Grezet, Matthias; Freij-Hollanti, Ragnar; Westerbäck, Thomas; Olmez, Oktay; Hollanti, Camilla

Bounds on Binary Locally Repairable Codes Tolerating Multiple Erasures

Published in:
The International Zurich Seminar on Information and Communication (IZS 2018) Proceedings

DOI:
10.3929/ethz-b-000245075

Published: 21/02/2018

Please cite the original version:
Bounds on Binary Locally Repairable Codes Tolerating Multiple Erasures

Author(s):
Grezet, Matthias; Freij-Hollanti, Ragnar; Westerbäck, Thomas; Olmez, Oktay; Hollanti, Camilla

Publication Date:
2018-02

Permanent Link:
https://doi.org/10.3929/ethz-b-000245075

Rights / License:
In Copyright - Non-Commercial Use Permitted
Bounds on Binary Locally Repairable Codes Tolerating Multiple Erasures

Matthias Grezet∗, Ragnar Freij-Hollanti †, Thomas Westerbäck∗, Oktay Olmez‡ and Camilla Hollanti∗

∗Department of Mathematics and Systems Analysis, Aalto University, Finland, Email: firstname.lastname@aalto.fi
†Department of Electrical and Computer Engineering, Technical University of Munich, Germany, Email: ragnar.freij@tum.de
‡Department of Mathematics, Ankara University, Turkey, Email: oolmez@ankara.edu.tr

Abstract—Recently, locally repairable codes have gained significant interest for their potential applications in distributed storage systems. However, most constructions in existence are over fields with size that grows with the number of servers, which makes the systems computationally expensive and difficult to maintain. Here, we study linear locally repairable codes over the binary field, tolerating multiple local erasures. We derive bounds on the minimum distance on such codes, and give examples of LRCs achieving these bounds. Our main technical tools come from matroid theory, and as a byproduct of our proofs, we show that the lattice of cyclic flats of a simple binary matroid is atomic.

I. INTRODUCTION

In modern distributed storage systems (DSSs) failures happen frequently, whence decreasing the number of connections required for node repair is crucial. Removing even one connection locally can easily imply huge gains in the overall system functionality, thanks to shortened queues and improved data availability. Consequently, locally repairable codes (LRCs) have gained a lot of interest in the past few years [1]–[3]. Namely, LRCs allow to repair a small number of failures locally, i.e., by only contacting few close-by nodes and hence avoiding congesting the system. Related Singleton-type bounds have been derived for various cases, see [4], [5]. The first bound on the minimum distance for fixed field size was obtained by Cadambe and Mazumdar in [6]. Recently, this bound was improved and generalized, via the observation that any log-convex bound on the “local rank” of a code can be blown up to obtain a bound on the global rank [7]. Interestingly, the bounds in [7] do not depend on the linearity of the code. However, all the bounds in [6], [7] are implicit, except for special classes of codes. For more details and tentative comparison, see the last section of this paper.

In this paper, we consider binary codes motivated by the fact that the computational complexity when retrieving a file or repairing a node grows with the field size. We derive new, improved Singleton-type bounds for this special case alongside with sporadic examples, in particular when the local repair sets can tolerate multiple failures. In contrast to the bounds in [6], [7], our bounds are explicit, and do not depend on any prior bounds on binary codes of shorter length.

As our main contribution, in Theorem 5, we obtain a closed-form bound on the minimum distance $d$ of a binary $(n,k)$-code of length $n$ and dimension $k$ and with all-symbol $(r,\delta)$-locality, where the local distance $\delta > 2$. Such bounds were previously only known when $\delta = 2$. The bound is in terms of the rank $\ell$ of the repair sets, but can easily be transformed to bounds in terms of the size $r+\delta-1$. Interestingly, while the two parameters $r$ and $\ell$ can be assumed to agree when $\delta = 2$, as well as when the field size is unbounded, this is no longer the case over the binary field with $\delta > 2$. While both parameters are of independent interest in applications, we have chosen to focus on the number of nodes $\ell$ that need to be contacted for local repair, rather than on the size of the local clouds.

In addition, in Section III we prove that every element of a non-degenerate binary locally repairable code without replication is contained in an atomic cyclic flat, and hence that the lattice of cyclic flats is atomic. From a practical point of view, this implies a hierarchy of failure tolerance, as explained in the end of Section III. In particular, whenever a symbol $c$ fails, we can start by downloading nodes in an atomic cyclic flat in order to repair $c$. If it turns out that some other nodes in this local set have failed as well, we can repair them while still keeping the part that we already downloaded, and simply contact some more nodes in the corresponding repair set to repair all the failed symbols. Thus, we do not have to restart from the beginning if we find out during the repair process that a small amount of other nodes have failed as well.

Several constructions are known for optimal LRCs over the binary field, for specified ranges of parameters, and almost exclusively in the case $\delta = 2$. The first such construction, for codes with exponentially low rate and locality $r = 2,3$, was obtained by deleting carefully chosen columns from the simplex code [8]. These constructions are also optimal when taking the availability $t$, i.e., the number of disjoint sets that can recover a given symbol, into consideration. A slightly more flexible family of codes, allowing for higher rate, was given in [9], [10], where also a slight improvement over the Cadambe–Mazumdar bound was given for linear codes. In the realm of multiple erasures, i.e., when $\delta > 2$, rate-optimal codes were studied in [11]. There, rate-optimal codes for short length codes were characterized when $\delta = 3$, and analogous constructions without optimality proof were given for $\delta > 3$. However, to the best of our knowledge, no previous work has studied bounds on the global minimum distance in the regime $\delta > 2$.

In the interest of space, we have relegated proofs to an extended version of this paper available on arXiv [12].
As is common practice, we say that $C$ is an $(n, k, d)$-code if it has length $n$, dimension $k$, and minimum Hamming distance $d$. A linear $(n, k, d)$-code $C$ over a finite field $\mathbb{F}$ is a non-degenerate storage code if $d \geq 2$ and there is no zero column in a generator matrix of $C$. For a fixed code $C$, we denote by $d_C$ the minimum Hamming distance of the punctured code $C|Y$ where $Y \subseteq [n]$ is a set of coordinates of the code $C$.

**Definition 1.** Let $C$ be an $(n, k, d)$-code over $\mathbb{F}^n$. A symbol $x \in [n]$ has locality $(r, \delta)$ if there exists a subset $R \subseteq [n]$, called repair set of $x$, such that $x \in R$, $|R| \leq r + \delta - 1$, and $d_C \geq \delta$.

**Definition 2.** A linear $(n, k, d, r, \delta)$-LRC over a finite field $\mathbb{F}^n$ is a non-degenerate linear $(n, k, d)$-code $C$ over $\mathbb{F}$ such that every coordinate $x \in [n]$ has locality $(r, \delta)$. In the literature, this is specifically called all-symbol locality.

The parameters $(n, k, d, r, \delta)$ can immediately be defined and studied for matroids in general, as in [2], [5], [13].

**a) Matroid fundamentals:** Matroids have many equivalent definitions in the literature. Here, we choose to present matroids via their rank functions.

**Definition 3.** (A finite) matroid $M = (E, \rho)$ is a finite set $E$ together with a rank function $\rho: 2^E \to \mathbb{Z}$ such that for all subsets $X, Y \subseteq E$:

1. $0 \leq \rho(X) \leq |X|$,  
2. $X \subseteq Y \Rightarrow \rho(X) \leq \rho(Y)$,  
3. $\rho(X) + \rho(Y) \geq \rho(X \cup Y) + \rho(X \cap Y)$.

A subset $X \subseteq E$ is called independent if $\rho(X) = |X|$. If $X$ is independent and $\rho(X) = \rho(E)$, then $X$ is called a basis. A subset that is not independent is called dependent. A circuit is a minimal dependent subset of $E$, that is, a dependent set whose proper subsets are all independent. Strongly related to the rank function is the nullity function $\eta : 2^E \to \mathbb{Z}$, defined by $\eta(X) = |X| - \rho(X)$ for $X \subseteq E$.

There is a straightforward connection between linear codes and matroids. Indeed, any linear code $C$ over a field $\mathbb{F}$ generates a matroid $M_C = (E, \rho_C)$, where $E$ is the set of coordinates of $C$, and $\rho_C(X)$ is the dimension of the punctured code $C|X$. For a given set $X \subseteq E$, we define the restriction of $M$ to $X$ to be the matroid $M|X = (X, \rho_X)$ by $\rho_X(Y) = \rho(Y)$ for all subsets $Y \subseteq X$.

Two matroids $M_1 = (E_1, \rho_1)$ and $M_2 = (E_2, \rho_2)$ are isomorphic if there exists a bijection $\psi : E_1 \rightarrow E_2$ such that $\rho_2(\psi(X)) = \rho_1(X)$ for all subsets $X \subseteq E_1$.

**Definition 4.** A matroid that is isomorphic to $M_C$ for some code $C$ over $\mathbb{F}$ is said to be representable over $\mathbb{F}$. We also say that such a matroid is $\mathbb{F}$-representable. A binary matroid is a matroid that is $\mathbb{F}_2$-representable.

**Definition 5.** A matroid is called simple if it has no circuits consisting of 1 or 2 elements. A element $e \in E$ is called a co-loop if $\rho(E - e) < \rho(E)$.

**b) Fundamentals on cyclic flats:** The main tool from matroid theory in this paper are the cyclic flats. We will define them using the closure and cyclic operators.

Let $M = (E, \rho)$ be a matroid. The closure operator $cl : 2^E \rightarrow 2^E$ and cyclic operator $cyc : 2^E \rightarrow 2^E$ are defined by

$$(i) \quad cl(X) = X \cup \{e \in E - X : \rho(X \cup e) > \rho(X)\},$$

$$(ii) \quad cyc(X) = \{e \in E : \rho(X - e) = \rho(X)\}.$$

A subset $X \subseteq E$ is a flat if $cl(X) = X$ and a cyclic set if $cyc(X) = X$. Therefore, $X$ is a cyclic flat if $\rho(X \cup y) > \rho(X)$ and $\rho(X - x) = \rho(X)$ for all $y \in E - X$ and $x \in X$. The collection of flats, cyclic sets, and cyclic flats of $M$ are denoted by $\mathcal{F}(M)$, $\mathcal{U}(M)$, and $\mathcal{Z}(M)$, respectively. Some more fundamental properties of flats, cyclic sets, and cyclic flats are given in [14].

If $C \subseteq \mathbb{F}^E$ is a linear code, then the cyclic flats of $M_C$ can be described as subsets $X \subseteq E$ such that $|C|(X \cup y) > |C|X$ and $|C|(X - x) = |C|X$ for all $y \in E - X$ and $x \in X$.

Before going deeper in the study of $\mathcal{Z}(M)$, we need a minimum background on poset and lattice theory. We will use the standard notation of $\land$ and $\lor$ for the meet and join operator, we will denote by $0_L$ and $1_L$ the bottom and top element of a lattice $L$, and we will denote by $\prec$ the covering relation, i.e., for $X, Y \in \mathcal{L}$, we say that $X \prec Y$ if $X \neq Y$ and there is no $Z \in \mathcal{L}$ with $X < Y < Z$.

The atoms and coatoms of a lattice $(\mathcal{L}, \subseteq)$ are defined as $A_L = \{X \in \mathcal{L} : 0_L \prec X\}$ and $coA_L = \{X \in \mathcal{L} : X \prec 1_L\}$, respectively. A lattice $\mathcal{L}$ is said to be atomic if every element of $\mathcal{L}$ is the join of atoms.

We can now give a crucial property of the set of cyclic flats.

**Proposition 1** (See [14]). Let $M = (E, \rho)$ be a matroid. Then

1) $\mathcal{Z}(M)$ is a lattice with $X \lor Y = cl(X \cup Y)$ and $X \land Y = cyc(X \cap Y)$.

2) $1_L = cyc(\emptyset)$ and $0_L = cl(\emptyset)$.

c) Relation between LRCs and the lattice of cyclic flats: Recently, some work has been done to emphasise the relation between cyclic flats and linear codes over finite fields. In [5], the authors proved that the minimum distance can be expressed in terms of the nullity of certain cyclic flats.

**Proposition 2.** Let $C$ be a non-degenerate $(n, k, d)$-code and $M = (E, \rho)$ the matroid associated to $C$. Then,

$$d = \eta(E) + 1 - \max \{\eta(Z) : Z \in \mathcal{Z}(M) - \{E\}\}.$$
• \( \rho(Y) - \rho(X) = l > 1 \) and \( \eta(Y) - \eta(X) = 1 \). We call such an edge in the Hasse diagram of \( Z(M) \) a rank edge and label it \( \rho = l \).
• \( \rho(Y) - \rho(X) = 1 \) and \( \eta(Y) - \eta(X) = l > 1 \). We call such an edge a nullity edge and label it \( \eta = l \).
• \( \rho(Y) - \rho(X) = 1 \) and \( \eta(Y) - \eta(X) = 1 \). We call such an edge a elementary edge.

**Theorem 1** (Announced in [15]). Let \( C \) be a non-degenerate, binary linear \((n, k, d, r, \delta)\)-LRC with \( d > 2 \) and without replication. Let \( M = (E, \rho) \) be the associated matroid. Then \( Z = Z(M) \) satisfies the following:

1) \( \emptyset \) and \( E \) are cyclic flats.
2) Every covering relation \( Z \prec E \) is a nullity edge labeled with a number \( \geq d - 1 \).
3) If \( \delta = 2 \), then for every \( i \in E \), there is \( X \in Z \) with \( i \in X \) such that \( \rho(X) \leq r \).
4) If \( \delta > 2 \), then for every \( i \in E \), there is \( X \in Z \) with \( i \in X \) such that
   a) Every covering relation \( Y \prec X \) is a nullity edge labeled with a number \( \geq d - 1 \).
   b) Every cyclic flat \( Y \) with \( Y \prec X \) has rank \( \leq r - 2 \).

**III. LATTICE STRUCTURE AND REPAIR PROPERTIES**

The first part of this section is devoted to understanding how restricting to binary linear codes affects the structure of the lattice of cyclic flats. The main result of this section consists of proving that the lattice of cyclic flats has the property of being atomic.

In the second part, we will discuss the meaning of these results for binary linear codes and LRCs. In particular, we will see that every non-degenerate binary \((n, k, d)\)-code with \( d \geq 2 \) and without replication is already a binary linear \((n, k, d, 2)\)-LRC for a certain \( \delta \). Furthermore, for LRCs with \( \delta > 2 \), we will see that these codes have a hierarchy in failure tolerance.

**Proposition 4.** Let \( C \) be a binary linear \((n, k, d)\)-code. Then \( C \) is non-degenerate with no replication if and only if the associated matroid \( M = (E, \rho) \) is simple and contains no co-loops.

Now that we have established the type of matroids that is relevant to our case, we can study the implications of Proposition 3 for the lattice of cyclic flats. The following lemma, in addition to being a crucial step towards proving Theorem 2, has even stronger implications, as it shows that any element in \( Z(M) \) is equal not only to the join, but also to the union of all the atoms it contains.

**Lemma 1.** Let \( M = (E, \rho) \) be a simple binary matroid that contains no co-loops, \( e \) an element of \( E \), and \( C \) a circuit of \( M \). We have the following results:

1) \( Z \) is an atom of \( Z(M) \) if and only if \( \eta(Z) = 1 \).
2) \( \text{cl}(C) \) is an atom of \( Z(M) \) if and only if \( \text{cl}(C) = C \).
3) If \( C \) is a circuit containing \( e \) of minimal length, then \( \text{cl}(C) = C \).
4) Every element \( e \in E \) is contained in an atom.

**Theorem 2.** Let \( M = (E, \rho) \) be a simple binary matroid that contains no co-loops. Then the lattice of cyclic flats \( Z(M) \) is atomic.

b) **Hierarchy of failure tolerance:** By Proposition 4, we can reinterpret Lemma 1 and Theorem 2 as statements about non-degenerate binary storage codes. Indeed, combining Lemma 1.4 with Proposition 2, we see that every symbol is directly contained in a small repair set with \( \delta = 2 \), i.e., in a repair set that can correct exactly one erasure. Hence we obtain the following theorem.

**Theorem 3.** For every non-degenerate binary linear \((n, k, d)\)-code \( C \) with no replication, \( C \) is also an \((n, k, d, r', 2)\)-LRC for some \( r' \in \{2, \ldots, k\} \).

Now, if we want to be able to correct more than one erasure, then the repair sets cannot be atoms of \( Z(M) \) as these have \( d_2 = 2 \). They have to be at least one level above some atoms. However, the previous theorem still holds, meaning that for every symbol, there is also an atom containing it. Thus, we get a natural hierarchy in failure tolerance. If one node fails, then we can contact the close-by nodes in the atom to repair it. If more nodes fail, but no more than \( \delta - 1 \), we can contact other repair sets to fix them. And if more than \( \delta - 1 \) nodes fail, then we need to use the global properties of the code.

Moreover, by the remark following Theorem 2, it follows that repair sets are unions of all the atoms below them. Since the collection of repair sets contains every symbol, we can choose the collection of atoms that will give us the 

**Corollary 1.** Let \( C \) be a non-degenerate binary linear \((n, k, d, r, \delta)\)-LRC with no replication and with \( \delta > 2 \). Let \( \{R_i\}_{i \in I} \) be the list of repair sets. Then, there exists a collection of sets \( \{X_j\}_{j \in J} \) such that for all \( X_j \), there exists \( R_i \) with \( X_j \subseteq R_i \) such that \( C \) is also an \((n, k, d, r', 2)\)-LRC.

From a practical point of view, this reinforces the usefulness of the failure tolerance hierarchy. For example, suppose that the symbol \( e \in E \) fails. We can start by downloading nodes in the atom \( Z^\text{at} \) in order to repair \( e \). Now, if we realize that some other nodes in \( Z^\text{at} \) have failed as well, we can keep the part that we already downloaded from \( Z^\text{at} \) and contact more nodes in the corresponding repair set to repair all the failed symbols in \( Z^\text{at} \). Thus, we do not have to restart from the beginning if we find out during the repair process that a small amount of other nodes have failed as well.

**IV. IMPROVING THE SINGLETON-TYPE BOUND FOR \( \delta > 2 \)**

The goal of this section is to improve the existing bound for non-degenerate linear \((n, k, d, r, \delta)\)-codes \( C \) when the codes...
are binary, contain no replication and $\delta > 2$. It has been proven in [16] that, for a linear $(n, k, d, r, \delta)$-code over $F_q$, we have
\[ d \leq n - k + 1 - \left(\left\lceil \frac{k}{r} \right\rceil - 1 \right) (\delta - 1). \] (1)
We start by defining a new parameter that represents the maximum rank of a repair set.

**Definition 6.** Let $C$ be an $(n, k, d, r, \delta)$-LRC. Let $\{R_i\}_{i \in I}$ be the list of repair sets and $M = (E, \rho)$ the associated matroid to $C$. We define $\ell$ to be $\ell := \max_{i \in I} \rho(R_i)$.

As mentioned in the introduction, even if the maximum rank and $r$ can be assumed to coincide over large field size, this is no longer the case for binary codes with $\delta > 2$. Indeed, as binary MDS codes do not exist when the minimum distance is greater than 2, we must have $\ell < r$ for such codes. Moreover, any bound on the rank of a binary code in terms of its size and minimum distance, gives a bound on $\ell$ in terms of $r$, and as a consequence a new translation of our bounds to a bound in terms of $r$. As our main interest is in small values of $\delta$, Proposition 5 is strong enough for the purposes of this paper.

**Proposition 5.** Let $C$ be a non-degenerate, binary linear $(n, k, d, r, \delta)$-LRC with $\delta > 2$ and without replication. Let $\{R_i\}_{i \in I}$ be the collection of repair sets. Then,
\[ \ell \leq r - 1 \quad \text{and} \quad \eta(R_i) \geq \delta \quad \text{for all } i \in I. \]

Since we will focus on the rank of the repair sets rather than on the size, we assume from now on that repair sets are maximal after fixing its rank or, in matroid terms, that repair sets are cyclic flats. We will denote these repair sets by $Z_i$ instead of $R_i$ to avoid confusion.

Remember that Proposition 2 links the minimum distance of a code to the coatoms among the cyclic flats. We would like to construct a cyclic flat that is close to the coatoms level, to give an accurate lower bound on $\max \{\eta(Z) : Z \in Z(M) - \{E\}\}$. We will do this by creating a chain in $Z(M)$ of joins of repair sets and we will call these type of chains repair-sets-chain or, for short, $rps$-chain.

**Definition 7.** Let $C$ be a non-degenerate $(n, k, d, r, \delta)$-code and $\{Z_i\}_{i \in I}$ the collection of repair sets. Let $M = (E, \rho)$ the associated matroid. An $rps$-chain
\[ \emptyset = Y_0 \subseteq Y_1 \subseteq \ldots \subseteq Y_m = E \]
is a chain in $Z(M)$ defined inductively by
1. Let $Y_0 = \emptyset$.
2. Given $Y_{i-1} \subseteq E$, we choose $x_i \in E \setminus Y_{i-1}$ and $Z_i$ with $x_i \in Z_i$ arbitrarily, and assign $Y_i = Y_{i-1} \cup Z_i$.
3. If $Y_i = E$, we set $m = i$.

Notice that this chain is not uniquely defined since we can choose symbols and corresponding repair sets freely.

Now, we are interested in how the rank and the nullity can increase at each step. To bound the rank and nullity difference at each step of the $rps$-chain $(Y_i)_{i=0}^m$, one can use the rank axioms to obtain $\rho(Y_i) - \rho(Y_{i-1}) \leq \ell$ and $\eta(Y_i) - \eta(Y_{i-1}) \geq \delta - 1$. However, this does not take binarity, and in particular Proposition 3, into account. Indeed, if $\rho(Y_i) - \rho(Y_{i-1}) = \ell$, then we must have $Y_{i-1} \cap Z_i = \emptyset$ and $\rho(Z_i) = \ell$. Hence, there is no code nor an $rps$-chain that can simultaneously achieve both bounds.

Next, we introduce an indicator function that will capture when the intersection is a coatom of a repair set. To be more concise, we will denote by $A_i$ the event that $Y_{i-1} \cap Z_i \in \text{co}A_{Z(M)|Z_i}$. First, this is a necessary condition to have $\eta(Y_i) - \eta(Y_{i-1}) = \delta - 1$. Secondly, this will also imply that $\rho(Y_i) - \rho(Y_{i-1}) = 1$ since every covering relation of a repair set has a nullity edge by Proposition 3. The following lemma summarizes these observations.

**Lemma 2.** Let $C$ be a non-degenerate, binary linear $(n, k, d, r, \delta)$-LRC with $\delta > 2$ and without replication, and let $M = (E, \rho)$ be the associated matroid. Let $\{Z_i\}_{i \in I}$ be the collection of repair sets and $(Y_i)_{i=0}^m$ an associated $rps$-chain. Then $(Y_i)_{i=0}^m$ has the following properties:
1. $\rho(Y_i) - \rho(Y_{i-1}) \leq \ell - (\ell - 1) \eta(A_i)$ for all $i = 2, \ldots, m$.
2. $\eta(Y_i) - \eta(Y_{i-1}) \geq \delta - \eta(A_i)$ for all $i = 2, \ldots, m$.

In order to use Lemma 2 to get a Singleton-type bound with these assumptions, we need to define a new parameter that will count the number of times the intersection $Y_{i-1} \cap Z_i$ is a coatom of $Z_i$.

**Definition 8.** Let $(Y_i)_{i=0}^m$ be an $rps$-chain. Define $0 \leq \alpha \leq 1$ by $\alpha m = \# \{i : Y_{i-1} \cap Z_i \in \text{co}A_{Z(M)|Z_i}\}$. We say that $(Y_i)_{i=0}^m$ is an $rps_\alpha$-chain.

We can now derive a new Singleton-type bound with the extra parameter $\alpha$.

**Theorem 4.** Let $C$ be a non-degenerate, binary linear $(n, k, d, r, \delta)$-LRC with $\delta > 2$ and without replication. Let $\alpha \in [0, 1]$ be such that $C$ has an $rps_\alpha$-chain. Then,
\[ d \leq n - k + 1 + \delta \left(\left\lceil \frac{k}{\ell - (\ell - 1)\alpha} \right\rceil (\delta - \alpha) \right). \]

Since this bound is valid for all $\alpha$, we can optimize $\alpha$ to get a bound for all types of $rps$-chain, i.e., a Singleton-type bound that only depends on the parameters $n, k, d, \ell$ and $\delta$.

**Theorem 5.** Let $C$ be a non-degenerate, binary linear $(n, k, d, r, \delta)$-LRC with $\delta > 2$ and without replication. Then,
\[ d \leq n - k + 1 - \left(\left\lceil \frac{k}{r - 1} \right\rceil - 1 \right) \delta + 1 \ell \ell(k-1) \quad \text{and} \quad r \neq k. \] (2)

In order to make the comparison to the previously known bound (1) easier and to emphasize the improvement provided by the new bound, we state the following corollary of Theorem 5 using Proposition 5. This gives us a bound that only depends on $n, k, d, r$ and $\delta$.

**Corollary 2.** Let $C$ be a non-degenerate, binary linear $(n, k, d, r, \delta)$-LRC with $\delta > 2$ and without replication. Then,
\[ d \leq n - k + 1 - \left(\left\lceil \frac{k}{r - 1} \right\rceil - 1 \right) \delta + 1 \ell \ell(k-1) \quad \text{and} \quad r \neq k. \] (3)

We provide one small example that achieves the bound from Corollary 2.
Example 1. Let $C$ be the binary linear $(10,4,4)$-code given by the following generator matrix,
\[
G = \begin{bmatrix}
1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\
0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1
\end{bmatrix}.
\]

We can define three repair sets by their corresponding columns in $G$, $Z_1 = \{1, 2, 3, 5, 6, 8\}$, $Z_2 = \{2, 3, 6, 7, 9, 10\}$, and $Z_3 = \{1, 4, 6, 7, 8, 10\}$.

For all $i \in \{1, 2, 3\}$, we have $|Z_i| = 6$, $\rho(Z_i) = 3$, and $d_{Z_i} = 3$, and hence we obtain a binary linear $(10, 4, 4, 3)$-LRC that achieves the bound from Corollary 2.

The following graph is a comparison of the previous known Singleton-type bound (1) and the new one from Corollary 2 for two different values of $r$. We can see that the new bound is always better than (or equivalently, smaller) or equal to the previous bound.

![Comparison of the previous Singleton-type bound (1) and the new bound (3) for $n = 50$ and $\delta = 3$.](image)

The most commonly used field-dependent distance bound for LRCs is the Cadambe–Mazumdar bound [6], which is only stated without reference to the local minimum distance (or equivalently, for $\delta = 2$). It can straightforwardly be generalized to codes tolerating more than one local erasure, as also noted in [17]:

**Proposition 6.** (Cf. [17, Remark 3]) Let $C$ be a non-degenerate linear $(n, k, d, r, \delta)$-LRC over $\mathbb{F}_q$ with maximal local rank $t$. Then,

\[
k \leq \min_{t\in\mathbb{Z}_{\geq 2}} \left( tf + k^{(q)}_{\text{opt}}(n - t(r + \delta - 1), d) \right),
\]

where $k^{(q)}_{\text{opt}}(n, d)$ is the maximum rank of a linear code of length $n$ and minimum distance $d$ over $\mathbb{F}_q$.

However, the determination of $k^{(q)}_{\text{opt}}$ is a classical open problem in coding theory. Moreover, even given a formula for $k^{(q)}_{\text{opt}}$, evaluating (4) may be a tedious task. In that sense, the bound (2) is more explicit than this one.

Comparing the bounds (2) and (4) represents a challenge since (2) is a bound on $d$ with, on the right-hand side, a ceiling function on $k$, while (4) is a bound on $k$ with a minimum over another unknown bound that includes $d$. One method is to transform (2) to have every term in $k$ on the left, so it will be bounded by the remaining terms on the right hand side and also by the left hand side after replacing $k$ by the maximum dimension given by (4). The best bound will be the one that gives the minimum of the two alternative right hand sides.

This allows for a partial but computable comparison of (2) and (4).

We estimated $k^{(q)}_{\text{opt}}$ in (4) via the Plotkin bound and took the approach described above and it turned out, for the values we tried, the extension of the Cadambe–Mazumdar bound (4) is better or equal than (2), i.e., gives the afore-mentioned minimum. However this is only a glimpse of the relation between the two bounds and the complete comparison is left for future work. The bound (2) improves the known Singleton-type bound for binary LRCs and highlights explicitly constraints on the minimum distance $d$. In conclusion, this work takes the first step toward an explicit bound for binary LRCs via matroidal techniques, and further improvements can be obtained by extending the techniques developed in this paper, via a more detailed (and technical) study of the cyclic flats. This is left to an extended version of this article.

**Acknowledgment.** This work is supported in part by the Academy of Finland (grants #276031, #282938 and #303819 to C. Hollanti), and by the TU Munich – Institute for Advanced Study, funded by the German Excellence Initiative and the EU 7th Framework Programme (grant #291763).

**References**


