SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

June 2024, Vol. 67, Iss. 6, 162306:1–162306:15
https://doi.org/10.1007/s11432-023-3863-y

Constructions of optimal binary locally repairable codes via intersection subspaces

Wenqin ZHANG, Deng TANG*, Chenhao YING & Yuan LUO*

School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240*, China*

Received 28 March 2023/Revised 25 July 2023/Accepted 7 September 2023/Published online 29 May 2024

Abstract Locally repairable codes (LRCs), which can recover any symbol of a codeword by reading only a small number of other symbols, have been widely used in real-world distributed storage systems, such as Microsoft Azure Storage and Ceph Storage Cluster. Since binary linear LRCs can significantly reduce coding and decoding complexity, constructions of binary LRCs are of particular interest. The aim of this paper is to construct dimensional optimal binary LRCs with disjoint local repair groups. We introduce a method to connect intersection subspaces with binary LRCs and construct dimensional optimal binary linear LRCs with locality 2^b ($b \geqslant 3$) and minimum distance $d \geqslant 6$ by employing intersection subspaces deduced from the direct sum. This method will sufficiently increase the number of possible repair groups of dimensional optimal LRCs, thus efficiently expanding the range of the construction parameters while keeping the largest code rates compared with all known binary linear LRCs with minimum distance $d \geqslant 6$ and locality 2^b.

Keywords locally repairable codes, disjoint local repair groups, distributed storage systems, intersection subspaces, direct sum

1 Introduction

Efficient distributed storage systems (DSSs) provide access to data by storing it in a distributed manner across several storage nodes. Data loss and unavailability could happen in a DSS due to the unreliability of individual nodes. A classical technique used in the storage system is replication schemes. In such a scheme, copies of data packets are stored across different nodes. This scheme provides high reliability and availability. A disadvantage of this scheme is that replication has very high storage overhead. In the case of accelerated and relentless data growth, a new technique is necessary.

Erasure coding has been widely used in distributed storage systems, such as Windows Azure Storage [1] and Facebook Analytics Hadoop cluster [2], because of its higher fault-tolerance values and lower storage overheads. The failure node can be repaired by calculating the redundancy out of the original data over the erasure channel. For an erasure code with length n, dimension k, and minimum distance d, any $d-1$ failures can be repaired by contacting at least k other nodes. Among these, the traditional maximum distance separable (MDS) erasure codes are optimal in terms of storage overhead. However, in the case of a single-node failure, the traditional MDS codes require connecting a large subset of surviving nodes, which will lead to an increase in the complexity of network traffic and the amount of input/output (I/O) operations. Consequently, regenerating codes [3] and codes with locality (known more commonly as locally repairable codes) [4] were introduced in such a scenario. Regenerating codes can efficiently repair a failure node by minimizing the number of transmitted symbols. There are some studies of regenerating codes in [5–7]. Nevertheless, the number of nodes contacted for repair can be a bottleneck for the system efficiency. Hence, locally repairable codes (LRCs) were introduced to optimize the number of disk reads required to repair a single-node failure. This paper is devoted to the construction of locally repairable codes with disjoint repair groups and good parameters.

* Corresponding author (email: dengtang@sjtu.edu.cn, luoyuan@cs.sjtu.edu.cn)

1.1 Code with locality and known results

Let q be a power of an arbitrary prime and $\mathbb{F}_q$ be the finite field with q elements. Let $\mathcal{C}$ be an $[n,k,d]_q$ linear code with length n, dimension k, and minimum distance d over $\mathbb{F}_q$. The code $\mathcal{C}$ is called an LRC with locality r if each code symbol c_i in a codeword $c \in \mathcal{C}$ can be recovered by downloading at most r other symbols. Let such a code $\mathcal{C}$ denote an r-LRC. When $q=2$, we omit q from the notation $[n,k,d]_q$. In addition, the set of such r symbols that can repair the ith symbol is called a "repair set".

LRCs are well studied and many studies have been done (see [4, 8–12]) to explore the relationship between parameters n,k,d, and r. For an $[n,k,d]_q$ LRC with locality r, Gopalan et al. [4] proved the well-known Singleton-like bound as

$$d \leqslant n-k-\left\lceil \frac{k}{r} \right\rceil + 2, \tag{1}$$

where $\lceil \cdot \rceil$ stands for the ceiling function. An LRC is said to be d-optimal if it satisfies (1) with equality for given n,k, and r. When $r=k$, the bound (1) specializes to the classical Singleton bound $d \leqslant n-k+1$. Over the past few years, many constructions of optimal LRCs which achieve bound (1) have been presented. Tamo et al. [13] proposed optimal LRCs over a finite field of size $q \geqslant n+1$ via subcodes of Reed-Solomon codes. Hao et al. [14] designed optimal LRCs with $d=3,4$ over a finite field of size $q \geqslant r+2$. By automorphism groups of elliptic curves, Ref. [15] constructed optimal locally repairable codes with length up to $q+\sqrt{q}$. In addition, the constructions of optimal LRCs based on Reed-Solomon codes, among other techniques, have been recently discovered in [16–18]. Notice that bound (1) has been proved to be tight for some special cases with large alphabet sizes according to the construction provided in [4]. When it comes to small fields, parameters of the optimal constructions become very restrictive [19, 20].

In practice, codes over small alphabets attract more attention particularly in the application of storage because of their ease for implementation. In 2013, Cadambe and Mazumdar [21] derived a new bound for $[n,k,d]_q$ LRCs which took the size of the alphabet into account. This bound is known as C-M bound. They showed that the dimension k of an $[n,k,d]_q$ LRC with locality r is upper bounded by

$$k \leqslant \min_{t \in \mathbb{Z}^+} \left\{ \mathrm{tr} + k_{\mathrm{opt}}^{(q)}(n-t(r+1),d) \right\}, \tag{2}$$

where $k_{\mathrm{opt}}^{(q)}(n,d)$ is the largest possible dimension of a code with length n for a given alphabet size q and a given minimum distance d. This bound applies to both linear and nonlinear codes. Later in [21, 22], explicit constructions of the family of binary LRCs are proposed which achieve the bound (2). However, because the exact value of $k_{\mathrm{opt}}^{(q)}(n,d)$ can only be obtained in a limited case with relatively short code length, it is difficult to apply the C-M bound to evaluate the optimality of general LRCs.

The original LRCs only support the repair of a single-node failure. To address the problem of multiple nodes failures in practical scenarios, the concept of original LRCs was further generalized to LRCs with (r,δ)-locality by Prakash et al. [23]. When $\delta=2$, an LRC with $(r,2)$-locality is reduced to an LRC with locality r. In 2019, Grezet et al. [24] used consecutive residual codes and Griesmer bound $\mathcal{G}(k,d)=\sum_{i=0}^{k-1}\lceil \frac{d}{q^i} \rceil \leqslant n$ to derive a new alphabet-dependent bound for an (r,δ)-LRC with parameters $[n,k,d]$:

$$k \leqslant \min_{\ell \in \mathbb{Z}^+} \left\{ \ell + k_{\mathrm{opt}}^{(q)}(n-(a+1)\mathcal{G}(\kappa,\delta)+\mathcal{G}(\kappa-b,\delta),d) \right\}, \tag{3}$$

where κ is the upper bound on dimension of local codes and $a,b \in \mathbb{Z}$ satisfy $\ell = a\kappa+b, 0 \leqslant b \leqslant \kappa$. Notably, the bound (3) is tighter than the bound C-M bound when $\kappa < r$ or $\delta > q$. Consequently, one tends to use the bound (3) for binary (r,δ)-LRCs with $\delta > 2$. In the past decade, many results have been obtained for (r,δ)-LRCs [25–29]. In addition, a lot of progress on the study of derivatives of LRC has been made, such as LRCs with locality and availability [30–32], maximally recoverable LRCs [1, 33], and scalable local reconstruction code [34].

Recently, Wang et al. [11] presented a sphere-packing bound for binary LRCs based on disjoint local repair groups, which serves as a generalization of the bounds in [4, 35]. Considering binary linear LRCs with minimum distance $d \geqslant 5$, the dimension k is actually upper bounded by the largest integer no greater than the following explicit bound [11] given in (4). For any $[n,k,d]$ binary linear LRCs with locality r

such that $d \geqslant 5$ and $2 \leqslant r \leqslant \frac{n}{2} - 2$, it follows that

$$k \leqslant \frac{rn}{r+1} - \min\left\{\log_2\left(1+\frac{rn}{2}\right), \frac{rn}{(r+1)(r+2)}\right\}. \tag{4}$$

We say that a binary linear LRC is k-optimal if it satisfies the bound (4) with equality for given n, d, and r. This paper will focus on a general assumption $n \geqslant 5(r+1)(r+2)$ that will be satisfied in the main results, and thus the bound (4) can be further simplified to be

$$k \leqslant \frac{rn}{r+1} - \log_2\left(1+\frac{rn}{2}\right).$$

Compared with q-ary LRCs, binary LRCs are known to be advantageous in terms of implementation complexity in practical systems. For optimal binary linear LRCs, various construction methods have been proposed, especially for the case of minimum distances 3 and 4. Nevertheless, constructing optimal binary LRCs becomes increasingly challenging as the minimum distance requirement grows. In 2017, by the partial spread, Nam et al. [36] constructed a class of binary linear LRCs with minimum distance at least 6 and showed some examples that are optimal with respect to the bound (2). Subsequently, Wang et al. [11] constructed an $[n = \frac{2^s-1}{2^t-1}, k = \frac{rn}{r+1} - s, d \geqslant 6]$ binary k-optimal LRC with locality $r = 2^b$ from generalized Hamming codes, where s and t are integers that satisfy $2t|s$ and $\frac{s}{2t} \geqslant 2$. Ma et al. [37] proposed a class of k-optimal binary linear LRCs for $d = 6$, which included the codes given in [11], and they also presented a new k-optimal construction for locality 3 and minimum distance 6 from a partial t-spread. For any fixed locality r and minimum distance d, the coding rate of optimal LRCs becomes larger as the code length becomes larger [3]. Note that most constructions of binary linear LRCs are based on the partial t-spread, namely a set of mutually disjoint t-dimensional subspaces. Owing to the mutually disjoint subspaces, it is easy to calculate the minimum distance of binary linear LRCs. However, a disadvantage of this approach is that the code length is limited. Actually, for a given locality, based on the intersection subspace, k-optimal binary linear LRCs can be constructed with code length larger than previously known, but few constructions exist. A more ingenious approach is necessary to cope with the intersection subspace in order to guarantee the minimum distance. Hence, it is a challenging and interesting problem to construct k-optimal binary linear LRCs by applying intersection subspaces.

1.2 Our results

This paper focuses on a single-node failure problem of LRCs. We consider binary linear LRCs of which all local repair groups have uniform size $r+1$ and are pairwise disjoint, i.e., $(r+1)|n$. Using parity check matrices, we present an explicit construction of binary linear LRCs based on intersection subspaces with minimum distance $d \geqslant 6$ and locality $r = 2^b$. These intersection subspaces are designed by the direct sum of subspaces. Our LRCs turned out to be k-optimal in terms of the bound (4). Precisely speaking, the following results are obtained.

An explicit construction of k-optimal binary linear LRCs with new parameters $[n = (r+1)\ell, k \geqslant n - s - \ell - m, d \geqslant 6]$ is proposed (see Construction 1 and Theorem 1 below), where $\ell = \frac{2^m-1}{2^{2b-s}-1}$, $b \geqslant 3$, $0 \leqslant s < b$, and $(2b-s)|m$. When ℓ belongs to a determined range, those binary linear LRCs all can attain the bound (4), so they are k-optimal (Theorem 2). In the case of $(2b-s) \nmid m$, we construct k-optimal binary linear LRCs with parameters $[n = (r+1)\ell, k = n-\ell-s-m, d \geqslant 6]$ where $\ell = \frac{2^{m-s}-2^{(2b-s)}(2^z-1)-1}{2^{(2b-s)}-1}$, $0 \leqslant s < b$, and $z \equiv (m-s) \bmod (2b-s) < b$ (Theorem 3). Similar to Theorem 2, a class of k-optimal LRCs with a wider code length can be obtained from Theorem 3 (see details in Theorem 4). All results of k-optimal binary linear LRCs in this paper are summarized in Table 1.

Moreover, we compare our results with the state-of-the-art approaches for a fixed locality r. The results show that the k-optimal LRCs in this study have more flexible parameters $[n, k]$ than those in [11, 37]. In other words, our construction can generate more repair groups, so with the same locality, the code length of k-optimal LRCs is larger. Additionally, by calculating code rates, it can be obtained that the code rate $R \triangleq k/n$ in our construction is higher than that in [11]. At the end of this paper, a shortening technique will yield the derivation of new binary linear LRCs. By deleting codewords in k-optimal binary linear LRCs with nonzero values in the last coordinates and then removing the last coordinates from the remaining codewords, we can suggest new parameters from the original binary linear LRCs (Theorem 5).

Table 1 k-optimal binary linear LRCs with $d \geqslant 6$ and $r = 2^b$

	n, k	ℓ	$b \geqslant 3, m \geqslant 4b, s$
Theorem 1	$n = (r+1)\ell$, $k = n - s - \ell - m$	$\ell = \frac{2^m-1}{2^{2b-s}-1}$	$0 \leqslant s < b$, $(2b-s) \mid m$
Theorem 2	$n = (r+1)\ell$, $k = n - s - \ell - m$	$\frac{2^{m+s-1}-1}{2^{b-1}(2^b+1)} < \ell \leqslant \frac{2^m-1}{2^{2b-s}-1}$	$0 \leqslant s < b$, $(2b-s) \mid m$
Theorem 3	$n = (r+1)\ell$, $k = n - \ell - m - s$	$\ell = \frac{2^m - 2^{(2b-s)}(2^z-1)-1}{2^{(2b-s)}-1}$	$0 \leqslant s < b$, $(2b-s) \nmid m$, $z \equiv m \bmod (2b-s) \leqslant b$
Theorem 4	$n = (r+1)\ell$, $k = n - \ell - m - s$	$\frac{2^{m+s-1}-1}{2^{b-1}(2^b+1)} < \ell \leqslant \frac{2^m - 2^{(2b-s)}(2^z-1)-1}{2^{(2b-s)}-1}$	$0 \leqslant s < b$, $(2b-s) \nmid m$, $z \equiv m \bmod (2b-s) \leqslant b$

1.3 Organization

In Section 2, some basic definitions and results on LRCs, partial t-spread, and intersection subspace are introduced. In Section 3, we present a definition of a desired matrix and an explicit construction of LRCs. Based on this construction, we obtain the main results in Theorems 1–4. We also give three examples to explain the corresponding construction and some tables to show the comparison. In Section 4, Theorem 5 proposes the result to shorten binary linear LRCs. Finally, Section 5 concludes the paper. In addition, the constructions of some desired matrices are displayed in Appendix A.

2 Preliminaries

In this section, we introduce some notations and basic results required later in this paper.

- Let $V_n(q)$ be the vector space with dimension n over $\mathbb{F}_q$. When $q = 2$, we omit q from the notation $V_n(q)$.
- Supposing that n is a positive integer, we write $[n] = \{1, \ldots, n\}$.
- For any $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ and $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ in $V_n(q)$, the Euclidean inner product of $\boldsymbol{x}$ and $\boldsymbol{y}$ is defined as $\boldsymbol{x} \cdot \boldsymbol{y} = \sum_{i=1}^{n} x_i y_i$.
- The support set of $\boldsymbol{x}$ is denoted by $\mathrm{supp}(\boldsymbol{x}) = \{i \mid x_i \neq 0\}$.

2.1 Locally repairable codes

Let $\mathcal{C}$ be an $[n, k, d]_q$ linear code. Then, $\mathcal{C}$ has a $k \times n$ generator matrix G and an $(n-k) \times n$ parity check matrix H. The dual of $\mathcal{C}$ is defined as

$$\mathcal{C}^{\perp} = \{\boldsymbol{w} \in V_n(q) : \boldsymbol{w} \cdot \boldsymbol{c} = 0 \text{ for all } \boldsymbol{c} \in \mathcal{C}\}.$$

The rows of H are codewords of $\mathcal{C}^{\perp}$. Hence, the $k \times n$ generator matrix G and $(n-k) \times n$ parity check matrix H satisfy $GH^{\mathrm{T}} = \mathbf{0}$, where T denotes the transpose of a matrix. There is a well-known distance property of linear codes as follows.

Lemma 1 ([38], Theorem 4.5.6)**.** Let $\mathcal{C}$ be a linear code and let H be a parity check matrix for $\mathcal{C}$. Then the minimum distance of $\mathcal{C}$ is not less than d if and only if any $d-1$ columns of H are linearly independent.

Now, we give the formal definition of linear LRCs.

Definition 1. The linear code $\mathcal{C}$ is a locally repairable code (LRC) with locality r if for any $i \in [n]$, there exists a subset $\mathcal{R}_i \subset [n] \backslash \{i\}$ with $|\mathcal{R}_i| \leqslant r$ such that the ith symbol c_i in each codeword $\boldsymbol{c} = (c_1, c_2, \ldots, c_n) \in \mathcal{C}$ can be recovered by $\{c_j\}_{j \in \mathcal{R}_i}$; i.e., c_i is a linear combination of $\{c_j\}_{j \in \mathcal{R}_i}$. The set $\mathcal{R}_i$ is called a repair set for c_i.

It is well known that two different approaches are used to construct LRC, the generator matrix approach [4], and the parity check matrix approach [14]. Next, we will introduce the parity check matrix approach to construct LRCs. In order to find a suitable parity check matrix involving locality, we begin with a simple lemma.

Lemma 2 ([10])**.** An LRC has locality r if and only if for every coded symbol there exists a codeword $\boldsymbol{x}$ in $\mathcal{C}^{\perp}$ whose support set $\mathrm{supp}(\boldsymbol{x})$ contains i and the size of $\mathrm{supp}(\boldsymbol{x})$ is at most $r+1$.

An LRC is said to have ℓ disjoint local repair groups if there exist $\ell \triangleq \frac{n}{r+1}$ vectors $\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_\ell$ of $\mathcal{C}^\perp$, such that $|\operatorname{supp}(\boldsymbol{h}_i)| = r+1$ and $\operatorname{supp}(\boldsymbol{h}_i) \cap \operatorname{supp}(\boldsymbol{h}_j) = \emptyset$ for any $1 \leqslant i \neq j \leqslant \ell$. Let $\mathcal{C}$ be an $[n,k]$ binary linear LRC with disjoint local repair groups. The parity check matrix H of $\mathcal{C}$ can be represented as follows:

$$H = \begin{pmatrix} H_L \\ H_G \end{pmatrix} = \begin{pmatrix} H_1 & H_2 & \cdots & H_\ell \\ H_G^1 & H_G^2 & \cdots & H_G^\ell \end{pmatrix}, \tag{5}$$

where H_i is an $\ell \times (r+1)$ submatrix of whose ith row is the all-one vector and the other rows are all-zero vectors and H_G^i is an ith $(n-k-\ell) \times (r+1)$ submatrix of H_G for $1 \leqslant i \leqslant \ell$. Note that, in the following sections, the column vector $\boldsymbol{h}_j^i$ denotes a column of H_G^i, which is different from the meaning of $\boldsymbol{h}_j$.

2.2 Intersection subspace

The set of all t-dimensional subspaces of $V_m(q)$ is denoted by $\mathcal{G}_q(m,t)$. Let U and V be subspaces of $V_m(q)$. Then $U \bigcap V = \{\boldsymbol{v} | \boldsymbol{v} \in U \text{ and } \boldsymbol{v} \in V\}$ is called the intersection of U and V. It is clear that the intersection $U \bigcap V$ is also a subspace.

In particular, two t-dimensional subspaces U and V which belong to $\mathcal{G}_q(m,t)$ are said to trivially intersect or disjoint if they only have a zero-dimensional intersection, i.e., $U \bigcap V = \{\mathbf{0}\}$. A partial t-spread of $V_m(q)$ is a collection $S = \{W_1, W_2, \ldots, W_\ell\}$ of t-dimensional subspaces from $\mathcal{G}_q(m,t)$ such that $W_i \bigcap W_j = \{\mathbf{0}\}$ for $1 \leqslant i < j \leqslant \ell$, where ℓ is the size of the partial t-spread S. If t divides m and $\operatorname{span}(\bigcup_{i=1}^{\ell} W_i) = V_m(q)$, the partial spread is called a t-spread. Let $\mu_q(m,t)$ denote the number of t-dimensional subspaces in the largest partial spread in $V_m(q)$. One challenging question is to find the maximum partial size of a t-spread. There are a few results related to $\mu_q(m,t)$, see below.

Lemma 3 ([39]). If t is a divisor of m and $\ell = \frac{q^m-1}{q^t-1}$, then there exists a t-spread of $V_m(q)$ with ℓ subspaces.

Lemma 4 ([40]). Let $m \equiv z \bmod t$. Then, for all q, we have

$$\mu_q(m,t) \geqslant \frac{q^m - q^t(q^z-1) - 1}{q^t - 1}. \tag{6}$$

Note that Lemma 3 is a special case of Lemma 4 if $z = 0$. In addition, a specific construction for a t-spread is given in [39] and a specific construction for a partial t-spread is given in [40].

3 Construction by the intersection subspace

In this section, we begin with a lemma that is essential to construct the parity check matrix of binary linear LRCs with minimum distance $d \geqslant 6$. Then a definition of a desired matrix is given for subsequent constructions of LRCs. Finally, combining the intersection subspaces with a desired matrix generates k-optimal binary linear LRCs with disjoint local repair groups. The parameters of k-optimal LRCs are derived in Theorems 1 and 3, respectively.

Lemma 5 ([36]). Consider a binary linear LRC defined by the parity check matrix H in (5). If the columns of H_G satisfy the following three conditions, then the LRC has minimum distance $d \geqslant 6$.

(1) No two column vectors from matrix H_G^i sum to zero for all $i \in [\ell]$;

(2) No four column vectors from matrix H_G^i sum to zero for all $i \in [\ell]$;

(3) No four column vectors consisting of two columns from matrix H_G^i and the other two columns from matrix H_G^j sum to zero for all distinct $i \neq j \in [\ell]$.

Hence, to obtain a binary linear LRC with minimum distance $d \geqslant 6$, we need to construct the parity check matrix H in (5) of which submatrix H_G^i satisfies the conditions of Lemma 5.

Suppose $W_1, W_2, \ldots, W_\ell$ are t-dimensional subspaces of a vector space V_m such that $W_i \bigcap W_j = \{\mathbf{0}\}$ for $i \neq j \in [\ell]$. Let $\{\boldsymbol{e}_1^i, \boldsymbol{e}_2^i, \ldots, \boldsymbol{e}_t^i\}$ be a basis of the subspace W_i, where $\boldsymbol{e}_j^i = (e_{1j}^i, e_{2j}^i, \ldots, e_{mj}^i)^{\mathrm{T}} \in V_m$ for $j \in [t]$. We shall write the coordinates of the vector $\boldsymbol{e}_j^i$ as the jth column of an $m \times t$ matrix G_{W_i}, i.e., $G_{W_i} = [\boldsymbol{e}_1^i, \boldsymbol{e}_2^i, \ldots, \boldsymbol{e}_t^i]$.

Definition 2. Let $G_U = [\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_s]$ be an $s \times s$ matrix over $\mathbb{F}_2$ with full column rank and column vectors $\boldsymbol{u}_i = (u_{1i}, u_{2i}, \ldots, u_{si})^{\mathrm{T}} \in V_s$ for $i \in [s]$. Let U be the span of the columns of $\begin{pmatrix} G_U \\ \mathbf{0}_{m \times s} \end{pmatrix}$, where

$\mathbf{0}_{m\times s}\in V_m$ is an $m\times s$ all-zero matrix. Define an $(s+m)\times(s+t)$ matrix G_{M_i} over $\mathbb{F}_2$ as follows:

$$G_{M_i}=\begin{pmatrix} G_U & \mathbf{0}_{s\times t}\\ \mathbf{0}_{m\times s} & G_{W_i}\end{pmatrix}=\begin{pmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_s & \mathbf{0}_{s\times 1} & \mathbf{0}_{s\times 1} & \cdots & \mathbf{0}_{s\times 1}\\ \mathbf{0}_{m\times 1} & \mathbf{0}_{m\times 1} & \cdots & \mathbf{0}_{m\times 1} & \boldsymbol{e}_1^i & \boldsymbol{e}_2^i & \ldots & \boldsymbol{e}_t^i\end{pmatrix}. \tag{7}$$

In the special case that $s=0$, we have $G_{M_i}=G_{W_i}$. Furthermore, let M_i denote the vector space spanned by the columns of G_{M_i}.

Remark 1. In this paper, we utilize the intersection subspaces M_i for $i\in[\ell]$ to construct the parity check matrix of a binary linear LRC. Furthermore, combining a t-spread with a matrix G_U, we give an explicit form of such intersection subspaces. From the above construction, each subspace M_i can be generated by a direct sum of two subspaces U and W_i for $i\in[\ell]$, i.e., $M_i=W_i\oplus U$. It is not hard to see that M_i is a vector subspace of V_{m+s} with dimension $s+t$. Hence, it is clear that the intersection of $M_1,M_2,\ldots,M_\ell$ only consists of the vector subspace U with dimension s.

Later, we will exploit these intersection subspaces and a matrix with a special structure to construct binary linear LRCs. In this paper, we call such a matrix a desired matrix and give the definition of a desired matrix as follows.

Definition 3. Let r,s, and t be integers such that $2^t>r>s+t$ and $t>s\geqslant 0$. A binary $(s+t)\times r$ matrix $A=\left(\begin{smallmatrix}A_1\\A_2\end{smallmatrix}\right)=(a_{i,j})\in V_{(s+t)\times r}$ is a desired matrix if it is a binary matrix with full column rank such that the submatrix A_2 has t rows and no two linearly dependent columns; furthermore, if $t\geqslant 3$, then any 4 columns of A must be linearly independent. Note that A can be viewed as the parity check matrix of an $[r,r-(s+t),\geqslant 5]$ binary linear code for $t\geqslant 3$ (see more details in Appendix A).

Remark 2. The key idea of intersection subspaces comes from the construction of matrix G_{M_i}, while matrix A will help to expand the range of the construction parameters and control the minimum distance of binary linear LRCs.

Let G_{M_i} be a matrix as defined in (7) for $i\in[\ell]$. Let A be an $(s+t)\times r$ desired matrix. Then the product of each matrix G_{M_i} and A is

$$G_{M_i}\cdot A=\begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1s} & & & & \\ u_{21} & u_{22} & \cdots & u_{2s} & & & & \\ \vdots & \vdots & & \vdots & & \mathbf{0}_{s\times t} & & \\ u_{s1} & u_{s2} & \cdots & u_{ss} & & & & \\ & & & & e_{11}^i & e_{12}^i & \cdots & e_{1t}^i\\ & & & & e_{21}^i & e_{22}^i & \cdots & e_{2t}^i\\ & \mathbf{0}_{m\times s} & & & \vdots & \vdots & & \vdots\\ & & & & e_{m1}^i & e_{m2}^i & \ldots & e_{mt}^i\end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1r}\\ a_{21} & a_{22} & \cdots & a_{2r}\\ \vdots & \vdots & & \vdots\\ a_{(s+t)1} & a_{(s+t)2} & \cdots & a_{(s+t)r}\end{pmatrix},$$

where $\mathbf{0}$ denotes a zero matrix. Let H_G^i denote an $(s+m)\times(r+1)$ matrix over $\mathbb{F}_2$ which is equal to the matrix $(\mathbf{0}_{(s+m)\times 1},\ G_{M_i}\cdot A)$. Note that every column $\boldsymbol{h}_j^i$ of H_G^i can be indexed by a pair (i,j) for $1\leqslant i\leqslant\ell$ and $0\leqslant j\leqslant r$, where $\boldsymbol{h}_0^i=(0,0,\ldots,0)^{\mathrm{T}}$ and

$$\boldsymbol{h}_{j\neq 0}^i=\left(\sum_{\ell=1}^{s}u_{1\ell}a_{\ell j},\ldots,\sum_{\ell=1}^{s}u_{s\ell}a_{\ell j},\sum_{\ell=1}^{t}e_{1\ell}^i a_{(s+\ell)j},\ldots,\sum_{\ell=1}^{t}e_{m\ell}^i a_{(s+\ell)j}\right)^{\mathrm{T}}.$$

Then, we propose the following Lemma 6 with regard to H_G^i, which plays an important role for the main results of Theorem 1.

Lemma 6. Let $H_G^i=\left(\mathbf{0}_{(s+m)\times 1},\ G_{M_i}\cdot A\right)$ for $i\in[\ell]$. The sum of any four columns of H_G, which consists of two columns from matrix $H_G^{i_1}$ and the other two columns from matrix $H_G^{i_2}$, is not equal to zero for all distinct $i_1,i_2\in[\ell]$.

Proof. Considering the four columns $\boldsymbol{h}_{j_1}^{i_1},\boldsymbol{h}_{j_2}^{i_1}\in H_G^{i_1}$, $\boldsymbol{h}_{j_3}^{i_2},\boldsymbol{h}_{j_4}^{i_2}\in H_G^{i_2}$, without loss of generality, we assume that $\boldsymbol{h}_{j_1}^{i_1}+\boldsymbol{h}_{j_2}^{i_1}+\boldsymbol{h}_{j_3}^{i_2}+\boldsymbol{h}_{j_4}^{i_2}=\mathbf{0}$.

Case (i): The four columns $\{\boldsymbol{h}_{j_1}^{i_1},\boldsymbol{h}_{j_2}^{i_1},\boldsymbol{h}_{j_3}^{i_2},\boldsymbol{h}_{j_4}^{i_2}\}$ contain the zero column vector.

(1) Assume that one of the columns is the zero column from $H_G^{i_1}$, i.e., $\boldsymbol{h}_{j_1}^{i_1} = \mathbf{0}$. Then we have

$$\begin{cases} \sum_{\ell=1}^{s} u_{1\ell}a_{\ell j_2} + \sum_{\ell=1}^{s} u_{1\ell}a_{\ell j_3} + \sum_{\ell=1}^{s} u_{1\ell}a_{\ell j_4} = 0, \\ \vdots \\ \sum_{\ell=1}^{s} u_{s\ell}a_{\ell j_2} + \sum_{\ell=1}^{s} u_{s\ell}a_{\ell j_3} + \sum_{\ell=1}^{s} u_{s\ell}a_{\ell j_4} = 0, \\ \sum_{\ell=1}^{t} e_{1\ell}^{i_1}a_{(s+\ell)j_2} + \sum_{\ell=1}^{t} e_{1\ell}^{i_2}a_{(s+\ell)j_3} + \sum_{\ell=1}^{t} e_{1\ell}^{i_2}a_{(s+\ell)j_4} = 0, \\ \vdots \\ \sum_{\ell=1}^{t} e_{m\ell}^{i_1}a_{(s+\ell)j_2} + \sum_{\ell=1}^{t} e_{m\ell}^{i_2}a_{(s+\ell)j_3} + \sum_{\ell=1}^{t} e_{m\ell}^{i_2}a_{(s+\ell)j_4} = 0. \end{cases} \tag{8}$$

From (8), the last t equations out of m equations show that

$$\sum_{\ell=1}^{t} a_{(s+\ell)j_2} \begin{pmatrix} e_{1\ell}^{i_1} \\ e_{2\ell}^{i_1} \\ \vdots \\ e_{m\ell}^{i_1} \end{pmatrix} + \sum_{\ell=1}^{t} (a_{(s+\ell)j_3} + a_{(s+\ell)j_4}) \begin{pmatrix} e_{1\ell}^{i_2} \\ e_{2\ell}^{i_2} \\ \vdots \\ e_{m\ell}^{i_2} \end{pmatrix} = \mathbf{0}.$$

Since W_{i_1} and W_{i_2} are disjoint t-dimensional subspaces, i.e., $W_{i_1} \bigcap W_{i_2} = \{\mathbf{0}\}$, which means that the linear combinations of the basis $\{\boldsymbol{e}_1^{i_1}, \boldsymbol{e}_2^{i_1}, \ldots, \boldsymbol{e}_t^{i_1}\}$ of W_{i_1} and $\{\boldsymbol{e}_1^{i_2}, \boldsymbol{e}_2^{i_2}, \ldots, \boldsymbol{e}_t^{i_2}\}$ of W_{i_2} are linearly independent. Thus, $\sum_{\ell=1}^{t} a_{(s+\ell)j_2}\boldsymbol{e}_\ell^{i_1} = \mathbf{0}$ and $\sum_{\ell=1}^{t}(a_{(s+\ell)j_3}+a_{(s+\ell)j_4})\boldsymbol{e}_\ell^{i_2} = \mathbf{0}$, implying that $a_{(s+\ell)j_2} = 0$ and $a_{(s+\ell)j_3} + a_{(s+\ell)j_4} = 0$ for all $\ell \in [t]$. A contradiction can be obtained from the definition of the desired matrix A, so $\boldsymbol{h}_{j_1}^{i_1} + \boldsymbol{h}_{j_2}^{i_1} + \boldsymbol{h}_{j_3}^{i_2} + \boldsymbol{h}_{j_4}^{i_2} \neq \mathbf{0}$.

(2) In the case that $\boldsymbol{h}_{j_1}^{i_1}$ from $H_G^{i_1}$ and $\boldsymbol{h}_{j_3}^{i_2}$ from $H_G^{i_2}$ are zero vector, respectively, by the similar analysis of case (1), the same result holds.

Case (ii): The four columns $\{\boldsymbol{h}_{j_1}^{i_1}, \boldsymbol{h}_{j_2}^{i_1}, \boldsymbol{h}_{j_3}^{i_2}, \boldsymbol{h}_{j_4}^{i_2}\}$ do not contain the zero column vector. Similar to case (i), we have

$$\sum_{\ell=1}^{s} (a_{\ell j_1} + a_{\ell j_2} + a_{\ell j_3} + a_{\ell j_4}) \begin{pmatrix} u_{1\ell} \\ u_{2\ell} \\ \vdots \\ u_{s\ell} \end{pmatrix} = \mathbf{0}, \tag{9}$$

and

$$\sum_{\ell=1}^{t} (a_{(s+\ell)j_1} + a_{(s+\ell)j_2}) \begin{pmatrix} e_{1\ell}^{i_1} \\ e_{2\ell}^{i_1} \\ \vdots \\ e_{m\ell}^{i_1} \end{pmatrix} + \sum_{\ell=1}^{t} (a_{(s+\ell)j_3} + a_{(s+\ell)j_4}) \begin{pmatrix} e_{1\ell}^{i_2} \\ e_{2\ell}^{i_2} \\ \vdots \\ e_{m\ell}^{i_2} \end{pmatrix} = \mathbf{0}.$$

Since the matrix $[\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_s]$ is an $s \times s$ matrix with full column rank, this formula (9) can be simplified as $a_{\ell j_1} + a_{\ell j_2} + a_{\ell j_3} + a_{\ell j_4} = 0$ for all $\ell \in [s]$. On the other hand, by the same argument as in the proof of case (i), $\sum_{\ell=1}^{t}(a_{(s+\ell)j_1} + a_{(s+\ell)j_2})\boldsymbol{e}_\ell^{i_1} = \mathbf{0}$ and $\sum_{\ell=1}^{t}(a_{(s+\ell)j_3} + a_{(s+\ell)j_4})\boldsymbol{e}_\ell^{i_2} = \mathbf{0}$ can be obtained, which forces $a_{(s+\ell)j_1} + a_{(s+\ell)j_2} = 0$ and $a_{(s+\ell)j_3} + a_{(s+\ell)j_4} = 0$ for all $\ell \in [t]$. Hence, it is clear that $\sum_{z=1}^{4} a_{\ell j_z} = 0$ for all $\ell \in [s+t]$. By definition, any four columns of the desired matrix A are linearly independent over $\mathbb{F}_2$. This contradiction completes the proof of Lemma 6.

With the above preparation, we give the following construction of binary linear LRCs with minimum distance $d \geqslant 6$.

Construction 1. Let b, s, t, and m be integers such that $s + t = 2b$ and $m \geqslant 4b$. Choose a desired matrix A with size $2b \times 2^b$ and a $(2b-s)$-spread $\{W_1, W_2, \ldots, W_\ell\}$ of V_m with size $\ell = \frac{2^m-1}{2^{2b-s}-1}$, where $b > s \geqslant 0$ and $(2b-s)|m$. The submatrix H_G^i is given by $H_G^i = (\mathbf{0}_{(s+m)\times 1},\ G_{M_i} \cdot A)$ for each $i \in [\ell]$. Then the linear code $\mathcal{C}$ is constructed by parity check matrix H given in (5) with submatrices H_G^i.

Note that a $(2b-s)$-spread $\{W_1, W_2, \ldots, W_\ell\}$ of V_m exists if and only if $(2b-s)$ divides m. By Lemma 3, it is known that the size of a $(2b-s)$-spread is $\ell = \frac{2^m-1}{2^{2b-s}-1}$. Additionally, for the existence of

the desired matrix A, it is required that $b > s \geqslant 0$. The submatrix A_2 of A can be viewed as a parity check matrix with parameters $[2^b, 2^b - 2b - s, 3]$. According to the Griesmer bound, the parameters of this code should satisfy $2^b \geqslant \sum_{i=0}^{2^b-2b-s} \lceil \frac{3}{2^i} \rceil$, which follows from $0 \leqslant s < b$.

Henceforth, we will consider a binary linear code obtained from Construction 1. We have the following theorem.

Theorem 1. Let b be an integer such that $b \geqslant 3$. The code $\mathcal{C}$ constructed by the parity check matrix H from Construction 1 is an $[n = (r+1)\ell, k = n - s - \ell - m, d \geqslant 6]$ binary linear LRC with locality $r = 2^b$, which is k-optimal and attains the bound (4).

Proof. This proof consists of two parts. In Part 1 we will prove that the dimension k of the corresponding codes achieves the bound (4), i.e., $k = n - s - \ell - m$. As to Part 2, we will show that the minimum distance $d \geqslant 6$.

Part 1: It is easy to determine the parameter $n = (r+1)\ell$, $k \geqslant \frac{rn}{(r+1)} - s - m$, and $r = 2^b$ by the parity check matrix H in Construction 1, where $\ell = \frac{2^m-1}{2^{2b-s}-1}$. Clearly,

$$\min\left\{\log_2\left(1+\frac{rn}{2}\right), \frac{rn}{(r+1)(r+2)}\right\} = \log_2\left(1+\frac{rn}{2}\right)$$

can be obtained when $m \geqslant 4b$ and $\ell = \frac{2^m-1}{2^{2b-s}-1}$. By the bound (4),

$$k \leqslant \frac{rn}{r+1} - \left\lceil \log_2\left(1+\frac{rn}{2}\right)\right\rceil = \frac{rn}{r+1} - \left\lceil \log_2(1+2^{b-1}(2^b+1)\ell)\right\rceil = \frac{rn}{r+1} - m - s$$

as a result of $2^{m+s-1} < 1 + 2^{b-1}(2^b+1)\ell \leqslant 2^{m+s}$. Hence, we obtain $k = n - \ell - s - m$.

Part 2: That code $\mathcal{C}$ has minimum distance $d \geqslant 6$ is equivalent to showing that the submatrix H_G^i of H satisfies the conditions in Lemma 5. Notably, the sum of the first ℓ rows of H is an all-one vector. Thus the minimum distance of $\mathcal{C}$ must be even.

Case (i): For any two columns $\boldsymbol{h}_{j_1}^i$ and $\boldsymbol{h}_{j_2}^i$ vectors from H_G^i, it is obvious that $\boldsymbol{h}_{j_1}^i + \boldsymbol{h}_{j_2}^i \neq \boldsymbol{0}$. Therefore, H_G^i satisfies condition (1) in Lemma 5.

Case (ii): Consider the four columns $\{\boldsymbol{h}_{j_z}^{i_z}\}_{z=1}^4$ from H belong to the same block, i.e., $i_1 = i_2 = i_3 = i_4$. Without loss of generality, we assume in the contradiction method that $\boldsymbol{h}_{j_1}^{i_1} + \boldsymbol{h}_{j_2}^{i_1} + \boldsymbol{h}_{j_3}^{i_1} + \boldsymbol{h}_{j_4}^{i_1} = \boldsymbol{0}$. Similar to the proof of Lemma 6, $\sum_{\ell=1}^{s}(\sum_{z=1}^{4} a_{\ell j_z})\boldsymbol{u}_\ell = \boldsymbol{0}$ and $\sum_{\ell=1}^{t}(\sum_{z=1}^{4} a_{(s+\ell)j_z})\boldsymbol{e}_\ell^{i_1} = \boldsymbol{0}$. Since $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_s, \boldsymbol{e}_{s+1}^{i_1}, \ldots,$ and $\boldsymbol{e}_{s+t}^{i_1}$ are linearly independent, $\sum_{z=1}^{4} a_{\ell j_z} = 0$ for all $\ell \in [s+t]$, which implies that $\boldsymbol{a}_{j_1}, \boldsymbol{a}_{j_2}, \boldsymbol{a}_{j_3}$, and $\boldsymbol{a}_{j_4}$ from A are linearly dependent. This result contradicts the definition of A. Thus for any four columns $\{\boldsymbol{h}_{j_z}^{i_z}\}_{z=1}^4$, $\boldsymbol{h}_{j_1}^{i_1} + \boldsymbol{h}_{j_2}^{i_1} + \boldsymbol{h}_{j_3}^{i_1} + \boldsymbol{h}_{j_4}^{i_1} \neq \boldsymbol{0}$. In particular, if $\boldsymbol{0} \in \{\boldsymbol{h}_{j_z}^{i_z}\}_{z=1}^4$, $\sum_{z=2}^{4} a_{\ell j_z} = 0$ also hold for all $\ell \in [s+t]$. By the definition of A, $\sum_{z=1}^{4} \boldsymbol{h}_{j_z}^{i_z} \neq \boldsymbol{0}$. Hence, H_G^i satisfies condition (2) in Lemma 5.

Case (iii): Two of $\{\boldsymbol{h}_{jz}^{iz}\}_{z=1}^4$ belong to one block and the other two lie in a different block. Then their sum is not equal to zero by Lemma 6, proving that H_G^i satisfies the condition (3) in Lemma 5.

As a consequence, the lower part H_G of H satisfies three conditions in Lemma 5. This completes the proof of Theorem 1.

Next, we give two examples to illustrate the corresponding construction in detail. Example 1, by Theorem 1, shows how to construct the k-optimal binary linear LRC from Construction 1; Example 2 is a special case when $s = 0$.

Example 1. Let $b = 3$, $s = 2$, and $m = 12$ in Construction 1. Let $\{W_1, W_2, \ldots, W_{273}\}$ be a 4-spread of V_{12}. Let $\{\boldsymbol{e}_1^i, \boldsymbol{e}_2^i, \boldsymbol{e}_3^i, \boldsymbol{e}_4^i\}$ for $i \in [273]$ denote a basis of W_i. Then we choose a matrix G_{M_i} and a desired matrix $A_{6\times 8}$ as follows:

$$G_{M_i} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{e}_1^i & \boldsymbol{e}_2^i & \boldsymbol{e}_3^i & \boldsymbol{e}_4^i \end{pmatrix}, \quad A_{6\times 8} = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix},$$

where

$$A_1 = \begin{pmatrix} 1&0&0&0&0&0&1&0 \\ 0&1&0&0&0&0&0&1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1&1&1&0&0&0&0&0 \\ 1&0&0&1&0&0&1&1 \\ 0&1&0&0&1&0&1&0 \\ 0&1&0&0&0&1&1&1 \end{pmatrix}.$$

For example, let α be a primitive element of $\mathbb{F}_{2^{12}}$ such that $\alpha^{12}+\alpha^7+\alpha^6+\alpha^5+\alpha^3+\alpha+1=0$. Let $\ell = \frac{2^{12}-1}{2^4-1}$ and $\gamma = \alpha^\ell$. We get a basis $\{\alpha^0, \alpha^0\gamma, \alpha^0\gamma^2, \alpha^0\gamma^3\}$ and a basis $\{\alpha^1, \alpha^1\gamma, \alpha^1\gamma^2, \alpha^1\gamma^3\}$ of subspaces W_1 and W_2, respectively. Then we have

$$G_{M_i}A = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ \boldsymbol{e}_1^i+\boldsymbol{e}_2^i & \boldsymbol{e}_1^i+\boldsymbol{e}_3^i+\boldsymbol{e}_4^i & \boldsymbol{e}_1^i & \boldsymbol{e}_2^i & \boldsymbol{e}_3^i & \boldsymbol{e}_4^i & \boldsymbol{e}_2^i+\boldsymbol{e}_3^i+\boldsymbol{e}_4^i & \boldsymbol{e}_2^i+\boldsymbol{e}_4^i \end{pmatrix},$$

where $\boldsymbol{e}^i$ is an element in $\mathbb{F}_{2^{12}}$. Columns of H_G^i are binary expansions of the each column vector $(\mathbf{0}, G_{M_i} \cdot A)$. For example, fixing a basis $\{\alpha^0, \alpha^1, \ldots, \alpha^{11}\}$, from the submatrix H_G^1, $\boldsymbol{e}_1^1+\boldsymbol{e}_2^1 = \alpha^0+\gamma = \alpha^{10}+\alpha^9+\alpha^8+\alpha^4+\alpha^3+\alpha^2+1 = (1, \alpha, \ldots, \alpha^{11}) \cdot (1,0,1,1,1,0,0,0,1,1,1,0)^{\mathrm{T}}$. Thus the binary expansion of the vector of $\boldsymbol{e}_1^1+\boldsymbol{e}_2^1$ with respect to the basis is $(1,0,1,1,1,0,0,0,1,1,1,0)$. Then we obtain the matrices H_G^1 and H_G^2 as follows by expanding the column vectors of the submatrices $G_{M_i} \cdot A$ $(i=1,2)$ with respect to the bases, respectively:

$$H_G^1 = \begin{pmatrix} 0&1&1&0&0&0&0&1&1 \\ 0&0&1&0&0&0&0&0&1 \\ 0&1&1&1&0&1&1&0&1 \\ 0&0&1&0&0&1&0&1&0 \\ 0&1&0&0&1&1&1&1&0 \\ 0&1&1&0&1&1&0&0&1 \\ 0&1&0&0&1&1&1&1&0 \\ 0&0&0&0&0&1&1&0&1 \\ 0&0&1&0&0&0&1&1&1 \\ 0&0&0&0&0&0&0&0&0 \\ 0&1&1&0&1&0&1&0&0 \\ 0&1&1&0&1&1&0&0&1 \\ 0&1&0&0&1&0&0&1&1 \\ 0&0&1&0&0&1&0&1&0 \end{pmatrix}, \quad H_G^2 = \begin{pmatrix} 0&1&1&0&0&0&0&1&1 \\ 0&0&1&0&0&0&0&0&1 \\ 0&0&1&0&1&0&0&1&0 \\ 0&1&1&1&0&0&1&0&1 \\ 0&0&1&0&1&0&0&1&0 \\ 0&0&0&1&0&1&0&1&1 \\ 0&1&0&0&1&1&0&1&1 \\ 0&0&0&1&0&1&0&1&1 \\ 0&1&1&1&0&0&0&1&0 \\ 0&1&0&1&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0 \\ 0&0&0&1&0&1&0&1&1 \\ 0&1&0&0&1&1&0&1&1 \\ 0&1&1&0&0&1&0&0&1 \end{pmatrix}.$$

It can be verified that any 5 columns of H in (5) are linearly independent, so parity check matrix H determines a $[2457, 2170, 6]$ k-optimal binary linear LRC with locality $r=8$ by Theorem 1.

Example 2. Taking $t=2$ and $s=0$ in the above example. Let $\{W_0, W_1, W_2, W_3, W_4\}$ be a 2-spread of V_4 and $\{\boldsymbol{e}_1^i, \boldsymbol{e}_2^i\}$ be a basis of subspace W_i. By choosing

$$A = \begin{pmatrix} 1&0 \\ 0&1 \end{pmatrix} \quad \text{and} \quad G_{M_i} = \begin{pmatrix} \boldsymbol{e}_1^i & \boldsymbol{e}_2^i \end{pmatrix},$$

we obtain a $[15, 6, 6]$ k-optimal binary linear LRC with locality $r=2$. Similarly, when $s=0$, more k-optimal binary linear LRCs are listed in Table 2.

Remark 3. Note that an LRC with the same parameters as in Example 2 was also constructed in Example 2 of [11]. Correspondingly, taking $s=0$ in Construction 1 and Theorem 1, we obtain an $[n=\frac{2^m-1}{2^t-1}, k=\frac{rn}{r+1}-m, d \geqslant 6]$ binary linear LRC with locality $r=2^t$, which includes the construction of k-optimal binary linear LRCs in [11].

Table 2 k-optimal binary linear LRCs with $d = 6$

	$r=2$			$r=4$		$r=5$	$r=6$
k	6	36	162	21	210	60	155
n	15	63	255	36	292	85	195

In fact, Construction 1 generates a family of k-optimal binary linear LRCs with locality $r = 2^b$ when ℓ belongs to a determined range. This point is presented in detail in Theorem 2.

Theorem 2. Assume that $r = 2^b$ for $b \geqslant 3$ and let m, s, and b be integers such that $m \geqslant 4b$ and $0 \leqslant s < b$. If

$$\frac{2^{m+s-1}-1}{2^{b-1}(2^b+1)} < \ell \leqslant \frac{2^m-1}{2^{2b-s}-1},$$

there exists an $[n = (r+1)\ell, k = r\ell - m - s, d \geqslant 6]$ binary linear LRC with locality $r = 2^b$, which is k-optimal with respect to the bound (4).

Proof. By Theorem 1, $\mathcal{C}$ is an $[n = (r+1)\ell, k \geqslant \frac{rn}{r+1} - m - s, d \geqslant 6]$ binary linear LRC. Hence, we need to show that $k \leqslant r\ell - m - s$. Due to the condition that

$$\frac{2^{m+s-1}-1}{2^{b-1}(2^b+1)} < \ell \leqslant \frac{2^m-1}{2^{2b-s}-1},$$

we have

$$2^{m+s-1} < 1 + 2^{b-1}(2^b+1)\ell \leqslant \frac{2^{b-1}(2^b+1)(2^m-1)}{2^{2b-s}-1} + 1 \leqslant 2^{m+s}. \tag{10}$$

Furthermore, by the bound (4) and the formula (10),

$$\begin{aligned} k \leqslant \frac{rn}{r+1} - \left\lceil \log_2\left(1+\frac{rn}{2}\right)\right\rceil &= \frac{rn}{r+1} - \left\lceil \log_2\left(1+\frac{r(r+1)\ell}{2}\right)\right\rceil \\ &= \frac{rn}{r+1} - \lceil \log_2(1+2^{b-1}(2^b+1)\ell)\rceil \\ &= \frac{rn}{r+1} - m - s. \end{aligned}$$

As $k \geqslant \frac{rn}{r+1} - m - s$ in Theorem 1, $k = \frac{rn}{r+1} - m - s$, which is k-optimal with respect to the bound (4).

Notice that a necessary condition for the existence of the $(2b-s)$-spread is $(2b-s) \mid m$ in Construction 1. This condition restricts the parameters of LRC codes constructed using intersection subspace. For the case of $(2b-s) \nmid m$, we utilize the partial $(2b-s)$-spread of V_m to replace the $(2b-s)$-spread. Although the size of a maximum partial spread of V_m is not known when $(2b-s) \nmid m$, an explicit construction for a partial t-spread of size $\frac{q^m - q^{(2b-s)}(q^z-1)-1}{q^{(2b-s)}-1}$ is presented in [40], where $z \equiv m \bmod (2b-s)$. Hence, we obtain the following theorem.

Theorem 3. Let $m \geqslant 4b$. There exists a k-optimal binary linear LRC with parameters $[n = (r+1)\ell, k = n - \ell - s - m, d \geqslant 6]$ and locality $r = 2^b$ if there exists a partial $(2b-s)$-spread of V_m for $(2b-s) \nmid m$, where $\ell = \frac{2^m - 2^{(2b-s)}(2^z-1)-1}{2^{(2b-s)}-1}$, $0 \leqslant s < b$ and $z \equiv m \bmod (2b-s) \leqslant b$.

Proof. By the method analogous to that used in the proof of Theorem 1, an LRC code has parameters $n = (r+1)\ell$, $k \geqslant n - \ell - m - s$, $d \geqslant 6$. Hence, we only need to show that its dimension k satisfies the bound (4):

$$k \leqslant \frac{rn}{r+1} - \min\left\{\log_2\left(1+\frac{rn}{2}\right), \frac{rn}{(r+1)(r+2)}\right\},$$

which is equivalent that $k \leqslant n - \ell - s - m$, i.e.,

$$m + s - 1 < \min\left\{\log_2\left(1+\frac{rn}{2}\right), \frac{rn}{(r+1)(r+2)}\right\} \leqslant m + s. \tag{11}$$

As $n > 5(r+1)(r+2)$, inequality (11) can be written as

$$2^{m+s-1} \leqslant 1 + \frac{rn}{2} \leqslant 2^{m+s}. \tag{12}$$

Table 3 k-optimal binary linear LRCs with $d \geqslant 6$

	$[n,k;k/n]$ from Theorem 1	$[n,k;k/n]$ from Theorem 3	$[n,k;k/n]$ in [11]
$r=8$	**[2457, 2170; 0.8832]**	**[1161, 1019; 0.8777]** **[9801, 8696; 0.8873]**	[585, 508; 0.8684]
	[10066329, 8947822; 0.8889]	**[38025, 33782; 0.8884]** **[1258281, 1118449; 0.8889]**	[2396745, 2130416; 0.8889]
$r=16$	**[4527185, 4260854; 0.9412]**	**[4369, 4096; 0.9375]** **[1122833, 1056760; 0.9412]**	[1118481, 1052664; 0.9412]
	[602957989425, 567489872357; 0.9412]	**[1150033, 1082360; 0.9412]** **[287458321, 270548976; 0.9412]**	[73300775185, 68988964840; 0.9412]
$r=32$	**[562436193, 545392638; 0.9697]**	**[33825, 32780; 0.9691]** **[34670625, 33619970; 0.9697]**	[34636833, 33587202; 0.9697]
$r=64$	**[4276545, 4210724; 0.9846]**	**[68290625, 67239968; 0.9846]** **[1091051585, 1074266140; 0.9846]**	[266305, 262184; 0.9845]

For the left side of the inequality, it has to be verified that $(2^{m+s-1}-1)(2^{2b-s}-1) < 2^{b-1}(2^b+1)(2^m - 2^{2b-s}(2^z-1)-1)$. Since $m \geqslant 4b$, $0 \leqslant z \leqslant b$, and $0 \leqslant s < b$, $2^{m+b-1}+2^{4b-s-1} = 2^{m-4b}\cdot 2^{3b-1}+2^{4b-s-1} > 2^{4b+z-s-1}+2^{3b-1-s+z}$. Then $(2^{m+s-1}-1)(2^{2b-s}-1) < 2^{b-1}(2^b+1)(2^m - 2^{2b-s}(2^z-1)-1)$ follows from $2^{m+s-1} \geqslant 2^{2b-1}$ and $2^{2b-s} > 2^{b-1}$, which supports the left side of (12).

The right side of inequality (12) holds by similar arguments used to prove the left side of inequality (12). The proof has been completed.

Similar to the above analysis of Theorem 2, we obtain a family of k-optimal LRCs with locality $r = 2^b$ from Theorem 3 when ℓ lies within a specific range.

Theorem 4. Let r, b, and s be integers such that $r = 2^b$, $b \geqslant 3$, and $0 \leqslant s < b$. Suppose that $m \geqslant 4b$ is an integer and $1 \leqslant z \equiv m \bmod (2b-s) \leqslant b$. When

$$\frac{2^{m+s-1}-1}{2^{b-1}(2^b+1)} < \ell \leqslant \frac{2^m - 2^{(2b-s)}(2^z-1)-1}{2^{(2b-s)}-1}, \tag{13}$$

the code $\mathcal{C}$ in Theorem 3 is a k-optimal binary linear LRC with parameters $[n=(r+1)\ell, k = n-\ell-m-s, d \geqslant 6]$.

Proof. It is easy to construct the code $\mathcal{C}$ with parameters $n=(r+1)\ell, d \geqslant 6$ and $k \geqslant n-\ell-m-s$ by using the partial $(2b-s)$-spread in Construction 1, where $\ell = \frac{2^m - 2^{(2b-s)}(2^z-1)-1}{2^{(2b-s)}-1}$. In the similar way provided in Theorem 2, we prove $k \leqslant n-\ell-m-s$ below.

Combining the proof of Theorem 3 with (12) and (13), we derive the following chain of inequalities:

$$2^{m+s-1} < 1 + 2^{b-1}(2^b+1)\ell \leqslant \frac{2^{b-1}(2^b+1)(2^m - 2^{(2b-s)}(2^z-1)-1)}{2^{(2b-s)}-1} + 1 \leqslant 2^{m+s},$$

which implies $k \leqslant n-\ell-m-s$ by the bound (4). Therefore, $k = n-\ell-m-s$ proves the theorem.

An example of Theorems 3 and 4 is presented below.

Example 3. Let $m=12$, $b=3$, and $s=1$ and let $\{W_1, W_2, \ldots, W_{129}\}$ be a partial 5-spread of V_{12}. Then there exists an $[n=1161, k=1019, d \geqslant 6]$ binary LRC with locality $r=8$ by Theorem 3. This code is k-optimal since it attains the bound (4). Moreover, taking $113 \leqslant \ell \leqslant 129$, the code is a k-optimal binary linear LRC with parameters $[n=(r+1)\ell, k=n-\ell-m-s, d \geqslant 6]$ by Theorem 4.

Here we list parameters of k-optimal binary linear LRCs with disjoint local repair groups given by Theorems 1 and 3 in Table 3 for $3 \leqslant b \leqslant 6$ and $12 \leqslant m \leqslant 40$, which achieves the maximum value obtained from the bound (4). The values highlighted in bold in Table 3 are new parameters of k-optimal binary linear LRCs in the current paper. The parameters of LRCs with the same locality r in [11] are also listed in Table 3.

Remark 4. Wang et al. [11] constructed the parity check matrix of binary linear LRCs based on a 2^{2b}-ary Hamming code with length $\frac{2^m-1}{2^{2b}-1}$. Then they obtained an $[n' = \frac{2^m-1}{2^b-1}, k' = \frac{rn'}{r+1} - m, d \geqslant 6]$ binary linear LRC with disjoint local repair groups and locality $r=2^b$. Furthermore, their code rate is $\frac{k'}{n'} = \frac{r}{r+1} - \frac{m}{n'}$. Compared with the code rate of binary linear LRCs for $r=2^b$ in [11], our constructions have a larger code rate. In this paper, taking $0 < s < b$, the length of k-optimal LRC is $n = (2^b+1)\frac{2^m-1}{2^{2b-s}-1}$,

Table 4 $[n=(r+1)\ell, k, d\geqslant 6]$ k-optimal binary linear LRCs with respect to the bound (4)

Ref.	r	The number of repair groups	Conditions
[41]	2	$\ell=\frac{2^m-1}{3}$	$2\vert m, m\geqslant 6$
[42]	3	$\frac{2^m-1}{6}\leqslant \ell<\frac{2^m-1}{3}$	$m\geqslant 6$
[11]	2^b	$\ell=\frac{2^m-1}{2^{2b}-1}$	$2b\mid m, m\geqslant 4b$
[37]	2^b	$\lfloor\frac{2^{m-1}-1}{2^{b-1}(2^b+1)}\rfloor+1\leqslant \ell\leqslant \mu_2(m,2b)$	$m\geqslant 4b$
Theorem 2	2^b	$\frac{2^{m+s-1}-1}{2^{b-1}(2^b+1)}<\ell\leqslant\frac{2^m-1}{2^{2b-s}-1}$	$(2b-s)\vert m, m\geqslant 4b, 0\leqslant s<b$
Theorem 4	2^b	$\frac{2^{m+s-1}-1}{2^{b-1}(2^b+1)}<\ell\leqslant\frac{2^m-2^{(2b-s)}(2^z-1)-1}{2^{(2b-s)}-1}$	$(2b-s)\nmid m, m\geqslant 4b, 0\leqslant s<b,$ $1\leqslant z\equiv m \bmod (2b-s)\leqslant b$

which is approximately 2^s times greater than n', and the dimension k is $\frac{rn}{r+1}-s-m$. Hence, for the same b and r, it is easy to show that the code rate $\frac{k}{n}=\frac{r}{r+1}-\frac{s}{n}-\frac{m}{n}$ is larger than $\frac{k'}{n'}$ because $\frac{s+m}{n}<\frac{m}{n'}$.

Table 4 gives the summary of k-optimal binary linear LRCs with disjoint local repair groups whose minimum distance $d\geqslant 6$. We also list results of Theorems 2 and 4. The comparison of the number of the disjoint local repair groups illustrates that k-optimal binary linear LRCs with a wider range of parameters can be obtained from Theorem 2. Here, $\mu_2(m,2b)$ denotes the size of a maximum partial $2b$-spread in V_m.

Remark 5. As a comparison, the k-optimal binary linear LRCs generated by this paper have more flexible parameters $[n,k]$ than those in [11, 37] for a fixed locality $r=2^b$. Particularly, if we take $s=0$ in Construction 1, we have $\frac{2^{m-1}-1}{2^{b-1}(2^b+1)}<\ell\leqslant\frac{2^m-1}{2^{2b}-1}$ in Theorem 2. Note that $\ell=\frac{2^m-1}{2^{2b}-1}$ in [11] and $\mu_2(m,2b)\leqslant\frac{2^m-1}{2^{2b}-1}$ in [37]. Hence, k-optimal binary linear LRCs in [11, 37] are included in our construction. For example, letting $b=3$ and $m=12$, we obtain $\ell=65$ in [11] and $56<\ell\leqslant 65$ in [37]. However, Theorems 2 and 4 yield $56<\ell\leqslant 65$ and $227<\ell\leqslant 273$, respectively, which shows that our method constructs more k-optimal binary linear LRCs with the same locality.

More specially, we concentrate on the value of ℓ; then we have the following corollary.

Corollary 1. Let $S=\bigcup_{\substack{m\geqslant 4b, 0\leqslant s<b\\ 1\leqslant z\equiv m \bmod (2b-s)\leqslant b}}([\frac{2^{m+s-1}-1}{2^{b-1}(2^b+1)}+1, \frac{2^m-2^{(2b-s)}(2^z-1)-1}{2^{(2b-s)}-1}])$. Supposing that $\ell\in S$, then there exists an $[n=(r+1)\ell, k=r-m-s, d\geqslant 6]$ binary linear LRC with locality $r=2^b$, which is k-optimal with respect to the bound (4).

Proof. Combining Theorem 2 with Theorem 4, this corollary can be obtained directly.

4 Shortening LRC

The shortening technique can be applied to the derivation of binary linear LRCs with new parameters. Let $\mathcal{C}$ be an $[n,k,d]$ code over $\mathbb{F}_q$ and let $\mathcal{S}$ be any set of $i\in[n]$ coordinates. Consider the set $(\mathcal{S})$ of codewords which are 0 on $\mathcal{S}$; this set is a subcode of $\mathcal{C}$. Deleting the same coordinate i for all $i\in\mathcal{S}$ in each codeword of $\mathcal{C}(\mathcal{S})$ gives a code over $\mathbb{F}_q$ of length $n-|\mathcal{S}|$ called the code shortened on $\mathcal{S}$ and denoted by $\mathcal{C}'$. Hence, we obtain the following theorem with respect to the shortened LRCs.

Theorem 5. Let $\mathcal{C}$ be an $[n,k,d]$ k-optimal binary linear LRC constructed in Theorems 1 or 3 such that $n\geqslant 2(r+1)$ and $k\geqslant 2r$.

(1) Suppose that a is an integer that satisfies $0\leqslant a\leqslant\frac{n}{r+1}$. An $[n',k',d']$ LRC $\mathcal{C}'$ with locality r can be obtained by shortening $\mathcal{C}$, where parameters of "$'$" satisfy $n'=n-a(r+1), k'\geqslant k-ar$, and $d'\geqslant d$.

(2) Let $H^i=\left(\begin{smallmatrix}H_i\\ i\\ H_G^i\end{smallmatrix}\right)$ for all $i\in[\ell]$. Removing a column of each distinct submatrix $H^{i_1}, H^{i_2}, \ldots, H^{i_\tau}$ from the parity check matrix H, respectively, for $i_\tau\in[\ell]$, then there exists a shortened LRC with parameters $[n'=n-\tau, k'=k-\tau, d'=d]$.

Proof. (1) Assume that H is a parity check matrix of $\mathcal{C}$. The first ℓ rows $\boldsymbol{h}_1,\ldots,\boldsymbol{h}_\ell$ from H form a set of locality rows of $\mathcal{C}$, where $\boldsymbol{h}_i\in V_n$ with $|\mathrm{supp}(\boldsymbol{h}_i)|=r+1$. Consider the first τ locality row of H, where $1\leqslant\tau\leqslant\ell$. By deleting the first τ locality rows $\boldsymbol{h}_1,\boldsymbol{h}_2,\ldots,\boldsymbol{h}_\tau$ and the corresponding column whose index belongs to the support of $\boldsymbol{h}_i$ for all $i\in[\tau]$, we obtain an $m'\times n'$ submatrix H' with $n'=n-\tau(r+1)$, $m'=n-k-\tau$. Let $\mathcal{C}'$ be the $[n',k',d']$ linear code with the parity check matrix H'. Due to $\mathrm{rank}(H')\leqslant(n'-m')$, $k'\geqslant k-\tau r$. Note that $\mathcal{C}'$ is a shortened code of $\mathcal{C}$; then $\mathcal{C}'$ is an LRC code with minimum distance $d'\geqslant d$. This completes the proof.

(2) Since each submatrix H^i is generated by a desired matrix A and a matrix G_{M_i} for $i \in [\ell]$. Note that A can be viewed as a parity check matrix of a linear code with minimum distance $d = 5$. Assume that A' is an $[r-1, r-1-(s+t), d]$ matrix obtained by deleting a column of A. Then we construct $H^{i_1}, \ldots, H^{i_\tau}$ of H_G by utilizing the matrix A' and the remaining $(\ell - \tau)$ submatrices of H_G by utilizing the matrix A in Construction 1, where $\tau \in [\ell]$. Hence, a linear LRC $\mathcal{C}'$ with parameters $[n' = n - \tau, k' = k - \tau, d' = d]$ can be obtained. In particular, when $\tau = \ell$, the locality of LRC is $r - 1$; otherwise, the locality of LRC is r.

Below, an example is given to show a shortened LRC in Theorem 5.

Example 4. The matrix A' is generated by removing the first column from A in Example 1, i.e.,

$$A' = \begin{pmatrix} 0\,0\,0\,0\,0\,1\,0 \\ 1\,0\,0\,0\,0\,0\,1 \\ 1\,1\,0\,0\,0\,0\,0 \\ 0\,0\,1\,0\,0\,1\,1 \\ 1\,0\,0\,1\,0\,1\,0 \\ 1\,0\,0\,0\,1\,1\,1 \end{pmatrix}.$$

Then the submatrix H_G^i is constructed by $(\mathbf{0},\ G_{M_i} \cdot A')$ for each $i \in [2]$, and the remaining submatrix H_G^i is constructed by $(\mathbf{0},\ G_{M_i} \cdot A)$ for $i \in \{3, 4, \ldots, \ell\}$. Thus we obtain a $[2455, 2168, 6]$ binary linear LRC in Theorem 5, which is k-optimal with respect to the bound (4).

5 Conclusion

In this paper, we present an explicit construction of k-optimal binary linear LRCs with minimum distance $d \geqslant 6$ by investigating parity check matrices. In general, k-optimal binary linear LRCs with minimum distance $d \geqslant 6$ and locality $r = 2^b$ are constructed by t-spread of an m-dimensional vector space over $\mathbb{F}_2$ which is a collection of t-dimensional subspaces with pairwise trivial. Of interest is the idea of using intersection subspaces to replace the method of t-spread. Based on this new idea, we efficiently enlarge the range of new parameters of k-optimal binary linear LRCs with minimum distance $d \geqslant 6$ and locality $r = 2^b$. In fact, it yields more repair groups such that the corresponding constructions have more flexible lengths and dimensions. Compared with the previous studies in [11, 37] with the same locality, the code lengths of our work are larger and the code rates are higher.

Acknowledgements This work was supported in part by National Key R&D Program of China (Grant Nos. 2022YFA1004900, 2022YFA1005000) and National Natural Science Foundation of China (Grant No. 62272303). We would like to thank Professor Chaoping XING for introducing us to this problem. During this work, he provided many valuable discussions and expert advices, which are very useful for improving the quality of this paper.

References

1 Huang C, Simitci H, Xu Y, et al. Erasure coding in Windows Azure storage. In: Proceedings of the USENIX Annual Technical Conference, 2012. 15–26
2 Sathiamoorthy M, Asteris M, Papailiopoulos D, et al. XORing elephants: novel erasure codes for big data. Proc VLDB Endow, 2013, 6: 325–336
3 Dimakis A G, Godfrey P B, Wu Y, et al. Network coding for distributed storage systems. IEEE Trans Inform Theor, 2010, 56: 4539–4551
4 Gopalan P, Huang C, Simitci H, et al. On the locality of codeword symbols. IEEE Trans Inform Theor, 2012, 58: 6925–6934
5 Rashmi K V, Shah N B, Kumar P V. Optimal exact-regenerating codes for distributed storage at the MSR and MBR points via a product-matrix construction. IEEE Trans Inform Theor, 2011, 57: 5227–5239
6 Hou H, Lee P P C, Shum K W, et al. Rack-aware regenerating codes for data centers. IEEE Trans Inform Theor, 2019, 65: 4730–4745
7 Zhou L Y, Zhang Z F. Explicit construction of minimum bandwidth rack-aware regenerating codes. Sci China Inf Sci, 2022, 65: 179301
8 Grezet M, Freij-Hollanti R, Westerbäck T, et al. Bounds on binary locally repairable codes tolerating multiple erasures. 2017. arXiv:1709.05801
9 Cadambe V R, Mazumdar A. Bounds on the size of locally recoverable codes. IEEE Trans Inform Theor, 2015, 61: 5787–5794
10 Guruswami V, Xing C, Yuan C. How long can optimal locally repairable codes be? IEEE Trans Inform Theor, 2019, 65: 3662–3670
11 Wang A, Zhang Z, Lin D. Bounds for binary linear locally repairable codes via a sphere-packing approach. IEEE Trans Inform Theor, 2019, 65: 4167–4179
12 Chen B, Fang W, Xia S T, et al. Improved bounds and Singleton-optimal constructions of locally repairable codes with minimum distance 5 and 6. IEEE Trans Inform Theor, 2021, 67: 217–231

13 Tamo I, Barg A. A family of optimal locally recoverable codes. IEEE Trans Inform Theor, 2014, 60: 4661–4676
14 Hao J, Xia S T, Shum K W, et al. Bounds and constructions of locally repairable codes: parity-check matrix approach. IEEE Trans Inform Theor, 2020, 66: 7465–7474
15 Ma L, Xing C. Constructive asymptotic bounds of locally repairable codes via function fields. IEEE Trans Inform Theor, 2020, 66: 5395–5403
16 Tamo I, Papailiopoulos D S, Dimakis A G. Optimal locally repairable codes and connections to matroid theory. IEEE Trans Inform Theor, 2016, 62: 6661–6671
17 Jin L. Explicit construction of optimal locally recoverable codes of distance 5 and 6 via binary constant weight codes. IEEE Trans Inform Theor, 2019, 65: 4658–4663
18 Zhang Y, Kan H B. Locally repairable codes from combinatorial designs. Sci China Inf Sci, 2020, 63: 122304
19 Zhang Z F, Xu J K, Liu M L. Constructions of optimal locally repairable codes over small fields. Sci Sin Math, 2017, 47: 1607–1614
20 Hao J, Xia S T, Chen B. On optimal ternary locally repairable codes. In: Proceedings of the IEEE International Symposium on Information Theory, 2017. 171–175
21 Cadambe V, Mazumdar A. An upper bound on the size of locally recoverable codes. In: Proceedings of the International Symposium on Network Coding, 2013. 1–5
22 Silberstein N, Zeh A. Optimal binary locally repairable codes via anticodes. In: Proceedings of the IEEE International Symposium on Information Theory, 2015. 1247–1251
23 Prakash N, Kamath G M, Lalitha V, et al. Optimal linear codes with a local-error-correction property. In: Proceedings of the IEEE International Symposium on Information Theory, 2012. 2776–2780
24 Grezet M, Freij-Hollanti R, Westerbäck T, et al. Alphabet-dependent bounds for linear locally repairable codes based on residual codes. IEEE Trans Inform Theor, 2019, 65: 6089–6100
25 Song W, Dau S H, Yuen C, et al. Optimal locally repairable linear codes. IEEE J Sel Areas Commun, 2014, 32: 1019–1036
26 Cai H, Miao Y, Schwartz M, et al. On optimal locally repairable codes with super-linear length. IEEE Trans Inform Theor, 2020, 66: 4853–4868
27 Xing C, Yuan C. Construction of optimal (r, δ)-locally recoverable codes and connection with graph theory. IEEE Trans Inform Theor, 2022, 68: 4320–4328
28 Luo G, Ezerman M F, Ling S. Three new constructions of optimal locally repairable codes from matrix-product codes. IEEE Trans Inform Theor, 2023, 69: 75–85
29 Cai H, Fan C, Miao Y, et al. Optimal locally repairable codes: an improved bound and constructions. IEEE Trans Inform Theor, 2022, 68: 5060–5074
30 Wang A, Zhang Z. Repair locality with multiple erasure tolerance. IEEE Trans Inform Theor, 2014, 60: 6979–6987
31 Tan P, Zhou Z, Sidorenko V, et al. Two classes of optimal LRCs with information (r, t)-locality. Design Code Cryptogr, 2020, 88: 1741–1757
32 Jin L, Kan H, Luo Y, et al. Binary locally repairable codes with large availability and its application to private information retrieval. IEEE Trans Inform Theor, 2022, 68: 2203–2210
33 Dhar M, Gopi S. A construction of maximally recoverable LRCs for small number of local groups. In: Proceedings of the IEEE International Symposium on Information Theory, 2023. 1753–1757
34 Zhang Z K, Gu S S, Zhang Q Y. Scalable local reconstruction code design for hot data reads in cloud storage systems. Sci China Inf Sci, 2022, 65: 222303
35 Zeh A, Yaakobi E. Optimal linear and cyclic locally repairable codes over small fields. In: Proceedings of the IEEE International Symposium on Information Theory, 2015. 1–5
36 Nam M Y, Song H Y. Binary locally repairable codes with minimum distance at least six based on partial t -spreads. IEEE Commun Lett, 2017, 21: 1683–1686
37 Ma J, Ge G. Optimal binary linear locally repairable codes with disjoint repair groups. SIAM J Discrete Math, 2019, 33: 2509–2529
38 Ling S, Xing C. Coding Theory: A First Course. Cambridge: Cambridge University Press, 2004
39 Bu T. Partitions of a vector space. Discrete Math, 1980, 31: 79–83
40 Etzion T, Vardy A. Error-correcting codes in projective space. IEEE Trans Inform Theor, 2011, 57: 1165–1173
41 Goparaju S, Calderbank R. Binary cyclic codes that are locally repairable. In: Proceedings of the IEEE International Symposium on Information Theory, 2014. 676–680
42 Kim C, No J S. New constructions of binary LRCs with disjoint repair groups and locality 3 using existing LRCs. IEEE Commun Lett, 2019, 23: 406–409

Appendix A Examples of desired matrix A

We provide an approach to construct the desired matrix A with the help of a computer search program. Note that a desired matrix A can be viewed as the parity check matrix of a $[2^b, 2^b - 2b, d \geqslant 5]$ binary linear code. Although Wang et al. [11] presented the explicit construction of binary linear code with parameters $[2^b, 2^b - 2b, d \geqslant 5]$ from a shortened nonprimitive cyclic code, which cannot be used directly here. A necessary condition for the desired matrix A is that it contains a submatrix in which any two distinct columns are linearly independent. This makes it difficult to give an explicit construction of the desired matrix A. In addition, LRCs constructed by using an arbitrary $t \times n$ matrix A have the same code length, dimension, and minimum distance, which implies that we can weaken its explicit construction. By the computer program MAGMA, we have found some examples of the desired matrix A, which also shows the existence of these desired matrices. However, how to construct more desired matrices A by using theoretical analysis and an effective search algorithm, remains an open problem.

We briefly recall the construction of a binary linear code with parameters $[2^b, 2^b - 2b, \geqslant 5]$ in [11]. Let $n = 2^b + 1$ and let α be a primitive root of $x^n - 1$ with minimal polynomial $M_\alpha(x)$. Clearly, the degree of $M_\alpha(x)$ is $2b$. Define $\mathcal{C}$ to be the binary cyclic code of length n generated by $(x-1)M_\alpha(x)$. It is not hard to show that $\{\alpha^i : i = -2, -1, 0, 1, 2\}$ forms a subset of the roots of the generator polynomial of $\mathcal{C}$, so $\mathcal{C}$ is an $[n = 2^b + 1, k = 2^b - 2b, d \geqslant 6]$ binary linear code. The code $\mathcal{C}$ can be punctured by deleting one of the check bits to yield a code $\mathcal{C}'$ of length 2^b with $2b$ check bits and $d \geqslant 5$. Hence, a $[2^b, 2b]$ parity check matrix A' of "$'$" can be obtained. Notice that a desired matrix A has the same parameters as matrix A' because A also can be viewed as a parity check matrix of an $[n = 2^b, k = 2^b - 2b, d \geqslant 5]$ linear code. Hence, applying the row transformation to A', A' can be transformed into a desired matrix A. Here, if a $z \times n$ submatrix of a $k \times n$ desired matrix satisfies that any two distinct column vectors from the submatrix are linearly independent, let an $[n, k]_z$ matrix denote this desired matrix A. See the following examples of the desired

matrix A.

- $[8,6]_4$ matrix A:

$$\begin{pmatrix} 10000010 \\ 01000001 \\ 11100000 \\ 10010011 \\ 01001010 \\ 01000111 \end{pmatrix}.$$

- $[16,8]_5$ matrix A:

$$\begin{pmatrix} 0010010001101011 \\ 0010000010001111 \\ 0100000100011110 \\ 1000000001001110 \\ 0101000011111100 \\ 0110100001011001 \\ 0110011000111110 \\ 0000010111011101 \end{pmatrix}.$$

- $[32,10]_6$ matrix A:

$$\begin{pmatrix} 10000000001001011101011010111010 \\ 01000000000100101110101101011101 \\ 00100000001100000001000101111101 \\ 00010000001000010110110001101101 \\ 10001010001000111010011001011101 \\ 01001001011001110100010111001101 \\ 00100001000111101010010010101111 \\ 00010000101111010001000000011000 \\ 00000110000000100010111110100010 \\ 00000000100101110101101011101001 \end{pmatrix}.$$

- $[64,12]_8$ matrix A:

$$\begin{pmatrix} 1000000000000101100111010010011010100010010001010110010010111 0011 \\ 0100000000000100001001110101011110010010101001001111010101110 0100 \\ 0010000000000100001001110101011110010010101001001111010101110 0010 \\ 0001000000000010000100111010101111001001010100100111101010111 0001 \\ 1000000110000110000001010101101110111001001010010011101110110 0101 \\ 0100000010100000010011010110100001100011111111000110000101101011 \\ 0010000010011011011010101011011001000001000110100011110001011 \\ 0001000110100011110001011000100000100110110110101010110110110010 \\ 0000100100110010010000011001111011010110101010101101011011110011 \\ 0000010100101110100000100111011000001101110010000010111010010100 \\ 0000011001011010011000000001011111001110111010111011100111111010 \\ 0000000001011001111010010011010100010010001010110010010111001 1010 \end{pmatrix}.$$

- $[128,14]_{10}$ matrix A:

$$\begin{pmatrix} 10000000000000010100100101110010110101111000111011010010000100010111110010011111010001000010010110111000111101011010011101001001 \\ 01000000000000100011011001011100110001001101101011001101001010100100011101110001001010100101100110101101100100011001110100110110 \\ 00100000000000010001101100101110011000100110110101100110100101010010001110111000100101010010110011010110110010001100111010011011 \\ 00010000000000100001111101110010100111000101110100000100000100101110111010101101111110011100111111010101011101110100100 \\ 10000000001110011101010111001110000000001000001011111001011011001101010011111110100100100101111110010101100110110100111110100 00 \\ 01000000000011010101100000000001001001101001011100010101111001000101000111101100001101111100010100010011110101000111010010111001 \\ 00100000001001110101101100101001011100011111101011111000111010010100110110101110010000000100000111011000000011110000000111011100 \\ 00010000001011100001010011100011001100011100101000011101000000100000100101110111010101101111110011100111111010101011101110100 10 \\ 00001000001001011101110101011011111100111001111110101010111011100100000100000101110000101001110011001100011100101000011101 0 \\ 00000100000101111100101101100110101001111111010010010101111110010101100110110100111110100000100000000011100111010101110011100 00 \\ 00000010001100110111001000011111000010011101100110001000000010100000110101011101001001001110110111001001001011101010110000010 1 \\ 00000001001000111011001011111010111010011010101011111001011110000100001111010011111010101011001011101011110110100111011100010001 \\ 00000000100100110100101110001010111100100010100011111011000011011111000101000100111101010001110100101100100100000000001101010110 \\ 00000000011100010011100101111001101011100001100001010111011110100100001100001110101100111101001110010001110000000000100110010 \end{pmatrix}.$$