Lecture 19

Lecturer: Eli Ben-Sasson *Scribe: Ben Wagner*

1 Linearity Testing Over $\mathbb{Z}_2$

1.1 Introduction

Today we will provide a tighter analysis of the BLR linearity test (defined in previous lecture) for the special case where $G = \mathbb{Z}_2^n$ and $H = \mathbb{Z}_2$. This analysis was originally shown by Bellare et al. [3]. We use the definitions from last lecture (specifically, the set of homomorphisms $\mathrm{Hom}(G, H)$, the BLR tester T^f and the definition of distance $\delta(f, \mathrm{Hom})$). Recall the BLR test T^f picks x, y uniformly at random from G, and accepts iff $f(x) + f(y) = f(x + y)$ and let $\mathbb{Z}_2$ be the additive group with two elements $0, 1$ and addition modulo two.

Theorem 1 *[3] For every function $f' : \mathbb{Z}_2^n \to \mathbb{Z}_2$, the BLR test $T^{f'}$ gives*

$$\delta(f', \mathrm{Hom}(\mathbb{Z}_2^n, \mathbb{Z}_2)) \leq \Pr\left[T^{f'} \textit{rejects}\right]$$

where

$$\mathrm{Hom}(\mathbb{Z}_2^n, \mathbb{Z}_2) = \left\{\ell_\alpha : \mathbb{Z}_2^n \to \mathbb{Z}_2 \,\middle|\, \alpha \in \mathbb{Z}_2^n,\ \ell_\alpha(x) = \sum \alpha_i x_i \bmod 2\right\}$$

The BLR Theorem proved in last lecture is stronger than Theorem 1, as it applies to any G, H, (not even necessarily Abelian). On the down side, the connection between rejection probability and distance to $\mathrm{Hom}(G, H)$ there is weaker, as the BLR theorem only holds when the rejection probability is small (less than 2/9) and even then only shows the distance is *twice* the rejection probability.

1.2 Adding Structure to the Problem

To obtain the tighter Theorem 1, we need to add structure to the set of homomorphisms from $\mathbb{Z}_2^n$ to $\mathbb{Z}_2$. This structure will be provided by working in the realm of real numbers, and we start by casting our problem in this world.

Definition 2 *The* real embedding *of a Boolean function $f' : \mathbb{Z}_2^n \to \mathbb{Z}_2$ is the function $f : \mathbb{Z}_2^n \to \{-1, +1\}$ defined by $f(x) = (-1)^{f'(x)}$. For $\ell_\alpha \in \mathrm{Hom}(\mathbb{Z}_2^n, \mathbb{Z}_2)$ let $\chi_\alpha : \mathbb{Z}_2^n \to \{-1, +1\}$ be the real embedding of ℓ_α.*

Using this change of representation, Theorem 1 is reduced to:

Theorem 3 *For every function $f' : \mathbb{Z}_2^n \to \mathbb{Z}_2$ and f its real embedding, the BLR test $T^{f'}$ gives*

$$\begin{aligned}
\delta(f', \mathrm{Hom}(\mathbb{Z}_2^n, \mathbb{Z}_2)) &= \min_\alpha \left(\Pr_x\left[f(x) \neq \chi_\alpha(x)\right]\right) && (*) \\
&\leq \Pr_{x,y}\left[T^{f'} \textit{ rejects}\right] \\
&= \Pr_{x,y}\left[f(x) \cdot f(y) \neq f(x + y)\right] && (**)
\end{aligned}$$

Proof (of Theorem 1 from Theorem 3): The first equality is our definition of distance, the second inequality is what we wish to prove, and the third equality is the definition of the BLR test. The labels (*) and (**) refer to the rhs of their respective equations. ■

From here on, we focus on proving Theorem 3.

1.3 The Fourier Basis and its elementary properties

It will be useful to think of all our functions as elements in $\mathbb{R}^{2^n}$. This is done by indexing the coordinates of $\mathbb{R}^{2^n}$ by $x \in \mathbb{Z}_2^n$ and letting the x^{th} coordinate of f have value $f(x)$. We define two bases for working in $\mathbb{R}^{2^n}$. The first is the "standard" one,

$$SB = \text{standard basis for } \mathbb{R}^{2^n} = \{e_x = (0, 0, \ldots, 1 \text{ (index x)}, \ldots, 0) : x \in \mathbb{Z}_2^n\}$$

In words, the x^{th} basis element of SB has 1 in position x and is zero elsewhere. The second is known as the Fourier basis,

$$FB = \text{Fourier Basis } = \{\chi_\alpha : \alpha \in \mathbb{Z}_2^n\}$$

In words, the α-basis element of FB (denoted χ_α) is the real embedding of the homomorphism ℓ_α as in Definition 2 (we will shortly prove that this collection of vectors is indeed a basis for $\mathbb{R}^{2^n}$).

We will make use of the linearity of the Fourier basis elements. The following fact follows from the definition of χ_α.

Fact 4 (Linearity) *For every $x, y, \alpha, \beta \in \mathbb{Z}_2^n$,*

$$\chi_{\alpha+\beta}(x) = \chi_\alpha(x) \cdot \chi_\beta(x), \tag{1}$$

$$\chi_\alpha(x + y) = \chi_\alpha(x) \cdot \chi_\alpha(y) \tag{2}$$

The next fact (that once again follows from the definition of χ_α) will be used several times in our proof.

Fact 5 *Denoting $\mathbf{0} = (0, \ldots, 0)$, notice $\chi_{\mathbf{0}} = (1, 1, \ldots, 1)$ (because $\ell_{\mathbf{0}}(x) = 0$ for all x). If $\alpha \neq \mathbf{0}$, then χ_α has half of its entries 1 and the other half -1.*

To prove that FB is a (orthonormal) basis for $\mathbb{R}^{2^n}$, we define an inner product (dot product) on functions.

Definition 6 (Inner Product) *For $f, g \in \mathbb{R}^{2^n}$, let*

$$\langle f, g\rangle = 2^{-n} \sum_x f(x) \cdot g(x)$$

We can relate this inner product to the distance between functions as follows:

Claim 7 *If $f, g \in \{\mp 1\}^{2^n}$ then $\langle f, g\rangle = 1 - 2\Pr_x[f(x) \neq g(x)]$.*

Proof We use the fact that the ranges of f and g are over $\{\mp 1\}$.

$$\begin{aligned} \langle f, g\rangle &= 2^{-n} \textstyle\sum_x f(x) \cdot g(x) \\ &= 2^{-n} \left(\textstyle\sum_{x, f(x)=g(x)} 1 - \sum_{x, f(x)\neq g(x)} 1\right) \\ &= \Pr_x[f(x) = g(x)] - \Pr[f(x) \neq g(x)] \\ &= (1 - \Pr[f(x) \neq g(x)]) - \Pr[f(x) \neq g(x)] \\ &= 1 - 2\Pr_x[f(x) \neq g(x)] \end{aligned}$$

■

Corollary 8 $\forall \alpha, \beta \in \{0, 1\}^n : \langle \chi_\alpha, \chi_\beta\rangle = \begin{cases} 1 & \textit{if } \alpha = \beta \\ 0 & \textit{otherwise} \end{cases}$

Proof Follows from Fact 5 and the Definition 6. The case when $\alpha = \beta$ is clear:

$$\langle \chi_\alpha, \chi_\alpha \rangle = 2^{-n} \sum_x \chi_\alpha^2(x) = 2^{-n} \sum_x 1 = 1$$

We sum over 2^n ones, since every element of χ_α is ± 1, and divide by 2^n to get 1. Now consider the case of $(\alpha \neq \beta)$,

$$\begin{aligned} \langle \chi_\alpha, \chi_\beta \rangle &= 2^{-n} \textstyle\sum_x \chi_\alpha(x) \cdot \chi_\beta(x) \\ &= 2^{-n} \textstyle\sum_x \chi_{\alpha+\beta}(x) = 0 \end{aligned}$$

The first equality is the definition of the inner product. The second equality is Equation (1), and the last equality follows from (the second part of) Fact 5. ■

A set of vectors is said to be orthonormal if the inner product between any two distinct vectors in the set is 0, and the inner product of a vector with itself is exactly 1. It is well-known that an orthonormal set of m vectors in $\mathbb{R}^m$ forms a basis for $\mathbb{R}^m$, called an orthonormal basis. Thus, the preceding discussion directly implies the following corollary.

Corollary 9 *FB is an orthonormal basis for $\mathbb{R}^{2^n}$.*

Since FB is a basis, every $f \in \mathbb{R}^{2^n}$ can be written as a sum of basis elements:

$$f(x) = \sum_\alpha \hat{f}_\alpha \cdot \chi_\alpha(x) \tag{3}$$

The coefficients of the different basis elements are called the *Fourier coefficients* (denoted $\hat{f}_\alpha$), and are calculated by projecting f onto the α basis element, i.e. taking its inner product with χ_α. Formally,

$$\hat{f}_\alpha = \langle f, \chi_\alpha \rangle$$

Parseval's Equality tells us that if we move from an orthonormal basis in $\mathbb{R}^n$ to another orthonormal basis, the norms don't change.

$$\textbf{Parseval's Equality:} \quad \forall f \in \mathbb{R}^{2^n}, \langle f, f \rangle = \sum_{\alpha \in \{0,1\}^n} \hat{f}_\alpha^2 \tag{4}$$

Using the fact that $f \in \{-1, +1\}^{2^n}$ we get:

$$\textbf{Corollary to Parseval's Equality:} \quad \forall f \in \{-1, +1\}^{2^n}, \quad \sum_{\alpha \in \{0,1\}^n} \hat{f}_\alpha^2 = \langle f, f \rangle = 1 \tag{5}$$

We are finally ready to prove Theorem 3.

1.4 Proof of Theorem 3

Simplifying (*): Using Claim 7, we can simplify (*) in the statement of Theorem 3. From Claim 7 we have:

$$\begin{aligned} \Pr_x \left[f(x) \neq \chi_\alpha(x) \right] &= \tfrac{1}{2} \left(1 - \langle f, \chi_\alpha \rangle \right) \\ &= \tfrac{1}{2} \left(1 - \hat{f}_\alpha \right) \end{aligned}$$

We substitute this expression in (*):

$$(*) = \min_\alpha \left(\Pr_x \left[f(x) \neq \chi_\alpha(x) \right] \right) = \frac{1}{2} \left(1 - \max_\alpha \hat{f}_\alpha \right) \tag{6}$$

Simplifying (**): Now we turn to simplifying (**). We use an indicator random variable to change the probability to an expectation.

$$\begin{aligned}(**) &= \Pr_{x,y}\left[f(x)\cdot f(y) \neq f(x+y)\right] \\ &= \mathrm{E}_{x,y}\left[\tfrac{1}{2}\left(1 - f(x)\cdot f(y)\cdot f(x+y)\right)\right] \\ &= \tfrac{1}{2}\left(1 - \mathrm{E}_{x,y}\left[f(x)\cdot f(y)\cdot f(x+y)\right]\right)\end{aligned} \tag{7}$$

The indicator represents the first event because $f(x)\cdot f(y)\cdot f(x+y)$ is 1 whenever $f(x)\cdot f(y) \neq f(x+y)$. Thus the indicator is 0 when this expression is -1 and 1 when it is +1.

We analyze the last expectation in the above equation:

$$\begin{aligned}\mathrm{E}_{x,y}\left[f(x)\cdot f(y)\cdot f(x+y)\right] &= 2^{-2n}\cdot \sum_{x,y}\left(\sum_\alpha \hat{f}_\alpha \cdot \chi_\alpha(x)\right)\left(\sum_\beta \hat{f}_\beta \cdot \chi_\beta(y)\right)\left(\sum_\gamma \hat{f}_\gamma \cdot \chi_\gamma(x+y)\right) \\ &= 2^{-2n}\cdot \sum_{\alpha,\beta,\gamma} \hat{f}_\alpha \hat{f}_\beta \hat{f}_\gamma \sum_{x,y} \chi_\alpha(x)\chi_\beta(y)\chi_\gamma(x+y) \\ &= 2^{-2n}\cdot \sum_{\alpha,\beta,\gamma} \hat{f}_\alpha \hat{f}_\beta \hat{f}_\gamma \left(\sum_x \chi_\alpha(x)\chi_\gamma(x)\right)\left(\sum_y \chi_\beta(y)\chi_\gamma(y)\right) \\ &= \sum_{\alpha,\beta,\gamma} \hat{f}_\alpha \hat{f}_\beta \hat{f}_\gamma \cdot \langle \chi_\alpha, \chi_\gamma\rangle \cdot \langle \chi_\beta, \chi_\gamma\rangle \\ &= \sum_\alpha \hat{f}_\alpha^3\end{aligned} \tag{8}$$

The first equality is the representation of f in the Fourier basis. The second equality is a change in the order of summation. The third equality uses Equation (2). The fourth equality is Definition 6. The last equality uses the orthonormality of the Fourier basis (Corollary 9).

Using the Corollary to Parseval's equality (Equation (5)) finishes our analysis.

$$\begin{aligned}\mathrm{E}_{x,y}\left[f(x)\cdot f(y)\cdot f(x+y)\right] &= \sum_\alpha \hat{f}_\alpha^3 \\ &\leq \max_\alpha \hat{f}_\alpha \sum_\beta \hat{f}_\beta^2 \\ &= \max_\alpha \hat{f}_\alpha\end{aligned} \tag{9}$$

Recalling Equation (6) and incorporating Equation (9) into Equation (7), we get:

$$(**) \geq \frac{1}{2}\left(1 - \max_\alpha \hat{f}_\alpha\right) = (*)$$

This completes the proof of Theorem 3. ■

2 PCPs of Proximity

In this section we show that *every* property that can be decided by a small circuit, can also be tested with constant query complexity, provided we are allowed to ask for a bit of additional information, in the form of a *probabilistically checkable proof of proximity* (denoted PCPP). Throughout this section, a property is a subset $P \subseteq \{0,1\}^n$ (the results can be generalized to arbitrary finite alphabet). We use the relative hamming distance as our measure of proximity to P. Namely, for $x \in \{0,1\}^n$ we set

$$\delta(x, P) = \min_{y\in P} \Pr_{i\in[n]}[x_i \neq y_i]$$

We are interested in a tester that will accept with high probability every string in P and reject with high probability every string that is far from P. Unfortunately, as we saw in a previous lecture, some properties (e.g. some $\mathbb{Z}_2$ linear spaces) that can be easily decided are very hard to test (i.e. require linear query complexity).

Thus, we relax our problem to a *non-deterministic* version. Namely, we allow our tester to access a *witness*, which is simply an auxiliary sequence of bits. We still restrict the query complexity of the tester (to both the input and witness oracles) to a constant number of bits, and wish to distinguish between good inputs (in P) and those far from it. Formally,

Definition 10 (PCPP) *[4] A (n, ℓ)-*verifier *is a randomized algorithm with two oracles, one of length n and the other of length ℓ, running in time polynomial in $n+\ell$. The verifier with oracles x, π is denoted $V^{(x,\pi)}$.*

A property $P \subseteq \{0,1\}^n$ is said to have a probabilistically checkable proof of proximity (PCPP) *with* length ℓ, randomness r, query complexity q *and* proximity parameter δ_0 *if there exists a (n, ℓ)-verifier with access to oracles $x \in \{0,1\}^n, \pi \in \{0,1\}^\ell$ with the following properties.*

Operation: *Verifier tosses r coins, makes q queries (to x and π) and outputs* accept/reject.

Completeness: *If $x \in P$ then $\exists \pi \in \{0,1\}^\ell$ such that $\Pr[V^{(x,\pi)} = \mathsf{accept}] = 1$. The probability is over the random coin tosses of the verifier.*

Soundness: *If $\delta(x, P) \geq \delta_0$ then $\forall \pi \in \{0,1\}^\ell$, $\Pr[V^{x,\pi} = \mathsf{accept}] \leq 1/3$. The probability is over the random coin tosses of the verifier.*

PCPPs are implicit in the original work on PCPs [1, 2]. They are also a special case of "PCP Spot Checkers" defined by Ergun, Kumar and Rubinfeld [6]. The definition above is from Ben-Sasson et al. [4] and a similar definition was given independently by Dinur and Reingold [5]. It turns out that PCPPs are also a good way to arrive at short (and simpler) PCPs (as defined in problem set 3), but this material is out of the scope of our course.

Does adding the witness π allow us to test more properties with constant query complexity? The answer is a resounding yes. A property $P \subseteq \{0,1\}^n$ is *decided* by a Boolean circuit C (with n inputs) if the set of assignments satisfying C (i.e. setting its output to 1) is precisely P. The following Theorem says that *any* property that is decided by a small circuit, is also testable with constant query complexity when given a short auxiliary proof. We end this lecture with the following Theorem from [4], given without proof. (Dinur and Reingold [5] prove a similar result with larger proof length that is however polynomial in s).

Theorem 11 *For every constant $\delta_0 > 0$ there exists a constant q such that the following holds. Let $P \subseteq \{0,1\}^n$ be a property decided by a Boolean circuit of size s. Then P has a PCPP with length $\ell = s^{1+o(1)}$, randomness $r = \log \ell + O(\log \log \ell)$ and query complexity q.*

References

[1] Arora, S., Lund, C., Motwani, R., Sudan, M., and Szegedy, M. Proof Verification and Hardness of Approximation Problems. In *Journal of ACM*, 45(3):501-555, 1998.

[2] Arora, S. and Safra, M. Probabilistic Checking of Proofs: A New Characterization of NP. In *Journal of ACM*, 45(1):70–122, 1998.

[3] Bellare, M., Coppersmith, D., Hastad, J., Kiwi, M. and Sudan, M. Linearity testing over characteristic two. *IEEE Transactions on Information Theory*, 42(6): 1781–1795, November 1996. Preliminary version in *Proceedings of the 36th Annual Symposium on Foundations of Computer Science*, pages 432-441, Milwaukee, Wisconsin, 23-25 October 1995.

[4] Ben-Sasson, E., Goldreich, O., Harsha, P., Sudan, M., Vadhan, S. Robust PCPs of proximity, shorter PCPs and applications to coding. *Proceedings of the 36th Annual ACM Symposium on Theory of Computing*, 2004: 1–10, June 2004.

[5] Dinur, I. and Reinglod, O. Assignment Testers: Towards a combinatorial proof of the PCP theorem. In *Proceedings of FOCS'04*, 2004.

[6] Ergun, F., Kumar, R., Rubinfeld, R. Fast approximate probabilistically checkable proofs. *Information and Computation*, 189(2): 135–159, 15 Mar 2004.