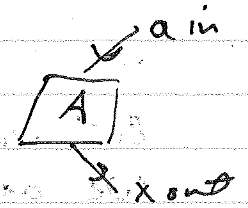
HIDDEN VARIABLES

3 Nov. 2008

What is a hidden variable theory? It is a model of quantum measurement based on the hypothesis that probabilities arise in quantum theory the same way they arise classically: because of ignorance.

Consider a single party (Alice). Imagine that she can set her apparatus to measure any one of m observables, labeled by a. And for each setting, she obtains one of v possible outcomes, labeled by x. But she can access the outcome x only for one setting at a time.



A state of Alice's system assigns a probability distribution of outcomes to each of Alice's measurement settings — i.e. here state is $P(x|a)$, the conditional probability of finding outcome x when setting a is chosen. In quantum mechanics, the measurements are POVMs $\{E_x^{(a)}\}$ i.e., for each value of the setting a, a POVM $M^{(a)} = \{E_x^{(a)}\}$ such that $\sum_x E_x^{(a)} = I$. And if ρ is the density operator of Alice's system, $P(x|a) = \mathrm{tr}\left(E_x^{(a)} \rho\right)$.

In QM, Alice cannot perform two (or more) measurements simultaneously because the measurements are noncommuting. Once she measures $M^{(a)}$, she can't undo the measurement to find out what would happen if she measured $M^{(b)}$ instead.

But in a HVT (a "classical" model) we suppose that if we had complete knowledge of the system's "configuration" then there would be <u>definite</u> outcome for each of the m measurements. These configurations are the <u>extremal</u> states of the HVT, and a general state is a convex combination of these extremal distributions. That is, a <u>probability</u> is assigned to each configuration.

How many configurations are there? For each of m settings we can choose any one of v outcomes; therefore

v^m configurations

What is the dimension of the state space?
We denote the space of states by Ω. Since for each a, $\sum_x P(x|a) = 1$, there are $v-1$ independent probabilities for each setting a, and therefore

$$\dim \Omega = m(v-1)$$

But what is the dimension of the space of HVTs?
We assign a probability to each of v^m configs.
So, if Ω_H denotes the space of HVTs

$$\dim \Omega_H = v^m - 1.$$

So Ω_H is larger than Ω — that is why we say the variables are "hidden." Much of the info about the prob distribution of configs remains inaccessible, even if we have complete knowledge of the state.

E.g. for $m = v = 2$, $\dim \Omega = 2$, $\dim \Omega_H = 3$
But disparity widens for more outcomes or settings:
$m=2 \quad v=3 \Rightarrow \dim \Omega = 4 \quad \dim \Omega_H = 8$
$m=3 \quad v=2 \Rightarrow \dim \Omega = 3 \quad \dim \Omega_H = 7$

It is easy to find a HVT that matches the predictions of QM for a single party. Each config has the form

$$(x_{a_1} x_{a_2} \ldots x_{a_m})$$

A length-m string in base-v

We can also denote this

$$(a_1, a_2, \ldots, a_m)$$

with understanding that the symbol a_j denotes outcome for jth setting

The Hidden state is

$$\{ P(a_1, a_2, \ldots, a_m) \} \text{ where}$$

$$\sum_{a_1} \sum_{a_2} \cdots \sum_{a_m} P(a_1, a_2, \ldots, a_m) = 1$$

each summed over v values

Actually, it might be a better notation to use place to indicate the measurement setting, so HVT is

$$\underline{P}(x_1, x_2, x_3, - x_m) \text{ or perhaps } \boxed{\underline{P}(a_1 = x_1, --, a_m = x_m)}$$

where x_j is outcome of measuring observable a_j and $\sum_{x_1 - x_m} \underline{P}(x_1, x_2, -, x_m) = 1$

Then the <u>observable</u> state is a marginal of the hidden state

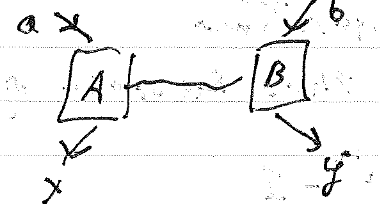
$$\underline{P}(x_j | a_j) = \sum_{\substack{x_1 - x_{j-1} \\ x_{j+1} - x_m}} \underline{P}(x_1, - x_{j-1}, x_j, x_{j+1}, -, x_m)$$

Question: given the marginals $\underline{P}(x_j | a_j)$, is there always a global probability distribution (a HVT) compatible with all marginals (the state). Yes, there is a trivial solution, in which the prob distributions for each setting are independent

$$\underline{P}(a_1 = x_1, -, a_m = x_m) = \underline{P}(x_1 | a_1) \underline{P}(x_2 | a_2) - \underline{P}(x_m | a_m)$$

Clearly $\underline{P}(a_1 = x_1, -, a_m = x_m)$ is normalized and has the right marginals.

Now consider the case of two (or more) parties. This case is more interesting because there is a notion of locality. E.g. two parties, Alice (A) and Bob (B) independently choose their measurement settings. They do not communicate, so Alice does not know Bob's setting, and Bob does not know Alice's



Each party finds an outcome. A state specifies prob distributions for Alice's and Bob's outcomes for each possible setting chosen by Alice and Bob: $\underline{P}(x, y | a, b)$

Alice outcome

Bob outcome

Alice setting

Bob setting

3 Nov 08

There are m^2 settings, and for each there are v^2 outcomes, whose probs of the outcomes sum to 1 so the 2-party state space Ω_2 has

$$\dim \Omega_2 = m^2(v^2-1)$$

And for n parties $\boxed{\dim \Omega_n = m^n(v^n-1)}$

E.g for $n=m=v=2$, $\dim \Omega = 12$

In QM, the A+B measurements are local POVMs

Alice measures $M_A^{(a)} = \{E_x^{(a)} \otimes I\}$

Bob measures $M_B^{(b)} = \{I \otimes E_y^{(b)}\}$

or — between A+B, they jointly perform a POVM with elements

$$\{E_x^{(a)} \otimes E_y^{(b)}\}$$

If ρ_{AB} is the shared density operator

$$P(x,y|a,b) = tr\left((E_x^{(a)} \otimes E_y^{(b)})\,\rho_{AB}\right)$$

What are the configurations? First consider the NLHVT (no locality constraint). Then for n parties, we allow the prob distribution governing the outcomes $(x_1, \ldots x_n)$ to depend on all the settings. This is a NL model, because eg. Alice's outcome might depend not just on her own setting, but also on Bob's. A config assigns one of the v^n possible outcomes to each of m^n possible settings. There are

$$(v^n)^{m^n} \text{ configurations}$$

i.e. $4^4 = 256$ for $n=m=v=2$. Thus the space of NLHVT has

$$\dim \Omega_H = (v^n)^{m^n} - 1$$

Given any state $P(x_1, \ldots x_n | a_1, \ldots a_n)$, there is a trivial NLHVT, i.e. an independent distribution for each of the m^n settings

3 Nov 08

But, in the case of a LHVT, we assume that a configuration specifies an outcome for Alice's measurement depends only on Alice's setting, that an outcome for Bob's measurement is determined by Bob's setting alone. There are v^m local configs for each party, and so a total of

$$(v^m)^n = v^{mn} \text{ local configs for } n \text{ parties}$$

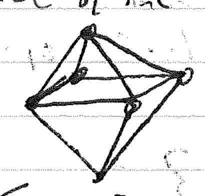
A LHVT assigns a probability to each of the local configs, so

$$\left.\dim \Omega_H^{local} = v^{mn} - 1\right] \text{(Bad notation?)}$$

E.g. $\dim \Omega_H = 14$ for $n=m=v=2$. Though there are far fewer local configs than configs, the state space has a much lower dimension than the space of LHVT when the number of parties and settings is large

$$\left.\begin{array}{l}\dim \Omega_H^{local} \simeq v^{mn} \\ \dim \Omega \simeq m^n v^n\end{array}\right\} \Rightarrow \frac{\dim \Omega_H^{local}}{\dim \Omega} \approx \left(\frac{v^{m-1}}{m}\right)^n$$

Each local config is an extremal local model, i.e. an extremal point in the space of states realized by local models. The space of LHVTs is a polytope embedded in Ω — the convex hull of the configs. We denote this space of local models as



$$\mathcal{L} = \text{convex hull of local configs}$$

(The "classical" or "local region" contained in Ω.)
A "polytope" means the convex hull of a finite set of points in a vector space.

Suppose that, after many repeated experiments, we determine a state $\xi \in \Omega$ to good statistical accuracy. Now we would like to know, is ξ contained in $\mathcal{L}$ / is it a "local" state described by a LHVT?

This may be a hard question to answer, even though we know all the extreme points of the polytope $\mathcal{C}$, because if there are many extreme points the geometry of the polytope is very complex. However there is a <u>dual</u> characterization of a polytope — in terms of its faces rather than its vertices. Or, in other words, in terms of extremal <u>inequalities</u> rather than extremal <u>vectors</u>

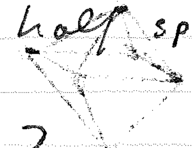
In d-dimensional inner product space, we may characterize a (d-1)-dim hyperplane that does not pass thru the origin by its <u>normal vector</u>

$$\{\xi\in\Omega \mid \xi\cdot\beta = 1\}$$

We may set the inner product to 1 by rescaling β (All the points in the plane have the same nonzero component along β.) Furthermore, a "half space" containing the plane and the points lying on one side of it is characterized by an inequality

$$\{\xi\in\Omega \mid \xi\cdot\beta\le 1\}$$ (This is the half space that <u>contains</u> the origin)

Instead of specifying the polytope as a convex hull of a set of points, i.e. in terms of its extreme points, we may consider the set of half spaces that contain the polytope



$$\mathcal{B} = \{\beta\in\Omega \mid \beta\cdot\xi\le 1 \quad \forall\, \xi\in\mathcal{C}\}$$

and $\mathcal{C}$ is related to $\mathcal{B}$ by

$$\mathcal{C} = \{\xi\in\Omega \mid \beta\cdot\xi\le 1 \quad \forall\, \beta\in\mathcal{B}\}$$

Clearly $\mathcal{B}$ is convex: If $\beta_1\cdot\xi\le 1$ and $\beta_2\cdot\xi\le 1$, then $[\lambda\beta_1+(1-\lambda)\beta_2]\cdot\xi\le 1$ for $\lambda\in[0,1]$

We say the "$\mathcal{B}$ is the polar of $\mathcal{C}$" and that "$\mathcal{B}$ and $\mathcal{C}$ are dual polytopes". $\mathcal{B}$ contains all normal vectors to planes ~~that do not intersect~~ such that the polytope $\mathcal{C}$ lies in a half space bounded by the plane. The extreme vectors in $\mathcal{B}$ specify planes that contain faces of the polytope.

So instead of characterizing $\mathcal{C}$ as the convex hull of the local configurations, we instead characterize it as the set of points ξ that satisfy

$$\xi \cdot \beta_j \le 1 \quad \text{where } \beta_j \text{ is an extremal vector in } \mathcal{B}$$

These extremal inequalities, $\{\xi \cdot \beta_j \le 1\}$, are called "Bell inequalities". If $\xi \in \mathcal{C}$ it obeys all Bell $\neq$'s and if $\xi \notin \mathcal{C}$, then ξ violates at least one Bell $\neq$. Furthermore once we know the $\{\beta_j\}$, the Bell $\neq$'s are easy to check, so we have an efficient algorithm for answering whether $\xi \in \mathcal{C}$

But, there might be exponentially many Bell $\neq$'s to check?

However, the catch is that, when there are many extremal points, finding the extreme points of the polar $\mathcal{B}$ of a polytope $\mathcal{C}$ with specified extreme points (i.e. finding all the faces of the polytope) is a hard problem in convex geometry. So for many parties and settings, finding the Bell $\neq$'s that characterize the classical local region is intractable. But for the case $n=m=v=2$ it is possible

Even here, since the geometry of the 12 dim state space is hard to visualize we'll simplify the problem by considering only the "correlator" of Alice's and Bob's observables. We denote Alice's two observables as a_s, $s \in \{0,1\}$ and Bob's observables by b_t, $t \in \{0,1\}$. Suppose that both observables take values in $\{+1,-1\}$. By the correlator we mean the expectation value

$\langle a_s b_t \rangle$, or in other words the four quantities

$$(\langle a_0 b_0\rangle, \langle a_0 b_1\rangle, \langle a_1 b_0\rangle, \langle a_1 b_1\rangle)$$

Thus the correlator lives in a 4-dim space.

If you flip the sign of all four, then correlator is the same

Now, if each of a_0, a_1, b_0, b_1 have definite values, these quantities are not independent – rather their product is $a_0^2 a_1^2 b_0^2 b_1^2 = 1$. Therefore there are 8 extreme points in the classical region of the correlator space – the number of (-1)'s must be even. The eight extreme vectors are

$$\begin{matrix} + & + & + & + \\ + & + & - & - \\ + & - & + & - \\ + & - & - & + \end{matrix} \qquad \text{and their complements} \qquad \begin{matrix} - & - & - & - \\ - & - & + & + \\ - & + & - & + \\ - & + & + & - \end{matrix}$$

$a_0 b_0 = 1$ $\qquad\qquad$ $a_0 b_0 = -1$

(here $+,-$ is shorthand for $+1, -1$.) A vector $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)$ is in the polar $\mathbb{B}$ iff it satisfies all eight of the inequalities

$$-1 \le \beta_0 + \beta_1 + \beta_2 + \beta_3 \le 1$$
$$-1 \le \beta_0 + \beta_1 - \beta_2 - \beta_3 \le 1$$
$$-1 \le \beta_0 - \beta_1 + \beta_2 - \beta_3 \le 1$$
$$-1 \le \beta_0 - \beta_1 - \beta_2 + \beta_3 \le 1$$

If β has an extreme value then all 4 are equalities, i.e.

$$\beta_0 + \beta_1 + \beta_2 + \beta_3 = f_0$$
$$\beta_0 + \beta_1 - \beta_2 - \beta_3 = f_2$$
$$\beta_0 - \beta_1 + \beta_2 - \beta_3 = f_1$$
$$\beta_0 - \beta_1 - \beta_2 + \beta_3 = f_3$$

where $f_0, f_1, f_2, f_3 \in \{+1, -1\}$

– so there are 16 ways to choose $\vec{f}$, each determines an extreme $\vec{\beta}$

So for each of the 16 choices for $\vec{f}$, we solve for $\vec{\beta}$, and determine an extreme point in $\mathcal{B}$ is a face of $\mathcal{E}$

~~~~ skip ahead to middle of next page.

It is convenient to use binary notation for the components of the vectors

[Or -- just skip right to the solution!]

$$\vec{\beta} = (\beta_{00}, \beta_{01}, \beta_{10}, \beta_{11})$$
$$\vec{f} = (f_{00}, f_{01}, f_{10}, f_{11})$$

Then for the extreme points in $\mathcal{E}$, if $a_0 a_1 = (-1)^{r_0}$, $b_0 b_1 = (-1)^{r_1}$

then $\xi^{(r)}_s = (-1)^{r_0 s_0}(-1)^{r_1 s_1} = (-1)^{r \cdot s}$ or this times (-1) for the complement vectors

where $r \cdot s = r_0 s_0 + r_1 s_1$

In the binary notation, our equalities are

$$f_r = \sum_s (-1)^{r \cdot s} \beta_s$$

These equations have the solution

$$\beta_s = \frac{1}{4} \sum_r (-1)^{r \cdot s} f_r$$

To check the solution, note that

$$f_r \stackrel{?}{=} \sum_s (-1)^{r \cdot s} \frac{1}{4} \sum_{r'} (-1)^{r' \cdot s} f_{r'}$$

$$= \sum_{r'} f_{r'} \frac{1}{4} \sum_s (-1)^{r \cdot s} (-1)^{r' \cdot s}$$

More generally $\frac{1}{2^n} \sum_{s \in \{0,1\}^n} (-1)^{r \cdot s} = \prod_{i=0}^{n-1} \left(\frac{1}{2} \sum_{s_i \in \{0,1\}} (-1)^{r_i s_i} \right)$

$$= \prod_{i=0}^{n-1} \left[\frac{1}{2} \left(1 + (-1)^{r_i}\right) \right] = \delta_{r,0}$$

But in fact the polytope has so much symmetry in this case, that there are really only 2 inequalities
~~~~

The symmetries are:

$a_0 \leftrightarrow a_1$ We can interchange Alice's two
$b_0 \leftrightarrow b_1$ settings, or Bob's two settings,
$A \leftrightarrow B$ or Alice and Bob, or the two outcomes
$1 \leftrightarrow -1$ All these mappings preserve the polytope

i.e. $a_0 \leftrightarrow a_1$ $\quad \xi_0 \leftrightarrow \xi_2 \quad \xi_1 \leftrightarrow \xi_3$ Flip one bit
$b_0 \leftrightarrow b_1$ $\quad \xi_0 \leftrightarrow \xi_1 \quad \xi_2 \leftrightarrow \xi_3$ Flip one bit
$A \leftrightarrow B$ $\quad \xi_1 \leftrightarrow \xi_2$ Exchange bits

i.e. $\quad r_1 r_0 \to \bar{r}_1 r_0$
$r_1 r_0 \to r_1 \bar{r}_0$
$r_1 r_0 \leftrightarrow r_0 r_1$

Also $1 \leftrightarrow -1$ means $\xi \to -\xi$

Solution is
$$\beta_0 = \tfrac{1}{4}(f_0 + f_1 + f_2 + f_3)$$
$$\beta_1 = \tfrac{1}{4}(f_0 - f_1 + f_2 - f_3)$$
$$\beta_2 = \tfrac{1}{4}(f_0 + f_1 - f_2 - f_3)$$
$$\beta_3 = \tfrac{1}{4}(f_0 - f_1 - f_2 + f_3)$$
we can see this solves eqns by inspection. We don't need the "Fourier Transform"

So – for $f = (1\,1\,1\,1) \to \beta = (1\,0\,0\,0)$
$f = (1,1,-1,-1) \to \beta = (0\,0\,1\,0)$
$f = (1,-1,1,-1) \to \beta = (0\,1\,0\,0)$
$f = (1,-1,-1,1) \to \beta = (0\,0\,0\,1)$

f has even parity $\Rightarrow$ β has one nonzero component

These are the "trivial" inequalities

$-1 \le \langle a_0 b_0 \rangle \le 1$
$-1 \le \langle a_1 b_0 \rangle \le 1$ $\quad a_0 \leftrightarrow a_1$
$-1 \le \langle a_0 b_1 \rangle \le 1$
$-1 \le \langle a_1 b_1 \rangle \le 1$ $\quad a_0 \leftrightarrow a_1$

$b_0 \leftrightarrow b_1$

3 Nov 08

$f = (1,1,1,-1) \longrightarrow \beta = (\frac{1}{2}\ \frac{1}{2}\ \frac{1}{2}\ -\frac{1}{2})$
$f = (1,1,-1,1) \longrightarrow \beta = (\frac{1}{2}\ \frac{1}{2}\ -\frac{1}{2}\ \frac{1}{2})$
$f = (1,-1,1,1) \longrightarrow \beta = (\frac{1}{2}\ -\frac{1}{2}\ \frac{1}{2}\ \frac{1}{2})$
$f = (-1,1,1,1) \longrightarrow \beta = (-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2})$

f has odd parity, all components of β are nonzero

Same symmetries $\Rightarrow$ we can pick any of the 4 components as having minus sign

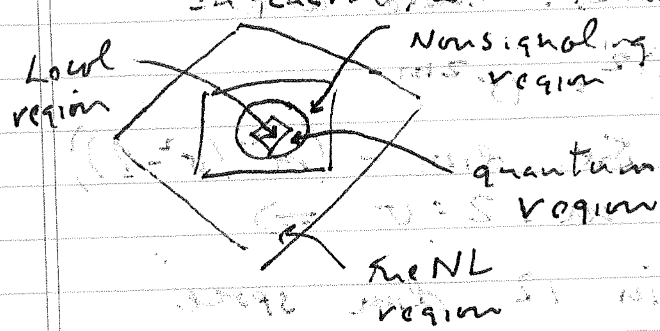
<u>This</u> is the CHSH inequality:

$$-1 \le \tfrac{1}{2}\langle a_0 b_0 + a_0 b_1 + a_1 b_0 - a_1 b_1 \rangle \le 1$$

plus the inequalities related to it by symmetry

Thus the CHSH inequality, plus the trivial $\neq$, completely characterize $\mathcal{C}$

In general, we have nested convex regions



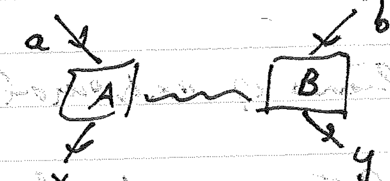
Physicist cherish the principle of locality – but LHVTs do not agree with the tested principles of QM

So – if we stick with QM and locality, then we must reject the ~~hidden~~ hidden variable concept – the randomness of quantum meas. outcomes is really intrinsic, <u>not</u> the result of ignorance.

5 NOV 2008
(O. WINS!)

<u>Local Hidden Variables</u>

Consider the two party case:



Alice and Bob choose from among m measurements, each with v outcomes. The shared state is $\underline{P}(x, y \mid a, b)$

This is what A & B can observ

– lies in a space Ω, $\dim \Omega = m^2(v^2-1)$

Hidden variable hypothesis: probability is due to ignorance. If system configuration is perfectly known, outcome is deterministic for each setting (a particular outcome occurs with probability 1).

Locality hypothesis: configurations are local: each parties outcome determined by that party's setting alone.

Configurations: $P(xy|ab) = \delta(x, x_a)\,\delta(y, y_b)$

Kronecker delta

In the space $\mathcal{P}$, there are extremal states, the deterministic states. There are

$$(v^m)^2 = v^{2m}$$

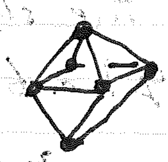
local extremal states in the dim-$(m^2(v^2-1))$ states space. E.g. $m = 2 = v \Rightarrow$

16 configs in 12 dim space

A LHVT is a prob. distribution on the set of configs. If $\{\mathcal{E}_i,\ i = 1, 2, \ldots, v^{2m}\}$ are the extremal states, then the local models are the states

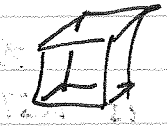
$$\mathcal{C} = \{\sum p_i \mathcal{E}_i,\ \text{where } \sum p_i = 1\}$$

— "the convex combinations of extremal states"



$\mathcal{C}$ is a polytope embedded in Ω

There is a dual polytope $\mathcal{B}$, the polar of $\mathcal{C}$



$\mathcal{B}$ defined relative to center of polytope

$$\mathcal{B} = \{\beta \in \Omega \mid \beta \cdot \xi \le 1 \quad \forall \xi \in \mathcal{C}\}$$

- These are normal vectors of planes bounding half spaces that contain $\mathcal{C}$. Also convex, and in fact a polytope, whose extremal vectors are faces of $\mathcal{C}$.

Bell inequality: $\xi \cdot \beta_i \le 1$ where β_i is extremal in $\mathcal{B}$

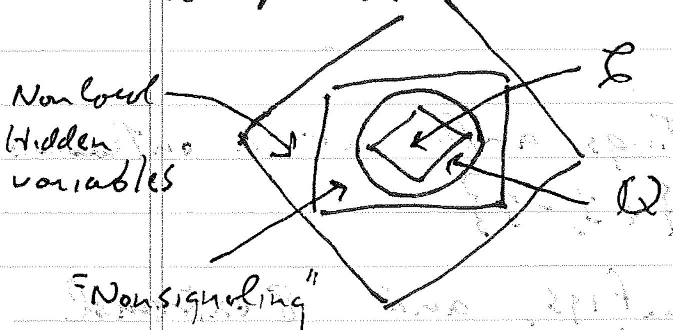
The set of all Bell inequalities completely characterizes the "classical" (or "local") region $\mathcal{C}$.

But there is also a quantum region

$$\mathcal{Q} = \{ \underline{P}(xy|ab) = \mathrm{tr}[(E_x^{(a)} \otimes E_y^{(b)})\rho] \}$$

for some choice of $H_A \otimes H_B$, ρ_{AB}, and measurements $M^{(a)} = \{E_x^{(a)}\}$ $M^{(b)} = \{E_y^{(b)}\}$

What's important is that there are points in $\mathcal{Q}$ that are not in $\mathcal{C}$. "Quantum correlations" are not compatible with local models.



There is a nested set of regions inside Ω as shown
- each region is convex, and each inclusion is proper.

We are stuck with this choice; Bell ≠ violation is observed.
- If HV models describe Nature, then must be nonlocal
- If physics is local, then no HV.

Thus, randomness in QM is intrinsic, not due to ignorance

That is, the outcome of a q. measurement is not predictable, even when our knowledge is the most complete possible!

Consider the correlator

$\langle a_s b_t \rangle$ s, t take m values, the measurement settings

Think of this as a vector in a space of dimension m^2

In a LHVT, $\langle \cdot \rangle$ is evaluated by summing over outcomes, weighted by probabilities i.e. sum over σ^2 outcomes for each fixed s and t

$$\langle a_s b_t \rangle = \sum_{xy} xy\, P(xy \mid st)$$

In the m^2 dimensional correlator space, each of σ^{2m} configs determines a point (though some of these points may coincide)

The local region $\mathcal{L}$ is a polytope in correlator space, with these vertices. What is the polar $\mathcal{B}$

Example: $m=2$ settings and $\sigma=2$ outcomes

$$a_0, a_1, b_0, b_1 \in \{+1, -1\}$$

There are $2^4 = 16$ configs, and 8 extremal points in correlator space

$$\{\langle a_0 b_0\rangle, \langle a_0 b_1\rangle, \langle a_1 b_0\rangle, \langle a_1 b_1\rangle\}$$

(flipping sign of all outcomes does not change correlator). In the last lecture we saw that there are 16 extremal vectors in the polar, parameterized as

$$\beta_0 = \tfrac{1}{4}(f_0 + f_1 + f_2 + f_3)$$
$$\beta_1 = \tfrac{1}{4}(f_0 - f_1 + f_2 - f_3)$$
$$\beta_2 = \tfrac{1}{4}(f_0 + f_1 - f_2 - f_3)$$
$$\beta_3 = \tfrac{1}{4}(f_0 - f_1 - f_2 + f_3)$$

where $f_0, f_1, f_2, f_3 \in \{+1, -1\}$

Thus, in the correlator space, there are 16 Bell $\neq$s that characterize the local region completely. Actually, the polytope has a lot of symmetry, so there are really only two different kinds of inequalities:

If f has even parity, e.g. $f = (1, 1, 1, 1)$, then

e.g. $\beta = (1, 0, 0, 0)$

For the eight even-parity fs we get β with one nonzero component, ± 1. So these are eight inequalities

$$-1 \le \langle a_0 b_0 \rangle \le 1$$
$$-1 \le \langle a_1 b_0 \rangle \le 1$$
$$-1 \le \langle a_0 b_1 \rangle \le 1$$
$$-1 \le \langle a_1 b_1 \rangle \le 1$$

these are "trivial" $\neq$s that follow from $a_s, b_t \in \{\pm 1\}$

If f has odd parity, e.g. $f = (1, 1, 1, -1)$, then e.g. $\beta = (\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2})$

For the eight odd-parity fs we get β with either three $+\tfrac{1}{2}$ components and one $-\tfrac{1}{2}$ component, or three $-\tfrac{1}{2}$ components and one $+\tfrac{1}{2}$ component.

So — there are eight inequalities

$$-1 \le \frac{1}{2}\langle a_0 b_0 + a_1 b_0 + a_0 b_1 - a_1 b_1 \rangle \le 1$$

(plus 3 other choices for the component with the minus sign.) These are the CHSH inequalities. As we have already discussed, in QM a violation can occur, with e.g.

$$\frac{1}{2}\langle a_0 b_0 + a_1 b_0 + a_0 b_1 - a_1 b_1 \rangle = \pm\sqrt{2}$$

Let's discuss the properties of the quantum region

$$\mathbb{Q} = \left\{ P(xy|ab) = tr\left[\left(E_x^{(a)} \otimes E_y^{(b)} \right) \rho \right] \right\}$$

① $\mathbb{Q}$ is convex (but not a polytope)

② $\mathbb{Q}$ is contained inside the NLHVT polytope (and in fact inside the "nonsignaling region")

③ $\mathbb{Q}$ contains $\mathcal{C}$

④ If ρ is separable (a convex combination of product states), then $\mathcal{E} \in \mathcal{C}$ (there is a local model of the correlation between Alice's and Bob's measurements)

⑤ If each party chooses among mutually commuting observables, then $\mathcal{E} \in \mathcal{C}$

Thus ④ and ⑤ imply that, for Bell ≠ violation, A and B must share an entangled state, and must choose from among noncommuting observables

Let's discuss these properties in greater detail

① convexity. Recall that a region R in a vector space is convex if $\xi_1, \xi_2 \in R \Rightarrow \lambda \xi_1 + (1-\lambda)\xi_2 \in R$ for $\lambda \in [0,1]$

So convexity of Q means

$$\left.\begin{array}{l} P^{(1)}(xy|ab) \in Q \\ P^{(2)}(xy|ab) \in Q \end{array}\right\} \Rightarrow \lambda P^{(1)}(xy|ab) + (1-\lambda) P^{(2)}(xy|ab) \in Q$$
for $\lambda \in [0,1]$

This property is not immediately obvious, since $P^{(1)}$ and $P^{(2)}$ could be realizes with different density operators and different sets of measurements.

But.. if $P^{(1)}$ realized by $\rho^{(1)}$, $M^{(1a)} = \{E_x^{(1,a)}\}$ on
$M^{(1b)} = \{E_y^{(1,b)}\}$ $H_A^{(1)} \otimes H_B^{(1)}$

and $P^{(2)}$ realized by
$\rho^{(2)}$, $M^{(2,a)} = \{E_x^{(2,a)}\}$, $M^{2,b} = \{E_y^{(2,b)}\}$
on $H_A^{(2)} \otimes H_B^{(2)}$

Then consider Hilbert space

$$\left(H_A^{(1)} \oplus H_A^{(2)}\right) \otimes \left(H_B^{(1)} \oplus H_B^{(2)}\right)$$

and state $\rho(\lambda) = \begin{pmatrix} \lambda\rho^{(1)} & 0 \\ 0 & (1-\lambda)\rho^{(2)} \end{pmatrix}$

Measurement

$$M^{(a)} = \left\{\begin{pmatrix} E_x^{(1,a)} & 0 \\ 0 & E_x^{(2,a)} \end{pmatrix}\right\} \quad M^{(b)} = \left\{\begin{pmatrix} E_y^{(1,b)} & 0 \\ 0 & E_y^{(2,b)} \end{pmatrix}\right\}$$

(state and POVM elements are block-diagonal).

Then

$$P(x,y|ab) = \mathrm{tr}\begin{pmatrix} E_x^{(1,a)} & 0 \\ 0 & E_x^{(2,a)} \end{pmatrix} \otimes \begin{pmatrix} E_y^{(1,b)} & 0 \\ 0 & E_y^{(2,b)} \end{pmatrix} \begin{pmatrix} \lambda\rho^{(1)} & 0 \\ 0 & (1-\lambda)\rho^{(2)} \end{pmatrix}$$

The full Hilbert space is $\mathcal{H} = \mathcal{H}_A^{(1)} \otimes \mathcal{H}_B^{(1)} \oplus \mathcal{H}_A^{(1)} \otimes \mathcal{H}_B^{(2)} \oplus \mathcal{H}_A^{(2)} \otimes \mathcal{H}_B^{(1)} \oplus \mathcal{H}_A^{(2)} \otimes \mathcal{H}_B^{(2)}$

E.g. $\rho^{(1)}$ has support on $\mathcal{H}_A^{(1)} \otimes \mathcal{H}_B^{(1)}$

$E_A^{(1)} \otimes E_B^{(2)}$ has support on $\mathcal{H}_A^{(1)} \otimes \mathcal{H}_B^{(2)}$ $\Rightarrow$ $\mathrm{tr}(E_A^{(1)} \otimes E_B^{(2)} \rho^{(1)}) = 0$

or $\quad \underline{P}(xy|ab) = \lambda \underline{P}^{(1)}(xy|ab) + (1-\lambda)\underline{P}^{(2)}(xy|ab)$

(for each x, y, a, b).

② There is a NLHVT for $\underline{P}(xy|ab)$

— we've already discussed this — in NLHVT, we are free to choose independent distribution for each choice of settings a, b.

Also $\underline{P}(xy|ab)$ is nonsignaling

$$\sum_y \underline{P}(xy|ab) = \sum_y \mathrm{tr}(E_x^{(a)} \otimes E_y^{(b)})\rho_{AB}$$

and $\quad \sum_y E_y^{(b)} = I \Rightarrow \quad = \mathrm{tr}(E_x^{(a)} \otimes I)\rho_{AB} = \mathrm{tr}(E_x^{(a)} \rho_A)$

— does not depend on Bob's choice b. Similarly

$\sum_x \underline{P}(xy|ab)$ does not depend on Alice's choice a

③ $\mathcal{Q}$ contains $\mathcal{L}$

We may choose each measurement to act on a different subsystem of $\mathcal{H}_A$ (and similar for $\mathcal{H}_B$)

$$\mathcal{H}_A = \boxed{\mathcal{H}_A^{(a_0)}} \otimes \boxed{\mathcal{H}_A^{(a_1)}} \otimes \cdots \otimes \boxed{\mathcal{H}_A^{(a_{m-1})}}$$

Each has $\dim = v$, and $\{E_x^{(a_j)}\}$ is an orthog measurement

A configuration is a basis state

$$|x\rangle = |x_0\rangle \otimes |x_1\rangle \otimes \cdots \otimes |x_{m-1}\rangle$$

and an arbitrary local model is realized by

$$\sum_x p_x |x\rangle\langle x|$$

⋮

④ Separable $\rho \Rightarrow \underline{P}(xy|ab)$ admits a local model

For a product state, $|\psi\rangle \otimes |\varphi\rangle$

$$\underline{P}(xy|ab) = \langle\psi|E_x^{(a)}|\psi\rangle\langle\varphi|E_y^{(b)}|\varphi\rangle$$
$$= \underline{P}(x|a)\,\underline{P}(y|b)$$ A and B meas. are uncorrelated

So, as we've seen, there is a NVT for $\underline{P}(x|a)$ and a NVT for $\underline{P}(y|b)$. The two together specify a LHVT. So product state $\Rightarrow \underline{P}(xy|ab) \in \mathcal{C}$

And a separable state is $\rho = \sum_{ij} p_{ij}\, |\psi_i\rangle\langle\psi_i| \otimes |\varphi_j\rangle\langle\varphi_j|$

$$\Rightarrow \underline{P}(xy|ab) = \sum_{ij} p_{ij}\, \underline{P}^{(i)}(x|a)\, \underline{P}^{(j)}(y|b)$$

This is just a convex combination of elements of $\mathcal{C}$; therefore in $\mathcal{C}$. (A separable state is one Alice and Bob can create locally, if they have "shared randomness")

⑤ If measurements $\{M^{(a)}\}$ are mutually commuting for each party, then $\{\underline{P}(x,y|ab)\} \in \mathcal{C}$

The statement holds for POVMs, where each $E_x^{(a_i)}$ commutes with each $E_x^{(a_j)}$, but to simplify notation, suppose all measurements are orthogonal measurements i.e. observables $\{a_j\}$ are mutually commuting and can be simultaneously diagonalized. We can expand in orthonormal basis states

or, consider orthog meas. on extended Hilbert space

$$\{\,|x_0, x_1, \ldots x_{m-1}, y_0, y_1, \ldots y_{m-1}\rangle\}$$

(there might also be an additional index, if simultaneous eigenstates of all observables are degenerate; I've suppressed this index

A pure state, expanded in this basis is

$$|\psi\rangle = \sum C(x,y)\ |x,y\rangle;$$

which specifies a prob. distribution on configurations

$$P(x,y) = |C(x,y)|^2$$

what if A's measurements commute but B's do not?

– this is a local model. And if ρ is a convex combination of pure states, we have a convex comb of elements of $\mathcal{E} \Rightarrow$ also in $\mathcal{E}$

The complete characterization of the convex region Q (e.g. all of its extremal points) is complicated in general, but is known for the $n=m=v=2$ case. Consider the correlator in a quantum model. The CHSH ≠ is a statement about

$$\langle C\rangle = \mathrm{tr}\, C\rho$$

where $$C = a_0 b_0 + a_1 b_0 + a_0 b_1 - a_1 b_1$$

and a_0, b_0, a_1, b_1 are Hermitian ops with eigenvalues ± 1

The <u>maximal</u> violation of CHSH is given by

$$\max_{\rho} |\mathrm{tr}(C\rho)|$$

Because of convexity of set of density operators, the max occurs for an extremal i.e. pure state. So we would like to determine

$$\max_{|\psi\rangle} |\langle\psi|C|\psi\rangle| \quad \text{where } C \text{ is Hermitian}$$

Thus, $$= \max|\lambda(C)| \equiv \|C\|_{sup}$$

eigenvalue

"sup norm" or L^∞ norm

The sup norm has the properties

$$\|M+N\| \le \|M\| + \|N\|$$

$$\|MN\| \le \|M\| \cdot \|N\|$$

So we can get an upper bound, using $\|a_s\| = I = \|b_t\|$

$$\|C\| \le 4$$

But we get a tighter bound by a more clever argument. Note that, for a hermitian operator C

$$\|C^2\| = \|C\|^2$$

So consider the square of $C = (a_0+a_1)b_0 + (a_0-a_1)b_1$

$$C^2 = (a_0+a_1)^2 b_0^2 + (a_0-a_1)b_1^2$$
$$+ (a_0+a_1)(a_0-a_1)b_0 b_1 + (a_0-a_1)(a_0+a_1)b_1 b_0$$

Now, $a_s^2 = I = a_t^2$, so

$$C^2 = 2(a_0^2 + a_1^2) + (a_1 a_0 - a_0 a_1)b_0 b_1 + (a_0 a_1 - a_1 a_0)b_1 b_0$$
$$= 4I + [a_1, a_0][b_0, b_1]$$

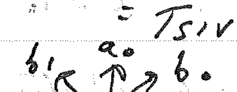
So — if a_0 and a_1 commute, or if b_0 and b_1 commute we have $\|C\| = 2$. Thus we recover the CHSH ≠

$$|\langle C \rangle| \le 2$$

But for noncommuting measurements

$$\|a_1 a_0 - a_0 a_1\| \le 2 \implies \|C^2\| \le 8$$
$$\|b_0 b_1 - b_1 b_0\| \le 2$$

$$\implies |\langle C \rangle| \le 2\sqrt{2}$$ this is the "Tsirelson bound"

saturated in our example 

$|\psi\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$

Experiments:

Done with photons e.g. qubit encoded in polarization $|H\rangle, |V\rangle$

or with ions e.g. qubit encoded in $|\text{ground}\rangle, |\text{excited}\rangle$

Convincing violations of CHSH seen.

Do the tests have "loopholes"

Locality loophole $\boxed{a_0, a_1}$ ~~~ • ~~~ $\boxed{b_0, b_1}$ spacelike separated "decisions"

Could info about A's decision reach B?

" " B's " " A?

Aspect 1982: 12 m apart (40 ns light travel time)

this time is >> difference in photon arrival time

>> switching time of polarization analyzer

5 σ violation of CHSH seen.

Since then CHSH violation seen for A, B separated by many km

Detection loophole

only a small fraction of photons are detected: {photons absorbed, miss detector, detector does not fire}

Detection at A+B in coincidence is a small fraction of all events. Are these selected events a fair sample?

What if the LHV determine whether detector fires as well as the measurement outcome

i.e. each pair has $x_0, x_1 \in \{\text{click, no click}\}$

$y_0, y_1 \in \{\text{click, no click}\}$

Fraction of events detected η_A, η_B

$\eta_A, \eta_B > 83\% \Rightarrow$ can do a "loophole-free test"

But not done for photons. Try ion traps

$\eta > 99\%$, but measurement not spacelike separated