Constructions of Codes with the Locality Property

Alexander Barg

University of Maryland

DIMACS Workshop “Network Coding: The Next 15 years”

Acknowledgment

Based on joint works with

Itzhak Tamo

Alexey Frolov

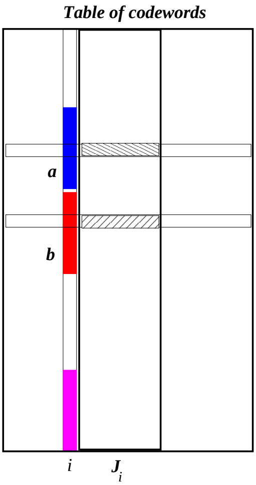
Serge Vlăduţ

Sreechakra Goparaju

Robert Calderbank

Locally recoverable codes

The code $\mathcal{C} \subset \mathbb{F}^n$ is locally recoverable with locality r if every symbol can be recovered by accessing some other r symbols in the encoding (recovery set of coordinate i)



(n, k, r) LRC code

Definition (LRC codes)

Code $\mathcal{C}$ has *locality* r if for every $i \in [n]$ there exists a subset $J_i \subset [n]\backslash i, |J_i| \leq r$ and a function ϕ_i such that for every codeword $c \in \mathcal{C}$

$$c_i = \phi_i(\{c_j, j \in J_i\})$$

J. Han and L. Lastras-Montano, *ISIT* 2007;
C. Huang, M. Chen, and J. Li, *Symp. Networks App.* 2007;
F. Oggier and A. Datta '10;
P. Gopalan, C. Huang, H. Simitci, and S. Yekhanin, *IEEE Trans. Inf. Theory,* Nov. 2012.

Linear index codes are duals of linear DS codes on graphs

(Mazumdar '14; Shanmugam-Dimakis '14)

(n, k, r) LRC code

Definition (LRC codes)

Code $\mathcal{C}$ has *locality* r if for every $i \in [n]$ there exists a subset $J_i \subset [n]\backslash i, |J_i| \le r$ and a function ϕ_i such that for every codeword $c \in \mathcal{C}$

$$c_i = \phi_i(\{c_j, j \in J_i\})$$

Examples:

Repetition codes, Single parity-check codes
$[n, r, n-r+1]$ RS code

Early constructions:

Prasanth, Kamath, Lalitha, Kumar, ISIT 2012
Silberstein, Rawat, Koyluoglu Vishwanath, ISIT 2013
Tamo, Papailiopoulos, Dimakis, ISIT 2013

Outline

- RS-like LRC codes
- Bounds on LRC codes
- LRC codes on curves
- Cyclic LRC codes

RS codes and Evaluation codes

Given a polynomial $f \in \mathbb{F}_q[x]$ and a set $A = \{P_1, \dots, P_n\} \subset \mathbb{F}_q$ define the map

$$ev_A : f \mapsto (f(P_i), i = 1, \dots, n)$$

Example: Let $q = 8$, $f(x) = 1 + \alpha x + \alpha x^2$

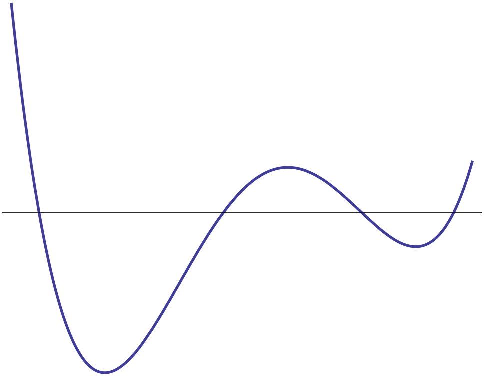
$$f(x) \mapsto (1, \alpha^4, \alpha^6, \alpha^4, \alpha, \alpha, \alpha^6)$$

Evaluation code $\mathcal{C}(A)$

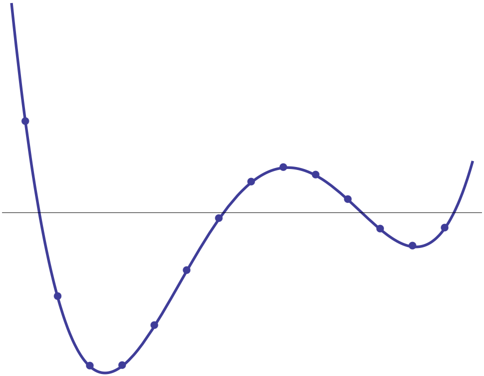
Let $V = \{f \in \mathbb{F}_q[x]\}$ be a set of polynomials, $\dim(V) = k$

$$\mathcal{C} : V \to \mathbb{F}_q^n$$
$$f \mapsto ev_A(f) = (f(P_i), i = 1, \dots, n)$$

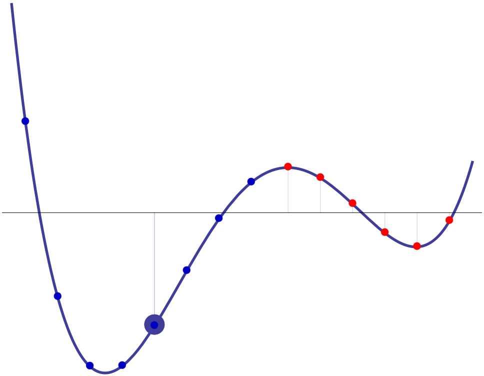
Reed-Solomon codes



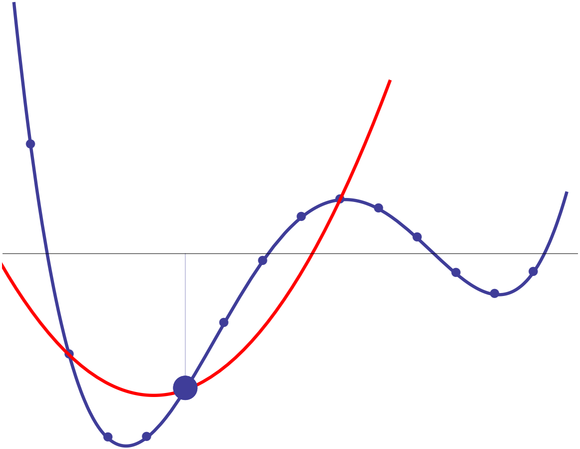
Reed-Solomon codes



Reed-Solomon codes



Evaluation codes with locality



Construction of (n, k, r) LRC codes: Example

Parameters: $n = 9, k = 4, r = 2, q = 13;$

Set of points: $A = \{P_1, \dots, P_9\} \subset \mathbb{F}_{13}$

$$\mathcal{A} = \{A_1 = (1, 3, 9), A_2 = (2, 6, 5), A_3 = (4, 12, 10)\}$$

Set of functions: $\mathcal{P} = \{f_a(x) = a_0 + a_1 x + a_3 x^3 + a_4 x^4\}$

Code construction:

$$ev_A : f_a \mapsto (f(P_i), i = 1, \dots 9)$$

E.g., $a = (1111)$ then $f_a(x) = 1 + x + x^3 + x^4$

$$c := ev_A(f_a) = (\underbrace{4, 8, 7}_{A_1} | \underbrace{1, 11, 2}_{A_2} | \underbrace{0, 0, 0}_{A_3})$$

$$f_a(x)|_{A_1} = a_0 + a_3 + (a_1 + a_4)x \ = 2 + 2x$$

$$f_a(x)|_{A_2} = a_0 + 8a_3 + (a_1 + 8a_4)x$$

Construction of (n, k, r) LRC codes

$$A = (P_1, \dots, P_n) \subset \mathbb{F}_q$$

$$A = A_1 \cup A_2 \cup \cdots \cup A_{\frac{n}{r+1}}$$

Basis of functions: Take $g(x)$ constant on A_i, $i = 1, \dots, \frac{n}{r+1}$ (above $g(x) = x^3$)

$$V = \left\langle g(x)^j x^i, i = 0, \dots, r-1; j = 0, \dots, \frac{k}{r} - 1 \right\rangle; \ \dim(V) = k$$

$$V = \left\{ f_a(x) = \sum_{i=0}^{r-1} \sum_{j=0}^{\frac{k}{r}-1} a_{ij} g(x)^j x^i \right\}$$

We obtain a family of optimal r-LRC codes

Erasure recovery by polynomial interpolation over r points.

I. Tamo and *A.B.*, *IEEE Trans. Inf. Theory,* Aug. 2014.

Extensions

- Codes with multiple disjoint recovery sets for every coordinate
- Codes that recover locally from $\rho \geq 2$ erasures: The local codes are $[r + \rho - 1, r, \rho]$ MDS
- Systematic encoding

Finite-length bounds

Let $\mathcal{C} \subset \mathbb{F}_q^n$ be an r-LRC code, $|\mathcal{C}| = q^k$, distance d

$$d \le n - k - \left\lceil \frac{k}{r} \right\rceil + 2$$

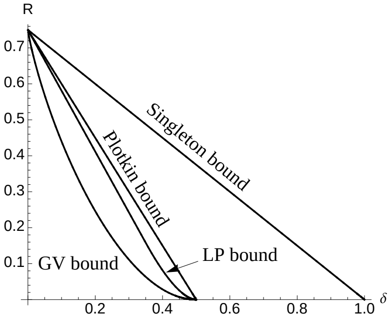
(*P. Gopalan* e.a. 2012)

$$k \le \min_{s \ge 1} \{ sr + k_q(n - s(r+1), d) \}$$

(*V. Cadambe* and *A. Mazumdar*, 2013-15)

Bounds for multiple recovery sets (work with I. Tamo, 2014)

Asymptotic bounds



Binary codes, $r = 3$; $n \to \infty$

$$R_q(r,\delta) > 0, \quad 0 \le \delta < (q-1)/q$$

$$R_q(r,0) = \frac{r}{r+1}, \quad R_q(r,\delta) = 0, \ \frac{q-1}{q} \le \delta \le 1$$

Geometric view of LRC codes

$$A = \{1, \ldots, 9\} \subset \mathbb{F}_{13}$$

$$A = A_1 \cup A_2 \cup A_3$$

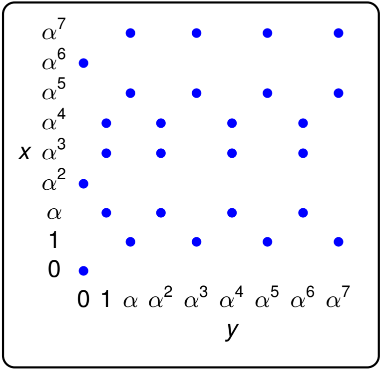
$$A_1 = (1, 3, 9)$$
$$A_2 = (2, 6, 5)$$
$$A_3 = (4, 12, 10)$$

$$g \colon A \to \mathbb{F}_{13}$$
$$x \mapsto x^3 - 1$$

$$g \colon \mathbb{F}_{13} \to \{0, 7, 8\} \subset \mathbb{F}_{13}$$
$$|g^{-1}(y)| = r + 1$$

LRC codes on curves

Consider the set of pairs $(x, y) \in \mathbb{F}_9$ that satisfy the equation $x^3 + x = y^4$



Affine points of the Hermitian curve $\mathcal{X}$ over $\mathbb{F}_9$; $\alpha^2 = \alpha + 1$

Hermitian codes

$$\begin{array}{rcl} g: \mathcal{X} & \to & \mathbb{P}^1 \\ (x,y) & \mapsto & y \end{array}$$

Space of functions $V := \langle 1, y, y^2, x, xy, xy^2 \rangle$

A={Affine points of the Hermitian curve over $\mathbb{F}_9$}; $n = 27, k = 6$

$$\mathcal{C} : V \to \mathbb{F}_9^n$$

E.g., message $(1, \alpha, \alpha^2, \alpha^3, \alpha^4, \alpha^5)$

$$F(x,y) = 1 + \alpha y + \alpha^2 y^2 + \alpha^3 x + \alpha^4 xy + \alpha^5 xy^2$$

$$F(0,0) = 1 \text{ etc.}$$

LRC codes on curves

x \ y	0	1	α	α^2	α^3	α^4	α^5	α^6	α^7
α^7			α		α^7		α^5		0
α^6	α^2								
α^5			α^6		α^4		α^2		0
α^4		α^7		α^3		α^5		α^5	
α^3		α^3		α^7		α		α	
α^2	α^3								
α		0		0		0		0	
1			1		α^6		α^4		0
0	1								

Hermitian LRC codes

x \ y	0	1	α	α^2	α^3	α^4	α^5	α^6	α^7
α^7			α		α^7		α^5		0
α^6	α^2								
α^5			α^6		α^4		α^2		0
α^4		α^7		α^3		α^5		α^5	
α^3		α^3		α^7		α		α	
α^2	α^3								
α		0 ✗		0		0		0	
1			1		α^6		α^4		0
0	1								

Let $P = (\alpha, 1)$ be the erased location.

Local recovery with Hermitian codes

x \ y	0	1	α	α^2	α^3	α^4	α^5	α^6	α^7
α^7			α		α^7		α^5		0
α^6	α^2								
α^5			α^6		α^4		α^2		0
α^4		α^7		α^3		α^5		α^5	
α^3		α^3		α^7		α		α	
α^2	α^3								
α		?		0		0		0	
1			1		α^6		α^4		0
0	1								

Let $P = (\alpha, 1)$ be the erased location. Recovery set $I_P = \{(\alpha^4, 1), (\alpha^3, 1)\}$
Find $f(x) : f(\alpha^4) = \alpha^7, f(\alpha^3) = \alpha^3$

$$\Rightarrow \quad f(x) = \alpha x - \alpha^2$$

$$f(\alpha) = 0 = F(P)$$

Hermitian codes

$q = q_0^2$, q_0 prime power

$$\mathcal{X} : x^{q_0} + x = y^{q_0+1}$$

$\mathcal{X}$ has $q_0^3 = q^{3/2}$ points in $\mathbb{F}_q$

Let $g : \mathcal{X} \to \mathcal{Y} = \mathbb{P}^1$, $g(P) = g(x, y) := y$

We obtain a family of q-ary codes of length $n = q_0^3$,

$$k = (t+1)(q_0 - 1), d \geq n - tq_0 - (q_0 - 2)(q_0 + 1)$$

with locality $r = q_0 - 1$.

It is also possible to take $g(P) = x$ (projection on x); we obtain LRC codes with locality q_0

General construction

Map of curves

X, Y smooth projective absolutely irreducible curves over $\mathbb{k}$

$$g : X \to Y$$

rational separable map of degree $r+1$

Lift the points of Y

$S = \{P_1, \dots, P_s\} \subset Y(\mathbb{k})$. Partition of points:

$$A := g^{-1}(S) = \{P_{ij}, i = 0, \dots, r, j = 1, \dots, s\} \subseteq X(\mathbb{k})$$

such that $g(P_{ij}) = P_j$ for all i, j

Basis of the function space:

$Q_\infty = \pi^{-1}(\infty)$, where $\pi : Y \to \mathbb{P}^1_{\mathbb{k}}$

$\{f_1, \dots, f_m\}$ span $L(tQ_\infty), t \geq 1$

$$\{f_j x^i, i = 0, \dots, r-1; j = 1, \dots, m\}$$

Construct LRC codes

Evaluation codes constructed on the set A are LRC codes with locality r

Asymptotically good sequences of codes

Let $q = q_0^2$, where q_0 is a prime power. Take Garcia-Stichtenoth towers of curves:

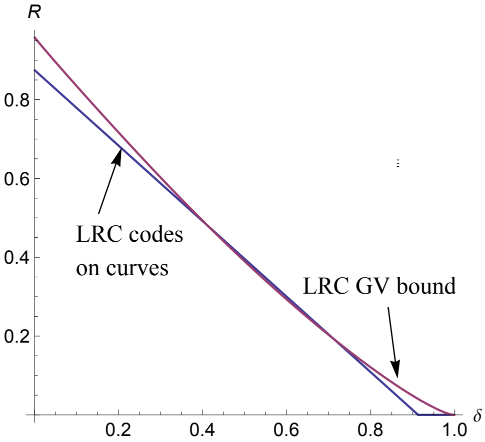
$$x_0 := 1;\ X_1 := \mathbb{P}^1, \Bbbk(X_1) = \Bbbk(x_1);$$
$$X_l : z_l^{q_0} + z_l = x_{l-1}^{q_0+1}, x_{l-1} := \frac{z_{l-1}}{x_{l-2}} \in \Bbbk(X_{l-1}) \text{ (if } l \geq 3)$$

There exist families of q-ary LRC codes with locality r whose *rate and relative distance* satisfy

$$R \geq \frac{r}{r+1}\Big(1 - \delta - \frac{3}{\sqrt{q}+1}\Big), \qquad r = \sqrt{q} - 1$$
$$R \geq \frac{r}{r+1}\Big(1 - \delta - \frac{2\sqrt{q}}{q-1}\Big), \qquad r = \sqrt{q}$$

$^{*)}$Recall the TVZ bound without locality: $R \geq 1 - \delta - \frac{1}{\sqrt{q}-1}$

LRC codes on curves better than the GV bound



Extensions

Common theme: Automorphism groups of curves

- LRC codes on curves with multiple recovery sets
- Asymptotically good codes with small locality
  Let $(r+1)|(q_0+1)$
  $k(Y_{l,r}) = k(x_1^{r+1}, z_2, \dots, z_l)$

$$\begin{aligned} g : X_l &\to Y_{l,r} \\ x_1 &\mapsto x_1^{r+1} \end{aligned}$$

- Local codes with distance $\rho \geq 3$

Work with *I. Tamo* and *S. Vlăduţ*, 2015; ongoing

Cyclic LRC codes

Consider the special case of the RS-like code family with $n|(q-1)$, $g(x)=\prod_{h\in H}(x-h)$, where H is a subgroup of $\mathbb{F}_q^*$

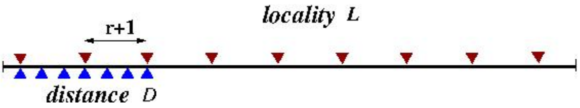
$$f_a(x)=\sum_{\substack{i=0\\ i\neq r \bmod (r+1)}}^{(k/r)(r+1)-2} a_i x^i$$

Theorem: Consider the following sets of elements of $\mathbb{F}_q$:

$$L=\{\alpha^i, i \bmod (r+1)=l\} \text{ and } D=\{\alpha^{j+s}, s=0,\dots,n-k(r+1)/r\},$$

where $\alpha^j\in L$. The cyclic code with the defining set of zeros $\mathscr{Z}=L\cup D$ is an optimal (n,k,r) q-ary cyclic LRC code.

Set of zeros



Subsets of zeros for distance (D) and locality (L)

Proposition: Let $t|n$. If $\mathscr{Z}$ contains some coset of the group of tth roots of unity, then

$d(\mathcal{C}^\perp) \le t$, i.e., $\mathcal{C}$ has locality $r = t - 1$.

$$\begin{array}{ccc} \mathcal{C} & \longleftrightarrow & \mathcal{C}^\perp \\ \downarrow \mathbb{F}_p & & \downarrow Tr \\ \mathcal{D} & \longleftrightarrow & \mathcal{D}^\perp \end{array}$$

(BCGT, 2015; ongoing)

Outlook

- Partial MDS codes (max recoverable codes)
- Cyclic codes
- Decoding
- Constructions on curves

Thank you!