CSCI5160 Approximation Algorithms — Spring 2022

Lecturer: Siu On Chan — *Scribe: Siu On Chan*

Notes 19 High-dimensional expander

1. Abstract simplicial complex

We want to apply the sampling algorithm based on random walk/Markov chain in the last lecture to other settings (spanning trees, d-paths, d-cliques, etc). Let us generalize those constructions.

Definition 1.1 (Abstract simplicial complex)**.** A set system $Y = (U, \mathcal{F})$ is a ground set U together with a family $\mathcal{F}$ of subsets over U. An abstract simplicial complex is downward closed set system: If $f \in \mathcal{F}$ and $g \subseteq f$, then $g \in \mathcal{F}$.

Abstract simplicial complex in combinatorics was originally proposed to describe the combinatorial structure of a (non-abstract) simplicial complex in algebraic topology. We need not worry about that motivation. Simply think of an abstract simplicial complex as a downward-closed set system.

Definition 1.2 (Level)**.** Level i of an abstract simplicial complex Y is the family of subsets of size i in Y, and is denoted $Y(i) = \{f \in \mathcal{F} \mid |f| = i\}$. The top level $Y(d)$ of Y is the non-empty level with the maximum d.

In the literature, $f \in \mathcal{F}$ of size i is also called a face of dimension $i - 1$. The collection of all such faces is denoted $X(i-1)$ (same as our $Y(i)$). I do not follow the standard terminology of "dimension", since this off-by-one is more confusing than helpful.

Definition 1.3 (Pure)**.** An abstract simplicial complex Y is pure if every face $f \in Y(i)$ is contained in some $g \in Y(d)$ in its top level.

Definition 1.4 (Weight)**.** Weight $w : Y(d) \to \mathbb{R}_+$ assigns positive weights to the maximal faces of a pure abstract simplicial complex Y. The weights induce a probability distribution π_d on $Y(d)$ by $\pi_d(f) = w(f)/\sum_{g \in Y(d)} w(g)$.

Random walk on pure abstract simplicial complex Y

> Let f_0 be an arbitrary face in the top level $Y(d)$
> For $t = 0, 1, 2, \ldots$
>   Remove an element from f_t uniformly at random to obtain $g_t \in Y(d-1)$
>   Among all $f_{t+1} \supset g_t$, pick the new $f_{t+1} \in Y(d)$ with probability proportional to $w(f_{t+1})$

This is a random walk/Markov chain on a weighted graph with vertex set $Y(d)$, and two nodes are adjacent if they share exactly $d-1$ elements.

An abstract simplicial complex $Y = (U, \mathcal{F})$ with top level $Y(d)$ represents a hypergraph, whose vertex set is U and whose set of hyperedges is $Y(d)$. When the top level is $Y(2)$, we get a graph (and weight is the usual edge weight).

From now on simply call the combinatorial set system a simplicial complex (without "abstract").

2. Inclusion graph

Definition 2.1 (Bipartite inclusion graph)**.** For $0 \leqslant k \leqslant d$, Γ_k has vertex set $Y(k) \cup Y(k-1)$. $t \in Y(k)$ is adjacent to $b \in Y(k-1)$ if $t \supset b$.

For every fixed $0 \leqslant k \leqslant d$, we again look at the random process:

(1) Pick a random t from $Y(d)$, then randomly remove all but k elements from t .

Denote by $\pi_k(f)$ the probability of obtaining $f \in Y(k)$.

The probability mass function π_k satisfies

$$\pi_k(f) = \frac{1}{k+1} \sum_{g \in Y(k+1), g \supset f} \pi_{k+1}(g) \ . \tag{2}$$

Since the set system is pure, every face at a lower level also gets positive probability mass.

As before, we will also look at $(d+1)$-partite inclusion graph $\Gamma_d \cup \cdots \cup \Gamma_0$.

3. Links

Recall Garland's method decomposed $\tilde{P}_k^\wedge$ into $\sum_b \tilde{P}_b^\wedge$ over $b \in Y(k-1)$, and $P_k^\vee = \sum_b P_b^\vee$. $\tilde{P}_b^\wedge$ corresponds to transitions in a weighted subgraph $H_b = (S_b, E_b)$, where

$$S_b = \{m \in Y(k) \mid m \supset b\} \qquad E_b = \{(m, m') \in S_b \times S_b \mid m \cup m' \in Y(k+1)\} .$$

In the literature, H_b is known as the 1-skeleton of the link of b:

Definition 3.1 (Link)**.** Given a simplicial complex $Y = (U, \mathcal{F})$ and a face $b \in \mathcal{F}$, the link of b is $Y_b = (U, \mathcal{F}_b)$, with faces

$$\mathcal{F}_b = \{f \setminus b \mid f \in \mathcal{F},\ f \supseteq b\} .$$

$\mathcal{F}_b$ consists of faces g that can extend b to remain in $\mathcal{F}$, so that $g \cup b \in \mathcal{F}$.
Every link Y_b in a pure simplicial complex Y is also a pure simplicial complex.

Definition 3.2 (Skeleton)**.** Given a simplicial complex $Y = (U, \mathcal{F})$, its k-skeleton $(U, \mathcal{F}_k)$ consists of faces in $\mathcal{F}$ of size at most $k+1$.

Think of $(U, \mathcal{F})$ as a hypergraph. 0- and 1-skeletons represent vertices and (non-hyper) edges.

Definition 3.3 (Link expander)**.** A pure simplicial complex Y with weight w is an α-link expander if $\lambda_2(\tilde{P}_b^\wedge) \leqslant \alpha$ for all $b \in Y(k-1)$ and $0 < k < d-1$.

In this convoluted language, the yet unproved lemma in last lecture becomes:

Lemma 3.4. *If Y is the pure simplicial complex of a matroid of rank d with uniform weight $w = \mathbb{1}$ on $Y(d)$, then Y is a 0-link expander.*

4. Spectral inequalities for transition matrices

Consider a weighted undirected graph with adjacency matrix A. Recall that its degree matrix Π is the diagonal matrix $\Pi = \mathrm{Diag}(\pi)$, whose vector π of diagonal entries encodes the vertex degrees:

$$\pi(v) = \deg(v) = \sum_u A(v, u) \qquad \text{for every } v .$$

In our application, A represents a distribution over edges, so that the sum of edge weights equal 1. In this case, one can easily show that the vector π represents the probability mass function of the following distribution:

(1) Pick an edge $e \in E$ according to the edge distribution A
(2) Uniformly select one of the two endpoints of e at random

Also, the random walk transition matrix is $P = \Pi^+ A$, where Π^+ represents the pseudo-inverse of Π. We use pseudo-inverse, as opposed to the usual inverse, because there may be isolated vertices, so that $\pi(v) = 0$. For $v \in \mathrm{supp}(\pi)$, $P(v, v)$ is simply $1/\pi(v)$. For $v \notin \mathrm{supp}(\pi)$, $P(v, v) = 0$.

Given two transition matrices P and Q with a common stationary distribution π and degree matrix $\Pi = \mathrm{Diag}(\pi)$, recall that the spectral comparison $\preccurlyeq_\Pi$ between P and Q is defined as

$$P \preccurlyeq_\Pi Q \qquad \Longleftrightarrow \qquad \Pi P \preccurlyeq \Pi Q .$$

Proposition 4.1. *Suppose a transition matrix P has a stationary distribution π and the corresponding degree matrix $\Pi = \mathrm{Diag}(\pi)$. Then for any $\gamma \leqslant 1$,*

$$\lambda_2(P) \leqslant \gamma \qquad \Longleftrightarrow \qquad P \preccurlyeq_\Pi \underbrace{\gamma I + (1-\gamma)\mathbb{1}\pi^\top}_{Q} .$$

The expression Q here represents a linear interpolation of two extremes: I and $\mathbb{1}\pi^\top$. This linear interpolation has two interpretations:

Independence/correlation interpretation:

- $\mathbb{1}\pi^\top$ is the transition matrix of a random walk that always picks the next vertex afresh from π, regardless of the current vertex
- I is the transition matrix of random walk that always stays at the current vertex

Spectral interpretation:

- $\mathbb{1}\pi^\top$ has only one nonzero left-eigenvector π, associated with the eigenvalue 1

- I not only has the same left-eigenvector π with eigenvalue 1, but also $n-1$ other left-eigenvectors associated with eigenvalue 1

Proof of Proposition 4.1. ($\Longleftarrow$): Q has top eigenvalue 1 (with left-eigenvector π) and all other eigenvalues γ. That means $\lambda_2(Q) = \gamma$ (since $\gamma \leqslant 1$). We proved in the last lecture notes that

$$P \preccurlyeq_\Pi Q \qquad \Longrightarrow \qquad \lambda_2(P) \leqslant \lambda_2(Q) \,,$$

so $\lambda_2(P) \leqslant \lambda_2(Q) = \gamma$.

($\Longrightarrow$): Consider the normalized adjacency matrix

$$\mathcal{A}_P = \sqrt{\Pi^+} A \sqrt{\Pi^+} = \sqrt{\Pi} P \sqrt{\Pi^+} \,. \tag{3}$$

$\sqrt{\pi}$ is an eigenvector of $\mathcal{A}_P$ with eigenvalue 1. Fix any set B eigenbasis of $\mathcal{A}$ associated with all the non-zero eigenvalues, so that B contains $\sqrt{\pi}$. (Here and below, we call a set of (left-)eigenvectors a (left-)eigenbasis if it spans the row space of the matrix, even without spanning the whole vector space.)

Let $\mathcal{A}_Q = \sqrt{\Pi} Q \sqrt{\Pi^+}$ be the normalized adjacency matrix for Q. We claim that B is also an eigenbasis for $\mathcal{A}_Q$. Indeed, Eq. (3) implies that $B' \stackrel{\text{def}}{=} \sqrt{\Pi} B$ is a left-eigenbasis of P, with the same eigenvalues, as we proved in the lecture on random walks. A vector in B' is π. So B' is a set of common left-eigenvectors of P, $\mathbb{1}\pi^\top$, and hence Q. Therefore $B = \sqrt{\Pi^+} B'$ is also an eigenbasis of $\mathcal{A}_Q$, with the same associated eigenvalues as the corresponding left-eigenvectors of Q in B'.

Since $\lambda_2(P) \leqslant \gamma = \lambda_2(Q)$, for every $y \in B$,

$$y^\top \mathcal{A}_P y \leqslant y^\top \mathcal{A}_Q y \,.$$

Indeed, the left hand side equals $\lambda_k(\mathcal{A}_P)\|y\|^2$, and the right hand side equals $\lambda_k(\mathcal{A}_Q)\|y\|^2$. Since this inequality holds for every vector in a common eigenbasis of $\mathcal{A}_P$ and $\mathcal{A}_Q$, it also holds for every vector. That means $\mathcal{A}_P \preccurlyeq \mathcal{A}_Q$, which is equivalent to $P \preccurlyeq_\Pi Q$. ☐

5. Oppenheim's theorem

Oppenheim found a way to translate eigenvalue bound on a higher layer links to that of a lower layer.

Theorem 5.1 (Oppenheim)**.** *Let Y be a pure simplicial complex with weight w. Suppose $\lambda_2(\tilde{P}_b^\wedge) \leqslant \alpha$ for every $b \in Y(1)$. Also, suppose its 1-skeleton graph $H = (Y(1), Y(2))$ is connected. Then H is an $\dfrac{\alpha}{1-\alpha}$-expander. Equivalently, the random walk P on H satisfies $\lambda_2(P) \leqslant \dfrac{\alpha}{1-\alpha}$.*

Applying Oppenheim's theorem inductively, we get:

Corollary 5.2. *Let Y be a pure simplicial complex with weight w. Suppose every link Y_b has a connected 1-skeleton graph. Also, suppose the 1-skeleton graph of every $b \in Y(d-2)$ is an α-expander. Then Y is an $\dfrac{\alpha}{1-(d-1)\alpha}$ expander.*

Before proving Oppenheim's Theorem 5.1, we first sketch the reasons that the hypotheses of the previous theorem holds for the matroid with uniform weight at the top level.

That every link Y_b is connected is due to the exchangable property of matroid (details omitted).

Given any $b \in Y(n-3)$, the 1-skeleton $H_b = (S_b, E_b)$ of Y_b has adjacency matrix

$$A_b(f, f') = \begin{cases} 1 & \text{if } b \cup f \cup f' \text{ is a spanning tree} \\ 0 & \text{otherwisew} \end{cases} .$$

Edges in b induces three connected components in G. Adding two more edges to these components yields a spanning tree, provided the two edges added are connecting different pairs of components. This partitions the 0-skeleton S_b of Y_b into three sets E_1, E_2, E_3.

The adjacency matrix A_b is of the form

$$A_b = \quad \begin{matrix} & \begin{matrix} E_1 & E_2 & E_2 \end{matrix} \\ \begin{matrix} E_1 \\ E_2 \\ E_3 \end{matrix} & \begin{pmatrix} O & \mathbb{1} & \mathbb{1} \\ \mathbb{1} & O & \mathbb{1} \\ \mathbb{1} & \mathbb{1} & O \end{pmatrix} \end{matrix} \qquad = J - \mathbb{1}_{E_1}\mathbb{1}_{E_1}^\top - \mathbb{1}_{E_2}\mathbb{1}_{E_2}^\top - \mathbb{1}_{E_3}\mathbb{1}_{E_3}^\top \,.$$

Here J denotes the all-one matrix of appropriate dimension.

The all-one matrix J on S_b has rank 1 and nonpositive second eigenvalue, so after subtracting three positive semidefinite matrices $\mathbb{1}_{E_i}\mathbb{1}_{E_i}^\top$ from J, A_b also has nonpositive second eigenvalue by Courant–Fishcer.

Therefore the normalized adjacency matrix of Y_b also has nonpositive second eigenvalue.

Proof of Theorem 5.1. Suppose someone picks $T \in Y(d)$ according to π_d and reveals its elements $T = \{e_1, \dots, e_d\}$ one by one to you (by picking a uniformly random ordering).

Scenario 1: Consider the first two elements revealed. Then the probability of $(e_1, e_2) = (f, g)$ is exactly the (f, g)-entry of the matrix

$$\Pi_1 P \,,$$

where $\Pi_1 = \mathrm{Diag}(\pi_1)$ and P is the transition matrix for random walk on the 1-skeleton graph of Y:

$$P(f, g) = \frac{\mathbb{P}[g \in T]}{\mathbb{P}[f \in T]} \,.$$

Scenario 2: This time consider the first three elements revealed. Then the probability of $(e_2, e_3) = (f, g)$ is exactly the (f, g)-entry of the matrix

$$\mathop{\mathbb{E}}_{a \sim \pi_1} \Pi_a \tilde{P}_a^\wedge \,,$$

where $\Pi_a = \mathrm{Diag}(\pi_a)$ (π_a is the conditional distribution of f conditioned on a) and $\tilde{P}_a^\wedge$ is the transition matrix for random walk on the the 1-skeleton graph of the link Y_a:

$$\tilde{P}_a^\wedge(f, g) = \frac{\mathbb{P}[g \in T \mid a \in T, f \in T]}{\mathbb{P}[f \in T \mid a \in T]} \,.$$

In fact, the ordered tuple (f, g) revealed in both scenarios above have the same marginal probabilities. We have thus proved

$$\Pi_1 P = \mathop{\mathbb{E}}_{a \sim \pi_1} \Pi_a \tilde{P}_a^\wedge \,. \tag{4}$$

Note that $\Pi_a P_a^\vee = \pi_a \pi_a^\top$ is a rank-1 matrix, and the assumption that $\lambda_2(\tilde{P}_a^\wedge) \leqslant \alpha$ is equivalent (by Proposition 4.1) to

$$\Pi_a \tilde{P}_a^\wedge \preccurlyeq \Pi_a(\alpha I + (1 - \alpha) P_a^\vee) \,.$$

Plugging the last inequality into Eq. (4), we get

$$\begin{aligned} \Pi_1 P &\preccurlyeq \mathop{\mathbb{E}}_{a \sim \pi_1} \Pi_a(\alpha I + (1 - \alpha) P_a^\vee) \\ &= \underbrace{\alpha \Pi_1}_{(*)} + \underbrace{(1 - \alpha) \Pi_1 P^2}_{(**)} \,, \end{aligned} \tag{5}$$

where the last equality will be explained next. The first term $(*)$ is due to

$$\mathop{\mathbb{E}}_{a \sim \pi_1} \Pi_a = \Pi_1 \,,$$

which is true because the second element revealed (e_2) has the same marginal distribution as the first (e_1). The second term $(**)$ is due to

$$\mathop{\mathbb{E}}_{a \sim \pi_1} \Pi_a P_a^\vee = \Pi_1 P^2 \,, \tag{6}$$

which holds because the (f, g)-entry of the matrix on the left equals the probability of:

$$\text{Sample } a \sim \pi_1 \text{, then independently sample } f \sim \pi_a \text{ and } g \sim \pi_a \,,$$

and the (f, g)-entry of the matrix on the right equals the probability of:

$$\text{Sample } f \sim \pi_1 \text{, then } a \sim P(f, \cdot) \text{, then } g \sim P(a, \cdot) \,.$$

It is possible to show that P has only non-negative eigenvalues (it is a non-lazy up-walk and is the product of two self-adjoint operators), and $\lambda_2(P^2) = \lambda_2(P)^2$, so Eq. (5) implies (via essentially the same proof as the first part of Proposition 4.1)

$$\lambda_2 \leqslant \alpha + (1 - \alpha)\lambda_2^2 \qquad \iff \qquad \lambda_2 - \lambda_2^2 \leqslant \alpha(1 - \lambda_2^2) \,,$$

where $\lambda_2 = \lambda_2(P)$. The assumption that the empty link is connected means $\lambda_2 < 1$. Divide both sides by $1 - \lambda_2$ to get

$$\lambda_2 \leqslant \alpha(1 + \lambda_2) \qquad \Longrightarrow \qquad \lambda_2 \leqslant \frac{\alpha}{1 - \alpha}\ . \qquad \square$$