How much entropy is in quantum non-locality?

Greg Kuperberg

UC Davis

July 23, 2009

(Related results by Wim van Dam and Richard Gill.)
Alice and Bob claim telepathy

A traditional example
Alice and Bob claim telepathy

A traditional example

Alice and Bob claim a form of telepathy. Alice is asked to choose between north or south, or between east and west.

- Interrogators choose the lines randomly.
- Questions continue round by round.
- Alice and Bob score 1 point for acute answers.
Alice and Bob claim telepathy

A traditional example

Alice and Bob claim a form of telepathy. Alice is asked to choose between north or south, or between east and west.

- Interrogators choose the lines randomly.
- Questions continue round by round.
- Alice and Bob score 1 point for acute answers.
Alice and Bob claim telepathy

A traditional example

Alice and Bob claim a form of telepathy. Alice is asked to choose between north or south, or between east and west.

- Interrogators choose the lines randomly.
- Questions continue round by round.
- Alice and Bob score 1 point for acute answers.
If Alice and Bob play randomly, they are 50% acute.

If they share data, i.e., plan ahead, they can score up to 75%.

Scoring more than 75% is classically impossible.
Alice and Bob claim telepathy

- If Alice and Bob play randomly, they are 50% acute.
- If they share data, *i.e.*, plan ahead, they can score up to 75%.
- Scoring more than 75% is classically impossible.
Alice and Bob claim telepathy

- If Alice and Bob play randomly, they are 50% acute.
- If they share data, i.e., plan ahead, they can score up to 75%.
- Scoring more than 75% is classically impossible.
- If they share quantum data, they can score

\[(\cos \frac{\pi}{8})^2 \approx 85.35\%\]
Quantum “telepathy” is real
(But it is not really telepathy.)

• Bell discovered this type of demonstration of quantum “non-locality.” The 75% classical bound is a Bell-type inequality. This one is the CHSH $\leq$ (Clauser, Horne, Shimony, and Holt).

• Alain Aspect’s experiment (and many since) confirmed quantum non-locality.

• The dynamics of wave functions, Schrödinger equations, etc., is not directly the point. Probability theory needs to change. Quantum probability is a more correct generalization.

• (In a sense, the dynamics doesn’t change at all.)
(Also, it is all quantumly local; it is not true non-locality.)
Quantum “telepathy” is real
(But it is not really telepathy.)

• Bell discovered this type of demonstration of quantum “non-locality.” The 75% classical bound is a Bell-type inequality. This one is the CHSH ≤ (Clauser, Horne, Shimony, and Holt).

• Alain Aspect’s experiment (and many since) confirmed quantum non-locality.

• The dynamics of wave functions, Schrödinger equations, etc., is not directly the point. Probability theory needs to change. Quantum probability is a more correct generalization.

• (In a sense, the dynamics doesn’t change at all.) (Also, it is all quantumly local; it is not true non-locality.)
What is quantum probability?

Answer: Non-commutative probability

In advanced probability, we see random variable algebras:

Ω - a σ-algebra of boolean variables
\( \mathcal{M} = L^\infty(\Omega) \) - algebra of bounded complex random variables

The algebra \( \mathcal{M} \) can be described by axioms:

- It is a commutative algebra with * (for \( \mathbb{C} \) conjugation).
- It is a Banach space, and \( ||A^*A|| = ||A||^2 \).
- It has a pre-dual \( \#\mathcal{M} \). (\( \#\mathcal{M} \cong L^1(\Omega) \))

This makes \( \mathcal{M} \) a commutative von Neumann algebra. Quantum probability is exactly the same, except that \( \mathcal{M} \) can be non-commutative.
What is quantum probability?
Answer: Non-commutative probability

In advanced probability, we see random variable algebras:

\[ \Omega - \text{a } \sigma\text{-algebra of boolean variables} \]
\[ M = L^\infty(\Omega) - \text{algebra of bounded complex random variables} \]

The algebra \( M \) can be described by axioms:

- It is a commutative algebra with \( * \) (for \( \mathbb{C} \) conjugation).
- It is a Banach space, and \( ||A^*A|| = ||A||^2 \).
- It has a pre-dual \( \#M \). (\( \#M \cong L^1(\Omega) \))

This makes \( M \) a commutative \textbf{von Neumann algebra}.
Quantum probability is exactly the same, except that \( M \) can be non-commutative.
What is quantum probability?

**Answer:** Non-commutative probability

In advanced probability, we see random variable algebras:

\[ \Omega - \text{a } \sigma\text{-algebra of boolean variables} \]
\[ M = L^\infty(\Omega) - \text{algebra of bounded complex random variables} \]

The algebra \( M \) can be described by axioms:

- It is a commutative algebra with \( \ast \) (for \( \mathbb{C} \) conjugation).
- It is a Banach space, and \( ||A^\ast A|| = ||A||^2 \).
- It has a pre-dual \( \#M \). (\( \#M \cong L^1(\Omega) \))

This makes \( M \) a commutative von Neumann algebra.

Quantum probability is exactly the same, except that \( M \) can be non-commutative.
The simplest example, and its states

Example: The $2 \times 2$ matrix algebra $\mathcal{M}_2$ is called a qubit.

As before, a state (= measure = distribution) is an expectation functional $\rho : \mathcal{M} \to \mathbb{C}$ which is $\geq 0$ on $\mathcal{M}_{\text{bool}}$, and s.t. $\rho(1) = 1$.

The state region of a classical trit $3\mathbb{C}$ vs that of a qubit $\mathcal{M}_2$:
Details of quantum probability

Probabilities and vectors

- The qubit probability of a boolean question is the height in the question’s direction.
- What about vector states $|\psi\rangle \in \mathbb{C}^d$? A qudit $\mathcal{M}_d$ has vector states $\rho(A) = \langle \psi | A | \psi \rangle$, but also other states (mixed states).

Joint systems

- If $A$ and $B$ are two algebras, their joint algebra is $A \otimes B$.
- In free probability, it is $A \ast B$. But this is less empirical.
- It is certainly not $A \times B$, either classically or quantumly.
Details of quantum probability

Probabilities and vectors

- The qubit probability of a boolean question is the height in the question’s direction.
- What about vector states $|\psi\rangle \in \mathbb{C}^d$? A qudit $\mathcal{M}_d$ has vector states $\rho(A) = \langle \psi | A | \psi \rangle$, but also other states (mixed states).

Joint systems

- If $A$ and $B$ are two algebras, their joint algebra is $A \otimes B$.
- In free probability, it is $A \ast B$. But this is less empirical.
- It is certainly not $A \times B$, either classically or quantumly.
How to obtain 85.35%

\[ |\psi_{AB}\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}} \]

- Alice and Bob should measure an entangled qubit pair in the requested directions.
- “Entanglement” is just correlation in quantum probability.
- This is not action at a distance. It is the same as if Alice “changed” Bob’s poker hand by reading her poker hand.
How to obtain 85.35%

\[ |\psi_{AB}\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}} \]

- Alice and Bob should measure an entangled qubit pair in the requested directions.
- “Entanglement” is just correlation in quantum probability.
- This is not action at a distance. It is the same as if Alice “changed” Bob’s poker hand by reading her poker hand.
How quickly are we persuaded?

The *Kullback-Leibler divergence* (= relative entropy) between classical states \( p \) and \( q \) expresses how quickly samples from \( p \) convince you that they are not from \( q \):

\[
D(p \| q) \overset{\text{def}}{=} \sum_{\alpha} p_{\alpha} \ln \frac{p_{\alpha}}{q_{\alpha}}.
\]

We want the evidence of non-locality in one round of CHSH:

\[
\frac{D(q \| c)}{\ln 2} = q_{\text{acute}} \log_2 \frac{q_{\text{acute}}}{c_{\text{acute}}} + q_{\text{obtuse}} \log_2 \frac{q_{\text{obtuse}}}{c_{\text{obtuse}}} \approx 4.63%
\]

This is not very big!
How quickly are we persuaded?

The *Kullback-Leibler divergence* (= relative entropy) between classical states $p$ and $q$ expresses how quickly samples from $p$ convince you that they are not from $q$:

$$D(p\|q) \overset{\text{def}}{=} \sum_{\alpha} p_{\alpha} \ln \frac{p_{\alpha}}{q_{\alpha}}.$$  

We want the evidence of non-locality in one round of CHSH:

$$\frac{D(q\|c)}{\ln 2} = q_{\text{acute}} \log_2 \frac{q_{\text{acute}}}{c_{\text{acute}}} + q_{\text{obtuse}} \log_2 \frac{q_{\text{obtuse}}}{c_{\text{obtuse}}} \approx 4.63\%$$

This is not very big!
How quickly can we be persuaded?

Proposition

Let $A = B = M_d$ be two qudits in a state $\rho$. If $q$ is any non-locality protocol, there exists a skeptical classical $c$ such that

$$D(q||c) \leq E_{RE}(\rho) \leq \ln d.$$ 

- The proof uses quantum relative entropy $D(\rho||\sigma)$, where $\sigma$ is the best “skeptical” separable state. For this $\sigma$,

  $$E_{RE} \overset{\text{def}}{=} D(\rho||\sigma)$$

  is the relative entropy of entanglement.

- The bound also applies to $\geq 2$ rounds or $\geq 3$ parties, and interrogators can confer.
Setting up the question

- Minimax of $D(q||c)$ can be viewed as a two-team game.
- $D(c||q)$ is much less interesting; it can be $\infty$ when $d = 3$.
- Allowing interrogators to talk between rounds is debatable.
- But if skeptics can share information, why not also interrogators?

- Our constructions are 1-round with correlated questions.
A shuffling principle

- If $\rho$ or $|\psi\rangle$ is a maximally entangled state on two qudits, it has $U(d)_{\Delta}$ symmetry.
- Interrogators should symmetrize or “shuffle” their questions. Then skeptics should too.
- Maximal questions should also be $S_d$-shuffled.

\[
U(d)_{\Delta} \times S_d \times S_d
\]

\[
U(d)_{\Delta} \times (S_d)_{\Delta}
\]
Two qubits

- Shuffling reduces interrogation to choosing an angle $\alpha$. CHSH has $\alpha = \frac{\pi}{4}$, but $\alpha = \frac{\pi}{8}$ is better.
- Skeptics can play Grothendieck’s (!) hemisphere strategy for all $\alpha$.
- We obtain $6.6167\% \lesssim D_2 \lesssim 6.6287\%$ (with $\alpha = 22.5^\circ$ and $\alpha \approx 23.81^\circ$.)

- The hemisphere strategy is not optimal for large $\alpha$. But is it optimal for all good $\alpha$?
- POVM questions are more general than these (projective) questions. Surely they yield worse protocols?
Two qubits

- Shuffling reduces interrogation to choosing an angle $\alpha$. CHSH has $\alpha = \frac{\pi}{4}$, but $\alpha = \frac{\pi}{8}$ is better.
- Skeptics can play Grothendieck’s (!) hemisphere strategy for all $\alpha$.
- We obtain $6.6167\% \lesssim D_2 \lesssim 6.6287\%$ (with $\alpha = 22.5^\circ$ and $\alpha \approx 23.81^\circ$).

- The hemisphere strategy is not optimal for large $\alpha$. But is it optimal for all good $\alpha$?
- POVM questions are more general than these (projective) questions. Surely they yield worse protocols?
Two qubits

- Shuffling reduces interrogation to choosing an angle $\alpha$. CHSH has $\alpha = \frac{\pi}{4}$, but $\alpha = \frac{\pi}{8}$ is better.
- Skeptics can play Grothendieck’s (!) hemisphere strategy for all $\alpha$.
- We obtain $6.6167\% \lesssim D_2 \lesssim 6.6287\%$ (with $\alpha = 22.5^\circ$ and $\alpha \approx 23.81^\circ$.)

- The hemisphere strategy is not optimal for large $\alpha$. But is it optimal for all good $\alpha$?
- POVM questions are more general than these (projective) questions. Surely they yield worse protocols?
Peres’ protocol

- Asher Peres defined a non-locality protocol for $d = 4$ qudits (ququats). It is based on the 24-cell in $\mathbb{R}^4 \subseteq \mathbb{C}^4$ and the dual 24-cell.

- The 12 diagonals of a 24-cell partition into 3 $\perp$ frames. Alice chooses 1 line from a random frame. Bob uses the dual 24-cell. Quantumly, $P[\perp] = 0$; classically, $P[\perp] \geq \frac{1}{9}$.

- Peres has much better divergence:

  \[
  \frac{D(q_{\text{Peres}} \mid\mid c)}{\ln 4} = \log_4 \frac{9}{8} \approx 8.50\%.
  \]

- Peres’ protocol can also be defined by $\otimes$ products of Pauli matrices. E.g., Alice measures $X \otimes X$ and $Z \otimes Z$. 
Peres’ protocol

• Asher Peres defined a non-locality protocol for $d = 4$ qudits (ququats). It is based on the 24-cell in $\mathbb{R}^4 \subseteq \mathbb{C}^4$ and the dual 24-cell.

• The 12 diagonals of a 24-cell partition into 3 $\perp$ frames. Alice chooses 1 line from a random frame. Bob uses the dual 24-cell. Quantumly, $P[\perp] = 0$; classically, $P[\perp] \geq \frac{1}{9}$.

• Peres has much better divergence:

$$D(q_{\text{Peres}}\|c) = \frac{\log_4 \frac{9}{8}}{\ln 4} \approx 8.50\%.$$  

• Peres’ protocol can also be defined by $\otimes$ products of Pauli matrices. E.g., Alice measures $X \otimes X$ and $Z \otimes Z$. 

Generalizing Peres

- We generalize Peres to $d = 2^n$ using larger $\otimes$ products:

$$A = Z \otimes I \otimes I, I \otimes Z \otimes I, I \otimes I \otimes Z \quad B = X \otimes I \otimes I, I \otimes Z \otimes I, I \otimes I \otimes Z$$

The frames make an “orthogonal spread” (Calderbank, Rains, Shor, Sloane).

- We obtain:

$$\frac{D(q_8 \| c)}{\ln 8} = \log_8 \frac{5}{4} \approx 10.7\% \quad \frac{D(q_{16} \| c)}{\ln 16} = \log_{16} \frac{45}{31} \approx 13.4\%.$$ 

- Actually $c$ is optimized by computer and non-rigorously.

- Only original Peres has uncorrelated questions.

- What happens as $n \to \infty$?
What I really think

Conjecture

\[
\lim_{d \to \infty} \frac{D(q_{\text{max}} \| c_{\text{min}})}{\ln d} = 1.
\]

This is suggested by the isoperimetric ≤ in high dimensions: A spherical region in \(d \to \infty\) dimensions is concentrated at its boundary.

Theorem

\[
\lim_{n \to \infty} \frac{D(q_n \| c_{\text{min}})}{E_{\text{ER}}(|\text{cat}_n\rangle)} = 1
\]

for cat states like

\[
|\text{cat}_5\rangle = \frac{|00000\rangle + |11111\rangle}{\sqrt{2}}.
\]
What I really think

Conjecture

\[ \lim_{d \to \infty} \frac{D(q_{\text{max}} \| c_{\text{min}})}{\ln d} = 1. \]

This is suggested by the isoperimetric inequality in high dimensions: A spherical region in \( d \to \infty \) dimensions is concentrated at its boundary.

Theorem

\[ \lim_{n \to \infty} \frac{D(q_n \| c_{\text{min}})}{E_{\text{ER}}(\ket{\text{cat}_n})} = 1 \]

for cat states like

\[ \ket{\text{cat}_5} = \frac{\ket{00000} + \sqrt{2} \ket{11111}}{\sqrt{2}}. \]
Asymptotic geometry problems

• A maximal question in $\mathcal{M}_d$ is a frame, i.e., an orthonormal line basis. The set of frames is a flag manifold. A lower bound on $D(q_d||c_{\min})$ would be an isoperimetric inequality for flag manifolds.

• The hemisphere strategy generalizes to a Voronoi strategy in $\mathbb{C}P^{d-1}$. Alice and Bob each pick the closest answer to a shared random line. I do not know how to compute or estimate its performance relative to a fixed “angle” between two frames, nor how to optimize or find asymptotics. Is it asymptotically optimal?
Asymptotic geometry problems

• A maximal question in $\mathcal{M}_d$ is a frame, i.e., an orthonormal line basis. The set of frames is a flag manifold. A lower bound on $D(q_d\|c_{\min})$ would be an isoperimetric inequality for flag manifolds.

• The hemisphere strategy generalizes to a Voronoi strategy in $\mathbb{CP}^{d-1}$. Alice and Bob each pick the closest answer to a shared random line. I do not know how to compute or estimate its performance relative to a fixed “angle” between two frames, nor how to optimize or find asymptotics. Is it asymptotically optimal?
Other topics

There are other ways in which quantum probability differs from classical probability, or is interesting in pure mathematics.

- Perpetual randomness. One qubit provides an $\infty$ sequence of Bernoulli variables.
- Quantum key distribution = non-cryptographic secrecy. If $\exists$ an undetected eavesdropper, then quantum probability is false.
- Quantum computation. Quantum probability yields a larger complexity class: BQP vs BPP.
- Quantum probability proofs of classical probability theorems.
- Is there a quantum probabilistic method?