Discrete Probability for Discrete Quantum Theories

May 7, 2016

1 Introduction

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 several variations that avoid using the real interval $[0, 1]$.

2.1 Real-Valued Probability Spaces

The conventional definition of a probability space [1, 2, 3] 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, $\mu(\bigcup_i E_i) = \sum_i \mu(E_i)$.

Example 1 (Two-coin probability space). 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:

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

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. □

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 finite set of elements computable with a fixed set of resources. We expand on this observation and then consider two approaches from the literature: set-valued probability measures [4, 5] and interval-valued probability measures [6, 7, 8, 9].

2.2 Buffon's Needle Problem

Suppose we drop a needle of length ℓ onto a floor made of equally spaced parallel lines a distance ℓ apart. It is a known fact that the probability of the needle crossing a line is $\frac{2}{\pi} = 0.6366197723675814$.

Yu-Tsung says: Add . . . at the end?

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 and observe the number X of needles that cross a line. The sample space can be expressed as the set $(X|-)^N$

Yu-Tsung says: Use more physical or mathematical notation?

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 event that exactly 2 needles cross lines is $\{-XX, X-X, XX-\}$: it has probability $\frac{2}{3}$. Generally, the event that exactly M needles out of the N total needles cross lines is $\frac{M}{N}$.

In an actual experiment with 500 needles, it was found that 321 crossed a line which gives a probability of 0.642. In a larger experiment with 3408 needles, the probability was 0.6366197193098592.

Yu-Tsung says: Add . . . at the end?

Comparing the values to the idealized mathematical value, we see that the observed probability approaches $\frac{2}{\pi}$ but only if *larger and larger resources* are expended.

These resource considerations thus lead us to replace the real interval $[0, 1]$ with rational numbers up to a certain precision. The intuition is that the precision depends on the amount of resources allocated for the measurement process.

There is however another more subtle source 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 ℓ. and the needles lengths are not all equal and are not all absolutely equal to ℓ. These perturbations make the events "fuzzy." For example, instead of talking about the idealized event that exactly M needles cross lines, we might instead talk about the event that $M - \delta$ needles evidently cross lines and $M + \delta'$ needles plausibly cross lines. This fuzzy notion of events leads us to consider intervals of probability reflecting the certainty of the event and its plausibility.

2.3 Set-valued Probability Measures

Instead of using every point in the real interval $[0, 1]$ as a potential value for a probability measure, we can lump points together in sets and consider probability measures up to set membership. The simplest such situation is to partition the interval $[0, 1]$ into two sets: the set $\{0\}$ (which we will call *impossible*) and the set containing all the points in the half-open interval $(0, 1]$ (which we will call *possible*). We will call the resulting collection $\{$*impossible*, *possible*$\}$, the collection $\mathscr{L}_2$. To completely specify a probability space, we also need to specify how to "add" the probabilities in $\mathscr{L}_2$. Given two subsets of the real interval, I_1 and I_2, the general formula for adding them is:[1]

$$I_1 \vee I_2 = \{x + y \mid x \in I_1, y \in I_2\} \ \cap \ [0,1] \ .$$

In the case of $\mathscr{L}_2$, this reduces to $x \vee y =$ *impossible* iff $x = y =$ *impossible*.

Amr says: From Puri and Ralescu's paper: Fix $n = 1$ (only 1-dimension). Conditions on μ:

- $\mu(E) \neq \emptyset$ for every $E \in \mathcal{E}$,
- $1 \in \mu(\Omega)$,
- for every disjoint family A_i we have, $\mu(\bigcup_i A_i) = \sum_i \mu(A_i)$

where the sum of subsets of $\mathbb{R}$ (not $\mathbb{R}^n$ as in the paper) is as follows. Let X and Y be subsets of $\mathbb{R}$, then $X + Y = \{x + y \mid x \in X, y \in Y\}$

[1] Notice that the set-valued measure in [4, 5], only require $I_1 + I_2 = \{a_1 + a_2 | a_1 \in I_1, a_2 \in I_2\}$ because they focused on measure theory. In probability theory, $2 \in \mu(E)$ doesn't really make sense so that we consider intersect $[0,1]$ in our definition.

Yu-Tsung says: Actually, using $X+Y=\{x+y \mid x\in X, y\in Y\}$ directly might be more confusing than I thought... Because our example will be follows:
For $\{0\} = \mathscr{impossible}$ and $(0,\infty) = \mathscr{possible}$, it will become

$$\begin{array}{rcl} \mu(\emptyset) &=& \{0\} \\ \mu(\{HH\}) &=& (0,\infty) \\ \mu(\{HT\}) &=& \{0\} \\ \mu(\{TH\}) &=& (0,\infty) \\ \mu(\{TT\}) &=& \{0\} \\ \mu(\{HH,HT\}) &=& (0,\infty) \\ \mu(\{HH,TH\}) &=& (0,\infty) \\ \mu(\{HH,TT\}) &=& (0,\infty) \end{array} \qquad \begin{array}{rcl} \mu(\{HT,TH\}) &=& (0,\infty) \\ \mu(\{HT,TT\}) &=& \{0\} \\ \mu(\{TH,TT\}) &=& (0,\infty) \\ \mu(\{HH,HT,TH\}) &=& (0,\infty) \\ \mu(\{HH,HT,TT\}) &=& (0,\infty) \\ \mu(\{HH,TH,TT\}) &=& (0,\infty) \\ \mu(\{HT,TH,TT\}) &=& (0,\infty) \\ \mu(\{HH,HT,TH,TT\}) &=& (0,\infty) \end{array}$$

For $\{0\} = \mathscr{impossible}$ and $(0,1] = \mathscr{possible}$, it will become the following and gives an extra-value $(0,2]$.

$$\begin{array}{rcl} \mu(\emptyset) &=& \{0\} \\ \mu(\{HH\}) &=& (0,1] \\ \mu(\{HT\}) &=& \{0\} \\ \mu(\{TH\}) &=& (0,1] \\ \mu(\{TT\}) &=& \{0\} \\ \mu(\{HH,HT\}) &=& (0,1] \\ \mu(\{HH,TH\}) &=& (0,2] \\ \mu(\{HH,TT\}) &=& (0,1] \end{array} \qquad \begin{array}{rcl} \mu(\{HT,TH\}) &=& (0,1] \\ \mu(\{HT,TT\}) &=& \{0\} \\ \mu(\{TH,TT\}) &=& (0,1] \\ \mu(\{HH,HT,TH\}) &=& (0,2] \\ \mu(\{HH,HT,TT\}) &=& (0,1] \\ \mu(\{HH,TH,TT\}) &=& (0,2] \\ \mu(\{HT,TH,TT\}) &=& (0,1] \\ \mu(\{HH,HT,TH,TT\}) &=& (0,2] \end{array}$$

For $\{0\} = \mathscr{impossible}$, $[0,\frac{1}{2}] = \mathscr{unlikely}$, and $[\frac{1}{2},1] = \mathscr{likely}$, it will become the following and gives an extra-value $[\frac{1}{2},\frac{3}{2}]$.

$$\begin{array}{rcl} \mu(\emptyset) &=& \{0\} \\ \mu(\{HH\}) &=& [0,\frac{1}{2}] \\ \mu(\{HT\}) &=& \{0\} \\ \mu(\{TH\}) &=& [\frac{1}{2},1] \\ \mu(\{TT\}) &=& \{0\} \\ \mu(\{HH,HT\}) &=& [0,\frac{1}{2}] \\ \mu(\{HH,TH\}) &=& [\frac{1}{2},\frac{3}{2}] \\ \mu(\{HH,TT\}) &=& [0,\frac{1}{2}] \end{array} \qquad \begin{array}{rcl} \mu(\{HT,TH\}) &=& [\frac{1}{2},1] \\ \mu(\{HT,TT\}) &=& \{0\} \\ \mu(\{TH,TT\}) &=& [\frac{1}{2},1] \\ \mu(\{HH,HT,TH\}) &=& [\frac{1}{2},\frac{3}{2}] \\ \mu(\{HH,HT,TT\}) &=& [0,\frac{1}{2}] \\ \mu(\{HH,TH,TT\}) &=& [\frac{1}{2},\frac{3}{2}] \\ \mu(\{HT,TH,TT\}) &=& [\frac{1}{2},1] \\ \mu(\{HH,HT,TH,TT\}) &=& [\frac{1}{2},\frac{3}{2}] \end{array}$$

Actually, this approach is more natural to me when I think in measure theory, but I guess the readers might be confused if write in this way?

To summarize, an $\mathscr{L}$-valued probability is a function $\mu : \mathcal{E} \to \mathscr{L}$ such that:

- $1 \in \mu(\Omega)$, and
- for a collection E_i of pairwise disjoint events, $\mu(\bigcup_i E_i) = \bigvee_i \mu(E_i)$.

Example 2. [Two-coin probability space with finite set-valued probability measure] Under the new set-valued requirement, the probability measure in the first example becomes:

$$
\begin{array}{rcl}
\mu(\emptyset) & = & \textit{impossible} \\
\mu(\{HH\}) & = & \textit{possible} \\
\mu(\{HT\}) & = & \textit{impossible} \\
\mu(\{TH\}) & = & \textit{possible} \\
\mu(\{TT\}) & = & \textit{impossible} \\
\mu(\{HH, HT\}) & = & \textit{possible} \\
\mu(\{HH, TH\}) & = & \textit{possible} \\
\mu(\{HH, TT\}) & = & \textit{possible}
\end{array}
\qquad
\begin{array}{rcl}
\mu(\{HT, TH\}) & = & \textit{possible} \\
\mu(\{HT, TT\}) & = & \textit{impossible} \\
\mu(\{TH, TT\}) & = & \textit{possible} \\
\mu(\{HH, HT, TH\}) & = & \textit{possible} \\
\mu(\{HH, HT, TT\}) & = & \textit{possible} \\
\mu(\{HH, TH, TT\}) & = & \textit{possible} \\
\mu(\{HT, TH, TT\}) & = & \textit{possible} \\
\mu(\{HH, HT, TH, TT\}) & = & \textit{possible}
\end{array}
$$

Despite the fact that we have lost all numeric information, the probability measure still reveals that the second coin is double-headed. We have however lost the information regarding the bias in the first coin. This information can be recovered with a more refined probability measure as we show next. □

Although $\mathscr{L}_2$-valued probability measure is quite intuitive, set-valued probability measure is not that intuitive if we extend to more values. For example, consider the following three *overlapping* closed subintervals: $[0, 0]$ is still call *impossible*, $[0, \frac{1}{2}]$ which we call *unlikely*, ad $[\frac{1}{2}, 1]$ which we call *likely*. The following example is a set-valued probability measure corresponding to the probability measure in the first example.

Example 3. [Two-coin probability space with set-valued probability measure, again]

$$
\begin{array}{rcl}
\mu(\emptyset) & = & \textit{impossible} \\
\mu(\{HH\}) & = & \textit{unlikely} \\
\mu(\{HT\}) & = & \textit{impossible} \\
\mu(\{TH\}) & = & \textit{likely} \\
\mu(\{TT\}) & = & \textit{impossible} \\
\mu(\{HH, HT\}) & = & \textit{unlikely} \\
\mu(\{HH, TH\}) & = & \textit{likely} \\
\mu(\{HH, TT\}) & = & \textit{unlikely}
\end{array}
\qquad
\begin{array}{rcl}
\mu(\{HT, TH\}) & = & \textit{likely} \\
\mu(\{HT, TT\}) & = & \textit{impossible} \\
\mu(\{TH, TT\}) & = & \textit{likely} \\
\mu(\{HH, HT, TH\}) & = & \textit{likely} \\
\mu(\{HH, HT, TT\}) & = & \textit{unlikely} \\
\mu(\{HH, TH, TT\}) & = & \textit{likely} \\
\mu(\{HT, TH, TT\}) & = & \textit{likely} \\
\mu(\{HH, HT, TH, TT\}) & = & \textit{likely}
\end{array}
$$

In this example, we can get the information that the first coin is weighted and the second coin is double-headed. However, we may notice that the set-valued probability of the whole space $\{HH, HT, TH, TT\}$ are different between these two examples. This problem can be fixed when we moved on to the next section. □

We will return to finite set-valued probability measures in Sec. **??**.

2.4 Interval-valued probability measures

The singleton closed interval $[0, 0]$ is a required spacial value for empty set, and a natural generalization is to consider another special closed interval $[1, 1]$ which we call *necessary*for the whole space. In general, if the probability of an event E is $[a, b]$, we think of the left-endpoint a as representing the strength of the evidence that supports E, and the right-endpoint b as the strength of the evidence that contradicts E.

Thus if we have an event E with probability $[a, b]$ where $a = 0.1$ and $b = 0.7$, we have that:

- the strength of evidence supporting E is 0.1; since either E or its complement must happen, we conclude that there is 0.9 evidence supporting the complement of E;
- the strength of evidence contradicting E is 0.7; again since either E or its complement must happen, we conclude that there is 0.3 evidence contradicting the complement of E.

Yu-Tsung says: Do we use the law of excluded middle here? You remind me Agda :) Did Homotopy Type Theory people said anything about the probability?

Turning things around, the strength of evidence that contradicts E is evidence supporting the complement of E. The complement of E must therefore have probability $[1 - b, 1 - a]$ which we abbreviate $1 - [a, b]$, so the probability measure μ should satisfy:

- $\mu(\emptyset) = [0,0]$,
- $\mu(\Omega) = [1,1]$, and
- $\mu\left(\Omega \backslash E\right) = 1 - \mu\left(E\right)$

However, if we want $\mu(\Omega) = [1,1]$, and the probability assignment for singleton sets $\mu(\{HH\}) = \mathscr{unlikely}$, $\mu(\{HT\}) = \mathscr{impossible}$, $\mu(\{TH\}) = \mathscr{likely}$, and $\mu(\{TT\}) = \mathscr{impossible}$ as in example 3, we only have $\mu(\Omega) \subsetneq \mu(\{HH\}) + \mu(\{HT\}) + \mu(\{TH\}) + \mu(\{TT\})$. In general, we only require

- for a collection E_i of pairwise disjoint events, we have $\mu\left(\bigcup_i E_i\right) \subseteq \sum_i \mu\left(E_i\right)$.

This statement means that the evidences of $\bigcup_i E_i$ is at least as strong as putting all the evidences of E_i together, but some evidence may only be acquired for $\bigcup_i E_i$ as the whole. Therefore, $\mu\left(\bigcup_i E_i\right)$ is a subset of $\sum_i \mu\left(E_i\right)$, but may not equal. In our example,

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

The above equation told us $\mu(\Omega) \subseteq \left[\frac{1}{2}, \frac{3}{2}\right]$. However, because of $\mu(\emptyset) = \mathscr{impossible}$, we have $\mu\left(\Omega\right) = 1 - \mu\left(\emptyset\right) = \mathscr{necessary}$, but $\mu(\emptyset) = \mathscr{impossible}$ cannot be used to reasoning the probability of $\{HH\}$, $\{HT\}$, $\{TH\}$ and $\{TT\}$ individually. The full probability assignment is shown in the following example.

Example 4. [Two-coin probability space with four intervals] 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 $\mathscr{necessary}$. 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:

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

This probability measure is more informative than the one in Ex. 2: not only does it reveal that the second coin is double-headed but it also reveals the bias in the first coin. $\square$

3 Quantum Probability Spaces

The mathematical framework above assumes that one has complete knowledge of the events and their relationships. 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 case, 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 5 (One-qubit quantum probability space). Consider a one-qubit Hilbert space with states $\alpha|0\rangle+\beta|1\rangle$ such that $|\alpha|^2 + |\beta|^2 = 1$. 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. In contrast, a collection of non-orthogonal events (e.g., $|0\rangle\langle 0|$ and $|+\rangle\langle +|$) is not itself an 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 and then define a quantum probability space.

Definition 1 (Projection Operators; Orthogonality; Commutativity [10, 11, 12, 13]). Given a Hilbert space $\mathcal{H}$, a projection operator P is a linear transformation from $\mathcal{H}$ to itself such that $P^2 = P = P^\dagger$. Projection operators have the following properties:[2]

- Projection operators P_1 and P_2 are orthogonal if $P_1P_2 = P_2P_1 = \mathbb{0}$;
- Projection operators P_1 and P_2 commute if $P_1P_2 = P_2P_1$;
- If the projections P_1 and P_2 are orthogonal then $P_1 + P_2$ is also a projection;
- If the projections P_1 and P_2 commute then P_1P_2 is also a projection.

Amr says: Here it would be good to refer to the notion of "quantum test" and define events as sums of quantum tests. This will automatically include everything except the products of commutative projections which we will have to explain that they can be expressed as sums of orthogonal projections.

Definition 2 (Quantum Probability Space [15, 16, 11, 17, 14]). Given a Hilbert space $\mathcal{H}$, a *quantum probability space* consists of a set of events $\mathcal{E}$ and a probability measure $\mu : \mathcal{E} \to [0, 1]$ such that:[3]

- The set of events consists of all projections. This set includes the empty projection, projection operators $|\psi\rangle\langle\psi|$ for each state $|\psi\rangle$, sums of *orthogonal* projections, and products of *commuting* projections;
- $\mu(\mathbb{1}) = 1$, and
- for mutually orthogonal projections E_i, we have $\mu\left(\sum_i E_i\right) = \sum_i \mu(E_i)$.

□

[2] "Projection" is sometimes called "orthogonal projection" or "self-adjoint projection" to emphasize $P^\dagger = P$ [14].

[3] It is possible to define a more general space of events consisting of all operators $\mathcal{A}$ on $\mathcal{H}$ and consider $\mu : \mathcal{A} \to \mathbb{C}$ [14, 13]. 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.

3.1 Quantum Probability Measures

For a given set of events $\mathcal{E}$, there are many possible probability measures $\mu : \mathcal{E} \to [0,1]$. The Born rule, a postulate of quantum mechanics, states that each quantum state $|\phi\rangle$ induces a probability measure μ_ϕ as follows:

$$\mu_\phi(E) = \langle\phi|E\phi\rangle$$

Conversely, Gleason's theorem states that given a probability measure μ, there exist a quantum state $|\phi\rangle$ that induces such a measure using the Born rule. The theorem is only valid in Hilbert spaces with dimension $d \geq 3$. It is instructive to study counterexamples in $d = 2$, i.e., the case of a one-qubit system. Consider five states $|\psi_0\rangle$ to $|\psi_4\rangle$ that form five orthogonal bases $\{|\psi_0\rangle, |\psi_1\rangle\}$, $\{|\psi_1\rangle, |\psi_2\rangle\}$, $\{|\psi_2\rangle, |\psi_3\rangle\}$, $\{|\psi_3\rangle, |\psi_4\rangle\}$, and $\{|\psi_4\rangle, |\psi_0\rangle\}$ and consider the probability measure defined as follows. For all $i \in \{0, 1, 2, 3, 4\}$, we have $\mu_X(|\psi_i\rangle\langle\psi_i|) = 1/2$. For each orthogonal basis, the probability is 1 as desired and yet it is impossible to find a single quantum state that realizes such a probability measure (see `http://tph.tuwien.ac.at/~svozil/publ/2006-gleason.pdf`)

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 $\langle 0|E|0\rangle$. In general, Born discovered each quantum state $|\psi\rangle \in \mathcal{H}\setminus\{0\}$ induces a probability measure $\tilde{\mu}_\psi : \mathcal{E} \to [0,1]$ on the space of events defined for any event $E \in \mathcal{E}$ as follows [18, 19]:

$$\tilde{\mu}_\psi(E) = \frac{\langle\psi|E|\psi\rangle}{\langle\psi|\psi\rangle} \tag{1}$$

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

$$\tilde{\mu}_\rho(E) = \mathrm{Tr}(\rho E) = \sum_{j=1}^{N} q_j \tilde{\mu}_{\Psi_j}(E) \ . \tag{2}$$

  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.

- $\tilde{\mu}_\rho$ is a probability measure for all mixed state ρ.
- $\langle\psi|\phi\rangle = 0 \Leftrightarrow \tilde{\mu}_\psi(|\phi\rangle\langle\phi|) = 0$.
- $\tilde{\mu}_\psi(E) = \tilde{\mu}_{\mathbf{U}|\psi\rangle}\left(\mathbf{U}E\mathbf{U}^\dagger\right)$,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 [16, 12, 11]. Furthermore, a simple corollary of Gleason's theorem can show the Born rule is the unique function satisfying conditions 1. to 3.

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

Proof. Assume there is another function $\tilde{\mu}'$ such that $\tilde{\mu}'_\rho$ is a quantum probability measure for all mixed state ρ. We are going to prove $\tilde{\mu}' = \tilde{\mu}$.

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

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

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

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

3.2 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 [20, 21, 22].
- 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, 5].

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 [18, 19] and the map is given by: $E \mapsto \langle\psi|E\psi\rangle$.
- given a probability measure μ mapping each event E to a probability, is there a *unique* state ψ such that $\mu(E) = \langle\psi|E\psi\rangle$? In the conventional case, the answer is yes by Gleason's theorem [16, 12, 11].

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

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

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] William G. Faris. Appendix: Probability in quantum mechanics. In *The infamous boundary : seven decades of controversy in quantum physics.* Boston : Birkhauser, 1995.

[2] R.L. Graham, D.E. Knuth, and O. Patashnik. *Concrete Mathematics: A Foundation for Computer Science.* Addison-Wesley, 1994.

[3] V.K. Rohatgi and A.K.M.E. Saleh. *An Introduction to Probability and Statistics.* Wiley Series in Probability and Statistics. Wiley, 2011.

[4] Zvi Artstein. Set-valued measures. *Transactions of the American Mathematical Society*, 165:103–125, 1972.

[5] 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.

[6] Arthur P Dempster. Upper and lower probabilities induced by a multivalued mapping. *The annals of mathematical statistics*, pages 325–339, 1967.

[7] Lotfi A Zadeh. A simple view of the dempster-shafer theory of evidence and its implication for the rule of combination. *AI magazine*, 7(2):85, 1986.

[8] Kurt Weichselberger. The theory of interval-probability as a unifying concept for uncertainty. *International Journal of Approximate Reasoning*, 24(2):149–170, 2000.

[9] 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.

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

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

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

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

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

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

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

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

[18] Max Born. Zur quantenmechanik der stoßvorgänge (1926). In *Die Deutungen der Quantentheorie*, pages 48–52. Springer, 1984.

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

[20] 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.

[21] 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.

[22] 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.