Note: Area Law for One-Dimensional Frustration-Free Systems

Ken Shiozaki

June 1, 2026

Abstract

We follow the proof of the area law for the entanglement entropy of gapped one-dimensional frustration-free systems.

1 Introduction

In a one-dimensional quantum many-body system whose local Hilbert space has finite dimension, if the system is gapped, namely if there is a finite energy gap between the ground state and the first excited state, then the von Neumann entanglement entropy obeys an area law: it is bounded from above by a constant independent of the system size. Let d be the dimension of the local Hilbert space and ϵ the spectral gap. The currently known results are as follows[1].

- A proof using the Lieb-Robinson bound [1][2]:

$$S \leq \exp\left[O(\frac{\log d}{\epsilon})\right]. \tag{1.1}$$

- A proof using the approximate ground-space projection (AGSP) [6]:

$$S \leq O(\frac{\log^3 d}{\epsilon}). \tag{1.2}$$

- For frustration-free systems, the AGSP result of [6] gives

$$S \leq O\left((\frac{\log d}{\epsilon})^3\right). \tag{1.3}$$

Here I take "frustration free" to include the assumption that the local Hamiltonian is k-local. The bound for frustration-free systems is obtained by combining the detectability lemma [2, 3] and the diluting lemma [3, 5], whose proofs are followed in detail in this note. The detectability lemma states that in a frustration-free system there exists an operator A, built as a product of local projections, which brings any state closer to the ground space. The diluting lemma guarantees that A can be redefined so that the increase of the Schmidt rank is suppressed. For general frustrated systems, the AGSP is constructed from the local Hamiltonian. Looking at later developments, AGSPs seem to be useful computational tools with many applications. With the goal of understanding AGSPs, this note follows the proofs of the detectability lemma and the diluting lemma in frustration-free systems. The area law proved in [5] is

$$S \leq O\left((\frac{\log d}{\epsilon})^3 \log^8(\frac{\log d}{\epsilon})\right). \tag{1.4}$$

[1] Check whether ground-state degeneracy is allowed in each result.

[2] I tried to follow the formulas in Hastings's original paper [1], but gave up. Reference [8] gives a detailed account including the infinite-volume setting.

As for the frustration-free condition, in general spatial dimension there is in fact a statement that, once a Hamiltonian is gapped, it can always be deformed into something frustration-free-like.

Proposition D.1 in [7]. Let $H = \sum_j H_j$ be a gapped local Hamiltonian. Here locality means that $||[H_j, O_i]||$ decays faster than any power. Let the energy eigenvalue of the ground state $|\Psi\rangle$ be zero. Let $\Delta > 0$ be the gap. Then one can express $|\Psi\rangle$ as a zero-energy eigenstate of some local term, $\tilde{H}_j |\Psi\rangle = 0$.

See App. B for an explanation. One caveat is that even if H_j is k-local, $\tilde{H}_j$ has tails. The question is how much the results change between the strictly k-local case and the case with tails. Nevertheless, for gapped systems it is natural to expect that many properties of the ground state can be learned from frustration-free systems.

Below, the base of log is 2.

2 Detectability lemma [3]

Following [3], I record the proof of the detectability lemma. The detectability lemma first appeared in [2]. Reference [4] apparently gives a simpler proof, but I have not checked it.

2.1 Setup

Consider the Hilbert space $\mathcal{H} = (\mathbb{C}^d)^{\otimes n}$ on a one-dimensional lattice. Here n is the number of sites. The local term H_i of the Hamiltonian $H = \sum_i H_i$ is assumed to have support only on two neighboring sites. For any finite-range interaction this assumption can be arranged. The lowest eigenvalue of H_i is 0, and the lowest eigenspace of H_i is assumed to be nondegenerate. We have $H \geq 0$, since for any state $|\psi\rangle$,

$$\langle \psi | H | \psi \rangle = \sum_i \langle \psi | H_i | \psi \rangle , \tag{2.1}$$

and each term satisfies $\langle \psi | H_i | \psi \rangle \geq 0$. **The ground state need not be unique, but if there is degeneracy the energy eigenvalue is assumed to be exactly degenerate.** Let $\mathcal{H}_0$ be the ground-state subspace and $\mathcal{H}^\perp$ its orthogonal complement. Let $\epsilon > 0$ be the spectral gap of H, that is, assume

$$|\psi\rangle \in \mathcal{H}^\perp \Rightarrow \langle \psi | H | \psi \rangle \geq \epsilon > 0. \tag{2.2}$$

The Hamiltonian is assumed to be frustration free:

$$|\Omega\rangle \in \mathcal{H}_0 \Leftrightarrow {}^\forall i, H_i |\Omega\rangle = 0. \tag{2.3}$$

Define the orthogonal projection obtained by "flattening" the local Hamiltonian H_i: the lowest eigenvalue of H_i is kept at 0, and all other eigenvalues are changed to 1. Concretely, if $H_i = \sum_{E_n^{(i)}>0} E_n^{(i)} |n^{(i)}\rangle \langle n^{(i)}|$, set

$$Q_i = \sum_{E_n^{(i)}>0} |n^{(i)}\rangle \langle n^{(i)}| . \tag{2.4}$$

This deformation preserves the frustration-free condition:

$$|\Omega\rangle \in \mathcal{H}_0 \Leftrightarrow {}^\forall i, Q_i |\Omega\rangle = 0. \tag{2.5}$$

Introduce $P_i = 1 - Q_i = |0^{(i)}\rangle \langle 0^{(i)}|$. Then

$$|\Omega\rangle \in \mathcal{H}_0 \Leftrightarrow {}^\forall i, P_i |\Omega\rangle = |\Omega\rangle . \tag{2.6}$$

Now define the operator obtained by multiplying projections onto the local ground spaces,

$$A := \Pi_{\text{even}}\Pi_{\text{odd}}, \tag{2.7}$$

$$\Pi_{\text{even}} = P_2 P_4 P_6 \cdots, \tag{2.8}$$

$$\Pi_{\text{odd}} = P_1 P_3 P_5 \cdots. \tag{2.9}$$

Neighboring projections such as P_{2i-1}, P_{2i} and P_{2i}, P_{2i+1} do not commute, while all other pairs commute. Also,

$$|\Omega\rangle \in \mathcal{H}_0 \Rightarrow A|\Omega\rangle = A^\dagger|\Omega\rangle = |\Omega\rangle. \tag{2.10}$$

Moreover,

$$|\psi\rangle \in \mathcal{H}^\perp \Rightarrow A|\psi\rangle \in \mathcal{H}^\perp. \tag{2.11}$$

Indeed, $\langle \Omega | A | \psi \rangle = \langle \Omega | \psi \rangle = 0$, so $A|\psi\rangle$ has no ground-state component. Thus, with respect to the decomposition $\mathcal{H}_0 \oplus \mathcal{H}^\perp$, A has the form

$$A = \begin{pmatrix} 1 & \\ & A|_{\mathcal{H}^\perp} \end{pmatrix}. \tag{2.12}$$

Since A preserves the ground state and reduces excited-state components, one expects A^l for sufficiently large l to be close to the projection onto the ground space. Let

$$H_Q = \sum_i Q_i \tag{2.13}$$

be the flattened Hamiltonian. Assume that this deformation leaves the ground space unchanged and keeps the spectral gap finite[3]. Let ϵ_Q be the spectral gap of H_Q. The detectability lemma says:

Theorem 2.1 (Detectability lemma [3]). *Consider a one-dimensional, frustration-free, 2-local, gapped system. Let $\epsilon_Q > 0$ be the spectral gap of the Hamiltonian $H = \sum_i Q_i$ made of projection operators. Then*

$$||A|_{\mathcal{H}^\perp}|| < \frac{1}{(\epsilon_Q/2+1)^{\frac{1}{3}}}. \tag{2.14}$$

Thus the action of A on $\mathcal{H}^\perp$ is strictly smaller than one when the spectral gap is finite.

Proof. Take a normalized state $|\psi\rangle \in \mathcal{H}^\perp$, and set $|\phi\rangle := A|\psi\rangle$. Since $|\phi\rangle \in \mathcal{H}^\perp$,

$$\langle \phi | H_Q | \phi \rangle \geq \epsilon_Q ||\phi||^2. \tag{2.15}$$

We estimate $\langle \phi | H_Q | \phi \rangle$ from above. First note that $Q_{i\in\text{odd}}\Pi_{\text{odd}} = 0$, so odd i do not contribute. Rewrite A as

$$A = \underbrace{(P_1 P_3 P_2)}_{\Delta_1}\underbrace{(P_5 P_7 P_6)}_{\Delta_2}\cdots\underbrace{(P_4 P_8 \cdots)}_{R} =: \Delta_1 \Delta_2 \cdots \Delta_m R. \tag{2.16}$$

Here $m \sim n/4$, and $R^\dagger = R$. We estimate $\langle \phi | Q_{4i-2} | \phi \rangle$:

$$\langle \phi | Q_{4i-2} | \phi \rangle = ||(1-P_{4i-2})A\psi||^2 \tag{2.17}$$

$$= ||(1-P_{4i-2})\Delta_1\Delta_2\cdots R\psi||^2 \tag{2.18}$$

$$= ||\Delta_1\cdots\Delta_{i-1}(1-P_{4i-2})\Delta_i\cdots R\psi||^2 \tag{2.19}$$

$$\leq ||(1-P_{4i-2})\Delta_i\cdots R\psi||^2. \tag{2.20}$$

[3] This is not shown here.

We used $||Av|| \leq ||A|| ||v||$[4]. Put

$$v_i = \Delta_i \Delta_{i+1} \cdots R\,|\psi\rangle\,, \quad v_{m+1} = R\,|\psi\rangle\,. \tag{2.21}$$

Note that

$$||v_i|| = ||\Delta_i v_{i+1}|| \leq ||\Delta_i||\,||v_{i+1}|| \leq ||v_{i+1}||. \tag{2.22}$$

Then

$$\langle \phi\,|\,Q_{4i-2}\,|\,\phi\rangle \leq ||(1-P_{4i-2})\Delta_i v_{i+1}||^2 = ||(1-P_{4i-2})P_{4i-3}P_{4i-1}P_{4i-2}v_{i+1}||^2, \tag{2.23}$$

with $v_i = \Delta_i v_{i+1}$. Using Eq. (A.1) with $X = P_{4i-3}P_{4i-1}$, $Y = P_{4i-2}$, and $v = v_{i+1}/||v_{i+1}||$, we get

$$||(1-P_{4i-2})\Delta_i v_{i+1}||^2 \leq \left(1 - \frac{||v_i||^2}{||v_{i+1}||^2}\right)\frac{||v_i||^2}{||v_{i+1}||^2}||v_{i+1}||^2 \leq 1 - \frac{||v_i||^2}{||v_{i+1}||^2}. \tag{2.24}$$

Here we used $||v_i|| \leq ||v_{m+1}|| = ||R\,|\psi\rangle\,|| \leq ||R|| = 1$. Set

$$a_i = \frac{||v_i||^2}{||v_{i+1}||^2}, \quad i = 1, \ldots, m, \quad a_{m+1} = ||R\psi||^2, \tag{2.25}$$

$$a_1 \cdots a_{m+1} = ||v_1||^2 = ||\phi||^2. \tag{2.26}$$

Then

$$\langle \phi\,|\,Q_{4i-2}\,|\,\phi\rangle \leq 1 - a_i. \tag{2.27}$$

Therefore

$$\langle \phi| \sum_i Q_{4i-2}|\phi\rangle = \sum_{i=2,6,\cdots}(1-a_i), \quad a_1 \cdots a_{m+1} = ||\phi||^2. \tag{2.28}$$

The maximum under this constraint is obtained at $a_i = ||\phi||^{2/(m+1)}$[5]. Hence

$$\langle \phi| \sum_i Q_{4i-2}|\phi\rangle \leq m(1 - ||\phi||^{\frac{2}{m}}). \tag{2.29}$$

We want an estimate independent of m. Expanding around $||\phi||^2 = 1 + \delta x$ gives

$$m(1-(1+\delta x)^{1/m}) \sim -\delta x + \frac{1}{2}\frac{1}{m}(1-\frac{1}{m})\delta x^2. \tag{2.30}$$

Since $1/m(1-1/m) = \frac{1}{4} - (\frac{1}{2} - \frac{1}{m})^2$, it is enough to bound the function from above by one of the form

$$-\delta x + c\delta x^2, \quad c \geq \frac{1}{8}. \tag{2.31}$$

One can show

$$m(1 - x^{\frac{1}{m}}) \leq \frac{1-x}{\sqrt{x}}, \quad x \in [0,1] \tag{2.32}$$

[6]. Thus

$$\langle \phi| \sum_i Q_{4i-2}|\phi\rangle \leq \frac{1 - ||\phi||^2}{||\phi||}. \tag{2.33}$$

[4] This follows from $||A|| := \max_{v\neq 0} ||Av||/||v|| \geq ||Av||/||v||$.

[5] Since the expression is symmetric, it is natural that the extremum occurs at constant a_i.

[6] Let $f_m(x) = m\sqrt{x}(1 - x^{1/m}) - (1-x)$. Then $f_m''(x) = \frac{(m^2-4)x^{1/m} - m^2}{4mx^{3/2}} \leq 0$ for $x \in [0,1]$ and $m \geq 2$. Thus $f_m'(x)$ is decreasing; together with $f_m'(1) = 0$ this gives $f_m'(x) \geq 0$ on $[0,1]$. Hence f_m is increasing, and $f_m(x) \leq f_m(1) = 0$.

Similarly, for Q_{4i}, for example by rewriting $A = P_2P_1(P_4P_6\cdots)(P_3P_5\cdots)$, one obtains

$$\langle\phi| \sum_{i=4,8,\dots} Q_{4i}|\phi\rangle \le \frac{1-||\phi||^2}{||\phi||} \tag{2.34}$$

[7].

Combining these estimates,

$$\epsilon_Q||\phi||^2 \le \langle\phi|H_Q|\phi\rangle \le 2\frac{1-||\phi||^2}{||\phi||}. \tag{2.35}$$

Therefore

$$\epsilon_Q||\phi||^3 \le 2(1-||\phi||^2) \le 2(1-||\phi||^3), \tag{2.36}$$

and hence

$$||\phi|| \le \frac{1}{(1+\epsilon_Q/2)^{1/3}}. \tag{2.37}$$

□

3 Diluting lemma [5]

Consider a one-dimensional lattice Hamiltonian whose local term H_i is supported on sites $i, i+1$, and assume that it is frustration free. Let d be the dimension of the local Hilbert space. For simplicity, assume that the Hamiltonian is a sum of projection operators:

$$H = \sum_i Q_i = \sum_i (1-P_i), \quad P_i\,|\Omega\rangle = |\Omega\rangle\,. \tag{3.1}$$

The Hamiltonian is assumed to be gapped, with finite spectral gap $\Delta > 0$. We do not impose uniqueness of the ground state. Let $\mathcal{H}_0$ be the ground-state subspace, P the orthogonal projection onto $\mathcal{H}_0$, and $\mathcal{H}^\perp$ the excited-state subspace. Define

$$\Pi_{\text{even}} = \prod_{i\in\text{even}} P_i, \quad \Pi_{\text{odd}} = \prod_{i\in\text{odd}} P_i, \tag{3.2}$$

$$A = \Pi_{\text{even}}\Pi_{\text{odd}}. \tag{3.3}$$

The operator A preserves the direct-sum decomposition $\mathcal{H}_0 \oplus \mathcal{H}^\perp$, namely

$$A = PAP + (1-P)A(1-P), \tag{3.4}$$

and by the detectability lemma

$$||A|_{\mathcal{H}^\perp}||^2 \le \Delta_0 := (1+\epsilon/2)^{-2/3}. \tag{3.5}$$

Fix the Schmidt cut between sites $2i$ and $2i+1$. Write $SR(\phi)$ for the Schmidt rank of a state $|\phi\rangle$. We use the following properties of the Schmidt rank.

Fact.

- $SR(\phi+\psi) \le SR(\phi) + SR(\psi)$.
- If an operator O has support on both sides of the cut and is k-local, then $SR(O\phi) \le d^k SR(\phi)$.
- If the Schmidt ranks at cuts $(i, i+1)$ and $(j, j+1)$ are r_i and r_j, then

$$d^{-|i-j|}r_j \le r_i \le d^{|i-j|}r_j. \tag{3.6}$$

[7] For a finite chain, one should carefully check what happens at the right boundary.

The third property follows because the degrees of freedom between i and j are added to the Schmidt legs. Since the local operators considered here are 2-local, we abbreviate the Schmidt-rank growth factor of a local operator as

$$D_0 := d^2. \tag{3.7}$$

We define an approximate ground-space projection (AGSP).

Definition 3.1 (AGSP, Definition III.1 in [5])**.** For a cut i^*, i^*+1, an operator K is a (D, Δ)-AGSP if it satisfies:

(a) It preserves the ground state: $K|\Omega\rangle = |\Omega\rangle$ for $|\Omega\rangle \in \mathcal{H}_0$.

(b) Shrinking: $||K|_{\mathcal{H}^\perp}||^2 \leq \Delta < 1$.

(c) Entangling: $SR(K|\phi\rangle) = D \cdot SR(\phi)$.

Starting from a state $|\phi\rangle$ with nonzero overlap with the ground state $|\Omega\rangle$, the AGSP K gives the ground-state approximation

$$K|\phi\rangle / ||K|\phi\rangle||. \tag{3.8}$$

There is a tradeoff between how much the nonground-state component is removed, measured by Δ, and how much the Schmidt rank grows, measured by D.

The main result of [5] is:

Theorem 3.2 (Diluting lemma, Lemma III.5 in [5])**.** *In a one-dimensional frustration-free gapped system, let $\epsilon > 0$ be the spectral gap, d the local Hilbert-space dimension, and $X = \log d/\epsilon$. There exists a (D, Δ)-AGSP satisfying*

- $D\Delta < 1/2$.
- $\log D \leq O(1) X^3 \log^8 X$.

Reference [5] proves Theorem 3.2 by explicitly constructing an operator $\hat{A}$ with the following properties.

Lemma 3.3 ([5])**.** *Write $X = \frac{\log d}{\epsilon}$. There exists an operator $\hat{A}$ satisfying:*

(a) For $|\Omega\rangle \in \mathcal{H}_0$, $\hat{A}|\Omega\rangle = |\Omega\rangle$.

(b) For $|\psi\rangle \in \mathcal{H}^\perp$, $\hat{A}|\psi\rangle \in \mathcal{H}^\perp$ and $||\hat{A}|\psi\rangle||^2 \leq \hat{\Delta}$.

(c) For any state $|\phi\rangle$ and any $l > 0$, $SR(\hat{A}^l|\phi\rangle) \leq D_I \hat{D}^l SR(\phi)$.

(d) The following estimates hold[8]:

$$\hat{D}\hat{\Delta} \leq 1/2, \tag{3.9}$$

$$\log \hat{D} \leq \log(\hat{\Delta}^{-1}) \leq O(1) \log^2 X, \tag{3.10}$$

$$\log D_I \leq O(1) X^3 \log^6 X. \tag{3.11}$$

Indeed, if such a $\hat{A}$ exists, then $K = \hat{A}^l$ is a $(D_I \hat{D}^l, \hat{\Delta}^l)$-AGSP. Choose l so that

$$D_I \hat{D}^l \hat{\Delta}^l = D_I (\hat{D}\hat{\Delta})^l < 1/2, \tag{3.12}$$

that is,

$$l > \frac{\log D_I - 1/2}{\log[(\hat{D}\hat{\Delta})^{-1}]}. \tag{3.13}$$

Since

$$\frac{\log D_I - 1/2}{\log[(\hat{D}\hat{\Delta})^{-1}]} \leq \frac{\log D_I}{\log 2} = \log D_I, \tag{3.14}$$

we take

$$l_0 = \lceil \log D_I \rceil. \tag{3.15}$$

Then

$$\begin{aligned}
\log(D_I \hat{D}^{l_0}) &\leq O(1) X^3 \log^6 X + O(1) l_0 \log^2 X &(3.16)\\
&= O(1) X^3 \log^6 X + O(1) \log D_I \log^2 X &(3.17)\\
&\leq O(1) X^3 \log^6 X + O(1) X^3 \log^8 X &(3.18)\\
&= O(1) X^3 \log^8 X, &(3.19)
\end{aligned}$$

which proves the theorem.

3.1 Proof of Lemma 3.3

For simplicity, write the set of m even sites centered at the cut as $I_m = \{1, \ldots, m\}$, and write the product of projections on this set as

$$\Pi_m = \prod_{i=1}^{m} P_i. \tag{3.20}$$

Let Π_{rest} be the product of even-site projections outside I_m. Then

$$A = \Pi_m \Pi_{\text{rest}} \Pi_{\text{odd}}. \tag{3.21}$$

Applying A to an arbitrary state with nonzero overlap with the ground state brings the state closer to the true ground state, but each application of an operator near the cut increases the Schmidt rank. The goal is to introduce an approximation $\hat{\Pi}_m$ to Π_m with a smaller number of effective applications, so that the increase of the Schmidt rank is suppressed while the convergence to the ground state is not much slower than that of Π_m.

We approximate Π_m by an operator obtained by applying a certain Hamiltonian several times. Let

$$N := \sum_{i=1}^{m} (1 - P_i). \tag{3.22}$$

The eigenvalues of N are $\{0, 1, \ldots, m\}$. The projections onto the sectors $N = 0$ and $N > 0$ are respectively

$$\Pi_m, \quad 1 - \Pi_m. \tag{3.23}$$

Fact. There exists a polynomial function C satisfying:

- Its degree is $\sqrt{m}$, and we may assume m is a square.
- $C(0) = 1$.
- For $x = 1, 2, \ldots, m$, $|C(x)| \leq 1/3$.

See [5] for an explicit construction[9]. Define

$$\hat{\Pi}_m := [C(N)]^q, \tag{3.24}$$

where q will be chosen later. Thus $\hat{\Pi}_m$ is a polynomial function of N of degree

$$j := q\sqrt{m}. \tag{3.25}$$

By construction, $\hat{A}$ preserves the decomposition $\mathcal{H}_0 \oplus \mathcal{H}^\perp$. Indeed, since $N|_{\mathcal{H}_0} = 0$, we have $C(N)|\Omega\rangle = |\Omega\rangle$ for $|\Omega\rangle \in \mathcal{H}_0$. For $|n\rangle \in \mathcal{H}^\perp$, the off-diagonal matrix element vanishes:

$$\left\langle n \middle| \hat{A} \middle| \Omega \right\rangle = 0. \tag{3.26}$$

Hence $\hat{A}$ is block diagonal,

$$\hat{A} = \begin{pmatrix} 1 & \\ & \hat{A}|_{\mathcal{H}^\perp} \end{pmatrix}. \tag{3.27}$$

Norm estimate. We estimate $||\hat{A}|_{\mathcal{H}^\perp}||$. For any excited state $|\Omega^\perp\rangle \in \mathcal{H}^\perp$,

$$\begin{aligned} \hat{A}|\Omega^\perp\rangle &= \hat{\Pi}_m \Pi_{\text{rest}} \Pi_{\text{odd}} |\Omega^\perp\rangle & (3.28) \\ &= \hat{\Pi}_m \Pi_m \Pi_{\text{rest}} \Pi_{\text{odd}} |\Omega^\perp\rangle + \hat{\Pi}_m (1 - \Pi_m) \Pi_{\text{rest}} \Pi_{\text{odd}} |\Omega^\perp\rangle & (3.29) \\ &=: \hat{\Pi}_m |\psi_0\rangle + \hat{\Pi}_m |\psi_1\rangle. & (3.30) \end{aligned}$$

Here $|\psi_0\rangle$ and $|\psi_1\rangle$ belong to the $N = 0$ and $N > 0$ sectors, respectively. For the first term, $\hat{\Pi}_m |\psi_0\rangle = |\psi_0\rangle$, and therefore

$$||\hat{\Pi}_m |\psi_0\rangle ||^2 = ||\psi_0||^2 = ||A|\Omega^\perp\rangle||^2 \leq \Delta_0 ||\Omega^\perp||^2. \tag{3.31}$$

For the second term, the definition of $\hat{\Pi}_m$ bounds the norm by $(1/9)^q$. Indeed, decomposing into eigenspaces of $N = 1, \ldots, m$ as

$$|\psi_1\rangle = \sum_{n=1}^{m} \alpha_n |n\rangle, \tag{3.32}$$

we have

$$\hat{\Pi}_m |\psi_1\rangle = \sum_{n=1}^{m} \alpha_n [C(n)]^q |n\rangle, \tag{3.33}$$

[9] It uses Chebyshev polynomials of the first kind. If the condition is replaced by $|C(x)| < 1/3$ for $1 \leq x \leq m$, then, surprisingly, according to [5], "this construction is optimal" in the sense that any polynomial satisfying the two properties above must have degree at least $\sqrt{m}$. Their Ref. 23 is R. Paturi, in Proceedings of the twenty-fourth annual ACM symposium on Theory of computing, STOC '92 (ACM, New York, NY, USA, 1992), pp. 468–474.

so

$$||\hat{\Pi}_m |\psi_1\rangle ||^2 \le \sum_{n=1}^{m} |\alpha_n|^2 |C(n)|^{2q} ||n||^2 \le (\frac{1}{9})^q \sum_{n=1}^{m} |\alpha_n|^2 ||n||^2 = (\frac{1}{9})^q ||\psi_1||^2. \tag{3.34}$$

Furthermore,

$$||\psi_1|| = ||(1 - \Pi_m)\Pi_{\text{rest}}\Pi_{\text{odd}} |\Omega^\perp\rangle || \le ||\Omega^\perp||. \tag{3.35}$$

Thus

$$||\hat{A} |\Omega^\perp\rangle ||^2 \le \hat{\Delta} ||\Omega^\perp||^2, \quad \hat{\Delta} = \Delta_0 + (1/9)^q. \tag{3.36}$$

Note that $(1/9)^q$ is independent of m and d. The parameter m will be chosen from the tradeoff with Schmidt-rank growth.

Schmidt-rank estimate. We estimate how much the Schmidt rank grows when $\hat{A}$ is applied l times. For notational simplicity, collect the operators that do not increase the Schmidt rank into

$$\Pi' = \Pi_{\text{rest}}\Pi_{\text{odd}}. \tag{3.37}$$

Then

$$\hat{A} = \hat{\Pi}_m \Pi'. \tag{3.38}$$

Since $\hat{\Pi}_m$ is a polynomial in N of degree j,

$$\hat{\Pi}_m = \sum_{i=0}^{j} c_i N^i. \tag{3.39}$$

Therefore $\hat{A}^l = (\hat{\Pi}_m \Pi')^l$ is a sum of terms of the form

$$(\sum_{i=0}^{j} c_i N^i)\Pi' = \sum_{i_1,\dots,i_l \in \{0,\dots,j\}^{\times l}} c_{i_1} \cdots c_{i_l} N^{i_l}\Pi' \cdots N^{i_2}\Pi' N^{i_1}\Pi'. \tag{3.40}$$

Only N can increase the Schmidt rank. Thus the term

$$(N^j \Pi')^l \tag{3.41}$$

causes the largest Schmidt-rank growth. Since there are $(j+1)^l$ terms in total, linearity of the Schmidt-rank upper bound gives an upper bound by multiplying the Schmidt-rank growth of $(N^j\Pi')^l$ by

$$(j+1)^l. \tag{3.42}$$

Let us first make a naive estimate. Only the projection $Q_{i^*} = 1 - P_{i^*}$ crossing the cut can increase the Schmidt rank. Since the terms of $N = \sum_i Q_i$ commute and each Q_i is a projection, one can write $N^j = \hat{O}_1 + \hat{O}_2 Q_{i^*}$, where $\hat{O}_1, \hat{O}_2$ have no support on sites i^*, i^*+1. Then

$$SR\left(N^j \Pi' |\phi\rangle\right) = SR\left((\hat{O}_1 + \hat{O}_2 Q_{i^*}) |\phi\rangle\right) \le (1 + d^2) SR(\phi). \tag{3.43}$$

Repeating l times and including the factor $(j+1)^l$ yields

$$SR(\hat{A}^l |\phi\rangle) \le (j+1)^l (1+d^2)^l SR(\phi). \tag{3.44}$$

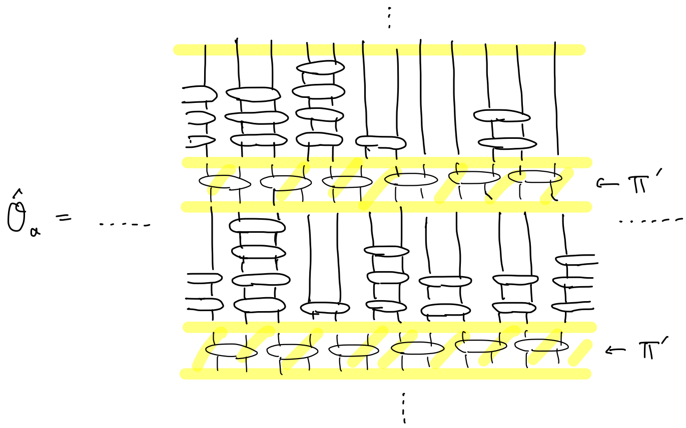

Figure 1: It is enough to find the sparsest Q_i.

This estimate is too large; it cannot be improved by tuning m and q.

A tighter estimate follows from Eq. (3.6): the Schmidt-rank growth may be estimated using any site $i \in I_m$. For any $i, k \in I_m$,

$$r_i \leq d^{|i-k|} r_k \leq D_0^m r_k. \tag{3.45}$$

Therefore an upper bound is obtained by choosing the smallest r_k and multiplying by at most the constant factor D_0^m.

Thus, when

$$(N^j \Pi')^l = \sum_a \hat{O}_a \tag{3.46}$$

is written as a sum of operators $\hat{O}_a$ supported on I_m, it suffices to estimate, for each $\hat{O}_a$, the site $i \in I_m$ where the degree of the projection Q_i is the smallest. See Fig. 1. We recursively divide I_m into two pieces and estimate such a minimal i.

First round. Separate $N = \sum_{i \in I_m} Q_i$ into the left and right halves of the cut:

$$N = N_L + N_R, \quad N_L = \sum_{i=1}^{m/2} Q_i, \quad N_R = \sum_{i=m/2+1}^{m} Q_i. \tag{3.47}$$

Then the expansion

$$(N^j \Pi')^l = \sum_{n_l=0}^{j} \binom{j}{n_l} N_L^{j-n_l} N_R^{n_l} \Pi' \cdots \sum_{n_2=0}^{j} \binom{j}{n_2} N_L^{j-n_2} N_R^{n_2} \Pi' \sum_{n_1=0}^{j} \binom{j}{n_1} N_L^{j-n_1} N_R^{n_1} \Pi' \tag{3.48}$$

has $(j+1)^l$ terms. In each term, the degrees of N_L and N_R are respectively $jl - \sum_{i=1}^{l} n_i$ and $\sum_{i=1}^{l} n_i$. The smaller of these two degrees is bounded by

$$\min_{\{n_i\}_i} \left(jl - \sum_{i=1}^{l} n_i, \sum_{i=1}^{l} n_i \right) \leq jl - \frac{jl}{2} = \frac{jl}{2}. \tag{3.49}$$

Thus $(N^j \Pi')^l$ is a sum of $(j+1)^l$ terms, each involving an operator made from either N_L or N_R with degree at most $jl/2$. Without loss of generality we call the lower-degree one N_R.

Second round. Extract a term in N_R of degree at most $jl/2$, ignoring coefficients:

$$N_R^{n_l}(*) \cdots N_R^{n_2}(*) N_R^{n_1}(*), \quad \sum_{i=1}^{l} n_i \le \frac{jl}{2}, \tag{3.50}$$

where $(*)$ denotes a product of operators other than N_R. Split

$$N_R = N_{RL} + N_{RR}. \tag{3.51}$$

Ignoring coefficients, the expansion gives

$$\sum_{m_l=0}^{n_l} N_{RL}^{n_l - m_l} N_{RR}^{m_l}(*) \cdots \sum_{m_2=0}^{n_2} N_{RL}^{n_2 - m_2} N_{RR}^{m_2}(*) \sum_{m_1=0}^{n_1} N_{RL}^{n_1 - m_1} N_{RR}^{m_1}(*). \tag{3.52}$$

The smaller degree among N_{RL} and N_{RR} is bounded by

$$\min_{\{m_i\}_i} \left(\sum_{i=1}^{l} (n_i - m_i), \ \sum_{i=1}^{l} m_i \right) \le \frac{1}{2} \sum_{i=1}^{l} n_i \le \frac{jl}{4}. \tag{3.53}$$

The number of terms is

$$\prod_{i=1}^{l} \sum_{m_i=0}^{n_i} = \prod_{i=1}^{l} (n_i + 1). \tag{3.54}$$

Under the condition $\sum_i n_i \le jl/2$, this is maximized when $n_i \equiv j/2$, and hence

$$\prod_{i=1}^{l} (n_i + 1) \le \prod_{i=1}^{l} (j/2 + 1) = (j/2 + 1)^l. \tag{3.55}$$

Therefore we obtain at most $(j/2 + 1)^l$ terms, each involving either N_{RL} or N_{RR} with degree at most $jl/4$.

Repeating this for $\log j$ rounds, one arrives at operators supported on an interval $I_{m/2^{\log j}} = I_{m/j} (= I_{\sqrt{m/q}})$, with degree at most

$$\frac{jl}{2^{\log j}} = \frac{jl}{j} = l, \tag{3.56}$$

and with the number of terms bounded by

$$\Bigg[\underbrace{(j+1)(j/2+1) \cdots (1+1)}_{\log j} \Bigg]^l. \tag{3.57}$$

Now estimate the remaining $\log(m/j)$ rounds. We split a term of the form

$$\tilde{N}^{n_l}(*) \cdots \tilde{N}^{n_1}(*), \quad \sum_{i=1}^{l} n_i \le l. \tag{3.58}$$

Writing $\tilde{N} = \tilde{N}_L + \tilde{N}_R$, the number of terms $\prod_{i=1}^{l} (n_i + 1)$ is maximized when all n_i are equal, namely $n_i \equiv 1$:

$$\prod_{i=1}^{l} (n_i + 1) \le 2^l. \tag{3.59}$$

The degree is bounded by $l/2$. In the next round, the number of terms is estimated under the condition

$$\tilde{N}^{n_l}(*)\cdots\tilde{N}^{n_1}(*), \quad \sum_{i=1}^{l} n_i \le l/2. \tag{3.60}$$

This time the product is maximized by taking as many n_i as possible equal to 1: for $l/2$ values of i, $n_i = 1$, and the rest are 0. Thus

$$\prod_{i=1}^{l}(n_i+1) \le 2^{l/2}. \tag{3.61}$$

The degree is at most $l/4$. Continuing similarly, the total number of terms after $\log(m/j)$ rounds is

$$2^l 2^{l/2} 2^{l/4}\cdots = 2^{l(1+1/2+1/4+\cdots)} \le 2^{2l} = 4^l. \tag{3.62}$$

The final degree is bounded by

$$\frac{l}{2^{\log(m/j)}} = \frac{l}{m/j} = \frac{lj}{m}. \tag{3.63}$$

This produces a Schmidt-rank growth factor $D_0^{lj/m}$.

Finally, multiplying by the constant factor from Eq. (3.45) and by Eq. (3.42), we get the Schmidt-rank growth factor

$$SR(\hat{A}^l\,|\phi\rangle) \le D_0^m\,[(j+1)(j/2+1)\cdots(1+1)]^l\,4^l D_0^{lj/m}(j+1)^l SR(\phi). \tag{3.64}$$

Using

$$4(j+1)^2(j/2+1)(j/4+1)\cdots(1+1) \le 20 j^{3/2} 2^{\frac{1}{2}\log^2 j}, \quad j\ge 2, \tag{3.65}$$

[10] we obtain

$$SR(\hat{A}^l\,|\phi\rangle) \le D_I \hat{D}^l\, SR(\phi), \tag{3.70}$$

$$D_I = D_0^m, \quad \hat{D} = 20 j^{3/2} 2^{\frac{1}{2}\log^2 j} D_0^{j/m}. \tag{3.71}$$

Coarse graining. The gap and the Schmidt-rank growth are related by

$$\hat{\Delta}\cdot\hat{D} = (\Delta_0 + (1/9)^q) 20 j^{3/2} 2^{\frac{1}{2}\log^2 j} D_0^{j/m}. \tag{3.72}$$

[10]

$$\prod_{k=0}^{\log j}(j/2^k+1) = j^{\frac{1}{2}(1+\log j)}\left(-\frac{1}{j};2\right)_{\log(2j)}. \tag{3.66}$$

Here

$$(a;q)_n = (a;q)_\infty/(aq^n;q)_\infty, \quad (a;q)_\infty = \prod_{k=0}^{\infty}(1-aq^k) \tag{3.67}$$

is the Pochhammer symbol. A Mathematica plot suggests

$$\frac{j+1}{j}\left(-\frac{1}{j};2\right)_{\log(2j)} \le 5. \tag{3.68}$$

Assuming this,

$$4(j+1)^2(j/2+1)(j/4+1)\cdots(1+1) \le 20j\, j^{\frac{1}{2}(1+\log j)} = 20 j^{3/2} 2^{\frac{1}{2}\log^2 j}. \tag{3.69}$$

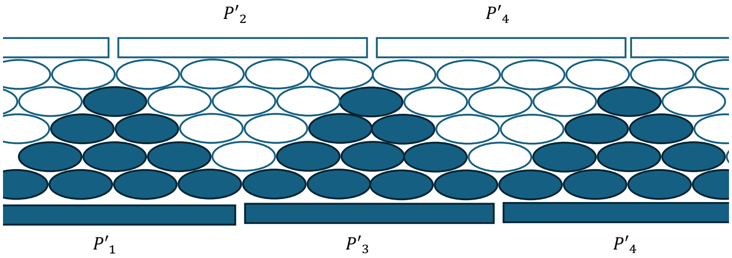

Figure 2: The figure is the same as in [5].

The desired condition $\hat{\Delta}\hat{D} < 1/2$ is not generally satisfied. We therefore perform a coarse graining in which k sites are grouped into one site. After grouping k sites into one site, regard the Hamiltonian as 2-local again, define projections, and introduce

$$A' = \Pi'_{\text{odd}}\Pi'_{\text{even}}, \tag{3.73}$$
$$\Pi'_{\text{odd}} = P'_1 P'_3 \ldots, \quad \Pi'_{\text{even}} = P'_2 P'_4 \ldots. \tag{3.74}$$

For example,

$$Q_1 + Q_2 + \cdots + Q_{2k-1} \xrightarrow{\text{flattening}} Q'_1 = 1 - P'_1, \tag{3.75}$$
$$Q_{k+1} + Q_{k+2} + \cdots + Q_{3k-1} \xrightarrow{\text{flattening}} Q'_2 = 1 - P'_2, \tag{3.76}$$
$$Q_{k(i-1)+1} + Q_{k(i-1)+2} + \cdots + Q_{ki+k-1} \xrightarrow{\text{flattening}} Q'_i = 1 - P'_i. \tag{3.77}$$

Let us assume that this flattening operation does not change the spectral gap ϵ very much[11]. Then

$$\underbrace{(P_k)(P_{k-1}P_{k+1})\cdots(P_3P_5\cdots P_{2k-3})(P_2P_4\cdots P_{2k-2})(P_1P_3\cdots P_{2k-1})}_{\text{``pyramid'' or ``light cone''}} P'_1 = P'_1. \tag{3.78}$$

The same holds for the other P'_i. This is because, on the subspace $P'_1 = 1$, the flattening assumption implies $P_1 = P_2 = \cdots = P_{2k-1} = 1$. Taking the Hermitian conjugate also gives

$$P'_1(P_1P_3\cdots P_{2k-1})(P_2P_4\cdots P_{2k-2})(P_3P_5\cdots P_{2k-3})\cdots(P_{k-1}P_{k+1})(P_k) = P'_1. \tag{3.79}$$

For example, when $k = 4$,

$$\Pi'_{\text{odd}}\Pi'_{\text{even}} = \Pi'_{\text{odd}}\Pi_{\text{even}}\Pi_{\text{odd}}\Pi_{\text{even}}\Pi_{\text{odd}}\Pi_{\text{even}}\Pi'_{\text{even}} \tag{3.80}$$
$$= \Pi'_{\text{odd}}\Pi_{\text{even}}\Pi_{\text{odd}}\Pi_{\text{odd}}\Pi_{\text{even}}\Pi_{\text{even}}\Pi_{\text{odd}}\Pi_{\text{odd}}\Pi_{\text{even}}\Pi'_{\text{even}} \tag{3.81}$$
$$= \Pi'_{\text{odd}}(\Pi_{\text{even}}\Pi_{\text{odd}}\Pi_{\text{odd}}\Pi_{\text{even}})^2\Pi'_{\text{even}} \tag{3.82}$$
$$= \Pi'_{\text{odd}}(AA^\dagger)^2\Pi'_{\text{even}}. \tag{3.83}$$

See Fig. 2. For general even k,

$$\Pi'_{\text{odd}}\Pi'_{\text{even}} = \Pi'_{\text{odd}}(AA^\dagger)^{k/2}\Pi'_{\text{even}}. \tag{3.84}$$

Therefore the detectability lemma gives

$$||A'|_{\mathcal{H}^\perp}||^2 \leq ||AA^\dagger|_{\mathcal{H}^\perp}||^{k/2} = \Delta_0^k, \quad \Rightarrow \quad \Delta'_0 = \Delta_0^k. \tag{3.85}$$

Thus coarse graining by k sites raises the convergence of A to the kth power. On the other hand, since the support of a 2-local projection changes from two sites to $2k$ sites, the Schmidt-rank growth changes as

$$D_0 = d^2 \to D'_0 = d^{2k} = D_0^k. \tag{3.86}$$

The change under k-site coarse graining is summarized as

[11] This is not estimated here. One should estimate how much the original spectral gap changes under the flattening operation.

$$D_0' = D_0^k, \quad \Delta_0' = \Delta_0^k. \tag{3.87}$$

Proof of the claim. Starting from the k-site coarse grained system and constructing $\hat{A}$, we have

$$D_I = D_0^{km}, \tag{3.88}$$

$$\hat{D} = 20j^{3/2}2^{\frac{1}{2}\log^2 j} D_0^{kj/m}, \quad j = q\sqrt{m}, \tag{3.89}$$

$$\hat{\Delta} = \Delta_0^k + (1/9)^q, \quad \Delta_0 = (1+\epsilon/2)^{-2/3}. \tag{3.90}$$

There are three free parameters:

- the number k of sites to coarse grain;
- the length m of the interval I_m introduced in the construction of $\hat{A}$;
- the power q in $(\hat{\Pi}_m)^q$.

Following [5], first take k large enough that $\hat{\Delta} \leq (1/8)^q$. This requires

$$\Delta_0^k \leq (1/8)^q - (1/9)^q, \tag{3.91}$$

or

$$k \geq \frac{-\log((1/8)^q - (1/9)^q)}{\frac{2}{3}\log(1+\epsilon/2)} = \frac{3q - \log(1-(8/9)^q)}{\frac{2}{3}\log(1+\epsilon/2)}. \tag{3.92}$$

Assume $q \geq q_0$ and $\epsilon \leq \epsilon_0$. Since $\log(1+\epsilon/2)/\epsilon > 0$ is monotonically decreasing,

$$\frac{3q - \log(1-(8/9)^q)}{\frac{2}{3}\log(1+\epsilon/2)} \leq \frac{3q - \log(1-(8/9)^{q_0})}{\frac{2}{3}\epsilon\log(1+\epsilon_0/2)/\epsilon_0} \tag{3.93}$$

$$= \frac{q}{\epsilon}\frac{3 - \log(1-(8/9)^{q_0})/q}{\frac{2}{3\epsilon_0}\log(1+\epsilon_0/2)} \leq \frac{q}{\epsilon}\frac{3 - \log(1-(8/9)^{q_0})/q_0}{\frac{2}{3\epsilon_0}\log(1+\epsilon_0/2)}. \tag{3.94}$$

For example, for $q_0 = 4$ and $\epsilon_0 = 10$,

$$\frac{3 - \log(1-(8/9)^{q_0})/q_0}{\frac{2}{3\epsilon_0}\log(1+\epsilon_0/2)} < 20. \tag{3.95}$$

Thus take

$$k := \frac{20q}{\epsilon}. \tag{3.96}$$

Then

$$\hat{\Delta} \leq (1/8)^q + (1/9)^q < 2(1/8)^q. \tag{3.97}$$

A sufficient condition for $\hat{D}\hat{\Delta} < 1/2$ is

$$20j^{3/2}2^{\frac{1}{2}\log^2 j} D_0^{\frac{1}{\epsilon}20qj/m} 2(1/8)^q \leq \frac{1}{2}, \tag{3.98}$$

namely

$$\frac{1}{2}\log j + \frac{1}{2}\log^2 j + 40Xqj/m - 3q < -\log(80), \tag{3.99}$$

where $D_0 = d^2$ and $X = \log d/\epsilon$ have been used. We now choose q and m. Since $m = (j/q)^2$, the sufficient condition is

$$\frac{1}{2}\log j + \frac{1}{2}\log^2 j + 40Xq^3/j - 3q < -\log(80). \tag{3.100}$$

Set

$$40Xq^2/j = 2 \quad \Leftrightarrow \quad j = 20Xq^2. \tag{3.101}$$

Then

$$\frac{1}{2}\log j + \frac{1}{2}\log^2 j - \underbrace{\sqrt{\frac{j}{20X}}}_{q} < -\log(80). \tag{3.102}$$

This is satisfied for sufficiently large j. Thus choices of k, m, q satisfying $\hat{D}\hat{\Delta} \leq 1/2$ exist. Estimating the order of such q gives

$$q \sim \log^2 j = \log^2(20Xq^2) \sim \log^2 X. \tag{3.103}$$

Therefore

$$\log(\hat{\Delta}^{-1}) = 3q - 1 \sim \log^2 X. \tag{3.104}$$

Also,

$$m = (20qX)^2 \sim X^2 \log^4 X, \tag{3.105}$$

and hence

$$\log D_I = \log D_0^{mk} = 2mk\log d = 2m\frac{20q}{\epsilon}\log d \sim (X^2\log^4 X)(\log^2 X)X = X^3\log^6 X. \tag{3.106}$$

This proves the lemma. □

4 Area law

> In this section, the ground state $|\Omega\rangle$ is assumed to be unique.

Let the Schmidt decomposition of the ground state be

$$|\Omega\rangle = \sum_{i\geq 1} \lambda_i |L_i\rangle |R_i\rangle . \tag{4.1}$$

The entanglement entropy is

$$S = -\sum_{i\geq 1} \lambda_i^2 \log \lambda_i^2. \tag{4.2}$$

When approximating $|\Omega\rangle$ by a state of Schmidt rank C, we need an upper bound on the tail

$$-\sum_{i>C} \lambda_i^2 \log \lambda_i^2. \tag{4.3}$$

Lemma 4.1 (Lemma III.3 in [5]). *Assume the existence of a product state $|\phi\rangle = |L\rangle \otimes |R\rangle$ with finite overlap $\mu = \langle\Omega|\phi\rangle \neq 0$ with the ground state $|\Omega\rangle$, and a (D, Δ)-AGSP K. Then*

$$S \leq O(1)\frac{\log|\mu|^{-1}}{\log\Delta^{-1}}\log D. \tag{4.4}$$

Proof. Introduce

$$|v_l\rangle := \frac{K^l|\phi\rangle}{||K^l|\phi\rangle||} = \frac{\mu}{||K^l|\phi\rangle||}|\Omega\rangle + \frac{\sqrt{1-|\mu|^2}}{||K^l|\phi\rangle||}K^l|\Omega^\perp\rangle. \tag{4.5}$$

Then

$$|\langle\Omega|v_l\rangle| = \frac{|\mu|}{||K^l|\phi\rangle||} = \frac{|\mu|}{\sqrt{|\mu|^2+(1-|\mu|^2)||K^l|\Omega^\perp\rangle||^2}} \geq \frac{|\mu|}{\sqrt{|\mu|^2+(1-|\mu|^2)\Delta^l}}. \tag{4.6}$$

The Schmidt rank satisfies

$$SR(v_l) \leq D^l. \tag{4.7}$$

Thus $|v_l\rangle$ has Schmidt rank at most D^l and approximates the ground state with weight at least

$$\frac{|\mu|}{\sqrt{|\mu|^2+(1-|\mu|^2)\Delta^l}}. \tag{4.8}$$

By the Eckart-Young theorem, the Schmidt decomposition gives the best rank-D^l approximation. Equivalently, the normalized state having the largest absolute overlap with $|\Omega\rangle$ is the normalization of $\sum_{i\leq D^l}\lambda_i|L_i\rangle|R_i\rangle$. Therefore

$$\sum_{i\leq D^l}\lambda_i^2 \geq |\langle\Omega|v_l\rangle|^2 \geq \frac{|\mu|^2}{|\mu|^2+(1-|\mu|^2)\Delta^l}. \tag{4.9}$$

Hence the tail is bounded by

$$\sum_{i>D^l}\lambda_i^2 = 1-\sum_{i\leq D^l}\lambda_i^2 \tag{4.10}$$

$$\leq 1-\frac{|\mu|^2}{|\mu|^2+(1-|\mu|^2)\Delta^l} \tag{4.11}$$

$$= \frac{(1-|\mu|^2)\Delta^l}{|\mu|^2+(1-|\mu|^2)\Delta^l} \leq \frac{1}{|\mu|^2}\Delta^l =: p_l. \tag{4.12}$$

It is enough to take this as a constant bound. For instance, choose l so that $p_{l_0} < 1$, namely

$$l > \frac{\log|\mu|^2}{\log\Delta}. \tag{4.13}$$

Set

$$l_0 = \left\lceil\frac{\log|\mu|^2}{\log\Delta}\right\rceil. \tag{4.14}$$

Here $\lceil x\rceil$ denotes the smallest integer not less than x. Equivalently, for $l > l_0$, $p_l \leq \Delta^{l-l_0}$.

Now the entropy contribution from the interval $[D^{2l}+1, D^{2(l+1)}]$ is maximized when the distribution $\{\lambda_i^2\}$ is uniform on that interval. The total weight in this interval is bounded by

$$\sum_{j=D^{2l}+1}^{D^{2(l+1)}}\lambda_j^2 \leq \sum_{j>D^{2l}}\lambda_j^2 \leq p_{2l} < p_l. \tag{4.15}$$

Therefore

$$-\sum_{j=D^{2l}+1}^{D^{2(l+1)}} \lambda_j^2 \log \lambda_j^2 \leq -p_l \log \frac{p_l}{D^{2(l+1)} - D^{2l}} \tag{4.16}$$

$$= p_l \log \frac{D^{2(l+1)} - D^{2l}}{p_l} \leq p_l \log \frac{D^{2(l+1)}}{p_l}. \tag{4.17}$$

The function $x \log(D^{2(l+1)}/x)$ is increasing for $0 < x < 1$. Thus, for $l > l_0$ and hence $p_l < \Delta^{l-l_0}$,

$$-\sum_{j=D^{2l}+1}^{D^{2(l+1)}} \lambda_j^2 \log \lambda_j^2 \leq \Delta^{l-l_0} \log \frac{D^{2(l+1)}}{\Delta^{l-l_0}} = \Delta^{l-l_0}(l-l_0) \log \frac{D^{2(l+1)/(l-l_0)}}{\Delta}. \tag{4.18}$$

Then

$$\sum_{l \geq 2l_0+1} \Delta^{l-l_0}(l-l_0) \log \frac{D^{2(l+1)/(l-l_0)}}{\Delta} = \sum_{l \geq l_0+1} l\Delta^l \log \frac{D^{2(1+\frac{l_0+1}{l})}}{\Delta} \tag{4.19}$$

$$\leq \sum_{l \geq l_0+1} l\Delta^l \log \frac{D^4}{\Delta} \tag{4.20}$$

$$= \frac{\Delta^{1+l_0}(1+l_0-l_0\Delta)}{(1-\Delta)^2} \log \frac{D^4}{\Delta} \tag{4.21}$$

$$\leq \frac{\Delta}{(1-\Delta)^2} \log \frac{D^4}{\Delta}. \tag{4.22}$$

This gives an upper bound on the contribution of Schmidt eigenvalues with $i > D^{2(2l_0+1)}$.

The contribution from Schmidt eigenvalues with $i \leq D^{2(2l_0+1)}$ is maximized when they are all equal, and is bounded by

$$-\sum_{i \leq D^{2(2l_0+1)}} \lambda_i^2 \log \lambda_i^2 \leq \log D^{2(2l_0+1)} = 2(2l_0+1) \log D = O(1) \frac{\log |\mu|^2}{\log \Delta} \log D. \tag{4.23}$$

Thus

$$S \leq O(1) \frac{\log |\mu|^2}{\log \Delta} \log D + \frac{\Delta}{(1-\Delta)^2} \log \frac{D^4}{\Delta}. \tag{4.24}$$

We simplify this estimate. For any $k \in \mathbb{Z}_{\geq 1}$, K^k is a (D^k, Δ^k)-AGSP. Choose

$$k = \left\lceil \frac{1}{\log(\Delta^{-1})} \right\rceil. \tag{4.25}$$

Then

$$\left\lceil \frac{1}{\log(\Delta^{-1})} \right\rceil \leq k \leq \left\lceil \frac{1}{\log(\Delta^{-1})} \right\rceil + 2, \tag{4.26}$$

so $\log \Delta^k \leq -1$ and $\log \Delta^{k-2} \geq -1$. It follows that

$$2^{-1-\frac{2}{k-2}} \leq \Delta^k \leq 2^{-1}. \tag{4.27}$$

Thus Δ^k and $\log \Delta^{-k}$ are $O(1)$. Replacing D, Δ by D^k, Δ^k gives

$$S \leq O(1) \frac{\log |\mu|^2}{\log \Delta^k} \log D^k + \frac{\Delta^k}{(1-\Delta^k)^2} \log \frac{D^{4k}}{\Delta^k} \tag{4.28}$$

$$= O(1) \log |\mu|^{-2} k \log D + O(1) k \log D \tag{4.29}$$

$$= O(1) \frac{\log |\mu|^{-2}}{\log \Delta^{-1}} \log D. \tag{4.30}$$

□

The strategy of the proof of the area law in [5] is to start from a product state of Schmidt rank one with finite overlap with the ground state, and then apply the AGSP K to approximate the ground state. The initial overlap with the ground state must be sufficiently large.

Lemma 4.2 (Lemma III.2 in [5]). *If there exists a (D, Δ)-AGSP K with $D\Delta \leq 1/2$, then there exists a product state $|\phi\rangle = |L\rangle \otimes |R\rangle$ such that*

$$\mu = \langle \Omega \,|\, \phi \rangle \,, \quad |\mu| \geq 1/\sqrt{2D}. \tag{4.31}$$

Proof. Let $|\phi\rangle = |L\rangle \otimes |R\rangle$ be a product state with $\mu = \langle \phi|\Omega\rangle$ and $|\mu| < 1/\sqrt{2D}$. Write

$$|\phi\rangle = \mu\,|\Omega\rangle + \sqrt{1-|\mu|^2}\,|\Omega^\perp\rangle\,, \tag{4.32}$$

$$\Rightarrow |\phi_1\rangle := K\,|\phi\rangle = \mu\,|\Omega\rangle + \delta\,|\Omega_1^\perp\rangle\,, \quad |\delta|^2 \leq \Delta, \quad SR(\phi_1) = D\,SR(\phi). \tag{4.33}$$

Let $|v\rangle = |\phi_1\rangle / ||\phi_1||$. Writing

$$|v\rangle = \sum_{i=1}^{D} \lambda_i\,|L_i\rangle\,|R_i\rangle\,, \quad \sum_{i=1}^{D} \lambda_i^2 = 1, \tag{4.34}$$

we obtain

$$\frac{|\mu|}{||\phi_1||} = |\,\langle \Omega|v\rangle\,| \leq \sum_{i=1}^{D} \lambda_i |\,\langle \Omega|\,|L_i\rangle\,|R_i\rangle\,| \tag{4.35}$$

$$\leq \sqrt{\sum_{i=1}^{D} |\,\langle \Omega|\,|L_i\rangle\,|R_i\rangle\,|^2} \leq \sqrt{D|\,\langle \Omega|\,|L_{i_{\max}}\rangle\,|R_{i_{\max}}\rangle\,|^2}. \tag{4.36}$$

Here $|\,\langle \Omega|\,|L_{i_{\max}}\rangle\,|R_{i_{\max}}\rangle\,| = \max_i |\,\langle \Omega|\,|L_i\rangle\,|R_i\rangle\,|$, and we used the Cauchy-Schwarz inequality[12]. This gives the lower bound

$$|\,\langle \Omega|\,|L_{i_{\max}}\rangle\,|R_{i_{\max}}\rangle\,|^2 \geq \frac{|\mu|^2}{D||\phi_1||^2} = \frac{|\mu|^2}{D(|\mu|^2 + |\delta|^2)} \geq \frac{|\mu|^2}{D(|\mu|^2 + \Delta)}. \tag{4.37}$$

By the assumptions $D\Delta < 1/2$ and $|\mu| < 1/\sqrt{2D}$,

$$D(|\mu|^2 + \Delta) < 1/2 + 1/2 = 1. \tag{4.38}$$

Thus the overlap $\mu_1 = \langle \Omega|\,|L_{i_{\max}}\rangle\,|R_{i_{\max}}\rangle$ has absolute value strictly larger than $|\mu|$. Moreover the ratio is bounded from below by a μ-independent constant:

$$|\mu_1|/|\mu| \geq \frac{1}{\frac{1}{2} + D\Delta}. \tag{4.39}$$

Therefore it is enough to repeat the AGSP at least n times with

$$\frac{1}{(1/2 + D\Delta)^n} \geq \frac{|\mu|}{\sqrt{2D}}. \tag{4.40}$$

□

[12] $|(x, y)|^2 \leq ||x||^2||y||^2$.

Corollary 4.3 (Corollary III.4 in [5]). *If there exists a* (D, Δ)*-AGSP* K *with* $D\Delta \le 1/2$*, then the entanglement entropy is bounded by*

$$S \le O(1) \log D. \tag{4.41}$$

Proof. We have

$$\log \Delta^{-1} \ge \log(2D). \tag{4.42}$$

Also, there exists a product state $|L\rangle |R\rangle$ with $|\mu|^{-2} \le 2D$. Thus

$$S \le O(1) \frac{\log |\mu|^{-2}}{\log \Delta^{-1}} \log D \le O(1) \frac{\log(2D)}{\log(2D)} \log D = O(1) \log D. \tag{4.43}$$

□

The estimate of $\log D$ in Theorem 3.2 immediately implies the area law.

Theorem 4.4 (Area law for frustration-free systems, Theorem I.1 in [5]). *Consider a one-dimensional, frustration-free, 2-local, gapped system with a unique ground state. Let* d *be the dimension of the local Hilbert space,* $\epsilon > 0$ *the spectral gap, and assume that the local terms of* $H = \sum_i H_i$ *satisfy* $||H_i|| \le J$*. Set*

$$X := \frac{J \log d}{\epsilon}. \tag{4.44}$$

For any cut, the von Neumann entanglement entropy is bounded by

$$S_{1\mathrm{D}} \le O(1) X^3 \log^8 X. \tag{4.45}$$

In the proof we set $||H_i|| \le 1$; here J has been restored.

As a personal comment, the proof in [5] proceeds by bounding the Schmidt rank. Since the contribution to the von Neumann entanglement entropy is $-\lambda_i^2 \log \lambda_i^2$, it seems that one could obtain an even better estimate if one could also bound the Schmidt eigenvalues in addition to the Schmidt rank. I have not yet examined [6], so this point may already be improved there.

A Miscellaneous lemma

Lemma A.1. *Let* $\mathcal{H}$ *be a Hilbert space, and let* X, Y *be orthogonal projections. Let* $v \in \mathcal{H}$ *be a vector of norm one. Assume* $||XYv|| = 1 - \epsilon$*. Then*

$$||(1 - Y)XYv||^2 \le \epsilon(1 - \epsilon). \tag{A.1}$$

First note that for any orthogonal projection P,

$$\begin{aligned} ||(1-P)v||^2 + ||Pv||^2 &= ((1-P)v, (1-P)v) + (Pv, Pv) && \text{(A.2)} \\ &= (v, (1-P)v) + (v, Pv) && \text{(A.3)} \\ &= (v, v) = ||v||^2. && \text{(A.4)} \end{aligned}$$

We need to show

$$||(1-Y)XYv||^2 = ||XYv||^2 - ||YXYv||^2 = 1-\epsilon - ||YXYv||^2 \le \epsilon(1-\epsilon). \tag{A.5}$$

It is enough to prove

$$||YXYv|| \ge 1-\epsilon. \tag{A.6}$$

But

$$||XYv||^2 = (XYv, XYv) = (v, YXYv) \le ||v||\,||YXYv|| = ||YXYv||. \tag{A.7}$$

□

B Proposition D.1 in [7]

Consider general spatial dimension. A Hamiltonian

$$H = \sum_j H_j \tag{B.1}$$

is called local if

$$||[H_j, \mathcal{O}_i]|| \le u(|j-i|) \tag{B.2}$$

holds, where u decays faster than any power. Exponential decay is not required; decay faster than any power is assumed.

> Let $H = \sum_j H_j$ be a gapped local Hamiltonian. Let the energy eigenvalue of the ground state $|\Psi\rangle$ be zero. Let $\Delta > 0$ be the gap. Then $|\Psi\rangle$ can be represented as a zero-energy eigenstate of a local term, $\tilde{H}_j |\Psi\rangle = 0$.

We show this. Define

$$\tilde{H}_j = \int_{-\infty}^{\infty} e^{iHt} H_j e^{-iHt} w(t)\, dt. \tag{B.3}$$

Then

$$\sum_j \tilde{H}_j = H \int_{-\infty}^{\infty} w(t)\, dt. \tag{B.4}$$

Choose $w(t)$ so that

$$\int_{-\infty}^{\infty} w(t)\, dt = 1. \tag{B.5}$$

Equivalently, for the Fourier transform

$$\hat{w}(\epsilon) = \int_{-\infty}^{\infty} dt\, w(t) e^{i\epsilon t}, \tag{B.6}$$

we require

$$\hat{w}(0) = 1. \tag{B.7}$$

If we also require $\tilde{H}_j$ to be Hermitian, then

$$w(t)^* = w(-t) \Leftrightarrow \hat{w}(\epsilon)^* = \hat{w}(-\epsilon). \tag{B.8}$$

For any excited state $|n\rangle$,

$$\left\langle n \middle| \tilde{H}_j \middle| \Psi \right\rangle = \int_{-\infty}^{\infty} e^{iE_n t} \langle n|H_j|\Psi\rangle \, w(t) \, dt = \hat{w}(E_n) \langle n|H_j|\Psi\rangle \, . \tag{B.9}$$

Therefore, choosing $\hat{w}(\epsilon)$ so that

$$\hat{w}(|\epsilon| > \Delta) = 0 \tag{B.10}$$

gives the desired property, with P the projection onto the ground space:

$$(1 - P)\tilde{H}_j P = 0. \tag{B.11}$$

Take a smooth C^∞ function $\hat{w}(\epsilon)$ satisfying

$$\hat{w}(0) = 1, \quad \hat{w}(|\epsilon| > \Delta) = 0. \tag{B.12}$$

Then $w(t)$ decays faster than any power.

It remains to ask whether $\tilde{H}_j$ can be constructed to be local. For any local operator O_i supported near site i,

$$||[\tilde{H}_j, O_i]|| \leq \int_{-\infty}^{\infty} ||[H_j(t), O_i]|| \, |w(t)| \, dt. \tag{B.13}$$

By the Lieb-Robinson bound, roughly,

$$||[H_j(t), O_i]|| \leq c||H_j||||O_i|| \max\{e^{-a(l-v|t|)}, 1\}, \quad l \sim |i - j|. \tag{B.14}$$

[13] Thus

$$||[\tilde{H}_j, O_i]|| \leq c||H_j||||O_i|| \, 2 \left[\int_0^{l/v} e^{-a(l-vt)} |w(t)| \, dt + \int_{l/v}^{\infty} |w(t)| \, dt \right]. \tag{B.15}$$

We want to show that this decays faster than any power in l.

The following proof is due to Kenji Shimomura[14]. Since $w(t)$ decays faster than any power, for any $n \in \mathbb{Z}_{\geq 0}$,

$$|w(t)|t^n \xrightarrow{t\to\infty} 0. \tag{B.16}$$

Since $w(t)$ is bounded, for any $n \in \mathbb{Z}_{\geq 0}$ there exists a constant M_n such that

$$|w(t)| \leq M_n t^{-n} \quad \text{for } t > 0. \tag{B.17}$$

Then

$$\int_{l/v}^{\infty} |w(t)| \, dt \leq \int_{l/v}^{\infty} M_{n+1} t^{-n-1} \, dt = M_{n+1}(-n)^{-1}(-(l/v)^{-n}) = O(l^{-n}), \tag{B.18}$$

and

$$\int_0^{l/v} e^{-a(l-vt)} |w(t)| \, dt \leq \int_0^{l/(2v)} e^{-a(l-vt)} M_0 \, dt + \int_{l/(2v)}^{l/v} e^{-a(l-vt)} M_{n+1} t^{-n-1} \, dt \tag{B.19}$$

$$= M_0 \frac{e^{-al/2} - e^{-al}}{av} + M_{n+1}(l/v)^{-n} \int_{1/2}^{1} e^{-a(l-ls)} s^{-n-1} \, ds \tag{B.20}$$

$$\leq M_0 \frac{e^{-al/2} - e^{-al}}{av} + M_{n+1}(l/v)^{-n} \int_{1/2}^{1} s^{-n-1} \, ds \tag{B.21}$$

$$= M_0 \frac{e^{-al/2} - e^{-al}}{av} + M_{n+1}(l/v)^{-n} \frac{2^n - 1}{n} = O(l^{-n}). \tag{B.22}$$

□

[13] Check, in the currently known results, how much power-law decay is allowed in the assumptions on H_j in the Lieb-Robinson bound, depending on spatial dimension.

[14] Kenji Shimomura, private communication.

C Eckart-Young-Mirsky theorem

For an $m \times n$ matrix $A \in \mathrm{Mat}_{m,n}(\mathbb{C})$, write its singular-value decomposition as

$$A = U\Sigma V^\dagger = \sum_{i=1}^{\mathrm{rank}(A)} \sigma_i u_i v_i^\dagger, \quad \sigma_1 \geq \sigma_2 \geq \dots . \tag{C.1}$$

The Frobenius norm is

$$||A||_F = \sqrt{\sum_{ij} |a_{ij}|^2} = \sqrt{\mathrm{tr}\,[A^\dagger A]} = \sqrt{\sum_{i=1}^{\mathrm{rank}(A)} \sigma_i^2}. \tag{C.2}$$

Here σ_i are the singular values of A. Let

$$A_k = \sum_{i=1}^{k} \sigma_i u_i v_i^\dagger. \tag{C.3}$$

The theorem is:

> In the operator norm and the Frobenius norm, the best approximation of A by a rank-k matrix is given by A_k. That is, for any rank-k matrix B_k,
>
> $$||A - B_k||_F \geq ||A - A_k||_F, \quad ||A - B_k||_2 \geq ||A - A_k||_2. \tag{C.4}$$

This is a note on the proof in [9]. By Weyl's inequality[15], for any i, j,

$$\sigma_{i+j-1}(X + Y) \leq \sigma_i(X) + \sigma_j(Y). \tag{C.7}$$

Let the rank of B be k. Since $\sigma_{i>k}(B) = 0$, setting $j = k+1$, $X = A - B$, and $Y = B$ gives

$$\sigma_{i+k}(A) \leq \sigma_i(A - B) + \sigma_{k+1}(B) = \sigma_i(A - B). \tag{C.8}$$

Thus

$$||A - B||_2 = \sigma_1(A - B) \geq \sigma_{k+1}(A) = \sigma_1(A - A_k). \tag{C.9}$$

Also,

$$||A - B||_F^2 = \sum_i \sigma_i^2(A - B) \geq \sum_i \sigma_{i+k}^2(A) = ||A - A_k||_F^2. \tag{C.10}$$

As a corollary, we obtain:

> Let $|\psi\rangle$ be a normalized state with Schmidt decomposition $|\psi\rangle = \sum_{i\geq 1} \lambda_i |L_i\rangle |R_i\rangle$. Let $|\phi_k\rangle$ be a normalized state of Schmidt rank k. Then
>
> $$|\langle\psi|\phi_k\rangle|^2 \leq \sum_{i\leq k} \lambda_i^2. \tag{C.11}$$

[15] For an $n \times n$ Hermitian matrix A, write its eigenvalues in decreasing order as

$$\lambda_1(A) \geq \lambda_2(A) \geq \cdots . \tag{C.5}$$

For any i, j,

$$\lambda_{i+j-1}(A + B) \leq \lambda_i(A) + \lambda_j(B) \leq \lambda_{i+j-n}(A + B). \tag{C.6}$$

Proof. Let the rank-k approximation be

$$|\psi_k\rangle = \sum_{i\leq k} \lambda_i |L_i\rangle |R_i\rangle , \tag{C.12}$$

and let its normalization be

$$|\tilde{\psi}_k\rangle = |\psi_k\rangle / ||\psi_k|| = |\psi_k\rangle / \sqrt{\sum_{i\leq k} \lambda_i^2}. \tag{C.13}$$

By Eq. (C.10),

$$|| \, |\psi\rangle - |\phi_k\rangle \, ||^2 \geq || \, |\psi\rangle - |\tilde{\psi}_k\rangle \, ||^2. \tag{C.14}$$

Therefore

$$2 - 2\mathrm{Re} \, \langle\psi|\phi_k\rangle \geq 2 - 2\mathrm{Re} \, \langle\psi|\tilde{\psi}_k\rangle , \tag{C.15}$$

and hence

$$\mathrm{Re} \, \langle\psi|\phi_k\rangle \leq \mathrm{Re} \, \langle\psi|\tilde{\psi}_k\rangle = \frac{\sum_{i\leq k} \lambda_i^2}{\sqrt{\sum_{i\leq k} \lambda_i^2}} = \sqrt{\sum_{i\leq k} \lambda_i^2}. \tag{C.16}$$

Since the phase of $|\phi_k\rangle$ is arbitrary, the claim follows.

References

[1] M. B. Hastings, *An Area Law for One Dimensional Quantum Systems*, arXiv:0705.2024.

[2] Dorit Aharonov, Itai Arad, Zeph Landau, Umesh Vazirani, *The Detectability Lemma and Quantum Gap Amplification*, arXiv:0811.3412.

[3] Dorit Aharonov, Itai Arad, Zeph Landau, Umesh Vazirani, *Quantum Hamiltonian complexity and the detectability lemma*, arXiv:1011.3445.

[4] Anurag Anshu, Itai Arad, Thomas Vidick, *A simple proof of the detectability lemma and spectral gap amplification*, arXiv:1602.01210.

[5] Itai Arad, Zeph Landau, Umesh Vazirani, *An improved 1D area law for frustration-free systems*, arXiv:1111.2970.

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

[7] Alexei Kitaev, *Anyons in an exactly solved model and beyond*, math-ph/0507008.

[8] Ayumi Ukai, *On Hastings factorization for quantum many-body systems in the infinite volume setting*, arXiv:2407.12324.

[9] Physics Stack Exchange, *Proof of Eckart-Young-Mirsky theorem*, url.