Quantum Information Theory: Lecture 4

March 20, 2017

1 Bell's theorem

In the previous lectures, we studied quantum entanglement and some of its applications for quantum information processing. The counter-intuitive aspect of this concept was discussed from a qualitative point of view. Here we formalize this intuition, and present the most striking demonstration of the nonlocal character of quantum mechanics: the correlations of local measurements performed on distant entangled particles. Following the approach initiated by John Bell in 1964, we will see that such correlations are so strong that they can in fact not be reproduced by any classical means. Thus we are lead to the conclusion that quantum mechanics is a nonlocal theory. This is the content of Bell's celebrated theorem.

1.1 Bell's inequality

We consider an experiment featuring two distant observers, Alice and Bob. A source, placed in-between A and B emits pairs of particles, one after the other. For each pair, one particle from the pair is sent to Alice, the other to Bob. Alice and Bob have measuring devices, with the help of which they can perform a measurement on each incoming particle.

In each run of the experiment, both Alice and Bob can choose between two possible measurements to perform. We denote the choice of measurement $i = 1, 2$ for Alice, and $j = 1, 2$ for Bob. The result of each measurement is binary. We note it $A_i = \pm 1$ for Alice, and $B_j = \pm 1$ for Bob.

Let's imagine now that Alice and Bob run the above experiment many times. In each round, each observer records which measurement she/he chose to perform, and which outcome she/he obtained. When the experiment is finished, both observers meet, and compare their results. Surprisingly, they notice that their data is correlated. Say, for instance, when both observers chose the same measurement, they obtained the same result with a probability higher than 1/2. Puzzled by this behavior, Alice and Bob try to find an explanation: how were these correlations generated?

The first explanation that comes to Alice's mind is that one device influenced the other. Imagine for instance that Alice's device upon receiving its input, sent a classical signal to Bob's device saying which input it received and which outcome it produced. Bob protests however, noticing that during the experiment Alice and Bob were space-like separated. Hence no classical communication (traveling at the speed of light) would have had enough time to travel from Alice to Bob. Thus Bob concludes (correctly) that the observed correlations cannot be explained by a classical signal.

Alice has then another idea. The correlated behavior can certainly be explained by the fact that A and B performed measurements on a pair of particles originating from a common source. Hence, the particles may have a common physical property (for instance they may have the same polarization) which would explain correlations in the measurements. Loosely speaking, one may imagine that the particles have been 'programmed' in the same way. They follow a common strategy.

Now comes the crucial question. Can we test whether the above mechanism is able to explain the observed correlations? A priori, this may seem difficult, since there exist many possible ways to prepare (or to 'program') the two particles. How can we be sure to take all possibilities into account? Remarkably, Bell showed that in any local theory, that is, where the correlations originate from a common program (or a common strategy), the correlations between distant events must satisfy certain constraints. Importantly, these constraints, which are formalized as inequalities (hence called Bell inequalities), are always satisfied in any local model, no matter how complicated this model is.

We will see now how to derive such a Bell inequality. Consider again the above experiment. A possible strategy, denoted λ, consists in attributing values to the outcomes of each measurement, that is, to define variables $A_1 = \pm 1$, $A_2 = \pm 1$, $B_1 = \pm 1$, and $A_2 = \pm 1$. It is easy to see that, no matter which values we choose, the following quantity:

$$S(\lambda) = A_1(B_1 + B_2) + A_2(B_1 - B_2) \quad (1)$$

takes values $S(\lambda) = \pm 2$. Note that we cannot measure $S(\lambda)$, since in each run of the experiment, A and B choose to perform either the first or the second measurement, but they cannot perform both at the same time.

Nevertheless, what can be measured is the average value of $S(\lambda)$ over many runs of the experiment. Indeed, we cannot exclude that the the strategy λ changes in each run, according a distribution $\rho(\lambda)$. Hence we must average over λ. We have that

$$S = \int d\lambda \rho(\lambda) S(\lambda) \quad (2)$$

where $\int d\lambda \rho(\lambda) = 1$. From equation (1), we get directly the inequality

$$|S| = |\langle A_1 B_1 \rangle + \langle A_1 B_2 \rangle + \langle A_2 B_1 \rangle - \langle A_2 B_2 \rangle| \leq 2. \quad (3)$$

This is Bell's inequality in its simplest form, which was derived in 1969 by Clauser, Horne, Shimony, and Holt. This inequality is thus referred to as the CHSH Bell inequality.

1.2 Bell locality

It is worth discussing in more details the assumptions that lie behind Bell's inequality. The main idea that is being tested here, is whether or not the observed correlations can be explained by a common cause. In order to formalize this idea, we introduce a variable λ that is shared between A and B. Imagine, that in the source, the particles are programmed in a specific manner. For instance, each particle may carry a program that indicates what outcome will occur for every possible measurements of Alice and Bob. This program is represented by the classical variable λ, which historically people refer to as a local hidden variable. Importantly, λ represents the most general classical program. That is, all possible classical strategies are taken into account here.

Now, it turns out that correlation that can be obtained via a shared variable λ can always be decomposed in a particular way. Consider an experiment as that described in the previous section, where Alice's measurement choice is denoted x, and its outcome is a. Similarly, we denote Bob's measurement choice by y, and its outcome by b. Note that x, y, a, b are not necessarily binary anymore. Then, if the correlations originate from a shared variable λ, they can always be decomposed in the following form

$$p(a, b|x, y) = \int d\lambda \mu(\lambda) p(a|x, \lambda) p(b|y, \lambda), \tag{4}$$

where $\mu(\lambda)$ represents the probability of having a certain λ, and we thus have that $\int d\lambda \mu(\lambda) = 1$. This means that on Alice's side, the outcome a depends on the choice of measurement x and on the shared variable λ, but not on Bob's measurement choice y. Similarly, on Bob's side, the outcome b depends only on y and λ. This type of correlations are referred to as local, in the sense of Bell; equation (4) captures the notion of Bell locality. When testing whether certain observed correlations can be explained by a common cause (i.e. by a shared classical program, or variable, λ), we test whether or not they admit a decomposition of the form (4). If such a decomposition can be found, we say that the correlations are local (or Bell local). If no decomposition of the form (4) exists, we say that the correlations are nonlocal (in the sense of Bell).

Nonlocality can be detected via Bell inequalities. We have discussed the simplest of such inequalities above, the CHSH Bell inequality. There exist however many other Bell inequalities (in fact an infinite number of them). Importantly, all correlations which are Bell local, i.e. of the form (4), will satisfy *all* Bell inequalities. Correlations which are nonlocal, will violate at least one Bell inequality.

1.3 Quantum nonlocality

Previously we have derived the CHSH Bell inequality (3) which holds for any physical model in which the particles obey a predetermined strategy.

We will see now that this inequality can be violated in quantum mechanics. That is, if we perform judiciously chosen measurements on a pair of entangled particles, we will get $S > 2$!

Let's imagine that the source described in the above experiment produces pairs of photons which are entangled in polarization. More precisely the state of the photon pair is given by

$$|\psi_-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle). \tag{5}$$

The measurement devices of Alice and Bob then perform measurements of the polarization characterized by vectors of the Bloch sphere. For a measurement direction $\hat{a}_i$, the corresponding observable is $A_i = \hat{a}_i \cdot \vec{\sigma}$, where $\vec{\sigma}$ is the vector containing the 3 Pauli matrices. The first measurements of Alice is given by $\hat{z}$, and the second by $\hat{x}$. For Bob, the first measurements is given by $(\hat{z} + \hat{x})/\sqrt{2}$, and the second by $(\hat{z} - \hat{x})/\sqrt{2}$.

We can now evaluate the quantity S. We have that

$$\langle A_i B_j \rangle = \langle \psi_- | A_i \otimes B_j | \psi_- \rangle. \tag{6}$$

With the above choice of observables, we get $\langle A_1 B_1 \rangle = \langle A_1 B_2 \rangle = \langle A_2 B_1 \rangle = -\langle A_2 B_2 \rangle = -\frac{1}{\sqrt{2}}$, and thus

$$S = 2\sqrt{2} \approx 2.83 > 2 \tag{7}$$

Therefore we see that quantum predictions violate Bell's inequality. We are forced to conclude that these quantum correlations are nonlocal. No classical mechanism can explain these correlations. On the one hand, a classical signal is excluded, since the particles can be put in a space-like separated configuration. On the other hand, a common cause is excluded via Bell's inequality violation. That is, the correlations are nonlocal in the sense of Bell, and cannot be decomposed in the form (4). This phenomenon is termed quantum nonlocality.

Following Bell's discovery, few pioneering physicists underwent the task of testing experimentally Bell's inequality. This became possible thank to immense progress in quantum optics. In the early 1970s, pairs of photons entangled in polarization could be experimentally produced, and the polarization of single photons could be measured using polarizers and single photon detectors. The main experimental confirmation came in 1982, when Alain Aspect and his team in Paris demonstrated violations of the CHSH Bell inequality with 5 standard deviations. Since then, numerous Bell experiments have been performed, using various physical systems, such as photons, atoms, and superconducting qubits. Remarkably all of these experiments demonstrated Bell inequality violations, and excellent agreement with quantum predictions. This represents very strong evidence that Nature is inherently nonlocal.

1.4 Maximal violation of the Bell's inequality in quantum theory

In our example above, we have seen that by performing measurements on a pair of entangled particles, we could violate the CHSH Bell inequality. In particular, we obtained $S = 2\sqrt{2}$. A natural question is whether this is the largest possible violation in quantum mechanics, or if one could achieve a larger value by choosing another entangled states and/or other measurements. We will see now that this is not the case. The value of $2\sqrt{2}$ represents the quantum limit for the CHSH Bell inequality. In other words, for any quantum state and measurements, we have that $S \leq 2\sqrt{2}$, which is known as Tsirelson's bound for the CHSH inequality.

This bound follows from the Hilbert space structure of quantum mechanics. It can be demonstrated as follows. Note first that the quantum value for the CHSH expression of equation (3) can be written as $\mathrm{tr}(\rho\mathcal{B})$, where $\rho \in H^d \otimes H^d$ is an arbitrary bipartite quantum state, and where we have defined the Bell operator

$$\begin{aligned} \mathcal{B} &= A_1 \otimes B_1 + A_1 \otimes B_2 + A_2 \otimes B_1 - A_2 \otimes B_2 && (8) \\ &= A_1 \otimes (B_1 + B_2) + A_2 \otimes (B_1 - B_2). && (9) \end{aligned}$$

Note that the observables A_i and B_j are self-adjoint operators with eigenvalues ± 1 acting on H^d. We are now interested in the largest eigenvalue of $\mathcal{B}$, which is given by the operator norm $||\mathcal{B}||$ (or spectral norm). It is convenient to first consider the quantity $||\mathcal{B}^2||$. It is straightforward to show that

$$||\mathcal{B}^2|| \leq 4 + ||[A_1, A_2]||\ ||[B_1, B_2]||. \tag{10}$$

Then, using the relation $||[A_1, A_2]|| \leq 2||A_1||\ ||A_2|| \leq 2$ (since A_i has eigenvalues ± 1), we get that $||\mathcal{B}^2|| \leq 8$, hence that $||\mathcal{B}|| \leq 2\sqrt{2}$, which finishes the proof.

1.5 Greenberger-Horne-Zeilinger paradox

It is interesting to note that quantum nonlocality can in fact be demonstrated without inequalities. Hence the demonstration is in some sense simpler than the CHSH inequality that we have seen above. The price to pay is that we need to consider a situation involving three observers.

Consider Alice, Bob, and Charlie, sharing the quantum state

$$|GHZ\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle) \tag{11}$$

Each observer can now perform two possible measurements on his/her particle. The first measurement A_1 (and similarly for B_1 and C_1) is represented by the observable σ_x, the second A_2 etc by the observable σ_y. It is straightforward to check that for the GHZ state one has that

$$A_1 B_1 C_1 = +1 \tag{12}$$

$$A_1 B_2 C_2 = -1 \tag{13}$$

$$A_2 B_1 C_2 = -1 \tag{14}$$

$$A_2 B_2 C_1 = -1 \tag{15}$$

Let us now see that no a local model can explain the above behavior. As we discussed above for the CHSH Bell inequality, the most general local model attributes a certain value to all observables. Hence we must specify $A_1 = \pm 1$, $A_2 = \pm 1$, $B_1 = \pm 1$ etc. Now, it is easy to see for any possible choice, we will have a contradiction with the above statistics. Take for instance the product of all left-hand side terms of the above 4 equations. This product is necessarily equal to +1, since $A_1^2 = A_2^2 = B_1^2 = ... = 1$. However, the product of all right hand side terms is equal to -1. Hence, the above quantum predictions are in full contradiction with those of any local model.