Codes With Minimum Bandwidth Cooperative Local Regeneration

by

Nikhil Krishnan M, Lalitha Vadlamani, Sreeranjani Didugu

Report No: IIIT/TR/2020/-1

Centre for Communications
International Institute of Information Technology
Hyderabad - 500 032, INDIA
January 2020

Codes With Minimum Bandwidth Cooperative Local Regeneration

M. Nikhil Krishnan, *Member, IEEE*, V. Lalitha, *Member, IEEE*, and Sreeranjani Didugu

***Abstract*— Locally recoverable codes (LRCs) are known to reduce the repair locality, i.e., the number of nodes accessed during node repair, in distributed storage systems. Regenerating codes, on the other hand, offer a decreased repair bandwidth, which is the repair traffic incurred during node repairs. Locally regenerating codes (LRGCs), which are constructed via intertwining regenerating codes of smaller block lengths, simultaneously offer a small repair locality and repair bandwidth. In this letter, we construct a family of LRGCs where the constituent codes are bandwidth-efficient, minimum bandwidth cooperative regenerating codes. The newly constructed LRGCs are optimal with respect to both minimum distance and rate.**

***Index Terms*— Locally regenerating codes (LRGCs), cooperative regenerating codes, locally minimum bandwidth cooperative regeneration (MBCR) codes, uniform rank accumulation (URA) codes.**

I. Introduction

The explosive growth in the amount of data generated necessitates storing data in a distributed manner, across multiple storage units. A commonplace occurrence in such distributed storage systems (DSSs), is the failure of an individual storage unit (node). In order to combat node failures, a naive approach is to replicate data across multiple nodes. However, erasure coding techniques, in comparison to the data replication approach, offer lower storage overheads, without compromising on the reliability. Typically, in order to apply erasure codes to reliably store data in DSSs, a given data file will be broken into k fragments. An additional $n-k$ parity fragments are then generated using the erasure code and the resultant n fragments are stored across n different nodes to better protect against node failures. If the erasure code of choice here is a maximum distance separable (MDS) code having dimension k and block length n, the DSS will tolerate up to any $n-k$ node failures without data loss.

In spite of being storage-efficient, traditional erasure coding techniques fall short in efficiently handling node failures in terms of metrics such as repair bandwidth and repair locality; repair bandwidth is the total amount of data download needed to repair a failed node, whereas repair locality is the number of helper nodes contacted for repair of a failed node. Clearly, it is desirable to minimize both the quantities. In response, researchers have come up with two different classes of codes, namely, regenerating codes [1] and locally recoverable codes (LRCs) [2], which are designed to reduce the repair bandwidth and repair locality, respectively. In this letter, our focus will be on a third class of codes, termed as locally regenerating codes (LRGCs) [3], [4], that is capable of reducing both repair bandwidth and repair locality, simultaneously.

Manuscript received July 22, 2020; revised August 18, 2020; accepted August 20, 2020. Date of publication August 24, 2020; date of current version December 10, 2020. The associate editor coordinating the review of this letter and approving it for publication was Y.-C. Huang. *(Corresponding author: M. Nikhil Krishnan.)*

M. Nikhil Krishnan is with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, M5S 1A1, Canada (e-mail: nikhilkrishnan.m@gmail.com).

V. Lalitha is with the Signal Processing and Communications Research Center, IIIT Hyderabad, Hyderabad 500 032, India (e-mail: lalitha.v@iiit.ac.in).

Sreeranjani Didugu is with the Department of Electrical and Computer Engineering, University of Wisconsin–Madison, Madison, WI 53706 USA (e-mail: didugu@wisc.edu).

Digital Object Identifier 10.1109/LCOMM.2020.3018937

A. Locally Recoverable Codes (LRCs)

For integers a, b, we use the notation $[a : b] \triangleq \{i : a \leq i \leq b\}$. Let $\mathcal{C}$ denote an $[n, k]$ linear code of block length n and dimension k over the finite field $\mathbb{F}_q$. The i-th code symbol c_i, $i \in [0 : n-1]$, of $\mathcal{C}$ is said to have (r, δ)-locality, if there exists a punctured code $\mathcal{C}_i$ (obtained by restricting $\mathcal{C}$ to some subset of coordinates, $S_i \subseteq [0 : n-1]$, where $i \in S_i$) such that $|S_i| \leq r + \delta - 1 \triangleq n_\ell$ and minimum distance of $\mathcal{C}_i$ is at least δ. The punctured code $\mathcal{C}_i$ will be referred to as a local code. A code where there exist k information symbols, all of which have (r, δ)-locality, will be referred to as a code with (r, δ)-information locality. Furthermore, if all the n code symbols have (r, δ)-locality, the code will be termed an (r, δ)-LRC. By virtue of the locality property, even in the presence of $\delta - 1$ erasures, every code symbol of an (r, δ)-LRC can be recovered by accessing at most r other code symbols. The following upper bound for minimum distance (d_{min}) of a code with (r, δ)-information locality is provided in [2], [3]:

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

B. Cooperative Regenerating Codes

In an (n, w, d, t) cooperative regenerating code, a file comprised of K message symbols from a finite field $\mathbb{F}_q$ is encoded into a set of $n\alpha$ code symbols and they are stored across n nodes in the network with each node storing α code symbols. The parameter α is called the subpacketization level of the code. A data collector can reconstruct the whole file, i.e., K message symbols, by connecting to any w nodes. In the event of t node failures, node repair is carried out in two phases. In the first phase, each of the replacement node connects to any $d \leq n - t$ nodes among the $n - t$ surviving nodes and downloads $\beta_1 \leq \alpha$ symbols from each node. In the second phase, the t replacement nodes exchange data among themselves. Every replacement node downloads β_2 symbols from every other replacement node. Hence, the repair bandwidth per replacement node is given by: $\gamma \triangleq d\beta_1 + (t-1)\beta_2$.

For fixed values of (n, w, d, t) and file-size K, there exists a tradeoff between α and γ. This is referred to as storage-repair bandwidth tradeoff. The point on the tradeoff where storage per node, α is minimum, is referred to as minimum storage cooperative regeneration (MSCR) point and the point where the repair bandwidth, γ is minimum, is referred to as minimum bandwidth cooperative regeneration (MBCR) point. The parameters of the MSCR and MBCR points are given below:

$$\text{MSCR: } (\alpha,\gamma) = \left(\frac{K}{w}, \frac{K(d+t-1)}{w(d+t-w)}\right), \tag{2}$$

$$\text{MBCR: } (\alpha,\gamma) = \left(\frac{K(2d+t-1)}{w(2d+t-w)}, \frac{K(2d+t-1)}{w(2d+t-w)}\right). \tag{3}$$

Cooperative regenerating codes reduce to regenerating codes by setting $t = 1$. MSCR and MBCR points become minimum storage regeneration (MSR) and minimum bandwidth regeneration (MBR) points, respectively, when $t = 1$. A formal study of the cooperative regenerating codes initially appeared in [5]–[8]. Explicit MBCR codes have been constructed in [9], [10]. Constructions of MSCR codes have been presented in [10]–[12].

C. Locally Regenerating Codes (LRGCs)

While LRCs tradeoff storage for repair locality, regenerating codes tradeoff storage for repair bandwidth. LRGCs are vector generalizations of LRCs (as we will see in the following discussion), where the local codes are regenerating codes themselves. The presence of local codes, as in the case of LRCs, ensures that LRGCs have a small repair locality. Furthermore, in LRGCs, repair bandwidth is low when repairing within the local code, as local codes are regenerating codes. LRGCs and cooperative regenerating codes belong to a general class of codes known as vector codes.

1) Vector Code: An $[n, K, d_{\min}, \alpha]$ linear vector code $\mathcal{C}_{\text{vec}}$ is a K-dimensional vector space over $\mathbb{F}_q$. Each element in $\mathcal{C}_{\text{vec}}$ (referred to as a codeword of $\mathcal{C}_{\text{vec}}$ from hereon) takes the form $(C_0, C_1, \ldots, C_{n-1})$, where $C_i \in \mathbb{F}_q^{\alpha}$, for some $\alpha \geq 1$. Naturally, each C_i will be referred to as a vector-code symbol. The minimum distance $d_{\min}$ of $\mathcal{C}_{\text{vec}}$ is the minimum number of vector-code symbols in which any two codewords differ or equivalently, minimum Hamming weight (at vector-code-symbol-level) among all non-zero codewords of $\mathcal{C}_{\text{vec}}$. When $\alpha = 1$, we refer to a vector code as a scalar code (e.g., LRCs are scalar codes). Clearly, an $[n, K, d_{\min}, \alpha = 1]$ scalar code is nothing but an $[n, K]$ linear code having minimum distance $d_{\min}$. Henceforth, we use the terms $[n, k]$ linear code and $[n, k]$ scalar code, interchangeably.

It is straightforward to note that, associated with $\mathcal{C}_{\text{vec}}$, there exists an $[n\alpha, K]$ scalar code $\mathcal{C}_{\text{sc}}$, which is obtained by expanding each vector-code symbol of a codeword in $\mathcal{C}_{\text{vec}}$ as α scalar symbols in some prescribed order.

2) Uniform Rank Accumulation Code: Let $\mathcal{C}_{\text{vec}}|_S$ denote the restriction of $\mathcal{C}_{\text{vec}}$ to the coordinates $S \subseteq [0 : n-1]$. $\mathcal{C}_{\text{vec}}$ is said to be a uniform rank accumulation (URA) code with rank accumulation profile $(a_1, a_2, \ldots, a_n)$ if dimension of $\mathcal{C}_{\text{vec}}|_S$, denoted by $\dim(\mathcal{C}_{\text{vec}}|_S)$, where $|S| = s$, is independent of the choice of S and $\dim(\mathcal{C}_{\text{vec}}|_S) = \sum_{j=1}^{s} a_j$. It has been shown in [13] that MSR and MBR codes are URA codes.

3) Locally URA Codes: Recall the definition of (r, δ)-locality introduced in Section I-A, in the context of LRCs (which are scalar codes). The definition naturally extends to vector codes as it is. A vector code $\mathcal{C}_{\text{vec}}$ is said to be a locally URA code if for every coordinate $i \in [0 : n-1]$, there exists an S_i satisfying $i \in S_i$ and $|S_i| \triangleq n_\ell$, such that $\mathcal{C}_{\text{vec}}|_{S_i}$ is a URA code. Clearly, it follows from the definition of (r, δ)-locality that i-th vector-code symbol of $\mathcal{C}_{\text{vec}}$ has $(n_\ell - d_{\min}(\mathcal{C}_{\text{vec}}|_{S_i}) + 1, d_{\min}(\mathcal{C}_{\text{vec}}|_{S_i}))$-locality.

4) Minimum Distance Bound for Locally URA Codes: Let $\mathcal{C}$ denote a locally URA code with parameters $[n, K, d_{\min}, \alpha]$. The local codes are all assumed to be identical, URA codes with parameters $[n_\ell, K_\ell, d'_{\min}, \alpha]$ and rank accumulation profile $(a_1, a_2, \ldots, a_{n_\ell})$. In order to state the minimum distance bound for $\mathcal{C}$ obtained in [3], we first define a function $P(.)$ based on the rank accumulation profile $(a_1, a_2, \ldots, a_{n_\ell})$ of the URA code. Construct the semi-infinite sequence $b_1, b_2, b_3, \ldots$, where $b_{i+jn_\ell} \triangleq a_i$, for $1 \leq i \leq n_\ell$ and $j \geq 0$. For $s \geq 1$, let

$$P(s) \triangleq \sum_{i=1}^{s} b_i. \tag{4}$$

For $x \geq 1$, set $P^{(\text{inv})}(x) = y$, where y is the smallest integer such that $P(y) \geq x$. The minimum distance of $\mathcal{C}$ is then upper bounded as [3]:

$$d_{\min} \leq n - P^{(\text{inv})}(K) + 1. \tag{5}$$

A code satisfying (5) with equality is said to be minimum-distance-optimal. Note that $K \leq P(n - d_{\min} + 1)$. Thus, a locally URA code is defined to be rate-optimal if $K = P(n - d_{\min} + 1)$. Locally URA codes where the local codes are MBR codes have been constructed in [3], [14]. Locally URA codes where the local codes are MSR codes have been presented in [3], [4], [14]. We will refer to these two classes of codes as locally MBR codes and locally MSR codes, respectively. Partial MDS codes where the local codes are MSR codes have been constructed in [15]. Locally MSR codes possessing distributed parities across local codes are presented in [16]. These codes allow for node repairs with lower repair locality even when the number of node failures are beyond the correction capability of local codes.

D. Our Contributions

In this letter, we present a family of minimum-distance-optimal, rate-optimal, locally MBCR codes (these are locally URA codes, where local codes are all MBCR codes) for all feasible values of $t \geq 1$. Our construction improves upon the existing works [3], [14] as follows:

- The letters [3], [14] construct locally MBCR codes for the special case of $t = 1$. It can be inferred from (3) that for a given file-size (dimension) K_ℓ, $(n_\ell, w, d, t > 1)$ MBCR codes offer a smaller storage per node (α) and repair bandwidth (γ) compared to an $(n_\ell, w, d, t = 1)$ code. Hence, by opting for $t > 1$, both storage and repair become more efficient.
- The present construction requires a finite field size only linear in block length and thus, generalizes the results

in [14]. The letter [3], on the other hand, requires a field size quadratic in code length at best.

II. PRELIMINARIES

A. Product-Matrix MBCR Codes

In this section, we give a summary of the product-matrix-based construction [10] of an MBCR code with parameters (n_ℓ, w, d, t), $\beta_1 = 2$, $\beta_2 = 1$ and $\alpha = 2d+t-1$. We refer the reader to [9] for the proof of data reconstruction and repair properties of the code. Consider a message matrix M formed of $w(2d-w+t)$ message symbols as follows:

$$M \triangleq \left[\begin{array}{c|c} A_{w\times w} & B_{w\times(d+t-w)} \\ \hline C_{(d-w)\times w} & 0_{(d-w)\times(d+t-w)} \end{array}\right]. \quad (6)$$

Let $\Gamma \triangleq \{\gamma_0, \gamma_1, \ldots, \gamma_{n_\ell-1}\}$ denote a set of n_ℓ distinct non-zero elements drawn from the finite field $\mathbb{F}_q$. For $l \in [0 : n_\ell - 1]$, the l-th vector-code symbol consists of the following $2d+t$ scalar symbols:

$$\begin{aligned} d \ \text{ symbols: } \ \mathbf{u}_l &= M[1 \ \gamma_l \ \gamma_l^2 \ \ldots, \ \gamma_l^{d+t-1}]^T, \\ d+t \ \text{ symbols: } \ \mathbf{v}_l &= M^T[1 \ \gamma_l \ \gamma_l^2 \ \ldots, \ \gamma_l^{d-1}]^T. \end{aligned}$$

There is a linear dependence relation among the $2d+t$ scalar code symbols within each vector-code symbol as given by:

$$[1 \ \gamma_l \ \gamma_l^2 \ \ldots, \ \gamma_l^{d-1}]\mathbf{u}_l = [1 \ \gamma_l \ \gamma_l^2 \ \ldots, \ \gamma_l^{d+t-1}]\mathbf{v}_l.$$

Hence, it suffices to store $\alpha = 2d+t-1$ scalar code symbols as the l-th vector-code symbol. However, in this letter, for simplicity in exposition, when we refer to the l-th vector-code symbol, we consider all the $2d+t$ symbols, i.e., $C_l \triangleq [\mathbf{u}_l^T \ \mathbf{v}_l^T]^T$.

It is shown in [9] that an (n_ℓ, w, d, t) MBCR code is a URA code with rank accumulation profile given by:

$$a_i = \begin{cases} (d-i+1)\beta_1 + (t-1)\beta_2, & 1 \le i \le w \\ 0, & (w+1) \le i \le n_\ell. \end{cases} \quad (7)$$

Interpretation Based on Polynomial Evaluations: For convenience of notation, let entries of M in (6) be given by $m_{i,j}$, $0 \le i \le d-1, 0 \le j \le d+t-1$, where clearly $m_{i,j} \triangleq 0$, if $\{i \ge w, j \ge w\}$. We denote row polynomials formed of the rows of the message matrix as follows: $p_i(x) = \sum_{j=0}^{d+t-1} m_{i,j}x^j, \ \ 0 \le i \le d-1$. Similarly, we denote column polynomials formed of the columns of the message matrix as follows: $q_j(x) = \sum_{i=0}^{d-1} m_{i,j}x^i, \ \ 0 \le j \le d+t-1$. It is straightforward to see that the d entries of $\mathbf{u}_l$ are evaluations of the d polynomials $\{p_i(x)\}$ at $x = \gamma_l$. Similarly, entries of $\mathbf{v}_l$ are essentially evaluations of the $d+t$ polynomials $\{q_j(x)\}$ at $x = \gamma_l$. Thus, we say the product-matrix-based MBCR code is obtained as evaluation of a message matrix M at Γ.

B. Tamo-Barg Codes

In this section, we describe the construction of a family of (r, δ)-LRCs, which is minimum-distance-optimal, i.e., these codes satisfy (1) with equality. Let n, k respectively denote the block length and dimension of the code. We set $\nu \triangleq \frac{n}{r+\delta-1}$. Let $\{\mathcal{A}_0, \ldots, \mathcal{A}_{\nu-1}\}$ be a partition of a set $\mathcal{A} \subset \mathbb{F}_q$, where $|\mathcal{A}| = n, |\mathcal{A}_i| = r+\delta-1 \triangleq n_\ell, 0 \le i \le \nu-1$. Consider a polynomial $g(x)$ which satisfies the following two properties:

- $g(x) \in \mathbb{F}_q[x]$ has degree n_ℓ and $g(x)$ evaluates to a constant μ_i on each set $\mathcal{A}_i$, i.e., $g(a) = \mu_i, \ \forall a \in \mathcal{A}_i$.

The k message symbols are arranged as coefficients of a 'message polynomial' as follows:

$$m(x) = \sum_{\substack{i=0 \\ i \bmod n_\ell = 0,1,\ldots,r-1}}^{k-1+(\lceil \frac{k}{r} \rceil -1)(\delta-1)} m_i g(x)^{\left\lfloor \frac{i}{n_\ell} \right\rfloor} x^{i \bmod n_\ell}.$$

The message polynomial is then evaluated at all the n elements in $\mathcal{A}$ to obtain the length-n codeword corresponding to the k message symbols. This completes the description of the code. One natural choice of the polynomial $g(x)$ and the sets $\mathcal{A}_i$ is as follows. Let field-size q be such that n divides $q-1$ and ζ be a primitive n-th root of unity in $\mathbb{F}_q$. The sets $\{\mathcal{A}_i\}$ can now be set to be: $\mathcal{A}_i \triangleq \zeta^i\{1, \lambda, \ldots, \lambda^{n_\ell-1}\}, 0 \le i \le \nu-1$, where $\lambda \triangleq \zeta^\nu$. The polynomial $g(x)$ can be chosen to be $g(x) \triangleq x^{n_\ell}$. Note that $a^{n_\ell} = \text{constant} \triangleq \mu_i, \ \forall \ a \in \mathcal{A}_i$. We will consider this version of the Tamo-Barg code in making a connection to our construction of locally MBCR code in the next section.

III. CONSTRUCTION OF OPTIMAL LOCALLY MBCR CODES

In this section, we will present our construction of locally MBCR codes. The construction is both minimum-distance-optimal and rate-optimal (Section I-C defines these notions of optimality). In our construction, each local code is an MBCR code obtained using the product-matrix-based construction (see Section II-A). A summary of notation is provided below.

- Let $\mathcal{C}$ denote the locally MBCR code to be constructed. $\mathcal{C}$ takes parameters $[n, K, d_{\min}, \alpha]$.
- Each local code $\mathcal{C}_\ell$ is an (n_ℓ, w, d, t) MBCR code with dimension $K_\ell \le K$. Thus, as per the vector code notation, the parameters of $\mathcal{C}_\ell$ are $[n_\ell, K_\ell, d'_{\min}, \alpha]$.
- We assume that n_ℓ divides n. Let $\nu \triangleq \frac{n}{n_\ell}$. $\mathcal{C}$ consists of ν disjoint local codes, which cover all the $n = \nu n_\ell$ nodes. Thus, $K \le \nu K_\ell$.

From Section II-A, we have $\alpha = 2d+t-1, \beta_1 = 2, \beta_2 = 1$ and $K_\ell = w(2d+t-w)$.

Construction 3.1: Let dimension K of $\mathcal{C}$ be expressed in terms of dimension K_ℓ of $\mathcal{C}_\ell$ as $K = aK_\ell + b$, where $1 \le a \le \nu-1$, $1 \le b \le K_\ell$. Here, b is of the form $b = s(2d+t-s), 1 \le s \le w$. Let field-size q be such that n divides $q-1$ and ζ be a primitive n-th root of unity in $\mathbb{F}_q$. Recall the definitions $\lambda \triangleq \zeta^\nu, \mathcal{A}_i \triangleq \zeta^i\{1, \lambda, \ldots, \lambda^{n_\ell-1}\}, 0 \le i \le \nu-1$. We will describe the locally MBCR code construction in three steps. The construction can be interpreted as obtained by stitching together Tamo-Barg code and product-matrix MBCR code. We will make the connection more precise after describing the construction.

Step-1 Construction of global message matrices $\{\tilde{M}^{(l)}\}_{l=0}^{\nu-1}$: Let $\tilde{M}^{(l)} = [\tilde{m}_{i,j}^{(l)}], 0 \le l \le \nu-1, 0 \le i \le d-1, 0 \le j \le d+t-1$. We set $\tilde{m}_{i,j}^{(l)} = 0$ if $\{i, j, l\}$ satisfies at least one of the following three cases: (i) $l > a$ (ii) $\{0 \le l < a, w \le i \le d-1, w \le j \le d+t-1\}$ (iii) $\{l = a, s \le i \le d-1, s \le j \le d+t-1\}$. It can now be inferred that there exist $K \triangleq aK_\ell + b$

non-trivial (i.e., not trivially zero) entries in total among all $\{\tilde{M}^{(l)}\}$, since $(a+1)(d)(d+t)-a(d-w)(d+t-w)-(d-s)(d+t-s)=aw(2d+t-w)+s(2d+t-s)=aK_\ell+b=K$. These K non-trivial entries of $\{\tilde{M}^{(l)}\}$ will be assigned the K message symbols.

Step-2 Construction of local message matrices $\{M^{(l)}\}_{l=0}^{\nu-1}$: Let $\mu_l \triangleq \zeta^{ln_\ell}$. The local message matrix $M^{(l)}$ corresponding to the l-th local code, $0 \le l \le \nu-1$ is obtained as $M^{(l)} = \sum_{i=0}^{\nu-1} \mu_l^i \tilde{M}^{(i)}$.

Step-3 Evaluation of local message matrices: The l-th local code will be obtained via evaluating the associated d row polynomials and $d+t$ column polynomials of $M^{(l)}$ at the n_ℓ points given by $\mathcal{A}_l$. This completes the construction.

We will make a precise connection of the above code with the Tamo-Barg codes. For $0 \le i \le d-1, 0 \le j \le d+t-1$, let $p_i^{(l)}(x), q_j^{(l)}(x)$ respectively denote the i-th row polynomial and j-th column polynomial corresponding to the local message matrix $M^{(l)}$. The collection of $n=\nu n_\ell$ code symbols $\cup_{l=0}^{\nu-1} \cup_{\theta_l \in \mathcal{A}_l} p_i^{(l)}(\theta_l)$ obtained via evaluating i-th row polynomials, and similarly, the n symbols $\cup_{l=0}^{\nu-1} \cup_{\theta_l \in \mathcal{A}_l} q_j^{(l)}(\theta_l)$ obtained via evaluating j-th column polynomials, can be shown to result in Tamo-Barg codes.

Example 1: Here, we consider the construction of a locally MBCR code with parameters $n=18, K=31$. Let the local MBCR code be specified as follows: $n_\ell=6, w=2, d=3, t=2$ and $K_\ell=12$. As $K=31$ and $K_\ell=12$, we have $K=2K_\ell+7$. Thus, $a=2$ and $b=7$. The rank accumulation profile of the MBCR code with parameters given above is $(7,12,12,12,12,12)$. Thus, we have $s=P^{\text{inv}}(b)=1$. Number of local codes, $\nu=\frac{n}{n_\ell}=3$. As n needs to divide $q-1$, we consider the finite field $\mathbb{F}_{19}$. Let $\zeta=2$, which is a primitive 18-th root of unity in $\mathbb{F}_{19}$. The ν sets of evaluation points and corresponding μ_i's are given by:

$$\begin{aligned}
\mathcal{A}_0 &= \{1,\zeta^3,\zeta^6,\zeta^9,\zeta^{12},\zeta^{15}\}, \quad \mu_0=1\\
\mathcal{A}_1 &= \{\zeta,\zeta^4,\zeta^7,\zeta^{10},\zeta^{13},\zeta^{16}\}, \quad \mu_1=\zeta^6\\
\mathcal{A}_2 &= \{\zeta^2,\zeta^5,\zeta^8,\zeta^{11},\zeta^{14},\zeta^{17}\}, \quad \mu_2=\zeta^{12}.
\end{aligned}$$

The ν global message matrices holding the K message symbols take the following form:

$$\tilde{M}^{(l)} = \begin{bmatrix} \tilde{m}_{0,0}^{(l)} & \tilde{m}_{0,1}^{(l)} & \tilde{m}_{0,2}^{(l)} & \tilde{m}_{0,3}^{(l)} & \tilde{m}_{0,4}^{(l)} \\ \tilde{m}_{1,0}^{(l)} & \tilde{m}_{1,1}^{(l)} & \tilde{m}_{1,2}^{(l)} & \tilde{m}_{1,3}^{(l)} & \tilde{m}_{1,4}^{(l)} \\ \tilde{m}_{2,0}^{(l)} & \tilde{m}_{2,1}^{(l)} & 0 & 0 & 0 \end{bmatrix}, \quad l=0,1, \tag{8}$$

$$\tilde{M}^{(2)} = \begin{bmatrix} \tilde{m}_{0,0}^{(2)} & \tilde{m}_{0,1}^{(2)} & \tilde{m}_{0,2}^{(2)} & \tilde{m}_{0,3}^{(2)} & \tilde{m}_{0,4}^{(2)} \\ \tilde{m}_{1,0}^{(2)} & 0 & 0 & 0 & 0 \\ \tilde{m}_{2,0}^{(2)} & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{9}$$

Note that there are $K=2*12+7=31$ entries which are trivially not zero, across the three global-message-matrices. The message matrices corresponding to the ν local MBCR codes are now given by:

$$M^{(l)} = \begin{bmatrix} m_{0,0}^{(l)} & m_{0,1}^{(l)} & m_{0,2}^{(l)} & m_{0,3}^{(l)} & m_{0,4}^{(l)} \\ m_{1,0}^{(l)} & m_{1,1}^{(l)} & m_{1,2}^{(l)} & m_{1,3}^{(l)} & m_{1,4}^{(l)} \\ m_{2,0}^{(l)} & m_{2,1}^{(l)} & 0 & 0 & 0 \end{bmatrix}, \tag{10}$$

where $l \in [0:2], m_{i,j}^{(l)} = \tilde{m}_{i,j}^{(0)} + \tilde{m}_{i,j}^{(1)}\mu_l + \tilde{m}_{i,j}^{(2)}\mu_l^2$.

Theorem 3.2: The code $\mathcal{C}$ obtained via Construction 3.1 is a locally MBCR code having dimension K, where the local MBCR codes have parameters (n_ℓ, w, d, t) and dimension K_ℓ.

Proof: In order to show that $\mathcal{C}$ indeed has dimension K, we first recall that the total number of message symbols embedded in the ν global message matrices $\tilde{M}^{(\ell)}, 0 \le \ell \le a$ is K. Note that $\{\mu_l\}_{l=0}^{\nu-1}$ can be verified to be distinct elements of $\mathbb{F}_q$ and hence, the following $\nu \times \nu$ Vandermonde matrix Λ is non-singular:

$$\Lambda \triangleq \begin{bmatrix} 1 & \mu_0 & \cdots & \mu_0^{\nu-1} \\ 1 & \mu_1 & \cdots & \mu_1^{\nu-1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \mu_{\nu-1} & \cdots & \mu_{\nu-1}^{\nu-1} \end{bmatrix}.$$

From the non-singularity of Λ, it follows that the mapping from $\{\tilde{M}^{(l)}\}_{l=0}^{\nu-1}$ to $\{M^{(l)}\}_{l=0}^{\nu-1}$ is a bijection. Given the vector-code symbols belonging to a local code-l (which is obtained via evaluating row and column polynomials of $M^{(l)}$ at $\mathcal{A}_l$), the coefficients of row and column polynomials of $M^{(l)}$ can be recovered. As these coefficients are nothing but entries of the message matrix $M^{(l)}$, the mapping of evaluating $\{M^{(l)}\}$ to create local codes is another bijection. It follows that the code $\mathcal{C}$ has dimension K.

In order to prove that the local codes are MBCR codes, we note that each $M^{(l)}$ has the structure depicted in (6). Together with the linearity of the code, it follows that each local code is a subcode of an (n_ℓ, w, d, t) MBCR code. We just need to show that as the K message symbols that $\mathcal{C}$ can encode vary over the q^K choices, $M^{(l)}$ takes all q^{K_ℓ} possible realizations. In order to see this, consider the entry $m_{i,j}^{(l)}$ of $M^{(l)}$, where $0 \le i \le d-1$, $0 \le j \le d+t-1$ and i,j are not both $\ge w$ (when both of them are $\ge w$, $m_{i,j}^{(l)}=0$). We have:

$$m_{i,j}^{(l)} = \sum_{h=0}^{\nu-1} \mu_l^h \tilde{m}_{i,j}^{(h)}. \tag{11}$$

It can be verified from Step 1 of Construction 3.1 that the global-message-matrix $\tilde{M}^{(0)}$ can take all the possible q^{K_ℓ} realizations (since $K \ge K_\ell$). Thus, for each i,j (in the non-trivial range) $\tilde{m}_{i,j}^{(0)}$ can independently take any value in $\mathbb{F}_q$. Combining this property of $\tilde{m}_{i,j}^{(0)}$ and (11), it follows that for a fixed l, $m_{i,j}^{(l)}$ can independently take any value in $\mathbb{F}_q$. Thus, $M^{(l)}$ lies in a space of dimension $d(d+t)-(d+t-w)(d-w)=w(2d+t-w)=K_\ell$. □

A. Distance and Rate Optimality of Locally MBCR Codes

In this section, we will prove that the locally MBCR code constructed in the previous section is minimum-distance optimal, i.e., it achieves the bound in (5) with equality. Also, we argue that the codes are rate-optimal.

Theorem 3.3: The code given in Construction 3.1 is minimum-distance-optimal, i.e.,

$$d_{min} = n - P^{inv}(K) + 1 = n - an_\ell - s + 1.$$

In addition, the code is also rate-optimal.

Proof: From (5), it is clear that the minimum distance of the code given in Construction 3.1 is upper bounded as:

$d_{\min} \leq n - P^{\text{inv}}(K) + 1 = n - an_\ell - s + 1$. In order to prove that $d_{\min} = n - P^{\text{inv}}(K) + 1 = n - an_\ell - s + 1$, we will show that the Hamming weight (at vector-code-symbol-level) of every non-zero codeword in $\mathcal{C}$ is at least $n - an_\ell - s + 1$. For $i \in [0 : d-1]$, consider the evaluation of the i-th row polynomial of $M^{(l)}$ at some $\theta_l \in \mathcal{A}_l$:

$$\begin{aligned} p_i^{(l)}(\theta_l) &= \sum_{j=0}^{d+t-1} \left(\sum_{h=0}^{\nu-1} \tilde{m}_{i,j}^{(h)} \mu_l^h \right) x^j \mid_{x=\theta_l} \\ &= \sum_{h=0}^{\nu-1} \sum_{j=0}^{d+t-1} \tilde{m}_{i,j}^{(h)} x^{hn_\ell + j} \mid_{x=\theta_l}, \end{aligned} \tag{12}$$

where we used the fact that $\mu_l = \theta_l^{n_\ell} \ \forall \theta_l \in \mathcal{A}_l$.

Similarly, for $j \in [0 : d+t-1]$, the evaluation of j-th column polynomial of $M^{(l)}$ at some $\theta_l \in \mathcal{A}_l$ can be expressed as:

$$q_j^{(l)}(\theta_l) = \sum_{h=0}^{\nu-1} \sum_{i=0}^{d-1} \tilde{m}_{i,j}^{(h)} x^{hn_\ell + i} \mid_{x=\theta_l} . \tag{13}$$

Consider any non-zero codeword $\underline{c} \in \mathcal{C}$. For the given codeword $\underline{c}$, recall that each vector-code symbol is obtained as evaluations of d row polynomials and $d+t$ column polynomials corresponding to some $M^{(l)}$ at some $\theta_l \in \mathcal{A}_l$. For any fixed $i \in [s : d-1]$, select from each vector-code symbol belonging to the l-th local code and associated with the evaluation point θ_l, the scalar symbol $p_i^{(l)}(\theta_l)$. Let the n scalar symbols obtained from n vector-code symbols be arranged as a length-n scalar codeword $\underline{c}_{p,i}$. From Step-1 of Construction 3.1, we have that $\tilde{m}_{i',j'}^{(h')} = 0$ when $h' > a, i' \in [0 : d-1], j' \in [0 : d+t-1]$. Also, $\tilde{m}_{i',j'}^{(a)} = 0$ when both $s \leq i' \leq d-1$ and $s \leq j' \leq d+t-1$ are true. Thus, for any $i \in [s : d-1]$, the largest possible degree the polynomial in (12) can take is $an_\ell + s - 1$. Hence, evaluations of (12) at $n = |\cup_{l \in [0:\nu-1]} \mathcal{A}_l|$ points results in at most $an_\ell + s - 1$ zeros or else, all the coefficients of (12) are zeros. These n evaluations of (12) are precisely the entries of $\underline{c}_{p,i}$. Thus, $\underline{c}_{p,i}$ has a Hamming weight of at least $n - (an_\ell + s - 1) = n - P^{\text{inv}}(K) + 1$ or else, all the coefficients of (12) are zeros.

Now, we can assume that $\tilde{m}_{i',j'}^{(h)} = 0$ if $i' \in [s : d-1]$, as otherwise, from the previous case, we know that the codeword $\underline{c}$ has a Hamming weight (at vector-code-symbol-level) of at least $n - P^{\text{inv}}(K) + 1$. Fix a $j \in [0 : d+t-1]$. Together with the fact that $\tilde{m}_{i',j'}^{(h')} = 0$ for $h' > a, i' \in [0 : d-1], j' \in [0 : d+t-1]$ by design, it can now be easily verified that for (13), the largest degree possible is $an_\ell + s - 1$. Hence, $\underline{c}_{q,j}$ (analogous to $\underline{c}_{p,i}$, obtained by selecting the scalar symbol corresponding to $q_j^{(l)}(\theta_l)$ from each vector-code symbol) has a Hamming weight of at least $n - (an_\ell + s - 1) = n - P^{\text{inv}}(K) + 1$, or else, all the coefficients of (13) are zeros. As $\underline{c}$ is a non-zero codeword, there exists a choice of $\{h', i', j'\}$, $h' \in [0 : a], i' \in [0 : d-1], j' \in [0 : d+t-1]$, for which $\tilde{m}_{i',j'}^{(h')} \neq 0$. As can be seen from the two cases discussed above, if $i' \in [s : d-1]$, Hamming weight of $\underline{c}_{p,i'}$ is at least $n - P^{\text{inv}}(K) + 1$. Otherwise, $\underline{c}_{q,j'}$ has a Hamming weight of at least $n - P^{\text{inv}}(K) + 1$. This ensures that Hamming weight (at vector-code-symbol-level) of $\underline{c}$ is also at least $n - P^{\text{inv}}(K) + 1$.

Rate-Optimality: As our codes are designed to be minimum-distance-optimal, we have: $P^{(\text{inv})}(K) = n - d_{\min} + 1 \implies K \leq P(P^{(\text{inv})}(K)) = P(n - d_{\min} + 1)$. It follows that the minimum-distance-optimal $\mathcal{C}$ can be rate-optimal iff $K = P(P^{(\text{inv})}(K))$. Since we have chosen $b = s(2d+t-s)$, we have that $P(s) = b$ and hence it follows that $P(P^{(\text{inv})}(K)) = K$. Thus, the codes constructed are rate-optimal. □

IV. Conclusion

This letter presents the first-known family of locally MBCR codes. Compared to locally MBR codes, these codes are more efficient in terms of storage and repair bandwidth. However, it is to be noted that the repair process in locally MBCR needs more coordination among failed nodes, because of the 2-step repair process inherent to MBCR codes. As a future direction, it is an open problem to construct optimal locally MSCR codes.

References

[1] A. G. Dimakis, P. B. Godfrey, Y. Wu, M. J. Wainwright, and K. Ramchandran, "Network coding for distributed storage systems," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4539–4551, Sep. 2010.

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

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

[4] A. S. Rawat, O. O. Koyluoglu, N. Silberstein, and S. Vishwanath, "Optimal locally repairable and secure codes for distributed storage systems," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 212–236, Jan. 2014.

[5] Y. Hu, Y. Xu, X. Wang, C. Zhan, and P. Li, "Cooperative recovery of distributed storage systems from multiple losses with network coding," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 2, pp. 268–276, Feb. 2010.

[6] X. Wang, Y. Xu, Y. Hu, and K. Ou, "MFR: Multi-loss flexible recovery in distributed storage systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2010, pp. 1–5.

[7] A.-M. Kermarrec, N. Le Scouarnec, and G. Straub, "Repairing multiple failures with coordinated and adaptive regenerating codes," in *Proc. Int. Symp. Netw. Coding*, Jul. 2011, pp. 1–6.

[8] K. W. Shum and Y. Hu, "Cooperative regenerating codes," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7229–7258, Nov. 2013.

[9] A. Wang and Z. Zhang, "Exact cooperative regenerating codes with minimum-repair-bandwidth for distributed storage," in *Proc. IEEE INFOCOM*, Apr. 2013, pp. 400–404.

[10] K. W. Shum and J. Chen, "Cooperative repair of multiple node failures in distributed storage systems," *Int. J. Inf. Coding Theory*, vol. 3, no. 4, pp. 299–323, 2016.

[11] N. Le Scouarnec, "Exact scalar minimum storage coordinated regenerating codes," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2012, pp. 1197–1201.

[12] M. Ye and A. Barg, "Cooperative repair: Constructions of optimal MDS codes for all admissible parameters," *IEEE Trans. Inf. Theory*, vol. 65, no. 3, pp. 1639–1656, Mar. 2019.

[13] N. B. Shah, K. V. Rashmi, P. V. Kumar, and K. Ramchandran, "Distributed storage codes with repair-by-transfer and nonachievability of interior points on the storage-bandwidth tradeoff," *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1837–1852, Mar. 2012.

[14] M. N. Krishnan, R. A. Narayanan, and P. V. Kumar, "Codes with combined locality and regeneration having optimal rate, d_{min} and linear field size," in *Proc. Int. Symp. Inf. Theory (ISIT)*, 2018, pp. 1196–1200.

[15] L. Holzbaur, S. Puchinger, E. Yaakobi, and A. Wachter-Zeh, "Partial MDS codes with local regeneration," Jan. 2020, *arXiv:2001.04711*. [Online]. Available: https://arxiv.org/abs/2001.04711

[16] J. Wang, Z. Yan, K.-C. Li, H. Xie, and X. Liu, "Local codes with cooperative repair in distributed storage of cyber-physical-social systems," *IEEE Access*, vol. 8, pp. 38622–38632, 2020.