Bell Inequalities

Roger Colbeck, *

(Dated: 23rd June, 2017)

This set of notes cover a slightly different set of topics to those of the summer school lectures, but broadly follow the same lines.

I. NONLOCALITY, EPR AND BELL'S THEOREM

In this section we will examine in detail one of the most fascinating departures from classical theory exhibited by our world, so-called non-locality. Although non-locality at first seems like an odd curiosity of the theory, it is in fact connected to a practical application: 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 (EPR)[1] to question the completeness of the theory. They considered a measurement on one half of a maximally entangled pair in the state $|\Psi_-\rangle$ 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 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

*roger.colbeck@york.ac.uk

[1] See *Can quantum-mechanical description of physical reality be considered complete?* Physical Review **47**, 777-780 (1935)

hidden variable, see later). The anti-correlated outcomes can then be alternatively explained using this shared randomness.

A. 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.

1. Example

Before discussing in more detail, here is an example that 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

$$(A, B, C) = (0,0,0),\ (0,1,1),\ (1,0,1),\ \text{and}\ (1,1,0)\,.$$

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 rules 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, $\ldots$, 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

$$x_0y_0z_0 = -1,\ x_0y_1z_1 = 1,\ x_1y_0z_1 = 1 \text{ and } x_1y_1z_0 = 1.$$

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 $\frac{3}{4}$. (To rigorously confirm this, we could run through each of the 2^6 assignments.)

However, we *can* always win the game using a quantum strategy! If the three players share the state $\frac{1}{\sqrt{2}}(|000\rangle - |111\rangle)$ and each player on receiving 0 measures the observable σ_x (i.e., with measurement operators $\{|+\rangle\langle+|, |-\rangle\langle-|\}$) and on receiving 1 measures the observable σ_y, giving the measurement outcome as their output, then they always win the game. 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. Local deterministic theories

In this section, we will consider bipartite correlations (although what we say can readily be generalized to more parties). Consider two distant particles and performing a measurement on each at spacelike separation. 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 unspecified theory. A general theory is based on some parameters that we denote using Λ, and gives us a prescription to compute the output distribution $P_{XY|AB\Lambda}$ based on the settings and the parameters of the theory.

Firstly, we make the important assumption that A and B can be chosen freely. We now explain what this means.

1. *Free choice*

It is difficult to imagine talking about physics without free choice. It is a notion that finds itself embedded within the usual language we use to describe physical scenarios. We ask questions of the form "What would happen if..." and reason about the consequences. Such statements are only meaningful under the premise that it is possible to set up the scenario in question.

Thus, free choice is a property of the way we describe the world, with an interventionist picture. We would not find a theory satisfying if there were particular scenarios (or sets of initial conditions) for which it would fail to compute the future evolution. Sometimes free choice has been called a no-conspiracy assumption, the idea being that if free choice were not to hold—i.e., were it impossible to set up a particular scenario—it would be a conspiracy on the part of nature preventing certain measurements being performed in certain situations. Note that although this may be considered a conspiracy, it isn't something that can be excluded experimentally.

Mathematically, free choice is the assertion that certain variables, "choices", are independent of certain other variables. When formulating a free choice assumption, one has to say how to specify these other variables. As we will explain, a natural way to make this specification is connected to the causal order.

Consider a simple setup where a system is prepared in state Λ and measured according to A, producing an outcome X. For A to be free it is natural to demand that it can be chosen independently of Λ, i.e., $P_{A|\Lambda} = P_A$. However, it would be too restrictive to also require independence from X, i.e., $P_{A|X\Lambda} = P_A$, as we expect the outcome of an experiment to depend on the measurement. This simple example illustrates that the notion of free choice is intrinsically connected to a causal order: we don't require that the free choice A is uncorrelated with X since X lies in the causal future of A.

This is the principle on which the general definition of free choice is based. For a collection of random variables, we give a causal order by drawing a set of arrows ($\rightsquigarrow$) between the variables.[2] These arrows should be interpreted as pointing to the future, so that, for example, $G \rightsquigarrow H$ indicates that H is in the causal future of G. Note that H can be in the causal future of G without G necessarily being a cause of H. We also use $G \not\rightsquigarrow H$ to indicate that H is not in the causal future of G. Note also that the relation $\rightsquigarrow$ is transitive, so $E \rightsquigarrow F$ and $F \rightsquigarrow G$ implies $E \rightsquigarrow G$. This fits well with our common understanding of "future".

[2] More precisely, a causal order is a binary relation ($\rightsquigarrow$) on those variables that is reflexive ($G \rightsquigarrow G$) and transitive ($G \rightsquigarrow H$ and $H \rightsquigarrow J$ implies $G \rightsquigarrow J$), in other words, a *preorder* relation.

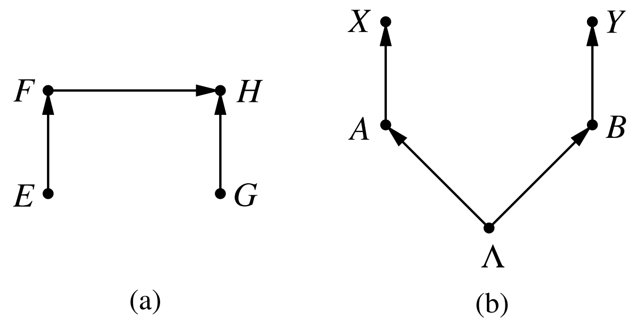

FIG. 1: Two examples of a causal order. In (a), F being free implies $P_{F|EG} = P_F$, while in (b), if A is free then $P_{A|BY\Lambda} = P_A$, for example.

With respect to a given causal order, free choice can be defined as follows:

> A random variable G is a **free choice** with respect to some causal order if G is uncorrelated with all variables of the causal order outside its causal future.

Said another way, G is free with respect to some causal order, if the only other variables it is correlated with are those that it could have caused.

For this definition, one can only make sense of free choice within a defined causal order. In principle, this may be selected arbitrarily. However, there is usually a well-motivated way to set this based on the physical situation. For example, it is natural to take $G \rightsquigarrow H$ to hold if and only if there exists a reference frame in which G occurs at an earlier time to H. In this way, the causal order is compatible with relativity theory. (Note, however, that doing this is not the same as assuming relativity theory. Instead, relativity simply acts as a physical motivation for the chosen causal order.) See Fig. 1 for examples.

Note that quantum theory is compatible with free choice with respect to the causal order in Fig. 1(b) (in the case of quantum theory, Λ is the quantum state Ψ). To see this, consider a measurement on the 2nd part of an entangled state $|\psi\rangle_{12}$. If the measurement $B = b$ has operators $\{M_y^b\}_y$[3], then on obtaining outcome $Y = y$, the post-measurement state is

$$|\phi_y^b\rangle = \frac{1}{\sqrt{P_{Y|b\psi}(y)}}(\mathbb{1} \otimes M_y^b)|\psi\rangle_{12}\,,$$

where

$$P_{Y|b\psi}(y) = \langle\psi|(\mathbb{1} \otimes (M_y^b)^\dagger M_y^b)|\psi\rangle\,.$$

[3] The condition for a valid measurements is that $\sum_y (M_y^b)^\dagger M_y^b = \mathbb{1}$ for all b.

Independently of the value of Y that occurred, we can still measure the 1st system with any distribution over the possible measurements, $P_{A|by\psi}$, we like, and quantum theory allows predict the distribution of outcomes. For measurement $\{N_x^a\}_x$, the outcome distribution is

$$P_{X|aby\psi}(x) = \langle\phi_y^b|((N_x^a)^\dagger N_x^a \otimes \mathbb{1})|\phi_y^b\rangle\,.$$

2. Connection to no signalling

In the bipartite scenario in question, the natural causal order is that of Fig. 1(b), which we henceforth call the bipartite causal order. The parameters of the higher theory, Λ, can be associated with the generation of the systems, and hence are naturally taken to be in the causal past of the measurements.

It turns out (see exercises) that free choice within this causal order implies

$$P_{X\Lambda|AB} = P_{X\Lambda|A} \text{ and } P_{Y\Lambda|AB} = P_{Y\Lambda|B}\,,$$

which are essentially no-signalling conditions (they say that if Alice holds one of the systems and has access to Λ, then Bob cannot signal to her, for example). (Note that quantum correlations, in spite of their apparent pseudo-telepathic properties mentioned above, do not allow signalling. Thus, if we look at quantum theory, which is the case where Λ is the quantum state, then we get compatibility with free choice in this causal order).

After this (slightly lengthy) aside, let us return to our discussion of general theories for some set of correlations $P_{XY|AB}$. A general theory based on parameters Λ specifies a distribution $P_{XY|AB\Lambda}$, as well as some distribution on Λ. The theory is a good one for the observed correlations if

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

Note that we have implicitly assumed free choice, since for an arbitrary distribution, by definition of conditional probability, we have

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

which reduces to the previous equation when $P_{\Lambda|AB} = P_\Lambda$ (which in turns follows from A and B being chosen freely wrt the causal order of Fig. 1).

Now let us specialize to the case of local deterministic theories. In such a theory, the outcomes X and Y are completely determined by the local measurement settings and the parameters of the theory. This means that

$$P_{XY|AB\lambda} = P_{X|A\lambda}P_{Y|B\lambda}\,,$$

with $P_{X|a\lambda}(x) \in \{0,1\}$ and $P_{Y|b\lambda}(y) \in \{0,1\}$ (such distributions are said to be local deterministic).

To summarize, we say that a set of correlations $P_{XY|AB}$ can be described in a **local deterministic theory** (with parameters Λ) if it can be expressed in the form

$$P_{XY|AB} = \sum_{\lambda} P_{\Lambda}(\lambda)P_{X|A\lambda}P_{Y|B\lambda}, \tag{1}$$

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

A note on terminology: Correlations that can be described by a local deterministic theory are also said to have a deterministic local hidden variable description, the idea being that the parameters of the deterministic theory must be unknown to us today (since we use quantum theory) and are therefore “hidden variables”. Strictly (1) describes correlations that can be described by a local deterministic theory compatible with free choice within the bipartite causal order. The free choice part is often left implicit for brevity, however, it is an important assumption that shouldn’t be forgotten.

Note also that correlations that can be described by a local deterministic theory can also be thought of as ones that have a “reasonable classical explanation”, or as correlations that shouldn’t be surprising if we live in a classical world. Consider again the EPR correlations, for example. These satisfy $P_{XY}(x, y) = \frac{1}{2}(1 - \delta_{x,y})$, where $x, y \in \{0, 1\}$. We can explicitly see that these correlations admit a local hidden variable description, as follows (in this case, there is no choice of input A and B).

Let Λ is a binary random variable that satisfies $P_{\Lambda}(0) = P_{\Lambda}(1) = \frac{1}{2}$, $P_{X|\lambda}(x) = \delta_{x,\lambda}$ and $P_{Y|\lambda}(y) = 1 - \delta_{y,\lambda}$. Then,

$$\sum_{\lambda} P_{\Lambda}(\lambda)P_{X|\lambda}(x)P_{Y|\lambda}(y) = \frac{1}{2}\sum_{\lambda} \delta_{x,\lambda}(1 - \delta_{y,\lambda}) = \frac{1}{2}(1 - \delta_{x,y}) = P_{XY}(x, y)\,.$$

C. Bell’s theorem

Having discussed local deterministic theories, we now would like a way to certify that particular correlations can be described in such a theory. A **Bell inequality** is a necessary condition for this to be possible. In other words, correlations that violate a Bell inequality cannot be described by a local deterministic theory. Roughly speaking, Bell's theorem says that there are quantum correlations that don't satisfy all Bell inequalities. We will proceed with a discussion of this.

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)

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

To connect to an alternative notation, note that $P(X = Y|a, b) = \sum_x P_{XY|ab}(x, x)$, and hence $P(X \neq Y|a, b) = 1 - \sum_x P_{XY|ab}(x, x)$.

If $P_{XY|AB}$ can be described by a local deterministic theory (i.e. can be written in the form of (1)) then $I_2 \geq 1$. This relation (a constraint on the correlations of any local deterministic theory) is called a Bell inequality[4] The proof of this inequality is a simple exercise in probability theory and will be worked out in the exercises. [Note that our definition of a theory includes that A and B are free choices wrt the causal order in Fig. 1.]

The amazing thing is that there exist quantum correlations that do not satisfy $I_2 \geq 1$. This is often stated as "quantum correlations are non-local", which is shorthand for the statement "there exist quantum correlations that cannot be described by any local deterministic theory". Again we will see this on the exercises. [Additional exercise: show that local measurements on unentangled quantum states also obey $I_2 \geq 1$.]

We can therefore state Bell's theorem more formally:

Theorem 1 (No locally deterministic theory can reproduce all quantum correlations)**.** *Let A, B, X, Y and Λ be RVs. Then at least one of the following cannot hold:*

- Freedom of choice: *A and B are free with respect to the causal order depicted in Figure 1;*
- Compatibility with quantum theory: *$P_{XY|AB\Lambda}$ is compatible with the predictions of quantum theory (in the sense of* (1)*) for all quantum realisable $P_{XY|AB}$;*
- Local determinism: *$P_{XY|ab\lambda}(x, y) = P_{X|a\lambda}(x)P_{Y|b\lambda}(y)\ \forall\ a, b, x, y, \lambda$, $P_{X|a\lambda}(x) \in \{0, 1\}\ \forall\ a, \lambda$ and $P_{Y|b\lambda}(y) \in \{0, 1\}\ \forall\ b, \lambda$.*

[4] This is a actually a rewriting of the CHSH inequality [*Proposed experiment to test local hidden-variable theories*, Physical Review Letters **23** 880–884 (1969)], see exercises.

This is the neutral way of stating the theorem. However, usually it is phrased as that the first and third conditions rule out the second. We can also specifically mention the quantum correlations that (maximally) violate $I_2 \geq 1$. These are formed by measurements on the maximally entangled 2-qubit state $|\Phi_+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, where $A = 0$ corresponds to measurement in the $\{|0\rangle, |1\rangle\}$ basis, $A = 2$ corresponds to measurement in the $\{|+\rangle, |-\rangle\}$ basis, while $B = 1$ corresponds to measurement in the $\{\cos\frac{\pi}{8}|0\rangle + \sin\frac{\pi}{8}|1\rangle, \sin\frac{\pi}{8}|0\rangle - \cos\frac{\pi}{8}|1\rangle\}$ basis and $B = 3$ is in the $\{\cos\frac{3\pi}{8}|0\rangle + \sin\frac{3\pi}{8}|1\rangle, \sin\frac{3\pi}{8}|0\rangle - \cos\frac{3\pi}{8}|1\rangle\}$ basis. If we use M_x^a for the respective measurement operators of Alice and N_y^b for those of Bob (e.g., $M_0^0 = |0\rangle\langle 0|$, $M_1^0 = |1\rangle\langle 1|$, $M_0^2 = |+\rangle\langle +|$ and $M_1^2 = |-\rangle\langle -|$ etc.) then $P_{XY|ab}(x, y) = \langle\Phi_+|(M_x^a \otimes N_y^b)|\Phi_+\rangle$ are the quantum predictions. [Note that all the measurements are projective so $(M_x^a)^2 = M_x^a$ etc.]

Theorem 2 (Bell's theorem reformulated)**.** *Suppose A, B, X, Y and Λ be RVs, that A and B are free with respect to the causal order depicted in Figure 1, and that $P_{XY|ab\lambda}(x, y) = P_{X|a\lambda}(x)P_{Y|b\lambda}(y)\ \forall\ a, b, x, y, \lambda$, $P_{X|a\lambda}(x) \in \{0, 1\}\ \forall\ a, \lambda$ and $P_{Y|b\lambda}(y) \in \{0, 1\}\ \forall\ b, \lambda$.*

Then it is not possible for $P_{XY|ab}(x, y)$ to be the distribution generated by the state and measurements specified above.

D. Other Bell inequalities

The term Bell inequality is usually used to refer to any (non-trivial) constraint on the outcome probabilities satisfied by correlations that admit a local hidden variable description. (The trivial constraints are that the probabilities are non-negative.) We have seen another example already, in the form of the three-party game discussed in Section I A 1. We will discuss a further family of bipartite inequalities here, called *chained Bell inequalities*.

It turns out that the minimum value of I_2 that can be achieved quantum mechanically is $4\sin^2\frac{\pi}{8} = 2 - \sqrt{2} \approx 0.59$. However, the minimum possible value of I_2 (for any correlations) is 0 (I_2 is the sum of positive quantities). There is a family of generalizations of I_2 that we call I_N. They maintain the classical minimum at 1, but there exist quantum correlations that get arbitrarily close to 0.

The generalizations of I_2 allow N measurement choices for each of the two parties (we denote these choices $A \in \{0, 2, \ldots 2N-2\}$ and $B \in \{1, 3, \ldots 2N-1\}$). We define

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

For any $N \geq 2$, $I_N \geq 1$ in a Bell inequality. There exist quantum correlations satisfying $I_N = 2N\sin^2(\frac{\pi}{4N})$, which tends to 0 for large N. This family of Bell inequalities is useful for cryptography.

II. LOOPHOLES

For a discussion of these I recomment the relevant sections of the review article *Bell Nonlocality* by Brunner et al. Revews of Modern Physics **86** (2014).

III. EXERCISES

A. Exercise 1

Consider a bipartite setting in which two spacelike separated parties make measurements (their choices being denoted $A \in \{0, 2\}$ and $B \in \{1, 3\}$) with respective outcomes $X \in \{0, 1\}$ and $Y \in \{0, 1\}$ giving rise to a joint distribution $P_{XY|AB}$. It will be convenient to write such distributions in the form of a table of values as follows:

$P_{XY\mid AB}$	B	1		3	
	Y	0	1	0	1
A	X				
0	0	$P_{00\mid 01}$	$P_{01\mid 01}$	$P_{00\mid 03}$	$P_{01\mid 03}$
	1	$P_{10\mid 01}$	$P_{11\mid 01}$	$P_{10\mid 03}$	$P_{10\mid 03}$
2	0	$P_{00\mid 21}$	$P_{01\mid 21}$	$P_{00\mid 23}$	$P_{01\mid 23}$
	1	$P_{10\mid 21}$	$P_{11\mid 21}$	$P_{10\mid 23}$	$P_{11\mid 23}$

This form is convenient because the fact that Alice cannot signal to Bob and vice versa can be seen by checking that within each row, the sum of the left two elements is equal to the sum of the right two elements, and analogously for the columns. We will also consider the quantity

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

(a) Draw the natural causal order associated with this setup (including allowance for an additional parameter Λ representing the parameters of a candidate alternative theory) and show that if A and B are free choices with respect to this causal order then it follows that $P_{X\Lambda|AB} = P_{X\Lambda|A}$ and $P_{Y\Lambda|AB} = P_{Y\Lambda|B}$. [Hint: consider expanding $P_{XB\Lambda|A}$ using the definition of conditional probability.] Show that this implies that Alice (holding the (A, X) system) cannot signal to Bob (holding the (B, Y) system) even if both parties know Λ.

(b) A local deterministic distribution is one in which $P_{X|a}(x) \in \{0,1\}$ and $P_{Y|b}(y) \in \{0,1\}$ for all a, b, x, y. Write down a local deterministic distribution in the form of such a table. How many such distributions are there?

(c) Consider now a general distribution $P_{XY|AB}$ that can be described in a local deterministic theory, i.e., that can be decomposed as a convex combination of deterministic distributions. Show that any such distribution satisfies $I_2 \geq 1$.

(d) Consider the distribution

$P_{XY\|AB}$	B	1		3	
	Y	0	1	0	1
A	X				
0	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$	0
	1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$	0
2	0	$\frac{1}{2}$	$\frac{1}{2}$	1	0
	1	0	0	0	0

Show that this can be described in a local deterministic theory.

(e) Consider the quantum correlations resulting from measurement on the state $|\Phi_+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, where $A = 0$ corresponds to measurement in the $\{|0\rangle, |1\rangle\}$ basis, $A = 2$ corresponds to measurement in the $\{|+\rangle, |-\rangle\}$ basis, while $B = 1$ corresponds to measurement in the $\{\cos\frac{\pi}{8}|0\rangle + \sin\frac{\pi}{8}|1\rangle, \sin\frac{\pi}{8}|0\rangle - \cos\frac{\pi}{8}|1\rangle\}$ basis and $B = 3$ is in the $\{\cos\frac{3\pi}{8}|0\rangle + \sin\frac{3\pi}{8}|1\rangle, \sin\frac{3\pi}{8}|0\rangle - \cos\frac{3\pi}{8}|1\rangle\}$ basis. (These measurements correspond to the observables σ_z and σ_x for one party and $\frac{1}{\sqrt{2}}(\sigma_x \mp \sigma_z)$ for the other party.) These correlations take the form

$P_{XY\|AB}$	B	1		3	
	Y	0	1	0	1
A	X				
0	0	$\frac{1}{2} - \varepsilon$	ε	ε	$\frac{1}{2} - \varepsilon$
	1	ε	$\frac{1}{2} - \varepsilon$	$\frac{1}{2} - \varepsilon$	ε
2	0	$\frac{1}{2} - \varepsilon$	ε	$\frac{1}{2} - \varepsilon$	ε
	1	ε	$\frac{1}{2} - \varepsilon$	ε	$\frac{1}{2} - \varepsilon$

.

Determine the value of ε and hence find the value of I_2 for these correlations. Comment on your answer in light of Part (c).

(f) Consider now an analogous set of correlations, but with a value of ε higher than that for the quantum correlations given above (but with $\varepsilon \leq \frac{1}{4}$). Give the form of a 1-parameter family of (mixed) quantum states which gives such correlations using the same measurements.

(g) At what value of ε do the correlations cease to violate $I_2 \geq 1$? What is the trace distance between the corresponding quantum state and $|\Phi_+\rangle$?

B. Exercise 2

Consider the Bell inequality $I_2 \geq 1$, as introduced in the lectures. In this exercise, we consider an alternative way to express this. Suppose that we relabel the measurement outcome $X = 0$ as $X = -1$, and similarly for Y, so that $X, Y \in \{-1, 1\}$ (rather than $\{0, 1\}$ as in the lectures). For measurement $A = 0$, define an observable A_0, and similarly for $A = 2$, $B = 1$ and $B = 3$. Show that $\langle A_0 \otimes B_1 \rangle = P(X = Y|0, 1) - P(X \neq Y|0, 1)$. By expressing $\langle A_0 \otimes B_3 \rangle$, $\langle A_2 \otimes B_3 \rangle$ and $\langle A_2 \otimes B_1 \rangle$ similarly, show that $I_2 \geq 1$ implies

$$\langle A_0 \otimes B_1 \rangle + \langle A_2 \otimes B_1 \rangle + \langle A_2 \otimes B_3 \rangle - \langle A_0 \otimes B_3 \rangle \leq 2$$

(this is a common form in which to express this Bell inequality, and the above is usually called the CHSH inequality).

The quantum bound on I_2 is $I_2 \geq 2 - \sqrt{2}$. Use this to calculate the maximum value of

$$\langle A_0 \otimes B_1 \rangle + \langle A_2 \otimes B_1 \rangle + \langle A_2 \otimes B_3 \rangle - \langle A_0 \otimes B_3 \rangle$$

that can be achieved quantum mechanically (this bound is called Tsirelson's bound).

IV. SOLUTIONS

A. Solution 1

(a) The causal order was given in the notes. Free choice gives $P_{A|BY\Lambda} = P_A$ and $P_{B|AX\Lambda} = P_B$. Now consider $P_{XB\Lambda|A}$. This can be expanded as $P_{XB\Lambda|A} = P_{X\Lambda|A} P_{B|AX\Lambda}$, or $P_{XB\Lambda|A} = P_{B|A} P_{X\Lambda|AB}$. Free choice gives that $P_{B|AX\Lambda} = P_{B|A} = P_B$, so we have $P_{X\Lambda|A} = P_{X\Lambda|AB}$. Analogously, by symmetry, we have $P_{Y\Lambda|B} = P_{Y\Lambda|AB}$. This implies $P_{Y|AB\Lambda} = P_{Y|B\Lambda}$, i.e., Alice cannot signal to Bob, even if he knows Λ. (If two distributions are equal, then so are their marginals.)

(b) Local deterministic strategies look as follows:

$P_{XY\|AB}$	B	1		3	
	Y	0	1	0	1
A	X				
0	0	1	0	1	0
	1	0	0	0	0
2	0	1	0	1	0
	1	0	0	0	0

and there are $16 = 4 \times 2 \times 2$ of them (4 ways to choose the 1 in the top left, then 2 in the top right, then two in the bottom left, which fixes the last position).

(c) It is easy to see that the smallest I_2 for a local deterministic strategy is 1 (if you're in doubt, there are only 16 cases to check). By definition, a distribution that admits a local hidden variable description is a convex combination of local deterministic strategies. Since I_2 is linear, it satisfies $I_2 \geq 1$.

(d) The distribution can be decomposed into the following convex combination of local deterministic distributions:

$$\frac{1}{4}\left(\begin{array}{|cc|cc|}\hline 1&0&1&0\\ 0&0&0&0\\ \hline 1&0&1&0\\ 0&0&0&0\\ \hline\end{array} + \begin{array}{|cc|cc|}\hline 0&1&1&0\\ 0&0&0&0\\ \hline 0&1&1&0\\ 0&0&0&0\\ \hline\end{array} + \begin{array}{|cc|cc|}\hline 0&0&0&0\\ 1&0&1&0\\ \hline 1&0&1&0\\ 0&0&0&0\\ \hline\end{array} + \begin{array}{|cc|cc|}\hline 0&0&0&0\\ 0&1&1&0\\ \hline 0&1&1&0\\ 0&0&0&0\\ \hline\end{array}\right).$$

Thus we can take $P_\Lambda(0) = P_\Lambda(1) = P_\Lambda(2) = P_\Lambda(3) = \frac{1}{4}$, and for example, if $\Lambda = 0$, then $X = 0$ and $Y = 0$, if $\Lambda = 1$, then $X = 0$, $Y = \begin{cases} 0 & \text{if } B = 1 \\ 1 & \text{if } B = 3 \end{cases}$, etc.

(e) We have

$$\begin{aligned}
P_{XY|A=0,B=1}(0,0) &= |\langle\Phi_+|(|0\rangle \otimes (\cos\frac{\pi}{8}|0\rangle + \sin\frac{\pi}{8}|1\rangle))|^2 = |\frac{1}{\sqrt{2}}\cos\frac{\pi}{8}|^2 = \frac{1}{2}\cos^2\frac{\pi}{8} = \frac{1}{2}(1-\sin^2\frac{\pi}{8}) \\
P_{XY|A=0,B=1}(0,1) &= |\langle\Phi_+|(|0\rangle \otimes (\sin\frac{\pi}{8}|0\rangle - \cos\frac{\pi}{8}|1\rangle))|^2 = \frac{1}{2}\sin^2\frac{\pi}{8} \\
&\vdots \\
P_{XY|A=0,B=3}(0,0) &= |\frac{1}{\sqrt{2}}\cos\frac{3\pi}{8}|^2 = \frac{1}{2}\sin^2\frac{\pi}{8} \\
&\vdots
\end{aligned}$$

etc. so that $\varepsilon = \frac{1}{2}\sin^2\frac{\pi}{8} = \frac{1}{4}(1-\cos\frac{\pi}{4}) = \frac{1}{4}(1-\frac{1}{\sqrt{2}}) = \frac{1}{8}(2-\sqrt{2}) \approx 0.073$.

The value of I_2 is hence $8\varepsilon = (2-\sqrt{2}) \approx 0.59$.

Since this is smaller than 1, these correlations violate the Bell inequality.

(f) Consider the state $\sigma = p|\Phi_+\rangle\langle\Phi_+| + (1-p)\frac{\mathbb{1}}{4}$. It has the form in the question where $\varepsilon = \frac{p}{8}(2-\sqrt{2}) + \frac{1-p}{4} = \frac{1}{8}(2-\sqrt{2}p)$, or $p = \sqrt{2}(1-4\varepsilon)$. You may alternatively have expressed this in terms of $p' = 1-p$, leading to $p' = 1-\sqrt{2}(1-4\varepsilon) = (1-\sqrt{2}) + 4\sqrt{2}\varepsilon$.

(g) Since $I_2 = 8\varepsilon$, we have $I_2 < 1$ for $\varepsilon < \frac{1}{8}$. $D(|\Phi_+\rangle\langle\Phi_+|, \sigma) = \frac{1}{2}\mathrm{tr}||\Phi_+\rangle\langle\Phi_+| - \sigma|$. However, both states are diagonal in the same (Bell) basis, so we can calculate the classical distance between the distributions $(1,0,0,0)$ and $(p+(1-p)/4, (1-p)/4, (1-p)/4, (1-p)/4)$. This is $1-p-(1-p)/4 = 3(1-p)/4$. For $\varepsilon = \frac{1}{8}$, we have $p = \sqrt{2}(1-4\varepsilon) = \frac{1}{\sqrt{2}}$, so we obtain distance $\frac{3}{4}(1-\frac{1}{\sqrt{2}}) \approx 0.22$.

B. Solution 2

$\langle A_0 \otimes B_1\rangle = \sum_{xy} xyP_{XY|0,1}(x,y) = P_{XY|01}(1,1) - P_{XY|01}(-1,1) - P_{XY|01}(1,-1) + P_{XY|01}(-1,-1) = P(X=Y|0,1) - P(X \neq Y|0,1)$

Similarly for the others, hence, noting that $P(X = Y|0,1) - P(X \neq Y|0,1) = 2P(X = Y|0,1) - 1 = 1 - 2P(X \neq Y|0,1)$,

$$\begin{aligned}
&\langle A_0 \otimes B_1\rangle + \langle A_2 \otimes B_1\rangle + \langle A_2 \otimes B_3\rangle - \langle A_0 \otimes B_3\rangle \\
&\qquad = 3 - (P(X \neq Y|0,1) + P(X \neq Y|2,1) + P(X \neq Y|2,3)) + 1 - 2P(X = Y|0,3) \\
&\qquad = 4 - 2I_2 \\
&\qquad \leq 2\,.
\end{aligned}$$

In the quantum case, we instead have $4-2I_2 \leq 4-2(2-\sqrt{2}) = 2\sqrt{2}$, so for quantum correlations

$$\langle A_0 \otimes B_1\rangle + \langle A_2 \otimes B_1\rangle + \langle A_2 \otimes B_3\rangle - \langle A_0 \otimes B_3\rangle \leq 2\sqrt{2}\,.$$