Classical probability measures as pullback of quantum probability measures?

April 10, 2022

As discussed in Sec. 1.1, the quantum probability spaces could be thought as a big space glued by "local" classical probability space defined by each orthonormal basis Ω according to the map $\varphi : 2^{\Omega} \to \mathcal{E}$ defined in Eq. (4) which sends a classical event $E \in 2^{\Omega}$ to a quantum event $\varphi(E) \in \mathcal{E}$. Then, for any quantum probability measure $\mu\colon \mathcal{E} \to [0,1]$, the function $(\lambda E.\mu(\varphi(E))) : 2^{\Omega} \to [0,1]$ defined by precomposition is a classical probability measure. Since that "precomposition" is called "pullback" in differential geometry [53][1], we will use their notation and write the precomposition $(\lambda E.\mu(\varphi(E)))$ as $\lambda E.(\varphi^{*}\mu)(E)$, i.e., $(\varphi^{*}\mu)(E) = \mu(\varphi(E))$, and $\varphi^{*}\mu\colon 2^{\Omega} \to [0,1]$ the pullback of μ by φ, and summarize how "pullback" preserves their structures in Table 1. We can then extend this table to compare between the classical and quantum measurement processes to get real numbers. It is clear that each step of the measurement process is exactly equal or just mapped by φ except we summarize classical measurement process as random variable, but summarize quantum measurement process as an observable. Notice that after we fix an orthonormal basis Ω, φ is actually an invertible map so that given an observable $\mathbf{O} = \sum_{i=0}^{D-1} \lambda_i P_i$, we can define an unique random variable $\sum_{i=0}^{D-1} \lambda_i \mathbf{1}_{\varphi^{-1}(P_i)}$ corresponding to $\mathbf{O}$. Therefore, we call this random variable as the pullback of $\mathbf{O}$ by φ, and denoted by $\varphi^{*}\mathbf{O}$.[2] According to Gleason's theorem, the table could be extended to simplify the formula computing expectation values and the the post-measurement-state postulates after we define a notation to send probability measures to states

[1] Whether this kind of "pullback" is categorical will need further investigate.

[2] We might also choose the other way around so that each random variable has a push forward observable. This is better in the sense that we don't have the ambiguity that $\varphi^{*}\mathbf{O}$ is might be different for different orthonormal basis, but every theorem for IVPMs starts from the quantum part with an observable, so we need the ambiguous $\varphi^{*}\mathbf{O}$ for each different orthonormal basis anyway.

Table 1: Comparison between the classical and quantum measurement processes with real-valued probability for finite dimensional cases

Name	Classical	Quantum	How does the structure preserve?
Event	$E \in 2^\Omega$	$\varphi(E) = \sum_{\lvert j\rangle \in E} \lvert j\rangle\langle j\rvert \in \mathcal{E}$	Eqs. (6) and (7)
Probability measure	$\varphi^*\mu\colon \begin{array}{ccc} 2^\Omega & \to & [0,1] \\ E & \mapsto & \mu(\varphi(E)) \end{array}$	$\mu\colon \begin{array}{ccc} \mathcal{E} & \to & [0,1] \\ P & \mapsto & \mu(P) \end{array}$	Eqs. (8) and (9)
Possible measured values	$\lambda_i \in \mathbb{R}$	$\lambda_i \in \mathbb{R}$	Equal
The event to measure λ_i	$E_i \in 2^\Omega$	$\varphi(E_i) \in \mathcal{E}$	Mapped by φ
The probability that the measured value is λ_i	$(\varphi^*\mu)(E_i) = \mu(\varphi(E_i)) \in [0,1]$	$\mu(\varphi(E_i)) \in [0,1]$	Equal
Summarize the measurement process as	random variable $\varphi^*\mathbf{O} = \sum_{i=0}^{D-1} \lambda_i \mathbf{1}_{E_i}$	observable $\mathbf{O} = \sum_{i=0}^{D-1} \lambda_i \varphi(E_i)$	
Expectation value	$\int (\varphi^*\mathbf{O})\,\mathrm{d}(\varphi^*\mu) = \sum_{i=0}^{D-1} \lambda_i (\varphi^*\mu)(E_i)$	$\langle \mathbf{O} \rangle_\mu = \sum_{i=0}^{D-1} \lambda_i \mu(\varphi(E_i))$	Equal, and proved in Lemma 1
Probability measure could be simplified as	probability mass function $m : \Omega \to [0,1]$ such that $(\varphi^*\mu)(E) = \sum_{\omega \in E} m(\omega)$	mixed state $\rho = \sum_{j=1}^{N} q_j \lvert\Phi_j\rangle\langle\Phi_j\rvert$ such that $\mu(P) = \mathrm{Tr}(\rho P)$	
	Need a notation for $\varphi^*\mu \mapsto m$ to keep going	Need a notation for $\mu \mapsto \rho$ to keep going	We may draw commuting diagram (?) here
The probability that the measured value is λ_i			
Expectation value			
Post-measurement state			

and the other way around.

By following the same process, Sec. 2 showed the corresponding results for interval-valued probability. For any QIVPM $\bar{\mu}\colon \mathcal{E} \to \mathscr{I}$, the function $\varphi^{*}\bar{\mu}\colon 2^{\Omega} \to \mathscr{I}$ defined by $(\varphi^{*}\bar{\mu})(E) = \bar{\mu}(\varphi(E))$ is a classical IVPM.

1 Quantum Probability

A *probability space* is a mathematical abstraction specifying the necessary conditions for reasoning coherently about collections of uncertain events [14, 15, 26, 38]. Although they can be used to describe an individual quantum experiment, to describe a family of quantum experiments, we would like to glue their probability spaces together to define a quantum probability space. The glued quantum probability space is well-behaved since not only we can define the expectation values, but the quantum probability measure defined on the whole space can be simply induced by the Born rule according to Gleason's theorem [12, 17, 40–42].

1.1 Classical and Quantum Probability Spaces

Given a finite sample space Ω representing all possible outcomes of a process, and its power set 2^{Ω} as the classical event space, a classical probability measure μ maps every event $E \subseteq \Omega$ to a number $\mu(E) \in [0,1]$ specifying how likely one of the outcomes in E will happen. To maintain the coherence, μ is subject to the following constraints: $\mu(\emptyset) = 0$, $\mu(\Omega) = 1$, and $\mu\left(\overline{E}\right) = 1 - \mu(E)$, where $\overline{E}$ is the complement of E. Moreover, we require $\mu(E_0 \cup E_1) = \mu(E_0) + \mu(E_1)$ for each pair of disjoint events $E_0 \subseteq \Omega$ and $E_1 \subseteq \Omega$.

Since the previous abstraction doesn't specify the process to generate the outcomes, this process could well be a quantum experiment. Let us prepare a beam of one kind of spin 1 particles whose state can be characterized by a vector in three-dimensional Hilbert space with basis vectors $|0\rangle$, $|1\rangle$, and $|2\rangle$. In principle, this beam can be split by a Stern-Gerlach type experiment according to the eigenvalues of the observable $\mathbf{O}_0$ with spectral decomposition

$$\mathbf{O}_0 = 0\ |0\rangle\langle 0| + 1\ |1\rangle\langle 1| + 2\ |2\rangle\langle 2|\ , \tag{1}$$

and the states corresponding to the split beams are $|0\rangle$, $|1\rangle$, and $|2\rangle$ [40, 50]. Instead of sending a beam of particles, if only one particle is sent to the beam splitter, the state of the particle after the experiment is one of $|0\rangle$, $|1\rangle$, and $|2\rangle$, and there is a probability to get each post-experimental state. In the language of our abstraction, the set of outcomes is $\Omega_0 = \{|0\rangle, |1\rangle, |2\rangle\}$. All

Table 2: Comparison between the classical and quantum measurement processes with interval-valued probability for finite dimensional cases

Name	Classical	Quantum	How does the structure preserve?
Probability measure	$\begin{aligned} \varphi^*\bar{\mu}\colon 2^{\Omega} &\to \mathscr{I} \\ E &\mapsto \bar{\mu}\left(\varphi\left(E\right)\right) \end{aligned}$	$\begin{aligned} \bar{\mu}\colon \mathcal{E} &\to \mathscr{I} \\ E &\mapsto \bar{\mu}\left(E\right) \end{aligned}$	

possible events are $\emptyset$, $\{|0\rangle\}$, $\{|1\rangle\}$, $\{|2\rangle\}$, $\{|0\rangle, |1\rangle\}$, $\{|0\rangle, |2\rangle\}$, $\{|1\rangle, |2\rangle\}$, and Ω_0. And this experiment defines a classical probability measure $\mu_0 \colon 2^{\Omega_0} \to [0,1]$.

One special feature of quantum experiments is that the probabilities in different experiments are correlated. Consider another experiment sending exactly the same particle as the previous one but using a different beam splitter corresponding to the observable $\mathbf{O}_1$ with spectral decomposition $\mathbf{O}_1 = 0\ |\text{+}\rangle\langle\text{+}| + 1\ |\text{-}\rangle\langle\text{-}| + 2\ |2\rangle\langle 2|$, where $|\text{+}\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$ and $|\text{-}\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}$. Although the sample space $\Omega_1 = \{|\text{+}\rangle, |\text{-}\rangle, |2\rangle\}$ and the probability measure $\mu_1 \colon 2^{\Omega_1} \to [0,1]$ defined by this experiment is different from the previous one, these two experiments may produce the same post-experimental state $|2\rangle$, and the probability of the common event $\{|2\rangle\}$ is believed to be the same, i.e.,

$$\mu_0\left(\{|2\rangle\}\right) = \mu_1\left(\{|2\rangle\}\right) . \tag{2}$$

In general, as long as sending the same particle, the probability of the same event in different experiments should always be the same. This fact is equivalent to the fact that commuting observables could be measured simultaneously which is essential to define contextuality and will be explained in Sec. 3.

Since the probability induced by the different beam splitters are correlated, it is more natural to define one quantum event space $\mathcal{E}$ containing all possible classical event spaces using different beam splitters. However, simply taking the union of all event spaces is a bad idea because two events might appear different but represent the same situation. For example, if we take the complement on both sides of Eq. (2), we will have

$$\mu_0\left(\{|0\rangle, |1\rangle\}\right) = 1 - \mu_0\left(\{|2\rangle\}\right) = 1 - \mu_1\left(\{|2\rangle\}\right) = \mu_1\left(\{|\text{+}\rangle, |\text{-}\rangle\}\right) , \tag{3}$$

that is, the probabilities of events $\{|0\rangle, |1\rangle\}$ and $\{|\text{+}\rangle, |\text{-}\rangle\}$ are always the same, and these events should be identified as the same quantum event to simplify our discussion. This identification can be achieved by mapping a classical event E to the projector generated by E,

$$\varphi\left(E\right) = \sum_{|j\rangle \in E} |j\rangle\langle j| \tag{4}$$

with the convention $\varphi\left(\emptyset\right) = \nvdash$, because $\varphi\left(\{|0\rangle, |1\rangle\}\right)$ is equal to $\varphi\left(\{|\text{+}\rangle, |\text{-}\rangle\}\right)$ as operators,

$$\varphi\left(\{|0\rangle, |1\rangle\}\right) = |0\rangle\langle 0| + |1\rangle\langle 1| = |\text{+}\rangle\langle\text{+}| + |\text{-}\rangle\langle\text{-}| = \varphi\left(\{|\text{+}\rangle, |\text{-}\rangle\}\right) . \tag{5}$$

In general, if two classical events E and E' are mapped to the same projector, i.e., $\varphi(E) = \varphi(E')$, then the probability of E is the same as the probability of E'. Therefore, for any classical event E, its corresponding quantum event is defined to be the projector $\varphi(E)$, and the set of all projectors on a given Hilbert space is called a quantum event space $\mathcal{E}$.

This function φ not only respects the probability of events but also naturally sends the set structure to the corresponding projector structure:

$$\varphi(\Omega) = \nVdash , \qquad \varphi\left(\overline{E}\right) = \nVdash - \varphi(E) \ . \tag{6}$$

Given two *commuting* projectors P_0 and P_1, there exists a pair of events E_0 and E_1 in the same sample space Ω such that $P_0 = \varphi(E_0)$ and $P_1 = \varphi(E_1)$. Conversely, given a pair of events E_0 and E_1 in the same sample space Ω, their corresponding quantum events $\varphi(E_0)$ and $\varphi(E_1)$ are commuting and satisfying the following properties:

$$\varphi(E_0 \cap E_1) = \varphi(E_0)\,\varphi(E_1) \ , \quad \varphi(E_0 \cup E_1) = \varphi(E_0) + \varphi(E_1) - \varphi(E_0)\,\varphi(E_1) \ . \tag{7}$$

Moreover, E_0 and E_1 are disjoint if and only if $\varphi(E_0)$ and $\varphi(E_1)$ are orthogonal, where two projectors P_0 and P_1 are called *orthogonal* if $P_0 P_1 = \nvdash$.

Then, a quantum probability space can be defined as a quantum event space $\mathcal{E}$ together with a quantum probability measure $\mu\colon \mathcal{E} \to [0, 1]$ subject to the corresponding constraints [1, 12, 28, 29, 41]:

$$\mu(\nvdash) = 0 \, , \qquad \mu(\nVdash) = 1 \, , \qquad \mu(\nVdash - P) = 1 - \mu(P) \ , \tag{8}$$

and for each pair of orthogonal projectors P_0 and P_1:

$$\mu(P_0 + P_1) = \mu(P_0) + \mu(P_1) \ . \tag{9}$$

Because φ respects the probability of events, if we restrict the domain of φ on a classical event space 2^Ω, the function $\varphi^*\mu\colon 2^\Omega \to [0, 1]$ defined by precomposition

$$(\varphi^*\mu)(E) = \mu(\varphi(E)) \tag{10}$$

is a classical probability measure and called the pullback of μ by $\varphi\colon 2^\Omega \to \mathcal{E}$.

Given a Hilbert space $\mathcal{H}$ of dimension D and a probability assignment for every projector P, we can define the expectation value of an observable $\mathbf{O}$ having spectral decomposition $\mathbf{O} = \sum_{i=0}^{D-1} \lambda_i P_i$, with eigenvalues $\lambda_i \in \mathbb{R}$, as [22, 38]:

$$\langle \mathbf{O} \rangle_\mu = \sum_{i=0}^{D-1} \lambda_i \mu(P_i) \ , \tag{11}$$

where the subscript μ might be omitted if it is clear according to the context. This definition is also consistent with the classical expectation values because we can pullback an observable to a classical random variable, and the expectation values are invariant.

Definition 1 (Pullback of Observables)**.** Consider an observable $\mathbf{O}$ diagonalizable by an orthonormal basis $\Omega = \{|0\rangle, |1\rangle, \ldots, |D-1\rangle\}$ so that $\mathbf{O}$ has spectral decomposition $\mathbf{O} = \sum_{i=0}^{D-1} \lambda_i |i\rangle\langle i|$. If we restrict φ on the classical event space 2^Ω and consider $\varphi\colon 2^\Omega \to \mathcal{E}$, then the pullback of $\mathbf{O}$ by φ is a random variable $\varphi^*\mathbf{O}\colon \Omega \to \mathbb{R}$ defined by $\varphi^*\mathbf{O} = \sum_{i=0}^{D-1} \lambda_i \mathbf{1}_{\{|i\rangle\}}$, where $\mathbf{1}_E$ is the indicator function defined by

$$\mathbf{1}_E(\omega) = \begin{cases} 1 & \text{if } \omega \in E\,; \\ 0 & \text{if } \omega \notin E\,. \end{cases} \tag{12}$$

Lemma 1. *Consider an observable* $\mathbf{O}$ *diagonalizable by an orthonormal basis* $\Omega = \{|0\rangle, |1\rangle, \ldots, |D-1\rangle\}$ *with* $\varphi\colon 2^\Omega \to \mathcal{E}$ *defined by Eq. (4). Given a quantum probability measure* $\mu\colon \mathcal{E} \to [0,1]$*, the expectation value of* $\mathbf{O}$ *relative to* μ *is exactly the expectation value of the pullback of* $\mathbf{O}$ *relative to the pullback of* μ*, i.e.,*

$$\langle \mathbf{O} \rangle_\mu = \int (\varphi^*\mathbf{O})\, \mathrm{d}(\varphi^*\mu)\ . \tag{13}$$

Proof. By Eqs. (11), (4), and (10), we have

$$\langle \mathbf{O} \rangle_\mu = \sum_{i=0}^{D-1} \lambda_i \mu(|i\rangle\langle i|) = \sum_{i=0}^{D-1} \lambda_i \mu(\varphi(\{|i\rangle\})) = \sum_{i=0}^{D-1} \lambda_i (\varphi^*\mu)(\{|i\rangle\}) = \int (\varphi^*\mathbf{O})\, \mathrm{d}(\varphi^*\mu)\ .$$

□

1.2 Gleason's Theorem and the Born Rule

After introducing quantum probability measures, we might follow the convention to introduce the Born rule which is the only way to relate a state to a quantum probability measure in CQT. However, since the variants of quantum probability measures in Sec. **??** and Chapter **??** could not be constructed by a "Born rule" easily, searching for a Born rule step-by-step here might provide a better idea of what we should do in Sec. **??** and Chapter **??**. While the situations will become more delicate in the later sections, we will fortunately find the unique Born rule for CQT here.

Although the quantum probability measure constructed by gluing together classical ones looks complex, it could be induced by an operator according to Gleason's theorem [12, 17, 40–42].

Theorem 1 (Gleason's theorem)**.** *In a Hilbert space $\mathcal{H}$ of dimension $D \geq 3$, given a quantum probability measure $\mu \colon \mathcal{E} \to [0,1]$, there exists a unique mixed state $\rho = \sum_{j=1}^{N} q_j \left|\Phi_j\right\rangle\!\left\langle\Phi_j\right|$ such that $\mu(P) = \mathrm{Tr}(\rho P)$ for any D-dimensional projector P, where $\left|\Phi_j\right\rangle \in \mathcal{H}$ are normalized, $q_j > 0$, and $\sum_{j=1}^{N} q_j = 1$.*

If we follow the discussion of Stern-Gerlach type experiments in the previous section to interpret Gleason's theorem, we can find Gleason's theorem doesn't specify whether its unique mixed state ρ characterizes the state of particle sending to the quantum experiment or not. Consider an extreme example: a quantum theory could ignore the input state and predict the experimental outcomes are always equally probable. Even if these predictions form a quantum probability measure, this kind of prediction is so different from the prediction of CQT that they can be easily distinguished experimentally.

Let a pure unnormalized state $|\Phi\rangle \in \mathcal{H}$ characterize the particle sending to a Stern-Gerlach type experiment, and μ_Φ^{B} be the quantum probability measure of the resulting events sensibly corresponding to $|\Phi\rangle$. For a correspondence $|\Phi\rangle \mapsto \mu_\Phi^{\mathrm{B}}$ to be sensible, we hope that if the state $|\Phi\rangle$ is one of the outcomes of a quantum event P, $P|\Phi\rangle = |\Phi\rangle$, then the event P always happens,

$$\mu_\Phi^{\mathrm{B}}(P) = 1\,, \tag{14}$$

and vice versa. Moreover, since the physical phenomena exist and should be the same no matter how we describe them, the probability of an event should be invariant despite how we choose the basis. Because changing to another basis is the same as applying a unitary map U, we should have

$$\mu_{U|\Phi\rangle}^{\mathrm{B}}\left(UPU^\dagger\right) = \mu_\Phi^{\mathrm{B}}(P)\,. \tag{15}$$

It is easy to check that the correspondence satisfying these conditions is unique,

$$\mu_\Phi^{\mathrm{B}}(P) = \frac{\langle\Phi|P|\Phi\rangle}{\langle\Phi|\Phi\rangle} \tag{16}$$

and called the Born rule [6, 22, 34]. If $|\Phi\rangle$ is normalized, Eq. (16) could be simplified as $\mu_\Phi^{\mathrm{B}}(P) = \langle\Phi|P|\Phi\rangle$. Since a mixed state $\rho = \sum_{j=1}^{N} q_j \left|\Phi_j\right\rangle\!\left\langle\Phi_j\right|$ is a weighted average of projectors $\left|\Phi_j\right\rangle\!\left\langle\Phi_j\right|$ with weights $q_j > 0$ and $\sum_{j=1}^{N} q_j = 1$, the generalized Born rule of ρ, $\mu_\rho^{\mathrm{B}}(P)$, is also a weighted average of $\mu_{\Phi_j}^{\mathrm{B}}(P)$,

$$\mu_\rho^{\mathrm{B}}(P) = \sum_{j=1}^{N} q_j \mu_{\Phi_j}^{\mathrm{B}}(P) = \mathrm{Tr}(\rho P)\,, \tag{17}$$

Table 3: Two fragments of valid probability measures μ_1 and μ_2.

$\lvert\Psi\rangle$	$\lvert 0\rangle$	$\lvert 1\rangle$	$\lvert 2\rangle$	$\frac{\lvert 0\rangle+\lvert 1\rangle}{\sqrt{2}}$	$\frac{\lvert 0\rangle+\mathrm{i}\lvert 1\rangle}{\sqrt{2}}$	$\frac{\lvert 0\rangle+\lvert 2\rangle}{\sqrt{2}}$	$\frac{\lvert 0\rangle+\mathrm{i}\lvert 2\rangle}{\sqrt{2}}$	$\frac{\lvert 1\rangle+\lvert 2\rangle}{\sqrt{2}}$	$\frac{\lvert 1\rangle+\mathrm{i}\lvert 2\rangle}{\sqrt{2}}$	$\dots$
$\mu_1\left(\lvert\Psi\rangle\langle\Psi\rvert\right)$	$\frac{1}{2}$	$\frac{1}{2}$	0	1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\dots$
$\mu_2\left(\lvert\Psi\rangle\langle\Psi\rvert\right)$	$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	$\dots$

which is also a quantum probability measure and consistent with Gleason's theorem.

As an example, consider a three-dimensional Hilbert space with orthonormal basis $\{\lvert 0\rangle, \lvert 1\rangle, \lvert 2\rangle\}$ and the observable $\mathbf{O}_0$ defined in Eq. (1). Two fragments of valid probability measures μ_1 and μ_2 that can be associated with this space are defined in Table 3. By the Born rule, the first probability measure corresponds to the quantum system being in the pure state $\lvert +\rangle = \frac{\lvert 0\rangle+\lvert 1\rangle}{\sqrt{2}}$ and the second corresponds to the quantum system being in the state $\frac{\lvert 0\rangle\langle 0\rvert+\lvert 2\rangle\langle 2\rvert}{2}$. The expectation values of the observable $\mathbf{O}$, $\langle\mathbf{O}\rangle_{\mu_{1,2}}$, are 1.5 in the first case and 2 in the second. The quantum expectation value can also be used to decide whether a state is entangled or not for multipartite systems as we describe in the following section.

2 Intervals of Uncertainty

2.1 Definitions of Classical and Quantum IVPMs

We will start by reviewing classical IVPMs and then propose our quantum generalization. In the classical setting, there are several proposals for "imprecise probabilities" [9, 11, 14, 21, 23, 30, 46, 52]. Although these proposals differ in some details, they all share the fact that the probability $\mu(E)$ of an event E is generalized from a single *real number* to an *interval* $[\ell, r]$, where ℓ intuitively corresponds to the strength of evidence for the event E and $1-r$ corresponds to the strength of the evidence against the same event. Under some additional assumptions, this interval could be interpreted as the Gaussian width of a probability distribution.

We next introduce probability axioms for IVPMs. First, for each interval $[\ell, r]$ we have the natural constraint $0 \le \ell \le r \le 1$ that guarantees that every element of the interval can be interpreted as a conventional probability.

We also include $\mathsf{F} = [0, 0]$ and $\mathsf{T} = [1, 1]$ as limiting intervals that refer, respectively, to the probability interval for impossible events and for events that are certain. We can write the latter as $\mu(\emptyset) = \mathsf{F}$ and $\mu(\Omega) = \mathsf{T}$, where $\emptyset$ is the empty set and Ω is the event covering the entire sample space. For each interval $[\ell, r]$, we also need the dual interval $[1 - r, 1 - \ell]$ so that if one interval refers to the probability of an event E, the dual refers to the probability of the event's complement $\overline{E}$. For example, if we discover as a result of an experiment that $\mu(E) = [0.2, 0.3]$ for some event E, we may conclude that $\mu\left(\overline{E}\right) = [0.7, 0.8]$ for the complementary event $\overline{E}$. In addition to these simple conditions, there are some subtle conditions on how intervals are combined, which we discuss next.

Let E_0 and E_1 be two disjoint events with probabilities $\mu(E_0) = [\ell_0, r_0]$ and $\mu(E_1) = [\ell_1, r_1]$. A first attempt at calculating the probability of the combined event that *either* E_0 or E_1 occurs might be $\mu(E_0 \cup E_1) = [\ell_0 + \ell_1, r_0 + r_1]$. In some cases, this is indeed a sensible definition. For example, if $\mu(E_0) = [0.1, 0.2]$ and $\mu(E_1) = [0.3, 0.4]$ we get $\mu(E_0 \cup E_1) = [0.4, 0.6]$. But consider an event E such that $\mu(E) = [0.2, 0.3]$ and hence $\mu\left(\overline{E}\right) = [0.7, 0.8]$. The two events E and $\overline{E}$ are disjoint; the naïve addition of intervals would give $\mu\left(E \cup \overline{E}\right) = [0.9, 1.1]$, which is not a valid probability interval. Moreover, the event $E \cup \overline{E}$ is the entire space; its probability interval should be T which is sharper than $[0.9, 1.1]$. The problem is that the two intervals are correlated: there is more information in the combined event than in each event separately so the combined event should be mapped to a sharper interval. In our example, even though the "true" probability of E can be anywhere in the range $[0.2, 0.3]$ and the "true" probability of $\overline{E}$ can be anywhere in the range $[0.7, 0.8]$, the values are not independent. Any value of $\mu(E) \leq 0.25$ will force $\mu\left(\overline{E}\right) \geq 0.75$. To account for such subtleties, the axioms of interval-valued probability do not use strict equality for the combination of disjoint events. The correct constraint enforcing coherence of the probability assignment for $E_0 \cup E_1$ when E_0 and E_1 are disjoint is taken to be:

$$\mu(E_0 \cup E_1) \subseteq [\ell_0 + \ell_1, r_0 + r_1]\,. \tag{18}$$

Note that for any event E with $\mu(E) = [\ell, r]$, we always have

$$\mu(\Omega) = \mathsf{T} \subseteq [\ell, r] + [1 - r, 1 - \ell] = \mu(E) + \mu\left(\overline{E}\right)\,. \tag{19}$$

When combining non-disjoint events, there is a further subtlety whose resolution will give us the final general condition for IVPMs. For events E_0 and E_1, not necessarily disjoint, we have:

$$\mu(E_0 \cup E_1) + \mu(E_0 \cap E_1) \subseteq \mu(E_0) + \mu(E_1)\,, \tag{20}$$

which is a generalization of the classical inclusion-exclusion principle that uses $\subseteq$ instead of $=$ for the same reason as before. The new condition, known as *convexity* [11, 14, 30, 31, 37, 47], reduces to the previously motivated Eq. (18) when the events are disjoint, i.e., when $\mu(E_0 \cap E_1) = 0$.

Previous discussions can be summarized as the following definition [23].

Definition 2 (IVPM)**.** Assume a collection of intervals $\mathscr{I}$ including F and T with addition and scalar multiplication defined as follows:

$$[\ell_0, r_0] + [\ell_1, r_1] = [\ell_0 + \ell_1, r_0 + r_1] \text{ and} \tag{21a}$$

$$x[\ell, r] = \begin{cases} [x\ell, xr] & \text{for } x \geq 0\,; \\ [xr, x\ell] & \text{for } x \leq 0\,. \end{cases} \tag{21b}$$

Given a finite sample space Ω, and its power set 2^Ω as the classical event space, a classical IVPM $\bar{\mu} \colon 2^\Omega \to \mathscr{I}$ is a function subject to the following constraints:

$$\bar{\mu}(\emptyset) = \mathsf{F}\,, \tag{22a}$$

$$\bar{\mu}(\Omega) = \mathsf{T}\,, \tag{22b}$$

$$\bar{\mu}\left(\overline{E}\right) = \mathsf{T} - \bar{\mu}\left(E\right)\,, \tag{22c}$$

and satisfying the convexity condition, Eq. (20), for each pair of events E_0 and E_1.

Note that the minus sign appearing in Eq. (22c) is accommodated by the $x \leq 0$ case in Eq. (21b).

We now have the necessary ingredients to define the quantum extension, QIVPMs, as a generalization of both classical IVPMs and conventional quantum probability measures in Sec. 1. We will show that QIVPMs reduce to classical IVPMs when the space of quantum events $\mathcal{E}$ is restricted to mutually commuting events $\mathcal{E}_{\mathrm{C}}$, i.e., to compatible events that can be measured simultaneously. In Sec. 4 we will discuss the connection between QIVPMs and conventional quantum probability measures in detail.

Definition 3 (QIVPM)**.** We take a QIVPM $\bar{\mu}$ to be an assignment of an interval to each event (projection operator P) subject to the following constraints:

$$\bar{\mu}(\nvdash) = \mathsf{F}\,, \tag{23a}$$

$$\bar{\mu}(\nVdash) = \mathsf{T}\,, \tag{23b}$$

$$\bar{\mu}\left(\nVdash - P\right) = \mathsf{T} - \bar{\mu}\left(P\right)\,, \tag{23c}$$

and satisfying for each pair of *commuting* projectors P_0 and P_1 with $P_0 P_1 = P_1 P_0$,

$$\bar{\mu}\left(P_0 + P_1 - P_0 P_1\right) + \bar{\mu}\left(P_0 P_1\right) \subseteq \bar{\mu}\left(P_0\right) + \bar{\mu}\left(P_1\right) \,. \tag{24}$$

The first three constraints, Eqs. (23), are the direct counterpart of the corresponding ones for classical IVPMs, Eqs. (22). With the understanding that the union of classical sets $E_0 \cup E_1$ is replaced by $P_0 + P_1 - P_0 P_1$ in the case of quantum projection operators [15], the last condition, Eq. (24), is a direct counterpart of the convexity condition of Eq. (20). Thus, our definition of QIVPMs merges aspects of both classical IVPMs and quantum probability measures.

2.2 States Consistent with Classical and Quantum IVPMs

Our definition of QIVPMs is consistent with classical IVPMs in the sense that a restriction of QIVPMs to mutually commuting subspaces of events $\mathcal{E}_{\mathrm{C}}$ recovers the definition of classical IVPMs. To see this, we first define a subspace of events: $\mathcal{E}'$ is called a *subspace* of the set of events $\mathcal{E}$ if $\mathcal{E}'$ contains the projectors $\mathbb{0}$ and $\mathbb{1}$ and is closed under complements, sums, and products. In particular, for any projector $P \in \mathcal{E}'$, we have $\mathbb{1} - P \in \mathcal{E}'$ and for each pair of commuting projectors $P_0 \in \mathcal{E}'$ and $P_1 \in \mathcal{E}'$, we have $P_0 P_1$ and $P_0 + P_1 - P_0 P_1 \in \mathcal{E}'$. Given a mutually commuting subspace $\mathcal{E}_{\mathrm{C}}$ whose elements are diagonalizable by a common orthonormal basis $\Omega = \{|0\rangle, |1\rangle, \ldots, |D-1\rangle\}$, the function $\varphi\colon 2^{\Omega} \to \mathcal{E}$ defined in Eq. (4) maps any set E to the sum of the projectors formed by elements in E, and preserves all operations used to define classical and quantum IVPMs as in Eqs. (6) and (7). According to the same reason as the classical counterpart, Eq. (10), given a QIVPM $\bar{\mu}\colon \mathcal{E} \to \mathscr{I}$, the function $\varphi^{*}\bar{\mu}\colon 2^{\Omega} \to \mathscr{I}$ defined by precomposition $\left(\varphi^{*}\bar{\mu}\right)(E) = \bar{\mu}\left(\varphi\left(E\right)\right)$ is the pullback of $\bar{\mu}$ by φ and is a classical IVPM naturally.

Since we can pull back a QIVPM to a classical one, known properties of classical IVPMs directly hold for QIVPMs when one restricts to mutually commuting events $\mathcal{E}_{\mathrm{C}}$, and one of them is having a *core*. Recall in Sec. 2.1, given a classical IVPM $\bar{\mu}\colon 2^{\Omega} \to \mathscr{I}$ and an event $E \subseteq \Omega$ such that $\bar{\mu}(E) = [0.2, 0.3]$ and $\bar{\mu}(E) = [0.7, 0.8]$, we discussed the "true" probabilities, $\mu(E)$ and $\mu\left(\overline{E}\right)$, can be anywhere in the range $[0.2, 0.3]$ and $[0.7, 0.8]$, respectively, as long as they satisfy

$$\mu\left(E\right) + \mu\left(\overline{E}\right) = 1 \,. \tag{25}$$

Eq. (25) guarantees the function mapping an event to its "true" probabilities, $\mu\colon 2^{\Omega} \to [0, 1]$, is a probability measure. Requiring the "true" probability of each event should be in the range spanned by the interval-valued probability of the same event, i.e.,

$$\mu(E) \in [0.2, 0.3] = \bar{\mu}(E) \ , \qquad \mu\left(\overline{E}\right) \in [0.7, 0.8] = \bar{\mu}\left(\overline{E}\right) \ , \tag{26}$$

$$\mu(\emptyset) = 0 \in \mathsf{F} = \bar{\mu}(\emptyset) \ , \qquad \mu(\Omega) = 1 \in \mathsf{T} = \bar{\mu}(\Omega) \ , \tag{27}$$

gives the following definition of core.

Definition 4 (Classical Consistency and Core)**.** We say an IVPM $\bar{\mu}\colon 2^{\Omega} \to \mathscr{I}$ is *consistent* with a probability measure $\mu\colon 2^{\Omega} \to [0,1]$ on an event $E \subseteq \Omega$ if the interval $\bar{\mu}(E)$ contains the precise probability calculated by μ, i.e., $\mu(E) \in \bar{\mu}(E)$. The *core* of an IVPM $\bar{\mu}$, $\mathrm{core}(\bar{\mu})$, is the collection of all probability measures μ that are *consistent* with $\bar{\mu}$ on every event, that is,

$$\mathrm{core}(\bar{\mu}) = \left\{\mu\colon 2^{\Omega} \to [0,1] \;\middle|\; \forall E \subseteq \Omega,\ \mu(E) \in \bar{\mu}(E)\right\} \ . \tag{28}$$

One fundamental question of these imprecise interval-valued probabilities is whether they *always* have underlying precise probabilities. A positive answer is given by the following theorem.

Theorem 2 (Shapley [11, 14, 37, 47])**.** *Every classical IVPM has a non-empty core.*

Although it is *impossible* in the classical world to have an empty core, it is possible in the quantum world for the imprecise probabilities associated with some events to be inconsistent with *any* quantum state, i.e., a QIVPM might have an empty core as we will show in Sec. 4. In that case, one cannot guarantee non-empty cores for finite-precision attempts at proving Gleason's theorem by extending the Born measure $\mu^{\mathrm{B}}_{\rho}(P)$ to QIVPMs $\bar{\mu}(P)$. However, if we restrict ourselves to the set $\mathcal{E}_{\mathrm{C}}$ of mutually commuting events, the situation reverts to the classical case in which probabilities always determine at least one state.

We now give the necessary technical definition and lemma to prove this non-empty core property.

Definition 5 (Quantum Consistency and Core)**.** We say a QIVPM $\bar{\mu}$ is *consistent* with a state ρ on a projector P if the interval $\bar{\mu}(P)$ contains the exact probability calculated by the Born rule [6, 22, 34], i.e.,

$$\mu^{\mathrm{B}}_{\rho}(P) = \mathrm{Tr}(\rho P) \in \bar{\mu}(P) \ . \tag{29}$$

The *core* $\overline{\mathcal{H}}(\bar{\mu}, \mathcal{E}')$ of $\bar{\mu}$ relative to a subspace of events $\mathcal{E}'$ is the collection of all states ρ that are *consistent* with $\bar{\mu}$ on every projector in $\mathcal{E}'$, that is,

$$\overline{\mathcal{H}}\left(\bar{\mu}, \mathcal{E}'\right) = \left\{\rho \;\middle|\; \forall P \in \mathcal{E}',\ \mu^{\mathrm{B}}_{\rho}(P) \in \bar{\mu}(P)\right\} \ . \tag{30}$$

In contrast with the classical Thm. 2, there is *no guarantee* that there exists a state ρ that satisfies Eq. (30) and therefore is in the core of a QIVPM. However, for the special case of commuting events, since the classical core corresponds to the quantum one by the following lemma, there will be a quantum theorem naturally corresponding to classical Thm. 2.

Lemma 2. *Consider a QIVPM $\bar{\mu}\colon \mathcal{E} \to \mathscr{I}$ and a commuting subspace of events $\mathcal{E}_{\mathrm{C}}$ diagonalizable by a common orthonormal basis $\Omega = \{|0\rangle, |1\rangle, \ldots, |D-1\rangle\}$ with the function $\varphi\colon 2^{\Omega} \to \mathcal{E}$ defined by Eq. (4). For any classical probability measure μ consistent with the pullback of $\bar{\mu}$, i.e., $\mu \in \mathrm{core}\,(\varphi^{*}\bar{\mu})$, there is a density matrix ρ consistent with $\bar{\mu}$ relative to $\mathcal{E}_{\mathrm{C}}$, $\rho \in \overline{\mathcal{H}}\,(\bar{\mu}, \mathcal{E}_{\mathrm{C}})$, such that the pullback of μ_{ρ}^{B} is μ, i.e.,*

$$\left(\varphi^{*}\mu_{\rho}^{\mathrm{B}}\right)(E) = \mu_{\rho}^{\mathrm{B}}\left(\varphi\left(E\right)\right) = \mu\left(E\right) \tag{31}$$

for all $E \subseteq \Omega$.

Proof. Consider $\rho = \sum_{j=0}^{D-1} \mu\left(\{|j\rangle\}\right) |j\rangle\langle j|$ which is a density matrix because ρ is a positive operator and has trace equal to one [38]:

$$\mathrm{Tr}\left(\rho\right) = \sum_{j=0}^{D-1} \mu\left(\{|j\rangle\}\right) \mathrm{Tr}\left(|j\rangle\langle j|\right) = \sum_{j=0}^{D-1} \mu\left(\{|j\rangle\}\right) = 1\,. \tag{32}$$

Then, Eq. (31) can be proved by induction on E. By Eq. (4) and the generalized Born rule, Eq. (17), we have the base case

$$\mu_{\rho}^{\mathrm{B}}\left(\varphi\left(\{|i\rangle\}\right)\right) = \mu_{\rho}^{\mathrm{B}}\left(|i\rangle\langle i|\right) = \sum_{j=0}^{D-1} \mu\left(\{|j\rangle\}\right) \mu_{|j\rangle}^{\mathrm{B}}\left(|i\rangle\langle i|\right) = \mu\left(\{|i\rangle\}\right) \tag{33}$$

for all i. By Eq. (7) and the induction hypothesis $\mu_{\rho}^{\mathrm{B}}\left(\varphi\left(E\right)\right) = \mu\left(E\right)$, the inductive case is also valid:

$$\mu_{\rho}^{\mathrm{B}}\left(\varphi\left(E \cup \{|i\rangle\}\right)\right) = \mu_{\rho}^{\mathrm{B}}\left(\varphi\left(E\right)\right) + \mu_{\rho}^{\mathrm{B}}\left(\varphi\left(\{|i\rangle\}\right)\right) = \mu\left(E\right) + \mu\left(\{|i\rangle\}\right) = \mu\left(E \cup \{|i\rangle\}\right) \tag{34}$$

for all $i \notin E$. After we inductively proved $\varphi^{*}\mu_{\rho}^{\mathrm{B}} = \mu$, we can finally prove $\rho \in \overline{\mathcal{H}}\,(\bar{\mu}, \mathcal{E}_{\mathrm{C}})$. Since μ is consistent with the pullback of $\bar{\mu}$, $\varphi^{*}\bar{\mu}$, we have

$$\mu_{\rho}^{\mathrm{B}}\left(\varphi\left(E\right)\right) = \mu\left(E\right) \in \left(\varphi^{*}\bar{\mu}\right)(E) = \bar{\mu}\left(\varphi\left(E\right)\right)\,. \tag{35}$$

Together with the fact that the image of φ contains $\mathcal{E}_{\mathrm{C}}$, we have $\mu_{\rho}^{\mathrm{B}}\left(P\right) \in \bar{\mu}\left(P\right)$ for all $P \in \mathcal{E}_{\mathrm{C}}$, i.e., $\rho \in \overline{\mathcal{H}}\,(\bar{\mu}, \mathcal{E}_{\mathrm{C}})$. $\square$

Then, the quantum version of Thm. 2 is a natural consequence of the previous lemma.

Theorem 3 (Non-empty Core for Compatible Measurements)**.** *For every QIVPM $\bar{\mu}\colon \mathcal{E} \to \mathscr{I}$, if a subspace of events $\mathcal{E}_{\mathrm{C}} \subseteq \mathcal{E}$ commutes, then $\overline{\mathcal{H}}\left(\bar{\mu}, \mathcal{E}_{\mathrm{C}}\right) \neq \emptyset$.*

Proof. Since $\mathcal{E}_{\mathrm{C}}$ is a set of mutually commuting projections, they can be diagonalized by a common orthonormal basis $\Omega = \{\left|0\right\rangle, \left|1\right\rangle, \ldots, \left|D-1\right\rangle\}$. We can then follow Eq. (4) to define $\varphi\colon 2^{\Omega} \to \mathcal{E}$ such that the pullback of $\bar{\mu}$ by φ, $\varphi^{*}\bar{\mu}$, is a classical IVPM. By Thm. 2, there is a classical probability measure $\mu\colon 2^{\Omega} \to [0,1]$ consistent with $\varphi^{*}\bar{\mu}$, and thus there must be a density matrix $\rho \in \overline{\mathcal{H}}\left(\bar{\mu}, \mathcal{E}_{\mathrm{C}}\right)$ satisfying $\varphi^{*}\mu_{\rho}^{\mathrm{B}} = \mu$ according to Lemma 2. □

2.3 Classical Choquet Integrals and Expectation Values of Observables

We conclude this section with a generalization of expectation values of observables in the context of QIVPMs. In conventional quantum mechanics, the expectation value of an observable as defined in Eq. (11) is a unique real number. The generalization to QIVPMs implies that this expectation value should be bounded by an interval. We will start from the classical notion of the *Choquet integral* which is used to calculate the expectation value of random variables as a weighted average [8, 11, 14]. Then, we will generalize this notation from classical IVPMs to QIVPMs and prove the generalized definition consistent with our intuition. For example, if $\bar{\mu}$ is a conventional (Born) probability measure induced by a state ρ, then the interval expectation value collapses to a point, thus reducing the interval expectation value to the conventional definition of Eq. (11). We will also show the expectation value of an observable relative to a QIVPM $\bar{\mu}$ lies between two possible outcomes, which themselves lie between the minimum and maximum bounds of the probability intervals associated with each state ρ that is consistent with $\bar{\mu}$ on every projector in the spectral decomposition of the observable.

Like how we pulled back classical ideas to the quantum world in Secs. 2.1 and 2.2, we start from defining the classical expectation value relative to IVPMs. Consider a classical IVPM $\bar{\mu}\colon 2^{\Omega} \to \mathscr{I}$ and an event $E \subseteq \Omega$ such that

$$\bar{\mu}\left(E\right) = [0.2, 0.3] \ , \qquad \bar{\mu}\left(\overline{E}\right) = [0.7, 0.8] \ , \tag{36}$$

and a random variable $X\colon \Omega \to \mathbb{R}$ defined by

$$X(\omega) = \begin{cases} 1 & \text{if } \omega \in E\,; \\ 2 & \text{if } \omega \notin E\,. \end{cases} \tag{37}$$

As we have discussed, the "true" probability is a probability measure $\mu\colon 2^{\Omega} \to [0,1]$ consistent with $\bar{\mu}$ on every event. On one extreme case, $\mu(E) = 0.2$, the expectation value of X relative to μ is

$$\int X \mathrm{d}\mu = 1 \cdot \mu(E) + 2 \cdot \mu\left(\overline{E}\right) = 1 \cdot 0.2 + 2 \cdot 0.8 = 1.8\,. \tag{38}$$

On the other extreme case, $\mu(E) = 0.3$, the expectation value of X relative to μ is

$$\int X \mathrm{d}\mu = 1 \cdot \mu(E) + 2 \cdot \mu\left(\overline{E}\right) = 1 \cdot 0.3 + 2 \cdot 0.7 = 1.7\,. \tag{39}$$

In other words, if μ is in the core of $\bar{\mu}$, the expectation value of X relative to μ, $\int X \mathrm{d}\mu$, belongs to $[1.7, 1.8]$ which should be the expectation value of X relative to $\bar{\mu}$.

Although it is intuitive to define the expectation value as the minimum and maximum bounds of the probability intervals associated with each probability measure in the core of $\bar{\mu}$, the core of a general IVPM is hard to be described and computed. Fortunately, the above description is equivalent to the well-known Choquet integral [8, 11, 14, 51], which can be computed step-by-step like a weighted average.

Definition 6 (Classical Expectation Values)**.** Consider a classical sample space Ω with a random variable $X\colon \Omega \to \mathbb{R}$. Since we only consider finite sample spaces, X can always be decomposed into the sum of step functions $\mathbf{1}_E$ defined in Def. 1. For conveniences to define the Choquet integral later, we order these step functions by their coefficients from the smallest to the largest, and express X as follow:

$$X = x_{N^-}^{-}\mathbf{1}_{E_{N^-}^{-}} + \cdots + x_1^{-}\mathbf{1}_{E_1^{-}} + x_1^{+}\mathbf{1}_{E_1^{+}} + \cdots + x_{N^+}^{+}\mathbf{1}_{E_{N^+}^{+}} = \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} x_i^s \mathbf{1}_{E_i^s}\,, \tag{40}$$

where $\left\{E_i^-\right\}_{i=1}^{N^-}$ and $\left\{E_i^+\right\}_{i=1}^{N^+}$ are all disjoint and $x_{N^-}^- < \cdots < x_1^- < 0 \leq x_1^+ < \cdots < x_{N^+}^+$. Then, the expectation values of X relative to real-valued and interval-valued probability measures can be defined as weighted averages as follows.

- Given a classical probability measure $\mu\colon 2^{\Omega} \to [0,1]$, the expectation value of X relative to μ is an average of x_i^s with the weight being the measure of their step functions $\mu\left(E_i^s\right)$, i.e.,

$$\int X \mathrm{d}\mu = \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} x_i^s \mu\left(E_i^s\right) . \tag{41}$$

- Consider a classical IVPM $\bar{\mu}\colon 2^{\Omega} \to \mathscr{I}$. The expectation value of X relative to $\bar{\mu}$ still looks like a weighted average of x_i^s, where s is either $+$ or $-$, but this time the "weights" are not just the measure of their step functions $\mu\left(E_i^s\right)$. Instead, by computing the cumulative sets $E_{i\uparrow}^s = \bigcup_{j\geq i} E_j^s$ and their interval-valued probabilities $\bar{\mu}\left(E_{i\uparrow}^s\right)$, the "weights" are the differences of the both-end of the interval probabilities, $\Delta\mu^t\left(E_{i\uparrow}^s\right) = \mu^t\left(E_{i\uparrow}^s\right) - \mu^t\left(E_{(i+1)\uparrow}^s\right)$, where $\mu^t\left(E_{i\uparrow}^s\right)$ is either the left-end $\mu^{\mathrm{L}}\left(E_{i\uparrow}^s\right)$ or right-end $\mu^{\mathrm{R}}\left(E_{i\uparrow}^s\right)$ of $\bar{\mu}\left(E_{i\uparrow}^s\right)$, that is, $\bar{\mu}\left(E_{i\uparrow}^s\right) = \left[\mu^{\mathrm{L}}\left(E_{i\uparrow}^s\right), \mu^{\mathrm{R}}\left(E_{i\uparrow}^s\right)\right]$. Then, the Choquet integral of X relative to $\bar{\mu}$ is averaging x_i^s with the given "weights,"

$$\int X \mathrm{d}\bar{\mu} = \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} x_i^s \left[\Delta\mu^{\mathrm{L}}\left(E_{i\uparrow}^s\right), \Delta\mu^{\mathrm{R}}\left(E_{i\uparrow}^s\right)\right] . \tag{42}$$

Suppose $\bar{\mu}$ is just a real-valued probability measure μ, i.e., $\bar{\mu}(E) = [\mu(E), \mu(E)]$ for all E. Since the "weights" in Eq. (42) are the difference on the probability of the cumulative sets, they have the same magnitude as the ones in Eq. (41), and the right-hand side of Eq. (42) can be easily simplified

Table 4: This table lists the immediate values to compute the Choquet integral $\int X \mathrm{d}\bar{\mu}$ according to Def. 6.

i	x_i^+	E_i^+	$E_{i\uparrow}^+$	$\mu^{\mathrm{L}}\left(E_{i\uparrow}^+\right)$	$\Delta\mu^{\mathrm{L}}\left(E_{i\uparrow}^+\right)$	$\mu^{\mathrm{R}}\left(E_{i\uparrow}^+\right)$	$\Delta\mu^{\mathrm{R}}\left(E_{i\uparrow}^+\right)$
1	1	E	Ω	1	0.3	1	0.2
2	2	$\overline{E}$	$\overline{E}$	0.7	0.7	0.8	0.8
3			$\emptyset$	0		0	

to the right-hand side of Eq. (41) as follow:

$$\begin{aligned}
\int X \mathrm{d}\bar{\mu} &= \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} x_i^s \left[\Delta\mu\left(E_{i\uparrow}^s\right), \Delta\mu\left(E_{i\uparrow}^s\right)\right] \\
&= \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} x_i^s \left[\mu\left(E_{i\uparrow}^s\right) - \mu\left(E_{(i+1)\uparrow}^s\right), \mu\left(E_{i\uparrow}^s\right) - \mu\left(E_{(i+1)\uparrow}^s\right)\right] \\
&= \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} x_i^s \left[\mu\left(\bigcup_{j\geq i} E_j^s\right) - \mu\left(\bigcup_{j\geq i+1} E_j^s\right), \mu\left(\bigcup_{j\geq i} E_j^s\right) - \mu\left(\bigcup_{j\geq i+1} E_j^s\right)\right] \\
&= \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} x_i^s \left[\sum_{j\geq i} \mu\left(E_j^s\right) - \sum_{j\geq i+1} \mu\left(E_j^s\right), \sum_{j\geq i} \mu\left(E_j^s\right) - \sum_{j\geq i+1} \mu\left(E_j^s\right)\right] \\
&= \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} x_i^s \left[\mu\left(E_i^s\right), \mu\left(E_i^s\right)\right] = \left[\int X \mathrm{d}\mu, \int X \mathrm{d}\mu\right].
\end{aligned} \tag{43}$$

The Choquet integral defined in Eq. (42) is consistent with not only Eq. (41) but also our previous intuition. For example, if $\bar{\mu}$ and X defined in Eqs. (36) and (37) have an intuitive expectation interval $[1.7, 1.8]$, their Choquet integral $\int X \mathrm{d}\bar{\mu}$ should give the same interval $[1.7, 1.8]$. To verify our idea, we first represent X as the sum $1 \cdot \mathbf{1}_E + 2 \cdot \mathbf{1}_{\overline{E}}$. Since the coefficients are all positive, we have $N^- = 0$, $N^+ = 2$, $x_1^+ = 1$, $E_1^+ = E$, $x_2^+ = 2$, and $E_2^+ = \overline{E}$ as listed in Table 4. Step-by-step we compute the cumulative sets $E_{i\uparrow}^+$, their measures $\mu^t\left(E_{i\uparrow}^+\right)$, and their differences $\Delta\mu^t\left(E_{i\uparrow}^s\right)$. These values can then be plugged into Eq. (42) giving the Choquet integral

$$\begin{aligned}\int X \mathrm{d}\bar{\mu} &= x_1^+ \left[\Delta\mu^{\mathrm{L}}\left(E_{1\uparrow}^+\right), \Delta\mu^{\mathrm{R}}\left(E_{1\uparrow}^+\right)\right] + x_2^+ \left[\Delta\mu^{\mathrm{L}}\left(E_{2\uparrow}^+\right), \Delta\mu^{\mathrm{R}}\left(E_{2\uparrow}^+\right)\right] \\ &= 1 \cdot [0.3, 0.2] + 2 \cdot [0.7, 0.8] = [1.7, 1.8] \ .\end{aligned} \tag{44}$$

In general, the Choquet integral is always equal to the minimum and maximum of expectation values relative to probability measures in the core according to the following theorem [11, 14, 44].

Theorem 4. *For every classical IVPM $\bar{\mu}\colon 2^{\Omega} \to \mathscr{I}$ and any random variable $X\colon \Omega \to \mathbb{R}$, we have*

$$\int X \mathrm{d}\bar{\mu} = \left[\min_{\mu \in \mathrm{core}(\bar{\mu})} \int X \mathrm{d}\mu, \max_{\mu \in \mathrm{core}(\bar{\mu})} \int X \mathrm{d}\mu\right] . \tag{45}$$

After we defined the expectation values of random variables relative to classical IVPMs, we can merge this definition with the expectation values to quantum probability measures, Eq. (11), and define the expectation values of observables relative to QIVPMs as follow.

Definition 7 (Expectation Values relative to QIVPMs)**.** Since we want to define the expectation value of observables relative to QIVPMs parallel to classical definition 6, we order the eigenvalues of an observable $\mathbf{O}$ in the spectral decomposition from the smallest to the largest as well

$$\mathbf{O} = \lambda_{N^-}^- P_{N^-}^- + \cdots + \lambda_1^- P_1^- + \lambda_1^+ P_1^+ + \cdots + \lambda_{N^+}^+ P_{N^+}^+ = \sum_{s \in \{+,-\}} \sum_{i=1}^{N^s} \lambda_i^s P_i^s \ , \tag{46}$$

where $\lambda_{N^-}^- < \cdots < \lambda_1^- < 0 \leq \lambda_1^+ < \cdots < \lambda_{N^+}^+$ and each P_i^s is the projector onto the eigenspace of distinct eigenvalues λ_i^s. Consider a QIVPM $\bar{\mu}\colon \mathcal{E} \to \mathscr{I}$ with its left-end and right-end denoted by $\mu^{\mathrm{L}}\colon \mathcal{E} \to [0,1]$ and $\mu^{\mathrm{R}}\colon \mathcal{E} \to [0,1]$, respectively, i.e., $\bar{\mu}(P) = \left[\mu^{\mathrm{L}}(P), \mu^{\mathrm{R}}(P)\right]$. Let s be either $+$ or $-$ in the superscript in the following discussion. We denote $P_{i\uparrow}^s$ as the cumulative projectors on the sub-index i, i.e., $P_{i\uparrow}^s = \sum_{j \geq i} P_j^s$, and $\Delta\mu^t\left(P_{i\uparrow}^s\right)$ as the difference of the measure of these cumulative projectors, i.e., $\Delta\mu^t\left(P_{i\uparrow}^s\right) = \mu^t\left(P_{i\uparrow}^s\right) - \mu^t\left(P_{(i+1)\uparrow}^s\right)$, where $t \in \{\mathrm{L}, \mathrm{R}\}$. Then, the expectation value of $\mathbf{O}$ relative to $\bar{\mu}$ is

$$\langle \mathbf{O} \rangle_{\bar{\mu}} = \sum_{s \in \{+,-\}} \sum_{i=1}^{N^s} \lambda_i^s \left[\Delta\mu^{\mathrm{L}}\left(P_{i\uparrow}^s\right), \Delta\mu^{\mathrm{R}}\left(P_{i\uparrow}^s\right)\right] , \tag{47}$$

which looks like a weighted average of the difference of the measure of these cumulative projectors.

Recall Lemma 1 proved real expectation values are invariant when pulling back observables and probability measures. The following lemma shows interval expectation values are invariant under pullback as well.

Lemma 3. *Consider an observable* $\mathbf{O}$ *diagonalizable by an orthonormal basis* $\Omega = \{|0\rangle, |1\rangle, \ldots, |D-1\rangle\}$ *with* $\varphi\colon 2^{\Omega} \to \mathcal{E}$ *defined by Eq. (4). Given a QIVPM* $\bar{\mu}\colon \mathcal{E} \to \mathscr{I}$*, the expectation value of* $\mathbf{O}$ *relative to* $\bar{\mu}$ *is exactly the expectation value of the pullback of* $\mathbf{O}$ *relative to the pullback of* $\bar{\mu}$*, i.e.,*

$$\langle \mathbf{O} \rangle_{\bar{\mu}} = \int (\varphi^* \mathbf{O}) \, \mathrm{d} (\varphi^* \bar{\mu}) \ . \tag{48}$$

Proof. Consider expressing the observable $\mathbf{O}$ in the spectral decomposition specialized to compute the expectation value relative to a QIVPM, $\sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} \lambda_i^s P_i^s$, where $\lambda_{N^-}^- < \cdots < \lambda_1^- < 0 \leq \lambda_1^+ < \cdots < \lambda_{N^+}^+$ and each P_i^s is the projector onto the eigenspace of distinct eigenvalue λ_i^s. Since $\mathbf{O}$ can be diagonalized by an orthonormal basis $\Omega = \{|0\rangle, |1\rangle, \ldots, |D-1\rangle\}$, each P_i^s can be expressed as the sum of some projectors formed by the elements in Ω, i.e., there is a subset $E_i^s \subseteq \Omega$ such that $P_i^s = \sum_{|j\rangle \in E_i^s} |j\rangle\langle j| = \varphi(E_i^s)$. To compute the expectation values, we want to not only map E_i^s to P_i^s but also map the other ingredients of the expectation values, including the cumulative sets or projectors, their measures, and their differences:

$$\begin{aligned}
P_{i\uparrow}^s &= \sum_{j\geq i} P_j^s = \sum_{j\geq i} \sum_{|k\rangle\in E_j^s} |k\rangle\langle k| = \sum_{|k\rangle \in \bigcup_{j\geq i} E_j^s} |k\rangle\langle k| = \varphi\left(\bigcup_{j\geq i} E_j^s\right) = \varphi\left(E_{i\uparrow}^s\right) \ , \\
\mu^t\left(P_{i\uparrow}^s\right) &= \mu^t\left(\varphi\left(E_{i\uparrow}^s\right)\right) = \left(\varphi^*\mu^t\right)\left(E_{i\uparrow}^s\right) \ , \qquad (49\text{a}) \\
\Delta\mu^t\left(P_{i\uparrow}^s\right) &= \mu^t\left(P_{i\uparrow}^s\right) - \mu^t\left(P_{(i+1)\uparrow}^s\right) = \left(\varphi^*\mu^t\right)\left(E_{i\uparrow}^s\right) - \left(\varphi^*\mu^t\right)\left(E_{(i+1)\uparrow}^s\right) \\
&= \Delta\left(\varphi^*\mu^t\right)\left(E_{i\uparrow}^s\right) \ , \qquad (49\text{b})
\end{aligned}$$

where $t \in \{\mathrm{L}, \mathrm{R}\}$ and $\left[\mu^{\mathrm{L}}(P), \mu^{\mathrm{R}}(P)\right] = \bar{\mu}(P)$ as in Defs. 6 and 7. Notice that the right-most sides of Eq. (49a) and (49b) are actually the pullback of μ^{L} and μ^{R} which can be combined into the pullback of $\bar{\mu}$ as follows:

$$\begin{aligned}
\left[\left(\varphi^*\mu^{\mathrm{L}}\right)(E), \left(\varphi^*\mu^{\mathrm{R}}\right)(E)\right] &= \left[\mu^{\mathrm{L}}(\varphi(E)), \mu^{\mathrm{R}}(\varphi(E))\right] = \bar{\mu}(\varphi(E)) \\
&= (\varphi^*\bar{\mu})(E) = \left[(\varphi^*\bar{\mu})^{\mathrm{L}}(E), (\varphi^*\bar{\mu})^{\mathrm{R}}(E)\right]
\end{aligned} \tag{50}$$

for all $E \subseteq \Omega$.

To compute the expectation value, we also need to know the pullback of $\mathbf{O}$. Since $\mathbf{O}$ can be expressed as the sum of one-dimensional projectors

$$\mathbf{O} = \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} \lambda_i^s P_i^s = \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} \lambda_i^s \sum_{|j\rangle\in E_i^s} |j\rangle\langle j| \ , \tag{51}$$

the pullback of $\mathbf{O}$ should be

$$\varphi^* \mathbf{O} = \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} \lambda_i^s \sum_{|j\rangle\in E_i^s} \mathbf{1}_{\{|j\rangle\}} = \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} \lambda_i^s \mathbf{1}_{E_i^s} \tag{52}$$

according to Def. 1. Then, the expectation value can be computed by spelling the definitions and applying Eqs. (49b), (50), and (52)

$$\begin{aligned}
\langle \mathbf{O} \rangle_{\bar{\mu}} &= \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} \lambda_i^s \left[\Delta\mu^{\mathrm{L}}\left(P_{i\uparrow}^s\right), \Delta\mu^{\mathrm{R}}\left(P_{i\uparrow}^s\right)\right] \\
&= \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} \lambda_i^s \left[\Delta\left(\varphi^*\mu^{\mathrm{L}}\right)\left(E_{i\uparrow}^s\right), \Delta\left(\varphi^*\mu^{\mathrm{R}}\right)\left(E_{i\uparrow}^s\right)\right] \\
&= \sum_{s\in\{+,-\}} \sum_{i=1}^{N^s} \lambda_i^s \left[\Delta\left(\varphi^*\bar{\mu}\right)^{\mathrm{L}}\left(E_{i\uparrow}^s\right), \Delta\left(\varphi^*\bar{\mu}\right)^{\mathrm{R}}\left(E_{i\uparrow}^s\right)\right] \\
&= \int \left(\varphi^*\mathbf{O}\right) \mathrm{d}\left(\varphi^*\bar{\mu}\right) \ .
\end{aligned} \tag{53}$$

□

Since Def. 7 is an extension of both quantum real expectation values and classical interval expectation values, not only the property of quantum real expectation values could be extended to the previous lemma, but the properties of classical interval expectation values, like Eq. (43) and Thm. 4, could be extended to quantum interval expectation values as well. In particular, Eq. (43) can be extended to the following theorem.

Theorem 5. *Given a quantum probability measure $\mu\colon \mathcal{E} \to [0,1]$ satisfying Eqs. (8) and (9), if we define a QIVPM $\bar{\mu}\colon \mathcal{E} \to \mathscr{I}$ by $\bar{\mu}(P) = [\mu(P), \mu(P)]$ for all projectors P, then the expectation value of any observable $\mathbf{O}$ relative to $\bar{\mu}$, $\langle \mathbf{O} \rangle_{\bar{\mu}}$, is just $\left[\langle \mathbf{O} \rangle_\mu, \langle \mathbf{O} \rangle_\mu\right]$.*

Although we can faithfully extend Eq. (43) to Thm. 5, this theorem is weaker than its classical counterpart, Thm. 4 because Lemma 2 hasn't established enough correspondence between classical and quantum cores.

Theorem 6. *Consider an observable* $\mathbf{O}$ *diagonalizable by an orthonormal basis* $\Omega = \{|0\rangle, |1\rangle, \dots, |D-1\rangle\}$ *and a QIVPM* $\bar{\mu}\colon \mathcal{E} \to \mathscr{I}$.

- *Express* $\mathbf{O}$ *in the spectral decomposition specialized to compute the expectation value relative to a QIVPM,* $\sum_{s\in\{+,-\}}\sum_{i=1}^{N^s} \lambda_i^s P_i^s$*, where* $\lambda_{N^-}^- < \cdots < \lambda_1^- < 0 \leq \lambda_1^+ < \cdots < \lambda_{N^+}^+$ *and each* P_i^s *is the projector onto the eigenspace of distinct eigenvalue* λ_i^s*. Given any commuting subspace of events* $\mathcal{E}_{\mathrm{C}}$ *containing all projectors* P_i^s*, we have*

$$\langle \mathbf{O} \rangle_{\bar{\mu}} \subseteq \left[\min_{\rho\in\overline{\mathcal{H}}(\bar{\mu},\mathcal{E}_{\mathrm{C}})} \langle \mathbf{O} \rangle_{\mu_\rho^{\mathrm{B}}}, \max_{\rho\in\overline{\mathcal{H}}(\bar{\mu},\mathcal{E}_{\mathrm{C}})} \langle \mathbf{O} \rangle_{\mu_\rho^{\mathrm{B}}} \right]. \tag{54}$$

- *Let* $\mathcal{E}_\Omega$ *be the set of projectors generated from* Ω*,* $\{\varphi(E) \mid E \subseteq \Omega\}$*, we have*

$$\langle \mathbf{O} \rangle_{\bar{\mu}} = \left[\min_{\rho\in\overline{\mathcal{H}}(\bar{\mu},\mathcal{E}_\Omega)} \langle \mathbf{O} \rangle_{\mu_\rho^{\mathrm{B}}}, \max_{\rho\in\overline{\mathcal{H}}(\bar{\mu},\mathcal{E}_\Omega)} \langle \mathbf{O} \rangle_{\mu_\rho^{\mathrm{B}}} \right]. \tag{55}$$

Proof. By Eq. (48) and Thm. 4, we have

$$\langle \mathbf{O} \rangle_{\bar{\mu}} = \int (\varphi^*\mathbf{O})\, \mathrm{d}(\varphi^*\bar{\mu}) = \left[\min_{\mu\in\mathrm{core}(\varphi^*\bar{\mu})} \int (\varphi^*\mathbf{O})\, \mathrm{d}\mu, \max_{\mu\in\mathrm{core}(\varphi^*\bar{\mu})} \int (\varphi^*\mathbf{O})\, \mathrm{d}\mu \right], \tag{56}$$

where $\varphi\colon 2^\Omega \to \mathcal{E}$ is defined by Eq. (4). Hence, to prove Eq. (54), it is sufficient to prove

$$\left[\min_{\mu\in\mathrm{core}(\varphi^*\bar{\mu})} \int (\varphi^*\mathbf{O})\, \mathrm{d}\mu, \max_{\mu\in\mathrm{core}(\varphi^*\bar{\mu})} \int (\varphi^*\mathbf{O})\, \mathrm{d}\mu \right] \subseteq \left[\min_{\rho\in\overline{\mathcal{H}}(\bar{\mu},\mathcal{E}_{\mathrm{C}})} \langle \mathbf{O} \rangle_{\mu_\rho^{\mathrm{B}}}, \max_{\rho\in\overline{\mathcal{H}}(\bar{\mu},\mathcal{E}_{\mathrm{C}})} \langle \mathbf{O} \rangle_{\mu_\rho^{\mathrm{B}}} \right]. \tag{57}$$

Since $\mathcal{E}_\Omega$ contains all projectors P_i^s no matter how Ω is picked, $\mathcal{E}_\Omega$ is one of the possible choices of $\mathcal{E}_{\mathrm{C}}$. Thus, if Eq. (57) is true, to prove Eq. (55), it is sufficient to verify

$$\left[\min_{\mu\in\mathrm{core}(\varphi^*\bar{\mu})} \int (\varphi^*\mathbf{O})\, \mathrm{d}\mu, \max_{\mu\in\mathrm{core}(\varphi^*\bar{\mu})} \int (\varphi^*\mathbf{O})\, \mathrm{d}\mu \right] \supseteq \left[\min_{\rho\in\overline{\mathcal{H}}(\bar{\mu},\mathcal{E}_\Omega)} \langle \mathbf{O} \rangle_{\mu_\rho^{\mathrm{B}}}, \max_{\rho\in\overline{\mathcal{H}}(\bar{\mu},\mathcal{E}_\Omega)} \langle \mathbf{O} \rangle_{\mu_\rho^{\mathrm{B}}} \right]. \tag{58}$$

Proof of Eq. (57) According to Lemma 2, for all $\mu \in \mathrm{core}(\varphi^*\bar{\mu})$, there is a density matrix $\rho \in \overline{\mathcal{H}}(\bar{\mu}, \mathcal{E}_{\mathrm{C}})$ such that $\mu = \varphi^*\mu_\rho^{\mathrm{B}}$. Together with the properties of extremum and Eq. (13), we have

$$\min_{\mu\in\mathrm{core}(\varphi^*\bar{\mu})}\int(\varphi^*\mathbf{O})\,\mathrm{d}\mu \geq \min_{\rho\in\overline{\mathcal{H}}(\bar{\mu},\mathcal{E}_{\mathrm{C}})}\int(\varphi^*\mathbf{O})\,\mathrm{d}\left(\varphi^*\mu_\rho^{\mathrm{B}}\right) = \min_{\rho\in\overline{\mathcal{H}}(\bar{\mu},\mathcal{E}_{\mathrm{C}})}\langle\mathbf{O}\rangle_{\mu_\rho^{\mathrm{B}}}, \tag{59}$$

$$\max_{\mu\in\mathrm{core}(\varphi^*\bar{\mu})}\int(\varphi^*\mathbf{O})\,\mathrm{d}\mu \leq \max_{\rho\in\overline{\mathcal{H}}(\bar{\mu},\mathcal{E}_{\mathrm{C}})}\int(\varphi^*\mathbf{O})\,\mathrm{d}\left(\varphi^*\mu_\rho^{\mathrm{B}}\right) = \max_{\rho\in\overline{\mathcal{H}}(\bar{\mu},\mathcal{E}_{\mathrm{C}})}\langle\mathbf{O}\rangle_{\mu_\rho^{\mathrm{B}}}, \tag{60}$$

which implies Eq. (57).

Proof of Eq. (58) Given a density matrix $\rho\in\overline{\mathcal{H}}\left(\bar{\mu},\mathcal{E}_\Omega\right)$, it must satisfy $\mu_\rho^{\mathrm{B}}(P)\in\bar{\mu}(P)$ for all $P\in\mathcal{E}_\Omega$. Since $\varphi(E)\in\mathcal{E}_\Omega$ for all $E\subseteq\Omega$, we have

$$\left(\varphi^*\mu_\rho^{\mathrm{B}}\right)(E)=\mu_\rho^{\mathrm{B}}\left(\varphi(E)\right)\in\bar{\mu}\left(\varphi(E)\right)=\left(\varphi^*\bar{\mu}\right)(E)\ , \tag{61}$$

i.e., $\varphi^*\mu_\rho^{\mathrm{B}}\in\mathrm{core}\left(\varphi^*\bar{\mu}\right)$.

Together with the properties of extremum and Eq. (13), the previous paragraph implies

$$\min_{\mu\in\mathrm{core}(\varphi^*\bar{\mu})}\int(\varphi^*\mathbf{O})\,\mathrm{d}\mu \leq \min_{\rho\in\overline{\mathcal{H}}(\bar{\mu},\mathcal{E}_\Omega)}\int(\varphi^*\mathbf{O})\,\mathrm{d}\left(\varphi^*\mu_\rho^{\mathrm{B}}\right) = \min_{\rho\in\overline{\mathcal{H}}(\bar{\mu},\mathcal{E}_\Omega)}\langle\mathbf{O}\rangle_{\mu_\rho^{\mathrm{B}}}, \tag{62}$$

$$\max_{\mu\in\mathrm{core}(\varphi^*\bar{\mu})}\int(\varphi^*\mathbf{O})\,\mathrm{d}\mu \geq \max_{\rho\in\overline{\mathcal{H}}(\bar{\mu},\mathcal{E}_\Omega)}\int(\varphi^*\mathbf{O})\,\mathrm{d}\left(\varphi^*\mu_\rho^{\mathrm{B}}\right) = \max_{\rho\in\overline{\mathcal{H}}(\bar{\mu},\mathcal{E}_\Omega)}\langle\mathbf{O}\rangle_{\mu_\rho^{\mathrm{B}}}, \tag{63}$$

which then implies Eq. (58). □

3 The Kochen-Specker Theorem and Contextuality

Our generalization of quantum probability measures to QIVPMs allows us to strengthen the scope of one of the fundamental theorems of quantum physics: the Kochen-Specker theorem [5, 20, 22, 25, 35, 40, 41]. Our finite-precision extension of that theorem will suggest a resolution to the debate initiated by Meyer and Mermin on the relevance of the Kochen-Specker to experimental, and hence finite-precision, quantum measurements [2–4, 7, 16, 18, 24, 27, 32, 33, 36, 48, 49]. Specifically, the original Kochen-Specker theorem is formulated using a model quantum mechanical system that *always has definite values* [20], i.e., its observables have infinitely precise values at all times. Our interval-valued probability framework will allow us to state and prove, a stronger version of the theorem that holds even if the observables have values that are only definite up to some precision specified by a parameter δ. Our approach provides a quantitative realization of Mermin's intuition [33]:

> . . . although the outcomes deduced from such imperfect measurements will occasionally differ dramatically from those allowed in the ideal case, if the misalignment is very slight, the statistical distribution of outcomes will differ only slightly from the ideal case.

3.1 Finite-Precision Extension of the Kochen-Specker Theorem

The first step in our formalization is to introduce a family of QIVPMs parameterized by an uncertainty δ, which we call δ-deterministic QIVPMs.

Definition 8 (δ-Determinism)**.** A QIVPM $\bar{\mu}\colon \mathcal{E} \to \mathscr{I}$ is δ-deterministic if, for every event $P \in \mathcal{E}$, we have that either $\bar{\mu}(P) \subseteq [0, \delta]$ or $\bar{\mu}(P) \subseteq [1 - \delta, 1]$.

This definition puts no restrictions on the set of intervals itself, only on which intervals are assigned to events. When $\delta = 0$, every event must be assigned a probability either in F or in T, i.e., whether every event happens is completely determined with certainty. As δ gets larger, the QIVPM allows for more indeterminate behavior.

The expectation value of an observable $\mathbf{O}$ in a Hilbert space $\mathcal{H}$ of dimension D relative to a 0-deterministic QIVPM is fully determinate and is equal to one of the eigenvalues λ_i of that observable. To see this, note that given an orthonormal basis $\Omega = \{|0\rangle, |1\rangle, \ldots, |D-1\rangle\}$, a 0-deterministic QIVPM must map exactly one of the projectors $|i\rangle\langle i|$ to T and all others to F. This is because, by Eq. (23b), we have $\bar{\mu}\left(\sum_{j=0}^{D-1} |j\rangle\langle j|\right) = \mathsf{T}$ and by inductively applying Eq. (24), we must have one of the $\bar{\mu}(|i\rangle\langle i|) = \mathsf{T}$ and all others mapped to F. This unique projector mapping to T will be denoted by $|\mathsf{T}\rangle\langle\mathsf{T}|$. Given any state ρ that is consistent with $\bar{\mu}$ on all the projectors in Ω, we have by Eq. (29) that μ_ρ^{B} must also map $|\mathsf{T}\rangle\langle\mathsf{T}|$ to 1 and all other projectors formed by elements in Ω to 0. If an observable has a spectral decomposition along Ω then, by Eq. (11), its expectation value relative to μ_ρ^{B} is the eigenvalue λ_{T} whose projector is $|\mathsf{T}\rangle\langle\mathsf{T}|$. It therefore follows, by Eq. (55), that the expectation value relative to the 0-deterministic $\bar{\mu}$ is fully determinate and lies in the interval $[\lambda_{\mathsf{T}}, \lambda_{\mathsf{T}}]$.

Given a Hilbert space $\mathcal{H}$ of dimension D, the expectation value of the product of a sequence of commuting observables $\{\mathbf{O}_j\}$ relative to the 0-deterministic QIVPM $\bar{\mu}$ is the product of the expectation value of individual $\mathbf{O}_j$. Since $\{\mathbf{O}_j\}$ is a sequence of commuting observables, they can be diagonalized by a common orthonormal basis $\Omega = \{|0\rangle, |1\rangle, \ldots, |D-1\rangle\}$ with spectral decompositions

$$\mathbf{O}_j = \sum_{i=0}^{D-1} \lambda_{j,i} \, |i\rangle\langle i| \ . \tag{64}$$

Because of the orthogonality of Ω, the product of Eqs. (64) can also be simplified as a spectral decomposition $\prod_j \mathbf{O}_j = \sum_{k=0}^{d-1} \prod_j \lambda_{j,i} \, |i\rangle\langle i|$. According to our discussion in the previous paragraph, there is a unique projector such that $\bar{\mu}\,(|\mathsf{T}\rangle\langle\mathsf{T}|) = \mathsf{T}$, and the expectation values relative to $\bar{\mu}$ are their eigenvalues of $|\mathsf{T}\rangle\langle\mathsf{T}|$, i.e., $\langle \mathbf{O}_j \rangle_{\bar{\mu}} = [\lambda_{j,\mathsf{T}}, \lambda_{j,\mathsf{T}}]$ and

$$\left\langle \prod_j \mathbf{O}_j \right\rangle_{\bar{\mu}} = \left[\prod_j \lambda_{j,\mathsf{T}}, \prod_j \lambda_{j,\mathsf{T}} \right] = \prod_j \langle \mathbf{O}_j \rangle_{\bar{\mu}} \tag{65}$$

with the understanding that the product of the singleton sets, $\prod_j [\lambda_{j,\mathsf{T}}, \lambda_{j,\mathsf{T}}]$, is defined to be the product of their elements, $\left[\prod_j \lambda_{j,\mathsf{T}}, \prod_j \lambda_{j,\mathsf{T}}\right]$.

We can now proceed with the main technical result of this section. We first observe that the original Kochen-Specker theorem is a statement regarding the non-existence of a 0-deterministic QIVPM and generalize to a corresponding statement about δ-deterministic QIVPMs.

Theorem 7 (0-Deterministic Variant of the Kochen-Specker Theorem)**.** *Given a Hilbert space $\mathcal{H}$ of dimension $D \geq 3$, there is no 0-deterministic measure $\bar{\mu}$ mapping every event to either* F *or* T.

To explain why this result is equivalent to the original Kochen-Specker theorem and to prove it at the same time, we proceed by assuming a 0-deterministic QIVPM $\bar{\mu}$ and derive the same contradiction as the original Kochen-Specker theorem. Instead of adapting the more complicated proof for $D = 3$, the counterexample presented below uses the simpler proof for a Hilbert space of dimension $D = 4$ and is constructed as follows.

Proof of Thm. 7. We consider a two spin-$\frac{1}{2}$ Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ of dimension $D = 4$. We use the same nine observables $\mathbf{O}_{ij}$ with i and j ranging over $\{0, 1, 2\}$ from the Mermin-Peres "magic square" used to prove the Kochen-Specker theorem [15, 35, 40]:

$\mathbf{O}_{ij}$	$j=0$	$j=1$	$j=2$
$i=0$	$\nVdash \otimes \sigma_z$	$\sigma_z \otimes \nVdash$	$\sigma_z \otimes \sigma_z$
$i=1$	$\sigma_x \otimes \nVdash$	$\nVdash \otimes \sigma_x$	$\sigma_x \otimes \sigma_x$
$i=2$	$\sigma_x \otimes \sigma_z$	$\sigma_z \otimes \sigma_x$	$\sigma_y \otimes \sigma_y$

The observables are constructed using the Pauli matrices $\{\nVdash, \sigma_x, \sigma_y, \sigma_z\}$ whose eigenvalues are all either 1 or -1 [15, 22, 34, 38, 41]. They are arranged such that in each row and column, *except the column* $j=2$, every observable is the product of the other two. In the $j=2$ column, we have instead that $(\sigma_z \otimes \sigma_z)(\sigma_x \otimes \sigma_x) = -\sigma_y \otimes \sigma_y$. Now assume a 0-deterministic QIVPM $\bar{\mu}$; the expectation values of the observables in each row relative to this 0-deterministic QIVPM are fully determinate and must lie in either the interval $[1,1]$ or the interval $[-1,-1]$ depending on which eigenvalue is the one whose associated projector is certain. Since the product of any two observables in a row is equal to the third, the product of any two expectation values in a row is also equal to the third by Eq. (65), and there must be an even number of occurrences of the interval $[-1,-1]$ in each row and hence in the entire table. However, looking at the expectation values of the observables in each column, by the same reason, there must be an even number of occurrences of the interval $[-1,-1]$ in the first two columns and an odd number in the $j=2$ column and hence in the entire table. The contradiction implies the non-existence of the assumed 0-deterministic QIVPM. $\square$

Our framework allows us to generalize the above theorem to state that for small enough δ, it is impossible to have δ-deterministic QIVPMs, which is a stronger statement of contextuality that includes the effects of finite-precision. Every QIVPM must map some events to truly uncertain intervals, not just "almost definite intervals." The proof requires two simple lemmas that we present first.

The first lemma shows a simpler way to prove the convexity condition. Recall that the convexity condition for a QIVPM $\bar{\mu}\colon \mathcal{E} \to \mathscr{I}$ states that for each pair of *commuting* projectors P_0 and P_1 with $P_0 P_1 = P_1 P_0$, the following equation holds:

$$\bar{\mu}(P_0 + P_1 - P_0 P_1) + \bar{\mu}(P_0 P_1) \subseteq \bar{\mu}(P_0) + \bar{\mu}(P_1) \ . \tag{66}$$

Lemma 4. *To verify the convexity condition of a QIVPM $\bar{\mu} \colon \mathcal{E} \to \mathscr{I}$, it is sufficient to check that:*

$$\bar{\mu}\left(P' + P''\right) = \bar{\mu}\left(P'\right) + \bar{\mu}\left(P''\right) \tag{67}$$

for all orthogonal projectors P' and P''.

Proof. The proof follows the outline of the proof of the classical inclusion-exclusion principle. From the commuting projectors P_0 and P_1, we construct the following three orthogonal projectors: $P_0 P_1$, $P_0\left(\mathbb{1} - P_1\right)$, and $\left(\mathbb{1} - P_0\right) P_1$. Then we proceed as follows:

$$\begin{aligned}
&\bar{\mu}\left(P_0 + P_1 - P_0 P_1\right) + \bar{\mu}\left(P_0 P_1\right) \\
&= \bar{\mu}\left(P_0 P_1 + P_0(\mathbb{1} - P_1) + P_1 - P_0 P_1\right) + \bar{\mu}\left(P_0 P_1\right) && \text{(because } P_0 = P_0 P_1 + P_0(\mathbb{1} - P_1)\text{)} \\
&= \bar{\mu}\left(P_0(\mathbb{1} - P_1) + P_1\right) + \bar{\mu}\left(P_0 P_1\right) \\
&= \bar{\mu}\left(P_0(\mathbb{1} - P_1) + P_0 P_1 + (\mathbb{1} - P_0) P_1\right) + \bar{\mu}\left(P_0 P_1\right) && \text{(because } P_1 = P_0 P_1 + (\mathbb{1} - P_0) P_1\text{)} \\
&= \bar{\mu}(P_0(\mathbb{1} - P_1)) + \bar{\mu}(P_0 P_1) + \bar{\mu}((\mathbb{1} - P_0) P_1) + \bar{\mu}\left(P_0 P_1\right) && \text{(using Eq. (67) twice)} \\
&= \bar{\mu}(P_0(\mathbb{1} - P_1) + P_0 P_1) + \bar{\mu}((\mathbb{1} - P_0) P_1 + P_0 P_1) && \text{(using Eq. (67) twice)} \\
&= \bar{\mu}(P_0) + \bar{\mu}(P_1)
\end{aligned}$$

□

The next lemma relates δ-deterministic QIVPMs with $\delta < \frac{1}{3}$ to 0-deterministic QIVPMs.

Lemma 5. *From any δ-deterministic QIVPM $\bar{\mu} \colon \mathcal{E} \to \mathscr{I}$ with $\delta < \frac{1}{3}$, we can construct a 0-deterministic QIVPM $\bar{\mu}^{\mathrm{D}} \colon \mathcal{E} \to \{\mathsf{F}, \mathsf{T}\}$ defined as follows:*

$$\bar{\mu}^{\mathrm{D}}\left(P\right) = \begin{cases} \mathsf{F} & \text{if } \bar{\mu}\left(P\right) \subseteq [0, \delta] \ ; \\ \mathsf{T} & \text{if } \bar{\mu}\left(P\right) \subseteq [1 - \delta, 1] \ . \end{cases} \tag{68}$$

Proof. The most important part of the proof is to verify the convexity condition for $\bar{\mu}^{\mathrm{D}}$. By Lemma 4, it is sufficient to verify the following equation for orthogonal projectors P' and P'',

$$\bar{\mu}^{\mathrm{D}}\left(P' + P''\right) = \bar{\mu}^{\mathrm{D}}\left(P'\right) + \bar{\mu}^{\mathrm{D}}\left(P''\right) \ , \tag{69}$$

for two cases, which we now examine in detail.

When one of $\bar{\mu}^{\mathrm{D}}\left(P'\right)$ and $\bar{\mu}^{\mathrm{D}}\left(P''\right)$ is T, say $\bar{\mu}^{\mathrm{D}}\left(P'\right) = \mathsf{F}$ and $\bar{\mu}^{\mathrm{D}}\left(P''\right) = \mathsf{T}$, we have $\bar{\mu}\left(P'\right) \subseteq [0, \delta]$ and $\bar{\mu}\left(P''\right) \subseteq [1 - \delta, 1]$ which implies $\bar{\mu}\left(P' + P''\right) \subseteq [1 - \delta, 1 + \delta]$. Since $\bar{\mu}\left(P' + P''\right)$ is a subset of $[0, 1]$, $\bar{\mu}\left(P' + P''\right)$ must be

Table 5: Possible probability measures on a Hilbert space of dimension $D = 3$, where $\bar{\mu}'_2$ and $\bar{\mu}_3$ are QIVPMs while $\bar{\mu}_0$, $\bar{\mu}_1$, and $\bar{\mu}_2$ are not. Events are listed in the column labeled by P.

P	$\bar{\mu}_0\,(P)$	$\bar{\mu}_1\,(P)$	$\bar{\mu}_2\,(P)$	$\bar{\mu}'_2\,(P)$	$\bar{\mu}_3\,(P)$
$\nvDash$	**F**	**F**	**F**	**F**	**F**
All one-dimensional projectors	$[0,0]$	$\left[0,\frac{1}{4}\right]$	$\left[0,\frac{1}{3}\right]$	$\left[\frac{1}{3},\frac{1}{3}\right]$	$\left[0,\frac{1}{2}\right]$
All two-dimensional projectors	$[1,1]$	$\left[\frac{3}{4},1\right]$	$\left[\frac{2}{3},1\right]$	$\left[\frac{2}{3},\frac{2}{3}\right]$	$\left[\frac{1}{2},1\right]$
$\nVdash$	**T**	**T**	**T**	**T**	**T**

a subset of $[1-\delta, 1]$, which implies $\bar{\mu}^{\mathrm{D}}\,(P' + P'')$ is also **T**, thus satisfying Eq. (69).

When both $\bar{\mu}^{\mathrm{D}}\,(P')$ and $\bar{\mu}^{\mathrm{D}}\,(P'')$ are **F**, we have both $\bar{\mu}\,(P')$ and $\bar{\mu}\,(P'') \subseteq [0,\delta]$ which implies $\bar{\mu}\,(P' + P'') \subseteq [0, 2\delta]$. Since we assume $\delta < \frac{1}{3}$, the intervals $[0, 2\delta]$ and $[1-\delta, 1]$ are disjoint, which implies $\bar{\mu}\,(P' + P'')$ and $[1-\delta, 1]$ are disjoint. Together with the fact that $\bar{\mu}\,(P' + P'')$ is a subset of either $[0,\delta]$ or $[1-\delta, 1]$, $\bar{\mu}\,(P' + P'')$ must be a subset of $[0,\delta]$, which implies $\bar{\mu}^{\mathrm{D}}(P' + P'') = \mathsf{F}$, and hence also Eq. (69) is again satisfied. □

Theorem 8 (Finite-precision Extension of the Kochen-Specker Theorem)**.** *Given a Hilbert space $\mathcal{H}$ of dimension $D \geq 3$, there is no δ-deterministic QIVPM for $\delta < \frac{1}{3}$.*

Proof by Contradiction. Suppose there is a δ-deterministic QIVPM $\bar{\mu}\colon \mathcal{E} \to \mathscr{I}$. By Lemma 5, we can construct a 0-deterministic QIVPM; however, by Thm. 7, such 0-deterministic QIVPMs do not exist. □

The bound $\delta < \frac{1}{3}$ is tight as it is possible to construct a $\frac{1}{3}$-deterministic QIVPM $\bar{\mu}\colon \mathcal{E} \to \mathscr{I}$. For example, $\bar{\mu}'_2$ defined in Table 5 is a valid $\frac{1}{3}$-deterministic QIVPM. When $\delta \geq \frac{1}{3}$, i.e., when the uncertainty in measurements becomes so large, it becomes possible to map every observable to some (quite inaccurate) probability interval, thus invalidating the Kochen-Specker theorem. We can summarize and illustrate the above arguments using Fig. 1.

As is the case for conventional, infinitely-precise, quantum probability measures, the theorem is only applicable to dimensions $D \geq 3$. Indeed, when the Hilbert space has dimension 2, it is straightforward to construct a 0-deterministic QIVPM as follows. Consider a non-contextual hidden variable model for $D = 2$ (e.g., as proposed by Bell or Kochen-Specker [5, 25]). Such a two-dimensional model always assigns definite values to all observables and hence assigns a *determinate* probability (0 or 1) to each event. This probability measure directly induces a 0-deterministic QIVPM by changing 0 to **F** and 1 to **T**. It follows that every 0-deterministic QIVPM is δ-deterministic.

3.2 Experimental Data and δ-determinism

We have thus quantified one important aspect of uncertainty in quantum mechanics—the effect of the imprecise nature of devices—which is a novel addition to the theory of measurement. Indeed, as Heisenberg emphasized in his famous microscope example [19], the conventional theory of measurement states that it is impossible to precisely measure any property of a system without disturbing it somewhat. Thus, there are fundamental limits to what one can measure and these limits have traditionally been attributed

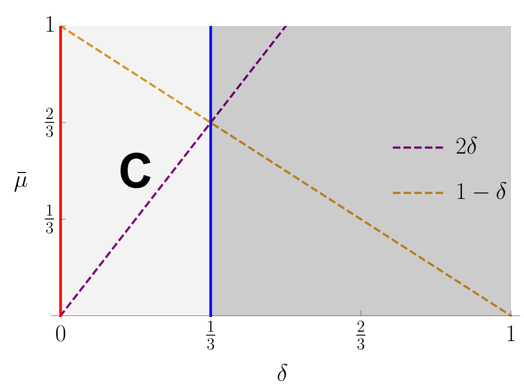

Figure 1: The region to the left of the vertical line at $\delta = \frac{1}{3}$ is where we assume small measurement degradation; in that region, our extension of the KS theorem demonstrates contextuality (**C**). In the region to the right, the degradation of the data is large, and our extension of the KS theorem no longer refutes other explanations for the experimental data.

to complementarity. Our imprecision represents an *additional* source of indeterminacy beyond the inherent probabilistic nature of quantum mechanics.

In an experimental setup, δ is calculated as follows. To determine the probability of any event, we typically repeat an experiment m times and count the number of times we witness the event. This assumes that for each run of the experiment we can determine, using our apparatus, whether the event occurred or not. Assume an event has an ideal mathematical probability of 0, and we repeat the experiment 100 times. In a perfect world, we should be able to refute the event 100 times and calculate that the probability is 0. We might also observe the event 2 times and refute it 98 times and therefore calculate the probability to be 0.02. Note that this situation assumes perfect measurement conditions and remains within the context of conventional (real-valued) probability theory. The question we focus on is what happens if we are only able to refute it 97 times and are *uncertain* 3 times? This is quite common in actual experiments. Mathematically we can model this idea by stating that the probability of the event is in the range $[0, 0.03]$ which says that the probability of the event could be 0, 0.01, 0.02, or 0.03 as each the three uncertain records could either be evidence for the event or against it. We just cannot nail it down given the current experimental results and therefore represent the evidence as a $(\delta =)0.03$-deterministic probability measure. The interesting observation is that the axioms of probability theory (like additivity and convexity) impose enough constraints on the structure of interval-valued quantum probability measures to make them robust in the face of small non-vanishing δ's.

To see this idea in the context of a quantum experiment, consider a three-dimensional Hilbert space with one-dimensional projectors P_ρ, two-dimensional projectors $P_\rho + P_\sigma$, and an experiment that is repeated 12 times. By the Kochen-Specker theorem, it is impossible to build a probability measure that maps every projection to either $0 = \frac{0}{12}$ or $1 = \frac{12}{12}$. That is, the assignment $\bar{\mu}_0$ defined in Table 5 is not a QIVPM.

Now consider what happens if $\frac{1}{4}$ of the data for *every* one-dimensional projector is uncertain. A potential account of this degradation is to assign to each event P the entire range of possibilities $\bar{\mu}_1(P)$ as defined in Table 5. This measure is not a valid QIVPM because it does not satisfy the convexity condition: for any two orthogonal one-dimensional events P_0 and P_1, the convexity condition requires $\bar{\mu}_1(P_0 + P_1) \subseteq \bar{\mu}_1(P_0) + \bar{\mu}_1(P_1)$, but $\bar{\mu}_1(P_0 + P_1) = \left[\frac{3}{4}, 1\right]$ which is not a subset of $\left[0, \frac{1}{2}\right] = \bar{\mu}_1(P_0) + \bar{\mu}_1(P_1)$. Interestingly, it is impossible to find any probability measure that would be consistent with these observations, as the interval $\left[\frac{3}{4}, 1\right]$ is completely disjoint from the interval $\left[0, \frac{1}{2}\right]$ and no amount of shifting of assumptions regarding the precise outcome of the uncertain observations could change that disjointness. However, as shown next, a sharp transition occurs when $\delta = \frac{1}{3}$.

When the proportion of uncertain data reaches $\frac{1}{3}$, the probability measure that assigns to each event the entire range of possibilities is $\bar{\mu}_2$ defined in Table 5. This is also not a valid probability measure by the same argument as above. However, in this case, $\bar{\mu}_2 (P_0 + P_1) = \left[\frac{2}{3}, 1\right]$ and $\left[0, \frac{2}{3}\right] = \bar{\mu}_2 (P_0) + \bar{\mu}_2 (P_1)$ have a *common point*. Hence, by assuming that the uncertain data for one-dimensional projectors always support the associated event, while those for two-dimensional projectors always refute the event, we can find the probability measure $\bar{\mu}_2'$ that can be verified as a valid QIVPM and is consistent with the experimental data.

A similar situation happens when more than $\frac{1}{3}$ of data is uncertain. In particular, if half of the data is uncertain, the probability measure $\bar{\mu}_3$ that assigns to each event the entire range of possibilities is already a QIVPM.

4 The Born Rule and Gleason's Theorem

A conventional quantum probability measure can be easily constructed from a state ρ according to the Born rule [6, 22, 34]. According to Gleason's theorem [12, 40, 41], this state ρ is also the unique state consistent with any possible probability measure.

4.1 Finite-Precision Extension of Gleason's Theorem

In order to re-examine these results in our framework, we first reformulate Gleason's theorem in QIVPMs using infinitely precise uncountable intervals $\mathscr{I}_\infty = \{[x,x] \mid x \in [0,1]\}$:

Theorem 9 ($\mathscr{I}_\infty$ Variant of the Gleason Theorem)**.** *In a Hilbert space $\mathcal{H}$ of dimension $D \geq 3$, given a QIVPM $\bar{\mu} \colon \mathcal{E} \to \mathscr{I}_\infty$, the state ρ consistent with $\bar{\mu}$ on every projector is unique, i.e., there exists a unique state ρ such that $\overline{\mathcal{H}} (\bar{\mu}, \mathcal{E}) = \{\rho\}$.*

Now let us consider relaxing $\mathscr{I}$ to a countable set of finite-width intervals. As the intervals in the image of a QIVPM become less and less sharp, we expect more and more states to be consistent with it. In the limit of minimal sharpness, all states ρ are consistent with the QIVPM

$$\bar{\mu}\left(P\right)=\begin{cases}\mathsf{F} & \text{if } P=\nvdash;\\ \mathsf{T} & \text{if } P=\nVdash;\\ \mathsf{U}=[0,1] & \text{otherwise}\end{cases}\tag{70}$$

mapping nearly all projections to the *unknown* interval U. There is however a subtlety: as we will show in Thm. 10 later, it is possible for an arbitrary assignment of intervals to projectors to be globally inconsistent, but before proving Thm. 10, we need the other two lemmas to simplify the proof of the convexity condition again.

Lemma 6. *Given a Hilbert space $\mathcal{H}$ of dimension 3, to verify $\bar{\mu}\colon\mathcal{E}\to\mathscr{I}$ is a QIVPM, it is sufficient to check Eqs. (23) and*

$$\bar{\mu}\left(P'+P''\right)\subseteq\bar{\mu}\left(P'\right)+\bar{\mu}\left(P''\right)\tag{71}$$

for each pair of orthogonal projectors P' and P''.

Proof. The most important part of the proof is to verify the convexity condition for $\bar{\mu}$. Given a pair of commuting projectors P_0 and P_1 on a three-dimensional Hilbert space, they can be diagonalized by a common orthonormal basis $\Omega=\{\left|0\right\rangle,\left|1\right\rangle,\left|2\right\rangle\}$. Consider the function $\varphi\colon 2^{\Omega}\to\mathcal{E}$ defined in Eq. (4), there are two sets of basis vectors E_0 and $E_1\subseteq\Omega$, such that $\varphi\left(E_0\right)=P_0$ and $\varphi\left(E_1\right)=P_1$. Since E_0 and E_1 are both subsets of a three-element set, their relation has only three possibilities. The first possibility is that one of them is a subset of the other one, $E_0\subseteq E_1$ or $E_1\subseteq E_0$. The second possibility is that they are disjoint, $E_0\cap E_1=\emptyset$. If neither of the previous possibilities is true, i.e., they have some intersections, but no subset relation, then $E_0\cap E_1$, $E_0\backslash E_1$, and $E_1\backslash E_0$ are all non-empty. Together with the fact that Ω has only three elements, they are all singleton sets. These three possibilities are going to be discussed as follows.

- When one of them is a subset of the other one, say $E_0\subseteq E_1$, we have $P_0P_1=\varphi\left(E_0\cap E_1\right)=P_0$ and $P_0+P_1-P_0P_1=P_1$. Thus,

$$\bar{\mu}\left(P_0+P_1-P_0P_1\right)+\bar{\mu}\left(P_0P_1\right)=\bar{\mu}\left(P_1\right)+\bar{\mu}\left(P_0\right).\tag{72}$$

- When $E_0\cap E_1=\emptyset$, we have $P_0P_1=\varphi\left(E_0\cap E_1\right)=\nvdash$ and

$$\bar{\mu}\left(P_0+P_1-P_0P_1\right)+\bar{\mu}\left(P_0P_1\right)=\bar{\mu}\left(P_0+P_1\right)\subseteq\bar{\mu}\left(P_0\right)+\bar{\mu}\left(P_1\right)\tag{73}$$

  by Eq. (71).

- When $E_0 \cap E_1$, $E_0 \backslash E_1$, and $E_1 \backslash E_0$ are all singleton sets, say $E_0 \backslash E_1 = \{|0\rangle\}$, $E_1 \backslash E_0 = \{|1\rangle\}$, and $E_0 \cap E_1 = \{|2\rangle\}$, proving an equivalent condition for the convexity condition, Eq. (24), is easier than proving Eq. (24) directly. Since one minus an interval maps this interval to its mirror image, and reflection preserves the subset relations, the convexity condition holds if and only if

$$\mathsf{T} - \bar{\mu}\left(P_0 P_1\right) + \mathsf{T} - \bar{\mu}\left(P_0 + P_1 - P_0 P_1\right) \subseteq \mathsf{T} - \bar{\mu}\left(P_0\right) + \mathsf{T} - \bar{\mu}\left(P_1\right) \quad (74)$$

which is equivalent to

$$\bar{\mu}\left(\nVdash - P_0 P_1\right) + \bar{\mu}\left(\nVdash - \left(P_0 + P_1 - P_0 P_1\right)\right) \subseteq \bar{\mu}\left(\nVdash - P_0\right) + \bar{\mu}\left(\nVdash - P_1\right) \quad (75)$$

because of Eq. (23c). The last equation holds because we can apply Eq. (71) on the following chain of equations:

$$\begin{aligned} &\bar{\mu}\left(\nVdash - P_0 P_1\right) + \bar{\mu}\left(\nVdash - \left(P_0 + P_1 - P_0 P_1\right)\right) = \bar{\mu}\left(|0\rangle\langle 0| + |1\rangle\langle 1|\right) + \bar{\mu}\left(\nVdash\right) \\ &\subseteq \bar{\mu}\left(|0\rangle\langle 0|\right) + \bar{\mu}\left(|1\rangle\langle 1|\right) = \bar{\mu}\left(\nVdash - P_0\right) + \bar{\mu}\left(\nVdash - P_1\right) . \end{aligned} \quad (76)$$

Since the convexity condition holds for all three possibilities, $\bar{\mu}$ is a QIVPM. □

Lemma 7. *Given a Hilbert space $\mathcal{H}$ of dimension 3, to verify $\bar{\mu}\colon \mathcal{E} \to \mathscr{I}$ is a QIVPM, it is sufficient to check Eqs. (23) and*

$$\bar{\mu}\left(\left|\psi'\right\rangle\left\langle\psi'\right| + \left|\psi''\right\rangle\left\langle\psi''\right|\right) \subseteq \bar{\mu}\left(\left|\psi'\right\rangle\left\langle\psi'\right|\right) + \bar{\mu}\left(\left|\psi''\right\rangle\left\langle\psi''\right|\right) \quad (77)$$

for each pair of orthogonal states $|\psi'\rangle$ and $|\psi''\rangle$.

Proof. Since any projectors can be expressed as the sum of orthogonal one-dimensional projectors, Eq. (77) implies Eq. (71) by induction, and this lemma holds because of Lemma 6. □

After we proved the lemmas, we can state and prove the theorem that some assignment of intervals to projectors can be globally inconsistent.

Theorem 10 (Empty Cores Exist for General QIVPMs)**.** *There exists a Hilbert space $\mathcal{H}$ and a QIVPM $\bar{\mu}\colon \mathcal{E} \to \mathscr{I}$ such that $\overline{\mathcal{H}}\left(\bar{\mu}, \mathcal{E}\right) = \emptyset$.*

Proof. To prove this theorem, we need to construct a QIVPM on some Hilbert space and verify that there are no states that are consistent (see Def. 5) with it on all possible events. Assume a Hilbert space of dimension

Table 6: QIVPM $\bar{\mu}\colon \mathcal{E} \to \mathscr{I}_0$ on a Hilbert space of dimension $D = 3$. Events are listed in the column labeled by P.

P	$\bar{\mu}(P)$
$\not{\mathbb{0}}$, $\|0\rangle\langle 0\|$, $\|+\rangle\langle +\|$, $\|+'\rangle\langle +'\|$	F
$\not{\mathbb{1}}$, $\not{\mathbb{1}} - \|0\rangle\langle 0\|$, $\not{\mathbb{1}} - \|+\rangle\langle +\|$, $\not{\mathbb{1}} - \|+'\rangle\langle +'\|$	T
All other projectors	U

$D = 3$ with orthonormal basis $\{|0\rangle, |1\rangle, |2\rangle\}$, let $|+\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$, $|+'\rangle = \frac{|0\rangle+|2\rangle}{\sqrt{2}}$, and assign

$$\mathscr{I}_0 = \{\mathsf{T}, \mathsf{F}, \mathsf{U}\} \,. \tag{78}$$

Consider the map $\bar{\mu}\colon \mathcal{E} \to \mathscr{I}_0$ defined in Table 6. We want to prove $\bar{\mu}$ is a QIVPM. Since it is easy to verify $\bar{\mu}$ satisfies Eqs. (23), it is sufficient by Lemma 7 to verify

$$\bar{\mu}\left(|\psi'\rangle\langle\psi'| + |\psi''\rangle\langle\psi''|\right) \subseteq \bar{\mu}\left(|\psi'\rangle\langle\psi'|\right) + \bar{\mu}\left(|\psi''\rangle\langle\psi''|\right) \tag{79}$$

for each pair of orthogonal states $|\psi'\rangle$ and $|\psi''\rangle$. Since $|0\rangle$, $|+\rangle$, and $|+'\rangle$ are not orthogonal to each other, at least one of $\bar{\mu}(|\psi'\rangle\langle\psi'|)$ and $\bar{\mu}(|\psi''\rangle\langle\psi''|)$ is unknown U, which implies $\mathsf{U} \subseteq \bar{\mu}(|\psi'\rangle\langle\psi'|) + \bar{\mu}(|\psi''\rangle\langle\psi''|)$. Together with the fact that every interval in $\mathscr{I}_0$ is a subset of U, Eq. (79) holds, and $\bar{\mu}$ is a QIVPM.

Next, we will prove by contradiction that $\overline{\mathcal{H}}(\bar{\mu}, \mathcal{E})$ is the empty set. Suppose there is a state $\rho = \sum_{j=1}^{N} q_j |\phi_j\rangle\langle\phi_j| \in \overline{\mathcal{H}}(\bar{\mu}, \mathcal{E})$, where $\sum_{j=1}^{N} q_j = 1$ and $q_j > 0$. Since we assumed the core $\overline{\mathcal{H}}(\bar{\mu}, \mathcal{E})$ is non-empty, so $\mu_\rho^{\mathrm{B}}(P) \in \bar{\mu}(P)$, and Table 6 tells us that $\bar{\mu}(|0\rangle\langle 0|) = \mathsf{F} = [0, 0]$, we must conclude that $\mu_\rho^{\mathrm{B}}(|0\rangle\langle 0|) = 0 \in [0, 0]$, and similarly for $|+\rangle\langle +|$ and $|+'\rangle\langle +'|$. If this is true, then $\langle 0|\phi_j\rangle = \langle +|\phi_j\rangle = \langle +'|\phi_j\rangle = 0$ for all j, and thus

$$\langle 1|\phi_j\rangle = \sqrt{2}\,\langle +|\phi_j\rangle - \langle 0|\phi_j\rangle = 0\,, \quad \langle 2|\phi_j\rangle = \sqrt{2}\,\langle +'|\phi_j\rangle - \langle 0|\phi_j\rangle = 0\,. \tag{80}$$

The above equations imply $|\phi_j\rangle = |0\rangle\langle 0|\phi_j\rangle + |1\rangle\langle 1|\phi_j\rangle + |2\rangle\langle 2|\phi_j\rangle = 0$, violating the assumption that $|\phi_j\rangle$ is a normalized state, and thus the theorem is proved. □

The fact that a collection of poor measurements on a quantum system cannot reveal the underlying state is not surprising. Under certain conditions, we can however guarantee that the uncertainty in measurements is consistent with *some* non-empty collection of quantum states. Furthermore, we can relate the uncertainty in measurements to the volume of quantum states such that, in the limit of infinitely precise measurements, the volume of states collapses to a single state.

To that end, we introduce the concept of *interval maps*, which we can use to construct a consistent family of QIVPMs. An interval map $f\colon [0,1] \to \mathscr{I}$ maps every real-valued probability $x \in [0,1]$ to a set of intervals $f(x) = [\ell, r]$ containing x, where $[0,1]$ denotes the set of real-valued probabilities (this should not be confused with the interval-valued probability U). We also need a notion of *norm* to quantify the uncertainty in measurements and the distance between (pure or mixed) states. The norm of a collection of intervals $\mathscr{I}$, $\|\mathscr{I}\|$, is defined as the maximum length of intervals in it. The norm of a pure state $\rho = |\psi\rangle\langle\psi|$ is defined as usual by $\|\psi\| = \sqrt{\langle\psi|\psi\rangle}$. For any given Hermitian operator A, we choose the operator norm $\|A\| = \max_{\|\psi\|=1} \|A\,|\psi\rangle\|$, which is also known as the 2-norm or the spectral norm [10, 13, 40, 43]. In fact, for any such matrix, including the density matrix ρ, this norm is the maximum absolute value of its eigenvalues. Then, a finite-precision extension of Gleason's theorem can be stated as follows.

Theorem 11 (Finite-Precision Extension of the Gleason Theorem)**.** *Let $f\colon [0,1] \to \mathscr{I}$ be an interval map and let the composition $f \circ \mu^{\mathrm{B}}_{\rho}$ be a QIVPM, where μ^{B}_{ρ} is the probability measure induced by the Born rule for a given state ρ. If a state ρ' is consistent with $f \circ \mu^{\mathrm{B}}_{\rho}$ on all events, i.e., $\rho' \in \overline{\mathcal{H}}\left(f \circ \mu^{\mathrm{B}}_{\rho}, \mathcal{E}\right)$, then the norm of their difference is bounded by $\|\mathscr{I}\|$, i.e., $\|\rho - \rho'\| \leq \|\mathscr{I}\|$.*

Proof. Given a state ρ' consistent with $f \circ \mu^{\mathrm{B}}_{\rho}$, we have $\mu^{\mathrm{B}}_{\rho'}(|\psi\rangle\langle\psi|) \in f\left(\mu^{\mathrm{B}}_{\rho}(|\psi\rangle\langle\psi|)\right)$ for any one-dimensional projector $P = |\psi\rangle\langle\psi|$. Since the maximum length of the intervals in $\mathscr{I}$ is $\|\mathscr{I}\|$, it is also the upper bound of the difference:

$$\left|\mu^{\mathrm{B}}_{\rho'}(|\psi\rangle\langle\psi|) - \mu^{\mathrm{B}}_{\rho}(|\psi\rangle\langle\psi|)\right| = \left|\left\langle\psi\middle|\rho - \rho'\middle|\psi\right\rangle\right| \leq \|\mathscr{I}\| \,. \tag{81}$$

Since $\rho - \rho'$ is Hermitian, $\max_{\|\psi\|=1} |\langle\psi|\rho - \rho'|\psi\rangle|$ is the maximum absolute value of the eigenvalues of $\rho - \rho'$ [38], and equal to $\|\rho - \rho'\|$ [10, 13]. Hence, $\|\rho - \rho'\| \leq \|\mathscr{I}\|$. □

4.2 Ultramodular Functions

Theorem 11 generalizes Gleason's theorem in the sense that it accounts for a larger class of probability measures that includes the conventional one as a limit. The theorem is however "special" in the sense that it only applies to the particular class of QIVPMs constructed by composing an interval map with a conventional quantum probability measure. QIVPMs constructed in this manner have some peculiar properties that we examine next.

An interval map is called *ultramodular* if it satisfies the following properties.

Definition 9 (Ultramodular Functions)**.** Given a collection of intervals $\mathscr{I}$ including F and T, an interval map $\mathcal{M}\colon\ [0,1] \to \mathscr{I}$ is called ultramodular if

$$\mathcal{M}(0) = \mathsf{F}\,, \qquad \mathcal{M}(1) = \mathsf{T}\,, \qquad \mathcal{M}\left(1-x\right) = \mathsf{T} - \mathcal{M}\left(x\right)\,, \tag{82}$$

and for any three numbers x_0, x_1, and $x_2 \in [0,1]$ such that $y = x_0+x_1+x_2 \in [0,1]$, we have

$$\mathcal{M}\left(y\right) + \mathcal{M}\left(x_2\right) \subseteq \mathcal{M}\left(x_0+x_2\right) + \mathcal{M}\left(x_1+x_2\right)\,. \tag{83}$$

The first three constraints, Eqs. (82), are the direct counterpart of the corresponding QIVPM constraints, Eqs. (23); the last condition, Eq. (83), is the direct counterpart of the convexity conditions, Eqs. (20) and (24) [8, 31, 37, 47]. Therefore, these conditions guarantee that for any conventional quantum probability measure μ, the composition $\mathcal{M} \circ \mu$ defines a valid QIVPM. Conversely, if for every quantum probability measure μ, it is the case that $f \circ \mu$ is a QIVPM, then the interval map f is an ultramodular function. Formally, we have the following result:

Theorem 12 (Equivalence of Ultramodular Functions and IVPMs)**.** *The following three statements are equivalent:*

1. *A function $\mathcal{M}\colon\ [0,1] \to \mathscr{I}$ is ultramodular.*
2. *The composite function $\mathcal{M} \circ \mu\colon 2^{\Omega} \to \mathscr{I}$ is a classical IVPM for all classical probability measures $\mu\colon 2^{\Omega} \to [0,1]$.*
3. *The composite function $\mathcal{M} \circ \mu\colon \mathcal{E} \to \mathscr{I}$ is a QIVPM for all quantum probability measures $\mu\colon \mathcal{E} \to [0,1]$.*

Proof. Statement 1 implies 2 and 3 as we have outlined above. Conversely, for the quantum case, we want to show that if $\mathcal{M}$ is not ultramodular, then for some quantum probability measure μ, the composite $\mathcal{M} \circ \mu$ might not be a QIVPM. Suppose there are three particular numbers x_0, x_1, and $x_2 \in [0,1]$ such that $y = x_0 + x_1 + x_2 \in [0,1]$, but they don't satisfy Eq. (83). Consider the state:

$$\rho = x_0 \left|0\right\rangle\!\left\langle 0\right| + x_1 \left|1\right\rangle\!\left\langle 1\right| + x_2 \left|2\right\rangle\!\left\langle 2\right| + (1-y) \left|3\right\rangle\!\left\langle 3\right| . \tag{84}$$

The induced map $\mathcal{M} \circ \mu_\rho^{\mathrm{B}}$ constructed using the Born rule and blurred by $\mathcal{M}$ fails to satisfy Eq. (24) when $P_0 = \left|0\right\rangle\!\left\langle 0\right| + \left|2\right\rangle\!\left\langle 2\right|$ and $P_1 = \left|1\right\rangle\!\left\langle 1\right| + \left|2\right\rangle\!\left\langle 2\right|$. In other words, this induced map fails to be a QIVPM.

For the classical case, if $\mathcal{M}$ is not ultramodular, we also want to find a classical probability measure $\mu\colon 2^\Omega \to [0,1]$ such that $\mathcal{M} \circ \mu$ is not a classical IVPM. Consider an orthonormal basis $\Omega = \{\left|0\right\rangle, \left|1\right\rangle, \ldots, \left|D-1\right\rangle\}$ and $\varphi\colon 2^\Omega \to \mathcal{E}$ defined by Eq. (4). Notice that the pullback of our previous quantum probability measure μ_ρ^{B}, $\varphi^* \mu_\rho^{\mathrm{B}}$, is a classical probability measure. If we pick μ as $\varphi^* \mu_\rho^{\mathrm{B}}$, then the induced map $\mathcal{M} \circ \mu$ fails to be a classical IVPM for the same reason as in the quantum case. □

In other words, the essential properties of QIVPMs constructed using interval maps can be gleaned from the properties of ultramodular functions. The following is the most interesting property in our setting.

Theorem 13 (Range of Ultramodular Functions)**.** *For any ultramodular function $\mathcal{M}\colon [0,1] \to \mathscr{I}$, either $\mathscr{I} = \mathscr{I}_0$ as defined in Eq. (78) or $\mathscr{I}$ contains uncountably many intervals.*

Proof. Since $\mathcal{M}$ maps to intervals, we can decompose it into two functions: its left-end and right-end, where $\left[\mathcal{M}^L(x), \mathcal{M}^R(x)\right] = \mathcal{M}(x)$. By Eq. (83), the left-end function $\mathcal{M}^L\colon [0,1] \to [0,1]$ is Wright-convex [39, 43, 54], i.e.,

$$\mathcal{M}^L(y) + \mathcal{M}^L(x_2) \geq \mathcal{M}^L(x_0 + x_2) + \mathcal{M}^L(x_1 + x_2) \tag{85}$$

for three numbers x_0, x_1, and $x_2 \in [0,1]$ with $y = x_0 + x_1 + x_2 \in [0,1]$. Together with the fact that $\mathcal{M}^L$ maps to a bounded interval $[0,1]$, the left-end function $\mathcal{M}^L$ must be continuous on the unit open interval $(0,1)$ [31]. Therefore, either $\mathcal{M}$ maps every number in $(0,1)$ to the same interval, or the number of intervals to which $\mathcal{M}$ maps must be uncountable. □

To summarize, a conventional quantum probability measure has an uncountable range $[0,1]$. A QIVPM constructed by blurring such a conventional quantum probability measure must also have an uncountable range of intervals. Of course, any particular QIVPM, or any particular experiment, will use a fixed collection of intervals appropriate for the resources and precision of the particular experiment.

5 Summary

Conventional quantum theory is based on the continuum of complex numbers, but we cannot distinguish two arbitrary complex numbers without unbounded resources. To explore alternative versions of quantum theory incorporating our limitation of distinguishability, two types of discrete quantum theories were described: *quantum theories and computing over finite fields* and *quantum interval-valued probability measures (QIVPMs)*. Examining the physical and computational consequences of such frameworks could yield new insights into not only the subtle properties of conventional quantum theory but also the power and capacity of quantum computing.

The theories over finite fields started with unrestricted discrete fields (Chapter **??**) and then advanced to a more reasonable framework based on complexifiable discrete fields (Secs. **??** to **??**), but both of them lack a notion of probability and support unnaturally efficient deterministic quantum algorithms. A still more plausible discrete theory with cardinal probabilities was proposed (Secs. **??** and **??**), where conventional quantum theory and computing emerge in a local sense, but lacking arithmetic operations among cardinal probabilities still posed difficulty to define expectation values. Since the axiomatic approach looked unlikely to provide sensible real-valued probability measures over finite fields (Sec. **??**), we shifted our attention to directly embed our limitation of distinguishability into the theory to define QIVPM.

As a natural extension of both conventional quantum probability measures and classical interval-valued probability measures (IVPMs) illustrated in Fig. 2, QIVPMs inherit definitions and properties from the both frameworks. While the expectation values with respect to QIVPMs can be pulled back to the classical ones and consistent with those with respect to quantum probability measures in the infinitely precise limit (Sec. 2), foundational concepts in quantum mechanics, such as the Kochen-Specker and Gleason

$$\begin{tikzcd}
\text{Classical Probability Measure} \arrow[r, "\text{glue events}"] \arrow[d, "\text{blur probability}"'] & \text{Quantum Probability Measure} \arrow[d, "\text{blur probability}"] \\
\text{Classical IVPM} \arrow[r, "\text{glue events}"'] & \text{QIVPM}
\end{tikzcd}$$

Figure 2: QIVPMs inherit from both quantum probability measures and classical IVPMs.

theorems, extended to QIVPMs in subtle ways. By carefully specifying experimental uncertainties, we established rigorous bounds on the validity of the Kochen-Specker theorem (Sec. 3). While there is a QIVPM not consistent with Gleason's unique state ρ on all projectors, we constructed a class of QIVPMs for which the original Gleason theorem could be recovered asymptotically (Sec. 4).

In the following further discussion, we will briefly explain why we only recovered Gleason's theorem on a class of QIVPMs, the possibility to further build a computational model over QIVPMs, and the possibility combining both approaches to consider QIVPMs over finite fields.

6 Gleason's Theorem for General QIVPMs

As we discussed in Sec. 4, Thm. 11 only applies to the QIVPMs constructed by composing an interval map with a conventional quantum probability measure, and the states consistent with the composite QIVPM collapse to a single state as the maximum length of intervals in $\mathscr{I}$, $\|\mathscr{I}\|$, shrinks to 0. In contrast, the globally inconsistent QIVPM defined in Table 6 has the least sharp range $\mathscr{I}_0$ with $\|\mathscr{I}_0\| = 1$. This suggests a possibility that shrinking the length $\|\mathscr{I}\|$ might help to regularize general QIVPMs, and it is natural to ask whether there is a short enough length ε such that QIVPMs mapping to intervals not longer than ε always have non-empty cores.

Question 1. Given a Hilbert space $\mathcal{H}$ of dimension $D \geq 3$, is there an $\varepsilon > 0$ such that for all QIVPM $\bar{\mu} \colon \mathcal{E} \to \mathscr{I}$ satisfying $\|\mathscr{I}\| \leq \varepsilon$, $\bar{\mu}$ must have a non-empty core, i.e., $\overline{\mathcal{H}}(\bar{\mu}, \mathcal{E}) \neq \emptyset$?

To better understand this question, consider the $D = 3$ situation, where any two-dimensional projectors can be expressed as the complement of a one-dimensional projector, and by Eq. (23c) so does their interval-valued probabilities, i.e.,

$$\bar{\mu}\left(\mathbb{1} - |\psi\rangle\langle\psi|\right) = \mathbf{T} - \bar{\mu}\left(|\psi\rangle\langle\psi|\right) \, . \tag{86}$$

Hence, a QIVPM is completely determined by its values on the one-dimensional projectors which are one-to-one corresponding to the irreducible states, and these irreducible states are encoded in the complex projective space $\mathbb{C}\mathbf{P}^2$ as we discussed in Sec. **??**. Therefore, to study a QIVPM $\bar{\mu}$, we just need to study a pair of functions $f^L \colon \mathbb{C}\mathbf{P}^2 \to [0, 1]$ and $f^R \colon \mathbb{C}\mathbf{P}^2 \to [0, 1]$ defined by $\left[f^L\left(|\psi\rangle\right), f^R\left(|\psi\rangle\right)\right] = \bar{\mu}\left(|\psi\rangle\langle\psi|\right)$ for any irreducible state $|\psi\rangle \in \mathbb{C}\mathbf{P}^2$.

According to Lemma 7, $\bar{\mu}\colon \mathcal{E} \to \mathscr{I}$ is a QIVPM if and only if $\bar{\mu}$ satisfies Eqs. (23) and

$$\bar{\mu}\left(\mathbb{1} - |\psi_0\rangle\langle\psi_0|\right) \subseteq \bar{\mu}\left(|\psi_1\rangle\langle\psi_1|\right) + \bar{\mu}\left(|\psi_2\rangle\langle\psi_2|\right) \tag{87}$$

for all orthonormal basis $\{|\psi_i\rangle\}_{i=0}^{2}$ because $|\psi_1\rangle\langle\psi_1|+|\psi_2\rangle\langle\psi_2| = \mathbb{1}-|\psi_0\rangle\langle\psi_0|$. By applying Eq. (86) on the left-hand side, Eq. (87) is equivalent to the following interval-inclusion

$$\begin{aligned} & \mathsf{T} - \left[f^L\left(|\psi_0\rangle\right), f^R\left(|\psi_0\rangle\right)\right] \subseteq \left[f^L\left(|\psi_1\rangle\right), f^R\left(|\psi_1\rangle\right)\right] + \left[f^L\left(|\psi_2\rangle\right), f^R\left(|\psi_2\rangle\right)\right] \quad (88) \\ \Leftrightarrow \quad & \left[1 - f^R\left(|\psi_0\rangle\right), 1 - f^L\left(|\psi_0\rangle\right)\right] \subseteq \left[f^L\left(|\psi_1\rangle\right) + f^L\left(|\psi_2\rangle\right), f^R\left(|\psi_1\rangle\right) + f^R\left(|\psi_2\rangle\right)\right] . \end{aligned}$$

This interval-inclusion can be rephrased as a long inequality

$$f^L\left(|\psi_1\rangle\right)+f^L\left(|\psi_2\rangle\right) \leq 1-f^R\left(|\psi_0\rangle\right) \leq 1-f^L\left(|\psi_0\rangle\right) \leq f^R\left(|\psi_1\rangle\right)+f^R\left(|\psi_2\rangle\right) . \tag{89}$$

Since $\|\mathscr{I}\| \leq \varepsilon$, the length of every interval $\left[f^L\left(|\psi\rangle\right), f^R\left(|\psi\rangle\right)\right]$ is bounded by ε as well, which implies the largest term in the previous inequality $f^R\left(|\psi_1\rangle\right) + f^R\left(|\psi_2\rangle\right)$ is bounded by $f^L\left(|\psi_1\rangle\right) + f^L\left(|\psi_2\rangle\right) + 2\varepsilon$. In other words, the left-end function f^L satisfies the following inequalities

$$\begin{aligned} & f^L\left(|\psi_1\rangle\right) + f^L\left(|\psi_2\rangle\right) \leq 1 - f^L\left(|\psi_0\rangle\right) \leq f^L\left(|\psi_1\rangle\right) + f^L\left(|\psi_2\rangle\right) + 2\varepsilon \quad (90a) \\ \Leftrightarrow \quad & 1 - 2\varepsilon \leq \sum_{i=0}^{2} f^L\left(|\psi_i\rangle\right) \leq 1 . \quad (90b) \end{aligned}$$

In this language, Gleason's theorem basically states that when $\varepsilon = 0$, given any function $f^L\colon \mathbb{C}\mathbf{P}^2 \to [0,1]$ satisfying Eq. (90b), there exists a unique mixed state ρ such that

$$f^L\left(|\psi\rangle\right) = \langle\psi|\rho|\psi\rangle \tag{91}$$

for any state $|\psi\rangle \in \mathbb{C}\mathbf{P}^2$. Our Question 1 then ask how f^L would look like with a positive ε.

With different settings, whether there is an approximate version of Gleason's theorem was asked by Sam Sanders in *constructivenews* on 2013 [45], and there is no clear answer for his question. To have an idea of how hard this question could be, recall in Sec. 1 we state that a quantum probability space is glued by a family of classical probability spaces. This is like the situation that a manifold is glued by many local coordinates. When each small piece has *exactly* the same and positive curvature, the Killing-Hopf theorem asserts this manifold is a sphere, but little can we say even if the curvature has a small deviation from constant. A similar situation might happen when approximating Gleason's theorem, but this time the whole space is glued by "local" classical probability space defined by each orthonormal basis $\{|\psi_i\rangle\}_{i=0}^{2}$. When the sum of f^L, $\sum_{i=0}^{2} f^L(|\psi_i\rangle)$, is *exactly* the same and equal to 1, Gleason's theorem asserts that f^L can be expressed as Eq. (91). However, when each local classical probability space becomes imprecise, a general f^L might be as wild as we can imagine, and we might need to know a bit more, like its QIVPM is a composite function, to deduce its global property.

7 And Beyond...

After we build the quantum interval-valued probability model, we might want to know how powerful a quantum computer could be based on this model. Since the conventional quantum circuit model manipulates the probability amplitudes instead of the measured probabilities, either a quantum computing model above QIVPMs needs to simultaneously manipulate all states in the core of a QIVPM, or we need to find a way to manipulate a QIVPM directly. However, both strategies are not straightforward. On one hand, as we proved in Thm. 10, a QIVPM might have an empty core which cannot be evolved over time. On the other hand, if we want to manipulate and compute QIVPMs directly for multi-qubit algorithms, we need to glue the QIVPM for each qubit or subsystem together to get the QIVPM of the whole system, and this is not straightforward either.

A successful interval-valued theory might be further extended over finite fields based on the following definition.

Definition 10 (Quantum Interval-valued Probability Measures over Finite Fields)**.** Consider a vector space $\mathcal{H}$ of dimension D over the complexified field $\mathbb{F}_{p^2}$, its set of events $\mathcal{E}_{p^2}$ as defined in Def. **??**, and a collection of intervals $\mathscr{I}$. A quantum interval-valued probability measure over finite field $\bar{\mu}\colon \mathcal{E}_{p^2} \to \mathscr{I}$ assigns an interval to each event P subject to $\bar{\mu}(\nvdash) = \mathsf{F}$, $\bar{\mu}(\nVdash) = \mathsf{T}$, $\bar{\mu}(\nVdash - P) = \mathsf{T} - \bar{\mu}(P)$, and satisfying for each pair of *commuting* events P_0 and P_1 with $P_0P_1 = P_1P_0$,

$$\bar{\mu}(P_0 + P_1 - P_0P_1) + \bar{\mu}(P_0P_1) \subseteq \bar{\mu}(P_0) + \bar{\mu}(P_1)\ . \tag{92}$$

Understanding the properties of these probability measures and whether we could define a "sensible" Born rule upon them combines two approaches for dealing with the continuous quantities used in the conventional quantum theory and will be the next natural extension for our discrete quantum theories and computing.

References

[1] Samson Abramsky. “Big toy models: Representing physical systems as Chu spaces”. In: *Synthese* 186.3 (2012), pp. 697–718.

[2] D. M. Appleby. “Existential contextuality and the models of Meyer, Kent, and Clifton”. In: *Phys. Rev. A* 65 (2 Jan. 2002), p. 022105.

[3] D. M. Appleby. “The Bell–Kochen–Specker theorem”. In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 36.1 (Mar. 2005), pp. 1–28.

[4] Jonathan Barrett and Adrian Kent. “Non-contextuality, finite precision measurement and the Kochen–Specker theorem”. In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 35.2 (2004), pp. 151–176.

[5] John S. Bell. “On the Problem of Hidden Variables in Quantum Mechanics”. In: *Rev. Mod. Phys.* 38.3 (3 July 1966), pp. 447–452.

[6] Max Born. “On the Quantum Mechanics of Collisions”. English. In: *Quantum Theory and Measurement.* Trans. by John Archibald Wheeler and Wojciech Hubert Zurek. Princeton University Press, 1983, pp. 52–55.

[7] Adán Cabello. “Finite-precision measurement does not nullify the Kochen-Specker theorem”. In: *Phys. Rev. A* 65 (5 Apr. 2002), p. 052101.

[8] Gustave Choquet. “Theory of capacities”. In: *Annales de l'institut Fourier* 5 (1954), pp. 131–295.

[9] A. P. Dempster. “Upper and Lower Probabilities Induced by a Multivalued Mapping”. In: *Ann. Math. Statist.* 38.2 (Apr. 1967), pp. 325–339.

[10] Simon Foucart. “Lecture 6: Matrix Norms and Spectral Radii”. lecture notes for the course NSTP187 at Drexel University, Philadelphia, PA, Fall 2012. 2012.

[11] Itzhak Gilboa and David Schmeidler. “Additive representations of non-additive measures and the Choquet integral.” In: *Annals of Operations Research* 52.1–4 (1994), pp. 43–65.

[12] Andrew Gleason. “Measures on the Closed Subspaces of a Hilbert Space”. In: *Indiana Univ. Math. J.* 6 (4 1957), pp. 885–893.

[13] Gene H. Golub and Charles F. Van Loan. *Matrix Computations.* 3rd ed. Johns Hopkins Studies in Mathematical Sciences. Johns Hopkins University Press, 1996.

[14] Michel Grabisch. *Set functions, games and capacities in decision making.* Theory and Decision Library C 46. Springer International Publishing, 2016.

[15] Robert B. Griffiths. *Consistent quantum theory.* Cambridge University Press, 2003.

[16] Otfried Gühne, Matthias Kleinmann, Adán Cabello, Jan-Åke Larsson, Gerhard Kirchmair, Florian Zähringer, Rene Gerritsma, and Christian F Roos. "Compatibility and noncontextuality for sequential measurements". In: *Phys. Rev. A* 81.2 (2 Feb. 2010), p. 022121.

[17] Jan Hamhalter. *Quantum Measure Theory.* Vol. 134. The Fundamental Theories of Physics. Springer Science & Business Media, Oct. 31, 2003. 420 pp.

[18] Hans Havlicek, Guenther Krenn, Johann Summhammer, and Karl Svozil. *On coloring the rational quantum sphere.* Nov. 9, 1999. arXiv: `quant-ph/9911040v1 [quant-ph]`. Published as Hans Havlicek, Günther Krenn, Johann Summhammer, and Karl Svozil. "Colouring the rational quantum sphere and the Kochen-Specker theorem". In: *J. Phys. A: Math. Gen.* 34.14 (2001), p. 3071.

[19] W. Heisenberg. *The actual content of quantum theoretical kinematics and mechanics.* English. Tech. rep. NASA-TM-77379. Washington, D.C.: National Aeronautics and Space Administration, Dec. 1, 1983. Trans. from "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik". In: *Zeitschrift für Physik* 43.3–4 (Mar. 1927), pp. 172–198.

[20] Carsten Held. "The Kochen-Specker Theorem". In: *The Stanford Encyclopedia of Philosophy.* Ed. by Edward N. Zalta. Fall 2016. Metaphysics Research Lab, Stanford University, 2016.

[21] Peter J. Huber and Elvezio M. Ronchetti. *Robust Statistics.* English. 2nd ed. Wiley Series in Probability and Statistics. John Wiley & Sons Inc., Mar. 6, 2009. 354 pp.

[22] Gregg Jaeger. *Quantum Information.* Springer New York, Apr. 3, 2007.

[23] Kenneth David Jamison and Weldon A. Lodwick. *Interval-Valued Probability Measures.* Tech. rep. 213. Center for Computational Mathematics, University of Colorado Denver, 2004.

[24] Adrian Kent. "Noncontextual hidden variables and physical measurements". In: *Phys. Rev. Lett.* 83.19 (1999), p. 3755.

[25] S. Kochen and E. Specker. "The Problem of Hidden Variables in Quantum Mechanics". In: *Indiana Univ. Math. J.* 17 (1 1968), pp. 59–87.

[26] Andrei Nikolaevich Kolmogorov. *Foundations of the Theory of Probability*. English. Trans. from the German by Nathan Morrison. New York: Chelsea Publishing Company, 1950.

[27] J.-Å. Larsson. "A Kochen-Specker inequality". In: *Europhys. Lett.* 58.6 (2002), p. 799.

[28] Hans Maassen. "Quantum probability and quantum information theory". In: *Quantum information, computation and cryptography*. Springer, 2010, pp. 65–108.

[29] George W. Mackey. "Quantum Mechanics and Hilbert Space". In: *The American Mathematical Monthly* 64.8 (1957), pp. 45–57.

[30] Massimo Marinacci. "Limit Laws for Non-additive Probabilities and Their Frequentist Interpretation". In: *Journal of Economic Theory* 84.2 (Feb. 1999), pp. 145–195.

[31] Massimo Marinacci and Luigi Montrucchio. "Ultramodular Functions". In: *Mathematics of Operations Research* 30.2 (2005), pp. 311–332.

[32] Michael D. Mazurek, Matthew F. Pusey, Ravi Kunjwal, Kevin J. Resch, and Robert W. Spekkens. "An experimental test of noncontextuality without unphysical idealizations". In: *Nat. Commun.* 7 (June 2016), ncomms11780.

[33] N. David Mermin. *A Kochen-Specker Theorem for Imprecisely Specified Measurement*. Dec. 16, 1999. arXiv: `quant-ph/9912081v1 [quant-ph]`.

[34] N. David Mermin. *Quantum Computer Science*. Cambridge University Press, 2007.

[35] N. David Mermin. "Simple unified form for the major no-hidden-variables theorems". In: *Phys. Rev. Lett.* 65.27 (27 Dec. 1990), pp. 3373–3376.

[36] David Meyer. "Finite Precision Measurement Nullifies the Kochen-Specker Theorem". In: *Phys. Rev. Lett.* 83 (19 Nov. 1999), pp. 3751–3754.

[37] Man-Chung() Ng(), Chi-Ping() Mo(), and Yeong-Nan() Yeh(). "On the Cores of Scalar Measure Games". In: *Taiwanese Journal of Mathematics* 1.2 (1997), pp. 171–180.

[38] Michael A. Nielsen and Isaac L. Chuang. *Quantum computation and quantum information*. New York, NY, USA: Cambridge University Press, 2000.

[39] Josip E. Pečarić, Frank Proschan, and Yung Liang Tong. *Convex Functions, Partial Orderings and Statistical Applications*. Vol. 187. Mathematics in Science and Engineering. Academic Press Inc, May 11, 1992. 467 pp.

[40] Asher Peres. *Quantum Theory: Concepts and Methods*. Fundamental Theories of Physics. Springer, Sept. 30, 1995. 464 pp.

[41] Michael Redhead. *Incompleteness, Nonlocality, and Realism: A Prolegomenon to the Philosophy of Quantum Mechanics*. Oxford University Press, 1987.

[42] Fred Richman and Douglas Bridges. "A Constructive Proof of Gleason's Theorem". In: *Journal of Functional Analysis* 162.2 (1999), pp. 287–312.

[43] A. Wayne Roberts and Dale E. Varberg. *Convex Functions*. Pure & Applied Mathematics. Academic Press Inc, 1973.

[44] Joachim Rosenmüller. "On core and value". In: *Operations Research-Verfahren. Methods of operations research* 9 (1971), pp. 84–104.

[45] Sam Sanders. *Gleason's Theorem*. constructivenews - Google Groups. Aug. 28, 2013. URL: `https://groups.google.com/forum/#!topic/constructivenews/R7oZqVuVrDM` (visited on 01/29/2017).

[46] Glenn Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Apr. 11, 1976. 314 pp.

[47] Lloyd S. Shapley. "Cores of convex games". In: *International Journal of Game Theory* 1.1 (1971), pp. 11–26.

[48] Christoph Simon, Časlav Brukner, and Anton Zeilinger. "Hidden-Variable Theorems for Real Experiments". In: *Phys. Rev. Lett.* 86 (20 May 2001), pp. 4427–4430.

[49] Robert W. Spekkens. "Contextuality for preparations, transformations, and unsharp measurements". In: *Phys. Rev. A* 71.5 (2005), p. 052108.

[50] Arthur R. Swift and Ron Wright. "Generalized Stern-Gerlach experiments and the observability of arbitrary spin operators". In: *Journal of Mathematical Physics* 21.1 (1980), pp. 77–82.

[51] Giuseppe Vitali. "Sulla definizione di integrale delle funzioni di una variabile". In: *Annali di Matematica Pura ed Applicata* 2.1 (Dec. 1925), pp. 111–121.

[52] Kurt Weichselberger. "The theory of interval-probability as a unifying concept for uncertainty". In: *Int. J. Approximate Reasoning* 24.2 (May 2000), pp. 149–170.

[53] Wikipedia contributors. *Pullback — Wikipedia, The Free Encyclopedia.* `https://en.wikipedia.org/w/index.php?title=Pullback&oldid=834071311`. [Online; accessed 4-August-2018]. 2018.

[54] E. M. Wright. "An Inequality for Convex Functions". In: *Amer. Math. Monthly* 61.9 (1954), pp. 620–622.