Designs, Codes and Cryptography (2023) 91:83–114
https://doi.org/10.1007/s10623-022-01099-z

The minimum locality of linear codes

Pan Tan[1] · Cuiling Fan[2] · Cunsheng Ding[3] · Chunming Tang[4] · Zhengchun Zhou[2]

Received: 9 October 2021 / Revised: 27 July 2022 / Accepted: 30 July 2022 /
Published online: 24 August 2022

Abstract
Locally recoverable codes (LRCs) were proposed for the recovery of data in distributed and cloud storage systems about nine years ago. A lot of progress on the study of LRCs has been made by now. However, there is a lack of general theory on the minimum locality of linear codes. In addition, the minimum locality of many known families of linear codes has not been studied in the literature. Motivated by these two facts, this paper develops some general theory about the minimum locality of linear codes, and investigates the minimum locality of a number of families of linear codes, such as q-ary Hamming codes, q-ary Simplex codes, generalized Reed-Muller codes, ovoid codes, maximum arc codes, the extended hyperoval codes, and near MDS codes. Many classes of both distance-optimal and dimension-optimal LRCs are presented in this paper. To this end, the concepts of linear locality and minimum linear locality are specified. The minimum linear locality of many families of linear codes are settled with the general theory developed in this paper.

Communicated by D. Panario.

✉ Cuiling Fan
fcl@swjtu.edu.cn

Pan Tan
ptan@ynnu.edu.cn

Cunsheng Ding
cding@ust.hk

Chunming Tang
tangchunmingmath@163.com

Zhengchun Zhou
zzc@home.swjtu.edu.cn

1 School of Information Science and Technology, Yunnan Normal University, Kunming 650500, China

2 School of Mathematics, Southwest Jiaotong University, Chengdu 611756, China

3 Department of Computer Science and Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China

4 School of Mathematics and Information, China West Normal University, Sichuan Nanchong 637002, China

Keywords Cyclic code · Linear code · Locally recoverable code · Near MDS code · Punctured code · Shortened code

Mathematics Subject Classification 94B05 · 94B15 · 94B25 · 05B05

1 Introduction of motivations, objectives and methodology

We first fix some notation and definitions that will be used throughout this paper.

- Let n be a positive integer and let q be a prime power.
- An $[n, k, d]$ *code* $\mathcal{C}$ over GF(q) is a k-dimensional subspace of GF$(q)^n$ with Hamming distance d.
- $A_i(\mathcal{C})$ or A_i, $\dim(\mathcal{C})$, $d(\mathcal{C})$ and $\mathcal{C}^\perp$ denote the number of codewords of Hamming weight i in $\mathcal{C}$, the dimension of $\mathcal{C}$, the minimum Hamming distance of $\mathcal{C}$, and the dual of $\mathcal{C}$, respectively.
- The *weight distribution* and *weight enumerator* of $\mathcal{C}$ are defined by the sequence $(A_0, \ldots, A_n)$ and the polynomial $\sum_{i=0}^{n} A_i z^i$, respectively.
- $\mathcal{C}$ is said to be a *t-weight code* if the sequence $(A_1, \ldots, A_n)$ has Hamming weight t.
- Denote $[n] = \{0, 1, \ldots, n-1\}$ for each positive integer n. We usually index the coordinates of the codewords in a linear code $\mathcal{C}$ of length n with the elements in $[n]$.
- An $[n, k, d]$ code $\mathcal{C}$ over GF(q) is called an $(n, k, d, q; r)$*-LRC* (i.e., locally recoverable code) if for each $i \in [n]$ there is a subset $R_i \subseteq [n] \setminus \{i\}$ of size r and a function $f_i(x_1, \ldots, x_r)$ on GF$(q)^r$ such that $c_i = f_i(\mathbf{c}_{R_i})$ for each codeword $\mathbf{c} = (c_0, \ldots, c_{n-1})$ in $\mathcal{C}$, where $\mathbf{c}_{R_i}$ is the projection of $\mathbf{c}$ at R_i. The symbol c_i is called the i-th *code symbol* and the set R_i is called the *repair set* or *recovering set* of the code symbol c_i. The minimum r such that $\mathcal{C}$ is an $(n, k, d, q; r)$-LRC is called the *minimum locality* of $\mathcal{C}$.

If the i-th coordinate c_i of each codeword $\mathbf{c}$ in a linear code $\mathcal{C}$ of length n is zero, we say that the i-th coordinate of $\mathcal{C}$ is zero. It is easily seen that a linear code $\mathcal{C}$ has a zero coordinate if and only if the dual distance $d(\mathcal{C}^\perp) = 1$. A linear code $\mathcal{C}$ is called *nontrivial* if $d(\mathcal{C}) \geq 2$ and $d(\mathcal{C}^\perp) \geq 2$ and *trivial* otherwise. In this paper, we consider only nontrivial linear codes, as trivial linear codes are not interesting for error correction. For each nontrivial linear code of length n over GF(q), since $d(\mathcal{C}^\perp) \geq 2$, it is easily seen that $\mathcal{C}$ is an $(n, k, d, q; r)$-LRC for some r with $1 \leq r \leq n$. Consequently, each nontrivial linear code $\mathcal{C}$ has a minimum locality.

In this definition of LRCs above, the degrees of the functions f_i are not restricted. If we require that each f_i be a homogeneous function of degree 1 in the definition above, then we say that $\mathcal{C}$ is $(n, k, d, q; r)$-LLRC (linearly local recoverable code) and has linear locality r. For the same reasons, each nontrivial linear code $\mathcal{C}$ has a minimum linear locality.

By definition, if a linear code has linear locality r, it has locality r. Hence, it is necessary to study the linear locality of linear codes. It will be proved in Sect. 3.1 that the minimum locality equals the minimum linear locality (see Lemma 1). This equality may be known to some people, but a reference pointing out this equality is missing in the literature. Although the minimum locality and minimum linear locality of any nontrivial linear code are identical, the complexity of recovering a code symbol c_i with a nonlinear function $f_i(x_1, \ldots, x_r)$ and a recovering set R_i may be more than that with a linear function $f_i'(x_1, \ldots, x_r) = a_1 x_1 + \cdots + a_r x_r$ and a recovering set R_i'. Hence, it would be better to distinguish the linear locality from the locality and study the linear locality and minimum linear locality of linear codes. In fact, the locality of the linear codes obtained in the literature is actually a linear locality. These facts show the necessity of studying the linear locality and minimum linear

locality of linear codes. It will be further justified in Sect. 3.1 that it is necessary and sufficient to study the linear locality and minimum linear locality of linear codes over finite fields. The following question is then fundamental.

Question 1 What is the minimum linear locality and how does one compute the minimum linear locality of a nontrivial linear code?

In the literature it was shown that many classes of linear codes have linear locality r for some r. However, the minimum linear locality of such linear codes is unknown, i.e., Question 1 is open. The first objective of this paper is to answer the question above. We will develop some general theory answering this question.

For any $(n, k, d, q; r)$-LLRC, Gopalan et al. proved the following upper bound on the minimum distance d [19]:

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

The bound in (1) is similar to the Singleton bound, so we call it the Singleton-like bound. If an $(n, k, d, q; r)$-LLRC meets the Singleton-like bound with equality, then we say that the $(n, k, d, q; r)$-LLRC is distance-optimal (d-optimal for short). If an $(n, k, d, q; r)$-LLRC meets the Singleton-like bound minus one with equality, then we say that the $(n, k, d, q; r)$-LLRC is almost distance-optimal (almost d-optimal for short). Note that the Singleton-like bound is not tight for codes over small finite fields, as it is independent of the alphabet size q.

For any $(n, k, d, q; r)$-LLRC, Cadambe and Mazumdar developed the following bound on the dimension k [5, 6]:

$$k \leq \min_{t \in \mathbb{Z}_+} [tr + k_{opt}^{(q)}(n - t(r + 1), d)], \tag{2}$$

where $\mathbb{Z}_+$ denotes the set of all positive integers, and $k_{opt}^{(q)}(n, d)$ is the largest possible dimension of a linear code with length n, minimum distance d, and alphabet size q. In this paper, we call the bound in (2) the CM bound. An $(n, k, d, q; r)$-LLRC that attains the CM bound with equality is said to be dimension-optimal (k-optimal for short). Note that the CM bound takes the alphabet size q into consideration, but the bound in (1) is independent of q. However, the CM bound involves a parameter $k_{opt}^{(q)}(n, d)$, which may be hard to determine in many cases. The two bounds may not be derived from each other.

While constructing new optimal LLRCs is an important task, searching for optimal LLRCs in the known families of linear codes is also important. The second objective of this paper is to study the minimum linear locality of certain known families of linear codes and try to find out d-optimal or k-optimal LLRCs. We focus on non-binary linear codes, as the linear locality of some families of binary codes were studied in [23]. Our methodology is combinatorial and group-theoretical.

Locally recoverable codes were proposed for the recovery of data in distributed and cloud storage systems by Gopalan, Huang, Simitci and Yikhanin [19]. In the past nine years, a lot of progress on the study of locally recoverable codes has been made. The reader is referred to [5–9, 19, 23, 26, 27, 29, 32–34, 40, 41] and the references therein for information. Despite of the good progress made by now, Question 1 looks still open, and there is a lack of general theory on the minimum linear locality of linear codes. In addition, the minimum linear locality of many known families of linear codes are not studied in the literature. Motivated by these two facts, this paper develops some general theory about the minimum linear locality of linear codes, and investigates the minimum linear locality of a number of families of linear

codes, such as q-ary Hamming codes, q-ary Simplex codes, generalized Reed-Muller codes, ovoid codes, maximum arc codes, the extended hyperoval codes, and near MDS codes. Many classes of both distance-optimal and dimension-optimal LRCs are presented in this paper. The minimum linear locality of many families of linear codes are settled with the general theory developed in this paper.

The rest of this paper is organized as follows. Section 2 introduces some basics of cyclic and linear codes and the support designs of linear codes. Section 3 develops some general theory about the minimum linear locality of nontrivial linear codes. Section 4 investigates the minimum linear locality of several families of famous linear codes, including the q-ary Hamming codes, the q-ary Simplex codes, the generalized Reed-Muller codes, the ovoid codes, and the maximum arc codes. Section 5 studies the minimum linear locality of near MDS codes. Section 6 summarizes the contributions of this paper and makes some concluding remarks.

2 Preliminaries

To study the minimum linear locality of linear codes, we need to introduce some basics of linear codes and cyclic codes. Since our methodology is combinatorial and group-theoretic, we have to introduce the automorphism groups of linear codes and combinatorial t-designs. The purpose of this section is to introduce these stuffs very briefly.

2.1 BCH and cyclic codes

An $[n, k, d]$ code $\mathcal{C}$ over GF(q) is said to be *cyclic* if for each $(c_0, c_1, c_2, \ldots, c_{n-1}) \in \mathcal{C}$ we have $(c_{n-1}, c_0, c_1, c_2, \ldots, c_{n-2}) \in \mathcal{C}$. We identify a vector $(c_0, c_1, c_2, \ldots, c_{n-1}) \in \mathrm{GF}(q)^n$ with the polynomial $c(x) = \sum_{i=0}^{n-1} c_i x^i \in \mathrm{GF}(q)[x]/\langle x^n - 1\rangle$. Then a code $\mathcal{C}$ of length n over GF(q) corresponds to a subset $\mathcal{C}(x)$ of the quotient ring GF$(q)[x]/\langle x^n - 1\rangle$, where

$$\mathcal{C}(x) := \left\{ \sum_{i=0}^{n-1} c_i x^i : c = (c_0, c_1, \ldots, c_{n-1}) \in \mathcal{C} \right\}.$$

It is easy to see that $\mathcal{C}$ is cyclic if and only if the set $\mathcal{C}(x)$ is an ideal of the ring GF$(q)[x]/\langle x^n - 1\rangle$.

It is well-known that each ideal of GF$(q)[x]/\langle x^n - 1\rangle$ is principal when $\gcd(n, q) = 1$. Let $\mathcal{C} = \langle g(x)\rangle$ be a cyclic code, where $g(x)$ is monic and has the smallest degree among all the generators of $\mathcal{C}$. Then $g(x)$ is unique and called the *generator polynomial,* and $h(x) = (x^n - 1)/g(x)$ is referred to as the *check polynomial* of $\mathcal{C}$.

Let n be a positive integer with $\gcd(n, q) = 1$, and let $m = \mathrm{ord}_n(q)$ be the order of q modulo n. Let α be a generator of the multiplicative group GF$(q^m)^*$. Put $\beta = \alpha^{(q^m-1)/n}$. Then β is a primitive n-th root of unity in GF(q^m). The minimal polynomial $\mathbb{M}_{\beta^s}(x)$ of β^s over GF(q) is defined to be the monic polynomial of the smallest degree over GF(q) with β^s as a root and is given by

$$\mathbb{M}_{\beta^s}(x) = \prod_{i \in C_s} (x - \beta^i) \in \mathrm{GF}(q)[x], \tag{3}$$

where $C_s = \{sq^i \bmod n : 0 \le i \le m - 1\}$ and is called the q-cyclotomic coset containing s.

Let δ be an integer with $2 \leq \delta \leq n$ and let h be an integer. A *BCH code* over GF(q) with length n and *designed distance* δ, denoted by $\mathcal{C}_{(q,n,\delta,h)}$, is a cyclic code with generator polynomial

$$g_{(q,n,\delta,h)} = \mathrm{lcm}(\mathbb{M}_{\beta^h}(x), \mathbb{M}_{\beta^{h+1}}(x), \ldots, \mathbb{M}_{\beta^{h+\delta-2}}(x)). \tag{4}$$

If $h = 1$, the code $\mathcal{C}_{(q,n,\delta,h)}$ with the generator polynomial in (4) is referred to as a *narrow-sense* BCH code. If $n = q^m - 1$, then $\mathcal{C}_{(q,n,\delta,h)}$ is called a *primitive* BCH code.

BCH codes form a subfamily of cyclic codes with attractive properties and applications. In many cases BCH codes are the best cyclic codes. For instance, among all binary cyclic codes of odd lengths n with $n \leq 125$ the best cyclic code is always a BCH code except for two special cases [11]. Reed-Solomon codes can be defined as punctured BCH codes and have been widely used in data storage systems, communication devices and consumer electronics.

2.2 Several basic operations on linear codes

Let $\mathcal{C}$ be a linear code with length n. Below we introduce several basic operations on $\mathcal{C}$ for obtaining new codes. Let T be a set of coordinate positions in $\mathcal{C}$ and let $\mathcal{C}^T$ denote the code obtained by puncturing $\mathcal{C}$ in all the coordinate positions in T, which has length $n - |T|$. Let $\mathcal{C}(T)$ denote the set of codewords whose coordinates are $\mathbf{0}$ on T, which is a subcode of $\mathcal{C}$. After puncturing $\mathcal{C}(T)$ on T, we get a linear code over GF(q) with length $n - |T|$, which is called a *shortened code* of $\mathcal{C}$, and is denoted by $\mathcal{C}_T$. It is known that $(\mathcal{C}^\perp)_T = (\mathcal{C}^T)^\perp$ and $(\mathcal{C}^\perp)^T = (\mathcal{C}_T)^\perp$. The *extended code* $\overline{\mathcal{C}}$ of $\mathcal{C}$ is defined by

$$\overline{\mathcal{C}} = \left\{ (c_0, c_1, \ldots, c_{n-1}, c_n) \in \mathrm{GF}(q)^{n+1} : (c_0, c_1, \ldots, c_{n-1}) \in \mathcal{C} \text{ with } \sum_{i=0}^{n} c_i = 0 \right\}.$$

Let G be a generator matrix of $\mathcal{C}$. Suppose that the all-1 vector is not a codeword of $\mathcal{C}$. Then the *augmented code*, denoted by $\widetilde{\mathcal{C}}$, of $\mathcal{C}$ is the linear code over GF(q) with generator matrix

$$\begin{bmatrix} G \\ \mathbf{1} \end{bmatrix},$$

where $\mathbf{1}$ denotes the all-1 vector. The augmented code has length n and dimension $k + 1$. Later in this paper, we will study the minimum linear locality of some punctured or shortened or augmented code of some linear codes.

2.3 Automorphism groups of linear codes

The *permutation automorphism group* of $\mathcal{C}$, denoted by PAut($\mathcal{C}$), is the set of coordinate permutations that map a code $\mathcal{C}$ to itself. A square matrix having exactly one nonzero element of GF(q) in each row and column is called a *monomial matrix* over GF(q). A monomial matrix M can be written in the form DP or the form PD_1, where P is a permutation matrix and D and D_1 are diagonal matrices. The *monomial automorphism group* of $\mathcal{C}$ refers to the set of monomial matrices that map $\mathcal{C}$ to itself. Obviously, PAut($\mathcal{C}$) $\subseteq$ MAut($\mathcal{C}$). The *automorphism group* of $\mathcal{C}$, denoted by Aut($\mathcal{C}$), is the set of maps of the form $M\gamma$ that map $\mathcal{C}$ to itself, where M is a monomial matrix and γ is a field automorphism. If $q = 2$, PAut($\mathcal{C}$), MAut($\mathcal{C}$) and Aut($\mathcal{C}$) are the same. If q is a prime, MAut($\mathcal{C}$) and Aut($\mathcal{C}$) are identical. In general, we have

$$\mathrm{PAut}(\mathcal{C}) \subseteq \mathrm{MAut}(\mathcal{C}) \subseteq \mathrm{Aut}(\mathcal{C}).$$

By the definitions above, each element in $\mathrm{Aut}(\mathcal{C})$ is of the form $DP\gamma$, where D is a diagonal matrix, P is a permutation matrix, and γ is an automorphism of $\mathrm{GF}(q)$. The automorphism group $\mathrm{Aut}(\mathcal{C})$ is said to be *t-transitive* if for every pair of t-element ordered sets of coordinates, there is an element $DP\gamma$ of the automorphism group $\mathrm{Aut}(\mathcal{C})$ such that its permutation part P sends the first set to the second set. The automorphism group $\mathrm{Aut}(\mathcal{C})$ is said to be *t-homogeneous* if for every pair of t-element sets of coordinates, there is an element $DP\gamma$ of the automorphism group $\mathrm{Aut}(\mathcal{C})$ such that its permutation part P sends the first set to the second set. If the automorphism group $\mathrm{Aut}(\mathcal{C})$ is t-transitive, then it must be t-homogeneous. But the converse may not be true. For simplicity, we say that $\mathrm{Aut}(\mathcal{C})$ is transitive (respectively, homogeneous) if $\mathrm{Aut}(\mathcal{C})$ is 1-transitive (respectively, 1-homogeneous).

2.4 The support designs of linear codes

Let $\mathcal{P}$ be a set of n elements, and let $\mathcal{B}$ be a set of k-subsets of $\mathcal{P}$, where $1 \leq k \leq n$. Let t be an integer with $1 \leq t \leq k$. The pair $\mathbb{D} := (\mathcal{P}, \mathcal{B})$ is an incidence structure, where the incidence relation is the set membership. The incidence structure $\mathbb{D} = (\mathcal{P}, \mathcal{B})$ is called a t-(n, k, λ) *design*, or simply *t-design*, if each t-subset of $\mathcal{P}$ is contained in λ elements of $\mathcal{B}$. The elements of $\mathcal{P}$ are referred to as points, and those of $\mathcal{B}$ are called blocks. If $\mathcal{B}$ does not contain any repeated blocks, then the t-design is called *simple*. This paper considers only simple t-designs. A t-(n, k, λ) design is referred to as a *Steiner system* if $t \geq 2$ and $\lambda = 1$, and is denoted by $S(t, k, n)$.

There are different ways to construct t-designs. A coding-theoretic construction of t-designs is briefly described below. Let $\mathcal{C}$ be a linear code over $\mathrm{GF}(q)$ with length n. For each k with $A_k \neq 0$, let $\mathcal{B}_k(\mathcal{C})$ denote the set of the supports of all codewords with Hamming weight k in $\mathcal{C}$, where the coordinates of a codeword are indexed by $(0, 1, \ldots, n-1)$. Let $\mathcal{P}(\mathcal{C}) = [n]$. The incidence structure $(\mathcal{P}(\mathcal{C}), \mathcal{B}_k(\mathcal{C}))$ may be a t-(n, k, λ) design for some positive integer λ, which is called a *support design* of the code $\mathcal{C}$, and is denoted by $\mathbb{D}_k(\mathcal{C})$. In such a case, we say that the codewords of weight k in $\mathcal{C}$ support or hold a t-(n, k, λ) design, and for simplicity, we say that $\mathcal{C}$ supports or holds a t-(n, k, λ) design.

The following theorem, called the Assmus-Mattson Theorem, demonstrates that the pair $(\mathcal{P}(\mathcal{C}), \mathcal{B}_k(\mathcal{C}))$ defined by a linear code $\mathcal{C}$ is a t-design under certain conditions [3].

Theorem 1 *Let $\mathcal{C}$ be an $[n, k, d]$ code over $\mathrm{GF}(q)$. Let $d^\perp$ denote the minimum distance of $\mathcal{C}^\perp$. Let w be the largest integer satisfying $w \leq n$ and*

$$w - \left\lfloor \frac{w+q-2}{q-1} \right\rfloor < d.$$

Define $w^\perp$ analogously using $d^\perp$. Let $(A_i)_{i=0}^n$ and $(A_i^\perp)_{i=0}^n$ denote the weight distribution of $\mathcal{C}$ and $\mathcal{C}^\perp$, respectively. Fix a positive integer t with $t < d$, and let s be the number of i with $A_i^\perp \neq 0$ for $1 \leq i \leq n - t$. Suppose $s \leq d - t$. Then

- *all the codewords of weight i in $\mathcal{C}$ support a simple t-design provided $A_i \neq 0$ and $d \leq i \leq w$, and*
- *all the codewords of weight i in $\mathcal{C}^\perp$ support a simple t-design provided $A_i^\perp \neq 0$ and $d^\perp \leq i \leq \min\{n - t, w^\perp\}$.*

The Assmus-Mattson Theorem above is a useful tool in constructing t-designs from linear codes (see, for example, [13]), but does not characterize all linear codes supporting t-designs. The reader is referred to [36] for a generalized Assmus-Mattson theorem.

Using the automorphism group of a linear code $\mathcal{C}$, the following theorem gives another sufficient condition for the code $\mathcal{C}$ to hold t-designs [24, p. 308].

Theorem 2 [24] *Let $\mathcal{C}$ be a linear code of length n over $\mathrm{GF}(q)$ such that $\mathrm{Aut}(\mathcal{C})$ is t-transitive or t-homogeneous. Then the codewords of any weight $i \geq t$ of $\mathcal{C}$ hold a t-design.*

3 General theory on the minimum locality of linear codes

The objective of this section is to develop some general theory about the minimum locality of linear codes over finite fields. In particular, we will answer Question 1 raised in Sect. 1. As made clear in Sect. 1, we consider only nontrivial linear codes $\mathcal{C}$ over finite fields, i.e., linear codes $\mathcal{C}$ with $d(\mathcal{C}) > 1$ and $d(\mathcal{C}^{\perp}) > 1$. Recall that we use the elements in $[n]$ to index the coordinate positions in a linear code of length n.

3.1 Minimum locality and minimum linear locality of nontrivial linear codes are identical

The minimum locality and minimum linear locality of linear codes were defined in Sect. 1. A referee of this paper has informed the authors of this paper that the proof of the Singleton bound for LRCs in [19] implies the minimum locality and minimum linear locality are equal. Since a complete proof of this equality is missing in the literature, we present a proof of the following lemma.

Lemma 1 *The minimum locality and minimum linear locality of a nontrivial linear code are identical.*

Proof Let $\mathcal{C}$ be a nontrivial linear code of length n over $\mathrm{GF}(q)$. Since $\mathcal{C}$ is nontrivial, $d(\mathcal{C}) > 1$ and $d(\mathcal{C}^{\perp}) > 1$. By definition, $\mathcal{C}$ has locality r if it has linear locality r. This means that the minimum locality is no more than the minimum linear locality of $\mathcal{C}$.

Suppose that $\mathcal{C}$ has minimum locality r. Consider any code symbol c_{i_0} with $i_0 \in [n]$. Let $1 \leq s \leq r$ be the smallest integer such that c_{i_0} can be recovered from s other code symbols, say, $c_{i_1}, \ldots, c_{i_s}$, where $i_1, \ldots, i_s$ are pairwise distinct and $\{i_1, \ldots, i_s\} \subseteq [n] \setminus \{i_0\}$. By definition, there is a function F from $\mathrm{GF}(q)^s$ to $\mathrm{GF}(q)$ such that

$$c_{i_0} = F(c_{i_1}, \ldots, c_{i_s}). \tag{5}$$

Let $\mathcal{C}(i_0, i_1, \ldots, i_s)$ denote the code obtained by puncturing $\mathcal{C}$ on the coordinate positions in $[n] \setminus \{i_0, i_1, \ldots, i_s\}$. Then $\mathcal{C}(i_0, i_1, \ldots, i_s)$ is a linear code of length $s + 1$ over $\mathrm{GF}(q)$. It then follows from (5) that

$$|\mathcal{C}(i_0, i_1, \ldots, i_s)| \leq q^s.$$

We then deduce that $\mathcal{C}(i_0, i_1, \ldots, i_s)$ is a proper subspace of $\mathrm{GF}(q)^{s+1}$. Consequently,

$$\mathcal{C}(i_0, i_1, \ldots, i_s)^{\perp} \neq \{\mathbf{0}\}.$$

Let $(a_0, a_1, \ldots, a_s)$ be a nonzero codeword in $\mathcal{C}(i_0, i_1, \ldots, i_s)^{\perp}$. Then

$$\sum_{j=0}^{s} a_j c_{i_j} = 0. \tag{6}$$

If $a_0 = 0$, then one of $a_1, \ldots, a_s$ must be nonzero. Without loss of generality, assume that $a_1 \neq 0$, then

$$c_{i_1} = -a_1^{-1}(a_2 c_{i_2} + \cdots + a_s c_{i_s}). \tag{7}$$

Combining (5) and (7), we deduce that the code symbol c_{i_0} can be recovered from $c_{i_2}, \ldots, c_{i_s}$, which is contrary to the minimality of s. Consequently, $a_0 \neq 0$.

Since $a_0 \neq 0$, (6) means that c_{i_0} can be linearly recovered from $c_{i_1}, \ldots, c_{i_s}$. Consequently, the minimum linear locality is no more than the minimum locality of $\mathcal{C}$. Summarizing the discussions above, we conclude that the minimum locality and minimum linear locality of a nontrivial linear code are identical. □

In general, the complexity of recovering a code symbol linearly from other code s symbols is no more than the complexity of recovering it from a nonlinear method. Consequently, it would be better to distinguish the linear locality from the general locality for linear codes. These justify the motivation of studying the linear locality and minimum linear locality of linear codes over finite fields. Lemma 1 allows us to settle the minimum locality of many families of linear codes by determining their minimum linear locality. It should be noticed that the minimum locality and the minimum linear locality of some nonlinear codes may be different.

3.2 Some general theory of the minimum locality of nontrivial linear codes

The following lemma follows from the definition of linear locality of linear codes. For completeness, we provide a proof of it below.

Lemma 2 *Let $\mathcal{C}$ be a nontrivial linear code of length n. Then $\mathcal{C}$ has linear locality r if and only if for each $i \in [n]$ the dual code $\mathcal{C}^\perp$ has a codeword $\mathbf{c}^\perp$ of Hamming weight at most $r+1$ such that $i \in \mathrm{supp}(\mathbf{c}^\perp)$, where $\mathrm{supp}(\mathbf{c}^\perp)$ denotes the support of the codeword $\mathbf{c}^\perp$.*

Proof Suppose that $\mathcal{C}$ is over $\mathrm{GF}(q)$ and has linear locality r. By definition, for each $i \in [n]$ there are a subset $\{i_1, \ldots, i_r\} \subseteq [n] \setminus \{i\}$ and r elements $a_1, \ldots, a_r$ in $\mathrm{GF}(q)$ such that $c_i = a_1 c_{i_1} + \cdots + a_r c_{i_r}$. Let $\mathbf{c}^\perp = (c_0^\perp, \ldots, c_{n-1}^\perp)$, where $c_i^\perp = 1$, $c_{i_j}^\perp = -a_j$ for $1 \leq j \leq r$ and $c_h^\perp = 0$ if $h \in [n] \setminus \{i, i_1, \ldots, i_r\}$. Then $\mathbf{c}^\perp$ has weight at most $r+1$ and is a codeword in $\mathcal{C}^\perp$.

Suppose that for each $i \in [n]$ there is a codeword $\mathbf{c}^\perp = (c_0^\perp, \ldots, c_{n-1}^\perp)$ in $\mathcal{C}^\perp$ such that $i \in \mathrm{supp}(\mathbf{c}^\perp)$. Assume that $\mathrm{supp}(\mathbf{c}^\perp) = \{i, i_1, \ldots, i_s\}$ with $s \leq r$. We have then

$$c_i = -(c_i^\perp)^{-1}(c_{i_1}^\perp c_{i_1} + \cdots + c_{i_s}^\perp c_{i_s}).$$

This means that c_i can be linearly recovered from $c_{i_1}, \ldots, c_{i_s}$. Hence, $\mathcal{C}$ has linear locality r. □

Theorem 3 *Let $\mathcal{C}$ be a nontrivial linear code of length n. Then there exists a positive integer w with $2 \leq w \leq n$ such that $A_w(\mathcal{C}^\perp) > 0$ and*

$$\bigcup_{j=1}^{w} \bigcup_{S \in \mathcal{B}_j(\mathcal{C}^\perp)} S = [n]. \tag{8}$$

Let w be the smallest integer such that (8) holds. Then $\mathcal{C}$ has minimum locality $w-1$.

Proof Suppose that there is an integer i in $[n]$ such that

$$i \notin \bigcup_{j=1}^{n} \bigcup_{S \in \mathcal{B}_j(\mathcal{C}^\perp)} S.$$

Then $c_i^\perp = 0$ for all codewords $\mathbf{c}^\perp = (c_0^\perp, \ldots, c_{n-1}^\perp)$ in $\mathcal{C}^\perp$. Consequently, the vector $\mathbf{c} = (0, \ldots, 0, 1, 0, \ldots, 0)$ of length n, which has only one nonzero coordinate 1 in coordinate position i, is a codeword in $\mathcal{C}$. This is contrary to the assumption that $d(\mathcal{C}) \geq 2$.

Let w be the smallest integer such that $A_w(\mathcal{C}^\perp) > 0$ and (8) holds. Then every integer $i \in [n]$ is contained in $\mathrm{supp}(\mathbf{c}^\perp)$, where $\mathbf{c}^\perp$ is some codeword with weight at most w in $\mathcal{C}^\perp$. Then the code symbol c_i in $\mathcal{C}$ can be recovered by a linear combination of the coordinates in the positions in $\mathrm{supp}(\mathbf{c}^\perp) \setminus \{i\}$. Then $\mathcal{C}$ has linear locality $w - 1$.

If the code symbol c_i in every codeword $\mathbf{c}$ in $\mathcal{C}$ can be recovered linearly by

$$c_i = u_1 c_{i_1} + \cdots + u_h c_{i_h}$$

where $u_i \neq 0$ and $R_i = \{i_1, \ldots, i_h\}$ is the corresponding recovering set of the code symbol c_i. Then $\mathcal{C}^\perp$ has a codeword with weight $h + 1$. Hence, $w - 1$ is the minimum linear locality, which is the minimum locality. □

Theorem 3 means that every nontrivial linear code has a minimum locality, and tells us how to calculate the minimum locality. In practice, it is also necessary and important to find a recovering set R_i for each code symbol c_i. But we will not deal with this problem in this paper.

3.3 Linear codes $\mathcal{C}$ with minimum locality $d(\mathcal{C}^\perp) - 1$

It follows from Lemma 2 that the minimum locality of a nontrivial linear code $\mathcal{C}$ is at least $d(\mathcal{C}^\perp) - 1$. Hence, nontrivial linear codes $\mathcal{C}$ with minimum locality $d(\mathcal{C}^\perp) - 1$ would be very interesting in both theory and practice. In this subsection, we develop some general results for such special codes. It will be seen later that there are indeed nontrivial linear codes $\mathcal{C}$ with minimum locality more than $d(\mathcal{C}^\perp) - 1$.

Corollary 1 *Let $\mathcal{C}$ be a nontrivial linear code of length n and put $d^\perp = d(\mathcal{C}^\perp)$. Then $\mathcal{C}$ has minimum locality $d^\perp - 1$ if and only if*

$$\bigcup_{S \in \mathcal{B}_{d^\perp}(\mathcal{C}^\perp)} S = [n]. \tag{9}$$

Proof The desired conclusion directly follows from Theorem 3 and Lemma 2.

The following result is well known in the literature [29]. We show that it is a corollary of Theorem 3.

Corollary 2 [29] *Let $\mathcal{C}$ be a nontrivial cyclic code of length n. Then $\mathcal{C}$ has minimum locality $d(\mathcal{C}^\perp) - 1$.*

Proof Put $d^\perp = d(\mathcal{C}^\perp)$. Let $i \in [n]$. Let $\mathbf{c}^\perp = (c_0^\perp, \ldots, c_{n-1}^\perp)$ be a minimum weight codeword in $\mathcal{C}^\perp$. By definition, the Hamming weight $\mathrm{wt}(\mathbf{c}^\perp) \geq 2$. Consequently, $\mathbf{c}^\perp$ has a nonzero coordinate. Since $\mathcal{C}^\perp$ is also cyclic, we can assume $c_i^\perp \neq 0$. We then deduce that

$$\bigcup_{S \in \mathcal{B}_{d^\perp}(\mathcal{C}^\perp)} S \supseteq \{i\}.$$

The desired conclusion then follows from Theorem 3. □

Corollary 3 *Let $\mathcal{C}$ be a nontrivial linear code of length n and put $d^\perp = d(\mathcal{C}^\perp)$. If $(\mathcal{P}(\mathcal{C}^\perp), \mathcal{B}_{d^\perp}(\mathcal{C}^\perp))$ is a 1-$(n, d^\perp, \lambda_1^\perp)$ design with $\lambda_1^\perp \geq 1$, then $\mathcal{C}$ has minimum locality $d^\perp - 1$.*

Proof By the definition of 1-designs, every $i \in \mathcal{P}(\mathcal{C}^\perp)$ is covered in $\lambda_1^\perp$ blocks in the block set $\mathcal{B}_{d^\perp}(\mathcal{C}^\perp)$. Hence,

$$\bigcup_{j=1}^{d^\perp} \bigcup_{S \in \mathcal{B}_j(\mathcal{C}^\perp)} S = \bigcup_{S \in \mathcal{B}_{d^\perp}(\mathcal{C}^\perp)} S = [n].$$

The desired conclusion then follows from Theorem 3. □

It should be noted that there are many nontrivial linear codes with minimum locality $d(\mathcal{C}^\perp) - 1$, but $(\mathcal{P}(\mathcal{C}^\perp), \mathcal{B}_{d^\perp}(\mathcal{C}^\perp))$ is not a 1-design. Hence, the converse of Corollary 3 is not true. Corollary 3 will be one of the tools for studying the minimum locality of some families of linear codes in this paper. Another tool is documented in the following corollary, which is a slight strengthening of Lemma 2 in [29].

Corollary 4 *Let $\mathcal{C}$ be a nontrivial linear code. If $\mathrm{Aut}(\mathcal{C})$ or $\mathrm{Aut}(\mathcal{C}^\perp)$ is transitive, then $\mathcal{C}$ has minimum locality $d(\mathcal{C}^\perp) - 1$ and $\mathcal{C}^\perp$ has minimum locality $d(\mathcal{C}) - 1$.*

Proof Put $d^\perp = d(\mathcal{C}^\perp)$. Let $\mathcal{C}$ be over $\mathrm{GF}(q)$ and have length n. Suppose that $\mathrm{Aut}(\mathcal{C}^\perp)$ is transitive. Let $\mathbf{c}^\perp$ be a minimum weight codeword in $\mathcal{C}^\perp$. Then $\mathrm{wt}(\mathbf{c}^\perp) \geq 2$. Let $i \in \mathrm{supp}(\mathbf{c}^\perp)$. For each $j \in [n] \setminus \{i\}$, there is an automorphism $DP\gamma$ in $\mathrm{Aut}(\mathcal{C}^\perp)$ such that the permutation part P sends i to j, as $\mathrm{Aut}(\mathcal{C}^\perp)$ is transitive. This means there is another minimum weight codeword $(\mathbf{c}')^\perp$ in $\mathcal{C}^\perp$ such that $j \in \mathrm{supp}((\mathbf{c}')^\perp)$. Consequently,

$$\bigcup_{j=1}^{d^\perp} \bigcup_{S \in \mathcal{B}_j(\mathcal{C}^\perp)} S = \bigcup_{S \in \mathcal{B}_{d^\perp}(\mathcal{C}^\perp)} S = [n].$$

It then follows from Theorem 3 that $\mathcal{C}$ has minimum locality $d(\mathcal{C}^\perp) - 1$.

In general, $\mathrm{Aut}(\mathcal{C})$ and $\mathrm{Aut}(\mathcal{C}^\perp)$ are different. However, it is straightforward to prove that $\mathrm{Aut}(\mathcal{C})$ is transitive if and only if $\mathrm{Aut}(\mathcal{C}^\perp)$ is so. Then the remaining desired conclusion follows from the first conclusion proved above. □

Lemma 2 in [29] can be employed to conclude that the minimum locality of $\mathcal{C}$ is at most $d(\mathcal{C}^\perp) - 1$ if $\mathrm{Aut}(\mathcal{C})$ is transitive. Hence, Corollary 4 is a slight strengthening of Lemma 2 in [29].

Note that combining Theorem 2 and Corollary 3 gives another proof of Corollary 4. Sometimes we may need to use Corollary 3, as the automorphism group of a code may be unknown. Sometimes it is more convenient to use Corollary 4. Sometimes both corollaries can be used to study the linear locality of some linear codes. In many cases, both corollaries cannot be used to do so. It looks impossible to find out all nontrivial linear codes with minimum locality $d(\mathcal{C}^\perp) - 1$. But Corollaries 3 and 4 can be employed to find many families of such codes. Most of the families of linear codes documented in the monograph [13] are such codes, as they support t-designs with $t \geq 2$ or their automorphism groups are doubly homogeneous. Other families of such linear codes are not documented in [13], as the monograph [13] does not include linear codes supporting 1-designs but not 2-designs.

The following result would also be useful in some cases.

Theorem 4 *Let $\mathcal{C}$ be a nontrivial linear code. If $\mathcal{C}^{\perp}$ is spanned by its minimum weight codewords, then $\mathcal{C}$ has minimum locality $d(\mathcal{C}^{\perp}) - 1$.*

Proof Let $\mathcal{C}$ have length n. Let $i \in [n]$. If

$$i \notin \bigcup_{S \in \mathcal{B}_{d^{\perp}}(\mathcal{C}^{\perp})} S,$$

then $(0, \ldots, 0, 1, 0, \ldots, 0)$ would be a codeword in $\mathcal{C}$, where the nonzero coordinate 1 is in coordinate position i, as all the minimum weight codewords in $\mathcal{C}^{\perp}$ span $\mathcal{C}^{\perp}$. This is contrary to the fact that $\mathcal{C}$ is nontrivial. The desired conclusion then follows from Corollary 1. □

3.4 The minimum locality of extended cyclic code

While any nontrivial cyclic code $\mathcal{C}$ has minimum locality $d(\mathcal{C}^{\perp}) - 1$, extended cyclic codes may not have such property. Note that even if $\mathcal{C}$ is nontrivial, the extended code $\overline{\mathcal{C}}$ may be trivial, as $d(\overline{\mathcal{C}}^{\perp})$ could be 1. The automorphism group of any cyclic code is transitive and each cyclic code supports 1-designs. But these may not be true for extended cyclic codes. In this section, we consider the linear locality of the extended cyclic codes and their duals.

Let H and $\overline{H}$ denote the parity-check matrix of $\mathcal{C}$ and $\overline{\mathcal{C}}$, respectively. Then we have the following well known lemma [24].

Lemma 3 *Let $\mathcal{C}$ be an $[n, \kappa, d]$ code over $\mathrm{GF}(q)$. Then $\overline{\mathcal{C}}$ is an $[n+1, \kappa, \overline{d}]$ linear code, where $\overline{d} = d$ or $d+1$. In the binary case, $\overline{d} = d$ if d is even, and $\overline{d} = d+1$ otherwise.*

In addition, the parity-check matrix $\overline{H}$ of $\overline{\mathcal{C}}$ can be deduced from that of $\mathcal{C}$ by

$$\overline{H} = \begin{bmatrix} \mathbf{1} & 1 \\ H & \mathbf{0} \end{bmatrix}, \tag{10}$$

where $\mathbf{1} = (1, 1, \ldots, 1)$ and $\mathbf{0} = (0, 0, \ldots, 0)^T$.

We now prove the following result, which will be needed later.

Theorem 5 *Let $\mathcal{C}$ be a nontrivial cyclic code. If $d(\overline{\mathcal{C}}) = d(\mathcal{C}) + 1$, then $(\overline{\mathcal{C}})^{\perp}$ has minimum locality $d(\mathcal{C})$.*

Proof Let $\mathcal{C}$ have length n. By definition, $d(\mathcal{C}) \geq 2$ and $d(\mathcal{C}^{\perp}) \geq 2$. Since $d(\overline{\mathcal{C}}) = d(\mathcal{C}) + 1$, we know that $d((\overline{\mathcal{C}})^{\perp}) \geq 2$. Therefore, $\overline{\mathcal{C}}$ is nontrivial. Let $\mathbf{c}_1, \ldots, \mathbf{c}_h$ be all the minimum weight codewords in $\mathcal{C}$, and let $\overline{\mathbf{c}_i}$ be the extended codeword of $\mathbf{c}_i$ in $\overline{\mathcal{C}}$. Since $d(\mathcal{C}) \geq 2$ and $\mathcal{C}$ is cyclic, we have

$$\bigcup_{i=1}^{h} \mathrm{supp}(\mathbf{c}_i) = [n].$$

Since $d(\overline{\mathcal{C}}) = d(\mathcal{C}) + 1$, the extended coordinate in each $\overline{\mathbf{c}_i}$ is nonzero. As a result, we get

$$\bigcup_{i=1}^{h} \mathrm{supp}(\overline{\mathbf{c}_i}) = [n+1].$$

The desired conclusion then follows from Corollary 1. □

Corollary 5 *Let $\mathcal{C}$ be a nontrivial binary cyclic code. If $d(\mathcal{C})$ is odd, then $(\overline{\mathcal{C}})^{\perp}$ has minimum locality $d(\mathcal{C})$.*

Proof The desired conclusion follows from Lemma 3 and Theorem 5. □

Corollary 5 has determined the minimum locality of $(\overline{\mathcal{C}})^{\perp}$ for all binary cyclic codes. Specifically, either $(\overline{\mathcal{C}})^{\perp}$ is a trivial code or nontrivial binary linear code with minimum locality $d(\mathcal{C})$ for each nontrivial cyclic code $\mathcal{C}$.

4 The minimum locality of some known families of linear codes

The objective of this section is to study the minimum locality of several families of linear codes which are geometric codes and their punctured and shortened codes. We wish to find out some families of optimal LLRCs.

4.1 The minimum locality of the q-ary Hamming codes and Simplex codes

A parity check matrix $H_{(q,m)}$ of the Hamming code $\mathcal{H}_{(q,m)}$ over GF(q) is defined by choosing for its columns a nonzero vector from each one-dimensional subspace of GF$(q)^m$. In terms of finite geometry, the columns of $H_{(q,m)}$ are the points of the projective geometry PG$(m-1, \mathrm{GF}(q))$. Hence $\mathcal{H}_{(q,m)}$ has length $n = (q^m-1)/(q-1)$ and dimension $n-m$. It is well known that $\mathcal{H}_{(q,m)}$ has minimum weight 3 and any $[(q^m-1)/(q-1), (q^m-1)/(q-1)-m, 3]$ code over GF(q) is monomially equivalent to the Hamming code $\mathcal{H}_{(q,m)}$ [24]. Note that the Hamming code $\mathcal{H}_{(q,m)}$ is permutation-equivalent to a cyclic code when $\gcd(m, q-1) = 1$. By Corollary 2, its minimum locality is known in this case. However, its minimum locality may not be known for the case that $\gcd(m, q-1) \neq 1$. The linear locality of the binary Hamming and Simplex codes was settled in [23]. Note that binary Hamming and Simplex codes are permutation-equivalent to cyclic codes. In this subsection, we investigate the minimum locality of the q-ary Hamming and Simplex codes.

The weight distribution of $\mathcal{H}_{(q,m)}$ is given in the following lemma [28].

Lemma 4 [28] *The weight distribution of $\mathcal{H}_{(q,m)}$ is given by*

$$q^m A_k(\mathcal{H}_{(q,m)}) = \sum_{\substack{0\le i\le \frac{q^{m-1}-1}{q-1} \\ 0\le j\le q^{m-1} \\ i+j=k}} \left[\binom{\frac{q^{m-1}-1}{q-1}}{i} \binom{q^{m-1}}{j} \left((q-1)^k + (-1)^j (q-1)^i (q^m-1) \right) \right]$$

for $0 \le k \le (q^m-1)/(q-1)$.

The duals of the Hamming codes $\mathcal{H}_{(q,m)}$ are called Simplex codes, denoted by $\mathcal{S}_{(q,m)}$, which have parameters $[(q^m-1)/(q-1), m, q^{m-1}]$. The nonzero codewords of the $[(q^m-1)/(q-1), m, q^{m-1}]$ Simplex codes all have weight q^{m-1}.

Theorem 6 *The Hamming code $\mathcal{H}_{(q,m)}$ is an $(n, n-m, 3, q; q^{m-1}-1)$-LLRC and the Simplex code $\mathcal{S}_{(q,m)}$ is an $(n, m, q^{m-1}, q; 2)$-LLRC. Furthermore, the Hamming code $\mathcal{H}_{(q,m)}$ and $\mathcal{S}_{(q,m)}$ are k-optimal.*

Proof The Hamming code $\mathcal{H}_{(q,m)}$ has parameters $[n, n-m, 3]$ and its dual code is a one-weight code. Then by the Assmus-Mattson Theorem, the codewords of minimum weight in the Hamming code and Simplex code both hold a 2-design. So $(\mathcal{P}(\mathcal{H}_{(q,m)}), \mathcal{B}_3(\mathcal{H}_{(q,m)}))$ and $(\mathcal{P}(\mathcal{H}^{\perp}_{(q,m)}), \mathcal{B}_{q^{m-1}}(\mathcal{H}^{\perp}_{(q,m)}))$ are 1-designs. Hence, the conclusions on the minimum locality of the two codes follow from Corollary 3.

We now prove the dimension optimality of $\mathcal{H}_{(q,m)}$. Putting $t = 1$ and the parameters of the $(n, n-m, 3, q; q^{m-1}-1)$-LLRC into the right-hand side of the CM bound in (2), we have

$$\begin{aligned} k &\le \min_{s\in\mathbb{Z}_+}\{rs + k^{(q)}_{opt}(n-(r+1)s, d)\} \\ &\le r + k^{(q)}_{opt}(n-(r+1), d) \\ &= q^{m-1} - 1 + k^{(q)}_{opt}(n-q^{m-1}, 3) \\ &\le n-m, \end{aligned}$$

where the last inequality holds due to the fact that $k^{(q)}_{opt}(n-q^{m-1}, 3) \le n-q^{m-1}-m+1$, which follows from the sphere packing bound. Therefore, the Hamming code $\mathcal{H}_{(q,m)}$ is k-optimal.

Putting $t = 1$ and the parameters of $(n, m, q^{m-1}, q; 2)$-LLRC into the right-hand side of the CM bound in (2), we have

$$\begin{aligned} k &\le r + k^{(q)}_{opt}(n-(r+1), d) \\ &= 2 + k^{(q)}_{opt}(n-3, q^{m-1}) \\ &\le m, \end{aligned}$$

where the last inequality holds due to the fact that $k^{(q)}_{opt}(n-3, q^{m-1}) \le m-2$, which follows from the Plotkin bound. Therefore, the Simplex code $\mathcal{S}_{(q,m)}$ is k-optimal. This completes the proof. □

According to the Singleton-like bound in (1), we can obtain the following family of d-optimal LLRCs.

Corollary 6 *When $m = 3$, the Hamming code $\mathcal{H}_{(q,3)}$ is a $(q^2+q+1, q^2+q-2, 3, q; q^2-1)$-LLRC and is both d-optimal and k-optimal.*

Proof The parameters and dimension optimality of $\mathcal{H}_{(q,3)}$ directly follow from Theorem 6. Hence, we only need to prove the distance optimality. It is easy to verify that the parameters of $\mathcal{H}_{(q,3)}$ satisfy the equality in (1). Hence, it is a d-optimal $(q^2+q+1, q^2+q-2, 3, q; q^2-1)$-LLRC. This completes the proof. □

If a linear code $\mathcal{C}$ supports 2-designs, then the punctured code $\mathcal{C}^{\{t_1\}}$ or shortened code $\mathcal{C}_{\{t_1\}}$ may support 1-designs. Then we can settle the minimum locality of $\mathcal{C}^{\{t_1\}}$ or $\mathcal{C}_{\{t_1\}}$. The parameters of some punctured codes and shortened codes of the Hamming code are given in the following lemma [28].

Lemma 5 [28] *Let $n = (q^m-1)/(q-1) \ge 4$, and let t_1 be any coordinate position of codewords in $\mathcal{H}_{(q,m)}$. Then the following hold:*

- *$(\mathcal{H}_{(q,m)})_{\{t_1\}}$ is an $[n-1, n-m-1, 3]$ code over $\mathrm{GF}(q)$ with*

$$A_k((\mathcal{H}_{(q,m)})_{\{t_1\}}) = \frac{n-k}{n} A_k(\mathcal{H}_{(q,m)})$$

for $0 \le k \le n-1$, *where* $A_k(\mathcal{H}_{(q,m)})$ *was given in Lemma 4.*

- $((\mathcal{H}_{(q,m)})_{\{t_1\}})^{\perp}$ *is an* $[n-1, m, q^{m-1}-1]$ *code over* GF(q) *with weight enumerator*

$$1 + (q-1)q^{m-1}z^{q^{m-1}-1} + (q^{m-1}-1)z^{q^{m-1}}.$$

- $(\mathcal{S}_{(q,m)})_{\{t_1\}}$ *is an* $[n-1, m-1, q^{m-1}]$ *code over* GF(q) *with weight enumerator* $1 + (q^{m-1}-1)z^{q^{m-1}}$.
- $((\mathcal{S}_{(q,m)})_{\{t_1\}})^{\perp}$ *is an* $[n-1, n-m, 2]$ *code over* GF(q) *with weight enumerator*

$$\frac{1}{q^{m-1}}[(1+(q-1)z)^{n-1} + (q^{m-1}-1)(1-z)^{q^{m-1}}(1+(q-1)z)^{n-1-q^{m-1}}].$$

With the Assmus-Mattson Theorem, we can deduce that the codewords of minimum weight in these codes in Lemma 5 hold a 1-design. Then by Corollary 3, we can settle the minimum locality of these codes in Lemma 5.

Theorem 7 *Let* $n = (q^m-1)/(q-1) \ge 4$, *and let* t_1 *be any coordinate position of codewords in* $\mathcal{H}_{(q,m)}$. *Then we have the following.*

- $(\mathcal{H}_{(q,m)})_{\{t_1\}}$ *is a k-optimal* $(n-1, n-m-1, 3, q; q^{m-1}-2)$*-LLRC.*
- $((\mathcal{H}_{(q,m)})_{\{t_1\}})^{\perp}$ *is a k-optimal* $(n-1, m, q^{m-1}-1, q; 2)$*-LLRC.*
- $(\mathcal{S}_{(q,m)})_{\{t_1\}}$ *is a k-optimal* $(n-1, m-1, q^{m-1}, q; 1)$*-LLRC.*
- $(\mathcal{S}_{(q,m)})_{\{t_1\}})^{\perp}$ *is a k-optimal* $(n-1, n-m, 2, q; q^{m-1}-1)$*-LLRC.*

Proof The conclusions on the parameters of the codes follow from Lemma 5, the Assmus–Mattson Theorem and Corollary 3. The proofs of the dimension optimality of $(\mathcal{H}_{(q,m)})_{\{t_1\}}$ and $(\mathcal{S}_{(q,m)})_{\{t_1\}})^{\perp}$ are similar to that in Theorem 6, and are omitted. □

Furthermore, we can obtain the following d-optimal LLRCs.

Corollary 7 *The code* $(\mathcal{H}_{(q,3)})_{\{t_1\}}$ *is a* $(q^2+q, q^2+q-3, 3, q; q^2-2)$*-LLRC and the code* $((\mathcal{S}_{(q,3)})_{\{t_1\}})^{\perp}$ *is a* $(q^2+q, q^2+q-2, 2, q; q^2-1)$*-LLRC. Furthermore, they are both d-optimal and k-optimal.*

Proof The parameters of the codes follow from Theorem 7. The d-optimality and k-opmality of the codes are with respect to the Singleton-like bound and CM bound and can be easily verified. □

Remark 1 Recently, Grezet and Hollanti [21, Corollary 1] also determined the locality of the punctured Simplex code using some tools from matroid theory. Our method is more generic in the sense that it also works for many other linear codes.

4.2 The minimum locality of the generalised Reed-Muller codes over GF(q)

The general affine group $\mathrm{GA}_1(\mathrm{GF}(q^m))$ is defined by

$$\mathrm{GA}_1(\mathrm{GF}(q^m)) = \{ax + b : a \in \mathrm{GF}(q^m)^*,\ b \in \mathrm{GF}(q^m)\},$$

which acts on GF(q^m) doubly transitively [13, Section 1.7].

We can index the coordinates of a linear code of length q^m with the elements of GF(q^m). When each permutation in $\mathrm{GA}_1(\mathrm{GF}(q^m))$ is applied to a codeword, it is applied to the indices of the coordinates. A linear code $\mathcal{C}$ of length q^m is said to be *affine-invariant* if $\mathrm{GA}_1(\mathrm{GF}(q))$ fixes $\mathcal{C}$. It follows from Theorem 2 that affine-invariant codes supports 2-designs. By Corollary

4, all affine-invariant codes $\mathcal{C}$ have minimum locality $d(\mathcal{C}^{\perp}) - 1$. There are many infinite families of affine-invariant codes [13, Chapter 6]. Our objective in this section is to study the minimum locality of the generalised Reed-Muller codes and obtain a class of either k-optimal or almost k-optimal LLRCs.

For any integer $j = \sum_{i=0}^{m-1} j_i q^i$, where $0 \le j_i \le q-1$ for all $0 \le i \le m-1$ and m is a positive integer, we define

$$\mathrm{wt}_q(j) = \sum_{i=0}^{m-1} j_i, \tag{11}$$

where the sum is taken over the ring of integers, and is called the q-weight of j. Let ℓ be a positive integer with $1 \le \ell < (q-1)m$. The ℓ-th order *punctured generalized Reed-Muller code* $\mathcal{R}_q(\ell, m)^*$ over GF(q) is the cyclic code of length $n = q^m - 1$ with generator polynomial

$$g(x) = \prod_{\substack{1 \le j \le n-1 \\ \mathrm{wt}_q(j) < (q-1)m - \ell}} (x - \alpha^j), \tag{12}$$

where α is a generator of GF$(q^m)^*$. Since $\mathrm{wt}_q(j)$ is a constant function on each q-cyclotomic coset modulo $n = q^m - 1$, $g(x)$ is a polynomial over GF(q).

The generalized Reed-Muller code $\mathcal{R}_q(\ell, m)$ is defined to be the extended code of $\mathcal{R}_q(\ell, m)^*$, and its parameters are given below [4].

Theorem 8 [4] *Let $0 \le \ell < q(m-1)$. Then the generalized Reed-Muller code $\mathcal{R}_q(\ell, m)$ has length $n = q^m$, dimension*

$$\kappa = \sum_{i=0}^{\ell} \sum_{j=0}^{m} (-1)^j \binom{m}{j} \binom{i - jq + m - 1}{i - jq}, \tag{13}$$

and minimum weight

$$d = (q - \ell_0) q^{m - \ell_1 - 1}, \tag{14}$$

where $\ell = \ell_1(q-1) + \ell_0$ and $0 \le \ell_0 < q - 1$.

The minimum locality of the generalized Reed-Muller code $\mathcal{R}_q(\ell, m)$ is given in the following theorem.

Theorem 9 *Let $0 \le \ell < q(m-1)$ and $m(q-1) - 1 - \ell = \ell_1'(q-1) + \ell_0'$ with $0 \le \ell_0' < q-1$. The generalized Reed-Muller code $\mathcal{R}_q(\ell, m)$ is a $[q^m, \kappa, d, q; (q - \ell_0')q^{m - \ell_1' - 1} - 1]$ LLRC, where κ and d were given in (13) and (14), respectively.*

Proof The dimension and minimum weight of the code were given in Theorem 8. We only prove its minimum locality. It is well known that the generalised Reed-Muller code $\mathcal{R}_q(\ell, m)$ is affine-invariant [13, Chapter 6]. By Corollary 4, the code has minimum locality $d(\mathcal{R}_q(\ell, m)^{\perp}) - 1$. It was proved in [4] that

$$\mathcal{R}_q(\ell, m)^{\perp} = \mathcal{R}_q(m(q-1) - 1 - \ell, m). \tag{15}$$

It then follows from Theorem 8 that the minimum locality is

$$r = d(\mathcal{R}_q(\ell, m)^{\perp}) - 1 = d(\mathcal{R}_q(m(q-1) - 1 - \ell, m)) - 1 = (q - \ell_0')q^{m - \ell_1' - 1} - 1.$$

This completes the proof. □

Corollary 8 *Let $q > 2$. Then $\mathcal{R}_q(1, m)$ is a $(q^m, 1+m, (q-1)q^{m-1}, q; 2)$-LLRC and its dual $\mathcal{R}_q(1, m)^\perp$ is a $(q^m, q^m - 1 - m, 3, q; (q-1)q^{m-1} - 1)$-LLRC. Both codes are k-optimal.*

Proof The parameters of the two codes were given in Theorem 9. Putting the parameters of $\mathcal{R}_q(1, m)^\perp$ into the right-hand side of the CM bound in (2), we have

$$\begin{aligned} k &\leq \min_{s\in\mathbb{Z}_+}\{rs + k_{opt}^{(q)}(n-(r+1)s, d)\} \\ &\leq r + k_{opt}^{(q)}(n-(r+1), d) \\ &= (q-1)q^{m-1} - 1 + k_{opt}^{(q)}(q^{m-1}, 3) \\ &\leq q^m - m - 1, \end{aligned}$$

where the last inequality holds due to the fact that $k_{opt}^{(q)}(q^{m-1}, 3) \leq q^{m-1} - m + 1$, which follows from the Sphere packing bound. Therefore, the code $\mathcal{R}_q(1, m)^\perp$ is k-optimal.

Putting the parameters of $\mathcal{R}_q(1, m)$ into the right-hand side of the CM bound in (2), we have

$$\begin{aligned} k &\leq \min_{s\in\mathbb{Z}_+}\{rs + k_{opt}^{(q)}(n-(r+1)s, d)\} \\ &\leq r + k_{opt}^{(q)}(n-(r+1), d) \\ &= 2 + k_{opt}^{(q)}(q^m - 3, (q-1)q^{m-1}) \\ &\leq m + 1, \end{aligned}$$

where the last inequality holds due to the fact that $k_{opt}^{(q)}(q^m - 3, (q-1)q^{m-1}) \leq m - 1$, which follows from the Plotkin bound. Therefore, $\mathcal{R}_q(1, m)$ is k-optimal. This completes the proof. □

In this section, we identified two classes of affine-invariant codes $\mathcal{R}_1(1, m)$ and $\mathcal{R}_q(1, m)^\perp$, which are k-optimal. It would be nice if other classes of d-optimal or k-optimal affine-invariant codes could be found. Notice that many classes of affine-invariant codes are documented in [13]. In addition, there are other classes of affine-invariant codes with very good locality properties and higher rate than RM codes [22].

4.3 The minimum locality of ovoid codes

In the projective space PG(3, GF(q)) with $q > 2$, an ovoid $\mathcal{V}$ is a set of q^2+1 points such that no three of them are collinear (i.e., on the same line). In other words, an ovoid is a (q^2+1)-cap (a cap with $q^2 + 1$ points) in PG(3, GF(q)), and thus a maximal cap. Two ovoids are said to be equivalent if there is a collineation (i.e., automorphism) of PG(3, GF(q)) that sends one to the other.

A classical ovoid $\mathcal{V}$ can be defined as the set of all points given by

$$\mathcal{V} = \{(0, 0, 1, 0)\} \cup \{(x, y, x^2 + xy + ay^2, 1) : x, y \in \mathrm{GF}(q)\}$$

where $a \in \mathrm{GF}(q)$ is such that the polynomial $x^2 + x + a$ has no root in GF(q). Such ovoid is called an elliptic quadric, as the points come from a non-degenerate elliptic quadratic form.

For $q = 2^{2e+1}$ with $e \geq 1$, there is an ovoid which is not an elliptic quadric, and is called the Tits ovoid. It is defined by

$$\mathcal{T} = \{(0, 0, 1, 0)\} \cup \{(x, y, x^\sigma + xy + y^{\sigma+2}, 1) : x, y \in \mathrm{GF}(q)\},$$

where $\sigma = 2^{e+1}$.

Let $\mathcal{V}$ be an ovoid in PG(3, GF(q)) with $q > 2$. Denote $\mathcal{V} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{q^2+1}\}$, where each $\mathbf{v}_i$ is a column vector in GF$(q)^4$. Let $\mathcal{C}_\mathcal{V}$ be the linear code over GF(q) with generator matrix $G_\mathcal{V} = [\mathbf{v}_1 \mathbf{v}_2 \ldots \mathbf{v}_{q^2+1}]$. It is known that $\mathcal{C}_\mathcal{V}$ is a $[q^2+1, 4, q^2-q]$ code over GF(q) with weight enumerator

$$1 + (q^2-q)(q^2+1)z^{q^2-q} + (q-1)(q^2+1)z^{q^2}$$

and its dual $\mathcal{C}_\mathcal{V}^\perp$ is a $[q^2+1, q^2-3, 4]$ almost MDS code over GF(q) [13, Chapter 13]. Conversely, the set of column vectors of a generator matrix of any $[q^2+1, 4, q^2-q]$ code over GF(q) is an ovoid in PG(3, GF(q)). Hence, ovoids in PG(3, GF(q)) and $[q^2+1, 4, q^2-q]$ codes over GF(q) are equivalent in the sense that one can be used to construct the other, and a $[q^2+1, 4, q^2-q]$ code over GF(q) is called an ovoid code over GF(q).

Ovoid codes are very interesting in combinatorics, as they support 3-designs, which are documented below [13, Chapter 13].

Lemma 6 [13] *The supports of all minimum weight codewords in an ovoid code form a 3-$(q^2+1, q^2-q, (q-2)(q^2-q-1))$ design and the supports of all codewords of weight 4 in the dual of the ovoid code form a 3-$(q^2+1, 4, q-2)$ design.*

The linear locality of an ovoid code and its dual is described in the next theorem.

Theorem 10 *An ovoid code $\mathcal{C}_o$ is a $(q^2+1, 4, q; q^2-q, 3)$-LLRC and its dual $\mathcal{C}_o^\perp$ is a $(q^2+1, q^2-3, 4, q; q^2-q-1)$-LLRC. Moreover, $\mathcal{C}_o$ is k-optimal and $\mathcal{C}_o^\perp$ is d-optimal and k-optimal.*

Proof The parameters follow from Lemma 6 and Corollary 3. It is easy to check the distance optimality of $\mathcal{C}_o^\perp$. Then we check the dimension optimality of $\mathcal{C}_o^\perp$. Putting $t = 1$ into the right-hand side of the CM bound in (2), one has

$$\begin{aligned} k &\le r + k_{opt}^{(q)}(n-(r+1), d) \\ &= (q-1)q - 1 + k_{opt}^{(q)}(q+1, 4) \\ &\le q^2 - 3, \end{aligned}$$

where the last inequality holds due to the fact that $k_{opt}^{(q)}(q+1, 4) \le q-2$, which follows from the classical Singleton bound. Therefore, the code $\mathcal{C}_o^\perp$ is k-optimal.

Putting the parameters of $\mathcal{C}_o$ into the right-hand side of the CM bound in (2), one has

$$\begin{aligned} k &\le \min_{t \in \mathbb{Z}_+} \{rt + k_{opt}^{(q)}(n-(r+1)t, d)\} \\ &\le r + k_{opt}^{(q)}(n-(r+1), d) \\ &= 3 + k_{opt}^{(q)}(q^2-3, q^2-q) \\ &\le 4, \end{aligned}$$

where the last inequality holds due to the fact that $k_{opt}^{(q)}(q^2-3, q^2-q) \le 1$, which follows from the Plotkin bound. Therefore, $\mathcal{C}_o$ is k-optimal. This completes the proof. □

In [28], Liu et al. studied some shortened and punctured codes of an ovoid code, and obtained the following results.

Lemma 7 [28] *Let $q \geq 4$, and let $\mathcal{C}_o$ be a $[q^2+1, 4, q^2-q]$ code over* GF(q). *For any coordinate position $\{t_1\}$, the following hold.*

- $(\mathcal{C}_o)_{\{t_1\}}$ *is a* $[q^2, 3, q^2-q]$ *code over* GF(q) *with weight enumerator*

$$1 + q(q^2-1)z^{q^2-q} + (q-1)z^{q^2}.$$

- $((\mathcal{C}_o)_{\{t_1\}})^{\perp}$ *is a* $[q^2, q^2-3, 3]$ *almost MDS code over* GF(q).
- $((\mathcal{C}_o^{\perp})_{\{t_1\}})^{\perp}$ *is a* $[q^2, 4, q^2-q-1]$ *code over* GF(q) *with weight enumerator*

$$1 + q^2(q-1)z^{q^2-q-1} + q(q^2-1)z^{q^2-q} + q^2(q-1)z^{q^2-1} + (q-1)z^{q^2}.$$

- $(\mathcal{C}_o^{\perp})_{\{t_1\}}$ *is a* $[q^2, q^2-4, 4]$ *almost MDS code over* GF(q).

Furthermore, these codes hold 2-design.

The minimum locality of these punctured and shortened codes of ovoid codes and their duals are documented in the following theorem.

Theorem 11 *Let $q \geq 4$. Then the code $(\mathcal{C}_o)_{\{t_1\}}$ is a k-optimal $(q^2, 3, q^2-q, q; 2)$-LLRC and the code $((\mathcal{C}_o^{\perp})_{\{t_1\}})^{\perp}$ is a k-optimal $(q^2, 4, q^2-q-1, q; 3)$-LLRC. The code $(\mathcal{C}_o^{\perp})_{\{t_1\}}$ is a $(q^2, q^2-4, 4, q; q^2-q-2)$-LLRC and the code $((\mathcal{C}_o)_{\{t_1\}})^{\perp}$ is a $(q^2, q^2-3, 3, q; q^2-q-1)$-LLRC. Furthermore, $(\mathcal{C}_o^{\perp})_{\{t_1\}}$ and $((\mathcal{C}_o)_{\{t_1\}})^{\perp}$ are both d-optimal and k-optimal.*

Proof The parameters of these codes follow from Lemma 7 and Corollary 3. It is easy to verify the distance optimality of $(\mathcal{C}_o^{\perp})_{\{t_1\}}$ and $((\mathcal{C}_o)_{\{t_1\}})^{\perp}$ with respect to the Singleton-like bound. The proofs of dimension optimality of $(\mathcal{C}_o^{\perp})_{\{t_1\}}$ and $((\mathcal{C}_o)_{\{t_1\}})^{\perp}$ are similar, so we just prove the dimension optimality of $(\mathcal{C}_o^{\perp})_{\{t_1\}}$. Putting $t = 1$ into the right-hand side of the CM bound in (2), one arrives at

$$\begin{aligned} k &\leq r + k_{opt}^{(q)}(n-(r+1), d) \\ &= (q-1)q - 2 + k_{opt}^{(q)}(q+1, 4) \\ &\leq q^2 - 4, \end{aligned}$$

where the last inequality holds due to the fact that $k_{opt}^{(q)}(q+1, 3) \leq q-2$, which follows from the classical Singleton bound. Therefore, the code $(\mathcal{C}_o^{\perp})_{\{t_1\}}$ is k-optimal. This completes the proof. □

Ovoid codes are very attractive in the sense that $\mathcal{C}_o^{\perp}$, $(\mathcal{C}_o^{\perp})_{\{t_1\}}$ and $((\mathcal{C}_o)_{\{t_1\}})^{\perp}$ are both d-optimal and k-optimal. Recall that ovoid codes are the same as ovoids in projective geometry. In addition, ovoid codes support 3-designs, which are related to inversive planes (also called Möbius planes) [13, Chapter 13]. Furthermore, the trace codes of some ovoid codes are also optimal [12]. These facts show that ovoid codes are really diamonds.

4.4 The minimum locality of maximal arc codes

Throughout this section, let $q = 2^m$ for some positive integer $m \geq 2$. A *maximal* (n, h)-arc $\mathcal{A}$ in the projective plane PG(2, GF(q)) is a subset of $n = hq + h - q$ points such that every line meets $\mathcal{A}$ in 0 or h points. A maximal (n, h)-arc $\mathcal{A}$ in PG(2, GF(q)) exists if and only if h divides q, where $2 \leq h < q$. Hence, in this section, we let $h = 2^i$ for some i with $1 \leq i < m$. There are several known families of maximal arcs and the reader is referred to [13, Section 12.7] for further information.

Let $\mathcal{A}$ be a maximal (n, h)-arc in PG$(2, \mathrm{GF}(q))$. Denote $\mathcal{A} = \{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}$, where each $\mathbf{a}_i$ is a column vector in $\mathrm{GF}(q)^3$. Let $\mathcal{C}(\mathcal{A})$ denote the linear code over $\mathrm{GF}(q)$ with generator matrix $G_{\mathcal{A}} = [\mathbf{a}_1 \mathbf{a}_2 \ldots \mathbf{a}_n]$. We call $\mathcal{C}(\mathcal{A})$ a *maximum arc code*. The following theorem was proved in [13, Section 12.7].

Theorem 12 [13] *Let $q = 2^m$ for any $m \geq 3$ and $h = 2^i$ with $2 \leq i < m$. Let $\mathcal{A}$ be a maximal (n, h)-arc in* PG$(2, \mathrm{GF}(q))$. *Then the maximum arc code $\mathcal{C}(\mathcal{A})$ has parameters $[n,\ 3,\ n-h]$ and weight enumerator*

$$1 + \frac{(q^2-1)n}{h} z^{n-h} + \frac{(q^3-1)h - (q^2-1)n}{h} z^n, \tag{16}$$

where $n = hq + h - q$. The dual code $\mathcal{C}(\mathcal{A})^{\perp}$ has parameters $[n, n-3, 3]$. Furthermore, the minimum weight codewords in both $\mathcal{C}(\mathcal{A})$ and $\mathcal{C}(\mathcal{A})^{\perp}$ support a 2-design.

Theorem 13 *Let $q = 2^m$ for any $m \geq 3$ and $h = 2^i$ with $2 \leq i < m$. Let $\mathcal{A}$ be a maximal (n, h)-arc in* PG$(2, \mathrm{GF}(q))$. *Then $\mathcal{C}(\mathcal{A})$ is a k-optimal $(n, 3, n-h, q; 2)$-LLRC and $\mathcal{C}(\mathcal{A})^{\perp}$ is a d-optimal and k-optimal $(n, n-3, 3, q; n-h-1)$-LLRC.*

Proof It follows from Corollary 3 and Theorem 12 that $\mathcal{C}(\mathcal{A})^{\perp}$ has minimum locality $d(\mathcal{C}(\mathcal{A})) - 1$ and $\mathcal{C}(\mathcal{A})$ has minimum locality $d(\mathcal{C}(\mathcal{A})^{\perp}) - 1$. The parameters of the two codes then follow from Theorem 12. It is straightforward to verify that the parameters of $\mathcal{C}(\mathcal{A})^{\perp}$ meet the Singleton-like bound. We now prove the dimension optimality of $\mathcal{C}(\mathcal{A})^{\perp}$. Putting $t = 1$ into the right-hand side of the CM bound in (2), we have

$$\begin{aligned} k &\leq r + k_{opt}^{(q)}(n - (r+1), d) \\ &= n - h - 1 + k_{opt}^{(q)}(h, 3) \\ &\leq n - 3, \end{aligned}$$

where the last inequality holds due to the fact that $k_{opt}^{(q)}(h, 3) \leq h - 2$, which follows from the classical Singleton bound. Therefore, the $\mathcal{C}(\mathcal{A})^{\perp}$ is k-optimal.

Plugging the parameters of $\mathcal{C}(\mathcal{A})$ into the right-hand side of the CM bound in (2), one has

$$\begin{aligned} k &\leq \min_{t \in \mathbb{Z}_+} \{rt + k_{opt}^{(q)}(n - (r+1)t, d)\} \\ &\leq r + k_{opt}^{(q)}(n - (r+1), d) \\ &= 2 + k_{opt}^{(q)}(n-3, n-h) \\ &\leq 3, \end{aligned}$$

where the last inequality holds due to the fact that $k_{opt}^{(q)}(n-3, n-h) \leq 1$, which follows from the Plotkin bound. Therefore, $\mathcal{C}(\mathcal{A})$ is k-optimal. This completes the proof. □

A family of extended cyclic codes with the parameters of the code in Theorem 12 were documented in [13, Section 12.8]. We are interested in maximal arc codes, as they are k-optimal LLRCs and their duals are d-optimal and k-optimal LLRCs.

5 The minimum locality of near MDS codes

5.1 Some general theory on the minimum locality of near MDS codes

The Singleton defect of an $[n, k, d]$ code $\mathcal{C}$ is defined by $\mathrm{def}(\mathcal{C}) = n - k + 1 - d$. Thus, MDS codes are codes with defect 0. A code $\mathcal{C}$ is said to be almost MDS (AMDS for short) if it has defect 1. Hence, AMDS codes have parameters $[n, k, n - k]$. A code is said to be near MDS (NMDS for short) if the code and its dual code both are AMDS. By definition, $\mathcal{C}$ is near MDS if and only if $\mathcal{C}^\perp$ is so. Then an $[n, k]$ code $\mathcal{C}$ over $\mathrm{GF}(q)$ is NMDS if and only if $d(\mathcal{C}) + d(\mathcal{C}^\perp) = n$ [15]. The following lemma will be needed later [15, 17].

Lemma 8 [15, 17] *Let $\mathcal{C}$ be an $[n, k, n - k]$ AMDS code over $\mathrm{GF}(q)$.*

- *If $k \geq 2$, then $\mathcal{C}$ is generated by its codewords of weight $n - k$ and $n - k + 1$.*
- *If $k \geq 2$ and $n - k > q$, then $\mathcal{C}$ is generated by its minimum weight codewords.*

Theorem 14 *Let $\mathcal{C}$ be a nontrivial NMDS code. Then the minimum locality of $\mathcal{C}$ is either $d(\mathcal{C}^\perp) - 1$ or $d(\mathcal{C}^\perp)$. In particular, the minimum locality of $\mathcal{C}$ is $d(\mathcal{C}^\perp) - 1$ if the minimum weight codewords in $\mathcal{C}^\perp$ generate $\mathcal{C}^\perp$.*

Proof Let n denote the length of $\mathcal{C}$. If $\mathcal{C}^\perp$ is generated by its minimum weight codewords, it then follows from Theorem 4 that $\mathcal{C}$ has minimum locality $d(\mathcal{C}^\perp) - 1$. Assume now that all the minimum weight codewords in $\mathcal{C}^\perp$ do not generate $\mathcal{C}^\perp$. Since $\mathcal{C}^\perp$ is nontrivial, $\dim(\mathcal{C}^\perp) \geq 2$. It then follows from Lemma 8 that $\mathcal{C}^\perp$ is generated by all the codewords of weights $d(\mathcal{C}^\perp)$ and $d(\mathcal{C}^\perp) + 1$. If the union of the supports of all the codes of weights $d(\mathcal{C}^\perp)$ and $d(\mathcal{C}^\perp) + 1$ does not contain $i \in [n]$, then $\mathcal{C}$ must be have a zero coordinate in position i. This means that $d(\mathcal{C}) = 1$, which contradicts to our assumption that $\mathcal{C}$ is nontrivial. It then follows from Theorem 3 that the minimum locality of $\mathcal{C}$ is either $d(\mathcal{C}^\perp) - 1$ or $d(\mathcal{C}^\perp)$. □

MDS codes are very interesting due to the following theorem whose proof is straightforward by following the assumptions and the parameters of NMDS codes.

Theorem 15 *If a nontrivial NMDS code $\mathcal{C}$ over $\mathrm{GF}(q)$ with parameters $[n, k, n - k]$ has minimum locality $d(\mathcal{C}^\perp) - 1$, then $\mathcal{C}$ is a d-optimal and k-optimal $(n, k, n - k, q; k - 1)$-LLRC with respect to the Singleton-like bound and the CM bound, respectively.*

If a nontrivial NMDS code $\mathcal{C}$ over $\mathrm{GF}(q)$ with parameters $[n, k, n - k]$ has minimum locality $d(\mathcal{C}^\perp)$, then $\mathcal{C}$ is an almost d-optimal and k-optimal $(n, k, n - k, q; k)$-LLRC with respect to the Singleton-like bound and the CM bound, respectively.

We will demonstrate later that some nontrivial NMDS codes $\mathcal{C}$ have minimum locality $d(\mathcal{C}^\perp) - 1$ and some nontrivial NMDS codes $\mathcal{C}$ indeed have minimum locality $d(\mathcal{C}^\perp)$. Of course, nontrivial NMDS codes $\mathcal{C}$ with minimum locality $d(\mathcal{C}^\perp) - 1$ are better. Therefore, we are more interested in nontrivial MDS code $\mathcal{C}$ with minimum locality $d(\mathcal{C}^\perp) - 1$.

Corollary 9 *Let $\mathcal{C}$ be a nontrivial NMDS code over $\mathrm{GF}(q)$ with parameters $[n, k, n - k]$. If $\mathcal{C}^\perp$ does not have a codeword of weight $d(\mathcal{C}^\perp) + 1$, then $\mathcal{C}$ is a d-optimal and k-optimal $(n, k, n - k, q; k - 1)$-LLRC.*

Proof By definition, $\mathcal{C}^\perp$ has parameters $[n, n - k, k]$. Since $\mathcal{C}$ is nontrivial, $d(\mathcal{C}^\perp) = k \geq 2$. By Lemma 8, $\mathcal{C}^\perp$ is generated by its codewords of weights $d(\mathcal{C}^\perp)$ and $d(\mathcal{C}^\perp) + 1$. Since $\mathcal{C}^\perp$ does not have a codeword of weight $d(\mathcal{C}^\perp) + 1$, $\mathcal{C}^\perp$ is generated by its codewords of weight $d(\mathcal{C}^\perp)$. By Theorem 4, $\mathcal{C}$ has minimum locality $d(\mathcal{C}^\perp) - 1$. The desired conclusion then follows from Theorem 15. □

We remark that under the condition of Corollary 9, it can be proved that the minimum weight codewords in $\mathcal{C}^{\perp}$ support a 1-design [14]. To prove another result about the minimum locality, we need the following lemma [17].

Lemma 9 [17] *Let $\mathcal{C}$ be an NMDS code. Then for every minimum weight codeword $\mathbf{c}$ in $\mathcal{C}$, there exists, up to a multiple, a unique minimum weight codeword $\mathbf{c}^{\perp}$ in $\mathcal{C}^{\perp}$ such that $\mathrm{supp}(\mathbf{c}) \cap \mathrm{supp}(\mathbf{c}^{\perp}) = \emptyset$. In particular, $\mathcal{C}$ and $\mathcal{C}^{\perp}$ have the same number of minimum weight codewords.*

We now provide the following result, which is useful in certain cases.

Theorem 16 *Let $\mathcal{C}$ be an NMDS code and let $d^{\perp} = d(\mathcal{C}^{\perp})$. If*

$$\bigcap_{S \in \mathcal{B}_{d^{\perp}}(\mathcal{C}^{\perp})} S = \emptyset,$$

then $\mathcal{C}^{\perp}$ has minimum locality $d(\mathcal{C}) - 1$.

Proof Let $\mathcal{C}$ have parameters $[n, k, d]$ with $d = n - k$. It follows from Lemma 9 that each set in $\mathcal{B}_d(\mathcal{C})$ is the complement of a set in $\mathcal{B}_{d^{\perp}}(\mathcal{C}^{\perp})$ and vice versa. If an integer $i \in [n]$ is not in $\cup_{S \in \mathcal{B}_d(\mathcal{C})}$, we then deduce that it must be in $\cap_{S \in \mathcal{B}_{d^{\perp}}(\mathcal{C}^{\perp})}$. This is contrary to the assumption that $\cap_{S \in \mathcal{B}_{d^{\perp}}(\mathcal{C}^{\perp})} S = \emptyset$. Consequently,

$$\bigcup_{S \in \mathcal{B}_d(\mathcal{C})} = [n].$$

The desired conclusion then follows from Corollary 1. □

5.2 The minimum locality of NMDS cyclic codes

According to Corollary 2 and Theorem 15, every nontrivial NMDS cyclic code $\mathcal{C}$ and its dual are both d-optimal and k-optimal LLRCs. In this subsection, we document such NMDS cyclic codes. We begin with the following example.

Example 1 The ternary Golay code $\mathcal{C}_{\mathrm{Golay}}$ has parameters $[11, 6, 5]$ and weight enumerator $1 + 132z^5 + 132z^6 + 330z^8 + 110z^9 + 24z^{11}$. The dual code $\mathcal{C}^{\perp}_{\mathrm{Golay}}$ has parameters $[11, 5, 6]$ and weight enumerator $1 + 132z^6 + 110z^9$. Hence, the ternary Golay code is an nontrivial NMDS. It is well known that $\mathcal{C}_{\mathrm{Golay}}$ is a BCH code, an irreducible cyclic code, and also a quadratic residue code. Therefore both codes are d-optimal and k-optimal LLRCs.

The following two theorems report infinite classes of d-optimal and k-optimal LLRCs from some known NMDS.

Theorem 17 *Let $q = 2^s$ with $s \geq 4$ being even or $q = 3^s$ with $s \geq 2$. Then the narrow-sense BCH code $\mathcal{C}_{(q,q+1,3,1)}$ over $\mathrm{GF}(q)$ is a d-optimal and k-optimal $(q, q-3, 4, q; q-4)$-LLRC. In addition, its dual code $\mathcal{C}^{\perp}_{(q,q+1,3,1)}$ is a d-optimal and k-optimal $(q, 4, q-3, q; 3)$-LLRC.*

Proof The desired conclusions follow directly from Corollary 2, Theorem 15, and Theorems 21 and 23 in [14]. □

Theorem 18 *Let $m \geq 5$ be odd and $q = 2^m$. Then the narrow-sense BCH code $\mathcal{C}_{(q,q+1,4,1)}$ over $\mathrm{GF}(q)$ is a d-optimal and k-optimal $(q+1, q-5, 6, q; q-6)$-LLRC. In addition, its dual code $\mathcal{C}^{\perp}_{(q,q+1,4,1)}$ a d-optimal and k-optimal $(q+1, 6, q-5, q; 5)$-LLRC.*

Proof The desired conclusions follow directly from Corollary 2, Theorem 15, and Theorems 34, 35, and 37 in [35]. □

Many NMDS codes have been constructed (see, for example, [1, 2, 10, 15–18, 25, 30, 37–39]). It is worthwhile to check if some of them are cyclic.

5.3 The minimum locality of the extended codes of some NMDS cyclic codes

In this section, we investigate the minimum locality of the extended codes of some NMDS cyclic codes, and will make use of Theorem 5.

Theorem 19 *Let $q = 3^s$ with $s \geq 2$. Then the extended code $\overline{\mathcal{C}_{(q,q+1,3,1)}}$ over* GF(q) *has parameters* $[q+2, q-3, 5]$*, and its dual code* $(\overline{\mathcal{C}_{(q,q+1,3,1)}})^{\perp}$ *has parameters* $[q+2, 5, q-3]$*. Furthermore,* $(\overline{\mathcal{C}_{(q,q+1,3,1)}})^{\perp}$ *is a d-optimal and k-optimal* $(q+2, 5, q-3, q; 4)$*-LLRC.*

Proof Put $n = q + 1$. Let α be a generator of GF$(q^2)^*$ and $\beta = \alpha^{q-1}$. Then β is an n-th root of unity in GF(q^2). Let $\mathbb{M}_{\beta}(x)$ and $\mathbb{M}_{\beta^2}(x)$ denote the minimal polynomial of β and β^2 over GF(q), respectively. Note that $\mathbb{M}_{\beta}(x)$ has only roots β and β^q and $\mathbb{M}_{\beta^2}(x)$ has roots β^2 and β^{q-1}. We deduce that $\mathbb{M}_{\beta}(x)$ and $\mathbb{M}_{\beta^2}(x)$ are distinct irreducible polynomials of degree 2. By definition, $g(x) := \mathbb{M}_{\beta}(x)\mathbb{M}_{\beta^2}(x)$ is the generator polynomial of $\mathcal{C}_{(q,q+1,3,1)}$. Put $\gamma = \beta^{-1}$. Then $\gamma^{q+1} = \beta^{-(q+1)} = 1$. It then follows from Delsarte's theorem that the trace expression of $\mathcal{C}^{\perp}_{(q,q+1,3,1)}$ is given by

$$\mathcal{C}^{\perp}_{(q,q+1,3,1)} = \{\mathbf{c}_{(a,b)} : a, b \in \mathrm{GF}(q^2)\},$$

where $\mathbf{c}_{(a,b)} = (\mathrm{Tr}_{q^2/q}(a\gamma^i + b\gamma^{2i}))_{i=0}^{q}$. Define

$$H = \begin{bmatrix} 1 & \gamma^1 & \gamma^2 & \dots & \gamma^q \\ 1 & \gamma^2 & \gamma^4 & \dots & \gamma^{2q} \end{bmatrix}.$$

It is easily seen that H is a parity-check matrix of $\mathcal{C}_{(q,q+1,3,1)}$, i.e.,

$$\mathcal{C}_{(q,q+1,3,1)} = \{\mathbf{c} \in \mathrm{GF}(q)^{q+1} : \mathbf{c}H^T = \mathbf{0}\}.$$

Note that $\{1, \gamma\}$ is a basis of GF(q^2) over GF(q). Every γ^j can be expressed as $\gamma^j = a_{j,0} + a_{j,1}\gamma$, where $a_{j,0} \in \mathrm{GF}(q)$ and $a_{j,1} \in \mathrm{GF}(q)$. Later, H refers to the $4 \times n$ matrix over GF(q) defined by these $a_{j,i}$.

From Lemma 3, $\overline{\mathcal{C}_{(q,q+1,3,1)}}$ has parameters $[q+2, q-3, d(\overline{\mathcal{C}_{(q,q+1,3,1)}})]$ and the parity-check matrix $\overline{H}$ of $\overline{\mathcal{C}_{(q,q+1,3,1)}}$ is

$$\overline{H} = \begin{bmatrix} \mathbf{1} & 1 \\ H & \mathbf{0} \end{bmatrix},$$

where $\mathbf{1} = (1, 1, \dots, 1)$ and $\mathbf{0} = (0, 0, \dots, 0)^T$.

Note that $d(\overline{\mathcal{C}_{(q,q+1,3,1)}}) = d(\mathcal{C}_{(q,q+1,3,1)}) + 1$ or $d(\overline{\mathcal{C}_{(q,q+1,3,1)}}) = d(\mathcal{C}_{(q,q+1,3,1)})$. Now we prove that $d(\overline{\mathcal{C}_{(q,q+1,3,1)}}) = d(\mathcal{C}_{(q,q+1,3,1)}) + 1 = 5$. Let U_{q+1} denote the set of all $(q+1)$-th roots of unity in GF(q^2). Suppose $d(\overline{\mathcal{C}_{(q,q+1,3,1)}}) = d(\mathcal{C}_{(q,q+1,3,1)}) = 4$. Then there are four pairwise distinct elements x, y, z, w in U_{q+1} such that

$$a\begin{bmatrix} 1 \\ x \\ x^2 \end{bmatrix} + b\begin{bmatrix} 1 \\ y \\ y^2 \end{bmatrix} + c\begin{bmatrix} 1 \\ z \\ z^2 \end{bmatrix} + d\begin{bmatrix} 1 \\ w \\ w^2 \end{bmatrix} = 0, \tag{17}$$

where $a, b, c, d \in \mathrm{GF}(q)^*$. Raising to the q-th power both sides of the equation $ax + by + cz + dw = 0$ yields

$$ax^{-1} + by^{-1} + cz^{-1} + dw^{-1} = 0. \tag{18}$$

Combining (17) and (18) gives

$$a\begin{bmatrix} x^{-1} \\ 1 \\ x \\ x^2 \end{bmatrix} + b\begin{bmatrix} y^{-1} \\ 1 \\ y \\ y^2 \end{bmatrix} + c\begin{bmatrix} z^{-1} \\ 1 \\ z \\ z^2 \end{bmatrix} + d\begin{bmatrix} w^{-1} \\ 1 \\ w \\ w^2 \end{bmatrix} = 0.$$

It then follows that

$$\begin{vmatrix} x^{-1} & y^{-1} & z^{-1} & w^{-1} \\ 1 & 1 & 1 & 1 \\ x & y & z & w \\ x^2 & y^2 & z^2 & w^2 \end{vmatrix} = \frac{(x-y)(x-z)(x-w)(y-z)(y-w)(z-w)}{xyzw} = 0.$$

This is contrary to our assumption that x, y, z, w are pairwise distinct. Hence,

$$d(\overline{\mathcal{C}_{(q,q+1,3,1)}}) = d(\mathcal{C}_{(q,q+1,3,1)}) + 1 = 5. \tag{19}$$

We now prove the following equalities:

$$(\overline{\mathcal{C}_{(q,q+1,3,1)}})^{\perp} = \widetilde{\overline{\mathcal{C}^{\perp}_{(q,q+1,3,1)}}}, \tag{20}$$

where $\widetilde{D}$ denotes the augmented code of a code D. It is easily verified that the sum of all coordinates in each codeword

$$\mathbf{c}_{(a,b)} = (\mathrm{Tr}_{q^2/q}(a\gamma^i + b\gamma^{2i}))_{i=0}^{q}$$

is zero, as both γ and γ^2 are n-th roots of unity. Consequently, $\overline{\mathcal{C}^{\perp}_{(q,q+1,3,1)}}$ is generated by the matrix $[H\mathbf{0}]$, where H is the $4 \times n$ matrix over $\mathrm{GF}(q)$ defined above. Then the equality in (20) follows from Lemma 3. By (19), we conclude that the all-one vector is not a codeword in $\mathcal{C}^{\perp}_{(q,q+1,3,1)}$. It then follows from (20) that

$$\dim((\overline{\mathcal{C}_{(q,q+1,3,1)}})^{\perp}) = 1 + \dim(\mathcal{C}^{\perp}_{(q,q+1,3,1)}) = 5.$$

Now we prove $d((\overline{\mathcal{C}_{(q,q+1,3,1)}})^{\perp}) = q - 3$. Note that $(\overline{\mathcal{C}_{(q,q+1,3,1)}})^{\perp}$ has generator matrix $\overline{H}$ and

$$\mathcal{C}^{\perp}_{(q,q+1,3,1)} = \{\mathbf{c}_{(a,b)} : a, b \in \mathrm{GF}(q^2)\},$$

where $\mathbf{c}_{(a,b)} = (\mathrm{Tr}_{q^2/q}(a\gamma^i + b\gamma^{2i}))_{i=0}^{q}$. It follows from (20) that the codewords in $(\overline{\mathcal{C}_{(q,q+1,3,1)}})^{\perp}$ have the form $(\mathbf{c}_{(a,b)} + c\mathbf{1}, c)$, where $c \in \mathrm{GF}(q)$. Let $u \in U_{q+1}$. Then

$$\begin{aligned} \mathrm{Tr}_{q^2/q}(au + bu^2) + c &= au + bu^2 + a^q u^{-1} + b^q u^{-2} + c \\ &= u^{-2}(bu^4 + au^3 + a^q u + b^q + cu^2). \end{aligned}$$

Hence, there are at most four $u \in U_{q+1}$ such that $\mathrm{Tr}_{q^2/q}(au + bu^2) + c = 0$ if $(a, b, c) \neq (0, 0, 0)$. As a result, for $(a, b, c) \neq (0, 0, 0)$, we have

$$\mathrm{wt}((\mathbf{c}_{(a,b)} + c\mathbf{1}, c)) = \mathrm{wt}(\mathbf{c}_{(a,b)} + c\mathbf{1}) + 1 \geq q + 1 - 4 + 1 = q - 2,$$

and for $(a, b) \neq (0, 0)$, $c = 0$, we have

$$\mathrm{wt}((\mathbf{c}_{(a,b)}, 0)) = \mathrm{wt}(\mathbf{c}_{(a,b)}) \geq q + 1 - 4 = q - 3.$$

This means that $d((\overline{\mathcal{C}_{(q,q+1,3,1)}})^{\perp}) \geq q-3$. If $d((\overline{\mathcal{C}_{(q,q+1,3,1)}})^{\perp}) = q-2$, then $(\overline{\mathcal{C}_{(q,q+1,3,1)}})^{\perp}$ would be an MDS code and $\overline{\mathcal{C}_{(q,q+1,3,1)}}$ would also be an MDS code, which leads to a contradiction. We then conclude that $d((\overline{\mathcal{C}_{(q,q+1,3,1)}})^{\perp}) = q - 3$. Now both $\overline{\mathcal{C}_{(q,q+1,3,1)}}$ and its dual are AMDS. Since $d(\overline{\mathcal{C}_{(q,q+1,3,1)}}) = 5 = d(\mathcal{C}_{(q,q+1,3,1)})+1$, by Theorem 5 we deduce that $(\overline{\mathcal{C}_{(q,q+1,3,1)}})^{\perp}$ has locality $d(\mathcal{C}_{(q,q+1,3,1)}) = 4$. The optimality of $(\overline{\mathcal{C}_{(q,q+1,3,1)}})^{\perp}$ then follows from Theorem 15. □

Theorem 20 *Let $q = 2^s$ with $s \geq 4$ being even. Then the extended code $\overline{\mathcal{C}_{(q,q+1,3,1)}}$ over* GF(q) *has parameters* $[q + 2, q - 3, 5]$*, and its dual code $(\overline{\mathcal{C}_{(q,q+1,3,1)}})^{\perp}$ has parameters* $[q + 2, 5, q - 3]$*. Furthermore, $(\overline{\mathcal{C}_{(q,q+1,3,1)}})^{\perp}$ is a d-optimal and k-optimal* $(q + 2, 5, q - 3, q; 4)$*-LLRC.*

Proof The proof of this theorem is similar to that of Theorem 19 and is omitted. □

Notice that the two theorems above provide not only two families of d-optimal and k-optimal LLRCs, but also two families of NMDS codes with new parameters.

5.4 The linear locality of some NMDS codes from oval polynomials

Oval polynomials were used to construct NMDS codes in [39]. These NMDS codes are not cyclic. In this subsection, we study the minimum locality of some of them. To introduce these codes, we need oval polynomials. Throughout this subsection, let $q = 2^m$, where $m \geq 3$.

An oval polynomial $f(x)$ on GF(q) is a polynomial such that

- f is a permutation polynomial of GF(q) with $\deg(f) < q$ and $f(0) = 0$, $f(1) = 1$; and
- for each $a \in$ GF(q), $g_a(x) := (f(x + a) + f(a))x^{q-2}$ is also a permutation polynomial of GF(q).

Every oval polynomial f can be used to construct a hyperoval in PG(2, GF(q)) [13, Chapter 12]. The following is a list of known infinite families of oval polynomials in the literature.

Theorem 21 [13] *Let $m \geq 4$ be an integer. The following are oval polynomials of* GF(q)*, where $q = 2^m$.*

- *The translation polynomial $f(x) = x^{2^h}$, where* $\gcd(h, m) = 1$.
- *The Segre polynomial $f(x) = x^6$, where m is odd.*
- *The Glynn oval polynomial $f(x) = x^{3\times 2^{(m+1)/2}+4}$, where m is odd.*
- *The Glynn oval polynomial $f(x) = x^{2^{(m+1)/2}+2^{(m+1)/4}}$ for $m \equiv 3 \pmod 4$.*
- *The Glynn oval polynomial $f(x) = x^{2^{(m+1)/2}+2^{(3m+1)/4}}$ for $m \equiv 1 \pmod 4$.*
- *The Cherowitzo oval polynomial $f(x) = x^{2^e} + x^{2^e+2} + x^{3\times 2^e+4}$, where $e = (m + 1)/2$ and m is odd.*
- *The Payne oval polynomial $f(x) = x^{\frac{2^{m-1}+2}{3}} + x^{2^{m-1}} + x^{\frac{3\times 2^{m-1}-2}{3}}$, where m is odd.*
- *The Subiaco polynomial*

$$f_a(x) = ((a^2(x^4 + x) + a^2(1 + a + a^2)(x^3 + x^2))(x^4 + a^2x^2 + 1)^{2^m-2} + x^{2^{m-1}},$$

where $\mathrm{Tr}_{q/2}(1/a) = 1$ *and* $a \notin$ GF(4) *if* $m \equiv 2 \bmod 4$.

– *The Adelaide oval polynomial*

$$f(x) = \frac{T(\beta^m)(x+1)}{T(\beta)} + \frac{T((\beta x + \beta^q)^m)}{T(\beta)(x + T(\beta)x^{2^{m-1}} + 1)^{m-1}} + x^{2^{m-1}},$$

where $m \geq 4$ *is even,* $\beta \in \mathrm{GF}(q^2) \setminus \{1\}$ *with* $\beta^{q+1} = 1$, $m \equiv \pm(q-1)/3 \pmod{q+1}$, *and* $T(x) = x + x^q$.

The following lemma will be needed later [31].

Lemma 10 [31] *A polynomial* f *over* $\mathrm{GF}(q)$ *with* $f(0) = 0$ *is an oval polynomial if and only if* $f_u := f(x) + ux$ *is* 2*-to-*1 *for every* $u \in \mathrm{GF}(q)^*$, *where* 2*-to-*1 *means that* $|f_u^{-1}(b)| = 2$ *for any element* b *in the image of* f_u.

5.4.1 NMDS codes with parameters $[q+3, 3, q]$ from oval polynomials

Let f be a polynomial over $\mathrm{GF}(q)$ with $f(0) = 0$ and $f(1) = 1$. Let α be a generator of $\mathrm{GF}(q)^*$. Define

$$\bar{B}_f = \begin{bmatrix} f(0) & f(\alpha^0) & f(\alpha^1) & \ldots & f(\alpha^{q-2}) & 1 & 0 & 1 \\ 0 & \alpha^0 & \alpha^1 & \ldots & \alpha^{q-2} & 0 & 1 & 1 \\ 1 & 1 & 1 & \ldots & 1 & 0 & 0 & 0 \end{bmatrix}.$$

Let $\bar{\mathcal{C}}_f$ denote the linear code over $\mathrm{GF}(q)$ with generator matrix $\bar{B}_f$. The following theorem was proved in [39].

Theorem 22 [39] *Let* $m \geq 3$, *and let* f *be an oval polynomial over* $\mathrm{GF}(q)$. *Then the code* $\bar{\mathcal{C}}_f$ *is an NMDS code over* $\mathrm{GF}(q)$ *with parameters* $[q+3, 3, q]$ *and weight enumerator*

$$1 + \frac{(q-1)(q+2)}{2} z^q + \frac{(q-1)q(q+2)}{2} z^{q+1} + \frac{(q-1)q}{2} z^{q+2} + \frac{(q-2)(q-1)q}{2} z^{q+3}.$$

Theorem 23 *The dual code* $(\bar{\mathcal{C}}_f)^\perp$ *is a* d*-optimal and* k*-optimal* $(q+3, q, 3, q; q-1)$*-LLRC.*

Proof Let $\mathbf{c}_1$, $\mathbf{c}_2$ and $\mathbf{c}_3$ denote the first, second and third rows of the generator matrix $\bar{B}_f$, respectively. By the definition of the polynomial f, it is easily seen that $\mathbf{c}_1 + \mathbf{c}_3$ and $\mathbf{c}_2 + \mathbf{c}_3$ are two minimum weight codewords in $\bar{\mathcal{C}}_f$. In addition, the supports of these two codewords are $[q+3] \setminus \{q+1\}$ and $[q+3] \setminus \{q\}$, respectively. Clearly,

$$([q+3] \setminus \{q+1\}) \cup ([q+3] \setminus \{q\}) = [q+3].$$

By Corollary 1, $(\bar{\mathcal{C}}_f)^\perp$ has minimum locality $d(\bar{\mathcal{C}}_f) - 1 = q - 1$. The desired conclusion then follows from Theorem 15. □

The minimum locality of $\bar{\mathcal{C}}_f$ is given below.

Theorem 24 *The NMDS code* $\bar{\mathcal{C}}_f$ *is a* d*-optimal and* k*-optimal* $(q+3, 3, q, q; 2)$*-LLRC.*

Proof Recall we use the elements in the set $[q+3] = \{0, 1, \ldots, q+2\}$ to index the coordinate positions of the code $\bar{\mathcal{C}}_f$ and its dual. Since all the codewords of weight 3 in $(\bar{\mathcal{C}}_f)^\perp$ were characterised in [39], we outline a proof here only. Notice that the union of the supports of all the codewords of weight 3 in $(\bar{\mathcal{C}}_f)^\perp$ specified in Case 1 of the proof of Theorem 8 in [39] is $\{0, 1, \ldots, q-1, q+2\}$, and the union of the supports of all the codewords of weight 3

in $(\bar{\mathcal{C}}_f)^\perp$ specified in Case 4 of the proof of Theorem 8 in [39] is $\{q, q+1, q+2\}$. It then follows that

$$\bigcup_{S \in \mathcal{B}_3((\bar{\mathcal{C}}_f)^\perp)} S = [q+3].$$

By Corollary 1, $\bar{\mathcal{C}}_f$ has minimum locality $d((\bar{\mathcal{C}}_f)^\perp) - 1 = 2$. The desired conclusion then follows from Theorem 15. □

We remark that the NMDS code $\bar{\mathcal{C}}_f$ is an extended hyperoval code (see for example [13, Section 12.2]). The reader is referred to [39] for detail.

5.4.2 NMDS codes with parameters $[q+1, 3, q-2]$ from oval polynomials

Let f be a polynomial over GF(q) with $f(0) = 0$ and $f(1) = 1$. Let α be a generator of GF(q)*. Define

$$G_f = \begin{bmatrix} f(\alpha^0) & f(\alpha^1) & \dots & f(\alpha^{q-2}) & 0 & 1 \\ \alpha^0 & \alpha^1 & \dots & \alpha^{q-2} & 1 & 0 \\ 1 & 1 & \dots & 1 & 1 & 1 \end{bmatrix}. \tag{21}$$

Let $\mathcal{C}_f$ denote the linear code over GF(q) with generator matrix G_f. The following result was proved in [39].

Theorem 25 [39] *Let $m \geq 3$ be odd and let $f(x)$ be an oval polynomial over* GF(q) *with coefficients in* GF(2). *Then $\mathcal{C}_f$ is a $[q+1, 3, q-2]$ NMDS code over* GF(q) *with weight enumerator*

$$A(z) = 1 + (q-1)(q-2)z^{q-2} + \frac{(q-1)(q^2-5q+12)}{2} z^{q-1} + (q-1)(4q-5)z^q + \frac{(q-1)(q^2-3q+4)}{2} z^{q+1}.$$

This class of NMDS codes are very important to us, as they demonstrate that some non-trivial NMDS codes $\mathcal{C}$ indeed have minimum locality $d(\mathcal{C}^\perp)$ rather than $d(\mathcal{C}^\perp) - 1$. We will need the following lemma later.

Lemma 11 *Let $m \geq 3$ be odd and let $f(x)$ be an oval polynomial over* GF(q) *with coefficients in* GF(2). *Then $f(x) + x + 1 = 0$ does not have a solution $x \in$ GF(q).*

Proof By the definition of oval polynomials, 0 and 1 are not not solutions of the equation $f(x) + x + 1 = 0$. Suppose that $f(a) + a + 1 = 0$ for some $a \in \text{GF}(q) \setminus \{0, 1\}$. Since $f(x)$ has coefficients in GF(2), we have $f(a^2) + a^2 + 1 = 0$ and $f(a^4) + a^4 + 1 = 0$. It then follows from Lemma 10 that the set $\{a, a^2, a^4\}$ has cardinality at most 2. If $a = a^2$, then $a \in \{0, 1\}$, which contradicts to the assumption that $a \in \text{GF}(q) \setminus \{0, 1\}$. If $a = a^4$, then $a = 0$ or $a^3 = 1$. If $a^3 = 1$, we have $a = 1$ as $\gcd(3, q-1) = 1$. Hence, $a = a^4$ implies that $a \in \{0, 1\}$, which contradicts to the assumption that $a \in \text{GF}(q) \setminus \{0, 1\}$. If $a^2 = a^4$, then $a \in \{0, 1\}$, which contradicts to the assumption that $a \in \text{GF}(q) \setminus \{0, 1\}$. This completes the proof. □

Theorem 26 *Let $m \geq 3$ be odd and let $f(x)$ be an oval polynomial over* GF(q) *with coefficients in* GF(2), *and let $\mathcal{C}_f$ be the code in Theorem 25.* Then $\mathcal{C}_f^\perp$ has minimum locality $d(\mathcal{C}) - 1$ and is a d-optimal and k-optimal $(q+1, q-2, 3, q; q-3)$-LLRC, and $\mathcal{C}_f$ has minimum locality $d(\mathcal{C}_f^\perp)$ and is an almost d-optimal and k-optimal $(q+1, 3, q-2, q; 3)$-LLRC.

Proof Since all the minimum weight codewords in $\mathcal{C}_f$ were not characterized in [39], we have to do this job here. Let $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ denote the first, second and third rows in the generator matrix G_f above.

We first consider all the codewords $\mathbf{v}_3+\mathbf{v}_2+b\mathbf{v}_1$, where $b \in \mathrm{GF}(q)^* \setminus \{1\}$. By definition,

$$\mathbf{v}_3+\mathbf{v}_2+b\mathbf{v}_1 = (b, 1+\alpha^1+bf(\alpha^1), \ldots, 1+\alpha^{q-2}+bf(\alpha^{q-2}), 0, 1+b).$$

Let $g(x) = \frac{1+x}{f(x)}$. By Lemmas 11 and 10, $g(x)$ is a 2-1 mapping from $\mathrm{GF}^*(q) \setminus \{1\}$ to itself. Therefore, there are totally $(q-2)/2$ b's such that $g(x) = b$ has two solutions in $\mathrm{GF}^*(q)\setminus\{1\}$, and such b's can be written as the form of $\frac{1+\alpha^i}{f(\alpha^i)}$, where $1 \le i \le q-2$. This means that there are $(q-2)/2$ b's such that

$$\mathrm{wt}(\mathbf{v}_3+\mathbf{v}_2+b\mathbf{v}_1) = q+1-3 = q-2.$$

In addition, the support of each codeword with minimum weight $q-2$ can be written as

$$\mathrm{supp}(\mathbf{v}_3+\mathbf{v}_2+b\mathbf{v}_1) = [q+1] \setminus \{i, j, q-1\},$$

where $1 \le i \ne j \le q-2$ vary with b. Since $\mathbf{v}_3$ is the all-one codeword, we have already characterized $(q-1)(q-2)/2$ minimum weight codewords of this form in $\mathcal{C}_f$.

We then consider all the codewords $\mathbf{v}_3+a\mathbf{v}_2+\mathbf{v}_1$, where $a \in \mathrm{GF}(q)^* \setminus \{1\}$. By definition,

$$\mathbf{v}_3+a\mathbf{v}_2+\mathbf{v}_1 = (a, 1+a\alpha^1+f(\alpha^1), \ldots, 1+a\alpha^{q-2}+f(\alpha^{q-2}), 1+a, 0).$$

Let $h(x) = \frac{1+f(x)}{x}$. Again, by Lemmas 11 and 10, $h(x)$ is also a 2-1 mapping from $\mathrm{GF}^*(q) \setminus \{1\}$ to itself. Therefore, there are totally $(q-2)/2$ a's such that $h(x) = a$ has two solutions in $\mathrm{GF}^*(q) \setminus \{1\}$, and such a's can be written as the form of $\frac{1+f(\alpha^i)}{\alpha^i}$, where $1 \le i \le q-2$. This implies that there are $(q-2)/2$ a's such that

$$\mathrm{wt}(\mathbf{v}_3+a\mathbf{v}_2+\mathbf{v}_1) = q+1-3 = q-2,$$

and the support of each codeword with minimum weight $q-2$ can be written as

$$\mathrm{supp}(\mathbf{v}_3+a\mathbf{v}_2+\mathbf{v}_1) = [q+1] \setminus \{i, j, q\},$$

where $1 \le i \ne j \le q-2$ vary with a. Since $\mathbf{v}_3$ is the all-one codeword, we have already characterized $(q-1)(q-2)/2$ minimum weight codewords of this form in $\mathcal{C}_f$.

In the first case above, the coordinate in position $q-1$ in the codewords $\mathbf{v}_3+\mathbf{v}_2+b\mathbf{v}_1$ is zero and the coordinate in position q in these codewords is nonzero. In the second case above, the coordinate in position $q-1$ in the codewords $\mathbf{v}_3+a\mathbf{v}_2+\mathbf{v}_1$ is nonzero and the coordinate in position q in these codewords is zero. Therefore, the minimum weight codewords in the two forms do not overlap. By Theorem 25, we have characterized all the minimum weight codewords in $\mathcal{C}_f$. From the discussions above, we have

$$\bigcup_{S \in \mathcal{B}_{q-2}(\mathcal{C}_f)} S = [q+1].$$

It then follows from Corollary 1 that $\mathcal{C}_f^\perp$ has minimum locality $d(\mathcal{C}_f)-1$.

The discussions above showed that the coordinate in position 0 in all the minimum weight codewords in $\mathcal{C}_f$ is nonzero. It then follows from Lemma 9 that the coordinate in position 0 in all the minimum weight codewords in $\mathcal{C}_f^\perp$ is zero. This means that

$$0 \notin \bigcup_{S \in \mathcal{B}_3(\mathcal{C}_f^\perp)} S.$$

Hence, $\mathcal{C}_f$ does not have minimum locality $d(\mathcal{C}_f^{\perp}) - 1$. It then follows from Theorem 14 that $\mathcal{C}_f$ has minimum locality $d(\mathcal{C}_f^{\perp})$. The remaining desired conclusions then follow from Theorems 25 and 15. □

The proof of Theorem 26 shows that it could be hard to determine the minimum locality of an NMDS code.

5.4.3 NMDS codes with parameters $[q+2, 3, q-1]$ from oval polynomials

Let f be a polynomial over GF(q) with $f(0) = 0$ and $f(1) = 1$. Let α be a generator of GF(q)*. Define

$$\bar{G}_f = \begin{bmatrix} f(0) & f(\alpha^0) & f(\alpha^1) & \dots & f(\alpha^{q-2}) & 0 & 1 \\ 0 & \alpha^0 & \alpha^1 & \dots & \alpha^{q-2} & 1 & 0 \\ 1 & 1 & 1 & \dots & 1 & 1 & 1 \end{bmatrix}. \tag{22}$$

By definition, $\bar{G}_f$ is a 3 by $q+2$ matrix over GF(q). Let $\bar{\mathcal{C}}_f$ denote the linear code over GF(q) with generator matrix $\bar{G}_f$.

The following theorem was presented in [39].

Theorem 27 [39] *Let $m \geq 3$ be odd and let $f(x)$ be an oval polynomial over* GF(q) *with coefficients in* GF(2). *Then $\bar{\mathcal{C}}_f$ is a $[q+2, 3, q-1]$ NMDS code over* GF(q) *with weight enumerator*

$$\bar{A}(z) = 1 + (q-1)(q-2)z^{q-1} + \frac{(q-1)(q^2-3q+14)}{2} z^q + 3(q-1)(q-2)z^{q+1} + \frac{(q-1)(q^2-3q+4)}{2} z^{q+2}.$$

Theorem 28 *Let $m \geq 3$ be odd and let $f(x)$ be an oval polynomial over* GF(q) *with coefficients in* GF(2), *and let $\bar{\mathcal{C}}_f$ be the code in Theorem 27. Then $\bar{\mathcal{C}}_f^{\perp}$ has minimum locality $d(\mathcal{C}_f) - 1$ and is a d-optimal and k-optimal $(q+2, q-1, 3, q; q-2)$-LLRC, and $\bar{\mathcal{C}}_f$ has minimum locality $d(\bar{\mathcal{C}}_f^{\perp})$ and is an almost d-optimal and k-optimal $(q+2, 3, q-1, q; 3)$-LLRC.*

Proof The proof is similar to that of Theorem 26 and is omitted here. However, we inform the reader that the coordinates in positions 0 and 1 in all the minimum weight codewords in $\bar{\mathcal{C}}_f$ are always nonzero. This means that

$$\{0, 1\} \bigcap \bigcup_{S \in \mathcal{B}_3(\bar{\mathcal{C}}_f^{\perp})} S = \emptyset.$$

To prove this theorem, one has to characterize all the minimum weight codewords in $\bar{\mathcal{C}}_f$. □

Notice that Theorem 26 documents the second class of nontrivial linear codes $\mathcal{C}$ with minimum locality more than $d(\mathcal{C}^{\perp}) - 1$. Many other infinite families of NMDS codes have been reported in the literature (see, for example, [1, 2, 10, 15–18, 25, 30, 37–39]). It is valuable to check which of the NMDS codes $\mathcal{C}$ have minimum locality $d(\mathcal{C}^{\perp}) - 1$, as they are d-optimal and k-optimal LLRCs. It has been observed that the analysis of the minimum locality of a linear code is harder than the determination of the minimum distance of the dual code.

Table 1 Some k-optimal LLRCs from known codes

$\mathcal{C}$	n	k	d	r	d_{opt}	k_{opt}
$\mathcal{H}_{(q,m)}$	n_h	$n_h - m$	3	$q^{m-1} - 1$	?	$\surd$
$\mathcal{S}_{(q,m)}$	n_h	m	q^{m-1}	2	?	$\surd$
$(\mathcal{H}_{(q,m)})_{\{t_1\}}$	$n_h - 1$	$n_h - m - 1$	3	$q^{m-1} - 2$	?	$\surd$
$((\mathcal{H}_{(q,m)})_{\{t_1\}})^{\perp}$	$n_h - 1$	m	$q^{m-1} - 1$	2	?	$\surd$
$(\mathcal{S}_{(q,m)})_{\{t_1\}}$	$n_h - 1$	$m - 1$	q^{m-1}	1	?	$\surd$
$((\mathcal{S}_{(q,m)})_{\{t_1\}})^{\perp}$	$n_h - 1$	$n_h - m$	2	$q^{m-1} - 1$	?	$\surd$
$\mathcal{R}_q(1, m)$	q^m	$m + 1$	$(q-1)q^{m-1}$	2	?	$\surd$
$\mathcal{R}_q(1, m)^{\perp}$	q^m	$q^m - m - 1$	3	$q^m - q^{m-1} - 1$	?	$\surd$
$\mathcal{C}_f$ (Thm. 26)	$2^m + 1$	3	$2^m - 2$	3	A	$\surd$
$\bar{\mathcal{C}}_f$ (Thm. 28)	$2^m + 2$	3	$2^m - 1$	3	A	$\surd$
$\mathcal{C}_o$	$q^2 + 1$	4	$q^2 - q$	3	?	$\surd$
$(\mathcal{C}_o)_{\{t_1\}}$	q^2	3	$q^2 - q$	2	?	$\surd$
$(\mathcal{C}_o)^{\{t_1\}}$	q^2	4	$q^2 - q - 1$	3	?	$\surd$
$\mathcal{C}(\mathcal{A})$ (Thm. 13)	n_a	3	$n_a - 2^i$	2	?	$\surd$

6 Summary and concluding remarks

The objectives of this paper are to develop some general theory for the minimum locality of linear codes and search for d-optimal or k-optimal LLRCs in known families of linear codes. Below is a summary of the major contributions of this paper.

1. We determined the minimum locality of $\overline{\mathcal{C}}^{\perp}$ under the condition that $d(\overline{\mathcal{C}}) = d(\mathcal{C}) + 1$ for each nontrivial linear code $\mathcal{C}$ (see Theorem 5), and settled the minimum locality of $\overline{\mathcal{C}}^{\perp}$ for each nontrivial binary linear code $\mathcal{C}$ (see Corollary 5).
2. We proved that the minimum locality of an NMDS code $\mathcal{C}$ is either $d(\mathcal{C}^{\perp}) - 1$ or $d(\mathcal{C}^{\perp})$ (see Theorem 14), and further proved that $\mathcal{C}$ is either a d-optimal and k-optimal or an almost d-optimal and k-optimal LLRC (see Theorem 15).

These general results have settled the minimum linear locality (also the minimum locality) of many families of nontrivial linear codes. Hence, we have reached our first objective.

After studying a number of families of known linear codes with the general theory developed, we have identified many classes of optimal LLRCs. These optimal LLRCs were not reported in the literature. Table 1 lists fourteen classes of k-optimal LLRCs. Table 2 lists nineteen classes of LLRCs which are both d-optimal and k-optimal and have different parameters. In both tables,

- $n_h = (q^m - 1)/(q - 1)$, $n_a = 2^{m+i} + 2^i - 2^m$, where $2 \le i \le m$,
- $\surd$ means that the code is optimal with the Singleton-like or CM bound,
- A means that the code is almost optimal with respect to the Singleton-like bound, and
- ? means that the optimality is open.

These classes of optimal LLRCs demonstrate that we have reached our second objective.

We remark that the locality of locally recoverable codes in the literature is actually the linear locality, but may not be the minimum linear locality. This paper has treated the minimum linear locality (also the minimum locality) of nontrivial linear codes. Availability is

Table 2 Both d-optimal and k-optimal LLRCs from known codes

$\mathcal{C}$	n	k	d	r	d_{opt}	k_{opt}
$\mathcal{H}_{(q,3)}$	q^2+q+1	q^2+q-2	3	q^2-1	$\surd$	$\surd$
$(\mathcal{H}_{(q,3)})_{\{t_1\}}$	q^2+q	q^2+q-3	3	q^2-2	$\surd$	$\surd$
$((\mathcal{S}_{(q,3)})_{\{t_1\}})^{\perp}$	q^2+q	q^2+q-2	2	q^2-1	$\surd$	$\surd$
$\mathcal{C}_o^{\perp}$	q^2+1	q^2-3	4	q^2-q-1	$\surd$	$\surd$
$(\mathcal{C}_o^{\perp})_{\{t_1\}}$	q^2	q^2-4	4	q^2-q-2	$\surd$	$\surd$
$(\mathcal{C}_o^{\perp})^{\{t_1\}}$	q^2	q^2-3	3	q^2-q-1	$\surd$	$\surd$
$\mathcal{C}(\mathcal{A})^{\perp}$ (Thm. 13)	n	$n-3$	3	$n-h-1$	$\surd$	$\surd$
$\mathcal{C}_{(3^s,3^s+1,3,1)}$	3^s+1	3^s-3	4	3^s-4	$\surd$	$\surd$
$\mathcal{C}^{\perp}_{(3^s,3^s+1,3,1)}$	3^s+1	4	3^s-3	3	$\surd$	$\surd$
$\mathcal{C}_{(2^s,2^s+1,3,1)}$	2^s+1	2^s-3	4	2^s-4	$\surd$	$\surd$
$\mathcal{C}^{\perp}_{(2^s,2^s+1,3,1)}$	2^s+1	4	2^s-3	3	$\surd$	$\surd$
$\mathcal{C}_{(2^s,2^s+1,4,1)}$	2^s+1	2^s-5	6	2^s-6	$\surd$	$\surd$
$\mathcal{C}^{\perp}_{(2^s,2^s+1,4,1)}$	2^s+1	6	2^s-5	5	$\surd$	$\surd$
$\mathcal{C}_f^{\perp}$ (Thm. 26)	2^m+1	2^m-2	3	2^m-3	$\surd$	$\surd$
$\bar{\mathcal{C}}_f^{\perp}$ (Thm. 28)	2^m+2	2^m-1	3	2^m-2	$\surd$	$\surd$
$\bar{\mathcal{C}}_f^{\perp}$ (Thm. 23)	2^m+3	2^m	3	2^m-1	$\surd$	$\surd$
$\bar{\mathcal{C}}_f$ (Thm. 24)	2^m+3	3	2^m	2	$\surd$	$\surd$
$(\overline{\mathcal{C}_{(2^s,2^s+1,3,1)}})^{\perp}$ (Thm 20)	2^s+2	5	2^s-3	4	$\surd$	$\surd$
$(\overline{\mathcal{C}_{(3^s,3^s+1,3,1)}})^{\perp}$ (Thm 19)	3^s+2	5	3^s-3	4	$\surd$	$\surd$

another interesting parameter of LLRCs. All the LLRCs presented in this paper naturally have availability 1, and some of them may have availability 2 or more. It is extremely hard to develop general theory for LLRCs with availability more than 1, although such LLRCs are available in the literature. To study the maximum availability of an LLRC code $\mathcal{C}$ with respect to the minimum locality $d(\mathcal{C}^{\perp})-1$, one has to characterize the minimum weight codewords in $\mathcal{C}^{\perp}$. This is a very hard problem in general. The reader is cordially invited to investigate the maximum availability of these optimal LLRCs documented in this paper. Finally, we point out that all the LLRCs presented in this paper are from known linear codes in the literature and our objective is to study their minimum locality and optimality with respect to the Singleton-like and CM bounds.

Acknowledgements The authors are very grateful to the reviewers and the Editor for their very detailed comments and suggestions that much improved the presentation and quality of this paper. The research of C. Fan and Z. Zhou was as supported by The National Natural Science Foundation of China (Grant Nos. 11971395, 62071397, and No. 62131016), and also by the Central Government Funds for Guiding Local Scientific and Technological Development under Grant 2021ZYD0001. The research of C. Ding was supported by the Research Grants Council of Hong Kong, under Grant No. 16301020. The research of C. Tang was supported by The National Natural Science Foundation of China (Grant No. 11871058) and China West Normal University (14E013, CXTD2014-4 and the Meritocracy Research Funds).

References

1. Abatangelo V., Larato B.: Near-MDS codes arising from algebraic curves. Discret. Math. **301**, 5–19 (2005).
2. Abatangelo V., Larato B.: Elliptic near-MDS codes over $\mathbf{F}_5$. Des. Codes Cryptgr. **46**, 167–174 (2008).
3. Assmus Jr., Mattson Jr.: New 5-designs. J. Comb. Theory, **6**(2), 122–151 (1969)
4. Assmus E.F. Jr., Key J.D.: Polynomial codes and finite geometries. In: Pless V.S., Huffman W.C. (eds.) Handbook of Coding Theory, pp. 1269–1343. Elsevier, Amsterdam (1998).
5. Cadambe, V.R., Mazumdar A.: An upper bound on the size of locally recoverable codes. International Symposium on Network Coding, Calgary, AB, Canada, pp. 1–5 (2013)
6. Cadambe V.R., Mazumdar A.: Bounds on the size of locally recoverable codes. IEEE Trans. Inf. Theory **61**(11), 5787–5794 (2015).
7. Cai, H., Fan, C., Miao, Y., Schwartz, M., Tang, X.: Optimal locally repairable codes: an improved bound and constructions. arXiv:2011.04966v1 [cs.IT]
8. Cai H., Cheng M., Fan C., Tang X.: Optimal locally repairable systematic codes based on packings. IEEE Trans. Commun. **67**(1), 39–49 (2019).
9. Chen B., Chen J.: A construction of optimal (r, δ)-locally recoverable codes. IEEE Access **7**, 180349–180353 (2019).
10. De Boer M.A.: Almost MDS codes. Des. Codes Cryptogr. **9**, 143–155 (1996).
11. Ding C.: Codes from Difference Sets. World Scientific, Singapore (2015).
12. Ding C., Heng Z.: The subfield codes of ovoid codes. IEEE Trans. Inf. Theory **65**(8), 4715–4729 (2019).
13. Ding C., Tang C.: Designs from Linear Codes, 2nd edn World Scientific, Singapore (2018).
14. Ding C., Tang C.: Infinite families of near MDS codes holding t-designs. IEEE Trans. Inf. Theory **66**(9), 5419–5428 (2020).
15. Dodunekov S., Landgev I.: On near-MDS codes. J. Geom. **54**(1–2), 30–43 (1995).
16. Dodunekov S.M., Landjev I.N.: Near-MDS codes over some small fields. Discret. Math. **213**(1–3), 55–65 (2000).
17. Faldum A., Willems W.: Codes of small defect. Des. Codes Cryptogr. **10**, 341–350 (1997).
18. Giulietti M.: On the extendibility of near-MDS elliptic codes. AAECC **15**, 1–11 (2004).
19. Gopalan P., Huang C., Simitci H., Yekhanin S.: On the locality of codeword symbols. IEEE Trans. Inf. Theory **58**(11), 6925–6934 (2012).
20. Grassl M.: Bounds on the minimum distance of linear codes and quantum codes. http://www.codetables.de, 2007. Accessed 07 Jan 2021
21. Grezet M., Hollanti C.: The complete hierarchical locality of the punctured Simplex code. Des. Codes Cryptogr. **89**, 1–21 (2021).
22. Guo, A., Kopparty, S., Sudan, M.M.: New affine-invariant codes from lifting. In: Proceedings of the 4th Conference on Innovations in Theoretical Computer Science, pp. 529–540 (2013)
23. Huang P., Yaakobi E., Uchikawa H., Siegel P.H.: Binary linear locally repairable codes. IEEE Trans. Inf. Theory **62**(11), 6268–6283 (2016).
24. Huffman W.C., Pless V.: Fundamentals of Error-Correcting Codes. Cambridge University Press, Cambridge (2003).
25. Jin L., Kan H.: Self-dual near MDS codes from elliptic curves. IEEE Trans. Inf. Theory **65**(4), 2166–2170 (2019).
26. Jin L., Kan H., Zhang Y.: Constructions of locally repairable codes with multiple recovering sets via rational function fields. IEEE Trans. Inf. Theory **66**(1), 202–209 (2020).
27. Liu J., Mesnager S., Tang D.: Constructions of optimal locally recoverable codes via Dickson polynomials. Des. Codes Cryptogr. **88**(2), 1759–1780 (2020).
28. Liu Y., Ding C., Tang C.: Shortened linear codes over finite fields. IEEE Trans. Inf. Theory **67**(8), 5119–5132 (2021).
29. Luo Y., Xing C., Yuan C.: Optimal locally repairable codes of distance 3 and 4 via cyclic codes. IEEE Trans. Inf. Theory **65**(2), 1048–1053 (2019).
30. Marcugini S., Milani A., Pambianco F.: NMDS codes of maximum length over $\mathbf{F}_q$, $8 \leq q \leq 11$,. IEEE Trans. Inf. Theory **48**(4), 963–966 (2002).
31. Maschietti A.: Difference set and hyperovals. Des. Codes Cryptogr. **14**, 89–98 (1998).
32. Micheli G.: Constructions of locally recoverable codes which are optimal. IEEE Trans. Inf. Theory **66**(1), 167–175 (2020).
33. Tamo I., Barg A.: A family of optimal locally recoverable codes. IEEE Trans. Inf. Theory **60**(8), 4661–4676 (2014).
34. Tan P., Zhou Z., Sidorenko V., Parampalli U.: Two classes of optimal LRCs with information (r, t)-locality. Des. Codes Cryptogr. **88**(2), 1741–1757 (2020).

35. Tang C., Ding C.: An infinite family of linear codes supporting 4-designs. IEEE Trans. Inf. Theory **67**(1), 244–254 (2021).
36. Tang C., Ding C., Xiong M.: Codes, differentially δ-uniform functions and t-designs. IEEE Trans. Inf. Theory **66**(6), 3691–3703 (2020).
37. Tong H., Ding Y.: Quasi-cyclic NMDS codes. Finite Fields Appl. **24**, 45–54 (2013).
38. Tsfasmann M.A., Vladut S.G.: Algebraic-Geometry Codes. Kluwer, Dordrecht (1991).
39. Wang Q., Heng Z.: Near MDS codes from overall polynomials. Discret. Math. **344**, 4 (2021).
40. Wang A., Zhang Z., Liu M.: Achieving arbitrary locality and availability in binary codes. Proc. ISIT **2016**, 1866–1870 (2015).
41. Xing, C., Yuan, C.: Construction of optimal locally recoverable codes and connection with hypergraph. Proceedings of the 46th International Colloquium on Automata, Languages, and Programming (ICALP 2019), pp. 1–98 (2019)