

Lecture 6, Part 2: BLR Linearity Test

Mathematical Methods in TCS: Boolean Functions (Fall 2026)

University of Toronto

Swastik Kopparty

Scribe: Ali Karim Lalani

1 Recap and notation

Let $G, H : \mathbb{F}_2^n \rightarrow \mathbb{R}$. For each $a \in \mathbb{F}_2^n$, the parity function and Fourier coefficient are defined by

$$\begin{aligned}\psi_a(x) &= (-1)^{\langle x, a \rangle}, \\ \widehat{G}(a) &= \mathbb{E}_{x \in \mathbb{F}_2^n} [G(x) \psi_a(x)].\end{aligned}$$

All expectations are taken uniformly at random. The parity functions form an orthonormal basis, giving the Fourier expansion $G(x) = \sum_{a \in \mathbb{F}_2^n} \widehat{G}(a) \psi_a(x)$.

The inner product is $\langle G, H \rangle = \mathbb{E}_{x \in \mathbb{F}_2^n} [G(x) H(x)]$. Plancherel's and Parseval's identities are,

$$\begin{aligned}\langle G, H \rangle &= \sum_{a \in \mathbb{F}_2^n} \widehat{G}(a) \widehat{H}(a), \\ \mathbb{E}_{x \in \mathbb{F}_2^n} [G(x)^2] &= \sum_{a \in \mathbb{F}_2^n} \widehat{G}(a)^2.\end{aligned}$$

respectively. If the output of G is always in $\{+1, -1\}$, then $\sum_{a \in \mathbb{F}_2^n} \widehat{G}(a)^2 = 1$.

The convolution of G and H is defined by $(G * H)(z) = \mathbb{E}_x [G(x) H(z + x)]$, and as a special case $(F * F)(x) = \mathbb{E}_y [F(y) F(x + y)]$.

And the key theorem we have about convolutions is $\widehat{G * H}(a) = \widehat{G}(a) \widehat{H}(a)$.

2 Linearity testing

Suppose we are given query access to a function $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$ that is claimed to be linear. We would like to test this whether this claim is true using as few queries to f as we can.

Definition 1 (Linear function). *A function $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$ is linear if $f(x + y) = f(x) + f(y)$ for every $x, y \in \mathbb{F}_2^n$. Equivalently, it is of the form $\ell_a(x) = \langle a, x \rangle = \sum_{i=1}^n a_i x_i$ for some $a \in \mathbb{F}_2^n$.*

If we insist on testing linearity exactly using a deterministic algorithm, then one can not hope to do better than querying all 2^n inputs in the worst case. To see why, suppose every query has returned 0 and some input z has not been queried. The function could be identically zero, which is linear, or

it could be zero everywhere except at z , (clearly not linear). Both functions give the same answers to the queries, so the algorithm cannot distinguish them.

In particular, checking only the standard basis is not enough. For example, $f(x) = x_1x_2$ is zero on every standard basis vector, but it is not linear.

Rather than deciding exactly whether f is linear, we want to distinguish a linear function from one that is far from every linear function. This is an example of a *property testing* problem.

Definition 2 (Distance between functions). *For two functions $f, g : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$, their distance $\Delta(f, g)$ is the probability that they disagree with each other i.e. $\Delta(f, g) = \Pr_x[f(x) \neq g(x)]$, where x is chosen uniformly from $\mathbb{F}_2^n$.*

Definition 3 (ε -far from linear). *Let $0 < \varepsilon < 1$. A function $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$ is ε -far from linear if $\Delta(f, \ell) \geq \varepsilon$ for every linear function ℓ . Equivalently, for every $a \in \mathbb{F}_2^n$, $\Pr_x[f(x) = \langle a, x \rangle] \leq 1 - \varepsilon$.*

For a given ε , we would like to find a randomized algorithm using few queries such that:

1. If f is linear, the algorithm accepts with probability 1.
2. If f is ε -far from linear, the algorithm accepts with probability at most 0.9.

If f is nonlinear but is not ε -far from linear, we make no guarantee about the algorithm's output.

2.1 The BLR linearity test

Surprisingly, the test below just uses three queries, regardless of what n is. If f is linear, then its values at x and y determine what its value at $x + y$ must be: $f(x + y) = f(x) + f(y)$. This suggests we choose x and y uniformly at random, querying these three values, and checking whether the equation holds. This is the Blum–Luby–Rubinfeld (BLR) linearity test.

Algorithm 1 BLR linearity test

Input: Query access to $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$.

1. Choose $x, y \in \mathbb{F}_2^n$ uniformly and independently.
 2. Query f at x , y , and $x + y$.
 3. Accept if $f(x) + f(y) = f(x + y)$; otherwise, reject.
-

Theorem 4. *If f is linear, the BLR test accepts with probability 1. If f is ε -far from linear, then $\Pr[\text{the BLR test accepts}] \leq 1 - \varepsilon$. The probability is over the uniformly and independently chosen inputs $x, y \in \mathbb{F}_2^n$.*

The theorem is only for one run of the test, using at most three queries. Repeating the test independently $O(1/\varepsilon)$ times and rejecting if any run rejects gives us the constant rejection probability we wanted.

3 Proof of the BLR theorem

We proceed by contrapositive. If $\Pr[\text{BLR accepts}] > 1 - \varepsilon$, then there exists $a \in \mathbb{F}_2^n$ such that f agrees with ℓ_a on more than a $1 - \varepsilon$ fraction of inputs. We will write agreement and acceptance in terms of Fourier coefficients, and then use Parseval to conclude.

To use Fourier analysis, we regard $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$ as a real-valued function $F : \mathbb{F}_2^n \rightarrow \{-1, +1\}$ defined by $F(x) = (-1)^{f(x)}$. Below, x, y are independent uniform inputs in $\mathbb{F}_2^n$.

3.1 Agreement

For $a \in \mathbb{F}_2^n$, let $\ell_a(x) = \langle x, a \rangle$. Write $\text{agree}(f, \langle \cdot, a \rangle) = \Pr_x[f(x) = \langle x, a \rangle]$ for the fraction of inputs on which f agrees with ℓ_a . Since $\frac{1 + (-1)^{f(x) - \langle x, a \rangle}}{2}$ is 1 when the two functions agree and 0 otherwise,

$$\begin{aligned} \text{agree}(f, \langle \cdot, a \rangle) &= \mathbb{E}_x \left[\frac{1 + (-1)^{f(x) - \langle x, a \rangle}}{2} \right] \\ &= \frac{1}{2} + \frac{1}{2} \mathbb{E}_x \left[(-1)^{f(x)} (-1)^{\langle x, a \rangle} \right] \\ &= \frac{1}{2} + \frac{1}{2} \mathbb{E}_x [F(x) \psi_a(x)] \\ &= \frac{1 + \widehat{F}(a)}{2}. \end{aligned}$$

Thus, a Fourier coefficient close to 1 gives high agreement with a linear function. In particular, finding an a with $\widehat{F}(a) > 1 - 2\varepsilon$ will give the desired agreement.

3.2 Acceptance probability

The BLR test accepts if and only if $F(x)F(y) = F(x + y)$, or equivalently $F(x)F(y)F(x + y) = 1$. Hence,

$$\frac{1 + F(x)F(y)F(x + y)}{2} = \begin{cases} 1, & \text{if the test accepts on } (x, y), \\ 0, & \text{if the test rejects on } (x, y). \end{cases}$$

Taking expectations, $\Pr[\text{BLR accepts}] = \frac{1 + \mathbb{E}_{x,y}[F(x)F(y)F(x+y)]}{2}$.

Taking the expectation over y first, and recognizing the convolution formula we have

$$\begin{aligned} \mathbb{E}_{x,y}[F(x)F(y)F(x + y)] &= \mathbb{E}_x [F(x) \mathbb{E}_y [F(y)F(x + y)]] \\ &= \mathbb{E}_x [F(x)(F * F)(x)] \\ &= \langle F, F * F \rangle. \end{aligned}$$

By Plancherel and the convolution theorem,

$$\begin{aligned}
\Pr[\text{BLR accepts}] &= \frac{1 + \langle F, F * F \rangle}{2} \\
&= \frac{1 + \sum_{a \in \mathbb{F}_2^n} \widehat{F}(a) \widehat{F * F}(a)}{2} \\
&= \frac{1 + \sum_{a \in \mathbb{F}_2^n} \widehat{F}(a)^3}{2}.
\end{aligned}$$

3.3 Finishing the proof

We want to show that if BLR accepts with high probability, then f is close to some linear function. More precisely,

$$\Pr[\text{BLR accepts}] > 1 - \varepsilon \implies \exists a^* \in \mathbb{F}_2^n \text{ s.t. } \text{agree}(f, \langle \cdot, a^* \rangle) > 1 - \varepsilon.$$

Assume $\Pr[\text{BLR accepts}] > 1 - \varepsilon$. Using the acceptance formula and Parseval, we obtain

$$\begin{aligned}
1 - 2\varepsilon &< 2 \Pr[\text{BLR accepts}] - 1 \\
&= \sum_{a \in \mathbb{F}_2^n} \widehat{F}(a)^3 \\
&= \sum_{a \in \mathbb{F}_2^n} \widehat{F}(a) \widehat{F}(a)^2 \\
&\leq \left(\max_{a \in \mathbb{F}_2^n} \widehat{F}(a) \right) \sum_{a \in \mathbb{F}_2^n} \widehat{F}(a)^2 \\
&= \max_{a \in \mathbb{F}_2^n} \widehat{F}(a).
\end{aligned}$$

The max step works even if the max is negative, since all the $\widehat{F}(a)^2$ are nonnegative.

Thereforw, there is an $a^* \in \mathbb{F}_2^n$ such that $\widehat{F}(a^*) > 1 - 2\varepsilon$. By the agreement formula,

$$\text{agree}(f, \langle \cdot, a^* \rangle) = \frac{1 + \widehat{F}(a^*)}{2} > \frac{1 + (1 - 2\varepsilon)}{2} = 1 - \varepsilon.$$

Thus, f disagrees with the $\ell_{a^*}(x) = \langle x, a^* \rangle$ on fewer than an ε fraction of inputs, so it is not ε -far from linear. This proves the contrapositive as desired.

Finally, if f is linear, the test accepts for every pair x, y , so its acceptance probability is 1. $\square$

3.4 An alternative calculation by direct Fourier expansion

Just for fun, we can also compute the acceptance probability by expanding the Fourier sums directly. We already showed that

$$\Pr[\text{BLR accepts}] = \frac{1 + \mathbb{E}_{x,y}[F(x)F(y)F(x+y)]}{2}.$$

Fourier expanding the three functions yields

$$\begin{aligned} & \mathbb{E}_{x,y}[F(x)F(y)F(x+y)] \\ &= \mathbb{E}_{x,y} \left[\left(\sum_{a \in \mathbb{F}_2^n} \widehat{F}(a) \psi_a(x) \right) \left(\sum_{b \in \mathbb{F}_2^n} \widehat{F}(b) \psi_b(y) \right) \left(\sum_{c \in \mathbb{F}_2^n} \widehat{F}(c) \psi_c(x+y) \right) \right] \\ &= \mathbb{E}_{x,y} \left[\sum_{a,b,c \in \mathbb{F}_2^n} \widehat{F}(a) \widehat{F}(b) \widehat{F}(c) \psi_a(x) \psi_b(y) \psi_c(x+y) \right] \\ &= \sum_{a,b,c \in \mathbb{F}_2^n} \widehat{F}(a) \widehat{F}(b) \widehat{F}(c) \mathbb{E}_{x,y}[\psi_a(x) \psi_b(y) \psi_c(x+y)]. \end{aligned}$$

Now, look at the expectation in each summand. Recall $\psi_c(x+y) = \psi_c(x)\psi_c(y)$, so

$$\begin{aligned} & \mathbb{E}_{x,y}[\psi_a(x) \psi_b(y) \psi_c(x+y)] \\ &= \mathbb{E}_{x,y}[\psi_a(x) \psi_c(x) \psi_b(y) \psi_c(y)] \\ &= \mathbb{E}_x[\psi_a(x) \psi_c(x)] \mathbb{E}_y[\psi_b(y) \psi_c(y)] \\ &= \begin{cases} 1, & a = b = c, \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

The second equality is because x and y are independent, and the last equality is by the orthonormality of parity functions. The first expectation is 1 when $a = c$ and 0 otherwise, while the second is 1 when $b = c$ and 0 otherwise.

Thus, only the terms with $a = b = c$ survive in the sum. Each contributes $\widehat{F}(a)^3$, giving

$$\mathbb{E}_{x,y}[F(x)F(y)F(x+y)] = \sum_{a \in \mathbb{F}_2^n} \widehat{F}(a)^3.$$

Plugging into the acceptance probability formula, we again have

$$\Pr[\text{BLR accepts}] = \frac{1 + \sum_{a \in \mathbb{F}_2^n} \widehat{F}(a)^3}{2}.$$