CMPUT 675: Computational Complexity Theory **Winter 2019**

Lecture 19 (Mar 19): Expander Graphs

Lecturer: Zachary Friggstad *Scribe: Noah Weninger*

19.1 Definitions

Definition 1 *Say G is a (n, d, λ)-expander if*

- *The number of nodes is n.*
- *G is d-regular. That is, every vertex touches exactly d edges. The graph may include parallel edges and loops, but loops only count once towards d.*
- *$|\lambda_2| \leq \lambda$ where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of G with $|\lambda_1| \geq \cdots \geq |\lambda_n|$.*

We will see this definition is most useful when $\lambda < 1$ and is independent of n.

Definition 2 *Say $\{G_n\}_{n \geq 1}$ is a family of (d, λ)-expanders if each G_n is a (n, d, λ)-expander.*

People frequently use the term "expander graph" to mean a graph from such a family for some constant d and some $\lambda < 1$.

Definition 3 *Say $\{G_n\}_{n \geq 1}$ is a strongly explicit family of (d, λ)-expanders if, given n and some vertex $v \in [n]$, we can compute the neighbours of v in G_n in polylog(n) time (i.e. polynomial in the logarithm of n).*

19.2 Random walks in expanders

There are quite a few "random walks in expanders" results that are useful for different applications. We will cover one that is particularly useful for applications we will discuss shortly.

Theorem 1 *Let $G = (V; E)$ be a (n, d, λ)-expander. Consider a random walk $v_0, \ldots, v_k$ where*

- *v_0 is uniformly chosen node of G.*
- *v_{i+1} is a random neighbour of v_i, selected by picking an edge incident to v_i uniformly at random. Note that v_i may equal v_{i+1} if the chosen edge is a loop.*

For any $B \subseteq V$, let $\beta = \frac{|B|}{|V|}$. Then the probability that the random walk stays in B satisfies

$$\Pr[v_0, \ldots, v_k \in B] \leq ((1 - \lambda) \cdot \sqrt{\beta} + \lambda)^k.$$

For example, say $\beta = \frac{1}{2}$. Then there is a 50% chance that $v_0 \in B$. If G wasn't an expander, then there might be only a single edge connecting B to $V - B$, so there is a high probability of staying in B after a single step. But in an expander graph, the number of edges leaving B is linear in $|B|$, so, intuitively, there is a constant probability of leaving the set in each step. This is not a precise statement, but what is true is that if $\lambda, \beta < 1$, then the probability of staying in B decreases geometrically with the length of the walk.

This theorem will be proved later on, but first we'll motivate it with a few applications.

19.3 Applications

19.3.1 Error reduction rates for RP

Take $L \in \mathbf{RP}$ and let M be a PTM using $r(n)$ random bits to decide L. So $\forall x$,

- $x \in L \Rightarrow \Pr_{y \sim \{0,1\}^{r(|x|)}}[M(x, y) = \text{ACCEPT}] = 1$
- $x \notin L \Rightarrow \Pr_{y \sim \{0,1\}^{r(|x|)}}[M(x, y) = \text{ACCEPT}] \leq \dfrac{1}{2}$

Recall that we could drive down the probability of accepting a "No" instance to 2^{-k} by independently repeating $M(x)$ k times. However, for k repetitions, this method requires $k \cdot r(n)$ random bits. Here, we will show that we can improve this to only require $O(k) + r(n)$ random bits.

Let G be a strongly explicit $(2^{r(|x|)}, d, \lambda)$-expander graph where d is a constant and $\lambda < 1$. We haven't yet seen how to construct one, but let's assume we have such a graph (represented implicitly). Notice there is a one-to-one correspondence between the nodes of G and all strings $y \in \{0,1\}^{r(|x|)}$. We can then use G to decide $x \in L$ with the following procedure:

Let $y_0, \ldots, y_k$ be a random walk in G as in the statement of Theorem 1.
Accept iff $\forall 0 \leq i \leq k,\ M(x, y_i) = \text{ACCEPT}$.

Proof. Creating the random walk takes $O(k)+r(n)$ random bits: $r(n)$ to select v_0, then $k \cdot O(\log d) = O(k)$ bits[1] to select a neighbour of v_i for all $0 \leq i < k$. Since G is strongly explicit, computing neighbouring vertices only takes polylogarithmic time in $2^{r(n)}$, which is polynomial in n. Let $B_x = \{y \in \{0,1\}^{r(|x|)} : M(x, y) = \text{ACCEPT}\}$. The procedure decides $x \in L$ because:

- If $x \in L$, $B_x = \{0,1\}^{r(|x|)}$, so $y_0, \ldots, y_k \in B_x$ and therefore all runs of $M(x, y_i)$ are accepted.
- If $x \notin L$, $|B_x| \leq \frac{1}{2} \cdot |\{0,1\}^{r(|x|)}|$ so by Theorem 1,

$$\Pr_{y_0,\ldots,y_k}[y_0, \ldots, y_k \in B_x] \leq ((1-\lambda) \cdot \frac{1}{\sqrt{2}} + \lambda)^k$$

  Since λ is a constant < 1, for some constant $c > 0$, we have $((1-\lambda) \cdot \frac{1}{\sqrt{2}} + \lambda)^k = 2^{-c \cdot k}$. So, we can achieve the same error bound as with independent repetitions using only $O(k) + r(n)$ random bits.

■

[1] If d is not a power of 2 there is a subtlety in how to do this with coin flips. We could (implicitly) add $\leq d$ loops to each vertex so each vertex has degree 2^k. It is easy to see the graph remains an expander: the parameter λ would change to $\frac{d}{2^k} \cdot \lambda + \frac{2^k - d}{2^k} \cdot 1 < 1$ as the new random walk matrix is the corresponding averaging of G's random walk matrix and the identity matrix I.

19.3.2 Approximability bounds

Claim 1 *There exists an* $(O(\log n), O(\log n))$-**PCP** *verifier* V *for SAT such that* $\forall\phi$*:*

- $\phi \in SAT \Rightarrow \exists\pi$ *such that* $\Pr_r[V(\phi,\pi,r)] = 1$
- $\phi \notin SAT \Rightarrow \forall\pi$ *such that* $\Pr_r[V(\phi,\pi,r)] \leq \frac{1}{n^s}$ *for some constant* s *that is independent of* n*.*

Proof. This proof is very similar to the previous statement about languages in **RP**, so here we only present the setup; the result follows almost exactly as before. By the PCP theorem, we know there exists an $(O(\log n), O(1))$-**PCP** verifier V' for SAT such that $\forall\phi$:

- $\phi \in SAT \Rightarrow \exists\pi$ such that $\Pr_r[V'(\phi,\pi,r)] = 1$
- $\phi \notin SAT \Rightarrow \forall\pi$ such that $\Pr_r[V'(\phi,\pi,r)] \leq \frac{1}{2}$

Now, using V', we will construct V.

V expects the same proof π as V'. Say V' uses $r(n) := c \cdot \log n$ random bits. First, for some $k \in \Theta(\log n)$, V samples $r_0, \ldots, r_k$ by a random walk in some $(2^{r(n)}, d, \lambda)$-expander where d is constant and $\lambda < 1$. Note such an expander does not need to be *strongly* explicit, we just need to construct it in poly(n) time. Then V accepts iff $\forall 0 \leq i \leq k,\ V'(x,\pi,r_i) = \textsc{accept}$.

Since each call to V' queries $O(1)$ bits of the proof, V queries only $O(\log n)$ bits. Note that if V didn't use expanders and just ran V' with k independent random bit strings, it would use $\Theta(\log^2 n)$ random bits. But here we only need $O(\log n)$: $r(n)$ bits to select r_0, and $\log n \cdot O(\log d) = O(\log n)$ bits to select a neighbour of r_i for all $0 \leq i < k$.

The justification for this is just as in the previous proof. ■

Corollary 1 *There exists a constant* $\gamma > 0$ *such that there is no* $\frac{1}{n^\gamma}$*-approximation for max independent set unless* P = NP*.*

Proof. Consider the following reduction from SAT to max independent set. Let V be a $(c \cdot \log n, q \cdot \log n)$-**PCP** verifier for SAT as in Claim 1 (so c, q and s are all constant). Then build a graph H by:

- Nodes are $(r, b) \in \{0,1\}^{c\cdot\log n} \times \{0,1\}^{q\cdot\log n}$ such that V would accept x given random string r if the queried bits of the proof were b.
- $[(r,b),(r',b')]$ is an edge if they disagree on a bit of the proof. Note that b and b' cannot be compared as strings: we need to first find all queried proof indices that are common to V with both random strings r and r', then check the corresponding positions in b and b' for consistency.

We would like to show that

- $\phi \in$ SAT $\Rightarrow$ max independent set size in H is $\geq n^c$,
- $\phi \notin$ SAT $\Rightarrow$ max independent set size in H is $\leq n^{c-s}$.

Consider both cases:

- $\phi \in$ SAT. Let π be a proof that is accepted by V for any random string r. Consider $\mathcal{I} = \{(r, b) : r \in \{0,1\}^{c \cdot \log n}$ and b agrees with π on the positions queried by V given $r\}$. There can only be one vertex in $\mathcal{I}$ for each random string r, because otherwise the two b's would have to disagree somewhere. Since every b in $\mathcal{I}$ agrees with π, there cannot be an edge between any two vertices in the set. Therefore $\mathcal{I}$ is a maximum independent set in H and it has size $2^{c \cdot \log n} = n^c$.
- $\phi \notin$ SAT. Let $\mathcal{I}$ be a max independent set of H. Form a proof string π as follows: $\forall (r, b) \in \mathcal{I}$, set the bits of π queried by V given r according to b. As $\mathcal{I}$ is an independent set, this does not give conflicting values to any bits of π. Any unspecified bit of π can be set arbitrarily. Then

$$\frac{|\mathcal{I}|}{2^{c \cdot \log n}} \leq \Pr_r[V(x, \pi, r) = \text{ACCEPT}] \leq \frac{1}{n^s}$$

  where s is as in Claim 1. This follows because the verifier would accept π for each random string r such that $(r, b) \in \mathcal{I}$ for some b. Therefore $|\mathcal{I}| \leq n^{c-s}$.

However, we wanted to our result in terms of the size of the graph, which may differ from n. Let N be the number of nodes in H, so $N \leq n^{c+q}$. Then

$$\frac{\text{max ind set in "No" case}}{\text{max ind set in "Yes" case}} \leq \frac{n^{c-s}}{n^c} = \frac{1}{n^s} \leq \frac{1}{N^{\frac{s}{c+q}}}.$$

Set $\gamma = \frac{s}{c+q}$. ∎

19.4 Proof of the expander random walk result

Before we can show Theorem 1, we need a few fundamentals from linear algebra.

Definition 4 *For a square matrix A, the spectral norm is defined as* $\|A\| := \max_{x : \|x\|_2 = 1} \|A \cdot x\|_2$.

Informally the spectral norm of A is the maximum scale by which A can stretch any vector. It has a few useful properties: $\forall$ matrices A, B and vectors x,

$$\|A \cdot B\| \leq \|A\| \cdot \|B\|$$
$$\|A + B\| \leq \|A\| + \|B\|$$
$$\|A \cdot x\|_2 \leq \|A\| \cdot \|x\|_2.$$

The proofs of these properties are left as an exercise (assignment 5, exercise 3).

We define $\mathbf{1}$ to be the n-element column vector of all $\frac{1}{n}$, with n inferred from context. Additionally, define J as the $n \times n$ matrix of all $\frac{1}{n}$. Intuitively, J is the random walk matrix of the n-clique graph where every node has a self loop, so a single random step could lead to any other vertex uniformly at random.

Theorem 2 (Cauchy-Schwarz inequality) *For $x, y \in \mathbb{R}^n$, $|\langle x, y \rangle| \leq \|x\|_2 \cdot \|y\|_2$.*

Proof. $\forall t \in \mathbb{R}$, $0 \leq \langle x \cdot t - y, x \cdot t - y \rangle = t^2 \cdot \|x\|_2^2 - 2t \cdot \langle x, y \rangle + \|y\|_2^2$. Let us pick t to put the most stress on this inequality: to minimize the quadratic.

If we take $t = \frac{\langle x,y\rangle}{\|x\|_2^2}$, then we have

$$\begin{aligned}
0 &\le \frac{\langle x,y\rangle^2}{\|x\|_2^2} - 2\cdot\frac{\langle x,y\rangle^2}{\|x\|_2^2} + \|y\|_2^2 = \|y\|_2^2 - \frac{\langle x,y\rangle^2}{\|x\|_2^2} \\
0 &\le \|x\|_2^2 \cdot \|y\|_2^2 - \langle x,y\rangle^2 \\
\langle x,y\rangle^2 &\le \|x\|_2^2 \cdot \|y\|_2^2 \\
|\langle x,y\rangle|^2 &\le \|x\|_2^2 \cdot \|y\|_2^2 \\
|\langle x,y\rangle| &\le \|x\|_2 \cdot \|y\|_2
\end{aligned}$$

■

Corollary 2 *Define* $\|x\|_1 := \sum_{i=1}^n |x_i|$*. Then* $\forall x \in \mathbb{R}^n$, $\|x\|_2 \le \|x\|_1 \le \sqrt{n}\cdot\|x\|_2$.

Proof. Let $\overline{x_i} = |x_i|$. Notice that $n\cdot\mathbf{1}$ is an column vector of ones, so $\|n\cdot\mathbf{1}\|_2 = \sqrt{n}$. Then $\|x\|_1 = \langle \overline{x}, n\cdot\mathbf{1}\rangle \le \|\overline{x}\|_2 \cdot \|n\cdot\mathbf{1}\|_2 = \|x\|_2\cdot\sqrt{n}$. For the lower bound, observe $\|x\|_2^2 = \sum_{i=1}^n x_i^2 \le (\sum_{i=1}^n x_i)^2 = \|x\|_1^2$. ■

Finally, we have all the tools we need to tackle Theorem 1.

Proof of Theorem 1. Recall:

- $G = (V;E)$ is an (n,d,λ)-expander.
- $v_0 \sim$ nodes of G.
- $v_{i+1} \sim$ neighbours of v_i by picking a random edge uniformly.

We want to to show $\forall B \subseteq V,\ \Pr[v_0,\dots,v_k \in B] \le ((1-\lambda)\sqrt{\beta}+\lambda)^k$, where $|B| = \beta\cdot n$.

Let $Z_{u,v} = \begin{cases} 1 & \text{if } u = v \text{ and } u \in B \\ 0 & \text{otherwise.}\end{cases}$

For any vector v, Zv will zero out the entries that are not in B. So $\|Z\cdot\mathbf{1}\|_1 = \beta = \Pr[v_0 \in B]$. After taking one random step, we have $\|(Z\cdot A)\cdot Z\cdot\mathbf{1}\|_1 = \Pr[v_0, v_1 \in B]$ where A is the random walk matrix of G. By induction, $\|(Z\cdot A)^k\cdot Z\cdot\mathbf{1}\|_1 = \Pr[v_0,\dots,v_k \in B]$. Note that by Corollary 2, $\|(Z\cdot A)^k\cdot Z\cdot\mathbf{1}\|_1 \le \sqrt{n}\cdot\|(Z\cdot A)^k\cdot Z\cdot\mathbf{1}\|_2$.

Define the notation $\lambda_2(A)$ to mean the 2nd largest eigenvalue of A (in absolute value). Recall that J is the uniform random walk matrix. We claim that $A = (1-\lambda_2(A))\cdot J + \lambda_2(A)\cdot C$ where $\|C\| \le 1$. The proof will be included in the next lecture. For now, we will assume it is true in order to finish the current proof.

Direct calculation shows $\|Z\cdot\mathbf{1}\|_2 = \frac{\sqrt{\beta}}{\sqrt{n}}$. By the properties of the spectral norm, $\|(Z\cdot A)^k\cdot Z\cdot\mathbf{1}\|_2 \le \|Z\cdot A\|^k\cdot\|Z\cdot\mathbf{1}\|_2 = \|Z\cdot A\|^k\cdot\frac{\sqrt{\beta}}{\sqrt{n}}$. Then, by the claim

$$\begin{aligned}
\|Z\cdot A\| &= \|Z\cdot(1-\lambda_2(A))\cdot J + Z\cdot\lambda_2(A)\cdot C\| \\
&\le (1-\lambda_2(A))\cdot\|Z\cdot J\| + \lambda_2(A)\cdot\|Z\cdot C\|
\end{aligned}$$

Since $\|C\| \le 1$, we have $\|Z\cdot C\| \le 1$. Now we need to show $\|Z\cdot J\| \le \sqrt{\beta}$. Let x be a unit vector that maximizes $\|Z\cdot J\cdot x\|_2$, so that $\|Z\cdot J\cdot x\|_2 = \|Z\cdot J\|$. Then $J\cdot x = \alpha\cdot\mathbf{1}$ for $\alpha = \sum_v x_v$, because all rows of J are $\mathbf{1}^\top$. Note

$|\alpha| \leq \|x\|_1$. With a bit of algebra we arrive at our bound:

$$\begin{aligned}\|Z \cdot J\| &= \|Z \cdot J \cdot x\|_2 \\ &= \|Z \cdot \alpha \cdot \mathbf{1}\|_2 \\ &= |\alpha| \cdot \|Z \cdot \mathbf{1}\|_2 \\ &= |\alpha| \cdot \frac{\sqrt{\beta}}{\sqrt{n}} \\ &\leq \|x\|_1 \cdot \frac{\sqrt{\beta}}{\sqrt{n}} \\ &\leq \|x\|_2 \sqrt{\beta} \\ &= \sqrt{\beta}\end{aligned}$$

Then because $\sqrt{\beta} \leq 1$ and $\lambda_2(A) \leq \lambda$, we have $\|Z \cdot A\| \leq (1-\lambda) \cdot \sqrt{\beta} + \lambda$. In conclusion,

$$\Pr[v_0, \ldots, v_k \in B] = \|(Z \cdot A)^k \cdot Z \cdot \mathbf{1}\|_2 \leq ((1-\lambda) \cdot \sqrt{\beta} + \lambda)^k \cdot \frac{\sqrt{\beta}}{\sqrt{n}} \leq ((1-\lambda) \cdot \sqrt{\beta} + \lambda)^k.$$

■

References

AB09 S. Arora and B. Barak, Computational Complexity: A Modern Approach, *Cambridge University Press, New York, NY, USA*, 2009.