Lecture 5

Approximate Ground State Projections

Scribes: Maria Okounkova and Fernando Granha Jeronimo

5.1 The Detectability lemma

We start by discussing the *detectability lemma*, which provides the basis for a local constraint satisfaction view of the local Hamiltonian problem. More formally, given a d-local Hamiltonian $H = \sum_{i=1}^{n} H_i$ and a state $|\psi\rangle$ that is "far" from being its ground state in the sense of having a high value for the energy $\langle\psi|H|\psi\rangle$, we want to say that we will be able to "detect" the high energy simply by making the local measurements $H_1, H_2, \ldots, H_n$ in sequence.[1]

The detectability lemma provides some sufficient conditions under which this is possible. We let the H_i be *projections*, and assume that H is *frustration free*, i.e., there is a state $|\Gamma\rangle$ that is simultaneously a ground state of all the H_i (and hence has energy 0). We further assume that the Hamiltonian is *gapped*, i.e., there is a positive constant δ such that any state $|\phi\rangle$ orthogonal to $|\Gamma\rangle$ has energy at least δ: $\langle\phi|H|\phi\rangle \geq \delta$.

Ideally, we would like to prove that whenever $|\phi\rangle$ is of the form $\langle\phi|H|\phi\rangle \geq \delta$, then the probability of failing to detect it using a sequential measurement as above is at most $1 - \Omega(\delta)$, i.e.,

$$\|(I - H_m)(I - H_{m-1}) \ldots (I - H_1)|\phi\rangle\|_2^2 = 1 - \Omega(\delta).$$

However, this statement cannot be true. Suppose for example that there exists a state $|u\rangle$ such that $\langle u|H_i|u\rangle = 1$ for $1 \leq i \leq m$. Then, a simple calculation shows that the state $|\phi\rangle = \sqrt{1-\delta/n}|\Gamma\rangle + \sqrt{\delta/n}|u\rangle$ satisfies $\langle\phi|H|\phi\rangle = \delta$, even though the probability of failure is $1 - \delta/n$. The actual detectability lemma gets around this by only considering states $|\phi\rangle$ that are orthogonal to the ground state $|\Gamma\rangle$.

Lemma 5.1 (Detectability lemma). *Let $H = \sum_{i=1}^{n} H_i$ be a frustration free Hamiltonian such that (i) all H_i are projections, and (ii) each H_i commutes with all but (at most) g other terms H_j. We further assume that $|\Gamma\rangle$ is the unique ground state of H with energy 0, and any state $|\psi\rangle$ orthogonal to $|\Gamma\rangle$ satisfies $\langle\psi|H|\psi\rangle \geq \delta > 0$. Then, for every $|\psi\rangle$ orthogonal to $|\Gamma\rangle$ we have*

$$\|(I - H_m)(I - H_{m-1}) \ldots (I - H_1)|\psi\rangle\|^2 \leq 1 - \Omega\left(\frac{\delta}{g^2}\right).$$

[1] The precise order of the measurements does not matter. The crucial point is that each of the n measurements involves only a small number of qubits, and is the non-commutative analog of classical tests of low query complexity.

Proof. Let P_i be the projection to the ground state of H_i (thus, $H_i = I - P_i$, since H_i are projections). We define

$$|\phi\rangle = P_n P_{n-1} \dots P_1 |\psi\rangle.$$

Note that our goal is to show that $\| |\phi\rangle \|^2 = 1 - \Omega(\delta/g^2)$. We now consider the term $\langle\phi|H_i|\phi\rangle$. Observe that if $P_m, P_{m-1}, \dots, P_{i+1}$ commute with H_i then $\langle\phi|H_i|\phi\rangle = 0$. So, let $i_1 > i_2 > \cdots > i_g$ be the (possible) indices greater than i such that H_{i_k}, $1 \le k \le g$ do not commute with H_i. We denote the set of these indices by $N(i)$. We then have

$$\begin{aligned} H_i|\phi\rangle &= P_m \dots P_{i_1+1} H_i P_{i_1-1} \dots P_1|\psi\rangle - P_m \dots P_{i_1+1} H_i H_{i_1} P_{i_1-1} \dots P_1 |\psi\rangle, \\ &\quad \text{using } I - H_{i_1} = P_{i_1}, \\ &= - \sum_{j \in N(i)} \Big(\prod_{\substack{k>j \\ k \notin N(i)}} P_k \Big) H_i H_j P_{j-1} \dots P_1 |\psi\rangle, \\ &\quad \text{by iterating the last step for each } i_k, 1 \le k \le g. \end{aligned}$$

Now, since H_i is a projection, we have

$$\begin{aligned} \langle\phi|H_i|\phi\rangle &= \|H_i|\phi\rangle\|^2 \\ &\le g \sum_{j \in N(i)} \left\| H_j P_{j-1} \dots P_1 |\psi\rangle \right\|^2, \\ &\quad \text{using the Schwarz inequality, followed by neglecting some projections,} \\ &= \sum_{j \in N(i)} \left[\left\| P_{j-1} \dots P_1 |\psi\rangle \right\|^2 - \left\| P_j P_{j-1} \dots P_1 |\psi\rangle \right\|^2 \right], \text{ since } H_j \text{ are projections.} \end{aligned}$$

Thus, summing over i, we have

$$\begin{aligned} \langle\phi|H|\phi\rangle &\le g \sum_{i=1}^{n} \sum_{j \in N(i)} \left[\left\| P_{j-1} \dots P_1 |\psi\rangle \right\|^2 - \left\| P_j P_{j-1} \dots P_1 |\psi\rangle \right\|^2 \right] \\ &\le g^2 \sum_{j=1}^{n} \left(\left\| P_{j-1} \dots P_1 |\psi\rangle \right\|^2 - \left\| P_j P_{j-1} \dots P_1 |\psi\rangle \right\|^2 \right), \quad (5.1) \\ &\quad \text{since each of the terms can appear for at most } d \text{ values of } i, \\ &= g^2 \left(1 - \|P_n \dots P_1 |\psi\rangle\|^2 \right) = d^2 \left(1 - \| |\phi\rangle \|^2 \right). \quad (5.2) \end{aligned}$$

Now, we note that $|\phi\rangle$ is orthogonal to the ground state $|\Gamma\rangle$:

$$\langle\phi|\Gamma\rangle = \langle\psi|P_1 \cdots P_n|\Gamma\rangle = \langle\psi|\Gamma\rangle = 0,$$

where for the second equality we used $P_i|\Gamma\rangle = |\Gamma\rangle$ for all i. Hence from the hypothesis of the theorem, we have $\langle\phi|H|\phi\rangle \ge \delta \| |\phi\rangle \|^2$. Using the above inequality, we then have

$$\| |\phi\rangle \|^2 \le \frac{1}{1 + \frac{\delta}{g^2}},$$

which completes the proof. $\square$

Remark 5.2. A partial converse to the main inequality, (5.2), is given by the "quantum union bound." Under the same assumptions about H, but without the restriction on limited non-commutation, this states that the inequality $\||\phi\rangle\|^2 \geq 1 - 4\langle\psi|H|\psi\rangle$. In case where all H_i are diagonal in the computational basis and ker H_i is interpreted as an "event" in the space of all bit strings, this recovers the classic union bound (up to the factor of 4...).

5.2 Decay of correlations

A consequence of the detectability lemma is decay of correlations, which we now discuss.

Theorem 5.3 (Decay of Correlations)**.** *Suppose H is a geometrically-local Hamiltonian on a n-qubit D-dimensional grid with spectral gap $\delta > 0$. Let X, Y be Hermitian operators on $(\mathbb{C}^2)^{\otimes n}$ such that $d(X, Y) \geq m$ (here $d(X, Y)$ is the length of the shortest path on the grid between the patches X and Y act on), then,*

$$\big|\langle\Gamma|X|\Gamma\rangle\langle\Gamma|Y|\Gamma\rangle - \langle\Gamma|X \otimes Y|\Gamma\rangle\big| \leq \|X\|\|Y\| \cdot e^{-\Omega(m\cdot\delta/D^3)} \tag{5.3}$$

In words, this theorem says that the further apart the operators X and Y are, the less correlated the results will be when we measure the ground state. Here is an example where such decay does *not* happen:

Example 5.4. Consider the CAT state $|\psi\rangle = \dfrac{1}{\sqrt{2}}(|0\ldots0\rangle + |1\ldots1\rangle)$ and let X be the Pauli Z operator on the ith quit, and Y be the Pauli Z operator on the jth qubit. Since our state is already expressed in the computation basis, we can easily see that $\langle\psi|X|\psi\rangle = -1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = 0$ and likewise $\langle\psi|Y|\psi\rangle = 0$. This corresponds to acting with X and Y independently. However, this state is highly correlated for these operators, as, acting on $|\psi\rangle$ first with X projects $|\psi\rangle$ into the state either $|0\ldots0\rangle$ or $|1\ldots1\rangle$, where the outcome of the Y measurement is fully determined (the outcomes are always perfectly correlated). Indeed, we can compute $\langle\psi|X \otimes Y|\psi\rangle = 1$. So te difference $|\langle\psi|X|\psi\rangle\langle\psi|Y|\psi\rangle - \langle\psi|X \otimes Y|\psi\rangle| = 1$ and does not decay exponentially: this state cannot be the ground state of a local, gapped Hamiltonian!

Now let's prove Theorem 5.3. We will use the detectability lemma, for a well-chosen ordering of the H_i terms.

Proof of Theorem 5.3. The proof is best described via a picture, see Figure 5.1. Consider a tensor network representation of $|\Gamma\rangle$with X and Y acting on sets of qubits at distance m. This is represented here in one dimension, but works in any number of dimensions. The picture, without the $1 - H_i$ terms inserted in the middle, corresponds to a tensor nertwork representation of $\langle\Gamma|X \otimes Y|\Gamma\rangle$. Let's consider inserting $1 - H_i$ terms (projectors). Any term which is not in the "causal" cone of X *and* Y (as pictured) can be "absorbed" into the top or bottom states $|\Gamma\rangle$, using $(1 - H_i)|\Gamma\rangle = |\Gamma\rangle$ for any i since $|\Gamma\rangle$ is the ground state and the Hamiltonian is frustration-free. A layer of corresponds to the product of pairwise commuting terms. Inserting all the Hamiltonians requires 2 layers in 1 dimension, and $2D$ layers in dimension D ($2D$ is just the maximum degree). How many layers can we insert? The spacing m between the operators X and Y. Hence overall we can insert $l = m/2D$ copies of the operator $A = \prod_i^m(\mathbb{I} - H_i)$ from the detectability lemma, without changing the value of the tensor network:

$$\langle\Gamma|X \otimes Y|\Gamma\rangle = \langle\Gamma|(X \otimes \mathbb{I})A^l(Y \otimes \mathbb{I})|\Gamma\rangle$$

The detectability lemma implies that

$$\|A^l - |\Gamma\rangle\langle\Gamma|\| \leq (1 - \Omega(\delta/D^2))^l \approx e^{-\Omega(\delta l/D^2)}.$$

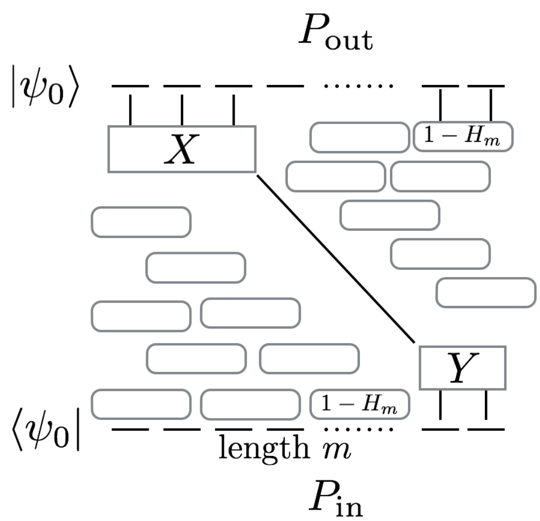

Figure 5.1: Visual for decay of correlations proof

Recall that we want to bound

$$\begin{aligned}
&|\langle\Gamma|X|\Gamma\rangle\langle\Gamma|Y|\Gamma\rangle - \langle\Gamma|X \otimes Y|\Gamma\rangle| \\
=&|\langle\Gamma|(X \otimes \mathbb{I})(A^{\frac{m}{2D}} - |\Gamma\rangle\langle\Gamma|)(\mathbb{I} \otimes Y)|\Gamma\rangle| \\
\leq&\|X\| \cdot \|Y\| \cdot \|A^{\frac{m}{2D}} - |\Gamma\rangle\langle\Gamma|\| \leq |X\|\|Y\|e^{-\Omega(m\delta/D^3)},
\end{aligned}$$

which completes the proof. □

Before we discuss the area law, which will essentially tell us that ground states of gapped local Hamiltonians have much less entanglement than the maximum allowed in "generic" states, let's introduce entropy, specifically Von Neumann entropy, as a measure of entanglement.

Definition 5.5 (von Neumann Entropy)**.** Consider any $|\psi\rangle \in H_A \otimes H_B$, and perform a Schmidt decomposition to obtain $|\psi\rangle = \sum_{i=1}^{d} \lambda_i |u_i\rangle_A |v_i\rangle_B$. The von Neumann entropy of $|\psi\rangle$ across region A is defined as

$$S(A)_{|\psi\rangle} = S\big((\lambda_i^2)_i\big) = \sum_i \lambda_i^2 \ln \frac{1}{\lambda_i^2}. \tag{5.4}$$

Example 5.6. If $|\psi\rangle = |L\rangle \otimes |R\rangle$ then $S(A)_{|\psi\rangle} = 0$ (the product state has no entanglement).

The maximally entangled state $|\psi\rangle = \sum_i^D \frac{1}{\sqrt{D}}|i\rangle_A \otimes |i\rangle_B$, has entropy $S(A)_{|\psi\rangle} = \ln(D)$. In general, $S(A)_{|\psi\rangle} \leq \ln(\dim \text{Hilbert space})$. This is a direct quantum generalization of the fact that a classical probability distribution on D elements has entropy at most $\ln D$, and this is achieved by the uniform distribution.

Using this definition, we can now discuss the Area law, stated as a conjecture.

Conjecture 5.7 (Area Law). *Let $|\Gamma\rangle$ be the ground state of H (the type of Hamiltonian discussed at the beginning of these notes). Then, for any region A,*

$$S(A)_{|\Gamma\rangle} \leq \ln\left(\dim \partial A\right), \tag{5.5}$$

where ∂A denotes the boundary *of region A, the set of qubits that interact (through H) with at least one qubit outside of region A.*

In other words, the area law conjectures that, in gound states of gapped local Hamiltonians, entropy scales are the surface area, rather than the maximum, which would be the volume. This has recently been proven for 1-dimensions, and we will see a complete proof in this and the next lecture. The conjecture is open for any dimension at least 2.

The proof makes essential use of the following ingredient called Approximate Ground-State Projection (AGSP):

Definition 5.8 (AGSP). Given a Hamiltonian H, a (D, Δ)-AGSP for H is an operator $K \in (\mathbb{C}^2)^{\otimes N}$ such that

- $K|\Gamma\rangle = |\Gamma\rangle$
- $\forall |\psi\rangle : \langle\psi|\Gamma\rangle = 0$, $\|K|\psi\rangle\|^2 \leq \Delta \| |\psi\rangle\|^2$
- If $|\psi\rangle$ has a tensor-network representation with bond dimension $\leq L$, then $K|\psi\rangle$ has a tensor-network representation with bond dimension $\leq D \cdot L$.

Example 5.9. For example, the operator $A = (\mathbb{I} - H_m) \ldots (\mathbb{I} - H_1)$ from the detectability lemma is an AGSP. The detectaibility lemma states that its Δ parameter is $\Delta = 1 - \Omega(\delta)$. To evaluate B, observe that applying any $(\mathbb{I} - H_i)$ can be done at the cost of a multiplicative blow-up by a factor at most 4 (the dimension of the space on which $(\mathbb{I} - H_i)$ acts non-trivially) across the bond on which H_i acts. Since each bond is acted on by only one operator, the total blow-up is $D = 4$.

We will prove a theorem which is a bit more general than the area law, namely

Theorem 5.10. *Suppose there exists a (D, Δ)-AGSP such that $D\Delta \leq \frac{1}{2}$. Then $|\Gamma\rangle$ satisfies an area law of the form $S(A)_{|\Gamma\rangle} \leq O(1) \log D$.*

First we prove the theorem, based on the following two lemmas. Then we will see a construction of the required (B, Δ)-AGSP. Indeed, note that the AGSP from the detectability lemma does not quite make the cut, since $B\Delta \approx 4$ for small constant δ.

Lemma 5.11. *Suppose $\exists(D, \Delta)$- AGSP such that $D\Delta \leq \frac{1}{2}$. Fix a partition $(A, \bar{A})$ of the space on which the Hamiltonian acts. Then there exists a product state $|\phi\rangle = |L\rangle_A \otimes |R\rangle_{\bar{A}}$ such that $|\langle\phi|\Gamma\rangle| = \mu \geq \frac{1}{\sqrt{2D}}$.*

Proof. Let $|\phi\rangle$ be a product state with the largest overlap on $|\Gamma\rangle$, meaning that is maximizes $\mu = |\langle\phi|\Gamma\rangle|$, and can be expressed as $|\phi\rangle = \mu|\Gamma\rangle + \sqrt{1-\mu^2}|\psi^\perp\rangle$ (where the latter is some state orthogonal to the ground state). Apply K to get $K|\phi\rangle = \mu|\Gamma\rangle + \delta|\psi'\rangle$ where $|\psi'\rangle$ is normalized and $|\delta|^2 \leq \Delta$. The Schmidt decomposition of $K|\Gamma\rangle$ has at most B terms, so we can express, using Cauchy-Schwarz,

$$\mu = |\langle\Gamma|K\phi\rangle| = |\sum_i \lambda_i \langle\Gamma|L\rangle_i|R\rangle_i| \leq \sqrt{\sum \lambda_i^2}\sqrt{\sum \langle\Gamma|L\rangle_i|R\rangle_i^2} \leq \sqrt{(\mu^2+\Delta)}\sqrt{D}\sqrt{\max_i \langle\Gamma|L\rangle_i|R\rangle_i}$$

Thus there exists a product state such that $|\langle\Gamma|L\rangle_i|R\rangle_i| \geq \frac{\mu}{\sqrt{D(\mu^2+\Delta)}}$. This must be $\leq \mu$ by assumption, and hence $\sqrt{D}\sqrt{\mu^2+\Delta} \geq 1$ meaning that $\mu^2 \geq \frac{1}{D} - \Delta \geq \frac{1}{D} - \frac{1}{2D} = \frac{1}{2D}$. □

Before we move to the next lemma, we need an auxiliary result known as Eckart-Young Theorem. This result bounds the maximum possible overlap between states $|\psi\rangle$ and $|\phi\rangle$ where the latter has a predefined Schmidt Rank.

Theorem 5.12 (Eckart-Young)**.** *Let $|\psi\rangle \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ be a normalized vector with Schmidt decomposition $|\psi\rangle = \sum \lambda_i |u_i\rangle|v_i\rangle$, where $\lambda_1 \geq \lambda_2 \geq \cdots$. Then for any normalized $|\phi\rangle \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2}$ with Schmidt rank at most B it holds that*

$$|\langle\psi|\phi\rangle| \leq \sqrt{\sum_{i=1}^{B} \lambda_i^2}.$$

Given a product state that has overlap μ with the ground state, the next lemma uses an AGSP to bound the entanglement across the associated cut $(i^\star, i^\star+1)$.

Lemma 5.13. *Suppose $\exists(D,\Delta)$-AGSP and $\exists$ a product state $|\phi\rangle = |L\rangle_A \otimes |R\rangle_{\bar{A}}$ such that $|\langle\phi|\Gamma\rangle| = \mu$. Then*

$$S((i^\star, i^\star+1))_{|\Gamma\rangle} \leq O(1)\frac{\log\mu}{\log\Delta}\log D \tag{5.6}$$

Proof. Denote by K the (D,Δ)-AGSP.Define $|\phi_l\rangle = \frac{K^l|\phi\rangle}{\|K^l|\phi\rangle\|}$. It is not hard to see that $|\phi_l\rangle$ is such that

(i) $SR(|\phi_l\rangle) \leq D^l$;

(ii) $|\langle\Gamma|\phi_l\rangle| \geq \frac{\mu}{\sqrt{\mu^2+\Delta^l(1-\mu^2)}}$.

Property (i) follows from $SR(|\phi\rangle) = 1$ and the AGSP having bond parameter D. Furthermore, property (ii) follows from the definition of the shrinking parameter Δ.

Let $|\Gamma\rangle = \sum \lambda_i |L_i\rangle|R_i\rangle$ be the Schmidt decomposition of the ground state relative to cut $(i^\star, i^\star+1)$, where $\lambda_1 \geq \lambda_2 \geq \ldots$. By the Eckart-Young Theorem we have

$$\sum_{i=1}^{D^l} \lambda_i^2 \geq |\langle\Gamma|\phi_l\rangle|^2 \geq \frac{\mu^2}{\mu^2+\Delta^l(1-\mu^2)}$$

or equivalently

$$\sum_{i>D^l} \lambda_i^2 \leq 1 - \frac{\mu^2}{\mu^2 + \Delta^l(1-\mu^2)} \leq 1 - \frac{\mu^2}{\mu^2 + \Delta^l} \leq \frac{\Delta^l}{\mu^2}.$$

We choose $l_0 = 2\frac{\log \mu}{\log \Delta} - \frac{\log 2}{\log \Delta}$ such that $\frac{\Delta^{l_0}}{\mu^2} \leq \frac{1}{2}$ and proceed to bound the worst case entropy across the AGSP cut. The first D^{l_0} Schmidt coefficients account for an entropy of at most $l_0 \log D$. For the remaining coefficients, we group them in chunks of size D^{l_0} in intervals $[D^{kl_0}+1, D^{(k+1)l_0}]$ indexed by k. For each of these intervals, the corresponding entropy can be upper bounded by

$$\frac{\Delta^{kl_0}}{\mu^2} \log D^{(k+1)l_0} = l_0 \frac{1}{2^k}(k+1) \log D,$$

where here $\frac{\Delta^{kl_0}}{\mu^2}$ is an upper bound on the total probability mass in the interval, and $D^{-(k+1)l_0}$ a lower bound on the size of any Schmidt coefficient (squared) in the interval, which follows from the fact that they are organized in descending order and must sum to 1. Therefore, the total entropy is

$$\begin{aligned} S((i^\star, i^\star+1)) &\leq l_0 \log D + \sum_{k\geq 1} l_0 \frac{1}{2^k}(k+1) \log D \\ &\leq l_0 \log D + l_0 \log D \sum_{k \geq 1} \frac{1}{2^k}(k+1) \\ &\leq O(1) l_0 \log D \\ &\leq O(1) \frac{\log \mu}{\log \Delta} \log D \end{aligned}$$

We note that this bound is sufficient to imply the area law as our AGSP construction will have constant Δ and B (and hence μ). □