9 Entanglement

When we have a composite system, the pure states of that system can not, in general, be written as a tensor product of pure states of the subsystems. We say that such states are *entangled*. When we talk about a **bipartite system** we mean a composite of two systems. We have already seen an example of an entangled state of a bipartite system AB, namely the state

$$|\phi^+\rangle_{\mathsf{AB}} = (|0\rangle_{\mathsf{A}} \otimes |0\rangle_{\mathsf{B}} + |1\rangle_{\mathsf{A}} \otimes |1\rangle_{\mathsf{B}})/\sqrt{2}$$

which was used in the quantum strategy for the CHSH game. We will see later that without entanglement, Alice and Bob can do no better than the best classical strategy in any game like the CHSH game. We will also study a number of other quite different uses for entanglement.

9.1 The Schmidt decomposition

In this section, we will prove that any *pure state of a bipartite system* can be written in a certain standard form called a **Schmidt decomposition**. The Schmidt decomposition makes certain features of pure states of bipartite systems apparent, such as the fact that the states of the two parts have the same eigenvalues. It also lets us determine when two pure states of a bipartite system can be reversibly transformed into one another by *local* operations.

Note that we can think of a composite of any number of systems as being bipartite if we specify a partition of its subsystems into two sets. Such a partition is called a **bipartition**. For example, we might consider the bipartition $\mathsf{A} : \mathsf{BR}$ of a composite system ABR.

Lemma 1. Suppose $L \in \mathcal{L}(\mathcal{H}_{\mathsf{A}}, \mathcal{H}_{\mathsf{B}})$ and $W_{\mathsf{A}} := L^\dagger L$ has eigendecomposition

$$W_{\mathsf{A}} = \sum_{j=1}^{d_{\mathsf{A}}} \lambda_j |\alpha_j\rangle\langle\alpha_j|_{\mathsf{A}} \text{ where } \lambda_1 \geq \ldots \geq \lambda_{d_{\mathsf{A}}},$$

and $r := \mathrm{rank}(W_{\mathsf{A}})$. Then

$$L = \sum_{j=1}^{r} \lambda_j^{1/2} |\phi_j\rangle_{\mathsf{B}} \langle\alpha_j|_{\mathsf{A}}$$

for some orthonormal set $\{|\phi_j\rangle_{\mathsf{B}} : 1 \leq j \leq r\}$.

Proof. From $W_{\mathsf{A}} \geq 0$ and the ordering of the eigenvalues λ_j we know that $\lambda_j > 0$ for $j \leq r$ and $\lambda_j = 0$ for $j > r$. Since W_{A} is hermitian, $\{|\alpha_j\rangle_{\mathsf{A}} : 1 \leq j \leq d_{\mathsf{A}}\}$ is an orthonormal basis for $\mathcal{H}_{\mathsf{A}}$. So, $L = \sum_{j=1}^{d_{\mathsf{A}}} |\tilde{\phi}_j\rangle_{\mathsf{B}}\langle\alpha_j|_{\mathsf{A}}$ where $|\tilde{\phi}_j\rangle_{\mathsf{B}} = L|\alpha_j\rangle_{\mathsf{A}}$ and

$$\langle\tilde{\phi}_k|\tilde{\phi}_j\rangle = \langle\alpha_k|L^\dagger L|\alpha_j\rangle = \langle\alpha_k|W_{\mathsf{A}}|\alpha_j\rangle = \lambda_j \delta_{kj}.$$

Letting $|\phi_j\rangle_{\mathsf{B}} := \lambda^{-1/2}|\tilde{\phi}_j\rangle_{\mathsf{B}}$ for $1 \leq j \leq r$ defines an orthonormal set, and the result follows. □

Remark 2. Recall that $\mathrm{vec}_{\mathsf{B}}|\psi\rangle_{\mathsf{A}}\langle\phi|_{\mathsf{B}} = |\psi\rangle_{\mathsf{A}} \otimes |\phi\rangle^*_{\mathsf{B}}$.

Theorem 3 (Schmidt decomposition)**.** Any vector $|\psi\rangle_{\mathsf{AB}}$ in the Hilbert space of a bipartite system AB has a **Schmidt decomposition** $|\psi\rangle_{\mathsf{AB}} = \sum_{j=1}^r \lambda_j^{1/2}|\alpha_j\rangle_{\mathsf{A}} \otimes |\beta_j\rangle_{\mathsf{B}}$ where:

1. The λ_j are real, strictly positive and decreasing $\lambda_1 \geq \ldots \geq \lambda_r > 0$.
2. $\{|\alpha_j\rangle_{\mathsf{A}} : 1 \leq j \leq r\}$ and $\{|\beta_j\rangle_{\mathsf{B}} : 1 \leq j \leq r)\}$ are orthonormal sets;
3. $r = \mathrm{rank}(\psi_{\mathsf{A}})$ and $\sum_{1\leq j\leq r} \lambda_j |\alpha_j\rangle\langle\alpha_j|_{\mathsf{A}}$ is an eigendecomposition of $\psi_{\mathsf{A}} := \mathrm{Tr}_{\mathsf{B}}|\psi\rangle\langle\psi|_{\mathsf{AB}}$.
4. $r = \mathrm{rank}(\psi_{\mathsf{B}})$ and $\sum_{1\leq j\leq r} \lambda_j |\beta_j\rangle\langle\beta_j|_{\mathsf{B}}$ is an eigendecomposition of $\psi_{\mathsf{B}} := \mathrm{Tr}_{\mathsf{A}}|\psi\rangle\langle\psi|_{\mathsf{AB}}$.

The numbers $\lambda_j^{1/2}$ are the (non-zero) **Schmidt coefficients** of $|\psi\rangle_{\mathsf{AB}}$, and r is the **Schmidt rank** of $|\psi\rangle_{\mathsf{AB}}$. If $|\psi\rangle_{\mathsf{AB}}$ is a state vector (i.e. a unit vector) then $\sum_{1\leq j\leq r} \lambda_j = 1$.

Proof. We expand $|\psi\rangle_{\mathsf{AB}}$ in the computational basis,

$$|\psi\rangle_{\mathsf{AB}} = \sum_{a=0}^{d_{\mathsf{A}}-1}\sum_{b=0}^{d_{\mathsf{B}}-1} x_{ab}|a\rangle_{\mathsf{A}} \otimes |b\rangle_{\mathsf{B}}, \text{ and let } L^\dagger := \mathrm{vec}_{\mathsf{B}}^{-1}|\psi\rangle_{\mathsf{AB}} = \sum_{a,b} x_{ab}|a\rangle_{\mathsf{A}}\langle b|_{\mathsf{B}}.$$

Using $\mathrm{Tr}_{\mathsf{B}}|b\rangle\langle b'|_{\mathsf{B}} = \langle b'|b\rangle = \delta_{b'b} = \langle b|b'\rangle$, we have

$$\psi_{\mathsf{A}} = \mathrm{Tr}_{\mathsf{B}}|\psi\rangle\langle\psi|_{\mathsf{AB}} = \sum_{a,b}\sum_{a',b'} x_{ab}x^*_{a'b'}|a\rangle_{\mathsf{A}}\langle a'|_{\mathsf{A}}\mathrm{Tr}_{\mathsf{B}}|b\rangle\langle b'|_{\mathsf{B}} = L^\dagger L. \tag{9.1}$$

Let $\psi_{\mathsf{A}} = \sum_{j=1}^r \lambda_j|\alpha_j\rangle\langle\alpha_j|_{\mathsf{A}}$ be any eigendecomposition for ψ_{A} with eigenvalues λ_j in decreasing order. By Lemma 1 we have $L = \sum_{j=1}^r \lambda_j^{1/2}|\phi_j\rangle_{\mathsf{B}}\langle\alpha_j|_{\mathsf{A}}$, where $r = \mathrm{rank}(\psi_{\mathsf{A}})$ and $\{|\phi_j\rangle_{\mathsf{B}} : 0 \leq j \leq r\}$ is an orthonormal set. Therefore, $L^\dagger = \sum_{j=1}^r \lambda_j^{1/2}|\alpha_j\rangle_{\mathsf{A}}\langle\phi_j|_{\mathsf{B}}$ and

$$|\psi\rangle_{\mathsf{AB}} = \mathrm{vec}_{\mathsf{A}}L^\dagger = \sum_{j=1}^r \lambda_j^{1/2}|\alpha_j\rangle_{\mathsf{A}} \otimes |\beta_j\rangle_{\mathsf{B}}, \text{ where } |\beta_j\rangle_{\mathsf{B}} := |\phi_j\rangle^*_{\mathsf{B}}.$$

Since complex conjugation of vectors preserves orthogonality and norm, we have established points (1) to (3). (4) follows by taking the partial trace of $|\psi\rangle\langle\psi|_{\mathsf{AB}}$ over A and using the orthonormality of the $\{|\alpha_j\rangle_{\mathsf{A}}\}$. □

9.2 Mixed state entanglement

We say ρ_{AB} is a **product state** if it is a tensor product of local states $\rho_{\mathsf{AB}} = \alpha_{\mathsf{A}} \otimes \beta_{\mathsf{B}}$. A *pure* state $\rho_{\mathsf{AB}} = |\psi\rangle\langle\psi|_{\mathsf{AB}}$ is a product state iff the following equivalent conditions hold

- Its state vector[1] $|\psi\rangle_{\mathsf{AB}}$ is a product vector i.e. $|\psi\rangle_{\mathsf{AB}} = |\alpha\rangle_{\mathsf{A}} \otimes |\beta\rangle_{\mathsf{B}}$;
- Its state vector has Schmidt rank one;
- Both of its marginal states are pure.

[1] Only unique up to global phase, but the conditions given are independent of this.

Among states which are not pure, however, we regard not only product states as unentangled, but rather any convex combination of product states.

Definition 4. We say an operator $M_{\mathsf{AB}} \in \mathcal{L}(\mathcal{H}_{\mathsf{A}} \otimes \mathcal{H}_{\mathsf{B}})$ is **separable** (with respect to the $\mathsf{A} : \mathsf{B}$ bipartition) iff it can be written

$$M_{\mathsf{AB}} = \sum_x F(x)_{\mathsf{A}} \otimes G(x)_{\mathsf{B}}, \text{ where}$$
$$\forall x : F(x)_{\mathsf{A}} \in \mathcal{L}(\mathcal{H}_{\mathsf{A}}), F(x)_{\mathsf{A}} \geq 0, G(x)_{\mathsf{B}} \in \mathcal{L}(\mathcal{H}_{\mathsf{B}}), G(x)_{\mathsf{B}} \geq 0.$$

We denote the set of all such operators by $\mathbf{sep}(\mathsf{A} : \mathsf{B})$. Note that a separable operator is necessarily positive.

A state (density operator) ρ_{AB} of AB belongs to $\mathbf{sep}(\mathsf{A} : \mathsf{B})$ if and only if it is a convex combination of product states, that is,

$$\rho_{\mathsf{AB}} = \sum_x P_X(x) \alpha(x)_{\mathsf{A}} \otimes \beta(x)_{\mathsf{B}} \tag{9.2}$$

where P_X is a probability distribution and the $\alpha(x)_{\mathsf{A}}$ and $\beta(x)_{\mathsf{B}}$ are density operators. Any state which is not separable, we call **entangled**.

If Alice and Bob both have access to a random variable X with distribution P_X, and Alice prepares A in the state $\alpha(X)_{\mathsf{A}}$ and Bob prepares B in the state $\beta(X)_{\mathsf{B}}$, then the state of AB will be the one given in (9.2).

9.2.1 A necessary condition for separability: PPT

Deciding whether a given state is separable or not is a computationally hard problem.[2] However, there is a simple, efficiently checkable necessary condition for separability, based on the fact that the transpose map $\mathbf{t}^{\mathsf{A} \leftarrow \mathsf{A}}$ is positive but not completely positive.

Definition 5. We say that an operator $M_{\mathsf{AB}} \in \mathcal{L}(\mathcal{H}_{\mathsf{A}} \otimes \mathcal{H}_{\mathsf{B}})$ is PPT (positive-partial transpose) with respect to the $\mathsf{A} : \mathsf{B}$ bipartition if $\mathbf{t}^{\mathsf{A} \leftarrow \mathsf{A}} M_{\mathsf{AB}} \geq 0$. We denote the set of positive operators which are also PPT by $\mathbf{ppt}(\mathsf{A} : \mathsf{B})$.

Remark 6. Since taking the (total) transpose of an operator does not change its eigenvalues, $\mathbf{t}^{\mathsf{A} \leftarrow \mathsf{A}} M_{\mathsf{AB}} \geq 0$ iff $\mathbf{t}^{\mathsf{AB} \leftarrow \mathsf{AB}} \mathbf{t}^{\mathsf{A} \leftarrow \mathsf{A}} M_{\mathsf{AB}} = \mathbf{t}^{\mathsf{B} \leftarrow \mathsf{B}} M_{\mathsf{AB}} \geq 0$. So, we could take the transpose on the B system rather than the A system in the definition and it would be equivalent. It is also easy to check that taking the transpose with respect to a different orthonormal basis doesn't change which states are PPT.

Proposition 7. $\mathbf{sep}(\mathsf{A} : \mathsf{B}) \subseteq \mathbf{ppt}(\mathsf{A} : \mathsf{B})$. That is, an operator which is separable with respect to a given bipartition is also PPT with respect to that bipartition.

Proof. M_{AB} is separable iff it can be written $M = \sum_j F(j)_{\mathsf{A}} \otimes G(j)_{\mathsf{B}}$ where $F(j)_{\mathsf{A}} \geq 0, G(j)_{\mathsf{B}} \geq 0$ for all j. Therefore, $F(j)_{\mathsf{A}}^{\mathrm{T}} \geq 0$ for all j and $\mathbf{t}^{\mathsf{A} \leftarrow \mathsf{A}} M_{\mathsf{AB}} = \sum_j F(j)_{\mathsf{A}}^{\mathrm{T}} \otimes G(j)_{\mathsf{B}} \geq 0$. □

Remark 8. It can be shown that if $d_{\mathsf{A}} + d_{\mathsf{B}} \leq 5$ then $\mathbf{sep}(\mathsf{A} : \mathsf{B}) = \mathbf{ppt}(\mathsf{A} : \mathsf{B})$.

[2] NP-hard, see e.g. `http://arxiv.org/abs/0810.4507`

9.3 Some questions for example class 2

♣♣Suppose we know that a qubit is either in the pure state $\psi_0 = |\psi_0\rangle\langle\psi_0|$ where $|\psi_0\rangle := \cos(\theta)|0\rangle + \sin(\theta)|1\rangle$ or in the pure state $\psi_1 = |\psi_1\rangle\langle\psi_1|$ where $|\psi_1\rangle := \cos(\theta)|0\rangle - \sin(\theta)|1\rangle$ (for some given $\theta \in [0, \pi/4]$) but we are completely unsure which it is, i.e. the state is ψ_X where X takes values in $\{0, 1\}$ and, $P_X(0) = P_X(1) = 1/2$.

1. If we measure a POVM with result $\hat{X}$ taking values in $\{0, 1\}$, what is the maximum success probability $\Pr(\hat{X} = X)$?

2. Argue that the maximum success probability in the previous part depends on the states only through the absolute value of their inner product $|\langle\psi_0|\psi_1\rangle|$.

3. Now suppose we are given n qubits, either all prepared in state ψ_0 or all prepared in state ψ_1, so the state is $\psi_X^{\otimes n}$. If we measure a POVM on the whole n qubit system with result $\hat{X}$ taking values in $\{0, 1\}$, what is the maximum success probability $\Pr(\hat{X} = X)$?

4. Now (for the original $n = 1$ case) suppose that we perform a POVM whose result Y takes values in $\{0, 1, ?\}$, where the outcome ? means that we don't know which state the system is in.

   (a) Give elements $E(0), E(1)$ and $E(?)$ for the POVM such that

   $$\Pr(Y = 0|X = 1) = \Pr(Y = 1|X = 0) = 0$$

   (i.e. the measurement never gets the wrong state) and

   $$\Pr(?) = \cos(2\theta).$$

   (Hint: Start by considering the which forms of $E(0)$ and $E(1)$ are allowed by the constraints.) Try also to show that this the *smallest possible* value of $\Pr(?)$.

♣♣ Consider a game similar to the CHSH game where a referee sends a question S to Alice and a question T to Bob, and Alice and Bob respond to the referee with answers X and Y respectively. As in the CHSH game the questions and answers were all bits, but now we make no assumption on the sets $\mathcal{A}_S$, $\mathcal{A}_T$, $\mathcal{A}_X$, $\mathcal{A}_Y$ except that they are all finite. There is some function $f : \mathcal{A}_S \times \mathcal{A}_T \times \mathcal{A}_X \times \mathcal{A}_Y \to \{0, 1\}$ and the players win the game iff $f(S, T, X, Y) = 1$.

Suppose that, as in the quantum strategy for CHSH, Alice has a system A and Bob has a system B; For each $s \in \mathcal{A}_S$ there is a POVM $E_s : \mathcal{A}_X \to \mathcal{L}(\mathcal{H}_\mathsf{A})$ and X is the result of Alice measuring E_S on A; for each $t \in \mathcal{A}_T$ there is a POVM $F_t : \mathcal{A}_Y \to \mathcal{L}(\mathcal{H}_\mathsf{B})$, and Y is the result of Bob measuring F_T on B. We make no assumption on the dimensions d_A and d_B except that they are finite.

Show that, if the state of ρ_{AB} prior to the measurements is *separable*, then there is a local hidden variables model for how X and Y depend on S and T, and therefore, the strategy will be no better than the best deterministic strategy.

♣♣ Let $d_\mathsf{A} = d_\mathsf{B} = 2$ and let ρ_{AB} be the state whose matrix in the computational basis is

$$\rho_{\mathsf{AB}} = \frac{1}{5}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Is ρ_{AB} separable?