Quantum Information Processing

Quantum Information Theory (III)

Richard Cleve

Institute for Quantum Computing & Cheriton School of Computer Science

University of Waterloo

December 5, 2021

Abstract

The goal of these notes is to explain the basics of quantum information processing, with intuition and technical definitions, in a manner that is accessible to anyone with a solid understanding of linear algebra and probability theory.

These are lecture notes for the third part of a course entitled "Quantum Information Processing" (with numberings QIC 710, CS 768, PHYS 767, CO 681, AM 871, PM 871 at the University of Waterloo). The other parts of the course are: a primer for beginners, quantum algorithms, and quantum cryptography. The course web site `http://cleve.iqc.uwaterloo.ca/qic710` contains other course materials, including video lectures.

I welcome feedback about errors or any other comments. This can be sent to `cleve@uwaterloo.ca` (with "Lecture notes" in subject, if at all possible).

Contents

1 Nonlocality

In this section, I will explain a phenomenon called *nonlocality*, which is a strange kind of behavior that quantum systems can exhibit. One way of explaining this behavior is in terms of games played by a team of cooperating players. They players must individually answer certain questions, and they must do this without communicating with each other. The lack of communication appears to restrict what the players can achieve. However, with quantum information, strange behaviors can occur, that defy what one might intuitively think is possible.

1.1 Entanglement and signalling

First of all, let's note that entangled states cannot be used to communicate instantaneously. What I mean by this is the following. Suppose that Alice has a quantum system in her lab and Bob has a quantum system in his lab, and the two labs are in physically separate locations. Alice and Bob's systems might be in an entangled state—for example one of the Bell states.

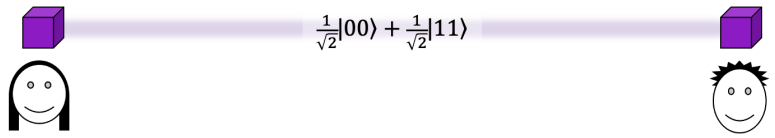

Figure 1: Is there anything Alice can do to *her* qubit that is detectable in Bob's qubit?

Then there is no quantum operation that Alice can perform on her system alone, whose effect is detectable by Bob. If Alice wants to communicate with Bob, she has to send something to him

To see why this is so, consider the density matrix of Bob's system (that is, the density matrix that results when Alice's system is traced out). If Alice performs a unitary operation or measurement on her system then this has no effect on the density matrix of Bob's system. So any kind of measurement that Bob performs on his system will have exactly the same outcome whether or not Alice performs the operation.

And this is not specific to two systems. If there are three or more parties then the same thing holds. Operations on one system have no detectable effect on the other systems.

1.2 GHZ game

Now let's consider a three-player game, commonly referred to as the GHZ game (named after it's inventors, Greenberger, Horne, and Zeillinger).

Let's call the three players Alice, Bob, and Carol.

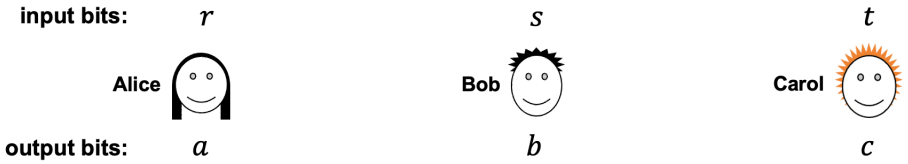

Figure 2: Alice, Bob, and Carol playing the GHZ game.

This game works as follows. Each player receives a 1-bit input. Call the respective input bits r, s, and t. And each player is required to produce a 1-bit output. Call the respective input bits a, b, and c.

The rules of this game are the following.

1. It is promised that the input bits, r, s, and t have an even number of 1s among them. In other words, $r \oplus s \oplus t = 0$. So there are actually three cases of inputs: 000, 011, 101, 110.

2. There is no communication allowed between Alice, Bob, and Carol once the game starts and their input bits are received. So, for example, although Alice will know what her input r is, she does not know what s and t are.

3. This rule defines the *winning conditions* of the output bits. The three players *win* the game if and if $a \oplus b \oplus c = r \vee s \vee t$.

The winning condition $a \oplus b \oplus c = r \vee s \vee t$ is just a condensed way of describing this table.

rst	$a \oplus b \oplus c$
000	0
011	1
101	1
110	1

Figure 3: For each input rst, the required value of $a \oplus b \oplus c$ to win.

For the first case, the XOR of the outputs should be 0 for the players to win. For the other three cases, the XOR of the outputs should be 1 for the players to win.

To get a feeling for this game, here is an example of a strategy that Alice, Bob and Carol could use:

Example of a strategy

Alice receives r as input and produces r as output.

Bob receives s as input and produces $\neg s$ as output.

Carol receives t as input and produces 1 as output.

So how well does this strategy perform? Here I've added the output bits arising from this strategy to the table.

rst	$a \oplus b \oplus c$	abc
000	0	011
011	1	001
101	1	111
110	1	101

Figure 4: Output bits produced by the example strategy (in red).

You can see that the first output bit a is r, the second output bit b is $\neg s$, and the third output bit c is always 1. Consider the first case of inputs 000. There the XOR of the output bits should be 0. And it is 0. So they win in that case. For the second case, the XOR of the output bits should be 1, and it is. So they win in that case too. They also win in the third case. So far, so good. But what happens in the fourth case? In that case, the XOR should be 1. But the output bits have XOR 0. So they lose in that case.

So this is an example of a strategy that wins in three out of the four cases. Is there a better strategy, that wins in all four cases?

1.2.1 Is there a perfect strategy for GHZ?

Call a strategy that wins in every case a *perfect* strategy. Let's see if we can find a perfect strategy for this game.

Alice's output bit is a function of her input bit. Let a_0 denote her output bit if her input bit is 0. And let a_1 denote her output bit if her input bit is 1. Similarly,

define b_0, b_1 as Bob's output bits in the two cases, and c_0, c_1 as Carol's output bits in the two cases. So we have six bits specifying a strategy.

We can express the winning conditions in terms of the four equations

$$a_0 \oplus b_0 \oplus c_0 = 0 \tag{1}$$
$$a_0 \oplus b_1 \oplus c_1 = 1 \tag{2}$$
$$a_1 \oplus b_0 \oplus c_1 = 1 \tag{3}$$
$$a_1 \oplus b_1 \oplus c_0 = 1. \tag{4}$$

The first equation says that, when all three inputs are 0, the XOR of the output bits should be 0. The second equation says that, when the input bits are 011, the XOR of the output bits should be 1. And so on.

So, to find a perfect strategy, we just have to solve this system of equations. Is there a solution?

In fact, there is no solution to this system of equations. These are linear equations in mod 2 arithmetic. Suppose we add the four equations. On the left side we get 0, because each variable appears exactly twice. On the right side, we get the XOR of three 1s, which is 1. So, summing the equations yields $0 = 1$, a contradiction.

It's possible to satisfy any three of the four equations, but not all four. Therefore, there does not exist a perfect deterministic strategy.

I say *deterministic*, because this analysis doesn't consider the case of probabilistic strategies. Could there exist a probabilistic perfect strategy?

There cannot exist a probabilistic perfect strategy either. This is because a probabilistic strategy is essentially a probability distribution over all the deterministic strategies, and the success probability is a weighted average of all the success probabilities of the deterministic strategies (weighted by the probabilities).

If the questions r, s, and t were selected randomly (with probability $\frac{1}{4}$ for each possible input) then the success probability for every deterministic strategy would be at most $\frac{3}{4}$. So the weighted average for any probability distribution on deterministic strategies cannot be higher than that.

Now imagine that you actually carried out this game with Alice, Bob, and Carol. You generate a random triple of questions and check whether their answers win or not. If you just play this once then they might win by luck. In fact, they can win with probability $\frac{3}{4}$. But suppose you play this several rounds in succession and they win every single round?

What if you play four rounds, once for each of the four input possibilities? Will the players necessarily fail in at least one of those rounds? No, not necessarily. Note

that the players might use a different deterministic strategy at each round. Their strategy can satisfy any three of the four equations, and they can arrange to have a different equation violated for each round.

In fact, the player can ensure that they win each round with probability $\frac{3}{4}$. So they would win any four rounds with probability $(\frac{3}{4})^4$, which is slightly more than 30%.

On the other hand, it's highly unlikely that the players would be lucky enough to win, say, 500 rounds. That success probability is $(\frac{3}{4})^{500}$, which less than 1 in a trillion trillion trillion trillion trillion.

So if you did play the game for a large number of rounds—with randomly selected questions at each round—and the players won *every single time*, what would you make of that?

1.2.2 Cheating by communicating

What makes the game non-trivial is that the players cannot communicate with each other. If they could communicate then there would be an easy perfect strategy. For example, suppose that Bob sent his input bit s to Alice. Then Alice knows r and s. She can also deduce t because of the condition that the parity of all three bits is promised to be 0. So Alice knows which of the four cases they're in. A winning strategy is for Alice to output the required parity to win, and Bob and Carol to output 0.

rst	$a \oplus b \oplus c$	abc
000	0	000
011	1	100
101	1	100
110	1	100

Figure 5: Output bits of the cheating-by-communicating strategy (in red).

This wins in all four cases. And there are ways of scrambling up the outputs that obscures this pattern—so that Bob and Carol don't always output 0, and yet the three players always win.

1.2.3 Enforcing no communication

So how can we ensure that they do not communicate? Maybe they each have a very well-concealed transmitter.

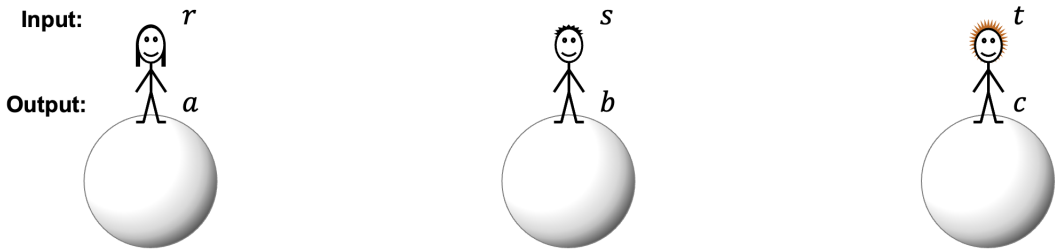

Figure 6: Alice, Bob, and Carol physically far apart.

You could ensure that there's a significant physical distance between the players (here I'm depicting them on separate planets), and also time their inputs and outputs tightly so that they cannot communicate fast enough for messages to reach each other before their deadlines for producing their outputs. For this, we assume that they cannot send signals faster than the speed of light. In physics-terminology we are making the input/output events *space-like separated*.

Then, assuming that the theory of relativity is correct, we can be sure they are not communicating their inputs to each other. But what if this is done and they *still* keep on winning, for every round?

If they players had quantum systems that were entangled, could that possibly help? Recall (as explained in section 1.1) that entanglement does not enable communication. There's no way that Bob can perform an operation on his system that can be detected by Alice. So is there any possible way that Alice, Bob, and Carol could keep on winning this game? If you're not already familiar with scenarios like this then I recommend that pause and think this over. The answer will come at the next page.

1.2.4 The "mystery" explained

The answer is yes, there is a way that they can always win, and I will show you how. They do use entanglement. Let them share the 3-qubit state

$$\tfrac{1}{2}|000\rangle - \tfrac{1}{2}|011\rangle - \tfrac{1}{2}|101\rangle - \tfrac{1}{2}|110\rangle \tag{5}$$

Alice possesses the first qubit, Bob possesses the second qubit and Carol possesses the third qubit.

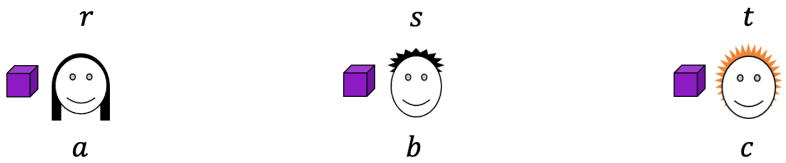

Figure 7: Alice, Bob, and Carol each possess a qubit of a tripartite entangled state.

First, I'll describe the strategy of the players.

Entangled strategy

Alice: if $r = 1$ then apply H; measure and output the result.

Bob: if $s = 1$ then apply H; measure and output the result.

Carol: if $t = 1$ then apply H; measure and output the result.

Now let's see how this strategy performs. There are four cases of inputs.

Let's begin by considering the first case, where $rst = 000$. In that case, neither player applies a Hadamard transform. Therefore, they measure the state in Eq. (5) with respect to the computational basis. The result is

$$\begin{cases} 000 & \text{with prob. } \frac{1}{4} \\ 011 & \text{with prob. } \frac{1}{4} \\ 101 & \text{with prob. } \frac{1}{4} \\ 110 & \text{with prob. } \frac{1}{4}. \end{cases} \tag{6}$$

For all four possibilities, the XOR of the three output bits is 0, which is what it's supposed to be for the 000 case.

Now let's consider the case where $rst = 011$. In that case, Bob and Carol apply Hadamard operations. It's straightforward to check that

$$\begin{aligned}\big(I \otimes H \otimes H\big)\big(\tfrac{1}{2}\,|000\rangle - \tfrac{1}{2}\,|011\rangle - \tfrac{1}{2}\,|101\rangle - \tfrac{1}{2}\,|110\rangle\big) \\ = \tfrac{1}{2}\,|001\rangle + \tfrac{1}{2}\,|010\rangle - \tfrac{1}{2}\,|100\rangle - \tfrac{1}{2}\,|111\rangle\end{aligned} \tag{7}$$

and when this state is measured in the computational basis the result is

$$\begin{cases} 001 & \text{with prob. } \frac{1}{4} \\ 010 & \text{with prob. } \frac{1}{4} \\ 100 & \text{with prob. } \frac{1}{4} \\ 111 & \text{with prob. } \frac{1}{4}. \end{cases} \tag{8}$$

So the XOR of the three output bits is 1, as required for that case.

The cases where $rst = 101$ and 110 are similar to the previous case, due to the symmetry of the state and the strategies. That's how Alice, Bob, and Carol can win with probability 1.

If they play several rounds of the game then they need to possess several copies of the entangled state in Eq. (5), and they consume one copy during each round.

1.2.5 Is the entangled strategy communicating?

So how is this entangled strategy working? Is it somehow communicating? If you look at the outcome distributions for the different cases, you can see that each individual output bit is an unbiased random bit. So the result of Alice's measurement contains absolutely no information about Bob and Carol's inputs. And similarly for the other players. In fact it can be shown that any perfect strategy using entanglement must have the property that each output bit by itself is a random unbiased bit.

Even if we consider pairs of output bits, they are uncorrelated random bits. It's only the *tripartite* correlations among all three output bits that contain information about the inputs.

1.2.6 GHZ conclusions

Let's summarize this GHZ game. It's a game played by a team of three cooperating players who cannot communicate with each other once the game starts. They each receive a bit as their input and are required to produce a bit as their output. There is

a well-defined winning condition for the output bits that depends on what the input bits are.

The following conditions hold:

- Any classical team can succeed with probability at most $\frac{3}{4}$.
- Allowing the players to communicate would enable them to boost their success probability to 1.
- Entanglement cannot be used to communicate.
- Nevertheless, entanglement is another way that the players can boost their success probability to 1. But not by using entanglement to communicate.
- Instead, entanglement enables the measurement outcomes to be correlated in ways that are impossible with classical information.

You might wonder why I showed you a three-player game. Are there two-player games for which the same phenomena occurs? I showed you a three-player game because it's the simplest game that that I'm aware of that illustrates the point.

1.3 Magic square game

Here's an example of a two-player game with a property similar to that of the GHZ game: that there is no perfect classical strategy, whereas there is a perfect strategy using entanglement. It's commonly called the *Magic Square Game* and I will give just a broad overview (without explaining how the entangled strategy works).

A good way of understanding how the Magic Square Game is defined is to first consider the following puzzle. Imagine a 3-by-3 array whose entries are bits.

b_1	b_2	b_3
b_4	b_5	b_6
b_7	b_8	b_9

Figure 8: Are there bits with even parity for each row and odd parity for each column?

Can you find values for the bits such that:

(a) the number of 1s in every row is even; and

(b) the number of 1s in every column is odd?

Please pause to think about how to do this.

You may have come to the realization that it is impossible to do this. Why? Consider the number of 1s among all the nine bits. The row condition implies that there are an odd number of 1s in total. The column condition implies that there are an even number of 1s in in total. This is a contradiction.

Now, keep this in mind, while I describe a two-player game. As with the three-player GHZ game, the players are collaborating and cannot communicate with each other once the game starts. Alice and Bob each receive a trit as input and are each required to return three bits. Think of Alice's input as specifying a row of the array, and think of Bob's input as specifying a column of the array.

Figure 9: Framework for playing the Magic Square Game.

The winning conditions are:

1. Alice's 3-bit output has even parity (think of these as the bits of one of the rows of the array).

2. Bob's 3-bit output has odd parity (think of these as the bits of one of the columns of the array).

3. Alice and Bob's outputs are consistent in the sense that, where Alice's row intersects Bob's column, the bits are the same. (For example, if Alice is queried the second row and Bob is queried the third column then Alice's third bit must be the same as Bob's second bit.)

What can we say about this game?

It turns out that there is no perfect classical strategy for this game. Of the nine possible question pairs, Alice and Bob's strategy must fail for at least one of them. The maximum success probability attainable is $\frac{8}{9}$. The proof of this is based on the fact that there is no way to set the bits of the 3-by-3 array that satisfy the parity conditions.

But there is a perfect entangled strategy that uses two Bell states as entanglement. It's a very interesting strategy, but I won't go into the details of it here.

1.4 Are nonlocal games useful?

So far, you might think that these games are weird curiosities, that have no conceivable application. But these games can be useful for enforcing certain kinds of behavior in cryptographic protocols. A simple example of this involves devices for generating random bits.

Suppose that you purchased such a device that purportedly generates a stream of random bits.

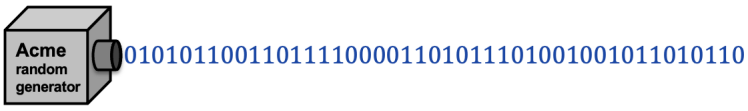

Figure 10: A supposed random number generator. Can you trust it?

How can you know that this is a good device? You could open the box and see if the internal mechanism looks legitimate. But, even if superficially it looks like some process that you think is generating randomness, it's possible that it's a fake random generator, and that the manufacturer is trying to trick you.

What the manufacturer could do is produce a long sequence of random bits in the factory and store those random bits in two memory devices, and hide one of those memory devices somewhere in the box and keep the other memory. And what the device could actually be doing is just outputting the bits stored in the memory device.

Note that the output would look like random bits to you. But the manufacturer would have a copy of those bits. They would know exactly what the next output bit is going to be. If you use such a fake random generator in a cryptographic context, for example to generate a random secret key, that that could be trouble. The manufacturer would know the key.

Unfortunately, in the context of classical information, there is no remedy for this, even in principle. If you don't trust the manufacturer of your devices, then there's no way that you can be sure that it's really generating random bits, that are unknown to the manufacturer.

But, with *quantum* devices it's possible to certify the randomness of untrusted devices. The "device" would actually consists of two components, that contain entangled quantum systems.

Figure 11: Generating random strings using untrusted devices.

You physically separate the two components and input a short random seed into each component. From this, each component outputs a long string of bits. You check if the inputs and outputs satisfy a certain function, called a test. If they pass the test, and if the input/output events are space-like separated, then you know for sure that the outputs are really random bits (within some precision ϵ, for some small $\epsilon > 0$).

I'm not claiming that such a system is practical to implement for wide usage. But it shows that, in principle, it's possible to actually certify randomness using quantum information. Something that's impossible, even in principle, with classical information.

Nonlocal games also have profound implications in the foundations of physics, which will be the topic of the next section.

2 Bell/CHSH inequality

This section is about the Bell inequality in physics, and its violation. The version that we'll consider was discovered shortly after John Bell's ground-breaking paper, and is called the CHSH inequality, after its authors, Clauser, Horne, Shimoney, and Holt.

2.1 Fresh randomness vs. stale randomness

I'd like to begin by making a distinction between "fresh" randomness and "stale" randomness. By *stale randomness*, I mean something like this. Suppose that I flipped a coin yesterday and I know what the outcome was, but I'm not telling you what it was. Then, from *your* perspective, the outcome is a probability distribution. From your perspective, outcome is "heads" with probability $\frac{1}{2}$ and "tails" with probability $\frac{1}{2}$. But the outcome is already determined. Your probabilities just reflect a lack of information on your part.

Contrast this situation with the case where the coin is spinning in the air right at this very moment. In that case, neither of us know the outcome. The outcome has not been determined yet. Let's call that *fresh randomness*.

But is a coin flip really a random process? Isn't the outcome determined by the present conditions? If we knew the exact shape of the coin, it's exact motion, and the positions of all the air molecules *and* we had an extremely powerful computer then maybe we could determine the value of the coin flip while it's spinning in the air.

Moreover, we should not conflate *forecasting* (being able to predict a future event) with *determinism* (a future event being determined by present conditions).

Consider the weather. Predicting, say, the outside temperature where I live one year for now is for all practical purposes impossible due to the chaotic nature of weather (the so-called *butterfly effect*). In terms of forecasting, this temperature is at best a probability distribution (a different distribution in the summer than in the winter). Nevertheless, it seems that the *precise* future weather is determined by the *precise* present conditions, even if we cannot how exactly what these are.

Whether the intuitive notion of fresh randomness actually exists or is just an illusion is an interesting question. I don't claim to know the answer. All of this discussion is a lead-in to the question:

> *If a qubit in state* $\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$ *is measured in the computational basis, will the outcome be fresh randomness or stale randomness?*

In the quantum information framework that has been the subject of these notes, the outcome is regarded as fresh randomness, that's spontaneously generated during the measurement process. It doesn't really make sense in our model for Alice to produce a qubit in state $\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$ and at the same time to know in advance the outcome of a future measurement (in the computational basis) of that state. Or does it?

2.2 Predetermined measurement outcomes of a qubit?

Let's explore the possibility that the outcomes for measuring a qubit in a state like $\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$ are predetermined. Think of a qubit as a physical entity, a particle (technically, a spin-$\frac{1}{2}$ particle) that was created at the big bang. Imagine that, *at the time of creation*, a predetermined outcome for each possible measurement outcome was embedded into the particle. So lurking within a qubit is some sort of table of predetermined outcomes. Let's visualize it as a literal table of outcome values.

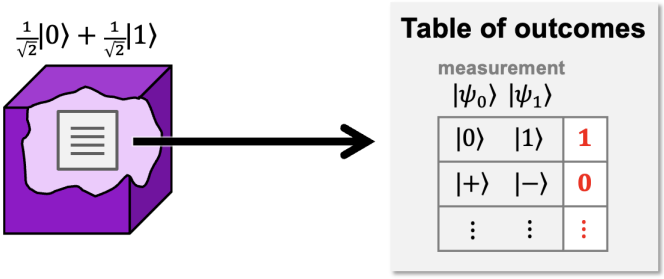

Figure 12: Are predetermined values of all measurement outcomes contained in a qubit?

Imagine that every entry of the table was created randomly so as to conform with the outcome probabilities of quantum measurements. For example, for a measurement in the computational basis, the outcome is an unbiased random bit. So, in that entry of the table, a random bit was inserted (in the figure, it was set to 1). On the other hand, for a measurement in the $|+\rangle/|-\rangle$ basis, the outcome should always be the first state $|+\rangle$. So that entry of the table was set the bit 0. And so on. For every other potential measurement, there is an entry in the table containing a bit that is sampled with the appropriate probability distribution for that measurement of the state. That's an infinitely large table. Maybe there's a compressed way of containing this information, but let's not concern ourselves with that issue. My point is that it's conceivable that the particle contains this table of predetermined measurement outcomes stored within it.

In physics, these are called *hidden variables*. The idea is that these hidden variables represent additional physical properties of systems that are yet to be discovered. When they are discovered, quantum mechanics will be tamed of its randomness. The randomness that arises in quantum theory as it currently exists could merely be a consequence of the fact that we don't know what these hidden variables are. In this way of thinking, a measurement merely extracts a predetermined value from the table of outcomes.

Let's continue developing this model. What happens if we apply a unitary operation to a qubit? This would rearrange the table of outcomes in some systematic way. For example, suppose that we apply a Hadamard transform to the qubit. That would swap the first two bits of the table. This is because, after applying a Hadamard to this $|+\rangle$ state, the state becomes $|0\rangle$, so now a measurement in the computational basis produces 0 for sure. And measuring in the $|+\rangle/|-\rangle$ basis is what produces a random bit. Without going into the details, the effect of any unitary operation can be captured by moving around the entries of the table of outcomes. Unitary operations

merely rearrange the stale randomness of the table of outcomes.

And, in this picture, every spin-$\frac{1}{2}$ particle has it's own separate table. If multiple particles are in the $|+\rangle$ state then each one contains an independent random bit for the first entry in its table. This is consistent with what happens when we measure several qubits that are each in the $|+\rangle$ state.

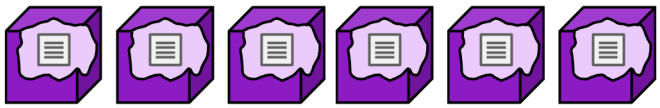

Figure 13: Multiple qubits, each containing a "table" of hidden variables.

What's interesting about this local hidden variables picture is that, so far, everything is consistent with quantum behavior.

We might imagine that this model can be extended to capture all of quantum information theory. For example, when a 2-qubit unitary gate entangles two particles, what might actually be happening is a rearrangement of both tables of outcomes, so that the entries are appropriately correlated for all possible measurements. If the unitary operation creates the state $\frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$ then the table entries for measuring each qubit in the computational basis should be: 0 for both particles with probability $\frac{1}{2}$; and 1 for both particles with probability $\frac{1}{2}$.

Let's continue exploring how a local hidden variable model would work.

2.3 CHSH inequality

Imagine a system consisting of two qubits (or two particles), and that there are two measurements, that we'll refer to as M_0 and M_1. We're supposing that each particle contains a full tables of outcomes, but here we'll only care about the parts of the tables that are associated with the measurements M_0 and M_1.

Figure 14: Two particles, with their hidden variables for two measurements, M_0 and M_1.

Call the two predetermined values for the first particle a_0 and a_1. And call the two predetermined values for the second particle b_0 and b_1.

Note that we making an assumption that the hidden variables are *local* hidden variables in the sense that each particle's predetermined outcomes depend only on the measurement performed on that particle, and not the measurements performed on the other particles. What's the justification for this?

The justification is that the particles might be far apart from each other and the timing of the measurements might be such that the two measurement events are space-like separated. This means that there isn't sufficient time for a signal to go from the second particle to the first particle with the information about which measurement is performed. For space-like separated measurement events, the first particle has no way of "knowing" what measurement is being performed on the second particle, even in principle. Therefore, for space-like separated measurements, it's impossible for one particle's outcome to depend on what measurement is performed on the other particle.

I will now describe a property that the bits a_0, a_1, b_0, b_1 must satisfy. It is convenient to describe this property in terms of bits that are expressed as $+1$ and -1 instead of 0 and 1. So we define such a conversion, into "uppercase" bits as

$$A_0 = (-1)^{a_0} \tag{9}$$

$$A_1 = (-1)^{a_1} \tag{10}$$

$$B_0 = (-1)^{b_0} \tag{11}$$

$$B_1 = (-1)^{b_1}. \tag{12}$$

I claim that inequality (13) holds, which is called the CHSH inequality.[1]

Theorem 2.1 (CHSH inequality)**.** *For any $A_0, A_1, B_0, B_1 \in \{+1, -1\}$, it holds that*

$$A_0B_0 + A_0B_1 + A_1B_0 - A_1B_1 \leq 2. \tag{13}$$

Note that each of the four terms on the left side can be $+1$ or -1. So we might imagine that the left side can be as large as 4. But the upper bound is 2.

Proof of Theorem 2.1. Let's see how to prove the CHSH inequality. We can write the left side of Eq. (13) as

$$A_0(B_0 + B_1) + A_1(B_0 - B_1). \tag{14}$$

[1]The original inequality along these lines is due to John Bell in 1964. All subsequent variations of it are loosely called *Bell inequalities*. This particularly nice version is due to Clauser, Horne, Shimony, and Holt and is also called the CHSH inequality.

Now consider the expressions in the parentheses, $B_0 + B_1$ and $B_0 - B_1$. Either B_0 and B_1 have the same sign or they have different signs. If B_0 and B_1 have the same sign then $B_0 + B_1$ can be as large as 2, but then $B_0 - B_1 = 0$. So in that case, the upper bound is 2. If B_0 and B_1 have the different signs then $B_0 - B_1$ can be as large as 2, but then $B_0 + B_1 = 0$. So in that case, the upper bound is also 2. This completes the proof. □

Why should we care about this CHSH inequality? The reason why is that the inequality can be experimentally tested, and if an experiment shows that it's violated then the possibility of a local hidden variable model is refuted.

First of all, let's consider how one could in principle design an experiment to verify that systems satisfy the CHSH inequality. There's some subtlety with this. The problem is that, for any single measurement of the two particles, only one of the four A_sB_t-terms in the inequality can be measured. Here again are the particles with their outcome tables, where I've taken the liberty of writing the outcomes with uppercase bits (in the ± 1 language).

Figure 15: Two particles, with their hidden variables for M_0 and M_1 specified as $\pm$ bits.

You can choose to perform any single measurement on the first system and get either A_0 or A_1 and then the state is disturbed, so you cannot measure again to get the original value of the other bit. Similarly, you can choose any single measurement on the second system and get B_0 or B_1 (but not both). So you can acquire only one of the four terms A_0B_0, A_0B_1, A_1B_0, A_1B_1 (whose value is $+1$ or -1). To verify the inequality, you would need to see all four terms.

However, the Bell inequality *can* be verified using statistical methods, by making several independent runs, using a separate pair of particles for each run. In each run, $st \in \{00, 01, 10, 11\}$ is chosen randomly and the $\pm$ bit $(-1)^{st}A_sB_t$ is calculated. The factor $(-1)^{st}$ means multiply by -1 in the $st = 11$ case.

If $st \in \{00, 01, 10, 11\}$ is sampled randomly according to the uniform distribution then the *expected value* of $(-1)^{st} A_s B_t$ is

$$\mathrm{E}_{s,t}\big[(-1)^{st} A_s B_t\big] = \tfrac{1}{4} A_0 B_0 + \tfrac{1}{4} A_0 B_1 + \tfrac{1}{4} A_1 B_0 - \tfrac{1}{4} A_1 B_1. \tag{15}$$

Does this expression look familiar? It's the left side of the CHSH inequality divided by 4. So we can deduce from the CHSH inequality that

$$\mathrm{E}_{s,t}\big[(-1)^{st} A_s B_t\big] \leq \tfrac{1}{2}. \tag{16}$$

The experiment to statistically verify the CHSH inequality is to make many separate runs, each on a separate pair of particles, of the procedure where you pick a random $st \in \{00, 01, 10, 11\}$ and then measure M_s and M_t to create a sample $(-1)^{st} A_s B_t \in \{+1, -1\}$. If local variables exist then the average over many runs should converge to $\frac{1}{2}$ or less.

Note that, in order to eliminate the possibility of a hidden variable model that is *not local*, the experiment should be implemented so that each pair of measurement events is space-like separated. If they are not space-like separated then the experiment does not eliminate the possibility that nature is behaving in a conspiratorial way, with signaling between pairs of particles, that permits the outcomes for each particle to depend on both measurements.

The fact that the CHSH inequality can be experimentally verified is remarkable because ...

2.4 Violating the CHSH inequality

... quantum systems can violate the CHSH inequality!

To see how, suppose that the two physically separated qubits are entangled in the Bell state $\frac{1}{\sqrt{2}} |00\rangle - \frac{1}{\sqrt{2}} |11\rangle$.

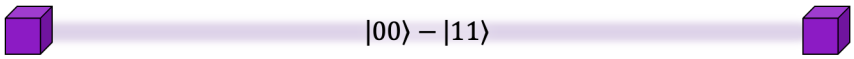

Figure 16: Two physically separated particles in the Bell state $\frac{1}{\sqrt{2}} |00\rangle - \frac{1}{\sqrt{2}} |11\rangle$.

Consider what happens if a rotation is performed on each qubit, by angle θ_s for the first qubit and θ_t for the second qubit.

Exercise 2.1 (straightforward)**.** *Check that, if* $R(\theta_s) \otimes R(\theta_t)$ *is applied to state* $\frac{1}{\sqrt{2}}|00\rangle - \frac{1}{\sqrt{2}}|11\rangle$ *then the result is*

$$\cos(\theta_s+\theta_t)\big(\tfrac{1}{\sqrt{2}}|00\rangle - \tfrac{1}{\sqrt{2}}|11\rangle\big) + \sin(\theta_s+\theta_t)\big(\tfrac{1}{\sqrt{2}}|01\rangle + \tfrac{1}{\sqrt{2}}|10\rangle\big). \tag{17}$$

If the state in Eq. (17) is measured in the computational basis then the outcome bits $(a_s, b_t \in \{0,1\})$ satisfy

$$\Pr[a_s \oplus b_t = 0] = \cos^2(\theta_s+\theta_t) \tag{18}$$
$$\Pr[a_s \oplus b_t = 1] = \sin^2(\theta_s+\theta_t). \tag{19}$$

It follows that, for the $\pm$ bits $A_s = (-1)^{a_s}$ and $B_t = (-1)^{b_t}$,

$$\begin{aligned}\mathrm{E}[A_sB_t] &= \cos^2(\theta_s+\theta_t) - \sin^2(\theta_s+\theta_t)\\ &= \frac{1+\cos\big(2(\theta_s+\theta_t)\big)}{2} - \frac{1-\cos\big(2(\theta_s+\theta_t)\big)}{2}\\ &= \cos\big(2(\theta_s+\theta_t)\big).\end{aligned} \tag{20}$$

Now define the measurements M_0 and M_1 as follows.

- M_0: rotate by $\theta_0 = -\frac{\pi}{16}$ and then measure in the computational basis.
- M_1: rotate by $\theta_1 = +\frac{3\pi}{16}$ and then measure in the computational basis.

Let's look at the various angles $\theta_s + \theta_t$ that arise for $st \in \{00, 01, 10, 11\}$.

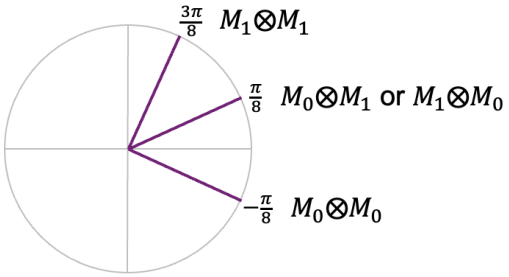

Figure 17: Angles $\theta_s + \theta_t$ that arise for $M_s \otimes M_t$, for the cases $st \in \{00, 01, 10, 11\}$.

When M_0 is performed on both sides, $\theta_0 + \theta_0 = -\frac{\pi}{8}$. When M_0 is performed on one side and M_1 on the other side, $\theta_0 + \theta_1 = \theta_1 + \theta_0 = +\frac{\pi}{8}$. And when M_1 is performed on both sides, $\theta_1 + \theta_1 = \frac{3\pi}{8}$.

Applying Eq. (20), this means that, for measurements M_s and M_t, the $\pm$ outcomes A_s and B_t have the property that

$$\mathrm{E}[A_s B_t] = \begin{cases} \cos(\pm\frac{\pi}{4}) & \text{if } st \in \{00, 01, 10\} \\ \cos(\frac{3\pi}{4}) & \text{if } st = 11 \end{cases} \tag{21}$$

$$= \begin{cases} \frac{1}{\sqrt{2}} & \text{if } st \in \{00, 01, 10\} \\ -\frac{1}{\sqrt{2}} & \text{if } st = 11. \end{cases} \tag{22}$$

It follows that

$$\mathrm{E}[(-1)^{st} A_s B_t] = \tfrac{1}{2}\sqrt{2}, \tag{23}$$

which clearly violates the CHSH inequality—explicitly the upper bound in Eq. (16).

So if we performed the aforementioned experiment of repeatedly picking a random $st \in \{00, 01, 10, 11\}$ and applying the measurements M_s and M_t to a pair of qubits in state $\frac{1}{\sqrt{2}}|00\rangle - \frac{1}{\sqrt{2}}|11\rangle$ to sample $(-1)^{st} A_s B_t$ then the average would exceed the bound of $\frac{1}{2}$, that was derived under the assumption of local hidden variables. Therefore, in the quantum information framework, local hidden variables cannot exist.

Summary and experimental implementations

Let's summarize the Bell inequality and its violation. Assuming that the measurement outcomes of quantum systems are predetermined by local hidden variables leads to the Bell inequality. But actual quantum quantum systems violate this inequality, by a factor of $\sqrt{2}$. Therefore, quantum systems cannot be based on local hidden variables.

And this behavior of quantum systems has been experimentally verified. The rough idea is to generate two particles in a Bell state and send them out in opposite directions to reach detectors, which are set to measure the particles in various ways.

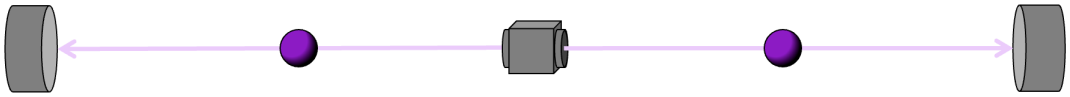

Figure 18: Form of test of the CHSH inequality violation.

In order for such an experiment to be *loophole-free*, the measurement events must be space-like separated; the precision in the state preparation and the measurements must exceed certain thresholds, and the random choices of $st \in \{00, 01, 10, 11\}$ must actually be random. This is non-trivial, but such experiments that are widely regarded as loophole-free have been performed, refuting the existence of local hidden variables.

2.5 Bell/CHSH inequality as a nonlocal game

Now let's look at the Bell/CHSH inequality and its violation in a different way, as a nonlocal game, similar to the ones that we saw in sections 1.2 (the GHZ game) and 1.3 (the Magic Square game).

Define the *CHSH game* as follows. Alice and Bob receive input bits, s and t, and they must produce output bits, a and b.

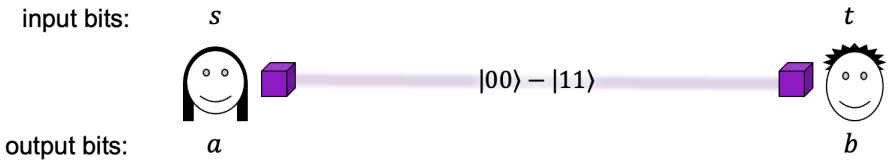

Figure 19: Alice and Bob playing the CHSH game.

The rules of the game are as follows. First, a question pair $st \in \{00, 01, 10, 11\}$ is randomly selected according to the uniform distribution. As usual for nonlocal games, there is no communication allowed between the players once the game starts. And the players *win* if and only if $a \oplus b = s \wedge t$, which is just a condensed way of specifying the following table.

st	$a \oplus b$
00	0
01	0
10	0
11	1

Figure 20: For each input st, the required value of $a \oplus b$ to win.

What's interesting about the CHSH game is that:

- The maximum winning probability for any *classical* strategy is $\frac{3}{4}$.
- There exists an *entangled* strategy that wins with probability

$$\cos^2\left(\tfrac{\pi}{8}\right) = \tfrac{1}{2}\left(1 + \tfrac{1}{\sqrt{2}}\right) = 0.853... \tag{24}$$

This is essentially the CHSH inequality and its violation, but where the outputs are $\{0, 1\}$-bits instead of $\{+1, -1\}$-bits. The upper bound on the winning probability corresponds to the CHSH inequality (in Eq. (16)), and the quantum strategy that

attains success probability $\cos^2\left(\frac{\pi}{8}\right)$ corresponds to the CHSH inequality violation (in section 2.4). However, I will provide a separate analysis of this game.

We can analyze classical strategies for the CHSH game in a manner similar to the way we analyzed the GHZ game in section 1.2.1. First note that any deterministic strategy can be described by four bits, a_0, a_1 (Alice's output bits for the two input possibilities), b_0, b_1 (Bob's output bits for the two input possibilities). And the winning condition can be expressed as the four equations

$$a_0 \oplus b_0 = 0 \tag{25}$$
$$a_0 \oplus b_1 = 0 \tag{26}$$
$$a_1 \oplus b_0 = 0 \tag{27}$$
$$a_1 \oplus b_1 = 1. \tag{28}$$

It's easy to show that at most three of these four equations can be satisfied, so the maximum success probability of any deterministic strategy is $\frac{3}{4}$. Since any classical probabilistic strategy is essentially a probability distribution on the set of all deterministic strategies, its winning probability cannot be higher than $\frac{3}{4}$.

The entangled strategy for the CHSH game whose probability of winning is $\cos^2(\frac{\pi}{8})$ is very similar to the CHSH inequality violation in section 2.4. Alice and Bob use the entangled state $\frac{1}{\sqrt{2}}|00\rangle - \frac{1}{\sqrt{2}}|11\rangle$ and they each perform the following on their qubit.

```
1: if input bit is 0 then
2:     apply R(−π/16) to qubit
3: else if input bit is 1 then
4:     apply R(+3π/16) to qubit
5: end if
6: measure qubit and output result
```

Figure 21: Alice and Bob's local behavior based on their input bit.

From Eqns. (17)(18)(19), we can deduce that, for input bits s and t, Alice and Bob's output bits a and b satisfy

$$\Pr[a \otimes b = s \wedge t] = \begin{cases} \cos^2\left(\pm\frac{\pi}{8}\right) & \text{if } st \in \{00, 01, 10\} \\ \sin^2\left(\frac{3\pi}{8}\right) & \text{if } st = 11 \end{cases} \tag{29}$$

$$= \cos^2\left(\frac{\pi}{8}\right). \tag{30}$$

Bell inequalities and nonlocal games can be thought of as different ways of expressing the same ideas about nonlocality. One perspective is to consider (and refute) the existence of local hidden variables. Another perspective is to consider communication protocols between separated parties, and what they can accomplish with and without the resource of entanglement.