

Some Cool things we can do:

① Estimate sum of powers of degree 1 Fourier Coeffts of a fctn.

$\hat{f}(s)$ for $|s|=1$:

$\hat{f}(\xi_i)$
denote as $\hat{f}(i)$

What is $\sum_{i=1}^n \hat{f}(i)^p$?

A random process:

Pick $X^{(1)} X^{(2)} \dots X^{(p-1)} \in \{\pm 1\}^n$

Pick noise vector μ st. each entry

is $+1$ with prob $\frac{1}{2} + \frac{1}{2}\eta$
 -1 " " $\frac{1}{2} - \frac{1}{2}\eta$

$y \leftarrow f(x^{(1)}) f(x^{(2)}) \dots f(x^{(p-1)}) f(x^{(1)} \odot x^{(2)} \odot \dots \odot x^{(p-1)} \odot \mu)$

Output y

Claim $E[y] = \sum_{S \subseteq [n]} \eta^{|S|} \hat{f}(S)^p$

Pf. since $f(x) = \sum_S \hat{f}(S) \chi_S(x)$

$$E[y] = E \left[\sum_{s_1, s_2, \dots, s_{p-1}, s_p} \hat{f}(s_1) \dots \hat{f}(s_{p-1}) \hat{f}(s_p) \chi_{s_1}(x^{(1)}) \chi_{s_2}(x^{(2)}) \dots \chi_{s_{p-1}}(x^{(p-1)}) \chi_{s_p}(x^{(p)}) \right]$$

$$= \sum_{s_1, s_2, \dots, s_{p-1}, s_p} \hat{f}(s_1) \dots \hat{f}(s_p) E \left[\chi_{s_1 \Delta s_p}(x^{(1)}) \chi_{s_2 \Delta s_p}(x^{(2)}) \dots \chi_{s_{p-1} \Delta s_p}(x^{(p-1)}) \chi_{s_p}(x^{(p)}) \right]$$

all independent so expectation of product = product of expectations

note if $s_1 = s_2 = \dots = s_p$ then $s_i \Delta s_p = \emptyset$ so $E[\chi_{s_i \Delta s_p}(x)] = 1$

$$\begin{aligned} \uparrow E[\chi_{s_p}(x)] &= 1 \cdot \Pr[\prod_{i \in S_p} M_i = 1] + (-1) \cdot \Pr[\prod_{i \in S_p} M_i = (-1)] \\ &= \eta^{|S_p|} \end{aligned}$$

proof by induction on $|S_p|$

Else some $s_i \Delta s_p \neq \emptyset$

$\uparrow E[\chi_{s_i \Delta s_p}(x)] = 0$

$$= \sum_S \hat{f}(S)^p \cdot \eta^{|S|}$$



Plan for estimating $\sum \hat{f}(i)^p$:

← all terms = 0 unless $s_1 = s_2 = \dots = s_p = \phi$

Chernoff/Hoeffding
↓
use $O(\frac{1}{\eta^4})$
Samples

1) estimate $E [f(x^{(1)}) f(x^{(2)}) \dots f(x^{(p)})] = \hat{f}(\phi)^p$

to additive $\pm \eta^2$ $(= \sum_{s=0} \eta \hat{f}(s)^p)$

2) estimate $E [f(x^{(1)}) f(x^{(2)}) \dots f(x^{(p)}) f(x^{(1)} \circ \dots \circ x^{(p-1)} \circ y)]$

$= \sum_s \eta^{|s|} \hat{f}(s)^p$

But $(2) - (1) = \sum_{|s| > 0} \underbrace{\eta^{|s|} \hat{f}(s)^p}_{\equiv \gamma}$

← we have estimate to within $\pm \eta^2$

Claim $\frac{\gamma}{\eta}$ is good estimate of $\sum_{|s|=1} \hat{f}(s)^p$

Why?

$$\frac{\sum_{|s|=1} \hat{f}(s)^p}{\eta} = \underbrace{\sum_{|s| > 0} \eta^{|s|} \hat{f}(s)^p}_{\frac{\gamma}{\eta}} - \underbrace{\sum_{|s| > 1} \eta^{|s|} \hat{f}(s)^p}_{\leq \eta^2 \sum_{|s| > 1} |\hat{f}(s)^p|}$$

So it is an additive estimate to within $\pm \eta$!!

⇒ for $p \geq 2$ can estimate $\sum_{|s|=1} \hat{f}(s)^p$ to w/in additive $\pm \eta$

with $O(p/\eta^4)$ queries

$$\frac{1}{\eta} \left\{ \begin{aligned} &\leq \eta^2 \sqrt{\sum_{|s| > 1} \hat{f}(s)^2} \sqrt{\sum_{|s| > 1} (\hat{f}(s)^{p-1})^2} \\ &\leq 1 \text{ by Bochner Parseval} \quad \leq \sum_{|s| > 1} \hat{f}(s)^2 \\ &\quad \text{since } |\hat{f}(s)| \leq 1 \\ &\leq \eta^2 \sqrt{\sum_{|s| > 1} \hat{f}(s)^2} \\ &\leq \eta^2 \end{aligned} \right.$$

2) Estimate max degree 1 Fourier coefficient

Fact if $\forall i \hat{f}(i) < \alpha$

then $\sum_{|s|=1} \hat{f}(s)^4 = \sum_{i=1}^n \hat{f}(i)^4 < \alpha^2 \sum_{i=1}^n \hat{f}(i)^2 \leq \alpha^2$

note naive bound: $\sum_{i=1}^n \hat{f}(i)^4 \leq n \cdot \alpha^4$
 ↓ this is much better!

def γ -regular fctn: $\forall i |\hat{f}(i)| < \gamma$

Thm $\exists O(\frac{1}{\gamma^4})$ -query test st.

1) if $\exists i$ st. $|\hat{f}(i)| \geq \gamma$ then $\Pr[\text{reject}] \geq 3/4$

Fail if not γ -regular

2) if $\forall i |\hat{f}(i)| < \frac{\gamma}{4}$ then $\Pr[\text{accept}] \geq 3/4$

Pass if $\frac{1}{4}$ -regular

PF

Algorithm:

estimate $\sum_i \hat{f}(i)^4$ to win $\pm \frac{\gamma^4}{4}$

accept if estimate $\leq \frac{\gamma^4}{2}$

Behavior:

if $\exists i$ st. $|\hat{f}(i)| \geq \gamma$ then $\sum \hat{f}(i)^4 \geq \gamma^4$

so estimate $\geq \gamma^4 - \frac{\gamma^4}{4} = \frac{3\gamma^4}{4}$

\Rightarrow algorithm accepts (unless estimate has really large error)

if $\forall i |\hat{f}(i)| < \frac{\gamma}{4}$ then $\sum \hat{f}(i)^4 \leq \frac{\gamma^4}{16}$ by note

so estimate $\leq \frac{\gamma^4}{16} + \frac{\gamma^4}{4} = \frac{5\gamma^4}{16} < \frac{\gamma^4}{2}$

so reject

Note this is much faster than learning the big $\hat{f}(i)$'s (which has a $\Omega(\log n)$ lower bound)

estimate max degree \Rightarrow dictator test

def f is a "dictator" if $f = \chi_{\{i\}}$ for some i
 $= X_i$

so $\hat{f}(i) = 1$
 $\hat{f}(j) = 0$ for $j \neq i$

∇ can use const γ to test.

Lots of other dictator tests,
 but this one is close to Hastad's test

Pick $x, y \in \{\pm 1\}^n$

$w \in \{\pm 1\}^n$ s.t. $\Pr[w_i = 1] = \delta$

Accept if $f(x) f(y) f(x \circ y \circ w) = 1$

here:

$y \leftarrow f(x) f(y) f(z) f(x \circ y \circ z \circ \mu)$

noise

Another model...



- Client sends computation on LARGE DATA SET to cloud
- Cloud returns result
- Why should the client trust it?

Can use Interactive proofs, probabilistically checkable proofs, ...
IP PCP

Today:

- Client sends x to cloud
- Cloud writes $f(x)$ + proof that $f(x)$ is approximately correct in stone

Example 1

How many website hits?

Setting: Cloud Prover wants to convince Client

that $\geq k$ clicks were made to Client's website.

Goal: Given proof,
 If # clicks $\geq k$, Client should pass } with prob $\geq 3/4$
 If # clicks $\leq (1-\epsilon)k$, Client should fail

Proof.

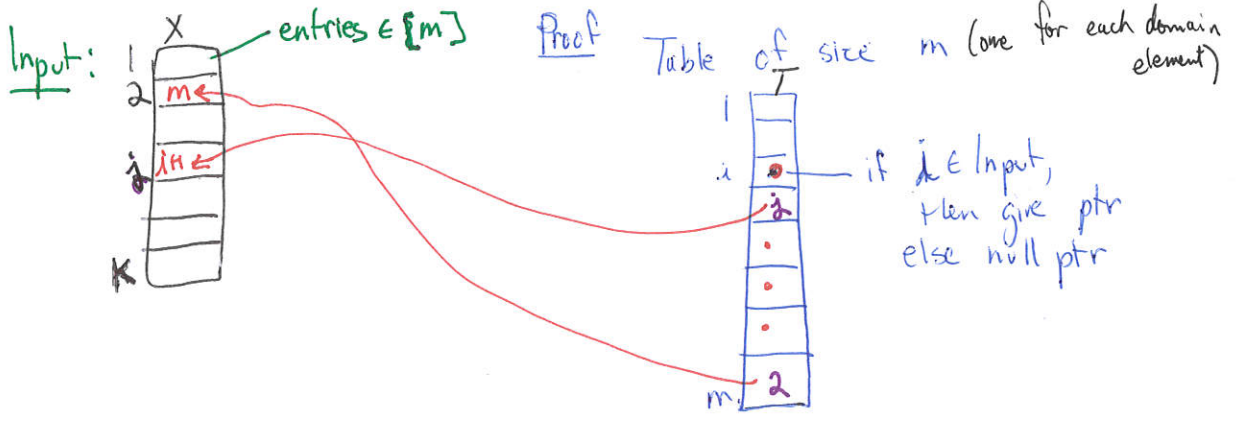
proof = list of k click descriptions
 + ? (duplicate detection) ← assume client can verify any description for "fake" or "real" in $O(1)$ steps

Client's Protocol:

not enough!
 what about duplicates?

- 1) Check $\leq \frac{\epsilon}{2}k$ click descriptions are fake $O(\frac{1}{\epsilon})$ queries
- 2) check $\leq \frac{\epsilon}{2}k$ duplicates $O(\frac{1}{\epsilon})$ queries

How to check for duplicates?



Protocol

Repeat $O(1/\epsilon)$ times
Pick random $j \in [k]$

$l \leftarrow X[j]$

look at locn $T(l)$:

if $T(l)$ contains j Pass + continue
else fail + halt

Proof of correctness:

if all k entries of X distinct, will always PASS

if $\exists \geq \epsilon k$ duplicates, \leftarrow (at least ϵk has appeared earlier in the list)

$\Pr[j \text{ is a duplicate, } \& \text{ not in } T] \geq \frac{\epsilon}{2}$

\Rightarrow Fail

\leftarrow since at least one duplicate isn't pointed to

Example 1'

Does G have cut of size $\geq k$?

proof =

• array A of size n

$A[i] = 0$ for nodes i on "left"
 1 "right"

• array B of k edges that cross cut

Client: • verifies that B contains $\geq (\frac{\epsilon}{2})k$ legal ^{cut} edges wrt A 's cut

• verifies that B contains $\leq \frac{\epsilon}{2}k$ duplicates

Example 2

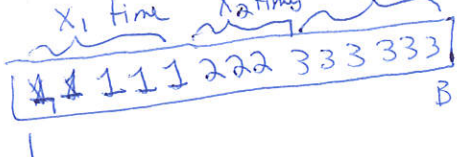
Bin packing

Setting $X_1, \dots, X_n \in [B]$ ← integers
 k bins of size B

Prover I can fit all of X_1, \dots, X_n in k bins of size B

Goal given proof
 if all X_i 's fit in k bins, pass
 if $\leq (1-\epsilon)n$ X_i 's fit in k bins, fail

Proof K arrays $[1..B]$
 in each array, put X_i 's for that bin
 ie. if X_1, X_2, X_3 assigned to bin 1



Protocol:
 Check $\geq (1-\epsilon)k$ X_i 's are allocated space $O(\frac{B}{\epsilon})$ queries