COM S 6810 Theory of Computing April 21, 2009

Lecture 24: Proving PCP Lite: Linearity Testing

Instructor: Rafael Pass *Scribe: Igor Gorodezky*

1 Linearity Testing

In this lecture we work towards proving the following lite version of the PCP theorem, due to Arora, Lund, Motwani, Sudan, and Szegedy (1992).

Theorem 1 (PCP Lite) $NP \subseteq PCP(\text{poly}(n), 1)$.

Recall that we aim to prove PCP Lite by showing that the (NP-complete) language of satisfiable systems of quadratic equations modulo 2 is in $\mathsf{PCP}(\text{poly}(n), 1)$. The proof strategy was described in the previous lecture. In this lecture we focus on the first step, linearity testing.

Recall that in our PCP protocol, the prover has a satisfying assignment $a = (a_1, \ldots, a_n)$ and sends to the verifier a proof π consisting of 2^{n^2} entries of the form

$$\pi(v) = \langle a^{(2)}, v \rangle$$

for $v \in \{0,1\}^{n^2}$ (see the previous lecture for a definition of $a^{(2)}$). In other symbols, π is the truth table for the linear function $v \mapsto \langle a^{(2)}, v \rangle$ on $\{0,1\}^{n^2}$.

In order to verify the proof, the verifier must first make sure that π describes a function that is indeed linear; this is the linearity testing step. The *naive linearity test* is as follows:

1. Choose $x, y \in \{0,1\}^{n^2}$ randomly and independently, and
2. accept π if $\pi(x) \oplus \pi(y) = \pi(x \oplus y)$.

This test clearly has perfect completeness; if π is linear then it will always be accepted. In general, however, it does not have non-negligible soundness.

Fortunately, it is the case that the naive linearity test has a constant probability of rejecting a function if it is far from linear, and as it turns out this will suffice for our purposes. Let's begin by formalizing this new notion.

Definition. Given f, g boolean functions on n bits, define their *distance* to be

$$d(f, g) = \text{Prob}_x[f(x) \neq g(x)].$$

We say that a boolean function f is *δ-far from linear* if $d(f,g) \geq \delta$ for every linear boolean function g.

The following theorem is due to Blum, Luby, and Rubinfeld (1990).

Theorem 2 *If f is δ-far from linear then the naive linearity test rejects f with probability at least δ.*

In order to prove this theorem we introduce discrete Fourier analysis.

2 Intermezzo: Discrete Fourier Analysis

Let us assume that bits take value in $\{-1,1\}$ where -1 is "true" and 1 is "false". Then boolean functions are maps from $\{-1,1\}^n$ to $\{-1,1\}$.

Consider the vector space of functions $f : \{-1,1\}^n \to \mathbb{R}$. Recall from linear algebra that the standard basis for this vector space is the set of functions

$$I_v(x) = \begin{cases} 1 & \text{if } x = v \\ 0 & \text{otherwise} \end{cases}$$

where $v \in \{-1,1\}^n$. There is a more useful basis for this vector space, however, which we will call the *Fourier basis*. Before defining it, let us redefine the scalar product by normalizing it; this is purely for convenience.

Definition. Given $f, g : \{-1,1\}^n \to \mathbb{R}$ define

$$\langle f, g \rangle = \frac{1}{2^n} \sum_x f(x)g(x).$$

Probabilistically speaking, we have defined

$$\langle f, g \rangle = \mathbb{E}_x[f(x)g(x)].$$

Now we can define the Fourier basis: it consists of the functions

$$\chi_v(x) = \prod_{i \,:\, v_i = 1} x_i$$

where $v \in \{-1,1\}^n$ and v_i is the ith bit of v. The empty product equals 1 by convention.

Proposition 3 *The collection $\{\chi_v\}_v$ is an orthonormal basis.*

Proof. Let us verify that the χ_v are orthonormal. Given vectors u, v we have

$$\begin{aligned}\langle \chi_u, \chi_v \rangle &= \mathbb{E}_x[\chi_u(x)\chi_v(x)] \\ &= \mathbb{E}_x\left[\prod_{i\,:\,u_i=1} x_i \prod_{i\,:\,v_i=1} x_i\right] \\ &= \mathbb{E}_x\left[\prod_{i\,:\,u_i\neq v_i} x_i\right]\end{aligned}$$

since $x_i^2 = 1$. If $u = v$ then the product is empty hence equals 1, and we have $\langle \chi_u, \chi_v \rangle = 1$. If $u \neq v$ then the product is not empty, and by the independence of bits we have

$$\langle \chi_u, \chi_v \rangle = \mathbb{E}_x\left[\prod_{i\,:\,u_i\neq v_i} x_i\right] = \prod_{i\,:\,u_i\neq v_i} \mathbb{E}_x[x_i] = 0$$

since $\mathbb{E}_x[x_i] = 0$. We conclude that the χ_v are orthonormal, hence linearly independent. Since there are 2^n of them, they form a basis. ■

Since the Fourier basis is indeed a basis, we can write any $f : \{-1,1\}^n \to \mathbb{R}$ as

$$f = \sum_v \langle f, \chi_v \rangle \chi_v.$$

It is traditional to write $\langle f, \chi_v \rangle$ as $\hat{f}(v)$, so that $f = \sum_v \hat{f}(v)\chi_v$. The $\hat{f}(v)$ are the *Fourier coefficients* of f. Interpreting the scalar product probabilistically, we note that

$$\hat{f}(v) = \mathbb{E}_x[f(x)\chi_v(x)].$$

The most basic result in Fourier analysis is Parseval's identity.

Theorem 4 (Parseval's Identity) *For any $f : \{-1,1\}^n \to \mathbb{R}$,*

$$\langle f, f \rangle = \sum_v \hat{f}(v)^2.$$

Proof. Using the Fourier expansion of f,

$$\begin{aligned}
\langle f, f\rangle &= \mathbb{E}_x[f(x)^2] \\
&= \mathbb{E}_x\left[\left(\sum_v \hat{f}(v)\chi_v(x)\right)^2\right] \\
&= \mathbb{E}_x\left[\sum_{u,v} \hat{f}(u)\hat{f}(v)\chi_u(x)\chi_v(x)\right] \\
&= \sum_{u,v} \hat{f}(u)\hat{f}(v)\,\mathbb{E}_x\left[\chi_u(x)\chi_v(x)\right] \\
&= \sum_{u,v} \hat{f}(u)\hat{f}(v)\langle\chi_u, \chi_v\rangle
\end{aligned}$$

and now the fact that the χ_v are orthonormal completes the proof. ■

Recalling that $\langle f, f\rangle = \mathbb{E}_x[f(x)^2]$ gives the following corollary.

Corollary 5 *If f is a boolean function then*

$$\sum_v \hat{f}(v)^2 = 1.$$

3 Proving the Theorem

We are now equipped to prove Theorem **??**. In order to use the results of the previous section, we will continue to assume that bits take value in $\{-1, 1\}$.

First, we observe that if $f : \{-1,1\}^n \to \{-1,1\}$ is δ-far from linear then $\hat{f}(v) \leq 1 - 2\delta$ for every $v \in \{-1,1\}^n$. This follows immediately from the following lemma.

Lemma 6 *If $d(f, g) = \delta$ then $\langle f, g\rangle = 1 - 2\delta$.*

Proof. We have

$$\begin{aligned}
\langle f, g\rangle &= \frac{1}{2^n}\sum_x f(x)g(x) \\
&= \frac{1}{2^n}\Big(|\{x \mid f(x) = g(x)\}| - |\{x \mid f(x) \neq g(x)\}|\Big) \\
&= 1 - d(f, g) - d(f, g) \\
&= 1 - 2\delta
\end{aligned}$$

as desired. ■

Proof of Theorem ??. We will show that if f is δ-far from linear then the naive linearity test accepts it with probability at most $1-\delta$. Observe that since our bits take value in $\{-1,1\}$, the naive linearity tests consists of sampling $x, y \in \{-1,1\}^n$ and accepting if $f(x)f(y) = f(x \times y)$, where $\times$ is the bit-wise product. Thus we have

$$\begin{aligned}
\text{Prob}[f \text{ accepted}] &= \text{Prob}_{x,y}[f(x)f(y) = f(x \times y)] \\
&= \mathbb{E}_{x,y}\left[\frac{1}{2} + \frac{1}{2}f(x)f(y)f(x \times y)\right] \\
&= \frac{1}{2} + \frac{1}{2}\,\mathbb{E}_{x,y}[f(x)f(y)f(x \times y)].
\end{aligned}$$

It suffices to prove that the expected value above, call it E, is at most $1-2\delta$. Using the Fourier expansion of f, we have

$$\begin{aligned}
E &= \mathbb{E}_{x,y}\left[\left(\sum_u \hat{f}(u)\chi_u(x)\right)\left(\sum_v \hat{f}(v)\chi_v(y)\right)\left(\sum_w \hat{f}(w)\chi_w(x \times y)\right)\right] \\
&= \mathbb{E}_{x,y}\left[\sum_{u,v,w} \hat{f}(u)\hat{f}(v)\hat{f}(w)\chi_u(x)\chi_v(y)\chi_w(x \times y)\right] \\
&= \sum_{u,v,w} \hat{f}(u)\hat{f}(v)\hat{f}(w)\,\mathbb{E}_{x,y}\left[\chi_u(x)\chi_v(y)\chi_w(x \times y)\right].
\end{aligned}$$

Now,

$$\begin{aligned}
\mathbb{E}_{x,y}\left[\chi_u(x)\chi_v(y)\chi_w(x \times y)\right] &= \mathbb{E}_{x,y}\left[\prod_{i\,:\,u_i=1} x_i \prod_{i\,:\,v_i=1} y_i \prod_{i\,:\,w_i=1} x_i y_i\right] \\
&= \mathbb{E}_{x,y}\left[\prod_{i\,:\,u_i \neq w_i} x_i \prod_{i\,:\,v_i \neq w_i} y_i\right] \\
&= \mathbb{E}_x\left[\prod_{i\,:\,u_i \neq w_i} x_i\right]\mathbb{E}_y\left[\prod_{i\,:\,v_i \neq w_i} y_i\right] \\
&= \langle \chi_u, \chi_w \rangle \langle \chi_v, \chi_w \rangle
\end{aligned}$$

and by the orthogonality of the χ_v this quantity is 1 when $u = v = w$ and 0 otherwise.

Therefore,

$$\begin{aligned} E &= \sum_{u,v,w} \hat{f}(u)\hat{f}(v)\hat{f}(w)\,\mathbb{E}_{x,y}\left[\chi_u(x)\chi_v(y)\chi_w(x\times y)\right] \\ &= \sum_{v} \hat{f}(v)^3 \\ &\leq (1-2\delta)\sum_{v} \hat{f}(v)^2 \\ &= 1-2\delta \end{aligned}$$

where the inequality follows from the observation with which we began this section, and the final equality follows from Corollary **??**. This completes the proof. ■

We conclude that in our PCP protocol, if the proof π is δ-far from linear, then the naive linearity test will reject it with probability at least δ.