CS369E: Expanders May 1, 2005

Lecture 5: Derandomization (Part II)

Lecturer: Prahladh Harsha *Scribe: Adam Barth*

Today we will use expanders to derandomize the algorithm for linearity test.

Before presenting the linearity testing algorithm and its derandomization, we review some expander preliminaries.

5.1 Expander Preliminaries

So far we have considered vertex expansion. For the derandomized linearity testing, we will need a notion of edge-expansion. Informally, edge-expansion implies that every small set of vertices has a large number of edges leaving the set. We show below that any graph that is an expander (i.e., has spectral expansion λ) is also an edge-expander.

Lemma 5.1 (Edge expansion). *If G has spectral expansion λ, then for all $A \subseteq V$ with $|A| \le n/2$,*

$$e(S, \overline{S}) \ge \frac{d(1-\lambda)}{2}|S|$$

where $e(S, \overline{S})$ denotes the number of edges between S and $\overline{S}$.

Proof. Given a vertex set G on n vertices with adjacency matrix A, recall that $\lambda = \max_{x \perp u} \langle Ax, x\rangle / \langle x, x\rangle$. For a subset $S \subseteq V$ of vertices, fix

$$x = \frac{\chi_S}{|S|} - \frac{\chi_{\overline{S}}}{|\overline{S}|}.$$

In other words, x is the n-dimension vector such that

$$x_v = \begin{cases} \frac{1}{|S|} & \text{if } v \in S \\ -\frac{1}{|\overline{S}|} & \text{if } v \notin S \end{cases}$$

Observe the following: $x \perp u$, and $\langle x, x\rangle = \frac{1}{|S|} + \frac{1}{|\overline{S}|}$, and

$$\begin{aligned}
\langle Ax, x\rangle &= \sum a_{ij} x_i x_j \\
&= \sum_{i,j \in S} \frac{2}{d} x_i x_j + \sum_{i,j \in \overline{S}} \frac{2}{d} x_i x_j + \sum_{i \in S, j \in \overline{S}} \frac{2}{d} x_i x_j \\
&= \frac{2}{d|S|^2}\left(\frac{d|S| - e(S,\overline{S})}{2}\right) + \frac{2}{d|\overline{S}|^2}\left(\frac{d|\overline{S}| - e(S,\overline{S})}{2}\right) - \frac{2e(S,\overline{S})}{d|S||\overline{S}|} \\
&= \left(\frac{1}{|S|} + \frac{1}{|\overline{S}|}\right)\left[1 - \frac{e(S,\overline{S})}{d}\left(\frac{1}{|S|} + \frac{1}{|\overline{S}|}\right)\right],
\end{aligned}$$

Since $\langle Ax, x\rangle \leq \lambda\langle x, x\rangle$, we have that

$$1 - \frac{e(S,\overline{S})}{d}\left[\frac{1}{|S|} + \frac{1}{|\overline{S}|}\right] \leq \lambda.$$

which implies

$$e(S,\overline{S}) \geq d(1-\lambda)\frac{|S||\overline{S}|}{|S| + |\overline{S}|} \geq \frac{d(1-\lambda)}{2}|S|$$

since $|S| \leq n/2$. $\square$

Suppose you remove a few edges from a graph. It is possible we might have partitioned the graph into several small (disconnected) pieces. However, if the graph is an expander, there must exist a huge connected component. This is captured in the following lemma.

Lemma 5.2. *For all $\delta \leq (1-\lambda)/12$, after removing any $2\delta dn$ edges from a graph G with spectral expansion at most λ, there exists a connected component of size at least*

$$\left(1 - \frac{4\delta}{1-\lambda}\right)n.$$

Proof. We prove this lemma in two steps. If removing the edges partitions the graph into two halves, these two halves must be unbalanced with the smaller side containing less than $n/3$ vertices. More precisely, fix a partition $S, \overline{S}$ of G such that the edges in $e(S,\overline{S})$ is contained in the set of removed edges and S is the smaller half (i.e., $|S| \leq n/2$). Then (by Lemma 5.1),

$$\frac{d(1-\lambda)|S|}{2} < e(S,\overline{S}) \leq 2\delta dn \text{ implies } |S| < \frac{4\delta}{1-\lambda}n \leq \frac{n}{3}.$$

Therefore, if there is a component of size at least $n/2$, then there is one of size $1-\frac{4\delta}{1-\lambda} \geq 2n/3$. On the other hand, the graph (on removal of the edges) could consist of several small components (of size less than $n/3$) and not have any large component. The following claim shows that this cannot be the case

Claim 5.3. *The union of all components of size less than $n/3$ is itself of size less than $n/3$.*

Proof. Consider two components C_1 and C_2, each of size less than $n/3$. Their union is of size at most $|C_1 \cup C_2| < 2n/3$. We know from above that any component is of size greater than $2n/3$ or less than $n/3$. Hence, $|C_1 \cup C_2| < n/3$. We could repeatedly do this for all components of size less than $n/3$ to show that their union is of size at most $n/3$. $\square$

Thus, there must exist a large sized component and we are done in this case. $\square$

5.2 Linearity testing

Linearity testing is an instance of the more general problem of property testing [RS, GGR]. In general property testing, the goal is to check whether a huge string has a specific property. The string is so huge that one can not afford to read it in its entirety.

For example, given the adjacency matrix of a huge graph G, suppose we wish to design an algorithm A (also called a (property) tester) to determine whether the graph is bipartite without reading the entire matrix. Clearly, A can not determine exactly whether the graph is bipartite without looking at the entire graph because a single edge may destroy the property of being bipartite. Therefore, we relax the requirements and require A to only distinguish between the cases when G is bipartite and when G is "far" from being bipartite[1] rather than the cases when G is bipartite and when it is non bipartite. Note that A must be randomized because otherwise an adversary could fool A by placing "bad" edges in parts of the matrix A does not inspect. We thus, require the following of the tester A:

- If the graph is bipartite, A must accept with probability at least 2/3.
- If the graph is "far" from bipartite, A must reject with probability at least 2/3.

We know to design testers A which satisfy the above properties and needs to probe at most a constant number of locations of the matrix (the precise constant depends on how "far" we want the graph to be from bipartite) [GGR].

We now consider the linearity testing of Blum, Luby and Rubinfeld [BLR]. In linearity testing, the string we wish to test is a function from $\mathbb{Z}_2^n$ to $\mathbb{Z}_2$, presented as a table. Our proofs will work for the more general case when $f : G \to H$ where G and H are arbitrary groups (not even abelian). For simplicity, we will assume that the groups G and H are abelian. Also it is a good idea to consider the case $G = \mathbb{Z}_2^n$ and $H = \mathbb{Z}_2$. The table lists the value of the function for each input value $x \in G$.

Definition 5.4.
- *A function $f : G \to H$ is said to be* linear *if for all $x, y \in G$, we have $f(x) + f(y) = f(x + y)$.*
- *A function $h : G \to H$ is said to be an* affine function, *if there exists a linear function $f : G \to H$ and a constant $a \in H$ such that for all $x \in G$, we have $h(x) = f(x) + a$.*

We say that a function is δ-far (δ-close) from linear, if the value of the function for at least (at most) δ-fraction of the points in G needs to be changed in order to make it linear.

The goal in linearity testing is to test whether the given function (specified as a table of values) is linear or far from linear, without reading the entire table. Linearity testing has applications to locally testable codes (Hadamard codes) and to probabilistically checkable proof constructions.

Blum, Luby and Rubinfeld proposed the following simple linearity testing algorithm (see Figure 1) [BLR].

We state, without proof, two properties of LT.

Proposition 5.5.
- Completeness: *If f is linear, then* $\Pr\left[\mathsf{LT} \text{ accepts } f\right] = 1$.

[1] We say G is "far" from being bipartite if a "lot of edges" need to be removed in order to make it bipartite

$\mathsf{LT}(G, H)$

Input: function $f : G \to H$, specified as a table of values.

1. Choose $x, y \in_R G$ uniformly at random.
2. Query the table for $f(x), f(y)$ and $f(x+y)$.
3. Check whether $f(x+y) = f(x) + f(y)$. If the check succeeds, LT accepts f, otherwise, LT rejects f.

Figure 1: Linearity Test of Blum, Luby and Rubinfeld [BLR]

- Soundness: *If f is δ-far from linear, then* $\Pr\left[\mathsf{LT} \textit{ accepts } f\right] < 1 - O(\delta)$.

The number of random bits used by LT is $2\log|G|$ ($= 2n$ in the case when $G = \mathbb{Z}_2^n$) because LT selects two element of G uniformly at random. The main question we will address is whether this randomness can be further reduced. Goldreich and Sudan showed that at least $log|G| - O(1)$ random bits is required [GS]. They also suggested that the random bits can be reduced by selecting the second point y from a smaller set S instead of the entire group G. Ben-Sasson et.al. showed that a set with the following properties suffices [BSVW].

1. $s \in S$ implies $-s \in S$.
2. The Cayley graph $G_S = (V_S, E_S)$, where $V_S = G$ and $E_S = \{(x, x+s) \mid x \in G, s \in S\}$, must have spectral expansion λ.

With such a set S in hand, we modify LT to choose $x \in_R G$ and $y \in_R S$ uniformly at random. We call the modified algorithm derand-LT. The modified derandomized linearity testing due to Ben-Sasson et. al. [BSVW] is shown in Figure 2.

derand-LT(G, H, S)

Input: function $f : G \to H$, specified as a table of values.

1. Choose $x \in_R G$ and $y \in_R S$ uniformly at random.
2. Query the table for $f(x), f(y)$ and $f(x+y)$.
3. Check whether $f(x+y) = f(x) + f(y)$. If the check succeeds, derand-LT accepts f, otherwise, derand-LT rejects f.

Figure 2: Derandomized Linearity Test of Ben-Sasson, Sudan, Vadhan and Wigderson [BSVW]

When $G = \mathbb{Z}_2^n$ there exist deterministic constructions for such sets S of size $\text{poly}\log|G|$. For general groups, deterministic constructions exist for sets S of size G^ϵ, for every $\epsilon > 0$. In terms of number of random bits, $G = \mathbb{Z}_2^n$ requires $\log|G| + \log\log|G|$ bits and generic groups G require $(1+\epsilon)\log|G|$ bits. This might not seem as a great savings in randomness –

a mere constant factor of 2; however, in PCP constructions and Locally testable codes, one is actually interested in reducing the constant before the leading term (which is typically $O(\log n)$).

Clearly, every linear function is accepted by derand-LT with probability 1. To prove the soundness of derand-LT, we follow the approach of Shpilka and Wigderson [SW].

Theorem 5.6 (Derandomized LT). *Let $\delta < (1-\lambda)/12$ where λ is the spectral expansion of the Cayley graph G_S. Then, if* derand-LT *accepts f with probability at least $1-\delta$, then f is $4\delta/(1-\lambda)$-close to an affine function.*

Note that we do not show that f is close to a linear function (this is in fact not true), but only the weaker statement that f is close to some affine function.

Proof. Suppose derand-LT rejects with probability $p \leq \delta$. That is,

$$\Pr_{x\in G, s\in S}\left[f(x+s) \neq f(x) + f(s)\right] \leq \delta.$$

Given $y \in G$, we define the "opinion of y about $f(x)$" as $f(x+y) - f(y)$. We define a function $\varphi : G \to G$ such that for all $x \in G$, $\varphi(x)$ is the plurality over $y \in G$ of the opinion of y about $f(x)$, that is $f(x+y) - f(y)$ i.e.,.

$$\varphi(x) \quad = \quad \text{plurality}_{y\in G}\left(f(x+y) - f(y)\right).$$

The following three claims prove the theorem.

Claim 5.7 (Popularity is majority). *For all x,*

$$\Pr_{y}\left[\varphi(x) = f(x+y) - f(y)\right] > 1 - \left(\frac{4\delta}{1-\lambda}\right).$$

Proof. Given $x \in G$, we remove the following edges from G_S. If $f(y+s) \neq f(y) + f(s)$, then remove edge $(y, y+s)$ from G_S. If $f(x+y+s) \neq f(x+y) + f(s)$, then remove edge $(y, y+s)$ from G_S. Notice we have removed at most $2\delta dn$ edges from G_S and hence by Lemma 5.2, there exists a huge connected component of size at least $1 - \frac{4\delta}{1-\lambda}$.

If an edge remains in the graph, then $f(x+y+s) - f(y+s) = f(x+y) - f(y)$ and therefore y and $y+s$ share the same opinion about $f(x)$. Hence, all vertices in the huge connected component share the same opinion about $f(x)$ which must agree with the plurality. □

Claim 5.8 (φ is linear). *For all $x, y \in G$, $\varphi(x+y) = \varphi(x) + \varphi(y)$.*

Proof. Let $x, y \in G$. We first show $\Pr_{z\in G}\left[\varphi(x+y) = \varphi(x) + \varphi(y)\right] > 0$. This will prove that $\varphi)x+y) = \varphi(x) + \varphi(y)$ since this event is independent of z. Hence, φ is linear. Consider the following events for a random z:

E_1: $\varphi(x+y) = f(x+y+z) - f(z)$.

E_2: $\varphi(x) = f(x+y+z) - f(y+z)$.

E_3: $\varphi(y) = f(y+z) - f(z)$.

Each of these events occurs with probability $1 - 4\delta/(1-\lambda)$ (by Claim 5.7). By the union bound, the probability that at least one of them fails to occur is at most $12\delta/(1-\lambda) < 1$. Hence, the events E_1, E_2 and E_3 occur simultaneously with non-zero probability. However, if E_1, E_2 and E_3 all occur, we then have that $\varphi(x+y) = \varphi(x) + \varphi(y)$. Thus, proved. □

Claim 5.9 (f is close to being affine). *f is $\frac{4\delta}{1-\lambda}$-close to an affine shift of φ.*

Proof. By Claim 5.7, for every $x \in G$, we have the following

$$\Pr_{y \in G}\Big[\varphi(x) = f(x+y) - f(y)\Big] > 1 - \frac{4\delta}{1-\lambda}.$$

Hence, by an averaging argument, there exists a $y \in G$ such that

$$\Pr_{x \in G}\Big[\varphi(x) = f(x+y) - f(y)\Big] > 1 - \left(\frac{4\delta}{1-\lambda}\right).$$

Therefore,

$$\Pr_{z \in G}\Big[f(z) = \varphi(z-y) + f(y) = \varphi(z) + (f(y) - \varphi(y))\Big] > 1 - \left(\frac{4\delta}{1-\lambda}\right),$$

and so $f(z)$ is $\frac{4\delta}{1-\lambda}$-close is φ with an affine shift of $f(y) - \varphi(y)$. □

□

References

[BSVW] Eli Ben-Sasson, Madhu Sudan, Salil P. Vadhan, Avi Wigderson: "Randomness-efficient low degree tests and short PCPs via epsilon-biased sets". STOC 2003: 612-621

[BLR] Manuel Blum, Michael Luby, Ronitt Rubinfeld: "Self-Testing/Correcting with Applications to Numerical Problems". J. Comput. Syst. Sci. 47(3): 549-595 (1993).

[GGR] Oded Goldreich, Shafi Goldwasser, Dana Ron: "Property Testing and its Connection to Learning and Approximation". J. ACM 45(4): 653-750 (1998)

[GS] Oded Goldreich, Madhu Sudan: "Locally Testable Codes and PCPs of Almost-Linear Length". FOCS 2002: 13-22

[RS] Ronitt Rubinfeld, Madhu Sudan: "Robust Characterizations of Polynomials with Applications to Program Testing". SIAM J. Comput. 25(2): 252-271 (1996)

[SW] Amir Shpilka, Avi Wigderson: "Derandomizing homomorphism testing in general groups". STOC 2004: 427-435