Notes on Contextuality

Davide Poderini

1 Introduction

To do.

2 Fundamentals

Suppose we can perform on a system measurements of some observables $A_1, A_2, \ldots, A_n$, whose outcomes are designated by $a_1, a_2, \ldots, a_n$, and suppose that not all of these observables can be measured at once without disturbing the system. We call the properties that can be measured together *compatible*, and *non compatible* the ones that cannot.

We then define as the *context* of the measurement of A_i the other *compatible* properties that are measured together with it. A theory is said to be *non contextual* if the predicted outcome a_i of some observable A_i does not depend on the context. If it does we say that the theory is *contextual*.

In quantum theory compatible measurements are represented by commuting operators on a Hilbert space $\mathcal{H}$. In this section we can limit ourselves to consider only dichotomic measurements, which are represented by projectors in QM. Compatible observables in this case are equivalent to orthogonal projectors.

In QM contextuality is said to be *not explicit*. This means that the outcome of a measure of an observable A_i may depend on the context, but the probability for that single observable $p(A_i = a_i)$, does not. This is also called the *no-disturbance* or *no-signaling* principle.

Nonetheless, as we will show in this section, a non-contextual HVT (hidden variable theory) cannot reproduce the prediction of QM.

2.1 Joint probability

The problem of the existence of a non-contextual HVT can be better expressed in terms of probability distributions. To simplify the notation we call $f(a_1, \ldots, a_k) = p(A_1 = a_1, \ldots, A_k = a_k)$ the probability that the measure of $A_1, \ldots, A_k$ gives outcomes $a_1, \ldots, a_k$.

Quantum mechanics gives us only the joint probability distribution for *compatible* observables. So if $\{A_1, A_2\}$ are compatible observables, but $\{A_1, A_2, A_3\}$ are not, QM gives us only $f(a_1, a_2)$, without any information about an hypothetical $f(a_1, a_2, a_3)$.

A non-contextual HVT instead, the presence of a global joint probability distribution for all observables $f(a_1, \ldots, a_n)$ is assured by the fact that every property of the system must have a definite value, and the uncertainty is only given by our ignorance about the initial conditions. From that global joint distribution one should be able to derive the quantum mechanical ones for compatible observables $f(a_1, \ldots, a_k)$, as marginals.

The possibility of the existence of a joint probability distribution depends on the compatibility relations between the observables. To better visualize them, we can construct a graph where every vertex represents an observable and two vertex are connected if the corresponding observables are compatible, some examples are shown in figure 1.

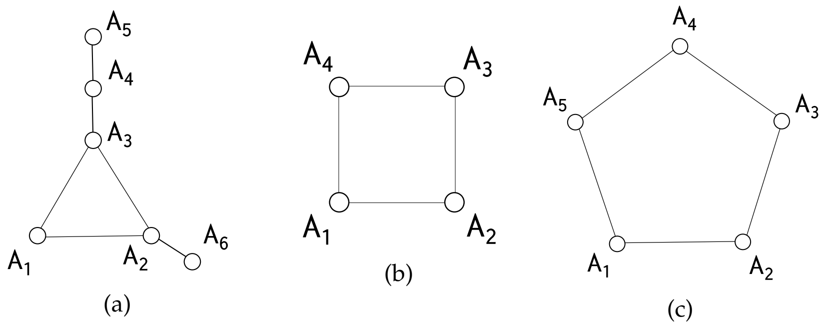


Figure 1: Examples of compatibility graphs

A sufficient condition for the existence of a joint probability distribution is that the corresponding graph should not contain *n-cycles* with $n > 3$. [1] This statement is easily proved by constructing the joint probability for such a graph.

Proof. If there are only 3-cycles, it is always possible to construct a global joint probability distribution from the joint distributions of compatible observables as follows: first multiply together the joint distribution for every 3-cycle [2] $f(a_i, a_j, a_k)$, then multiply those by the joint distributions for all the other edges, and if an observable is present in more than one distribution divide by $f(a_l)^{n-1}$, where n is the number of terms in which it appears. For example, for the first graph in figure 1 we have:

$$f(a_1, \ldots, a_6) = \frac{f(a_1, a_2, a_3) f(a_2, a_6) f(a_3, a_4) f(a_4, a_5)}{f(a_2) f(a_3) f(a_4)}$$

from which all the joint distributions for every edge can be obtained by summing over the other variables. □

So for example it is certainly possible to construct such a distribution for graph 1a, but not for graphs 1b and 1c.

While this gives no information about the existence of a joint distribution for other graphs, the next two theorems will prove that a similar construction is impossible in general.

[1]The definition of *n-cycle*, and other terms related to graph theory, is given in the appendix A.

2.2 Gleason's theorem

In the course of an analysis on the axioms of quantum theory, Gleason proved an important theorem that poses a limit to form that the distribution of an hypothetical HVT could have.

As we said in quantum theory every yes-no test about a system can be represented by a projector P in a Hilbert space $\mathcal{H}$. A *quantum probability distribution* is a function $f(P)$ that assign to every such projector a value such that:

$$\begin{aligned} & f(P) > 0 \\ & \sum_i f(P_i) = 1 \qquad\qquad \text{for every orthogonal base} \quad \{P_i\} \end{aligned}$$

What Gleason proved is that every quantum probability distribution f on an Hilbert space of dimension greater than three can be written as:

$$f(\mathbf{P}) = \mathrm{Tr}(\rho P) \tag{1}$$

for a suitable positive semi-definite, hermitian operator ρ, such that $\mathrm{Tr}(\rho) = 1$. This means that the density matrix formalism is enough to describe the most general kind of quantum probability distribution in an Hilbert space.

Proof. To do. □

With this theorem is possible to deduce quantum theory from these three postulates:

1. dichotomic tests are represented by projectors in a Hilbert space.
2. Compatible test are represented by commuting projectors.
3. If P and Q are orthogonal projector then

$$\langle P + Q \rangle = \langle P \rangle + \langle Q \rangle$$

For a non-contextual HVT to exist, it has to be possible then to define a function $f(P)$ with values 0 or 1 for every projector P in $\mathcal{H}$. Such a function cannot be written in the form (1), required by Gleason theorem for a probability distribution in $\mathcal{H}$, so for a HVT we are forced to abandon axiom (3). But axiom (3) is exactly the one that guarantees the no-disturbance principle. If it wasn't true we could in fact choose four projectors P, Q, R, S such that

$$P + Q = R + S \quad \text{but} \quad \langle P \rangle + \langle Q \rangle \neq \langle R \rangle + \langle S \rangle$$

We can then choose another projector T such that P, Q, T and R, S, T are orthogonal bases, so that they represent a complete test, and it must be true that:

$$\begin{aligned} \langle T \rangle &= 1 - \langle P \rangle - \langle Q \rangle \\ \langle T \rangle &= 1 - \langle R \rangle - \langle S \rangle \end{aligned}$$

but given that $\langle P \rangle + \langle Q \rangle \neq \langle R \rangle + \langle S \rangle$ the value of $\langle T \rangle$ depends on the choice of measuring P and Q or R and S, thus violating the *no-disturbance* principle.

2.3 Kochen-Specker theorem

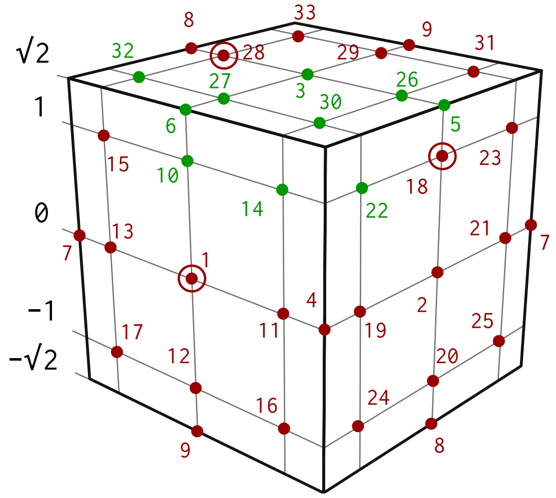
This fundamental incompatibility is better expressed by the Kochen-Specker theorem.

This theorem affirms that in an Hilbert space $\mathcal{H}$ of dimension $d \geqslant 3$, there is no function h from the set of projectors of $\mathcal{H}$ to $\{0, 1\}$ such that:

$$\sum_i P_i = \mathbf{I} \implies \sum_i h(P_i) = 1 \tag{2}$$

The proofs of this theorem are usually constructive: they present a set of projectors and show how such a function h cannot be defined. The original proof used 117 projectors in a three dimensional space, but smaller sets exists. For example this proof, due to A. Peres uses only 33 rays:

Proof. The 33 rays in question are shown in figure and listed in the table below. We will show that is impossible to assign value 1 or 0 to these rays without violating (2). To better visualize the argument we will use the colors *green* and *red* to represent the values 1 and 0 respectively.



Peres KS set

1	$(1, 0, 0)$	12	$(\sqrt{2}, 0, -1)$	23	$(-1, \sqrt{2}, 1)$
2	$(0, 1, 0)$	13	$(\sqrt{2}, -1, 0)$	24	$(1, \sqrt{2}, -1)$
3	$(0, 0, 1)$	14	$(\sqrt{2}, 1, 1)$	25	$(-1, \sqrt{2}, -1)$
4	$(1, 1, 0)$	15	$(\sqrt{2}, -1, 1)$	26	$(0, 1, \sqrt{2})$
5	$(0, 1, 1)$	16	$(\sqrt{2}, 1, -1)$	27	$(1, 0, \sqrt{2})$
6	$(1, 0, 1)$	17	$(\sqrt{2}, -1, -1)$	28	$(0, -1, \sqrt{2})$
7	$(-1, 1, 0)$	18	$(0, \sqrt{2}, 1)$	29	$(-1, 0, \sqrt{2})$
8	$(0, -1, 1)$	19	$(1, \sqrt{2}, 0)$	30	$(1, 1, \sqrt{2})$
9	$(1, 0, -1)$	20	$(0, \sqrt{2}, -1)$	31	$(-1, 1, \sqrt{2})$
10	$(\sqrt{2}, 0, 1)$	21	$(-1, \sqrt{2}, 0)$	32	$(1, -1, \sqrt{2})$
11	$(\sqrt{2}, 1, 0)$	22	$(1, \sqrt{2}, 1)$	33	$(-1, -1, \sqrt{2})$

As we can see the system is symmetric under the interchange or the reversal of the axes x, y, z. This means that without losing generality we can paint green any of the rays in the orthogonal triad $\{1, 2, 3\}$. Similarly in $\{1, 5, 8\}$ we can choose between 5 or 8 (reversal of the y axis), between 6 or 9 in $\{2, 6, 9\}$ (reversal of the x axis) and between 31 or 32 in $\{4, 31, 32\}$ (interchange of x with y).

If we choose for example to paint $3, 5, 6$ and 32 green, the value of the other rays will be fixed by the procedure shown below:

$$
\begin{array}{lllll}
3 & \Longrightarrow & 1,2,4,7,11,13,19,21 & & \\
5 & \Longrightarrow & 1,8,15,16 & & \\
6 & \Longrightarrow & 2,9,23,24 & & \\
32 & \Longrightarrow & 4,12,18,31 & & \\
\hline
12,2 & \Longrightarrow & 27 & \Longrightarrow & 17 \\
8,17 & \Longrightarrow & 14 & \Longrightarrow & 29 \\
29,2 & \Longrightarrow & 10 & \Longrightarrow & 33 \\
33,7 & \Longrightarrow & 30 & \Longrightarrow & 20 \\
20,1 & \Longrightarrow & 26 & \Longrightarrow & 25 \\
25,8 & \Longrightarrow & 22 & \Longrightarrow & 28
\end{array}
$$

But with this coloring the rays $\{1, 18, 28\}$, which form an orthogonal triad, are all red: we have a contradiction, and since the initial assignments where completely general the theorem is proved. □

Clearly this means that an HVT, that assigns definite values 0 or 1 to every yes-no test regardless of the context, is fundamentally incompatible with QM. In fact every such assignment has to satisfy (2), since a set of projectors that sums to identity represent a complete test, so the assigned values must sum to unity. But we have just that proved that this is not possible in general for $d \geqslant 3$.

Other kinds of proofs exists which relies on general observables (not just projectors), like the Peres-Mermin square described in 3.3, or on inequalities as shown in 3.4.

2.4 Experimental Test of KS theorem

What we have shown is just that the QM predictions are impossible to reproduce with a classically behaving HVT. The next obvious step would be to see which way nature prefers by testing it experimentally. While it is possible to test the Kochen-Specker directly using the proof presented in the last section (or other equivalent proofs), this approach is known to be problematic.

One of the biggest problem is that every observable must be measured in more than one, incompatible contexts. So one could not discard the possibility that the results are influenced by the different measurement's procedure. The inequalities, described in the next section, offer a better alternative for experimental verification.

Another difficulty lies in the precision with which we can measure an observable: there is always the chance that our measured observable could differ slightly from the one we wanted to measure, this is called the *finite precision problem*. Since it can be shown that it is possible to approximate a KS set of rays with a non KS one, this is a serious problem in testing the KS theorem. The way out of it is to create new versions of the KS theorem for imprecisely defined observables, often expressed as the already mentioned inequalities.

The impossibility to precisely define an observable leads also to the *problem of compatibility*: we cannot assume that the observable are compatible anymore, and as we have seen, this is a crucial assumption for testing contextuality. To deal with this problem, we need to add new terms to the classical bound in inequalities, as will be described in the next section.

3 Non-contextual inequalities

In some systems, assuming classical behavior implies that certain combinations of correlations between observables are bounded by an inequality that is violated QM. This suggests another way to prove and to test the fundamental incompatibility between QM and non-contextual theories.

The most famous of these inequality is certainly the *Bell theorem*, which showed how QM cannot be described by a local HVT. Since non-locality is a particular form of contextuality, the same inequality can be used to settle the question of the existence of non-contextual HVT, and is described in section 3.1 (in the form of the CHSH inequality).

Anyway, in the more general setting of contextuality we don't necessarily need two subsystems or entanglement, as we showed in the KS theorem, a spin-1 system is enough.

It is useful to divide non-contextual inequalities into two classes:

- *State-dependent*, where the inequality is violated only when the system is in some particular state.
- *State-independent*, where the inequality is always violated regardless of the state of the system.

Usually state-dependent tests involve less measurements, but obviously need a fixed initial state, while state-independent does not.

As in section 2.1, we can associate a graph to every non-contextual inequality. There are two ways to construct a graph for an inequality test:

- The *Compatibility graph* is the graph constructed from the set of observables: in such a graph every observable is represented by a vertex and two vertex are connected if the corresponding observables are compatible.
- The *Exclusivity graph* constructed from the single dichotomic exclusive events, where every vertex represent a yes-no measurement (a projector in the quantum formalism), and edges represent exclusivity (orthogonality in QM).

In the next sections we will describe some of the most common non-contextual inequalities.

3.1 CHSH

This inequality was first devised by Clauser, Horne, Shimony and Holt as an alternative to the Bell inequality to test non-locality. For this reason the system is composed by two 2-dimensional subsystems (normally spin-$\frac{1}{2}$ particles) in entanglement, that can be separated by a space-like distance.

The experimenters can measure the spin on their particles in two different directions. The chosen observable are:

$$A_1 = \sigma_x \otimes \mathbf{I} \qquad B_1 = \sigma_z \otimes \mathbf{I}$$
$$A_2 = -\mathbf{I} \otimes \frac{\sigma_z + \sigma_x}{\sqrt{2}} \qquad B_2 = \mathbf{I} \otimes \frac{\sigma_z - \sigma_x}{\sqrt{2}}$$

and the corresponding compatibility graph is shown in figure 2a.

Since the observables measured can takes only ± 1 values if we suppose a classical, non-contextual(or local) behavior we expect:

$$\langle (A_1 + B_1)A_2 + (B_1 - A_1)B_2 \rangle = \langle A_1A_2 + B_1A_2 + B_1B_2 - A_1B_2 \rangle \leqslant 2 \tag{3}$$

But if, say, we prepare the system in the entangled state

$$|\psi\rangle = \frac{|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle}{\sqrt{2}}$$

QM predicts:

$$\langle A_1A_2 \rangle + \langle B_1A_2 \rangle + \langle B_1B_2 \rangle - \langle A_1B_2 \rangle = 2\sqrt{2}$$

so (3) is violated by quantum mechanics. Notice that for the inequality to be violated the system has to be in a particular entangled state, so this is an example of a state-dependent inequality.

We can also write (3) using directly probabilities for exclusive events:

$$p(11|A_1A_2) + p(-1-1|A_1A_2) + p(11|B_1B_2) + p(-1-1|B_1B_2) + \\ p(11|B_1A_2) + p(-1-1|B_1A_2) + p(-11|A_1B_2) + p(1-1|A_1B_2) \leqslant 3$$

where $p(ab|AB)$ represents the probability of having oucomes a and b measuring the properties A and B. The events involved are represented in the exclusivity graph 2b.

It can be shown that in QM the maximum value of this sum is $2+\sqrt{2}$, so the inequality is again violated.

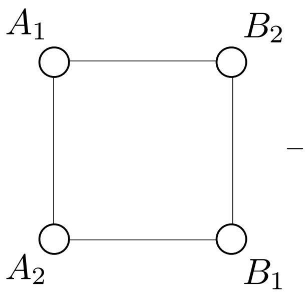
3.2 KCBS

A more recent example of a state-dependent non-contextual inequality is called the KCBS inequality, from Klyachko, Can, Binicioglu and Shumovsky.

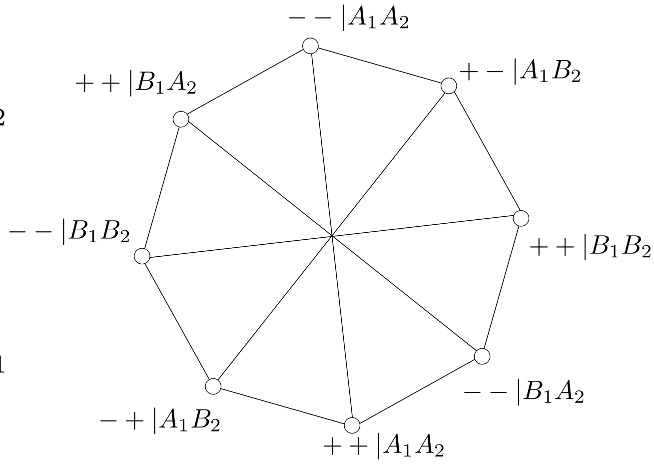
Here we have five dichotomic observables which can take values $\{0, 1\}$ that satisfy the exclusivity relations shown in figure 3.

Classically, since only 2 observables can be assigned the value 1 to satisfy the exclusivity graph, we have the inequality.

$$\sum_{i=1}^{5} a_i \leqslant 2 \tag{4}$$



(a) Compatibility graph



(b) Exclusivity graph

Figure 2: Graphs for the CHSH inequality

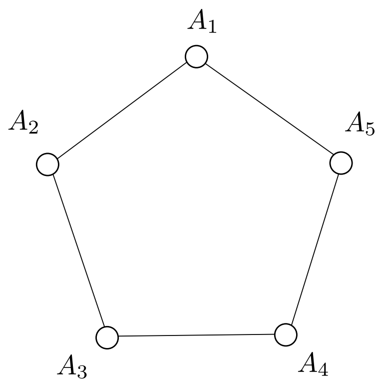


Figure 3: Exclusivity graph for the *KCBS* inequality

This inequality is often presented in another form, using the observables $B_i = 2A_i - 1$ instead that takes the values $\{1, -1\}$, so that (4) becomes:

$$\langle B_1 B_2 \rangle + \langle B_2 B_3 \rangle + \langle B_3 B_4 \rangle + \langle B_4 B_5 \rangle + \langle B_5 B_1 \rangle \geqslant -3 \tag{5}$$

In QM the $A_i = |a_i\rangle \langle a_i|$ are projectors in a 3-dimensional Hilbert space, with orthogonality relations given by the graph 3. This can be implemented measuring $A_i = 1 - S_i^2$, where the S_i are the spin observables for a spin-1 particle in five different direction. For example we can use the rays:

$$|a_1\rangle = \left(1, 0, \sqrt{\cos(\pi/5)}\right)$$
$$|a_2\rangle = \left(-\cos(\pi/5), \sin(\pi/5), \sqrt{\cos(\pi/5)}\right)$$
$$|a_3\rangle = \left(-\cos(\pi/5), -\sin(\pi/5), \sqrt{\cos(\pi/5)}\right)$$
$$|a_4\rangle = \left(-\cos(2\pi/5), \sin(2\pi/5), \sqrt{\cos(\pi/5)}\right)$$
$$|a_4\rangle = \left(-\cos(2\pi/5), -\sin(2\pi/5), \sqrt{\cos(\pi/5)}\right)$$

directed along the vertices of a pentagram as in figure 4. If we measure them on

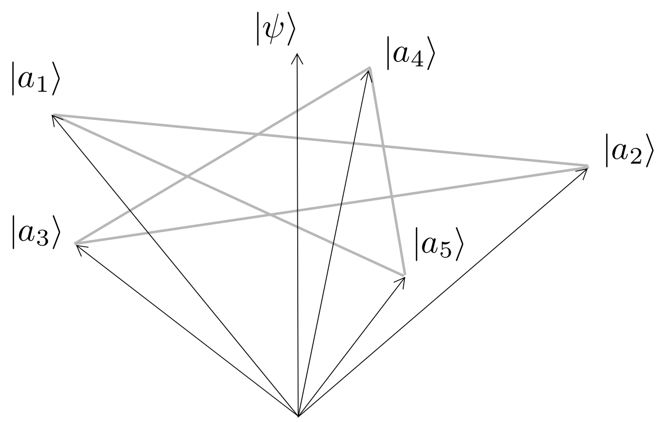


Figure 4: Rays corresponding to the projectors $A_1, \dots, A_5$

the state $|\psi\rangle = (1, 0, 0)$ we have

$$\sum_{i=1}^{5} |\langle a_i \,|\, \psi\rangle|^2 = \sqrt{5} > 2$$

which correspond to a violation of $5 - 4\sqrt{5}$ for the inequality (5).

3.3 Peres-Mermin square

This state-independent inequality has its origin in another proof of the Kochen-Specker theorem.

Suppose to have a system similar to the one used for the CHSH inequality: two particles of spin-$\frac{1}{2}$, and consider the observables:

$$\begin{array}{lll} A_{11} = \mathbf{I} \otimes \sigma_z & A_{12} = \sigma_z \otimes \mathbf{I} & A_{13} = \sigma_z \otimes \sigma_z \\ A_{21} = \sigma_x \otimes \mathbf{I} & A_{22} = \mathbf{I} \otimes \sigma_x & A_{23} = \sigma_x \otimes \sigma_x \\ A_{31} = \sigma_x \otimes \sigma_z & A_{32} = \sigma_z \otimes \sigma_x & A_{33} = \sigma_y \otimes \sigma_y \end{array} \tag{6}$$

The last element of every row is the product of the other two, and the same is true for the columns, except for the last one since:

$$(\sigma_z \otimes \sigma_z)(\sigma_x \otimes \sigma_x) = -\sigma_y \otimes \sigma_y$$

For this reason a fixed assignment of values $\{1, -1\}$ to every observable in (6) is not possible, and this rules out the possibility of having a non-contextual HVT to explain those measures.

This can also be rephrased in the form of an inequality:

$$\begin{aligned}\langle A_{11}A_{12}A_{13}\rangle + \langle A_{21}A_{22}A_{23}\rangle + \langle A_{31}A_{32}A_{33}\rangle + \\ + \langle A_{11}A_{21}A_{31}\rangle + \langle A_{12}A_{22}A_{32}\rangle - \langle A_{13}A_{23}A_{33}\rangle \leqslant 4\end{aligned} \tag{7}$$

Which holds classically, since the third element in every correlation is the product of the other two.

But in QM we have:

$$\begin{aligned}\langle A_{11}A_{12}A_{13}\rangle + \langle A_{21}A_{22}A_{23}\rangle + \langle A_{31}A_{32}A_{33}\rangle + \\ + \langle A_{11}A_{21}A_{31}\rangle + \langle A_{12}A_{22}A_{32}\rangle - \langle A_{13}A_{23}A_{33}\rangle = 6\end{aligned}$$

so the inequality (7) is violated for every state of the system.

3.4 Yu-Oh's 13 rays inequality

The proofs of the Kochen-Specker theorem we have seen are all based on the impossibility of assigning definite value to a groups of observable.

This proof involves instead a set of 13 rays, where KS-type assignments are possible, but every such assignment satisfy an inequality violated by QM.

The rays used in the proof are:

$$\begin{array}{llll} a_1 = (1,0,0) & a_2 = (0,1,0) & a_3 = (0,0,1) & \\ b_1 = (0,1,1) & b_2 = (1,0,1) & b_3 = (1,1,0) & \\ c_1 = (0,1,-1) & c_2 = (1,0,-1) & c_3 = (1,-1,0) & \\ d_1 = (-1,1,1) & d_2 = (1,-1,1) & d_3 = (1,1,-1) & d_4 = (1,1,1) \end{array} \tag{8}$$

and they satisfy the exclusivity conditions showed in the graph in figure 5.

If we try to assign values $\{0, 1\}$ to these rays we notice that only one of the rays d_i can be assigned the value 1.

Proof. Like we did in section 2.3, let's use the colours green and red instead the values 1 and 0 respectively. We first show that if d_4 is green, none of the other d_i can be green.

If for instance d_4 and d_1 are green then b_3, c_3 and b_2, c_2 must be red, which means that a_2 and a_3 have to be green but this is not allowed. The same is true for d_2 and d_3.

Moreover only one between d_1, d_2, d_3 can be green at once, for example if d_1 and d_2 were both green, then b_1, c_1 and b_2, c_2 have to be red which

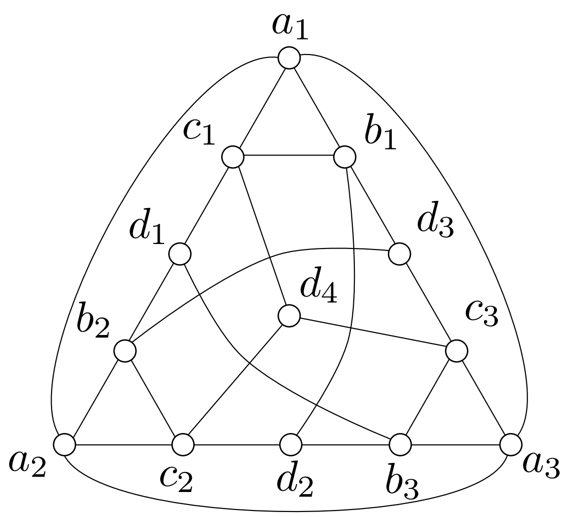


Figure 5: Exclusivity graph for the 13 rays

means that a_1 and a_2 must be green and this is again not possible. For symmetry reason, the same is true for d_2, d_3 and d_1, d_3, so we must conclude that only one between $\{d_1, d_2, d_3, d_4\}$ can be green in every assignment. $\square$

From this we have that every non-contextual HVT must obey the inequality

$$\sum_{i=1}^{4} \langle d_i \rangle \leqslant 1$$

In QM instead we have:

$$\sum_{i=1}^{4} d_i = \frac{4}{3}\mathbf{I} \quad \Longrightarrow \quad \sum_{i=1}^{4} \langle d_i \rangle = \frac{4}{3}$$

and no non-contextual HVT can reproduce this result.

If we use the observable $A_i = 2\mathbf{I} - P_i$, where P_i are the projectors associated with the rays (8), that takes the values $\{1, -1\}$, we can write the inequality

$$\sum_i \langle A_i \rangle - \frac{1}{4} \sum_{(i,j)} \langle A_i A_j \rangle \leqslant 8 \tag{9}$$

where the sum in the second term is extended only to i, j connected by an edge in the graph 5. As it can easily be proved in QM we have:

$$\sum_i A_i - \frac{1}{4} \sum_{(i,j)} A_i A_j = \frac{25}{3}\mathbf{I}$$

so the inequality (9) is violated for every state.

3.5 Graph theoretical approach

The *exclusivity* graph is a powerful instrument to analyze inequalities. In fact from the graph, we can easily deduce if the inequality is violated by QM, and how much is the violation.

As we did in section 3.1 for the CHSH inequality, every non-contextual inequality can be expressed as a positive linear combination of probabilities:

$$A = \sum_i w_i P_i$$

To every such a linear combination we can associate an exclusivity graph as described at the beginning of this section[3], where now the dichotomic observable A_i that takes values $\{0, 1\}$, is the one associated with the ith event, so that $\langle A_i \rangle = P_i$. Once we construct this graph we can derive the inequality and its violation in QM (if present) using the following theorems.

Theorem 3.1. *Given a graph* G *associated with a non-contextual inequality* A*, the maximum value attainable by a non-contextual* HVT *is equivalent to the independence number* $\alpha(G)$ *of the graph while the maximum for quantum theory is the Lovàsz number* $\theta(G)$ *of the graph, and*

$$\alpha(G) \leqslant \theta(G)$$

Proof. In a non-contextual HVT, being deterministic, each events has probability 0 or 1. The maximum of A is attained when the value 1 is assigned to the maximum possible number of events, A in this case corresponds to the cardinality of the largest set of independent vertices, that is *independence number* $\alpha(G)$ of the graph.

In QM every event corresponds to a unitary ray in a given Hilbert space, and connected vertices correspond to orthogonal rays, this is exactly an *orthonormal representation* of the graph G. Since the probabilities are $P_i = |\langle a_i|\psi\rangle|$ for a given $|\psi\rangle$, the maximum of A is clearly given by the *Lovàsz number* of the graph G as defined in (10).

The inequality follows directly by the Lovàsz's sandwich theorem A.2. □

Notice that $\alpha(G)$ is only an upper bound for the maximum value attainable for a non-contextual HVT, since it does not impose the condition that in every complete test the probabilities sums to 1. For example for the graph in figure 5 the independence number is $\alpha(G) = 5$, but every independent set with that number of vertices is one in which all the vertices d_i are included (i.e. have value 1) which as we said is impossible, at least when is implemented in dimension $d = 3$.

Theorem 3.2. *The inequality associated with the graph* G *is violated if and only if* G *contain as a subgraph a cycle graph* C_n *of* $n \geqslant 5$ *or its complement.*

[3]The graph is weighted if the coefficients $w_i \neq 1$, the following theorems are valid also in this case.

Proof. Indeed if the graph have got no such odd cycles or their complement then is perfect, by the strong perfect graph theorem A.1. In this case we have

$$\alpha(G) = \omega(\bar{G}) = \chi(\bar{G})$$

so for the sandwich theorem A.2

$$\alpha(G) = \theta(G)$$

and we have no violation. □

A Basic glossary of graph theory

In this small section we define all the terms connected to graph theory used in the text.

Definition A.1 (Graph)**.** A graph $G(V, E)$ is a collection of elements called *vertices* $v \in V$, and *edges* $e \in E$, where every edge is associated with two vertex, which are said to be *connected* by the edge.

A graph is usually represented as a series of dots, that represent vertices, connected by lines, that represent edges.

A graph is called *complete* when every vertex is connected to every other vertex in the graph.

Definition A.2 (Cycle graph)**.** A cycle graph C_n is a graph with n vertices joint by edges that form a single closed path. So if $V = \{v_1, \ldots, v_n\}$ are the vertices of C_n, the edges are

$$E = \{(v_1, v_2), (v_2, v_3), \ldots, (v_n, v_1)\}$$

For example the graph in figure 3 is the cycle graph with 5 vertex C_5.

Definition A.3 (Subgraph)**.** A subgraph $S(U, F)$ of a graph $G(V, E)$ is a graph that contains a subset of vertex and edges, UV FE of G. An *induced* subgraph of G is a graph that contains a subset of vertex of G and only the edges that join those vertex.

A complete subgraph of a graph is called a *clique*. The number of vertices that constitutes the largest clique in a graph G is called the *clique number* $\omega(G)$ of the graph.

Definition A.4 (Complementary graph)**.** The *complementary graph* $\bar{G}$ of a graph G is a graph with the same vertices as G in which every vertex v_i is connected to v_j if and only if v_i is *not* connected to v_j in G.

For example the complementary graph of a complete graph is a null graph (a graph with no edges).

Definition A.5 (Independence set)**.** An *independent set* is a set of vertex not connected by any edge. The number of vertex contained in the largest independent set of a graph is called the *independence number* $\alpha(G)$ of the graph.

Clearly the clique number of a graph is the independence number of the complementary graph:

$$\omega(G) = \alpha(\bar{G})$$

A vertex coloring of a graph is a way of assigning labels (or "colors") to the vertices of a graph such that no two adjacent vertices has the same one.

Definition A.6 (Chromatic number)**.** The minimum number of color needed for a vertex coloring of a graph G is called the *chromatic number* $\chi(G)$ of the graph.

A graph G where

$$\omega(G) = \chi(G)$$

for every induced subgraph is called a *perfect graph*

Theorem A.1 (Strong perfect graph theorem)**.** *A graph is perfect if and only if it does not contain odd cycles C_n with $n \geqslant 5$ or their complements as induced subgraphs*

Definition A.7 (Orthogonal representation)**.** An *orthogonal representation* of a graph $G(V, E)$ assigns to every vertex $v_i \in V$ a vector $u_i \in \mathbb{R}^d$ such that $u_i \cdot u_j = 0$ when $v_i i$ and v_j are connected.

If the vectors u_i are also of unit length the representation is called *orthonormal*.

The above definition does not require to assign different vertices to different vectors, an orthogonal representation with that property is said to be *faithful*.

Definition A.8 (Lovàsz number)**.** Given a graph $G(V, E)$ the *Lovàsz number* is defined as

$$\theta(G) = \max \sum_{v_i \in V} |h \cdot u_i|^2 \tag{10}$$

where the maximum is taken over all the orthonormal representations of G with vectors u_i, and all the vectors $h \in \mathbb{R}^d$.

Theorem A.2 (Lovàsz's sandwich theorem)**.** *The Lovàsz number always lies between the independence number $\alpha(G)$ and the chromatic number of the complementary graph $\chi(\bar{G})$*

$$\alpha(G) \leqslant \theta(G) \leqslant \chi(\bar{G})$$

<++>