mathematics

MDPI

Article

Optimal Binary Locally Repairable Codes with Locality and Availability from Latin Squares

Nanyuan Cao [1], Yu Zhang [1,*], Xiangqiong Zeng [2] and Li Zhang [1]

[1] School of Computing and Artificial Intelligence, Shanghai Lixin University of Accounting and Finance, Shanghai 201209, China; caony@lixin.edu.cn (N.C.); 20180087@lixin.edu.cn (L.Z.)
[2] School of Materials and Chemistry, University of Shanghai for Science and Technology, Shanghai 200093, China; zengxq@usst.edu.cn
* Correspondence: 20200042@lixin.edu.cn

Abstract

The rapid development of machine learning, large language models, and related technologies has greatly increased the demand for data storage capacity. Therefore, the role of distributed storage systems in such applications has become more prominent. However, it is inevitable that a single node fails in a distributed storage system during long-term use. Being able to repair failed nodes in a timely manner is extremely important for the stable operation of distributed storage systems, and a specific encoding scheme is required to meet the needs of efficiently repairing failed nodes. This research presents a novel family of binary locally repairable codes (LRCs) developed using multiple disjoint recovery sets constructed based on mutually orthogonal Latin squares (MOLS). The proposed constructions achieve distance optimality under the Singleton-like bound for LRCs with availability. Specifically, the codes are parameterized as $(n = r^2 + tr, k = r^2, r, t)$ and $(n = rm + tm, k = rm, r, t)$, where n is the block length, k is the dimension, r is the locality, and t is the availability. These codes achieve minimum distance $d = t + 1$, guaranteeing efficient recovery with t disjoint repair sets, each of size r. Compared to existing constructions, the proposed codes offer significant improvements in terms of code rate $\mathcal{R} = \frac{r}{r+t}$, support for larger block lengths, and reduced finite field size requirements (field size $q = 2$). Additionally, a method is introduced to improve the minimum distance of codes with even availability t, constructing codes with parameters $(n + 1, k, r, t)$ and $d = t + 2$, while preserving optimality. These properties make the proposed codes particularly suitable for distributed storage systems, where efficient and parallel repair of failed nodes is critical.

Keywords: locally repairable codes; erasure codes; codes for distributed storage systems; multiple disjoint repair sets; mutually orthogonal Latin squares

MSC: 94B05

Check for updates

Academic Editors: Yansheng Wu, Shangdong Yang and Marjan Mernik

Received: 26 February 2026
Revised: 25 March 2026
Accepted: 14 April 2026
Published: 15 April 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

With the development of artificial intelligence, machine learning, and big data, traditional single-node storage systems can no longer meet the needs of data storage. In recent years, distributed storage systems have been widely adopted to meet the growing demands of managing and storing large volumes of data. In the long-term operation of the system, node failure or disk malfunction is inevitable. In order to maintain data integrity, some distributed storage systems use a three-replication mechanism to maintain data integrity and fault tolerance [1]. When a node fails, all data can be copied from any

of the two surviving copies to repair it. However, the storage overhead of this method is large. Erasure codes, including $[n, k]$ Reed–Solomon codes, are extensively employed in distributed storage systems to address this issue. For example, Facebook [2] utilizes a $(14, 10)$ Reed–Solomon (RS) code within its storage infrastructure. The system is capable of withstanding the simultaneous failure of four nodes. In the case of a node failure, data from all 10 surviving nodes must be transmitted to restore the lost data of the failed node. The repair bandwidth is equivalent to ten times the size of the data stored in the failed node. As a result, designing an erasure coding scheme to minimize the repair bandwidth becomes essential. Locally repairable codes (LRCs) [3] were introduced as a solution to this problem.

In recent years, distributed storage systems have increasingly relied on locally repairable codes (LRCs), which have drawn considerable attention in the field [3–5]. Locally repairable codes (LRCs) have found significant applications in distributed storage systems such as Amazon S3, Google [1], and HDFS [2]. In these systems, data is distributed across multiple nodes, and failures are common due to hardware malfunctions or network issues. The proposed LRCs, with their efficient recovery properties, can minimize repair bandwidth and improve overall system reliability, making them particularly suitable for such applications. These codes are designed to recover a failed node by utilizing data from no more than $r \leq k$ other nodes. This method significantly decreases the bandwidth required for the recovery process, enhancing overall efficiency [6]. The nodes used to repair failed nodes are referred to as the repair set or recovery set, and the size r of this set is defined as the code's locality. Cai et al. [7] introduced a construction technique for locally repairable codes featuring (r, δ)-locality, while also minimizing the required field size. Tan et al. [8] proposed optimal binary locally repairable codes. Cai et al. [9] developed a framework for designing optimal locally repairable codes using packing structures. In addition, Cai et al. [10] proposed a series of constructions that attain the bound described in (2). Yavari et al. [11] proposed a construction of LRCs with joint sequential–parallel repair property. They give an extension of this construction using back-circulant Latin squares [12]. Hao et al. [13] provided various bounds and constructions for locally repairable codes, focusing on the parity-check matrix representation of LRCs.

In a distributed storage system, sometimes external users are accessing data in the failed node at the same time as the node fails. This problem is called degraded reading. Moreover, in distributed storage systems, certain data is frequently accessed simultaneously by multiple users. This type of data is called hot data. The parallel reading of hot data and degraded reading problems of hot data can be solved by adding disjoint recovering sets to a single node [14]. Wang et al. [15] proposed LRCs with disjoint recovery sets and established an upper bound on their minimum distance:

$$d \leq n - k - \left\lceil \frac{t(k-1)+1}{t(r-1)+1} \right\rceil + 2. \tag{1}$$

The availability of a locally repairable code (LRC) is defined as the number of disjoint recovery sets, represented by t. Rawat et al. [16] established an upper bound on the minimum distance of LRCs characterized by locality r and availability t, under the assumption that each recovery set contains exactly one parity symbol:

$$d \leq n - k - \left\lceil \frac{kt}{r} \right\rceil + t + 1. \tag{2}$$

Hao et al. [17] developed a construction framework for LRCs characterized by locality r and availability t, in which each information symbol is associated with t subcodes possessing a minimum distance of δ. Su et al. [18] constructed locally repairable codes

using resolvable configurations. In addition, Su [19] proposed distance-optimal locally repairable codes that achieve the bound (2), also derived from resolvable configurations.

Prakash et al. [20] introduced an upper limit for the rate of locally repairable codes under the sequential repair model:

$$\frac{k}{n} \leq \frac{r}{r+2}. \tag{3}$$

Further constraints on the code rate and minimum distance of locally repairable codes are discussed in [21–24].

In practical implementations of distributed storage systems, binary locally repairable codes offer significant advantages over those defined on larger finite fields, primarily due to their reduced computational complexity. Zhang et al. [25] developed a construction for binary LRCs, achieving a minimum distance of six.

Hao et al. [26] proposed a technique for designing binary locally repairable codes that feature multiple repair sets, utilizing finite geometries and circulant permutation matrices. Zhang et al. [27] proposed a different approach, developing locally repairable codes through the use of linear functions. In addition, Zhang et al. [28] developed a class of distance-optimal binary LRCs using combinatorial designs. Zhang et al. [29] presented a family of binary locally repairable codes featuring nonuniform locality and availability, constructed using pairwise balanced designs. Wang et al. [30] provided distance-optimal and rate-optimal constructions of binary LRCs for cases where $r = 2$ or $t = 2$.

This paper introduces constructions of binary locally repairable codes that attain an optimal minimum distance. A summary of the parameters for all proposed constructions is provided in Table 1. These codes are tailored to enable practical binary implementations, ensuring optimal performance with respect to rate, minimum distance, and repair efficiency. The main contributions of this work are summarized as follows:

1. This article gives a method for constructing binary locally repairable codes with information-symbol locality, which is based on mutually orthogonal Latin squares. This construction includes the codes from [28] and produces a distance-optimal binary LRC that achieves the bound in (2). Furthermore, the construction is rate-optimal when $t = 2$;
2. We introduce a construction for binary locally repairable codes with information-symbol locality, derived from mutually orthogonal Latin squares. This approach achieves large block lengths while maintaining small locality. The proposed construction is distance-optimal, meeting the bound specified in (2). Furthermore, when $t = 2$, it also achieves rate optimality.
3. We introduce a family of distance-optimal binary locally repairable codes, constructed using mutually orthogonal Latin rectangles, which support arbitrary values of locality. This construction attains (3) when $t = 2$ and encompasses specific codes from [8,9,28]. Moreover, the rate achieved by this construction surpasses that of the codes proposed in [9,10,19].
4. We propose a family of locally repairable codes with varying localities, which are constructed using mutually orthogonal Latin squares. These codes attain distance-optimal performance and support nonuniform localities. The rate of this construction exceeds that of the codes proposed in [9,10,19] and operates with significantly smaller field sizes. Furthermore, this construction encompasses certain codes from [8,9,28] and produces codes with parameters not addressed in [8–10,19,28]. Moreover, this construction ensures information-symbol (r, t)-availability while meeting the bound specified in (2).
5. Finally, we introduce a method to improve the minimum distance of locally repairable codes that feature multiple recovery sets. For any binary distance-optimal locally

repairable code $\mathcal{C}$ with information-symbol (r,t)-availability, where t is an even integer, we construct a new code $\mathcal{C}'$. The resulting $(n+1,k,r,t)$ locally repairable code $\mathcal{C}'$ preserves the information-symbol (r,t)-availability and attains an optimal minimum distance of $t+2$.

Table 1. Summary of parameters for all constructions introduced in this paper.

Ref.	Field Size	Structure Used	Code Parameters	Min. Distance $d(\mathcal{C})$	Rate $\mathcal{R}(\mathcal{C})$	Key Features
Theorem 1	2	$N(r)$ MOLS of order r	$(n=r^2+tr, k=r^2, r, t\le N(r)+2)_2$	$d(\mathcal{C})=t+1$, distance-optimal	$\mathcal{R}(\mathcal{C})=\frac{r}{r+t}$, rate-optimal for $t=2$	Supports arbitrary locality r
Remark 1	2	A Latin square of order r	$(n=r^2+tr, k=r^2, r, t\le 3)_2$	$d(\mathcal{C})=t+1$, distance-optimal	$\mathcal{R}(\mathcal{C})=\frac{r}{r+t}$, rate-optimal for $t=2$	Supports arbitrary locality r
Theorem 2	2	$N(t)$ MOLS of order t	$(n=t^2+rt, k=rt, r\le N(t)+2, t)_2$	$d(\mathcal{C})=t+1$, distance-optimal	$\mathcal{R}(\mathcal{C})=\frac{r}{r+t}$, rate-optimal for $t=2$	Supports arbitrary availability t
Theorem 3	2	$N(r)$ MOLS of order r	$(n=mr^2+mtr, k=mr^2, r, t\le N(r)+2)_2$	$d(\mathcal{C})=t+1$, distance-optimal	$\mathcal{R}(\mathcal{C})=\frac{r}{r+t}$, rate-optimal for $t=2$	Large block size supported
Theorem 4	2	$N(r)$ MOLR of dimensions $r\times m$	$(n=rm+tm, k=rm, r, t\le N(r)+1)_2$	$d(\mathcal{C})=t+1$, distance-optimal	$\mathcal{R}(\mathcal{C})=\frac{r}{r+t}$, rate-optimal for $t=2$	Arbitrary locality r, small locality supported
Theorem 5	2	$s\le N(r)$ MOLS of order r	$(n=r^2-u+tr, k=r^2-u, r, t<\frac{r}{u})_2$, $t\le s+2, u<r$	$d(\mathcal{C})=t+1$, distance-optimal	$\mathcal{R}(\mathcal{C})=\frac{r^2-u}{r^2-u+tr}$	Nonuniform locality supported

2. Preliminaries

In this section, we introduce several concepts related to locally repairable codes and Latin squares. The set $\{1,2,\ldots,n\}$ is denoted by $[n]$. For $i<j$, the interval $[i,j]$ denotes the set $\{i,i+1,\ldots,j\}$, whereas $(i,j]$ represents $\{i+1,\ldots,j\}$, and $[i,j)$ corresponds to $\{i,\ldots,j-1\}$. The notation $[n,k,d]_q$ is used to describe a linear code $\mathcal{C}$ over the finite field $\mathbb{F}_q$, where n, k, and d represent the block length, dimension, and minimum distance, respectively. For any codeword $\mathbf{c}$ belonging to $\mathcal{C}$, the Hamming weight, denoted by $wt(\mathbf{c})$, is defined as the count of nonzero entries in $\mathbf{c}$. The minimum distance of the code $\mathcal{C}$, written as $d(\mathcal{C})$, is the smallest Hamming weight among all nonzero codewords in $\mathcal{C}$. Additionally, the rate of the code $\mathcal{C}$, represented by $\mathcal{R}(\mathcal{C})$, is also defined. The support of an n-dimensional vector $\mathbf{c}$ is defined as $supp(\mathbf{c})=\{i\in[n]: c_i\neq 0\}$. In order to simplify the explanation of proof process in this article, when the supports of two vectors intersect at some elements, we say that the two vectors intersect at some elements.

2.1. Locally Repairable Codes

In this section, we formally define (n,k,r,t) locally repairable codes over the binary field. This paper mainly considers systematic linear locally repairable codes. Systematic linear codes can be considered as a linear mapping from $\mathbb{F}_q^k$ to $\mathbb{F}_q^n$. The encoding process maps k message symbols $(m_1,\ldots,m_k)$ to a codeword $(c_1,\ldots,c_n)$ of length n. In the resulting codeword, the first k symbols correspond exactly to the original message symbols, also referred to as information symbols. Thus, the codeword c can be expressed as $(m_1,\ldots,m_k,p_1,\ldots,p_{n-k})$. The initial k symbols of the codeword are known as message symbols or information symbols, while the remaining $n-k$ symbols are called parity symbols.

The encoding process of a linear code $\mathcal{C}$ can be expressed as a matrix multiplication $\mathbf{k}G = \mathbf{c}$, where $\mathbf{k}$ is a k-dimensional row vector over $\mathbb{F}_q$ representing the information symbols, and $\mathbf{c}$ is an n-dimensional row vector over $\mathbb{F}_q$ representing the corresponding codeword. The matrix G of size $k \times n$ is referred to as the generator matrix of the code $\mathcal{C}$. For systematic codes, the generator matrix G can be written in the form $G = [I_k|P]$, where I_k is the $k \times k$ identity matrix. The parity-check matrix of $\mathcal{C}$, denoted by H, is defined as the generator matrix of the dual code of $\mathcal{C}$. The matrix $H = [-P^T|I_{n-k}]$ is an $(n-k) \times n$ matrix, and the conditions $GH^T = \mathbf{0}$ and $HG^T = \mathbf{0}$ hold.

Definition 1. *An element c_i in a codeword $\mathbf{c} \in \mathcal{C}$ is said to have locality r if there exists a subset $R(i) \subset [n] \setminus \{i\}$ such that c_i can be expressed as a linear combination of the code symbols in $\mathbf{c}$ indexed by the elements of $R(i)$, where $|R(i)| \leq r$.*

An element c_i in a codeword $\mathbf{c} \in \mathcal{C}$ is said to have locality r and availability t if there exist t disjoint subsets $R_1(i), R_2(i), \ldots, R_t(i) \subset [n] \setminus \{i\}$. Each subset $R_j(i)$, for $j \in [t]$, allows c_i to be expressed as a linear combination of the code symbols in $\mathbf{c}$ indexed by $R_j(i)$, where $|R_j(i)| \leq r$ and $R_j(i) \cap R_l(i) = \emptyset$ whenever $j \neq l$.

Definition 2. *A systematic $[n, k, d]$ linear code $\mathcal{C}$ is called a locally repairable code with all-symbol locality r and availability t (AS-LRC) if every code symbol in $\mathcal{C}$ has locality r and availability t.*

A systematic $[n, k, d]$ linear code $\mathcal{C}$ is called a locally repairable code with information-symbol locality r and availability t (IS-LRC) if every message symbol in $\mathcal{C}$ has locality r and availability t. An $(n, k, r, t)_q$ LRC is defined as a linear $[n, k, d]$ code $\mathcal{C}$ over the finite field $\mathbb{F}_q$ that has information-symbol locality r and availability t. In this paper, all constructions are over the binary field $\mathbb{F}_2$. Therefore, the parameter q is omitted.

2.2. Latin Squares and Latin Rectangles

Definition 3. *A Latin square of order n [31] is an $n \times n$ matrix in which each row and each column contain every element from the set $[n] = \{1, 2, \ldots, n\}$ exactly once. This means that every row and column is a permutation of the elements in $[n]$. The entry in the cell located at row i and column j of the Latin square L is denoted by $L(i, j)$.*

Two Latin squares L_1 and L_2 of size n are defined as orthogonal if, for every pair $(i, j) \in [n] \times [n]$, there exists exactly one corresponding pair (k, l), such that $L_1(k, l) = i$ and $L_2(k, l) = j$. A collection of t Latin squares $L_1, L_2, \ldots, L_t$ of order n is described as a set of mutually orthogonal Latin squares (commonly abbreviated as MOLS) if any two distinct squares L_i and L_j, with $i \neq j$, maintain orthogonality. The maximum number of MOLS for a given order n is denoted as $N(n)$. For any q that is a prime power, the maximum is given by $N(q) = q - 1$ [32].

When n can be expressed in the form $n = p_1^{e_1} p_2^{e_2} \ldots p_r^{e_r}$, where $p_1, p_2, \ldots, p_r$ are distinct primes and $e_1, e_2, \ldots, e_r$ are positive integers, the largest possible number of MOLS is determined by $N(n) = \min\{p_1^{e_1}, p_2^{e_2}, \ldots, p_r^{e_r}\} - 1$ [32]. MOLS are essential in the design and construction of locally repairable codes. To define additional parameter sets needed for our construction, we propose a generalization of the Latin square.

Definition 4. *A Latin rectangle L [31] is an $m \times n$ array, where $m \leq n$, such that each row is a permutation of the elements in the set $[n] = \{1, 2, \ldots, n\}$, and no element is repeated within any column.*

Two Latin rectangles L_1 and L_2 are considered orthogonal if, for each $(i, j) \in [n] \times [n]$, there is at most one distinct pair (k, l) satisfying $L_1(k, l) = i$ and $L_2(k, l) = j$. Analogous to the notion of MOLS, a collection of t Latin rectangles $L_1, \ldots, L_t$ of dimensions $m \times n$ is

said to form a set of mutually orthogonal Latin rectangles (MOLR) if, for any pair of Latin rectangles L_i and L_j, where $i \neq j$, the orthogonality condition is satisfied [33].

Assume there exist s mutually orthogonal Latin squares (MOLS) of order n. By selecting the same m rows, where $m \leq n$, from each of these MOLS, one can construct s mutually orthogonal Latin rectangles with dimensions $m \times n$. Therefore, given any set of n MOLS, one can construct a minimum of n mutually orthogonal Latin rectangles with dimensions $m \times n$, where $m \leq n$.

3. Optimal Binary Locally Repairable Codes from MOLS

In this section, we introduce a framework of locally repairable codes that leverages the partitioning of sets using mutually orthogonal Latin squares.

Let $\Omega = \{1, \ldots, n^2\}$ represent a set containing n^2 integers. Arrange these n^2 elements in an $n \times n$ array following their natural order. s MOLS of order n can generate $s+2$ distinct partitions of the n^2 cells in the array into n subsets, each containing n elements. These subsets are called blocks. A partition of these blocks is referred to as a parallel class, where any two blocks within the same parallel class are disjoint, and the union of all blocks is Ω. Another important property of these partitions is that any two blocks from distinct parallel classes share exactly one common element.

The partitions consist of the row partition P^{row}, the column partition P^{col}, and s partitions derived from the s MOLS, denoted as $P^1, \ldots, P^s$. In the partitions P^i, each block corresponds to the cells in the array labeled with the numbers $1, 2, \ldots, n$.

Here we give a simple proof of the above properties. For partition of rows and partition of columns, each row and column intersect exactly once, and different blocks from partition of rows (partition of colums) are disjoint. Now consider two blocks, $B_k \in P^i$ and $B_l \in P^j$, where $i \in [s]$, $j \in [s]$, and $i \neq j$. Let B_i represent the cells in the array labeled with i. Based on the properties of MOLS, for every $(i, j) \in [n] \times [n]$, there is a unique pair (k, l) satisfying $L_1(k, l) = i$ and $L_2(k, l) = j$. Consequently, the blocks B_k and B_l have exactly one common element. Now consider two blocks $B_k \in P^{row}$, $B_l \in P^i$, $i \in [s]$. From the property of Latin squares, in each row, there is exactly one cell numbered l, so each block from P^{row} and each block from P^i intersects on exactly one element. The proof is complete.

Example 1.

$$L_1 = \begin{pmatrix} 1 & 3 & 2 \\ 3 & 2 & 1 \\ 2 & 1 & 3 \end{pmatrix}, L_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{pmatrix} \tag{4}$$

The above are two MOLS of order 3. *We get* 4 *partitions of* 9 *elements into* 3 *blocks of size* 3.

$$\begin{aligned} P^{row} &= \{[1,2,3], [4,5,6], [7,8,9]\}, \\ P^{col} &= \{[1,4,7], [2,5,8], [3,6,9]\}, \\ P^1 &= \{[1,6,8], [3,5,7], [2,4,9]\}, \\ P^2 &= \{[1,5,9], [2,6,7], [3,4,8]\} \end{aligned} \tag{5}$$

The design of our locally repairable codes is accomplished by constructing a parity-check matrix for the code. To create an (n, k, r, t) LRC over the binary field, we start by choosing any $t \leq s+2$ partitions (parallel classes) from the $s+2$ partitions of r^2 elements, which are associated with the s MOLS of order r. We construct an $r \times r^2$ submatrix M_l for partition P^l. Each row of M_l is indexed by a block from the parallel class associated with P^l, and each column is indexed by the r^2 elements. Let $M_l(i, j)$ denote the entry located in the i-th row and j-th column of M_l. The value of $M_l(i, j)$ is equal to 1 if and only if the

element j is contained in the block B_i. Subsequently, t parallel classes are chosen to form a $tr \times r^2$ matrix R.

$$R = \begin{pmatrix} M_1 \\ \dots \\ M_t \end{pmatrix}. \tag{6}$$

Note that the support of any two distinct rows in R intersects in at most one element. This comes from blocks in different partitions intersecting at most once and blocks from the same partition not intersecting. The supports of any two columns intersect at most once too. Since R is composed of t parallel classes, the weight of any column in R is t. Moreover, because each block in the partitions mentioned earlier contains r elements, the weight of every row in R is r. The parity-check matrix H for this construction is defined as

$$H = \left(\begin{array}{c|c} R & I_{(n-k)\times(n-k)} \end{array} \right). \tag{7}$$

Generator matrix for our construction is defined as

$$G = \left(\begin{array}{c|c} I_{k\times k} & -R^T \end{array} \right) = \left(\begin{array}{c|c} I_{k\times k} & R^T \end{array} \right). \tag{8}$$

The second equation comes from our construction over $\mathbb{F}_2$. From the above construction, we can deduce that $n = r^2 + tr, k = r^2$. When r is a prime power, there are $r - 1$ MOLS, which implies that $t \leq r + 1$. For other scenarios, we have $t \leq N(r) + 2$, where $N(r)$ denotes the maximum number of mutually orthogonal Latin squares (MOLS) of order r.

The following theorem establishes that the linear code, defined by H as its parity-check matrix and G as its generator matrix, is a distance-optimal LRC.

Theorem 1. *The locally repairable code $\mathcal{C}$, defined by the parity-check matrix H and the generator matrix G, is an $(n = r^2 + tr, k = r^2, r, t \leq N(r) + 2)_2$ IS-LRC. This code achieves the optimal minimum distance that satisfies the bound in (2), where $N(r)$ represents the maximum number of MOLS of order r. Additionally, the rate of the code is $\mathcal{R}(\mathcal{C}) = \frac{r}{r+t}$, and when $t = 2$, the construction attains rate optimality.*

Proof. To begin, we show that in our construction, each message symbol has t disjoint repair sets, where each set contains r elements. Note that the first k columns of H correspond to the information symbols of the code. From the construction of H, every row has the same Hamming weight, which is $r + 1$.

Since R consists of t parallel classes, there is a nonzero entry in each parallel class for any coordinate within the first k columns of H. Consequently, the Hamming weight of each column in R corresponding to the first k columns of H is equal to t. Each of these t nonzero entries corresponds to a repair set of size r.

Next, we demonstrate that these t repair sets are mutually disjoint. The rows of H associated with the t nonzero entries in a given column are derived from blocks in t different parallel classes. As any pair of blocks from distinct parallel classes have exactly one common element and because the final $n - k$ columns of H form an identity matrix, the nonzero elements in the rows corresponding to the t nonzero entries in any column of H are disjoint. Therefore, the t repair sets for each information symbol are mutually disjoint.

We now prove that the minimum distance of $\mathcal{C}$ is $t + 1$. Assume that the first nonzero coordinate of a codeword $\mathbf{c} \in \mathcal{C}$ is j.

1. Case 1: $j \in [k]$: Each information symbol in $\mathcal{C}$ has t disjoint repair sets. Each repair set contributes one additional nonzero element to $\mathbf{c}$. Thus, the Hamming weight of $\mathbf{c}$ is at least $t + 1$.

2. Case 2: $j \in [k+1, n]$: In this case, there is a nonzero entry in the j-th column of H, which implies that there is at least one nonzero entry in the corresponding row within the first k columns of H. Let this nonzero entry be located in the j'-th column of H, where $j' \in [k]$. The j'-th information symbol has t disjoint repair sets, so at least one nonzero symbol of $\mathbf{c}$ lies in these t repair sets. As a result, $\mathbf{c}$ has at least $2 + (t-1) = t+1$ nonzero entries.

Therefore, the minimum distance of $\mathcal{C}$ is $d \geq t+1$. Finally, consider the bound (2) for locally repairable codes where each repair set contains exactly one parity symbol:

$$d \leq n - k - \left\lceil \frac{kt}{r} \right\rceil + t + 1. \tag{9}$$

In our construction,

$$\begin{aligned} d(\mathcal{C}) &\leq n - k - \left\lceil \frac{kt}{r} \right\rceil + t + 1 \\ &= (r^2 + tr) - r^2 - \left\lceil \frac{tr^2}{r} \right\rceil + t + 1 \\ &= t + 1. \end{aligned} \tag{10}$$

Thus, the minimum distance of the code is $d(\mathcal{C}) = t+1$, achieving distance optimality. Finally, the rate of the code is calculated as

$$\mathcal{R}(\mathcal{C}) = \frac{k}{n} = \frac{r^2}{r^2 + tr} = \frac{r}{r+t}. \tag{11}$$

In the special case where $t = 2$, the rate of $\mathcal{C}$ meets the bound (3), making the construction rate-optimal. □

Remark 1. *The construction outlined above does not require the use of a set of MOLS. In fact, as long as there is a Latin square of order r, it is feasible to construct an $(n, k, r, t = 3)$ LRC. By utilizing a Latin square of order r, it is possible to generate three partitions of the r^2 elements, each corresponding to a distinct parallel class.*

The first construction presented in [16] introduced a family of $(N + \frac{kt}{r}, k, r, 1 \leq t \leq \frac{k-1}{r-1})_q$ LRCs. This construction was based on a $(k, r, 1)$ resolvable balanced incomplete block design (RBIBD) [32] combined with an $(N + t, k)_q$ MDS code. However, it is worth noting that a $(36, 6, 1)$ RBIBD does not exist [32], which implies that the construction in [16] cannot be used to create $(n, 36, 6, t)_q$ LRCs. In contrast, the approach described in this section utilizes a Latin square of order 6 to construct a $(36 + 6t, 36, 6, t \leq 3)_2$ LRC.

Hao et al. [26] introduced a method for constructing $(n = q^2 + tq, k = q^2, r = q, t \leq q)_2$ LRCs, where q is an odd prime. The construction described in this section generalizes and builds upon their work. When q is an odd prime, the parameters of our construction encompass those of the codes in [26], allowing t to reach up to $q+1$. Moreover, our construction extends their results to include cases where q is either a prime power or a general integer. When q is a prime power, our construction yields a family of $(n = q^2 + tq, k = q^2, r = q, t \leq q+1)_2$ LRCs. For arbitrary integers q, the construction produces $(n = q^2 + tq, k = q^2, r = q, t \leq N(q) + 2)_2$ LRCs, where $N(q)$ represents the maximum number of MOLS of order q.

Hao et al. also proposed a construction for $(n = 2m, k = m, r, t = r)$ LRCs with information (r, t)-availability, where $r = 2^s$ and $m = 4^s - 1$. Based on Theorem 1 in this section, when q is a prime power, our construction produces $(n = 2q^2, k = q^2, r = q, t = q)$ LRCs with message symbol (r, t)-availability. Therefore, the codes generated through our construction include most of the codes introduced in [26].

Remark 2. *In our construction, $R(i)_t$ denotes the t-th repair set of the i-th symbol in the code. Each repair set $R(i)_t$ is capable of repairing the corresponding code symbol. Specifically,*

1. If $i \in [1,k]$, the repair set is used to repair the i-th information symbol.

2. $R(i)_t$ is represented as a set of integers, where each element in the set corresponds to the index of a parity symbol in the code that participates in the repair process.

3. The largest index in the set $R(i)_t$ directly identifies the position of the parity symbol in the code that is used for repairing the i-th symbol. This definition ensures that the relationship between repair sets, information symbols, and parity symbols is clearly established in the context of binary LRCs.

Example 2. *For the 2 MOLS of order 3 in Example 1, the correspondence submatrices of the 4 parallel classes are*

$$\begin{aligned} M_{row} &= \begin{pmatrix} 1&1&1&0&0&0&0&0&0\\ 0&0&0&1&1&1&0&0&0\\ 0&0&0&0&0&0&1&1&1 \end{pmatrix}, M_{col} = \begin{pmatrix} 1&0&0&1&0&0&1&0&0\\ 0&1&0&0&1&0&0&1&0\\ 0&0&1&0&0&1&0&0&1 \end{pmatrix} \\ M_1 &= \begin{pmatrix} 1&0&0&0&0&1&0&1&0\\ 0&0&1&0&1&0&1&0&0\\ 0&1&0&1&0&0&0&0&1 \end{pmatrix}, M_2 = \begin{pmatrix} 1&0&0&0&1&0&0&0&1\\ 0&1&0&0&0&1&1&0&0\\ 0&0&1&1&0&0&0&1&0 \end{pmatrix} \end{aligned} \tag{12}$$

Without loss of generality, we select two submatrices, M_1 and M_2, to construct the matrix R.

$$R = \begin{pmatrix} 1&0&0&0&0&1&0&1&0\\ 0&0&1&0&1&0&1&0&0\\ 0&1&0&1&0&0&0&0&1\\ 1&0&0&0&1&0&0&0&1\\ 0&1&0&0&0&1&1&0&0\\ 0&0&1&1&0&0&0&1&0 \end{pmatrix}, \tag{13}$$

$$H = \begin{pmatrix} 1&0&0&0&0&1&0&1&0&1&0&0&0&0&0\\ 0&0&1&0&1&0&1&0&0&0&1&0&0&0&0\\ 0&1&0&1&0&0&0&0&1&0&0&1&0&0&0\\ 1&0&0&0&1&0&0&0&1&0&0&0&1&0&0\\ 0&1&0&0&0&1&1&0&0&0&0&0&0&1&0\\ 0&0&1&1&0&0&0&1&0&0&0&0&0&0&1 \end{pmatrix}, \tag{14}$$

$$G = \begin{pmatrix} 1&0&0&0&0&0&0&0&0&1&0&0&1&0&0\\ 0&1&0&0&0&0&0&0&0&0&0&1&0&1&0\\ 0&0&1&0&0&0&0&0&0&0&1&0&0&0&1\\ 0&0&0&1&0&0&0&0&0&0&0&1&0&0&1\\ 0&0&0&0&1&0&0&0&0&0&1&0&1&0&0\\ 0&0&0&0&0&1&0&0&0&1&0&0&0&1&0\\ 0&0&0&0&0&0&1&0&0&0&1&0&0&1&0\\ 0&0&0&0&0&0&0&1&0&1&0&0&0&0&1\\ 0&0&0&0&0&0&0&0&1&0&0&1&1&0&0 \end{pmatrix}. \tag{15}$$

This code is a $(15,9,3,2)$ LRC with a minimum distance of $d(\mathcal{C}) = 3$ and a rate of $\mathcal{R} = \frac{3}{5}$. The code achieves both distance optimality and rate optimality. Each message symbol can be recovered using three other coordinates, with two disjoint recovery sets of size three. Without loss of generality, consider the 4-th information symbol. From H or G, we can conclude that $R_1(4) = \{2,9,12\}, R_2(4) = \{3,8,15\}$, which means the 4-th message symbol can be recovered by adding the coordinates indexed by $R_1(4)$ or $R_2(4)$, and these two repair sets are disjoint.

By utilizing $N(r)$ MOLS of order r, the construction outlined in Theorem 1 produces an $(n = r^2 + tr, k = r^2, r, t \leq N(r) + 2)$ LRC that provides information (r,t)-availability.

We now introduce the construction of locally repairable codes based on MOLS for any given locality r.

First, choose $r \leq N(t) + 2$ parallel classes (partitions) from $N(t) + 2$ partitions from $N(t)$ MOLS of order t. Each partition is associated with a $t \times t^2$ submatrix M_l. Then construct an $rt \times t^2$ matrix R^T. Its transpose is a $t^2 \times rt$ matrix R.

$$R^T = \begin{pmatrix} M_1 \\ \dots \\ M_r \end{pmatrix}. \tag{16}$$

$$R = \begin{bmatrix} M_1 & \dots & M_r \end{bmatrix}. \tag{17}$$

From the properties of MOLS, we conclude that each row in R has a weight of r, every column in R has a weight of t, and the supports of any two distinct rows in R intersect in at most one element. Using this matrix R, we construct both a parity-check matrix H and a generator matrix G. The parity-check matrix H in our construction is defined as

$$H = \left(\begin{array}{c|c} R & I_{(n-k)\times(n-k)} \end{array} \right), \tag{18}$$

$$G = \left(\begin{array}{c|c} I_{k\times k} & -R^T \end{array} \right) = \left(\begin{array}{c|c} I_{k\times k} & R^T \end{array} \right), \tag{19}$$

where $n - k = t^2, k = rt$.

Theorem 2. *The locally repairable code $\mathcal{C}$, designed using H as its parity-check matrix and G as its generator matrix, is an $(n = t^2 + rt, k = rt, r \leq N(t) + 2, t)_2$ IS-LRC. This code achieves an optimal minimum distance that satisfies the bound in (2), where $N(t)$ denotes the maximum number of MOLS of order t. Additionally, the rate of the code is $\mathcal{R}(\mathcal{C}) = \frac{r}{r+t}$, and when $t = 2$, this construction achieves rate-optimality.*

Proof. We start by showing that each information symbol in this code has t disjoint repair sets, where each repair set contains r elements. Observe that the first k columns of H correspond to the information symbols in the code. From the given construction, every row in H has a Hamming weight of $r + 1$. Additionally, for any coordinate in the first k columns of H, there exist exactly t nonzero entries, corresponding to the t repair sets.

To confirm that these t repair sets are disjoint, note that the supports of any two distinct rows in H overlap in at most one element. Furthermore, the final $n - k$ columns of H form an identity matrix, ensuring that the repair sets for each information symbol are disjoint. Thus, the information symbols in $\mathcal{C}$ exhibit (r, t)-availability.

Next, we prove that the minimum distance of $\mathcal{C}$ is $t + 1$. Let $\mathbf{c} \in \mathcal{C}$ be a randomly selected codeword, and let j denote the index of its first nonzero coordinate.

1. Case 1: $j \in [k]$: Since each information symbol in $\mathcal{C}$ has (r, t)-availability, there must exist one additional nonzero symbol in $\mathbf{c}$ for each of the t repair sets. As a result, the Hamming weight of $\mathbf{c}$ is at least $t + 1$.

2. Case 2: $j \in [k + 1, n]$: In this case, the j-th column of H contains a nonzero entry. Consequently, there must also be at least one nonzero entry in the same row within the first k columns of H. Let this nonzero entry be located in the i'-th row and the j'-th column of H, where $j' \in [k]$. Since every information symbol in $\mathcal{C}$ has (r, t)-availability, there must be one nonzero symbol in $\mathbf{c}$ for each repair set. Therefore, the codeword $\mathbf{c}$ includes at least two nonzero elements at the j-th and j'-th coordinates, in addition to $t - 1$ other nonzero elements elsewhere. This demonstrates that $d(\mathcal{C}) \geq t + 1$.

Finally, consider the bound (2):

$$\begin{aligned} d(\mathcal{C}) &\leq n-k-\left\lceil \frac{kt}{r} \right\rceil + t + 1 \\ &= (t^2+rt) - rt - \left\lceil \frac{rt^2}{r} \right\rceil + t + 1 \\ &= t+1. \end{aligned} \tag{20}$$

Thus, the minimum distance of our construction is $d(\mathcal{C}) = t+1$, achieving distance optimality.

Finally, the rate of the code is calculated as

$$\mathcal{R}(\mathcal{C}) = \frac{k}{n} = \frac{rt}{rt+t^2} = \frac{r}{r+t}. \tag{21}$$

For the case where $t = 2$, the rate of $\mathcal{C}$ achieves the bound in (3), rendering this construction rate-optimal when $t = 2$. □

Remark 3. *In a binary linear code, it is possible for the sum of the symbols in a repair set to be zero. However, due to the properties of binary codes and the structure of the parity-check matrix, the sum (i.e., XOR) of a code symbol and all the symbols in any of its repair sets is always zero.*

This property ensures that to repair a failed symbol, it suffices to read all the symbols in any one of its repair sets and perform an XOR operation on these elements. The result of this operation will recover the value of the failed code symbol.

Example 3. *Similar to Example 2, we choose* 2 *parallel classes from* 2 *MOLS of order* 3 *to form* R^T *and* R.

$$R^T = \begin{pmatrix} 1&0&0&0&0&1&0&1&0 \\ 0&0&1&0&1&0&1&0&0 \\ 0&1&0&1&0&0&0&0&1 \\ 1&0&0&0&1&0&0&0&1 \\ 0&1&0&0&0&1&1&0&0 \\ 0&0&1&1&0&0&0&1&0 \end{pmatrix}, \tag{22}$$

$$R = \begin{pmatrix} 1&0&0&1&0&0 \\ 0&0&1&0&1&0 \\ 0&1&0&0&0&1 \\ 0&0&1&0&0&1 \\ 0&1&0&1&0&0 \\ 1&0&0&0&1&0 \\ 0&1&0&0&1&0 \\ 1&0&0&0&0&1 \\ 0&0&1&1&0&0 \end{pmatrix}. \tag{23}$$

$$G = \left(\begin{array}{cccccc|cccccccc} 1&0&0&0&0&0&1&0&0&0&0&1&0&1&0 \\ 0&1&0&0&0&0&0&0&1&0&1&0&1&0&0 \\ 0&0&1&0&0&0&0&1&0&1&0&0&0&0&1 \\ 0&0&0&1&0&0&1&0&0&0&1&0&0&0&1 \\ 0&0&0&0&1&0&0&1&0&0&0&1&1&0&0 \\ 0&0&0&0&0&1&0&0&1&1&0&0&0&1&0 \end{array}\right), \tag{24}$$

$$H = \left(\begin{array}{cccccc|cccccccc} 1&0&0&1&0&0&1&0&0&0&0&0&0&0\\ 0&0&1&0&1&0&0&1&0&0&0&0&0&0\\ 0&1&0&0&0&1&0&0&1&0&0&0&0&0\\ 0&0&1&0&0&1&0&0&0&1&0&0&0&0\\ 0&1&0&1&0&0&0&0&0&0&1&0&0&0\\ 1&0&0&0&1&0&0&0&0&0&0&1&0&0\\ 0&1&0&0&1&0&0&0&0&0&0&0&1&0\\ 1&0&0&0&0&1&0&0&0&0&0&0&0&1\\ 0&0&1&1&0&0&0&0&0&0&0&0&0&1 \end{array}\right). \tag{25}$$

This code is a $(15, 6, 2, 3)$ *LRC with a minimum distance of* $d(\mathcal{C}) = 4$ *and a rate of* $\mathcal{R} = \frac{2}{5}$. *The code achieves distance optimality. Each message symbol can be repaired using two other coordinates, and there are three repair sets of size two.*

Without any loss of generality, let us examine the third information symbol. From the matrices H or G, we observe that $R_1(3) = \{5, 8\}$, $R_2(3) = \{6, 10\}$, *and* $R_3(3) = \{4, 15\}$. *This indicates that the third message symbol can be reconstructed by summing the coordinates indexed by* $R_i(3)$ *for* $i \in [3]$, *and these three repair sets are mutually disjoint.*

4. Optimal Binary Locally Repairable Codes with Large Block Length

Notice that in our construction from Section 3, the block length n of locally repairable codes are restricted to $r^2 + rt$ or $t^2 + rt$. When the block length gets larger, locality r grows. In this section, we present a class of constructions in which the locality remains unchanged irrespective of the code length.

The construction described in this section is an extension of the codes proposed in Section 3. Similar to the construction in the previous section, we select any $t \leq s + 2$ parallel classes from $s + 2$ partitions derived from s MOLS of order r to form a $tr \times r^2$ matrix R'.

$$R' = \begin{pmatrix} M_1 \\ \dots \\ M_t \end{pmatrix}. \tag{26}$$

The construction presented in this section is based on the Kronecker product of two matrices. Let A be an $m \times n$ matrix and B a $p \times q$ matrix. The Kronecker product $A \otimes B$ generates a block matrix, which is defined as

$$A \otimes B = \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{pmatrix}, \tag{27}$$

Here, a_{ij} denotes the entry located in the i-th row and j-th column of A. The resulting matrix $A \otimes B$ has dimensions $mp \times nq$.

Next, we define an $mtr \times mr^2$ matrix R as

$$R = R' \otimes I_m, \tag{28}$$

where I_m denotes the $m \times m$ identity matrix over $\mathbb{F}_2$. The parity-check matrix H is defined as follows:

$$G = \left(\begin{array}{c|c} I_{k\times k} & -R^T \end{array} \right) = \left(\begin{array}{c|c} I_{k\times k} & R^T \end{array} \right), \tag{29}$$

$$H = \left(\begin{array}{c|c} R & I_{(n-k)\times(n-k)} \end{array} \right), \tag{30}$$

where $k = mr^2$ and $n - k = mtr$.

The following theorem demonstrates that the linear code, constructed using H as its parity-check matrix and G as its generator matrix, is an $(n, k, r, t)_2$ attaining the bound (2).

Theorem 3. *The locally repairable code $\mathcal{C}$, designed using H as its parity-check matrix and G as its generator matrix, is an $(n = mr^2 + mtr, k = mr^2, r, t \leq N(r) + 2)_2$ IS-LRC that achieves the bound (2). Here, $N(r)$ denotes the maximum number of MOLS of order r. Additionally, the rate of the code is $\mathcal{R}(\mathcal{C}) = \frac{r}{r+t}$, and when $t = 2$, this construction is rate-optimal.*

Proof. From the construction described above, every column in R has weight t, while every row has weight r. Furthermore, the structure of the matrix R ensures that the supports of any two rows intersect at most once. Since R comes from the tensor product of R' and I_m, any two rows in R intersect more than once if and only if there exist two rows in R' intersecting more than once, which would violate the properties of MOLS. Hence, the matrix R is constructed such that 1. Each row has uniform weight r. 2. Each column has uniform weight t. 3. The supports of any two rows intersect at most once.

The first k columns of H correspond to the information symbols of the code, and all rows in H have weight $r + 1$. Now, consider the i-th coordinate of a randomly selected codeword $\mathbf{c} \in \mathcal{C}$.

1. Case 1: $i \in [k]$: In this scenario, the i-th column contains exactly t nonzero entries. Since the supports of any two rows in H intersect at most once, the supports of the rows corresponding to these t nonzero entries overlap only at coordinate i. As a result, every coordinate i in $\mathbf{c} \in \mathcal{C}$ has (r, t)-availability.

2. Case 2: $i \in [k + 1, n]$: In this case, the i-th column of H contains a nonzero entry. This implies that there exists at least one nonzero entry in the same row within the first k columns of H. Let this nonzero entry be located in the i'-th row and the j'-th column of H, where $j' \in [k]$. Since the information symbol corresponding to j' is associated with t disjoint repair sets, the symbol $c_{j'}$ has at least t additional nonzero entries. Therefore, the codeword $\mathbf{c}$ contains at least $t + 1$ nonzero entries.

Thus, in both cases, we have established that $d(\mathcal{C}) \geq t + 1$.

Next, consider the bound (2) for locally repairable codes where each repair set contains exactly one parity symbol:

$$d \leq n - k - \left\lceil \frac{kt}{r} \right\rceil + t + 1. \tag{31}$$

In our construction

$$\begin{aligned} d(\mathcal{C}) &\leq n - k - \left\lceil \frac{kt}{r} \right\rceil + t + 1 \\ &= (mr^2 + mtr) - mr^2 - \left\lceil \frac{mr^2 t}{r} \right\rceil + t + 1 \\ &= t + 1. \end{aligned} \tag{32}$$

Therefore, the minimum distance of our construction is $d(\mathcal{C}) = t + 1$, achieving distance optimality.

Finally, the rate of the code is calculated as

$$\mathcal{R}(\mathcal{C}) = \frac{k}{n} = \frac{mr^2}{mr^2 + mtr} = \frac{r}{r + t}. \tag{33}$$

When $t = 2$, the rate of $\mathcal{C}$ achieves the bound in (3), making the construction rate-optimal. □

Remark 4. *Notice that the tensor product method can also be applied to the construction in Theorem 2.*

Example 4. *Similar to Example 2, we choose* 2 *parallel classes from* 2 *MOLS of order* 3 *to form* R' *and* $R = R' \otimes I_2$.

$$R' = \begin{pmatrix} 1&0&0&0&0&1&0&1&0\\ 0&0&1&0&1&0&1&0&0\\ 0&1&0&1&0&0&0&0&1\\ 1&0&0&0&1&0&0&0&1\\ 0&1&0&0&0&1&1&0&0\\ 0&0&1&1&0&0&0&1&0 \end{pmatrix}. \tag{34}$$

$$R = \left(\begin{array}{cc|cc|cc|cc|cc|cc|cc|cc} 1&0&0&0&0&0&0&0&1&0&0&0&1&0&0&0\\ 0&1&0&0&0&0&0&0&0&1&0&0&0&1&0&0\\ \hline 0&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0\\ 0&0&0&0&0&1&0&0&0&1&0&0&0&1&0&0\\ \hline 0&0&1&0&0&0&1&0&0&0&0&0&0&0&1&0\\ 0&0&0&1&0&0&0&1&0&0&0&0&0&0&0&1\\ \hline 1&0&0&0&0&0&0&0&1&0&0&0&0&0&1&0\\ 0&1&0&0&0&0&0&0&0&1&0&0&0&0&0&1\\ \hline 0&0&1&0&0&0&0&0&0&0&1&0&1&0&0&0\\ 0&0&0&1&0&0&0&0&0&0&0&1&0&1&0&0\\ \hline 0&0&0&0&1&0&1&0&0&0&0&0&0&0&1&0\\ 0&0&0&0&0&1&0&1&0&0&0&0&0&0&0&1 \end{array}\right) \tag{35}$$

$$H = \left(\begin{array}{cccccccccccccccc|cccccccccccc} 1&0&0&0&0&0&0&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0\\ 0&1&0&0&0&0&0&0&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0\\ 0&0&1&0&0&0&1&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&1&0&0&0&0&0&0&0\\ 0&0&0&1&0&0&0&1&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&1&0&0&0&0&0&0\\ 1&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&1&0&0&0&0&0\\ 0&1&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&1&0&0&0&0\\ 0&0&1&0&0&0&0&0&0&0&1&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0\\ 0&0&0&1&0&0&0&0&0&0&0&1&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0\\ 0&0&0&0&1&0&1&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0\\ 0&0&0&0&0&1&0&1&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&1 \end{array}\right), \tag{36}$$

$$G = \left(\begin{array}{cccccccccccccccc|cccccccccccc} 1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&1&0&0&0&0&0\\ 0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&1&0&0&0&0\\ 0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&1&0&0&0\\ 0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&1&0&0\\ 0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&1&0\\ 0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&1\\ 0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&1&0\\ 0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&1\\ 0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&1&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&1&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&1&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&1&0&0&0&0&0&1&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0&1&0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&1&0&0&0&0&0&0&0&0&0&1&0\\ 0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&1&0&0&0&0&0&0&0&0&0&1\\ 0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&1&0&1&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&1&0&1&0&0&0&0&0&0 \end{array}\right). \tag{37}$$

H is a parity-check matrix for a $(30, 18, 3, 2)$ LRC $\mathcal{C}$. The code $\mathcal{C}$ has a minimum distance of $d(\mathcal{C}) = 3$ and a rate of $\mathcal{R}(\mathcal{C}) = \frac{3}{5}$. The value of $d(\mathcal{C})$ satisfies the bound in (2), and $\mathcal{R}(\mathcal{C})$ meets the bound in (3). Thus, $\mathcal{C}$ is both distance-optimal and rate-optimal. Suppose the symbol c_1 in $\mathbf{c} \in \mathcal{C}$ is erased. Erased symbol c_1 can be represented by elements in $\mathbf{c}$ from $R_i(1)$. Specifically, $R_1(1) = \{11, 15, 19\}$ and $R_2(1) = \{9, 17, 25\}$.

5. Optimal Binary Locally Repairable Codes Derived from MOLR with General Locality

In the codes constructed from MOLS in Theorem 1, the dimension $k = r^2$, and the locality r increase as the dimension k grows. In a locally repairable code, repairing a failed node involves transmitting data from r other nodes. Therefore, it is desirable to minimize the locality r as much as possible. In this section, we introduce a construction of optimal binary locally repairable codes with message symbol (r, t)-availability and arbitrary locality r, which is based on mutually orthogonal Latin rectangles.

Given s MOLR of size $r \times m$, $\Omega = \{1, \ldots, rm\}$ is a set of rm integers. Then place these rm integers in an $r \times m$ array in natural order. Then s MOLR of size $r \times m$ can give $s + 1$ different partitions of the rm cells into m sets of size r. The sets in the partitions are called blocks. Each partition is called a parallel class, wherein every pair of blocks within the same parallel class is disjoint, and the union of all blocks in a single parallel class equals the set Ω. Additionally, these partitions satisfy the property that any two blocks belonging to different parallel classes intersect at most once. The partitions include the column partition P^{column} and the partitions derived from s MOLR, denoted as $P^1, \ldots, P^s$ (where the blocks in P^i correspond to the cells in the array labeled $1, 2, \ldots, m$).

We provide a simple proof of the above properties. Consider two blocks $B_k \in P^i$ and $B_l \in P^j$, where $i \in [s]$, $j \in [s]$, and $i \neq j$. Let B_i represent the cells in the array labeled with i. By the properties of MOLR, for any $(i, j) \in [r] \times [m]$, there is at most one pair $(L_1(k, l), L_2(k, l))$ satisfying $L_1(k, l) = i$ and $L_2(k, l) = j$. Thus, the intersection of B_k and B_l contains at most one element. Now consider two blocks $B_k \in P^{col}$, $B_l \in P^i$, $i \in [s]$. From the property of Latin rectangles, in each column, there is exactly one cell numbered l, so each block from P^{col} and each block from P^i intersects at most once. The proof is complete.

Example 5.

$$L_1 = \begin{pmatrix} 1 & 3 & 2 \\ 3 & 2 & 1 \end{pmatrix}, L_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \tag{38}$$

The above are two MOLR of size 2×3. We get 3 partitions of 6 elements into 3 blocks of size 2.

$$\begin{aligned} P^{col} &= \{[1, 4], [2, 5], [3, 6]\}, \\ P^1 &= \{[1, 6], [3, 5], [2, 4]\}, \\ P^2 &= \{[1, 5], [2, 6], [3, 4]\} \end{aligned} \tag{39}$$

In this section, to construct an (n, k, r, t) LRC over the binary field, we start by selecting any $t \leq s + 1$ partitions (parallel classes) from the $s + 1$ partitions of rm elements, which correspond to s MOLR of size $r \times m$. For each partition P^l, a submatrix M_l of size $m \times rm$ is constructed. Each row of M_l is indexed by a block in the parallel class associated with P^l, while each column corresponds to one of the rm elements. Let $M_l(i, j)$ denote the entry located in the i-th row and j-th column of M_l. The value of $M_l(i, j)$ is 1 if and only if the

element j is a member of block B_i. Finally, t parallel classes are chosen from these $s+1$ partitions to construct a $tm \times rm$ matrix R.

$$R = \begin{pmatrix} M_1 \\ \dots \\ M_t \end{pmatrix}. \tag{40}$$

Since any two blocks intersect at most once in $P^{col}, P^1, \dots, P^s$, it follows that the supports of any two rows in R intersect at most once. Similarly, the supports of any two columns in R also intersect at most once. As R is composed of t parallel classes, each column in R has weight t. Furthermore, because the size of each block in the partitions is r, every row in R has a weight of r. H is defined as follows:

$$H = \left(\begin{array}{c|c} R & I_{(n-k)\times(n-k)} \end{array} \right). \tag{41}$$

Generator matrix G is given by

$$G = \left(\begin{array}{c|c} I_{k\times k} & -R^T \end{array} \right) = \left(\begin{array}{c|c} I_{k\times k} & R^T \end{array} \right). \tag{42}$$

The equivalence of the two expressions for G follows from the fact that the construction is over $\mathbb{F}_2$. From the matrices H and G, we can deduce that $n = tm + rm$ and $k = rm$.

When m is a prime power, there exist at least $m-1$ MOLR, which implies that $t \leq m$. For all other cases, $t \leq N(r)+1$, where $N(r)$ represents the maximum number of MOLS of size m.

The following theorem establishes that the linear code constructed using H as its parity-check matrix and G as its generator matrix is a distance-optimal (n, k, r, t) LRC over the binary field.

Theorem 4. *The locally repairable code $\mathcal{C}$, designed using H as its parity-check matrix and G as its generator matrix, is an $(n = rm + tm, k = rm, r, t \leq N(r)+1)_2$ IS-LRC. This code achieves the optimal minimum distance that satisfies the bound (2). Here, $N(r)$ represents the maximum number of MOLS of order r. The rate of the code is given by $\mathcal{R}(\mathcal{C}) = \frac{r}{r+t}$. Moreover, when $t = 2$, this construction achieves rate-optimality.*

Proof. First, we prove that every message element from $\mathcal{C}$ has (r, t)-availability. The first k columns of H correspond to the k information symbols in $\mathcal{C}$. Since each column in R (i.e., the first k columns of H) has a weight of t, every message symbol in $\mathcal{C}$ is associated with t recovery sets. Furthermore, because the supports of any two rows in R overlap at most once, these t recovery sets for any message symbol are disjoint. Additionally, since the weight of each row in H is $r+1$, the size of each recovery set is r.

Then we demonstrate that $\mathcal{C}$ is distance-optimal and its distance is $t+1$. Consider any nonzero codeword $\mathbf{c} \in \mathcal{C}$, and let c_i be a nonzero symbol in $\mathbf{c}$. First, assume that c_i lies within the first k coordinates of $\mathbf{c}$. Since every message entry from $\mathcal{C}$ has (r, t)-availability, the symbol c_i has t disjoint recovery sets. Each recovery set contributes at least one additional nonzero symbol. Hence, the codeword $\mathbf{c}$ must have $t+1$ nonzero entries.

Now, suppose c_i is located in one of the last $n-k$ positions of $\mathbf{c}$. From the structure of H, it can be concluded that there is at least one nonzero symbol c_j within the first k coordinates of $\mathbf{c}$. Since c_j possesses (r, t)-availability, the codeword $\mathbf{c}$ must contain $t+1$ nonzero symbols. As $\mathbf{c}$ was chosen arbitrarily, it follows that $d(\mathcal{C}) \geq t+1$.

Notice that each recovery set consists of exactly one parity symbol. Therefore, the minimum distance $d(\mathcal{C})$ satisfies

$$\begin{aligned} d(\mathcal{C}) &\leq n - k - \left\lceil \frac{kt}{r} \right\rceil + t + 1 \\ &= rm + tm - rm - \left\lceil \frac{rmt}{r} \right\rceil + t + 1 \\ &= tm - tm + t + 1 = t + 1. \end{aligned} \tag{43}$$

The minimum distance of $\mathcal{C}$ meets the bound given in (2). Thus, the code $\mathcal{C}$ constructed here is distance-optimal. The rate $\mathcal{R}$ of this code is expressed as

$$\mathcal{R}(\mathcal{C}) = \frac{k}{n} = \frac{rm}{rm + tm} = \frac{r}{r + t}. \tag{44}$$

When $t = 2$, $\mathcal{R}(\mathcal{C})$ reaches the bound (3), making $\mathcal{C}$ rate-optimal for $t = 2$. □

Example 6. *We choose all* 2 *parallel classes from* 2 *MOLR of size* 2×3 *in Example 5. The matrix R, as well as the parity-check matrix H and the generator matrix G, are presented below.*

$$R = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}, \tag{45}$$

$$H = \left(\begin{array}{cccccc|cccccc} 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right), \tag{46}$$

$$G = \left(\begin{array}{cccccc|cccccc} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \end{array}\right). \tag{47}$$

H is a parity-check matrix for $(12, 6, 2, 2)$ *LRC* $\mathcal{C}$*, while G represents its corresponding generator matrix. The code* $\mathcal{C}$ *has a minimum distance of* $d(\mathcal{C}) = 3$ *and a rate of* $\mathcal{R}(\mathcal{C}) = \frac{1}{2}$*. The minimum distance* $d(\mathcal{C})$ *satisfies the bound given in (2), and the rate* $\mathcal{R}(\mathcal{C})$ *meets the bound in (3). Therefore, the code* $\mathcal{C}$ *is both distance-optimal and rate-optimal. Now, suppose the symbol* c_2 *is erased. The erased symbol* c_2 *can be represented by all elements from* $R_1(2) = \{4, 9\}$ *(*$R_2(2) = \{6, 11\}$*).*

6. Optimal Binary Locally Repairable Codes with Nonuniform Locality

In this section, we introduce a construction for locally repairable codes that feature nonuniform locality. We take the construction in Theorem 1 as an example. M is a $tr \times r^2$ matrix from $t \leq s + 2$ partitions (parallel classes) from the $s + 2$ partitions of r^2 elements associated with s MOLS of order r. In this section $t < \frac{r}{u}$, $t \leq s + 2$, $u < r$.

$$M = \begin{pmatrix} M_1 \\ \dots \\ M_t \end{pmatrix}. \tag{48}$$

Note that the supports of any two rows in M overlap at most once. The matrix M is composed of t parallel classes, which ensures that each column in R' has a weight of t. Furthermore, since the size of each block in the partitions is r, every row in M has a weight of r. Next, remove $u < r$ points from M, resulting in a $tr \times (r^2 - u)$ matrix R given by

$$R = \begin{pmatrix} M_1 \\ \vdots \\ M_t \end{pmatrix}. \tag{49}$$

The parity-check matrix H for this construction is specified as

$$H = \left(\begin{array}{c|c} R & I_{(n-k)\times(n-k)} \end{array} \right). \tag{50}$$

Generator matrix G for this construction is defined as

$$G = \left(\begin{array}{c|c} I_{k\times k} & -R^T \end{array} \right) = \left(\begin{array}{c|c} I_{k\times k} & R^T \end{array} \right). \tag{51}$$

Theorem 5. *A code $\mathcal{C}$, constructed using G as its generator matrix and H as its parity-check matrix, is a distance-optimal (2) $(n = r^2 - u + tr, k = r^2 - u, r, t)_2$ locally repairable code, where $t < \frac{r}{u}$, $t \leq s + 2$, $u < r$. This construction provides nonuniform localities while maintaining the same availability across message symbols. $d(\mathcal{C}) = t + 1$, $\mathcal{R}(\mathcal{C}) = \frac{r^2-u}{r^2-u+tr}$.*

Proof.

$$H = \left(\begin{array}{c|c} R & I_{tr\times tr} \end{array} \right) = \left(\begin{array}{c|c} R & I_{(n-k)\times(n-k)} \end{array} \right), \tag{52}$$

where $n - k = tr$, $n = r^2 - u + tr$, and $k = r^2 - u$. Based on the structure of H and R, it can be concluded that the supports of any two rows in H intersect exactly once in initial k coordinates in $\mathcal{C}$. Each information symbol in $\mathcal{C}$ is associated with t disjoint recovery sets, where the size of each recovery set is at most r. Consequently, $\mathcal{C}$ is an IS-LRC. Next, we demonstrate that the distance of $\mathcal{C}$ is $t + 1$. Let $\mathbf{c}$ be an arbitrary codeword in $\mathcal{C}$, and let c_i represent the first nonzero symbol in $\mathbf{c}$. First, consider the scenario where c_i lies within the first k coordinates of $\mathbf{c}$. Since c_i is associated with t disjoint recovery sets, each recovery set contributes at least one additional nonzero symbol to $\mathbf{c}$. Therefore, $\mathbf{c}$ must have at least $t + 1$ nonzero elements.

Now, consider the case where c_i is located in the last $n - k$ coordinates of $\mathbf{c}$. Based on the structure of H, there must be at least one nonzero symbol c_j present in the first k coordinates of $\mathbf{c}$. As a result, $\mathbf{c}$ has $t + 1$ nonzero entries. Hence, we have proven that $d(\mathcal{C}) \geq t + 1$.

$$\mathcal{R}(\mathcal{C}) = \frac{k}{n} = \frac{r^2 - u}{r^2 - u + tr}. \tag{53}$$

From Theorem 1, it follows that a binary locally repairable code $\mathcal{C}'$ with parameters $(n' = n + u = r^2 + tr, k' = k + u = r^2, r' = r, t' = t)$ attains (2), achieving $d(\mathcal{C}') = t' + 1$. Since each recovery set contains exactly one parity symbol, $d(\mathcal{C}')$ satisfies bound in (2)

$$d(\mathcal{C}') = n' - k' - \left\lceil \frac{k't'}{r'} \right\rceil + t' + 1 = t' + 1 = t + 1. \tag{54}$$

In this construction, every recovery set contains exactly one parity symbol. Therefore, we have

$$d(\mathcal{C}) \leq n - k - \left\lceil \frac{kt}{r} \right\rceil + t + 1. \tag{55}$$

If $d(\mathcal{C}) = n - k - \left\lceil \frac{kt}{r} \right\rceil + t + 1$, then the code $\mathcal{C}$ is distance-optimal. By comparing $\lceil \frac{kt}{r} \rceil$ with $\lceil \frac{k't'}{r'} \rceil$, we observe that $\frac{kt}{r} < \frac{k't'}{r'}$. Note that $\frac{k't'}{r'}$ is an integer. When $\frac{k't'}{r'} - \frac{kt}{r} < 1$, it implies $\lceil \frac{kt}{r} \rceil = \lceil \frac{k't'}{r'} \rceil$:

$$\begin{aligned} \frac{k't'}{r'} - \frac{kt}{r} &= \frac{(k+u)t}{r} - \frac{kt}{r} \\ &= \frac{ut}{r}. \end{aligned} \tag{56}$$

Since $t < \frac{r}{u}$, it follows that $\frac{k't'}{r'} - \frac{kt}{r} < 1$. Therefore,

$$\begin{aligned} d(\mathcal{C}) &\leq n - k - \left\lceil \frac{kt}{r} \right\rceil + t + 1 \\ &= n' - k' - \left\lceil \frac{k't'}{r'} \right\rceil + t' + 1 \\ &= t' + 1 \\ &= t + 1. \end{aligned} \tag{57}$$

Finally, since it has already been established that $d(\mathcal{C}) \geq t + 1$, then $d(\mathcal{C}) = t + 1$, meeting (2). Therefore, $\mathcal{C}$ is distance-optimal. □

Remark 5. *In Theorem 5, we gave a construction of locally repairable codes that supports nonuniform locality as a distinguishing feature. Now we give characterization of the locality distribution. In previous sections, the locality of all symbols in locally repairable codes is r, i.e., the hamming weight of all rows in the parity-check matrix H is r + 1. In Theorem 5, the locality distribution depends on the points removed, and the distribution varies when different points are removed. However, we can still provide the average locality of our construction. This is calculated by determining the average weight of each row in the parity-check matrix. In the* $(n = r^2 - u + tr, k = r^2 - u, r, t)_2$ *LRC constructed in Theorem 5, the average weight of each row in H is* $\frac{(r^2-u)t+tr}{tr} = \frac{r^2-u}{r} + 1$. *Therefore, the average locality r in the construcion from Theorem 5 is* $\frac{r^2-u}{r}$.

Example 7. M_1 *and* M_2 *are* 2 *parallel classes from* 2 *MOLS of order* 3 *in Example 1.*

$$M_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, M_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \tag{58}$$

M comes from the above two parallel classes.

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \tag{59}$$

$$R = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{60}$$

The matrix R is obtained from M by removing the last point from each block in M.

$$H = \left(\begin{array}{cccccccc|cccccc} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right), \tag{61}$$

$$G = \left(\begin{array}{cccccccc|cccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right). \tag{62}$$

The resulting code $\mathcal{C}$ is a $(14, 8, 3, 2)$ LRC, with a minimum distance of $d(\mathcal{C}) = 3$ and a rate of $\mathcal{R} = \frac{4}{7}$. This code is distance-optimal. Each message symbol in $\mathcal{C}$ can be repaired by accessing no more than three other coordinates, and there are two repair sets of size at most three.

Now, consider the scenario where the second message symbol, c_2, is erased. We can conclude that $R_1(2) = \{4, 10\}$ and $R_2(2) = \{6, 7, 12\}$. This means that c_2 can be reconstructed by summing the coordinates indexed by $R_1(2)$ or $R_2(2)$, and these two recovery sets are disjoint. Note that c_2 can be repaired by reading only two other symbols.

7. Techniques for Enhancing the Minimum Distance of Locally Repairable Codes

The minimum distance of locally repairable codes is strongly related to their ability to tolerate failures. In this section, we present a set of methods designed to enhance the minimum distance of existing locally repairable codes. The proposed methodology is built upon the following lemma.

Lemma 1. *Let $\mathcal{C}$ be a binary $[n, k]$ code with a minimum distance d, where d is odd. Then, there exists a binary $[n + 1, k]$ code $\mathcal{C}'$ with a minimum distance of $d + 1$ [34].*

Proof. Assume that the parity-check matrix of $\mathcal{C}$ is H. Let $\mathcal{C}'$ denote an extended version of $\mathcal{C}$. In this case, the parity-check matrix of $\mathcal{C}'$ can be expressed as

$$H' = \left(\begin{array}{ccccc|c} & & H & & & 0 \\ & & & & & \vdots \\ & & & & & 0 \\ \hline 1 & 1 & \cdots & & 1 & 1 \end{array}\right). \tag{63}$$

Here, $\mathcal{C}'$ is a linear code with parameters $[n' = n+1, k' = k]$. Next, we show that the minimum distance of $\mathcal{C}'$ is $d+1$. Since d is an odd integer, the weight of $\mathbf{c} \in \mathcal{C}'$ is even. As $\mathcal{C}'$ is formed by extending $\mathcal{C}$, it follows that $d(\mathcal{C}') \geq d(\mathcal{C})$. Consequently, we conclude that $d(\mathcal{C}') = d+1$. □

Theorem 6. *For any distance-optimal binary (n, k, r, t) locally repairable code $\mathcal{C}$ with information symbol (r, t)-availability, where $t+1$ is an odd number, it is possible to construct a distance-optimal $(n+1, k, r, t)$ locally repairable code $\mathcal{C}'$. $\mathcal{C}'$ is an IS-LRC, $d(\mathcal{C}') = t+2$. Hence, $\mathcal{C}'$ is distance-optimal.*

Proof. Let H represent the parity-check matrix of $\mathcal{C}$, and let H' denote the parity-check matrix of $\mathcal{C}'$. Then

$$H' = \left(\begin{array}{ccccc|c} & & H & & & \begin{matrix}0\\ \vdots \\ 0\end{matrix} \\ \hline 1 & 1 & \cdots & 1 & & 1 \end{array}\right). \tag{64}$$

In this construction, the information symbols in $\mathcal{C}'$ continue to be protected by t disjoint repair sets, where each repair set is of size at most r. Furthermore, the positions in $\mathcal{C}'$ corresponding to the parity symbols of $\mathcal{C}$ have a repair set of size r. The main modification in $\mathcal{C}'$ compared to $\mathcal{C}$ is the addition of a new parity bit. This newly added parity bit can be repaired by accessing all n remaining symbols. Additionally, every other symbol in $\mathcal{C}'$ gains an extra repair set of size n. Consequently, $\mathcal{C}'$ is an $(n+1, k, r, t)_2$ IS-LRC that retains (r, t)-availability.

Now, suppose the minimum distance d of $\mathcal{C}$ is odd. By Lemma 1, the minimum distance of $\mathcal{C}'$ becomes $d+1$. First, consider the case where $\mathcal{C}$ satisfies the minimum distance bound (1), given by

$$d = n - k - \left\lceil \frac{t(k-1)+1}{t(r-1)+1} \right\rceil + 2. \tag{65}$$

For $\mathcal{C}'$, the left-hand side increases to $d+1$, while the right-hand side also increments by 1 due to the inclusion of the additional bit. Thus, $\mathcal{C}'$ remains distance-optimal.

Next, consider the case where $\mathcal{C}$ satisfies the minimum distance bound (2), given by

$$d = n - k - \left\lceil \frac{kt}{r} \right\rceil + t + 1. \tag{66}$$

For $\mathcal{C}'$, the left-hand side increases to $d+1$, and the right-hand side also increases by 1. Therefore, $\mathcal{C}'$ continues to be distance-optimal. □

Remark 6. *Our method of increasing the minimum distance of LRCs can be applied to all constructions in this paper when $t+1$ is odd (t is even). Moreover, this method gives an additional repair set for all parity-check symbols in the codes proposed in this paper. When t is even, the parity symbol added in the extended code is the XOR of all symbols in the original code $\mathcal{C}$. Since t is even, there is an even number of parallel classes in the codeword. Thus, in the extended code $\mathcal{C}'$ derived from $\mathcal{C}$, the final parity symbol is equal to the sum of all blocks within any given parallel class. For each parity symbol in $\mathcal{C}$, there will be two disjoint repair sets: one repair set comes from $\mathcal{C}$, and the other recovering set is all other blocks in the same parallel class and the last parity symbol in the extended code.*

Remark 7. *In Theorem 6, the extended code C' is constructed as an $(n+1, k, r, t)$ IS-LRC that retains information-symbol (r, t)-availability. A potential concern arises regarding the newly*

appended parity symbol in C', as it has a repair set of size n, which appears to violate the locality constraint r. However, this does not contradict the definition of an IS-LRC with (r,t)-availability, as stated in Definition 2.

The bound (2) in this paper comes from [16], where the requirement for an IS-LRC is that each information symbol must have a repair set of size at most r. Similarly, Definition 2 in this work also focuses solely on the repair sets of information symbols. Therefore, while the newly appended parity symbol has a repair set of size n, this does not affect the size of the repair sets for the information symbols, which remain consistent with the locality constraint r.

This distinction ensures that the extended code C' satisfies the definition of an IS-LRC with (r,t)-availability, as the locality constraint applies only to the information symbols.

Example 8.

$$H = \left(\begin{array}{cccccccc|cccccc} 1&0&0&0&0&1&0&1&1&0&0&0&0&0\\ 0&0&1&0&1&0&1&0&0&1&0&0&0&0\\ 0&1&0&1&0&0&0&0&0&0&1&0&0&0\\ 1&0&0&0&1&0&0&0&0&0&0&1&0&0\\ 0&1&0&0&0&1&1&0&0&0&0&0&1&0\\ 0&0&1&1&0&0&0&1&0&0&0&0&0&1 \end{array}\right), \tag{67}$$

$$G = \left(\begin{array}{cccccccc|cccccc} 1&0&0&0&0&0&0&0&1&0&0&1&0&0\\ 0&1&0&0&0&0&0&0&0&0&1&0&1&0\\ 0&0&1&0&0&0&0&0&0&1&0&0&0&1\\ 0&0&0&1&0&0&0&0&0&0&1&0&0&1\\ 0&0&0&0&1&0&0&0&0&1&0&1&0&0\\ 0&0&0&0&0&1&0&0&1&0&0&0&1&0\\ 0&0&0&0&0&0&1&0&0&1&0&0&1&0\\ 0&0&0&0&0&0&0&1&1&0&0&0&0&1 \end{array}\right). \tag{68}$$

The matrices H and G serve as the parity-check matrix and generator matrix, respectively, for the $(14,8,3,2)$ LRC $\mathcal{C}$ described in Example 7. This code has a minimum distance of $d(\mathcal{C}) = 3$ and a rate of $\mathcal{R} = \frac{4}{7}$. This code is distance-optimal. By applying the approach outlined in Theorem 6, we construct a parity-check matrix H' and a generator matrix G' for the $(15,8,3,2)$ LRC $\mathcal{C}'$.

$$H' = \left(\begin{array}{cccccccc|cccccc|c} 1&0&0&0&0&1&0&1&1&0&0&0&0&0&0\\ 0&0&1&0&1&0&1&0&0&1&0&0&0&0&0\\ 0&1&0&1&0&0&0&0&0&0&1&0&0&0&0\\ 1&0&0&0&1&0&0&0&0&0&0&1&0&0&0\\ 0&1&0&0&0&1&1&0&0&0&0&0&1&0&0\\ 0&0&1&1&0&0&0&1&0&0&0&0&0&1&0\\ \hline 1&1&1&1&1&1&1&1&1&1&1&1&1&1&1 \end{array}\right), \tag{69}$$

$$G' = \left(\begin{array}{cccccccc|cccccc|c} 1&0&0&0&0&0&0&0&1&0&0&1&0&0&1\\ 0&1&0&0&0&0&0&0&0&0&1&0&1&0&1\\ 0&0&1&0&0&0&0&0&0&1&0&0&0&1&1\\ 0&0&0&1&0&0&0&0&0&0&1&0&0&1&1\\ 0&0&0&0&1&0&0&0&0&1&0&1&0&0&1\\ 0&0&0&0&0&1&0&0&1&0&0&0&1&0&1\\ 0&0&0&0&0&0&1&0&0&1&0&0&1&0&1\\ 0&0&0&0&0&0&0&1&1&0&0&0&0&1&1 \end{array}\right). \tag{70}$$

The code C' is a $(15,8,3,2)$ LRC with information (r,t)-availability. Suppose the second message symbol c_2 is erased. From H' or G', we find that $R_1(2) = \{4,10\}$ and $R_2(2) = \{6,7,12\}$.

Additionally, the parity symbol in $\mathcal{C}'$ is associated with two disjoint recovery sets. If c_9 in a codeword $\mathbf{c} \in \mathcal{C}'$ is erased, it can be reconstructed using either $R_1(9) = \{1, 6, 8\}$ or $R_2(9) = \{10, 11, 15\}$.

Remark 8. *In this work, we propose a construction of binary LRCs inspired by combinatorial techniques. This can also be applied to q-ary LRCs. It is important to note that while multi-bit errors in q-ary LRCs can be interpreted as single errors due to the field structure, similar constructions can be analogously applied in binary LRCs to repair multi-bit errors. However, one key distinction lies in the repair bandwidth: in q-ary LRCs, even a single-bit error in the underlying storage requires repairing the entire symbol, resulting in higher repair bandwidth.*

In contrast, the binary construction proposed in this paper retains the advantage of low repair bandwidth, as repairs are performed at the bit level. This characteristic makes the proposed binary LRCs more efficient in scenarios where repair bandwidth is critical. We emphasize this distinction to highlight the practical advantages of the proposed construction.

8. Discussion

In this section, we analyze and compare the parameters of our proposed constructions with those from related works. Table 1 provides an overview of the parameters for all constructions presented in this paper. Additionally, Tables 2 and 3 outline the parameters of various locally repairable codes with information-symbol (r, t)-availability, including those proposed in this paper as well as other related constructions.

Here is the analysis of the percentage of efficient recovery for inclusion in the paper. Efficient recovery refers to the repair scenario where, when i nodes fail, the total repair bandwidth is less than or equal to ir. Since most node failures involve single-node and double-node failures [20], we provide the percentage of efficient recovery for these two scenarios. In the case of single-node failure, all repairs are efficient; thus, the percentage of efficient recovery is 100% (the repair of the additional parity bits requires the remaining n nodes in Theorem 6, so the percentage of efficient recovery is $\frac{n}{n+1}$). Next, consider the case of two-node failure. There are $\binom{n}{2}$ possible combinations of failed nodes, where inefficient recovery occurs if and only if the two failed nodes belong to the same repair set. Therefore, in the case of two-node failure, the percentage of efficient recovery is at least $\frac{\binom{n}{2} - \frac{nt(r-1)}{2}}{\binom{n}{2}} = \frac{n-1-t(r-1)}{n-1}$. Next, we analyze the complexity of the proposed codes. In LRCs, each parity bit is computed as a linear combination of data symbols in a local group. The encoding complexity of the proposed LRCs is $O(kr)$, where k is the number of data symbols and r is the size of each local repair group. The decoding complexity for a single-node failure is $O(r)$, as the repair only requires accessing symbols from a single local group. For comparison, traditional Reed–Solomon codes require $O(k)$ complexity for both encoding and decoding.

Then we compare our constructions with the codes introduced in [28,29]. Codes constructed in our paper contain some constructions in code constructed in [28]. The construction outlined in Theorem 5 produces locally repairable codes with nonuniform localities, and its parameters cannot be derived from those in [28]. Moreover, all codes obtained through our constructions are distance-optimal, whereas not all constructions in [28] achieve this property. Additionally, we propose a method to improve the minimum distance of (n, k, r, t) locally repairable codes when t is even. The codes constructed using Theorem 6 result in distance-optimal LRCs with an enhanced minimum distance.

Since the constructions in this paper are based on MOLS, the proposed parameter set includes configurations that cannot be obtained using Balanced Incomplete Block Design (BIBD). Therefore, these parameters cannot be derived from the constructions presented in Zhang et al.'s work [29].

Next, we compare constructions in this article with several distance-optimal LRCs introduced in [7–10,19]. Locally repairable codes with (r, δ)-locality, as presented in [7], aim to improve fault tolerance within each repair group. In these codes, each symbol is associated with multiple repair group options; however, these repair groups are not required to be disjoint. In contrast, LRCs with multiple disjoint recovery sets, such as those in [8–10,19], enable parallel data recovery for frequently accessed symbols.

Su [19] proposed a series of distance-optimal locally repairable codes with information symbol (r, t)-availability. These were based on (k, b, c, r) resolvable configurations combined with either an $(N + t, k)$ MDS code or an $(N + t - 1, k)$ Gabidulin code. The first construction in [19] produced $(N + \frac{kt}{r}, k, r, t)_q$ IS-LRCs without all-symbol locality, with a minimum distance of $N - k + t + 1$, where $q \geq N + t$ and $r \mid k$. The second construction generated distance-optimal $(N + \frac{N}{r} + (t-1)\frac{kt}{r}, k, r, t)_{q^M}$ IS-LRCs with all-symbol locality, where $M \geq N + t - 1$. Compared to these constructions, our distance-optimal codes have higher code rates while requiring a field of size 2, whereas the codes [19] require the field size to grow exponentially with n. Although our constructions may have a slightly lower minimum distance, they still satisfy the bound in (2) and offer advantages such as smaller field sizes, higher rates, and all-symbol locality.

In [16], there is an open problem regarding whether $r \mid k$ is necessary for (n, k, r, t) to meet (2). This question was addressed in [9,29], who constructed distance-optimal LRCs when $r \nmid k$. Cai [9] introduced construction from $(k, R, 1)$-packings, which extend PBDs by ensuring that each pair of points appears in at most one block. They provided distance-optimal constructions with parameters $(n = k + \frac{kt}{r}, k, r, t)$, $d = t + 1$, and $\mathcal{R} = \frac{r}{r+t}$ when $r \mid kt$. Our constructions achieve similar parameters in these cases. Additionally, Cai et al. proposed constructions for cases where $r \nmid kt$, with parameters summarized in Table 3. The codes from Theorem 5 also give constructions of optimal locally repairable codes when $r \nmid kt$. Furthermore, we addressed the open question posed in [16].

They [10] gave constructions of distance-optimal LRCs. The codes attain (1) and are $(n = k(1 + rt), k, r, t)_q$ IS-LRCs with information (r, t)-availability for any r and t. These codes were the first explicit constructions achieving the bound in (1). Although the minimum distance of our constructions is smaller compared to [10], their code rate is $\frac{1}{rt+1}$, and they require a much larger field size than 2. In contrast, our binary distance-optimal constructions achieve higher code rates and are better suited for practical applications with the price of a slightly lower distance attaining bound (2). Notice that the codes in [10] target a strictly stronger distance bound (1), which inherently results in lower code rates.

Tan et al. [8] obtains optimal LRCs based on linear algebra and partial geometries. Our constructions encompass some parameters from [8]. Specifically, their codes derived from linear algebra have parameters $(n = r^2 + tr, k = r^2, r, t)$, where $t \leq r$ and r is a prime. These codes are included in our constructions based on MOLS. Additionally, when $r \nmid k$, our code is also distance-optimal. This is not achievable in [19] or [8].

Table 2. Summary of partial known LRCs with availability, part 1.

Reference	Field Size	Underlying Structure	Code Parameters	Minimum Distance $d(\mathcal{C})$	Rate $\mathcal{R}(\mathcal{C})$	Special Features
[28]	2	$(\hat{v},\hat{b},\hat{r},\hat{k},1)$-BIBD	$n=\hat{v}+\hat{b}, k=\hat{b}, r=\hat{r}, t=\hat{k}$	$d(\mathcal{C})=t+1$, distance-optimal	$\mathcal{R}(\mathcal{C})=\frac{r}{r+t}$, rate-optimal for $t=2$	All-symbol locality
[28]	2	$(\hat{v},\hat{b},\hat{r},\hat{k},1)$-BIBD	$n=\hat{v}+\hat{b}, k=\hat{v}, r=\hat{k}, t=\hat{r}$	$d(\mathcal{C})=t+1$, distance-optimal	$\mathcal{R}(\mathcal{C})=\frac{r}{r+t}$, rate-optimal for $t=2$	All-symbol locality
[28]	2	$(\hat{v},\hat{b},\hat{r},\hat{k},1)$-RBIBD	$n=\hat{v}+t\frac{\hat{v}}{\hat{k}}, k=\hat{v}, r=\hat{k}, t\leq\hat{r}$	$d(\mathcal{C})=t+1$, distance-optimal	$\mathcal{R}(\mathcal{C})=\frac{r}{r+t}$, rate-optimal for $t=2$	All-symbol locality
[19]	$\geq N+t$	$(N+t,k)$ MDS code, (k,b,c,r) resolvable configuration	$(N+\frac{kt}{r},k,r,t)$	$d(\mathcal{C})=N-k+t+1$, distance-optimal	$\mathcal{R}(\mathcal{C})=\frac{k}{N+\frac{kt}{r}}$	Without all-symbol locality
[19]	$\geq q^{N+t-1}$	$(N+t-1,k)$ Gabidulin code, (k,b,c,r) resolvable configuration	$(N+\frac{N}{r}+(t-1)\frac{kt}{r},k,r,t)$	$d(\mathcal{C})=N+\frac{N}{r}-k-\frac{k}{r}+t+1$, distance-optimal	$\mathcal{R}(\mathcal{C})=\frac{k}{N+\frac{N}{r}+(t-1)\frac{kt}{r}}$	All-symbol locality
[29]	2	$(\hat{v}-1,\{\hat{k},\hat{k}-1\},1)$-PBD from $(\hat{v},\hat{b},\hat{r},\hat{k},1)$-RBIBD	$n=\hat{v}-1+t\frac{\hat{v}}{\hat{k}}, k=\hat{v}-1, r=\hat{k}, t<r$	$d(\mathcal{C})=t+1$, distance-optimal	$\mathcal{R}(\mathcal{C})=\frac{\hat{v}-1}{\hat{v}-1+t\frac{\hat{v}}{\hat{k}}}$	Nonuniform locality, all-symbol locality
[29]	2	$(\hat{v},\hat{b},\hat{r},\hat{k},1)$-RBIBD	$n=\hat{v}-s+t\frac{\hat{v}}{\hat{k}}, k=\hat{v}-s, s+1\leq r=\hat{k}, t<\frac{r}{s}$	$d(\mathcal{C})=t+1$, distance-optimal	$\mathcal{R}(\mathcal{C})=\frac{\hat{v}-s}{\hat{v}-s+t\frac{\hat{v}}{\hat{k}}}$	Nonuniform locality, all-symbol locality
Theorem 1	2	$N(r)$ MOLS of order r	$(n=r^2+tr, k=r^2, r, t\leq N(r)+2)_2$	$d(\mathcal{C})=t+1$, distance-optimal	$\mathcal{R}(\mathcal{C})=\frac{r}{r+t}$, rate-optimal for $t=2$	Arbitrary locality r
Remark 1	2	A Latin square of order r	$(n=r^2+tr, k=r^2, r, t\leq 3)_2$	$d(\mathcal{C})=t+1$, distance-optimal	$\mathcal{R}(\mathcal{C})=\frac{r}{r+t}$, rate-optimal for $t=2$	Arbitrary locality r
Theorem 2	2	$N(t)$ MOLS of order t	$(n=t^2+rt, k=rt, r\leq N(t)+2, t)_2$	$d(\mathcal{C})=t+1$, distance-optimal	$\mathcal{R}(\mathcal{C})=\frac{r}{r+t}$, rate-optimal for $t=2$	Arbitrary availability t
Theorem 3	2	$N(r)$ MOLS of order r	$(n=mr^2+mtr, k=mr^2, r, t\leq N(r)+2)_2$	$d(\mathcal{C})=t+1$, distance-optimal	$\mathcal{R}(\mathcal{C})=\frac{r}{r+t}$, rate-optimal for $t=2$	Large block length
Theorem 4	2	$N(r)$ MOLR of size $r\times m$	$(n=rm+tm, k=rm, r, t\leq N(r)+1)_2$	$d(\mathcal{C})=t+1$, distance-optimal	$\mathcal{R}(\mathcal{C})=\frac{r}{r+t}$, rate-optimal for $t=2$	Arbitrary locality r, small locality
Theorem 5	2	$s\leq N(r)$ MOLS of order r	$(n=r^2-u+tr, k=r^2-u, r, t<\frac{r}{u})_2$, $t\leq s+2, u<r$	$d(\mathcal{C})=t+1$, distance-optimal	$\mathcal{R}(\mathcal{C})=\frac{r^2-u}{r^2-u+tr}$	Nonuniform locality

Table 3. Summary of partial known LRCs with availability, part 2.

Reference	Field Size	Underlying Structure	Code Parameters	Minimum Distance $d(\mathcal{C})$	Rate $\mathcal{R}(\mathcal{C})$	Special Features
[9]	2	$(k, r, 1)$-packing	$n = k + \frac{kt}{r}, k, r, t$	$d(\mathcal{C}) = t + 1$, distance-optimal	$\mathcal{R}(\mathcal{C}) = \frac{r}{r+t}$, rate-optimal for $t = 2$	All-symbol locality, $r \mid kt$
[9]	$\geq k + d - 1$	$(rm, r, 1)$-packing, $(k + d - 1, k, d)$-MDS code	$n = k + tm - t + d - 1, k = rm - v, r, t$	$d(\mathcal{C}) = d$, distance-optimal	$\mathcal{R}(\mathcal{C}) = \frac{k}{k+tm-t+d-1}$	Without all-symbol locality, $v \in [m - 1], r, t \in [m]$
[9]	$\geq k + d - 1$	$(k, \{p^a, p^a - 1\}, 1)$-packing, $(k + d - 1, k, d)$-MDS code	$n = k + (\lceil \frac{k}{r} \rceil - 1)t + d - 1, k, r = p^a, t$	$d(\mathcal{C}) = d$, distance-optimal	$\mathcal{R}(\mathcal{C}) = \frac{k}{k+(\lceil \frac{k}{r} \rceil - 1)t+d-1}$	Without all-symbol locality, p is a prime
[9]	$\geq k + d - 1$	$(k, \{r, r - 1\}, 1)$-packing, $(k + d - 1, k, d)$-MDS code	$n = k + (\lceil \frac{k}{r} \rceil - 1)t + d - 1, k, r, t$	$d(\mathcal{C}) = d$, distance-optimal	$\mathcal{R}(\mathcal{C}) = \frac{k}{k+(\lceil \frac{k}{r} \rceil - 1)t+d-1}$	Without all-symbol locality, $r = 3, 4$
[10]	$\geq q_1^{k(1+(r-1)t)}$	Gabidulin Code	$n = k(1 + rt), k, r, t$	$d(\mathcal{C}) = n - k - \left\lceil \frac{t(k-1)+1}{t(r-1)+1} \right\rceil + 2$	$\mathcal{R}(\mathcal{C}) = \frac{1}{1+rt}$	$q_1 \geq r + 1$
[10]	$\geq q_1^{1+(r-1)t}$	Linearized Reed–Solomon Code	$n = k(1 + rt), k, r, t$	$d(\mathcal{C}) = n - k - \left\lceil \frac{t(k-1)+1}{t(r-1)+1} \right\rceil + 2$	$\mathcal{R}(\mathcal{C}) = \frac{1}{1+rt}$	$q_1 \geq k + 1$
[8]	2	Linear algebra	$n = r^2 + tr, k = r^2, r, t$	$d(\mathcal{C}) = t + 1$, distance-optimal	$\mathcal{R}(\mathcal{C}) = \frac{r}{r+t}$, rate-optimal for $t = 2$	r is a prime, $t \leq r$
[8]	2	Linear algebra	$n = r^t + tr^{t-1}, k = r^t, r, t$	$d(\mathcal{C}) = t + 1$, distance-optimal	$\mathcal{R}(\mathcal{C}) = \frac{r}{r+t}$, rate-optimal for $t = 2$	None
[8]	2	$\hat{b} \times \hat{v}$ incidence matrix of a partial geometry $PG(s + 1, u + 1)$	$n = \hat{v} + \hat{b}, k = \hat{b}, r = u + 1, t = s + 1$	$d(\mathcal{C}) = t + 1$, distance-optimal	$\mathcal{R}(\mathcal{C}) = \frac{r}{r+t}$, rate-optimal for $t = 2$	None
[8]	2	Dual code of codes from $PG(s + 1, u + 1)$	$n = \hat{v} + \hat{b}, k = \hat{v}, r = s + 1, t = u + 1$	$d(\mathcal{C}) = t + 1$, distance-optimal	$\mathcal{R}(\mathcal{C}) = \frac{r}{r+t}$, rate-optimal for $t = 2$	None

9. Conclusions

This paper presents a series of constructions for locally repairable codes (LRCs) utilizing mutually orthogonal Latin squares (MOLS). We propose optimal binary LRCs with information-symbol locality, which include previously known codes from [28]. All constructions introduced in this work guarantee distance optimality under the bound specified in (2). Furthermore, we give a construction of LRCs with nonuniform localities. When $t = 2$, they are rate-optimal.

In addition, we give LRCs with nonuniform localities with multiple disjoint repair sets. This construction solves the open problem from [16]. This construction generalizes certain codes proposed in [8,9,28] and achieves higher code rates compared to the codes in [9,10,19], while requiring significantly smaller finite field sizes. Additionally, our construction generates codes with parameters not covered by [8–10,19,28].

By conducting comprehensive comparisons with other constructions [8–10,19,28,29], we demonstrate that our framework extends earlier codes. Furthermore, the binary structure of our codes makes them particularly well suited for practical deployment in dis-

tributed storage systems. In conclusion, this paper offers a unified and versatile framework for designing high-performance locally repairable codes, with both theoretical and practical significance.

Author Contributions: Conceptualization, Y.Z. and X.Z.; methodology, Y.Z.; writing—original draft preparation, N.C., Y.Z., X.Z. and L.Z.; writing—review and editing, N.C., Y.Z., X.Z. and L.Z.; supervision, Y.Z.; funding acquisition, X.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China (Grant No. 22478245) and Shanghai Educational Research Project (Grant No. C2023068).

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Ghemawat, S.; Gobioff, H.; Leung, S.-T. The Google File System. In *Proceedings of the Nineteenth ACM Symposium on Operating Systems Principles*; Association for Computing Machinery: New York, NY, USA, 2003; pp. 29–43. [CrossRef]
2. Sathiamoorthy, M.; Asteris, M.; Papailiopoulos, D.; Dimakis, A.G.; Vadali, R.; Chen, S.; Borthakur, D. XORing Elephants: Novel Erasure Codes for Big Data. *Proc. VLDB Endow.* **2013**, *6*, 325–336. [CrossRef]
3. Gopalan, P.; Huang, C.; Simitci, H.; Yekhanin, S. On the Locality of Codeword Symbols. *IEEE Trans. Inform. Theory* **2012**, *58*, 6925–6934. [CrossRef]
4. Jin, L.; Kan, H.; Zhang, Y. Constructions of Locally Repairable Codes with Multiple Recovering Sets via Rational Function Fields. *IEEE Trans. Inform. Theory* **2020**, *66*, 202–209. [CrossRef]
5. Jin, L. Explicit Construction of Optimal Locally Recoverable Codes of Distance 5 and 6 via Binary Constant Weight Codes. *IEEE Trans. Inform. Theory* **2019**, *65*, 4658–4663. [CrossRef]
6. Prakash, N.; Kamath, G.M.; Lalitha, V.; Kumar, P.V. Optimal Linear Codes with a Local-Error-Correction Property. In *2012 IEEE International Symposium on Information Theory Proceedings*; IEEE: Cambridge, MA, USA, 2012; pp. 2776–2780. [CrossRef]
7. Cai, H.; Miao, Y.; Schwartz, M.; Tang, X. A Construction of Maximally Recoverable Codes with Order-Optimal Field Size. *IEEE Trans. Inf. Theory* **2022**, *68*, 204–212. [CrossRef]
8. Tan, P.; Zhou, Z.; Sidorenko, V.; Parampalli, U. Two Classes of Optimal LRCs with Information (r, t)-Locality. *Des. Codes Cryptogr.* **2020**, *88*, 1741–1757. [CrossRef]
9. Cai, H.; Cheng, M.; Fan, C.; Tang, X. Optimal Locally Repairable Systematic Codes Based on Packings. *IEEE Trans. Commun.* **2019**, *67*, 39–49. [CrossRef]
10. Cai, H.; Miao, Y.; Schwartz, M.; Tang, X. On Optimal Locally Repairable Codes with Multiple Disjoint Repair Sets. *IEEE Trans. Inf. Theory* **2020**, *66*, 2402–2416. [CrossRef]
11. Yavari, E.; Esmaeili, M. Locally Repairable Codes: Joint Sequential–Parallel Repair for Multiple Node Failures. *IEEE Trans. Inf. Theory* **2020**, *66*, 222–232. [CrossRef]
12. Yavari, E.; Esmaeili, M.; Rifà, J. Locally Repairable Codes Based on Permutation Cubes and Latin Squares. *IEEE Access* **2022**, *10*, 9686–9702. [CrossRef]
13. Hao, J.; Xia, S.-T.; Shum, K.W.; Chen, B.; Fu, F.-W.; Yang, Y. Bounds and Constructions of Locally Repairable Codes: Parity-Check Matrix Approach. *IEEE Trans. Inf. Theory* **2020**, *66*, 7465–7474. [CrossRef]
14. Teng, J.; Jin, L. Constructions of Binary Locally Repairable Codes with Multiple Recovering Sets. *IEEE Access* **2021**, *9*, 92239–92245. [CrossRef]
15. Wang, A.; Zhang, Z. Repair Locality with Multiple Erasure Tolerance. *IEEE Trans. Inform. Theory* **2014**, *60*, 6979–6987. [CrossRef]
16. Rawat, A.S.; Papailiopoulos, D.S.; Dimakis, A.G.; Vishwanath, S. Locality and Availability in Distributed Storage. *IEEE Trans. Inf. Theory* **2016**, *62*, 4481–4493. [CrossRef]
17. Hao, J.; Shum, K.W.; Xia, S.-T.; Li, D. Optimal Locally Repairable Codes for Parallel Reading. *IEEE Access* **2020**, *8*, 80447–80453. [CrossRef]
18. Su, Y.-S. Design of Membership Matrices for (r, t)-Availability in Distributed Storage. In Proceedings of the 2016 IEEE International Symposium on Information Theory (ISIT), Barcelona, Spain, 10–15 July 2016; pp. 998–1002. [CrossRef]
19. Su, Y.-S. On the Construction of Local Parities for (r, t) -Availability in Distributed Storage. *IEEE Trans. Commun.* **2017**, *65*, 2332–2344. [CrossRef]

20. Prakash, N.; Lalitha, V.; Balaji, S.B.; Vijay Kumar, P. Codes with Locality for Two Erasures. *IEEE Trans. Inf. Theory* **2019**, *65*, 7771–7789. [CrossRef]
21. Tamo, I.; Barg, A.; Frolov, A. Bounds on the Parameters of Locally Recoverable Codes. *IEEE Trans. Inf. Theory* **2016**, *62*, 3070–3083. [CrossRef]
22. Zeh, A.; Yaakobi, E. Bounds and Constructions of Codes with Multiple Localities. In *2016 IEEE International Symposium on Information Theory (ISIT)*; IEEE: Barcelona, Spain, 2016; pp. 640–644. [CrossRef]
23. Balaji, S.B.; Kumar, P.V. Bounds on the Rate and Minimum Distance of Codes with Availability. In *2017 IEEE International Symposium on Information Theory (ISIT)*; IEEE: Aachen, Germany, 2017; pp. 3155–3159. [CrossRef]
24. Fang, W.; Tao, R.; Fu, F.-W.; Chen, B.; Xia, S.-T. Bounds and Constructions of Singleton-Optimal Locally Repairable Codes with Small Localities. *IEEE Trans. Inf. Theory* **2024**, *70*, 6842–6856. [CrossRef]
25. Zhang, W.; Luo, Y.; Wang, L. Optimal Binary and Ternary Locally Repairable Codes with Minimum Distance 6. *Des. Codes Cryptogr.* **2024**, *92*, 1251–1265. [CrossRef]
26. Hao, J.; Xia, S.-T. Constructions of Optimal Binary Locally Repairable Codes with Multiple Repair Groups. *IEEE Commun. Lett.* **2016**, *20*, 1060–1063. [CrossRef]
27. Zhang, Y.; Kan, H. Locally Repairable Codes with Strict Availability from Linear Functions. *Sci. China Inf. Sci.* **2018**, *61*, 109304. [CrossRef]
28. Zhang, Y.; Kan, H. Locally Repairable Codes from Combinatorial Designs. *Sci. China Inf. Sci.* **2020**, *63*, 122304. [CrossRef]
29. Zhang, Y.; Zeng, X. Construction of Binary Locally Repairable Codes with Nonuniform Locality and Availability Using Combinatorial Designs. *Entropy* **2025**, *27*, 269. [CrossRef] [PubMed]
30. Wang, J.; Shen, K.; Liu, X.; Yu, C. Construction of Binary Locally Repairable Codes with Optimal Distance and Code Rate. *IEEE Commun. Lett.* **2021**, *25*, 2109–2113. [CrossRef]
31. Keedwell, A.D.; Dénes, J. *Latin Squares and Their Applications*, 2nd ed.; Elsevier: Amsterdam, The Netherlands, 2015; ISBN 978-0-444-63555-6. [CrossRef]
32. Wan, Z.-X. *Design Theory*; World Scientific/Higher Education Press: Beijing, China, 2009; ISBN 978-981-4287-42-5. [CrossRef]
33. Jäger, G.; Öhman, L.-D.; Markström, K.; Shcherbak, D. Enumeration of Sets of Mutually Orthogonal Latin Rectangles. *Electron. J. Comb.* **2024**, *31*, P1.53. [CrossRef] [PubMed]
34. Ling, S.; Xing, C. *Coding Theory: A First Course*; Cambridge University Press: Cambridge, UK, 2010; ISBN 978-0-521-52923-5. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.