HAL open science

Bounds and Constructions of Codes with Multiple Localities

Alexander Zeh, Eitan Yaakobi

▶ To cite this version:

Alexander Zeh, Eitan Yaakobi. Bounds and Constructions of Codes with Multiple Localities. 2016. hal-01253384

HAL Id: hal-01253384

https://inria.hal.science/hal-01253384v1

Preprint submitted on 10 Jan 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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 as well as codes that achieve the extended alphabet-dependent bound. We give a family of binary ML-LRCs based on generalized code concatenation that is optimal with respect to the alphabet-dependent bound.

Index Terms

Alphabet-dependent bound, distributed storage, generalized code concatenation, locally repairable codes, Pyramid code, Singleton bound, shortening

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.* [6], Oggier and Datta [13] and Papailiopoulos and Dimakis [14]. The majority of the constructions of LRCs requires a large field size (see, e.g., [8, 19, 16]). Tamo and Barg gave a family of optimal LRCs for which the required field size is slightly larger than the code length in [18]. Cadambe and Mazumdar [3] 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 [7, 9, 20, 17].

We generalize the Singleton bound of Gopalan *et al.* [6] as well as the alphabet-dependent Cadambe–Mazumdar bound [3] 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- respectively alphabet-optimal LRC (i.e., a LRC that achieves the corresponding bound with equality and where all n symbols have the same locality), a Singleton- respectively alphabet-optimal ML-LRC is obtained. For the case of two localities, we give explicit algorithms that return a parity-check matrix of an optimal linear code for both bounds. Both algorithms can be easily extended to the general case of more than two localities. The adaption of the Pyramid code construction [8] to the case of multiple localities is outlined, because they require a (slightly) smaller field-size as the construction of [18] even though the locality-property is only guaranteed for the information symbols). Furthermore, we identify a family of binary alphabet-optimal ML-LRCs using generalized code concatenation (see e.g., [1, 21]).

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. The proof is given in the appendix. An explicit construction (Algorithm 1) of a Singleton-optimal ML-LRC with two different localities is proven in Thm. 13. Basically, Algorithm 1 shortens a given Singleton-optimal LRC code. The adaption of Pyramid codes [8] for different localities among the information symbols is outlined at the end of Section III. Similar to Section III, we start the description of the extended alphabet-dependent bound (for non-linear) LRCs in Thm. 20 in Section IV for two different localities. The general case ($s \geq 2$) is contained in Thm. 22 and the proof can be found in the appendix. 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 [18, Proof of Thm. A.1]. Thm. 25 proves an explicit construction (Algorithm 3) of an ML-LRC that is optimal with respect to new alphabet-dependent bound (for two different

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.

localities). Similar to Algorithm 1, Algorithm 3 shortens a given alphabet-optimal LRC. Both algorithms (Algorithm 1 and 3) can be adapted to the case of an ML-LRC with more than two different localities.

Section V contains a family of r-local binary LRCs based on generalized code concatenation. It is shown that this family is optimal with respect to the Cadambe–Mazumdar bound for $r = 2, 3$ (see Construction 27 and Thm. 28). We conclude the paper in Section VI.

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 with 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 all n code symbols are a linear combination of at most r other code symbols.*

The following generalization of the Singleton bound for LRCs was among others proven in [10, Thm. 3.1], [18, Construction 8 and Thm. 5.4] and [15, 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*

$$d \leq n - k + 2 - \left\lceil \frac{k}{r} \right\rceil .$$

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 Thm. 2, we extend the bound on the dimension of an LRC given by Cadambe and Mazumdar [3, 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*

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

where

$$k_t \stackrel{\text{def}}{=} tr + k_{opt}^{(q)}(n - t(r+1), d), \tag{2}$$

where $k_{opt}^{(q)}(n, d)$ 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 $t^\star = \min\left(\lceil n/(r+1) \rceil, \lceil k/r \rceil\right)$.

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. Afterwards, Singleton-optimal all-symbol ML-LRCs are constructed and the construction of Pyramid codes [8] with information-symbol locality is extended.

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 all n_i code symbols within $\mathcal{T}_i$ are a linear combination of at most r_i other code symbols for $i = 1, 2$.*

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.

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. 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*

$$d \leq n - k + 2 - \left\lceil \frac{n_1}{r_1 + 1} \right\rceil - \left\lceil \frac{k - r_1 \left\lceil \frac{n_1}{r_1+1} \right\rceil}{r_2} \right\rceil. \tag{3}$$

Proof. The proof follows the idea of the proof of the Singleton bound of an r-local code as in [11]. Let $\mathbf{G}$ be a $k \times n$ generator matrix of $\mathcal{C}$. We assume that $r_1 \lceil n_1/(r_1 + 1) \rceil < k - 1$.

1) Choose $\kappa_1 \stackrel{\text{def}}{=} \lceil n_1/(r_1 + 1) \rceil$ columns of $\mathbf{G}$ indexed by $j_1^{\langle 1 \rangle}, j_2^{\langle 1 \rangle}, \ldots, j_{\kappa_1}^{\langle 1 \rangle}$, 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 $j_1^{\langle 1 \rangle}, \ldots, j_{\kappa_1}^{\langle 1 \rangle}$, but without the indexes themselves. We have $|\mathcal{I}^{\langle 1 \rangle}| \leq r_1 \cdot \kappa_1 < k - 1$.
3) Choose
$$\kappa_2 \stackrel{\text{def}}{=} \left\lfloor \frac{k - 1 - r_1 \left\lceil \frac{n_1}{r_1+1} \right\rceil}{r_2} \right\rfloor$$
columns of $\mathbf{G}$ indexed by $j_1^{\langle 2 \rangle}, j_2^{\langle 2 \rangle}, \ldots, j_{\kappa_2}^{\langle 2 \rangle}$, 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 $j_1^{\langle 2 \rangle}, j_2^{\langle 2 \rangle}, \ldots, j_{\kappa_2}^{\langle 2 \rangle}$, but without the indexes themselves.
4) Let $\mathcal{I} \stackrel{\text{def}}{=} \mathcal{I}^{\langle 1 \rangle} \cup \mathcal{I}^{\langle 2 \rangle}$. 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 $\left\{ j_1^{\langle i \rangle}, j_2^{\langle i \rangle}, \ldots, j_{\kappa_i}^{\langle i \rangle} \right\}_{i=1,2}$.
6) Then, define
$$\mathcal{U} \stackrel{\text{def}}{=} \left\{ \mathcal{I}' \cup \left\{ j_1^{\langle 1 \rangle}, j_2^{\langle 1 \rangle}, \ldots, j_{\kappa_1}^{\langle 1 \rangle}, j_1^{\langle 2 \rangle}, j_2^{\langle 2 \rangle}, \ldots, j_{\kappa_2}^{\langle 2 \rangle} \right\} \right\},$$
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 (3) 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 $r_1, r_2, \ldots, r_s$ with $r_1 < r_2 < \cdots < r_s$ be given. Denote by $\mathcal{T}_i \subset [n]$ for all $i \in [s]$ pairwise disjoint sets. Let $\cup_{i\in[s]}\mathcal{T}_i = [n]$. A linear $[n, k, d]_q$ code is $((n_1, r_1), (n_2, r_2), \ldots, (n_s, r_s))$-local if all n_i code symbols within a set $\mathcal{T}_i$ with $|\mathcal{T}_i| = n_i$ are a linear combination of at most r_i other code symbols within $\mathcal{T}_i$ for all $i \in [s]$.*

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

Theorem 8 (Singleton Bound for ML-LRCs)**.** *Let $\mathcal{C}$ be an $[n, k, d]_q$ $((n_1, r_1), (n_2, r_2), \ldots, (n_s, r_s))$-local code as in Definition 7. Then, the minimum Hamming distance of $\mathcal{C}$ is upper bounded by*

$$d \leq n\text{-}k + 2\text{-} \sum_{i\in[s-1]} \left\lceil \frac{n_i}{r_i + 1} \right\rceil \text{-} \left\lceil \frac{k\text{-} \sum\limits_{i\in[s-1]} r_i \left\lceil \frac{n_i}{r_i+1} \right\rceil}{r_s} \right\rceil . \tag{4}$$

Proof. See appendix. □

We call an $((n_1, r_1), (n_2, r_2), \ldots, (n_s, r_s))$-local ML-LRC with minimum Hamming distance that fulfills (4) 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-local Singleton-optimal LRC. Then, we analyze the shortening of an $((n_1, r_1), (n_2, r_2), \ldots, (n_s, r_s))$-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. Furthermore, we adapt the construction of information-symbol-local Pyramid codes to the case of multiple localities.

We consider the case of shortening the ith information symbol (see, e.g., [12, 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*

$$\begin{aligned}\mathcal{C}^{\langle i\rangle} \stackrel{\text{def}}{=} \{&(c_1 \ c_2 \ \ldots \ c_{i-1} \ c_{i+1} \ \ldots \ c_n) : \\ &(c_1 \ c_2 \ \ldots \ c_{i-1} \ 0 \ c_{i+1} \ \ldots \ c_n) \in \mathcal{C}\}.\end{aligned}$$

For $i \in [k]$, the generator matrix of $\mathcal{C}^{\langle i\rangle}$ is obtained through deleting the ith column of $\mathbf{I}_k$ (and the corresponding ith row) of $\mathbf{G}$. A parity-check of $\mathcal{C}^{\langle i\rangle}$ is obtained by deleting the ith 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 shortening to the case where $i \in [k]$.

Clearly, if $\mathcal{C}$ is an MDS code, then the shortened code $\mathcal{C}^{\langle i\rangle}$ is also an MDS code. 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 and is Singleton-optimal.*

Proof. Clearly, the shortening by one symbol affects at least one repair-set and its cardinality is decreased to r. The locality of the contained code symbols is then $r-1$. We obtain from (3) with $n_1 = r$, $r_1 = r-1$ and $r_2 = r$ for the shortened code of length $n-1$ and dimension $k-1$ the following bound on the minimum Hamming distance:

$$\begin{aligned} d &\leq n\text{-}1\text{-}(k\text{-}1) + 2\text{-} \left\lceil \frac{n_1}{r_1 + 1} \right\rceil \text{-} \left\lceil \frac{(k\text{-}1)\text{-}r_1 \left\lceil \frac{n_1}{r_1+1} \right\rceil}{r_2} \right\rceil \\ &= n - k + 2 - \left\lceil \frac{r}{r} \right\rceil - \left\lceil \frac{k - 1 - (r-1) \left\lceil \frac{r}{r} \right\rceil}{r} \right\rceil \\ &= n - k + 2 - \left\lceil \frac{k}{r} \right\rceil , \end{aligned}$$

which coincides with the minimum Hamming distance of the given $[n, k, d]_q$ r-local Singleton-optimal code. □

Example 11 (Shortened Singleton-optimal LRC)**.** *We shorten the $[12, 6, 6]_{13}$ 3-local Singleton-optimal code of Example 2 in [18] 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*

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

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

$$\begin{aligned}(n'_i,r'_i)&=(n_i,r_i),\qquad \forall i\in[s]\setminus\{\alpha\},\\(n'_\iota,r'_\iota)&=(r_\alpha,r_\alpha-1),\\(n'_\alpha,r'_\alpha)&=(n_\alpha-r_\alpha-1,r_\alpha).\end{aligned}\tag{6}$$

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 $((n'_1,r'_1),(n'_2,r'_2),$ ... $,(n'_\iota,r'_\iota),$ ... $,(n'_s,r'_s))$-local ML-LRC is Singleton-optimal.

Proof. See appendix. □

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.

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 $\mathcal{R}_1,\mathcal{R}_2,\dots,\mathcal{R}_\rho$.

for $i\in[\frac{n_1}{r_1+1}]$ **do**
  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.

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

$$n'=n_2+(r_2+1)\frac{n_1}{r_1+1},\tag{7}$$

$$k'=k+(r_2-r_1)\frac{n_1}{r_1+1},\tag{8}$$

equals

$$d'=n'-k'+2-\left\lceil\frac{k'}{r_2}\right\rceil.\tag{9}$$

Inserting the expression of the length and the dimension as in (7) respectively (8) into (9) leads to:

$$\begin{aligned}
d' &\leq n_2 + (r_2+1)\frac{n_1}{r_1+1} - k - (r_2-r_1)\frac{n_1}{r_1+1} + 2 \\
&\quad - \left\lceil \frac{k + (r_2-r_1)\frac{n_1}{r_1+1}}{r_2} \right\rceil \\
&= n_2 + \frac{n_1}{r_1+1} - k + r_1\frac{n_1}{r_1+1} + 2 - \left\lceil \frac{k - r_1\frac{n_1}{r_1+1}}{r_2} \right\rceil \\
&\quad - \frac{n_1}{r_1+1} \\
&= n_2 - k + r_1\frac{n_1}{r_1+1} + 2 - \left\lceil \frac{k - r_1\frac{n_1}{r_1+1}}{r_2} \right\rceil \\
&= n_2 - k + (r_1+1)\frac{n_1}{r_1+1} + 2 - \frac{n_1}{r_1+1} - \left\lceil \frac{k - r_1\frac{n_1}{r_1+1}}{r_2} \right\rceil \\
&= n_1 + n_2 - k + 2 - \frac{n_1}{r_1+1} - \left\lceil \frac{k - r_1\frac{n_1}{r_1+1}}{r_2} \right\rceil, \qquad (10)
\end{aligned}$$

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:

$$k' \leq \frac{r_2}{r_2+1} n', \qquad (11)$$

and inserting the expressions of (7) and (8) in (11) 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. □

Pyramid codes [8] are a class of Singleton-optimal *information-symbol*-local codes. The construction is based on an MDS code and the maximal code length of a Pyramid code scales linearly with the field size. Let $\mathbf{a}\cdot\mathbf{b} = \sum_{i\in[n]} a_i b_i$ denote the dot-product for two given vectors $\mathbf{a}, \mathbf{b} \in \mathbb{F}_q^n$.

Construction 14 (ML-Pyramid Code)**.** *Let s integers $r_1, r_2, \ldots, r_s$ with $r_1 < r_2 < \cdots < r_s$ be given. Denote by $\mathcal{T}_i \subseteq [k]$ with $k_i \stackrel{\text{def}}{=} |\mathcal{T}_i|$ for all $i \in [s]$ pairwise disjoint sets and let $\cup_{i\in[s]}\mathcal{T}_i = [k]$. We partition $\mathcal{T}_i$ into $\kappa_i = \lceil k_i/r_i \rceil$ disjoint subsets $\mathcal{I}_{i,1}, \mathcal{I}_{i,2}, \ldots, \mathcal{I}_{i,\kappa_i}$, such that $\mathcal{T}_i = \cup_{j\in[\kappa_i]}\mathcal{I}_{i,j}$ and $|\mathcal{I}_{i,j}| \leq r_i$ for all $i \in [s]$. Let a $[k+d-1, k, d]_q$ MDS code $\mathcal{C}$ and an information vector $\mathbf{x} \in \mathbb{F}_q^k$ be given. Any codeword $\mathbf{c} \in \mathcal{C}$ is of the following form:*

$$\mathbf{c} = \left(\mathbf{x}, \mathbf{p}^{\langle 1\rangle}\cdot\mathbf{x}, \mathbf{p}^{\langle 2\rangle}\cdot\mathbf{x}, \ldots, \mathbf{p}^{\langle d-1\rangle}\cdot\mathbf{x}\right).$$

The modified codeword of an information-symbol-local ML-LRC is then:

$$\begin{aligned}
\Big(\mathbf{x}, \mathbf{p}^{\langle 1\rangle}_{\mathcal{I}_{1,1}}\cdot\mathbf{x}_{\mathcal{I}_{1,1}}, \ldots, \mathbf{p}^{\langle 1\rangle}_{\mathcal{I}_{1,\kappa_1}}\cdot\mathbf{x}_{\mathcal{I}_{1,\kappa_1}}, \mathbf{p}^{\langle 1\rangle}_{\mathcal{I}_{2,1}}\cdot\mathbf{x}_{\mathcal{I}_{2,1}}, \ldots, \\
\mathbf{p}^{\langle 1\rangle}_{\mathcal{I}_{s,\kappa_s}}\cdot\mathbf{x}_{\mathcal{I}_{s,\kappa_s}}, \mathbf{p}^{\langle 2\rangle}\cdot\mathbf{x}, \ldots, \mathbf{p}^{\langle d-1\rangle}\cdot\mathbf{x}\Big).
\end{aligned}$$

Clearly, the obtained multiple-locality Pyramid code has minimum Hamming distance d and length $k + \sum_{i\in[s]}\lceil k_i/r_i\rceil + d - 2$.

Lemma 15 (Singleton-Optimal Construction)**.** *The ML-Pyramid code as in Construction 14 has $n_i = k_i + \lceil k_i/r_i\rceil = \lceil k_i(r_i+1)/r_i\rceil$ symbols with locality r_i for all $i \in [s]$. We get from the bound in Thm. 8:*

$$\begin{aligned}
d &\leq n - k + 2 - \sum_{i\in[s-1]}\left\lceil \frac{n_i}{r_i+1}\right\rceil - \left\lceil \frac{k - \sum\limits_{i\in[s-1]} r_i\left\lceil \frac{n_i}{r_i+1}\right\rceil}{r_s}\right\rceil \\
&= n - k + 2 - \sum_{i\in[s-1]}\left\lceil \frac{k_i}{r_i}\right\rceil - \left\lceil \frac{k - \sum\limits_{i\in[s-1]} r_i\left\lceil \frac{k_i}{r_i}\right\rceil}{r_s}\right\rceil. \qquad (12)
\end{aligned}$$

With $k_{s-1} = k - \sum_{i\in[s-1]} k_i$, we obtain from (12)*:*

$$d \leq n - k + 2 - \sum_{i\in[s-1]} \left\lceil \frac{k_i}{r_i} \right\rceil - \left\lceil \frac{k_{s-1}}{r_s} \right\rceil,$$

which coincides with the minimum Hamming distance of the constructed ML-Pyramid 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 [4]

$$H(\mathcal{I}) \stackrel{\text{def}}{=} \frac{\log |\{\mathbf{x}_{\mathcal{I}} : \mathbf{x} \in \mathcal{C}\}|}{\log q}. \tag{13}$$

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

Lemma 16 ([4, 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* (13) *exists (and is constructed explicitly).*

Lemma 17 ([4, 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 16 slightly and generalize it to an $((n_1, r_1), (n_2, r_2), \ldots, (n_s, r_s))$-local ML-LRC.

Lemma 18. *Let $\mathcal{T}_i$ with $|\mathcal{T}_i| = n_i$ for all $i \in [s]$ be the locality sets of a given $(n, k, d)_q$ $((n_1, r_1), (n_2, r_2), \ldots, (n_s, r_s))$-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*

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

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

Proof. Similar to [4, 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 [4, 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 [4, Lemma 1], it is shown that $H(\mathcal{I}_i) = H(\mathcal{I}_i \setminus \{a_0, a_1, \ldots, a_{t_i-1}\})$. 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

$$\begin{aligned} H(\mathcal{I}_i) &\leq n_i - \lceil n_i/(r_i+1) \rceil \\ &= \left\lceil \frac{r_i n_i}{r_i+1} \right\rceil . \end{aligned} \tag{14}$$

□

The distinction between Cases a) and b) in Lemma 18 is not relevant for the bound of an r-local code, but becomes necessary if we want to bound the dimension of an $((n_1,r_1),(n_2,r_2),\ldots,(n_s,r_s))$-local ML-LRC.

Lemma 19. *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 17 consecutively s times and the statement follows. □

From Lemma 18 and Lemma 19 the following theorem follows.

Theorem 20 (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*

$$\begin{aligned} &k_{t_1,t_2} \stackrel{\text{def}}{=} \\ &t_1 r_1 + t_2 r_2 + k_{opt}^{(q)}\left(n - \min(n_1, t_1(r_1+1)) - \min(n_2, t_2(r_2+1)), d\right). \end{aligned} \tag{15}$$

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

$$k \leq \min_{\substack{t_1\in[t_1^\star],\\ t_2\in[t_2^\star]}} k_{t_1,t_2}, \tag{16}$$

where

$$t_1^\star \stackrel{\text{def}}{=} \left\lceil \frac{n_1}{r_1+1} \right\rceil, \tag{17}$$

$$t_2^\star \stackrel{\text{def}}{=} \left\lfloor \frac{k-1-t_1 r_1}{r_2} \right\rfloor, \quad \forall t_1 \in [t_1^\star]. \tag{18}$$

Proof. The expressions as in (15) and (16) follow from Lemma 18 and Lemma 19 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 $r_1\lceil n_1/(r_1+1)\rceil < k-1$ and therefore we obtain (17). The maximal value for t_2 follows from the fact that for $t_2 > t_2^*$ the expression in (16) 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:

$$\left\lfloor \frac{k-1-r_1 t_1}{r_2} \right\rfloor \leq \left\lceil \frac{n_2}{r_2+1} \right\rceil,$$

and therefore the maximal $t_2^\star$ as in (18) follows. □

Corollary 21 (Singleton vs. Alphabet-Dependent Bound)**.** *The Singleton bound as in Thm. 6 for an $[n,k,d]_q$ $((n_1,r_1),(n_2,r_2))$-local code is weaker than the bound given in Thm. 20.*

Proof. We first bound the dimension $k_{\text{opt}}^{(q)}$ by the locality-unaware Singleton bound.

$$\begin{aligned} k &\leq \min_{t_1,t_2}\Big(t_1 r_1 + t_2 r_2 + k_{\text{opt}}^{(q)}(n - t_1(r_1+1) - t_2(r_2+1), d)\Big) \\ &\leq t_1 r_1 + t_2 r_2 + n - t_1(r_1+1) - t_2(r_2+1) - d + 1 \\ &= n - d + 1 - t_1 - t_2. \end{aligned} \tag{19}$$

We assume that $d > n - k + 2 - \lceil n_1/(r_1+1)\rceil - \lceil (k - r_1\lceil n_1/(r_1+1)\rceil)/r_2\rceil$ and inserted in (19) leads to:

$$\begin{aligned} k &< n - n + k - 2 + \left\lceil \frac{n_1}{r_1+1} \right\rceil + \left\lceil \frac{k - r_1 \left\lceil \frac{n_1}{r_1+1} \right\rceil}{r_2} \right\rceil \\ &\quad + 1 - t_1 - t_2 \\ &= k - 1 + \left\lceil \frac{n_1}{r_1+1} \right\rceil + \left\lceil \frac{k - r_1 \left\lceil \frac{n_1}{r_1+1} \right\rceil}{r_2} \right\rceil - t_1 - t_2. \end{aligned} \tag{20}$$

For the maximal values of t_1^* and t_2^* as in (17) and (18), (20) gives a contradiction. □

Note that in the classical case of an r-local LRC the Cadambe–Mazumdar bound (Thm. 3) gives also a contradiction (if $d > n - k + 2 - \lceil k/r\rceil$) for the maximal value of $t = \lceil k/r\rceil$.

The following theorem generalizes Thm. 20 to the case of $s > 2$ different localities.

Theorem 22 (Alphabet-Dependent Bound for ML-LRCs)**.** *Let $\mathcal{C}$ be an $(n,k,d)_q$ $((n_1,r_1),(n_2,r_2),\ldots,(n_s,r_s))$-local code as in Definition 7. Furthermore, assume $\sum_{i\in[s]} r_i\lceil n_i/(r_i+1)\rceil < k-1$. Define*

$$t_i^\star \stackrel{\text{def}}{=} \left\lceil \frac{n_i}{r_i+1} \right\rceil, \quad \forall i \in [s-1], \tag{21}$$

$$t_s^\star \stackrel{\text{def}}{=} \left\lfloor \frac{k - 1 - \sum\limits_{i\in[s-1]} t_i r_i}{r_s} \right\rfloor, \quad \forall t_i \in [t_i^\star], \tag{22}$$

and let

$$k_{t_1,t_2,\ldots,t_s} \stackrel{\text{def}}{=} \sum_{i\in[s]} t_i r_i + k_{opt}^{(q)}\left(n - \sum_{i\in[s]} \min(n_i, t_i(r_i+1)), d\right). \tag{23}$$

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

$$k \le \min_{\substack{t_i\in[t_i^\star] \\ \forall i\in[s]}} k_{t_1,t_2,\ldots,t_s}. \tag{24}$$

Proof. See appendix. □

The following corollary generalizes Corollary 21 to the case of $s > 2$ different localities.

Corollary 23 (Singleton vs. Alphabet-Dependent Bound)**.** *The Singleton bound as in Thm. 8 for an $[n,k,d]_q$ $((n_1,r_1),(n_2,r_2),$ ... $,(n_s,r_s))$-local code is weaker than the bound given in Thm. 22.*

Proof. See appendix. □

If there exist s parameters $t_i \in [t_i^\star], \forall i \in [s]$ such that the dimension k of an $[n,k,d]_q$ $((n_1,r_1),(n_2,r_2),\ldots,(n_s,r_s))$-local ML-LRC equals $k_{t_1,t_2,\ldots,t_s}$ as in (23), then we call the ML-LRC *alphabet-optimal*.

B. Shortening and Construction

Similar to Lemma 10, we now prove that shortening an r-local LRC, which is optimal with respect to the bound given in Thm. 3, by one coordinate gives an $((r,r-1),(n-1-r,r))$-local alphabet-optimal ML-LRC. We give an explicit construction (in Algorithm 3) of a ML-LRC with two different localities based on a given an alphabet-optimal LRC.

Lemma 24 (Shortening an r-local Code)**.** *Let an $[n,k,d]_q$ r-local alphabet-optimal code be given. Then, the shortened $[n-1,k-1,d]_q$ $((r,r-1),(n-1-r,r))$-local code is alphabet-optimal.*

Proof. Clearly, the shortening affects one repair-set and its cardinality is r and the locality of the contained code symbols is $r-1$. Let t' be the index of the minimal value $k_{t'}$ of (2) such that the dimension of the given code is

$$k = k_{t'} = t'r + k_{\text{opt}}^{(q)}(n - t'(r+1), d).$$

We obtain from Thm. 20 with $n_1 = r$, $r_1 = r-1$, $r_2 = r$, $t_1 = 1$ and $t_2 = t'-1$ for the shortened code of length $n-1$ and dimension $k-1$:

$$\begin{aligned} k_{1,t'-1} &= r\text{-}1 + (t'\text{-}1)r + k_{\text{opt}}^{(q)}(n-1-r-(t'\text{-}1)(r+1), d) \\ &= t'r - 1 + k_{\text{opt}}^{(q)}(n - t'(r+1), d), \end{aligned}$$

which equals $k_{t'} - 1$. Therefore, the dimension of the shortened code meets the bound of Thm. 20 and is alphabet-optimal. □

Algorithm 3 returns an ML-LRC with two localities that is optimal with respect to the bound given in Thm. 20.

Algo 3: Alphabet-Optimal ML-LRC over small field size.

Input:
Parity-check matrix $\mathbf{H}$ of an $[n,k,d]_q$ r_2-local alphabet-optimal code with $\rho = \lceil n/(r_2+1) \rceil$ repair sets $\mathcal{R}_1, \mathcal{R}_2, \ldots, \mathcal{R}_\rho$.
$\alpha \in [\rho]$.

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

Output: Parity-check matrix $\mathbf{H}$ of an $[n-\alpha(r_2-r_1), k-\alpha(r_2-r_1), d]_q$ $((\alpha(r_1+1), r_1), (n-\alpha(r_2+1), r_2))$-local code.

For a sufficiently large field-size a Singleton-optimal LRC can be constructed for a wide choice of parameters, but an alphabet-optimal LRC for a (aimed) smaller field-size are not known for all parameters. Therefore, the parameters in Algorithm 3 are expressed in terms of the given parameters n, k, d of the given alphabet-optimal r-local code.

Theorem 25 (Alphabet-Optimal ML-LRC)**.** *Given an $[n,k,d]_q$ r_2-local alphabet-optimal LRC and a integer $\alpha \in [\lceil n/(r_2+1) \rceil]$, Algorithm 3 returns a parity-check matrix of an $[n-\alpha(r_2-r_1), k-\alpha(r_2-r_1), d]_q$ $((\alpha(r_1+1), r_1), (n-\alpha(r_2+1), r_2))$-local ML-LRC that is alphabet-optimal.*

Proof. For the given $[n,k,d]_q$ r_2-local alphabet-optimal code, there exists an integer t' such that

$$k = k'_t = t'r_2 + k_{\text{opt}}^{(q)}(n - t'(r_2+1), d). \tag{25}$$

The shortened code is an $[n-\alpha(r_2-r_1), k-\alpha(r_2-r_1), d]_q$ $((\alpha(r_1+1), r_1), (n-\alpha(r_2+1), r_2))$-local code. With $t_1 = \alpha$ and $t_2 = t' - \alpha$, we obtain with the new alphabet-dependent bound of Thm. 20 that the dimension of the shortened code is upper bounded by

$$\begin{aligned} k_{\alpha,t'\text{-}\alpha} = \alpha r_1 + (t'\text{-}\alpha)r_2 + \\ k_{\text{opt}}^{(q)}(n\text{-}\alpha(r_2\text{-}r_1)\text{-}\alpha(r_1+1)\text{-}(t'\text{-}\alpha)(r_2+1), d), \end{aligned}$$

and this leads to:

$$\begin{aligned} k_{\alpha,t'\text{-}\alpha} &= t'r_2 - \alpha(r_2-r_1) + k_{\text{opt}}^{(q)}(n - t'(r_2+1), d), \\ &= k_{t'} - \alpha(r_2 - r_1). \end{aligned} \tag{26}$$

The expression in (26) equals the dimension of the shortened $((\alpha(r_1+1), r_1), (n-\alpha(r_2+1), r_2))$-local ML-LRC and therefore it is optimal with respect to the new alphabet-dependent bound of Thm. 20. □

Similar to Algorithm 1, Algorithm 3 can be easily be extended to $s \geq 2$ localities.

V. Alphabet-Optimal Binary LRCs and ML-LRCs

This section provides a family of binary r-local LRCs. Their construction is based on generalized (or multilevel) code concatenation introduced for linear components by Blokh and Zyablov [1] and described for not necessarily linear components by Zinoviev [21]. Multilevel code concatenation generalizes classical concatenation introduced by Forney [5]. Furthermore, we give an example of a shortened r-local code and show that it is optimal with respect to our new alphabet-dependent bound for ML-LRCs from Thm. 20.

Let us recall some definitions and facts before we define a generalized concatenated code. The Kronecker product of two matrices $\mathbf{A} = (a_{ij})_{i\in[m]}^{j\in[m]}$ and $\mathbf{B}$ is defined as $\mathbf{A} \otimes \mathbf{B} = (a_{ij}\mathbf{B})_{i\in[m]}^{j\in[m]}$. For two linear codes $\mathcal{A}$ and $\mathcal{B}$ of same length, define the direct sum code as $\mathcal{A} \oplus \mathcal{B} \stackrel{\text{def}}{=} \{\mathbf{a} \oplus \mathbf{b} | \mathbf{a} \in \mathcal{A}, \mathbf{b} \in \mathcal{B}\}$. For a linear subcode $\mathcal{B}^{\langle 2 \rangle}$ of $\mathcal{B}^{\langle 1 \rangle}$ with generator matrix $\mathbf{G}_{\mathcal{B}^{\langle 2 \rangle}}$, let

$\mathcal{B}^{\langle 2\rangle} \setminus \mathcal{B}^{\langle 1\rangle}$ be such that $\mathcal{B}^{\langle 1\rangle} = \mathcal{B}^{\langle 2\rangle} \oplus \mathcal{B}^{\langle 1\rangle} \setminus \mathcal{B}^{\langle 2\rangle}$. It is well-known (see, e.g., [2]) that there exists a generator matrix of $\mathcal{B}^{\langle 1\rangle}$ that can be written as

$$\mathbf{G}_{\mathcal{B}^{\langle 1\rangle}} = \begin{pmatrix} \mathbf{G}_{\mathcal{B}^{\langle 2\rangle}} \\ \mathbf{G}_{\mathcal{B}^{\langle 2\rangle}\setminus\mathcal{B}^{\langle 1\rangle}} \end{pmatrix}.$$

Now, we can define generalized code concatenation.

Definition 26 (Generalized Concatenated Code (GCC))**.** *Let $\mathcal{A}^{\langle i\rangle}$ be an outer $[n_a, k_{a,i}, d_{a,i}]_{q^{l_i}}$ code with a generator matrix $\mathbf{G}_{\mathcal{A}^{\langle i\rangle}} \in \mathbb{F}_{q^{l_i}}^{k_{a,i}\times n_a}$ and let $\mathcal{B}^{\langle i\rangle}$ be an $[n_b, k_{b,i}, d_{b,i}]_q$ inner code for all $i \in [s]$. Furthermore, let*

$$\mathcal{B}^{\langle 1\rangle} \supset \mathcal{B}^{\langle 2\rangle} \supset \cdots \supset \mathcal{B}^{\langle s\rangle},$$

and let $\mathbf{G}_{\mathcal{B}^{\langle i\rangle}\setminus\mathcal{B}^{\langle i+1\rangle}} \in \mathbb{F}_q^{(k_{b,i}-k_{b,i+1})\times n_b}$ denote a generator matrix of the code $\mathcal{B}^{\langle i\rangle}\setminus\mathcal{B}^{\langle i+1\rangle}$ for all $i \in [s-1]$. Let $k_{b,i}-k_{b,i+1} = \lambda_i l_i$ for all $i \in [s-1]$ and let $k_{b,s} = \lambda_s l_s$.

The matrix $\overline{\mathbf{G}}_{\mathcal{A}^{\langle i\rangle}} \in \mathbb{F}_q^{k_{a,i}l_i\times n_a}$ is the given $k_{a,i}\times n_a$ generator matrix $\mathbf{G}_{\mathcal{A}^{\langle i\rangle}}$ represented over $\mathbb{F}_q$. The code with generator matrix

$$\mathbf{G} = \begin{pmatrix} \overline{\mathbf{G}}_{\mathcal{A}^{\langle 1\rangle}} \otimes \mathbf{G}_{\mathcal{B}^{\langle 1\rangle}\setminus\mathcal{B}^{\langle 2\rangle}} \\ \overline{\mathbf{G}}_{\mathcal{A}^{\langle 2\rangle}} \otimes \mathbf{G}_{\mathcal{B}^{\langle 2\rangle}\setminus\mathcal{B}^{\langle 3\rangle}} \\ \vdots \\ \overline{\mathbf{G}}_{\mathcal{A}^{\langle s-1\rangle}} \otimes \mathbf{G}_{\mathcal{B}^{\langle s-1\rangle}\setminus\mathcal{B}^{\langle s\rangle}} \\ \overline{\mathbf{G}}_{\mathcal{A}^{\langle s\rangle}} \otimes \mathbf{G}_{\mathcal{B}^{\langle s\rangle}} \end{pmatrix} \tag{27}$$

is an $[n = n_a n_b, k = \sum_{i\in[s]} k_{a,i}\lambda_i l_i, d]_q$ s-level concatenated code $\mathcal{C}$.

It is well-known that the minimum Hamming distance of a generalized concatenated code is $d \geq \min_{i\in[s]} d_{a,i}d_{b,i}$ (see, e.g., [1, 21]). Every column of $\mathbf{G}$ of (27) belongs to $\mathcal{B}^{\langle 1\rangle}$ (assuming that the inner codewords are represented as rows in the codeword matrix). The following construction is based on this fact and generalized concatenated code inherits the locality property from $\mathcal{B}^{\langle 1\rangle}$.

Construction 27 (2-Level r-local GCC)**.** *Let $r > 1$ be the given locality parameter and let a integer be $j < 2^{r-1} - r + 1$. Let the two outer codes and the two inner codes of a 2-level concatenated code as in Definition 26 be a:*

- $\mathcal{A}^{\langle 1\rangle}$: $[2^{r-1}+1-j, 2^{r-1}-r+1-j, r+1]_{2^{r-1}}$ *MDS code,*
- $\mathcal{A}^{\langle 2\rangle}$: $[2^{r-1}+1-j, 2^{r-1}-j, 2]_2$ *single-parity check code,*
- $\mathcal{B}^{\langle 1\rangle}$: $[r+1, r, 2]_2$ *single-parity check code,*
- $\mathcal{B}^{\langle 2\rangle}$: $[r+1, 1, r+1]_2$ *repetition code.*

Then the 2-level concatenated code is an

$$[(r+1)(2^{r-1}+1-j), r(2^{r-1}-r+2-j)-1, 2(r+1)]_2$$

r-local code.

The following theorem shows the optimality of Construction 27 with respect to the bound given in Thm. 3 for $r = 2, 3$.

Theorem 28 (Alphabet-Optimal Binary LRC)**.** *For $r = 2, 3$ Construction 27 gives alphabet-optimal LRCs.*

Proof. It is sufficient to show it for the case of $j = 0$. The general case $j > 0$ follows directly, because the length of the code is reduced by $j(r+1)$ and the dimension by jr and Thm. 3 gives the same bound for $t = t' - j$ as t' for the code with $j = 0$.

The upper bound on the dimension of an r-local code from Thm. 3 gives for a binary code as in Construction 27 for $t = 2^{r-1} - r + 1$:

$$\begin{aligned} &k_{2^{r-1}-r+1} \\ &= r(2^{r-1}-r+1) \\ &+ k_{opt}^{(2)}\big((r+1)(2^{r-1}+1)\text{-}(2^{r-1}\text{-}r+1)(r+1), 2(r+1)\big) \\ &= r(2^{r-1}-r+1) + k_{opt}^{(2)}\big(r(r+1), 2(r+1)\big). \end{aligned} \tag{28}$$

For $r = 2, 3$, the maximal dimension of a binary code of length $r(r+1)$ and minimum Hamming distance $2(r+1)$ is $k_{opt}^{(2)}\big(r(r+1), 2(r+1)\big) = r - 1$ and therefore $k_{2^{r-1}-r+1}$ equals the dimension of the 2-level generalized concatenated code of Construction 27. ∎

We now give an example of a binary alphabet-optimal ML-LRC obtained via shortening an alphabet-optimal LRC as in Thm. 28.

Example 29 (Alphabet-Optimal Binary ML-LRC)**.** *Let $r = 3$ in Construction 27, then the two outer and inner codes of Construction 27 are:*

$\mathcal{A}^{\langle 1 \rangle}$: $[5, 2, 4]_{2^2}$ *MDS code,*
$\mathcal{A}^{\langle 2 \rangle}$: $[5, 4, 2]_2$ *single-parity check code,*
$\mathcal{B}^{\langle 1 \rangle}$: $[4, 3, 2]_2$ *single-parity check code,*
$\mathcal{B}^{\langle 2 \rangle}$: $[4, 1, 4]_2$ *repetition code.*

The 2-level concatenated code $\mathcal{C}$ is a $[20, 8, 8]_2$ 3-local alphabet-optimal code.

The bound as in Thm. 3 gives the following upper bound on the dimension of such a binary code:

$$\begin{aligned} k_2 &= 2 \cdot 3 + k_{opt}^{(2)}(20 - 2(3+1), 8) \\ &= 6 + k_{opt}^{(2)}(12, 8), \end{aligned}$$

where the maximal dimension of a binary code of length 12 and with minimum Hamming distance 8 is $k_{opt}^{(2)}(12, 8) = 2$ and therefore $k_2 = 8$, which is the dimension of our constructed code $\mathcal{C}$. Now we shortened the $[20, 8, 8]_2$ 3-local alphabet-optimal code by one position and obtain an $[19, 7, 8]_2$ $((3, 2), (16, 3))$-local code and we obtain from Thm. 20 that:

$$\begin{aligned} k_{1,1} &= 1 \cdot 2 + 1 \cdot 3 + k_{opt}^{(2)}(19 - 3 - 4, 8) \\ &= 5 + k_{opt}^{(2)}(12, 8) = 7, \end{aligned}$$

which equals the dimension of the shortened $((3, 2), (16, 3))$-local ML-LRC.

VI. Conclusion and Outlook

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. Furthermore, we gave constructions of optimal codes with respect to the two new bounds based on shortening a Singleton respectively alphabet-optimal LRC. An adapted Pyramid code construction for the case of different information-symbol locality was outlined. Moreover, we gave a class of binary r-local LRCs based on generalized code concatenation and showed that they are alphabet-optimal for $r = 2, 3$ and also provide optimal ML-LRCs over the binary alphabet.

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).

Acknowledgment

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

References

[1] E. L. Blokh and V. V. Zyablov. "Coding of Generalized Concatenated Codes". *Problems of Information Transmission* 10.3 (1974), pp. 45–50.
[2] M. Bossert, H. Grießer, J. Maucher, and v. v. Zyablov. "Some Results on Generalized Concatenation of Block Codes". *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes*. Ed. by M. Fossorier, H. Imai, S. Lin, and A. Poli. Lecture Notes in Computer Science 1719. Springer Berlin Heidelberg, Jan. 1999, pp. 181–190.
[3] V. Cadambe and A. Mazumdar. "An Upper Bound On the Size of Locally Recoverable Codes". *International Symposium on Network Coding (NetCod)*. Calgary, Canada, June 2013, pp. 1–5. DOI: 10.1109/NetCod.2013.6570829.
[4] V. Cadambe and A. Mazumdar. "An Upper Bound On the Size of Locally Recoverable Codes". *arXiv:1308.3200 [cs, math]* (Aug. 2013).
[5] G. D. Forney. "Concatenated Codes". PhD thesis. Cambridge, USA: Massachusetts Institute of Technology (MIT), 1966.
[6] 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. DOI: 10.1109/TIT.2012.2208937.
[7] 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. DOI: 10.1109/ISIT.2014.6874918.
[8] 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. DOI: 10.1109/NCA.2007.37.
[9] 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. DOI: http://dx.doi.org/10.1109/ITW.2015.7133128.
[10] 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. DOI: 10.1109/TIT.2014.2329872.
[11] M. Kuijper and D. Napp. "Erasure codes with simplex locality". *arXiv:1403.2779 [cs, math]* (Mar. 2014).
[12] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes*. North Holland Publishing Co., June 1988.
[13] F. Oggier and A. Datta. "Self-repairing Homomorphic Codes for Distributed Storage Systems". *IEEE INFOCOM*. Apr. 2011, pp. 1215–1223. DOI: 10.1109/INFCOM.2011.5934901.
[14] D. S. Papailiopoulos and A. G. Dimakis. "Locally Repairable Codes". *IEEE Transactions on Information Theory* 60.10 (Oct. 2014), pp. 5843–5855. DOI: 10.1109/TIT.2014.2325570.

[15] 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. DOI: 10.1109/ISIT.2012.6284028.
[16] 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. DOI: 10.1109/ISIT.2013.6620541.
[17] 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. DOI: 10.1109/ISIT.2015.7282655.
[18] I. Tamo and A. Barg. "A Family of Optimal Locally Recoverable Codes". *IEEE Transactions on Information Theory* 60.8 (Aug. 2014), pp. 4661–4676. DOI: 10.1109/TIT.2014.2321280.
[19] 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. DOI: 10.1109/ISIT.2013.6620540.
[20] 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. DOI: 10.1109/ITW.2015.7133123.
[21] V. A. Zinoviev. "Generalized Cascade Codes". *Problems of Information Transmission* 12.1 (1976), pp. 5–15.

APPENDIX

Proof of Thm. 8. Let $\mathbf{G}$ be a $k \times n$ generator matrix of the $[n,k,d]_q$ $((n_1,r_1),(n_2,r_2),\ldots,(n_s,r_s))$-local code. We assume that $\sum_{i\in[s-1]} r_i\lceil n_i/(r_i+1)\rceil < k-1$.

i) Choose $\kappa_i \stackrel{\text{def}}{=} \lceil n_i/(r_i+1)\rceil$ columns of $\mathbf{G}$ indexed by $j_1^{\langle i\rangle}, j_2^{\langle i\rangle},\ldots,j_{\kappa_i}^{\langle i\rangle}$, where $j_\iota^{\langle i\rangle} \in \mathcal{T}_i$ for all $i \in [s-1]$. Each column is a linear combination of at most r_i other columns.

s) Let $\mathcal{I}^{\langle 1\rangle},\mathcal{I}^{\langle 2\rangle},\ldots,\mathcal{I}^{\langle s-1\rangle}$ be the $s-1$ sets of indexes of all repair columns indexed by $j_1^{\langle i\rangle}, j_2^{\langle i\rangle},\ldots,j_{\kappa_i}^{\langle i\rangle}$ but without the indexes itself. We have $|\cup_{i\in[s-1]}\mathcal{I}^{\langle i\rangle}| \le \sum_{i\in[s-1]} r_i\cdot\kappa_i < k-1$.

s+1) Choose

$$\kappa_s \stackrel{\text{def}}{=} \left\lfloor \frac{k-1-\sum\limits_{i\in[s-1]} r_i\left\lceil\frac{n_i}{r_i+1}\right\rceil}{r_s}\right\rfloor$$

columns of $\mathbf{G}$ indexed by $j_1^{\langle s\rangle}, j_2^{\langle s\rangle},\ldots,j_{\kappa_s}^{\langle s\rangle}$, where $j_i^{\langle s\rangle} \in \mathcal{T}_s$. Each column is a linear combination of at most r_s other columns. Let $\mathcal{I}^{\langle s\rangle} \subset \mathcal{T}_s$ be the set of indexes of all repair columns indexed by $j_1^{\langle s\rangle}, j_2^{\langle s\rangle},\ldots,j_{\kappa_s}^{\langle s\rangle}$ but without the indexes themselves.

s+2) Let $\mathcal{I} \stackrel{\text{def}}{=} \cup_{i\in[s]}\mathcal{I}^{\langle i\rangle}$. Then $|\mathcal{I}| < k$ and we have $\operatorname{rank}(\mathbf{G}_{\mathcal{I}}) < k$.

s+3) Enlarge $\mathcal{I}$ to the set $\mathcal{I}'$, s.t. $\operatorname{rank}(\mathbf{G}_{\mathcal{I}'}) = k-1$, but without using $\left\{j_1^{\langle i\rangle}, j_2^{\langle i\rangle},\ldots,j_{\kappa_i}^{\langle i\rangle}\right\}_{i\in[s]}$.

s+4) Define

$$\mathcal{U} \stackrel{\text{def}}{=} \left\{\mathcal{I}' \cup \left\{j_1^{\langle 1\rangle}, j_2^{\langle 1\rangle},\ldots,j_{\kappa_s}^{\langle s\rangle}\right\}\right\}, \tag{29}$$

where $|\mathcal{U}| \ge k-1+\sum_{i\in[s]}\kappa_i$.

We have $\operatorname{rank}(\mathbf{G}_{\mathcal{U}}) \le k-1$. From Lemma 5, we know that $|\mathcal{U}|$ can be upper bounded, i.e.:

$$k-1+\sum_{i\in[s]}\kappa_i \le n-d,$$

and this implies for the minimum Hamming distance of an $((n_1,r_1),(n_2,r_2),\ldots,(n_s,r_s))$-local code

$$\begin{aligned} d &\le n\text{-}k\text{+}1 - \sum_{i\in[s-1]}\left\lceil\frac{n_i}{r_i+1}\right\rceil \text{-} \left\lfloor\frac{k\text{-}1-\sum\limits_{i\in[s-1]} r_i\left\lceil\frac{n_i}{r_i+1}\right\rceil}{r_s}\right\rfloor \\ &= n\text{-}k\text{+}2 - \sum_{i\in[s-1]}\left\lceil\frac{n_i}{r_i+1}\right\rceil \text{-} \left\lceil\frac{k-\sum\limits_{i\in[s-1]} r_i\left\lceil\frac{n_i}{r_i+1}\right\rceil}{r_s}\right\rceil. \end{aligned}$$

In case the assumption does not hold in Step j), i.e., $\sum_{i\in[j]} r_i\lceil n_i/(r_i+1)\rceil \ge k-1$, the bound becomes the bound for an $((n_1,r_1),(n_2,r_2),\ldots,(\sum_{i\in[j,s]} n_i, r_j))$-local code. Clearly, an $((n_1,r_1),(n_2,r_2),\ldots,(\sum_{i\in[j,s]} n_i, r_j))$-local code is also an $((n_1,r_1),(n_2,r_2),\ldots,(n_s,r_s))$-local code, because $r_1 < r_2 < \cdots < r_j < \cdots < r_s$. $\square$

Proof of Lemma 12. For the first case, we obtain from (5) with $r_{\alpha-1} = r_\alpha - 1$:

$$\begin{aligned}&\sum_{i\in[s-1]} \left\lceil \frac{n'_i}{r'_i+1} \right\rceil \\ &= \sum_{\substack{i\in[s-1]\setminus \\ \{\alpha-1,\alpha\}}} \left\lceil \frac{n'_i}{r'_i+1} \right\rceil + \left\lceil \frac{n'_{\alpha-1}}{r'_{\alpha-1}+1} \right\rceil + \left\lceil \frac{n'_\alpha}{r'_\alpha+1} \right\rceil .\end{aligned} \tag{30}$$

Inserting the expressions of n'_i, $n'_{\alpha-1}$ and n'_α as in (5) into (30) leads to:

$$\begin{aligned}&= \sum_{\substack{i\in[s-1]\setminus \\ \{\alpha-1,\alpha\}}} \left\lceil \frac{n_i}{r_i+1} \right\rceil + \left\lceil \frac{n_{\alpha-1}+r_\alpha}{r_{\alpha-1}+1} \right\rceil + \left\lceil \frac{n_\alpha - r_\alpha - 1}{r_\alpha+1} \right\rceil \\ &= \sum_{\substack{i\in[s-1]\setminus \\ \{\alpha-1,\alpha\}}} \left\lceil \frac{n_i}{r_i+1} \right\rceil + \left\lceil \frac{n_{\alpha-1}}{r_{\alpha-1}+1} + \frac{r_\alpha}{r_\alpha} \right\rceil + \left\lceil \frac{n_\alpha}{r_\alpha+1} - \frac{r_\alpha+1}{r_\alpha+1} \right\rceil \\ &= \sum_{i\in[s-1]} \left\lceil \frac{n_i}{r_i+1} \right\rceil .\end{aligned} \tag{31}$$

The second relevant term can be expressed as:

$$\begin{aligned}&\sum_{i\in[s-1]} r'_i \left\lceil \frac{n'_i}{r'_i+1} \right\rceil \\ &= \sum_{\substack{i\in[s-1]\setminus \\ \{\alpha-1,\alpha\}}} r'_i \left\lceil \frac{n'_i}{r'_i+1} \right\rceil + r'_{\alpha-1} \left\lceil \frac{n'_{\alpha-1}}{r'_{\alpha-1}+1} \right\rceil + r'_\alpha \left\lceil \frac{n'_\alpha}{r'_\alpha+1} \right\rceil ,\end{aligned} \tag{32}$$

and also with (5) for the expressions of n'_i, $n'_{\alpha-1}$ and n'_α inserted in (32) leads to:

$$\begin{aligned}&= \sum_{\substack{i\in[s-1]\setminus \\ \{\alpha-1,\alpha\}}} r_i \left\lceil \frac{n_i}{r_i+1} \right\rceil + r_{\alpha-1} \left\lceil \frac{n_{\alpha-1}+r_\alpha}{r_{\alpha-1}+1} \right\rceil \\ &\quad + r_\alpha \left\lceil \frac{n_\alpha - r_\alpha - 1}{r_\alpha+1} \right\rceil \\ &= \sum_{\substack{i\in[s-1]\setminus \\ \{\alpha-1,\alpha\}}} r_i \left\lceil \frac{n_i}{r_i+1} \right\rceil + (r_\alpha - 1) \left\lceil \frac{n_{\alpha-1}}{r_{\alpha-1}+1} + \frac{r_\alpha}{r_\alpha} \right\rceil \\ &\quad + r_\alpha \left\lceil \frac{n_\alpha}{r_\alpha+1} - \frac{r_\alpha+1}{r_\alpha+1} \right\rceil \\ &= \sum_{i\in[s-1]} r_i \left\lceil \frac{n_i}{r_i+1} \right\rceil - 1.\end{aligned} \tag{33}$$

For the second case, i.e., $r_{\alpha-1} \neq r_\alpha - 1$, we get for

$$\begin{aligned}&\sum_{i\in[s-1]\cup\{\iota\}} \left\lceil \frac{n'_i}{r'_i+1} \right\rceil \\ &= \sum_{\substack{i\in[s-1]\setminus \\ \{\alpha\}}} \left\lceil \frac{n'_i}{r'_i+1} \right\rceil + \left\lceil \frac{n'_\alpha}{r'_\alpha+1} \right\rceil + \left\lceil \frac{n'_\iota}{r'_\iota+1} \right\rceil\end{aligned} \tag{34}$$

Inserting the expressions of n'_i, n'_α and n'_ι as in (6) into (34) leads to:

$$
\begin{aligned}
&= \sum_{\substack{i\in[s-1]\setminus\\ \{\alpha\}}} \left\lceil \frac{n_i}{r_i+1} \right\rceil + \left\lceil \frac{n_\alpha - r_\alpha - 1}{r_\alpha + 1} \right\rceil + \left\lceil \frac{r_\alpha}{r_\alpha - 1 + 1} \right\rceil \\
&= \sum_{\substack{i\in[s-1]\setminus\\ \{\alpha\}}} \left\lceil \frac{n_i}{r_i+1} \right\rceil + \left\lceil \frac{n_\alpha}{r_\alpha+1} - \frac{r_\alpha+1}{r_\alpha+1} \right\rceil + \left\lceil \frac{r_\alpha}{r_\alpha} \right\rceil \\
&= \sum_{i\in[s-1]} \left\lceil \frac{n_i}{r_i+1} \right\rceil . \qquad (35)
\end{aligned}
$$

The second relevant expression is:

$$
\begin{aligned}
&\sum_{i\in[s-1]\cup\{\iota\}} r'_i \left\lceil \frac{n'_i}{r'_i+1} \right\rceil \\
&= \sum_{\substack{i\in[s-1]\setminus\\ \{\alpha\}}} r'_i \left\lceil \frac{n'_i}{r'_i+1} \right\rceil + r'_\alpha \left\lceil \frac{n'_\alpha}{r'_\alpha+1} \right\rceil + r'_\iota \left\lceil \frac{n'_\iota}{r'_\iota+1} \right\rceil , \qquad (36)
\end{aligned}
$$

and with (6) for the expressions of n'_i, n'_α and n'_ι inserted in (36) leads to:

$$
\begin{aligned}
&= \sum_{\substack{i\in[s-1]\setminus\\ \{\alpha\}}} r_i \left\lceil \frac{n_i}{r_i+1} \right\rceil + r_\alpha \left\lceil \frac{n_\alpha - r_\alpha - 1}{r_\alpha+1} \right\rceil \\
&\quad + (r_\alpha - 1) \left\lceil \frac{r_\alpha}{r_\alpha - 1 + 1} \right\rceil \\
&= \sum_{\substack{i\in[s-1]\setminus\\ \{\alpha\}}} r_i \left\lceil \frac{n_i}{r_i+1} \right\rceil + r_\alpha \left\lceil \frac{n_\alpha}{r_\alpha+1} \right\rceil - r_\alpha + (r_\alpha - 1) \left\lceil \frac{r_\alpha}{r_\alpha} \right\rceil \\
&= \sum_{i\in[s-1]} r_i \left\lceil \frac{n_i}{r_i+1} \right\rceil - 1. \qquad (37)
\end{aligned}
$$

The minimum Hamming distance of the shortened code for both cases is with (3):

$$
d \leq n'\text{-}k'\text{+2-} \sum_{i\in[s-1]} \left\lceil \frac{n'_i}{r'_i\text{+1}} \right\rceil \text{-} \left\lceil \frac{k'\text{-} \left(\sum_{i\in[s-1]} r'_i \left\lceil \frac{n'_i}{r'_i+1} \right\rceil \text{-1} \right)}{r'_s} \right\rceil . \qquad (38)
$$

With (31), (33), (35) and (37) inserted in (38) gives:

$$
\begin{aligned}
d &\leq n - 1 - (k-1) + 2 - \sum_{i\in[s-1]} \left\lceil \frac{n_i}{r_i+1} \right\rceil \\
&\quad - \left\lceil \frac{(k-1) - \left(\sum_{i\in[s-1]} r_i \left\lceil \frac{n_i}{r_i+1} \right\rceil - 1 \right)}{r_s} \right\rceil \\
&= n\text{-}k\text{+2-} \sum_{i\in[s-1]} \left\lceil \frac{n_i}{r_i\text{+1}} \right\rceil \text{-} \left\lceil \frac{k\text{-} \sum_{i\in[s-1]} r_i \left\lceil \frac{n_i}{r_i+1} \right\rceil}{r_s} \right\rceil ,
\end{aligned}
$$

which is the minimum Hamming distance of the original $[n,k,d]_q$ $((n_1,r_1),(n_2,r_2),\ldots,(n_s,r_s))$-local Singleton-optimal ML-LRC. □

Proof of Thm. 22. It follows from Lemma 18 and Lemma 19 for s disjoint sets and the expression as in (23) follows. The values of t_i can be bounded by $\min(\lceil n_i/(r_i+1)\rceil, \lfloor (k-1-\prod_{j\in[s-1]\setminus\{i\}} t_j r_j)/r_i\rfloor)$ for all $i \in [s-1]$. We assume that $\sum_{i\in[s-1]} r_i\lceil n_i/(r_i+1)\rceil < k-1$ and therefore we obtain (21). The minimizing value for t_s follows from the fact that for $t_s > t_s^*$ is greater than k and therefore:

$$t_s \leq \left\lfloor \frac{k-1-\sum\limits_{i\in[s-1]} t_i r_i}{r_s} \right\rfloor.$$

□

Proof of Corollary 23. We again first bound the dimension $k_{\text{opt}}^{(q)}$ by the locality-unaware Singleton bound.

$$\begin{aligned} k &\leq \min_{t_1,t_2,\ldots,t_s} \left(\sum_{i\in[s]} t_i r_i + k_{\text{opt}}^{(q)} \left(n - \sum_{i\in[s]} \min(n_i, t_i(r_i+1)), d \right) \right) \\ &= \sum_{i\in[s]} t_i r_i + n - \sum_{i\in[s]} \min(n_i, t_i(r_i+1)) - d + 1 \\ &= n - d + 1 - \sum_{i[s]} t_i. \end{aligned} \tag{39}$$

Assume that the minimum Hamming distance contradicts the Singleton-like bound, i.e.:

$$d > n - k + 2 - \sum_{i\in[s-1]} \left\lceil \frac{n_i}{r_i+1} \right\rceil - \left\lceil \frac{k - \sum\limits_{i\in[s-1]} r_i \left\lceil \frac{n_i}{r_i+1} \right\rceil}{r_s} \right\rceil,$$

and inserted into (39) gives:

$$\begin{aligned} & n - d + 1 - \sum_{i\in[s]} t_i \\ & \leq n - n + k - 2 + \sum_{i\in[s-1]} \left\lceil \frac{n_i}{r_i+1} \right\rceil \\ & \quad - \left\lceil \frac{k - \sum\limits_{i\in[s-1]} r_i \left\lceil \frac{n_i}{r_i+1} \right\rceil}{r_s} \right\rceil + 1 - \sum_{i\in[s]} t_i \\ & = k - 1 + \sum_{i\in[s-1]} \left\lceil \frac{n_i}{r_i+1} \right\rceil - \left\lceil \frac{k - \sum\limits_{i\in[s-1]} r_i \left\lceil \frac{n_i}{r_i+1} \right\rceil}{r_s} \right\rceil - \sum_{i\in[s]} t_i, \end{aligned}$$

which is a contradiction for the maximal values of $t_i^*, \forall i \in [s]$ as in (21) and (22). □