Interval Probability and Measurement for Fuzzy Quantum Theories

November 17, 2016

1 Introduction

Fuzzy quantum mechanics:

- http://cds.cern.ch/record/518511/files/0107054.pdf
- http://link.springer.com/chapter/10.1007%2F978-3-642-35644-5_18#page-1
- http://link.springer.com/chapter/10.1007%2F978-3-540-93802-6_20#page-1
- http://www.du.edu/nsm/departments/mathematics/media/documents/preprints/m0412.pdf
- http://www.space-lab.ru/files/pages/PIRT_VII-XII/pages/text/PIRT_X/Bobola.pdf
- http://www.vub.ac.be/CLEA/aerts/publications/1993LiptovskyJan.pdf

Pseudo-randomness:

- https://people.csail.mit.edu/silvio/Selected%20Scientific%20Papers/Pseudo%20Randomness/How_To_Generate_Cryptographically_Strong_Sequences_Of_Pseudo-Random_Bits.pdf: "the randomness of an event is relative to a specific model of computation with a specified amount of computing resources."
- Another version https://pdfs.semanticscholar.org/3e9c/5f6f48d9ef426655dc799e9b287d754e86c1.pdf

1.1 Plan

In the remainder of the paper, we consider variations of quantum probability spaces motivated by computation of numerical quantities in a world with limited resources:

- Instead of the Hilbert space $\mathcal{H}$ (constructed over the uncountable and uncomputable complex numbers $\mathbb{C}$), we will consider variants constructed over finite fields [1, 2].
- Instead of real-valued probability measures producing results in the uncountable and uncomputable interval $[0, 1]$, we will consider finite set-valued probability measures [3].

We will then ask if it is possible to construct variants of quantum probability spaces under these conditions. The main question is related to the definition of probability measures: is it possible to still define a probability measure as a function that depends on a single state? Specifically,

- given a state $|\psi\rangle$, is there a probability measure mapping events to probabilities that only depends on $|\psi\rangle$? In the conventional quantum probability space, the answer is yes by the Born rule [4, 5] and the map is given by: $P \mapsto \langle\psi|P\psi\rangle$.

- given a probability measure μ mapping each event P to a probability, is there a *unique* state ψ such that $\mu(P) = \langle\psi|P\psi\rangle$? In the conventional case, the answer is yes by Gleason's theorem [6, 7, 8].

Andy says: Quantum meeting: the basics of classical and quantum standard and interval probabilities are becoming clearer (except need clearer exposition). The key step is the replacement of "sum" and "=" in the rule $\mu\left({}_{\mathcal{O}}\sum_{i=0}^{N-1} P_i \right) = {}_{\mathbb{R}}\sum_{i=0}^{N-1} \mu\left(P_i\right)$ by various logical operations on sets instead of actually summing unit norm partitions of unity in $\mathbb{R}$. What remains is to determine how the Meyer/Mermin debate plays out for finite precision or uncertain measurements of events, and its implication for the validity of the Kochen-Specker theorem, and hence for the validity of Gleason's theorem. Do we have anything ANALOGOUS to a Gleason theorem for interval probability quantum mechanics? If so, what are the implications for Kochen-Specker and Bell analogs for interval probabilities, and what are the implications for the Meyer/Mermin debate? And if not, what are the consequences? Finally, given whatever remains of Gleason/Kochen-Specker for uncertain event measurements, what happens when we redefine "sum" yet again for $\mathbb{F}_{p^2}$ valued quantum theories (and degenerate states, and density matrices) to create a non-wrapping extension of interval probability calculus to Galois fields? (The objective is to put possible/impossible and its extensions on a solid footing for DQC, and also to set up a seamless transition back to a continuous CQC limit that is consistent, while incorporating such issues as the cost of measurement precision.)

Yu-Tsung says: A mathematical reason why Gleason's theorem should not be valid in the interval-valued probability quantum mechanics. Every classical probability measure has a one-to-one correspondence to its Radon-Nikodym derivative. The Radon-Nikodym theorem extends to Gleason's theorem in quantum, while the role of Radon-Nikodym derivative is replaced by the density matrix. If there is no easy Radon-Nikodym theorem for classical interval-valued probability measure, it is natural that there is no easy Gleason's theorem for quantum interval-valued probability measure...
If Gleason's theorem is not valid, a non-Born quantum probability measure might exist in reality...
If there exists a non-Born quantum probability measure which cannot correspond to any state vectors, then what's the post-measurement quantum probability measure of it after measurement? If the updated post-measurement postulate is strange enough, then commuting observables may not interchange their measurement order... And commuting observables is compatible is the fundamental assumption for the Kochen-Specker theorem, the Bell theorem, and everything... If commuting observables becomes non-compatible, everything might need to be rewritten...

2 Classical Probability Spaces

A *probability space* specifies the necessary conditions for reasoning coherently about collections of uncertain events [9, 10, 11, 12]. We review the conventional presentation of probability spaces and then discuss the computational resources needed to estimate probabilities.

2.1 Classical Real-Valued Probability Spaces

The conventional definition of a probability space builds upon the field of real numbers. In more detail, a probability space consists of a *sample space* Ω, a space of *events* $\mathcal{E}$, and a *probability measure* μ mapping events in $\mathcal{E}$ to the real interval $[0,1]$. In this paper, we will only consider *finite* sets of events: we therefore restrict our attention to non-empty finite sets Ω as the sample space. The space of events $\mathcal{E}$ includes every possible subset of Ω: it is the powerset $2^\Omega = \{E | E \subseteq \Omega\}$.

Definition 1 (Probability Measure). Given the set of events $\mathcal{E}$, a *probability measure* is a function $\mu : \mathcal{E} \to [0,1]$ such that:

- $\mu(\varnothing) = 0$.

- $\mu(\Omega) = 1$.
- For any event E,
$$\mu\left(\Omega\backslash E\right) = 1 \,{}_{\mathbb{R}}\!- \mu\left(E\right) \ , \tag{1}$$
where $\Omega\backslash E$ is the complement event, and $1 \,{}_{\mathbb{R}}\!- \mu\left(E\right)$ explicitly specifies 1 and $\mu\left(E\right) \in \mathbb{R}$. Besides $\mathbb{R}$, we will prepose other symbols later to specify the type of operations, and they may be dropped when there is no ambiguity.
- For a collection $\{E_i\}_{i=0}^{N-1}$ of pairwise disjoint events, $\mu\left(\bigcup_{i=0}^{N-1} E_i\right) = {}_{\mathbb{R}}\sum_{i=0}^{N-1} \mu(E_i)$. Similar to the above condition, ${}_{\mathbb{R}}\sum_{i=0}^{N-1} \mu(E_i)$ explicitly specifies $\mu(E_i) \in \mathbb{R}$.

□

Notice that $\mu(\varnothing) = 0$ and equation (1) can be induced by other conditions. They are listed in here so that we can compare definition 1, lemma 1, definition 5, and definition 6 easily. We will compare these lemma and definitions when we formulate quantum interval-valued probability measures later.

Example 1 (Two-coins experiment). Consider an experiment that tosses two coins. We have four possible outcomes that constitute the sample space $\Omega = \{HH, HT, TH, TT\}$. There are 16 total events including for example the event $\{HH, HT\}$ that the first coin lands heads up, the event $\{HT, TH\}$ that the two coins land on opposite sides, and the event $\{HT, TH, TT\}$ that at least one coin lands tails up. Here is a possible probability measure for these events:

$$\begin{array}{rcl@{\qquad\qquad}rcl}
\mu(\varnothing) & = & 0 & \mu(\{HT,TH\}) & = & 2/3 \\
\mu(\{HH\}) & = & 1/3 & \mu(\{HT,TT\}) & = & 0 \\
\mu(\{HT\}) & = & 0 & \mu(\{TH,TT\}) & = & 2/3 \\
\mu(\{TH\}) & = & 2/3 & \mu(\{HH,HT,TH\}) & = & 1 \\
\mu(\{TT\}) & = & 0 & \mu(\{HH,HT,TT\}) & = & 1/3 \\
\mu(\{HH,HT\}) & = & 1/3 & \mu(\{HH,TH,TT\}) & = & 1 \\
\mu(\{HH,TH\}) & = & 1 & \mu(\{HT,TH,TT\}) & = & 2/3 \\
\mu(\{HH,TT\}) & = & 1/3 & \mu(\{HH,HT,TH,TT\}) & = & 1
\end{array}$$

The assignment satisfies the two constraints for probability measures: the probability of the entire sample space is 1, and the probability of every collection of disjoint events (e.g., $\{HT\} \cup \{TH\} = \{HT, TH\}$) is the sum of the individual probabilities. The probability of collections of non-disjoint events (e.g., $\{HT, TH\} \cup \{TH, TT\} = \{HT, TH, TT\}$) may add to something different than the probabilities of the individual events. It is useful to think that this probability measure is completely determined by the two coins in question and their characteristics, in the sense that each pair of coins induces a measure, and each measure must correspond to some pair of coins. The measure above is induced by two coins such that the first coin is twice as likely to land tails up than heads up and the second coin is double-headed. □

Although specifying a probability measure for every event looks complex, it can be simply constructed by
$$\mu(E) = {}_{\mathbb{R}}\sum\nolimits_{\omega \in E} m(\omega) \ , \tag{2}$$
where $m : \Omega \to [0, 1]$, ${}_{\mathbb{R}}\sum_{\omega\in\Omega} m(\omega) = 1$, and m is called the Radon-Nikodym derivative of μ (with respect to the counting measure) [13, 14]. The converse is also valid and called the Radon-Nikodym theorem.

Theorem 1 (Radon-Nikodym theorem for finite probability space [9, 12]). For every probability measure μ, there exists a unique Radon-Nikodym derivative m such that equation (2) holds. □

For example, the Radon-Nikodym derivative of μ in example 1 is:
$$m\left(HH\right) = 1/3, \quad m\left(HT\right) = 0, \quad m\left(TH\right) = 2/3, \quad m\left(TT\right) = 0 \ .$$

In a strict computational or experimental setting, one may question the reliance of the definition of probability space on the uncountable and uncomputable real interval $[0, 1]$. This interval includes numbers like

$0.h_1h_2h_3\ldots$ where h_i is 1 or 0 depending on whether Turing machine M_i halts or not. Such numbers cannot be computed. This interval also includes numbers like $\frac{\pi}{4}$ which can only be computed with increasingly large resources as the precision increases. Therefore, in a resource-aware computational or experimental setting, it is more appropriate to consider probability measures that map events to a set of elements computable with a fixed set of resources. We expand on this observation and then consider interval-valued probability measures [15, 16] in detail.

2.2 Measuring Probabilities: Buffon's Needle Problem

In the previous section, the probability $\mu(E)$ of each event E is known a priori. In reality, although each event is assumed to have a probability, the exact value of $\mu(E)$ may not be known. If we want to know its value, we can run N independent trials. Let $x_i = 1$ or 0 denote whether the event E occurs in the i-th trial or not, respectively, then $\mu(E)$ could be approximated to given accuracy $\epsilon > 0$ by the relative frequency $\frac{1}{N}\,{}_{\mathbb{R}}\sum_{i=0}^{N-1} x_i$ with the probability converging to one as N goes to infinity, i.e.,

$$\forall\epsilon > 0, \lim_{N\to\infty} \mu\left(\left|\mu(E) - \frac{1}{N}\,{}_{\mathbb{R}}\sum_{i=0}^{N-1} x_i\right| < \epsilon\right) = 1\ .$$

This fact is called the law of large numbers [17, 9, 18, 10, 19].

Suppose we drop a needle of length ℓ onto a floor made of equally spaced parallel lines a distance h apart, where $\ell < h$. It is a known fact that the probability of the needle crossing a line is $\frac{2\ell}{\pi h}$ [20, 21, 22, 18]. We analyze this situation in the mathematical framework of probability spaces paying special attention to the resources needed to estimate the probability computationally or experimentally.

To formalize the experiment, we consider an experimental setup consisting of a collection of N identical needles of length ℓ. We throw the N needles one needle at a time, and observe the number M of needles that cross a line. The sample space can be expressed as the set $\{X,-\}^N$ of sequences of characters of length N where each character is either X to indicate a needle crossing a line or $-$ to indicate a needle not crossing a line. If $N = 3$, the probability of the event that exactly 2 needles cross lines $\{-XX, X{-}X, XX-\}$ can be estimated by the relative frequency $\frac{2}{3}$. Generally, the probability of the event that exactly M needles out of the N total needles cross lines can be estimated by $\frac{M}{N}$.

In an actual experiment with 500 needles and the ratio $\frac{\ell}{h} = 0.75$ [22], it was found that 236 crossed a line so the relative frequency is 0.472 whereas the idealized mathematical probability is $0.4774\ldots$. In a larger experiment with 5000 needles and the ratio $\frac{\ell}{h} = 0.8$ [18], the relative frequency was calculated to be 0.5064 whereas the idealized mathematical probability is $0.5092\ldots$. We see that the observed probability approaches $\frac{2\ell}{\pi h}$ but only if *larger and larger resources* are expended. These resource considerations suggest that it is possible to replace the real interval $[0,1]$ with rational numbers up to a certain precision related to the particular experiment in question. There is clearly a connection between the number of needles and the achievable precision: in the hypothetical experiment with 3 needles, it is not sensible to retain 100 digits in the expansion of $\frac{2\ell}{\pi h}$.

There is however another more subtle assumption of unbounded computational power in the experiment. We are assuming that we can always determine with certainty whether a needle is crossing a line. But "lines" on the the floor have thickness, their distance apart is not exactly h, and the needles' lengths are not all absolutely equal to ℓ. These perturbations make the events "fuzzy." Thus, in an experiment with limited resources, it is not possible to talk about the idealized event that exactly M needles cross lines as this would require the most expensive needles built to the most precise accuracy, laser precision for drawing lines on the floor, and the most powerful microscopes to determine if a needle does cross a line. Instead we might talk about the event that $M - \delta$ needles evidently cross lines and $M + \delta'$ needles plausibly cross lines where δ and δ' are resource-dependent approximations. This fuzzy notion of events leads to probabilities being only calculable within intervals of confidence reflecting the certainty of events and their plausibility. This is indeed consistent with published experiments: in an experiment with 3204 needles and the ratio $\frac{\ell}{h} = 0.6$ [21], 1213 needles clearly crossed a line and 11 needles were close enough to plausibly be considered as crossing the line: we would express the probability in this case as the interval $\left[\frac{1213}{3204}, \frac{1224}{3204}\right]$ expressing that we are certain that the event has probability at least $\frac{1213}{3204}$ but it is possible that it would have probability $\frac{1224}{3204}$.

Recall that the relative frequency will approximate the probability of an event if the event has a probability. What if the event doesn't have infinity precise probability because of the experimental limit? In this case, two sequences of independent copies of experimental results can have their relative frequencies converge almost surely to different limits [23, 24]. In other words, to get a better approximation of the probability, the quality of the experimental equipment cannot be compensated by the number of independent trials.

2.3 Classical Interval-valued probability measures

As motivated above, an event E_1 may have an interval of probability $[l_1, r_1]$. Assume that another disjoint event E_2 has interval probability $[l_2, r_2]$, what is the interval probability of the event $E_1 \cup E_2$? The answer is somewhat subtle: although it is possible to use the sum of the intervals $[l_1 \,_{\mathbb{R}}\!+ l_2, r_1 \,_{\mathbb{R}}\!+ r_2]$ as the combined probability, one can do find a much tighter interval if information *against* the event (i.e., information about the complement event) is also taken into consideration. Formally, for a general event E with probability $[l, r]$, the evidence that contradicts E is an evidence supporting the complement of E. The complement of E must therefore have probability $[1 \,_{\mathbb{R}}\!- r, 1 \,_{\mathbb{R}}\!- l]$ which we abbreviate $[1,1] \,_{\mathscr{I}}\!- [l, r]$, where the preposing $\mathscr{I}$ specifies we subtract intervals. Given a collection of intervals $\mathscr{I}$, an $\mathscr{I}$–interval-valued probability measure is a function $\bar{\mu} : \mathcal{E} \to \mathscr{I}$ such that [16]:[1]

- $\bar{\mu}(\varnothing) = [0, 0]$.
- $\bar{\mu}(\Omega) = [1, 1]$.
- For every collection of pairwise disjoint events $\{E_i\}_{i=0}^{N-1}, \{E'_i\}_{i=0}^{N'-1} \subseteq \mathcal{E}$ with $\Omega = \left(\bigcup_{i=0}^{N-1} E_i\right) \cup \left(\bigcup_{i=0}^{N'-1} E'_i\right)$, we have

$$\bar{\mu}\left(\bigcup_{i=0}^{N-1} E_i\right) \subseteq \left[\max\left\{1 - {}_{\mathbb{R}}\sum_{i=0}^{N'-1} r'_i,\ {}_{\mathbb{R}}\sum_{i=0}^{N-1} l_i\right\}, \min\left\{1 - {}_{\mathbb{R}}\sum_{i=0}^{N'-1} l'_i,\ {}_{\mathbb{R}}\sum_{i=0}^{N-1} r_i\right\}\right] ,$$

  where $\bar{\mu}(E_i) = [l_i, r_i]$ and $\bar{\mu}(E'_i) = [l'_i, r'_i]$ for all i.

In order to understand this definition, we tear the last condition apart and provide an equivalent definition of interval-valued probability measures.

Lemma 1. Given a sample space Ω and its event $\mathcal{E}$, a function $\bar{\mu} : \mathcal{E} \to [0, 1]$ is a classical interval-valued probability measure if and only if $\bar{\mu}$ satisfies the following conditions:

- $\bar{\mu}(\varnothing) = [0, 0]$.
- $\bar{\mu}(\Omega) = [1, 1]$.
- For any event E,

$$\bar{\mu}(\Omega \backslash E) = [1, 1] \,_{\mathscr{I}}\!- \bar{\mu}(E) \ . \tag{3}$$

- For a collection $\{E_i\}_{i=0}^{N-1}$ of pairwise disjoint events, we have $\bar{\mu}\left(\bigcup_{i=0}^{N-1} E_i\right) \subseteq {}_{\mathscr{I}}\sum_{i=0}^{N-1} \bar{\mu}(E_i)$, where ${}_{\mathscr{I}}\sum_{i=0}^{N-1} [l_i, r_i] = \left[{}_{\mathbb{R}}\sum_{i=0}^{N-1} l_i,\ {}_{\mathbb{R}}\sum_{i=0}^{N-1} r_i\right]$. We may drop the preposing $\mathscr{I}$ when summands are clearly intervals.

□

We will explain why the last condition is expressed using $\subseteq$ by the following.

[1] The right-end of the interval-valued probability measure is a special case of (sub-additive) capacity [25, 26, 27] or outer measure [14]. The left-end of the interval-valued probability measure is a special case of (super-additive cooperative) game [28, 29].

Example 2 (Two-coin experiment with interval probability). We split the unit interval $[0,1]$ in the following four closed sub-intervals: $[0,0]$ which we call $\mathscr{impossible}$, $[0,\frac{1}{2}]$ which we call $\mathscr{unlikely}$, $[\frac{1}{2},1]$ which we call $\mathscr{likely}$, and $[1,1]$ which we call $\mathscr{certain}$. Using these new values, we can modify the probability measure of Ex. 1 by mapping each numeric value to the smallest sub-interval containing it to get the following:

$$
\begin{array}{rcl@{\qquad}rcl}
\bar{\mu}(\varnothing) &=& \mathscr{impossible} & \bar{\mu}(\{HT,TH\}) &=& \mathscr{likely} \\
\bar{\mu}(\{HH\}) &=& \mathscr{unlikely} & \bar{\mu}(\{HT,TT\}) &=& \mathscr{impossible} \\
\bar{\mu}(\{HT\}) &=& \mathscr{impossible} & \bar{\mu}(\{TH,TT\}) &=& \mathscr{likely} \\
\bar{\mu}(\{TH\}) &=& \mathscr{likely} & \bar{\mu}(\{HH,HT,TH\}) &=& \mathscr{certain} \\
\bar{\mu}(\{TT\}) &=& \mathscr{impossible} & \bar{\mu}(\{HH,HT,TT\}) &=& \mathscr{unlikely} \\
\bar{\mu}(\{HH,HT\}) &=& \mathscr{unlikely} & \bar{\mu}(\{HH,TH,TT\}) &=& \mathscr{certain} \\
\bar{\mu}(\{HH,TH\}) &=& \mathscr{certain} & \bar{\mu}(\{HT,TH,TT\}) &=& \mathscr{likely} \\
\bar{\mu}(\{HH,TT\}) &=& \mathscr{unlikely} & \bar{\mu}(\{HH,HT,TH,TT\}) &=& \mathscr{certain}
\end{array}
$$

Despite the absence of any numeric information, the probability measure is quite informative: it reveals that the second coin is double-headed and that the first coin is biased. To understand the $\subseteq$-condition, consider the following calculation:

$$
\begin{aligned}
& \bar{\mu}(\{HH\}) \;{}_{\mathscr{I}}\!+ \bar{\mu}(\{HT\}) \;{}_{\mathscr{I}}\!+ \bar{\mu}(\{TH\}) \;{}_{\mathscr{I}}\!+ \bar{\mu}(\{TT\}) \\
= \; & \mathscr{impossible} \;{}_{\mathscr{I}}\!+ \mathscr{unlikely} \;{}_{\mathscr{I}}\!+ \mathscr{impossible} \;{}_{\mathscr{I}}\!+ \mathscr{likely} \\
= \; & [0,0] \;{}_{\mathscr{I}}\!+ \left[0,\frac{1}{2}\right] \;{}_{\mathscr{I}}\!+ [0,0] \;{}_{\mathscr{I}}\!+ \left[\frac{1}{2},1\right] = \left[\frac{1}{2},\frac{3}{2}\right]
\end{aligned}
$$

If we were to equate $\bar{\mu}(\Omega)$ with the sum of the individual probabilities, we would get that $\bar{\mu}(\Omega) = \left[\frac{1}{2},\frac{3}{2}\right]$. However, using the fact that $\bar{\mu}(\varnothing) = \mathscr{impossible}$, we have $\bar{\mu}(\Omega) = 1 \;{}_{\mathscr{I}}\!- \bar{\mu}(\varnothing) = \mathscr{certain} = [1,1]$. This interval is tighter and a better estimate for the probability of the event Ω, and of course it is contained in $\left[\frac{1}{2},\frac{3}{2}\right]$. However it is only possible to exploit the information about the complement when all four events are combined. Thus the $\subseteq$-condition allows us to get an estimate for the combined event from each of its constituents and then gather more evidence knowing the aggregate event. □

In contrast to real-valued probability measures, there is no easy Radon-Nikodym theorem to simplify an interval-valued probability measure to its Radon-Nikodym derivative. So far, the interval-valued Radon-Nikodym Theorems are either too restricted to apply to all interval-valued probability measures [25, 26, 27] or corresponding an interval-valued probability measure $\bar{\mu} : \mathcal{E} \to \mathscr{I}$ to another set function $\bar{m} : \mathcal{E} \to \bar{\mathscr{I}}$ which is not as simple as the original Radon-Nikodym derivative [30, 27].

Although there is no easy Radon-Nikodym theorem, an interval-valued probability measure can still be understood by its relation with real-valued probability measures. For example, the real-valued probability measure in example 1 can construct the interval-valued probability measure in example 2. Conversely, we could understand the interval-valued probability measure in example 2 by the real-valued probability measure and its Radon-Nikodym derivative in example 1. The relation between example 1 and 2 can be abstracted into the definition of core. Besides of core, we also define another property, convex, which will be used later.

Definition 2. Given an interval-valued probability measure $\bar{\mu} : \mathcal{E} \to \mathscr{I}$:

- A *core*[2] of $\bar{\mu}$ is the set $\operatorname{core}(\bar{\mu}) = \{\text{probability measure } \mu : \mathcal{E} \to [0,1] \mid \forall E \in \mathcal{E}. \mu(E) \in \bar{\mu}(E)\}$ [28, 30, 23, 24].
- $\bar{\mu}$ is called *convex*, *2-monotone*, or *supermodular* if $\bar{\mu}(E_0 \cup E_1) \;{}_{\mathscr{I}}\!+ \bar{\mu}(E_0 \cap E_1) \subseteq \bar{\mu}(E_0) \;{}_{\mathscr{I}}\!+ \bar{\mu}(E_1)$ for

[2]The core is also called a structure [15].

all $E_0, E_1 \in \mathcal{E}$ [28, 30, 23, 24, 29].[3]

□

It is common to consider convex interval-valued probability measures. For example, a special case of convex interval-valued probability measures consists the Dempster-Shafer belief function as the left-end and the Dempster-Shafer plausibility as the right-end [31, 10, 30, 27, 23, 24]. Dempster-Shafer functions, also called *completely* or *totally monotone*, plays an important role in the theory of non-additive belief. Furthermore, convex interval-valued probability measures have a good property that they can always be associated with real-valued probability measures.

Theorem 2 (Shapley [28]). Every convex interval-valued probability measure has a nonempty core. □

For general interval-valued probability measures, the following theorem gives an algorithmic way to decide whether it has a non-empty core.

Theorem 3 (Bondareva-Shapley [32, 33]). Let l be the left-end of an interval-valued probability measure $\bar{\mu}$, $\bar{\mu}$ has a non-empty core if and only if for any minimal balanced sets $\{E_i\}_{i=0}^{N-1}$, we have ${}_{\mathbb{R}}\sum_{i=0}^{N-1} \gamma_i l(E_i) \leq 1$, where $\{E_i\}_{i=0}^{N-1}$ is called minimal balanced sets if they are distinct, $\varnothing \neq E_i \subsetneq \Omega$, and there exist a unique positive coefficient $\{\gamma_i\}_{i=0}^{N-1}$ such that ${}_{\mathbb{R}}\sum_{i=0 \text{ and } \omega \in E_i}^{N-1} \gamma_i = 1$ for all $\omega \in \Omega$. □

The Bondareva-Shapley theorem can be used to algorithmically decide non-empty core because the minimal balanced sets can be generated by an algorithm [34, 29]. However, the number of minimal balanced sets grows very rapidly with $|\Omega|$, the size of Ω [32, 35][4], so that we have only verified every interval-valued probability measure has a non-empty core when $|\Omega| \leq 5$ so far.[5] Based on the previous fact, we have the following conjecture.

Conjecture 1. Every interval-valued probability measure has a nonempty core. □

3 Quantum Probability Spaces

The mathematical framework above assumes that there exists a predetermined set of events that are independent of the particular experiment. However, in many practical situations, the structure of the event

[3] Yu-Tsung says: In order to simplify the discussion, let l be the left-end of an interval-valued probability measure $\bar{\mu}$. Because of equation (3), $\bar{\mu}$ is convex if and only if

$$l(E_0 \cup E_1) \;{}_{\mathbb{R}}\!+ l(E_0 \cap E_1) \geq l(E_0) \;{}_{\mathbb{R}}\!+ l(E_1)\ , \tag{4}$$

which is a special case of the Dempster-Shafer condition [10, 29]

$$l\left(\bigcup_{i=0}^{N-1} E_i\right) \geq {}_{\mathbb{R}}\sum_{I \subseteq \{0,\ldots,N-1\}} (-1)^{|I|+1}\, l\left(\bigcap_{i\in I} E_i\right)\ . \tag{5}$$

Therefore, in order to understand the meaning of equation (4), we could try to understand meaning of the Dempster-Shafer belief function. Unfortunately, when people introduce the Dempster-Shafer theory, they usually don't start from equation (5). Instead, they introduce a basic belief assignment $m(E) = {}_{\mathbb{R}}\sum_{E' \subseteq E} (-1)^{|E \setminus E'|}\, l(E')$ first. $m(E)$ represents the belief committed to the event E and only to it, in the sense that it could not be committed to any proper subset of E. Then, equation (5) is equivalent to $m(E) \geq 0$ for all E which is intuitive.
Comparing to the Dempster-Shafer condition (5), the equivalent condition for convexity is more complex. At least, when l is convex, we have

$$m(\{\omega_0, \omega_1\}) \geq 0 \tag{6}$$

for all $\{\omega_0, \omega_1\} \subseteq \Omega$, although equation (6) is not strong enough to imply l is convex.

[4] Yu-Tsung says: Prove the number of minimal balanced sets grow exponentially or more?

[5] Yu-Tsung says: The program for $|\Omega| = 6$ blows up the memory again. Hence, currently I only run a extremely slow backtrack program searching for a IVPM with an empty core...
Although the computer simulation supports the idea that every IVPM has a nonempty core, if we look the equations, $\bar{\mu}$ is an interval-valued probability measure if and only if its left-end l satisfies equation (4) when $E_0 \cap E_1 = \varnothing$ or $E_0 \cup E_1 = \Omega$. This condition is much weaker than convex, and I will be surprise if this condition is strong enough to imply a nonempty core.

space is only partially known and the precise dependence of two events on each other cannot, a priori, be determined with certainty. In the quantum framework, this partial knowledge is compounded by the fact that there exist non-commuting events which cannot happen simultaneously. To accommodate these more complex situations, we abandon the sample space Ω and reason directly about events. A quantum probability space therefore consists of just two components: a set of events $\mathcal{E}$ and a probability measure $\mu : \mathcal{E} \to [0,1]$. To properly explain the quantum probability, we will first discuss projection operators as quantum events, provide an easy example, define quantum probability measures, and extend to quantum interval-valued probability.

3.1 Quantum Events

Definition 3 (Projection Operators; Orthogonality [36, 8, 7, 11, 12]). Given a Hilbert space $\mathcal{H}$, an event[6] mathematically is represented as a projection operator $P : \mathcal{H} \to \mathcal{H}$ onto a linear subspace S of $\mathcal{H}$. The set of all events $\mathcal{E}$ can be defined recursively as follow: [7]

- $\mathbb{0}$ is a projection.
- For any pure state $|\psi\rangle$, $|\psi\rangle\langle\psi|$ is a projection operator.
- Projection operators P_0 and P_1 are *orthogonal* if $P_0 P_1 = P_1 P_0 = \mathbb{0}$. The sum of two projection operators $P_0 \;{}_{\mathscr{O}}{+}\; P_1$ is also a projection operator if and only if they are orthogonal, where the preposing subscript $\mathscr{O}$ means ${}_{\mathscr{O}}{+}$ is an operation between operators.
- Conversely, every projection P can be expressed as ${}_{\mathscr{O}}\sum_{j=0}^{N-1} |\psi_j\rangle\langle\psi_j|$, where P actually projects onto the linear subspace S with an orthonormal basis $\{|\psi_j\rangle\}_{j=0}^{N-1}$.

□

Because quantum events are projection operators, operations and properties of quantum events can be written in terms of those of operators.

Definition 4 (Ideal Measurement; Complement; Commutativity; disjunction).

- A set of projections $\{P_i\}_{i=0}^{N-1}$ is called an *ideal measurement* if it is a partition of the identity, i.e., ${}_{\mathscr{O}}\sum_{i=0}^{N-1} P_i = \mathbb{1}$ [12]. In this case, projections $\{P_i\}_{i=0}^{N-1}$ must be mutually orthogonal [11].
- If P is a projection operator, then $\mathbb{1} \;{}_{\mathscr{O}}{-}\; P$ is also a projection operator, called *complement*. It is orthogonal to P, and corresponds to the complement event $\Omega \backslash E$ in classical probability [11].
- Projection operators P_0 and P_1 *commute* if $P_0 P_1 = P_1 P_0$. The product of two projection operators $P_0 P_1$ is also a projection operator if and only if they commute. This corresponds to the classical intersection between events [7, 11].
- For two commuting projection operators P_0 and P_1, their *disjunction* $P_0 \;{}_{\mathscr{O}}{\vee}\; P_1$ is defined to be $P_0 \;{}_{\mathscr{O}}{+}\; P_1 \;{}_{\mathscr{O}}{-}\; P_0 P_1$ [11].

□

Based on the above definition and properties of quantum events, we can present an example before giving the formal definition of quantum probability measures.

Example 3 (One-qubit quantum probability space). Consider a one-qubit Hilbert space with states $|\phi\rangle$, which can be expressed as a linear combination of $|0\rangle$ and $|1\rangle$, i.e., $|\phi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$ such that $|\alpha|^2 + |\beta|^2 = 1$, $\alpha, \beta \in \mathbb{C}$. The set of events associated with this Hilbert space consists of all projection operators. Each event is interpreted as a possible post-measurement state of a quantum system in current state $|\phi\rangle$. For

[6] An event is formally called an experimental proposition [37], a question [36, 38], or an elementary quantum test [7].
[7] "Projection" is sometimes called "orthogonal projection" or "self-adjoint projection" to emphasize $P^\dagger = P$ [11, 39].

example, the event $|0\rangle\langle 0|$ indicates that the post-measurement state will be $|0\rangle$; the event $|1\rangle\langle 1|$ indicates that the post-measurement state will be $|1\rangle$; the event $|+\rangle\langle +|$ where $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ indicates that the post-measurement state will be $|+\rangle$; the event $\mathbb{1} = |0\rangle\langle 0| \;{}_{\mathscr{O}}\!+ |1\rangle\langle 1|$ indicates that the post-measurement state will be a linear combination of $|0\rangle$ and $|1\rangle$; and the empty event $\mathbb{0}$ states that the post-measurement state will be the empty state. As in the classical case, a probability measure is a function that maps events to $[0,1]$. Here is a partial specification of a possible probability measure:

$$\mu(\mathbb{0}) = 0, \quad \mu(\mathbb{1}) = 1, \quad \mu(|0\rangle\langle 0|) = 1, \quad \mu(|1\rangle\langle 1|) = 0, \quad \mu(|+\rangle\langle +|) = 1/2, \quad \ldots$$

Note that, similarly to the classical case, the probability of $\mathbb{1}$ is 1 and the probability of collections of orthogonal events (e.g., $|0\rangle\langle 0| + |1\rangle\langle 1|$) is the sum of the individual probabilities. A collection of non-orthogonal events (e.g., $|0\rangle\langle 0|$ and $|+\rangle\langle +|$) is however not even a valid event. In the classical example, we argued that each probability measure is uniquely determined by two actual coins. A similar (but much more subtle) argument is valid also in the quantum case. By postulates of quantum mechanics and Gleason's theorem, it turns out that for large enough quantum systems, each probability measure is uniquely determined by an actual quantum state. □

3.2 Quantum Real-valued Probability Measures

Definition 5 (Quantum Probability Measure [36, 6, 8, 39]). Given a Hilbert space $\mathcal{H}$ with its set of events $\mathcal{E}$, a *quantum probability measure* is a function $\mu : \mathcal{E} \to [0,1]$ such that:[8]

- $\mu(\mathbb{0}) = 0$.
- $\mu(\mathbb{1}) = 1$.
- For any projection P,
$$\mu(\mathbb{1} \;{}_{\mathscr{O}}\!- P) = 1 \;{}_{\mathbb{R}}\!- \mu(P) \ . \tag{7}$$
- For a set of mutually orthogonal projections $\{P_i\}_{i=0}^{N-1}$, we have
$$\mu\left({}_{\mathscr{O}}\!\sum_{i=0}^{N-1} P_i \right) = {}_{\mathbb{R}}\!\sum_{i=0}^{N-1} \mu(P_i) \ . \tag{8}$$

□

Similar to classical definition 1, $\mu(\mathbb{0}) = 0$ and equation (7) can be induced by other conditions. A set of events $\mathcal{E}$ together with a quantum probability measure is called a *quantum probability space*. Comparing to classical probability spaces, the empty set $\varnothing$ corresponds to the empty projection $\mathbb{0}$, and the event of whole space Ω corresponds to the identity projection $\mathbb{1}$. In contrast, the union $\cup$ of any two events always gives an event classically, but the operator addition ${}_{\mathscr{O}}\!+$ of two projections may not be a projection. As the result, the classical condition $\mu(E_0 \cup E_1) = \mu(E_0) \;{}_{\mathbb{R}}\!+ \mu(E_1)$ is always defined, and it is true when E_0 and E_1 are disjoint; however, $\mu(P_0 \;{}_{\mathscr{O}}\!+ P_1) = \mu(P_0) \;{}_{\mathbb{R}}\!+ \mu(P_1)$ is always true whenever the left-handed side is defined.[9]

Recall that classical probability measure can be constructed from its Radon-Nikodym derivative. Similarly, a quantum probability measure can be easily constructed by states according to the Born rule [4, 5, 40]. For each pure normalized ($\langle\phi|\phi\rangle = 1$) quantum state $|\phi\rangle$, the Born rule induces a probability measure μ_ϕ^B as follows:

$$\mu_\phi^B(P) = \langle\phi|P\phi\rangle \ .$$

[8] It is possible to define a more general space of events consisting of all operators $\mathcal{A}$ on $\mathcal{H}$ and to extend μ linearly to $\mathcal{A}$ and consider $\mu : \mathcal{A} \to \mathbb{C}$ [39, 12]. When an operator $A \in \mathcal{A}$ is Hermitian, $\mu(A)$ is the expectation value of A which could bigger than one because μ is linear. We does not take this approach because we want to focus only on probability.

[9] Yu-Tsung says: The definition of independence is interesting, and we definitely need to discuss it when we want to discuss Bell's theorem and the Kochen-Specker theorem. However, the definition diverges. So we just leave it so far, and we will go back if we really need it in this paper. (Maybe when discussing repeating experiments?)

Moreover, the Born rule can be extended to a mixed state. When we prepare a state, suppose the probability of preparing state $|\phi_j\rangle$ is q_j, the state of the system can be expressed as a density matrix $\rho = {}_{\mathscr{O}}\sum_{j=0}^{N-1} q_j |\phi_j\rangle\langle\phi_j|$, where ${}_{\mathbb{R}}\sum_{j=0}^{N-1} q_j = 1$. Then, the quantum probability measure introduced by ρ is the combination of $\mu_{\phi_j}^B$ with respect to probability q_j [7, 19, 40]:

$$\mu_\rho^B(P) \quad = \quad \mathrm{Tr}(\rho P) = {}_{\mathbb{R}}\sum_{j=0}^{N-1} q_j \mu_{\phi_j}^B(P) \ . \tag{9}$$

3.3 Measuring Quantum Probabilities Ideally: From Radon-Nikodym to Gleason's Theorem

Similar to the classical case, by applying the law of large numbers, quantum probabilities can be estimated by relative frequencies. For example, if we want to know the probability of the spin up in the Stern-Gerlach experiment [41, 7, 19, 11], we can put a beam of silver atoms in a highly inhomogeneous magnetic field, and counting the number of atoms deflects up. Ideally, if the local field of strength directs to the z-axis, and all particles have the same velocity, the Stern-Gerlach experiment only produces two spots corresponding to $|0\rangle$ and $|1\rangle$. Notice that whenever the magnetic field of the Stern-Gerlach is fixed, i.e., an ideal measurement is picked, measuring the spin is exactly the same as tossing a coin. In another word, if somebody claims she is tossing a coin behind a veil, and only shows the resulting heads or tails, we cannot distinguish whether she has really tossed a coin, or she has run an Stern-Gerlach experiment, and shows us the head, the tail, or the side of coin if the silver atoms is spin up, spin down, or just hit the middle, respectively.

Mathematically, to fix the magnetic field of the Stern-Gerlach experiment is to pick a orthonormal basis $\Omega = \{|\psi_j\rangle\}_{j=0}^{d-1}$ for a given Hilbert space $\mathcal{H}$ of dimension d, and the state of the silver atom can be represented by a density matrix ρ written in the same basis, i.e., $\rho = {}_{\mathscr{O}}\sum_{i=0}^{d-1} {}_{\mathscr{O}}\sum_{i=0}^{d-1} r_{ij} |\psi_i\rangle\langle\psi_j|$. As we discussed in the last paragraph, the Born rule of the mixed state ρ induces a classical probability measure $\mu^\Omega : 2^\Omega \to [0,1]$ on Ω. Let $E \in 2^\Omega$ be a classical event and $P = {}_{\mathscr{O}}\sum_{|\psi_j\rangle\in E} |\psi_j\rangle\langle\psi_j|$ be the quantum event corresponding to subspace span by E, we can have

$$\mu^\Omega(E) = \mu_\rho^B(P) \overset{(9)}{=} \mathrm{Tr}(\rho P) = {}_{\mathbb{R}}\sum_{|\psi_j\rangle\in E} r_{jj} \ . \tag{10}$$

By equation (2), the Radon-Nikodym derivative of μ' is $m(|\psi_j\rangle) = r_{jj}$, which is just the diagonal elements of ρ. In this sense, the Radon-Nikodym derivative m is a special case of the density matrix ρ, and the Radon-Nikodym theorem is a special case of Gleason's theorem [42, 8, 40].

Theorem 4 (Gleason's theorem [6, 8, 7]). In a Hilbert space $\mathcal{H}$ of dimension $d \geq 3$, given a quantum probability measure $\mu : \mathcal{E} \to [0,1]$, there exists a unique mixed state ρ such that $\mu = \mu_\rho^B$.

It is instructive to study a counterexample when $d = 2$, i.e., the case of a one-qubit system.

Example 4 (One-qubit quantum probability measure). Consider a quantum probability measure $\mu : \mathcal{E} \to [0,1]$ defined as follow:

$$\mu(P) = \begin{cases} 1 & \text{, if } P = |+\rangle\langle+| \ ; \\ 0 & \text{, if } P = |-\rangle\langle-| \ ; \\ \mu_{|0\rangle}^B(P) & \text{, otherwise.} \end{cases}$$

On one hand, μ is a quantum probability measure. Because μ is almost the same as the quantum probability measure $\mu_{|0\rangle}^B$, we only need to check the orthogonal pair $|+\rangle\langle+|$ and $|-\rangle\langle-|$: $\mu(|+\rangle\langle+|) \,{}_{\mathbb{R}}\!+ \mu(|-\rangle\langle-|) = 1 \,{}_{\mathbb{R}}\!+ 0 = 1$. On the other hand, μ cannot be induced by any mixed state because $\mu(|+\rangle\langle+|) = \mu(|0\rangle\langle0|) = 1$, and $\mu_\rho^B(P) = 1$ if and only if ρ represents a pure state and $\rho = P$. □

3.4 Measuring Quantum Probabilities in Reality: Quantum Interval-valued Probability Measures

In reality, a Stern-Gerlach experiment requires a lot of resource to keep the local field of strength directing to the z-axis precisely, to keep the atoms having almost the same velocity, and to point out the exact position each particle landed on. If the field of strength does not perfectly direct to the z-axis, we do not really test the quantum event we want to test. This problem will be handled later when we introduce the discrete quantum theory. Even if the field of strength perfectly directs to the z-axis, the variant velocity makes spots broader and more washed out. Together with the precision limit of the detector, it may sometimes be hard to decide a particle corresponding to which state[10].

Similar to Buffon's needles, this kind of fuzziness can be taken into account by associating each quantum event with an interval-valued probability $[l,r]$. Therefore, we should plug in the definition of interval-valued probability measure into the definition of quantum probability space for each ideal measurement. More specifically, for classical probability measures, definition 1 and lemma 1 are almost the same except the last condition using the equal sign = in real-valued definition 1 and the subset equal sign $\subseteq$ in interval-valued lemma 1. Therefore, for quantum probability measures, replacing the the equal sign = in real-valued definition 5 by the subset equal sign $\subseteq$ gives the only reasonable definition for quantum interval-valued probability measures.

Definition 6 (Quantum Interval-valued Probability Measure). Given a Hilbert space $\mathcal{H}$ with the set of quantum events $\mathcal{E}$, and a collection of intervals $\mathscr{I}$, a *quantum $\mathscr{I}$-interval-valued probability measure* is a function $\bar{\mu} : \mathcal{E} \to \mathscr{I}$ such that:

- $\bar{\mu}(\mathbb{0}) = [0,0]$.
- $\bar{\mu}(\mathbb{1}) = [1,1]$.
- For any projection P,
$$\bar{\mu}\left(\mathbb{1} \;_{\mathscr{O}}\!- P\right) = [1,1] \;_{\mathscr{I}}\!- \bar{\mu}\left(P\right) \ . \tag{11}$$
- For a set of mutually orthogonal projections $\{P_i\}_{i=0}^{N-1}$, we have
$$\bar{\mu}\left({}_{\mathscr{O}}\sum_{i=0}^{N-1} P_i \right) \subseteq {}_{\mathscr{I}}\sum_{i=0}^{N-1} \bar{\mu}\left(P_i\right) \ . \tag{12}$$

□

Similar to the classical interval-valued probability measure, a quantum interval-valued probability measure can be constructed from a real-valued one.

Example 5 (Quantum three-value interval-valued probability measure). We consider three intervals $[0,0]$, $[1,1]$ and $[0,1]$, where $[0,0]$ and $[1,1]$ are called *impossible* and *certain* as before, and $[0,1]$ is called *unknown* because it provides no information. For any Hilbert space and any quantum probability measure $\mu : \mathcal{E} \to [0,1]$, we can define a quantum interval-valued probability measure $\bar{\mu} : \mathcal{E} \to \mathscr{I}$ by $\bar{\mu}(P) = \iota\left(\mu(P)\right)$, where $\iota : [0,1] \to \mathscr{I}$ is defined by

$$\iota(x) = \begin{cases} \mathit{certain} & \text{, if } x = 1 \ ; \\ \mathit{impossible} & \text{, if } x = 0 \ ; \\ \mathit{unknown} & \text{, otherwise.} \end{cases}$$

ι has two interesting properties $\iota\left(1 \;_{\mathbb{R}}\!- x\right) = [1,1] \;_{\mathscr{I}}\!- \iota\left(x\right)$ and $\iota\left({}_{\mathbb{R}}\sum_{i=0}^{N-1} x_i \right) \subseteq {}_{\mathscr{I}}\sum_{i=0}^{N-1} \iota\left(x_i\right)$, where x and ${}_{\mathbb{R}}\sum_{i=0}^{N-1} x_i \in [0,1]$. By applying these two properties, it is easy to verify $\bar{\mu}$ is a quantum $\mathscr{I}$-interval-valued probability measure. □

Similar to classical interval-valued probability measures which have no easy Radon-Nikodym theorem, quantum interval-valued probability measures have no easy Gleason's theorem. Although we could define core and convex for quantum interval-valued probability measures, in contrast to theorem 2, a convex quantum interval-valued probability measure may have a empty core, i.e., it cannot correspond to a quantum real-valued probability measure or a mixed state.

[10] Yu-Tsung says: Add citations to support the idea... Haven't found suitable ones...

Amr says: preparation fuzzy, device fuzzy, Meyer [43]

Definition 7. Given a quantum interval-valued probability measure $\bar{\mu} : \mathcal{E} \to \mathscr{I}$:

- A *core* of $\bar{\mu}$ is the set $\mathrm{core}(\bar{\mu}) = \{\text{quantum probability measure } \mu : \mathcal{E} \to [0,1] | \forall E \in \mathcal{E}. \mu(E) \in \bar{\mu}(E)\}$.
- $\bar{\mu}$ is called convex if

$$\bar{\mu}(P_0 \,{}_{\mathscr{O}}\!\vee P_1) \,{}_{\mathscr{I}}\!+ \bar{\mu}(P_0 P_1) \subseteq \bar{\mu}(P_0) \,{}_{\mathscr{I}}\!+ \bar{\mu}(P_1) \tag{13}$$

for all commuting $P_0, P_1 \in \mathcal{E}$.

□

Theorem 5. There is a convex quantum interval-valued probability measure $\bar{\mu} : \mathcal{E} \to \mathscr{I}$ such that $\mathrm{core}(\bar{\mu}) = \emptyset$. □

This theorem can be proved by the following example.

Example 6 (Three-dimensional quantum three-value interval-valued probability measure). Given a three dimensional Hilbert space with an orthonormal basis $\{|0\rangle, |1\rangle, |2\rangle\}$. Consider $\mathscr{I} = \{\mathit{certain}, \mathit{impossible}, \mathit{unknown}\}$ and $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, and $|+'\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |2\rangle)$. Let

$$\begin{aligned} \bar{\mu}(\mathbb{0}) = \bar{\mu}(|0\rangle\langle 0|) = \bar{\mu}(|+\rangle\langle +|) = \bar{\mu}(|+'\rangle\langle +'|) &= \mathit{impossible}, \\ \bar{\mu}(\mathbb{1}) = \bar{\mu}(\mathbb{1} \,{}_{\mathscr{O}}\!- |0\rangle\langle 0|) = \bar{\mu}(\mathbb{1} \,{}_{\mathscr{O}}\!- |+\rangle\langle +|) = \bar{\mu}(\mathbb{1} \,{}_{\mathscr{O}}\!- |+'\rangle\langle +'|) &= \mathit{certain}, \\ \bar{\mu}(P) &= \mathit{unknown}, \text{ otherwise.} \end{aligned}$$

To verify $\bar{\mu}$ is a quantum interval-valued probability measure, first notice that $\bar{\mu}(P_0 \,{}_{\mathscr{O}}\!+ P_1) \subseteq \bar{\mu}(P_0) \,{}_{\mathscr{I}}\!+ \bar{\mu}(P_1)$ implies equation (12) by induction. Second, equation (11) implies $\bar{\mu}(\mathbb{1}) \subseteq \bar{\mu}(P) \,{}_{\mathscr{I}}\!+ \bar{\mu}(\mathbb{1} \,{}_{\mathscr{O}}\!- P)$. Therefore, to prove equation (12), it is sufficient to check

$$\bar{\mu}(|\psi_0\rangle\langle\psi_0| \,{}_{\mathscr{O}}\!+ |\psi_1\rangle\langle\psi_1|) \subseteq \bar{\mu}(|\psi_0\rangle\langle\psi_0|) \,{}_{\mathscr{I}}\!+ \bar{\mu}(|\psi_1\rangle\langle\psi_1|)$$

for orthogonal $|\psi_0\rangle$ and $|\psi_1\rangle$. This equation always holds because we always have $\mathit{unknown} \subseteq \bar{\mu}(|\psi_0\rangle\langle\psi_0|) \,{}_{\mathscr{I}}\!+ \bar{\mu}(|\psi_1\rangle\langle\psi_1|)$.

Next, we need to verify that $\bar{\mu}$ is convex. First, for orthogonal projections P_0 and P_1, equation (12) implies equation (13). Second, when $P_0 = P_1 \,{}_{\mathscr{O}}\!+ P_2$, equation (13) holds trivially. Hence, we only need to verify the case which $P_0 = |\psi_0\rangle\langle\psi_0| \,{}_{\mathscr{O}}\!+ |\psi_2\rangle\langle\psi_2|$ and $P_1 = |\psi_1\rangle\langle\psi_1| \,{}_{\mathscr{O}}\!+ |\psi_2\rangle\langle\psi_2|$, where $\{|\psi_0\rangle, |\psi_1\rangle, |\psi_2\rangle\}$ is an orthonormal basis. Because we have

$$\begin{aligned} \bar{\mu}(P_0 \,{}_{\mathscr{O}}\!\vee P_1) &= \bar{\mu}(\mathbb{1}) = \mathit{certain} \\ \bar{\mu}(P_0 P_1) &= \bar{\mu}(|\psi_2\rangle\langle\psi_2|) \subseteq \mathit{unknown} \end{aligned}$$

and $\bar{\mu}(P_0) \,{}_{\mathscr{I}}\!+ \bar{\mu}(P_1)$ is either $\mathit{certain} \,{}_{\mathscr{I}}\!+ \mathit{unknown}$ or $\mathit{unknown} \,{}_{\mathscr{I}}\!+ \mathit{unknown}$. Therefore, we have $\bar{\mu}(P_0 \,{}_{\mathscr{O}}\!\vee P_1) \,{}_{\mathscr{I}}\!+ \bar{\mu}(P_0 P_1) \subseteq \bar{\mu}(P_0) \,{}_{\mathscr{I}}\!+ \bar{\mu}(P_1)$.

Finally, we will prove that $\bar{\mu}$ cannot correspond to any quantum probability measure by contradiction. Suppose $\mu(P) \in \bar{\mu}(P)$ for some quantum probability measure $\mu : \mathcal{E} \to [0,1]$. We must have

$$\mu(|0\rangle\langle 0|) = \mu(|+\rangle\langle +|) = \mu(|+'\rangle\langle +'|) = 0 \ . \tag{14}$$

By Gleason's theorem, there is a mixed state $\rho = {}_{\mathscr{O}}\!\sum_{j=0}^{N-1} q_j |\phi_j\rangle\langle\phi_j|$ such that $\mu(P) = {}_{\mathbb{R}}\!\sum_{j=0}^{N-1} q_j \langle\phi_j|P\phi_j\rangle$, where ${}_{\mathbb{R}}\!\sum_{j=0}^{N-1} q_j = 1$ and $q_j > 0$. However, no pure state $|\phi\rangle$ satisfies $\langle\phi|0\rangle = \langle\phi|+\rangle = \langle\phi|+'\rangle = 0$ so that we have $\mu(P) \notin \bar{\mu}(P)$ for all quantum probability measure μ. □

Recall that when approximating the probability classically, the quality of the experimental equipment cannot be compensated by the number of independent trials. If the precision of the intervals $\mathscr{I}$ represents the quality of the quantum experimental equipment, example 6 tells us that we cannot identify a particular state if the experimental equipment reveals too little information. For example, $\bar{\mu}$ might correspond to $|2\rangle$ because $\bar{\mu}(|0\rangle\langle 0|) = \bar{\mu}(|+\rangle\langle +|) = \mathit{impossible}$, $|1\rangle$ because $\bar{\mu}(|0\rangle\langle 0|) = \bar{\mu}(|+'\rangle\langle +'|) = \mathit{impossible}$, or the density matrix $\frac{\mathbb{1}}{3}$ because $\bar{\mu}(|\phi\rangle\langle\phi|) \neq \mathit{certain}$ for all $|\phi\rangle$. Also, example 6 is no longer valid if our measurement equipment is a little more precise, say $\mathscr{I} = \{\mathit{impossible}, \mathit{unlikely}, \mathit{likely}, \mathit{certain}\}$, because example 6 really use the property of $\mathit{unknown}$.

If we consider $\mathscr{I} = \{\mathit{impossible}, \mathit{unlikely}, \mathit{likely}, \mathit{certain}\}$, example 7 in appendix A provides another example with an empty core. Comparing to example 6, example 7 is closer to the Born rule in a sense that it might only correspond to $|2\rangle$ or $\frac{\mathbb{1}}{3}$, but not $|1\rangle$. Example 7 is also no longer valid for more precise measurement equipment as well. In general, we believe if the measurement equipment is more and more precise, its interval-valued probability measure will be closer and closer to the the Born rule. In the limit case, $\mathscr{I} = \{\{a\} | a \in [0, 1]\}$ recovers conventional Gleason's theorem and the Born rule.
[11]

References

[1] Andrew J Hanson, Gerardo Ortiz, Amr Sabry, and Yu-Tsung Tai. "Discrete quantum theories". In: *Journal of Physics A: Mathematical and Theoretical* 47.11 (2014), p. 115305. URL: http://stacks.iop.org/1751-8121/47/i=11/a=115305.

[2] Andrew J Hanson, Gerardo Ortiz, Amr Sabry, and Yu-Tsung Tai. "Geometry of discrete quantum computing". In: *Journal of Physics A: Mathematical and Theoretical* 46.18 (2013), p. 185301. DOI: 10.1088/1751-8113/46/18/185301. Erratum "Corrigendum: Geometry of discrete quantum computing". In: *Journal of Physics A: Mathematical and Theoretical* 49.3 (2015), p. 039501.

[3] Madan L Puri and Dan A Ralescu. "Strong law of large numbers with respect to a set-valued probability measure". In: *The Annals of Probability* (1983), pp. 1051–1054.

[4] 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.

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

[6] Andrew Gleason. "Measures on the Closed Subspaces of a Hilbert Space". In: *Indiana Univ. Math. J.* 6 (4 1957), pp. 885–893. ISSN: 0022-2518.

[7] Asher Peres. *Quantum Theory: Concepts and Methods*. Fundamental Theories of Physics. Springer, 1995. ISBN: 9780792336327. URL: http://books.google.com/books?id=rMGqMyFBcL8C.

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

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

[10] Glenn Shafer. *A mathematical theory of evidence*. Princeton University Press, 1976.

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

[12] Jan Swart. "Introduction to Quantum Probability". Lecture Notes for the course NSTP187 at Charles University, Prague, Summer 2013. June 1, 2013. URL: http://staff.utia.cas.cz/swart/dict13_6.pdf.

[13] Otton Nikodym. "Sur une généralisation des intégrales de M. J. Radon". In: *Fundamenta Mathematicae* 15.1 (1930), pp. 131–179.

[11] Yu-Tsung says: Notice that Gleason used his conditions to prove quantum probability measures are continuous first, and then prove they can be got from the Born rule. Since the current examples are all discontinuous, maybe "add continuous" could give us a interval-valued Gleason's theorem (?)
Continuity definitely has physical and computer science meaning. For example, we could said something like, "no matter the underlining physical reality, human doesn't have the ability to distinguish discontinuous function," and cite [44, 45]. However, we might need to rewrite all the interpretation if this approach works...

[14] Gerald B Folland. *Real analysis: modern techniques and their applications*. 2nd ed. Pure and applied mathematics. John Wiley & Sons, 1999. ISBN: 0-471-31716-0.

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

[16] Kenneth David Jamison and Weldon A Lodwick. *Interval-Valued Probability Measures*. 213. Center for Computational Mathematics, University of Colorado Denver, 2004. URL: http://ccm.ucdenver.edu/reports/rep213.pdf.

[17] Jacob Bernoulli. *The Art of Conjecturing*. English. Trans. and comm., with a forew., by Edith Dudley Sylla. Baltimore: Johns Hopkins University Press, 2006.

[18] James Victor Uspensky. *Introduction to mathematical probability*. McGraw-Hill Book Co., Inc., 1937.

[19] Michael A. Nielsen and Isaac L. Chuang. *Quantum computation and quantum information*. New York, NY, USA: Cambridge University Press, 2000. ISBN: 0-521-63503-9.

[20] Comte de Buffon Georges Louis Leclerc. "Essai d'Arithmétique Morale". In: *Histoire naturelle, générale et particulière, Supplément*. Vol. 4. Paris: De l'Imprimerie royale, an 12, 1777, pp. 46–109.

[21] Augustus De Morgan. *A Budget of Paradoxes*. Longmans, Green, and Co., 1872.

[22] Asaph Hall. "On an experimental determination of π". In: *Messeng. Math* 2 (1873), pp. 113–114.

[23] Massimo Marinacci. "Limit laws for non-additive probabilities and their frequentist interpretation". In: *Journal of Economic Theory* 84.2 (1999), pp. 145–195.

[24] Pedro Terán. "Laws of large numbers without additivity". In: *Transactions of the American Mathematical Society* 366.10 (2014), pp. 5431–5451.

[25] Gustave Choquet. "Theory of capacities". In: *Annales de l'institut Fourier*. Vol. 5. Institut Fourier. 1954, pp. 131–295.

[26] Siegfried Graf. "A Radon-Nikodym theorem for capacities". In: *Journal für die reine und angewandte Mathematik* 320 (1980), pp. 192–214.

[27] Irwin R Goodman, Ronald P Mahler, and Hung T Nguyen. *Mathematics of data fusion*. Theory and Decision Library 37. Kluwer Academic Publishers, 1997. ISBN: 0-7923-4674-2.

[28] Lloyd S. Shapley. "Cores of convex games". In: *International Journal of Game Theory* 1.1 (1971), pp. 11–26. ISSN: 1432-1270. DOI: 10.1007/BF01753431. URL: http://dx.doi.org/10.1007/BF01753431.

[29] Michel Grabisch. *Set functions, games and capacities in decision making*. Theory and Decision Library C 46. Springer International Publishing, 2016. ISBN: 978-3-319-30688-9. DOI: 10.1007/978-3-319-30690-2. URL: http://www.springer.com/us/book/9783319306889.

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

[31] Arthur P Dempster. "Upper and lower probabilities induced by a multivalued mapping". In: *The Annals of Mathematical Statistics* (1967), pp. 325–339.

[32] Lloyd S Shapley. "On balanced sets and cores". In: *Naval research logistics quarterly* 14.4 (1967), pp. 453–460.

[33] Olga N Bondareva. "Several applications of linear programming methods to the theory of cooperative games". English. In: *Selected Russian Papers on Game Theory. 1959 - 1965*. Trans. from the Russian by Eugene Wesley. Princeton, New Jersey: Econometric Research Program, 1968, pp. 79–114. URL: http://www.princeton.edu/~erp/ERParchives/archivepdfs/WP0.pdf.

[34] Bezalel Peleg. "An inductive method for constructing mimmal balanced collections of finite sets". In: *Naval Research Logistics Quarterly* 12.2 (1965), pp. 155–162.

[35] Jingang Zhao. “The maximal payoff and coalition formation in coalitional games”. In: 13th Coalition Theory Network Workshop. (Venice, Italy, Jan. 24, 2008). Fondazione Eni Enrico Mattei (FEEM). 2008. URL: http://www.feem.it/getpage.aspx?id=1785&sez=Publications&padre=73.

[36] George W. Mackey. “Quantum Mechanics and Hilbert Space”. In: *The American Mathematical Monthly* 64.8 (1957), pp. 45–57. ISSN: 00029890, 19300972. URL: http://www.jstor.org/stable/2308516.

[37] Garrett Birkhoff and John Von Neumann. “The logic of quantum mechanics”. In: *Annals of mathematics* (1936), pp. 823–843.

[38] Samson Abramsky. “Big Toy Models: Representing Physical Systems as Chu Spaces”. In: *CoRR* abs/0910.2393 (2009).

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

[40] Gregg Jaeger. *Quantum Information.* Springer New York, Apr. 3, 2007. ISBN: 978-0-387-36944-0. URL: http://www.ebook.de/de/product/24986917/gregg_jaeger_quantum_information.html.

[41] Otto Stern. “A way towards the experimental examination of spatial quantisation in a magnetic field”. In: *Zeitschrift für Physik D Atoms, Molecules and Clusters* 10.2 (1988), pp. 114–116.

[42] Samuel S Holland Jr. “The current interest in orthomodular lattices”. In: *Trends in lattice theory.* Ed. by James Crawford Abbott. Van Nostrand Reinhold Company, 1970, pp. 41–126.

[43] David Meyer. “Finite Precision Measurement Nullifies the Kochen-Specker Theorem”. In: *Phys. Rev. Lett.* 83 (19 Nov. 1999), pp. 3751–3754. DOI: 10.1103/PhysRevLett.83.3751. URL: http://link.aps.org/doi/10.1103/PhysRevLett.83.3751.

[44] Martin Ziegler. “Real Computability and Hypercomputation”. habilitation. Dec. 2007. URL: http://www2.cs.uni-paderborn.de/cs/ag-madh/WWW/ziegler/DFG/Habilitation_Ziegler.pdf.

[45] K. Weihrauch. *Computable Analysis: An Introduction.* Texts in Theoretical Computer Science. An EATCS Series. Springer Berlin Heidelberg, 2012. ISBN: 9783642569999. URL: https://books.google.com.tw/books?id=Ir6qCAAAQBAJ.

[46] Roger Cooke, Michael Keane, and William Moran. “An elementary proof of Gleason’s theorem”. In: *Mathematical Proceedings of the Cambridge Philosophical Society* 98.1 (July 1985), pp. 117–128. DOI: 10.1017/S0305004100063313. URL: https://www.cambridge.org/core/article/an-elementary-proof-of-gleason-s-theorem/89A7AB2F467BD015B2115B637867FB32.

[47] Fred Richman and Douglas Bridges. “A Constructive Proof of Gleason’s Theorem”. In: *Journal of Functional Analysis* 162.2 (1999), pp. 287–312. ISSN: 0022-1236. DOI: http://dx.doi.org/10.1006/jfan.1998.3372. URL: http://www.sciencedirect.com/science/article/pii/S0022123698933729.

[48] Jan Hamhalter. *Quantum measure theory.* Vol. 134. Springer Science & Business Media, 2013.

[49] Allen Hatcher. *Algebraic Topology.* Cambridge University Pr., 2001. 556 pp. ISBN: 0521795400. URL: http://www.ebook.de/de/product/3255852/allen_hatcher_algebraic_topology.html.

[50] Ian L. Dryden. “Statistical analysis on high-dimensional spheres and shape spaces”. In: *Ann. Statist.* 33.4 (Aug. 2005), pp. 1643–1665. DOI: 10.1214/009053605000000264. URL: http://dx.doi.org/10.1214/009053605000000264.

[51] THEODORE GAMELIN. *Complex Analysis.* Springer, July 17, 2003. 500 pp. ISBN: 0387950699. URL: http://www.ebook.de/de/product/3654178/theodore_gamelin_complex_analysis.html.

[52] Richard L. Wheeden and Antoni Zygmund. *Measure and Integral: An Introduction to Real Analysis.* Monographs and textbooks in Pure and Applied Mathematics. New York and Basel: Marcel Dekker, Inc., 1977. ISBN: 9780824764999. URL: https://books.google.com/books?id=YDkDmQ%5C_hdmcC.

A Examples for Quantum Interval-valued Probability Measures

We are going to provide an quantum 4-value interval-valued probability measures corresponding to no quantum probability measure, i.e., for all mixed state ρ there is a $P \in \mathcal{E}$ such that $\mu^B_\rho(P) \notin \bar{\mu}(P)$. Before the example, we need a notation of cross product.

Definition 8. For $|\psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle + \alpha_2|2\rangle = \begin{pmatrix}\alpha_0\\ \alpha_1\\ \alpha_2\end{pmatrix}$ and $|\phi\rangle = \beta_0|0\rangle + \beta_1|1\rangle + \beta_2|2\rangle = \begin{pmatrix}\beta_0\\ \beta_1\\ \beta_2\end{pmatrix}$, their cross product is defined as follow:

$$|\psi \times \phi\rangle = |\psi\rangle \times |\phi\rangle = \begin{vmatrix}\alpha_1^* & \beta_1^*\\ \alpha_2^* & \beta_2^*\end{vmatrix}|0\rangle + \begin{vmatrix}\alpha_2^* & \beta_2^*\\ \alpha_0^* & \beta_0^*\end{vmatrix}|1\rangle + \begin{vmatrix}\alpha_0^* & \beta_0^*\\ \alpha_1^* & \beta_1^*\end{vmatrix}|2\rangle \ ,$$

where $\begin{vmatrix}\alpha & \beta\\ \alpha' & \beta'\end{vmatrix} = \alpha\beta' - \alpha'\beta$. □

Notice that we put the complex conjugate in the cross product, so that the result of cross product is orthogonal to $|\psi\rangle$ and $|\phi\rangle$.

$$\begin{aligned}\langle\psi \times \phi|\psi\rangle &= (\alpha_1\beta_2 - \alpha_2\beta_1)\,\alpha_0 + (\alpha_2\beta_0 - \alpha_0\beta_2)\,\alpha_1 + (\alpha_0\beta_1 - \alpha_1\beta_0)\,\alpha_2 = 0 \ ,\\ \langle\psi \times \phi|\phi\rangle &= (\alpha_1\beta_2 - \alpha_2\beta_1)\,\beta_0 + (\alpha_2\beta_0 - \alpha_0\beta_2)\,\beta_1 + (\alpha_0\beta_1 - \alpha_1\beta_0)\,\beta_2 = 0 \ .\end{aligned}$$

Also, even if $|\psi\rangle$ and $|\phi\rangle$ are normalized, their cross product $|\psi\rangle \times |\phi\rangle$ need not be normalized as usual.

Example 7 (Three-dimensional quantum 4-value interval-valued probability measure). [12]Given a three dimensional Hilbert space with an orthonormal basis $\{|0\rangle, |1\rangle, |2\rangle\}$. Consider $\mathscr{I} = \{\mathit{impossible}, \mathit{unlikely}, \mathit{likely}, \mathit{certain}\}$, $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, and a quantum $\mathscr{I}$-interval-valued probability measure $\bar{\mu} : \mathcal{E} \to \mathscr{I}$ defined as follow, where the state vectors corresponding to 1-dimensional projector are also illustrated in figure 1.

1. Let

$$\begin{aligned}\bar{\mu}(\mathbb{0}) = \bar{\mu}(|0\rangle\langle 0|) = \bar{\mu}(|+\rangle\langle +|) &= \mathit{impossible}\\ \bar{\mu}(\mathbb{1}) = \bar{\mu}(\mathbb{1}_{\mathscr{O}} - |0\rangle\langle 0|) = \bar{\mu}(\mathbb{1}_{\mathscr{O}} - |+\rangle\langle +|) &= \mathit{certain}\ ,\end{aligned}$$

   where $|0\rangle$ and $|+\rangle$ is illustrated as the red and green dotted vector, respectively.

2. Consider the red and green circles in figure 1. They are the states orthogonal to $|0\rangle$ and $|+\rangle$, respectively, and can be parametrized as $|0^\perp_{\theta,\gamma}\rangle = e^{i\gamma}\sin\theta|1\rangle + \cos\theta|2\rangle$ and $|+^\perp_{\theta,\gamma}\rangle = -e^{i\gamma}\sin\theta|-\rangle + \cos\theta|2\rangle$, where $0 \le \theta \le \frac{\pi}{2}$ and $0 \le \gamma < 2\pi$. The dotted half of those states need special treatment, i.e., whenever $0 \le \theta < \frac{\pi}{2}$ and $0 \le \gamma < \pi$, we define

$$\begin{aligned}\bar{\mu}\left(|0^\perp_{\theta,\gamma}\rangle\langle 0^\perp_{\theta,\gamma}|\right) = \bar{\mu}\left(|+^\perp_{\theta,\gamma}\rangle\langle +^\perp_{\theta,\gamma}|\right) &= \mathit{likely}\\ \bar{\mu}\left(\mathbb{1}_{\mathscr{O}} - |0^\perp_{\theta,\gamma}\rangle\langle 0^\perp_{\theta,\gamma}|\right) = \bar{\mu}\left(\mathbb{1}_{\mathscr{O}} - |+^\perp_{\theta,\gamma}\rangle\langle +^\perp_{\theta,\gamma}|\right) &= \mathit{unlikely}\ .\end{aligned}$$

   [13]

3. Otherwise, $\bar{\mu}(|\psi\rangle\langle\psi|) = \mathit{unlikely}$ and $\bar{\mu}(\mathbb{1}_{\mathscr{O}} - |\psi\rangle\langle\psi|) = \mathit{likely}$.

[12] Yu-Tsung says: Need to verify whether this example is convex or not.

[13] Yu-Tsung says: This case might be simplify if we consider 5 values. Then, every state orthogonal to $|0\rangle$ and $|+\rangle$ can map to $\left[\frac{1}{4}, \frac{3}{4}\right]$. Not sure whether we need to simplify the example or not, because this example is used in example 9. Or we may just keep example 9 in the later version if that example is more important...

That $\bar{\mu}$ has an empty core can be proved by contradiction. Assume there is a real-valued probability measure satisfying $\mu_\rho^B(P) \in \bar{\mu}(P)$ for all $P \in \mathcal{E}$. Because $\mu_\rho^B(|0\rangle\langle 0|) \in \bar{\mu}(|0\rangle\langle 0|) = \mathit{impossible}$ and $\mu_\rho^B(|+\rangle\langle +|) \in \bar{\mu}(|+\rangle\langle +|) = \mathit{impossible}$, we must have $\mu_\rho^B(|0\rangle\langle 0|) = \mu_\rho^B(|+\rangle\langle +|) = 0$ so that $\mu_\rho^B = \mu_{|2\rangle}^B$. However,

$$\mu_{|2\rangle}^B(|2\rangle\langle 2|) = 1 \notin \mathit{unlikely} = \bar{\mu}(|2\rangle\langle 2|) \ .$$

Similar to example 6, to verify $\bar{\mu}$ is a quantum interval-valued probability measure, we only need to verify equation (12), and it is sufficient to check

$$\begin{aligned}
\bar{\mu}\left(\mathbb{1} \mathbin{{}_{\mathscr{O}}-} |\psi_0\rangle\langle\psi_0|\right) &\subseteq \bar{\mu}\left(|\psi_1\rangle\langle\psi_1|\right) \mathbin{{}_{\mathscr{I}}+} \bar{\mu}\left(|\psi_2\rangle\langle\psi_2|\right) \\
\bar{\mu}\left(\mathbb{1} \mathbin{{}_{\mathscr{O}}-} |\psi_1\rangle\langle\psi_1|\right) &\subseteq \bar{\mu}\left(|\psi_2\rangle\langle\psi_2|\right) \mathbin{{}_{\mathscr{I}}+} \bar{\mu}\left(|\psi_0\rangle\langle\psi_0|\right) \\
\bar{\mu}\left(\mathbb{1} \mathbin{{}_{\mathscr{O}}-} |\psi_2\rangle\langle\psi_2|\right) &\subseteq \bar{\mu}\left(|\psi_0\rangle\langle\psi_0|\right) \mathbin{{}_{\mathscr{I}}+} \bar{\mu}\left(|\psi_1\rangle\langle\psi_1|\right)
\end{aligned} \tag{15}$$

for every orthonormal basis $\mathcal{B} = \{|\psi_0\rangle, |\psi_1\rangle, |\psi_2\rangle\}$. We are going to enumerate all possible orthonormal bases to verify equation (15), and case 1 and 2 are illustrated in figure 1.

1. When $|\psi_0\rangle$ is $|0\rangle$, then either $|\psi_1\rangle$ or $|\psi_2\rangle$ must be $\left|0_{\theta,\gamma}^\perp\right\rangle$ for some θ and γ. Because of

$$|0\rangle \times \left|0_{\theta,\gamma}^\perp\right\rangle = \begin{pmatrix}1\\0\\0\end{pmatrix} \times \begin{pmatrix}0\\ \mathrm{e}^{\mathrm{i}\gamma}\sin\theta \\ \cos\theta\end{pmatrix} = \begin{pmatrix}0\\ -\cos\theta \\ \mathrm{e}^{-\mathrm{i}\gamma}\sin\theta\end{pmatrix} = \mathrm{e}^{-\mathrm{i}\gamma}\begin{pmatrix}0\\ \mathrm{e}^{\mathrm{i}(\pi+\gamma)}\sin\left(\frac{\pi}{2}-\theta\right) \\ \cos\left(\frac{\pi}{2}-\theta\right)\end{pmatrix} = \mathrm{e}^{-\mathrm{i}\gamma}\left|0_{\frac{\pi}{2}-\theta,\pi+\gamma}^\perp\right\rangle$$

the pair $|\psi_1\rangle = \left|0_{\theta,\gamma}^\perp\right\rangle$ and $|\psi_2\rangle = \left|0_{\frac{\pi}{2}-\theta,\pi+\gamma}^\perp\right\rangle$ for $(0 < \theta < \frac{\pi}{2}$ and $0 \le \gamma < \pi)$ or $\theta = 0$ enumerate all possible orthonormal bases. Equation (15) can then be verified as follow.

$$\begin{aligned}
\bar{\mu}\left(\mathbb{1} \mathbin{{}_{\mathscr{O}}-} |0\rangle\langle 0|\right) = \mathit{certain} \subseteq \mathit{likely} \mathbin{{}_{\mathscr{I}}+} \mathit{unlikely} &= \bar{\mu}\left(\left|0_{\theta,\gamma}^\perp\right\rangle\left\langle 0_{\theta,\gamma}^\perp\right|\right) \mathbin{{}_{\mathscr{I}}+} \bar{\mu}\left(\left|0_{\frac{\pi}{2}-\theta,\pi+\gamma}^\perp\right\rangle\left\langle 0_{\frac{\pi}{2}-\theta,\pi+\gamma}^\perp\right|\right) \\
\bar{\mu}\left(\mathbb{1} \mathbin{{}_{\mathscr{O}}-} \left|0_{\theta,\gamma}^\perp\right\rangle\left\langle 0_{\theta,\gamma}^\perp\right|\right) = \mathit{unlikely} = \mathit{unlikely} \mathbin{{}_{\mathscr{I}}+} \mathit{impossible} &= \bar{\mu}\left(\left|0_{\frac{\pi}{2}-\theta,\pi+\gamma}^\perp\right\rangle\left\langle 0_{\frac{\pi}{2}-\theta,\pi+\gamma}^\perp\right|\right) \mathbin{{}_{\mathscr{I}}+} \bar{\mu}\left(|0\rangle\langle 0|\right) \\
\bar{\mu}\left(\mathbb{1} \mathbin{{}_{\mathscr{O}}-} \left|0_{\frac{\pi}{2}-\theta,\pi+\gamma}^\perp\right\rangle\left\langle 0_{\frac{\pi}{2}-\theta,\pi+\gamma}^\perp\right|\right) = \mathit{likely} = \mathit{impossible} \mathbin{{}_{\mathscr{I}}+} \mathit{likely} &= \bar{\mu}\left(|0\rangle\langle 0|\right) \mathbin{{}_{\mathscr{I}}+} \bar{\mu}\left(\left|0_{\theta,\gamma}^\perp\right\rangle\left\langle 0_{\theta,\gamma}^\perp\right|\right)
\end{aligned} \tag{16}$$

Similarly, when $|\psi_0\rangle$ is $|+\rangle$, equation (15) holds.

2. When $|0\rangle \notin \mathcal{B}$ and $|\psi_0\rangle = \left|0_{\theta,\gamma}^\perp\right\rangle$ for $0 \le \theta < \frac{\pi}{2}$ and $0 \le \gamma < \pi$, we want to prove $\bar{\mu}\left(|\psi_1\rangle\langle\psi_1|\right) = \bar{\mu}\left(|\psi_2\rangle\langle\psi_2|\right) = \mathit{unlikely}$. Then, we have

$$\begin{aligned}
\bar{\mu}\left(\mathbb{1} \mathbin{{}_{\mathscr{O}}-} \left|0_{\theta,\gamma}^\perp\right\rangle\left\langle 0_{\theta,\gamma}^\perp\right|\right) = \mathit{unlikely} \subseteq \mathit{unlikely} \mathbin{{}_{\mathscr{I}}+} \mathit{unlikely} &= \bar{\mu}\left(|\psi_1\rangle\langle\psi_1|\right) \mathbin{{}_{\mathscr{I}}+} \bar{\mu}\left(|\psi_2\rangle\langle\psi_2|\right) \\
\bar{\mu}\left(\mathbb{1} \mathbin{{}_{\mathscr{O}}-} |\psi_1\rangle\langle\psi_1|\right) = \mathit{likely} \subseteq \mathit{unlikely} \mathbin{{}_{\mathscr{I}}+} \mathit{likely} &= \bar{\mu}\left(|\psi_2\rangle\langle\psi_2|\right) \mathbin{{}_{\mathscr{I}}+} \bar{\mu}\left(\left|0_{\theta,\gamma}^\perp\right\rangle\left\langle 0_{\theta,\gamma}^\perp\right|\right)
\end{aligned} \tag{17}$$

In order to verify $\bar{\mu}\left(|\psi_1\rangle\langle\psi_1|\right) = \bar{\mu}\left(|\psi_2\rangle\langle\psi_2|\right) = \mathit{unlikely}$, it is sufficient to prove that if $\left|+_{\theta',\gamma'}^\perp\right\rangle \in \mathcal{B}$, then $(0 < \theta' < \frac{\pi}{2}$ and $\pi \le \gamma' < 2\pi)$ or $\theta' = \frac{\pi}{2}$. Recall $\left\langle + \middle| +_{\theta',\gamma'}^\perp\right\rangle = 0$. Hence, if $\left|+_{\theta',\gamma'}^\perp\right\rangle \in \mathcal{B}$, we have $\left|+_{\theta',\gamma'}^\perp\right\rangle$ parallel to $|+\rangle \times \left|0_{\theta,\gamma}^\perp\right\rangle$.

$$|+\rangle \times \left|0_{\theta,\gamma}^\perp\right\rangle = \begin{pmatrix}1\\1\\0\end{pmatrix} \times \begin{pmatrix}0\\ \mathrm{e}^{\mathrm{i}\gamma}\sin\theta \\ \cos\theta\end{pmatrix} = \begin{pmatrix}\cos\theta\\ -\cos\theta \\ \mathrm{e}^{-\mathrm{i}\gamma}\sin\theta\end{pmatrix} = \mathrm{e}^{-\mathrm{i}\gamma}\begin{pmatrix}-\mathrm{e}^{\mathrm{i}(\pi+\gamma)}\sin\left(\frac{\pi}{2}-\theta\right)\\ \mathrm{e}^{\mathrm{i}(\pi+\gamma)}\sin\left(\frac{\pi}{2}-\theta\right) \\ \cos\left(\frac{\pi}{2}-\theta\right)\end{pmatrix} = \sqrt{2}\mathrm{e}^{-\mathrm{i}\gamma}\left|+_{\frac{\pi}{2}-\theta,\pi+\gamma}^\perp\right\rangle \tag{18}$$

Because of $0 \le \theta < \frac{\pi}{2}$ and $0 \le \gamma < \pi$, we have $(0 < \theta' = \frac{\pi}{2} - \theta < \frac{\pi}{2}$ and $\pi \le \gamma' = \pi + \gamma < 2\pi)$ or $\theta' = \frac{\pi}{2} - \theta = \frac{\pi}{2}$. Similarly, when $|\psi_0\rangle$ is $\left|+_{\theta,\gamma}^\perp\right\rangle$ and $|+\rangle \notin \mathcal{B}$, equation (15) holds.

3. Finally, when $|0\rangle \notin \mathcal{B}$, $|+\rangle \notin \mathcal{B}$, $\left|0_{\theta,\gamma}^\perp\right\rangle \notin \mathcal{B}$, and $\left|+_{\theta,\gamma}^\perp\right\rangle \notin \mathcal{B}$ for $0 \le \theta < \frac{\pi}{2}$ and $0 \le \gamma < \pi$, i.e., the "otherwise" case. Then, equation (15) can easily be verified.

$$\bar{\mu}\left(\mathbb{1} \mathbin{{}_{\mathscr{O}}-} |\psi_0\rangle\langle\psi_0|\right) = \mathit{likely} \subseteq \mathit{unlikely} \mathbin{{}_{\mathscr{I}}+} \mathit{unlikely} = \bar{\mu}\left(|\psi_1\rangle\langle\psi_1|\right) \mathbin{{}_{\mathscr{I}}+} \bar{\mu}\left(|\psi_2\rangle\langle\psi_2|\right) \tag{19}$$

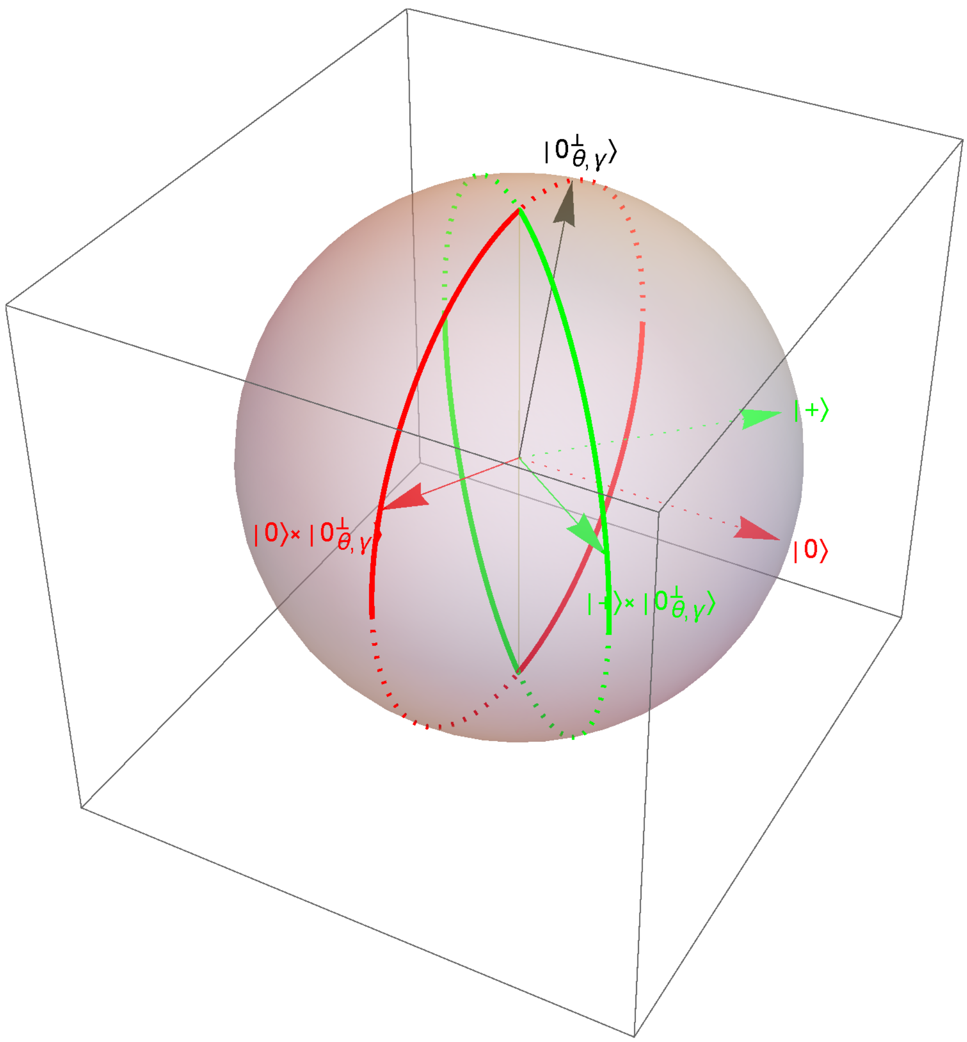

Figure 1: This figure illustrates case 1 and 2 in $\mathbb{R}^3$ in example 7. The red and green dotted vector are $|0\rangle$ and $|+\rangle$ respectively. All possible vectors of $\left|0^{\perp}_{\theta,\gamma}\right\rangle$ and $\left|+^{\perp}_{\theta,\gamma}\right\rangle$ are drawn in the red and green circles, respectively. Within the circles, a vector $|\psi\rangle$ in dotted part means $\bar{\mu}(|\psi\rangle\langle\psi|) = \mathit{likely}$; otherwise, $\bar{\mu}(|\psi\rangle\langle\psi|) = \mathit{unlikely}$. The gray vector is a generic vector $\left|0^{\perp}_{\theta,\gamma}\right\rangle$, and the red and green solid vectors are normalized $|0\rangle \times \left|0^{\perp}_{\theta,\gamma}\right\rangle$ and $|+\rangle \times \left|0^{\perp}_{\theta,\gamma}\right\rangle$, respectively.

[14] $\square$

The following example verifies example 7 cannot hold with more precise intervals.

Example 8 (Three-dimensional quantum 4-value interval-valued probability measure (Continue)). Consider $\mathscr{I} = \left\{\mathit{impossible}, [l_1, r_1], [1 -_{\mathbb{R}} r_1, 1 -_{\mathbb{R}} l_1], \mathit{certain}\right\}$, where $[l_1, r_1]$ and $[1 -_{\mathbb{R}} r_1, 1 -_{\mathbb{R}} l_1]$ will be used to replace every $\mathit{unlikely}$ and likely in example 7, respectively. The bound of l_1 and r_1 can then be estimated by the equations corresponding to equation (16), (17), and (19).

When we replace $\mathit{unlikely}$ and likely by $[l_1, r_1]$ and $[1 -_{\mathbb{R}} r_1, 1 -_{\mathbb{R}} l_1]$, equation (16) will become

$$\begin{aligned} \mathit{certain} &\subseteq [1 -_{\mathbb{R}} r_1, 1 -_{\mathbb{R}} l_1] +_{\mathscr{I}} [l_1, r_1] \\ [l_1, r_1] &\subseteq [l_1, r_1] +_{\mathscr{I}} \mathit{impossible} \\ [1 -_{\mathbb{R}} r_1, 1 -_{\mathbb{R}} l_1] &\subseteq \mathit{impossible} +_{\mathscr{I}} [1 -_{\mathbb{R}} r_1, 1 -_{\mathbb{R}} l_1] \end{aligned} \tag{20}$$

which are all tautologies. Equation (17) will become

$$\begin{aligned} [l_1, r_1] &\subseteq [l_1, r_1] +_{\mathscr{I}} [l_1, r_1] \\ [1 -_{\mathbb{R}} r_1, 1 -_{\mathbb{R}} l_1] &\subseteq [l_1, r_1] +_{\mathscr{I}} [1 -_{\mathbb{R}} r_1, 1 -_{\mathbb{R}} l_1] \end{aligned} \tag{21}$$

which implies $l_1 = 0$. Finally, equation (19) will become

$$[1 -_{\mathbb{R}} r_1, 1 -_{\mathbb{R}} l_1] \subseteq [l_1, r_1] +_{\mathscr{I}} [l_1, r_1] \tag{22}$$

which implies $\frac{1}{2} \le r_1 \le 1$. Therefore, $l_1 = 0$ and $r_1 = \frac{1}{2}$ provides the most precise intervals which make example 7 work, and in this case, the intervals are $\left\{\mathit{impossible}, \mathit{unlikely}, \mathit{likely}, \mathit{certain}\right\}$. $\square$

Although it is easier to deduce an interval-valued probability measure has an empty core by considering the state mapped to $\mathit{impossible}$ as in example 6 and 7, it is not the only way to prove an interval-valued probability measure has an empty core. The following family of interval-valued probability measures satisfy that $\bar{\mu}(P) = \mathit{impossible}$ if and only if $P = \mathbb{0}$, but we can still prove them have empty cores.

Although it is easier to deduce an interval-valued probability measure has an empty core by considering the state mapped to $\mathit{impossible}$ as in example 6 and 7, it is not the only way to prove an interval-valued probability measure has an empty core. The following family of interval-valued probability measures satisfy that $\bar{\mu}(P) = \mathit{impossible}$ if and only if $P = \mathbb{0}$, but we can still prove them have empty cores.

Example 9 (Another three-dimensional quantum 6-value interval-valued probability measure). Given $r_0 \in \left[\frac{1}{3}, \frac{1}{2}\right]$, we are going to construct an interval-valued probability measures for each r_0, and it has an empty core whenever $r_0 \in \left(\frac{5_{\mathbb{R}}+2\sqrt{2}}{17}, \frac{1}{2}\right]$.

Consider $\mathscr{I} = \left\{\mathit{impossible}, \mathit{certain}, \mathit{unlikely}_0, \mathit{likely}_0, \mathit{unlikely}_1, \mathit{likely}_1\right\}$, where $\mathit{unlikely}_0 = [1 -_{\mathbb{R}} 2r_0, r_0]$, $\mathit{likely}_0 = [1 -_{\mathbb{R}} r_0, 2r_0]$, $\mathit{unlikely}_1 = [2 -_{\mathbb{R}} 4r_0, 1 -_{\mathbb{R}} r_0]$, and $\mathit{likely}_1 = [r_0, 4r_0 -_{\mathbb{R}} 1]$. This is well-defined since the left-end is less or equal to the right-end whenever $r_0 \ge \frac{1}{3}$. If $|0\rangle$, $|1\rangle$, $|2\rangle$, $|+\rangle$, $|-\rangle$, $\left|0^{\perp}_{\theta,\gamma}\right\rangle$, and $\left|+^{\perp}_{\theta,\gamma}\right\rangle$ are defined as in example 7, a quantum $\mathscr{I}$-interval-valued probability measure $\bar{\mu} : \mathcal{E} \to \mathscr{I}$ can be defined as follow:

1. $\bar{\mu}(\mathbb{0}) = \mathit{impossible}$ and $\bar{\mu}(\mathbb{1}) = \mathit{certain}$.
2. For the dotted half of the red and green circles in figure 1, i.e., whenever $0 \le \theta < \frac{\pi}{2}$ and $0 \le \gamma < \pi$, we define

$$\begin{aligned} \bar{\mu}\left(\left|0^{\perp}_{\theta,\gamma}\right\rangle\left\langle 0^{\perp}_{\theta,\gamma}\right|\right) = \bar{\mu}\left(\left|+^{\perp}_{\theta,\gamma}\right\rangle\left\langle +^{\perp}_{\theta,\gamma}\right|\right) &= \mathit{likely}_1 \\ \bar{\mu}\left(\mathbb{1} -_{\mathcal{O}} \left|0^{\perp}_{\theta,\gamma}\right\rangle\left\langle 0^{\perp}_{\theta,\gamma}\right|\right) = \bar{\mu}\left(\mathbb{1} -_{\mathcal{O}} \left|+^{\perp}_{\theta,\gamma}\right\rangle\left\langle +^{\perp}_{\theta,\gamma}\right|\right) &= \mathit{unlikely}_1 \end{aligned}.$$

3. Otherwise, $\bar{\mu}(|\psi\rangle\langle\psi|) = \mathit{unlikely}_0$ and $\bar{\mu}(\mathbb{1} -_{\mathcal{O}} |\psi\rangle\langle\psi|) = \mathit{likely}_0$.

[14] Yu-Tsung says: This example is a little bit too long to human-verify, and we need more examples. Maybe we need to write a program for further verification.

We now prove $\bar{\mu}$ has an empty core by contradiction. Assume there is a real-valued probability measure satisfying $\mu_\rho^B(P) \in \bar{\mu}(P)$ for all $P \in \mathcal{E}$. We first claim that for all $0 \le \theta < \frac{\pi}{2}$ and $0 \le \gamma < \pi$, we have $\mu_\rho^B\left(\left|0^\perp_{\theta,\gamma}\right\rangle\left\langle 0^\perp_{\theta,\gamma}\right|\right) = r_0$. On one hand, $\mu_\rho^B\left(\left|0^\perp_{\theta,\gamma}\right\rangle\left\langle 0^\perp_{\theta,\gamma}\right|\right) \in \bar{\mu}\left(\left|0^\perp_{\theta,\gamma}\right\rangle\left\langle 0^\perp_{\theta,\gamma}\right|\right) = [r_0, 4r_0 {}_{\mathbb{R}}- 1]$ implies $\mu_\rho^B\left(\left|0^\perp_{\theta,\gamma}\right\rangle\left\langle 0^\perp_{\theta,\gamma}\right|\right) \ge r_0$. On the other hand, by figure 1, it is easy to see for any neighborhood of $\left|0^\perp_{\theta,\gamma}\right\rangle$, these is always a state $|\psi\rangle$ such that $\mu_\rho^B(|\psi\rangle\langle\psi|) \in \bar{\mu}(|\psi\rangle\langle\psi|) = \mathscr{unlikely}_0$, i.e., $\mu_\rho^B(|\psi\rangle\langle\psi|) \le r_0$. Because μ_ρ^B is continuous, we have $\mu_\rho^B\left(\left|0^\perp_{\theta,\gamma}\right\rangle\left\langle 0^\perp_{\theta,\gamma}\right|\right) = r_0$. Similarly, $\mu_\rho^B\left(\left|+^\perp_{\theta,\gamma}\right\rangle\left\langle +^\perp_{\theta,\gamma}\right|\right) = r_0$.

Consider a general density matrix $\rho = {}_{\mathscr{O}}\sum_{i=0}^2 {}_{\mathscr{O}}\sum_{j=0}^2 r_{ij} |i\rangle\langle j|$. Then, we have

$$\begin{aligned}
r_0 &= \mu_\rho^B\left(\left|0^\perp_{\theta,\gamma}\right\rangle\left\langle 0^\perp_{\theta,\gamma}\right|\right) \\
&\overset{(9)}{=} \mathrm{Tr}\left(\rho\left|0^\perp_{\theta,\gamma}\right\rangle\left\langle 0^\perp_{\theta,\gamma}\right|\right) \\
&= \mathrm{Tr}\left(\left({}_{\mathscr{O}}\sum_{i=0}^2 {}_{\mathscr{O}}\sum_{j=0}^2 r_{ij}|i\rangle\langle j|\right)\left(\mathrm{e}^{\mathrm{i}\gamma}\sin\theta|1\rangle + \cos\theta|2\rangle\right)\left(\mathrm{e}^{\mathrm{i}\gamma}\sin\theta|1\rangle + \cos\theta|2\rangle\right)^\dagger\right) \\
&= \mathrm{Tr}\left({}_{\mathscr{O}}\sum_{i=0}^2 \left(r_{i1}\mathrm{e}^{\mathrm{i}\gamma}\sin\theta + r_{i2}\cos\theta\right)|i\rangle\left(\mathrm{e}^{-\mathrm{i}\gamma}\sin\theta\langle 1| + \cos\theta\langle 2|\right)\right) \\
&= r_{11}\sin^2\theta + r_{12}\cos\theta\mathrm{e}^{-\mathrm{i}\gamma}\sin\theta + r_{21}\mathrm{e}^{\mathrm{i}\gamma}\sin\theta\cos\theta + r_{22}\cos^2\theta
\end{aligned}$$

Because the above equation holds for all $0 \le \theta < \frac{\pi}{2}$ and $0 \le \gamma < \pi$, we must have $r_{12} = r_{21} = 0$ and $r_{11} = r_{22} = r_0$. We also have

$$\begin{aligned}
r_0 &= \mu_\rho^B\left(\left|+^\perp_{\theta,\gamma}\right\rangle\left\langle +^\perp_{\theta,\gamma}\right|\right) \\
&\overset{(9)}{=} \mathrm{Tr}\left(\rho\left|+^\perp_{\theta,\gamma}\right\rangle\left\langle +^\perp_{\theta,\gamma}\right|\right) \\
&= \mathrm{Tr}\left(\left({}_{\mathscr{O}}\sum_{i=0}^2 {}_{\mathscr{O}}\sum_{j=0}^2 r_{ij}|i\rangle\langle j|\right)\left(-\mathrm{e}^{\mathrm{i}\gamma}\sin\theta|-\rangle + \cos\theta|2\rangle\right)\left(-\mathrm{e}^{\mathrm{i}\gamma}\sin\theta|-\rangle + \cos\theta|2\rangle\right)^\dagger\right) \\
&= \mathrm{Tr}\left({}_{\mathscr{O}}\sum_{i=0}^2\left(-\frac{r_{i0}\mathrm{e}^{\mathrm{i}\gamma}\sin\theta}{\sqrt{2}} + \frac{r_{i1}\mathrm{e}^{\mathrm{i}\gamma}\sin\theta}{\sqrt{2}} + r_{i2}\cos\theta\right)|i\rangle\left(-\frac{\mathrm{e}^{-\mathrm{i}\gamma}\sin\theta}{\sqrt{2}}\langle 0| + \frac{\mathrm{e}^{-\mathrm{i}\gamma}\sin\theta}{\sqrt{2}}\langle 1| + \cos\theta\langle 2|\right)\right) \\
&= -\frac{\mathrm{e}^{-\mathrm{i}\gamma}\sin\theta}{\sqrt{2}}\left(-\frac{r_{00}\mathrm{e}^{\mathrm{i}\gamma}\sin\theta}{\sqrt{2}} + \frac{r_{01}\mathrm{e}^{\mathrm{i}\gamma}\sin\theta}{\sqrt{2}} + r_{02}\cos\theta\right) \\
&\quad + \frac{\mathrm{e}^{-\mathrm{i}\gamma}\sin\theta}{\sqrt{2}}\left(-\frac{r_{10}\mathrm{e}^{\mathrm{i}\gamma}\sin\theta}{\sqrt{2}} + \frac{r_0\mathrm{e}^{\mathrm{i}\gamma}\sin\theta}{\sqrt{2}}\right) + \cos\theta\left(-\frac{r_{20}\mathrm{e}^{\mathrm{i}\gamma}\sin\theta}{\sqrt{2}} + r_0\cos\theta\right) \\
&= \left(\frac{r_{00}\sin^2\theta}{2} - \frac{r_{01}\sin^2\theta}{2} - \frac{r_{02}\mathrm{e}^{-\mathrm{i}\gamma}\sin\theta\cos\theta}{\sqrt{2}}\right) + \left(-\frac{r_{10}\sin^2\theta}{2} + \frac{r_0\sin^2\theta}{2}\right) + \left(-\frac{r_{20}\mathrm{e}^{\mathrm{i}\gamma}\cos\theta\sin\theta}{\sqrt{2}} + r_0\cos^2\theta\right) \\
&= \left(\frac{r_{00}}{2} - \frac{r_{01}}{2} - \frac{r_{10}}{2} + \frac{r_0}{2}\right)\sin^2\theta + r_0\cos^2\theta - \left(r_{02}\mathrm{e}^{-\mathrm{i}\gamma} + r_{20}\mathrm{e}^{\mathrm{i}\gamma}\right)\frac{\cos\theta\sin\theta}{\sqrt{2}}
\end{aligned}$$

Because of the same reason, we have $r_{02} = r_{20} = 0$ and $r_{00} = r_0 + r_{01} + r_{10}$. Moreover, because ρ is a density matrix, we should have $\rho^\dagger = \rho$, $\mathrm{Tr}(\rho) = 1$, and ρ is a positive operator [19]. These implies that we can express $r_{01} = a + \mathrm{i}b$ and $r_{10} = a - \mathrm{i}b$ for some a, $b \in \mathbb{R}$. After plugging into $\mathrm{Tr}(\rho) = 1$, we have $r_{00} = 1 - 2r_0$, $r_{01} = \frac{1}{2} - \frac{3}{2}r_0 + \mathrm{i}b$, and $r_{01} = \frac{1}{2} - \frac{3}{2}r_0 - \mathrm{i}b$. Next, since ρ is a positive operator, ρ should have all non-negative eigenvalues, which are the solutions of the characteristic polynomial

$$\begin{aligned}
0 &= \det\left(\rho {}_{\mathscr{O}}- \lambda\mathbb{1}\right) \\
&= \begin{vmatrix} 1 - 2r_0 - \lambda & \frac{1}{2} - \frac{3}{2}r_0 + \mathrm{i}b & 0 \\ \frac{1}{2} - \frac{3}{2}r_0 - \mathrm{i}b & r_0 - \lambda & 0 \\ 0 & 0 & r_0 - \lambda \end{vmatrix} \\
&= (r_0 - \lambda)\left[(1 - 2r_0 - \lambda)(r_0 - \lambda) - \left(\frac{1}{2} - \frac{3}{2}r_0 + \mathrm{i}b\right)\left(\frac{1}{2} - \frac{3}{2}r_0 - \mathrm{i}b\right)\right] \\
&= (r_0 - \lambda)\left[\lambda^2 - (1 - r_0)\lambda + r_0 - 2r_0^2 - \frac{1}{4} + \frac{3}{2}r_0 - \frac{9}{4}r_0^2 - b^2\right] \\
&= (r_0 - \lambda)\left[\lambda^2 - (1 - r_0)\lambda - \frac{1}{4} + \frac{5}{2}r_0 - \frac{17}{4}r_0^2 - b^2\right]
\end{aligned}$$

Hence, the solutions of λ are r_0 and $1 - r_0 \pm \sqrt{(1-r_0)^2 - 4\cdot 1\cdot\left(-\frac{1}{4}+\frac{5}{2}r_0 - \frac{17}{4}r_0^2\mathbb{R} - b^2\right)}$. If they are all non-negative, we want

$$\begin{aligned} 0 &\geq -4\cdot 1\cdot\left(-\frac{1}{4}+\frac{5}{2}r_0-\frac{17}{4}r_0^2-b^2\right) \\ &= 1-10r_0+17r_0^2+4b^2 \\ &= 17\left(r_0-\frac{5+\sqrt{25-17\cdot(4b^2+1)}}{17}\right)\left(r_0-\frac{5-\sqrt{25-17\cdot(4b^2+1)}}{17}\right) \end{aligned}$$

Therefore, when $r_0 > \frac{5+2\sqrt{2}}{17}$ or $r_0 < \frac{5-2\sqrt{2}}{17}$, ρ always has a negative eigenvalue for all $b \in \mathbb{R}$. This contradiction implies $\bar{\mu}$ has an empty core when $r_0 \in \left(\frac{5_{\mathbb{R}}+2\sqrt{2}}{17}, \frac{1}{2}\right]$.

Similar to example 7, to verify $\bar{\mu}$ is a quantum interval-valued probability measure for $r_0 \in \left[\frac{1}{3}, \frac{1}{2}\right]$, we are going to enumerate all possible orthonormal bases $\mathcal{B} = \{|\psi_0\rangle, |\psi_1\rangle, |\psi_2\rangle\}$ to verify equation (15).

1. When $|\psi_0\rangle = \left|0^{\perp}_{\theta,\gamma}\right\rangle$ for $0 \leq \theta < \frac{\pi}{2}$ and $0 \leq \gamma < \pi$, we want to prove $\bar{\mu}(|\psi_1\rangle\langle\psi_1|) = \bar{\mu}(|\psi_2\rangle\langle\psi_2|) = \mathit{unlikely}_0$. Then, we have

$$\begin{aligned} &\bar{\mu}\left(\mathbb{1}_{\mathcal{O}} - \left|0^{\perp}_{\theta,\gamma}\right\rangle\left\langle 0^{\perp}_{\theta,\gamma}\right|\right) = \mathit{unlikely}_1 = [2_{\mathbb{R}} - 4r_0, 1_{\mathbb{R}} - r_0] \\ \subseteq &[2_{\mathbb{R}} - 4r_0, 2r_0] = \mathit{unlikely}_0 +_{\mathscr{I}} \mathit{unlikely}_0 = \bar{\mu}(|\psi_1\rangle\langle\psi_1|) +_{\mathscr{I}} \bar{\mu}(|\psi_2\rangle\langle\psi_2|) \\ &\bar{\mu}(\mathbb{1}_{\mathcal{O}} - |\psi_1\rangle\langle\psi_1|) = \mathit{likely}_0 = [1_{\mathbb{R}} - r_0, 2r_0] \\ \subseteq &[1_{\mathbb{R}} - r_0, 5r_0 {}_{\mathbb{R}} - 1] = \mathit{unlikely}_0 +_{\mathscr{I}} \mathit{likely}_1 = \bar{\mu}(|\psi_2\rangle\langle\psi_2|) +_{\mathscr{I}} \bar{\mu}\left(\left|0^{\perp}_{\theta,\gamma}\right\rangle\left\langle 0^{\perp}_{\theta,\gamma}\right|\right) \end{aligned} \tag{23}$$

   In order to verify $\bar{\mu}(|\psi_1\rangle\langle\psi_1|) = \bar{\mu}(|\psi_2\rangle\langle\psi_2|) = \mathit{unlikely}_0$, it is sufficient to prove that if $\left|+^{\perp}_{\theta',\gamma'}\right\rangle \in \mathcal{B}$, then $(0 < \theta' < \frac{\pi}{2}$ and $\pi \leq \gamma' < 2\pi)$ or $\theta' = \frac{\pi}{2}$. This is easy because of equation (18). Similarly, when $|\psi_0\rangle$ is $\left|+^{\perp}_{\theta,\gamma}\right\rangle$, equation (15) holds.

2. Otherwise, when $\left|0^{\perp}_{\theta,\gamma}\right\rangle \notin \mathcal{B}$, and $\left|+^{\perp}_{\theta,\gamma}\right\rangle \notin \mathcal{B}$ for $0 \leq \theta < \frac{\pi}{2}$ and $0 \leq \gamma < \pi$, equation (15) can easily be verified.

$$\begin{aligned} &\bar{\mu}(\mathbb{1}_{\mathcal{O}} - |\psi_0\rangle\langle\psi_0|) = \mathit{likely}_0 = [1_{\mathbb{R}} - r_0, 2r_0] \\ \subseteq &[2_{\mathbb{R}} - 4r_0, 2r_0] = \mathit{unlikely}_0 +_{\mathscr{I}} \mathit{unlikely}_0 = \bar{\mu}(|\psi_1\rangle\langle\psi_1|) +_{\mathscr{I}} \bar{\mu}(|\psi_2\rangle\langle\psi_2|) \end{aligned} \tag{24}$$

□

Yu-Tsung says: Since we have a family of interval-valued probability measures with empty cores, we need to find a more systematically way to how close these interval-valued probability measures to the Born rule. Need to check whether people defined the distance between games or not... Possible candidates are sup-norm or total variation distance(?)

B Proofs

Because the importance of Gleason's theorem, its proof have been rewritten to be more accessible [46, 7, 47, 48]. Since the additivity of the quantum probability measure (8), a quantum probability measure μ is completely determined by its value on the one-dimensional projections $\mu(|\psi\rangle\langle\psi|)$.

Definition 9. [49, 50] The unit sphere in $\mathbb{R}^d$ is denoted by $S^{d-1} = \left\{|\psi\rangle \in \mathbb{R}^d \middle| \langle\psi|\psi\rangle = 1\right\}$, while the unit sphere in $\mathbb{C}^d$ is denoted by $\mathbb{C}S^{d-1} = \left\{|\psi\rangle \in \mathbb{C}^d \middle| \langle\psi|\psi\rangle = 1\right\}$. Since $\mathbb{C}$ is one-to-one correspondence to $\mathbb{R}^2$ [51], $\mathbb{C}S^{d-1}$ has the same structure as S^{2d-1}. However, different symbols emphasis that they are embedded in different spaces. □

Definition 10. A function $f : S^{d-1} \to \mathbb{R}$ or $f : \mathbb{C}S^{d-1} \to \mathbb{R}$ is called a *real* or *complex approximation frame function* of weight (W_l, W_r) if

- $W_l \leq {}_{\mathbb{R}}\sum_{i=0}^{d-1} f(|\psi_i\rangle) \leq W_r$ for any orthonormal basis $\{|\psi_i\rangle\}_{i=0}^{d-1}$ of $\mathbb{R}^d$ or $\mathbb{C}^d$.
- $f(\lambda|\psi\rangle) = f(|\psi\rangle)$ for any $|\psi\rangle \in S^{d-1}$ or $|\psi\rangle \in \mathbb{C}S^{d-1}$ and $\lambda \in S^0$ or $\lambda \in \mathbb{C}S^0$.

Also, we will use the same letter f to denote the function defined on nonzero vector $|\psi\rangle$ by $f(|\psi\rangle) = \||\psi\rangle\|^2 f\left(\||\psi\rangle\|^{-1}|\psi\rangle\right)$and $f(0) = 0$ [47]. In this case,

- $f(\lambda|\psi\rangle) = f(|\psi\rangle)$ for any $|\psi\rangle \in \mathbb{R}^d$ or $|\psi\rangle \in \mathbb{C}^d$ and $\lambda \in \mathbb{R} \setminus \{0\}$ or $\lambda \in \mathbb{C} \setminus \{0\}$.

□

Definition 11. [52] For any function f, the modulus of continuity of f is $\omega_f(\delta) = \sup_{\||\psi\rangle - |\phi\rangle\| < \delta} |f(|\psi\rangle) - f(|\phi\rangle)|$. □

Definition 12. Given a vector $|\psi\rangle \in \mathbb{C}^3$, $\sigma_{|\psi\rangle}$ is the clockwise rotation by $\frac{\pi}{2}$ about $|\psi\rangle$. More explicitly, for any vector $|\phi\rangle \in \mathbb{C}S^2$, $\sigma_{|\psi\rangle}|\phi\rangle = \frac{|\psi\rangle \times |\phi\rangle}{\||\psi\rangle \times |\phi\rangle\|}$. □

Definition 13. [47] If $f : S^2 \to \mathbb{R}$ is a approximation frame function of weight (W_l, W_r), then the *symmetrization* of f with respect to $|0\rangle$ is the function $g_{f,|0\rangle} : S^2 \to \mathbb{R}$ defined by

$$g_{f,|0\rangle}(|\psi\rangle) = f(|\psi\rangle) + f\left(\sigma_{|0\rangle}|\psi\rangle\right) .$$

□

Definition 14. [6, 46, 47, 48] Given a vector $|\psi\rangle \in S^2$ with a specified north pole $|0\rangle \in S^2$, then

- $\theta(|\psi\rangle) = \arccos\langle 0|\psi\rangle$ denotes the the latitude of $|\psi\rangle$.
- The great circle on S^2 is the intersection of S^2 with a plane containing the center of S^2.
- If $0 < \theta(|\psi\rangle) < \frac{\pi}{2}$, there is a unique great circle tangent to the circle of latitude through $|\psi\rangle$; we shall call it a East-West great circle through $|\psi\rangle$, and denoted by $C(|\psi\rangle)$.
- We say taking a small step along an East-West great circle to go from $|\psi\rangle$ to a point $|\psi'\rangle$ if $|\psi'\rangle \in C(|\psi\rangle)$ and
$$\cos\theta(|\psi'\rangle) = \langle 0|\psi'\rangle \leq 2\langle 0|\psi\rangle = 2\cos\theta(|\psi\rangle) .$$

□

Lemma 2. Let
$$n(\theta) = \begin{cases} 7 & \text{, if } \cos\theta \leq \frac{1}{32} \text{ ;} \\ k+2 & \text{, if } \frac{1}{2^{k-1}} < \cos\theta \leq \frac{1}{2^k} \text{ for } k \geq 6 \text{ .} \end{cases}$$
If $|\psi\rangle$ and $|\psi'\rangle$ are points in the northern hemisphere with the same latitude $\theta(|\psi\rangle) = \theta(|\psi'\rangle)$, and the difference of their longitude is a, then we could takes at least $n(\theta(|\psi\rangle))$ close steps along East-West great circles to go from $|\psi'\rangle$ to a point $|\psi_s\rangle$ due south of $|\psi\rangle$. When taking $n \geq n(\theta(|\psi\rangle))$ steps, the latitude will go down by at most $\frac{a^2}{n}$. □

Proof. .... □

The following theorem is an analog to Lemma 20 in [47].

Lemma 3. If $f : S^2 \to [0, \infty)$ is an approximation frame function of weight (W_l, W_r), and $\eta = f(|0\rangle)$, then the oscillation on any latitude of the symmetrization of f with respect to $|0\rangle$ is at most

$$\sup_{\theta \in [0, \frac{\pi}{2})} \min\left\{ \min_{n \geq n(\theta)} 2n\eta + 4n(W_r - W_l) + 2\omega_f\left(2 - 2\cos\left(\frac{\pi^2}{n}\right)\right), 2\omega_f(2\cos\theta) \right\} .$$

The exact n could be computed to minimize the oscillation when the function f is given. □

Proof. Consider $|0\rangle$ to be the north pole of S^2. Let g be the symmetrization of f with respect to $|0\rangle$, then $\omega_g(\delta) = 2\omega_f(\delta)$[15]. Let $|\psi\rangle$ and $|\psi'\rangle$ be points in the northern hemisphere with $\theta(|\psi\rangle) = \theta(|\psi'\rangle)$. Because the maximum distance between points in latitude $\theta(|\psi\rangle)$ is $2\cos\theta(|\psi\rangle)$, then

$$|g(|\psi\rangle) - g(|\psi'\rangle)| \le \omega_g(2\cos\theta(|\psi\rangle)) = 2\omega_f(2\cos\theta(|\psi\rangle)) \ .$$

Let a be the difference in longitudes of $|\psi\rangle$ and $|\psi'\rangle$. By lemma 2, we can take $n \ge n(\theta(|\psi\rangle))$ close steps along East-West great circles to go from $|\psi'\rangle$ to a point $|\psi_s\rangle$ due south of $|\psi\rangle$. The difference in latitude between $|\psi\rangle$ and $|\psi_s\rangle$ is at most $\frac{a^2}{n}$, while, from Lemma [16],

$$g(|\psi_s\rangle) \ge g(|\psi'\rangle) - n[2f(|0\rangle) + 4(W_r - W_l)] = g(|\psi'\rangle) - 2n\eta - 4n(W_r - W_l) \ .$$

When the difference in latitude between $|\psi\rangle$ and $|\psi_s\rangle$ is at most $\frac{a^2}{n}$, the length between $|\psi\rangle$ and $|\psi_s\rangle$ is at most $2 - 2\cos\left(\frac{a^2}{n}\right)$, and hence

$$g(|\psi\rangle) \ge g(|\psi_s\rangle) - \omega_g\left(2 - 2\cos\left(\frac{a^2}{n}\right)\right) \ge g(|\psi'\rangle) - 2n\eta - 4n(W_r - W_l) - 2\omega_f\left(2 - 2\cos\left(\frac{\pi^2}{n}\right)\right) \ .$$

[17] ∎

The following theorem is an analog to Theorem 21 in [47].

Theorem 6.

- Let $f : S^2 \to [0,\infty)$ be a continuous[18] approximation frame function of weight (W_l, W_r), $W = \frac{W_l + W_r}{2}$, $\eta = f(|0\rangle)$,
- $\ell(|\psi\rangle) = \langle 0|\psi\rangle^2$ and $Q(|\psi\rangle) = W(1 - \ell(|\psi\rangle))$, and
- $\delta(\eta, W_\Delta, \omega_f) = \sup_{\theta\in[0,\frac{\pi}{2})} \min\left\{\min_{n\ge n(\theta)} 2n\eta + 4nW_\Delta + 2\omega_f\left(2 - 2\cos\left(\frac{\pi^2}{n}\right)\right), 2\omega_f(2\cos\theta)\right\}$.

Then, $\left|Q(|\psi\rangle) - g_{f,|0\rangle}(|\psi\rangle)\right| \le 9\delta(\eta, W_r - W_l, \omega_f) + 3(W_r - W_l)$ for all $|\psi\rangle \in S^2$. ∎

Proof. Consider $|0\rangle$ to be the north pole of S^2. Note that $g_{f,|0\rangle}$ is a nonnegative frame function of weight $(2W_l, 2W_r)$, and that $W_l - \eta \le g_{f,|0\rangle}(|\psi\rangle) \le W_r - \eta$ for all $|\psi\rangle$ on the equator. From Lemma 3, the oscillation of $g_{f,|0\rangle}$ on latitudes is at most $\delta(\eta, W_r - W_l, \omega_f)$.

If $|\psi_0\rangle$, $|\psi_1\rangle$, and $|\psi_2\rangle$ are orthogonal vectors on the closed northern hemisphere, then $\ell(|\psi_0\rangle) + \ell(|\psi_1\rangle) + \ell(|\psi_2\rangle) = 1$. Conversely, if $a_0 + a_1 + a_2 = 1$, for $a_i \in [0,1]$, then there exist orthogonal vectors $|\psi_0\rangle$, $|\psi_1\rangle$, and $|\psi_2\rangle$ on the closed northern hemisphere such that $a_i = \ell(|\psi_i\rangle)$.

For any point $|\psi\rangle$ in the closed northern hemisphere, consider its longitude φ, and a function $F(a) = W - g_{f,|0\rangle}(|\psi'\rangle)$, where $|\psi'\rangle$ is a point in the closed northern hemisphere with longitude φ, and $a = \ell(|\psi'\rangle)$. More specifically, let $|\psi_i'\rangle$ be a point in the closed northern hemisphere with longitude φ such that $a_i = \ell(|\psi_i'\rangle)$. Hence, $|\psi_i'\rangle$ and $|\psi_i\rangle$ are in the same latitude so that

$$\left|F(a_i) - \left(W - g_{f,|0\rangle}(|\psi_i\rangle)\right)\right| = \left|g_{f,|0\rangle}(|\psi_i\rangle) - g_{f,|0\rangle}(|\psi_i'\rangle)\right| \overset{\text{Lemma 3}}{\le} \delta(\eta, W_r - W_l, \omega_f) \ .$$

[15] Yu-Tsung says: Notice that [47] derives the modulus of continuity $\omega_f(\delta)$ in their Corollary 18(?). Maybe need to upgrade and derive the modulus of continuity in our case later.

[16] Yu-Tsung says: Analog to Lemma 16 in [47]... To be typed...

[17] Yu-Tsung says: We are going to verify this proof by set $W_r - W_l = 0$. Since f is continuous in this case, we could solve $n \ge 2\pi$ such that $2\omega_f\left(2 - 2\cos\left(\frac{\pi^2}{n}\right)\right) \le \frac{\delta}{2}$ for any $\delta > 0$, and then solve $\eta \ge 0$ such that $2n\eta \le \frac{\delta}{2}$.

[18] Yu-Tsung says: We need this "continuous" to apply Lemma 19 in [47], and need to rewrite Lemma 19 if we want to take away this "continuous"...

for all i. Whence,

$$\begin{aligned}
& \left|F(a_0)+F(a_1)+F(a_2)-W\right| \\
\leq\ & \left|F(a_0)+F(a_1)+F(a_2)-3W+g_{f,|0\rangle}(|\psi_0\rangle)+g_{f,|0\rangle}(|\psi_1\rangle)+g_{f,|0\rangle}(|\psi_2\rangle)\right| \\
& +\left|2W-\left(g_{f,|0\rangle}(|\psi_0\rangle)+g_{f,|0\rangle}(|\psi_1\rangle)+g_{f,|0\rangle}(|\psi_2\rangle)\right)\right| \\
\leq\ & \sum_{i=0}^{2}\left|F(a_i)-W+g_{f,|0\rangle}(|\psi_i\rangle)\right|+\frac{2W_r-2W_l}{2} \\
\leq\ & 3\delta(\eta,W_r-W_l,\omega_f)+W_r-W_l\ .
\end{aligned}$$

Therefore, by Lemma 19 in [47],

$$\left|Q(|\psi\rangle)-g_{f,|0\rangle}(|\psi\rangle)\right|=\left|W-g_{f,|0\rangle}(|\psi\rangle)-W\ell(|\psi\rangle)\right|\leq 9\delta(\eta,W_r-W_l,\omega_f)+3(W_r-W_l)\ .$$

□

The following theorem is an analog to Corollary 22 in [47].

Corollary 1.

- Let $f:S^2\to\mathbb{R}$ be a bounded continuous approximation frame function of weight (W_l,W_r), $W=\frac{W_l+W_r}{2}$, $m=\min_{|\psi\rangle\in S^2}f(|\psi\rangle)$, $\eta=f(|0\rangle)-m$,
- $Q(|\psi\rangle)=W-m+(3m-W)\langle 0|\psi\rangle^2=2m\langle 0|\psi\rangle^2+(W-m)\left(1-\langle 0|\psi\rangle^2\right)$, and
- $\delta(\eta,W_\Delta,\omega_f)=\sup_{\theta\in[0,\frac{\pi}{2})}\min\left\{\min_{n\geq n(\theta)}2n\eta+4nW_\Delta+2\omega_f\left(2-2\cos\left(\frac{\pi^2}{n}\right)\right),2\omega_f(2\cos\theta)\right\}$.

Then, $\left|Q(|\psi\rangle)-g_{f,|0\rangle}(|\psi\rangle)\right|\leq 9\delta(\eta,W_r-W_l,\omega_f)+3(W_r-W_l)$ for all $|\psi\rangle\in S^2$. □

Proof. Clearly, $f'=f-m$ is a nonnegative frame function of weight $(W_l',W_r')=(W_l-3m,W_r-3m)$ and $W'=\frac{W_l'+W_r'}{2}=W-3m$. Then, $\eta=f'(|0\rangle)$ and $g_{f',|0\rangle}=g_{f,|0\rangle}-2m$. Let $Q'(|\psi\rangle)=Q(|\psi\rangle)-2m=W'\left(1-\langle 0|\psi\rangle^2\right)$, and we have $\left|Q(|\psi\rangle)-g_{f,|0\rangle}(|\psi\rangle)\right|=\left|Q'(|\psi\rangle)-g_{f',|0\rangle}(|\psi\rangle)\right|\leq 9\delta(\eta,W_r-W_l,\omega_f)+3(W_r-W_l)$ from theorem 6. □

Corollary 22 in [47] is used to prove theorem 23 which is the final step to prove Gleason's theorem in [47]. Theorem 23 reads as the following.

Theorem 7.

- Let $f:S^2\to\mathbb{R}$ be a bounded continuous approximation frame function of weight (W_l,W_r), $W=\frac{W_l+W_r}{2}$,
- $M=\max_{|\psi\rangle\in S^2}f(|\psi\rangle)=f(|0\rangle)$, $m=\min_{|\psi\rangle\in S^2}f(|\psi\rangle)=f(|\bar{0}\rangle)$,
- $|2\rangle=\sigma_{|0\rangle\times|\bar{0}\rangle}|0\rangle$, $|1\rangle=|2\rangle\times|0\rangle$,
- $Q(|\psi\rangle)=M\langle 0|\psi\rangle^2+(W-M-m)\langle 1|\psi\rangle^2+m\langle 2|\psi\rangle^2$,
- $\delta(\eta,W_\Delta,\omega_f)=\sup_{\theta\in[0,\frac{\pi}{2})}\min\left\{\min_{n\geq n(\theta)}2n\eta+4nW_\Delta+2\omega_f\left(2-2\cos\left(\frac{\pi^2}{n}\right)\right),2\omega_f(2\cos\theta)\right\}$, and
- $f'(|\psi\rangle)=f(|\psi\rangle)-Q(|\psi\rangle)$.

Then, $|Q(|\psi\rangle)-f(|\psi\rangle)|\leq 27\delta(W_r-W_l,W_r-W_l,\omega_f)+27\delta(W_r-W_l,W_r-W_l,\omega_{f'})+18(W_r-W_l)$ for all $|\psi\rangle\in S^2$. □

Proof. Let $|\bar{2}\rangle = \sigma_{|0\rangle\times|\bar{0}\rangle}|\bar{0}\rangle$. Then

$$|(f(|0\rangle) + f(|2\rangle)) - (f(|\bar{0}\rangle) + f(|\bar{2}\rangle))| \le W_r - W_l$$

so

$$f(|\bar{2}\rangle) - f(|2\rangle) \ge f(|0\rangle) - f(|\bar{0}\rangle) - (W_r - W_l) = M - m - (W_r - W_l)$$

whence $f(|2\rangle) \le m + W_r - W_l$. Let $(x, y, z) = (\langle\psi|0\rangle, \langle\psi|1\rangle, \langle\psi|2\rangle)$ denote the coordinates of a point $|\psi\rangle$ with respect to the frame $\{|i\rangle\}_{i=0}^{2}$, then $Q(|\psi\rangle) = Mx^2 + (W - M - m)y^2 + mz^2$.

Then, [19]

$$\begin{aligned} \sigma_{|0\rangle}(x, y, z) &= (x, -z, y) \\ \sigma_{|2\rangle}(x, y, z) &= (-y, x, z) \end{aligned}$$

Note that

$$\begin{aligned} Q(|\psi\rangle) + Q\left(\sigma_{|0\rangle}|\psi\rangle\right) &= 2Mx^2 + (W - M)\left(y^2 + z^2\right) \\ Q(|\psi\rangle) + Q\left(\sigma_{|2\rangle}|\psi\rangle\right) &= 2mz^2 + (W - m)\left(x^2 + y^2\right) \end{aligned}$$

which is the same as the quadratic form defined in corollary 1. Therefore, by corollary 1, we should have

$$\begin{aligned} \left|f'(|\psi\rangle) + f'\left(\sigma_{|0\rangle}|\psi\rangle\right)\right| &= \left|g_{f,|0\rangle} - \left(2Mx^2 + (W - M)\left(y^2 + z^2\right)\right)\right| \le 9\delta(0, W_r - W_l, \omega_f) + 3(W_r - W_l) \\ \left|f'(|\psi\rangle) + f'\left(\sigma_{|2\rangle}|\psi\rangle\right)\right| &= \left|g_{f,|2\rangle} - \left(2mz^2 + (W - m)\left(x^2 + y^2\right)\right)\right| \le 9\delta(W_r - W_l, W_r - W_l, \omega_f) + 3(W_r - W_l) \end{aligned}$$

Follow the same procedure in the proof in [47]. We have

$$|f'(|\psi\rangle)| = |f(|\psi\rangle) - Q(|\psi\rangle)| \le \frac{3}{2}\left(9\delta(W_r - W_l, W_r - W_l, \omega_f) + 3(W_r - W_l)\right)$$

on the six great circles $x = \pm y$, $y = \pm z$, and $z = \pm x$ on S^2. In particular,

$$\left|f(|\psi\rangle) - M\left(x^2 - z^2\right) - Wz^2\right| \le \frac{3}{2}\left(9\delta(W_r - W_l, W_r - W_l, \omega_f) + 3(W_r - W_l)\right)$$

on the great circle $y = z$.

Let $M' = \max_{|\psi\rangle\in S^2} f'(|\psi\rangle)$, $m' = \min_{|\psi\rangle\in S^2} f'(|\psi\rangle)$. Note that the weight of f' is $(W_l', W_r') = (W_l - W, W_r - W)$ and $W' = \frac{W_l' + W_r'}{2} = 0$. Choose an orthonormal basis $\{|i'\rangle\}_{i=0}^{2}$ such that $f'(|0'\rangle) = M'$ and $f(|2'\rangle) \le m' + W_r' - W_l' = m' + W_r - W_l$. In this case, the approximating form is$Q'(|\psi\rangle) = Mx'^2 - (M + m)y'^2 + mz'^2$. By the previous paragraph, for t on the great circle $y' = z'$, we have

$$\left|f'(|\psi\rangle) - M'\left(x'^2 - z'^2\right)\right| \le \frac{3}{2}\left(9\delta(W_r - W_l, W_r - W_l, \omega_{f'}) + 3(W_r - W_l)\right)$$

Notice that the great circles $x = \pm y$, $y = \pm z$, and $z = \pm x$ on S^2 divide the sphere into 24 triangles. The 14 vertices are

$$\left(\pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}\right), \quad (\pm 1, 0, 0), \quad (0, \pm 1, 0), \quad (0, 0, \pm 1).$$

All 24 triangles are congruent to

$$(1, 0, 0), \quad \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \quad \left(\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right).$$

The maximum distance between two vertices is $\frac{2}{\sqrt{3}}$. So any great circle through a point in the interior of that triangle must intersect the boundary at a distance at most $\frac{1}{\sqrt{3}}$ from the point.

[19] Yu-Tsung says: Verify the details. Especially the left-handed or right-handed of the coordinate system, and the rotation matrix... >"<

In particular, because $|0'\rangle$ is on the great circle $y' = z'$, the great circle $y' = z'$ must intersect the boundary of one of the previous triangles. Let $|\psi\rangle$ be the intersection, the distance between $|0'\rangle$ and $|\psi\rangle$ is at most $\frac{1}{\sqrt{3}}$. That is

$$(1-x')^2 + 2z'^2 = (1-x')^2 + y'^2 + z'^2 \leq \frac{1}{3} \ .$$

Also $x'^2 + 2z'^2 = 1$. So $2(1-x') \leq \frac{1}{3}$ or $x' \geq \frac{5}{6}$ so $2z'^2 \leq \frac{11}{36}$ whence $x'^2 - z'^2 \geq \frac{13}{24} > \frac{1}{2}$. Now $\left|M'\left(x'^2 - z'^2\right)\right| \leq \frac{3}{2}\left(9\delta\left(W_r - W_l, W_r - W_l, \omega_f\right) + 3\left(W_r - W_l\right)\right) + \frac{3}{2}\left(9\delta\left(W_r - W_l, W_r - W_l, \omega_{f'}\right) + 3\left(W_r - W_l\right)\right)$ so

$$\begin{aligned}|M'| &\leq 3\left(9\delta\left(W_r - W_l, W_r - W_l, \omega_f\right) + 3\left(W_r - W_l\right)\right) + 3\left(9\delta\left(W_r - W_l, W_r - W_l, \omega_{f'}\right) + 3\left(W_r - W_l\right)\right) \\ &= 27\delta\left(W_r - W_l, W_r - W_l, \omega_f\right) + 27\delta\left(W_r - W_l, W_r - W_l, \omega_{f'}\right) + 18\left(W_r - W_l\right)\end{aligned}$$

Same for m', so

$$|Q(|\psi\rangle) - f(|\psi\rangle)| \leq 27\delta\left(W_r - W_l, W_r - W_l, \omega_f\right) + 27\delta\left(W_r - W_l, W_r - W_l, \omega_{f'}\right) + 18\left(W_r - W_l\right)$$

□

Lemma 4.

- Let $f : \mathbb{C}S^1 \to \mathbb{R}$ be a bounded continuous approximation frame function of weight (W_l, W_r), $W = \frac{W_l + W_r}{2}$. Recall f could be uniquely extended to the whole space by $f(0) = 0$ and $f(|\psi\rangle) = \||\psi\rangle\|^2 f\left(\||\psi\rangle\|^{-1}|\psi\rangle\right)$ for nonzero vector $|\psi\rangle$.
- $M = \max_{|\psi\rangle \in \mathbb{C}S^1} f(|\psi\rangle) = f(|0\rangle)$.
- $|1\rangle \in \mathbb{C}S^1$ is orthogonal to $|0\rangle$.
- $Q(|\psi\rangle) = \langle\psi|\left[M|0\rangle\langle 0| + (W - M)|1\rangle\langle 1|\right]|\psi\rangle$.
- $f'(|\psi\rangle) = f(|\psi\rangle) - Q(|\psi\rangle)$.
- $\delta(\eta, W_\Delta, \omega_f) = \sup_{\theta \in [0, \frac{\pi}{2})} \min\left\{\min_{n \geq n(\theta)} 2n\eta + 4nW_\Delta + 2\omega_f\left(2 - 2\cos\left(\frac{\pi^2}{n}\right)\right), 2\omega_f(2\cos\theta)\right\}$.
- For any $|1'\rangle \in \mathbb{C}S^1$ is orthogonal to $|0\rangle$, we have $|Q(|\psi\rangle) - f(|\psi\rangle)| \leq 27\delta\left(W_r - W_l, W_r - W_l, \omega_f\right) + 27\delta\left(W_r - W_l, W_r - W_l, \omega_{f'}\right) + 18\left(W_r - W_l\right)$ for all $|\psi\rangle \in \mathbb{C}S^1$ such that $\langle 0|\psi\rangle, \langle 1'|\psi\rangle \in \mathbb{R}$.[20]

Then, $|Q(|\psi\rangle) - f(|\psi\rangle)| \leq 27\delta\left(W_r - W_l, W_r - W_l, \omega_f\right) + 27\delta\left(W_r - W_l, W_r - W_l, \omega_{f'}\right) + 18\left(W_r - W_l\right)$ for all $|\psi\rangle \in \mathbb{C}S^1$. □

Proof. For all $|\psi\rangle \in \mathbb{C}S^1$,

$$\begin{aligned}&|Q(|\psi\rangle) - f(|\psi\rangle)| \\ =\ &|Q(|\psi\rangle) - f(|0\rangle\langle 0|\psi\rangle + |1\rangle\langle 1|\psi\rangle)| \\ =\ &\left|Q\left(\frac{|\langle 0|\psi\rangle|}{\langle 0|\psi\rangle}|\psi\rangle\right) - f\left(\frac{|\langle 0|\psi\rangle|}{\langle 0|\psi\rangle}\left(|0\rangle\langle 0|\psi\rangle + |1\rangle\langle 1|\psi\rangle\right)\right)\right| \\ =\ &\left|Q\left(\frac{|\langle 0|\psi\rangle|}{\langle 0|\psi\rangle}|\psi\rangle\right) - f\left(|0\rangle|\langle 0|\psi\rangle| + |1\rangle\frac{|\langle 0|\psi\rangle|}{\langle 0|\psi\rangle}\frac{\langle 1|\psi\rangle}{|\langle 1|\psi\rangle|}|\langle 1|\psi\rangle|\right)\right| \\ \leq\ &27\delta\left(W_r - W_l, W_r - W_l, \omega_f\right) + 27\delta\left(W_r - W_l, W_r - W_l, \omega_{f'}\right) + 18\left(W_r - W_l\right)\end{aligned}$$

where the last inequality holds because $|1\rangle\frac{|\langle 0|\psi\rangle|}{\langle 0|\psi\rangle}\frac{\langle 1|\psi\rangle}{|\langle 1|\psi\rangle|} \in \mathbb{C}S^1$ is orthogonal to $|0\rangle$, $\langle 0|\left(\frac{|\langle 0|\psi\rangle|}{\langle 0|\psi\rangle}|\psi\rangle\right) = 1 \in \mathbb{R}$, and

$$\left(|1\rangle\frac{|\langle 0|\psi\rangle|}{\langle 0|\psi\rangle}\frac{\langle 1|\psi\rangle}{|\langle 1|\psi\rangle|}\right)^\dagger\left(\frac{|\langle 0|\psi\rangle|}{\langle 0|\psi\rangle}|\psi\rangle\right) = \frac{\langle\psi|1\rangle}{|\langle\psi|1\rangle|}\frac{|\langle\psi|0\rangle|}{\langle\psi|0\rangle}\langle 1|\frac{|\langle 0|\psi\rangle|}{\langle 0|\psi\rangle}|\psi\rangle \in \mathbb{R} \ .$$

□

[20] Yu-Tsung says: This lemma cannot fit with theorem 7 because the quadratic form required there may not be the quadratic form we used here... Or we should estimate the difference between two quadratic forms?

Definition 15. [6, 47, 40] On any complex vector space $\mathbb{C}^d$, a (positive-definite) Hermitian form $\langle\langle . | . \rangle\rangle : \mathbb{C}^d \times \mathbb{C}^d \to \mathbb{C}$ is a function satisfying:

- $\langle\langle \phi | \alpha_0 \psi_0 + \alpha_1 \psi_1 \rangle\rangle = \alpha_0 \langle\langle \phi | \psi_0 \rangle\rangle + \alpha_1 \langle\langle \phi | \psi_1 \rangle\rangle$, where α_0 and $\alpha_1 \in \mathbb{C}$; $|\phi\rangle$, $|\psi_0\rangle$, and $|\psi_1\rangle \in \mathbb{C}^d$[21].
- $\langle\langle \phi | \psi \rangle\rangle = \langle\langle \psi | \phi \rangle\rangle^*$ for all $|\phi\rangle$ and $|\psi\rangle \in \mathbb{C}^d$.
- $\langle\langle \phi | \phi \rangle\rangle > 0$ for all $|\phi\rangle \in \mathbb{C}^d \setminus \{0\}$.

□

Yu-Tsung says:

Theorem 8.

- Let $f : \mathbb{C}S^2 \to \mathbb{R}$ be a bounded continuous approximation frame function of weight (W_l, W_r), $W = \frac{W_l + W_r}{2}$. Recall f could be uniquely extended to the whole space $\mathbb{C}^3$ by $f(0) = 0$ and $f(|\psi\rangle) = \||\psi\rangle\|^2 f\left(\||\psi\rangle\|^{-1} |\psi\rangle\right)$ for nonzero vector $|\psi\rangle$.

Then, $|Q(|\psi\rangle) - f(|\psi\rangle)| \leq 27\delta(W_r - W_l, W_r - W_l, \omega_f) + 27\delta(W_r - W_l, W_r - W_l, \omega_{f'}) + 18(W_r - W_l)$ for all $|\psi\rangle \in \mathbb{C}S^2$. □

[21] Notice that in the mathematical literature [47], a Hermitian form is usually defined to be linear in the first complement, and conjugate linear in the second complement.

Yu-Tsung says:

Proof. This proof may not really work, because I haven't find a way to prove $|\bar{0}\rangle \in S^2$ in the real vector space spanned by $|0\rangle$, $|1'\rangle$, and $|2'\rangle$... >”<

Claim 1.

- $M = \max_{|\psi\rangle \in \mathbb{C}S^2} f(|\psi\rangle) = f(|0\rangle)$, $m = \min_{|\psi\rangle \in \mathbb{C}S^2} f(|\psi\rangle) = f(|\bar{0}\rangle)$,
- $|2\rangle = \sigma_{|0\rangle \times |\bar{0}\rangle} |0\rangle$, $|1\rangle = |2\rangle \times |0\rangle$,
- $Q(|\psi\rangle) = \langle\psi| \left[M |0\rangle\langle 0| + (W - M - m) |1\rangle\langle 1| + m |2\rangle\langle 2|\right] |\psi\rangle$,
- $\delta(\eta, W_\Delta, \omega_f) = \sup_{\theta \in [0, \frac{\pi}{2})} \min\left\{\min_{n \geq n(\theta)} 2n\eta + 4nW_\Delta + 2\omega_f\left(2 - 2\cos\left(\frac{\pi^2}{n}\right)\right), 2\omega_f(2\cos\theta)\right\}$, and
- $f'(|\psi\rangle) = f(|\psi\rangle) - Q(|\psi\rangle)$.

Then, $|Q(|\psi\rangle) - f(|\psi\rangle)| \leq 27\delta(W_r - W_l, W_r - W_l, \omega_f) + 27\delta(W_r - W_l, W_r - W_l, \omega_{f'}) + 18(W_r - W_l)$ for all $|\psi\rangle \in S^2$. □

Proof. For all vector $|\psi\rangle \in \mathbb{C}S^1$,

$$\begin{aligned}
& |Q(|\psi\rangle) - f(|\psi\rangle)| \\
= & \left|Q(|\psi\rangle) - f\left(\sum_{i=0}^{2} |i\rangle \langle i|\psi\rangle\right)\right| \\
= & \left|Q\left(\frac{|\langle 0|\psi\rangle|}{\langle 0|\psi\rangle} |\psi\rangle\right) - f\left(\frac{|\langle 0|\psi\rangle|}{\langle 0|\psi\rangle} \sum_{i=0}^{2} |i\rangle \langle i|\psi\rangle\right)\right| \\
= & \left|Q\left(\frac{|\langle 0|\psi\rangle|}{\langle 0|\psi\rangle} |\psi\rangle\right) - f\left(|0\rangle |\langle 0|\psi\rangle| + \sum_{i=1}^{2} |i\rangle \frac{|\langle 0|\psi\rangle|}{\langle 0|\psi\rangle} \frac{\langle i|\psi\rangle}{|\langle i|\psi\rangle|} |\langle i|\psi\rangle|\right)\right|
\end{aligned}$$

In order to estimate the last equation, we need to verify the conditions listed in theorem 7. First, $|0\rangle$, $|1'\rangle = |1\rangle \frac{|\langle 0|\psi\rangle|}{\langle 0|\psi\rangle} \frac{\langle 1|\psi\rangle}{|\langle 1|\psi\rangle|}$, and $|2'\rangle = |2\rangle \frac{|\langle 0|\psi\rangle|}{\langle 0|\psi\rangle} \frac{\langle 2|\psi\rangle}{|\langle 2|\psi\rangle|}$ form an orthonormal basis, and the restriction of f on the real vector space spanned by $|0\rangle$, $|1'\rangle$, and $|2'\rangle$ is still a approximation frame function of weight (W_l, W_r). Because $f(|0\rangle) = M$, we must also have $f(|0\rangle) = \max_{|\psi\rangle \in S^2} f(|\psi\rangle)$. □

□

Yu-Tsung says:

Proof.

Proof. This proof will be separated into two parts. The first part will estimate the error between the approximation frame function and the quadratic form on every complex two-space. The second part will then verify the quadratic forms in the first part could be connected and get a quadratic form on the whole space. □

Claim 2. For any pair of orthogonal vectors $|\psi_0\rangle, |\psi_1\rangle \in \mathbb{C}S^2$.

- Consider the complex vector space spanned by $|\psi_0\rangle$ and $|\psi_1\rangle$, and $\mathbb{C}S^1$ be its unit sphere, i.e.,
$$\left\{\alpha_0 |\psi_0\rangle + \alpha_1 |\psi_1\rangle \middle| \alpha_0, \alpha_1 \in \mathbb{C} \text{ and } |\alpha_0|^2 + |\alpha_1|^2 = 1\right\} .$$
- $|\psi_2\rangle = |\psi_0\rangle \times |\psi_1\rangle \in \mathbb{C}S^2$.
- $M = \max_{|\psi\rangle \in \mathbb{C}S^1} f(|\psi\rangle) = f(|0\rangle)$.
- $|1\rangle \in \mathbb{C}S^1$ is orthogonal to $|0\rangle$.
- $Q(|\psi\rangle) = M \langle 0|\psi\rangle^2 + (W - M) \langle 1|\psi\rangle^2$.
- $f'(|\psi\rangle) = f(|\psi\rangle) - Q(|\psi\rangle)$.
- $\delta(\eta, W_\Delta, \omega_f) = \sup_{\theta \in [0, \frac{\pi}{2})} \min\left\{\min_{n \geq n(\theta)} 2n\eta + 4nW_\Delta + 2\omega_f\left(2 - 2\cos\left(\frac{\pi^2}{n}\right)\right), 2\omega_f(2\cos\theta)\right\}$.

Then, $|Q(|\psi\rangle) - f(|\psi\rangle)| \leq 27\delta(W_r - W_l, W_r - W_l, \omega_f) + 27\delta(W_r - W_l, W_r - W_l, \omega_{f'}) + 18(W_r - W_l)$ for all $|\psi\rangle \in \mathbb{C}S^1$. □

□