Lecture B.2

Linearity Testing

(Locality of the Hadamard code)

Tom Gur

Summer Graduate School on
Foundations and Frontiers of Probabilistic Proofs
July 27, 2021

Recap

Checking a proof without reading it?

Proofs have many interpretations:

- "logical derivations from axioms" - Zermelo
- "approximation of understanding" - Dinur
- "Whatever that convinces me" - Even


Science Times

Biologists Tell a Tale Of Interfering In-Laws

Asteroid Defense: 'Risk Is Real,' Planners Say

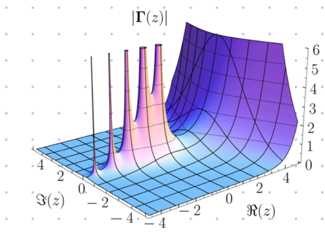
New Short Cut Found For Long Math Proofs

Key idea: check proofs probablistically allowing negligable error

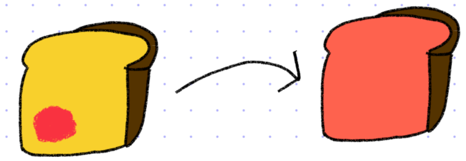
Caveat: What if only one line of the proof is wrong ?

Local-to-global phenomena

Idea: endow proof with a rich structure that allows checking global properties via local constraints!



aka, the "Jam principle"

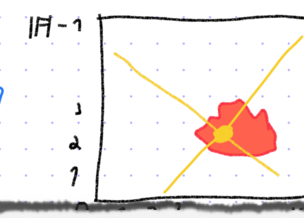


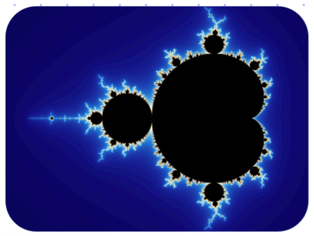
Informal example: low-degree polynomials

$p \in \mathbb{F}[X,Y]$, $\deg(p) \le d$

$p_\ell(t) := p(\bar{x}+t\bar{y}) \in \mathbb{F}[X]$

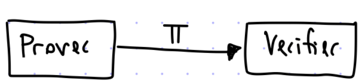
$\deg(p_\ell) \le d$



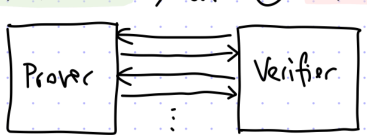


A New Model: Probabilistically Checkable Proofs

- NP represents proofs having a deterministic polynomial-time verifier

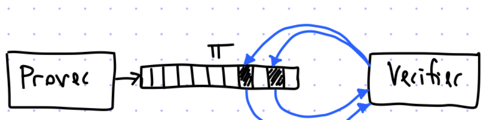


- IP represents proofs where the polynomial-time verifier has two new resources: ① randomness, and ② interaction



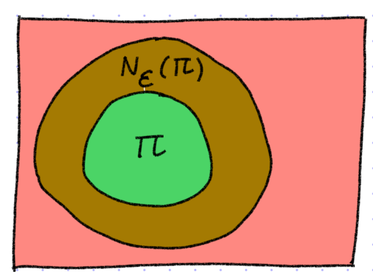
Today we study a new model:

- PCP represents proofs where the polynomial-time verifier has two new resources: ① randomness, and ② oracle access to proof



Conceptual perspectives

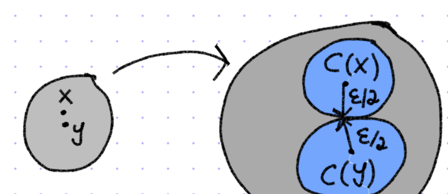
Property testing



Is $f \in \Pi$ (e.g. $\Pi = \mathbb{F}^{\le d}[X]$)

or $\delta(f,\Pi) > \varepsilon$?

Coding theory



For every $x \neq y$, we have

$\delta(C(x), C(y)) > \varepsilon$

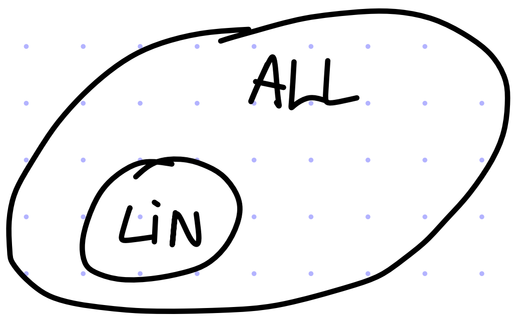
e.g., univariate polynomials

low-degree polynomials

Linear func. on hypercube

Linearity Testing

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



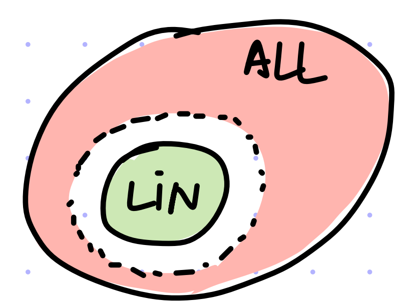
$ALL = \{ f: \mathbb{F}^n \to \mathbb{F} \}$ $\qquad |ALL| = |\mathbb{F}|^{|\mathbb{F}|^n}$

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

We want a $O(1)$-query test that, given $f \in ALL$, says YES if $f \in LIN$ and NO if $f \notin LIN$.

But this is impossible: if f differs in 1 location from $\bar{f} \in LIN$ then $f \notin LIN$
but we cannot detect this with constant soundness error.

So we relax the question: given oracle access to $f \in ALL$,
say YES if $f \in LIN$ and NO if f is far from LIN



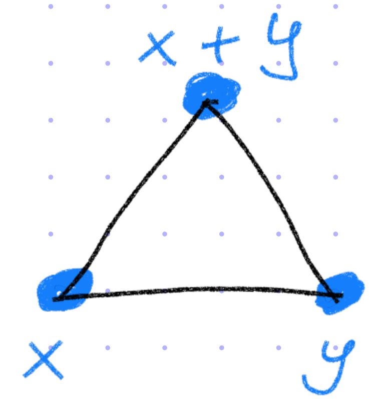
an instance of a problem in Property Testing

We count in Hamming distance:

$$\Delta(f,g) := \Pr_{x \in \mathbb{F}^n}[f(x) \neq g(x)] \quad \text{and} \quad \Delta(f,S) := \min_{g \in S} \Delta(f,g).$$

Q1: can we solve the relaxed problem?

A simple yet important idea: duality

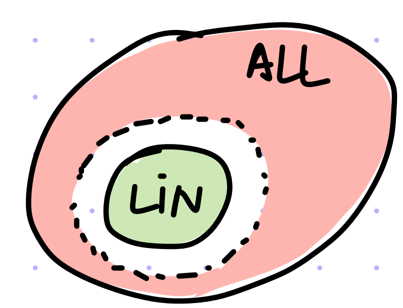


A function $f: \mathbb{F}^n \to \mathbb{F}$ is linear if $\exists\, c \in \mathbb{F}^n$ s.t. $f(x) = \sum_{i=1}^{n} c_i x_i$.
Equivalently, if $\forall x, y \in \mathbb{F}^n$ $f(x) + f(y) = f(x+y)$.

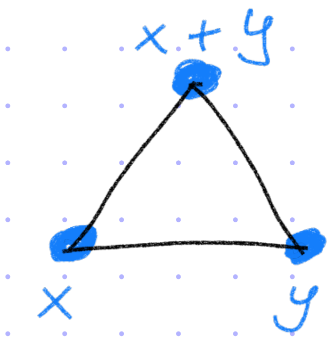
$\text{ALL} = \{ f: \mathbb{F}^n \to \mathbb{F} \}$

$\text{LiN} = \{ f: \mathbb{F}^n \to \mathbb{F} \text{ is linear} \}$

So we relax the question: given oracle access to $f \in \text{ALL}$, say YES if $f \in \text{LiN}$ and NO if f is far from LiN



The Blum-Luby-Rubinfeld Test



A $O(1)$-query test for linearity testing:

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

randomness: $2n$ field elts
queries: 3 locations of f

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

Soundness: non-trivial.

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

Proof intuition:

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

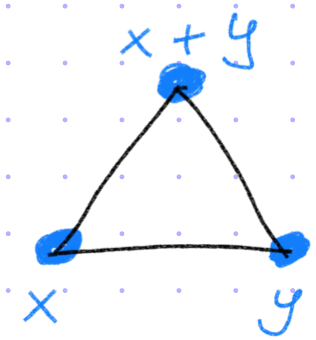
$g_f : \mathbb{F}^n \to \mathbb{F}$ is defined as $g_f(x) := \underbrace{\arg\max_{v \in \mathbb{F}} \left| \{ y \in \mathbb{F}^n \mid v = f(x+y) - f(y) \} \right|}_{\text{this is the plurality value}}$

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

$g_f : \mathbb{F}^n \to \mathbb{F}$ is defined as $\underbrace{g_f(x) := \arg\max_{v\in\mathbb{F}} \left|\left\{ y\in\mathbb{F}^n \mid v = f(x+y)-f(y) \right\}\right|}_{\text{this is the plurality value}}$



Step 1: ※Bad-triangle is captured by distance from plurality vote

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

Step 2: plurality implies overwhelming majority

$$\Pr_{y\in\mathbb{F}^n}\left[g_f(x)=f(x+y)-f(y)\right] \geq 1-2\cdot\Pr\left[V_{BLR}^{f}=0\right] \geq \frac{2}{3}$$

Step 3: plurality vote (g_f) is a linear function

if $\Pr\left[V_{BLR}^{f}=0\right]<\frac{1}{6}$ then $\forall x,y \;\; g_f(x)+g_f(y)=g_f(x+y)$

Analysis of BLR Test - Part 1

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

If g_f is far from f then V_{BLR}^f must reject with high probability:

claim: $\Pr[V_{BLR}^f = 0] \geq \frac{1}{2} \cdot \Delta(g_f, f)$

proof: Letting $S = \{ x \in \mathbb{F}^n \text{ s.t. } \Pr_{y \leftarrow \mathbb{F}^n}[f(x) \neq f(x+y) - f(y)] \geq 1/2 \}$, we get

$$\Pr[V_{BLR}^f = 0] = \Pr_x[x \in S] \Pr_{x,y}[V_{BLR}^f = 0 \mid x \in S] + \Pr_x[x \notin S] \Pr_{x,y}[V_{BLR}^f = 0 \mid x \notin S]$$

$$\geq \frac{|S|}{|\mathbb{F}|^n} \cdot \min_{x \in S} \left\{ \Pr_y[f(x) \neq f(x+y) - f(y)] \right\} + 0 \geq \frac{|S|}{|\mathbb{F}|^n} \cdot \frac{1}{2} .$$

Also, for every $x \notin S$ we have $\Pr_{y \leftarrow \mathbb{F}^n}[f(x) = f(x+y) - f(y)] > \frac{1}{2}$ so $f(x) = g_f(x)$.
This tells us that $\frac{|S|}{|\mathbb{F}|^n} \geq \Delta(g_f, f)$. ∎

Analysis of BLR Test - Part 2

$$\Pr_{y\leftarrow\mathbb{F}^n}\left[g_f(x)=f(x+y)-f(y)\right] \geq \max_{v\in\mathbb{F}} 2 \Pr_{y\leftarrow\mathbb{F}^n}\left[\left[V^f_{BLR}\ f(x+y) \tfrac{2}{3} f(y)\right]\right]$$

$$\geq \sum_{v\in\mathbb{F}} \Pr_{y\leftarrow\mathbb{F}^n}\left[v=f(x+y)-f(y)\right]^2$$

$\sum_i p_i^2 \leq \max_i \{p_i\} \cdot \sum_i p_i$

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

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

Next we analyze the collision probability:

<u>claim</u>: $\forall x\in\mathbb{F}^n$, $\Pr_{y,z}\left[f(x+y)-f(y)=f(x+z)-f(z)\right] \geq 1-2\cdot\Pr\left[V^f_{DLR}=0\right]$

<u>proof</u>: Define $T := \left\{(y,z)\in\mathbb{F}^n\times\mathbb{F}^n \;\middle|\; \begin{array}{l} f(z)=f(y)+f(z-y) \\ f(x+z)=f(x+y)+f(z-y) \end{array}\right\}$.

- if $(y,z)\in T$ then $f(x+y)-f(y)=\left[f(x+y)+f(z-y)\right]-\left[f(z-y)+f(y)\right]=f(x+z)-f(z)$.
- $\Pr_{y,z}\left[(y,z)\notin T\right]\leq 2\cdot\Pr\left[V^f_{BLR}=0\right]$ because $(y,z-y)$ and $(x+y,z-y)$ are random in $\mathbb{F}^2$ ∎

Analysis of BLR Test - Part 3

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

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

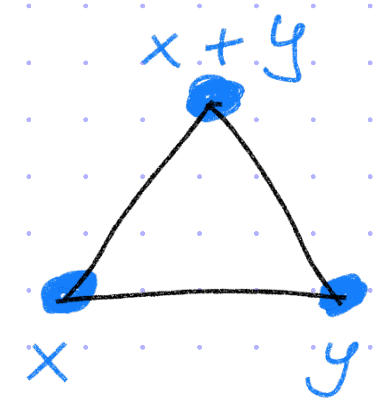
We established that $\Pr[V_{BLR}^f = 0] \geq \frac{1}{2} \cdot \Delta(g_f, f)$ & $\Pr_{y \in \mathbb{F}^n}[g_f(x) = f(x+y) - f(y)] \geq 1 - 2 \cdot \Pr[V_{BLR}^f = 0]$.

$> 2/3$

If $\Pr[V_{BLR}^f = 0] \geq 1/6$ then we are done. So assume that $\Pr[V_{BLR}^f = 0] < 1/6$.

We prove that $g_f \in \mathrm{LIN}$, so we are done as $\Pr[V_{BLR}^f = 0] \geq \frac{1}{2} \Delta(g_f, f) = \frac{1}{2} \cdot \Delta(f, \mathrm{LIN})$.

<u>claim:</u> if $\Pr[V_{BLR}^f = 0] < \frac{1}{6}$ then $\forall x, y \;\; g_f(x) + g_f(y) = g_f(x+y)$



<u>proof:</u>

$$\Pr_z[g_f(x) = f(x+z) - f(z)] \geq 1 - 2 \cdot \Pr[V_{BLR}^f = 0] > \frac{2}{3}$$

$$\Pr_z[g_f(y) = f(y+z) - f(z)] \geq 1 - 2 \cdot \Pr[V_{BLR}^f = 0] > \frac{2}{3}$$

$z \leftrightarrow z-y$

$$\Pr_z[g_f(y) = f(z) - f(z-y)]$$

$$\Pr_z[g_f(x+y) = f(x+y+z) - f(z)] \geq 1 - 2 \Pr[V_{BLR}^f = 0] > \frac{2}{3}$$

$z \leftrightarrow z-y$

$$\Pr_z[g_f(x+y) = f(x+z) - f(z-y)]$$

$\exists\, z^*$ s.t.

$g_f(x) = f(x+z^*) - f(z^*)$

$g_f(y) = f(z^*) - f(z^* - y)$

$g_f(x+y) = f(x+z^*) - f(z^* - y)$

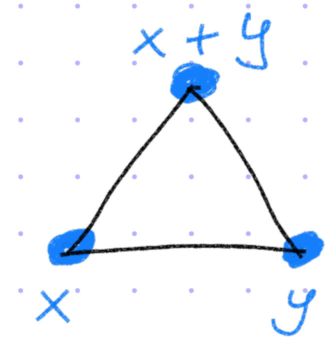
$\Rightarrow g_f$ linear at $(x, y) \in \mathbb{F}^n$!

Digest

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

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

$g_f : \mathbb{F}^n \to \mathbb{F}$ is defined as $g_f(x) := \underbrace{\arg\max_{v \in \mathbb{F}} \left|\left\{ y \in \mathbb{F}^n \mid v = f(x+y) - f(y) \right\}\right|}_{\text{this is the plurality value}}$



Step 1: ※Bad-triangle is captured by distance from plurality vote

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

Step 2: plurality implies overwhelming majority

$$\Pr_{y \leftarrow \mathbb{F}^n}\left[g_f(x) = f(x+y) - f(y)\right] \geq 1 - 2\cdot\Pr\left[V_{BLR}^{f}=0\right] \geq \frac{2}{3}$$

Step 3: plurality vote (g_f) is a linear function

if $\Pr\left[V_{BLR}^{f}=0\right] < \frac{1}{6}$ then $\forall x, y \;\; g_f(x)+g_f(y)=g_f(x+y)$

Discussion

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 $f(\boldsymbol{x})f(\boldsymbol{y}) = f(\boldsymbol{x}+\boldsymbol{y})$. Since

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

we conclude

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

We rearrange this equality and then continue:

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