Frontiers of Quantum Complexity *Instructor: Anand Natarajan, Anurag Anshu*

Lecture 13: Monogamy of entanglement and barriers to qPCP

October 16, 2025 *TF: Yeongwoo Hwang, Tina Zhang*

The qPCP conjecture is all about distinguishing NP from QMA: if NP = QMA, then the conjecture is automatically true because of the classical PCP theorem. So we need to find Hamiltonians where all low-energy states are highly nonclassical, in the sense of having no efficient classical description. One important class of states with classical description is *product states*, so showing that a Hamiltonian family has a low-energy product states implies that it cannot be a candidate for qPCP hardness. Today we will see a result of Brandão and Harrow that shows that 2-local Hamiltonians whos interaction graph has large degree have product state approximations.

1 Monogamy of entanglement

The intuition to the Brandao-Harrow theorem is the phenomenon called monogamy of entanglement. Very basically, this says that there is no maximally entangled state on more than two qubits. For example, for $|\psi\rangle_{ABC}$ where A, B are maximally entangled, then $|\psi\rangle_{ABC} = |\psi'\rangle_{AB} \otimes |\psi''\rangle_C$. Such a statement is not true for classical correlation, where we can have a distribution with maximal correlations between any subsystems.

We can make more quantitatively versions of this phenomenon. The de Finetti theorem(s) says that a *permutation-invariant* quantum state is close to product if you trace out some of the subsystems.

Theorem 1.1. *Let $|\psi\rangle_{A_1 \ldots A_n}$ be a permutation-invariant state. Then tracing out $n-k$ subsystems gives approximately a separable state σ:*

$$\|\psi_{A_1 \ldots A_k} - \sigma_{A_1 \ldots A_k}\| \leq O(k(d+k)/(n+k)). \tag{1}$$

Intuition: permutation invariance requires that, e.g., if A_1 and A_2 are highly entangled, than so are A_1 and A_i for other i's, but this is not possible by monogamy of entanglement. So there is low entanglement to begin with.

Proof idea: conditioning on a few systems destroys correlations between the rest (very roughly).

2 The Brandão-Harrow theorem

The permutation invariance requirement is very strong, so for product state approximations of local Hamiltonians, we want a similar statement without this requirement.

Theorem 2.1. *Let ρ be a state on n particles of dimension d and $G = (V, E)$ be a D-regular graph on the vertex set $V = [n]$. Then there is a separable state σ such that*

$$\mathbb{E}_{(i,j)\in E} \|\rho_{ij} - \sigma_{ij}\|_1 \leq 12 \left(\frac{d^2 \ln d}{D}\right)^{1/3}.$$

This implies as a simple corollary that any local Hamiltonian has a product state achieving an additive approximation of the ground energy, with the same error.

Corollary 2.2. *For any 2-local Hamiltonian on the graph G, there exists a product state σ such that* $\operatorname{Tr}(\sigma) H \leq E_0(H) + 12\left(\frac{d^2 \ln d}{D}\right)^{1/3}$.

There's also fancier version of this theorem for non-regular graphs, and for graphs that "cluster" well. But we'll focus on the simple version above. Interestingly, most of the proofs are classical arguments, based on ideas from a classical paper of Raghavendra and Tan (SODA 2012). The only quantum component is the first step.

2.1 Informationally complete measurements

The idea will be to prove the BH theorem by passing from quantum states to classical distributions. This is done by using *informationally complete measurements* (IC-POVM).

For every d, there's an informationally complete measurement Λ on d-dimensional systems such that for all ξ traceless on k particles,

$$\|\xi\|_1 \leq (18d)^{k/2} \|\Lambda^{\otimes k}(\xi)\|_1.$$

In particular think of ξ as the difference between two density matrices. This is saying that Λ distorts (shrinks) differences by a bounded amount when it maps states to probability distributions.

For us, k will be 2. An example of IC-POVM for $d = 2$ (qubits) is the set of projectors on the 6 eigenvectors of Pauli X, Y, Z, normalized appropriately.

2.2 Information theory and decoupling

Now we move on to the classical part of the proof. This relies on information-theoretic tools, so we need some definitions.

To measure correlation between quantum systems we use mutual information.

$$I(X:Y)_\rho = S(\rho_{XY} \| \rho_X \otimes \rho_Y) = S(\rho_X) + S(\rho_Y) - S(\rho_{XY}).$$

This is nonnegative because relative entropy is nonnegative, and is equal to 0 only when the mixed state ρ is product.

There's also a *classically conditioned* version of this: if R is an additional system and XYR is in a CQ state:

$$\rho_{XYR} = \sum_i p_i \rho^i_{XY} \otimes |i\rangle \langle i|_R ,$$

then

$$I(X:Y|R)_\rho = \sum_i p_i I(X:Y)_{\rho_i}.$$

Mutual information satisfies a few important properties:

- Chain rule: $I(A:BR) = I(A:R) + I(A:B|R)$.
- Monotonically non-increasing under local operations: $I(X:Y)_\rho \geq I(X:H)_{(\mathcal{N}_X \otimes \mathcal{N}_Y)(\rho)}$.
- Pinsker's inequality
$$I(X:Y) \geq \frac{1}{2} \|\rho_{XY} - \rho_X \otimes \rho_Y\|_1^2.$$
- Upper bounds: $I(X:Y) \leq \min(\ln |X|, \ln |Y|)$ for classical version and $I(X:Y) \leq 2\min(\ln |X|, \ln |Y|)$ for quantum.

2.3 Decoupling

Suppose we have n classical systems $X_1, \ldots, X_n$ and we condition on a subset of up k of them as indicated below. Then we have a bound on the mutual information of the remaining systems.

$$\mathbb{E}_{0\leq k'<k} \mathbb{E}_{(a,b,\underbrace{c_1,\ldots,c_{k'}}_{C})\sim\mu} I(X_a : X_b|\underbrace{X_{c_1},\ldots,X_{c_{k'}}}_{Z}) \leq \frac{1}{k}\mathbb{E}_i I(X_i : X_{-i}),$$

where μ is the distribution of tuples of indices of whatever length without replacement, and X_{-i} is the complement of X_i. To prove this,

$$\begin{aligned}
\frac{1}{k}\mathbb{E}_i I(X_i : X_{-i}) &\geq \frac{1}{k}\mathbb{E}_{(i,j_1,\ldots,j_k)\sim\mu} I(X_i : X_{j_1}\ldots X_{j_k}) && (2)\\
&= \mathbb{E}_{(i,j_1,\ldots,j_k)\sim\mu}\frac{1}{k}\sum_{0\leq k'<k} I(X_i : X_{j_{k'+1}}|X_{j_1}\ldots X_{j_{k'}}) && (3)\\
&= \mathbb{E}_{(i,j_1,\ldots,j_k)\sim\mu}\mathbb{E}_{0\leq k'<k} I(X_i : X_{j_{k'+1}}|X_{j_1}\ldots X_{j_{k'}}) && (4)\\
&= \mathbb{E}_{0\leq k'<k}\mathbb{E}_{a,b,c_1\ldots,c_{k'}} I(X_a : X_b|X_{c_1}\ldots X_{c_{k'}}). && (5)
\end{aligned}$$

This is kind of a baby version of de Finetti theorem. The first step is by monotonicity under randomly throwing away subsystems in the complement of i. The second is by the chain rule (with $j_0 := i$). The third and fourth are just notational rewriting.

This statement is useful because we can bound $I(X_i : X_{-i})$ by the trivial bound of $\ln d$.

$$\mathbb{E}_{0\leq k'<k} \mathbb{E}_{(a,b,\underbrace{c_1,\ldots,c'_k}_{C})\sim\mu} I(X_a : X_b|\underbrace{X_{c_1},\ldots,X_{c'_k}}_{Z}) \leq \frac{\ln(d)}{k},$$

where μ is the distribution of tuples of indices without replacement.

2.4 The proof

Measure ρ with the IC-POVM $\Lambda^{\otimes n}$ and let the resulting classical systems be $X_1, \ldots, X_n$, with joint probability distribution p.

Now, define

$$\Delta_z(a,b) = \|p_{X_aX_b}|_{Z=z} - (p_{X_a}|_{Z=z}) \otimes (p_{X_b}|_{Z=z})\|$$

if $a \neq b$ and $a, b \notin C$, and set $\Delta_z(a,b)$ to be 0 otherwise. Pinsker tells us

$$\Delta_z^2(a,b) \leq 2I(X_a : X_b|Z = z).$$

So we get for each fixed C,

$$\mathbb{E}_{z\sim p_Z}\mathbb{E}_{a,b\in[n]}\Delta_z(a,b)^2 \leq 2\,\mathbb{E}_{(a,b)\sim\mu_C} I(X_a : X_b|Z),$$

where μ_C is the distribution over pairs a, b that are distinct and not contained in C. The next step is just to use a very simplistic bound: if I sample a random pair $a, b \in [n]$ (with replacement), the chance that it forms an edge is $(nD)/n^2 = D/n$. So we can just blow up the average above by a factor of n/D:

$$\mathbb{E}_{z\sim p_Z}\mathbb{E}_{(a,b)\in E}\Delta_z(a,b)^2 \leq \frac{n}{D}\cdot\mathbb{E}_{z\sim p_Z}\mathbb{E}_{a,b\in[n]}\Delta_z(a,b)^2 \leq \frac{2n}{D}\cdot\mathbb{E}_{(a,b)\sim\mu_C} I(X_a:X_b|Z).$$

We wanted to control the average of Δ, not Δ^2. But we can use Jensen's inequality for this:

$$\mathbb{E}_x\sqrt{f(x)} \leq \sqrt{\mathbb{E}_x f(x)}.$$

Applying this, we have

$$\mathbb{E}_{0\leq k'<k}\mathbb{E}_C\mathbb{E}_{z\sim p_Z}\mathbb{E}_{(a,b)\in E}\Delta_z(a,b) \leq \sqrt{\mathbb{E}_{0\leq k'<k}\mathbb{E}_C\frac{2n}{D}\cdot\mathbb{E}_{(a,b)\sim\mu_C} I(X_a:X_b|Z)} \leq \sqrt{\frac{2n\ln(d)}{kD}}.$$

Now, how do we get the product state out? Just pick a random C of random size $0\leq k'<k$, measure and get a random z. Define the CQ state

$$\tau = \mathbb{E}_{0\leq k'<k}\mathbb{E}_C\mathbb{E}_{z\sim p_Z}(\rho_z)_{\bar{C}}\otimes\left(\bigotimes_{i\in C}(I/d)_i\right),$$

where ρ_z is the normalized state after conditioning the C subsystem of ρ on value z. In other words, we measured and destroyed the C registers to get outcome z, then replaced the C register in the maximally mixed state (any state on C would work). Our product state approximation will be

$$\sigma = \mathbb{E}_{0\leq k'<k}\mathbb{E}_C\mathbb{E}_{z\sim p_Z}\left(\bigotimes_{i\in\bar{C}}(\rho_z)_i\right)\otimes\left(\bigotimes_{i\in C}(I/d)_i\right)$$

The point is as follows. On random $(a,b)\in E$ not contained in C, applying the IC-POVM $\Lambda^{\otimes 2}$ on τ,σ and compare, we get exactly $E_{0\leq k'<k}\mathbb{E}_C\mathbb{E}_{z\sim p_Z}\mathbb{E}_{(a,b)\in E}\Delta_z(a,b)\leq\sqrt{\frac{2n\ln(d)}{kD}}$. Using the $18d$ distortion factor of the IC-POVM, we can convert this into the error on the two-particle marginal. We also have a factor $2k/n$ from the event the chosen edge edge overlaps with qubits in register C, since we will have no control on the approximation there. So overall we get

$$\mathbb{E}_{(a,b)\in E}\|\rho_{a,b}-\sigma_{a,b}\|_1 \leq O\left(18d\sqrt{\frac{2n\ln(d)}{kD}}+\frac{2k}{n}\right).$$

We can now pick k to minimize the RHS.

$$\begin{aligned}
\frac{d}{dk}(a/\sqrt{k}+bk) &= 0\\
-\frac{1}{2}ak^{-3/2}+b &= 0\\
bk^{3/2} &= \frac{1}{2}a\\
k &= (a/2b)^{2/3}\\
\frac{a}{\sqrt{k}}+bk &= a\cdot(2b/a)^{1/3}+b\cdot(a/2b)^{2/3}\\
&= (2^{1/3}+2^{-2/3})\cdot a^{2/3}b^{1/3}.
\end{aligned}$$

So we get

$$\mathbb{E}_{(a,b)\in E} \|\rho_{a,b} - \sigma_{a,b}\|_1 \leq O\left(\left(\frac{d^2 \ln(d)}{D}\right)^{1/3}\right)$$

3 Implications for gap amplification

Suppose we had a procedure that preserves the gap of 2-local H while giving a new 2-local H', with $n \to n' = n^{O(t)}$, $D \to D' \geq D^t$, and $d \to d' = d^t$, for every positive integer t. (a sparse growing sequence, like $t = 2^s$, would also be enough.) Then quantum PCP is false for 2-local Hamiltonians (unless NP = QMA).

Here's the reason: start with a qPCP-hard Hamiltonian H with gap ε. First, blow up the degree to some parameter D by adding dummy edges. This creates H_{pad} with degree D and energy gap at least ε/D. Now if we apply our procedure, we'll get an H' for which the BH theorem says we can approximate the energy up to error

$$O\left(\left(\frac{(d')^2 \ln(d')}{D'}\right)^{1/3}\right) = O\left(\left(\frac{d^{2t} \cdot \ln(d^t)}{D^t}\right)^{1/3}\right).$$

If we picked our initial $D = 8d^3$, then the numerator can be crudely upper bounded by d^{3t}, so overall we get an approximation of

$$O(2^{-t}).$$

We just need to pick t to be big enough for this be smaller than ε/D, which means $t = O(\log D) = O(\log d)$ which is a constant.

But you may say this kind of transformation is too much to ask for, and maybe even this is classically not possible. It turns out that classically this can be done (and in fact increasing the gap) via parallel repetition. If $H = \mathbb{E}_{(i,j)\sim G} h_{ij}$, a two-fold repetition forms

$$H' \;=\; \mathbb{E}_{(i,j),(k,\ell)\sim G}\big(h_{ij} \wedge h_{k\ell}\big)_{(i,k),(j,\ell)}, \tag{6}$$

acting on the paired vertices (i,k) and (j,ℓ). Parameters square: $n' = n^2$, $d' = d^2$, $D' = D^2$. We can take higher fold repetitions.