(67902) Topics in Theory and Complexity Nov 2, 2006

Lecture 7

Lecturer: Irit Dinur *Scribe: Rani Lekach*

1 Lecture overview

This Lecture consists of two parts. In the first part we will refresh the definition of the zigzag product of graphs and the expansion property of these graph. We will also see another operation on graphs which is the replacement graph.
In the second part we will introduce the notion of probabilistically checkable proofs, define the class PCP, and make some easy remarks.

2 ZigZag

Recall we have defined the zigzag product between two graphs as follows.

Definition 1 *Let G_1 be a $(n, d_1, \alpha) - graph$ and H_1 be a $(d_1, d_2, \beta) - graph$. Let $\{1, .., d_1\}$ be the numbering of vertices in H, and for each vertex $v \in G$ fix an ordering of v's neighbors. We define the zigzag product $G \textcircled{z} H$ as follows:*

- $V(G \textcircled{z} H) = V(G) \times V(H)$
- $((v, i), (u, j)) \in E(G \textcircled{z} H)$ *if exists i', j' such that:*
  - $(i, i') \in E(H)$ *and u is the $i' - th$ neighbor of v in $V(G)$*
  - $(j, j') \in E(H)$ *and v is the $j' - th$ neighbor of u in $V(G)$*
  - *this means that $e_v^{i'} = e_u^{j'}$*

Remark *A way to look at this graph is as taking the graph G and replacing each vertex $v \in V(G)$ with a "cloud" that is the graph H (thus creating the vertices $(v, i) \forall i \in V(H)$), we will refer to this "cloud" as H_v. Now, with respect to the edges, we can look at the edges of each instance of H and refer to them as "in-cloud" edges. from each cloud there are "between-clouds" edges that are connected according to the edges of G and the ordering of the edges for each vertex in G. (This is actually the replacement-product graph edges - see below). This way, the edges of the zigzag-product graph are edges that connect vertices that have a "in-cloud" $\rightarrow$ "between-clouds" $\rightarrow$ "in-cloud" path between them.*

The ZigZag product allows us to create a family of expander graphs as we have seen in the previous lecture. We will now see a proof of the theorem that states this.

Theorem 2 *Let G_1 be a (n, d_1, α)-graph and H_1 be a (d_1, d_2, β)-graph.*
Then $G \textcircled{z} H$ is a $(n \cdot d_1, d_2^2, \gamma) - graph$ where $\gamma \leq \alpha + \beta + \beta^2$.

Proof

- The size and the degree of the graph are trivial and derive from the construction.
- Proving the bound on γ.
  Proof Idea: We will see that the normalized adjacency matrix of $G \textcircled{z} H$ can be expressed with the adjacency matices of G and H. We will then use this in order to bound the eigenvalue of that matrix.

Let A, B be the normalized adjacency matrices of G and H respectively. A describes a step in a random walk on G and B describes a step in a random walk on H

A single step in $G(z)H$ can be thought of as three steps:
Zig - a step in the "cloud" of H_v - can be described as a step in H.
Zag - a step in the graph G (going from H_v to some other cloud $H_{v'}$) -can be described as a step in G.
Zig - a step in the "cloud" of $H_{v'}$ - another step that can be described as a step in H.

Let $Z_{n\cdot d_1 \times n\cdot d_1}$ be the normalized adjacency matrix of $G(z)H$. we would like to know what does Z look like with respect to A and B.

Observation 3 *We can think of Z as a $n \times n$ block matrix where each block is of size $d_1 \times d_1$. This way, the first and last steps are limited to steps within the respective blocks of the vertices in the diagonal blocks of the matrix.*

Definition 4 *The tensor product of two matices A and B is the matrix $C = A \otimes B$. This matrix can be thought of as a block matrix, in which the i, j block is $A_{i,j}B$:*

$$A \otimes B = \begin{pmatrix} A_{1,1}B, & A_{1,2}B, & \cdots \\ \vdots & \vdots & \vdots \\ A_{n,1}B, & A_{n,2}B, & \cdots \end{pmatrix}$$

We define $\tilde{B} = I_n \otimes B$, to be the matrix that describes the the steps within the "clouds". We define $\tilde{A}$ to be the following $nd_1 \times nd_1$ matrix defined by

$$A_{(u,i),(v,j)} = 1 \quad iff \quad u \text{ is the } j\text{-th neighbor of } v \text{ and } v \text{ is the } i\text{-th neighbor of } u$$

and $A_{(u,i),(v,j)} = 0$ otherwise.

It is immediate that $\tilde{A}$ has a single 1 in each row, and since it is a symmetric matrix it holds that in each column there is a single 1. Thus, $\tilde{A}$ is a permutaion matrix, meaninig that the second step (Zag) is deterministic. We can now represent Z in the following manner

$$Z = \tilde{B}\tilde{A}\tilde{B}$$

Now, let γ be the second largest eigenvalue of Z. Let $x \perp \mathbf{1}$, and we can use the Rayleigh quotient to obtain:

$$\gamma = \max_{x \perp \mathbf{1}, x \neq 0} \frac{\langle x, Zx \rangle}{\langle x, x \rangle}$$

We will represent x in the following manner: $x = x_0 + x_{\mathbf{1}}$ where $x_0 \perp \mathbf{1}$ and for each cloud its values are identical.

$$x_0 = \begin{pmatrix} a_{v_1} \\ \vdots \\ a_{v_1} \\ \text{—} \\ a_{v_2} \\ \vdots \\ a_{v_2} \\ \text{—} \\ \vdots \end{pmatrix}$$

We can construct this vector by letting a_{v_i} to be the average value x on each cloud. The vector x_1 is a vector that is orthogonal to $\mathbf{1}$, and such that on each cloud it is orthogonal to the uniform vector on that cloud.

We now examine the Rayleigh quotient:

$$\langle Zx, x\rangle = \left\langle \tilde{B}\tilde{A}\tilde{B}x, x\right\rangle$$

Since $\tilde{B}$ is a symmetric and real matrix:

$$\begin{aligned} &= \left\langle \tilde{A}\tilde{B}x, \tilde{B}x\right\rangle \\ &= \left\langle \tilde{A}\tilde{B}(x_0+x_1), \tilde{B}(x_0+x_1)\right\rangle \\ &= \left\langle \tilde{A}\tilde{B}x_0 + \tilde{A}\tilde{B}x_1, \tilde{B}x_0 + \tilde{B}x_1\right\rangle \end{aligned}$$

Since each segment of x_0 is orthogonal to $\mathbf{1}_{d_1}$ then x_0 is invariant on $\tilde{B}$, i.e., $\tilde{B}x_0 = x_0$. Thus,

$$= \left\langle \tilde{A}(x_0 + \tilde{B}x_1), x_0 + \tilde{B}x_1\right\rangle$$

$$= \left\langle \tilde{A}x_0, x_0\right\rangle + \left\langle \tilde{A}\tilde{B}x_1, x_0\right\rangle + \left\langle \tilde{A}x_0, \tilde{B}x_1\right\rangle + \left\langle \tilde{A}\tilde{B}x_1, \tilde{B}x_1\right\rangle = (*)$$

According to the definition of x_0, there is a vector $\omega \in \mathbb{R}^n, \omega \perp \mathbf{1}_n$ such that $x_0 = \omega \otimes \mathbf{1}_{d_1}$ (the vector $\mathbf{1}$ with d_1 entries. This way:

$$\left\langle \tilde{A}x_0, x_0\right\rangle = \sum_{(v,i)\sim(u,j)} x_{0_{(v,i)}} x_{0_{(u,j)}} = \sum_{u\sim v} \omega_u\omega_v = d_1 \langle A\omega, \omega\rangle \le d_1\alpha \langle \omega, \omega\rangle \le \alpha \langle x_0, x_0\rangle$$

For the remaining three terms, let us observe that $\left\|\tilde{B}x_1\right\| \le \beta \|x_1\|$. This follows since in each cloud, the vector x_1 is orthogonal to the uniform distribution on that cloud. Let $y = \tilde{B}x_1$. By the Cauchy-Schwartz inequality we get

$$\left\langle \tilde{A}\tilde{B}x_1, \tilde{B}x_1\right\rangle = \left\langle \tilde{A}y, y\right\rangle \le \|y\|^2 \le \left\|\tilde{B}x_1\right\|^2 \le \beta^2 \|x\|^2 .$$

Next observe that $\tilde{A}$ cannot increase norm, so $\left\|\tilde{A}\tilde{B}x_1\right\| \le \left\|\tilde{B}x_1\right\| \le \beta\|x_1\|$. Now, using the Cauchy-Schwartz inequality it also holds that

$$\left\langle \tilde{A}\tilde{B}x_1, x_0\right\rangle \le \left\|\tilde{A}\tilde{B}x_1\right\| \|x_0\| \le \beta \|x_1\| \|x_0\|$$

and similarly,

$$\left\langle \tilde{A}x_0, \tilde{B}x_1\right\rangle \le \beta \|x_0\| \|x_1\|$$

Now we can get back to where we left:

$$\begin{aligned} (*) &\le \alpha \|x_0\|^2 + \beta \|x_1\| \|x_0\| + \beta \|x_0\| \|x_1\| + \beta^2 \|x_1\|^2 \\ &= \alpha \|x_0\|^2 + 2\beta \|x_1\| \|x_0\| + \beta^2 \|x_1\|^2 = (**) \end{aligned}$$

Now, Since $x_0 \perp x_1$ and $(\|x_0\| + \|x_1\|)^2 > 0$:

$$2\beta \|x_1\| \|x_0\| = \beta((\|x_0\|^2 + \|x_1\|^2) - (\|x_0\| + \|x_1\|)^2) = \beta(\|x\|^2 - (\|x_0\| + \|x_1\|)^2) \le \beta \|x\|^2$$

And since we can always increase the norms in the the equation to be $\|x\|$ we get that

$$(**) \leq (\alpha + \beta + \beta^2) \|x\|^2$$

This concludes the proof since we have shown that:

$$\langle Zx, x \rangle \leq (\alpha + \beta + \beta^2) \|x\|^2$$

meaning that

$$\gamma = \max_{x \perp \mathbf{1}, x \neq 0} \frac{\langle x, Zx \rangle}{\langle x, x \rangle} \leq \frac{(\alpha + \beta + \beta^2) \|x\|^2}{\|x\|^2} = \alpha + \beta + \beta^2$$

■

Corollary 5 *If α and β are small then so is γ. Meaning that if G and H are expanders so is $G ⓩ H$.*

This theorem can be used to prove that the replacement graphs (metioned above) are also expanders. We will start with the definition of the replacement product and then see that these are also expanders.

Definition 6 *Let G_1 be a (n, d_1, α)-graph and H_1 be a (d_1, d_2, β)-graph. We define the replacement product $G ⓡ H$ as follows:*

- $V(G ⓡ H) = V(G) \times V(H)$
- $E(G ⓡ H) = E_1 \cup E_2$ *where*
  - $E_1 = \{ ((v, i), (v, j)) \mid (i, j) \in E(H) \}$
  - $E_2 = \{ ((v, i), (u, j) \mid u \sim_i v, v \sim_j u \}$

  The notation $v \sim_j u$ states that v is the $i-th$ neighbour of u

Claim 7 *The replacement product graph is an expander.*

Proof Let $R = G ⓡ H$ be a graph created by the replacement product. Consider the graph G^3, which is the graph on the same vertex set but whose edges correspond to length-3 paths (the adjacency matrix of this graph is simply A^3). It is immediate that $E(G^3) \supset E(G ⓩ H)$. This means that G^3 has at least the same edge expansion as $G ⓩ H$, meaning that its second eigenvalue is also bounded away from 1. Since $\lambda(G)^3 = \lambda(G^3)$ we deduce that $\lambda(G)$ is also bounded away from 1, and thus it is also an expander.
■

We conclude this section by reminding that the construction of a family of expander graphs using the zigzag product require some initial expander graph to begin with. We have seen last week, that once we have this initial graph, we can create a whole family of expanders. Since this expander is only on a constant number of vertices, it can be found in $O(1)$ time.

3 PCP - Probabilistically Checkable Proofs

3.1 Proof systems and the class NP

Definition 8 *Given a set of axioms and deduction rules. A Proof is defined to be a series of formulas $\phi_1, \phi_2, ..., \phi_n$ where the each of the formulas are either an axiom or was deduced by the deduction rules by previous formulas.*

Definition 9 *A system which holds the above is called a Proof System.*

A proof system must have the two following properties:

- Completeness: if ϕ is true, then there exists a valid proof $\phi_1, \phi_2, ..., \phi_m$. And it holds that $\phi = \phi_m$.
- Soundness: if ϕ is false, then there is no proof $\phi_1, \phi_2, ..., \phi_m$, such that $\phi = \phi_m$.

Definition 10 *The Class NP is the set of languages L for which there exists a polynomial algorithm A and a constant c such that:*

1. $x \in L \Rightarrow \exists y, |y| \leq |x|^c; A(x, y) = 1$
2. $x \notin L \Rightarrow \forall y, |y| \leq |x|^c; A(x, y) = 0$

Remark

1. A is a verification algorithm.
2. A can be viewed as a proof system for L

3.2 Randomized Verifiers

Definition 11 *It is said that an algorithm (receiving input x and a proof y) has oracle access to the string y, if it can access a specific index within y with a cost of one step. This is denoted by $A^y(x)$.*

We use this oracle access to define randomized algorithms which access only a limited number of bits in the string y.

Definition 12 *An algorithm A is a (r, q)-verifier for a language L if for an input string x and a proof string y, A uses at most r random bits, reads at most q bits from y, runs in polynomial time, such that:*

1. $x \in L \Rightarrow \exists y, |y| \leq |x|^c; \Pr_{\rho \in_R \{0,1\}^r}[A^y(x, \rho) = 1] = 1$
2. $x \notin L \Rightarrow \forall y, |y| \leq |x|^c; \Pr_{\rho \in_R \{0,1\}^r}[A^y(x, \rho) = 1] < \frac{1}{2}$

The notation $A^y(x, \rho)$ means that A has regular access to the input x and the random bits ρ, and oracle access to the proof string y.

Definition 13 *We define the class $PCP[r, q]$ to be the collection of all languages L for which there is an (r, q)-verifier.*

Remark Note that by this definition, $NP = \cup_c PCP[0, n^c]$.

Remark $P = PCP(0, O(1))$.

Remark $NP = \cup_c PCP(\log n, n^c)$. The $\subseteq$ direction is immediate. For the $\supseteq$ direction, observe that any verifier which uses $O(\log n)$ random bits can be made deterministic by enumerating over all possible random strings (there are at most $2^{O(\log n)}$ of them) and outputting 1 unless we encounter any string for which the algorithm returns 0.

Theorem 14 (PCP Theorem) $PCP[O(\log n), O(1)] = NP$

This is a very surprising theorem since it states that every language in NP has a proof format that can be verified by reading only a constant number of bits! These must be randomly chosen of course.