Cooperative Locality and Availability of the MacDonald Codes for Multiple Symbol Erasures

Zhi Jing and Hong-Yeop Song

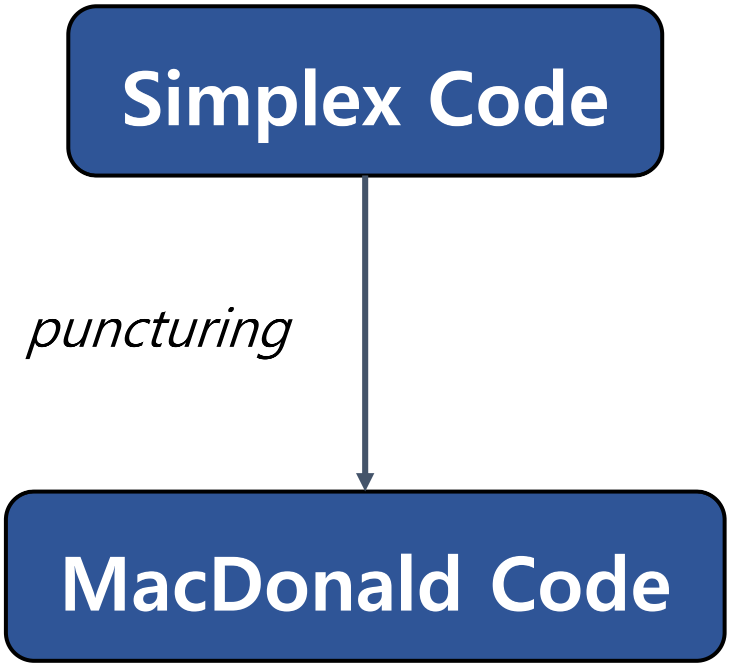
Yonsei University

2020

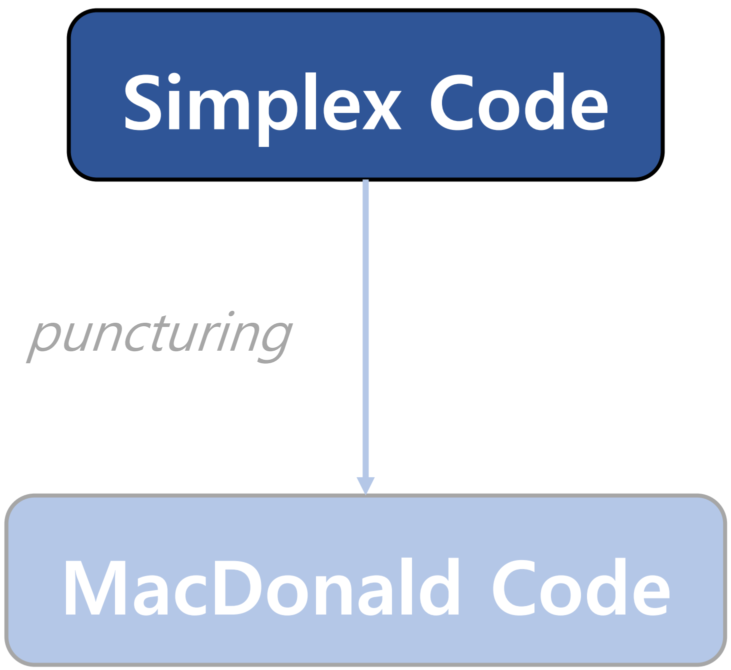
ISITA

Contents

MacDonald Codes



MacDonald Codes



For $(2^k - 1, k)$ Simplex code:

- S_k: generator matrix
- Initialize $S_1 = (1)$, and then

$$S_k = \begin{pmatrix} S_{k-1} & \mathbf{0}_{k-1}^T & S_{k-1} \\ \mathbf{0}_{2^{k-1}-1} & 1 & \mathbf{1}_{2^{k-1}-1} \end{pmatrix}$$

* $\mathbf{0}_n$ and $\mathbf{1}_n$ be all-zero and all one row vector of length n.

MacDonald Codes

Simplex Code

puncturing

MacDonald Code

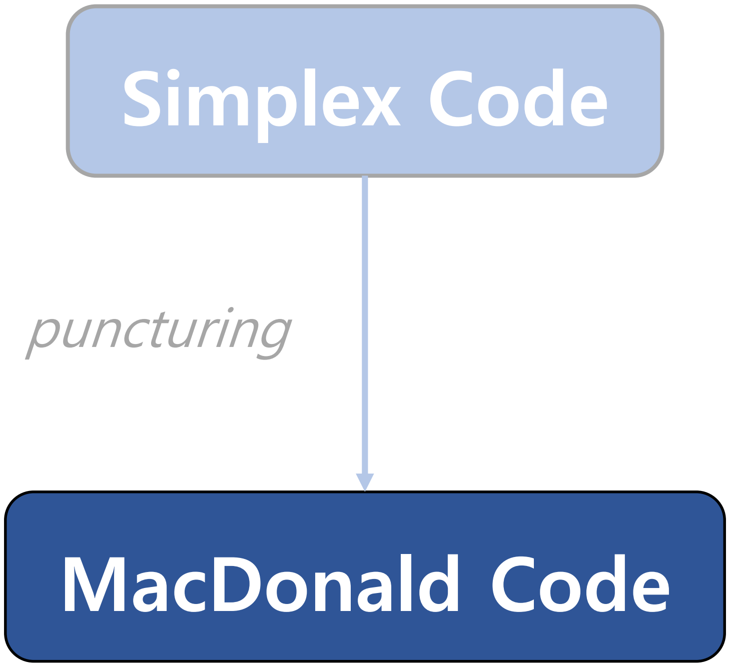
$$S_k = \begin{pmatrix} S_{k-1} & \mathbf{0}_{k-1}^T & S_{k-1} \\ \mathbf{0}_{2^{k-1}-1} & 1 & \mathbf{1}_{2^{k-1}-1} \end{pmatrix}$$

deleting the first $2^l - 1$ *columns* $(1 \le l \le k-1)$

$$G_k(l) = \begin{pmatrix} B & \mathbf{0}_{k-1}^T & A & B \\ \mathbf{0}_{2^{k-1}-2^l} & 1 & \mathbf{1}_{2^l-1} & \mathbf{1}_{2^{k-1}-2^l} \end{pmatrix}$$

* $\mathbf{0}_n$ and $\mathbf{1}_n$ be all-zero and all one row vector of length n.

MacDonald Codes



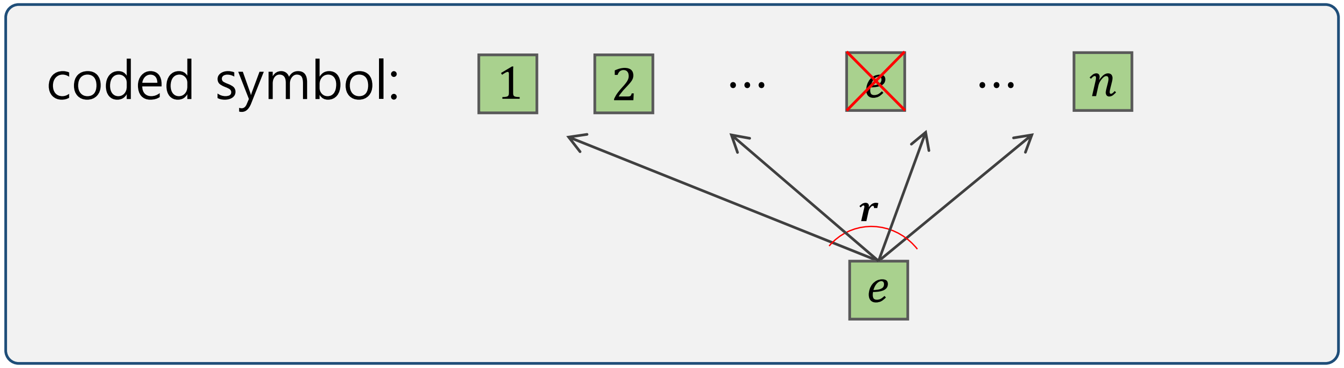
The $(2^k - 2^l, k)$ MacDonald code $M_k(l)$:

$$G_k(l) = \begin{pmatrix} B & \mathbf{0}_{k-1}^T & A & B \\ \mathbf{0}_{2^{k-1}-2^l} & 1 & \mathbf{1}_{2^l-1} & \mathbf{1}_{2^{k-1}-2^l} \end{pmatrix}$$

* $\mathbf{0}_n$ and $\mathbf{1}_n$ be all-zero and all one row vector of length n.

Locality

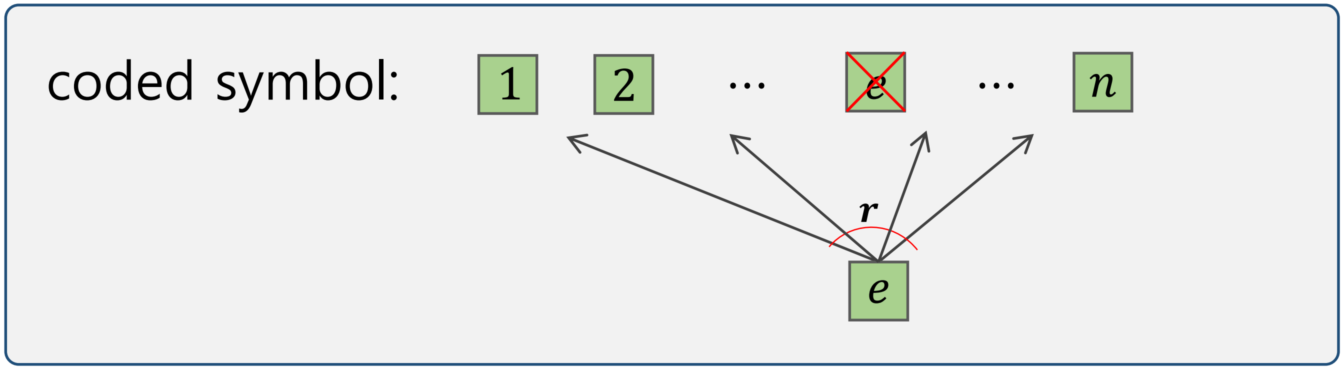
- For an $[n, k]$ code C:



  - **Symbol locality:**
    the smallest number of symbols needed to repair the failed symbol.
  - $[\boldsymbol{n}, \boldsymbol{k}, \boldsymbol{r}]_{\boldsymbol{a}}$ **(**$[\boldsymbol{n}, \boldsymbol{k}, \boldsymbol{r}]_{\boldsymbol{i}}$**) code** $\boldsymbol{C}$**:**
    All coded (information) symbol has the locality at most r.

Locality

- For an $[n, k, r]_a$ code C:



Let $\boldsymbol{u}$ be the nonzero information vector.

$$c_e = c_{i_1} + c_{i_2} + \cdots + c_{i_r}$$
$$\Updownarrow$$
$$\boldsymbol{u} \cdot g_e = \boldsymbol{u} \cdot (g_{i_1} + g_{i_2} + \cdots + g_{i_r})$$

Linear combination of g_j

* g_j, $1 \le j \le n$, is the j^{th} column of the generator matrix of C.

Locality

Lemma 1 [1]:

The locality of the MacDonald code $M_k(l)$ is

$$r = \begin{cases} 2, & l < k-1 \\ 3, & l = k-1 \end{cases}$$

- When $l < k-1$,

$$G_k(l) = \begin{pmatrix} B & \mathbf{0}_{k-1}^T & A & B \\ \mathbf{0}_{2^{k-1}-2^l} & 1 & \mathbf{1}_{2^l-1} & \mathbf{1}_{2^{k-1}-2^l} \end{pmatrix}$$

- When $l = k-1$,

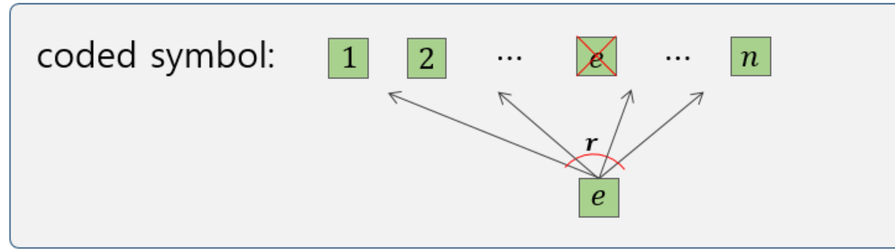
$$G_k(l) = \begin{pmatrix} \mathbf{0}_{k-1}^T & A & B \\ 1 & \mathbf{1}_{2^l-1} & \mathbf{1}_{2^{k-1}-2^l} \end{pmatrix}$$

[1]. Q. Fu, R. Li, L. Guo, and L. Lv, "Locality of optimal binary codes," Finite Fields and Their Applications, vol. 48, pp. 371-394, 2017.

Erasures

- Single erasure
  - ✓ Locality r
- Multiple erasures
  - ✓ Cooperative locality r_h
  - ✓ Availability t

Cooperative Locality



Generalize the locality $r \triangleq r_1$

- **Cooperative locality r_h:**
  - ✓ The smallest number of symbols needed to repair **$h \geq 1$ erased symbols**.
  - ✓ $r_h \leq r_1 \cdot h$

Cooperative Locality

- Code locality:

  All coded (information) symbol has the locality at most r_1.

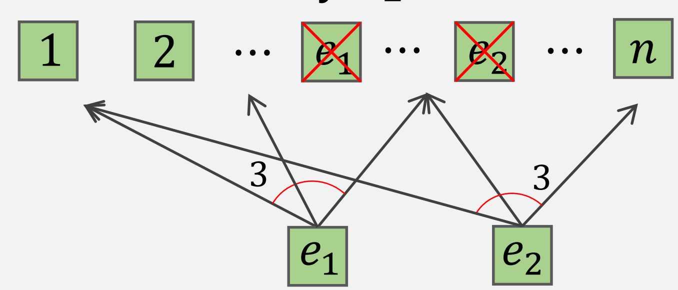
Cooperative Locality

- Code ~~locality:~~ cooperative locality:

  ~~All coded (information) symbol~~ Any h coded (information) symbol has the locality at most ~~r_1~~ r_h.

Eg: (C has the cooperative locality $r_2 = 4 < 6$)

coded symbol:



Cooperative Locality

Theorem 1:

The cooperative locality r_2 of the MacDonald code $M_k(l)$ is

$$r_2 = \begin{cases} 3 \ (< 4 = 2r_1), & l < k-1 \\ 4 \ (< 6 = 2r_1), & l = k-1 \end{cases}$$

Cooperative Locality

Proof:

Let e_i and e_j be two erased symbols.

1) When $l < k-1$,

Case 1:

$$G_k(l) = \begin{pmatrix} B & \mathbf{0}_{k-1}^T & A & B \\ \mathbf{0}_{2^{k-1}-2^l} & 1 & \mathbf{1}_{2^l-1} & \mathbf{1}_{2^{k-1}-2^l} \end{pmatrix}$$

$$g_i = (\boldsymbol{u}\ 0)$$

$$g_j = (\boldsymbol{v}\ 0)$$

Cooperative Locality

Proof:

Let e_i and e_j be two erased symbols.

1) When $l < k - 1$,

Case 1:

$$G_k(l) = \begin{pmatrix} B & \mathbf{0}_{k-1}^T & A & B \\ \mathbf{0}_{2^{k-1}-2^l} & 1 & \mathbf{1}_{2^l-1} & \mathbf{1}_{2^{k-1}-2^l} \end{pmatrix}$$

$$g_i = (\boldsymbol{u}\ 0)$$
$$g_j = (\boldsymbol{v}\ 0)$$

$$(\boldsymbol{w}\ 1)$$

Cooperative Locality

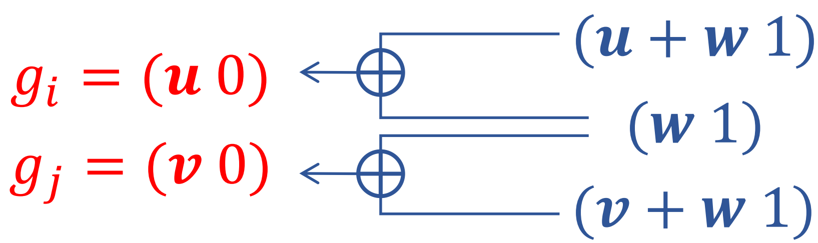
Proof:

Let e_i and e_j be two erased symbols.

1) When $l < k - 1$,

Case 1:

$$G_k(l) = \begin{pmatrix} B & \mathbf{0}_{k-1}^T & A & B \\ \mathbf{0}_{2^{k-1}-2^l} & 1 & \mathbf{1}_{2^l-1} & \mathbf{1}_{2^{k-1}-2^l} \end{pmatrix}$$



$$\therefore \boldsymbol{r_2 = 3}$$

Cooperative Locality

Proof:

Let e_i and e_j be two erased symbols.

1) When $l < k-1$,

Case 2:

$$G_k(l) = \begin{pmatrix} B & \mathbf{0}_{k-1}^T & A & B \\ \mathbf{0}_{2^{k-1}-2^l} & 1 & \mathbf{1}_{2^l-1} & \mathbf{1}_{2^{k-1}-2^l} \end{pmatrix}$$

$$g_i = (\boldsymbol{u}\ 1)$$
$$g_j = (\boldsymbol{v}\ 1)$$

Cooperative Locality

Proof:

Let e_i and e_j be two erased symbols.

1) When $l < k-1$,

Case 2:

$$G_k(l) = \begin{pmatrix} B & \mathbf{0}_{k-1}^T & A & B \\ \mathbf{0}_{2^{k-1}-2^l} & 1 & \mathbf{1}_{2^l-1} & \mathbf{1}_{2^{k-1}-2^l} \end{pmatrix}$$

$g_i = (\boldsymbol{u}\ 1)$

$g_j = (\boldsymbol{v}\ 1)$

$(\boldsymbol{w}\ 0)$

and $\boldsymbol{w} \neq \boldsymbol{u} + \boldsymbol{v}$

Cooperative Locality

Proof:

Let e_i and e_j be two erased symbols.

1) When $l < k-1$,

Case 2:

$$G_k(l) = \begin{pmatrix} B & \mathbf{0}_{k-1}^T & A & B \\ \mathbf{0}_{2^{k-1}-2^l} & 1 & \mathbf{1}_{2^l-1} & \mathbf{1}_{2^{k-1}-2^l} \end{pmatrix}$$

$$(\boldsymbol{u}+\boldsymbol{w}\ 1) \oplus (\boldsymbol{w}\ 0) \rightarrow g_i = (\boldsymbol{u}\ 1)$$

$$(\boldsymbol{w}\ 0) \oplus (\boldsymbol{v}+\boldsymbol{w}\ 1) \rightarrow g_j = (\boldsymbol{v}\ 1)$$

$$\therefore \boldsymbol{r_2} = \mathbf{3}$$

Cooperative Locality

Proof:

Let e_i and e_j be two erased symbols.

1) When $l < k-1$,

Case 3:

$$G_k(l) = \begin{pmatrix} B & \mathbf{0}_{k-1}^T & A & B \\ \mathbf{0}_{2^{k-1}-2^l} & 1 & \mathbf{1}_{2^l-1} & \mathbf{1}_{2^{k-1}-2^l} \end{pmatrix}$$

$g_i = (\boldsymbol{u}\ 0)$ $\quad$ $g_j = (\boldsymbol{v}\ 1)$

Cooperative Locality

Proof:

Let e_i and e_j be two erased symbols.

1) When $l < k-1$,

Case 3:

$$G_k(l) = \begin{pmatrix} B & \mathbf{0}_{k-1}^T & A & B \\ \mathbf{0}_{2^{k-1}-2^l} & 1 & \mathbf{1}_{2^l-1} & \mathbf{1}_{2^{k-1}-2^l} \end{pmatrix}$$

$g_i = (\boldsymbol{u}\ 0)$ $\qquad$ $g_j = (\boldsymbol{v}\ 1)$

$(\boldsymbol{v} + \boldsymbol{w}\ 0)$

Cooperative Locality

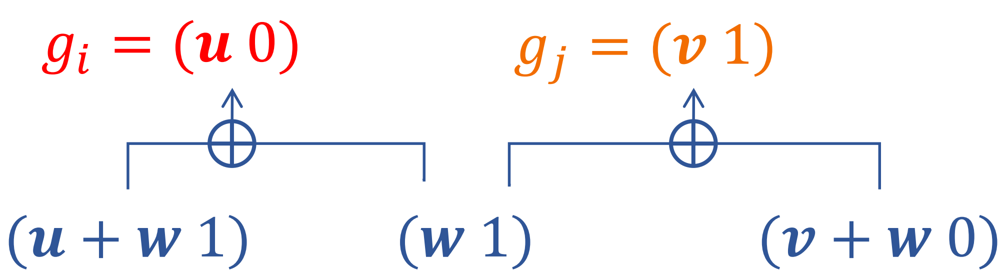
Proof:

Let e_i and e_j be two erased symbols.

1) When $l < k - 1$,

Case 3:

$$G_k(l) = \begin{pmatrix} B & \mathbf{0}_{k-1}^T & A & B \\ \mathbf{0}_{2^{k-1}-2^l} & 1 & \mathbf{1}_{2^l-1} & \mathbf{1}_{2^{k-1}-2^l} \end{pmatrix}$$



$$\therefore \boldsymbol{r_2} = \mathbf{3}$$

Cooperative Locality

Proof:

Let e_i and e_j be two erased symbols.

2) When $l = k - 1$,

$$G_k(l) = \begin{pmatrix} \mathbf{0}_{k-1}^T & A & B \\ 1 & \mathbf{1}_{2^l-1} & \mathbf{1}_{2^{k-1}-2^l} \end{pmatrix}$$

$g_i = (\boldsymbol{u}\ 1)$ $g_j = (\boldsymbol{v}\ 1)$

Cooperative Locality

Proof:

Let e_i and e_j be two erased symbols.

2) When $l = k-1$,

$$G_k(l) = \begin{pmatrix} \mathbf{0}_{k-1}^T & A & B \\ 1 & \mathbf{1}_{2^l-1} & \mathbf{1}_{2^{k-1}-2^l} \end{pmatrix}$$

$g_i = (\boldsymbol{u}\ 1)$

$(\boldsymbol{a}\ 1)$

$(\boldsymbol{b}\ 1)$

and $\boldsymbol{a} \neq \boldsymbol{b} \neq \boldsymbol{u}, \boldsymbol{v}$

$\boldsymbol{a} + \boldsymbol{b} \neq \boldsymbol{u} + \boldsymbol{v}$

$g_j = (\boldsymbol{v}\ 1)$

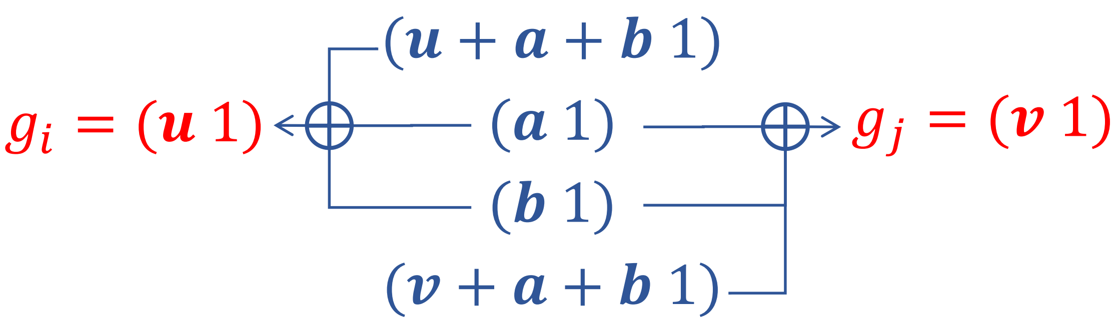
Cooperative Locality

Proof:

Let e_i and e_j be two erased symbols.

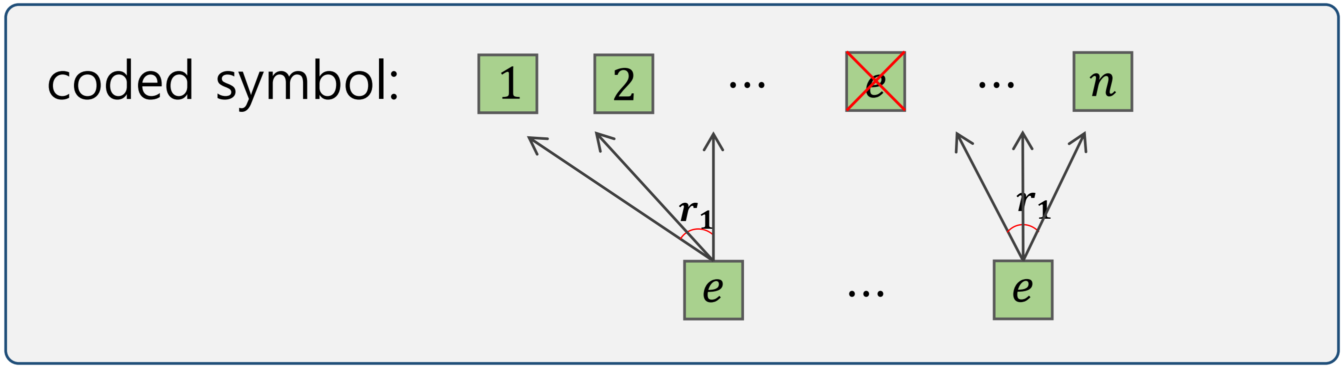
2) When $l = k - 1$,

$$G_k(l) = \begin{pmatrix} \mathbf{0}_{k-1}^T & A & B \\ 1 & \mathbf{1}_{2^l-1} & \mathbf{1}_{2^{k-1}-2^l} \end{pmatrix}$$



$$\therefore \boldsymbol{r_2 = 4}$$

Availability



- **Availability t of symbol c_i:**
  - ✓ The largest number of the **disjoint** repair sets.
  - ✓ $|R_\tau(i)| \leq r_1, 1 \leq \tau \leq t$
- **Code availability:**
  - ✓ All coded (information) symbol has at least t disjoint repair set at most r_1.
  - ✓ $(r_1, t)_a$ / $(r_1, t)_i$

Availability

Theorem 2:

The MacDonald code $M_k(l)$, $k \geq 3$, are LRCs with all-symbol availability

$$\begin{cases} (r_1, t)_a = \left(2, 2^{k-1} - 2^l\right)_a, & l < k-1 \\ (r_1, t)_a = \left(3, \dfrac{2^{k-1}-1}{3}\right)_a, & l = k-1, k \text{ is odd} \end{cases}$$

Lemma 2

Availability

Proof:

1) $l < k - 1$:

$[2^k - 1, k]$ Simplex code:

$$(r_1, t)_a = \left(2, 2^{k-1} - 1\right)_a \text{ [2]}$$

For any symbol s_i,

$$\left|\{i\} \cup R_1(i) \cup \cdots \cup R_{2^{k-1}-1}(i)\right|$$

$$= 1 + 2 \times \left(2^{k-1} - 1\right) = 2^k - 1$$

The repair sets cover all other symbols.

[2]. M. Kuijper and D, Napp, "Erasure codes with simplex locality," [Online.] Available:http://arxiv.org/abs/1403.2779

Availability

Proof:

1) $l < k-1$:

$$S_k = \begin{pmatrix} S_{k-1} & \mathbf{0}_{k-1}^T & S_{k-1} \\ \mathbf{0}_{2^{k-1}-1} & 1 & \mathbf{1}_{2^{k-1}-1} \end{pmatrix}$$

$$= \begin{pmatrix} A & B & \mathbf{0}_{k-1}^T & A & B \\ \mathbf{0}_{2^l-1} & \mathbf{0}_{2^{k-1}-2^l} & 1 & \mathbf{1}_{2^l-1} & \mathbf{1}_{2^{k-1}-2^l} \end{pmatrix}$$

the last $k-l$ elements of the column are 0

the last $k-l$ elements of the column have at least one 1

Availability

Proof:

1) $l < k-1$:

$$S_k = \begin{pmatrix} S_{k-1} & \mathbf{0}_{k-1}^T & S_{k-1} \\ \mathbf{0}_{2^{k-1}-1} & 1 & \mathbf{1}_{2^{k-1}-1} \end{pmatrix}$$

$$= \begin{pmatrix} A & B & \mathbf{0}_{k-1}^T & A & B \\ \mathbf{0}_{2^l-1} & \mathbf{0}_{2^{k-1}-2^l} & 1 & \mathbf{1}_{2^l-1} & \mathbf{1}_{2^{k-1}-2^l} \end{pmatrix}$$

the last $k-l$ elements of the column are 0

the last $k-l$ elements of the column have at least one 1

each repair set of the symbol contains at most one element that belongs to $[2^l-1]$.

Availability

Proof:

1) $l < k-1$:

$$S_k = \begin{pmatrix} S_{k-1} & \mathbf{0}_{k-1}^T & S_{k-1} \\ \mathbf{0}_{2^{k-1}-1} & 1 & \mathbf{1}_{2^{k-1}-1} \end{pmatrix}$$

$$= \begin{pmatrix} A & B & \mathbf{0}_{k-1}^T & A & B \\ \mathbf{0}_{2^l-1} & \mathbf{0}_{2^{k-1}-2^l} & 1 & \mathbf{1}_{2^l-1} & \mathbf{1}_{2^{k-1}-2^l} \end{pmatrix} = \boldsymbol{G_k(l)}$$

each symbol has $2^{k-1} - 2^l$ repair sets

Availability

Proof:

2) $l = k - 1$ and k is odd integer:

All the symbol have the same number of disjoint repair sets.

The availability of code

$\Downarrow$

The availability of the first symbol of $M_k(k-1)$

Availability

Proof:

2) $l = k - 1$ and k is odd integer:

Base case:

When $k = 3$, the generator matrix of $M_3(2)$ is

$$G_3(2) = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

Obviously, $(\mathrm{r}_1, t) = (3,1 = \frac{2^{3-1}-1}{3})$

Availability

Proof:

2) $l = k - 1$ and k is odd integer:

Induction step:

Assume $M_m(m-1)$ has the availability $(r_1, t)_a = \left(3, h = \frac{2^{m-1}-1}{3}\right)_a$ for a given odd m.

$$g_1 = g_{\alpha_i} + g_{\beta_i} + g_{\gamma_i}, \text{ where } 1 \leq i \leq h$$

Availability

Proof:

2) $l = k - 1$ and k is odd integer:

Induction step:

Assume $M_m(m-1)$ has the availability $(r_1, t)_a = \left(3, h = \frac{2^{m-1}-1}{3}\right)_a$ for a given odd m.

$$g_1 = g_{\alpha_i} + g_{\beta_i} + g_{\gamma_i}, \text{ where } 1 \le i \le h$$

Availability

Proof:

2) $l = k - 1$ and k is odd integer:

Induction step:

$$\begin{pmatrix} 0 \\ 0 \\ g_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ g_{\alpha_i} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ g_{\beta_i} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ g_{\gamma_i} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ g_{\alpha_i} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ g_{\beta_i} \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ g_{\gamma_i} \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 0 \\ g_{\alpha_i} \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ g_{\beta_i} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ g_{\gamma_i} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ g_{\alpha_i} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ g_{\beta_i} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ g_{\gamma_i} \end{pmatrix}$$

Availability

Proof:

2) $l = k - 1$ and k is odd integer:

Induction step:

$$\begin{pmatrix} 0 \\ 0 \\ g_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ g_1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ g_1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ g_1 \end{pmatrix}$$

$\begin{pmatrix} 0 \\ 0 \\ g_1 \end{pmatrix}$ has $4h + 1$ disjoint linear combination.

$3 \times (4h + 1) = 2^{m+1} - 1 \Longrightarrow$ No more repair sets

$$4h + 1 = 4 \times \frac{2^{m-1} - 1}{3} + 1 = \frac{2^{m+1} - 1}{3}$$

Optimal

Theorem 3:

The MacDonald code $M_k(l)$ are both the optimal LRCs with all-symbol availability and the optimal LRCs with information availability only when $l \leq k-1$, $k=3$ and $l=3$, $k=4$.

All-symbol availability [3]:

$$d \leq n - \sum_{i=0}^{t} \left\lfloor \frac{k-1}{r_1^i} \right\rfloor$$

Information availability [4]:

$$d \leq n - k + 2 - \left\lceil \frac{(k-1)t+1}{(r_1-1)t+1} \right\rceil$$

[3]. I. Tamo and A. Barg, "Bounds on locally recoverable codes with multiple recovering sets," in Proc. IEEE Int. Symp. Inf. Theory (ISIT), pp. 691–695, Jun./Jul. 2014.

[4]. A. Wang and Z. Zhang, "Repair locality with multiple erasure tolerance," IEEE Trans. Inf. Theory, vol. 60, no. 11, pp. 6979–6987, Nov. 2014.

Conclusion

In this paper,

- Calculate the cooperative locality r_2 and the availability t of the MacDonald codes.
- Show its optimization when $k = 3$ and 4.