Interval Probability for Fuzzy Quantum Theories

May 21, 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

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 [1, 2, 3, 4] 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:

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

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. $\square$

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 [5, 2, 6, 7] in detail.[1]

2.2 Measuring Probabilities: Buffon's Needle Problem

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 $\frac{2\ell}{\pi h}$ [10, 11, 12, 13]. 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 $\frac{M}{N}$.

[1] There is another possible approach that can be used to split the real interval $[0, 1]$ into a collection of subsets [8, 9]

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

In an actual experiment with 500 needles and the ratio $\frac{\ell}{h} = 0.75$ [12], 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 $\frac{\ell}{h} = 0.8$ [13], 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 $\frac{\ell}{h} = 0.6$ [11], 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 $\frac{1224}{3204}$.

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 $\mu : \mathcal{E} \to \mathscr{I}$ such that:

- $\mu(\emptyset) = [0, 0]$,
- $\mu(\Omega) = [1, 1]$,
- $\mu(\Omega \backslash E) = 1 \,{}_{\mathscr{I}}\!- \mu(E)$, and
- for a collection E_i of pairwise disjoint events, we have $\mu\left(\bigcup_i E_i\right) \subseteq {}_{\mathscr{I}}\sum_i \mu(E_i)$, 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 *impossible*, $[0, \frac{1}{2}]$ which we call *unlikely*, $[\frac{1}{2}, 1]$ which we call *likely*, and $[1, 1]$ which we call *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:

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

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

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

If we were to equate $\mu(\Omega)$ with the sum of the individual probabilities we would get that $\mu(\Omega) = \left[\frac{1}{2}, \frac{3}{2}\right]$. However, using the fact that $\mu(\emptyset) = \mathit{impossible}$, we have $\mu(\Omega) = 1 - \mu(\emptyset) = \mathit{certain} = [1,1]$. 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. □

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 and then define a quantum probability space.

Definition 1 (Projection Operators; Orthogonality [14, 15, 16, 3, 4]). Given a Hilbert space $\mathcal{H}$, an event[2] 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: [3]

- $\mathbb{0}$ 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 $P_1P_2 = P_2P_1 = \mathbb{0}$. 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 $\{|\psi_j\rangle\}$.

□

Based on the set of events, we can define the quantum probability space.

Definition 2 (Quantum Probability Space [14, 20, 15, 19]). 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:[4]

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

□

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, $\mu\left(P_1 \,{}_{\mathscr{O}}\!+ P_2\right) = \mu(P_1) \,{}_{\mathbb{R}}\!+ \mu(P_2)$ is always true whenever the left-handed side is defined. Whether an equation is defined also effects the further properties of quantum probability space.

Definition 3 (Complement; Commutativity [16, 3]).

- If P is a projection operator, then $\mathbb{1} \,{}_{\mathscr{O}}\!- P$ is also a projection operator, called *complement*. It is orthogonal to P, has probability $\mu\left(\mathbb{1} \,{}_{\mathscr{O}}\!- P\right) = 1 \,{}_{\mathbb{R}}\!- \mu\left(P\right)$, and corresponds to the complement event $\Omega \backslash E$ in classical probability.
- Projection operators P_1 and P_2 *commute* if $P_1P_2 = P_2P_1$. The product of two projection operators P_1P_2 is also a projection operator if and only if they commute. This corresponds to the classical intersection between events.

□

[2]An event is formally called an experimental proposition [17], a question [14, 18], or an elementary quantum test [16].

[3]"Projection" is sometimes called "orthogonal projection" or "self-adjoint projection" to emphasize $P^\dagger = P$ [3, 19].

[4]It is possible to define a more general space of events consisting of all operators $\mathcal{A}$ on $\mathcal{H}$ and consider $\mu : \mathcal{A} \to \mathbb{C}$ [19, 4]. When an operator $A \in \mathcal{A}$ is Hermitian, $\mu\left(A\right)$ is the expectation value of A. We does not take this approach because we want to focus only on probability.

Yu-Tsung says: The definition of independence is different from different sources:

1. In Swart's [4], given a $*$-algebra $\mathscr{A}$, two commuting sub-$*$-algebra $\mathscr{A}_1$ and $\mathscr{A}_2$ are *logically independent* if for all probability measure μ_1 on $\mathscr{A}_1$ and μ_2 on $\mathscr{A}_2$, there is an unique probability measure μ on the smallest sub-$*$-algebra of $\mathscr{A}$ containing both $\mathscr{A}_1$ and $\mathscr{A}_2$ such that $\mu(A_1A_2) = \mu_1(A_1)\mu_2(A_2)$ for all $A_1 \in \mathscr{A}_1$ and $A_2 \in \mathscr{A}_2$. If two sub-$*$-algebra $\mathscr{A}_1$ and $\mathscr{A}_2$ are logically independent, the smallest sub-$*$-algebra containing both $\mathscr{A}_1$ and $\mathscr{A}_2$ is denoted by $\mathscr{A}_1 \otimes \mathscr{A}_2$, and its probability measure μ is denoted by $\mu_1 \otimes \mu_2$
   Then, Swart proved the Bell inequality on logically independence $*$-algebras with a product state which gives a clear physical meaning. However, this only defines the independence of algebras instead of events.

2. In Yuan's [21], given a μ on a $*$-algebra $\mathscr{A}$, two sub-$*$-algebra $\mathscr{A}_1$ and $\mathscr{A}_2$ are *independent* if $\mu(A_1A_2) = \mu(A_1)\mu(A_2)$ for all $A_1 \in \mathscr{A}_1$ and $A_2 \in \mathscr{A}_2$. The same problem as the previous one: this defines the independent of algebras, but not events. Morever, I haven't found a clear physical meaning of this definition, yet...

3. In order to understand the definition of independence in [22], we need to first understand their notation because they focused on the hidden-variable (HV) models. Let λ is the HV, A_i is the measurement of the observable A at the position i in the sequence. For example, $A_1B_2C_3$ denotes the sequence of measuring A first, then B, and finally C. Given a fixed λ, we assume the outcome of an observable is deterministic. For example, the outcome of B_2 from the preceding sequence is denoted by $v(B_2|A_1B_2C_3)$. Then, the outcome of A is *independent* of whether B is measured before or after A is $v(A_1) = v(A_2|B_1A_2)$.
   If the outcomes of compatible observables are independent for any HV λ, the HV model is non-contextual. (Notice that they have their own definition of "compatible observables" for their HV model...)
   Compare to the previous definitions, this definition is appealing in a sense that projection operators are special cases of observables, and this definition has a physical meaning related to contextuality. However, before adopting their definition, we need to reformulate the definition to the usual quantum model....

4. For the definition we used last time, given a quantum probability measure $\mu : \mathcal{E} \to [0,1]$, two commuting projections P_1 and P_2 are *independent* if $\mu(P_1P_2) = \mu(P_1)\mu(P_2)$. We have two examples:

Yu-Tsung says:

Example 4 (Independence of product state). Consider the 2-qubit Hilbert space with a probability measure $\mu(P) = \langle{++}|P|{++}\rangle$ with the following projections:

$$P_1 = |00\rangle\langle 00| + |01\rangle\langle 01|, \quad P_2 = |00\rangle\langle 00| + |10\rangle\langle 10|, \quad P_3 = |10\rangle\langle 10| + |11\rangle\langle 11|, \quad P_4 = |00\rangle\langle 00|$$

and their probabilities:

$$\mu(\mathbb{0}) = 0, \quad \mu(P_1) = 1/2, \quad \mu(P_2) = 1/2, \quad \mu(P_3) = 1/2, \quad \mu(P_4) = 1/4 \ .$$

On one hand, P_1 and P_3 are complement to each other so we have

$$\begin{aligned} \mu(P_1) + \mu(P_3) &= \frac{1}{2} + \frac{1}{2} = 1 \\ \mu(P_1)\,\mu(P_3) &= \frac{1}{4} \neq 0 = \mu(\mathbb{0}) = \mu(P_1 P_3) \end{aligned}$$

Therefore, P_1 and P_3 commute, but are not independent. On the other hand, P_1 also commutes with P_2, and

$$\mu(P_1)\,\mu(P_2) = \frac{1}{4} = \mu(P_4) = \mu(P_1 P_2)$$

so that P_1 and P_2 are independent. □

Example 5 (Independence of any states). In general, for any orthonormal basis $\{|\psi_j\rangle\}_{j=0}^{3}$, we can always pick

$$P_1 = |\psi_0\rangle\langle\psi_0| + |\psi_1\rangle\langle\psi_1|, \quad P_2 = |\psi_0\rangle\langle\psi_0| + |\psi_2\rangle\langle\psi_2|, \quad P_3 = |\psi_0\rangle\langle\psi_0| \ .$$

With the probability measure $\mu(P) = \langle\psi|P|\psi\rangle$, where $|\psi\rangle = \frac{1}{2}\sum_{j=0}^{3}|\psi_j\rangle$, their probabilities are

$$\mu(P_1) = 1/2, \quad \mu(P_2) = 1/2, \quad \mu(P_3) = 1/4 \ ,$$

and we have

$$\mu(P_1)\,\mu(P_2) = \frac{1}{4} = \mu(P_3) = \mu(P_1 P_2) \ .$$

In particular, the following four Bell states form an orthonormal basis [23]:

$$|\phi^\pm\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle \pm |11\rangle\right), \quad |\psi^\pm\rangle = \frac{1}{\sqrt{2}}\left(|01\rangle \pm |10\rangle\right) \ .$$

Their sum is also a Bell state

$$|\psi\rangle = \frac{1}{2}\left(|\phi^+\rangle + |\phi^-\rangle + |\psi^+\rangle + |\psi^-\rangle\right) = \frac{1}{\sqrt{2}}\left(|00\rangle + |01\rangle\right) \ .$$

Thus, $|\phi^+\rangle\langle\phi^+| + |\psi^+\rangle\langle\psi^+|$ and $|\phi^+\rangle\langle\phi^+| + |\psi^-\rangle\langle\psi^-|$ are independent projections with respect to the probability measure $\mu(P) = \langle\psi|P|\psi\rangle$. □

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, and none of the source I have checked support our definition, yet. We may need to check more papers to find a source to support our definition or to find the majority definition (if there is one)...

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 [24, 25, 26], a postulate of quantum mechanics, states that each pure quantum state $|\phi\rangle$ induces a probability measure μ_ϕ as follows:

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

Moreover, the Born rule can be extended to a mixed state. Suppose we want to prepare a quantum system. As we discussed in classical probability, our ability of preparing a state might not be prefect. If we want to prepare $|\phi\rangle$, we may turn out preparing a set of state $|\phi_j\rangle$ each with probability q_j, then 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 μ_{ϕ_j} with respect to probability q_j [16, 27, 26]:

$$\mu_\rho\left(E\right) \quad = \quad \mathrm{Tr}\left(\rho E\right) = \sum_{j=1}^{N} q_j \mu_{\phi_j}\left(E\right) \ . \tag{1}$$

Conversely, Gleason's theorem states that given a probability measure μ, there exist a mixed state ρ that induces such a measure using the Born rule [20, 15, 16]. 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.

Example 6 (One-qubit quantum probability measure). Consider a quantum probability measure $\mu : \mathcal{E} \to [0,1]$ defined as follow:

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

On one hand, μ is a probability measure. Because μ is almost the same as a probability measure $\mu_{|0\rangle}$, 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

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

However, $\mu_\rho(E) = 1$ if and only if ρ represents a pure state and $\rho = E$. □

Amr says: The idea will be the following. First describe quantum probability spaces conventionally. Then talk about the following:

- the dimension of the Hilbert space is a parameter that is like the number of needles; it gives an upper bound on the accuracy of the numbers that are relevant in expressing probabilities
- the intervals will come from two things: the fact that states can only be prepared to a certain accuracy so when we say the state is $|\psi\rangle$ we really mean a neighborhood of states close to $|\psi\rangle$
- similarly when we do an experiment with $|\phi\rangle\langle\phi|$ we are really testing a family of projections that are near $|\phi\rangle\langle\phi|$; this fuziness will cause the probability to only be specifiable as intervals

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.

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 [24, 25]:

$$\tilde{\mu}_\psi(E) = \frac{\langle\psi|E|\psi\rangle}{\langle\psi|\psi\rangle} \tag{2}$$

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

$$\tilde{\mu}_\rho(E) \quad = \quad \mathrm{Tr}(\rho E) = \sum_{j=1}^{N} q_j \tilde{\mu}_{\Psi_j}(E) \ . \tag{3}$$

  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 [20, 16, 15]. 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

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

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 [28, 29, 30].

- Instead of real-valued probability measures producing results in the uncountable and uncomputable interval $[0,1]$, we will consider finite set-valued probability measures [8, 9].

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 [24, 25] 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 [20, 16, 15].

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

[2] Glenn Shafer. *A mathematical theory of evidence*. Princeton university press, 1976.

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

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

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

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

[7] 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.

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

[9] 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.

[10] G Buffon. Essai d'arithmétique morale. In *Histoire naturelle, générale et particulière, Supplément*, volume 4, pages 46–109. 1777.

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

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

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

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

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

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

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

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

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

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

[21] Qiaochu Yuan. Noncommutative probability. Blog, August 2012.

[22] Otfried Gühne, Matthias Kleinmann, Adán Cabello, Jan-Åke Larsson, Gerhard Kirchmair, Florian Zähringer, Rene Gerritsma, and Christian F Roos. Compatibility and noncontextuality for sequential measurements. *Physical Review A*, 81(2):022121, 2010.

[23] Adán Cabello. Bell's theorem without inequalities and without probabilities for two observers. *Phys. Rev. Lett.*, 86:1911–1914, Mar 2001.

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

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

[26] 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.

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

[28] 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.

[29] 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.

[30] 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.