6 Local hidden variable (LHV) theories and Bell inequalities

Aim. Description in terms of classical probability theory + physically reasonable assumptions.

Assumptions.

(i) One can assign definite values to quantities even before they are measured. These values may, however, be 'hidden' or unknown to the observer so that a probabilistic description may be necessary.

(ii) The properties of one subsystem should not immediately depend on what happens to a very distant other subsystem.

Example scenario. A source emits pairs of particles to distant observers (Alice and Bob) who perform ± 1 valued measurements each.

Assume that both observers choose one out of $m \in \mathbb{N}$ measurement devices denoted by A_x, B_x, with $x, y \in \{1, \ldots, m\}$.
Let $p(a, b \mid x, y)$ be the probability that Alice obtains outcome $a \in \{\pm 1\}$ while Bob obtains outcome $b \in \{\pm 1\}$ in the same run of the experiment, if they have used devices A_x and B_y, respectively.
Consider the expectation value of their product, i.e., $\langle A_x B_y \rangle := \sum_{a,b \in \{\pm 1\}} ab\, p(a, b \mid x, y)$.

LHV Ansatz.

$$\langle A_x B_y \rangle = \int_\Omega A_x(\omega) B_y(\omega)\, dP(\omega) \tag{6.1}$$

with random variables $A_x, B_y \colon \Omega \to \{\pm 1\}$ and a probability measure P.
$\omega \in \Omega$ is the *hidden variable* and *locality* is expressed in the fact that A_x does not depend on y and B_y does not depend on x.

Bell inequalities.

Definition 29. Consider an $m \times m$ tuple $C \in \mathbb{R}^{m \times m}$ of empirically obtained expectation values $\langle A_x B_y \rangle =: C_{xy}$ for pairs of ± 1 valued measurements. Let $\mathcal{C} \subseteq \mathbb{R}^{m \times m}$ be the set of all such C's for which there exists an LHV description as in (6.1).
An inequality for C is called *Bell inequality* if it holds for all $C \in \mathcal{C}$.

Remarks.

- $\mathcal{C}$ is a closed convex polytope, so for a complete description of $\mathcal{C}$ a finite set of linear inequalities suffices.
- There are trivial Bell inequalities of the form $|\langle A_x B_y \rangle| \leq 1$.
- *Stochastic LHV theories* allow the random variables in (6.1) to have ranges in $[-1, 1]$ instead of $\{-1, +1\}$. It turns out, however, that this does not change $\mathcal{C}$ ($\rightarrow$ exercise).

Corollary 30. *Let $\rho \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be classically correlated and $M_x \colon \mathbb{B}(\{\pm 1\}) \to \mathcal{B}(\mathcal{H}_A)$ and similarly $M'_y \colon \mathbb{B}(\{\pm 1\}) \to \mathcal{B}(\mathcal{H}_B)$ be POVMs for $x, y \in \{1, \dots, m\}$.*
Then $C \in \mathcal{C}$ holds for

$$C_{xy} := \sum_{a,b \in \{\pm 1\}} ab \operatorname{tr}\left[\rho\, M_x(a) \otimes M'_y(b)\right].$$

Proof. Define $A_x := \sum_{a \in \{\pm 1\}} a\, M_x(a) = M_x(1) - M_x(-1)$ and $B_y := \sum_{b \in \{\pm 1\}} b\, M'_y(b)$.
Then $-\mathbb{1} \leq A_x, B_y \leq \mathbb{1}$.
Suppose $\rho = \sum_\omega p_\omega\, \rho_A^{(\omega)} \otimes \rho_B^{(\omega)}$ is a convex combination of product states.
Then $C_{xy} = \operatorname{tr}[\rho\, A_x \otimes B_y] = \sum_\omega p_\omega \underbrace{\operatorname{tr}\left[\rho_A^{(\omega)} A_x\right]}_{=:A_y(\omega)} \underbrace{\operatorname{tr}\left[\rho_B^{(\omega)} B_y\right]}_{=:B_y(\omega)}$ is a stochastic LHV description with discrete probability space. □

Consequently, unentangled quantum states can never violate any Bell inequality.
The simplest and most famous non-trivial Bell inequality is the Causer-Horne-Shimony-Holt (CHSH) inequality:

Theorem 31 (CHSH)**.** *Every LHV theory for the description of ± 1 valued measurements satisfies:*

$$\left|\langle A_1 B_1 \rangle + \langle A_1 B_2 \rangle + \langle A_2 B_1 \rangle - \langle A_2 B_2 \rangle\right| \leq 2. \tag{6.2}$$

Proof. L.h.s. $= \left|\int_\Omega A_1(\omega)\big(B_1(\omega) + B_2(\omega)\big) + A_2(\omega)\big(B_1(\omega) + B_2(\omega)\big)\, dP(\omega)\right|$.
For a fixed $\omega \in \Omega$ we can distinguish two cases:

1. $B_1(\omega) = B_2(\omega)$. Then $B_1(\omega) - B_2(\omega) = 0$ and $B_1(\omega) + B_2(\omega) \in \{\pm 2\}$.
2. $B_1(\omega) \neq B_2(\omega)$. Then $B_1(\omega) + B_2(\omega) = 0$ and $B_1(\omega) - B_2(\omega) \in \{\pm 2\}$.

In either case the integrand is in $\{-2, 2\}$ so that the average is in $[-2, 2]$. □

The remarkable thing is, that this can be violated within quantum theory. For the quantum description, we assign POVMs M_x and M'_y to the measurement devices and introduce again $A_x := M_x(1) - M_x(-1), B_y := M'_y(1) - M'_y(-1)$.
With $\hat{\beta} := A_1 \otimes (B_1 + B_2) + A_2 \otimes (B_1 - B_2) \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ the l.h.s. of (6.2) is then given by $\left|\operatorname{tr}\left[\rho\hat{\beta}\right]\right|$.

Theorem 32 (CHSH violation & Cirelson bound). *If $\hat{\beta}$ is as above, then for all density operators $\rho \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$:*

$$\left|\operatorname{tr}\left[\rho\hat{\beta}\right]\right| \leq 2\sqrt{2}.$$

That is, quantum theory violates the CHSH inequality at most by a factor $\sqrt{2}$.
Moreover, there exists a pure state on $\mathbb{C}^2 \otimes \mathbb{C}^2$ and observables $A_x, B_y \in \mathcal{B}(\mathbb{C}^2)$ with eigenvalues ± 1, s.t. equality holds in this equation.

Proof. Consider the map $A_1 \mapsto \operatorname{tr}\left[\rho\hat{\beta}\right]$. Since this is an affine functional the supremum and infimum over the convex set $-\mathbb{1} \leq A_1 \leq \mathbb{1}$ is attained for some extreme point for which $\operatorname{spec}(A_1) = \{\pm 1\}$ and thus $A_1^2 = \mathbb{1}$.
By the same reasoning we can assume that A_1, A_2, B_1, B_2 are all such that their square is $\mathbb{1}$. Using this property, direct computation leads to

$$\hat{\beta}^2 = 4\,\mathbb{1} \otimes \mathbb{1} + [A_2, A_1] \otimes [B_1, B_2].$$

We exploit this via positivity of the variance and obtain

$$\operatorname{tr}\left[\rho\hat{\beta}\right]^2 \leq \operatorname{tr}\left[\rho\hat{\beta}^2\right] = 4 + \operatorname{tr}\left[\rho\,[A_2, A_1] \otimes [B_1, B_2]\right] \tag{6.3}$$
$$\leq 4 + \left\|[A_2, A_1] \otimes [B_1, B_2]\right\| \tag{6.4}$$
$$\leq 8, \tag{6.5}$$

where the last inequality uses that $\|[A_1, A_2]\| \leq \|A_1A_2\| + \|A_2A_1\| \leq 2\|A_1\|\|A_2\| = 2$ and similarly for the B's.
(6.3)-(6.5) shows that $\operatorname{tr}\left[\rho\hat{\beta}\right] \leq 2\sqrt{2}$ as claimed.
In order to prove that equality can be achieved, assume that $\rho = |\psi\rangle\langle\psi|$ where $|\psi\rangle$ is an eigenvector of $\hat{\beta}$ with eigenvalue ν. Then equality holds in (6.3).
Now take $A_1 = B_1 = \sigma_1$ and $A_2 = B_2 = \sigma_2$ Pauli matrices, so that $\hat{\beta}^2 = 4\,\mathbb{1} + 4\,\sigma_3 \otimes \sigma_3$ has eigenvalues 0 and 8. Hence, ν can be chosen s.t. $\operatorname{tr}\left[\rho\hat{\beta}\right]^2 = \nu^2 = 8$. □

Remarks.

- In the early 80's, the violation of CHSH by a factor $\sqrt{2}$ has been verified experimentally. This was done using *down conversion* in a non-linear crystal, which produces entangled pairs of photons, whose polarization degrees of freedom violate CHSH.
- This is a remarkable step in the history of science, since a debate (initially mainly between Einstein and Bohr) that was originally considered metaphysical has eventually been decided by an experiment.

The argumentation can be generalized to more than two observables:

Consider $\langle A_x B_y \rangle =: C_{xy},\ x, y \in \{1, \ldots, m\}$ as before, $\gamma \in \mathbb{R}^{m\times m}$ and define

$$\|\gamma\|_{\text{LHV}} := \sup_{a,b\in\{\pm1\}^m} \left|\sum_{x,y} \gamma_{xy} a_x b_y\right|$$
$$\|\gamma\|_{\text{quantum}} := \sup_{\rho,\{A_x,B_y\}} \left|\sum_{x,y} \gamma_{xy} \operatorname{tr}[\rho\, A_x \otimes B_y]\right|$$

where $\rho \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is a density operator and $-\mathbb{1} \leq A_x, B_y \leq \mathbb{1}$.
The Cirelson bound for the CHSH inequality then reads:

$$\nu(\gamma) := \frac{\|\gamma\|_{\text{quantum}}}{\|\gamma\|_{\text{LHV}}} = \sqrt{2} \text{ for } \gamma = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Theorem 33 (General Cirelson bounds)**.**

(i) $\gamma \in \mathbb{R}^{2\times2} \Rightarrow \nu(\gamma) \leq \sqrt{2}$.

(ii) $\gamma \in \mathbb{R}^{m\times m}, m \in \mathbb{N} \Rightarrow \nu(\gamma) \leq K_G < \frac{\pi}{2\ln(1+\sqrt{2})} \approx 1.782$.

Remark.

- (i) means that the choice of coefficients in CHSH is optimal.
- (ii) is a non-trivial statement whose proof is based on a deep result of Grothendieck. K_G is called *Grothendieck's constant*, which is unknown but equal to the supremum of $\nu(\gamma)$ over all $\gamma \in \mathbb{R}^{m\times m}$ and all $m \in \mathbb{N}$.