

Lecture 6

Lecturer: Irit Dinur

Scribe: Elad Eban

1 Lecture Overview

This lecture has two parts, both of which deal with expanders. In the first part we analyze the fraction of paths of length t which stays solely in the boundaries of a subset $S \subset V$. We shall see both a lower and upper bound for this fraction and give a proof only for the later.

In the second part we shall define the celebrated **zigzag** product of graphs and explore some properties of this operation. The main point about the zigzag product is that it is a straightforward explicit construction of expanders (possibly in future lectures we shall see that this construction is **very** explicit).

2 Hardness Amplification

We start by stating a theorem that follows directly from the PCP theorem (as done in the previous lecture)

Theorem 1 *There exists $\alpha < 1$ and $\epsilon > 0$ for which the gap problem of deciding, for a graph G of size n , between the two options:*

1. $\omega(G) < \alpha n$
2. $\omega(G) > (1 + \epsilon)\alpha n$

is NP hard.

The purpose of this section is to prove the following stronger theorem.

Theorem 2 *There exist $0 < c_1 < c_2 < 1$ for which the gap problem of deciding, for a graph H of size m , between the two options:*

1. $\omega(G) < m^{c_1}$
2. $\omega(G) > m^{c_2}$

is NP hard.

As an immediate corollary we obtain,

Corollary 3 *There is a constant $\delta > 0$ such that no polynomial-time algorithm that approximates the maximum-clique in a graph $G = (V, E)$ to within factor $|V|^\delta$, unless $P = NP$.*

Proof (of theorem) We shall prove Theorem 2 using Theorem 1. Given a graph G of size n input for the problem of theorem 1, we want to construct a graph H which will hold theorem 2.

Let us set $t \in \mathbb{N}$, and X a (n, d, λ) expander. We consider the vertices X and G as having the same labels.

$$V(H) \stackrel{\text{def}}{=} \{(v_0, v_1, \dots, v_t) \mid (v_0, v_1, \dots, v_t) \text{ is a legal walk in } X\}$$

$$E(H) \stackrel{\text{def}}{=} \left\{ ((v_0, v_1, \dots, v_t), (u_0, u_1, \dots, u_t)) \mid \{u_i\}_{i=0}^t \cup \{v_i\}_{i=0}^t \text{ is a clique in } G \right\}$$

What is the size of $V(H)$? Well this is fairly simple, a walk in X can start from any of its n vertices, for each further step there are exactly d adjacent vertices the walk can choose (going back is allowed). Thus, we conclude that $m \stackrel{\text{def}}{=} |V(H)| = nd^t$.

Claim 4 *There exists constants $\beta_1 = \alpha(1 + \frac{\epsilon}{4})$ and $\beta_2 = \alpha(1 + \frac{3\epsilon}{4}) > \beta_1$ for which:*

1. *If $\omega(G) < \alpha n$ then $\omega(H) < \beta_1^t m$*
2. *If $\omega(G) > \alpha(1 + \epsilon)n$ then $\omega(H) > \beta_2^t m$*

It is easy to see that this claim concludes the proof. We set $t = \log(n)$ so $m = nd^{\log(n)} = n^{1+\log(d)}$ or in another way $n = m^\delta$ for some constant δ . Substituting this in the first part of the claim:

$$\omega(H) < \beta_1^t m = n^{\log(\beta_1)} m = m^{1+\delta \log(\beta_1)} \stackrel{\text{def}}{=} m^{c_1}$$

And in the same way for the second part:

$$\omega(H) > \beta_2^t m = n^{\log(\beta_2)} m = m^{1+\delta \log(\beta_2)} \stackrel{\text{def}}{=} m^{c_2}$$

■

Proof of Claim 4: Let A be a maximal clique in H then it is immediate that the set

$$S = \bigcup_{(v_0, v_1, \dots, v_t) \in A} \{v_0, v_1, \dots, v_t\} \quad (1)$$

is a clique in G . We assume that $|S| < \alpha n$, and bound from above the value of $|A|$.

We first claim that A contains all the paths of length $t + 1$ which use only vertices from S . This holds since any two paths which use vertices from S are connected when we considered as vertices in H , thus these paths form a clique in H . On the other hand, a walk corresponding to a vertex in A uses only vertices from S (from the definition of S). Thus we can write

$$A = \{(v_0, v_1, \dots, v_t) \in V(H) \mid \forall i : v_i \in S\}.$$

This leads us to a more general and interesting problem of bounding, in an expander graph X , the number of paths of length $t + 1$ that stay inside a given set $S \subset V(X)$.

Lemma 5 *Let $X = (V, E)$ be an expander, and let $S \subset V$. We define $\mu = \frac{|S|}{n}$. The number of paths of length $t + 1$ which remain in S (denoted $p(S, t)$) is bounded by:*

$$\mu n d^t (\mu + \frac{\lambda_{n-1}}{d} (1 - \mu))^t \leq p(S, t) \leq \mu n d^t (\mu + \frac{\lambda_1}{d} (1 - \mu))$$

where $\lambda_0 > \lambda_1 > \dots \lambda_{n-1}$ are the eigenvalues of the adjacency matrix of X .

This lemma is just what we need in order to prove the claim: Assuming $\lambda(X) = \frac{\alpha\epsilon}{4}$.

If $\omega(G) < \alpha n$ then $\mu < \alpha$ and making some substitutions:

$$\omega(H) = |A| = p(S, t) \leq m \mu (\mu + (1 - \mu) \lambda / \mu)^t < m \alpha (\alpha + \frac{\alpha\epsilon}{4})^t \leq m (\beta_1)^t$$

for $\beta_1 = \alpha(1 + \frac{\epsilon}{4})$.

For the second case, if $\omega(G) > \alpha(1 + \epsilon)n$ then let S be a maximum clique in G and now $\mu = \frac{|S|}{n} > \alpha(1 + \epsilon)$. Also let A be a clique in H defined by Equation (1). We have (assuming $\lambda_{n-1} < 0$),

$$\omega(H) \geq |A| \geq m \cdot \mu (\mu + \frac{\lambda_{n-1}}{d} (1 - \mu))^t \geq m \mu \cdot (\mu - \frac{\epsilon\alpha}{4})^t \geq m \cdot \alpha \cdot (\alpha(1 + \frac{3\epsilon}{4}))^t \geq m (\beta_2)^t$$

for $\beta_2 = \alpha(1 + \frac{3\epsilon}{4})$.

■

We will prove only the first part of Lemma 5, the second is left as an exercise in problem set #2.

Proof of Lemma 5: Let A be the normalized adjacency matrix of X . We define P to be the projection of the subspace spanned by $v_i \in S$. That is $P_{i,j} = \begin{cases} 1 & i = j \in S \\ 0 & \text{otherwise} \end{cases}$. Additionally, we define $N = PAP$ and notice that N results from A by erasing all rows and columns corresponding to vertices outside S .

Claim 6 For each t and for each i it holds that $(N^t Pu)_i$ is the probability of a $t+1$ length walk, to end in v_i and to remain in S for all its length.

Proof of Claim 6: If $v_i \notin S$ then it is clear that both the probability and the i 'th entry of the vector are 0. Now assume $i \in S$, we prove the claim by induction.

For $t = 0$, $(Pu)_i = 1/n$ which is exactly the probability for a one-vertex-walk to equal v_i .

The probability for a walk to end up on v_i after $t+1$ steps is exactly the sum over all vertices $v_j \in S$ adjacent to v_i of the probability to reach v_j after t steps and then move to v_i in the last step. Assume by induction for t , we denote $y = N^t Pu$.

$$(Ny)_i = \sum_j N_{i,j} y_j = \sum_{j: (v_j, v_i) \in E} N_{i,j} y_j$$

which is exactly the sum we are looking for. ■

Claim 7 Let γ be the largest eigenvalue of N . Then $\gamma \leq \mu + \frac{\lambda_1}{d}(1 - \mu)$.

Proof of Claim 7: In order to prove this claim we use the Rayleigh (hard to spell) quotient, which gives us an explicit formula for γ .

$$\gamma = \max_{y \neq 0} \frac{\langle y, Ny \rangle}{\langle y, y \rangle}$$

Let y be a vector that maximizes the above expression, and note that for all $i \notin S$ $y_i = 0$ (otherwise by zeroing out these coordinates in y we would get a higher Rayleigh quotient). Now we decompose y in respect to the eigenvectors of A . We write $y = y^0 + y^1$ where y^0 is y 's projection on u (the uniform vector is an eigenvector of A and y^1 is some vector in the subspace orthogonal to u). Next, note that

$$\langle y, Ny \rangle = \langle y, Ay \rangle$$

holds since as previously noted N is identical to A restricted to S and $y_i = 0$ for all $i \notin S$.

$$\langle y, Ay \rangle = \langle y^0 + y^1, y^0 + Ay^1 \rangle = \langle y^0, y^0 \rangle + \langle y^1, Ay^1 \rangle$$

where the first equality follows from the fact that y^0 is an eigenvector of A with eigenvalue 1; and the second equality follows from the fact that y^0 is orthogonal to both y^1 and Ay^1 .

We recall that $\frac{\lambda}{d}$ is the second largest eigenvalue of A . The above expression is bounded by:

$$\langle y, Ny \rangle \leq \|y^0\|^2 + \frac{\lambda}{d} \|y^1\|^2 = \|y^0\|^2 \left(1 - \frac{\lambda}{d}\right) + \frac{\lambda}{d} (\|y^1\|^2 + \|y^0\|^2)$$

From Pythagoras' Theorem we know that $\|y\|^2 = \|y^1\|^2 + \|y^0\|^2$ and so:

$$\langle y, Ny \rangle \leq \|y^0\|^2 \left(1 - \frac{\lambda}{d}\right) + \frac{\lambda}{d} \|y\|^2 \quad (2)$$

We recall that the y^0 is the projection of y on u , and use the Cauchy Schwartz inequality:

$$\|y^0\|^2 = \left\langle u, \frac{\bar{1}}{\sqrt{n}} \right\rangle^2 = \frac{1}{n} \langle y, \bar{1}_S \rangle^2 \leq \frac{1}{n} \|y\|^2 |S| = \mu \|y\|^2 \quad (3)$$

Plugging this result in equation 2:

$$\langle y, Ny \rangle \leq \|y\|^2 \left(\frac{\lambda}{d} + \mu \left(1 - \frac{\lambda}{d} \right) \right)$$

And we simply get:

$$\gamma = \frac{\langle y, Ny \rangle}{\langle y, y \rangle} \leq \frac{\|y\|^2 \left(\frac{\lambda}{d} + \mu \left(1 - \frac{\lambda}{d} \right) \right)}{\|y\|^2} = \frac{\lambda}{d} + \mu - \mu \frac{\lambda}{d}$$

■

Let us get back to the proof of Lemma 5. We start by an application of claim 6 :

$$Pr(\text{A walk of length } t+1 \text{ remains in } S) = \sum_i N^t P u = \langle N^t P u, \bar{1} \rangle = \sum_i N^t P u = \langle N^t P u, \bar{1}_S \rangle$$

Now Using Cauchy Schwartz inequality:

$$\leq \|N^t P u\| \|\bar{1}_S\| \leq \gamma^t \|P u\| \|\bar{1}_S\| = \gamma^t \sqrt{\frac{|S|}{n^2}} \sqrt{|S|} = \gamma^t \frac{|S|}{n} = \gamma^t \mu$$

And finally :

$$P(S, t) = |\{t+1 \text{ length paths in } X\}| Pr(\text{A } t+1 \text{ length walk which remains in } S)$$

and thus:

$$\begin{aligned} P(S, t) &= n d^t \gamma^t \mu \implies \\ P(S, t) &\leq \mu n d^t \left(\mu + \frac{\lambda_1}{d} (1 - \mu) \right) \end{aligned}$$

follows from Claim 7. ■

3 The Zigzag Product

The zigzag product of a graph comes to answer the very interesting problem of explicit construction of expanders. Some other explicit constructions exist for expanders but, the beauty of zigzag is in its simplicity.

In this section for simplicity we let $1 = \lambda_0 > \lambda_1 \dots$ be the eigenvalues of the *normalized* adjacency matrix. Also, an (n, d, λ) -graph is a d -regular graph on n vertices, whose normalized second largest eigenvalue (in absolute value) is at most λ .

Let G be a (n, d, λ) graph, we can consider G^2 which is a (n, d^2, λ^2) graph. We recall that the k power of a graph G is the graph of all paths of length exactly k from G . This nicely coincide with raising the adjacency matrix of G by the power of k . The graph G^2 has an appealing expansion but a much larger degree.

Definition 8 Let G be (n, d_1, λ_1) and H a (d_1, d_2, λ_2) graphs. The zigzag product $G \mathbin{\text{\textcircled{Z}}} H$ is defined in the following way:

- $V(G \mathbin{\text{\textcircled{Z}}} H) = \{(v, k) : v \in G, k \in H\}$
- $((u, i), (v, j)) \in E(G \mathbin{\text{\textcircled{Z}}} H)$ if there exists i', j' such that $(i, i'), (j, j') \in E(H)$, v is the j' th neighbor of u and u is the i' th neighbor of v .

Theorem 9 [RVW] Let G be (n, d_1, λ_1) and H a (d_1, d_2, λ_2) then $G \textcircled{Z} H$ is a $(nd_1, d_2^2, f(\lambda_1, \lambda_2))$ graph. And $f(\lambda_1, \lambda_2) \leq \lambda_1 + \lambda_2 + \lambda_2^2$

Corollary 10 Let H be a (D^4, D, λ_0) graph for $\lambda_0 < \frac{1}{5}$. We define the family of graphs G_i in the following way:

- $G_1 = H^2$
- $G_{i+1} = G_i^2 \textcircled{Z} H$

Then G_i is a (D^{4i}, D^2, λ) graph for $\lambda < \frac{2}{5}$

Proof of Corollary 10: We prove the corollary by induction on i . The case for G_1 is clear from the properties of H . Let us assume the claim for G_i and analyze G_{i+1} .

By definition of \textcircled{Z} we know that G_{i+1} is $(D^{4i}D^4, (D^2)^2, f(\lambda(G_i), \lambda_0))$ clearly both the size of the graph and its degree satisfies the corollary. What is left to check is the value of $\lambda(G_{i+1})$.

$$\lambda(G_{i+1}) = f(\lambda(G_i^2), \lambda_0) \leq (\lambda(G_i)^2 + \lambda_0 + (\lambda_0)^2) \leq \frac{4}{25} + \frac{1}{5} + \frac{1}{25} = \frac{10}{25} = \frac{2}{5}$$

As claimed! ■