CS 6815: Lecture 14

Instructor: Eshan Chattopadhyay Scribes: Jason Gaitonde, Shawn Ong

October 16, 2018

1 a-Expanding Graphs

We will repeat the results from the end of last class for future reference:

Definition 1.1. *Graph G is (n, d, a)-**expanding** if G is undirected, d-regular, $|V(G)| = n$, and for every $S, T \subseteq V$, if $|S| = |T| = a$, then there is an edge between some $s \in S$ and $t \in T$.*

We also had the following bounds on the degree of such expanders:

1. In general, there is a lower bound $d \geq \frac{n}{a}$.
2. A random graph with $d = O\left(\frac{n}{a}\log n\right)$ is an a-expander with high probability.
3. By the expander mixing lemma, any spectral expander satisfies $d \geq \frac{1}{2}\left(\frac{n}{a}\right)^2$.

So spectral expanders are insufficient for reaching the probabilistic bound of $O\left(\frac{n}{a}\log n\right)$ (in particular, the best we can do is quadratic in n). Instead, we will use extractors to achieve this bound. To this end, let $Ext : \{0,1\}^r \times \{0,1\}^d \to \{0,1\}^m$ be a $\left(\delta r, \frac{1}{4}\right)$-seeded extractor.

Proposition 1.2. *Ext can be used to build an a-expanding graph.*

Recall that Ext can interpreted as a function labelling the endpoints of the edges of a left-D-regular bipartite graph with vertex sets $[R]$ and $[M]$ (we use the convention that uppercase letters take value exponential in their lowercase versions, e.g. $X = 2^x$). Let $S \subseteq [R]$, $|S| = R^\delta$, and let $\Gamma(S)$ denote the set of neighbors of S, i.e.,

$$\Gamma(S) = \{z \in [M] : \exists x \in S, \exists y \in [D], Ext(x, y) = z\} \tag{1}$$

Claim 1.3. $|\Gamma(S)| \geq \frac{3M}{4}$

Suppose otherwise; let S be such that $|\Gamma(S)| < \frac{3M}{4}$. Let X be the distribution flat on S (recall that a distribution X is flat on S if $\mathbf{Pr}[X = x] = \frac{1}{|S|}$ if $x \in S$ and 0 otherwise). Then $\underset{z\sim U_m}{\mathbf{Pr}}[z \in \Gamma(S)] = \frac{|\Gamma(S)|}{M} < \frac{3}{4}$. Observe that $\underset{x\sim U_r, y\sim U_d}{\mathbf{Pr}}[Ext(x, y) \in \Gamma(S)] = 1$ by construction. But we have:

$$H_\infty(X) = -\log\left(\frac{1}{|S|}\right) = \delta r \tag{2}$$

Since Ext is a $\left(\delta r, \frac{1}{4}\right)$-seeded extractor, it must be the case that $|Ext(X, U_d) - U_m| \leq \frac{1}{4}$ if $H_\infty(X) \geq \delta r$. But this is a contradiction, since:

$$\left|\underset{x\sim U_r, y\sim U_d}{\mathbf{Pr}}[Ext(x, y) \in \Gamma(S)] - \underset{z\sim U_m}{\mathbf{Pr}}[z \in \Gamma(S)]\right| > 1 - \frac{3}{4} = \frac{1}{4} \tag{3}$$

Claim 1.4. *If $|S|, |T| \geq R^\delta$, then $|\Gamma(S) \cap \Gamma(T)| \geq \frac{M}{2}$.*

This follows immediately from applying Claim 1.3 to both S and T, then using inclusion-exclusion on $\Gamma(S)$ and $\Gamma(T)$ (the size of their union is at most M).

Construction 1.5. *Construct an expanding graph G' from an extractor (represented by graph G) by taking $V(G') = [R]$ and adding edges between any i, j sharing a common neighbor in G.*

However, this straightforward attempt at constructing an expanding graph does not guarantee d-regularity for sufficiently small d. As a worst-case example, it could be possible that every vertex in $[R]$ has an edge to the same vertex in $[M]$. Then G would be complete, with $d = n - 1$. To prevent this and similar problems, we will delete vertices in $[M]$ with high degree, i.e. degree at least $2\frac{RD}{M}$, before performing the construction. Observe that $\frac{RD}{M}$ is the average degree of vertices in $[M]$, as there are RD edges in G. We claim that applying Construction 1.5 after removing such vertices still generates an expanding graph, but with lower degree. To be precise, let M' be the number of vertices remaining after this removal.

Claim 1.6. *$M' \geq \frac{3M}{4}$*

Suppose towards contradiction that $M' < \frac{3M}{4}$, that is, a set of vertices BAD of size $|\mathsf{BAD}| > \frac{M}{4}$ was removed, with every element of BAD having degree at least $2\frac{RD}{M}$. Then:

$$\mathop{\mathbf{Pr}}_{x \sim U_r, y \sim U_d}[Ext(x, y) \in \mathsf{BAD}] \geq 2\frac{|\mathsf{BAD}|}{M} \tag{4}$$

Note that the LHS indicates the probability that the right endpoint of a randomly chosen edge is in BAD. Since each vertex in BAD has degree at least $2\frac{RD}{M}$, and there are a total of RD edges, summing over the vertices gives the RHS.

Now, since $|\mathsf{BAD}| > \frac{M}{4}$, we can write:

$$\left| \mathop{\mathbf{Pr}}_{x \sim U_r, y \sim U_d}[Ext(x, y) \in \mathsf{BAD}] - \mathop{\mathbf{Pr}}_{z \sim U_m}[z \in \mathsf{BAD}] \right| = 2\frac{|\mathsf{BAD}|}{M} - \frac{|\mathsf{BAD}|}{M} > \frac{1}{4} \tag{5}$$

By assumption, Ext was a $\left(\delta r, \frac{1}{4}\right)$-seeded extractor. However, $H_\infty(U_r) = r \geq \delta r$, giving a contradiction. Hence, at most $\frac{M}{4}$ vertices are removed.

Proposition 1.7. *G', as constructed from G via Construction 1.5, is a $\left(R, D \cdot \frac{2RD}{M}, R^\delta\right)$-expanding graph.*

The fact that G' has R vertices is obvious. To show that $a = R^\delta$, consider 1.4; before removal, every $S, T \subseteq [R]$ with $|S|, |T| = R^\delta$ satisfied $|\Gamma(S) \cap \Gamma(T)| \geq \frac{M}{2}$. Then after the removal, at least common $\frac{M}{4}$ vertices must remain, so S and T will still share an edge after applying Construction 1.5. After removal, we can also bound the number of paths of length 2 in G with one endpoint $v \in [R]$. v has degree D and each such path has a midpoint remaining in $[M]$; each remaining vertex has degree less than $\frac{2RD}{M}$. Hence, there are no more than $D \cdot \frac{2RD}{M}$ such paths, so the degree of G' is bounded by $D \cdot \frac{2RD}{M}$ (every edge on v corresponds to such a path in G).

Proposition 1.8. *It is possible to create $(n, n^{1-\delta+o(1)}, n^\delta)$-expanding graphs.*

If we take $n = R$, it turns out that it is possible to construct extractors that will yield $\frac{2D^2}{M} \leq n^{o(1)-\delta}$. Specifically, it is possible to construct extractors with $d = O(\log r)$ and $M = n^{\delta - O(1)}$, giving $D = poly(r)$; intuitively, since D is the degree of left vertices in G, it should be substantially smaller than R. This gives the desired construction. The only issue is that these graphs are not quite regular; we can fix this by relaxing the definition of expanding graphs to ignore regularity, or by letting expanding graphs be multigraphs and add enough self-loops to enforce regularity.

While this works, for large δ, it seems likely that there should be more straightforward ways to construct expanding graphs, leading to the following open problem:

Open problem 1.9. *Is there an easy construction for $\left(n, n^{1/2-\delta}, \sqrt{n}\right)$-expanding graphs? [Kleinberg 18]*

2 Condensers, Expanders, and List-Decodable Codes

We begin with a slightly different definition than what we have been dealing with so far:

Definition 2.1. *A distribution D is ϵ-close to min-entropy k if there exists a distribution X with $H_\infty(X) \geq k$ and $|D - X| \leq \epsilon$.*

The property of being ϵ-close to min-entropy k is a little bit different than actually having min-entropy close to k; the difference is that min-entropy is a global statement that bounds the probability of all x in the support of X. On the other hand, being ϵ-close to min-entropy k allows a distribution to have a relatively high mass on a small subset, which would cause the min-entropy to rise.

Recall that the goal of extractors was to take two weak sources in terms of min-entropy and get a distribution with better min-entropy. We now consider a different kind of operation, where we want to take weak sources and output a shorter source without much entropy loss.

Definition 2.2. *A function $Con : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a $(n,k) \to_\epsilon (m,k')$ condenser if for every (n,k)-source X, $Con(X,U_d)$ is ϵ-close to min-entropy k'. We say a condenser is lossless if $k' = k$. A strong condenser is a $(n,k) \to_\epsilon (m,k')$ condenser such that $(C(X,U_d),U_d)$ is ϵ-close to a distribution with min-entropy $k'+d$ on $m+d$ bits.*

Typically, we will want $\frac{k'}{m} > \frac{k}{n}$; this means that we are getting more entropy per bit after condensing. It turns out that lossless expanders are equivalent to vertex expanders. Specifically:

Theorem 2.3. *Con is a strong, lossless $(n,k) \to_\epsilon (m+d, k+d)$ condenser if and only if the corresponding bipartite graph $([N],[D]\times[M],E)$ is a $(K,(1-\epsilon)D)$ vertex expander with left degree D, where $K = 2^k$, $N = 2^N$, $M = 2$, and $D = 2^d$, where d is the seed-length.*

Proof. In the forward direction, let $S \subseteq [N]$ such that $|S| = K$. Let X be uniform on S. Then $|\Gamma(S)| = |(Con(X,U_d),U_d)|$. We want this to be at least $KD(1-\epsilon)$ for this to be the claimed vertex expander. Suppose for a contradiction that this is false, and let Y be any $(m+d,k+d)$-source on $[M]\times[D]$. Then

$$Pr(Y \in \Gamma(S)) \leq \frac{|\Gamma(S)|}{2^{k+d}} < \frac{KD(1-\epsilon)}{KD} = 1-\epsilon, \tag{6}$$

which contradicts the assumption that there exists some $(m+d,k+d)$-source that $(Con(X,U_d),U_D)$ is ϵ-close to.

In the reverse direction, it suffices to show $(Con(X,U_D),U_D)$ is ϵ-close to a source distribution with min-entropy $k+d$ for any flat source X. Let $S \subseteq N$ be such that $|S| = K$ and take $X = U_S$. By the fact that the corresponding bipartite graph is a $([N],[D]\times[M],E)$ vertex expander, it follows that $|\Gamma(S)| \geq (1-\epsilon)KD$. But as the graph is left D-regular, there are exactly KD edges leaving S; therefore, by redirecting just ϵKD edges, we could ensure that all edges from S go to distinct neighbors. This gives a uniform distribution on KD vertices, which is a $(m+d,k+d)$-source, and therefore $(Con(X,U_d),U_d)$ is ϵ-close to a $(k+d)$-source. □

Next class, we will see that there is a connection between list-decodable codes and strong-seeded extractors. That is, we will show:

Theorem 2.4. *Let* $C : [N] \to [M]^D$ *be a* $(1 - \frac{1}{M} - \epsilon, L)$ *list-decodable code. Then* $Ext : [N] \times [D] \to [M]$ *defined by*

$$Ext(x, y) = C(x)|_y \tag{7}$$

is a strong-seeded extractor for min-entropy $k = \log(L) + \log(1/\epsilon)$ *with error* $M\epsilon$.