

Lecture 16: Generalization of Extractors

November 14, 2004

Scribe: Habil Zare

1 Conductors

We generalize the definition of extractors in this way:

Definition 1 $\text{Con} : \{0, 1\}^n \times \{0, 1\}^d \rightarrow \{0, 1\}^m$ is a (k, ϵ, a) -conductor if, $\forall k$ -source X :

$$\text{Con}(X, U_d) \stackrel{\epsilon}{\approx} (k + a) - \text{source}$$

Remark 2 In the above definition we want $m < n$. Also notice that:

- (1) If $k + a = m$, we get a standard (k, ϵ) -extractor.
- (2) If $a = d$, we get a lossless conductor.

The following theorem shows that there is a two-sided relation between conductors and left D -regular bipartite graphs.

conductors \longleftrightarrow left D -regular bipartite graphs

Theorem 3 Con is a (k', ϵ, a) -conductor, $\forall k' < k$, iff its bipartite graph is a $(k, (1 - \epsilon)D)$ -expander.

Proof: (\implies) Take any left set S , $|S| = k' \leq k$. Let $T = N(S)$ be the set of all neighbors of S .

$$|\Pr[\text{Con}(S, U_A) \in T] - \Pr[Y \in T]| \leq \epsilon$$

where Y is a $(k' + d)$ -source which is ϵ -close to Con .

$$1 - \Pr[Y \in T] \leq \epsilon$$

$$1 - \frac{|T|}{k'D} \leq \epsilon \implies \frac{|N(S)|}{|S|} \geq (1 - \epsilon)D$$

This means that we have almost D expansion ($(1 - \epsilon)D$ indeed) in a D -regular graph. ■

According to the above proof, the expansion is better than Ramanujan graphs. Let's compare it with extractors:

$$\begin{aligned} M &\leq KD\epsilon^2 \\ \implies \frac{M}{K} &\leq D\epsilon^2 \end{aligned}$$

since a set of size K can not expand to a set of size more than M . Mainly, because extractors have significant entropy loss, the best expansion for them is $\epsilon^2 D$ while here we have $(1 - \epsilon)D$ using conductors.

Remark 4 Doing zig-zag product on conductors, we can achieve: $\forall \alpha, \epsilon > 0, M = \alpha N, D = O(1), K = \Omega(N)$.

Open problem: Construct non-bipartite expanders with vertex expansion more than $\frac{\Delta}{2}$. ($\frac{\Delta}{2}$ is achieved by Ramanujan graphs.)

Open problem: Construct expanders with vertex expansion more than $(1 - \epsilon)D$ (say: $D - O(1)$).

2 Pseudorandom Generators

Intuitively, pseudorandom generators are efficient algorithms $G : \{0, 1\}^{l(n)} \rightarrow \{0, 1\}^n$ such that $G(U_l)$ is *computationally distinguishable* from U_m .

Definition 5 Random variables X_n, Y_n are $\epsilon(n)$ -indistinguishable for time $t(n)$, denoted by $X_n \stackrel{(\epsilon, t)}{\approx} Y_n$, if for any probability time $t(n)$ algorithm T and for all large n :

$$|\Pr[T(X_n) = 1] - \Pr[T(Y_n) = 1]| \leq \epsilon$$

where the probabilities are over the random choices of T as well as X_n and Y_n .

Note: Here $T^{-1}(1)$ is our statistical test.

Notation: We write $X_n \stackrel{c}{\equiv} Y_n$ if for all polynomial $t(n)$ and $\epsilon = 1/t(n)$, $X_n \stackrel{(\epsilon, t)}{\approx} Y_n$.

Definition 6 (Non-uniform indistinguishability) Random variables X_n, Y_n are $\epsilon(n)$ -indistinguishable for non-uniform time $t(n)$, if for any non-uniform time $t(n)$ algorithm T (e.g. a boolean circuit of size $t(n)$) and for all large n :

$$|\Pr[T(X_n) = 1] - \Pr[T(Y_n) = 1]| \leq \epsilon$$

Definition 7 X_n is pseudorandom if $X_n \stackrel{c}{\equiv} Y_n$.

Definition 8 $G : \{0, 1\}^{l(n)} \rightarrow \{0, 1\}^n$ is a pseudorandom generator if $G(U_m) \stackrel{c}{\equiv} U_m$ (or more formally, $G(1^n, U_{l(n)}) \stackrel{c}{\equiv} U_m$) and G is computable in polynomial time. Usually $l(n) \ll n$.

Theorem 9 If $X_n \stackrel{c}{\equiv} Y_n$ non-uniformly, $\forall k = \text{poly}(n)$, $X_n^k \stackrel{c}{\equiv} Y_n^k$ non-uniformly. Here $X_n^k = X_n \times \dots \times X_n$.

Proof: Notation, will drop subscript n . Suppose there is a poly-time non-uniform algorithm T such that:

$$|\Pr[T(X^k) = 1] - \Pr[T(Y^k) = 1]| > \epsilon$$

with out loss of generality, we can drop the absolute sign. (otherwise consider $-T$). Interpolate between X^k and Y^k :

$$\begin{aligned} X \dots XX &= H_0 \\ X \dots XY &= H_1 \\ &\vdots \\ XY \dots Y &= H_{k-1} \\ Y \dots Y &= H_k \end{aligned}$$

To show that:

$$\begin{aligned}
& \left| \sum_{i=1}^k \Pr[T(H_{i-1}) = 1] - \Pr[T(H_i) = 1] \right| \\
&= \left| \Pr[T(H_0) = 1] - \Pr[T(H_k) = 1] \right| \\
&= \left| \Pr[T(X^k) = 1] - \Pr[T(Y^k) = 1] \right| > \epsilon
\end{aligned}$$

which implies there is at list one $0 \leq i \leq k$ such that:

$$\epsilon_i \stackrel{\text{def}}{=} \left| \Pr[T(X^{k-i}XY^{i-1}) = 1] - \Pr[T(X^{k-i}YY^{i-1}) = 1] \right| > \frac{\epsilon}{k}$$

Look at this averaging:

$$\text{Ave} \left| \Pr[T(X^{k-i}XY^{i-1}) = 1] - \Pr[T(X^{k-i}YY^{i-1}) = 1] \right| > \frac{\epsilon}{k}$$

where the average is taken over all $x_1, \dots, x_{k-i} \leftarrow X^{k-i}$ and $Y_1, \dots, Y_{i-1} \leftarrow Y^{i-1}$. So there exists $(x_1, \dots, x_{k-i}, y_1, \dots, y_{i-1})$ such that the above equality is preserved. Define:

$$T'(z) = T(x_1, \dots, x_{k-i}, z, y_1, \dots, y_{i-1})$$

T' can distinguish between X and Y , a contradiction! ($\text{time}(T') = \text{time}(T) + O(nk)$) ■

Theorem 10 *If $X_n \stackrel{c}{\equiv} Y_n$ uniformly, then $X_n^k \stackrel{c}{\equiv} Y_n^k$, assuming X and Y are efficiently sam-pleable. (e.g. there is a probabilistic polynomial time algorithm M such that $M(1^n) \equiv X_n$.)*

Proof: Recall that in the previous proof, we had: $\sum_{i=1}^k \epsilon_i \geq \epsilon$.

$T'(z)$: randomly pick $0 \leq i \leq k$,

randomly sample $x_1, \dots, x_{k-i} \leftarrow X^{k-i}$

randomly sample $Y_1, \dots, Y_{i-1} \leftarrow Y^{i-1}$

and output $T(x_1, \dots, x_{k-i}, z, y_1, \dots, y_{i-1})$.

Similarly we get:

$$\left| \Pr[T'(X) = 1] - \Pr[T'(Y) = 1] \right| \geq \frac{1}{k} \sum_{i=1}^k \epsilon_i > \frac{\epsilon}{k}$$

A contradiction! ■