8.S372/18.S996 Quantum Information Science III **Fall 2020**

Lecture 16: October 27, 2020

Lecturer: Aram Harrow *Scribe: Joshua Lin, Andrey Boris Khesin*

In this lecture, we will be covering random states and unitaries. One application of random unitaries is that we can use them to destroy information in a certain sense (see the last lecture). A certain theme that we will see repeat is 'concentration of measure', as we take the limit of high dimensions (e.g. if we have large quantum systems), certain natural measures will concentrate along physically interesting subspaces.

16.0.4 Scalar Random Variables

Lemma 8 (Markov's Inequality)

$$X \geq 0 \implies Pr[X \geq a\mathbb{E}X] \leq \frac{1}{a}$$

To get a tighter bound, use higher moments:

Lemma 9 (Chebyshev's Inequality)

$$\mathbb{E}X = \mu,\ \mathbb{E}(X-\mu)^2 = \sigma^2 \implies Pr[(X-\mu)^2 \geq a\sigma] \leq \frac{1}{a^2}$$

This is an example of the "Bernstein trick":

$$f(x) > 0,\ f'(x) \geq 0 \implies \Pr[X \geq a] = \Pr[f(x) \geq f(a)] \leq \frac{E[f(x)]}{f(a)}$$

in other words, to derive Chebyshev's Inequality, we apply Markov's Inequality to a transformed version of the random variable, where we take $f(x)$ to be squared difference from expectation (caveat, f is not monotone here).

Note that Chebyshev's inequality is really weak for things like Gaussian random variables, it tells us that the probably of having a gaussian r.v. 5σ above average is only bounded by $1/25$ (we know that it's very small in reality). Instead, use $f(x) = e^{\lambda x}$ in Bernstein trick, and we get the much better:

$$X \sim \mathcal{N}(\mu = 0, \sigma^2 = 1),\ \Pr[X \geq a] = e^{\lambda^2/2 - \lambda a}$$

This bound is funny because λ is a parameter we are free to choose. Clearly, some values of λ give better bounds than others, optimal bound at

$$\lambda = a \implies \Pr[X \geq a] \leq e^{-a^2/2}$$

Note that this is a much better bound than just using Chebyshev's Inequality. Actually this bound is pretty much optimal; it turns out it's only off by roughly a constant factor in this particular example - and the only thing we really needed to know is $\mathbb{E}e^{\lambda X}$ called the 'moment generating function', :

$$\mathbb{E}e^{\lambda X} = 1 + \lambda \mathbb{E}X + \frac{\lambda^2}{2}\mathbb{E}X^2 + \dots$$

Lemma 10 (Chernoff bound) *$X_1, \dots, X_n$ i.i.d $Pr[X_i = 1] = Pr[X_i = -1] = \frac{1}{2}$, and let $X = \sum_i X_i$.*

$$\mathbb{E}[e^{\lambda X}] = \mathbb{E}[e^{\lambda X_1}]^n = (\cosh(\lambda))^n \leq e^{n\lambda^2/2}$$
$$Pr[X \geq \delta n] \leq e^{n\lambda^2/2 - n\delta\lambda} = e^{n\delta^2/2} \quad \textit{if} \quad \lambda = \delta$$

We get similar bounds if $|X_i| \leq 1$ and $\mathbb{E}X_i = 0$. Intuitively the same bound still applies, because we will only get less deviation if we allow the random variables to be between -1 and 1.

16.0.5 Random Vectors

Gaussian random vectors look like:

$$|g\rangle = \begin{pmatrix} g_1 \\ \vdots \\ g_d \end{pmatrix}, \quad g_i \in \mathcal{N}_{\mathbb{C}}\left(0, \frac{1}{d}\right), g_i = x_i + iy_i, \quad x_i, y_i \in \mathcal{N}\left(0, \frac{1}{2d}\right)$$

which gives us $\mathbb{E}\langle g|g\rangle = 1$, and $p(|g\rangle) = \left(\frac{d}{\pi}\right)^d e^{-d\langle g|g\rangle}$. We can use this to generate a random unit vector:

$$|v\rangle = \frac{|g\rangle}{\sqrt{\langle g|g\rangle}}.$$

Gaussian vectors are great because the entries are independent, and the distribution is rotationally symmetric (i.e. it is symmetric under actions of $U(d)$). It turns out that the Gaussian distribution is the only distribution with these properties.

Note that we have $\mathbb{E}[g_i] = 0$ since the distribution of g_i is invariant under phase rotations due to the unitary invariance of the gaussian vector. The only way to get nonzero expectations is to write down 'scalar quantities' under the group invariance:

$$\mathbb{E}[g_i g_j^*] = \frac{\delta_{ij}}{d}, \quad \mathbb{E}[|g\rangle\langle g|] = \sum_{ij} \mathbb{E}[g_i g_j^*]|i\rangle\langle j| = \frac{I}{d}$$

$$\mathbb{E}[g_i g_j g_k^* g_l^*] = \mathbb{E}[g_i g_k^*]\mathbb{E}[g_j g_l^*] + \mathbb{E}[g_i g_l^*]\mathbb{E}[g_j g_k^*] = \frac{\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}}{d^2}$$

by Isserlis' (/Wick's) theorem - since we know the only way to get nonzero answers out is to pair up the g factors.

$$\mathbb{E}[|g,g\rangle\langle g,g|] = \frac{1}{d^2}\sum_{ij} |i,j\rangle\langle i,j| + |i,j\rangle\langle j,i| = \frac{I + \text{SWAP}}{d^2}$$

where the SWAP operator swaps the two registers. In general, using Wick's theorem:

$$\mathbb{E}[g_{i_1} \dots g_{i_n} g_{j_1}^* \dots g_{j_n}^*] = \frac{1}{d^n}\sum_{\pi \in S_n}\prod_{l=1}^{n} \delta_{i_l, j_{\pi(l)}}$$

$$\mathbb{E}[|g\rangle\langle g|^{\otimes n}] = \frac{1}{d^n}\sum_{\pi \in S_n} P_\pi, \quad P_\pi = \sum_{i_1,\dots,i_n} |i_1,\dots,i_n\rangle\langle i_{\pi(1)},\dots,i_{\pi(n)}|$$

where S_n is the symmetric group, π is a permutation, and we are essentially just summing over ways of matching up the g factors. Now, how do we interpret this quantity?

$$\Pi_{\text{sym}} := \frac{1}{n!}\sum_{\pi \in S_n} P_\pi = \text{ projector onto symmetric subspace}$$

$$\text{Sym}^n\mathbb{C}^d = \{|\psi\rangle \in (\mathbb{C}^d)^{\otimes n} : P_\pi|\psi\rangle = |\psi\rangle \quad \forall \pi \in S_n\}$$

Note that we can prove the above fact in much bigger generality, suppose I have an arbitrary finite group G with unitary rep $r : G \to U(V)$, let V^G be the G-invariant vectors in V, then the claim is that:

$$\Pi = \frac{1}{|G|}\sum_{g \in G} r(g) \quad \text{projects onto} \quad V^G$$

First, note that Π itself is invariant under G-action:

$$r(h)\Pi = r(h)\frac{1}{|G|}\sum_g r(g) = |G|^{-1}\sum_g r(hg) = |G|^{-1}\sum_g r(g) = \Pi$$

note that group acts freely on itself in the sense that its action is 1-to1. So, that means $r(h)\Pi|\psi\rangle = \Pi|\psi\rangle$, so $\text{Im}(\Pi) \subset V^G$, convsersely $|\psi\rangle \in V^G \implies \Pi|\psi\rangle \in \text{Im}(\Pi)$, finally, to prove that it's a projector:

$$\Pi^\dagger\Pi = \mathbb{E}_h r(h^{-1})\Pi = \Pi$$

so, going back to what we had before,

$$\mathbb{E}[|g\rangle\langle g|^{\otimes n}] = \frac{1}{d^n}\sum_{\pi \in S_n} P_\pi = \frac{n!}{d^n}\Pi_{\text{sym}}$$

Now, what about random unit vectors? Note that $|g\rangle = r|u\rangle$ with $r, |u\rangle$ independent, so that:

$$\mathbb{E}[|g\rangle\langle g|^{\otimes n}] = \mathbb{E}[r^{2n}]\mathbb{E}[|u\rangle\langle u|^{\otimes n}]$$

so, we know in fact that:

$$\mathbb{E}[|u\rangle\langle u|^{\otimes n}] = \frac{\Pi_{\text{sym}}}{\text{tr}\Pi_{\text{sym}}} \implies \text{Sym}^n\mathbb{C}^d = \text{span}\{|\psi\rangle^{\otimes n} : |\psi\rangle \in \mathbb{C}^d\}$$

If we have another basis $p \in P_n$ where p describes a partition, describing a 'type' of unit vector:

$$|p\rangle = \binom{n}{np}^{-1/2} \sum_{x \in T_p^n} |x\rangle$$

$$\mathbb{E}[|u\rangle\langle u|^{\otimes n}] = \frac{\Pi_{\text{sym}}}{\binom{d+n-1}{n}} = \frac{\sum_\pi P_\pi}{d(d+1)\dots(d+n-1)}$$

$$1 \le \mathbb{E}[r^{2n}] = \frac{d(d+1)\dots(d+n-1)}{d^n} \le e^{n^2/2d}$$

So, if $d \gg n$, we have a concentration of measure of the gaussian vectors around the unit vectors.

16.0.6 Applications to Entanglement of random states

Suppose that we have a uniformly random $|\psi\rangle \in \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$, then how entangled is ψ, in other words, what is $\mathbb{E}S(A)_\psi$?

$$\mathbb{E}S(A)_\psi \ge \mathbb{E}S_2(\psi_A) = -\mathbb{E}\log\ \text{tr}\ \psi_A^2$$

where $S_2(\rho) = -\log \text{tr}\ \rho^2$ is the Renyi entropy. But, by concavity of log, we have:

$$\mathbb{E}S(A)_\psi \ge -\log \mathbb{E}\ \text{tr}\ \psi_A^2$$

Now, note this cool fact:

$$\text{tr}(X \otimes Y)\text{SWAP} = \text{tr}[XY]$$

which we can draw as a circuit. This gives us:

$$\text{tr}\psi_A^2 = \text{tr}(\psi_A \otimes \psi_A)\text{SWAP} = \text{tr}[(\psi_{A_1B_1} \otimes \psi_{A_2B_2})(\text{SWAP}_{A_1A_2} \otimes I_{B_1B_2})]$$

But now, we have that:

$$\mathbb{E}\text{tr}(\psi_A^2) = \text{tr}\Big[\mathbb{E}(\psi_{A_1B_1} \otimes \psi_{A_2B_2})\,(\text{SWAP}_{A_1A_2} \otimes I_{B_1B_2})\Big]$$

$$= \operatorname{tr}\frac{\mathrm{SWAP}_{A_1A_2} + \mathrm{SWAP}_{B_1B_2}}{d_Ad_B(d_Ad_B+1)} = \frac{d_Ad_B^2 + d_A^2d_B}{d_Ad_B(d_Ad_B+1)} = \frac{d_A + d_B}{d_Ad_B+1}$$

So, putting all this together, we get that

$$\mathbb{E}S(A)_\psi \geq \log\left(\frac{d_Ad_B+1}{d_A+d_B}\right)$$

So, if the dimensions are equal, our bound is roughly $\log(d) - 1$, and if we have $d_A \ll d_B$, then our bound looks instead like $\log(d_A) - \frac{d_A}{d_B}$ (which means you have a small correction to that of the maximally entangled state).

The takeaway is that random states are close to maximally entangled, with small corrections due to the finiteness of the dimensions.