# Lecture 5

Lecturer: Irit Dinur
Preliminary Scriber: Karthik C. S.
Scribe Upgrader: Inbal Livni

Previously, we discussed the following theorem.

**Theorem 1.** *Fix* $\Sigma = \{0, 1\}$. *The complete complex (k-dimensional, n-vertices) is an agreement expander, i.e.,* $\exists c > 0,\ \gamma_{\frac{1}{2}}(X(n, k)) > c$.

We explained that it is tight, and outlined an approach toward proving it. Today we will give a proof.

Given $f \in \mathfrak{a}_{k-1}$, and $\alpha_{\frac{k}{10}}(f) > 1 - \delta$, we will find $g \in \mathfrak{a}_0$ such that the following holds:

$$\Pr_{S \in X(k-1)}[f(S) = g|_S] \geq 1 - O(\delta).$$

Let $A_1 \in X(t-1)$, and let $A_1 \subset A \in X(t+r-1)$. We define

$$\beta_{t,r}(A_1, A) = \Pr_{A \subset S, S' \in X(k-1)}[f(S)_{A_1} = f(S')_{A_1}], \qquad \beta_{t,r} = \mathbb{E}_{A_1 \subset A} \beta_{t,r}(A_1, A).$$

Let $L(A_1, A, \sigma)$ be the set of all $k$-sets containing $A$ such that the assignment of $A_1$ given by $f$ is equal to $\sigma$,

$$L(A_1, A, \sigma) = \{S \in X(k-1) \mid S \supset A,\ f(S)|_{A_1} = \sigma\}$$

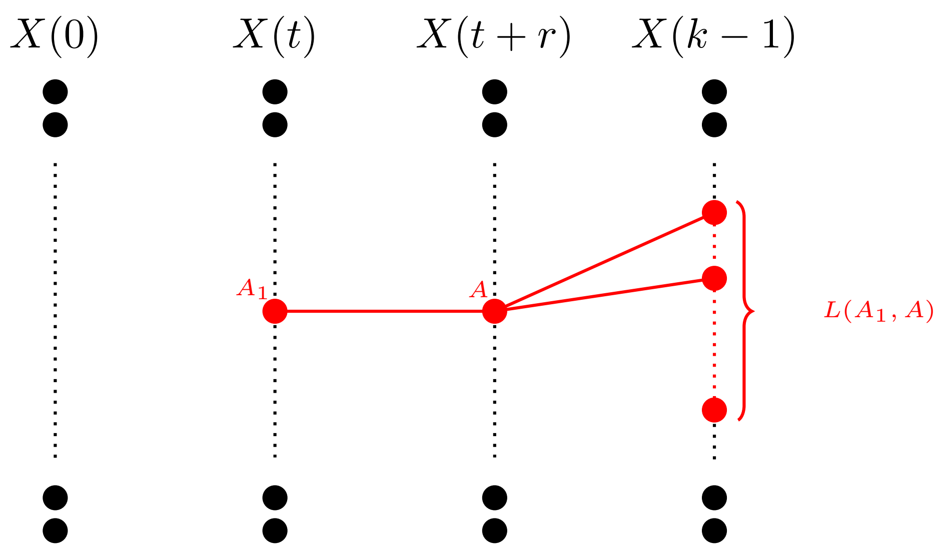


Figure 1: $L(A_1, A)$ are all $k$-sets containing $A$, edges denotes containment.

We have the following lemmas:

**Lemma 2.** *There are* $\frac{k}{20} \le t \le \frac{k}{10}$ *and* $r \le \frac{k}{10}$ *such that* $\beta_{t,r} \approx \beta_{t,r+1} \approx \beta_{t+1,r}$*, where* $\approx$ *is up to multiplicative factors of order* $1+\eta$ *for* $\eta = O(\frac{\delta}{k})$*.*

For each $A_1, A$ and each $\sigma$ a partial coloring of $A_1$, we define a coloring of $X(0)$ by taking the most popular value among $S \in L(A_1, A, \sigma)$ that contain $x$:

$$\forall x \in X(0), \quad g_{A_1,A,\sigma}(x) = argmax_{a\in\Sigma} \Pr_S[f(S)(x) = a \mid S \in L(A_1, A, \sigma)]$$

**Lemma 3.** *If* $A_1 \subset A$ *are such that* $0.9 \le \beta_{t,r}(A_1, A) \approx \beta_{t,r+1}(A_1, A) \approx \beta_{t+1,r}(A_1, A)$ *then, there exists unique* $\sigma$ *such that* $g_{A_1,A,\sigma}$ *"explains* $L(A_1, A, \sigma)$*",i.e.,*

$$\Pr_{S\in X(k-1)}[f(S) = g_{A_1,A,\sigma}|_S| \; S \in L(A_1, A, \sigma)] \ge 1 - O(\delta).$$

**Lemma 4.** *There is one* $g : X(0) \to \Sigma$*, such that the following holds for at least* $1 - O(\delta)$ *fraction of the sets* $(A_1, A)$*:*

$$\Pr_{x\in X(0)}[g(x) = g'_{A_1,A}(x)] \ge 1 - \frac{\delta}{k}.$$

*Proof of Lemma 2.* The lemma follows from two steps: the monotonicity of $\beta_{t,r}$ and a "potential function" argument.

**Claim 5** (Monotonicity)**.** $\beta_{t+1,r} \le \beta_{t,r+1}$ *and* $\beta_{t,r} \le \beta_{t,r+1}$*.*

*Proof.* We first observe that for every $A_1 \subset A$ and $x \notin A$,

$$\beta_{t+1,r}(A_1 \cup \{x\}, A \cup \{x\}) \le \beta_{t,r+1}(A_1, A \cup \{x\}).$$

This is true because the distributions on pairs $S, S'$ in the above expressions are identical, and also for every pair of subsets $S, S' \supset A \cup \{x\}$ if $f(X)|_{A_1\cup\{x\}} = f(S')|_{A_1\cup\{x\}}$ then surely $f(X)|_{A_1} = f(S')|_{A_1}$. This establishes $\beta_{t+1,r} \le \beta_{t,r+1}$.

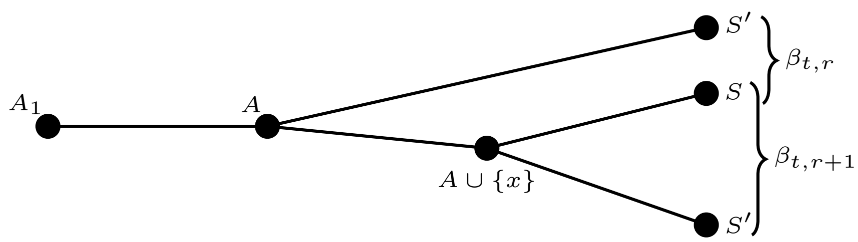


Figure 2: $\beta_{t,r}$ compares to $\beta_{t,r+1}$.

To see that $\beta_{t,r} \le \beta_{t,r+1}$ we fix $A_1, A$ and show that

$$\beta_{t,r}(A_1, A) \le \mathop{\mathbb{E}}_{x\notin A} \beta_{t,r+1}(A_1, A \cup \{x\}),$$

where the expression on the right should be thought of as "$\beta_{t,r+1}(A_1, A)$".

For each $\sigma \in \Sigma^t$, let $p(\sigma) = \Pr_{S\supset A}[S \in L(A_1, A, \sigma)]$ and observe that

$$\beta_{t,r}(A_1, A) = \sum_{\sigma\in\Sigma^t} p(\sigma)^2$$

Let us denote for each $x \notin A$ and each $\sigma \in \Sigma^t$, $p_x(\sigma) = \Pr_{S \supset \{x\} \cup A}[S \in L(A_1, A, \sigma)]$.Then $p(\sigma) = \mathbb{E}_{x \notin A}\, p_x(\sigma)$ and

$$\beta_{t,r+1}(A_1, A) = \sum_{\sigma \in \Sigma^t} \mathop{\mathbb{E}}_{x \notin A}[p_x(\sigma)^2]$$

Now the proof follows because for each $\sigma$, by convexity,

$$\mathop{\mathbb{E}}_{x \notin A}[p_x(\sigma)^2] \geq [\mathop{\mathbb{E}}_{x \notin A}[p_x(\sigma)]]^2 = p(\sigma)^2.$$

□

We complete the proof of the lemma using a potential function argument. Note that $\alpha_t = \beta_{t,0}$ for $t = k/10$, we try to either increase $r$, or increase $r$ and decrease $t$, and continue as long as $\beta_{t,r}$ increases by a multiplicative factor of at least $(1 + \eta)$. We stop when such an increase cannot be obtained. We know that $1 - \delta \leq \beta_{t,0} \leq 1$ so after $O(k)$ steps we must stop.

Formally, define $(a_0, b_0) = (k/10, 0)$. At each step $i = 1, \cdots, k/10$, we have that either $(a_{i+1}, b_{i+1}) \leftarrow (a_i, b_i + 1)$ or $(a_{i+1}, bi + 1) \leftarrow (a_i - 1, b_i + 1)$, as long as the growth is greater than a multiplicative factor of $1 + O(\frac{\delta}{k})$. By the monotonicity claim $\beta_{a_{i+1}, b_{i+1}} \geq \beta_{a_i, b_i}$ always, and until we halt we also have $\beta_{a_i, b_i} \geq (1+\eta)^i(1-\delta)$ and this must not exceed 1 so must halt after $k/20$ steps (for an appropriate choice of $\eta$). □

*Proof of Lemma 3.* Let $A_1 \subset A$, and let $\sigma$ be the value that maximizes $p(\sigma) = \Pr_{S \supset A}[S \in L(A_1, A, \sigma)]$. We first claim that there must be some $\sigma$ for which $p(\sigma) \geq \beta_{t,r}(A_1, A)$. This is because

$$\beta_{t,r}(A_1, A) = \sum_\sigma p(\sigma)^2 \leq \sum_\sigma p(\sigma) \cdot \max_\sigma p(\sigma) = \max_\sigma p(\sigma)$$

where we have used $\sum_\sigma p(\sigma) = 1$. If $0.9 \leq \beta_{t,r}$ then clearly this $\sigma$ is unique.

Next, we claim that if

$$\begin{aligned} \beta_{t,r}(A_1, A) &\approx \mathbb{E}_{x \not\subseteq A} \beta_{t,r+1}(A_1, A \cup x) \\ &\approx \mathbb{E}_{x \not\subseteq A} \beta_{t+1,r}(A_1 \cup \{x\}, A \cup \{x\}) \end{aligned}$$

then $g = g_{A_1, A, \sigma}$ satisfies the claim of the lemma, i.e.

$$\Pr_{S \in L(A_1, A, \sigma)}[f(S) = U^{k-1} g(S) \mid S \in L(A_1, A, \sigma)] \geq 1 - O(\delta).$$

We will estimate the following to be greater than $1 - \frac{\delta}{k}$:

$$\Pr_{S \in L(A_1, A\, \sigma), x \in S}[f(S)_x = g_{A_1, A}(x)] \geq 1 - \delta/k$$

and the proof follows by a union bound over $k$ elements $x \in S$.

For each $x \notin A$, we consider the minority set of all $S$'s that belong to $L(A_1, A \cup \{x\}, \sigma)$ for which $f(S)_x \neq g(x)$. These sets are the ones that contribute to the relative difference between $\beta_{t+1,r}(A_1, A \cup \{x\})$ and $\beta_{t,r+1}(A_1, A \sup\{x\})$. The claim will follow since on an average $x$ this difference is on the order of $\delta/k$.

...FILL IN... □

*Proof of Lemma 4.* Consider the bipartite graph with all $(A_1, A)$ as the left vertices and all subsets $S$ of $X(k-1)$ as the right vertices. We have the following:

$$\Pr_{A \cup A' \subset S}[g_{A_1, A}|_S = f(S) = g_{A'_1, A'}|_S] \geq 1 - 2\delta.$$

This is simply because each of the two equalities holds with probability $1-\delta$ and we take the union of the bad events.

Fix two pairs $(A_1, A)$ and $(A'_1, A')$. If $g_{A_1,A}$ differs from $g_{A'_1,A'}$ on an $\epsilon$ fraction of its inputs, then the probability that they agree on a random $k$ set is at most $(1-\epsilon)^k$. Note that for these estimates we rely on the fact that $n \ll k$ so the elements of a set can be thought of as chosen independently without replacement. Since for a typical such pair we know that this probability is at least $1-\delta$, we deduce that $\epsilon = O(\delta/k)$.

So a typical pair of functions agree on $1-\delta/k$ of their domains, and we shall define $g = g_{A_1,A}$ for $(A_1, A)$ that maximizes this.

...FILL IN...

□