Today

- Multiplicity Codes - IV

(Multivariate Setting)

Locally Decodable Codes.

CSS.318.1

Coding Theory

Lecture 26 (2022-12-5)

Instructor: Prahladh Harsha.

Recall:

Question: Do there exist locally decodable codes

- recover from constant fraction of errors
- query complexity - sublinear
- rate approaching 1. ?

What do Reed-Muller codes achieve?

RM codes

$RM_q(m,d): \mathbb{F}_q[x_1, \ldots, x_m] \longrightarrow \mathbb{F}^{(\mathbb{F}^m)}$

$P \longmapsto (P(a))_{a \in \mathbb{F}^m}$

Distance: $1 - \frac{d}{q} \triangleq \delta$ (i.e., $d = (1-\delta)q$)

Rate: $\frac{\log_\Sigma |C|}{n} = \frac{k}{n} = \frac{\binom{d+m}{m}}{q^m} \simeq \frac{d^m}{m!} (1 \pm o(1)) \frac{1}{q^m}$

(for constant m).

$= \frac{(1-\delta)^m}{m!}(1+o(1)) < 1/2$ (since $m \geq 2$)

Conclusion: RM codes even though locally decodable (in fact even correctable) using $q = O(n^{1/m})$ queries

However, $R < 1/2$ (and goes to 0 as $m \to \infty$)

Today: See how multivariate multiplicity codes get around this rate $1/2$ barrier.

Multivariate Multiplicity Codes.

$MMULT_q(m, s, d) : \mathbb{F}_q[X_1 .. X_m] \to (\Sigma)^{q^m}$

where $\Sigma = \mathbb{F}^{\binom{s+m-1}{m}}$

$P \longmapsto (P^{<s}(a))_{a \in \mathbb{F}^m}$

$m = 2$; $s = 2$

Rate $= \frac{\log_\Sigma |C|}{n} = \frac{\log_{(q^2)} q^{\binom{d+2}{2}}}{q^2} = \frac{1}{3}\frac{\binom{d+2}{2}}{q^2} \approx \frac{1}{6}\frac{d^2}{q^2}$

Distance: $1 - d/sq$ (Why?).

Distance

Schwartz-Zippel Lemma: If $P \in \mathbb{F}_{\leq d}[x_1 .. x_m]$ & $P \neq 0$, then for any $S \subseteq \mathbb{F}$

$$\mathop{\mathbb{E}}_{a \leftarrow S^m}[\mathbb{1}_{P(a)=0}] = \Pr_{a \leftarrow S^m}[P(a)=0] \leq \frac{d}{|S|}$$

Extension to multiplicities:

$$\mathop{\mathbb{E}}_{a \leftarrow S^m}[\text{mult}(P,a)] \leq \frac{d}{|S|}.$$

Returning to distance of MMult$(m=2, s=2)$

$P \neq 0 \Rightarrow \Pr_{a \leftarrow \mathbb{F}^2}[P^{(<2)}(a) = 0] = \mu$ then $\mu \leq \frac{d}{2q}$.

If we want dist $= \delta$; $\frac{d}{2q} = 1-\delta$

i.e., $d = 2(1-\delta)q$

—

Distance $= \delta$ (i.e., $d = 2(1-\delta)q$)

Rate $= R \cong \frac{1}{6}\frac{d^2}{q^2} = \frac{1}{6}4(1-\delta)^2 = \frac{2}{3}(1-\delta)^2$

Plan for the remaining lecture

① Proof of Multiplicity Schwartz-Zippel Lemma

(2) Local Corrector for MMult.

Summarizing, discussion on distance

$$P \neq 0 \Rightarrow \Pr_{a \leftarrow S^m}\left[P^{(<k)}(a) = 0\right] \leq \frac{d}{k|S|}$$

Distance of $MMult_q(m, k, d) \geq 1 - \frac{d}{q}$.

Multiplicity Schwartz-Zippel Lemma (MSZ)

Facts (about Hasse Derivatives)

(1) $P \in \mathbb{F}[X_1 \ldots X_m]$; $a \in \mathbb{F}^m$

$\forall\, e = (e_1 \ldots e_m) \in \mathbb{Z}_{\geq 0}^m$

$$mult(P^{(e)}, a) \geq mult(P, a) - wt(e)$$

Pf: $\left(P^{(e)}\right)^{(f)} = \binom{e+f}{e} P^{(e+f)}$

For all f s.t $wt(f) < mult(P,a) - wt(e)$

$\left(P^{(e)}\right)^{(f)}(a) = 0$ since $P^{(e+f)}(a) = 0$

(since $wt(e+f) = wt(e) + wt(f) < m$)

□

(2) $P \in \mathbb{F}[X_1, \ldots X_m]$

$Q \in \mathbb{F}[X_1, \ldots X_m]^m$, $a \in \mathbb{F}^m$

$$\text{mult}(P \circ Q, a) \geq \text{mult}(P, Q(a)) \cdot \text{mult}(Q - Q(a), a)$$

where $\text{mult}((R_1, \ldots R_k), a) = \min_{j \in [k]} \text{mult}(R_j, a)$

In particular,

$$\text{mult}(P \circ Q, a) \geq \text{mult}(P, Q(a))$$

Pf: $P \circ Q(a+Z) = P(Q(a+Z))$

$= P(Q(a) + h(Z))$ where $\deg h \geq \text{mult}(Q - Q(a), a)$

$= P \cdot Q(a) + \underbrace{\sum \ldots}$

every term will have deg $\geq \text{mult}(P, Q(a)) \cdot \text{mult}(Q - Q(a), a)$.

□

Proof of MSZ Lemma:

MSZ: $\mathop{\mathbb{E}}_{(a_1 \ldots a_m) \leftarrow S^m} [\text{mult}(P, a)] \leq \frac{d}{|S|}$ if $P \neq 0$.

Proof by induction on dim m.

$m = 1$. $\mathop{\mathbb{E}}_{a \leftarrow S}[\text{mult}(P, a)] \le \frac{d}{|S|}$ ✓

Assume it is true for $m = 1 \ldots m-1$.

$$P(x_1 \ldots x_m) = \sum_{j=0}^{t} P_j(x_1 \ldots x_{m-1}) x_m^j$$ for some $0 \le j \le d$

- $\deg(P_j) \le d - j$
- $P_t \neq 0$

Claim: For each $(a_1 \ldots a_{m-1}) \in S^{m-1}$

$$\mathop{\mathbb{E}}_{a_m \leftarrow S}[\text{mult}(P, (a_1 \ldots a_m))] \le \text{mult}(P_t, (a_1 \ldots a_{m-1})) + \frac{t}{|S|}$$

Given claim

$$\mathop{\mathbb{E}}_{a \leftarrow S^m}[\text{mult}(P, a)] \le \mathop{\mathbb{E}}_{a_{-m} \leftarrow S^{m-1}}\left[\text{mult}(P_t, a_{-m})\right] + \frac{t}{|S|}$$

$$\le \frac{d-t}{|S|} + \frac{t}{|S|} = \frac{d}{|S|}.$$

Proof of Claim:

$r = \text{mult}(P_t, (a_1 \ldots a_{m-1}))$

$\forall f$, $\text{wt}(f) < r$, $P_t^{(f)}(a_1 \ldots a_{m-1}) = 0$

$\exists e$, $\text{wt}(e) = r$ $P_t^{(e)}(a_1 \ldots a_{m-1}) \neq 0$.

$$P^{(e,0)}(X_1 \ldots X_m) = \sum_{j=0}^{t} P_j^{(e)}(X_1 \ldots X_{m-1}) X_m^j$$

· (Note: $P^{(e,0)}(a_1 \ldots a_{m-1}, X_m)$ - univariate poly in X_m of exact degree t)

$$\text{mult}(P(X_1 \ldots X_m), (a_1 \ldots a_m)) \leq \text{wt}(e,0) + \text{mult}(P^{(e,0)}(X_1 \ldots X_m), (a_1 \ldots a_m))$$

What we know is

$$\text{mult}(P^{(e,0)}(a_1 \ldots a_{m-1}, X_m), a_m) \leq t$$

But $\text{mult}(P^{(e,0)}(a_1 \ldots a_{m-1}, X_m), a_m) \geq \text{mult}(P^{(e,0)}(X_1 \ldots X_m), (a_1 \ldots a_m)) \cdot m'$ where $m' \geq 1$.

Hence

$$\text{mult}(P(X_1 \ldots X_m), (a_1 \ldots a_m)) \leq \text{mult}(P_e, (a_1 \ldots a_{m-1})) + t.$$

Claim is obtained by taking an average over t. ☆.

Local Decoder for MMult.

Simple Case: $m = 2$ (bivariate)

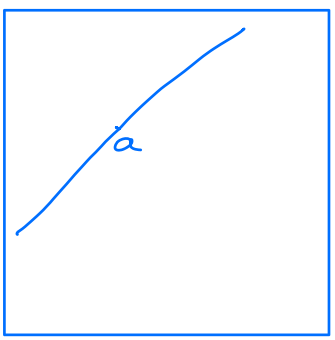
$s = 2$ (i.e., eval + first order derivatives)

At each point $a = (a_1, a_2) \in \mathbb{F}^2$

$$\left(P(a),\ P^{(1,0)}(a),\ P^{(0,1)}(a)\right)$$

Received Word: $w : \mathbb{F}^2 \to (\mathbb{F}^3)$

$$\Delta\left(w, \mathrm{Enc}(P)\right) < \frac{\delta}{1000} \quad \text{where } \delta = \text{distance of code.}$$



$\mathbb{F}^2$

$w : \mathbb{F}^2 \to (\mathbb{F}^3)$

$$\ell_{a,b} = \{a + bt \mid t \in \mathbb{F}\}$$

$$b = (b_1, b_2) \in \mathbb{F}^2$$

$$Q_{a,b}(t) = P(a + bt)$$

$$Q_{a,b}(t) = P(a + bt)$$

$$Q_{a,b}^{(1)}(t) = b_1 P^{(1,0)}(a + bt) + b_2 P^{(0,1)}(a + bt)$$

Apply univariate mult decoder on line $\ell_{a,b}$

$$Q_{a,b}(0) = P(a)$$

$$Q_{a,b}^{(1)}(0) = b_1 P^{(1,0)}(a) + b_2 P^{(0,1)}(a).$$

At this point, we have decoded

$$\left(P(a),\ b_1 P^{(1,0)}(a) + b_2 P^{(0,1)}(a)\right).$$

Use another line to extract $P^{(1,0)}(a)$, $P^{(0,1)}(a)$.

Parameters of Code.

$m=2$; $\delta \geq 2$.

Distance $= \delta$: $d = 2(1-\delta)q$. | $d = \delta(1-\delta)q$.

Rate $= \dfrac{\log_{\Sigma} |C|}{q^m} = \dfrac{1}{3}\dfrac{\binom{d+2}{2}}{q^2} \approx \dfrac{1}{6}\dfrac{d^2}{q^2}$

$= \dfrac{2}{3}(1-\delta)^2$ | $R = \dfrac{2\binom{d+2}{2}}{\delta(\delta+1)\,q^2} \cong \dfrac{\delta}{\delta+1}(1-\delta)^2$

Query Complexity $= 2q = O(\sqrt{n})$ | Query $= O(\delta q)$ $= O(\delta\sqrt{n})$

Concl: Codes (w Rate $> 1/2$, infact $\to 1$) which are LCC w/ sub-linear # queries.

Ref: Survey by Kopparty "Remarks on Multiplicity Codes"