

# Bounds on parameters of LRC-t codes

Petar Hlad Colic

Universitat Politècnica de Catalunya

May 2018

- 1 Preliminaries on LRC codes
  - Definition and properties of LRC
- 2 Construction of LRC codes
- 3 Bounds on LRC-t codes
  - Recovery graph
  - Statements of the bounds
  - Definitions and examples for the proof
  - Proof for the rate bound
  - Proof of the minimum distance bound

# Notation

- $\mathcal{C}$  denotes a code over the finite field  $\mathbb{F}_q$ .
- The triple of parameters  $(n, k, r)$  refers to a code of:
  - length  $n$
  - cardinality  $q^k$
  - locality  $r$
- $[n] := \{1, \dots, n\}$
- A *restriction*  $\mathcal{C}_I$  of the code  $\mathcal{C}$  to a subset of coordinates  $I \subset [n]$  is the code obtained by removing from each vector the coordinates outside  $I$ .

## Definition of LRC Codes

Given  $a \in \mathbb{F}_q$  consider the sets of codewords of  $\mathcal{C}$  with fixed value  $a$  at the symbol  $x_i$ :

$$\mathcal{C}(i, a) = \{x \in \mathcal{C} : x_i = a\}, \quad i \in [n]$$

### Definition 1.1.

A code  $\mathcal{C}$  of length  $n$  has **locality  $r$**  if  $\forall i \in [n]$  there exists a subset  $I_i \subset [n] \setminus i$ ,  $|I_i| \leq r$  such that the restrictions of the sets  $\mathcal{C}(i, a)$  to the coordinates in  $I_i$  for different  $a$  are disjoint:

$$\mathcal{C}_{I_i}(i, a) \cap \mathcal{C}_{I_i}(i, a') = \emptyset, \quad a \neq a'.$$

## Bounds on rate and min. distance

Let  $\mathcal{C}$  be an  $(n, k, r)$  LRC code of cardinality  $q^k$  over an alphabet of size  $q$ . Then:

### Theorem 1.2 (Upper bound on the rate).

*The rate of  $\mathcal{C}$  satisfies*

$$\frac{k}{n} \leq \frac{r}{r+1}$$

### Theorem 1.3 (Generalization of Singleton bound).

*The minimum distance of  $\mathcal{C}$  satisfies*

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

*A code that achieves the bound with equality will be called an optimal LRC code*

# Construction of LRC codes

We want to construct a linear  $(n, k, r)$ -LRC code. Assume  $r|k$  and  $(r+1)|n$ .

We need:

- $A_1, \dots, A_{\frac{n}{r+1}}$  disjoint subsets of the field  $\mathbb{F}_q$ , s.t.  $|A_i| = r+1$
- $g(x) \in \mathbb{F}_q[x]$  a polynomial s.t.
  - 1  $\deg(g) = r+1$
  - 2  $g$  is constant on each set  $A_i$ :  $g(\alpha) = g(\beta)$  for  $\alpha, \beta \in A_i$

We will call  $g$  a good polynomial.

## Construction of LRC codes

Let  $A = \bigcup_{i=1}^{\frac{n}{r+1}} A_i \subset \mathbb{F}_q$ ,  $|A| = n$ .

We write now message vectors  $a \in \mathbb{F}_q^k$  as  $r \times \frac{k}{r}$  matrices.

$$a = \begin{pmatrix} a_{0,0} & a_{0,1} & \cdots & a_{0,\frac{k}{r}-1} \\ a_{1,0} & a_{1,1} & \cdots & a_{1,\frac{k}{r}-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r-1,0} & a_{r-1,1} & \cdots & a_{r-1,\frac{k}{r}-1} \end{pmatrix}$$

# Construction of LRC codes

## Encoding polynomial

Given the message vector  $a \in \mathbb{F}_q^k$ , define the **encoding polynomial** as:

$$f_a(x) = \sum_{i=0}^{r-1} x^i \cdot f_i(x)$$

where

$$f_i(x) = \sum_{j=0}^{\frac{k}{r}-1} a_{ij} g(x)^j$$



$$\begin{aligned}
 f_a(x) &= (x^0 \quad \dots \quad x^{r-1}) \begin{pmatrix} a_{0,0} & \cdots & a_{0,\frac{k}{r}-1} \\ \vdots & \ddots & \vdots \\ a_{r-1,0} & \cdots & a_{r-1,\frac{k}{r}-1} \end{pmatrix} \begin{pmatrix} g(x)^0 \\ \vdots \\ g(x)^{\frac{k}{r}-1} \end{pmatrix} = \\
 &= (x^0 \quad \dots \quad x^{r-1}) \begin{pmatrix} f_0(x) \\ \vdots \\ f_{r-1}(x) \end{pmatrix}
 \end{aligned}$$

The codeword for  $a \in \mathbb{F}_q^k$  is found as the evaluation vector of  $f_a$  at all the points of  $A$ .

### LRC code

The  $(n, k, r)$  LRC code  $\mathcal{C}$  is defined as the set of  $n$ -dimensional vectors

$$\mathcal{C} = \{(f_a(\alpha), \alpha \in A) : a \in \mathbb{F}_q^k\}$$

## Remark 2.1.

$$x \in A_i \Rightarrow g(x) \text{ constant}$$

$$\Rightarrow f_\ell(x) = \sum_{j=0}^{\frac{k}{r}-1} a_{\ell j} g(x)^j \text{ constant in } A_i$$

$$\Rightarrow \deg(f_a(x)) = \deg\left(\sum_{j=0}^{r-1} x^j \cdot f_j(x)\right) \leq r-1 \text{ in } A_i$$

## Recovery of the erased symbol

Suppose erased symbol:  $\alpha \in A_j$ .

Let  $(c_\beta, \beta \in A_j \setminus \alpha)$  denote the remaining  $r$  symbols of the recovering set.

To find the value  $c_\alpha = f_a(\alpha)$ , find the unique polynomial  $\delta(x)$  s.t.

- $\deg(\delta(x)) \leq r$
- $\delta(\beta) = c_\beta \quad \forall \beta \in A_j \setminus \alpha$

This polynomial is:

$$\delta(x) = \sum_{\beta \in A_j \setminus \alpha} c_\beta \prod_{\beta' \in A_j \setminus \{\alpha, \beta\}} \frac{x - \beta'}{\beta - \beta'}$$

Finally, set  $c_\alpha = \delta(\alpha)$ .

## Theorem 2.2.

*The linear code  $\mathcal{C}$  defined has dimension  $k$  and is an optimal  $(n, k, r)$  LRC code.*

## Proof of dimension.

For  $i \in \{0, \dots, r-1\}; j \in \{0, \dots, \frac{k}{r-1}\}$  the  $k$  polynomials  $g(x)^j x^i$  all are of distinct degrees, i.e. linearly independent over  $\mathbb{F}$ .

$\Rightarrow$  The mapping  $a \mapsto f_a$  is injective.

$$\begin{aligned}\deg(f_a(x)) &\leq \deg(x^{r-1}) + \deg(g(x)^{\frac{k}{r}-1}) = r-1 + (r+1)\left(\frac{k}{r} - 1\right) \\ &= k + \frac{k}{r} - 2 \leq n - 2\end{aligned}$$

This means that two distinct encoding polynomials give rise to two distinct codewords.  $\Rightarrow$  The dimension of the code is  $k$ .  $\square$

## Proof of optimality.

Since the encoding is linear:

$$d(\mathcal{C}) \geq n - \max_{f_a, a \in \mathbb{F}_q^k} \deg(f_a) = n - k - \frac{k}{r} + 2 \geq n - k - \left\lceil \frac{k}{r} \right\rceil + 2$$

But we have that  $d(\mathcal{C}) \leq n - k - \left\lceil \frac{k}{r} \right\rceil + 2$ . Therefore, we have equality and thus it is an optimal LRC Code.  $\square$

## Example: $(9,4,2)$ LRC code

We will now construct a  $(n = 9, k = 4, r = 2)$  LRC code over the field  $\mathbb{F}_q$ .

$$q = |\mathbb{F}_q| \geq n \Rightarrow q \geq 9$$

Choose  $q = 13$

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

.

$$g(x) = x^3 = \begin{cases} 1 & \text{if } x \in A_1 \\ 8 & \text{if } x \in A_2 \\ 12 & \text{if } x \in A_3 \end{cases}$$



For  $a = \begin{pmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{pmatrix} \in \mathbb{F}_{13}^4$  define the encoding polynomial:

$$f_a(x) = \begin{pmatrix} 1 & x \end{pmatrix} \begin{pmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{pmatrix} \begin{pmatrix} 1 \\ x^3 \end{pmatrix} = a_{00} + a_{10}x + a_{01}x^3 + a_{11}x^4$$

E.g.  $a = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ .  $f_a(x) = 1 + x + x^3 + x^4$

$$\begin{aligned} c &= (f_a(1), f_a(3), f_a(9), f_a(2), f_a(6), f_a(5), f_a(4), f_a(12), f_a(10)) \\ &= (4, 8, 7, 1, 11, 2, 0, 0, 0) \end{aligned}$$

$$\delta(x) = \sum_{\beta \in A_j \setminus \alpha} c_\beta \prod_{\beta' \in A_j \setminus \{\alpha, \beta\}} \frac{x - \beta'}{\beta - \beta'}$$

$$(f_a(1), f_a(3), f_a(9), f_a(2), f_a(6), f_a(5), f_a(4), f_a(12), f_a(10)) \\ (4, 8, 7, 1, 11, 2, 0, 0, 0)$$

$$\delta(x) = \sum_{\beta \in A_j \setminus \alpha} c_\beta \prod_{\beta' \in A_j \setminus \{\alpha, \beta\}} \frac{x - \beta'}{\beta - \beta'}$$

$$(\cancel{f_a(1)}, f_a(3), f_a(9), f_a(2), f_a(6), f_a(5), f_a(4), f_a(12), f_a(10))$$

$$(\cancel{4}, 8, 7, 1, 11, 2, 0, 0, 0)$$

$$\delta(x) = \sum_{\beta \in A_j \setminus \alpha} c_\beta \prod_{\beta' \in A_j \setminus \{\alpha, \beta\}} \frac{x - \beta'}{\beta - \beta'}$$

$$(\cancel{f_a(1)}, f_a(3), f_a(9), f_a(2), f_a(6), f_a(5), f_a(4), f_a(12), f_a(10))$$

$$(\cancel{4}, 8, 7, 1, 11, 2, 0, 0, 0)$$

$$1 \in A_1 = \{1, 3, 9\}$$

$$\delta(x) = \sum_{\beta \in A_j \setminus \alpha} c_\beta \prod_{\beta' \in A_j \setminus \{\alpha, \beta\}} \frac{x - \beta'}{\beta - \beta'}$$

$$(\cancel{f_a(1)}, f_a(3), f_a(9), f_a(2), f_a(6), f_a(5), f_a(4), f_a(12), f_a(10))$$

$$(\cancel{4}, 8, 7, 1, 11, 2, 0, 0, 0)$$

$$1 \in A_1 = \{1, 3, 9\}$$

$$\Rightarrow \delta(x) = c_3 \frac{x - 9}{3 - 9} + c_9 \frac{x - 3}{9 - 3} = 2x + 2$$

$$\delta(x) = \sum_{\beta \in A_j \setminus \alpha} c_\beta \prod_{\beta' \in A_j \setminus \{\alpha, \beta\}} \frac{x - \beta'}{\beta - \beta'}$$

$$(\cancel{f_a(1)}, f_a(3), f_a(9), f_a(2), f_a(6), f_a(5), f_a(4), f_a(12), f_a(10))$$

$$(\cancel{4}, 8, 7, 1, 11, 2, 0, 0, 0)$$

$$1 \in A_1 = \{1, 3, 9\}$$

$$\Rightarrow \delta(x) = c_3 \frac{x - 9}{3 - 9} + c_9 \frac{x - 3}{9 - 3} = 2x + 2$$

$$\delta(1) = 4$$

$$\delta(x) = \sum_{\beta \in A_j \setminus \alpha} c_\beta \prod_{\beta' \in A_j \setminus \{\alpha, \beta\}} \frac{x - \beta'}{\beta - \beta'}$$

$$(f_a(1), f_a(3), f_a(9), \cancel{f_a(2)}, f_a(6), f_a(5), f_a(4), f_a(12), f_a(10))$$
$$(4, 8, 7, \cancel{11}, 11, 2, 0, 0, 0)$$

$$\delta(x) = \sum_{\beta \in A_j \setminus \alpha} c_\beta \prod_{\beta' \in A_j \setminus \{\alpha, \beta\}} \frac{x - \beta'}{\beta - \beta'}$$

$$(f_a(1), f_a(3), f_a(9), \cancel{f_a(2)}, f_a(6), f_a(5), f_a(4), f_a(12), f_a(10))$$

$$(4, 8, 7, \cancel{11}, 11, 2, 0, 0, 0)$$

$$2 \in A_2 = \{2, 6, 5\}$$



$$\delta(x) = \sum_{\beta \in A_j \setminus \alpha} c_\beta \prod_{\beta' \in A_j \setminus \{\alpha, \beta\}} \frac{x - \beta'}{\beta - \beta'}$$

$$(f_a(1), f_a(3), f_a(9), \cancel{f_a(2)}, f_a(6), f_a(5), f_a(4), f_a(12), f_a(10))$$

$$(4, 8, 7, \cancel{11}, 11, 2, 0, 0, 0)$$

$$2 \in A_2 = \{2, 6, 5\}$$

$$\Rightarrow \delta(x) = c_6 \frac{x-5}{6-5} + c_5 \frac{x-6}{5-6} = 9x + 9$$

$$\delta(x) = \sum_{\beta \in A_j \setminus \alpha} c_\beta \prod_{\beta' \in A_j \setminus \{\alpha, \beta\}} \frac{x - \beta'}{\beta - \beta'}$$

$$(f_a(1), f_a(3), f_a(9), \cancel{f_a(2)}, f_a(6), f_a(5), f_a(4), f_a(12), f_a(10))$$

$$(4, 8, 7, \cancel{11}, 11, 2, 0, 0, 0)$$

$$2 \in A_2 = \{2, 6, 5\}$$

$$\Rightarrow \delta(x) = c_6 \frac{x-5}{6-5} + c_5 \frac{x-6}{5-6} = 9x + 9$$

$$\delta(2) = 1$$

$$\delta(x) = \sum_{\beta \in A_j \setminus \alpha} c_\beta \prod_{\beta' \in A_j \setminus \{\alpha, \beta\}} \frac{x - \beta'}{\beta - \beta'}$$

$$(f_a(1), f_a(3), f_a(9), f_a(2), f_a(6), f_a(5), \cancel{f_a(4)}, f_a(12), f_a(10))$$

$$(4, 8, 7, 1, 11, 2, \cancel{0}, 0, 0)$$

$$\delta(x) = \sum_{\beta \in A_j \setminus \alpha} c_\beta \prod_{\beta' \in A_j \setminus \{\alpha, \beta\}} \frac{x - \beta'}{\beta - \beta'}$$

$$(f_a(1), f_a(3), f_a(9), f_a(2), f_a(6), f_a(5), \cancel{f_a(4)}, f_a(12), f_a(10))$$

$$(4, 8, 7, 1, 11, 2, \cancel{0}, 0, 0)$$

$$4 \in A_3 = \{4, 12, 10\}$$

$$\delta(x) = \sum_{\beta \in A_j \setminus \alpha} c_\beta \prod_{\beta' \in A_j \setminus \{\alpha, \beta\}} \frac{x - \beta'}{\beta - \beta'}$$

$$(f_a(1), f_a(3), f_a(9), f_a(2), f_a(6), f_a(5), \cancel{f_a(4)}, f_a(12), f_a(10))$$

$$(4, 8, 7, 1, 11, 2, \cancel{0}, 0, 0)$$

$$4 \in A_3 = \{4, 12, 10\}$$

$$\Rightarrow \delta(x) = c_{12} \frac{x - 10}{12 - 10} + c_{10} \frac{x - 12}{10 - 12} = 0$$

$$\delta(x) = \sum_{\beta \in A_j \setminus \alpha} c_\beta \prod_{\beta' \in A_j \setminus \{\alpha, \beta\}} \frac{x - \beta'}{\beta - \beta'}$$

$$(f_a(1), f_a(3), f_a(9), f_a(2), f_a(6), f_a(5), \cancel{f_a(4)}, f_a(12), f_a(10))$$

$$(4, 8, 7, 1, 11, 2, \cancel{0}, 0, 0)$$

$$4 \in A_3 = \{4, 12, 10\}$$

$$\Rightarrow \delta(x) = c_{12} \frac{x - 10}{12 - 10} + c_{10} \frac{x - 12}{10 - 12} = 0$$

$$\delta(4) = 0$$

## Example of LRC-2 code

Let  $\mathbb{F} = \mathbb{F}_{13}$ ,  $A = \mathbb{F} \setminus \{0\}$

$\mathcal{A} = \{\{1, 5, 12, 8\}, \{2, 10, 11, 3\}, \{4, 7, 9, 6\}\}$

$\mathcal{A}' = \{\{1, 3, 9\}, \{2, 6, 5\}, \{4, 12, 10\}, \{7, 8, 11\}\}$

$f_a(x) = a_0 + a_1x + a_2x^4 + a_3x^6$

$$a = (1, 1, 1, 1) \longrightarrow c = (4, 8, 7, 5, 2, 6, 2, 2, 2, 3, 9, 1)$$

As already seen:  $\delta(x) = 2x + 2$ ;  $\delta(1) = 4$ .

$$\delta'(x) = c_5 \frac{x-12}{5-12} \frac{x-8}{5-8} + c_{12} \frac{x-5}{12-5} \frac{x-8}{12-8} + c_8 \frac{x-5}{8-5} \frac{x-12}{8-12}$$

$$= 6 \cdot 5 \cdot (x^2 + 6x + 5) + 2 \cdot 7 \cdot (x^2 + 1) + 9 \cdot 1 \cdot (x^2 + 9x + 8)$$

$$= x^2 + x + 2 \longrightarrow \delta'(1) = 4$$

# Recovery graph

Assume every coordinate  $i$  has  $t$  disjoint recovering sets  $\mathcal{R}_i^1, \dots, \mathcal{R}_i^t$ , each of size  $r$ , where  $\mathcal{R}_i^j \subset [n] \setminus i$ .

## Definition

The **recovery graph** of a  $(n, k, r, t)$  LRC code  $\mathcal{C}$  is a directed graph  $G = (V, E)$  where:

- $V = [n]$ . (Vertices  $\leftrightarrow$  coordinates of  $\mathcal{C}$ ).
- $(i, j) \in E \iff j \in \mathcal{R}_i^\ell$  for some  $\ell \in [t]$ .

There is an edge  $i \rightarrow j$  if  $j$  is in a recovering set of  $i$ .

Note that  $N(i) = \bigcup_{\ell=1}^t \mathcal{R}_i^\ell$



Let  $C$  be an  $(n, k, r, t)$  LRC code with  $t$  disjoint recovering sets of size  $r$ . Then:

### Theorem 3.1.

*The rate of  $C$  satisfies*

$$\frac{k}{n} \leq \frac{1}{\prod_{j=1}^t (1 + \frac{1}{jr})}$$

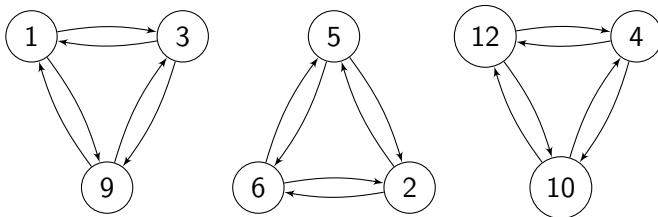
### Theorem 3.2.

*The minimum distance of  $C$  is bounded above as follows*

$$d \leq n - \sum_{i=0}^t \left\lfloor \frac{k-1}{r^i} \right\rfloor$$

Recovery graph for the  $(9, 4, 2)$ -LRC code.

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



Color the edges with  $t$  distinct colors to differentiate recovering sets.

Let  $F$  be a coloring function of the edges:

$$\begin{aligned} F : E(G) &\longrightarrow [t] \\ (i, j) &\longmapsto \ell \quad \text{iff } j \in \mathcal{R}_i^\ell \end{aligned}$$

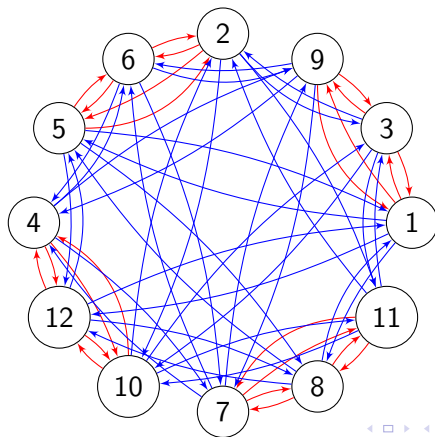
Remark: the out-degree of any vertex  $i \in V$  is  $\sum_{\ell} |\mathcal{R}_i^\ell| = tr$ , and the edges leaving  $i$  are colored in  $t$  colors.

Recovery graph for the  $(12, 4, \{2, 3\})$ -LRC code with edge coloring.

Recall:

$$\mathcal{A} = \{\{1, 5, 12, 8\}, \{2, 10, 11, 3\}, \{4, 7, 9, 6\}\}$$

$$\mathcal{A}' = \{\{1, 3, 9\}, \{2, 6, 5\}, \{4, 12, 10\}, \{7, 8, 11\}\}$$

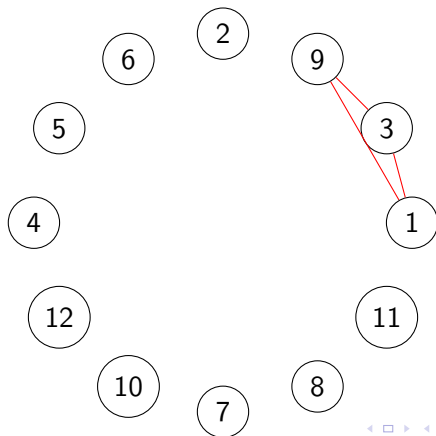


Recovery graph for the  $(12, 4, \{2, 3\})$ -LRC code with edge coloring.

Recall:

$$\mathcal{A} = \{\{1, 5, 12, 8\}, \{2, 10, 11, 3\}, \{4, 7, 9, 6\}\}$$

$$\mathcal{A}' = \{\{1, 3, 9\}, \{2, 6, 5\}, \{4, 12, 10\}, \{7, 8, 11\}\}$$

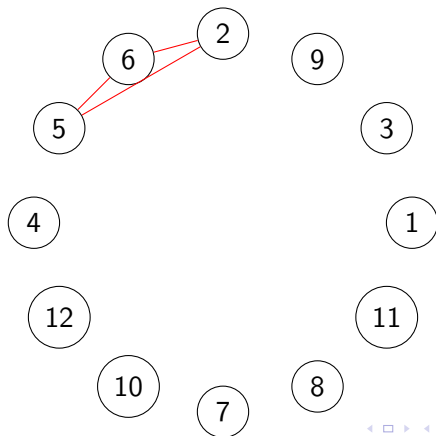


Recovery graph for the  $(12, 4, \{2, 3\})$ -LRC code with edge coloring.

Recall:

$$\mathcal{A} = \{\{1, 5, 12, 8\}, \{2, 10, 11, 3\}, \{4, 7, 9, 6\}\}$$

$$\mathcal{A}' = \{\{1, 3, 9\}, \{2, 6, 5\}, \{4, 12, 10\}, \{7, 8, 11\}\}$$

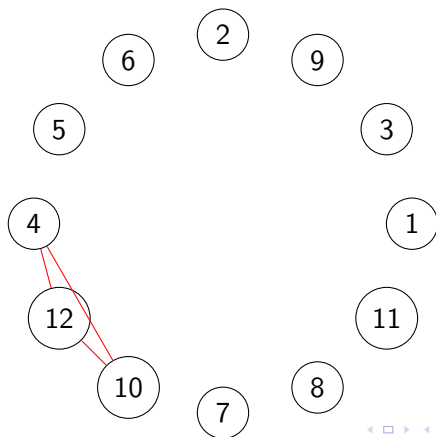


Recovery graph for the  $(12, 4, \{2, 3\})$ -LRC code with edge coloring.

Recall:

$$\mathcal{A} = \{\{1, 5, 12, 8\}, \{2, 10, 11, 3\}, \{4, 7, 9, 6\}\}$$

$$\mathcal{A}' = \{\{1, 3, 9\}, \{2, 6, 5\}, \{4, 12, 10\}, \{7, 8, 11\}\}$$

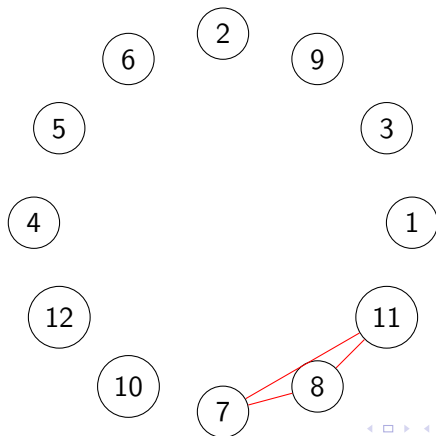


Recovery graph for the  $(12, 4, \{2, 3\})$ -LRC code with edge coloring.

Recall:

$$\mathcal{A} = \{\{1, 5, 12, 8\}, \{2, 10, 11, 3\}, \{4, 7, 9, 6\}\}$$

$$\mathcal{A}' = \{\{1, 3, 9\}, \{2, 6, 5\}, \{4, 12, 10\}, \{7, 8, 11\}\}$$



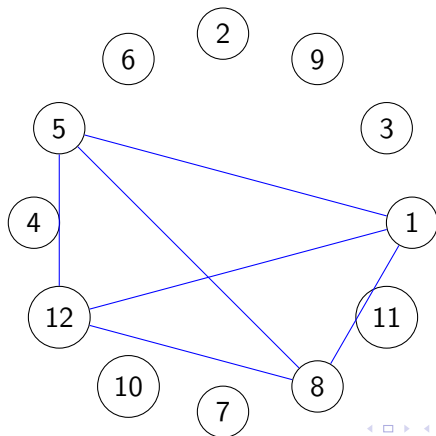


Recovery graph for the  $(12, 4, \{2, 3\})$ -LRC code with edge coloring.

Recall:

$$\mathcal{A} = \{\{1, 5, 12, 8\}, \{2, 10, 11, 3\}, \{4, 7, 9, 6\}\}$$

$$\mathcal{A}' = \{\{1, 3, 9\}, \{2, 6, 5\}, \{4, 12, 10\}, \{7, 8, 11\}\}$$

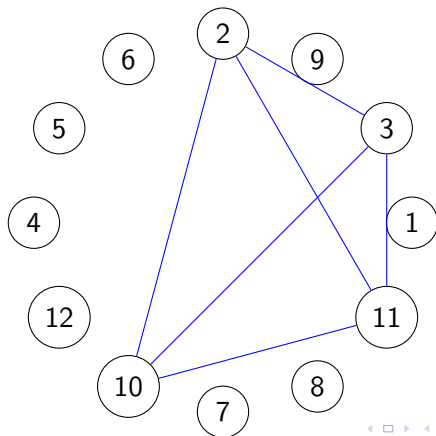


Recovery graph for the  $(12, 4, \{2, 3\})$ -LRC code with edge coloring.

Recall:

$$\mathcal{A} = \{\{1, 5, 12, 8\}, \{2, 10, 11, 3\}, \{4, 7, 9, 6\}\}$$

$$\mathcal{A}' = \{\{1, 3, 9\}, \{2, 6, 5\}, \{4, 12, 10\}, \{7, 8, 11\}\}$$

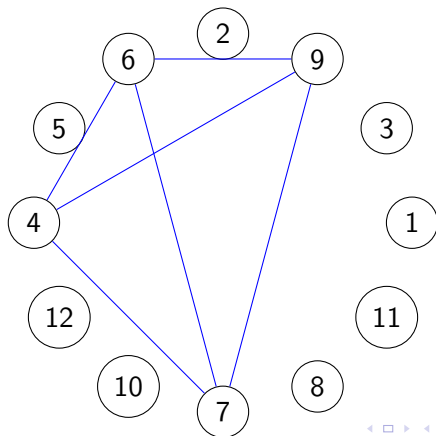


Recovery graph for the  $(12, 4, \{2, 3\})$ -LRC code with edge coloring.

Recall:

$$\mathcal{A} = \{\{1, 5, 12, 8\}, \{2, 10, 11, 3\}, \{4, 7, 9, 6\}\}$$

$$\mathcal{A}' = \{\{1, 3, 9\}, \{2, 6, 5\}, \{4, 12, 10\}, \{7, 8, 11\}\}$$

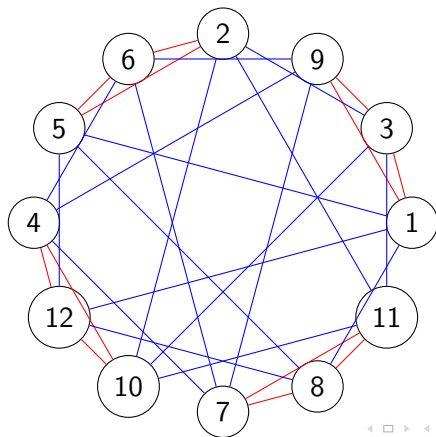


Recovery graph for the  $(12, 4, \{2, 3\})$ -LRC code with edge coloring.

Recall:

$$\mathcal{A} = \{\{1, 5, 12, 8\}, \{2, 10, 11, 3\}, \{4, 7, 9, 6\}\}$$

$$\mathcal{A}' = \{\{1, 3, 9\}, \{2, 6, 5\}, \{4, 12, 10\}, \{7, 8, 11\}\}$$



### Lemma 3.3.

*There exists a subset of the vertices  $U \subseteq V$  of size at least*

$$|U| \geq n \left( 1 - \frac{1}{\prod_{j=1}^t \left( 1 + \frac{1}{jr} \right)} \right)$$

*such that for any  $U' \subseteq U$ ,  $G_{U'}$  has at least one vertex  $v \in U'$  such that its set of outgoing edges is missing at least one color.*

For a given permutation  $\tau$  of the set of vertices  $V$ , define the coloring of some of the vertices.

The color  $j \in [t]$  is assigned to  $v$  if

$$\tau(v) > \tau(m) \quad \forall m \in \mathcal{R}_v^j$$

If this condition is satisfied for several values of  $m$ , the vertex  $v$  is assigned any of these colors.

If this condition is not satisfied at all, the vertex  $v$  is not colored.

Let  $U$  be the set of colored vertices. Consider one of its subsets  $U' \subseteq U$ .

Assume toward a contradiction that every vertex of  $G_{U'}$  has outgoing edges of all  $t$  colors.

Choose a vertex  $v \in U'$  and construct the following walk: if the path ends at some vertex with color  $j$ , choose one of its outgoing edges colored in  $j$ .

$$v_1 \xrightarrow{\text{red}} v_2 \xrightarrow{\text{blue}} v_3 \xrightarrow{\text{green}} \dots \xrightarrow{\text{orange}} v_\ell$$

We can extend this path indefinitely as every vertex has outgoing edges of all  $t$  colors.

Since the graph  $G_{U'}$  is finite, there will be a vertex (call it  $v_1$ ) that is encountered twice.

$$v_1 \xrightarrow{\text{red}} v_2 \xrightarrow{\text{blue}} v_3 \xrightarrow{\text{green}} \dots \xrightarrow{\text{orange}} v_\ell \longrightarrow v_1$$

But:

$$\left. \begin{array}{l} v_i \\ (v_i, v_{i+1}) \end{array} \right\} \begin{array}{l} \text{color } j \\ \text{color } j \end{array} \Bigg\} \implies \tau(v_i) > \tau(v_{i+1})$$

$$\tau(v_1) > \tau(v_2) > \dots > \tau(v_\ell) > \tau(v_1)$$

Contradiction!!!

Then, there must be a vertex in  $U'$  such that its set of outgoing edges is missing one color.



To show that there exists such  $U$  of large cardinality, we choose  $\tau$  uniformly at random among  $\mathfrak{S}_n$  and compute  $E(|U|)$ .

Let  $A_{v,j}$  be the event that the equation  $\tau(v) > \tau(m) \quad \forall m \in \mathcal{R}_v^j$  holds for the vertex  $v$  and the color  $j$ . Since all vertices have  $r$  outgoing edges for each of the  $t$  different colors, the probability  $Pr(A_{v,j})$  does not depend on  $v$ , we write  $A_j := A_{v,j}$ .

Recall that  $v \in U$  if  $v$  is colored with some color  $\ell$ , and is colored with  $\ell$  if  $A_\ell$  happens.

$$Pr(v \in U) = Pr\left(\bigcup_{j=1}^t A_j\right)$$

Inclusion-exclusion formula:

$$Pr(\bigcup_{j=1}^t A_j) = \sum_{\emptyset \neq S \subseteq [t]} (-1)^{|S|-1} Pr(\bigcap_{j \in S} A_j)$$

For that, we need, for any set  $S \subseteq [t]$  the probability that all the  $A_j, j \in S$  occur simultaneously.

$$Pr(\bigcap_{j \in S} A_j) = \frac{1}{|S| r + 1}$$

**Proof.**

For any assignation of values given to  $\{v\} \cup (\bigcup_{j \in S} R_v^j)$ , only 1 out of  $|S| r + 1$  is the maximum. □

$$\begin{aligned} Pr\left(\bigcup_{j=1}^t A_j\right) &= \sum_{\emptyset \neq S \subseteq [t]} (-1)^{|S|-1} Pr\left(\bigcap_{j \in S} A_j\right) \\ &= \sum_{\emptyset \neq S \subseteq [t]} (-1)^{|S|-1} \frac{1}{|S|^{r+1}} = \sum_{j=1}^t (-1)^{j-1} \binom{t}{j} \frac{1}{|S|^{r+1}} \\ &= 1 - \frac{1}{\prod_{j=1}^t \left(1 + \frac{1}{j^r}\right)} \end{aligned}$$

Let  $X_v$  be the indicator for the event that  $v \in U$ .

$$\begin{aligned} E(|U|) &= \sum_{v \in V} E(X_v) = \sum_{v \in V} Pr(v \in U) = n \cdot Pr\left(\bigcup_{j=1}^t A_j\right) \\ &= n\left(1 - \frac{1}{\prod_{j=1}^t \left(1 + \frac{1}{jr}\right)}\right) \end{aligned}$$

Observation: there exists  $\tau \in \mathfrak{S}_n$  for which  $|U| \geq E(|U|)$ .

## Proof of Theorem 3.1

Let  $U \subseteq [n]$  be the set of vertices of cardinality as in the one constructed in Lemma 3.3.

Claim: every coordinate  $i \in U$  can be recovered by accessing the coordinates in  $\bar{U} = [n] \setminus U$ .

By Lemma 3.3, for any  $U' \subseteq U$ ,  $\exists v \in U'$  that is missing one color, say  $\ell$ , in its outgoing edges in  $G'_U$ .

$$\Rightarrow \mathcal{R}_v^\ell \subseteq \bar{U}' \quad \Rightarrow \quad v \text{ can be recovered from } \bar{U}'$$

Suppose the values of the coordinates in  $\bar{U}$  are known.

Consider:

- $U^{(0)} = U$
- $U^{(i+1)} = U^{(i)} \setminus \{v_i\}$   
 s.t.  $R_{v_i}^{\ell_i} \subseteq \overline{U^{(i)}}$  for some  $\ell_i \in [t]$

Every  $v_i$  can be recovered from  $\overline{U^{(i)}}$  which consists of the known values of  $\bar{U}$  and the  $i$  previously recovered values  $v_0, v_1, \dots, v_{i-1}$ .

Conclusion: Every coordinate  $i \in U$  can be recovered from the coordinates in  $\bar{U}$ .

From the last claim, we deduce

$$k \leq |\overline{U}| \leq \frac{n}{\prod_{j=1}^t (1 + \frac{1}{jr})}$$

From where we get a bound on the rate

$$\frac{k}{n} \leq \frac{1}{\prod_{j=1}^t (1 + \frac{1}{jr})}$$



- Definition and properties of LRC

### 3 Bounds on LRC-t codes

- Recovery graph
- Statements of the bounds
- Definitions and examples for the proof
- Proof for the rate bound
- Proof of the minimum distance bound



To be continued ...