# Entanglement in a Nutsell

Daniel Freeman
(Dated: November 21st, 2014)

The following is an overview of entanglement as discussed in class and in discussion sections.

## I. WARMUP

This simplest states considered in quantum mechanics are separable pure states. These states live in some Hilbert space with tensor product structure:

$$\mathcal{H} = \mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_{\mathcal{B}} \otimes \mathcal{H}_{\mathcal{C}} \otimes ... \tag{1}$$

where we have become familiar with writing a states in this space as:

$$|\psi_{\text{separable}}\rangle = |\psi_A\rangle \otimes |\psi_B\rangle \otimes |\psi_C\rangle \otimes ... \tag{2}$$

The emphasis on 'separable' here is meant to reinforce that this state has only one term. That it is factorizable in this way distinguishes it from other states we will consider momentarily. Should we want to complicate our lives, we can likewise define separable mixed states:

$$\rho = \sum_i p_i \rho_A^i \otimes \rho_B^i \otimes \rho_C^i \otimes ... \tag{3}$$

Where $\sum_i p_i = 1$ and where $\rho_A$, $\rho_B$,etc. are densities matrices formed from the separable states in the form above. That is, $\rho_A = |\psi_A\rangle\langle\psi_A|$.

A fairly simple definition of entanglement follows: entangled states are those which are not separable. We have already seen some examples of these states such as the Bell states:

$$|\psi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle) \tag{4}$$

$$|\phi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle) \tag{5}$$

But these examples don't give us a good operational meaning for entanglement: What can we do to it? Can we quantify it? What does it allow us to do? We'll answer each of these questions in turn.

## II. WHAT CAN WE DO TO ENTANGLEMENT?

Suppose Alice and Bob each have a qubit. We'd like to investigate to what extent Alice and Bob can alter the presence of any entanglement shared by their qubits. Suppose that Alice and Bob are well separated, are only allowed to do local operations on their qubits, and are allowed to communicate classically (in the literature, this set of constraints is called LOCC, or Local Operations Classical Communication). Suppose Alice and Bob's qubit is in the state $\psi_{AB}$. If, through local operations and classical communication, Alice and Bob can transform their pair into another state, call it $\phi_{AB}$, and then transform their state back to $\psi_{AB}$, then we call $\psi_{AB}$ and $\phi_{AB}$ LOCC equivalent.

It's fairly easy to see that the Bell states above are LOCC equivalent. If Alice and Bob know ahead of time which Bell state they have, a single qubit rotation by either of them will take $|\psi^+\rangle$ to $|\phi^+\rangle$ (i.e., let Alice apply the unitary $\sigma_X \otimes 1_D$). Another application of the unitary will rotate back to the original Bell state. Note however that the state $|00\rangle + |11\rangle$ is not LOCC equivalent to the state $|00\rangle$. While measurement followed by a conditional application of unitaries can map $|00\rangle + |11\rangle \to |00\rangle$ via LOCC, the other direction is not possible under LOCC. This example is meant to demonstrate that there's something inherently nonlocal about entanglement. That is, local operations can keep you within a given manifold of LOCC equivalent states, but you can't create an entangled state via LOCC operations. Further, LOCC operations can even destroy entanglement, as in the previous example.

Quite a bit more could be said about ways in which entanglement can be transformed so as to be useful, which I won't talk much more about. But if we have any hope to transform entanglment controllably, we need a good measure of entanglement.

## III. CAN WE QUANTIFY ENTANGLEMENT?

Yes. In fact, there are well over a dozen entanglement measures with widespread use in the literature. The first, and probably most widespread candidate is Schmidt rank.

Suppose we have an arbitrary 2-qubit state: $|\psi_{AB}\rangle$. For the Hilbert space $\mathcal{H_A} \otimes \mathcal{H_B}$, let $\{|e_i\rangle \otimes |e_j\rangle\}$ be an orthonormal basis for the space. Then, in general:

$$|\psi_{AB}\rangle = \sum_{i=0}^{d_A-1} \sum_{j=0}^{d_B-1} A_{ij} |e_i\rangle \otimes |e_j\rangle \tag{6}$$

Written this way, if the matrix $A_{ij}$ is rank-1, then $\psi_{AB}\rangle$ is a separable state (remember that the rank of a matrix is just the number of linearly dependent vectors that make up (interchangeably) the columns or rows of the matrix). This should be intuitive–if a Matrix is rank 1, then upon diagonalizing the matrix, there's only 1 nonzero entry on the diagonal, or, in other words, there's only 1 term to write down. This supplies a simple entanglement measure: the measure is defined to be the number of nonzero terms in the diagonalization of $A_{ij}$, or equivalently, the rank of $A_{ij}$. This is the Schmidt Rank. In the literature, Schmidt rank is called: Schmidt Rank (unsurprisingly), $r(\psi)$, or sometimes $K$.

In general, if we've diagonalized an arbitrary quantum state, we can write it as:

$$|\psi_{AB}\rangle = \sum_{i=0}^{r(\psi)} a_i |\tilde{e}_i\rangle \otimes |\tilde{e}_i\rangle \tag{7}$$

Where the tilde over the $e$s is just meant to emphasize that we're working in the basis in which $\psi_{AB}\rangle$ is diagonal. Notice the matrix $A_{ij}$ has been replaced with a set of scalars $a_i$ which are nothing other than the diagonal entries of the diagonalization of $A_{ij}$. These diagonal entries also have a nice operational interpretation in terms of reduced density matrices. Letting $\rho_{AB} \equiv |\psi_{AB}\rangle\langle\psi_{AB}|$:

$$\begin{aligned}
\rho_{AB} &= \sum_{i=0}^{K-1} \sum_{j=0}^{K-1} a_i a_j |\tilde{e}_i\rangle \otimes |\tilde{e}_i\rangle\langle\tilde{e}_j| \otimes \langle\tilde{e}_j| \\
\mathrm{Tr_B}[\rho_{\mathrm{AB}}] &= \sum_{l=0}^{K-1} 1_D \otimes \langle\tilde{e}_k|(\sum_{i=0}^{K-1} \sum_{j=0}^{K-1} a_i a_j |\tilde{e}_i\rangle \otimes |\tilde{e}_i\rangle\langle\tilde{e}_j| \otimes \langle\tilde{e}_j|) 1_D \otimes |\tilde{e}_k\rangle \\
&= \sum_{i=0}^{K-1} a_i^2 |\tilde{e}_i\rangle\langle\tilde{e}_i|
\end{aligned} \tag{8}$$

But these coefficients are the same as if we had traced over $A$:

$$\begin{aligned}
\rho_{AB} &= \sum_{i=0}^{K-1} \sum_{j=0}^{K-1} a_i a_j |\tilde{e}_i\rangle \otimes |\tilde{e}_i\rangle\langle\tilde{e}_j| \otimes \langle\tilde{e}_j| \\
\mathrm{Tr_A}[\rho_{\mathrm{AB}}] &= \sum_{l=0}^{K-1} \langle\tilde{e}_k| \otimes 1_D(\sum_{i=0}^{K-1} \sum_{j=0}^{K-1} a_i a_j |\tilde{e}_i\rangle \otimes |\tilde{e}_i\rangle\langle\tilde{e}_j| \otimes \langle\tilde{e}_j|)|\tilde{e}_k\rangle \otimes 1_D \\
&= \sum_{j=0}^{K-1} a_j^2 |\tilde{e}_j\rangle\langle\tilde{e}_J|
\end{aligned} \tag{9}$$

It's important to emphasize that the states are different in the two resulting partial traces, but that the coefficients are the same. Identiftying $a_i^2$ as a probability draws a nice correspondence between the Schmidt coefficients and the terms in the above reduced density matrices.

The second most widely studied type of entanglement is the Entanglement Entropy. If you've studied classical information theory, or seen a statistical mechanics oriented thermodynamics course, you've probably seen the expression for Shannon entropy:

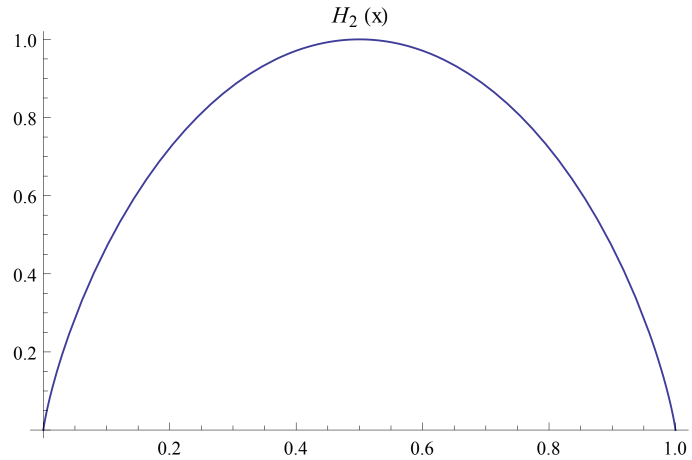


FIG. 1: The Shannon entropy of a random variable which is 1 with probability x and 0 with probability 1-x.

$$H_b(x) = -\sum_i p(x_i) log_b(p(x_i)) \tag{10}$$

Where $b$ specifies the base of the log, and $p(x_i)$ are probabilities of outcomes for some random variable. For the case of a binary random variable (like a coin flip), this is easy to visualize, and is depicted in Figure 1.

Intuitively, entropy is maximized when the random variable is 1 with probability .5 and 0 with probability .5. In this way, the Shannon entropy is giving some sort of measure on our uncertainty about the outcomes in the system. As the probability of x tends towards 0 or 1, the system starts becoming more predictable, and this measure commensurately decreases. After all, if you have a heavily biased coin (i.e., $x \rightarrow 1$), you will be able to predict outcomes more often. Predictability and entropy go hand in hand in classical information theory, where these sorts of measures are used to study compressibility of signals, as well as the content of encrypted data.

Quantum mechanically, a related entropy measure will tell us something similar. The (sort of) obvious extension of Shannon entropy to the classical realm is the Von Neumann Entropy:

$$S = -\mathrm{Tr}[\rho \mathrm{log} \rho] = -\sum_{\mathrm{i}} \lambda_{\mathrm{i}} \mathrm{log} \lambda_{\mathrm{i}} \tag{11}$$

Where the $\lambda_i$ are just the squares of the Schmidt coefficients, as discussed previously. Before we understand how this object behaves for entangled states, let's look at how it classifies pure and mixed states:

If we have a pure state, then there's only 1 term in the sum, namely $\lambda_1 = 1$. In this case, the Von Neumann Entropy is zero! Note that this is independent of whether the state is separable or not–the Von Neumann entropy is zero for all pure states. From our discussion of density matrices, this should seem sensible, as we discussed in class how pure density matrices can be thought of as quantum systems which we have perfect information about. But this can't be the only ingredient of our entanglement measure, as the Von Neumann entropy alone can't distinguish between an entangled pure state and a separable pure state.

Now, suppose we want to measure the Von Neumann entropy of a totally mixed state:

$$\begin{aligned} S_{\text{totally mixed}} &= -\mathrm{Tr}[\sum_{\mathrm{i}=0}^{2^{\mathrm{n}}-1} \frac{1}{2^{\mathrm{n}}} \mathrm{log}_2(2^{-\mathrm{n}})] \\ &= +\mathrm{Tr}[\sum_{\mathrm{i}=0}^{2^{\mathrm{n}}-1} \frac{\mathrm{n}}{2^{\mathrm{n}}}] = \mathrm{n} \end{aligned} \tag{12}$$

Now we see that a totally mixed state has a nonvanishing amount of entropy (which scales linearly in the number of qubits). This is also somewhat sensible, as we just finished talking about how entropy is a measure of uncertainty, and totally mixed states are the maximum-uncertain states in quantum mechanics. But this doesn't quite get at an entanglement measure. For that, we'll need to define one more object which looks like the Von Neumann entropy with one caveat:

$$S(\rho_B) = -\mathrm{Tr}[\rho_\mathrm{B}\mathrm{log}\rho_\mathrm{B}] \tag{13}$$

This is the Entanglement Entropy. More verbosely, this is sometimes called the Entanglement Entropy of subsystem B with respect to A. Even for pure states (which have vanishing Von Neumann entropy), this measure is not necessarily vanishing, and tells something meaningful about the amount of entanglement shared between two subsystems. You should be able to convince yourself that this measure is still vanishing for separable states. You should also fairly easily see that something interesting happens for entangled states. For bipartite (i.e., 2 qubits) pure states, one can show fairly easily that $S(\rho_A) = S(\rho_B) = 1$.

Another measure is the Entanglement of Formation:

$$E_f(\rho) = \mathrm{min}_\epsilon \mathrm{p_i E_E}(|\psi\rangle) \tag{14}$$

Where $\epsilon$ is some ensemble of pure states and probabilities: $\epsilon \equiv \{(p_i, |\psi_i\rangle)\}$, where $\rho = \sum_i p_i|\psi\rangle\langle\psi|$, and where $E_E(|\psi_i\rangle)$ is the Entanglement Entropy of $\psi_i$. This definition requires a little bit of unpacking: Remember that density matrices are not a unique representation of a quantum state. $|0\rangle\langle 0| + |1\rangle\langle 1|$ is the same as $|+\rangle\langle +| + |-\rangle\langle -|$. So, in general, there's some ensemble of pure states that can be used to generate $\rho$, and the minimization is done over this ensemble. This entropy measure has a nice operational meaning related to the number of Bell states needed to generate a given $\rho$. For more information, see `http://www.quantiki.org/wiki/Entanglement_of_formation`.

Another measure favored by experimentalists is called the Concurrence. For pure states, it's simply:

$$C \equiv \sqrt{2(1 - \mathrm{Tr}[\rho_\mathrm{A}^2])} \tag{15}$$

Where $\rho_A$ is the reduced density matrix for some bipartite state $\rho_{AB}$. You are to imagine the experimentalist has a two-qubit system, and wants to measure the entanglement between the two systems. If the system is in a separable state, the concurrence is clearly 0. However, if the pair of qubits is in a Bell state, it's easy to show that the Concurrence is 1. Experimentalists like to show that they can make things close to Bell states, so this measure caught on. There's also a nice generalization to mixed states:

$$C(\rho) = \mathrm{min}_\epsilon \sum_\mathrm{i} \mathrm{p_i C}(|\psi_\mathrm{i}\rangle) \tag{16}$$

Where we again minimize over ensembles of states which realize a given $\rho$. It can be shown that for 2-qubit states, this is the same as:

$$C(\rho) = \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\} \tag{17}$$

Where the $\lambda_i$ are Schmidt coefficients in decreasing order.

The last measure we will consider is a multipartite distance measure (where all of the previous measures are bipartite measures) called Quantum Relative Entropy. First we define a distance measure between $\rho_1$ and $\rho_2$ to be:

$$D(\rho_1||\rho_2) = \mathrm{Tr}[\rho_1 \mathrm{log}\rho_1 - \rho_1 \mathrm{log}\rho_2] \tag{18}$$

The relative entropy is then:

$$\mathrm{RE}(\rho) = \min_{\sigma\in\mathrm{S}}(\rho||\sigma) \tag{19}$$

Where $S$ is some ensemble of separable states. Note that this measure is extremely difficult to calculate in general.

Hopefully, this will have given you an idea of the many different ways in which quantum information scientists try and quantify entanglement. The takeaway from this list, if anything, is that there's not really a single overarching entanglement measure because different measures are useful for different contexts.

## IV. WHAT DOES ENTANGLEMENT ALLOW US TO DO?

I talked about this quite a bit in my last couple of discussion sections. Instead of rehashing those discussion sections, I'll point to the two papers that I covered and briefly summarize them:
First: http://arxiv.org/abs/1204.3107

In this paper, Van den Nest considers to what extent different amounts of entropy during a calculation allow the ability to perform universal quantum computation. The punchlines are: 1) If the Schmidt rank of all your qubits is logarithmically small in the number of qubits at every step of a quantum computation, then whatever computations you try to perform are efficiently classically simulatable. 2) You can perform arbitrary quantum computation with polynomially (vanishingly) small Renyi Entropy. I did not discuss Renyi entropy here, but it is a standard generalization of Entanglement Entropy.

Second: http://arxiv.org/abs/1401.4174 or if you want the fancy nature version: http://www.nature.com/nature/journal/v510/n7505/full/nature13460.html

In this paper, Howard et al. demonstrate that a property called Contextuality is fundamentally important in allowing for quantum speedup. This paper is fairly technically heavy, but the gist of the problem is that a) we know how to make transversal gates really well, and b) we know how to make stabilizer codes really well, but c) If you only ever use transversal gates and stabilizer codes, you can't actually perform arbitrary quantum computation. The authors spend a long time showing that this special resource called Contextuality is related to the necessary ingredient to extend Transversal/Stabilizer computing to universal quantum computation. They do this by relating Contextuality to Magic State Distillation in a fairly clever way. The punchlines of this paper are: 1) If a state is distillable via magic state distillation, then it is necessarily a Contextual state 2) It might be the case (in dimension greater than 3) that Contextual states are necessarily distillable via magic state distillation. (It is already known than in dimension 2, this is false–i.e., there are Contextual states which are not distillable via magic state distillation).

If you would like to know more about these last two topics, I encourage you to read about these topics more. This last subheading is very much still an active research area.