Foundations of Probabilistic Proofs

A course by **Alessandro Chiesa**

Lecture 07

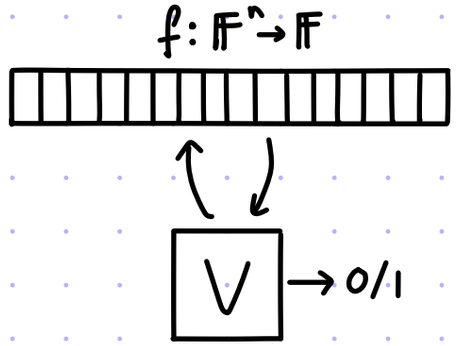
Linearity Testing

Warmup 1: All-Zero Testing

A function $f:\mathbb{F}^n \to \mathbb{F}$ is ALL-ZERO if $\forall x \in \mathbb{F}^n \; f(x)=0$.

QUESTION: is there a $O(1)$-query test V s.t. $\forall f:\mathbb{F}^n \to \mathbb{F}$

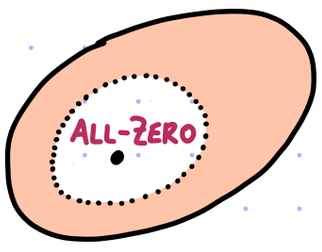
- COMPLETENESS: if f is ALL-ZERO then $\Pr[V^f=1]=1$
- SOUNDNESS: if f is not ALL-ZERO then $\Pr[V^f=1] \le 1/2$



ANSWER: No. (If f is $\neq 0$ at a single location, how can a $O(1)$-query V detect this?)

RELAXED QUESTION: is there a $O(1)$-query test V s.t. $\forall f:\mathbb{F}^n \to \mathbb{F}$

- COMPLETENESS: if f is ALL-ZERO then $\Pr[V^f=1]=1$
- SOUNDNESS: if f is far from ALL-ZERO then $\Pr[V^f=1] \le 1/2$



We use (relative) Hamming distance: $\Delta(f,g) := \Pr_{x \in \mathbb{F}^n}[f(x) \neq g(x)]$ and $\Delta(f,S) := \min_{g \in S} \Delta(f,g)$.

ANSWER: YES! V_0^f := sample random $x \in \mathbb{F}^n$ and check that $f(x)=0$

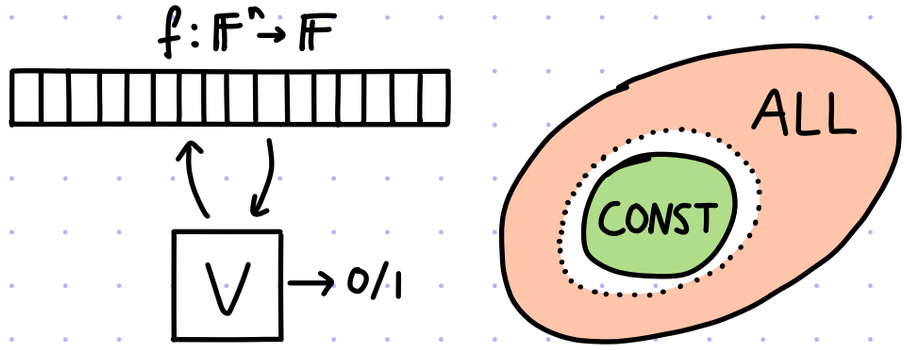
- if f is ALL-ZERO then $\Pr[V^f=1] = \Pr_x[f(x)=0] = 1$
- if $\Delta(f, \text{ALL-ZERO}) \ge \delta$ then $\Pr[V^f=1] = 1-\Pr[V^f=0] = 1-\Pr_x[f(x)\neq 0] = 1-\Delta(f,\text{ALL-ZERO}) \le 1-\delta$.

Warmup 2: Constant Testing

Define the set $\text{CONST} := \{f : \mathbb{F}^n \to \mathbb{F} : \exists c \in \mathbb{F} \text{ s.t. } \forall x \in \mathbb{F}^n \ f(x) = c\}$ (constant functions).

We cannot expect to distinguish (with small error) $f \in \text{CONST}$ vs $f \notin \text{CONST}$ by querying f at few locations.

What about distinguishing $f \in \text{CONST}$ vs $\Delta(f, \text{CONST}) \geq \delta$?



$V^f :=$ Sample random $x \in \mathbb{F}^n$ and check that $f(x) = f(0^n)$.

- If $f \in \text{CONST}$ then $\exists c \in \mathbb{F}$ s.t. $\forall x \in \mathbb{F}^n \ f(x) = c$

  so $\Pr[V^f = 1] = \Pr_x[f(x) = f(0^n)] = \Pr_x[c = c] = 1$.

- If $\Delta(f, \text{CONST}) \geq \delta$ then $\Pr[V^f = 1] = 1 - \Pr[V^f = 0]$

$$= 1 - \Pr_x[f(x) \neq f(0^n)]$$

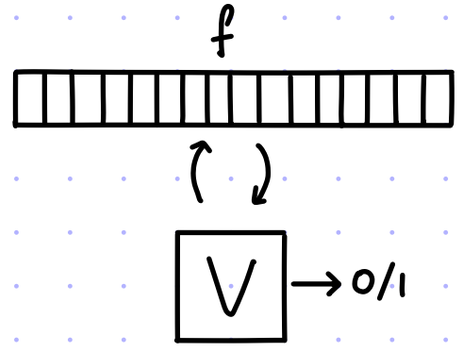
$$\leq 1 - \min_{c \in \mathbb{F}} \Pr_x[f(x) \neq c]$$

$$= 1 - \Delta(f, \text{CONST}) \leq 1 - \delta.$$

Proximity Testing

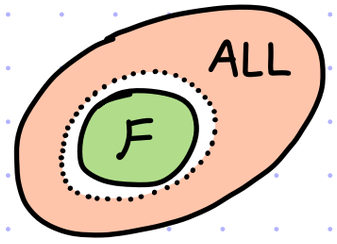
Both warmups are examples of problems in PROPERTY TESTING.

A PROPERTY is a set of functions $F = \{f : D \to \Sigma\}$.



def: V is a test for the property F with proximity error ε if

- COMPLETENESS: $\forall f \in F \quad \Pr[V^f = 1] = 1$
- SOUNDNESS: $\forall \delta \in [0,1] \; \forall f$ with $\Delta(f,F) \geq \delta \quad \Pr[V^f = 1] \leq \varepsilon(\delta)$



The distance Δ is usually (relative) Hamming distance.

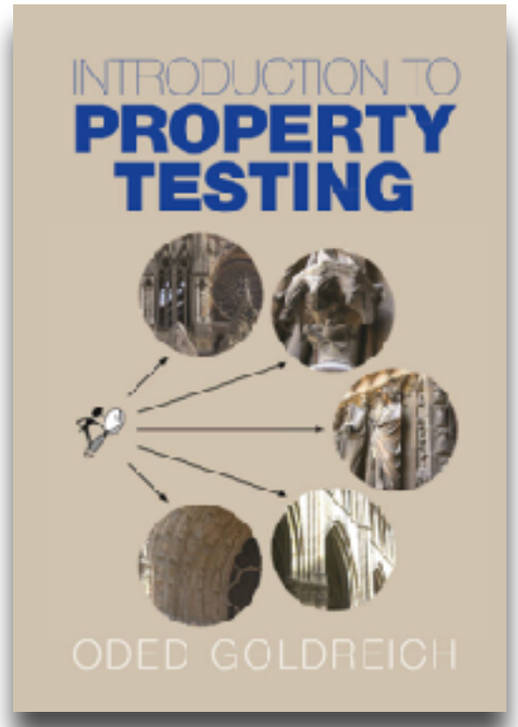
Sometimes Δ is ℓ_p distance (e.g. when $\Sigma = [0,1]$) or other metrics.

MAIN GOAL: minimize the query complexity q of the test V

In warmup 1: $F = \{\text{ALL-ZERO}\}$, $\varepsilon(\delta) = 1-\delta$, $q = 1$

In warmup 2: $F = \text{CONST}$, $\varepsilon(\delta) = 1-\delta$, $q = 2$

See Goldreich's book for an introduction to property testing. →

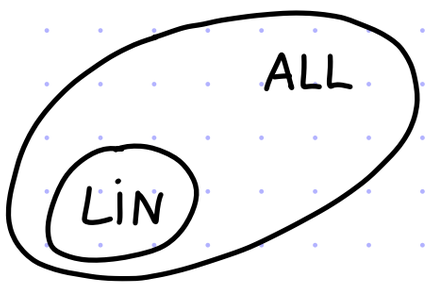


Today we are interested in testing an algebraic property:

LINEARITY TESTING

Linear Functions

A function $f: \mathbb{F}^n \to \mathbb{F}$ is linear if $\exists c \in \mathbb{F}^n$ s.t. $f(x) = \sum_{i=1}^{n} c_i x_i$.



$ALL = \{ f: \mathbb{F}^n \to \mathbb{F} \}$ $\quad |ALL| = |\mathbb{F}|^{|\mathbb{F}|^n}$

$LIN = \{ f: \mathbb{F}^n \to \mathbb{F} \text{ is linear} \}$ $\quad |LIN| = |\mathbb{F}|^n$

The set LIN is known as the HADAMARD CODE.

LIN is a linear error-correcting code (since LIN is an $\mathbb{F}$-linear space).

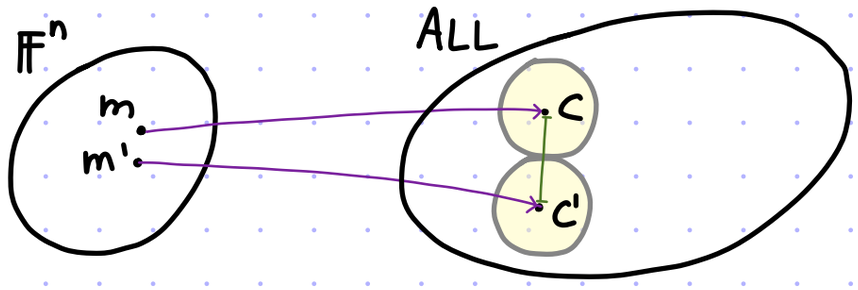
A message $m \in \mathbb{F}^n$ is mapped to the codeword $c := (\langle a, m \rangle)_{a \in \mathbb{F}^n}$.

Parameters of the code: message length = n

block length = $|\mathbb{F}|^n$

relative distance = $1 - \frac{1}{|\mathbb{F}|}$

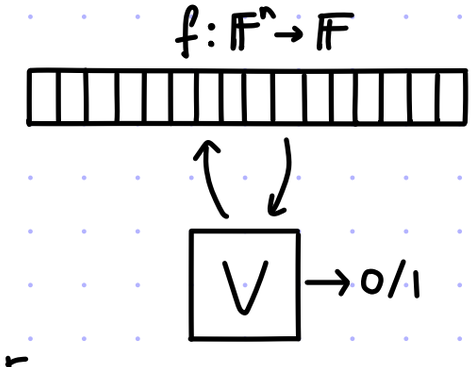
$(\forall \text{ distinct } f, g \in LIN \quad \Pr_{x \in \mathbb{F}^n}[f(x) = g(x)] \leq \frac{1}{|\mathbb{F}|})$



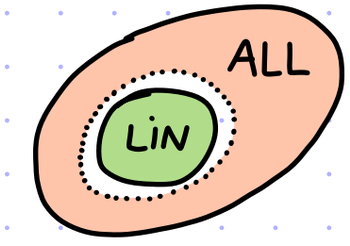
We CANNOT distinguish (with small error) $f \in LIN$ vs $f \notin LIN$ by querying f at few locations.

If $f \notin LIN$ differs in 1 location from $\overline{f} \in LIN$,

no $O(1)$-query test V detects that $f \notin LIN$ with constant soundness error.



Linearity Testing

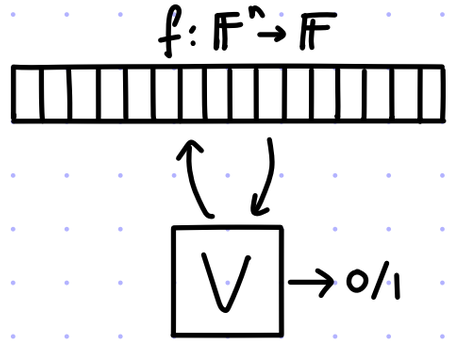


GOAL: decide between "$f \in LIN$" and "f is far from LIN".

def: V is a LINEARITY TEST (for the field $\mathbb{F}$) with proximity error ε if

- COMPLETENESS: $\forall f \in LIN \quad \Pr[V^f = 1] = 1$
- SOUNDNESS: $\forall \delta \in [0,1] \; \forall f$ with $\Delta(f, LIN) \geq \delta \quad \Pr[V^f = 1] \leq \varepsilon(\delta)$

(relative) Hamming distance



EXAMPLE: $V^{f:\mathbb{F}^n \to \mathbb{F}}$:= Sample random $x \in \mathbb{F}^n$ and query f at $x, e_1, \ldots, e_n$.

Check that $f(x) = \sum_{i=1}^n f(e_i) \cdot x_i$.

$e_i := 0^{i-1} 1 0^{n-i}$

n+1 queries

- if $f \in LIN$ then $\exists c \in \mathbb{F}^n$ s.t. $f(x) = \sum_{i=1}^n c_i x_i$

  so $\Pr[V^f = 1] = \Pr_x[f(x) = \sum_{i=1}^n f(e_i) x_i] = \Pr_x[\sum_{i=1}^n c_i x_i = \sum_{i=1}^n c_i x_i] = 1$.
- if $\Delta(f, LIN) \geq \delta$ then

  $\Pr[V^f = 1] = 1 - \Pr[V^f = 0] = 1 - \Pr_x[f(x) \neq \sum_{i=1}^n f(e_i) x_i] \leq 1 - \min_{g \in LIN} \Pr_x[f(x) \neq g(x)] = 1 - \Delta(f, LIN) \leq 1 - \delta$.

PROBLEM: query complexity is LARGE

In fact everything we did so far is TRIVIAL: queries = {queries to determine which function $f \in F$} ∪ {1 random query}

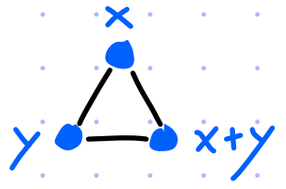
- no queries for ALL-ZERO
- 1 query for CONST
- n queries for LIN

Q: Is there a non-trivial (e.g. constant query) linearity test?

The Blum–Luby–Rubinfeld Test

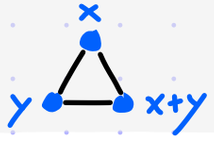


The idea is to leverage DUALITY: $f \in LIN \leftrightarrow \underbrace{\forall x,y \in \mathbb{F}^n \quad f(x)+f(y)=f(x+y)}_{|\mathbb{F}|^{2n} \text{ local constraints}}$



The BLR test for linearity:

$V_{BLR}^{f:\mathbb{F}^n\to\mathbb{F}}$:=
1. Sample $x,y \in \mathbb{F}^n$
2. Check that $f(x)+f(y)=f(x+y)$



randomness: $2n$ field elements

queries: 3 locations of f

Completeness: if $f \in LIN$ then $\forall x,y \in \mathbb{F}^n \quad f(x)+f(y)=f(x+y)$ so $\Pr\left[V_{BLR}^f = 1\right] = 1$

Soundness: non-trivial. (An example of a LOCAL-TO-GLOBAL PHENOMENON.)

theorem: $\Pr\left[V_{BLR}^f = 0\right] \geq \min\left\{1/6, \frac{1}{2}\cdot\Delta(f,LIN)\right\}$

Equivalently:

$\Pr\left[V_{BLR}^f = 1\right] \leq \max\left\{5/6, 1-\frac{1}{2}\Delta(f,LIN)\right\}$

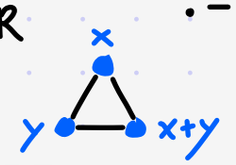
Intuition:

- if f is linear then each $y \in \mathbb{F}^n$ "votes" for the same value of x: $\forall y \in \mathbb{F}^n$, $f(x+y)-f(y)=f(x)$
- if f is not linear then we can still consider, for every x, the most popular value:

the plurality correction $g_f:\mathbb{F}^n\to\mathbb{F}$ is $g_f(x) := \arg\max_{v\in\mathbb{F}} \left|\left\{ y\in\mathbb{F}^n \;\middle|\; v = f(x+y)-f(y) \right\}\right|$

Proof overview

$V_{BLR}^{f:\mathbb{F}^n\to\mathbb{F}}$:= 1. Sample $x,y\in\mathbb{F}^n$



2. Check that $f(x)+f(y)=f(x+y)$

theorem: $\Pr\left[V_{BLR}^{f}=0\right] \geq \min\left\{1/6, \frac{1}{2}\cdot\Delta(f,\mathrm{LIN})\right\}$

The plurality correction is

$g_f:\mathbb{F}^n\to\mathbb{F}$ where $g_f(x) := \arg\max_{v\in\mathbb{F}} \left|\left\{y\in\mathbb{F}^n \mid v=f(x+y)-f(y)\right\}\right|$

- Part 1: $\Pr\left[V_{BLR}^{f}=0\right] \geq \frac{1}{2}\cdot\Delta(f,g_f)$ — far from plurality correction → many bad triangles
- Part 2: $\Pr\left[V_{BLR}^{f}=0\right] < \frac{1}{6} \rightarrow g_f\in\mathrm{LIN}$ — few bad triangles → plurality correction is linear

Conclusion:

- If $\Pr\left[V_{BLR}^{f}=0\right] \geq 1/6$ then we are done.
- If $\Pr\left[V_{BLR}^{f}=0\right] < 1/6$ then (by Part 2) g_f is linear and (by Part 1) we get

$$\Pr\left[V_{BLR}^{f}=0\right] \geq \frac{1}{2}\cdot\Delta(f,g_f) \geq \frac{1}{2}\cdot\Delta(f,\mathrm{LIN}).$$ ■

Analysis of BLR Test - Part 1

The plurality correction of f is $g_f(x) := \arg\max_{v \in \mathbb{F}} \left| \{ y \in \mathbb{F}^n \mid v = f(x+y) - f(y) \} \right|$.

If g_f is far from f then V^f_{BLR} rejects with high probability:

<u>claim:</u> $\Pr\left[V^f_{BLR} = 0 \right] \geq \frac{1}{2} \cdot \Delta(f, g_f)$.

But $\Delta(f, g_f) = 0 \not\Rightarrow \Pr[V^f_{BLR} = 0] = 0$.
Exercise: find a counterexample.

<u>proof:</u> Define $S := \left\{ x \in \mathbb{F}^n \mid \Pr_{y \leftarrow \mathbb{F}^n}\left[f(x) \neq f(x+y) - f(y) \right] \geq \frac{1}{2} \right\}$.

For every $x \notin S$, $\Pr_{y \leftarrow \mathbb{F}^n}\left[f(x) = f(x+y) - f(y) \right] > \frac{1}{2}$ (more than half of y's vote for $f(x)$), so $f(x) = g_f(x)$.

Hence $\Delta(f, g_f) \leq \frac{|S|}{|\mathbb{F}|^n}$ ($\forall x$ if $f(x) \neq g(x)$ then $x \in S$).

$$\begin{aligned}
\text{So } \Pr\left[V^f_{BLR} = 0 \right] &= \Pr_x[x \in S] \cdot \Pr_{x,y}\left[V^f_{BLR} = 0 \mid x \in S \right] + \Pr_x[x \notin S] \cdot \Pr_{x,y}\left[V^f_{BLR} = 0 \mid x \notin S \right] \\
&\geq \frac{|S|}{|\mathbb{F}|^n} \cdot \min_{x \in S} \left\{ \Pr_y\left[f(x) \neq f(x+y) - f(y) \right] \right\} + 0 \\
&\geq \frac{|S|}{|\mathbb{F}|^n} \cdot \frac{1}{2} \geq \Delta(f, g_f) \cdot \frac{1}{2}.
\end{aligned}$$

■

Analysis of BLR Test - Collision Lemma

We show that few bad triangles imply many votes for the plurality correction.

claim: $\forall\, x \in \mathbb{F}^n \quad \Pr_{y \leftarrow \mathbb{F}^n}\left[\, g_f(x) = f(x+y) - f(y) \,\right] \geq 1 - 2 \cdot \Pr\left[V^f_{BLR} = 0 \right].$

proof: $\Pr_{y \leftarrow \mathbb{F}^n}\left[\, g_f(x) = f(x+y) - f(y) \,\right] = \max_{v \in \mathbb{F}} \Pr_{y \leftarrow \mathbb{F}^n}\left[v = f(x+y) - f(y) \right]$

$\Sigma_i p_i^2 \leq \max\{p_i\}_i \cdot \Sigma_i p_i$ $\geq \sum_{v \in \mathbb{F}} \Pr_{y \leftarrow \mathbb{F}^n}\left[v = f(x+y) - f(y) \right]^2 = \sum_{v \in \mathbb{F}} \Pr_{y \leftarrow \mathbb{F}^n}\left[v = f(x+y) - f(y) \right] \Pr_{z \leftarrow \mathbb{F}^n}\left[v = f(x+z) - f(z) \right]$

independence $= \sum_{v \in \mathbb{F}} \Pr_{y,z \leftarrow \mathbb{F}^n}\left[v = f(x+y) - f(y) \wedge v = f(x+z) - f(z) \right] = \Pr_{y,z \leftarrow \mathbb{F}^n}\left[f(x+y) - f(y) = f(x+z) - f(z) \right]$

$\geq 1 - 2 \cdot \Pr\left[V^f_{BLR} = 0 \right].$

We now analyze the COLLISION PROBABILITY.

Note that $f(x+y) - f(y) \neq f(x+z) - f(z) \leftrightarrow f(x+y) + f(z) \neq f(x+z) + f(y)$

$\rightarrow f(x+y) + f(z) \neq f(x+y+z)$ OR $f(x+z) + f(y) \neq f(x+y+z)$.

Hence $\Pr_{y,z}\left[f(x+y) - f(y) \neq f(x+z) - f(z) \right]$

$\leq \Pr_{y,z}\left[f(x+y) + f(z) \neq f(x+y+z) \right] + \Pr_{y,z}\left[f(x+z) + f(y) \neq f(x+y+z) \right]$

$\leq 2 \cdot \Pr\left[V^f_{BLR} = 0 \right]$ (because $(x+y, z)$ and $(x+z, y)$ are random in $\mathbb{F}^n \times \mathbb{F}^n$). ■

Analysis of BLR Test - Part 2

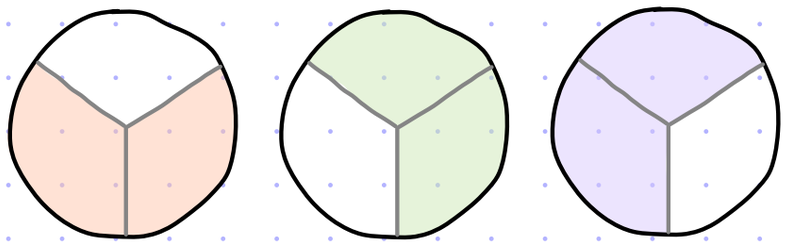
claim: if $\Pr[V^f_{BLR} = 0] < \frac{1}{6}$ then $g_f \in LIN$

proof: Fix $x, y \in \mathbb{F}^n$. We show that $g_f(x) + g_f(y) = g_f(x+y)$.

- $\Pr_z[g_f(x) = f(x+z) - f(z)] \geq 1 - 2 \cdot \Pr[V^f_{BLR} = 0] > \frac{2}{3}$ (collision lemma)
- $\Pr_z[g_f(y) = f(z) - f(z-y)] = \Pr_z[g_f(y) = f(y+z) - f(z)] \geq 1 - 2 \cdot \Pr[V^f_{BLR} = 0] > \frac{2}{3}$ (rename $z-y$ to z; collision lemma)
- $\Pr_z[g_f(x+y) = f(x+z) - f(z-y)] = \Pr_z[g_f(x+y) = f(x+y+z) - f(z)] \geq 1 - 2 \cdot \Pr[V^f_{BLR} = 0] > \frac{2}{3}$ (rename $z-y$ to z; collision lemma)

Hence $\exists z^* \in \mathbb{F}^n$ s.t.

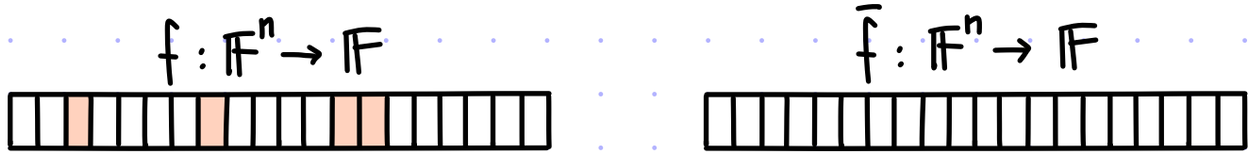
$$\left\{ \begin{array}{lll} g_f(x) & = f(x+z^*) & - f(z^*) \\ g_f(y) & = f(z^*) & - f(z^*-y) \\ g_f(x+y) & = f(x+z^*) & - f(z^*-y) \end{array} \right\}.$$



We conclude that $g_f(x) + g_f(y) = f(x+z^*) - f(z^*-y) = g_f(x+y)$. ■

Local Correction of Linear Functions [1/2]

Suppose that $f: \mathbb{F}^n \to \mathbb{F}$ is ε-close to linear: $\exists\, \bar{f} \in LIN$ s.t. $\Delta(f, \bar{f}) \leq \varepsilon$.



GOAL: given $x \in \mathbb{F}^n$ and oracle access to f, output $\bar{f}(x)$.

SOLUTION: leverage the fact that $\forall\, y \in \mathbb{F}^n \;\; \bar{f}(x) = \bar{f}(x+y) - \bar{f}(y)$.

$A^f(x)$: Sample $y \in \mathbb{F}^n$ and output $f(x+y) - f(y)$.

claim: $\Pr[A^f(x) \neq \bar{f}(x)] \leq 2 \cdot \varepsilon$

proof: Since y is random in $\mathbb{F}^n$, we know that:

- $\Pr_y[f(y) \neq \bar{f}(y)] \leq \varepsilon$
- $\Pr_y[f(x+y) \neq \bar{f}(x+y)] \leq \varepsilon$

We conclude that $\Pr_y[f(x+y) - f(y) \neq \bar{f}(x)] = \Pr_y[f(x+y) - f(y) \neq \bar{f}(x+y) - \bar{f}(y)]$

$\leq \Pr_y[f(y) \neq \bar{f}(y) \vee f(x+y) \neq \bar{f}(x+y)] \leq 2\varepsilon$. ■

Local Correction of Linear Functions [2/2]

We can reduce error via REPETITION:

$A^f(x,t)$: Sample $y_1,\dots,y_t \in \mathbb{F}^n$ and output $\operatorname{plurality}_{i\in[t]}\{f(x+y_i)-f(y_i)\}$.

(most frequent value in the set)

claim: if $\bar{f}\in \text{LIN}$ and $\Delta(f,\bar{f})\le\varepsilon\le\frac{1}{4}$ then $\Pr[A^f(x,t)\neq\bar{f}(x)]\le 2\cdot e^{-\frac{t}{4}\cdot(\frac{1}{2}-2\varepsilon)^2}$.

proof: By a CONCENTRATION argument.

Recall (one version of) the Chernoff Bound:

Let $(Z_i)_{i\in[t]}$ be independent random variables in $[0,1]$.

Define $Z:=\frac{1}{t}\sum_{i\in[t]}Z_i$. Then $\forall\gamma\ge 0\quad \Pr[|Z-\mathbb{E}[Z]|\ge\gamma]\le 2\cdot e^{-\frac{t}{4}\cdot\gamma^2}$.

Define $Z_i := "f(x+y_i)-f(y_i)\neq\bar{f}(x)"$. Note that $\mathbb{E}[Z]\le 2\cdot\varepsilon\le\frac{1}{2}$.

If $\operatorname{plurality}_{i\in[t]}\{f(x+y_i)-f(y_i)\}\neq\bar{f}(x)$ then $\sum_{i\in[t]}Z_i\ge t/2$.

Hence
$$\Pr[A^f(x,t)\neq\bar{f}(x)]\le\Pr[\textstyle\sum_{i\in[t]}Z_i\ge t/2]=\Pr[Z\ge 1/2]$$
$$=\Pr[Z-\mathbb{E}[Z]\ge 1/2-\mathbb{E}[Z]]$$
$$\le\Pr[|Z-\mathbb{E}[Z]|\ge 1/2-\mathbb{E}[Z]]\le 2\cdot e^{-\frac{t}{4}\cdot(1/2-\mathbb{E}[Z])^2}\le 2\cdot e^{-\frac{t}{4}\cdot(1/2-2\varepsilon)^2}.$$

($\mathbb{E}[Z]\le\frac{1}{2}$; Chernoff Bound) ■

Homomorphism Testing

The analysis we saw is the COMBINATORIAL ANALYSIS of the BLR test.

It extends, essentially with no changes, to achieve HOMOMORPHISM TESTING.

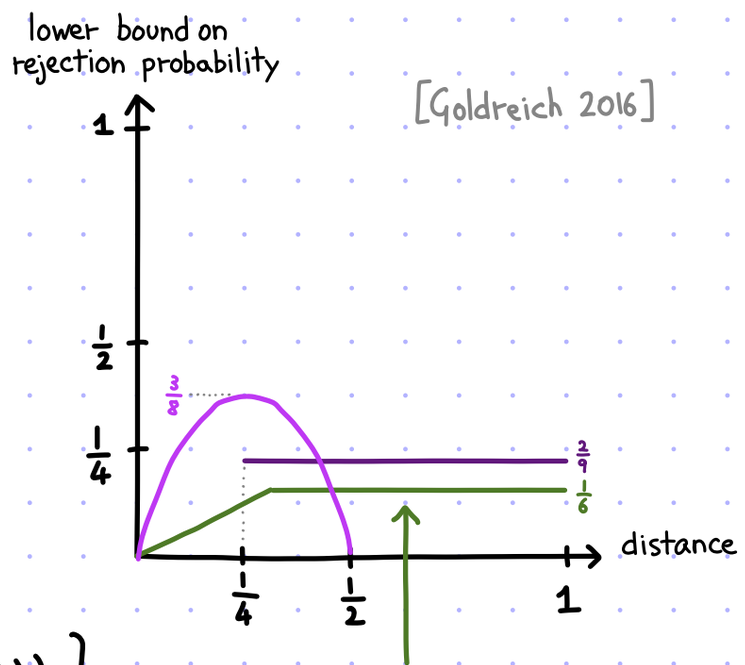
Let G, H be groups. The set of group homomorphisms from G to H is

$$\mathrm{Hom}(G,H) := \{ f: G \to H \mid \forall x,y \in G \quad f(x) +_H f(y) = f(x +_G y) \}.$$

Example:
$\mathrm{Hom}((\mathbb{F}^n,+),(\mathbb{F},+)) = \mathrm{LIN}$

The BLR test extends naturally:

$V_{BLR}^{f:G\to H} :=$ 1. Sample $x,y \in G$.
2. Check that $f(x) +_H f(y) = f(x +_G y)$.



Completeness: if $f \in \mathrm{Hom}(G,H)$ then $\Pr[V_{BLR}^f = 1] = 1$

Soundness: $\Pr[V_{BLR}^f = 0] \geq \min\left\{\frac{1}{6}, \frac{1}{2}\cdot\Delta(f,\mathrm{Hom}(G,H))\right\}$ ← today's analysis

The lower bound can be somewhat improved (see diagram).

OPEN: determine the function $\varepsilon:[0,1]\to[0,1]$ s.t. $\Pr[V_{BLR}^f = 0] = \varepsilon(\Delta(f,\mathrm{Hom}(G,H)))$.

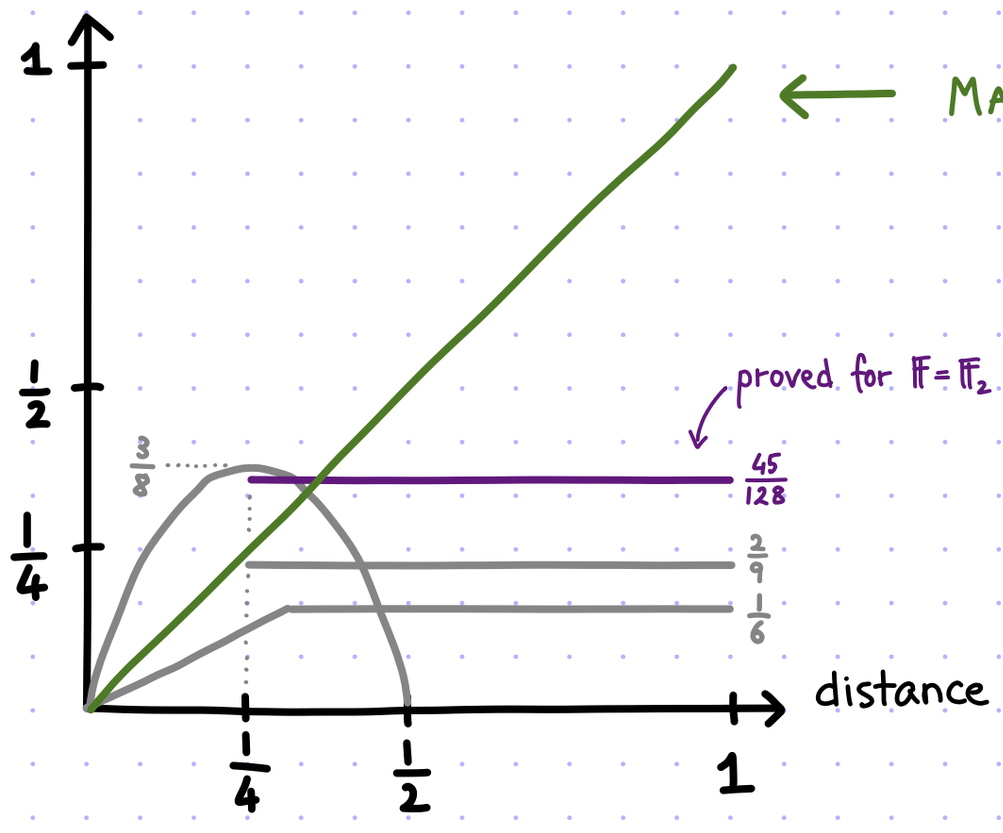
Improved Analysis for Finite Fields [1/4]

A different analysis, for linear functions over finite fields, achieves an improved bound.

Theorem: $\forall\, f:\mathbb{F}^n \to \mathbb{F},\ \Pr\left[V_{BLR}^f = 0\right] \geq \Delta(f, LIN)$

lower bound on rejection probability

$V_{BLR}^{f:\mathbb{F}^n\to\mathbb{F}}$:=
1. Sample $x, y \in \mathbb{F}^n$ and $a, b \in \mathbb{F}\setminus\{0\}$.
2. Check that $af(x)+bf(y) = f(ax+by)$.



← MAIN TOOL: Fourier analysis

The gray curves carry over from the generic analyses for homomorphism testing (including today's $\min\{\frac{1}{6}, \frac{1}{2}\cdot\delta\}$ curve).

Improved Analysis for Finite Fields [2/4]

Theorem: $\forall f: \mathbb{F}^n \to \mathbb{F}$, $\Pr\left[V_{BLR}^f = 0\right] \geq \Delta(f, LIN)$

$V_{BLR}^{f:\mathbb{F}^n\to\mathbb{F}}$:=

1. Sample $x, y \in \mathbb{F}^n$ and $a, b \in \mathbb{F}\setminus\{0\}$.
2. Check that $af(x)+bf(y) = f(ax+by)$.

We outline the proof approach.

Let $q = p^e$ be the prime-power size of $\mathbb{F}$.

For $\alpha \in \mathbb{F}_q^n$, the character function $\chi_\alpha: \mathbb{F}_q^n \to \mathbb{C}$ is $\chi_\alpha(x) := \omega_p^{Tr(\langle \alpha, x\rangle)}$.

field trace map $Tr: \mathbb{F}_q \to \mathbb{F}_p$, $Tr(x) := \sum_{i=0}^{e-1} x^{p^i} \bmod p$

p-th root of unity $\omega_p := e^{\frac{2\pi i}{p}}$

(If $q=p$ then $Tr(x)=x$ so χ_α simplifies to $\chi_\alpha(x) = \omega_p^{\langle \alpha, x\rangle}$.)

The set $\{\chi_\alpha\}_{\alpha\in\mathbb{F}_q^n}$ is an orthonormal basis for the functions $\{g: \mathbb{F}_q^n \to \mathbb{C}\}$

with the inner product $\langle g, h\rangle := \mathop{\mathbb{E}}_{x\leftarrow\mathbb{F}_q^n}[g(x)\overline{h(x)}]$. (complex conjugate)

Hence, every $g: \mathbb{F}_q^n \to \mathbb{C}$ can be written uniquely as $g(x) = \sum_{\alpha\in\mathbb{F}_q^n} \hat{g}(\alpha)\chi_\alpha(x)$

where $\{\hat{g}(\alpha)\}_{\alpha\in\mathbb{F}_q^n} := \{\langle g, \chi_\alpha\rangle\}_{\alpha\in\mathbb{F}_q^n}$ are g's Fourier coefficients.

Useful for derivations:

- Parseval's identity: $\langle g, g\rangle = \sum_{\alpha\in\mathbb{F}_q^n} \hat{g}(\alpha)\overline{\hat{g}(\alpha)} = \sum_{\alpha\in\mathbb{F}_q^n} |\hat{g}(\alpha)|^2$
- Plancherel's identity: $\langle g, h\rangle = \sum_{\alpha\in\mathbb{F}_q^n} \hat{g}(\alpha)\overline{\hat{h}(\alpha)}$

Improved Analysis for Finite Fields [3/4]

The Fourier set of $f:\mathbb{F}_q^n \to \mathbb{F}_q$ is $\Phi(f) := \{\varphi_c : \mathbb{F}_q^n \to \mathbb{C},\ \varphi_c(x) := \omega_p^{Tr(c\cdot f(x))}\}_{c\in\mathbb{F}_q^*}$.

Note that $|\Phi(f)| = q-1$.

EXAMPLE: If $\ell:\mathbb{F}_q^n \to \mathbb{F}_q$ is the linear function $\ell(x)=\langle\alpha,x\rangle$ then $\Phi(f) = \{\chi_{c\alpha}\}_{c\in\mathbb{F}_q^*}$.

Note that $\hat{\chi}_{c\alpha}(c\alpha)=1$ and, $\forall\,\beta\in\mathbb{F}_q^n\setminus\{c\alpha\}$, $\hat{\chi}_{c\alpha}(\beta)=0$. ⊛

The distance between two functions can be expressed in terms of their Fourier sets:

<u>lemma:</u> $\forall f,g:\mathbb{F}_q^n\to\mathbb{F}_q$ with Fourier sets $\{\varphi_c\}_{c\in\mathbb{F}_q^*}$ and $\{\gamma_c\}_{c\in\mathbb{F}_q^*}$

$$\Delta(f,g) = \Pr_{x\leftarrow\mathbb{F}_q^n}[f(x)\neq g(x)] = 1-\frac{1}{q}\cdot\left(1+\sum_{c\in\mathbb{F}_q^*}\langle\varphi_c,\gamma_c\rangle\right).$$

Let $f:\mathbb{F}_q^n\to\mathbb{F}_q$ with Fourier set $\{\varphi_c\}_{c\in\mathbb{F}_q^n}$.

The distance to linear functions is

Plancherel's identity & ⊛:
$\langle\varphi_c,\chi_{c\alpha}\rangle = \sum_{\beta\in\mathbb{F}_q^n}\hat{\varphi}_c(\beta)\overline{\hat{\chi}_{c\alpha}(\beta)} = \hat{\varphi}_c(c\alpha)$

$$\Delta(f,LIN) = \min_{\alpha\in\mathbb{F}_q^n}\left\{1-\frac{1}{q}\cdot\left(1+\sum_{c\in\mathbb{F}_q^*}\langle\varphi_c,\chi_{c\alpha}\rangle\right)\right\} = 1-\frac{1}{q}\left(1+\max_{\alpha\in\mathbb{F}_q^n}\sum_{c\in\mathbb{F}_q^*}\langle\varphi_c,\chi_{c\alpha}\rangle\right) = 1-\frac{1}{q}\left(1+\max_{\alpha\in\mathbb{F}_q^n}\sum_{c\in\mathbb{F}_q^*}\hat{\varphi}_c(c\alpha)\right).$$

Then more analysis yields the desired bound:

$$\Pr[V_{BLR}^f=1] = \frac{1}{q}\left(1+\frac{1}{(q-1)^2}\sum_{\alpha\in\mathbb{F}_q^n}\left(\sum_{c\in\mathbb{F}_q^*}\hat{\varphi}_c(c\alpha)\right)^3\right) \leq \frac{1}{q}\left(1+\max_{\alpha\in\mathbb{F}_q^n}\sum_{c\in\mathbb{F}_q^*}\hat{\varphi}_c(c\alpha)\cdot\frac{1}{(q-1)^2}\sum_{\alpha\in\mathbb{F}_q^n}\left(\sum_{c\in\mathbb{F}_q^*}\hat{\varphi}_c(c\alpha)\right)^2\right)$$
$$\leq \frac{1}{q}\left(1+\max_{\alpha\in\mathbb{F}_q^n}\sum_{c\in\mathbb{F}_q^*}\hat{\varphi}_c(c\alpha)\cdot 1\right) = 1-\Delta(f,LIN).$$

↑ most of the work

←
$$\sum_{\alpha\in\mathbb{F}_q^n}\hat{\varphi}_c(c\alpha)\hat{\varphi}_d(d\alpha)$$
$$= \sum_{\alpha\in\mathbb{F}_q^n}\sum_{x,y\in\mathbb{F}_q^n} q^{-2n}\varphi_c(x)\varphi_d(y)\,\omega_p^{-\langle c\alpha,x\rangle-\langle d\alpha,y\rangle}$$
$$= q^{-n}\sum_{x\in\mathbb{F}_q^n}\varphi_c(x)\varphi_d(-cd^{-1}x)$$
$$\leq q^{-n}\sum_{x\in\mathbb{F}_q^n}|\varphi_c(x)\varphi_d(-cd^{-1}x)|$$
$$= q^{-n}\cdot q^n = 1$$

$$\sum_{\alpha\in\mathbb{F}_q^n}\left(\sum_{c\in\mathbb{F}_q^*}\hat{\varphi}_c(c\alpha)\right)^2$$
$$= \sum_{c\in\mathbb{F}_q^*}\sum_{d\in\mathbb{F}_q^*}\left(\sum_{\alpha\in\mathbb{F}_q^n}\hat{\varphi}_c(c\alpha)\hat{\varphi}_d(d\alpha)\right)$$
$$\leq \sum_{c\in\mathbb{F}_q^*}\sum_{d\in\mathbb{F}_q^*}1 = (q-1)^2$$

Improved Analysis for Finite Fields [4/4]

The special case $q=2$ corresponds to linearity testing for boolean functions $f: \mathbb{F}_2^n \to \mathbb{F}_2$.

The analysis of the BLR test simplifies to an elegant and concise computation,

now via Fourier analysis of boolean functions.

Theorem 1.30. *Suppose the BLR Test accepts $f : \mathbb{F}_2^n \to \mathbb{F}_2$ with probability $1-\epsilon$. Then f is ϵ-close to being linear.*

Proof. In order to use the Fourier transform we encode f's output by $\pm 1 \in \mathbb{R}$; thus the acceptance condition of the BLR Test becomes $f(x)f(y) = f(x+y)$. Since

$$\tfrac{1}{2} + \tfrac{1}{2} f(x)f(y)f(x+y) = \begin{cases} 1 & \text{if } f(x)f(y) = f(x+y), \\ 0 & \text{if } f(x)f(y) \neq f(x+y), \end{cases}$$

we conclude

$$\begin{aligned} 1-\epsilon = \mathbf{Pr}[\text{BLR accepts } f] &= \mathop{\mathbf{E}}_{x,y}[\tfrac{1}{2} + \tfrac{1}{2} f(x)f(y)f(x+y)] \\ &= \tfrac{1}{2} + \tfrac{1}{2} \mathop{\mathbf{E}}_{x}[f(x) \cdot \mathop{\mathbf{E}}_{y}[f(y)f(x+y)]] \\ &= \tfrac{1}{2} + \tfrac{1}{2} \mathop{\mathbf{E}}_{x}[f(x) \cdot (f * f)(x)] && \text{(by definition)} \\ &= \tfrac{1}{2} + \tfrac{1}{2} \sum_{S \subseteq [n]} \widehat{f}(S)\widehat{f * f}(S) && \text{(Plancherel)} \\ &= \tfrac{1}{2} + \tfrac{1}{2} \sum_{S \subseteq [n]} \widehat{f}(S)^3 && \text{(Theorem 1.27).} \end{aligned}$$

We rearrange this equality and then continue:

$$\begin{aligned} 1-2\epsilon &= \sum_{S \subseteq [n]} \widehat{f}(S)^3 && (1.10) \\ &\leq \max_{S \subseteq [n]} \{\widehat{f}(S)\} \cdot \sum_{S \subseteq [n]} \widehat{f}(S)^2 \\ &= \max_{S \subseteq [n]} \{\widehat{f}(S)\} && \text{(Parseval).} \end{aligned}$$

But $\widehat{f}(S) = \langle f, \chi_S \rangle = 1 - 2\mathrm{dist}(f, \chi_S)$ (Proposition 1.9). Hence there exists some $S^* \subseteq [n]$ such that $1-2\epsilon \leq 1-2\mathrm{dist}(f, \chi_{S^*})$; i.e., f is ϵ-close to the linear function χ_{S^*}. □

Source: Analysis of Boolean Functions
Ryan O'Donnell, 2014

Bibliography

Linearity testing

- [BLR 1990]: Self-testing/correcting with applications to numerical problems, by Manuel Blum, Michael Luby, Ronitt Rubinfeld.
- [BCHKS 1996]: Linearity testing in characteristic two, by Mihir Bellare, Don Coppersmith, Johan Håstad, Marcos A. Kiwi, Madhu Sudan.
- [GS 2002]: Locally testable codes and PCPs of almost-linear length, by Oded Goldreich, Madhu Sudan.
- [BCLR 2004]: Non-abelian homomorphism testing, and distributions close to their self-convolutions, by Michael Ben Or, Don Coppersmith, Mike Luby, Ronitt Rubinfeld.
- [Goldreich 2016]: Lecture notes on linearity (group homomorphism) testing, by Oded Goldreich.
- [Goldreich 2017]: Introduction to property testing, by Oded Goldreich.
- (▸A comedy of errors), by Ronitt Rubinfeld.
- (▸Two decades of property testing), by Madhu Sudan.

Tighter bound using Fourier analysis