Today
- Local decoding of Hadamard Code [Goldreich-Levin Algorithm]

CS5.318.1
Coding Theory
Lecture 17 (2022-10-31)
Instructor: Prahladh Harsha.

Recall the Hadamard code
$[2^k, k, 2^{k-1}]$-code

a (k-bit) $\longmapsto$ ℓ_a (2^k-bit)

$\ell_a : \{0,1\}^k \to \{0,1\}$
$x \mapsto \langle x, a\rangle = \sum x_i a_i \pmod 2$

Exact Decoding: Can recover using $f(e_1), \ldots f(e_k)$ (or any k-linearly independent locations)

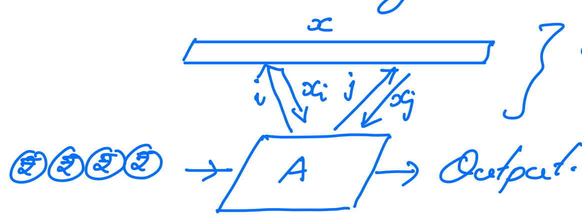
Unique Decoding: Given $f : \{0,1\}^k \to \{0,1\}$ s.t.
$\exists a$, $\Pr_x[f(x) = \ell_a(x)] > \frac{3}{4}$, find a?

List-decoding: Given $f : \{0,1\}^k \to \{0,1\}$ find all a's such that
$$\Pr_x[f(x) = \ell_a(x)] \geq \frac{1}{2} + \varepsilon.$$

If Decoding Alg is allowed to run poly(2^k), can go over all messages (ie 2^k of them) and decode.

Can the decoder run in time poly(message length) rather than poly(codeword length)?

- don't have time to read entire input.
- sub-linear time algorithms



} oracle access to input

Time:

#queries:

random coins:

Typically, sublinear refers to running time but sometime we might just focus on #queries.

Local- sublinear algorithms (i.e. read input locally)

Local Exact Decoding: Query $f(e_1), \ldots, f(e_k)$

Local Unique Decoding:

Problem: Given $f: \{0,1\}^n \to \{0,1\}$ (as an oracle), w/ promise that $\exists\, a \in \{0,1\}^k$ s.t.

$$\Pr_x[f(x) = \ell_a(x)] \geq \frac{3}{4} + \varepsilon$$

find a?

$$a_i = \ell_a(e_i) = \ell_a(e_i + x) - \ell_a(x)$$

$$\cong f(e_i + x) - f_a(x) \quad \text{w/ prob}$$

$$1 - \left(\left(\tfrac{1}{4} - \varepsilon\right) + \left(\tfrac{1}{4} - \varepsilon\right)\right)$$

$$= \tfrac{1}{2} + 2\varepsilon$$

Run majority to boost succ

Pick $x_1 \ldots x_t \in \{0,1\}^k$

$$a_i \leftarrow \operatorname{maj}_j \{f(e_i + x_j) - f(x_j)\}$$

$GLW^f(\cdot)$ (Goldreich Levin Algorithm)

- Pick $x_1 \ldots x_t \in \{0,1\}^k$ at random $t = O\left(\frac{\log k}{\varepsilon^2}\right)$
- For $i \leftarrow 1$ to k

  $a_i \leftarrow \operatorname{maj}_j \{f(e_i + x_j) - f(x_j)\}$
- Output $(a_1 \ldots a_k)$.

Chernoff Bound: $X_1 \ldots$ - be independent 0/1 r.vs

$$\Pr\left[\sum X_i \geq \mathbb{E} X_i + \varepsilon t\right] \leq \exp(-2\varepsilon^2 t).$$

Choose t s.t. $\exp(-2\varepsilon^2 t) \leq \frac{1}{k^2}$

ie, $t = \frac{2 \log k}{2\varepsilon^2} = \frac{\log k}{\varepsilon^2}$

What about list-decoding?

Prob: Given $f:\{0,1\}^k \to \{0,1\}$ s.t. there exists an $a \in \{0,1\}^k$

$$\Pr[f(x) = \ell_a(x)] \geq \frac{1}{2} + \varepsilon$$

output a list that contains a?

Idea: Assume (guess) the value of $\ell_a(r_j)$, $j = 1 .. t$ and obtain $\ell_a(e_i)$ w/ high probability.

GL (first attempt)$^f(\cdot)$

0. List $\leftarrow \phi$
1. Pick $r_1 \dots r_t \in \{0,1\}^k$ where $t = O(\frac{1}{\varepsilon^2})$
2. For $\bar{b} = (b_1 \dots b_t) \in \{0,1\}^t$
   (a) Set $f_{\bar{b},\bar{r}}(x) = \text{maj}_j\{f(x+r_j) - b_j\}$
   (b) Apply GLW to $f_{\bar{b},\bar{r}}$ to obtain a
   (c) Add a to List
3. Output List.

Fix an a s.t. $\Pr[f(x) = \ell(x)] \geq \frac{1}{2} + \varepsilon$.

Obs.

$$\Pr_{\bar{r}, \bar{x}}\left[f_{\ell_a(\bar{r}),\bar{r}}(x) = \ell_a(x)\right] \geq 1 - \exp(-\varepsilon^2 t)$$

$$\geq \frac{31}{32} \quad \text{(by setting } t = O(\tfrac{1}{\varepsilon^2})\text{)}$$

$\#$queries $= O\left(\frac{k \log k}{\varepsilon^2}\right) \cdot 2^{1/\varepsilon^2}$ | list-size $= 2^{1/\varepsilon^2}$

running time $= O(k^2 \log k) \cdot 2^{1/\varepsilon^2} \cdot \frac{1}{\varepsilon^2}$

Analysis: By setting $t = C/\varepsilon^2$.

$$\Pr_{\bar{r}, x}\left[f_{l_a(\bar{r}), \bar{r}}(x) = l_a(x) \right] \geq \frac{31}{32}$$

By Markov,

$$\Pr_{\bar{r}}\left[\Pr_{x}\left[f_{l_a(\bar{r}), \bar{r}}(x) = l_a(x) \right] \geq \frac{7}{8} \right] \geq \frac{3}{4}$$

Reducing list-size from $\exp(\frac{t}{\varepsilon})$ to $\mathrm{poly}(\frac{t}{\varepsilon})$

Idea: Choose $r_1 \ldots r_t$ such that they form a subspace. Hence to guess values of l_a at $\bar{r}$, sufficient to guess on the basis.

$GL^f(\cdot)$:

0. List $\leftarrow \emptyset$
1. Pick $r_1 \ldots r_t \leftarrow \{0,1\}^k$ where $t = O(\log 1/\varepsilon^2)$
2. For each $\bar{b} = (b_1, \ldots b_t) \in \{0,1\}^t$

    (a) For each $S \subseteq \{1, 2, \ldots t\}$ $(S \neq \phi)$

    $$b_S := \sum_{i \in S} b_i$$

    (b). Define $\tilde{f}_{\bar{r}, \bar{b}}(x) = \mathrm{maj}_S \left\{ f(x + r_S) - b_S \right\}$

(c) Apply GLW on $\tilde{f}_{\bar{x},\bar{b}}$ to get a

(d) Add a to LIST

3. Output LIST.

queries: $O(k \log k) \cdot 2^t \cdot 2^t = O\left(\frac{k \log k}{\varepsilon^4}\right)$

running time $: O(k^2 \log k / \varepsilon^4)$

List size $= 2^t = O(1/\varepsilon^2)$

Analysis: Fix a s.t. $\Pr[f(x) = \ell_a(x)] \geq 1/2 + \varepsilon$

By previous analysis, sufficient to prove

$$\Pr_{\bar{x}, x}\left[\tilde{f}_{\bar{x}, \ell_a(\bar{x})}(x) = \ell_a(x)\right] \geq \frac{31}{32} \quad \ldots (*)$$

Since $\Pr_{\bar{x}}\left[\Pr_{x}\left[f_{\bar{x}, \ell_a(\bar{x})}(x) = \ell_a(x)\right] \geq \frac{7}{8}\right] \geq 3/4$

& then GLW will output a correctly).

Rewriting (*)

$$\Pr_{x_1..x_t, x}\left[\ell_a(x) = \operatorname{maj}_S\{f(x_S + x) - \ell_a(x_S)\}\right]$$

- Can't apply Chernoff bound as the x_S' are not independent

- they are only pairwise independent (use Chebyshev instead)

$\frac{1}{2}+\varepsilon \longrightarrow 1-\delta$: Chernoff $O\left(\frac{1}{\varepsilon^2}\log\frac{1}{\delta}\right)$

Chebyshev $O\left(\frac{1}{\varepsilon^2}\frac{1}{\delta}\right)$

$\frac{1}{2}+\varepsilon \longrightarrow \frac{31}{32}$: Since δ is constant, ok to use Chebyshev. ✓

$\boxtimes$