Optimal Linear and Cyclic Locally Repairable Codes over Small Fields

Alexander Zeh, Eitan Yaakobi

Optimal Linear and Cyclic Locally Repairable Codes over Small Fields

Alexander Zeh and Eitan Yaakobi
Computer Science Department
Technion—Israel Institute of Technology
alex@codingtheory.eu, yaakobi@cs.technion.ac.il

***Abstract*—We consider locally repairable codes over small fields and propose constructions of optimal cyclic and linear codes in terms of the dimension for a given distance and length.**

Four new constructions of optimal linear codes over small fields with locality properties are developed. The first two approaches give binary cyclic codes with locality two. While the first construction has availability one, the second binary code is characterized by multiple available repair sets based on a binary Simplex code.

The third approach extends the first one to q-ary cyclic codes including (binary) extension fields, where the locality property is determined by the properties of a shortened first-order Reed–Muller code. Non-cyclic optimal binary linear codes with locality greater than two are obtained by the fourth construction.

***Index Terms*—Availability, distributed storage, locally repairable codes, Reed–Muller code, Simplex code, sphere-packing bound**

I. Introduction

Locally repairable codes (LRC) can recover from erasure(s) by accessing a small number of erasure-free code symbols and therefore increase the efficiency of the repair-process in large-scale distributed storage systems. Basic properties and bounds of LRCs were identified by Gopalan *et al.* [1], Oggier and Datta [2] and Papailiopoulos and Dimakis [3]. The majority of the constructions of LRC requires a large field size (see e.g. [4]–[6]). The work of Kuijper and Napp [7] considers binary LRCs (and over binary extension field). Cadambe and Mazumdar [8] gave an upper bound on the dimension of a (nonlinear) code with locality which takes the field size into account. Goparaju and Calderbank [9] proposed binary cyclic LRCs with optimal dimension (among linear codes) for distances 6 and 10 and locality 2.

Our paper is based on the work of Goparaju and Calderbank [9] and we use their projection to an additive code without locality (see Calderbank *et al.* [10], Gaborit *et al.* [11], Kim *et al.* [12] for additive codes). We construct a new family of optimal binary codes (with distance 10 and locality 2) and generalize the approach to q-ary alphabets. Furthermore, we give a construction of optimal binary cyclic codes with availability greater than one based on Simplex codes (see Pamies-Juarez *et al.* [13], Rawat *et al.* [14] for the definition of availability and Kuijper and Napp [7] for a Simplex code based construction).

This work has been supported by German Research Council (Deutsche Forschungsgemeinschaft, DFG) under grant ZE1016/1-1.

This paper is structured as follows. Section II gives necessary preliminaries on linear and cyclic codes, defines LRCs, recalls the generalized Singleton bound, the Cadambe–Mazumdar bound [8] as well as the definition of availability for LRCs. The concept of a locality code and the projection to an additive code are discussed in Section III based on the work of Goparaju and Calderbank [9]. Two new constructions of optimal binary codes are given in Section IV and a construction based on a q-ary shortened cyclic first-order Reed–Muller code is given in Section V. The fourth construction in Section VI uses code concatenation and provides optimal linear binary codes. Section VII concludes this contribution.

II. Preliminaries

Let $[a,b)$ denote the set of integers $\{a, a+1, \ldots, b-1\}$ and $[b)$ be the shorthand notation for $[0,b)$. Let $\mathbb{F}_q$ denote the finite field of order q and $\mathbb{F}_q[X]$ the polynomial ring over $\mathbb{F}_q$ with indeterminate X. A linear $[n,k,d]_q$ code of length n, dimension k and minimum Hamming distance d over $\mathbb{F}_q$ is denoted by a calligraphic letter like $\mathcal{C}$ as well as a non-linear $(n,M,d)_q$ of length n, cardinality M and minimum distance d.

An $[n,k,d]_q$ q-ary cyclic code $\mathcal{C}$ with distance d is an ideal in the ring $\mathbb{F}_q[X]/(X^n-1)$ generated by $g(X)$. The generator polynomial $g(X)$ has roots in the splitting field $\mathbb{F}_{q^s}$, where $n \mid (q^s-1)$.

A q-cyclotomic coset $M_{i,n}$ is defined as

$$M_{i,n} \stackrel{\text{def}}{=} \left\{ iq^j \mod n \,|\, j \in [a) \right\}, \tag{1}$$

where a is the smallest positive integer such that $iq^a \equiv i \bmod n$. The minimal polynomial in $\mathbb{F}_q[X]$ of the element $\alpha^i \in \mathbb{F}_{q^m}$ is given by $m_i(X) = \prod_{j \in M_{i,n}} (X - \alpha^j)$. The defining set $D_{\mathcal{C}}$ of an $[n,k,d]_q$ cyclic code $\mathcal{C}$ is

$$D_{\mathcal{C}} = \left\{ 0 \le i \le n-1 \,|\, g(\alpha^i) = 0 \right\}. \tag{2}$$

For visibility we sometimes mark a position i with a $\square$ if $g(\alpha^i) \neq 0$. Furthermore, let $D_{\mathcal{C}}^{[z]}$ be the short-hand notation for $\{(i+z) \mid i \in D_{\mathcal{C}}\}$ for a given $z \in \mathbb{Z}$. Let us recall the definition of linear locally repairable codes.

Definition 1 (Locally Repairable Code (LRC))**.** *A linear $[n,k,d]_q$ code $\mathcal{C}$ is said to have (r,δ)-locality if for all n code symbols $c_i, \forall i \in [n)$, there exists a punctured subcode of*

$\mathcal{C}$ with support containing i, whose length is at most $r+\delta-1$, and whose minimum distance is at least δ.

A code $\mathcal{C}$ is called *r-local* if it has $(r,2)$-locality.

The following generalization of the Singleton bound for LRCs was among others proven in [15, Thm. 3.1], [6, Construction 8 and Thm. 5.4] and [16, Thm. 2].

Theorem 2 (Generalized Singleton Bound)**.** *The minimum distance d of an $[n,k,d]_q$ linear (r,δ)-locally repairable code $\mathcal{C}$ (as in Def. 1) is upper bounded by*

$$d \leq n-k+1-\left(\left\lceil \frac{k}{r} \right\rceil -1\right)(\delta-1). \tag{3}$$

For $\delta=2$ and $r=k$ it coincides with the classical Singleton bound. Throughout this contribution we call a code *Singleton-optimal* if its distance meets the bound in Thm. 2 with equality. The generalized Singleton bound as in Thm. 2 does not take the field size into account. We compare our constructions with the bound given by Cadambe and Mazumdar [8, Thm. 1] which depends on the alphabet size. In general it holds also for nonlinear codes, but we state it only for linear codes in the following.

Theorem 3 (Cadambe–Mazumdar (CM Bound))**.** *The dimension k of an r-local repairable code $\mathcal{C}$ of length n and minimum Hamming distance d is upper bounded by*

$$k \leq \min_{t\in\mathbb{Z}} \left\{ tr + k_{opt}^{(q)}\big(n-t(r+1),d\big) \right\}, \tag{4}$$

where $k_{opt}^{(q)}(n,d)$ is the largest possible dimension of a code of length n, for a given alphabet size q and a given minimum distance d.

In the following we use Thm. 3 to bound the dimension of linear codes. Another important parameter for LRCs is the availability e.g. considered in Kuijper and Napp [7] and Cadambe and Mazumdar [8] and therefore we define it in the following.

Definition 4 (Availability)**.** *An $[n,k,d]_q$ linear code $\mathcal{C}$ is called t-available-r-local locally repairable if every code symbol $c_i, \forall i \in [n)$, has at least t parity-checks of weight $r+1$ which intersect pairwise in (and only in) $\{i\}$.*

III. Locality Code and Additive Code

In this section we shortly recall the approach of Goparaju and Calderbank [9] and extend it to what we call a *locality code*.

Let us first describe the idea of Constructions 1 and 2 of [9] in terms of a locality code. Construction 1 of [9] gives a binary cyclic code $\mathcal{C}$ of length $n=2^m-1$ with locality r, where the code length n is divisible by $r+1$. The defining set is $D_{\mathcal{C}} = \{i \mod (r+1), \forall i \in [n)\}$. This equals the union of $n/(r+1)$ shifted defining sets $D_{\mathcal{L}}=\{0\}$ of the binary cyclic $[r+1,r,2]_2$ single-parity check code, which is able to correct one erasure within a block of length $r+1$ by "accessing" only r other code symbols. The code $\mathcal{C}$ inherits the properties of $\mathcal{L}$, also the minimum distance of two, which is Singleton-optimal, but does not increase the overall erasure-correction capability. In general, let $\mathcal{C}$ be the aimed $[n,k,d]_q$ code with locality properties that are inherited from an $[n_l,k_l,d_l]_q$ locality code $\mathcal{L}$. Namely the code constructions of [9] and our (cyclic) codes are subcodes of the $[n,k,d]_q$ cyclic product code $\mathcal{L}\otimes\mathcal{T}$, where $\mathcal{T}$ is the trivial $[n/n_l, n/n_l, 1]_q$ code (see [17]), i.e., a cyclic code with defining set:

$$D_{\mathcal{C}} = \left\{ D_{\mathcal{L}} \cup D_{\mathcal{L}}^{[n_l]} \cup \cdots \cup D_{\mathcal{L}}^{[n-n_l-1]} \cup R \right\}. \tag{5}$$

In Construction 2 ($R=M_{1,n}$) and 3 ($R=M_{1,n}\cup M_{-1,n}$) of [9], the locality code $\mathcal{L}$ is a $[3,2,2]_2$ single-parity check code with defining set $D_{\mathcal{L}}=\{0\}$. The optimality among binary codes with locality $r=2$ is shown via the projection to an additive code.

Lemma 5 (Projection to Additive Code)**.** *Let $\mathcal{L}$ be an $[n_l,k_l,d_l]_q$ locality code and let $\mathcal{C}$ be an $[n,k,d]_q$ code with defining set as in* (5)*. Then, we can project each sub-block of n_l symbols of a codeword in $\mathcal{C}$ to one symbol in $\mathbb{F}_{q^{k_l}}$. The obtained $(n',q^k,d')_{q^{k_l}}$ additive code $\mathcal{A}$ has parameters:*

$$n' = n/n_l \quad \textit{and} \quad d' \geq \lceil d/\omega \rceil, \tag{6}$$

where ω is the maximum weight of a codeword in $\mathcal{L}$.

Proof: The length n' and the alphabet-size follow directly from the projection of the coordinates. The cardinality of $\mathcal{A}$ equals the one of $\mathcal{C}$. The distance follows from the fact that in the worst-case ω non-zero symbols of $\mathcal{C}$ are projected to one symbol over $\mathbb{F}_{q^{k_l}}$ (see Fig. 1). ∎

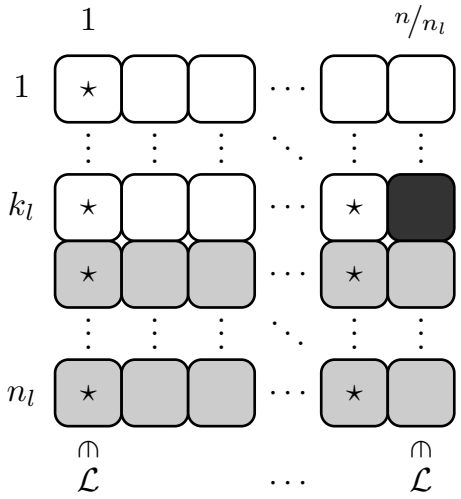

Fig. 1. Illustration of a nonzero minimum-weight codeword of weight seven of the $[n,k,d]_q$ code $\mathcal{C}$ arranged in n/n_l blocks of length n_l. The $\star$ marks a nonzero symbol in $\mathbb{F}_q$. The corresponding codeword of the additive code $\mathcal{A}$ over $\mathbb{F}_{q^{k_l}}$ has length n/n_l and has at least weight $\lceil d/\omega \rceil$. Here $\omega=4$ and therefore at least two symbols in $\mathcal{A}$ are nonzero (first and before last column). The redundancy added by the $[n_l,k_l,d_l]_q$ locality code $\mathcal{L}$ is illustrated as gray symbols, while the black symbol marks the additional redundancy to obtain a distance that is higher than the one given by the Singleton bound.

For a cyclic binary $[r+1,r,2]_2$ single-parity check code of odd length, the maximum weight of a codeword is $\omega=r$.

Lemma 6 (Locality Code)**.** *If $\mathcal{L}$ is an $[n_l = r+\delta-1, r, \delta]_q$ MDS locality code, then the cyclic code $\mathcal{C}$ with defining set as in* (5) *has (r,δ)-locality and its distance is $d \geq \delta$.*

Proof: The (r,δ)-locality of $\mathcal{C}$ follows directly from the construction. The distance of $\mathcal{C} \subseteq \mathcal{L} \otimes \mathcal{T}$ is at least the distance of the product code $\mathcal{L} \otimes \mathcal{T}$. ∎

Note that for $R = \emptyset$, the code $\mathcal{C} = \mathcal{L} \otimes \mathcal{T}$ is Singleton-optimal, i.e., the dimension of $\mathcal{C}$ is $k = nr/(r+\delta-1)$ and from (3) we obtain:

$$\begin{aligned} d &\leq n - \frac{nr}{r+\delta-1} + 1 - \left(\frac{n}{r+\delta-1} - 1\right)(\delta-1) \\ &\leq \frac{n(r+\delta-1) - nr - n(\delta-1)}{r+\delta-1} + \delta = \delta. \end{aligned}$$

IV. Binary Cyclic Codes With Locality Two

Construction 1 in [9] gives Singleton-optimal binary cyclic codes (these codes are of lowest-rate for $d = 2$ see [18, Prop. 2]). The following construction gives a new class of binary cyclic 2-local codes.

Construction 7 (Binary Reversible Codes)**.** *Let $n = 2^m + 1$ and $3|n$ and therefore m odd. Let the locality $r = 2$, i.e. let $\mathcal{L}$ be a $[3,2,2]_2$ single-parity check code with $D_{\mathcal{L}} = \{0\}$. Let the defining set be*

$$\begin{aligned} D_{\mathcal{C}} &= \big\{\{\dots,-6,-3,0,3,6,9,12,\dots\} \cup M_{1,n}\big\}, \\ &= \{\dots,-6,-4,-3,-2,-1,0,1,2,3,4,6,\dots\}. \end{aligned}$$

Then $\mathcal{C}$ has dimension $k = \frac{2}{3}(2^m+1) - 2m$ and distance $d \geq 10$.

Due to the length, the coset $M_{1,n}$ is reversible (see [19] for reversible codes), i.e., $M_{1,n} = \{1,2,\dots,2^{m-1},2^m = -1,-2,\dots,-2^{m-1}\}$ and has cardinality $2m$. The distance follows from the BCH bound [20], [21], where the consecutive sequence is $-4,-3,\dots,3,4$.

Theorem 8. *Let a binary linear code with parameters as in Construction 7 be given. Then its dimension satisfies:*

$$k \leq \frac{2}{3}(2^m+1) - 2m. \tag{7}$$

Proof: Equivalent to the proof of [9, Thm. 2], we have via sphere-packing bound (see [22, Ch. 1 §5]) for the $(2n/3, 2^k, 5)_{2^2}$ additive code that

$$2^k \leq \frac{4^{n'}}{1 + 3n' + \frac{9n'(n'-1)}{2}} \tag{8}$$

and therefore

$$\begin{aligned} k &\leq \log_2(4^{n/3}) - \log_2\left(1 - \frac{1}{2}n + \frac{1}{2}n^2\right) \\ &= \frac{2n}{3} + 1 - \lceil \log_2\left(2 - n + n^2\right)\rceil. \end{aligned} \tag{9}$$

With $n = 2^m + 1$, we obtain from (9)

$$\begin{aligned} k &= \frac{2n}{3} + 1 - \lceil \log_2\left(2 - (2^m+1) + (2^m+1)^2\right)\rceil \\ &= \frac{2n}{3} + 1 - \lceil \log_2\left(2 + 2^m + 2^{2m}\right)\rceil \\ &= \frac{2n}{3} + 1 - (2m+1) = \frac{2n}{3} - 2m. \end{aligned}$$

∎

Remark 1: A binary cyclic code as in Construction 7 without $M_{1,n}$ in the defining set is Singleton-optimal and the distance equals $d = 2$ (for $k = 2n/3$, $r = 2$ and $\delta = 2$), which is the smallest minimum distance possible for a binary cyclic code with rate $2/3$ (see [18, Prop. 2]).

Remark 2: The nth root of unity is in $\mathbb{F}_{2^{2m}-1}$ (which is twice the extension order of the code obtained via [9, Construction 3]) and therefore the (non-local) decoding complexity is higher than [9, Construction 3].

Example 9 (Optimal 2-Local Binary Code)**.** *Let $n = 2^5+1 = 33$ and via Construction 7 we obtain a binary cyclic code of minimum distance $d = 10$, with*

$$D_{\mathcal{C}} = \big\{\{0,3,6,9,\dots,30\} \cup \{1,2,4,\dots,32\}\big\}$$

and dimension $k = 12$, which is 2-local. The CM bound (see Thm. 3) based on the best-known linear codes give $k \leq 13$.

Let us consider Construction 4 of [9] based on the $[7,3,4]_2$ locality code $\mathcal{L}$ with locality $r = 2$, availability $t = 3$ and with defining set $D_{\mathcal{L}} = \big\{0,\square,\square,3,\square,5,6\big\}$. The code $\mathcal{C}$ with defining set as in (5), but with $R = \emptyset$ is an $[n = 2^m - 1, k = 3n/7, 4]_2$ cyclic code, where m is a multiple of three.

We extend Construction 4 of [9] to obtain a higher distance and small reduction of the rate as follows.

Construction 10 (Sphere-Packing Optimal Binary Code with Locality Two and Increased Availability)**.** *Let $n = 2^m - 1$ and be divisible by 7 and therefore $3|m$. Let the defining set be:*

$$\begin{aligned} D_{\mathcal{C}} = \big\{\{\dots,-9,-8,|-7,\square,\square,-4,\square,-2,-1,|0,\square,\square, \\ 3,\square,5,6,|7,\square,\square,10,\square,\square,12,13|,\dots\} \\ \cup M_{1,n} = \{1,2,4,\dots,2^{m-1}\}\big\}. \end{aligned}$$

Then $d \geq 12$ (via BCH bound [20], [21], where the consecutive sequence is $-2,1,\dots,8$). The constructed $[n = 2^m - 1, k = 3n/7 - m, d \geq 12]_2$ cyclic code $\mathcal{C}$ is a 2-local code and has availability $t = 3$ as defined in Def. 4.

Example 11 (2-Local Binary Code with Availability Three)**.** *Let $n = 9 \cdot 7 = 63$, $D_{\mathcal{L}} = \big\{0,\square,\square,3,\square,5,6\big\}$ and let the defining set be*

$$\begin{aligned} D_{\mathcal{C}} &= \{D_{\mathcal{L}} \cup D_{\mathcal{L}}^{[7]} \cup \dots \cup D_{\mathcal{L}}^{[56]} \cup M_{1,63}\} \\ &= \{\{..,59,61,62,|0,3,5,6,|7,10,..\} \cup \{1,2,4,..,32\}\}. \end{aligned}$$

The constructed code $\mathcal{C}$ is an $[63,21,12]_2$ code and the corresponding additive code according to Lemma 5 is a $(9,2^7,3)_{2^3}$ code. The BCH bound is tight and the consecutive sequence ranges is $61,62,0,1,\dots,8$.

Let us prove the optimality of Construction 10 in the following theorem.

Theorem 12. *Let $3|m$ and let $\mathcal{C}$ be an $[2^m-1,k,12]_2$ linear code, let $\mathcal{L}$ be the $[7,3,4]_2$ Simplex code and let $\mathcal{C}$ have the locality inherited from $\mathcal{L}$ as in Lemma 5. Then,*

$$k \leq \frac{3}{7}\left(2^m-1\right)-m. \tag{10}$$

Proof: From (6) we have an $(n'=n/7, 2^k, d'=\lceil 12/4\rceil = 3)_{2^3}$ additive code $\mathcal{A}$. The $[7,3,4]_2$ Simplex code is a constant-weight code and therefore $\omega = 4$. The code $\mathcal{A}$ is defined over $\mathbb{F}_{2^3}$ and has dimension

$$k' = k/3 = \frac{1}{3}\left(\frac{3}{7}n - m\right) = n' - m'.$$

and therefore the parameters of a $[((2^3)^{m'}-1)/(2^3-1), n'-m', 3]_{2^3}$ Hamming code, which is optimal w.r.t. the sphere-packing bound. ∎

Construction 10 can be extended to the case where the locality code $\mathcal{L}$ is the $[2^a-1, a, 2^{a-1}]_2$ cyclic Simplex code, which is 2-local and has availability $t = 2^{a-1}-1$ (see e.g. Kuijper–Napp [7, Lemma 3.1] and Wang–Zhang [23]).

Construction 13 (Binary Code with Simplex Locality)**.** *Let $n = 2^m-1$ and be divisible by 2^a-1 and therefore $a|m$. Let $\mathcal{L}$ be the $[2^a-1, a, 2^{a-1}]_2$ cyclic Simplex code with defining set*

$$D_{\mathcal{L}} = \left\{0, \square, \square, 3, \square, 5, 6, .., \square, 2^{a-1}+1, .., 2^a-1\right\}. \tag{11}$$

Let the defining set of the code $\mathcal{C}$ be:

$$\begin{aligned} D_{\mathcal{C}} &= \left\{ D_{\mathcal{L}} \cup D_{\mathcal{L}}^{[2^a-1]} \cup D_{\mathcal{L}}^{[2(2^a-1)]} \cup \cdots \cup M_{1,n} \right\} \\ &= \left\{ \ldots, -2^{a-1}+1, \ldots, -1 | , 0, 1, \ldots, 2^a-1, | 2^a, \ldots \right\}. \end{aligned}$$

Then $d \geq 2^a + 2^{a-1}$ (via BCH bound for the consecutive sequence from $-(2^{a-1}-1)$ to 2^a) and the dimension is $k = \frac{a}{2^a-1}(2^m-1) - m$.

We have the following theorem on the optimality of the dimension of linear codes.

Theorem 14 (Simplex Locality)**.** *Let $a|m$ and let $\mathcal{C}$ be an $[2^m-1, k, 2^{a-1}\cdot 3]_2$ linear code, let $\mathcal{L}$ be the $[2^a-1, a, 2^{a-1}]_2$ binary Simplex code and let $\mathcal{C}$ have the locality properties according to $\mathcal{L}$ as in Lemma 5. Then,*

$$k \leq \frac{a}{2^a-1}\left(2^m-1\right) - m. \tag{12}$$

Proof: From (6) we have an $(n/(2^a-1), 2^k, \lceil d/(2^{a-1})\rceil)_{2^a}$ additive code $\mathcal{A}$, where $\omega = 2^a-1$, because the simplex code is a constant-weight code. The additive code has the parameters of a Hamming code over $\mathbb{F}_{2^a}$ with dimension

$$k' = k/a = \frac{1}{a}\left(\frac{a}{2^a-1}n - m\right) = n' - m',$$

and distance

$$d' = \left\lceil \frac{d}{2^{a-1}} \right\rceil = \left\lceil \frac{2^{a-1}(1+2)}{2^{a-1}} \right\rceil = 3.$$

∎

V. Q-Ary Case: First-Order Shortened RM Code as Locality Code

We extend the previous approach for cyclic codes to the q-ary case and use as locality code $\mathcal{L}$ the q-ary

$$[q^2-1, 2, (q-1)q^{2-1}]_q = [q^2-1, 2, q^2-q]_q \tag{13}$$

cyclic shortened first-order Reed–Muller (RM, see [24, Problem 2.17] and [25, Section 6.11]) code. Its dual code is the $[q^2-1, q^2-3, 2]_q$ code with defining set $\{1, q\}$. A $[q^a-1, a, q^{a-1}(q-1)]_q$ shortened first-order RM code is the q-ary pendant of the Simplex code and also a constant-weight code with $\omega = q^{a-1}(q-1)$. RM codes have the highest minimum distance possible for the given parameters among q-ary linear codes. Furthermore, first-order RM codes and their locality properties were investigated by Rawat and Vishwanath in [26].

Construction 15 (Reed–Muller Code Locality)**.** *Assume $q > 2$. Let $\mathcal{C}$ be an $[q^m-1, k, d]_q$ code. Let $\mathcal{L}$ be an $[q^2-1, 2, q^2-q]_q$ cyclic RM code with defining set*

$$D_{\mathcal{L}} = \left\{0, \square, 2, 3, \ldots, q-1, \square, q+1, \ldots, q^2-2\right\}.$$

Let the defining set of $\mathcal{C}$ be:

$$\begin{aligned} D_{\mathcal{C}} &= \left\{ D_{\mathcal{L}} \cup D_{\mathcal{L}}^{[q-1]} \cdots \cup \{1, q, q^2, \ldots, q^{m-1}\} \right\} \\ &= \{\square, (q^2-q-2), .., 0, .., q^2+q-2, \square, ..\}. \end{aligned}$$

Then the dimension of $\mathcal{C}$ is $k = \frac{2n}{q^2-1} - m$ and the distance is $d \geq q^2-q-2+q^2+q-2+1+1 = 2q^2-2$ via BCH bound [20], [21] (from $-(q^2-q-2)..+(q^2+q-2)$).

Theorem 16 (Locality Code: Shortened First-Order RM Code)**.** *Let $\mathcal{C}$ be an $[q^m-1, k, (q^2-q)\cdot 3]_2$ linear code, let $\mathcal{L}$ be the $[q^2-1, 2, q^2-q]_q$ RM code and let $\mathcal{C}$ have the locality properties according to $\mathcal{L}$ as in Lemma 5. Then,*

$$k \leq \frac{2(q^m-1)}{q^2-1} - m.$$

Proof: The additive code $\mathcal{A}$ has parameters

$$\begin{aligned} n' &= \frac{q^m-1}{q^2-1}, \\ k' &= k/2, \\ d' &= \left\lceil \frac{d}{q^2-q} \right\rceil, \end{aligned}$$

and has alphabet-size $\mathbb{F}_{q^2}$. More explicitly, the dimension is:

$$k' = k/2 = \frac{1}{2}\left(\frac{2(q^m-1)}{q^2-1} - m\right) = n' - m',$$

and the distance is

$$d' = \left\lceil \frac{d}{q^2-q} \right\rceil \geq \left\lceil \frac{2q^2-2}{q-1} \right\rceil = \left\lceil \frac{q^2(2-2/q^2)}{q^2(1-\frac{1}{q})} \right\rceil = 3.$$

Therefore the additive code has the parameters of an q^2-ary Hamming code, which is optimal w.r.t. to the sphere-packing bound. ∎

Example 17 (Optimal Ternary Code)**.** *Let $q = 3$ and let $n = 3^4 - 1 = 80$. Let $\mathcal{L}$ be the $[8, 2, 6]_3$ shortened first-order cyclic RM code with defining set is $D_{\mathcal{L}} = \{0, \square, 2, \square, 4, 5, 6, 7\}$. Then, the defining set of $\mathcal{C}$ according to Construction 15 is $D_{\mathcal{C}} = \{\{.., -4, .., 0, \square, 2, \square, 4, .., 8, \square, 10, \square, ..\} \cup \{1, 3, 9, 27\}\}$ and therefore the BCH bound gives $d \geq 16$ ($-4..10$). And thus $d' = \lceil 16/6 \rceil = 3$.*

Construction 15 is also valid for extension fields and therefore let us give another example over a binary extension field.

Example 18 (Cyclic Optimal Code over Binary Extension Field)**.** *Let $q = 2^2$ and let $n = 4^4 - 1 = 255$. Let $\mathcal{L}$ be the $[15, 2, 12]_4$ RM code with defining set $D_{\mathcal{L}} = \{0, \square, 2, 3, \square, 5, \ldots, 14\}$. Then, the defining set of $\mathcal{C}$ according to Construction 15 is $D_{\mathcal{C}} = \{\{.., -10, .., 0, \square, 2, 3, \square, 5, .., 15, \square, 17, 18, \square, ..\} \cup \{1, 4, 16, 64\}\}$. The BCH bound gives $d \geq 30$ (from $-10..18$). The additive code over $\mathbb{F}_{2^2}$ has length $n' = 255/15 = 17$, dimension $k' = 17 - 4 = 13$ and distance $d' = \lceil 30/12 \rceil = 3$.*

The real distance of the codes via Construction 15 is $3(q^2 - q)$, but the BCH bound is not tight. Other bounds for cyclic codes can deliver a better result and a more advanced algebraic decoders for the non-local erasure-decoding can be applied.

VI. Optimal Linear Codes

Based on concatenated codes [27], [28], we propose a construction, where the row-code is a linear Hamming code over the binary extension field $\mathbb{F}_{2^r}$.

Construction 19 (Binary Linear r-Local Code)**.** *Let the row code be a $[(2^{2r}-1)/(2^r-1) = 2^r+1, 2^r+1-2 = 2^r-1, 3]_{2^r}$ binary Hamming code and let the column-code $\mathcal{L}$ be a linear (not necessarily cyclic) $[r+1, r, 2]_2$ single-parity check code. Then the concatenated coded code is a $[(2^r+1)(r+1), (2^r-1)r, 6]_2$ r-local code.*

Construction 19 gives an r-local linear code, with highest possible dimension k. (The additive codes is a binary Hamming code.)

Example 20 (Binary Code with Locality $r = 3$)**.** *Let the $[((2^3)^2 - 1)/(2^3 - 1) = 9, 9 - 2 = 7, 3]_{2^3}$ Hamming code be the row code of a concatenated code and let the column code be the $[4, 3, 2]_2$ single-parity check code that corresponds to locality $r = 3$. Then $\mathcal{C}$ is a $[36, 21, 6]_2$ 3-local linear code. The CM bound (Thm. 3) gives $k \leq 21$.*

VII. Conclusion and Outlook

We proposed new constructions of optimal binary and q-ary cyclic (and linear) codes with locality and availability.

The following future work seems fruitful. The extension of Construction 13 to codes with higher distance and optimal dimension, the usage of other first-order cyclic Reed–Muller codes as locality code similar to Construction 15, the usage of improved bounds on the minimum distance for the cyclic codes obtained via Construction 15 and the extension of Construction 19 to q-ary linear codes.

References

[1] P. Gopalan, C. Huang, H. Simitci, and S. Yekhanin, "On the locality of codeword symbols," *IEEE Trans. Inform. Theory*, vol. 58, no. 11, pp. 6925–6934, 2012.

[2] F. Oggier and A. Datta, "Self-repairing homomorphic codes for distributed storage systems," in *IEEE INFOCOM*, 2011, pp. 1215–1223.

[3] D. S. Papailiopoulos and A. G. Dimakis, "Locally repairable codes," *IEEE Trans. Inform. Theory*, vol. 60, no. 10, pp. 5843–5855, 2014.

[4] I. Tamo, D. S. Papailiopoulos, and A. G. Dimakis, "Optimal locally repairable codes and connections to matroid theory," *arXiv:1301.7693 [cs, math]*, 2013. [Online]. Available: http://arxiv.org/abs/1301.7693

[5] N. Silberstein, A. Rawat, O. Koyluoglu, and S. Vishwanath, "Optimal locally repairable codes via rank-metric codes," in *IEEE Intern. Symp. on Inf. Theory (ISIT)*, 2013, pp. 1819–1823.

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

[7] M. Kuijper and D. Napp, "Erasure codes with simplex locality," *arXiv:1403.2779 [cs, math]*, 2014. [Online]. Available: http://arxiv.org/abs/1403.2779

[8] V. Cadambe and A. Mazumdar, "An upper bound on the size of locally recoverable codes," in *IEEE Intern. Symp. on Network Coding (NetCod)*, 2013, pp. 1–5.

[9] S. Goparaju and R. Calderbank, "Binary cyclic codes that are locally repairable," in *IEEE Intern. Symp. on Inf. Theory (ISIT)*, 2014, pp. 676–680.

[10] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, "Quantum error correction via codes over GF(4)," *IEEE Trans. Inform. Theory*, vol. 44, no. 4, pp. 1369–1387, 1998.

[11] P. Gaborit, W. C. Huffman, J.-l. Kim, and V. Pless, "On additive GF(4) codes," in *DIMACS Series in Discrete Math. and Theoret. Computer Science*. American Mathematical Society, 2001, pp. 135–149.

[12] J.-L. Kim, K. E. Mellinger, and V. Pless, "Projections of binary linear codes onto larger fields," *SIAM J. Discrete Math.*, vol. 16, no. 4, pp. 591–603, 2003.

[13] L. Pamies-Juarez, H. D. L. Hollmann, and F. Oggier, "Locally repairable codes with multiple repair alternatives," in *IEEE Intern. Symp. on Inf. Theory (ISIT)*, 2013, pp. 892–896.

[14] A. S. Rawat, D. S. Papailiopoulos, A. G. Dimakis, and S. Vishwanath, "Locality and availability in distributed storage," in *IEEE Intern. Symp. on Inf. Theory (ISIT)*, 2014, pp. 681–685.

[15] G. Kamath, N. Prakash, V. Lalitha, and P. Kumar, "Codes with local regeneration and erasure correction," *IEEE Trans. Inform. Theory*, vol. 60, no. 8, pp. 4637–4660, 2014.

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

[17] H. Burton and E. J. Weldon, "Cyclic product codes," *IEEE Trans. Inform. Theory*, vol. 11, no. 3, pp. 433–439, 1965.

[18] A. Zeh and S. V. Bezzateev, "A new bound on the minimum distance of cyclic codes using small-minimum-distance cyclic codes," *Des. Codes Cryptogr.*, vol. 71, no. 2, pp. 229–246, 2014.

[19] J. L. Massey, "Reversible codes," *Inf. Control*, vol. 7, no. 3, pp. 369–380, 1964.

[20] A. Hocquenghem, "Codes correcteurs d'erreurs," *Chiffres (Paris)*, vol. 2, pp. 147–156, 1959.

[21] R. C. Bose and D. K. Ray-Chaudhuri, "On a class of error correcting binary group codes," *Inf. Control*, vol. 3, no. 1, pp. 68–79, 1960.

[22] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. North Holland Publishing Co., 1988.

[23] A. Wang and Z. Zhang, "Repair locality from a combinatorial perspective," *arXiv:1401.2607 [cs, math]*, 2014. [Online]. Available: http://arxiv.org/abs/1401.2607

[24] R. M. Roth, *Introduction to Coding Theory*. Cambridge University Press, 2006.

[25] J. van Lint, *Introduction to Coding Theory*. Springer, 1999.

[26] A. S. Rawat and S. Vishwanath, "On locality in distributed storage systems," *arXiv:1204.6098 [cs, math]*, 2012. [Online]. Available: http://arxiv.org/abs/1204.6098

[27] G. D. Forney, "Concatenated codes," 1966.

[28] E. L. Blokh and V. V. Zyablov, "Coding of generalized concatenated codes," *Probl. Inf. Transm.*, vol. 10, no. 3, pp. 45–50, 1974.