

Improved Maximally Recoverable LRCs using Skew Polynomials

Sivakanth Gopi*

Venkatesan Guruswami†

Abstract

An (n, r, h, a, q) -LRC is a linear code over \mathbb{F}_q of length n , whose codeword symbols are partitioned into n/r local groups each of size r . Each local group satisfies ‘ a ’ local parity checks to recover from ‘ a ’ erasures in that local group and there are further h global parity checks to provide fault tolerance from more global erasure patterns. Such an LRC is Maximally Recoverable (MR), if it can correct all erasure patterns which are information-theoretically correctable given this structure—these are precisely patterns with up to ‘ a ’ erasures in each local group and an additional h erasures anywhere in the codeword.

We give an explicit construction of (n, r, h, a, q) -MR LRCs with field size q bounded by $(O(\max\{r, n/r\}))^{\min\{h, r-a\}}$. This significantly improves upon known constructions in most parameter ranges. Moreover, it matches the best known lower bound from [GGY20] in an interesting special case when $r = \Theta(\sqrt{n})$ and h, a are constants with $h \leq a + 2$, achieving the optimal field size of $\Theta_{a,h}(n^{h/2})$. Our construction is based on the theory of skew polynomials.

1 Introduction

In distributed storage such as in data centers, data is partitioned and stored in individual servers; each with a small storage capacity of a few terabytes. A server can crash any time losing all the data it contains. More often than a crash, a server might become temporarily unavailable either due to system updates, network bottlenecks or it might be busy serving requests of some other user. Thus there are two design objectives for a distributed storage system. The first one is to never lose user data in the event of crashes (or at least make it highly improbable). The second is to service user requests with low latency despite some servers becoming temporarily unavailable. Instead of just replicating data which is wasteful, distributed storage systems use erasure codes. Using a Reed-Solomon code, if we add $n - k$ parity check servers to k data servers, we can recover user data from any k available servers. But as k gets larger, this doesn’t satisfy our second objective of servicing user requests with low latency. Local Reconstruction Codes (LRCs) were invented precisely for achieving both the objectives while still maintaining storage efficiency and have been implemented in several large scale systems such as Microsoft Azure [HSX⁺12] and Hadoop [SAP⁺13]. These codes have *locality* which means that they can recover quickly from a small number of erasures by reading only a small number of available servers. But at the same time, they can also recover from the unlikely event of a large number of erasures (but can do so less efficiently). Locality in distributed storage was first introduced in [HCL07, CHL07], but LRCs were first formally defined and studied in [GHSY12] and [PD14]. We will now define them formally.

*Microsoft Research. Email: sigopi@microsoft.com.

†Carnegie Mellon University. Email: venkatg@cs.cmu.edu. Research supported in part by NSF grants CCF-1563742 and CCF-1814603.

An (n, r, h, a, q) -LRC is a linear code over \mathbb{F}_q of length n , whose codeword symbols are partitioned into n/r local groups each of size r . The coordinates in each local group satisfy ‘ a ’ local parity checks and there are further h global parity checks that all the n coordinates satisfy. The local parity checks are used to recover from up to ‘ a ’ erasures in a local group by reading at most $r - a$ symbols in that local group. The h global parities are used to correct more global erasure patterns which involve more than a erasures in each local group. The parity check matrix H of an (n, r, h, a, q) -LRC has the structure shown in Equation 1.

$$H = \left[\begin{array}{c|c|c|c} A_1 & 0 & \cdots & 0 \\ \hline 0 & A_2 & \cdots & 0 \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline 0 & 0 & \cdots & A_g \\ \hline B_1 & B_2 & \cdots & B_g \end{array} \right]. \quad (1)$$

Here $g = n/r$ is the number of local groups. A_1, A_2, \dots, A_g are $a \times r$ matrices over \mathbb{F}_q which correspond to the local parity checks that each local group satisfies. B_1, B_2, \dots, B_g are $h \times r$ matrices over \mathbb{F}_q and together they represent the h global parity checks that the codewords should satisfy.

Equivalently, from an encoding point of view, an (n, r, h, a, q) -LRC is obtained by adding h global parity checks to k data symbols, partitioning these $k + h$ symbols into local groups of size r , and then adding ‘ a ’ local parity checks for each local group. As a result we have $n = k + h + a \cdot \frac{k+h}{r}$ codeword symbols. This is shown in Figure 1.

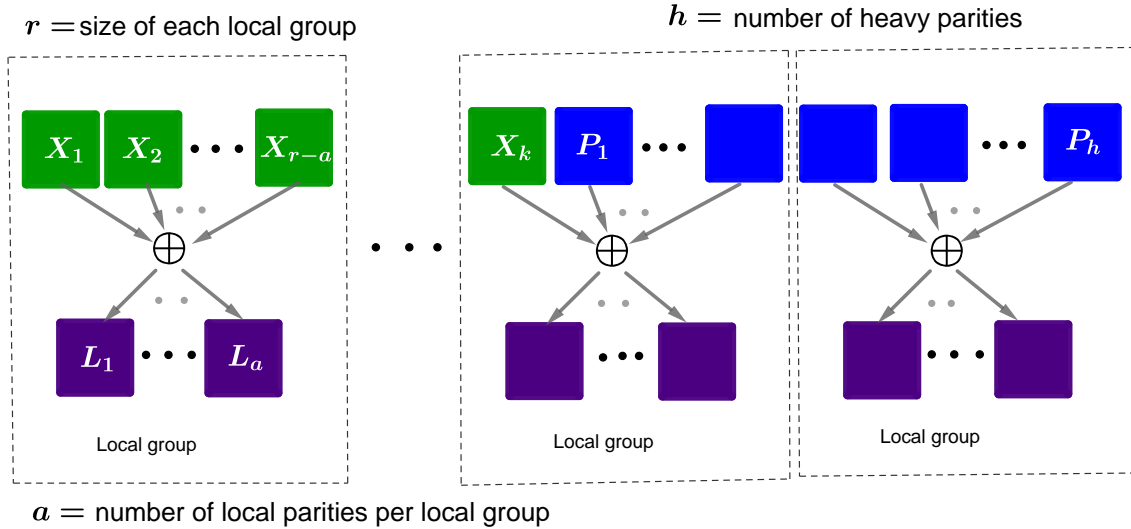


Figure 1: An LRC with k data symbols, h heavy parities and ‘ a ’ local parities per local group. The length of the code $n = k + h + a \cdot \frac{k+h}{r}$.

Information-theoretically, one can show that we can at best hope to correct an additional h erasures distributed across global groups on top of the ‘ a ’ erasures in each local group. LRCs which can correct all such erasure patterns which are information-theoretically possible to correct are called *Maximally Recoverable (MR) LRCs*. The notion of maximal recoverability was first introduced by [CHL07, HCL07] and extended to more general settings in [GHJY14]. But MR LRCs

were specifically studied first by [BHH13, Bla13] where they are called *Partial-MDS (Maximum Distance Separable) codes*.

Definition 1.1. Let C be an arbitrary (n, r, h, a, q) -local reconstruction code. We say that C is *maximally recoverable* if:

1. Any set of ‘a’ erasures in a local group can be corrected by reading the rest of the $r - a$ symbols in that local group.
2. Any erasure pattern $E \subseteq [n]$, $|E| = ga + h$, where E is obtained by selecting a symbols from each of g local groups and h additional symbols arbitrarily, is correctable by the code C .

For a code C with parity check matrix H , an erasure pattern E is correctable iff the submatrix of H formed by columns corresponding the coordinates in E has full column rank. Therefore, we have the following characterization of an MR LRC in terms of its parity check matrix.

Proposition 1.2. An (n, r, h, a, q) -LRC with parity check matrix given by H from Equation 1 is *maximally recoverable* iff:

1. Each of the local parity check matrices A_i are the parity check matrices of an MDS code, i.e., any a columns of A_i are linearly independent.
2. Any submatrix of H which can be formed by selecting a columns in each local group and an additional h columns has full column rank.

It is known that MR-LRCs exist over exponentially large fields [GHSY12]. This can be seen by instantiating the parity check matrix H from Equation 1 randomly from an exponentially large field and verifying that the condition in Proposition 1.2 is satisfied with high probability by Schwartz-Zippel lemma. But codes deployed in practice require small fields for computational efficiency, typically fields such as \mathbb{F}_{2^8} or $\mathbb{F}_{2^{16}}$ are preferred. Therefore a lot of prior work focused on explicit constructions of MR LRCs over small fields.

1.1 Prior Work

Upper Bounds. There are several known constructions of MR LRCs which are incomparable to each other in terms of the field size [GHJY14, GYBS17, GJX20, MK19, GGY20, Bla13, TPD16, HY16, GHK⁺17, CK17, BPSY16]. Some constructions are better than others based on the range of parameters. A few of the important ones are shown in Table 1.1. The table is divided into two parts. The first part shows constructions which work for all ranges of parameters and the second part shows constructions which work for some special cases. The first bound by [GYBS17] is good when r is close to n . The second bound by [GJX20] is better when $h \ll r \ll n$. The bound by [MK19] is better when $r - a \leq h$. The construction in [MK19] is also significantly different from the previous constructions and our construction is inspired by the construction in [MK19]. Finally, the bound in [GHJY14] is best when $a = 1$ and $h \leq r = O(1)$ are constants (we note that the implicit constant hidden in $O_r(\cdot)$ has an exponential dependence in r). In the special case when $h = 2$, a construction over linear sized fields for all ranges of other parameters is given in [GGY20].

Lower Bounds. The best known lower bounds on the field size required for (n, r, h, a, q) -MR LRCs (with $g = n/r$ local groups) is from [GGY20] who show that

$$q \geq \Omega_{h,a}(n \cdot r^\alpha) \text{ where } \alpha = \frac{\min\{a, h - 2\lceil h/g \rceil\}}{\lceil h/g \rceil}. \quad (2)$$

Field size q	
$O(r \cdot n^{(a+1)h-1})$	[GYBS17]
$\max(O(n/r), O(r)^{\min\{r, h+a\}})^{\min\{h, g\}}$	[GJX20]
$(O(\max\{n/r, r\}))^{r-a}$	[MK19]
$O_r\left(n^{\lceil (h-1)(1-1/2^r) \rceil}\right)$ when $a = 1$ and $r = O(1)$	[GHJY14]
$O(r)$ when $h = 0$ or $h = 1$	[BHH13]
$O(n)$ when $h = 2$	[GGY20]
$\tilde{O}(n)$ when $h = 3, a = 1, r = 3$	[GGY20]

Table 1: Table showing the best known upper bounds on the field size of (n, r, h, a, q) -MR LRCs.

The lower bound (2) simplifies to

$$q \geq \Omega_{h,a} \left(n r^{\min\{a, h-2\}} \right) \quad (3)$$

when $g = n/r \geq h$. The hidden constant in (3) have exponential dependence in $\min\{a, h\}$.

1.2 Our Results

We are now ready to present our main result.

Theorem 1.3 (Main). *Let $q_0 \geq \max\{g + 1, r - 1\}$ be any prime power where $g = n/r$ is the number of local groups. Then there exists an explicit (n, r, h, a, q) -MR LRC with $q = q_0^{\min\{h, r-a\}}$. Asymptotically, the field size satisfies*

$$q \leq (O(\max\{r, n/r\}))^{\min\{h, r-a\}}. \quad (4)$$

Our construction is better than the first three bounds in Table 1.1 for *all* parameter ranges. Moreover when a, h are constants with $h \leq a + 2$ and $r = \Theta(\sqrt{n})$, our construction matches the lower bound in Equation (3), achieving the optimal field size of $\Theta_h(n^{h/2})$. This is first non-trivial case (other than when $h = 2$ [GGY20]) where we know the optimal field size for MR LRCs.

Corollary 1.4. *Suppose $r = g = \Theta(\sqrt{n})$ and $h \leq a + 2$ with a, h being fixed constants independent of n . Then the optimal field size of an $(n, r = \Theta(\sqrt{n}), h, a, q)$ -LRC is $q = \Theta_{a,h}(n^{h/2})$.*

We note that our construction is worse compared to the constructions in the second part of Table 1.1 which work for some special setting of parameters.

MR LRCs used in practice typically have only 2 or 3 local groups i.e. $g = n/r$ is typically a constant [HSX⁺12]. We can further improve the construction from Theorem 1.3 in this regime, in the special case when the number of local parities $a = 1$.

Theorem 1.5. *Suppose the number of local groups $g = n/r$ is some fixed constant and the number of local parities $a = 1$. Let $q_0 \geq g + 1$ be any prime power and let s be such that $q_0^s \geq r$. Then there exists an explicit $(n, r, h, a = 1, q)$ -LRC with field size $q = q_0^{s \lceil \min\{h, r-1\}(1-1/q_0) \rceil}$. Asymptotically, the field size satisfies*

$$q \leq (O(n))^{\lceil \min\{h, r-1\}(1-1/q_0) \rceil}.$$

Our Techniques. Our constructions are based on the theory of skew polynomials and is inspired by the construction from [MK19]. Skew polynomials are a non-commutative generalization of polynomials, but they retain many of the familiar and important properties of polynomials. Just as Reed-Solomon codes are constructed using the fact that a degree d polynomial can have at most d roots, our codes will use an analogous theorem that a degree d skew polynomial can have at most d roots *when counted appropriately* (see Theorem 2.21). Unlike the roots of the usual degree d polynomials which do not have any structure, the roots of degree d skew polynomials have an interesting linear-algebraic structure which we exploit in our constructions. The construction from [MK19] is also implicitly based on skew polynomials. In this paper, we make this connection explicit in the hope that the theory of skew polynomials will lead to further developments in the constructions of MR LRCs and coding theory more broadly. Readers familiar with the theory of skew polynomials or who directly want to get to the construction can skip most of the preliminaries except for Section 2.4.

Related Work. Shortly before we published our results, we learned that [CMST20] have independently obtained a result analogous to Theorem 1.3 with a very similar construction. They construct (n, r, h, a, q) -MR LRCs with a field size of

$$q = (O(\max\{r, n/r\}))^h. \quad (5)$$

Compared to this, we have a $\min\{h, r - a\}$ in the exponent in our field size bound (4).

Soon after [CMST20], two more constructions of MR LRCs were published by [Mar20] with the following field sizes:

$$q \leq \left(\max\left\{ (2r)^{r-a}, \frac{g}{r} \right\} \right)^{\min\{h, \lfloor g/r \rfloor\}}, \quad (6)$$

$$q \leq (2r)^{r-a} \left(\left\lfloor \frac{g}{r} \right\rfloor + 1 \right)^{h-1}. \quad (7)$$

The constructions in (6) and (7) are incomparable to our construction in (4). For example when $r = O(1)$, the construction (7) achieves $O(n)^{h-1}$ field size, whereas our construction achieves $O(n)^{\min\{h, r-a\}}$ field size. In the regime when $r = g = \Theta(\sqrt{n})$ and $h \leq a + 2$, our construction achieves the optimal field size of $\Theta_{a,h}(n^{h/2})$, whereas the constructions from [Mar20] require fields of size $n^{\Theta(\sqrt{n})}$.

2 Preliminaries

2.1 Skew polynomial ring

Skew polynomials generalize polynomials while inheriting many of the nice properties of polynomials. Skew polynomials can be defined over division rings* and most of the results about skew polynomials are true in this more general setting. It is known that every finite division ring is a field. Since we will only work with skew polynomial rings defined over fields, we will only define them over fields for simplicity. Most of the theory of skew polynomials presented here is from [LL88, Lam85].

Let \mathbb{K} be a field and let $\sigma : \mathbb{K} \rightarrow \mathbb{K}$ be an endomorphism. For example if $\mathbb{K} = \mathbb{F}_{q^m}$, then $\sigma(x) = x^q$ is an endomorphism. We will now define the concept of a derivation.

Definition 2.1 (Derivation). *A map $\delta : \mathbb{K} \rightarrow \mathbb{K}$ is called a σ -derivation if:*

*Rings where every non-zero element has a multiplicative inverse, but multiplication may not be commutative.

1. δ is a linear map i.e. $\delta(a + b) = \delta(a) + \delta(b)$ for all $a, b \in K$ and
2. $\delta(ab) = \sigma(a)\delta(b) + \delta(a)b$ for all $a, b \in K$.

Example 2.2. 1. The simplest derivation is the zero map i.e. $\delta(a) = 0$ for all $a \in \mathbb{K}$. Skew polynomials are interesting even in this case, and in fact the constructions in this paper only use skew polynomials with $\delta \equiv 0$. So the reader can imagine that the derivation is the zero map on a first reading. We include the general case in the hope that skew polynomial rings with non-zero derivations will find applications in future.

2. Let $\mathbb{K} = \mathbb{F}(x)$ and σ be the identity map. Then $\delta(f(x))$ defined as the Hasse derivate is a σ -derivation. Note that the Hasse derivative of a polynomial function $f(x)$ is the coefficient of y in $f(x + y)$. This can be extended to rational functions in a consistent way using power series.

We will now define the skew polynomial ring.

Definition 2.3 (Skew polynomial ring). Let σ be an endomorphism of \mathbb{K} and δ be a σ -derivation. The skew polynomial ring in variable t , denoted by $\mathbb{K}[t; \sigma, \delta]$, is a non-commutative ring of skew polynomials in t of the form $\{\sum_{i=0}^d a_i t^i : d \geq 0, a_i \in \mathbb{K}\}$ (where we always write the coefficients to the left). Degree of a polynomial $f(t) = \sum_i a_i t^i$, denoted by $\deg(f)$, is the largest d such that $a_d \neq 0$.[†] Addition in $\mathbb{K}[t; \sigma, \delta]$ is component wise. But multiplication is distributive and done according to the following rule:

$$\text{For } a \in \mathbb{K}, t \cdot a = \sigma(a)t + \delta(a). \quad (8)$$

To multiply $f(t)g(t)$, we can first use distributivity to get $f(t)g(t) = \sum_{ij} f_i t^i \cdot g_j t^j$ where $f_i, g_j \in \mathbb{K}$ are coefficients of f, g respectively. Then we use the rule 8 for i times to move the coefficient g_j to the left of t^i . This multiplication turns out to be associative, but may not be commutative. Also $\deg(f \cdot g) = \deg(f) + \deg(g)$. Therefore the skew polynomial ring has no zero divisors.

Example 2.4. The simplest example is when σ is the identity map and δ is the zero map. In this case, skew polynomials coincide with the usual notion of polynomials.

When δ is the zero map, the skew ring is denoted by $\mathbb{K}[t; \sigma]$ and is said to be of endomorphism type. When σ is the identity map, the skew ring is denoted by $\mathbb{K}[t; \delta]$ and is said to be of derivation type.

We will now collect some simple facts about skew polynomials rings. Let $\mathbb{K}[t; \sigma, \delta]$ be a skew polynomial ring.

Lemma 2.5 ([LL88]). $t^n a = \sum_{i=0}^n f_i^n(a) t^i$ where $f_0^n = \delta^n$, $f_1^n = \delta^{n-1}\sigma + \delta^{n-2}\sigma\delta + \dots + \sigma\delta^{n-1}$, \dots , $f_n^n = \sigma^n$ are linear maps.

It turns out that the skew polynomial ring has Euclidean algorithm for right division.

Lemma 2.6 (Euclidean algorithm for right division [LL88]). For every two polynomial $f, g \in \mathbb{K}[t; \sigma, \delta]$, there exist unique polynomials $q(t), r(t)$ such that $f = q \cdot g + r$ where $\deg(r) < \deg(g)$ or $r = 0$.

[†]We will define the degree of the zero polynomial to be ∞ .

This brings us to the most important definition about skew polynomial rings. In the usual polynomial world, we can define the evaluation of a polynomial $f(t) = \sum_i f_i t^i$ at $t = a$ as $\sum_i f_i a^i$. With this definition, it is true that $f(t) = q(t)(t - a) + f(a)$. But for skew polynomials, these two notions of evaluation differ with each other. And the right definition is the second one.

Definition 2.7 (Evaluation). *The evaluation of a polynomial $f \in \mathbb{K}[t; \sigma, \delta]$ at a point $a \in \mathbb{K}$, denoted by $f(a)$, is defined as the remainder obtained when we divide f by $t - a$ on the right i.e. $f(t) = q(t)(t - a) + f(a)$.*

Note that evaluation is a linear map i.e. $(f + g)(a) = f(a) + g(a)$. But it is not always true that $(fg)(a) = f(a)g(a)$. We will see shortly how to compute $(fg)(a)$. The evaluation map can be expressed using “power functions”, which are the evaluations of monomials of the form t^i .

Definition 2.8 (Power functions). *The power functions are defined inductively as follows. For every $a \in \mathbb{K}$*

1. $N_0(a) = 1$ and
2. $N_{i+1}(a) = \sigma(N_i(a))a + \delta(N_{i-1}(a))$.

When $\delta \equiv 0$, we have $N_i(a) = \sigma^{i-1}(a)\sigma^{i-2}(a) \cdots \sigma(a)a$. Additionally if $\sigma \equiv \text{Id}$, then $N_i(a) = a^i$ which explains the terms “power functions”.

Lemma 2.9. *Let $f = \sum_i f_i t^i$. Then $f(a) = \sum_i f_i N_i(a)$.*

Proof. It is easy to prove by induction that evaluation of t^i at a is $N_i(a)$. The general claim follows by linearity of evaluation. \square

We now come to the problem of evaluating $(fg)(a)$. For this, it is useful to define the notion of *conjugates*, which play a big role in this theory.

2.2 Conjugation and Product Rule

Definition 2.10 (Conjugation). *Let $a \in \mathbb{K}$ and $c \in \mathbb{K}^*$. We define the c -conjugate of a , denoted by ${}^c a$, as*

$${}^c a = \sigma(c)ac^{-1} + \delta(c)c^{-1}.$$

We say that b is a conjugate of a if there exists some $c \in \mathbb{K}^$ such that $b = {}^c a$.*

Lemma 2.11. 1. ${}^d({}^c a) = {}^{dc} a$

2. *Conjugacy is an equivalence relation, i.e., we can partition \mathbb{K} into conjugacy classes where elements in each part are conjugates of each other, but elements in different parts are not conjugates.*

Proof. (1) follows easily from the definition of conjugation and the using the fact that $\delta(cd) = \sigma(c)\delta(d) + \delta(c)d$.

$$\begin{aligned} {}^d({}^c a) &= \sigma(d) \cdot {}^c a \cdot d^{-1} + \delta(d)d^{-1} \\ &= \sigma(d)(\sigma(c)ac^{-1} + \delta(c)c^{-1})d^{-1} + \delta(d)d^{-1} \\ &= \sigma(dc)ac^{-1}d^{-1} + \sigma(d)\delta(c)c^{-1}d^{-1} + \delta(d)d^{-1} \\ &= \sigma(dc)a(dc)^{-1} + (\sigma(d)\delta(c) + \delta(d)c)c^{-1}d^{-1} \\ &= \sigma(dc)a(dc)^{-1} + \delta(dc)(dc)^{-1} \\ &= {}^{dc} a. \end{aligned}$$

We now prove (2). Suppose a is a conjugate of b , i.e., $a = {}^x b$ for some $x \in \mathbb{K}^*$. Then $x^{-1}a = x^{-1}({}^x b) = x^{-1}x b = b$. Therefore b is a conjugate of a . Suppose a is a conjugate of b , with $a = {}^x b$, and c is a conjugate of b , with $b = {}^y c$. Then $a = {}^x b = {}^x({}^y c) = {}^{xy} c$. So a is a conjugate of c . \square

So \mathbb{K} will get partitioned into conjugacy classes. To understand the structure of each conjugacy class, we need the notion of *centralizer*.

Definition 2.12 (Centralizer). *The centralizer of $a \in \mathbb{K}$ is defined as:*

$$\mathbb{K}_a = \{c \in \mathbb{K}^* : {}^c a = a\} \cup \{0\}.$$

Lemma 2.13. 1. \mathbb{K}_a is a subfield of \mathbb{K} .[‡]

2. If $a, b \in \mathbb{K}$ are conjugates, then $\mathbb{K}_a = \mathbb{K}_b$.[§]

Proof. (1) Let $x, y \in \mathbb{K}_a \setminus \{0\}$ i.e. ${}^x a = {}^y a = a$. Then

$$\begin{aligned} {}^{x+y} a(x+y) &= \sigma(c+d)a + \delta(c+d) \\ &= \sigma(c)a + \sigma(d)a + \delta(c) + \delta(d) \\ &= {}^c a + {}^d a \\ &= ac + ad = a(c+d). \end{aligned}$$

Therefore ${}^{x+y} a = a$. Also ${}^{yx} a = {}^y({}^x a) = a$. And finally $x^{-1}a = x^{-1}({}^x a) = x^{-1}x a = a$.

(2) Suppose $b = {}^d a$ and let $c \in \mathbb{K}_a$. Then ${}^c b = {}^c({}^d a) = {}^{cd} a = {}^{dc} a = {}^d({}^c a) = {}^d a = b$. Therefore $\mathbb{K}_a \subset \mathbb{K}_b$. By symmetry, $\mathbb{K}_b \subset \mathbb{K}_a$. \square

Because of the above lemma, we can associate a centralizer subfield to each conjugacy class.

Example 2.14. Let $\mathbb{K} = \mathbb{F}_{q^m}$, $\sigma(a) = a^q$ and $\delta \equiv 0$. Then ${}^c a = c^{q-1}a$. Suppose γ is a generator for $\mathbb{F}_{q^m}^*$. There are q equivalence classes, $E_{-1}, E_0, E_1, \dots, E_{q-2}$, where $E_\ell = \{\gamma^i : i \equiv \ell \pmod{(q-1)}\}$ and $E_{-1} = \{0\}$. The centralizer of an element $a \in \mathbb{K}^*$ is

$$\mathbb{K}_a = \{c : c^{q-1}a = a\} \cup \{0\} = \{c : c^{q-1} = 1\} \cup \{0\} = \mathbb{F}_q.$$

Therefore the centralizer of every non-zero element is \mathbb{F}_q and the centralizer of 0 is $\mathbb{K}_0 = \mathbb{K}$.

We will now show how to evaluate $(fg)(a)$. And conjugates play a key role.

Lemma 2.15 (Product evaluation rule). *If $g(a) = 0$, then $(fg)(a) = 0$. If $g(a) \neq 0$ then*

$$(fg)(a) = f\left({}^{g(a)} a\right) g(a).$$

Using the product rule, we can show the following interpolation theorem for polynomials.

Lemma 2.16 (Interpolation [LL88]). *Let $A \subset \mathbb{K}$ be of size n . Then there exists a non-zero degree $\leq n$ polynomial $f \in \mathbb{K}[t; \sigma, \delta]$ which vanishes on A .*

Note that this shows some similarity between polynomials and skew polynomials. We will later need the following lemma.

[‡]When \mathbb{K} is a division ring, \mathbb{K}_a will be a sub-division ring of \mathbb{K} .

[§]When \mathbb{K} is a division ring and not a field, we have $\mathbb{K}_{(x a)} = x \mathbb{K}_a x^{-1}$.

Lemma 2.17. *Let f be any skew polynomial. Then $D_{f,a}(y) = f(ya)y$ is an \mathbb{K}_a -linear map from $\mathbb{K} \rightarrow \mathbb{K}$.*

Proof. Linearity follows from the fact that $N_n(ya)y = \sum_{i=0}^n f_i^n(y)N_i(a)$ and the fact that f_i^n are linear functions.

Alternatively, it follows since $f(ya)y$ is equal to the evaluation of the polynomial $f(t)y$ at a by Lemma 2.15. And clearly the evaluation is linear in y .

\mathbb{K}_a -linearity follows since $\forall c \in \mathbb{K}_a$, $D_{f,a}(yc) = f(y^c a)yc = f(y^c a)yc = f(ya)yc = D_{f,a}(y)c$. \square

2.3 Roots of skew polynomials

The most important and useful fact about usual polynomials is that a degree d non-zero polynomial can have at most d roots. It turns out that this statement is false for skew polynomials! **A skew polynomial can have many more roots than its degree.** But when counted in the right way, we can recover an analogous statement for skew polynomials. In this section, we will prove the “fundamental theorem” about roots of skew polynomials which shows that a degree d skew polynomial cannot have more than d roots when counted the right way. We will begin with showing that any non-zero degree d skew polynomial can have at most d roots in distinct conjugacy classes.

Lemma 2.18. *Let $f \in \mathbb{K}[t; \sigma, \delta]$ be a degree d non-zero polynomial. Then f can have **at most d roots in distinct conjugacy classes.***

Proof. We will prove it using induction on the degree. For the base case, it is clear that a degree 0 polynomial which is a non-zero constant cannot have any roots. Suppose $a_0, a_1, \dots, a_d \in \mathbb{K}$ be roots of f in distinct conjugacy classes. Since $f(a_0) = 0$, we can write $f(t) = h(t)(t - a_0)$ where $\deg(h) = d-1$. By Lemma 2.15, $f(a_i) = h(a_i^{a_0} a_i)(a_i - a_0)$. Therefore $b_i = a_i^{a_0} a_i$ for $i \in \{1, \dots, d\}$ are d roots of h and they lie in distinct conjugacy classes because a_i lie in distinct conjugacy classes. Thus by induction $h = 0$ and therefore $f = 0$ which is a contradiction. \square

Now let us try to understand, the **roots of a skew polynomial in the same conjugacy class.** The following lemma shows that they **form a vector space over a subfield of \mathbb{K} .**

Lemma 2.19. *Let $f \in \mathbb{K}[t; \sigma, \delta]$ be a non-zero polynomial and fix some $a \in \mathbb{K}$ and let $\mathbb{F} = \mathbb{K}_a$ be the centralizer of a (which is a subfield of \mathbb{K}). Define $V_f(a) = \{y \in \mathbb{K}^* : f(ya) = 0\} \cup \{0\}$. Then $V_f(a)$ is a vector space over \mathbb{F} .*

Proof. For any $\lambda \in \mathbb{F}$ and $y \in V_f(a)$, $f(\lambda ya) = f(y(\lambda a)) = f(ya) = 0$. Therefore $\lambda y \in V_f(a)$. If $y_1, y_2 \in V_f(a)$ where $y_1 + y_2 \neq 0$, then by Lemma 2.17, $f(y_1 + y_2)a = 0$. Therefore $y_1 + y_2 \in V_f(a)$. \square

The next lemma shows that the dimension of $V_f(a)$ can be at most $\deg(f)$.

Lemma 2.20. *Let $f \in \mathbb{K}[t; \sigma, \delta]$ be a degree d non-zero polynomial and fix some $a \in \mathbb{K}$ and let $\mathbb{F} = \mathbb{K}_a$ be the centralizer subfield of a . Define $V_f(a) = \{y \in \mathbb{K}^* : f(ya) = 0\} \cup \{0\}$. Then $V_f(a)$ is a vector space over \mathbb{F} of dimension at most d .*

Proof. We will use induction on the degree. For the base case, it is clear that for a degree 0 polynomial, which is a non-zero constant, $\dim_{\mathbb{F}}(V_f(a)) = 0$. Suppose for contradiction that there exists $y_0, y_1, \dots, y_d \in V_f(a)$ which are linearly independent over \mathbb{F} . WLOG, we can assume that $y_0 = 1$ (by redefining a to be equal to $y_0 a$). Since $f(a) = 0$, we can write $f(t) = h(t)(t - a)$ where $\deg(h) = d - 1$. By Lemma 2.15, $f(y_i a) = h(y_i(y_i a - a))(y_i a - a)$. Since $y_0 = 1$ and y_i is linearly independent from y_0 over \mathbb{F} , $y_i \notin \mathbb{F}$. Therefore $y_i a - a \neq 0$, and so $b_i = y_i(y_i a - a)a$ for $i \in \{1, \dots, d\}$

are d roots of h . If we show that $y_i(y_i a - a)$ for $i \in \{1, \dots, d\}$ are linearly independent over \mathbb{F} , then we are done by induction.

Suppose they are not independent. Then there exists $c_1, \dots, c_d \in \mathbb{F}$ s.t. $\sum_{i=1}^d c_i y_i (y_i a - a) = 0$. Therefore,

$$\begin{aligned} a \sum_{i=1}^d c_i y_i &= \sum_{i=1}^d c_i y_i \cdot y_i a \\ &= \sum_{i=1}^d c_i y_i \cdot c_i y_i a && (c_i \in \mathbb{F} = \mathbb{K}_a) \\ &= \left(\sum_{i=1}^d c_i y_i \right) \left(\sum_{i=1}^d c_i y_i \right) a && (x+y a(x+y) = x a x + y a y \text{ for all } x, y \in \mathbb{K}^*) \end{aligned}$$

Since y_1, \dots, y_d are independent over \mathbb{F} , $\sum_{i=1}^d c_i y_i \neq 0$. Therefore $\left(\sum_{i=1}^d c_i y_i \right) a = a$ i.e. $\sum_{i=1}^d c_i y_i \in \mathbb{K}_a = \mathbb{F}$. But this contradicts the fact that $\{y_0 = 1, y_1, \dots, y_d\}$ are linearly independent over \mathbb{F} . \square

The following theorem is the “fundamental theorem” about roots of skew polynomials. It immediately implies Lemma 2.18 and Lemma 2.20 as corollaries. But we have proved them before, just to convey some intuition.

Theorem 2.21. *Let $f \in \mathbb{K}[t; \sigma, \delta]$ be a degree d non-zero polynomial. Let A be the set of roots of f in \mathbb{K} and let $A = \cup_i A_i$ be a partition of A into conjugacy classes. Fix some representatives $a_i \in A_i$. Let $V_i = \{y : y a_i \in A_i\} \cup \{0\}$ which is a linear subspace over $\mathbb{F}_i = \mathbb{K}_{a_i}$ by Lemma 2.19. Then*

$$\sum_i \dim_{\mathbb{F}_i}(V_i) \leq d.$$

The proof of Theorem 2.21 is given in Appendix A.

2.4 Vandermonde matrix

Definition 2.22 (Vandermonde matrix). *Let $A = \{a_1, \dots, a_n\} \subset \mathbb{K}$. The Vandermonde matrix formed by A , denoted by $V(a_1, \dots, a_n)$, is defined as:*

$$V_d(a_1, \dots, a_n) = \begin{bmatrix} N_0(a_1) & N_0(a_2) & \cdots & N_0(a_n) \\ N_1(a_1) & N_1(a_2) & \cdots & N_1(a_n) \\ \vdots & \vdots & \ddots & \vdots \\ N_{d-1}(a_1) & N_{d-1}(a_2) & \cdots & N_{d-1}(a_n) \end{bmatrix}$$

Lemma 2.23. *Let $a_1, \dots, a_d \in \mathbb{K}$ be in distinct conjugacy classes. Then $V_d(a_1, \dots, a_d)$ is full-rank.*

Proof. Follows from Lemma 2.9 and Lemma 2.18. \square

Corollary 2.24. *Let $\gamma \in \mathbb{F}_{q^m}^*$ be a generator of the multiplicative group. Let $d \leq q - 1$ and $\ell_1, \dots, \ell_d \in \{0, 1, 2, \dots, q - 2\}$ be distinct. Then the following matrix M is full rank.*

$$M = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \gamma^{\ell_1} & \gamma^{\ell_2} & \cdots & \gamma^{\ell_d} \\ \gamma^{\ell_1(1+q)} & \gamma^{\ell_2(1+q)} & \cdots & \gamma^{\ell_d(1+q)} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma^{\ell_1(1+q+\cdots+q^{d-2})} & \gamma^{\ell_2(1+q+\cdots+q^{d-2})} & \cdots & \gamma^{\ell_d(1+q+\cdots+q^{d-2})} \end{bmatrix}$$

Proof. Let $\mathbb{K} = \mathbb{F}_{q^m}$, $\sigma(a) = a^q$ and $\delta \equiv 0$. Then $N_i(a) = a^{1+q+q^2+\dots+q^{i-1}}$. By Lemma 2.23, it is enough to show that ℓ_1, \dots, ℓ_d fall in distinct conjugacy classes. This is shown in Example 2.14. \square

Note that when $m = 1$, the matrix in the above corollary reduces to the usual Vandermonde matrix one is familiar with.

In general we would want to compute the rank of $V_n(a_1, \dots, a_n)$ for any given a_1, \dots, a_n . The following lemma generalizes Lemma 2.23.

Lemma 2.25. *Let $A = \{a_1, \dots, a_n\} \subset \mathbb{K}$. Let $A = A_1 \cup A_2 \cup \dots \cup A_r$ be the partition of A into conjugacy classes. Then $\text{rank}(V_n(A)) = \sum_i \text{rank}(V_n(A_i))$.*

Proof. Follows from Theorem 2.21. \square

By the above lemmas, we reduced the problem to computing $\text{rank}(V_n(A))$ when all elements of A belong to the same conjugacy class. The following lemma shows how to compute this.

Lemma 2.26. *Let $a \in \mathbb{K}$ and $\mathbb{F} = \mathbb{K}_a$ which is a subfield of \mathbb{K} . Then for any $\{c_1, \dots, c_n\} \subset \mathbb{K}^*$, we have*

$$\text{rank}(V_n(c_1 a, \dots, c_n a)) = \dim_{\mathbb{F}} \text{span}_{\mathbb{F}}\{c_1, \dots, c_n\}.$$

In particular, $V_n(c_1 a, \dots, c_n a)$ is full-rank iff $\{c_1, \dots, c_n\}$ are linearly independent over \mathbb{F} .

Proof. Follows from Lemma 2.20. \square

Corollary 2.27. *Let $\gamma \in \mathbb{F}_{q^m}^*$ be a generator of the multiplicative group and let $\ell \in \{0, 1, \dots, q-2\}$. Let $\beta_1, \dots, \beta_m \in \mathbb{F}_{q^m}$ be linearly independent over \mathbb{F}_q . Then the following matrix M is full rank.*

$$M = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \gamma^\ell \beta_1^{q-1} & \gamma^\ell \beta_2^{q-1} & \dots & \gamma^\ell \beta_m^{q-1} \\ \gamma^{\ell(1+q)} \beta_1^{q^2-1} & \gamma^{\ell(1+q)} \beta_2^{q^2-1} & \dots & \gamma^{\ell(1+q)} \beta_m^{q^2-1} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma^{\ell(1+q+\dots+q^{m-2})} \beta_1^{q^{m-1}-1} & \gamma^{\ell(1+q+\dots+q^{m-2})} \beta_2^{q^{m-1}-1} & \dots & \gamma^{\ell(1+q+\dots+q^{m-2})} \beta_m^{q^{m-1}-1} \end{bmatrix}$$

Proof. Let $\mathbb{K} = \mathbb{F}_{q^m}$, $\sigma(a) = a^q$ and $\delta \equiv 0$. Then $N_i(a) = a^{1+q+q^2+\dots+q^{i-1}}$. Let $a = \gamma^\ell$ then $M = V_m(\beta_1 a, \dots, \beta_m a)$. Therefore M is full rank by Lemma 2.26. \square

3 Skew polynomials based MR LRC constructions

Let us recall that an (n, r, h, a, q) -LRC admits a parity check matrix H of the following form

$$H = \left[\begin{array}{c|c|c|c} A_1 & 0 & \dots & 0 \\ \hline 0 & A_2 & \dots & 0 \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline 0 & 0 & \dots & A_g \\ \hline B_1 & B_2 & \dots & B_g \end{array} \right]. \quad (9)$$

Here A_1, A_2, \dots, A_g are $a \times r$ matrices over \mathbb{F}_q which represent the local parity checks, B_1, B_2, \dots, B_g are $h \times r$ matrices over \mathbb{F}_q which together represent the h global parity checks. The rest of the matrix is filled with zeros. By Proposition 1.2, C is an MR LRC iff (1) any ' a ' columns of each matrix A_i are linearly independent and (2) any submatrix of H formed by selecting a columns in each local group and any h additional columns is full rank.

3.1 Construction: Proof of Theorem 1.3

In this section, we will prove Theorem 1.3 by presenting a construction of MR LRCs over fields of size $q = O(\max(g, r))^{\min\{h, r-a\}}$. The construction presented here is inspired from [MK19], where they achieve a field size of $O(\max(g, r))^{r-a}$.[¶]

Let $q_0 \geq \max\{g+1, r\}$ be a prime power. Choose $\alpha_1, \alpha_2, \dots, \alpha_r \in \mathbb{F}_{q_0}$ to be distinct. Define

$$A_\ell = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \alpha_1 & \alpha_2 & \dots & \alpha_r \\ \alpha_1^2 & \alpha_2^2 & \dots & \alpha_r^2 \\ \vdots & \vdots & & \vdots \\ \alpha_1^{a-1} & \alpha_2^{a-1} & \dots & \alpha_r^{a-1} \end{bmatrix}.$$

Note that $A_1 = A_2 = \dots = A_g$. Let $m = \min\{r-a, h\}$ and let γ be a generator for $\mathbb{F}_{q_0}^*$. Our codes will be defined over the field $\mathbb{F}_q = \mathbb{F}_{q_0^m}$. Define $\beta_1, \beta_2, \dots, \beta_r \in \mathbb{F}_{q_0^m}$ as

$$\beta_i = \begin{bmatrix} \alpha_i^a \\ \alpha_i^{a+1} \\ \vdots \\ \alpha_i^{a+m-1} \end{bmatrix},$$

where we are expressing β_i in some basis for $\mathbb{F}_{q_0^m}$ (which is a \mathbb{F}_{q_0} -vector space of dimension m). Define

$$B_\ell = \begin{bmatrix} \beta_1 & \beta_2 & \dots & \beta_r \\ \gamma^\ell \beta_1^{q_0} & \gamma^\ell \beta_2^{q_0} & \dots & \gamma^\ell \beta_r^{q_0} \\ \gamma^{\ell(1+q_0)} \beta_1^{q_0^2} & \gamma^{\ell(1+q_0)} \beta_2^{q_0^2} & \dots & \gamma^{\ell(1+q_0)} \beta_r^{q_0^2} \\ \vdots & \vdots & & \vdots \\ \gamma^{\ell(1+q_0+\dots+q_0^{h-2})} \beta_1^{q_0^{h-1}} & \gamma^{\ell(1+q_0+\dots+q_0^{h-2})} \beta_2^{q_0^{h-1}} & \dots & \gamma^{\ell(1+q_0+\dots+q_0^{h-2})} \beta_r^{q_0^{h-1}} \end{bmatrix}.$$

To prove that the above construction is an MR LRC, we will use **properties of the skew field $\mathbb{F}_{q_0^m}[x; \sigma]$** where $\sigma(a) = a^{q_0}$. We know that $\mathbb{F}_{q_0^m}$ will get partitioned into $q_0 - 1$ conjugacy classes as shown in Example 2.14. If $\gamma \in \mathbb{F}_{q_0^m}^*$ is a generator of $\mathbb{F}_{q_0^m}^*$, then $\{1, \gamma, \gamma^2, \dots, \gamma^{q_0-2}\}$ fall in distinct conjugacy classes. Intuitively, in the construction each local group corresponds to one conjugacy class. This is possible since we chose $q_0 \geq g+1$. The stabilizer subfield of each conjugacy class is \mathbb{F}_{q_0} as shown in Example 2.14. Therefore **we choose the matrices B_i for local group i as a (skew) Vandermonde matrix** where the evaluation points β_1, \dots, β_r are from the conjugacy class of γ^i , but are linearly independent over the stabilizer subfield \mathbb{F}_{q_0} .

Claim 3.1. *The above construction is an MR LRC over fields of size $q = q_0^{\min\{h, r-a\}}$.*

Proof. Given an erasure pattern E of size $|E| = ag + h$, composed of a erasures in each local group and h additional erasures, we want to argue that the submatrix $H(E)$, H restricted columns in E , is full rank. WLOG, assume that the h additional erasures happen in local groups $1, 2, \dots, t \in [g]$ for $t \leq h$. Let E_i be the set of erasures that happen in the i^{th} local group. Let $S_i \subset E_i$ be an arbitrary subset of size $|S_i| = a$ and let $T_i = E_i \setminus S_i$. Note that $|T_i| \leq m$ for all i . For a matrix M

[¶]The improvement comes from choosing β_1, \dots, β_r carefully in our construction. Moreover [MK19] constructs a generator matrix for the code, whereas we construct a parity check matrix.

and a subset X of its columns, we will use $M(X)$ to denote the submatrix of M formed by columns in X . We need to show that $H(E)$ (which is an $(ag + h) \times (ag + h)$ matrix) is full rank where

$$H(E) = \left[\begin{array}{c|c|c|c} A_1(S_1 \cup T_1) & 0 & \cdots & 0 \\ \hline 0 & A_2(S_2 \cup T_2) & \cdots & 0 \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline 0 & 0 & \cdots & A_g(S_g \cup T_g) \\ \hline B_1(S_1 \cup T_1) & B_2(S_2 \cup T_2) & \cdots & B_g(S_g \cup T_g) \end{array} \right].$$

Note that $A_1(S_1), A_2(S_2), \dots, A_g(S_g)$ are $a \times a$ matrices of full rank. By doing column operations on $H(E)$, in each local group we can use the columns of $A_i(S_i)$ to remove the columns of $A_i(T_i)$. This results in the lower block $B_i(T_i)$ to change into a Schur complement as follows:

$$\left[\begin{array}{c|c} A_i(S_i) & A_i(T_i) \\ \hline B_i(S_i) & B_i(T_i) \end{array} \right] \rightarrow \left[\begin{array}{c|c} A_i(S_i) & 0 \\ \hline B_i(S_i) & B_i(T_i) - B_i(S_i)A_i(S_i)^{-1}A_i(T_i) \end{array} \right].$$

Note that $T_i = \phi$ for $i > t$. So by doing row and column operations on $H(E)$, we can set it in a block diagonal form, where the diagonal blocks are given by $A_1(S_1), A_2(S_2), \dots, A_g(S_g)$ and one additional $h \times h$ block given by

$$C = \left[\begin{array}{c|c|c} B_1(T_1) - B_1(S_1)A_1(S_1)^{-1}A_1(T_1) & \cdots & B_t(T_t) - B_t(S_t)A_t(S_t)^{-1}A_t(T_t) \end{array} \right].$$

Note that all the entries in $A(S_i)^{-1}A_i(T_i)$ are in the base field \mathbb{F}_{q_0} . Also column operations on B_i with \mathbb{F}_{q_0} coefficients retain its structure with β 's replaced by their corresponding \mathbb{F}_{q_0} -linear combinations. Therefore by Lemma 2.25 and Lemma 2.26, it is enough to show that the following t matrices D_1, D_2, \dots, D_t are full rank:

$$D_i = \left[\beta(T_i) - \beta(S_i)A_i(S_i)^{-1}A_i(T_i) \right]$$

where $\beta = [\beta_1, \dots, \beta_r]$ is a $m \times r$ matrix over \mathbb{F}_{q_0} . Note that $[D_1 | D_2 | \dots | D_t]$ is just the first row of C (with entries in $\mathbb{F}_{q_0^m}$) expressed as a matrix over \mathbb{F}_{q_0} . Consider following matrices given by

$$F_i = \left[\begin{array}{c|c} A_i(S_i) & A_i(T_i) \\ \hline \beta(S_i) & \beta(T_i) \end{array} \right]$$

where each F_i is of size $(a + m) \times (a + |T_i|)$. Each F_i is a Vandermonde matrix by construction. Since $|T_i| \leq m$, each F_i is full rank. Now if we do column operations to get F_i into block diagonal form we get:

$$\left[\begin{array}{c|c} A_i(S_i) & 0 \\ \hline \beta(S_i) & \beta(T_i) - \beta(S_i)A_i(S_i)^{-1}A_i(T_i) \end{array} \right] = \left[\begin{array}{c|c} A_i(S_i) & 0 \\ \hline \beta(S_i) & D_i \end{array} \right].$$

This implies that D_1, D_2, \dots, D_t are full rank over \mathbb{F}_{q_0} which completes the proof. \square

A slightly better construction which only requires $q_0 \geq \max\{g + 1, r - 1\}$ can be obtained by choosing

$$A_\ell = \begin{bmatrix} 1 & \alpha_2^{m+a-1} & \alpha_3^{m+a-1} & \cdots & \alpha_r^{m+a-1} \\ 0 & \alpha_2^{m+a-2} & \alpha_3^{m+a-2} & \cdots & \alpha_r^{m+a-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & \alpha_2^{m+1} & \alpha_3^{m+1} & \cdots & \alpha_r^{m+1} \\ 0 & \alpha_2^m & \alpha_3^m & \cdots & \alpha_r^m \end{bmatrix}$$

and $\beta_1, \beta_2, \dots, \beta_r \in \mathbb{F}_{q_0}^m$ as:

$$\beta_1 = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \text{ and } \beta_i = \begin{bmatrix} \alpha_i^{m-1} \\ \vdots \\ \alpha_i \\ 1 \end{bmatrix} \text{ for } i \in \{2, 3, \dots, r\}.$$

3.2 Construction: Proof of Theorem 1.5

When $a = 1$ and g is a fixed constant, we can improve the construction from the previous section using ideas from BCH codes. Let $q_0 \geq g + 1$ be a prime power. Define

$$A_\ell = \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}.$$

Note that $A_1 = A_2 = \dots = A_g$. We will construct $\beta_1, \beta_2, \dots, \beta_r$ similarly as in the previous construction, but since we will only need \mathbb{F}_{q_0} linear independence of β 's, we can improve the construction by using BCH codes. Let $q_1 = q_0^s$ where $s = \lceil \log_{q_0}(r) \rceil$, note that $r \leq q_1 \leq q_0 r = O(gr) = O(n)$. Let $\alpha_1, \alpha_2, \dots, \alpha_r \in \mathbb{F}_{q_1}$ be distinct. Let $m = \min\{r-1, h\}$ and let $m' = m - \lceil m/q_0 \rceil$ and define $\beta_1, \beta_2, \dots, \beta_r \in \mathbb{F}_{q_1^{m'}}$ as

$$\beta_i = \begin{bmatrix} \alpha_i \\ \alpha_i^2 \\ \vdots \\ \alpha_i^{q_0-1} \\ \alpha_i^{q_0+1} \\ \vdots \\ \alpha_i^m \end{bmatrix},$$

where we are expressing β_i in some basis for $\mathbb{F}_{q_1^{m'}}$ (which is a \mathbb{F}_{q_1} -vector space of dimension m'). Note that we are skipping powers of α_i which are divisible by q_0 . Therefore the dimension of β_i with entries in \mathbb{F}_{q_1} is $m' = m - \lceil m/q_0 \rceil$. Let γ be a generator of $\mathbb{F}_{q_1^{m'}}^* = \mathbb{F}_{q_0^{sm'}}^*$. Define

$$B_\ell = \begin{bmatrix} \beta_1 & \beta_2 & \dots & \beta_r \\ \gamma^\ell \beta_1^{q_0} & \gamma^\ell \beta_2^{q_0} & \dots & \gamma^\ell \beta_r^{q_0} \\ \gamma^{\ell(1+q_0)} \beta_1^{q_0^2} & \gamma^{\ell(1+q_0)} \beta_2^{q_0^2} & \dots & \gamma^{\ell(1+q_0)} \beta_r^{q_0^2} \\ \vdots & \vdots & & \vdots \\ \gamma^{\ell(1+q_0+\dots+q_0^{h-2})} \beta_1^{q_0^{h-1}} & \gamma^{\ell(1+q_0+\dots+q_0^{h-2})} \beta_2^{q_0^{h-1}} & \dots & \gamma^{\ell(1+q_0+\dots+q_0^{h-2})} \beta_r^{q_0^{h-1}} \end{bmatrix}.$$

Claim 3.2. *The above construction is an MR LRC over fields of size*

$$q = q_1^{m'} \leq (O(n))^{m - \lceil m/q_0 \rceil}$$

where $q_0 \geq g + 1$ is any prime power and $m = \min\{r-1, h\}$.

Proof. The proof is analogous to the proof of Theorem 1.3. We only need \mathbb{F}_{q_0} -linear independence of any $m+1$ columns of

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ \beta_1 & \beta_2 & \dots & \beta_r \end{bmatrix}.$$

This follows from the properties of the BCH code construction. Since we only care about \mathbb{F}_{q_0} -linear independence, it is enough to show linear independence of any $m + 1$ columns of the $(m + 1) \times r$ matrix over \mathbb{F}_{q_1} given by

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_r \\ \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_r^2 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^m & \alpha_2^m & \cdots & \alpha_r^m \end{bmatrix}$$

where we added back all the rows where the powers are multiples of q_0 . This follows trivially, since this is a Vandermonde matrix. \square

Acknowledgment

We thank Sergey Yekhanin for several illuminating discussions about LRCs and MR codes.

References

- [BHH13] Mario Blaum, James Lee Hafner, and Steven Hetzler. Partial-MDS codes and their application to RAID type of architectures. *IEEE Transactions on Information Theory*, 59(7):4510–4519, 2013.
- [Bla13] Mario Blaum. Construction of PMDS and SD codes extending RAID 5. Arxiv 1305.0032, 2013.
- [BPSY16] Mario Blaum, James Plank, Moshe Schwartz, and Eitan Yaakobi. Construction of partial MDS and sector-disk codes with two global parity symbols. *IEEE Transactions on Information Theory*, 62(5):2673–2681, 2016.
- [CHL07] Minghua Chen, Cheng Huang, and Jin Li. On maximally recoverable property for multi-protection group codes. In *IEEE International Symposium on Information Theory (ISIT)*, pages 486–490, 2007.
- [CK17] Gokhan Calis and Ozan Koyluoglu. A general construction fo PMDS codes. *IEEE Communications Letters*, 21(3):452–455, 2017.
- [CMST20] Han Cai, Ying Miao, Moshe Schwartz, and Xiaohu Tang. A construction of maximally recoverable codes with order-optimal field size. *arXiv preprint arXiv:2011.13606*, 2020.
- [GGY20] Sivakanth Gopi, Venkatesan Guruswami, and Sergey Yekhanin. Maximally recoverable lrcs: A field size lower bound and constructions for few heavy parities. *IEEE Trans. Inf. Theory*, 66(10):6066–6083, 2020.
- [GHJY14] Parikshit Gopalan, Cheng Huang, Bob Jenkins, and Sergey Yekhanin. Explicit maximally recoverable codes with locality. *IEEE Transactions on Information Theory*, 60(9):5245–5256, 2014.
- [GHK⁺17] Parikshit Gopalan, Guangda Hu, Swastik Kopparty, Shubhangi Saraf, Carol Wang, and Sergey Yekhanin. Maximally recoverable codes for grid-like topologies. In *28th Annual Symposium on Discrete Algorithms (SODA)*, pages 2092–2108, 2017.

- [GHSY12] Parikshit Gopalan, Cheng Huang, Huseyin Simitci, and Sergey Yekhanin. On the locality of codeword symbols. *IEEE Transactions on Information Theory*, 58(11):6925–6934, 2012.
- [GJX20] Venkatesan Guruswami, Lingfei Jin, and Chaoping Xing. Constructions of maximally recoverable local reconstruction codes via function fields. *IEEE Trans. Inf. Theory*, 66(10):6133–6143, 2020.
- [GYBS17] Ryan Gabrys, Eitan Yaakobi, Mario Blaum, and Paul Siegel. Construction of partial MDS codes over small finite fields. In *2017 IEEE International Symposium on Information Theory (ISIT)*, pages 1–5, 2017.
- [HCL07] Cheng Huang, Minghua Chen, and Jin Li. Pyramid codes: flexible schemes to trade space for access efficiency in reliable data storage systems. In *6th IEEE International Symposium on Network Computing and Applications (NCA 2007)*, pages 79–86, 2007.
- [HSX⁺12] Cheng Huang, Huseyin Simitci, Yikang Xu, Aaron Ogus, Brad Calder, Parikshit Gopalan, Jin Li, and Sergey Yekhanin. Erasure coding in Windows Azure Storage. In *USENIX Annual Technical Conference (ATC)*, pages 15–26, 2012.
- [HY16] Guangda Hu and Sergey Yekhanin. New constructions of SD and MR codes over small finite fields. In *2016 IEEE International Symposium on Information Theory (ISIT)*, pages 1591–1595, 2016.
- [Lam85] Tsit-Yuen Lam. *A general theory of Vandermonde matrices*. Center for Pure and Applied Mathematics, University of California, Berkeley, 1985.
- [LL88] Tsit-Yuen Lam and André Leroy. Vandermonde and wronskian matrices over division rings. *Journal of Algebra*, 119(2):308–336, 1988.
- [Mar20] Umberto Martínez-Peñas. A general family of MSRD codes and PMDS codes with smaller field sizes from extended Moore matrices. *CoRR*, abs/2011.14109, 2020.
- [MK19] Umberto Martínez-Peñas and Frank R. Kschischang. Universal and dynamic locally repairable codes with maximal recoverability via sum-rank codes. *IEEE Trans. Inf. Theory*, 65(12):7790–7805, 2019.
- [PD14] Dimitris Papailiopoulos and Alexandros Dimakis. Locally repairable codes. *IEEE Transactions on Information Theory*, 60(10):5843–5855, 2014.
- [SAP⁺13] Maheswaran Sathiamoorthy, Megasthenis Asteris, Dimitris S. Papailiopoulos, Alexandros G. Dimakis, Ramkumar Vadali, Scott Chen, and Dhruva Borthakur. XORing elephants: novel erasure codes for big data. In *Proceedings of VLDB Endowment (PVLDB)*, pages 325–336, 2013.
- [TPD16] Itzhak Tamo, Dimitris Papailiopoulos, and Alexandros G. Dimakis. Optimal locally repairable codes and connections to matroid theory. *IEEE Transactions on Information Theory*, 62:6661–6671, 2016.

A Proof of Theorem 2.21

We restate Theorem 2.21 for convenience.

Theorem A.1. *Let $f \in \mathbb{K}[t; \sigma, \delta]$ be a degree d non-zero polynomial. Let A be the set of roots of f in \mathbb{K} and let $A = \cup_i A_i$ be a partition of A into conjugacy classes. Fix some representatives $a_i \in A_i$. Let $V_i = \{y : {}^y a_i \in A_i\} \cup \{0\}$ which is a linear subspace over $\mathbb{F}_i = \mathbb{K}_{a_i}$ by Lemma 2.19. Then*

$$\sum_i \dim_{\mathbb{F}_i}(V_i) \leq d.$$

Proof. We will use induction on the degree. For the base case, it is clear that for a degree 0 polynomial, which is a non-zero constant, $\dim_{\mathbb{F}_i}(V_i) = 0$ for every i . We will now show the induction step.

For each i , let $d_i = \dim_{\mathbb{F}_i}(V_i)$. Fix some basis $y(i, 1), y(i, 2), \dots, y(i, d_i) \in \mathbb{K}^*$ which span V_i with coefficients in $\mathbb{F}_i = \mathbb{K}_{a_i}$. WLOG, we can assume that $y(i, 1) = 1$ for every i , by reassigning $a_i = y(i, 1)a_i$.

Fix some conjugacy class i^* s.t. $d_{i^*} \geq 1$. Since $f(a_{i^*}) = 0$, we can write $f(t) = h(t)(t - a_{i^*})$ where $\deg(h) = d - 1$. Now let A'_i be the roots of h in conjugacy class i and $V'_i = \{y : {}^y a_i \in A'_i \cup \{0\}\}$. We claim that $\dim_{\mathbb{F}_i}(V'_i) \geq \dim_{\mathbb{F}_i}(V_i)$ for every $i \neq i^*$ and $\dim_{\mathbb{F}_{i^*}}(V'_{i^*}) \geq \dim_{\mathbb{F}_{i^*}}(V_{i^*}) - 1$. By induction $\sum_i \dim_{\mathbb{F}_i}(V'_i) \leq d - 1$. Therefore we have $\sum_i \dim_{\mathbb{F}_i}(V_i) \leq d$. We will now prove the claim in two parts.

Claim A.2. $\dim_{\mathbb{F}_i}(V'_i) \geq \dim_{\mathbb{F}_i}(V_i)$ for every $i \neq i^*$.

Proof. Fix some conjugacy class $i \neq i^*$. By Lemma 2.15,

$$f\left(y(i, j)a_i\right) = h\left(y(i, j)(y(i, j)a_i - a_{i^*})a_i\right)\left(y(i, j)a_i - a_{i^*}\right).$$

Since a_i, a_{i^*} are in different conjugacy classes, $y(i, j)a_i - a_{i^*} \neq 0$. So $b_j = y(i, j)(y(i, j)a_i - a_{i^*})a_i$ for $j \in \{1, \dots, d_i\}$ are d_i roots of h in the i^{th} conjugacy class A'_i . If we show that $y(i, j)(y(i, j)a_i - a_{i^*})$ for $j \in \{1, \dots, d_i\}$ are linearly independent over \mathbb{F}_i , then this proves the claim.

Suppose they are not independent. Then there exists $c_1, \dots, c_{d_i} \in \mathbb{F}_i$ s.t. $\sum_{j=1}^{d_i} c_j y(i, j)(y(i, j)a_i - a_{i^*}) = 0$. Therefore,

$$\begin{aligned} a_{i^*} \sum_{j=1}^{d_i} c_j y(i, j) &= \sum_{j=1}^{d_i} c_j y(i, j) \cdot y(i, j)a_i \\ &= \sum_{j=1}^{d_i} c_j y(i, j) \cdot c_j y(i, j)a_i && (c_j \in \mathbb{F}_i = \mathbb{K}_{a_i}) \\ &= \left(\sum_{j=1}^{d_i} c_j y(i, j) \right) \left(\sum_{j=1}^{d_i} c_j y(i, j) \right) a_i && (x + {}^y a(x + y) = {}^x a x + {}^y a y \text{ for all } x, y \in \mathbb{K}^*) \end{aligned}$$

Since $y(i, 1), \dots, y(i, d_i)$ are independent over \mathbb{F}_i , $\sum_{j=1}^{d_i} c_j y(i, j) \neq 0$. Therefore $\left(\sum_{j=1}^{d_i} c_j y(i, j) \right) a_i = a_{i^*}$. This is a contradiction because a_i, a_{i^*} are in different conjugate classes. \square

Claim A.3. $\dim_{\mathbb{F}_{i^*}}(V'_{i^*}) \geq \dim_{\mathbb{F}_{i^*}}(V_{i^*}) - 1$.

Proof. The proof is exactly similar to that of the previous claim, up until the last. Let $j \in \{2, 3, \dots, d_{i^*}\}$. By Lemma 2.15,

$$f\left(y(i^*, j)a_{i^*}\right) = h\left(y(i^*, j)\left(y(i^*, j)a_{i^*} - a_{i^*}\right)a_{i^*}\right)\left(y(i^*, j)a_{i^*} - a_{i^*}\right).$$

Since $y(i^*, 1) = 1$ and $y(i^*, j)$ are linearly independent over \mathbb{F}_{i^*} , $y(i^*, j) \notin \mathbb{F}_{i^*}$. Therefore $y(i^*, j)a_{i^*} - a_{i^*} \neq 0$. So $b_j = y(i^*, j)\left(y(i^*, j)a_{i^*} - a_{i^*}\right)a_{i^*}$ for $j \in \{2, \dots, d_{i^*}\}$ are $d_{i^*} - 1$ roots of h in the i^{*th} conjugacy class A'_{i^*} . If we show that $y(i^*, j)\left(y(i^*, j)a_{i^*} - a_{i^*}\right)$ for $j \in \{2, \dots, d_{i^*}\}$ are linearly independent over \mathbb{F}_{i^*} , then this proves the claim.

Suppose they are not independent. Then there exists $c_2, \dots, c_{d_{i^*}} \in \mathbb{F}_{i^*}$ s.t. $\sum_{j=2}^{d_{i^*}} c_j y(i^*, j)\left(y(i^*, j)a_{i^*} - a_{i^*}\right) = 0$. Therefore,

$$\begin{aligned} a_{i^*} \sum_{j=2}^{d_{i^*}} c_j y(i^*, j) &= \sum_{j=2}^{d_{i^*}} c_j y(i^*, j) \cdot y(i^*, j)a_{i^*} \\ &= \sum_{j=2}^{d_{i^*}} c_j y(i^*, j) \cdot c_j y(i^*, j)a_{i^*} \quad (c_j \in \mathbb{F}_{i^*} = C(a_{i^*})) \\ &= \left(\sum_{j=2}^{d_{i^*}} c_j y(i^*, j)\right) \left(\sum_{j=2}^{d_{i^*}} c_j y(i^*, j)\right)a_{i^*} \quad (x+y)a(x+y) = xax + yay \text{ for all } x, y \in \mathbb{K}^*) \end{aligned}$$

Since $y(i^*, 1), \dots, y(i^*, d_{i^*})$ are independent over \mathbb{F}_{i^*} , $\sum_{j=2}^{d_{i^*}} c_j y(i^*, j) \neq 0$. Therefore $\left(\sum_{j=2}^{d_{i^*}} c_j y(i^*, j)\right)a_{i^*} = a_{i^*}$. Therefore $\sum_{j=2}^{d_{i^*}} c_j y(i^*, j) \in C(a_{i^*}) = \mathbb{F}_{i^*}$. But this contradicts the fact that $\{y(i^*, 1) = 1, y(i^*, 2), \dots, y(i^*, d_{i^*})\}$ are linearly independent over \mathbb{F}_{i^*} . \square

The above two claims finish the proof of Theorem 2.21. \square