3/2/2017

CS 229 - Lecture 12

FOLDED REED-SOLOMON CODES

~~Intuition~~
Main Result today:

- Over poly(n)-sized alphabets, can get codes that have rate $\to 1-\delta$, list decoding from δ-fraction errors.
  [Not $\delta/2$ ~ which Reed-Solomon does with unique decoding]

- More precisely
  $\forall \epsilon > 0, \delta > 0$ $\exists$ poly $p()$ s.t. $\forall n$ $\exists$ codes of length n, over Σ with $|\Sigma| = p(n)$, of rate $1-\delta-\epsilon$, list-decoding (in poly(n) time) from δ-fraction errors
  [$|L| \leq p(n)$]

Intuition / Motivation / Context

Welch-Berlekamp / List-decoding algorithm illustrate power of two variables.

" $\forall$ subset of n points in plane $\exists$ deg $2\sqrt{n}$ polynomials vanishing on those points "

— x —

Things would be better if there were a <u>third</u> variable

$$2\sqrt{n} \rightsquigarrow 3n^{1/3}$$

How can we arrange this?

— x —

$$X \to \alpha_1 \quad \alpha_2 \quad \dots \quad \alpha_n$$

$$Y \to y_1 \quad y_2 \quad \dots \quad y_n$$

$$Z \to z_1 \quad \dots \quad z_n = ?$$

Maybe message = two polynomials (P_1, P_2)
Encoding = evaluation pairs $((P_1(\alpha_i), P_2(\alpha_i)))_i^n$

1. But this can't help increase list-decoding radius ....
2. One-poly-reconstruction $\leq$ Two-poly-reconstruction

[Set $P_2(\alpha_i) = 0 \quad \forall i$]

3. Even worse: $Q(\ )$ might give us information only about P_1, or only about P_2 or only about relationship between P_1 & P_2

Examples

$$\left.\begin{matrix} x_1 & \cdots & x_n \\ y_1 & \cdots & y_n \\ 0 & \cdots & 0 \end{matrix}\right\} \Rightarrow Q(x,y,z) = z$$

$$\left.\begin{matrix} x_1 & \cdots & x_n \\ 0 & 0 & 0\ 0 \\ z_1 & \cdots & z_n \end{matrix}\right\} \Rightarrow Q(x,y,z) = y$$

$$\left.\begin{matrix} x_1 & \cdots & x_n \\ y_1 & \cdots & y_n \\ y_1 & \cdots & y_n \end{matrix}\right\} \quad Q(x,y,z) = z-y$$

How to avoid all these "trivial" statements.

Idea 1: [Parvaresh-Vardy 2005]

Relate P_1 & P_2 so that

(i) Any information about one of them yields both.

(ii) No ~~relat~~ low degree relation exists among them.

$$\text{E.g.} \quad P_2 = P_1^N \pmod{f(x)}$$

Good News: No immediate obstacles; indeed can prove that degree k poly recoverable from $O(k^{2/3} N^{1/3})$ agreement.

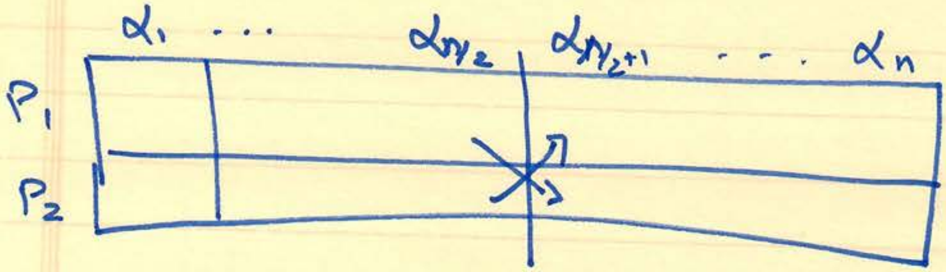
Bad News: Rate $\leq \frac{1}{2}$!

[Encoding P_1 twice → first as P_1

→ then as $P_1^D \pmod{f(x)}$

Idea 2: [Guruswami-Rudra 2006]

Insist on an (impossible) miracle



insist that $P_1(\alpha_i) = P_2(\alpha_{n-i})$

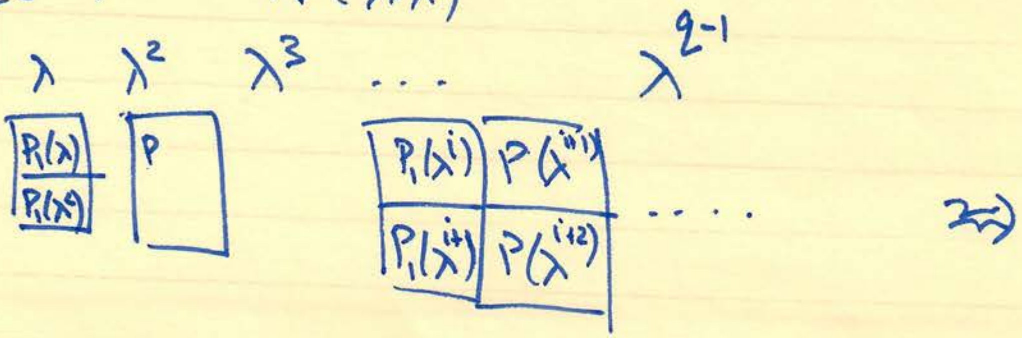
[Will lead to "trivial" relationship between P_1 & P_2, but ignore this]

~~Also~~ Good News: Rate ↑ by 2.

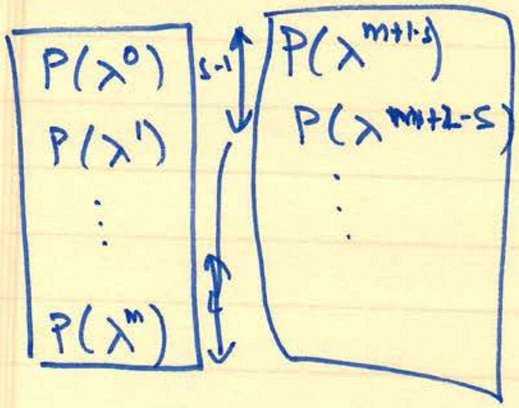
if second half = first half then just throw it away!

Less Impossible miracle

$P_2(x) = P_1(\lambda . x)$ $\qquad \lambda \in \mathbb{F}_q$



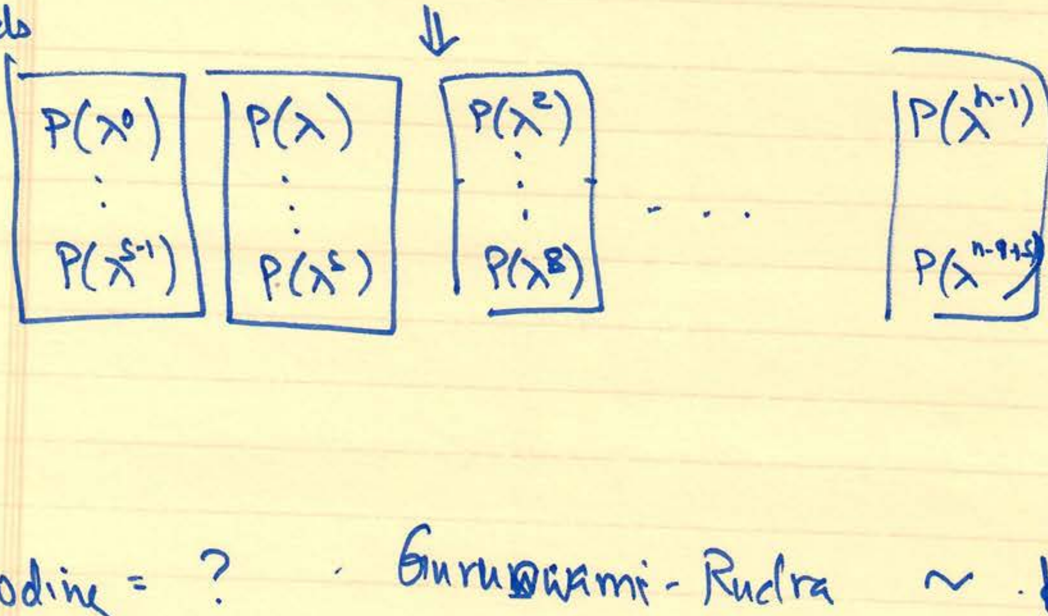
(m,s) folded - RS-code



Rate $= \left(1 - \frac{s}{m}\right)\left(\frac{k}{n}\right)$ $\qquad n = q-1$

Yields

fraction of errors preserved.



Decoding = ? · Guruswami - Rudra $\sim k^{\frac{s-1}{s}} n^{1/s}$ ~~errors~~ agreements

· Guruswami $\sim \frac{n}{s} + \left(\frac{s-1}{s}\right) \cdot k$ agreem

[simpler]

Decoding Algorithm $[s=2]$

(1) Find $A(x)$, $B(x)$, $C(x)$ $\qquad \deg A \le \frac{n}{3} + \frac{2k}{3}$

$$\deg \begin{matrix} B \\ C \end{matrix} \le \frac{n}{3} - \frac{k}{3}$$

s.t. $A(x_i) + y_i B(x_i) + z_i C(x_i) = 0 \quad \forall i$

$$(A, B, C) \not\equiv (0,0,0)$$

$\boxed{\text{Usual Claim: Exists (counting) \& Can be found (Linear system)}}$

(2) Challenge: Find all $P(x) \in \mathbb{F}_q^{<k}[x]$ s.t.

$$A(x) + p(x) \cdot B(x) + \underbrace{p(\lambda \cdot x)}_{\text{Ugly}} C(x) = 0$$

Ugly [But not linear system]

Instead will find p s.t.

$$A(x) + p(x) \cdot B(x) + p(\lambda \cdot x) C(x) = 0 \pmod{X^{q-1} - \lambda}$$

Claim 1:

$$* \quad P(\lambda X) = P(X)^q \pmod{X^{q-1} - \lambda}$$

Claim 2: $X^{q-1} - \lambda$ is irreducible if λ primitive in $\mathbb{F}_q$.

Claim 3:

$\#\ p$ s.t. $A + p \cdot B + p^q \cdot C = 0$

in $\mathbb{K} = \mathbb{F}_q[X] / (X^{q-1} - \lambda)$

is at most q.

[So almost done: # solutions to (2) is at most q; can we find them]

Claim 4: solutions can be found by factoring over $\mathbb{K}$;

Claim 5: Solutions can be found by solving linear system over $\mathbb{F}_q$.

Larger s :

- No different.
- set $P_i(x) = P(\lambda^{i-1} x)$

$$P_i(x) = P(x)^{q^i} \pmod{x^{q-1} - \lambda}$$

$\Rightarrow$ So Step 1 : find $A_0, A_1 \dots A_s$ s.t.

$$A_0(x_i) + \sum y_{ij} A_j(x_i) = 0 \quad \forall i$$

$$\deg A_0 \le \frac{n}{s+1} + \frac{s}{s+1} k$$

$$\deg A_i \le \frac{n}{s+1} - \frac{k}{s+1}$$

Step 2 : Find p s.t.

$$A_0(x) + \sum_{j=1}^{s} p(\lambda^{j-1} x) \cdot A_j(x) = 0 \qquad - \ (\star)$$

List size Bounded by q^s.

since solution to $(\star)$ is also solution to

$$A_0 + p^j \cdot A_j = 0 \quad \text{in} \quad K = \mathbb{F}_q[x] / x^{q-1} - \lambda$$

State of art on list size

[Guruswami]: Use subset ~~of~~ $M \subseteq \mathbb{F}_q^k$ such that $|M \cap A| \le$ ~~poly(s)~~ ~~small~~ small $\forall$ s-dimensional affine subspace in $\mathbb{F}_q^k$.

Encode only messages in M. [Reduces rate but by $1-o(1)$ factor]

[Guruswami]: Random subset M is subspace evasive.

$\Rightarrow$ list size "small"; but run time $\gtrsim q^s$.

[Dvir Lovett]: Explicit construction of M.

$\Rightarrow$ list size "SMALL"; but run time $= o(q^s)$.

Conclusions

- Can get to rate $1-\delta$ for δ-fraction errors.

- Getting right list size, alphabet, explicitness ... still work in progress.

- Reducing Alphabet size - concatenation ++

  [Now we really need to learn graph theory.]

- Linearity: Very subtle concept?
  - GR codes are not linear !!
  - I don't know any linear capacity achieving codes.
  - [Dvir-Lovett] construction is linear $(\mathbb{F}_2)$ subspace that is subspace evasive !! how? $(\mathbb{F}_{2^k})$

See green