Bounds on parameters of LRC-t codes

Petar Hlad Colic

Universitat Politècnica de Catalunya

May 2018

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
- $[n] := \{1, \ldots, n\}$
- 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

Let $\mathcal{C}$ be a code of length n over an alphabet A.

$\mathcal{C}$ is a Locally Recoverable Code (LRC) with locality r if the value of every coordinate $i \in [n]$ of the code is determined by the value of (at most) r other coordinates.

LRC codes can be linear or non linear.

Formal definition

Given $a \in A$ consider the set of codewords with fixed value a at coordinate 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*

$$\forall i \in [n] \text{ there exists } I \subseteq [n] \setminus i, \quad |I| \leq r \qquad s.t.$$

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

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

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

If equality holds, we call $\mathcal{C}$ 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:

- $A_1, \ldots, A_{\frac{n}{r+1}}$ 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

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

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

$$= \begin{pmatrix} x^0 & \dots & x^{r-1} \end{pmatrix} \begin{pmatrix} f_0(x) \\ \vdots \\ f_{r-1}(x) \end{pmatrix}$$

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 1.4.

$$x \in A_i \Rightarrow \quad g(x) \ constant$$

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

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

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 1.5.

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.

$$\deg(f_a(x)) \leq \deg(x^{r-1}) + \deg(g(x)^{\frac{k}{r}-1}) = r - 1 + (r+1)(\frac{k}{r} - 1)$$

$$= k + \frac{k}{r} - 2 \leq n - 2$$

This means that two distinct encoding polynomials give rise to two distinct codevectors. $\Rightarrow$ 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 $\mathbb{F}_q$.

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

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

$$c = (f_a(1), f_a(3), f_a(9), f_a(2), f_a(6), f_a(5), f_a(4), f_a(12), f_a(10))$$

$$= (4, 8, 7, 1, 11, 2, 0, 0, 0)$$

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

$$(f_a(1), f_a(3), f_a(9), f_a(2), f_a(6), f_a(5), f_a(4), f_a(12), f_a(10))$$

$$(4, 8, 7, 1, 11, 2, 0, 0, 0)$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$\delta(1) = 4$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$\delta(2) = 1$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$\delta(4) = 0$$

Example of LRC-2 code

Let $\mathbb{F} = \mathbb{F}_{13}$, $A = \mathbb{F} \setminus \{0\}$
$\mathcal{A} = \{\{1, 5, 12, 8\}, \{2, 10, 11, 3\}, \{4, 7, 9, 6\}\}$
$\mathcal{A}' = \{\{1, 3, 9\}, \{2, 6, 5\}, \{4, 12, 10\}, \{7, 8, 11\}\}$
$f_a(x) = a_0 + a_1 x + a_2 x^4 + a_3 x^6$

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

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

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

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

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

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

Theorem 2.1.

The rate of C satisfies

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

Theorem 2.2.

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

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

Recovery graph

Assume every coordinate i has t disjoint recovering sets $\mathcal{R}_i^1, \ldots, \mathcal{R}_i^t$, 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:

Recovery graph

Assume every coordinate i has t disjoint recovering sets $\mathcal{R}_i^1, \ldots, \mathcal{R}_i^t$, 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 $\mathcal{C}$).

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 $\ell \in [t]$.

Recovery graph

Assume every coordinate i has t disjoint recovering sets $\mathcal{R}_i^1, \ldots, \mathcal{R}_i^t$, 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 $i \to j$ if j is in a recovering set of i.

Recovery graph

Assume every coordinate i has t disjoint recovering sets $\mathcal{R}_i^1, \ldots, \mathcal{R}_i^t$, 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 $i \to j$ if j is in a recovering set of i.

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

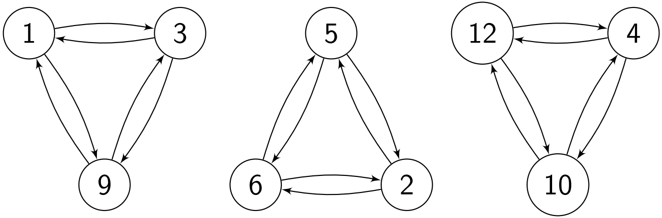
Recovery graph for the $(9, 4, 2)$-LRC code.

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

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

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 differenciate recovering sets.

Color the edges with t distinct colors to differenciate recovering sets.

Let F be a coloring function of the edges:

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

Color the edges with t distinct colors to differenciate recovering sets.

Let F be a coloring function of the edges:

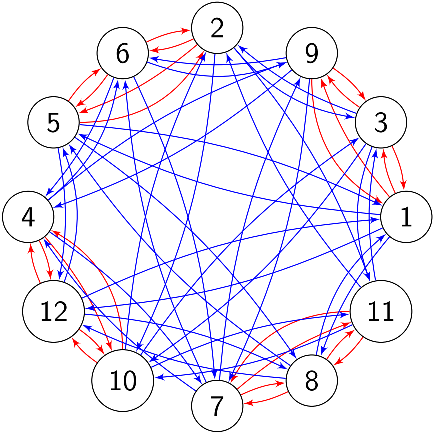
$$\begin{array}{rccll} F: & E(G) & \longrightarrow & [t] & \\ & (i,j) & \longmapsto & \ell & \text{iff } j \in \mathcal{R}_i^\ell \end{array}$$

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.

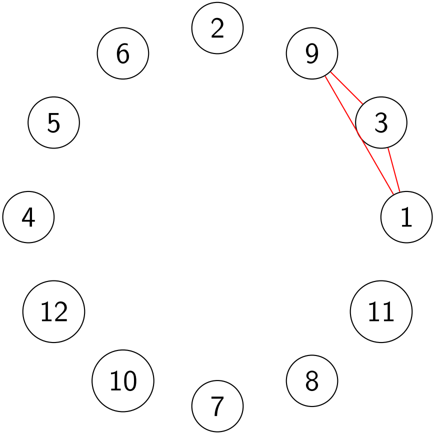
Recovery graph for the $(12, 4, \{2, 3\})$-LRC code with edge coloring.
Recall:

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

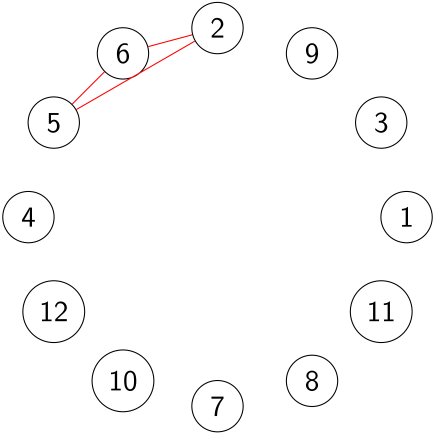
$\mathcal{A}' = \{\{1, 3, 9\}, \{2, 6, 5\}, \{4, 12, 10\}, \{7, 8, 11\}\}$



Recovery graph for the $(12, 4, \{2, 3\})$-LRC code with edge coloring.
Recall:
$\mathcal{A} = \{\{1, 5, 12, 8\}, \{2, 10, 11, 3\}, \{4, 7, 9, 6\}\}$
$\mathcal{A}' = \{\{1, 3, 9\}, \{2, 6, 5\}, \{4, 12, 10\}, \{7, 8, 11\}\}$



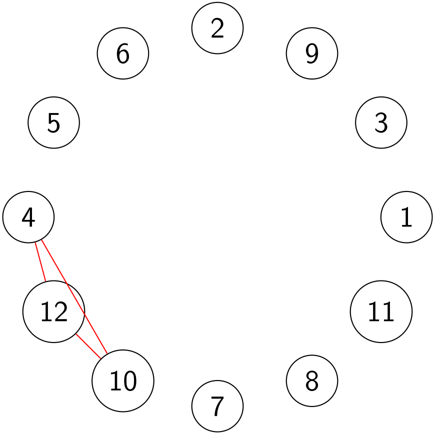
Recovery graph for the $(12, 4, \{2, 3\})$-LRC code with edge coloring.
Recall:
$\mathcal{A} = \{\{1, 5, 12, 8\}, \{2, 10, 11, 3\}, \{4, 7, 9, 6\}\}$
$\mathcal{A}' = \{\{1, 3, 9\}, \{2, 6, 5\}, \{4, 12, 10\}, \{7, 8, 11\}\}$



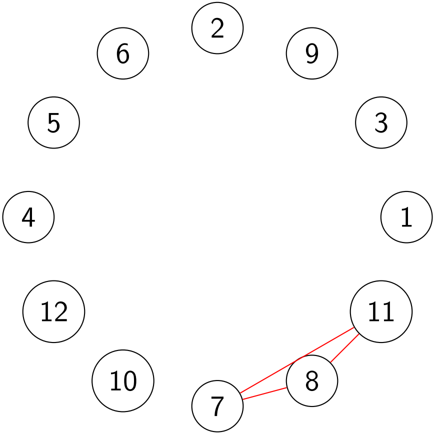
Recovery graph for the $(12, 4, \{2, 3\})$-LRC code with edge coloring.
Recall:

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

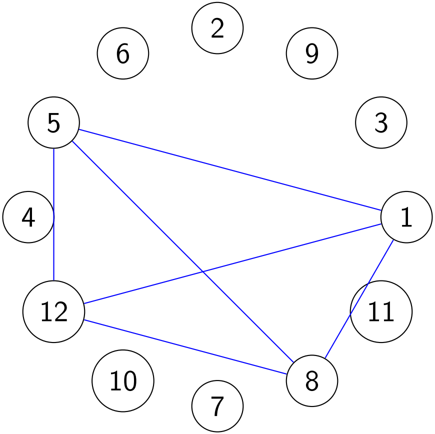
$\mathcal{A}' = \{\{1, 3, 9\}, \{2, 6, 5\}, \{4, 12, 10\}, \{7, 8, 11\}\}$



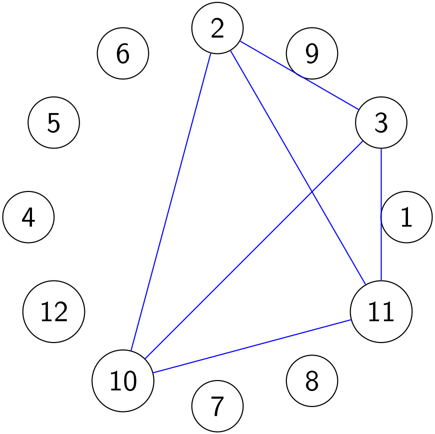
Recovery graph for the $(12, 4, \{2, 3\})$-LRC code with edge coloring.
Recall:
$\mathcal{A} = \{\{1, 5, 12, 8\}, \{2, 10, 11, 3\}, \{4, 7, 9, 6\}\}$
$\mathcal{A}' = \{\{1, 3, 9\}, \{2, 6, 5\}, \{4, 12, 10\}, \{7, 8, 11\}\}$



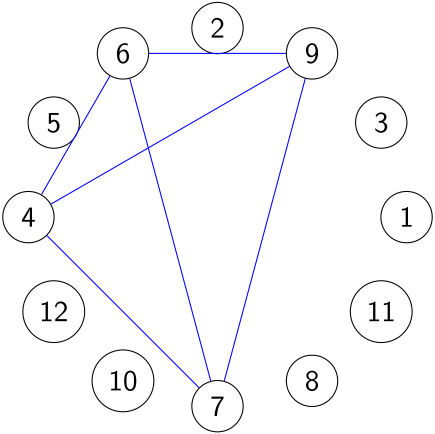
Recovery graph for the $(12, 4, \{2, 3\})$-LRC code with edge coloring.
Recall:
$\mathcal{A} = \{\{1, 5, 12, 8\}, \{2, 10, 11, 3\}, \{4, 7, 9, 6\}\}$
$\mathcal{A}' = \{\{1, 3, 9\}, \{2, 6, 5\}, \{4, 12, 10\}, \{7, 8, 11\}\}$



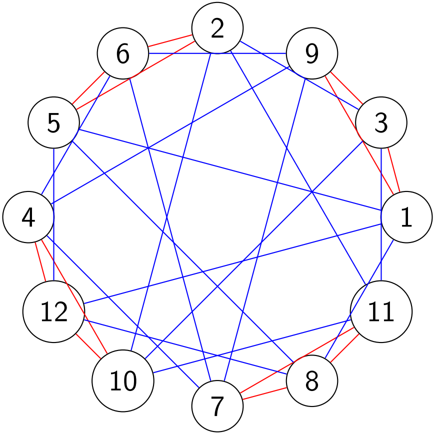
Recovery graph for the $(12, 4, \{2, 3\})$-LRC code with edge coloring.
Recall:
$\mathcal{A} = \{\{1, 5, 12, 8\}, \{2, 10, 11, 3\}, \{4, 7, 9, 6\}\}$
$\mathcal{A}' = \{\{1, 3, 9\}, \{2, 6, 5\}, \{4, 12, 10\}, \{7, 8, 11\}\}$



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



Proof for the rate bound

To prove bound on max. rate:

1. Construct set U of coordinates that can be recovered from $\overline{U} := [n] \setminus U$.
2. Compute lower bound on $|U| \quad \longrightarrow \quad$ upper bound on $\left|\overline{U}\right|$
3. u. bound on $\left|\overline{U}\right| \rightarrow$ u. bound on $k \rightarrow$ u. bound on max. rate

Lemma 2.3.

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.

Implications of Lemma 2.3

1.
Outgoing edges of $v \in U'$ missing color ℓ
$\Rightarrow \mathcal{R}_v^\ell \subseteq V \setminus U'$
$\Rightarrow$ v can be recovered from $V \setminus U'$

2.
$\forall U' \subseteq U, \quad \exists v \in U'$ that can be recovered from $V \setminus U'$
$\Rightarrow$ There is a chain $U = U_0 \supset U_1 \supset \cdots \supset U_{|U|} = \emptyset$
where $U_i \setminus U_{i+1} = v_i$ s.t. v_i can be recovered from $V \setminus U_i$
$\Rightarrow$ U can be recovered from $\overline{U}$

Proof of Lemma 2.3

To construct U:

1. Define a coloring on V in which not all vertices are colored.

Proof of Lemma 2.3

To construct U:

1. Define a coloring on V in which not all vertices are colored.
2. Define U as the set of colored vertices

Proof of Lemma 2.3

To construct U:

1. Define a coloring on V in which not all vertices are colored.
2. Define U as the set of colored vertices
3. Check that U has the desired properties

Proof of Lemma 2.3 (I)

For a given permutation $\tau \in \mathfrak{S}_n$ of 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 not satisfied at all, the vertex v is not colored.

Proof of Lemma 2.3 (II)

Let U be the set of colored vertices. Consider any $U' \subseteq U$.

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

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

$$v_1 \longrightarrow v_2 \longrightarrow v_3 \longrightarrow \ldots \longrightarrow v_\ell$$

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

Proof of Lemma 2.3 (III)

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

$$v_1 \longrightarrow v_2 \longrightarrow v_3 \longrightarrow \dots \longrightarrow v_\ell \longrightarrow v_1$$

Proof of Lemma 2.3 (III)

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

$$v_1 \longrightarrow v_2 \longrightarrow v_3 \longrightarrow \ldots \longrightarrow v_\ell \longrightarrow v_1$$

$$v_i \text{ color j} \qquad \Rightarrow \quad \tau(v_i) > \tau(m) \quad m \in \mathcal{R}^j_{v_i}$$

Proof of Lemma 2.3 (III)

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

$$v_1 \longrightarrow v_2 \longrightarrow v_3 \longrightarrow \dots \longrightarrow v_\ell \longrightarrow v_1$$

$$\begin{array}{rcc} v_i \text{ color j} & \Rightarrow & \tau(v_i) > \tau(m) \quad m \in \mathcal{R}^j_{v_i} \\ (v_i, v_{i+1}) \text{ color j} & \Rightarrow & v_{i+1} \in \mathcal{R}^j_{v_i} \end{array}$$

Proof of Lemma 2.3 (III)

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

$$v_1 \longrightarrow v_2 \longrightarrow v_3 \longrightarrow \ldots \longrightarrow v_\ell \longrightarrow v_1$$

$$\left.\begin{array}{ccc} v_i \text{ color j} & \Rightarrow & \tau(v_i) > \tau(m) \quad m \in \mathcal{R}^j_{v_i} \\ (v_i, v_{i+1}) \text{ color j} & \Rightarrow & v_{i+1} \in \mathcal{R}^j_{v_i} \end{array}\right\} \Rightarrow \tau(v_i) > \tau(v_{i+1})$$

Proof of Lemma 2.3 (III)

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

$$v_1 \longrightarrow v_2 \longrightarrow v_3 \longrightarrow \dots \longrightarrow v_\ell \longrightarrow v_1$$

$$\left.\begin{array}{lcc} v_i \text{ color j} & \Rightarrow & \tau(v_i) > \tau(m) \quad m \in \mathcal{R}^j_{v_i} \\ (v_i, v_{i+1}) \text{ color j} & \Rightarrow & v_{i+1} \in \mathcal{R}^j_{v_i} \end{array}\right\} \Rightarrow \tau(v_i) > \tau(v_{i+1})$$

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

Proof of Lemma 2.3 (III)

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

$$v_1 \longrightarrow v_2 \longrightarrow v_3 \longrightarrow \dots \longrightarrow v_\ell \longrightarrow v_1$$

$$\left.\begin{array}{rcc} v_i \text{ color j} & \Rightarrow & \tau(v_i) > \tau(m) \quad m \in \mathcal{R}^j_{v_i} \\ (v_i, v_{i+1}) \text{ color j} & \Rightarrow & v_{i+1} \in \mathcal{R}^j_{v_i} \end{array}\right\} \Rightarrow \tau(v_i) > \tau(v_{i+1})$$

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

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

Proof of Lemma 2.3 (IV)

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

Proof of Lemma 2.3 (IV)

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

Let $A_{v,j}$ be the event that $\tau(v) > \tau(m) \quad \forall m \in \mathcal{R}_v^j$ holds.

Proof of Lemma 2.3 (IV)

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

Let $A_{v,j}$ be the event that $\tau(v) > \tau(m) \quad \forall m \in \mathcal{R}_v^j$ holds.

Since the probability $Pr(A_{v,j})$ does not depend on v, we write $A_j := A_{v,j}$.

Proof of Lemma 2.3 (IV)

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

Let $A_{v,j}$ be the event that $\tau(v) > \tau(m) \quad \forall m \in \mathcal{R}_v^j$ holds.

Since the probability $Pr(A_{v,j})$ does not depend on v, we write $A_j := A_{v,j}$.

Note that $v \in U$ if A_ℓ happens for some $\ell \in [t]$.

Proof of Lemma 2.3 (IV)

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

Let $A_{v,j}$ be the event that $\tau(v) > \tau(m) \quad \forall m \in \mathcal{R}_v^j$ holds.

Since the probability $Pr(A_{v,j})$ does not depend on v, we write $A_j := A_{v,j}$.

Note that $v \in U$ if A_ℓ happens for some $\ell \in [t]$.

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

Proof of Lemma 2.3 (V)

Inclusion-exclusion formula:

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

Proof of Lemma 2.3 (V)

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 the probability of simultaneous events:

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

Proof of Lemma 2.3 (VI)

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

Proof of Lemma 2.3 (VII)

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

$$E(|U|) = \sum_{v \in V} E(X_v) = \sum_{v \in V} Pr(v \in U) = n \cdot Pr(\bigcup_{j=1}^{t} A_j)$$

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

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

Proof of Theorem 2.1

Let $U \subseteq [n]$ be the set of vertices of cardinality as in the one constructed in Lemma 2.3.

Proof of Theorem 2.1

Let $U \subseteq [n]$ be the set of vertices of cardinality as in the one constructed in Lemma 2.3.

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

Proof of Theorem 2.1

Let $U \subseteq [n]$ be the set of vertices of cardinality as in the one constructed in Lemma 2.3.

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

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

Proof of Theorem 2.1

Let $U \subseteq [n]$ be the set of vertices of cardinality as in the one constructed in Lemma 2.3.

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

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

From where we get a bound on the rate

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

To be continued ...

Consider the recovery graph $G = (V, E)$ of an (n, k, r, t) LRC code $\mathcal{C}$.
Consider the following coloring procedure of V.
Start with $S \subseteq V$ and color it in som fixed color.
For the remaining vertices: a vertex is colored if at least one of its recovering sets is completely colored.
I.e., we color all vertices that can be recovered from S. We denote the set of colored vertices by $\mathrm{Cl}(S)$ and call it the closure of S in G.
Call $|\mathrm{Cl}(S)| \,/\, |S|$ the expansion ratio of the set S.