Bell's theorem and the connection to cryptography

Roger Colbeck

(Dated: July 21, 2011)

Please report any errors/typos to me at rcolbeck@perimeterinstitute.ca.

I. INTRODUCTION

In this course we will examine in detail one of the most fascinating departures from classical theory exhibited by our world, so-called non-locality, and connect it to a practical application: device-independent cryptography.

Quantum mechanics is a probabilistic theory: it does not assign precise values to experimental outcomes, but instead prescribes the distribution over the outcomes, *even* with the most complete description of the state within the theory. For example, when a particle in the state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ is measured in the $\{|0\rangle, |1\rangle\}$ basis, according to quantum theory each outcome occurs with probability $\frac{1}{2}$. (Note that, since the state is pure, this is the most complete description we can have about it.)

This is already a stark departure from classical theory which is fundamentally deterministic: any uncertainty we may have about the outcomes of future events is purely due to a lack of knowledge about the initial configuration.

This inherent randomness was widely discussed in the early days of quantum mechanics, and led Einstein, Podolsky and Rosen [1] to question the completeness of the theory. They considered a measurement on one half of a maximally entangled pair whose outcome (according to the theory) allows perfect prediction of the outcome of the analogous measurement on the other half (the outcomes are always anti-correlated). Furthermore, this conclusion holds no matter how far apart the two particles are.

A natural way to explain such correlations, is to imagine that the quantum state is not the most complete description of the system, but that, in fact, some additional shared randomness was given to the particles by the source (this type of additional information is often termed local hidden variable, see later). The anti-correlated outcomes can then be alternatively explained using this shared randomness.

II. NON-LOCALITY

As argued above, perfectly anti-correlated outcomes do not present any mystery. If sleeping beauty awakes one day and sees daylight, she immediately knows it is night on the opposite side of the planet, and there is nothing surprising about that. However, quantum theory does contain mysterious correlations, unexplainable in such a classical way. Such correlations are said to be non-local and are important in the context of Bell's theorem.

A. Example

Before discussing in more detail, here is an example which neatly shows the power of quantum correlations, in the form of a game between three co-operating players and a referee. The three players are going to play the following game. They will be isolated from one another and each will be asked one of two questions by the referee (the first party's question is denoted $A \in \{0, 1\}$, the second party's $B \in \{0, 1\}$ and the third party's $C \in \{0, 1\}$) to which they must answer either $+1$ or -1. The set of questions is picked uniformly from $(A, B, C) = (0, 0, 0)$, $(0, 1, 1)$, $(1, 0, 1)$, and $(1, 1, 0)$. If the set $(0, 0, 0)$ is asked, they win the game if the product of their outputs is -1, while for each of the remaining sets of questions, they win the game if the product of their outcomes is $+1$. The three players know the form of the game and are allowed to meet beforehand to discuss their strategy. However, no communication is allowed once they are isolated at the start of the game.

Let's think about ways to win this game. Imagine that, at the start of the game, the players agree on two values each, one of which is output if they receive input 0, and the other is output if they receive input 1 (let us denote these values x_0, x_1 for the first player y_0, y_1 for the second and z_0, z_1 for the third, so that $x_0 \in \pm 1$ is the output made by the first player if asked question $A = 0$). We call such a strategy an assignment strategy. One possible choice is for all of the values x_0, x_1, y_0, …, z_1 to be $+1$. This strategy wins the game with probability $\frac{3}{4}$, since the product of the outputs is $+1$, which wins the game unless $(A, B, C) = (0, 0, 0)$.

So if we play the game once and the players win, we shouldn't be very surprised. But what if we play the game 100 times and the players always win? Using the above strategy, the probability of this is $(\frac{3}{4})^{100} \sim 10^{-13}$, but is there a better strategy?

It turns out that no assignment strategy can always win the game. To see this note that the requirements on the values are that $x_0 y_0 z_0 = -1$, $x_0 y_1 z_1 = 1$, $x_1 y_0 z_1 = 1$ and $x_1 y_1 z_0 = 1$. The

product of the left-hand-sides is $(x_0x_1y_0y_1z_0z_1)^2$, while the product of the right-hand-sides in -1, a contradiction. In fact, the best assignment strategy has success probability $\frac{3}{4}$.

However, we *can* always win the game using a quantum strategy! If the three players share a GHZ state [2], $\frac{1}{\sqrt{2}}(|000\rangle - |111\rangle)$ and each player on receiving 0 measures the observable σ_x and on receiving 1 measures the observable σ_y, giving the measurement outcome as their output, then they always win the game (Exercise: check this). The resulting correlations are called the GHZ correlations, and the above game is called the GHZ pseudo-telepathy game (the reason for calling it pseudo-telepathy is that the ability to always win the game cannot be explained from a classical point-of-view, and hence might appear telepathic).

That no assignment strategy works is the idea behind the potential power of GHZ (or indeed other) correlations in cryptography. One reverses the role of the game, replacing the referee with an adversary who is asked to send states to Alice, Bob and Charlie which together always win the GHZ game. The idea is that in order to do so, the adversary cannot use an assignment strategy, and hence has limited information about the outcomes. We will come back to this point later.

B. Bell Locality

In this section, we will consider arbitrary bipartite correlations (the definitions can be readily generalized to more parties). Consider two spacelike separated particles and performing a measurement on each. We denote the choice of measurement using A and B, and the respective outcomes X and Y. The distribution of outcomes given the measurement choices is denoted $P_{XY|AB}$. For the moment, we can forget quantum theory, and think of this setup in an arbitrary theory. All the distributions we will consider will be non-signalling, i.e. they satisfy $P_{X|AB} = P_{X|A}$ (the setting B does not affect the distribution of X given A) and $P_{Y|AB} = P_{Y|B}$. (Note that quantum correlations, in spite of their apparent pseudo-telepathic properties mentioned above, do not allow signalling.)

In a local hidden variable theory, one attempts to explain such correlations in terms of an additional parameter, Λ, which one can intuitively think of as being attached to the particles by the source. One then considers the distribution $P_{XY|AB\Lambda}$ such that the original correlations are recovered after averaging over Λ, i.e.[1]

$$P_{XY|AB} = \sum_{\lambda} P_{\Lambda|AB}(\lambda) P_{XY|AB\lambda} \tag{1}$$

[1] A note on notation. I will use upper case to denote random variables and lower case to denote particular instances of such variables. To connect with another common notation, $P_{\Lambda|AB}(\lambda)$ is often written $P(\Lambda = \lambda|A, B)$.

(this equation holds by definition of conditional probability distributions).

The first assumption that we will use is that the measurement settings can be chosen freely. This means that they are independent of all variables outside their future lightcone. In other words, we say that A is free if $P_{A|\Gamma} = P_A$, where Γ is the collection of all variables outside the future lightcone of A.

In the present case, that the settings A and B can be chosen freely allows us to write $P_{\Lambda|AB} = P_\Lambda$, so that (1) becomes

$$P_{XY|AB} = \sum_\lambda P_\Lambda(\lambda) P_{XY|AB\lambda}. \tag{2}$$

We also consider the case that the outcomes X and Y are completely determined from the settings and the hidden variables, so that $P_{XY|AB\lambda} = P_{X|A\lambda} P_{Y|B\lambda}$, with $P_{X|a\lambda}(x) \in \{0,1\}$ and $P_{Y|b\lambda}(y) \in \{0,1\}$.

This coincides with Bell's definition of locality (see e.g. [3]). In other words, we say that a set of correlations $P_{XY|AB}$ admits a local hidden variable description (with local hidden variable Λ) if it can be expressed in the form

$$P_{XY|AB} = \sum_\lambda P_\Lambda(\lambda) P_{X|A\lambda} P_{Y|B\lambda}, \tag{3}$$

with $P_{X|a\lambda}(x) \in \{0,1\}$ and $P_{Y|b\lambda}(y) \in \{0,1\}$. Correlations which cannot be expressed in this way are said to be *non-local*.

Remark: there exist stronger results that do not rely on X and Y being completely determined from the settings and the hidden variables [4].

C. Bell's theorem

We consider again the bipartite setting above and imagine that on each particle one of two measurements is made, i.e. $A \in \{0,2\}$ and $B \in \{1,3\}$, giving one of two outcomes $X, Y \in \{0,1\}$. We then characterize the quantum correlations in terms of the following quantity (the reason for calling this I_2 will become clear later)

$$I_2 = P(X = Y|0,3) + P(X \neq Y|0,1) + P(X \neq Y|2,1) + P(X \neq Y|2,3).$$

Bell's theorem (this is a actually a rewriting of the CHSH version of the theorem) then states that if $P_{XY|AB}$ admits a local hidden variable description (i.e. can be written in the form of (3)), then $I_2 \geq 1$. Interestingly, there are correlations which do not obey this inequality.

The proof of Bell's theorem will be worked out in the exercises, where we also examine a set of quantum correlations that do not satisfy $I_2 \geq 1$.

D. Other Bell inequalities

We will use the term Bell inequality to refer to any (non-trivial) constraint on the outcome probabilities satisfied by correlations which admit a local hidden variable description. (The trivial constraints are those such as that the probabilities are non-negative and sum to 1.)

We have seen another example already, in the form of the three-party game discussed in Section II A. We will discuss a further family of inequalities here, called *chained Bell inequalities* [5, 6], which we will come back to later on in a cryptographic context.

These are an extension of the CHSH inequality discussed above, where we allow N measurement choices for each of the two parties (we denote these $A \in \{0, 2, \ldots 2N-2\}$ and $B \in \{1, 3, \ldots 2N-1\}$). We define

$$I_N = P(X = Y|0, 2N-1) + \sum_{\substack{a,b \\ |a-b|=1}} P(X \neq Y|a, b). \tag{4}$$

For any $N \geq 2$, $I_N \geq 1$ in a Bell inequality. It turns out that there exist quantum correlations satisfying $I_N = 2N \sin^2 \frac{\pi}{4N}$, which tends to 0 for large N. We will come back to this family of Bell inequalities later.

E. More general non-local correlations

So far we have discussed correlations that admit a local hidden variable description and seen that there exist quantum correlations that are more general. In fact we can consider correlations which are more general still. Like in the previous section, most of what is said here could be generalized to more parties, but we focus on the bipartite case for simplicity.

One might ask, why study these? There are several reasons. For one, they shed light on quantum theory itself: one can examine why quantum theory has the correlations it does by studying theories with more general correlations and showing they are "paradoxical" in some way (I use paradoxical in quotes because such paradoxes can be subjective). In addition, as we will see there are sets of post-quantum correlations which are easier to characterize, and hence allowing post-quantum correlations can make security proofs simpler. Another reason is simply that these allow stronger security proofs, and ones that will stand even if quantum theory is one day

superseded. There is an additional subtle point regarding security proofs. Most are constructed under the implicit assumption that quantum theory is the most complete description of the world. However, the theory was developed to explain a set of observed phenomena, and does not contain any guarantee of completeness (hence the work of EPR, Bell, Kochen-Specker etc.) Fortunately, recent results have shown that it cannot be extended in a way which provides more information about experimental outcomes [4].

Let us denote the number of choices of measurement A by $|A|$ and the number of choices of outcome X by $|X|$, etc. For simplicity, we will take $|A| = |B|$ and $|X| = |Y|$, and hence define certain sets of distribution, e.g. $\mathcal{P}^{(2,|A|,|X|)}$, where 2 denotes the bipartite case[2]. This is the set of all "bipartite"[3] probability distributions with $|A|$ measurement settings for each party and $|X|$ outcomes for each (this is straightforwardly generalized), i.e.

$$\mathcal{P}^{(2,|A|,|X|)} = \{P_{XY|AB} : P_{XY|ab}(x,y) \geq 0 \ \forall \ a,b,x,y, \ \sum_{xy} P_{XY|ab}(x,y) = 1 \ \forall \ a,b\}.$$

We then define the set of non-signalling distributions, denoted $\mathcal{P}_{NS}^{(2,|A|,|X|)}$:

$$\mathcal{P}_{NS}^{(2,|A|,|X|)} = \{P_{XY|AB} \in \mathcal{P}^{(2,|A|,|X|)} : P_{X|AB} = P_{X|A}, P_{Y|AB} = P_{Y|B}\}.$$

Quantum correlations form a subset of $\mathcal{P}_{NS}^{(2,|A|,|X|)}$ which we denote $\mathcal{P}_{QM}^{(2,|A|,|X|)}$. These correlations are those for which there exists a bipartite quantum state (positive semi-definite operator with trace 1) ρ and sets of POVMs, $\{E_x^a\}$ and $\{F_y^b\}$ (i.e. positive semi-definite operators with $\sum_x E_x^a = \mathbb{1}$ for all a, and $\sum_y F_y^b = \mathbb{1}$ for all b) such that $P_{XY|ab}(x,y) = \mathrm{tr}((E_x^a \otimes F_y^b)\rho)$.

These sets are distinct. One example of a non-signalling distribution that is not quantum is the set of correlations for which $I_2 = 0$ with $P_{X|a}(0) = P_{Y|b}(0) = \frac{1}{2}$ for all a and b. This set of correlations is often called a non-local box. It is extremal amongst the set of non-signalling distributions (i.e. it cannot be decomposed as a convex mixture of other non-signalling correlations). This distribution will be discussed further in the exercises, where we will prove that it is not in $\mathcal{P}_{QM}^{(2,|A|,|X|)}$.

The third set we consider is the set that admits a local hidden variable description, which I here call $\mathcal{P}_{LHV}^{(2,|A|,|X|)}$. As is shown in Exercise 1, this set is distinct from the quantum set, so that $\mathcal{P}_{LHV}^{(2,|A|,|X|)} \subset \mathcal{P}_{QM}^{(2,|A|,|X|)} \subset \mathcal{P}_{NS}^{(2,|A|,|X|)}$. Note also that each of the sets $\mathcal{P}_{NS}^{(2,|A|,|X|)}$, $\mathcal{P}_{QM}^{(2,|A|,|X|)}$ and $\mathcal{P}_{LHV}^{(2,|A|,|X|)}$ is convex.

---

[2] More generally $\mathcal{P}^{(n,|A|,|B|,\ldots,|X|,|Y|,\ldots)}$ is the set of n-party distributions with the stated number of settings and outcomes for each party.

[3] I use bipartite in quotes here, since the concept of subsystems breaks down if the correlations are signalling (which is allowed within this set).

These distributions can be represented as vectors in a vector space. In the bipartite case where each system has 2 inputs and 2 outputs, i.e. $|A| = |X| = 2$, there are 16 probabilities $P_{XY|AB}$, but because of the non-signalling conditions and normalization, the set of distributions is 8-Dimensional (Exercise: show this). These sets can be characterized by their extreme points (any member of each set is a convex combination of the extreme points of that set). The sets $\mathcal{P}_{NS}^{(2,2,2)}$ and $\mathcal{P}_{LHV}^{(2,2,2)}$ form polytopes within the 8-Dimensional space (i.e., these sets have a finite number of extremal points, the vertices of the polytope). Bell inequalities, such as $I_2 \geq 1$ correspond to faces of $\mathcal{P}_{LHV}^{(2,2,2)}$. The set $\mathcal{P}_{QM}^{(2,2,2)}$ does not form a polytope. For a review of these properties, see [7].

The boundaries of these three sets all come together at local deterministic points (points where $P_{X|a} \in \{0,1\}$ and $P_{Y|b} \in \{0,1\}$ for all a and b), which form the set of extreme points of $\mathcal{P}_{LHV}^{(n,|A|,|X|)}$, but there can also be points where quantum correlations exist on the boundary of $\mathcal{P}_{NS}^{(n,|A|,|X|)}$, but not on the boundary of $\mathcal{P}_{LHV}^{(n,|A|,|X|)}$, for example, GHZ correlations discussed in Section II A. (Note that the probability of winning the game with GHZ correlations is 1, which is equal to maximum possible probability of winning the game, and therefore to the best non-signalling strategy.)

F. Tests for quantum correlations

It is reasonably straightforward to test whether or not a given set of correlations, $P_{XY|AB}$, is in $\mathcal{P}_{NS}^{(2,|A|,|X|)}$ or $\mathcal{P}_{LHV}^{(2,|A|,|X|)}$ (again, for notational simplicity, I'll take $|A| = |B|$ and $|X| = |Y|$). However, deciding whether or not a given set of correlations is quantum is less straightforward, in spite of the simple definition. One reason for this is that the quantum set is not a polytope, so cannot be characterized in terms of a finite number of linear inequalities.

One way to certify that a set of correlations $P_{XY|AB}$ is quantum is to search for sets of positive semi-definite operators $\{E_x^a\}$, $\{F_y^b\}$ and ρ such that $\sum_x E_x^a = \mathbb{1}$ for all a, and $\sum_y F_y^b = \mathbb{1}$ for all b, $\mathrm{tr}\rho = 1$, and $P_{XY|ab}(x,y) = \mathrm{tr}((E_x^a \otimes F_y^b)\rho)$. However, any computational search is restricted in the maximum dimension of the spaces $\mathcal{H}_A$ and $\mathcal{H}_B$. So, although finding that such a set exists in some finite dimensions certifies that the correlations are in $\mathcal{P}_{QM}^{(2,|A|,|X|)}$, not finding any such set is not, in general, sufficient to establish that the correlations are not quantum.

Conversely, there is a technique which provides a certificate that a distribution is not in $\mathcal{P}_{QM}^{(2,|A|,|X|)}$. This is achieved through a set of necessary conditions for quantum correlations [8]. Suppose a set of correlations $P_{XY|AB}$ is in $\mathcal{P}_{QM}^{(2,|A|,|X|)}$. We define the set $\{A_1, A_2, \ldots\} = \{\mathbb{1} \otimes \mathbb{1}\} \cup \{E_x^a \otimes \mathbb{1}\}_{a,x} \cup \{\mathbb{1} \otimes F_y^b\}_{b,y}$, and construct the matrix M for which $M_{i,j} = \mathrm{tr}(A_i^\dagger A_j \rho)$. (Note that this matrix contains elements defined by the product of operators on the same system,

e.g. $\mathrm{tr}((E_0^0 E_1^2 \otimes \mathbb{1})\rho)$.) This matrix is always positive semi-definite, since[4]

$$\langle v|M|v\rangle = \sum_{i,j} \mathrm{tr}(v_i^\dagger A_i^\dagger A_j v_j \rho) = \mathrm{tr}((\sum_i A_i v_i)^\dagger (\sum_i A_i v_i)\rho) \geq 0,$$

where the last inequality follows because the trace of the product of two positive semi-definite operators cannot be negative.

Now imagine we are given a set of correlations $P_{XY|AB}$, and want to establish whether they are in $\mathcal{P}_{QM}^{(2,|A|,|X|)}$. We can imagine a partial construction of M, where we fill in any entries we can, e.g. $\mathrm{tr}((E_0^0 \otimes F_0^1)\rho) = P_{XY|01}(0,0)$, or $\mathrm{tr}((E_0^0 \otimes \mathbb{1})\rho) = P_{X|0}(0)$ etc. Note that some entries cannot be filled in directly, such as that corresponding to $\mathrm{tr}((E_0^0 E_1^2 \otimes \mathbb{1})\rho)$, which is not an element of $P_{XY|AB}$, and is not a measurable quantity. Having filled in as many entries as possible, we ask: does there exist a completion of the matrix such that it is positive semi-definite? If there does not, then we know that the distribution is not in $\mathcal{P}_{QM}^{(2,|A|,|X|)}$. However, if there does, then the problem is undecided (the condition that M is positive semi-definite is necessary for the distribution to be in $\mathcal{P}_{QM}^{(2,|A|,|X|)}$, but not sufficient).

There is another characterization of the set of quantum correlations, $\mathcal{P}_{QM}^{(2,|A|,|X|)}$ that is often more suited to this purpose. One can equivalently define the set $\mathcal{P}_{QM}^{(2,|A|,|X|)}$ as the set of distributions for which there exists a pure state $|\Psi\rangle$ and orthogonal projective measurements, $\{E_x^a\}$ and $\{F_y^b\}$ (for $\{E_x^a\}$, this means $E_x^a E_{x'}^a = E_x^a \delta_{x,x'}$ and $\sum_x E_x^a = \mathbb{1}$ for all a) such that $P_{XY|ab}(x,y) = \langle\Psi|E_x^a \otimes F_y^b|\Psi\rangle$. This follows because any mixed state can be purified on a larger system, and any POVM can be viewed as a projective measurement on a larger system (this is proven in [9]). Thus, since we have not set any limit on the size of the space required to realize these correlations, this is an equivalent description. This characterization allows us to fill in the entry of M corresponding to $\langle\Psi|E_0^a E_1^a \otimes \mathbb{1}|\Psi\rangle$ as 0, for example.

In addition, this technique can be used for other sets of operators (instead of $\{E_x^a \otimes \mathbb{1}\}$ etc.), and an example is given in the exercises.

As mentioned above, if M can be completed in a positive semi-definite way, it does not imply that the distribution is quantum. However, the set $\{A_i\}$ can be extended to including other operators, building a hierarchy of conditions to test for quantum correlations. These are beyond

[4] Recall that an operator A is positive semi-definite if for all vectors v, $\langle v|A|v\rangle \geq 0$. Equivalently, A is positive semi-definite if all its eigenvalues are greater than or equal to 0.

the scope of this module. Further details can be found in [10].

---

[1] Einstein, A., Podolsky, B. & Rosen, N. Can quantum-mechanical description of physical reality be considered complete? *Physical Review* **47**, 777–780 (1935).

[2] Greenberger, D. M., Horne, M. & Zeilinger, A. Going beyond Bell's theorem. In Kafatos, M. (ed.) *Bell's Theorem, Quantum Mechanics and Conceptions of the Universe*, 69–72 (Kluwer Academic, Dordrecht, The Netherlands, 1989).

[3] Bell, J. S. *Speakable and unspeakable in quantum mechanics* (Cambridge University Press, 1987).

[4] Colbeck, R. & Renner, R. Quantum theory cannot be extended. e-print `arXiv:1005.5173` (2010).

[5] Pearle, P. M. Hidden-variable example based upon data rejection. *Physical Review D* **2**, 1418–1425 (1970).

[6] Braunstein, S. L. & Caves, C. M. Wringing out better Bell inequalities. *Annals of Physics* **202**, 22–56 (1990).

[7] Tsirelson, B. Some results and problems on quantum Bell-type inequalities. *Hadronic Journal Supplement* **8**, 329–345 (1993).

[8] Navascués, M., Pironio, S. & Acín, A. Bounding the set of quantum correlations. *Physical Review Letters* **98**, 010401 (2007).

[9] Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).

[10] Navascués, M., Pironio, S. & Acín, A. A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. *New Journal of Physics* **10**, 073013 (2008).