On Optimal Locally Repairable Codes and Generalized Sector-Disk Codes

Han Cai*, and Moshe Schwartz*
*School of Electrical and Computer Engineering, Ben-Gurion University of the Negev,
Beer Sheva 8410501, Israel, hancai@aliyun.com, schwartz@ee.bgu.ac.il

***Abstract*—Optimal locally repairable codes with information locality are considered. Optimal codes are constructed, whose length is also order-optimal with respect to a new bound on the code length derived in this paper. The length of the constructed codes is super-linear in the alphabet size, which improves upon the well known pyramid codes, whose length is only linear in the alphabet size. The recoverable erasure patterns are also analyzed for the new codes. Based on the recoverable erasure patterns, we construct generalized sector-disk (GSD) codes, which can recover from disk erasures mixed with sector erasures in a more general setting than known sector-disk (SD) codes. Additionally, the number of sectors in the constructed GSD codes is super-linear in the alphabet size, compared with known SD codes, whose number of sectors is only linear in the alphabet size.**

I. INTRODUCTION

In the large distributed storage systems of today, an $[n,k]$ storage code encodes k information symbols to n symbols and stores them across n disks in a storage system. Generally speaking, among all storage codes, maximum distance separable (MDS) codes are preferred for practical systems both in terms of redundancy and in terms of reliability. However, as pointed out in [26], MDS codes such as Reed-Solomon codes suffer from a high repair cost. This is mainly because, for an $[n,k]$ MDS code, whenever one wants to recover a symbol, one needs to contact k surviving symbols, which is costly, especially in large-scale distributed file systems.

To improve the repair efficiency, the concept of r-locality for a code $\mathcal{C}$ was initially studied in [14] to ensure that a failed symbol can be recovered by only accessing $r \ll k$ other symbols which form a repair set.

In the past decade, the original definition has been generalized in different aspects. Firstly, to guarantee that the system can recover locally from multiple erasures, the notion of r-locality was generalized to (r,δ)-locality. Secondly, to let code symbols have good availability, the notion of locality has been generalize to (r,δ)-availability [24] (or $(r,\delta)_c$-locality [30]), in which case a code symbol has more than one repair set. Finally, to satisfy differing locality requirements, the notion of locality has been generalized to the hierarchical [25] and the unequal [19], [31] locality cases. For theoretic upper bounds and constructions, the reader may refer to [6], [7], [10], [17], [18], [21], [23], [28] for (r,δ)-locality, [8], [27], [29], [30] for (r,δ)-availability, [25] for hierarchical locality, and [19], [20], [31] for unequal locality.

This work was supported in part by a German Israeli Project Cooperation (DIP) grant under grant no. PE2398/1-1.

Based on the observation given in [13], locally repairable codes may recover from some special erasure patterns beyond their minimum Hamming distance. Thus, another research branch for locally repairable codes is the study of their recoverable erasure patterns. In this aspect, two special kinds of codes have received most of the attention. One is the $(\delta-1,\gamma)$-maximally recoverable code first introduced in [2], [13], that can recover from erasure patterns that include any $\delta-1$ erasures from each repair set, and any other γ erasures. The $(\delta-1,\gamma)$-maximally recoverable codes are equivalent to $(\delta-1,\gamma)$-partial MDS codes – a special kind of array codes that was introduced to improve the storage efficiency of redundant arrays of independent disks (RAIDs) [2]. The other is $(\delta-1,\gamma)$-sector-disk (SD) codes [22], that can recover from erasure patterns that include any $\delta-1$ erasures from each repair set with consistent indices (i.e., whole disk erasures) and any other γ erasures (i.e., sector erasures). For construction of SD codes the reader may refer to [2], [3], [5], [12], [13], [21], [22] for examples. The main drawback of all of the reported constructions for SD codes is the requirement for a large finite field.

In this paper, we focus on both (r,δ)-locality and recoverable erasure patterns beyond the minimum Hamming distance. For (r,δ)-locality we propose constructions of locally repairable codes whose information symbols have (r,δ)-locality and their length is super-linear in the field size. The codes generated by our constructions have new parameters compared with known locally repairable codes. In particular, our codes have a smaller requirement on the field size compared with pyramid codes. Additionally, we consider the following fundamental problem: how long can a locally repairable code be, whose information symbols have (r,δ)-locality? We propose a new upper bound on the length of optimal locally repairable codes. Based on this bound, we prove that the codes generated by our construction may have order-optimal length. We also analyze recoverable erasure patterns beyond the minimum Hamming distance in the codes we construct. Based on this analysis, we construct array codes that can recover special erasure patterns which mix whole disk erasures together with additional sector erasures beyond the minimum Hamming distance. These codes generalize SD codes, and we therefore call them generalized sector-disk (GSD) codes.

Due to space limitations we omit all proofs, which are available in a full version of this work [9].

II. Preliminaries

Throughout this paper, the following notations are used:

- For a positive integer n, let $[n]$ denote the set $\{1, 2, \cdots, n\}$;
- For any prime power q, let $\mathbb{F}_q$ denote the finite field with q elements;
- An $[n, k]_q$ linear code $\mathcal{C}$ over $\mathbb{F}_q$ is a k-dimensional subspace of $\mathbb{F}_q^n$ with a $k \times n$ generator matrix $G = (\mathbf{g}_1, \mathbf{g}_2, \cdots, \mathbf{g}_n)$, where $\mathbf{g}_i$ is a column vector of length k for all $1 \leqslant i \leqslant n$. Specifically, it is called an $[n, k, d]_q$ linear code if the minimum Hamming distance is d;
- For a subset $S \subseteq [n]$, let $|S|$ denote the cardinality of S, $\mathrm{Span}(S)$ be the linear space spanned by $\{\mathbf{g}_i : i \in S\}$ over $\mathbb{F}_q$ and $\mathrm{Rank}(S)$ be the dimension of $\mathrm{Span}(S)$.

Throughout this paper we assume that $\mathcal{C}$ be an $[n, k, d]_q$ linear code with generator matrix $G = (\mathbf{g}_1, \mathbf{g}_2, \cdots, \mathbf{g}_n)$.

Definition 1 ([17], [23])**:** The ith code symbol of an $[n, k, d]_q$ linear code $\mathcal{C}$, is said to have (r, δ)-locality if there exists a subset $S_i \subseteq [n]$ (a *repair set*) such that

- $i \in S_i$ and $|S_i| \leqslant r + \delta - 1$; and
- The minimum Hamming distance of the punctured code $\mathcal{C}|_{S_i}$, obtained by deleting the code symbols c_j for all $j \in [n] \setminus S_i$, is at least δ.

Furthermore, an $[n, k, d]_q$ linear code $\mathcal{C}$ is said to have information (r, δ)-locality (denoted as $(r, \delta)_i$-locality) if there exists a k-subset $I \subseteq [n]$ with $\mathrm{Rank}(I) = k$ such that for each $i \in I$, the ith code symbol has (r, δ)-locality and all symbol (r, δ)-locality (denoted as $(r, \delta)_a$-locality) if all the n code symbols have (r, δ)-locality.

Lemma 1 ([14], [23])**:** The minimum distance of an $[n, k, d]_q$ code $\mathcal{C}$ with $(r, \delta)_i$-locality is upper bounded by

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

Definition 2: A linear code with $(r, \delta)_i$-locality is said to be an *optimal locally repairable code* if its minimum Hamming distance meets the Singleton-type bound of Lemma 1 with equality.

By (1), even for an optimal $[n, k, d]_q$ linear code with $(r, \delta)_i$-locality (or $(r, \delta)_a$-locality), $d < n-k+1$ under the nontrivial case $k > r$. Thus, for a linear code with $(r, \delta)_i$-locality, it is natural to ask if it is possible for an erasure pattern $E \subset [n]$ with size $d \leqslant |E| \leqslant n - k$ to be recoverable [14]. Generally this problem is still open. However, two special settings of this problem received special attention in previous works.

Setting I: (e.g., [2], [22]) For a linear code with $(r, \delta)_a$-locality, let $(r + \delta - 1)|n$ and $|\{S_i : i \in [n]\}| = \frac{n}{r+\delta-1}$, i.e., all the n symbols are divided into $\frac{n}{r+\delta-1}$ repair sets. Let $s = \frac{n}{r+\delta-1} r - k$ and assume the elements of S_i are denoted by $\{s_{i,1}, s_{i,2}, \ldots, s_{i,r+\delta-1}\}$. An erasure pattern E is required to be recoverable if there exists a $(\delta - 1)$-subset of $[r + \delta - 1]$, $\{j_1, j_2, \cdots, j_{\delta-1}\}$, and there exists a set $E^* \subseteq E \subseteq [n]$, $|E^*| \leqslant \frac{n}{r+\delta-1} r - k$ and

$$(E \backslash E^*) \cap S_i \subseteq \{s_{i,j_1}, s_{i,j_2}, \ldots, s_{i,j_{\delta-1}}\} \text{ for each } i \in [n].$$

Setting II: (e.g., [13]) For a linear code with $(r, \delta)_a$-locality, let $(r + \delta - 1)|n$ and $|\{S_i : i \in [n]\}| = \frac{n}{r+\delta-1}$, i.e., all the n symbols are divided into $\frac{n}{r+\delta-1}$ repair sets. Let $s = \frac{n}{r+\delta-1} r - k$. An erasure pattern E is required to be recoverable if there exists a set $E^* \subseteq E \subseteq [n]$, $|E^*| \leqslant s$ and

$$|(E \backslash E^*) \cap S_i| \leqslant \delta - 1 \text{ for each } 1 \leqslant i \leqslant \frac{n}{r + \delta - 1}.$$

Definition 3: An $[n, k, d]_q$ linear code that satisfies the conditions of Setting I is said to be a sector-disk code ($(\delta - 1, s)$-SD).

As an intuition, we make the following analogies between a distributed storage system and Setting I. In this analogy, we have a total of $r + \delta - 1$ disks, each containing $\frac{n}{r+\delta-1}$ sectors, with a total number of sectors in the system which is n. The ith stripe, i.e., the set containing the ith sector from each disk, is an (r, δ)-repair set, for each i. Finally, an SD code is capable of correcting $\delta - 1$ whole disk erasures, as well as an extra s erased sectors.

Definition 4: An $[n, k, d]_q$ linear code that satisfies the conditions of Setting II is said to be a maximally recoverable code (MR code).

III. Constructions of Locally Repairable Codes

In this section, we introduce a general construction of locally repairable codes with information locality. Let $k = r\ell + v$ with $0 < v \leqslant r$ and $n = k + (\ell + 1)(\delta - 1) + h$ with $h \geqslant 0$, where all parameters are integers.

Construction A: Let the k information symbols be partitioned into $\ell + 1$ sets, say,

$$I^{(i)} = \{I_{i,1}, I_{i,2}, \ldots, I_{i,r}\}, \quad \text{for } i \in [\ell],$$
$$I^{(\ell+1)} = \{I_{\ell+1,1}, I_{\ell+1,2}, \ldots, I_{\ell+1,v}\}.$$

Let S be an h-subset of $\mathbb{F}_q$ and denote $A \triangleq \mathbb{F}_q \backslash S$. Let $\mathcal{A} = \{A_i : 1 \leqslant i \leqslant \ell + 1\}$ be a family of subsets of A with $|A_i| = r + \delta - 1$ and $|A_{\ell+1}| = v + \delta - 1$. Define

$$g_i(x) = \prod_{\theta \in A_i} (x - \theta) \text{ for } 1 \leqslant i \leqslant \ell + 1 \tag{2}$$

and

$$\Delta(x) = \prod_{1 \leqslant i \leqslant \ell+1} g_i(x). \tag{3}$$

A linear code with length n can be generated by defining a linear map from the information $\boldsymbol{I} = (I_{1,1}, \ldots, I_{\ell,v}) \in \mathbb{F}_q^k$ to a codeword $\boldsymbol{C}(\boldsymbol{I}) = (c_{1,1}, \ldots, c_{\ell,r+\delta-1}, c_{\ell+1,1}, \ldots, c_{\ell+1,v+\delta-1}, c_{\ell+2,1}, \ldots, c_{\ell+2,h}) \in \mathbb{F}_q^n$, thus the $[n, k]_q$ linear code is $\mathcal{C} = \{\boldsymbol{C}(\boldsymbol{I}) : \boldsymbol{I} \in \mathbb{F}_q^k\}$. This mapping is performed by the following two steps:

a) Step 1: For $1 \leqslant j \leqslant \ell+1$, by polynomial interpolation, there exists a unique $f_j(x) \in \mathbb{F}_q[x]$ with $\deg(f_j) < |A_j| - \delta + 1$ such that $f_j(\theta_{j,t}) = I_{j,t}$ for $1 \leqslant t \leqslant |A_j| - \delta + 1$, where $A_j = \{\theta_{j,t} \ : \ 1 \leqslant t \leqslant |A_j|\}$. For $1 \leqslant j \leqslant \ell+1$ and $1 \leqslant t \leqslant |A_j|$, set $c_{j,t} = f_j(\theta_{j,t})$.

b) Step 2: Let

$$f_I(x) = \Delta(x) \sum_{1 \leqslant i \leqslant \ell+1} \frac{f_i(x)}{g_i(x)}. \tag{4}$$

Set $c_{\ell+2,i} = f_I(s_i)$ for $1 \leqslant i \leqslant h$, where $S = \{s_i \ : \ 1 \leqslant i \leqslant h\}$.

Lemma 2: The code $\mathcal{C}$ generated by Construction A is an $[n,k]_q$ linear code with $(r,\delta)_i$ locality.

For ease of presentation, we use the evaluation points (instead of the indices of code symbols) to denote erasure patterns. Additionally, we shall group the erased positions by the index of the repair set they hit. Namely, we shall use $\mathcal{E} = \{E_1, \ldots, E_{\ell+2}\}$ to denote an erasure pattern, where $E_j \subseteq A_j$ is the set of erasure points in A_j, $1 \leqslant j \leqslant \ell+1$, and $E_{\ell+2} \subseteq S$ is the set of erasure points in S.

Theorem 1: Let $\mathcal{C}$ be the linear code generated by Construction A. Assume $\mathcal{E} = \{E_i \ : \ 1 \leqslant i \leqslant \ell+2\}$ is an erasure pattern, with $E_i \subseteq A_i$ for $1 \leqslant t \leqslant \ell+1$ and $E_{\ell+2} \subseteq S$. For $1 \leqslant i \leqslant \ell+1$, assume that, in $\mathcal{E}$, there exist $w \leqslant \ell+1$ sets with $|E_{i_t}| \geqslant \delta$ for $1 \leqslant t \leqslant w$ and $1 \leqslant i_t \leqslant \ell+1$. If the erasure pattern $\mathcal{E}$ satisfies

$$\left| \bigcup_{1 \leqslant t \leqslant w} E_{i_t} \right| + |E_{\ell+2}| \leqslant h + \delta - 1, \tag{5}$$

and for any $1 \leqslant j \leqslant w$

$$\left| A_{i_j} \cap \left(\bigcup_{j \neq t \in [w]} A_{i_t} \right) \right| \leqslant \delta - 1, \tag{6}$$

then the erasure pattern $\mathcal{E}$ can be recovered.

Remark 1: In Theorem 1, we want to highlight that the size $|(\bigcup_{1 \leqslant t \leqslant w} E_{i_t}) \cup E_{\ell+2}|$ dictates whether an erasure pattern is recoverable, and not the number of erased coordinates, i.e., $\sum_{1 \leqslant t \leqslant w} |E_{i_t}| + |E_{\ell+2}|$. This is to say, if there are erasures that share the same evaluation point (even in different coordinates), then those erasures as a whole will only increase the discriminant value by one. In such a case we may recover more than $h + \delta - 1$ erasures that are guaranteed to be recoverable by the value of the Singleton-type bound, i.e., $h + \delta$.

Corollary 1: If the set system $\mathcal{A}$ of Construction A satisfies that for any μ-subset D of $[\ell+1]$

$$\left| A_i \cap \left(\bigcup_{j \neq i, j \in D} A_j \right) \right| \leqslant \delta - 1 \quad \text{for } i \in D, \tag{7}$$

then the code $\mathcal{C}$ generated by Construction A is an $[n,k,d]_q$ linear code with $(r,\delta)_i$ locality and $d \geqslant \min\{(\mu+1)\delta, h+\delta\}$. Furthermore, if $h + \delta \leqslant (\mu+1)\delta$, then the code $\mathcal{C}$ is optimal with respect to the bound in Lemma 1.

Based on Corollary 1, to construct optimal locally repairable codes we only need to find $\mathcal{A}$ such that (7) holds.

Theorem 2: Assume the setting of Construction A. Let $\mathcal{A}$ be a set system formed by subsets of $\mathbb{F}_q \setminus S$, where S is an h-subset of $\mathbb{F}_q$. If there exists a positive integer a such that $|A_i \cap A_j| \leqslant a$ for all $i \neq j$, then the code $\mathcal{C}$ generated by Construction A is an $[n,k,d \geqslant \min\{h+\delta, (\lceil \frac{\delta}{a} \rceil + 1)\delta\}]_q$ linear code with $(r,\delta)_i$-locality. If additionally, $h \leqslant \lceil \frac{\delta}{a} \rceil \delta$, then the code $\mathcal{C}$ generated by Construction A is an optimal $[n,k,d = h+\delta]_q$ linear code with $(r,\delta)_i$-locality.

A. Optimal Locally Repairable Codes with Order-Optimal Length: $(r,\delta)_i$-Locality

Finding the maximal length of optimal locally repairable codes with $(r,\delta)_a$-locality was the subject of [16] and [7], for the cases of $\delta = 2$ and $\delta \geqslant 2$, respectively. It is therefore natural to further ask how long can optimal locally repairable codes with $(r,\delta)_i$-locality be. This question is also important to us in order to analyze the performance of Construction A.

Theorem 3: Let $n = k + \ell(\delta-1) + h$, $\delta \geqslant 2$, $k = \ell r$. Assume there exists an optimal $[n,k,d]_q$ linear code $\mathcal{C}$ with $(r,\delta)_i$-locality. For any given integer $0 \leqslant a \leqslant h$ define $T(a) = \lfloor (d-a-1)/\delta \rfloor$. If $T(a) \geqslant 2$, then

$$n \leqslant \begin{cases} \frac{r+\delta-1}{r} \left(\frac{T(a)-1}{2(q-1)} q^{\frac{2(h-a-1)}{T(a)-1}} + a + 1 \right) - \frac{h(\delta-1)}{r}, & \text{if } 2 \nmid T(a), \\ \frac{r+\delta-1}{r} \left(\frac{T(a)}{2(q-1)} q^{\frac{2(h-a)}{T(a)}} + a \right) - \frac{h(\delta-1)}{r}, & \text{if } 2 \mid T(a), \end{cases}$$

where h can be rewritten as $h = d - \delta$.

Definition 5 ([11], VI. 40)**:** Let $n_1 \geqslant 2$ be an integer and t a positive integer. A τ-$(n_1,t,1)$-*packing* is a pair $(\mathcal{X}, \mathcal{S})$, where $\mathcal{X}$ is a set of n_1 elements (called points) and $\mathcal{S} \subseteq 2^{\mathcal{X}}$ is a collection of t-subsets of $\mathcal{X}$ (called blocks), such that each τ-subset of $\mathcal{X}$ is contained in at most one block of $\mathcal{S}$. An τ-$(n_1,t,1)$-packing is said to be *regular* if each element of X appears in exactly w blocks, denoted by w-regular τ-$(n_1,t,1)$-packing. Furthermore, if each τ-subset of $\mathcal{X}$ is contained in exactly one block of $\mathcal{S}$, then $(\mathcal{X}, \mathcal{S})$ is also called a (τ,t,n_1)-*Steiner system*.

Corollary 2: Let $n_1 = q - h$. If there exists a $(\tau+1, r+\delta-1, n_1)$-Steiner system and $0 \leqslant h \leqslant \lceil \frac{\delta}{\tau} \rceil \delta$, then there exists an optimal $[n,k,d]_q$ linear code with $(r,\delta)_i$-locality, where

$$n = \frac{\binom{n_1}{\tau+1}(r+\delta-1)}{\binom{r+\delta-1}{\tau+1}} + h, \quad k = \frac{r\binom{n_1}{\tau+1}}{\binom{r+\delta-1}{\tau+1}},$$

and $d = h + \delta$. In particular, for the case $h \geqslant \delta + 1$ and $\tau = \delta - 1$, the code based on the $(\delta, r+\delta-1, q-h)$-Steiner system has order-optimal length, where h, r and δ are regarded as constants.

Remark 2: One well known construction for optimal locally repairable codes with $(r,\delta)_i$-locality is that of pyramid codes. The pyramid code is based on an MDS code whose length is upper bounded by $q+d-2$ (and by the MDS conjecture this may be reduced to $q+1$ for q odd [1]). Thus, the length of pyramid code is upper bounded by $q+d-1-\delta+\lceil\frac{k}{r}\rceil(\delta-1) \leqslant q+d-1-\delta+\lceil\frac{q-1}{r}\rceil(\delta-1)$ ($q+2-\delta+\lceil\frac{k}{r}\rceil(\delta-1) \leqslant q+2-\delta+\lceil\frac{q-d+2}{r}\rceil(\delta-1)$ according to MDS conjecture for the case of q odd), where $d \geqslant \delta$. According to our construction and bound (in Theorem 3), it follows that the pyramid code is sub-optimal in terms of asymptotic length, since we construct locally repairable codes with $(r,\delta)_i$-locality and length $n=\Omega(q^\delta)$.

Example 1: Set $n=24$, $k=14$, $\delta=2$, $r=2$, and $h=3$. Let $\mathcal{A}=\{A_i \;:\; A_i \triangleq \{3,6,5\}+i \subseteq \mathbb{Z}_7, i\in\mathbb{Z}_7\}$. According to Construction A, we can construct a $[24,14,5]_{11}$ optimal linear code with $(2,2)_i$-locality, consistent with the result in Theorem 2. The reader may refer to [9] for the parity check matrix of $\mathcal{C}$. Note that, to construct a code sharing the same parameters via the pyramid code, we need an MDS code with parameters $[18,14,5]_q$. However, according to the MDS conjecture this MDS code exists only under the condition that $q\geqslant 17$. Without the help of MDS conjecture, based on the result proposed in [1], we have $q\geqslant 16$ for this special setting.

Remark 3: For the case $\delta=2$ and $d=5,6$, optimal linear codes with all symbol $(r,2)$-locality and order-optimal length $\Theta(q^2)$ have been introduced in [4], [16], [18]. One can verify that our construction yields codes for more general cases $d>6$ even if we only consider the case $\delta=2$.

Remark 4: For the case $\delta\geqslant 2$ and $d=2\delta+1$, optimal linear codes with all symbol $(r,2)$-locality and order-optimal length $\Theta(q^\delta)$ have been introduced in [7].

IV. Generalized Sector-Disk Codes

By Theorem 1, we may have extra benefits if $|\bigcup_{|E_i|\geqslant\delta, i\in[\ell+1]} E_i| < \sum_{|E_i|\geqslant\delta,i\in[\ell+1]}|E_i|$. In this section, we are going to use this property to construct array codes that can recover from special erasure patterns beyond the minimum Hamming distance. The basic idea of those construction is to let all the code symbols share the same evaluation point in step 1 of Construction A in the same column of an array code. Then for this array code, one erased column may only increase the value $|\bigcup_{|E_i|\geqslant\delta,i\in[\ell+1]} E_i|$ by one. Hence, when we consider sector-disk-like erasure patterns, we may get some extra benefit beyond the minimum Hamming distance. We begin with some definitions.

Definition 6: Let $\mathcal{C}$ be an optimal $[n,k,d]_q$ linear code with $(r,\delta)_i$-locality. Then the code $\mathcal{C}$ is said to be an (s,γ)-*generalized sector-disk code* (GSD code) if the codewords can be arranged into an array

$$C=\begin{pmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,a}\\ c_{2,1} & c_{2,2} & \cdots & c_{2,a}\\ \vdots & \vdots & \ddots & \vdots\\ c_{b,1} & c_{b,2} & \cdots & c_{b,a}\end{pmatrix} \tag{8}$$

such that:

(I) all the erasure patterns that contain any s columns and additional γ cells can be recovered; and

(II) $sb+\gamma>d-1$.

Remark 5: If the code $\mathcal{C}$ has $(r,\delta)_a$-locality, the repair sets are exactly the rows, and then the $(\delta-1,d-\delta)$-GSD code is exactly the $(\delta-1,d-\delta)$-SD code [22]. Compared with SD codes, GSD codes relax the conditions in the following three aspects:

- GSD codes only require $(r,\delta)_i$-locality, whereas SD codes require $(r,\delta)_a$-locality;
- A row in an array codeword of a GSD code is not necessary a repair set;
- The number of column erasures is not restricted to $\delta-1$ as in SD codes.

In the following construction, we use Construction A to generate GSD codes.

Construction B: Let S be an h-subset of $\mathbb{F}_q$ and let $(X=\mathbb{F}_q\setminus S, \mathcal{A}=\{A_i \;:\; 1\leqslant i\leqslant \ell+1\})$ be a t-regular $(m,r+\delta-1,1)$-packing, where $A_i=\{\theta_{i,j} \;:\; 1\leqslant j\leqslant r+\delta-1\}$ for $1\leqslant i\leqslant \ell+1$. Based on $\mathcal{A}$ and S, we can generate a locally repairable code $\mathcal{C}$ according to Construction A. Define column vectors $V_\tau\in\mathbb{F}_q^t$ for $\tau\in\mathbb{F}_q$ as

$$V_\tau^{\mathsf{T}}=(c_{i_{\tau,1},j_{\tau,1}}, c_{i_{\tau,2},j_{\tau,2}},\ldots,c_{i_{\tau,t},j_{\tau,t}}),$$

where $\theta_{i_{\tau,v},j_{\tau,v}}=\tau$, for $1\leqslant v\leqslant t$. Arrange the h global parity symbols as the last $\lceil\frac{h}{t}\rceil$ columns.

Theorem 4: Let $\mathcal{C}$ be the $t\times(m+\lceil\frac{h}{t}\rceil)$ array code generated by Construction B. Then each element of the first m columns has (r,δ)-locality. If $h\leqslant\delta^2$, then the code can recover from any $h+\delta-1$ erasures. Furthermore:

(I) The code $\mathcal{C}$ can recover from any erasure pattern of $y\leqslant 2$ columns from the first m columns and any other $h-y-1$ erasures.

(II) If $\binom{y}{2}\leqslant\delta$, then the code $\mathcal{C}$ can recover from any erasure pattern of y columns from the first m columns and any other $h-2-\binom{y}{2}$ erasures.

(III) The code $\mathcal{C}$ can recover from any erasure pattern of $y<\frac{(\delta+1)\delta}{2}-1$ columns from the first m columns and any other $\min\{\frac{(\delta+1)\delta}{2}-y-1, h+\delta-1-y\}$ erasures.

For the case $r\nmid k$ and $h=r-v$, we may modify Construction B as follows.

Construction C: Let S be an $(r-v)$-subset of $\mathbb{F}_q$ and let $(X\subseteq\mathbb{F}_q\setminus S, \mathcal{B}=\{B_i \;:\; 1\leqslant i\leqslant\ell+1\})$ be a t-regular $(m,r+\delta-1,1)$-packing. Let $A_i=B_i$ for $1\leqslant i\leqslant\ell$ and $A_{\ell+1}\subseteq B_{\ell+1}$. Let $n=t|X|=t\rho$ and $k=\ell r+v$, then based on $\mathcal{A}$ and S, we can generate a locally repairable code $\mathcal{C}$ according to Construction A. List the elements of $B_{\ell+1}\setminus A_{\ell+1}$ as $(x_1,x_2,\ldots,x_{r-v})$ and X as $(x_1,x_2,\cdots,x_\rho)$. Define column vectors $V_{x_a}\in\mathbb{F}_q^v$ for $a\in[\rho]$ as

$$V_{x_a}^{\mathsf{T}}=\begin{cases} (c_{i_{x_a,1},j_{x_a,1}}, c_{i_{x_a,2},j_{x_a,2}},\ldots,c_{i_{x_a,t-1},j_{x_a,t-1}}, c_{\ell+2,a}), & \text{if } 1\leqslant a\leqslant r-v,\\ (c_{i_{x_a,1},j_{x_a,1}}, c_{i_{x_a,2},j_{x_a,2}},\ldots,c_{i_{x_a,t},j_{x_a,t}}), & \text{otherwise,}\end{cases}$$

where $\theta_{i_{x_a,b},j_{x_a,b}}=x_a$, $1\leqslant b\leqslant t-1$ for $1\leqslant a\leqslant r-v$ and $1\leqslant b\leqslant t$ for $r-v+1\leqslant a\leqslant\rho$.

Corollary 3: Let $\mathcal{C}$ be the $t \times \rho$ array code generated by Construction C. Then $\mathcal{C}$ has $(r,\delta)_i$-locality. If $h \leqslant \delta^2$, then the code can recover any $h+\delta-1$ erasures. Furthermore:

(I) The code $\mathcal{C}$ can recover from any erasure pattern of $y \leqslant 2$ columns and any other $h-2y-1$ erasures.

(II) If $\binom{y}{2} \leqslant \delta$, then the code $\mathcal{C}$ can recover from any erasure pattern of y columns and any other $h-2-\binom{y}{2}-y$ erasures.

(III) The code $\mathcal{C}$ can recover from any erasure pattern of $y < \frac{(\delta+1)\delta}{2}-1$ columns and any other $\min\{\frac{(\delta+1)\delta}{2}-2y-1, h+\delta-1-2y\}$ erasures.

As examples, we use two well known classes of Steiner systems that are the affine geometries and projective geometries.

Lemma 3 ([11]): Let $\beta \geqslant 2$ be an integer and q_1 a prime power, then there exists a $(2, q_1, q_1^{\beta})$-Steiner system.

Based on affine geometries and Construction C, we have the following conclusion for GSD codes.

Corollary 4: Let $\beta \geqslant 2$ be an integer and q_1 a prime power. Set $q_1 = r+\delta-1$, $n = \frac{q_1^{\beta}(q_1^{\beta}-1)}{q_1-1}$, $\delta \geqslant 2$, $k = (\frac{q_1^{\beta-1}(q_1^{\beta}-1)}{q_1-1}-1)r+v$ with $1 \leqslant v \leqslant r-1$, and $h = r-v = q_1-\delta-v+1$. Let $\mathcal{C}$ be the $\frac{q_1^{\beta}-1}{q_1-1} \times q_1^{\beta}$ array code generated by Construction C using a $(2, q_1, q_1^{\beta})$-Steiner system from Lemma 3. If $h \leqslant \delta^2$, then the code $\mathcal{C}$ is an $[n, k, h+\delta-1]_q$ optimal locally repairable code with $(r,\delta)_i$-locality, where $q \geqslant q_1^{\beta}+h$. Furthermore:

(I) If $y \leqslant 2$ and $y(\frac{q_1^{\beta}-1}{q_1-1}-2) > \delta$, then the code $\mathcal{C}$ is a $(y, h-2y-1)$-GSD code.

(II) If $\binom{y}{2} \leqslant \delta$ and $y\frac{q_1^{\beta}-1}{q_1-1}-1-\binom{y}{2}-y > \delta$, then the code $\mathcal{C}$ is a $(y, h-2-\binom{y}{2}-y)$-GSD code.

(III) If $y < \frac{(\delta+1)\delta}{2}-1$ and $y\frac{q_1^{\beta}-1}{q_1-1}+\gamma > h+\delta-1$, then the code $\mathcal{C}$ is a (y,γ)-GSD code, where $\gamma = \min\{\frac{(\delta+1)\delta}{2}-2y-1, h+\delta-1-2y\}$ erasures.

Herein, we highlight that the second restriction of each item comes from the requirement in Definition 6-(II).

Lemma 4 ([11]): Let $\beta \geqslant 2$ be an integer and q_1 a prime power, then there exists a $(2, q_1+1, \frac{q_1^{\beta+1}-1}{q_1-1})$-Steiner system.

Based on projective geometries and Construction C, we have the following conclusion for GSD codes.

Corollary 5: Let $\beta \geqslant 2$ be an integer and q_1 a prime power. Set $q_1+1 = r+\delta-1$, $n = \frac{(q_1^{\beta+1}-1)(q_1^{\beta}-1)}{(q_1-1)^2}$, $\delta \geqslant 2$, $k = (\frac{(q_1^{\beta+1}-1)(q_1^{\beta}-1)}{(q_1-1)(q_1^2-1)}-1)r+v$ with $1 \leqslant v \leqslant r-1$, and $h = r-v = q_1-\delta-v+2$. Let $\mathcal{C}$ be the $\frac{q_1^{\beta}-1}{q_1-1} \times \frac{q_1^{\beta+1}-1}{q_1-1}$ array code generated by Construction C using a $(2, q_1+1, \frac{q_1^{\beta+1}-1}{q_1-1})$-Steiner system from Lemma 3. If $h \leqslant \delta^2$, then the code $\mathcal{C}$ is an $[n, k, h+\delta-1]_q$ optimal locally repairable code with $(r,\delta)_i$-locality, where $q \geqslant \frac{q_1^{\beta+1}-1}{q_1-1}+h$ is a prime power. Furthermore:

(I) If $y \leqslant 2$ and $y(\frac{q_1^{\beta}-1}{q_1-1}-2) > \delta$, then the code $\mathcal{C}$ is a $(y, h-2y-1)$-GSD code.

(II) If $\binom{y}{2} \leqslant \delta$ and $y\frac{q_1^{\beta}-1}{q_1-1}-1-\binom{y}{2}-y > \delta$, then the code $\mathcal{C}$ is a $(y, h-2-\binom{y}{2}-y)$-GSD code.

(III) If $y < \frac{(\delta+1)\delta}{2}-1$ and $y\frac{q_1^{\beta}-1}{q_1-1}+\gamma > h+\delta-1$, then the code $\mathcal{C}$ is a (y,γ)-GSD code, where $\gamma = \min\{\frac{(\delta+1)\delta}{2}-2y-1, h+\delta-1-2y\}$ erasures.

Remark 6: In what follows, we list some known results about SD codes and MR code (PMDS codes) as a comparison, where $n = m(r+\delta-1)$ is the total number of sectors for a codeword, k is the number of sectors for information symbols, $r+\delta-1$ is the number of columns (also means that the code have $(r,\delta)_a$-locality), and q stands for the field size. To keep in the same starting line with our results in Corollaries 4 and 5, we regard r, δ, γ as a constant when we consider the relationship between n and q.

- For $\gamma = 1$, there exist (s,γ)-MR codes for $n = \Theta(q)$ [2].
- For $\gamma = 2$ and $s \in \{1,2,3\}$, there exist (s,γ)-SD codes for $n = \Theta(q)$ [22].
- For $\gamma = 2$, there exist (s,γ)-SD codes for $n = \Theta(q)$ [3].
- For $s = 1$, there exist (s,γ)-MR codes when $n = \Theta\left(q^{\frac{1}{\gamma-1}}\right)$ and $r \mid (k+\gamma)$ [13].
- For $\gamma = 3$, there exist (s,γ)-MR codes when $n = \Theta(q^{1/3})$ [15].
- There exist $(s = \delta-1, \gamma)$-MR codes when $n = \Theta(\log q)$ [5].
- There exist $(s = \delta-1, \gamma)$-MR codes when $n = \Theta(q^{\frac{1}{\gamma}})$ [12].
- There exist $(s = \delta-1, \gamma)$-MR codes when $n = \Theta(q^{\frac{1}{r}})$ [21].

Remark 7: By Corollaries 4 and 5, there exist GSD codes with $n = \Theta(q^2)$, where h, r, and δ are regarded as constants, i.e., q_1 is a constant. Note that if we regard $\beta \geqslant 2$ as a constant then $n = \Theta(q^{\frac{2\beta-1}{\beta}})$ with $q = \Theta(q_1^{\beta})$. In addition, for general cases by using Steiner systems with parameters $(\tau, r+\delta-1, n_1)$, Steiner systems are capable of yielding optimal locally repairable codes (similarly, GSD codes) with length $n = \Theta(q^{\tau})$ as shown in Corollary 2 and Remark 2.

References

[1] S. Ball, "On sets of vectors of a finite vector space in which every subset of basis size is a basis," *J. Eur. Math. Soc.*, vol. 14, no .3, pp. 733-748, 2012.

[2] M. Blaum, J. L. Hafner, and S. Hetzler, "Partial-MDS codes and their application to RAID type of architectures," *IEEE Trans. Inf. Theory,* vol. 59, no. 7, pp. 4510-4519, Jul. 2013.

[3] M. Blaum, J. Plank, M. Schwartz, and E. Yaakobi, "Construction of partial MDS and sector-disk codes with two global parity symbols," *IEEE Trans. Inf. Theory,* vol. 62, no. 5, pp. 2673-2681, May 2016.

[4] A. Beemer, R. Coatney, V. Guruswami, H. H. López, and F. Piñero, "Explicit optimal-length locally repairable codes of distance 5," arXiv:1810.03980, 2018.

[5] G. Calis and O. O. Koyluoglu, "A general construction for PMDS codes," *IEEE Commun. Lett.,* vol. 21, no. 3, pp. 452-455, Mar. 2017.

[6] V. R. Cadambe and A. Mazumdar, " Bounds on the size of locally recoverable codes," *IEEE Trans. Inf. Theory,* vol. 61, no. 11, pp. 5787-5794, Nov. 2015.

[7] H. Cai, Y. Miao, M. Schwartz, and X. Tang, "On optimal locally repairable codes with super-linear length," to appear in *IEEE Trans. Inf. Theory*.

[8] H. Cai, Y. Miao, M. Schwartz, and X. Tang, "On optimal locally repairable codes with multiple disjoint repair sets," vol. 66, no. 4, pp. 2402-2416, Apr. 2020.

[9] H. Cai and M. Schwartz, " On optimal locally repairable codes and generalized sector-disk codes " *arXiv:2001.02909*, 2020.

[10] B. Chen, S. Xia, J. Hao, and F. Fu, "Constructions of optimal cyclic (r, δ) locally repairable codes," *IEEE Trans. Inf. Theory*, vol. 64, no. 4, pp. 2499-2511, Apr. 2018.

[11] C. J. Colbourn and J. H. Dinitz, *Handbook of Combinatorial Designs*, Chapman & Hall/CRC, vol. 42, 2006.

[12] R. Gabrys, E. Yaakobi, M. Blaum, and P.H. Siegel, "Constructions of partial MDS codes over small fields," *In Proc. of IEEE ISIT*, 2017.

[13] P. Gopalan, C. Huang, B. Jenkins, and S. Yekhanin, "Explicit maximally recoverable codes with locality," *IEEE Trans. Inf. Theory,* vol. 60, no.9, pp. 5245-5256, Sept. 2014.

[14] P. Gopalan, C. Huang, H. Simitci, and S. Yekhanin, "On the locality of codeword symbols," *IEEE Trans. Inf. Theory,* vol. 58, no. 11, pp. 6925-6934, Nov. 2012.

[15] S. Gopi, V. Guruswami, and S. Yekhanin, "Maximally recoverable LRCs: A field size lower bound and constructions for few heavy parities," to appear in *IEEE Trans. Inf. Theory*.

[16] V. Guruswami, C. Xing, and C. Yuan, "How long can optimal locally repairable codes be?" *IEEE Trans. Inf. Theory,* vol. 65, no. 6, pp. 3662-3670, Jun. 2019.

[17] C. Huang, M. Chen, and J. Li, "Pyramid codes: Flexible schemes to trade space for access efficiency in reliable data storage systems," In NCA, 2007.

[18] L. Jin, "Explicit construction of optimal locally recoverable codes of distance 5 and 6 via binary constant weight codes," *IEEE Trans. Inf. Theory,* vol. 65, no. 8, pp. 4658-4663, Aug. 2019.

[19] S. Kadhe and A. Sprintson, "Codes with unequal locality," *In Proc. of IEEE ISIT*, 2016.

[20] G. Kim and J. Lee, "Locally repairable codes with unequal locality requirements," *IEEE Trans. Inf. Theory,* vol. 64, no. 11, pp. 7137-7152, Nov. 2018.

[21] U. Martínez-Peñas and F. R. Kschischang, "Universal and dynamic locally repairable codes with maximal recoverability via sum-rank codes," *IEEE Trans. Inf. Theory,* vol. 65, no. 12, pp. 7790-7805, Dec. 2019.

[22] J. S. Plank and M. Blaum, "Sector-disk (SD) erasure codes for mixed failure modes in RAID systems," *ACM Transactions on Storage,* vol. 10, no. 1, article 4, 2014.

[23] N. Prakash, G. M. Kamath, V. Lalitha, and P. V. Kumar, "Optimal linear codes with a local-error-correction property," *In Proc. of IEEE ISIT*, 2012.

[24] A. S. Rawat, D. S. Papailopoulos, A. G. Dimakis, and S. Vishwanath, "Locality and availability in distributed storage," *IEEE Trans. Inf. Theory,* vol. 62, no. 8, pp. 4481-4493, 2016.

[25] B. Sasidharan, G. K. Agarwal, and P. V. Kumar, "Codes with hierarchical locality," *In Proc. of IEEE ISIT*, 2015.

[26] M. Sathiamoorthy, M. Asteris, D. Papailiopoulos, A. G. Dimakis R. Vadali, S. Chen, and D. Borthakur "Xoring elephants: Novel erasure codes for big data," *In Proc. of VLDB*, 2013.

[27] N. Silberstein, T. Etzion, and M. Schwartz, "Locality and availability of array codes constructed from subspaces," *IEEE Trans. Inf. Theory,* vol. 65, no. 5, pp. 2648-2660, May 2019.

[28] I. Tamo and A. Barg, "A family of optimal locally recoverable codes," *IEEE Trans. Inf. Theory,* vol. 60, no. 8, pp. 4661-4676, 2014.

[29] I. Tamo, A. Barg, and A. Frolov, "Bounds on the parameters of locally recoverable codes," *IEEE Trans. Inf. Theory,* vol. 62, no. 6, pp. 3070-3083, Jun. 2016.

[30] A. Wang, and Z. Zhang, "Repair locality with multiple erasure tolerance," *IEEE Trans. Inf. Theory,* vol. 60, no. 11, pp. 6979-69878, Nov. 2014.

[31] A. Zeh and E. Yaakobi, "Bounds and constructions of codes with multiple localities," *In Proc. of IEEE ISIT*, 2016.