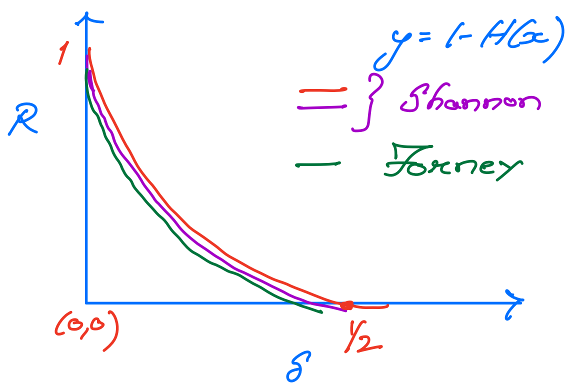
Today
- List decoding
  * Combinatorics
  * Johnson Radius

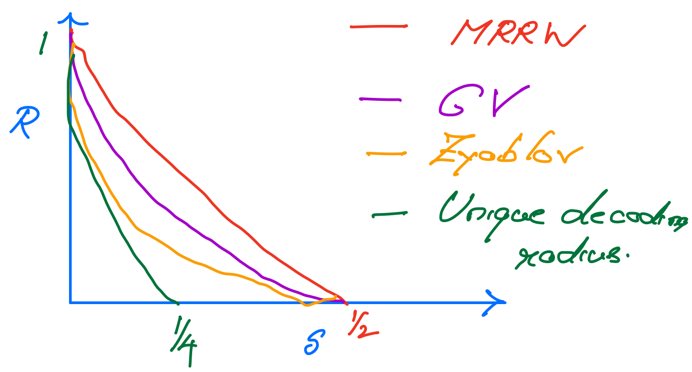
CS5.318.1
Coding Theory
Lecture 15 (2022-10-21)
Instructor: Prahladh Harsha.

Where we are (w.r.t Rate vs errors)?

Shannon Model (Random Errors)

Hamming Model (worst-case errors)





Hamming's bound: $2e < d$

Relax the algorithmic challenge to return a list of at most L codewords instead of just 1.

(ρ, L)-list-decodable : $\rho \in (0,1)$, $L \in \mathbb{Z}_{>0}$

$\mathcal{C} \subseteq \Sigma^n$ is (ρ, L)-list decodable if there exists a decoder $D : \Sigma^n \to \binom{\mathcal{C}}{\leq L}$ s.t.

$$\forall c \in \mathcal{C},\ \eta \in Ball(0, \rho n)$$
$$c \in D(c+\eta)$$

Equivalently:

$\mathcal{C}$ is (ρ, L)-list decodable if
$\forall y \in \Sigma^n$, $\quad |Ball(y, \rho n) \cap \mathcal{C}| \le L$.

Remarks

(1) Not a computational defn, but combinatorial.

(2) Is it a reasonable?

(i) random errors, list-size is typically 1

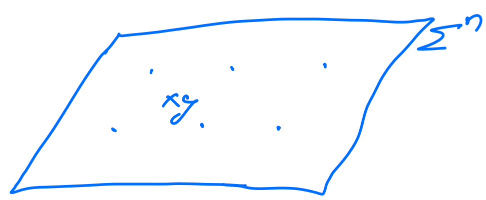
(ii) Side information to disambiguate from the list

(iii) Cryptographically bounded channels.

Limits on Rate of (ρ, L)-list-decodable codes

<u>Thm</u>: Suppose $\mathcal{C} \subseteq \Sigma^n$ is (ρ, L)-list-decodable &
$R \geq 1 - H_q(\rho) + \varepsilon$ where $q = |\Sigma|$
then $L \geq 2^{\Omega(\varepsilon n)}$. if $\rho \le 1 - \frac{1}{q}$

<u>Pf</u>: Choose $y \in_R \Sigma^n$



Fix $c \in \mathcal{C}$.

$$\Pr_y[c \in Ball(y, \rho n)] = \Pr_y[y \in Ball(c, \rho n)]$$
$$= \frac{Vol_q(n, \rho n)}{q^n} \quad \ldots \quad (*)$$
$$\geq q^{-n(1 - H_q(\rho)) - o(n)}$$

$$\mathbb{E}_y[|\mathcal{C} \cap Ball(y, \rho n)|] = \sum_{c \in \mathcal{C}} \Pr[c \in Ball(y, \rho n)]$$
$$= q^{Rn} \cdot (*)$$
$$\geq q^{-n(1 - H_q(\rho) - R) - o(n)}$$
$$\geq q^{\Omega(\epsilon n)} \quad \text{if } R \geq 1 - H_q(\rho) + \epsilon.$$

Hence, $\exists y$ (in fact a random y) has exp. many codewords in a ρn-ball around y.

$\boxtimes$

Theorem: Let $L \in \mathbb{Z}_{>0}$, then there exist (ρ, L)-list-decodable codes with rate $R \geq 1 - H_q(\rho) - \frac{1}{L}$

Proof: Pick $\mathcal{C}$ at random

For each $i = 1 \ldots q^{Rn}$, pick $c_i \leftarrow_R \Sigma^n$ independently.

BAD Event: $\exists y \in \Sigma^n$: $(L+1)$ codewords $c^{(0)}, c^{(1)}, \ldots, c^{(L)}$
s.t. $c^{(j)} \in \mathrm{Ball}(y, \rho n)$ $\forall\, 0 \le j \le L$.

Fix y.

$$\Pr_{c^{(j)}}\left[c^{(j)} \in \mathrm{Ball}(y, \rho n)\right] = \frac{\mathrm{Vol}_q(n, \rho n)}{q^n} \le q^{-n(1-H_q(\rho))}$$

$$\Pr_{c^{(0)} \ldots c^{(L)}}\Big[\underbrace{\bigwedge_{j=0}^{L}\left[c^{(j)} \in \mathrm{Ball}(y, \rho n)\right]}_{(**)}\Big] \le q^{-n(L+1)(1-H_q(\rho))}$$

$$\Pr[\text{BAD event}] = \Pr[\exists y \; \exists c^{(0)} \ldots c^{(L)}, (**) \text{ holds}]$$

$$\le q^n \binom{q^{Rn}}{L+1} \cdot q^{-n(L+1)(1-H_q(\rho))}$$

$$\le q^{n[1 + R(L+1) - (L+1)(1-H_q(\rho))]}$$

$$\le q^{n[1-1]} = 1 \qquad \left(\text{Setting } R \le 1 - H_q(\rho) - \frac{1}{L+1}\right)$$

$\boxtimes$

Hence, impossibility & achievability curves for (ρ, L)-list-decodable codes match.

Next Question: (1) Are there explicit codes which are (ρ, L)-list-decodable

(2) Can the list-decoders made efficient?

Recall:

Johnson Radius: $J_2(\delta) = \frac{1}{2}\left(1-\sqrt{1-2\delta}\right)$

<u>Lemma</u>: Given any $\delta \in (0, \frac{1}{2})$ & $C = (n, Rn, \delta n)_2$-code then for $\tau = J_2(\delta)$

$$\forall y \in \{0,1\}^n, \quad |Ball(y, \tau n) \cap C| < n+1$$

q-ary version: $J_q(\delta) = \left(1-\frac{1}{q}\right)\left(1-\sqrt{1-\frac{q\delta}{q-1}}\right)$

Alphabet-independent version:

$$J_q(\delta) = \left(1-\frac{1}{q}\right)\left(1-\sqrt{1-\frac{q\delta}{q-1}}\right) \geq 1-\sqrt{1-\delta} =: J(\delta)$$

Now, give an alternate (combinatorial) proof of the alphabet-independent Johnson bound.

<u>Lemma</u>: C - $(n, Rn, \delta n)_q$-code then for $\tau \leq 1-\sqrt{1-\delta}$

$$\forall y \in [q]^n, \quad |Ball(y, \tau n) \cap C| < \tau(1-\tau)n+1$$

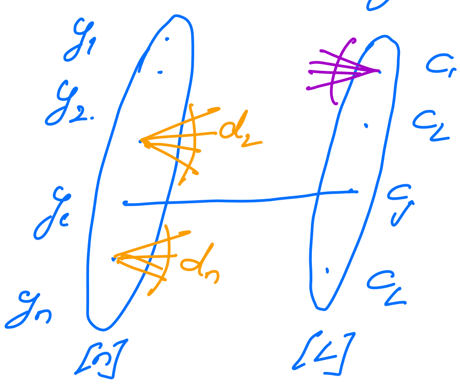
$\boxtimes$

Proof (Jaikumar Radhakrishnan).

Suppose $\exists y \in [q]^n$ & $c_1 \ldots c_L \in C$ s.t.

$$y \in Ball(c_i, \tau n) \quad \forall\ 1 \leq i \leq L.$$

Consider the following bipartite graph.



(i,j) exists

$\Updownarrow$

$(C_j)_i = y_i$

Every right vertex has degree at least $n(1-\tau)$

$$d_i = \#\{j \in L \mid (C_j)_i = y_i\}.$$

$$\bar{d} = \frac{\sum d_i}{n}$$

Counting #edges from both sides

$$\bar{d} \cdot n \geq n \cdot (1-\tau) \cdot L$$

$$\Rightarrow \bar{d} \geq (1-\tau) \cdot L$$

$i \in [n]$

$$\Pr_{j_1 \neq j_2}\left[i \text{ - adj to both } j_1 \,\&\, j_2 \right] = \frac{\binom{d_i}{2}}{\binom{L}{2}}$$

$$\mathbb{E}_{j_1 \neq j_2}\left[\#\text{common nbrs of } j_1 \,\&\, j_2 \right] = \sum_{i=1}^{n} \frac{\binom{d_i}{2}}{\binom{L}{2}}$$

$$\geq \frac{n \cdot \binom{\bar{d}}{2}}{\binom{L}{2}}$$

On the other hand for every $j_1 \neq j_2$

$$\#\text{common nbr} \leq n(1-\delta) - 1$$

Putting these two together.

$$n\binom{\bar{d}}{2} \leq \binom{L}{2}\left(n(1-\delta)-1\right)$$

$$n\frac{((1-\tau)L)((1-\tau)L-1)}{2} \leq \frac{L(L-1)}{2}\left(n(1-\delta)-1\right) \quad \left(\text{since } \bar{d} \geq (1-\tau)L\right)$$

$$(1-\tau)^2 Ln - (1-\tau)n \leq (L-1)\left(n(1-\delta)-1\right)$$

$$(1-\tau)^2 Ln - (1-\tau)n \leq (L-1)\left(n(1-\tau)^2-1\right) \quad \left(\tau \leq 1-\sqrt{1-\delta}\right)$$

$$\Rightarrow L \leq (1-\tau)n - (1-\tau)^2 n + 1$$

$$= (1-\tau)\tau \cdot n + 1$$

$$= n/4 + 1$$

$\boxtimes$