Non-commutative probability theory and its applications

Zheyuan Wu

Washington University in St. Louis

June 23, 2025

Table of Contents

Introduction

Non-commutative probability theory is a branch of generalized probability theory that studies the probability of events in non-commutative algebras (e.g. the algebra of observables in quantum mechanics).

Light polarization and the violation of Bell's inequality

The light which comes through a polarizer is polarized in a certain direction. If we fixed the first filter and rotate the second filter, we will observe the intensity of the light will change.
The light intensity decreased with α (the angle between the two filters).
The light should vanished when $\alpha = \pi/2$.

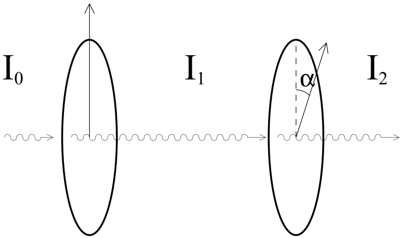

FIG. 1

Light polarization and the violation of Bell's inequality

By experimental measurement, the intensity of the light passing the first filter is half the beam intensity (Assume the original beam is completely unpolarized).
Then $I_1 = I_0/2$, and

$$I_2 = I_1 \cos^2 \alpha$$

Claim: there exist a smallest package of monochromatic light, which is a photon.
We can model the behavior of each individual photon passing through the filter with direction α with random variable P_α. The $P_\alpha(\omega) = 1$ if the photon passes through the filter, and $P_\alpha(\omega) = 0$ if the photon does not pass through the filter.
Then, the probability of the photon passing through the two filters with direction α and β is given by

$$\mathbb{E}(P_\alpha P_\beta) = \text{Prob}(P_\alpha = 1 \text{ and } P_\beta = 1) = \frac{1}{2} \cos^2(\alpha - \beta)$$

Light polarization and the violation of Bell's inequality

However, for system of 3 polarizing filters F_1, F_2, F_3, having direction $\alpha_1, \alpha_2, \alpha_3$. If we put them on the optical bench in pairs, Then we will have three random variables P_1, P_2, P_3.

Bell's 3 variable inequality

For any three random variables P_1, P_2, P_3 in a classical probability space, we have

$$\mathrm{Prob}(P_1 = 1, P_3 = 0) \leq \mathrm{Prob}(P_1 = 1, P_2 = 0) + \mathrm{Prob}(P_2 = 1, P_3 = 0) \tag{1}$$

Proof of Bell's 3 variable inequality

Proof:
By the law of total probability, (The event that the photon passes through the first filter but not the third filter is the union of the event that the photon did not pass through the second filter and the event that the photon passed the second filter and did not pass through the third filter) we have

$$\begin{aligned}\mathrm{Prob}(P_1 = 1, P_3 = 0) &= \mathrm{Prob}(P_1 = 1, P_2 = 0, P_3 = 0) \\ &+ \mathrm{Prob}(P_1 = 1, P_2 = 1, P_3 = 0) \\ &\leq \mathrm{Prob}(P_1 = 1, P_2 = 0) + \mathrm{Prob}(P_2 = 1, P_3 = 0)\end{aligned}$$

Light polarization and the violation of Bell's inequality

However, according to our experimental measurement, for any pair of polarizers F_i, F_j, by the complement rule, we have

$$\begin{aligned}\mathrm{Prob}(P_i = 1, P_j = 0) &= \mathrm{Prob}(P_i = 1) - \mathrm{Prob}(P_i = 1, P_j = 1) \\ &= \frac{1}{2} - \frac{1}{2}\cos^2(\alpha_i - \alpha_j) \\ &= \frac{1}{2}\sin^2(\alpha_i - \alpha_j)\end{aligned}$$

This leads to a contradiction if we apply the inequality to the experimental data.

$$\frac{1}{2}\sin^2(\alpha_1 - \alpha_3) \leq \frac{1}{2}\sin^2(\alpha_1 - \alpha_2) + \frac{1}{2}\sin^2(\alpha_2 - \alpha_3)$$

Light polarization and the violation of Bell's inequality

If $\alpha_1 = 0, \alpha_2 = \frac{\pi}{6}, \alpha_3 = \frac{\pi}{3}$, then

$$\begin{aligned}
\frac{1}{2}\sin^2(-\frac{\pi}{3}) &\leq \frac{1}{2}\sin^2(-\frac{\pi}{6}) + \frac{1}{2}\sin^2(\frac{\pi}{6} - \frac{\pi}{3}) \\
\frac{3}{8} &\leq \frac{1}{8} + \frac{1}{8} \\
\frac{3}{8} &\leq \frac{1}{4}
\end{aligned}$$

Other revised experiments (eg. Aspect's experiment, Calcium entangled photon experiment) are also conducted and the inequality is still violated.

The true model of light polarization

The full description of the light polarization is given belows:
State of polarization of a photon: $\psi = \alpha|0\rangle + \beta|1\rangle$, where $|0\rangle$ and $|1\rangle$ are the two orthogonal polarization states in $\mathbb{C}^2$.
Polarization filter (generalized 0,1 valued random variable): orthogonal projection P_α on $\mathbb{C}^2$ corresponding to the direction α. (operator satisfies $P_\alpha^* = P_\alpha = P_\alpha^2$.)
The matrix representation of P_α is given by

$$P_\alpha = \begin{pmatrix} \cos^2(\alpha) & \cos(\alpha)\sin(\alpha) \\ \cos(\alpha)\sin(\alpha) & \sin^2(\alpha) \end{pmatrix}$$

Probability of a photon passing through the filter P_α is given by $\langle P_\alpha \psi, \psi\rangle$, this is $\cos^2(\alpha)$ if we set $\psi = |0\rangle$.
Since the probability of a photon passing through the three filters is not commutative, it is impossible to discuss $\mathrm{Prob}(P_1 = 1, P_3 = 0)$ in the classical setting.

Let $\mathscr{H}$ be a Hilbert space. $\mathscr{H}$ consists of complex-valued functions on a finite set $\Omega = \{1, 2, \cdots, n\}$. and that the functions $(e_1, e_2, \cdots, e_n)$ form an orthonormal basis of $\mathscr{H}$. (We use Dirac notation $|k\rangle$ to denote the basis vector e_k.)

Definition 1.1 (non-commutative probability space)

A non-commutative probability space is a pair $(\mathscr{B}(\mathscr{H}), \mathscr{P})$, where $\mathscr{B}(\mathscr{H})$ is the set of all bounded linear operators on $\mathscr{H}$ and $\mathscr{P}$ is the set of all orthogonal projections on $\mathscr{B}(\mathscr{H})$.

The set $\mathscr{P} = \{P \in \mathscr{B}(\mathscr{H}) : P^* = P = P^2\}$

Non-commutative probability theory

Definition 1.2 (state)

A state on $(\mathscr{B}(\mathscr{H}), \mathscr{P})$ is a map $\mu : \mathscr{P} \to [0,1]$ such that:

- $\mu(O) = 0$, where O is the zero projection.
- If $P_1, P_2, \cdots, P_n$ are pairwise disjoint orthogonal projections, then $\mu(P_1 \vee P_2 \vee \cdots \vee P_n) = \sum_{i=1}^{n} \mu(P_i)$.

Definition 1.3 (density operator)

A density operator ρ on the finite-dimensional Hilbert space $\mathscr{H}$ is:

- self-adjoint ($A^* = A$, that is $\langle Ax, y\rangle = \langle x, Ay\rangle$ for all $x, y \in \mathscr{H}$)
- positive semi-definite (all eigenvalues are non-negative)
- $\mathrm{Tr}(\rho) = 1$.

Example of density operator

If $(|\psi_1\rangle, |\psi_2\rangle, \cdots, |\psi_n\rangle)$ is an orthonormal basis of $\mathscr{H}$ consisting of eigenvectors of ρ, for the eigenvalue $p_1, p_2, \cdots, p_n$, then $p_j \geq 0$ and $\sum_{j=1}^n p_j = 1$.
We can write ρ as

$$\rho = \sum_{j=1}^n p_j |\psi_j\rangle\langle\psi_j|$$

(under basis $|\psi_j\rangle$, it is a diagonal matrix with p_j on the diagonal)

Non-commutative probability theory

Born's rule

Let ρ be a density operator on $\mathscr{H}$. then

$$\mu(P) \coloneqq \mathrm{Tr}(\rho P) = \sum_{j=1}^{n} p_j \langle \psi_j | P | \psi_j \rangle$$

Defines a probability measure on the space $\mathscr{P}$.

Non-commutative probability theory

Definition 1.4 (random variable, or observable)

Let $\mathscr{B}(\mathbb{R})$ be the set of all Borel sets on $\mathbb{R}$.
A random variable on the Hilbert space $\mathscr{H}$ is a projection valued map (measure) $P : \mathscr{B}(\mathbb{R}) \to \mathscr{P}$.
With the following properties:

- $P(\emptyset) = O$ (the zero projection)
- $P(\mathbb{R}) = I$ (the identity projection)
- For any sequence $A_1, A_2, \cdots, A_n \in \mathscr{B}(\mathbb{R})$. the following holds:
  - $P(\bigcup_{i=1}^n A_i) = \bigvee_{i=1}^n P(A_i)$
  - $P(\bigcap_{i=1}^n A_i) = \bigwedge_{i=1}^n P(A_i)$
  - $P(A^c) = I - P(A)$
  - If A_j are mutually disjoint (that is $P(A_i)P(A_j) = P(A_j)P(A_i) = O$ for $i \neq j$), then $P(\bigcup_{j=1}^n A_j) = \sum_{j=1}^n P(A_j)$

Definition 1.5 (probability of a random variable)

For a system prepared in state ρ, the probability of the random variable by the projection-valued measure P is in the Borel set A is $\mathrm{Tr}(\rho P(A))$.

When operators commute, we recover classical probability measures.

Summary of analog of classical probability theory and non-commutative (*quantum*) probability theory

Classical probability	**Non-commutative probability**
Sample space Ω, cardinality $\lvert\Omega\rvert = n$, example: $\Omega = \{0, 1\}$	Complex Hilbert space $\mathscr{H}$, dimension $\dim \mathscr{H} = n$, example: $\mathscr{H} = \mathbb{C}^2$
Common algebra of $\mathbb{C}$ valued functions	Algebra of bounded operators $\mathscr{B}(\mathscr{H})$
$f \mapsto \bar{f}$ complex conjugation	$P \mapsto P^*$ adjoint
Events: indicator functions of sets	Projections: space of orthogonal projections $\mathscr{P} \subseteq \mathscr{B}(\mathscr{H})$
functions f such that $f^2 = f = \overline{f}$	orthogonal projections P such that $P^* = P = P^2$
$\mathbb{R}$-valued functions $f = \overline{f}$	self-adjoint operators $A = A^*$
$\mathbb{I}_{f^{-1}(\{\lambda\})}$ is the indicator function of the set $f^{-1}(\{\lambda\})$	$P(\lambda)$ is the orthogonal projection to eigenspace
$f = \sum_{\lambda \in \mathrm{Range}(f)} \lambda \mathbb{I}_{f^{-1}(\{\lambda\})}$	$A = \sum_{\lambda \in \mathrm{sp}(A)} \lambda P(\lambda)$
Probability measure μ on Ω	Density operator ρ on $\mathscr{H}$
Delta measure δ_ω	Pure state $\rho = \lvert\psi\rangle\langle\psi\rvert$
μ is non-negative measure and $\sum_{i=1}^n \mu(\{i\}) = 1$	ρ is positive semi-definite and $\mathrm{Tr}(\rho) = 1$
Expected value of random variable f is $\mathbb{E}_\mu(f) = \sum_{i=1}^n f(i)\mu(\{i\})$	Expected value of operator A is $\mathbb{E}_\rho(A) = \mathrm{Tr}(\rho A)$
Variance of random variable f is $\mathrm{Var}_\mu(f) = \sum_{i=1}^n (f(i) - \mathbb{E}_\mu(f))^2 \mu(\{i\})$	Variance of operator A is $\mathrm{Var}_\rho(A) = \mathrm{Tr}(\rho A^2) - \mathrm{Tr}(\rho A)^2$
Covariance of random variables f and g is $\mathrm{Cov}_\mu(f, g) = \sum_{i=1}^n (f(i) - \mathbb{E}_\mu(f))(g(i) - \mathbb{E}_\mu(g))\mu(\{i\})$	Covariance of operators A and B is $\mathrm{Cov}_\rho(A, B) = \mathrm{Tr}(\rho A \circ B) - \mathrm{Tr}(\rho A)\,\mathrm{Tr}(\rho B)$

DeepSeek: Non-Commutative Probability under Von Neumann Algebra Framework

Definition 2.1 (Von Neumann algebra)

A Von Neumann algebra $\mathcal{M}$ (weakly closed *-subalgebra of bounded operators $\mathscr{B}(\mathscr{H})$ on a Hilbert space $\mathscr{H}$) replaces Ω. Projections in $\mathcal{M}$ represent "events".

- *-subalgebra:
- A *-subalgebra A of $B(H)$ is a linear subspace that is closed under multiplication (so if T, S are in A, then TS is in A) and closed under taking adjoints (so if T is in A, then T^* is in A).
- Weakly closed:
- The weak operator topology on $B(H)$ is the topology defined by the seminorms: $T \mapsto |\langle Tx, y\rangle|$ for all x, y in H.
- In other words, a net of operators T_α converges to T in the weak operator topology if for every x, y in H, the net of complex numbers $\langle T_\alpha x, y\rangle$ converges to $\langle Tx, y\rangle$.

DeepSeek: Non-Commutative Probability under Von Neumann Algebra Framework

Why is weak closure important? Because of von Neumann's double commutant theorem:

Theorem 2.2 (Von Neumann's double commutant theorem)

A *-subalgebra A of $B(H)$ that contains the identity operator is a von Neumann algebra if and only if it is equal to its double commutant, i.e., $A = (A')'$ where A' is the commutant of A.

The commutant A' is defined as the set of all bounded operators on H that commute with every operator in A.
So, the weak closure is equivalent to being equal to the double commutant (when the algebra contains the identity).

DeepSeek: Non-Commutative Probability under Von Neumann Algebra Framework

Definition 2.3 (Random variables)

Self-adjoint operators $X = X^*$ in a von Neumann algebra $\mathcal{M}$ represent random variables.

- Spectrum $\sigma(X)$ generalizes the range of f in classical probability theory.
- Spectral measure $E_X(B)$ replaces the preimage $f^{-1}(B)$ in classical probability theory.

DeepSeek: Non-Commutative Probability under Von Neumann Algebra Framework

Concept	**Classical Probability**	**Non-Commutative Probability under Von Neumann Algebra Framework**
Sample Space	(Ω, Σ)	Von Neumann algebra $\mathcal{M}$
Observable	Measurable function $f : \Omega \to \mathbb{R}$	Self-adjoint operator $X \in \mathcal{M}$
State	Probability measure μ	Normal state $\phi : \mathcal{M} \to \mathbb{C}$
Event	Measurable set $A \in \Sigma$	Projection $P \in \mathcal{M}$
Effect	Fuzzy event $[0, 1]$-function	Operator $A \in \mathcal{M}$ with $0 \leq A \leq I$ (POVMs)
Transformation	Stochastic map $K : \Omega \to \Omega'$	Completely positive map $\Phi : \mathcal{M} \to \mathcal{N}$ (sub-algebra of $\mathcal{M}$)
Independence	Product measure $\mu_1 \otimes \mu_2$	Tensor state $\phi_1 \otimes \phi_2$ on $\mathcal{M}_1 \overline{\otimes} \mathcal{M}_2$ (spatial tensor product)

Future Plans

- Quantum channels and information theory build on that
- Quantum circuits and proofs of quantum algorithms
- Quantum error correction codes
- Other ideas and applications on the non-commutative probability theory
  - Positive Operator Valued Measures (POVMs)
  - Kraus operators
  - Dynamical systems in non-commutative probability theory
- Other interesting topics in computer science
  - CUDA Quantum Simulation Backends
  - Quantum error correction below the surface code threshold

References I

Feres, R.
Math 444 lecture notes – the mathematics of quantum theory.

Kümmer, B. and Maassen, H.
Elements of quantum probability, pages 73–100.

Parthasarathy, K. R. (1992).
An Introduction to Quantum Stochastic Calculus, volume 85 of *Monographs in Mathematics*.
Birkhäuser Basel.

Q&A