CS 294–66 **Lecture 7: Area Laws and Tensor Networks** **October 14, 2016**

Lecturer: U. Vazirani **Scribe: B. O'Gorman**
A. Szasz

1 Preliminaries

Consider a bipartite system whose Hilbert space is $\mathcal{H}_{AB} = \mathcal{H} \otimes \mathcal{H}_B$, where $\mathcal{H}_A$ and $\mathcal{H}_B$ are m- and n-dimensional, respectively. An arbitrary state $|\psi\rangle \in \mathcal{H}_{AB}$ has a unique *Schmidt decomposition*

$$|\psi\rangle = \sum_{k=1}^{\min\{m,n\}} c_k |v_k\rangle \otimes |w_k\rangle, \tag{1}$$

where $\{|v_k\rangle | k \in [m]\} \subset \mathcal{H}_A$ is an orthonormal basis of $\mathcal{H}_A$, and similarly for B. The *Schmidt rank* of a state $|\psi\rangle$ is the number of non-zero c_j in its Schmidt decomposition. Specifically, a seperable state $|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$ has Schmidt rank 1. At the other extreme, a state with maximal Schmidt rank is known as a *generalized cat state*, so called because it generalizes the (bipartite) *cat state*

$$|\text{cat}\rangle = m^{-1/2} \sum_{j=1}^{m} |j\rangle \otimes |j\rangle, \tag{2}$$

when $m = n$.

The Schmidt decomposition of an arbitrary state $|\psi\rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{i,j} |i\rangle \otimes |j\rangle$ is the solution (V, W, K) to the system of equations

$$VAW^\dagger = C, \tag{3}$$

where

$$V = \sum_{i=1}^{m} |i\rangle \langle v_i|, \quad W = \sum_{i=1}^{n} |i\rangle \langle w_i|, \quad A = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{i,j} |i\rangle \langle j|, \quad C = \sum_{i=1}^{\min\{m,n\}} c_i |i\rangle \langle i|. \tag{4}$$

Note that this is equivalent to

$$A = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{i,j} |i\rangle \langle j| = V^\dagger C W = \sum_{i=1}^{\min\{m,n\}} c_i |v_i\rangle \langle w_i|, \tag{5}$$

which we could also write as

$$|\psi\rangle \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{i,j} |i\rangle \otimes |j\rangle = \sum_{i=1}^{\min\{m,n\}} c_i |v_i\rangle \otimes |w_i\rangle. \tag{6}$$

That is, the bases $\{|v_i\rangle\}$ of $\mathcal{H}_A$ and $\{|w_i\rangle\}$ of $\mathcal{H}_B$ are such that the state $|\psi\rangle$ is "diagonal" in the composite basis $\{|v_i\rangle \otimes |w_j\rangle\}$, in the sense that

$$(V \otimes W) |\psi\rangle = \sum_{i=1}^{\min\{m,n\}} |i\rangle \otimes |i\rangle. \tag{7}$$

The *entanglement entropy* of a state $|\psi\rangle = \sum_i c_i |v_i\rangle \otimes |w_i\rangle$ is defined as $H(\{|c_i|^2\}) = -\sum_i |c_i|^2 \log(|c_i|^2)$, which is equivalent to the von Neumann entropy of the reduced state of either subsystem.

2 Area laws: Introduction

Consider a D-dimensional lattice of qudits and a Hamiltonian $H = \sum_{H_j}$, where each H_j only acts nontrivially on pairs of adjacent qudits. Let E_i be the ith lowest eigenvalue of H, $\epsilon = E_1 - E_0 > 0$, and $|\Omega\rangle$ be the ground state. We say that H obeys an *area law* if the entanglement entropy across a cut dividing the lattice into a set of qubits A and its complement B is at most proportional to the surface area of the cut, where the "area" of a cut a is the number of terms H_j acting on qubits in both A and B.

Now, we focus on the 1-dimensional case. First, note that we can reduce a (geometrically) k-local Hamiltonian on a line of qudits into a 2-local Hamiltonian on a line of larger-dimensional qudits by combining adjacent chunks of qudits into "super-qudits". Therefore, we henceforth assume without loss of generality that we have 2-local Hamiltonian on a line of n d-dimensional qudits:

$$H = \sum_{i=1}^{n-1} H_{j,j+1} = \sum_{i=1}^{n-1} H_j. \tag{8}$$

We say that a family of Hamiltonians is *gapped* if the gap ϵ between the two lowest eigenvalues is independent of the system size. Hastings showed that the entanglement entropy across any cut of the ground state of a 1D gapped Hamiltonian is at most $\exp(O(\log d/\epsilon))$ [1]. (Note that this is an area law because in 1D, the surface area of a cut is constant.)

In 1D, the area law implies that there is a succinct classical description of the ground state (using a tensor network). Thus finding the ground state of a *gapped* 1D local Hamiltonian is in NP. Recall, however, that without the gap condition, finding the ground state of a 1D local Hamiltonian is QMA-complete.

More recently, Arad et al. proved the tighter upper bound $\tilde{O}(\log^3 d/\epsilon)$ [2]. In the remainder of these notes, we will sketch the proof of this result.

3 Proof that 1D, gapped H implies area law entanglement

This proof is due to Arad, Kitaev, and Landau. The result is that the entanglement scales as

$$S \leq \mathcal{O}\left(\frac{\log^3 d}{\epsilon}\right) \tag{9}$$

where d is the dimensionality of the Hilbert space on each lattice site.

> **Note:** If you could get $\mathcal{O}(\log d/\epsilon)$ instead, then you could actually prove the area law in any number of dimensions! In spacial dimension D, a shell of sites of dimension d would have total dimension proportional to $d^{L^{D-1}}$ (if the linear dimension of the shell, eg the radius for a (hyper)sphere, is L) and we could think of this as a single 1D site. Then $\mathcal{O}(\log d)$ would become $\mathcal{O}(L^{D-1})$, ie proportional to the surface area of the shell. This is precisely area law entanglement!
>
> In two dimensions, a volume law would be $\mathcal{O}(m^2)$, so even a 1D result finding $\mathcal{O}(\log^{2-\mu} d)$ for any $\mu > 0$ would indicate sub-volume law entanglement scaling. (In fact $\mathcal{O}(\log^{3-\mu} d)$ would also lead to sub-volume law scaling, though that cannot be proven directly from the result for 1D but rather involves some details of the proof. This means that the bound we prove here for 1D is the best you can do without implying sub-volume law scaling in dimensions higher than 1.) **End of Note**

We begin the proof by writing the ground state $|\Omega\rangle$ via a Schmidt decomposition:

$$|\Omega\rangle = \sum_j c_j |\psi_A^{(j)}\rangle|\psi_B^{(j)}\rangle \tag{10}$$

We want to show that the entanglement entropy of the ground state across a cut is proportional to l, the surface area of the cut (ie a constant in 1D).

Aside: The most general area law we could try to prove would be defined on a graph as follows:

- Represent a system as a graph G with vertices for the lattice sites/particles and edges for the interactions
- The generalized area law would be true if for an arbitrary cut through the graph with a local, gapped Hamiltonian, the ground state would have area law entanglement across the cut (ie proportional to the number of cut edges)

Unfortunately there exists a counterexample, so the generalized area law is not true. **End of Aside**

Some assumptions:

For the remainder of the lecture, we assume that the Hamiltonian is frustration free. So we can assume that $H|\Omega\rangle = 0$. (Note that a similar, albeit more complicated, approach will work for H with frustration.) We can approach this by starting with a product state and iteratively modifying it to get constant overlap with Ω.

Additionally, for the remainder of the lecture any state denoted by $|\Omega^\perp\rangle$ will be assuemd to be both normalize and perpendicular to the ground state $|\Omega\rangle$, so that

$$\langle\Omega^\perp|\Omega\rangle = 0 \ \text{ and } \ \langle\Omega^\perp|\Omega^\perp\rangle = 1 \tag{11}$$

A note on our proof strategy:

As part of the proof, we will show that the ground state is "similar to" a product state, in the sense that there exists a product state that has constant (with respect to the system size) overlap with the ground state. Since we are looking to prove area law entanglement entropy, while product states have actually 0 entanglement, we might worry that a large constant overlap with a product state is proving "too much" and that an easier proof is available. However, for a lattice with n sites there must be a single Schmidt coefficient (ie overlap with some product state) that is larger than $2^{-n/4}$ to avoid volume-law entanglement entropy: for a lattice with n sites, the max Schmidt rank across any one cut is $2^{n/2}$, and thus the state with maximum entanglement entropy across a cut has all Schmidt coefficients equal to $2^{-n/4}$, giving $2^{n/2}\left(-2^{-n/2}\log(2^{-n/2})\right) = n\log(2)/2$ for the entanglement entropy; to avoid this volume-law scaling, at least one Schmidt coefficient must be larger than $2^{-n/4}$.

More generally, suppose that the Schmidt coefficients across a cut are ordered such that $|c_1| > |c_2| > \cdots$. Then the minimum entanglement entropy across the cut, given the value of c_1, occurs when the remaining weight is distributed as unevenly as possible, ie with $c_1 = c_2 = c_3 = \cdots = c_m$ up to some $m < 2^{n/2}$ and $c_r = 0$ for $r > m+1$. (c_{m+1} is somewhere between, with the "leftover" weight needed to sum to 1.) The total entanglement entropy in that case is between $-m\,c_1^2\log(c_1^2)$ and $-(m+1)c_1^2\log(c_1^2)$. Since the absolute squares of the Schmidt coefficients sum to 1, we have $m \approx c_1^{-2}$, so that the minimum entanglement entropy, given c_1, is approximately $-\log(c_1^2)$. Then if there is area law entanglement in 1D, so that the entanglement does not grow with system size, c_1 cannot shrink below some minimum value as the system size increases. Thus a constant (wrt system size) overlap with some product state is indeed a necessary condition to have area law entanglement. (Note that it is not a sufficient condition: set c_2 through $c_{2^{n/2}}$ to be all equal, and the entanglement will obey a volume law.)

3.1 Introducing the idea of an AGSP

To show that the ground state has low (in fact constant) entanglement entropy, we will start with a state that has low entanglement, but which may have only a small overlap with the ground state $|\Omega\rangle$. We then iteratively apply an Approximate Ground State Projector, or AGSP. Each application of this operator will shrink the part of the state that is perpendicular to $|\Omega\rangle$ by at least a factor Δ, while increasing the Schmidt rank of the state by up to a factor D. Repeated iterations will produce a state that both has high overlap with the ground state and a Schmidt rank that is not too large, since we require a favorable tradeoff between overlap and Schmidt rank, as guaranteed by the condition $D\Delta \le 1/2$.

Such a (D, Δ)-AGSP, which we call K, must in particular have the following properties:

(a) $K|\Omega\rangle = |\Omega\rangle$

(b) $K|\Omega^\perp\rangle$ remains perpendicular to $|\Omega\rangle$, ie $\langle\Omega|\Omega^\perp\rangle = 0$, and $\left\|K|\Omega^\perp\rangle\right\|^2 \le \Delta$ for some small Δ. Taken together with (a), this ensures that K doesn't affect the part proportional to $|\Omega\rangle$ while shrinking the part that is perpendicular. After renormalizing the resulting state, it becomes closer to the ground state $|\Omega\rangle$.

(c) The Schmidt rank of K is less than or equal to D, where the Schmidt rank of an operator is defined by writing

$$K = \sum_{j=1}^{D} K_A^{(j)} \otimes K_B^{(j)} \tag{12}$$

for D as small as possible.

3.2 Consequences of the existence of an AGSP

The proof of the area law can be broken into two parts: first, we must actually find such a (D, Δ)-AGSP, and second, we must use it to prove the area law. In practice, we will do the proof in the opposite order, first proving two consequences of having such an AGSP, and then actually constructing the desired operator. The two claims we prove are as follows:

Claim 1: The existence of such a (D, Δ)-AGSP implies that there exists a product state $|\phi_A\rangle \otimes |\phi_B\rangle$ such that $\langle\Omega|\phi_A \otimes \phi_B\rangle \geq \frac{1}{\sqrt{2D}}$

Claim 2: If such a (D, Δ)-AGSP exists, then the entanglement entropy in $|\Omega\rangle$ across the cut between A and B is $\mathcal{O}(\log D)$.

Proof of claim 1: Consider the product state $|\phi_A\rangle \otimes |\phi_B\rangle$ that has the maximum possible overlap with the ground state $|\Omega\rangle$. Suppose that this overlap has the value a, $\Omega|\phi_A \otimes \phi_B\rangle = a$, so that

$$|\phi_A\rangle \otimes |\phi_B\rangle = a|\Omega\rangle + \sqrt{1-a^2}|\Omega^\perp\rangle \tag{13}$$

Now apply the (D, Δ)-AGSP, K, to both sides of the equation to get

$$K|\phi_A \otimes \phi_B\rangle = a|\Omega\rangle + b|\Omega^\perp\rangle \tag{14}$$

for some $b \leq \sqrt{\Delta}\sqrt{1-a^2}$. Normalizing this state gives

$$K|\phi_A \otimes \phi_B\rangle \to \frac{a}{\sqrt{a^2+b^2}}|\Omega\rangle + \frac{b}{\sqrt{a^2+b^2}}|\Omega^\perp\rangle \tag{15}$$

so that after normalizing the overlap with Ω is

$$\frac{a}{\sqrt{a^2+b^2}} \geq \frac{a}{\sqrt{a^2+\Delta}} \tag{16}$$

Unfortunately, the state $K|\phi_A \otimes \phi_B\rangle$ may have a Schmidt rank as large as D.

Proof: We can write out K using its Schmidt decomposition so that

$$K|\phi_A \otimes \phi_B\rangle = \sum_{j=1}^{D} \left(K_A^{(j)}|\phi_A\rangle\right) \otimes \left(K_B^{(j)}|\phi_B\rangle\right) \tag{17}$$

which is a state that explicitly has rank up to but no more than the rank of the operator K.

$K|\phi_A \otimes \phi_B\rangle$ can thus be decomposed into D Schmidt vectors. As these are orthonormal, we can conclude that at least one of the Schmidt vectors has an overlap with $|\Omega\rangle$ that is at least $1/\sqrt{D}$ times as large.

Proof: Let $|\phi\rangle$ be the normalized state proportional to $K|\phi_A \otimes \phi_B\rangle$. If we write $|\phi\rangle$ as

$$|\phi\rangle = \sum_{j=1}^{D} c_j|\phi_A^{(j)}\rangle|\phi_B^{(j)}\rangle, \tag{18}$$

from the Cauchy-Schwarz inequality we have

$$|\langle\phi|\Omega\rangle|^2 \leq \langle\phi|\phi\rangle \cdot \langle\Omega|\Omega\rangle = \langle\phi|\phi\rangle = \sum_{j=1}^{D} c_j^2 \leq D \times \max(\{c_j^2\}) \tag{19}$$

so that the largest of the c_j is at least $1/\sqrt{D}$ times $|\langle\phi|\Omega\rangle|$.

This gives a new product state (the Schmidt vector of $K|\phi_A \otimes \phi_B\rangle$ with the largest overlap with $|\Omega\rangle$) whose overlap with the ground state is at least

$$\frac{1}{\sqrt{D}} \frac{a}{\sqrt{a^2+\Delta}} \tag{20}$$

By assumption this overlap cannot be larger than the maximum overlap a, so that

$$\frac{1}{\sqrt{D}} \frac{a}{\sqrt{a^2+\Delta}} \leq a \Rightarrow D(a^2+\Delta) \geq 1 \Rightarrow Da^2 \geq 1 - D\Delta \geq \frac{1}{2} \Rightarrow a \geq \frac{1}{\sqrt{2D}} \tag{21}$$

where we used the fact that $D\Delta \leq 1/2$. This completes the proof of claim 1.

Proof of claim 2: We begin from the results of claim 1, so that there exists some product state $|\phi_A \otimes \phi_B\rangle$ satisfying

$$\langle\Omega|\phi_A \otimes \phi_B\rangle \geq \frac{1}{\sqrt{2D}}. \tag{22}$$

Since the overlap of $|\Omega\rangle$ with $|\phi_A \otimes \phi_B\rangle$ is at least $1/\sqrt{2D}$, we can write $|\phi_A \otimes \phi_B\rangle$ as

$$|\phi_A \otimes \phi_B\rangle = a|\Omega\rangle + b|\Omega^\perp\rangle \tag{23}$$

where $a \geq 1/\sqrt{2D}$ and $b \leq \sqrt{1 - 1/(2D)}$. If we now apply the operator K to the state, the second term shrinks by at least a factor of Δ, so that the new value of b is less than or equal to $\sqrt{\Delta}\sqrt{1 - 1/2D}$. After normalizing, the overlap between $K|\phi_A \otimes \phi_B\rangle$ and $|\Omega\rangle$ is at least

$$\frac{1/\sqrt{2D}}{\sqrt{\frac{1}{2D} + \Delta\left(\frac{2D-1}{2D}\right)}} = \frac{1}{\sqrt{1+\Delta(2D-1)}} = \frac{1}{\sqrt{1-\Delta+2\Delta D}} \geq \frac{1}{\sqrt{2-1/(2D)}} \geq \frac{1}{\sqrt{2}} \tag{24}$$

where in the second-to-last inequality we used the fact that $\Delta \leq 1/2D$.

We have thus found a state (namely the normalized $K|\phi_A \otimes \phi_B\rangle$) which has Schmidt rank at most D and has overlap with Ω that is at least a constant $1/\sqrt{2}$ regardless of the value of D. If we continue iteratively applying K, we get states with Schmidt ranks at most D^2, D^3, $\cdots$, that have constant and increasing overlaps with the ground state.

For instance, we can work out the effect of applying K a second time. We start with a state with overlap $a \geq 1/\sqrt{2}$ with the ground state:

$$|\phi\rangle = a|\Omega\rangle + \sqrt{1-a^2}|\Omega^\perp\rangle \tag{25}$$

and then apply K to get

$$K|\phi\rangle = a|\Omega\rangle + c|\Omega^\perp\rangle \tag{26}$$

with $c \leq \sqrt{\Delta}\sqrt{1-a^2}$. Normalizing this state and taking the overlap with $|\Omega\rangle$ gives

$$\frac{a}{\sqrt{a^2+c^2}} \geq \frac{a}{\sqrt{a^2+(1-a^2)\Delta}} \geq \frac{1/\sqrt{2}}{\sqrt{1/2+\Delta/2}} = \frac{1}{\sqrt{1+\Delta}} \geq \frac{1}{\sqrt{1+1/(2D)}} \geq \sqrt{\frac{2}{3}} \tag{27}$$

In the same way, we can show that the normalized state in the direction of $K^m|\phi_A \otimes \phi_B\rangle$ has Schmidt rank at most D^m and overlap with the ground state of at least $\sqrt{2^{m-1}/(2^{m-1}+1)}$. (This can, for instance, be proven by induction.)

We next put lower bounds on the overlap of the ground state with truncations of its Schmidt decomposition. That Schmidt decomposition looks like

$$|\Omega\rangle = \sum_{j=1}^{\infty} c_j |\psi_A^{(j)}\rangle |\psi_B^{(j)}\rangle, \tag{28}$$

where we make no assumptions on the Schmidt rank of the state. The entanglement entropy is given by

$$H(\{c_j^2\}) = -\sum_j c_j^2 \log(c_j^2) \tag{29}$$

and we want to bound this above by an expression which is logarithmic in D. For this purpose we use the Eckart-Young Lemma (aka the Eckart-Young-Mirsky theorem), which states that the best approximation to $|\Omega\rangle$ with rank D is given by truncating the Schmidt decomposition of $|\Omega\rangle$ after the first D terms (assuming the terms are ordered from largest to smallest singular value),

$$|\Omega\rangle = \sum_{i=1}^{\infty} c_i |\psi_A^{(i)}\rangle |\psi_B^{(i)}\rangle \rightarrow \sum_{i=1}^{D} c_i |\psi_A^{(i)}\rangle |\psi_B^{(i)}\rangle \tag{30}$$

This state has overlap $\sum_{j=1}^{D} c_j^2$ with the ground state.

Furthermore, we said above that the normalized $K|\phi_A \otimes \phi_B\rangle$ has Schmidt rank $\leq D$ and overlap with the ground state of at least $1/\sqrt{2}$, which means that $\sum_{j=1}^{D} c_j^2 \geq 1/\sqrt{2}$ since the truncated Schmidt decomposition is the best approximation with rank $\leq D$ and therefore must have overlap at least as large. Similarly, applying K twice tells us that $\sum_{j=1}^{D^2} c_j^2 \geq \sqrt{2/3}$. Continuing to apply K many times, we find

$$\sum_{j=1}^{D^m} c_j^2 \geq \sqrt{\frac{2^{m-1}}{2^{m-1}+1}} \tag{31}$$

Finally, we will show that the entanglement in the ground state is $\mathcal{O}(\log D)$. If the sum of n nonnegative numbers, x_1 through x_n, is fixed at $z \leq 1$, then the entropy $-\sum x_i \log(x_i)$ is maximized when all the x_i are equal, $x_i = z/n$. This gives an entropy of $-n \times z/n \times \log(z/n) = z \log(n/z)$.

Then the maximum entanglement entropy possible in the rank-D Schmidt approximation to $|\Omega\rangle$ is

$$H(D) \leq \frac{1}{\sqrt{2}} \log\left(D\sqrt{2}\right) \leq \frac{1}{\sqrt{2}} \log\left(D\right) \tag{32}$$

and the maximum entanglement entropy possible in the rank-D^2 Schmidt approximation to $|\Omega\rangle$ is

$$\begin{aligned} H(D^2) &\leq \frac{1}{\sqrt{2}} \log\left(D\sqrt{2}\right) + \frac{2-\sqrt{3}}{\sqrt{6}} \log\left(\frac{D^2 - D}{(2-\sqrt{3})/\sqrt{6}}\right) && (33) \\ &\leq \frac{1}{\sqrt{2}} \log\left(D\sqrt{2}\right) + \frac{2-\sqrt{3}}{\sqrt{6}} \log\left(\frac{D^2}{(2-\sqrt{3})/\sqrt{6}}\right) && (34) \\ &\leq \frac{1}{\sqrt{2}} \log\left(D\right) + 2\left(\frac{2-\sqrt{3}}{\sqrt{6}}\right) \log\left(D\right) && (35) \end{aligned}$$

Generalizing this approach, we find that

$$H(D^m) \leq \log(D) \times \left[\frac{1}{\sqrt{2}} + \sum_{j=2}^{m} j \left(\sqrt{\frac{2^{j-1}}{2^{j-1}+1}} - \sqrt{\frac{2^{j-2}}{2^{j-2}+1}}\right)\right] \tag{36}$$

but for large j the coefficient becomes

$$j\left(\sqrt{\frac{2^{j-1}}{2^{j-1}+1}}-\sqrt{\frac{2^{j-2}}{2^{j-2}+1}}\right)=j\left(\sqrt{\frac{1}{1+2^{-(j-1)}}}-\sqrt{\frac{1}{1+2^{-(j-2)}}}\right) \tag{37}$$

$$\approx j\left[\left(1-2^{-j}\right)-\left(1-2^{-(j-1)}\right)\right] \tag{38}$$

$$\approx \frac{j}{2^j} \tag{39}$$

and the sum of $j/2^j$ converges, so that as $m \to \infty$, $H(D^\infty)$ is bounded above by $\log(D)$ times a positive constant. (Note that if the sum over j had diverged, the proof of the area law would have failed, since the maximum possible Schmidt rank of $|\Omega\rangle$ is $d^{N/2}$ where N is the total number of sites in the lattice, and thus the upper limit of the sum would have grown with the system size so that a divergent sum would as well.)

This completes the proof of claim 2, namely that the entanglement entropy in the ground state across any cut is bounded by $\mathcal{O}(\log d)$.

3.3 Finding the AGSP

We have now proven some consequences of the existence of a (D,Δ)-AGSP, most importantly (**Claim 2**) that the entanglement in the ground state across any cut scales as $\mathcal{O}(\log D)$. Then to prove the desired result, that the entanglement scales as $\mathcal{O}(\log^3(d)/\epsilon)$, all we need to do is construct a (D,Δ)-AGSP with an appropriately small value of D, namely $D=\mathcal{O}\left(\exp\left(\frac{\log^3 d}{\epsilon}\right)\right)$.

A naive first guess for such an AGSP looks something like $I-H$ since for the ground state we have $H|\Omega\rangle=0$ and hence this acts as the identity and satisfies the requirement that $K|\Omega\rangle=|\Omega\rangle$. We have to be a bit more careful, though. If we assume that H has n local terms, each satisfying $0\le|H_i|\le 1$, then $0\le|H|\le n$, so to shrink the part of the state perpendicular to $|\Omega\rangle$ we must divide H by n. We thus try

$$K=\left(I-\frac{H}{n}\right) \tag{40}$$

which again clearly satisfies $K|\Omega\rangle=|\Omega\rangle$. This candidate for K is unfortunately not actually very good. The Schmidt rank D can in general be quite large, and a state perpendicular to $|\Omega\rangle$ but which has a nonzero expectation value for only one term of H would have $\Delta=1-1/n$, which is not even remotely small! We clearly need to do better.

As it turns out, we will construct K as something like a polynomial in H, $K=\mathrm{poly}(H)$, and in particular it will be closely related to a Chebyshev polynomial of H.

3.3.1 Truncating the Hamiltonian

Recall that each local term of H satisfies $0\le|H_i|\le 1$, so that $0\le|H|\le n$, where n is the number of terms. We would like a new Hamiltonian with a smaller norm but with the same ground state, $|\Omega\rangle$. We can construct this Hamiltonian in reference to the location of the cut between the two parts of the system, A and B. We label s local terms of the Hamiltonian in the vicinity of the cut by H_1 through H_s, as shown in figure 1. We keep these terms of the Hamiltonian the same. All terms to the left of this we group into H_L', which we then truncate by reducing each nonzero eigenvalue, so that any value larger than 1 is reduced to 1. (In fact, we need to truncate to values slightly different from 1, but we will ignore that subtlety here.) This truncated Hamiltonian we call H_L. We likewise truncate the part of the Hamiltonian to the right of H_1 through H_s, and get H_R. In total the new Hamiltonian is

$$H'=H_L+H_1+H_2+\cdots+H_s+H_R \tag{41}$$

This new Hamiltonian satisfies

$$|H'|\le s+4 \tag{42}$$

It would be tempting now to try $K=I-H'/(s+4)$, but while smaller than $1-1/n$, $\Delta=1-1/(s+4)$ is still too large.

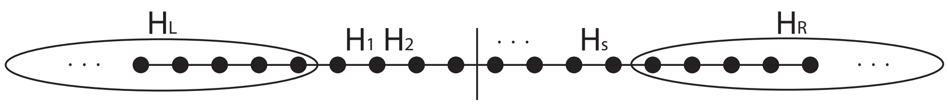

Figure 1: Partitioning the system into regions S, left, and right.

3.3.2 Finding the AGSP as a polynomial in H'

We now want to find the AGSP K as a polynomial of degree l in H', $K = P(H')$. Here are some useful facts and considerations:

- If $|\phi\rangle$ is an eigenvector of H' with eigenvalue λ, then $|\phi\rangle$ is also an eigenvector of $P(H')$ with eigenvalue $P(\lambda)$.
- The eigenvalues of H' are 0 (for the ground state), ϵ (for the first excited state, just above the energy gap; this doesn't depend on system size), $\cdots$, $\leq s+4$ (ie bounded by the max of $|H'|$)
- Since the polynomial will be our AGSP, $P(H')$ should take $|\Omega\rangle$ to itself, which is accomplished if P maps the ground state energy to 1. Thus we must have $P(0) = 1$.
- Likewise, all other states must be reduced by a factor Δ, so $|P(\lambda)| \leq \Delta$ for all $\lambda \in [\epsilon, s+4]$.
- The Schmidt rank D of $P(H')$ cannot be too large.

For P to satisfy the condition that $P(0) = 1$ and $|P(\lambda)| \leq \Delta$ for $\lambda \in [\epsilon, s+4]$, it must look something like what is shown in figure 2. To satisfy the final condition, that D be as small as possible, we want to minimize the degree of the polynomial. For a polynomial of the type shown in figure 2, the degree is minimized by using a Chebyshev polynomial, which can achieve an initial slope of $-1/\epsilon$ with degree only $1/\sqrt{\epsilon}$. (Note that the number of nontruncated terms in the Hamiltonian, s, also comes into play here, as detailed below.)

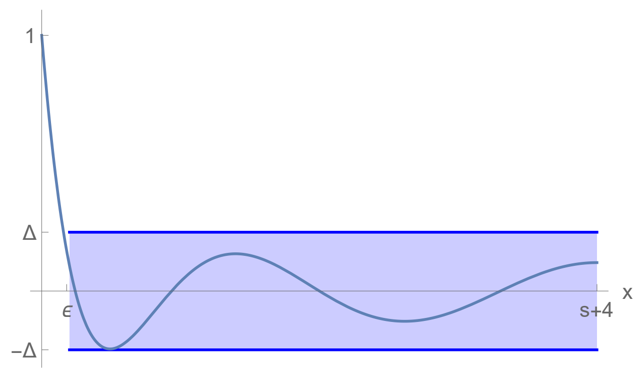

Figure 2: We want a polynomial that takes $0 \mapsto 1$ and all other values in the range $[\epsilon, s+4]$ to a region around 0 of width 2Δ (shaded).

3.3.3 Analyzing this AGSP

To apply the results of claims 1 and 2 above, we need to know the values of D and Δ for this prospective AGSP, and we need to confirm that $D\Delta \leq 1/2$.

In this case, the Chebyshev polynomial (of degree l) of H' gives

$$\Delta = e^{-l\sqrt{\epsilon/s}} \tag{43}$$

and

$$D = d^{(l/s)+s} \begin{pmatrix} l \\ s \end{pmatrix}. \tag{44}$$

To make D scale with as small a power of d as possible, we choose $l = s^2$. Then D becomes $d^{2s}\begin{pmatrix} s^2 \\ s \end{pmatrix}$. Simplifying both this and Δ, we have

$$D = d^{2s}s^s \tag{45}$$

$$\Delta = e^{-s^{3/2}\epsilon^{1/2}} \tag{46}$$

Finally we can choose s in order to guarantee that $D\Delta \leq 1/2$. This works if $s = \mathcal{O}(\log^2 d)$, in which case we get

$$\begin{aligned} &s \log d \ll s^{3/2}\epsilon^{1/2} \\ \Rightarrow &s^{1/2}\epsilon^{1/2} \gg \log d \\ \Rightarrow &s \gg \frac{\log^2 d}{\epsilon} \\ \Rightarrow &D \approx d^{\log^3(d)/\epsilon} \end{aligned}$$

This proves that $P(H')$ is the desired (D, Δ)-AGSP satisfying

$$D \approx \exp\left(\frac{\log^3 d}{\epsilon}\right) \tag{47}$$

3.4 Finishing the proof and some discussion

From our **Claim 2** above, we know that the entanglement across a cut in the ground state, $|\Omega\rangle$, is proportional to $\log(D)$. Since we have just found an AGSP which has $D \approx \exp\left(\frac{\log^3 d}{\epsilon}\right)$, this completes the proof that the entanglement across the cut is order

$$\frac{\log^3 d}{\epsilon} \tag{48}$$

Why exactly does this AGSP work? The key is to look at where the entanglement we see when we make a cut is actually coming from. If it were entirely from entanglement between the two sites across the cut, as in figure 3a, then the maximum entanglement entropy would be $-d \times d^{-1}\log(d^{-1}) = \log d$. Then if applying the Hamiltonian (or $K = P(H')$) creates a large amount of entanglement across the cut, any of it that exceeds $\log d$ must be coming from entanglement between sites that are farther from the cut (figure 3b). Then either

(a) There is not much entanglement across the cut from applying $P(H')$, in which case D is small, or

(b) there is a large amount of entanglement, but then there is already a lot of entanglement on other sites near the cut as well and future applications of $P(H')$ cannot further increase the entanglement very much.

Either way, the total amount of entanglement is limited.

For another perspective on this, we can consider the optimal size s of our nontruncated region for H'. If s is too large, then $|H'|$ is large and hence Δ will not be small enough. But if s is too small, then $P(H')$ creates entanglement only in the region s, and that entanglement can flow out to parts of the system farther from the cut and allow for more entanglement to be added with future actions of $P(H')$. This tradeoff is what gives rise to the factor $d^{s+l/s}$ in our expression for D above.

References

[1] Matthew B Hastings. An area law for one-dimensional quantum systems. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(08):P08024, 2007.

[2] Itai Arad, Alexei Kitaev, Zeph Landau, and Umesh Vazirani. An area law and sub-exponential algorithm for 1d systems. *arXiv preprint arXiv:1301.1162*, 2013.

(a) Entanglement just across the cut

(b) Entanglement moved away from the cut

Figure 3: Entanglement as a result of applying the AGSP