Interval Probability for Fuzzy Quantum Theories

July 18, 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, 3].
- Instead of real-valued probability measures producing results in the uncountable and uncomputable interval $[0, 1]$, we will consider finite set-valued probability measures [4].

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 [5, 6] 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 [7, 8, 9].

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 $\mu\left({}_{\mathcal{O}}\sum_i P_i\right) = {}_{\mathbb{R}}\sum_i \mu\left(P_i\right)$ 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 ([10] or other sources. Need verify!), 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. We review the conventional presentation of probability spaces and then discuss the computational resources needed to estimate probabilities.

2.1 Real-Valued Probability Spaces

The conventional definition of a probability space [11, 12, 13, 14] builds upon the 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^Ω. Given the set of events $\mathcal{E}$, a *probability measure* is a function $\mu : \mathcal{E} \to [0, 1]$ such that:

- $\mu(\Omega) = 1$, and

- for a collection E_i of pairwise disjoint events, $\mu(\bigcup_i E_i) = {}_{\mathbb{R}}\sum_i \mu(E_i)$, where ${}_{\mathbb{R}}\sum_i \mu(E_i)$ explicitly specifies $\mu(E_i) \in \mathbb{R}$. 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.

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:

$$
\begin{array}{rclcrcl}
\mu(\emptyset) & = & 0 & & \mu(\{HT, TH\}) & = & 2/3 \\
\mu(\{HH\}) & = & 1/3 & & \mu(\{HT, TT\}) & = & 0 \\
\mu(\{HT\}) & = & 0 & & \mu(\{TH, TT\}) & = & 2/3 \\
\mu(\{TH\}) & = & 2/3 & & \mu(\{HH, HT, TH\}) & = & 1 \\
\mu(\{TT\}) & = & 0 & & \mu(\{HH, HT, TT\}) & = & 1/3 \\
\mu(\{HH, HT\}) & = & 1/3 & & \mu(\{HH, TH, TT\}) & = & 1 \\
\mu(\{HH, TH\}) & = & 1 & & \mu(\{HT, TH, TT\}) & = & 2/3 \\
\mu(\{HH, TT\}) & = & 1/3 & & \mu(\{HH, HT, TH, TT\}) & = & 1
\end{array}
$$

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 induced 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 for every event looks complex, a probability measure can be simply constructed by

$$\mu(E) = {}_{\mathbb{R}}\sum_{\omega \in E} \rho(\omega) \ , \tag{1}$$

where $\rho : \Omega \to [0, 1]$ and ${}_{\mathbb{R}}\sum_{\omega \in E} \rho(\omega) = 1$. ρ is called the Radon-Nikodym derivative of μ with respect to the counting measure, and the Radon-Nikodym theorem states that the converse is also true.

Theorem 1. [Radon-Nikodym theorem for finite probability space [15, 11, 16, 14]] For every probability measure μ, there exists an unique Radon-Nikodym derivative ρ with respect to the counting measure such that equation (1) holds. □

For example, the Radon-Nikodym derivative of μ in example 1 is:

$$\rho\,(HH) = 1/3, \quad \rho\,(HT) = 0, \quad \rho\,(TH) = 2/3, \quad \rho\,(TT) = 0 \ .$$

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 [17, 12, 18] in detail.[1]

[1] There is another possible approach that can be used to split the real interval $[0, 1]$ into a collection of subsets [4]

Amr says: need to explain the connection and why we are not using it.

2.2 Measuring Probabilities: Buffon's Needle Problem

In previous section, the probability of each event is known a priori. In reality, we seldom know much about events, but we could still assume each event E has a probability $\mu(E)$. If we want to know the probability $\mu(E)$, we could run N independent trials. Let x_i denote whether the event E occurs in the i-th trial or not, then $\mu(E)$ could be approximated to given accuracy by the relative frequency $\frac{1}{N}\sum_i x_i$ with the probability converging to one as N goes to infinity. This fact is called the law of large numbers [19, 11, 20, 12, 21].

Suppose we drop a needle of length ℓ onto a floor made of equally spaced parallel lines a distance h apart. It is a known fact that the probability of the needle crossing a line is $\frac{2\ell}{\pi h}$ [22, 23, 24, 20]. 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 X 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 $\frac{\ell}{h} = 0.75$ [24], 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 $\frac{\ell}{h} = 0.8$ [20], 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 $\frac{\ell}{h} = 0.6$ [23], 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 [25, 26]. In another word, to get a better approximation of the probability, the quality of experimental equipment cannot be compensated by the number of independent trials.

2.3 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 $[l_1 + l_2, r_1 + r_2]$ 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 $[1-r, 1-l]$ which we abbreviate $1 \ {}_{\mathscr{I}}\!- [l, r]$, where the preposing $\mathscr{I}$ specifies we subtracts 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 [18]:[2]

- $\bar{\mu}(\emptyset) = [0, 0]$,
- $\bar{\mu}(\Omega) = [1, 1]$,
- for any event E, $\bar{\mu}(\Omega \backslash E) = 1 \ {}_{\mathscr{I}}\!- \bar{\mu}(E)$, and
- for a collection E_i of pairwise disjoint events, we have $\bar{\mu}\left(\bigcup_i E_i\right) \subseteq \ {}_{\mathscr{I}}\!\sum_i \bar{\mu}(E_i)$, where ${}_{\mathscr{I}}\!\sum_i [l_i, r_i] = [\,{}_{\mathbb{R}}\!\sum_i l_i,\ {}_{\mathbb{R}}\!\sum_i r_i]$. We may drop the preposing $\mathscr{I}$ when summands are clearly intervals.

We will explain why the last condition is expressed using $\subseteq$ by a small example.

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 $\mathit{impossible}$, $[0, \frac{1}{2}]$ which we call $\mathit{unlikely}$, $[\frac{1}{2}, 1]$ which we call likely, and $[1, 1]$ which we call $\mathit{certain}$. 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:

$$
\begin{array}{rclrcl}
\bar{\mu}(\emptyset) & = & \mathit{impossible} & \bar{\mu}(\{HT, TH\}) & = & \mathit{likely} \\
\bar{\mu}(\{HH\}) & = & \mathit{unlikely} & \bar{\mu}(\{HT, TT\}) & = & \mathit{impossible} \\
\bar{\mu}(\{HT\}) & = & \mathit{impossible} & \bar{\mu}(\{TH, TT\}) & = & \mathit{likely} \\
\bar{\mu}(\{TH\}) & = & \mathit{likely} & \bar{\mu}(\{HH, HT, TH\}) & = & \mathit{certain} \\
\bar{\mu}(\{TT\}) & = & \mathit{impossible} & \bar{\mu}(\{HH, HT, TT\}) & = & \mathit{unlikely} \\
\bar{\mu}(\{HH, HT\}) & = & \mathit{unlikely} & \bar{\mu}(\{HH, TH, TT\}) & = & \mathit{certain} \\
\bar{\mu}(\{HH, TH\}) & = & \mathit{certain} & \bar{\mu}(\{HT, TH, TT\}) & = & \mathit{likely} \\
\bar{\mu}(\{HH, TT\}) & = & \mathit{unlikely} & \bar{\mu}(\{HH, HT, TH, TT\}) & = & \mathit{certain}
\end{array}
$$

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:

$$
\begin{array}{rl}
 & \bar{\mu}(\{HH\}) + \bar{\mu}(\{HT\}) + \bar{\mu}(\{TH\}) + \bar{\mu}(\{TT\}) \\
= & \mathit{impossible} + \mathit{unlikely} + \mathit{impossible} + \mathit{likely} \\
= & [0, 0] + \left[0, \dfrac{1}{2}\right] + [0, 0] + \left[\dfrac{1}{2}, 1\right] = \left[\dfrac{1}{2}, \dfrac{3}{2}\right]
\end{array}
$$

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 $\bar{\mu}(\emptyset) = \mathit{impossible}$, we have $\bar{\mu}(\Omega) = 1 - \bar{\mu}(\emptyset) = \mathit{certain} = [1, 1]$. This interval is tighter and a better estimate for the probability of the event Ω and of course it is contained in $[\frac{1}{2}, \frac{3}{2}]$. 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. $\square$

[3]

[2] The lower probability (belief) and the upper probability (plausibility) defined by Dempster [17] and Shafer [12] is a special case of the left-end and the right-end of the interval-valued probability measure, and the right-end of the interval-valued probability measure is a special case of capacity [27, 28, 29].

[3] Yu-Tsung says: The Radon-Nikodym Theorem for interval-valued probability?[10]?

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 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 $\mu : \mathcal{E} \to [0,1]$. We give an example before giving the formal definition.

Example 3 (One-qubit quantum probability space)*.* Consider a one-qubit Hilbert space with states $\alpha|0\rangle+\beta|1\rangle$ 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 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 $\mathbb{1} = |0\rangle\langle 0| + |1\rangle\langle 1|$ 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. □

To properly explain the previous example and generalize to arbitrary quantum systems, we formally discuss projection operators, define quantum probability measures, and extend to quantum interval-valued probability.

3.1 Quantum Events

Definition 1 (Projection Operators; Orthogonality [30, 9, 8, 13, 14])*.* Given a Hilbert space $\mathcal{H}$, an event[4] 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 can be defined recursively as follow: [5]

- $\mathbb{0}$ is a projection.
- For any pure state $|\psi\rangle$, $|\psi\rangle\langle\psi|$ is a projection operator.
- Projection operators P_1 and P_2 are *orthogonal* if $P_1 P_2 = P_2 P_1 = \mathbb{0}$. The sum of two projection operators $P_1 \,{}_{\mathscr{O}}\!+ P_2$ 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 ${}_{\mathscr{O}}\!\sum_j |\psi_j\rangle\langle\psi_j|$, where P actually projects onto the linear subspace S which has an orthonormal basis $\{|\psi_j\rangle\}$.

□

[4] An event is formally called an experimental proposition [31], a question [30, 32], or an elementary quantum test [8].

[5] "Projection" is sometimes called "orthogonal projection" or "self-adjoint projection" to emphasize $P^\dagger = P$ [13, 33].

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

Definition 2 (Ideal Measurement; Complement; Commutativity [14, 8, 13]).

- A set of mutually orthogonal projections P_i is called an *ideal measurement* if it is a partition of the identity, i.e., ${}_{\mathscr{O}}\sum_i P_i = \mathbb{1}$.
- 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.
- Projection operators P_1 and P_2 *commute* if $P_1 P_2 = P_2 P_1$. The product of two projection operators $P_1 P_2$ is also a projection operator if and only if they commute. This corresponds to the classical intersection between events.

□

3.2 Quantum Probability Measures

Definition 3 (Quantum Probability Measure [30, 7, 9, 33]). 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:[6]

- $\mu(\mathbb{1}) = 1$, and
- for mutually orthogonal projections P_i, we have $\mu\left({}_{\mathscr{O}}\sum_i P_i\right) = {}_{\mathbb{R}}\sum_i \mu(P_i)$.

□

In order to motivate the definition of quantum interval-valued probability measures later, we provide an equivalent definition of quantum probability measures.

Lemma 1. Given a Hilbert space $\mathcal{H}$, a function $\mu : \mathcal{E} \to [0, 1]$ is a quantum probability measure if and only if μ satisfies the following conditions:

- $\mu(\mathbb{0}) = 0$,
- $\mu(\mathbb{1}) = 1$,
- for any projection P, $\mu(\mathbb{1} \;{}_{\mathscr{O}}\!- P) = 1 \;{}_{\mathbb{R}}\!- \mu(P)$, and
- for a set of mutually orthogonal projections P_i, we have $\mu\left({}_{\mathscr{O}}\sum_i P_i\right) = {}_{\mathbb{R}}\sum_i \mu(P_i)$.

□

A set of events $\mathcal{E}$ together with quantum probability measure is called a *quantum probability space*. Comparing to the classical probability space, the empty set $\emptyset$ 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 $\mu(E_1 \cup E_2) = \mu(E_1) \;{}_{\mathbb{R}}\!+ \mu(E_2)$ is always defined, and it is true when E_1 and E_2 are disjoint; however, $\mu(P_1 \;{}_{\mathscr{O}}\!+ P_2) = \mu(P_1) \;{}_{\mathbb{R}}\!+ \mu(P_2)$ is always true whenever the left-handed side is defined.
[7]

[6] It is possible to define a more general space of events consisting of all operators $\mathcal{A}$ on $\mathcal{H}$ and consider $\mu : \mathcal{A} \to \mathbb{C}$ [33, 14]. When an operator $A \in \mathcal{A}$ is Hermitian, $\mu(A)$ is the expectation value of A. We does not take this approach because we want to focus only on probability.

[7] 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?)

A quantum probability measure can be easily constructed by states according to the Born rule [5, 6, 34]. For each pure normalized ($\langle\phi|\phi\rangle = 1$) quantum state $|\phi\rangle$, the Born rule induces a probability measure μ^B_ϕ as follows:

$$\mu^B_\phi(P) = \langle\phi|P\phi\rangle \ .$$

Moreover, the Born rule can be extended to a mixed state. Suppose we prepares a set of state $|\phi_j\rangle$ each with probability q_j, the state of the system can be expressed as a density matrix $\rho = \sum_j q_j|\phi_j\rangle\langle\phi_j|$, where $\sum_j q_j = 1$. It is natural that the quantum probability measure introduced by ρ is the combination of $\mu^B_{\phi_j}$ with respect to probability q_j [8, 21, 34]:

$$\mu^B_\rho(P) \quad = \quad \mathrm{Tr}\,(\rho P) = \sum_{j=1}^{N} q_j \mu^B_{\phi_j}(P) \ . \tag{2}$$

Recall that the Radon-Nikodym theorem claims every probability measure has an unique Radon-Nikodym derivative with respect to the counting measure. When we generalize the Radon-Nikodym theorem to quantum, the mixed state ρ can be considered as the Radon-Nikodym derivative with respect to the trace [35, 9, 36], and gives Gleason's theorem: in Hilbert spaces of dimension $d \geq 3$, given a quantum probability measure μ, there exist a mixed state ρ that induces such a measure using the Born rule [7, 9, 8]. By applying Gleason's theorem, we can extend lemma 1 to the following corollary.

Corollary 1. Given a Hilbert space $\mathcal{H}$ of dimension $d \geq 3$, for any function $\mu : \mathcal{E} \to [0, 1]$ satisfying the conditions listed in lemma 1, there exists an unique mixed state ρ such that $\mu = \mu^B_\rho$.

It is instructive to study counterexamples 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:

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

On one hand, μ is a quantum probability measure. Because μ is almost the same as a quantum probability measure $\mu^B_{|0\rangle}$, we only need to check the orthogonal pair $|+\rangle\langle+|$ and $|-\rangle\langle-|$:

$$\mu(|+\rangle\langle+|) + \mu(|-\rangle\langle-|) = 1 + 0 = 1 \ .$$

On the other hand, μ cannot be induced by any mixed state because

$$\mu(|+\rangle\langle+|) = \mu(|0\rangle\langle0|) = 1 \ .$$

However, $\mu^B_\rho(P) = 1$ if and only if ρ represents a pure state and $\rho = P$. □

3.3 Measuring Quantum Probabilities

Similar to the classical case, by applying the law of large number, 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 [37, 8, 21, 13], 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 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 $|1\rangle$. In reality, we need 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 does 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 is perfectly direct 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[8]. Similar to Buffon's needles, this kind of fuzziness can be taken into account by associating each quantum event with an interval-valued probability $[l, r]$.

Amr says: preparation fuzzy, device fuzzy, Meyer [38]

3.4 Quantum Interval-valued Probability Measures

As mentioned above, each quantum event P will be associated with an interval-valued probability $[l, r]$. 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 claimed she is tossing a coin behind a veil, and only show the resulting heads or tails, we cannot distinguish whether she has really tossed a coin, or she has run an Stern-Gerlach experiment, and show 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. Therefore, we should plug in the definition of interval-valued probability measure into the definition of quantum probability space for each ideal measurement, and get the following definition.

Definition 4 (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, $\bar{\mu}\left(\mathbb{1} \;_{\mathscr{E}}\!- P\right) = 1 \;_{\mathscr{I}}\!- \bar{\mu}\left(P\right)$, and
- for a set of mutually orthogonal projections P_i, we have $\bar{\mu}\left({}_{\mathscr{E}}\!\sum_i P_i\right) \subseteq {}_{\mathscr{I}}\!\sum_i \bar{\mu}\left(P_i\right)$.

As before, we will only define a quantum interval-valued probability measure with preposing $\mathscr{I}$ when we want to compare a quantum probability measure and a quantum interval-valued probability measure. □

Notice that if the last condition is replaced by the equal sign, these conditions will be exactly the same as the conditions in lemma 1 which is equivalent to the Born rule as in corollary 1. Therefore, it is reasonable to believe these conditions should be equivalent to a interval-valued Born rule.

When we pick $\mathscr{I} = \{[x, x] | x \in [0, 1]\}$, for any quantum probability measure $\mu : \mathcal{E} \to [0, 1]$, we can define a corresponding quantum $\mathscr{I}$-interval-valued probability measure $\bar{\mu} : \mathcal{E} \to \mathscr{I}$ by $\bar{\mu}\left(P\right) = \left[\mu\left(P\right), \mu\left(P\right)\right]$. In this sense, every usual quantum probability measure can be considered as a quantum interval-valued probability measure. There are more interesting examples as we now show.

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 $\mathscr{impossible}$ and $\mathscr{certain}$ as before, and $[0, 1]$ is called $\mathscr{unknown}$ 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

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

ι has two interesting properties

$$\begin{aligned} \iota\left(\mathbb{1} \;_{\mathbb{R}}\!- x\right) &= 1 \;_{\mathscr{I}}\!- \iota\left(x\right) \\ \iota\left({}_{\mathbb{R}}\!\sum_i x_i\right) &\subseteq {}_{\mathscr{I}}\!\sum_i \iota\left(x_i\right) \ , \end{aligned}$$

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

where x and ${}_{\mathbb{R}}\sum_i x_i \in [0,1]$. By applying these two properties, it is easy to verify $\bar{\mu}$ is a quantum $\mathscr{I}$-interval-valued probability measure. $\square$

Although $\iota : [0,1] \to \mathscr{I}$ maps every quantum probability measure to a quantum interval-valued probability measure, there is a quantum $\mathscr{I}$-interval-valued probability measure which does not come from a quantum probability measure.

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 $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. Let

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

To verify $\bar{\mu}$ is a quantum interval-valued probability measure, it is sufficient to check $\bar{\mu}\left(|\psi_1\rangle\langle\psi_1| \ {}_{\mathscr{O}}\!+ |\psi_2\rangle\langle\psi_2|\right) \subseteq \bar{\mu}\left(|\psi_1\rangle\langle\psi_1|\right) \ {}_{\mathscr{I}}\!+ \bar{\mu}\left(|\psi_2\rangle\langle\psi_2|\right)$ for orthogonal $|\psi_1\rangle$ and $|\psi_2\rangle$. This is easily verified by enumerating all possible cases.

$\bar{\mu}$ cannot correspond to any quantum probability measure. Suppose $\bar{\mu}(P) = \iota\left(\mu(P)\right)$ for some $\mu : \mathcal{E} \to [0,1]$. We must have

$$\mu(|0\rangle\langle 0|) = \mu(|+\rangle\langle +|) = 0\ . \tag{3}$$

By Gleason's theorem, there is a mixed state $\rho = \sum_j q_j |\phi_j\rangle\langle\phi_j|$ such that $\mu\left(P\right) = \sum_{j=1}^N q_j \langle\phi_j|P\phi_j\rangle$, where $\sum_j q_j = 1$ and $q_j > 0$. The only pure state $|\phi\rangle$ satisfies

$$\langle\phi|0\rangle = \langle\phi|+\rangle = 0$$

is $|2\rangle$, but clearly $\bar{\mu}(P) \neq \iota\left(\mu^B_{|2\rangle}(P)\right)$. $\square$

Notice that in the above two examples, quantum interval-valued probability measures are all come from quantum probability measures because the numbers in $[0,1]$ can be mapped to the chosen intervals $\mathscr{I}$ naturally. This is not always the case. For example, if $\mathscr{I} = \{\mathit{impossible}, \mathit{unlikely}, \mathit{likely}, \mathit{certain}\}$, there is no natural way to map from $\frac{1}{2}$, as both $\frac{1}{2} \in \left[0, \frac{1}{2}\right] = \mathit{unlikely}$ and $\frac{1}{2} \in \left[\frac{1}{2}, 1\right] = \mathit{likely}$. However, we still can find a quantum interval-valued probability measure corresponding to a quantum probability measure as in the following example.

Example 7 (One-qubit quantum interval-valued probability measure). Given a two dimensional Hilbert space, a vector $|\psi\rangle$ can be normalized and written in $\begin{pmatrix} \cos\theta \\ \mathrm{e}^{\mathrm{i}\gamma}\sin\theta \end{pmatrix}$, where $0 \leq \theta \leq \frac{\pi}{2}$ and $0 \leq \gamma < 2\pi$. Consider a quantum interval-valued probability measure $\bar{\mu} : \mathcal{E} \to \mathscr{I}$ defined by

$$\begin{aligned} \bar{\mu}(\mathbb{0}) &= \mathit{impossible}\,, \\ \bar{\mu}(\mathbb{1}) &= \mathit{certain}\,, \\ \bar{\mu}(|\psi\rangle\langle\psi|) &= \begin{cases} \mathit{certain} & \text{, if } \theta = 0\ ; \\ \mathit{likely} & \text{, if } 0 < \theta < \frac{\pi}{4}\ ; \\ \mathit{likely} & \text{, if } \theta = \frac{\pi}{4} \text{ and } 0 \leq \gamma < \pi\ ; \\ \mathit{unlikely} & \text{, if } \theta = \frac{\pi}{4} \text{ and } \pi \leq \gamma < 2\pi\ ; \\ \mathit{unlikely} & \text{, if } \frac{\pi}{4} < \theta < \frac{\pi}{2}\ ; \\ \mathit{impossible} & \text{, if } \theta = \frac{\pi}{2}\ . \end{cases} \end{aligned}$$

Let $|\psi^\perp\rangle$ be the state perpendicular to $|\psi\rangle$. Then, the vector representation of $|\psi^\perp\rangle$ can be simplified as

$$\begin{pmatrix} -\mathrm{e}^{-\mathrm{i}\gamma}\sin\theta \\ \cos\theta \end{pmatrix} = -\mathrm{e}^{-\mathrm{i}\gamma} \begin{pmatrix} \cos\left(\frac{\pi}{2} - \theta\right) \\ \mathrm{e}^{\mathrm{i}(\pi+\gamma)}\sin\left(\frac{\pi}{2} - \theta\right) \end{pmatrix}$$

so we have

$$\bar{\mu}\left(\mathbb{1} \underset{\mathscr{O}}{-} |\psi\rangle\langle\psi|\right) = \bar{\mu}\left(|\psi^{\perp}\rangle\langle\psi^{\perp}|\right) = 1 \underset{\mathscr{I}}{-} \bar{\mu}\left(|\psi\rangle\langle\psi|\right) \ .$$

Hence, $\bar{\mu}$ is a quantum $\mathscr{I}$-interval-valued probability measure. Notice that for all projection P, $\mu^{B}_{|0\rangle}(P) \in \bar{\mu}(P)$ implies that $\bar{\mu}(P)$ is one of quantum $\mathscr{I}$-interval-valued probability measures corresponding to $\mu^{B}_{|0\rangle}(P)$. However, $\mu^{B}_{|0\rangle}(P)$ can correspond to infinity many quantum $\mathscr{I}$-interval-valued probability measures by twisting the definition involving γ. □

We are going to provide an quantum 4-value interval-valued probability measures corresponding to no quantum probability measure, i.e., $\mu^{B}_{\rho}(P) \notin \bar{\mu}(P)$. In general, we believe this kind of quantum $\mathscr{I}$-interval-valued probability measure exists for any finite set of interval $\mathscr{I}$.

Example 8 (Three-dimensional quantum 4-value interval-valued probability measure). Given a three dimensional Hilbert space with an orthonormal basis $\{|0\rangle, |1\rangle, |2\rangle\}$. Consider $\mathscr{I} = \left\{\mathit{impossible}, \mathit{unlikely}, \mathit{likely}, \mathit{certain}\right\}$, $|{+}\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, $|{-}\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, and a quantum $\mathscr{I}$-interval-valued probability measure $\bar{\mu}$ defined as follow:

1. $\bar{\mu}(\mathbb{0}) = \bar{\mu}(|0\rangle\langle 0|) = \bar{\mu}(|{+}\rangle\langle{+}|) = \mathit{impossible}$ and $\bar{\mu}(\mathbb{1}) = \bar{\mu}(\mathbb{1} - |0\rangle\langle 0|) = \bar{\mu}(\mathbb{1} - |{+}\rangle\langle{+}|) = \mathit{certain}$.
2. Consider the states orthogonal to $|0\rangle$ and $|{+}\rangle$ respectively, $|0^{\perp}_{\theta,\gamma}\rangle = \mathrm{e}^{\mathrm{i}\gamma}\sin\theta|1\rangle + \cos\theta|2\rangle$ and $|{+}^{\perp}_{\theta,\gamma}\rangle = -\mathrm{e}^{\mathrm{i}\gamma}\sin\theta|{-}\rangle + \cos\theta|2\rangle$, where $0 \leq \theta \leq \frac{\pi}{2}$ and $0 \leq \gamma < 2\pi$. Half of those states need special treatment, i.e., when where $0 \leq \theta < \frac{\pi}{2}$ and $0 \leq \gamma < \pi$, we define $\bar{\mu}(|0^{\perp}_{\theta,\gamma}\rangle\langle 0^{\perp}_{\theta,\gamma}|) = |{+}^{\perp}_{\theta,\gamma}\rangle = \mathit{likely}$ and $\bar{\mu}(\mathbb{1} - |0^{\perp}_{\theta,\gamma}\rangle\langle 0^{\perp}_{\theta,\gamma}|) = \bar{\mu}(\mathbb{1} - |{+}^{\perp}_{\theta,\gamma}\rangle\langle{+}^{\perp}_{\theta,\gamma}|) = \mathit{unlikely}$.
3. Otherwise, $\bar{\mu}(|\psi\rangle\langle\psi|) = \mathit{unlikely}$ and $\bar{\mu}(\mathbb{1} - |\psi\rangle\langle\psi|) = \mathit{likely}$.

$\bar{\mu}$ cannot correspond to any quantum probability measure because of the exact same reason as in example 6. Before we verify that $\bar{\mu}$ is a quantum interval-valued probability measure, we need a notation for cross product. 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 $|\psi \times \phi\rangle$ need not be normalized as usual.

To verify $\bar{\mu}$ is a quantum interval-valued probability measure, it is sufficient to check

$$\begin{aligned}\bar{\mu}\left(\mathbb{1} \underset{\mathscr{O}}{-} |\psi_1\rangle\langle\psi_1|\right) &\subseteq \bar{\mu}\left(|\psi_2\rangle\langle\psi_2|\right) \underset{\mathscr{I}}{+} \bar{\mu}\left(|\psi_3\rangle\langle\psi_3|\right)\\ \bar{\mu}\left(\mathbb{1} \underset{\mathscr{O}}{-} |\psi_2\rangle\langle\psi_2|\right) &\subseteq \bar{\mu}\left(|\psi_3\rangle\langle\psi_3|\right) \underset{\mathscr{I}}{+} \bar{\mu}\left(|\psi_1\rangle\langle\psi_1|\right)\\ \bar{\mu}\left(\mathbb{1} \underset{\mathscr{O}}{-} |\psi_3\rangle\langle\psi_3|\right) &\subseteq \bar{\mu}\left(|\psi_1\rangle\langle\psi_1|\right) \underset{\mathscr{I}}{+} \bar{\mu}\left(|\psi_2\rangle\langle\psi_2|\right)\end{aligned} \tag{4}$$

for every orthonormal basis $\mathcal{B} = \{|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle\}$. We are going to enumerate all possible orthonormal bases to verify the equations, and case 1 and 2 are illustrated in figure 1.

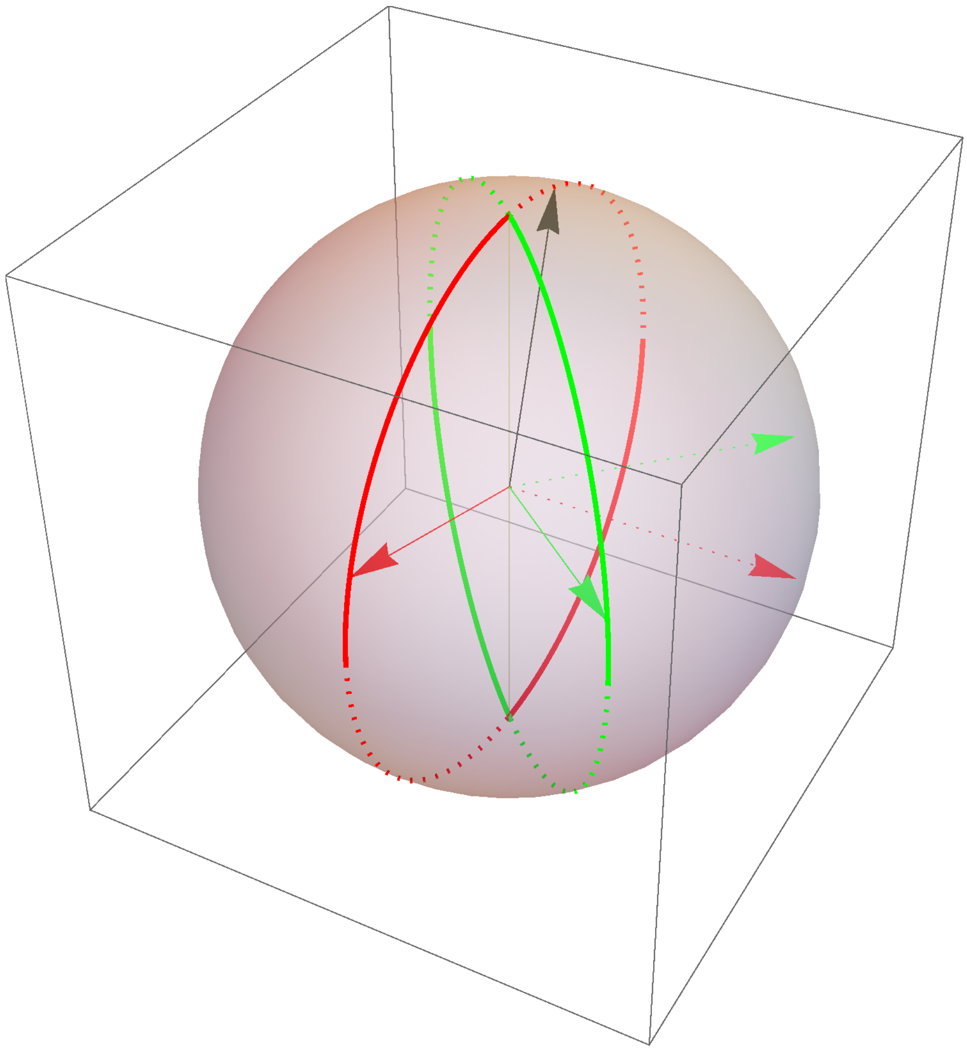

Figure 1: This figure illustrates case 1 and 2 in $\mathbb{R}^3$ in example 8. The red and green dotted vector are $|0\rangle$ and $|+\rangle$ respectively. All possible vectors of $|0^{\perp}_{\theta,\gamma}\rangle$ and $|+^{\perp}_{\theta,\gamma}\rangle$ are drawn in the red and green circles, respectively. 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 $|0^{\perp}_{\theta,\gamma}\rangle$, and the red and green solid vectors are normalized $|0\rangle \times |0^{\perp}_{\theta,\gamma}\rangle$ and $|+\rangle \times |0^{\perp}_{\theta,\gamma}\rangle$, respectively.

1. When $|\psi_1\rangle$ is $|0\rangle$, then $|\psi_2\rangle$ and $|\psi_3\rangle$ must be $|0^{\perp}_{\theta,\gamma}\rangle$ for some θ and γ. Because of

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

the pair $|\psi_2\rangle = |0^{\perp}_{\theta,\gamma}\rangle$ and $|\psi_3\rangle = |0^{\perp}_{\frac{\pi}{2}-\theta,\pi+\gamma}\rangle$ for ($0 < \theta < \frac{\pi}{2}$ and $0 \leq \gamma < \pi$) or $\theta = 0$ enumerate all possible orthonormal bases. Equation (4) can be verified as follow.

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

Similarly, when $|\psi_1\rangle$ is $|+\rangle$, equation (4) holds.

2. When $|0\rangle \notin \mathcal{B}$ and $|\psi_1\rangle = |0^{\perp}_{\theta,\gamma}\rangle$ for $0 \leq \theta < \frac{\pi}{2}$ and $0 \leq \gamma < \pi$, we want to prove $\bar{\mu}\left(|\psi_2\rangle\langle\psi_2|\right) = \bar{\mu}\left(|\psi_3\rangle\langle\psi_3|\right) = \mathit{unlikely}$. Then, we have

$$\begin{aligned}
\bar{\mu}\left(\mathbb{1}_{\mathscr{O}} - |0^{\perp}_{\theta,\gamma}\rangle\langle 0^{\perp}_{\theta,\gamma}|\right) = \mathit{unlikely} \quad &\subseteq \quad \mathit{unlikely}\ {}_{\mathscr{I}}\!+ \mathit{unlikely} = \bar{\mu}\left(|\psi_2\rangle\langle\psi_2|\right)\ {}_{\mathscr{I}}\!+ \bar{\mu}\left(|\psi_3\rangle\langle\psi_3|\right)\\
\bar{\mu}\left(\mathbb{1}_{\mathscr{O}} - |\psi_2\rangle\langle\psi_2|\right) = \mathit{likely} \quad &\subseteq \quad \mathit{unlikely}\ {}_{\mathscr{I}}\!+ \mathit{likely} = \bar{\mu}\left(|\psi_3\rangle\langle\psi_3|\right)\ {}_{\mathscr{I}}\!+ \bar{\mu}\left(|0^{\perp}_{\theta,\gamma}\rangle\langle 0^{\perp}_{\theta,\gamma}|\right)
\end{aligned}$$

In order to verify $\bar{\mu}\left(|\psi_2\rangle\langle\psi_2|\right) = \bar{\mu}\left(|\psi_3\rangle\langle\psi_3|\right) = \mathit{unlikely}$, it is sufficient to prove that $|+^{\perp}_{\theta',\gamma'}\rangle \in \mathcal{B}$ implies ($0 < \theta' < \frac{\pi}{2}$ and $\pi \leq \gamma' < 2\pi$) or $\theta' = \frac{\pi}{2}$. Recall $\langle +|+^{\perp}_{\theta',\gamma'}\rangle = 0$. Hence, if $|+^{\perp}_{\theta',\gamma'}\rangle \in \mathcal{B}$, we have $|+^{\perp}_{\theta',\gamma'}\rangle$ parallel to $|+\rangle \times |0^{\perp}_{\theta,\gamma}\rangle$.

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

Because of $0 \leq \theta < \frac{\pi}{2}$ and $0 \leq \gamma < \pi$, we have ($0 < \theta' = \frac{\pi}{2} - \theta < \frac{\pi}{2}$ and $\pi \leq \gamma' = \pi + \gamma < 2\pi$) or $\theta' = \frac{\pi}{2} - \theta = \frac{\pi}{2}$. Similarly, when $|\psi_1\rangle$ is $|+^{\perp}_{\theta,\gamma}\rangle$ and $|+\rangle \notin \mathcal{B}$, equation (4) holds.

3. Finally, when $|0\rangle \notin \mathcal{B}$, $|+\rangle \notin \mathcal{B}$, $|0^{\perp}_{\theta,\gamma}\rangle \notin \mathcal{B}$, and $|+^{\perp}_{\theta,\gamma}\rangle \notin \mathcal{B}$ for $0 \leq \theta < \frac{\pi}{2}$ and $0 \leq \gamma < \pi$, i.e., the "otherwise" case. Then, equation (4) can easily be verified.

$$\bar{\mu}\left(\mathbb{1}_{\mathscr{O}} - |\psi_1\rangle\langle\psi_1|\right) = \mathit{likely} \quad \subseteq \quad \mathit{unlikely}\ {}_{\mathscr{I}}\!+ \mathit{unlikely} = \bar{\mu}\left(|\psi_2\rangle\langle\psi_2|\right)\ {}_{\mathscr{I}}\!+ \bar{\mu}\left(|\psi_3\rangle\langle\psi_3|\right)$$

$\square$

Amr says: We can use DQC if we have some kind of topology (distances). The idea will be that we want to prepare state PSI but because of errors etc we prepare a close state. Well the next closest state will be the next state in our discrete grid. I am sure that a state that's very close to PSI can involve some wrapping around.

Yu-Tsung says: About continuity, one of the confusion comes from the subtle correspondence between the measurement processes and observables. Consider the following observables and their measurement processes:

- For any $x \in \mathbb{R}$, consider the observable operator $\mathbf{O}_x = |1\rangle\langle 1| + x|-1\rangle\langle -1|$. Except $x = 1$, given the system in the state $|\phi\rangle$, we can denote the measurement process for $\mathbf{O}_x$ by $\mathcal{O}_x$:
  1. With probability $\mu_\phi(|1\rangle\langle 1|)$, return the measurement result 1 with the post-measurement state $|1\rangle$;
  2. With probability $\mu_\phi(|0\rangle\langle 0|)$, return the measurement result 0 with the post-measurement state $|0\rangle$;
  3. With probability $\mu_\phi(|-1\rangle\langle -1|)$, return the measurement result x with the post-measurement state $|-1\rangle$.
- The observable operator $\mathbf{J}_z^2 = |1\rangle\langle 1| + |-1\rangle\langle -1|$, and its measurement process $\mathcal{J}_z^2$:
  1. With probability $\mu_\phi(|1\rangle\langle 1| + |-1\rangle\langle -1|)$, return the measurement result 1 with the post-measurement state $|1\rangle\langle 1|\phi\rangle + |-1\rangle\langle -1|\phi\rangle$;
  2. With probability $\mu_\phi(|0\rangle\langle 0|)$, return the measurement result 0 with the post-measurement state $|0\rangle$.

As an operator, it is clear that $\lim_{x\to 1} \mathbf{O}_x = \mathbf{J}_z^2$, and we should have $\lim_{x\to 1} \mathcal{O}_x = \mathcal{O}_1$ for the measurement process. However, it is not obvious why $\mathcal{O}_1 = \mathcal{J}_z^2$. Although they have the same probability for the measurement results, but their post-measurement states are different. This might be a defect of representing a measurement process as an observable. As $\mathbf{O}_1 = \mathbf{J}_z^2$, there is no observable corresponding to the measurement process $\mathcal{O}_1$.

Amr says: the rest needs cleaning up and perhaps does not even belong in this section

Although it seems that we need an infinite long table to specify the quantum probability measure μ, our μ is actually given by a simple formula $\langle 0|P|0\rangle$. In general, Born discovered each quantum state $|\psi\rangle \in \mathcal{H}\backslash\{0\}$ induces a probability measure $\mu_\psi^B : \mathcal{E} \to [0,1]$ on the space of events defined for any event $P \in \mathcal{E}$ as follows [5, 6]:

$$\mu_\psi^B(P) = \frac{\langle\psi|P|\psi\rangle}{\langle\psi|\psi\rangle} \tag{5}$$

The Born rule satisfies the following properties:

- It can be extend to mixed states. Given a mixed state represented by a density matrix $\rho = \sum_{j=1}^N q_j \frac{|\psi_j\rangle\langle\psi_j|}{\langle\psi_j|\psi_j\rangle}$, where $\sum_{j=1}^N q_j = 1$, i.e., $\mathrm{Tr}(\rho) = 1$, then the Born rule can be extended to ρ by

  $$\mu_\rho^B(P) \quad = \quad \mathrm{Tr}(\rho P) = \sum_{j=1}^N q_j \mu_{\Psi_j}^B(P) \ . \tag{6}$$

  Notice that $\left(\{1,\ldots,N\}, 2^{\{1,\ldots,N\}}, \mu(J) = \sum_{j\in J} q_j\right)$ is a classical probability space. Therefore, when we discretize the Hilbert space later, we may need to discretize this probability space as well.
- μ_ρ^B is a probability measure for all mixed state ρ.
- $\langle\psi|\phi\rangle = 0 \Leftrightarrow \mu_\psi^B(|\phi\rangle\langle\phi|) = 0$.
- $\mu_\psi^B(P) = \mu_{\mathbf{U}|\psi\rangle}^B\left(\mathbf{U}P\mathbf{U}^\dagger\right)$,where $\mathbf{U}$ is any unitary map, i.e., $\mathbf{U}^\dagger\mathbf{U} = \mathbb{1}$.

Naturally, we may ask: is every probability measure induced from a state by the Born rule? The answer is yes by Gleason's theorem when the dimension ≥ 3 [7, 8, 9]. Furthermore, a simple corollary of Gleason's theorem can show the Born rule is the unique function satisfying conditions 1. to 3.

Corollary 2. The Born rule is the unique function satisfying conditions 1. to 3.

Proof. Assume there is another function μ'^B such that μ'^B_ρ is a quantum probability measure for all mixed state ρ. We are going to prove $\mu'^B = \mu^B$.

Fix a pure normalized state ϕ, μ'^B_ϕ is a quantum probability measure by condition 2. By Gleason's theorem, there is a mixed state ρ', such that $\mu'^B_\phi(P) = \mathrm{Tr}(\rho' P) = \sum_{j=1}^N q_j \mu^B_{\psi_j}(P)$ for all event P.

Consider the event $P' = \mathbb{1} - |\phi\rangle\langle\phi|$, we have

$$\begin{aligned} 0 &\overset{\text{Condition 3}}{=} \mu^B_\phi(P') \\ &= \sum_{j=1}^N q_j \mu^B_{\psi_j}(P') \end{aligned}$$

Because $q_j > 0$, we have $\mu^B_{\psi_j}(P) = 0$, i.e., ψ_j is orthogonal to a co-dimension-1 subspace P'. However, the only subspace orthogonal to P' is span by $|\phi\rangle$. Hence, $\mu'^B_\phi = \mu^B_\phi$. □

4 All Continuous or All Discrete

Before we turn to the main part of the paper, we quickly dismiss the possibility of having one but not the other of the discrete variations. Specifically, it is impossible to maintain the Hilbert space and have a finite set-valued probability measure and it is also impossible to have a vector space constructed over a finite field with a real-valued probability measure.

4.1 Hilbert Space with Finite Set-Valued Probability Measure

However, there is a $\mathscr{L}_2$-valued probability measure

$$\hat{\mu}_1(P) = \begin{cases} \mathit{impossible} & \text{, if } P = |+\rangle\langle+|; \\ \bar{\mu}(P) & \text{, otherwise.} \end{cases}$$

such that $\hat{\mu}_1 \neq \bar{\mu}_\psi$ for all mixed state $|\psi\rangle$.

4.2 Discrete Vector Space with Real-Valued Probability Measure

References

[1] Andrew J Hanson, Gerardo Ortiz, Amr Sabry, and Yu-Tsung Tai. Corrigendum: Geometry of discrete quantum computing. *Journal of Physics A: Mathematical and Theoretical*, 49(3):039501, 2015.

[2] Andrew J Hanson, Gerardo Ortiz, Amr Sabry, and Yu-Tsung Tai. Discrete quantum theories. *Journal of Physics A: Mathematical and Theoretical*, 47(11):115305, 2014.

[3] Andrew J Hanson, Gerardo Ortiz, Amr Sabry, and Yu-Tsung Tai. Geometry of discrete quantum computing. *Journal of Physics A: Mathematical and Theoretical*, 46(18):185301, 2013.

[4] Madan L Puri and Dan A Ralescu. Strong law of large numbers with respect to a set-valued probability measure. *The Annals of Probability*, pages 1051–1054, 1983.

[5] Max Born. "zur quantenmechanik der stoßvorgänge," zeitschrift für physik, 37, 863-67 (1926). *English Translation, Wheeler and Zurek, in Quantum Theory and Measurement*, pages 48–52, 1983.

[6] N. D. Mermin. *Quantum Computer Science*. Cambridge University Press, 2007.

[7] Andrew Gleason. Measures on the closed subspaces of a Hilbert space. *Indiana Univ. Math. J.*, 6:885–893, 1957.

[8] A. Peres. *Quantum Theory: Concepts and Methods*. Fundamental Theories of Physics. Springer, 1995.

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

[10] Itzhak Gilboa and David Schmeidler. Additive representations of non-additive measures and the Choquet integral. *Annals of Operations Research*, 52(1):43–65, 1994.

[11] Andrei Nikolaevich Kolmogorov. Grundbegriffe der wahrscheinlichkeitsrechnung, Berlin, 1933. *English translation, Chelsea, New York*, 1950.

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

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

[14] Jan Swart. Introduction to quantum probability. *Lecture Notes*, 2013.

[15] Otton Nikodym. Sur une généralisation des intégrales de M. J. Radon. *Fundamenta Mathematicae*, 15(1):131–179, 1930.

[16] Gerald B Folland. *Real analysis: modern techniques and their applications*. John Wiley & Sons, 2013.

[17] Arthur P Dempster. Upper and lower probabilities induced by a multivalued mapping. *The Annals of Mathematical Statistics*, pages 325–339, 1967.

[18] Kenneth David Jamison and Weldon A Lodwick. *Interval-Valued Probability Measures*, volume 213 of *Center for Computational Mathematics Reports Series*. Department of Mathematics, University of Colorado at Denver, 2004.

[19] J Bernoulli. Ars conjectandi [the art of conjecturing], 1713. *English translation and Notes, Sylla, Baltimore*, 2006.

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

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

[22] Comte de Buffon, Georges Louis Leclerc. Essai d'arithmétique morale. In *Histoire naturelle, générale et particulière, Supplément*, volume 4, pages 46–109. De l'Imprimerie royale, an 12, 1777.

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

[24] Asaph Hall. On an experimental determination of π. *Messeng. Math*, 2:113–114, 1873.

[25] Massimo Marinacci. Limit laws for non-additive probabilities and their frequentist interpretation. *Journal of Economic Theory*, 84(2):145–195, 1999.

[26] Pedro Terán. Laws of large numbers without additivity. *Transactions of the American Mathematical Society*, 366(10):5431–5451, 2014.

[27] Gustave Choquet. Theory of capacities. In *Annales de l'institut Fourier*, volume 5, pages 131–295. Institut Fourier, 1954.

[28] Siegfried Graf. A radon-nikodym theorem for capacities. *Journal für die reine und angewandte Mathematik*, 320:192–214, 1980.

[29] Irwin R Goodman, Ronald P Mahler, and Hung T Nguyen. *Mathematics of data fusion*, volume 37. Springer Science & Business Media, 2013.

[30] George W. Mackey. Quantum mechanics and Hilbert space. *The American Mathematical Monthly*, 64(8):45–57, 1957.

[31] Garrett Birkhoff and John Von Neumann. The logic of quantum mechanics. *Annals of mathematics*, pages 823–843, 1936.

[32] Samson Abramsky. Big toy models: Representing physical systems as Chu spaces. *CoRR*, abs/0910.2393, 2009.

[33] Hans Maassen. Quantum probability and quantum information theory. In *Quantum information, computation and cryptography*, pages 65–108. Springer, 2010.

[34] Sebastian Rieder and Karl Svozil. Probability distributions and Gleason's theorem. In *Foundations of Probability and Physics- 4(AIP Conference Proceedings Volume 889)*, volume 889, pages 235–242, 2007.

[35] Samuel S Holland Jr. The current interest in orthomodular lattices. In James Crawford Abbott, editor, *Trends in lattice theory*, pages 41 – 126. Van Nostrand Reinhold Company, 1970.

[36] Gregg Jaeger. *Quantum information*. Springer, 2007.

[37] Otto Stern. A way towards the experimental examination of spatial quantisation in a magnetic field. *Zeitschrift für Physik D Atoms, Molecules and Clusters*, 10(2):114–116, 1988.

[38] David Meyer. Finite precision measurement nullifies the Kochen-Specker theorem. *Phys. Rev. Lett.*, 83:3751–3754, Nov 1999.