**18.405J/6.841J: Advanced Complexity Theory** Spring 2016

# Lecture 16: Linearity Testing

*Prof. Dana Moshkovitz* *Scribes: Dana Moshkovitz & Anonymous Student*
*Scribe Date: Spring 2015*

## 1 Introduction

In the last lecture we saw that proving the NP-hardness of distinguishing the case where a MAX-3SAT instance can be fully satisfied from the case where only 0.99 fraction of its clauses can be satisfied is equivalent to proving that any NP language can be tested locally probabilistically (the PCP Theorem). This is because the former implies a reduction from any NP language $L$ to the distinguishing problem. Inputs $x \in L$ are mapped to fully satisfiable MAX-3SAT instances, whereas inputs $x \notin L$ are mapped to instances where only 0.99 fraction of their clauses can be satisfied. Given an assignment to the MAX-3SAT instance, a verifier can pick a random clause, query its 3 variables, and check whether the clause is satisfied. Even though the test of the verifier is *local* (depends on only 3 locations in the proof), we'll be able to infer (except for some constant error probability), a *global* property: the membership of $x$ in $L$.

In the next couple of lectures we'll prove a weak version of the PCP Theorem (the probabilistically checkable proof will be of exponential, rather than polynomial, length). We'll also hint at how the full PCP Theorem is proved. In this lecture we develop the mathematical machinery that will allow us to relate the results of local tests to global properties.

We'll prove an important theorem in *property testing*, showing that closeness to *linearity* can be tested locally. We say that functions $f, g : \{0,1\}^n \to \{0,1\}$ are $(1-\delta)$-close if $f(x) = g(x)$ for at least $1-\delta$ fraction of the $x \in \{0,1\}^n$. We say that $f$ is $(1-\delta)$-*close* to having a property $\mathcal{P} \subseteq \{\{0,1\}^n \to \{0,1\}\}$ if there is $\tilde{f} \in \mathcal{P}$ that is $(1-\delta)$-close to $f$. We say that $f, g$ are $\delta$-far if they are not $(1-\delta)$-close. Similarly we say that $f$ is $\delta$-far from $\mathcal{P}$ if $f$ is not $(1-\delta)$-close to $\mathcal{P}$. The properties $\mathcal{P}$ we'll consider are such that it is impossible to check locally whether $f \in \mathcal{P}$, and the relaxation to checking closeness to $\mathcal{P}$ is needed.

## 2 Linear Functions

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 $\{e_1, \dots, e_n\}$ be the standard basis, i.e.

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

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 only makes 3 queries to $f$. This is the minimal number of queries needed 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 $\delta$-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.** $\Pr[$*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* $\Pr[$*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* $\Pr[\textit{BLR rejects } f] < \frac{2}{9}$*, 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

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

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

## 5 Random Self Reducibility

We finish this lecture by pointing out a useful technique, called *random self-reducibility.* Suppose that $f : \{0,1\}^n \to \{0,1\}$ is $(1-\delta)$-close to a linear function $\tilde{f} : \{0,1\}^n \to \{0,1\}$. Note that for small $\delta$ there can be only one such $\tilde{f}$, since if there are two $\tilde{f}_1 \neq \tilde{f}_2$, then by the triangle inequality, $dist(\tilde{f}_1, \tilde{f}_2) \leq 2\delta$, i.e., $\Pr_x\left[(\tilde{f}_1 - \tilde{f}_2)(x) = 1\right] \leq 2\delta$, but $\tilde{f}_1 - \tilde{f}_2$ is a non-zero linear function, hence $\Pr_x\left[(\tilde{f}_1 - \tilde{f}_2)(x) = 1\right] = 1/2$.

This means that if we pick uniformly at random $x \in \{0,1\}^n$, then $f(x) = \tilde{f}(x)$ with probability at least $1 - \delta$. What if we'd like to evaluate $\tilde{f}(x_0)$ for an arbitrary $x_0 \in \{0,1\}^n$? We can write $x_0 = (x_0 + x) + x$ for a uniformly random $x \in \{0,1\}^n$, and evaluate $\tilde{f}(x_0)$ by computing $f(x_0+x)+f(x)$. Since both $x_0+x$ and $x$ are uniformly random, with probability at least $1-2\delta$ we have $f(x_0 + x) = \tilde{f}(x_0 + x)$ and $f(x) = \tilde{f}(x)$, and hence $f(x_0 + x) + f(x) = \tilde{f}(x)$.

MIT OpenCourseWare
https://ocw.mit.edu

18.405J / 6.841J Advanced Complexity Theory
Spring 2016

For information about citing these materials or our Terms of Use, visit: https://ocw.mit.edu/terms.