Bounded independence plus noise fools products

Chin Ho Lee

Northeastern University

Elad Haramaty

Harvard University

Emanuele Viola

Northeastern University

Outline

1. **Bounded independence, noise, product tests**
2. Main Result
3. Complexity of Decoding
4. Pseudorandom generators
5. Proof Sketch
6. Open questions

Bounded independence

Definition:

A distribution D over $\{0,1\}^m$ is *b-wise independent* if every b bits of D are uniform

- Introduced by [Carter-Wegman77] as hash functions
- Used everywhere in TCS

Bounded independence

Major research direction:

- Understand what tests f are *fooled* by bounded independence
- i.e., $\mathrm{E}[f(D)]$ is close to $\mathrm{E}[f(U)]$

f	
Combinatorial rectangles	[Even-Goldreich-Luby-Nisan-Velickovic98]
Bounded depth circuits	[Bazzi09], [Razborov09], [Braverman10], [Tal14]
Halfspaces	[Diakonikolas-Gopalan-Jaiswal-Servedio-Viola10], [Gopalan-O'Donnell-Wu-Zuckerman10], [Diakonikolas-Kane-Nelson10]

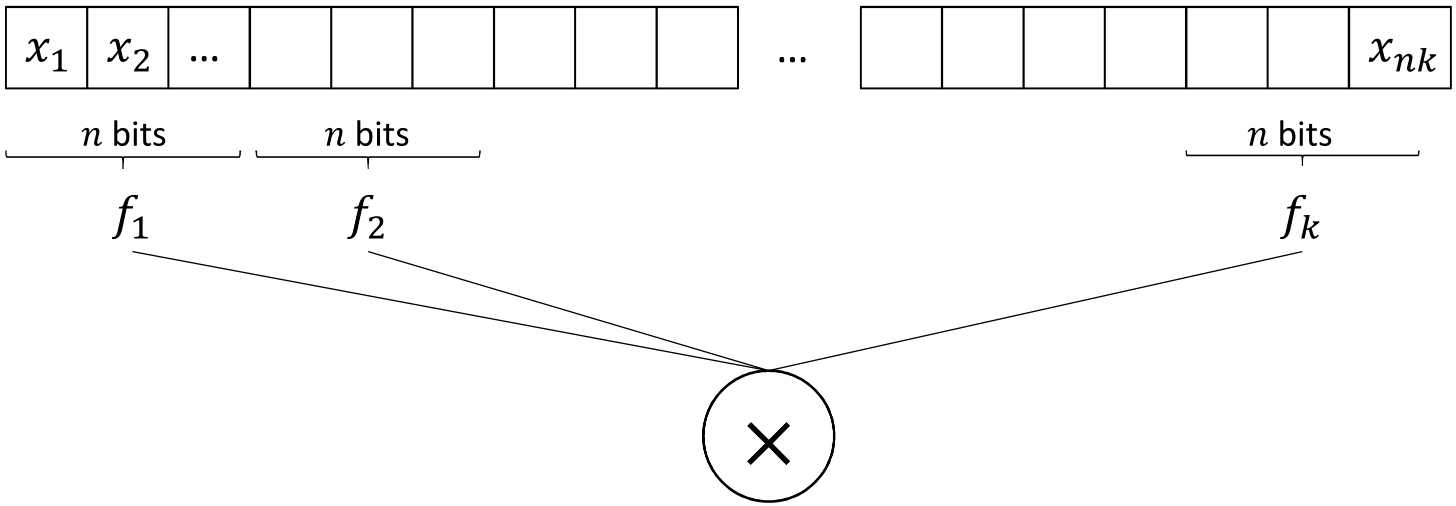
Product tests

<u>Definition:</u>

$F: (\{0,1\}^n)^k \to [-1,1]$ is a *product test if*

$$F(x_1, \dots, x_k) := \prod_i f_i(x_i), \text{ where}$$

$f_1, \dots, f_k \colon \{0,1\}^n \to [-1,1]$ are k arbitrary functions on *disjoint* n bits.



Bounded independence cannot fool product tests

Product test $(m := nk)$
$F: (\{0,1\}^n)^k \to [-1,1]$
$F(x_1, \dots, x_k) := \prod_i f_i(x_i)$

Fact:

$(nk - 1)$-wise independence cannot fool product tests

Proof:

- Parity on nk bits is a product over $\{-1, 1\}$
- Uniform over the same parity is $(nk - 1)$-wise independent

Bounded independence cannot fool product tests

Same example gives error 2^{-k} over product tests over $\{0,1\}$

- So bounded independence cannot fool combinatorial rectangles with error better than 2^{-k}
- Error not good enough for some applications
  - e.g. communication lower bounds
- Too large to sum over 2^{k} rectangles

Small-bias cannot fool product tests

Product test $(m := nk)$
$F : (\{0,1\}^n)^k \to [-1,1]$
$F(x_1, \dots, x_k) := \prod_i f_i(x_i)$

Same issue with small-bias distributions [Naor-Naor]

Fact:

$2^{-\Omega(nk)}$-bias cannot fool product tests

Proof:

- Inner product (IP) on nk bits is a product
- Uniform over IP = 1 is $2^{-\Omega(nk)}$-biased

Our starting observation

All these examples break when few bits of D are perturbed

- one bit of noise fools parity completely

Our main result shows this is a general phenomenon

- Bounded independence plus noise fools product tests with good error bound

Original motivation [L Viola]: sum of small-bias distributions

Outline

1. Bounded independence, noise, product tests
2. **Main Result**
3. Complexity of Decoding
4. Pseudorandom generators
5. Proof Sketch
6. Open questions

Main Result

Product test
$F: (\{0,1\}^n)^k \to [-1,1]$
$F(x_1, \dots, x_k) := \prod_i f_i(x_i)$

Theorem:

Let

- D := n-wise independent on nk symbols
- E := set each symbol to uniform independently with probability η

For any product test F,

$$|\mathrm{E}[F(D+E)] - \mathrm{E}[F(U)]| \leq (1-\eta)^{\Omega\left(\frac{n}{k}\right)}$$

Main Result

Product test
$F: (\{0,1\}^n)^k \to [-1,1]$
$F(x_1, \dots, x_k) := \prod_i f_i(x_i)$

Theorem:
D := n-wise independent on nk symbols
E := set each symbol to uniform independently with probability η

$$|\mathrm{E}[F(D+E)] - \mathrm{E}[F(U)]| \le (1-\eta)^{\Omega\left(\frac{n}{k}\right)}$$

1. Tight when $k = O(1)$
2. Is false for independence $< n$
3. D is not even pairwise independent over blocks
   - Different from previous works
4. Similar result holds when D is $2^{-\Omega(n)}$-almost n-wise independent or $2^{-\Omega(n)}$-biased

Main Result

Product test

$F: (\{0,1\}^n)^k \to [-1,1]$

$F(x_1, \ldots, x_k) := \prod_i f_i(x_i)$

Theorem:

D := n-wise independent on nk symbols

E := set each symbol to uniform independently with probability η

$$|\mathrm{E}[F(D+E)] - \mathrm{E}[F(U)]| \leq (1-\eta)^{\Omega\left(\frac{n}{k}\right)}$$

5. Makes sense for wide range of η
   1. $\eta = c/n, k = O(1)$, error 0.01

      Constant number of noise symbols

   2. $\eta = \Omega(1), k = O(1)$, error $2^{-\Omega(n)}$

      Constant fraction of noise symbols

- Critical for our applications

Noise $\equiv$ Random Restrictions

Can interpret our result as:

On average, a product test becomes simpler under a random restriction [Subbotovskaya61]

- it can be fooled by bounded independence

Differences:

Our results hold for

- *arbitrary* functions
- *arbitrary* η, useful for our applications

Outline

1. Bounded independence, noise, product tests
2. Main Result
3. **Complexity of Decoding**
4. Pseudorandom generators
5. Proof Sketch
6. Open questions

Complexity of decoding

Error-correcting codes

- a fundamental concept in computer science
- many applications in TCS

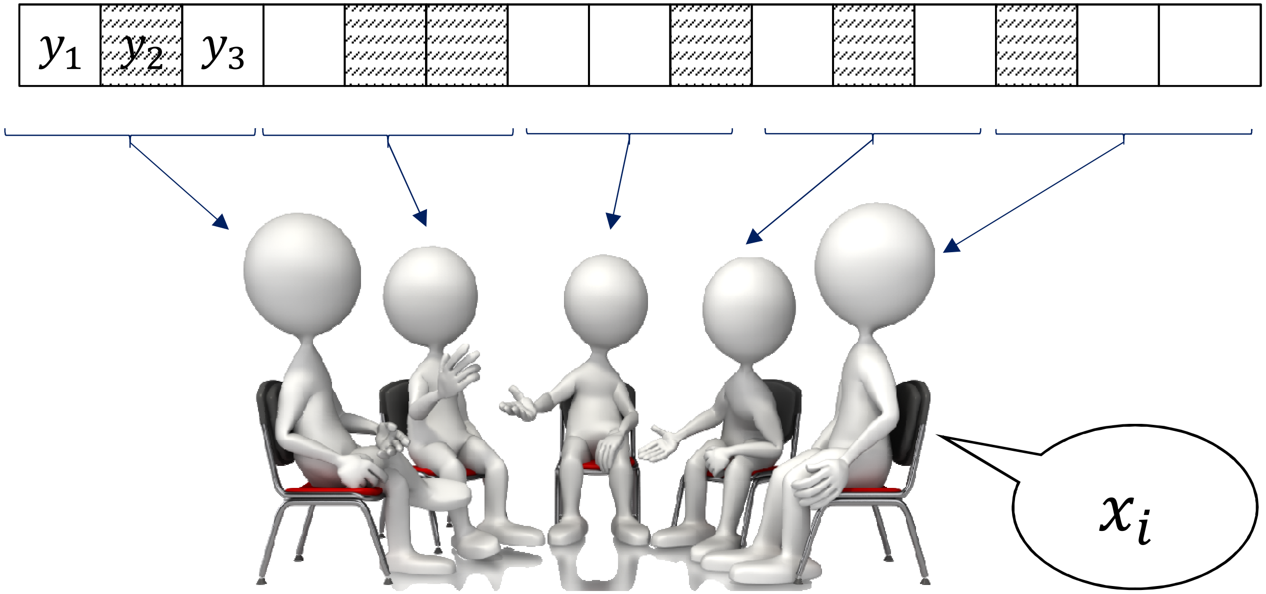
Natural to ask

- What is the complexity of encoding and decoding?
  - [Bar-Yossef—Reingold—Shaltiel—Trevisan02]
  - [Bazzi—Mitter05]
  - [Gronemeier06]

The complexity of decoding 1 symbol

A number-in-hand multiparty communication problem

- Given $y = Enc(x) + noise$ split among $k = O(1)$ parties
- Compute x_i



Our results

This talk: *Code* $:= \left[q, \frac{q}{100}\right]$-Reed—Solomon over F_{q}

- evaluations of degree-$\frac{q}{100}$ polynomials at q positions
- linear rate and linear minimum distance

η = fraction of noise symbols

Theorem:

For most encodings and positions, any $k = O(1)$ parties, $\Omega(\eta q)$ bits of communication is required to decode 1 symbol better than random guessing

- This is essentially tight

Our results

Previous lower bounds	Our lower bounds
Streaming	Communication
For computing the entire message	For computing one symbol of the message
No better for decoding than encoding	Stronger for decoding than encoding

Outline

1. Bounded independence, noise, product tests
2. Main Result
3. Complexity of Decoding
4. **Pseudorandom generators**
5. Proof Sketch
6. Open questions

Pseudorandom generators (PRGs)

Definition:

$G\colon \{0,1\}^{\ell} \to (\{0,1\}^n)^k$ is a pseudorandom generator for test f, if

$$\left|\mathrm{E}\left[f\left(G(U_\ell)\right)\right] - \mathrm{E}[f(U_{nk})]\right| \le 1/3$$

Major line of research: constructing PRGs for one-way space bounded algorithms

- RL vs L
- State of the art [Nisan92, Impagliazzo-Nisan-Wigderson94, Nisan-Zuckerman96]

Pseudorandom generators (PRGs)

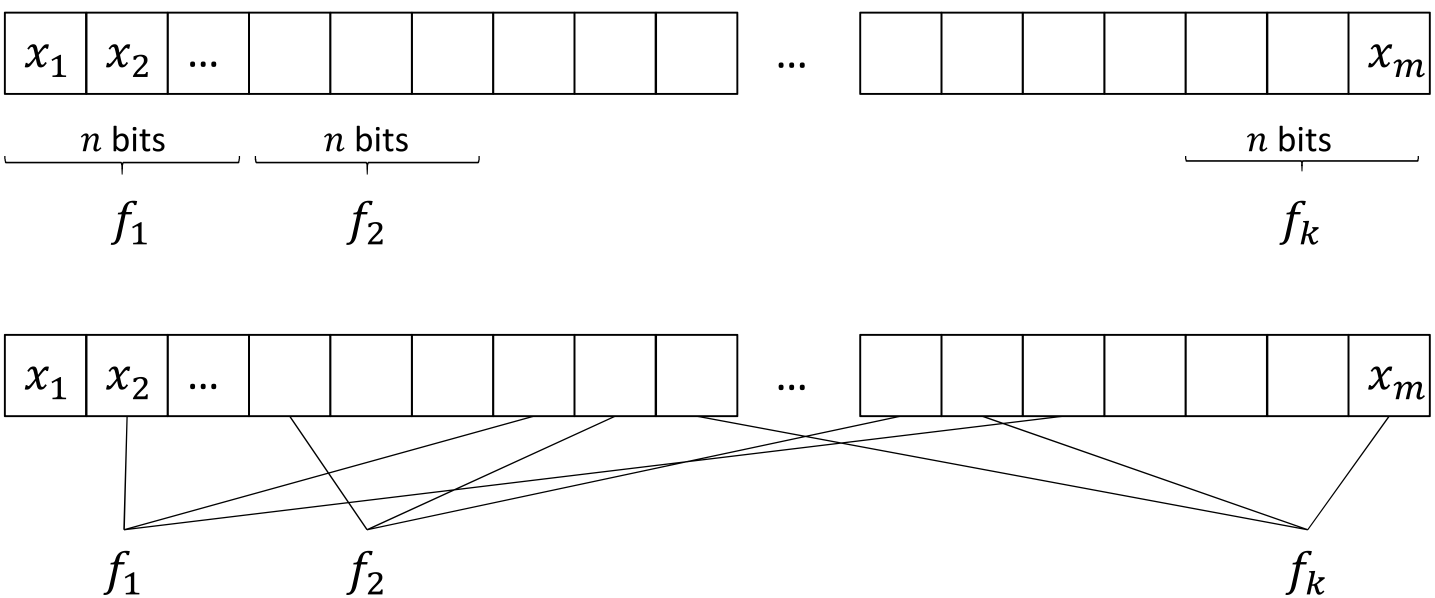
Better PRGs are known on fooling special cases

- Combinatorial rectangles
  - [Even-Goldreich-Luby-Nisan-Velickovic98]
  - [Lu02]
  - [Gopalan-Meka-Reingold-Trevisan-Vadhan12]
- Combinatorial shapes
  - [Gopalan-Meka-Reingold-Zuckerman13]
  - [De15]
- Product tests (aka. Fourier shapes)
  - [Gopalan-Kane-Meka15]

Fixed-order vs any-order products

[Bogdanov-Papakonstantinou-Wan11], [Impagliazzo-Meka-Zuckerman12], [Reingold-Steinke-Vadhan13]

What if input bits are read in *any* order?



Previous results

For k = 2

- [BPW11] gives PRGs with seed length $1.99n$

For larger k

- [Reingold-Steinke-Vadhan13]
- seed length $\tilde{O}(\sqrt{m} \log w)$ for read-once width-w branching programs
- implies seed length $\tilde{O}(n^{3/2} \sqrt{k})$ for rectangles

Our Results

Product test
$F: (\{0,1\}^n)^k \to [-1,1]$
$F(x_1, \dots, x_k) := \prod_i f_i(x_i)$

Theorem

New PRGs for *any-order product tests* with k functions on n bits

- For $k \leq \sqrt{n}$, seed length $2n + \tilde{O}(k^2)$

  Close to optimal when $k = O(1)$

- For $k \geq \sqrt{n}$, seed length $O(n) + \tilde{O}(\sqrt{nk})$

  Improves on [RSV13]'s $\tilde{O}(n^{3/2}\sqrt{k})$ by $O(n)$

For $k = 2$, [BPW11] remains the best known for rectangles

PRGs for other models

Our theorem holds for product tests where each f_i has output in the *complex unit disk* = $\{z \in \mathbb{C}:\ |z| \leq 1\}$

- aka. Fourier shapes in [Gopalan-Kane-Meka15]

[GKM15] shows PRGs for products implies PRGs for

  - generalized halfspaces, combinatorial shapes, ...

We obtain PRGs with seed length $\tilde{O}\left(n\sqrt{k}\right)$ for these models that read bits in *any order*

Bounded Independence plus noise fools space

Our main result also gives a simple PRG for one-way space algorithms

Theorem:

- D: $m^{2/3} \log m$-wise independent on m bits
- E: set each bit to uniform independent with probability 0.01

For any *one-way logspace algorithm* $A: \{0,1\}^m \to \{0,1\}$,

$$|\mathrm{E}[A(D+E)] - \mathrm{E}[A(U)]| \leq o(1)$$

Outline

1. Bounded independence, noise, product tests
2. Main Result
3. Complexity of Decoding
4. Pseudorandom generators
5. **Proof Sketch**
6. Open questions

D := n-wise independent on $3n$ bits
E := set each bit to uniform independently with probability η

Proof Sketch (k = 3)

For any f, g, h: $\{0,1\}^n \to [-1,1]$ on disjoint n bits,

$$|\mathrm{E}[(fgh)(D+E)] - \mathrm{E}[f]E[g]E[h]| \leq 3(1-\eta)^{n/6}$$

Fourier Analysis

1. Noise damps high order Fourier coefficients
2. Independence fools low degree terms

Proof Sketch

D := n-wise independent on $3n$ bits
E := set each bit to uniform independently with probability η

$|\mathrm{E}[(fgh)(D+E)] - \mathrm{E}[f]E[g]E[h]| \leq 3(1-\eta)^{n/6}$

Decompose f into $f(x) = f_L(x) + f_H(x)$, where

- $f_L(x) \coloneqq \sum_{|\alpha| \leq t} \hat{f}_\alpha \chi_\alpha(x)$
- $f_H(x) \coloneqq \sum_{|\alpha| > t} \hat{f}_\alpha \chi_\alpha(x)$
- $t = n/6$

Similarly for g and h

Write $fgh = fgh_H + fgh_L$

$$= fgh_H + fg_H h_L + fg_L h_L$$

$$= fgh_H + fg_H h_L + f_H g_L h_L + f_L g_L h_L$$

Proof Sketch

D := n-wise independent on $3n$ bits
E := set each bit to uniform independently with probability η

$|\mathrm{E}[(fgh)(D+E)] - \mathrm{E}[f]E[g]E[h]| \leq 3(1-\eta)^{n/6}$

$\mathrm{E}[(fgh)(D+E)] - E[f]E[g]E[h]$

$= \mathrm{E}[fgh_H] + \mathrm{E}[fg_H h_L] + \mathrm{E}[f_H g_L h_L] +$
$\mathrm{E}[f_L g_L h_L] - \mathrm{E}[f]\mathrm{E}[g]\mathrm{E}[h]$

- $f_L g_L h_L$ has degree $\leq n$
- $\mathrm{E}[(f_L g_L h_L)(D+E)] - \mathrm{E}[f]\mathrm{E}[g]\mathrm{E}[h] = 0$

- Bound each of $|\mathrm{E}[fgh_H]|, |\mathrm{E}[fg_H h_L]|, |\mathrm{E}[f_H g_L h_L]|$ under $D+E$ by $(1-\eta)^t$

Bounding $|\mathrm{E}[f\, g_H h_L]|$

$f(x) = f_L(x) + f_H(x)$
$f_L(x) := \sum_{|\alpha|\le t} \hat{f}_\alpha \chi_\alpha(x)$
$f_H(x) := \sum_{|\alpha|> t} \hat{f}_\alpha \chi_\alpha(x)$
$t = n/6$

$$\left|\mathrm{E}_{D,E}[f(D_1+E_1) g_H(D_2+E_2) h_L(D_3+E_3)]\right|$$
$$\le \mathrm{E}_D\left[|\mathrm{E}_{E_1}[f(D_1+E_1)]|\left|\mathrm{E}_{E_2}[g_H(D_2+E_2)]\right||\mathrm{E}_{E_3}[h_L(D_3+E_3)]|\right]$$
$$\le \mathrm{E}_D\left[\left|\mathrm{E}_{E_2}[g_H(D_2+E_2)]\right||\mathrm{E}_{E_3}[h_L(D_3+E_3)]|\right]$$

- $\mathrm{E}_{E_2}[g_H(D_2+E_2)]\mathrm{E}_{E_3}[h_L(D_3+E_3)]$ has degree $> n$
- But we can apply Cauchy-Schwarz, and bound instead
  - $\mathrm{E}_U\left[\left|\mathrm{E}_{E_2}[g_H(U+E_2)]\right|^2\right]$ by $(1-\eta)^{2t}$, and
  - $\mathrm{E}_U\left[\left|\mathrm{E}_{E_3}[h_L(U+E_3)]\right|^2\right]$ by 1

- For $k \leq \sqrt{n}$, seed length $2n$ + $\tilde{O}(k^2)$
- For $k \geq \sqrt{n}$, seed length O(n) + $\tilde{O}(\sqrt{nk})$

PRG constructions

For $k \leq \sqrt{n}$,

1. D = $O(2^{-n})$-biased distribution on nk bits
2. E = Set each bit to uniform with prob. η = $\tilde{O}(k/n)$

(1) takes $2n$ + $O(1)$ bits

(2) takes $nkH(\eta)$ = $\tilde{O}(k^2)$ bits to sample E' $\approx$ E

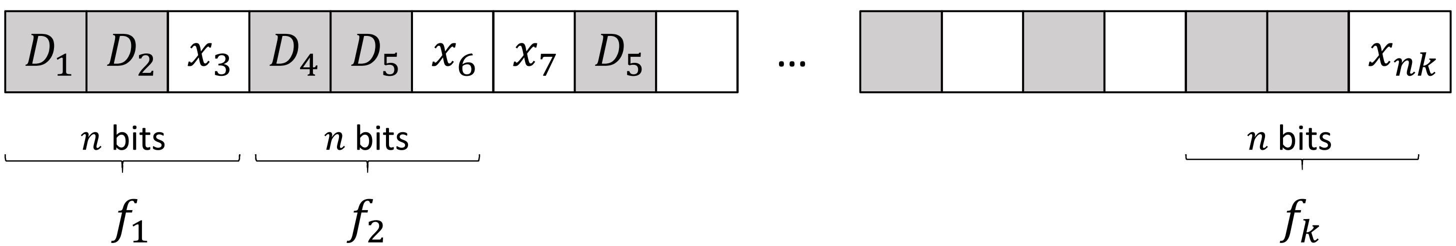
For $k \geq \sqrt{n}$,

- we apply the PRGs recursively
- similar to [RSV13], originated from [Gopalan-Meka-Reingold-Trevisan-Vadhan12]

Recursive construction

Sample E by

1. T: setting each bit to 1 with probability $\eta = 1/8$
2. Setting the 1-positions to uniform



- For every fixed $d \in D, t \in T$, F becomes a product test $F' = \prod_i f_i'$ on $|t|$ bits
- With high probability, each f_i has input length $\leq n/4$
- remains true when T is almost n-wise independent

Outline

1. Bounded independence, noise, product tests
2. Main Result
3. Complexity of Decoding
4. Pseudorandom generators
5. Proof Sketch
6. **Open questions**

Open Questions

Product test
$F: (\{0,1\}^n)^k \to [-1,1]$
$F(x_1, \dots, x_k) := \prod_i f_i(x_i)$

Theorem:

Let

- D := n-wise independent on nk symbols
- E := set each symbol to uniform independently with probability η

For any product test F,

$$|\mathrm{E}[F(D+E)] - \mathrm{E}[F(U)]| \leq (1-\eta)^{\Omega\left(\frac{n}{k}\right)}$$

Can we remove the $1/k$ in the exponent?

- Could give much better PRGs for any-order product tests

Open Questions

Product test
$F: (\{0,1\}^n)^k \to [-1,1]$
$F(x_1, \dots, x_k) := \prod_i f_i(x_i)$

Theorem:

- D: $m^{2/3}\log m$-wise independent on m bits
- E: set each bit to uniform independent with probability 0.01

For any logspace algorithm $A: \{0,1\}^m \to \{0,1\}$,

$$|\mathrm{E}[A(D+E)] - \mathrm{E}[A(U)]| \le o(1)$$

Can we use less independence?

Thank you!