Lecturer: Tongyang, scribed by Shuo Zhou
Date: November 24, 2023

Lecture 15

Sparse Hamiltonian Simulation

- Sparse Hamiltonian Simulation
- Continuous-time Quantum Walk

1 Sparse Hamiltonian Simulation

Last time: Efficient simulation of Hamiltonians, product formulas. k-local Hamiltonian can be efficiently simulated for $k = O(1)$.

More general Hamiltonians? Further extensions that contain general classes while efficiency.

We say that a $N \times N$ Hermitian is sparse (in a fixed basis) if, in any row, there are only $\text{poly}(\log N) = \text{poly}(n)$ non-zero entries.

Note that a k-local Hamiltonian is $2^k C_n^k$-sparse Hamiltonian. This is because a k-local Hamiltonian has at most C_n^k terms, each of which is 2^k-sparse.

Therefore, the overall Hamiltonian is $2^k C_n^k$-sparse.

As long as $k = O(1)$, a $O(1)$-local Hamiltonian is sparse.

Assumption: **Query model of sparse matrix:** Given a row index a, one can determine all of bs for which $\langle a| H |b\rangle$ is non-zero.

Our specific research on the simulation of sparse Hamiltonians is motivated by the following three reasons:

- On the one hand, sparse Hamiltonian could be generally efficiently simulated.
- On the other hand, dense Hamiltonian simulation is in general difficult.
- Furthermore, sparse Hamiltonian simulation has wide application in computing.

Observation: We can regard the entries of H as the adjacency matrix of a graph.

$$G = (V, E),\ |V| = N = 2^n,\ (i, j) \in E \text{ if and only if } H_{ij} \neq 0$$

$$H \text{ is } d\text{-sparse} \iff G \text{ is a graph of maximum degree} \leq d.$$

Idea: Decompose d-sparse Hamiltonian into 1-sparse Hamiltonians.

Vizing's theorem: A graph of max degree d has an edge coloring with $\leq d + 1$ colors.

However, Vizing's theorem is inefficient in general.

We only have the information of neighbors of vertex (local information).

Observation 1. Given H, $X \otimes H = \begin{pmatrix} & H \\ H & \end{pmatrix}$

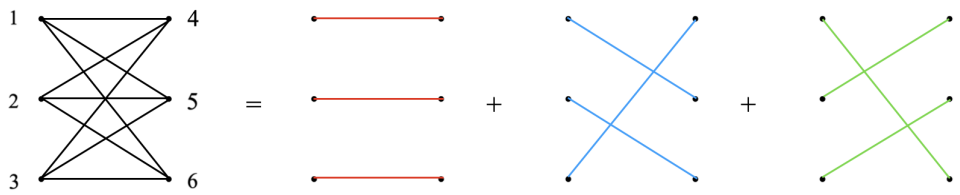


Figure 1: $N = 6,\ d = 3$

This is a d-sparse bipartite graph: Vertices in $\{1, 2, \cdots, N\}$ only have edges to $\{N+1, N+2, \cdots, 2N\}$, and vertices in $\{N+1, N+2, \cdots, 2N\}$ only have edges to $\{1, 2, \cdots, N\}$.

Note that we have:

$$\begin{aligned} e^{-i(X\otimes H)t} &= \begin{bmatrix} I & \\ & I \end{bmatrix} + (-it)\begin{bmatrix} & H \\ H & \end{bmatrix} + \frac{(-it)^2}{2!}\begin{bmatrix} H^2 & \\ & H^2 \end{bmatrix} + \cdots \\ &= I_2 \otimes \left(I_n + \frac{(-it)^2}{2!}H^2 + \frac{(-it)^4}{4!}H^4 + \cdots \right) + X \otimes \left(-itH + \frac{(-it)^3}{3!}H^3 + \cdots \right) \end{aligned}$$

In addition, $I_2|+\rangle = |+\rangle, X|+\rangle = |+\rangle$. Therefore, for any state $|\psi\rangle \in \mathbb{C}^N$

$$\begin{aligned} e^{-i(x\otimes H)t}|+\rangle|\psi\rangle &= |+\rangle \otimes \left(I_n + \frac{(-it)^2}{2!}H^2 + \cdots \right)|\psi\rangle + |+\rangle \otimes \left(-itH + \frac{(-it)^3}{3!}H^3 + \cdots \right)|\psi\rangle \\ &= |+\rangle \otimes e^{-iHt}|\psi\rangle. \end{aligned}$$

Conclusion: WLOG. we are assuming we're simulating a bipartite Hamiltonian H.

Observation 2. Suppose we are given:

- an undirected bipartite graph G, with N vertices and max degree d.
- we have a query oracle that can compute the neighbors of any vertex.

There is an edge coloring of G with at most d^2 colors, the color of an edge can be computed in $O(d)$ time.

Proof. For any vertex α, let $\mathrm{idx}(\alpha, \beta)$ denote the index of vertex β in the list of neighbors of α.

$$\begin{aligned} &\mathrm{idx}(1,4) = 1, \mathrm{idx}(1,5) = 2, \mathrm{idx}(1,6) = 3 \\ &\mathrm{idx}(5,1) = 1, \mathrm{idx}(5,2) = 2, \mathrm{idx}(5,3) = 3 \end{aligned}$$

Define the color of edge (α, β), where α is from the left and β is from the right, to be

$$d \cdot \mathrm{idx}(\alpha, \beta) + \mathrm{idx}(\beta, \alpha) - d.$$

Note that

$$\left.\begin{aligned} d \cdot \mathrm{idx}(\alpha, \beta) + \mathrm{idx}(\beta, \alpha) - d \geqslant d \cdot 1 + 1 - d \geqslant 1 \\ d \cdot \mathrm{idx}(\alpha, \beta) + \mathrm{idx}(\beta, \alpha) - d \leqslant d \cdot d + d - d \leqslant d^2 \end{aligned}\right\} \Rightarrow \text{ at most } d^2 \text{ colors}$$

Furthermore, if (α, β) and (α, δ) have the same coloring, then

$$\begin{aligned} & d \cdot \mathrm{idx}(\alpha, \beta) + \mathrm{idx}(\beta, \alpha) = d \cdot \mathrm{idx}(\alpha, \delta) + \mathrm{idx}(\delta, \alpha). \\ \Rightarrow \quad & d(\mathrm{idx}(\alpha, \beta) - \mathrm{idx}(\alpha, \delta)) = \mathrm{idx}(\delta, \alpha) - \mathrm{idx}(\beta, \alpha) \end{aligned}$$

$$\text{RHS} \in [-(d-1), d-1]. \text{ LHS: multiple of } \alpha \Rightarrow \begin{cases} \text{idx}(\alpha, \beta) = \text{idx}(a, \delta) \\ \text{idx}(\delta, \alpha) = \text{idx}(\beta, \alpha) \end{cases} \Rightarrow \beta = \delta \ .$$

□

Observation 3. A 1-sparse Hamiltonian H (for color c) can be efficiently simulated.

Proof. For vertex x, denote $v_c(x)$ to be its neighbor in H_c. Since H_c is 1-sparse, v_c is well-defined, and $v_c(v_c(x)) = x$.

Essentially, H_c is a directed-sum of 2-dimensional blocks.

By our assumption, we have an oracle $U_c |x, 0, 0\rangle = |x, v_c(x), H_{x,v_c(x)}\rangle$, and each query takes $O(d)$ cost to query to the original H.

$$\begin{cases} U_c |x, 0, 0\rangle = |x, v_c(x), H_{x,v_c(x)}\rangle \\ U_c |v_c(x), 0, 0\rangle = \left|v_c(x), v_c\left(v_c(x)\right), H_{v_c(x),}, v_c\left(v_c(x)\right)\right\rangle = |v_c(x), x, \bar{H}_{x,v_c(x)}\rangle \end{cases}$$

Based on the Solovay-Kitaev Theorem, time-evolution operator can be implemented with poly-cost.

Final algorithm: (Assuming input state is$\sum_i a_i |x_i\rangle$)

1. Apply U_c resulting in $\sum_i a_i |x_i\rangle |v_c(x_i)\rangle |H_{x,v_c(x_i)}\rangle$
2. Using value in the 3-rd register as a controlled rotation on the 1-st register.
3. Apply $U_c^\dagger$ to uncompute the junk.

□

Putting Observation 1,2,3 together, we get an efficient quantum algorithm for simulating sparse Hamiltonians.

2 Continuous-time Quantum Walk

Random walks come into two flavors: discrete-time (studied) and continuous-time.

Given an undirected simple (no self-loop, no parallel edges) graph $G = (V, E)$. Its Laplacian is

$$L_{jk} = \begin{cases} -\deg(j) & j = k \\ 1 & (j,k) \in E \\ 0 & \text{otherwise} \end{cases} \qquad \text{If weighted:} \qquad L_{jk} = \begin{cases} -\sum_{l \neq j} \omega_{jl} & j = k \\ \omega_{jk} & (j,k) \in E \\ 0 & \text{otherwise} \end{cases}$$

The classical random walk on G is defined as the solution of the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t} p_j(t) = \sum_{k \in V} L_{jk} p_k(t).$$

Here $p_j(t)$ denotes the probability associated with vertex j at time t. This can be viewed as a diffusion process. Note that

$$\frac{\mathrm{d}}{\mathrm{d}t} \sum_{j \in V} p_j(t) = \sum_{j,k \in V} L_{jk} p_k(t) = 0.$$

Since the columns/rows of L sum to 0.

Therefore, an initially normalized distribution in l_1-norm remains normalized $\Rightarrow$ the evolution of the classical random walk for any time t is a stochastic process.

In closed form: $p(t) = e^{Lt}p(0)$.

This is very similar to the Schrödinger equation: $i\frac{\mathrm{d}|\psi(t)\rangle}{\mathrm{d}t} = H|\psi\rangle \Rightarrow |\psi(t)\rangle = e^{-iHt}|\psi(0)\rangle$.

Take $H = L$, and denote the amplitudes $q_j(t) = \langle j|\psi(t)\rangle$, we obtain the equation to define quantum walk:

$$i\frac{\mathrm{d}}{\mathrm{d}t}q_j(t) = \sum_{k\in V} L_{jk}q_k(t)$$

Only difference: A factor of i. But this gives fundamental difference:

$$L\begin{cases}\text{row/column sum } 0 \Rightarrow e^{Lt} \text{ preserves } l_1\text{-norm of vector (stochastic matrix).}\\ \text{Hermitian } \Rightarrow e^{-iLt} \text{ preserves } l_2\text{-norm of vector (unitary matrix).}\end{cases}$$

Our focus: Continuous-time quantum walks using Hamiltonian represents graph structures.

2.1 Classical and quantum walks on the hybercube

Boolean hypercube: $V = \{0,1\}^n$ $E = \{(x,y) \in V^2 : H(x,y) = 1\}$, where $H(x,y)$ denotes the Hamming distance between the strings x and y.

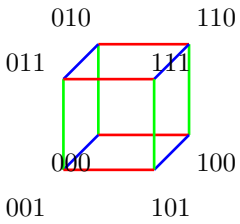
When $n = 1$, the hypercube is simply an edge:

Adjacency matrix: $\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} = X$.

For general n, the graph has adjacency matrix $A = \sum_{j=1}^{n} X^{(j)}$, where $X^{(j)}$ denotes the operator acting on X on the j^{th} qubit, and as I on other qubits.

Formally:

$$A = X\otimes I\otimes\cdots\otimes I + I\otimes X\otimes I\otimes\cdots\otimes I + \cdots + I\otimes\cdots\otimes I\otimes X$$



Fact: For n-regular graphs, $L = -nI + A$

$$e^{-iLt} = e^{-it(-nI+A)=e^{int}e^{-itA}}$$

In other words, up to a global phase, we can consider the continuous-time quantum walk with Hamiltonian given by the adjacency matrix for n-dim Boolean hypercube.

Note that $[X\otimes I, I\otimes X] = X\otimes X - X\otimes X = 0$. Therefore, $X^{(i)}$ and $X^{(j)}$ commute for any $i,j\in[n]$. Therefore:

$$e^{-itA} = \Pi_{j=1}^{n} e^{-iX^{(j)}t} = \bigotimes_{j=1}^{n}\begin{pmatrix}\cos t & i\sin t\\ -i\sin t & \cos t\end{pmatrix}$$

After time $t = \frac{\pi}{2}$, this is $\begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix}$ on each qubit.

$$\Rightarrow e^{-iA\frac{\pi}{2}} |0^n\rangle = \left(\begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right)^{\otimes n} = (-i)^n |1^n\rangle$$

If we measure in the computational basis, $\Pr[1^n] = 1$ (and all other probabilities $= 0$.)
We can generate this conclusion to the following two cases:

- (Different initial state) In fact, any computational basis state $|x\rangle$ is mapped to the state $|\overline{x}\rangle$ corresponds to the opposite vertex of the hypercube (up to a global phase).
- (Different evolution time) $t =$ multiple of $\frac{\pi}{2}$ will result in shifts between $|0^n\rangle$ and $|1^n\rangle$.

How about classical random walk on the hypercube? We can prove:

- It mixes to the uniform distribution $p_i = \frac{1}{2^n}$ $\forall i \in \{0,1\}^n$ in polynomial time (poly$(n, 1/\epsilon)$, where ϵ is the precision to the distribution. i.e., $\sum |p_i - \frac{1}{2^n} \leq \epsilon|$.
- The probability of reaching 1^n to 0^n is exponentially small in n at any time.

Conclusion: Classical and quantum walks can exhibit radically different behavior.
Next: A black-box problem can be solved exponentially faster by a quantum walk than any classical algorithm.

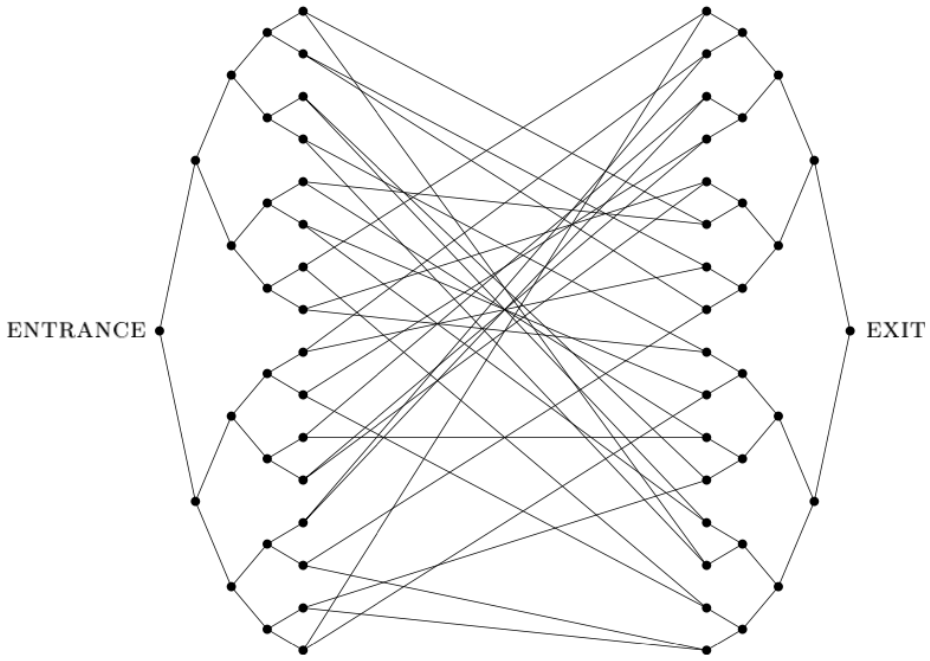


Figure 2: glued tree