Module II, Part 2

Entropy, information and entanglement

PH 534 QIC

Himadri Shekhar Dhar
himadri.dhar@iitb.ac.in

B. Entanglement, majorization and Nielsen’s theorem

In this section we begin by briefly discussing the origin story of entanglement (and quantum correlations in general) before going on to ask the more fundamental question about what is entanglement. We then try to come up with a more operational approach to define the all-conquering physical entity in quantum physics.

i) Background

The first instance of confrontation with the classical viewpoint of physics, was raised by Einstein, Podolsky and Rosen (EPR), who were troubled by the violation of the conjunction of “objective reality” and “locality” in the quantum description of a physical system with spatially separated subsystems, as mentioned in their seminal paper of 1935[1]. The main contention was the presence of weird and spooky quantum correlations, and in general they proposed that quantum mechanics was incomplete.

Schrödinger often discussed the weirdness of quantum theory and quantum correlations in general with a very philosophical tone. In different texts, he touched upon different aspects of nonclassicality:

In this statement, “…like a scholar in an examination, cannot possibly know which of the two questions I am going to ask first: it so seems that our scholar is prepared to give the right answer to the first question… Therefore, he must know both answers; which is an amazing knowledge,” he touches upon the notion of nonlocality.

In another, “It is rather discomforting that the theory should allow a system to be steered or piloted into one or the other type of state at the experimenter's mercy in spite of his having no access to it,” he hints at what would later be termed as the phenomena of quantum steering.

[1] A. Einstein, B. Podolsky, and N. Rosen, *Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?* Phys. Rev. **47**, 777 (1935).

And finally, “When two systems… enter into temporary physical interaction due to known forces between them… they can no longer be described … by endowing each of them with a representative of its own. I would not call that one but rather the characteristic trait of quantum mechanics,” where he finally calls upon entanglement.

To reiterate, Schrödinger called the “characteristic trait of quantum mechanics” being the complete description of a composite system without providing all the information about its subsystems, and originally referred to this “Verschränkung” or entanglement

The signature of nonlocal quantum correlations was first quantified through the seminal derivation of Bell inequalities. In 1964, John Bell showed that for all theoretical description of quantum mechanics, which additionally account for “objective reality and locality” by means of some “hidden variable”, all bipartite correlations must be statistically constrained by a set of inequalities. Further, he pointed out that certain quantum states did not satisfy these inequalities. In the following years, experiments demonstrated that quantum states can violate these Bell inequalities, thus confirming the impossibility of using only “objectively real and local” or local hidden-variable description of quantum phenomena[2]. Over the years, important theoretical and experimental results have supported and enriched the quantum viewpoint of the physical world and established quantum theory as one of the foundational cornerstones of modern physics. Significantly, the enigmatic trait of the world arising due to quantum correlations is at the heart of major technological developments in the 21st Century and is the fundamental resource for quantum information processing.

The violation of Bell inequalities led to a critical interest in quantum correlations for future development of concepts such as quantum communication and the possibility of developing computational devices with no classical analogue. However, it was not until the late 20th Century, that the quantum correlation, in its quintessential form of entanglement, was established in terms of local quantum operations and classical

[2] The notion of nonlocal hidden variables to describe quantum mechanics was first proposed by David Bohm in his now seminal work. But by invoking nonlocality it essentially violates EPR’s original argument against the lack of “objective reality” and “locality” in quantum mechanics.

operations. In subsequent years, various theoretical approaches for studying entanglement, such as inequalities derived from asymptotic rates of information compression, distillation of entangled states, majorization conditions, witnesses and resource theories were introduced and studied[3].

ii) Entanglement

The fundamental basis of what constitutes nonclassicality or how quantum correlations are conceptually formulated is not limited to a single theoretical framework. In a broad sense, nonclassicality arises when composite physical systems or degrees of freedom are correlated in ways that are inaccessible to classical objects. The earliest forms of quantifiable nonclassicality in two-party or bipartite quantum states (bipartite quantum states) arose from the violation of Bell inequality. The conceptualization of quantum correlation has evolved over the years. The fundamental idea that violation of Bell inequality was the key principle of defining quantum correlation suffered a setback when it was shown that there exist certain entangled states that do not violate the Bell inequalities and hence, violation of Bell inequality is a sufficient but not necessary condition for entanglement. This led to the realization that states without entanglement are those which can be prepared using *local quantum operations and classical communication* (LOCC). The set of states that cannot be prepared using LOCC are then called entangled. Consequently the set of separable states is smaller than the set of states that do not violate Bell inequalities.

iii) Local operations and classical communication (LOCC)

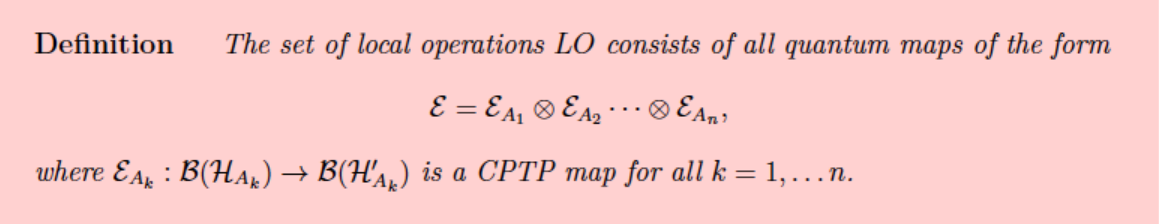
Note: This subsection is not necessary for exams.

a) Local operations (LO):

Let us consider the bipartite case, with Aditi and Bharat sharing a state located at two spatially distant points. The joint system is described by the composite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, and the class of possible local operations are described by $\mathcal{E} = \mathcal{E}_A \otimes \mathcal{E}_B$,

[3] Although we will not cover all these topics in the course, I will provide additional material for those who are interested in learning a bit more. Of course, excellent books and papers exist on the Internet.

where $\mathcal{E}_A$ is a CPTP map applied by Aditi locally on her system, and $\mathcal{E}_B$ is a CPTP operator applied by Bharat locally on his state. Here, we note that addition of an ancilla state or tracing out of a state are valid CPTP maps that can be applied locally by both Aditi and Bharat.



Definition *The set of local operations LO consists of all quantum maps of the form*

$$\mathcal{E} = \mathcal{E}_{A_1} \otimes \mathcal{E}_{A_2} \cdots \otimes \mathcal{E}_{A_n},$$

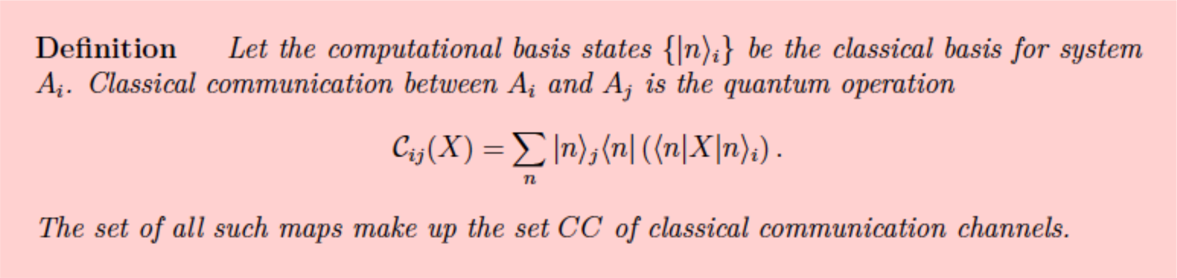
where $\mathcal{E}_{A_k} : \mathcal{B}(\mathcal{H}_{A_k}) \to \mathcal{B}(\mathcal{H}'_{A_k})$ *is a CPTP map for all* $k = 1, \ldots n.$

b) Classical communication (CC):

Classical communication is a bit more subtle. What we term as classical information is essentially any information that can be encoded in some orthonormal basis (say the computation basis) with no superpositions. In the density matrix language, this looks like a diagonal matrix in an orthonormal basis or the eigenbasis. This is deemed classical because the orthogonality of states gives information a strong classical flavour. For instance, we know that information communicated using orthogonal states are perfectly distinguishable. Take the example of qubits: in the absence of coherent superpositions (or off-diagonal terms), the states $|0\rangle$ and $|1\rangle$, behave for all purposes like classical bits 0 and 1, even if they represent the spin of a single atom. This is a very different perspective than how we think of quantumness in other disciplines of physics, say condensed matter or many-body physics. What is termed as classical or quasi-classical in quantum information theory are quantities that have clear classical analogue, even though the systems on which they are encoded can be microscopic and quantum in nature. On the other hand, quantum features, such as coherence and entanglement, have no classical analogue but may in principle arise in large macroscopic quantum systems.[4]

[4] One may argue that coherence surely arises in classical theory of waves, say electromagnetic (EM) theory. To place a simplified argument here, quantum mechanics actually puts the classical EM theory and the first quantization due to Schrodinger's equation on an equal footing, giving us the first step of the famous but often misunderstood "wave-particle" duality. So quantum mechanics in first quantization forces you to think of everything as waves and coherence. In the second step, EM fields are quantized to get the notion of photons, and you now have the second quantization in quantum field theory, which gives rise to the interaction of relativistic particles. Now everything is described in some sort of particle-like objects.

So, getting back to classical communication, we can think of it as a kind of re-labelling of who owns the information stored in an orthonormal basis. Say, Aditi has a classical information stored in her register or orthogonal state $|k\rangle\langle k|_A$, which she simply communicates to Bharat, who then owns $|k\rangle\langle k|_B$. In an operational sense, classical communication typically implies that Bharat can coordinate any local operation on his state conditioned on some local operation done by Aditi.



Definition *Let the computational basis states $\{|n\rangle_i\}$ be the classical basis for system A_i. Classical communication between A_i and A_j is the quantum operation*

$$\mathcal{C}_{ij}(X) = \sum_n |n\rangle_j\langle n| \left(\langle n|X|n\rangle_i\right).$$

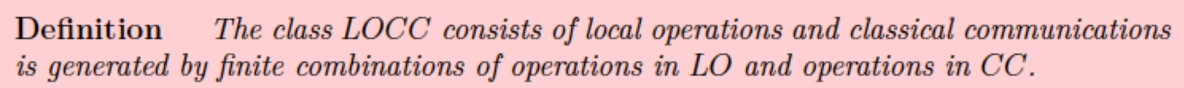
The set of all such maps make up the set CC of classical communication channels.

c) LOCC:

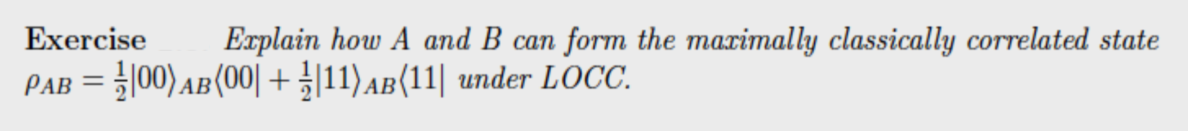
So LOCC operations typically combine and coordinate all local operations with some classical communication. The standard LOCC for bipartite pure systems can be reduced to the following: Aditi does some measurement $\mathcal{M} = \{M_i\}$ on her system, with outcomes $\{m_i\}$. She communicates the value of m_i to Bharat, who then performs a unitary transformation U_i on his system. Mathematically we can write this as:

$$\rho_{AB} \rightarrow \sum_i M_i \otimes \mathbb{I}\, \rho_{AB}\, M_i^{\dagger} \otimes \mathbb{I} \rightarrow \sum_i (\mathbb{I} \otimes U_i)\,(M_i \otimes \mathbb{I}\,)\rho_{AB}\,\left(M_i^{\dagger} \otimes \mathbb{I}\right)\left(\mathbb{I} \otimes U_i^{\dagger}\right)$$

$$= \sum_i (M_i \otimes U_i\,)\, \rho_{AB}\, \left(M_i^{\dagger} \otimes U_i^{\dagger}\right)$$



Definition *The class LOCC consists of local operations and classical communications is generated by finite combinations of operations in LO and operations in CC.*



Exercise *Explain how A and B can form the maximally classically correlated state $\rho_{AB} = \frac{1}{2}|00\rangle_{AB}\langle 00| + \frac{1}{2}|11\rangle_{AB}\langle 11|$ under LOCC.*

iv) Transforming entangled states using LOCC

Now that we have discussed what is LOCC and briefly touched upon its connection to entanglement, we look at how quantum states can be transformed using LOCC protocols. In particular, we seek to delve deeper into the connection between LOCC and entanglement. A general study of LOCC protocols in entangled states is a very difficult proposition, so we stick to the simple case of bipartite pure quantum states shared between Aditi and Bharat (say, $|\psi\rangle_{AB}$), and we ask under what conditions can one use LOCC to transform, $|\psi\rangle_{AB} \to |\phi\rangle_{AB}$.

Look at the following example (please check the calculations):

Example: Suppose we want to convert the state,

$$|\psi\rangle_{AB} = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle),$$

to the state, $|\phi\rangle_{AB} = \cos\theta\,|00\rangle + \sin\theta\,|11\rangle$, using LOCC in a deterministic manner.

Let, Aditi perform measurement, $\mathcal{M} = \left\{M_1 = \begin{pmatrix} \cos\theta & 0 \\ 0 & \sin\theta \end{pmatrix};\ M_1 = \begin{pmatrix} 0 & \cos\theta \\ \sin\theta & 0 \end{pmatrix}\right\}$. Depending on her measurement, she classically communicates her outcome to Bharat classically, who action is conditioned upon inputs from Aditi.

If Aditi performs: $(M_1 \otimes \mathbb{I})|\psi\rangle_{AB} \to |\tilde{\psi}\rangle_{AB} = \begin{pmatrix} \cos\theta & 0 & 0 & 0 \\ 0 & \cos\theta & 0 & 0 \\ 0 & 0 & \sin\theta & 0 \\ 0 & 0 & 0 & \sin\theta \end{pmatrix} |\psi\rangle_{AB},$

$$= \frac{1}{\sqrt{2}}(\cos\theta\,|00\rangle + \sin\theta\,|11\rangle) = \frac{1}{\sqrt{2}}|\phi\rangle_{AB}$$

So, for outcome m_1, Bharat does nothing (or applies the identity operator) to get the state $|\phi\rangle_{AB}$, thus completing the LOCC protocol.

If Aditi performs: $(M_2 \otimes \mathbb{I})|\psi\rangle_{AB} \to |\tilde{\psi}\rangle_{AB} = \begin{pmatrix} 0 & 0 & \cos\theta & 0 \\ 0 & 0 & 0 & \cos\theta \\ \sin\theta & 0 & 0 & 0 \\ 0 & \sin\theta & 0 & 0 \end{pmatrix} |\psi\rangle_{AB},$

$$= \frac{1}{\sqrt{2}}(\cos\theta\,|01\rangle + \sin\theta\,|10\rangle)\,.$$

For outcome m_2, Bharat applies the Pauli σ_x operator on his qubit to get $|\phi\rangle_{AB}$,

$$(\mathbb{I} \otimes \sigma_x)|\tilde{\psi}\rangle_{AB} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} |\tilde{\psi}\rangle_{AB} = \frac{1}{\sqrt{2}}(\cos\theta\,|00\rangle + \sin\theta\,|11\rangle) = \frac{1}{\sqrt{2}}|\phi\rangle_{AB}.$$

The factor $\frac{1}{\sqrt{2}}$ simply tells us that each occur with probability $\left|\frac{1}{\sqrt{2}}\right|^2 = \frac{1}{2}$.

Exercise: Show that the state $|\psi^-\rangle_{AB} \to |\phi\rangle_{AB}$ is possible under LOCC, where $|\psi^-\rangle_{AB} = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$, is the singlet and $|\phi\rangle_{AB}$ is any two-qubit state?

The question that now arises is whether we can always do such a thing. Can we transform any two-qubit state to another state? Is $|\psi\rangle_{AB} \to |\phi\rangle_{AB}$ under LOCC for any $|\psi\rangle_{AB}$ and $|\phi\rangle_{AB}$?

As you may have guessed, the answer is resoundingly NO!! We know that LOCC cannot create entanglement, so you cannot transform a separable state to an entangled state. It so happens that LOCC defines a partial order on set of pure bipartite quantum states, depending on whether we can reach certain states from a given initial entangled state. This also gives us the tool to define entanglement measures. But before that, to fully describe the interconversion of pure, bipartite quantum states we need to introduce the important concept of majorization, which captures the notion of order and disorder in a range of areas that involve some kind of irreversibility.

v) Majorization theory

Majorization is a relation defined between vectors $\boldsymbol{x} = \{x_1, x_2, x_3, \dots, x_n\}$ and $\boldsymbol{y} = \{y_1, y_2, y_3, \dots, y_N\}$. The ordering of the components does not matter, but just the distribution of the values and in particular the deviation of these components from being uniform. Since the ordering doesn't matter it is convenient to define the sorted vector $\boldsymbol{x}^{\downarrow} = \{x_1^{\downarrow}, x_2^{\downarrow}, x_3^{\downarrow}, \dots, x_N^{\downarrow}\}$, where $\boldsymbol{x}^{\downarrow} :=$ re-order components of $\boldsymbol{x}$ so elements decrease in size. For example: $\boldsymbol{x} = \{2, 3, 9, 5, 0, 1\}$ and $\boldsymbol{x}^{\downarrow} = \{9, 5, 3, 2, 1, 0\}$. The majorization relation between two vectors is then related by the following statement: For two vectors x and y, x is majorized by y, which is written as $x \prec y$, if and only if:

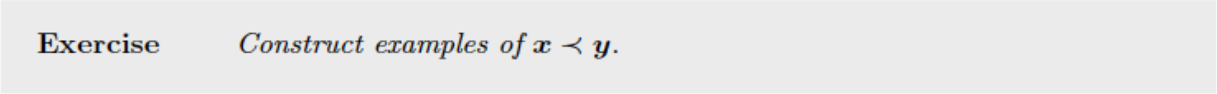
$$\text{I)} \sum_{i=1}^{j} x_i^{\downarrow} \leq \sum_{i=1}^{j} y_i^{\downarrow} \text{ for all } j = 1,\ 2,\ \ldots, N-1.$$

$$\text{II)} \sum_{i=1}^{N} x_i^{\downarrow} = \sum_{i=1}^{N} y_i^{\downarrow}$$

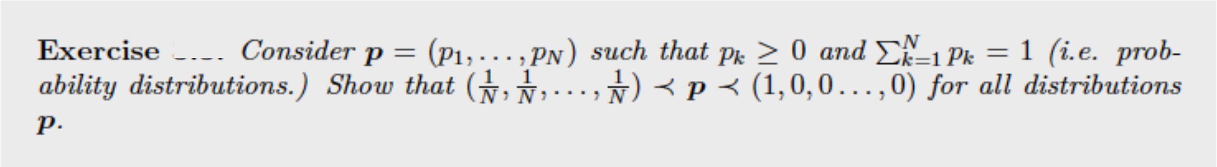
These conditions can be rewritten as:

$$\begin{gathered} x_1^{\downarrow} \leq y_1^{\downarrow} \\ x_1^{\downarrow} + x_2^{\downarrow} \leq y_1^{\downarrow} + y_2^{\downarrow} \\ x_1^{\downarrow} + x_2^{\downarrow} + x_3^{\downarrow} \leq y_1^{\downarrow} + y_2^{\downarrow} + y_3^{\downarrow} \\ . \\ . \\ x_1^{\downarrow} + x_2^{\downarrow} + x_3^{\downarrow} + \cdots + x_N^{\downarrow} = y_1^{\downarrow} + y_2^{\downarrow} + y_3^{\downarrow} + \cdots + y_N^{\downarrow} \end{gathered} \tag{5}$$

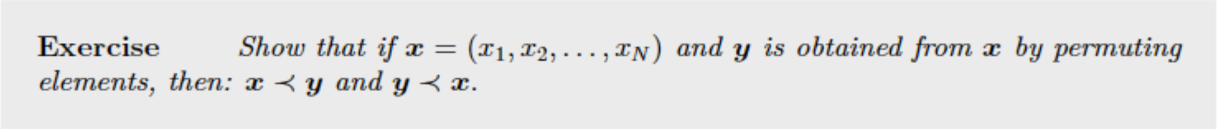
The condition $x \prec y$, implies that x is more disordered than y, or "x is more uniform than y".



Exercise *Construct examples of $\boldsymbol{x} \prec \boldsymbol{y}$.*



Exercise *Consider $\boldsymbol{p} = (p_1, \ldots, p_N)$ such that $p_k \geq 0$ and $\sum_{k=1}^{N} p_k = 1$ (i.e. probability distributions.) Show that $(\frac{1}{N}, \frac{1}{N}, \ldots, \frac{1}{N}) \prec \boldsymbol{p} \prec (1, 0, 0 \ldots, 0)$ for all distributions $\boldsymbol{p}$.*



Exercise *Show that if $\boldsymbol{x} = (x_1, x_2, \ldots, x_N)$ and $\boldsymbol{y}$ is obtained from $\boldsymbol{x}$ by permuting elements, then: $\boldsymbol{x} \prec \boldsymbol{y}$ and $\boldsymbol{y} \prec \boldsymbol{x}$.*

vi) Nielsen's theorem

Now that we have the majorization theorem, we can now make a strong statement about the central result for pure, bipartite state conversions using the LOCC protocol. The following theorem by Nielsen gives the necessary and sufficient condition for all such possible LOCC conversions.

Consider the state, $|\psi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_A$, where

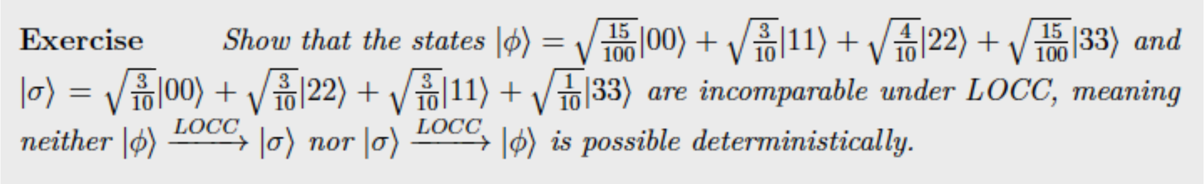
$$\lambda(\psi) := \{\text{eigenvalues of } \rho_A = \mathrm{Tr}_B[|\psi\rangle\langle\psi|_{AB}]\}$$

then the transformation, $|\psi\rangle_{AB} \xrightarrow{LOCC} |\phi\rangle_{AB}$ is possible deterministically, if and only if,

$$\lambda(\psi) \prec \lambda(\phi).$$

In other words, the partial order induced by LOCC on the set of pure bipartite states coincides with the partial order from majorization theory for the marginal spectra $\lambda(\psi)$ of the pure states.

For a given pure state there will exist states which can be reached in a reversible way under LOCC, and states that can be reached (or can be arrived from) in an irreversible way, and then there are states that are incomparable under LOCC.



Exercise *Show that the states* $|\phi\rangle = \sqrt{\frac{15}{100}}|00\rangle + \sqrt{\frac{3}{10}}|11\rangle + \sqrt{\frac{4}{10}}|22\rangle + \sqrt{\frac{15}{100}}|33\rangle$ *and* $|\sigma\rangle = \sqrt{\frac{3}{10}}|00\rangle + \sqrt{\frac{3}{10}}|22\rangle + \sqrt{\frac{3}{10}}|11\rangle + \sqrt{\frac{1}{10}}|33\rangle$ *are incomparable under LOCC, meaning neither* $|\phi\rangle \xrightarrow{LOCC} |\sigma\rangle$ *nor* $|\sigma\rangle \xrightarrow{LOCC} |\phi\rangle$ *is possible deterministically.*

vii) Catalysis

Consider the situation where Aditi and Bharat share a quantum state consisting by a pair of qudits ($d = 4$), given by

$$|\psi\rangle_{AB} = \sqrt{2/5}\,|00\rangle + \sqrt{2/5}\,|11\rangle + \sqrt{1/10}\,|22\rangle + \sqrt{1/10}\,|33\rangle.$$

Using Nielsen's theorem, it is clear to see that Aditi and Bharat cannot use LOCC to deterministically transform $|\psi\rangle_{AB}$ to the state below:

$$|\phi\rangle_{AB} = \sqrt{1/2}\,|00\rangle + \sqrt{1/4}\,|11\rangle + \sqrt{1/4}\,|22\rangle.$$

Let, $\lambda(\phi) := \{\text{eigenvalues of } \rho_A = \mathrm{Tr}_B[|\phi\rangle\langle\phi|_{AB}]\}$ (similarly for $\lambda(\psi)$). Since, both $|\psi\rangle_{AB}$ and $|\phi\rangle_{AB}$ are written in their Schmidt representation, it is straightforward to estimate both $\lambda(\psi)$ and $\lambda(\phi)$:

$$\lambda(\psi) = \{x\} = \{2/5, 2/5, 1/10, 1/10\};\ \lambda(\phi) = \{y\} = \{1/2, 1/4, 1/4, 0\}.$$

Applying the majorization criteria we have;

$$x_1 \leq y_1;\ x_1 + x_2 \geq y_1 + y_2.$$

Therefore, $\lambda(\psi) \nprec \lambda(\phi)$ and $\lambda(\phi) \nprec \lambda(\psi)$.

But if they can find an ancilla state somewhere:

$$|\zeta\rangle_{A'B'} = \sqrt{{}^{3}/_{5}}\,|00\rangle + \sqrt{{}^{2}/_{5}}\,|11\rangle.$$

it can be shown that $|\psi\rangle_{AB} \otimes |\zeta\rangle_{A'B'} \xrightarrow{LOCC} |\phi\rangle_{AB} \otimes |\zeta\rangle_{A'B'}$. Therefore the state $|\zeta\rangle_{A'B'}$ acts like a catalyst.

Let, $\lambda(\psi) := \{\text{eigenvalues of } \rho_{AA'} = \mathrm{Tr}_{BB'}[|\psi\rangle\langle\psi|_{AB} \otimes |\zeta\rangle\langle\zeta|_{A'B'}]\}$. The same for $\lambda(\phi)$. The eigenvalues are obtained by simply multiplying and squaring the Schmidt coefficients:

$$\lambda(\psi) = \{{}^{6}/_{25}, {}^{6}/_{25}, {}^{4}/_{25}, {}^{4}/_{25}, {}^{3}/_{50}, {}^{3}/_{50}, {}^{2}/_{50}, {}^{2}/_{50}\};$$

$$\lambda(\phi) = \{{}^{3}/_{10}, {}^{2}/_{10}, {}^{3}/_{20}, {}^{3}/_{20}, {}^{2}/_{20}, {}^{2}/_{20}, 0, 0\}.$$

It is easy to verify that now, $\lambda(\psi) \prec \lambda(\phi)$, which implies that

$$|\boldsymbol{\psi}\rangle_{AB} \otimes |\boldsymbol{\zeta}\rangle_{A'B'} \xrightarrow{LOCC} |\boldsymbol{\phi}\rangle_{AB} \otimes |\boldsymbol{\zeta}\rangle_{A'B'}.$$

viii) Majorization, the notion of disorder and its connection to LOCC

Note: This subsection is not necessary for exams.

As stated earlier the notion that $x \prec y$, implies that x is more disordered than y, is because $x \prec y$, if and only if $x = \sum_i q_i P_i y$, where P_i are permutation matrices. In other words, x is more disordered then y as it can always be represented as a convex mixture of vectors obtained after permuting the elements of y.

Interestingly, a doubly stochastic matrix D (with non-negative elements, and each row and column sum equal to 1) can always be written as $D = \sum_i q_i P_i$. Therefore, for $x \prec y$, we have $x = Dy$, for some doubly-stochastic matrix. Now, Horn's lemma[5] connects these double stochastic matrices to unitary matrices and majorization. For

[5] Please see the paper: *Majorization and the interconversion of bipartite states*, by Nielsen and Vidal, Quantum Information & Computation 1, 76 (2001). Also see: Chapter 12 of Nielsen and Chuang.

instance, if U is a unitary matrix with elements u_{ij}, then the matrix $D_{ij} \equiv |u_{ij}|^2$ is a doubly stochastic. So, for $x \prec y$, we have $x = Dy$, where D is a unitary-stochastic matrix. Uhlmann's theorem connects the Horn's lemma to the eigenvalues of Hermitian operators, X and Y, given by say $\lambda(X)$ and $\lambda(Y)$, by stating that $\lambda(X) \prec \lambda(Y)$, if and only if, $X = \sum_i q_i U_i Y U_i^\dagger$, where U_i are unitary matrices. This shows that operator X (reduced density matrices in Nielsen's theorem) is more disordered that Y, as X can be obtained by the convex mixing of Y acted upon by unitaries.

So, all that remains is to connect LOCC to Uhlmann's lemma. As shown earlier in the example on LOCC, for bipartite pure states shared between two parties, the LOCC protocol can be described in terms of measurement $\mathcal{M} = \{M_i\}$ performed by Aditi and a set of conditioned unitaries $\{U_i\}$ performed by Bharat locally on his system. Again, consider the transformation, $|\psi\rangle_{AB} \rightarrow |\phi\rangle_{AB}$.

At Aditi's end she starts with a reduced system be ρ_ψ and ends up with ρ_ϕ after measurements on her subsystem, i.e., $M_i \rho_\psi {M_i}^\dagger = p_i \rho_\phi$, where p_i is the probability of the measurement outcome, m_i.

Now, $M_i \rho_\psi {M_i}^\dagger = \sqrt{M_i \rho_\psi {M_i}^\dagger}\sqrt{M_i \rho_\psi {M_i}^\dagger} = \sqrt{M_i \rho_\psi {M_i}^\dagger}\, U_i {U_i}^\dagger \sqrt{M_i \rho_\psi {M_i}^\dagger}$ and $M_i \rho_\psi {M_i}^\dagger = M_i \sqrt{\rho_\psi}\sqrt{\rho_\psi} {M_i}^\dagger$, which gives us the polar decomposition: $M_i \sqrt{\rho_\psi} = \sqrt{M_i \rho_\psi {M_i}^\dagger} U_i = \sqrt{p_i \rho_\phi}\, U_i$. Multiplying with the adjoint gives us:

$$\sqrt{\rho_\psi} {M_i}^\dagger M_i \sqrt{\rho_\psi} = p_i\, {U_i}^\dagger \rho_\phi U_i.$$

Summing, we have: $\rho_\psi = \sum_i p_i\, {U_i}^\dagger \rho_\phi U_i$, which gives us the Uhlmann's theorem for convex mixing and thus connects LOCC with majorization.

ix) Entanglement monotone

Now one can attach a measure of entanglement E for pure bipartite quantum states under LOCC, by at the very least demanding that if $|\psi\rangle_{AB} \rightarrow |\phi\rangle_{AB}$, we have the relation, $E(|\psi\rangle_{AB}) \geq E(|\phi\rangle_{AB})$. This condition demands that the measure is an entanglement monotone under LOCC, meaning its value can never increase under local operations and classical communications. This makes sense as we know LOCC cannot create nor increase entanglement. For such an entanglement monotone we also demand that if the reversible transformation $|\phi\rangle_{AB} \rightarrow |\psi\rangle_{AB}$ is permissible, then $E\left(|\psi\rangle_{AB}\right) = E(|\phi\rangle_{AB})$, and the states must be equally entangled with respect to any measure. If $|\phi\rangle_{AB}$ and $|\psi\rangle_{AB}$ are incomparable then there is, a priori, no constraint on the measure of entanglement. How would you find such a measure?

Now Nielsen's theorem requires $\lambda(\psi) \prec \lambda(\phi)$ for $|\psi\rangle_{AB} \rightarrow |\phi\rangle_{AB}$, where $\lambda(\psi)$ are the eigenvalues of ρ_A $(\rho_A = \mathrm{Tr}_B\left[|\psi\rangle\langle\psi|_{AB}\right])$. From Section viii) above, we know that $\rho_\psi = \sum_i p_i\, {U_i}^\dagger \rho_\phi U_i = \sum_i p_i\, \sigma_i$. Now for any symmetric, concave function $\mathcal{E}$ of the density matrix, we know that if $\rho_\psi = \sum_i p_i\, \sigma_i$, then

$$\mathcal{E}\left(\rho_\psi\right) \geq \sum_i p_i\, \mathcal{E}(\sigma_i) = \sum_i p_i\, \mathcal{E}({U_i}^\dagger \rho_\phi U_i).$$

Now, if $\mathcal{E}$ is invariant under local rotations, then $\mathcal{E}({U_i}^\dagger \rho_\phi U_i) = \mathcal{E}\left(\rho_\phi\right)$, and therefore we have the required monotone $\mathcal{E}$, where $\mathcal{E}\left(\rho_\psi\right) \geq \mathcal{E}\left(\rho_\phi\right)$.

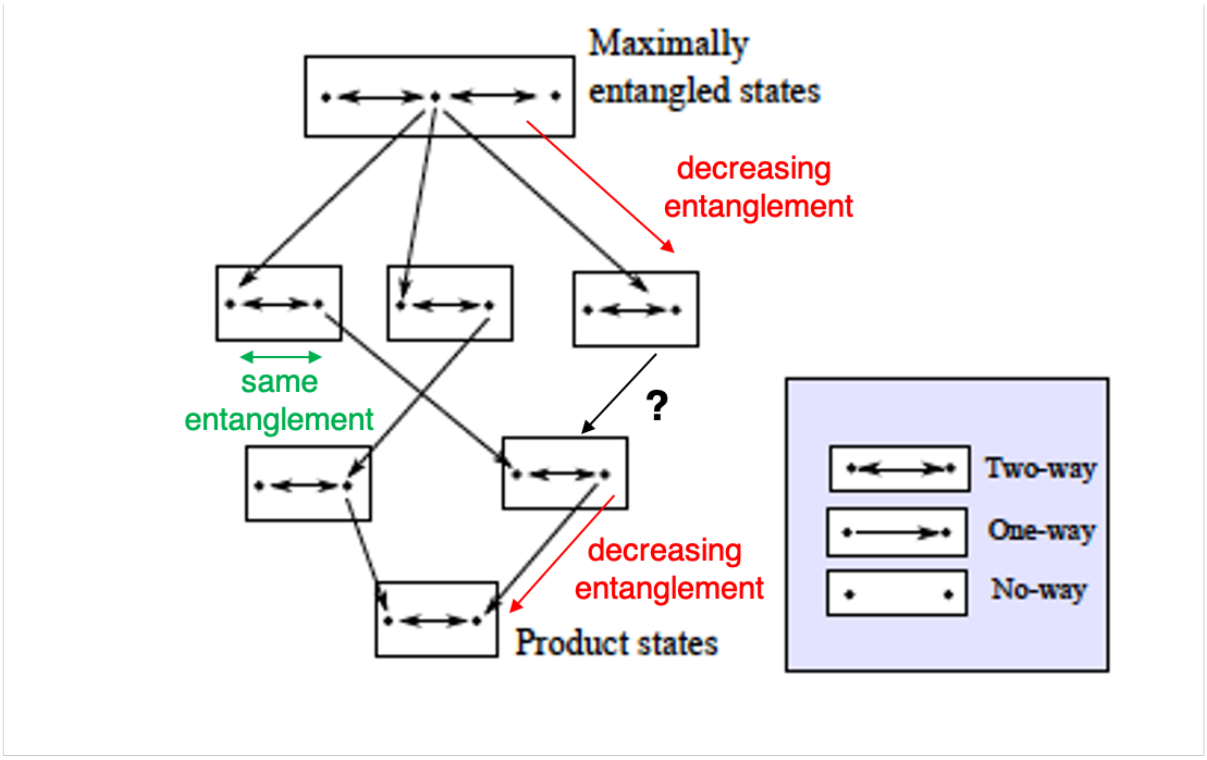
This is related to Schur-concave functions, where $f\left(\lambda(\psi)\right) \geq f\left(\lambda(\phi)\right)$ for $\lambda(\psi) \prec \lambda(\phi)$. One such Schur-concave function is the von Neumann entropy of entanglement $S\left(\rho_\psi\right) = S\left(\lambda(\psi)\right) = -\sum_i \lambda_i \log \lambda_i$. In the next part, we will see how this quantity is an operational measure of entanglement for pure states.

C. Entanglement measures

We briefly discussed the possibility of obtaining a measure of entanglement starting simply from Nielsen's theorem, and demanding that such a measure should be a monotone under LOCC, meaning its value can never increase under local operations and classical communications. Here, we will look at the "asymptotic regime" and by adding a few more very natural and intuitive properties for any potential measure we will attempt to arrive at a unique measure of entanglement.

i) Maximally entangled states

Before we start with asymptotic conversion of quantum states, we revisit Nielsen's theorem once more to discuss the set of maximally entangled states. The figure below captures the different interconversion as defined by Nielsen's theorem:



Let us consider the pure bipartite state in the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$,

$$|\Psi\rangle_{AB} = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |i\rangle_A \otimes |i\rangle_B .$$

Here, $d = \dim(\mathcal{H}_A) = \dim(\mathcal{H}_B)$ and the marginally states are maximally mixed (all states are equally probable), i.e., $\mathrm{Tr}_A[|\Psi\rangle\langle\Psi|] = \mathrm{Tr}_B[|\Psi\rangle\langle\Psi|] = \mathbb{I}_d/d$. So, the eigenvalues of the marginals are $\lambda(|\Psi\rangle_{AB}) = \left\{{}^1\!/_d, {}^1\!/_d, {}^1\!/_d, \dots, {}^1\!/_d\right\}$.

Now for any arbitrary state $|\varphi\rangle_{AB}$, say the marginals are $\lambda(|\varphi\rangle_{AB}) = \{x_1, x_2, x_3 \dots, x_d\}$.

Exercise: Prove that $\lambda(|\Psi\rangle_{AB}) \prec \lambda(|\varphi\rangle_{AB})$ using majorization conditions.

Therefore, $\lambda(|\Psi\rangle_{AB}) \prec \lambda(|\varphi\rangle_{AB})$, from Nielsen's theorem implies that:

$$|\Psi\rangle_{AB} \longrightarrow |\varphi\rangle_{AB} \text{ and } E(|\Psi\rangle_{AB}) \geq E(|\varphi\rangle_{AB}), \qquad (2)$$

for all bipartite stays $|\varphi\rangle_{AB}$. This means all bipartite states can be created from $|\Psi\rangle_{AB}$ using LOCC, and hence has entanglement less than it. See the illustration in the previous page. Hence, $|\Psi\rangle_{AB}$ is a maximally entangled state in $\mathcal{H}_A \otimes \mathcal{H}_B$.

ii) Asymptotic conversion rates – Distillation of entanglement

First, we focus on the scenario of transforming a large number of copies of a state into a large number of copies of some other state. The standard form of this is where we start with N copies of some bipartite state ρ_{AB} and wish to know how many copies of the state σ_{AB} can we create using only LOCC? In other words, determine M in $\rho_{AB}^{\otimes N} \longrightarrow \sigma_{AB}^{\otimes M}$, as a function of N.

However, it makes sense to standardize things some more, and instead of leaving the final σ_{AB} unspecified we shall pick a natural unit of entanglement. In particular we take the singlet state $|\psi^-\rangle$ to be our standard measure, since it is the simplest bipartite maximally entangled pure state. We shall take the singlet to comprise one unit of entanglement, and can then consider all LOCC operations on N copies of some initial state ρ_{AB} and "count" how much entanglement we can extract. We consider LOCC transformations $\rho_{AB}^{\otimes N} \longrightarrow (|\psi^-\rangle\langle\psi^-|)_{AB}^{\otimes M}$ and see what is the largest M that can be extracted. The process of obtaining maximally entangled singlets from multiple copies of a mixed state is called *distillation of entanglement*.

Conversely, we can also consider the reverse process, and begin with multiple copies of our standard, $|\psi^-\rangle\langle\psi^-|_{AB}^{\otimes k}$ and using only LOCC try to form as many copies of ρ_{AB} as possible. This is sometimes called a dilution of entanglement.

iii) Quantifying pure bipartite entanglement

Now we look at the interconversion between pure, bipartite quantum states in the asymptotic regime, i.e., when there are $N \to \infty$ copies of the state $|\varphi\rangle$ and in the process derive the central measure of pure, bipartite entanglement - the von Neumann entropy of entanglement, which we introduced earlier.

The central result here is, given $N \to \infty$ copies of any state $|\varphi\rangle$ there exist reversible interconversions $|\varphi\rangle\langle\varphi|_{AB}^{\otimes N} \longrightarrow |\psi^-\rangle\langle\psi^-|_{AB}^{\otimes S(\rho_A)N}$, where $\rho_A = \mathrm{Tr}_B[|\varphi\rangle\langle\varphi|]$ is the reduced state on A of a single copy $|\varphi\rangle$ and $S(\rho_A) = -\rho_A \log \rho_A$ is the von Neumann entropy of the state reduced state ρ_A.

Note: This proof is not necessary for exams.

Proof: Let $|\varphi\rangle_{AB}$ be a pair of qubits for simplicity. The Schmidt decomposition for this state is, $|\varphi\rangle_{AB} = \sqrt{1-p}\,|0\rangle_A|0\rangle_B + \sqrt{p}|1\rangle_A|1\rangle_B$. We will first look at the distillation protocol $|\varphi\rangle\langle\varphi|_{AB}^{\otimes N} \longrightarrow |\psi^-\rangle\langle\psi^-|_{AB}^{\otimes M}$ and determine M as a function of N, as N gets larger and larger.

Let us start with the simple case: $N = 3$,

$$|\varphi\rangle_{AB}^{\otimes N} = \left(\sqrt{1-p}\,|0\rangle_A|0\rangle_B + \sqrt{p}|1\rangle_A|1\rangle_B\right)^{\otimes 3}$$

$$|\varphi\rangle_{AB}^{\otimes 3} = (1-p)^{3/2}|000\rangle_A|000\rangle_B + (1-p)p^{1/2}\begin{Bmatrix}(|001\rangle+|010\rangle+|100\rangle)_A\otimes \\ (|001\rangle+|010\rangle+|100\rangle)_B\end{Bmatrix}$$

$$+\,p(1-p)^{1/2}\begin{Bmatrix}(|011\rangle+|101\rangle+|111\rangle)_A\otimes \\ (|011\rangle+|101\rangle+|111\rangle)_B\end{Bmatrix} + p^{3/2}|111\rangle_A|111\rangle_B.$$

For general N, the terms in the state can be arranged in a similar manner:

$$|\varphi\rangle_{AB}^{\otimes N}=\sum_{k=1}^{N}(1-p)^{k/2}\,p^{\frac{N-k}{2}}\sum_{\mathcal{P}}\left\{\mathcal{P}\left(|0\rangle^{\otimes k}\;|1\rangle^{\otimes(N-k)}\right)_A\otimes\mathcal{P}\left(|0\rangle^{\otimes k}\;|1\rangle^{\otimes(N-k)}\right)_B\right\}$$

where, $\mathcal{P}$ is the permutation operator that contains the sum of all possible combinations of k zeros and $(N-k)$ ones, such that (for example in $N=3$):

$$\sum_{\mathcal{P}}\mathcal{P}\left(|1\rangle^{\otimes 2}\;|0\rangle^{\otimes 1}\right)=|011\rangle+|101\rangle+|111\rangle.$$

For, every k, and the corresponding term $\mathcal{P}\left(|1\rangle^{\otimes k}\;|0\rangle^{\otimes(N-k)}\right)$, there are $\binom{N}{k}=\frac{N!}{k!(N-k)!}$ terms in the permutation. Now this allows us to make use of the typical subspaces that we discussed in our first lecture.

For, $|\varphi\rangle_{AB}^{\otimes N}=\left(\sqrt{1-p}\;|0\rangle_A|0\rangle_B+\sqrt{p}|1\rangle_A|1\rangle_B\right)^{\otimes N}$, and $N\longrightarrow\infty$, the "typical subspace" will have $k=pN$ and the normalized state is a uniform superposition of $\binom{N}{pN}$ states. The rest of the states are unlikely to appear and form the atypical subspace.

Now how can we create an LOCC protocol for the distillation process in the asymptotic regime.

- (**Local measurement**) Aditi collapses the state to one of her k-subspaces and obtains an outcome, 0≤ $k\leq 1$. Let us assume for the moment that k is not the typical subspace. Her collapsed state now is:

  $$|\varphi\rangle_{AB}^{\otimes N}=\frac{1}{\binom{N}{k}}\sum_{\mathcal{P}}\left\{\mathcal{P}\left(|0\rangle^{\otimes k}\;|1\rangle^{\otimes(N-k)}\right)_A\otimes\mathcal{P}\left(|0\rangle^{\otimes k}\;|1\rangle^{\otimes(N-k)}\right)_B\right\}$$

  This outcome happens with the binomial probability $(1-p)^{k/2}\,p^{\frac{N-k}{2}}\binom{N}{k}$.

  If the measured subspace by Aditi is the typical subspace, as is the case when $N\longrightarrow\infty$, she has $k=pN$, and then from earlier calculations for typical subspaces we know that

  $$\log\binom{N}{pN}=\log\frac{N!}{pN!\,(N-pN)!}\cong NS(\rho_A);\rho_A=\mathrm{Tr}_B[|\varphi\rangle\langle\varphi|].$$

  So Aditi now has $|\varphi\rangle_{AB}^{\otimes N}$ which is a uniform superposition of $2^{NS(\rho_A)}$ orthogonal product states.
- (**Classical communication**) She communicates her result to Bharat, saying that she has $k=pN$. Maybe she does not really need to communicate.
- (**Local unitaries**) Both parties now know that the pure entangled state they possess corresponds to the typical subspace, $k=pN$. They can rearrange their

respective qubits such that $2^{NS(\rho_A)}$ orthogonal product states such that they can be compressed to $S(\rho_A)N$ maximally entangled qubits, as shown below:

$$|111\dots100\rangle_A \longrightarrow |0\dots00\rangle_A\,|1\dots100\rangle_A$$

$$|111\dots100\rangle_A \longrightarrow |0\dots01\rangle_A\,|1\dots100\rangle_A$$

$$\vdots$$

$$|111\dots100\rangle_A \longrightarrow \underbrace{|1\dots11\rangle_A}_{NS(\rho_A)}\,\underbrace{|1\dots100\rangle_A}_{\text{throw}}$$

Now, Bharat does the same. They effectively compress the state to the form:

$$\left(|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B\right)^{\otimes NS(\rho_A)} \otimes |0\dots00\rangle_A \otimes |1\dots100\rangle_B.$$

Upon discarding the extra qubits, Aditi and Bharat are left with $S(\rho_A)N$ maximally entangled qubits, which they can convert to equal number of singlets using LOCC. This completes the distillation protocol.

We can also go in the opposite direction. One can dilute $S(\rho_A)N$ singlets into N copies of $|\varphi\rangle_{AB}$. The protocol basically relies on teleportation (which we will discuss later) and proceeds along these lines.

- (Local operations) Aditi creates N copies of $|\varphi\rangle_{AA'}$ locally.
- (LOCC) She compresses the state of N copies in the A' subspace to $NS(\rho_A)$ qubits and teleports these to Bharat, consuming $NS(\rho_A)$ number of qubits.
- (Local operations) Bharat decompresses these to obtain N copies of $|\varphi\rangle_{AB}$.

An important aspect we discussed in the last two sections was that asymptotic distillation is reversible.

$$|\varphi\rangle\langle\varphi|_{AB}^{\otimes N} \xleftrightarrow{\text{LOCC}} |\psi^-\rangle\langle\psi^-|_{AB}^{\otimes S(\rho_A)N}.$$

This makes the rate of distillation optimal, i.e., we can utmost distil, $S(\rho_A)N$ singlets from N copies of $|\varphi\rangle_{AB}$. Why is this the case? Because, LOCC cannot create entanglement for free, and distilling anything more than $S(\rho_A)N$ singlets will allow one to do so. For instance, if we could distill $S(\rho_A)N + k$ singlets from $|\varphi\rangle_{AB}^{\otimes N}$. Then one can construct N copies of $|\varphi\rangle_{AB}$ and teleport them using $S(\rho_A)N$ singlet pairs. This implies that at the end of teleportation we are not only left with $|\varphi\rangle_{AB}^{\otimes N}$ states that we began with but also an additional k singlets, which we now have created using LOCC. This is not permissible and therefore $S(\rho_A)N$ is the optimal distillation rate.

iv) The von Neumann entropy of entanglement

The asymptotic result and the optimality allow us to define a unique and natural measure for pure bipartite state entanglement – the von Neumann entropy. It's key features as an entanglement monotone include:

1) It is an LOCC monotone, i.e.,

$$|\psi\rangle_{AB} \longrightarrow |\varphi\rangle_{AB} \text{ and } E\big(|\psi\rangle_{AB}\big) \geq E\big(|\varphi\rangle_{AB}\big).$$

2) It is equal to one for the maximally entangled two-qubit or singlet state:

$$E\big(|\psi^-\rangle_{AB}\big) = 1;\ |\psi^-\rangle_{AB} = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle).$$

3) The measure is extensive: $E\big(|\varphi\rangle_{AB}^{\otimes N}\big) = n\, E\big(|\varphi\rangle_{AB}\big)$. For singlets: $E\big(|\psi^-\rangle_{AB}^{\otimes k}\big) = k$.

D. Mixed state entanglement – convex sets and geometry of state space

In the last couple of chapters, we discussed entanglement of pure, bipartite quantum states starting from an operational perspective, which allows for the definition of entanglement as a *resource*[6] that cannot be created by the set of local operations and classical communication (LOCC). Moreover, we also discussed how von Neumann entropy emerges as a unique measure of pure state entanglement when we study asymptotic conversion rates between quantum states under LOCC protocols.

The general extension of our study at this stage is to study entanglement in mixed states. It turns out that even for the case of bipartite mixed states, determining whether the state is entangled or not is quite challenging and not always completely well-understood. As a start, we first depart from an operational perspective and quantify entanglement in terms of the geometry and convexity of the relevant state space.

Before, that let us take a quick look at how entanglement becomes quirky even in simple bipartite mixed states if we stick to our naive pure state understanding.

i) Entangled mixed states

Let us consider the mixed state given by equal mixture of maximally entangled states:

$$\rho_{AB} = \sum_i p_i |\varphi_i\rangle\langle\varphi_i|_{AB} = \frac{1}{4}\,|\phi^+\rangle\langle\phi^+| + \frac{1}{4}\,|\phi^-\rangle\langle\phi^-| + \frac{1}{4}\,|\psi^+\rangle\langle\psi^+| + \frac{1}{4}\,|\psi^-\rangle\langle\psi^-|, \quad (1)$$

where, the Bell states are given by:

$$|\phi^\pm\rangle_{AB} = \tfrac{1}{\sqrt{2}}\left(|00\rangle_{AB} \pm |11\rangle_{AB}\right) \text{ and } |\psi^\pm\rangle_{AB} = \tfrac{1}{\sqrt{2}}\left(|01\rangle_{AB} \pm |10\rangle_{AB}\right).$$

$$\rho_{AB} = \frac{1}{4}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \frac{1}{4}\,\mathbb{I}_{4\,\mathrm{x}\,4}.$$

Now, an intuitive way of calculating entanglement is by taking its average in some sense, as given by the following relation,

[6] Entanglement can be formulated as a resource theory. I will provide additional notes for those interested.

$$E(\rho_{AB}) = \sum_i p_i \, E\big(|\varphi_i\rangle_{AB}\big),$$

where, $E\big(|\varphi_i\rangle_{AB}\big)$ is the von Neumann entanglement entropy.

For a general, $\rho_{AB} = \sum_i p_i |\varphi_i\rangle\langle\varphi_i|_{AB}$, one can invoke the amount of dilution needed to create ρ_{AB}. For instance, the number of singlets needed by Aditi and Bharat to create n copies of the pure state $|\varphi_i\rangle_{AB}$ by LOCC in the asymptotic limit is given by $nE\big(|\varphi_i\rangle_{AB}\big)$. So, one needs $\sum_i p_i nE\big(|\varphi_i\rangle_{AB}\big)$ number of singlets in the ensemble to create n copies of ρ_{AB}. But this is not really the same as distillation (as was the case for pure states), and is called the entanglement of formation (which we discuss later).

Therefore, for the bipartite mixed state ρ_{AB}, which is a mixture of four maximally entangled Bell states, where we have $E\big(|\varphi_i\rangle_{AB}\big) = 1$ for $|\varphi_i\rangle_{AB} = \big\{|\phi^{\pm}\rangle_{AB}, |\psi^{\pm}\rangle_{AB}\big\}$ and therefore the entanglement $E(\rho_{AB}) = 1$. Therefore, the bipartite mixed state ρ_{AB} also appears to be maximally entangled. Is this correct? No. Why?

Because the identity matrix $\mathbb{I}_{4\,\mathrm{x}\,4}$, and therefore the state ρ_{AB}, can also be written using the computational basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$, such that:

$$\rho_{AB} = \frac{1}{4}\,|00\rangle\langle 00| + \frac{1}{4}\,|01\rangle\langle 01| + \frac{1}{4}\,|10\rangle\langle 10| + \frac{1}{4}\,|11\rangle\langle 11| = \frac{1}{4}\,\mathbb{I}_{4\,\mathrm{x}\,4}.$$

However, these basis states are all product states of the form $|i\rangle \otimes |j\rangle$, and therefore, require no singlets to create, i.e., $E(\rho_{AB}) = 0$. So simply by using a different pure state decomposition for the same density matrix, the entanglement in the state vanishes. The anomaly it seems is that ρ_{AB} can be created using four maximally entangled states, which are very expensive in terms of resource, and on the other hand, it can also be created simply using four product or separable states. The final density matrix using both the methods of preparation are indistinguishable. An important way to define entanglement is to then use the following definition, often called the convex roof construction:

$$E(\rho_{AB}) = \min_{\{p_i, |\varphi_i\rangle\}} \sum_i p_i \, E\big(|\varphi_i\rangle_{AB}\big).$$

So, entanglement in ρ_{AB} is defined as the average over any pure state decomposition, minimized over all possible decompositions $\{p_i, |\varphi_i\rangle\}$. Therefore, for the state ρ_{AB}, the entanglement, $E(\rho_{AB}) = 0$.

However, in general this is a very difficult task. For any arbitrary dimensional bipartite state ρ_{AB}, it is nigh impossible to find all possible pure state decompositions. So, we make use of the geometry and convexity properties of the set of quantum states.

Exercise: Verify that the state $\rho_{AB} = {}^1/_4 \; \mathbb{I}_{4\,\mathrm{x}\,4}$, does indeed have two or more indistinguishable pure state decompositions.

ii) Separable states

Let us begin by defining the set of states that can be created using LOCC i.e., the set of separable states, as those that can be written as a sum of product states. Given a bipartite quantum system in the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, the set of separable states $\mathcal{D}_{sep}$ is given by[7]:

$$\mathcal{D}_{sep} = \left\{ \sigma_{AB}\colon \sigma_{AB} = \sum_i p_i \; \sigma_i^A \otimes \sigma_i^B \right\},$$

where σ_i^A and σ_i^B are local density matrices and p_i is the probability distribution, such that $\sum_i p_i \; = 1$. Think of the following LOCC protocol between Aditi and Bharat:

ADITI

Aditi creates σ_1^A and Bharat σ_1^B with prob p_1
σ_2^A and σ_2^B with prob p_2
⋮
σ_N^A and σ_N^B with prob p_N

BHARAT

We note that $\mathcal{D}_{sep}$ is a convex set; and belongs to the larger set of bipartite quantum states, $\mathcal{D} \coloneqq \{\sigma_{AB} \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)\}\colon \mathrm{Tr}(\sigma_{AB}) = 1; \sigma_{AB} \geq 0$.

Exercise: Show that $\mathcal{D}_{sep}$ is a convex set.

[7] Werner, R. F. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. Phys. Rev. A **40**, 4277 (1989).

iii) Entanglement witness

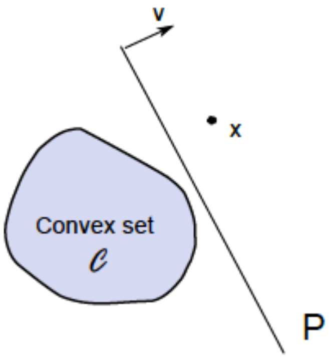
The convexity of the set of separable states is a very important result. We now want to make use of this convexity property to be able to distinguish these "classical" separable states from the genuinely "quantum" entangled states. This can be done using a special form of the Hahn-Banach theorem or the hyperplane separation theorem, which we state without proof:

Given a closed convex set $\mathcal{C}$ in a finite Banach space and a point x outside $\mathcal{C}$, there exists a hyperplane P that separates x from $\mathcal{C}$.

What this really means is that one can divide the whole space into a region that does not contain $\mathcal{C}$ and one that does, by simply using a hyperplane.

A hyperplane P in an n-dimensional vector space is a vector subspace of codimension 1 (or dimension $n-1$), either lying on the origin or shifted by a vector. So, one can define the hyperplane $P = \{\boldsymbol{x}: \langle \boldsymbol{x}, \boldsymbol{v} \rangle = 0\}$, where the vector $\boldsymbol{v}$ characterizes the hyperplane (this is similar to how one defines a plane in co-ordinate geometry). Now all vectors on the hyperplane are orthogonal to $\boldsymbol{v}$, with respect to some inner product $\boldsymbol{x} \cdot \boldsymbol{v} = \langle \boldsymbol{x}, \boldsymbol{v} \rangle$.

It is easy to check that there will always exist a set of points $\boldsymbol{x}$, where $\langle \boldsymbol{x}, \boldsymbol{v} \rangle > 0$ on one side of P and $\langle \boldsymbol{x}, \boldsymbol{v} \rangle < 0$ on the other.



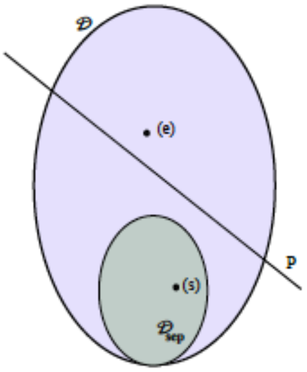
Exercise: Determine the hyperplanes in the real vector space defined by the vector $\boldsymbol{v}$, where (a) $\boldsymbol{v} = (0, 1, 0)$, (b) $\boldsymbol{v} = (-1, 0, 0)$, and (c) $\boldsymbol{v} = (1, 1, -1)$. Find the positive and negative sides of the planes.

An important consequence of the special form of the Hahn-Banach theorem is that it can now be applied to the convex set of separable states, $\mathcal{D}_{sep}$, if only we can find the suitable vector space and an inner product that can be used to define hyperplane.

The vector space is the $d^2 \times d^2$ space of Hermitian matrices in $\mathcal{H}_A \otimes \mathcal{H}_B$, where all the quantum states and observables lie. The inner product is the Hilbert-Schmidt inner product in quantum information theory, which is defined as $x.v = \langle x|v\rangle = \mathrm{Tr}[X^\dagger Y]$, for two complex matrices X and Y. The Hahn-Banach theorem can now be used to detect an entangled state.

Consider the state $\rho_{AB} \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ which is entangled and therefore lies outside the convex set of separable states $\mathcal{D}_{sep}$. Therefore, there must exist a hyperplane in the relevant vector space of Hermitian matrices that separates ρ_{AB} from the separable convex set $\mathcal{D}_{sep}$. This hyperplane can be defined using a Hermitian operator $W \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$, called the entanglement witness. We define it more precisely as follows.

An entanglement witness $W \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is a Hermitian observable such that $\mathrm{Tr}[W\sigma_{AB}] \geq 0$ for all separable states $\sigma_{AB} \in \mathcal{D}_{sep}$.



Schematic of the convex set of all bipartite quantum states. Here $\mathcal{D}_{sep}$ is the convex set of separable states. For any entangled state (e), there always exists a separating hyperplane P which separates (e) from $\mathcal{D}_{sep}$. This hyperplane is defined by some Hermitian entanglement witness W.

Exercise: Determine the hyperplane that separates the separable set from the entangled state.

Therefore, if there exists an observable W, which has positive expectation values for the set of all separable states, and we experimentally find that $\mathrm{Tr}[W\rho_{AB}] < 0$, for some quantum state ρ_{AB} then that state must be entangled. Conversely, we know that for any entangled quantum state ρ_{AB} there must exist at least one entanglement witness that "detects" its entanglement in the sense that $\mathrm{Tr}[W\rho_{AB}] < 0$.

> *Exercise:* From the schematic of the convex set of all bipartite quantum states and separable states above, prove using geometric arguments that the condition $\mathrm{Tr}[W\sigma_{AB}] \geq 0$ is necessary but not sufficient for σ_{AB} to be separable.

iv) Necessary and sufficient criteria for entanglement

In the last section, we observed how the convexity of the set of separable states can be exploited to come up with the notion of special observables – called entanglement witnesses – that mainly detect entanglement in general quantum states through negative expectation values. Now, we know that the "Choi-Jamiolkowski isomorphism" also provides us with necessary tools to convert observables (remember Choi operators $\mathcal{J}(\mathcal{E})$) on two systems to an operator ($\mathcal{E}$) acting on a single system

i.e., $\mathcal{J}(\mathcal{E}) = (\mathcal{E} \otimes \mathbb{I})\, |vec(\mathbb{I})\rangle\langle vec(\mathbb{I})|$, where $\mathbb{I}$ is the identity operator.

If $W = \mathcal{J}(\mathcal{E})$ is a Hermitian operator, then from the Choi-Jamiolkowski isomorphism we know that there then exists an Hermiticity preserving operator $\mathcal{E}$. Additionally, if $W = \mathcal{J}(\mathcal{E})$ is also an entanglement witness, i.e., it satisfies $\mathrm{Tr}[W\sigma_{AB}] \geq 0 \;\forall\; \sigma_{AB} \in \mathcal{D}_{sep}$, then $\mathcal{E}$ must also be positive.

Note: This proof is not necessary for exams.

Proof: $\sigma_{AB} = \sum_i p_i\, \sigma_i^A \otimes \sigma_i^B, \forall\, \sigma_{AB} \in \mathcal{D}_{sep}$. Now, $\mathrm{Tr}[W\sigma_{AB}] \geq 0$ will hold if $\mathrm{Tr}[W\sigma_i^A \otimes \sigma_i^B] \geq 0$. In fact, we can generalize this to pure states such that $\mathrm{Tr}[W|\phi_A\rangle\langle\phi_A| \otimes |\phi_B\rangle\langle\phi_B|] \geq 0$, which implies that $\langle\phi_A|\langle\phi_B|W|\phi_A\rangle|\phi_B\rangle \geq 0$. For $W = \mathcal{J}(\mathcal{E})$, we need to show that $\langle\phi_A|\langle\phi_B|((\mathcal{E}\otimes\mathbb{I})|v\rangle\langle v|)|\phi_A\rangle|\phi_B\rangle \geq 0$, where we have $|v\rangle\langle v| = |vec(\mathbb{I})\rangle\langle vec(\mathbb{I})| = \sum_{i,j}|i\rangle\langle i| \otimes |j\rangle\langle j|$. Now, one can write $|\phi_A\rangle = \sum_k \alpha_k|k\rangle$ and $|\phi_B\rangle = \sum_k \beta_k|k\rangle$, such that $\langle i|\phi_A\rangle = \alpha_i$ and $\langle j|\phi_b\rangle = \beta_j$.

$$\langle\phi_A|\langle\phi_B|\big((\mathcal{E}\otimes\mathbb{I})|v\rangle\langle v|\big)|\phi_A\rangle|\phi_B\rangle \geq 0 \;\Rightarrow \sum_{i,j}\langle\phi_A|\mathcal{E}|i\rangle\langle i|\phi_A\rangle \otimes \langle\phi_B|j\rangle\langle j|\phi_B\rangle \geq 0 \Rightarrow \sum_i\langle\phi_A|\mathcal{E}|i\rangle\langle i|\phi_A\rangle \geq 0.$$

From the last term we get $\langle\phi_A|\mathcal{E}|\phi_A\rangle \geq 0$, which proves that $\mathcal{E}$ is positive if $\langle\phi_A|\langle\phi_B|W|\phi_A\rangle|\phi_B\rangle \geq 0$, or W is an entanglement witness such that $\mathrm{Tr}[W\sigma_{AB}] \geq 0$.

Bipartite operator $W = \mathcal{J}(\mathcal{E})$	Linear operator $\mathcal{E}$ on a single system
Hermitian	Hermiticity preserving
Hermitian and $\mathrm{Tr}[W\sigma_{AB}] \geq 0 \; \forall \; \sigma_{AB} \in \mathcal{D}_{sep}$	Positive
Positive	Completely positive

Importantly, if $W = \mathcal{J}(\mathcal{E})$ is an entanglement witness with negative values, then $\mathcal{E}$ cannot be completely positive, which means $\mathcal{E}$ is positive but $(\mathcal{E} \otimes \mathbb{I})$ is not. Therefore, we can now convert the bipartite observable that defines the entanglement witness into an equivalent statement in terms of operator maps that detect the presence of entanglement through the generation of negative eigenvalues. To reiterate, this implies that the corresponding operator is not completely positive, but it is positive and hermiticity preserving (this is due to the fact that the original entanglement witness is an Hermitian operator).

A state $\rho_{AB} \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is entangled if and only if there exists a positive operator $\mathcal{E}$, such that $(\mathcal{E} \otimes \mathbb{I})\rho_{AB}$ has a negative eigenvalue.

Proof: It is quite straightforward to prove the above statement. Let us start with a direct proof. Let us say, σ_{AB} is a separable state, such that: $\sigma_{AB} = \sum_i p_i \; \sigma_i^A \otimes \sigma_i^B$. So, we have, $(\mathcal{E} \otimes \mathbb{I}) \, \sigma_{AB} = \sum_i p_i \; \mathcal{E}(\sigma_i^A) \otimes \mathbb{I} \, \sigma_i^B = \sum_i p_i \; \mathcal{E}(\sigma_i^A) \otimes \sigma_i^B$. Since, $\mathcal{E}$ is a positive operator, therefore $\mathcal{E}(\sigma_i^A) = \sigma'^A_i$ is also a valid quantum density matrix. Therefore, $(\mathcal{E} \otimes \mathbb{I}) \, \sigma_{AB} = \sigma'_{AB}$, where σ'_{AB} is a valid density matrix and therefore has non-negative eigenvalues.

Therefore, $(\mathcal{E} \otimes \mathbb{I}) \, \rho_{AB} < 0$, if and only if $\rho_{AB} \neq \sum_i p_i \; \sigma_i^A \otimes \sigma_i^B$, i.e., ρ_{AB} is entangled.

Alternative proof: If the state ρ_{AB} is entangled then we have $\mathrm{Tr}[W\rho_{AB}] < 0$, where we know that an entanglement witness W exists that detects entanglement. Now, we can write the bipartite Hermitian operator W acting on the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ as a Choi operator, such that: $W = \mathcal{J}(\mathcal{E}) = (\mathcal{E} \otimes \mathbb{I}) \, |vec(\mathbb{I})\rangle\langle vec(\mathbb{I})|$. This gives us,

$$\mathrm{Tr}[(\mathcal{E} \otimes \mathbb{I}) \, |vec(\mathbb{I})\rangle\langle vec(\mathbb{I})| \, \sigma_{AB}] < 0.$$

Since $\mathcal{J}(\mathcal{E})$ is Hermitian, $(\mathcal{E} \otimes \mathbb{I})|vec(\mathbb{I})\rangle\langle vec(\mathbb{I})| \, \rho_{AB} = \rho_{AB}|vec(\mathbb{I})\rangle\langle vec(\mathbb{I})|(\mathcal{E}^* \otimes \mathbb{I})$, and using the cyclicity of the trace operation, we have

$$\mathrm{Tr}[|vec(\mathbb{I})\rangle\langle vec(\mathbb{I})|(\mathcal{E} \otimes \mathbb{I})\ \rho_{AB}] < 0 \Longrightarrow \langle vec(\mathbb{I})|\ (\Phi \otimes \mathbb{I})\ \sigma_{AB}|vec(\mathbb{I})\rangle\ < 0.$$

Importantly, we should note that $\mathcal{E}$ cannot be completely positive, but only positive, since a completely positive $\mathcal{E}$ would imply a positive W, which cannot detect entanglement.

v) Peres-Horodecki criterion for two-qubit mixed entangled states

Armed with the necessary and sufficient condition for entangled bipartite mixed states in terms of positive, but not completely positive operators, we are equipped to dig further to find concrete measures and quantifiers of entanglement. However, it is important to clarify at this point that finding out whether a state is separable or not is a very difficult problem. It is impossible to find our whether a generic quantum ρ_{AB} state can be written in the form $\rho_{AB} = \sum_i p_i\ \sigma_i^A \otimes \sigma_i^B$, or the even harder problem of proving that all entanglement witnesses are positive or all positive operators have only nonnegative eigenvalues. This is a search problem over a huge set of parameters.

However, the problem is significantly simpler for two-qubit states as every positive operator $\mathcal{E}$ in the two-qubit Hilbert space can be written as: $\mathcal{E}(X) = \mathcal{E}_1(X) + \mathcal{E}_2(X^T)$, where $\mathcal{E}_1$ and $\mathcal{E}_2$ are a pair of completely positive operators and X^T is positive transpose operation. Therefore, all positive but not necessarily completely positive operators essentially boil down to the transpose operator. This can then be used to define a necessary and sufficient condition for detecting separability or entanglement, and is called the Peres-Horodecki criteria.

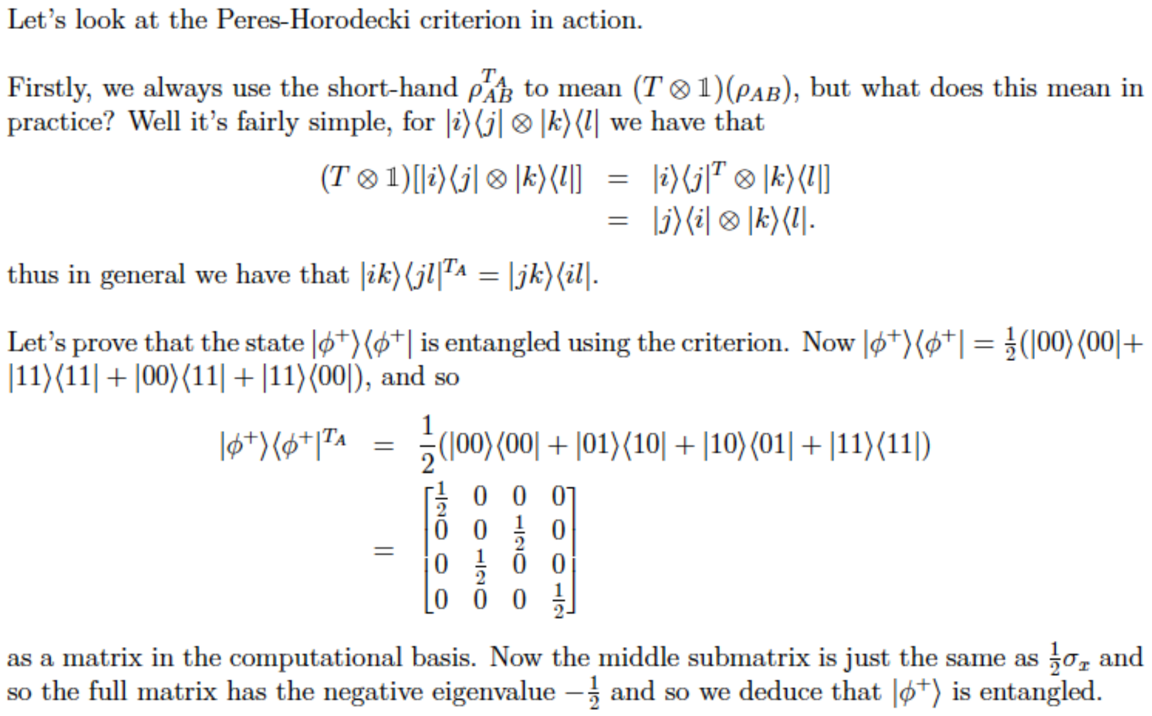
A two-qubit state ρ_{AB} is entangled if and only if $(T\ \otimes\ \mathbb{I})\ \rho_{AB}$ has a negative eigenvalue, where T is the positive transpose operator.

For qubits $\mathcal{E}(X) = \mathcal{E}_1(X) + \mathcal{E}_2(X^T)$, where $\mathcal{E}_1$ and $\mathcal{E}_2$ are completely positive maps. Therefore, ρ_{AB} is entangled if and only if there exist a positive operator $\mathcal{E}$, such that $(\mathcal{E}\ \otimes\ \mathbb{I})\ \rho_{AB} < 0$. For qubits, given the definition of $\mathcal{E}$, it is sufficient to show that $(T\ \otimes\ \mathbb{I})\ \rho_{AB}\ =\ \rho_{AB}^{T_A} < 0$, where T is the transpose operator.

The quantity $(T \otimes \mathbb{I})\,\rho_{AB} = \rho_{AB}^{T_A}$, is called the partial transpose of ρ_{AB}, with respect to the subsystem A. Importantly, the Peres-Horodecki criterion holds not only 2×2 dimensional system, but also 2×3 dimensional systems. However, beyond that the Peres-Horodecki criteria is no longer an "if and only if" statement, which means there can be entangled states for which $(\mathcal{E} \otimes \mathbb{I})\,\rho_{AB} \geq 0$. Such states are called "bound entangled states" and have a very interesting relation to distillable entanglement.

Exercise: Explain why for general system dimensions d, $\rho_{AB}^{T_A} < 0$ means that the state is entangled.

Exercise: Explain why one can equally have stated the partial transpose condition by applying the partial transpose to Bharat's side, i.e., $\rho_{AB}^{T_B} < 0$ means that the state is entangled. (Hint: does swapping the partial transpose change whether a state is separable or entangled?)



Let's look at the Peres-Horodecki criterion in action.

Firstly, we always use the short-hand $\rho_{AB}^{T_A}$ to mean $(T \otimes \mathbb{1})(\rho_{AB})$, but what does this mean in practice? Well it's fairly simple, for $|i\rangle\langle j| \otimes |k\rangle\langle l|$ we have that

$$\begin{aligned}(T \otimes \mathbb{1})[|i\rangle\langle j| \otimes |k\rangle\langle l|] &= |i\rangle\langle j|^T \otimes |k\rangle\langle l|] \\ &= |j\rangle\langle i| \otimes |k\rangle\langle l|.\end{aligned}$$

thus in general we have that $|ik\rangle\langle jl|^{T_A} = |jk\rangle\langle il|$.

Let's prove that the state $|\phi^+\rangle\langle\phi^+|$ is entangled using the criterion. Now $|\phi^+\rangle\langle\phi^+| = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 11| + |00\rangle\langle 11| + |11\rangle\langle 00|)$, and so

$$\begin{aligned}|\phi^+\rangle\langle\phi^+|^{T_A} &= \frac{1}{2}(|00\rangle\langle 00| + |01\rangle\langle 10| + |10\rangle\langle 01| + |11\rangle\langle 11|) \\ &= \begin{bmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{bmatrix}\end{aligned}$$

as a matrix in the computational basis. Now the middle submatrix is just the same as $\frac{1}{2}\sigma_x$ and so the full matrix has the negative eigenvalue $-\frac{1}{2}$ and so we deduce that $|\phi^+\rangle$ is entangled.

Exercise: Construct the entanglement witness W corresponding to the positive transpose operator.

Exercise: Show that $\rho_{AB}^{T_B}$ is another valid quantum state if and only if ρ_{AB} is separable. Also show that it is separable by applying the partial transposition to the other party.

vi) Computable measure of entanglement – Negativity

The partial transpose criterion provides us a clear method of detecting entanglement in mixed quantum states. For two qubit states, it can be computed using a measure called negativity. Given a bipartite state ρ_{AB}, its negativity is defined as

$$\mathcal{N}(\rho_{AB}) = \frac{1}{2}\sum_k [|\lambda_k| - \lambda_k],$$

where $\{\lambda_k\}$ are the eigenvalues of $\rho_{AB}^{T_A}$. Here, $\mathcal{N}(\rho_{AB})$ is simply the absolute sum of all negative eigenvalues of the partially transposed density matrix.

Another measure is the logarithm negativity, which is defined by:

$$E_{LN}(\rho_{AB}) = \log(2\mathcal{N}(\rho_{AB}) + 1).$$

Exercise: Plot $\mathcal{N}$ of $\rho_{AB}(t) = \cos^2 t\ |\phi^+\rangle\langle\phi^+| + \sin^2 t\ |\psi^+\rangle\langle\psi^+|$ as a function of t.

vii) Other computable measures of entanglement

a) Entanglement of formation and concurrence

The entanglement of formation (EOF) is the entanglement measure for bipartite quantum states ρ_{AB}, given by the convex roof construction, we discussed earlier:

$$E_F(\rho_{AB}) = \min_{\{p_i, |\varphi_i\rangle\}} \sum_i p_i\, E_F(|\varphi_i\rangle\langle\varphi_i|),$$

where, $\rho_{AB} = \sum_i p_i\, |\varphi_i\rangle\langle\varphi_i|$ is the pure-state decomposition and $E_F(|\varphi_i\rangle\langle\varphi_i|)$ is the pure state entanglement, which is the von Neumann entropy of entanglement. From an operational perspective, one can define a r*egularized* EOF, equal to the entanglement cost $E_c(\rho_{AB})$, which is defined as: $(|\psi^-\rangle\langle\psi^-|)_{AB}^{\otimes M} \longrightarrow \rho_{AB}^{\otimes N}$, as the minimum number of copies required to create the N copies of the state ρ_{AB}, in the asymptotic limit, i.e.,

$$E_C(\rho_{AB}) = \lim_{N\to\infty} \frac{M_{min}}{N} = \lim_{N\to\infty} \frac{E_F\left(\rho_{AB}^{\otimes N}\right)}{N}.$$

For pure states, this is the same as the reversible entanglement distillation (or dilution) $E_D(\rho_{AB})$. In general, for arbitrary quantum states, $E_F(\rho_{AB}) \geq E_D(\rho_{AB})$. However, the convex roof construction to compute the entanglement of formation for arbitrary ρ_{AB} is a hard task.

For two-qubit systems the EOF can be exactly computed using another entanglement monotone called concurrence, $C(\rho_{AB})$. For a pure state, say $|\phi\rangle$, the concurrence is defined as $C(|\phi\rangle) = |\langle\phi|\tilde{\phi}\rangle|$, where $|\tilde{\phi}\rangle$ is the spin flip operation: $|\tilde{\phi}\rangle = (\sigma_y \otimes \sigma_y)|\phi^*\rangle$, where $|\phi^*\rangle$ is the complex conjugate of $|\phi\rangle$ and $\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ is one the Pauli matrices. In the computational basis, it is easy to check that $|\tilde{\phi}\rangle$ is orthogonal to $|\phi\rangle$, if $|\phi\rangle$ is a product state. On the other hand, Bell states, $|\phi^{\pm}\rangle_{AB} = \frac{1}{\sqrt{2}}(|00\rangle_{AB} \pm |11\rangle_{AB})$ and $|\psi^{\pm}\rangle_{AB} = \frac{1}{\sqrt{2}}(|01\rangle_{AB} \pm |10\rangle_{AB})$, are left invariant. Therefore, $C(|\phi\rangle) = 0$ for product and $C(|\phi\rangle) = 1$ for maximally entangled Bell state. Hence, $C(|\phi\rangle) \neq 0$, for all entangled states. The EOF can be derived from the concurrence by the relation,

$$E_F(|\phi\rangle) = h(C(|\phi\rangle)), \text{ where } h(C) = H\left(\frac{1+\sqrt{1-C^2}}{2}\right),$$

and $H(x) = -x\log x - (1-x)\log(1-x)$ is the Shannon-entropy. Therefore, $C(|\phi\rangle)$ is an algebraic measure of entanglement.

For mixed states, one can again write a convex roof construction,

$$C(\rho_{AB}) = \min_{\{p_i, |\varphi_i\rangle\}} \sum_i p_i \, C(|\varphi_i\rangle).$$

Without proof, we show that the quantity $C(\rho_{AB})$ for mixed states is given by:

$$C(\rho_{AB}) = \max[0, \sqrt{\lambda_1} - \sqrt{\lambda_2} - \sqrt{\lambda_3} - \sqrt{\lambda_4}],$$

where, $\lambda_i's$ are the eigen values of the $\rho_{AB}\tilde{\rho}_{AB}$, arranged in decreasing order, and once again $\tilde{\rho}_{AB}$ results from the spin-flip operation, $\tilde{\rho}_{AB} = (\sigma_y \otimes \sigma_y)\,\rho_{AB}\,(\sigma_y \otimes \sigma_y)$.

The EOF can then be derived using the previous relation.

d) Distance based entanglement measures

The distance-based measures as suggested by the name measures the distance between state ρ and the closest state σ in the convex set $\mathcal{S}$ of separable states.

$$E_D(\rho) = \inf_{\sigma\in\mathcal{S}} D(\rho,\sigma).$$

The distance should be monotonic:

$$D(\rho,\sigma) \geq D\big(\mathcal{E}(\rho),\mathcal{E}(\sigma)\big).$$

Two such measures are: Relative entropy of entanglement: $D(\rho,\sigma) = S(\rho||\sigma)$, where $S(\rho||\sigma)$ is the relative entropy and the Bures entanglement: $D(\rho,\sigma) = 2 - 2\,F(\rho,\sigma)$, where $F(\rho,\sigma)$ is the quantum fidelity.

viii) Distillation of mixed state entanglement, positive partial transpose (PPT) states and bound entanglement

Note: This subsection is not necessary for exams.

We now return to look at distillation of singlets out of several copies of the bipartite mixed state, ρ_{AB}. In terms of the allowed set of local operation and classical communication (LOCC) protocols, we have $\rho_{AB}^{\otimes N} \longrightarrow (|\psi^-\rangle\langle\psi^-|)_{AB}^{\otimes M}$; where $|\psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$, is the maximally entangled singlet state. So, we can define the distillation of entanglement measure.

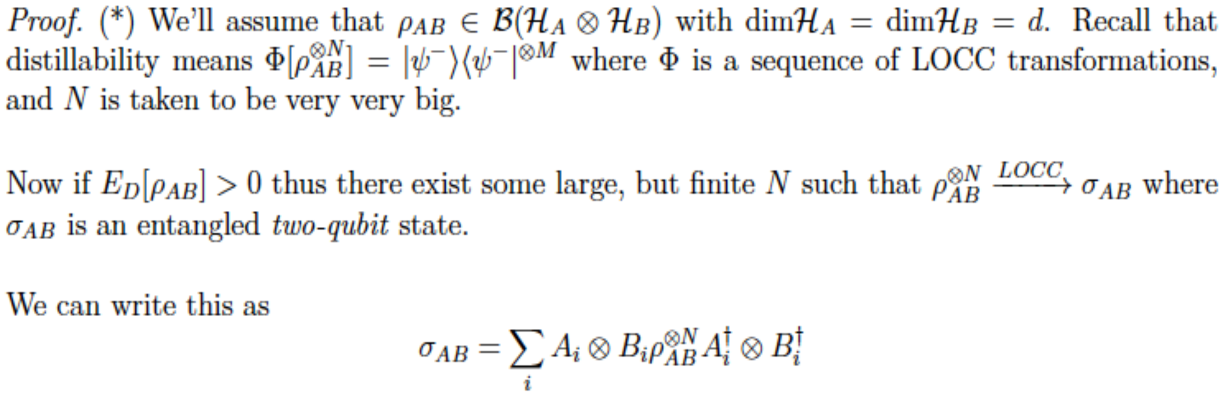
Given a quantum state ρ_{AB}, the entanglement of distillation $E_D(\rho_{AB})$ is the largest value of M/N attainable for the LOCC transformation $\rho_{AB}^{\otimes N} \longrightarrow (|\psi^-\rangle\langle\psi^-|)_{AB}^{\otimes M}$ as $N \longrightarrow \infty$.

The distillation of entanglement is in general a complex matter for mixed quantum states, however it is possible to say some general things about it. We saw that for pure states, if you build a state asymptotically with M singlets, then you can reversibly extract the same number of states through dilution, i.e., the distillation cost equals the formation cost. Does the same happen for mixed states? Or do we have states for which $E_D(\rho_{AB})$ is strictly less than the amount of entanglement required to build the mixed state? It turns out that this is exactly what does happen. States exist for which $E_D(\rho_{AB}) = 0$, even though the state is entangled, which is called *bound* entanglement. In the next section we connect distillability with the partial transpose criterion to find out more about these nonintuitive bound entangled states.

Here, we connect the two perspectives of entanglement that we have discussed so far. First, the operational perspective used to define LOCC protocols and asymptotic conversions, leading to the definition of distillable entanglement. Second, we looked at the convexity of separable states to define entanglement witnesses and then corresponding positive operators, we arrived at the partial transposition criterion to detect and quantify entanglement. Here, we establish some interesting connections between distillability of entanglement in mixed bipartite state and the set of bipartite states that have a positive partial transpose (PPT), which we call PPT states for simplicity. Note that for two-qubit systems the set of PPT states is nothing but the set of separable states, $\mathcal{D}_{sep}$.

We note that, regardless of system size, we measured entanglement in units of the gold standard of entanglement, which is a singlet. Now a singlet is a two-qubit state and so we should be able to connect general cases with the Peres-Horodecki criterion in some way. As discussed in the previous section, we defined the distillation of entanglement as the asymptotic conversion under LOCC, as $\rho_{AB}^{\otimes N} \longrightarrow (|\psi^-\rangle\langle\psi^-|)_{AB}^{\otimes M}$. We can now make the following statement (the proof is not part of assessment).

If $E_D(\rho_{AB}) > 0$ for a bipartite quantum state ρ_{AB} then we must have $\rho_{AB}^{T_A} \ngeq 0$. In other words, if a state has a positive partial transpose, $\rho_{AB}^{T_A} \geq 0$, then $E_D(\rho_{AB}) = 0$ and no entanglement can be distilled from it using LOCC in the asymptotic regime.



Proof. (*) We'll assume that $\rho_{AB} \in \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B)$ with $\dim\mathcal{H}_A = \dim\mathcal{H}_B = d$. Recall that distillability means $\Phi[\rho_{AB}^{\otimes N}] = |\psi^-\rangle\langle\psi^-|^{\otimes M}$ where Φ is a sequence of LOCC transformations, and N is taken to be very very big.

Now if $E_D[\rho_{AB}] > 0$ thus there exist some large, but finite N such that $\rho_{AB}^{\otimes N} \xrightarrow{LOCC} \sigma_{AB}$ where σ_{AB} is an entangled *two-qubit* state.

We can write this as

$$\sigma_{AB} = \sum_i A_i \otimes B_i \rho_{AB}^{\otimes N} A_i^\dagger \otimes B_i^\dagger$$

where A_i and B_i are mapping into single qubit spaces, and we will ignore issues of normalization for simplicity. Since σ_{AB} is entangled it must be that $\tilde{\sigma}_k = A_k \otimes B_k \rho_{AB}^{\otimes N} A_k^\dagger \otimes B_k^\dagger$ must be entangled for at least one value of k.

Now the operators A_k and B_k map from the large spaces $\mathcal{H}_A^{\otimes N}$ and $\mathcal{H}_B^{\otimes N}$ into 2-D spaces. Thus we can write them simply as

$$\begin{aligned} A_k &= |0\rangle_A\langle\psi_A| + |1\rangle_A\langle\phi_A| \qquad |\psi_A\rangle, |\phi_A\rangle \in \mathcal{H}_A^{\otimes N} \\ B_k &= |0\rangle_B\langle\psi_B| + |1\rangle_B|\phi_B\rangle \qquad |\psi_B\rangle, |\phi_B\rangle \in \mathcal{H}_B^{\otimes N.} \end{aligned}$$

However we can re-write $\tilde{\sigma}_k$ as

$$\tilde{\sigma}_k = A_k \otimes B_k \rho_{AB}^{\otimes N} A_k^\dagger \otimes B_k^\dagger = (A_k \otimes B_k)(\Pi_A \otimes \Pi_B \rho_{AB}^{\otimes N} \Pi_A \otimes \Pi_B) A_k^\dagger \otimes B_k^\dagger$$

where Π_A and Π_B are projectors onto the spaces spanned by $\{|\psi_A\rangle, |\phi_A\rangle\}$ and $\{|\psi_B\rangle, |\phi_B\rangle\}$, respectively. Since $A_k \otimes B_k$ cannot generate entanglement from nothing, we deduce that $(\Pi_A \otimes \Pi_B \rho_{AB}^{\otimes N}(\Pi_A \otimes \Pi_B)$ must be entangled.

Define orthonormal bases $\{|a_1\rangle, |a_2\rangle\}$ and $\{|b_1\rangle, |b_2\rangle\}$ for the two subspaces that Π_A and Π_B project onto. In particular, we have that

$$\begin{aligned} \Pi_A &= |a_1\rangle\langle a_1| + |a_2\rangle\langle a_2| \\ \Pi_B &= |b_1\rangle\langle b_1| + |b_2\rangle\langle b_2|. \end{aligned}$$

Extending these to a full basis for $\mathcal{H}_A^{\otimes N} \otimes \mathcal{H}_B^{\otimes N}$ we have that $(\Pi_A \otimes \Pi_B)\rho_{AB}^{\otimes N}(\Pi_A \otimes \Pi_B)$ has matrix

$$\begin{bmatrix} * & * & * & * & * & 0 & \cdots & 0 \\ * & * & * & * & * & 0 & \cdots & 0 \\ * & * & * & * & * & 0 & \cdots & 0 \\ * & * & * & * & * & 0 & \cdots & 0 \\ 0 & \cdots & & & & & & \\ \vdots & & & & & & & \end{bmatrix}$$

in this full basis.

Since the state is entangled, but is effectively a *two-qubit* state with support on the 4 dimensional subspace, we deduce that the partial transpose of the 4×4 sub-density matrix in (49) must have negative eigenvalue. That implies that $[\Pi_A \otimes \Pi_B \rho_{AB}^{\otimes N} \Pi_A \otimes \Pi_B]^{T_A} < 0$, and we have a state $|\kappa\rangle$ such that $\langle\kappa|[\Pi_A \otimes \Pi_B \rho_{AB}^{\otimes N} \Pi_A \otimes \Pi_B]^{T_A}|\kappa\rangle < 0$. That is equivalent to saying $\langle\varphi| \left(\rho_{AB}^{\otimes N}\right)^{T_A} |\varphi\rangle < 0$ which implies $\left[\rho_{AB}^{\otimes N}\right]^{T_A} \ngeq 0$. But $\left[\rho_{AB}^{\otimes N}\right]^{T_A} \ngeq 0$ if and only if $\rho_{AB}^{T_A} \ngeq 0$.

This means that if $E_D[\rho_{AB}] > 0$ then it must be that $\rho_{AB}^{T_A} \ngeq 0$. In other words, if a state is asymptotically distillable then the partial transpose must detect its entanglement. □

ix) Bound entanglement

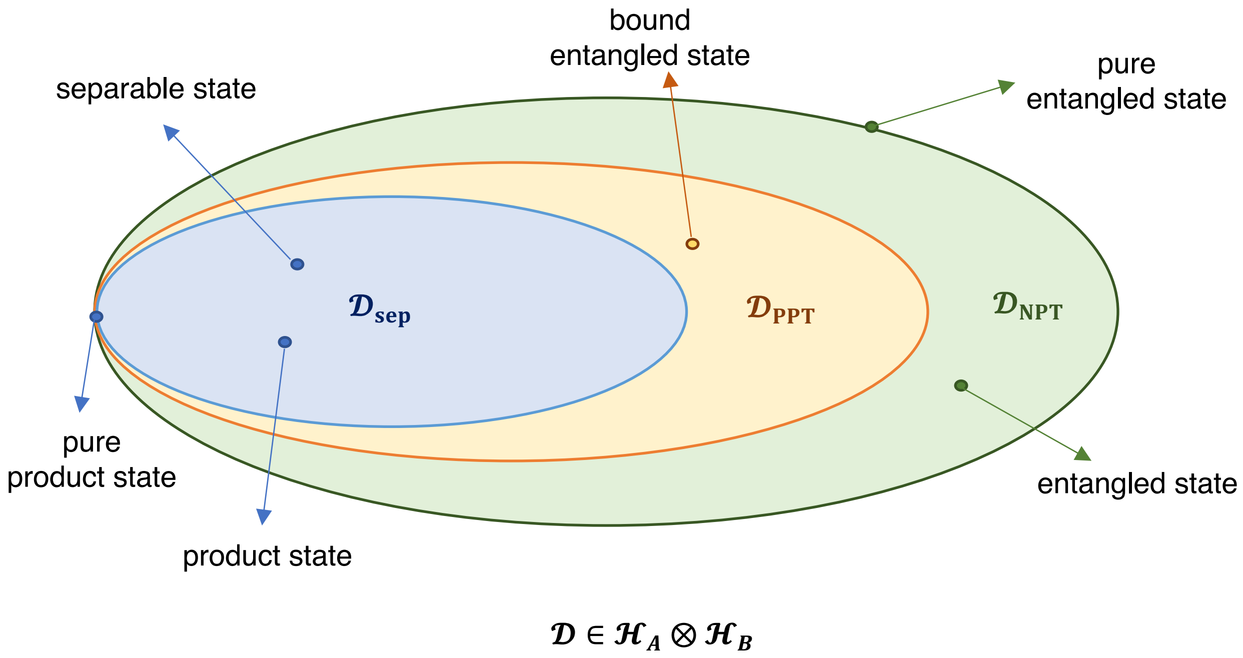
Note: This subsection is not necessary for exams.

Importantly, the previous statement and proof highlights an important gap that arises when we try to connect the two different perspectives of entanglement. The result shows that if we can asymptotically distill entanglement from multiple copies of the bipartite mixed state ρ_{AB}, then a single of the state must make this entanglement clear to see and easy to detect through the partial transpose operator, i.e., it must satisfy the condition, $(T \otimes \mathbb{I}_2)\,\rho_{AB} = \rho_{AB}^{T_A} < 0$. However, we know that beyond the Peres-Horodecki criterion for two qubit states (and dimension, $d = 2 \times 3$) the negative partial transpose criteria is not necessary for entangled states.

So, what happens if we discover an entangled state with positive partial transpose? Well, the previous result implies that we cannot distill any entanglement from it! In other words, if $\rho_{AB}^{T_A} \geq 0$, then we have that $E_D(\rho_{AB}) = 0$. The tricky question that now arises is: DO states exist for which it costs entanglement to build the state ρ_{AB}, but they have a positive partial transpose and therefore we can never get the entanglement back via distillation? The non-intuitive answer is YES, such entanglement is called bound entanglement, and its discovery was a big surprise.

Therefore, there exist quantum states ρ_{AB} such that $\rho_{AB} \neq \sum_i p_i\ \sigma_i^A \otimes \sigma_i^B$, i.e., ρ_{AB} is genuinely entangled, but still satisfies the condition $\rho_{AB}^{T_A} \geq 0$. It takes entanglement to build the state, but then this entanglement is locked in the state and is fully inaccessible under any asymptotic LOCC operations. The set of positive partial transpose (PPT) states as $\mathcal{D}_{\text{PPT}} = \{\rho_{AB}\colon \rho_{AB}^{T_A} \geq 0\}$, and is larger than the set $\mathcal{D}_{\text{sep}}$.

We now know that states in $\mathcal{D}_{\text{PPT}} - \mathcal{D}_{\text{sep}}$ are bound entanglement states, but we don't know if these are the only bound entangled states! Do non-distillable states exist with $\rho_{AB}^{T_A} \ngeq 0$? In other words, do we have $E_D(\rho_{AB}) = 0$ if and only if $\rho_{AB}^{T_A} \geq 0$?


bound
entangled state
separable state
pure
entangled state
$\mathcal{D}_{\text{sep}}$
$\mathcal{D}_{\text{PPT}}$
$\mathcal{D}_{\text{NPT}}$
pure
product state
entangled state
product state
$\mathcal{D} \in \mathcal{H}_A \otimes \mathcal{H}_B$