Lecture Notes for Ph219/CS219: Quantum Information and Computation Chapter 4

John Preskill
California Institute of Technology

November 2, 2001

Contents

4
Quantum Entanglement

4.1 Nonseparability of EPR pairs

4.1.1 Hidden quantum information

The deep ways that quantum information differs from classical information involve the properties, implications, and uses of *quantum entanglement.* Recall from §2.4.1 that a bipartite pure state is *entangled* if its Schmidt number is greater than one. Entangled states are interesting because they exhibit correlations that have no classical analog. We will study these correlations in this chapter.

Recall, for example, the *maximally entangled* state of two qubits (or *EPR pair*) defined in §3.4.1:

$$|\phi^+\rangle_{AB} = \frac{1}{\sqrt{2}}\left(|00\rangle_{AB} + |11\rangle_{AB}\right) \ . \tag{4.1}$$

"Maximally entangled" means that when we trace over qubit B to find the density operator $\boldsymbol{\rho}_A$ of qubit A, we obtain a multiple of the identity operator

$$\boldsymbol{\rho}_A = \mathrm{tr}_B(|\phi^+\rangle\langle\phi^+|) = \frac{1}{2}\boldsymbol{I}_A \ , \tag{4.2}$$

(and similarly $\boldsymbol{\rho}_B = \frac{1}{2}\boldsymbol{I}_B$). This means that if we measure spin A along *any* axis, the result is completely random — we find spin up with probability 1/2 and spin down with probability 1/2. Therefore, if we perform any local measurement of A or B, we acquire no information about the preparation of the state, instead we merely generate a random bit. This situation contrasts sharply with case of a single qubit in a pure state; there we can store a bit by preparing, say, either $|\uparrow_{\hat{n}}\rangle$ or $|\downarrow_{\hat{n}}\rangle$, and we can recover that bit reliably by measuring along the $\hat{n}$-axis. With two qubits, we ought to be able to store two bits, but in the state $|\phi^+\rangle_{AB}$ this information is *hidden*; at least, we can't acquire it by measuring A or B.

In fact, $|\phi^+\rangle$ is one member of a basis of four mutually orthogonal states for the two qubits, all of which are maximally entangled — the basis

$$\begin{aligned} |\phi^\pm\rangle &= \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle) \ , \\ |\psi^\pm\rangle &= \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle) \ , \end{aligned} \tag{4.3}$$

introduced in §3.4.1. Imagine that Alice and Bob play a game with Charlie. Charlie prepares one of these four states, thus encoding two bits in the state of the two-qubit system. One bit is the *parity* bit ($|\phi\rangle$ or $|\psi\rangle$): are the two spins aligned or antialigned? The other is the *phase* bit ($+$ or $-$): what superposition was chosen of the two states of like parity. Then Charlie sends qubit A to Alice and qubit B to Bob. To win the game, Alice (or Bob) has to identify which of the four states Charlie prepared.

Of course, if Alice and Bob bring their qubits together, they can identify the state by performing an orthogonal measurement that projects onto the $\{|\phi^+\rangle, |\phi^-\rangle, |\psi^+\rangle, |\psi^-\rangle\}$ basis. But suppose that Alice and Bob are in different cities, and that they are unable to communicate at all. Acting locally, neither Alice nor Bob can collect any information about the identity of the state.

What they *can* do locally is *manipulate* this information. Alice may apply $\boldsymbol{\sigma}_3$ to qubit A, flipping the relative phase of $|0\rangle_A$ and $|1\rangle_A$. This action flips the phase bit stored in the entangled state:

$$\begin{aligned} &|\phi^+\rangle \leftrightarrow |\phi^-\rangle \ , \\ &|\psi^+\rangle \leftrightarrow |\psi^-\rangle \ . \end{aligned} \tag{4.4}$$

On the other hand, she can apply $\boldsymbol{\sigma}_1$, which flips her spin ($|0\rangle_A \leftrightarrow |1\rangle_A$), and also flips the parity bit of the entangled state:

$$\begin{aligned} &|\phi^+\rangle \leftrightarrow |\psi^+\rangle \ , \\ &|\phi^-\rangle \leftrightarrow -|\psi^-\rangle \ . \end{aligned} \tag{4.5}$$

Bob can manipulate the entangled state similarly. In fact, as we discussed in §2.4, either Alice or Bob can perform a local unitary transformation that changes one maximally entangled state to any other maximally entangled state.* What their local unitary transformations *cannot* do is alter

* But of course, this does not suffice to perform an arbitrary unitary transformation on the four-dimensional space $\mathcal{H}_A \otimes \mathcal{H}_B$, which contains states that are not maximally entangled. The maximally entangled states are *not* a subspace — a superposition of maximally entangled states typically is *not* maximally entangled.

$\boldsymbol{\rho}_A = \boldsymbol{\rho}_B = \frac{1}{2}\boldsymbol{I}$ — the information they are manipulating is information that neither one can read.

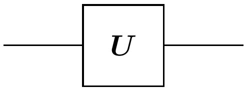
But now suppose that Alice and Bob are able to exchange (classical) messages about their measurement outcomes; together, then, they can learn about how their measurements are correlated. The entangled basis states are conveniently characterized as the simultaneous eigenstates of two commuting observables:

$$\begin{aligned} &\boldsymbol{\sigma}_1^{(A)} \otimes \boldsymbol{\sigma}_1^{(B)} \ , \\ &\boldsymbol{\sigma}_3^{(A)} \otimes \boldsymbol{\sigma}_3^{(B)} \ ; \end{aligned} \tag{4.6}$$

the eigenvalue of $\boldsymbol{\sigma}_3^{(A)} \otimes \boldsymbol{\sigma}_3^{(B)}$ is the parity bit, and the eigenvalue of $\boldsymbol{\sigma}_1^{(A)} \otimes \boldsymbol{\sigma}_1^{(B)}$ is the phase bit. Since these operators commute, they can in principle be measured simultaneously. But they cannot be measured simultaneously if Alice and Bob perform localized measurements. Alice and Bob could both choose to measure their spins along the z-axis, preparing a simultaneous eigenstate of $\boldsymbol{\sigma}_3^{(A)}$ and $\boldsymbol{\sigma}_3^{(B)}$. Since $\boldsymbol{\sigma}_3^{(A)}$ and $\boldsymbol{\sigma}_3^{(B)}$ both commute with the parity operator $\boldsymbol{\sigma}_3^{(A)} \otimes \boldsymbol{\sigma}_3^{(B)}$, their orthogonal measurements do not disturb the parity bit, and they can combine their results to infer the parity bit. But $\boldsymbol{\sigma}_3^{(A)}$ and $\boldsymbol{\sigma}_3^{(B)}$ do *not* commute with phase operator $\boldsymbol{\sigma}_1^{(A)} \otimes \boldsymbol{\sigma}_1^{(B)}$, so their measurement disturbs the phase bit. On the other hand, they could both choose to measure their spins along the x-axis; then they would learn the phase bit at the cost of disturbing the parity bit. But they can't have it both ways. To have hope of acquiring the parity bit without disturbing the phase bit, they would need to learn about the product $\boldsymbol{\sigma}_3^{(A)} \otimes \boldsymbol{\sigma}_3^{(B)}$ without finding out anything about $\boldsymbol{\sigma}_3^{(A)}$ and $\boldsymbol{\sigma}_3^{(B)}$ separately. That cannot be done locally.

Now let us bring Alice and Bob together, so that they can operate on their qubits jointly. How might they acquire both the parity bit and the phase bit of their pair? By applying an appropriate unitary transformation, they can rotate the entangled basis $\{|\phi^\pm\rangle, |\psi^\pm\rangle\}$ to the unentangled basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. Then they can measure qubits A and B separately to acquire the bits they seek. How is this transformation constructed?

This is a good time to introduce notation that will be used heavily in later chapters, the quantum circuit notation. Qubits are denoted by horizontal lines, and the single-qubit unitary transformation $\boldsymbol{U}$ is denoted:



A particular single-qubit unitary we will find useful is the *Hadamard transform*

$$\boldsymbol{H} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{\sqrt{2}}(\boldsymbol{\sigma}_1 + \boldsymbol{\sigma}_3) \ , \tag{4.7}$$

which has the properties

$$\boldsymbol{H}^2 = \boldsymbol{I} \ , \tag{4.8}$$

and

$$\begin{aligned} \boldsymbol{H}\boldsymbol{\sigma}_1\boldsymbol{H} &= \boldsymbol{\sigma}_3 \ , \\ \boldsymbol{H}\boldsymbol{\sigma}_3\boldsymbol{H} &= \boldsymbol{\sigma}_1 \ . \end{aligned} \tag{4.9}$$

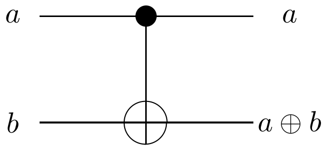
(We can envision $\boldsymbol{H}$ (up to an overall phase) as a $\theta = \pi$ rotation about the axis $\hat{n} = \frac{1}{\sqrt{2}}(\hat{n}_1 + \hat{n}_3)$ that rotates $\hat{x}$ to $\hat{z}$ and vice-versa; we have

$$\boldsymbol{U}(\hat{n},\theta) = \boldsymbol{I}\cos\frac{\theta}{2} + i\hat{n}\cdot\vec{\boldsymbol{\sigma}}\sin\frac{\theta}{2} = i\frac{1}{\sqrt{2}}(\boldsymbol{\sigma}_1 + \boldsymbol{\sigma}_3) = i\boldsymbol{H} \ .) \tag{4.10}$$

Also useful is the two-qubit transformation known as the reversible XOR or controlled-NOT transformation; it acts as

$$\mathbf{CNOT} : |a,b\rangle \rightarrow |a, a\oplus b\rangle \ , \tag{4.11}$$

on the basis states $a, b = 0, 1$, where $a \oplus b$ denotes addition modulo 2. The **CNOT** is denoted:

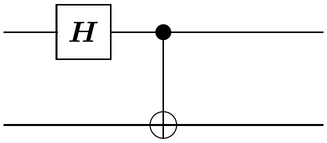


Thus this transformation flips the second bit if the first is 1, and acts trivially if the first bit is 0; it has the property

$$(\mathbf{CNOT})^2 = \boldsymbol{I}\otimes\boldsymbol{I} \ . \tag{4.12}$$

We call a the *control* (or source) bit of the **CNOT**, and b the *target* bit.

By composing these "primitive" transformations, or quantum *gates*, we can build other unitary transformations. For example, the "circuit"



(to be read from left to right) represents the product of $\boldsymbol{H}$ applied to the first qubit followed by **CNOT** with the first bit as the source and the second bit as the target. It is straightforward to see that this circuit transforms the standard basis to the entangled basis,

$$\begin{aligned} |00\rangle &\to \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|0\rangle \to |\phi^+\rangle, \\ |01\rangle &\to \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|1\rangle \to |\psi^+\rangle, \\ |10\rangle &\to \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)|0\rangle \to |\phi^-\rangle, \\ |11\rangle &\to \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)|1\rangle \to |\psi^-\rangle, \end{aligned} \tag{4.13}$$

so that the first bit becomes the phase bit in the entangled basis, and the second bit becomes the parity bit.

Similarly, we can invert the transformation by running the circuit backwards (since both **CNOT** and $\boldsymbol{H}$ square to the identity); if we apply the inverted circuit to an entangled state, and then measure both bits, we can learn the value of both the phase bit and the parity bit.

Of course, $\boldsymbol{H}$ acts on only one of the qubits; the "nonlocal" part of our circuit is the controlled-NOT gate — this is the operation that establishes or removes entanglement. If we could only perform an "interstellar **CNOT**," we would be able to create entanglement among distantly separated pairs, or extract the information encoded in entanglement. But we can't. To do its job, the **CNOT** gate must act on its target without revealing the value of its source. Local operations and classical communication will not suffice.

4.1.2 Einstein locality and hidden variables

Einstein was disturbed by quantum entanglement. Eventually, he along with Podolsky and Rosen (EPR) sharpened their discomfort into what they regarded as a paradox. As later reinterpreted by Bohm, the situation they described is really the same as that discussed in §2.5.3. Given a maximally entangled state of two qubits shared by Alice and Bob, Alice can choose one of several possible measurements to perform on her spin that will realize different possible ensemble interpretations of Bob's density matrix; for example, she can prepare either $\boldsymbol{\sigma}_1$ or $\boldsymbol{\sigma}_3$ eigenstates.

We have seen that Alice and Bob are unable to exploit this phenomenon for faster-than-light communication. Einstein knew this but he was still dissatisfied. He felt that in order to be considered a *complete* description of physical reality a theory should meet a stronger criterion, that might be called *Einstein locality* (also sometimes known as *local realism*):

> Suppose that A and B are spacelike separated systems. Then in a *complete* description of physical reality an action performed on system A must not modify the description of system B.

But if A and B are entangled, a measurement of A is performed and a *particular* outcome is known to have been obtained, then the density matrix of B *does* change. Therefore, by Einstein's criterion, the description of a quantum system by a wavefunction or density operator cannot be considered complete.

Einstein seemed to envision a more complete description that would remove the indeterminacy of quantum mechanics. A class of theories with this feature are called *local hidden-variable theories*. In a hidden-variable theory, measurement is actually fundamentally deterministic, but appears to be probabilistic because some degrees of freedom are not precisely known. For example, perhaps when a spin is prepared in what quantum theory would describe as the pure state $|\uparrow_{\hat{z}}\rangle$, there is actually a deeper theory in which the state prepared is parametrized as $(\hat{z}, \lambda)$ where λ $(0 \leq \lambda \leq 1)$ is the hidden variable. Suppose that with present-day experimental technique, we have no control over λ, so when we prepare the spin state, λ might take any value — the probability distribution governing its value is uniform on the unit interval.

Now suppose that when we measure the spin along an axis $\hat{n}$ rotated by θ from the $\hat{z}$ axis, the outcome will be

$$\begin{aligned} &|\uparrow_{\hat{n}}\rangle\ , \quad \text{for } 0 \leq \lambda \leq \cos^2\frac{\theta}{2}\ , \\ &|\downarrow_{\hat{n}}\rangle\ , \quad \text{for } \cos^2\frac{\theta}{2} < \lambda \leq 1\ . \end{aligned} \tag{4.14}$$

If we know λ, the outcome is deterministic, but if λ is completely unknown, then the probability distribution governing the measurement will agree with the predictions of quantum theory. In a hidden-variable theory, the randomness of the measurement outcome is not intrinsic; rather, it results from ignorance — our description of the system is not the most complete possible description.

Now, what about entangled states? When we say that a hidden-variable theory is *local*, we mean that it satisfies the Einstein locality constraint. A measurement of A does not modify the values of the variables that govern the measurements of B. Rather, when Alice measures her half of an entangled state that she shares with Bob, she gains information about the values of the hidden variables, sharpening her ability to predict what Bob will find when he measures the other half. This seems to be what Einstein had in mind when he envisioned a more complete description.

4.2 The Bell inequality

4.2.1 Three quantum coins

Is a local hidden-variable theory merely a reformulation of quantum mechanics, or is it a testable hypothesis? John Bell's fruitful idea was to test Einstein locality by considering the quantitative properties of the correlations between measurement outcomes obtained by two parties, Alice and Bob, who share an entangled state. Let us consider an example of the sort of correlations that Alice and Bob would like to explain.

The system that Alice and Bob are studying might be described this way: Alice, in Pasadena, has in her possession three coins laid out on a table, labeled $1, 2, 3$. Each coin has either its heads (H) or tails (T) side facing up, but it is hidden under an opaque cover, so that Alice is not able to tell whether it is an H or a T. Alice can uncover any one of the three coins, and so learn its value (H or T). However, as soon as that one coin is uncovered, the other two covered coins instantly disappear in a puff of smoke, and Alice never gets an opportunity to uncover the other coins. She has many copies of the three-coin set, and eventually she learns that, no matter which coin she exposes, she is just as likely to find an H as a T. Bob, in Chicago, has a similar set of coins, also labeled $1, 2, 3$. He too finds that each one of his coins, when revealed, is as likely to be an H as a T.

In fact, Alice and Bob have many identical copies of their shared set of coins, so they conduct an extensive series of experiments to investigate how their coin sets are correlated with one another. They quickly make a remarkable discovery: Whenever Alice and Bob uncover coins with the *same* label (whether 1, 2, or 3), they *always* find coins with the same value — either both are H or both are T. They conduct a million trials, just to be sure, and it works every single time! Their coin sets are perfectly correlated.

Alice and Bob suspect that they have discovered something important, and they frequently talk on the phone to brainstorm about the implications of their results. One day, Alice is in an especially reflective mood:

Alice: You know, Bob, sometimes it's hard for me to decide which of the three coins to uncover. I know that if I uncover coin 1, say, then coins 2 and 3 will disappear, and I'll never have a chance to find out the values of those coins. Once, just once, I'd love to be able to uncover two of the three coins, and find out whether each is an H or a T. But I've tried and it just isn't possible — there's no way to look at one coin and prevent the other from going poof!

Bob: [*Long pause*] Hey ... wait a minute Alice, I've got an idea ... Look, I think there *is* a way for you to find the value of two of your coins, after all! Let's say you would like to uncover coin 1 and coin 2. Well, I'll uncover *my* coin 2 here in Chicago, and I'll call you to tell you what I found, let's say its an H. We know, then, that you are certain to find an H if you uncover your coin 2 also. There's absolutely no doubt about that, because we've checked it a million times. Right?

Alice: Right . . .

Bob: But now there's no reason for you to uncover your coin 2; you know what you'll find anyway. You can uncover coin 1 instead. And then you'll know the value of both coins.

Alice: Hmmm . . . yeah, maybe. But we won't be sure, will we? I mean, yes, it always worked when we uncovered the same coin before, but this time you uncovered your coin 2, and your coins 1 and 3 disappeared, and I uncovered my coin 1, and my coins 2 and 3 disappeared. There's no way we'll ever be able to check anymore what would have happened if we had both uncovered coin 2.

Bob: We don't have to check that anymore, Alice; we've already checked it a million times. Look, your coins are in Pasadena and mine are in Chicago. Clearly, there's just no way that my decision to uncover my coin 2 can have any *influence* on what you'll find when you uncover your coin 2. That's not what's happening. It's just that when I uncover my coin 2 we're collecting the information we need to predict with certainty what will happen when you uncover your coin 2. Since we're already certain about it, why bother to do it!

Alice: Okay, Bob, I see what you mean. Why don't we do an experiment to see what really happens when you and I uncover different coins?

Bob: I don't know, Alice. We're not likely to get any funding to do such a dopey experiment. I mean, does anybody really care what happens when I uncover coin 2 and you uncover coin 1?

Alice: I'm not sure, Bob. But I've heard about a theorist named Bell. They say that he has some interesting ideas about the coins. He might have a theory that makes a prediction about what we'll find. Maybe we should talk to him.

Bob: Good idea! And it doesn't really matter whether his theory makes any sense or not. We can still propose an experiment to test his prediction, and they'll probably fund us.

So Alice and Bob travel to CERN to have a chat with Bell. They tell Bell about the experiment they propose to do. Bell listens closely, but for a long time he remains silent, with a faraway look in his eyes. Alice and Bob are not bothered by his silence, as they rarely understand anything that theorists say anyway. But finally Bell speaks.

Bell: I think I have an idea . . . . When Bob uncovers his coin in Chicago, that can't exert any *influence* on Alice's coin in Pasadena. Instead, what Bob finds out by uncovering his coin reveals some *information* about what will happen when Alice uncovers her coin.

Bob: Well, that's what I've been saying . . .

Bell: Right. Sounds reasonable. So let's assume that Bob is right about that. Now Bob can uncover any one of his coins, and know for sure what Alice will find when she uncovers the corresponding coin. He isn't *disturbing* her coin in any way; he's just finding out about it. We're forced to conclude that there must be some *hidden variables* that specify the condition of Alice's coins. And if those variables are completely known, then the value of each of Alice's coins can be unambiguously predicted.

Bob: [*Impatient with all this abstract stuff*] Yeah, but so what?

Bell: When your correlated coin sets are prepared, the values of the hidden variables are not completely specified; that's why any one coin is as likely to be an H as a T. But there must be some probability distribution $P(x, y, z)$ (with $x, y, z \in \{H, T\}$) that characterizes the preparation and governs Alice's three coins. These probabilities must be nonnegative, and they sum to one:

$$\sum_{x,y,z\in\{H,T\}} P(x, y, z) = 1 \; . \tag{4.15}$$

Alice can't uncover all three of her coins, so she can't measure $P(x, y, z)$ directly. But with Bob's help, she can in effect uncover any two coins of her choice. Let's denote with $P_{\text{same}}(i, j)$, the probability that coins i and j $(i, j = 1, 2, 3)$ have the same value, either both H or both T. Then we see that

$$\begin{aligned} P_{\text{same}}(1,2) = &\; P(HHH) + P(HHT) + P(TTH) + P(TTT) \; , \\ P_{\text{same}}(2,3) = &\; P(HHH) + P(THH) + P(HTT) + P(TTT) \; , \\ P_{\text{same}}(1,3) = &\; P(HHH) + P(HTH) + P(THT) + P(TTT) \; , \end{aligned} \tag{4.16}$$

and it immediately follows from eq. (4.15) that

$$\begin{aligned} &P_{\text{same}}(1,2) + P_{\text{same}}(2,3) + P_{\text{same}}(1,3) \\ &= 1 + 2\; P(HHH) + 2\; P(TTT) \geq 1 \; . \end{aligned} \tag{4.17}$$

So that's my prediction: P_{same} should obey the inequality

$$P_{\mathrm{same}}(1,2) + P_{\mathrm{same}}(2,3) + P_{\mathrm{same}}(1,3) \geq 1 \ . \tag{4.18}$$

You can test it my doing your experiment that "uncovers" two coins at a time.

Bob: Well, I guess the math looks right. But I don't really get it. Why does it work?

Alice: I think I see .... Bell is saying that if there are three coins on a table, and each one is either an H or a T, then at least two of the three have to be the *same*, either both H or both T. Isn't that it, Bell?

Bell stares at Alice, a surprised look on his face. His eyes glaze, and for a long time he is speechless. Finally, he speaks:

Bell: Yes

So Alice and Bob are amazed and delighted to find that Bell is that rarest of beasts — a theorist who makes sense. With Bell's help, their proposal is approved and they do the experiment, only to obtain a shocking result. After many careful trials, they conclude, to very good statistical accuracy that

$$P_{\mathrm{same}}(1,2) \simeq P_{\mathrm{same}}(2,3) \simeq P_{\mathrm{same}}(1,3) \simeq \frac{1}{4} \ , \tag{4.19}$$

and hence

$$P_{\mathrm{same}}(1,2) + P_{\mathrm{same}}(2,3) + P_{\mathrm{same}}(1,3) \simeq 3 \cdot \frac{1}{4} = \frac{3}{4} < 1 \ . \tag{4.20}$$

The correlations found by Alice and Bob flagrantly violate Bell's inequality!

Alice and Bob are good experimenters, but dare not publish so disturbing a result unless they can find a plausible theoretical interpretation. Finally, they become so desperate that they visit the library to see if quantum mechanics can offer any solace ...

4.2.2 Quantum entanglement vs. Einstein locality

What Alice and Bob read about is quantum entanglement. Eventually, they learn that their magical coins are governed by a maximally entangled state of two qubits. What Alice and Bob really share are many copies of the state $|\psi^-\rangle$. When Alice uncovers a coin, she is measuring her qubit along one of three possible axes, no two of which are orthogonal. Since the measurements don't commute, Alice can uncover only one of her three coins. Similarly, when Bob uncovers his coin, he measures his member of the entangled pair along any one of three axes, so he too is limited to uncovering just one of his three coins. But since Alice's measurements commute with Bob's, they can uncover one coin each, and study how Alice's coins are correlated with Bob's coins.

To help Alice and Bob interpret their experiment, let's see what quantum mechanics predicts about these correlations. The state $|\psi^-\rangle$ has the convenient property that it remains invariant if Alice and Bob each apply the same unitary transformation,

$$\boldsymbol{U} \otimes \boldsymbol{U} |\psi\rangle = |\psi\rangle \ . \tag{4.21}$$

For infinitesimal unitaries, this becomes the property

$$\left(\vec{\boldsymbol{\sigma}}^{(A)} + \vec{\boldsymbol{\sigma}}^{(B)}\right) |\psi^-\rangle = 0 \tag{4.22}$$

(the state has vanishing total angular momentum, as you can easily check by an explicit computation). Now consider the expectation value

$$\langle\psi^-| \left(\vec{\boldsymbol{\sigma}}^{(A)} \cdot \hat{a}\right) \left(\vec{\boldsymbol{\sigma}}^{(B)} \cdot \hat{b}\right) |\psi^-\rangle \ , \tag{4.23}$$

where $\hat{a}$ and $\hat{b}$ are unit 3-vectors. Acting on $|\psi^-\rangle$, we can replace $\vec{\boldsymbol{\sigma}}^{(B)}$ by $-\vec{\boldsymbol{\sigma}}^{(A)}$; therefore, the expectation value can be expressed as a property of Alice's system, which has density operator $\boldsymbol{\rho}_A = \frac{1}{2}\boldsymbol{I}$:

$$\begin{aligned} &- \langle\psi^-| \left(\vec{\boldsymbol{\sigma}}^{(A)} \cdot \hat{a}\right) \left(\vec{\boldsymbol{\sigma}}^{(A)} \cdot \hat{b}\right) |\psi^-\rangle \\ &= -a_i b_j \mathrm{tr}\left(\boldsymbol{\rho}_A \boldsymbol{\sigma}_i^{(A)} \boldsymbol{\sigma}_j^{(A)}\right) = -a_i b_j \delta_{ij} = -\hat{a} \cdot \hat{b} = -\cos\theta \ , \end{aligned} \tag{4.24}$$

where θ is the angle between the axes $\hat{a}$ and $\hat{b}$. Thus we find that the measurement outcomes are always perfectly anticorrelated when we measure both spins along the same axis $\hat{a}$, and we have also obtained a more general result that applies when the two axes are different.

The projection operator onto the spin up (spin down) states along $\hat{n}$ is

$\boldsymbol{E}(\hat{n},\pm) = \frac{1}{2}(\boldsymbol{I} \pm \hat{n}\cdot\vec{\boldsymbol{\sigma}})$; we therefore obtain

$$\begin{aligned}
P(++) &= \langle\psi^-|\boldsymbol{E}^{(A)}(\hat{a},+)\boldsymbol{E}^{(B)}(\hat{b},+)|\psi^-\rangle = \frac{1}{4}(1-\cos\theta)\ ,\\
P(--) &= \langle\psi^-|\boldsymbol{E}^{(A)}(\hat{a},-)\boldsymbol{E}^{(B)}(\hat{b},-)|\psi^-\rangle = \frac{1}{4}(1-\cos\theta)\ ,\\
P(+-) &= \langle\psi^-|\boldsymbol{E}^{(A)}(\hat{a},+)\boldsymbol{E}^{(B)}(\hat{b},-)|\psi^-\rangle = \frac{1}{4}(1+\cos\theta)\ ,\\
P(-+) &= \langle\psi^-|\boldsymbol{E}^{(A)}(\hat{a},-)\boldsymbol{E}^{(B)}(\hat{b},+)|\psi-\rangle = \frac{1}{4}(1+\cos\theta)\ ;
\end{aligned} \tag{4.25}$$

here $P(++)$ is the probability that Alice and Bob both obtain the spin-up outcome when Alice measures along $\hat{a}$ and Bob measures along $\hat{b}$, etc. The probability that their outcomes are the same is

$$P_{\text{same}} = P(++) + P(--) = \frac{1}{2}(1-\cos\theta)\ , \tag{4.26}$$

and the probability that their outcomes are opposite is

$$P_{\text{opposite}} = P(+-) + P(-+) = \frac{1}{2}(1+\cos\theta)\ . \tag{4.27}$$

Now suppose that Alice measures her spin along one of the three symmetrically distributed axes in the $x-z$ plane,

$$\begin{aligned}
\hat{a}_1 &= (0,0,1)\ ,\\
\hat{a}_2 &= \left(\frac{\sqrt{3}}{2},0,-\frac{1}{2}\right)\ ,\\
\hat{a}_3 &= \left(-\frac{\sqrt{3}}{2},0,-\frac{1}{2}\right)\ ,
\end{aligned} \tag{4.28}$$

so that

$$\hat{a}_1\cdot\hat{a}_2 = \hat{a}_2\cdot\hat{a}_3 = \hat{a}_3\cdot\hat{a}_1 = -\frac{1}{2}\ . \tag{4.29}$$

And suppose that Bob measures along one of three axes that are diametrically opposed to Alice's:

$$\hat{b}_1 = -\hat{a}_1\ ,\quad \hat{b}_2 = -\hat{a}_2\ ,\quad \hat{b}_3 = -\hat{a}_3\ . \tag{4.30}$$

When Alice and Bob choose opposite axes, then $\theta = 180^\circ$ and $P_{\text{same}} = 1$. But otherwise $\theta = \pm 60^\circ$ so that $\cos\theta = 1/2$ and $P_{\text{same}} = 1/4$. This is just the behavior that Alice and Bob found in their experiment, in violation of Bell's prediction.

Bell's logic seemed compelling but something went wrong, so we are forced to reconsider his tacit assumptions. First, Bell assumed that there is a joint probability distribution that governs the possible outcomes of all measurements that Alice and Bob might perform. This is the hidden-variable hypothesis. He imagines that if the values of the hidden variables are exactly known, then the outcome of any measurement can be predicted with certainty — measurement outcomes are described probabilistically because the values of the hidden variables are drawn from an ensemble of possible values. Second, Bell assumed that Bob's decision about what to measure in Chicago has no effect on the hidden variables that govern Alice's measurement in Pasadena. This is the assumption that the hidden variables are local. If we accept these two assumptions, there is no escaping Bell's conclusion. We have found that the correlations predicted by quantum theory are incompatible with theses assumptions.

What are the implications? Perhaps the moral of the story is that it can be dangerous to reason about what might have happened, but didn't actually happen — what are sometimes called *counterfactuals*. Of course, we do this all the time in our everyday lives, and we usually get away with it; reasoning about counterfactuals seems to be acceptable in the classical world, but sometimes it gets us into trouble in the quantum world. We claimed that Alice knew what would happen when she measured along $\hat{a}_1$, because Bob measured along $-\hat{a}_1$, and every time we have ever checked, their measurement outcomes are always perfectly correlated. But Alice did *not* measure along $\hat{a}_1$; she measured along $\hat{a}_2$ instead. We got into trouble by trying to assign probabilities to the outcomes of measurements along $\hat{a}_1, \hat{a}_2$, and $\hat{a}_3$, even though Alice can perform just one of those measurements. In quantum theory, assuming that there is a probability distribution that governs the outcomes of all three measurements that Alice might have made, even though she was able to carry out only one of these measurements, leads to mathematical inconsistencies, so we had better not do it. We have affirmed Bohr's principle of *complementary* — we are forbidden to consider simultaneously the possible outcomes of two mutually exclusive experiments.

One who rejects the complementarity principle may prefer to say that violations of the Bell inequalities (confirmed experimentally) have exposed an essential nonlocality built into the quantum description of Nature. *If* we do insist that it is legitimate to talk about outcomes of mutually exclusive experiments *then* we are forced to conclude that Bob's choice of measurement actually exerted a subtle *influence* on the outcome of Alice's measurement. Thus advocates of this viewpoint speak of "quantum nonlocality."

By ruling out local hidden variables, Bell demolished Einstein's dream that the indeterminacy of quantum theory could be eradicated by adopting a more complete, yet still local, description of Nature. If we accept locality as an inviolable principle, then we are forced to accept randomness as an unavoidable and intrinsic feature of quantum measurement, rather than a consequence of incomplete knowledge.

To some, the peculiar correlations unmasked by Bell's inequality call out for a deeper explanation than quantum mechanics seems to provide. They see the EPR phenomenon as a harbinger of new physics awaiting discovery. But they may be wrong. We have been waiting over 65 years since EPR, and so far no new physics.

The human mind seems to be poorly equipped to grasp the correlations exhibited by entangled quantum states, and so we speak of the weirdness of quantum theory. But whatever your attitude, experiment forces you to accept the existence of the weird correlations among the measurement outcomes. There is no big mystery about how the correlations were established — we saw that it was necessary for Alice and Bob to get together at some point to create entanglement among their qubits. The novelty is that, even when A and B are distantly separated, we cannot accurately regard A and B as two separate qubits, and use classical information to characterize how they are correlated. They are more than just correlated, they are a single *inseparable* entity. They are *entangled*.

4.3 More Bell inequalities

4.3.1 CHSH inequality

Experimental tests of Einstein locality typically are based on another form of the Bell inequality, which applies to a situation in which Alice can measure either one of two observables $\boldsymbol{a}$ and $\boldsymbol{a}'$, while Bob can measure either $\boldsymbol{b}$ or $\boldsymbol{b}'$. Suppose that the observables $\boldsymbol{a}$, $\boldsymbol{a}'$, $\boldsymbol{b}$, $\boldsymbol{b}'$ take values in $\{\pm 1\}$, and are functions of hidden random variables.

If $\boldsymbol{a}, \boldsymbol{a}' = \pm 1$, it follows that either $\boldsymbol{a}+\boldsymbol{a}' = 0$, in which case $\boldsymbol{a}-\boldsymbol{a}' = \pm 2$, or else $\boldsymbol{a} - \boldsymbol{a}' = 0$, in which case $\boldsymbol{a} + \boldsymbol{a}' = \pm 2$; therefore

$$\boldsymbol{C} \equiv (\boldsymbol{a} + \boldsymbol{a}')\boldsymbol{b} + (\boldsymbol{a} - \boldsymbol{a}')\boldsymbol{b}' = \pm 2 \ . \tag{4.31}$$

(Here is where the local hidden-variable assumption sneaks in — we have imagined that values in $\{\pm 1\}$ can be assigned simultaneously to all four observables, even though it is impossible to measure both of $\boldsymbol{a}$ and $\boldsymbol{a}'$, or both of $\boldsymbol{b}$ and $\boldsymbol{b}'$.) Evidently

$$|\langle \boldsymbol{C} \rangle| \leq \langle |\boldsymbol{C}| \rangle = 2, \tag{4.32}$$

so that

$$|\langle \boldsymbol{a}\boldsymbol{b} \rangle + \langle \boldsymbol{a}'\boldsymbol{b} \rangle + \langle \boldsymbol{a}\boldsymbol{b}' \rangle - \langle \boldsymbol{a}'\boldsymbol{b}' \rangle| \leq 2. \tag{4.33}$$

This result is called the *CHSH inequality* (for Clauser-Horne-Shimony-Holt). It holds for any random variables $\boldsymbol{a}, \boldsymbol{a}', \boldsymbol{b}, \boldsymbol{b}'$ taking values in ±1 that are governed by a joint probability distribution.

To see that quantum mechanics violates the CHSH inequality, let $\boldsymbol{a}, \boldsymbol{a}'$ denote the Hermitian operators

$$\boldsymbol{a} = \vec{\boldsymbol{\sigma}}^{(A)} \cdot \hat{a} \ , \quad \boldsymbol{a}' = \vec{\boldsymbol{\sigma}}^{(A)} \cdot \hat{a}' \ , \tag{4.34}$$

acting on a qubit in Alice's possession, where $\hat{a}, \hat{a}'$ are unit 3-vectors. Similarly, let $\boldsymbol{b}, \boldsymbol{b}'$ denote

$$\boldsymbol{b} = \vec{\boldsymbol{\sigma}}^{(B)} \cdot \hat{b} \ , \quad \boldsymbol{b}' = \vec{\boldsymbol{\sigma}}^{(B)} \cdot \hat{b}' \ , \tag{4.35}$$

acting on Bob's qubit. Each observable has eigenvalues ±1 so that an outcome of a measurement of the observable takes values in ±1.

Recall that if Alice and Bob share the maximally-entangled state $|\psi^-\rangle$, then

$$\langle \psi^- | \left(\vec{\boldsymbol{\sigma}}^{(A)} \cdot \hat{a} \right) \left(\vec{\boldsymbol{\sigma}}^{(B)} \cdot \hat{b} \right) | \psi^- \rangle = -\hat{a} \cdot \hat{b} = -\cos\theta \ , \tag{4.36}$$

where θ is the angle between $\hat{a}$ and $\hat{b}$. Consider the case where $\hat{a}', \hat{b}, \hat{a}, \hat{b}'$ are coplanar and separated by successive 45° angles. so that the quantum-mechanical predictions are

$$\begin{aligned} \langle \boldsymbol{a}\boldsymbol{b} \rangle = \langle \boldsymbol{a}'\boldsymbol{b} \rangle = \langle \boldsymbol{a}\boldsymbol{b}' \rangle &= -\cos\frac{\pi}{4} = -\frac{1}{\sqrt{2}}, \\ \langle \boldsymbol{a}'\boldsymbol{b}' \rangle &= -\cos\frac{3\pi}{4} = \frac{1}{\sqrt{2}} \ . \end{aligned} \tag{4.37}$$

The CHSH inequality then becomes

$$4 \cdot \frac{1}{\sqrt{2}} = 2\sqrt{2} \le 2 \ , \tag{4.38}$$

which is clearly violated by the quantum-mechanical prediction.

4.3.2 Maximal violation

In fact the case just considered provides the largest possible quantum-mechanical violation of the CHSH inequality, as we can see by the following argument. Suppose that $\boldsymbol{a}, \boldsymbol{a}', \boldsymbol{b}, \boldsymbol{b}'$ are Hermitian operators with eigenvalues ±1, so that

$$\boldsymbol{a}^2 = \boldsymbol{a}'^2 = \boldsymbol{b}^2 = \boldsymbol{b}'^2 = \boldsymbol{I} \ , \tag{4.39}$$

and suppose that "Alice's observables" $\boldsymbol{a}, \boldsymbol{a}'$ commute with "Bob's observables" $\boldsymbol{b}, \boldsymbol{b}'$:

$$0 = [\boldsymbol{a}, \boldsymbol{b}] = [\boldsymbol{a}, \boldsymbol{b}'] = [\boldsymbol{a}', \boldsymbol{b}] = [\boldsymbol{a}', \boldsymbol{b}'] \ . \tag{4.40}$$

Defining

$$\boldsymbol{C} = \boldsymbol{ab} + \boldsymbol{a}'\boldsymbol{b} + \boldsymbol{ab}' - \boldsymbol{a}'\boldsymbol{b}' \ , \tag{4.41}$$

we evaluate

$$\boldsymbol{C}^2 = \begin{array}{cccc} \boldsymbol{I} & +\boldsymbol{aa}' & +\boldsymbol{bb}' & -\boldsymbol{aa}'\boldsymbol{bb}' \\ +\boldsymbol{a}'\boldsymbol{a} & +\boldsymbol{I} & +\boldsymbol{a}'\boldsymbol{abb}' & -\boldsymbol{bb}' \\ +\boldsymbol{b}'\boldsymbol{b} & +\boldsymbol{aa}'\boldsymbol{b}'\boldsymbol{b} & +\boldsymbol{I} & -\boldsymbol{aa}' \\ -\boldsymbol{a}'\boldsymbol{ab}'\boldsymbol{b} & -\boldsymbol{b}'\boldsymbol{b} & -\boldsymbol{a}'\boldsymbol{a} & +\boldsymbol{I} \end{array} \ , \tag{4.42}$$

using eq. (4.39). All the quadratic terms cancel pairwise, so that we are left with

$$\begin{aligned} \boldsymbol{C}^2 &= 4\boldsymbol{I} - \boldsymbol{aa}'\boldsymbol{bb}' + \boldsymbol{a}'\boldsymbol{abb}' + \boldsymbol{aa}'\boldsymbol{b}'\boldsymbol{b} - \boldsymbol{a}'\boldsymbol{ab}'\boldsymbol{b} \\ &= 4\boldsymbol{I} - [\boldsymbol{a}, \boldsymbol{a}'][\boldsymbol{b}, \boldsymbol{b}'] \ . \end{aligned} \tag{4.43}$$

Now recall that the *sup norm* $\| \boldsymbol{M} \|_{\text{sup}}$ of a bounded operator $\boldsymbol{M}$ is defined by

$$\| \boldsymbol{M} \|_{\text{sup}} = \begin{array}{c}\text{sup} \\ |\psi\rangle\end{array} \left(\frac{\| \boldsymbol{M}|\psi\rangle \|}{\| |\psi\rangle \|} \right) \ ; \tag{4.44}$$

that is, the sup norm of $\boldsymbol{M}$ is the maximum eigenvalue of $\sqrt{\boldsymbol{M}^\dagger \boldsymbol{M}}$. It is easy to verify that the sup norm has the properties

$$\begin{aligned} \| \boldsymbol{MN} \|_{\text{sup}} &\leq \| \boldsymbol{M} \|_{\text{sup}} \cdot \| \boldsymbol{N} \|_{\text{sup}} \ , \\ \| \boldsymbol{M} + \boldsymbol{N} \|_{\text{sup}} &\leq \| \boldsymbol{M} \|_{\text{sup}} + \| \boldsymbol{N} \|_{\text{sup}} \ . \end{aligned} \tag{4.45}$$

A Hermitian operator with eigenvalues ±1 has unit sup norm, so that

$$\| \boldsymbol{C}^2 \|_{\text{sup}} \leq 4 + 4 \| \boldsymbol{a} \|_{\text{sup}} \cdot \| \boldsymbol{a}' \|_{\text{sup}} \cdot \| \boldsymbol{b} \|_{\text{sup}} \cdot \| \boldsymbol{b}' \|_{\text{sup}} = 8 \ . \tag{4.46}$$

Because $\boldsymbol{C}$ is Hermitian,

$$\| \boldsymbol{C}^2 \|_{\text{sup}} = \| \boldsymbol{C} \|_{\text{sup}}^2 \ , \tag{4.47}$$

and therefore

$$\| \boldsymbol{C} \|_{\text{sup}} \leq 2\sqrt{2} \ , \tag{4.48}$$

which is known as Cirel'son's inequality.

The CHSH inequality is the statement $|\langle \boldsymbol{C} \rangle| \leq 2$. Quantum mechanically, the absolute value of the expectation value of the Hermitian operator $\boldsymbol{C}$ can be no larger than its largest eigenvalue,

$$|\langle \boldsymbol{C} \rangle| \leq \| \boldsymbol{C} \|_{\text{sup}} \leq 2\sqrt{2} \ . \tag{4.49}$$

We saw that this upper bound is saturated in the case where $\boldsymbol{a}', \boldsymbol{b}, \boldsymbol{a}, \boldsymbol{b}'$ are separated by successive 45^o angles. Thus the violation of the CHSH inequality that we found is the largest violation allowed by quantum theory.

4.3.3 Quantum strategies outperform classical strategies

The CHSH inequality is a limitation on the strength of the correlations between the two parts of a bipartite classical system, and the Cirel'son inequality is a limitation on the strength of the correlations between the two parts of a bipartite quantum system. We can deepen our appreciation of how quantum correlations differ from classical correlations by considering a game for which quantum strategies outperform classical strategies.

Alice and Bob are playing a game with Charlie. Charlie prepares two bits $x, y \in \{0, 1\}$; then he sends x to Alice and y to Bob. After receiving the input bit x, Alice is to produce an output bit $a \in \{0, 1\}$, and after receiving y, Bob is to produce output bit $b \in \{0, 1\}$. But Alice and Bob are not permitted to communicate, so that Alice does not know y and Bob does not know x.

Alice and Bob win the game if their output bits are related to the input bits according to

$$a \oplus b = x \wedge y \ , \tag{4.50}$$

where $\oplus$ denotes the sum modulo 2 (the XOR gate) and $\wedge$ denotes the product (the AND gate). Can Alice and Bob find a strategy that enables them to win the game every time, no matter how Charlie chooses the input bits?

No, it is easy to see that there is no such strategy. Let a_0, a_1 denote the value of Alice's output if her input is $x = 0, 1$ and let b_0, b_1 denote Bob's output if his input is $y = 0, 1$. For Alice and Bob to win for all possible inputs, their output bits must satisfy

$$\begin{aligned} a_0 \oplus b_0 &= 0 \ , \\ a_0 \oplus b_1 &= 0 \ , \\ a_1 \oplus b_0 &= 0 \ , \\ a_1 \oplus b_1 &= 1 \ . \end{aligned} \tag{4.51}$$

But this is impossible, since by summing the four equations we obtain 0=1.

Suppose that Charlie generates the input bits at random. Then there is a very simple strategy that enables Alice and Bob to win the game three times our of four: they always choose the output $a = b = 0$ so that they lose only if the input is $x = y = 1$. The CHSH inequality can be regarded as the statement that, if Alice and Bob share no quantum entanglement, then there is no better strategy.

To connect this statement with our previous formulation of the CHSH inequality, define random variables taking values ± 1 as

$$\begin{aligned} \boldsymbol{a} &= (-1)^{a_0} , & \boldsymbol{a}' &= (-1)^{a_1} , \\ \boldsymbol{b} &= (-1)^{b_0} , & \boldsymbol{b}' &= (-1)^{b_1} . \end{aligned} \tag{4.52}$$

Then the CHSH inequality says that for any joint probability distribution governing $\boldsymbol{a}, \boldsymbol{a}', \boldsymbol{b}, \boldsymbol{b}' \in \{\pm 1\}$, the expectation values satisfy

$$\langle \boldsymbol{a}\boldsymbol{b}\rangle + \langle \boldsymbol{a}\boldsymbol{b}'\rangle + \langle \boldsymbol{a}'\boldsymbol{b}\rangle - \langle \boldsymbol{a}'\boldsymbol{b}'\rangle \le 2 . \tag{4.53}$$

Furthermore, if we denote by p_{xy} the probability that eq. (4.51) is satisfied when the input bits are (x, y), then

$$\begin{aligned} \langle \boldsymbol{a}\boldsymbol{b}\rangle &= 2p_{00} - 1 , \\ \langle \boldsymbol{a}\boldsymbol{b}'\rangle &= 2p_{01} - 1 , \\ \langle \boldsymbol{a}'\boldsymbol{b}\rangle &= 2p_{10} - 1 , \\ \langle \boldsymbol{a}'\boldsymbol{b}'\rangle &= 1 - 2p_{11} ; \end{aligned} \tag{4.54}$$

for example $\langle \boldsymbol{a}\boldsymbol{b}\rangle = p_{00} - (1 - p_{00}) = 2p_{00} - 1$, because the value of $\boldsymbol{a}\boldsymbol{b}$ is $+1$ when Alice and Bob win and -1 when they lose. The CHSH inequality eq. (4.53) becomes

$$2\left(p_{00} + p_{01} + p_{10} + p_{11}\right) - 4 \le 2 , \tag{4.55}$$

or

$$\langle p\rangle \equiv \frac{1}{4}\left(p_{00} + p_{01} + p_{10} + p_{11}\right) \le \frac{3}{4} , \tag{4.56}$$

where $\langle p\rangle$ denotes the probability of winning averaged over a uniform ensemble for the input bits. Thus, if the input bits are random, Alice and Bob cannot attain a probability of winning higher than 3/4.

It is worthwhile to consider how the assumption that Alice and Bob take actions governed by "*local* hidden variables" limits their success in playing the game. Although Alice and Bob do not share any quantum entanglement, they are permitted to share a table of random numbers that they may consult to produce their output bits. Thus we may imagine that hidden variables drawn from an ensemble of possible values guide Alice and Bob to make correlated decisions. These correlations are limited by locality — Alice does not know Bob's input and Bob does not know Alice's. In fact, we have learned that for playing this game their shared randomness is of no value — their best strategy does not use the shared randomness at all.

But if Alice and Bob share quantum entanglement, they can devise a better strategy. Based on the value of her input bit, Alice decides to measure one of two Hermitian observables with eigenvalues ±1: $\boldsymbol{a}$ if $x = 0$ and $\boldsymbol{a}'$ is $x = 1$. Similarly, Bob measures $\boldsymbol{b}$ if $y = 0$ and $\boldsymbol{b}'$ if $y = 1$. Then the quantum-mechanical expectation values of these observables satisfy the Cirel'son inequality

$$\langle \boldsymbol{a}\boldsymbol{b}\rangle + \langle \boldsymbol{a}\boldsymbol{b}'\rangle + \langle \boldsymbol{a}'\boldsymbol{b}\rangle - \langle \boldsymbol{a}'\boldsymbol{b}'\rangle \le 2\sqrt{2}\ , \tag{4.57}$$

and the probability that Alice and Bob win the game is constrained by

$$2\left(p_{00} + p_{01} + p_{10} + p_{11}\right) - 4 \le 2\sqrt{2}\ , \tag{4.58}$$

or

$$\langle p\rangle \equiv \frac{1}{4}\left(p_{00} + p_{01} + p_{10} + p_{11}\right) \le \frac{1}{2} + \frac{1}{2\sqrt{2}} \approx .853\ . \tag{4.59}$$

Furthermore, we have seen that this inequality can be saturated if Alice and Bob share a maximally entangled state of two qubits, and the observables $\boldsymbol{a}, \boldsymbol{a}', \boldsymbol{b}, \boldsymbol{b}'$ are chosen appropriately.

Thus we have found that Alice and Bob can play the game more successfully with quantum entanglement than without it. At least for this purpose, shared quantum entanglement is a more powerful resource than shared classical randomness. But even the power brought by entanglement has limits, limits embodied by the Cirel'son inequality.

4.3.4 All entangled pure states violate Bell inequalities

Separable states do not violate Bell inequalities. For example, in the case of a separable *pure* state, if $\boldsymbol{a}$ is an observable acting on Alice's qubit, and $\boldsymbol{b}$ is an observable acting on Bob's, then

$$\langle \boldsymbol{a}\boldsymbol{b}\rangle = \langle \boldsymbol{a}\rangle\langle \boldsymbol{b}\rangle. \tag{4.60}$$

No Bell-inequality violation can occur, because we have already seen that a (local) hidden-variable theory *does* exist that correctly reproduces the predictions of quantum theory for a pure state of a single qubit. A general separable state is just a probabilistic mixture of separable pure states, so that the correlations between the subsystems are entirely classical, and the Bell inequalities apply.

On the other hand, we have seen that a maximally entangled state such as $|\psi^-\rangle$ *is* Bell-inequality violating. But what about pure states that are only partially entangled such as

$$|\phi\rangle = \alpha|00\rangle + \beta|11\rangle \ ? \tag{4.61}$$

Any pure state of two qubits can be expressed this way in the Schmidt basis; with suitable phase conventions, α and β are real and nonnegative.

Suppose that Alice and Bob both measure along an axis in the x-z plane, so that their observables are

$$\begin{aligned} \boldsymbol{a} &= \boldsymbol{\sigma}_3^{(A)} \cos\theta_A + \boldsymbol{\sigma}_1^{(A)} \sin\theta_A \ , \\ \boldsymbol{b} &= \boldsymbol{\sigma}_3^{(B)} \cos\theta_B + \boldsymbol{\sigma}_1^{(B)} \sin\theta_B \ . \end{aligned} \tag{4.62}$$

The state $|\phi\rangle$ has the properties

$$\begin{aligned} &\langle\phi|\boldsymbol{\sigma}_3 \otimes \boldsymbol{\sigma}_3|\phi\rangle = 1 \ , \quad \langle\phi|\boldsymbol{\sigma}_1 \otimes \boldsymbol{\sigma}_1|\phi\rangle = 2\alpha\beta \ , \\ &\langle\phi|\boldsymbol{\sigma}_3 \otimes \boldsymbol{\sigma}_1|\phi\rangle = 0 = \langle\phi|\boldsymbol{\sigma}_1 \otimes \boldsymbol{\sigma}_3|\phi\rangle \ , \end{aligned} \tag{4.63}$$

so that the quantum-mechanical expectation value of $\boldsymbol{ab}$ is

$$\langle\boldsymbol{ab}\rangle = \langle\phi|\boldsymbol{ab}|\phi\rangle = \cos\theta_A \cos\theta_B + 2\alpha\beta \sin\theta_A \sin\theta_B \tag{4.64}$$

(and we recover $\cos(\theta_A - \theta_B)$ in the maximally entangled case $\alpha = \beta = 1/\sqrt{2}$). Now let us consider, for simplicity, the (nonoptimal!) special case

$$\theta_A = 0, \quad \theta_A' = \frac{\pi}{2}, \quad \theta_B' = -\theta_B, \tag{4.65}$$

so that the quantum predictions are:

$$\begin{aligned} \langle\boldsymbol{ab}\rangle &= \cos\theta_B = \langle\boldsymbol{ab}'\rangle \ , \\ \langle\boldsymbol{a}'\boldsymbol{b}\rangle &= 2\alpha\beta \sin\theta_B = -\langle\boldsymbol{a}'\boldsymbol{b}'\rangle \ . \end{aligned} \tag{4.66}$$

Plugging into the CHSH inequality, we obtain

$$|\cos\theta_B - 2\alpha\beta \sin\theta_B| \le 1 \ , \tag{4.67}$$

and we easily see that violations occur for θ_B close to 0 or π. Expanding to linear order in θ_B, the left-hand side is

$$\simeq 1 - 2\alpha\beta\theta_B \ , \tag{4.68}$$

which surely exceeds 1 for $\alpha\beta > 0$ and θ_B negative and small.

We have shown that *any* entangled pure state of two qubits violates some Bell inequality. It is not hard to generalize the argument to an arbitrary bipartite pure state. For bipartite pure states, then, "entangled" is equivalent to "Bell-inequality violating." For bipartite mixed states, however, we will see later that the situation is more subtle.

4.3.5 Photons

Experiments that test the Bell inequality usually are done with entangled photons, not with spin-$\frac{1}{2}$ objects. What are the quantum-mechanical predictions for photons?

Recall from §2.2.2 that for a photon traveling in the $\hat{z}$ direction, we use the notation $|x\rangle$, $|y\rangle$ for the states that are linearly polarized along the x and y axes respectively. In terms of these basis states, the states that are linearly polarized along "horizontal" and "vertical" axes that are rotated by angle θ relative to the x and y axes can be expressed as

$$|H(\theta)\rangle = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} , \quad |V(\theta)\rangle = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix} . \tag{4.69}$$

We can construct a 2×2 matrix whose eigenstates are $|H(\theta)\rangle$ and $|V(\theta)\rangle$, with respective eigenvalues ±1; it is

$$\boldsymbol{\tau}(\theta) \equiv |H(\theta)\rangle\langle H(\theta)| - |V(\theta)\rangle\langle V(\theta)| = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix} . \tag{4.70}$$

The generator of rotations about the $\hat{z}$ axis is $\boldsymbol{J} = \boldsymbol{\sigma}_2$, and the eigenstates of $\boldsymbol{J}$ with eigenvalues ±1 are the circularly polarized states

$$|+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix} , \quad |-\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} i \\ 1 \end{pmatrix} . \tag{4.71}$$

Suppose that an excited atom emits two photons that come out back to back, with vanishing angular momentum and even parity. The two-photon states

$$\begin{array}{l} |+\rangle_A|-\rangle_B \\ |-\rangle_A|+\rangle_B \end{array} \tag{4.72}$$

are invariant under rotations about $\hat{z}$. The photons have opposite values of J_z, but the same *helicity* (angular-momentum along the axis of propagation), since they are propagating in opposite directions. Under a reflection in the $y-z$ plane, the polarization states are modified according to

$$|x\rangle \to -|x\rangle \ , \quad |y\rangle \to |y\rangle \ , \tag{4.73}$$

or

$$|+\rangle \to +i|-\rangle \ , \quad |-\rangle \to -i|+\rangle \ ; \tag{4.74}$$

therefore, the parity eigenstates are *entangled* states

$$\frac{1}{\sqrt{2}}(|+\rangle_A|-\rangle_B \pm |-\rangle_A|+\rangle_B) \ . \tag{4.75}$$

The state with $J_z = 0$ and even parity, then, expressed in terms of the linear polarization states, is

$$\begin{aligned} &-\frac{i}{\sqrt{2}}(|+-\rangle_{AB} + |-+\rangle_{AB}) \\ &= \frac{1}{\sqrt{2}}(|xx\rangle_{AB} + |yy\rangle_{AB}) \equiv |\phi^+\rangle_{AB} \ . \end{aligned} \tag{4.76}$$

Because of invariance under rotations about $\hat{z}$, the state has this form irrespective of how we orient the x and y axes.

Alice or Bob can use a polarization analyzer to project the polarization state of a photon onto the basis $\{|H(\theta)\rangle, |V(\theta)\rangle\}$, and hence measure $\boldsymbol{\tau}(\theta)$. For two photons in the state $|\phi^+\rangle$, if Alice orients her polarizer with angle θ_A and Bob with angle θ_B, then the correlations of their measurement outcomes are encoded in the expectation value

$$\langle\phi^+|\boldsymbol{\tau}^{(A)}(\theta_A)\boldsymbol{\tau}^{(B)}(\theta_B)|\phi^+\rangle. \tag{4.77}$$

Using rotational invariance:

$$\begin{aligned} &= \ \langle\phi^+|\boldsymbol{\tau}^{(A)}(0)\boldsymbol{\tau}^{(B)}(\theta_B-\theta_A)|\phi^+\rangle \\ &= \frac{1}{2}\langle x|\boldsymbol{\tau}^{(B)}(\theta_B-\theta_A)|x\rangle - \frac{1}{2}\langle y|\boldsymbol{\tau}^{(B)}(\theta_B-\theta_A)|y\rangle \\ &= \cos 2(\theta_B-\theta_A) \ . \end{aligned} \tag{4.78}$$

Recall that for the measurement of qubits on the Bloch sphere, we found the similar expression $\cos\theta$, where θ is the angle between Alice's polarization axis and Bob's. Here we have $\cos 2\theta$ instead, because photons have spin-1 rather than spin-$\frac{1}{2}$.

If Alice measures one of the two observables $\boldsymbol{a} = \boldsymbol{\tau}^{(A)}(\theta_A)$ or $\boldsymbol{a'} = \boldsymbol{\tau}^{(A)}(\theta'_A)$ and Bob measures either $\boldsymbol{b} = \boldsymbol{\tau}^{(B)}(\theta_B)$ or $\boldsymbol{b'} = \boldsymbol{\tau}^{(B)}(\theta_B)$, then under the local hidden-variable assumption the CHSH inequality applies.

If we plug in the quantum predictions for the expectation values, we obtain

$$\left|\cos 2(\theta_B - \theta_A) + \cos 2(\theta_B - \theta'_A) + \cos 2(\theta'_B - \theta_A) - \cos 2(\theta'_B - \theta'_A)\right| \leq 2 \ . \tag{4.79}$$

The maximal violation that saturates Cirel'son's inequality — left-hand side equal to $2\sqrt{2}$ — occurs when θ'_A, θ_B, θ_A and θ'_B are separated by successive $22\frac{1}{2}^\circ$ angles, so that

$$\begin{aligned} \frac{1}{\sqrt{2}} &= \cos 2(\theta_B - \theta_A) = \cos 2(\theta_B - \theta'_A) \\ &= = \cos 2(\theta'_B - \theta_A) = -\cos 2(\theta'_B - \theta'_A) \ . \end{aligned} \tag{4.80}$$

4.3.6 Experiments and loopholes

Locality loophole. Experiments with entangled pairs of photons have tested the CHSH inequality in the form eq. (4.79). The experiments confirm the quantum predictions, and demonstrate convincingly that the CHSH inequality is violated. These experiments, then, seem to show that Nature cannot be accurately described by a local hidden-variable theory.

Or do they? A skeptic might raise objections. For example, in the derivation of the CHSH inequality, we assumed that after Alice decides to measure either $\boldsymbol{a}$ or $\boldsymbol{a}'$, no information about Alice's decision reaches Bob's detector before Bob measures (and likewise, we assumed that if Bob measures first, no information about Bob's decision reaches Alice before she measures). Otherwise, the marginal probability distribution for Bob's outcomes could be updated after Alice's measurement and before Bob's, so that the CHSH inequality need not apply. The assumption that no such update can occur is justified by relativistic causality if Alice's decision and measurement are events spacelike separated from Bob's decision and measurement. The skeptic would insist that the experiment fulfill this condition, which is called the *locality loophole.*

In 1982, Aspect and collaborators conducted an experiment that addressed the locality loophole. Two entangled photons were produced in the decay of an excited calcium atom, and each photon was directed by a switch to one of two polarization analyzers, chosen pseudo-randomly. The photons were detected about 12m apart, corresponding to a light travel time of about 40 ns. This time was considerably longer than either the cycle time of the switch, or the difference in the times of arrival of the two photons. Therefore the "decision" about which observable to measure was made after the photons were already in flight, and the events that selected the axes for the measurement of photons A and B were spacelike separated. The results were consistent with the quantum predictions, and violated the CHSH inequality by five standard deviations. Since Aspect, many other experiments have confirmed this finding, including ones in which detectors A and B are kilometers apart.

Detection loophole. Another objection that the skeptic might raise is called the *detection loophole.* In experiments with photons, the detection efficiency is low. Most entangled photon pairs do not result in detections at both A and B. Among the things that can go wrong: a photon might be absorbed before reaching the detector, a photon might miss the detector, or a photon might arrive in the detector but fail to trigger it. Data is accepted by the experiment only when two photons are detected in coincidence, and in testing the CHSH inequality, we must assume that the data collected is a fair sample of all the entangled pairs.

But, what if the local hidden variables govern not just *what* polarization state is detected, but also *whether* the detector fires at all? Then the data collected might be a biased sample, and the CHSH inequality need not apply.

In Exercise 4.??, we will show that the detection loophole can be closed if the photons are detected with an efficiency above 82.84%. Current experiments with photons don't approach the necessary efficiency. Experiments that use ion traps have tested the CHSH inequality with detection efficiency close to 100%, but these experiments do not address the locality loophole. No experiment that simultaneously closes the locality and detection loopholes has yet been done.

Free-will loophole. Suppose that an experiment is done in which the photon detection efficiency is perfect, and in which Alice and Bob appear to make spacelike-separated decisions. A skeptic might still resist the conclusion that local hidden-variable theories are ruled out, by invoking the *free-will loophole.* Conceivably, the decisions that Alice and Bob make about what to measure are themselves governed by the local hidden variables. Then their decisions might be correlated with the values of the hidden variables that determine the measurement outcomes, so that they are unable to obtain a fair sample of the distribution of the hidden variables, and the CHSH inequality might be violated.

All of us have to decide for ourselves how seriously to take this objection.

4.4 Using entanglement

After Bell's work, quantum entanglement became a subject of intensive study, among those interested in the foundations of quantum theory.

Gradually, a new viewpoint evolved: entanglement is not just a unique tool for exposing the weirdness of quantum mechanics, but also a potentially valuable *resource*. By exploiting entangled quantum states, we can perform tasks that are otherwise difficult or impossible.

4.4.1 Dense coding

Our first example is an application of entanglement to communication. Alice wants to send messages to Bob. She might send classical bits (like dots and dashes in Morse code), but let's suppose that Alice and Bob are linked by a *quantum* channel. For example, Alice can prepare qubits (like photons) in any polarization state she pleases, and send them to Bob, who measures the polarization along the axis of his choice. Is there any advantage to sending qubits instead of classical bits?

In principle, if their quantum channel has perfect fidelity, and Alice and Bob perform the preparation and measurement with perfect efficiency, then they are no *worse* off using qubits instead of classical bits. Alice can prepare, say, either $|\uparrow_z\rangle$ or $|\downarrow_z\rangle$, and Bob can measure along $\hat{z}$ to infer the choice she made. This way, Alice can send one classical bit with each qubit. But in fact, that is the best she can do. Sending one qubit at a time, no matter how she prepares it and no matter how Bob measures it, no more than one classical bit can be carried by each qubit (even if the qubits are entangled with one another). This statement, a special case of the Holevo bound on the classical information capacity of a quantum channel, will be derived in Chapter 5.

But now, let's change the rules a bit — let's suppose that Alice and Bob share an entangled pair of qubits in the state $|\phi^+\rangle_{AB}$. The pair was prepared last year; one qubit was shipped to Alice and the other to Bob, in the hope that the shared entanglement would come in handy someday. Now, use of the quantum channel is very expensive, so Alice can afford to send only one qubit to Bob. Yet it is of the utmost importance for Alice to send Bob *two* classical bits of information.

Fortunately, Alice remembers about the entangled state $|\phi^+\rangle_{AB}$ that she shares with Bob, and she carries out a protocol that she and Bob had arranged for just such an emergency. On her member of the entangled pair, she can perform one of four possible unitary transformations:

1) $\boldsymbol{I}$ (she does nothing) ,

2) $\boldsymbol{\sigma}_1$ (180^o rotation about $\hat{x}$-axis) ,

3) $\boldsymbol{\sigma}_2$ (180^o rotation about $\hat{y}$-axis) ,

4) $\boldsymbol{\sigma}_3$ (180^o rotation about $\hat{z}$-axis) .

As we have seen, by doing so, she transforms $|\phi^+\rangle_{AB}$ to one of 4 mutually orthogonal states:

1) $|\phi^+\rangle_{AB}$,

2) $|\psi^+\rangle_{AB}$,

3) $|\psi^-\rangle_{AB}$ (up to a phase) ,

4) $|\phi^-\rangle_{AB}$.

Now, she sends her qubit to Bob, who receives it and then performs an orthogonal collective measurement on the pair that projects onto the maximally entangled basis. The measurement outcome unambiguously distinguishes the four possible actions that Alice could have performed. Therefore the single qubit sent from Alice to Bob has successfully carried 2 bits of classical information! Hence this procedure is called "dense coding."

A nice feature of this protocol is that, if the message is highly confidential, Alice need not worry that an eavesdropper will intercept the transmitted qubit and decipher her message. The transmitted qubit has density matrix $\boldsymbol{\rho}_A = \frac{1}{2}\boldsymbol{I}_A$, and so carries no information at all. All the information is in the correlations between qubits A and B, and this information is inaccessible unless the adversary is able to obtain both members of the entangled pair. (Of course, the adversary *can* "jam" the channel, preventing the information from reaching Bob.)

From one point of view, Alice and Bob really *did* need to use the channel twice to exchange two bits of information. For example, we can imagine that Alice prepared the state $|\phi^+\rangle$ herself. Last year, she sent half of the state to Bob, and now she sends him the other half. So in effect, Alice has sent two qubits to Bob in one of four mutually orthogonal states, to convey two classical bits of information as the Holevo bound allows.

Still, dense coding is rather weird, for several reasons. First, Alice sent the first qubit to Bob long before she knew what her message was going to be. Second, each qubit by itself carries no information at all; all the information is encoded in the correlations between the qubits. Third, it would work just as well for Bob to prepare the entangled pair and send half to Alice; then two classical bits are transmitted from Alice to Bob by sending a single qubit from Bob to Alice and back again.

Anyway, when an emergency arose and two bits had to be sent immediately while only one use of the channel was possible, Alice and Bob could exploit the pre-existing entanglement to communicate more efficiently. They used entanglement as a resource.

4.4.2 Quantum teleportation

In dense coding, quantum information could be exploited to enhance the transmission of classical information. Specifically, if Alice and Bob share entanglement, then sending one qubit is sufficient to convey two classical bits. Now one wonders about the converse. If Alice and Bob share entanglement, can sending two classical bits suffice to convey a qubit?

Imagine that Charlie has prepared for Alice a qubit in the state $|\psi\rangle$, but Alice doesn't know anything about what state Charlie prepared. Bob needs this qubit desperately, and Alice wants to help him. But that darn quantum channel is down again! Alice can send only *classical* information to Bob.

She could try measuring $\vec{\boldsymbol{\sigma}} \cdot \hat{n}$, projecting her qubit to either $|\uparrow_{\hat{n}}\rangle$ or $|\downarrow_{\hat{n}}\rangle$. She could send the one-bit measurement outcome to Bob who could then proceed to prepare the state that Alice found. But you showed in Exercise ?? that Bob's state $|\varphi\rangle$ will not be a perfect copy of Alice's; on the average it will match Alice's qubit with fidelity

$$F = |\langle\varphi|\psi\rangle|^2 = \frac{2}{3}, \tag{4.81}$$

This fidelity is better than could have been achieved if Bob had merely chosen a state at random ($F = \frac{1}{2}$), but it is not nearly as good as the fidelity that Bob requires. Furthermore, as we will see in Chapter 5, there is no protocol in which Alice measures the qubit and sends classical information to Bob that achieves a fidelity better than 2/3.

Fortunately, Alice and Bob recall that they share the maximally entangled state $|\phi^+\rangle_{AB}$, which they prepared last year. Why not use the entanglement as a *resource*? If they are willing to consume the shared entanglement and communicate classically, can Alice send her qubit to Bob with fidelity better than 2/3?

In fact they can achieve fidelity $F = 1$, by carrying out the following protocol: Alice unites the unknown qubit $|\psi\rangle_C$ she wants to send to Bob with her half of the $|\phi^+\rangle_{AB}$ pair that she shares with Bob. She measures the two commuting observables

$$\boldsymbol{\sigma}_1^{(C)} \otimes \boldsymbol{\sigma}_1^{(A)} \ , \quad \boldsymbol{\sigma}_3^{(C)} \otimes \boldsymbol{\sigma}_3^{(A)} \ , \tag{4.82}$$

thus performing *Bell measurement* — a projection of the two qubits onto one of the four maximally entangled states $|\phi^\pm\rangle_{CA}, |\psi^\pm\rangle_{CA}$. She sends her measurement outcome (two bits of classical information) to Bob over the classical channel. Upon receiving this information, Bob performs one of four operations on his qubit

$$
\begin{array}{llll}
\text{Alice measures} & |\phi^+\rangle_{CA} & \rightarrow & \text{Bob applies} \quad \boldsymbol{I}^{(B)}\ , \\
\text{Alice measures} & |\psi^+\rangle_{CA} & \rightarrow & \text{Bob applies} \quad \boldsymbol{\sigma}_1^{(B)}\ , \\
\text{Alice measures} & |\psi^-\rangle_{CA} & \rightarrow & \text{Bob applies} \quad \boldsymbol{\sigma}_2^{(B)}\ , \\
\text{Alice measures} & |\phi^-\rangle_{CA} & \rightarrow & \text{Bob applies} \quad \boldsymbol{\sigma}_3^{(B)}\ .
\end{array}
\tag{4.83}
$$

This action transforms Bob's qubit (his member of the entangled pair that he initially shared with Alice) into a perfect copy of $|\psi\rangle_C$. This magic trick is called *quantum teleportation*.

How does it work? We merely note that for $|\psi\rangle = a|0\rangle + b|1\rangle$, we may write

$$
\begin{aligned}
|\psi\rangle_C|\phi^+\rangle_{AB} &= (a|0\rangle_C + b|1\rangle_C)\frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB}) \\
&= \frac{1}{\sqrt{2}}(a|000\rangle_{CAB} + a|011\rangle_{CAB} + b|100\rangle_{CAB} + b|111\rangle_{CAB}) \\
&= \frac{1}{2}a(|\phi^+\rangle_{CA} + |\phi^-\rangle_{CA})|0\rangle_B + \frac{1}{2}a(|\psi^+\rangle_{CA} + |\psi^-\rangle_{CA})|1\rangle_B \\
&+ \frac{1}{2}b(|\psi^+\rangle_{CA} - |\psi^-\rangle_{CA})|0\rangle_B + \frac{1}{2}b(|\phi^+\rangle_{CA} - |\phi^-\rangle_{CA})|1\rangle_B \\
&= \frac{1}{2}|\phi^+\rangle_{CA}(a|0\rangle_B + b|1\rangle_B) \\
&+ \frac{1}{2}|\psi^+\rangle_{CA}(a|1\rangle_B + b|0\rangle_B) \\
&+ \frac{1}{2}|\psi^-\rangle_{CA}(a|1\rangle_B - b|0\rangle_B) \\
&+ \frac{1}{2}|\phi^-\rangle_{CA}(a|0\rangle_B - b|1\rangle_B) \\
&= \frac{1}{2}|\phi^+\rangle_{CA}|\psi\rangle_B + \frac{1}{2}|\psi^+\rangle_{CA}\boldsymbol{\sigma}_1|\psi\rangle_B \\
&+ \frac{1}{2}|\psi^-\rangle_{CA}(-i\boldsymbol{\sigma}_2)|\psi\rangle_B + \frac{1}{2}|\phi^-\rangle_{CA}\boldsymbol{\sigma}_3|\psi\rangle_B.
\end{aligned}
\tag{4.84}
$$

Thus we see that when Alice performs the Bell measurement on qubits C and A, all four outcomes are equally likely. Once Bob learns Alice's measurement outcome, he possesses the pure state $\boldsymbol{\sigma}|\psi\rangle$, where $\boldsymbol{\sigma}$ is a known Pauli operator, one of $\{\boldsymbol{I}, \boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3\}$. The action prescribed in eq. (4.83) restores Bob's qubit to the initial state $|\psi\rangle$.

Quantum teleportation is a curious procedure. Initially, Bob's qubit is completely uncorrelated with the unknown qubit $|\psi\rangle_C$, but Alice's Bell measurement establishes a correlation between A and C. The measurement outcome is in fact completely random, so Alice (and Bob) actually acquire no information at all about $|\psi\rangle$ by making this measurement. And that is a good thing, as we know that if they were to collect any information about the state they would unavoidably disturb the state.

How then does the quantum state manage to travel from Alice to Bob? It is a bit puzzling. On the one hand, we can hardly say that the two classical bits that were transmitted carried this information — the bits were random. So we are tempted to say that the shared entangled pair made the teleportation possible. But remember that the entangled pair was actually prepared last year, long before Alice ever dreamed that she would be sending the qubit to Bob ...

We should also note that the teleportation procedure is fully consistent with the no-cloning principle. True, a copy of the state $|\psi\rangle_B$ appeared in Bob's hands. But the original $|\psi\rangle_C$ had to be destroyed by Alice's measurement before the copy could be created.

Our findings about dense coding and quantum teleportation can be summarized as statements about how one type of resource can simulate another. Let us introduce the terminology *ebit* for an entangled pair of qubits shared by two parties (e for *entangled*), and *cbit* for a classical bit (c for *classical*). We teleport one qubit from Alice to Bob by consuming one ebit and sending two cbits, and we send two cbits from Alice and Bob via dense coding by consuming one ebit and transporting one qubit. Thus we may say

$$\begin{aligned} &1 \text{ ebit } + \text{ 2 cbits} &&\rightarrow &&\text{1 qubit} , \\ &1 \text{ ebit } + \text{ 1 qubit} &&\rightarrow &&\text{2 cbits} , \end{aligned} \tag{4.85}$$

meaning that the resources on the left suffice to simulate the resources on the right. Entanglement is essential in these protocols. Without ebits, a qubit is worth only one cbit, and without ebits, a "teleported" qubit has fidelity $F \leq 2/3$.

4.4.3 Quantum teleportation and maximal entanglement

The teleportation concept has an air of mystery. One would like to understand more deeply why it works. A helpful clue is that to teleport with fidelity $F = 1$ the entangled state consumed in the protocol must be *maximally* entangled. And the crucial feature of bipartite maximally entangled states is that *either Alice or Bob* can transform one maximally entangled state to another by applying a local unitary transformation.

To see more clearly how quantum teleportation works, consider teleporting an N-dimensional system using an $N \times N$ maximally entangled state of the form

$$|\Phi\rangle = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} |i\rangle \otimes |i\rangle \ . \tag{4.86}$$

A useful property of this state is

$$\begin{aligned} {}_{CA}\langle \Phi | \Phi \rangle_{AB} &= \frac{1}{N} \sum_{i,j} \left({}_C\langle i| \otimes {}_A\langle i| \right) \left(|j\rangle_A \otimes |j\rangle_B \right) \\ &= \frac{1}{N} \sum_i |i\rangle_B \ {}_C\langle i| \equiv \frac{1}{N} (\boldsymbol{T})_{BC} \end{aligned} \tag{4.87}$$

Here we have defined the *transfer operator* $(\boldsymbol{T})_{BC}$ which has the property

$$\boldsymbol{T}_{BC} |\varphi\rangle_C = \boldsymbol{T}_{BC} \left(\sum_i a_i |i\rangle_C \right) = \sum_i a_i |i\rangle_B = |\varphi\rangle_B \ ; \tag{4.88}$$

it maps a state in C to the identical state in B. This property has no invariant meaning independent of the choice of basis in B and C; rather $\boldsymbol{T}_{BC}$ just describes an arbitrary way to relate the orthonormal bases of the two systems. Of course, Alice and Bob would need to align their bases in some way to verify that teleportation has really succeeded.

Now recall that any other $N \times N$ maximally entangled state has a Schmidt decomposition of the form

$$\frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} |i'\rangle \otimes |i\rangle \ , \tag{4.89}$$

and so can be expressed as

$$|\Phi(\boldsymbol{U})\rangle \equiv \boldsymbol{U} \otimes \boldsymbol{I} |\Phi\rangle \ , \tag{4.90}$$

where

$$\boldsymbol{U}|i\rangle = |i'\rangle = \sum_j |j\rangle U_{ji} \ . \tag{4.91}$$

Writing

$$|\Phi(\boldsymbol{U})\rangle_{AB} = \frac{1}{\sqrt{N}} \sum_{i,j} |j\rangle_A \otimes |i\rangle_B \ U_{ji} \ , \tag{4.92}$$

we can easily verify that

$${}_{CA}\langle \Phi(\boldsymbol{U}) | \Phi(\boldsymbol{V}^T) \rangle_{AB} = \frac{1}{N} \left(\boldsymbol{V} \boldsymbol{U}^{-1} \right)_B \boldsymbol{T}_{BC} \ , \tag{4.93}$$

where $\boldsymbol{V}^T$ denotes the transpose of $\boldsymbol{V}$ in the standard basis ($V^T_{ij} = V_{ji}$); in particular, then, the transfer operator can be expressed as

$$\frac{1}{N}\boldsymbol{T}_{BC} = {}_{CA}\langle\Phi(\boldsymbol{U})|\Phi((\boldsymbol{U}^T)\rangle_{AB} \ , \tag{4.94}$$

for any unitary $\boldsymbol{U}$.

Now suppose that Alice and Bob share $|\Phi\rangle_{AB}$, and that Charlie has prepared the state $|\psi\rangle_C$ and has deposited it in Alice's laboratory. Alice performs a measurement that projects CA onto a maximally entangled basis, obtaining the outcome $|\Phi(\boldsymbol{U}_a)\rangle_{CA}$ for some unitary $\boldsymbol{U}_a$. Then we know from eq. (4.94) that *if* Alice and Bob had shared the state $|\Phi((\boldsymbol{U}^T_a)\rangle_{AB}$ instead of $|\Phi\rangle_{AB}$, then Alice's measurement would have prepared in Bob's lab a perfect replica of the state $|\psi\rangle$. Unfortunately, they did not have the foresight to share the right state to begin with. But it's not too late! Bob realizes that

$$|\Phi(\boldsymbol{U}^T_a)\rangle = \boldsymbol{I}_A \otimes (\boldsymbol{U}_a)_B \, |\Phi\rangle_{AB} \ , \tag{4.95}$$

and of course $(\boldsymbol{U}_a)_B$ commutes with Alice's measurement. Hence, when Bob hears from Alice that her measurement outcome was $|\Phi((\boldsymbol{U}^T_a)\rangle_{AB}$, he applies $(\boldsymbol{U}_a)_B$ to his half of the state he had shared with Alice. Then the protocol is equivalent to one in which they had shared the right maximally entangled state to begin with, and Bob's state has been transformed into $|\psi\rangle_B$!

This approach to teleportation has some conceptual advantages. For one, we can easily see that Alice is not required to perform an orthogonal measurement. To achieve teleportation with fidelity $F = 1$ it suffices that she perform a POVM with operation elements $\boldsymbol{M}_a$, where each $\boldsymbol{M}_a$ has the property

$$\boldsymbol{M}^\dagger_a\boldsymbol{M}_a \propto |\Phi(\boldsymbol{U}_a)\rangle\langle\Phi(\boldsymbol{U}_a)| \tag{4.96}$$

for some unitary $\boldsymbol{U}_a$. Also, we can easily see how the teleportation protocol should be modified if the initial maximally entangled state that Alice and Bob share is not $|\Phi\rangle_{AB}$ but rather

$$|\Phi(\boldsymbol{V}^T)\rangle_{AB} = \boldsymbol{I}_A \otimes \boldsymbol{V}_B|\Phi\rangle_{AB} \ . \tag{4.97}$$

If Alice's measurement outcome is $|\Phi(\boldsymbol{U}_a)\rangle_{CA}$, then eq. (4.93) tells us that the state Bob receives is

$$\boldsymbol{V}\boldsymbol{U}^{-1}_a|\psi\rangle_B \ . \tag{4.98}$$

To recover $|\psi\rangle_B$, Bob must apply $\boldsymbol{U}_a\boldsymbol{V}^{-1}$.

The operator ordering in eq. (4.98) may seem counterintuitive at first — it seems as though Alice's measurement ($\boldsymbol{U}_a$) precedes the preparation of the shared entangled state ($\boldsymbol{V}$). But this "time reversal" has a straightforward interpretation. If Alice's measurement outcome is $|\Phi(\boldsymbol{U}_a)\rangle_{CA}$, then Bob would have received a perfect copy of $|\psi\rangle$ if the initial entangled state had been $\boldsymbol{I}_A \otimes (\boldsymbol{U}_a)_B\, |\Phi\rangle_{AB}$. To simulate the situation in which the entangled state had been chosen properly from the start, Bob first applies $\boldsymbol{V}^{-1}$ to undo the "twist" in $|\Phi(\boldsymbol{V}^T)\rangle_{AB}$, recovering $|\Phi\rangle_{AB}$, and then applies $\boldsymbol{U}_a$ to transform the entangled state to the desired one.

There is a more fanciful interpretation of eq. (4.98) which, while not necessary, is nonetheless irresistable. We might "explain" how quantum information is transferred from Alice and Bob by following the world line of a qubit traveling in spacetime. The qubit moves forward in time from Charlie's preparation to Alice's measurement, then backward in time from the measurement to the initial preparation of the entangled pair, and finally forward in time again from the preparation of the pair to Bob's laboratory. Since this world line visits Alice's measurement before arriving at the preparation of the entanglement, $\boldsymbol{U}_a^{-1}$ acts "first" and $\boldsymbol{V}$ acts "later on."

4.4.4 Quantum software

Teleportation has some interesting applications. For example, imagine that Alice and Bob wish to apply the "quantum gate" $\boldsymbol{V}$ to the unknown state $|\psi\rangle_C$. But applying $\boldsymbol{V}$ requires sophisitcated hardware that they can't afford.

A more economical alternative is to purchase *quantum software* from a vendor. The software is a bipartite state that the vendor certifies to be

$$|\Phi(\boldsymbol{V}^T)\rangle_{AB} = \boldsymbol{I}_A \otimes \boldsymbol{V}_B |\phi\rangle_{AB} \ . \tag{4.99}$$

Alice's hardware is powerful enough for her to perform a measurement that projects onto the basis $\{|\Phi(\boldsymbol{U}_a)\rangle_{CA}\}$; once the outcome a is known, the state $\boldsymbol{V}\boldsymbol{U}_a^{-1}|\psi\rangle_B$ has been prepared. Bob can then complete the execution of $\boldsymbol{V}$ to $|\psi\rangle$ by applying $\boldsymbol{V}\boldsymbol{U}_a\boldsymbol{V}^{-1}$

This procedure may seem silly — why assume that Bob is able to apply $\boldsymbol{V}\boldsymbol{U}_a\boldsymbol{V}^{-1}$ but unable to apply $\boldsymbol{V}$? In fact it is not so silly, and has important applications to fault-tolerant quantum computation that we will explore further in Chapter 8. In some cases, executing $\boldsymbol{V}\boldsymbol{U}_a\boldsymbol{V}^{-1}$ really is a lot easier than applying $\boldsymbol{V}$. Furthermore, Alice and Bob might be able to prepare the quantum software themselves, instead of buying it, even though they can't apply $\boldsymbol{V}$ reliably. This is possible because it is easier to verify that a *known* quantum state has been properly prepared than to verify that a known unitary transformation has been successfully applied to an unknown state. If the hardware that applies $\boldsymbol{V}$ cannot be trusted, then we prefer to use it to prepare software offline, and then subject the software to quality assurance, rather than risk causing irrevocable damage to our unknown state through a faulty execution of $\boldsymbol{V}$.

Each application of $\boldsymbol{V}$ consumes one copy of the quantum software. Thus, this protocol for executing $\boldsymbol{V}$ with the help of quantum software uses entanglement as a resource.

4.5 Quantum cryptography

4.5.1 EPR quantum key distribution

Everyone has secrets, including Alice and Bob. Alice needs to send a highly private message to Bob, but Alice and Bob have a very nosy friend, Eve, who they know will try to listen in. Can they communicate with assurance that Eve is unable to eavesdrop?

Obviously, they should use some kind of code. Trouble is, aside from being very nosy, Eve is also very smart. Alice and Bob are not confident that they are clever enough to devise a code that Eve cannot break.

Except there is one coding scheme that is surely unbreakable. If Alice and Bob share a *private key*, a string of random bits known only to them, then Alice can convert her message to ASCII (a string of bits no longer than the key) *add* each bit of her message (module 2) to the corresponding bit of the key, and send the result to Bob. Receiving this string, Bob can add the key to it to extract Alice's message.

This scheme is secure because even if Eve should intercept the transmission, she will not learn anything because the transmitted string itself carries no information — the message is encoded in a correlation between the transmitted string and the *key* (which Eve doesn't know).

There is still a problem, though, because Alice and Bob need to establish a shared random key, and they must ensure that Eve can't know the key. They could meet to exchange the key, but that might be impractical. They could entrust a third party to transport the key, but what if the intermediary is secretly in cahoots with Eve? They could use "public key" distribution protocols, but the security of such protocols is founded on assumptions about the computational resources available to a potential adversary. Indeed, we will see in Chapter 6 that public key protocols are vulnerable to attack by an eavesdropper who is equipped with a quantum computer.

Can Alice and Bob exploit *quantum* information (and specifically entanglement) to solve the key exchange problem? They can! *Quantum key distribution* protocols can be devised that are invulnerable to any attack allowed by the laws of physics.

Let's suppose that Alice and Bob share a supply of entangled pairs, each prepared in the state $|\phi^+\rangle$. To establish a shared private key, they may carry out this protocol:

For each qubit in her/his possession, Alice and Bob decide to measure either $\boldsymbol{\sigma}_1$ or $\boldsymbol{\sigma}_3$. The decision is pseudo-random, each choice occuring with probability 1/2. Then, after the measurements are performed, both Alice and Bob publicly announce what observables they measured, but do not reveal the outcomes they obtained. For those cases (about half) in which they measured their qubits along different axes, their results are discarded (as Alice and Bob obtained uncorrelated outcomes). For those cases in which they measured along the same axis, their results, though random, are *perfectly correlated.* Hence, they have established a shared random key.

But, is this protocol really invulnerable to a sneaky attack by Eve? In particular, Eve might have clandestinely tampered with the pairs at some time in the past. Then the pairs that Alice and Bob possess might be (unbeknownst to Alice and Bob) not perfect $|\phi^+\rangle$'s, but rather pairs that are entangled with qubits in Eve's possession. Eve can then wait until Alice and Bob make their public announcements, and proceed to measure her qubits in a manner designed to acquire maximal information about the results that Alice and Bob obtained. Alice and Bob must protect themselves against this type of attack.

If Eve has indeed tampered with Alice's and Bob's pairs, then the most general possible state for an AB pair and a set of E qubits has the form

$$\begin{aligned} |\Upsilon\rangle_{ABE} = {} & |00\rangle_{AB}|e_{00}\rangle_E + |01\rangle_{AB}|e_{01}\rangle_E \\ & + |10\rangle_{AB}|e_{10}\rangle_E + |11\rangle_{AB}|e_{11}\rangle_E \ , \end{aligned} \tag{4.100}$$

where Eve's states $|e_{ij}\rangle_E$ are neither normalized nor mutually orthogonal. Now recall that the defining property or $|\phi^+\rangle$ is that it is an eigenstate with eigenvalue +1 of both $\boldsymbol{\sigma}_1^{(A)}\boldsymbol{\sigma}_1^{(B)}$ and $\boldsymbol{\sigma}_3^{(A)}\boldsymbol{\sigma}_3^{(B)}$. Suppose that A and B are able to verify that the pairs in their possession have this property. To satisfy $\boldsymbol{\sigma}_3^{(A)}\boldsymbol{\sigma}_3^{(B)} = 1$, we must have

$$|\Upsilon\rangle_{AB} = |00\rangle_{AB}|e_{00}\rangle_E + |11\rangle_{AB}|e_{11}\rangle_E \ , \tag{4.101}$$

and to also satisfy $\boldsymbol{\sigma}_1^{(A)}\boldsymbol{\sigma}_1^{(B)} = 1$, we must have

$$|\Upsilon\rangle_{ABE} = \frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB})|e\rangle_E = |\phi^+\rangle_{AB}|e\rangle_E \ . \tag{4.102}$$

We see that it is possible for the AB pairs to be eigenstates of $\boldsymbol{\sigma}_1^{(A)}\boldsymbol{\sigma}_1^{(B)}$ and $\boldsymbol{\sigma}_3^{(A)}\boldsymbol{\sigma}_3^{(B)}$ only if they are completely unentangled with Eve's qubits.

Therefore, Eve will not be able to learn anything about Alice's and Bob's measurement results by measuring her qubits. The random key is secure.

To verify the properties $\boldsymbol{\sigma}_1^{(A)}\boldsymbol{\sigma}_1^{(B)} = 1 = \boldsymbol{\sigma}_3^{(A)}\boldsymbol{\sigma}_3^{(B)}$, Alice and Bob can sacrifice a portion of their shared key, and publicly compare their measurement outcomes. They should find that their results are indeed perfectly correlated. If so they will have high statistical confidence that Eve is unable to intercept the key. If not, they have detected Eve's nefarious activity. They may then discard the key, and make a fresh attempt to establish a secure key.

As I have just presented it, the quantum key distribution protocol seems to require entangled pairs shared by Alice and Bob, but this is not really so. We might imagine that Alice prepares the $|\phi^+\rangle$ pairs herself, and then measures one qubit in each pair before sending the other to Bob. This is completely equivalent to a scheme in which Alice prepares one of the four states

$$|\uparrow_z\rangle, |\downarrow_z\rangle, |\uparrow_x\rangle, |\downarrow_x\rangle, \tag{4.103}$$

(chosen at random, each occuring with probability 1/4) and sends the qubit to Bob. Bob's measurement and the verification are then carried out as before. This scheme (known as the BB84 quantum key distribution protocol) is just as secure as the entanglement-based scheme.†

Another intriguing variation is called the "time-reversed EPR" scheme. Here both Alice and Bob prepare one of the four states in eq. (4.103), and they both send their qubits to Charlie. Then Charlie performs a Bell measurement on the pair — that is, he measures $\boldsymbol{\sigma}_1^{(A)}\boldsymbol{\sigma}_1^{(B)}$ and $\boldsymbol{\sigma}_3^{(A)}\boldsymbol{\sigma}_3^{(B)}$, orthogonally projecting out one of $|\phi^\pm\rangle|\psi^\pm\rangle$, and he publicly announces the result. Since all four of these states are simultaneous eigenstates of $\boldsymbol{\sigma}_1^{(A)}\boldsymbol{\sigma}_1^{(B)}$ and $\boldsymbol{\sigma}_3^{(A)}\boldsymbol{\sigma}_3^{(B)}$, when Alice and Bob both prepared their spins along the same axis (as they do about half the time) they share a single bit.‡ Of course, Charlie could be allied with Eve, but Alice and Bob can verify that Charlie and Eve have acquired no information as before, by comparing a portion of their key. This scheme has the advantage that Charlie could operate a central switching station by storing qubits received from many parties, and then perform his Bell measurement when two of the parties request a secure communication link. (Here we assume that Charlie has a stable quantum memory in which qubits can be stored

† Except that in the EPR scheme, Alice and Bob can wait until just before they need to talk to generate the key, thus reducing the risk that Eve might at some point burglarize Alice's safe to learn what states Alice prepared (and so infer the key).

‡ Until Charlie makes his measurement, the states prepared by Bob and Alice are totally uncorrelated. A definite correlation (or anti-correlation) is established after Charlie performs his measurement.

accurately for as long as necessary.) A secure key can be established even if the quantum communication line is down temporarily, as long as both parties had the foresight to send their qubits to Charlie on an earlier occasion (when the quantum channel was open).

So far, we have made the unrealistic assumption that the quantum communication channel is perfect, but of course in the real world errors will occur. Therefore even if Eve has been up to no mischief, Alice and Bob will sometimes find that their verification test will fail. But how are they to distinguish errors due to imperfections of the channel from errors that occur because Eve has been eavesdropping?

To address this problem, Alice and Bob can enhance their protocol in two ways. First they implement (classical) error correction to reduce the effective error rate. For example, to establish each bit of their shared key they could actually exchange a block of three random bits. If the three bits are not all the same, Alice can inform Bob which of the three is different than the other two; Bob can flip that bit in his block, and *then* use majority voting to determine a bit value for the block. This way, Alice and Bob share the same key bit even if an error occured for one bit in the block of three.

However, error correction alone does not suffice to ensure that Eve has acquired negligible information about the key — error correction must be supplemented by (classical) privacy amplification. For example, after performing error correction so that they are confident that they share the same key, Alice and Bob might extract a bit of "superkey" as the *parity* of n key bits. To know *anything* about the parity of n bits, Eve would need to know *something* about each of the bits. Therefore, the parity bit is considerably more secure, on the average, than each of the individual key bits.

If the error rate of the channel is low enough, one can show that quantum key distribution, supplemented by error correction and privacy amplification, is invulnerable to any attack that Eve might muster (in the sense that the information acquired by Eve can be guaranteed to be arbitrarily small). We will return to the problem of proving the security of quantum key distribution in Chapter 7.

4.5.2 No cloning

The security of quantum key distribution is based on an essential difference between quantum information and classical information. It is not possible to acquire information that *distinguishes* between nonorthogonal quantum states without *disturbing* the states.

For example, in the BB84 protocol, Alice sends to Bob any one of the four states $|\uparrow_z\rangle|\downarrow_z\rangle|\uparrow_x\rangle|\downarrow_x\rangle$, and Alice and Bob are able to verify that none of their states are perturbed by Eve's attempt at eavesdropping. Suppose, more generally, that $|\varphi\rangle$ and $|\psi\rangle$ are two nonorthogonal states in $\mathcal{H}$ ($\langle\psi|\varphi\rangle \neq 0$) and that a unitary transformation U is applied to $\mathcal{H}\otimes\mathcal{H}_E$ (where $\mathcal{H}_E$ is a Hilbert space accessible to Eve) that leaves both $|\psi\rangle$ and $|\varphi\rangle$ undisturbed. Then

$$\begin{aligned} U: \quad & |\psi\rangle \otimes |0\rangle_E \to |\psi\rangle \otimes |e\rangle_E \ , \\ & |\varphi\rangle \otimes |0\rangle_E \to |\varphi\rangle \otimes |f\rangle_E \ , \end{aligned} \tag{4.104}$$

and unitarity implies that

$$\begin{aligned} \langle\psi|\phi\rangle &= ({}_E\langle 0| \otimes \langle\psi|)(|\varphi\rangle \otimes |0\rangle_E) \\ &= ({}_E\langle e| \otimes \langle\psi|)(|\varphi\rangle \otimes |f\rangle_E) \\ &= \langle\psi|\varphi\rangle\langle e|f\rangle \ . \end{aligned} \tag{4.105}$$

Hence, for $\langle\psi|\varphi\rangle \neq 0$, we have $\langle e|f\rangle = 1$, and therefore since the states are normalized, $|e\rangle = |f\rangle$. This means that no measurement in $\mathcal{H}_E$ can reveal any information that distinguishes $|\psi\rangle$ from $|\varphi\rangle$. In the BB84 case this argument shows that, if Eve does not disturb the states sent by Alice, then the state in $\mathcal{H}_E$ is the same irrespective of which of the four states $|\uparrow_z\rangle, |\downarrow_z\rangle, |\uparrow_x\rangle, |\downarrow_x\rangle$ is sent by Alice, and therefore Eve learns nothing about the key shared by Alice and Bob. On the other hand, if Alice is sending to Bob one of the two orthogonal states $|\uparrow_z\rangle$ or $|\downarrow_z\rangle$, there is nothing to prevent Eve from acquiring a copy of the information (as with classical bits).

We have noted earlier that if we have many identical copies of a qubit, then it is possible to measure the mean value of noncommuting observables like $\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2$, and $\boldsymbol{\sigma}_3$ to completely determine the density matrix of the qubit. Inherent in the conclusion that nonorthogonal state cannot be distinguished without disturbing them, then, is the implicit provision that it is not possible to make a perfect copy of a qubit. (If we could, we would make as many copies as we need to find $\langle\boldsymbol{\sigma}_1\rangle$, $\langle\boldsymbol{\sigma}_2\rangle$, and $\langle\boldsymbol{\sigma}_3\rangle$ to any specified accuracy.) Let's now make this point explicit: there is no such thing as a perfect quantum Xerox machine.

Orthogonal quantum states (like classical information) *can* be reliably copied. For example, the unitary transformation that acts as

$$\begin{aligned} U: \quad & |0\rangle_A|0\rangle_E \to |0\rangle_A|0\rangle_E \ , \\ & |1\rangle_A|0\rangle_E \to |1\rangle_A|1\rangle_E \ , \end{aligned} \tag{4.106}$$

copies the first qubit onto the second if the first qubit is in one of the states $|0\rangle_A$ or $|1\rangle_A$. But if instead the first qubit is in the state $|\psi\rangle = a|0\rangle_A + b|1\rangle_A$, then

$$\begin{aligned} U: \quad & (a|0\rangle_A + b|1\rangle_A)|0\rangle_E \\ & \to a|0\rangle_A|0\rangle_E + b|1\rangle_A|1\rangle_E \ . \end{aligned} \tag{4.107}$$

This is *not* the state $|\psi\rangle \otimes |\psi\rangle$ (a tensor product of the original and the copy); rather it is something very different – an entangled state of the two qubits.

To consider the most general possible quantum Xerox machine, we allow the full Hilbert space to be larger than the tensor product of the space of the original and the space of the copy. Then the most general "copying" unitary transformation acts as

$$\begin{aligned} U: \quad & |\psi\rangle_A|0\rangle_E|0\rangle_F \rightarrow |\psi\rangle_A|\psi\rangle_E|e\rangle_F \\ & |\varphi\rangle_A|0\rangle_E|0\rangle_F \rightarrow |\varphi\rangle_A|\varphi\rangle_E|f\rangle_F \ . \end{aligned} \tag{4.108}$$

Unitarity then implies that

$$\langle\psi|\varphi\rangle = \langle\psi|\varphi\rangle\langle\psi|\varphi\rangle\langle e|f\rangle \ ; \tag{4.109}$$

therefore, if $\langle\psi|\varphi\rangle \neq 0$, then

$$1 = \langle\psi|\varphi\rangle\langle e|f\rangle. \tag{4.110}$$

Since the states are normalized, we conclude that

$$|\langle\psi|\varphi\rangle| = 1, \tag{4.111}$$

so that $|\psi\rangle$ and $|\varphi\rangle$ actually represent the same ray. No unitary machine can make a copy of both $|\varphi\rangle$ and $|\psi\rangle$ if $|\varphi\rangle$ and $|\psi\rangle$ are *distinct*, *nonorthogonal* states. This result is called the no-cloning theorem.

4.6 Mixed-state entanglement

The crucial property of quantum entanglement is that it cannot be created *locally*. Up to now in this chapter we have limited our attention to the properties of entangled *pure* states, but it is important to recognize that mixed states can be entangled, too.

Recall that a bipartite pure state $|\Psi\rangle_{AB}$ is *separable* if and only if it is a product state $|\Psi\rangle_{AB} = |\alpha\rangle_A \otimes |\beta\rangle_B$. We say that a bipartite mixed state $\boldsymbol{\rho}_{AB}$ is separable if and only if it can be realized as an ensemble of separable pure states,

$$\boldsymbol{\rho}_{AB} = \sum_i p_i \left(|\alpha_i\rangle\langle\alpha_i|\right)_A \otimes \left(|\beta_i\rangle\langle\beta_i|\right)_B \ , \tag{4.112}$$

where the p_i's are positive and sum to one. Alternatively, we may say that $\boldsymbol{\rho}_{AB}$ is separable if and only if it can be expressed as

$$\boldsymbol{\rho}_{AB} = \sum_{i,j} p_{ij} \boldsymbol{\rho}_{A,i} \otimes \boldsymbol{\rho}_{B,j} \ , \tag{4.113}$$

where each $\boldsymbol{\rho}_{A,i}$ and $\boldsymbol{\rho}_{B,j}$ is a density operator, and the p_{ij}'s are positive and sum to one. Thus if a state is separable, the correlations between the state of part A and the state of part B are entirely classical, and embodied by the joint probability distribution p_{ij}. The two criteria eq. (4.112) and eq. (4.113) are equivalent because $\boldsymbol{\rho}_{A,i}$ and $\boldsymbol{\rho}_{B,j}$ can be realized as an ensemble of pure states.

Of course, it may be possible to realize a separable mixed state as an ensemble of entangled pure states as well. A simple example is that the random state $\boldsymbol{\rho} = \frac{1}{4}\boldsymbol{I} \otimes \boldsymbol{I}$ of two qubits can be expressed as either

$$\boldsymbol{\rho} = \frac{1}{4}\left(|00\rangle\langle 00| + |01\rangle\langle 10| + |10\rangle\langle 01| + |11\rangle\langle 11|\right) \tag{4.114}$$

(an ensemble of product states) or

$$\boldsymbol{\rho} = \frac{1}{4}\left(|\phi^+\rangle\langle\phi^+| + |\phi^-\rangle\langle\phi^-| + |\psi^+\rangle\langle\psi^+| + |\psi^-\rangle\langle\psi^-|\right) \tag{4.115}$$

(an ensemble of maximally entangled states). The state is separable if and only if there is *some* way to represent is as an ensemble of product states. As for a pure state, if a mixed state is not separable, we say that it is *inseparable* or *entangled.*

Consider two distantly separated parties Alice and Bob who carry out a protocol involving *local operations and classical communication.* That is, Alice is permitted to perform quantum operations on her system A, Bob is permitted to perform quantum operations on his system B, and Alice and Bob are permitted to exchange classical bits as many times as they want. But no exchange of qubits is permitted. Then if Alice and Bob share a separable state to start with, their state will still be separable at the end of the protocol. The reason is that neither a local operation nor exchange of a classical bit can increase the Schmidt number of a bipartite pure state from the value 1 to a value greater than 1. Of course, Alice and Bob might have a mixed state, but in each step of the protocol an ensemble of product states is transformed to another ensemble of product states. Alice and Bob cannot create entanglement locally if they have none to begin with. In discussions of entanglement, the concept of a protocol that uses only Local Operations and Classical Communication is so prevalent that we will find it convenient to use the abbreviation *LOCC.*

On the other hand, with LOCC, Alice and Bob can prepare any separable state. To prepare $\boldsymbol{\rho}_{AB}$ in eq. (4.112), Alice generates random numbers to sample the probability distribution $\{p_i\}$; if outcome i is found, she informs Bob, and Alice prepares the $|\alpha_i\rangle_A$ while Bob prepares $|\beta_i\rangle_B$.

4.6.1 Positive-partial-transpose criterion for separability

Now, consider a bipartite density operator $\boldsymbol{\rho}_{AB}$ on $\mathcal{H}_A \otimes \mathcal{H}_B$, presented as (say) a matrix in some basis. We would like to know whether $\boldsymbol{\rho}_{AB}$ is separable. How do we decide? It is not obvious how to devise an efficient algorithm that will definitively answer whether $\boldsymbol{\rho}_{AB}$ can be realized as an ensemble of product states. However, it is useful to note that there are necessary conditions for separability that are easy to check.

Recall that relative to a specified orthonormal basis $\{|i\rangle\}$ for a Hilbert space $\mathcal{H}$, a transpose operation T can be defined — the transpose acts on a basis for the linear operators according to

$$T : |i\rangle\langle j| \to \left(|i\rangle\langle j|\right)^T = |j\rangle\langle i| \ ; \qquad (4.116)$$

its action on a matrix M_{ij} expressed in this basis is

$$\left(M^T\right)_{ij} = M_{ji} \ . \qquad (4.117)$$

Evidently transposition preserves the trace of the matrix $\boldsymbol{M}$. If $\boldsymbol{M}$ is Hermitian, then its transpose is its complex conjugate, which has the same (real) eigenvalues. Therefore, the transpose of a density operator is another density operator with the same eigenvalues — the transpose is a trace-preserving positive map.

But we saw in §3.?? that the transpose, although positive, is not completely positive; that is, the *partial transpose* $I \otimes T$ can map a bipartite positive operator to an operator that is not positive. For example, the maximally entangled state

$$|\Phi\rangle = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} |i\rangle_A \otimes |i\rangle_B \qquad (4.118)$$

has density operator

$$\boldsymbol{\rho} = \frac{1}{N} \sum_{i,j} |ii\rangle\langle jj| \ . \qquad (4.119)$$

Its *partial transpose* is

$$(I \otimes T)(\boldsymbol{\rho}) = \frac{1}{N} \sum_{i,j} |ij\rangle\langle ji| = \frac{1}{N}\,(\mathrm{SWAP}) \ ; \qquad (4.120)$$

the SWAP operator has eigenstates with eigenvalue $+1$ (symmetric states) and eigenstates with eigenvalue -1 (antisymmetric states) — hence it is not positive. We will use the notation

$$\boldsymbol{\rho}^{PT} = (I \otimes T)(\boldsymbol{\rho}) \qquad (4.121)$$

for the partial transpose of the bipartite density operator $\boldsymbol{\rho}$.

While the partial transpose is not a positive map in general, it *is* positive acting on separable states. The partial transpose of $\boldsymbol{\rho}_{AB}$ in eq. (4.113) is

$$\boldsymbol{\rho}_{AB}^{PT} = \sum_{i,j} p_{ij} \boldsymbol{\rho}_{A,i} \otimes \boldsymbol{\rho}_{B,j}^{T} \ ; \tag{4.122}$$

since $\boldsymbol{\rho}_{B,j}^{T}$ is a density operator, so is $\boldsymbol{\rho}_{AB}^{PT}$. Thus we arrive at a useful necessary condition for separability.

> *Positive partial-transpose criterion for separability*: If $\boldsymbol{\rho}_{AB}$ is separable, then $\boldsymbol{\rho}_{AB}^{PT}$ is nonnegative.

We will say that a bipartite density operator is *PPT* (for "positive partial transpose) if its partial transpose is nonnegative.

Thus, if we are presented with a density operator $\boldsymbol{\rho}_{AB}$, we may compute the eigenvalues of $\boldsymbol{\rho}_{AB}^{PT}$; if negative eigenvalues are found, then $\boldsymbol{\rho}_{AB}$ is known to be inseparable. But because the PPT condition is necessary but not sufficient for separability, if $\boldsymbol{\rho}_{AB}^{PT}$ is found to be nonnegative, then whether $\boldsymbol{\rho}_{AB}$ is separable remains unsettled. The PPT criterion is sometimes called the *Peres-Horodecki* criterion for separability.

Let's apply the PPT criterion to a two-qubit state of the form

$$\boldsymbol{\rho}(\lambda) = \lambda |\phi^+\rangle\langle\phi^+| + \frac{1}{4}(1-\lambda)\boldsymbol{I} \ . \tag{4.123}$$

This state may also be expressed as

$$\begin{aligned} \boldsymbol{\rho}(F) &= F|\phi^+\rangle\langle\phi^+| \\ &\quad + \frac{1}{3}(1-F)\Big(|\phi^-\rangle\langle\phi^-| + |\psi^+\rangle\langle\psi^+| + |\psi^-\rangle\langle\psi^-|\Big) \ , \end{aligned} \tag{4.124}$$

where $(1-F) = \frac{3}{4}(1-\lambda)$, and as we saw in §3.??, it results from subjecting half of the state $|\phi^+\rangle$ to the depolarizing channel with error probability $p = 1-F$. This state is sometimes called a *Werner state* with fidelity F.

Now

$$\Big(|\phi^+\rangle\langle\phi^+|\Big)^{PT} = \frac{1}{2}(\mathrm{SWAP}) = \frac{1}{2}\boldsymbol{I} - |\psi^-\rangle\langle\psi^-| \ , \tag{4.125}$$

where the second equality follows from the property that $|\phi^\pm\rangle$, $|\psi^+\rangle$ (which are symmetric) are eigenvalues of SWAP with eigenvalue 1, and $|\psi^-\rangle$ (which is antisymmetric) is an eigenstate of SWAP with eigenvalue -1. Since also $\boldsymbol{I}^{PT} = \boldsymbol{I}$, we see that the partial transpose of a Werner state is

$$\begin{aligned} \boldsymbol{\rho}(\lambda)^{PT} &= \lambda\Big(\frac{1}{2}\boldsymbol{I} - |\psi^-\rangle\langle\psi^-|\Big) + \frac{1}{4}(1-\lambda)\boldsymbol{I} \\ &= \frac{1}{4}(1+\lambda)\boldsymbol{I} - \lambda|\psi^-\rangle\langle\psi^-| \ . \end{aligned} \tag{4.126}$$

This operator has a negative eigenvalue if $\lambda > 1/3$, and we conclude that the Werner state is inseparable for $\lambda > 1/3$. Therefore, if half of the maximally entangled state $|\phi^+\rangle$ is subjected to the depolarizing channel with error probability $p < 1/2$, the resulting state remains entangled.

Although we won't prove it here, it turns out that for the case of two-qubit states, the PPT criterion is both necessary and sufficient for separability. Thus the Werner state with $\lambda < 1/3$ (or $F < 1/2$) is separable.

While we found that a bipartite pure state is entangled if and only if it violates some Bell inequality, this equivalence does not hold for mixed states. You will show in Exercise 4.?? that for a Werner state with $\lambda = 1/2$ (or any smaller value of λ) there is a local hidden-variable theory that fully accounts for the correlations between measurements of Alice's qubit and Bob's. Thus, Werner states with $1/3 < \lambda < 1/2$ are inseparable states that violate no Bell inequality.

Oddly, though a Werner state with $1/3 < \lambda < 1/2$ is not Bell-inequality violating, it is nonetheless a shared resource more powerful than classical randomness. You will also show in Exercise 4.?? that by consuming a Werner state Alice and Bob can teleport a qubit in an unknown state with fidelity

$$F_{\text{teleport}} = \frac{1}{2}(1+\lambda) \ . \tag{4.127}$$

This fidelity exceeds the maximal fidelity $F_{\text{teleport}} = 2/3$ that can be achieved without any shared entanglement, for any $\lambda > 1/3$ — that is, for any inseparable Werner state, whether Bell-inequality violating or not. Even if well described by local hidden variables, an entangled mixed state can be useful.

It seems rather strange that shared entangled states described by local hidden-variable theory should be a more powerful resource than classical shared randomness. Further observations to be discussed in §5.?? will deepen our grasp of the situation. There we will find that if Alice and Bob share *many copies* of the Werner state $\boldsymbol{\rho}(\lambda)$ with $1/3 < \lambda < 1/2$, then while local hidden variables provide an adequate description of the correlations if Alice and Bob are restricted to measuring the pairs *one at a time*, violations of Bell inequalities still arise if they are permitted to perform more general kinds of measurements. These observations illustrate that mixed-state entanglement is a surprisingly subtle and elusive concept.

4.7 Nonlocality without entanglement

Quantum entanglement typifies the principle that there are bipartite quantum operations that cannot be implemented using only local operations and classical communication (LOCC). For example, if Alice and Bob share no prior entanglement, they cannot perform Bell measurement or prepare the entangled state $|\phi^+\rangle_{AB}$ unless they get together. Now we will encounter an interesting surprise: some things that Alice and Bob are unable to do with LOCC do *not* involve quantum entanglement, at least not directly.

Consider a game played by Alice, Bob, and Charlie. Charlie prepares a state $|\psi_i\rangle_{AB}$ selected from an ensemble of mutually orthogonal bipartite states, and distributes $|\psi_i\rangle_{AB}$ to Alice and Bob. To win the game, Alice and Bob must identify the state that Charlie prepared. Of course, if Alice and Bob were permitted to unite, they could perform an orthogonal measurement that would identify the state with certainty, and they would be able to win every time. But the rules of the game require Alice and Bob to stay separated, and they are forbidden to exchange quantum information — only LOCC is allowed. Thus, if Charlie's ensemble includes entangled states, Alice and Bob won't be able to win in general.

To make things easier for Alice and Bob, let's impose a new rule: Charlie is required to prepare a product state

$$|\psi\rangle_{AB} = |\alpha_i\rangle_A \otimes |\beta_i\rangle_B \ . \tag{4.128}$$

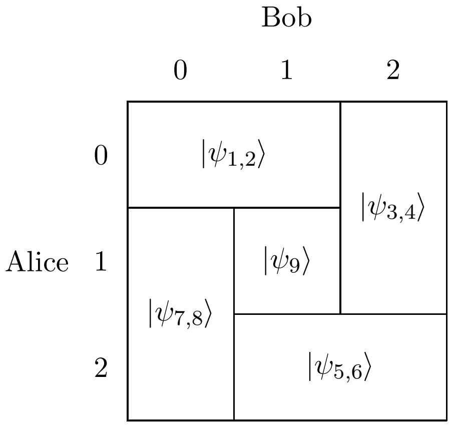
Now, since Alice has a pure state, and so does Bob, we might expect them to be able to devise a winning strategy. But on further reflection, this is not so obvious. Though the states $\{|\psi_i\rangle_{AB}\}$ in Charlie's ensemble are mutually orthogonal, the states $\{|\alpha_i\rangle_A\}$ that Alice could receive need not be mutually orthogonal, and the same is true of the states $\{|\beta_i\rangle_B\}$ that Bob could receive.

Indeed, even under the new rules, there is no winning strategy for Alice and Bob in general. Though Charlie sends a pure state to Alice and a pure state to Bob, there is no way for Alice and Bob, using LOCC, to fully decipher the message that Charlie has sent to them. This phenomenon is called *nonlocality without entanglement.*

The best way to understand nonlocality without entanglement is to consider an example. Suppose that Alice and Bob share a pair of *qutrits* (3-level quantum systems), and denote the three elements of an orthonormal basis for the qutrit by $\{|0\rangle, |1\rangle, |2\rangle\}$. In a streamlined notation, Charlie's ensemble of nine mutually orthogonal states is

$$\begin{aligned}
|\psi\rangle_{1,2} &= |0, 0 \pm 1\rangle \ , \\
|\psi\rangle_{3,4} &= |0 \pm 1, 2\rangle \ , \\
|\psi\rangle_{5,6} &= |2, 1 \pm 2\rangle \ , \\
|\psi\rangle_{7,8} &= |1 \pm 2, 0\rangle \ , \\
|\psi\rangle_{9} &= |1, 1\rangle \ .
\end{aligned} \tag{4.129}$$

(Here, $|0, 0 \pm 1\rangle$ denotes $|0\rangle_A \otimes \frac{1}{\sqrt{2}}(|0\rangle_B \pm |1\rangle_B)$, etc.) For ease of visualization, it is very convenient to represent this basis pictorially, as a tiling of a square by rectangles:



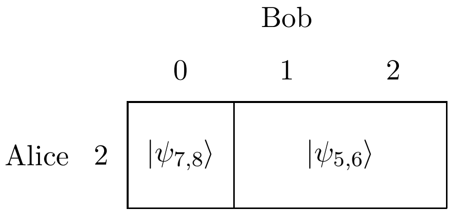
In the picture, the mutual orthogonality of the elements of Charlie's basis is reflected in the property that the rectangles are nonoverlapping.

When Charlie prepares one of these 9 states and distributes it, Alice receives one of the states

$$|0\rangle, |1\rangle, |2\rangle, |0 \pm 1\rangle, |1 \pm 2\rangle \ , \tag{4.130}$$

and similarly for Bob. These states are *not* mutually orthogonal, and so cannot be perfectly distinguished by the recipient.

For example, Alice might perform an incomplete orthogonal measurement that distinguishes the state $|2\rangle$ from its orthogonal complement. Pictorially, this measurement "cuts" the square into two nonoverlapping parts. If Charlie prepared one of $|\psi_{5,6}\rangle$, $|\psi_{7,8}\rangle$, then Alice's outcome could be $|2\rangle\langle 2|$; in that case the state prepared by her measurement can be represented as:



After learning Alice's measurement outcome, Bob can perform an orthogonal measurement that projects on the basis

$$\{|0\rangle, |1+2\rangle, |1-2\rangle\} \ . \tag{4.131}$$

If his outcome is $|1+2\rangle$ or $|1-2\rangle$, then Alice and Bob have successfully identified Charlie's state as $|\psi_5\rangle$ or $|\psi_6\rangle$. But if Bob's outcome is $|0\rangle$, then Alice and Bob remain uncertain whether Charlie prepared $|\psi_7\rangle$ of $|\psi_8\rangle$. On the other hand, if Charlie prepared one of $|\psi_{1,2}\rangle, |\psi_{3,4}\rangle, |\psi_{7,8}\rangle, |\psi_9\rangle$, then Alice's outcome could be $|0\rangle\langle 0| + |1\rangle\langle 1|$; in that case the state prepared by her measurement can be represented as:

<table>
<tr><td></td><td></td><td colspan="3">Bob</td></tr>
<tr><td></td><td></td><td>0</td><td>1</td><td>2</td></tr>
<tr><td rowspan="2">Alice</td><td>0</td><td colspan="2">$|\psi_{1,2}\rangle$</td><td rowspan="2">$|\psi_{3,4}\rangle$</td></tr>
<tr><td>1</td><td>$|\psi_{7,8}\rangle$</td><td>$|\psi_9\rangle$</td></tr>
</table>

Once again, Alice and Bob have lost any hope of distinguishing $|\psi_7\rangle$ from $|\psi_8\rangle$, but in a few more rounds of LOCC, they can successfully identify any of the other five states. Bob projects onto $|2\rangle$ or its complement; if he finds $|2\rangle\langle 2|$, then Alice projects onto $|0\pm 1\rangle$ to complete the protocol. If Bob's outcome is $|0\rangle\langle 0| + |1\rangle\langle 1|$, then Alice projects onto $\{|0\rangle, |1\rangle\}$; finally Bob measures in either the $|0\pm 1\rangle$ basis (if Alice found $|0\rangle$) or the $\{|0\rangle, |1\rangle\}$ basis (if Alice found $|1\rangle$).

By choosing one of nine mutually orthogonal product states, Charlie has sent two trits of classical information to Alice and Bob. But their LOCC protocol, which fails to distinguish $|\psi_7\rangle$ from $|\psi_8\rangle$, has not been able to recover all of the information in Charlie's message. Of course, this is just one possible protocol, but one can prove (we won't here) that no LOCC protocol can extract two trits of classical information. The trouble is that with LOCC, Alice and Bob cannot fully "dissect" the square into nonoverlapping rectangles. This is nonlocality without entanglement.

4.8 Multipartite entanglement

Up until now, we have mostly limited our attention to quantum states shared by two parties. We will conclude this chapter with some observations about the properties of entanglement shared by three or more parties: *multipartite entanglement.*

Consider for example the case of a pure state $|\psi\rangle_{A_1,A_2,\ldots A_n}$ shared by n parties $A_1, A_2, \ldots A_n$, and suppose that there is no way to divide the parties into two smaller camps, where each camp shares a pure state.

Thus the state is entangled, and furthermore, it can't be expressed as a product of states each involving fewer than n parties. Hence we might say that the state exhibits n-party entanglement. If the parties start out with an n-fold product state $|\psi_1\rangle_{A_1} \otimes |\psi_2\rangle_{A_2} \otimes \cdots |\psi_n\rangle_{A_n}$, then there is no way for them to assemble the state $|\psi\rangle_{A_1,A_2,\ldots A_n}$ using LOCC alone — quantum communication is required. Indeed, no matter how we divide the n parties into two subsystems A and B, quantum communication between A and B is needed.

What if we disallow quantum communication, but we do equip the parties with *pairwise* entanglement that has been established in advance? Then for the purpose of constructing the state $|\psi\rangle_{A_1,A_2,\ldots A_n}$, it clearly sufficies for the first party A_1 to share bipartite entanglement with each of the other $n-1$ parties. Party A_1 can build the state $|\psi\rangle_{A_1,A_2,\ldots A_n}$ in her own laboratory, and then teleport the corresponding share of the state to each of the $n-1$ other parties. In this sense, then, bipartite entanglement and LOCC is as powerful a resource as multiparty entanglement.

Nonetheless, multipartite entangled states exhibit some qualitatively new phenomena that we don't encounter in the study of bipartite entanglement, such as nonprobabilistic tests of Einstein locality, and entanglement-enhanced multiparty communication.

4.8.1 Three quantum boxes

In the wake of the wildly successful experiment with the three coins on the table, Alice and Bob are now world famous. They are both tenured professors, Alice at Caltech, and Bob at Chicago. They are much too important to spend much time in the lab, but they have many graduate students and remain scientifically active.

Their best student, Charlie, who did all the hard work on the coin experiment, has graduated and is now an assistant professor at Princeton. Alice and Bob would like to nurture Charlie's career, and help him earn tenure. One day, Alice and Bob are chatting on the phone:

Alice: You know, Bob, we really ought to help Charlie. Can you think of a neat experiment that the three of us can do together?

Bob: Well, I dunno, Alice. There are a lot of experiments I'd like to do with our entangled pairs of qubits. But in each experiment, there's one qubit for me and one for you. It looks like Charlie's the odd man out.

Alice: [*Long pause*] Bob .... Have you ever thought of doing an experiment with *three* qubits?

Bob's jaw drops and his pulse races. In a sudden epiphany, his whole future career seems mapped out before him. Truth be told, Bob was beginning to wonder if pairs of qubits were getting to be old hat. Now he knows that for the next five years, he will devote himself slavishly to performing the definitive three-qubit experiment. By that time, he, Alice, and Charlie will have trained another brilliant student, and will be ready for a crack at four qubits. Then another student, and another qubit. And so on to retirement.

Here is the sort of three-qubit experiment that Alice and Bob decide to try: Alice instructs her technician in her lab at Caltech to prepare carefully a state of three quantum boxes. (But Alice doesn't know exactly how the technician does it.) She keeps one box for herself, and she ships the other two by quantum express, one to Bob and one to Charlie. Each box has a ball inside that can be either black or white, but the box is sealed tight shut. The only way to find out what is inside is to open the box, but there are two different ways to open it — the box has two doors, clearly marked X and Y. When either door opens, a ball pops out whose color can be observed. It isn't possible to open both doors at once.

Alice, Bob, and Charlie decide to study how the boxes are correlated. They conduct many carefully controlled trials. Each time, one of the three, chosen randomly, opens door X, while the other two open door Y. Lucky as ever, Alice, Bob, and Charlie make an astonishing discovery. They find that every single time they open the boxes this way, the number of black balls they find is *always* odd.

That is, Alice, Bob and Charlie find that when they open door X on one box and door Y on the other two, the colors of the balls in the boxes are guaranteed to be one of

$$0_A 0_B 1_C\ , \quad 0_A 1_B 0_C\ , \quad 1_A 0_B 0_C\ , \quad 1_A 1_B 1_C\ , \tag{4.132}$$

(0 for white, 1 for black); They *never* see any of

$$1_A 1_B 0_C\ , \quad 1_A 0_B 1_C\ , \quad 0_A 1_B 1_C\ , \quad 0_A 0_B 0_C\ . \tag{4.133}$$

It makes no difference which of the three boxes is opened through door X.

After a while, Alice, Bob, and Charlie catch on that after opening two of the boxes, they can always predict what will happen before they open the third box. If the first two balls are the same color, the last ball is sure to be black, and if the first two are different colors, the last ball is sure to be white. They've tried it a zillion times, and it always works!

Even after all the acclaim showered upon the three-coin experiment, Alice, Bob, and Charlie have never quite shaken their attachment to Einstein locality. One day they are having a three-way conference call:

Alice: You know, guys, sometimes I just can't decide whether to open door X or door Y of my box. I know I have to choose carefully ... If I open door X, that's sure to disturb the box; so I'll never know what would have happened if I had opened door Y instead. And if I open door Y, I'll never know what I would have found if I had opened door X. It's frustrating!

Bob: Alice, you're so wrong! Our experiment shows that you can have it both ways. Don't you see? Let's say that you want to know what will happen when you open door X. Then just ask Charlie and me to open door Y of our boxes and to tell you what we find. You'll know absolutely for sure, without a doubt, what's going to happen when you open door X. We've tested that over and over again, and it always works. So why bother to open door X? You can go ahead and open door Y instead, and see what you find. That way, you really do know the result of opening *both* doors!

Charlie: But how can you be sure? If Alice opens door Y, she passes up the opportunity to open door X. She can't really ever have it both ways. After she opens door Y, we can never check whether opening door X would have given the result we expected.

Bob: Oh come on, how can it be otherwise? Look, you don't really believe that what you do to your box in Princeton and I do to mine in Chicago can exert any *influence* on what Alice finds when she opens her box in Pasadena, do you? When we open our boxes, we can't be changing anything in Alice's box; we're just finding the information we need to predict with certainty what Alice is going to find.

Charlie: Well, maybe we should do some more experiments to find out if you're right about that.

Indeed, the discovery of the three-box correlation has made Alice and Bob even more famous than before, but Charlie hasn't gotten the credit he deserves — he still doesn't have tenure. No wonder he wants to do more experiments! He continues:

Charlie: Here's something we can try. In all the experiments we've done up to now, we have always opened door Y on two boxes and door X on the other box. Maybe we should try something different.

Like, maybe we should see what happens if we open the same door on all three boxes. We could try opening three X doors.

Bob: Oh, come on! I'm tired of three boxes. We already know all about three boxes. It's time to move on, and I think Diane is ready to help out. Let's do four boxes!

Alice: No, I think Charlie's right. We can't really say that we know everything there is to know about three boxes until we've experimented with other ways of opening the doors.

Bob: Forget it. They'll never fund us! After we've put all that effort into opening two Y's and an X, now we're going to say we want to open three X's? They'll say we've done whiffnium and now we're proposing whaffnium ... We'll sound ridiculous!

Alice: Bob has a point. I think that the only way we can get funding to do this experiment is if we can make a prediction about what will happen. Then we can say that we're doing the experiment to test the prediction. Now, I've heard about some theorists named Greenberger, Horne, Zeilinger, and Mermin (GHZM). They've been thinking a lot about our three-box experiments; maybe they'll be able to suggest something.

Bob: Well, these boxes are my life, and they're just a bunch of theorists. I doubt that they'll have anything interesting or useful to say. But I suppose it doesn't really matter whether their theory makes any sense ... If we can test it, then even I will accept that we have a reason for doing another three-box experiment.

And so it happens that Alice, Bob, and Charlie make the pilgrimage to see GHZM. And despite Bob's deep skepticism, GHZM make a very interesting suggestion indeed:

GHZM: Bob says that opening a box in Princeton and a box in Chicago can't possibly have any influence on what will happen when Alice opens a box in Pasadena. Well, let's suppose that he's right. Now you guys are going to do an experiment in which you all open your X doors. No one can say what's going to happen, but we can reason this way: Let's just *assume* that if you had opened three Y doors, you would have found three white balls. Then we can use Bob's argument to see that if you open three X doors instead, you will have to find three black balls. It goes like this: if Alice opens X, Bob opens Y, and Charlie opens Y, then you know for certain that the number of black balls has to be odd. So, if we know that Bob and Charlie both would find white when they open door Y, then Alice *has* to find black when she opens door X. Similarly, if Alice and Charlie both would find white when they open Y, then Bob has to find black when he opens X, and if Alice and Bob both would find white when they open Y, then Charlie must find black when he opens X. So we see that[§]

$$Y_A Y_B Y_C = 000 \longrightarrow X_A X_B X_C = 111 \ . \tag{4.134}$$

Don't you agree?

Bob: Well, maybe that's logical enough, but what good is it? We don't know what we're going to find inside a box until we open it. You've assumed that we know $Y_A Y_B Y_C = 000$, but we never know that ahead of time.

GHZM: Sure, but wait. Yes, you're right that we can't know ahead of time what we would find if we opened door Y on each box. But there are only eight possibilities for three boxes, and we can easily list them all. And for each of those eight possibilities for $Y_A Y_B Y_C$ we can use the same reasoning as before to infer the value of $X_A X_B X_C$. We obtain a table, like this:

$$\begin{aligned}
Y_A Y_B Y_C = 000 &\longrightarrow \quad X_A X_B X_C = 111 \\
Y_A Y_B Y_C = 001 &\longrightarrow \quad X_A X_B X_C = 001 \\
Y_A Y_B Y_C = 010 &\longrightarrow \quad X_A X_B X_C = 010 \\
Y_A Y_B Y_C = 100 &\longrightarrow \quad X_A X_B X_C = 100 \\
Y_A Y_B Y_C = 011 &\longrightarrow \quad X_A X_B X_C = 100 \\
Y_A Y_B Y_C = 101 &\longrightarrow \quad X_A X_B X_C = 010 \\
Y_A Y_B Y_C = 110 &\longrightarrow \quad X_A X_B X_C = 001 \\
Y_A Y_B Y_C = 111 &\longrightarrow \quad X_A X_B X_C = 111
\end{aligned} \tag{4.135}$$

Bob: Okay, but so what?

GHZM: There's something interesting about the table, Bob! Look at the values for $X_A X_B X_C$ ... Every single entry has an *odd* number of 1's. That's our prediction: when you all open door X on your boxes, you'll always find an odd number of black balls! Could be one, or could be three, but always *odd*.

Naturally, Alice, Bob, and Charlie are delighted by this insight from GHZM. They proceed to propose the experiment, which is approved and

[§] Here 0 stands for white and 1 stands for black; Y_A is what Alice finds when she opens door Y on her box, and so on.

generously funded. Finally the long awaited day arrives when they are to carry out the experiment for the first time. And when Alice, Bob, and Charlie each open door X on their boxes, can you guess what they find? Three white balls. Whaaaa??!!

Suspecting an error, Alice and Bob and Charlie repeat the experiment, very carefully, over and over and over again. And in every trial, every single time, they find an even number of black balls when they open door X on all three boxes. Sometimes none, sometimes two, but never one and never three. What they find, every single time, is just the opposite of what GHZM had predicted would follow from the principle of Einstein locality!

Desperation once again drives Alice, Bob, and Charlie into the library, seeking enlightenment. After some study of a quantum mechanics textbook, and a thorough interrogation of Alice's lab technician, they realize that their three boxes had been prepared in a GHZM quantum state

$$|\psi\rangle_{ABC} = \frac{1}{\sqrt{2}}\left(|000\rangle_{ABC} + |111\rangle_{ABC}\right) \ , \tag{4.136}$$

a simultaneous eigenstate with eigenvalue one of the three observables

$$\boldsymbol{Z}_A \otimes \boldsymbol{Z}_B \otimes \boldsymbol{I}_C \ , \quad \boldsymbol{I}_A \otimes \boldsymbol{Z}_B \otimes \boldsymbol{Z}_C \ , \quad \boldsymbol{X}_A \otimes \boldsymbol{X}_B \otimes \boldsymbol{X}_C \ . \tag{4.137}$$

And since $\boldsymbol{Z}\boldsymbol{X} = i\boldsymbol{Y}$, they realize that this state has the properties

$$\begin{aligned}
\boldsymbol{Y}_A \otimes \boldsymbol{Y}_B \otimes \boldsymbol{X}_C &= -1 \\
\boldsymbol{X}_A \otimes \boldsymbol{Y}_B \otimes \boldsymbol{Y}_C &= -1 \\
\boldsymbol{Y}_A \otimes \boldsymbol{X}_B \otimes \boldsymbol{Y}_C &= -1 \\
\boldsymbol{X}_A \otimes \boldsymbol{X}_B \otimes \boldsymbol{X}_C &= \ \ 1 \ .
\end{aligned} \tag{4.138}$$

In opening the box through door X or door Y, Alice, Bob, and Charlie are measuring the observable $\boldsymbol{X}$ or $\boldsymbol{Y}$, where the outcome 1 signifies a white ball, and the outcome -1 a black ball. Thus if the three qubit state eq. (4.136) is prepared, eq. (4.138) says that an odd number of black balls will be found if door Y is opened on two boxes and door X on the third, while an even number of black balls will be found if door X is opened on all three boxes. This behavior, unambiguously predicted by quantum mechanics, is just what had seemed so baffling to Alice, Bob, and Charlie, and to their fellow die-hard advocates of Einstein locality.

After much further study of the quantum mechanics textbook, Alice, Bob, and Charlie gradually come to recognize the flaw in their reasoning. They learn of Bohr's principle of complementarity, of the irreconcilable incompatibility of noncommuting observables. And they recognized that to arrive at their prediction, they had *postulated* an outcome for the measurement of $\boldsymbol{YYY}$, and then proceeded to infer the consequences for a measurement of $\boldsymbol{XXX}$. By failing to heed the insistent admonitions of Niels Bohr, they had fallen prey to the most pernicious of fallacies.

As they had hoped, the experiment of the three boxes brings even further acclaim to Alice and Bob, and tenure to Charlie. Of course, the three-coin experiment had already convincingly struck down Einstein locality; even so, the three-box experiment had a different character. In the coin experiment, Alice and Bob could uncover any two of the three coins, finding any one of four possible configurations: HH, HT, TH, TT. Only by carrying out many trials could they amass a convincing statistical case for the violation of the Bell inequality. In contrast, in the three-box experiment, Alice, Bob, and Charlie had found a result inconsistent with Einstein locality in every single trial in which they opened door X on all three boxes!

4.8.2 Cat states

The GHZM state studied by Alice, Bob, and Charlie is a natural three-qubit generalization of the maximally entangled Bell pair. A Bell pair can be characterized as the simultaneous eigenstate of the two commuting operators ZZ, whose eigenvalue is the "parity bit" of the pair, and XX, whose eigenvalue is the phase bit. (Here we use a compressed notation in which the tensor product symbol $\otimes$ is suppressed — *e.g.*, XX denotes the operator that simultaneously applies X to both Alice's qubit and Bob's.) The GHZM state is the simultaneous eigenstate of ZZI, IZZ, and XXX.

An n-qubit generalization of the GHZM state can be defined, which is the simultaneous eigenstate of the n commuting operators

$$\begin{aligned} &ZZIII\dots I\ ,\\ &IZZII\dots I\ ,\\ &IIZZI\dots I\ ,\\ &\qquad\dots\\ &III\dots IZZ\ ,\\ &XX\dots XX\ . \end{aligned} \tag{4.139}$$

Each such simultaneous eigenstate has the form

$$\frac{1}{\sqrt{2}}\Big(|x\rangle \pm |\neg x\rangle\Big)\ , \tag{4.140}$$

where $\neg x$ denotes the complement of the binary string x. Since for large n this state is a coherent superposition of two "macroscopically distinguish-able" states, it is called an n-qubit *cat state*, in homage to Schrödinger's cat. The n-qubit cat state has $n-1$ parity bits, and just one phase bit.

Some noteworthy properties of cat states are:

- Each qubit is maximally entangled with the rest. That is, if we trace over the other $n-1$ qubits, the qubit's density operator is $\boldsymbol{\rho} = \frac{1}{2}\boldsymbol{I}$. For this reason, it is sometimes said that a cat state is a maximally entangled state of n qubits.

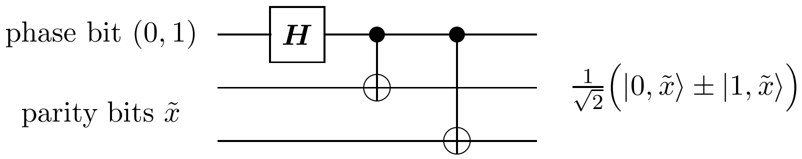
- But this is a rather misleading locution. Because its parity and phase bits are treated quite asymmetrically, the cat is not so profoundly entangled as some other multiqubit states that we will encounter in Chapter 7. For example, for the cat state with $x = 000\ldots0$, if we trace over $n-2$ qubits, the density operator of the remaining two is
$$\boldsymbol{\rho}_{2-\text{qubit}} = \frac{1}{2}\Big(|00\rangle\langle 00| + |11\rangle\langle 11|\Big) \ , \tag{4.141}$$
which has rank two rather than four. Correspondingly, we can acquire a bit of information about a cat state (one of its parity bits) by observing only two of the qubits in the state. Other multiqubit states, which might be regarded as more highly entangled than cat states, have the property that the density operator of two (or more) qubits is proportional to the identity, if we trace over the rest.

- Suppose that Charlie prepares one of the 2^n possible cat states and distributes it to n parties. Using LOCC, the parties can determine all $n-1$ parity bits of the state — each pary measures Z and all broadcast their results. But by measuring Z they destroy the phase bit. Alternatively, they can all measure X to determine the phase bit, but at the cost of destroying all the parity bits.

- Each party, by applying one of $\{\boldsymbol{I}, \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}\}$ can transform a given cat state to any one of four other cat states; that is, the party can modify the phase bit and one of the $n-1$ parity bits. All n parties, working together, can transform one cat state to any one of the 2^n mutually orthogonal cat states; for example, one party can manipulate the phase bit while each of the the other $n-1$ parties controls a parity bit.

- If the parties unite, the phase bit and all parity bits can be simultaneously measured.

If the parties start out with a product state, the three-qubit cat state (for example) can be prepared by executing the quantum circuit:


phase bit $(0,1)$
H
parity bits $\tilde{x}$
$\frac{1}{\sqrt{2}}\Big(|0,\tilde{x}\rangle \pm |1,\tilde{x}\rangle\Big)$

For the n-party case, a similar circuit with $n-1$ CNOT gates does the job. Thus, to prepare the state, it suffices for the first party to visit each of the other $n-1$ parties. By running the circuit in reverse, a cat state can be transformed to a product state that can be measured locally.

4.8.3 Entanglement-enhanced communication

An intriguing property of the n-qubit cat state is that its phase bit can be manipulated by each one of the n parties that share the state. One wonders how this shared resource might be exploited.

We will describe a setting in which possession of a cat state reduces the amount of communication that is required to accomplish a distributed information processing task. Suppose that each one of n parties labeled by index $i = 1, 2, 3, \ldots, n$ resides on a separate planet, and that party i possesses some data (a string of bits x_i) known only to that party. The goal of the parties is to compute a function f (with a one-bit output) that depends on all the data:

$$f(x_1, x_2, x_3, \ldots, x_n) \in \{0, 1\}\ . \tag{4.142}$$

In this universe, computation is cheap, and communication is expensive. Each party has unlimited computational power at her disposal, but since no party knows the full input of the function f, no one can compute f unless the parties communicate. For this purpose, they are equipped with a broadcast channel — if any party speaks, all the others can hear her. However, use of the broadcast channel is very expensive, so that the parties wish to compute f while making minimal use of the channel.

With this motivation, we define the *classical communication complexity* $CCC[f]$ of the function f:

> $CCC[f]$= the minimum of bits that must be broadcast (in the worst case) for all the parties to know the value of $f(x_1, x_2, x_3, \ldots, x_n)$.

Here "in the worst case" means that we maximize the number of bits of communication required over all possible values for the input strings $x_1, x_2, x_3, \ldots, x_n$.

We are interested in whether using quantum information can reduced the amount of communication required to compute a function. Hence we contrast the function's classical communication complexity with its *quantum communication complexity.* There are actually several different natural ways to generalize a classical communication setting to a quantum setting. In one, to which we return in Chapter 6, the parties are allowed to exchange qubits instead of classical bits. Here, we consider a scenario in which all communication is via the classical broadcast channel, but the parties are allowed to share entangled states that have been prepared in advance, and to manipulate their shared entanglement locally. Thus we define the *quantum communication complexity* $QCC[f]$ as

> $QCC[f]$= the minimum of bits that must be broadcast (in the worst case) for all the parties to know the value of $f(x_1, x_2, x_3, \ldots, x_n)$, where the parties are permitted to share prior quantum entanglement.

One way to argue that multipartite entanglement can be a useful resource is to establish that there are functions f such that

$$QCC[f] < CCC[f] \ . \tag{4.143}$$

Here is an example of such a function: Each party holds an m-bit string, and they are to compute

$$\sum_{i=1}^{n} x_i \ (\text{mod } 2^m) \ . \tag{4.144}$$

Except that they have been promised that the answer is either 0 or 2^{m-1}; therefore, their function has just a one-bit output.

First consider what strategy the parties should play if they share no entanglement. Suppose that parties 2 through n broadcast their data, and that the first party computes f and broadcasts the result. But note that it is not necessary for the parties to broadcast all of their bits, since some of the bits cannot affect the answer. Indeed, the k least significant bits are irrelevant as long as

$$(n-1)\left(2^k - 1\right) < 2^{m-1} \ , \tag{4.145}$$

which is satisfied provided that

$$(n-1)2^k \le 2^{m-1} \ . \tag{4.146}$$

It suffices then, for parties 2 through n to broadcast their $m-k$ most significant bits, where

$$m - k \ge \log_2(n-1) + 1 \ ; \tag{4.147}$$

including one more bit for the first party to broadcast the answer, we conclude that

$$CCC[f] \leq (n-1)\Big(\log_2(n-1)+1\Big)+1\ . \tag{4.148}$$

In fact, this protocol is close to optimal — it can be proved that

$$CCC[f] > n\log_2 n - n\ . \tag{4.149}$$

But the amount of communication required can be reduced if the parties share an n-qubit cat state, for they can imprint the answer on their shared phase bit! Each applies to her qubit the transformation

$$\begin{aligned} |0\rangle &\rightarrow |0\rangle\ ,\\ |1\rangle &\rightarrow e^{2\pi i(x_i/2^m)}|1\rangle\ . \end{aligned} \tag{4.150}$$

Thus the cat state

$$|\mathrm{cat}\rangle = \frac{1}{\sqrt{2}}\Big(|000\ldots0\rangle + |111\ldots1\rangle\Big) \tag{4.151}$$

is transformed to

$$|\mathrm{cat}'\rangle = \frac{1}{\sqrt{2}}\Big(|000\ldots0\rangle + \eta|111\ldots1\rangle\Big)\ , \tag{4.152}$$

where the phase η is

$$\eta = \exp\left(2\pi i\left(\sum_{i=1}^n x_i\right)/2^m\right) = (-1)^{f(x_1,x_2,\ldots,x_n)}\ . \tag{4.153}$$

Thus the fu

4.9 Manipulating entanglement

4.10 Summary

Summary 1.
Summary 2.
Summary 3.

4.11 Bibliographical notes

4.12 Exercises

4.1 Hardy's theorem

Bob (in Boston) and Claire (in Chicago) share many identically prepared copies of the two-qubit state

$$|\psi\rangle = \sqrt{(1-2x)}\ |00\rangle + \sqrt{x}\ |01\rangle + \sqrt{x}\ |10\rangle\ , \tag{4.154}$$

where x is a real number between 0 and 1/2. They conduct many trials in which each measures his/her qubit in the basis $\{|0\rangle, |1\rangle\}$, and they learn that if Bob's outcome is 1 then Claire's is always 0, and if Claire's outcome is 1 then Bob's is always 0.

Bob and Claire conduct further experiments in which Bob measures in the basis $\{|0\rangle, |1\rangle\}$ and Claire measures in the orthonormal basis $\{|\varphi\rangle, |\varphi^\perp\rangle\}$. They discover that if Bob's outcome is 0, then Claire's outcome is always φ and never $\varphi^\perp$. Similarly, if Claire measures in the basis $\{|0\rangle, |1\rangle\}$ and Bob measures in the basis $\{|\varphi\rangle, |\varphi^\perp\rangle\}$, then if Claire's outcome is 0, Bob's outcome is always φ and never $\varphi^\perp$.

a) Express the basis $\{|\varphi\rangle, |\varphi^\perp\rangle\}$ in terms of the basis $\{|0\rangle, |1\rangle\}$.

Bob and Claire now wonder what will happen if they both measure in the basis $\{|\varphi\rangle, |\varphi^\perp\rangle\}$. Their friend Albert, a firm believer in local realism, predicts that it is impossible for both to obtain the outcome $\varphi^\perp$ (a prediction known as *Hardy's theorem*). Albert argues as follows:

> When both Bob and Claire measure in the basis $\{|\varphi\rangle, |\varphi^\perp\rangle\}$, it is reasonable to consider what might have happened if one or the other had measured in the basis $\{|0\rangle, |1\rangle\}$ instead.
>
> So suppose that Bob and Claire both measure in the basis $\{|\varphi\rangle, |\varphi^\perp\rangle\}$, and that they both obtain the outcome $\varphi^\perp$. Now if Bob had measured in the basis $\{|0\rangle, |1\rangle\}$ instead, we can be certain that his outcome would have been 1, since experiment has shown that if Bob had obtained 0 then Claire could not have obtained $\varphi^\perp$. Similarly, if Claire had measured in the basis $\{|0\rangle, |1\rangle\}$, then she certainly would have obtained the outcome 1. We conclude that if Bob and Claire both measured in the basis $\{|0\rangle, |1\rangle\}$, both would have obtained the outcome 1. But this is a contradiction, for experiment has shown that it is not possible for both Bob and Claire to obtain the outcome 1 if they both measure in the basis $\{|0\rangle, |1\rangle\}$.
>
> We are therefore forced to conclude that if Bob and Claire both measure in the basis $\{|\varphi\rangle, |\varphi^\perp\rangle\}$, it is impossible for both to obtain the outcome $\varphi^\perp$.

Though impressed by Albert's reasoning, Bob and Claire decide to investigate what predictions can be inferred from quantum mechanics.

b) If Bob and Claire both measure in the basis $\{|\varphi\rangle, |\varphi^\perp\rangle\}$, what is the quantum-mechanical prediction for the probability $P(x)$ that both obtain the outcome $\varphi^\perp$?

c) Find the "maximal violation" of Hardy's theorem: show that the maximal value of $P(x)$ is $P[(3-\sqrt{5})/2] = (5\sqrt{5}-11)/2 \approx .0902$.

d) Bob and Claire conduct an experiment that confirms the prediction of quantum mechanics. What was wrong with Albert's reasoning?

4.2 Closing the detection loophole

Recall that the *CHSH inequality*

$$|\langle ab\rangle + \langle a'b\rangle + \langle ab'\rangle - \langle a'b'\rangle| \leq 2 \tag{4.155}$$

holds if the random variables $a, b, a'b'$ take values ±1 and are governed by a joint probability distribution. The maximal violation of this inequality by the quantum-mechanical predictions occurs when the left-hand-side is $2\sqrt{2}$, which is achieved if Alice and Bob share the maximally entangled state $|\phi^+\rangle$, a, a' are measurements of Alice's qubit along axes $\hat{x}$ and $\hat{z}$, and b, b' are measurements of Bob's qubit along axes $(\hat{x}+\hat{z})/\sqrt{2}$ and $(\hat{x}-\hat{z})/\sqrt{2}$.

Alice and Bob have done a beautiful experiment measuring the polarizations of entangled photon pairs, and have confirmed the CHSH inequality violation predicted by quantum mechanics. Albert is skeptical. He points out that the detectors used by Alice and Bob in their experiment are not very efficient. Usually, when Alice detects a photon, Bob does not, and when Bob detects a photon, Alice does not. Therefore, they discard the data for most of the photon pairs, and retain the results only in the case when two photons are detected in coincidence. In their analysis of the data, Alice and Bob assume that their results are based on a fair sample of the probability distribution governing the measured variables. But Albert argues that their conclusions could be evaded if *whether* a photon is detected is correlated with the *outcome* of the polarization measurement.

Alice and Bob wonder how much they will need to improve their detector efficiency to do an experiment that will impress Albert.

Alice can choose to orient her detector along any axis, and if she aligns the detector with the axis a, then ideally the detector will click when her qubit's spin is pointing up along a, but because of detector inefficiencies it sometimes fails to click even though the qubit points up. For pair number i, let $x_i \in \{0,1\}$ be a variable indicating whether Alice's detector would click when aligned with a — if there would be a click then $x_i = 1$, and if there would be no click then $x_i = 0$. Since the detector is imperfect, x_i may be 0 even though the qubit points up along a. Similarly, $x'_i \in \{0,1\}$ indicates whether Alice's detector would click if aligned with a', $y_i \in \{0,1\}$ indicates whether Bob's detector would click if aligned with b and $y'_i \in \{0,1\}$ indicates whether Bob's detector would click if aligned with b'. Under the assumption of local realism, each pair can be assigned values of x, x', y, y' that are determined by local hidden variables.

Alice and Bob are free to decide how to align their detectors in each measurement; therefore they can fairly sample the values of x, x', y, y' and infer from their measurements the values of

$$\begin{aligned}
P_{++}(ab) &= N^{-1}\sum_{i=1}^{N} x_i y_i \ , \\
P_{++}(a'b) &= N^{-1}\sum_{i=1}^{N} x'_i y_i \ , \\
P_{++}(ab') &= N^{-1}\sum_{i=1}^{N} x_i y'_i \ , \\
P_{++}(a'b') &= N^{-1}\sum_{i=1}^{N} x'_i y'_i \ ,
\end{aligned} \tag{4.156}$$

where N is the total number of pairs tested. Here *e.g* $P_{++}(ab)$ is the probability that both detectors click when Alice and Bob orient their detectors along a and b (including the effects of detector inefficiency).

a) If $x, x', y, y' \in \{0,1\}$, show that

$$xy + xy' + x'y - x'y' \leq x + y \ . \tag{4.157}$$

b) Show that

$$P_{++}(ab) + P_{++}(a'b) + P_{++}(ab') - P_{++}(a'b') \leq P_{+\cdot}(a) + P_{\cdot +}(b) \ ; \tag{4.158}$$

here $P_{+\cdot}(a)$ denotes the probability that Alice's detector clicks if oriented along a, and $P_{\cdot +}(b)$ denotes the probability that Bob's detector clicks if oriented along b.

c) Now compare with the predictions of quantum mechanics, where Alice's detector has efficiency η_A and Bob's detector has efficiency η_B. This means that Alice's detector clicks with probability $P = \eta_A P_{\text{perf}}$, where P_{perf} is the probability of a click if her detector were perfect, and similarly for Bob. Choosing the a, a', b, b' that maximally violate the CHSH inequality, show that the quantum-mechanical predictions violate eq. (4.158) only if

$$\frac{\eta_A \eta_B}{\eta_A + \eta_B} > \frac{1}{1+\sqrt{2}} \ . \qquad (4.159)$$

Thus, if $\eta_A = \eta_B$, Alice and Bob require detectors with efficiency above 82.84% to overcome Albert's objection.

4.3 Teleportation with continuous variables

One complete orthonormal basis for the Hilbert space of two particles on the real line is the (separable) position eigenstate basis $\{|q_1\rangle \otimes |q_2\rangle\}$. Another is the entangled basis $\{|Q, P\rangle\}$, where

$$|Q, P\rangle = \frac{1}{\sqrt{2\pi}} \int dq \ e^{iPq} |q\rangle \otimes |q + Q\rangle \ ; \qquad (4.160)$$

these are the simultaneous eigenstates of the relative position operator $\boldsymbol{Q} \equiv \boldsymbol{q}_2 - \boldsymbol{q}_1$ and the total momentum operator $\boldsymbol{P} \equiv \boldsymbol{p}_1 + \boldsymbol{p}_2$.

a) Verify that

$$\langle Q', P' | Q, P\rangle = \delta(Q' - Q)\delta(P' - P) \ . \qquad (4.161)$$

b) Since the states $\{|Q, P\rangle\}$ are a basis, we can expand a position eigenstate as

$$|q_1\rangle \otimes |q_2\rangle = \int dQ dP \ |Q, P\rangle\langle Q, P| \left(|q_1\rangle \otimes |q_2\rangle\right) \ . \qquad (4.162)$$

Evaluate the coefficients $\langle Q, P| \left(|q_1\rangle \otimes |q_2\rangle\right)$.

c) Alice and Bob have prepared the entangled state $|Q, P\rangle_{AB}$ of two particles A and B; Alice has kept particle A and Bob has particle B. Now Alice has received an unknown single-particle wavepacket $|\psi\rangle_C = \int dq \ |q\rangle_C \ {}_C\langle q|\psi\rangle_C$ that she intends to teleport to Bob. Design a protocol that they can execute to achieve the teleportation. What should Alice measure? What information should she send to Bob? What should Bob do when he receives this information, so that particle B will be prepared in the state $|\psi\rangle_B$?

4.4 Teleportation with mixed states.

An operational way to define entanglement is that an entangled state can be used to teleport an unknown quantum state with better fidelity than could be achieved with local operations and classical communication only. In this exercise, you will show that there are mixed states that are entangled in this sense, yet do not violate any Bell inequality. Hence, for mixed states (in contrast to pure states) "entangled" and "Bell-inequality-violating" are not equivalent.

Consider a "noisy" entangled pair with density matrix.

$$\boldsymbol{\rho}(\lambda) = (1-\lambda)|\psi^-\rangle\langle\psi^-| + \lambda\frac{1}{4}\mathbf{1}. \tag{4.163}$$

a) Find the fidelity F that can be attained if the state $\boldsymbol{\rho}(\lambda)$ is used to teleport a qubit from Alice to Bob. [Hint: Recall that you showed in an earlier exercise that a "random guess" has fidelity $F = \frac{1}{2}$.]

b) For what values of λ is the fidelity found in (a) better than what can be achieved if Alice measures her qubit and sends a classical message to Bob? [Hint: Earlier, you showed that $F = 2/3$ can be achieved if Alice measures her qubit. In fact this is the best possible F attainable with classical communication.]

c) Compute

$$\text{Prob}(\uparrow_{\hat{n}}\uparrow_{\hat{m}}) \equiv \text{tr}\left(\boldsymbol{E}_A(\hat{n})\boldsymbol{E}_B(\hat{m})\boldsymbol{\rho}(\lambda)\right), \tag{4.164}$$

where $\boldsymbol{E}_A(\hat{n})$ is the projection of Alice's qubit onto $|\uparrow_{\hat{n}}\rangle$ and $\boldsymbol{E}_B(\hat{m})$ is the projection of Bob's qubit onto $|\uparrow_{\hat{m}}\rangle$.

d) Consider the case $\lambda = 1/2$. Show that in this case the state $\boldsymbol{\rho}(\lambda)$ violates no Bell inequalities. Hint: It suffices to construct a local hidden-variable model that correctly reproduces the spin correlations found in (c), for $\lambda = 1/2$. Suppose that the hidden variable $\hat{\alpha}$ is uniformly distributed on the unit sphere, and that there are functions f_A and f_B such that

$$\text{Prob}_A(\uparrow_{\hat{n}}) = f_A(\hat{\alpha}\cdot\hat{n}), \quad \text{Prob}_B(\uparrow_{\hat{m}}) = f_B(\hat{\alpha}\cdot\hat{m}). \tag{4.165}$$

The problem is to find f_A and f_B (where $0 \le f_{A,B} \le 1$) with the properties

$$\int_{\hat{\alpha}} f_A(\hat{\alpha}\cdot\hat{n}) = 1/2, \quad \int_{\hat{\alpha}} f_B(\hat{\alpha}\cdot\hat{m}) = 1/2,$$
$$\int_{\hat{\alpha}} f_A(\hat{\alpha}\cdot\hat{n}) f_B(\hat{\alpha}\cdot\hat{m}) = \text{Prob}(\uparrow_{\hat{n}}\uparrow_{\hat{m}}). \tag{4.166}$$

4.5 Quantum key distribution

Alice and Bob want to execute a quantum key distribution protocol. Alice is equipped to prepare either one of the two states $|u\rangle$ or $|v\rangle$. These two states, in a suitable basis, can be expressed as

$$|u\rangle = \begin{pmatrix}\cos\alpha \\ \sin\alpha\end{pmatrix} \ , \quad |v\rangle = \begin{pmatrix}\sin\alpha \\ \cos\alpha\end{pmatrix} \ , \tag{4.167}$$

where $0 < \alpha < \pi/4$. Alice decides at random to send either $|u\rangle$ or $|v\rangle$ to Bob, and Bob is to make a measurement to determine what she sent. Since the two states are not orthogonal, Bob cannot distinguish the states perfectly.

a) Bob realizes that he can't expect to be able to identify Alice's qubit every time, so he settles for a procedure that is successful only some of the time. He performs a POVM with three possible outcomes: $\neg u$, $\neg v$, or DON'T KNOW. If he obtains the result $\neg u$, he is certain that $|v\rangle$ was sent, and if he obtains $\neg v$, he is certain that $|u\rangle$ was sent. If the result is DON'T KNOW, then his measurement is inconclusive. This POVM is defined by the operators

$$\begin{aligned} \boldsymbol{f}_{\neg u} = A(\boldsymbol{I} - |u\rangle\langle u|) \ , \quad \boldsymbol{f}_{\neg v} = A(\boldsymbol{I} - |v\rangle\langle v|) \ , \\ \boldsymbol{f}_{\mathrm{DK}} = (1-2A)\boldsymbol{I} + A\left(|u\rangle\langle u| + |v\rangle\langle v|\right) \ , \end{aligned} \tag{4.168}$$

where A is a positive real number. How should Bob choose A to minimize the probability of the outcome DK, and what is this minimal DK probability (assuming that Alice chooses from $\{|u\rangle, |v\rangle\}$ equiprobably)? [**Hint:** If A is too large, $\boldsymbol{f}_{\mathrm{DK}}$ will have negative eigenvalues, and eq.(4.168) will not be a POVM.]

b) Design a quantum key distribution protocol using Alice's source and Bob's POVM.

c) Of course, Eve also wants to know what Alice is sending to Bob. Hoping that Alice and Bob won't notice, she intercepts each qubit that Alice sends, by performing an orthogonal measurement that projects onto the basis $\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}$. If she obtains the outcome $\begin{pmatrix}1\\0\end{pmatrix}$, she sends the state $|u\rangle$ on to Bob, and if she obtains the outcome $\begin{pmatrix}0\\1\end{pmatrix}$, she sends $|v\rangle$ on to Bob. Therefore each time Bob's POVM has a conclusive outcome, Eve knows with certainty what that outcome is. But Eve's tampering causes detectable errors; sometimes Bob obtains a "conclusive" outcome that actually differs from what Alice sent. What is the probability of such an error?

4.6 Minimal disturbance

In Exercise 2.1, we studied a game in which Alice decides at random (equiprobably) whether to prepare one of two possible pure states of a single qubit, either

$$|\psi\rangle = \begin{pmatrix}\cos\alpha \\ \sin\alpha\end{pmatrix} , \qquad \text{or} \qquad |\tilde{\psi}\rangle = \begin{pmatrix}\sin\alpha \\ \cos\alpha\end{pmatrix} , \tag{4.169}$$

and sends the state to Bob. By performing an orthogonal measurement in the basis $\{|0\rangle, |1\rangle\}$, Bob can identify the state with minimal error probability

$$(p_{\text{error}})_{\text{optimal}} = \sin^2\alpha = \frac{1}{2}(1 - \sin\theta) , \tag{4.170}$$

where we have defined θ by

$$\langle\psi|\tilde{\psi}\rangle \equiv \cos\theta = \sin(2\alpha) . \tag{4.171}$$

But now let's suppose that Eve wants to *eavesdrop* on the state as it travels from Alice to Bob. Like Bob, she wishes to extract optimal information that distinguishes $|\psi\rangle$ from $|\tilde{\psi}\rangle$, and she also wants to minimize the disturbance introduced by her eavesdropping, so that Alice and Bob are not likely to notice that anything is amiss.

Eve realizes that the optimal POVM can be achieved by measurement operators

$$\boldsymbol{M}_0 = |\phi_0\rangle\langle 0| , \qquad \boldsymbol{M}_1 = |\phi_1\rangle\langle 1| , \tag{4.172}$$

where the vectors $|\phi_0\rangle$, and $|\phi_1\rangle$ are arbitrary. If Eve performs this measurement, then Bob receives the state

$$\rho' = \cos^2\alpha|\phi_0\rangle\langle\phi_0| + \sin^2\alpha|\phi_1\rangle\langle\phi_1| , \tag{4.173}$$

if Alice sent $|\psi\rangle$, and the state

$$\tilde{\rho}' = \sin^2\alpha|\phi_0\rangle\langle\phi_0| + \cos^2\alpha|\phi_1\rangle\langle\phi_1| , \tag{4.174}$$

if Alice sent $|\tilde{\psi}\rangle$.

Eve wants the average fidelity of the state received by Bob to be as large as possible. The quantity that she wants to minimize, which we will call the "disturbance" D, measures how close this average fidelity is to one:

$$D = 1 - \frac{1}{2}(F + \tilde{F}) \ , \tag{4.175}$$

where

$$F = \langle\psi|\rho'|\psi\rangle \ , \qquad \tilde{F} = \langle\tilde{\psi}|\tilde{\rho}'|\tilde{\psi}\rangle \ . \tag{4.176}$$

The purpose of this exercise is to examine how effectively Eve can reduce the disturbance by choosing her measurement operators properly.

a) Show that $F + \tilde{F}$ can be expressed as

$$F + \tilde{F} = \langle\phi_0|A|\phi_0\rangle + \langle\phi_1|B|\phi_1\rangle \ , \tag{4.177}$$

where

$$\begin{aligned} A &= \begin{pmatrix} 1 - 2\cos^2\alpha\sin^2\alpha & \cos\alpha\sin\alpha \\ \cos\alpha\sin\alpha & 2\cos^2\alpha\sin^2\alpha \end{pmatrix} \ , \\ B &= \begin{pmatrix} 2\cos^2\alpha\sin^2\alpha & \cos\alpha\sin\alpha \\ \cos\alpha\sin\alpha & 1 - 2\cos^2\alpha\sin^2\alpha \end{pmatrix} \ . \end{aligned} \tag{4.178}$$

b) Show that if $|\phi_0\rangle$ and $|\phi_1\rangle$ are chosen optimally, the minimal disturbance that can be attained is

$$D_{\text{min}}(\cos^2\theta) = \frac{1}{2}(1 - \sqrt{1 - \cos^2\theta + \cos^4\theta}) \ . \tag{4.179}$$

[**Hint**: We can choose $|\phi_0\rangle$ and $|\phi_1\rangle$ to maximize the two terms in eq. (4.177) independently. The maximal value is the maximal eigenvalue of A, which since the eigenvalues sum to 1, can be expressed as $\lambda_{\text{max}} = \frac{1}{2}\left(1 + \sqrt{1 - 4 \cdot \det A}\right)$.] Of course, Eve could reduce the disturbance further were she willing to settle for a less than optimal probability of guessing Alice's state correctly.

c) Sketch a plot of the function $D_{\text{min}}(\cos^2\theta)$. Interpret its value for $\cos\theta = 1$ and $\cos\theta = 0$. For what value of θ is D_{min} largest? Find D_{min} and $(p_{\text{error}})_{\text{optimal}}$ for this value of θ.

4.7 Approximate cloning

The no-cloning theorem shows that we can't build a unitary machine that will make a perfect copy of an unknown quantum state. But suppose we are willing to settle for an *imperfect* copy — what fidelity might we achieve?

Consider a machine that acts on three qubit states according to

$$\begin{aligned}|000\rangle_{ABC} &\rightarrow \sqrt{\frac{2}{3}}|00\rangle_{AB}|0\rangle_C + \sqrt{\frac{1}{3}}|\psi^+\rangle_{AB}|1\rangle_C \\ |100\rangle_{ABC} &\rightarrow \sqrt{\frac{2}{3}}|11\rangle_{AB}|1\rangle_C + \sqrt{\frac{1}{3}}|\psi^+\rangle_{AB}|0\rangle_C \ . \end{aligned} \tag{4.180}$$

a) Is such a device physically realizable, in principle?

If the machine operates on the initial state $|\psi\rangle_A|00\rangle_{BC}$, it produces an pure entangled state $|\Psi\rangle_{ABC}$ of the three qubits. But if we observe qubit A alone, its final state is the density operator $\boldsymbol{\rho}'_A = \mathrm{tr}_{BC}\left(|\Psi\rangle_{ABC}\ {}_{ABC}\langle\Psi|\right)$. Similarly, the qubit B, observed in isolation, has the final state $\boldsymbol{\rho}'_B$. It is easy to see that $\boldsymbol{\rho}'_A = \boldsymbol{\rho}'_B$ — these are the identical, but imperfect, copies of the input pure state $|\psi\rangle_A$.

b) The mapping from the initial state $|\psi\rangle_A\ {}_A\langle\psi|$ to the final state $\boldsymbol{\rho}'_A$ of qubit A defines a superoperator \$. Find an operator-sum representation of \$.

c) For $|\psi\rangle_A = a|0\rangle_A + b|1\rangle_A$, find $\boldsymbol{\rho}'_A$, and compute its fidelity $F \equiv {}_A\langle\psi|\boldsymbol{\rho}'_A|\psi\rangle_A$.

4.8 We're so sorry, Uncle Albert

Consider the n-qubit "cat" state

$$|\psi\rangle_n = \sqrt{\frac{1}{2}}\left(|000\ldots0\rangle + |111\ldots1\rangle\right) \ . \tag{4.181}$$

This state can be characterized as the simultaneous eigenstate (with eigenvalue 1) of the n operators

$$\begin{gathered}\boldsymbol{\sigma}_3 \otimes \boldsymbol{\sigma}_3 \otimes I \otimes I \otimes \cdots \otimes I \otimes I \otimes I \\ I \otimes \boldsymbol{\sigma}_3 \otimes \boldsymbol{\sigma}_3 \otimes I \otimes \cdots \otimes I \otimes I \otimes I \\ \ldots \\ I \otimes I \otimes I \otimes I \otimes \cdots \otimes I \otimes \boldsymbol{\sigma}_3 \otimes \boldsymbol{\sigma}_3 \\ \boldsymbol{\sigma}_1 \otimes \boldsymbol{\sigma}_1 \otimes \boldsymbol{\sigma}_1 \otimes \cdots \otimes \boldsymbol{\sigma}_1 \otimes \boldsymbol{\sigma}_1 \otimes \boldsymbol{\sigma}_1 \end{gathered} \tag{4.182}$$

a) Show that $|\psi\rangle_n$ is an eigenstate of the operator

$$(\boldsymbol{\sigma}_1 + i\boldsymbol{\sigma}_2)^{\otimes n} + (\boldsymbol{\sigma}_1 - i\boldsymbol{\sigma}_2)^{\otimes n} \ , \tag{4.183}$$

and compute its eigenvalue.

b) If we believe in local hidden variables, then we believe that, for each of the n qubits, both $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_2$ have definite values once the hidden variables are specified. If so, then what can we say about the *modulus* of $(\boldsymbol{\sigma}_1 + i\boldsymbol{\sigma}_2)^{\otimes n}$ or $(\boldsymbol{\sigma}_1 - i\boldsymbol{\sigma}_2)^{\otimes n}$, assuming definite values for the hidden variables?

c) From (*b*), derive an upper bound on

$$\frac{1}{2}\left|(\boldsymbol{\sigma}_1 + i\boldsymbol{\sigma}_2)^{\otimes n} + (\boldsymbol{\sigma}_1 - i\boldsymbol{\sigma}_2)^{\otimes n}\right| \tag{4.184}$$

that follows from the local hidden-variable hypothesis.

d) Compare with (*a*). What would Einstein say?

4.9 Entanglement manipulation

a) Twenty-five players on the New York Yankees, and twenty-five players on the San Diego Padres, want to share a 50-qubit cat state. The Yankees prepare a 26-qubit cat state, and give one of the qubits to Alice; so do the Padres. Now Alice is to sew the states together and prepare the 50-qubit state. What should she do? [**Hint**: Think about stabilizers.]

b) After joining the Yankees, Alice assumed custody of one of the qubits in their 25-qubit cat state. But today she has been traded! Alice is ordered to pull her qubit out of the cat state, leaving an undamaged 24-qubit cat state for the other players. What should she do? [**Hint**: Think about stabilizers.]

4.10 Peres-Horodecki criterion in d dimensions

Recall that a *Werner state* of a pair of qubits can be expressed as

$$\boldsymbol{\rho}(\lambda) = \lambda|\phi^+\rangle\langle\phi^+| + \frac{1}{4}(1-\lambda)\boldsymbol{I} \ , \tag{4.185}$$

and that the *partial transpose* $\boldsymbol{\rho}_{AB}^{PT}$ of a bipartite density operator $\boldsymbol{\rho}_{AB}$ is defined as

$$\boldsymbol{\rho}_{AB}^{PT} \equiv (I_A \otimes T_B)(\boldsymbol{\rho}_{AB}) \tag{4.186}$$

where T is the transpose operation that acts in the computational basis $\{|i\rangle\}$ as

$$T\left(|i\rangle\langle j|\right) = |j\rangle\langle i| \ . \tag{4.187}$$

We saw in class that the partial transpose of the Werner state $\boldsymbol{\rho}(\lambda)$ is negative for $\lambda > 1/3$; therefore, by the *Peres-Horodecki criterion*, the Werner state is inseparable for $\lambda > 1/3$.

a) One natural way to generalize the Werner state to a pair of d-dimensional systems is to consider

$$\boldsymbol{\rho}_\Phi(\lambda) = \lambda|\Phi\rangle\langle\Phi| + \frac{1}{d^2}(1-\lambda)\boldsymbol{I} \ , \tag{4.188}$$

where $|\Phi\rangle$ is the maximally entangled state

$$|\Phi\rangle = \frac{1}{\sqrt{d}}\sum_{i=1}^{d}|i\rangle \otimes |i\rangle \ . \tag{4.189}$$

Show that

$$(|\Phi\rangle\langle\Phi|)^{PT} = \frac{1}{d}\left(\boldsymbol{I} - 2\boldsymbol{E}_{\text{antisym}}\right) \ , \tag{4.190}$$

where $\boldsymbol{E}_{\text{antisym}}$ is the projector onto the space that is antisymmetric under interchange of the two systems A and B.

b) For what values of λ does the state $\boldsymbol{\rho}_\Phi(\lambda)$ have a negative partial transpose?

c) If the Werner state for two qubits is chosen to be

$$\boldsymbol{\rho}(\lambda) = \lambda|\psi^-\rangle\langle\psi^-| + \frac{1}{4}(1-\lambda)\boldsymbol{I} \ , \tag{4.191}$$

then another natural way to generalize the Werner state to a pair of d-dimensional systems is to consider

$$\boldsymbol{\rho}_{\text{anti}}(\lambda) = \lambda\left(\frac{1}{\frac{1}{2}d(d-1)}\right)\boldsymbol{E}_{\text{antisym}} + \frac{1}{d^2}(1-\lambda)\boldsymbol{I} \ . \tag{4.192}$$

For what values of λ does $\boldsymbol{\rho}_{\text{anti}}(\lambda)$ have a negative partial transpose?