Selected topics in Mathematical Physics: Quantum Information Theory

Almost all pure quantum states are almost maximally entangled

Thomas Gläßle

December 3, 2013

Abstract

It is shown that in a bipartite system almost all pure quantum states are close to maximally entangled. The prior result about the expected trace of the squared local density operator is employed to estimate the expected distinguishability of the local quantum state from the maximally mixed state. Levy's Lemma is then used to prove that most states are near the expectation value.

1 Introduction

Consider pure states $|\psi\rangle \in A \otimes B$ in a bipartite system. Denote the partial trace over one of the subsystems $\rho_A := \operatorname{tr}_B |\psi\rangle\langle\psi|$. It has been shown that the expectation value of this local density matrix can be written in terms of the dimensions of the constituent systems

$$\mathbb{E}_\psi \operatorname{tr} \rho_A^2 = \frac{|A| + |B|}{|A| \cdot |B| + 1}. \tag{1}$$

For small subsystems $|A| \ll |B|$ this implies

$$\mathbb{E}_\psi \operatorname{tr} \rho_A^2 \approx \frac{1}{|A|}. \tag{2}$$

This leads to the suspection that most states are almost maximally entangled. However, this is not an obvious conclusion from Eq. (2). First, observe that the square of a density matrix $\operatorname{tr} \rho_A^2$ does not have an immediate physical interpretation. From this result alone it is unclear how to quantify the physical distance between the expected local state and the maximally mixed state. Second, note that Eq. (1) makes no statement about higher order moments. It does not answer how probable it is to find random states with $\operatorname{tr} \rho_A^2$ close by or far from the expectation value.

To fix the first issue a physically meaningful distance measure will be introduced in the following. The result Eq. (1) can be used to give an estimate for the distance between local state and maximally mixed state. Using Levy's Lemma will allow to restrict the probability for finding a random state with a larger-than-expected distance to be exponentially low.

2 Distance measures

Recall that for vectors $\boldsymbol{x} \in \mathbb{R}^n$ one can define their norm by

$$\|\boldsymbol{x}\|_p = \sqrt[p]{\sum_{i=1}^{n} |x_i|^p}. \tag{3}$$

To generalize this expression for matrices $M \in M(m,n)$ one can start by defining their modulus

$$|M| := \sqrt{M^\dagger M}.$$

Note that $M^\dagger M$ is a self adjoint and positive semi-definite $n \times n$ matrix. Hence, it can be diagonalized with non-negative real eigenvalues. Its square root can be defined via the positive square roots of its eigenvalues, i.e.

$$|M| = U^\dagger \begin{pmatrix} \sqrt{\sigma_1} & & \\ & \ddots & \\ & & \sqrt{\sigma_n} \end{pmatrix} U.$$

Definition The p-norm of a matrix is

$$\|M\|_p := \sqrt[p]{\operatorname{tr}|M|^p} \tag{4}$$

Denote $\boldsymbol{\sigma} = \left(\sqrt{\sigma_i}\right)_i$ the vector of eigenvalues of $|M|$. It becomes clear how Eq. (4) is indeed a generalization of Eq. (3) for matrices:

$$\|M\|_p = \|\boldsymbol{\sigma}\|_p.$$

2.1 The matrix 2-norm

The 2-norm is the best suited norm for many computations. One example making use of the 2-norm has already been mentioned. Eq. (1) contains the trace of the squared density matrix. That this corresponds to its 2-norm can be seen when taking $M = M^\dagger$ a self-adjoint matrix. Eq. (4) then becomes

$$\begin{aligned} \|M\|_2 &= \sqrt{\operatorname{tr} M^\dagger M} \\ &= \sqrt{\operatorname{tr} M^2} \\ &= \sqrt{\sum_{i,j} |M_{ij}|^2}. \end{aligned} \tag{5}$$

This is just the norm of all the matrix components.

2.2 The matrix 1-norm and trace distance

Another useful case of the p-norm is the 1-norm. As shown in the following it corresponds to the so called trace distance, which is a capable measure of distinguishability for quantum states.

Definition The trace distance of two density matrices ρ, σ can be defined as

$$D(\rho, \sigma) = \max_P \{\operatorname{tr}(P\rho) - (P\sigma)\} \tag{6}$$

where the maximization is over all projectors or alternatively over all POVMs P.

Physical interpretation Consider a two-outcome experiment using a projective measurement P. When starting with the state ρ the probability to measure 1 is given by $\operatorname{tr} P\rho$ and for the state σ by $\operatorname{tr} P\sigma$, respectively. This leads to the following clear interpretation of Eq. (6): The trace distance of ρ and σ is a measure for the distinguishability of the states ρ and σ using the most distinguishing measurement P.

$$D(\rho, \sigma) = \begin{cases} 0 & \rho, \sigma \text{ indistinguishable} \\ 1 & \rho, \sigma \text{ perfectly distinguishable} \end{cases}$$

Lemma The trace distance of two density matrices can be expressed as the 1-norm of their difference:

$$D(\rho, \sigma) = \frac{1}{2} \|\rho - \sigma\|_1. \tag{7}$$

Proof for the case of maximizing over projection operators P. Since ρ and σ are density matrices their difference $\rho - \sigma$ is self-adjoint. Therefore it can be diagonalized with real eigenvalues:

$$\begin{aligned} \rho - \sigma &= U^\dagger D U \\ &= U^\dagger (D_+ - D_-) U \\ &= Q - S, \end{aligned} \tag{8}$$

such that Q and S are positive semi-definite operators with orthogonal non-zero eigenspaces. The modulus $|\rho - \sigma|$ contains the absolute values of D, thus

$$|\rho - \sigma| = Q + S. \tag{9}$$

Since $\operatorname{tr} \rho = \operatorname{tr} \sigma = 1$ it follows from the definition Eq. (8) of Q and S that $\operatorname{tr} Q = \operatorname{tr} S$. Now denote P_Q the projection operator onto the non-zero eigenspaces of Q. Then

$$\begin{aligned} \frac{1}{2} \|\rho - \sigma\|_1 &= \frac{1}{2} \operatorname{tr} |\rho - \sigma| \\ &= \frac{1}{2} \operatorname{tr}(Q + S) \\ &= \operatorname{tr} Q \\ &= \operatorname{tr} P_Q Q \\ &= \operatorname{tr} P_Q (Q - S) \\ &= \operatorname{tr} P_Q (\rho - \sigma). \end{aligned} \tag{10}$$

On the other hand when using an arbitrary projection operator P the trace contribution of the PQ term cannot be increased, while the contribution of the PS term cannot be decreased. This is due to the fact that Q and S are non-negative and have orthogonal non-zero eigenspaces. Therefore

$$\begin{aligned} \operatorname{tr} P_Q (Q - S) &\geq \operatorname{tr} P (Q - S) \\ &= \operatorname{tr} P (\rho - \sigma). \end{aligned} \tag{11}$$

This, together with Eq. (10), proves the Lemma.

Lemma The trace distance is non-increasing under trace preserving operators ϕ.

$$D(\phi(\rho), \phi(\sigma)) \leq D(\rho, \sigma) \tag{12}$$

Proof Using the decomposition $\rho - \sigma = Q - S$ as before. Then

$$\begin{aligned} D(\rho, \sigma) &= \frac{1}{2}\operatorname{tr}|\rho - \sigma| \\ &= \operatorname{tr}(Q) \\ &= \operatorname{tr}(\phi(Q)) \\ &\geq \operatorname{tr}(P\phi(Q)) \\ &\geq \operatorname{tr}(P(\phi(Q) - \phi(S)) \\ &= \operatorname{tr}(P(\phi(\rho) - \phi(\sigma)) \\ &= D(\phi(\rho), \phi(\sigma)), \end{aligned} \tag{13}$$

for the appropriate projection operator P.

Corollary Consider a composite system $A \otimes B$. The partial trace $\phi : \rho \mapsto \rho_A = \operatorname{tr}_B(\rho)$ is a trace preserving operator. For two states ρ and σ it follows

$$D(\rho_A, \sigma_A) \leq D(\rho, \sigma). \tag{14}$$

This fits the physical intuition: Looking at only one constituent of a composite system removes part of the information and can therefore not increase the distinguishability of two quantum states.

Lemma The trace distance can be estimated from the 2-norm. Given two $n \times n$ matrices ρ,σ it follows that

$$\|\rho - \sigma\|_1 \leq \sqrt{n}\|\rho - \sigma\|_2. \tag{15}$$

In general the 2-norm is much easier to calculate than the 1-norm. With this equivalence, however, it is possible to transfer any results to the physically more useful trace distance.

3 Local distinguishability from the maximally mixed state

3.1 Expected distance

Using the trace distance as measure of distinguishability a more meaningful version of Eq. (2) can now be formulated.

Theorem Consider pure states in a bipartite system $|\psi\rangle \in A \otimes B$. The expected trace distance from the maximally mixed state locally obeys the relation

$$\mathbb{E}_\psi \left\| \rho_A - \frac{\mathbb{1}_A}{|A|} \right\|_1 \leq \sqrt{\frac{|A|}{|B|}}. \tag{16}$$

Proof Using only Eq. (1) the square Euclidean distance of ρ_A to the maximally mixed state can be calculated

$$\begin{aligned}
\mathbb{E}_\psi \left\| \rho_A - \frac{\mathbb{1}_A}{|A|} \right\|_2^2 &= \mathbb{E}_\psi \operatorname{tr} \left(\rho_A - \frac{\mathbb{1}_A}{|A|} \right)^2 \\
&= \mathbb{E}_\psi \operatorname{tr} \left(\rho_A^2 - \frac{2}{|A|}\rho_A + \frac{\mathbb{1}_A}{|A|^2} \right) \\
&= \frac{|A| + |B|}{|A| \cdot |B| + 1} - \frac{1}{|A|} \\
&\leq \frac{|A| + |B|}{|A| \cdot |B|} - \frac{1}{|A|} \\
&= \frac{1}{|B|}.
\end{aligned} \tag{17}$$

At this point Jensen's inequality can be used to remove the undesired square root. It states that for convex functions ϕ and random variables X the expectation value fulfills

$$\phi(\mathbb{E}_X[X]) \leq \mathbb{E}_X[\phi(X)]. \tag{18}$$

The function $\phi : x \mapsto -\sqrt{x}$ is convex. Thus

$$\begin{aligned}
\mathbb{E}_\psi \left\| \rho_A - \frac{\mathbb{1}_A}{|A|} \right\|_1 &\leq \sqrt{|A|}\, \mathbb{E}_\psi \left\| \rho_A - \frac{\mathbb{1}_A}{|A|} \right\|_2 \\
&\leq \sqrt{|A|} \sqrt{\mathbb{E}_\psi \left\| \rho_A - \frac{\mathbb{1}_A}{|A|} \right\|_2^2} \\
&\leq \sqrt{\frac{|A|}{|B|}}.
\end{aligned} \tag{19}$$

In the first step the relation Eq. (15) between 1-norm and 2-norm was used. This finishes the proof for Eq. (16).

3.2 Number of states near the expectation value

With the above findings and Levy's Lemma the probability to find states near the expectation value can be estimated. Recall that for a slowly varying function on the unit sphere, Levy's Lemma gives an estimate how much measure is contained within an ϵ surrounding of its expectation value.

Levy's Lemma Let $\phi : S^{2n-1} \to \mathbb{R}$ be a Lipschitz continuous function on the unit sphere, i.e. $|\phi(\boldsymbol{x}) - \phi(\boldsymbol{y})| \leq \eta \cdot \|\boldsymbol{x} - \boldsymbol{y}\|_2$. Then

$$\operatorname{Prob}\left\{ |\phi(\boldsymbol{x}) - \mathbb{E}_{\boldsymbol{x}}\phi| \geq \epsilon \right\} \leq 2 \exp\left(\frac{2n\epsilon^2}{9\pi^3\eta^2} \right). \tag{20}$$

Lemma The local trace distance

$$\phi : |\psi\rangle \mapsto \left\| \rho_A - \frac{\mathbb{1}_A}{|A|} \right\|_1 \tag{21}$$

is Lipschitz continuous.

Proof First note that for pure states $\rho = |\psi\rangle\langle\psi|$, $\sigma = |\varphi\rangle\langle\varphi|$ their trace distance assumes the nice form

$$D(\rho,\sigma)^2 = 1 - |\langle\psi|\varphi\rangle|^2. \tag{22}$$

A proof can be found in [1]. Then

$$\begin{aligned} D(\rho,\sigma)^2 &= (1 - |\langle\psi|\varphi\rangle|)(1 + |\langle\psi|\varphi\rangle|) \\ &\leq 2(1 - |\langle\psi|\varphi\rangle|) \\ &\leq 2(1 - \Re\langle\psi|\varphi\rangle) \\ &= \langle\psi - \varphi|\psi - \varphi\rangle \\ &= \big\||\psi\rangle - |\varphi\rangle\big\|^2. \end{aligned} \tag{23}$$

By the triangular inequality

$$\begin{aligned} |\phi(\psi) - \phi(\varphi)| &= \left| \left\| \rho_A - \frac{\mathbb{1}_A}{|A|} \right\|_1 - \left\| \sigma_A - \frac{\mathbb{1}_A}{|A|} \right\|_1 \right| \\ &\leq \left\| \left(\rho_A - \frac{\mathbb{1}_A}{|A|} \right) - \left(\sigma_A - \frac{\mathbb{1}_A}{|A|} \right) \right\|_1 \\ &= \|\rho_A - \sigma_A\|_1 \\ &= 2D(\rho_A, \sigma_A) \\ &\leq 2D(\rho,\sigma) \\ &\leq 2\big\||\psi\rangle - |\varphi\rangle\big\|. \end{aligned} \tag{24}$$

This proves that the local trace distance is in fact Lipschitz continuous with $\eta = 2$.

Theorem The number of states with a locally large distance from the maximally mixed state decreases exponentially as

$$\mathrm{Prob}\left\{ \left\| \rho_A - \frac{\mathbb{1}_A}{|A|} \right\|_1 \geq \sqrt{\frac{|A|}{|B|}} + \epsilon \right\} \leq 2\exp\left(\frac{|A| \cdot |B| \epsilon^2}{18\pi^3} \right). \tag{25}$$

Identifying the state space $A \otimes B \equiv \mathbb{R}^{2n} \supset S^{2n-1}$, the proof follows directly from Levy's Lemma and the above shown Lipschitz continuity of the local trace distance.

Thus, if a state $|\psi\rangle \in A \otimes B$ is randomly selected from the set all states and $|A| \ll |B|$, then ρ_A is highly probable to be almost indistinguishable from the maximally mixed state. In other words, it is found with high probability that

$$D\left(\rho_A, \frac{\mathbb{1}_A}{|A|} \right) \ll 1. \tag{26}$$

3.3 Fannes' inequality

Suppose ρ and σ are density matrices with $D(\rho,\sigma) \leq 1/e$. Then, Fannes' inequality relates their entropy difference via their trace distance

$$|S(\rho) - S(\sigma)| \leq D(\rho,\sigma) \cdot (\log n - \log D(\rho,\sigma)). \tag{27}$$

For a short proof, see [1].

Application Apply this to the case where a state $|\psi\rangle \in A \otimes B$ is randomly drawn from a bipartite system with $|A| \ll |B|$. With high probability, $D\left(\rho_A, \frac{\mathbb{1}_A}{|A|}\right) \ll 1$. Thus, the difference in entropy for ρ_A and $\frac{\mathbb{1}_A}{|A|}$ becomes small

$$\left|S(\rho_A) - S\left(\frac{\mathbb{1}_A}{|A|}\right)\right| \leq \epsilon. \tag{28}$$

In other words, ρ_A has almost maximal entropy

$$S(\rho_A) \geq \log|A| - \epsilon. \tag{29}$$

Audenaert's improvement Audenaert has established an improved version of Fannes' inequality that holds true for arbitrary $D(\rho, \sigma)$. Furthermore his version gives the optimal bounds. The improved statement is

$$|S(\rho_A) - S(\sigma_A)| \leq D(\rho, \sigma) \log(n-1) + H(\{D(\rho, \sigma), 1 - D(\rho, \sigma)\}), \tag{30}$$

where $H(\{p_i\})$ is the Shannon entropy. The proof is given in [2].

4 Conclusion

When looking at small subsystems, most pure states are locally almost indistinguishable from the maximally mixed. Furthermore the number of states with considerably higher distinguishability decreases exponentially. It can be noted that this finding reappears in the expectation that the entropy of the subsystem will be almost maximal.

References

[1] Michael A. Nielsen and Isaac L. Chuang. *Quantum computation and quantum information*. Cambridge Univ. Press, Cambridge [u.a.], 9. print. edition, 2007.

[2] Koenraad M R Audenaert. A sharp continuity estimate for the von neumann entropy. *Journal of Physics A: Mathematical and Theoretical*, 40(28):8127, 2007.