CS 6815 Pseudorandomness and Combinatorial Constructions **Fall 2019**

Lecture 7: September 19th

Lecturer: Eshan Chattopadhyay *Scribe: Alexander Frolov*

7.1 Proof of Expander Chernoff Bound

Let $G = (V, E)$ be an (N, γ) spectral expander that is D-regular.

Define a random walk of length t as $l_1 \to l_2 ... \to l_t$, such that l_1 is chosen randomly from V and each l_{i+1} is a random element of the neighbors of l_i.

Claim 7.1 *For all* $f : [N] \to [0, 1]$, $Pr\left[\left|\frac{1}{t}\sum_i f(l_i) - \mu_f\right| > (\delta + \lambda)\right] \leq e^{-\Omega(\delta^2 t)}$, *where* $\mu_i = \mathbb{E}_{j\sim[n]}(f(j))$.

Proof: Define $X_i = f(l_i)$. Define $X = \sum_i^t X_i$.
The goal of this proof will be to bound $\Delta = Pr[X \geq (\delta + \lambda + \mu_f)t]$.
Mirroring the proof of the Chernoff bound, $\Delta = Pr[e^{rx} \geq e^{r(\delta+\lambda+\mu_f)t}]$.
Applying Markov's inequality, $Pr[e^{rx} \geq e^{r(\delta+\lambda+\mu_f)t}] \leq \frac{\mathbb{E}[e^{rx}]}{e^{r(\delta+\lambda+\mu_f)t}}$.
Next, the goal of this proof is to express $\mathbb{E}[e^{rx}]$ linear algebraically.

Define the diagonal matrix $D_f = \begin{pmatrix} e^{rf(1)} & & & \\ & e^{rf(2)} & & \\ & & \ddots & \\ & & & e^{rf(N)} \end{pmatrix}$, with zeros elsewhere.

Let $u = \frac{1}{N}\overrightarrow{1}$. $\mathbb{E}[e^{rx}] = ||uD_f(AD_f)^{t-1}||_1$ (this equality is a linear algebraic interpretation of expectation).
Next, we will attempt to bound this quantity. $||uD_f(AD_f)^{t-1}||_1 = ||uAD_f(AD_f)^{t-1}||_1 = ||u(AD_f)^t||_1$ since $uA = A$, as $\overrightarrow{1}$ is an eigenvalue of this matrix.
Applying Cauchy-Schwarz and then the submultiplicativity of the L_2 norm:
$||u(AD_f)^t||_1 \leq \sqrt{N}||u(AD_f)^t||_2 \leq \sqrt{N}||u||_2||(AD_f)^t||_2 = \sqrt{N} * \frac{1}{\sqrt{N}}||(AD_f)^t||_2 = ||(AD_f)^t||_2$. Next, applying submultiplicativity again: $||(AD_f)^t||_2 \leq (||AD_f||_2)^t$. Now, we aim to bound $||AD_f||_2$.
We will apply the matrix decomposition lemma from a previous lecture:

Lemma 7.2 The adjacency matrix A of a spectral expander with expansion $\gamma = 1 - \lambda$ can be decomposed as $A = \gamma J + \lambda E$. such that J is $\frac{1}{N}$ times the all ones matrix, $||E||_2 \leq 1$, and γ, λ are scalars.

This lemma has an interesting interpretation in the context of expander graphs. Since J is the normalized adjacency matrix of a fully connected graph, random walks on such a graph converge to uniformity very quickly. Since the adjacency matrix A deviates from J by a small amount (a matrix E with spectral norm less than or equal to 1), this means that an expander is not far from one whose random walks converge to uniformity quickly.

Continuing with the proof of the Expander Chernoff bound, and specifically bounding $||AD_f||_2$, $||AD_f||_2 \leq \gamma||JD_f||_2 + \lambda||ED_f||_2$ (using the triangle inequality and the matrix decomposition lemma). $||JD_f||_2, ||ED_f||_2$ will be bound individually to bound $||AD_f||$.
$||ED_f||_2 \leq ||E_2||_2||D_f||_2 \leq 1 * ||D_f||_2 \leq max_{i\in[N]}(e^{rf(i)})$. Since $f(i) \in \{0, 1\}$, $e^{r*f(i)} \leq e^r \leq 1 + r + O(r^2)$ using a Taylor expansion.

$||JD_f||_2 = max_{x \in R_n, ||x||_2=1} ||JD_f x||_2$. Looking at $JD_f x$, it is equal to $J * (x_1 e^{rf(1)}, x_2 e^{rf_2}, ... x_n e^{rf(n)}) = \frac{1}{N} \sum_{i=1}^{N} x_i e^{rf(i)} \cdot \overrightarrow{1}$. Hence, $||JD_f x||_2 \leq \frac{1}{N} | \sum_{i=1}^{N} x_i e^{rf(i)} | * \sqrt{N} = \frac{1}{\sqrt{N}} | \sum_{i=1}^{N} x_i e^{rf(i)} |$. This is bounded above by $\frac{1}{\sqrt{N}} (\sum x_i^2)^{\frac{1}{2}} (\sum e^{2rf(i)})^{\frac{1}{2}}$. Applying a Taylor expansion, this is bounded above by $\frac{1}{\sqrt{N}} (\sum_{i=1}^{N} (1 + 2rf(i) + O(r^2)))^{\frac{1}{2}}$. Since $f(i) \in [0,1]$, this can be again bounded by $\frac{1}{\sqrt{N}} (N + 2r\mu_f N + NO(r^2))^{\frac{1}{2}} \leq 1 + r\mu_f + O(r^2)$ (the last step is justified by the Taylor series for $f(x) = \sqrt{x}$.

Using these bounds (and that $\lambda + \gamma = 1$) $||AD_f||_2 \leq \gamma(1 + r\mu_f) + \lambda(1+r) + O(r^2)$. $\leq 1 + r(\lambda + \mu_f) + O(r^2)$.

Next, $\mathbb{E}[e^{rx}] \leq (||AD_f||_2)^t \leq e^{rt(\lambda+\mu_f)+O(r^2)t}$. Plugging this bound into the bound from Markov's inequality, $\Delta \leq \frac{\mathbb{E}[e^{rx}]}{e^{r(\delta+\lambda+\mu_f)t}} \cdot \leq \frac{e^{rt(\lambda+\mu_f)+O(tr^2)}}{e^{rt(\lambda+\mu_f)+rt\delta}} = e^{O(r^2 t) - rt\delta}$. Choosing $r = \frac{\delta}{C}$, such that C is a large constant, this yields the desired bound of $e^{-\Omega(\delta^2 t)}$.

7.2 Using the Expander Mixing Lemma for Error reduction in BPP

Let A be an algorithm for $L \in BPP$ that uses R bits of randomness and time T. Remember that $\forall x \in L, y \sim [0,1]^R, Pr[A(x,y) = 1] \geq \frac{2}{3}$, and $\forall x \notin L, y \sim [0,1]^R, Pr[A(x,y) = 0] \geq \frac{2}{3}$

Take an expander G on $[2^R]$ nodes, $\lambda = \delta = \frac{1}{20}$ (and $D = O(1)$ probabilistically). Now, for a fixed x, define the function $f : [2^R] \to [0,1]$, where $f(y)$ returns 1 if $A(x,y)$ is correct, and 0 otherwise. Since the two sided success probability of A is $\frac{2}{3}$, $f(y) = 1$ for at least $\frac{2}{3}$ of random bit strings, so $\mu_f \geq \frac{2}{3}$.

Define A' to run $A(x, l_1), A(x, l_2), ... A(x, l_t)$, where $l_1, l_2, ..., l_t$ are random bitstrings of length R taken from a random walk of expander G, and output the majority answer from these t executions.

Now, since δ, λ, are constant, $\mu_f \geq \frac{2}{3}$, and $\mu_f - \delta - \lambda \geq \frac{1}{2}$, the error probability of A' is bounded by $e^{-\Omega(t)}$. Defining this error probability as ϵ, reducing the error probability to ϵ requires $O(log(\frac{1}{\epsilon}))$ executions of A. The random bits required for this algorithm are R bits to sample the first vertex in the walk, and then $O(log(\frac{1}{e}))$ random bits (a constant number of bits per iteration of the A to take a step in the expander walk.)

Making a table of time vs random bits required for various error reduction techniques for BPP algorithms shows that this approach is both efficient in terms of random bits and time.

Error Reduction Technique	*Time*	*Random Bits*
Expander Walk	$T \cdot O(log(\frac{1}{\epsilon}))$	$R + O(log(\frac{1}{\epsilon}))$
I.i.d Repetitions	$T \cdot O(log(\frac{1}{\epsilon}))$	$R \cdot O(log(\frac{1}{\epsilon}))$
Pairwise Independent Repetitions	$T \cdot O(\frac{1}{\epsilon})$	$R + 2log(\frac{1}{\epsilon}) + O(1)$

7.3 Bounds on Expansion of Graphs

Definition 7.3 *The spectral gap, γ of a graph is $1 - \lambda$, where λ is the second largest eigenvalue of the graph's normalized adjacency matrix.*

Theorem 7.4 *[Alon-Boppana]: $\lambda \geq \frac{1}{D} 2\sqrt{D-1} - o_N(1)$, where G is D regular.*

A result of Ramanujan is that there exist graphs (Ramanujan expanders) for which $\lambda \leq \frac{2\sqrt{D-1}}{D}$, which nearly achieve the bound of Alon and Boppana.

Claim 7.5 *We will show a weaker result, that* $\lambda \geq \sqrt{D-1} - o_N(1)$.

Let M be the unnormalized adjacency matrix of expander G. By inspection, $tr(M^2)$ is equal to the number of length 2 walks that start and end at the same vertex of G. This quantity is bounded above by ND, since each vertex can have a maximum of D length 2 walks for each of the D vertices leaving it.

By the properties of eigenvalues, $tr(M^2) = D^2 \sum_{i=1}^{N} \lambda_i^2$, where λ_i is the ith eigenvalue of A, the normalized adjacency matrix of G.

Combining these bounds, $tr(M^2) = D^2 \sum_{i+1}^{N} \lambda_i^2 \leq ND$. Subtracting the largest eigenvalue and dividing yields $D^2 \sum_{i=2}^{N} \lambda_i^2 \leq ND - D^2$ and $\sum_{i=2}^{N} \lambda_i^2 \leq \frac{N-D}{D}$. Finally, bounding each eigenvalue by λ_2 yields that $\lambda_2 \geq \sqrt{\frac{N-D}{DN}}$.

Finally, if $D \ll N$ (which is achievable since D can be $O(1)$), this yields the desired bound.

7.4 Explicit Constructions of Expanders

Definition 7.6 *The Margulis-Gabbard-Galil Construction is a graph where* $V = \mathbb{Z}_m \times \mathbb{Z}_m$, *and the edges satisfy* $(x, y) \to (x \pm y, y)$, $(x, y) \to (x \pm (y + 1), y)$, $(x, y) \to (x, x \pm y)$, *or* $(x, y) \to (x, y \pm (x + 1))$.

This construction achieves (m^2, γ) expansion, where γ is a constant greater than 0. This graph is 8-regular and strongly-explicit (see definition below). The proof that this is an expander is beyond the scope of this class.

Definition 7.7 *We will say that the construction of a graph on* N *nodes is strongly explicit if there exists an algorithm that determines whether there is an edge between any two vertices in* $poly(log(N))$ *time.*

Definition 7.8 *We will say that the construction of a graph on* N *nodes is mildly explicit if there exists an algorithm that runs in time poly(N) and outputs the adjacency matrix of the graph.*

We will study a more combinatorial construction of an expander. The approach will be to start with a constant sized expander (which we know exists from the probabilistic method, and hence can be found by a brute-force search in constant time) and create larger expanders graphs via using various graph products, such as graph squaring, the tensor product, and the zig-zag product.