CS 766/QIC 820 Theory of Quantum Information (Fall 2011)

Lecture 17: Measures of entanglement

The topic of this lecture is *measures of entanglement*. The underlying idea throughout this discussion is that entanglement may be viewed as a resource that is useful for various communication-related tasks, such as teleportation, and we would like to quantify the amount of entanglement that is contained in different states.

17.1 Maximum inner product with a maximally entangled state

We will begin with a simple quantity that relates to the amount of entanglement in a given state: the maximum inner product with a maximally entangled state. It is not necessarily an interesting concept in its own right, but it will be useful as a mathematical tool in the sections of this lecture that follow.

Suppose $\mathcal{X}$ and $\mathcal{Y}$ are complex Euclidean spaces, and assume for the moment that $\dim(\mathcal{X}) \geq \dim(\mathcal{Y}) = n$. A *maximally entangled state* of a pair of registers (X, Y) to which these spaces are associated is any pure state uu^* for which

$$\operatorname{Tr}_{\mathcal{X}}(uu^*) = \frac{1}{n}\mathbb{1}_{\mathcal{Y}};$$

or, in other words, tracing out the larger space leaves the completely mixed state on the other register. Equivalently, the maximally entangled states are those that may be expressed as

$$\frac{1}{n}\operatorname{vec}(U)\operatorname{vec}(U)^*$$

for some choice of a linear isometry $U \in \mathrm{U}(\mathcal{Y}, \mathcal{X})$. If it is the case that $n = \dim(\mathcal{X}) \leq \dim(\mathcal{Y})$ then the maximally entangled states are those states uu^* such that

$$\operatorname{Tr}_{\mathcal{Y}}(uu^*) = \frac{1}{n}\mathbb{1}_{\mathcal{X}},$$

or equivalently those that can be written

$$\frac{1}{n}\operatorname{vec}(U^*)\operatorname{vec}(U^*)^*$$

where $U \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$ is a linear isometry. Quite frequently, the term *maximally entangled* refers to the situation in which $\dim(\mathcal{X}) = \dim(\mathcal{Y})$, where the two notions coincide. As is to be expected when discussing pure states, we sometimes refer to a unit vector u as being a maximally entangled state, which means that uu^* is maximally entangled.

Now, for an arbitrary density operator $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$, let us define

$$M(\rho) = \max\{\langle uu^*, \rho\rangle \,:\, u \in \mathcal{X} \otimes \mathcal{Y} \text{ is maximally entangled}\}.$$

Clearly it holds that $0 < M(\rho) \leq 1$ for every density operator $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$. The following lemma establishes an upper bound on $M(\rho)$ based on the min-rank of ρ.

Lemma 17.1. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, and let $n = \min\{\dim(\mathcal{X}), \dim(\mathcal{Y})\}$. It holds that*

$$M(\rho) \leq \frac{\text{min-rank}(\rho)}{n}$$

for all $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$.

Proof. Let us assume $\dim(\mathcal{X}) \geq \dim(\mathcal{Y}) = n$, and note that the argument is equivalent in case the inequality is reversed.

First let us note that M is a convex function on $\mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$. To see this, consider any choice of $\sigma, \xi \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$ and $p \in [0,1]$. For $U \in \mathrm{U}(\mathcal{Y}, \mathcal{X})$ we have

$$\begin{aligned}
&\frac{1}{n} \operatorname{vec}(U)^* \left(p\sigma + (1-p)\xi\right) \operatorname{vec}(U) \\
&\qquad\qquad = p\, \frac{1}{n} \operatorname{vec}(U)^* \sigma \operatorname{vec}(U) + (1-p)\, \frac{1}{n} \operatorname{vec}(U)^* \xi \operatorname{vec}(U) \\
&\qquad\qquad \leq pM(\sigma) + (1-p)M(\xi).
\end{aligned}$$

Maximizing over all $U \in \mathrm{U}(\mathcal{Y}, \mathcal{X})$ establishes that M is convex as claimed.

Now, given that M is convex, we see that it suffices to prove the lemma by considering only pure states. Every pure density operator on $\mathcal{X} \otimes \mathcal{Y}$ may be written as $\operatorname{vec}(A)\operatorname{vec}(A)^*$ for $A \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ satisfying $\|A\|_2 = 1$. We have

$$M\left(\operatorname{vec}(A)\operatorname{vec}(A)^*\right) = \frac{1}{n} \max_{U \in \mathrm{U}(\mathcal{Y},\mathcal{X})} \left|\langle U, A\rangle\right|^2 = \frac{1}{n} \|A\|_1^2.$$

Given that

$$\|A\|_1 \leq \sqrt{\operatorname{rank}(A)}\, \|A\|_2$$

for every operator A, the lemma follows. □

17.2 Entanglement cost and distillable entanglement

We will now discuss two fundamental measures of entanglement: the *entanglement cost* and the *distillable entanglement*. For the remainder of the lecture, let us take

$$\mathcal{Y}_A = \mathbb{C}^{\{0,1\}} \quad \text{and} \quad \mathcal{Y}_B = \mathbb{C}^{\{0,1\}}$$

to be complex Euclidean spaces corresponding to single qubits, and let $\tau \in \mathrm{D}(\mathcal{Y}_A \otimes \mathcal{Y}_B)$ denote the density operator

$$\tau = \frac{1}{2}(e_0 \otimes e_0 + e_1 \otimes e_1)(e_0 \otimes e_0 + e_1 \otimes e_1)^*,$$

which may be more recognizable to some when expressed in the Dirac notation as

$$\tau = \left|\phi^+\right\rangle\left\langle\phi^+\right| \qquad \text{for} \qquad \left|\phi^+\right\rangle = \frac{1}{\sqrt{2}}\left|00\right\rangle + \frac{1}{\sqrt{2}}\left|11\right\rangle.$$

We view that the state τ represents one unit of entanglement, typically called an *e-bit* of entanglement.

17.2.1 Definition of entanglement cost

The first measure of entanglement we will consider is called the *entanglement cost*. Informally speaking, the entanglement cost of a density operator $\rho \in \mathrm{D}(\mathcal{X}_A \otimes \mathcal{X}_B)$ represents the number of e-bits Alice and Bob need to share in order to create a copy of ρ by means of an LOCC operation with high fidelity. It is an information-theoretic quantity, so it must be understood to be asymptotic in nature—where one amortizes over many parallel repetitions of such a conversion. The following definition states this more precisely.

Definition 17.2. The *entanglement cost* of a density operator $\rho \in \mathrm{D}(\mathcal{X}_A \otimes \mathcal{X}_B)$, denoted $E_c(\rho)$, is the infimum over all real numbers $\alpha \geq 0$ for which there exists a sequence of LOCC channels $\{\Phi_n : n \in \mathbb{N}\}$, where

$$\Phi_n \in \mathrm{LOCC}\left(\mathcal{Y}_A^{\otimes \lfloor \alpha n \rfloor}, \mathcal{X}_A^{\otimes n} : \mathcal{Y}_B^{\otimes \lfloor \alpha n \rfloor}, \mathcal{X}_B^{\otimes n}\right),$$

such that

$$\lim_{n \to \infty} \mathrm{F}\left(\Phi_n\left(\tau^{\otimes \lfloor \alpha n \rfloor}\right), \rho^{\otimes n}\right) = 1.$$

The interpretation of the definition is that, in the limit of large n, Alice and Bob are able to convert $\lfloor \alpha n \rfloor$ e-bits into n copies of ρ with high fidelity for any $\alpha > E_c(\rho)$. It is not entirely obvious that there should exist any value of α for which there exists a sequence of LOCC channels $\{\Phi_n : n \in \mathbb{N}\}$ as in the statement of the definition, but indeed there always does exist a suitable choice of α.

17.2.2 Definition of distillable entanglement

The second measure of entanglement we will consider is the *distillable entanglement*, which is essentially the reverse of the entanglement cost. It quantifies the number of e-bits that Alice and Bob can extract from the state in question, again amortized over many copies.

Definition 17.3. The *distillable entanglement* of a density operator $\rho \in \mathrm{D}(\mathcal{X}_A \otimes \mathcal{X}_B)$, denoted $E_d(\rho)$, is the supremum over all real numbers $\alpha \geq 0$ for which there exists a sequence of LOCC channels $\{\Phi_n : n \in \mathbb{N}\}$, where

$$\Phi_n \in \mathrm{LOCC}\left(\mathcal{X}_A^{\otimes n}, \mathcal{Y}_A^{\otimes \lfloor \alpha n \rfloor} : \mathcal{X}_B^{\otimes n}, \mathcal{Y}_B^{\otimes \lfloor \alpha n \rfloor}\right),$$

such that

$$\lim_{n \to \infty} \mathrm{F}\left(\Phi_n\left(\rho^{\otimes n}\right), \tau^{\otimes \lfloor \alpha n \rfloor}\right) = 1.$$

The interpretation of the definition is that, in the limit of large n, Alice and Bob are able to convert n copies of ρ into $\lfloor \alpha n \rfloor$ e-bits with high fidelity for any $\alpha < E_d(\rho)$. In order to clarify the definition, let us state explicitly that the operator $\tau^{\otimes 0}$ is interpreted to be the scalar 1, implying that condition in the lemma is trivially satisfied for $\alpha = 0$.

17.2.3 The distillable entanglement is at most the entanglement cost

At an intuitive level it is clear that the entanglement cost must be at least as large as the distillable entanglement, for otherwise Alice and Bob would be able to open an "entanglement factory" that would violate the principle that LOCC channels cannot create entanglement out of thin air. Let us now prove this formally.

Theorem 17.4. *For every state $\rho \in \mathrm{D}(\mathcal{X}_A \otimes \mathcal{X}_B)$ we have $E_d(\rho) \leq E_c(\rho)$.*

Proof. Let us assume that α and β are nonnegative real numbers such that the following two properties are satisfied:

1. There exists a sequence of LOCC channels $\{\Phi_n : n \in \mathbb{N}\}$, where
$$\Phi_n \in \mathrm{LOCC}\left(\mathcal{Y}_A^{\otimes \lfloor \alpha n \rfloor}, \mathcal{X}_A^{\otimes n} : \mathcal{Y}_B^{\otimes \lfloor \alpha n \rfloor}, \mathcal{X}_B^{\otimes n}\right),$$
such that
$$\lim_{n\to\infty} \mathrm{F}\left(\Phi_n\left(\tau^{\otimes \lfloor \alpha n \rfloor}\right), \rho^{\otimes n}\right) = 1.$$

2. There exists a sequence of LOCC channels $\{\Psi_n : n \in \mathbb{N}\}$, where
$$\Psi_n \in \mathrm{LOCC}\left(\mathcal{X}_A^{\otimes n}, \mathcal{Y}_A^{\otimes \lfloor \beta n \rfloor} : \mathcal{X}_B^{\otimes n}, \mathcal{Y}_B^{\otimes \lfloor \beta n \rfloor}\right),$$
such that
$$\lim_{n\to\infty} \mathrm{F}\left(\Psi_n\left(\rho^{\otimes n}\right), \tau^{\otimes \lfloor \beta n \rfloor}\right) = 1.$$

Using the Fuchs–van de Graaf inequalities, along with triangle inequality for the trace norm, we conclude that
$$\lim_{n\to\infty} \mathrm{F}\left((\Psi_n \Phi_n)\left(\tau^{\otimes \lfloor \alpha n \rfloor}\right), \tau^{\otimes \lfloor \beta n \rfloor}\right) = 1. \tag{17.1}$$
Because min-rank$(\tau^{\otimes k}) = 2^k$ for every choice of $k \geq 1$, and LOCC channels cannot increase min-rank, we have that
$$\mathrm{F}\left((\Psi_n \Phi_n)\left(\tau^{\otimes \lfloor \alpha n \rfloor}\right), \tau^{\otimes \lfloor \beta n \rfloor}\right)^2 \leq 2^{\lfloor \alpha n \rfloor - \lfloor \beta n \rfloor} \tag{17.2}$$
by Lemma 17.1. By equations (17.1) and (17.2), we therefore have that $\alpha \geq \beta$.

Given that $E_c(\rho)$ is the infimum over all α, and $E_d(\rho)$ is the supremum over all β, with the above properties, we have that $E_c(\rho) \geq E_d(\rho)$ as required. □

17.3 Pure state entanglement

The remainder of the lecture will focus on the entanglement cost and distillable entanglement for bipartite pure states. In this case, these measures turn out to be identical, and coincide precisely with the von Neumann entropy of the reduced state of either subsystem.

Theorem 17.5. *Let $\mathcal{X}_A$ and $\mathcal{X}_B$ be complex Euclidean spaces and let $u \in \mathcal{X}_A \otimes \mathcal{X}_B$ be a unit vector. It holds that $E_c(uu^*) = E_d(uu^*) = S(\mathrm{Tr}_{\mathcal{X}_A}(uu^*)) = S(\mathrm{Tr}_{\mathcal{X}_B}(uu^*))$.*

Proof. The proof will start with some basic observations about the vector u that will be used to calculate both the entanglement cost and the distillable entanglement. First, let
$$u = \sum_{a\in\Sigma} \sqrt{p(a)}\, v_a \otimes w_a$$
be a Schmidt decomposition of u, so that
$$\mathrm{Tr}_{\mathcal{X}_B}(uu^*) = \sum_{a\in\Sigma} p(a)\, v_a v_a^* \qquad \text{and} \qquad \mathrm{Tr}_{\mathcal{X}_A}(uu^*) = \sum_{a\in\Sigma} p(a)\, w_a w_a^*.$$

We have that $p \in \mathbb{R}^\Sigma$ is a probability vector, and $S(\operatorname{Tr}_{\mathcal{X}_A}(uu^*)) = H(p) = S(\operatorname{Tr}_{\mathcal{X}_B}(uu^*))$. Let us assume hereafter that $H(p) > 0$, for the case $H(p) = 0$ corresponds to the situation where u is separable (in which case the entanglement cost and distillable entanglement are both easily seen to be 0).

We will make use of concepts regarding compression that were discussed in Lecture 9. Recall that for each choice of n and $\varepsilon > 0$, we denote by $T_{n,\varepsilon} \subseteq \Sigma^n$ the set of ε-typical sequences of length n with respect to the probability vector p:

$$T_{n,\varepsilon} = \left\{ a_1 \cdots a_n \in \Sigma^n \,:\, 2^{-n(H(p)+\varepsilon)} < p(a_1) \cdots p(a_n) < 2^{-n(H(p)-\varepsilon)} \right\}.$$

For each $n \in \mathbb{N}$ and $\varepsilon > 0$, let us define a vector

$$x_{n,\varepsilon} = \sum_{a_1 \cdots a_n \in T_{n,\varepsilon}} \sqrt{p(a_1) \cdots p(a_n)}\, (v_{a_1} \otimes w_{a_1}) \otimes \cdots \otimes (v_{a_n} \otimes w_{a_n}) \in (\mathcal{X}_A \otimes \mathcal{X}_B)^{\otimes n}.$$

We have

$$\|x_{n,\varepsilon}\|^2 = \sum_{a_1 \cdots a_n \in T_{n,\varepsilon}} p(a_1) \cdots p(a_n),$$

which is the probability that a random choice of $a_1 \cdots a_n$ is ε-typical with respect to the probability vector p. It follows that $\lim_{n\to\infty} \|x_{n,\varepsilon}\| = 1$ for any choice of $\varepsilon > 0$. Let

$$y_{n,\varepsilon} = \frac{x_{n,\varepsilon}}{\|x_{n,\varepsilon}\|}$$

denote the normalized versions of these vectors.

Next, consider the vector of eigenvalues

$$\lambda\left(\operatorname{Tr}_{\mathcal{X}_B^{\otimes n}} \left(x_{n,\varepsilon} x_{n,\varepsilon}^* \right) \right).$$

The nonzero eigenvalues are given by the probabilities for the various ε-typical sequences, and so

$$2^{-n(H(p)+\varepsilon)} \;<\; \lambda_j \left(\operatorname{Tr}_{\mathcal{X}_B^{\otimes n}} \left(x_{n,\varepsilon} x_{n,\varepsilon}^* \right) \right) \;<\; 2^{-n(H(p)-\varepsilon)}$$

for $j = 1, \ldots, |T_{n,\varepsilon}|$. (The remaining eigenvalues are 0.) It follows that

$$\frac{2^{-n(H(p)+\varepsilon)}}{\|x_{n,\varepsilon}\|^2} \;<\; \lambda_j \left(\operatorname{Tr}_{\mathcal{X}_B^{\otimes n}} \left(y_{n,\varepsilon} y_{n,\varepsilon}^* \right) \right) \;<\; \frac{2^{-n(H(p)-\varepsilon)}}{\|x_{n,\varepsilon}\|^2}$$

for $j = 1, \ldots, |T_{n,\varepsilon}|$, and again the remaining eigenvalues are 0.

Let us now consider the entanglement cost of uu^*. We wish to show that for every real number $\alpha > H(p)$ there exists a sequence $\{\Phi_n \,:\, n \in \mathbb{N}\}$ of LOCC channels such that

$$\lim_{n\to\infty} \mathrm{F}\left(\Phi_n \left(\tau^{\otimes \lfloor \alpha n \rfloor} \right), (uu^*)^{\otimes n} \right) = 1. \tag{17.3}$$

We will do this by means of Nielsen's theorem. Specifically, let us choose $\varepsilon > 0$ so that $\alpha > H(p) + 2\varepsilon$, from which it follows that $\lfloor \alpha n \rfloor \geq n(H(p) + \varepsilon)$ for sufficiently large n. We have

$$\lambda_j \left(\operatorname{Tr}_{\mathcal{Y}_B^{\otimes \lfloor \alpha n \rfloor}} \left(\tau^{\otimes \lfloor \alpha n \rfloor} \right) \right) = 2^{-\lfloor \alpha n \rfloor}$$

for $j = 1, \ldots, 2^{\lfloor \alpha n \rfloor}$. Given that

$$\frac{2^{-n(H(p)+\varepsilon)}}{\|x_{n,\varepsilon}\|^2} \geq 2^{-n(H(p)+\varepsilon)} \geq 2^{-\lfloor \alpha n \rfloor},$$

for sufficiently large n, it follows that

$$\operatorname{Tr}_{\mathcal{Y}_B^{\otimes \lfloor \alpha n \rfloor}} \left(\tau^{\otimes \lfloor \alpha n \rfloor}\right) \prec \operatorname{Tr}_{\mathcal{X}_B^{\otimes n}} \left(y_{n,\varepsilon} y_{n,\varepsilon}^*\right). \tag{17.4}$$

This means that $\tau^{\otimes \lfloor \alpha n \rfloor}$ can be converted to $y_{n,\varepsilon} y_{n,\varepsilon}^*$ by means of an LOCC channel Φ_n by Nielsen's theorem. Given that

$$\lim_{n\to\infty} \mathrm{F}\left(y_{n,\varepsilon} y_{n,\varepsilon}^*, (uu^*)^{\otimes n}\right) = 1$$

this implies that the required equation (17.3) holds. Consequently $E_c(uu^*) \leq H(p)$.

Next let us consider the distillable entanglement, for which a similar argument is used. Our goal is to prove that for every $\alpha < H(p)$, there exists a sequence $\{\Psi_n : n \in \mathbb{N}\}$ of LOCC channels such that

$$\lim_{n\to\infty} \mathrm{F}\left(\Psi_n\left((uu^*)^{\otimes n}\right), \tau^{\otimes \lfloor \alpha n \rfloor}\right) = 1. \tag{17.5}$$

In this case, let us choose $\varepsilon > 0$ small enough so that $\alpha < H(p) - 2\varepsilon$. For sufficiently large n we have

$$2^{\lfloor \alpha n \rfloor} \leq \|x_{n,\varepsilon}\|^2 \, 2^{n(H(p)-\varepsilon)}.$$

Similar to above, we therefore have

$$\operatorname{Tr}_{\mathcal{X}_B^{\otimes n}} \left(y_{n,\varepsilon} y_{n,\varepsilon}^*\right) \prec \operatorname{Tr}_{\mathcal{Y}_B^{\otimes \lfloor \alpha n \rfloor}} \left(\tau^{\otimes \lfloor \alpha n \rfloor}\right),$$

which implies that the state $y_{n,\varepsilon} y_{n,\varepsilon}^*$ can be converted to the state $\tau^{\otimes \lfloor \alpha n \rfloor}$ by means of an LOCC channel Ψ_n for sufficiently large n. Given that

$$\lim_{n\to\infty} \left\|(uu^*)^{\otimes n} - y_{n,\varepsilon} y_{n,\varepsilon}^*\right\|_1 = 0$$

it follows that

$$\begin{aligned}&\lim_{n\to\infty} \left\|\Psi_n\left((uu^*)^{\otimes n}\right) - \tau^{\otimes \lfloor \alpha n \rfloor}\right\|_1 \\ &\qquad \leq \lim_{n\to\infty} \left(\left\|\Psi_n\left((uu^*)^{\otimes n}\right) - \Psi_n\left(y_{n,\varepsilon} y_{n,\varepsilon}^*\right)\right\|_1 + \left\|\Psi_n\left(y_{n,\varepsilon} y_{n,\varepsilon}^*\right) - \tau^{\otimes \lfloor \alpha n \rfloor}\right\|_1\right) = 0,\end{aligned}$$

which establishes the above equation (17.5). Consequently, $E_d(uu^*) \geq H(p)$.

We have shown that

$$E_c(uu^*) \leq H(p) \leq E_d(uu^*).$$

As $E_d(uu^*) \leq E_c(uu^*)$, the equality in the statement of the theorem follows. □