Received January 15, 2022, accepted April 6, 2022, date of publication April 14, 2022, date of current version April 27, 2022.
Digital Object Identifier 10.1109/ACCESS.2022.3167411

Some New Sequential-Recovery LRCs Based on Good Polynomials

ZHI JING, (Student Member, IEEE), AND HONG-YEOP SONG, (Senior Member, IEEE)
School of Electrical and Electronic Engineering, Yonsei University, Seoul 03722, South Korea
Corresponding author: Hong-Yeop Song (hysong@yonsei.ac.kr)

This work was supported by the National Research Foundation of Korea (NRF) Grant by the Korean Government through Ministry of Sciences and ICT (MSIT) under Grant 2020R1A2C2011969.

ABSTRACT We propose a new construction of sequential-recovery Locally Repairable Codes (LRCs) of length n with even locality r for two erasures, based on some 'good' polynomials, over a relatively small alphabet of size $q \approx \frac{(r+1)n}{r+2}$, which becomes rate optimal in some cases. We also derive an explicit form of the upper bound on the minimum distance of these codes with some additional constraints. The minimum distance of the proposed sequential-recovery LRCs for $r = 2$ achieves this explicit bound when $k = \frac{n}{2}$ and is one less than the bound when $k < \frac{n}{2}$.

INDEX TERMS Distance bound, generalized Hamming weight, locally repairable codes, sequential-recovery LRCs.

I. INTRODUCTION

For the reliability of the distributed storage systems (DSSs), the locally repairable codes (LRCs) has drawn much attention, since any erased symbol can be repaired by only a few other symbols. Let C be an $[n, k, d]$ linear code over $\mathbb{F}_q$, whose length is n, dimension is k, and minimum distance is d. Let $\boldsymbol{c} = (c_0, c_1, \ldots, c_{n-1})$ be a codeword of C. The code C is said to be an *LRC with locality* r [1] if, for each $i = 0, 1, \ldots, n-1$, the coded symbol c_i is a linear combination of r other symbols, and denoted by an $[n, k, d, r]$ LRC.

In this paper, we will consider LRCs with multiple erasures. They are divided into sequential- and parallel-recovery LRCs based on whether the repair process is either sequential or parallel. Now, let C be an $[n, k, d, r]$ LRC with a codeword $\boldsymbol{c}$. Then it is said to be a t-sequential-recovery (t-seq) LRC [2] if, for any s ($\leq t$) erased symbols, there exists an arrangement of s erased positions given by $(j_0, j_1, \ldots, j_{s-1})$ such that, for each $l = 0, 1, \ldots, s-1$, there is a subset $R_l \subset \{0, 1, \ldots, n-1\}$ satisfying

1) $j_l \in R_l$ and $|R_l| \leq r + 1$,
2) $R_l \cap \{j_{l+1}, j_{l+2}, \ldots, j_{s-1}\} = \emptyset$, and
3) $c_{j_l} = \sum_{i \in R_l \setminus j_l} a_i c_i$, for some $a_i \in \mathbb{F}_q$.

The t-parallel-recovery (t-para) $[n, k, d, r]$ LRCs [2] is defined similarly as the sequential-recovery LRCs by replacing the second condition with

$$R_l \cap \{j_0, j_1, \ldots, j_{s-1}\} = j_l.$$

It is obvious that the t-para LRCs is also the t-seq LRCs since it can be locally repaired by any arrangement. As shown in Fig. 1, various types of parallel-recovery LRCs are proposed.

- LRCs with availability (r, t) [3]: Let C be an $[n, k, d]$ code over $\mathbb{F}_q$ and $\boldsymbol{c} = (c_1, c_2, \ldots, c_n) \in C$. The code C is said to have the availability (r, t) if, for each $i \in \{1, 2, \ldots, n\}$, there exist at least t disjoint subsets $R_1(i), R_2(i), \ldots, R_t(i) \subset \{1, 2, \ldots, n\} \setminus \{i\}$ satisfying, for $j = 1, 2, \ldots, t$
  1) $|R_j(i)| \leq r$,
  2) $c_i = \sum_{l \in R_j(i)} a_l c_l$, where $a_l \in \mathbb{F}_q$.
- (r, δ) LRCs [4]: Let C be an $[n, k, d]$ code over $\mathbb{F}_q$ and $\boldsymbol{c} = \{c_1, c_2, \ldots, c_n\}$ be a codeword of C. The code C is said to have locality (r, δ) if, for each $i \in \{1, 2, \ldots, n\}$, there exists a subset $S_i \subset \{1, 2, \ldots, n\}$ satisfying
  1) $i \in S_i$ and $|S_i| \leq r + \delta - 1$,
  2) $d_{min}(C\mid_{S_i}) \geq \delta$,

  where $C\mid_{S_i}$ is the punctured subcode of C by deleting code symbols $c_j, j \in \{1, 2, \ldots, n\} \setminus S_i$.
- Hierarchical LRCs (HLRCs) [5]: Let C be an $[n, k, d]$ code and $\boldsymbol{c} = (c_1, c_2, \ldots, c_n) \in C$. The code C is said to have the hierarchical locality $[(r_1, \delta_1), (r_2, \delta_2)]$ if, for

The associate editor coordinating the review of this manuscript and approving it for publication was Zilong Liu.

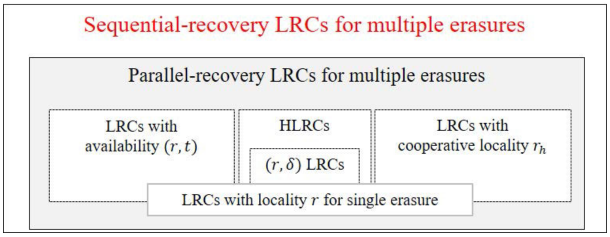

FIGURE 1. Classification of LRCs.

each $i \in \{1, 2, \ldots, n\}$, there exists a punctured subcode C_i such that $c_i \in supp(C_i)$, $dim(C_i) \leq r_1$, $d_{min}(C_i) \geq \delta_1$ and C_i is a (r_2, δ_2) LRCs.

- LRCs with cooperative locality [6]: Let C be an $[n, k, d]$ code and $\boldsymbol{c} = (c_1, c_2, \ldots, c_n) \in C$. The code C is said to have the cooperative locality r_h if, for each $S \subset \{1, 2, \ldots, n\}$ of size h, there exists a subset $R(S)$ of $\{1, 2, \ldots, n\} \setminus S$ satisfying
  1) $|R(S)| \leq r_h$,
  2) for each $i \in S$, $c_i = \sum_{l \in R(S)} a_l c_l$, where $a_l \in \mathbb{F}_q$.

Compared with the sequential-recovery LRCs, parallel-recovery LRCs needs less time to repair all erasures. Various bounds [3]–[8] and constructions [9]–[17] have been proposed for each type of parallel-recovery LRCs, even some are binary codes [9], [10], [14], [16], [17].

The sequential-recovery LRCs have a much more advantage on the erasure tolerance than the parallel-recovery LRCs with the same n and k [18]. The bounds on the code rate and/or the block-length of t-seq LRCs were proposed in [2], [19]–[22]. In [2], binary rate optimal 2-seq LRCs is constructed based on the regular graph. For any t, three classes of t-seq LRCs have been proposed: 1) graph based construction [19]; 2) resolvable configurations based construction [20]; 3) the generalized direct product construction [20]. The first two constructions are rate optimal when $t = 2, 3$, and the last one is rate optimal for any t. Binary 2-seq and 3-seq length optimal LRCs are proposed based on the graph [2], [21].

The upper bound on the minimum distance of 2-seq LRCs was proposed in [2], which is the only bound on the distance of the t-seq LRCs as far as we know. Four explicit constructions of 2-seq LRCs for any r that achieve the upper bound on the minimum distance were also proposed in [2]. Note that $r = 2$ is the most interesting situation in practice. The parameters of all the known distance-optimal 2-seq $[n, k, d, 2]$ LRCs are shown as follows [2].

1) Field size of 2: [6, 3, 3, 2], [8, 4, 3, 2];
2) Field size of $O(r)$: [8, 2, 6, 2];
3) Field size of $O(n^{d/2})$: $[6, k, d, 2]$, $[8, k, d, 2]$;

From the above parameters, we know that the options for the values of n and k are very limited. It would be better if we could have a large family of 2-seq LRCs with a larger minimum distance over a relatively small field size. In this paper, we focus on the linear 2-seq LRCs with $k > r > 1$. Our first contribution of this paper is a construction of 2-seq $[n, k, d, r]$ LRCs for even r, based on some good polynomials, and show several properties. Our second contribution of this paper is the derivation of an explicit upper bound on the minimum distance of a certain class of 2-seq LRCs. We also prove the proposed 2-seq LRCs for $r = 2$ is optimal or near-optimal in the sense of attaining the upper bound on the minimum distance that we derived.

Section II introduces some preliminaries about the good polynomial-based LRCs for single erasure and several bounds of 2-seq LRCs. Section III describes two main contributions of the paper in detail. An open problem for the near future is given in Section IV.

II. PRELIMINARIES

A. LRC BASED ON GOOD POLYNOMIAL

Let $g(x) \in \mathbb{F}_q[x]$ be a polynomial of degree $r + 1$. If there exist t_n disjoint subsets $A_0, A_1, \ldots, A_{t_n-1}$ of $\mathbb{F}_q$, each of size $r + 1$, such that $g(x)$ is constant on each subset, then $g(x)$ is called good [1]. Note that a subset of size j is called a j-subset.

Known Fact 1: (Construction 1 in [1]) *Let r be a positive integer. Let k and n be positive integers with $r|k$, $(r+1)|n$, and $\frac{k}{n} \leq \frac{r}{r+1}$. Let $\mu = \frac{n}{r+1}$ and $A_0, A_1, \ldots, A_{\mu-1}$ be μ disjoint $(r+1)$-subsets of $\mathbb{F}_q$, and let $g(x)$ be a good polynomial with respect to these disjoint subsets. Then, the good polynomial-based $[n, k, d, r]$ LRC C_1 over $\mathbb{F}_q$ is defined as the set of codewords given as follows:*

$$C_1 = \{(f_{\boldsymbol{a}}(\gamma), \gamma \in \cup_{l=0}^{\mu-1} A_l) | \boldsymbol{a} \in \mathbb{F}_q^k\},$$

where $\boldsymbol{a} \in \mathbb{F}_q^k$ is an information vector written as $\boldsymbol{a} = (a_{i,j}, i = 0, 1, \ldots, r-1; j = 0, 1, \ldots, \frac{k}{r} - 1)$, and $f_{\boldsymbol{a}}(x)$ is the encoding polynomial of $\boldsymbol{a}$, given as

$$f_{\boldsymbol{a}}(x) = \sum_{i=0}^{r-1} \sum_{j=0}^{\frac{k}{r}-1} a_{i,j} x^i g(x)^j. \tag{1}$$

The good polynomial-based LRCs is a class of the optimal $[n, k, d, r]$ LRCs for single erasure in the sense that the minimum distance $d = n - k - \frac{k}{r} + 2$ over the field of size $q \approx n$. Suppose the symbol $c_{\gamma(j)} \triangleq f_{\boldsymbol{a}}(\gamma)$ of a codeword is erased, where $\gamma \in A_j$ for some $j = 0, 1, \ldots, \mu - 1$, then its decoding polynomial is given as [1]

$$\delta(x) = \sum_{\beta \in A_j \setminus \gamma} f_{\boldsymbol{a}}(\beta) \prod_{\beta' \in A_j \setminus \{\beta, \gamma\}} \frac{x - \beta'}{\beta - \beta'}. \tag{2}$$

Then, $c_{\gamma(j)} = \delta(\gamma)$.

B. RECURSIVE UPPER BOUND OF 2-SEQ LRCs

The support set of a vector $\boldsymbol{u} = (u_0, u_1, \ldots, u_{n-1})$ is defined as $supp(\boldsymbol{u}) = \{i | u_i \neq 0\}$, and $w(\boldsymbol{u}) = |supp(\boldsymbol{u})|$ is the weight of $\boldsymbol{u}$. The support set of a subcode D of a code C is defined as $supp(D) = \cup_{\boldsymbol{c} \in D} supp(\boldsymbol{c})$. For $i = 1, 2, \ldots, k$, the i^{th} generalized Hamming weight (GHW) of an $[n, k, d]$ linear code C is defined as [2], [23]

$$d_i(C) = \min_{dim(D)=i} |supp(D)|.$$

It is well-known that

$$d = d_1(C) < d_2(C) < \cdots < d_k(C) \le n.$$

The $n-k$ remaining numbers when $d_1(C), d_2(C), \ldots, d_k(C)$ are removed from $\{1, 2, \ldots, n\}$ are called the gap numbers $g_1(C), g_2(C), \ldots, g_{n-k}(C)$, with $g_1(C) < g_2(C) < \cdots < g_{n-k}(C)$. The gap number is also called the gap for simplicity. For a given 2-seq $[n, k, d, r]$ LRC C, its local dual subcode $C_\perp$ is defined as [2]

$$C_\perp = span\left(\{\boldsymbol{c} \in C^\perp \mid |supp(\boldsymbol{c})| \le r+1\}\right),$$

where $C^\perp$ is the dual code of C. It is also called the local dual for simplicity. It is well-known that [2]

$$dim(C_\perp) \ge \left\lceil \frac{2n}{r+2} \right\rceil. \tag{3}$$

The necessary and sufficient condition of the equality in (3) is widely open. It is known that the code becomes rate-optimal when the equality is satisfied [2].

Known Fact 2: (Theorem 9 in [2]) *Let C be a 2-seq $[n, k, d, r]$ LRC and $b \triangleq \lceil \frac{2n}{r+2} \rceil$. Then, the upper bounds on the first b GHWs of $C_\perp$ are given by, for $i = 1, 2, \ldots, b$,*

$$d_i(C_\perp) \le e_i,$$

where e_i can be obtained recursively as follows:

$$e_i = \begin{cases} n, & i = b, \\ e_{i+1} - \left\lceil \frac{2e_{i+1}}{i+1} \right\rceil + (r+1), & i = b-1, \ldots, 2, 1. \end{cases} \tag{4}$$

Furthermore, if there exists a unique integer l such that

$$e_l < k + l < e_{l+1}, \tag{5}$$

then the upper bound on the minimum distance of C is given as

$$d_{min}(C) \le n + 1 - (k + l). \tag{6}$$

The 2-seq LRC C is said to be distance optimal if (6) holds with equality, and is said to be distance near-optimal if $d_{min}(C)$ is one less than the RHS of (6).

Known Fact 3: (Theorem 2 in [2]) Let C be a 2-seq $[n, k, d, r]$ LRC. The upper bound on the rate of C is given by

$$\frac{k}{n} \le \frac{r}{r+2}. \tag{7}$$

We say that the 2-seq LRC C is rate optimal when its code rate achieves the equality in (7).

III. MAIN RESULT

A. NEW CONSTRUCTION OF 2-SEQ LRCs

In this subsection, we propose a construction of 2-seq LRCs with even locality r over $\mathbb{F}_q$, which can be seen as an "extended code" of the good polynomial-based LRCs. Furthermore, we calculate its minimum distance for $r = 2$ and show that it is either optimal or near-optimal in terms of the minimum distance.

Let r be an even integer and q be a prime or a prime power with $r+1 | q-1$. Let t_n and t_k be positive integers with $t_k \le t_n \le \frac{q-1}{r+1}$. For $i = 0, 1, \ldots, t_n - 1$, let $A_i = \{\alpha^i, \alpha^i\beta, \ldots, \alpha^i\beta^r\}$ where α is a primitive element of $\mathbb{F}_q$ and $\beta = \alpha^{\frac{q-1}{r+1}}$. An $[(r+1)t_n, rt_k, d_1, r]$ good polynomial-based LRC C_1 over $\mathbb{F}_q$ can be constructed using the good polynomial $g(x) = x^{r+1}$ with respect to $A_0, A_1, \ldots, A_{t_n-1}$ as described in Known Fact 1 of Subsection II-A. It is well-known that $d_1 = (r+1)(t_n - t_k) + 2$ since C_1 is distance-optimal [1]. For the construction below, any codeword $\boldsymbol{u} \in C_1$ is now written as an array $\boldsymbol{u} = (u_{i,j})$ for $i = 0, 1, \ldots, t_n - 1$ and $j = 0, 1, \ldots, r$, where $u_{i,j} = f_{\boldsymbol{a}}(\alpha^i\beta^j)$ where $\alpha^i\beta^j$ is the j^{th} element of A_i.

Construction 1: We assume the same r (even), q, t_k, and t_n as above for the LRC C_1 over $\mathbb{F}_q$ with parameters $[(r+1)t_n, rt_k, d_1, r]$. The code C_2 is an $[(r+2)t_n, rt_k, d_2, r_2]$ LRC over $\mathbb{F}_q$ with codewords $\boldsymbol{c}$ as the following array

$$\boldsymbol{c} = \begin{pmatrix} c_{0,0} & c_{0,1} & \cdots & c_{0,r} & c_{0,r+1} \\ c_{1,0} & c_{1,1} & \cdots & c_{1,r} & c_{0,r+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_{t_n-1,0} & c_{t_n-1,1} & \cdots & c_{t_n-1,r} & c_{t_n-1,r+1} \end{pmatrix}$$

where, for $i = 0, 1, \ldots, t_n - 1$,

$$c_{i,j} = \begin{cases} u_{i,j}, & j = 0, 1, \ldots, r, \\ \sum_{l=0}^{r-1} u_{i,l}, & j = r+1. \end{cases} \tag{8}$$

Remark 1: The construction will work if $c_{i,r+1}$ is the sum of ANY r symbols among $u_{i,0}, u_{i,1}, \ldots, u_{i,r}$. We call this an "overall" parity bit.

Theorem 1: The code C_2 from Construction 1 is a 2-seq $[n = (r+2)t_n, rt_k, d_2, r_2]$ LRC over $\mathbb{F}_q$ with $d_2 \ge d_1$, $r_2 = r$ and $q \approx n\frac{r+1}{r+2}$.

Proof: The "overall" parity bit $c_{i,r+1}$ will not decrease its minimum distance, and hence, $d_2 \ge d_1$. For C_1 we have $q \approx (r+1)t_n$. The same value of q is used for C_2 with $t_n = n/(r+2)$. Therefore, $q \approx n\frac{r+1}{r+2}$.

We now prove that any 2 erasures of C_2 can be repaired locally and sequentially. We write any codeword of C_2 as an array $\boldsymbol{c} = (c_{i,j})$ for $i = 0, 1, \ldots, t_n - 1$ and $j = 0, 1, \ldots, r+1$. That is, we may view the codeword written as a matrix of size $t_n \times (r+2)$. When two erasures occur in two distinct rows, each erasure can be repaired in any order one by one individually because they belong to different, and hence, disjoint repair sets. We now consider the case with two erasures. Without loss of generality, assume that these two erasures belong to the top row of a codeword of C_2. We will denote $c_j \triangleq c_{0,j}$ for $0 \le j \le r+1$ for simplicity and convenience. Assume c_x and c_y are two erasures, and defined by e_x and e_y, respectively. We will distinguish the following two cases: 1) $0 \le x \le r$, $y = r+1$ and 2) $0 \le x < y \le r$.

The case 1) is easy since e_x can be repaired by other r symbols from (2) first and then e_y is the sum of the first r symbols.

The case 2) has two subcases: $y = r$ and $y < r$. When $y = r$, the sequential recovery is easy since e_x can be repaired by r symbols first as

$$e_x = c_{r+1} - \sum_{\substack{l=0 \\ l \neq x}}^{r-1} c_l, \tag{9}$$

and then e_y can be repaired by r symbols from (2). We now consider the case where $0 \leq x < y < r$ in the following, which is repaired by the remaining r unerased symbols in the top row. Using the decoding polynomial (2) for the code C_1, we have

$$e_x = \delta(\beta^x) \quad \text{and} \quad e_y = \delta(\beta^y),$$

where the polynomial $\delta(\beta^j)$ is determined as

$$\delta(\beta^j) = \sum_{\substack{l=0 \\ l \neq j}}^{r} c_l \cdot \prod_{\substack{\tau=0 \\ \tau \neq j,l}}^{r} \frac{\beta^j - \beta^\tau}{\beta^l - \beta^\tau}.$$

Adding these two relations, we have

$$e_x + e_y = \delta(\beta^x) + \delta(\beta^y). \tag{10}$$

Using the relation of C_2 in (8), we have

$$e_x + e_y + \sum_{\substack{0 \leq j \leq r-1 \\ j \neq x,y}} c_j = c_{r+1}. \tag{11}$$

It is now enough to show that two equations (10) and (11) in two unknowns e_x and e_y have a unique solution. The first equation (10) can be written as

$$(1 - \prod_{\substack{\tau=0 \\ \tau \neq x,y}}^{r} \frac{\beta^y - \beta^\tau}{\beta^x - \beta^\tau})e_x + (1 - \prod_{\substack{\tau=0 \\ \tau \neq x,y}}^{r} \frac{\beta^x - \beta^\tau}{\beta^y - \beta^\tau})e_y = u_1$$

for some $u_1 \in \mathbb{F}_q$. Similarly, (11) can be written as

$$e_x + e_y = u_2$$

for some $u_2 \in \mathbb{F}_q$. Some simple row-operations give the following:

$$\begin{cases} e_x + e_y = u_2 \\ e_x \prod_{\substack{\tau=0 \\ \tau \neq x,y}}^{r} \frac{\beta^y - \beta^\tau}{\beta^x - \beta^\tau} + e_y \prod_{\substack{\tau=0 \\ \tau \neq x,y}}^{r} \frac{\beta^x - \beta^\tau}{\beta^y - \beta^\tau} = u_3 \end{cases} \tag{12}$$

for some constants u_2 and $u_3 \in \mathbb{F}_q$. This equation will have a unique solution if the coefficient matrix is non-singular, or

$$\left(\prod_{\substack{\tau=0 \\ \tau \neq x,y}}^{r} \frac{\beta^y - \beta^\tau}{\beta^x - \beta^\tau} \right)^2 \neq 1.$$

It is straightforward to show that, for any $0 \leq x < y \leq r$, we have

$$\prod_{\substack{\tau=0 \\ \tau \neq x,y}}^{r} \frac{\beta^y - \beta^\tau}{\beta^x - \beta^\tau} = \pm \beta^m,$$

for $(r+1) \nmid m$, and the result is also an element of A_0, since A_0 is a multiplicative subgroup. Then we can get that

$$\left(\prod_{\substack{\tau=0 \\ \tau \neq x,y}}^{r} \frac{\beta^y - \beta^\tau}{\beta^x - \beta^\tau} \right)^2 = (\pm \beta^m)^2 = \beta^{2m} \neq 1, \tag{13}$$

since $2m \not\equiv 0 \pmod{r+1}$ for even r. Therefore, (12) has a unique solution. ■

From Known Fact 2, we have the following condition for the rate optimality of the proposed LRCs.

Corollary 1: The 2-seq LRC C_2 over $\mathbb{F}_q$ in Theorem 1 is rate optimal when $t_k = t_n$.

Example 1: Let $r = 2$, $q = 13$ and hence, $\frac{q-1}{r+1} = 4$. Choose $t_k = 3$ and $t_n = 4$ for a 2-seq LRC C_2 over $\mathbb{F}_{13}$ with parameters $[16, 6, d_2 = 6, 2]$. *We will prove shortly that $d_2 = 4(t_n - t_k) + 2 = 6$ when $t_k < t_n$ and $r = 2$ (Theorem 2). We first construct a* $[12, 6, 5, 2]$ *good polynomial-based LRC C_1 over $\mathbb{F}_{13}$ using $A_0 = \{1, 3, 9\}$, $A_1 = \{2, 6, 5\}$, $A_2 = \{4, 12, 10\}$, $A_3 = \{8, 11, 7\}$ where we use $\alpha = 2$ and $\beta = \alpha^4 = 3$. For any information $\boldsymbol{a}$, the corresponding codeword $\boldsymbol{c}$ of C_2 (as an array) becomes*

$$\boldsymbol{c} = \begin{pmatrix} f_{\boldsymbol{a}}(1) & f_{\boldsymbol{a}}(3) & f_{\boldsymbol{a}}(9) & f_{\boldsymbol{a}}(1) + f_{\boldsymbol{a}}(3) \\ f_{\boldsymbol{a}}(2) & f_{\boldsymbol{a}}(6) & f_{\boldsymbol{a}}(5) & f_{\boldsymbol{a}}(2) + f_{\boldsymbol{a}}(6) \\ f_{\boldsymbol{a}}(4) & f_{\boldsymbol{a}}(12) & f_{\boldsymbol{a}}(10) & f_{\boldsymbol{a}}(4) + f_{\boldsymbol{a}}(12) \\ f_{\boldsymbol{a}}(8) & f_{\boldsymbol{a}}(11) & f_{\boldsymbol{a}}(7) & f_{\boldsymbol{a}}(8) + f_{\boldsymbol{a}}(11) \end{pmatrix}.$$

Theorem 2: Consider the 2-seq $[4t_n, 2t_k, d_2, 2]$ *LRC C_2 over $\mathbb{F}_q$ in Theorem 2 with $r = 2$. The minimum distance d_2 of C_2 is given by*

$$d_2 = \begin{cases} 4(t_n - t_k) + 2, & t_k < t_n, \\ 3, & t_k = t_n. \end{cases}$$

We will provide the proof in Appendix A at the end for the readability and convenience.

B. EXPLICIT UPPER BOUND ON THE MINIMUM DISTANCE

In this subsection, we derive an explicit form of the upper bound on the minimum distance of C, and show in Corollary 3 that the proposed 2-seq LRCs with $r = 2$ is distance near-optimal when $t_k < t_n$, and is distance optimal when $t_k = t_n$.

Lemma 1: Let C be a 2-seq $[n, k, d, r]$ LRC and $C_\perp$ be its local dual. Denote by $d_i(C_\perp)$ the i^{th} GHW of $C_\perp$ as defined in Subsection II-B. *When $n = (r+2)t$ for some t and $(r-1) \mid 2(t-1)$, the upper bound of $d_i(C_\perp)$ is given as, for $i = 1, 2, \ldots, 2t$, with $h \triangleq \frac{2(t-1)}{r-1}$,*

$$d_i(C_\perp) \leq (r+1)i - h \sum_{j=1}^{\lfloor \frac{i-1}{h} \rfloor} j - \left\lceil \frac{i-1}{h} \right\rceil \left(i - 1 - h \cdot \left\lfloor \frac{i-1}{h} \right\rfloor \right). \tag{14}$$

Proof: Observe that $\lceil \frac{2n}{r+2} \rceil = 2t$. Denote by ψ_i the RHS of (14), and we will prove that ψ_i satisfies the same recursion of e_i in (4) for $i = 1, 2, \ldots, 2t$. We will distinguish the cases where $h > 1$ and $h = 1$.

TABLE 1. Comparison of the various parameters of the 2-seq LRCs.

Reference	Parameters					Note	Property
	n	k	r	q	d		
[2]	$\frac{(r+\beta)(r+2)}{2}$	any k	any r	$\mathcal{O}(n^{\frac{d_{min}-1}{2}})$	d_{min}	$1\le\beta\le r,\ \beta\|r$	Distance optimal
	$\frac{(r+\beta)(r+2)}{2}$	$\frac{(r+\beta)r}{2}-2$	any r	$\mathcal{O}(r)$	6	$1\le\beta\le r,\ \beta\|r$	Distance optimal
	$r(r+2)$	r^2-3	any r	$\mathcal{O}(n)$	8		Distance optimal
	$\frac{(r+2)k}{r}$	$\ge\frac{r(r+1)}{2}$	any r	2			Rate optimal
[20]	$(1+\frac{2}{r})r^m$	r^m	any r	2		$m\in\mathbb{N}^+$	Rate optimal
[21]	$\frac{(r+2)k}{r}$	$\lceil\frac{k}{r}\rceil\ge r$	any r	2			Rate optimal
This paper	$4t_n$	$2t_k$	2	$q\approx\frac{3n}{4}$	$4(t_n-t_k)+2$	$t_k<t_n$	Distance near-optimal
	$4t_n$	$2t_k$	2	$q\approx\frac{3n}{4}$	3	$t_k=t_n$	Distance optimal
	$(r+2)t_n$	rt_k	even r	$q\approx\frac{(r+1)n}{r+2}$		$t_k=t_n$	Rate optimal

Assume $h=\frac{2(t-1)}{r-1}>1$. When $i=2t$, we have $2t-1=2(t-1)+1$. Therefore, $\frac{2t-1}{h}=r-1+\frac{1}{h}$. Now, it is straightforward to check that $\psi_{2t}=t(r+2)=n$. For any integer $m=1,2,\ldots,2t-1$, assume that ψ_{m+1} is given as the RHS of (14) for $i=m+1$. Now, we will check that $\psi_{m+1}-\lceil\frac{2\psi_{m+1}}{m+1}\rceil+(r+1)$ becomes ψ_m. Let δ and ϵ be positive integers, such that $m=\delta h+\epsilon$ and $0\le\epsilon<h$. When $1<\epsilon<h$, $\lfloor\frac{m}{h}\rfloor=\lfloor\frac{m-1}{h}\rfloor=\delta$ and $\lceil\frac{m}{h}\rceil=\lceil\frac{m-1}{h}\rceil=\delta+1$. Therefore, we have

$$\frac{2\psi_{m+1}}{m+1}=2(r+1)-(\delta+1)-\frac{\delta\epsilon+\epsilon-(\delta+1)}{\delta h+\epsilon+1},$$
$$\left\lceil\frac{2\psi_{m+1}}{m+1}\right\rceil=2(r+1)-\left\lceil\frac{m}{h}\right\rceil,$$

and

$$\psi_{m+1}-\left\lceil\frac{2\psi_{m+1}}{m+1}\right\rceil+(r+1)=(r+1)m-h\sum_{j=1}^{\lfloor\frac{m-1}{h}\rfloor}j-\left\lceil\frac{m-1}{h}\right\rceil\left(m-1-h\cdot\left\lfloor\frac{m-1}{h}\right\rfloor\right)=\psi_m.$$

It is not difficult to check the same recursion is satisfied when $\epsilon=0$ or $\epsilon=1$.

For $h=1$, the recursive relationship of ψ_m can be proved similarly. ■

LRCs with $r=k$ can be seen as the maximum distance separable code, so we do not consider this case in this paper. Further, if $r(r-1)\mid 2(t-1)$ then $(r-1)\mid 2(t-1)$. Therefore, the explicit upper bound on the minimum distance of 2-seq $[n,k,d,r]$ LRC C with $r+2|n$ and $r<k$ can be derived based on the above explicit upper bound on the GHW of its local dual $C_\perp$.

Theorem 3: Let C be a 2-seq $[n,k,d,r]$ LRC. When $n=(r+2)t$ for some t and $r(r-1)\mid 2(t-1)$, the upper bound on the minimum distance d of C is explicitly given as

$$d\le n+1-k-\left(jh+\lceil\frac{k-\Gamma}{r-(j+1)}\rceil\right), \tag{15}$$

where $h=\frac{2(t-1)}{r-1}$ and j is the largest nonnegative integer that satisfies $k>h\sum_{a=1}^{j}(r-a)+r\triangleq\Gamma$.

Proof: From (5) and (6), it is enough to show that $l\triangleq jh+\lceil\frac{k-\Gamma}{r-(j+1)}\rceil$ satisfies that $\psi_l<k+l<\psi_{l+1}$. Let j be the largest nonnegative integer that satisfies $k>\Gamma$. $\Gamma<k\le\Gamma+h(r-(j+1))$ and hence

$$0<\lceil\frac{k-\Gamma}{r-(j+1)}\rceil\le h.$$

Let $m\triangleq\lceil\frac{k-\Gamma}{r-(j+1)}\rceil$. $\Gamma+(m-1)(r-(j+1))<k\le\Gamma+m(r-(j+1))$. By (14), it is staightforward to get

$$\psi_{jh+m}=h\sum_{s=1}^{j}(r-s)+r+(m-1)(r-(j+1))+jh+m<k+jh+m.$$

Similarly, we can get $\psi_{jh+m+1}>k+jh+m$. ■

Corollary 2: For the 2-seq LRC C in Theorem 3, when $r=2$ and $k=2t_k$ for some $t_k\le t$, we have,

$$d\le 4(t-t_k)+3.$$

Proof: We note that $h=2(t-1), j=0$ and $r=2$ hence $\Gamma=2$. ■

Corollary 3: The 2-seq LRC C_2 over $\mathbb{F}_q$ in Theorem 2 is distance near-optimal when $t_k<t_n$, and is distance optimal when $t_k=t_n$.

IV. CONCLUDING REMARK

This paper constructed the near-optimal 2-seq $[(r+2)t_n,rt_k,d,r]$ LRCs for even r over a relatively small alphabet of size $q\approx\frac{(r+1)n}{r+2}$, where $t_k\le t_n\le\frac{q-1}{r+1}$. For comparison, we show the various parameters of the 2-seq LRCs in Table 1. The proposed 2-seq LRCs is rate optimal or distance optimal or distance near-optimal for some cases. In the future, it may be important to find a construction for the optimal 2-seq LRC with any locality $r \geq 2$ over a smaller alphabet.

.

APPENDIX A THE PROOF OF THEOREM 2

We will fix the notations for C_2 in Theorem 2 and hence the corresponding C_1 also. Any codeword $\boldsymbol{c} \in C_2$ in the array representation of size $t_n \times (r+2)$ consists of the codeword $\boldsymbol{u} \in C_1$ of size $t_n \times (r+1)$ on the left and the right-most column of length t_n. Here, we write $u_i = (u_{i,0}, u_{i,1}, u_{i,2})$ as i^{th} row of $\boldsymbol{u}$ and $c_{i,3} = u_{i,0} + u_{i,1}$ for $i = 0, 1, \ldots, t_n - 1$. Therefore, we may write, as an array,

$$\boldsymbol{c} = (\boldsymbol{u}|\boldsymbol{u}_{add}) \tag{16}$$

and

$$w(\boldsymbol{c}) = w(\boldsymbol{u}) + w(\boldsymbol{u}_{add}), \tag{17}$$

where $\boldsymbol{u}_{add}$ is the last column of $\boldsymbol{c}$ which consists of the "overall" parity bits.

We first take a look at the encoding polynomial of element $u_{i,j}$ of $\boldsymbol{u} \in C_1$. From (1) with $r = 2$,

$$u_{i,j} = f_{\boldsymbol{a}}(\alpha^i \beta^j) = \left(\sum_{l=0}^{t_k-1} a_{0,l} g(x)^l \right) + x \left(\sum_{l=0}^{t_k-1} a_{1,l} g(x)^l \right) \Bigg|_{x=\alpha^i\beta^j}.$$

We note that $g(\alpha^i\beta^j) = \alpha^{3i}$ for all $j = 0, 1, 2$. Therefore, for each $i = 0, 1, \ldots, t_n - 1$, the encoding polynomial $f_{\boldsymbol{a},i}(x)$ of the i^{th} row u_i is given as

$$f_{\boldsymbol{a},i}(x) = H_i + B_i x, \tag{18}$$

where

$$H_i = \sum_{l=0}^{t_k-1} a_{0,l}\alpha^{3il}, \quad B_i = \sum_{l=0}^{t_k-1} a_{1,l}\alpha^{3il}.$$

We will denote by τ_m the number of rows of $\boldsymbol{u}$ with weight m in $\boldsymbol{u}$ for $m = 0, 1, 2, 3$.

Lemma 2: Consider C_1 and C_2 in Theorem 2 and assume all the notation in the discussion after Theorem 2, leading to $f_{\boldsymbol{a},i}(x)$ in (18). Let $\boldsymbol{a} \in \mathbb{F}_q^{2t_k}$ be a non-zero information vector, and let $\boldsymbol{u}$ be its codeword of C_1. Let $\boldsymbol{u}_{add}$ be the last column of the corresponding codeword $\boldsymbol{c} \in C_2$. Then, the following holds:

1) *For each $i = 0, 1, \ldots, t_n - 1$,*
   a) *$w(u_i) = 0$ if and only if $H_i = 0$ and $B_i = 0$. It is obvious that $u_{i,0} + u_{i,1} = 0$ if $w(u_i) = 0$;*
   b) *$w(u_i) = 1$ is impossible; hence, $\tau_1 = 0$ for any $\boldsymbol{u}$.*
   c) *$w(u_i) = 2$ if and only if $-\frac{H_i}{B_i} \in A_i$ with $B_i \neq 0$. If $w(u_i) = 2$, then $u_{i,0} + u_{i,1} \neq 0$;*
   d) *$w(u_i) = 3$ if and only if $H_i \neq 0$ (when $B_i = 0$) or $-\frac{H_i}{B_i} \notin A_i$ (when $B_i \neq 0$). For a row u_i of weight 3, $u_{i,0} + u_{i,1} = 0$ if and only if $2H_i - \alpha^i\beta^2 B_i = 0$.*
2) *In any $\boldsymbol{u}$, the number of rows with $w(u_i) = 3$ and $u_{i,0} + u_{i,1} = 0$ is at most $\min(\tau_3, 2t_k - 2\tau_0 - 1)$.*
3) *For any $\boldsymbol{c}$, $w(\boldsymbol{u}_{add}) \geq \tau_2 + \tau_3 - \min(\tau_3, 2t_k - 2\tau_0 - 1)$.*

Proof: We recall that $|A_i| = 3$ for all i. We skip the proof of a) of Case 1). For the subcase b), we note that $w(u_i) = 1$ implies $f_{\boldsymbol{a},i}(x)$ in (18) of degree at most 1 must have two roots. For the subcase c), we note that $w(u_i) = 2$ if and only if $u_{i,j} = H_i + B_i\alpha^i\beta^j = 0$ or equivalently, $-\frac{H_i}{B_i} = \alpha^i\beta^j \in A_i$ for some j. For the second assertion, we assume that $w(u_i) = 2$ for some i. Then, $-\frac{H_i}{B_i} \in A_i$ implies $H_i \neq 0$. Suppose that $u_{i,0} + u_{i,1} = 0$ on the contrary. Then, $u_{i,0}, u_{i,1}$ must be both non-zero and $u_{i,2} = 0$, and hence, $u_{i,0} + u_{i,1} + u_{i,2} = 0$, which contradicts to the following:

$$\begin{aligned} u_{i,0} + u_{i,1} + u_{i,2} &= f_{\boldsymbol{a},i}(\alpha^i) + f_{\boldsymbol{a},i}(\alpha^i\beta) + f_{\boldsymbol{a},i}(\alpha^i\beta^2) \\ &= 3H_i + B_i\alpha^i(1 + \beta + \beta^2) = 3H_i. \end{aligned}$$

For the subcase d), we note that $w(u_i) = 3$ if and only if B_i and H_i satisfy the condition which is the complement of the union of the previous cases. For the second assertion, we observe first that $u_{i,2} = H_i + B_i\alpha^i\beta^2$. Then, $u_{i,0} + u_{i,1} = 0 \Leftrightarrow u_{i,0} + u_{i,1} + u_{i,2} = u_{i,2} \Leftrightarrow 3H_i = H_i + B_i\alpha^i\beta^2 \Leftrightarrow 2H_i - B_i\alpha^i\beta^2 = 0$.

For the proof of Case 2), we observe the following: the codeword $\boldsymbol{u}$ has τ_0 rows of weight 0 if and only if there exists $2\tau_0$ linear relations on the elements of the information $\boldsymbol{a}$ as

$$H_i = 0 \text{ and } B_i = 0, \tag{19}$$

for some τ_0 values of i. This is equivalent to saying that some $2\tau_0$ elements in such $\boldsymbol{a}$ are linear combinations of the remaining $2t_k - 2\tau_0$ elements. When $2t_k - 2\tau_0 < \tau_3$, therefore, the number of additional linear dependencies of elements in such $\boldsymbol{a}$ can be at most $2t_k - 2\tau_0 - 1$. The necessary and sufficient condition

$$2H_i - \alpha^i\beta^2 B_i = 0, \tag{20}$$

for $u_{i,0} + u_{i,1} = 0$, is also a linear dependency in $\boldsymbol{a}$. Therefore, the number of rows u_i with weight 3 satisfying $2H_i - \alpha^i\beta^2 B_i = 0$ can be at most $2t_k - 2\tau_0 - 1$. When $2t_k - 2\tau_0 \geq \tau_3$ on the other hand, it is obvious that the number of rows u_i with weight 3 satisfying $2H_i - \alpha^i\beta^2 B_i = 0$ can be at most τ_3.

Case 3) comes easily from the previous cases. ■

Remark 2: The equality of Case 3) in Lemma 2 holds when, for information $\boldsymbol{a} \in \mathbb{F}_q^{2t_k}$, there is only one choice of freedom in $\boldsymbol{a}$ and the remaining $2t_k - 1$ elements of $\boldsymbol{a}$ are decided by $2\tau_0$ equations of the type in (19) and $2t_k - 2\tau_0 - 1$ equations of the type in (20).

Now, we continue the proof of Theorem 2. Let $\boldsymbol{u}, \boldsymbol{u}' \in C_1$ be codewords with the same weight such that $2\tau_2 + 3\tau_3 = 2\tau_2' + 3\tau_3'$. If $\tau_2 > \tau_2'$ then $\tau_2 - \tau_2' = 3l$, $\tau_0 - \tau_0' = -l$ and $\tau_3 - \tau_3' = -2l$, for some positive integer l, and hence, we have $w(\boldsymbol{u}_{add}) > w(\boldsymbol{u}'_{add})$ by Lemma 2, which implies that $w(\boldsymbol{c}) > w(\boldsymbol{c}')$.

We now claim that

$$w(\boldsymbol{c}) \geq 4(t_n - t_k) + 3 - \min(t_n - t_k, 1) \triangleq w_{min} \tag{21}$$

for any $\boldsymbol{c} \in C_2$. Observe that it is enough to prove (21) for all the codewords $\boldsymbol{c}$ corresponding to $\boldsymbol{u} \in C_1$ with $\tau_2 \in \{0, 1, 2\}$. We will prove this by induction on the weight of $\boldsymbol{u}$.

When $\boldsymbol{u} \in C_1$ be a non-zero codeword with the minimum weight $d_1 = 3(t_n - t_k) + 2$, we can get $\tau_3 = t_n - t_k$, $\tau_2 = 1$, and $\tau_0 = t_k - 1$. So, the weight of the corresponding codeword $\boldsymbol{c}$ is, from 3) of Lemma 2,

$$\begin{aligned} w(\boldsymbol{c}) &\geq 3(t_n - t_k) + 2 \\ &\quad + (\tau_2 + \tau_3 - \min(\tau_3, 2t_k - 2\tau_0 - 1)) \\ &= 4(t_n - t_k) + 3 - \min(t_n - t_k, 1). \end{aligned}$$

The weight of codeword $\boldsymbol{c}$ achieves the bound when the corresponding information $\boldsymbol{a}$ satisfies the condition in Remark 2, which obviously exists in $\mathbb{F}_q^{2t_k}$. Let $\boldsymbol{u} \in C_1$ have $w(\boldsymbol{u}) > d_1$ and $\boldsymbol{c} \in C_2$ be its corresponding codeword. If there exists a codeword $\boldsymbol{u'} \in C_1$ with $w(\boldsymbol{u'}) = w(\boldsymbol{u}) - 1$, then

$$\begin{aligned} 2\tau_2' + 3\tau_3' &= 2\tau_2 + 3\tau_3 - 1 \\ &\Rightarrow 2(\tau_2 - \tau_2') = 3(\tau_3' - \tau_3) + 1 \\ &\Rightarrow \tau_2 - \tau_2' = 2 \text{ or } -1. \end{aligned}$$

If $\tau_2 - \tau_2' = 2$ (the other case can be treated similarly), then $\tau_3 - \tau_3' = \tau_0 - \tau_0' = -1$. In this case, $w(\boldsymbol{u}_{add}) > w(\boldsymbol{u}'_{add})$. Therefore, $w(\boldsymbol{u}) + w(\boldsymbol{u}_{add}) > w(\boldsymbol{u}') + w(\boldsymbol{u}'_{add})$, which implies $w(\boldsymbol{c}) > w(\boldsymbol{c}') \geq w_{min}$ by induction. If there exists a codeword $\boldsymbol{u'} \in C_1$ with $w(\boldsymbol{u'}) \leq w(\boldsymbol{u}) - 2$, we can prove similarly.

REFERENCES

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

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

[3] A. Wang and Z. Zhang, "Repair locality with multiple erasure tolerance," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 6979–6987, Nov. 2014.

[4] N. Prakash, G. M. Kamath, V. Lalitha, and P. V. Kumar, "Optimal linear codes with a local-error-correction property," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2012, pp. 2776–2780.

[5] B. Sasidharan, G. K. Agarwal, and P. V. Kumar, "Codes with hierarchical locality," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2015, pp. 1257–1261.

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

[7] J. Hao, K. W. Shum, S.-T. Xia, and Y.-X. Yang, "Classification of optimal ternary (r, δ)-locally repairable codes attaining the singleton-like bound," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2019, pp. 2828–2832.

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

[9] J. Hao and S.-T. Xia, "Constructions of optimal binary locally repairable codes with multiple repair groups," *IEEE Commun. Lett.*, vol. 20, no. 6, pp. 1060–1063, Jun. 2016.

[10] J.-H. Kim and H.-Y. Song, "Hypergraph-based binary locally repairable codes with availability," *IEEE Commun. Lett.*, vol. 21, no. 11, pp. 2332–2335, Nov. 2017.

[11] Z. Sun, S. Zhu, and L. Wang, "Optimal constacyclic locally repairable codes," *IEEE Commun. Lett.*, vol. 23, no. 2, pp. 206–209, Feb. 2019.

[12] B. Chen, W. Fang, S.-T. Xia, and F.-W. Fu, "Constructions of optimal (r, δ) locally repairable codes via constacyclic codes," *IEEE Trans. Commun.*, vol. 67, no. 8, pp. 5253–5263, Aug. 2019.

[13] H. Cai, Y. Miao, M. Schwartz, and X. Tang, "On optimal locally repairable codes with multiple disjoint repair sets," *IEEE Trans. Inf. Theory*, vol. 66, no. 4, pp. 2402–2416, Apr. 2020.

[14] Y.-S. Su, "On constructions of a class of binary locally repairable codes with multiple repair groups," *IEEE Access*, vol. 5, pp. 3524–3528, 2017.

[15] J. Hao and B. Chen, "On the generalized Hamming weights of (r, δ)-locally repairable codes," *IEEE Access*, vol. 8, pp. 149706–149713, 2020.

[16] J. Teng and L. Jin, "Constructions of binary locally repairable codes with multiple recovering sets," *IEEE Access*, vol. 9, pp. 92239–92245, 2021.

[17] J.-H. Kim, M. K. Song, and H.-Y. Song, "Block-punctured binary simplex codes for local and parallel repair in distributed storage systems," *IEICE Trans. Fundam. Electron., Commun. Comput. Sci.*, vol. E101.A, no. 12, pp. 2374–2381, Dec. 2018.

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

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

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

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

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

[23] V. K. Wei, "Generalized Hamming weights for linear codes," *IEEE Trans. Inf. Theory*, vol. 37, no. 5, pp. 1412–1418, Sep. 1991.

ZHI JING (Student Member, IEEE) received the B.S. degree in information engineering from the China University of Geoscience, Beijing, China, in 2014, and the M.S. degree in information technology management from Hong Kong Baptist University, Hong Kong, China, in 2015. She is currently pursuing the Ph.D. degree with the Channel Coding Laboratory, Yonsei University, Seoul, South Korea. Her research interests include digital communications and coding theory and techniques for distributed storage and/or mobile communications.

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