Constructions of ternary LRCs with optimal distance

Youliang Zheng¹,a, Ruihu Li¹,b*, Jingjie Lv¹,c, Mao Zhang¹,d
¹Department of Basic Sciences Air Force Engineering University Xi’an, Shaanxi, 710051, P. R. China
b*ldlv@xidian.edu.cn
aE-mail: zxy951783247@163.com, b* Corresponding author’s e-mail: llzsy2015@163.com, cE-mail: juxianljj@163.com, dE-mail: 18092166304@163.com

Abstract—As a new class of erasure codes, locally repairable codes (LRCs) can provide data repair capability for distributed storage system. In this paper, we firstly propose an effective construction method to obtain optimal LRCs based on ternary optimal codes. Then, three classes of optimal LRCs are constructed by ternary cyclic codes and constant cyclic codes. All codes we gained have the optimal distance and reach the Cadambe-Mazumdar (C-M) bound.

1. INTRODUCTION
With the coming era of big data, the amount of data in the world increases rapidly, and the requirements for storage system are also gradually improving. Distributed storage system has been widely used because of its extensible structure, which improves storage efficiency of the system. In order to improve the fault-tolerant capability of distributed storage system, Gopalan et al. [1] proposed the concept of locally repairable codes in 2012. For a linear code with length $n$, dimension $k$ and distance $d$, if each code symbol of the code can be repaired by no more than other $r$ symbols, the code is called locally repair code with locality $r$ and denoted by $[n, k, d; r]$ in this paper. Cadambe and Mazumdar [2] proposed a field size dependent bound about LRCs as follows:

$$k \leq \min_{m \in Z^+} \left\{ \lfloor tr + k_{opt}^q (n - t(r + 1), d) \rfloor \right\},$$

(1)

where $Z^+$ denotes positive integer and $k_{opt}^q (n, d)$ is the largest possible dimension of a code with length $n$, distance $d$ and alphabet size $q$.

In this paper, a code that has maximum distance given code length $n$ and dimension $k$ is called $d$-optimal code. And we use $[d, r]$-optimal LRCs to denote the $d$-optimal LRCs which reach C-M bound. The $[d, r]$-optimal LRCs have great practical value with the advantages of both $d$-optimal codes and LRCs, can be well suited for distributed storage systems. So it is interesting to construct $[d, r]$-optimal LRCs using $d$-optimal codes. The $d$-optimal codes studied in this paper come from code tables which are collected by Grassl [3]. On the other hand, cyclic codes can also construct $d$-optimal codes. Since the special structure of cyclic code, it can be effectively used in the coding and decoding process of shift registers, and can design the locality of codes well. Therefore, constructing LRCs using cyclic codes is interesting. Zeh and Yaakobi studied three constructions of optimal linear codes over small field in [4]. Reference [5] obtained some LRCs with $d \geq 4$ by analyzing cyclic codes in the binary and ternary fields. Hao, Xia and Chen constructed some optimal binary LRCs meeting the Singleton-like bound [6]. In references [7-8], scholars considered to construct LRCs through constant cyclic codes.
Rao et al. [9] proposed a construction method of cyclic codes with locality 2 in binary field. Tan et al. studied some optimal cyclic LRCs via cyclotomic polynomials in [10]. Reference [11] constructs LRCs by using subfield subcodes of cyclic codes.

In terms of ternary \(d\)-optimal codes, people mainly studied LRCs by low-dimensional generator matrix. There are still few researches for the high-dimensional case at present. This paper constructs high-dimensional LRCs with the method of replacing the rows of parity check matrix when \(n-k\) is small. Many \(d\)-optimal codes can also be constructed through cyclic codes. In the previous research, people mainly constructed cyclic LRCs in the binary field and general fields, the studies of ternary codes are not enough. In this paper, BCH bound is used to select appropriate 3-cyclotomic cosets with codelength \(n\) as the defining set of dual codes, so as to pre-construct some codes with small dual distances, and then Hamming bound and Griesmer bound are used to further determine the parameters of codes, thus we construct some LRCs with small localities. The codes constructed above all reach C-M bound and are \([d,r]\)-optimal LRCs.

An outline of this paper is as follows. In Section II, we introduce some mathematical notations and definitions. Section III gives the detailed construction process of \([d,r]\)-optimal LRCs by \(d\)-optimal code. In Section IV, we construct three classes \([d,r]\)-optimal LRCs by cyclic codes and constant cyclic codes. Section V concludes the paper.

2. PLIMINARIES

First we give some mathematical notations and definitions which will be used later:

- Let \(v = (v_1, v_2, \cdots, v_n)\) and \(v_i\) be the \(i\)-th coordinate of \(v\). The support of \(v\) is denoted as \(\text{supp}(v) = \{i \mid v_i \neq 0\}\) and the Hamming weight of \(v\) is denoted as \(\text{wt}(v) = |\text{supp}(v)|\).

- Let \(0_n\), \(1_n\) and \(2_n\) denote the all-zero row vector, all-one row vector and all-two row vector of length \(n\), respectively.

- Let \([i] = \{1, 2, \cdots, i\}\) and \([a, b] = \{a, a+1, \cdots, b \mid a \leq b\}\), where \(i\), \(a\) and \(b\) are three positive integers.

- For a linear code with an \(m \times n\) \((m = n - k)\) parity-check matrix \(H\), denote the \(i\)-th row of \(H\) as \(h_i (i \in \mathbb{N})\).

- \(cnr(n, l)\) denotes an \(l\)-bit \(n\)-base array which is from \([1, 2, \cdots, l]\) to \([n - l, n - l + 1, \cdots, n - 1]\).

- \(|V|\) denotes the number of vectors in the vector space \(V\).

**Definition 1** A generator matrix of 2-dimensional Simplex code is

\[
S_2 = \begin{pmatrix}
1011 \\
0112
\end{pmatrix}.
\]

A generator matrix of \(k\)-dimensional Simplex code can be constructed through recursion formula

\[
S_k = \begin{pmatrix}
S_{k-1} & 0_{k-1} & S_{k-1} \\
0_{n_{k-1}} & 1 & 2_{n_{k-1}}
\end{pmatrix},
\]

where \(n_{k-1}\) is the number of columns in \(S_{k-1}\). It is easy to know that \(S_k\) has \((3^k - 1) / 2\) columns, and Simplex code can be still generated by changing the order of the columns in \(S_k\).

**Definition 2** A set \(S\) of \(F_q^n\) is cyclic if \((a_{n+1}, a_0, a_1, \cdots, a_{n-2}) \in S\) whenever \((a_0, a_1, \cdots, a_{n-2}) \in S\). A linear code \(C\) is called a cyclic code if \(C\) is a cyclic set.

**Definition 3** If \(\gcd(m, 3) = 1\), the 3-cyclotomic coset modulo \(n\) containing \(x\) is defined as

\[
C_x = \{x, 3 \cdot x, 3^2 \cdot x, \cdots, 3^{a-1} \cdot x\} \pmod{n}
\]

where \(a\) is the smallest positive integer such that \(3^a \cdot x \equiv x \pmod{n}\). The smallest integer in \(C_x\) is called the coset leader of \(C_x\).
Let $\alpha$ be a primitive element in the splitting field $F_{q^m}$ with $n | (3^m - 1)$, the defining set of a cyclic code $C = \langle n,k,d \rangle$ is $T_C = C_{t_1} \cup C_{t_2} \cup \ldots \cup C_{t_l}$, where $t_1,t_2,\ldots,t_l$ are coset leaders belonging to distinct 3-cyclotomic cosets modulo $n$.

**Lemma 1**\(^{[12]}\) Let $C = \langle n,k,d \rangle$ be a linear code, if the maximum weight of the rows that reach the cover condition in the parity-check matrix of the code is $r + 1$, then the locality of $C$ is $r$.

**Lemma 2**\(^{[12]}\) Let $C = \langle n,k,d \rangle$ be a cyclic code, and $D$ be the dual code of $C$. If the distance of $D$ is $d^\perp$, then the locality of $C$ is $r = d^\perp - 1$.

Since constant cyclic codes have similar algebraic structure to cyclic codes, Lemma 2 is also applicable to constant cyclic codes.

**Lemma 3**\(^{[13]}\) (BCH bound) Let $C = \langle n,k,d \rangle$ be a cyclic code over $F_q$ with defining set $T$, if $T$ contains $\delta$ length consecutive elements, then the minimum distance of $C$ is at least $\delta + 1$.

**Lemma 4**\(^{[13]}\) (Constant cyclic codes BCH bound) Let $C = \langle n,k,d \rangle$ be a ternary 2-constant cyclic code with defining set $T$, if $T$ contains length consecutive elements, then the minimum distance of $C$ is at least $\delta + 1$, where two adjacent numbers of consecutive elements differ by 2.

### 3. Construction of LRCS using optimal distance codes

The $[d,r]$-optimal LRCS constructed in this section are based on the ternary $d$-optimal codes, which satisfy the condition of Lemma 2 by replacing the parity-check matrix of the codes. First, we divide the parity-check matrix into two parts:

$$H = \begin{pmatrix} H_L \\ H_G \end{pmatrix} \quad (3)$$

where $H_L$ needs to satisfy the cover condition to determine the locality, and $H_G$ is composed of other dual codewords. Satisfying the cover condition means that $\text{supp}[H_L] = [n]$.

To construct the LRCS, first, the $d$-optimal codes with maximum code length is selected under the same distance of $C$ and dimension of $H$. Then the codewords of $C^\perp$ are analyzed to construct the $H_L$ matrix with the highest deleting efficiency. After that, add the $H_L$ matrix to $m$ rows to construct a new parity-check matrix $H$. Finally, we regard the new check matrix as seed matrix and construct the rest code length $[d,r]$-optimal LRCS by deleting the seed matrix.

#### 3.1. Construction of $H_L$ matrix

After determining the parameters of the $d$-optimal codes, the construction steps of $H_L$ are as follows:

1. Choose $V = \{ h_i | \text{wt}(h_i) = d^\text{min}, h_i \in C^\perp \}$
2. Let $v \leftarrow |V|$, full $\leftarrow 0$
3. For $l = 2 : m$
4. Let $\text{cnr}(v,l) \leftarrow [1,2,\ldots,l-1,l]$;
5. While $\text{cnr}(v,l) = [v-l,v-l+1,\ldots,v-1,v]$
6. Let $V_{\text{cnr}(v,l)} = [h_i,h_2,\ldots,h_l]$;
7. Let $H'_{L} \leftarrow V_{\text{cnr}(v,l)}$;
8. If $\text{rank}(H'_{L}) = l$ & $\text{supp}[H'_{L}] = [n]$
9. Calculate $\text{full}(H'_{L})$
10. If $\text{full}(H'_{L}) > \text{full}$
11. $H_L \leftarrow H'_{L}$, full $\leftarrow \text{full}(H'_{L})$
12. $\text{cnr} \leftarrow \text{cnr} + 1$
13. $l \leftarrow l + 1$
where \( f \) \textit{full} denotes the number of all non-zero element columns in the \( H_L \). The \( H_L \) matrix consists of dual codewords, and the rows of the matrix ranges general from \( 2 \) to \( m \). Since LRCs have higher repair efficiency with small locality, the codewords with minimum weight are first selected in the above program to generate vectors space \( V \) (Step 1). Then, combining the vectors in \( V \) to construct all matrices with rows number from \( 2 \) to \( m \), and the \( f \) \textit{full} of the matrices with full rank and satisfying cover condition is calculated (Step 8, 9). According to Lemma 2, it is easy to know that each column deleted in \( H \) which corresponding to the full weight column in \( H_L \), the locality of \( C \) reduce 1. Therefore, we need to find the \( H_L \) with maximum \( f \) \textit{full}, so that the localities of codes with other code length will decrease fastest by deleting \( H \). And when \( d=3 \), we find that the \( H_L \) of codes can be obtained directly by the generator matrix of Simplex code.

3.2. Construction of seed matrix from \( H_L \)

After obtaining \( H_L \) matrix, \( m-1 \) rows other dual codewords should be added to \( H_L \) to form a new parity-check matrix \( H \) of \( C \). Then, seed matrix can be obtained by arranging the columns in \( H \) from left to right according to the weight of the columns from small to large, where equal weight columns containing zero elements should be staggered zero elements for sorting. Thus, the parity-check matrix of \([d,r]\)-optimal LRCs within a code length range can be obtained by deleting the seed matrix.

\textbf{Example 1:} When \( d=3,m=4,11 \leq n \leq 40 \), \( H_L \) can be constructed by \( S_4 \) matrix. After sorting the columns of the matrix, the following seed matrix can be obtained (the seed matrix has been segmented with dashed lines according to different deletion efficiencies):

\[
H_{40} = \begin{bmatrix}
10001001010 & 01020 & 2201101210121021 & 21122112 \\
0100011010 & 20102 & 2201101210121021 & 21122112 \\
00101120000 & 20102 & 201120221011022 & 12112122 \\
00010001122 & 00102 & 011011102220222 & 11221122 \\
\end{bmatrix}
\]

by deleting \( i \) column from the back to the front of \( H_{40} \) \((0 \leq i \leq 8)\), LRCs \([40-i,36-i,3;26-i]\) can be obtained; then by successively deleting \( j \) columns \((0 \leq j \leq 15)\), we can get LRCs \([31-j,27-j,3,18-[3j+3/4]]\); finally, by successively deleting \( f \) columns \((0 \leq f \leq 4)\), LRCs \([15-f,11-f,3,6-[f+1/2]]\) can be obtained.

\textbf{Example 2:} When \( d=7,m=10,16 \leq n \leq 22 \). By analyzing the dual code of the optimal code with parameter \([22,12,7]\), we can select 4 rows dual codewords that satisfy the cover condition to construct \( H_L \), and then seed matrix can be get by adding 6 rows dual codewords to the \( H_L \) \((H_L \text{ and } H_G \text{ have been distinguished by dashed line})\):

\[
H_{10} = \begin{bmatrix}
00000000002012100 & 20 & 1111 \\
00020000100020 & 02 & 2121 \\
2120002000000000 & 10 & 1111 \\
0002000100000000 & 01 & 2121 \\
1000000000012211 & 00 & 1000 \\
0010000000101112 & 20 & 0200 \\
000000100121211 & 10 & 1010 \\
0000001002021101 & 00 & 1210 \\
0000000012211101 & 20 & 1000 \\
0000100001201011 & 00 & 0210 \\
\end{bmatrix}
\]

by deleting \( i \) columns from the back to the front of \( H_{10} \) \((0 \leq i \leq 4)\), LRCs \([22-i,12-i,7,8-i]\) can be obtained; then deleting the last 1 and 2 columns, we can get LRCs with parameters of \([17,7,7;4]\) and \([16,6,7;3]\), respectively.
4. CONSTRUCTION OF LRCs USING CYCLE CODES AND CONSTANT CYCLIC CODES

LRCs with small locality have higher repair efficiency in practical application. According to Lemma 2, the locality of cyclic code is \(d_{⊥} = d - 1\), and cyclic code with small dual distance can be designed by BCH bound, optimal LRCs with small localities can be constructed by studying cyclic codes. First calculate the code length \(n\) corresponding 3-cyclotomic cosets. Then, according to BCH bound, the union of 3-cyclotomic cosets with fewer continuous elements is selected as the defining sets of dual codes, so, some codes with small dual distance are prefabricated. Finally, the Hamming bound and Griesmer bound are used to further determine the parameters of codes, thus some classes of \([d,r]\)-optimal LRCs with small localities are obtained. Constant cyclic codes have similar structures to cyclic codes, LRCs can be constructed by similar methods based on constant cyclic codes.

**Construction 1:** Suppose \(m\) is a positive integer greater than 2, then there exists an explicit construction of ternary cyclic \([d,r]\)-optimal LRCs with parameters \([3^m - 1, m, 2 : 3^{m-1} - 1]\).

**Proof:** Let \(C = [3^m - 1, m, d]\), and suppose the defining set of \(C^⊥\) is \(T_{C^⊥} = C_1 = \{1, 3^2, \ldots, 3^{m-1}\}\). According to BCH bound, we have that the distance \(d \geq 2 \cdot 3^{m-1}\). Since \(n = 3^m - 1\) and \(k = m\), applying Griesmer bound, then we can obtain the distance \(d \leq 2 \cdot 3^{m-1}\). Naturally, \(C\) is a \(d\)-optimal code with \(d = 2 \cdot 3^{m-1}\). Moreover, by virtue of Hamming bound and BCH bound, the dual distance \(d^⊥ = 2\) can be gained similarly. By Lemma 2, locality of \(C\) is 1. As a consequence, \(C = [3^m - 1, m, 2 : 3^{m-1} - 1]\) is a class of \([d,r]\)-optimal LRCs derived from the C-M bound.

**Construction 2:** Suppose \(m\) is a positive integer and \(\gcd(m, 3) = 1\), then there exists an explicit construction of ternary cyclic \([d,r]\)-optimal LRCs with parameters \([8m, 2, 6m; 1]\).

**Proof:** Let \(C = [8m, 2, d]\), and suppose the defining set of \(C^⊥\) is \(T_{C^⊥} = C_m = \{m, 3m\}\). According to BCH bound, we have that the distance \(d \geq 6m\). Since \(n = 8m\) and \(k = 2\), applying Griesmer bound, then we can obtain the distance \(d \leq 6m\). Naturally, \(C\) is a \(d\)-optimal code with \(d = 6m\). Moreover, by virtue of Hamming bound and BCH bound, the dual distance \(d^⊥ = 2\) can be gained similarly. By Lemma 2, locality of \(C\) is 1. As a consequence, \(C = [8m, 2, 6m; 1]\) is a class of \([d,r]\)-optimal LRCs derived from the C-M bound.

**Construction 3:** Suppose \(m\) is a positive integer greater than 5 and \(\gcd(m, 3) = 1\), then there exists an explicit construction of ternary 2-constant cyclic \([d,r]\)-optimal LRCs with parameters \([4m, 2, 3m; 1]\).

**Proof:** Let \(C = [4m, 2, d]\), and suppose the defining set of \(C^⊥\) is \(T_{C^⊥} = C_m = \{m, 3m\}\). According to BCH bound, we have that the distance \(d \geq 3m\). Since \(n = 4m\) and \(k = 2\), applying Griesmer bound, then we can obtain the distance \(d \leq 3m\). Naturally, \(C\) is a \(d\)-optimal code with \(d = 3m\). Moreover, by virtue of Hamming bound and BCH bound, the dual distance \(d^⊥ = 2\) can be gained similarly. By Lemma 2, locality of \(C\) is 1. As a consequence, \(C = [4m, 2, 3m; 1]\) and \(C^⊥ = [4m, 4m - 2, 2; 3m - 1]\) are \([d,r]\)-optimal LRCs derived from the C-M bound.

5. CONCLUSIONS

In this paper, we present an algorithm to construct LRCs for the known optimal ternary distance codes from the code table. For ternary cyclic codes and constant cyclic codes, we construct three kinds of codes with small locality by analyzing the 3-cyclotomic cosets of codes. These codes all reach C-M bound and are ternary \([d,r]\)-optimal LRCs. In the future, we will further study how to use cyclic codes to construct other LRCs with good parameters.

ACKNOWLEDGMENTS

This work is supported by National Natural Science Foundation of China (Nos.11801564, 11901579).
REFERENCES

[1] P. Gopalan, C. Huang, H. Simitci and S. Yekhanin, “On the Locality of Codeword Symbols,” IEEE Transactions on Information Theory, vol. 58, no. 11, pp. 6925-6934, Nov. 2012.

[2] V. Cadambe and A. Mazumdar, “An upper bound on the size of locally recoverable codes,” 2013 International Symposium on Network Coding, Calgary, AB, 2013, pp. 1-5.

[3] M. Grassl, Bounds on the minimum distance of linear codes. http://www.codetables.de, 2015.

[4] A. Zeh and E. Yaakobi, “Optimal linear and cyclic locally repairable codes over small fields,” 2015 IEEE Information Theory Workshop, Jerusalem, 2015, pp. 1-5.

[5] C. Kim and J. No, “New Constructions of Binary and Ternary Locally Repairable Codes Using Cyclic Codes,” IEEE Communications Letters, vol. 22, no. 2, pp. 228-231, Feb. 2018.

[6] J. Hao, S. Xia and B. Chen, “Some results on optimal locally repairable codes,” IEEE International Symposium on Information Theory, Barcelona, 2016, pp. 440-444.

[7] Z. Sun, S. Zhu and L. Wang, “Optimal constacyclic locally repairable codes,” IEEE Communications Letters, vol. 23, no. 2, pp. 206-209, Feb. 2019.

[8] B. Chen, W. Fang, S. Xia and F. Fu, “Constructions of Optimal (r,δ) Locally Repairable Codes via Constacyclic Codes,” IEEE Transactions on Communications, vol. 67, no. 8, pp. 5253-5263, Aug. 2019.

[9] Y. Rao, L. Ruihu, “On Binary Cyclic Locally Repairable Codes with Locality 2,” IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, vol. 100, no. 7, pp. 1588-1591, 2017.

[10] P. Tan, Z. Zhou, H. Yan and U. Parampalli, “Optimal Cyclic Locally Repairable Codes via Cyclotomic Polynomials,” IEEE Communications Letters, vol. 23, no. 2, pp. 202-205, Feb. 2019.

[11] I. Tamo, A. Barg, S. Goparaju and R. Calderbank, “Cyclic LRC codes and their subfield subcodes,” IEEE International Symposium on Information Theory, Hong Kong, 2015, pp. 1262-1266.

[12] P. Huang, E. Yaakobi, H. Uchikawa and P. H. Siegel, “Binary Linear Locally Repairable Codes,” IEEE Transactions on Information Theory, vol. 62, no. 11, pp. 6268-6283, Nov. 2016.

[13] W. Huffman, V. Pless, Fundamentals of Error-Correcting Codes. Cambridge University Press, London, 1975.