

Linear functions:

f is "linear" iff $\forall x, y \quad f(x) + f(y) = f(x+y)$
← actually these are homomorphisms

will consider $f: \{0,1\}^d \rightarrow \{0,1\}$
 here, "linear fncts" are the parity fncts *(XOR)*

observation $\forall x, y \quad f(x) + f(y) = f(x+y)$
 iff

$$f(x) = \bigoplus_{i \in S} x_i \quad \text{for some } S \subseteq [d]$$

K -linear fncts:

f is " K -linear" if

- (1) linear
- (2) depends on $=K$ variables

ie. $|S| = K$

← also called "K-junta fnctn"

linearity testing:

given $f: \{0,1\}^n \rightarrow \{0,1\}$ is f linear?

$$\text{i.e. } \forall x, y \quad f(x) + f(y) = f(x+y) \quad ?$$

e.g. $\forall x \quad f(x) = 0$ is linear

$\forall x \quad f(x) = 1$ is not

$\forall x \quad f(x) = x \cdot b$ inner prod of $x \cdot b$
is linear

Thm Can properly test linearity in $O(1)$ queries:

linearity test:

Pick random x, y + fail if $f(x) + f(y) \neq f(x+y)$

Proof later lecture

Consider functions $f: \{0,1\}^d \rightarrow \{0,1\}$ here, domain size = $2^d \equiv n$

Testing k -linear functions: e.g. $f(x) = \bigoplus_{i \in S} x_i$ s.t. $|S| \leq k$

related to testing if fctn is k -junta (depends only on k vars), low Fourier degree, computable by small depth decision trees, ...

First Algorithm: ("learns" f) wlog assume $f(\vec{0}) = 0$

$O(d) = O(\log n)$ {

Given

Query f on all $e_i = (0 \dots 0 \underset{i^{th} \text{ locn}}{1} 0 \dots 0)$ for $i = 1 \dots d$ \uparrow $\log n$

+ $(0 \dots 0)$

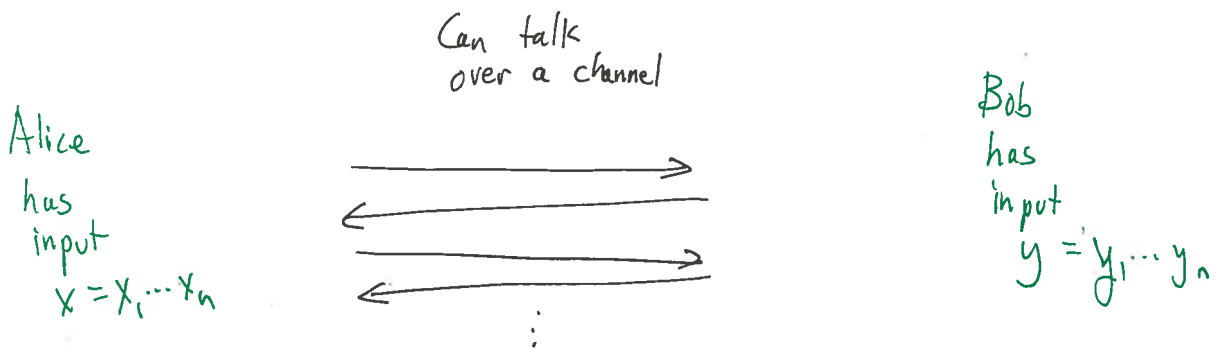
if $f(e_i) = 1$ for $\neq k$ i 's then fail

else, test if $f(x) = \bigoplus_{i \text{ s.t. } f(e_i)=1} x_i$ for most x via sampling

Can we do better?

What is Communication Complexity?

Setting:



Goal Compute $f(x,y)$ ← how many bits, rounds of communication required?

examples:

1) $f(x,y) = (\bigoplus_i x_i) \oplus (\bigoplus_i y_i)$

- requires 2 bits/round of communication
- A → B $\bigoplus_i x_i$
- B → A $f(x,y)$ (or $\bigoplus_i y_i$)

2) $f(x,y) = \sum x_i + \sum y_i$

- requires $O(\log n)$ bits

3) $f(x,y) = \begin{cases} 1 & \text{if } x=y \\ 0 & \text{o.w.} \end{cases}$

- requires $O(\log n)$ bits
 - A → B $\sum x_i$
 - B → A $\sum y_i$ (or $f(x,y)$)
- ← can we do better?

4) $f(x,y) = \text{"do } x+y \text{ agree on any bit?"}$

- requires $\Theta(n)$ bits

Communication Complexity (CC) lower bounds (we have these!)

cc/lb
n

⇒ Property testing (PT) lower bounds

Idea give reduction from CC problem to PT problem

⇒ L.B. for CC. problem yields

L.B. for P.T. problem

a lot of great work done in this area

so we get this almost for free!!

Example:

• A hard CC problem
SET DISJOINTNESS

Alice

$$x \in \{0,1\}^n$$

Bob

$$y \in \{0,1\}^n$$

$$\text{Disj}(x,y) = \bigvee_{i=1}^n (x_i \wedge y_i)$$

do A+B agree on any bit?

Known lb.: $\Omega(n)$ bits of communication required to solve it.

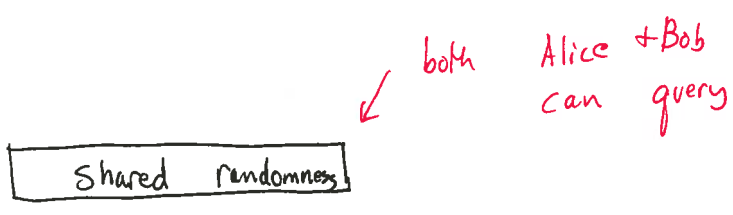
even if allow many rounds, probabilistic protocols

Sparse Set disjointness: A+B have at most k 1's

needs $\Omega(k)$ bits communication (even if guaranteed that intersect only once or not at all)

How can we use this to lower bound PT problems?

A reduction from sparse set disjointness to PT for k -linearity:



Alice

Bob

Set A { n bit vector $\{0,1\}^n$
with exactly k 1's
in it
describing k -linear fctn f
(ie. f is XOR of
bits with indices
in A)



n bit vector $\{0,1\}^n$
with k 1's
describing k -linear
fctn g

Question:

does $h = f \oplus g$
 have $2k$ -linearity property?

note:

if $A \cap B = \emptyset$ then h is $2k$ -linear

if $A \cap B \neq \emptyset$ then h is j -linear
 for $j \leq 2k - 2$.

e.g. if $A = \{x_1, x_2\}$ + $B = \{x_3, x_4\}$

$$A \cap B = \emptyset$$

$$f = x_1 \otimes x_2$$

$$g = x_3 \otimes x_4$$

$$h = x_1 \otimes x_2 \otimes x_3 \otimes x_4 \leftarrow 4 \text{ linear}$$

if $A = \{x_1, x_2\}$ $B = \{x_2, x_3\}$

$$A \cap B = \{x_2\}$$

$$f = x_1 \otimes x_2$$

$$g = x_2 \otimes x_3$$

$$h = x_1 \otimes \underbrace{x_2 \otimes x_2}_{=1} \otimes x_3$$

$$= x_1 \otimes x_3 \leftarrow 2 \text{ linear}$$

for all x_i in $A \cap B$,

two variables drop out of h

so h is $(k - 2|A \cap B|)$ -linear

Fact if $h_1 \neq h_2$ are 2 linear fctns (for any k)

$$\text{then } \frac{\#x \text{ st. } h_1(x) \neq h_2(x)}{2^n} = \frac{1}{2}$$

We will prove this in h.w.

\Rightarrow if $A \cap B = \emptyset$, h is $\frac{1}{2}$ -far from $2k$ -linear

Why is this interesting?

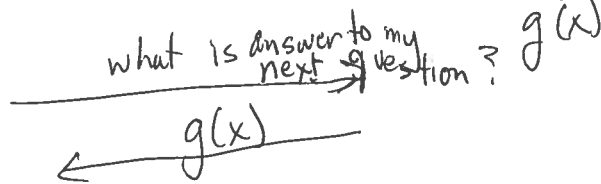
protocol for testing $2k$ -linearity of h
with q queries \Rightarrow C.C. protocol for set disjointness of A, B

Shared random string which contains random bits for Alice's queries $\{R\}$

A runs prop test alg. When needs

$$h(x) = f(x) \oplus g(x):$$

- 1) compute $f(x)$
- 2) ask Bob for $g(x)$
- 3) output $f(x) \oplus g(x)$ as $h(x)$



Bob simulates A's run on R .
Bob computes x & then $g(x)$

Note: Alice doesn't need to send x 's

$$\text{Total communication} = 2q \text{ bits}$$

$$\Rightarrow q = \Omega(k)$$

Thm k -linearity testing requires $\Omega(k)$ queries!
Interesting, since linearity testing only needs $O(1)$!