Freie Universität Berlin
Tutorials on Quantum Information Theory
Summer Term 2018

Problem Sheet 7
Transforming and quantifying entanglement

Discussed in Tutorial: 14/06/2018 J. Eisert, D. Hangleiter, I. Roth

1. **Local operations and classical communication (LOCC).**

   At the heart of entanglement theory lies the notion of LOCC. To see why, imagine two parties that are a large distance apart from each other, say, Alice is in Berlin and Bob in New York. While they may obtain access to shared entanglement from a third party, it is unreasonable to assume that they are able to perform global operations on the state they share. On the other hand, it is perfectly conceivable that they transmit classical messages, for example, to communicate measurement results.

   In the LOCC paradigm, each party is allowed to measure and perform unitary operations on their part of the shared state, and communicate via a classical channel. While general LOCC protocols may involve several rounds of interaction (so Alice does something, sends the result to Bob, then Bob does something, who then communicates with Alice, and so forth) it is often sufficient to consider only a single round of interaction.

   So suppose, Alice and Bob share a state $|\psi\rangle$ with Schmidt decomposition $|\psi\rangle = \sum_i \sqrt{\lambda_i} \left|i^{(A)}\right\rangle \left|i^{(B)}\right\rangle$. We will now show that any measurement $\{M_j\}_j$ on Bob's side can be simulated as follows: Alice performs a measurement $\{N_j\}_j$ on her side, sends the result to Bob, who applies a corresponding unitary transformation.

   a) Expand M_j in the Schmidt basis $\{\left|i^{(B)}\right\rangle\}_i$ and define the measurement operator N_j in terms of the expansion coefficients. Determine the post-measurement state of Bob $|\psi_j\rangle$ (who performs $\{M_j\}$), and of Alice $|\phi_j\rangle$ (who performs $\{N_j\}$).

      Solution: If we want to get the result j with the same probability for Bob measuring $\{M_j\}$ or Alice measuring $\{N_j\}$, the measurement operators of both two should have the same expansion coefficients in their respective Schmidt basises, i.e. $M_j = \sum_{ik} M_j^{(ik)} \left|i^{(B)}\right\rangle \left\langle k^{(B)}\right|$ and $N_j = \sum_{ik} M_j^{(ik)} \left|i^{(A)}\right\rangle \left\langle k^{(A)}\right|$.

      Then, $p_j^A = \sum_{ik} \lambda_i |M_j^{(ik)}|^2 = p_j^B$. The post-measurement state of Bob is given by $(\mathbb{1}\otimes M_j) |\psi\rangle = \sum_{ikl} \sqrt{\lambda_l} M_j^{(ik)} (\mathbb{1}\otimes \left|i^{(B)}\right\rangle\!\left\langle k^{(B)}\right|) \left|l^{(A)}\right\rangle \left|l^{(B)}\right\rangle = \sum_{ik} \sqrt{\lambda_k} M_j^{(ik)} \left|k^{(A)}\right\rangle \left|i^{(B)}\right\rangle$. Similarly, Alice post-measurement state is $(N_j \otimes \mathbb{1}) |\psi\rangle = \sum_{ik} \sqrt{\lambda_k} M_j^{(ik)} \left|i^{(A)}\right\rangle \left|k^{(B)}\right\rangle$.

   b) Show that $|\phi_j\rangle$ is local-unitary equivalent to $|\psi_j\rangle$.

      Solution: Both post-measurement state have the same Schmidt-decomposition and are, thus, related by a local unitary transformation.

   c) Summarise the LOCC protocol.

2. **Majorisation and transforming quantum states by local unitaries.**

   In this problem we will look at the task of transforming a given copy of a pure bipartite quantum state $|\psi\rangle$ to another quantum state $|\phi\rangle$ using LOCC. The question is: Under which conditions is the transition $|\psi\rangle \xrightarrow{\text{LOCC}} |\phi\rangle$ possible?

   The key to the answer of this question is the concept of majorisation. We say that a real vector $x \in \mathbb{R}^n$ majorises $y \in \mathbb{R}^n$ ($x \succ y$) if for all $k = 1, \dots, n$ $\sum_{j=1}^k x_j^\downarrow \geq \sum_{j=1}^k y_j^\downarrow$ and

$\sum_{j=1}^n x_j^\downarrow = \sum_{j=1}^d y_j^\downarrow$. Here, $x^\downarrow$ denotes the sorted version of x, i.e., a permutation of the elements of x such that $x_1^\downarrow \geq x_2^\downarrow \geq \ldots \geq x_n^\downarrow$. So from now on, let $\sum_{j=1}^n x_j^\downarrow = \sum_{j=1}^d y_j^\downarrow$

a) Show that $x \succ y$ if and only if for all $t \in \mathbb{R}$

$$\sum_{j=1}^n \max(x_j - t, 0) \geq \sum_{j=1}^n \max(y_j - t, 0)$$

(and equality for $t = 0$).

Solution: Assume $x \succ y$. First, for $t \geq y_0^\downarrow$ the RHS vanishes and the inequality holds. If we choose $t \in [y_{k+1}^\downarrow, y_k^\downarrow)$, then

$$\sum_{j=1}^n \max\{y_j - t, 0\} = \sum_{j=1}^k y_j - kt \leq \sum_{j=1}^k x_j - kt \leq \sum_{j=1}^n \max\{x_j - t, 0\}. \qquad (1)$$

Now the last possibility is $t \leq y_n^\downarrow$ then the RHS is $\sum_{j=1}^n y_j - nt = \sum_{j=1}^n x_j - nt \leq \sum_{j=1}^n \max\{x_j - t, 0\}$.

Now, the converse direction: We assume that the inequality holds and choose $t = x_{k+1}^\downarrow$

$$\sum_{j=1}^k y_j - kt \leq \sum_{j=1}^k \max\{y_j - t, 0\} \leq \sum_{j=1}^n \max\{y_j - t, 0\} \qquad (2)$$

$$\leq \sum_{j=1}^n \max\{x_j - t, 0\} = \sum_{j=1}^k x_j - kt, \qquad (3)$$

from which majorisation follows.

b) Use the characterisation from (a) to show that the set $\{x : x \prec y\}$ is convex for any given y.

Solution: Let $x_1, x_2 \prec y$ and $\lambda \in (0,1)$. The map $\mathbb{R} \to \mathbb{R}$, $a \mapsto \max\{x - t, 0\}$ is convex for all t. Thus, also $x \mapsto \sum_{i=1}^n \max\{x_j - t, 0\}$ is convex. Then, for all t

$$\sum_{j=1}^n \max\{\lambda(x_1)_j + (1-\lambda)(x_2)_j - t, 0\} \qquad (4)$$

$$\leq \lambda \sum_{j=1}^n \max\{(x_1)_j - t, 0\} + (1-\lambda) \sum_{j=1}^n \max\{(x_2)_j - t, 0\} \qquad (5)$$

$$\leq \sum_{j=1}^n \max\{y_j - t, 0\}. \qquad (6)$$

Furthermore, $\sum_{j=1}^n [\lambda x_1 + (1-\lambda)x_2] = \sum_{j=1}^n y_j$ by linearity. In conclusion, we have established that $\lambda x_1 + (1-\lambda)x_2 \prec y$.

One can now show that $x \prec y$ if and only if $x = \sum_j p_j \Pi_j y$ for a probability distribution p and permutation matrices Π_j. By Birkhoff's theorem, which lies at the heart of majorisation theory, that statement is equivalent to saying that $x \prec y$ if and only if $x = Dy$ for some doubly stochastic matrix D[1].

For two Hermitian operators $X, Y \in L(\mathbb{C}^d)$ we say that $X \prec Y$ if $\mathrm{spec}(X) \prec \mathrm{spec}(Y)$.

[1] A matrix D is called doubly stochastic if $\forall i, j D_{ij} \geq 0$ and $\forall i \sum_j D_{ij} = \sum_j D_{ji} = 1$, i.e., all rows and columns are probability distributions.

c) Show that $X \prec Y$ if and only if there exists a probability distribution p and unitary matrices U_j such that

$$X = \sum_j p_j U_j Y U_j^\dagger.$$

Solution: First, we assume $X \prec Y$. We denote the eigenvalue decomposition of X and Y by $X = U\Lambda_X U^\dagger$ and $Y = V\Lambda_Y V^\dagger$, respectively. Since $\lambda_X \prec \lambda_Y$, there exists a probability distribution p and permutation matrices Π_j, such that $\lambda_X = \sum_j p_j \Pi_j \lambda_y$. Correspondingly, we have $\Lambda_X = \sum_j p_j \Pi_j \Lambda_Y \Pi_j^\dagger$, which may be seen from writing down Λ_X in coordinates. Thus, $X = U\Lambda_X U^\dagger = \sum_j p_j U \Pi_j \Lambda_Y \Pi_j^\dagger U^\dagger = \sum_j p_j U \Pi_j V^\dagger Y V \Pi_j^\dagger U^\dagger = \sum_j p_j U_j Y U_j^\dagger$ with $U_j = U\Pi_j V^\dagger$.

Conversely, if $X = \sum_j p_j U_j Y U_j^\dagger$, then $\Lambda_X = \sum_j p_j U^\dagger U_j V \Lambda_Y V^\dagger U_j U$. Thus, defining new unitary matrices $V_j = U^\dagger U_j V$, we have $(\lambda_X)_i = \sum_{jk} p_j (V_j)_{ik} (\lambda_Y)_k (V_j^\dagger)_{ki} = \sum_{jk} p_j |(V_j)_{ik}|^2 (\lambda_X)_k$. We arrive at $\lambda_X = D\lambda_Y$ with $D_{ik} = \sum_j p_j |(V_j)_{ik}|^2$. Obviously, $D_{ik} \geq 0$ for all i, k. Furthermore, since the columns and rows of a unitary matrix V_j have unit ℓ_2-norm, it holds that $\sum_i D_{ik} = \sum_{ij} p_j |(V_j)_{ik}|^2 = \sum_j p_j = 1$ and, analogously, $\sum_k D_{ik} = 1$. Altogether, we see that D is doubly stochastic. Therefore, by Birkhoff's theorem we have $\lambda_X \prec \lambda_Y$ and, thus, $X \prec Y$.

We are now ready to prove the main theorem:

Theorem 1 $|\psi\rangle \xrightarrow{LOCC} |\phi\rangle$ *if and only if* $\mathrm{Tr}_B[|\psi\rangle\langle\psi|] \prec \mathrm{Tr}_B[|\phi\rangle\langle\phi|]$.

d) Show the forward direction using the result of Problem 1. You may assume that the transformation is effected by a measurement on Alice's side and a corresponding unitary on Bob's side. In other words, from Alice's point of view it must be the case that [2]

$$M_j \, \mathrm{Tr}_B[|\psi\rangle\langle\psi|] M_j^\dagger = p_j \, \mathrm{Tr}_B[|\phi\rangle\langle\phi|].$$

Hint: Use the polar decomposition of $M_j \sqrt{\mathrm{Tr}_B[|\psi\rangle\langle\psi|]}$.

Solution: Let's call $\rho_\psi = \mathrm{Tr}_B[|\psi\rangle\langle\psi|]$ and ρ_ϕ, similarly. Using the polar decomposition $M_j\sqrt{\rho_\psi} = |M_j\sqrt{\rho_\psi}|V_j = \sqrt{p_j\rho_\phi}V_j$. Now,

$$\rho_\psi = \sum_j \sqrt{\rho_\psi} M_j^\dagger M_j \sqrt{\rho_\psi} = \sum_j p_j V_j^\dagger \rho_\phi V_j, \tag{7}$$

from which $\rho_\psi \prec \rho_\phi$ follows from the assertion of (c).

e) Now show the backward direction by proceeding analogously.

Solution: Run the proof of (d) backwards: By assumption and (c), we know the existence of V_js and p_j. Define the measurement operators such that $M_j\sqrt{\rho_\psi} = \sqrt{p_j\rho_\phi}V_j$ and check the completeness. *It is not obvious that this definiton is possible in general. So one actually has to work harder here having a closer look at the respective ranges of the states ...*

Thus, we have an LOCC scheme doing the job.

[2] This is because the transition from $|\psi\rangle$ to $|\phi\rangle$ comes about as a post-measurement state with probability p_j.

3. **Distilling and diluting entanglement.**

Now instead of being supplied with a single copy of an entangled state $|\psi\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$ Alice and Bob have access to a large number of copies $|\psi\rangle^{\otimes m}$. We now ask two questions (that were already asked in the lecture): (1) How many copies of the Bell state $(|00\rangle + |11\rangle)/\sqrt{2}$, or *ebits* can be 'distilled' from $|\psi\rangle^{\otimes m}$? (2) Into how many 'less entangled states' $|\phi\rangle$ can $|\psi\rangle$ be diluted?

To begin with, recall the definition of ϵ-typical sequences: Given $\epsilon > 0$, a sequence $x = (x_1, x_2, \ldots, x_n)$ is called ϵ-typical with respect to a distribution p if

$$2^{-n(H(p)+\epsilon)} \leq p(x_1)\cdots p(x_n) \leq 2^{-n(H(p)-\epsilon)}, \tag{8}$$

where $H(p) = -\sum_i p_i \log p_i$ is the Shannon entropy of p. Denote by $T(\epsilon, n)$ the set of length-n ϵ-typical sequences with respect to p. Also, recall the theorem of ϵ-typical sequences:

Theorem 2 *(i) Let* $\epsilon > 0$. *Then for any* $\delta > 0\, \exists n \in \mathbb{N} : \Pr_{x_1,\ldots,x_n\sim p}[(x_1, \ldots, x_n) \in T(\epsilon, n)] \geq 1 - \delta$.

(ii) $\forall \epsilon, \delta > 0\, \exists n \in \mathbb{N} : (1-\delta)2^{n(H(p)-\epsilon)} \leq |T(\epsilon, n)| \leq 2^{n(H(p)+\epsilon)}$

We will now apply this theorem to the problem of diluting and distilling entanglement from $|\psi\rangle$. To this end, suppose that $|\psi\rangle = \sum_x \sqrt{p(x)}\, |x_A\rangle\, |x_B\rangle$ in Schmidt decomposition. Moreover, define by $|\phi_m\rangle$ the state obtained from $|\psi\rangle^{\otimes m}$ by omitting all terms that are not ϵ-typical and renormalising.

a) Show that the number of terms in $|\phi_m\rangle$ is at most $2^{m(S(\rho_\psi)+\epsilon)}$ where $\rho_\psi = \mathrm{Tr}_B[|\psi\rangle\langle\psi|]$.

Solution: We have that

$$|\psi\rangle^{\otimes m} = \sum_{x_1,\ldots,x_m} \sqrt{p(x_1)\cdots p(x_m)}\, |x_{A,1}\cdots x_{A,m}\rangle\, |x_{B,1}\cdots x_{B,m}\rangle$$

from which we obtain $|\phi_m\rangle$ by omitting all basis states indexed by strings that are not ϵ-typical, i.e., strings $x_1, \ldots, x_m$ for which Eq. (8) holds. This yields an unnormalized state

$$|\phi'_m\rangle = \sum_{(x_1,\ldots,x_m)\,\epsilon-\text{typical}} \sqrt{p(x_1)\cdots p(x_m)}\, |x_{A,1}\cdots x_{A,m}\rangle\, |x_{B,1}\cdots x_{B,m}\rangle\,,$$

which we normalize to obtain $|\phi_m\rangle = |\phi'_m\rangle / \sqrt{\langle\phi'_m \mid \phi'_m\rangle}$. By Thm. 2 (ii) the number of terms is upper bounded by $2^{m(H(p)+\epsilon)}$. But on the other hand, we have that

$$|\psi\rangle\langle\psi| = \sum_{x,y} \sqrt{p(x)p(y)}\, |x_A\rangle\, |x_B\rangle\, \langle y_A|\, \langle y_B|$$

and hence

$$\mathrm{Tr}_B[|\psi\rangle\langle\psi|] = \sum_{x,y} \sqrt{p(x)p(y)}\, \langle x_B|\, |x_A\rangle\, |x_B\rangle\, \langle y_A|\, \langle y_B|\, |x_B\rangle = \sum_x p(x)\, |x_A\rangle\langle x_A|\,.$$

Since ρ_ψ is diagonal in the Schmidt basis we have that $S(\rho_\psi) = H(p)$ and the claim follows.

Let us now look at the following entanglement dilution protocol: Alice and Bob share n Bell states. Alice locally prepares $|\phi_m\rangle$ and teleports one half of $|\phi_m\rangle$ to Bob.

b) How many Bell states are required such that after the dilution protocol Alice and Bob share $|\phi_m\rangle$.

Solution: Recall from quantum teleportation that a qubit can be teleported using one bell state. Correspondingly, a state $|\psi\rangle \in \mathbb{C}^{2n}$ having an n-qubit representation needs n Bell states for telportation.

From part a) we know, that $|\phi\rangle$ is in a subspace $\mathcal{C} \subset \mathbb{C}^{2m}$ of dimension $d_{\mathsf{eff}} = 2^{m(S(\rho_\psi)+\epsilon)}$, i.e. $\mathcal{C} \cong \mathbb{C}^{d_{\mathsf{eff}}}$. Thus, we need $\lceil m(S(\rho_\psi)+\epsilon)\rceil$ Bell states for teleportation.

An explicit representation of $|\phi\rangle$ in a computational basis of the smaller vector space $\mathcal{C}$ can be achieved by performing Schumacher compression.

c) Use Theorem 2 (i) to find a lower bound on the fidelity $|\langle\phi_m \mid \psi\rangle^{\otimes m}|^2$ for a suitably chosen m.

Solution: The theorem gives us a lower bound on the probability that any given string chosen according to p is ϵ-typical. Rewriting this condition, we have that

$$\Pr[(x_1,\ldots,x_m) \in T(\epsilon,m)] = \sum_{x_1,\ldots,x_m\ \epsilon-\text{typical}} p(x_1)\cdots p(x_m) \geq 1-\delta$$

But on the other hand

$$|\langle\phi_m \mid \psi\rangle^{\otimes m}|^2 = \frac{1}{\langle\phi'_m \mid \phi'_m\rangle}|\langle\phi'_m \mid \psi\rangle^{\otimes m}|^2 = \langle\phi'_m \mid \phi'_m\rangle$$
$$= \sum_{x_1,\ldots,x_m\ \epsilon-\text{typical}} p(x_1)\cdots p(x_m),$$

which shows the claim.

An entanglement concentration protocol proceeds along similar lines: Suppose Alice and Bob share n copies of $|\psi\rangle$. Alice performs a measurement that projects onto the ϵ-typical subspace of $|\psi\rangle$ to convert $|\psi\rangle^{\otimes m}$ into $|\phi_m\rangle$.

d) Determine an upper bound on the Schmidt coefficients of $|\phi_m\rangle$.

Solution: By the definition and Theorem 2 (i) we have the upper bound $2^{-m(S(\rho_\psi)-\epsilon)}/(1-\delta)$.

e) Determine n as a function of m such that the state $|\phi_m\rangle$ obtained from the projective measurement can be transformed into n Bell pairs using LOCC.

Solution: Choose n such that

$$\frac{2^{-m(S(\rho_\psi)-\epsilon)}}{1-\delta} \leq 2^{-n}$$

Then the vector of eigenvalues of $\mathrm{Tr}_B[|\phi_m\rangle\langle\phi_m|]$ is majorized by $(2^{-n},\ldots,2^{-n})$ and by the previous exercise there exists an LOCC-protocol transforming $|\phi_m\rangle$ into $|\omega\rangle^{\otimes n}$, recalling that the Schmidt coefficients of $|\omega\rangle^{\otimes n}$ are just given by 2^{-n}

f) Show that the scaling of resources required for the distillation procedure is optimal.

Hint: Argue by means of a contradiction.

Solution: Assume it was possible to distil more than $n \approx mS(\rho_\psi)$ many ebits from $|\phi_m\rangle$. Then one could take $S(\rho_\psi)$ many Bell states, dilute them into a copy of $|\psi\rangle$ and from $|\psi\rangle$ distil $S > S(\rho_\psi)$ many Bell states again by means of LOCC. But one can convince oneself that under LOCC the Schmidt number, i.e. the rank of the reduced density matrix, is preserved.

There are many more interesting issues that arise in the context of entanglement transformation, some of which we want to mention here but cannot go into here[3].

(i) There exist entangled states that cannot be distilled. These are precisely the positive-partial-transpose entangled states (Horodecki et al., 1998).

(ii) Some transformations between quantum states $|\psi\rangle$ and $|\phi\rangle$ using LOCC become possible only through a so-called catalyst state $|c\rangle$, that is, while $|\psi\rangle \xrightarrow{\text{LOCC}} |\phi\rangle$ is impossible, there exists a state $|c\rangle$ such that $|\psi\rangle |c\rangle \xrightarrow{\text{LOCC}} |\phi\rangle |c\rangle$ is possible.

(iii) For the original paper on concentrating entanglement see (Bennett et al., 1996).

References

Bennett, C. H., H. J. Bernstein, S. Popescu, and B. Schumacher (1996, April). Concentrating partial entanglement by local operations. *Phys. Rev. A 53*(4), 2046–2052.

Horodecki, M., P. Horodecki, and R. Horodecki (1998, June). Mixed-State Entanglement and Distillation: Is there a "Bound" Entanglement in Nature? *Physical Review Letters 80*(24), 5239–5242.

[3]... because the sheet is already a bit long...