Low degree tests at large distances

Alex Samorodnitsky

Hebrew University

April 7, 2008

Notions in CS

Informal - some suspension of disbelief is asked for

- NP is the class of mathematical statements with easily verifiable (short) proofs.
- CL'71: Reduction to verifying that a given 3-CNF boolean formula is satisfiable.
- A...S'92, D'05 PCP: Reduction to distinguishing between a satisfiable 3-CNF boolean formula, and a significantly unsatisfiable formula - an optimal assignment leaves a positive fraction of terms unsatisfied.
- R' 95 Parallel Repetition: Invalid statement is translated into a very unsatisfiable formula - an optimal assignment leaves a $(1-\epsilon)$-fraction of terms unsatisfied.

Notions in CS - continued

- BGS'95, H'97: A format for proving satisfiability which allows verification by looking at tiny randomized samples from the proof.
- A proof is partitioned into 0-1 strings of length 2^n viewed as functions $f_i : \{0,1\}^n \rightarrow \{0,1\}$.
- In a valid proof all the functions f_i are structured. In any proof of an invalid statement, many of the functions are not structured.

Important building block

$\mathbb{F}$ is a finite field. Given $f : \mathbb{F}^n \to \mathbb{F}$, determine if

- f is a low-degree n-variate polynomial. (structured)
- f is ϵ-far from all low-degree polynomials:

$$Pr_x\{f(x) \neq g(x)\} \geq \epsilon$$

  for any degree-d polynomial g. (not structured)

- Allowed only local tests - may query the function only at a few points.
- May use randomization.

Generalization - Property testing

Given a large combinatorial object G, determine if

- G has a global property P.
- f is ϵ-far from all objects with property P

- Only randomized local queries to G are allowed.
- Ex. Given a graph G on k vertices determine whether G is bi-partite or requires removal of at least ϵk^2 edges to become bi-partite, by querying a small number of edges of G (AK '02).

Specification - extremal polynomiality testing

Main question

Given $f : \mathbb{F}_2^n \to \mathbb{F}_2$, determine if

- f is a low-degree n-variate polynomial.
- f is very far from all low-degree polynomials:

$$Pr_x\{f(x) \neq g(x)\} \geq \frac{1}{2} - \epsilon$$

for any degree-d polynomial g.

- Allowed only local tests - may query the function only at a few points.

Linear polynomials

- Distinguishing between linear and far from linear functions.
- This case is known. Plays an important role in PCP constructions.
- BLR'93, BCHKS'96 - A local test, querying f at 3 points and returning 1 bit, which behaves
  - Deterministically for linear functions
  - Randomly for functions far from linear
- The test makes 3 queries and distinguishes linear and far from linear functions w.p. $1/2$.
- H'97: Can be "lifted" to a PCP construction with same parameters.

Pseudorandomness

- Point of view: Linear functions are structured, functions far from linear are pseudorandom - allowing to extract one random bit.
- In fact, this definition of pseudorandomness for a function f is equivalent to the usual one: f has small Fourier coefficients.
- Need to distinguish between pseudorandomness and structure.

Motivation: stronger linearity tests

- Want to optimize the ratio

$$\rho = \frac{q}{\log_2 1/p}$$

where q is the number of queries and p is the probability the test succeeds.

- For the previous test $\rho = 3/\log(2) = 3$.
- Want to have a test with $\rho = 1 + o_q(1)$.

Motivation: stronger linearity tests

- Want to have a test with $\rho = 1 + o_q(1)$.
- ST'00 - A local test, querying f at q points and returning $q - \sqrt{2q}$ bits, which behaves
  - Deterministically for linear functions
  - Randomly for pseudorandom (far from linear) functions
- The test makes q queries and distinguishes linear and pseudorandom functions w.p. $2^{-q+\sqrt{2q}}$.
- $$\rho = \frac{q}{q - \sqrt{2q}} = 1 + o(1)$$

Motivation: stronger linearity tests

- ST'00 - A local test, querying f at q points and returning $q - \sqrt{2q}$ bits, which behaves
  - Deterministically for linear functions
  - Randomly for pseudorandom (far from linear) functions
- The test makes q queries and distinguishes linear and pseudorandom functions w.p. $2^{-q+\sqrt{2q}}$.
- Lifting to a PCP construction with similar parameters.
- Can we squeeze out even more randomness? How powerful is this notion of pseudorandomness?

Local tests for pseudorandomness

Structure vs. pseudorandomness

Let $f: \{0,1\}^n \to \{0,1\}$ be a boolean function.

- BLR'93, BCHKS'96: Choose $x, y \in \{0,1\}^n$ at random.

Compute

$$f(x) + f(y) + f(x+y)$$

Makes 3 queries, returns 1 useful bit.

- ST'00: Graph tests. Let $G = (V, E)$ be a graph on k vertices.

Choose $x_1 ... x_k \in \{0,1\}^n$ at random. For all $(i,j) \in E$ compute

$$f(x_i) + f(x_j) + f(x_i + x_j)$$

Makes $|E| + |V|$ queries, returns $|E|$ useful bits.

Local tests for pseudorandomness

Structure vs. pseudorandomness

Let $f: \{0,1\}^n \to \{0,1\}$ be a boolean function.

- BLR'93, BCHKS'96: Choose $x, y \in \{0,1\}^n$ at random.

Compute

$$f(x) + f(y) + f(x+y)$$

Makes 3 queries, returns 1 useful bit.

- ST'00: Graph tests. Let $G = (V, E)$ be a graph on k vertices.

Choose $x_1 ... x_k \in \{0,1\}^n$ at random. For all $(i,j) \in E$ compute

$$f(x_i) + f(x_j) + f(x_i + x_j)$$

Makes $|E| + |V|$ queries, returns $|E|$ useful bits.

- If G is the complete graph: makes q queries, returns $q - \sqrt{2q}$ bits.

Even better tests for pseudorandomness

- Let $f: \{0,1\}^n \to \{0,1\}$ be a boolean function.
- ST'00: Hypergraph tests. Let $G = (V, E)$ be a hypergraph on k vertices. Choose $x_1 ... x_k \in \{0,1\}^n$ at random. For all $e = (x_i)_{i \in e} \in E$ compute

$$\sum_{i \in e} f(x_i) + f\left(\sum_{i \in e} x_i\right)$$

Makes $|E| + |V|$ queries, returns $|E|$ bits.

Even better tests for pseudorandomness

Structure vs. pseudorandomness

- Let $f : \{0,1\}^n \to \{0,1\}$ be a boolean function.
- ST'00: Hypergraph tests. Let $G = (V, E)$ be a hypergraph on k vertices. Choose $x_1 ... x_k \in \{0,1\}^n$ at random. For all $e = (x_i)_{i \in e} \in E$ compute

$$\sum_{i \in e} f(x_i) + f\left(\sum_{i \in e} x_i\right)$$

Makes $|E| + |V|$ queries, returns $|E|$ bits.

- If G is the complete hypergraph: makes q queries, returns $q - \log q$ bits.

Even better ?? tests for pseudorandomness

Doesn't work...

- Let $f : \{0,1\}^n \to \{0,1\}$ be a boolean function.
- ST'00: Hypergraph tests. Let $G = (V, E)$ be a hypergraph on k vertices. Choose $x_1 ... x_k \in \{0,1\}^n$ at random. For all $e = (x_i)_{i \in e} \in E$ compute

$$\sum_{i \in e} f(x_i) + f\left(\sum_{i \in e} x_i\right)$$

Makes $|E| + |V|$ queries, returns $|E|$ useless bits.

- If G is the complete hypergraph: makes q queries, returns $q - \log q$ bad bits.

An inconvenient example

Let n be even, and let

$$f(x) = x(1) \cdot x(2) + x(3) \cdot x(4) + \ldots + x(n-1) \cdot x(n)$$

- f is bent (maximally far from all linear functions).
- ST'00: Any hypergraph linearity test with q queries that accepts linear functions accepts f with probability at least $2^{-q+\sqrt{2q}}$.

An inconvenient example

Let n be even, and let

$$f(x) = x(1) \cdot x(2) + x(3) \cdot x(4) + \ldots + x(n-1) \cdot x(n)$$

- f is bent (maximally far from all linear functions).
- ST'00: Any hypergraph linearity test with q queries that accepts linear functions accepts f with probability at least $2^{-q+\sqrt{2q}}$.
- ST'06: Any non-adaptive linearity test with q queries that accepts linear functions accepts f with probability at least $2^{-q+\sqrt{2q}}$.

An inconvenient example

Let n be even, and let

$$f(x) = x(1) \cdot x(2) + x(3) \cdot x(4) + \ldots + x(n-1) \cdot x(n)$$

- f is bent (maximally far from all linear functions).
- ST'00: Any hypergraph linearity test with q queries that accepts linear functions accepts f with probability at least $2^{-q+\sqrt{2q}}$.
- ST'06: Any non-adaptive linearity test with q queries that accepts linear functions accepts f with probability at least $2^{-q+\sqrt{2q}}$.
- +BHR'03, L'07: Any linearity test with q queries that accepts linear functions accepts f with probability at least $2^{-q+\sqrt{2q}}$.

An inconvenient example

Let n be even, and let

$$f(x) = x(1) \cdot x(2) + x(3) \cdot x(4) + \ldots + x(n-1) \cdot x(n)$$

- f is bent (maximally far from all linear functions).
- +BHR'03, L'07: Any linearity test with q queries that accepts linear functions accepts f with probability at least $2^{-q+\sqrt{2q}}$.
- What's going on? The function f must have a hidden structure.

Property testing

Low degree polynomials

Let $\mathbb{F}$ be a finite field. Given $f : \mathbb{F}^n \to \mathbb{F}$, determine if

- f is a polynomial of (low) degree at most d.
- f is ϵ-far from all degree-d polynomials.

- Usually the field is large.

BFL'91: If $|\mathbb{F}| > d + 1$ - restrict f to a random line and check it's a degree-d univariate polynomial.

- Always accepts degree-d polynomials.
- If f is ϵ-far from degree-d polynomials, rejects after $T(\epsilon, d)$ random restrictions.
- Self-correction aka a decoding algorithm for generalized Reed Muller codes

Property testing

Low degree polynomials

Let $\mathbb{F}$ be a finite field. Given $f : \mathbb{F}^n \to \mathbb{F}$, determine if

- f is a polynomial of (low) degree at most d.
- f is ϵ-far from all degree-d polynomials.

- Usually the field is large.

BFL'91: If $|\mathbb{F}| > d + 1$ - restrict f to a random line and check it's a degree-d univariate polynomial.

- Always accepts degree-d polynomials.
- If f is ϵ-far from degree-d polynomials, rejects after $T(\epsilon, d)$ random restrictions.
- Self-correction aka a decoding algorithm for generalized Reed Muller codes

- AKKLR'03: What if the field is small, $\mathbb{F} = \mathbb{F}_2$?

Let $\mathbb{F}$ be a finite field. Given $f: \mathbb{F}^n \to \mathbb{F}$, determine if

- f is a polynomial of (low) degree at most d.
- f is ϵ-far from all degree-d polynomials.

- AKKLR'03: What if the field is small, $\mathbb{F} = \mathbb{F}_2$?

Restrict f to a random $(d+1)$-dimensional affine subspace and check it's a degree-d polynomial.

- Always accepts degree-d polynomials.
- If f is ϵ-far from degree-d polynomials, rejects after $T(\epsilon, d)$ random restrictions.
- Self-correction aka a decoding algorithm for Reed Muller codes

Low-degree testing over $\mathbb{F}_2$

Following AKKLR'03

- To test if f is a degree-d polynomial, compute a random $(d+1)$-st directional derivative of f.
- If f is degree-d this derivative is always zero.
- If it's zero with high probability, then f is close to a degree-d polynomial.

Back to the obstructing function

- The value of the k-th directional derivative of AKKLR is one of the bits computed by a hypergraph test with hyperedges of size k.

Back to the obstructing function

- The value of the k-th directional derivative of AKKLR is one of the bits computed by a hypergraph test with hyperedges of size k.
- Hypergraph tests compute higher derivatives of a function in many of the bits they return.

Back to the obstructing function

- The value of the k-th directional derivative of AKKLR is one of the bits computed by a hypergraph test with hyperedges of size k.
- Hypergraph tests compute higher derivatives of a function in many of the bits they return.
- The function

$$f(x) = x(1) \cdot x(2) + x(3) \cdot x(4) + \ldots + x(n-1) \cdot x(n)$$

is a quadratic polynomial

- f is pseudorandom for linearity tests but structured for higher-degree tests

Back to the obstructing function

- The value of the k-th directional derivative of AKKLR is one of the bits computed by a hypergraph test with hyperedges of size k.
- Hypergraph tests compute higher derivatives of a function in many of the bits they return.
- The function

$$f(x) = x(1) \cdot x(2) + x(3) \cdot x(4) + \ldots + x(n-1) \cdot x(n)$$

is a quadratic polynomial

- f is pseudorandom for linearity tests but structured for higher-degree tests
- Want stronger notion of pseudorandomness

Pseudorandomness I: Balanced derivatives

A technical notion

- A function is *d-pseudorandom* if the probability that its random $(d+1)$-st derivative is zero is very close to $1/2$.
- An analytic pseudorandomness measure for a boolean function f:

d-pseudorandomness of f :

$$(2P(f)-1)^{1/2^d}$$

where $P(f)$ is the probability that f restricted to a random $(d+1)$-dimensional affine subspace of the cube is a degree-d polynomial.

Pseudorandomness I: Balanced derivatives

A technical notion

- A function is d-pseudorandom if the probability that its random $(d+1)$-st derivative is zero is very close to $1/2$.
- An analytic pseudorandomness measure for a boolean function f:

  d-pseudorandomness of f :

  $$(2P(f) - 1)^{1/2^d}$$

  where $P(f)$ is the probability that f restricted to a random $(d+1)$-dimensional affine subspace of the cube is a degree-d polynomial.
- The $1/2^d$-root is to deal with various notions of derivatives.

Pseudorandomness I: Gowers Uniformity

A technical notion

- Defined in G'01 - for functions on $\mathbb{Z}_n$.
- An analytic pseudorandomness measure for a boolean function f:

Gowers uniformity of degree d of f:

$$\left(\mathbb{E}_{x,y_1,\ldots,y_d}(-1)^{\sum_{S\subseteq[d]} f\left(x+\sum_{i\in S} y_i\right)}\right)^{1/2^d}$$

- A function is pseudorandom if its Gowers uniformity is small.

A stronger linearity test given low Gowers uniformity

- S'05, ST'06 - A local test, querying q bits and returning $q - q^{1/d}$ bits, which behaves
  - Deterministically for linear functions
  - Randomly for pseudorandom (low degree-d Gowers uniformity) functions.
- The test makes q queries and distinguishes linear and pseudorandom functions w.p. $2^{-q+q^{1/d}}$.
- ST'06 Conditional lifting to a PCP construction with similar parameters.

Pseudorandomness II - Polynomial pseudorandomness

Structure

Definition: A function is degree-d pseudorandom if it is far from degree-d polynomials.

- The "right" notion we seem to be looking for.
- Additional dividends: explicit degree-d pseudorandom functions for large d lead to interesting lower bounds and pseudorandom generator constructions.

Pseudorandomness II - Polynomial pseudorandomness

Structure

Definition: A function is degree-d pseudorandom if it is far from degree-d polynomials.

- The "right" notion we seem to be looking for.
- Additional dividends: explicit degree-d pseudorandom functions for large d lead to interesting lower bounds and pseudorandom generator constructions.
- Can we compare the two notions of pseudorandomness?
- G'01, GT'05: Low Gowers Uniformity of degree d implies polynomial degree-d pseudorandomness.
- The other direction?

Lack of pseudorandomness should imply structure

Inverse claims

- What if a function f has a non-negligible Gowers Uniformity?

$$\|f\|_{U_d} > \epsilon$$

- $d = 2$: In this case f is $1/2 - \epsilon$ close to a linear function BLR'93, BCHKS'96.
- ϵ is BIG, $\epsilon = 1 - \delta$. In this case f is δ'-close to a degree-$(d-1)$ polynomial AKKLR'03.
- $d = 3$: In this case f is $1/2 - \epsilon'$ close to a degree-2 polynomial GT'05, S'05.
- Any d: f has a variable whose influence is at least $\epsilon'/2^d$ ST'06.

An inverse conjecture for Gowers uniformity

Conjecture T'07 (GT'05), S'05: The two notions of pseudorandomness are equivalent: $\|f\|_{U_{d+1}} > \epsilon$ implies f is $1/2 - \epsilon'$ close to a degree-d polynomial.

An inverse conjecture for Gowers uniformity

Conjecture T'07 (GT'05), S'05: The two notions of pseudorandomness are equivalent: $\|f\|_{U_{d+1}} > \epsilon$ implies f is $1/2 - \epsilon'$ close to a degree-d polynomial.

- Discussion
  - If this conjecture were true, this would give a concise description of Gowers uniformity.
  - It is "equivalent" to low-degree testing at large distances.

An inverse conjecture for Gowers uniformity

Conjecture T'07 (GT'05), S'05: The two notions of pseudorandomness are equivalent: $\|f\|_{U_{d+1}} > \epsilon$ implies f is $1/2 - \epsilon'$ close to a degree-d polynomial.

- BV'07: A weaker conjecture, useful for constructing pseudorandom generators: May also assume f is a polynomial of degree $d + 1$.

The conjecture is false

- GT'07, LMS'07: The conjecture is false, even for $d = 4$ and for f a polynomial of degree 4.
- GT'07: Partial positive results for larger fields.

The conjecture is false

- GT'07, LMS'07: The conjecture is false, even for $d = 4$ and for f a polynomial of degree 4.
- GT'07: Partial positive results for larger fields.
- Counterexample: $f = S_4$ is a symmetric polynomial of degree 4.

$$f(x) = \sum_{|S|=4} \prod_{i \in S} x(i)$$

- $\|f\|_{U_4} > 0.9$
- f is $\left(\frac{1}{2} - \exp\{-cn\}\right)$-far from cubic polynomials.

The conjecture is false

- GT'07, LMS'07: The conjecture is false, even for $d = 4$ and for f a polynomial of degree 4.
- GT'07: Partial positive results for larger fields.
- Question. Assume a big family of degree-4 derivatives of f are non-negligibly imbalanced. Does this imply f is somewhat close to a cubic polynomial?
- BL'08: There is a version of a degree-4 derivative which is negligible for S_4.

Some details

S_4 has large 4-uniformity.

- The directional derivative of S_4 in directions y_1, y_2, y_3, y_4 is:

$$\sum_{|S|=4} Det_S\,(y_1 \cdots y_4) = \sum_{|S|=4} Det_S^2\,(y_1 \cdots y_4) = Det\,(\langle y_i, y_j \rangle)$$

- The behavior of a random 4×4 matrix $(\langle y_i, y_j \rangle)$ is not hard to analyze.

Some details

S_4 is far from cubics.

• A correlation between functions is upperbounded by average correlation between their derivatives.

$$\langle f, g \rangle^8 \leq \mathbb{E}_{y,z} \langle f_{y,z}, g_{y,z} \rangle^2$$

• Let $f = S_4$, g a cubic. Second derivative of f is quadratic, depending on y, z. Second derivative of g is linear.

• By Dixon's theorem know the Fourier spectrum of quadratic polynomials. Need multilinear algebra to wrap this together.