additivity postulate*: The player is indifferent between receiving a single reward $(a + b)$ in one game and receiving separate rewards a and b in two games.

This additivity assumption seems trivial and reasonable. However, Deutsch actually uses it to justify another important

additivity, or strong normalization,

$$\sum_i \mu(|\psi\rangle, |x_i\rangle\langle x_i|) = 1. \quad (1)$$

The original value function $\mathcal{V}$ proposed by Deutsch does not take a single projective operator $|x_k\rangle\langle x_k|$ as input. However, we can easily define it by considering a game in which we only receive the reward x_k when the corresponding eigenstate is measured, and all other x_i 's are discarded.

It is from this additivity assumption that the whole proof starts. It can be obtained from Deutsch's additivity postulate if we consider the situation of a game in real life—it is trivial that, in a game, the sum of probabilities of all possible events should be 1. However, here Deutsch already introduces a probabilistic assumption. We will discuss in detail below how Eq. (1) is implicitly introduced and used.

(D2) *Non-negativity postulate*: $\mu(|\psi\rangle, |x_i\rangle\langle x_i|) \geq 0$.

This assumption is not stated explicitly by Deutsch. However, when he discusses the result for real probabilities, he effectively invokes it by claiming that if the eigenvalue for a certain eigenstate is increased while the remaining eigenvalues are unchanged, the player will prefer the new game to the old one. Although Deutsch's proof for the result for any real probability is problematic, this assumption does guarantee the continuity of the function μ and thus the Born rule for real probabilities. This is because it implies the non-negativity of the probability: If the probability μ of obtaining a particular eigenstate were negative, then the player should not want the corresponding eigenvalue to be larger.

(D3) *Transitivity postulate*: If the player prefers reward $a > b$ and $b > c$, then he or she also prefers $a > c$.

(D4) *Substitutability postulate*: In a composite game, the player is indifferent if any subgame is replaced by another game with the same value.

This assumption is not very useful, although it appears essential at the step where the result for unequal probabilities is proven. However, the real key point is to determine when two games have the same value. This is precisely where Deutsch's hidden noncontextuality assumption enters.

(D5) *Zero-Sum Rule postulate*: If the player always receives opposite rewards (x and $-x$) for two games, then the values of the two games sum to 0.

(D6) *State noncontextuality and ONC postulate*: $\mu(|\psi\rangle, |x_i\rangle\langle x_i|)$ only depends on the state $|\psi\rangle$ and the projective operator $|x_i\rangle\langle x_i|$.

This hidden ONC assumption is pointed out by Logiurato and Smerzi [18]. It is only essential at the step of proving the result for unequal probabilities, where compatible measurements on an ancillary system are used. However, that step also requires infinite dimensionality. Therefore, for finite-dimensional Hilbert spaces, there is no ONC assumption in Deutsch's proof, and contextual counterexamples exist [see more below, especially Eq. (4)].

The structure of Deutsch's proof is as follows [12]. He first proves that for the simplest two-dimensional symmetric pure state $|\psi\rangle = \frac{1}{\sqrt{2}}(|x_1\rangle + |x_2\rangle)$, $\mu(|\psi\rangle, |x_i\rangle\langle x_i|)$ must be equal to the modulus square of the amplitude of the eigenstate $|x_i\rangle$. He then extends this result to the general equal-amplitude pure state $|\psi\rangle = \frac{1}{\sqrt{2}}(|x_1\rangle + |x_2\rangle + \dots + |x_n\rangle)$ for any positive

integer n . Next, for pure states with unequal amplitudes, Deutsch starts from states with rational amplitudes,

$$|\psi\rangle = \sqrt{\frac{m}{n}}|x_1\rangle + \sqrt{\frac{n-m}{n}}|x_2\rangle.$$

Using his additivity (D1), Deutsch proves that the value for this game is equal to another game $\mathcal{V}[|\psi_{XY}\rangle, X \otimes Y]$, where

$$|\psi_{XY}\rangle = \sum_{j=1}^m \sqrt{\frac{1}{n}}|x_1\rangle|y_j\rangle + \sum_{j=m+1}^n \sqrt{\frac{1}{n}}|x_2\rangle|y_j\rangle \quad (2)$$

and

$$\sum_{j=1}^m |y_j\rangle = \sum_{j=m+1}^n |y_j\rangle = 0.$$

The result then follows from the equal-amplitude case. It is then straightforward to extend the result to any $|\psi\rangle = \sum \sqrt{q_i}|x_i\rangle$ with $q_i \in \mathbb{Q}$.

The extension to real probabilities proposed by Deutsch is problematic. He claims that we can always find unitary transformations U that transform each $|x_i\rangle$ into an eigenstate with a higher eigenvalue and that the player will prefer this new game to the old one. Deutsch also notes that transforming the eigenvalue is not relevant to the amplitude. He therefore introduces a more complicated operation that transforms each $|x_i\rangle$ into a superposition of itself and other eigenstates, all with higher eigenvalues. He claims that it is possible to find a U , such that the state after the transformation, $U|\psi\rangle$, has amplitude q_i with $|q_i|^2 \in \mathbb{Q}$ for each eigenstate of X . Then the value of this new game becomes an upper bound for the initial game. Since rational numbers are dense in the reals, he argues that we can prove the result for any amplitude $\sqrt{r_i}$ ($r_i \in \mathbb{R}$) by squeezing it between upper and lower bounds obtained from limits of the $\sqrt{q_i}$'s.

However, even if this complex operation is possible, obtaining upper or lower bounds for each amplitude λ_i in the original state $|\psi\rangle$ means making each amplitude λ_i smaller or larger, which contradicts normalization condition for a quantum state.

Nevertheless, this step of the proof can still be made to work if we can prove that μ is continuous with respect to the amplitudes, as we will show below.

Deutsch then extends the result to the case where λ_i is any complex number by arguing that applying unitary transformations that change only certain partial phase before and after the measurement will give the player equal rewards. This ends his proof [12].

We now analyze the problem of continuity raised by Barnum [41]. We have shown that Deutsch's proof of continuity does not work. However, since countable additivity and non-negativity are both implicitly present in Deutsch's derivation, the continuity of μ is in fact guaranteed by Gleason's theorem. We first show where Deutsch introduces the hidden assumption of countable additivity.

Deutsch claims that the following equation follows from Deutsch's additivity:

$$\mathcal{V}\left[\sum_i \lambda_i |x_i + k\rangle\right] = k + \mathcal{V}\left[\sum_i \lambda_i |x_i\rangle\right]. \quad (3)$$

The idea is simply that if we add k to each reward, then the final expected reward also increases by k . From our definition of $\mu(|\psi\rangle, |x_i\rangle\langle x_i|)$, we have

$$\mathcal{V}\left[\sum_i \lambda_i |x_i\rangle\right] = \sum_i \mu(|\psi\rangle, |x_i\rangle\langle x_i|)x_i,$$

and Eq. (3) is only valid when

$$\sum_i \mu(|\psi\rangle, |x_i\rangle\langle x_i|) = 1.$$

This is the truly important additivity assumption in Deutsch's proof, and it is exactly Gleason's additivity, or strong normalization, except without ANC.

Although we could conclude that function μ must be continuous from Gleason's theorem, we cannot say that this continuity can be proved in the same way as in Gleason's proof. Gleason's notion of continuity concerns eigenvectors rotating by a small angle [10] rather than small changes in amplitudes. These differences between the various forms of continuity are easy to confuse and can lead to misunderstandings (see more in Appendix A5, especially Remark A3).

At last, we expand the discussion by Caves concerning dimensionality [42]. This issue is similar to the exception for $\dim(\mathcal{H}) = 2$ in Gleason's theorem, but Deutsch's derivation leaves space for more exceptions.

Taking a closer look at the proof, we notice that only the first step is valid for a genuine two-dimensional Hilbert space—this means that the two-dimensional space is not a subspace contained in a larger Hilbert space. The second step, for $|\psi\rangle = \frac{1}{\sqrt{2}}(|x_1\rangle + |x_2\rangle + \dots + |x_n\rangle)$, uses induction, and for odd n we need to derive the result from an $(n+1)$ -dimensional space. This means, for example, that we need a four-dimensional space to prove the Born rule for $|\psi\rangle = \frac{1}{\sqrt{2}}(|x_1\rangle + |x_2\rangle + |x_3\rangle)$. To derive the Born rule for any rational m/n , we need an auxiliary system of arbitrary integer dimension. This can only be guaranteed when the auxiliary system has countably infinite dimension, which indicates $\dim(\mathcal{H}) = \aleph_0$. A $\aleph_0$ -dimensional space is sufficient for the rest of the proof. Thus, Deutsch's proof actually shows that, for a state in any subspace of a $\aleph_0$ -dimensional Hilbert space, the Born rule is the only solution for the value function $\mathcal{V}$.

We present a counterexample for $d = 2$ here:

$$\mu(|\psi\rangle, |x_i\rangle\langle x_i|) = \frac{|\langle\psi|x_i\rangle|^4}{\sum_{j=1}^2 |\langle\psi|x_j\rangle|^4}. \quad (4)$$

This counterexample satisfies all of Deutsch's postulates, but it apparently generates results different from the Born rule. Notice that when we have a two-dimensional symmetric pure state, this μ function gives the same result as the Born rule—namely $\mu(\frac{1}{\sqrt{2}}(|x_1\rangle + |x_2\rangle), |x_1\rangle\langle x_1|) = \mu(\frac{1}{\sqrt{2}}(|x_1\rangle + |x_2\rangle), |x_2\rangle\langle x_2|) = 1/2$. This is the only result that can be derived for a two-dimensional space from Deutsch's proof. For $d > 2$, counterexamples of the same form are not valid counterexamples to Deutsch's proof, because higher dimensionality imposes more constraints. However, as long as $d < \aleph_0$, there could be states $|\psi\rangle$ for which $\mu(|\psi\rangle, X)$ do not obey the Born rule, and other counterexamples can still be constructed.

Note that this counterexample still satisfies both Deutsch's additivity postulate and strong normalization without ANC, but it does not satisfy noncontextuality. This is reasonable, because Deutsch introduces an ONC postulate only implicitly when proving the result for unequal amplitudes [18]. If ONC does not hold, we can have

$$\mathcal{V}\left(\sqrt{\frac{m}{n}}|x_1\rangle + \sqrt{\frac{n-m}{n}}|x_2\rangle, X\right) \neq \mathcal{V}(|\psi_{XY}\rangle, X \otimes Y),$$

where $|\psi\rangle_{XY}$ is as in Eq. (2). In that case, the proof does not work. The ONC assumption is introduced only at this step, but the infinite dimensionality of the auxiliary system Y is also introduced at the same step. Therefore, counterexamples exist for any finite-dimensional Hilbert space.

D. Zurek's derivation with enviance

People are unsatisfied with Gleason's theorem because of its lack of physical intuition [7,18,30,31], and Deutsch's appeal to decision theory and rationality also seems suspicious to many [13,37–40]. By contrast, Zurek aims to provide a more physically motivated derivation of the Born rule [7,13,43]. The structure of Zurek's proof is actually very similar to Deutsch's, except for the first few steps. Zurek proves the Born rule for the equal-amplitude qubit state using a property he calls enviance, which was first abbreviated as “environment-assisted invariance” in his 2003 paper [43] but was later updated to the more appropriate term “entanglement-assisted invariance” [7]. This grounds Zurek's proof in the physically meaningful property of quantum entanglement.

Zurek gives special attention to the additivity assumption in his works. He claims that Gleason's additivity can be derived from a “weak” additivity assumption [13]. However, the consequence is that Zurek's additivity is in fact weaker than Gleason's additivity or Deutsch's implicit countable additivity and thus cannot guarantee the continuity of the measurement function μ . Due to the similarity between Zurek's and Deutsch's derivations, it also suffers from the same problem of dimensionality. We will first briefly introduce Zurek's proof and postulates and then explain these two problems.

There is a very interesting innovation in Zurek's assumptions. He does not assume the orthogonality of eigenstates or the expression of observables from the Observable postulate (QM3) but instead derives them from the State postulate (QM1), the Evolution postulate (QM2), and the Eigenstate postulate (QM4) [7]. In order not to distract from our main focus, we rephrase this proof in Appendix A3. Nevertheless, this is a very inspiring result worth noting. Zurek also uses another not-so-common nonprobabilistic quantum postulate: The representation of a composite state is a vector in the tensor product space of the components' Hilbert spaces [7]. The term “probability” is used directly by Zurek to describe the measurement result, so we can conclude that the Real Function postulate (M1) is also assumed. Since the observable X or the Schmidt decomposition of the state is given at the beginning, we can also represent the measurement function used by Zurek as $\mu(|\psi\rangle, X)$. However, we do not need to assume all properties of a probability measure here, because

the additivity and normalization are assumed independently by Zurek, and the non-negativity is not useful in the proof.

The additional postulates beyond nonprobabilistic quantum mechanics used by Zurek are as follows:

(E1) *Zurek's weak additivity and normalization postulate*: For two mutually exclusive events κ and $\kappa^\perp$,

$$\mu(\kappa \cup \kappa^\perp) = 1 = \mu(\kappa) + \mu(\kappa^\perp).$$

This means that the event κ either occurs or does not occur. We also use measurement outcomes x_i to represent events for simplification.

Zurek proves that the usual countable additivity can be derived from this weak additivity [13]. However, it actually ensures additivity only for states with rational probabilities for all outcomes (see detailed proof in Appendix A4). That is, if outcome events $\{x_i\}_i$ are mutually disjoint, and each $\mu(|x_i\rangle\langle x_i|) \in \mathbb{Q}$, then

$$\mu\left(\sum_i |x_i\rangle\langle x_i|\right) = \sum_i \mu(|x_i\rangle\langle x_i|).$$

Therefore, Zurek's additivity is not equivalent to Gleason's additivity or Deutsch's hidden countable additivity but is weaker than both. The continuity of μ cannot be guaranteed even with non-negativity.

(E2) *Zurek's locality postulate*: The state of system S cannot be affected by any operation on the environment E alone and vice versa.

(E3) *State noncontextuality postulate*: For any observable, the measurement results fully depend on the quantum state $|\psi\rangle$. Moreover, if the measured system is a subsystem of a larger composite system, then the composite state determines the state of the subsystem.

This corresponds to “fact 2” and “fact 3” in Zurek's paper [7]. Note that this assumption appears to be a form of noncontextuality, but it is not the same as the usual ONC and cannot play the role of ONC. We refer to this assumption as *state noncontextuality*. Deutsch also adopts this state noncontextuality postulate [12]. The main use of this postulate is to rule out any possible influence from hidden variables. Gleason's and Busch's approaches do not include this type of noncontextuality, because they do not assume the State postulate (QM1) at the beginning [10,11].

(E4) *ONC postulate*: For any $|\psi\rangle$, $\mu(|\psi\rangle, |x_i\rangle\langle x_i|)$ depends only on the projective operator $|x_i\rangle\langle x_i|$, regardless of the whole observable.

The proof for states with unequal amplitudes given by Zurek is exactly the same as that of Deutsch. Thus, both of them implicitly introduce the ONC postulate at that step.

The most important concept in Zurek's derivation, enviance, is not a postulate but a symmetry property of an entangled state under certain unitary transformations.

Definition 12 (Ref. [13]). We call a state $|\Psi_{SE}\rangle$ envariant under a unitary transformation U_S if

$$(I \otimes U_E)(U_S \otimes I)|\Psi_{SE}\rangle = |\Psi_{SE}\rangle.$$

This means that a unitary transformation U_S acting only on the measured system can be undone by another unitary transformation acting only on the environment.

Zurek's proof then follows this structure [13]: First, he proves the Born rule for equal-amplitude states using the envariance property. Consider an equal-amplitude state

$$|\Psi_{SE}\rangle = \frac{1}{\sqrt{2}}(|x_1\rangle|E_1\rangle + |x_2\rangle|E_2\rangle).$$

We can swap the eigenstates of the measured system using the unitary operator

$$U_S \otimes I = (|x_1\rangle\langle x_2| + |x_2\rangle\langle x_1|) \otimes I.$$

Let the initial probabilities for outcomes x_1 and x_2 be p_1 and p_2 . Before the swap, when the environment is in eigenstate $|E_1\rangle$, the corresponding eigenstate for the system is $|x_1\rangle$; after the swap, it becomes $|x_2\rangle$. Thus, for the state after the first swap, we have $\mu(x_1) = \mu(E_2) = p_2$ and $\mu(x_2) = \mu(E_1) = p_1$. We then perform another swap acting only on the environment,

$$I \otimes U_E = I \otimes (|E_1\rangle\langle E_2| + |E_2\rangle\langle E_1|).$$

The final composite state is then the same as the initial one and $\mu(x_1) = p_1$ and $\mu(x_2) = p_2$ for the final state after the two swaps. However, from Zurek's locality postulate, an operation on the environment E alone should not change the probability distribution of the system S . Therefore, we must conclude that $\mu(x_1) = p_2 = p_1$. From Zurek's weak additivity and normalization postulate, it follows that $p_1 = p_2 = 1/2$.

This result can be extended to any equal-amplitude state with n outcomes, yielding $p(x_i) = 1/n$, by performing the swap multiple times.

The remainder of Zurek's proof parallels that of Deutsch [7,13]. Similarly, extending the result to real probabilities requires continuity of μ . However, unlike in Deutsch's proof, this continuity cannot be guaranteed by Zurek's postulates. We analyze this problem below.

Zurek discusses the continuity problem explicitly in his paper [13]. His proof is as follows: If event $\kappa_1 \subseteq \kappa_2$, then the probability must satisfy $\mu(\kappa_1) \leq \mu(\kappa_2)$. Then the real probability of event κ_1 can be approached by events κ'_2 with rational probabilities from above. Continuity then follows by taking limits.

This statement is very confusing. In quantum measurements, events correspond to measurement outcomes or eigenvalues of an observable. It is unclear what an inclusion relation between two different eigenvalues or eigenspaces would mean. We could possibly let κ_2 correspond to more than one eigenstate, including the one corresponding to κ_1 . However, in that case the notion of "approaching" is not very meaningful. For example, let the quantum state be

$$|\psi\rangle = \sqrt{p_1}|x_1\rangle + \sqrt{p_2}|x_2\rangle + \dots,$$

where $p_1 \notin \mathbb{Q}$ and $p_1 + p_2 \in \mathbb{Q}$. Let $\mu(\kappa_1) = \mu(x_1)$ and $\mu(\kappa_2) = \mu(x_1 \vee x_2)$. Approaching $\mu(\kappa_1) = p_1$ by decreasing p_2 cannot prove the continuity of μ with respect to coefficient of $|x_1\rangle$. Zurek's mistake is similar to that of Deutsch. Deutsch attempts to take limits by increasing eigenvalues, while Zurek attempts to take limits by expanding the eigenspace. Neither approach satisfies our objective of proving continuity with respect to the amplitude or the probability.

Unlike Deutsch's proof, the continuity of μ becomes a genuine problem here. This issue has been pointed out in the

literature [41], but we present a clearer explanation of why continuity cannot be derived from Zurek's assumptions and proof.

Although Zurek uses the term probability several times in his proof, and probabilities are generally expected to be non-negative, non-negativity is not required at any step of his proof. Moreover, even if non-negativity is assumed, Zurek's additivity is still too weak to guarantee continuity. For example, we can let $\mu(x_i)$ follow the Born rule when the square of the amplitude $|\lambda_i|^2 \in \mathbb{Q}$, and let $\mu(x_i) = \sqrt{2}$ (or any irrational number) when $|\lambda_i|^2 \notin \mathbb{Q}$. This counterexample satisfies Zurek's additivity and non-negativity, but not the usual countable additivity (see more in Appendix A4), and is clearly discontinuous. This completes our proof of Observation 2.

Zurek's proof also lacks the ONC postulate until the step dealing with unequal amplitudes, since all eigenstates of the observable are fixed at the beginning when the entangled state is specified. Therefore, it suffers from the same dimensionality problem as Deutsch's proof, namely that the result is only valid for $\aleph_0$ -dimensional Hilbert spaces.

E. Hartle's frequentist derivation

In this subsection, we justify the following observation:

Observation 3. The lack of an additivity assumption in the Finkelstein-Hartle theorem causes an inconsistency in the core frequentist interpretation on which the proof is based: pure states and mixed states are not treated consistently within the method. As a result, there are two reasonable results when a mixed state is measured: One is the Born rule, and the other is an excrescent result.

We first introduce Finkelstein-Hartle theorem: Another early and important derivation of the Born rule was given by Hartle in 1967 using a frequentist method based on the infinite ensemble [14]. The same method was independently proposed earlier by Finkelstein in 1963 [15], and it is therefore also known as the *Finkelstein-Hartle theorem* [8]. This derivation is rather controversial due to its use of infinity [8,17,18,28,44]. More seriously, it contains some mathematical self-incoherence: The infinite ensemble state $|\psi\rangle^{\otimes \aleph_0}$ appears to be both an eigenstate of the frequency operator f^k and orthogonal to all initial eigenstates of f^k [8,44]. We will point out a new incoherence, which can be resolved by adding an additivity assumption that is lacking in the original derivation. To begin with, we review the postulates and the proof given by Hartle [14].

Hartle assumes only a small number of postulates. He adopts the State postulate (QM1) and the Observable postulate (QM3) from nonprobabilistic quantum mechanics. The Eigenstate postulate (QM4), however, is not assumed, because $|\psi\rangle^{\otimes \aleph_0}$ is generally not one of the initial eigenstates of f^k . Hartle describes the measurement result as a frequency, so we can think of the Real Function postulate (M1) as being assumed as well. We can denote the frequency function as $\mu(|\psi\rangle, |x_k\rangle\langle x_k|)$. Normalization and non-negativity can be deduced trivially from Hartle's description, but they are not useful for the proof. Apart from these, Hartle uses only one additional noncontextuality postulate:

(F1) *State noncontextuality and ONC postulate:* The frequency of obtaining outcome x_i in a quantum measurement

depends only on the measured state $|\psi\rangle$ and the projective operator $|x_i\rangle\langle x_i|$.

The structure of Hartle's proof is as follows [14]. He aims to show that the frequency of obtaining an outcome x_k in infinitely many measurements of the pure state $|\psi\rangle$ is equal to the probability given by the Born rule, $|\langle\psi|x_k\rangle|^2$. He defines a frequency operator f_N^k , which measures the number of times the outcome x_k occurs in N measurements:

$$f_N^k = \sum_{i_1=1}^d \sum_{i_2=1}^d \cdots \sum_{i_N=1}^d \left(\frac{1}{N} \sum_{\alpha=1}^N \delta_{k i_\alpha} \right) |x_{i_1}\rangle_1 \dots |x_{i_N}\rangle_N \langle x_{i_N}|_N \dots \langle x_{i_1}|_1. \quad (5)$$

This operator f_N^k is a composition of projections onto N eigenspaces with degenerate eigenvalues ranging from $1/N$ to 1 [8,14]. Hartle then shows that, in the limit $N \rightarrow \aleph_0$,

$$\|f^k |\psi\rangle^{\otimes \aleph_0} - |\langle x_k | \psi \rangle|^2 |\psi\rangle^{\otimes \aleph_0}\| = 0.$$

Certainly, some special treatment is required to define f^k as the limit of f_N^k when $N \rightarrow \aleph_0$. We omit the discussion of this step here, since it is not directly relevant to our focus. Hartle's treatment can be found in Ref. [14], and criticisms by others can be found in Refs. [8,28].

We want to point out a new self-incoherence in Hartle's proof that has not been discussed in the literature: Hartle bases his proof on the conjecture that the infinite ensemble $|\psi\rangle^{\otimes \aleph_0}$ of any pure state $|\psi\rangle$ is an eigenstate of the frequency operator f^k , and he proves this conjecture at the end [14]. If this is true, then it seems trivial to conclude that the infinite ensemble $\rho^{\otimes \aleph_0}$ of any mixed state ρ should also be an eigenstate of f^k —strictly speaking, $\rho^{\otimes \aleph_0}$ is the projector onto an eigenstate. We refer to it simply as an eigenstate for simplicity.

We now check whether this can be true. Let a mixed state be

$$\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|.$$

In his proof, Hartle sets $f^k |S\rangle = 0$ if $|S\rangle \notin \mathcal{J}$, where $\mathcal{J} = \overline{\text{Span}\{|\psi\rangle^{\otimes \aleph_0} \mid |\psi\rangle \in \mathcal{H}, \langle\psi|\psi\rangle = 1\}}$ [14]. This indicates that he intentionally avoids discussing mixed states. From this definition, we have

$$f^k \rho^{\otimes \aleph_0} = f^k \sum_j p_j^{\aleph_0} (|\psi_j\rangle\langle\psi_j|)^{\otimes \aleph_0} + 0 + 0 + \dots$$

Therefore, $f^k \rho^{\otimes \aleph_0} = 0$ unless $\exists p_j = 1$, which means that ρ is a pure state. On the contrary, if we calculate the result for a mixed state by applying the result for pure states, then we obtain

$$f^k \sum_j p_j (|\psi_j\rangle\langle\psi_j|)^{\otimes \aleph_0} = \sum_j p_j |\langle a_k | \psi_j \rangle|^2 (|\psi_j\rangle\langle\psi_j|)^{\otimes \aleph_0}.$$

Thus, these two calculations, both consistent with Hartle's key idea, lead to different results.

As mentioned above, Hartle avoids discussing mixed states in his proof. From his analysis, we can conclude that he does not regard $\rho^{\otimes \aleph_0}$ as an eigenstate of f^k . However, this is inconsistent with the frequentist conjecture of the proof. Moreover, the reason why the infinite state corresponding to ρ should be $\sum_j p_j (|\psi_j\rangle\langle\psi_j|)^{\otimes \aleph_0}$ rather than $\rho^{\otimes \aleph_0}$ is not

explained. Clearly, these two infinite states describe different experimental setups.

We can resolve this problem by adding a countable additivity postulate to Hartle's proof.

(F2') *Additivity postulate for states*: The frequency function μ satisfies the following countable additivity property:

$$\mu\left(\sum_j p_j |\psi_j\rangle\langle\psi_j|, |x_k\rangle\langle x_k|\right) = \sum_j p_j \mu(|\psi_j\rangle, |x_k\rangle\langle x_k|)$$

for any $p_j \geq 0$ and disjoint pure states $|\psi_j\rangle$'s.

Although additivity is usually assumed for compatible observables, it can also be applied to disjoint states, since the roles of the two are interchangeable in experiments. We can always regard the experiment as measuring some property of the measurement device using a given quantum state rather than measuring the state with the device. We can therefore retain Hartle's proof for pure states while assuming this additivity property. In that case, infinite ensembles of mixed states need not be eigenstates of f^k , and the contradiction is removed. This proves Observation 3.

V. CONCLUSION

In this work, we demonstrate the significance of additivity as an additional assumption in deriving the Born rule. In Sec. III, we have proved that additivity cannot be derived from the other two additional assumptions, noncontextuality and normalization. In Sec. IV, we have shown its crucial role in five important existing derivations of the Born rule. In Gleason's theorem and Busch's extended Gleason theorem, it is the core postulate. In Deutsch's derivation, it is implicitly assumed and also plays a central role. In Zurek's derivation, an attempt is made to weaken the additivity assumption, but this leads to the problem of continuity. In Hartle's derivation, the lack of an additivity assumption causes incoherence in his main idea when mixed states are considered. To summarize, we not only prove the necessity of the additivity assumption in these five derivations but also show that none of the proposed equivalent assumptions in these works can serve as a substitute for it.

Additivity is an assumption with a clear probabilistic nature, whereas noncontextuality and normalization are not. Our results not only facilitate the goal of determining the minimal set of additional postulates required to derive the Born rule but also help to better understand the origin of probability in quantum mechanics. In the existing literature, some authors argue that additional assumptions beyond the postulates of standard quantum mechanics are essential for deriving the Born rule [17], and some even believe that these additional assumptions must themselves have probabilistic properties [17,20]. Our analysis of the five seminal derivations supports both viewpoints: The Born rule cannot be derived solely from the nonprobabilistic postulates of quantum mechanics, and all attempts to avoid using the probabilistic assumption of additivity fail in the derivations we examine. This suggests that we may need to reconsider the hope of deriving the probabilistic quantum postulate of the Born rule solely from deterministic principles.

We have also given an organized set of quantum mechanics postulates in Sec. II. It clearly shows which assumptions are regarded as additional. It is also worth noting that different derivations can use different postulates from nonprobabilistic quantum mechanics. For example, the State postulate (QM1) is not used in Gleason's or Busch's proofs but is used in the other three; the Observable postulate (QM3) or the POVM postulate (QM3') is used in all other four derivations except Zurek's, and as a compensation, Zurek's derivation is also the only one that assumes the Evolution postulate (QM2). This difference is interesting and deserves further exploration.

In Sec. II and Appendix A5, we also give clear and distinguishable definitions for various types of additivity, noncontextuality, normalization, and continuity. In the literature we discuss, ambiguity in these concepts can easily cause confusion, leading people to incorrectly conclude that two assumptions are equivalent or that one method can be applied to prove a property in another situation. Our clarification helps to eliminate these confusions. This fills a gap in the literature and is beneficial for related research in the field.

Our analyses of the five derivations are mostly based on the first version of each method. This is because we include many different methods of deriving the Born rule, and it is impossible to examine every version of every method in a single paper. Although these initial versions may contain loopholes that are better addressed in later work, they represent the origins of the method chains with the clearest physical intuitions. Discussion of these versions is the most important first step in further investigating each method. In the future, we plan to discuss Wallace's proof of the Deutsch-Wallace theorem [36], and the improved version of Finkelstein-Hartle theorem by Farhi, Goldstone, and Gutmann [28]. We also plan to discuss derivations of the Born rule proposed by supporters of Bohmian mechanics [45] and QBism [23].

Readers may have noticed the importance of noncontextuality in the derivations we discussed. We do not provide a more complete discussion of the role of noncontextuality in this work, but this point will also be addressed in subsequent works.

ACKNOWLEDGMENTS

We thank Vlatko Vedral for valuable discussions and continuous support throughout this work. We are also grateful to Chiara Marletto for inspiring discussions on the Deutsch-Wallace theorem and to Fabrizio Logiurato for helpful clarifications concerning his work on noncontextuality and the Born rule. We also thank Qin Li for constructive suggestions that helped improve the manuscript.

DATA AVAILABILITY

There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

APPENDIX

1. Tables for important concepts

In this work, we discuss several different definitions of three important properties of a measurement function μ : addi-

tivity, noncontextuality, and continuity. Although all of these definitions are presented clearly in the paper, we believe that providing tables comparing their differences can be helpful to readers. Therefore, we summarize the notations used throughout this paper in Table I. We also summarize the various definitions of additivity, noncontextuality, and continuity in Tables II, III, and IV, respectively.

2. Equivalence between Gleason's additivity assumption and strong normalization

We prove that, for projective operators, the additivity assumption (or the frame function assumption) used by Gleason [10] is equivalent to strong normalization when $W = 1$.

The definition of a frame function μ of weight W can be written in the following equivalent form: if a countable set $\{|i\rangle\}_i$ forms an orthonormal basis of $\mathcal{H}$ —the definition of a separable Hilbert space guarantees the existence of such bases—then

$$\sum_i \mu(|i\rangle\langle i|) = W,$$

where W is a fixed real value for a given μ and $\mathcal{H}$.

To prove our result, we first introduce two lemmas.

Lemma A1. We can decompose any orthogonal projection A on a separable real or complex Hilbert space into a sum of disjoint rank-one orthogonal projections $|i\rangle\langle i|$, i.e.,

$$A = \sum_{i \in \mathcal{I}_A} |i\rangle\langle i|.$$

Proof. An orthogonal projection A always has eigenvalues equal to 0 or 1. Therefore, it can be transformed into a diagonal matrix, where the diagonal entries with value 1 correspond to the $|i\rangle\langle i|$'s. ■

Lemma A2 (Ref. [10]). If we have a frame function μ on $\mathcal{H}$ with weight W , then it is also a (sub-)frame function when restricted to any closed subspace of $\mathcal{H}$. The weight of the subframe function is usually different from W .

Proof. We can fix one $|i_k\rangle$. Then

$$\sum_{i \neq i_k} f(|i\rangle) = W - f(|i_k\rangle).$$

The right-hand side of the equation is constant for all choices of $\{|i\rangle\}_{i \neq i_k}$, which form bases of the closed subspace. By induction, we can prove this result for any subspace of any dimension. ■

Lemma A3. A function μ on projective operators is a frame function with weight $W = 1$ if and only if it is strongly normalized.

Proof. If μ is strongly normalized, then by definition it is a frame function. Conversely, if μ is a frame function, to test whether μ is strongly normalized, then we need to define $\mu(A)$ for arbitrary projective operator A . Using the additivity property of the frame function and Lemma A2, it is natural to define

$$\mu(A) = W_A = W - \sum_{j \notin \mathcal{I}_A} \mu(|j\rangle\langle j|) = \sum_{i \in \mathcal{I}_A} \mu(|i\rangle\langle i|).$$

Since $W = 1$, μ is strongly normalized. ■

TABLE I. Notations used in this paper and appendixes.

Notations	Explanation
MWI	Many worlds interpretation
POVM $\{A_i\}$	Positive operator valued measure (see Definition 1)
POVM effect A_i	One operator A_i in POVM $\{A_i\}$ (see Definition 1)
Measurement function μ	A real-valued function assumed to represent the result of a quantum measurement (see Postulate (M1))
Frame function μ	A measurement function μ assumed by Gleason satisfying strong normalization (see Definition 8)
Deutsch's value function $\mathcal{V}[\psi\rangle, X]$	The expectation value of the reward obtained by a rational player in a quantum measurement game, where the measurement outcome is the reward
$g \circ \mu$	Composite function: $(g \circ \mu)(x) = g(\mu(x))$
ONC	Observable noncontextuality (see Definition 4)
ANC	Additivity noncontextuality (see Definition 5)
Normalization	$\mu(I) = 1$ (see Definition 6)
Strong normalization	$\sum_i \mu(A_i) = 1$ if $\sum_i A_i = I$ (see Definition 7)
$\sup$	Supremum
$\inf$	Infimum
Span	Closed linear span
$\mathbb{R}$	The set of the real numbers
$\mathbb{Q}$	The set of the rational numbers
$\mathbb{N}$	The set of the natural numbers
$\aleph_0$	Cardinality of the set of the natural numbers, i.e., $\aleph_0 = \mathbb{N} $
$ \mathcal{I}_m $ for set $\mathcal{I}_m$	Cardinality of set $\mathcal{I}_m$
$\langle \psi x_i \rangle$	Inner product of $ \psi\rangle$ and $ x_i\rangle$
$ \langle \psi x_i \rangle $	Absolute value of $\langle \psi x_i \rangle$
Event κ	A subset of the sample space of an experiment, or a possible outcome of the experiment, usually assigned a probability
Event $\kappa_1 \subseteq \kappa_2$	Inclusion of events, meaning that if κ_1 occurs, then κ_2 certainly occurs
Event $x_1 \vee x_2$	Union of events, meaning that event x_1 occurs or event x_2 occurs
$X \otimes Y$	Tensor product of operators X and Y
$ \psi\rangle^{\otimes \aleph_0}$	$ \psi\rangle^{\otimes \aleph_0} = \underbrace{ \psi\rangle \otimes \psi\rangle \otimes \psi\rangle \otimes \cdots \otimes \psi\rangle}_{\aleph_0}$
Frequency operator f_N^k	The operator acting on a quantum state that measures the number of times the outcome x_k occurs in N repeated measurements on the state [see Eq. (5)]
Frequency operator f^k	The operator acting on a quantum state that measures the number of times the outcome x_k occurs in $\aleph_0$ repeated measurements on the state

Remark A1. Although ANC is often assumed together with additivity, both Gleason's additivity and strong normalization admit versions that do not require ANC.

3. Zurek's proof for the observable postulate

Zurek does not assume the observable postulate (QM3). Instead, he derives it from the state postulate (QM1), the Evolution postulate (QM2), and the eigenstate postulate (QM4). We rephrase his proof as follows [7].

Suppose the only two possible eigenstates of a quantum state $|\psi\rangle$ measured by an observable X are $|x_1\rangle$ and $|x_2\rangle$. From the MWI and (QM2), the measurement process is described by an entangling unitary evolution U_X . Then, by the Eigenstate postulate (QM4),

$$U_X |x_1\rangle |E_0\rangle = |x_1\rangle |E_1\rangle$$

$$U_X |x_2\rangle |E_0\rangle = |x_2\rangle |E_2\rangle,$$

where $|E_0\rangle$ denotes the initial state of the measurement device and $|E_1\rangle$ and $|E_2\rangle$ denote the states corresponding to $|x_1\rangle$ and

$|x_2\rangle$, respectively. Then,

$$\langle E_0 | \langle x_1 | U_X^\dagger U_X | x_2 \rangle | E_0 \rangle = \langle E_1 | \langle x_1 | x_2 \rangle | E_2 \rangle,$$

which implies

$$\langle x_1 | x_2 \rangle (1 - \langle E_1 | E_2 \rangle) = 0.$$

Zurek then concludes that if we do not want $\langle E_1 | E_2 \rangle = 1$, which would mean that the outcomes are indistinguishable, we must have $\langle x_1 | x_2 \rangle = 0$, namely the eigenstates are orthogonal. This argument applies to any two distinct eigenstates in any dimension. Thus, the orthogonality of eigenstates is established [7].

If we further assume that the eigenvalues are real values, then the Observable postulate (QM3) follows. This assumption is natural, since all readings on a measurement device are real values.

4. Zurek's weak additivity assumption and the usual countable additivity

Zurek shows how his weak additivity (E1) leads to our familiar countable additivity as follows [13].

TABLE II. Comparison of different additivity assumptions and properties. We list their positions of definition in the paper, the objects to which they apply, and whether the ONC property can be derived from them. The phrase “Usually Yes” means that the corresponding additivity is often implicitly assumed to contain the ANC property (and therefore the ONC property), although in our analysis we list the ONC postulate separately. The strengths of these additivity assumptions are compared in Fig. 1.

	Position of definition	Additivity for	Include ONC?
Countable additivity (CA)	Definition 2	Disjoint projections or compatible POVM effects	Usually Yes
Finite additivity (FA)	Definition 3	Disjoint projections or compatible POVM effects	Usually Yes
Strong normalization (SN)	Definition 7	Disjoint projections or compatible POVM effects	Usually Yes
Gleason’s frame function (GA)	Definition 8	Disjoint projections	Usually Yes
Busch’s additivity (BA)	Postulate (G-B1)	Compatible POVM effects	Usually Yes
Detusch’s additivity (DA)	Postulate (D1)	Rewards of games	Yes (at the step of proving the result for unequal probabilities)
Detusch’s hidden additivity (DHA)	In Postulate (D1)	Disjoint $ x_i\rangle\langle x_i $ ’s corresponding to eigenstates of an observable X	No
Zurek’s weak additivity (ZA)	Postulate (E1)	Any two mutually exclusive events κ and $\kappa^\perp$	No
Additivity added for Finkelstein-Hartle theorem (FHA)	Postulate (F2’)	Disjoint pure states	(This additivity is not for observables)

In the situations of interest, events are described by measurement outcomes. For equal-amplitude states, if there are n eigenstates, then we know from Zurek’s preceding proof that $\mu(x_i) = 1/n$ for any outcome x_i .

By Zurek’s weak additivity (E1),

$$\mu(x_1 + \cdots + x_{n-1}) + \mu(x_n) = 1$$

$$\mu(x_1 + \cdots + x_{n-1}) = 1 - \frac{1}{n}.$$

Zurek then claims that the conditional probability satisfies

$$\mu(x_1 + \cdots + x_{n-2} | x_1 + \cdots + x_{n-1}) = 1 - \frac{1}{n-1}.$$

We clarify the reasoning here: since the envariant swaps still apply to the outcomes $x_1, \dots, x_{n-1}$ when x_n is ruled out, we can still conclude that $\mu(x_1) = \mu(x_2) = \cdots = \mu(x_{n-1}) = \frac{1}{n-1}$. Then,

$$\mu(x_1 + \cdots + x_{n-2}) = \left(1 - \frac{1}{n}\right) \left(1 - \frac{1}{n-1}\right) = \frac{n-2}{n}.$$

We can permute x_i ’s arbitrarily and repeat this multiplication multiple times. This yields

$$\mu\left(\sum_{j \in \mathcal{I}_m} x_j\right) = \frac{m}{n},$$

where $m = |\mathcal{I}_m|$.

Thus, we obtain the additivity property for states with equal amplitudes [13]. We can easily extend this additivity to states with unequal but rational probabilities for the outcomes using the auxiliary system method. That is, if the outcome events $\{x_i\}_i$ are mutually disjoint and each $\mu(x_i) \in \mathbb{Q}$, then

$$\mu\left(\sum_i x_i\right) = \sum_i \mu(x_i).$$

However, there is no way to extend this additivity to states with real probabilities using this method. Therefore, Zurek’s additivity is weaker than Gleason’s additivity or the hidden countable additivity in Deutsch’s proof. We cannot guarantee the continuity of μ from Zurek’s weak additivity assumption, even when combined with non-negativity (a counterexample is presented in Sec. IV D).

5. Measure, additive function, and continuity

We discuss the continuity issue in this section. This clarifies why questions about continuity frequently appear in derivations of the Born rule. First, we introduce the definition and some important properties of measures and additive functions from mathematics. Second, we distinguish between five different notions of continuity that appear in the literature (Remark A3): Although Gleason’s theorem starts from a measure, the notion of continuity used in its proof is not the same

TABLE III. Comparison of different noncontextuality assumptions and properties. We list their positions of definition in the paper, the objects to which they apply, and whether the ONC property can be derived from them.

	Position of definition	Noncontextuality for	Include ONC?
ONC	Definition 4	Operator A for $\mu(A)$	Yes
ANC	Definition 5	Some additivity property of μ	Yes
State noncontextuality	Postulate (E3)	State ρ for $\mu(\rho)$	No

TABLE IV. Comparison of different continuity properties. We list their positions of definition in the paper and the objects to which they apply.

	Position of definition	Continuity for
Continuity for an additive function $\mu : \mathbb{R}^N \rightarrow \mathbb{R}$		variable $x \in \mathbb{R}^N$
Continuity for Busch's derivation		From rational to real coefficients for a POVM effect A
Continuity required in Deutsch's and Zurek's derivations	Remark A3	From rational to real amplitudes for an eigenstate $ x_i\rangle$
Gleason's continuity for a frame function		A unit vector x rotating by a small angle
Continuity of measure concerning inclusion of sets		Sets $A_1 \subset A_2 \subset \dots$

as the usual continuity discussed in measure theory [10,19]. These two are also very different from the continuity used in Busch's derivation [11]. Continuity for Busch's derivation, together with the continuity required in Deutsch's [12] and Zurek's proofs [13], are closer to each other and to the continuity for additive functions [46], but they are still not identical. This clarification is useful because it is easy to confuse these different notions of continuity and mistakenly assume that a proof for one kind of continuity will work for another. Third, we explain why an additive function can be discontinuous. The conditions for an additive function to be continuous are given in Theorem A3 and Theorem A4—this helps explain why Gleason introduces the non-negativity assumption to ensure continuity [10]. Finally, in Theorem A5 we show how Busch's additivity implies his continuity (a step that is not explicit in his original paper) [11]. Our proof makes clear how the flexibility of scaling POVM effects makes a simple proof for continuity possible. This flexibility is absent for additivity restricted to projective operators, as used in Gleason's, Deutsch's, and Zurek's derivations.

Definition A1 (Ref. [19]). Let X be a set and Σ a σ algebra on X . A *measure* μ on X is a function $\mu : \Sigma \rightarrow \mathbb{R} \cup \{+\infty\}$ such that

- (i) If $\{A_i\}$ is a sequence of disjoint sets in Σ , then $\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$ (countable additivity)
- (ii) $\mu(\emptyset) = 0$
- (iii) $\mu \geq 0$

Definition A2 ([46]). An *additive function* is a function $f : \mathbb{R}^N \rightarrow \mathbb{R}$ that satisfies

$$f(|x\rangle + |y\rangle) = f(|x\rangle) + f(|y\rangle)$$

for any $|x\rangle, |y\rangle \in \mathbb{R}^N$. We will also write these vectors as x, y for simplicity. This equation is called the *Cauchy functional equation*.

Remark A2. We should note a subtle but important difference in the definitions of measure and additive function: The codomain of a measure may include $+\infty$, while the codomain of an additive function does not. Therefore, properties of additive functions do not always apply to measures. In Gleason's [10] and Busch's [11] derivations, which are based on measures or probability measures, the normalization and non-negativity assumptions are essential to rule out the output value $+\infty$.

Also, for our purposes, we cannot always directly use results for additive functions, since our function μ takes an operator A as input rather than a vector $|x\rangle$. In the end, we will find that μ is linear with respect to any $|x\rangle\langle x|$ rather than with respect to $|x\rangle$. We analyze this in more detail in Remark A5 below.

We now list some important properties of additive functions.

Lemma A4. If μ is an additive function, then μ satisfies finite additivity but not necessarily countable additivity.

Proof. It is immediate from the definition that an additive function satisfies finite additivity. For countable additivity, there is a well-known counterexample. Let A be a subset of the natural numbers $\mathbb{N}$. Define $\mu(A) = 0$ if $|A|$ is finite and $\mu(A) = 1$ if $|A|$ is infinite. This function μ is finitely additive but not countably additive [47]. ■

Lemma A5 (Ref. [46]). For an additive function f ,

$$f(qx) = qf(x)$$

for any $q \in \mathbb{Q}$ and $x \in \mathbb{R}^N$.

Proof. From finite additivity, we have $f(my) = mf(y)$ for $m \in \mathbb{N}$ and $y \in \mathbb{R}^N$.

Let $y = x/n$ for $n \in \mathbb{N}$. Then

$$f\left(m\frac{x}{n}\right) = mf\left(\frac{x}{n}\right) = \frac{m}{n}f(x) \quad (m, n \in \mathbb{N}, x, y \in \mathbb{R}^N)$$

$$f(qx) = qf(x) \quad \left(q = \frac{m}{n} \in \mathbb{Q}, x \in \mathbb{R}^N\right).$$

Notice Lemma A5 shows that an additive function is linear for any rational coefficient, but it is not always linear for real coefficients and thus is not always a linear function.

We will show how an additive function can be nonlinear and discontinuous in Lemma A6. To do that, we first distinguish between five different continuities.

Remark A3. There exist different definitions of *continuity* in the literature which indicate very different properties. It is therefore easy to confuse them. We list five different notions of continuity closely related to our topic below:

(1) *Continuity for an additive function:* An additive function $\mu : \mathbb{R}^N \rightarrow \mathbb{R}$ is continuous if

$$\lim_{x \rightarrow y} f(x) - f(y) = 0.$$

Here the equal sign indicates the $(\epsilon-\delta)$ definition of limit. We omit an explanation of this elementary definition.

(2) *Continuity for Busch's derivation* [11] *from rational to real coefficients:* A real-valued function $\mu(A)$ (where A is a POVM effect) is continuous if for any $q \in \mathbb{Q}$ and $r \in \mathbb{R}$,

$$\lim_{q \rightarrow r} \mu(qA) - \mu(rA) = 0.$$

Certainly, for genuine continuity we should consider arbitrary $q, r \in \mathbb{R}$. However, for the purpose of extending the linearity of μ from rational to real coefficients in front of A , the notion

of continuity discussed here is enough. A proof of continuity for Busch's derivation from his postulates is given below in Theorem A5.

(3) *Continuity required in Deutsch's [12] and Zurek's [13] derivations:* The value function μ is continuous if, for two different quantum states $|\psi_1\rangle$ and $|\psi_2\rangle$, in which a common eigenstate $|x_i\rangle$ has amplitudes $q \in \mathbb{Q}$ and $r \in \mathbb{R}$ in the two states, respectively, we always have

$$\lim_{q \rightarrow r} \mu(|\psi_1\rangle, |x_i\rangle\langle x_i|) - \mu(|\psi_2\rangle, |x_i\rangle\langle x_i|) = 0.$$

These three continuities for additive functions, Busch's derivation, and Deutsch's and Zurek's proofs, are related but distinct. Note that in Busch's notion the coefficients q, r multiply a POVM effect A , while in Deutsch's and Zurek's notion the coefficients q, r are amplitudes in front of an eigenstate $|x_i\rangle$.

(4) *Gleason's continuity for a frame function:* Let μ be a frame function. μ is continuous if for any x we can find a neighborhood W of x such that [10]

$$\sup\{f(y)|y \in W\} - \inf\{f(y)|y \in W\} \leq \epsilon \quad (\epsilon \rightarrow 0).$$

In Gleason's paper [10] he actually mentions two different continuity notions. The first one appears at the beginning of the paper when he discusses the existence of discontinuous additive functions. The second one is the notion we show here, which he uses as a crucial step of the proof: Any continuous frame function on $\mathbb{R}^3$ is regular (Theorem 2.3 in Ref. [10]).

(5) *Continuity of measure concerning inclusion of sets:* In measure theory, a measure μ is continuous from below if, for sets $A_1 \subset A_2 \subset \dots$, we have [19]

$$\mu(\cup_{i=1}^{\infty} A_i) = \lim_{j \rightarrow \infty} \mu(A_j).$$

The definition for continuity from above is analogous.

There is a classical result in measure theory that if μ is countably additive, then it is also continuous in this sense [19]. However, this continuity is very different from continuities for Gleason's or Busch's derivation, which are the types used to derive the Born rule. Gleason's continuity concerns how μ changes when we slightly rotate $|x\rangle$, while the continuity of measure concerns how μ changes when we enlarge the set $A = |x\rangle\langle x|$.

We can now show how an additive function can be discontinuous. The following theorems are important results from the theory of additive functions.

Theorem A1 (Ref. [48]). $\mathbb{R}^N$ is a vector space over the field $\mathbb{Q}$. Its basis is called a *Hamel basis*, which has cardinality $\aleph_1$ and cannot be written down explicitly. There are many different Hamel bases, as long as the basis vectors are linearly independent.

Theorem A2. (Ref. [46]). If H is a Hamel basis of $\mathbb{R}^N$, then for any function $g : H \rightarrow \mathbb{R}$ there exists a unique additive function $f : \mathbb{R}^N \rightarrow \mathbb{R}$ such that

$$f(h_i) = g(h_i)h_i \in H.$$

f is continuous if and only if for every $h_i \in H$, we have $\frac{g(h_i)}{h_i} = C_0$, where C_0 is a constant.

Proof. The proof is omitted here and can be found in Ref. [46]. ■

Lemma A6. An additive function is not always continuous.

Proof. From Theorem A1 and Theorem A2, we can construct a Hamel basis for $\mathbb{R}^N$ which contains two basis vectors h_1, h_2 such that $h_1 \rightarrow h_2$. They are guaranteed to be Hamel basis vectors as long as $\frac{h_1}{h_2} \notin \mathbb{Q}$. Define

$$\begin{aligned} g(h_1) &= C_1 h_1 \\ g(h_2) &= C_2 h_2, \end{aligned}$$

where $C_1, C_2 \in \mathbb{R}$. We can let $C_1 \ll C_2$, for example, $C_1 = 1$, $C_2 = 10\,000$. Then we can construct the additive function f based on g . For this f one has

$$h_1 \rightarrow h_2, \quad \text{but } f(h_1) = h_1 \ll f(h_2) = 10\,000h_2.$$

So additive function f is not continuous. ■

Remark A4. The measure function μ used by Gleason takes operators as inputs, not vectors. This is why Gleason introduces another additive function $g : \mathbb{R} \rightarrow \mathbb{R}$ and claims that $g \circ \mu$ is still a frame function and can be wildly discontinuous [10]. All such pathological frame functions are unbounded (see Theorem A3 below). Gleason then introduces the non-negativity assumption to guarantee boundedness and thus continuity.

We now present the conditions for an additive function to be a linear functional.

Theorem A3 (Ref. [46]). An additive function $f : \mathbb{R}^N \rightarrow \mathbb{R}$ has the form

$$f(|x\rangle) = \langle c|x\rangle = \sum_{i=1}^N c_i x_i$$

if it satisfies any of the conditions below:

- (a) f is continuous.
- (b) f is continuous at some point x .
- (c) f is bounded below or above on some interval.
- (d) $f \geq 0$ or $f \leq 0$ for all x .

Proof. (a) If $N = 1$, since we already know $f(qx) = qf(x)$ when $q \in \mathbb{Q}$ from Lemma A5, continuity implies $f(rx) = rf(x)$ for all $r \in \mathbb{R}$. Let $c = f(1)$; then $f(x) = xf(1) = cx$ for any $x \in \mathbb{R}$ [46].

If $N > 1$, because of additivity we can write $f(|x\rangle) = f(x_1) + \dots + f(x_N)$. From the one-dimensional result, each $f(x_i) = c_i x_i$, so $f(|x\rangle) = \langle c|x\rangle = \sum_{i=1}^N c_i x_i$ [46].

(b) Continuity at a single point extends to the whole domain (see Theorem A4 below). The result follows from (a).

(c) The full proof is omitted here. The key idea is that any additive function is a convex function, which satisfies

$$f\left(\frac{x+y}{2}\right) \leq \frac{f(x) + f(y)}{2}$$

for all $x, y \in \mathbb{R}^N$ [46]. It can be shown that a convex function that is bounded above or below at any point is bounded above or below everywhere and that a bounded convex function is continuous. Readers can see Ref. [49] for more details. The result then follows from (a).

(d) The result follows directly from (c). ■

Theorem A4. If an additive function f is continuous at some $x \in \mathbb{R}^N$, then it is continuous on the whole domain.

Proof. From Lemma A5, $f(0) = 0$. If f is continuous at a point x , then by additivity

$$\lim_{\sigma \rightarrow 0} f(x + \sigma) = f(x) + \lim_{\sigma \rightarrow 0} f(\sigma).$$

Hence $\lim_{\sigma \rightarrow \vec{0}} f(\sigma) = 0$, so f is continuous at $\vec{0}$. It is then trivial to prove f is continuous at any point. ■

Remark A5. If $f(|x\rangle) = \langle c|x\rangle$, then f is a linear functional. However, we cannot directly apply this result to derive the Born rule. We would like to point out two important differences between the additive function f and our real-valued function μ here: First, f takes $|x\rangle$ as input, and not all operators A in the domain of μ can be written as $A = |x\rangle\langle x|$. Only projective operators have such property. Second, by definition, an additive function f takes $|x\rangle \in \mathbb{R}^N$ as input, whereas the eigenstates of A may contain complex entries.

Thus, for our purpose, we can do the following modifications: Any POVM effect A can be written as $A = \sum_i \eta_i |x_i\rangle\langle x_i|$ ($\eta_i \geq 0$). Because of additivity assumed separately, we have

$$\mu(A) = \sum_i \eta_i \mu(|x_i\rangle\langle x_i|).$$

From Theorem A3, it is easy to show that if an additive function f takes $|x\rangle\langle x|$ as input and satisfies conditions such as continuity, then

$$f(|x\rangle\langle x|) = \langle x| \left(\sum_j \lambda_j |c_j\rangle\langle c_j| \right) |x\rangle (\lambda_j \geq 0).$$

This is also true when we extend the domain of $|x\rangle$ to $\mathbb{C}^N$. Clearly, $(\sum_j \lambda_j |c_j\rangle\langle c_j|) \geq 0$. We can let $\rho = (\sum_j \lambda_j |c_j\rangle\langle c_j|)$ and rescale it to obtain a trace-one operator. Combining the two equations above yields the usual Born rule,

$$\mu(\rho, A) = \text{Tr}[\rho A].$$

Note that this is not a rigorous proof. We only intend to explain the idea of how one can obtain the function μ for the Born rule from an additive function f when their domains differ. A rigorous proof is just Busch's derivation [11].

We now show how continuity for Busch's derivation can be derived from Busch's postulates: countable additivity, normalization, and non-negativity [11]. Note that this proof does not work for Gleason's derivation [10], because Gleason's countable additivity degenerates to finite additivity when the Hilbert space is finite dimensional.

Theorem A5. If real-valued function $\mu(A)$ (A is any POVM effect) satisfies the following conditions:

(i) If effects $A + B + \dots \leq I$, $\mu(A + B + \dots) = \mu(A) + \mu(B) + \dots$.

(ii) $\mu(I) = 1$.

(iii) For any $A \geq 0$, $\mu(A) \geq 0$.

Then μ satisfies the continuity for Busch's derivation, i.e., for any $q \in \mathbb{Q}$ and $r \in \mathbb{R}$,

$$\lim_{q \rightarrow r} \mu(qA) - \mu(rA) = 0.$$

Proof. We first prove $\lim_{\sigma \rightarrow 0} \mu(\sigma A) = 0$ for any POVM effect A . From countable additivity, we have

$$\mu(A) = \mu\left(\sum_{i=1}^{\infty} \frac{A}{2^i}\right) = \sum_{i=1}^{\infty} \mu\left(\frac{A}{2^i}\right).$$

Since $\mu(A) \in \mathbb{R}$, the series on the right-hand side must converge. This guarantees that

$$\lim_{i \rightarrow +\infty} \mu\left(\frac{A}{2^i}\right) = 0$$

and hence

$$\lim_{\sigma \rightarrow 0} \mu(\sigma A) = 0.$$

Notice this is not true for μ that is only finitely additive. In that case $\lim_{\sigma \rightarrow 0} \mu(\sigma A)$ can be nonzero because there is no infinite terms in the sum.

We can now prove the continuity of μ . To prove that μ is continuous at rA ($r \in \mathbb{R}$), we need to show

(i) $p(rA)$ exists for any $r \in \mathbb{R}$.

(ii) $p(rA) = \lim_{q \rightarrow r} p(qA)$, where $q \in \mathbb{Q}$.

The first step (i) is trivial since rA is an effect. For (ii), if q approaches r from below, then we write $rA = qA + \sigma A$ with $\sigma \geq 0$. By additivity,

$$p(rA) = p(qA + \sigma A) = p(qA) + p(\sigma A).$$

Thus,

$$\lim_{q \rightarrow r} p(qA) = p(rA) - \lim_{\sigma \rightarrow 0} p(\sigma A) = p(rA).$$

A similar proof applies if q approaches r from above. Therefore μ is continuous in the stated notion. ■

-
- [1] T. Norsen, The measurement problem, in *Foundations of Quantum Mechanics: An Exploration of the Physical Meaning of Quantum Theory* (Springer International, Cham, 2017), pp. 59–85.
 - [2] A. A. Tomaz, R. S. Mattos, and M. Barbatti, The quantum measurement problem: A review of recent trends, *Philos. Mag.* **1** (2025).
 - [3] M. Born, On the quantum mechanics of collision processes [English translation], *Terry's Archive Online* **2021**, 0828 (2021).
 - [4] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, Cambridge, UK, 2010).
 - [5] P. A. M. Dirac, *The Principles of Quantum Mechanics* (Clarendon Press, Oxford, 1958).
 - [6] J. von Neumann and R. T. Beyer, *Mathematical Foundations of Quantum Mechanics: New Edition* (Princeton University Press, Princeton, NJ, 2018).
 - [7] W. H. Zurek, Quantum theory of the classical: Quantum jumps, Born's rule and objective classical reality via quantum Darwinism, *Philos. Trans. R. Soc. A* **376**, 20180107 (2018).
 - [8] C. M. Caves and R. Schack, Properties of the frequency operator do not imply the quantum probability postulate, *Ann. Phys.* **315**, 123 (2005).
 - [9] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski, Information causality as a physical principle, *Nature (London)* **461**, 1101 (2009).
 - [10] A. M. Gleason, Measures on the closed subspaces of a Hilbert space, *J. Math. Mech.* **6**, 885 (1957).

- [11] P. Busch, Quantum states and generalized observables: A simple proof of Gleason's theorem, *Phys. Rev. Lett.* **91**, 120403 (2003).
- [12] D. Deutsch, Quantum theory of probability and decisions, *Proc. R. Soc. Lond. Ser. A* **455**, 3129 (1999).
- [13] W. H. Zurek, Probabilities from entanglement, Born's rule $p_k = |\psi_k|^2$ from envariance, *Phys. Rev. A* **71**, 052105 (2005).
- [14] J. B. Hartle, Quantum mechanics of individual systems, *Am. J. Phys.* **36**, 704 (1968).
- [15] D. Finkelstein, Section of physical sciences: The logic of quantum physics, *Trans. N. Y. Acad. Sci.* **25**, 621 (1963).
- [16] H. Everett, "Relative state" formulation of quantum mechanics, *Rev. Mod. Phys.* **29**, 454 (1957).
- [17] L. Vaidman, Derivations of the Born rule, in *Quantum, Probability, Logic: The Work and Influence of Itamar Pitowsky*, edited by M. Hemmo and O. Shenker (Springer International, Cham, 2020), pp. 567–584.
- [18] F. Logiurato and A. Smerzi, Born rule and noncontextual probability, *J. Mod. Phys.* **03**, 1802 (2012).
- [19] G. Folland, *Real Analysis: Modern Techniques and Their Applications*, Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts (Wiley, New York, 2013).
- [20] H. Barnum, C. M. Caves, J. Finkelstein, C. A. Fuchs, and R. Schack, Quantum probability from decision theory? *Proc. R. Soc. Lond. Ser. A* **456**, 1175 (2000).
- [21] C. M. Caves, C. A. Fuchs, and R. Schack, Quantum probabilities as Bayesian probabilities, *Phys. Rev. A* **65**, 022305 (2002).
- [22] C. M. Caves, C. A. Fuchs, K. K. Manne, and J. M. Renes, Gleason-type derivations of the quantum probability rule for generalized measurements, *Found. Phys.* **34**, 193 (2004).
- [23] J. B. DeBroya, C. A. Fuchs, J. L. Pienaar, and B. C. Stacey, Born's rule as a quantum extension of bayesian coherence, *Phys. Rev. A* **104**, 022207 (2021).
- [24] A. Esmaili Zaghi, Natural emergence of the Born rule from relational quantum dynamics (RQD) (2025), https://doi.org/10.31219/osf.io/j73ak_v1.
- [25] A. Proxmire, The Born rule as a forced theorem of event density: A Gleason-Busch reconstruction from first principles (2026), <https://philpapers.org/rec/PROTBR>.
- [26] S. Saunders, Derivation of the Born rule from operational assumptions, *Proc. R. Soc. A* **460**, 1771 (2004).
- [27] A. J. Short, Probability in many-worlds theories, *Quantum* **7**, 971 (2023).
- [28] E. Farhi, J. Goldstone, and S. Gutmann, How probability arises in quantum mechanics, *Ann. Phys.* **192**, 368 (1989).
- [29] J. S. Bell, On the problem of hidden variables in quantum mechanics, *Rev. Mod. Phys.* **38**, 447 (1966).
- [30] M. Schlosshauer and A. Fine, On Zurek's derivation of the Born rule, *Found. Phys.* **35**, 197 (2005).
- [31] J. Lawrence, The Born rule: Axiom or result? *Am. J. Phys.* **92**, 371 (2024).
- [32] R. Cooke, M. Keane, and W. Moran, An elementary proof of Gleason's theorem, *Math. Proc. Cambr. Philos. Soc.* **98**, 117 (1985).
- [33] A. Grudka and P. Kurzyński, Is there contextuality for a single qubit? *Phys. Rev. Lett.* **100**, 160401 (2008).
- [34] G. Hermann, Determinism and quantum mechanics, in *Grete Hermann—Between Physics and Philosophy*, edited by E. Crull and G. Bacciagaluppi (Springer Netherlands, Dordrecht, 2016), pp. 223–237.
- [35] S. Axler, Linear maps, in *Linear Algebra Done Right* (Springer International, Cham, 2024), pp. 51–118.
- [36] D. Wallace, How to prove the Born rule, in *Many Worlds? Everett, Quantum Theory, and Reality*, edited by S. Saunders, J. Barrett, A. Kent, and D. Wallace (Oxford University Press, Oxford, 2010), pp. 227–263.
- [37] L. Vaidman, Probability in the many-worlds interpretation of quantum mechanics, in *Probability in Physics*, edited by Y. Ben-Menahem and M. Hemmo (Springer, Berlin, 2012), pp. 299–311.
- [38] H. Price, Decisions, decisions, decisions: Can Savage salvage Everettian probability? in *Many Worlds? Everett, Quantum Theory, and Reality*, edited by S. Saunders, J. Barrett, A. Kent, and D. Wallace (Oxford University Press, Oxford, 2010), pp. 369–390.
- [39] R. Dawid and K. P. Thébault, Against the empirical viability of the Deutsch–Wallace–Everett approach to quantum mechanics, *Stud. Hist. Philos. Sci. Part B* **47**, 55 (2014).
- [40] M. Ridley, Many retrocausal worlds: A foundation for quantum probability, *Eur. J. Philos. Sci.* **15**, 70 (2025).
- [41] H. Barnum, No-signalling-based version of Zurek's derivation of quantum probabilities: A note on "environment-assisted invariance, entanglement, and probabilities in quantum physics", [arXiv:quant-ph/0312150](https://arxiv.org/abs/quant-ph/0312150).
- [42] C. M. Caves, Notes on Zurek's derivation of the quantum probability rule (2005), <http://info.phys.unm.edu/caves/reports/ZurekBornDerivation.pdf>.
- [43] W. H. Zurek, Decoherence, einselection, and the quantum origins of the classical, *Rev. Mod. Phys.* **75**, 715 (2003).
- [44] E. J. Squires, On an alleged "proof" of the quantum probability law, *Phys. Lett. A* **145**, 67 (1990).
- [45] D. Dörr, S. Goldstein, and N. Zanghí, Quantum equilibrium and the origin of absolute uncertainty, *J. Stat. Phys.* **67**, 843 (1992).
- [46] Additive functions and convex functions, in *An Introduction to the Theory of Functional Equations and Inequalities: Cauchy's Equation and Jensen's Inequality*, edited by A. Gilányi (Birkhäuser Basel, Basel, 2009), pp. 117–142.
- [47] R. Ash and C. Doleans-Dade, *Probability and Measure Theory* (Elsevier Science, Amsterdam, 2000).
- [48] Algebra, in *An Introduction to the Theory of Functional Equations and Inequalities: Cauchy's Equation and Jensen's Inequality*, edited by A. Gilányi (Birkhäuser Basel, Basel, 2009), pp. 75–114.
- [49] Elementary properties of convex functions, in *An Introduction to the Theory of Functional Equations and Inequalities: Cauchy's Equation and Jensen's Inequality*, edited by A. Gilányi (Birkhäuser Basel, Basel, 2009), pp. 143–159.