Tensor Norms for Quantum Entanglement

(Part 1)

Keywords: Quantum Entanglement, Tensor Norms, Separability, Nuclear Norm, Partial Transpose

Abstract

This document provides an introduction to the concept of quantum entanglement, beginning with an analogy to classical statistical independence and proceeding to formal definitions for pure and mixed quantum states. We explore the limitations of simple rank measures and introduce tensor norms, particularly the nuclear norm, as a more powerful tool for detecting entanglement. The utility of this approach is demonstrated through the Positive Partial Transpose (PPT) criterion, which provides a computable test for entanglement in mixed states.

1 Recall

1.1 Definition. A **State** ρ is a self-adjoint operator on a complex Hilbert space (e.g., $\mathbb{C}^d, L^2(\mathbb{R})$) where $d < \infty$, satisfying $\rho \geq 0$ and $\mathrm{Tr}(\rho) = 1$. **For this note, the Hilbert space will be $\mathbb{C}^d$.**

1.2 Definition. A **Pure State** is a rank-1 projection, i.e., $\rho = |\psi\rangle\langle\psi|$, where $\langle\psi|\psi\rangle = 1$ and $|\psi\rangle \in \mathbb{C}^d$.

1.3 Definition. A **Mixed State** is a convex combination of pure states, $\rho = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$.

2 Entanglement Testing

- Classical notion of entanglement (dependence)
- Entanglement of pure states
- Entanglement of mixed states
  - Tensor norms
- Story of norms
- Graphical tensor notation in Quantum Computing

2.1 Classical Notion of entanglement (dependence)

Consider X, Y discrete random variables with finite support.

2.1 Definition. "X, Y are **Independent**" if for $x_i \in \mathrm{Range}(X)$, $y_j \in \mathrm{Range}(Y)$:

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j)$$

"X, Y are **Dependent**" if not independent.

Goal
From the matrix $M_{ij} = P(X = x_i, Y = y_j)$, decide whether X, Y are independent.

2.2 Observation. If X, Y are independent, then $M = \mathbf{u}\mathbf{v}^T$ where
$\mathbf{u} = (P(X = x_1), P(X = x_2), \ldots)^T$
$\mathbf{v} = (P(Y = y_1), P(Y = y_2), \ldots)^T$.
M is of rank 1.

2.3 Theorem. *X, Y are independent $\iff$ M has rank 1.*

Proof. □

($\implies$) Done (by Observation 2.2).

($\impliedby$) If M has rank 1, then $M = \mathbf{u}\mathbf{v}^T$. Since M represents a joint probability distribution with marginals $P(X = \cdot)$ and $P(Y = \cdot)$, we get that

- $u_i v_j \geq 0$
- $\sum_i \sum_j u_i v_j = 1$
- $u_i \sum_j v_j = P(X = x_i)$
- $v_j \sum_i u_i = P(Y = y_j)$

Hence $P(X = x_i)P(Y = y_j) = u_i v_j = M_{ij}$, showing X, Y are independent.

3 Entanglement of Pure States

Let $\mathcal{H}_X = \mathrm{span}\{|x_1\rangle, \ldots, |x_m\rangle\}$ and $\mathcal{H}_Y = \mathrm{span}\{|y_1\rangle, \ldots, |y_n\rangle\}$. We can map the classical probability matrix M to a quantum state vector:

$$|\psi\rangle = \sum_{i,j} \sqrt{P(X = x_i, Y = y_j)}|x_i\rangle \otimes |y_j\rangle$$

3.1 Proposition. $M = uv^T$ iff $|\psi\rangle = |\psi_X\rangle \otimes |\psi_Y\rangle$ for some state vectors $|\psi_X\rangle \in \mathcal{H}_X, |\psi_Y\rangle \in \mathcal{H}_Y$.

This means that X and Y are independent iff the state vector $|\psi\rangle$ is a pure tensor.

3.2 Definition (Tensor Rank Decomposition). Given Hilbert Spaces $\mathcal{H}_1, \ldots, \mathcal{H}_k$, a state vector $|\psi\rangle \in \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_k$ can be written as:

$$|\psi\rangle = \sum_{a=1}^{r} |\varphi_1^{(a)}\rangle \otimes |\varphi_2^{(a)}\rangle \otimes \cdots \otimes |\varphi_k^{(a)}\rangle$$

3.3 Definition (Tensor Rank). $R(|\psi\rangle) = \min\{r \geq 1 \mid |\psi\rangle \text{ has a decomposition into } r \text{ pure tensors}\}$.

- $R(|\psi\rangle) = 1$ "$|\psi\rangle$ is a separable pure state"
- $R(|\psi\rangle) > 1$ "$|\psi\rangle$ is an entangled pure state"

3.1 Examples (Computation of Tensor Rank)

① Bipartite Quantum States ($k = 2$)

If $|\psi\rangle = \sum_{ij} \Psi_{ij}|x_i\rangle \otimes |y_j\rangle$, then its tensor rank $R(|\psi\rangle)$ is equal to the matrix rank of the coefficient matrix $\Psi = (\Psi_{ij})$. This is also known as the Schmidt rank.

3.4 Caution. In our classical-to-quantum mapping, $R(|\psi\rangle) = \mathrm{rank}((\sqrt{M_{ij}}))$, which is not equal to $\mathrm{rank}(M)$ in general. However, one of them has rank 1 iff the other does. Hence, either can be used for testing entanglement.

In practice, we can compute the rank of a matrix using Singular Value Decomposition (SVD). Let's recall its definition as we will need it later as well.

3.5 Definition (Singular Value Decomposition). For a bounded linear operator M on $\mathcal{H}^d$, there exist orthonormal bases $\{|\alpha_i\rangle\}, \{|\beta_i\rangle\}$ and singular values $\sigma_1 \geq \cdots \geq \sigma_r > 0$ ($r = \mathrm{rank}(M)$) such that:

$$M = \sum_{i=1}^{r} \sigma_i |\alpha_i\rangle\langle\beta_i|$$

The singular values are unique upto permutation.

② The GHZ state - $|\mathrm{GHZ}\rangle = \frac{1}{\sqrt{2}}|000\rangle + \frac{1}{\sqrt{2}}|111\rangle$

The decomposition shows $R(|\mathrm{GHZ}\rangle) \leq 2$. To show it is entangled, we assume it is separable for contradiction, i.e., $R(|\mathrm{GHZ}\rangle) = 1$. Then $|\mathrm{GHZ}\rangle = |\phi_A\rangle \otimes |\phi_B\rangle \otimes |\phi_C\rangle$. The density operator would be $\rho_{\mathrm{GHZ}} = \rho_A \otimes \rho_B \otimes \rho_C$. Note that

$$\rho_{\mathrm{GHZ}} = \frac{1}{2}(|000\rangle\langle000| + |111\rangle\langle111| + |000\rangle\langle111| + |111\rangle\langle000|)$$

while $\rho_A \otimes \rho_B \otimes \rho_C = |\phi_A\phi_B\phi_C\rangle\langle\phi_A\phi_B\phi_C|$. Now it's tempting to apply the uniqueness of SVD to conclude that the LHS has non zero singular values $\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}$, while the RHS has non zero singular value 1, giving a contradiction. However, note that the vectors $|000\rangle, |111\rangle, |000\rangle, |111\rangle$ are not orthogonal, so we cannot apply the uniqueness of SVD directly.

However, we can "cheat" a bit by rearranging the last two qubits between the input and output, as follows:

$$\frac{1}{2}|000\rangle\langle000| + \frac{1}{2}|111\rangle\langle111| + \frac{1}{2}|011\rangle\langle100| + \frac{1}{2}|100\rangle\langle011| = |\phi_A\phi_B\phi_C\rangle\langle\phi_A\phi_B\phi_C|$$

Now we can safely apply the uniqueness of SVD, as the vectors $|000\rangle, |111\rangle, |011\rangle, |100\rangle$ are orthogonal. This gives us a contradiction, hence $R(|\mathrm{GHZ}\rangle) > 1$.

The trick we used above is called the **partial transpose**, which we will come back to later while discussing entanglement testers.

3.1 Remark. Tensor rank is called a discrete norm. It satisfies some but not all properties of a norm: $R(|\psi\rangle) \geq 0$, $R(|\psi\rangle) = 0 \iff |\psi\rangle = 0$, and $R(|\psi\rangle + |\phi\rangle) \leq R(|\psi\rangle) + R(|\phi\rangle)$. However, it fails positive homogeneity, as $R(\lambda|\psi\rangle) = R(|\psi\rangle)$ for $\lambda \neq 0$.

4 Entanglement of Mixed Quantum States

We now consider general density operators ρ.

4.1 Definition (Separable Mixed State). A density operator $\rho \in \mathcal{M}_{d_1} \otimes \cdots \otimes \mathcal{M}_{d_k}$ is **separable** if it can be written as a convex combination of product density operators:

$$\rho = \sum_i p_i(\rho_i^{(1)} \otimes \rho_i^{(2)} \otimes \cdots \otimes \rho_i^{(k)})$$

where $p_i > 0, \sum p_i = 1$, and each $\rho_i^{(j)}$ is a density operator.

4.2 Clarifying Notation (Digression). The notation $\rho_i^{(1)} \otimes \rho_i^{(2)}$ means the tensor product of operators. If $\rho_i^{(1)} = |\psi_i^{(1)}\rangle\langle\psi_i^{(1)}|$ and $\rho_i^{(2)} = |\psi_i^{(2)}\rangle\langle\psi_i^{(2)}|$, then their product is $|\psi_i^{(1)}\rangle\langle\psi_i^{(1)}| \otimes |\psi_i^{(2)}\rangle\langle\psi_i^{(2)}|$, which lives in $\mathbb{C}^{d_1} \otimes (\mathbb{C}^{d_1})^* \otimes \mathbb{C}^{d_2} \otimes (\mathbb{C}^{d_2})^*$. But it can also be viewed as an element of $\mathbb{C}^{d_1 d_2} \otimes (\mathbb{C}^{d_1 d_2})^*$ using the notation $|\psi_i^{(1)}\psi_i^{(2)}\rangle\langle\psi_i^{(1)}\psi_i^{(2)}|$, or as an element of $\mathbb{C}^{d_1 d_2} \otimes \mathbb{C}^{d_1 d_2}$ using the notation $|\psi_i^{(1)}\psi_i^{(2)}\rangle \otimes |\psi_i^{(1)}\psi_i^{(2)}\rangle$.

Question: How useful is (matrix) rank for mixed quantum states?

Example State (ρ)	Rank of ρ
$\|00\rangle\langle00\|$	1
$\|01\rangle\langle01\|$	1
$\|10\rangle\langle10\|$	1
$\|11\rangle\langle11\|$	1
$\frac{1}{4}(\|00\rangle\langle00\| + \|01\rangle\langle01\| + \|10\rangle\langle10\| + \|11\rangle\langle11\|)$	4

But the last state is separable by definition. This demonstrates that rank is not a good measure of entanglement for mixed states: it does not behave well with respect to taking convex combinations.

5 Criteria for Entanglement

5.1 Desirable Properties of an Entanglement Measure

We seek a function Φ that can test for entanglement. A key property is convexity. If $\rho = \sum_i p_i \rho_i$, we want $\Phi(\rho) \leq \sum_i p_i \Phi(\rho_i)$. If Φ is convex and normalized such that $\Phi(\rho_{prod}) \leq C$ for any product state, then for any separable state ρ_{sep}:

$$\Phi(\rho_{sep}) = \Phi\left(\sum_a p_a \rho_{prod,a}\right) \leq \sum_a p_a \Phi(\rho_{prod,a}) \leq \sum_a p_a \cdot C = C$$

Thus, if we find $\Phi(\rho) > C$, the state ρ must be entangled. This motivates the use of norms.

5.1 Definition (Convex Envelope). The **convex envelope** of a function f on a convex set C is the largest convex function g such that $g(x) \leq f(x)$ for all $x \in C$.

5.2 Theorem. *The convex envelope of the rank function $R(A)$ (on the set of matrices with operator norm $\|A\|_{S_\infty} \leq 1$) is the nuclear norm $\|A\|_{S_1} = \sum \sigma_i(A)$. (Sometimes, denoted by $\|A\|_*$)*

5.3 Lemma (Variational characterization of Nuclear Norm). For a bounded linear operator A on $\mathcal{H}^d$:

$$\|A\|_{S_1} = \max_{U \text{ unitary}} |\mathrm{Tr}(AU)|$$

Proof sketch.

- First, recall that $A = \sum_{i=1}^r \sigma_i |\alpha_i\rangle\langle\beta_i|$ for some orthonormal bases. Then $|\mathrm{Tr}(AU)| = |\sum_{i=1}^r \sigma_i \langle\beta_i|U\alpha_i\rangle| \leq \sum_{i=1}^r \sigma_i \|\beta_i\|\|U\alpha_i\|$ by Cauchy-Schwarz.
- Then for U unitary, the RHS is just $\sum_{i=1}^r \sigma_i \|\beta_i\|\|\alpha_i\| = \sum_i^r \sigma_i$
- Finally, equality is achieved by taking $U = A(A^*A)^{-1/2}$

5.4 Theorem (Nuclear Norm is a Norm). The nuclear norm $\|\cdot\|_{S_1}$ satisfies all properties of a norm.

Proof sketch The non trivial part is triangle inequality. Using the variational characterization from **Lemma 5.3**, for any unitary operator U :

$$|\mathrm{Tr}((A + B)U)| \leq |\mathrm{Tr}(AU)| + |\mathrm{Tr}(BU)|$$

Taking max over U and using sub-additivity of max gives $\|A + B\|_{S_1} \leq \|A\|_{S_1} + \|B\|_{S_1}$.

6 The PPT Entanglement Test

Now we come to the reason why we introduced the nuclear norm: entanglement testing.

6.1 Theorem (Entanglement Test using Partial Transpose). *Let ρ be a $d_A \times d_B$ dimensional bipartite density operator. If $\|\rho^{\Gamma_B}\|_{S_1} > 1$, then ρ is entangled.*

Proof. Next day. □