Grezet, Matthias; Hollanti, Camilla

The complete hierarchical locality of the punctured Simplex code

The complete hierarchical locality of the punctured Simplex code

Matthias Grezet[1] [ID] · **Camilla Hollanti[1]** [ID]

Abstract

This paper presents a new alphabet-dependent bound for codes with hierarchical locality. Then, the complete list of possible localities is derived for a class of codes obtained by deleting specific columns from a Simplex code. This list is used to show that these codes are optimal codes with hierarchical locality.

1 Introduction

In modern distributed storage systems (DSSs) failures happen frequently, whence decreasing the number of connections required for node repair is crucial. Locally repairable codes (LRCs) are a subclass of erasure-correcting codes, which allow a small number of failed nodes to be repaired by accessing only a few other nodes. LRCs were introduced in [6], [11] where the codes can locally repair one failure. They were later extended in [12], [8] to be able to repair more failures locally.

An $[n, k, d]$ linear code $\mathcal{C}$ of length n, dimension k, and minimum Hamming distance d, has all-symbol locality (r, δ) if for all code symbols $i \in [n] = \{1, \ldots, n\}$, there exists a set $R_i \subseteq [n]$ containing i such that $|R_i| \leq r + \delta - 1$ and the minimum distance of the restriction of $\mathcal{C}$ to R_i is at least δ. We refer to $\mathcal{C}$ as an (n, k, d, r, δ)-LRC and to the sets R_i as

---

Communicated by T. Etzion.

✉ Matthias Grezet
matthias.grezet@aalto.fi

Camilla Hollanti
camilla.hollanti@aalto.fi

[1] Department of Mathematics and Systems Analysis, Aalto University, 00076 AALTO (Espoo), Finland

repair sets or local sets. Related Singleton-type bounds have been derived for various cases in [6,11,12] and the first bound with a fixed code alphabet was obtained in [3] for $\delta = 2$. Constructions achieving the Singleton-type bounds and the bound in [3] for $\delta = 2$ were provided in [8,9,11–14,16–18,20].

The authors of [1] proposed the first alphabet-dependent bound on LRCs over an alphabet Q with $|Q| = q$ using an upper bound $B(n, d)$ on the cardinality of a code of length n and minimum distance d. The global bound is as follows:

$$k \leq \left(\left\lceil \frac{n - d + 1}{r + \delta - 1} \right\rceil + 1 \right) \log_q B(r + \delta - 1, \delta). \tag{1}$$

Recently, [7] provided a different alphabet-dependent bound for LRCs of the same type as the bound in [3] using the Griesmer bound $\mathcal{G}_q(k, d) := \sum_{i=0}^{k-1} \lceil d/q^i \rceil$. The bound has the following form : Any linear (n, k, d, r, δ)-LRC $\mathcal{C}$ with κ the upper bound on the dimension of the restriction of $\mathcal{C}$ to a repair set satisfies

$$k \leq \min_{\lambda \in \mathbb{Z}_+} \left\{ \lambda + k_{\text{opt}}^{(q)}(n - \mu, d) \right\} \tag{2}$$

where $a, b \in \mathbb{Z}$ such that $\lambda = a\kappa + b, 0 \leq b < \kappa$ and $\mu = (a + 1)\mathcal{G}_q(\kappa, \delta) - \mathcal{G}_q(\kappa - b, \delta)$.

In [15], the authors introduced the notion of codes with hierarchical locality (H-LRCs), which optimizes futher the number of nodes contacted for repair according to the number of failures. A 2-level H-LRC is a code where the restrictions to the repair sets are themselves LRCs, thus providing an extra layer of locality. If an H-LRC has locality $[(r_1, \delta_1), (r_2, \delta_2)]$, then the number of nodes contacted to repair up to $\delta_2 - 1$ failures is at most r_2 and the number of nodes contacted for repair is at most r_1 if the number of failures is $\geq \delta_2$ and $\leq \delta_1 - 1$. This concept can be easily generalized to an arbitrary level of hierarchy. A Singleton-type bound for codes with hierarchical locality was derived in [15] and constructions attaining the bound were provided in [2,15].

In Sect. 3, we show how we can adapt the construction algorithms provided in [15] to obtain an alphabet-dependent bound for H-LRCs of the same type as in [3]. By construction, this bound is at least as good as the Singleton-type bound derived in [15].

In Sect. 4, we study the locality of one particular construction of LRCs presented in [16]. The general idea of this construction is to remove a Simplex code from another Simplex code of higher dimension. It was shown in [16] that these codes achieve the Griesmer bound and are LRCs with $\delta = 2$. The goal in this section is to prove the locality for every dimension and higher δ and show that this construction leads to optimal LRCs for every locality by the bound (2) and to optimal H-LRCs by the new bound derived in Sect. 3.

As a first step, we describe the restrictions of dimension $k - 1$ and prove the locality for this dimension using combinatorial techniques. These results allow us to derive the weight enumerator of the constructed codes. As a second step, we use a recursive argument to get all the restrictions of these codes to closed sets. Our main contribution is the complete list of possible localities for these codes. In particular, this shows that the constructed codes are alphabet-optimal H-LRCs. Finally, since a special case of this construction leads to the Reed–Muller codes $RM(1, m)$, we obtain as a corollary to our result that the Reed–Muller codes $RM(1, m)$ are H-LRCs and we derive their locality parameters.

2 Preliminaries

We denote the set $\{1, 2, \ldots, n\}$ by $[n]$ and the set of all subsets of $[n]$ by $2^{[n]}$. The set of all positive integers including 0 is denoted by $\mathbb{Z}_+$.

The Gaussian coefficient, which counts the number of subspaces of dimension k in the vector space $\mathbb{F}_q^n$, is denoted by

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \prod_{i=0}^{k-1} \frac{q^{n-i} - 1}{q^{i+1} - 1}.$$

For a length-n vector $\mathbf{v}$ and a set $I \subseteq [n]$, the vector $\mathbf{v}_I$ denotes the restriction of the vector $\mathbf{v}$ to the coordinates in the set I. A generator matrix of a linear code $\mathcal{C}$ is $G_\mathcal{C} = (\mathbf{g_1} \cdots \mathbf{g_n})$ where $\mathbf{g_i} \in \mathbb{F}_q^k$ is a column vector for $i \in [n]$. The shortening of a code $\mathcal{C}$ to the set of coordinates $I \subseteq [n]$ is defined by $\mathcal{C}/I = \{\mathbf{c}_{[n]\setminus I} : \mathbf{c} \in \mathcal{C}$ such that $c_i = 0$ for all $i \in I\}$ and the restriction of a code $\mathcal{C}$ to I is defined by $\mathcal{C}|_I = \{\mathbf{c}_I : \mathbf{c} \in \mathcal{C}\}$. For convenience, we call the codes obtained by a restriction *restricted codes*. Two linear codes $\mathcal{C}$ and $\mathcal{C}'$ are called isomorphic if $\mathcal{C}'$ can be obtained by a permutation on the coordinates of the codewords of $\mathcal{C}$.

The support of a codeword $\mathbf{c} \in \mathcal{C}$ is $\mathrm{supp}(\mathbf{c}) = \{i \in [n] : c_i \neq 0\}$ and its weight is $\mathrm{wt}(\mathbf{c}) = |\mathrm{supp}(\mathbf{c})|$. The weight enumerator of $\mathcal{C}$ is defined as

$$W_\mathcal{C}(x, y) = \sum_{\mathbf{c} \in \mathcal{C}} x^{n - \mathrm{wt}(\mathbf{c})} y^{\mathrm{wt}(\mathbf{c})}.$$

The Simplex code $\mathcal{S}(m)$, or sometimes $\mathcal{S}_q(m)$, is a linear code over $\mathbb{F}_q$ obtained via the generator matrix G_m consisting of all pairwise linearly independent vectors in $\mathbb{F}_q^m$. The parameters of $\mathcal{S}(m)$ are therefore $[(q^m - 1)/(q - 1), m, q^{m-1}]$.

Since most of the work in this paper is done on subsets of coordinates and restricted codes, it is preferable to use the notion of an entropy function on the subsets $I \subseteq [n]$ that corresponds to the dimension of the restriction to I. We only state here the definition for linear codes but it can be generalize to bigger class of codes (see [19]).

Definition 1 Let $\mathcal{C}$ be a linear code of length n and $I \subseteq [n]$. The entropy associated to $\mathcal{C}$ is the function $H_\mathcal{C} : 2^{[n]} \to \mathbb{Z}$ with $H_\mathcal{C}(I) = \dim(\mathcal{C}|_I)$.

For ease of notation, if the underlying code of $H_\mathcal{C}$ is clear, we drop the specification to $\mathcal{C}$. For a subset $I \subseteq [n]$, $H_\mathcal{C}(I)$ is equivalent to the rank of the submatrix of the generator matrix formed by the columns $\mathbf{g}_i$ with $i \in I$ or the rank function of I in the associated matroid of $\mathcal{C}$. As such, it has the following standard properties.

Proposition 1 *Let $\mathcal{C}$ be a linear code of length n and H the entropy function associated to $\mathcal{C}$. For $I, J \subseteq [n]$, we have*

1. $H(I) \leq |I|$,
2. *If $I \subseteq J$ then $H(I) \leq H(J)$,*
3. $H(I \cup J) + H(I \cap J) \leq H(I) + H(J)$.

The entropy function also behaves nicely for restricted codes. Let $I \subseteq [n]$ and $\mathcal{C}|_I$ be the restriction of $\mathcal{C}$ to the set I. Then for $J \subseteq I$, we have $H_{\mathcal{C}|_I}(J) = H_\mathcal{C}(J)$.

Finally, we define a closure operation on the subsets of $[n]$ for linear codes.

Definition 2 Let $\mathcal{C}$ be a linear code of length n and $I \subseteq [n]$. The *closure operator* $\mathrm{cl} : 2^{[n]} \to 2^{[n]}$ is $\mathrm{cl}(I) = \{e \in [n] : H(I \cup e) = H(I)\}$. A set $I \subseteq [n]$ is a *closed set* if $\mathrm{cl}(I) = I$.

One can think of the closure operator via the generator matrix $G_\mathcal{C}$ of $\mathcal{C}$ where $\mathrm{cl}(I)$ is the set of all columns in $G_\mathcal{C}$ contained in the linear span of the columns indexed by I.

2.1 Preliminaries on matroids

Since we work on Simplex codes which have a lot of combinatorial properties, we use some tools coming from matroid theory. Matroids have many equivalent definitions in the literature. Here, we choose to present matroids via their rank functions. Much of the contents in this part can be found in more detail in [5].

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

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

Any matrix G over a field $\mathbb{F}$ generates a matroid $M_G = (E, \rho)$, where E is the set of indices of the columns of G, and $\rho(X)$ is the rank of the submatrix of G formed by the columns indexed by X. As elementary row operations preserve the row space of $G(X)$ for all $X \subseteq E$, it follows that row-equivalent matrices generate the same matroid. Thus, there is a straightforward connection between linear codes and matroids. Let $\mathcal{C}$ be a linear code of length n and H the entropy function associated to $\mathcal{C}$. Since the entropy function fits the requirements of a rank function of a matroid, we can associate a matroid to $\mathcal{C}$ by considering the matroid $M_\mathcal{C} = ([n], H)$.

Two matroids $M_1 = (E_1, \rho_1)$ and $M_2 = (E_2, \rho_2)$ are *isomorphic* if there exists a bijection $\psi : E_1 \to E_2$ such that $\rho_2(\psi(X)) = \rho_1(X)$ for all subsets $X \subseteq E_1$. We denote two isomorphic matroids by $M_1 \cong M_2$. This implies in particular that if $\mathcal{C}_1$ and $\mathcal{C}_2$ are two linear codes and $M_{\mathcal{C}_1}$ is isomorphic to $M_{\mathcal{C}_2}$, then $\mathcal{C}_1$ is isomorphic to $\mathcal{C}_2$.

One way of defining a new matroid from an existing one is obtained by restricting the matroid to one of its subsets. For a given set $Y \subseteq E$, we define the *restriction* of M to Y to be the matroid $M|Y = (Y, \rho_{|Y})$ by $\rho_{|Y}(X) = \rho(X)$ for all subsets $X \subseteq Y$. The restriction of M to $E - Y$ is called the *deletion* of Y and is denoted by $M \setminus Y$. The two previous operations correspond to the restriction and puncturing of a linear code $\mathcal{C}$.

Let $M = (E, \rho)$ be a matroid. The *closure* operator $\mathrm{cl} : 2^E \to 2^E$ is defined by $\mathrm{cl}(X) = \{e \in E : \rho(X \cup e) = \rho(X)\}$. A subset $F \subseteq E$ is a *flat* if $\mathrm{cl}(F) = F$. The collection of flats of M is denoted by $\mathcal{F}(M)$ and forms a lattice under the inclusion. For $F_1, F_2 \in \mathcal{F}(M)$, the meet is $F_1 \wedge F_2 = F_1 \cap F_2$ and join is $F_1 \vee F_2 = \mathrm{cl}(F_1 \cup F_2)$. We denote by $\lessdot$ the covering relation, i.e., $F_1 \lessdot F_2$ if $F_1 \subsetneq F_2$ and there is no $F_3 \in \mathcal{F}(M)$ with $F_1 \subsetneq F_3 \subsetneq F_2$. For $Y \subseteq E$, we can express the flats of $M|Y$ via the flats of M by the relation $\mathcal{F}(M|Y) = \{F \cap Y : F \in \mathcal{F}(M)\}$. A *hyperplane* H is a flat of M with $\rho(H) = \rho(E) - 1$ and the collection of all hyperplanes is denoted by $\mathcal{H}(M)$. One interesting property of $\mathcal{H}(M)$ is the following.

Proposition 2 ([10, Proposition 1.7.8]) *Let F be a flat in M and suppose that $\rho(F) = \rho(E) - m$ where $m \geq 1$. Then there is a set $\{H_1, H_2, \ldots, H_m\}$ of hyperplanes such that $F = \cap_{i=1}^m H_i$.*

Proof We argue by induction on m. The result is immediate for $m = 1$. Assume it is true for m and let F a flat with $\rho(F) = \rho(E) - (m + 1)$. Clearly, there exists an element $y \in E - F$. As F is a flat, $\mathrm{cl}(F \cup y)$ is a flat of M covering F. Thus, $\rho(\mathrm{cl}(F \cup y)) = \rho(E) - m$ and by the induction assumption, there are m hyperplanes $H_1, H_2, \ldots, H_m$ such that $\cap_{i=1}^m H_i = \mathrm{cl}(F \cup y)$. Now, either $E - y$ is a hyperplane of M or not. In the former case, let $H_{m+1} = E - y$. In the latter case, there is a hyperplane that contains F and is contained in $E - y$. Let H_{m+1} be this hyperplane.

In both cases, since $y \in \cap_{i=1}^{m} H_i$ but $y \notin H_{m+1}$, we have $\cap_{i=1}^{m+1} H_i \neq \cap_{i=1}^{m+1} H_i$. Thus,

$$F \subseteq \cap_{i=1}^{m+1} H_i \subsetneq \cap_{i=1}^{m} H_i = \mathrm{cl}(F \cup y).$$

But $\mathrm{cl}(F \cup y)$ covers F, so $F = \cap_{i=1}^{m+1} H_i$. □

3 Bound for H-LRCs

Codes with hierarchical locality were introduced in [15] to optimize further the number of nodes contacted for repair. In this section, we first focus on H-LRCs with 2-level hierarchy and derive an alphabet-dependent bound for these codes based on [3]. Then, we extend this bound to H-LRCs with h-level hierarchy.

Definition 4 Let $r_2 \leq r_1$ and $\delta_2 < \delta_1$. An $[n, k, d]$ linear code $\mathcal{C}$ is a code with hierarchical locality having locality parameters $[(r_1, \delta_1), (r_2, \delta_2)]$ if for all code symbols $i \in [n]$, there exists a set $M_i \subseteq [n]$ such that

1. $i \in M_i$,
2. $H(M_i) \leq r_1$,
3. The minimum distance of $\mathcal{C}|_{M_i}$ is at least δ_1,
4. $\mathcal{C}|_{M_i}$ is an LRC with (r_2, δ_2)-locality.

The codes $\mathcal{C}|_{M_i}$ are called *middle codes* and their restrictions of dimension $\leq r_2$ and minimum distance $\geq \delta_2$ are called *local codes*. Similarly, the *middle sets* and *local sets* are the sets M_i and the sets such that the restrictions to them give the local codes. Notice that, contrary to the standard definition of LRCs, the authors of [15] bound the dimension of the restricted codes instead of the size. H-LRCs, as storage codes, allow a more optimal repair process. If less than δ_2 failures occurred inside a local set, then the corresponding local code can be used to repair those failures by contacting at most r_2 nodes in the storage system. When the number of failures is in between δ_2 and $\delta_1 - 1$, the middle codes guarantee a repair where at most r_1 nodes are contacted. Finally, if the number of failures is in between δ_1 and $d - 1$, those failures can be repaired by contacting at most k nodes in the global code.

It was proven in [15, Theorem 2.1] that any $[n, k, d]$ linear code with hierarchical locality $[(r_1, \delta_1), (r_2, \delta_2)]$ satisfies the following Singleton-type bound

$$d \leq n - k + 1 - \left\lfloor \frac{k-1}{r_2} \right\rfloor (\delta_2 - 1) - \left\lfloor \frac{k-1}{r_1} \right\rfloor (\delta_1 - \delta_2). \tag{3}$$

To obtain the alphabet-dependent bound in [3], the authors proved two results: the construction of a restricted code with small dimension and large size and a lemma about shortened codes. The lemma is the following.

Lemma 1 ([3, Lemma 2]) *Let $\mathcal{C}$ be an $[n, k, d]$ linear code over $\mathbb{F}_q$ and $I \subseteq [n]$ such that $H(I) < k$. Then the shortened code $\mathcal{C}/I$ has parameters $[n - |I|, k - H(I), d' \geq d]$.*

Therefore, to get an alphabet-dependent bound for H-LRCs, we need to construct a set with an upper bound on its entropy and such that its size uses the hierarchical locality property to be as large as possible. To achieve this requirement, we modify the construction algorithm used in the proof of the Singleton-type bound in [15].

Lemma 2 *Let $\mathcal{C}$ be an $[n, k, d]$ H-LRC with locality $[(r_1, \delta_1), (r_2, \delta_2)]$ and $\lambda \in \mathbb{Z}_+$ with $0 \leq \lambda \leq k$. Then, there exists a set I_c such that*

$-\; H(I_c) \le \lambda,$

$-\; |I_c| \ge \lambda + \left\lfloor \frac{\lambda}{r_2} \right\rfloor (\delta_2 - 1) + \left\lfloor \frac{\lambda}{r_1} \right\rfloor (\delta_1 - \delta_2).$

Algorithm 1 For the proof of Lemma 2

1: Let $i_1 = 0, i_2 = 0, I = \emptyset$.
2: **while** $\exists$ middle set M_j such that $M_j \nsubseteq I$ and $H(I \cup M_j) \le \lambda$ **do**
3:      **while** $\exists$ local set $L_i \subset M_j$ such that $L_i \nsubseteq I$ **do**
4:          $i_2 = i_2 + 1$
5:          $a_{i_2} = H(I \cup L_i) - H(I)$
6:          $s_{i_2} = a_{i_2} + \delta_2 - 1$
7:          $I = \mathrm{cl}(I \cup L_i)$
8:      **end while**
9:      $i_1 = i_1 + 1$
10:      $s_{i_2} = s_{i_2} - \delta_2 + \delta_1$
11: **end while**
12: Let M_x a middle set with $M_x \nsubseteq I$.
13: **while** $\exists$ local set $L_i \subset M_x$ such that $L_i \nsubseteq I$ and $H(I \cup L_i) \le \lambda$ **do**
14:      $i_2 = i_2 + 1$
15:      $a_{i_2} = H(I \cup L_i) - H(I)$
16:      $s_{i_2} = a_{i_2} + \delta_2 - 1$
17:      $I = \mathrm{cl}(I \cup L_i)$
18: **end while**

Proof We use the construction given in Algorithm 1 to build the set I_c. In the algorithm, a_{i_2} denotes the incremental entropy and s_{i_2} is a lower bound on the incremental size. Indeed if $J \subseteq [n]$ and L_i is a local set such that $L_i \nsubseteq J$ and $H(J \cup L_i) > H(J)$, then $|\mathrm{cl}(J \cup L_i)| - |J| \ge H(J \cup L_i) - H(J) + \delta_2 - 1$. Similarly, if $J \subseteq [n]$ and M_j is a middle set such that $M_j \nsubseteq J$ and $H(J \cup M_j) > H(J)$, then $|\mathrm{cl}(J \cup M_j)| - |J| \ge H(J \cup M_j) - H(J) + \delta_1 - 1$.

Denote by I the set obtained at the end of Algorithm 1 and by i_1 and j_2 the last indices. The total number of local sets visited in this algorithm is $i_2 \ge \lfloor \frac{\lambda}{r_2} \rfloor$ since the union of $\lfloor \frac{\lambda}{r_2} \rfloor$ arbitrary local sets has an entropy less than λ. By the same argument, we have $i_1 \ge \lfloor \frac{\lambda}{r_1} \rfloor$.

Let $s_e = \lambda - H(I)$ and $A \subseteq [n]$ such that $H(A) = s_e$ and $I \cap A = \emptyset$ which is always possible to construct. Finally, let $I_c = I \cup A$. We prove now that I_c has the desired properties.

By construction, we have $H(I_c) = \lambda$. For the size, we have

$$|I_c| = s_e + |I| \ge s_e + \sum_{i=1}^{i_2} s_i$$

$$= s_e + \sum_{i=1}^{i_2} (a_i + \delta_2 - 1) + i_1(\delta_1 - \delta_2)$$

$$= \lambda + i_2(\delta_2 - 1) + i_1(\delta_1 - \delta_2)$$

$$\ge \lambda + \left\lfloor \frac{\lambda}{r_2} \right\rfloor (\delta_2 - 1) + \left\lfloor \frac{\lambda}{r_1} \right\rfloor (\delta_1 - \delta_2).$$

Hence, I_c has the desired entropy and size. $\qquad\square$

Thus, we obtain the following bound for 2-level H-LRCs.

Theorem 1 *Let $\mathcal{C}$ be an $[n, k, d]$ H-LRC over $\mathbb{F}_q$ with locality $[(r_1, \delta_1), (r_2, \delta_2)]$. Then we have*

$$k \leq \min_{\lambda \in \mathbb{Z}_+} \left\{ \lambda + k_{\text{opt}}^{(q)}(n - \nu, d) \right\} \tag{4}$$

where $\nu = \lambda + \lfloor \frac{\lambda}{r_2} \rfloor (\delta_2 - 1) + \lfloor \frac{\lambda}{r_1} \rfloor (\delta_1 - \delta_2)$ and $k_{\text{opt}}^{(q)}(n, d)$ is the largest possible dimension of a code of length n, for a given alphabet size q and minimum distance d.

Proof Let $\lambda \in \mathbb{Z}_+$ with $0 \leq \lambda \leq k$. By Lemma 2, there exists a set $I_c \subseteq [n]$ such that $H(I_c) \leq \lambda$ and $|I_c| \geq \nu$. Therefore, we have $k = H(I_c) + k - H(I_c) \leq \lambda + k - H(I_c)$. Now, $k - H(I_c)$ is the dimension of the shortened code $\mathcal{C}/I_c$. By definition of $k_{\text{opt}}^{(q)}$ and by Lemma 1, we have

$$\dim(\mathcal{C}/I_c) \leq k_{\text{opt}}^{(q)}(n - |I_c|, d') \leq k_{\text{opt}}^{(q)}(n - \nu, d).$$

The statement now follows by taking the minimum over all $\lambda \in \mathbb{Z}_+$. $\qquad\square$

Lemma 2 can be seen as a proof of concept that we can modify the algorithms in [15] to obtain an alphabet-dependent bound. Indeed, the algorithm presented here and the one presented in [15] are equivalent in the sense that if $\lambda = k - 1$ and I_c is the set obtained by the algorithm in the proof of Lemma 2, we obtain again the Singleton-type bound (3) via the relation $d \leq n - |I_c|$. This implies that the bound (4) is at least as good as the Singleton-type bound (3).

3.1 Extension of the bound to arbitrary levels of hierarchy

In this part, we extend the alphabet-dependent bound to H-LRCs with h-level hierarchical locality. The definition and notation of these codes is the following.

Definition 5 An $[n, k, d]$ linear code $\mathcal{C}$ is a code with *h-level hierarchical locality* having locality parameters $[(r_1, \delta_1), (r_2, \delta_2), \ldots, (r_h, \delta_h)]$ if for all code symbols $i \in [n]$, there exists a collection of sets $\{L_{1_i}, L_{2_i}, \ldots, L_{h_i}\}$ such that for every $j \in [h]$, we have

1. $i \in L_{j_i} \subseteq [n]$,
2. $H(L_{j_i}) \leq r_j$
3. The minimum distance of $\mathcal{C}|_{L_{j_i}}$ is at least δ_j;
4. $\mathcal{C}|_{L_{j_i}}$ is a code with $(h - j)$-level hierarchical locality having locality parameters $[(r_{j+1}, \delta_{j+1}), (r_{j+2}, \delta_{j+2}), \ldots, (r_h, \delta_h)]$.

A set L_{j_i} for $j \in [h]$ is referred to as a *level-j* set. It was proven in [15, Theorem 3.1] that any $[n, k, d]$ H-LRC with hierarchical locality $[(r_1, \delta_1), \ldots, (r_h, \delta_h)]$ satisfies the following Singleton-type bound

$$d \leq n - k + 1 - \left\lfloor \frac{k-1}{r_h} \right\rfloor (\delta_h - 1) - \sum_{l=1}^{h-1} \left\lfloor \frac{k-1}{r_l} \right\rfloor (\delta_l - \delta_{l+1}). \tag{5}$$

Obtaining an alphabet-dependent bound for these codes is done in the same manner as in the case of $h = 2$.

Lemma 3 *Let $\mathcal{C}$ be an $[n, k, d]$ H-LRC with locality $[(r_1, \delta_1), \ldots, (r_h, \delta_h)]$ and $\lambda \in \mathbb{Z}$ with $0 \leq \lambda \leq k$. Then, there exists a set I_c such that*

– $H(I_c) \leq \lambda$,

Algorithm 2 For the proof of Lemma 3

1: Let $i_1 = 0, i_2 = 0, \ldots, i_h = 0; I = \emptyset$;
2: $M_0 = [n], M_1 = \emptyset, M_2 = \emptyset, \ldots, M_h = \emptyset$.
3: **for** $j \leftarrow 1, h$ **do**
4:     **while** $\exists$ set L_{j_i} such that $L_{j_i} \nsubseteq I$ and $H(I \cup L_{j_i}) \leq \lambda$ **do**
5:         $M_{j-1} = L_{j_i}, M_j = L_{j_i}, l = j$.
6:         **while** $\mathrm{cl}(I \cup L_{j_i}) \neq I$ **do**
7:             **if** $\exists$ a level-l set $L_{l_i} \subseteq M_{l-1}$ such that $L_{l_i} \nsubseteq I$ **then**
8:                 $M_l = L_{l_i}$.
9:                 **if** l equals h **then**
10:                     $i_h = i_h + 1, a_{i_h} = H(I \cup M_h) - H(I), s_{i_h} = a_{i_h} + \delta_h - 1$.
11:                     $I = \mathrm{cl}(I \cup M_h)$
12:                 **else**
13:                     $l = l + 1$
14:                 **end if**
15:             **else**
16:                 $l = l - 1$
17:                 $i_l = i_l + 1, s_{i_h} = s_{i_h} + \delta_l - \delta_{l+1}$.
18:             **end if**
19:         **end while**
20:         **for** $l \leftarrow j, h - 1$ **do**
21:             $i_l = i_l + 1$
22:         **end for**
23:         $s_{i_h} = s_{i_h} + \delta_j - \delta_h$
24:     **end while**
25: **end for**

$$- |I_c| \geq \lambda + \left\lfloor \frac{\lambda}{r_h} \right\rfloor (\delta_h - 1) + \sum_{l=1}^{h-1} \left\lfloor \frac{\lambda}{r_l} \right\rfloor (\delta_l - \delta_{l+1}).$$

Proof We use the construction given in Algorithm 2 to build the set I_c. The basic idea is the same as in the proof of Lemma 2. First, the algorithm identifies the smallest j such that a level-j set L_{j_i} has $H(L_{j_i}) \leq \lambda$. This is important when $\lambda < r_1$ or during the process of the algorithm when $\lambda - H(I)$ becomes small. After this set is found, the algorithm visits recursively the sets L_{l_i} with $l \in [j, h-1]$ such that each of them is contained in the previous one and identifies a level-$(h-1)$ set $L_{(h-1)_i}$. Then, it starts adding to I the level-h sets contained in $L_{(h-1)_i}$. When every symbol of $L_{(h-1)_i}$ has been added to I, it steps back one level, finds a new $L_{(h-1)_{i'}}$ that contains new symbols not in I, and adds them in the same manner using level-h sets. At some point, the algorithm adds to I a level-h set $L_{h_{i''}}$ containing the last remaining symbols of L_{j_i}. The second **while** loop is not satisfied anymore and because the last added set was a level-h set, the algorithm adds one to the count of the visited sets per level between j and $h - 1$. Next, it pursues by identifying another level-j set $L_{j_{i'}}$ satisfying $H(I \cup L_{j_{i'}}) \leq \lambda$ or, if all level-j sets exceed λ, the algorithm searches for valid level-$(j + 1)$ sets.

As in the proof of Lemma 2, a_{i_h} denotes the incremental entropy and s_{i_h} is a lower bound on the incremental size corrected accordingly to the level of the added set. The counters i_j for $j \in [h]$ are the number of level-j sets visited by the algorithm.

When the algorithm terminates, the counters are lower-bounded by

$$i_l \geq \left\lfloor \frac{\lambda}{r_l} \right\rfloor, \quad 1 \leq l \leq h$$

since the union of $\lfloor \frac{\lambda}{r_l} \rfloor$ arbitrary level-l sets L_{j_l} has an entropy less than λ.

Denote by I the set obtained at the end of the algorithm and by $i_1, \ldots, i_h$ the last values of the counters. Let $s_e = \lambda - H(I)$ and $A \subseteq [n]$ such that $H(A) = s_e$ and $I \cap A = \emptyset$. Finally, let $I_c = I \cup A$. We prove now that I_c has the desired properties. By construction, we have $H(I_c) = \lambda$. For the size, we have

$$|I_c| = s_e + |I| \geq s_e + \sum_{i=1}^{i_h} s_i$$

$$= s_e + \sum_{i=1}^{i_h} a_i + \sum_{i=1}^{i_h} (\delta_h - 1) + i_{h-1}(\delta_{h-1} - \delta_h) + i_{h-2}(\delta_{h-2} - \delta_{h-1}) + \ldots + i_1(\delta_1 - \delta_2)$$

$$= \lambda + i_h(\delta_h - 1) + \sum_{l=1}^{h-1} i_l(\delta_l - \delta_{l+1})$$

$$\geq \lambda + \left\lfloor \frac{\lambda}{r_h} \right\rfloor (\delta_h - 1) + \sum_{l=1}^{h-1} \left\lfloor \frac{\lambda}{r_l} \right\rfloor (\delta_l - \delta_{l+1}).$$

Hence I_c has the desired entropy and size. $\qquad\square$

Thus, we obtain the following extension of the bound (4) for h-level H-LRCs.

Theorem 2 *Let $\mathcal{C}$ be an $[n, k, d]$ H-LRC over $\mathbb{F}_q$ with locality $[(r_1, \delta_1), \ldots, (r_h, \delta_h)]$. Then, we have*

$$k \leq \min_{\lambda \in \mathbb{Z}_+} \left\{ \lambda + k_{\text{opt}}^{(q)}(n - \nu, d) \right\} \tag{6}$$

where $\nu = \lambda + \lfloor \frac{\lambda}{r_h} \rfloor (\delta_h - 1) + \sum_{l=1}^{h-1} \lfloor \frac{\lambda}{r_l} \rfloor (\delta_l - \delta_{l+1})$ and $k_{\text{opt}}^{(q)}(n, d)$ is the largest possible dimension of a code of length n, for a given alphabet size q and minimum distance d.

As in the case $h = 2$, for $\lambda = k - 1$ and I_c the set obtained by Algorithm 2, we obtain again the Singleton-type bound (5) via the relation $d \leq n - |I_c|$, which shows that the bound (6) is at least as good as the Singleton-type bound (5). Moreover, the bound (6) yields that H-LRCs achieving any bound on the parameters $[n, k, d]$ only are directly alphabet-optimal H-LRCs by setting $\lambda = 0$.

4 Hierarchical locality of $\mathcal{S}(m) - \mathcal{S}(s)$

In [16], the authors presented four different constructions of linear LRCs with small locality and high availability. The constructions are based on a method developed in [4] where the generator matrix of a code is obtained by deleting specific columns from the generator matrix of a Simplex code. In this section, we are interested in the locality of one particular construction in [16] where the deleted columns form again a Simplex code. This construction is highly combinatorial and yields optimal codes achieving the Griesmer bound. The objective is to describe the locality parameters for $\delta > 2$ and all dimensions. We show that these codes are locally repairable codes for every dimension implying a complete optimization of the number of nodes contacted for repair according to the number of failures. Moreover, using combinatorial techniques, we establish the complete list of possible locality and show how the local sets can be arranged to form a hierarchical locality. Finally, we prove that these codes are optimal LRCs for all localities and alphabet-optimal H-LRCs by the new bound (6).

We start by formally defining the construction of these linear LRCs.

Construction 1 ([16, Construction IV]) Let G_m be an $m \times \frac{q^m-1}{q-1}$ generator matrix of the Simplex code $\mathcal{S}(m)$. Let $G_\mathcal{C}$ be the $m \times \frac{q^m-q^s}{q-1}$ matrix obtained by deleting the $\frac{q^s-1}{q-1}$ columns in G_m that begin with $m-s$ zeros. Notice that these $\frac{q^s-1}{q-1}$ columns generate a code isomorphic to $\mathcal{S}(s)$, the Simplex code of dimension s. Then $G_\mathcal{C}$ generates a linear code $\mathcal{C}$ over $\mathbb{F}_q$ denoted by $\mathcal{S}(m) - \mathcal{S}(s)$.

It was proven in [16, Theorem 14] that the code $\mathcal{S}(m) - \mathcal{S}(s)$ with $m \geq 3$ and $s \in [2, m-1]$ is a $[\frac{q^m-q^s}{q-1}, m, q^{m-1} - q^{s-1}]$ linear LRC over $\mathbb{F}_q$ with locality $(r = 2, \delta = 2)$ if $q > 2$ or if $q = 2$ and $s < m - 1$, and with locality $(r = 3, \delta = 2)$ when $q = 2$ and $s = m - 1$. Moreover, $\mathcal{S}(m) - \mathcal{S}(s)$ achieves the Griesmer bound by [16, Lemma 16]. Notice that the code $\mathcal{S}(m) - \mathcal{S}(m-1)$ is isomorphic to the Reed–Muller code RM$(1, m-1)$.

The following example illustrates Construction 1.

Example 1 Let G_4 and G_2 be the generator matrices of the binary Simplex codes $\mathcal{S}_2(4)$ and $\mathcal{S}_2(2)$ respectively. Then $\mathcal{C} = \mathcal{S}_2(4) - \mathcal{S}_2(2)$ is a binary $[12, 4, 6]$ code generated by the matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

where the shadowed columns 3, 4, and 10 are deleted.

4.1 Locality of $\mathcal{S}(m) - \mathcal{S}(s)$ with dimension $m - 1$

The goal of this subsection is to obtain the locality of $\mathcal{S}(m) - \mathcal{S}(s)$ where the local codes have a dimension of $m - 1$. For this, we make a detour to matroid theory by studying the relation between the hyperplanes of the matroid associated to the Simplex code $\mathcal{S}(m)$ and the hyperplanes of the matroid associated to $\mathcal{S}(m) - \mathcal{S}(s)$. Indeed, the Simplex code has intrinsically a lot of useful combinatorial structures and Construction 1 corresponds to a deletion in matroid theory. Therefore, matroid theory is used here as a tool to understand the closed sets of $\mathcal{S}(m) - \mathcal{S}(s)$ of dimension $m - 1$ and to construct the local set for every code symbol. We start by presenting a lemma that gives the relation between a flat and the hyperplanes of the matroid associated to the Simplex code $\mathcal{S}(m)$.

Lemma 4 *Let $M_\mathcal{S}$ be the matroid associated to the Simplex code $\mathcal{S}(m)$ and $Y \in \mathcal{F}(M_\mathcal{S})$ a flat with $\rho(Y) \leq m - 1$. Then, for all hyperplanes $H \in \mathcal{H}(M_\mathcal{S})$ either $Y \subseteq H$ or $H \cap Y \lessdot Y$.*

Proof This claim can be proven using arguments from linear algebra. In fact, for $\mathcal{S}_q(m)$, $M_\mathcal{S}$ can be viewed as the matroid associated to the projective geometry $PG(m-1, q)$. Therefore, every element of $M_\mathcal{S}$ spans a 1-subspace of $\mathbb{F}_q^m$ and the rank of a flat is the dimension of the subspace spanned by the elements of this flat.

Assume now that $Y \not\subseteq H$. Using the familiar vector space identity, we have

$$\rho(H \cap Y) = \dim(\text{span}(H \cap Y)) = \dim(\text{span}(Y)) + \dim(\text{span}(H)) - \dim(\text{span}(H \cup Y))$$
$$= \dim(\text{span}(Y)) + m - 1 - m$$
$$= \rho(Y) - 1.$$

Hence, we have indeed that $H \cap Y \lessdot Y$. $\qquad \square$

Because the Simplex code $\mathcal{S}(s)$ in Construction 1 is a closed set, we can use the previous lemma applied to $\mathcal{S}(s)$ to understand the hyperplanes of $\mathcal{S}(m) - \mathcal{S}(s)$ via the hyperplanes of $\mathcal{S}(m)$ and the deletion corresponding to removing the columns in the generator matrix of $\mathcal{S}(m)$.

Proposition 3 *Let $M_{\mathcal{S}} = (E_{\mathcal{S}}, \rho_{\mathcal{S}})$ be the matroid associated to the Simplex code $\mathcal{S}(m)$ and $M_{\mathcal{C}} = (E_{\mathcal{C}}, \rho_{\mathcal{C}})$ the matroid associated to the code $\mathcal{C} = \mathcal{S}(m) - \mathcal{S}(s)$. Let also $Y \subseteq E_{\mathcal{S}}$ be such that $M_{\mathcal{S}} \setminus Y = M_{\mathcal{C}}$. Then, the map*

$$\phi : \{H \in \mathcal{H}(M_{\mathcal{S}}) \text{ with } Y \neq H\} \to \mathcal{H}(M_{\mathcal{C}})$$
$$H \mapsto H - H \cap Y$$

is a bijection.

Proof By construction, the image of ϕ is a subset of $2^{E_{\mathcal{C}}}$. Firstly, we prove that $\mathcal{H}(M_{\mathcal{C}})$ is contained in the image of ϕ. Secondly, we prove that ϕ is well-defined which implies with the first part that ϕ is a surjection. Finally, we prove that ϕ is an injection.

The flats of $M_{\mathcal{C}}$ can be obtained by the flats of $M_{\mathcal{S}}$ via the relation

$$\mathcal{F}(M|(E - Y)) = \{F \cap (E - Y) : F \in \mathcal{F}(M)\}$$

and $F \cap (E - Y) = F - Y = F - F \cap Y$.

Therefore, we have that $\phi(H) \in \mathcal{F}(M_{\mathcal{C}})$. Moreover, since $\rho_{\mathcal{C}}(E_{\mathcal{C}}) = m$, all hyperplanes of $M_{\mathcal{C}}$ have a rank equal to $m - 1$. Combining this with the fact that $\rho_{\mathcal{C}}(F - F \cap Y) = \rho_{\mathcal{S}}(F - F \cap Y) \leq \rho_{\mathcal{S}}(F)$ implies that if $H_{\mathcal{C}} \in \mathcal{H}(M_{\mathcal{C}})$, then there exists $H \in \mathcal{H}(M_{\mathcal{S}})$ such that $H_{\mathcal{C}} = H - H \cap Y$. Thus $\mathcal{H}(M_{\mathcal{C}})$ is contained in the image of ϕ.

Let $H \in \mathcal{H}(M_{\mathcal{S}})$ with $H \neq Y$. To show that ϕ is well-defined, it is enough to prove that $\rho_{\mathcal{C}}(\phi(H)) = m - 1$, since $\phi(H) \in \mathcal{F}(M_{\mathcal{C}})$. Let $M_H = M_{\mathcal{S}}|H$. By restriction and deletion definitions, we have

$$\rho_{\mathcal{C}}(\phi(H)) = \rho_{\mathcal{S}}(\phi(H)) = \rho_{M_H}(\phi(H)).$$

Since the restricted code $\mathcal{S}(m)|_H$ is isomorphic to the Simplex code $\mathcal{S}(m-1)$, its minimum distance is equal to q^{m-2}. One equivalent definition of the minimum distance states that a code of length n and dimension k has a minimum distance d if $n - d$ is the maximal number of columns of the generator matrix of the code that span a $(k-1)$-dimensional vector subspace. By rephrasing this definition in terms of the matroid M_H, its rank ρ_{M_H}, and the minimum distance $d_{M_H} = q^{m-2}$, we obtain that

$$d_{M_H} = \min\{|A| : A \subseteq H \text{ and } \rho_{M_H}(H - A) < \rho_{M_H}(H)\}.$$

We distinguish two cases. Assume that $s < m - 1$. By construction of $\mathcal{C}$ and $M_{\mathcal{C}}$, the size of Y is $|Y| = \frac{q^s - 1}{q - 1}$. Since $|H \cap Y| \leq |Y| = \frac{q^s - 1}{q - 1} < q^{m-2} = d_{M_H}$, we have $\rho_{M_H}(H - H \cap Y) = \rho_{M_H}(H) = m - 1$. Thus, $\rho_{\mathcal{C}}(\phi(H)) = m - 1$.

Assume now that $s = m - 1$. Then, there is a unique hyperplane of $M_{\mathcal{S}}$ that contains Y which is Y itself. By Lemma 4, we have $H \cap Y \lessdot Y$. Since the restricted code $\mathcal{S}(m)|_{H \cap Y}$

is isomorphic to the Simplex code $S(m-2)$, we have $|H \cap Y| = \frac{q^{m-2}-1}{q-1} < q^{m-2}$. By the same argument as before, we have

$$\rho_{\mathcal{C}}(\phi(H)) = \rho_{M_H}(H - H \cap Y) = \rho_{M_H}(H) = m - 1.$$

Hence, ϕ is well-defined and since the image of ϕ contains $\mathcal{H}(M_{\mathcal{C}})$, ϕ is a surjection.

To prove the injection, suppose that $\phi(H_1) = \phi(H_2)$ for $H_1, H_2 \in \mathcal{H}(M_{\mathcal{S}})$ with $Y \neq H_1, H_2$. Then, $H_1 - H_1 \cap Y = H_2 - H_2 \cap Y$. Now we look at these sets in $M_{\mathcal{S}}$. Since we showed that $\rho_{\mathcal{S}}(H_1 - H_1 \cap Y) = m - 1$, we have that $\mathrm{cl}(H_1 - H_1 \cap Y) = H_1$ because H_1 is a hyperplane containing $H_1 - H_1 \cap Y$. Similarly, we have $\mathrm{cl}(H_2 - H_2 \cap Y) = H_2$. Hence $H_1 = H_2$ and ϕ is an injection. □

The map in Proposition 3 gives us the relation between the hyperplanes of $M_{\mathcal{C}}$ and the hyperplanes of the matroid $M_{\mathcal{S}}$. We can now completely describe the restrictions of $\mathcal{C}$ to hyperplanes and see that these restrictions are in fact isomorphic to certain codes obtained by Construction 1.

Proposition 4 *Let $\mathcal{C}$ be the code $\mathcal{C} = S(m) - S(s)$, $M_{\mathcal{C}}$ the matroid associated to $\mathcal{C}$, and $H_{\mathcal{C}} \in \mathcal{H}(M_{\mathcal{C}})$. Then $\mathcal{C}|_{H_{\mathcal{C}}}$ is either isomorphic to the code $S(m-1) - S(s-1)$ or to $S(m-1) - S(s)$.*

Proof Let $M_{\mathcal{S}}$ be the matroid associated with $S(m)$ and $Y \in \mathcal{F}(M_{\mathcal{S}})$ such that $M_{\mathcal{C}} = M_{\mathcal{S}} \setminus Y$. By Proposition 3, there exists $H \in \mathcal{H}(M_{\mathcal{S}})$ such that $H_{\mathcal{C}} = H - H \cap Y$. We prove the isomorphism by distinguishing two cases depending on Lemma 4.

Suppose that $Y \subseteq H$. By restriction properties, we have

$$M_{\mathcal{C}}|H_{\mathcal{C}} = M_{\mathcal{S}} \setminus Y|H_{\mathcal{C}} = M_{\mathcal{S}} \setminus Y|(H - Y) = M_{\mathcal{S}}|H \setminus Y.$$

By construction of the Simplex code, the matroid $M_{\mathcal{S}}|H$ is isomorphic to $M(S(m-1))$, the matroid associated to the Simplex code $S(m-1)$. Since $Y \in \mathcal{F}(M_{\mathcal{S}}|H)$, we have that $(M_{\mathcal{S}}|H)|Y = M_{\mathcal{S}}|Y$ is isomorphic to $M(S(s))$.

Let E_H and $E_{S(m-1)}$ be the ground sets of $M_{\mathcal{S}}|H$ and $M(S(m-1))$ respectively. Then, there exists a bijection $\varphi : E_H \to E_{S(m-1)}$ that preserves the rank. Let $G_{S(m-1)}$ be a matrix associated to $M(S(m-1))$, $\{\mathbf{g}_i : i \in E_{S(m-1)}\}$ the set of columns of $G_{S(m-1)}$ and $\hat{Y} = \{\varphi(e) : e \in Y\}$. Then $M_{\mathcal{S}}|H \setminus Y$ is isomorphic to $M(G_{S(m-1)} - \{\mathbf{g}_i : i \in \hat{Y}\})$. Since the submatrix formed by the columns indexed in $\hat{Y}$ has a rank equal to s, we can perform some suitable row operations on $G_{S(m-1)}$ to transform the columns $\mathbf{g}_i$ with $i \in \hat{Y}$ such that they are of the form $\tilde{\mathbf{g}}_i = (0, 0, \ldots, 0, \tilde{g}_{i_1}, \ldots, \tilde{g}_{i_s}) \in \mathbb{F}_q^{m-1}$. Let $\tilde{G}_{S(m-1)}$ be the matrix obtained after the row operations. Then the matrix $\tilde{G}_{S(m-1)} - \{\tilde{\mathbf{g}}_i : i \in \hat{Y}\}$ is exactly the generator matrix of the code $S(m-1) - S(s)$. By the previous isomorphism, we have $M(\tilde{G}_{S(m-1)} - \{\tilde{\mathbf{g}}_i : i \in \hat{Y}\}) \cong M(G_{S(m-1)} - \{\mathbf{g}_i : i \in \hat{Y}\}) \cong M_{\mathcal{S}}|H \setminus Y \cong M_{\mathcal{C}}|H_{\mathcal{C}}$. Hence we have indeed that $\mathcal{C}|_{H_{\mathcal{C}}}$ is isomorphic to $S(m-1) - S(s)$.

For the other case, assume now that $Y \not\subseteq H$. By Lemma 4, we have that $H_{\mathcal{C}} = H - H \cap Y$ with $H \cap Y \subsetneq Y$. Now this case follows the previous case in a similar manner by replacing Y by $H \cap Y$ and s by $s - 1$ since $M_{\mathcal{S}}|(H \cap Y)$ is isomorphic to $S(s-1)$. Therefore the same type of isomorphisms yields that $\mathcal{C}|_{H_{\mathcal{C}}}$ is isomorphic to $S(m-1) - S(s-1)$. □

It remains to show the existence of such closed sets for every code symbol in order to prove that the code $S(m) - S(s)$ is an LRC with locality obtained by restrictions to closed sets of dimension $m - 1$.

Theorem 3 *Let C be the linear code $C = S(m) - S(s)$ of length n with $m \geq 2$ and $s \geq 0$. Then, for all $e \in [n]$ we have the following.*

- *If $s = 0$, there is a set $H \subseteq [n]$ containing e such that $C|_H$ is isomorphic to $S(m-1)$.*
- *If $s = m - 1$, there is a set $H \subseteq [n]$ containing e such that $C|_H$ is isomorphic to $S(m-1) - S(m-2)$.*
- *If $1 \leq s \leq m - 2$, there exist two sets H_1, H_2 containing e with $H_1 \not\subseteq H_2$ such that $C|_{H_1}$ is isomorphic to $S(m-1) - S(s)$ and $C|_{H_2}$ is isomorphic to $S(m-1) - S(s-1)$.*

Theorem 3 shows that every coordinate of C is contained in a restriction isomorphic to a Simplex code or to a code obtained by Construction 1. To prove that C is an LRC, it only remains to show that the minimum distance of the restrictions is greater than 1. This is done later in Corollary 1 where we compute the minimum distance for many restrictions of different dimensions.

Proof Let $M_S = (E_S, \rho_S)$ be the matroid associated with the Simplex code $S(m)$ and $Y \subseteq E_S$ such that $M_C = M_S \setminus Y$. Let $e \in E_S \setminus Y$. The general idea of the proof is the following. First, we construct a specific hyperplane in $\mathcal{H}(M_S)$ containing e. Secondly, we use Proposition 3 to get a hyperplane of M_C. Finally, we apply Proposition 4 to this hyperplane to obtain the isomorphism.

If $s = 0$, then the code C is the Simplex code $S(m)$ and there is therefore a hyperplane $H \in \mathcal{H}(M_C)$ containing e such that $C|_H$ is isomorphic to $S(m-1)$.

If $s = m - 1$, then $Y \in \mathcal{H}(M_S)$ and is the only hyperplane containing Y. We know that there exists at least one hyperplane in $\mathcal{H}(M_S)$ that contains e. Let H_S be such a hyperplane. By Proposition 3, $H := H_S - H_S \cap Y$ is a hyperplane of M_C containing e. Applying Proposition 4 yields that $C|_H$ is isomorphic to $S(m-1) - S(s-1)$ since Y is not contained in H_S.

Suppose now that $1 \leq s \leq m - 2$. Since $\rho_S(Y \cup \{e\}) \leq m - 1$, there is a hyperplane $H_{Y,e}$ in $\mathcal{H}(M_S)$ that contains $Y \cup e$. Therefore, Proposition 3 and Proposition 4 yield that $H_1 := H_{Y,e} - Y$ is a hyperplane of M_C containing e and $C|_{H_1}$ is isomorphic to $S(m-1) - S(s)$.

Let $X \in \mathcal{F}(M_S)$ such that $X \lessdot Y$ and let $X_e = \mathrm{cl}(X \cup \{e\}) \in \mathcal{F}(M_S)$. By the hyperplanes property, we have $X_e = \bigcap \{H_S : X_e \subset H_S \text{ and } H_S \in \mathcal{H}(M_S)\}$. Thus, there exists $H_{X,e} \in \mathcal{H}(M_S)$ such that $X_e \subseteq H_{X,e}$ and $Y \not\subseteq H_{X,e}$. Applying Proposition 3 and Proposition 4 yields that $H_2 := H_{X,e} - H_{X,e} \cap Y = H_{X,e} - X$ is a hyperplane of M_C containing e and $C|_{H_2}$ is isomorphic to $S(m-1) - S(s-1)$. □

Theorem 3 can be seen as showing the existence of certain hyperplanes while Proposition 4 is of the form of a uniqueness statement on the parameters size, dimension and minimum distance of the hyperplanes. Therefore, the two combined with Proposition 3 yield the complete characterization of all the hyperplanes of the matroid associated to $S(m) - S(s)$ and thus the characterization of all restrictions of $S(m) - S(s)$ to closed sets of dimension $m - 1$.

4.2 Weight enumerator of $S(m) - S(s)$

Before we continue deriving the localities with dimension less than $m - 1$, the results developed so far allow us to compute the weight enumerator of the codes obtained by Construction 1. For this, we use a theorem from [10] that links the hyperplanes and the codeword supports.

Theorem 4 ([10, Theorem 9.2.4]) *For each linear code C, the hyperplanes of M_C are precisely the complements of the minimal non-empty codeword supports of C.*

In order to compute the weight enumerator of $\mathcal{C} = \mathcal{S}(m) - \mathcal{S}(s)$, the idea is to associate the codewords of $\mathcal{C}$ with the codewords of $\mathcal{S}(m)$ and the hyperplanes of $M_{\mathcal{S}}$. Then, we can use Lemma 4 to understand the effect of the puncturing of $\mathcal{S}(m)$ on the hyperplanes of $M_{\mathcal{S}}$. We start by a lemma that links the codewords of $\mathcal{S}(m)$ and the hyperplanes of $M_{\mathcal{S}}$.

Lemma 5 *Let $\mathcal{S}_q(m)$ be the Simplex code of dimension m over $\mathbb{F}_q$ and $M_{\mathcal{S}}$ the associated matroid. Define $\sim$ as the equivalence relation on the non-zero codewords of $\mathcal{S}(m)$ given by $\mathbf{c} \sim \mathbf{c}'$ if $\mathbf{c} = a\mathbf{c}'$ with $a \in \mathbb{F}_q^*$. Then the map*

$$\varphi : \{\mathbf{c} : \mathbf{c} \in \mathcal{S}(m), \mathbf{c} \neq \mathbf{0}\}/\sim \;\to\; \mathcal{H}(M_{\mathcal{S}})$$
$$[\mathbf{c}] \mapsto [n] - \text{supp}(\mathbf{c})$$

is a bijection.

Proof It is clear that $\sim$ is indeed an equivalence relation. Furthermore, the map is well-defined since all multiples of a codewords share the same support. Then, Theorem 4 implies that φ is a surjection. Now, there are $\frac{q^m - 1}{q - 1}$ equivalence classes of codewords and $\begin{bmatrix} m \\ m-1 \end{bmatrix}_q = \frac{q^m - 1}{q - 1}$ hyperplanes in $M_{\mathcal{S}}$ since they are exactly the linear spaces of dimension $m - 1$ in $\mathbb{F}_q^m$. Hence φ is a bijection. $\qquad\square$

We can now state the formula for the weight enumerator of $\mathcal{C}$.

Theorem 5 *The weight enumerator of the code $\mathcal{C} = \mathcal{S}_q(m) - \mathcal{S}_q(s)$ over $\mathbb{F}_q$ is*

$$W_{\mathcal{C}}(x, y) = x^{\frac{q^m - q^s}{q - 1}} + (q^m - q^{m-s})x^{\frac{q^{m-1} - q^{s-1}}{q - 1}} y^{q^{m-1} - q^{s-1}} + (q^{m-s} - 1)x^{\frac{q^{m-1} - q^s}{q - 1}} y^{q^{m-1}}.$$

Proof Let $\mathbf{c} \in \mathcal{C}$ be a non-zero codeword of $\mathcal{C}$ and $Y \subset [n]$ such that $\mathcal{S}(m)|_{[n]-Y} = \mathcal{C}$, where n is the length of $\mathcal{S}(m)$. Since both codes $\mathcal{C}$ and $\mathcal{S}(m)$ have the same dimension, there is a bijection $\pi : \{\mathbf{c}' : \mathbf{c}' \in \mathcal{S}(m)\} \to \{\mathbf{c} : \mathbf{c} \in \mathcal{C}\}$ given by $\pi(\mathbf{c}') = \mathbf{c}'_{|[n]-Y}$. Let $\mathbf{c}' \in \mathcal{S}(m)$ be such that $\pi(\mathbf{c}') = \mathbf{c}$. We have

$$\text{wt}(\mathbf{c}) = \text{wt}(\pi(\mathbf{c}')) = \text{wt}(\mathbf{c}'_{|[n]-Y}) = \text{wt}(\mathbf{c}') - \text{wt}(\mathbf{c}'_{|Y}).$$

Let $H \in \mathcal{H}(M_{\mathcal{S}})$ be the hyperplane obtained by $H = \varphi(\mathbf{c}')$ in Lemma 5. Then, $\text{wt}(\mathbf{c}') = |\text{supp}(\mathbf{c}')| = n - |H|$ and $\text{wt}(\mathbf{c}'_{|Y}) = |Y| - |H \cap Y|$, since H is the set of coordinates where $\mathbf{c}'_i = 0$ for $i \in H$. Hence, if $n_{\mathcal{C}}$ denotes the length of $\mathcal{C}$, we obtain

$$\text{wt}(\mathbf{c}) = n_{\mathcal{C}} - |H - H \cap Y|.$$

This shows that the weight of $\mathbf{c}$ can be computed from the hyperplanes of $\mathcal{S}(m)$ and their relation with Y. By Lemma 4, the hyperplanes of $M_{\mathcal{S}}$ split into two disjoint sets depending on whether they contain Y. We consider these two cases separately.

Suppose first that $Y \subset H$. Then, we have

$$\text{wt}(\mathbf{c}) = n_{\mathcal{C}} - |H| + |Y|$$
$$= \frac{q^m - q^s}{q - 1} - \frac{q^{m-1} - 1}{q - 1} + \frac{q^s - 1}{q - 1}$$
$$= q^{m-1}.$$

Now, the number of different hyperplane containing Y in $M_{\mathcal{S}}$ is $\begin{bmatrix} m-s \\ m-s-1 \end{bmatrix}_q = \frac{q^{m-s} - 1}{q - 1}$. By Lemma 5, each hyperplanes yields a different equivalence class of codewords. Since there

are $q - 1$ codewords in each class, the number of codewords of weight q^{m-1} is equal to $q^{m-s} - 1$.

Suppose now that $Y \nsubseteq H$. By Lemma 4 and the flats of the Simplex code, we know that $|H \cap Y| = \frac{q^{s-1}-1}{q-1}$. Then, we have

$$\text{wt}(\mathbf{c}) = n_\mathcal{C} - |H| + |H \cap Y|$$
$$= \frac{q^m - q^s}{q-1} - \frac{q^{m-1} - 1}{q-1} + \frac{q^{s-1} - 1}{q-1}$$
$$q = q^{m-1} - q^{s-1}.$$

The number of hyperplanes not containing Y is $\begin{bmatrix} m \\ m-1 \end{bmatrix}_q - \begin{bmatrix} m-s \\ m-s-1 \end{bmatrix}_q = \frac{q^m - q^s}{q-1}$. By the same previous argument, the number of codewords of weight $q^{m-1} - q^{s-1}$ is then equal to $q^m - q^s$ and this concludes the proof. $\qquad\square$

### 4.3 The complete locality of $\mathcal{S}(m) - \mathcal{S}(s)$

In this subsection, we describe the restrictions of $\mathcal{C} = \mathcal{S}(m) - \mathcal{S}(s)$ of dimension less than $m - 1$ to obtain the rest of the possible localities. The main observation is that Theorem 3 was carefully written as a recursive statement on the same type of construction already considered, i.e., a Simplex code removed from another Simplex code. Therefore, we can apply Theorem 3 again on the restriction $\mathcal{C}|_H$ to obtain, up to isomorphism, the restricted codes of dimension $m - 2$ of $\mathcal{C}|_H$ and thus of $\mathcal{C}$ as well. Moreover, since every code symbol is contained in a restriction of $\mathcal{C}$ of dimension $m - 1$, applying Theorem 3 again on the restricted codes implies that every code symbol is also contained in a restriction of $\mathcal{C}$ of dimension $m - 2$. This is crucial when considering the locality of $\mathcal{C}$. We start with an example.

Example 2 Let $\mathcal{C} = \mathcal{S}_2(4) - \mathcal{S}_2(2)$ be the $[12, 4, 6]$ binary linear code of Example 1. Since $s < m - 1$, Theorem 3 implies that for all code symbols $e \in [n]$, there exist two sets $H_1, H_2 \subseteq [12]$ containing e such that $\mathcal{C}|_{H_1}$ is isomorphic to $\mathcal{S}_2(3) - \mathcal{S}_2(1)$ and $\mathcal{C}|_{H_2}$ is isomorphic to $\mathcal{S}_2(3) - \mathcal{S}_2(2)$. In other words, there are two restrictions $\mathcal{C}|_{H_1}, \mathcal{C}|_{H_2}$ containing e with parameters $[6, 3, 3]$ and $[4, 3, 2]$ respectively. Therefore, $\mathcal{C}$ is an LRC with locality $(r = 4, \delta = 3)$ and also an LRC with locality $(r = 3, \delta = 2)$. Notice that $r = 4$ for the first locality even if the dimension of the restricted codes is equal to 3. This is due to the fact that r must satisfy $|R_i| \leq r + \delta - 1$ as opposed to the definition of H-LRCs where $H(R_i) \leq r_1$.

If we apply Theorem 3 to $\mathcal{S}_2(3) - \mathcal{S}_2(1)$, we obtain by isomorphism that there exist $H_3, H_4 \subseteq H_1$ containing e such that $\mathcal{C}|_{H_3}$ is isomorphic to $\mathcal{S}_2(2) - \mathcal{S}_2(1)$ and $\mathcal{C}|_{H_4}$ is isomorphic to $\mathcal{S}_2(2)$. $\mathcal{S}_2(2) - \mathcal{S}_2(1)$ is a $[2, 2, 1]$ code and thus does not provide an extra locality. On the other hand, $\mathcal{S}_2(2)$ is a $[3, 2, 2]$ code which implies that $\mathcal{C}$ is also an LRC with locality $(r = 2, \delta = 2)$. Furthermore, by construction of the local sets, $\mathcal{C}$ is an H-LRC with locality $[(3, 3), (2, 2)]$.

Example 2 illustrates how Theorem 3 can be used to obtain the locality for different dimensions. Moreover, because of the recursive form of Theorem 3, the local sets can be arranged in such a way that we obtain a hierarchical locality. We break down what happens to the restrictions when we iterate Theorem 3 by considering the restriction types, i.e., the different isomorphic restrictions. Suppose for simplicity that m is sufficiently large and s is close to half of m. As illustrated in Fig. 1, applying Theorem 3 on the two restriction types of dimension $m - 1$ gives three new restriction types of dimension $m - 2$, as two of them lead to the same isomorphic code.

Fig. 1 Illustration of the top level locality of the code $\mathcal{S}(m) - \mathcal{S}(s)$ after appropriate isomorphisms
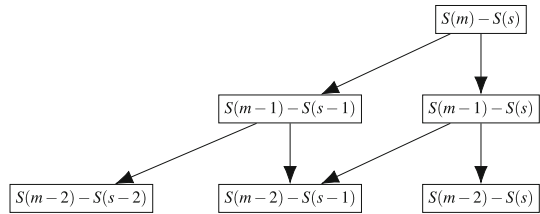
Suppose now that, after some iterations of Theorem 3, we obtain the restriction types of dimension κ. Let $a, b \in \mathbb{Z}_+$ such that all restriction types are of the form $\mathcal{S}(\kappa) - \mathcal{S}(i)$ with $a \leq i \leq b$. Now, the restriction types of dimension $\kappa - 1$ can be obtained by applying Theorem 3 on the restriction types of dimension κ. Two extremal cases need to be taken into account. If $a = 0$, the only restriction type that we get from $\mathcal{S}(\kappa) - \mathcal{S}(a)$ is $\mathcal{S}(\kappa - 1)$. If $b = \kappa - 1$, the only restriction type is $\mathcal{S}(\kappa - 1) - \mathcal{S}(\kappa - 2)$. This is illustrated in Fig. 2 where the two dashed boxes represent the conditional new restriction types that exist only if $a > 0$ or $b < \kappa - 1$. This sketches the high level idea of the proof of the next theorem describing precisely the different restriction types for a given dimension κ.

Theorem 6 *Let $\mathcal{C} = \mathcal{S}(m) - \mathcal{S}(s)$ with $m \geq 3$ and $0 \leq s \leq m - 1$. Let $\kappa \in [2, m - 1]$ and i be an integer such that $\max\{0, s - m + \kappa\} \leq i \leq \min\{s, \kappa - 1\}$. Then for all code symbols $e \in [n]$, there is a set $F_i \subseteq [n]$ containing e such that $\mathcal{C}|_{F_i}$ is isomorphic to $\mathcal{S}(\kappa) - \mathcal{S}(i)$.*

Remark 1 Notice that $\max\{0, s - m + \kappa\} \leq \min\{s, \kappa - 1\}$. Indeed we have $\min\{s, \kappa - 1\} \geq 0$ since $\kappa \geq 3$ and $s \geq 0$. Since $\kappa \leq m - 1$ and $s \leq m - 1$, then $s - m + \kappa \leq s - 1$ and $s - m + \kappa \leq \kappa - 1$. Thus, $\min\{s, \kappa - 1\} \geq s - m + \kappa$ and we have that $\max\{0, s - m + \kappa\} \leq \min\{s, \kappa - 1\}$. Therefore, the claim of Theorem 6 is that there always exist such restrictions for all dimensions $\kappa \in [2, m - 1]$.

Proof We denote by I_κ the set $I_\kappa = \{i \in \mathbb{Z}_+ : \max\{0, s - m + \kappa\} \leq i \leq \min\{s, \kappa - 1\}\}$. Let also $e \in [n]$ be an arbitrary code symbol.

The case $\kappa = m - 1$ is exactly Theorem 3 with the set I_{m-1} being equal to $\{0\}$ when $s = 0$, $I_{m-1} = \{m - 2\}$ if $s = m - 1$, and $I_{m-1} = \{s - 1, s\}$ if $1 \leq s \leq m - 2$.

We now prove the theorem by doing a reverse induction on κ as long as $\kappa \geq 3$. Assume that the claim is true for κ, i.e., for all $j \in I_\kappa$ and for all code symbol $e \in [n]$, there exists a set $W_j \subseteq [n]$ containing e such that $\mathcal{C}|_{W_j}$ is isomorphic to $\mathcal{S}(\kappa) - \mathcal{S}(j)$. The goal is to construct the restrictions of dimension $\kappa - 1$.

Let $b = \min\{s, \kappa - 1\}$. By the induction hypothesis, there exists a set $W_b \subseteq [n]$ containing e such that $\mathcal{C}|_{W_b}$ is isomorphic to $\mathcal{S}(\kappa) - \mathcal{S}(b)$.

If $b = \kappa - 1$ then applying Theorem 3 on $\mathcal{C}' = \mathcal{S}(\kappa) - \mathcal{S}(\kappa - 1)$ yields that for all symbols $e' \in [n']$, with n' being the length of the code $\mathcal{C}'$, there is a set $F \subseteq [n']$ containing e' such that $\mathcal{C}'|_F$ is isomorphic to $\mathcal{S}(\kappa - 1) - \mathcal{S}(\kappa - 2)$. Therefore, there exists a set $F_{b-1} \subseteq [n]$

Fig. 2 Sketch of an iteration of Theorem 3 on the different restrictions types

containing e such that $\mathcal{C}|_{F_{b-1}}$ is isomorphic to $\mathcal{S}(\kappa - 1) - \mathcal{S}(\kappa - 2)$ since we have that $\mathcal{C}|_{F_{b-1}} = (\mathcal{C}|_{W_b})|_{F_{b-1}}$. Since $b = \kappa - 1$, then we have that $\kappa - 2 = b - 1 = \min\{s, \kappa - 2\}$.

If $b = s < \kappa - 1$, then the third case of Theorem 3 applied to $\mathcal{S}(\kappa) - \mathcal{S}(s)$ and the corresponding isomorphisms yields that there exists a set F_b such that $\mathcal{C}|_{F_b}$ is isomorphic to $\mathcal{S}(\kappa) - \mathcal{S}(b)$. Moreover, since $b < \kappa - 1$, then $b = \min\{s, \kappa - 2\}$.

Let $j \in [\max\{0, s - m + \kappa\}, \min\{s, \kappa - 1\} - 1]$. Notice that this interval might be empty for example if $s = m - 1$. If the interval is not empty, let W_j the set containing e by the induction hypothesis such that $\mathcal{C}|_{W_j}$ is isomorphic to $\mathcal{S}(\kappa) - \mathcal{S}(j)$. Since by definition of the interval, we have that $j < \kappa - 1$, we can apply Theorem 3 to obtain a set $F_i \subseteq [n]$ containing e such that $\mathcal{C}|_{F_i}$ is isomorphic to $\mathcal{S}(\kappa - 1) - \mathcal{S}(j)$.

Finally, if $\max\{0, s - m + \kappa\} > 0$, then let $a = s - m + \kappa$ and W_a containing e such that $\mathcal{C}|_{W_a}$ is isomorphic to $\mathcal{S}(\kappa) - \mathcal{S}(a)$. Since $a \geq 1$, we can use the third case of Theorem 3 applied to $\mathcal{S}(\kappa) - \mathcal{S}(a)$ to obtain a set $F_{a-1} \subseteq [n]$ such that $\mathcal{C}|_{F_{a-1}}$ is isomorphic to $\mathcal{S}(\kappa - 1) - \mathcal{S}(a - 1)$. Moreover, we have $a - 1 = \max\{0, s - m + \kappa - 1\}$.

Hence, we have constructed a set for e isomorphic to $\mathcal{S}(\kappa - 1) - \mathcal{S}(i)$ for all i in the set

$$[\max\{0, s - m + \kappa\}, \min\{s, \kappa - 1\} - 1]$$
$$\cup \min\{s, \kappa - 2\} \cup \max\{0, s - m + \kappa - 1\}$$
$$= [\max\{0, s - m + \kappa - 1\}, \min\{s, \kappa - 2\}]$$
$$= I_{\kappa - 1}.$$

Since e was an arbitrary symbol, this concludes the proof. $\qquad\square$

As a corollary of Theorem 6, we have that $\mathcal{C}$ is an LRC as long as the restriction types have a minimum distance greater than or equal to 2. We choose to give here the length, dimension and minimum distance of the local codes to avoid confusion on the parameter r and r_1 in the two definitions of LRCs and H-LRCs.

Corollary 1 *Let $\mathcal{C} = \mathcal{S}(m) - \mathcal{S}(s)$ with $m \geq 3$ and $0 \leq s \leq m - 1$. Let $\kappa \in [3, m - 1]$ and i be an integer such that $\max\{0, s - m + \kappa\} \leq i \leq \min\{s, \kappa - 1\}$. Then $\mathcal{C}$ is an LRC with local code parameters*

$$\begin{cases} \left[\frac{q^\kappa - q^i}{q - 1}, \kappa, q^{\kappa - 1} - q^{i - 1}\right] & \text{if } i > 0, \\ \left[\frac{q^\kappa - 1}{q - 1}, \kappa, q^{\kappa - 1}\right] & \text{if } i = 0. \end{cases}$$

Furthermore, for $\kappa = 2$, we have the following local parameters:

- *If $0 \leq s \leq m - 2$, then $\mathcal{C}$ is an LRC with local parameters $[q + 1, 2, q]$*
- *If $q > 2$ and $1 \leq s \leq m - 1$ then $\mathcal{C}$ is an LRC with local parameters $[q, 2, q - 1]$.*

Proof When $\kappa \geq 3$, Theorem 6 guaranties that for every code symbol $e \in [n]$, there is a restriction containing e isomorphic to $\mathcal{S}(\kappa) - \mathcal{S}(i)$. Therefore $\mathcal{C}$ is an LRC code with the parameters of the local sets given by the parameters of $\mathcal{S}(\kappa) - \mathcal{S}(i)$ with a minimum distance greater than or equal to 2.

When $\kappa = 2$, we need to distinguish some cases as the minimum distance might be smaller than 2. Let $I_2 = \{i : \max\{0, s - m + 2\} \leq i \leq \min\{s, 1\}\}$ be the range of the parameter i when $\kappa = 2$. If $s = 0$, then $I_2 = \{0\}$ and by Theorem 6, there is a restriction containing e isomorphic to $\mathcal{S}(2)$. The parameters of the restriction is then $[q + 1, 2, q]$ and the minimum distance is thus greater than or equal to 2.

If $1 \leq s \leq m - 2$, then $s - m + 2 \leq 0$ so $\max\{0, s - m + 2\} = 0$ and $I_2 = \{0, 1\}$. By Theorem 6, every code symbol is contained in a restriction isomorphic to $\mathcal{S}(2)$ and in another restriction isomorphic to $\mathcal{S}(2) - \mathcal{S}(1)$. Since the parameters of $\mathcal{S}(2) - \mathcal{S}(1)$ are $[q, 2, q - 1]$, then the minimum distance is greater than or equal to 2 if and only if $q > 2$.

Finally, if $s = m - 1$, then $s - m + 2 = 1$ and $I_2 = \{1\}$. Thus, $\mathcal{C}$ is an LRC code with locality parameters $[q, 2, q - 1]$ if $q > 2$. $\square$

Notice that when $\kappa = 2$, we obtain again the locality proven in [16]. We explain next why the list of localities of Corollary 1 is the complete list of possible localities with closed local sets. In Corollary 1, the list of different localities is established by removing the restrictions leading to codes with minimum distance 1 from the list of Theorem 6. But the list of Theorem 6 relies on consecutive iterations of Theorem 3, where Proposition 4 can be applied at each iteration guaranteeing the uniqueness of the restriction types. Therefore, Theorem 6 contains the complete list of restriction types and Corollary 1 contains the complete list of possible localities with closed local sets.

We will now look at two special cases of Theorem 6 and Corollary 1. In the first example, we study the case when $s = 0$.

Example 3 Let $\mathcal{C} = \mathcal{S}(m) - \mathcal{S}(0) = \mathcal{S}(m)$ with $m \geq 3$. Let $I_\kappa = \{i : \max\{0, s - m + \kappa\} \leq i \leq \min\{s, \kappa - 1\}\}$ be the range of the parameter i in Theorem 6. Since $s = 0$, we have that $I_\kappa = 0$. Then Theorem 6 and Corollary 1 imply that $\mathcal{C}$ is an LRC code with local parameters $[\frac{q^\kappa - 1}{q - 1}, \kappa, q^{\kappa - 1}]$ for all $\kappa \in [2, m - 1]$. This result was already proven in [7].

For the second example, we look at the case when $s = m - 1$ and $\mathcal{C}$ corresponds to the Reed–Muller code $RM(1, m - 1)$.

Example 4 Let $\mathcal{C} = \mathcal{S}(m) - \mathcal{S}(m - 1)$ with $m \geq 3$ corresponding to the Reed–Muller code $RM(1, m - 1)$. Let I_κ the set $I_\kappa = \{i : \max\{0, s - m + \kappa\} \leq i \leq \min\{s, \kappa - 1\}\}$. Since $s = m - 1$, we have $I_\kappa = \{\kappa - 1\}$. Hence Theorem 6 and Corollary 1 imply that $RM(1, m - 1)$ is an LRC with local parameters $[q^{\kappa - 1}, \kappa, q^{\kappa - 1} - q^{\kappa - 2}]$ for all $\kappa \in [3, m - 1]$ and the local codes are isomorphic to the Reed–Muller code $RM(1, \kappa - 1)$. Furthermore, if $q > 2$, then $RM(1, m - 1)$ is also an LRC with local parameters $[q, 2, q - 1]$. Moreover, since every locality is obtained by restrictions on the previous local sets, we get that the Reed–Muller codes $RM(1, m - 1)$ are H-LRCs with $(m - 3)$-level hierarchy over the binary field and with $(m - 2)$-level hierarchy over $\mathbb{F}_q$ with $q > 2$.

We conclude this section by showing the optimality of $\mathcal{S}(m) - \mathcal{S}(s)$ with respect to the bounds (2) and (6) and present a table containing binary codes obtained by Construction 1 and their hierarchical localities.

Theorem 7 *Let $\mathcal{C} = \mathcal{S}(m) - \mathcal{S}(s)$ with $m \geq 3$ and $0 \leq s \leq m - 1$. Then $\mathcal{C}$ is an optimal LRC for every locality described in Corollary 1. Moreover, $\mathcal{C}$ is an optimal H-LRC with*

$$\begin{cases} (m - 2)\text{-level hierarchical locality} & \text{when } q > 2 \text{ or } q = 2 \text{ and } s < m - 1, \\ (m - 3)\text{-level hierarchical locality} & \text{when } q = 2 \text{ and } s = m - 1. \end{cases}$$

Proof Since $\mathcal{C}$ already achieves the Griesmer bound on the parameters $[n, k, d]$, it will achieves the bound (2) for $\lambda = 0$ and thus be an optimal LRC code for each locality described in Corollary 1.
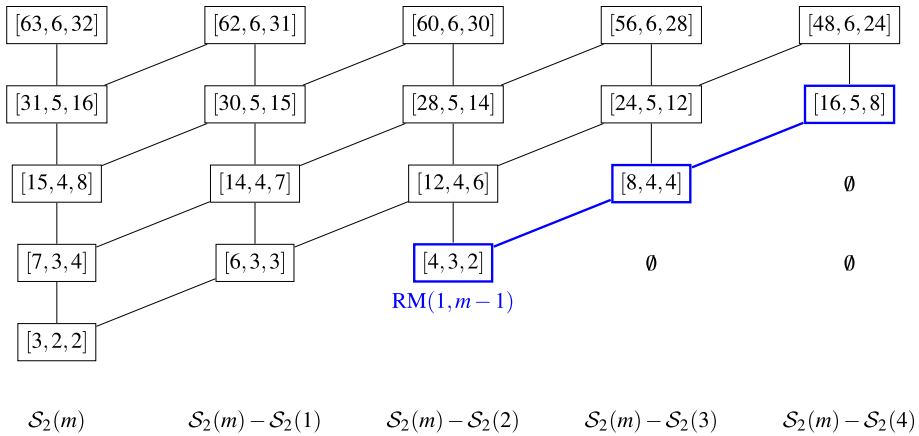
$\mathcal{S}_2(m)$	$\mathcal{S}_2(m)-\mathcal{S}_2(1)$	$\mathcal{S}_2(m)-\mathcal{S}_2(2)$	$\mathcal{S}_2(m)-\mathcal{S}_2(3)$	$\mathcal{S}_2(m)-\mathcal{S}_2(4)$
$[63,6,32]$	$[62,6,31]$	$[60,6,30]$	$[56,6,28]$	$[48,6,24]$
$[31,5,16]$	$[30,5,15]$	$[28,5,14]$	$[24,5,12]$	$[16,5,8]$
$[15,4,8]$	$[14,4,7]$	$[12,4,6]$	$[8,4,4]$	$\emptyset$
$[7,3,4]$	$[6,3,3]$	$[4,3,2]$	$\emptyset$	$\emptyset$
$[3,2,2]$			$\mathrm{RM}(1,m-1)$	

Fig. 3 Table of the complete hierarchical locality for binary codes $\mathcal{S}_2(m)-\mathcal{S}_2(s)$ with $m \in [2,6]$ and $s \in [0,4]$

We exhibit now a particular chain of subsets that yields the hierarchical locality. Vaguely speaking, the chain is obtained by taking the left foremost diagonal restriction types in Figs. 1 and 2 .

Formally, let $e \in [n]$ be an arbitrary code symbol. By Theorem 6, for all $\kappa \in [2, m-1]$, there exists a set $F_\kappa \subseteq [n]$ containing e such that $\mathcal{C}|_{F_\kappa}$ is isomorphic to $\mathcal{S}(\kappa) - \mathcal{S}(\max\{0, s - m + \kappa\})$. Define $F_m = [n]$. Each F_κ was obtained in the proof of Theorem 6 by applying Theorem 3 to $\mathcal{C}|_{F_{\kappa+1}}$. Therefore, we have $F_2 \subseteq F_3 \subseteq \cdots \subseteq F_{m-1} \subseteq [n]$. Moreover, because $\mathcal{C}|_{F_\kappa}$ is isomorphic to $\mathcal{S}(\kappa) - \mathcal{S}(\max\{0, s - m + \kappa\})$, for every $a \in F_\kappa$, there exists such a chain $F_{2,a} \subseteq F_{3,a} \subseteq \cdots \subseteq F_{\kappa-1,a}$ with $a \in F_{l,a}$ and $\mathcal{C}|_{F_{l,a}}$ is isomorphic to $\mathcal{S}(l) - \mathcal{S}(\max\{0, s - m + l\})$ for $l \in [2, \kappa - 1]$.

Now if d is the minimum distance of $\mathcal{C}$, then the minimum distance of $\mathcal{S}(\kappa) - \mathcal{S}(\max\{0, s - m + \kappa\})$ is

$$
\delta_\kappa = \begin{cases} q^{\kappa-1} - q^{s-m+k-1} = \dfrac{d}{q^{m-\kappa}} & \text{if } \kappa > m - s, \\ q^{\kappa-1} & \text{if } \kappa \leq m - s. \end{cases}
$$

Notice that δ_κ is less than 2 if and only if $q = 2$, $s = m - 1$, and $\kappa = 2$ by Corollary 1.

Hence, if $q > 2$ or $q = 2$ and $s < m - 1$, then $\mathcal{C}$ is an H-LRC with $(m-2)$-level hierarchical locality having locality parameters

$$
[(m - 1, \delta_{m-1}), \ldots, (2, \delta_2)].
$$

When $q = 2$ and $s = m - 1$, $\mathcal{C}$ is an H-LRC with $(m-3)$-level hierarchical locality having locality parameters

$$
[(m - 1, \delta_{m-1}), \ldots, (3, \delta_3)].
$$

Finally, $\mathcal{C}$ achieves the new alphabet-dependent bound (6) for H-LRC codes when $\lambda = 0$ and k_{opt}^q is taken to be the Griesmer bound. □

The table in Fig. 3 represents the binary linear codes obtained by the construction $\mathcal{S}_2(m) - \mathcal{S}_2(s)$. The columns are sorted by $s \in [0, 4]$ and the rows are sorted by $m \in [2, 6]$. Moreover, the lines describes the locality of dimension $m - 1$ obtained by Theorem 3. Therefore, if $\mathcal{C}$

and $\mathcal{C}'$ are two codes in the table such that there exists a path from $\mathcal{C}'$ to $\mathcal{C}$, then $\mathcal{C}$ has locality $\mathcal{C}'$, i.e., each symbol of $\mathcal{C}$ is contained in a restriction isomorphic to $\mathcal{C}'$. Figure 3 gives also the hierarchical locality of a binary code via the paths to smaller codes. Finally, the codes in blue are the binary Reed–Muller codes RM$(1, m - 1)$.

5 Conclusion

In this paper, we presented a new alphabet-dependent bound for codes with hierarchical locality. Then, we worked on a class of codes obtained by deleting a Simplex code from another Simplex code of higher dimension. We derived the weight enumerator of these codes and the complete list of possible localities with closed repair sets. Finally, we used this list to show that these codes are optimal LRCs and optimal H-LRCs by the new bound.

References

1. Agarwal A., Barg A., Hu S., Mazumdar A., Tamo I.: Combinatorial alphabet-dependent bounds for locally recoverable codes. IEEE Trans. Inf. Theory **64**, 3481–3492 (2018).
2. Ballentine S., Barg A., Vladuts S.: Codes with hierarchical locality from covering maps of curves. arXiv preprint. arXiv:1807.05473 (2018).
3. Cadambe V., Mazumdar A.: An upper bound on the size of locally recoverable codes. In: International Symposium on Network Coding, pp. 1–5 (2013).
4. Farrell P.: Linear binary anticodes. Electron. Lett. **6**(13), 419–421 (1970).
5. Freij-Hollanti R., Hollanti C., Westerbäck T.: Matroid theory and storage codes: bounds and constructions. In: Network Coding and Subspace Designs, pp. 385–425. Springer (2018).
6. Gopalan P., Huang C., Simitci H., Yekhanin S.: On the locality of codeword symbols. IEEE Trans. Inf. Theory **58**(11), 6925–6934 (2012).
7. Grezet M., Freij-Hollanti R., Westerbäck T., Hollanti C.: Alphabet-dependent bounds for linear locally repairable codes based on residual codes. arXiv preprint. arXiv:1810.08510 (2018).
8. Kamath G.M., Prakash N., Lalitha V., Kumar P.V.: Codes with local regeneration. In: 2013 Information Theory and Applications Workshop (ITA), pp. 1–5 (2013).
9. Kamath G.M., Prakash N., Lalitha V., Kumar P.V., Silberstein N., Rawat A.S., Koyluoglu O.O., Vishwanath S.: Explicit MBR all-symbol locality codes. In: 2013 IEEE International Symposium on Information Theory, pp. 504–508 (2013).
10. Oxley J.: Matroid theory. Oxford University Press, New York (1992).
11. Papailiopoulos D., Dimakis A.: Locally repairable codes. In: International Symposium on Information Theory, pp. 2771–2775. IEEE (2012).

12. Prakash N., Kamath G.M., Lalitha V., Kumar P.V.: Optimal linear codes with a local-error-correction property. In: International Symposium on Information Theory, pp. 2776–2780. IEEE (2012).
13. Rawat A.S., Koyluoglu O.O., Silberstein N., Vishwanath S.: Optimal locally repairable and secure codes for distributed storage systems. IEEE Trans. Inf. Theory **60**, 212–236 (2014).
14. Rawat A.S., Papailiopoulos D.S., Dimakis A.G., Vishwanath S.: Locality and availability in distributed storage. IEEE Trans. Inf. Theory **62**(8), 4481–4493 (2016). https://doi.org/10.1109/tit.2016.2524510.
15. Sasidharan B., Agarwal G.K., Kumar P.V.: Codes with hierarchical locality. In: 2015 IEEE International Symposium on Information Theory (ISIT), pp. 1257–1261 (2015).
16. Silberstein N., Zeh A.: Anticode-based locally repairable codes with high availability. Des. Codes Cryptogr. **86**(2), 419–445 (2018).
17. Tamo I., Barg A.: A family of optimal locally recoverable codes. IEEE Trans. Inf. Theory **60**(8), 4661–4676 (2014).
18. Westerbäck T., Freij-Hollanti R., Ernvall T., Hollanti C.: On the combinatorics of locally repairable codes via matroid theory. IEEE Trans. Inf. Theory **62**, 5296–5315 (2016).
19. Westerbäck T., Grezet M., Freij-Hollanti R., Hollanti C.: On the polymatroidal structure of quasi-uniform codes with applications to heterogeneous distributed storage. In: International Symposium on Mathematical Theory of Networks and Systems (2018).
20. Zeh A., Yaakobi E.: Optimal linear and cyclic locally repairable codes over small fields. In: 2015 IEEE Information Theory Workshop (ITW), pp. 1–5 (2015).