UNIVERSITY OF OXFORD

Spring School --

Quantum Structures in Physics and Computer Science

Mathematical Institute, University of Oxford, 19-22 May 2014

Correlations and contexts: the quantum no-go theorems

Jonathan Barrett

Introduction

- The three lectures will focus on three well known no-go theorems for quantum theory:
  - Bell's theorem (nonlocality)
  - The Kochen-Specker theorem (contextuality)
  - The PBR theorem (psi-ontology)

- Roughly speaking, a no-go theorem states that some class of theories/models either doesn't exist (KS) or must make different predictions from quantum theory (Bell, PBR). Experimental results have always confirmed quantum theory.

- Significant for two reasons:
  - Foundational. There is no underlying theory waiting to be discovered that is locally causal/non-contextual/psi-epistemic.
  - Simulation. Quantum theory cannot be simulated by a model that is locally causal/etc. This can lead to quantum advantages in information processing tasks.

Table of contents

Part 1: Nonlocality

The CHSH game

$X, Y, a, b \in \{0, 1\}$

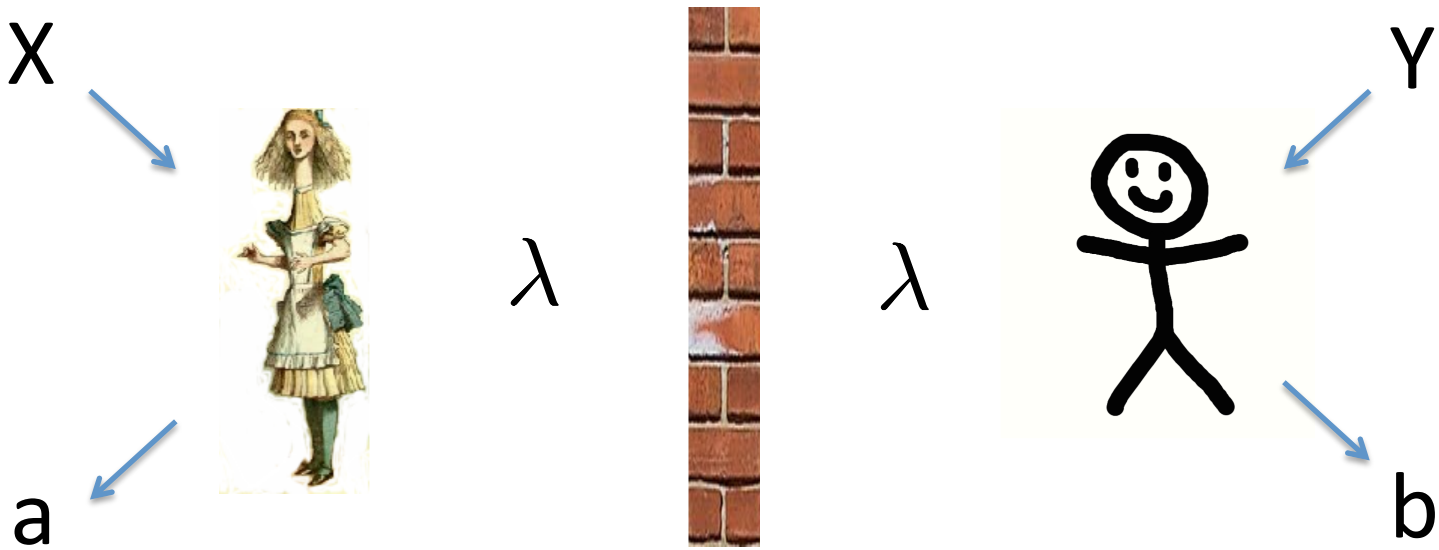
Alice, Bob win if $a \oplus b = XY$.

The CHSH game

$X, Y, a, b \in \{0, 1\}$
Alice, Bob win if $a \oplus b = XY$.

Equivalently, win if
$XY = 00, 01, 10 \rightarrow a = b$
$XY = 11 \rightarrow a \neq b$

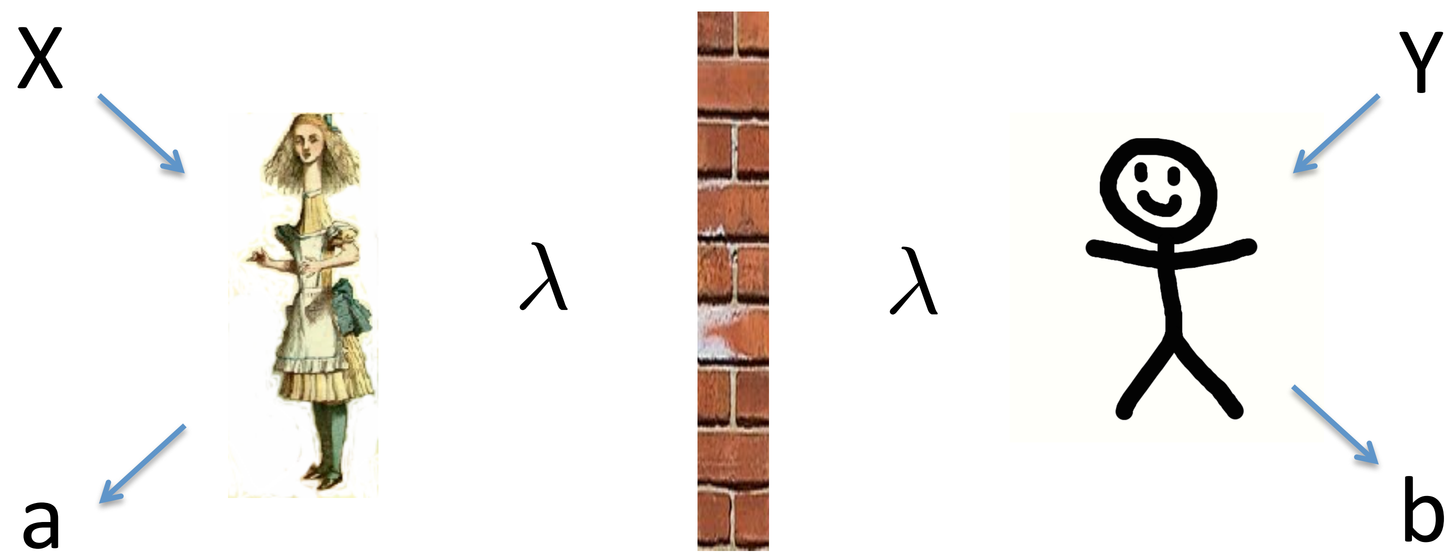
The CHSH game



$X, Y, a, b \in \{0, 1\}$
Alice, Bob win if $a \oplus b = XY$.

Equivalently, win if
$XY = 00, 01, 10 \rightarrow a = b$
$XY = 11 \rightarrow a \neq b$

The CHSH game



$X, Y, a, b \in \{0, 1\}$
Alice, Bob win if $a \oplus b = XY$.

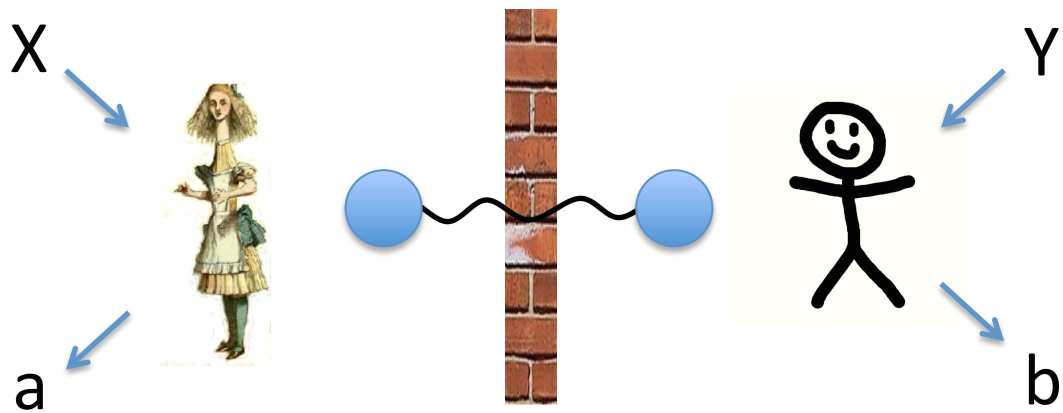
$$\mathrm{P}(\text{win}) \leq \frac{3}{4}$$

Equivalently, win if
$XY = 00, 01, 10 \rightarrow a = b$
$XY = 11 \rightarrow a \neq b$

The CHSH game

Quantum players…

$$|\phi_+\rangle = \frac{1}{2}(|0\rangle|0\rangle + |1\rangle|1\rangle)$$



$$\begin{array}{llll} X = 0 & \rightarrow & \text{Alice measures} & \{|0\rangle, |1\rangle\} \\ X = 1 & \rightarrow & \text{Alice measures} & \{|+\rangle, |-\rangle\} \end{array}$$

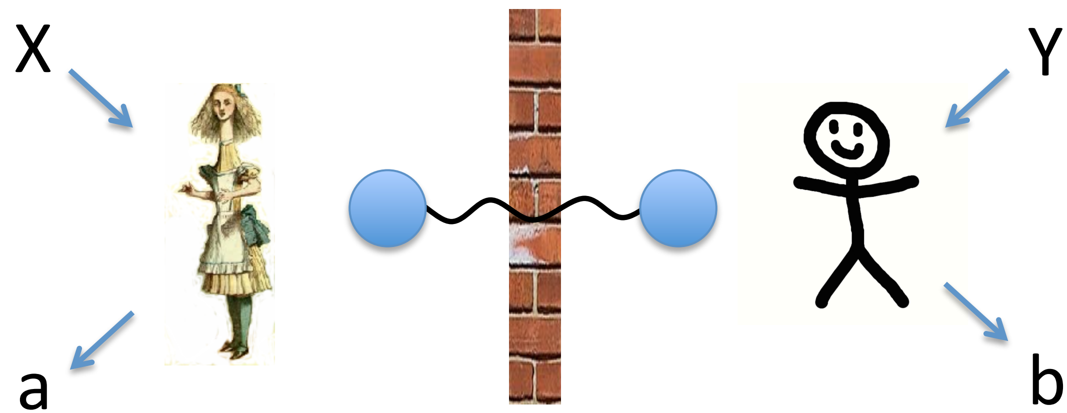
$$\begin{array}{llll} Y = 0 & \rightarrow & \text{Bob measures} & \{|\phi_0\rangle, |\phi_0^\perp\rangle\} \\ Y = 1 & \rightarrow & \text{Bob measures} & \{|\phi_1\rangle, |\phi_1^\perp\rangle\} \end{array}$$

$$\begin{array}{lcl} |+\rangle & = & 1/\sqrt{2}(|0\rangle + |1\rangle) \\ |-\rangle & = & 1/\sqrt{2}(|0\rangle - |1\rangle) \\ |\phi_0\rangle & = & \cos\pi/8|0\rangle + \sin\pi/8|1\rangle, \\ |\phi_1\rangle & = & \cos 3\pi/8|0\rangle + \sin 3\pi/8|1\rangle \end{array}$$

The CHSH game

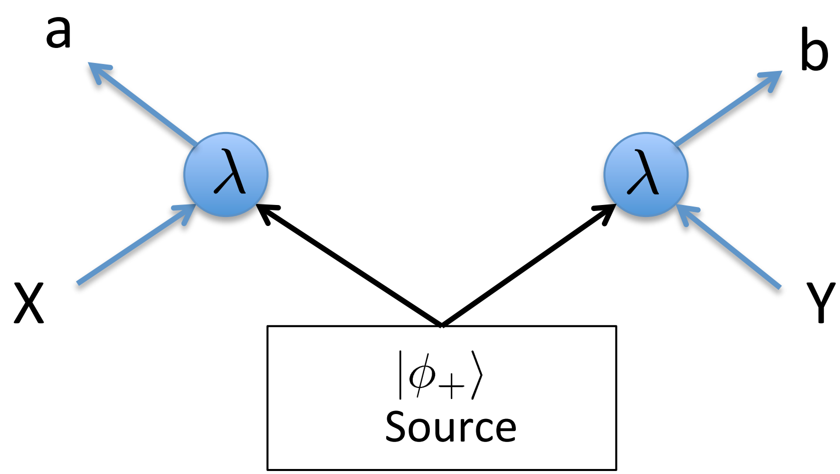
Quantum players…

$$|\phi_+\rangle = \frac{1}{2}\left(|0\rangle|0\rangle + |1\rangle|1\rangle\right)$$



$$\mathrm{P(win)} = \cos^2 \pi/8 = \frac{2+\sqrt{2}}{4} \approx 0.85 > 3/4$$

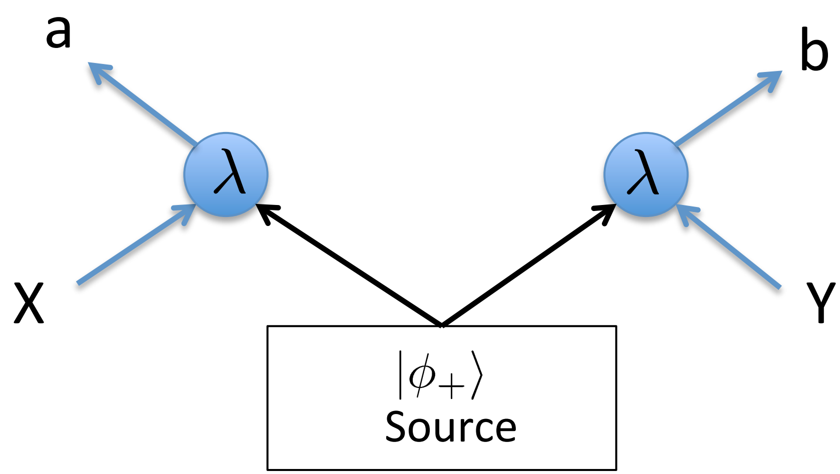
In terms of hidden variables…



- Particles are produced at a source in the quantum state $|\phi_+\rangle$.
- Measurements X,Y are performed.
- Imagine that $|\phi_+\rangle$ is an *incomplete description* of the particles, reflecting our ignorance of the values of some underlying variables λ. The quantum state determines a probability distribution $\mu(\lambda)$. Probabilities for outcomes a,b are determined by λ.

$$\Rightarrow \quad P(ab|XY) = \int d\lambda\ \mu(\lambda) P(ab|XY\lambda)$$

In terms of hidden variables…



Bell locality (*) is the condition that: $P(ab|XY\lambda) = P(a|X\lambda)P(b|Y\lambda)$

But winning the CHSH game with probability > 3/4 implies that the correlations *cannot* be written in the form:

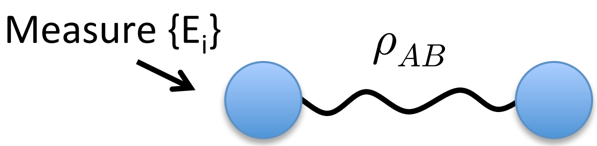
$$P(ab|XY) = \int d\lambda\ \mu(\lambda)\ P(a|X\lambda)\ P(b|Y\lambda)$$

(*) J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics* (CUP, Cambridge 2004).

What is going on?

- It is “as if” the quantum particles “talk to each other”.
- There are models, e.g., the de Broglie-Bohm theory that define an explicit space of hidden variables λ, and where Bell locality fails: $P(a|XY\lambda) \neq P(a|XY'\lambda)$
  Such models are *nonlocal*.
- Standard quantum theory evades the question: correct predictions are obtained for the observed correlations, but without describing any kind of explicit mechanism that would mediate the correlations.
- The effect remains even when the measurements are spacelike separated.
- **But** cannot be used directly to signal.

The no-signalling principle



$$\rho_{AB} \in \mathcal{B}(H_A \otimes H_B)$$
$$\rho_{AB} \geq 0$$
$$\mathrm{Tr}(\rho_{AB}) = 1$$

If Alice gets outcome i, then Bob's collapsed state is: $\rho_B^i \equiv \dfrac{\mathrm{Tr}(E_i \otimes I \rho_{AB})}{\mathrm{Prob}(i)}$

Averaging over Alice's outcomes,

$$\sum_i \mathrm{Prob}(i)\rho_B^i = \sum_i \mathrm{Tr}_A(E_i \otimes I \rho_{AB}) = \mathrm{Tr}_A(I \otimes I \rho_{AB}) = \mathrm{Tr}_A(\rho_{AB}) \equiv \rho_B$$

Hence if Bob performs a measurement and is ignorant of Alice's outcome, his outcome probabilities are *independent of which measurement Alice chose to do*.

Tsirelson's Bound

Can Alice and Bob do any better using quantum systems?

No. Tsirelson showed (*) that for any shared state ρ_{AB}, and whichever measurements X=0,1 and Y=0,1 correspond to, then P(win) $\leq \frac{2+\sqrt{2}}{4}$

(*) B. S. Tsirelson, Lett Math Phys **4**, 93 (1980).

Proof

Let Alice measure $\{P_X, I - P_X\}$, for P_X a projector, $X = 0, 1$.
Let Bob measure $\{Q_Y, I - Q_Y\}$, for Q_Y a projector, $Y = 0, 1$.
Let

$$\begin{aligned} \hat{A}_0 &= 2P_0 - I, \hat{A}_1 = 2P_1 - I \\ \hat{B}_0 &= 2Q_0 - I, \hat{B}_1 = 2Q_1 - I \end{aligned}$$

Let $\mathcal{G} = \hat{A}_0\hat{B}_0 + \hat{A}_0\hat{B}_1 + \hat{A}_1\hat{B}_0 - \hat{A}_1\hat{B}_1$.
NB $\mathrm{P(win)} = \langle \mathcal{G} \rangle/8 + 1/2$.
Now, $\mathcal{G}^2 = 4I - [\hat{A}_0, \hat{A}_1][\hat{B}_0, \hat{B}_1]$.
But $||[\hat{A}_0, \hat{A}_1]|| \leq 2||\hat{A}_0||||\hat{A}_1|| = 2$, hence $\langle \mathcal{G} \rangle \leq 2\sqrt{2}$.

The PR box

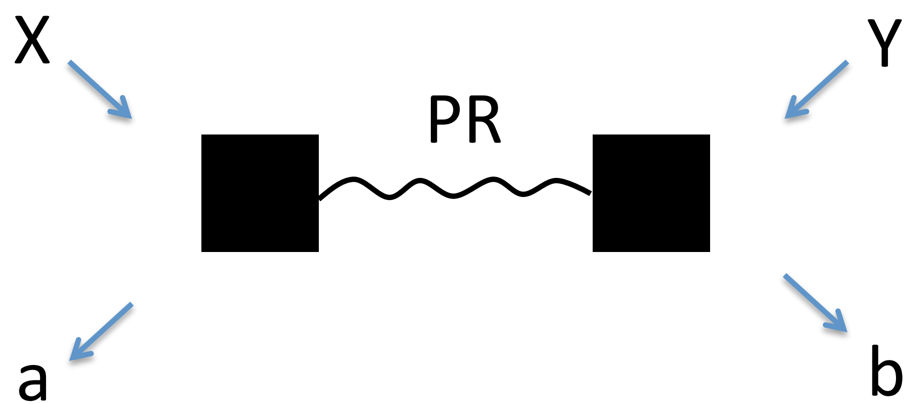
- In quantum theory, can only win the CHSH game with probability $\leq (2+\sqrt{2})/4 \approx 0.85$
- Popescu and Rohrlich wondered, why can't we do better? Could it be that if you win the game with greater probability, then you inevitably violate the no-signalling principle?
- They found that the answer is no: it is possible to imagine a fictional device, which *respects the no-signalling principle*, but which would *allow Alice and Bob to win the game with certainty*.

The PR box

PR – Popescu and Rohrlich

Box – as in "black box"

$$a, b, X, Y \in \{0, 1\}$$



Defined by a set of conditional probability distributions P(ab|XY), with:

$$\begin{array}{lll} XY = 00, 01, 10 & \rightarrow & P(ab = 00|XY) = P(ab = 11|XY) = 1/2 \\ XY = 11 & \rightarrow & P(ab = 01|XY) = P(ab = 10|XY) = 1/2 \end{array}$$

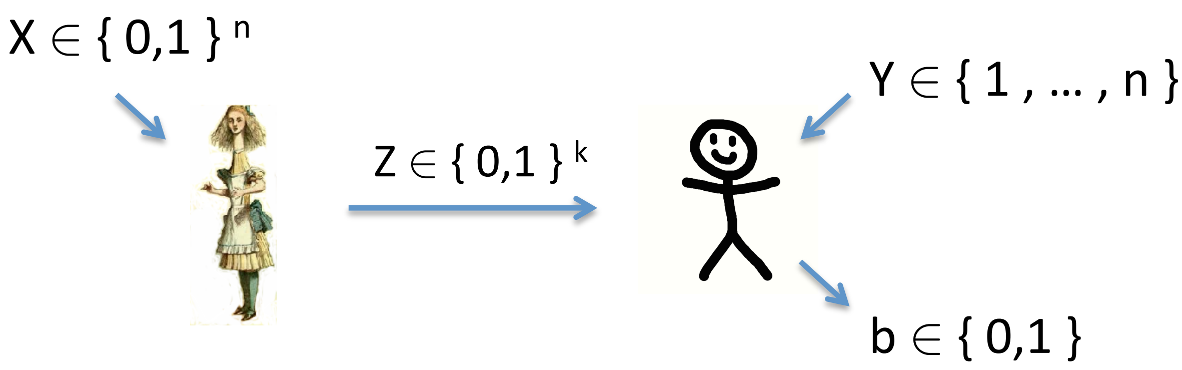
- Outputs always satisfy $a \oplus b = XY$
- Non-signalling

The PR box

These stronger-than-quantum correlations are logically possible, and would be very useful for solving communication complexity problems!

Why doesn't nature allow them?

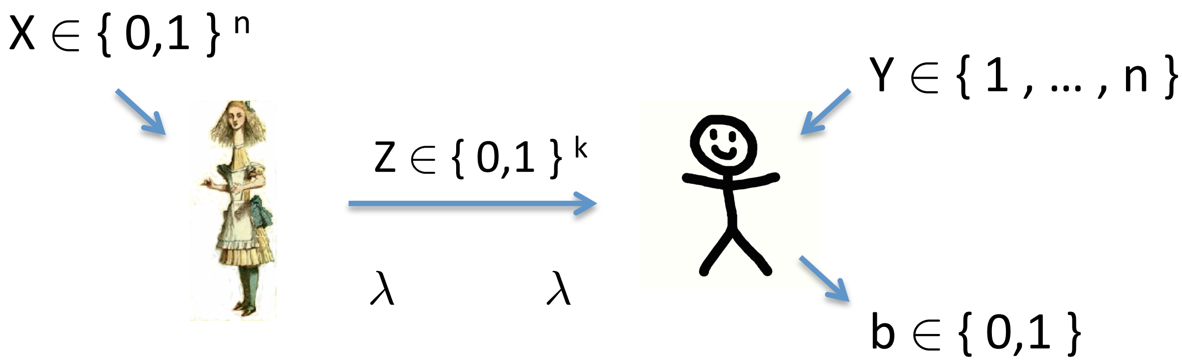
Deriving the quantum bound from a physical principle



(*) See M. Pawlowski et al., Nature **461**, 1101 (2009).

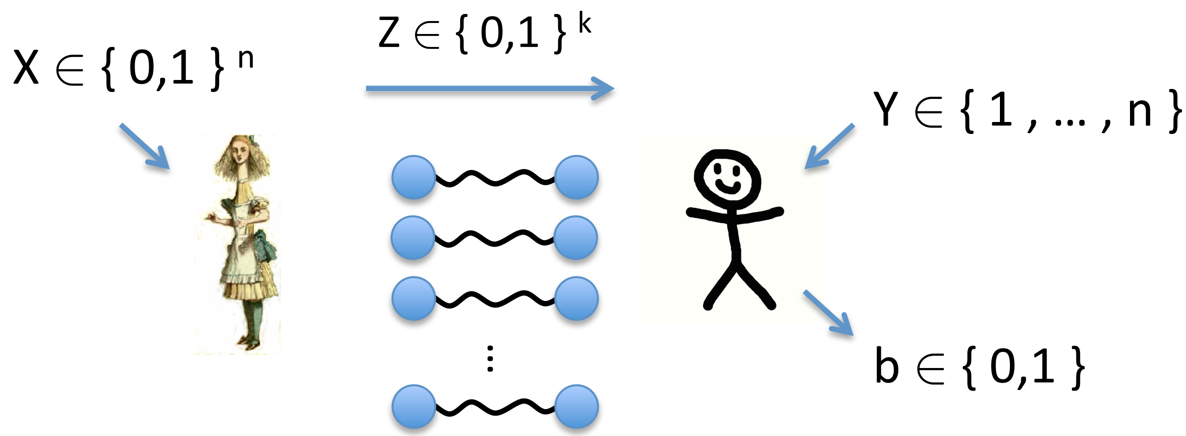
- Alice is given an n bit string X.
- Alice can communicate k classical bits to Bob.
- Bob is given Y, which takes values 1, … , n.
- Bob outputs a single bit b.
- For Y=i, Bob's aim is for b to equal X_i, the *i*th bit of Alice's string X.

Deriving the quantum bound from a physical principle



Classical players have access to pre-shared random data.

Deriving the quantum bound from a physical principle



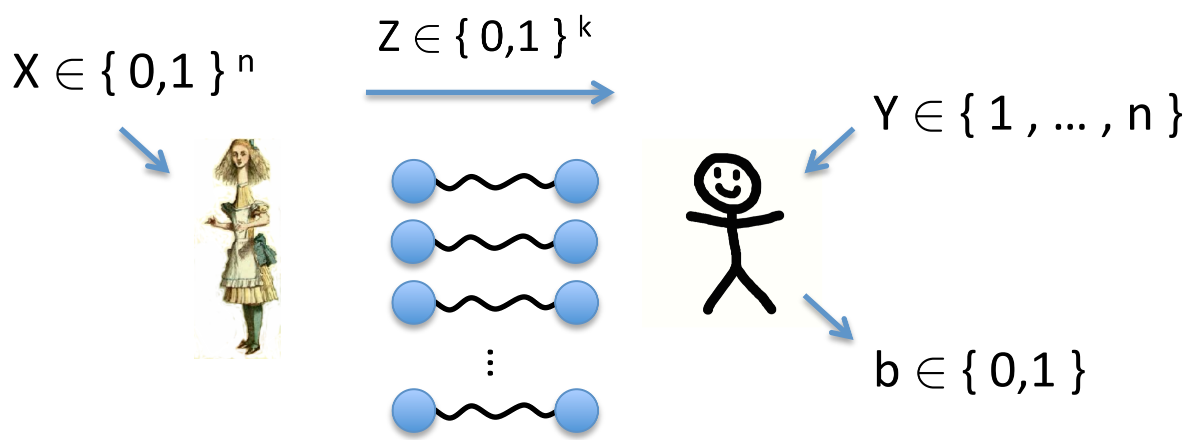
$$|\phi_+\rangle = \frac{1}{2}\left(|0\rangle|0\rangle + |1\rangle|1\rangle\right)$$

Quantum players have pre-shared entanglement.

A protocol goes as follows:

- Alice receives X from a referee, where X is drawn randomly from the set of all n bit strings.
- Alice performs a measurement on her quantum systems.
- Alice sends Z to Bob, where Z depends on X and on the outcome of Alice's measurement.
- Bob measures his quantum systems. His choice of measurement can depend on Y,Z. Bob outputs b, where b depends on Z, Y, and on the outcome of Bob's measurement.

Quantifying the success of a protocol



$$|\phi_+\rangle = \frac{1}{2}\left(|0\rangle|0\rangle + |1\rangle|1\rangle\right)$$

How successful is a given protocol? We wish to quantify the amount of information, on average, that b contains about X_i. Consider the following quantity:

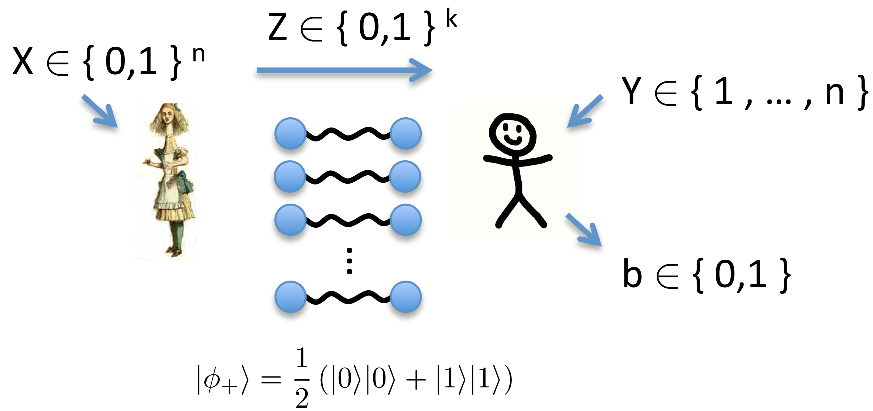
$$\sum_i I(X_i : b|Y = i),$$

where I(X_i : b | Y=i) is the *mutual information* between the variables X_i and b, conditioned on Y taking the value i.

Information causality

Information causality is the principle that:

$$\sum_i I(X_i : b | Y = i) \leq k$$



Roughly: Information causality states that no matter what preshared resources they might have, if Alice communicates k bits to Bob, then the total information access that Bob gains to her data is not more than k.

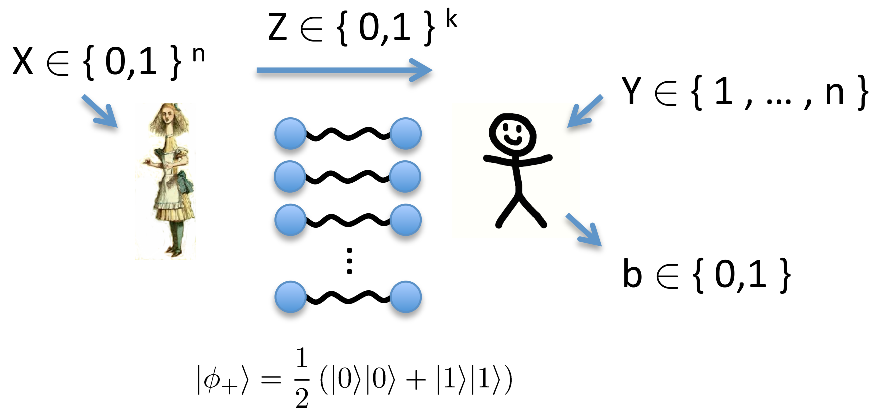
NB For k=0, information causality reduces to the no-signalling principle.

So information causality is a more refined notion of the no-signalling principle, that takes into account a finite amount of communication.

Information causality

Information causality is the principle that:

$$\sum_i I(X_i : b|Y = i) \leq k$$



Theorem: Any classical or quantum protocol satisfies information causality.

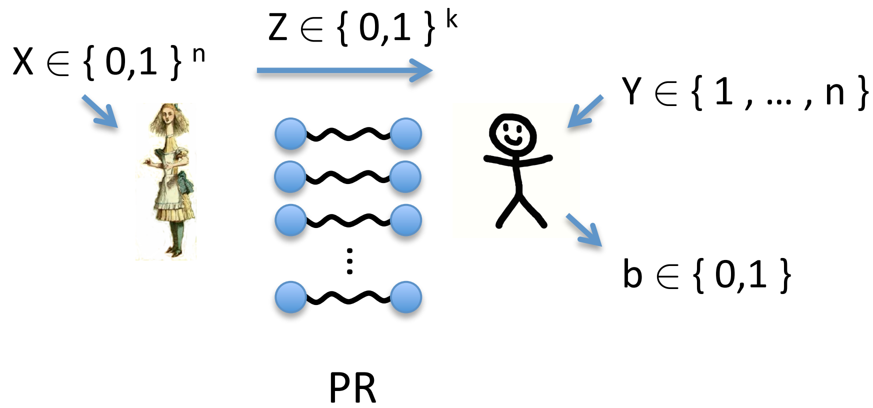
NB Even classical players can saturate the information causality bound.

Simply let Z be equal to the first k bits of X. Bob gets lucky if $Y \in \{1, \ldots, k\}$, and can output the correct value. Otherwise Bob must simply guess.

PR box players

Information causality is the principle that:

$$\sum_i I(X_i : b|Y = i) \leq k$$



What if the players have PR boxes?

Consider the following protocol. Alice and Bob share n PR boxes.

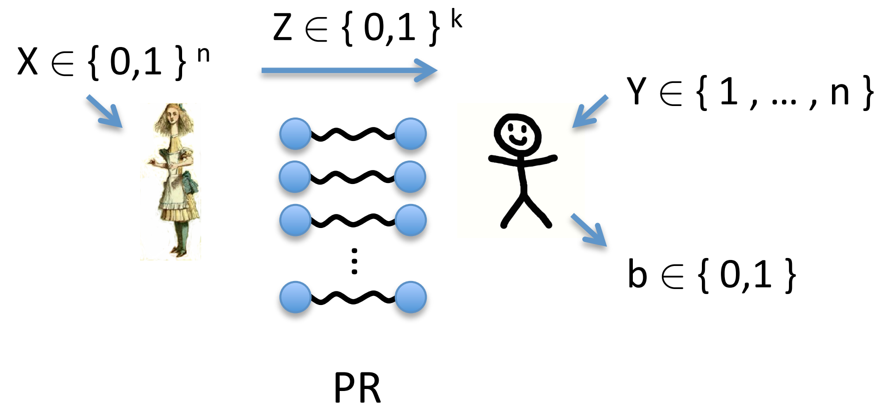
- Alice inputs X_1 , … , X_n into the n PR boxes, obtaining outputs a_1 , … , a_n .
- Alice evaluates $Z = a_1 \oplus \cdots \oplus a_n$ and sends Z to Bob.
- For Y = i, Bob inputs 0 into all PR boxes, except for the ith, which has input 1. Bob obtains outputs c_1 , … , c_n .
- Bob evaluates $b = c_1 \oplus \cdots \oplus c_n \oplus Z$.
- Check: $b = X_i$

$$\Rightarrow \quad \sum_i I(X_i : b|Y = i) = n$$

Here's the interesting bit

Information causality:

$$\sum_i I(X_i : b|Y = i) \leq k$$



What if the players have *noisy* PR boxes?

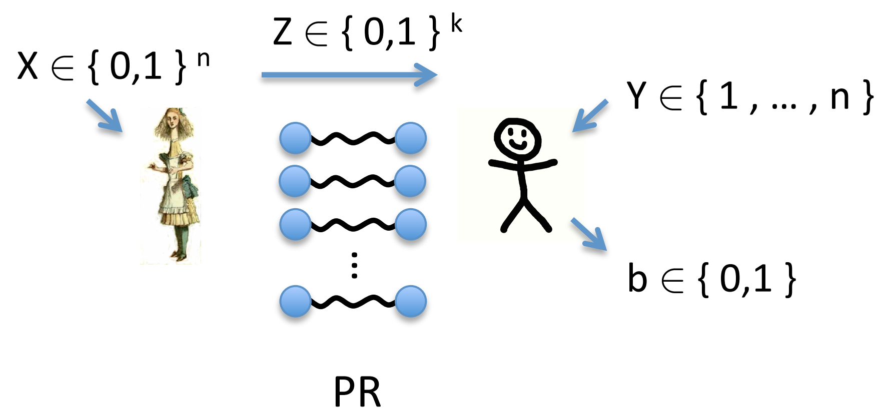
A noisy PR box is defined by a set of conditional probability distributions P(ab|XY), with:

- P(a=0|XY) = P(a=1|XY) = P(b=0|XY) = P(b=1|XY) = 1/2.
- P(a$\oplus$ b = XY) = q , where ½ $\leq$ q < 1 .
- Super-quantum iff q > $(2+\sqrt{2})/4$

Here’s the interesting bit

Information causality:

$$\sum_i I(X_i : b|Y = i) \leq k$$



Theorem (*). With noisy PR boxes, information causality can be violated iff the boxes are super-quantum.

So the principle of information causality picks out the quantum bound exactly.

(*) M. Pawlowski et al., *ibid*.

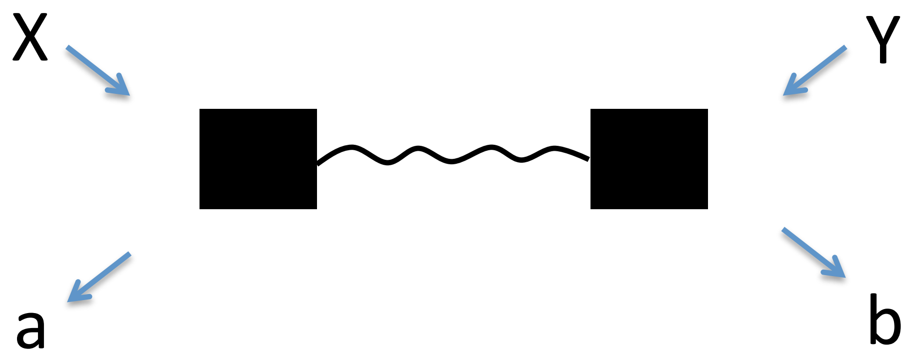
Pause

- So we've seen that quantum systems can produce *nonlocal correlations*.
- Logically, one can imagine correlations that are *more nonlocal* than is possible in quantum theory, yet which still respect the no-signalling principle.
- But with correlations that are even a little bit more nonlocal than quantum, information causality is violated.

Pause

- Now we turn to some different problems:
- Can we characterize completely the problem of distinguishing local and nonlocal correlations?
- Can we generalize the Bell scenario that we have been considering?

Characterizing nonlocal correlations

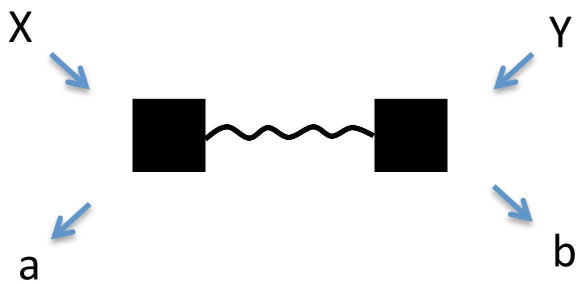


$$
\begin{aligned}
X &\in \{0,\ldots,k-1\}\\
Y &\in \{0,\ldots,k-1\}\\
a &\in \{0,\ldots,d-1\}\\
b &\in \{0,\ldots,d-1\}
\end{aligned}
$$

$$
\mathbf{V} = \begin{pmatrix} P(ab=00|XY=00)\\ P(ab=01|XY=00)\\ \vdots \\ P(ab=(d-1)(d-1)|XY=00)\\ \vdots \\ P(ab=(d-1)(d-1)|XY=(k-1)(k-1)) \end{pmatrix}
$$

Given **V**, how do we tell if these correlations are local or nonlocal?
How do we tell if they can be produced by measuring quantum states?

Characterizing nonlocal correlations



$$\mathbf{V} = \begin{pmatrix} P(ab=00|XY=00) \\ P(ab=01|XY=00) \\ \vdots \\ P(ab=(d-1)(d-1)|XY=00) \\ \vdots \\ P(ab=(d-1)(d-1)|XY=(k-1)(k-1)) \end{pmatrix}$$

$\mathbf{V}$ lives in a $k^2 d^2$-dimensional real vector space.

However, $\mathbf{V}$ satisfies a number of equalities:

Normalization:
$\sum_{ij} P(ab=ij|XY) = 1 \quad \forall X, Y$

No-signalling:
$\sum_j P(ab=ij|XY) = \sum_j P(ab=ij|XY') \quad \forall i, X, Y, Y'$
$\sum_i P(ab=ij|XY) = \sum_i P(ab=ij|X'Y) \quad \forall j, X, X', Y$

Characterizing nonlocal correlations

Hence the allowed V span an affine subspace of the k^2d^2-dimensional vector space. Taking into account the number of independent equalities, the dimension of the affine subspace spanned by allowed $\mathbf{V}$ is:

$$k^2(d-1)^2 + 2k(d-1)$$

$\mathbf{V}$ also satisfies inequalities: $P(ab|XY) \geq 0 \quad \forall a, b, X, Y.$

The set of all non-signalling $\mathbf{V}$ is defined by a finite number of (non-strict) inequalities. Normalization ensures that it is bounded. Hence the set is a polytope – the convex hull of a finite number of extremal points. Let's call it $\mathcal{P}$.

Characterizing nonlocal correlations

A set of correlations is *local* if $P(ab|XY)$ can be written in the form:

$$P(ab|XY) = \int d\lambda \; \mu(\lambda) \; P(a|X\lambda)P(b|Y\lambda).$$

Otherwise the correlations are *nonlocal*.

Theorem. *A set of correlations* $P(ab|XY)$ *is local iff* $P(ab|XY)$ *can be written as a convex combination of deterministic correlations, i.e.,*

$$P(ab|XY) = \sum_{\lambda} q(\lambda) D(a|X\lambda) D(b|Y\lambda),$$

where $D(a|X\lambda), D(b|Y\lambda) \in \{0, 1\}$.

Characterizing nonlocal correlations

So, the set $\mathcal{L}$ of all $\mathbf{V}$ that describe local correlations is a polytope. $\mathcal{L}$ is the convex hull of those $\mathbf{V}$ that satisfy:

$$P(ab|XY) = D(a|X)D(b|Y),$$

where $D(a|X), D(b|Y) \in \{0,1\}$.

The facets of $\mathcal{L}$ correspond to inequalities of the form:

$$\sum_{abXY} c_{abXY} P(abXY) \leq 0.$$

These are the *Bell inequalities*. A set of correlations is nonlocal if and only if $\mathbf{V}$ violates a Bell inequality.

Characterizing nonlocal correlations

Correlations are *quantum* if $P(ab|XY)$ can be written

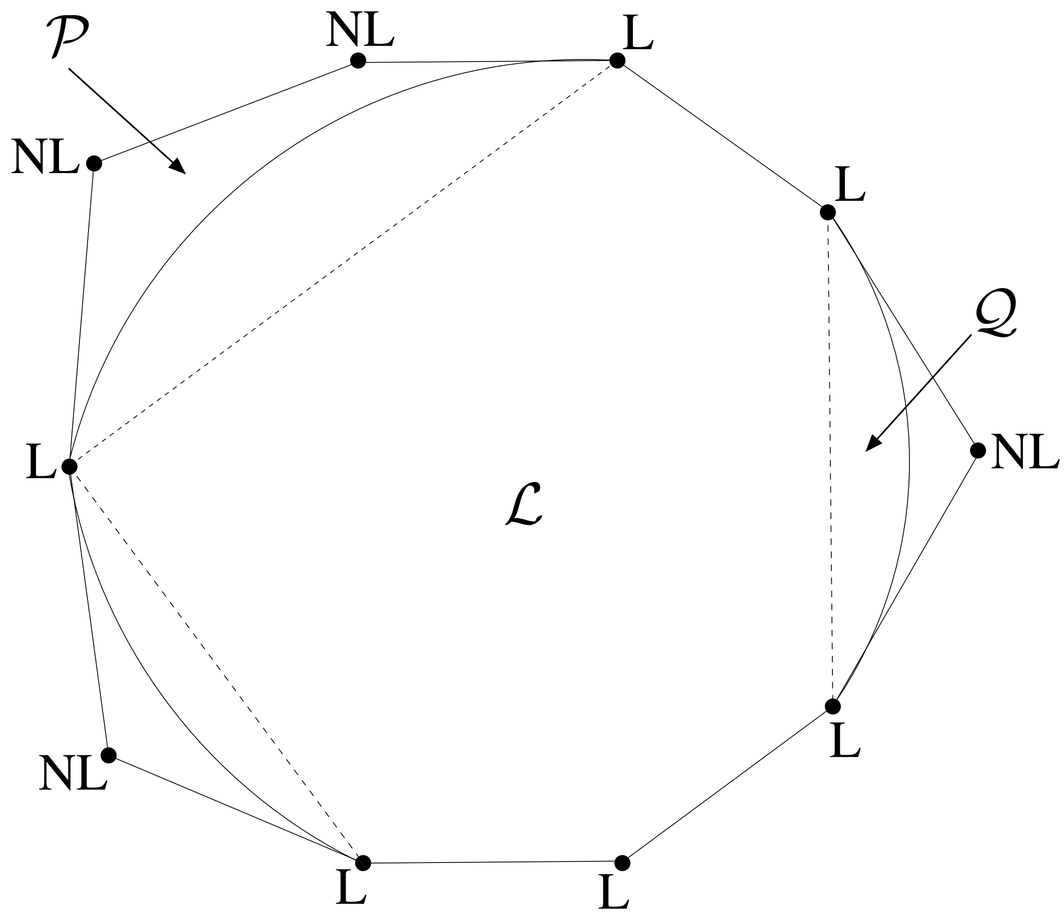
$$P(ab|XY) = \mathrm{Tr}(E_a^X \otimes E_b^Y \rho),$$

with

$$\rho \text{ a density operator } \in \mathcal{B}(H_A \otimes H_B),$$

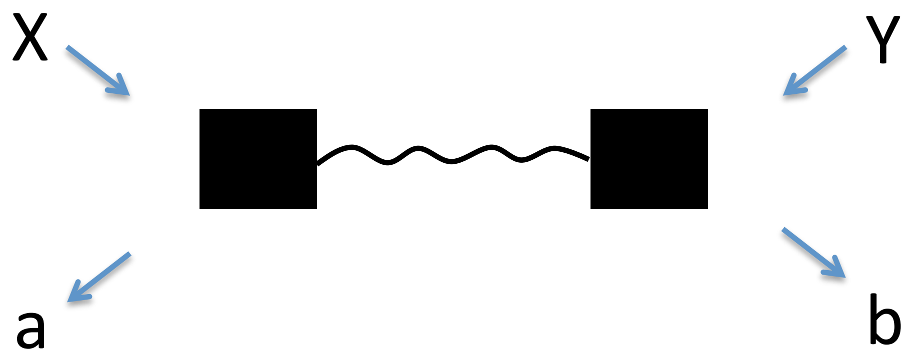
$$E_a^X, E_b^Y \geq 0, \ \sum_a E_a^X = I, \ \sum_b E_b^Y = I.$$

Let $\mathcal{Q}$ be the set of $\mathbf{V}$ describing quantum correlations. It is known that $\mathcal{Q}$ is convex but not a polytope.



From what we've seen already, $\mathcal{L} \subset \mathcal{Q} \subset \mathcal{P}$

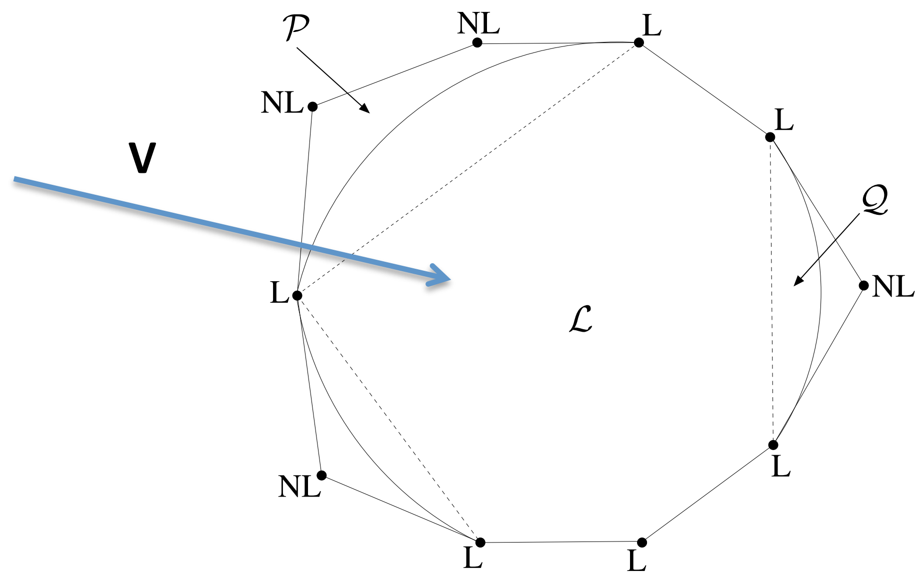
Characterizing nonlocal correlations



$$
\begin{aligned}
X &\in \{0,\dots,k-1\} \\
Y &\in \{0,\dots,k-1\} \\
a &\in \{0,\dots,d-1\} \\
b &\in \{0,\dots,d-1\}
\end{aligned}
$$

$$
\mathbf{V} = \begin{pmatrix}
P(ab=00|XY=00) \\
P(ab=01|XY=00) \\
\vdots \\
P(ab=(d-1)(d-1)|XY=00) \\
\vdots \\
P(ab=(d-1)(d-1)|XY=(k-1)(k-1))
\end{pmatrix}
$$

Given **V**, how do we tell if these correlations are local or nonlocal?
How do we tell if they can be produced by measuring quantum states?



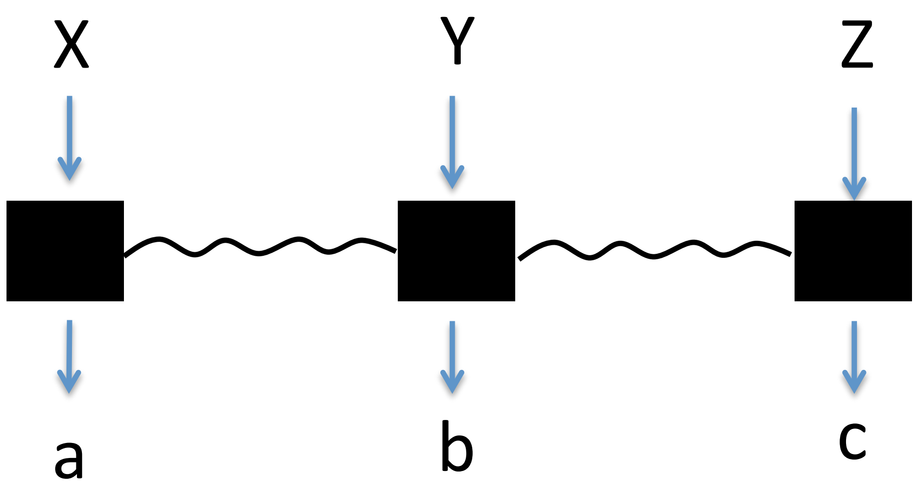
Need to determine whether $\mathbf{V}$ is $\in \mathcal{L}$ (local) or $\in \mathcal{Q}$ (quantum).

Determining membership of a vector in a polytope with specified vertices is a linear programming problem. This is "easy", except NB the number of vertices increase exponentially with the number of settings available to Alice and Bob.

Determining membership in $\mathcal{Q}$ (the Tsirelson probem) seems harder and is not solved in general.

Generalizing the Bell scenario

The first and most obvious way in which the scenario I've described can be generalized is to allow more parties.

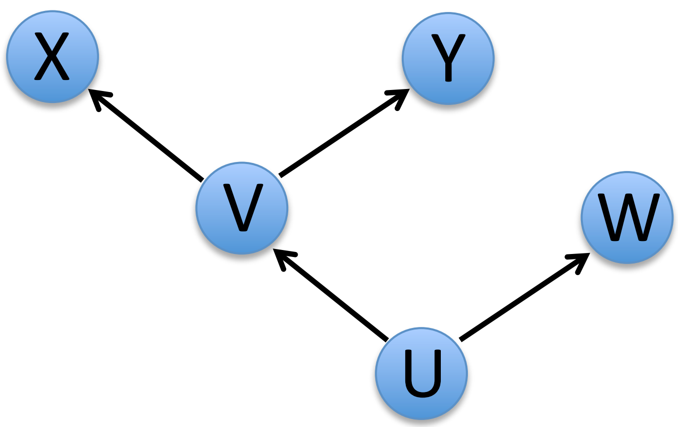


P(abc|XYZ)

Won't say much about this, but the basic formalism is similar: local polytope, non-signalling polytope, quantum set etc.

Generalizing the Bell scenario

Causal networks…



Directed acyclic graph, with random variables on nodes.

Arrows indicate causal relationships.

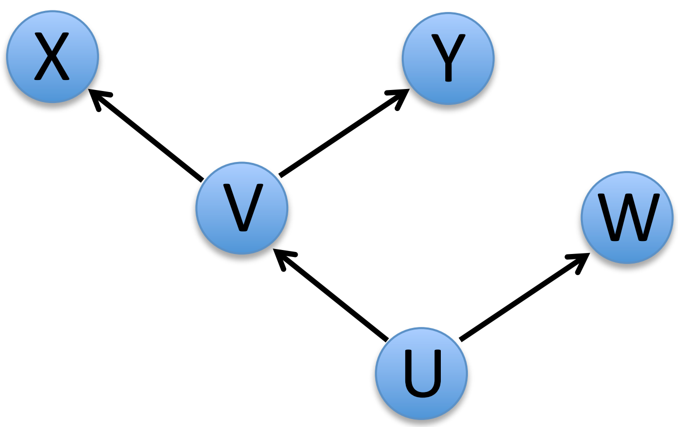
Here, for example:
-- V is not a cause of W. But they can be correlated via a common cause U.
-- X can depend on U, but only indirectly through V.

Here, a joint distribution is compatible with the graph if:

$$P(UVWXY) = P(X|YVWU)\,P(Y|VWU)\,P(V|WU)\,P(W|U)\,P(U)$$
$$= P(X|V)\,P(Y|V)\,P(V|U)\,P(W|U)\,P(U)$$

Generalizing the Bell scenario

Causal networks…



Directed acyclic graph, with random variables on nodes.

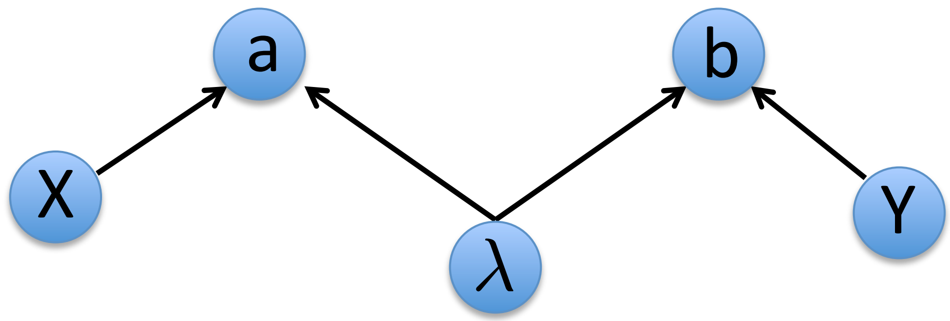
Arrows indicate causal relationships.

Here, for example:
-- V is not a cause of W. But they can be correlated via a common cause U.
-- X can depend on U, but only indirectly through V.

In general, given a graph with nodes $X_1, \dots, X_n$, a joint distribution $P(X_1, \dots, X_n)$ is compatible if $P(X_i \mid pa(X_i), nd(X_i)) = P(X_i \mid pa(X_i))$,

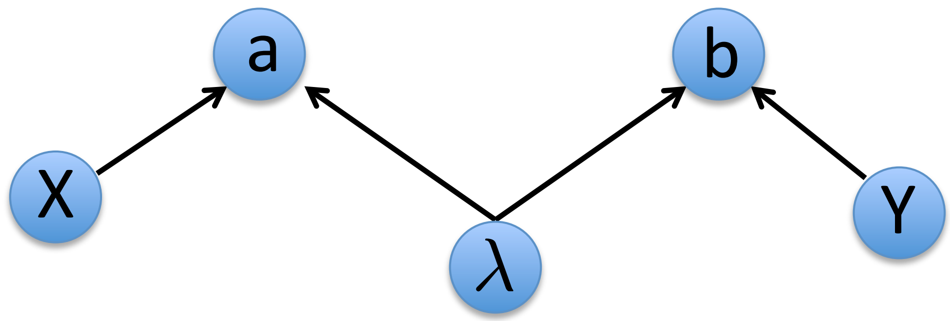
where $pa(X_i)$ are the parents of X_i and $nd(X_i)$ are the non-descendants of X_i.

Generalizing the Bell scenario



- The Bell scenario corresponds to the graph above. Measurement settings are now treated as random variables.
- We didn't talk about this before, but when establishing nonlocality, it is crucial that measurement settings are independent of the putative hidden variable λ . This is encoded by the above graph.
- The locality condition is also encoded – e.g., given X and λ, a is independent of b and Y.

Generalizing the Bell scenario



Which joint distributions P(abXY) can arise as the marginal of a distribution P(abXYλ), which is compatible with the above graph?

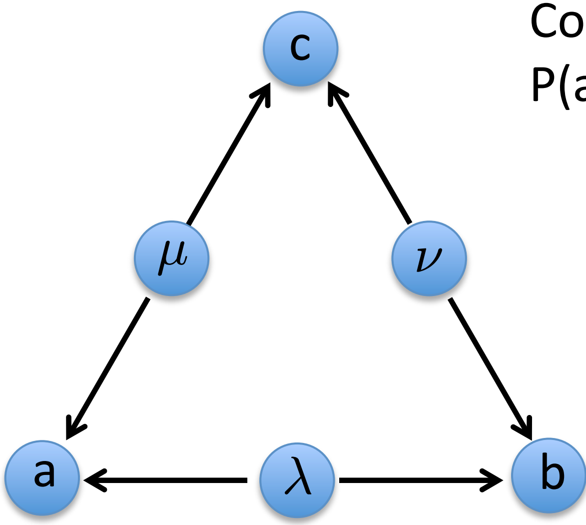
We have answered this already. It is those such that

$$P(abXY) = P(X)\ P(Y)\ P(ab|XY),$$

where P(ab|XY) is a set of local correlations (**V** lies in the local polytope).

The triangle

- A different graph



Compatible joint distribution:

$P(abc\mu\nu\lambda) = P(\lambda)\, P(\mu)\, P(\nu)\, P(a|\mu\lambda)\, P(b|\lambda\nu)\, P(c|\mu\nu)$

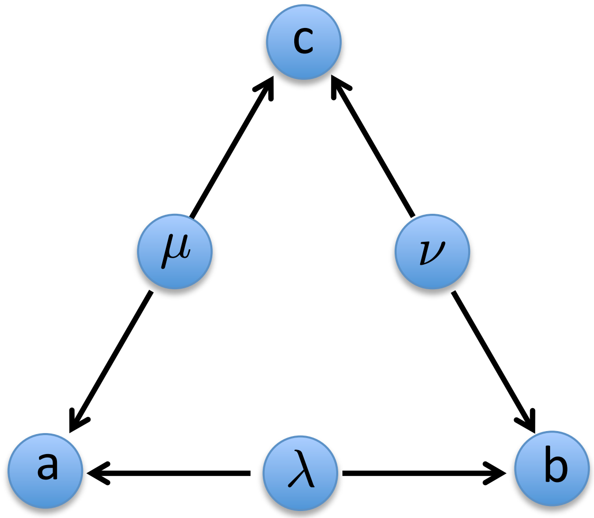
- Three unobserved nodes (hidden variables).
- Three observed nodes (measurement outcomes). No measurement choices!
- A joint distribution P(abc) is *classical* if it can arise as the marginal of a joint distribution $P(abc\lambda\mu\nu)$ that is compatible with the graph.

The triangle

Consider P(abc=000) = P(abc=111) = 1/2.

Is this classical? No!

Rough argument: We have P(a=b)=1. Hence a is independent of μ . But then there is no way for a to be correlated with c.



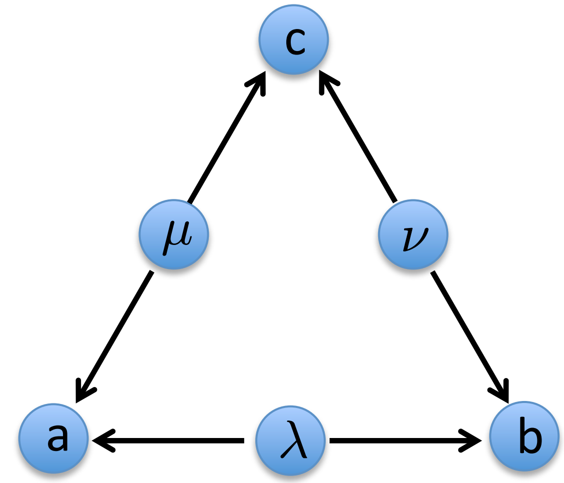
More precisely…

Therorem(*): Any classical distribution P(abc) satisfies

$$I(a{:}b) + I(a{:}c) \leq H(a)$$

(*) T. Fritz, New. J. Phys. **14**1003001 (2012).

The triangle



Therorem: Any classical distribution P(abc) satisfies

$$I(a:b) + I(a:c) \leq H(a).$$

Proof:

$I(a:b) + I(a:c) \leq I(a:\lambda) + I(a:\mu) = 2\,H(a) + H(\lambda) + H(\mu) - H(a\ \mu) - H(a\ \lambda)$

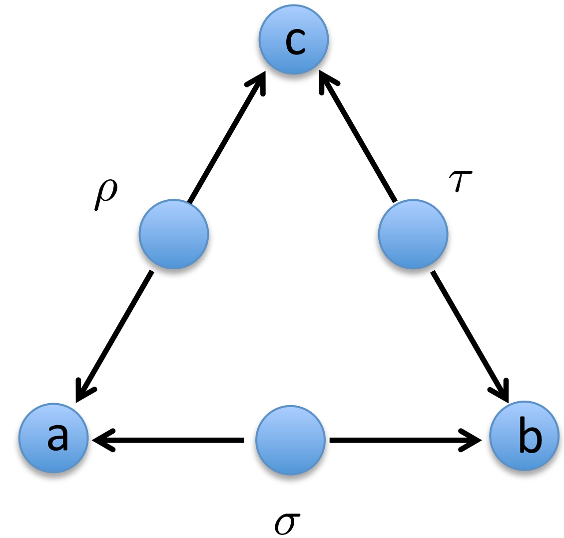
But by submodularity of the Shannon entropy:

$H(a\ \mu) + H(a\ \lambda) \geq H(a) + H(a\ \lambda\ \mu).$

Hence $I(a:b) + I(a:c) \leq H(a) + H(\lambda\ ) + H(\mu) - H(a,\lambda,\mu)\ \leq H(a) + I(\lambda : \mu\ )$

Since λ and μ must be independent, we get $I(a:b) + I(a:c) \leq H(a)$

The triangle



In the quantum case, unobserved hidden variables are replaced with quantum sources.

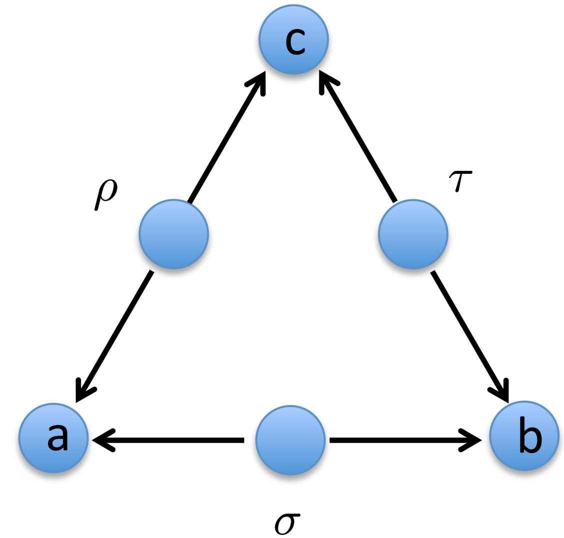
Here, the sources are independent, hence the joint state of the quantum particles must be $\rho \otimes \sigma \otimes \tau$.

Observer Alice can perform a joint measurement on the two quantum particles she receives, before outputting a classical variable *a*. Similarly Bob, Charles.

A distribution P(abc) is *quantum* if it can be written in the form

$$P(abc) = \mathrm{Tr}\left((E_a^{12} \otimes I^{3456})(F_b^{34} \otimes I^{1256})(G_c^{56} \otimes I^{1234})\rho^{16} \otimes \sigma^{23} \otimes \tau^{45}\right).$$

The triangle



Theorem(*): Any quantum distribution P(abc) satisfies

$$I(a:b) + I(a:c) \leq H(a).$$

(*) J. Henson, R. Lal, M. Pusey, arXiv:1405.2572.

Does there exist a quantum distribution on the triangle which is not a classical distribution? Or a distribution obtained from PR boxes which is not quantum?

- If a,b,c are 4-valued then yes (to both)! (For details see Fritz, *ibid*.)
- If a,b,c are binary … no one knows!

Summary

- In Bell scenarios, the problem of classifying local vs nonlocal correlations is solved. Reduces to membership of a vector in a polytope.

- In Bell scenarios, the problem of classifying quantum vs non-quantum correlations is more difficult (but for interesting progress, see M. Navascues, S. Pironio, A. Acin, arXiv: 0803.4290).

- In more general scenarios, classifying the sets of classical and quantum correlations is a difficult problem.

Applications

There are many applications of quantum nonlocality, e.g.,

- Communication complexity
- Device-independent QKD
- Device-independent randomness generation

- All of these are active research areas, but sadly no time in these lectures.
- The study of generalized Bell scenarios should lead to new applications.

Part 2: Contextuality

The Kochen-Specker theorem

Quantum observables correspond to Hermitian operators.

The expectation value of an observable corresponding to G is given by Tr (ρ G).

Suppose that a quantum state is an incomplete description of a quantum system, and that there are some underlying hidden variables.

Previously, we made a similar assumption, and showed that if the hidden variables obey Bell locality, then we cannot recover the predictions of quantum theory.

Now we forget about locality. Suppose that the hidden variables are deterministic, i.e., the hidden variables specify a definite value for each quantum observable, such that measuring the observable simply reveals that value.

Suppose that the outcome does not depend on the *context* of the measurement, i.e., only depends on the Hermitian operator, and not on how the measurement was carried out, nor on what other observables are measured at the same time.

The Kochen-Specker theorem

Hidden variables satisfying the above assumptions are *non-contextual*. Restricting to projection operators, non-contextual hidden variables can be thought of as defining a map:

$$\lambda : P \to \{0, 1\},$$

such that

- if P,Q project onto orthogonal subspaces then λ(P + Q) = λ(P) + λ(Q),
- λ(I) = 1 .

As a colouring problem

Equivalently, non-contextual hidden variables would define a *KS-colouring* of the set of unit vectors in a complex Hilbert space.

Assume finite dimensions. A *KS-colouring* is a map

$$\lambda : v \to \{0, 1\},$$

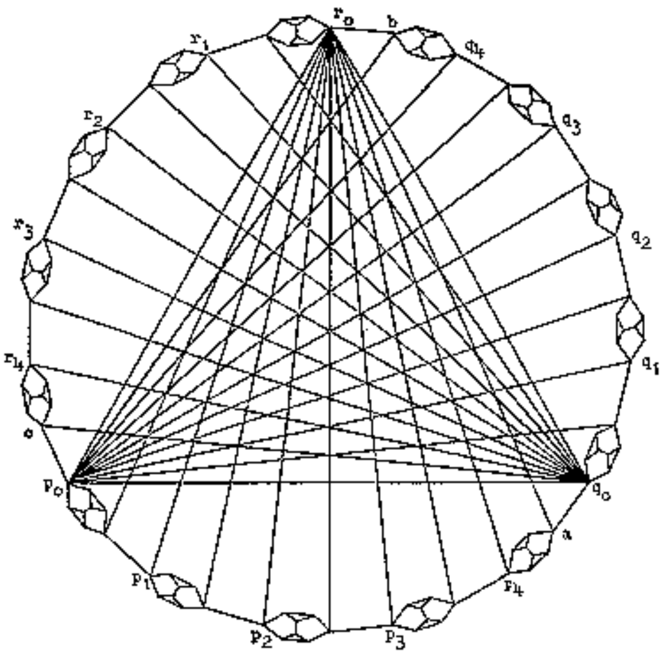
such that

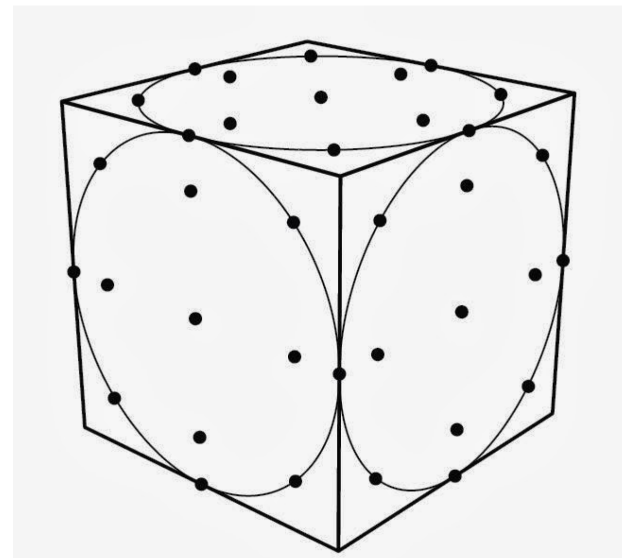
- If $\{ v_1 , \dots , v_d \}$ form an orthonormal basis, then $\sum_\iota \lambda(v_i) = 1$.

Theorem (*) If the dimension of the Hilbert space is ≥ 3, then a KS-colouring of the unit vectors does not exist.

(*) S. Kochen and E. P. Specker, J. Math. Mech. **17**, 59 (1967)

As a colouring problem





The set of 117 vectors originally constructed by Kochen and Specker, and shown to be KS-uncolourable

A smaller set of 33 vectors, constructed by Peres, and shown to be KS-uncolourable.

Image from http://plato.stanford.edu/entries/kochen-specker/

Image from http://quantumgazette.blogspot.co.uk/2014/03/orthogonal-states-and-quantum.html

Finite precision

- So: there is no non-contextual hidden variable interpretation of quantum theory.
- However, a curious loophole in the argument was discovered by David Meyer.

Theorem(*) Let S be the set of unit vectors in R^3 with rational entries. Then S admits a KS-colouring.

NB S is *dense* in the set of unit vectors in R^3, and furthermore the set of orthogonal triads in S is *dense* in the set of orthogonal triads in R^3.

Proof Each unit vector v, with rational entries, can be associated uniquely with a triple of integers (x,y,z) such that v = c (x,y,z), and (x,y,z) have no common divisor. Note that $x^2 + y^2 + z^2$ is a square, hence exactly one of (x,y,z) must be odd. If v and v' are orthogonal, then the corresponding triples must differ in which entry is odd, since xx'+yy'+zz'=0. Therefore, a KS-colouring is obtained by letting $\lambda(v) = 1$ iff the corresponding integer z is odd.

(*) D. Meyer, arXiv:quant-ph/9905080

Finite precision

This was extended by R. Clifton and A. Kent, who give colourings of dense subsets of complex Hilbert spaces in arbitrary finite dimensions (*).

(*) Clifton and Kent, Proc. Roy. Soc. Lond. A, **456** 2101 (2000)

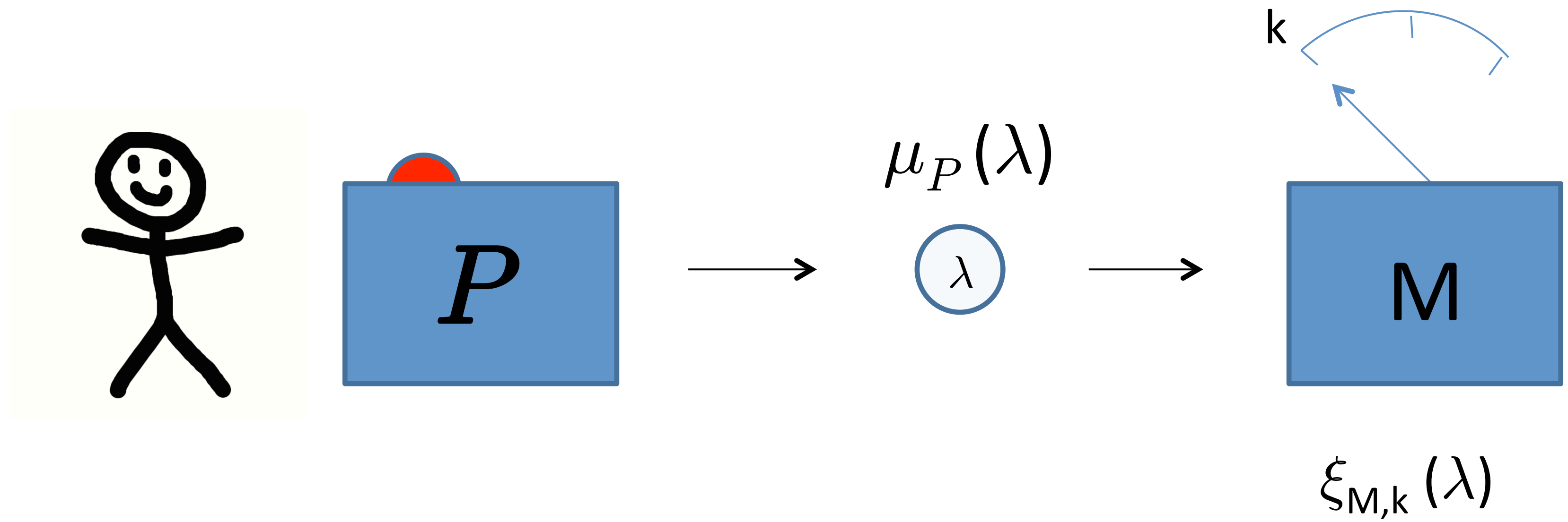
Operational notions of contextuality

The standard Kochen-Specker notion of contextuality is closely tied to the quantum formalism – it concerns colourings of the vectors in complex Hilbert space. It also assumes deterministic hidden variables.

Bell's theorem, by contrast, involves only empirically observable probabilities.

Is there a more operational notion of contextuality?

Operational notions of contextuality



Consider a quantum system of dimension d. An *ontic model* defines:

- A space of ontic states λ
- For each preparation P, a probability distribution $\mu_P(\lambda)$
- For each measurement M and outcome k, a *response function* $\xi_{M,k}(\lambda)$, understood as the probability of getting outcome k when measurement M is performed and the ontic state is λ.

Recover quantum predictions: $\int d\lambda\ \mu_P(\lambda)\ \xi_{M,k}(\lambda) = \mathrm{Tr}(\rho_P E_{M,k})$

Measurement contextuality

We do not assume that ontic states are deterministic.

An ontic state λ is *measurement non-contextual* (*) if $\xi_{M,k}(\lambda)$ only depends on the positive operator $E_{M,k}$ corresponding to the outcome.

A measurement non-contextual ontic state defines a map that takes positive operators $0 \leq E \leq 1$ to probabilities:

$$\lambda : E \to [0, 1]$$

An ontic model is measurement non-contextual if all ontic states are measurement non-contextual.

(*) R. W. Spekkens, Phys. Rev. A 71, 052108 (2005)

Preparation contextuality

An ontic model is *preparation non-contextual* if the distribution $\mu_P(\lambda)$ only depends on the quantum state ρ that is produced by the preparation P.

For example: the qubit state I/2 can be prepared by either of the following methods:

1) Flip a coin, and prepare $|0\rangle$ on heads, $|1\rangle$ on tails.
2) Flip a coin, and prepare $|+\rangle$ on heads, $|-\rangle$ on tails.

Preparation non-contextuality would demand that the resulting distribution $\mu(\lambda)$ is the same for both methods.

Contextuality

An ontic model is *fully non-contextual* iff it is both measurement and preparation non-contextual.

Aside: Why is this a natural constraint?

An argument goes like this. Consider Bell's theorem again. One can always write down a nonlocal hidden variable model which recovers the quantum predictions. However, one is faced with a new problem. Generically, a nonlocal model will lead to *signalling* at the operational level. To prevent this, the model will need to be *finely tuned*.

A similar argument applies to (either) measurement or preparation contextual ontic models. There is a difference in the model, e.g., $\mu_P(\lambda) \neq \mu_P'(\lambda)$, but if P and P' prepare the same quantum state, then there is no corresponding difference at the operational level.

Contextuality

A theory is *contextual* if (some system) does not admit a fully non-contextual ontic model.

Contextuality

Theorem. There exists a measurement non-contextual ontic model for a quantum system of any dimension.

Proof sketch: Simply let the ontic state be identical with a quantum pure state. The response function is defined by the Born rule.

Contextuality

What about preparation contextuality? First, we impose a natural constraint:

Consider a set of preparation procedures P1,…,Pr. Suppose that an experimenter selects one of these at random, perhaps by rolling dice. Let the probability of selecting procedure Pi be q_i .

Viewed as a whole, this defines a new preparation procedure P.

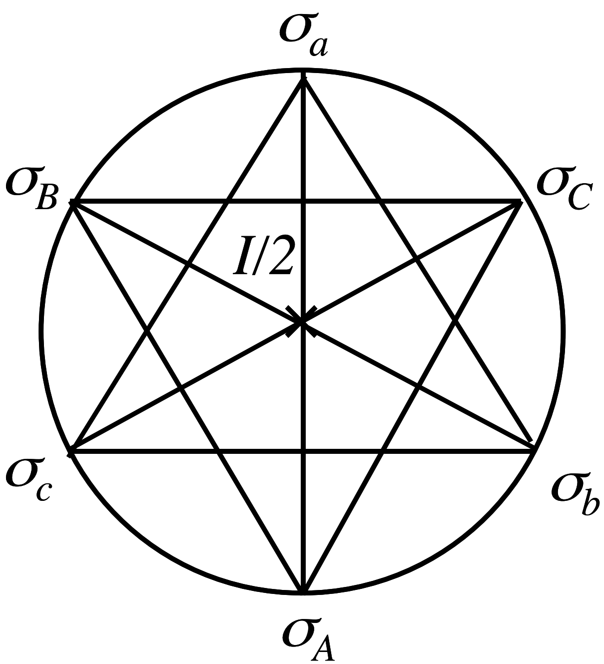
Impose: $\mu_P(\lambda) = \sum_i q_i \, \mu_{Pi}(\lambda)$

Theorem (*). Given the above constraint, there does not exist a preparation non-contextual ontic model for any quantum system of d $\geq$ 2 .

(*) R. W. Spekkens, *ibid.*

Contextuality

Proof(*):



σ_a, σ_A are orthogonal.
Similarly σ_b, σ_B , σ_c , σ_C

Hence $\mu_a(\lambda)\,\mu_A(\lambda) = 0$
$\mu_b(\lambda)\,\mu_B(\lambda) = 0$
$\mu_c(\lambda)\,\mu_C(\lambda) = 0$

$\nu = ½\,(\mu_a + \mu_A)$
$= ½\,(\mu_b + \mu_B)$
$= ½\,(\mu_c + \mu_C)$
$= 1/3\,(\mu_a + \mu_b + \mu_c)$
$= 1/3(\mu_A + \mu_B + \mu_C)$

For each λ , either $\mu_a(\lambda)$ or $\mu_A(\lambda) = 0$.
8 cases to consider…
Conclude that $\nu(\lambda)$ is zero everywhere.

(*) R. W. Spekkens, *ibid.*

Generalized probabilistic theories

Consider a system in a generalized probabilistic theory…

Suppose that the set of allowed measurements is the complete set with respect to the set of states.

Then

Theorem(*): The system is contextual iff it is non-classical.

(*) JB, unpublished

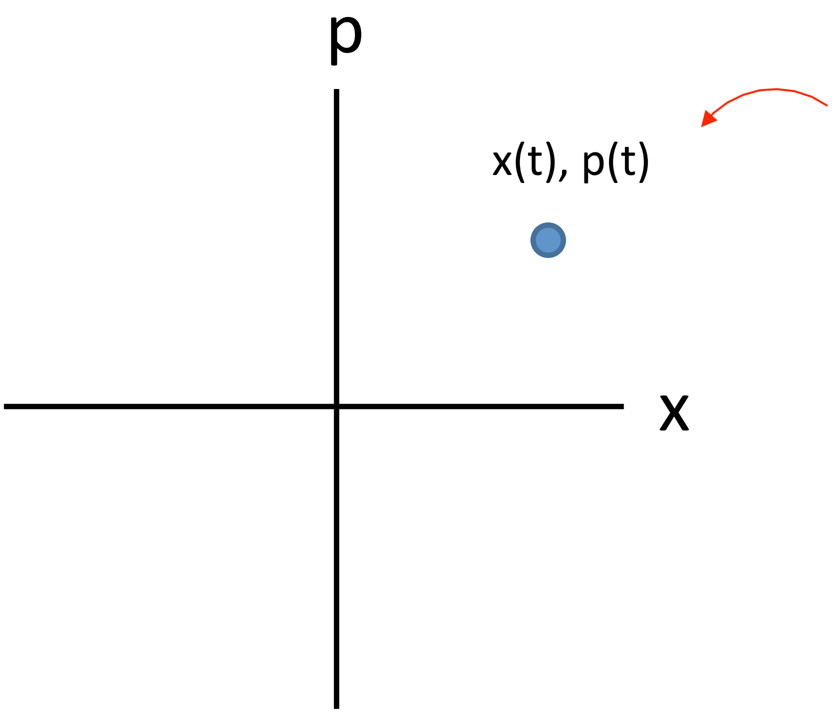
Part 3: Psi-ontology

But our present QM formalism is not purely epistemological; it is a peculiar mixture describing in part realities of Nature, in part incomplete human information about Nature --- all scrambled up by Heisenberg and Bohr into an omelette that nobody has seen how to unscramble. Yet we think that the unscrambling is a prerequisite for any further advance in basic physical theory. **For, if we cannot separate the subjective and objective aspects of the formalism, we cannot know what we are talking about; it is just that simple.**

E. T. Jaynes

Classical Mechanics

- Consider a single particle in 1 dimension.
- Particle has position and momentum. State of particle is completely determined by the values of x,p.
- Other physical properties of the particle are functions of x,p, e.g., energy H(x,p).



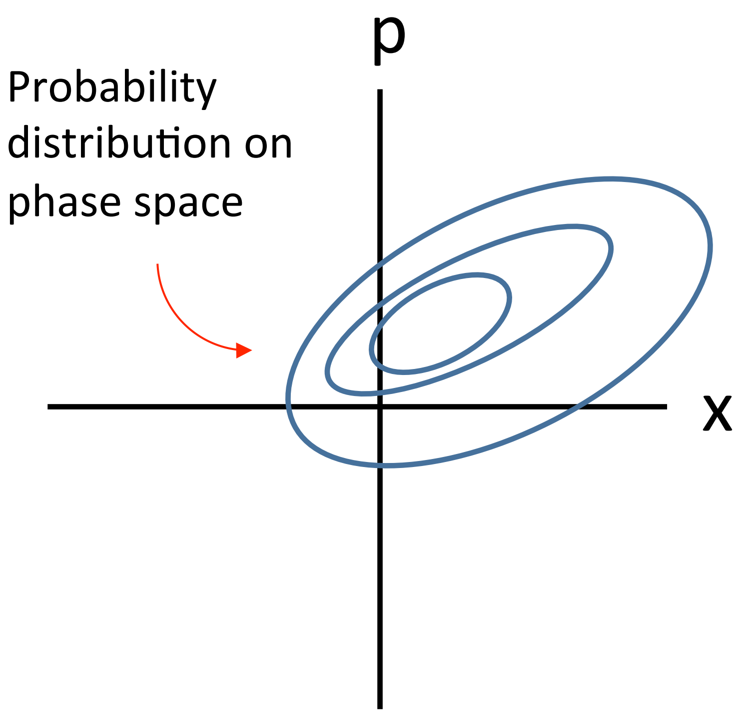
State of system at time t is a point in phase space.

Motion determined by Hamilton's equations

$$\dot{q} = \frac{\partial H}{\partial p}$$
$$\dot{p} = -\frac{\partial H}{\partial q}$$

Liouville Mechanics

- Sometimes we don't know the exact microstate of a classical system.
- The information we have defines a probability distribution ρ over phase space.
- ρ is not a physical property of the particle. The particle occupies a definite point in phase space and does not care what probabilities I have assigned to different states.

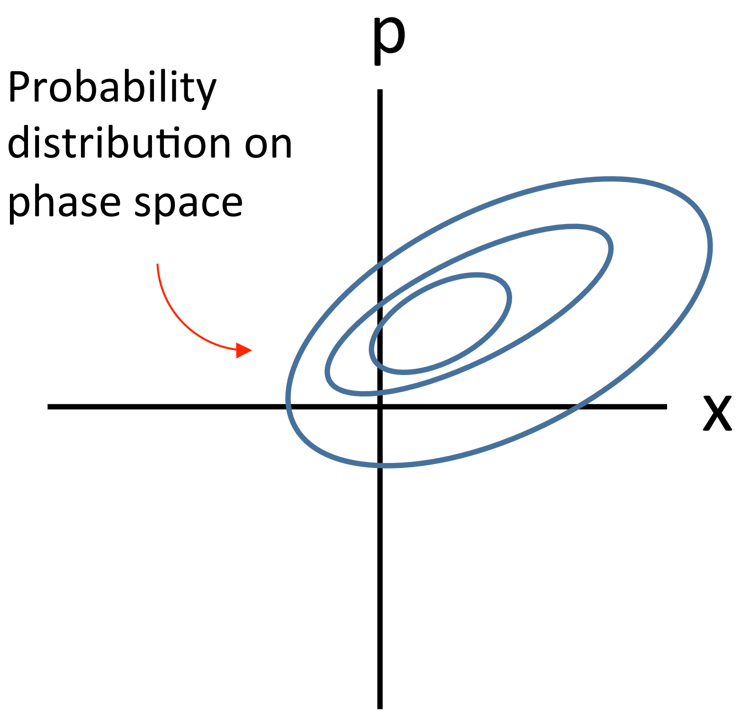


Evolution of the probability distribution is given by the Liouville equation:

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \sum_{i=1}^{n} \left(\frac{\partial \rho}{\partial q^i} \dot{q}^i + \frac{\partial \rho}{\partial p_i} \dot{p}_i \right) = 0.$$

Liouville Mechanics

- Sometimes we don't know the exact microstate of a classical system.
- The information we have defines a probability distribution ρ over phase space.
- ρ is not a physical property of the particle. The particle occupies a definite point in phase space and does not care what probabilities I have assigned to different states.



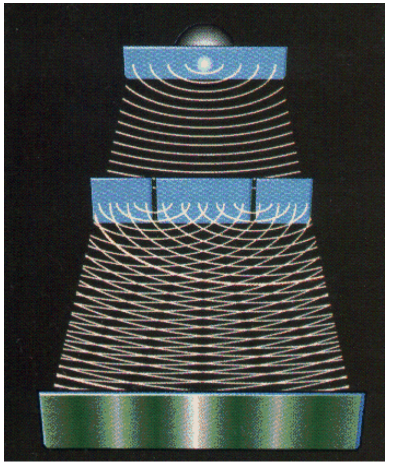
Terminology:

(x,p) *ontic state*
ρ *epistemic state*

What is the quantum state?

Ontic ?

- A quantum wave function is a *real physical wave*.
- Quantum interference most easily understood this way.
- But it is defined on configuration space…



What is the quantum state?

Epistemic ?

- A quantum state encodes an experimenter's *knowledge* or *information* about some aspect of reality.

Arguments for ψ being epistemic

Collapse! ⟶ just Bayesian updating

The wave function is not a thing which lives in the world. It is a tool used by the theory to make those inferences from the known to the unknown. Once one knows more, the wave function changes, since it is only there to reflect within the theory the knowledge one assumes one has about the world.
-----Bill Unruh

Arguments for ψ being epistemic

- Non-orthogonal quantum states cannot reliably be distinguished – just like probability distributions.

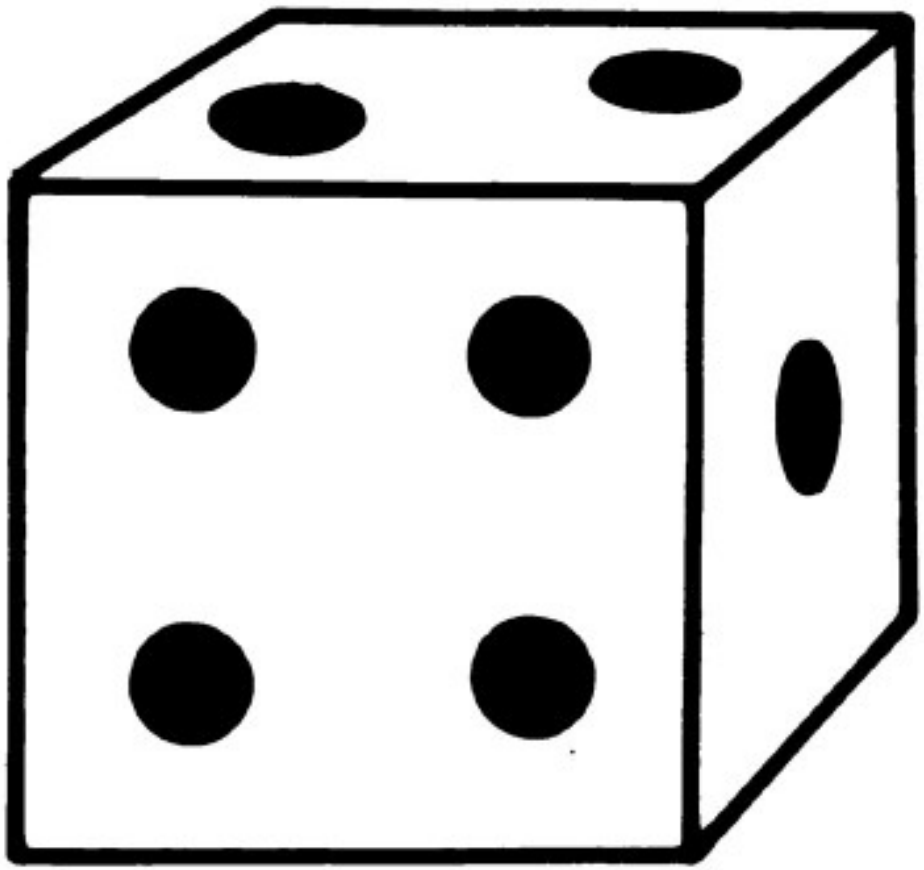
- Quantum states are exponential in the number of systems – just like probability distributions.

- Quantum states cannot be cloned, can be teleported etc – just like probability distributions.

We will show that...

- If ψ merely represents information about the objective physical state of a system, then predictions are obtained that contradict quantum theory.

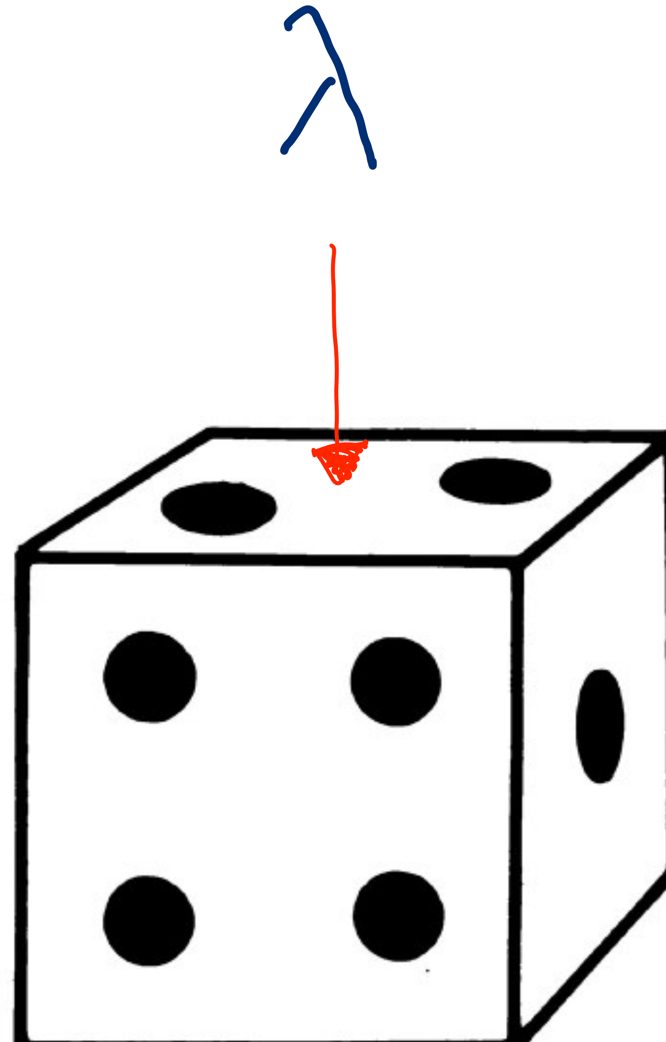
DISCRETE ONTIC STATE SPACE



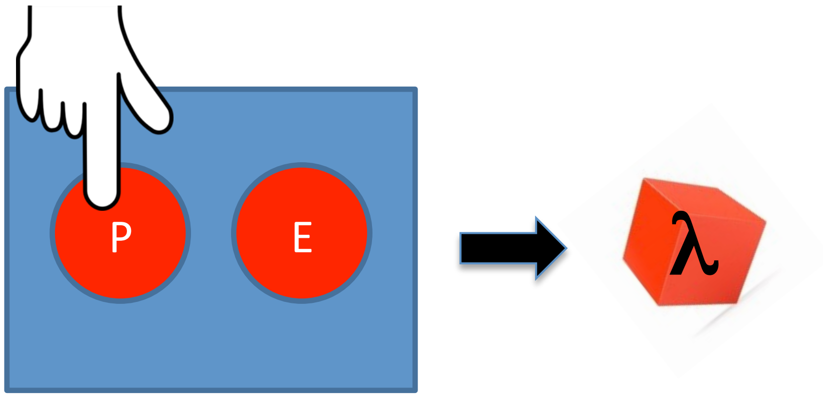
EPISTEMIC STATE

$$\rho = (p_1, p_2, p_3, p_4, p_5, p_6)$$

ONTIC STATE

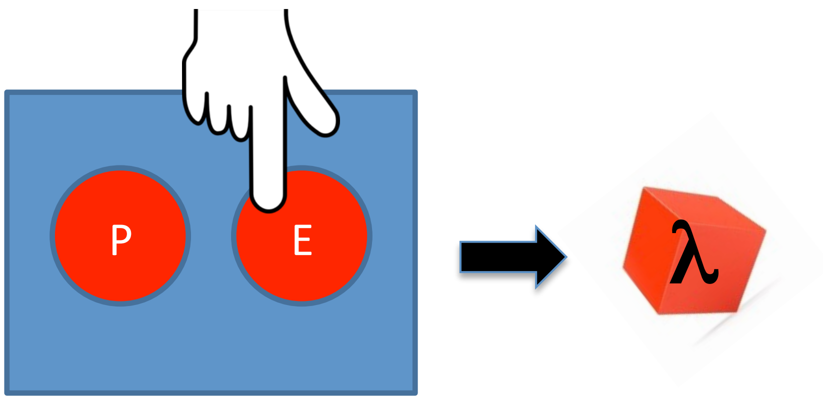


A biased die rolling device



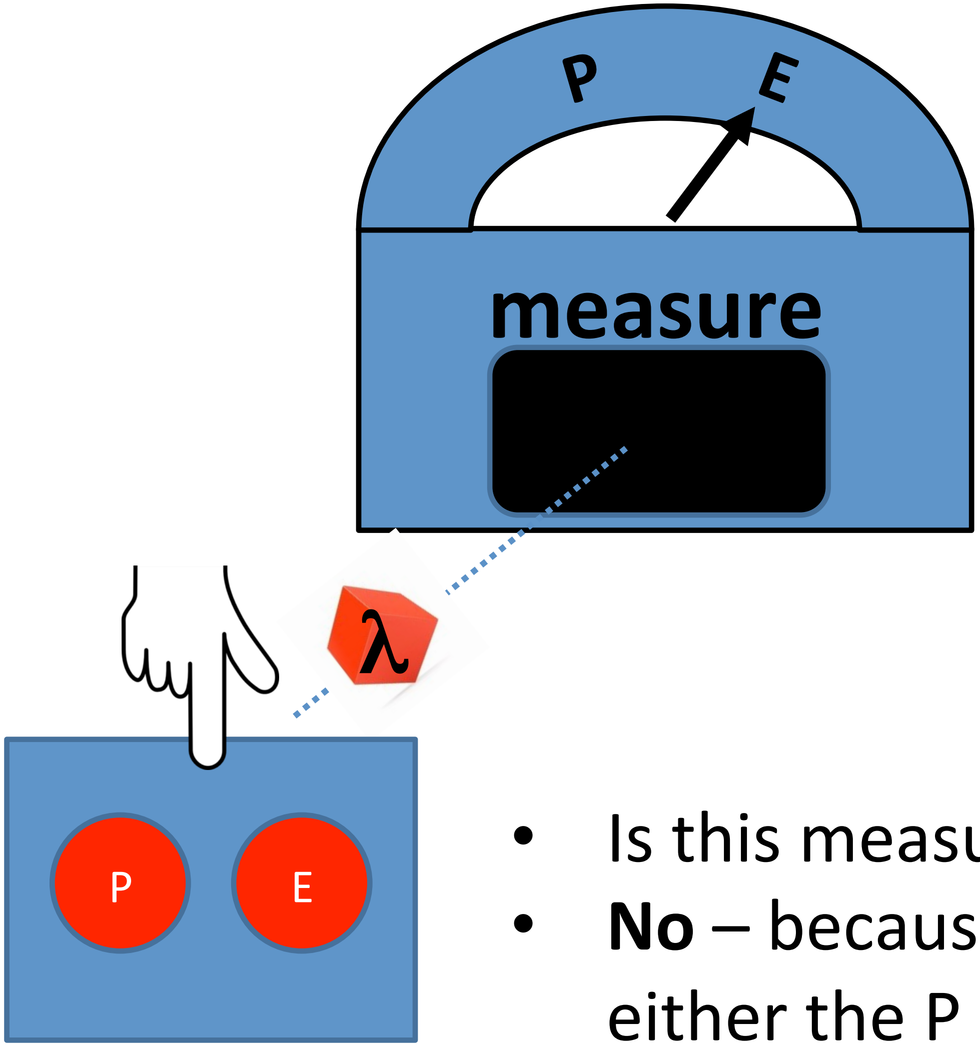
$$\rho_P = \left(0, \frac{1}{3}, \frac{1}{3}, 0, \frac{1}{3}, 0\right)$$

Always get λ = prime

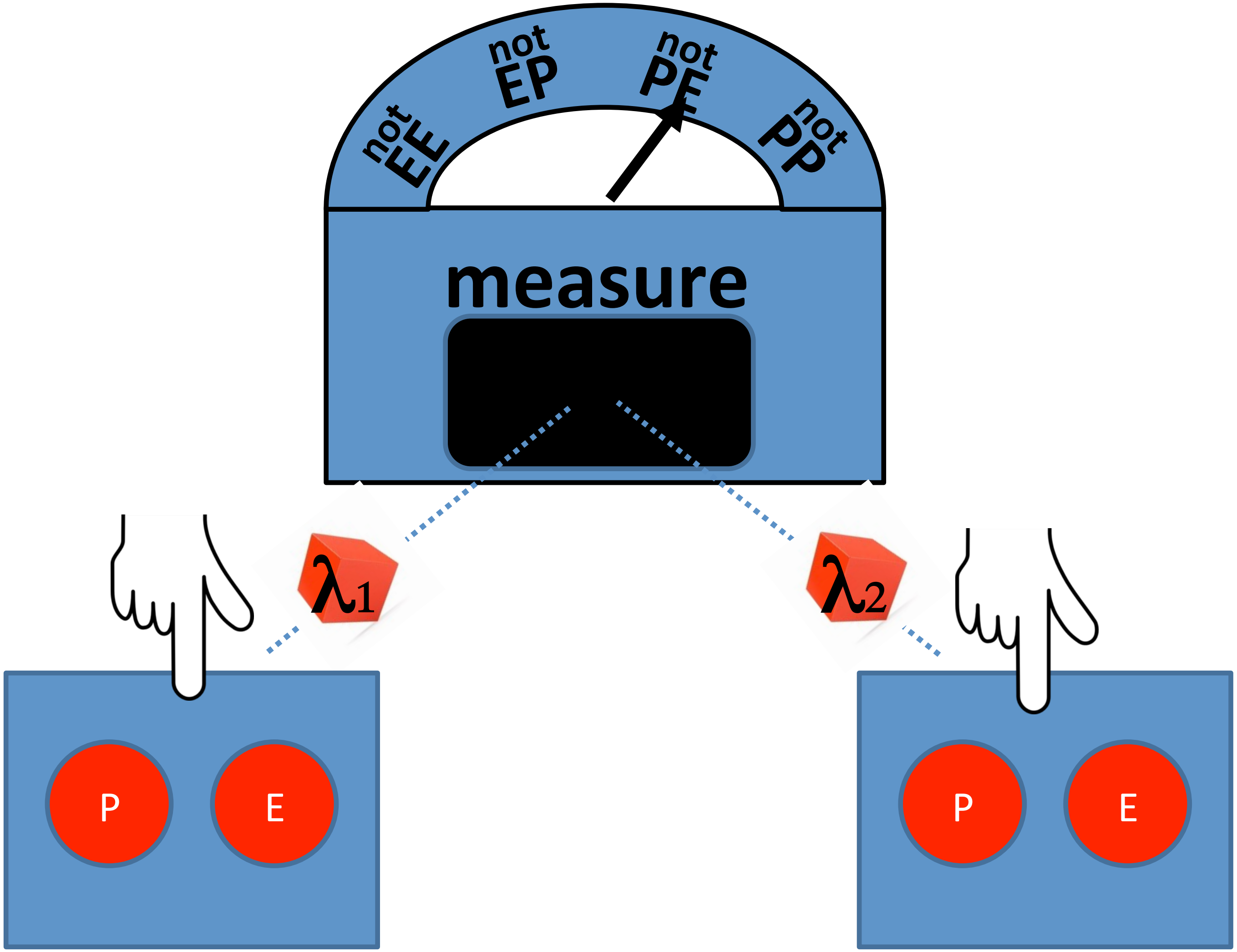


$$\rho_E = \left(0, \frac{1}{3}, 0, \frac{1}{3}, 0, \frac{1}{3}\right)$$

Always get λ = even

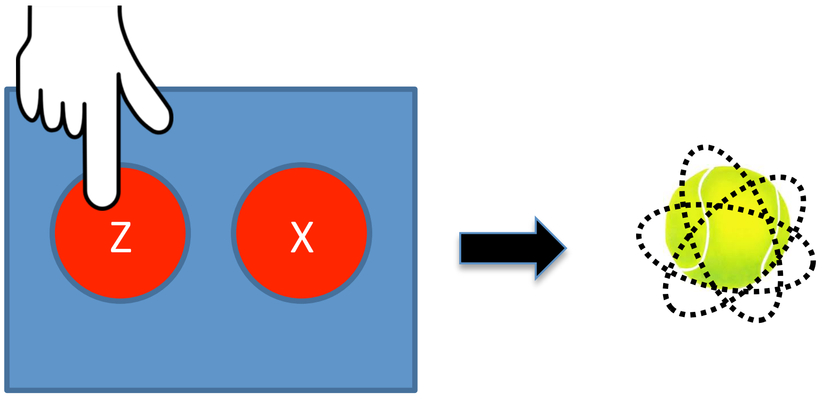


- Is this measurement possible?
- **No** – because λ = 2 can happen with either the P or the E preparation.
- Cannot reliably distinguish overlapping probability distributions.

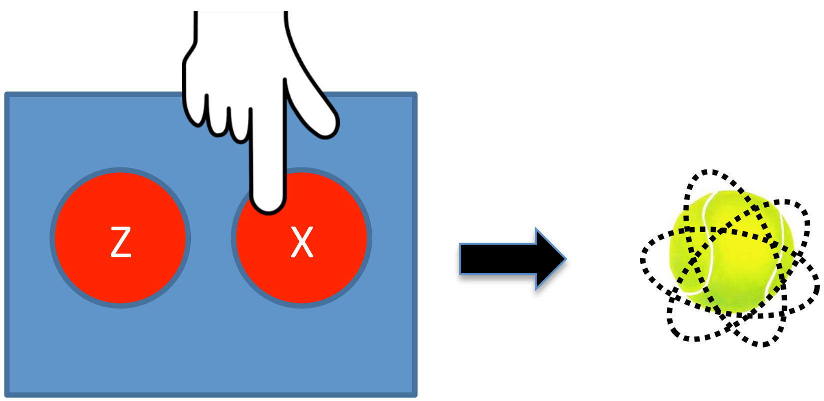


Is this measurement possible? No again.

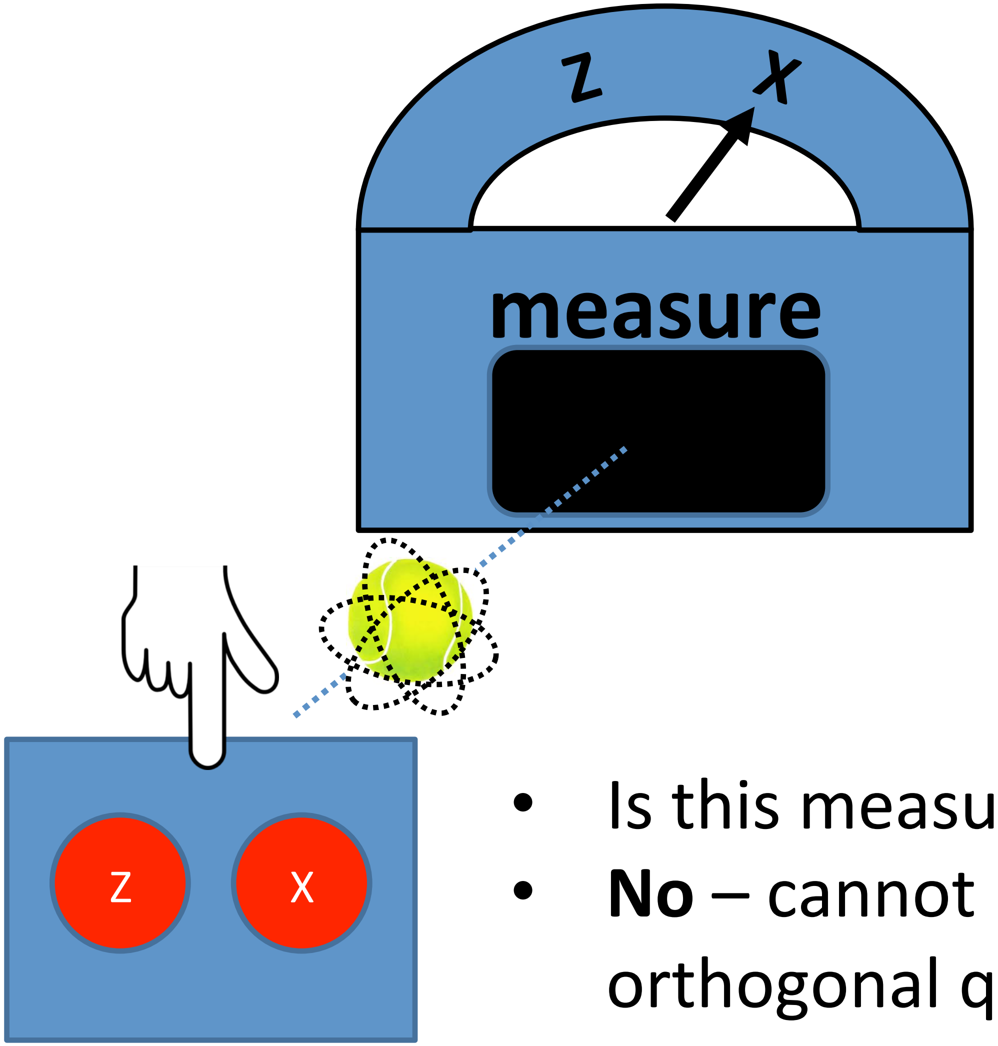
Quantum systems



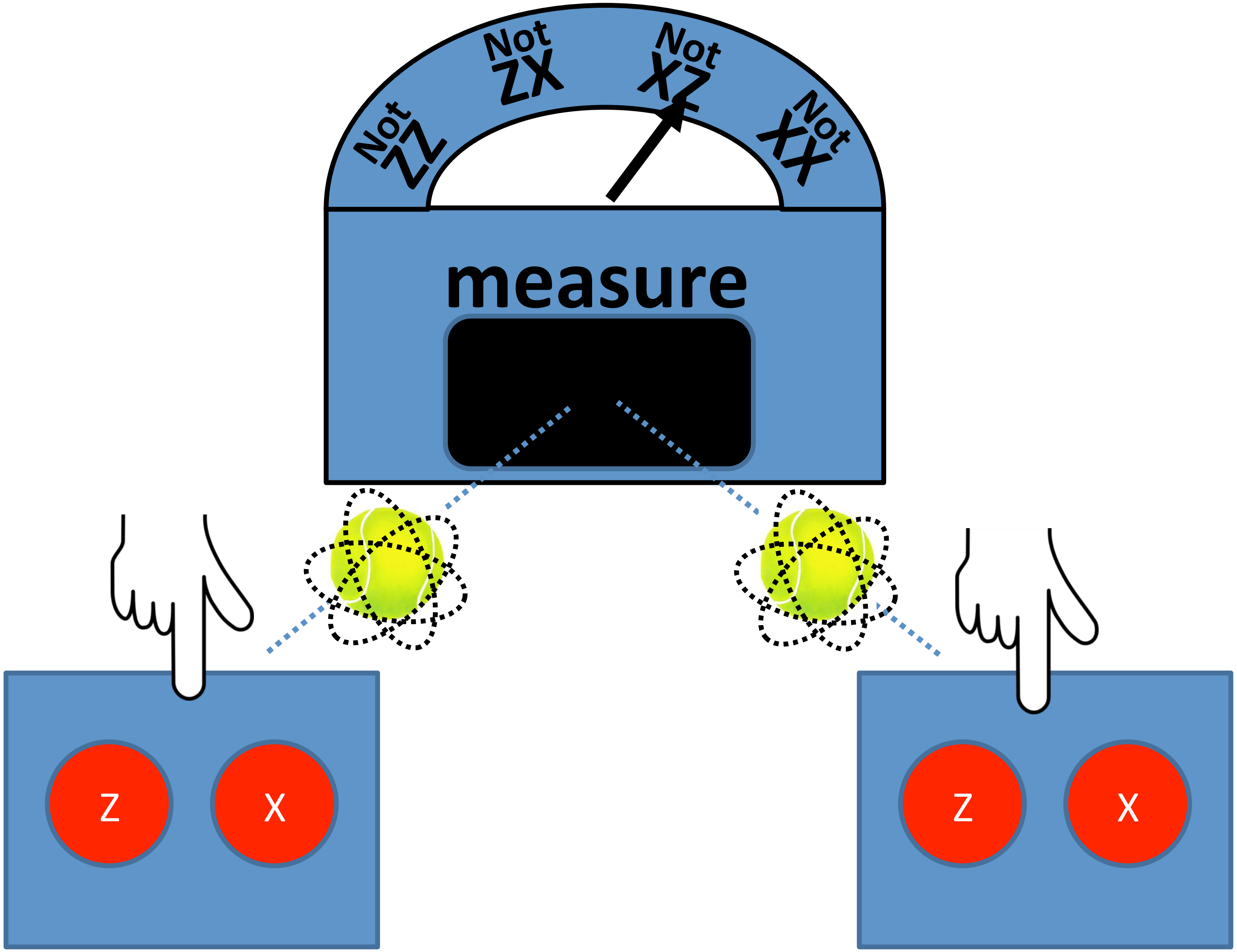
$$\psi = |\uparrow\rangle$$



$$\psi = |\rightarrow\rangle$$



- Is this measurement possible?
- **No** – cannot reliably distinguish non-orthogonal quantum states.
- A very natural *explanation* of this would be that Z,X sometimes prepare the same ontic state.



But this measurement exists !!

Project onto the basis:

$$|1\rangle = 1/\sqrt{2}\,(|\uparrow\rangle|\downarrow\rangle + |\downarrow\rangle|\uparrow\rangle)$$
$$|2\rangle = 1/\sqrt{2}\,(|\uparrow\rangle|\leftarrow\rangle + |\downarrow\rangle|\rightarrow\rangle)$$
$$|3\rangle = 1/\sqrt{2}\,(|\leftarrow\rangle|\uparrow\rangle + |\rightarrow\rangle|\downarrow\rangle)$$
$$|4\rangle = 1/\sqrt{2}\,(|\rightarrow\rangle|\leftarrow\rangle + |\leftarrow\rangle|\rightarrow\rangle)$$

Project onto the basis:

$$|1\rangle = 1/\sqrt{2}\,(|\uparrow\rangle|\downarrow\rangle + |\downarrow\rangle|\uparrow\rangle)$$
$$|2\rangle = 1/\sqrt{2}\,(|\uparrow\rangle|\leftarrow\rangle + |\downarrow\rangle|\rightarrow\rangle)$$
$$|3\rangle = 1/\sqrt{2}\,(|\leftarrow\rangle|\uparrow\rangle + |\rightarrow\rangle|\downarrow\rangle)$$
$$|4\rangle = 1/\sqrt{2}\,(|\rightarrow\rangle|\leftarrow\rangle + |\leftarrow\rangle|\rightarrow\rangle)$$

$$\langle 1|(|\uparrow\rangle|\uparrow\rangle) = 0$$

Project onto the basis:

$$|1\rangle = 1/\sqrt{2}\,(|\uparrow\rangle|\downarrow\rangle + |\downarrow\rangle|\uparrow\rangle)$$
$$|2\rangle = 1/\sqrt{2}\,(|\uparrow\rangle|\leftarrow\rangle + |\downarrow\rangle|\rightarrow\rangle)$$
$$|3\rangle = 1/\sqrt{2}\,(|\leftarrow\rangle|\uparrow\rangle + |\rightarrow\rangle|\downarrow\rangle)$$
$$|4\rangle = 1/\sqrt{2}\,(|\rightarrow\rangle|\leftarrow\rangle + |\leftarrow\rangle|\rightarrow\rangle)$$

$$\langle 2|(|\uparrow\rangle|\rightarrow\rangle) = 0$$

Project onto the basis:

$$|1\rangle = 1/\sqrt{2}\,(|\uparrow\rangle|\downarrow\rangle + |\downarrow\rangle|\uparrow\rangle)$$
$$|2\rangle = 1/\sqrt{2}\,(|\uparrow\rangle|\leftarrow\rangle + |\downarrow\rangle|\rightarrow\rangle)$$
$$|3\rangle = 1/\sqrt{2}\,(|\leftarrow\rangle|\uparrow\rangle + |\rightarrow\rangle|\downarrow\rangle)$$
$$|4\rangle = 1/\sqrt{2}\,(|\rightarrow\rangle|\leftarrow\rangle + |\leftarrow\rangle|\rightarrow\rangle)$$

$$\langle 3|(|\rightarrow\rangle|\uparrow\rangle) = 0$$

Project onto the basis:

$$|1\rangle = 1/\sqrt{2}\,(|\uparrow\rangle|\downarrow\rangle + |\downarrow\rangle|\uparrow\rangle)$$
$$|2\rangle = 1/\sqrt{2}\,(|\uparrow\rangle|\leftarrow\rangle + |\downarrow\rangle|\rightarrow\rangle)$$
$$|3\rangle = 1/\sqrt{2}\,(|\leftarrow\rangle|\uparrow\rangle + |\rightarrow\rangle|\downarrow\rangle)$$
$$|4\rangle = 1/\sqrt{2}\,(|\rightarrow\rangle|\leftarrow\rangle + |\leftarrow\rangle|\rightarrow\rangle)$$

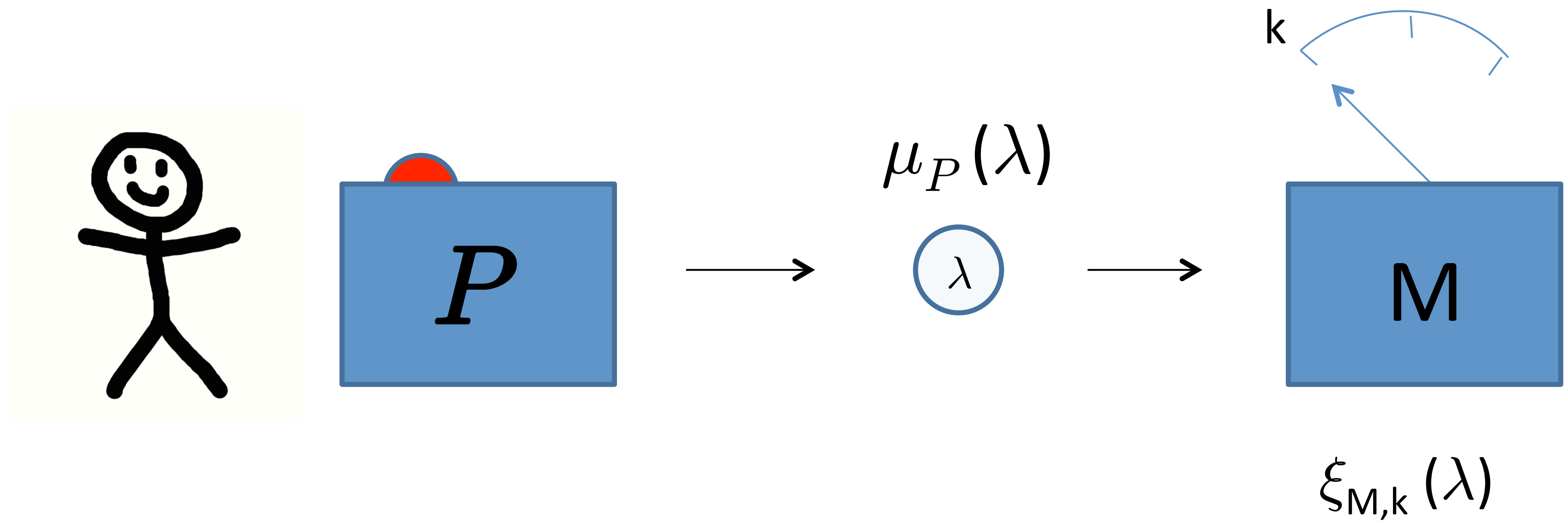
$$\langle 4|(|\rightarrow\rangle|\rightarrow\rangle) = 0$$

What have we learned?

It cannot be the case that preparation of spin-up-Z and spin-up-X can sometimes result in the same underlying ontic state.

Now show that a similar argument goes through for *any* pair of distinct quantum states…

Recall: Ontic models



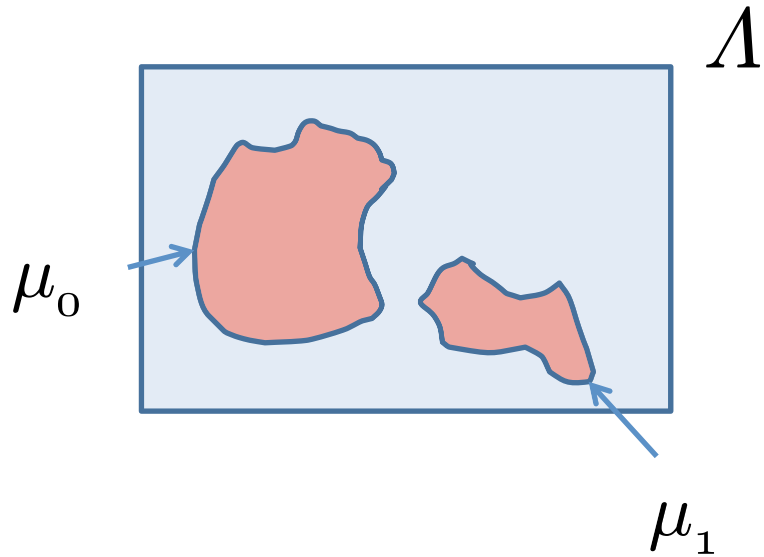
Consider a quantum system of dimension d. An *ontic model* defines:

- A space of ontic states λ
- For each preparation P, a probability distribution $\mu_P(\lambda)$
- For each measurement M and outcome k, a *response function* $\xi_{M,k}(\lambda)$, understood as the probability of getting outcome k when measurement M is performed and the ontic state is λ.

Recover quantum predictions: $$\int d\lambda\ \mu_P(\lambda)\ \xi_{M,k}(\lambda) = \mathrm{Tr}(\rho_P E_{M,k})$$

The ψ-ontic case

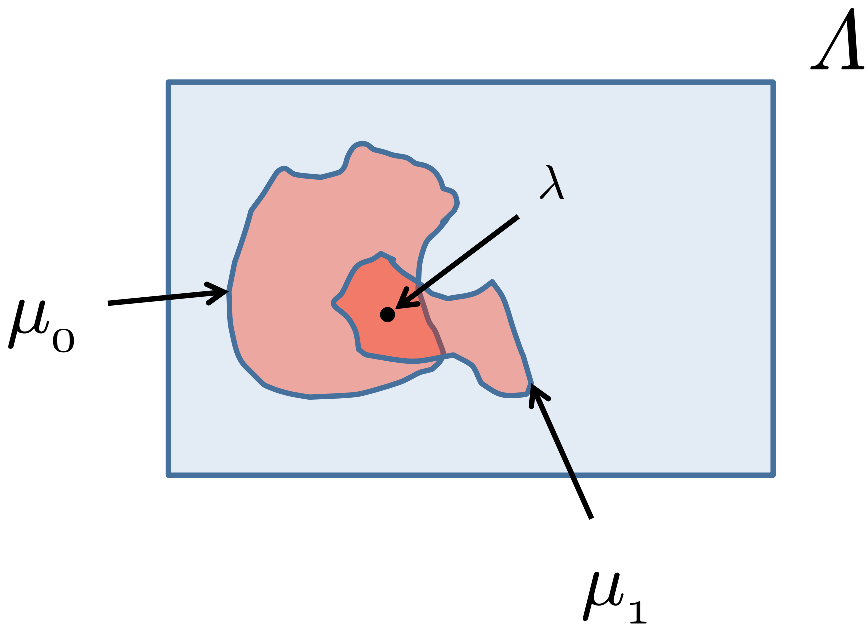
Suppose that for every pair of distinct quantum states ϕ_0 and ϕ_1, the distributions μ_0 and μ_1 do not overlap:



- The quantum state can be inferred from the ontic state.
- The quantum state is a *physical property* of the system, and is not mere information.

See N. Harrigan and R. W. Spekkens, Found. Phys. 40, 125 (2010).

The ψ-epistemic case



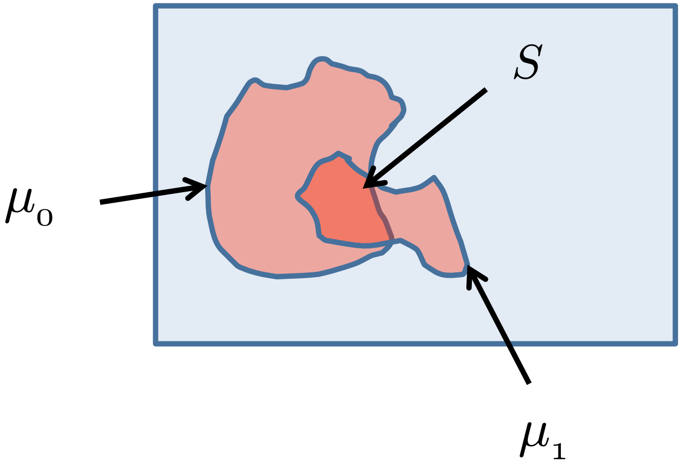
- μ_0 and μ_1 can overlap.
- Given the ontic state λ above, cannot infer whether the quantum state ϕ_0 or ϕ_1 was prepared.

See N. Harrigan and R. W. Spekkens, Found. Phys. 40, 125 (2010).

A no-go theorem

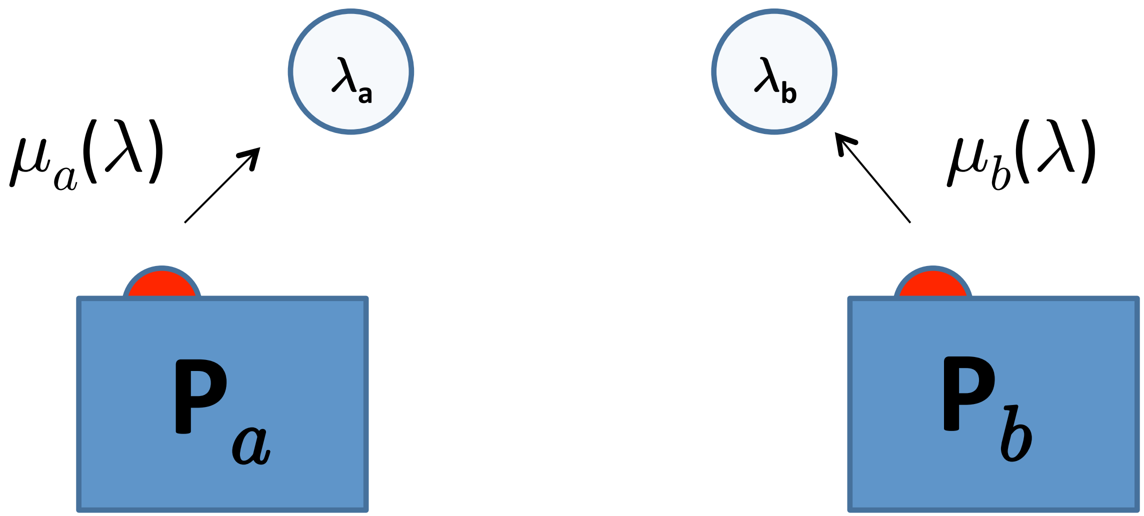
Suppose there are distinct quantum states ϕ_0 and ϕ_1, and a subset S of the ontic states such that

$\mu_0(\lambda) > 0$ and $\mu_1(\lambda) > 0$ for all $\lambda \in$ S
μ_0(S) $\geq$ q > 0
μ_1(S) $\geq$ q > 0 .



Preparation Independence

Consider independent preparations, of quantum states ϕ_a and ϕ_b, producing a joint state $\phi_a \otimes \phi_b$.

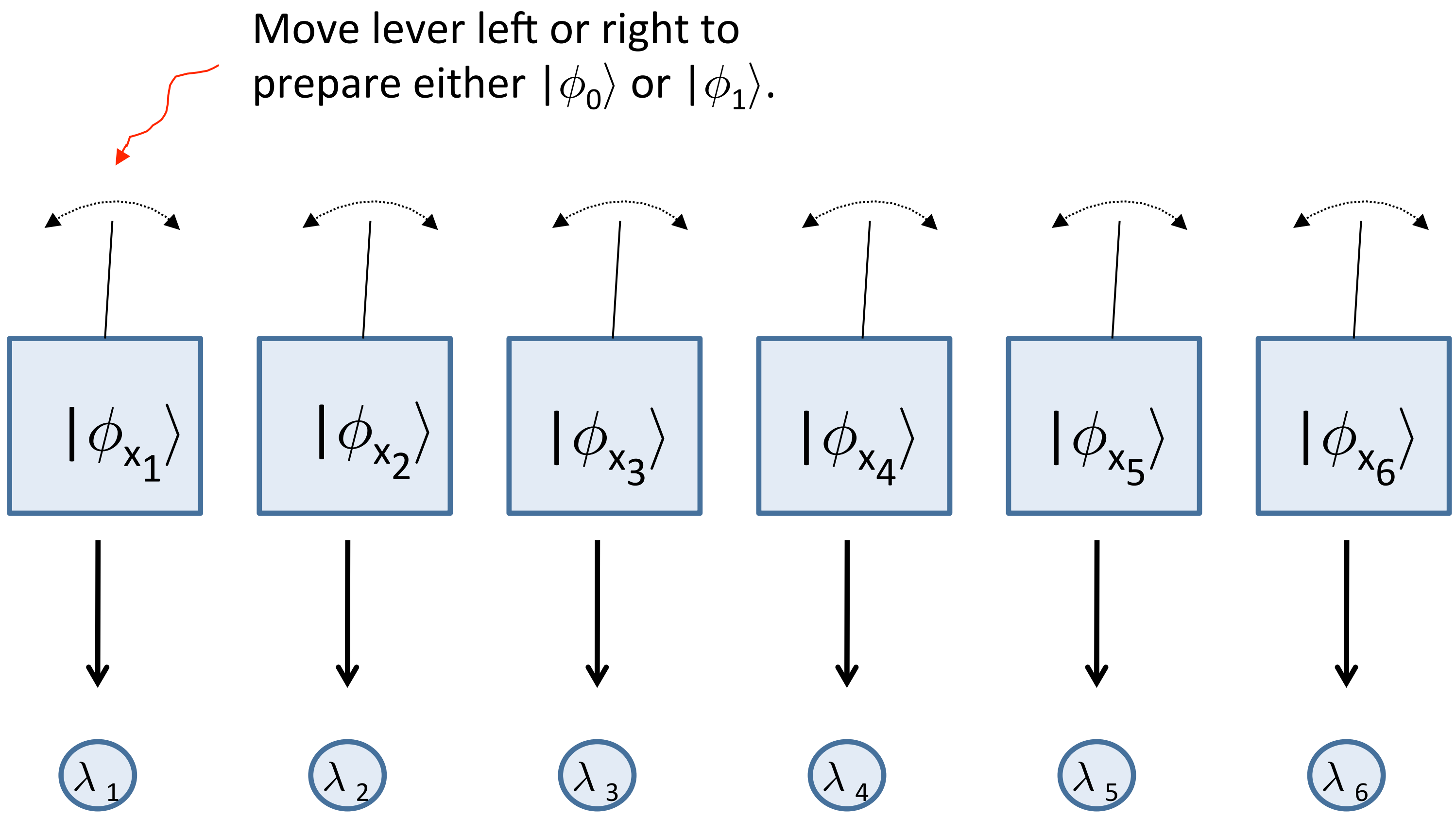


Assume *preparation independence*: the joint distribution over λ_a and λ_b, corresponding to preparation of the product state $\phi_a \otimes \phi_b$ satisfies:

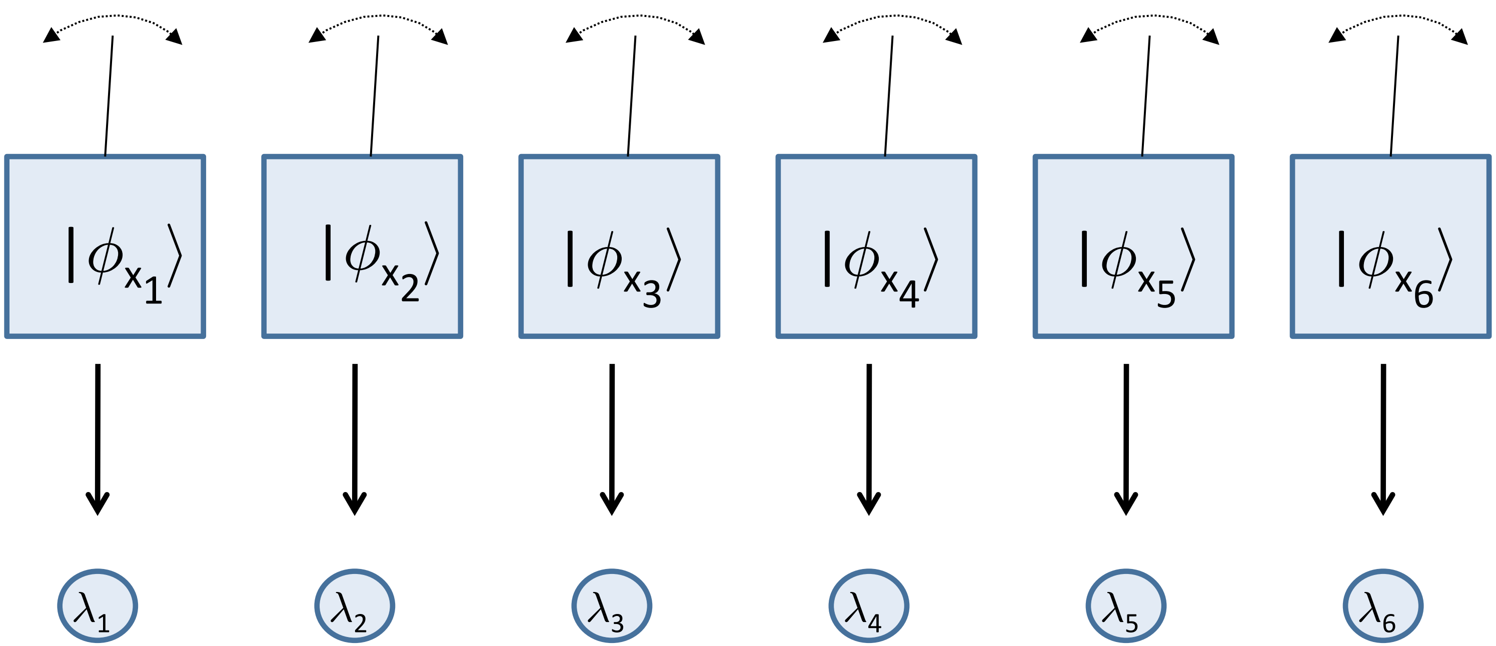
$$\mu(\lambda_a, \lambda_b) = \mu_a(\lambda_a) \times \mu_b(\lambda_b)$$

Prepare n systems independently...

- Each is prepared in either the state $|\phi_0\rangle$ or the state $|\phi_1\rangle$.
- 2^n possible joint states: $|\phi_{x_1}\rangle \otimes |\phi_{x_2}\rangle \otimes \cdots \otimes |\phi_{x_n}\rangle$



For any $|\phi_{x_1}\rangle \otimes |\phi_{x_2}\rangle \otimes \cdots \otimes |\phi_{x_n}\rangle$ there is some chance that for every one of the n systems, the ontic state is in the region S.



$\Pr(\lambda_i \in S \text{ for all } i) \geq q^n$

- **Now here's the problem...**

A `*PP-measurement*`
Cf Caves, Fuchs, Schack, Phys. Rev. A **66**, 062111 (2002).

- For large enough n there is an entangled measurement across the n systems, with 2^n outcomes corresponding to projectors P_1, ... , P_{2^n} and

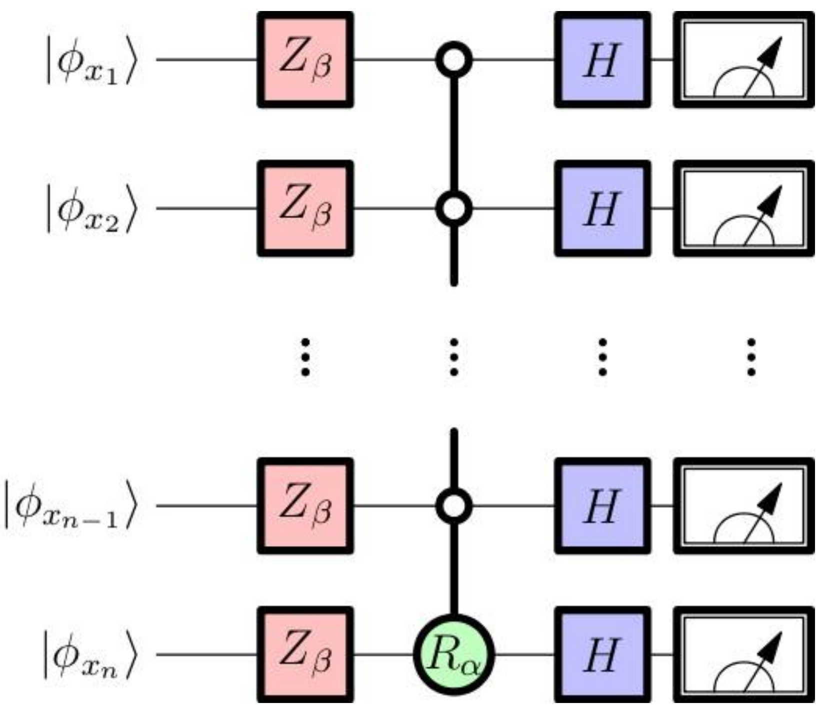
$$\langle\phi_0|\otimes\cdots\otimes\langle\phi_0|\otimes\langle\phi_0| \quad P_1 \quad |\phi_0\rangle\otimes\cdots\otimes|\phi_0\rangle\otimes|\phi_0\rangle = 0$$
$$\langle\phi_0|\otimes\cdots\otimes\langle\phi_0|\otimes\langle\phi_1| \quad P_2 \quad |\phi_0\rangle\otimes\cdots\otimes|\phi_0\rangle\otimes|\phi_1\rangle = 0$$
$$\vdots$$
$$\langle\phi_1|\otimes\cdots\otimes\langle\phi_1|\otimes\langle\phi_1| \quad P_{2^n} \quad |\phi_1\rangle\otimes\cdots\otimes|\phi_1\rangle\otimes|\phi_1\rangle = 0$$

- For any of these 2^n joint preparations there is a non-zero probability that the ontic state $(\lambda_1, \ldots, \lambda_n) \in S \times \cdots \times S$.
- In this case, must have $\xi_{M,i}(\lambda_1 \times \cdots \times \lambda_n) = 0$ for any i. But probs must sum to 1!

The measurement

Choose n such that $2^{1/n} - 1 \leq \tan(\theta/2)$.

Wlog, write $|\phi_0\rangle = \cos(\theta/2)\ |0\rangle - \sin(\theta/2)\ |1\rangle$
$|\phi_1\rangle = \cos(\theta/2)\ |0\rangle + \sin(\theta/2)\ |1\rangle$



$$Z_\beta = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\beta} \end{pmatrix}$$

$$\begin{aligned} R_\alpha|00\cdots0\rangle &= e^{i\alpha}|00\cdots0\rangle \\ R_\alpha|b\rangle &= |b\rangle, \end{aligned}$$

on all other basis states $|b\rangle$.

Approximate case

Suppose that in a real experiment, the measured probabilities are within ϵ of the quantum predictions. Then

$$\delta(\mu_0, \mu_1) \geq 1 - 2\sqrt[n]{\epsilon} \qquad (*)$$

Classical trace distance

(*) M. Pusey, JB, T. Rudolph, Nature Physics **8**, 475 (2012).

Approximate case

Suppose that in a real experiment, the measured probabilities are within ϵ of the quantum predictions. Then

$$\delta(\mu_0, \mu_1) \geq 1 - 2\sqrt[n]{\epsilon} \qquad (*)$$

Classical trace distance

So with a well-performed experiment, we can put an upper bound on how much the classical distributions overlap.

(*) M. Pusey, JB, T. Rudolph, Nature Physics **8**, 475 (2012).

Psi-epistemic models for single systems exist

- An explicit construction is given in P. Lewis, D. Jennings, JB, T. Rudolph, PRL **109**, 150404 (2012).

- But μ_ϕ and μ_ψ only overlap for *some* pairs of quantum states.

- Aaronson et al. arXiv:1303.2834 go further. Provide an explicit construction such that μ_ϕ overlaps with μ_ψ for *any* pair of non-orthogonal state vectors $|\phi\rangle$ and $|\psi\rangle$.

How much can μ_ϕ and μ_ψ overlap?

Distinguishing probability distributions

Consider two preparations of a classical system, corresponding to distributions $\mu_1(\lambda)$ and $\mu_2(\lambda)$.

A priori probability for each preparation is ½.

With a single-shot measurement on the system, must guess which preparation method was used.

Prob(guess correctly) = ½ (1 + D(μ_1, μ_2)),

where D(μ_1,μ_2) is the classical trace distance between μ_1 and μ_2,
$D(\mu_1,\mu_2) = 1/2 \int d\lambda\, |\mu_1(\lambda) - \mu_2(\lambda)|$.

Distinguishing quantum states

Consider two preparations of a quantum system, corresponding to state vectors:
$|\phi\rangle, |\psi\rangle$

A priori probability for each preparation is ½.

With a single-shot measurement on the system, must guess which preparation method was used. With an optimal measurement:

Prob(guess correctly) = ½ (1 + $D_Q(|\phi\rangle, |\psi\rangle)$),

where $D_Q(|\phi\rangle, |\psi\rangle)$ is the quantum trace distance between $|\phi\rangle$ and $|\psi\rangle$,

$$D_Q(|\phi\rangle, |\psi\rangle) = \sqrt{1 - |\langle\phi|\psi\rangle|^2}.$$

Maximally psi-epistemic models

Theorem(*)

In any ontic model that reproduces the predictions for a *d* dimensional quantum system:

$$D(\mu_\phi, \mu_\psi) \geq D_Q(|\phi\rangle, |\psi\rangle) \qquad \forall |\phi\rangle, |\psi\rangle$$

Proof sketch

Consider the optimal quantum measurement for guessing the preparation. Guess correctly with P = ½ (1 + $D_Q(|\phi\rangle, |\psi\rangle)$). But measurement device only has access to λ, and must distinguish μ_ϕ from μ_ψ. Cannot achieve this success rate if μ_ϕ , μ_ψ overlap too much!

Maximally epistemic models

An ontic model is *maximally psi-epistemic* if

$$D(\mu_\phi, \mu_\psi) = D_Q(|\phi\rangle, |\psi\rangle) \qquad \forall |\phi\rangle, |\psi\rangle$$

Why is this natural?

- Recall that overlap of μ_ϕ , μ_ψ would explain why $|\phi\rangle$, $|\psi\rangle$ cannot be distinguished with certainty.
- In a maximally psi-epistemic model, failure to distinguish $|\phi\rangle$, $|\psi\rangle$ is *entirely due* to the ordinary classical difficulty in distinguishing the probability distributions μ_ϕ , μ_ψ . No other limitations or uniquely quantum effects need be invoked.

Overlap bounds

Theorem(*)

Define the *classical overlap* $\omega\,(\mu_\phi\,, \mu_\psi) \equiv 1 - D(\mu_\phi\,, \mu_\psi)$.
Similarly the *quantum overlap* $\omega_Q\,(|\phi\rangle\,, |\psi\rangle) \equiv 1 - D_Q\,(|\phi\rangle\,, |\psi\rangle)$.

Consider an ontic model that reproduces quantum predictions in dimension *d* and satisfies:

$$\omega(\mu_\phi, \mu_\psi) \geq \alpha\ \omega_Q(|\phi\rangle, |\psi\rangle)) \qquad \forall |\phi\rangle, |\psi\rangle$$

Then α < 4/(d-1). If d is power prime then α < 2/d.

Hence no maximally psi-epistemic model can recover the quantum predictions in d $\geq$ 3.

(*) JB, E. Cavalcanti, R. Lal, O.Maroney, PRL to appear

Additionally

- There is also a noise-tolerant version. Maximally psi-epistemic models cannot *approximately* recover quantum predictions in d $\geq$ 3.

- An explicit maximally psi-epistemic model exists in d=2 (constructed by Kochen and Specker).

- An improved bound, exponentially small in d, is obtained in M.S.Leifer, Phys. Rev. Lett. **112**, 160404 (2014).

Summary

- Given preparation independence, psi-epistemic models cannot reproduce the predictions of quantum theory.

- Without preparation independence, can still derive interesting results by considering single systems. For at least some pairs of quantum states, the overlap of the classical distributions must be small, as d gets large.

- Are there information-theoretic applications of these new theorems?

Conclusions

All the no-go theorems we have considered begin with the idea of a *classical*, or *hidden variable*, or *ontic* model for a quantum system.

Bell → Model cannot be locally causal.

Extensions of Bell → In various other scenarios, model cannot respect underlying causal structure.

Kochen-Specker → A deterministic ontic model cannot be KS-non-contextual ($d \geq 3$).

Spekkens → Model cannot be preparation non-contextual.

PBR → Model cannot be psi-epistemic.

Recent results → Even for single systems, model cannot be maximally psi-epistemic.

Conclusions

One approach is simply to accept the conclusion of each of these theorems.

E.g., de Broglie-Bohm theory is nonlocal, preparation contextual, psi-ontic etc.

But many quantum scientists would take the view that the theorems are evidence that the *assumptions* need to be given up.

The primary assumption that all these theorems share is that there is an underlying theory, whose schematic form is well captured by the notion of an ontic model.

There remains the possibility that there is an underlying theory to be discovered, which is structured quite differently from the way in which ontic models describe the world.

Conclusions

Another way of thinking about the no-go theorems is as constraints on classical simulation of quantum systems.

Bell nonlocality already has many applications, in communication complexity, key distribution, measurement-based computation.

It remains to be seen what applications might emerge from the generalized Bell scenarios, or from preparation contextuality, or the PBR argument.

Conclusions

All these things are active research areas. I have missed out a great deal of topics.