Low Degree Testing

Madhu Sudan
Harvard University

This talk

- Mostly … historical tour of the low-degree testing problem: Results, motivations, some proof insights

Problem Definition

- Given oracle access to $f: S^m \to \mathbb{F}$ and d is f close to a degree d polynomial? (Usually $S = \mathbb{F} = \mathbb{F}_q$)
- Considerations:
  - Minimize query complexity (#queries to f)!
    - Independent of m?
  - Query structured sets (querying a line/plane/subspace, better than arbitrary queries (Why?))
  - Detect even small correlations between f and degree d polynomials.
  - Reduce randomness?

Why study this question?

- Historically:
  - Mathematical curiosity … natural question!
  - Has applications …
    - To PCPs
    - (un-)Bias amplification
    - Explicit small set expanders …

Brief History

- Phase 1: Blum-Luby-Rubinfeld, Babai-Fortnow-Lund, Babai-Fortnow-Levin-Szegedy.
  - Special cases.
- Phase 2: Rubinfeld-S.
  - General definition/setup
- Phase 3: Arora-Safra, ALMSS, Polishchuk-Spielman Applications$\Leftarrow$ Strengthenings
- Multiple directions:
  - Correlation detection:
  - Randomness reduction:
  - "Moderate degree": ($d > |S|$)

Linearity Testing

- Say $f: \mathbb{F}_2^m \to \mathbb{F}_2$; test for "$d = 1$ + homogeneity".
- The BLR test: Pick $a, b \in \mathbb{F}_2^m$ unif. ind.
  - Accept iff $f(a) + f(b) = f(a+b)$
- Def: $\mathrm{Rej}(f) \coloneqq \Pr_{a,b}[f(a) + f(b) \neq f(a+b)]$
- Clearly: f linear $\Leftrightarrow \mathrm{Rej}(f) = 0$
- Closeness: $\delta(f,g) \coloneqq \Pr_{a}[f(a) \neq g(a)]$

  $$\delta(f) \coloneqq \min_{\{g \text{ linear}\}} \{\delta(f,g)\}$$
- BLR Theorem: $\mathrm{Rej}(f) < \frac{2}{9} \Rightarrow \delta(f) \leq 2\mathrm{Rej}(f)$
- [Bellare-Coppersmith-Hastad-Kiwi-S.] $\delta(f) \leq \mathrm{Rej}(f) \leq 3\delta(f)$

Role of the Linearity Test in PCPs

- Note: Growing space of functions (size 2^m); Query complexity $O(1)$
- First glimpse of $O(1)$ query PCPs.
- Leads to (relatively simple) PCPs of exponential size with $O(1)$ queries. (Non-trivial as a MIP)
- Yields poly size PCPs in [Arora-Lund-Motwani-S.-Szegedy]
- [Bellare-Goldreich-S.]: Improved analysis improves PCP query complexity … motivating BCHKS.
- [Hastad]: Tests long codes using noisy BLR-test … leads to optimal query complexity.

Proof Ideas

- Proof 2: Fourier Analysis
  - Viewed properly: linear functions form orthogonal basis of all functions $\mathbb{F}_2^m \to \mathbb{F}_2$
  - $\left\{\widehat{f_g} \coloneqq 1 - 2\delta(f,g)\right\}_{\{g \text{ linear}\}}$: coordinates of f in this basis.
  - Miraculous Identity: $\mathrm{Rej}(f) = \sum_g \widehat{f_g}^{\,3}$
- Proof 1: "Original" BLR proof (due to Coppersmith)

  $\mathrm{Vote}_a^f(r) \coloneqq f(a+r) - f(r); \qquad h(a) \coloneqq \mathrm{Maj}_r\left\{\mathrm{Vote}_a^f(r)\right\}$
  - $\delta(h,f) \leq 2\,\mathrm{Rej}(f)$
  - $\mathrm{Rej}(f) < \frac{2}{9} \Rightarrow h$ linear.
    - Key step: $\forall a \Pr_{r,s}[\mathrm{Vote}_a(r) \neq \mathrm{Vote}_a(s)] \leq 2\mathrm{Rej}(f)$

      $f(a+r) + f(s) \approx f(a+r+s) \approx f(a+s) + f(r)$

Beyond linearity?

- Proof 2?
  - No luck in this direction … orthogonality is very special
  - Nearest attempts to extend:
    - [Kiwi] (other fields)
    - [Kaufman-Litsyn], [Kaufman-S.]: any sparse high dist. linear code … use MacWilliams Identity, Krawtchouk …
    - [Kopparty-Saraf]: Above reduces to linearity test.

Beyond Linearity

- First studied in [Gemmell-Lipton-Rubinfeld-S.-Wigderson]:
- Proof 1 Extends:
  - $f: \mathbb{Z}_p^m \to \mathbb{Z}_p$ is of deg $d \Leftrightarrow \forall a, b\ \sum_{i=0}^{d+1} (-1)^i \binom{d+1}{i} f(a+ib) = 0$
  - Leads to natural test: #Queries = $d+2$
  - Can define $\mathrm{Vote}_a(r)$
  - Can prove the magic identity:

$$\Pr_{r,s}[\mathrm{Vote}_a(r) \neq \mathrm{Vote}_b(s)] \leq 2(d+1)\mathrm{Rej}(f)$$

  - Independent of the number of variables!!
  - Actually …

… thanks to [Sasha Shen]!

Summary of State of Knowledge in '91

- Have a low degree test …
- Analysis OK-ish:
  - $\mathrm{Rej}(f) \geq \frac{\delta(f)}{d^3}$
- No geometry
- No symmetry
- No intuition …

'91-'92: Rubinfeld-S, ALMSS

- Tests rely on the fact that f restricted to affine subspace (line) $A \subseteq \mathbb{F}^m$ does not increase in degree.
- $\delta^t(f) \coloneqq \mathbb{E}_{\{\text{affine } A:\, \dim(A)=t\}}[\delta(f|_A)]$
- Fact: $\delta^t(f) \leq \mathrm{Rej}(f) \leq d \cdot \delta^t(f)$
- In fact $\delta^t(f)$ more important than $\mathrm{Rej}(f)$ …
  - Corresponds to query complexity of q^t but morally $O(1)$
- Question [ALMSS]: Is $\delta(f) = \Theta\big(\delta^{\text{one}}(f)\big)$?
- [RS]: Yes, provided this is true for $m = 2^*$
- [Arora-Safra]: It is true for $m = 2^*$!
- Thm [ALMSS]: $d = o(q^{1/3}) \Rightarrow \delta(f) = \Theta\big(\delta^{\text{one}}(f)\big)$

Polynomials and PCPs

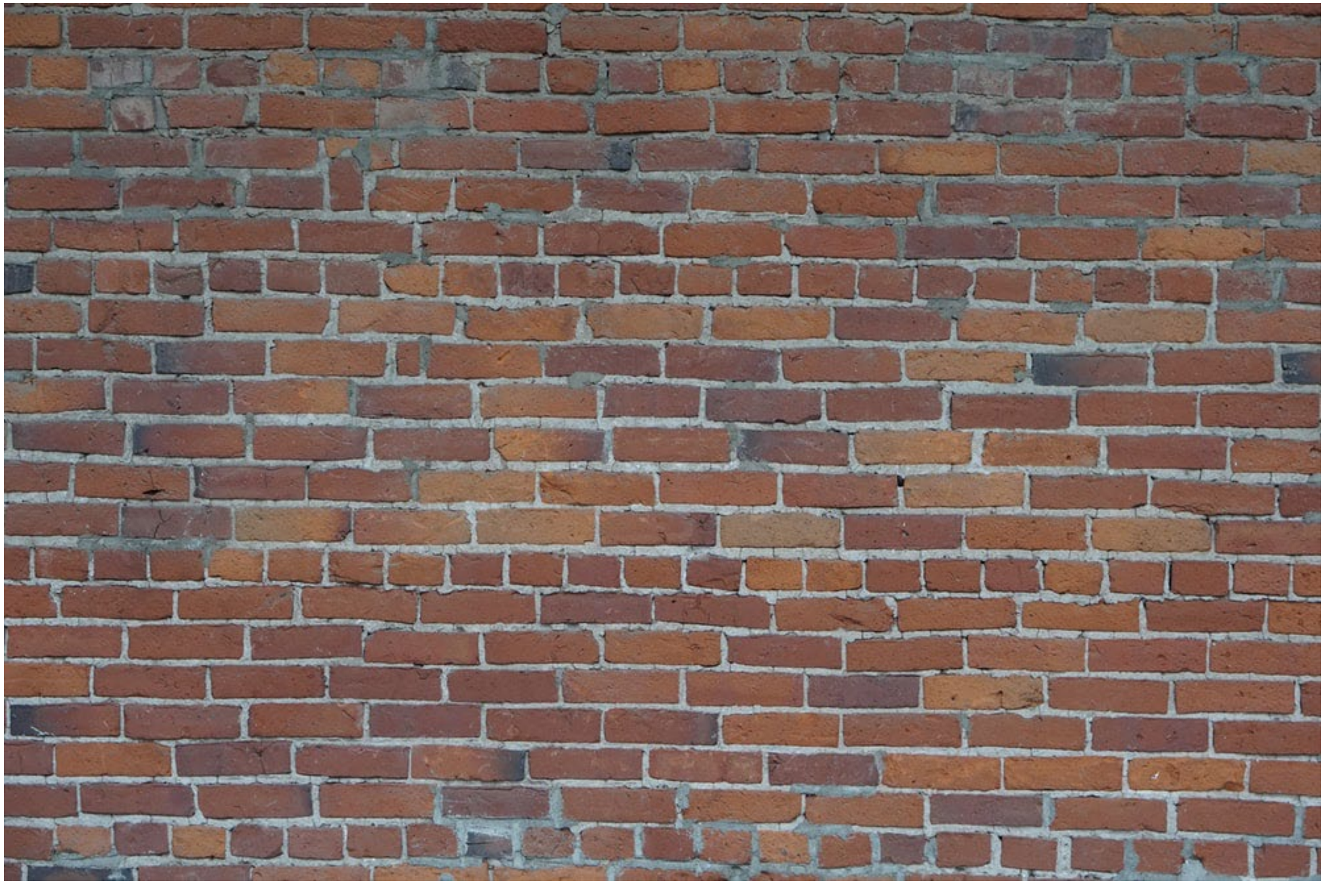
- PCP: Format for proving general statements (e.g., "G is 3-colorable") verifiable by few queries.
- Initial constructions + currently best-known constructions: Depend on polynomials and low-degree testing.
- Why polynomials?
  1. Polynomials are error-correcting codes!
  2. Polynomial are expressive

Polynomials = Walls

- Data/Proof = zillions of bits … each bit acting independently

Polynomials = walls

- Proof = zillions of bits … each bit acting independently
- Polynomial = glue that binds them together.



An Analogy

- Inspecting a building:
  - "Building = $O(n)$ atoms" … OR
  - "Building = $O(1)$ rooms = $O(1)$ walls"
- Former view:
  - Verifying stability takes $\Omega(n)$-checks.
- Latter view:
  - Verifying stability takes $O(1)$-checks +
  - $O(1)$-"stability of wall-checks".

- Polynomials ≡ Walls!

Polynomials = Walls?

- A (NP-)complete statement:
  - Graph $G \in \{0,1\}^{n \times n}$ is 3-colorable.
  - Proof: Coloring ($\Theta(n)$-bits).
  - Verification: Read entire coloring.
- Equivalent (NP-)complete statement:
  - Given: Φ local map from poly's to poly's
  - $\exists$ poly's A, B, C, D s.t. $\Phi(A, B, C, D) \equiv 0$
  - Verification:
    - Step 1: Test A, B, C, D are polynomials
    - Step 2: Verify $\Phi(A, B, C, D)[r] = 0$ for random r.

Polynomials = Wall - II

- Reduction from 3-coloring to polynomial satisfiability [Ben-Sasson-S.'04]
- $\Phi(A,B,C,D)[x_0,\boldsymbol{x},\boldsymbol{y}] = \Phi_E(A,B,C,D)[x_0,\boldsymbol{x},\boldsymbol{y}]$

$$= (A[\boldsymbol{x}](A[\boldsymbol{x}]-1)(A[\boldsymbol{x}]-2)\ - B[\boldsymbol{x}]\Pi_{v\in V}(\boldsymbol{x}-v))$$

$$+\, x_0 \cdot (E(\boldsymbol{x},\boldsymbol{y}) \cdot \Pi_{i\in\{-2,-1,1,2\}}(A[\boldsymbol{x}]-A[\boldsymbol{y}]-i)$$

$$-\, C[\boldsymbol{x},\boldsymbol{y}]\Pi_{v\in V}(\boldsymbol{x}-v) - D(\boldsymbol{x},\boldsymbol{y})\Pi_{v\in V}(\boldsymbol{y}-v))$$

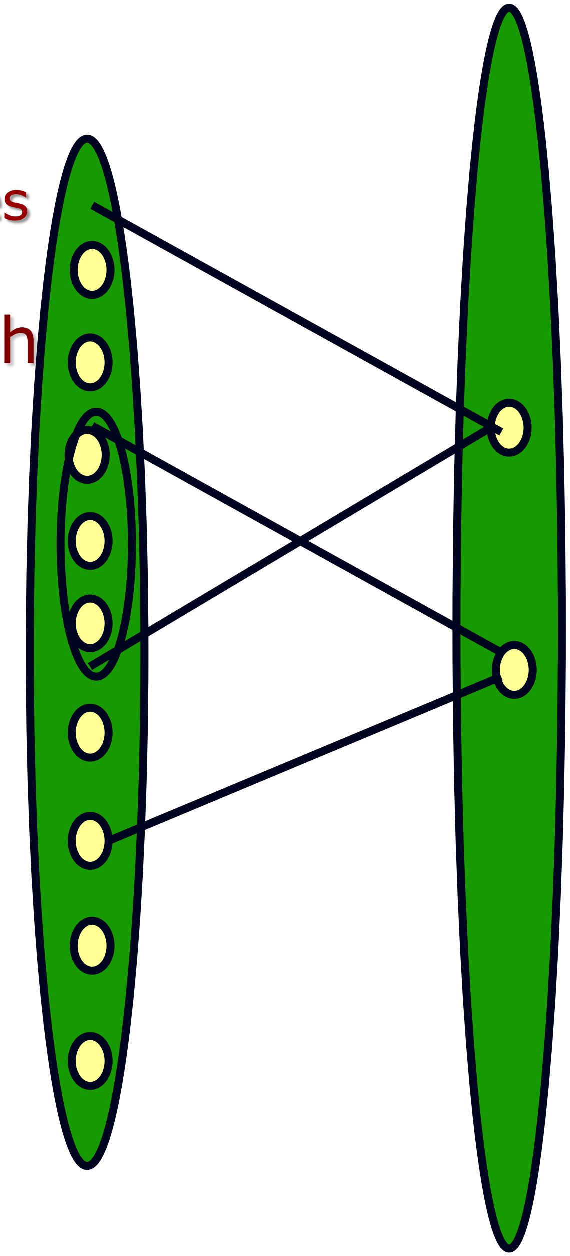
Finer questions: Degree vs. Field Size

- If $d < q$, then distance of code = $1 - \frac{d}{q}$ (want $\Omega(1)$)
- Message size $k \approx \left(\frac{d}{m}\right)^m$; Codeword size $n = q^m$
- Want $n = k^{1+o(1)}$ (implies PCP blowup $n^{1+o(1)}$)
  $\Leftarrow q = O\left(d^{1+o(1)}\right)$
- ALMSS Thm: Needed $q = \Omega(d^3)$ (inherited from [AS])
- Resolved by Polishchuk-Spielman … $q = O(d)$ suffices. (Used Berlekamp-Welch decoder, Introduced polynomial method? Derivatives/multiplicity?)
- Aside: Proofs still need $q > 2d$. Necessary?

Small Correlations

- Motivation: Get PCPs with 2 queries, $\omega(1)$ answer size and $o(1)$ error.
- Need to say something about $\delta(f)$ even if $\delta^t(f) = .999$
  - f is showing $.001$ correlation with low-degree polynomials on t dim subspaces.
  - Is f correlated with some fixed low-degree polynomial everywhere?
- Hope:
  - If $\delta(f,g) = \epsilon$ then for most A, $\delta(f|_A, g|_A) \approx \epsilon$
  - Problem: What about $\delta(f|_A, g^A)$?
- [Raz-Safra]: $d = o(q) \quad \Rightarrow \quad \delta(f) = \big(1 + o(1)\big)\delta^2(q)$
- [Arora-S.]: $d = o\big(q^{1/4}\big) \Rightarrow \quad \delta(f) = \big(1 + o(1)\big)\delta^1(q)$

Key insights [RazSafra]



- Consider structure of testing graph
- Neighboring tests overlap …
  … on entire line
- Restriction to line is a code!!

- Powerful tool: E.g. used to test tensor product of codes [BenSasson-S.]
- Built in to C^3-LTCs of [DinurEvraLivneLubotzkyMozes]

Key insights [Arora-S.]

- Two steps:
  - General reduction from m dim. to 3-dim.
  - 3-dim. extension of Polishchuk-Spielman replacing [Berlekamp-Welch] with list-decoder.

Reducing Randomness

- Randomness of native low-deg. test:
  - Picks two random points of code (line).
  - Code-length = n ⇒ Randomness $\geq 2 \log n$ ⇒ PCP blowup = n^2
- Polishchuk-Spielman: Use axis parallel lines
  - m variables ⇒ $O(m)$ queries, PCP size $O\left(n^{1+1/m}\right)$
  - Inherent?
- Goldreich-S.: Use $O(n)$ random lines … works, but what does it mean?
- BenSasson-S.-Vadhan-Wigderson: Use $\tilde{O}(n)$ lines in ϵ-biased directions.
  - ⇒ PCPs of length $O\left(n^{1+o(1)}\right)$ [BenSassonGoldreichHarshaS.Vadhan]
  - Ultimate Result: [Moshkovitz-Raz] Use lines in "subfield" directions

$d \geq q$

- Summary: '91 – first LDT – no motivation.
- '91-2000 … many improvements all motivated by PCPs.
  - $\delta(f) = \Theta(\delta^1(f))$ provided $q = \Omega(d)$
  - $\delta(f) = \big(1 + o(1)\big) \cdot \delta^2(f)$ provided $q = \Omega(d)$
  - $\delta(f) = \big(1 + o(1)\big) \cdot \delta^1(f)$ provided $q = \Omega(d^4)$
- '2001: [AlonKaufmanKrivelevichLitsynRon] "Return to LDTs without motivation" … $d \to \infty, q = 2$
- (Restrict individual degrees to $\leq q - 1$)
- Inherent query complexity 2^d - indep. of m !
- AKKLR Thm: $q = 4^d$ suffices (rejects f that is $\Omega(1)$ far w.p. $\Omega(1)$)

AKKLR Test

- Pick random $d+1$ dim. subspace A;
- Reject iff $\deg(f|_{\mathrm{A}}) > d$.

- I.e., $\mathrm{Rej}(f) \coloneqq 2^{d+1} \cdot \delta^{d+1}(f)$.

- Why $d+1$: Smallest dim. where some function is not of degree d

- AKKLR Analysis: Extends BLR naturally. (need to extend magical step).

XOR Lemma for Bias of Polynomials

- $\mathrm{Bias}(f) \coloneqq 1 - \left(\frac{q}{q-1}\right)\delta(f, g)$
- XOR question: Let $F(x^1, \ldots, x^n) = f(x^1) + \cdots + f(x^n)$.
  If $\mathrm{Bias}(f) < 1 - \epsilon$ then is $\mathrm{Bias}(F) \leq \exp(-n)$?

- Viola-Wigderson Thm: Yes .. $\mathrm{Bias}(F) \leq \exp\left(-\frac{\epsilon n}{2^d}\right)$
- Key ingredients in proof:
  - Bias is not directly multiplicative
  - But something called Gowers Norm is …
  - Gowers norm nicely relates to AKKLR rejection probability, which relates to distance/bias …

Extending + improving AKKLR

- Extending beyond $q = 2$: [KaufmanRon] + [JutlaPatthakRudraZuckerman]: Query complexity $q^{O\left(\frac{d}{q}\right)}$
  - (magic becomes more complex)
- Improving analysis ($q = 2$):
  - Natural test: $O(2^d)$ queries
  - Analysis weak: Shows only $\delta(f) \leq 2^d \cdot \mathrm{Rej}(f) = 4^d \cdot \delta^{d+1}(f)$
  - Optimal Analysis: [BhattacharyyaKoppartySchoenebeckS.Zuckerman]

    $\delta(f) = O\left(\delta^{d+1}(f)\right)$ unless $\delta^{d+1}(f) = \Omega\left(2^{-d}\right)$
- Improved analysis $q > 2$: [HaramatyShpilkaS.][KaufmanMinzer]

Main Insight

- [BKSSZ, HSS]: If $\delta(f)$ large then on all but $c_{q,d}$ hyperplanes H, $\delta(f|_H)$ large.
- (No explicit local decoder …)
  - Example: $d = 1$:"Linearity test"
  - Inductive Claim: $\delta(f) > .01 \Rightarrow \delta^{10}(f) > \ .001 \ - \frac{5}{2^m}$
  - Helpful claim: $\delta^{10}(f) > 3\delta(f)\big(1 \ - \delta(f)\big)$
  - If $H_1, \ldots H_5$ s.t $\delta\left(f|_{H_i}\right) \leq .01$ then $\delta(f) \leq \frac{1}{8} + .03$
- [KM]: Small set expansion properties of the testing graph. (Many ideas adapted from the proof of 2-2 games conjecture [KhotMinzerSafra])

Some introspection: What made polys testable?

- Form linear error-correcting code; satisfy local constraints.
  - Motivated [Sipser-Spielman]!! But alas insufficient [BenSasson-Harsha-Raskhodnikova]
- Contained in nice tensor-codes.
  - Motivated [BenSasson-S] to study tensor codes
  - plays role in recent C^3-LTCs of [DinurEvraLivneLubotzkyMozes]
- Enjoys Nice symmetries: “Affine-invariance”
  - [KaufmanS.] Seems to explain much of the success (a la BLR, RS, AKKLR, KR, JPRZ).
    - Demystifies “magic” $\Leftarrow$ “Row rank=column rank”
    - Can even extend BKSSZ/HSS [Haramaty-RonZewi-S.]
    - Not necessary. Can test $f: S^n \to \mathbb{F}$ [Bafna-Srinivasan-S.], [Amireddy-Srinivasan-S.]

Summary/Future questions

- Over three decades … have developed a deep understanding of the local-global structure of polynomials

- Some future directions:
  - Explain everything in terms of affine-invariance?
  - Generalize to product property codes
  - LTCs to PCPs … abstract expressivity?

Thank You!