U.C. Berkeley — CS294: PCP and Hardness of Approximation — Professor Luca Trevisan
Handout N20 — 10 April 2006 — Scribe: Cameron Hill

Notes for Lecture 20

In this lecture, we are still caught up in the "alphabet-reduction phase" of Dinur's proof of the PCP theorem. In all of the following, assume that $\Sigma = \{1, ..., k\}$ and $\Gamma = \{1, ..., h\}$ with $h \geq k$.

We need an encoding procedure E, a decoding procedure D, and a testing procedure T as follows:

- $E : \Sigma \to \big\{\{0,1\}^\Sigma \to \{0,1\}\big\}$, and $E : \Sigma \to \big\{\{0,1\}^\Gamma \to \{0,1\}\big\}$ defined by
$$E(a)(x_1, ..., x_k) = x_a$$
- $D(A) = a$, where a minimizes the Hamming distance between $E(a)$ and A or $-A$ (whichever is closer);
- And T is has the following properties:

  1. If $A = E(a)$, $B = E(b)$ and $\pi(b) = a$, then $T(A, B, \pi)$ accepts w.p. 1;
  2. If $T(A, B, \pi)$ accepts w.p. $\geq .99$, then $\pi(D(B)) = D(A)$.
  3. T uses only $O(1)$ queries into A, B

Initially, we think of T implemented as follows:

> Testing procedure, T: ($\{0,1\}$–version)
> INPUT: $A : \{0,1\}^\Sigma \to \{0,1\}, B : \{0,1\}^\Gamma \to \{0,1\}$, and $\pi : \Gamma \to \Sigma$
>
> Choose $x \in \{0,1\}^\Sigma$ and $w, y, z \in \{0,1\}^\Gamma$ u.a.r.
> **accept** iff
> (1) $A(x) + B(y) + B(z) = B(x \circ \pi + y + z)$, and
> (2) $B(z) = 0 \Rightarrow B(y) = B(z \wedge w + y)$

In the algorithm, $+$ denotes bitwise *xor* (addition in $\mathbb{Z}/2\mathbb{Z}$), and $\wedge$ denotes bitwise conjunction. To interpret the first condition, suppose $A(x) = x_a$, $B(y) = y_b$ (i.e. A, B are encodings of a, b, respectively) and $\pi(b) = a$. The latter implies that $x_{\pi(b)} = x_a$ for each $x \in \{0,1\}^\Sigma$. The former implies that $B(x \circ \pi) = x_{\pi(b)}$. Then $A(x) = B(x \circ \pi)$ implies $x_a = x_{\pi(b)}$, so T should accept. As we are only checking a small number of bits of A, B, the following criterion will be more useful: $A(x) = B(x \circ \pi + y) - B(y)$ (equivalently, $x_a = x_{\pi(b)} + y_b - y_b$). For technical reasons (that will become clear once we get to the analysis) it will be even more convenient to check $A(x) = B(x \circ \pi + y + z) - B(y) - B(z)$, which is true for the same reason if A and B are consistent long codes.

1 Harmonic analysis

For the sake of the argument, we will exploit an alternative notation for the objects we've already been talking about. First of all, note that $0 \mapsto 1$ and $1 \mapsto -1$ defines a group isomorphism $\mathbb{Z}/2\mathbb{Z} \to (\{-1,1\}, \cdot)$, where $\cdot$ is the multiplication inherited from $\mathbb{R}$. It's convenient, then, to think of encodings $A : \{0,1\}^\Sigma \to \{0,1\}$ as real-valued functions $A : \{0,1\}^\Sigma \to \mathbb{R}$. Moreover, $V = \left\{\{0,1\}^\Sigma \to \mathbb{R}\right\}$ is an $\mathbb{R}$-vector space and admits the following inner product:

$$\langle F, G\rangle = \frac{1}{2^{|\Sigma|}} \sum_{x\in\{0,1\}^\Sigma} F(x)G(x) = \mathop{\mathbb{E}}_{x\in\{0,1\}^\Sigma} [F(x)G(x)]$$

Definition 1 *For each $\alpha \subseteq \Sigma$, define $u_\alpha : \{0,1\}^\Sigma \to \mathbb{R}$ by $u_\alpha(x_1, ..., x_k) = (-1)^{\sum_{i\in\alpha} x_i}$.*

For example, $u_\emptyset$ is identically 1, and $u_{\{2,4\}}(x_1, ..., x_k) = (-1)^{x_2+x_4}$

Claim 2 *$\{u_\alpha : \alpha \subseteq \Sigma\}$ is an orthonormal set of vectors in V.*

Proof: (1) $\langle u_\alpha, u_\alpha\rangle = \mathbb{E}_x[u_\alpha^2(x)] = 1$ for any $\alpha \subseteq \Sigma$.
(2) If $\alpha, \beta \subseteq \Sigma$ and $\alpha \neq \beta$, then

$$\begin{aligned}\langle u_\alpha, u_\beta\rangle &= \mathop{\mathbb{E}}_x\left[(-1)^{\sum_{i\in\alpha} x_i + \sum_{j\in\beta} x_j}\right]\\ &= \mathop{\mathbb{E}}_x\left[(-1)^{\sum_{i\in\alpha\Delta\beta} x_i}\right]\\ &= \prod_{i\in\alpha\Delta\beta} \mathop{\mathbb{E}}_x\left[(-1)^{x_i}\right] = 0\end{aligned}$$

□

It's not hard to see, then, that $\{u_\alpha\}_{\alpha\subseteq\Sigma}$ forms an orthonormal basis for V. Hence, if $F \in V$, then

$$F = \sum_{\alpha\subseteq\Sigma} \langle F, u_\alpha\rangle\, u_\alpha$$

For $\alpha \subseteq \Sigma$, let $\hat{F}(\alpha) = \langle F, \alpha\rangle$, so that $F = \sum_{\alpha\subseteq\Sigma} \hat{F}(\alpha)u_\alpha$. Let $\{u'_\beta\}_{\alpha\subseteq\Gamma}$ be the orthonormal basis of $\left\{\{0,1\}^\Gamma \to \mathbb{R}\right\}$ defined in the same fashion.
Obviously, $\{-1,1\} \subseteq \mathbb{R}$, and if $F : \{0,1\}^\Sigma \to \{-1,1\}$, then

$$\begin{aligned}\hat{F}(\alpha) &= \mathop{\mathbb{E}}_x\left[F(x)u_\alpha(x)\right]\\ &= \mathop{\mathbf{Pr}}_x\left[F(x) = u_\alpha(x)\right] - \mathop{\mathbf{Pr}}_x\left[F(x) \neq u_\alpha(x)\right]\\ &= 2\mathop{\mathbf{Pr}}_x\left[F(x) = u_\alpha(x)\right] - 1\end{aligned}$$

With these insights, we modify the encoding and testing procedures as follows. The encoding, $E : \Sigma \to \{\{0,1\}^\Sigma \to \{-1,1\}\}$, is defined by $E(a)(x_1, ..., x_k) = (-1)^{x_a}$. And

Testing procedure, T: ($\{-1,1\}$–version) INPUT: $A : \{0,1\}^\Sigma \to \{-1,1\}, B : \{0,1\}^\Gamma \to \{-1,1\}$, and $\pi : \Gamma \to \Sigma$ Assume $\mathbb{E}_x[A(x)] = 0$ and $\mathbb{E}_x[B(x)] = 0$. Choose $x \in \{0,1\}^\Sigma$ and $w, y, z \in \{0,1\}^\Gamma$ u.a.r. **accept** iff (1) $A(x) \cdot B(y) \cdot B(z) = B(x \circ \pi + y + z)$, and (2) $B(z) = 1 \Rightarrow B(y) = B(z \wedge w + y)$

We can ensure $\mathbb{E}_x[A(x)] = 0$ by, if necessary, replacing A with A', where $A'(0, x_2, ..., x_k) = A(0, x_2, ..., x_k)$ and $A'(1, x_2, ..., x_k) = -A(0, \overline{x_2}, ..., \overline{x_k})$, and adjusting T's queries as necessary. This procedure is called *folding*. As A and B are long codes, this does not degrade performance.

We will also make use of the following

Claim 3 *If* $F : \{0,1\}^\Delta \to \mathbb{R}$, *then* $\sum_{\alpha \subseteq \Delta} \hat{F}(\alpha) = \mathbb{E}_x \left[F(x)^2\right]$.

PROOF:

$$\begin{aligned}
\mathop{\mathbb{E}}_x \left[F(x)^2\right] &= \mathop{\mathbb{E}}_x \left[\left(\sum_{\alpha \subseteq \Delta} \hat{F}(\alpha) u_\alpha(x)\right)^2\right] \\
&= \sum_{\alpha \subseteq \Delta} \sum_{\beta \subseteq \Delta} \mathop{\mathbb{E}}_x \left[u_\alpha(x) u_\beta(x)\right] \hat{F}(\alpha)\hat{F}(\beta) \\
&= \sum_{\alpha \subseteq \Delta} \sum_{\beta \subseteq \Delta} \langle u_\alpha, u_\beta \rangle \hat{F}(\alpha)\hat{F}(\beta) \\
&= 2 \sum_{\alpha,\beta \subseteq \Delta : \alpha \neq \beta} \langle u_\alpha, u_\beta \rangle \hat{F}(\alpha)\hat{F}(\beta) + \sum_{\alpha \subseteq \Delta} \hat{F} \langle u_\alpha, u_\alpha \rangle (\alpha)^2 \\
&= \sum_{\alpha \subseteq \Delta} \langle u_\alpha, u_\alpha \rangle \hat{F}(\alpha)^2
\end{aligned}$$

□

And the next, which is trivial.

Claim 4 *If* $F : \{0,1\}^\Delta \to \{-1,1\}$, *then* $\mathbb{E}_x \left[F(x)^2\right] = 1$

2 Behavior of T

Suppose $T(A,B,\pi)$ accepts with probability $\geq .99$, and suppose $B(z)=1$. Observe that

$$\begin{aligned}\mathop{\mathbf{Pr}}_{x\in\{0,1\}^\Sigma, y\in\{0,1\}^\Gamma, z\in\{0,1\}^Gamma}[A(x)B(y)B(z)=B(x\circ\pi+y+z)] &= \mathop{\mathbf{Pr}}_{x,y,z}[A(x)B(y)B(z)B(x\circ\pi+y+z)=1] \\ &= \mathop{\mathbb{E}}_{x,y}\left[\frac{1}{2}+\frac{1}{2}A(x)B(y)B(x\circ\pi+y+z)\right] \\ &= \frac{1}{2}+\frac{1}{2}\mathop{\mathbb{E}}_{x,y}[A(x)B(y)B(x\circ\pi+y+z)]\end{aligned}$$

Thus, $\mathbf{Pr}_{x,y,z}[A(x)B(y)B(z)=B(x\circ\pi+y+z)]\geq .99$ if and only if $\mathbb{E}_{x,y}[A(x)B(y)B(z)B(x\circ\pi+y+z)]\geq .98$, and it suffices to compute the expectation.

First,

$$\mathop{\mathbb{E}}_{x,y,z}[A(x)B(y)B(z)B(x\circ\pi+y+z)]=$$

$$\begin{aligned}&= \mathop{\mathbb{E}}_{x,y,z}\left[\left(\sum_{\alpha\subseteq\Sigma}\hat{A}(\alpha)u_\alpha(x)\right)\left(\sum_{\beta\subseteq\Gamma}\hat{B}(\beta)u_\beta(y)\right)\left(\sum_{\gamma\subseteq\Gamma}\hat{B}(\gamma)u_\gamma(z)\right)\left(\sum_{\delta\subseteq\Gamma}\hat{B}(\delta)u_\delta(x\circ\pi+y+z)\right)\right] \\ &= \sum_{\alpha\subseteq\Sigma}\sum_{\beta,\gamma,\delta\subseteq\Gamma}\hat{A}(\alpha)\hat{B}(\beta)\hat{B}(\gamma)\hat{B}(\delta)\mathop{\mathbb{E}}_{x}[u_\alpha(x)u_\delta(x\circ\pi)]\mathop{\mathbb{E}}_{y}[u_\beta(y)u_\delta(y)]\mathop{\mathbb{E}}_{z}[u_\gamma(z)u_\delta(z)] \\ &= \sum_{\alpha\subseteq,\Sigma\beta\subseteq\Gamma}\hat{A}(\alpha)\hat{B}(\beta)^3\mathop{\mathbb{E}}_{x}[u_\alpha(x)u_\beta(x\circ\pi)] \\ &= \sum_{\alpha\subseteq,\Sigma\beta\subseteq\Gamma}\hat{A}(\alpha)\hat{B}(\beta)^3\mathop{\mathbb{E}}_{x}\left[(-1)^{\sum_{a\in\alpha}x_a}(-1)^{\sum_{b\in\beta}x_{\pi(b)}}\right]\end{aligned}$$

Now, if we define $\pi_2: 2^\Gamma\to 2^\Sigma$ by $\pi_2(\beta)=\{a\in\Sigma : |\pi^{-1}[a]\cap\beta| \text{ is odd}\}$, then

$$\sum_{b\in\beta}x_{\pi(b)}=\sum_{a\in\pi_2(\beta)}x_a$$

(using addition modulo 2). It follows that

$$\mathop{\mathbb{E}}_{x,y,z}[A(x)B(y)B(z)B(x\circ\pi+y+z)]=\sum_{\alpha\subseteq\Sigma,\beta\subseteq\Gamma}\hat{A}(\alpha)\hat{B}(\beta)^3\mathop{\mathbb{E}}_{x}\left[(-1)^{\sum_{a\in\alpha}x_a}(-1)^{\sum_{a\in\pi_2(\beta)}x_a}\right]=$$

$$=\sum_{\beta\subseteq\Gamma}\hat{A}(\pi_2(\beta))\hat{B}(\beta)^3$$

because

$$\mathop{\mathbb{E}}_{x}\left[(-1)^{\sum_{a\in\alpha}x_a}\cdot(-1)^{\sum_{a\in\pi_2(\beta)}x_a}\right]=\begin{cases}1 & \text{if } \alpha=\pi_2(\beta)\\ 0 & \text{otherwise}\end{cases}$$

Thus, if $T(A, B, \pi)$ accepts with probability $\geq .99$, then

$$\begin{aligned} .98 &\leq \sum_{\beta \subseteq \Gamma} \hat{A}(\pi_2(\beta))\hat{B}(\beta)^3 \\ &\leq \max_{\beta} \hat{A}(\pi_2(\beta))\hat{B}(\beta) \sum_{\beta \subseteq \Gamma} \hat{B}(\beta)^2 \\ &= \max_{\beta} \hat{A}(\pi_2(\beta))\hat{B}(\beta) \end{aligned}$$

And so there exists a β_0 such that $|\hat{A}(\pi_2(\beta_0))| \geq .98$ and $|\hat{B}(\beta_0)| \geq .98$.

In the coming lecture, we will show that β_0 is a set of cardinality 1, that is, there exists a $b \in \Gamma$ such that $|\hat{A}(\{\pi(b)\})| \geq .98$ and $|\hat{B}(\{b\})| \geq .98$. From this, it will follow that $D(A) = \pi(b)$ and $D(B) = b$, so that $D(A) = \pi(D(B))$ as desired.