CSE 533: The PCP Theorem and Hardness of Approximation **(Autumn 2005)**

Lecture 16: Hardness of E3-LIN2, Håstad's 3-query PCP

Nov. 23, 2005

Lecturer: Venkatesan Guruswami *Scribe: Chris Ré*

1 Overview

In this lecture, we first prove a result about functions which pass the long code test. We will then use this to construct Håstad's PCP, in which we blend all the checks we need into one small (3 Query) check. The rest of the lecture is spent proving this result.

2 The Long Code Test

We begin with a long code test, where our goal is that functions which pass this test with high probability are close to a linear function with small support. We do this by examining our old linearity test but add in a bit of noise, in the form a string μ. This little bit of noise will allow us to conclude that the support for the function is small and further is under the control of out ϵ and δ parameters.

Definition 2.1. *We say a string $\mu \in \{-1,1\}^m$ is picked randomly with ϵ bias if for each i.*
$\mu_i = \begin{cases} 1 & \text{with prob } 1 - \frac{\epsilon}{2} \\ -1 & \text{with prob } \frac{\epsilon}{2} \end{cases}$.

Recall the long code, $\texttt{LONG} : \{1 \dots m\} \to (\{-1,1\}^m \to \{-1,1\})$. so that $\texttt{LONG}(a)(x) = x_a$ And the long code test, pick $x, y \in \{0,1\}^m$ at random and then pick $\mu \in \{-1,1\}^m$ with ϵ-bias.

Check that $A(x)A(y)A(xy\mu) = 1$. Now we analyze the probability that the test accepts.

$$\begin{aligned}
\mathbf{Pr}[\text{accepts}] &= \mathop{\mathbf{E}}_{x,y,\mu}[1 + \frac{1}{2}A(x)A(y)A(xy\mu)] \quad \text{(Indicator test)} \\
&= \frac{1}{2} + \frac{1}{2}\mathop{\mathbf{E}}_{x,y,\mu}[A(x)(y)(xy\mu)] \quad \text{(Linearity of Expectation)} \\
&= \frac{1}{2} + \frac{1}{2}\mathop{\mathbf{E}}_{x,y,\mu}[\sum_{S,T,U}\hat{A}(S)\hat{A}(T)\hat{A}(U)\chi_S(x)\chi_T(y)\chi_U(xy\mu)] \quad \text{(Fourier expansion)} \\
&= \frac{1}{2} + \frac{1}{2}\sum_{S,T,U}\hat{A}(S)\hat{A}(T)\hat{A}(U)\mathop{\mathbf{E}}_{x,y,\mu}[\chi_S(x)\chi_T(y)\chi_U(xy\mu)] \quad \text{(Linearity of expectation)} \\
&= \frac{1}{2} + \frac{1}{2}\sum_{S,T,U}\hat{A}(S)\hat{A}(T)\hat{A}(U)\mathop{\mathbf{E}}_{x}[\chi_{S\Delta U}(x)]\mathop{\mathbf{E}}_{y}[\chi_{T\Delta U}(y)]\mathop{\mathbf{E}}_{\mu}[\chi_U(\mu)] \\
&= \frac{1}{2} + \frac{1}{2}\sum_{S}\hat{A}(S)^3\ \mathbf{E}[\chi_U(\mu)] \quad \text{(terms survive only when } S = T = U\text{)}
\end{aligned}$$

Now $\mathbf{E}_\mu[\chi_U(\mu)] = \mathbf{E}_\mu[\prod \mu_i] = \prod \mathbf{E}_\mu[\mu_i] = (1-\epsilon)^{|S|}$. We used the fact here that the μ_i bits are selected independently and are able to conclude that its expected value only depends on the support size, $|S|$.

$$\begin{aligned}
\mathbf{Pr}[\text{test accepts}] &= \frac{1}{2} + \frac{1}{2}\sum_S \hat{A}(S)^3(1-\epsilon)^{|S|} \quad \text{(plugging in above)} \\
&\leq \frac{1}{2} + \frac{1}{2}\max \hat{A}(S)(1-\epsilon)^{|S|}\sum_S \hat{A}(S)^2 \\
&= \frac{1}{2} + \frac{1}{2}\max \hat{A}(S)\ (1-\epsilon)^{|S|} \quad \text{(by Parseval's).}
\end{aligned}$$

Now suppose our accept probability is at least $\frac{1}{2} + \delta$. This implies there exists a subset S with $|S| = O(\frac{1}{\epsilon}\log\frac{1}{\delta})$ and $\hat{A}(S) \geq 2\delta$. So the table A is close to χ_S where the size of S is small and determined by our choice of ϵ and δ.

3 Håstad's 3-Query PCP

We want to develop a 3-Query PCP, whose test is xoring the bits we query. That is all tests are of the form $x_1 \oplus x_2 \oplus x_3 = b$. The big idea here is that you can bundle all of the necessary checks into one check. This is significant because before prior to Håstad's work, the constructions would separately check closeness to a long code, and if this check passes then check the additional projection property of the label cover instance. This is similar to the approach we took for the Hadamard code based assignment tester. Since we now care for the optimal trade-off between number of queries and the soundness, such a two-step check does not quite suffice.

Now consider an instance of $\mathsf{GAP\text{-}LC}_{1,\gamma}(\Sigma)$. Verifying an assignment here amounts to checking that an edge $e = (u, v)$ in this bipartite graph is correctly labeled according to some relation π_e. To perform this verification, we get the label for u, $\sigma(u)$, and the label for v, $\sigma(v)$, then test $\pi_e(\sigma(u)) = \sigma_v$. This test makes only 2 queries, for each node label. However, the queries are over the larger alphabet Σ. We would like to reduce the number of *bits* read to exactly 3, though we can increase the soundness to $1/2$ (from the arbitrarily low soundness for the label cover instance).

Once the verifier picks $e = (u, v) \in E$ at random, for notational convenience, let A and B be the long codes of $\sigma(v)$ and $\sigma(u)$ and let $\pi = \pi_e$ with $e = (u, v)$. Let $m = |\Sigma|$. Pick further pick $x, y \in \{-1, 1\}^m$ at random. Suppose now that B really $\mathsf{LONG}(\sigma(u))$ and A really is $\mathsf{LONG}(\sigma(v))$. Then $A(x)$ should be $x_{\sigma(v)} = x_{\pi(\sigma(u))}$ and $B(x \circ \pi) = (x \circ \pi)_{\sigma(u)} = x_{\pi(\sigma(u))}$. Thus the following test:

$$A(x)B(y)B((x \circ \pi)y) = 1$$

would actually be testing $A(x)B(y)B((x \circ \pi)y) = (x_{\sigma(v)})(y_{\sigma(u)})(x_{\sigma(\pi(u))}y_{\sigma(u)}) = 1$ where $\sigma(v) = \pi(\sigma(u))$. Thus the test will accept and thus has completeness 1. We know from the problem set that such a test is doomed to failure. Drawing on Sec. 2, we add in a little bit of noise. We therefore use the following test:

$$A(x)B(y)B((x \circ \pi)y\mu) = 1$$

where $\mu \in \{1, -1\}^m$ is picked with ϵ bias.

3.1 Completeness

Here we look at what happens when we write down the correct long codes, $A = \mathsf{LONG}(\sigma(v))$, $B = \mathsf{LONG}(\sigma(u))$ and $\pi(\sigma(u)) = \sigma(v)$. Substituting in and noting that the long codes are linear, we get:

$$\begin{aligned} A(x)B(y)B((x \circ \pi)y\mu) &= x_{\sigma(v)}\ y_{\sigma(u)}\ (x \circ \pi)_{\sigma(u)}\ y_{\sigma(u)}\ \mu_{\sigma(u)} \\ &= \mu_{\sigma(u)} \end{aligned}$$

Therefore we succeed when $\mu_{\sigma(u)} = 1$ and this happens with probability $1 - \frac{\epsilon}{2} > 1 - \epsilon$.

4 Soundness

4.1 Derailed by all ones in the table

Now we would like to continue our proof by arguing about soundness, but there is a problem. To see the problem we think about what happens to our test if instead of long codes, the malicious prover writes out blocks of all 1s.

Unfortunately, we see that our current test passes every time - so before we can tackle soundness we have to deal with the problem of all 1s in the table. We observe that such tables could not

possibly be real long codes because Long Codes are balanced, that is $\mathsf{LONG}_a(x) = -\mathsf{LONG}_a(-x)$. But how can we ensure this without making an extra query?

One idea is to just force the tables to balanced. More specifically, we can define a convention to never read half the table and substitute any query of bits from the other half of the table using the above formula. So for example, our query location within the long code is $q = (q_1, \ldots, q_n)$ and $q_1 = 1$ then read location q and return its value. If $q_1 = -1$ then read location $q = (q_1, q_2, \ldots, q_n)$ and return the negation of what we read.

This trick is called enforcing the "Folding property". When a function expressed as a table of values has the property that $A(x) = -A(-x)$, we say it's folded. It should be clear that our convention above ensures the folding property.

4.2 Soundness Continued

We will ensure the folding property of the previous section, and so should correctly speak about access to tables A' and B' which are accessed like above to ensure the folded property. However, we will just keep A and B and assume that they have the folded property in what follows to keep our notation free from excess clutter.

One can show the following lemma about functions which have a folded representation:

Lemma 4.1. *If A is a folded table, then $\hat{A}(S) = 0$ when $|S|$ is even. In particular, $\hat{A}(\emptyset) = 0$.*

However, we only need that $\hat{A}(\emptyset) = 0$. This term is the expected value of $A(x)$ for a random x, which clearly equals 0 for folded functions (since $A(x) + A(-x) = 0$ for every x).

Before proceeding with the soundness analysis, we first make one definition and show a fact about it to aid in our proof:

Definition 4.2. $\pi_2(S) = \{j \in \{1 \ldots m\} | \textit{exists odd number of } i \in S \textit{ s.t. } \pi(i) = j\}$.

Fact 4.3. $\chi_S(x \circ \pi) = \chi_{\pi_2(S)}(x)$

Proof. First, break up the terms in to a convenient form.

$$\chi_S(x \circ \pi) = \prod_{i \in S} (x \circ \pi)_i = \prod_{i \in S} x_{\pi(i)} = \prod_{j \in \pi(S)} \Big(\prod_{i \in S : \pi(j) = i} x_j \Big) .$$

Notice that x_j contributes $x_j^{|\{i | \pi(i) = j\}|}$. So if x_j has an even number of i's for which $\pi(i) = j$, it will contribute 1 otherwise it contributes x_j. This implies

$$\prod_{j \in \pi(S)} \left(\prod_{i \in S : \pi(j) = i} x_j \right) = \prod_{j \in \pi_2(S)} x_j = \chi_{\pi_2(S)}(x) .$$

□

Now we return to the soundness portion in earnest. Below, the expectation is also taken over the choice of the edge (u, v), but we omit mentioning this after the first line for notational convenience.

$$\begin{aligned}
\mathbf{Pr}[\text{accepts}] &= \frac{1}{2} + \frac{1}{2} \mathop{\mathbf{E}}_{u,v,x,y,\mu}[A(x)B(y)B((x \circ \pi)y\mu)] \quad \text{(Indicator and Linearity)} \\
&= \frac{1}{2} + \frac{1}{2} \sum_{S,T,U} \hat{A}(S)\hat{B}(T)\hat{B}(U). \mathop{\mathbf{E}}_{x,y,\mu}[\chi_U(x)\chi_T(y)\chi_S((x \circ \pi)y\mu)] \\
&= \frac{1}{2} + \frac{1}{2} \sum_{S,U} \hat{A}(U)\hat{B}(S)^2 \mathop{\mathbf{E}}_{x}[\chi_U(x)\chi_S(x \circ \pi)] \mathop{\mathbf{E}}_{\mu}[\chi_S(\mu)] \quad (S = T) \\
&= \frac{1}{2} + \frac{1}{2} \sum_{S,U} \hat{A}(U)\hat{B}(S)^2 (1-\epsilon)^{|S|} \mathop{\mathbf{E}}_{x}[\chi_U(x)\chi_S(x \circ \pi)] \\
&= \frac{1}{2} + \frac{1}{2} \sum_{S,U} \hat{A}(U)\hat{B}(S)^2 (1-\epsilon)^{|S|} \mathop{\mathbf{E}}_{x}[\chi_U(x)\chi_{\pi_2(S)}(x)] \quad \text{(Using fact 4.3)} \\
&= \frac{1}{2} + \frac{1}{2} \sum_{S} \hat{A}(\pi_2(S))\hat{B}(S)^2 (1-\epsilon)^{|S|} \quad \text{(Conclude } \pi_2(S) = U\text{)}
\end{aligned}$$

Note that if the tables were not folded, for the all 1's tables, we would have $\hat{A}(\emptyset) = \hat{B}(\emptyset) = 1$ and the above acceptance probability equals 1.

Now suppose that $\mathbf{Pr}[\text{Test accepts}] \geq \frac{1}{2} + \delta$, this implies

$$\mathop{\mathbf{E}}_{u,v}\left[\sum_{S} \hat{A}(\pi_2(S))\hat{B}(S)^2 (1-\epsilon)^{|S|}\right] \geq 2\delta \, .$$

By an averaging argument, for at least a fraction δ of edges (u, v) we have

$$\sum_{S} \hat{A}(\pi_2(S))\hat{B}(S)^2 (1-\epsilon)^{|S|} \geq \delta \, . \tag{1}$$

Let us call such edges for which the above holds to be *good*. Our goal is to extract a labeling that satisfies a large fraction of the good edges.

Decoding an Assignment. For each $u \in V_1$ pick some set S with probability $\hat{B}(S)^2$ then choose $\sigma(u)$ be a random element of S. For each $v \in V_2$, pick some set T with prob $\hat{A}(T)^2$ and set $\sigma(v)$ to be a random element of T. First, notice that $\sum_S \hat{B}(S)^2 = 1$ and $\sum_T \hat{A}(T)^2 = 1$ by Parseval's identity, so these are indeed valid probability distributions. Also, by the folding property, we have $\hat{B}(\emptyset) = 0$, which implies that we never pick $S = \emptyset$ above, and likewise for T, and so the above procedure is well defined.

Fix and edge (u, v). The following lower bounds the probability that (u, v) is satisfied by the above randomized assignment.

$$\mathop{\mathbf{Pr}}_{\sigma}[(u,v) \text{ is satisfied}] = \mathop{\mathbf{Pr}}_{\sigma}[\pi(\sigma(u)) = \sigma(v)] \geq \sum_{S \neq \emptyset} \sum_{T \subseteq \pi(S)} \hat{B}(S)^2 \hat{A}(T)^2 \frac{1}{|S|}$$

To see this, whenever S, T are picked so that $T \subseteq \pi(S)$, irrespective of the label $\sigma(v)$ for v that we pick from $T \subseteq \pi(S)$, it is guaranteed to have a preimage under π in S, and therefore we have at least a $\frac{1}{|S|}$ chance of picking a label for u from S such that $\pi(\sigma(u)) = \sigma(v)$.

$$\begin{aligned}
\mathop{\mathbf{Pr}}_{\sigma}[(u,v) \text{ is satisfied}] &\geq \sum_{S} \hat{A}(\pi_2(S))^2 \hat{B}(S)^2 \frac{1}{|S|} \quad \text{(Drop the (positive) other terms)} \\
&\geq \sum_{S} \left(\hat{A}(\pi_2(S)) \hat{B}(S) \frac{1}{\sqrt{|S|}} \right)^2 \cdot \sum_{S} \hat{B}(S)^2 \quad \text{(by Parseval's)} \\
&\geq \left(\sum_{S} \hat{A}(\pi_2(S)) \hat{B}(S)^2 \frac{1}{\sqrt{|S|}} \right)^2 \quad \text{(Cauchy-Schwartz).}
\end{aligned}$$

First, a moment to show a useful bound for our situation

Proposition 4.4. $|S|^{-\frac{1}{2}} \geq (2\epsilon)^{1/2}(1-\epsilon)^{|S|}$

Proof.

$$\begin{aligned}
(2\epsilon|S|)^{-\frac{1}{2}} &\geq e^{-\epsilon|S|} \quad \text{(because } x^{-1} \geq e^{-x}\text{, and hence } x^{-1/2} \geq e^{-x/2} \text{ for } x > 0\text{)} \\
&\geq (1-\epsilon)^{|S|} \quad \text{(because } e^{-y} \geq (1-y) \text{ for } y > 0\text{)}
\end{aligned}$$

□

Now, combining Proposition 4.4 with the bound (1) for good edges, we conclude that whenever edge (u,v) is good, we have

$$\mathop{\mathbf{Pr}}_{\sigma}[(u,v) \text{ is satisfied}] \geq 2\epsilon \left(\sum_{S} \hat{A}(\pi_2(S)) \hat{B}(S)^2 (1-\epsilon)^{|S|} \right)^2 \geq 2\epsilon\delta^2 \, .$$

Therefore the randomized assignment σ satisfies at least a fraction $\delta \cdot (2\epsilon\delta^2) = 2\epsilon\delta^3$ of edges in expectation. In particular, this implies there exists an assignment sastifying at least a fraction $2\epsilon\delta^3$ of the edges. By starting with a $\mathsf{GAP\text{-}LC}_{1,\gamma}(\Sigma)$ instance where $\gamma < 2\epsilon\delta^3$, we conclude that the test must accept with probability less than $1/2 + \delta$. This establishes the soundness and concludes the construction of the 3-query PCP.