CS 6815 Pseudorandomness and Combinatorial Constructions **Fall 2018**

Lecture 16: October 23

Instructor: Eshan Chattopadhyay *Scribe: Wei-Kai Lin (wl572), Jyun-Jie Liao (jl3825)*

Optimal Vertex Expander. In this lecture, we first construct an "optimal" vertex expander. Recall that a graph G is a (K, A)-vertex expander iff for all vertex set $|S| \leq K$, it holds that $|\Gamma(S)| \geq A \cdot |S|$.

Let q be a power of prime, $\mathbb{F}_q$ be the field of q, $n, m, h \in \mathbb{N}$ be parameters that will be chosen later. We represent $f \in \mathbb{F}_q^n$ as a polynomial over $\mathbb{F}_q$ that has degree at most $n-1$, and we choose an irreducible polynomial E over $\mathbb{F}_q$ of degree n. For all $f \in \mathbb{F}_q$, let $f_i := f^{h^i} \mod E$.

The construction is a bipartite graph G consisting of $N = q^n$ vertices of the left and $M \times D$ vertices on the right, where $M = q^m$ and $D = q$, and the vertex degree on the left side is D. The left and right vertex sets are chosen to be $\mathbb{F}_q^n$ and $[D] \times \mathbb{F}_q^m$, and we describe the edge set by the mapping $e : \mathbb{F}_q^n \times [D] \mapsto [D] \times \mathbb{F}_q^m$ such that for a left vertex $f \in \mathbb{F}_q^n$, the edge $y \in [D]$ goes to the following right vertex,

$$e(f, y) = (y, f_0(y), f_1(y), \ldots, f_{m-1}(y)).$$

Claim 1. *G is a $(h^m, q - (n-1)(h-1)m)$-vertex expander.*

Proof. Let L and R be the set of left and right vertices on G correspondingly. For any $T \subseteq R$, define $\text{LIST}(T) := \{x \in L : \Gamma(x) \subseteq T\}$. To show a graph is a (K, A)-vertex expander, it suffices to show that for all subset T of the right vertex set such that $|T| = AK - 1$, it holds that $|\text{LIST}(T)| \leq K - 1$.

Given such $T \subseteq [D] \times \mathbb{F}_q^m$, let

$$Q(Y, Y_0, Y_1, \ldots, Y_{m-1})$$

be a polynomial on $\mathbb{F}_q$ that vanishes on T. Observe that Q can be decomposed into the following summation of monomials,

$$\sum_{i=0}^{A-1} \sum_{j=0}^{K-1} \beta_{i,j} Y^i M_j(Y_0, \ldots, Y_{m-1}),$$

where $K = h^m$, $A = q - (n-1)(h-1)m$, and $\beta_{i,j} \in \mathbb{F}_q$ are coefficients. Rearranging, we get

$$\sum_{j=0}^{K-1} P_j(Y) M_j(Y_0, \ldots, Y_{m-1}),$$

where there exists j s.t. $P_j(Y)$ such that is not divisible by $E(Y)$. On the other side, fix any $f \in \text{LIST}(T)$. Let $R_f(Y) := Q(Y, f_0(Y), \ldots, f_{m-1}(Y))$. Then, by Q vanishes on T and $f \in \text{LIST}(T)$, it holds that $R_f(y) = 0$ for all $y \in \mathbb{F}_q$. In addition, the degree of R_f is $(A-1) + m(n-1)(h-1) < q$, which implies that $R_f(Y)$ is a zero polynomial. Now, define

$$W(Z) := Q\left(Y, Z, Z^h, Z^{h^2}, \ldots, Z^{h^{m-1}}\right)$$

on $\mathbb{F}_q$. Then, $f(Y)$ is a root of $W(Z)$ as $W(f(Y)) = R_f(Y) = 0$. Finally, note that

$$W(Z) = \sum_{j=0}^{K-1} P_j(Y) Z^j \mod E(Y).$$

It follows that $|\text{LIST}(T)|$ is at most the degree of $W(Z)$, which is at most $K - 1$. $\square$

Theorem 1. *For any N, $K \leq N$, $\epsilon, \alpha > 0$, there exists a (K, A)-vertex bipartite expander such that consists of N left vertices, M right vertices, and left vertex degree D, where $A \geq (1-\epsilon)D$, $M \leq D^2 K^{1+\alpha}$, and $D = \left(\frac{\log N \cdot \log K}{\epsilon}\right)^{1+1/\alpha}$.*

Proof. In the graph G of Claim 1, pick $h = \left(\frac{\log N \cdot \log K}{\epsilon}\right)^{1/\alpha}$, $q \in [h^{1+\alpha}, 2h^{1+\alpha}]$, n such that $q^n = N$, and m such that $q^{m+1} = M$. □

Lossless Condenser from Vertex Expander. Note that the vertex expansion A of the (K, A)-vertex expander is ϵ close to the degree D, and that is why we called it "optimal". Next, we recall the theorem from Lecture 14 that states a $(K, (1-\epsilon))$-vertex expander implies (and also the converse) a strong lossless condenser, we claim the following corollary.

Corollary 1. *For all n, k, d, any constant $\epsilon, \alpha > 0$, there exists a lossless condenser $Con : \{0,1\}^n \times \{0,1\}^d \mapsto \{0,1\}^m$ such that takes an (n, k)-source and outputs an $(m, k+d)$-source with error ϵ, where $d = (1+1/\alpha)(\log n + \log k + \log \frac{1}{\epsilon}) + O(1)$, $k \leq m \leq 2d + (1+\alpha)k$.*

Extractor for Arbitrary Entropy. Now we have a condenser. Combining a condenser with an extractor for high-entropy source, we can get an extractor for arbitrary entropy.

Theorem 2. *Suppose $Con : \{0,1\}^n \times \{0,1\}^{d_1} \to \{0,1\}^{k'}$ is a $(n, k) \to_{\epsilon_1} (k', (1-\delta)k')$ condenser, and $Ext^\delta : \{0,1\}^{k'} \times \{0,1\}^{d_2} \to \{0,1\}^m$ is a $((1-\delta)k', \epsilon_2)$ extractor, Then $Ext : \{0,1\}^n \times \{0,1\}^{d_1+d_2} \to \{0,1\}^m$, defined by*

$$Ext(x, (y_1, y_2)) = Ext^\delta(Con(x, y_1), y_2),$$

is a $(k, \epsilon_1 + \epsilon_2)$ extractor.

We give a construction of Ext^δ with seed length $O(\log n)$ for $\delta = O(\epsilon^2)$ below.

Theorem 3. *Suppose G is a $([2^m], 2^c, \lambda)$ expander for some constant c, λ, and $\epsilon > 0$ is an error parameter. Define a function $Ext^\delta : \{0,1\}^n \times [L] \to \{0,1\}^m$, where $L = \frac{n-m}{c}$ and $Ext^\delta(x, y)$ is generated as follow:*

1. *Take the first m bits of x to be $z_0 \in [2^m]$, and parse the remaining $(n-m)$ bits as $e_1, e_2, \ldots, e_{\frac{n-m}{c}}$, where each e_i is a c-bit integer.*
2. *Start a y-step walk from u on G, following the sequence of edge labels $(e_1, \ldots, e_y)$. Output the destination $v \in [2^m]$.*

Then Ext^δ is a $((1-\delta)n, \epsilon)$ extractor, where $\delta = O(\epsilon^2)$ and $m = O(n)$.

Proof. Consider any statistical test for the output $Ext^{0.99}$, interpreted as a set $S \subseteq \{0,1\}^m$. Consider a uniformly random source $X \in \{0,1\}^n$ and a uniformly random seed Y. Let Z_i denote the indicator for the event $Ext(X, i) \in S$. Then $\Pr[Ext(X, Y) \in S] = \frac{1}{L}\sum_{i \in [L]} Z_i$. By the Chernoff bound on expander graph,

$$\Pr\left[\left|\frac{1}{L}\sum_{i \in [L]} Z_i - \frac{|S|}{2^m}\right| > \epsilon/2\right] \leq 2\exp\left(\frac{-(1-\lambda)\epsilon^2 L}{16}\right)$$

Now consider any $(1-\delta)n$-source X', and define Z'_i similarly as above. Then

$$\Pr\left[\left|\frac{1}{L}\sum_{i\in[L]} Z'_i - \frac{|S|}{2^m}\right| > \epsilon\right] \leq 2\exp\left(\frac{-(1-\lambda)\epsilon^2 L}{16}\right)\cdot 2^{\delta n} \leq \epsilon/2$$

for proper choice of δ and m. Therefore S cannot distinguish $Ext(X', Y)$ from uniform with advantage more than ϵ. □