Optimal 5-Seq LRCs With Availability From Golomb Rulers

Hyojeong Choi, *Graduate Student Member, IEEE*, and Hong-Yeop Song, *Senior Member, IEEE*

***Abstract*— In this paper, we propose a simple construction for binary (n, k) linear codes using s-mark Golomb rulers. We prove that these codes are sequential-recovery locally repairable codes (LRCs) with availability 2, which can sequentially recover 5 erased symbols. We prove the necessary and sufficient condition for the proposed codes to be rate-optimal. We also prove the necessary and sufficient condition for the proposed codes to be dimension-optimal. Finally, we propose some variations of this constructions to obtain some 5-sequential recovery LRCs with availability 3. The proposed codes have higher rates and have more flexible choice for the lengths than other previously reported constructions.**

***Index Terms*— Locally repairable codes, sequential recovery, Golomb rulers, cyclic planar difference sets.**

I. Introduction

TO STABLY store big data in distributed storage systems (DSSs) and to increase their reliability, locally repairable codes (LRCs) that repair a single node failure with only a small number of nodes have been proposed by Gopalan et al. [15], [25]. The LRCs with length n, dimension k, and locality r is denoted as an (n, k, r)-LRC. An important parameter for LRCs is the locality r [15], [22], [25], which is the minimum number of symbols in the codeword required to repair a single erasure symbol. If each symbol has locality at most r, then the code is said to have the locality r.

For multiple erasures, one refers to the parallel-recovery or sequential-recovery LRCs, depending on whether the process of recovering multiple erasures is simultaneously parallel or sequentially one by one.

Parallel-recovery LRCs have been extensively studied in various types, including LRCs with joint locality [22], LRCs with cooperative locality [30], and LRCs with availability [31], [38], [39], [41], [42], [43]. Let C be an (n, k) linear code and $c = (c_0, c_1, \ldots, c_{n-1}) \in C$. Then C is said to be an (n, k, r)-LRC with availability t if, for each $i \in \{0, 1, \ldots, n-1\}$, there exist at least t pairwise disjoint repair sets $R_1(i), R_2(i), \ldots, R_t(i) \subseteq \{0, 1, \ldots, n-1\} \setminus \{i\}$, such that for each $j = 1, 2, \ldots, t$, we have (i) $|R_j(i)| \le r$, (ii) the symbol c_i is a linear combination of c_l's for $l \in R_j(i)$. The concept of availability is used for 'hot data' that needs to be repaired frequently. By definition, an LRC with availability can repair a single erasure using multiple disjoint repair sets, and hence it can also repair multiple erasures in parallel. Note that the LRC has availability $t \ge 2$ implies t parallel recovery process is possible for any single erasure of the code. However, t-parallel-recovery LRC does not necessarily have availability t.

Sequential recovery means that erased symbols are recovered sequentially one by one, and some previously recovered symbols can also be used to recover the remaining erased symbols. In contrast to the extensively studied parallel-recovery LRCs, this important method has only recently received significant attention. Sequential-recovery LRC was first proposed by Prakash et al. [27]. Let C be an (n, k) linear code, $c = (c_0, c_1, \ldots, c_{n-1}) \in C$ and u be a positive integer. Then C is said to be a u-sequential-recovery LRC (u-seq LRC) if there exists a sequential arrangement $(i_0, i_1, \ldots, i_{u-1})$ for any u erased positions such that, for each $l = 0, 1, \ldots, u-1$, there is a subset $R_l \subset \{0, 1, \ldots, n-1\}$ satisfying (i) $i_l \in R_l$ and $|R_l| \le r + 1$, (ii) $R_l \cap \{i_{l+1}, i_{l+2}, \ldots, i_{u-1}\} = \emptyset$ and (iii) the symbol c_{i_l} is a linear combination of c_j's for $j \in R_l \setminus i_l$. One important metric for u-seq LRCs is a repair time, which defines the maximum number of steps required to repair u erasures [36], [42]. In general, u-seq LRC repairs multiple erasures sequentially one by one, so the repair time is at most u. Designing u-seq LRCs with the repair time less than u is a challenging problem.

The u-seq LRCs for $u = 2$ or 3 have been extensively studied in [3], [7], [17], [26], [34], and [35]. Prakash et al. [26] established upper bounds on the code rate and minimum distance for 2-seq LRCs and proposed graph-based constructions that achieve these bounds. Jing and Song [17] constructed 2-seq LRCs that are either rate-optimal or distance-optimal, based on good polynomials for relatively small alphabets. For 3-seq LRCs, Balaji et al. [3] proposed some general constructions with short block lengths, as well as a bound on the block length. Subsequently, Song et al. [34] were the first to derive a tight upper bound on the code rate for any 3-seq LRC and proposed rate-optimal constructions based on resolvable configurations. In [7], the upper bound for dimension k of some 3-seq LRCs with availability 2 was established. Additionally, in [35], a length bound for 3-seq LRCs was proposed and the existence of LRCs that achieve this bound was also discussed.

Some results on u-seq LRCs for some larger u have been reported in [2], [3], [18], [36], and [42]. Balaji et al. [3] constructed high-rate u-seq LRCs for $u = 4, 5, 6, 7$ with $r = 2$ that approach the rate bound. The case for $u = 5$ is further

Received 29 July 2024; revised 11 November 2024; accepted 26 December 2024. Date of publication 3 January 2025; date of current version 19 February 2025. This work was supported by the National Research Foundation of Korea (NRF) Grant by the Korean Government through the Ministry of Sciences and ICT (MSIT) under Grant RS-2023-00209000. *(Corresponding author: Hong-Yeop Song.)*

The authors are with the School of Electrical and Electronic Engineering, Yonsei University, Seoul 03722, South Korea (e-mail: hjchoi3022@yonsei.ac.kr; hysong@yonsei.ac.kr).

Communicated by Y. Miao, Associate Editor for Coding and Decoding.
Digital Object Identifier 10.1109/TIT.2025.3525668

reviewed in Section IV. Song and Yuen [36] constructed (2^m-1)-seq LRCs that is equivalent to the direct product of m copies of the single-parity-check code. These constructions do not achieve the rate bound for general u-seq LRCs established later. Subsequently, Balaji et al. [2] derived upper bounds on the code rate and dimension for u-seq LRCs for any u and $r \geq 3$ as follows:

$$\frac{k}{n} \leq \frac{r^{\sigma+1}}{r^{\sigma+1} + 2\sum_{i=1}^{\sigma} r^i + (u-2\sigma)}\,, \tag{1}$$

$$k \leq \left\lfloor \frac{nr^{\sigma+1}}{r^{\sigma+1} + 2\sum_{i=1}^{\sigma} r^i + (u-2\sigma)} \right\rfloor, \tag{2}$$

where $\sigma \triangleq \lfloor \frac{u-1}{2} \rfloor$. Here, the authors designed rate-optimal binary u-seq LRCs that achieve the bound (1) by using the incidence matrix of a tree-like graph with girth $\geq u+1$. The detailed construction for $u = 5$ is also reviewed in Section IV.

An u-seq LRC that achieves the bound (1) with equality is defined to be a rate-optimal code [2]. It is well-known [2] that the bound cannot be achieved when the code length n is not a multiple of the denominator of RHS of (1). In this case, the focus shifts to achieving the bound given in (2). An u-seq LRC of length n that achieves the bound (2) with equality is defined to be a dimension-optimal code [2]. We note that a rate-optimal u-seq LRC is always dimension-optimal, but not conversely. That is, a dimension-optimal u-seq LRC is not always rate-optimal.

Yavari and Esmaeili [42] constructed u-seq LRCs with availability t ($u \geq t$) and introduced the concept of *joint sequential-parallel recovery LRCs*, which uses availability to recover some erased symbols simultaneously. This concept is aimed at reducing repair time, making the recovery process more efficient. By introducing this concept, u-seq LRCs with availability t can recover t symbols simultaneously, while the remaining $u-t$ symbols can be sequentially recovered, resulting in a repair time of at most $u-t+1$. The construction for $u = 5$ here is also reviewed in Section IV.

In this paper, we will design 5-sequential-recovery LRCs with availability 2 or 3 using Golomb rulers. Our main contributions are the following:

- We propose some new constructions for 5-seq LRCs with availability 2 using Golomb rulers. When the length of the code is sM for some parameters s and M, the dimension of the code is determined to be $sM-2M+1$. The repair time of this code is at most 3 due to the availability 2.
- We prove the necessary and sufficient conditions for obtaining rate-optimal LRCs in the proposed construction.
- We prove the necessary and sufficient conditions for obtaining dimension-optimal LRCs in the proposed construction.
- We propose 5-seq LRCs with availability 3 as a variation of the above. The repair time of this code is at most 2.

Section II provides some preliminary background information. Four main results above are discussed in Section III and it has four subsections described as above. Finally, Section IV, we conclude the paper by comparing the constructions with those in [2], [3], and [42] as a table.

II. Preliminaries

A set of integers $G = \{g_1 < g_2 < \cdots < g_s\}$ is called an *s-mark Golomb ruler* if the differences $g_j - g_i$ for $i < j$ are all distinct [5], [10], [11], [12], [14], [23], [29], [32], [36]. Subtracting g_1 from all the $g_i (1 \leq i \leq s)$ above gives also an s-mark Golomb ruler, and we may assume that $g_1 = 0$ in this paper.

We denote an s-mark Golomb ruler G by a strictly increasing integer sequence $0 = g_1 < g_2 < \cdots < g_s$ and also denote the set of positive distances of all the mark-pairs of G by

$$D = \{g_j - g_i | i < j\}. \tag{3}$$

Note that $|D| = s(s-1)/2$.

Our main theorem essentially constructs a parity check matrix H of a Quasi-Cyclic Low Density Parity Check (QC-LDPC) code with some interesting properties which lead to the conclusion that it is indeed a parity check matrix of a u-seq LRC. The interesting such relation was recently found in [18] and quoted as:

Known-fact 1: [18]

1) A linear block code is a u-seq LRC with locality r if its parity check matrix satisfies the following:
   (i) the girth is $2(u+1)$.
   (ii) the column weight is at least 2, and
   (iii) the row weight is at most $r+1$.
2) The repair time of this u-seq LRC is at most $\lceil u/2 \rceil$.

Tanner graph of a parity check matrix H or Tanner graph representation of H is a bipartite graph consisting of check nodes and variable nodes corresponding to rows and columns of H respectively, and a check node and a variable node is connected if and only if the corresponding row and column intersect with the value 1 in H [13], [18]. In Tanner graph, a cycle is a closed path and 2α-cycle is a cycle of length 2α. The length of a cycle is the total number of edges consisting of the cycle. In a bipartite graph, any cycle must have an even length. The girth of a Tanner graph is the minimum integer Γ such that Γ-cycle exists in the graph [13], [18].

As a method for constructing a parity check matrix for QC-LDPC codes, the approach based on the exponent matrix was introduced in [13] and used a lot, for example, in [19], [20], and [21].

Construct an integer matrix $E = [e(i,j)]$, called an exponent matrix. Then the parity check matrix H is constructed by replacing each entry of E with an appropriate Circular Permutation Matrix (CPM) of fixed size $M \times M$. For the position (i,j) in E, the appropriate CPM is the $M \times M$ binary identity matrix with its columns circularly shifted to the left by $e(i,j)$. The direction of the shift does not matter much, but for clarity and simplicity, we will fix it to the LEFT in this paper. We will always consider this type of H matrix in this paper.

An interesting relation between the girth of the resulting QC-LDPC code and some structural properties of E is given in [13]. We quote this as:

Known-fact 2: [13]

1) If a 2α-cycle exists in the Tanner graph of H matrix based on an exponent matrix $E = [e(i,j)]$ and CPMs,

then

$$\sum_{l=0}^{\alpha-1}(e(i_l,j_l)-e(i_l,j_{l+1}))\equiv 0 \pmod{M} \tag{4}$$

for some $i_0, i_1, \ldots, i_{\alpha-1}$ and $j_0, j_1, \ldots, j_\alpha = j_0$ such that $i_l \neq i_{l+1}$ for $0 \leq l < \alpha - 1$ and $j_l \neq j_{l+1}$ for $0 \leq l < \alpha$.

2) When the exponent matrix E has only 2 rows, the girth of H must be a multiple of 4.

Consider a binary vector $h = (h_1, h_2, \ldots, h_n)$ of length n. The $supp(h)$ is defined to be a subset of $\{1, 2, \ldots, n\}$ such that $i \in supp(h)$ if and only if $h_i \neq 0$ [26]. Given an LRC with its parity check matrix H, any row vector h of H determines a repair group of the symbols in codewords as a subset $supp(h)$ of $\{1, 2, \ldots, n\}$, whose size is equal to the weight of h [25]. If all these repair groups by the rows of H are pairwise disjoint and their union becomes $\{1, 2, \ldots, n\}$, the LRC is said to have disjoint (local) repair groups [7], [24], [37], [40].

In fact, in the parity check matrix H constructed from a $1 \times s$ exponent matrix and various CPMs substituted as in the previous paragraph, the repair groups corresponding to the rows of H are pairwise disjoint, and the union of their column indices must become $\{1, 2, \ldots, n\}$. Proposition 1 gives a necessary and sufficient condition for the LRC with H based on the $t \times s$ exponent matrix and various CPMs substituted will have availability t.

Proposition 1: Let $t \geq 2$ be an integer. Assume that we have t distinct parity check matrices H_i, for $i = 1, 2, \ldots, t$, all of size $M \times (r+1)M$ with $M \geq r+1$. Assume that each H_i has disjoint repair groups of the constant size $r+1$. Then, the linear code C with the parity check matrix given as follows which is the intersection of constituent codes corresponding to H_i $(1 \leq i \leq t)$:

$$H = \begin{pmatrix} H_1 \\ \hdashline H_2 \\ \hdashline \vdots \\ \hdashline H_t \end{pmatrix} \tag{5}$$

of size $tM \times (r+1)M$ has availability t if and only if

$$|supp(h_i) \cap supp(h_j)| \leq 1 \tag{6}$$

where h_i and h_j are two distinct row vectors of H.

Proof: Recall that each H_i determines M disjoint repair groups all of constant size $r+1$, and that every column of H_i has weight 1. Therefore, every column of H has weight t.

For sufficiency, let $v \in \{1, 2, \ldots, n\}$ and we pick up the v-th column of H of weight t. Here, we use row indices $h_1, h_2, \ldots, h_t$ for these t 1's in this column. If the condition in (6) is satisfied, then $supp(h_i) \cap supp(h_j) = \{v\}$ for $i, j \in \{1, 2, \ldots, t\}$. This gives the availability t for the symbol c_v. The necessary condition is easy and straightforward. ■

It is known [7] that when H consists of H_1 and H_2, the code with H becomes a 3-seq LRC with availability 2 even if its individual constituent code with H_i is a (1-seq) LRC.

In this paper, our main construction in Section III demonstrates that such a code becomes a 5-seq LRC with availability 2 when H_1 and H_2 are appropriately selected using Golomb rulers. Furthermore, these 5-seq LRCs can achieve rate-optimality in certain constructions, and with availability 2, they can reduce the repair time for 5 erased symbols to at most 3. Additionally, we will apply this proposition for $t = 3$ in Section IV to construct 5-seq LRCs with availability 3.

III. Some New Constructions and Rate-Optimality

A. Main Construction

Theorem 1: Let $s \geq 3$ be an integer and $G = \{g_1 = 0, g_2, \ldots, g_s\}$ with $0 = g_1 < g_2 < \cdots < g_s$ be an s-mark Golomb ruler and $D = \{g_j - g_i | i < j\}$. Let E be the $2 \times s$ integer matrix of the form

$$E = (e(i,j)) = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & g_2 & g_3 & \cdots & g_s \end{pmatrix}, \tag{7}$$

where $i = 0, 1,\ j = 0, 1, \ldots, s-1$. Now, construct a binary $2M \times sM$ matrix H by substituting a circular permutation matrix (CPM) of size $M \times M$ into the position (i, j) of E for all i, j. Here, each CPM for the position (i, j) is obtained by taking the circular left-shift of the columns of the identity matrix by the integer $e(i, j)$.

Then, H becomes a parity check matrix of a binary linear (n, k) code with length $n = sM$ and dimension $k = sM - 2M + 1$. Furthermore, this binary linear (n, k) code becomes a 5-seq LRC with availability 2, locality $r = s - 1$ and the repair time at most 3, provided that the positive integer M satisfies the following three conditions:

(M1) $g_i \not\equiv g_j \pmod{M}$ for all $i \neq j$;

(M2) $d + d' \not\equiv 0 \pmod{M}$ for all $d, d' \in D$ where d and d' are not necessarily distinct members of D; and

(M3) all the members of D and M are collectively relatively prime.

Proof: Let I be the $M \times M$ identity matrix and $I^{(\lambda)}$ be the λ-shifted version of I circularly to the left by the integer λ. The matrix H in the theorem becomes

$$H = \begin{pmatrix} I & I & \cdots & I \\ I & I^{(g_2)} & \cdots & I^{(g_s)} \end{pmatrix}, \tag{8}$$

which defines a binary linear (n, k) code with length $n = sM$ and dimension $k \geq sM - 2M$.

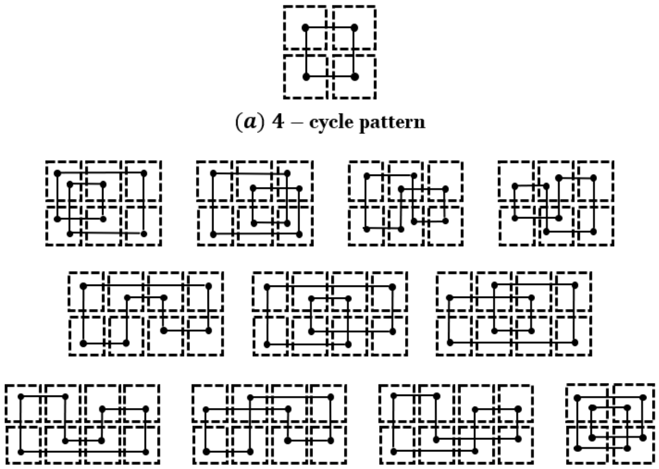
Now, we first prove the dimension $k = sM - 2M + 1$ of this linear code by using **M3** to show that $\text{rank}(H) = 2M - 1$.

To prove this, we first observe that all the rows of H sum to the zero vector and hence the $2M$ rows of H are not linearly independent. This implies that the maximum number of independent rows is at most $2M - 1$, or $\text{rank}(H) \leq 2M - 1$. This also implies that the top row is the sum of the remaining $2M - 1$ rows of H. Next, we claim that,

Main Claim: the top row of H cannot be made by less than the remaining $2M - 1$ rows of H.

This implies that we need at least these $2M - 1$ rows of H (except for the top row) to span the row space of H. Therefore, $\text{rank}(H) \geq 2M - 1$.

To prove the main claim, we first shuffle the columns of H so that all the 1's at the top row appear in the left-most s positions. We then shuffle the remaining $(s-1)M$ right-most columns of H so that all the 1's at the next row appear in the



(a) **4 − cycle pattern**

(b) **8 − cycle patterns**

Fig. 1. Patterns of 4-cycle and 8-cycle.

positions $s, s+1, \ldots, 2s-1$. We repeat this for the remaining columns of H so that all the 1's appear in the next s columns, and finally arrive at the following form, in which the result contains M groups of columns in which 1 repeats s times in the successive rows of the upper half. We call this H' in the following:

$$H' = \left(\begin{array}{c|c|c|c} 1\cdots 1 & 0\cdots 0 & & 0\cdots 0 \\ 0\cdots 0 & 1\cdots 1 & & 0\cdots 0 \\ & & \ddots & \\ 0\cdots 0 & 0\cdots 0 & & 1\cdots 1 \\ \hdashline 1 & & & 1 \\ \;\;1 & 1 & & \\ & \;\;1 & & \;\;1 \\ \cdots & \cdots & \cdots & \cdots \\ \;\;1 & & & \\ & \;\;\;1 & & 1 \end{array}\right). \tag{9}$$

For simplicity, denote by $0, 1, 2, \ldots, M-1$ the row indices of the lower half of H' in (9). The lower half of H' is now partitioned into M column groups each consisting of s columns. Here, the column group indices also use $0, 1, 2, \ldots, M-1$.

We now observe that the left-most s columns in the lower half of H' contain 1's in exactly the rows indexed by $G = \{g_1, g_2, \ldots, g_s\}$. This is a simple and direct consequence of using various CPMs in (8) which are some circularly **left-shifted** versions of the $M \times M$ identity matrix. We furthermore observe that the j-th column group (of s columns), for $j = 0, 1, \ldots, M-1$, has 1's in the rows indexed by the set $G + j \pmod{M}$. Recall the definition of D of G in (3). We now claim that, for any given $d \in D$, row indices of two column groups which are d apart has some intersections of row indices:

$$(G + j) \cap (G + j + d) \neq \phi \quad \text{for any } j \text{ and } d \in D. \tag{10}$$

If $d \in D$ then $d = g_k - g_l$ for some $l < k$, and hence, $G+j+d$ contains (for $i = 1, 2, \ldots, s$) $g_i + j + d = g_i + j + (g_k - g_l)$ which becomes $g_k + j \in G + j$ when $i = l$.

Now we try to make the top row by adding some remaining rows of H'. Since the top row has 1's only in all the columns of the left-most (0-th) column group, we have to add all the rows in the lower half indexed by G. Observe now that these rows indexed by G in the lower half of H' may have 1's in other positions (of different column groups). From (10), we see that for any $d \in D$, the d-th column group has 1's in the rows indexed by $G+d$ and the intersection of G and $G+d$ is non-empty. To make the top row, this 1 must be cancelled by adding the corresponding row (d-th row) in the upper half of H', which forces to add all the s rows in the lower half of H' indexed by $G + d \pmod{M}$.

We repeat this process and conclude that, if $d' \in D$, then $(G + d) \cap (G + d + d') \neq \phi$, and hence, we have to add $d + d'$-th row in the upper half and also all the rows in the lower half indexed by $G + d + d'$. We may have to continue this process indefinitely, and we have to add the rows in the lower half indexed by G, $G + d$, $G + d + d'$, $G + d + d' + d''$, $\ldots$ for all $d, d', d'', \ldots \in D$ as well as the rows in the upper half indexed by $d, d + d', d + d' + d'', \ldots$. It is now sufficient for **Main Claim** to confirm that the sequence

$$0, d, d + d', d + d' + d'', \ldots \pmod{M}$$

contains all the residues (indices) $0, 1, 2, \ldots, M-1 \pmod{M}$, where $d, d', d'', \ldots$ are not necessarily distinct in D. This is equivalent to the fact that any member in $\{0, 1, 2, \ldots, M-1\}$ is an integer linear combination mod M of the members in D, which comes from the condition **M3**.

Now we prove that this binary (n, k) code defined by H in (8) is an LRC with prescribed properties using the conditions **M1** and **M2**.

We first show that the code is a 5-seq LRC. Observe that every column of H has weight 2 and every row has weight $s = r + 1$. From **Known-fact 1**, therefore, it is sufficient to show that H has girth at least 12. Also, the LRC having this H as a parity check matrix becomes a 5-seq LRC with repair time of at most 3 by **Known-fact 1**.

Now, we claim that the girth is at least 12. We first observe that the 6-cycle and 10-cycle cannot occur in the Tanner graph (bi-partite graph) representation of the matrix H from 2) in **Known-fact 2**. Therefore, we only need to rule out the existence of 4- and 8-cycles.

By 1) in **Known-fact 2**, the cycles in H defined from the exponent matrix E can be analysed as the cycle patterns of E as shown in Fig. 1. It shows the unique pattern of 4-cycles in (a) and all possible (distinct) patterns of 8-cycles in (b).

We first consider the 4-cycle pattern (a) in Figure 1. We see that the existence of the pattern (a) implies that, when two blocks correspond to $I^{(g_i)}$ and $I^{(g_j)}$, we have the situation shown below in the exponent matrix E:

		$j_0 = j_2$		j_1	
$i_0 = 0$	$\cdots$	0	$\cdots$	0	$\cdots$
$i_1 = 1$	$\cdots$	g_i	$\cdots$	g_j	$\cdots$

Therefore, from (4) with $\alpha = 2$, we have

$$\begin{aligned} & e(i_0, j_0) - e(i_0, j_1) + e(i_1, j_1) - e(i_1, j_0) \\ \equiv\; & g_j - g_i \\ \equiv\; & 0 \pmod{M}, \end{aligned}$$

which is impossible, since M and G satisfy **M1**.

Now, we consider all the 8-cycle patterns in (b) of Fig. 1. The existence of a pattern in (b), for example, the very first one in (b), implies that, when three blocks correspond to $I^{(g_i)}$, $I^{(g_j)}$ and $I^{(g_k)}$, we have the situation shown below in the exponent matrix E:

$$\begin{array}{c|ccccccc} & & j_0=j_2=j_4 & & j_3 & & j_1 & \\ \hline i_0=i_2=0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots \\ i_1=i_3=1 & \cdots & g_i & \cdots & g_j & \cdots & g_k & \cdots \end{array}$$

Therefore, from (4) with $\alpha = 4$ we have

$$(g_k - g_i) + (g_j - g_i) \equiv 0 \pmod{M},$$

which is impossible, since M and G satisfy **M2**. All other remaining cases of the 8-cycle patterns in (b) of Fig. 1 can be ruled out similarly, by the condition **M2**. This proves that H has girth at least 12.

Now we have confirmed that the girth of H is 12, and hence, there are no 4-cycles. Consequently, the availability $t = 2$ is guaranteed. ■

Corollary 1: Let $s \geq 3$ be an integer and $G = \{g_1 = 0, g_2, \ldots, g_s\}$ with $0 = g_1 < g_2 < \cdots < g_s$ be an s-mark Golomb ruler and $D = \{g_j - g_i | i < j\}$. Assume that an integer M satisfies three conditions mentioned in Theorem 1. Following variations of the construction in Theorem 1 are possible and they all result in a 5-seq LRC with the same parameter $(n = sM, k = sM - 2M + 1, r = s - 1, t = 2)$.

1) Constant addition to the first row of E in (7).
2) Constant addition to the second row of E in (7).
3) Constant multiplication to the second row of E in (7). Here, the constant must be relatively prime to M.
4) The first row of E in (7) is added by the Golomb ruler $g_1, g_2, \ldots, g_s$ and the second row is multiplied by a constant y where $y - 1$ is relatively prime to M.

Proof:

1) When some positive constant c is added to the first row, the resulting new exponent matrix becomes the following:

$$\begin{pmatrix} c & c & c & \cdots & c \\ 0 & g_2 & g_3 & \cdots & g_s \end{pmatrix}.$$

Since the second row G is unchanged, we use the same integer M satisfying the three conditions. Then, the resulting parity check matrix becomes the following:

$$\begin{pmatrix} I^{(c)} & I^{(c)} & \cdots & I^{(c)} \\ I & I^{(g_2)} & \cdots & I^{(g_s)} \end{pmatrix}.$$

Now, it is obvious that the remaining proof works the the same as the proof of Theorem 1. Similarly, for the case of a negative constant c.

2) New second row is the same Golomb ruler except that marks are shifted by the constant. Therefore, the same integer M can be used for the construction and the remaining proof works the same.

3) This construction is the same as that in Theorem 1 with the Golomb ruler $G' = \{xg_1, xg_2, \ldots, xg_s\}$ in (8). Therefore, the set D' of positive distances is xD where $G = \{g_1, g_2, \ldots, g_s\}$ with D. We now have to check whether the parameter M for G satisfying **M1**, **M2** and **M3** also satisfies these three conditions for the new Golomb ruler $G' = xG$ with $D' = xD$. This is straightforward and the remaining part of the proof works the same.

4) Consider the exponent matrix of the following form:

$$\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & (y-1)g_2 & (y-1)g_3 & \cdots & (y-1)g_s \end{pmatrix}.$$

Here, the second row is a Golomb ruler since $y - 1$ is relatively prime to M. Therefore, we may use the same integer M from here to construct the parity check matrix of the form:

$$\begin{pmatrix} I & I & \cdots & I \\ I & I^{((y-1)g_2)} & \cdots & I^{((y-1)g_s)} \end{pmatrix}. \quad (11)$$

Now, we shift first M columns cyclically to the left by $g_1 = 0$, and then shift next M columns cyclically to the left by g_2, etc., and finally shift final M columns cyclically to the left by g_s, and obtain the following:

$$\begin{pmatrix} I^{(g_1)} & I^{(g_2)} & \cdots & I^{(g_s)} \\ I^{(yg_1)} & I^{(yg_2)} & \cdots & I^{(yg_s)} \end{pmatrix}.$$

This is the same as the one obtained by substituting appropriate CPM of size $M \times M$ to the proposed exponent matrix of the item 4) in the beginning. Therefore, using the parity check matrix in (11), the proof works the same as in the proof of Theorem 1. ■

■

Example 1: Use the Golomb ruler $\{0, 1, 4, 6\}$ and select a positive integer $M = 13$ satisfying **M1**, **M2** and **M3**. Then the following are some pairs of E and H from the above discussion. Here, $I^{(\lambda)}$ is the $M \times M$ identity matrix circularly left-shifted by λ.

1) Theorem 1 gives the following:

$$E = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 4 & 6 \end{pmatrix}, \ H = \begin{pmatrix} I & I & I & I \\ I & I^{(1)} & I^{(4)} & I^{(6)} \end{pmatrix}.$$

2) Corollary 1 with $c = 2$ in item 3) gives the following:

$$E = \begin{pmatrix} 2 & 2 & 2 & 2 \\ 0 & 1 & 4 & 6 \end{pmatrix}, \ H = \begin{pmatrix} I^{(2)} & I^{(2)} & I^{(2)} & I^{(2)} \\ I & I^{(1)} & I^{(4)} & I^{(6)} \end{pmatrix}.$$

All the linear codes defined by the above 26×52 matrices H are 5-seq $(n = 52, k = 27, r = 3)$-LRC with availability $t = 2$ and the repair time at most 3.

In addition, we observe that these codes have rate $27/52$, which attains the rate bound (1).

We now discuss some simple sufficient conditions for the three conditions **M1**, **M2** and **M3** for the positive integer M with respect to a given Golomb ruler G.

The first condition **M1** is that

$$g_i \neq g_j \pmod{M} \quad \text{for all} \quad i \neq j.$$

We simply take M to be bigger than the length of the Golomb ruler, which is $g_s - g_1$, which is g_s when $g_1 = 0$. This guarantees that this condition is satisfied with any G. In the following, since we have assumed in the beginning

TABLE I
5-SEQ LRCs FROM **Theorem 1** USING SOME GOLOMB RULERS

s	s-mark Golomb ruler	smallest $M > g_s$ for **M2** and **M3**	n	k	code rate	rate bound (1)
3	$0, 1, 3$	$2g_s < 7$	21	8	0.38095	–
4	$0, 1, 4, 6$	$2g_s < 13$	52	27	0.51923	$=$ 0.51923
	$0, 2, 3, 7$	$13 < 2g_s$				
5	$0, 1, 4, 9, 11$	$2g_s < 23$	115	70	0.60870	$<$ 0.60952
	$0, 2, 7, 8, 11$	$21 < 2g_s$	105	64	0.60952	$=$ 0.60952
6	$0, 1, 4, 10, 12, 17$	$31 < 2g_s$	186	125	0.67204	$=$ 0.67204
	$0, 1, 3, 8, 12, 18$					
7	$0, 2, 3, 10, 16, 21, 25$	$49 < 2g_s$	343	246	0.71720	$<$ 0.71761
	$0, 2, 7, 13, 21, 22, 25$					
8	$0, 1, 4, 9, 15, 22, 32, 34$	$2g_s < 69$	552	415	0.75181	$<$ 0.75219
	$0, 4, 5, 17, 19, 25, 28, 35$	$57 < 2g_s$	465	343	0.75219	$=$ 0.75219
	$0, 1, 3, 13, 32, 36, 43, 52$					
9	$0, 1, 5, 12, 25, 27, 35, 41, 44$	$2g_s < 89$	801	624	0.77903	$<$ 0.77930
	$0, 2, 10, 24, 25, 29, 36, 42, 45$	$73 < 2g_s$	657	512	0.77930	$=$ 0.77930
	$0, 1, 3, 7, 15, 31, 36, 54, 63$					
10	$0, 1, 6, 10, 23, 26, 34, 41, 53, 55$	$91 < 2g_s$	910	729	0.80110	$=$ 0.80110
	$0, 1, 3, 9, 27, 49, 56, 61, 77, 81$					

that $g_1 = 0$, we will assume the parameter $M > g_s$ unless otherwise specified.

The second condition **M2** is the following:

$$d + d' \not\equiv 0 \pmod{M} \quad \text{for all} \quad d, d' \in D$$

where we allow $d = d'$. Note that **M2** is satisfied when $M > 2g_s$, since $d \le g_s$ for $d \in D$. The value $M = 2g_s$ will not satisfy **M2**, since $g_s + g_s = 0 \pmod{M}$, where $g_s = g_s - g_1 \in D$. Some values in the range $g_s < M < 2g_s$ could satisfy the condition also, and this must be checked individually.

The third condition **M3** is that all the members in D and M are collectively relatively prime. This condition will be trivially satisfied for any M if all the members in D themselves are already relatively prime. We simply point out here some obvious sufficient conditions on D for any positive integer M in **M3**:

1) D contains 1; or
2) D contains $M - 1$; or
3) D contains two integers that are relatively prime.

In the following, we discuss the importance of choosing the smallest possible value of $M > g_s$ to achieve the best possible (largest) code rate from Theorem 1.

Corollary 2: Let $M < M'$ be two positive integers satisfying **M1**, **M2** and **M3** for a given Golomb ruler. The code rate of the code from Theorem 1 using the value M becomes larger than those using M'.

Proof:

$$\frac{sM - (2M-1)}{sM} - \frac{sM' - (2M'-1)}{sM'} = \frac{M' - M}{sMM'} > 0.$$

■

By Corollary 2, we are interested in the smallest M that satisfies **M1**, **M2** and **M3** for a given Golomb ruler. It is interesting to see that some cases are rate-optimal and some others are not. Table I shows the smallest possible M in the range $M > g_s$ and the resulting code rate from the construction of Theorem 1.

Remark 1: From Table I, it can be observed that the decrease in the length g_s of the Golomb ruler does not necessarily decrease in the smallest M for the construction. For example, for $s = 8$ in Table I, the Golomb ruler with $g_s = 35$ or 52 has the smallest $M = 57$, while those with $g_s = 34$ has the smallest $M = 69 > 57$ and results in a non-rate-optimal LRC.

Remark 2: All of the non-rate-optimal codes in Table I will satisfy some other optimality, called dimension-optimality. We will discuss this in detail in Subsection III-C.

B. Condition on the Golomb Rulers for Rate-Optimality

We identify the necessary condition for the rate-optimal examples shown in Table I. Recall that the rate bound (1) can be written with $r = s - 1$ and $u = 5$ (so $\sigma = 2$) in our construction as,

$$\frac{k}{n} \le \frac{r^3}{r^3 + 2(r + r^2) + 1} = \frac{s^3 - 3s^2 + 3s - 1}{s(s^2 - s + 1)}.$$

On the other hand, the code rate of the 5-seq LRC from Theorem 1 is given as

$$\frac{k}{n} = \frac{sM - (2M-1)}{sM},$$

where the positive integer s comes from the number of marks of the chosen Golomb ruler G and a positive integer M is selected at random but it has to satisfy three conditions **M1**,

M2 and **M3** with G. If the code is rate-optimal, then

$$\frac{sM-2M+1}{sM}=\frac{s^3-3s^2+3s-1}{s(s^2-s+1)},$$

which implies that

$$M=s^2-s+1. \tag{12}$$

Conversely, if, for an s-mark Golomb ruler G, the value $M=s^2-s+1$ satisfies the three conditions **M1**, **M2** and **M3** with G, then the resulting 5-seq LRC from Theorem 1 is rate-optimal. Furthermore, the rate-optimality implies that the value M must be the smallest possible due to Corollary 2. This can be summarized as:

Lemma 1: The 5-seq LRC from Theorem 1 will be rate-optimal if and only if $M=s^2-s+1$ is the (smallest possible) positive integer satisfying **M1**, **M2** and **M3** with the s-mark Golomb ruler G for the construction.

Remark 3: The rate of the code from Theorem 1 can be stated as

$$\frac{k}{n}=\frac{sM-(2M-1)}{sM}>\frac{s-2}{s}.$$

Now, we would like to characterize those s-mark Golomb rulers G so that the value $M=s^2-s+1$ satisfies **M1**, **M2** and **M3** with respect to G. We will eventually prove that ONLY the s-mark Golomb rulers from "some" (m,s)-modular Golomb rulers satisfy this property.

In the meantime, we would like to note that, for an s-mark Golomb ruler G, the value $M=s^2-s+1$ may not always satisfy the condition **M2**. An example is the 4-mark Golomb ruler $\{0,1,4,9\}$ with $D=\{1,3,4,5,8,9\}$. Here, $s^2-s+1=13$ and $8+5\equiv 0 \pmod{13}$.

An (m,s)-*modular Golomb ruler* (MGR) is a set of s residues, $g_1,g_2,\ldots,g_s \pmod m$, such that all the differences $g_i-g_j \pmod m$ for $i\neq j$ are distinct and nonzero [1], [6], [8], [9], [28], [33]. If we take an (m,s)-MGR $G \pmod m$ as a straight integer set in the range from 0 to $m-1$, the result becomes an s-mark Golomb ruler. We will use this method of getting an s-mark Golomb ruler from an (m,s)-MGR in this paper. We also assume $3\le s<m$ for some non-triviality.

The following three types are well-known systematic constructions for (m,s)-MGRs [1], [28], [33]:

Known-fact 3: 1) (Singer [33]) For any prime power q, there is a $(q^2+q+1,q+1)$-MGR.

2) (Bose [1]) For any prime power q, there is a (q^2-1,q)-MGR.

3) (Ruzsa [28]) For any prime q, there is a $(q^2-q,q-1)$-MGR.

Consider any (m,s)-MGR and its straight version of the s-mark Golomb ruler G. We will claim that the value $M=m$ satisfies **M1** and **M2** with respect to G.

Consider an MGR $\{g_1,g_2,\ldots,g_s \pmod m\}$ of distinct s residues mod m and its straight version $G=\{g_1,g_2,\ldots,g_s\}$ where we assume that $0\le g_i<m$ for all i, and that $g_1<g_2<\cdots<g_s$. Since an MGR contains distinct residues mod m, we see that $g_i\neq g_j \pmod m$, which is **M1** with $M=m$ for G. For **M2** we recall the defining property of an MGR: $g_i-g_j \pmod m$ are all distinct residues for $i\neq j$. Recall the definition of $D=\{g_j-g_i|i<j\}$ which contains all the positive distances of the (straight) G. Since G is a Golomb ruler, we know that D contains all distinct positive integers. When we denote by D_m the set of residues $d \pmod m$ for each integer $d\in D$, then the defining property of an MGR implies that $-D_m\cup D_m$ contains all distinct residues mod m, where $-D_m$ is the set of negative residues of the members of D_m. Therefore, D_m and $-D_m$ are disjoint. If $d+d'=0 \pmod m$ for some $d,d'\in D_m$, then $d=-d' \pmod m$ with $d\in D_m$ and $-d'\in -D_m$, which is impossible, which implies **M2** with $M=m$. This can be summarized as

Lemma 2: We consider an s-mark Golomb ruler G from an (m,s)-MGR by taking the marks as straight integers in the range from 0 to $m-1$. Then, the value $M=m$ satisfies **M1** and **M2** with respect to G.

Now, if we take $M=m$ from an (m,s)-MGR, then we only have to check **M3** with respect to its straight version G. We are able to find a necessary and sufficient condition on (m,s)-MGRs for the value $M=m$ to satisfy **M3**. They are the MGRs which come from the $(m,s,1)$-cyclic difference set [6], and mentioned as an item 1) of **Known-fact 3**.

The $(m,s,1)$-cyclic difference set is the set of s residues $g_1,g_2,\ldots,g_s \pmod m$ such that the set of all the differences $g_i-g_j \pmod m$ is the set of all the non-zero residues mod m [4]. It is also called a cyclic planar difference set [4]. It is well-known that the $(m,s,1)$-cyclic difference set is equivalent to an (m,s)-MGR [6]. One of the well-known $(m,s,1)$-cyclic difference set comes from Singer with parameters $(q^2+q+1,q+1,1)$ when q is a power of a prime [4], [8], [9], [33]. If we let $q+1=s$, then $q^2+q+1=s^2-s+1$ and the parameters become

$$(q^2+q+1,q+1,1)=(s^2-s+1,s,1).$$

We are considering exactly those s-mark Golomb rulers from (m,s)-MGRs that are equivalent to $(m,s,1)$-cyclic difference sets with $m=s^2-s+1$.

Theorem 2: The resulting 5-seq LRC from Theorem 1 using an s-mark Golomb ruler G and some positive integer M is rate-optimal if and only if $G \pmod M$ with $M=s^2-s+1$ is an $(M,s,1)$-cyclic difference set.

Proof: We first note that $M=s^2-s+1$ by Lemma 1. We also note that G and M used in Theorem 1 must satisfy **M2**. For the necessary condition, recall the notation that $D=\{g_j-g_i|i<j\}$ is the set of positive differences of marks of G. Observe that $\frac{M-1}{2}=\frac{s(s-1)}{2}=|D|$.

Claim that

$$D\cup(-D)\equiv\{1,2,\ldots,M-1\} \pmod M.$$

If $-d\in D \pmod M$, then $d+(-d)=0 \pmod M$ which is impossible by **M2**. Therefore, $D \pmod M$ and $-D \pmod M$ are disjoint subsets of the integers mod M, each of size $(M-1)/2$, and hence, their union becomes all the non-zero residues mod M. This proves that $G \pmod M$ is an $(M,s,1)$-cyclic difference set.

For sufficiency, we have already done for M to satisfy **M1** and **M2** in Lemma 2, since an $(M,s,1)$-cyclic difference set is an (M,s)-MGR and G is its straight integer version. Now, the fact that $G \pmod{M}$ is an $(M, s, 1)$-cyclic difference set implies that D contains 1 or $M-1$, and hence the value M also satisfies **M3**. Now, it is known that if an (M, s)-MGR exists, then $M \geq s^2-s+1$ [6]. Therefore, the value $M = s^2-s+1$ is the smallest possible value satisfying all three conditions **M1**, **M2** and **M3**, and hence, the 5-seq LRC from Theorem 1 with this G and M becomes rate-optimal. ■

It turned out that the Golomb rulers in Table I for rate-optimal codes are all derived from some $(s^2-s+1, s, 1)$ cyclic difference sets, and the smallest value of M turns out to be $M = s^2-s+1$. These are the cases $s = 4, 5, 6, 8, 9, 10$ with $M = 13, 21, 31, 57, 73, 91$, respectively.

The case $s = 7$ in Table I is particularly interesting. It is known that a $(43, 7, 1)$-cyclic difference set does not exist since $s-1 = 6$ is not a prime power [16]. Therefore, one cannot construct a 7-mark Golomb ruler from a $(43, 7, 1)$-cyclic difference set. On the other hand, there are lots of 7-mark Golomb rulers in general, two of which are selected in Table I, both of which give some non-optimal codes with $M = 49 > 43$.

We now apply the s-mark Golomb rulers from the (m, s)-MGRs in the items 2) and 3) of **Known-fact 3**. The value $M = m$ in both cases is known to satisfy **M1** and **M2** by Lemma 2. We confirmed that it also satisfies **M3**.

Table II shows the application of s-mark Golomb rulers with $M = m$ from the items 2) and 3) of **Known-fact 3** [1], [28]. When M is chosen as $m = s^2-1$ or s^2+s, the resulting code rate cannot achieve the bound (1). This is actually the result we expected, since $s^2+s > s^2-1 > s^2-s+1$. However, if there exists a smallest M that is smaller than s^2-1 or s^2+s while satisfying **M2** and **M3** for the corresponding Golomb rulers, it would be possible to generate a code with a code rate closer to the bound (1). For example, when $s = 4$ in Ruzsa's type, the $m = 20$ from the parameter (m, s). However, the smallest M that satisfies **M2** and **M3** is $13 = 4^2-4+1$, which yields a rate-optimal code. In this case, $\{0, 1, 3, 9\} \pmod{M = 13}$ indeed forms a $(13, 4, 1)$-cyclic difference set.

C. Condition for Dimension-Optimality

It is obvious that the rate-optimal LRC is dimension-optimal. Therefore, an easy and obvious sufficient condition for dimension-optimality is the same condition for the rate-optimality. We note that some non-rate-optimal LRC can be dimension-optimal. Therefore, it would be interesting to find some general conditions for the code (whether it is rate-optimal or not) from Theorem 1 to be dimension-optimal. Recall that the dimension bound (2) can be written with $r = s-1$ and $u = 5$ (so $\sigma = 2$) in our construction as,

$$k \leq \left\lfloor \frac{nr^3}{r^3+2r^2+2r+1} \right\rfloor = \left\lfloor \frac{n(s-1)^3}{s(s^2-s+1)} \right\rfloor.$$

On the other hand, the 5-seq LRC according to Theorem 1 has length $n = sM$ and dimension $k = sM - 2M + 1$. Therefore,

$$\begin{aligned} k &= sM - 2M + 1 \\ &\leq \left\lfloor \frac{sM(s-1)^3}{s(s^2-s+1)} \right\rfloor \\ &= sM - 2M + \left\lfloor \frac{M}{(s^2-s+1)} \right\rfloor, \end{aligned}$$

where the equality holds if and only if

$$s^2-s+1 \leq M < 2(s^2-s+1). \tag{13}$$

Theorem 3: The resulting 5-seq LRC from Theorem 1 is dimension-optimal if and only if the selected positive integer M in the construction is in the range of (13).

We remark that the equality $M = s^2-s+1$ in (13) is the necessary condition for rate-optimality as discussed in Lemma 1.

Recall that the each of the values $M = m, m+1, \ldots, 2(s^2-s+1)-1$ where m is from (m, s)-MGR in the items 2) and 3) of **Known-fact 3** satisfies **M1**, **M2** and **M3** with its straight version s-mark Golomb ruler. For the value $m = s^2-1$ or $m = s^2+s$, the above values of M are all in the range (13), the resulting codes from Ruzsa and Bose type (m, s)-MGRs with any of the above M are all dimension-optimal. The cases with $M = m$ are shown in Table II.

Finally, we would like to argue that, according to Remark 2, when the LRC is not rate-optimal but dimension-optimal, it can also be called rate-optimal in general in the sense that there does not exist an LRC of larger dimension and hence there does not exist an LRC of larger rate for the given length. In particular, using the $(s^2-s+1, s, 1)$-cyclic difference set, the resulting codes with any M in the range (13) except for $M = s^2-s+1$ are not rate-optimal but dimension-optimal (Theorems 2 and 3). All these cases also can be said to be rate-optimal in general in the same sense.

D. Some Variations for Availability 3

We now construct 5-seq LRCs with availability 3 by combining several non-trivial variations of Theorem 1, ensuring that 5 erased symbols can be recovered with at most 2 repair time.

Theorem 4: Let $\{g_1, g_2, \ldots, g_s\}$ with $0 = g_1 < g_2 < \cdots < g_s$ be an s-mark Golomb ruler and $D = \{g_j - g_i | i < j\}$. Let $M > g_s$ be a positive integer that satisfies **M2**. Consider a positive integer $2 \leq x < M$ that satisfies

$$gcd(x-1, M) = gcd(x, M) = 1. \tag{14}$$

Let E be the $3 \times s$ integer matrix of the form

$$E = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & g_2 & \cdots & g_s \\ 0 & xg_2 & \cdots & xg_s \end{pmatrix}. \tag{15}$$

Now, construct a binary $3M \times sM$ matrix H by substituting a circular permutation matrix (CPM) of size $M \times M$ into the position (i, j) of E for all i, j, where $i = 0, 1, 2$ and $j = 0, 1, \ldots, s-1$. Here, each CPM for the position (i, j) is obtained by taking the circular left-shift of the columns of the identity matrix by the integer $e(i, j)$.

Then, H becomes a parity check matrix of a binary linear (n, k) code with length $n = sM$ and dimension $k \geq sM - 3M + 2$. Furthermore, this binary linear (n, k) code becomes a 5-seq LRC with availability 3, locality $r = s-1$ and the repair time at most 2.

TABLE II
5-SEQ LRCS FROM **Theorem 1** USING SOME GOLOMB RULERS GENERATED BY VARIOUS MODULAR GOLOMB RULERS

type	(m,s)-MGR	s-mark Golomb ruler from (m,s)-MGR	smallest $M > g_s$ for **M2** and **M3**	M used for construction	n	k	dimension bound (2)	code rate	rate bound (1)
Bose [1]	$(8,3)$	$0,1,3$	$2g_s < 7 = M$	$m = 8$	24	9	–	0.37500	–
				7	21	8	–	0.38095	–
Bose [1]	$(15,4)$	$0,1,3,7$	$2g_s < 15 = M$	15	60	31	$= 31$	0.51667	$<$ 0.51923
Bose [1]	$(24,5)$	$0,1,4,9,11$	$2g_s < 23 = M$	$m = 24$	120	73	$= 73$	0.60833	$<$ 0.60952
				23	115	70	$= 70$	0.60870	$<$ 0.60952
Bose [1]	$(48,7)$	$0,5,7,18,19,22,28$	$M = 48 < 2g_s$	48	336	241	$= 241$	0.71726	$<$ 0.71761
Bose [1]	$(63,8)$	$0,2,8,21,22,25,32,37$	$M = 63 < 2g_s$	63	504	379	$= 379$	0.75198	$<$ 0.75219
Ruzsa [28]	$(20,4)$	$0,1,3,9$	$M = 13 < 2g_s$	$m = 20$	80	41	$= 41$	0.51250	$<$ 0.51923
				13	52	27	$= 27$	0.51923	$=$ 0.51923
Ruzsa [28]	$(42,6)$	$0,1,3,11,16,20$	$2g_s < 41 = M$	$m = 42$	252	169	$= 169$	0.67063	$<$ 0.67204
				41	246	165	$= 165$	0.67073	$<$ 0.67204
Ruzsa [28]	$(110,10)$	$0,13,16,17,25,31,52,54,59,78$	$M = 110 < 2g_s$	110	1100	881	$= 881$	0.80091	$<$ 0.80110

TABLE III
COMPARISON OF THE VARIOUS PARAMETERS OF THE 5-SEQ LRCS

Reference	Parameters n	k	r	t	repair time	Conditions	Comments
[3]	$21k/8$	k	2			$8 \mid k$	rate $= \frac{8}{21}$
	r^2+4r+1	r^2	≥ 3			two $r \times r$ orthogonal Latin squares	n,k are fixed from r ; rate $> \frac{r-4}{r}$ eventually
[2]	$a_0(r^2+r+1)$	$a_0 r^2 - a_0 r \frac{r}{r+1}$	≥ 3			$a_0 \geq r+1$	$r+1 \mid a_0$ for rate-optimality
[42]	$(r+1)v$	$\geq (r-1)v+1$	≥ 2	2	3	A $(v, \geq r+1, 1)$ symmetric block design	v and r are tightly connected by the symmetric block design
	$(r+1)^2$	$\geq r(r-1)$	≥ 4	3	3	An $(r+1)\times(r+1)$ Latin square not containing a 2×2 sublatin square	n,k are fixed from r
This paper: Main Construction	sM	$sM-2M+1$	≥ 2	2	3	An s-mark Golomb ruler and a positive integer M with three conditions. (Theorem 1)	rate-optimal iff $M = s^2-s+1$ and G from $(M,s,1)$-CDS. (Theorem 2)
							dimension-optimal iff $(s^2-s+1) \leq M < 2(s^2-s+1)$. (Theorem 3)
							rate $> \frac{s-2}{s}$ (Remark 2)
This paper: some variation	sM	$\geq sM-3M+2$	≥ 2	3	2	similar to above. (Theorem 4)	rate $> \frac{s-3}{s}$ (Remark 4)

Proof: The matrix H in the theorem becomes

$$H = \begin{pmatrix} H_1 \\ H_2 \\ H_3 \end{pmatrix} = \begin{pmatrix} I^{(0)} & I^{(0)} & \cdots & I^{(0)} \\ I^{(0)} & I^{(g_2)} & \cdots & I^{(g_s)} \\ I^{(0)} & I^{(xg_2)} & \cdots & I^{(xg_s)} \end{pmatrix} \quad (16)$$

which defines an LRC of length $n = sM$ and locality $r = s-1$.

Now, we claim that $\mathrm{rank}(H) \leq 3M-2$. For $1 \leq i \leq 3$, each row-block H_i contains exactly single 1 in each column. Therefore, the sum of all the rows in each H_i results in the all-one vector. From the proof of the dimension in Theorem 1, the first row of H_2 is the sum of all other rows in their row-block plus the sum of all the rows of H_1. So is the first row of H_3, similarly. Therefore, $3M-2$ rows from H except for, for example, the first rows of H_2 and H_3, can span the whole row space of H. Therefore, the $\mathrm{rank}(H) \leq 3M-2$ and thus $k \geq sM-(3M-2)$.

Now, we will show the availability $t = 3$. We need to show that any two rows in H satisfies condition (6). According to Theorem 1, the submatrix composed of H_1 and H_2 satisfies condition (6). By Corollary 1, the submatrix composed of H_2 and H_3 satisfies condition (6). By Corollary 1, the submatrix composed of H_1 and H_3 satisfies condition (6). Therefore, H in (16) satisfies the condition (6).

Now, the LRC by only the H_1 and H_2 is already a 5-seq LRC, so is the code by H. That is, inclusion of H_3 only adds up some additional (disjoint) repair sets so that the repair time is decreased to at most 2. ■

Corollary 3: Any non-zero constant can be used in the first row of the exponent matrix (15) in Theorem 4.

Remark 4: The rate of the code from Theorem 4 can be stated as

$$\frac{k}{n} \geq \frac{sM - 3M + 2}{sM} = \frac{s-3}{s} + \frac{2}{sM} > \frac{s-3}{s}.$$

Remark 5: To achieve the rate upper bound in (1), Corollary 3 of [2] implies that each column of H must have a weight of 1 or 2, which implies that ANY sequential-recovery LRC with availability 3 cannot achieve this bound. As a result, the resulting LRC from Theorem 4 does not achieve the upper bound in (1). Its dimension-optimality is not discussed only because its dimension cannot be exactly determined.

IV. Concluding Remarks

We will review some previous constructions for 5-seq LRCs and compare these with the proposed constructions in this paper.

In 2016 [3], u-seq LRCs for $u = 4, 5, 6, 7$ with $r = 2$ have been constructed in a graphical form (First row of Table III). The first construction is restricted to the cases $r = 2$ and $8|k$ and the rate $8/21$. The second construction designs u-seq LRC for $r \geq 3$ when there exist two or more $r \times r$ orthogonal Latin squares. Here, the length and the dimension are determined by the value r. Both constructions result in u-seq LRCs without any availability.

In 2020 [2], Some parity-check matrices are constructed using the tree-like graphs with girth $\geq u+1$, which are, in fact, u-seq LRCs (Second row of Table III). For $u = 5$, to achieve the rate bound (1), the code length must be a multiple of $r^2 + r + 1$, and the parameter a_0 in the table must be a multiple of $r + 1$. These constructions also result in u-seq LRCs without any availability. The rate-optimal 5-seq LRCs proposed in this paper (when $a_0 = r+1$) have the same length n as the proposed rate-optimal code (when $M = s^2 - s + 1$ and $s = r + 1$) from Theorems 1 and 2.

In 2020 [42], the concept of joint sequential-parallel recovery LRCs is first introduced (Third row of Table III). The first construction produces 5-seq LRCs with availability 2 using a Latin rectangle derived from a $(v, \geq r+1, 1)$ symmetric block design. Here, the authors left the determination of the exact dimension for future work and mentioned, based on experimental results, that if $(r^2 + r + 1, r + 1, 1)$ symmetric block designs exist, it would be possible to generate rate-optimal codes with some shorter block lengths. The second construction produces 5-seq LRCs with availability 3 using a Latin square not containing a 2×2 sublatin square. Here, the length is determined from r as $n = (r + 1)^2$.

Besides the comments in the last column of Table III, some other properties of the proposed constructions (Fourth row of Table III) can be summarized as follows:

1) The proposed construction achieves a higher code rate (almost $\frac{s-2}{s}$) than those from [3]. It will be higher if a smaller M can chosen.
2) It produces the code with availability 2 or 3 in a simpler manner than those from [2].
3) It has much more flexible choice for the length since there are a lot of choices for the value M than those from [42].
4) We provide necessary and sufficient conditions for the proposed code to be rate-optimal and/or dimension-optimal by calculating the exact value of the dimension of the code.
5) The proposed construction uses any s-mark Golomb ruler, which is not necessarily an optimal one (shortest Golomb ruler). There are a lot of s-mark Golomb rulers for a given integer $s \geq 3$, and the length g_s does not matter expect for the conditions for M.

References

[1] R. C. Bose, "An affine analogue of singer's theorem," *Ind. Math. Soc.*, vol. 6, pp. 1–5, Jan. 1942.

[2] S. B. Balaji, G. R. Kini, and P. V. Kumar, "A tight rate bound and matching construction for locally recoverable codes with sequential recovery from any number of multiple erasures," *IEEE Trans. Inf. Theory*, vol. 66, no. 2, pp. 1023–1052, Feb. 2020.

[3] S. B. Balaji, K. P. Prasanth, and P. V. Kumar, "Binary codes with locality for multiple erasures having short block length," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 655–659.

[4] L. D. Baumert, *Cyclic Difference Sets*. New York, NY, USA: Springer, 1971.

[5] G. S. Bloom and S. W. Golomb, "Applications of numbered undirected graphs," *Proc. IEEE*, vol. 65, no. 4, pp. 562–570, Apr. 1977.

[6] M. Buratti and D. R. Stinson, "New results on modular Golomb rulers, optical orthogonal codes and related structures," 2020, *arXiv:2007.01908*.

[7] H. Choi, Z. Jing, G. Kim, and H.-Y. Song, "Some intersections of two binary LRCs with disjoint repair groups," in *Proc. 10th Int. Workshop Signal Design Appl. Commun. (IWSDA)*, Aug. 2022, pp. 1–5.

[8] C. J. Colbourne and J. H. Dinitz, *Handbook of Combinatorial Designs*, 2nd ed., Boca Raton, FL, USA: CRC Press, 2007.

[9] A. Dimitromanolakis, "Analysis of the Golomb ruler and the Sidon set problems, and determination of large, near-optimal Golomb rulers," M.S. thesis, Dept. Electron. Comput. Eng., Tech. Univ. Crete, Chania, Greece, 2002.

[10] A. Dollas, W. T. Rankin, and D. McCracken, "A new algorithm for Golomb ruler derivation and proof of the 19 mark ruler," *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 379–382, Jan. 1998.

[11] K. Drakakis, "A review of the available construction methods for Golomb rulers," *Adv. Math. Commun.*, vol. 3, no. 3, pp. 235–250, 2009.

[12] P. Erdös and P. Turán, "On a problem of Sidon in additive number theory, and on some related problems," *J. London Math. Soc.*, vols. s1–16, no. 4, pp. 212–215, Oct. 1941.

[13] M. P. C. Fossorier, "Quasi-cyclic low-density parity-check codes from circulant permutation matrices," *IEEE Trans. Inf. Theory*, vol. 50, no. 8, pp. 1788–1793, Aug. 2004.

[14] M. Gardner, "Mathematical games: The graceful graphs of Solomon Golomb, or how to number a graph parsimoniously," *Sci. Amer.*, vol. 226, no. 3, pp. 108–113, Mar. 1972.

[15] 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.

[16] D. M. Gordon, "The prime power conjecture is true for $n < 2,000,000$," *Electron. J. Combinatorics*, vol. 1, no. 1, pp. 1–7, Aug. 1994.

[17] Z. Jing and H.-Y. Song, "Some new sequential-recovery LRCs based on good polynomials," *IEEE Access*, vol. 10, pp. 42844–42850, 2022.

[18] Z. Jing and H.-Y. Song, "Girth-based sequential-recovery LRCs," *IEEE Access*, vol. 10, pp. 126156–126160, 2022.

[19] W. Kim, H. Cho, H.-Y. Song, and M. K. Song, "Some variations of Tanner's construction for short length QC-LDPC codes," *Electron. Lett.*, vol. 60, 2024, Art. no. e13088, doi: 10.1049/ell2.13088.

[20] I. Kim, T. Kojima, and H.-Y. Song, "Some short-length Girth-8 QC-LDPC codes from primes of the form $t^2 + 1$," *IEEE Commun. Lett.*, vol. 26, no. 6, pp. 1211–1215, Jun. 2022.

[21] I. Kim and H. Song, "A construction for girth-8 QC-LDPC codes using Golomb rulers," *Electron. Lett.*, vol. 58, no. 15, pp. 582–584, Jul. 2022.

[22] J.-H. Kim and H.-Y. Song, "Alphabet-dependent bounds for locally repairable codes with joint information availability," *IEEE Commun. Lett.*, vol. 21, no. 8, pp. 1687–1690, Aug. 2017.

[23] B. Lindström, "On B2-sequences of vectors," *J. Number Theory*, vol. 4, no. 3, pp. 261–265, Jun. 1972.

[24] J. Ma and G. Ge, "Optimal binary linear locally repairable codes with disjoint repair groups," *SIAM J. Discrete Math.*, vol. 33, no. 4, pp. 2509–2529, Dec. 2019.

[25] D. S. Papailiopoulos and A. G. Dimakis, "Locally repairable codes," *IEEE Trans. Inf. Theory*, vol. 60, no. 10, pp. 5843–5855, Oct. 2014.

[26] N. Prakash, V. Lalitha, S. B. Balaji, and P. Vijay Kumar, "Codes with locality for two erasures," *IEEE Trans. Inf. Theory*, vol. 65, no. 12, pp. 7771–7789, Dec. 2019.

[27] N. Prakash, V. Lalitha, and P. V. Kumar, "Codes with locality for two erasures," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Sep. 2014, pp. 1962–1966.

[28] I. Ruzsa, "Solving a linear equation in a set of integers I," *Acta Arithmetica*, vol. 65, no. 3, pp. 259–282, 1993.

[29] W. T. Rankin, "Optimal Golomb rulers: An exhaustive parallel search implementation," Ph.D. dissertation, Dept. Elect. Eng., Duke Univ., Durham, NC, USA, 1993.

[30] A. S. Rawat, A. Mazumdar, and S. Vishwanath, "On cooperative local repair in distributed storage," *EURASIP J. Adv. Signal Process.*, vol. 2015, no. 1, p. 107, Dec. 2015.

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

[32] J. B. Shearer, "Some new optimum Golomb rulers," *IEEE Trans. Inf. Theory*, vol. 36, no. 1, pp. 183–184, Jan. 1990.

[33] J. Singer, "A theorem in finite projective geometry and some applications to number theory," *Trans. Amer. Math. Soc.*, vol. 43, no. 3, pp. 377–385, 1938.

[34] W. Song, K. Cai, C. Yuen, K. Cai, and G. Han, "On sequential locally repairable codes," *IEEE Trans. Inf. Theory*, vol. 64, no. 5, pp. 3513–3527, May 2018.

[35] W. Song and C. Yuen, "Locally repairable codes with functional repair and multiple erasure tolerance," 2015, *arXiv:1507.02796*.

[36] W. Song and C. Yuen, "Binary locally repairable codes–sequential repair for multiple erasures," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Washington, DC, USA, Dec. 2016, pp. 1–6.

[37] Y.-S. Su, "On the construction of local parities for (r, t) -Availability in distributed storage," *IEEE Trans. Commun.*, vol. 65, no. 6, pp. 2332–2344, Jun. 2017.

[38] I. Tamo and A. Barg, "Bounds on locally recoverable codes with multiple recovering sets," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2014, pp. 691–695.

[39] 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.

[40] A. Wang, Z. Zhang, and D. Lin, "Bounds for binary linear locally repairable codes via a sphere-packing approach," *IEEE Trans. Inf. Theory*, vol. 65, no. 7, pp. 4167–4179, Jul. 2019.

[41] A. Wang, Z. Zhang, and M. Liu, "Achieving arbitrary locality and availability in binary codes," in *Proc. IEEE Int. Symp. Inf. Theory*, Hong Kong, Jun. 2015, pp. 1866–1870.

[42] E. Yavari and M. Esmaeili, "Locally repairable codes: Joint sequential–parallel repair for multiple node failures," *IEEE Trans. Inf. Theory*, vol. 66, no. 1, pp. 222–232, Jan. 2020.

[43] E. Yavari, M. Esmaeili, and J. Rifà, "Locally repairable codes based on permutation cubes and Latin squares," *IEEE Access*, vol. 10, pp. 9686–9702, 2022.

Hyojeong Choi (Graduate Student Member, IEEE) received the B.S. degree in communication and information engineering and the M.S.E.E. degree in electrical engineering from Yonsei University, Seoul, South Korea, in 2018 and 2021, respectively, where she is currently pursuing the Ph.D. degree, under the supervision of Prof. Hong-Yeop Song. Her research interests include coding theory, PN sequences, related discrete mathematics, and their applications to digital communication systems.

Hong-Yeop Song (Senior Member, IEEE) received the B.S. degree in electronic engineering from Yonsei University, Seoul, South Korea, in 1984, and the M.S.E.E. and Ph.D. degrees from the University of Southern California (USC), Los Angeles, CA, USA, in 1986 and 1991, respectively. He spent two years as a Research Associate with USC and then two years as a Senior Engineer with the Standard Team, Qualcomm Inc., San Diego, CA, USA. Since September 1995, he has been with the Department of Electrical and Electronic Engineering, Yonsei University. His research interests include digital communications and channel coding, design and analysis of various pseudo-random sequences for communications, and cryptography. He is a member of the National Academy of Engineering of Korea (NAEK), the Mathematical Association of America (MAA), the Korean Mathematical Society (KMS), KICS, IEIE, and KIISC. He was awarded the 2017 Special Contribution Award from Korean Mathematical Society and the 2021 S. J. Choi Award from the Korean Government, both for his contribution to the global wide-spread of the fact that S. J. Choi (1646-1715) from South Korea had discovered a pair of orthogonal Latin squares of order nine much earlier than Euler. He served for the IEEE IT Society Seoul Chapter as the Chair from 2009 to 2016. He served as the General Co-Chair for IEEE ITW 2015, Jeju, South Korea.