QI Lecture 10

Quantum Information II

Why Von Neumann Entropy?

There are several reasons for talking about Von Neumann entropy. It's a useful measure for entanglement as well as following our intuition about how information should behave.

Entropy is a good measure of entanglement.

On the composite space $A \otimes B$ the Bell states are pure states. With access to both qubits, a measurement in the Bell basis produces a single reliable result (which is what pure states do). But in either subspace, A or B, the reduced density matrices of the Bell states are mixed states and therefore always random.

This difference, between being a pure state in $A \otimes B$ and a mixed state in just A or B, provides us with a way to measure entanglement in general. For a pure state with density matrix $\rho \in A \otimes B$, the entanglement between systems A and B is

$$\boxed{Entanglement = S[\rho_a] = S[\rho_b]}$$

Just to literally draw a line under it, this way of quantifying entanglement assumes that the bipartite state, $\rho = \rho_{ab}$, is a pure state.

Example Find the entropy of $|\Psi_+\rangle$ in $A \otimes B$ and in the subspace A.

The entropy[1] of $|\Psi_+\rangle = \frac{|01\rangle+|10\rangle}{\sqrt{2}}$ is $S[|\Psi_+\rangle\langle\Psi_+|]$. Because $|\Psi_+\rangle\langle\Psi_+|$ is a pure state, this is like sending the signal "AAAAAAA…": zero entropy. But just to be sure, we'll express the density matrix as an actual matrix and find its eigenvalues. In the computational basis, $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$,

[1] Generally speaking, whether we're talking about Shannon entropy or Von Neumann entropy should be obvious from context. Here we're talking about Von Neumann entropy.

$$\rho = |\Psi_+\rangle\langle\Psi_+| \sim \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \begin{pmatrix} 0, & \frac{1}{\sqrt{2}}, & \frac{1}{\sqrt{2}}, & 0 \end{pmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The eigenvalues of this matrix are solutions of

$$\begin{aligned} C(\lambda) &= \begin{vmatrix} -\lambda & 0 & 0 & 0 \\ 0 & \frac{1}{2}-\lambda & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2}-\lambda & 0 \\ 0 & 0 & 0 & -\lambda \end{vmatrix} \\ &= -\lambda \begin{vmatrix} \frac{1}{2}-\lambda & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2}-\lambda & 0 \\ 0 & 0 & -\lambda \end{vmatrix} \\ &= \lambda^2 \begin{vmatrix} \frac{1}{2}-\lambda & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2}-\lambda \end{vmatrix} \\ &= \lambda^2 \left(\left(\tfrac{1}{2}-\lambda\right)^2 - \tfrac{1}{4} \right) \\ &= \lambda^2 \left(\lambda^2 - \lambda + \tfrac{1}{4} - \tfrac{1}{4} \right) \\ &= \lambda^2 \left(\lambda^2 - \lambda \right) \\ &= \lambda^3 (\lambda - 1) \end{aligned}$$

The eigenvalues are $\lambda = 0, 0, 0, 1$ and therefore the entropy[2] is

$$S[|\Psi_+\rangle\langle\Psi_+|] = -0\log_2(0) - 0\log_2(0) - 0\log_2(0) - 1\log_2(1) = 0$$

Now to find the entropy of the reduced density matrix, $\rho_a = Tr_b\left[|\Psi_+\rangle\langle\Psi_+|\right]$. Once again we'll use the trick that "the trace turns outer products into inner products".

[2]Remember that we use $0\log(0) \equiv 0$.

$$
\begin{aligned}
\rho_a &= Tr_b\left[|\Psi_+\rangle\langle\Psi_+|\right] \\
&= Tr_b\left[\left(\frac{|0\rangle_a|1\rangle_b+|1\rangle_a|0\rangle_b}{\sqrt{2}}\right)\left(\frac{\langle 0|_a\langle 1|_b+\langle 1|_a\langle 0|_b}{\sqrt{2}}\right)\right] \\
&= \left(\frac{|0\rangle_a\langle 1|_b+|1\rangle_a\langle 0|_b}{\sqrt{2}}\right)\left(\frac{\langle 0|_a|1\rangle_b+\langle 1|_a|0\rangle_b}{\sqrt{2}}\right) \\
&= \tfrac{1}{2}|0\rangle_a\langle 0|_a\langle 1|1\rangle_b + \tfrac{1}{2}|1\rangle_a\langle 0|_a\langle 0|1\rangle_b + \tfrac{1}{2}|0\rangle_a\langle 1|_a\langle 1|0\rangle_b + \tfrac{1}{2}|1\rangle_a\langle 1|_a\langle 0|0\rangle_b \\
&= \tfrac{1}{2}|0\rangle_a\langle 0|_a + \tfrac{1}{2}|1\rangle_a\langle 1|_a
\end{aligned}
$$

Finally, to find $S[\rho_a]$ we first need to find the eigenvalues. Since ρ_a is already written in its spectral decomposition here, we know that the eigenvalues are $\lambda = \frac{1}{2}, \frac{1}{2}$. But just to make sure, we'll turn it into a matrix and find the eigenvalues the old way. In the computational basis, $\rho_a = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$ and so

$$C(\lambda) = \begin{vmatrix} \frac{1}{2} - \lambda & 0 \\ 0 & \frac{1}{2} - \lambda \end{vmatrix} = \left(\lambda - \frac{1}{2}\right)^2$$

So the entropy of the reduced density matrix is

$$S[\rho_a] = -\frac{1}{2}\log_2\left(\frac{1}{2}\right) - \frac{1}{2}\log_2\left(\frac{1}{2}\right) = \frac{1}{2} + \frac{1}{2} = 1$$

■

The big take-home here is that pure states, even when they're on composite spaces, have zero entropy, but if they're entangled, then the entropy in the subspaces is non-zero because the results of measurements will always have some fundamental randomness. This is to be expected whenever we measure a state in a basis that doesn't include that state (e.g., if you measure the state $|+\rangle$ in the $\{|0\rangle, |1\rangle\}$ basis, you're going to see unpredictable results). Since states that entangle A and B straddle both spaces, no measurement in just one space will ever be in the "correct" basis.

Notice that we just declared that for $\rho = |\psi\rangle_{ab}\langle\psi|_{ab} \in A \otimes B$, $S[\rho_a] = S[\rho_b]$. This isn't obvious and is certainly isn't true for mixed states. So let's make it a theorem and prove it.

Theorem (Equal entropies for reduced pure states)**.** *If $\rho = |\psi\rangle_{ab}\langle\psi|_{ab} \in A \otimes B$ is a pure state, then the entropies of the reduced density matrices, $\rho_a = Tr_b[\rho]$ and $\rho_b = Tr_a[\rho]$, are equal: $S[\rho_a] = S[\rho_b]$.*

Given any bases $\{|u_j\rangle\} \in A$ and $\{|r_k\rangle\} \in B$ we can write

$$|\psi\rangle_{ab} = \sum_{jk} c_{jk}|u_j\rangle_a|r_k\rangle_b = \sum_j |u_j\rangle_a|\overline{r_j}\rangle_b$$

where $|\overline{r_j}\rangle \equiv \sum_k c_{jk}|r_k\rangle$. Generally speaking, $\{|\overline{r_k}\rangle\}$ is not an orthonormal set. But there is a cute trick we can use. Since ρ_a is hermitian (like all density matrices) it has a spectral decomposition

$$\rho_a = \sum_j \lambda_j |u_j\rangle_a \langle u_j|_a$$

Using this basis from the spectral decomposition of ρ_a when we write $|\psi\rangle_{ab}$ actually forces $\{|\overline{r_k}\rangle\}$ to be orthogonal. We can see that by taking the partial trace of ρ over B.[3]

$$\sum_j \lambda_j |u_j\rangle_a \langle u_j|_a = \rho_a = Tr_b\left[\sum_{jk} |u_j\rangle_a |\overline{r_j}\rangle_b \langle u_k|_a \langle \overline{r_k}|_b\right] = \sum_{jk} |u_j\rangle_a \langle u_k|_a \langle \overline{r_k}|\overline{r_j}\rangle_b \qquad \Rightarrow \qquad \langle \overline{r_k}|\overline{r_j}\rangle = \delta_{jk}\lambda_j$$

Finally, if we define $|w_j\rangle \equiv \frac{1}{\sqrt{\lambda_j}}|\overline{r_j}\rangle$ (meaning that $\langle w_k|w_j\rangle = \delta_{jk}$), then we can write the original pure state as

$$|\psi\rangle_{ab} = \sum_j \sqrt{\lambda_j}|u_j\rangle_a |w_j\rangle_b$$

This form for $|\psi\rangle_{ab}$ is the whole point of all that manipulation above. We needed to carefully chose bases for A and B that would allow this form to exist. The density matrix is

$$\rho = \sum_{jk} \sqrt{\lambda_j \lambda_k}|u_j\rangle_a |w_j\rangle_b \langle u_k|_a \langle w_k|_b$$

we already know the density matrix for ρ_a and

$$\begin{aligned}
\rho_b &= Tr_a\left[\textstyle\sum_{jk} \sqrt{\lambda_j \lambda_k}|u_j\rangle_a |w_j\rangle_b \langle u_k|_a \langle w_k|_b\right] \\
&= \textstyle\sum_{jk} \sqrt{\lambda_j \lambda_k}\langle u_k|u_j\rangle_a |w_j\rangle_b \langle w_k|_b \\
&= \textstyle\sum_{jk} \delta_{jk}\sqrt{\lambda_j \lambda_k}|w_j\rangle_b \langle w_k|_b \\
&= \textstyle\sum_k \lambda_k |w_k\rangle_b \langle w_k|_b
\end{aligned}$$

We have just shown that, for a pure state, the reduced density matrices have the same spectrum of eigenvalues, $\{\lambda_k\}$! This means that in particular

$$S[\rho_a] = S[\rho_b] = -\sum_k \lambda_k \log(\lambda_k)$$

QED

[3] Recall that the easiest way to take the partial trace is to "turn outer products into inner products".

The Entanglement of Independent Systems is Additive

If you have two completely unrelated density matrices, ρ_a and ρ_b, on two different spaces, A and B, then the density matrix for the composite system is $\rho_a \otimes \rho_b$. The total information in two random books is the sum of the information in each individually. Happily, the same intuition applies here:

$$\boxed{S\left[\rho_a \otimes \rho_b\right] = S\left[\rho_a\right] + S\left[\rho_b\right]}$$

Defining

$$\rho_a = \sum_j \lambda_j |j\rangle_a \langle j|_a \qquad \rho_b = \sum_j \sigma_k |k\rangle_b \langle k|_b$$

we can quickly calculate

$$\begin{aligned}
S\left[\rho_a \otimes \rho_b\right] &= S\left[\left(\sum_j \lambda_j |j\rangle_a \langle j|_a\right)\left(\sum_k \sigma_k |k\rangle_b \langle k|_b\right)\right] \\
&= S\left[\sum_{jk} \lambda_j \sigma_k |j\rangle_a |k\rangle_b \langle j|_a \langle k|_b\right] \\
&= -\sum_{jk} \lambda_j \sigma_k \log\left(\lambda_j \sigma_k\right) \\
&= -\sum_{jk} \lambda_j \sigma_k \log\left(\lambda_j\right) - \sum_{jk} \lambda_j \sigma_k \log\left(\sigma_k\right) \\
&= -\left(\sum_k \sigma_k\right)\left(\sum_j \lambda_j \log\left(\lambda_j\right)\right) - \left(\sum_j \lambda_j\right)\left(\sum_k \sigma_k \log\left(\sigma_k\right)\right) \\
&= -\sum_j \lambda_j \log\left(\lambda_j\right) - \sum_k \sigma_k \log\left(\sigma_k\right) \\
&= S[\rho_a] + S[\rho_b]
\end{aligned}$$

QED

Entanglement as a Resource

When we describe something as a resource, it's because we can talk about how much we need for a job, how much we have, and the amount is generally conserved. We saw this in the context of classical information and compression; Shannon entropy tells us how much a string of characters can be compressed while retaining the ability to recreate the original string (the two strings have the same information).

We can either "**distill**" weakly entangled systems in fewer strongly entangled systems or "**dilute**" strongly entangled systems into a larger number of weakly entangled systems. Like Shannon describing the limits of data compression, Von Neumann entropy tells us that if we have k systems with entanglement E, then we can distill or dilute them into n systems with entanglement E' such that

$$\boxed{kE = nE'}$$

in the limit as we use more and more states. So this doesn't hold true for a *single* entangled state, but it does hold *on average* for many, many entangled states considered together.

In other words, the total Von Neumann entropy is conserved on average. In the same way that data compression can be used to convey the same amount of information (as quantified by Shannon entropy) using fewer bits, entanglement distillation can be used to convey the same total entanglement (as quantified by Von Neumann entropy) using fewer qubits.

For the same reason that we skipped the details in the classical information theory lecture (there's just too much to go over), we'll skip the details here, but we will have a few examples of very rough distillation.

Local Operations and Classical Communication (LOCC)

When dealing with composite systems we often find ourselves in a situation where two of the subspaces cannot directly interact (often because of physical distance). For this very common situation we introduce the LOCC rules. "Classical communication" doesn't need a lot of explanation.[4]

If Alice controls system A and Bob controls system B, then a "local operation" is one that Alice and Bob only perform on their own space. These include:

- i) Append an ancillary system, that's not (initially) entangled with either A or B.
- ii) Local unitary operators.
- iii) Local projective measurements.
- iv) Throw away part of the system.

One of the most important, simple rules to keep in mind when considering isolated systems is

nothing that's done to one subsystem ever has any direct impact on the density matrix of the other.

After all, if Alice could do something to A and Bob could detect it by measuring statistics on B, then instantaneous (faster than light) communication would be possible.[5]

[4] If you can imagine hearing or reading an explanation about what classical communication is, then you already know what it is.

[5] It isn't.

Teleportation gets around this by including classical communication (which is necessarily at least as slow as light).

i) Append an Ancillary System Even though Alice and Bob already share a state, $|\psi\rangle_{ab} \in A \otimes B$, there's a lot more to the universe. Maybe Bob has another quantum system, perhaps some more qubits, in the state $|\eta\rangle_c \in C$. The state of the shared state along with Bob's extra state is

$$|\psi\rangle_{ab}|\eta\rangle_c \in A \otimes B \otimes C$$

Technically speaking, there's already an infinite number of these ancillary systems floating around and, technically speaking, they're already part of the composite state $(|\psi\rangle_{ab}|\eta\rangle_c|\zeta\rangle_d|\chi\rangle_e \ldots)$. Typically you only bother to mention an ancillary system when Bob is about to apply some joint operation on his half of the entangled state (B) and the ancillary system (C)

$$(I_a \otimes U_{bc})\,|\psi\rangle_{ab}|\eta\rangle_c$$

ii) Local Unitary Operators. Everything that you can do to a quantum system is described by unitary operations, U. However, if Alice and Bob's systems are isolated from each other (often by being in different places), then one cannot affect the other's system. "Alice does nothing to Bob's system" means applying I_b to Bob's system. So if Alice and Bob share $|\psi\rangle_{ab}$, then the only kinds of operations that Alice can perform are

$$(U_a \otimes I_b)\,|\psi\rangle_{ab}$$

iii) Local Projective Measurements. Whether or not Alice's system is entangled with Bob's, there's nothing to stop her from making any measurement she might normally make, $\{P_a^{(m)}\}_m$, on her system. Once again, this entails doing nothing to Bob's system so the measurement operation $P_a^{(m)}$ in space A becomes $P_a^{(m)} \otimes I_b$ in $A \otimes B$. Other than that, the rules for measurements are essentially the same.

$$p(m) = \langle\psi|_{ab}\left(P_a^{(m)} \otimes I_b\right)|\psi\rangle_{ab} \qquad |\psi\rangle_{ab} \longrightarrow |\psi^{(m)}\rangle_{ab} = \frac{\left(P_a^{(m)} \otimes I_b\right)|\psi\rangle_{ab}}{\sqrt{p(m)}}$$

iv) Throw Away Part of the System. There's nothing to stop Alice or Bob from just "walking out of the lab" or more likely literally losing their qubit. Losing track of a system, however it's done, means taking the partial trace over that system.

Entanglement is Conserved Under Local Unitary Operations

Unitary operators maintain the "relationships" of a state, both to itself and to everything else, and that includes entanglement.

You can also think of this as being like translating a book into code; it still carries the same information, just written in a different way.

$$\boxed{S\left[U_a \rho_a U_a^\dagger\right] = S\left[\rho_a\right]}$$

Proving this is a homework assignment.

Local Measurements' Effect on Entanglement

When Alice does a projective measurement of a pure entangled state, $\rho = |\psi\rangle_{ab}\langle\psi|_{ab}$, one of many projections, $\{P_a^{(m)} \otimes I_b\}$,[6] are applied to her portion of the state . The probabilities and effects of those measurements are:

$$p(m) = \langle\psi|_{ab}\left(P_a^{(m)} \otimes I_b\right)|\psi\rangle_{ab}$$

$$|\psi\rangle_{ab} \longrightarrow |\psi^{(m)}\rangle_{ab} = \frac{\left(P_a^{(m)} \otimes I_b\right)|\psi\rangle_{ab}}{\sqrt{p(m)}}$$

$$\rho \longrightarrow \rho^{(m)} = \frac{\left(P_a^{(m)} \otimes I_b\right)\rho\left(P_a^{(m)} \otimes I_b\right)}{p(m)}$$

And the effect on an unreported measurement (like flipping and covering a coin) is a mixed state, because without knowing the result we're forced to sum over the states with their probabilities.

$$\boxed{\rho \longrightarrow \rho' = \sum_m p(m)\rho^{(m)} = \sum_m \left(P_a^{(m)} \otimes I_b\right)\rho\left(P_a^{(m)} \otimes I_b\right)}$$

Measurements by Alice can have no impact on Bob's reduced density matrix.[7] If they did, Bob would be able to detect those differences and we'd have access to "spooky action at a distance" (and faster than light communication). Measurement projections have the property that they cover all possible results from the system under consideration (like the probabilities they generate), so $I = \sum_m P^{(m)}$. Before Alice's measurement

[6]What would normally be subscripts, "P_m", are now superscripts, "$P^{(m)}$", because we need the subscript space to indicate subspace. If we need even more indexes later, we'll have to start being clever.

[7]Entanglement is a lot of things, but it's definitely not like a "voodoo doll" where something done to one magically has an effect on the other. Take this to heart; pop science gets this one wrong almost universally.

$$\rho_b = Tr_a[\rho]$$

and afterward

$$\begin{aligned}
\rho_b &= Tr_a\left[\sum_m p(m)\rho^{(m)}\right] \\
&= Tr_a\left[\sum_m (P_m \otimes I_b)\rho(P_m \otimes I_b)\right] \\
&= \sum_m Tr_a\left[(P_m \otimes I_b)\rho(P_m \otimes I_b)\right] \\
&= \sum_m Tr_a\left[(P_m \otimes I_b)(P_m \otimes I_b)\rho\right] \\
&= \sum_m Tr_a\left[(P_m^2 \otimes I_b^2)\rho\right] \\
&= \sum_m Tr_a\left[(P_m \otimes I_b)\rho\right] \\
&= Tr_a\left[(\sum_m P_m \otimes I_b)\rho\right] \\
&= Tr_a\left[(I_a \otimes I_b)\rho\right] \\
&= Tr_a\left[\rho\right]
\end{aligned}$$

Therefore, Alice's measurement does not impact Bob's density matrix (it's nice to see the math working out). Notice that we can express ρ_b in terms of these individual measurement results:

$$\rho_b = Tr_a\left[\sum_m p(m)\rho^{(m)}\right] = \sum_m p(m) Tr_a\left[\rho^{(m)}\right] = \sum_m p(m)\rho_b^{(m)}$$

So if Alice says nothing, then ρ_b cannot change. On the other hand, as we'll see in quantum teleportation, if Alice communicates with Bob over classical channels, then they can change Bob's ρ_b. What we've done so far is consider the intrinsic effect that Alice's measurement has on Bob's density matrix (none!). But of course, Alice's measurement produces a result, m, for Alice and if she picks up the phone and calls Bob, then his density matrix changes.[8] By telling Bob the result of her measurement, they can agree on what new state they share, $\rho^{(m)}$.

$$\rho^{(m)} = \frac{\left(P_a^{(m)} \otimes I_b\right)\rho\left(P_a^{(m)} \otimes I_b\right)}{p(m)}$$

Measurement operators, like $P^{(m)}$, map states to states, $P^{(m)}|\psi\rangle = |\psi^{(m)}\rangle$. So if $\rho = |\psi\rangle\langle\psi|$ is a pure state, then $\rho^{(m)} = \frac{|\psi^{(m)}\rangle_{ab}\langle\psi^{(m)}|_{ab}}{p(m)}$ is also a pure state. It's possible that this

[8]Remember that the density matrix describes probabilities from both classical and quantum sources, and you can certainly change someone's probability distribution by giving them new information. In fact, that's arguably how communication is defined.

new state will actually be more entangled (we'll see examples of this later), but on average entanglement decreases under local measurements.

Since $\rho^{(m)}$ is a pure state, if $\rho_b^{(m)} = Tr_a\left[\rho^{(m)}\right]$, then $S\left[\rho_b^{(m)}\right]$ is the entanglement between A and B. The average entanglement after Alice's measurement is

$$\sum_m p(m) S\left[\rho_b^{(m)}\right]$$

At the same time we know that $\rho_b = \sum_m p(m)\rho_b^{(m)}$ and by subadditivity ($\sum_k p_k S[\rho_k] \leq S\left[\sum_k p_k \rho_k\right]$) we know that

$$\boxed{\sum_m p(m) S\left[\rho_b^{(m)}\right] \leq S\left[\sum_m p(m)\rho_b^{(m)}\right] = S\left[\rho_b\right]}$$

In other words, the average entanglement after a local measurement, $\sum_m p(m) S\left[\rho_b^{(m)}\right]$, is less than the entanglement before, $S\left[\rho_b\right]$.

Example The simplest example of a pure entangled bipartite state is always $|\Phi_+\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$ with the usual arrangement: Alice and Bob each have one qubit.

We'll look at what happens when Alice makes a measurement of $\rho = |\Phi_+\rangle\langle\Phi_+|$ in the computational basis and verify that $\sum_m p(m) S\left[\rho_b^{(m)}\right] \leq S\left[\rho_b\right]$.

$$\begin{aligned}
\rho_b &= Tr_a\left[\left(\frac{|0\rangle_a|0\rangle_b+|1\rangle_a|1\rangle_b}{\sqrt{2}}\right)\left(\frac{\langle 0|_a\langle 0|_b+\langle 1|_a\langle 1|_b}{\sqrt{2}}\right)\right] \\
&= \left(\frac{\langle 0|_a|0\rangle_b+\langle 1|_a|1\rangle_b}{\sqrt{2}}\right)\left(\frac{|0\rangle_a\langle 0|_b+|1\rangle_a\langle 1|_b}{\sqrt{2}}\right) \\
&= \frac{1}{2}\langle 0|0\rangle_a|0\rangle_b\langle 0|_b + \frac{1}{2}\langle 1|1\rangle_a|1\rangle_b\langle 1|_b \\
&= \frac{1}{2}|0\rangle_b\langle 0|_b + \frac{1}{2}|1\rangle_b\langle 1|_b \\
&\sim \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}
\end{aligned}$$

Before Alice makes a measurement, the entanglement is

$$S[\rho_b] = -\frac{1}{2}\log_2\left(\frac{1}{2}\right) - \frac{1}{2}\log_2\left(\frac{1}{2}\right) = 1$$

The measurement operators for the computational basis are $P^{(0)} = |0\rangle\langle 0|$ and $P^{(1)} = |1\rangle\langle 1|$. The probability of each of Alice's measurement results are

$$
\begin{aligned}
p(0) &= Tr\left[\left(P_a^{(0)} \otimes I_b\right)\rho\right] \\
&= Tr\left[\left(|0\rangle_a\langle 0|_a \otimes I_b\right)\left(\frac{|0\rangle_a|0\rangle_b+|1\rangle_a|1\rangle_b}{\sqrt{2}}\right)\left(\frac{\langle 0|_a\langle 0|_b+\langle 1|_a\langle 1|_b}{\sqrt{2}}\right)\right] \\
&= Tr\left[\left(\frac{|0\rangle_a|0\rangle_b}{\sqrt{2}}\right)\left(\frac{\langle 0|_a\langle 0|_b+\langle 1|_a\langle 1|_b}{\sqrt{2}}\right)\right] \\
&= \left(\frac{\langle 0|_a\langle 0|_b}{\sqrt{2}}\right)\left(\frac{|0\rangle_a|0\rangle_b+|1\rangle_a|1\rangle_b}{\sqrt{2}}\right) \\
&= \frac{1}{2}
\end{aligned}
$$

and

$$
p(1) = Tr\left[\left(|1\rangle_a\langle 1|_a \otimes I_b\right)\rho\right] = \ldots = \frac{1}{2}
$$

The effects of these measurements are

$$
\begin{aligned}
\rho^{(0)} &= \frac{P^{(0)}\rho P^{(0)}}{p(0)} \\
&= \frac{\left(|0\rangle_a\langle 0|_a\otimes I_b\right)\left(\frac{|0\rangle_a|0\rangle_b+|1\rangle_a|1\rangle_b}{\sqrt{2}}\right)\left(\frac{\langle 0|_a\langle 0|_b+\langle 1|_a\langle 1|_b}{\sqrt{2}}\right)\left(|0\rangle_a\langle 0|_a\otimes I_b\right)}{\frac{1}{2}} \\
&= 2\left(\frac{|0\rangle_a|0\rangle_b}{\sqrt{2}}\right)\left(\frac{\langle 0|_a\langle 0|_b}{\sqrt{2}}\right) \\
&= |0\rangle_a|0\rangle_b\langle 0|_a\langle 0|_b
\end{aligned}
$$

and

$$
\rho^{(1)} = \ldots = |1\rangle_a|1\rangle_b\langle 1|_a\langle 1|_b
$$

The partial trace of these density matrices are

$$
\rho_b^{(0)} = Tr_a\left[|0\rangle_a|0\rangle_b\langle 0|_a\langle 0|_b\right] = \langle 0|0\rangle_a|0\rangle_b\langle 0|_b = |0\rangle_b\langle 0|_b
$$

and

$$
\rho_b^{(1)} = \ldots = |1\rangle_b\langle 1|_b
$$

These are pure states, meaning that the entropy of these states is

$$
S\left[\rho_b^{(0)}\right] = S\left[\rho_b^{(1)}\right] = 0
$$

and so finally, plugging into $\sum_m p(m)S\left[\rho_b^{(m)}\right] \leq S\left[\rho_b\right]$,

$$
0 = \frac{1}{2}0 + \frac{1}{2}0 = p(0)S\left[\rho_b^{(0)}\right] + p(1)S\left[\rho_b^{(1)}\right] \leq S[\rho_b] = 1
$$

$\blacksquare$

Exercises

#1) Local Unitary Operations.

The most general local unitary operations are

$$U_a \otimes U_b$$

Show that this operation has no impact on entanglement, by showing that for $\rho \in A \otimes B$:

a) $Tr_b \left[U_a \otimes U_b \rho U_a^\dagger \otimes U_b^\dagger \right] = U_a \rho_a U_a^\dagger$

b) $S \left[U_a \rho_a U_a^\dagger \right] = S\left[\rho_a \right]$

#2) Two Sides of the Same Coin.

$$|\psi\rangle = \frac{1}{\sqrt{3}}|1\rangle_a|1\rangle_b + \frac{1}{\sqrt{6}}|1\rangle_a|2\rangle_b + \frac{1}{2}|3\rangle_a|3\rangle_b + \frac{1}{2}|4\rangle_a|5\rangle_b$$

a) Find the density matrix and reduced density matrices: ρ, ρ_a, and ρ_b.

b) Find the entanglement of this state in bits using ρ_a and then ρ_b.

#3) The Problem With Having Too Much Rope.

Alice and Bob each have one particle from an entangled qutrit[9] in the state

$$|\mu\rangle = \frac{1}{4}|1\rangle_a|2\rangle_b + \frac{1}{4}|2\rangle_a|1\rangle_b + \sqrt{\frac{7}{8}}|3\rangle_a|3\rangle_b$$

a) What is the entanglement between A and B in bits?

b) Alice makes a “Is it 3 or not?” measurement, with measurement operators $P_a^{(3)} = |3\rangle_a\langle 3|_a$ and $P_a^{(\not{3})} = |1\rangle_a\langle 1|_a + |2\rangle_a\langle 2|_a$. For each result, what is the entanglement in bits after Alice’s measurement?

c) What is the average entanglement after Alice’s measurement and what is the probability that the entanglement increases?

d) How will Alice be able to tell if the entanglement has increased? How can Bob find out?

[9] “Qubits” have two states, “qutrits” have three, and “qudits” have d. Nobody has seriously tried to get a name for four-state systems to catch on.