Quantum Complexity Theory Sevag Gharibian
Summer 2019, University of Paderborn

Lecture 6: The quantum Cook-Levin theorem

> *"Steve Cook was primarily in math but also in the new CS Department. It is to our everlasting shame that we were unable to persuade the math department to give him tenure. Perhaps they would have done so if he had published his proof of the NP-completeness of satisfiability a little earlier."*
> — Richard Karp, speaking about UC Berkeley Computer Science in the late 1960's

Contents

Introduction. Among the many advances of theoretical computer science during the 20th century, three are unquestionably among the crown jewels of the field: First, that there exist *unsolvable* computational problems (Turing's proof that the Halting Problem is undecidable, with Gödel's incompleteness theorem acting as an important precursor). Second, even among solvable problems, not all of them can be solved *efficiently* (the Cook-Levin theorem, which spawned the theory of NP-completeness and arguably founded the field of complexity theory). Third, some problems are not only hard to solve exactly, but are even hard to solve *approximately* (the PCP theorem of the 1990's).

In this lecture, we focus on the Cook-Levin theorem, and in particular its quantum analogue. The latter roughly says that the quantum analogue of Boolean Constraint Satisfaction, known as the *Local Hamiltonian problem*, is QMA-complete. Thus, just as the Cook-Levin theorem says that (assuming P $\neq$ NP) 3-SAT cannot be solved in polynomial time, the *quantum* Cook-Levin theorem implies the Local Hamiltonian problem has no efficient classical or quantum solution.

This statement marks a striking departure from the realm of computation theory into the realm of quantum physics. For the quantum Cook-Levin theorem implies that, at least in principle, there exist quantum many-body systems which *cannot be "cooled to their lowest energy configuration in polynomial time"*. This is particularly enlightening given the original motivation for quantum computation, at least from the perspective of Richard Feynman — that quantum physics seems "difficult" to study with classical computers due to the exponential blowup in dimension. Indeed, the quantum Cook-Levin formally confirms Feynman's intuition, by showing that (assuming QMA $\neq$ BQP) certain properties of low-temperature quantum systems simply cannot be computed efficiently in polynomial time.

This lecture begins by briefly reviewing the classical Cook-Levin theorem, whose techniques will inspire its quantum generalization. We then introduce and motivate "local Hamiltonians", followed by a proof of the quantum Cook-Levin Theorem. For the latter, the soundness proof technique introduces the Geometric Lemma, a lemma with applications beyond quantum complexity theory.

1 The Cook-Levin theorem

Recall that the Cook-Levin theorem states the following, for SAT the generalization of k-SAT in which clauses need not be of size at most k.

Theorem 1 (Cook-Levin Theorem). *SAT is NP-complete.*

In other words, any instance x of a decision problem L in NP can be efficiently encoded into a CNF Boolean formula $\phi : \{0,1\}^m \mapsto \{0,1\}$, such that $x \in L$ if and only if ϕ is satisfiable. The high-level outline of the proof is also exploited for the quantum Cook-Levin theorem; let us hence sketch a proof of Theorem 1 now.

Proof. Let $L \subseteq \{0,1\}^*$ be a language with NP verifier (i.e. deterministic TM) $M = (Q, \Sigma, \Gamma, \delta, q_0, q_{\text{accept}}, q_{\text{reject}})$, and $z \in \{0,1\}^n$ an input. Recall here that Q is the set of states for M, δ its transition function, and Γ its tape alphabet (which contains the input alphabet, Σ).

We sketch a polynomial-time many-one reduction to SAT, i.e. to a CNF formula $\phi : \{0,1\}^m$ which is satisfiable if and only if $z \in L$. To do so, we design ϕ so that it "simulates" M. Any such simulation must presumably capture the following three key properties:

- **Tape initialization.** Before M begins, its tape must be initialized correctly with the input x and blank symbols elsewhere.
- **Correct propagation.** Step $i+1$ of the computation must legally follow from step i according to the rules of the transition function δ for M.
- **Correct output.** Once the final step of the computation ends, the state of M should be q_{accept} if and only if $z \in L$.

With this high-level view in mind, we shall design ϕ to consist of four "components", $\phi_{\text{in}}, \phi_{\text{prop}}, \phi_{\text{out}}, \phi_{\text{alpha}}$, the first three of which correspond to the three bullet points above, respectively.

Construction sketch. We begin by viewing the computation of M as a sequence of *configurations* of M arranged as rows of a table or *tableau*. Specifically, the ith row of the tableau encodes the ith configuration entered by M. Here, recall that a *configuration* is a snapshot in time of M, and is given by string xqy for $x, y \in \Gamma^*$ and $q \in Q$, where xy denote the current tape contents and q the current state of M. The placement of q to the left of y indicates that the head of M is on the first bit of y.

Exercise. What does the starting configuration for M look like on input z? How about the accepting configuration?

As a first step, we require some way to map the symbols which can appear in a configuration (i.e. $Q \cup \Gamma$) to the single-bit literals which ϕ can use. This is achieved by the clever idea of defining, for any symbol $s \in Q \cup \Gamma$, an indicator variable $x_{ijs} \in \{0,1\}$, which is set to 1 if and only if cell (i,j) in the tableau contains symbol s. Thus, for example, we can "simulate" a particular cell (i,j) containing at least one valid symbol via formula $\bigvee_{s \in Q \cup \Gamma} x_{ijs}$.

Exercise. Fix a cell position (i,j). Give a CNF formula encoding the constraint "cell (i,j) contains *precisely one* symbol from $Q \cup \Gamma$".

Exercise. Use the previous exercise to build ϕ_{alpha}, which should enforce that *all* cells (i,j) contain precisely one symbol from $Q \cup \Gamma$.

With our mapping from $Q \cup \Gamma$ to binary literals in place, we can now sketch the remaining components of ϕ.

- **Tape initialization:** ϕ_{in}**.** This CNF formula enforces that the first row of the tableau contains the correct starting configuration for M on input z (where recall $q_0 \in Q$ is the start state of M):

$$\phi_{\text{in}} = x_{11q_0} \wedge x_{12z_1} \wedge x_{13z_2} \cdots$$

Exercise. Fill in the rest of ϕ_{in}. Note that, obviously, we cannot encode all blank symbols on the infinite length tape explicitly into ϕ_{in}. Rather, we may truncate the tape at a finite length — how many cells of the tape are sufficient to keep around?

- **Correct output: ϕ_{out}.** This CNF formula checks whether there exists a time step in which M enters its accepting state:

  $$\phi_{\text{out}} = \bigvee_{(i,j)} x_{ijq_{\text{accept}}}.$$

- **Correct propagation: ϕ_{prop}.** Finally, we must enforce that row $i+1$ of the tableau correctly follows from row i. We omit the full details of this construction, but the key observation is that *computation is local* — from time step i to $i+1$, the string encoding the configuration of M can *only change* at the positions directly adjacent to the head's location. For example, if given configuration $c_i = 000q111$, M writes 0, moves the head right, and enters state q', our new configuration is $c_{i+1} = 0000q'11$ — note that only *two* symbols changed between configurations from i to $i+1$.

  Using this observation, it turns out that to ensure that configuration c_{i+1} follows from c_i, it suffices to check all 2×3 "windows" between rows i and $i+1$ of the tableau. Continuing our example above, rows i and $i+1$ of our tableau

  $$\begin{array}{ccccccc} \hline 0 & 0 & 0 & q & 1 & 1 & 1 \\ \hline 0 & 0 & 0 & 0 & q' & 1 & 1 \\ \hline \end{array}$$

  are fully characterized by the set of five 2×3 windows:

  $$\begin{array}{ccc} \hline 0 & 0 & 0 \\ \hline 0 & 0 & 0 \\ \hline \end{array} \quad \begin{array}{ccc} \hline 0 & 0 & q \\ \hline 0 & 0 & 0 \\ \hline \end{array} \quad \begin{array}{ccc} \hline 0 & q & 1 \\ \hline 0 & 0 & q' \\ \hline \end{array} \quad \begin{array}{ccc} \hline q & 1 & 1 \\ \hline 0 & q' & 1 \\ \hline \end{array} \quad \begin{array}{ccc} \hline 1 & 1 & 1 \\ \hline q' & 1 & 1 \\ \hline \end{array} \tag{1}$$

  Here, for example, the first window depicts the first three symbols of rows i and $i+1$, the second window symbols 2 to 4 of rows i and $i+1$, and so forth. Encoding each of these windows into a CNF formula is done analogously to (e.g.) ϕ_{out}.

  Exercise. Give a CNF formula encoding the constraint that all windows of Equation (1) contain the specified symbols.

  Exercise. Checking 2×3 windows of rows i and $i+1$ seems a bit funny — why does the seemingly more natural idea of simply checking *all* of row i and $i+1$ simultaneously not work? (Hint: How many possible ways are there to validly fill out rows i and $i+1$ of the tableau? How many terms would this lead to in your CNF formula encoding that row $i+1$ correctly follows from row i?)

Finally, the output of the construction is CNF formula $\phi = \phi_{\text{alpha}} \wedge \phi_{\text{in}} \wedge \phi_{\text{out}} \wedge \phi_{\text{prop}}$.

Exercise. Assuming ϕ_{prop} correctly enforces valid propagation from row i to row $i+1$ of the tableau, why is ϕ satisfiable if and only if $z \in L$? □

It is worth pausing to reflect on the crucial fact that made the Cook-Levin theorem possible — that computation (in the TM model) is *local*, meaning only bits around the head can change in any give time step. Remarkably, it turns out that this is no coincidence; in Section 2, we shall see that Nature itself also behaves in a local fashion.

2 Local Hamiltonians and Boolean Constraint Satisfaction

We now move from Boolean constraint satisfaction to "quantum constraint satisfaction". Our motivating theme, foreshadowed by the end of Section 1, shall be that like "computation", Nature is "local".

Hamiltonians. To appreciate the connection between quantum constraint satisfaction and quantum mechanics, we must return to one of the defining equations of the theory: *Schrödinger's equation.* This equation prescribes how quantum systems evolve in time. Namely, given a starting state $|\psi\rangle \in (\mathbb{C}^2)^{\otimes n}$, the rate of change of $|\psi\rangle$ with respect to time t is given by the differential equation (ignoring Planck's constant)

$$i\frac{d|\psi\rangle}{dt} = H|\psi\rangle,$$

where $H \in \text{Herm}((\mathbb{C}^2)^{\otimes n})$ is known as a *Hamiltonian.* In other words, the change in $|\psi\rangle$ through time t is fully specified by H. By solving this equation, one obtains that after time t, our new state $|\psi_t\rangle$ is given by

$$|\psi_t\rangle = e^{-iHt}|\psi\rangle.$$

Exercise. What type of operator is e^{-iHt}? How does this explain the time evolution postulate of quantum mechanics, which states that the set of allowed operations on quantum states is precisely the set of unitary operations?

***Local* Hamiltonians.** Schrödinger's equation tells us that, in principle, any Hamiltonian $H \in \text{Herm}(\mathbb{C}^2)^{\otimes n}$ describes the evolution of *some* quantum system. The natural question is now: *Which Hamiltonians actually arise in Nature?* It is here that Nature takes a page from theoretical computation's book (or perhaps it is theoretical computation which has taken a page from Nature's book), in that essentially all known naturally occurring quantum systems evolve according to *local* Hamiltonians, which we now define.

Definition 2 (k-local Hamiltonian)**.** *A Hermitian operator $H \in \text{Herm}((\mathbb{C}^2)^{\otimes n})$ acting on n qubits is a k-local Hamiltonian if it can be written*

$$H = \sum_{\{S \mid S \subseteq [n] \text{ s.t. } |S|=k\}} H_S \otimes I_{[n]\setminus S},$$

where each operator $H_S \in \text{Herm}((\mathbb{C}^2)^{\otimes k})$ acts on the subset S of qubits. (Note that we allow $H_S = 0$.) The eigenvalues of H denote energy levels *of the system described by H, with $\lambda_{\min}(H)$ denoting the* ground state energy*. The eigenvectors corresponding to $\lambda_{\min}(H)$ are* ground states*.)*

Let us dissect this definition, as it will play a crucial role in the remainder of this course.

- Definition 2 says that the action of H on all n qubits is fully specified by a set of *local* operators H_S, each acting non-trivially on some subset S of k out of n qubits. For example, the following are 2-local Hamiltonians (subscripts denote qubit indexes acted on by the respective operators, and Z is the Pauli operator):

  $$H_1 = Z_1 \otimes Z_2, \qquad H_2 = Z_1 \otimes Z_2 \otimes I_{3,4} + I_1 \otimes Z_2 \otimes Z_3 \otimes I_4 + I_{1,2} \otimes Z_3 \otimes Z_4.$$

  The first of these acts on a 2-qubit system, and the second on a 4-qubit system. For brevity, we typically omit the identity terms and simply write $H_2 = Z_1 \otimes Z_2 + Z_2 \otimes Z_3 + Z_3 \otimes Z_4$.

  Exercise. What are the matrix representations of H_1 and H_2 above?

- Hamiltonians describing physical quantum systems are typically[1] k-local for *constant* k, and there is something very special about this, which you will explore in the next exercise.

[1] This is an idealization — for any actual many-body quantum system, one can only guess at what the defining Hamiltonian should be, and attempt to corroborate this via experiment. Thus, what we write down as "the defining local Hamiltonian" for a system is really a well-motivated model or *approximation* to the true dynamics of the system. It just so happens that setting $k \in O(1)$ typically suffices to accurately reproduce the desired local physical properties of quantum systems.

Exercise. How many bits are required to specify an *arbitrary* Hamiltonian H? How about a k-local Hamiltonian for $k \in O(1)$? What does this tell us about our ability to efficiently represent the dynamics of typical physical systems in Nature?

Thus, although in principle, one requires exponential space to write down the Hamiltonian governing an arbitrary many-body quantum system, in *practice* this is one place we catch a break — for physical systems, we can at least succinctly describe the rules governing their time evolution. However, let us be clear that this is just a minor concession by quantum physics — for even though we can *describe* the dynamics, *computing properties* of the evolution is complexity theoretically hard.

Physical motivation. The eigenvalues of a local Hamiltonian (more generally, of any Hamiltonian) are known as *energy levels*, literally because they represent the energy levels the system may settle into. The quantum state of the system at energy level λ is none other than the eigenvector $|\lambda\rangle$ satisfying $H|\lambda\rangle = \lambda|\lambda\rangle$.

Remark. Despite the fact that quantum time evolution is continuous (i.e. described by unitary maps), the Schrödinger equation highlights that a quantum system may settle into only a *discrete* or "quantized" set of energy values $\{\lambda_i\}$. Indeed, this type of "quantization" phenomenon is precisely what gives *quantum* mechanics its name.

The *ground state energy* $\lambda_{\min}(H)$ plays a particularly important role — it describes the energy level the system will relax into when cooled to very low temperature (think billionths of a degree above absolute zero[2]). This regime is particularly important, as it gives rise to exotic phenomena such as superconductivity and superfluidity. It is thus of utmost importance to fields such as materials design to be able to understand and predict the properties of such low temperature systems; in particular, this means one wishes to understand the properties of $\lambda_{\min}(H)$ and its corresponding eigenvector, the ground state $|\lambda_{\min}(H)\rangle$. Unfortunately, it turns out that $\lambda_{\min}(H)$ is hard to estimate, not least of which because one can embed the answers to NP-complete problems such as 3-SAT into it, as we now discuss.

Embedding k-SAT into local Hamiltonians. We said above that even though a k-local Hamiltonian $H \in \text{Herm}((\mathbb{C}^2)^{\otimes n})$ has a succinct representation in n, computing properties of H is hard. The intuitive reason for this is that H is really the quantum analogue of a k-SAT formula $\phi : \{0,1\}^n \mapsto \{0,1\}$; even though ϕ describes a truth table of size 2^n, it also has a succinct representation of size $\text{poly}(n)$, and computing its properties is NP-complete. We may formalize this connection as follows.

Consider a 2-SAT clause $c = (x_1 \vee \overline{x}_2)$, which has unique unsatisfying assignment $|01\rangle$. We may embed this into a 2-local Hamiltonian term

$$H = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

so that any *satisfying assignment* $x \in \{0,1\}^2$ to c is a ground state of H. More generally, we think of the rows of H as labelled by binary strings 00, 01, 10, 11, and denote the label of row r as $l(r)$. Then, for a 2-SAT constraint c, define H to be all-zeroes except on the diagonal, where we place a 1 if and only if $l(r)$ is a non-satisfying assignment to c.

Exercise. What is the ground state energy of H?

[2] Again, this is an excellent opportunity to procrastinate by heading to the Wikipedia page for "absolute zero". For example, a 1999 experiment cooled nuclear spins in rhodium metal to 0.0000000001 Kelvin. (Recall absolute zero is defined as 0 Kelvin.)

Exercise. Prove that for any $x \in \{0,1\}^2$, $\langle x|H|x\rangle = 0$ if $c(x) = 1$, and $\langle x|H|x\rangle = 1$ if $c(x) = 0$. Conclude that any satisfying assignment to c is a ground state of H. Given the structure of H, why can we assume without loss of generality that the best assignment is a standard basis state?

This construction generalizes directly to *any* Boolean function $c : \{0,1\}^k \mapsto 0, 1$, so that (say) a 3-SAT clause c is embedded into a $2^k \times 2^k$ quantum constraint H.

Exercise. Give the quantum constraint H encoding the 3-SAT clause $c = (\overline{x_1} \vee x_2 \vee \overline{x_3})$.

Suppose now we have three clauses $c_1 = (x_1 \vee x_2)$, $c_2 = (\overline{x_2} \vee x_3)$, $c_3 = (x_3 \vee x_4)$. The full CNF formula $\phi = c_1 \wedge c_2 \wedge c_3$ is given by adding all quantum constraints H_{c_i} for each clause c_i:

$$H = H_{c_1} \otimes I_{3,4} + I_1 \otimes H_{c_2} \otimes I_4 + I_{1,2} \otimes H_{c_3}.$$

Exercise. Prove that $|x\rangle$ for $x \in \{0,1\}^4$ satisfies $\langle x|H|x\rangle = 0$ if and only if $\phi(x) = 1$. More generally, prove that $\langle x|H|x\rangle$ counts the number of unsatisfied clauses for x with respect to ϕ.

Exercise. We are almost ready to conclude that "the ground state energy of H equals 0 if and only if ϕ is satisfiable", thus yielding that estimating $\lambda_{\min}(H)$ is NP-hard if ϕ is a 3-SAT formula. It only remains to prove that if ϕ is unsatisfiable, without loss of generality the *best* assignment $|\psi\rangle \in (\mathbb{C}^2)^{\otimes n}$ (i.e. minimizing $\langle\psi|H|\psi\rangle$) is an n-bit string $|x\rangle$, as opposed to a genuinely quantum state on n qubits; show this.

Exercise. The ground state energy of H above encodes something more precise than whether ϕ is satisfiable — what does it actually *count*?

Exercise. Give a 2-local Hamiltonian whose ground state energy encodes the problem MAX CUT. Recall the latter is defined as: Given a simple, undirected graph $G = (V, E)$, partition E into sets E_1 and E_2 so that the maximum number of edges possible crosses between E_1 and E_2. (Hint: For each edge $(i, j) \in E$, start by thinking about 2-local quantum constraint $H_{ij} = I - Z_i \otimes Z_j$; what is the matrix representation of H_{ij}?)

3 The quantum Cook-Levin theorem

We have seen in Section 2 that, given a 3-local Hamiltonian H as input, estimating its ground state energy $\lambda_{\min}(H)$ allows us to decide 3-SAT instances ϕ, and is thus NP-hard. The problem of estimating $\lambda_{\min}(H)$ more generally (i.e. for non-diagonal H) is sufficiently important to warrant its own name.

Definition 3 (k-local Hamiltonian problem (k-LH))**.** *Fix a polynomial $p : \mathbb{N} \mapsto \mathbb{R}^+$. The promise problem k-LH is defined as follows.*

- *Input:*
  - *A k-local Hamiltonian $H = \sum_S H_S \in \text{Herm}((\mathbb{C}^2)^{\otimes n})$.*
  - *Efficiently computable threshold functions $\alpha(n), \beta(n) \in \mathbb{R}$ satisfying promise gap $\alpha(n) - \beta(n) \geq 1/p(n)$ $\forall n \geq 1$.*
- *Output:*
  - *If $\lambda_{\min}(H) \leq \alpha(n)$, accept.*
  - *If $\lambda_{\min}(H) \geq \beta(n)$, reject.*
  - *Else, accept or reject arbitrarily.*

A few remarks are in order: (1) Unlike the examples of k-SAT from Section 2, in Definition 3 H need not be diagonal in the standard basis, and the sets of qubits S acted on H_S need not be constrained in any particular geometric fashion (say, on a 1D chain). (2) Technically, one should write k-LH(p), since the problem is parameterized by the promise gap polynomial p. We will implicitly set p to the polynomial arising from the QMA-hardness reduction for k-LH below, and henceforth simply write k-LH. (3) The fact that the promise gap $\alpha(n) - \beta(n)$ is at least inverse polynomial is crucial; making the gap, say, inverse *exponential* yields a much more difficult PSPACE-complete problem, rather than a QMA-complete problem for k-LH as defined here. (4) In principle, the number of terms H_S, denoted m, need not be polynomial in n, the number of qubits. Thus, p, α, and β should more generally depend on both m and n. However, for simplicity and due to physical motivation, the community typically assumes $m \in \text{poly}(n)$ and thus drops the m parameter.

Exercise. Why is it not a problem if, say, $m \in \Theta(2^n)$? (Hint: To a computer scientist, relative to which input parameter do we typically define run-times? (It's not just the number of qubits, n.))

Statement of the quantum Cook-Levin theorem. The quantum Cook-Levin theorem of Kitaev states that, just as k-SAT is NP-complete for $k \geq 3$, k-LH is QMA-complete for $k \geq 5$. In other words, Feynman's intuition was right — estimating the ground state energy of a local Hamiltonian, and hence more generally properties of quantum many-body systems, is provably *hard* (assuming QMA $\neq$ BQP).

Theorem 4 (Quantum Cook-Levin Theorem). *There exists an efficiently computable polynomial* $p : \mathbb{N} \mapsto \mathbb{R}^+$ *such that* 5-LH *with promise gap p is QMA-complete.*

The QMA-hardness result above can be improved to $k \geq 2$, even if one has (higher-dimensional) quantum systems on a 1D chain. Due to time constraints and for pedagogical reasons, however, we shall restrict our exposition to Theorem 4. As with any completeness proof, we proceed by showing containment in QMA in Section 3.1, followed by a proof of QMA-hardness in Section 3.2.

3.1 Containment in QMA

Lemma 5. *For any polynomial* $p : \mathbb{N} \mapsto \mathbb{R}^+$ *and* $k \in O(\log n)$*,* k-LH $\in$ QMA*.*

Proof. Let (H, α, β) be an instance of k-LH. We wish to decide in QMA whether $\lambda_{\min}(H) \leq \alpha$ or $\lambda_{\min}(H) \geq \beta$, assuming one of the two is the case. In the YES case, the proof is obvious — the prover sends the ground state $|\psi\rangle$ satisfying $H|\psi\rangle = \lambda_{\min}(H)|\psi\rangle$. The question is: *How do we verify that* $\langle\psi|H|\psi\rangle \leq \alpha$*?*

The key insight is that H is Hermitian, and hence may be viewed as an observable. Recall that for any observable H with spectral decomposition $H = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$, H encodes a projective measurement in basis $\{|\psi_i\rangle\}$ with corresponding outcomes λ_i. Moreover, the expected value of measuring $|\psi\rangle$ with H is

$$\text{Tr}(H|\psi\rangle\langle\psi|) = \langle\psi|H|\psi\rangle.$$

In other words, if we could simulate a measurement in the eigenbasis of H, then we could estimate $\langle\psi|H|\psi\rangle$.

Exercise. What does it mean to "simulate a measurement in a given basis B"? (Hint: Which unitary operation must we be able to perform if we restrict our circuits to only perform measurements in the standard basis, as we did for BQP?)

Unfortunately, the eigenbasis of H may be quite complicated; thus, we cannot directly measure with respect to it. However, the *local* structure of H allows us to approximate such a measurement in a simple fashion: Since by linearity

$$\langle\psi|H|\psi\rangle = \sum_S \langle\psi|H_S|\psi\rangle,$$

intuitively we may equivalently measure the local terms H_S.

Exercise. Suppose $H_S \in \mathrm{Herm}(\mathbb{C}^2)^{\otimes k}$ for $k \in O(\log n)$. How can one efficiently simulate a measurement of $|\psi\rangle$ with respect to the eigenbasis of H_S? (Hint: Use the previous exercise on simulating arbitrary measurements via standard basis measurements.)

We hence apply the following procedure, denoted V: Suppose there are m terms H_S. Pick a term H_S uniformly at random from H, and simulate a measurement with respect to observable $\langle\psi|H_S|\psi\rangle$. Since expectation is linear, the total expectation for this procedure with respect to the random choice of S is

$$\sum_S \Pr[\text{picking } S] \cdot \langle\psi|H_S|\psi\rangle = \frac{1}{m}\sum_S \langle\psi|H_S|\psi\rangle = \frac{1}{m}\langle\psi|\left(\sum_S H_S\right)|\psi\rangle = \frac{1}{m}\langle\psi|H|\psi\rangle.$$

Exercise. In principle, m may be exponential with respect to the number of qubits n. Why is this not a problem for the protocol above? (Hint: With respect to which parameter must we run efficiently?)

Recalling the Courant-Fischer variational characterization of eigenvalues from Lecture 4, we conclude that in the YES case, there exists a quantum proof $|\psi\rangle$ such that the expected output value of V is at most α/m, and in the NO case, any proof $|\psi\rangle$ has expected value at least β/m.

There is only one thing left to do — the definition of QMA says nothing about *expectation values* of the verification. Rather, we must strengthen our expectation bounds for V to a *high-probability statement*: In the YES case, our verifier must accept with probability at least 2/3, and in the NO case with probability at most 1/3.

Exercise. Give an example of a probability distribution over an appropriate sample space so that the expected value is 0, and yet the probability of obtaining any outcome with value v satisfying $|v| \leq \epsilon$ (for some fixed $\epsilon > 0$) is zero. In other words, a statement about expectation is in general *not* sufficient to yield a high-probability statement.

To give a high probability statement, we employ the *Höffding bound*, which is worth knowing in its own right. Let $\Omega \subseteq \mathbb{R}$ denote our sample space, meaning the union over all eigenvalues of all terms H_S. Let $X_i \in \Omega$ denote the random variable corresponding to the measurement outcome of the ith run of V, given state $|\psi\rangle$. Intuitively, if we repeat V N times and take the average measurement result, $A := (\sum_{i=1}^N X_i)/N$, we might expect A to approximate the true expected value, $\langle\psi|H|\psi\rangle$. To formalize this, the Höffding bound says that if $a_i \leq X_i \leq b_i$ for all i, and if the X_i are independent, then

$$\Pr[|A - E[A]| \geq t] \leq 2^{-\frac{2N^2t^2}{\sum_{i=1}^N (b_i - a_i)^2}}.$$

Thus, the final verification procedure, denoted V', is precisely this: Take in N copies of the proof $|\psi\rangle$. For each copy $|\psi_i\rangle$, independently repeat the procedure V to obtain outcome $X_i \in \Omega$. If $A := (\sum_{i=1}^N X_i)/N \leq \alpha + (1/4p(n))$, accept, and if $A \geq \beta - (1/4p(n))$, reject.

Exercise. Use the Höffding bound to prove that in the YES case, for sufficiently large polynomial N, V' accepts with probability exponentially close to 1. For simplicity, you may assume $0 \preceq H_S \preceq I$ for all H_S, i.e. all H_S have eigenvalues between 0 and 1, and so $a_i = 0$ and $b_i = 1$ in the statement of the Höffding bound.

Exercise. In the NO case, the prover can again cheat by sending a large entangled state over the N copies of the proof space; why does V' still reject with high probability in this case? □

3.2 Hardness for QMA

In Section 3.1, we showed k-LH $\in$ QMA for $k \in O(\log n)$. Theorem 4 thus immediately follows from the following lemma.

Lemma 6. k-LH *is QMA-hard under polynomial-time many-one reductions for $k \geq 5$.*

Proof. The proof is based on an old trick of Feynman, employed in a clever way by Kitaev; a similar trick was used in the proof of BQP-completeness for the Matrix Inversion problem from Lecture 4 (which was discovered after Theorem 4). To begin, let $A = (A_{\text{yes}}, A_{\text{no}}, A_{\text{inv}})$ denote a QMA promise problem. Let $x \in \{0,1\}^n$ be an input, with corresponding QMA verifier $V = V_m \cdots V_1 \in (\mathbb{C}^2)^{\otimes n} \otimes (\mathbb{C}^2)^{\otimes p(n)} \otimes (\mathbb{C}^2)^{\otimes q(n)}$. Recall V is a uniformly generated quantum verification circuit consisting of 1- and 2-qubit unitary gates, acting on registers A (n qubits containing the input x), B ($p(n)$ qubits containing the proof $|\psi\rangle$), and C ($q(n)$ ancilla qubits initialized to all zeroes). Assume without loss of generality that the completeness and soundness parameters for V are $1-\epsilon$ and ϵ, so that $1-2\epsilon \in \Omega(1/\operatorname{poly}(n))$. Our goal is to construct an instance (H, α, β) of 5-local Hamiltonian H such that, if $x \in A_{\text{yes}}$, then $\lambda_{\min}(A) \leq \alpha$, and if $x \in A_{\text{no}}$, then $\lambda_{\min}(A) \geq \beta$.

Construction. The high level setup is analogous to that of the classical Cook-Levin theorem (Theorem 1), in that we will track a sequence of "quantum configurations"' $|\psi_t\rangle$ over time, and use "local Hamiltonian checks" to ensure the propagation from configuration $|\psi_t\rangle$ to $|\psi_{t+1}\rangle$ proceeds correctly. The main difference is that instead of encoding each configuration as a row of a tableau, we shall encode it as a term in a superposition, i.e. as $\sum_t |\psi_t\rangle$. Unfortunately, in doing so, we lose our notion of *time*, in that for a tableau, time was encoded by *position* — the row index of a configuration in the tableau gave away the time step during which the configuration was entered by the verifier. To recover this, we use an idea of Feynman and attach a new ancilla register to track time, D, denoted the "clock" register. Thus, we aim to encode the computation as a state of the form $\sum_t |\psi_t\rangle_{A,B,C}|t\rangle_D$.

It remains to specify what a "quantum configuration" $|\psi_t\rangle$ for time t should be — but this is easy, since at time t we have applied the first t gates. In other words, defining $|\psi_t\rangle := V_t \cdots V_1 |x\rangle_A |\psi\rangle_B |0\cdots 0\rangle_C$, we arrive at the so-called *history state*

$$|\psi_{\text{hist}}\rangle = \frac{1}{\sqrt{m+1}} \sum_{t=0}^{m} V_t \cdots V_1 |x\rangle_A |\psi\rangle_B |0\cdots 0\rangle_C |t\rangle_D.$$

Just as Theorem 1 used local Boolean checks to force a tableau to have certain properties, we now use local Hamiltonian terms to force a quantum state to look like $|\psi_{\text{hist}}\rangle$. Specifically, we will design $H = H_{\text{in}} + H_{\text{out}} + H_{\text{prop}}$ with the hope of making $|\psi_{\text{hist}}\rangle$ its ground state.

- **Ancilla initialization:** H_{in}**.** This constraint enforces that at time step $t = 0$, the A register reads $|x\rangle$ for x the input, and the ancilla register C is all zeroes:

  $$H_{\text{in}} = (I - |x\rangle\langle x|)_A \otimes I_B \otimes I_C \otimes |0\rangle\langle 0|_D + I_A \otimes I_B \otimes (I - |0\cdots 0\rangle\langle 0\cdots 0|)_C \otimes |0\rangle\langle 0|_D.$$

  Exercise. Prove $H_{\text{in}} \succeq 0$.

  Exercise. Let $|\phi(y)\rangle := |x\rangle_A |\psi\rangle_B |y\rangle_C |0\rangle_D$. Prove that $\langle \phi(y)|H_{\text{in}}|\phi(y)\rangle$ equals 0 if $y = 0^{q(n)}$, and equals 1 if y has Hamming weight at least one.

  Exercise. As stated, H_{in} is not local — the projector in C, for example, acts non-trivially and simultaneously on $q(n)$ qubits. Show that replacing $\Delta := (I - |0\cdots 0\rangle\langle 0\cdots 0|)_C$ with $\Delta' := \sum_{i=1}^{q(n)} |1\rangle\langle 1|_{C_i}$ in H_{in} obeys the same properties regarding $|\phi(y)\rangle$ as in the previous exercise, and that Δ' is 1-local. How can we similarly reduce the locality of $(I - |x\rangle\langle x|)_A$ to 1 in H_{in}?

- **Correct output:** H_{out}**.** This constraint checks whether, at time step m, the verifier accepted (recall the verifier's output qubit is C_1):

  $$H_{\text{out}} = I_A \otimes I_B \otimes |0\rangle\langle 0|_{C_1} \otimes |m\rangle\langle m|_D.$$

Exercise. Why do we project onto $|0\rangle\langle 0|_{C_1}$ above, as opposed to $|1\rangle\langle 1|_{C_1}$?

Exercise. Prove $H_{\text{out}} \succeq 0$.

- **Correct propagation:** H_{prop}**.** Interestingly, specifying the propagation Hamiltonian H_{prop} is simpler than specifying ϕ_{prop} classically. Namely,

$$H_{\text{prop}} = \sum_{t=0}^{m-1} -V_{t+1} \otimes |t+1\rangle\langle t|_D - V_{t+1}^{\dagger} \otimes |t\rangle\langle t+1|_D + I \otimes |t\rangle\langle t|_D + I \otimes |t+1\rangle\langle t+1|_D,$$

  where recall V_t acts on A, B, C. The intuition is best captured by the first term above, which encodes the idea that in going from time step t to $t+1$, we must apply V_{t+1}.

  Exercise. For any state of the form $|\phi\rangle = \frac{1}{\sqrt{m+1}} \sum_{t=0}^{m} V_t \cdots V_1 |\eta\rangle_{A,B,C} |t\rangle_D$, prove that $H_{\text{prop}}|\phi\rangle = 0$. Conclude that vectors encoding "correct propagation according to V" fall into the null space of H_{prop}.

Finally, the output of the construction is Hamiltonian $H = H_{\text{in}} + H_{\text{prop}} + H_{\text{out}}$. We will choose α and β as needed based on the correctness analysis below.

Exercise. Is H as specified 5-local? (Hint: Think about the clock register.) We will revisit this question at the end of the proof.

Correctness. We now show correctness.

Completeness. Assume first $x \in A_{\text{yes}}$. Then, there exists a proof $|\psi\rangle$ accepted by V with probability at least $1 - \epsilon$. We must prove that $\lambda_{\min}(H) \leq \alpha$, or equivalently, there exists $|\phi\rangle$ such that $\langle\phi|H|\phi\rangle \leq \alpha$. The obvious choice is to choose $|\phi\rangle$ as the history state $|\psi_{\text{hist}}\rangle$. Then,

$$\langle\psi_{\text{hist}}|H|\psi_{\text{hist}}\rangle = \langle\psi_{\text{hist}}|H_{\text{in}}|\psi_{\text{hist}}\rangle + \langle\psi_{\text{hist}}|H_{\text{out}}|\psi_{\text{hist}}\rangle + \langle\psi_{\text{hist}}|H_{\text{prop}}|\psi_{\text{hist}}\rangle = \langle\psi_{\text{hist}}|H_{\text{out}}|\psi_{\text{hist}}\rangle,$$

which follows by the previous exercises in this proof.

Exercise. Prove that $\langle\psi_{\text{hist}}|H_{\text{out}}|\psi_{\text{hist}}\rangle = \frac{1}{m+1} \Pr[V \text{ rejects } |\psi\rangle] \leq \frac{\epsilon}{m+1}$.

Thus, setting $\alpha := \frac{\epsilon}{m+1}$, we have shown the YES case.

Soundness. Assume now $x \in A_{\text{no}}$. Then, for all proofs $|\psi\rangle$, V accepts $|\psi\rangle$ with probability at most ϵ. We must prove that $\lambda_{\min}(H) \geq \beta$, or equivalently, for all states $|\phi\rangle$, $\langle\phi|H|\phi\rangle \geq \beta$. Unfortunately, due to the universal quantifier on $|\psi\rangle$ we can no longer give a simple constructive proof as in the YES case. Things are further complicated by the fact that the terms $H_{\text{in}}, H_{\text{out}}, H_{\text{prop}}$ do not pairwise commute, so it is not in general true that $\lambda_{\min}(H) = \lambda_{\min}(H_{\text{in}}) + \lambda_{\min}(H_{\text{out}}) + \lambda_{\min}(H_{\text{prop}})$.

Exercise. If $[A, B] = 0$ for normal operators A, B, prove that $\lambda_{\min}(A + B) = \lambda_{\min}(A) + \lambda_{\min}(B)$. (Hint: Recall the simultaneous diagonalization theorem, which states that normal operators A and B commute if and only if they diagonalize in a common eigenbasis.)

As lower bounding $\lambda_{\min}(H)$ is rather involved, let us state the result below as a lemma, and prove it in Section 3.2.1.

Lemma 7. *If $x \in A_{\text{no}}$, it holds that $\lambda_{\min}(H_{\text{in}} + H_{\text{out}} + H_{\text{prop}}) \geq \frac{\pi^2(1-\sqrt{\epsilon})}{2(m+1)^3}$.*

Thus, setting $\beta = \frac{\pi^2(1-\sqrt{\epsilon})}{2(m+1)^3}$ completes the proof of the NO case, with the exception of one observation.

Exercise. Is it true in general that $\alpha - \beta \geq 1/\operatorname{poly}(n)$? What value of ϵ suffices for this to hold, and why can we let ϵ take this value without loss of generality?

The last stand: Locality. We have shown correctness for our many-one reduction to Hamiltonian $H = H_{\text{in}} + H_{\text{out}} + H_{\text{prop}}$. Congratulations! There is one tiny problem, however; as foreshadowed in a previous exercise, H is technically not $O(1)$-local. This is because implicitly we have assumed that the clock register D is encoded in *binary*, and is hence $\Theta(\log m)$ qubits in size. To correctly identify a time step written in binary, we must read *all* the bits in D; thus, the projectors onto D in $H_{\text{in}}, H_{\text{out}}, H_{\text{prop}}$ are all $O(\log m)$-local.

The solution is rather simple; let us encode D in *unary*. Specifically, encode time $t \in \{0, \ldots, m\}$ as $1^t 0^{m-t}$. There is a tradeoff here — now the D register is m qubits, up from $\Theta(\log m)$ qubits. However, to check the current time step, we can do a *local* check — namely, we just have to identify the position of the leading 1 in $1^t 0^{m-t}$. And this identification can be made by checking at most 3 qubits of D at a time. With respect to this new encoding, one can rewrite the terms of $H_{\text{in}}, H_{\text{out}}, H_{\text{prop}}$ acting on D to obtain a new Hamiltonian H (we omit this for brevity). Since it now takes three qubit checks to identify the time in register D, and since H_{prop} now pairs such 3-local clock checks with 2-local gates V_t, the new Hamiltonian H we get is 5-local as claimed.

The final hitch is that we now have to add yet more constraints to H to enforce that the states appearing in D are indeed valid unary time encodings of form $1^t 0^m - t$ (e.g. we want to disallow setting D to $|01^{m-1}\rangle$). This is accomplished by adding a fourth Hamiltonian term to H:

$$H_{\text{stab}} := I_{A,B,C} \otimes \sum_{i=1}^{m-1} |0\rangle\langle 0|_i \otimes |1\rangle\langle 1|_{i+1}.$$

Exercise. Why does H_{stab} correctly enforce unary time encodings in register D?

Our final Hamiltonian is $H = H_{\text{in}} + H_{\text{out}} + H_{\text{prop}} + H_{\text{stab}}$, with the time register encoded in unary. Of course, now one must revisit the soundness analysis to account for the H_{stab} term (we do not need to repeat the completeness analysis, since in the YES case an honest prover will correctly encode D anyway). It turns out this can be done rather easily; nevertheless, again we shall omit it for brevity. □

3.2.1 Proof of soundness via Geometric Lemma

We now prove the eigenvalue lower bound of Lemma 7 required to complete the soundness analysis of Theorem 4.

Proof of Lemma 7. It will be enlightening to first rewrite H_{prop} via a unitary change of basis, i.e. to instead consider $UH_{\text{prop}}U^\dagger$ for some cleverly chosen U.

Exercise. Why is $\lambda_{\min}(H_{\text{in}} + H_{\text{prop}} + H_{\text{out}}) = \lambda_{\min}(UH_{\text{in}}U^\dagger + UH_{\text{prop}}U^\dagger + UH_{\text{out}}U^\dagger)$? Conclude that there is no loss of generality in applying a unitary change of basis for the purposes of our proof. More generally, observe that for eigenvalue problems, a trick one should always keep in mind is the potential for viewing the problem in a different basis (which may simplify the matrices involved).

A convenient change of basis. The intuition for choosing U is this: H_{prop}, reproduced below for convenience,

$$H_{\text{prop}} = \sum_{t=0}^{m-1} -V_{t+1} \otimes |t+1\rangle\langle t|_D - V_{t+1}^\dagger \otimes |t\rangle\langle t+1|_D + I \otimes |t\rangle\langle t|_D + I \otimes |t+1\rangle\langle t+1|_D,$$

"feels" a lot like a random walk on a line. It has four equally weighted terms, two of which correspond to staying in the same spot (i.e. time step) and doing nothing (i.e. I), one of which corresponds to moving to right one step on the line (i.e. forward one time step) and applying V_{i+1}, and one of which corresponds to moving left one step on the line (i.e. backward one time step) and applying $V_{i+1}^\dagger$. In fact, if it wasn't for those pesky V_{i+1} terms, it would essentially be a random walk on a line, where with probability 1/2 we stay put, and otherwise we flip a fair coin and move right one step if we get heads and left one step if we get tails. And indeed, we can get rid of those V_{i+1} terms by conjugating H_{prop} by unitary

$$U = \sum_{t=0}^{m} V_1^\dagger \cdots V_t^\dagger \otimes |t\rangle\langle t|_D.$$

Exercise. Prove U is unitary.

Exercise. Prove that

$$UH_{\text{prop}}U^\dagger = \sum_{t=0}^{m-1} -I \otimes |t+1\rangle\langle t|_D - I \otimes |t\rangle\langle t+1|_D + I \otimes |t\rangle\langle t|_D + I \otimes |t+1\rangle\langle t+1|_D.$$

The exercise above shows that under the change of basis U, H_{prop} acts non-trivially *only on the clock register*, D. Thus, it has a nice matrix representation via $UH_{\text{prop}}U^\dagger = I_{A,B,C} \otimes \Lambda_D$ for

$$\Lambda := \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & \cdots \\ -1 & 2 & -1 & 0 & 0 & \cdots \\ 0 & -1 & 2 & -1 & 0 & \cdots \\ 0 & 0 & -1 & 2 & -1 & \cdots \\ 0 & 0 & 0 & -1 & \ddots & \ddots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix}.$$

Exercise. Prove that Λ indeed has the matrix representation above.

Exercise. Write down the transition matrix for a random walk on a line of length $m+1$, where with probability 1/2 we stay put, and otherwise we flip a fair coin and move right one step if we get heads and left one step if we get tails. How does Λ compare with this transition matrix?

The advantage of this representation for $UH_{\text{prop}}U^\dagger$ is that it is easier to understand its spectral decomposition (indeed, the matrix is "almost diagonal" now). Using tools from the analysis of 1D random walks, one may now show that the eigenvalues of Λ are $\lambda_k(\Lambda) = 2(1 - \cos(\pi k/(m+1)))$, a fact we will use later.

Exercise. What is the full set of eigenvalues for H_{prop}?

Exercise. Prove that the *unique* null vector (and in this case, ground state) of Λ is

$$|\gamma\rangle := \frac{1}{\sqrt{m+1}} \sum_{t=0}^{m} |t\rangle.$$

(Hint: Checking that $\Lambda|\gamma\rangle = 0$ is straightforward. For uniqueness, use the fact that if a normal operator has a non-degenerate spectrum (i.e. all eigenvalues are distinct), then it has a unique eigenbasis.)

By applying the change of basis U, we hence have a full understanding of H_{prop}. It remains to understand how U affects the other important quantities in this proof: H_{in}, H_{out}, and $|\psi_{\text{hist}}\rangle$.

Exercise. Prove that $UH_{\text{in}}U^\dagger = H_{\text{in}}$.

Exercise. Prove that $UH_{\text{out}}U^\dagger = (V^\dagger_{A,B,C} \otimes I_D)H_{\text{out}}(V_{A,B,C} \otimes I_D)$.

Exercise. Prove that $U|\psi_{\text{hist}}\rangle = |x\rangle_A|\psi\rangle_B|0\cdots 0\rangle_C|\gamma\rangle_D$.

In the remainder of this proof, we hence work in this new basis and refer to quantities $H'_{\text{in}} := H_{\text{in}}$, $H'_{\text{out}} := (V^\dagger_{A,B,C} \otimes I_D)H_{\text{out}}(V_{A,B,C} \otimes I_D)$, $H'_{\text{prop}} := I_{A,B,C} \otimes \Lambda_D$, and $|\psi'_{\text{hist}}\rangle := |x\rangle_A|\psi\rangle_B|0\cdots 0\rangle_C|\gamma\rangle_D$.

The Geometric Lemma. Recall that our goal is now to prove $\lambda_{\min}(H'_{\text{in}} + H'_{\text{out}} + H'_{\text{prop}}) \geq \beta$, and the difficulty is that the matrices involved do not all pairwise commute. However, one pair *does* commute.

Exercise. Prove that $[H_{\text{in}}, H_{\text{out}}] = 0$.

Let us hence write $G := H'_{\text{in}} + H'_{\text{out}}$, so that we want to lower bound $\lambda_{\min}(G + H'_{\text{prop}})$. For this, we will need a technical tool known as the *Geometric Lemma*, whose intuition we now explain. We know from our previous discussion that $G \succeq 0$ and $H'_{\text{prop}} \succeq 0$. Thus, for any $|\phi\rangle$ the minimum expectation we can hope for against G and H_{prop} is $\langle\phi|G|\phi\rangle = 0$ and $\langle\phi|H'_{\text{prop}}|\phi\rangle = 0$. The question is: *Can we attain zero for both expressions simultaneously, i.e. do G and H'_{prop} share a common null vector?* Unfortunately, the answer is no.

Exercise. Show that $\text{Null}(H'_{\text{prop}}) = (\mathbb{C}^2)^{\otimes n}_A \otimes (\mathbb{C}^2)^{\otimes p(n)}_B \otimes (\mathbb{C}^2)^{\otimes q(n)}_C \otimes |\gamma\rangle_D$.

Exercise. Show that $\text{Null}(G)$ is the direct sum of 3 orthogonal spaces, i.e. $\text{Null}(G) = N_1 \oplus N_2 \oplus N_2$ for

$$\begin{aligned}
N_1 &:= |x\rangle_A \otimes (\mathbb{C}^2)^{\otimes p(n)}_B \otimes |0\cdots 0\rangle_C \otimes |0\rangle_D\\
N_2 &:= (\mathbb{C}^2)^{\otimes n}_A \otimes (\mathbb{C}^2)^{\otimes p(n)}_B \otimes (\mathbb{C}^2)^{\otimes q(n)}_C \otimes \text{Span}(|1\rangle, \ldots, |m-1\rangle)_D\\
N_3 &:= \{|\eta\rangle \mid V|\eta\rangle \text{ has qubit } C_1 \text{ set to } |1\rangle\}_{A,B,C} \otimes |m\rangle_D
\end{aligned}$$

Exercise. Conclude that $\text{Null}(H'_{\text{prop}}) \cap \text{Null}(G) = \emptyset$, as claimed.

We thus have that for any choice of $|\phi\rangle$, at most one of $\langle\phi|G|\phi\rangle$ and $\langle\phi|H'_{\text{prop}}|\phi\rangle$ can be zero; thus minimizing $\langle\phi|G|\phi\rangle + \langle\phi|H'_{\text{prop}}|\phi\rangle$ is a balancing act between "not upsetting" either G or H'_{prop} too much. Intuitively, the "minimum joint unhappiness" experienced by G and H'_{prop} relative to a state $|\phi\rangle$ should depend on "how close" their null spaces are — if there is a $|\phi\rangle$ that is "close" to both $\text{Null}(G)$ and $\text{Null}(H'_{\text{prop}})$, then we might expect to minimize $\langle\phi|G|\phi\rangle + \langle\phi|H'_{\text{prop}}|\phi\rangle$ "well". Conversely, if $\text{Null}(G)$ and $\text{Null}(H'_{\text{prop}})$ are "far", then we might expect $\lambda_{\min}(G + H'_{\text{prop}})$ to be "large". This is precisely the intuition captured by the Geometric Lemma (whose proof uses elementary Linear Algebra, and which we omit for brevity).

Lemma 8 (Geometric Lemma)**.** *Let $A_1, A_2 \succeq 0$, and let v lower bound the minimum* non-zero *eigenvalues of both A_1 and A_2. Then,*

$$\lambda_{\min}(A_1 + A_2) \geq 2v \sin^2 \frac{\angle(\text{Null}(A_1), \text{Null}(A_2))}{2},$$

where the angle *between spaces $\mathcal{X}$ and $\mathcal{Y}$ is defined as* $\angle(\mathcal{X}, \mathcal{Y}) := \arccos\left[\max_{\substack{|x\rangle\in\mathcal{X}, |y\rangle\in\mathcal{Y}\\ \| \, |x\rangle \, \|_2 = \| \, |y\rangle \, \|_2 = 1}} |\langle x|y\rangle|\right]$.

Thus, the correct notion of "closeness" for $\text{Null}(G)$ and $\text{Null}(H'_{\text{prop}})$ is the *angle* between the two spaces.

Exercise. Suppose $\text{Null}(A_1) \cap \text{Null}(A_2) \neq \emptyset$. Why is the lower bound given by the Geometric Lemma trivial in this case, and why is this expected?

We now have a clear approach for trying to show $\lambda_{\min}(G + H'_{\text{prop}}) \geq \beta$ — apply the Geometric Lemma with $A_1 = G$ and $A_2 = H'_{\text{prop}}$. It just remains to figure out v and $\angle(\text{Null}(G), \text{Null}(H'_{\text{prop}}))$.

Exercise. Prove that the smallest non-zero eigenvalue of G is 1. (Hint: Recall H_{in} and H_{out} are commuting projectors.)

Exercise. Use the formula for the eigenvalues of Λ to show that the smallest non-zero eigenvalue of H'_{prop} is $2(1 - \cos(\pi/(m+1))) \geq \pi^2/(m+1)^2$. (Hint: Taylor series.)

By the two exercises above, we may set $v = \pi^2/(m+1)^2$. The next lemma suffices to finish the proof of Lemma 7.

Lemma 9. *It holds that*

$$\sin^2 \frac{\angle(\text{Null}(G), \text{Null}(H'_{\text{prop}}))}{2} \geq \frac{1-\sqrt{\epsilon}}{4(m+1)}.$$

Proof. Since $\sin^2 x = 1 - \cos^2 x$, it suffices to show the bound

$$\cos^2 \angle(\text{Null}(G), \text{Null}(H'_{\text{prop}})) \leq 1 - \frac{1-\sqrt{\epsilon}}{m+1}. \tag{2}$$

Exercise. Prove that Lemma 9 follows from Equation (2).

To show Equation (2), we must upper bound

$$\max_{\substack{|x\rangle \in \text{Null}(G), |y\rangle \in \text{Null}(H'_{\text{prop}}) \\ \| |x\rangle \|_2 = \| |y\rangle \|_2 = 1}} |\langle x|y\rangle|^2 = \max_{\substack{|x\rangle \in \text{Null}(G), |y\rangle \in \text{Null}(H'_{\text{prop}}) \\ \| |x\rangle \|_2 = \| |y\rangle \|_2 = 1}} \langle y|x\rangle\langle x|y\rangle = \max_{\substack{|y\rangle \in \text{Null}(H'_{\text{prop}}) \\ \| |y\rangle \|_2 = 1}} \langle y|\Pi_{\text{Null}(G)}|y\rangle,$$

for $\Pi_{\mathcal{X}}$ the projector onto a space $\mathcal{X}$. Since $\text{Null}(G) = N_1 \oplus N_2 \oplus N_3$, we may write $\Pi_{\text{Null}(G)} = \Pi_{N_1} + \Pi_{N_2} + \Pi_{N_3}$. Handling Π_{N_2} can be done directly.

Exercise. Prove that $\max_{\substack{|y\rangle \in \text{Null}(H'_{\text{prop}}) \\ \| |y\rangle \|_2 = 1}} \langle y|\Pi_{N_2}|y\rangle = \frac{m-1}{m+1}$.

To next relate our bound to ϵ, which encodes the probability of acceptance, we must consider $\Pi_{N_1} + \Pi_{N_3}$ jointly. Defining $\mathcal{X} := |x\rangle_A \otimes (\mathbb{C}^2)_B^{\otimes p(n)} \otimes |0 \cdots 0\rangle_C$ and $\mathcal{Y} := \{|\eta\rangle \mid V|\eta\rangle$ has qubit C_1 set to $|1\rangle\}_{A,B,C}$, and recalling that any $|y\rangle \in \text{Null}(H_{\text{prop}})$ has form $|\eta\rangle_{A,B,C} \otimes |\gamma\rangle_D$, we have

$$\langle y|\Pi_{N_1} + \Pi_{N_3}|y\rangle = \langle y|(\Pi_{\mathcal{X}} \otimes |0\rangle\langle 0|_D + \Pi_{\mathcal{Y}} \otimes |m\rangle\langle m|_D)|y\rangle = \frac{1}{m+1}\langle \eta|(\Pi_{\mathcal{X}} + \Pi_{\mathcal{Y}})|y\rangle.$$

In other words, maximizing the left hand side has been reduced to upper bounding $\lambda_{\max}(\Pi_{\mathcal{X}} + \Pi_{\mathcal{Y}})$. Luckily, the main statement derived in the proof of the Geometric Lemma (which, recall, we've omitted) gives us exactly the requisite tool for this.

Fact 10. *For spaces $\mathcal{A}$ and $\mathcal{B}$, $\lambda_{\max}(\Pi_{\mathcal{A}} + \Pi_{\mathcal{B}}) \leq 1 + \cos(\angle(\mathcal{A}, \mathcal{B}))$.*

Exercise. Use Fact 10 to prove that $\cos^2(\angle(\mathcal{X}, \mathcal{Y})) \leq \epsilon$. Conclude that $\langle y|\Pi_{N_1} + \Pi_{N_3}|y\rangle \leq \frac{1+\sqrt{\epsilon}}{m+1}$.

Exercise. Combine the bounds we have computed to conclude that Equation (2) holds, thus completing the proof of Lemma 9. □

Exercise. Combine our choice of v and Lemma 9 to obtain Lemma 7.

$\square$