8.372 Quantum Information Science III **Fall 2024**

Lecture 14: 10/24/2024

Scribe: Gary Yang *k-designs, nets and concentration*

14.1 Entanglement of Random States

14.1.1 Continuation from Last Lecture

A main theme of this section is to continue illustrating that random states are likely to be highly entangled. Recall from last lecture, if have a random state

$$|\psi\rangle \in C^{d_A} \otimes C^{d_B} \tag{14.1}$$

then to quantify is entanglement, we have the expectation value of

$$\mathbb{E}\,\mathrm{tr}_A\,\psi_A^2 = \frac{d_A + d_B}{d_A d_B + 1} \approx \frac{1}{d_A} + \frac{1}{d_B} \tag{14.2}$$

To illustrate a consequence of this, let us imagine we have a random state on system A and B, with total of n qubits, where we allocate n_A qubits to the first and n_B qubits to the second. This means $d_A = 2^{n_A}$, $d_B = 2^{n_B}$, with $n_A + n_B = n$. If we plot the entropy of subsystem A as a function of qubits in that system n_A between 0 and n, we will find what is called the "Page Curve".

We can write down a straightforward upperbound, which is

$$S(A) \leq \min(n_A, n_B), \tag{14.3}$$

because the entropy cannot exceed $\log_2(d_A) = n_A$. Also we stipulate that the entropy of the two system to be the same, entropy also cannot exceed the entropy of system B. The actual plot is shown below, where the upperbound is mostly saturated except by a small amount in the middle, which is off by a constant amount.

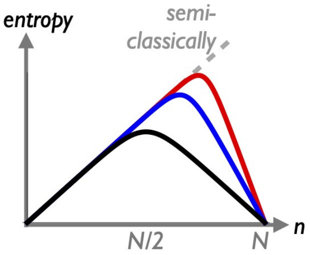

Figure 14.1: Page Curve (black), taken from Markus P. Müller's website

If we used

$$\mathbb{E}\,\mathrm{tr}_A\,\psi_A^2 \approx \frac{1}{d_A} + \frac{1}{d_B} \tag{14.4}$$

to calculate the entropy, we would have gotten $S(A) = n_A - 1$.

14.1.2 Another Way to Gauge Entanglement

After studying entanglement using entropic measurements, let us explore how to use our old friend distance measure to measure entanglement. Say,

$$\|\psi_A - \frac{I}{d}\|_1 \tag{14.5}$$

We found that it will be easier to convert the 1−norm in Trace Distance to a 2−norm squared, because we want our tools of moments to show up cleanly in this expression. In problem set, we have shown that, using Cauchy-Schwartz,

$$\|\psi_A - \frac{I}{d_A}\|_1 \leq \sqrt{d_A}\|\psi_A - \frac{I}{d}\|_2 \tag{14.6}$$

We then calculate

$$\|\psi_A - \frac{I}{d}\|_2^2 = \operatorname{tr}\psi_A^2 + \operatorname{tr}\left(\psi_A \frac{I}{d_A}\right) - \operatorname{tr}\left(\psi_A \frac{I}{d_A}\right) - \operatorname{tr}\left(\frac{I}{d_A}\right)^2 \tag{14.7}$$

$$= \operatorname{tr}\psi_A^2 - \frac{1}{d_A} \tag{14.8}$$

$$\implies \mathbb{E}\|\psi_A - \frac{I}{d}\|_2^2 \approx \frac{1}{d_B} \tag{14.9}$$

Using this, we get

$$\mathbb{E}\|\psi_A - \frac{I}{d}\|_1^2 \leq d_A \mathbb{E}\|\psi_A - \frac{I}{d}\|_2^2 \approx \frac{d_A}{d_B} = 2^{n_A - n_B} \tag{14.10}$$

$$\tag{14.11}$$

which is more useful if n_A and n_B are different.

$$\mathbb{E}\|\psi_A - \frac{I}{d}\|_1 \leq \sqrt{\mathbb{E}\|\psi_A - \frac{I}{d}\|_1^2} \leq \sqrt{\frac{d_A}{d_B}} \tag{14.12}$$

14.2 Pseudo-Random States and k-design

Now that we have seen random states are highly entangled, how would we as mortal human get close its power? We will use Pseudo-Random States, which hopefully are informationally indistinguishable from true random states, at least up to a certain moments. Note that there are computationally pseudo-random states, which we won't get into.

To define this better, we say

Definition 14.2.1. *If μ is a measure on $\mathbb{C}^d$, we say μ is a state $k-$design if*

$$\mathbb{E}_{|\psi\rangle\in\mu}|\psi\rangle\langle\psi|^{\otimes k} = \mathbb{E}_{|\psi\rangle\in\mathbb{C}^d}|\psi\rangle\langle\psi|^{\otimes k}$$

This ensures that the first $k-$moments match the true random states. Later we will talk $k-$design for random circuits and random unitaries.

14.2.1 $k = 1$

In this case, $1-$ design needs

$$\mathbb{E}_{|\psi\rangle\in\mu}|\psi\rangle\langle\psi|^{\otimes k} = \frac{I}{d}$$

We can just draw states from any orthonormal basis, to achieve $1-$desgin. Note we only need a discrete set of things to achieve this. Note this means that, if we have an observable M,

$$\mathbb{E}\ \langle|\psi|M|\psi\rangle \tag{14.13}$$

will be the same as if we drawn ψ from true random states, but the variance might be different.

However, if we know measure in the orthonormal basis we draw states from, with many samples we may be able to distinguish $1-$designs we ture random states.

What is worse, these basis does not reproduce the entanglement properties of random states, which we derived from second-moment calculations.

14.2.2 $k = 2$

The condition now is

$$\mathbb{E}_{|\psi\rangle\in\mu}|\psi\rangle\langle\psi|^{\otimes 2} = \frac{I+F}{d(d+1)} \tag{14.14}$$

An example of this $2-$design is Random Stabilizer States:

$$\psi_S = |\psi_S\rangle\langle\psi_S| = 2^{-n}\sum_{\sigma_p\in S}\sigma_p \tag{14.15}$$

Some bookkepping is in place:

1. S is an abelian subgroup of Pauli's on $n-$qubits with $|S| = 2^n$. S does not include $-I$.
2. Using these states, can specify a state with n^2 bits (instead of 2^n for uniformly random states).
3. These states are also highly entangled, which is what we want.
4. They can be efficiently simulated classically using the Gottesman-Knill theorem.
5. I also found a GitHub Repo that claim it does sampling of random stabilizer states efficiently (https://github.com/qotlabs/randstab).

In this case, μ is a probability distribution on $|\psi_S\rangle$, where we choose S uniformly to specify a state from all the random stabilizer states, liek this:

$$\mathbb{E}_S\psi_S\otimes\psi_S = \frac{I+F}{2^n(2^n+1)} \tag{14.16}$$

To prove this statement, let us recall that

$$\operatorname{tr} F(X\otimes Y) = \operatorname{tr} XY \tag{14.17}$$

$$\tag{14.18}$$

and we can plug in pauli's into the above expression as X, Y, and we can learn the Pauli's expansion of the swap operator F:

$$\operatorname{tr} F(\sigma_p \otimes \sigma_q) = \operatorname{tr} \sigma_p \sigma_q = 2^n \delta_{pq} \tag{14.19}$$

$$\tag{14.20}$$

This way, we get:

$$F = 2^{-n} \sum_{p \in \{0,1,2,3\}^n} \sigma_p \otimes \sigma_p \tag{14.21}$$

Now let us calculate,

$$\mathbb{E}_S 4^{-n} \sum_{\sigma_p \in S} \sigma_p \otimes \sum_{\sigma_q \in S} \sigma_p = 4^{-n} \sum_p \Pr(\sigma_p \in S) \sigma_p \otimes \sigma_p \tag{14.22}$$

Because S is uniformly distributed, we can say that if $p = 0^n$, $\Pr(\sigma_p \in S) = 1$ but if $p \neq 0^n$, $\Pr(\sigma_p \in S) = \frac{1}{2^n+1}$, which is the same for any other Pauli, after counting how many non-identity Pauli's. We draw $2^n - 1$ non identity Paulis. The $p = 0^n$ part shall give us the identity part, and the rest shall give us the swap operator part of the expression.

14.2.3 Other 2-designs

Fermionic Gaussian States (free fermion states) are also 2-designs. It is a continuous distribution that have symmetry of orthogonal group. We will discuss them later.

14.3 More calculation on Entropy

Next, let us what other ways of characterizing entanglement in random states. For example, let consider what is the probability that a random state is low in entanglement. This is

$$\Pr\Big[\operatorname{tr} \psi_A^2 \geq 2^{l-n_A}\Big] \cong n_A - l \tag{14.23}$$

We can use the Markov's inequality to do this,

$$\Pr\Big[\operatorname{tr} \psi_A^2 \geq 2^{l-n_A}\Big] \leq \frac{\mathbb{E} \operatorname{tr} \psi_A^2}{2^{l-n_A}} \approx 2^{-l} \tag{14.24}$$

Assuming ψ_A is pretty random. We can see the larger the deficit of entanglement (l) the smaller the probability. If we want to calculate the variance, which appear in Chebyshev's inequality, we would have need $\mathbb{E}(tr\psi_A^2)^2 = \mathbb{E}(tr\psi_A^{\otimes 4}(F_{12} \otimes F_{34})$, which requires 4-th moment, so a 2-design wouldn't have gave us the accurate information.

14.4 Characterizing Random States without Moments

Moving on to do some calculation on uniform state. The largest eigenvalue of ψ_A is

$$\|\psi_A\|_\infty = \max_{|\alpha\rangle \in \mathbb{C}^d, |\beta\rangle \in \mathbb{C}^d} [\langle \psi | (|\alpha\rangle_A \otimes |\beta\rangle_B)] \tag{14.25}$$

Note if ψ_A is uniform this expression will be $\frac{1}{d}$. Let us see how close a random states can get to this value.

If we fix $|\alpha\rangle, |\beta\rangle$ and calculate $\langle\psi|\alpha,\beta\rangle$, we argue that there is a small probability that $\|\psi_A\|_\infty$ will be large. We can use result from last lecture with the replica trick that

$$\mathbb{E}|\langle\psi|\alpha,\beta\rangle|^{2k} = \frac{k!}{d\cdots(d+k-1)} \leq \frac{k!}{d^k}, \tag{14.26}$$

where $|\psi_A\rangle$ is random.

A trick to bound this is relate $|\psi\rangle$ to random Gaussian states $|g\rangle$

$$\mathbb{E}_g\left[|\langle g \mid \alpha,\beta\rangle|^{2k}\right] = \mathbb{E}_r r^{2k} \cdot \mathbb{E}_\psi|\langle\psi \mid \alpha,\beta\rangle|^{2k} \leq \mathbb{E}|\langle\psi|\alpha,\beta\rangle|^{2k}$$

where $|g\rangle = r|\psi\rangle$ and we normalize $\mathbb{E}r^2 = 1$. This means:

$$\Pr\left[|\langle\psi \mid \alpha,\beta\rangle|^2 \geq \gamma\right] \leq \Pr\left[|\langle g \mid \alpha,\beta\rangle|^2 \geq \gamma\right] = e^{-\frac{d^2}{\gamma}}, \tag{14.27}$$

because every moment on the LHS is smaller than that of Gaussian. To illustrate the equality to be exponential, we write explicitly

$$|g\rangle = \begin{pmatrix} x_1 + iy_1 \\ x_2 + iy_2 \\ \vdots \\ x_d + iy_d \end{pmatrix} \tag{14.28}$$

and do a Gaussian integral

$$\Pr\left[x_i^2 + y_i^2 \geq \gamma^2\right] = \int_{r\geq\gamma}\int d\theta e^{-\frac{r^2}{2\sigma^2}} \tag{14.29}$$

Applying this, we see

$$\Pr\left[|\langle\psi|\alpha,\beta\rangle|^2 \geq \frac{a}{d}\right] \leq e^{-da} \tag{14.30}$$

which says the fluctuation of the inner product is very small. Next, let us consider the worst case scenario using the Union Bound.

$$\Pr\left[\max_{\alpha,\beta}|\langle\psi|\alpha,\beta\rangle|^2 \geq \frac{a}{d}\right] \leq \sum_{\alpha,\beta}\Pr\left[|\langle\psi|\alpha,\beta\rangle|^2 \geq \frac{a}{d}\right]$$
$$= (\text{number of } \{\alpha,\beta\})\, e^{-da}$$

But exactly how many α,β are out there? To pinpoint them exactly, we need infinite precision. However, we can use the concept of δ-net to specify them to good precision (not exactly). We define

$$\mathcal{N} = \delta\text{-net} \subset \mathbb{C}^d$$

Now we just maximize all α,β that lives in the $\delta-$net, with size $|\mathcal{N}|$.

$$\Pr\left[\max_{\alpha,\beta\in\mathcal{N}}|\langle\psi|\alpha,\beta\rangle|^2\geq\frac{a}{d}\right]\leq\sum_{\alpha,\beta\in\mathcal{N}}\Pr\left[|\langle\psi|\alpha,\beta\rangle|^2\geq\frac{a}{d}\right]=|\mathcal{N}|^2e^{-da}$$

Let us see how big $\mathcal{N}$ is then. Ideally we don't want them to be too big. We claim that

$$\exists\mathcal{N}\text{ such that }|\mathcal{N}|\leq\left(1+\frac{2}{\delta}\right)^{2d}\text{, and }\mathcal{N}\text{ is a }\delta\text{-covering of unit sphere in }\mathbb{C}^d.$$

To prove this, we can construct the net this way. If $\mathbb{N}$ is not a $\delta-$net, at step i, you can add $|\beta_i\rangle$, who are at least δ far away from the set, to the net. Suppose we have a measure where volume of a ball of radius 1 is 1, so a ball of radius $\delta/2$ would have a volume of $(\delta/2)^{2d}$. As long as the following is true, we cannot add more points to the net. So we can consider an $|\mathcal{N}|$ that barely passes this inequality, which proves the statement on the size of $\mathcal{N}$

$$|\mathcal{N}|\cdot\left(\frac{\delta}{2}\right)^{2d}<\left(1+\frac{\delta}{2}\right)^{2d}\tag{14.31}$$

Final Stretch now! Let us define

$$\max_{\alpha,\beta\in\mathbb{C}^d}|\langle\psi\mid\alpha,\beta\rangle|^2=A$$
$$\max_{\alpha,\beta\in\mathbb{N}}|\langle\psi\mid\alpha,\beta\rangle|^2=B$$

We have proven that B is not that big from the $\delta-$net argument, say

$$B\leq\frac{10}{d}$$

By triangle inequality, we have a straightforward bound:

$$A\leq B+\delta$$

However, a better bound is

$$A\leq B+2\delta A$$

which comes from analyzing the error by restricting to the $\delta-$net more carefully. We have

$$\text{error }=(|x\rangle-|x_N\rangle)\otimes|\beta\rangle+(|\alpha\rangle\otimes|\beta\rangle-|\beta_N\rangle)=\mathcal{O}(\delta)$$

We then take the inner produc between this error term with Ψ, which from the definition of A is at most to get δA.

So we have

$$A\leq\frac{B}{1-\delta}$$

As a result, the probability of not enough entanglement goes down exponentially in d, who is exponential in the number of bits:

$$\Pr[\text{low entanglement}]\leq e^{-2^n}$$

This bound is stronger than that we calculated from second moments, which shows how strong the entanglment of uniformly random states is.