# Bounds and Constructions of Codes with Multiple Localities

Alexander Zeh and Eitan Yaakobi
Computer Science Department
Technion—Israel Institute of Technology
alex@codingtheory.eu, yaakobi@cs.technion.ac.il

***Abstract*—This paper studies bounds and constructions of locally repairable codes (LRCs) with multiple localities so-called multiple-locality LRCs (ML-LRCs). In the simplest case of two localities some code symbols of an ML-LRC have a certain locality while the remaining code symbols have another one.**

**We extend two bounds, the Singleton and the alphabet-dependent upper bound on the dimension of Cadambe–Mazumdar for LRCs, to the case of ML-LRCs with more than two localities. Furthermore, we construct Singleton-optimal ML-LRCs codes.**

## I. Introduction

Locally repairable codes (LRCs) can recover data from erasure(s) by accessing a small number $r$ of erasure-free code symbols and therefore increase the efficiency of the repair-process in large-scale distributed storage systems. For a classical LRC, all $n$ code symbols depend on at most $r$ other symbols. Basic properties and bounds of LRCs were identified by Gopalan *et al.* [2], Oggier and Datta [9] and Papailiopoulos and Dimakis [10]. The majority of the constructions of LRCs requires a large field size (see, e.g., [4, 15, 12]). Tamo and Barg gave a family of optimal LRCs for which the required field size is slightly larger than the code length in [14]. Cadambe and Mazumdar [1] gave an upper bound on the dimension of an $r$-local code which takes the field size into account. LRCs with small alphabet size were constructed in [3, 5, 16, 13].

We generalize the Singleton bound of Gopalan *et al.* [2] as well as the alphabet-dependent Cadambe–Mazumdar bound [1] to so-called multiple-locality LRCs (ML-LRCs), i.e., LRCs where the *locality* among the code symbols can be different. There exist several practical scenarios, where different localities are relevant, e.g., a distributed storage system, where a part of the stored data is accessed more frequently and therefore requires a smaller locality. We show that shortening a given Singleton-optimal LRC (i.e., an LRC that achieves the corresponding bound with equality and where all $n$ symbols have the same locality), a Singleton-optimal ML-LRC is obtained. For the case of two localities, we give an explicit algorithm that returns a parity-check matrix of an optimal linear code.

This paper is structured as follows. In Section II, we introduce notation and recall the existing Singleton and alphabet-dependent bound of Cadambe–Mazumdar for LRCs (Thm. 2 and Thm. 3). Section III provides the definition of ML-LRCs and the Singleton bound for the case of $s = 2$ different localities (Thm. 6). Thm. 8 is the Singleton bound for the general case of linear ML-LRCs with $s \geq 2$ different localities. An explicit construction (Algorithm 1) of a Singleton-optimal ML-LRC with two different localities is proven in Thm. 13. Similar to Section III, we start the description of the extended alphabet-dependent bound (for non-linear) LRCs in Thm. 18 in Section IV for two different localities. Our proof requires linearity, but it is straightforward to extend the proof-idea to the non-linear case similar to the graph approach as in Tamo–Barg [14, Proof of Thm. A.1]. We conclude the paper in Section V.

This work has been supported by German Research Council (Deutsche Forschungsgemeinschaft, DFG) under grant Ze1016/1-1 and in part by the Israel Science Foundation (ISF) Grant No. 1624/14.

## II. Preliminaries

For two integers $a, b$ with $a < b$, let $[a, b]$ denote the set of integers $\{a, a+1, \ldots, b\}$ and let $[b]$ be the shorthand notation for $[1, b]$. Let $\mathbb{F}_q$ denote the finite field of order $q$, where $q$ is a prime power. A linear $[n, k, d]_q$ code of length $n$, dimension $k$ and minimum Hamming distance $d$ over $\mathbb{F}_q$ is denoted by a calligraphic letter like $\mathcal{C}$. The parameters of a non-linear code $\mathcal{C}$ of length $n$, over an alphabet of size $q$ with dimension $k = \log|\mathcal{C}|/\log q$ are identified by $(n, k, d)_q$. For a set $\mathcal{I} \subseteq [n]$ and a matrix $\mathbf{G} \in \mathbb{F}_q^{k \times n}$, let $\mathbf{G}_{\mathcal{I}}$ denote the submatrix of $\mathbf{G}$ that consists of the columns indexed by $\mathcal{I}$. Denote by $\mathbf{I}_n$ the $n \times n$ identity matrix.

**Definition 1** (Locally Repairable Code (LRC))**.** *A linear $[n, k, d]_q$ code $\mathcal{C}$ is $r$-local if for every $i \in [n]$ there exists a subset $\mathcal{R}_i \subset [n] \setminus \{i\}, |\mathcal{R}_i| \leq r$ such that for every codeword $(c_1 \ c_2 \ \ldots \ c_n) \in \mathcal{C}$, $c_i = \sum_{j \in \mathcal{R}_i} \lambda_{i,j} c_j$, where $\lambda_{i,j} \in \mathbb{F}_q, \forall j \in \mathcal{R}_i$.*

The following generalization of the Singleton bound for LRCs was among others proven in [6, Thm. 3.1], [14, Construction 8 and Thm. 5.4] and [11, Thm. 2].

**Theorem 2** (Generalized Singleton Bound)**.** *The minimum Hamming distance $d$ of an $[n, k, d]_q$ $r$-local code $\mathcal{C}$ is bounded from above by $d \leq n - k + 2 - \lceil k/r \rceil$.*

Throughout this paper we call an $r$-local code *Singleton-optimal* if its minimum Hamming distance achieves the bound in Thm. 2 with equality. For $r \geq k$ Thm. 2 coincides with the classical (locality-unaware) Singleton bound and then a Singleton-optimal code is called maximum distance separable (MDS). In addition to the generalization of the bound in

978-1-5090-1806-2/16/$31.00 ©2016 IEEE

Thm. 2, we extend the bound on the dimension of an LRC given by Cadambe and Mazumdar [1, Thm. 1], which takes the alphabet into account.

**Theorem 3** (Cadambe–Mazumdar Bound). *The dimension $k$ of an $r$-local code $\mathcal{C}$ of length $n$ and minimum Hamming distance $d$ as in Definition 1 is upper bounded by*

$$k \leq \min_{t\in[t^\star]} k_t, \quad (1)$$

*where $k_t \overset{\text{def}}{=} tr + k_{opt}^{(q)}(n - t(r+1), d)$ and $k_{opt}^{(q)}(n,d)$ is the largest possible dimension of a code of length $n$, for a given alphabet size $q$ and a given minimum Hamming distance $d$ and $t^\star = \min(\lceil n/(r+1)\rceil, \lceil k/r\rceil)$.*

An $r$-local code is called *alphabet-optimal* if its dimension meets the bound in Thm. 3 with equality.

## III. Multiple-Locality LRCs

In this section we first define ML-LRCs with two localities and provide a Singleton-like bound. We generalize the definition to ML-LRCs with more than two localities and extend the bound. For a linear code of length $n$, we consider in a first step the case where all code symbols in $\mathcal{T}_1 \subset [n]$ have locality $r_1$ and the remaining code symbols in $\mathcal{T}_2 = [n] \setminus \mathcal{T}_1$ have locality $r_2$.

**Definition 4** (ML-LRC with Two Localities). *Let $\mathcal{T}_1 \subset [n]$ and $\mathcal{T}_2 = [n] \setminus \mathcal{T}_1$ be two distinct sets with $|\mathcal{T}_i| = n_i$, for $i = 1, 2$. Let two integers $r_1, r_2$ with $r_1 < r_2$ be given. A linear $[n,k,d]_q$ code is $((n_1,r_1),(n_2,r_2))$-local if for every $\iota \in \mathcal{T}_i$ there exists a subset $\mathcal{R}_\iota^{\langle i\rangle} \subset \mathcal{T}_i \setminus \{\iota\}, |\mathcal{R}_\iota^{\langle i\rangle}| \leq r_i$ such that for every codeword $(c_1\ c_2\ \dots\ c_n) \in \mathcal{C}$, $c_\iota = \sum_{j\in\mathcal{R}_\iota^{\langle i\rangle}} \lambda_{\iota,j}^{\langle i\rangle} c_j$, where $\lambda_{\iota,j}^{\langle i\rangle} \in \mathbb{F}_q, \forall j \in \mathcal{R}_\iota^{\langle i\rangle}, i = 1,2$.*

Note, that a code symbol belongs to $\mathcal{T}_1$ if it is a linear combination of at most $r_1$ other symbols. Clearly, an $((n_1,r_1),(n_2,r_2))$-local code is also an $r_2$-local code and an $r_1$-local code is an $((n_1,r_1),(n_2,r_2))$-local code as well.

### A. Singleton Bound

The following lemma is needed to prove the Singleton-like bound on the minimum Hamming distance of an $((n_1,r_1),(n_2,r_2))$-local code.

**Lemma 5** (Rank of Generator Matrix). *Let $\mathcal{I} \subseteq [n]$. All $k \times |\mathcal{I}|$ submatrices $\mathbf{G}_\mathcal{I}$ of a generator matrix $\mathbf{G}$ of an $[n,k,d]_q$ code $\mathcal{C}$ of rank smaller than $k$ must have at most $n-d$ columns.*

*Proof.* W.l.o.g. we assume $\mathcal{I} = [|\mathcal{I}|]$ and let $t = \operatorname{rank}\mathbf{G}_\mathcal{I}$. Clearly, $|\mathcal{I}| \geq t$. The generator matrix can be transformed to the following form:

$$\mathbf{G} = \begin{pmatrix} \mathbf{I}_t & \mathbf{A}_1 & \mathbf{A}_2 \\ & \mathbf{0} & \mathbf{A}_3 \end{pmatrix},$$

where $\mathbf{A}_1 \in \mathbb{F}_q^{t\times(|\mathcal{I}|-t)}$, $\mathbf{A}_2 \in \mathbb{F}_q^{t\times(n-|\mathcal{I}|)}$, and $\mathbf{A}_3 \in \mathbb{F}_q^{(k-t)\times(n-|\mathcal{I}|)}$. The $(t+1)$th row of $\mathbf{G}$ (first $|\mathcal{I}|$ positions are zero) has Hamming weight at least $d$ (because it is a codeword of $\mathcal{C}$) and therefore $|\mathcal{I}| \leq n-d$. □

**Theorem 6** (Singleton Bound for ML-LRCs with Two Localities). *Let the parameters as in Definition 4 be given and assume that $r_1\lceil n_1/(r_1+1)\rceil < k-1$. Then, the minimum Hamming distance of an $[n,k,d]_q$ $((n_1,r_1),(n_2,r_2))$-local code $\mathcal{C}$ is upper bounded by*

$$d \leq n - k + 2 - \left\lceil \frac{n_1}{r_1+1} \right\rceil - \left\lceil \frac{k - r_1 \left\lceil \frac{n_1}{r_1+1} \right\rceil}{r_2} \right\rceil. \quad (2)$$

*Proof.* The proof follows the idea of the proof of the Singleton bound of an $r$-local code as in [7]. Let $\mathbf{G}$ be a $k \times n$ generator matrix of $\mathcal{C}$.

1) Choose $\kappa_1 \overset{\text{def}}{=} \lceil n_1/(r_1+1)\rceil$ columns of $\mathbf{G}$ indexed by $j_1^{\langle 1\rangle}, j_2^{\langle 1\rangle}, \dots, j_{\kappa_1}^{\langle 1\rangle}$, where $j_i^{\langle 1\rangle} \in \mathcal{T}_1$. Each column is a linear combination of at most $r_1$ other columns.
2) Let $\mathcal{I}^{\langle 1\rangle}$ be the set of indexes of all repair columns of the columns indexed by $j_1^{\langle 1\rangle}, \dots, j_{\kappa_1}^{\langle 1\rangle}$, but without the indexes themselves. We have $|\mathcal{I}^{\langle 1\rangle}| \leq r_1 \cdot \kappa_1 < k-1$.
3) Choose
$$\kappa_2 \overset{\text{def}}{=} \left\lfloor \frac{k - 1 - r_1\left\lceil \frac{n_1}{r_1+1}\right\rceil}{r_2} \right\rfloor$$
columns of $\mathbf{G}$ indexed by $j_1^{\langle 2\rangle}, j_2^{\langle 2\rangle}, \dots, j_{\kappa_2}^{\langle 2\rangle}$, where $j_i^{\langle 2\rangle} \in \mathcal{T}_2$. Now, each of these columns is a linear combination of at most $r_2$ other columns. Let $\mathcal{I}^{\langle 2\rangle}$ be the set of indexes of all repair columns of the columns indexed by $j_1^{\langle 2\rangle}, j_2^{\langle 2\rangle}, \dots, j_{\kappa_2}^{\langle 2\rangle}$, but without the indexes themselves.
4) Let $\mathcal{I} \overset{\text{def}}{=} \mathcal{I}^{\langle 1\rangle} \cup \mathcal{I}^{\langle 2\rangle}$. Then $|\mathcal{I}| < k$ and we have $\operatorname{rank}(\mathbf{G}_\mathcal{I}) < k$.
5) Enlarge $\mathcal{I}$ to a set $\mathcal{I}'$, such that $\operatorname{rank}(\mathbf{G}_{\mathcal{I}'}) = k-1$, but without using $\left\{j_1^{\langle i\rangle}, j_2^{\langle i\rangle}, \dots, j_{\kappa_i}^{\langle i\rangle}\right\}_{i=1,2}$.
6) Then, define
$$\mathcal{U} \overset{\text{def}}{=} \left\{\mathcal{I}' \cup \left\{j_1^{\langle 1\rangle}, j_2^{\langle 1\rangle}, \dots, j_{\kappa_1}^{\langle 1\rangle}, j_1^{\langle 2\rangle}, j_2^{\langle 2\rangle}, \dots, j_{\kappa_2}^{\langle 2\rangle}\right\}\right\},$$
where $|\mathcal{U}| \geq k-1+\kappa_1+\kappa_2$. We have $\operatorname{rank}(\mathbf{G}_\mathcal{U}) \leq k-1$.

Using Lemma 5, we know that $|\mathcal{U}|$ can be upper bounded by $k-1+\kappa_1+\kappa_2 \leq n-d$ and therefore we obtain the following bound on the minimum Hamming distance:

$$\begin{aligned} d &\leq n-k+1 - \left\lceil \frac{n_1}{r_1+1} \right\rceil - \left\lfloor \frac{k-1-r_1\left\lceil \frac{n_1}{r_1+1}\right\rceil}{r_2} \right\rfloor \\ &= n-k+2 - \left\lceil \frac{n_1}{r_1+1} \right\rceil - \left\lceil \frac{k-r_1\left\lceil \frac{n_1}{r_1+1}\right\rceil}{r_2} \right\rceil. \end{aligned}$$

□

In case $r_1\lceil n_1/(r_1+1)\rceil \geq k-1$, set $\kappa_1 = \lfloor (k-1)/r_1 \rfloor$ in Step 1) and $\kappa_2 = 0$ in Step 2). Then, we will obtain the Singleton bound of an $r_1$-local code. For $r_1 = r_2 = r$, we obtain from (2) the Singleton bound of an $r$-local LRC as in Thm. 2.

We extend Definition 4 to ML-LRCs with $s \geq 2$ localities and generalize the Singleton-like bound on the minimum Hamming distance from Thm. 6.

**Definition 7** (ML-LRC)**.** *Let $s$ integers $r_1, r_2, \dots, r_s$ with $r_1 < r_2 < \dots < r_s$ be given. Denote by $\mathcal{T}_i \subset [n]$ for all $i \in [s]$ pairwise disjoint sets, where $\cup_{i\in[s]}\mathcal{T}_i = [n]$. A linear $[n, k, d]_q$ code is $((n_1, r_1), (n_2, r_2), \dots, (n_s, r_s))$-local if for every $\iota \in \mathcal{T}_i$ there exist a subset $\mathcal{R}_\iota^{\langle i\rangle} \subset \mathcal{T}_i \setminus \{\iota\}$ such that for every codeword $(c_1\ c_2\ \dots\ c_n) \in \mathcal{C}$, $c_\iota = \sum_{j\in\mathcal{R}_\iota^{\langle i\rangle}} \lambda_{\iota,j}^{\langle i\rangle} c_j$, where $\lambda_{\iota,j}^{\langle i\rangle} \in \mathbb{F}_q, \forall j \in \mathcal{R}_\iota^{\langle i\rangle}, i \in [s]$.*

Now, we give a Singleton-like bound for an $((n_1, r_1), (n_2, r_2), \dots, (n_s, r_s))$-local code as in Definition 7.

**Theorem 8** (Singleton Bound for ML-LRCs)**.** *Let $\mathcal{C}$ be an $[n, k, d]_q$ $((n_1, r_1), (n_2, r_2), \dots, (n_s, r_s))$-local code as in Definition 7 and assume that $\sum_{i\in[s-1]} r_i \lceil n_i/(r_i+1)\rceil < k-1$. Then, the minimum Hamming distance of $\mathcal{C}$ is upper bounded by*

$$d \leq n\text{-}k + 2\text{-} \sum_{i\in[s-1]} \left\lceil \frac{n_i}{r_i+1} \right\rceil \text{-} \left\lceil \frac{k\text{-} \sum\limits_{i\in[s-1]} r_i \left\lceil \frac{n_i}{r_i+1} \right\rceil}{r_s} \right\rceil . \quad (3)$$

In case $\sum_{i\in[j]} r_i \lceil n_i/(r_i + 1)\rceil \geq k-1$, the bound becomes the bound for an $((n_1, r_1), (n_2, r_2), \dots, (\sum_{i\in[j,s]} n_i, r_j))$-local code. Clearly, an $((n_1, r_1), (n_2, r_2), \dots, (\sum_{i\in[j,s]} n_i, r_j))$-local code is also an $((n_1, r_1), (n_2, r_2), \dots, (n_s, r_s))$-local code, because $r_1 < \dots < r_j < \dots < r_s$. We call an $((n_1, r_1), (n_2, r_2), \dots, (n_s, r_s))$-local ML-LRC with minimum Hamming distance that fulfills (3) with equality *Singleton-optimal*.

### *B. Shortening and Constructions*

In this subsection, we first show that an $((n_1, r_1), (n_2, r_2))$-local Singleton-optimal ML-LRC can be obtained through shortening an $r_2$-local Singleton-optimal LRC. Then, we analyze the shortening of an $((n_1, r_1), (n_2, r_2), \dots, (n_s, r_s))$-local Singleton-optimal ML-LRC. We give an explicit construction of an $((n_1, r_1), (n_2, r_2))$-local Singleton-optimal code for the ease of notation. The construction can be easily extended to $s \geq 2$ localities.

We consider the case of shortening the $i$th information symbol (see, e.g., [8, Ch. 18 §9]).

**Definition 9** (Shortening)**.** *Let $\mathcal{C}$ be an $[n, k, d]_q$ code with generator matrix in systematic form, i.e., $\mathbf{G} = (\mathbf{I}_k\ \mathbf{A})$. A parity-check matrix is then $\mathbf{H} = (-\mathbf{A}^T\ \mathbf{I}_{n-k})$. The shortened code is defined as*

$$\mathcal{C}^{\langle i\rangle} \overset{\text{def}}{=} \{(c_1 \dots c_{i\text{-}1} c_{i+1} \dots c_n) : (c_1 \dots c_{i\text{-}1} 0 c_{i+1} \dots c_n) \in \mathcal{C}\}.$$

*For $i \in [k]$, the generator matrix of $\mathcal{C}^{\langle i\rangle}$ is obtained through deleting the $i$th column of $\mathbf{I}_k$ (and the corresponding $i$th row) of $\mathbf{G}$. A parity-check of $\mathcal{C}^{\langle i\rangle}$ is obtained by deleting the $i$th column in $\mathbf{A}^T$ of $\mathbf{H}$. The shortened code is an $[n-1, k-1, \geq d]_q$ code.*

Throughout this paper we refer to shortening to the case where $i \in [k]$. Clearly, if $\mathcal{C}$ is MDS, then the shortened code $\mathcal{C}^{\langle i\rangle}$ is also MDS. The following lemma shows that this holds similarly for an $r$-local Singleton-optimal LRC.

**Lemma 10** (Shortening an $r$-local LRC)**.** *Let an $[n, k, d]_q$ $r$-local Singleton-optimal LRC be given. Then, the shortened $[n - 1, k - 1, d]_q$ code is an $((r, r - 1), (n - 1 - r, r))$-local code is Singleton-optimal.*

**Example 11** (Shortened Singleton-optimal LRC)**.** *We shorten the $[12, 6, 6]_{13}$ 3-local Singleton-optimal code of Example 2 in [14] by one position. We obtain an $[11, 5, 6]_{13}$ $((3, 2), (8, 3))$-local code and Thm. 6 gives $d \leq 11 - 5 + 2 - \lceil 3/(2 + 1)\rceil - \lceil(5 - 2)/3\rceil = 6$.*

**Lemma 12** (Shortening an ML-LRC)**.** *Let $\mathcal{C}$ be an $[n, k, d]_q$ $((n_1, r_1), (n_2, r_2), \dots, (n_s, r_s))$-local Singleton-optimal code. Let $\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_s$ denote the locality sets and let $\alpha \in [s]$. Shortening $\mathcal{C}$ by a coordinate that is contained in $\mathcal{T}_\alpha$ gives if $r_{\alpha-1} = r_\alpha - 1$ an $[n' = n - 1, k' = k - 1, d]_q$ $((n'_1, r'_1), (n'_2, r'_2), \dots, (n'_s, r'_s))$-local code $\mathcal{C}'$ with*

$$\begin{aligned} (n'_i, r'_i) &= (n_i, r_i), \qquad \forall i \in [s] \setminus \{\alpha - 1, \alpha\}, \\ (n'_{\alpha-1}, r'_{\alpha-1}) &= (n_{\alpha-1} + r_\alpha, r_{\alpha-1}), \\ (n'_\alpha, r'_\alpha) &= (n_\alpha - r_\alpha - 1, r_\alpha), \end{aligned} \quad (4)$$

*else (i.e., $r_{\alpha-1} \neq r_\alpha-1$) shortening gives an $[n' = n-1, k' = k - 1, d]_q$ $((n'_1, r'_1), (n'_2, r'_2), \dots, (n'_\iota, r'_\iota), \dots, (n'_s, r'_s))$-local code $\mathcal{C}'$ with*

$$\begin{aligned} (n'_i, r'_i) &= (n_i, r_i), \qquad \forall i \in [s] \setminus \{\alpha\}, \\ (n'_\iota, r'_\iota) &= (r_\alpha, r_\alpha - 1), \\ (n'_\alpha, r'_\alpha) &= (n_\alpha - r_\alpha - 1, r_\alpha). \end{aligned} \quad (5)$$

*Then, the shortened $[n' = n - 1, k' = k - 1, d]_q$ $((n'_1, r'_1), (n'_2, r'_2), \dots, (n'_s, r'_s))$-local respectively the $((n'_1, r'_1), (n'_2, r'_2), \dots, (n'_\iota, r'_\iota), \dots, (n'_s, r'_s))$-local ML-LRC is Singleton-optimal.*

---

**Algo 1:** Singleton-optimal ML-LRC with $r_2 > r_1$.

---

**Input**:
Parity-check matrix $\mathbf{H}$ of an $[n_2 + (r_2 + 1)\frac{n_1}{r_1+1}, k + (r_2 - r_1)\frac{n_1}{r_1+1}, d]_q$ $r_2$-local Singleton-optimal code with $\rho = \frac{n_2}{r_2+1} + \frac{n_1}{r_1+1}$ repair sets $\mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_\rho$.

1 **for** $i \in [\frac{n_1}{r_1+1}]$ **do**
2 Delete $r_2 - r_1$ columns of $\mathbf{H}$ that are contained in $\mathcal{R}_i$.

**Output**: Parity-check matrix $\mathbf{H}$ of an $[n_1 + n_2, k, d]_q$ $((n_1, r_1), (n_2, r_2))$-local code.

---

Algorithm 1 provides a parity-check matrix of a Singleton-optimal $[n_1 + n_2, k, d]_q$ $((n_1, r_1), (n_2, r_2))$-local code. We assume that $(r_1 + 1) \mid n_1$, $(r_2 + 1) \mid n_2$, and that an $r_2$-local Singleton-optimal code is given.

**Theorem 13** (Singleton-optimal ML-LRC). *Algorithm 1 returns a parity-check matrix of a Singleton-optimal* $[n_1 + n_2, k, d]_q$ $((n_1, r_1), (n_2, r_2))$*-local code that covers the possible rate-regime.*

*Proof.* The minimum Hamming distance of the given $[n', k', d']_q$ $r_2$-local Singleton-optimal LRC, where

$$n' = n_2 + (r_2+1)\frac{n_1}{r_1+1}, \tag{6}$$

$$k' = k + (r_2 - r_1)\frac{n_1}{r_1+1}, \tag{7}$$

equals

$$d' = n' - k' + 2 - \left\lceil \frac{k'}{r_2} \right\rceil. \tag{8}$$

Inserting the expression of the length and the dimension as in (6) respectively (7) into (8) leads to:

$$\begin{aligned} d' &\leq n_2 + (r_2+1)\frac{n_1}{r_1+1} - k - (r_2 - r_1)\frac{n_1}{r_1+1} + 2 \\ &\quad - \left\lceil \frac{k + (r_2-r_1)\frac{n_1}{r_1+1}}{r_2} \right\rceil \\ &= n_1 + n_2 - k + 2 - \frac{n_1}{r_1+1} - \left\lceil \frac{k - r_1\frac{n_1}{r_1+1}}{r_2} \right\rceil, \end{aligned} \tag{9}$$

which is the minimum Hamming distance of an $[n_1+n_2, k, d]_q$ $((n_1, r_1), (n_2, r_2))$-local code that is optimal with respect to the bound given in Thm. 6.

To prove the achievable rate-regime, we recall that the original $r_2$-local LRC has to satisfy:

$$k' \leq \frac{r_2}{r_2+1}n', \tag{10}$$

and inserting the expressions of (6) and (7) in (10) gives

$$\begin{aligned} k + (r_2 - r_1)\frac{n_1}{r_1+1} &\leq \frac{r_2}{r_2+1}n_2 + (r_2+1)\frac{n_1}{r_1+1} \\ \Leftrightarrow k &\leq \frac{r_1}{r_1+1}n_1 + \frac{r_2}{r_2+1}n_2, \end{aligned}$$

which is clearly the rate-restriction for an $((n_1, r_1), (n_2, r_2))$-local code. □

## IV. Alphabet-Dependent Bound for ML-LRCs

### A. Bound and Shortening

For a given $(n, k, d)_q$ code $\mathcal{C}$ and a subset $\mathcal{I} \subseteq [n]$ define as in [1]

$$H(\mathcal{I}) \overset{\text{def}}{=} \frac{\log |\{\mathbf{x}_{\mathcal{I}} : \mathbf{x} \in \mathcal{C}\}|}{\log q}. \tag{11}$$

The bound given in Thm. 3 follows from the following two lemmas.

**Lemma 14** ([1, Lemma 1]). *For a given* $(n, k, d)_q$ *r-local code* $\mathcal{C}$*, an integer* $t \in [\lceil k/r \rceil]$*, a set* $\mathcal{I} \subseteq [n]$ *with* $|\mathcal{I}| = t(r+1)$ *and* $H(\mathcal{I}) \leq tr$ *as defined in* (11) *exists (and is constructed explicitly).*

**Lemma 15** ([1, Lemma 2]). *For an* $(n, k, d)_q$ *code* $\mathcal{C}$ *with* $\mathcal{I} \subseteq [n]$ *and* $H(\mathcal{I}) \leq m$ *an* $(n - |\mathcal{I}|, k - m, d)_q$ *code exists.*

In the following lemma, we refine Lemma 14 slightly and generalize it to an $((n_1, r_1), (n_2, r_2), \dots, (n_s, r_s))$-local ML-LRC.

**Lemma 16.** *Let* $\mathcal{T}_i$ *with* $|\mathcal{T}_i| = n_i$ *for all* $i \in [s]$ *be the locality sets of a given* $(n, k, d)_q$ $((n_1, r_1), (n_2, r_2), \dots, (n_s, r_s))$*-local code* $\mathcal{C}$ *as in Definition 7. Then, there exist s sets* $\mathcal{I}_i \subseteq \mathcal{T}_i$ *for all* $i \in [s]$ *with*

$$|\mathcal{I}_i| = \begin{cases} t_i(r_i+1) & \textit{for } t_i \leq n_i/(r_i+1), \textit{ if } (r_i+1) \mid n_i \\ n_i & \textit{for } t_i \leq \lceil n_i/(r_i+1) \rceil, \textit{ if } (r_i+1) \nmid n_i, \end{cases}$$

*and* $H(\mathcal{I}_i) \leq t_i r_i$ *(for both cases) and for all* $i \in [s]$.

*Proof.* Similar to [1, Proof of Lemma 1], we construct the set $\mathcal{I}_i$ explicitly. Let $\mathcal{R}_i^{\langle j \rangle}$ denote the repair set of the $j$-th coordinate that belongs to $\mathcal{T}_i$. Clearly $|\mathcal{R}_i^{\langle j \rangle}| \leq r_i$. Algorithm 2 constructs the set $\mathcal{I}_i$.

**Algo 2:** Construction of $\mathcal{I}_i \subseteq \mathcal{T}_i$.

**Input**: Set $\mathcal{T}_i$ with $|\mathcal{T}_i| = n_i$ and locality $r_i$.
Integer $t_i \in [\lceil n_i/(r_i+1) \rceil]$.

1 Select $a_0$ arbitrarily from $\mathcal{T}_i$.
2 $\mathcal{S}^{\langle 0 \rangle} \leftarrow \emptyset$.
3 **for** $m \in [t_i - 1]$ **do**
4 Select $a_m$ in $\mathcal{T}_i \setminus \left( \bigcup_{l=0}^{m-1} \{a_m \cup \mathcal{R}_i^{\langle a_l \rangle} \cup \mathcal{S}^{\langle l \rangle} \right)$.
5 $\mathcal{I}_i^{\langle m \rangle} \leftarrow \bigcup_{l=0}^{m-1} \{a_l\} \cup \mathcal{R}_i^{\langle a_l \rangle} \cup \mathcal{S}^{\langle l \rangle}$.
6 Select $\mathcal{S}^{\langle m \rangle} \subseteq \mathcal{T}_i \setminus \left( \{a_m\} \cup \mathcal{R}_i^{\langle a_m \rangle} \cup \mathcal{I}_i^{\langle m \rangle} \right)$ with
7 $|\mathcal{S}^{\langle m \rangle}| = \min((m+1)(r_i+1), n_i) - |\{a_m\} \cup \mathcal{R}_i^{\langle a_m \rangle} \cup \mathcal{I}_i^{\langle m \rangle}|$
8 $\mathcal{I}_i \leftarrow \bigcup_{m=0}^{t_i-1} \{a_m\} \cup \mathcal{R}_i^{\langle a_m \rangle} \cup \mathcal{S}^{\langle m \rangle}$.

**Output**: Return the set $\mathcal{I}_i$.

Algorithm 2 differs from the algorithm used in [1, Proof of Lemma 1] only in Line 7, which ensures that the constructed set cannot have a cardinality greater than $n_i$. In the proof of [1, Lemma 1], it is shown that $H(\mathcal{I}_i) = H(\mathcal{I}_i \setminus \{a_0, a_1, \dots, a_{t_i-1}\})$. Clearly, for:
Case a): If $t_i \leq \lfloor n_i/(r_i+1) \rfloor$, then $|\mathcal{I}_i| = t_i(r_i+1)$ and $H(\mathcal{I}_i) \leq t_i r_i$.
Case b): If $t_i = \lceil n_i/(r_i+1) \rceil$ and $(r_i+1) \nmid n_i$, then $|\mathcal{I}_i| = n_i$ and

$$H(\mathcal{I}_i) \leq n_i - \lceil n_i/(r_i+1) \rceil = \left\lceil \frac{r_i n_i}{r_i+1} \right\rceil. \tag{12}$$

□

The distinction between Cases a) and b) in Lemma 16 is not relevant for the bound of an $r$-local code, but becomes necessary if we want to bound the dimension of an $((n_1, r_1), (n_2, r_2), \dots, (n_s, r_s))$-local ML-LRC.

**Lemma 17.** *For an* $(n, k, d)_q$ *code* $\mathcal{C}$ *with s pairwise disjoint sets* $\mathcal{I}_i \subseteq [n]$ *with* $H(\mathcal{I}_i) \leq m_i$ *for all* $i \in [s]$*, there exists an* $(n - \sum_{i \in [s]} |\mathcal{I}_i|, k - \sum_{i \in [s]} m_i, d)$ *code.*

*Proof.* Apply Lemma 15 consecutively $s$ times and the statement follows. □

From Lemma 16 and Lemma 17 the following theorem follows.

**Theorem 18** (Alphabet-Dependent Bound for ML-LRCs with Two Different Localities)**.** *Let $\mathcal{C}$ be an $(n,k,d)_q$ $((n_1,r_1),(n_2,r_2))$-local code. Define*

$$k_{t_1,t_2} \stackrel{\text{def}}{=} t_1r_1+t_2r_2+ k_{opt}^{(q)}\left(n-\min(n_1,t_1(r_1+1))-\min(n_2,t_2(r_2+1)),d\right). \quad (13)$$

*The dimension of $\mathcal{C}$ is bounded from above by*

$$k \leq \min_{\substack{t_1\in[t_1^\star],\\ t_2\in[t_2^\star]}} k_{t_1,t_2}, \quad (14)$$

*where*

$$t_1^\star \stackrel{\text{def}}{=} \left\lceil \frac{n_1}{r_1+1} \right\rceil, \quad (15)$$

$$t_2^\star \stackrel{\text{def}}{=} \left\lfloor \frac{k-1-t_1r_1}{r_2} \right\rfloor, \quad \forall t_1 \in [t_1^\star]. \quad (16)$$

*Proof.* The expressions as in (13) and (14) follow from Lemma 16 and Lemma 17 for $s=2$.

The value of $t_1$ can be bounded by $\min(\lceil n_1/(r_1+1)\rceil, \lfloor (k-1)/r_1 \rfloor)$ similar as in the case of an $r$-local LRC. We assume that $r_1\lceil n_1/(r_1+1)\rceil < k-1$ The maximal value for $t_2$ follows from the fact that for $t_2 > t_2^*$ the expression in (14) is at least $k$. Clearly, $t_2 \leq \lceil n_2/(r_2+1)\rceil$. It is well-known that the rate of an $r$-local code is at most $r/(r+1)$ and due to the fact that an $((n_1,r_1),(n_2,r_2))$-local code is also an $r_2$-local code, we have $\left\lfloor \frac{k-1-r_1t_1}{r_2} \right\rfloor \leq \left\lceil \frac{n_2}{r_2+1} \right\rceil$ and therefore the maximal $t_2^\star$ as in (16) follows. □

The following theorem generalizes Thm. 18 to the case of $s>2$ different localities.

**Theorem 19** (Alphabet-Dependent Bound for ML-LRCs)**.** *Let $\mathcal{C}$ be an $(n,k,d)_q$ $((n_1,r_1),(n_2,r_2),\dots,(n_s,r_s))$-local code as in Definition 7. Furthermore, assume $\sum_{i\in[s]} r_i\lceil n_i/(r_i+1)\rceil < k-1$. Define*

$$t_i^\star \stackrel{\text{def}}{=} \left\lceil \frac{n_i}{r_i+1} \right\rceil, \quad \forall i \in [s-1], \quad (17)$$

$$t_s^\star \stackrel{\text{def}}{=} \left\lfloor \frac{k-1-\sum\limits_{i\in[s-1]} t_ir_i}{r_s} \right\rfloor, \quad \forall t_i \in [t_i^\star], \quad (18)$$

*and let*

$$k_{t_1,\dots,t_s} \stackrel{\text{def}}{=} \sum_{i\in[s]} t_ir_i + k_{opt}^{(q)}\left(n-\sum_{i\in[s]}\min(n_i,t_i(r_i+1)),d\right). \quad (19)$$

*Then the dimension of $\mathcal{C}$ is upper bounded by*

$$k \leq \min_{\substack{t_i\in[t_i^\star]\\ \forall i\in[s]}} k_{t_1,\dots,t_s}. \quad (20)$$

If there exist $s$ parameters $t_i \in [t_i^\star], \forall i \in [s]$ such that the dimension $k$ of an $[n,k,d]_q$ $((n_1,r_1),(n_2,r_2),\dots,(n_s,r_s))$-local ML-LRC equals $k_{t_1,t_2,\dots,t_s}$ as in (19), then we call the ML-LRC *alphabet-optimal*.

## V. Conclusion

We introduced the class of multiple-locality LRCs. The Singleton-like upper bound on the minimum Hamming distance and the alphabet-dependent bound of Cadambe–Mazumdar on the dimension of an LRC were generalized to the case of ML-LRCs. Future work are (direct) constructions of ML-LRC without shortening and the adaption of existing bounds and constructions for LRC where the parameter for every code symbol is also variable (e.g., availability). Proofs can be found in an extended version of the paper on arxiv http://arxiv.org/abs/1601.02763.

The authors thank I. Tamo for fruitful discussions (especially for Thm. 6).

## References


[1] V. R. Cadambe and A. Mazumdar. "Bounds on the Size of Locally Recoverable Codes". *IEEE Transactions on Information Theory* 61.11 (Nov. 2015), pp. 5787–5794.

[2] P. Gopalan, C. Huang, H. Simitci, and S. Yekhanin. "On the Locality of Codeword Symbols". *IEEE Transactions on Information Theory* 58.11 (Nov. 2012), pp. 6925–6934.

[3] S. Goparaju and R. Calderbank. "Binary Cyclic Codes that are Locally Repairable". *IEEE International Symposium on Information Theory (ISIT).* Honolulu, USA, June 2014, pp. 676–680.

[4] C. Huang, M. Chen, and J. Li. "Pyramid Codes: Flexible Schemes to Trade Space for Access Efficiency in Reliable Data Storage Systems". *IEEE International Symposium on Network Computing and Applications (NCA).* Cambridge, USA, July 2007, pp. 79–86.

[5] P. Huang, E. Yaakobi, H. Uchikawa, and P. H. Siegel. "Cyclic Linear Binary Cocally Repairable Codes". *IEEE Information Theory Workshop (ITW).* Jerusalem, Israel, Apr. 2015, pp. 1–5.

[6] G. Kamath, N. Prakash, V. Lalitha, and P. Kumar. "Codes With Local Regeneration and Erasure Correction". *IEEE Transactions on Information Theory* 60.8 (Aug. 2014), pp. 4637–4660.

[7] M. Kuijper and D. Napp. "Erasure codes with simplex locality". *arXiv:1403.2779 [cs, math]* (Mar. 2014).

[8] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes.* North Holland Publishing Co., June 1988.

[9] F. Oggier and A. Datta. "Self-repairing Homomorphic Codes for Distributed Storage Systems". *IEEE INFOCOM.* Apr. 2011, pp. 1215–1223.

[10] D. S. Papailiopoulos and A. G. Dimakis. "Locally Repairable Codes". *IEEE Transactions on Information Theory* 60.10 (Oct. 2014), pp. 5843–5855.

[11] N. Prakash, G. M. Kamath, V. Lalitha, and P. V. Kumar. "Optimal Linear Codes with a Local-Error-Correction Property". *IEEE International Symposium on Information Theory (ISIT).* Cambridge, USA, July 2012, pp. 2776–2780.

[12] N. Silberstein, A. S. Rawat, O. O. Koyluoglu, and S. Vishwanath. "Optimal Locally Repairable Codes via Rank-Metric Codes". *IEEE International Symposium on Information Theory (ISIT).* Istanbul, Turkey, July 2013, pp. 1819–1823.

[13] N. Silberstein and A. Zeh. "Optimal Binary Locally Repairable Codes via Anticodes". *IEEE International Symposium on Information Theory (ISIT).* Hong Kong, China, June 2015, pp. 1247–1251.

[14] I. Tamo and A. Barg. "A Family of Optimal Locally Recoverable Codes". *IEEE Transactions on Information Theory* 60.8 (Aug. 2014), pp. 4661–4676.

[15] I. Tamo, D. S. Papailiopoulos, and A. G. Dimakis. "Optimal Locally Repairable Codes and Connections to Matroid Theory". *IEEE International Symposium on Information Theory Proceedings (ISIT).* Istanbul, Turkey, July 2013, pp. 1814–1818.

[16] A. Zeh and E. Yaakobi. "Optimal Linear and Cyclic Locally Repairable Codes over Small Fields". *IEEE Information Theory Workshop (ITW).* Jerusalem, Israel, Apr. 2015, pp. 1–5.