5TO8 LECTURE 15 — 4/2/2008

Note Title — 3/31/2008

TODAY (+ NEXT LECTURE)

- [PARVARESH - VARDY '05]
  + [GURUSWAMI - RUDRA '06]

"Rate - Optimal, Polytime - list - decodable Codes (over large alphabets)"

What

- Reed-Solomon codes + list-decoder, gives codes of rate $(1-p)^2$ correcting p fraction errors, over alphabet of size $q(n) = n$.

(How? Set $k = (1-p)^2 n$; RS decoder corrects $1 - \sqrt{k/n}$ fraction errors.

$$1 - \sqrt{\frac{k}{n}} = 1 - \sqrt{(1-p)^2} = p.)$$

- But is this the best we can do?
- <u>Existentially</u>: There exist codes of rate $1-p-\epsilon$ over alphabet of size $f(\epsilon)$, that are (p, poly)-list-decodable
- <u>Constructively</u>: No "explicit" codes known till 2006.
- [PV + GR] Explicit codes + polytime list decoder, with $q = q(n,\epsilon) = n^{f(\epsilon)}$.

Folded Reed-Solomon Codes [GR '06]

WARNING: NOT ALPHABET SIZE!

- Let $n+1 = q$ - prime power
- Let $c = c(\epsilon)$ be a constant (ind. of n)
- Let $\alpha \in \mathbb{F}_q$ be a primitive element

  (i.e. $\mathbb{F}_q^* = \{\alpha, \alpha^2, \alpha^3, \ldots, \alpha^{q-1}\}$)

- $FRS_{n,k,\epsilon} : \Sigma^{k'} \to \Sigma^{n'}$

  where $n' = \frac{n}{c}$; $k' = \frac{k}{c}$; $\Sigma = \mathbb{F}_q^c$

given $m = m_1, \ldots, m_{k'} \in \mathbb{F}_q^c$

view m as deg. $k-1$ poly $M \in \mathbb{F}_q[x]$

Encode $M \to$

$$\boxed{M(\alpha), M(\alpha^2) \ldots M(\alpha^c)}, \boxed{M(\alpha^{c+1}) \ldots M(\alpha^{2c})} \cdots \boxed{M(\alpha^n)}$$

(yields n' elements of Σ)!

Theorem [GR '06]: An algorithm of [PV'05] can be used to list-decode this code from $\left(1 - \frac{k'}{n'} - \epsilon\right)$ fraction errors !!!

Rest of these lectures

- Development of these codes / decoder
- Decoding Algorithm
- Analysis

ACCIDENTAL DISCOVERIES

- [KIAYIAS + YUNG] : Reed-Solomon decoding from more than $1 - \sqrt{\frac{k}{n}} + \epsilon$ fraction errors appears hard. Maybe can build some cryptographic primitives from this hard problem?

- EXAMPLE (Not from [KY] but useful for us):
  - Suppose A & B share secret $S \subseteq [n]$ which is not known to E
  - Then to send p to B, A sends $y_1 \ldots y_n$ to B, where $y_i = p(x_i)$, $i \in S$
    $= \text{random o.w.}$

- B knows S & so finding p is just interpolation
- E does not know S, so has to recover message from $1 - \frac{|S|}{n}$ errors ... hard (by assumption)!

- WEAKNESS: Useful for one-time key exchange, but what happens when A & B use same S to exchange P_1 & P_2?

- Leads to new-code + decoding problem

$\boxed{\text{"Interleaved RS Code"}}$

Message $\qquad\qquad\qquad$ Encoding

$$(P_1, P_2) \longmapsto \{(P_1(\alpha), P_2(\alpha))\}_{\alpha \in F}$$

- maps $((\mathbb{F}_q)^k)^2$ to $(\mathbb{F}_q{}^2)^n$

- Error Model : - Some symbols in $\mathbb{F}_q{}^2$ received OK
  - Others corrupted at random.
  - Is it still hard to recover from $1-\sqrt{\frac{k}{n}}+\epsilon$ errors?

- [Bliechenbakher Kiayias Yung]

Can recover from $1-\frac{2k+n}{3n}$ random errors!

- [Coppersmith + S.]

Can recover from $1-O\left(\left(\frac{k}{n}\right)^{2/3}\right)$ random errors!

[CS '04] Algorithm

Idea: Now we have triples $\{(x_i, y_i, z_i)\}_{i=1}^{n}$ & want to find p_1, p_2 s.t. for many $i \in [n]$, $y_i = p_1(x_i)$, $z_i = p_2(x_i)$

Maybe should fit 3-variate poly ?

deg. $= 3n^{1/3}$ ... very good !

1^{st} Attempt:

- Find coefficients of Q by solving some big linear system $A \cdot v_Q = 0$
- Stare at v_Q.

Conclusion of [CS] : Eyes get tired

2^{nd} Attempt:

- Solve $A^T \cdot w = 0$ where A as before
- $w_i = 0 \Rightarrow$ Erase (x_i, y_i, z_i).

Theorem: [CS]: Corrects $1 - O\left(\frac{k}{n}\right)^{2/3}$ random errors, if $q \gg n$.

Motivating questions for [PV]

- Is Big-Oh in $1 - O\left(\frac{k}{n}\right)^{2/3}$ necessary?
- Is random errors, the best we can deal with
- Can we stare at 1^{st} Attempt any better?

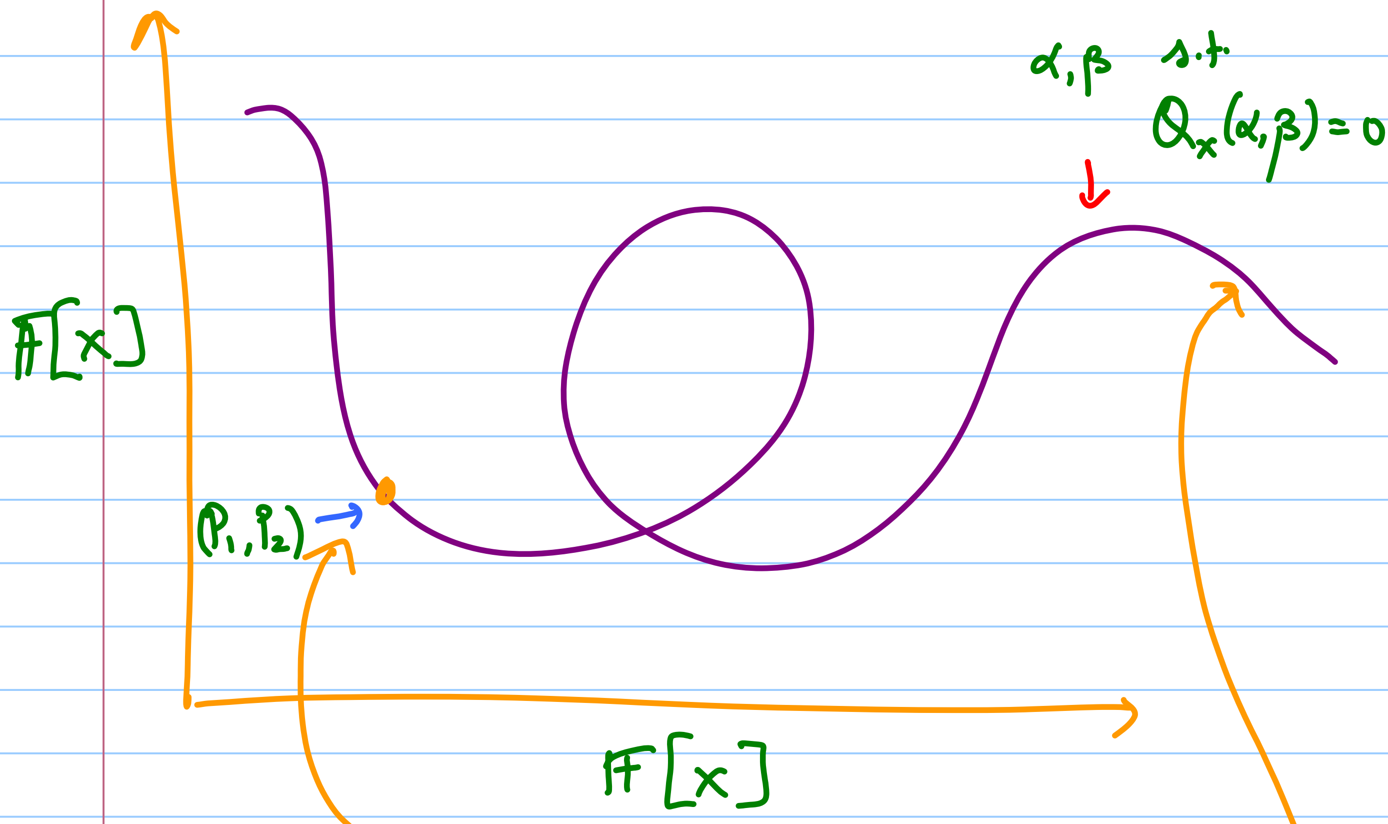
Some Obstacles

1. Can't correct more worst-case errors unless one can decode more errors in RS codes.
(else, just pad RS decoding instance with $z_i = 0$, $\forall i$)

2. Problem with $A \cdot v_Q = 0$ approach.
If lucky we find Q s.t.

$$Q(x,y,z) = A(x,y,z)\cdot(y - p_1(x)) + B(x,y,z)\cdot(z - p_2(x))$$

Viewing $Q \in \mathbb{F}[x][y,z]$ and plotting all its zeroes, we get a picture like the following

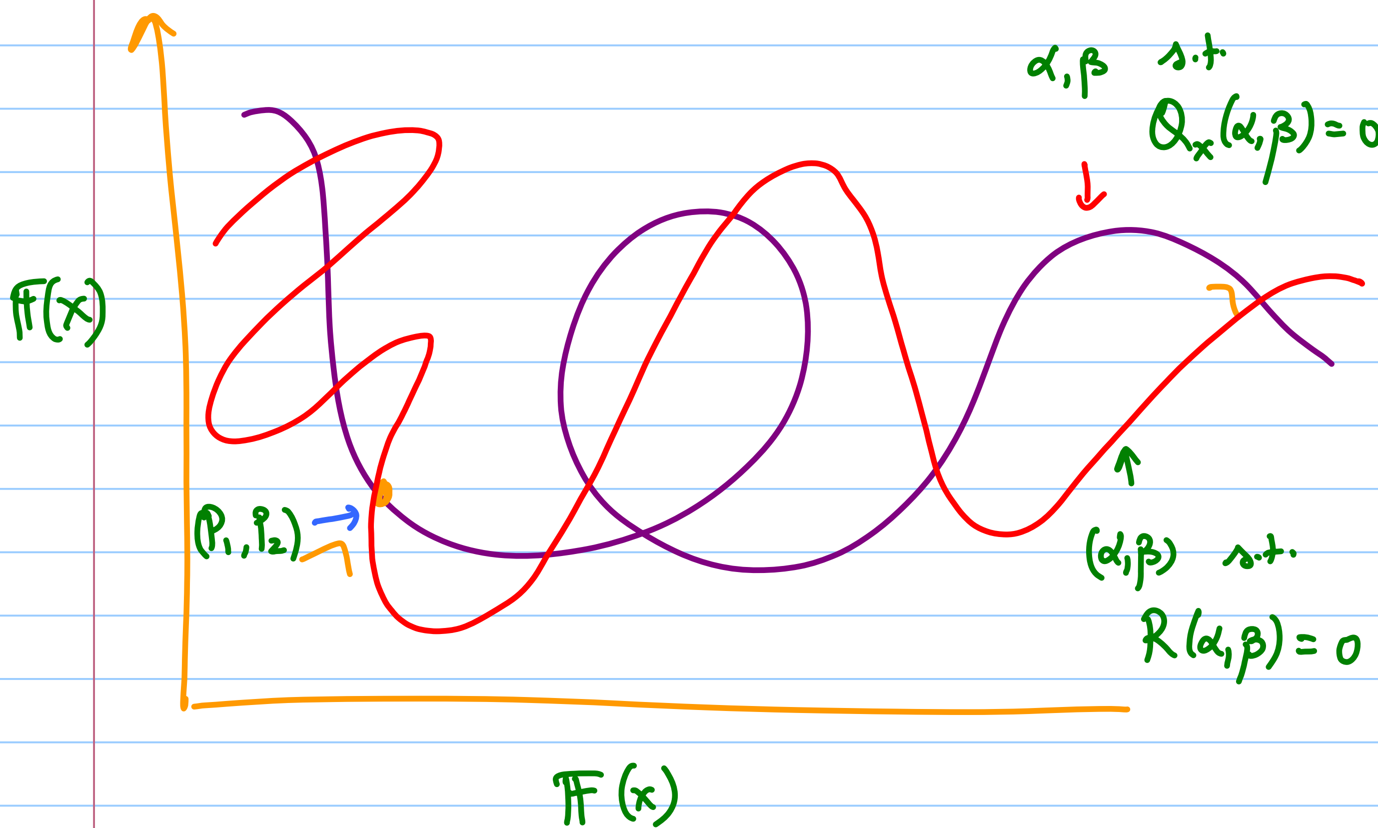


- We are trying to find this point.
- Only information about it we have (if we throw away data) is this...

... (p_1, p_2) is somewhere on this curve.

[PV '05] Ingenious Idea

Impose a relation on (p_1, p_2) a priori !



- Specifically treat p_1 as message & let p_2 be such that $R(p_1, p_2) = 0$.
- Now have only few points that could be (p_1, p_2).
- GOOD NEWS : list-decodable ....

BAD NEWS : lost in rate ... rate $= \frac{k'}{2n'}$.

Some issues

- For arbitrary $R_x(y,z)$, given P_1 finding P_2 s.t. $R(P_1, P_2) = 0$ could be non-trivial.
- Even if we find it P_2 may have large degree.
- [PV '05] Idea (only "clever" compared to their other idea of introducing $R(y,z)$)
  - Reduce $\mathbb{F}[x]$ mod $h(x)$ of deg. k.
  - Reduces degree of P_2 !
  - Makes coefficient ring nice (a field if $h(x)$ irreducible).
  - $R(y,z) = Z - y^D$ works!!

[PV '05]: CODE + DECODING (fix $\mathbb{F}_q$, $h(x)$ monic irred. deg. k, integer D)

$\Sigma = \mathbb{F}_q^2$; $n = q$;

- Given message $= p_1(x) \in \mathbb{F}_q[x]$ of deg $< k$.

- Let $p_2(x) = p_1(x)^D \bmod h(x)$

- Encoding

$$p_1 \longmapsto \left\{ \left(p_1(\alpha), p_2(\alpha) \right) \right\}_{\alpha \in \mathbb{F}_q}$$

- Rate $= \dfrac{k}{2n}$

Decoding Problem

Given: $\{(\alpha_i, \beta_i, \gamma_i)\}_{i=1}^{n}$

Find: A list of all deg $< k$ polys P_1

s.t.

$$\left| \left\{ i \;\middle|\; \begin{array}{l} \beta_i = P_1(\alpha_i) \\ \gamma_i = P_2(\alpha_i) \end{array} \right\} \right| \geq t$$

for $P_2(x) = P_1(x)^{D} \bmod h(x)$.

Decoding Algorithm

Step 1: Find Q of deg $\leq k^{2/3} n^{1/3}$ in x

$\leq \left(\frac{n}{k}\right)^{1/3}$ in y

$\leq \left(\frac{n}{k}\right)^{1/3}$ in z

s.t. $Q(\alpha_i, \beta_i, \gamma_i) = 0 \quad \forall i$

[if $h(x)^i \mid Q(x,y,z)$, use Q/h^i instead; since $h(\alpha) \neq 0$, $\frac{Q}{h^i}(\alpha_i,\beta_i,\gamma_i) = 0$ ...]

(Can/Should throw in multiplicities as well.)

Step 2: Let $Q_x(y,z) = Q(x,y,z) \bmod h(x)$

Let $P_x(Y) = Q_x(Y, Y^D)$

Report all "roots" $p_1(x)$ in $E = \mathbb{F}[x]/h(x)$ satisfying $P_x(p_1(x)) = 0$

Claim 1: Such a poly Q exists & can be found (Step 1).

Claim 2: If P_1 & $P_2 = P_1^D \mod h(x)$

satisfy $\left| S = \{ i \mid \beta_i = P_1(\alpha_i),\ \gamma_i = P_2(\alpha_i) \} \right| > 3k^{2/3} n^{1/3}$

then $Q_x(P_1, P_2) = 0 \mod h(x)$

Proof: Let $g(x) = Q(x, P_1(x), P_2(x))$

Then $\deg(g) \leq 3 \cdot k^{2/3} n^{1/3}$

But $g(\alpha_i) = 0 \quad \forall i \in S$

$\Rightarrow g \equiv 0 \Rightarrow g \mod h(x) \equiv 0$

$\Rightarrow Q_x(P_1, P_2) = 0 \mod h(x)$.

[Note: $Q_x(Y, Z) \not\equiv 0$.]

Claim 3: P_1 is a root of $P_x(Y)$

(Immediate from Claim 3)

Claim 4: # roots of P_x is bounded by ... .

provided $D > \ldots \left(\left(\frac{n}{k}\right)^{1/3} \right)$

Proof:

- First, $Q \not\equiv 0$ - by constraint on Step 1.
- Next, $Q_x = Q \bmod h(x) \not\equiv 0$ since we divided out by h^i
- Note y-degree of $Q_x \le \left(\frac{n}{k}\right)^{1/3}$, so if $D > \left(\frac{n}{k}\right)^{1/3}$ then $P_x(y) \neq 0$ (since no pair of monomials cancel each other.)
- But $\deg P_x \le D \cdot \left(\frac{n}{k}\right)^{1/3}$

$$\Rightarrow \# \text{roots} \le D\left(\frac{n}{k}\right)^{1/3} = \left(\frac{n}{k}\right)^{2/3}$$

$\boxtimes$

Conclusions

- Can correct $n - O(k^{2/3} n^{1/3})$ errors.
- With multiplicities $n - k^{2/3} n^{1/3}$ errors.
- But $R = \frac{1}{2} \cdot \frac{k}{n}$
- So only getting codes of rate $\frac{1}{2}(1-p)^{3/2}$ correcting p fraction errors.
- CS perspective: Exponent of $1-p$ more important than constant in front. (so this is important)
- Proved formally in [GR '06]: Next lecture!

$\boxtimes$