Lecture 13: October 22, 2024

Scribe: Louis Marquis and Ruohan Shen *More on Random States*

Unfortunately, there was no the audio in the recording of the lecture, so some details explained verbally might be missed in these notes. This lecture will continue on random states where the previous lecture began. Random states are useful to generate states with desireable properties when the probability of not having that property is low-dimensional. Low-dimensional, in this case, means that the probability is essentially zero. For example, non-invertibility of a matrix is low-dimensional, as it requires the determinant to be exactly zero, while there are essentially infinite real values it could take. This lecture covers various properties of random states, including moments, anticoncentration, Renyi entropy, and entanglement.

13.1 Moments of Random States

As previously introduced, the the nth moment of a distribution over states is the expected value of the density matrix of n copies of the state. A very important such distribution is the uniform distribution over all unit-1 complex vectors (we will write this condition as $|u\rangle \in \mathbb{C}^d$, with the unit-1 implicit). We can calculate the nth moment of this random state.

$$E_{|u\rangle \in \mathbb{C}^d} |u\rangle \langle u|^{\otimes n} = \frac{\Pi_{sym}^{(n,d)}}{\text{tr}(\Pi_{sym})} \tag{13.1}$$

$$= \frac{\frac{1}{n!} \sum_{\Pi \in S_n} P_\Pi}{\binom{d+n-1}{n}} \tag{13.2}$$

$$= \frac{\sum_{\Pi \in S_n} P_\Pi}{\prod_{i=0}^{n-1} d+i} \tag{13.3}$$

We used the definition of Π_{sym} from last lecture, whose trace can be calculated via combinatorics (the stars and bars method).

$$\Pi_{sym} = \sum_{t \in \text{types}} |T_t\rangle \langle T_t| \tag{13.4}$$

$$|T_t\rangle = \frac{1}{\sqrt{\binom{n}{t}}} \sum_{x^n \in T_t} |x^n\rangle \tag{13.5}$$

$$T_t = \{x^n : \text{type}(x^n) = t\} \tag{13.6}$$

We can estimate $E(|u\rangle \langle u|^{\otimes n})$ by sampling over the Gaussian state $|g\rangle \in N_{\mathbb{C}(0,\frac{1}{d})^d}$ instead.

$$E_{|g\rangle \in N_{\mathbb{C}(0,\frac{1}{d})^d}} |g\rangle \langle g|^{\otimes n} = \frac{\Pi_{sym}^{(n,d)}}{d^n} \tag{13.7}$$

Notice that this value differs from the other moment by a normalization. Also notice that higher n are more sensitive to fluctuations. Also, it's useful to know that the Gaussian is the only moment that is rotationally (unitarily) invariant.

Remark 13.1.1. *We can conclude that*

$$\text{i.i.d. + rotational invariance} \Leftrightarrow \text{Gaussian}$$

We now provide a concrete proof for this statement. It is straightforward to verify that an i.i.d. Gaussian distribution is rotationally invariant, so we will focus on proving the converse. Given an i.i.d. distribution, it can be written as:

$$p(x_1, \dots, x_n) = f(x_1) \dots f(x_n) \tag{13.8}$$

Rotational invariance imposes the condition that the joint probability only depends on the radial distance, meaning:

$$p(x_1, \dots, x_n) = g\left(\sqrt{x_1^2 + \dots + x_n^2}\right) \tag{13.9}$$

To connect g with f , we choose $x_2 = \dots = x_n = 0$, which allow us to express g in terms of f:

$$g(x) = f(x) f^{n-1}(0) \tag{13.10}$$

Furthermore, we can get the functional equation for f:

$$\sum_i \ln \frac{f(x_i)}{f(0)} = \ln \frac{f(\sqrt{x_1^2 + \dots + x_n^2})}{f(0)} \tag{13.11}$$

The only solution to this functional equation is the Gaussian distribution, which concludes the proof.

We can use the moment to calculate the expected value of an exponent of the overlap of a random state with an arbitrary given vector.

$$E_{|u\rangle} |\langle u|0\rangle|^{2k} = E_{|u\rangle} |\langle 0|u\rangle \langle u|0\rangle|^{k} \tag{13.12}$$

$$= \operatorname{tr}\Big(|0\rangle\langle 0|^{\otimes k} \otimes E_{|u\rangle}(|u\rangle\langle u|^{\otimes k})\Big) \tag{13.13}$$

$$= E_{|u\rangle} \operatorname{tr}\Big(|u\rangle\langle u|^{\otimes k}\Big) \tag{13.14}$$

$$= \frac{1}{\binom{d+k-1}{k}} \tag{13.15}$$

This value approximates to $\frac{k!}{d^k}$ if $d >> k$. We can then calculate the mean and standard deviation of $|\langle u|0\rangle|^2$.

$$E_{|u\rangle} |\langle u|0\rangle|^2 = \frac{1}{d} \tag{13.16}$$

$$E_{|u\rangle} |\langle u|0\rangle|^4 = \frac{2}{d(d+1)} \tag{13.17}$$

$$\approx \frac{2}{d^2} \tag{13.18}$$

$$\sigma(|\langle u|0\rangle|^2) = \sqrt{E_{|u\rangle} |\langle u|0\rangle|^4 - (E_{|u\rangle} |\langle u|0\rangle|^2)^2} \tag{13.19}$$

$$\approx \frac{1}{d} \tag{13.20}$$

Notice that the mean and standard deviation are comparable. There's also a tail bound $\Pr\big(|\langle u|0\rangle|^2 \geq \frac{r}{d}\big) \approx e^{-r}$ but we aren't proving this this lecture.

13.2 Representation

Denote R as a representation that maps elements in group G to unitaries that operate on the vector space V. V^G as the set of all states that are invariant under all $R(g)$.

$$R : G \to U(V) \tag{13.21}$$

$$V^G = \{\psi \in V : R(g) |\psi\rangle = |\psi\rangle , \forall g \in G\} \tag{13.22}$$

For example, suppose $G = U(d)$ and $R(g) = g$. Then $V^G = \{0\}$. If instead $R(g) = g \otimes g^*$, then $V^G = \mathbb{C} |\phi\rangle = \mathbb{C}\text{vec}(I)$.

Theorem 13.2.1. *Define* Π *as the average of all* $R(g)$*. We claim that this is the projector onto* V^G*.*

$$\Pi = \frac{1}{|G|} \sum_{g \in G} R(g) = \text{proj } V^G$$

Proof. We first prove that Π is a projector. Let h be an arbitrary element of G.

$$R(h)\Pi = \frac{1}{|G|} \sum_{g \in G} R(h)R(g) \tag{13.23}$$

$$= \frac{1}{|G|} \sum_{g \in G} R(hg) \tag{13.24}$$

$$= \frac{1}{|G|} \sum_{g' \in G} R(g') \tag{13.25}$$

$$= \Pi \tag{13.26}$$

$$\Pi^\dagger \Pi = \frac{1}{|G|} \sum_{h \in G} R(h)^\dagger \Pi \tag{13.27}$$

$$\Pi^\dagger \Pi = \frac{1}{|G|} \sum_{h \in G} R(h^{-1}) \Pi \tag{13.28}$$

$$\Pi^\dagger \Pi = \frac{1}{|G|} |G| \Pi \tag{13.29}$$

$$\Pi^\dagger \Pi = \Pi \tag{13.30}$$

$$\tag{13.31}$$

Therefore, Π is a projector. We also show that it projects onto V^G.

By definition, each $|\psi\rangle \in V^G$ is invariant under Π, since Π is a linear combination of $R(g)$. We also see that $R(g)\Pi |\psi\rangle = \Pi |\psi\rangle$, so $\Pi |\psi\rangle \in V^G$. This proves that Π projects onto V^G. □

13.3 Anticoncentration

Ideally, we would like to sample random norm-1 states. A first strategy that might come to mind is simply to apply random gates to a initial state $|0\rangle$. WLOG assume that these are two-qubit gates, since any multi-qubit gate can be made from two-qubit gates.

$$|\psi\rangle = U_T ... U_1 |0\rangle \tag{13.32}$$

It's clear that this fails to achieve uniform randomness with a polynomial number of gates, as $|\psi\rangle$ has an exponential number of degrees of freedom. However, it is possible to sample from distributions that aren't exactly the uniform distribution but whose moments are close to that of the uniform. We'd like a metric of the anticoncentration (or how un-uniform it is) of such an approximate distribution, so that we can minimize it. Entropy $H(p)$ would work but it's harder to calculate with log. Instead, we define a new one. Let $p(z) = |\langle z|\psi\rangle|^2$ be the probability of sampling the value z from the a state $|\psi\rangle$.

$$E_{|\psi\rangle} \sum_z p(x)^2 = E_{|\psi\rangle} \sum_z |\langle z|\psi\rangle|^4 = \frac{2}{2^n + 1} \tag{13.33}$$

We see that this metric is 1 if the distribution is deterministic. Otherwise, it is $\sum_z \frac{2}{2^n(2^n+1)} = \frac{2}{2^n+1}$, which is exponentially decreasing. We can then define a different metric that directly calculates the sum-of-squared differences of a distribution's probabilities from the uniform.

$$\sum_z (p(z) - \frac{1}{2^n})^2 = \sum_z p(z)^2 - \frac{1}{2^n} \tag{13.34}$$

We can also use (13.33) to bound the average Shannon entropy of the measurement outcomes.

$$\begin{aligned}
\mathbb{E}_{|\psi\rangle} H(p) &\geq \mathbb{E}_{|\psi\rangle} H_2(p) & (13.35)\\
&= \mathbb{E}_{|\psi\rangle} \left[-\ln \sum_z p(z)^2 \right] & (13.36)\\
&\geq -\ln \mathbb{E}_{|\psi\rangle} \left[\sum_z p(z)^2 \right] & (13.37)\\
&\approx n - 1 & (13.38)
\end{aligned}$$

13.4 Renyi Entropy

Besides Shannon entropy, we can define Renyi entropy to quantify our ignorance about a classical distribution. The classical Renyi entropy is defined as:

$$H_\alpha(X) = \frac{1}{1-\alpha} \log \left(\sum_x p(x)^\alpha \right) \tag{13.39}$$

where $\alpha \geq 0$. The classical Rényi entropy possesses several properties that are straightforward to verify:

1. $H_\alpha(\text{uniform distribution}) = \log d$
2. $0 \leq H_\alpha \leq \log d$
3. $\frac{d}{d\alpha} H_\alpha \leq 0$

Similarly, the quantum Renyi entropy can be defined as:

$$S_\alpha(\rho) = \frac{1}{1-\alpha} \log \left(\operatorname{tr} \rho^\alpha \right) \tag{13.40}$$

We are particularly interested in specific values of α:

1. $S_\infty(\rho) = -\log \lambda_{max} = -\log ||\rho||_\infty$
2. $S_2(\rho) = -\log\left(\operatorname{tr}\rho^2\right)$
3. $S_0(\rho) = \log(\operatorname{rank}\rho)$

We can see that for $\alpha < 1$, the Renyi entropy is more sensitive to small eigenvalues of ρ; whereas for $\alpha > 1$, it becomes more sensitive to large eigenvalues. Here's a concrete example to illustrate this sensitivity. Consider the following density matrix on 10^6 qubits:

$$\rho = \frac{1}{2}\left(\frac{I}{2}\right)^{\otimes 10^2} \otimes \left(|0\rangle\langle 0|\right)^{\otimes 10^6 - 10^2} + \frac{1}{2}\left(\frac{I}{2}\right)^{\otimes 10^6} \tag{13.41}$$

This density matrix describes a system that, with probability 1/2 , is maximally mixed over all 10^6 qubits, and with probability 1/2 , is only maximally mixed on the first 10^2 qubits while the remaining $10^6 - 10^2$ qubits are in the pure state $|0\rangle$. Approximately, this state has 2^{10^6} eigenvalues equal to 2^{-10^6} and 2^{10^2} eigenvalues equal to 2^{-10^2} . Using this information, we can calculate the Rényi entropy for different values of α :α:

$$S_\alpha(\rho) \approx \begin{cases} 10^6 & \alpha < 1 \\ 1 + \frac{10^6+10^2}{2} & \alpha \to 1 \\ 10^2 & \alpha > 1 \end{cases} \tag{13.42}$$

The quantum Renyi entropy can be related to the von Neumann entropy by taking the limit of α:

$$\lim_{\alpha \to 1} S_\alpha(\rho) = S(\rho) \tag{13.43}$$

Unlike the von Neumann entropy, which requires computing the logarithm of the density matrix, the Renyi entropy instead only involves calculating the moments. This makes the Renyi entropy more tractable when dealing with random ensembles.

With these tools in hand, we are now ready to prove the statement that *most pure states are highly entangled.* Consider a quantum state $|\psi\rangle \in \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$. Without loss of generality, we assume that $d_A \le d_B$. Our goal is to establish a lower bound on the expected von Neumann entropy of the reduced state by relating it to the second Renyi entropy, and then to directly calculate the Renyi entropy.

$$\mathbb{E}S(A) \ge \mathbb{E}S_2(A) = \mathbb{E}\operatorname{tr}\left(-\log \psi_A^2\right) \tag{13.44}$$
$$\ge -\log \mathbb{E}\operatorname{tr}\psi_A^2 \tag{13.45}$$

the second line comes from the convexity of the function $-\log$.

To further calculate $\mathbb{E}\operatorname{tr}\psi_A^2$, we need the following Lemma:

Lemma 13.4.1 (SWAP trick)**.** *Let F_{AB} be the swap operator on system A and B, then*

$$\operatorname{tr} F_{AB}(X_A \otimes Y_B) = \operatorname{tr} X_A Y_B \tag{13.46}$$

Proof. The two operators X_A and Y_B can be written as:

$$X_A = \sum_{ij} X_{ij} |i\rangle\langle j| \tag{13.47}$$

$$Y_B = \sum_{kl} Y_{kl} |k\rangle\langle l| \tag{13.48}$$

Then we have

$$\operatorname{tr} F_{AB}(X_A \otimes Y_B) = \sum_{ijkl} \operatorname{tr} F_{AB} X_{ij} Y_{kl} \left|ik\right\rangle \left\langle jl\right| \tag{13.49}$$

$$= \sum_{ijkl} \operatorname{tr} X_{ij} Y_{kl} \left|ki\right\rangle \left\langle jl\right| \tag{13.50}$$

$$= \sum_{ik} \left\langle ki\right| X_{ik} Y_{ki} \left|ki\right\rangle \tag{13.51}$$

$$= \operatorname{tr} X_A Y_B \tag{13.52}$$

□

Then we can directly calculate that

$$\mathbb{E} \operatorname{tr}_A \psi_A^2 = \mathbb{E} \operatorname{tr}_{A_1 A_2} \left[F_{A_1 A_2} \left(\psi_{A_1} \otimes \psi_{A_2}\right)\right] \tag{13.53}$$

$$= \operatorname{tr}_{A_1 A_2 B_1 B_2} F_{A_1 A_2} \mathbb{E} \left(\psi_{A_1 B_1} \otimes \psi_{A_2 B_2}\right) \tag{13.54}$$

$$= \operatorname{tr}_{A_1 A_2 B_1 B_2} F_{A_1 A_2} \frac{I + F_{A_1 A_2} F_{B_1 B_2}}{d_A d_B (d_A d_B + 1)} \tag{13.55}$$

$$= \operatorname{tr}_{A_1 A_2} \frac{F_{A_1 A_2} d_B^2}{d_A d_B (d_A d_B + 1)} + \operatorname{tr}_{B_1 B_2} \frac{F_{B_1 B_2} d_A^2}{d_A d_B (d_A d_B + 1)} \tag{13.56}$$

$$= \frac{d_A + d_B}{d_A d_B + 1} = \frac{1}{d_A} + \frac{1}{d_B} \tag{13.57}$$

The first line applies the SWAP trick. The second line extends the reduced density matrix ψ_A to $\operatorname{tr}_B \psi_{AB}$. The third line computes the second moment of the random state. The fourth line comes from the fact that $F_{AB} F_{AB} = I$ and $\operatorname{tr}_{B_1 B_2} I_{B_1 B_2} = d_B^2$. The last line comes from that $\operatorname{tr}_{A_1 A_2} F_{A_1 A_2} = d_A$. Now we can conclude that:

$$\mathbb{E} S(A) \geq -\log \left(\frac{1}{d_A} + \frac{1}{d_B}\right) \tag{13.58}$$

In the limit $d_B \to \infty$, system B behaves like a huge heat bath, and $\mathbb{E} S(A) \sim \log d_A = \operatorname{tr} I_A / d_A$. This shows that, in the typical case, region A is maximally entangled with system B.