# Bounds on the Maximal Minimum Distance of Linear Locally Repairable Codes

Antti Pöllänen*, Thomas Westerbäck*, Ragnar Freij-Hollanti†, and Camilla Hollanti*
*Department of Mathematics and Systems Analysis, Aalto University, P.O.Box 11100, FI-00076 Aalto, Finland
Emails: {firstname.lastname}@aalto.fi
†Department of Communications and Networking, Aalto University, P.O.Box 13000, FI-00076 Aalto, Finland
Email: ragnar.freij@aalto.fi



***Abstract*—Locally repairable codes (LRCs) are error correcting codes used in distributed data storage. Besides a global level, they enable errors to be corrected locally, reducing the need for communication between storage nodes. There is a close connection between almost affine LRCs and matroid theory which can be utilized to construct good LRCs and derive bounds on their performance.**

**A generalized Singleton bound for linear LRCs with parameters $(n, k, d, r, \delta)$ was given in [N. Prakash *et al.*, "Optimal Linear Codes with a Local-Error-Correction Property", IEEE Int. Symp. Inf. Theory]. In this paper, a LRC achieving this bound is called *perfect*. Results on the existence and nonexistence of linear perfect $(n, k, d, r, \delta)$-LRCs were given in [W. Song *et al.*, "Optimal locally repairable codes", IEEE J. Sel. Areas Comm.]. Using matroid theory, these existence and nonexistence results were later strengthened in [T. Westerbäck *et al.*, "On the Combinatorics of Locally Repairable Codes", Arxiv: 1501.00153], which also provided a general lower bound on the maximal achievable minimum distance $d_{\max}(n, k, r, \delta)$ that a linear LRC with parameters $(n, k, r, \delta)$ can have. This article expands the class of parameters $(n, k, d, r, \delta)$ for which there exist perfect linear LRCs and improves the lower bound for $d_{\max}(n, k, r, \delta)$. Further, this bound is proved to be optimal for the class of matroids that is used to derive the existence bounds of linear LRCs.**

## I. Introduction

In modern times, the need for large scale data storage is swiftly increasing. This need is present for example in large data centers and in cloud storage. The large scale of these distributed data storage systems makes hardware failures common. However, the data should be preserved regardless of failures, and error correcting codes can be utilized to prevent data loss.

A traditional approach is to look for codes which simultaneously maximize error tolerance and minimize storage space consumption. However, this tends to yield codes for which error correction requires an unrealistic amount of communication between storage nodes. *Locally repairable codes* (LRCs) solve this problem by allowing errors to be corrected locally, in addition to the global level.

This work was partially supported by the Academy of Finland grants #276031, #282938, #283262, and by Magnus Ehrnrooth Foundation, Finland. The support from the European Science Foundation under the COST Action IC1104 is also gratefully acknowledged. The first author would like to thank the Department of Mathematics and Systems Analysis at Aalto University for financial support during the summer of 2015.

Besides the parameters $(n, k, d)$ referring to the length, dimension, and minimum distance of a regular linear code, respectively, a LRC is characterized by two additional parameters, $r$ and $\delta$. Informally speaking, the local error correction is enabled by dividing the code symbols into locality sets whose size is at most $r + \delta - 1$ and inside which any $\delta - 1$ symbols can be recovered using the rest of the symbols in the locality set.

### A. Related Work

The notion of a LRC was first introduced in [1]. The generalized Singleton bound for linear $(n, k, d, r, \delta)$-LRCs states that

$$d \leq n - k + 1 - (\lceil k/r \rceil - 1)(\delta - 1). \tag{1}$$

This bound was given in [2] for $\delta = 2$ and in [3] for a general $\delta$. This bound has then been generalized for both linear and nonlinear codes in several ways, see *e.g.* [4], [5], [6] and [7].

The class of *almost affine* codes is a generalization of the class of linear codes. In [8] it was proved that every almost affine code induces a matroid. Many important properties (but not all) of almost affine codes are *matroid invariants* in the sense that the properties only depend on the matroid structure of the code. Matroid theory was used in [9] in order to prove that the minimum distance of a class of linear LRCs achieves the generalized Singleton bound. It was proved in [10] that every almost affine LRC induces a *matroid* such that the parameters $(n, k, d, r, \delta)$ of the LRC appear as matroid invariants. Consequently, the parameters $(n, k, d, r, \delta)$ were generalized to matroids and the bound (1) was proven to also hold for all matroids, which is nontrivial since not all matroids are induced by almost affine codes. An even more general Singleton bound was given for polymatroids in [11], motivated by the fact that all general LRCs induce a polymatroid.

Results on the existence and non-existence of linear $(n, k, d, r, \delta)$-LRCs achieving the generalized Singleton bound were given in [12]. Codes or matroids achieving the generalized Singleton bound are here called *perfect*. Using the *lattice of cyclic flats* of matroids, the non-existence results of [12] were strengthened in [10].

There are many different constructions of perfect LRCs, e.g. see [3], [9], [12] [13], [14]. Using a matroid-based construction in [10], classes of linear LRCs with a large span on the parameters $(n,k,d,r,\delta)$ and local repair sets were given. By this construction, linear perfect $(n,k,d,r,\delta)$-LRCs were constructed for all the parameters from the existence results given in [12]. Further, again by the matroid-based construction, a general lower bound was given on the maximal achievable minimum distance $d_{\max}(n,k,r,\delta)$ that a linear LRC with parameters $(n,k,r,\delta)$ can have.

### *B. Contributions*

This paper strengthens several results given in [10]. Firstly, using the matroid-based construction we extend the class of linear perfect $(n,k,d,r,\delta)$-LRCs with $\lceil k/r \rceil = 2$. Secondly, we improve the general lower bound on $d_{\max}(n,k,r,\delta)$ for linear LRCs and prove that the new bound is optimal for the matroid-based construction. The results of this paper were originally presented in the bachelor thesis of the first author [15], which provides a more comprehensive account as well as full proofs.

## II. Preliminaries

### *A. Almost Affine Locally repairable codes*

In this section, we will define an almost affine $(n,k,d,r,\delta)$-LRC. As usual, $n$ denotes the length of a codeword and $d$ its minimum (Hamming) distance. An almost affine code is defined as follows:

*Definition 2.1:* A code $C \subseteq \Sigma^n$, where $\Sigma$ is a finite set of size $s \geq 2$, is *almost affine* if for each $X \subseteq [n]$ we have $\log_s(|C_X|) \in \mathbb{Z}$.

Here $[n] = \{1,2,...,n\}$ and $C_X$ denotes the projection of the code $C$ to $\Sigma^{|X|}$, *i.e.*, $C_X = \{(c_{i_1},...,c_{i_m}) : \mathbf{c} = (c_1,...,c_n) \in C\}$, where $X = \{i_1,\ldots,i_m\} \subseteq [n]$. The parameter $k$ is, as usual, defined as $k = \log_s(|C|)$.

The local error correction of a LRC is performed inside $(r,\delta)$-*locality sets*:

*Definition 2.2:* When $1 \leq r \leq k$ and $\delta \geq 2$, an $(r,\delta)$-locality set of $C$ is a subset $S \subseteq [n]$ such that

(i) $|S| \leq r + \delta - 1$,
(ii) $d(C_S) \geq \delta$, where $d(C_S)$ is the min. distance of $C_S$.

We say that $C$ is a *locally repairable code* with *all-symbol locality* $(r,\delta)$ if every code symbol $l \in [n]$ is included in an $(r,\delta)$-locality set.

### *B. Matroids*

Matroids are combinatorial structures that capture, in an abstract sense, a certain kind of dependence common to various mathematical structures. Of the numerous equivalent matroid definitions, we will use the one utilizing the *rank function* $\rho$. In the following, $2^E$ denotes the set of all subsets of $E$.

*Definition 2.3:* A matroid $M = (E,\rho)$ is a finite set $E$ along with a rank function $\rho : 2^E \to \mathbb{Z}$ satisfying the following conditions for every subsets $X, Y \subseteq E$:

(i) $0 \leq \rho(X) \leq |X|$,
(ii) $X \subseteq Y \Rightarrow \rho(X) \leq \rho(Y)$,
(iii) $\rho(X) + \rho(Y) \geq \rho(X \cup Y) + \rho(X \cap Y)$.

This definition is for instance satisfied by the set of column vectors $E$ of a matrix over a field, and $\rho(X)$ being equal to the rank of the submatrix consisting of the column vectors indexed by $X$. If $E$ is the set of edges of an undirected graph, then a matroid is obtained by letting $\rho(X)$ be the size of a minimal spanning tree of the subgraph with edges $X$.

Next, we define some matroid concepts relevant to us. A subset $X \subseteq E$ is said to be *independent* if $\rho(X) = |X|$. The *nullity* of a set $X \subseteq E$ is defined by $\eta(X) = |X| - \rho(X)$.

A *circuit* is a dependent set $X \subseteq E$ whose all proper subsets are independent, *i.e.*, $\rho(X \setminus \{x\}) = \rho(X) = |X| - 1$ for every $x \in X$. A set $X \subseteq E$ is *cyclic* if it is a union of circuits. We denote the sets of circuits and cyclic sets of a matroid by $\mathcal{C}(M)$ and $\mathcal{U}(M)$, respectively.

The *closure* of a set $X \subseteq E$ is defined by $\mathrm{cl}(X) = \{x \in E : \rho(X \cup \{x\}) = \rho(X)\}$. A set $X \subseteq E$ is a *flat* if $X = \mathrm{cl}(X)$. A *cyclic flat* is a flat that also is a cyclic set.

The *restriction of* $M = (E,\rho)$ *to* $X$ is the matroid $M|X = (X, \rho_{|X})$ where $\rho_{|X}(Y) = \rho(Y)$ for $Y \subseteq X$.

A *lattice* is a partially ordered set for which every pair of two elements has a unique infimum, *meet*, and a unique supremum, *join*. The cyclic flats of a matroid have the property that they form a finite lattice $(\mathcal{Z}, \subseteq)$ with meet $X \wedge Y = \bigcup_{C \in \mathcal{C}(M): C \subseteq X \cap Y} C$ and join $X \vee Y = \mathrm{cl}(X \cup Y)$, for $X, Y \in \mathcal{Z}$ [16].

The least element of the lattice is the element $0_{\mathcal{Z}} \in \mathcal{Z}$ such that $X \subseteq 0_{\mathcal{Z}} \Rightarrow X = 0_{\mathcal{Z}}$ for every $X \in \mathcal{Z}$. Correspondingly, the greatest element is the element $1_{\mathcal{Z}} \in \mathcal{Z}$ such that $1_{\mathcal{Z}} \subseteq X \Rightarrow X = 0_{\mathcal{Z}}$ for every $X \in \mathcal{Z}$.

The sets of the atoms $A_{\mathcal{Z}}$ and coatoms $coA_{\mathcal{Z}}$ are defined by $A_{\mathcal{Z}} = \{X \in \mathcal{Z} \setminus \{0_{\mathcal{Z}}\} : \nexists Y \in \mathcal{Z} \text{ such that } 0_{\mathcal{Z}} \subsetneq Y \subsetneq X\}$ and $coA_{\mathcal{Z}} = \{X \in \mathcal{Z} \setminus \{1_{\mathcal{Z}}\} : \nexists Y \in \mathcal{Z} \text{ such that } X \subsetneq Y \subsetneq 1_{\mathcal{Z}}\}$, respectively.

Matroids can also be defined via this *lattice of cyclic flats*, which is our main tool for constructing and analyzing matroids in this paper. The associated axioms are presented in the following theorem:

*Theorem 2.1 ([16]):* Let $\mathcal{Z} \subseteq 2^E$ and let $\rho$ be a function $\rho : \mathcal{Z} \to \mathbb{Z}$. There is a matroid $M$ on $E$ for which $\mathcal{Z}$ is the set of cyclic flats and $\rho$ is the rank function restricted to the sets in $\mathcal{Z}$ if and only if

$(Z0)$ $\mathcal{Z}$ is a lattice under inclusion,
$(Z1)$ $\rho(0_{\mathcal{Z}}) = 0$,
$(Z2)$ $X, Y \in \mathcal{Z}$ and $X \subsetneq Y \Rightarrow$
$0 < \rho(Y) - \rho(X) < |Y| - |X|$,
$(Z3)$ $X, Y \in \mathcal{Z} \Rightarrow \rho(X) + \rho(Y) \geq$
$\rho(X \vee Y) + \rho(X \wedge Y) + |(X \cap Y) \setminus (X \wedge Y)|$.

## III. Matroids and LRCs

### A. Relationship between matroids and almost affine LRCs

The following theorem defines the associated matroid $M_C$ of an almost affine code $C$.

*Theorem 3.1 ([8]):* Let $C \subseteq \sum^n$ be an almost affine code, where $|\sum| = s$. Then $M_C = ([n], \rho_C)$ is a matroid, where

$$\rho_C(X) = \log_s(|C_X|), \text{ for } X \subseteq [n].$$

The following result can be viewed as a definition of the parameters $(n, k, d, r, \delta)$ for a matroid from the viewpoint of its cyclic flats. Hence, the parameters $(n, k, d, r, \delta)$ of an almost affine LRC $C$ can be analyzed using its associated matroid $M_C = (\rho_C, [n])$ in the theorem below.

*Theorem 3.2 ([10]):* Let $M = (E, \rho)$ be a matroid with $0 < \rho(E)$ and $1_{\mathcal{Z}} = E$. Then

- $(i)$ $n = |1_{\mathcal{Z}}|$,
- $(ii)$ $k = \rho(1_{\mathcal{Z}})$,
- $(iii)$ $d = n - k + 1 - \max\{\eta(Z) : Z \in coA_{\mathcal{Z}}\}$,
- $(iv)$ $M$ has locality $(r, \delta)$ if and only if for each $x \in E$ there exists a cyclic set $S_x \in \mathcal{U}(M)$ such that
  a) $x \in S_x$,
  b) $|S_x| \leq r + \delta - 1$,
  c) $d(M|S_x) =$
  $\eta(S_x) + 1 - \max\{\eta(Z) : Z \in coA_{\mathcal{Z}(M|S_x)}\} \geq \delta$.

### B. Matroid-based constructions of linear LRCs

The matroid-based construction of linear LRCs that is used in the constructive proofs of both [10] and this article is the following:

*Construction 1* [10]: Let $F_1, ..., F_m$ be a collection of subsets of a finite set $E$, $k$ a positive integer, and $\rho : \{F_i\}_{i\in[m]} \to \mathbb{Z}$ a function such that

$$\begin{aligned} &(i) && 0 < \rho(F_i) < |F_i| \text{ for } i \in [m], \\ &(ii) && F_{[m]} = E, \\ &(iii) && k \leq F_{[m]} - \sum_{i\in[m]} \eta(F_i), \\ &(iv) && |F_{[m]\setminus\{j\}} \cap F_j| < \rho(F_j) \text{ for all } j \in [m], \end{aligned} \tag{2}$$

where for every element $i \in [m]$ and subset $I \subseteq [m]$,

$$\begin{aligned} &\text{(a)} && \eta(F_i) = |F_i| - \rho(F_i)\,, \\ &\text{(b)} && F_I = \bigcup_{i\in I} F_i\,. \end{aligned}$$

Further, we extend $\rho$ to a function for subsets $I \subseteq [m]$ by

$$\rho(F_I) = \min\{|F_I| - \sum_{i\in I} \eta(F_i), k\}\,.$$

*Theorem 3.3 ([10]):* The previous construction defines a matroid $M(F_1, ..., F_m; k; \rho)$ which equals $M_C = ([n], \rho_C)$ for some linear LRC $C$ over a sufficiently large $\mathbb{F}_q$ such that

- (i) $\mathcal{Z} = \{F_I : I \subseteq [m], \rho(F_I) < k\} \cup E\,,$
- (ii) $n = |E|\,,$
- (iii) $k = \rho(E)\,,$
- (iv) $d = n - k + 1 - \max\{\sum_{i\in I} \eta(F_i) : F_I \in \mathcal{Z} \setminus E\}\,,$
- (v) $\delta - 1 = \min_{i\in[m]}\{\eta(F_i)\}\,,$
- (vi) $r = \max_{i\in[m]}\{\rho(F_i)\}\,.$

For each $i \in [m]$, any subset $S \subseteq F_i$ with $|S| = \rho(F_i)+\delta-1$ is a locality set of the matroid.

The motivation to use this construction comes from the fact that a matroid from it has a maximal $d$, given the matroid's set of atoms $\{F_i\}$, rank function $\rho : \{F_i\} \to \mathbb{Z}$ restricted to the atoms, and dimension $k$. This follows from the fact that its cyclic flats $F_I$ have minimal size and maximal rank, achieving the bound in Z3 when $\rho(F_I) < k$.

In a proof given later, we will use the following more specialized version of the matroid-based construction given above.

*Graph construction 1: ([10, v2])* Let $G = G(\alpha, \beta, \gamma; k, r, \delta)$ be a graph with vertices $[m]$ and edges $W$, where $(\alpha, \beta)$ are two functions $[m] \to \mathbb{Z}$, $\gamma : W \to \mathbb{Z}$, and $(k, r, \delta)$ are three integers with $0 < r < k$ and $\delta \geq 2$, such that

$$\begin{aligned} &\text{(i)} && G \text{ is a graph with no 3-cycles}, \\ &\text{(ii)} && 0 \leq \alpha(i) \leq r - 1 \text{ for } i \in [m], \\ &\text{(iii)} && \beta(i) \geq 0 \text{ for } i \in [m], \\ &\text{(iv)} && \gamma(w) \geq 1 \text{ for } w \in W, \\ &\text{(v)} && k \leq rm - \sum_{i\in[m]} \alpha(i) - \sum_{w\in W} \gamma(w), \\ &\text{(vi)} && r - \alpha(i) > \sum_{w=\{i,j\}\in W} \gamma(w) \text{ for } i \in [m]. \end{aligned} \tag{3}$$

*Theorem 3.4 ([10], v2):* Let $G(\alpha, \beta, \gamma; k, r, \delta)$ be a graph on $[m]$ such that the conditions (i)-(vi) given in (3) are satisfied. Then there is an $(n, k, r, d, \delta)$-matroid $M(F_1, ..., F_m; k; \rho)$ given by Theorem 3.3 with

(i) $n = (r + \delta - 1)m - \sum_{i\in[m]} \alpha(i) + \sum_{i\in[m]} \beta(i) - \sum_{w\in W} \gamma(w),$

(ii) $d = n - k + 1 - \max_{I\in V_{<k}} \{(\delta - 1)|I| + \sum_{i\in I} \beta(i)\},$

where

$$V_{<k} = \{I \subseteq [m] : r|I| - \sum_{i\in I} \alpha(i) - \sum_{i,j\in I, w=\{i,j\}\in W} \gamma(w) < k\}.$$

## IV. Main results

Our first result is an expanded class of parameters $(n, k, r, \delta)$ for which the generalized Singleton bound (1) can be achieved for linear LRCs. The previous bound in [10] was identical to this bound for $2a \leq r - 1$ but weaker otherwise. The parameter restrictions $0 < r < k \leq n - \lceil k/r \rceil (\delta - 1)$ and $\delta \geq 2$ are required for $(n, k, d, r, \delta)$-matroids to exist [10].

*Theorem 4.1:* Define $a = r\lceil k/r \rceil - k$ and $b = (r+\delta-1)\lceil \frac{n}{r+\delta-1} \rceil - n$, and let $(n,k,r,\delta)$ be integers such that $0 < r < k \le n - \lceil k/r \rceil (\delta-1)$, $\delta \ge 2$, $b > a \ge \lceil k/r \rceil - 1$, and $\lceil k/r \rceil = 2$. If

$$\left\lceil \frac{n}{r+\delta-1} \right\rceil \ge \lceil b/a \rceil + 1\,, \tag{4}$$

then the maximal achievable minimum distance for linear LRCs with parameters $(n,k,r,\delta)$ is

$$d_{\max} = n - k + 1 - (\lceil k/r \rceil - 1)(\delta - 1)\,.$$

*Proof:* We prove our result by giving an explicit construction of perfect matroids $M(F_1,\dots,F_m;k;\rho)$ of Thm. 3.3 for the desired parameter values.

*A matroid construction.* Let $n'$, $r'$, $\delta'$, and $k$ be integers such that $0 < r' < k \le n' - \lceil k/r' \rceil (\delta'-1)$, $\delta' \ge 2$, $b' > a'$, and $m \ge \lceil b'/a' \rceil + 1$, where we define

$$\begin{aligned} b' &= \left\lceil \frac{n'}{r'+\delta'-1} \right\rceil (r'+\delta'-1) - n', \\ a' &= \lceil k/r' \rceil r' - k, \\ m &= \left\lceil \frac{n'}{r'+\delta'-1} \right\rceil. \end{aligned}$$

Let $F_1,...,F_m = \{F_i\}_{i\in[m]}$ be a collection of finite sets with $E = \bigcup_{i\in m} F_i$ and $X \subseteq E$ a set such that

(i) $F_i \cap F_j \subseteq X$ for $i,j \in [m]$ with $i \ne j$,
(ii) $|X| = a'$,
(iii) $|F_i| = r' + \delta' - 1$ for $i \in [m]$,
(iv) $|F_i \cap X| = a'$ for $1 \le i \le \lceil b'/a' \rceil$,
(v) $|F_i \cap X| = b' - (\lceil b'/a' \rceil - 1)\,a'$ for $i = \lceil b'/a' \rceil + 1$,
(vi) $|F_i \cap X| = 0$ for $i > \lceil b'/a' \rceil + 1$.

Let $\rho$ be a function $\rho : \{F_i\}_{i\in[m]} \to \mathbb{Z}$ such that $\rho(F_i) = r'$ for each $i \in [m]$.

For the rest of the proof, we first check that this construction satisfies the conditions in (2). Then we use Theorem 3.3 to show that it yields perfect matroids (and thus linear LRCs) for the desired class of parameters $(n,k,r,\delta)$, which are shown to equal their primed counterparts. The details of this can be found in [15].

■

Our second main result is an improved lower bound for $d$. The actual improvement is the bound (6) as the bound (5) is identical to what was used in [10].

*Theorem 4.2:* Let $(n,k,r,\delta)$ be integers such that $0 < r < k \le n - \lceil k/r \rceil (\delta-1)$, $\delta \ge 2$, and $b > a$. Also let $m = \left\lceil \frac{n}{r+\delta-1} \right\rceil - 1$ and $v = r+\delta-1-b-\left\lfloor \frac{r+\delta-1-b}{m} \right\rfloor m$. Then for linear LRCs with parameters $(n,k,r,\delta)$:

If $\delta - 1 \le (\lceil k/r \rceil - 1)\left\lfloor \frac{r+\delta-1-b}{m} \right\rfloor + \min\{v, \lceil k/r \rceil - 1\}$, we have

$$d_{\max} \ge n - k + 1 - \lceil k/r \rceil (\delta - 1). \tag{5}$$

Otherwise, if $\delta - 1 > (\lceil k/r \rceil - 1)\left\lfloor \frac{r+\delta-1-b}{m} \right\rfloor + \min\{v, \lceil k/r \rceil - 1\}$, then

$$\begin{aligned} d_{\max} \ge{}& n - k + 1 - \min\{v, \lceil k/r \rceil - 1\} \\ &- (\lceil k/r \rceil - 1)\left( \left\lfloor \frac{r+\delta-1-b}{m} \right\rfloor + \delta - 1 \right). \end{aligned} \tag{6}$$

We denote the right side of the bound (6) by $d_{\text{new}}$. This bound is an improvement over its counterpart $d_{\text{old}} = n - k + 1 - \lceil k/r \rceil (\delta-1) + (b - r)$ in [10], since

$$d_{\text{new}} - d_{\text{old}} \ge \left\lfloor \frac{r+\delta-1-b}{m} \right\rfloor (m - \lceil k/r \rceil + 1) \ge 0. \tag{7}$$

*Proof:* Let $n' \in \mathbb{Z}$ be such that it satisfies the conditions for $n$ in Theorem 4.2.

*A graph construction.* Let $G(\alpha,\beta,\gamma;k,r,\delta)$ be intended as an instance of Graph construction 1 with

$$\begin{aligned} &\text{(a)} \quad m = \left\lceil \frac{n'}{r+\delta-1} \right\rceil - 1, \\ &\text{(b)} \quad W = \emptyset, \\ &\text{(c)} \quad \alpha(i) = 0 \text{ for } i \in [m], \\ &\text{(d)} \quad \beta(i) = \begin{cases} \left\lceil \frac{r+\delta-1-b'}{m} \right\rceil & \text{for } 1 \le i \le v', \\ \left\lfloor \frac{r+\delta-1-b'}{m} \right\rfloor & \text{for } v' < i \le m, \end{cases} \end{aligned} \tag{8}$$

where $b' = \left\lceil \frac{n'}{r+\delta-1} \right\rceil (r+\delta-1) - n'$ and $v' = r+\delta-1-b'-\left\lfloor \frac{r+\delta-1-b'}{m} \right\rfloor m$.

The rest of the proof consists of checking that the conditions in (3) are satisfied and using Theorem 3.4 to show that the construction yields the expected $d$ for all desired parameter sets $(n,k,r,\delta)$. Finally, the inequalities in (7) will be proved. A full version of the proof can be found in [15].

■

*Example 4.1:* To see that the difference $d_{\text{new}} - d_{\text{old}}$ is not identically zero, consider for instance the graph construction used in the proof with parameter values $n' = 139, k = 60, r = 20, \delta = 21$.

Lastly, we show that the bound in Thm. 4.2 for matroids (linear LRCs) from Construction 1 is tight for parameter sets $(n,k,r,\delta)$ for which there exists no perfect matroid (linear LRC) from Construction 1.

*Theorem 4.3:* Let $(n,k,r,\delta)$ be integers such that there exists no perfect $(n,k,d',r,\delta)$-matroid from Construction 1. Let $M$ be an $(n,k,d,r,\delta)$-matroid from Construction 1 and let us denote the bound in Theorem 4.2 by $d_b = d_b(n,k,r,\delta)$. Then $d \le d_b$.

*Proof:* A more detailed proof can be found in [15]. Let $M = M(F_1,...,F_m;k;\rho)$ be a matroid from Construction 1 for which there exists no perfect matroid from the same construction with the same parameters $(n,k,r,\delta)$.

Assume that $\max\{|I| : F_I \in \mathcal{Z}_{<k}\} \ge \lceil k/r \rceil$. Using Theorem 3.3 (iii), we then obtain $d \le n-k+1-\lceil k/r \rceil(\delta-1)$, as $\eta(F_i) \ge \delta - 1$ for every $i \in [m]$.

Thus the theorem holds in this case and we are only left with the case $\max\{|I| : F_I \in \mathcal{Z}_{<k}\} = \lceil k/r \rceil - 1$, as we easily see that $\max\{|I| : F_I \in \mathcal{Z}_{<k}\} < \lceil k/r \rceil - 1$ is impossible.

There must be an atom $F_i$ with $\eta(F_i) > \delta - 1$, since otherwise the matroid would be perfect. Next we show that our current assumptions imply $m < \lceil \frac{n}{r+\delta-1} \rceil$. We do this by showing that $m \geq \lceil \frac{n}{r+\delta-1} \rceil$ would allow the existence of perfect matroids, which is a contradiction. The perfect matroids are constructed by, roughly speaking, repeatedly decreasing the nullity of atoms $F_u$ with $\eta(F_u) > \delta - 1$ by an element of $F_u$ to another atom $F_i$. which either has $\rho(F_i) < r$ or overlaps with another atom $F_k$. In the former case, $\rho(F_i)$ will be increased by one, and in the latter case, the element in the intersection will no longer be part of $F_i$.

Let us denote $s = \sum_{i\in[m]} \eta(F_i)$. Let us distribute this nullity evenly among the atoms $F_i$, *i.e.*, set

$$\eta(F_i) = \begin{cases} \lceil s/m \rceil & \text{for } 1 \leq i \leq s - \lfloor s/m \rfloor m, \\ \lfloor s/m \rfloor & \text{for } s - \lfloor s/m \rfloor m < i \leq m. \end{cases}$$

For minimizing $\max\left\{\sum_{i\in I} \eta(F_i) : |I| = \lceil k/r \rceil - 1\right\}$, this setup is clearly optimal and yields the bound

$$\max\left\{\sum_{i\in I} \eta(F_i) : |I| = \lceil k/r \rceil - 1\right\} \\ \geq (\lceil k/r \rceil - 1) \lfloor s/m \rfloor + \min\{\lceil k/r \rceil - 1, s - \lfloor s/m \rfloor m\}. \tag{9}$$

The bound in (9) is clearly increasing as a function of $s$, and $s$ is bounded by $s \geq n - rm$. Thus we obtain the bound

$$\max\left\{\sum_{i\in I} \eta(F_i) : |I| = \lceil k/r \rceil - 1\right\} \geq (\lceil k/r \rceil - 1) \left\lfloor \frac{n-rm}{m} \right\rfloor \\ + \min\left\{\lceil k/r \rceil - 1, n - rm - \left\lfloor \frac{n-rm}{m} \right\rfloor m\right\}. \tag{10}$$

This bound is in turn decreasing as a function of $m$ and we can obtain a new bound by substituting $m = \lceil \frac{n}{r+\delta-1} \rceil - 1$. By additionally substituting $v$ and $b$ by their definitions in (6), we can see that the bounds (6) and (10) are equal.

We have thus proved that the value of $d$ for non-perfect matroids is always bounded from above by either the bound (5) or the bound (6). This proves the theorem.

■

*Remark 4.1:* The class of matroids constructed in (2) constitutes a small subclass of the class of matroids called gammoids [10]. A method of constructing linear codes from gammoids can be extracted by using [17]. The smallest field size required by LRCs is an important issue, since it affects the computational complexity of the code. In general for gammoids there is a known upper bound for the field size, $2^n$ [17]. However, we are convinced that this bound is not tight for the construction given in (2). We have ongoing research on explicit constructions of linear LRCs over small fields obtained from (2) and conjecture an upper bound on the smallest field size that is polynomial with $n$. However, explicit constructions of LRCs for the matroid-based construction given in (2) are out of the scope of this paper.

## V. Conclusions

In this paper, we provided an expanded class of parameters for which perfect linear LRCs exist (Thm. 4.1). We also gave a general lower bound for the maximal minimum distance $d_{max}$ (Thm. 4.2), which we proved to be optimal for sub-perfect LRCs from Construction 1 (Thm. 4.3).

These theorems suggest the following two-stage approach for solving $d_{max}(n, k, r, \delta)$ for almost affine LRCs: The first goal is to derive an expression for $d_{max}$ restricted to sub-perfect LRCs. Then, full knowledge of $d_{max}$ would be achieved by determining the class of parameters $(n, k, r, \delta)$ for which perfect LRCs exist.

Theorem 4.3 is an attempt at accomplishing the first task. It is only a partial result towards this goal as it is limited to matroids from Construction 1. However, matroids from Construction 1 have a maximal $d$ given their setup of atoms, which suggests that the bound in Theorem 4.2 is tight or almost tight in the general case for sub-perfect matroids.

Theorem 4.1 in turn is an addition to the existing results on for which parameter values perfect matroids exist. A complete solution of this second question would seem to require solving hard problems of extremal set theory.

## References


[1] D. S. Papailiopoulos and A. G. Dimakis, “Locally repairable codes,” *IEEE Trans. Inf. Theory*, vol. 60, no. 10, pp. 5843–5855, 2014.

[2] P. Gopalan, C. Huang, H. Simitci, and S. Yekhanin, “On the locality of codeword symbols,” *IEEE Trans. Inf. Theory*, 58(11), pp. 6925–6934, 2012.

[3] N. Prakash, G. M. Kamath, V. Lalitha, and P. V. Kumar, “Optimal linear codes with a local-error-correction property,” in *2012 IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 2776–2780.

[4] D. S. Papailiopoulos, and A. G. Dimakis, “Locally repairable codes,” *2012 IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 2771–2775.

[5] V. Cadambe and A. Mazumdar, “An upper bound on the size of locally recoverable codes”, In *Proc. IEEE Symp. Netw. Coding*, pp. 1–5, Jun. 2013.

[6] A. S. Rawat, A. Mazumdar and S. Vishwanath “Cooperative local repair in distributed storage,” EURASIP J. Adv. Sign. Proc, online 2015.

[7] I. Tamo, A. Barg and A. Frolov, "Bounds on the parameters of locally recoverable codes", arXiv: 1506.07196.

[8] J. Simonis and A. Ashikhmin, “Almost affine codes”, *Design, codes and cryptography*, 14, pp. 179–197, 1998.

[9] I. Tamo, D. S. Papailiopoulos, and A. G. Dimakis, “Optimal locally repairable codes and connections to matroid theory,” in *2013 IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 1814–1818.

[10] T. Westerbäck, R. Freij-Hollanti, T. Ernvall, and C. Hollanti, “On the combinatorics of locally repairable codes via matroid theory.” arXiv: 1501.00153.

[11] T. Westerbäck, R. Freij-Hollanti, and C. Hollanti, “Applications of polymatroid theory to distributed storage systems,” *in proc. 53rd Annual Allerton Conf. on Comm. Control*, 2015.

[12] W. Song, S. H. Dau, C. Yuen, and T. J. Li, “Optimal locally repairable linear codes,” *IEEE J. Sel. Areas Commun.*, vol. 32, no. 5, pp. 1019–1036, 2014.

[13] N. Silberstein, A. S. Rawat, O. O. Koyluoglu, and S. Vishwanath, “Optimal locally repairable codes via rank-metric codes,” in *2013 IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 1819–1823.

[14] I. Tamo and A. Barg, “A family of optimal locally recoverable codes,” in *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 4661–4676, 2014.

[15] A. Pöllänen, “Locally repairable codes and matroid theory,” bachelor thesis, Aalto University, arXiv: 1512.05325, 2015.

[16] J. E. Bonin and A. De Mier, “The lattice of cyclic flats of a matroid,” *Annals of Combinatorics*, vol. 12, no. 2, pp. 155–170, 2008.

[17] B. Lindström, “On the vector representations of induced matroids,” *Bull. London Math. Soc.*, vol. 5, no. 1, pp. 85–90, 1973.