Fundamental Techniques in Pseudorandomness

Part 1

Motivations and goals

(and previews of future lectures)

Goals:

Introduce fundamental problems

Introduce basic techniques

Show how basic techniques can be composed to provide solutions

Message:

everything follows from

- elementary algebra
- good definitions
- composition

How elementary is the linear algebra:

In every field:

- k linearly independent linear equations in n variables have a solution space of dimension $n-k$
- A non-zero degree-d polynomial in one variable has $\leq d$ roots

What problems do we want to solve?

1. Construct <u>efficiently</u> and deterministically objects whose existence is guaranteed by proofs based on the probabilistic method

2. Convert a randomized algorithm for a problem of interest into a deterministic algorithm of <u>comparable</u> <u>complexity</u>

3. <u>Efficiently</u> construct deterministically (or with little randomness) objects having many of the useful properties of random objects

Probabilistic Method : Example 1

Ramsey Theorem (Erdős, Szekeres)

Every n-vertex graph has either a clique or an independent set of size $\geq \frac{1}{2}\log n$

Erdős

There is an n-vertex graph in which max clique $\leq 2\log n$ and max i.s. $\leq 2\log n$

80-year old problem: match Erdős existence proof with an explicit (say, polynomial time) construction

Recent breakthrough (Chattopadhyay, Zuckerman, Cohen):

max clique, max i.s. $\leq \exp((\log\log n)^c)$

c absolute constant

Probabilistic Method: Example 2
Shannon's Second Theorem

Pick $C : \{0,1\}^n \to \{0,1\}^m$, known to both A and B

A computes and send $C(x)$

B receives corrupted transmission y

B guesses message as

$$x' = \operatorname*{argmin}_{z} d_H(C(z), y)$$

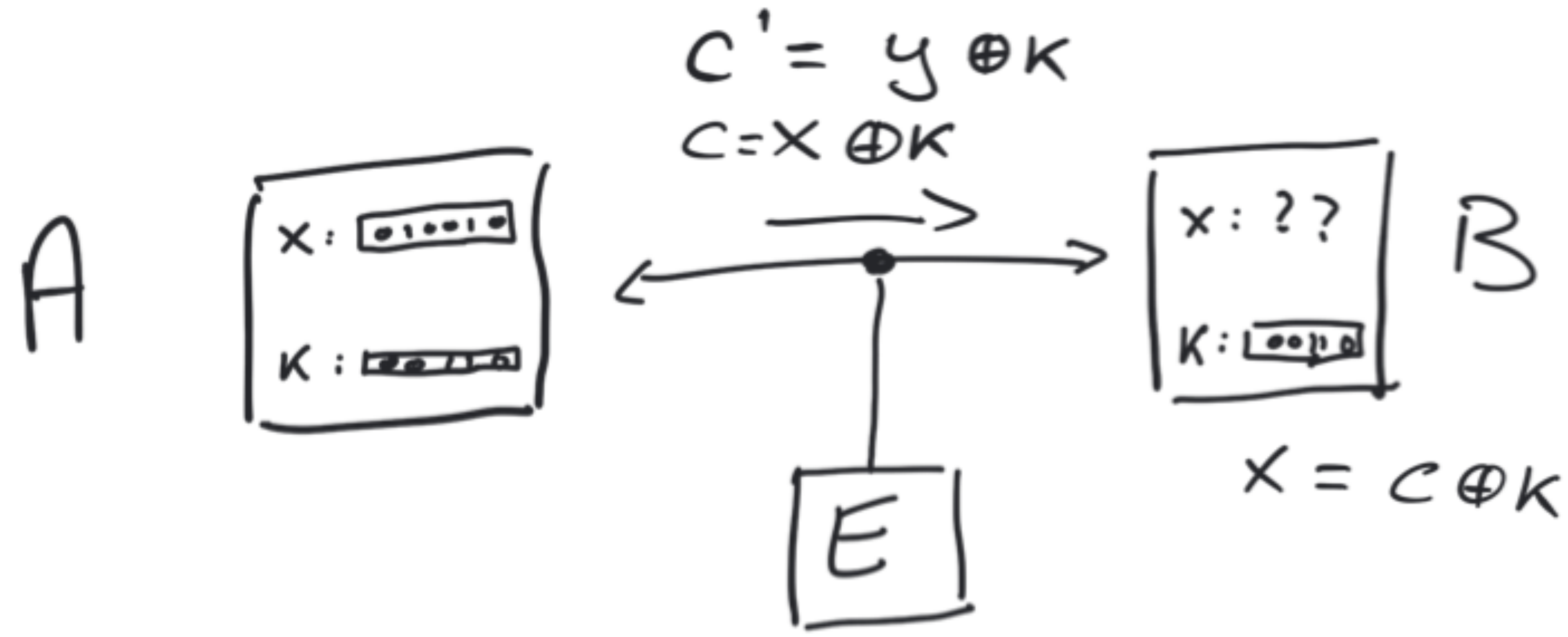
Thm (very informally)

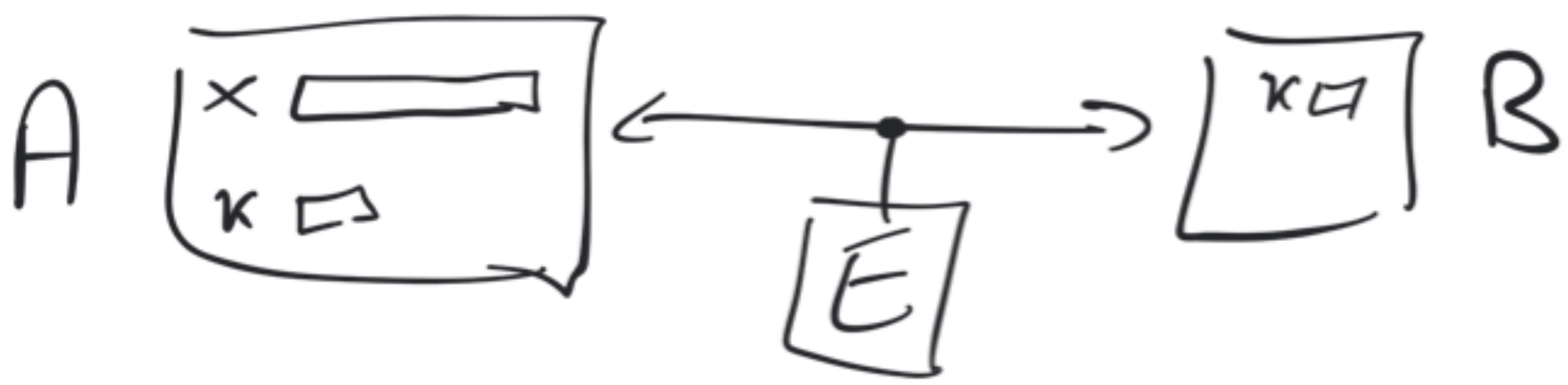
Either this works whp, or it is impossible for any coding scheme to reliably send an n-bits message by transmitting m bits over channel

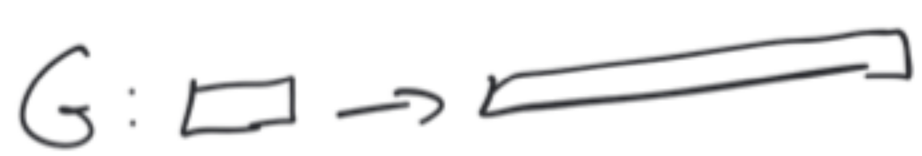
Note: - C has doubly exponential size
- Brute force decoding takes exponential time

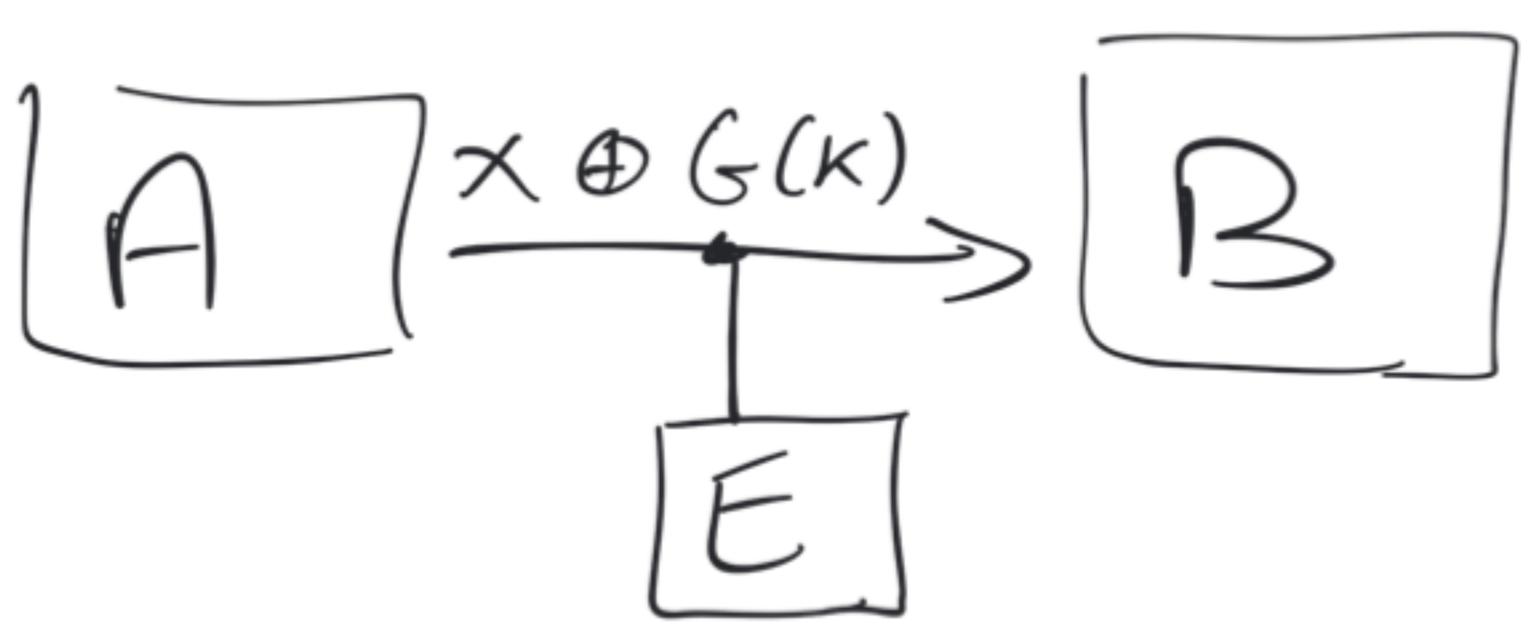
Lots of work in past 70 years toward making code explicit, and encoding and decoding polynomial (or even linear) time

Example 3 : One Time Pad









Derandomization of Algorithms

Example 1: Primality testing

Idea of Miller-Rabin and Solovay-Strassen randomized algorithms:

If N is composite, it has 'certificates of compositeness' (e.g. violations of Fermat's Little Theorem, or of Quadratic Reciprocity) that exist and can be found with high probability using randomness

To find a deterministic algorithm: construct certificates in polynomial time

AKS: - new randomized algorithm.

- given N, define polynomial P_N that is efficiently evaluable and non-zero iff N is composite
- find non-zero point of P_N randomly if one exists

- derandomization: show how to construct a non-zero point of P_N if one exists

Derandomization of Algorithms
Example 2: Undirected Reachability

Problem: given undirected graph G, nodes s, t, is there a path from s to t?

Aleliunas, Karp, Lipton, Lovász, Rackoff:

start at s
do a $100 \cdot n^3$-step random walk
if t is encountered, output YES
else output NO

Memory use: $O(1)$ variables, $O(\log n)$ bits

Reingold: same performance, deterministically

Derandomization of Algorithms

Example 3: Finding Large Primes

Problem: given N, find a prime between
N and $2N$

Randomized algorithm: pick $O(\log N)$
random integers

Deterministic algorithm: ??

Derandomization of Algorithms

Example 4: Polynomial Identity Testing

Problem: given two multivariate polynomials p, q, e.g. as formulas or as arithmetic circuits, is $p = q$?

Randomized Algorithm: check a random point in a suitably large discrete range

Deterministic Algorithm: ??
(c.f. AKS)

Derandomization of Algorithms
Example 5: Approximating the Permanent

Problem: given square 0/1 matrix M, approximate

$$\mathrm{perm}(M) := \sum_{\pi} \prod_{i} M_{i,\pi(i)}$$

Equivalently: given bipartite graph, approximate
the number of perfect matchings

(Approximate: achieve, say, 1% multiplicative
approximation)

Randomized Algorithm: long story
(Jerrum, Sinclair, Vigoda)

Deterministic Algorithm: ??

Derandomization of Algorithms

Example 6: Circuit acceptance

Problem: given boolean circuit C with one bit output, find a number in the range

$$\mathop{\mathbb{P}}_{x \sim U_n} [\, C(x) = 1 \,] \pm \frac{1}{10}$$

Randomized algorithm: evaluate C at 1,000 random inputs, output fraction of times you see 1

Deterministic algorithm: ??

Note: a derandomization of this algorithm implies a derandomization of all algorithms

Complexity-Theoretic Questions

P = BPP?

Can we solve in polynomial time, deterministically, all problems that we can solve in polynomial time probabilistically whp?

Note: • implied by derandomization of circuit approx problem

• also implied by plausible circuit-complexity conjectures

• it implies unproven (and hard) circuit complexity conjectures

L = BPL ?

Is every problem solvable in polynomial time and logarithmic space prob. whp also solvable in poly time and log space deterministically ?

• Non-trivial result: can do in space $O((\log n)^{1.5})$, time $n^{O(\sqrt{\log n})}$

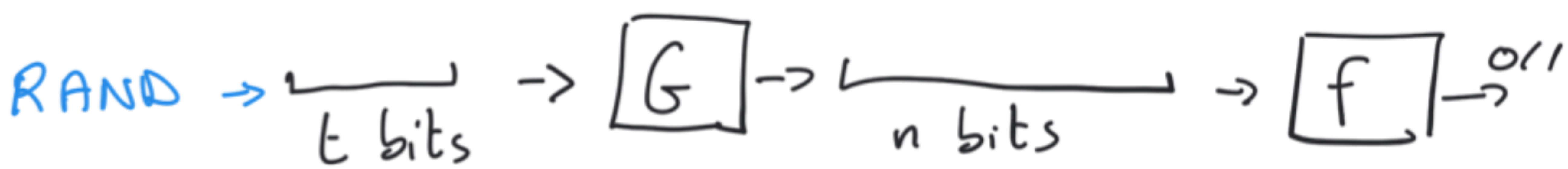
• No known barrier

Unifying Notion:
Pseudorandom Generator

$G : \{0,1\}^t \to \{0,1\}^n$ ε-fools

a family of tests $\mathcal{F}$,

where each $f \in \mathcal{F}$ is $f : \{0,1\}^n \to \{0,1\}$



$$\forall f \in \mathcal{F}. \left| \underset{x \sim U_n}{\mathbb{P}}[f(x)=1] - \underset{z \sim U_t}{\mathbb{P}}[f(G(z))=1] \right| \le \varepsilon$$

Ideal Case of Existence of PRGs

Suppose we have poly(n)-time computable

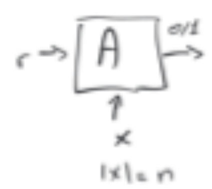
$$G_n : \{0,1\}^{O(\log n)} \to \{0,1\}^n$$

that $\frac{1}{10}$-fools all f computable by circuits of size $\leq n^2$

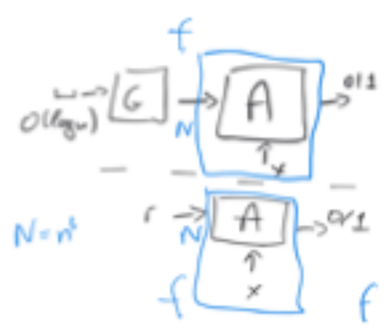
Then:

- $P = BPP$
- All applications of prob method to construct objects with poly(n)-time checkable properties can be made constructive
- All randomized approximation algorithms can be derandomized

---



A runs in n^3 time
$|r| = n^3$
correct with 90% prob



correct with ~~< 80%~~ prob
> 80%

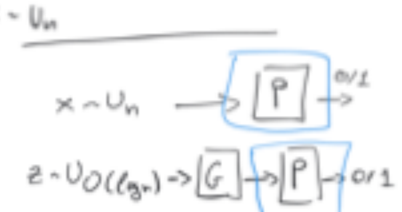
correct with > 90% prob

f computable in size $\ll N^2$, not $\frac{1}{10}$-fooled

Interested in constructing $\in \{0,1\}^n$ that satisfies property P

P is checkable in n^3 time

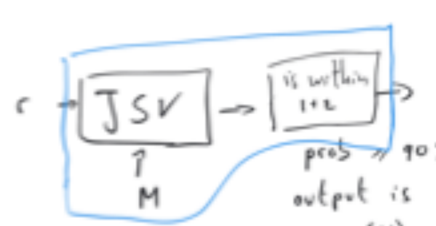
$\mathbb{P}_{x \sim U_n}[P(x) \text{ true}] \geq 20\%$



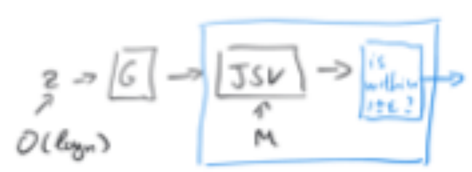
1 with prob ≥ 20%

1 with prob ≥ 10% ~~< 10%~~

$G : \{0,1\}^{O(\log n)} \to \{0,1\}^{n^2}$
look at first n bits



prob ≥ 90%
output is
$\text{perm}(M) \cdot (1 \pm \varepsilon)$



prob ≥ 80%
~~prob~~ 80%
output is
$\text{perm}(M) \cdot (1 \pm \varepsilon)$

Part 2

Simple constructions of Pseudorandom Generators

Tests that look at only one bit

construct:

$$G : \{0,1\}^t \to \{0,1\}^N$$

such that, for uniform x,

each bit of $G(x)$ is uniformly distributed

Tests that look at 2 bits

construct

$$G : \{0,1\}^t \to \{0,1\}^N$$

such that, for uniform x, the bits of $G(x)$ are uniformly distributed and pairwise independent

$$N = 2^t - 1$$

$$z \to \langle z, a_1 \rangle, \langle z, a_2 \rangle, \ldots, \langle z, a_{2^t-1} \rangle$$

$z \in \mathbb{F}_2^t$ where $a_1, \ldots, a_t$ is an enumeration of $\mathbb{F}_2^t - \{\underline{0}\}$

$$\langle z, a \rangle = 0 \wedge \langle z, a' \rangle = 0$$

$$z_1 z_2 z_3 \to z_1, z_2, z_3, z_1 + z_2, z_1 + z_3,$$

$$z_2 + z_3, z_1 + z_2 + z_3$$

Tests that look at two values

construct

$$G : \{0,1\}^t \to (\{0,1\}^n \to \{0,1\}^m)$$

such that, for uniform x, the outputs of $h_x := G(x)$ are pairwise independent

---

$n = m$

$\mathbb{F}_{2^n}$ $\quad (\mathbb{Z}_2)^n$

input of genarator $a, b \in \mathbb{F}_{2^n}$

$t = 2 \cdot n$

$h_{a,b}(x) = ax + b$

For every $x \neq y$

over the rand. of a, b

$ax + b$
$ay + b$ are uniform and indep.

---

what is prob. for fixed v, w

$ax + b = v$
$ay + b = w$ prob over a, b

We can generate

$h : \{0,1\}^n \to \{0,1\}^m$ $\quad m < n$ ✓

pairwise indep.

input is $2 \cdot n$ bits $\quad m > n$

$2 \cdot \max\{m, n\}$ bits

Tests that look at k values

construct

$$G : \{0,1\}^t \to (\{0,1\}^n \to \{0,1\}^m)$$

such that, for uniform z, $h_z := G(z)$

has k-wise independent outputs

$n = m$
$k = 3$

$t = 3 \cdot n$

input of G $\quad a, b, c \in \mathbb{F}_{2^n}$

$$h_{a,b,c}(x) = ax^2 + bx + c$$

Fix x, y, z

Fix u, v, w

$$\mathbb{P}_{a,b,c}\left[\begin{array}{l} h_{a,b,c}(x) = u \\ h_{a,b,c}(y) = v \\ h_{a,b,c}(z) = w \end{array}\right] \overset{?}{=} \left(\frac{1}{2^n}\right)^3$$

$$\| $$

$$\frac{\#a,b,c : \left\{\begin{array}{l} ax^2 + bx + c = u \\ ay^2 + by + c = v \\ az^2 + bz + c = w \end{array}\right.}{(2^n)^3} = \frac{1}{(2^n)^3}$$

In general

$$\{0,1\}^{kn} \to (\{0,1\}^n \to \{0,1\}^n)$$

k-wise $\quad \{0,1\}^n \to \{0,1\}^m$

$t = k \cdot \max\{n, m\}$

Applications

Suppose $G : \{0,1\}^t \to \{0,1\}^n$ is k-wise indep.

[Recall: we can have $t = k\log n$]

What functions $f : \{0,1\}^n \to \{0,1\}$ are guaranteed to be ε-fooled by G? For what ε?

① $\boxed{f : \{0,1\}^n \to \mathbb{R} \text{ is a polynomial of degree } \le k}$

Then

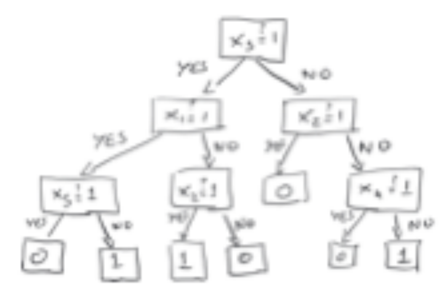
$$\mathop{\mathbb{E}}_{x \sim U_n} f(x) = \mathop{\mathbb{E}}_{z \sim U_t} f(G(z))$$

So if $f : \{0,1\}^n \to \{0,1\}$ has Fourier transform in which only coefficients of degree $\le k$ are non-zero,

$$\mathbb{P}[f(x)=1] = \mathbb{P}[f(G(z))=1]$$

Example: decision tree of depth $\le k$

$f(x_1, x_2, x_3, x_4, x_5) =$



$f(1,0,0,1,1) = 0$

A depth-k decision tree is 0-fooled by a k-wise indep. distribution

---

② $f(x_1, x_2, \dots, x_n) = x_1 \wedge x_2 \wedge \dots \wedge x_n$

Fourier degree is n

$$\mathop{\mathbb{P}}_{x \sim U_n}[f(x) = 1] = 1/2^n$$

if G is k-wise indep.

$$\mathop{\mathbb{P}}_{z \sim U_t}[f(G(z)) = 1] \le 1/2^k$$

A k-wise independent distribution fools AND with $\varepsilon \le 1/2^k$

---

③ $f : \{0,1\}^n \to \{0,1\}$ such that $g : \{0,1\}^n \to \{0,1\}$ has degree $\le d$ and

$$\mathop{\mathbb{P}}_{x \sim U_n}[f(x) \ne g(x)] \le \varepsilon$$

$G : \{0,1\}^{k\log n} \to \{0,1\}^n$ $\qquad f(x) = \begin{cases} 1 & x \in \mathrm{Im}(G) \\ 0 & x \notin \mathrm{Im}(G) \end{cases}$

---

④ $f, u, \ell : \{0,1\}^n \to \mathbb{R}$ [Bazzi]

where:

- $u(\cdot), \ell(\cdot)$, have degree $\le k$
- $\forall x.\ \ell(x) \le f(x) \le u(x)$
- $\mathop{\mathbb{E}}_{x \sim U_n} u(x) - \ell(x) \le \varepsilon$
- suppose $G : \{0,1\}^t \to \{0,1\}^n$ k-wise indep.

Then $\left| \mathop{\mathbb{E}}_{x \sim U_n} f(x) - \mathop{\mathbb{E}}_{z \sim U_t} f(G(z)) \right| \le \varepsilon$

$$\mathbb{E} f - \mathbb{E} f(G(\cdot))$$

$$\le \mathop{\mathbb{E}}_{x \sim U_n} u(x) - \mathop{\mathbb{E}}_{z \sim U_t} \ell(G(z))$$

$$= \mathop{\mathbb{E}}_{x \sim U_n} u(x) - \mathop{\mathbb{E}}_{x \sim U_n} \ell(x) \le \varepsilon$$

Pseudorandomness for bounded-depth Circuits

Suppose $f : \{0,1\}^n \to \{0,1\}$

is computed by a Boolean circuit with:

- AND, OR, NOT gates
- AND, OR gates have no bound on fan-in, fan-out
- depth $\leq d$
- size $= s$

Then [Braverman]

There are $\ell, u : \{0,1\}^n \to \mathbb{R}$ of degree $\leq (\log \frac{s}{\varepsilon})^{O_d(1)}$ such that

$\forall x.$ $\quad \ell(x) \leq f(x) \leq u(x)$

$$\mathbb{E}\ u(x) - \ell(x) \leq \varepsilon$$

Constructing Optimal Ramsey Graphs in $n^{\text{polylog}\, n}$ time

Erdős: if $c = 2\log n$, then with positive probability a $G_{n,\frac{1}{2}}$ graph has no clique of size c and no ind. set of size c

Proof: $\mathbb{E}(\#\text{ cliques of size } c)$
$+ \mathbb{E}(\#\text{ ind. set of size } c)$
< 1

Fix $Gen: \{0,1\}^t \to \{0,1\}^{\binom{n}{2}}$
$\binom{c}{2}$- wise independent

Enough to have $t = \binom{2\log n}{2} \cdot \log \binom{n}{2}$
$= O(\log^3 n)$

Interpret output of Gen as n-vertex graph, call resulting distribution $\tilde{G}_{n,\frac{1}{2}}$

$\mathbb{E}[\#\text{ clique of size } c]$
$+ \mathbb{E}[\#\text{ ind. of size } c]$
< 1

Find one: $2^{O(\log^3 n)} \cdot n^{2\log n} =$
$n^{O(\log^2 n)}$

Tests that take XORs

construct

$$G : \{0,1\}^t \to \{0,1\}^N$$

such that, for every $a \in \{0,1\}^N$,

$$\mathbb{P}_{x \sim U_t} \left[\langle G(x), a \rangle = 1 \right] \in \frac{1}{2} \pm \varepsilon$$

where operations are in $\mathbb{F}_2$

Input $a \in \mathbb{F}_{2^t}$ $\quad b \in (\mathbb{Z}_2)^t$

$a, b \to \langle a, b \rangle, \langle a^2, b \rangle, \ldots \langle a^N, b \rangle$

take

$$\sum_{i \in S} \langle a^i, b \rangle$$

$$= \langle \sum_{i \in S} a^i, b \rangle$$

$$\mathbb{P}_{a,b \in \{0,1\}^t} \left[\langle \sum_{i \in S} a^i, b \rangle = 0 \right]$$

$$= \mathbb{P}_a \left[\sum_{i \in S} a^i = 0 \right] + \frac{1}{2} \left(1 - \mathbb{P}[\Sigma = 0] \right)$$

$$= \frac{1}{2} + \frac{1}{2} \mathbb{P}_a \left[\sum_{i \in S} a^i = 0 \right]$$

$$\le \frac{1}{2} + \frac{1}{2} \frac{N}{2^t}$$

$$\{0,1\}^{2t} \to \{0,1\}^N$$

ε-fool lin. test

$$\varepsilon \sim N/2^t$$

$$t = O\left(\log \frac{N}{\varepsilon}\right)$$

Applications

Take $f : \{0,1\}^n \to \mathbb{R}$

write as

$$f(x) = \sum_s \hat{f}(s) \cdot (-1)^{\langle a, x \rangle}$$

Suppose $G : \{0,1\}^t \to \{0,1\}^n$ is ε-biased

Then

$$\forall a. \quad \left| \mathop{\mathbb{E}}_{U_n} (-1)^{\langle a, x \rangle} - \mathop{\mathbb{E}}_{U_t} (-1)^{\langle a, G(z) \rangle} \right|$$

$$\leq 2\varepsilon$$

So

$$\left| \mathop{\mathbb{E}}_{U_n} f(x) - \mathop{\mathbb{E}}_{U_t} f(G(z)) \right| \leq 2\varepsilon \cdot \sum_s |\hat{f}(s)|$$

---

k-wise independence: fools f of degree k (no error)

ε-biased: fools f w/ small L_1 (error $\varepsilon \cdot \|\hat{f}\|_1$)

Cauchy-Schwarz

If $f : \{0,1\}^n \to \{0,1\}$ depends on $\leq k$ vars

$$\sum_s |\hat{f}(s)| \leq \sqrt{2^k} \cdot \sqrt{\sum_s \hat{f}^2(s)} \leq \sqrt{2^k}$$

Often, if f is fooled by k-wise independence, it is also fooled by ε-biased generators with $\varepsilon \approx 1/2^k$

Better: seed goes from $k \log n$ to $\frac{1}{2}k + \log n$

$f =$ Depth k decision tree

k-wise ind. fool f

$t = k \cdot \log n$

$f =$ decision tree of size S

$\|\hat{f}\|_1 \leq S$

$f =$ depth k d.t.

size $\leq 2^k$

$\varepsilon/2^k$-bias generator, ε-fool f

$t = \log \frac{n \cdot 2^k}{\varepsilon} = k + \log n/\varepsilon$

---

f comp. d.t. of depth $O(\log n)$

size poly(n)

$O(\log n)$-wise indep. $\qquad t = O(\log^2 n)$

$1/\text{poly}(n)$-bias $\qquad t = O(\log n)$

Optimal Ramsey graphs in time $n^{O(\log n)}$

Take $\mathrm{Gen}: \{0,1\}^t \to \{0,1\}^{\binom{n}{2}}$

to be ε-biased with $\varepsilon = \exp(-100 \log^2 n)$

Intepret output of Gen as graph,

call $\tilde{G}_{n,\frac{1}{2}}$ resulting distribution

$$\mathop{\mathbb{E}}_{G \sim \tilde{G}_{n,\frac{1}{2}}} \#(2\log n)\text{-cliques in } G$$

$$= \sum_{\substack{S \subseteq V \\ |S| \le 2\log n}} \mathop{\mathbb{E}}_{\tilde{G}} I_{S \text{ is a clique}}(G)$$

$$\le \sum_{\substack{S \subseteq V \\ |S| \le 2\log n}} \left(\mathop{\mathbb{E}}_{G_{n,\frac{1}{2}}} I_{S \text{ is a clique}} + \varepsilon \cdot 2^{2\log^2 n} \right)$$

$$\le \underbrace{\text{Erdős}}_{G_{n,\frac{1}{2}}} + \varepsilon \cdot n^{2\log n} \cdot n^{2\log n}$$

$$\le \boxed{\text{Erdős}} + 2^{-96 \log^2 n}$$

$f : \{0,1\}^n \to \{0,1\}$

$f(x) = \sum_S c_S x^S$

ε-bias distrib

$\varepsilon \cdot \sum_S |c_S|$ - fool f

Error-correcting codes

$$C : \Sigma^k \to \Sigma^n$$

is an error-correcting code of minimum distance $\geq \Delta$ if

$$\forall x \neq y \in \Sigma^k$$

$$d_H(C(x), C(y)) \geq \Delta$$

Motivating application

find x'
$d_H(C(x'), z) < \Delta/2$

Channel: · can be used to transmit n elements of Σ

· is guaranteed to make $< \frac{\Delta}{2}$ errors

Linear Error-Correcting Codes

$C : \mathbb{F}^k \to \mathbb{F}^n$

- is linear
- is an error-correcting code of minimum distance $\geq \Delta$

Possible to use linear algebra to reason about encoding, decoding

E.g.

call $|y| =$ # non-zero entries of y

Then

(1) $d_H(y, z) = |y - z|$

(2) C has min distance $\geq \Delta$

iff

$\forall x \neq \underline{0}.\ |C(x)| \geq \Delta$

Proof

$$\min_{x \neq y} |C(x) - C(y)| = \min_{x \neq y} |C(x-y)|$$
$$= \min_{x \neq 0} |C(x)|$$

---

(3) There is matrix M such that

$C(x) = M \cdot x$

Also, there is matrix P such that

$y \in \{C(x) : x \in \mathbb{F}^k\} \iff P \cdot y = \underline{0}$

(either P or M determines C)

Note: C has min-distance $\geq \Delta$

iff

every $\leq \Delta - 1$ rows of P are linearly independent

Reed-Solomon Codes

$C : \mathbb{F}^K \to \mathbb{F}^n \quad [\ n \leq |\mathbb{F}|\ ]$

choose $a_1, \ldots, a_n \in \mathbb{F}$

given $x_0 \ldots x_{K-1}$ Let $P_x(z) = z^{K-1} x_{K-1} + \cdots + z x_1 + x_0$

$C(x) = P_x(a_1), P_x(a_2), \ldots, P_x(a_n)$

min distance $\geq n - K + 1$

'Hadamard' code

$$C : \mathbb{F}_2^t \to \mathbb{F}_2^{2^t}$$

Let $a_1, \dots, a_{2^t}$ be the elements of $\mathbb{F}_2^t$

$$C(x) = \langle x, a_1 \rangle, \dots, \langle x, a_{2^t} \rangle$$

min distance $= \frac{1}{2} \cdot 2^t$

$C(x) - C(y)$

$= \langle x-y, a_1 \rangle, \dots, \langle x-y, a_{2^t} \rangle$

Concatenation

Suppose we have

$$C_1 : \Sigma^K \to \Sigma^N$$

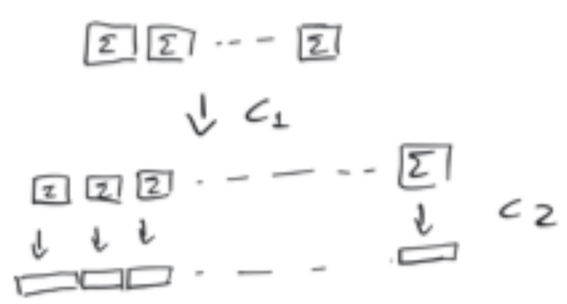
min distance $\geq \Delta_1$

$$C_2 : \Sigma \to \{0,1\}^n$$

min distance $\geq \Delta_2$

Then

$$C : \Sigma^K \to \{0,1\}^{n \cdot N}$$



has min distance $\geq \Delta_1 \cdot \Delta_2$

Eg Concatenate Reed-Solomon

$$RS : \mathbb{F}_{2^t}^K \to \mathbb{F}_{2^t}^n$$

with $H : \mathbb{F}_{2^t} \to \mathbb{F}_2^{2^t}$

---

we get

$$C : \{0,1\}^{t \cdot K} \to \{0,1\}^{2^t \cdot n}$$

with min distance $\geq \frac{1}{2} \cdot (n-K) \cdot 2^t$

choose $n = \frac{K}{\varepsilon}$, $2^t = 2n$

$$C : \{0,1\}^{K \log \frac{2K}{\varepsilon}} \to \{0,1\}^{2K^2/\varepsilon^2}$$

$$C : \{0,1\}^K \to \{0,1\}^n \qquad n \doteq \frac{K^2}{\varepsilon^2}$$

min distance $\geq n \cdot (\frac{1}{2} - \varepsilon)$

max distance $\leq \frac{n}{2}$

Linear Error-Correcting Codes vs k-wise independence

Suppose

$$C : \mathbb{F}^t \to \mathbb{F}^n$$

is linear e.c.c. of min dist $\geq k+1$

Let P be $(n-t) \times n$ matrix s.t.

$$Py = \underline{0} \quad \text{iff} \quad y \in \mathrm{Im}(C)$$

If $0 < |y| \leq k$, then $P \cdot y \neq \underline{0}$

Every $\leq k$ columns of P are linearly ind.

Consider $G : \mathbb{F}^{n-t} \to \mathbb{F}^n$ given by

$$x \to P^T x$$

$$x \to P_1^T \cdot x \,,\, P_2^T x \,,\, \dots, P_n^T \cdot x$$

This is k-wise indep. because every k output bits correspond to multiplying x by linearly indep. vectors

---

Code of min distance 2

$$x_1, \dots, x_{n-1} \to x_1, \dots x_{n-1}, \sum_{i=1}^{n-1} x_i$$

$$C : \{0,1\}^{n-1} \to \{0,1\}^n$$

$$P = (1, \dots, 1)$$

$$b \to b, b, \dots, b$$

Linear Error-Correcting Codes vs ε-biased spaces

Suppose $C : \mathbb{F}_2^k \to \mathbb{F}_2^n$
is such that $\forall\, y \in \mathrm{Im}(C)$, $y \neq \underline{0}$

$$\left(\tfrac{1}{2} - \varepsilon\right) \cdot n \leq |y| \leq \left(\tfrac{1}{2} + \varepsilon\right) \cdot n$$

if $C(x) = M \cdot x$

$$M = n\left\{\underbrace{\begin{pmatrix} - M_1 - \\ \vdots \\ - M_n - \end{pmatrix}}_{k}\right.$$

picking at random a row of M gives an
an ε-biased ~~space~~ generator

$$\{0,1\}^{\log n} \to \{0,1\}^k$$

Fix $a \in \{0,1\}^k$
pick at random $i \in \{1, \dots, n\}$ — ε-close to unbiased
consider $\langle M_i, a \rangle$

— — — —

pick a random i — ε-close to unbiased
consider i-th bit of $M \cdot a$

— — — —

consider vector $M \cdot a = C(a)$
n-dim vector
$\geq \left(\frac{1}{2} - \varepsilon\right) \cdot n$ ones

$\leq \left(\frac{1}{2} + \varepsilon\right) \cdot n$ one

Samplers

Given oracle access to $f : \{0,1\}^n \to [0,1]$

Output an estimate A of $\mathbb{E}_{x \sim U_n} f(x)$

such that

$$\mathbb{P}_{\substack{\text{randomness} \\ \text{of algorithm}}} \left[\left| A - \mathbb{E}_{x \sim U_n} f(x) \right| > \varepsilon \right] \leq \delta$$

How many queries and how much randomness do we need as a function of n, ε, δ ?

Solution 1: independent queries

$\boxed{\text{Make } t \text{ independent queries } x_1, \ldots, x_t, \text{ output } A := \frac{1}{t}\sum_i f(x_i)}$

Chernoff bound:

$$\mathbb{P}\left[\, \left| A - \mathop{\mathbb{E}}_{U_n} f(x) \right| > \varepsilon \right] \le 2 e^{-\varepsilon^2 t/2}$$

choose $t = \Theta\left(\frac{1}{\varepsilon^2} \log \frac{1}{\delta}\right)$

Complexity [ignore constants]

Queries: $\frac{1}{\varepsilon^2} \cdot \log \frac{1}{\delta}$

Randomness: $n \cdot \frac{1}{\varepsilon^2} \log 1/\delta$

Solution 2: Pairwise Independent Queries

Generate t pairwise independent elements $x_1 \ldots x_t$ of $\{0,1\}^n$

Output $A := \frac{1}{t} \sum_i x_i$

Chebyshev Inequality

$$\mathbb{P}\left[\, \left| A - \mathop{\mathbb{E}}_{x \sim U_n} f(x) \right| > \varepsilon \right] \leq \frac{1}{\varepsilon^2 t}$$

Take $t = \frac{1}{\varepsilon^2 \delta}$

Complexity

Method	Queries	Randomness
Independent	$\frac{1}{\varepsilon^2} \log \frac{1}{\delta}$	$n \cdot \frac{1}{\varepsilon^2} \cdot \log \frac{1}{\delta}$
Pairwise ind.	$\frac{1}{\varepsilon^2} \cdot \frac{1}{\delta}$	$n + \log \frac{1}{\varepsilon} + \log \frac{1}{\delta}$

New tool: Random Walks on Expanders

Let $H = (V, E)$ be a d-regular graph

$d = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ be e-values of adjacency matrix

call $\lambda = \max\{|\lambda_2|, |\lambda_3|, \dots, |\lambda_n|\}$

[we can get $\lambda = \Theta(\sqrt{d})$]

Let $f : V \to [0,1]$ be arbitrary

Let $x_1 \dots x_t$ be the sequence of vertices encountered by taking a $(t-1)$-step random walk in H (x_1 is uniform)

[$\log|V| + (t-1) \cdot \log d$ random bits used]

THEN (Chernoff bound on expanders)

$$\mathop{\mathbb{P}}_{\text{random walk}}\left[\left| \frac{1}{t} \sum_i f(x_i) - \mathop{\mathbb{E}}_{x \sim V} f(x) \right| > \varepsilon + \frac{\lambda}{d} \right] \leq e^{-\Omega(\varepsilon^2 t)}$$

Solution 3: Random Walk on Expanders

Construct $H = (\{0,1\}^n, E)$ such that $\lambda/d \le \frac{\varepsilon}{2}$. (Enough to take $d = \Theta(1/\varepsilon^2)$)

Pick a random walk $x_1 \dots x_t$ in H

Output $A := \frac{1}{t} \sum_i f(x_i)$

$$\mathbb{P}[\,|A - \mathbb{E}f| > \varepsilon\,]$$

$$= \mathbb{P}[\,|A - \mathbb{E}f| > \tfrac{\varepsilon}{2} + \tfrac{\lambda}{d}\,] \le e^{-\Omega(\varepsilon^2 t)}$$

enough to take $t = O(\frac{1}{\varepsilon^2} \log \frac{1}{\delta})$

Method	Queries	Randomness
Independent	$\frac{1}{\varepsilon^2} \log \frac{1}{\delta}$	$n \cdot \frac{1}{\varepsilon^2} \log \frac{1}{\delta}$
Pairwise ind.	$\frac{1}{\varepsilon^2} \cdot \frac{1}{\delta}$	$n + \log \frac{1}{\varepsilon} + \log \frac{1}{\delta}$
Expander r.w.	$\frac{1}{\varepsilon^2} \cdot \log \frac{1}{\delta}$	$n + \frac{1}{\varepsilon^2} \cdot \log \frac{1}{\delta} \cdot \log \frac{1}{\varepsilon}$

Solution 4 : composition