Measure concentration in complex projective space and quantum entanglement

Zheyuan Wu

Washington University in St. Louis

March 23, 2026

Table of Contents

Light polarization and non-commutative probability

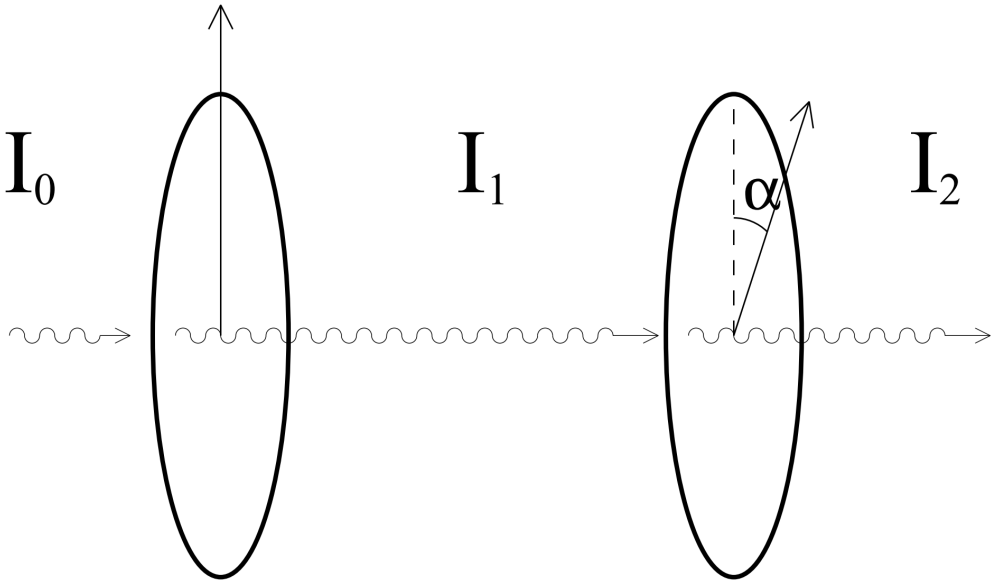


FIG. 1

- Light passing through a polarizer becomes polarized in the direction of that filter.
- If two filters are placed with relative angle α, the transmitted intensity decreases as α increases.
- In particular, the transmitted intensity vanishes when $\alpha = \pi/2$.

Polarization experiment

Now consider three filters F_1, F_2, F_3 with directions

$$\alpha_1, \alpha_2, \alpha_3.$$

Testing them pairwise suggests introducing three 0–1 random variables

$$P_1, P_2, P_3,$$

where $P_i = 1$ means that the photon passes filter F_i.

If these were classical random variables on one probability space, they would satisfy a Bell-type inequality.

A classical Bell-type inequality

Bell-type inequality

For any classical random variables $P_1, P_2, P_3 \in \{0, 1\}$,

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

Proof.

The event $\{P_1 = 1, P_3 = 0\}$ splits into two disjoint cases according to whether $P_2 = 0$ or $P_2 = 1$:

$$\{P_1 = 1, P_3 = 0\} = \{P_1 = 1, P_2 = 0, P_3 = 0\} \sqcup \{P_1 = 1, P_2 = 1, P_3 = 0\}.$$

Therefore,

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

Experimental law

For unpolarized incoming light, the **observed transition law** for a pair of filters is

$$\mathrm{Prob}(P_i = 1, P_j = 0) = \mathrm{Prob}(P_i = 1) - \mathrm{Prob}(P_i = 1, P_j = 1).$$

Using the polarization law,

$$\mathrm{Prob}(P_i = 1) = \frac{1}{2}, \qquad \mathrm{Prob}(P_i = 1, P_j = 1) = \frac{1}{2}\cos^2(\alpha_i - \alpha_j),$$

hence

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

So the experimentally observed probabilities depend only on the angle difference $\alpha_i - \alpha_j$.

Violation of the classical inequality

Substituting the experimental law into the classical inequality gives

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

Choose

$$\alpha_1 = 0, \qquad \alpha_2 = \frac{\pi}{6}, \qquad \alpha_3 = \frac{\pi}{3}.$$

Then

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

which is false.

Therefore the pairwise polarization data cannot come from one classical probability model with random variables P_1, P_2, P_3.

The quantum model of polarization

The correct model uses a Hilbert space rather than classical events.

- A pure polarization state is a vector

$$\psi = \alpha|0\rangle + \beta|1\rangle \in \mathbb{C}^2.$$

- A filter at angle α is represented by the orthogonal projection

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

- For a pure state ψ, the probability of passing the filter is

$$\langle P_\alpha\psi, \psi\rangle.$$

The key point is that sequential measurements are described by *ordered products* of projections, and these need not commute.

Recovering the observed law from the operator model

Assume the incoming light is unpolarized, so its state is the density matrix

$$\rho = \frac{1}{2} I.$$

The probability of passing the first filter P_{α_i} is

$$\mathrm{Prob}(P_i = 1) = \mathrm{tr}(\rho P_{\alpha_i}) = \frac{1}{2}\,\mathrm{tr}(P_{\alpha_i}) = \frac{1}{2}.$$

If the photon passes the first filter, the post-measurement state is

$$\rho_i = \frac{P_{\alpha_i} \rho P_{\alpha_i}}{\mathrm{tr}(\rho P_{\alpha_i})} = P_{\alpha_i}.$$

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

Therefore

$$\mathrm{Prob}(P_j = 1 \mid P_i = 1) = \mathrm{tr}(\rho_i P_{\alpha_j}) = \mathrm{tr}(P_{\alpha_i} P_{\alpha_j}) = \cos^2(\alpha_i - \alpha_j).$$

Recovering the observed law from the operator model (cont.)

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

This matches the experiment exactly.

Conclusion

- The classical model predicts a Bell-type inequality for three 0–1 random variables.
- The polarization experiment violates that inequality.
- The resolution is that the quantities measured are *sequential probabilities*, not joint probabilities of classical random variables.
- In quantum probability, events are modeled by projections on a Hilbert space, and measurement order matters.

This is one of the basic motivations for passing from classical probability to non-commutative probability.

Quantum states: pure vs. mixed

- A finite-dimensional quantum system is modeled by a complex Hilbert space (a complete inner product space)

$$\mathcal{H} \cong \mathbb{C}^{n+1}.$$

- A **pure state** is represented by a unit vector

$$\psi \in \mathcal{H}, \qquad \|\psi\| = 1.$$

- A **mixed state** is represented by a density matrix

$$\rho \geq 0, \qquad \mathrm{tr}(\rho) = 1.$$

- Pure states describe maximal information; mixed states describe probabilistic mixtures or partial information.

Key distinction

Pure states form a curved geometric space; mixed states form a convex set inside the space of matrices.

Why pure states are not vectors

- Two nonzero vectors that differ by a nonzero complex scalar represent the same physical state:

$$\psi \sim \lambda\psi, \qquad \lambda \in \mathbb{C}^{\times}.$$

- In particular, multiplying by a phase $e^{i\theta}$ does not change any physical predictions.
- Therefore the physical pure state is not a single vector, but the *complex line* spanned by that vector.

Hence the space of pure states is

$$\mathbb{P}(\mathcal{H}) = (\mathcal{H} \setminus \{0\})/\mathbb{C}^{\times}.$$

After choosing a basis $\mathcal{H} \cong \mathbb{C}^{n+1}$, this becomes

$$\mathbb{P}(\mathcal{H}) \cong \mathbb{C}P^{n}.$$

Relation with the sphere

- Every nonzero vector can be normalized, so each pure state has a representative on the unit sphere

$$S^{2n+1} \subset \mathbb{C}^{n+1}.$$

- Two unit vectors represent the same pure state exactly when they differ by a phase:

$$z \sim e^{i\theta} z.$$

- Therefore

$$\mathbb{C}P^n = S^{2n+1}/S^1.$$

The quotient map

$$p : S^{2n+1} \to \mathbb{C}P^n, \qquad p(z) = [z] = \{\lambda z : \lambda \in \mathbb{C}^\times\},$$

is the **Hopf fibration**.

How the metric descends to $\mathbb{C}P^n$

- The sphere S^{2n+1} inherits the round metric from the Euclidean metric on
$$\mathbb{C}^{n+1} \cong \mathbb{R}^{2n+2}.$$
- The fibers of the Hopf map are circles
$$p^{-1}([z]) = \{e^{i\theta} z : \theta \in \mathbb{R}\}.$$
- Tangent vectors split into:
  - **vertical directions**: tangent to the S^1-fiber,
  - **horizontal directions**: orthogonal complement to the fiber.
- The differential dp identifies horizontal vectors on the sphere with tangent vectors on $\mathbb{C}P^n$.

This allows the round metric on S^{2n+1} to define a metric on $\mathbb{C}P^n$.

The induced metric: Fubini–Study metric

- The metric on $\mathbb{C}P^n$ obtained from the Hopf quotient is the **Fubini–Study metric**.
- So the geometric picture is:

$$S^{2n+1} \xrightarrow{\text{Hopf fibration}} \mathbb{C}P^n$$

$$\text{round metric} \rightsquigarrow \text{Fubini–Study metric.}$$

- The normalized Riemannian volume measure induced by this metric gives the natural probability measure on pure states.

Proof roadmap

To prove this carefully, one usually shows:

1. $p : S^{2n+1} \to \mathbb{C}P^n$ is a smooth surjective submersion,
2. the vertical space is the tangent space to the S^1-orbit,
3. horizontal lifts are well defined,
4. the quotient metric is exactly the Fubini–Study metric.

Maxwell-Boltzmann Distribution Law

Consider the orthogonal projection

$$\pi_{n,k} : S^n(\sqrt{n}) \to \mathbb{R}^k.$$

Its push-forward measure converges to the standard Gaussian:

$$(\pi_{n,k})_* \sigma^n \to \gamma^k.$$

This explains why Gaussian behavior emerges from high-dimensional spheres and supports the proof strategy for Levy concentration.

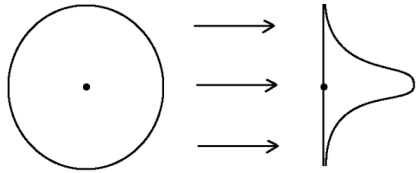


Figure 3.9 The projective central limit theorem: the projection of the uniform distribution on the sphere of radius $\sqrt{n}$ onto a line converges to the normal distribution $N(0,1)$ as $n \to \infty$.

Levy Concentration

Levy's theorem

If $f : S^n \to \mathbb{R}$ is 1-Lipschitz, then there exists a median a_0 such that

$$\mu\{x \in S^n : |f(x) - a_0| \geq \epsilon\} \leq 2 \exp\left(-\frac{(n-1)\epsilon^2}{2}\right).$$

- In high dimension, most Lipschitz observables are almost constant.
- This is the geometric mechanism behind generic entanglement.

Generic Entanglement Theorem

Hayden–Leung–Winter

Let $\psi \in \mathcal{P}(A \otimes B)$ be Haar-random and define

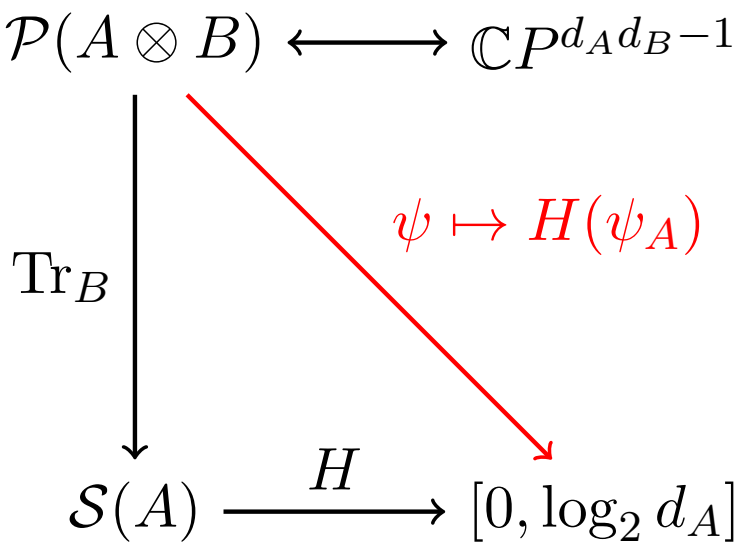
$$\beta = \frac{1}{\ln(2)} \frac{d_A}{d_B}.$$

For $d_B \geq d_A \geq 3$,

$$\Pr[H(\psi_A) < \log_2(d_A) - \alpha - \beta] \leq \exp\left(-\frac{1}{8\pi^2 \ln(2)} \frac{(d_A d_B - 1)\alpha^2}{(\log_2 d_A)^2}\right).$$

With overwhelming probability, a random pure state is almost maximally entangled.

How the Entropy Observable Fits In



- The red arrow is the observable to which concentration is applied.
- The projective description is natural because global phase does not change the physical state.

Ingredients Behind the Tail Bound

Page-type lower bound

$$\mathbb{E}[H(\psi_A)] \geq \log_2(d_A) - \frac{1}{2\ln(2)}\frac{d_A}{d_B}.$$

Lipschitz estimate

$$\|H(\psi_A)\|_{\mathrm{Lip}} \leq \sqrt{8}\,\log_2(d_A), \qquad d_A \geq 3.$$

Levy concentration plus these two estimates produces the exponential entropy tail bound.

Observable diameter: the inner definition

Partial diameter on $\mathbb{R}$

Let ν be a Borel probability measure on $\mathbb{R}$ and let $\alpha \in (0, 1]$. The **partial diameter** of ν at mass level α is

$$\operatorname{diam}(\nu; \alpha) := \{\operatorname{diam}(A) : A \subseteq \mathcal{B}(\mathbb{R}), \nu(A) \geq \alpha\},$$

where

$$\operatorname{diam}(A) := \sup_{x,y \in A} |x - y|.$$

- This asks for the shortest interval-like region containing at least α of the total mass.
- So $\operatorname{diam}(\nu; 1 - \kappa)$ measures how tightly we can capture *most* of the distribution, allowing us to discard a set of mass at most κ.

Observable diameter of a metric-measure space

Definition

Let $X = (X, d_X, \mu_X)$ be a metric-measure space and let $\kappa > 0$. The **observable diameter** of X is

$$\mathrm{ObsDiam}_{\mathbb{R}}(X; -\kappa) := \sup_{f \in \mathrm{Lip}_1(X,\mathbb{R})} \mathrm{diam}(f_*\mu_X; 1-\kappa),$$

where $\mathrm{Lip}_1(X, \mathbb{R})$ is the set of all 1-Lipschitz functions $f : X \to \mathbb{R}$, and $f_*\mu_X$ is the pushforward measure on $\mathbb{R}$.

- Each 1-Lipschitz function f is viewed as an **observable** on X.
- The pushforward measure $f_*\mu_X$ is the distribution of the values of that observable.
- If $\mathrm{ObsDiam}_{\mathbb{R}}(X; -\kappa)$ is small, then *every* 1-Lipschitz observable is strongly concentrated.

A Geometric Consequence

In this thesis, entropy functions are used as concrete observables to estimate observable diameter, and the Hopf fibration helps transfer information between S^{2n+1} and $\mathbb{C}P^n$.

Projective-space estimate

For $0 < \kappa < 1$,

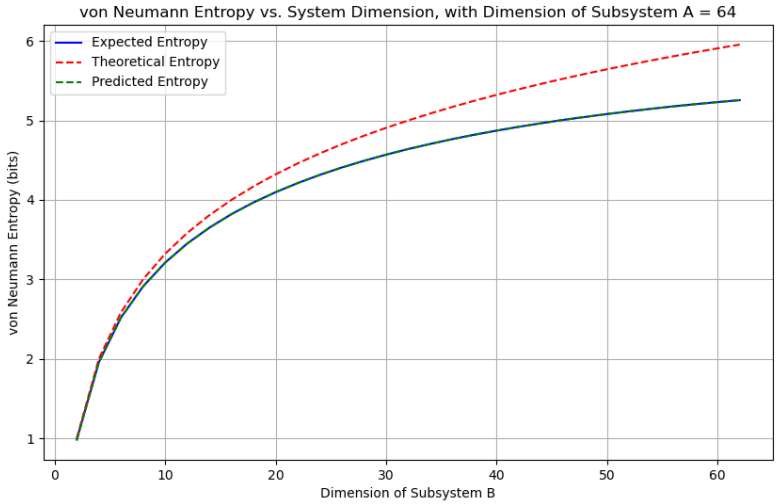
$$\mathrm{ObsDiam}(\mathbb{C}P^n(1); -\kappa) \leq O(\sqrt{n}).$$

- First estimate observable diameter on spheres via Gaussian limits.
- Then use the Hopf map $S^{2n+1}(1) \to \mathbb{C}P^n$.
- This gives a geometric explanation for why many projective-space observables concentrate.

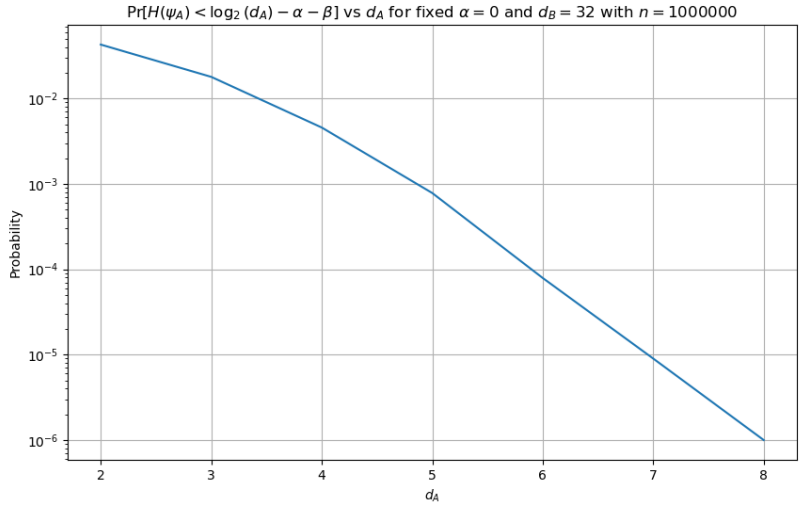
Entropy-Based Simulations

- Sample Haar-random pure states in $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$.
- Compute reduced density matrices and entanglement entropy.
- Measure shortest intervals containing mass $1 - \kappa$ in the entropy distribution.
- Compare concentration across:
  - real spheres,
  - complex projective spaces,
  - symmetric states via Majorana stellar representation.

What the Data Suggests



Entropy vs. ambient dimension



Entropy vs. subsystem dimension

As dimension increases, the entropy distribution concentrates near the maximal value.

Conclusion and Outlook

- Concentration of measure explains generic high entanglement in large bipartite systems.
- Complex projective space provides the natural geometric setting for pure quantum states.
- Observable diameter gives a way to phrase concentration geometrically.
- Ongoing directions:
  - sharper estimates for $\mathbb{C}P^n$,
  - deeper use of Fubini–Study geometry,
  - Majorana stellar representation for symmetric states.

References I


Bengtsson, I. and Życzkowski, K. (2017).
Geometry of Quantum States: An Introduction to Quantum Entanglement.
Cambridge University Press, 2 edition.


Gromov, M. (1981).
Metric structures for Riemannian and non-Riemannian spaces.
Birkhäuser.


Hayden, P. (2010).
Concentration of measure effects in quantum information.
In *Quantum Information Science and Its Contributions to Mathematics*, volume 68 of *Proceedings of Symposia in Applied Mathematics*, pages 211–260. American Mathematical Society.

References II


Hayden, P., Leung, D. W., and Winter, A. (2006).
Aspects of generic entanglement.
Communications in Mathematical Physics, 265(1):95–117.


Shioya, T. (2014).
Metric measure geometry.


Vershynin, R. (2018).
High-dimensional probability: an introduction with applications in data science.
Cambridge University Press.

Q&A