6.895 PCP and Hardness of Approximation MIT, Fall 2010

Lecture 5: Linearity Testing

Lecturer: *Dana Moshkovitz* Scribe: *Gregory Minton and Dana Moshkovitz*

In the last lecture, we proved a weak PCP Theorem, namely,

$$NP \subseteq PCP_{1,0.99}[O(\log n), \text{polylog}(n)],$$

under the low degree testing assumption, i.e., the assumption that given the table of a function $f : \mathbb{F}^m \to \mathbb{F}$, a degree d and access to an auxiliary proof, a probabilistic verifier can test whether f corresponds to a polynomial of degree at most d using a poly-logarithmic number of queries.

In this lecture we discuss what should be the exact formulation of the low degree testing assumption. Then we prove this formulation for the special case of $d = 1$ and $\mathbb{F} = GF(2)$. This is the Blum-Luby-Rubinfeld linearity test.

In the next lecture we deal with the harder case of general degree d and field $\mathbb{F}$.

1 Formulating The Low Degree Testing Problem

Can a local algorithm distinguish polynomials of degree at most d from functions that are not polynomials of degree at most d? The answer to this question is *no*. For let p be a polynomial of degree at most d, and let p' be the function

$$p'(x) = \begin{cases} p(x) & \text{if } x \neq x_0 \\ p(x_0) + 1 & \text{if } x = x_0. \end{cases}$$

The function $p - p'$ is not the zero function, but it has a root everywhere except one point. Because a nonzero polynomials of degree at most d have roots in at most $d/|\mathbb{F}|$ fraction of the space, we deduce that $p - p'$, and so p', cannot be polynomials of degree at most d. Thus a verifier which purports to test for low degree polynomials should pass p but fail p'. But of course the probability that any verifier can distinguish p and p' is bounded above by the probability that it queries location x_0! If the verifier makes few queries, and we choose x_0 appropriately, then this probability can be made quite small. (To be more precise, if the proof has length ℓ and the verifier makes q queries, then there exists some x_0 for which the probability that the verifier queries x_0 is $\leq q/\ell$. We are interested in verifiers in which the proof length grows faster than the number of queries, so this is $o(1)$.)

With this in mind, we cannot hope for an exact test of low degree polynomials. What we ask for instead is testing whether the function is *close* to a low degree polynomial in Hamming distance.

Definition 1 (Far/close to low degree). *We say that a function $f : \mathbb{F}^m \to \mathbb{F}$ is δ-*far *from degree d if for any m-variate polynomial p of degree at most d over $\mathbb{F}$, we have $\Delta(f, p) \geq \delta$ where we view f and p as vectors (so $\Delta(f, p) = \Pr_{x \in \mathbb{F}^m}[f(x) \neq p(x)]$).*

*We say that f is δ-*close *to degree d, if there exists an m-variate polynomial p of degree at most d over $\mathbb{F}$, such that $\Delta(f, p) \leq \delta$.*

Note that, by definition, if we query a function that is $(1-\gamma)$-close to a low degree polynomial on a uniformly random point, then with probability at least $(1-\gamma)$ we get the same answer as if we queried the low degree polynomial. We will use this observation (as well as additional ideas) to argue that the low degree testing assumption we formulate suffices to prove the weak PCP Theorem.

Assumption 1.1 (Low Degree Testing)**.** *There are constants $\delta, \gamma, \gamma' > 0$, such that given the table of a function $f : \mathbb{F}^m \to \mathbb{F}$ and a degree $d \leq \delta |\mathbb{F}|$, there is a probabilistic verifier for the statement "$\deg f \leq d$". The verifier is given access to f and to an auxiliary proof and satisfies the following:*

- Completeness: *If $\deg f \leq d$, then there is a proof that the verifier always accepts.*
- Soundness: *If f is γ-far from degree d, then for any proof, the verifier rejects with probability at least γ'.*

The verifier uses $O(\log(|\mathbb{F}|^m))$ random bits. It makes only $|\mathbb{F}|^{O(1)}$ queries to f.

Put differently, the soundness condition says that if, for some auxiliary proof, the verifier accepts with probability more than $1-\gamma'$, then there is a polynomial p of degree at most d that agrees with f on $1-\gamma$ fraction of the points in $\mathbb{F}^m$.

2 Linear Functions

In this lecture we consider the case of $d = 1$, $\mathbb{F} = GF(2)$. That is, given a function $f : \{0,1\}^n \to \{0,1\}$, we want to test (with high probability) whether f is (close to) a linear function. This special case is not sufficient for proving the PCP Theorem, but it is simpler and its ideas – important.

We first define what we mean by a linear function. Two different definitions come to mind, so let us show that they are equivalent.

Definition. *A function $f : \{0,1\}^n \to \{0,1\}$ is linear if for all $x, y \in \{0,1\}^n$, $f(x) + f(y) = f(x+y)$.*

Claim. *A function $f : \{0,1\}^n \to \{0,1\}$ is linear iff there exists a vector $a = (a_1, \ldots, a_n) \in \{0,1\}^n$ such that $f(x) = \sum_{i=1}^n a_i x_i = \langle a, x \rangle$. (As we are working in $GF(2)$, addition is of course taken modulo 2.)*

Proof. The $(\Leftarrow)$ direction is clear. For the $(\Rightarrow)$ direction, let $\{e_1, \ldots, e_n\}$ be the standard basis, i.e.

$$e_i = \underbrace{(0, \ldots, 0, 1, 0, \ldots, 0)}_{\text{1 in position } i},$$

and define $a_i = f(e_i)$. Then the stated formula follows by linearity. □

Note that the functions we are calling "linear" are not affine functions $y = ax + b$; we require that the "constant term" be zero.

3 The Blum-Luby-Rubinfeld Linearity Test

The test we show does not use an auxiliary proof; it only makes queries to f. In this case, the tester has to make at least three queries, as there exist functions f which are far from linear but which have the property that for any two points x, y, there exists a linear function g such that $f(x) = g(x)$ and $f(y) = g(y)$. We will show that three queries suffices, by studying the properties of the following simple test.

BLR Test (Blum, Luby, Rubinfeld)**.** *Choose uniformly random points* $x, y \in \{0,1\}^n$*. Test if* $f(x) + f(y) = f(x+y)$*.*

This algorithm uses $2n$ random bits. It makes $q = 3$ queries. The completeness of this test is 1, because obviously a linear function passes with probability 1. Analyzing the soundness is the interesting part; the answer is given by the following theorem.

Theorem (Soundness of BLR)**.** *If f is δ-far from linear, then*

$$\Pr[\textit{BLR test rejects } f] \geq \min\left(\frac{2}{9}, \frac{\delta}{2}\right) \geq \frac{2\delta}{9}.$$

Before proving this theorem, we make some remarks.

- This result is not tight, as we can prove via Fourier analysis that $\Pr[\text{rejection}] \geq \delta$. (We will return to this later in the course.) And even that result is not tight in the low order terms!

- The definition of linearity generalizes to any group G; in the setting of group theory such a map is known as a *homomorphism*. In fact, the BLR test generalizes to testing homomorphisms on groups. In this setting, the soundness theorem above is tight. For instance, define $f : (\mathbb{Z}/9)^n \to \mathbb{Z}/9$ by $f(u) = 3k$ if $u_1 \in \{3k-1, 3k, 3k+1\}$. (That is, f rounds the first coordinate to the nearest multiple of 3.) This is not linear; one can check that f is (2/3)-far from linear. Hence the soundness theorem tells us that BLR should reject f with probability at least 2/9. In fact, that is exactly the rejection probability, because

$$f(x) + f(y) \neq f(x+y) \iff x_1 \equiv y_1 \equiv 1 \bmod 3 \text{ or } x_1 \equiv y_1 \equiv -1 \bmod 3.$$

4 Analysis of the Linearity Test

The rest of this lecture is devoted to proving the soundness theorem.

4.1 Majority Correction

The proof uses the useful idea of majority correction. Fix a function $f : \{0,1\}^n \to \{0,1\}$ and a point $x \in \{0,1\}^n$. If f is linear, then for any $y \in \{0,1\}^n$ we have $f(x) = f(y) + f(x-y)$. Thus we may think of each of the 2^n values of y as offering the "vote" $f(y) + f(x-y)$ for $f(x)$. As there are only two possible values for $f(x)$, 0 and 1, one of them must get a majority of the votes. We define a function g by setting $g(x)$ to be the value that receives the most votes.

More formally, $g : \{0,1\}^n \to \{0,1\}$ is defined by

$$g(x) = \begin{cases} 1 & \text{if } \Pr_y[f(y) + f(x-y) = 1] \geq 1/2 \\ 0 & \text{otherwise.} \end{cases}$$

(We have chosen to always break a tie with the value 1; this was arbitrary, and it will turn out that the definition of $g(x)$ in the case $\Pr_y[f(y) + f(x-y) = 1] = 1/2$ is unimportant.) It will be useful later to define the probabilities

$$P_x = \Pr_y[g(x) = f(y) + f(x-y)].$$

By definition of $g(x)$, $P_x \geq 1/2$ for all x.

4.2 Majority Correction Works

We will obtain the soundness theorem by proving three claims relating properties of the BLR test to the function g.

Claim. $\Pr[\mathit{BLR\ rejects\ f}] \geq \frac{1}{2} \cdot \mathrm{dist}(g,f)$.

Proof. Conditioning on whether $g(x) = f(x)$ or not, we have

$$\begin{aligned} \Pr[\text{rejection}] &= \Pr[g(x) \neq f(x)] \cdot \Pr[\text{rejection} \mid g(x) \neq f(x)] \\ &\quad + \Pr[g(x) = f(x)] \cdot \Pr[\text{rejection} \mid g(x) = f(x)]. \end{aligned}$$

We get a lower bound on Pr[rejection] by ignoring the second term. In the first term, notice that $\Pr[g(x) \neq f(x)] = \mathrm{dist}(g,f)$ by definition of the distance. By definition of g, if $g(x) \neq f(x)$ then $f(x) = f(y) + f(x-y)$ for $1 - P_x \leq 1/2$ of the possible values of y. But because we are working over the binary field (so addition and subtraction are the same), the equation $f(x) = f(y)+f(x-y)$ is equivalent to the BLR test $f(x+y) = f(x)+f(y)$. Hence, given $g(x) \neq f(x)$, the BLR test fails with probability at least 1/2. Putting this together, $\Pr[\text{rejection}] \geq \frac{1}{2} \cdot \mathrm{dist}(g,f)$, as desired. □

Claim. *If* $\Pr[\mathit{BLR\ rejects\ f}] < \frac{2}{9}$, *then for all x we have* $P_x > \frac{2}{3}$.

Proof. Fix x. We compute

$$A = \Pr_{y,z}[f(y) + f(x+y) = f(z) + f(x+z)]$$

in two different ways. First, notice that $f(y) + f(x+y)$ equals $g(x)$ with probability P_x, and the same is true of $f(z) + f(x+z)$. Using independence of y and z, the probability that both expressions are equal to $g(x)$ is P_x^2, and the probability that they are both equal to $g(x)+1$ is $(1-P_x)^2$. Hence $A = P_x^2 + (1-P_x)^2$.

We can also bound A using the probability of BLR rejection. First, rewrite the condition $f(y) + f(x+y) = f(z) + f(x+z)$ as $f(y) + f(z) = f(x+y) + f(x+z)$. By definition of the BLR test, $f(y) + f(z)$ equals $f(y+z)$ with probability $1 - \Pr[\text{BLR rejects } f] > 7/9$. As y and z are independent and uniformly sampled, the same is true of $x+y$ and $x+z$, and so the same argument shows that $f(x+y) + f(x+z) = f((x+y) + (x+z)) = f(y+z)$ with probability $> 7/9$. Thus

$$f(y) + f(z) = f(y+z) = f(x+y) + f(x+z)$$

with probability $> 5/9$, so certainly $A = \Pr[f(y) + f(z) = f(x+y) + f(x+z)] > 5/9$.

Combining the results of the last two paragraphs, we deduce that

$$P_x^2 + (1 - P_x)^2 > 5/9.$$

This implies either $P_x < 1/3$ or $P_x > 2/3$. Of course the first case is impossible because $P_x \geq 1/2$, so we must have $P_x > 2/3$ as desired. □

Claim. *If* $\Pr[\mathit{BLR\ rejects}\ f] < \frac{2}{9}$*, then* g *is linear.*

Proof. By the previous claim, we must have $P_x > 2/3$ for all x. Now fix x, y and consider choosing z uniformly at random. Then $g(x)$ equals $f(z) + f(x+z)$ with probability larger than 2/3. Similarly, $g(y)$ equals $f(z)+f(y+z)$ with probability larger than 2/3. The same argument says that $g(x+y)$ equals $f(z) + f(z+x+y)$ with probability larger than 2/3; we can of course replace the uniformly sampled value z by $z+x$, finding that $g(x+y) = f(z+x)+f(z+y)$ with the same probability (more than 2/3). As each of these three conditions holds with probability larger than 2/3, they hold simultaneously with positive probability. That is, there exists at least one z_0 such that

$$\begin{aligned} g(x) &= f(z_0) + f(x + z_0), \\ g(y) &= f(z_0) + f(y + z_0), \text{ and} \\ g(x+y) &= f(z_0 + x) + f(z_0 + y) \end{aligned}$$

all hold. But this shows that

$$g(x) + g(y) = f(z_0) + f(x + z_0) + f(z_0) + f(y + z_0) = f(x + z_0) + f(y + z_0) = g(x + y).$$

This holds for all x, y, so g is linear, as desired. □

Putting the last few claims together, we immediately get the soundness theorem. Specifically, we find that either Pr[rejection] $\geq 2/9$, or else g is linear and so

$$\Pr[\text{rejection}] \geq \frac{1}{2} \cdot \operatorname{dist}(g, f) \geq \frac{1}{2} \operatorname{dist}(f, \text{linear}).$$

That is exactly what the soundness theorem asserts, so we are done.