Repairing Schemes for Tamo-Barg Codes

Han Cai, *Member, IEEE*, Ying Miao, Moshe Schwartz, *Fellow, IEEE*, and Xiaohu Tang, *Senior Member, IEEE*

***Abstract*— In this paper, the repair problem for erasures beyond locality in locally repairable codes is explored under a practical system setting, where a rack-aware storage system consists of racks, each containing a few parity checks. This is referred to as a rack-aware system with locality. Two repair schemes are devised to reduce the repair bandwidth for Tamo-Barg codes under the rack-aware model by setting each repair set as a rack. Additionally, a cut-set bound for locally repairable codes under the rack-aware model with locality is introduced. Using this bound, the second repair scheme is proven to be optimal. Furthermore, the partial-repair problem is considered for locally repairable codes under the rack-aware model with locality, and both repair schemes and bounds are introduced for this scenario.n this paper, the repair problem for erasures beyond locality in locally repairable codes is explored under a practical system setting, where a rack-aware storage system consists of racks, each containing a few parity checks. This is referred to as a rack-aware system with locality. Two repair schemes are devised to reduce the repair bandwidth for Tamo-Barg codes under the rack-aware model by setting each repair set as a rack. Additionally, a cut-set bound for locally repairable codes under the rack-aware model with locality is introduced. Using this bound, the second repair scheme is proven to be optimal. Furthermore, the partial-repair problem is considered for locally repairable codes under the rack-aware model with locality, and both repair schemes and bounds are introduced for this scenario.**

***Index Terms*— Distributed storage, locally repairable codes, Tamo-Barg codes, rack-aware system with locality, regenerating codes.**

Received 20 December 2023; revised 25 July 2024; accepted 6 November 2024. Date of publication 14 November 2024; date of current version 26 December 2024. The work of Han Cai was supported in part by the National Key Research and Development Program of China under Grant 2022YFA1005000 and in part by the National Natural Science Foundation of China under Grant 62271421. The work of Ying Miao was supported in part by Japan Society for the Promotion of Science Grant-in-Aid for Scientific Research (C) under Grant 24K06824. An earlier version of this paper was presented in part at the 2024 IEEE International Symposium on Information Theory [DOI: 10.1109/ISIT57864.2024.10619492]. *(Corresponding author: Han Cai.)*

Han Cai and Xiaohu Tang are with the Information Coding & Transmission Key Laboratory of Sichuan Province, CSNMT Int. Coop. Res. Centre (MoST), Southwest Jiaotong University, Chengdu 611756, China (e-mail: hancai@aliyun.com; xhutang@swjtu.edu.cn).

Ying Miao is with the Faculty of Engineering, Information and Systems, University of Tsukuba, Tsukuba 305-8573, Japan (e-mail: miao@sk.tsukuba.ac.jp).

Moshe Schwartz is with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4K1, Canada, on leave from the School of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Beer-Sheva 8410501, Israel (e-mail: schwartz.moshe@mcmaster.ca).

Communicated by I. Tamo, Associate Editor for Coding and Decoding.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TIT.2024.3498033.

Digital Object Identifier 10.1109/TIT.2024.3498033

I. Introduction

WITH the expanding volume of data in large-scale cloud storage and distributed file systems like Windows Azure Storage and Google File System (GoogleFS), disk failures have become a norm rather than an exception. To protect data from such failures, the simplest solution is to replicate data packets across different disks. However, this approach suffers from large storage overhead. Consequently, coding techniques have been developed as an alternative solution.

Several performance metrics have been introduced to assess the codes' performance, taking into account various aspects of the storage system that are of interest. In order to maximize failure tolerance and minimize redundancy, *maximum distance separable (MDS) codes* have been considered. To minimize the number of bits communicated during the repair procedure, codes called *regenerating codes with optimal repair bandwidth* have been developed [11]. At the same time, *locally repairable codes* have been proposed [13], to reduce the number of nodes participating in the repair process. To ensure that data can be frequently accessed by multiple processes in parallel, codes that support parallel reads were introduced [38], [45]. Finally, to improve the update and access efficiency, codes with optimal access and update properties have also been considered [33] and [42]. Over the past decade, many results have been obtained regarding codes for distributed storage systems according to these metrics, e.g., see [8], [11], [14], [18], [20], [25], [28], [31], [42], [43], [44], [46], [48], [49], and [50] for codes with optimal repair bandwidth, see [2], [5], [6], [7], [13], [15], [21], [24], [29], [37], [39], [40], and [47] for optimal locally repairable codes, see [3], [4], [22], [38], and [41] for codes with good availability, and see [9], [27], [33], and [42] for codes with optimal update and access properties.

In this paper, we focus on locally repairable codes and codes with optimal repair bandwidth. For a locally repairable code, all code symbols are partitioned into repair sets, each containing some redundancy to allow local repair. When a small prescribed number of erasures affect a repair set, the repair process is designed to be as easy as repairing a short MDS code, usually accessing far less data than the amount of the original data encoded into the entire codeword. However, when the erasure patterns exceed the local-repair ability, the repair problem is still open, which may result in a repair scheme similar to that of MDS codes, requiring a bandwidth as large as the original data. Our motivation is to consider the repair problem for erasures beyond locality for locally repairable codes, i.e., to find an efficient repair scheme for this kind of erasure patterns.

In the literature, a known approach to combine locality and regenerating codes is to include redundancies in each repair set, allowing the codes in each set to form regenerating codes, e.g., [17], [23], and [26]. By doing so, the repair bandwidth required can be reduced when the system performs local erasure repairs.

However, this method has a drawback: the repair property only works for the punctured codes in the repair sets. This means that if there are erasures beyond the local repair capability in one repair set, the repair scheme and the locality cannot simultaneously reduce the repair bandwidth. To address this issue, we propose a new combination strategy to repair erasures beyond the locality. Our approach involves repair schemes for locally repairable codes that can handle erasures beyond local recoverability. In addition, from a practical perspective, our idea is motivated by the observation that repair sets may be located on the same server or physically nearby servers, and communication within a repair set may be less expensive than communication across repair sets, which is the practical setting for rack-aware storage systems. Therefore, desirable repairing schemes should be able to reduce the bandwidth required for communication across repair sets.

Specifically, in this paper, we propose repair schemes for the well-known Tamo-Barg codes [40], which are optimal locally repairable codes with respect to the Singleton-type bound [13], [36]. We present two proposed schemes. Firstly, in a rack-aware model where each repair set is one rack, we introduce an optimal repair scheme for the case of one failed rack, i.e., one erased repair set. Secondly, for the scenario where there are erasures within a repair set that cannot be recovered locally, we introduce a repair scheme that reduces the repair bandwidth required for recovering those failures. We prove the optimality of our schemes by modifying the well-known cut-set bound [11] to incorporate locality. Our proposed schemes generalize the rack-aware model regenerating codes [18], [19].

The remainder of this paper is organized as follows. Section II introduces some preliminaries about locally repairable codes and regenerating codes. Section III introduces the basic system setting for rack-aware systems with locality. Section IV describes some basic properties for Tamo-Barg codes. Section V introduces schemes for repairing a whole rack erasure (repair set). Section VI presents a partial repair scheme for Tamo-Barg codes, which is capable of repairing an erased fraction of racks. Additionally, the scheme is proved to be optimal. Section VII concludes this paper with some remarks.

II. Preliminaries

We start by introducing basic notation and definitions. For any $n \in \mathbb{N}$ we denote $[n] \triangleq \{1,2,\ldots,n\}$. For a prime power q, let $\mathbb{F}_q$ denote the finite field of size q, $\mathbb{F}_q^* \triangleq \mathbb{F}_q \setminus \{0\}$, and let $\mathbb{F}_q[x]$ denote the set of polynomials in the indeterminate x with coefficients from $\mathbb{F}_q$. An $[n,k]_q$ linear code $\mathcal{C}$ over $\mathbb{F}_q$ is a k-dimensional subspace of $\mathbb{F}_q^n$ with a $k \times n$ generator matrix $G = (\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_n)$, where $\mathbf{g}_i$ is a column vector of length k for all $i \in [n]$. More specifically, it is called an $[n,k,d]_q$ linear code if its minimum Hamming distance is d. For a subset $S \subseteq [n]$, we use $\mathrm{span}(S)$ to denote the linear space spanned by $\{\mathbf{g}_i : i \in S\}$ over $\mathbb{F}_q$, and $\mathrm{rank}(S)$ to denote the dimension of $\mathrm{span}(S)$.

A. Generalized Reed-Solomon Codes

Let $\theta = (\theta_1, \ldots, \theta_n) \in \mathbb{F}_q^n$ contain distinct entries, where we assume $q \geqslant n$. Then the well-known *generalized Reed-Solomon (GRS) code* with parameters $[n, k, n-k+1]_q$ can be defined as

$$\mathrm{GRS}_k(\boldsymbol{\theta}, \boldsymbol{\alpha}) \triangleq \{(\alpha_1 f(\theta_1), \alpha_2 f(\theta_2), \ldots, \alpha_n f(\theta_n)) : f(x) \in \mathbb{F}_q[x] \text{ with } \deg(f(x)) < k\},$$

where $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in (\mathbb{F}_q^*)^n$. It is well known that the dual of an $[n,k,n-k+1]_q$ GRS code is an $[n, n-k, k+1]_q$ GRS code (e.g., see [32]).

B. Regenerating Codes

An important problem in distributed storage systems is to repair an erasure by downloading as little data as possible. Dimakis et al. [11] introduced *repair bandwidth*, the amount of data downloaded during a node repair, as a metric to measure the procedure's efficiency.

Definition 1: Let $\mathcal{C}$ be an $[N,K]_q$ array code with *sub-packetization* L, that is, $c_i \in \mathbb{F}_q^L$ for any codeword $(c_1, c_2, \ldots, c_N) \in \mathcal{C}$. For an erasure pattern $\mathcal{E} \subseteq [N]$ and a D-subset $\mathcal{R} \subseteq [N] \setminus \mathcal{E}$ (whose entries are called *helper nodes*), define $B(\mathcal{C}, \mathcal{E}, \mathcal{R})$ as the *minimum repair bandwidth* for $c_i \in \mathbb{F}_q^L$ stored in node $i \in \mathcal{E}$, i.e., the smallest total number of symbols of $\mathbb{F}_q$ helper nodes need to send in order to recover c_i (where each helper node $j \in \mathcal{R}$ may send symbols that depend solely on $c_j \in \mathbb{F}_q^L$).

Definition 2 [35]: Let $\mathcal{C}$ be an $L \times N$ array code over $\mathbb{F}_q$, i.e., $\mathcal{C} \subseteq \mathbb{F}_q^{L \times N}$. Then the dimension of the array code is defined as $K = \log_{q^L}(|\mathcal{C}|)$. Furthermore, the code is said to be *maximum distance separable (MDS)* code if

$$\min\{d_c(C_1, C_2) : C_1 \neq C_2, \text{ and } C_1, C_2 \in \mathcal{C}\} = N - K + 1,$$

where the column Hamming distance is defined as

$$d_c(C_1, C_2) = |\{i \in [N] : C_{1,i} \neq C_{2,i}, C_1 = (C_{1,1}, C_{1,2}, \cdots, C_{1,N}), \text{and} C_2 = (C_{2,1}, C_{2,2}, \cdots, C_{2,N})\}|.$$

In [11], the well-known cut-set bound was first derived for the minimum download bandwidth.

Theorem 1 (Cut-Set Bound, [1], [11]): Let $\mathcal{C}$ be an $[N,K]_q$ MDS array code with sub-packetization L.Let D be an integer with $K \leqslant D \leqslant N-1$. For any non-empty$\mathcal{E} \subseteq [N]$ with $|\mathcal{E}| \leqslant N - D$ and any D-subset $\mathcal{R} \subseteq [N] \setminus \mathcal{E}$, we have

$$B(\mathcal{C}, \mathcal{E}, \mathcal{R}) \geqslant \frac{DL}{D - K + |\mathcal{E}|}.$$

Definition 3: For $K < D \leqslant N - \tau$, an $[N,K]_q$ MDS array code is said to be an $[N,K]_q$ *minimum storage regenerating* (MSR) code with repair degree D, if for each $\mathcal{I} = \{i_1, i_2, \ldots, i_\tau\} \subset [N]$ and any D-subset $\mathcal{R}_\mathcal{I} \subseteq [N] \setminus \mathcal{I}$, the repair bandwidth $B(\mathcal{C}, \mathcal{I}, \mathcal{R}_\mathcal{I})$ meets the cut-set bound described above with equality. Throughout this paper, such codes are also said to have (τ, D) *optimal repair* property.

C. Locally Repairable Codes

Another important figure of merit is symbol locality [13], [36].

Definition 4: Let $\mathcal{C}$ be an $[n,k,d]_q$ linear code, and let G be a generator matrix for it. For $j \in [n]$, the j-th code symbol, c_j, of $\mathcal{C}$, is said to have (r,δ)-locality if there exists a subset $S_j \subseteq [n]$ such that:

- $j \in S_j$ and $|S_j| \leqslant r+\delta-1$; and
- the minimum Hamming distance of the punctured code $\mathcal{C}|_{S_j}$ is at least δ.

In that case, the set S_j is also called a *repair set* of c_j. The code $\mathcal{C}$ is said to have information (r,δ)-locality (denoted as $(r,\delta)_i$-locality) if there exists $S \subseteq [n]$ with $\operatorname{rank}(S)=k$ such that for each $i \in S$, the i-th code symbol has (r,δ)-locality. Similarly, $\mathcal{C}$ is said to have all symbol (r,δ)-locality (denoted as $(r,\delta)_a$-locality) if all the code symbols have (r,δ)-locality.

In [36] (and for the case $\delta = 2$, originally [13]), the following upper bound on the minimum Hamming distance of linear codes with information (r,δ)-locality is derived.

Lemma 1 ([13], [36]): The minimum distance, d, of an $[n,k,d]_q$ code with $(r,\delta)_i$-locality, is upper bounded by

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

Definition 5: A code is said to be an optimal locally repairable code (LRC) with $(r,\delta)_i$-locality (or $(r,\delta)_a$-locality) if its minimum distance d attains the bound of Lemma 1 with equality.

We remark that in this paper, we focus only on linear codes. For nonlinear codes with locality the reader may refer to [12] and the references therein.

III. Rack-Aware Distributed Storage System With Locality

In this section, we introduce some basic model settings for a rack-aware distributed storage system with local parity checks. In general, nodes or servers in a distributed storage system are placed on different racks, where these racks may be geographically isolated, and each rack has independent data processing capabilities. Typically, these racks contain a small number of local parity nodes (such as those with a parity-check disk). Additionally, to protect the data, several global parity-check racks are usually added to the system. However, existing systems are usually designed using a layered approach, for instance, employing MDS codes for encoding within the rack and then applying erasure codes at the rack level for cross rack encoding. In what follows, we will model these systems as a rack system with locality and define the repair problem for these systems. In our setting, the main idea is to consider the two-layer encoding system as a whole, that is, as a locally repairable code.

Specifically, we consider a system containing k original files, which are encoded into n files stored on n nodes (or servers). The $n = N \times L$ nodes are divided into N racks, and each rack contains L nodes. In each rack, the data of the servers form a codeword with length L and a minimum Hamming distance of at least δ.

Due to the data processing capabilities of each rack, in each rack, we assume that all nodes within a rack have data processing capabilities and can access data from other nodes. Denote this system as the $(n,k;L,\delta)$ rack-aware system with locality or $(n,k;L,\delta)$-RASL. If among these n nodes, k of them store original information and are named as information nodes then the system is said to be systematic. In this case, all the remaining nodes are parity checks, including the local parity checks and cross-rack parity checks. Formally, we define array codes for $(n=NL, k \leqslant Kr; L=r+\delta-1,\delta)$-RASL as follows.

Definition 6: An $L \times N$ array code $\mathcal{C}$ over $\mathbb{F}_q$ is said to be an $(N,K,k;L,\delta)$-RASL code when the code satisfies that:

- Any K columns are capable of recovering the whole codeword;
- For any given $i \in [N]$, the code formed by the i-th columns for all codewords, i.e., the punctured code over the i-th column has minimum Hamming distance at least δ;
- The parameter k defined as $k \triangleq \log_q|\mathcal{C}| \leqslant Kr$, where $r = L-\delta+1$ and the inequality follows from the well-known Singleton bound.

Remark 1: In the literature, there are two different kinds of codes related to the rack-aware distributed storage system model with locality.The first one is rack-aware regenerating codes, which are usually array codes with efficient repair schemes for failed nodes.The other is locally repairable codes which are usually scalar codes. In this paper we consider codes for rack-aware distributed storage systems, which are related to both of these kinds. In the view of rack-aware regenerating codes, we are going to investigate rack-aware regenerating codes such that each rack contains local parity checks. To avoid a “three dimensional system”, we only consider the scalar case, that is,each node is only an element of a given finite field $\mathbb{F}_q$ as in Definition 6 but not a vector over $\mathbb{F}_q^\ell$ as usual. When viewed differently, codes for rack-aware distributed storage systems can be regarded as array codes formed by rearranging repair sets of scalar locally repairable codes as columns. In this viewpoint, our motivation is to build efficient repair schemes across repair sets. Thus, in what follows,we use the notation “L”, to denote the number of nodes in each rack and the length of columns for the array codes, for these two different viewpoints.

As an example, a systematic $(N,K,k=Kr;r+\delta-1,\delta)$-RASL is depicted in Fig. 1.

In this system, data is distributed across several racks, each containing multiple nodes and $\delta-1$ local parity checks, i.e., the data in each rack is erasure coded with $\delta-1$ parity checks. Similar to RAID systems that utilize disk parity checks, this storage system employs $N-K$ racks for across rack parity checks to handle rack erasures effectively. In this example, in total there are $n-k$ parity checks including $N(\delta-1)$ local parity checks and $(N-K)r$ cross-rack parity checks.

In each rack of the rack-aware system, communication between racks is more costly than communication within a rack. Therefore, similar to rack regenerating codes [18], we only consider the communication bandwidth across racks and disregard the inner rack bandwidth. Since, in each rack, the

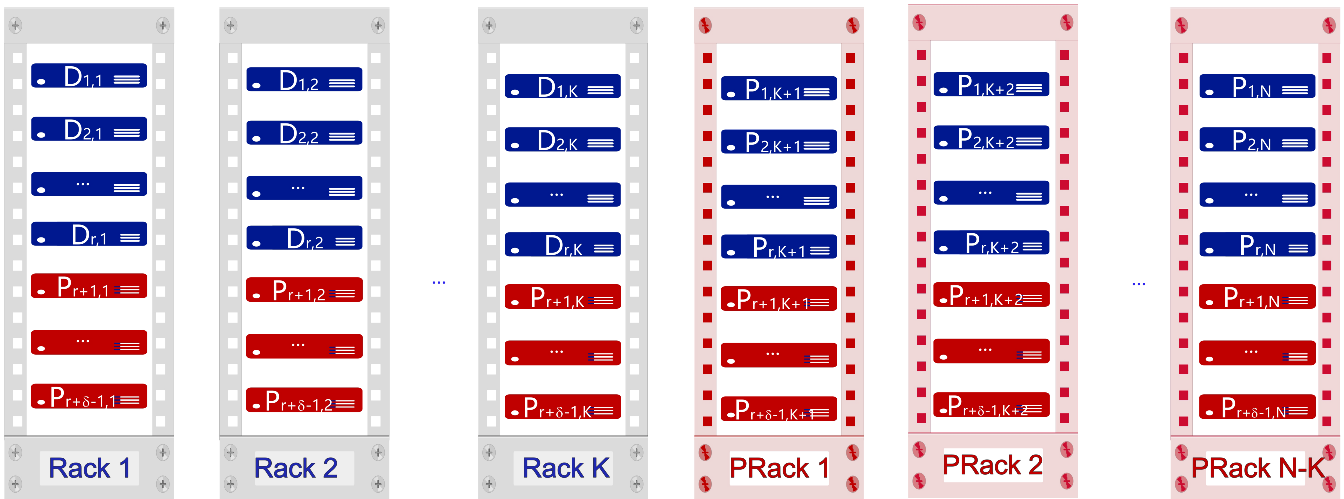

Fig. 1. The distributed storage system is organized with a racks-servers structure, where each rack consists of $r+\delta-1$ servers. Within each rack, $\delta-1$ parity checks are included to ensure data integrity. Additionally, to safeguard against rack erasures, $N-K$ rack parity checks (red columns) are stored.

data of servers form a codeword with a minimum Hamming distance of at least δ, when the rack suffers $\tau \leqslant \delta-1$ erasures, the rack can recover the erasures locally, where we disregard the inner rack bandwidth. Thus, we focus on racks that experience more than $\delta-1$ server erasures. Specifically, we consider the following two patterns of erasures:

- Rack erasures: there are λ racks erased in the systems;
- Partial erasures: there are λ racks that suffer more than $\delta-1$ nodes erasures(each).

An intriguing challenge for these systems involves minimizing cross-rack bandwidth when facing erasures that exceed the capabilities of the local parity checks within each rack. This challenge serves as the motivation for the subsequent part of our discussion. That is, we would like to determine the minimum amount of data we need download from help racks to repair erasures and the explicit construction of codes with optimal repair schemes.

Remark 2: The aforementioned rack-aware system with locality is a generalization of the rack-aware model in [18]. For explicit constructions of regenerating codes for rack-aware system the reader may refer to [8], [18], [20], [46], and [50]. This extension is primarily driven by the practical observation that modern storage systems incorporate both parity checks and controllers with data processing capabilities within each rack. In Fig. 2, we compare the rack-aware model and rack-aware model with locality for $n=18$, $k=10$, and $r=5$, where we set $\lambda=1$ for both rack erasure and partial erasure cases.

Remark 3: The term "locality" is derived from the condition that, for (n,k) locally repairable codes with $(r,\delta)_a$-locality, if the repair sets $S_1, S_2, \cdots, S_N$ constitute a partition of $[n]$ and each set has a uniform size, i.e., $|S_i| = r+\delta-1$ for $1 \leqslant i \leqslant N$, then each repair set can be arranged as a rack to construct the desired code for a rack-aware system. Consequently, when the repair sets form a partition, the repair problems for the rack-aware system with locality are equivalent to the repair problems of locally repairable codes, assuming that we disregard the bandwidth within the rack or the repair set. Therefore, in the subsequent discussion, we will use these notations interchangeably. As in Figure 2, we may apply a $(18,10)$ locally repairable code with $(5,2)$-locality in the $(18,10;6,2)$-RASL, when there are three disjoint repair sets with size 6 corresponding to three racks.

Remark 4: To simplify the system setting, we focus on the regular case, where each rack contains an equal number of servers.

IV. Tamo-Barg Codes and Redundant Residue Codes

In [40], Tamo and Barg proposed constructions of locally repairable codes. Among them, one of the constructions yields codes now named Tamo-Barg codes, and another is via the Chinese Remainder Theorem, namely, redundant residue codes. Tamo-Barg codes are in fact a special case of locally repairable codes via the Chinese Remainder Theorem [40]. However, in general, the locality of parity-check symbols for locally repairable codes based on the Chinese Remainder Theorem is not well understood. In this section, we review the construction of locally repairable codes using redundant residue codes and prove that it can, in fact, explain Tamo-Barg codes. We begin with the Chinese Remainder Theorem for polynomials over finite fields.

Lemma 2 [34]: Let $h_1(x), \ldots, h_t(x) \in \mathbb{F}_q[x]$ be pairwise co-prime polynomials. Then for any t polynomials $m_1(x), \ldots, m_t(x) \in \mathbb{F}_q[x]$, there exists a unique polynomial $f(x) \in \mathbb{F}_q[x]$ of degree less than $\sum_{i=1}^{t} \deg(h_i(x))$ such that

$$f(x) \equiv m_i(x) \pmod{h_i(x)} \quad \text{for all } i \in [t].$$

Construction A [40]: Let $h(x) \in \mathbb{F}_q[x]$ and denote $\deg(h(x)) = w$. For $y \in \mathbb{F}_q$,define $\mathrm{Roots}(y) \triangleq \{x \in \mathbb{F}_q : h(x) = y\}$ and $t_y \triangleq |\mathrm{Roots}(y)|$. Assume $m_1 \leqslant k$ and r_i, for all $i \in [m_1]$, are positive integers such that $\sum_{i=1}^{m_1} r_i = k$. We further assume that there exist two disjoint subsets of $\mathbb{F}_q$, $\{y_i\}_{i=1}^{m_1}$ and $\{y_i\}_{i=m_1+1}^{m_1+m_2}$, satisfying $t_{y_i} > r_i$ for all $i \in [m_1]$, and m_2 is a non-negative integer.

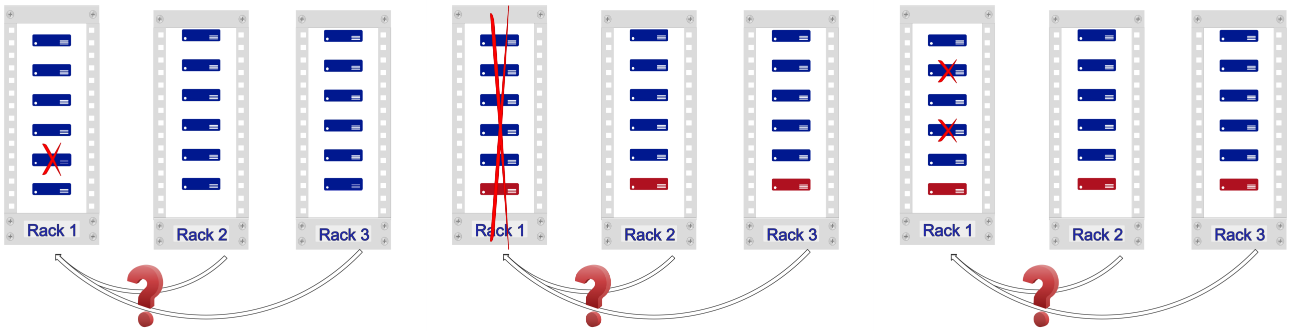

Fig. 2. As a comparison, setting $n = 18$, $k = 10$, the first figure shows the rack-aware model with one erasure, the second one is corresponding to the $(18, 10; 6, 2)$-RASL for one rack erasure and the last is for the $(18, 10; 6, 2)$-RASL with one partial erasure, where the red nodes refer to the local parity checks.

Denote $m \triangleq m_1 + m_2$, and $\text{Roots}(y_i) = \{\beta_{i,1}, \beta_{i,2}, \ldots, \beta_{i,t_{y_i}}\}$ for all $i \in [m]$. Let $\boldsymbol{a} = (a_{1,1}, a_{1,2}, \ldots, a_{1,r_1}, a_{2,1}, \ldots, a_{m_1,r_{m_1}}) \in \mathbb{F}_q^k$ be the information vector. Define $f_{\boldsymbol{a},i}$ as the polynomial with degree less than r_i such that $f_{\boldsymbol{a},i}(\beta_{i,j}) = a_{i,j}$ for all $i \in [m_1]$ and $j \in [r_i]$. Let $F_{\boldsymbol{a}}(x) \in \mathbb{F}_q[x]$ be a polynomial with degree less than $m_1 w$ satisfying

$$F_{\boldsymbol{a}}(x) \equiv f_{\boldsymbol{a},i}(x) \pmod{h(x) - y_i} \quad \text{for all } i \in [m_1]. \quad (1)$$

Then the code we construct is

$$\mathcal{C} = \{C_{\boldsymbol{a}} = (F_{\boldsymbol{a}}(\beta_{1,1}), F_{\boldsymbol{a}}(\beta_{1,2}), \ldots, F_{\boldsymbol{a}}(\beta_{1,t_{y_1}}) \\ \ldots, F_{\boldsymbol{a}}(\beta_{m,t_{y_m}})) \;:\; \boldsymbol{a} \in \mathbb{F}_q^k\}.$$

Remark 5: By Lemma 2, the fact that $\gcd(h(x)-y_i, h(x)-y_j) = 1$ for distinct $i, j \in [m_1]$ implies that $F_{\boldsymbol{a}}(x)$ and $C_{\boldsymbol{a}}$ are well defined.

According to (1), determining the locality of information symbols in $\mathcal{C}$ is straightforward, as discussed in [40, Theorem5.3]. In the following, we present a lemma that is useful for determining the locality of global parity-check symbols.

Lemma 3: Consider the setting of Construction A, and let $0 \leqslant r \leqslant w$ be an integer. Suppose that there exist m_1 distinct constants $y_1, y_2, \ldots, y_{m_1} \in \mathbb{F}_q$ such that

$$F_{\boldsymbol{a}}(x) \pmod{h(x) - y_i} = f_{\boldsymbol{a},i}(x) = \sum_{j=0}^{r-1} e_{i,j} x^j \quad \text{for all } i \in [m_1].$$

Then for any $y \in \mathbb{F}_q$,

$$F_{\boldsymbol{a}}(x) \pmod{h(x) - y} = \sum_{j=0}^{r-1} H_{\boldsymbol{a},j}(y) x^j,$$

where $H_{\boldsymbol{a},j}(x)$ is a polynomial satisfying $\deg(H_{\boldsymbol{a},j}(x)) \leqslant m_1 - 1$ and

$$H_{\boldsymbol{a},j}(y_i) = e_{i,j} \text{ for all } i \in [m_1] \text{ and } 0 \leqslant j \leqslant r-1.$$

Proof: Let us divide $F_{\boldsymbol{a}}(x)$ by $h(x)$ to obtain a quotient $U_1(x)$ and a remainder $R_1(x)$,

$$F_{\boldsymbol{a}}(x) = h(x)U_1(x) + R_1(x),$$

with $\deg(R_1(x)) < w = \deg(h(x))$, $\deg(U_1(x)) = \deg(F_{\boldsymbol{a}}(x)) - w < (m_1 - 1)w$ and $U_1(x), R_1(x) \in \mathbb{F}_q[x]$. We also have

$$F_{\boldsymbol{a}}(x) = (h(x) - y)U_1(x) + yU_1(x) + R_1(x).$$

Similarly, we can write $U_1(x)$ as

$$U_1(x) = h(x)U_2(x) + R_2(x).$$

If $\deg(U_1(x)) \geqslant w$, then $\deg(R_2(x)) < w$ and $\deg(U_2(x)) < (m_1 - 2)w$. However, if $\deg(U_1(x)) < w$, we set $U_2(x) = 0$ and $R_2(x) = U_1(x)$. Then

$$\begin{aligned} &F_{\boldsymbol{a}}(x) \\ &= (h(x) - y)U_1(x) + y(h(x)U_2(x) + R_2(x)) + R_1(x) \\ &= (h(x) - y)(U_1(x) + yU_2(x)) + y^2U_2(x) + yR_2(x) + R_1(x). \end{aligned}$$

Repeating this procedure $m_1 - 1$ times, we conclude that

$$U_{m_1-2}(x) = h(x)U_{m_1-1}(x) + R_{m_1-1}(x),$$

with $\deg(U_{m_1-1}(x)) < w$, $\deg(R_{m_1-1}(x)) < w$, and

$$\begin{aligned} &F_{\boldsymbol{a}}(x) \\ &= (h(x) - y)(U_1(x) + yU_2(x)) + y^2U_2(x) + yR_2(x) + R_1(x) \\ &= (h(x) - y)(U_1(x) + yU_2(x) + \cdots + y^{m_1-2}U_{m_1-1}(x)) \\ &\quad + y^{m_1-1}U_{m_1-1}(x) + y^{m_1-2}R_{m_1-1}(x) \\ &\quad + \cdots + yR_2(x) + R_1(x). \end{aligned}$$

Note that $\deg(U_{m_1-1}(x)) < \deg(h(x)) = w$ and $\deg(R_i(x)) < w$ for all $i \in [m_1 - 1]$. Thus,

$$\begin{aligned} &F_{\boldsymbol{a}}(x) \pmod{h(x) - y} \\ &= y^{m_1-1}U_{m_1-1}(x) + y^{m_1-2}R_{m_1-1}(x) + \cdots + yR_2(x) + R_1(x) \\ &\triangleq H_{\boldsymbol{a}}(x, y). \end{aligned} \quad (2)$$

Let us now rewrite $H_{\boldsymbol{a}}(x, y)$ as

$$\begin{aligned} &H_{\boldsymbol{a}}(x, y) \\ &= H_{\boldsymbol{a},w-1}(y)x^{w-1} + H_{\boldsymbol{a},w-2}(y)x^{w-2} + \cdots + H_{\boldsymbol{a},0}(y)x^0, \end{aligned} \quad (3)$$

where $H_{\boldsymbol{a},j}(y) = \sum_{i=0}^{m_1-1} h_{i,j} y^i$ and $h_{i,j}$ is the coefficient of x^j for $R_{i+1}(x)$ if $0 \leqslant i \leqslant m_1 - 2$, and the coefficient of x^j for

$U_{m_1-1}(x)$ if $i = m_1 - 1$. Thus, $H_{\boldsymbol{a},j}(y)$ for $0 \leqslant j \leqslant w-1$ can be regarded as polynomials in y. According to the assumption, there exist m_1 constants $y_1, y_2, \ldots, y_{m_1}$ such that $H_{\boldsymbol{a}}(x, y_i)$ for $i \in [m_1]$ are polynomials with degree less than r in x. Hence,

$$H_{\boldsymbol{a},j}(y_i) = 0 \qquad \text{for } r \leqslant j \leqslant w-1, \text{ and } i \in [m_1],$$

which means $H_{\boldsymbol{a},j}(y) \equiv 0$ for $r \leqslant j \leqslant w-1$ since $\deg(H_{\boldsymbol{a},j}(y)) < m_1$ for $r \leqslant j \leqslant w-1$, and then

$$F_{\boldsymbol{a}}(x) \pmod{h(x) - y} = \sum_{j=0}^{r-1} H_{\boldsymbol{a},j}(y) x^j.$$

The first claimed result follows directly from (2) and (3). For the second part, since $\deg(H_{\boldsymbol{a},j}(x)) \leqslant m_1 - 1$ for $0 \leqslant j \leqslant r-1$, the rest of the lemma follows from the fact that there exists a unique polynomial $H_{\boldsymbol{a},j}(y)$ for $0 \leqslant j \leqslant r-1$ such that

$$H_{\boldsymbol{a},j}(y_i) = e_{i,j} \qquad \text{for all } i \in [m_1],$$

where $f_{\boldsymbol{a},i}(x) = \sum_{j=0}^{r-1} e_{i,j} x^j = F_{\boldsymbol{a}}(x) \pmod{h(x) - y_i}$. This completes the proof. ■

Corollary 1: In the setting of Construction A, let $\Gamma = \bigcup_{i\in[m]} \mathrm{Roots}(y_i) \subseteq \mathbb{F}_q$. If $|\mathrm{Roots}(y_i)| > r$ for any $i \in [m]$, then the code constructed by Construction A is a locally repairable code with all symbol (r, δ)-locality, where $\delta = \min\{|\mathrm{Roots}(y_i)| + 1 - r \ : \ i \in [m]\}$.

Corollary 1 follows directly from Lemma 3, and we omit its proof. In general, the code may not be optimal for the simple reason that there may not be enough roots in $\mathbb{F}_q$ for some $h(x) - y_i$, where $i \in [m]$, to serve as evaluation points. Therefore, to attain optimality, we consider the case where $\mathbb{F}_q$ contains the splitting field of $h(x) - y_i$, for all $i \in [m]$.

Definition 7: In the setting of Construction A, let $\Gamma = \bigcup_{i\in[m]} \mathrm{Roots}(y_i) \subseteq \mathbb{F}_q$. If $|\mathrm{Roots}(y_i)| = \deg(h(x)) = w$ for all $i \in [m]$, then the polynomial $h(x)$ is said to be a *good polynomial* over Γ.

For more details about good polynomials the readers may refer to [30] and [40].

Corollary 2 (Tamo-Barg codes, [40]): Consider the setting of Construction A and let $h(x)$ be a good polynomial over $\Gamma = \bigcup_{i\in[m]} \mathrm{Roots}(y_i) \subseteq \mathbb{F}_q$. If $r_i = r < w$ and $k = rm_1$, then the resulting code is an optimal $[n = mw, m_1 r, (m_2+1)w - r + 1]_q$ locally repairable code with all symbol $(r, w-r+1)$-locality (where optimality is with respect to the bound in Lemma 1).

Corollary 2 follows directly from Lemma 3 and Corollary 1, which also can be directly derived from [40, Construction 1]. We omit its proof. This result was first introduced in [40].

V. Repairing Tamo-Barg Codes: Rack Erasures

In this section, we consider the repair problem for Tamo-Barg codes for the rack erasure case, where the repair sets are arranged as racks. We begin with an array form of the Tamo-Barg code, where each repair set is arranged as a column in the array.

Construction B: Let $h(x) \in \mathbb{F}_q[x]$ be a good polynomial over $\Gamma = \bigcup_{i\in[m]} \mathrm{Roots}(y_i) \subseteq \mathbb{F}_q$ with $\deg(h(x)) = r + \delta - 1$, and let $\boldsymbol{a} = (a_{1,1}, a_{1,2}, \ldots, a_{1,r}, a_{2,1}, \ldots a_{m_1,r}) \in \mathbb{F}_q^k$ be the information vector, where $k = rm_1$. Define $f_{\boldsymbol{a},i}(x)$ as the polynomial with degree less than r such that $f_{\boldsymbol{a},i}(\beta_{i,j}) = a_{i,j}$ for all $i \in [m_1]$ and $j \in [r]$, where we assume that $\mathrm{Roots}(y_i) = \{\beta_{i,1}, \beta_{i,2}, \ldots, \beta_{i,r+\delta-1}\}$ for all $i \in [m_1]$. Then for any $\boldsymbol{a} \in \mathbb{F}_q^k$ we can find a polynomial $F_{\boldsymbol{a}}(x) \in \mathbb{F}_q[x]$ with degree less than $m_1(r + \delta - 1)$ satisfying

$$F_{\boldsymbol{a}}(x) \equiv f_{\boldsymbol{a},i}(x) \pmod{h(x) - y_i} \qquad \text{for all } i \in [m_1].$$

Construct an array code as follows $\mathcal{A} \triangleq$, shown at the bottom of the next page, where we define $m \triangleq m_1 + m_2$.

Now, we are going to present repair schemes for the array-form Tamo-Barg codes, whose main idea is explained in Fig. 3.

Theorem 2: Consider the setting of Construction B. Let q_0 be a prime power, and $\mathbb{F}_q = \mathbb{F}_{q_0}(y_1, y_2, \ldots, y_m)$, $\mathbb{F}_{q_i} = \mathbb{F}_{q_0}(y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_m)$. Define $w_i \triangleq [\mathbb{F}_q : \mathbb{F}_{q_i}]$ for each $i \in [m]$.

I. Define $w_i^* \triangleq \frac{r}{\gcd(w_i, r)}$. If $w_i^* \leqslant w_i$ and $m_2 \geqslant \max\{2, w_i^*\}$, then we can recover $\boldsymbol{A}_i$ by downloading $(w_i^* + m_1 - 1)r$ elements of $\mathbb{F}_{q_i}$, i.e., $\frac{(w_i^* + m_1 - 1)r}{w_i^*}$ symbols in $\mathbb{F}_q$ from any other $w_i^* + m_1 - 1$ racks (columns), where each symbol is an element of $\mathbb{F}_q$.

II. Assume w_i^* is a positive integer. If $w_i^* | w_i$ and $m_2 \geqslant \max\{2, w_i^*\}$, then we can recover $\boldsymbol{A}_i$ by downloading $(w_i^* + m_1 - 1)r$ elements of $\mathbb{F}_{q_i}$, i.e., $\frac{(w_i^* + m_1 - 1)r}{w_i^*}$ symbols in $\mathbb{F}_q$ from any other $w_i^* + m_1 - 1$ racks (columns), where each symbol denotes an element of $\mathbb{F}_q$.

Proof: We begin by proving claim I. To recover

$$\begin{aligned} \boldsymbol{A}_i &= (F_{\boldsymbol{a}}(\beta_{i,1}), F_{\boldsymbol{a}}(\beta_{i,2}), \ldots, F_{\boldsymbol{a}}(\beta_{i,r+\delta-1}))^\top \\ &= (f_{\boldsymbol{a},i}(\beta_{i,1}), f_{\boldsymbol{a},i}(\beta_{i,2}), \ldots, f_{\boldsymbol{a},i}(\beta_{i,r+\delta-1}))^\top, \end{aligned}$$

it suffices that we recover $f_{\boldsymbol{a},i}(x) \equiv F_{\boldsymbol{a}}(x) \pmod{h(x) - y_i}$. By Lemma 3, we only need to figure out $H_{\boldsymbol{a},j}(y_i)$ for $0 \leqslant j \leqslant r-1$. Note that for $t \in [m] \setminus \{i\}$ we have

$$\begin{aligned} \boldsymbol{A}_t &= (F_{\boldsymbol{a}}(\beta_{t,1}), F_{\boldsymbol{a}}(\beta_{t,2}), \ldots, F_{\boldsymbol{a}}(\beta_{t,r+\delta-1}))^\top \\ &= (f_{\boldsymbol{a},t}(\beta_{t,1}), f_{\boldsymbol{a},t}(\beta_{t,2}), \ldots, f_{\boldsymbol{a},t}(\beta_{t,r+\delta-1}))^\top \end{aligned}$$

and $f_{\boldsymbol{a},t}(x) = \sum_{j=0}^{r-1} H_{\boldsymbol{a},j}(y_t) x^j$. Thus, based on $\boldsymbol{A}_t$, we can calculate $H_{\boldsymbol{a},j}(y_t)$ for $t \in [m] \setminus \{i\}$ and $0 \leqslant j \leqslant r-1$.

We observe that $w_i^* | r$, which implies that we can divide $H_{\boldsymbol{a},j}(y_t)$ for $0 \leqslant j \leqslant r-1$ into r/w_i^* vectors of length w_i^*, say

$$\begin{aligned} &\boldsymbol{H}_{\boldsymbol{a}}(y_t, \tau) \\ &\triangleq \left(H_{\boldsymbol{a},\tau w_i^*}(y_t), H_{\boldsymbol{a},\tau w_i^*+1}(y_t), \ldots, H_{\boldsymbol{a},(\tau+1)w_i^*-1}(y_t)\right), \end{aligned}$$

for $0 \leqslant \tau \leqslant r/w_i^* - 1$.Let $\mathbb{F} = \mathbb{F}_q(\beta)$ such that $[\mathbb{F} : \mathbb{F}_q] = w_i^*$. Then we have $[\mathbb{F} : \mathbb{F}_{q_i}] = w_i^* w_i$ and $\{\beta^t y_i^j \ : \ 0 \leqslant t \leqslant w_i^* - 1, 0 \leqslant j \leqslant w_i - 1\}$ is a basis of $\mathbb{F}$ over $\mathbb{F}_{q_i}$. Let Ψ be the bijection from $\mathbb{F}_q^{w_i^*}$ to $\mathbb{F}$

$$\Psi(V) = \sum_{j=0}^{w_i^*-1} v_j \beta^j,$$

where $V = (v_0, v_1, \ldots, v_{w_i^*-1}) \in \mathbb{F}_q^{w_i^*}$. Now, for any $0 \leqslant \tau \leqslant r/w_i^* - 1$, we may regard

$$(\Psi(\boldsymbol{H}_{\boldsymbol{a}}(y_1, \tau)), \Psi(\boldsymbol{H}_{\boldsymbol{a}}(y_2, \tau)), \ldots, \Psi(\boldsymbol{H}_{\boldsymbol{a}}(y_{m_1+m_2}, \tau)))$$

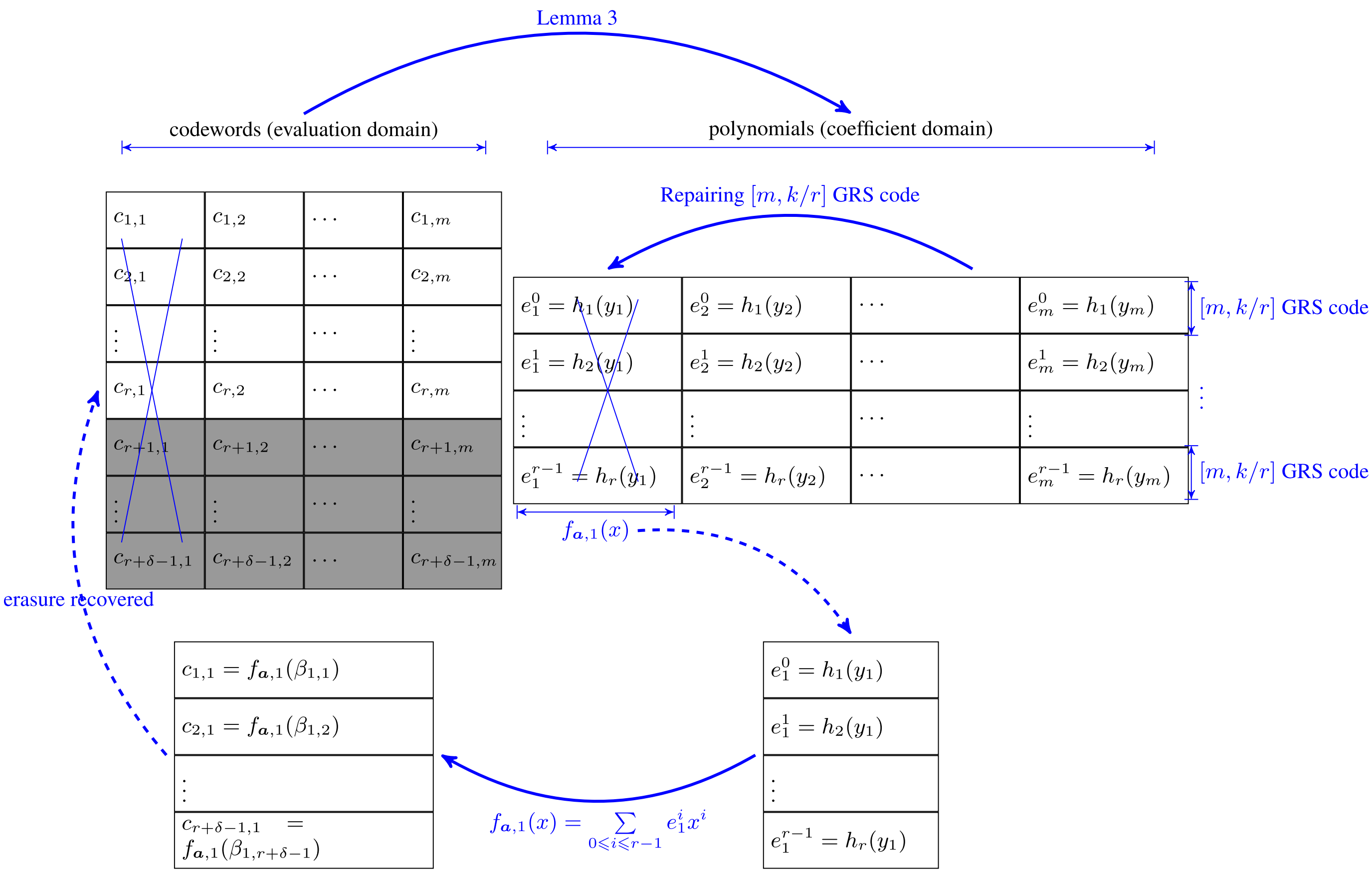

Fig. 3. Repairing Tamo-Barg codes. The repair problem of the first column is reduced to the repair problem for the corresponding polynomial remainder. Secondly, by Lemma 3, this is equivalent to the repair problem of certain codewords within a Reed-Solomon code, where each of them suffers one erasure. Finally, the repairing of these codewords ensures the recovery of the erased column.

as a codeword of a GRS code $\mathcal{C}_1$ with parameters $[m_1 + m_2, m_1, m_2 + 1]_{q^{w_i^*}}$, since it corresponds to evaluations of a polynomial

$$H_{\boldsymbol{a}}^{(\tau)}(x) = \sum_{j=0}^{w_i^*-1} \beta^j H_{\boldsymbol{a},\tau w^*+j}(x) \in \mathbb{F}[x],$$

with $\deg(H_{\boldsymbol{a}}^{(\tau)}(x)) < m_1$. Now the repair problem of $H_{\boldsymbol{a},j}(y_i)$ (i.e., $\boldsymbol{H}_{\boldsymbol{a}}(y_i, \tau)$) is exactly the repair problem of a GRS code. For completeness, we include some known methods for repairing GRS codes [14], [43], [44].

Let $\Theta = \{y_{j_t} : 1 \leqslant t \leqslant w_i^*+m_1-1\}$ be any $(w_i^*+m_1-1)$-subset of $\Gamma_i = \{y_j : j \in [m]\} \setminus \{y_i\}$. Define

$$g_\Theta(x) \triangleq \prod_{\theta\in\Gamma_i\setminus\Theta} (x - \theta). \tag{4}$$

By (4), we have $\deg(x^u g_\Theta(x)) < m_2$ for $0 \leqslant u \leqslant w_i^* - 1$. Note that the dual code of $\mathcal{C}_1$ is also a GRS code, with parameters $[m_1 + m_2, m_2, m_1 + 1]_{q^{w_i^*}}$. Thus, there exists a vector $(v_1, v_2, \ldots, v_{m_1+m_2}) \in (\mathbb{F}^*)^{m_1+m_2}$ such that

$$(v_1 y_1^u g_\Theta(y_1), v_2 y_2^u g_\Theta(y_2), \ldots, v_{m_1+m_2} y_{m_1+m_2}^u g_\Theta(y_{m_1+m_2})) \in \mathcal{C}_1^\perp \text{ for all } 0 \leqslant u \leqslant w_i^* - 1,$$

i.e.,

$$\sum_{j=1}^{m_1+m_2} v_j y_j^u g_\Theta(y_j)\Psi(\boldsymbol{H}_{\boldsymbol{a}}(y_j, \tau)) = 0 \quad \text{for all } 0 \leqslant u \leqslant w_i^* - 1 \text{ and } 0 \leqslant \tau \leqslant \frac{r}{w_i^*} - 1. \tag{5}$$

For $\{y_i^u : 0 \leqslant u \leqslant w_i^* - 1\}$, from Lemma 1 in [44], we can find a set $\{\gamma_t : 0 \leqslant t \leqslant w_i - 1\}$ such that $\{y_i^u\gamma_t : 0 \leqslant u \leqslant w_i^* - 1, 0 \leqslant t \leqslant w_i - 1\}$, just as $\{\beta^t y_i^j : 0 \leqslant t \leqslant w_i^* - 1, 0 \leqslant j \leqslant w_i - 1\}$, is a basis of $\mathbb{F}$ over $\mathbb{F}_{q_i}$. By (5), for $0 \leqslant u \leqslant w_i^* - 1$, $0 \leqslant \tau \leqslant r/w_i^* - 1$, and $0 \leqslant t \leqslant w_i - 1$,

$$\mathrm{Tr}_{\mathbb{F}/\mathbb{F}_{q_i}} (\gamma_t v_i y_i^u g_\Theta(y_i)\Psi(\boldsymbol{H}_{\boldsymbol{a}}(y_i, \tau)))$$

$$\mathcal{A} \triangleq \left\{ A_{\boldsymbol{a}} = (\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_m) = \begin{pmatrix} F_{\boldsymbol{a}}(\beta_{1,1}) & F_{\boldsymbol{a}}(\beta_{2,1}) & \ldots & F_{\boldsymbol{a}}(\beta_{m,1}) \\ F_{\boldsymbol{a}}(\beta_{1,2}) & F_{\boldsymbol{a}}(\beta_{2,2}) & \ldots & F_{\boldsymbol{a}}(\beta_{m,2}) \\ \vdots & \vdots & & \vdots \\ F_{\boldsymbol{a}}(\beta_{1,r+\delta-1}) & F_{\boldsymbol{a}}(\beta_{2,r+\delta-1}) & \ldots & F_{\boldsymbol{a}}(\beta_{m,r+\delta-1}) \end{pmatrix} : \boldsymbol{a} \in \mathbb{F}_q^k \right\}$$

$$=-\sum_{\substack{j\in[m_1+m_2]\\ j\neq i}} \mathrm{Tr}_{\mathbb{F}/\mathbb{F}_{q_i}}\left(\gamma_t v_j y_j^u g_\Theta(y_j)\Psi(\boldsymbol{H}_{\boldsymbol{a}}(y_j,\tau))\right)$$
$$=-\sum_{\substack{j\in[m_1+m_2]\\ j\neq i}} y_j^u g_\Theta(y_j)\,\mathrm{Tr}_{\mathbb{F}/\mathbb{F}_{q_i}}\left(\gamma_t v_j \Psi(\boldsymbol{H}_{\boldsymbol{a}}(y_j,\tau))\right)$$
$$=-\sum_{\substack{j\in[m_1+m_2]\\ j\neq i,\, y_j\in\Theta}} y_j^u g_\Theta(y_j)\,\mathrm{Tr}_{\mathbb{F}/\mathbb{F}_{q_i}}\left(\gamma_t v_j \Psi(\boldsymbol{H}_{\boldsymbol{a}}(y_j,\tau))\right),$$

where the last equality holds by (4). As already mentioned, $\{y_i^u\gamma_t\ :\ 0\leqslant u\leqslant w_i^*-1,\, 0\leqslant t\leqslant w_i-1\}$ is a basis of $\mathbb{F}$ over $\mathbb{F}_{q_i}$, so $\{\mathrm{Tr}_{\mathbb{F}/\mathbb{F}_{q_i}}(\gamma_t v_i y_i^u g_\Theta(y_i)\Psi(\boldsymbol{H}_{\boldsymbol{a}}(y_i,\tau)))\ :\ 0\leqslant u\leqslant w_i^*-1, 0\leqslant t\leqslant w_i-1\}$ is a set of $w_i^* w_i$ independent traces, which can uniquely determine $\Psi(\boldsymbol{H}_{\boldsymbol{a}}(y_i,\tau))$ by means of its dual basis. Thus, to recover $\boldsymbol{H}_{\boldsymbol{a}}(y_i,\tau)$, i.e., $\Psi(\boldsymbol{H}_{\boldsymbol{a}}(y_i,\tau))$ (since $\Psi(\cdot)$ is a bijection), we only need the code symbols with evaluation points y_j such that $y_j\in\Theta\setminus\{y_i\}$ to transmit $\mathrm{Tr}_{\mathbb{F}/\mathbb{F}_{q_i}}(\gamma_t v_j\Psi(\boldsymbol{H}_{\boldsymbol{a}}(y_j,\tau)))$ for $0\leqslant\tau\leqslant r/w_i^*-1$ and $0\leqslant t\leqslant w_i-1$. Therefore, we can recover $\boldsymbol{A}_i$ by downloading $\frac{(w_i^*+m_1-1)r}{w_i^*}$ symbols of $\mathbb{F}_q$ from any other $w_i^*+m_1-1$ columns.

To prove claim II, we consider each row individually, i.e., the recovery problem for $H_{\boldsymbol{a},\tau}(y_i)$ for $0\leqslant\tau\leqslant r-1$. By a similar analysis, $H_{\boldsymbol{a},\tau}(y_i)$ for $0\leqslant\tau\leqslant r-1$ can be determined by the following equations:

$$\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_{q_i}}\left(y_i^{w_i^*t}v_i y_i^u g_\Theta(y_i)H_{\boldsymbol{a},\tau}(y_i))\right)$$
$$=-\sum_{\substack{j\in[m_1+m_2]\\ j\neq i}} \mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_{q_i}}\left(y_i^{w_i^*t}v_j y_j^u g_\Theta(y_j)H_{\boldsymbol{a},\tau}(y_j))\right)$$
$$=-\sum_{\substack{j\in[m_1+m_2]\\ j\neq i}} y_j^u g_\Theta(y_j)\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_{q_i}}\left(y_i^{w_i^*t}v_j H_{\boldsymbol{a},\tau}(y_j))\right)$$
$$=-\sum_{\substack{j\in[m_1+m_2]\\ j\neq i,\, y_j\in\Theta}} y_j^u g_\Theta(y_j)\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_{q_i}}\left(y_i^{w_i^*t}v_j H_{\boldsymbol{a},\tau}(y_j))\right),$$

where $0\leqslant u\leqslant w_i^*-1$, $0\leqslant t\leqslant w_i/w_i^*-1$, and $g_\Theta(\cdot)$ is also defined by (4). Thus, the fact that $\{y_i^t\ :\ 0\leqslant t\leqslant w_i-1\}$ is a basis of $\mathbb{F}_q$ over $\mathbb{F}_i$ means that we can recover $H_{\boldsymbol{a},\tau}(y_i)$ for $0\leqslant\tau\leqslant r-1$ by downloading $\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_{q_i}}(y_i^{w_i^*t}v_j H_{\boldsymbol{a},\tau}(y_j)))$ for $y_j\in\Theta$ and $0\leqslant t\leqslant w_i/w_i^*-1$. Namely, we can recover $\boldsymbol{A}_i$ by downloading $\frac{(w_i^*+m_1-1)r}{w_i^*}$ symbols of $\mathbb{F}_q$ from any other $w_i^*+m_1-1$ columns. ■

To analyze the repair bandwidth performance for the scheme introduced in Theorem 2, we need to modify the cut-set bound to the case that each code symbol in the array code can be locally repaired.

Theorem 3: Let $\mathcal{C}$ be an $(N,K,k=Kr;L=r+\delta-1,\delta)$-RASL code with $(r,L-r+1)_a$-locality in which each column corresponds to a repair set, where $0<r\leqslant L-1$. Let D be an integer with $K\leqslant D\leqslant N-1$. For any $i\in[N]$ and any D-subset $\mathcal{R}\subseteq[N]\setminus\{i\}$, we have

$$B(\mathcal{C},\{i\},\mathcal{R})\geqslant\frac{Dr}{D-K+1}.$$

Proof: The fact that each column corresponds to a repair set of $(r,L-r+1)$-locality means that the punctured code over each column has distance at least $L-r+1$. That is, any r symbols in the i-th column are capable of recovering the entire i-th column. Consider the array code $\mathcal{C}'$ formed by deleting an arbitrary set of $L-r$ rows from the array code $\mathcal{C}$, say the last $L-r$ rows. Since $\mathcal{C}$ is anan$(N,K,k=Kr;L=r+\delta-1,\delta)$-RASL code in which each column corresponds to a repair set, it is easy to check that $\mathcal{C}'$ is also an $[N,K,N-K+1]_q$ MDS array code, but with sub-packetization r. Even though each column of $\mathcal{C}'$ is only a part of the corresponding column in $\mathcal{C}$, due to the locality, each column may compute its original column from $\mathcal{C}$. Thus, any repair procedure on $\mathcal{C}$ may also be run on $\mathcal{C}'$, and therefore we have $B(\mathcal{C},\{i\},\mathcal{R})\geqslant B(\mathcal{C}',\{i\},\mathcal{R})$ for any $i\in[N]$ and any D-subset $\mathcal{R}\subseteq[N]\backslash\{i\}$. Now the desired result follows from Theorem 1, that is, for any $i\in[N]$ and any D-subset $\mathcal{R}\subseteq[N]\setminus\{i\}$, we have

$$B(\mathcal{C},\{i\},\mathcal{R})\geqslant B(\mathcal{C}',\{i\},\mathcal{R})\geqslant\frac{Dr}{D-K+1}.$$

■

Remark 6: In the proof of Theorem 3, we used $B(\mathcal{C},\{i\},\mathcal{R})\geqslant B(\mathcal{C}',\{i\},\mathcal{R})$. In fact, the reverse inequality, $B(\mathcal{C},i,\mathcal{R})\leqslant B(\mathcal{C}',i,\mathcal{R})$ is trivially true since each column of $\mathcal{C}'$ is a part of the original column from $\mathcal{C}$, and therefore any repair procedure running on $\mathcal{C}'$ may be run on $\mathcal{C}$.

Based on Theorems 2 and 3, we have the following corollary:

Corollary 3: Let q_0 be a prime power, and $\mathbb{F}_q=\mathbb{F}_{q_0}(y_1,y_2,\ldots,y_m)$, $\mathbb{F}_{q_i}=\mathbb{F}_{q_0}(y_1,\ldots,y_{i-1},y_{i+1},\ldots,y_m)$. Define $w_i\triangleq[\mathbb{F}_q:\mathbb{F}_{q_i}]$ for all $i\in[m]$. Furthermore, let $m_2\geqslant\max\{2,r\}$, where $m=m_1+m_2$ and $k=m_1r$. If $0<r\leqslant\min\{w_1,w_2,\ldots,w_m\}$, and $\gcd(r,w_i)=1$ for all $i\in[m]$, then for any $i\in[m]$ we can recover $\boldsymbol{A}_i$ by downloading $r+m_1-1$ symbols in $\mathbb{F}_q$ from any other $r+m_1-1$ racks (columns), which is exactly the optimal bandwidth with respect to the cut-set bound of Theorem 3.

Proof: According to Theorem 2-I, the facts that $\gcd(w_i,r)=1$ for $i\in[m]$ and $r\leqslant\min\{w_1,w_2,\ldots,w_m\}$ mean that $w_i^*=r$ and we can recover $\boldsymbol{A}_i$ by downloading $r+m_1-1$ symbols from any other $r+m_1-1$ columns. Note that $\mathcal{C}$ is an $(m,m_1,k=m_1r;L=r+\delta-1,\delta)$-RASL code in which each column corresponds to a repair set, where $0<r\leqslant L-1$. By Theorem 3,

$$B(\mathcal{C},\{i\},\mathcal{R})\geqslant\frac{(r+m_1-1)r}{r},$$

for any $(r+m_1-1)$-subset of $[n]\setminus\{i\}$. Therefore, in this case, the code $\mathcal{C}$ has optimal repair bandwidth. ■

Remark 7: For multiple rack erasures, the method used in Theorem 2 may not achieve the optimal bandwidth by Theorem 3. The key aspect may be to design a repair scheme that allows some downloaded data to be repeatedly used when repairing multiple rack erasures, similar to the method used in [44]. Finding optimal schemes for repairing multiple rack-erasures is still an open problem.

VI. Repairing Tamo-Barg Codes: Partial Erasures

In the previous section, we demonstrated that Tamo-Barg codes may exhibit optimal repair properties when represented in an array form. In this section, we further explore the partial repair problem for the rack-aware system with locality, i.e., array codes under the assumption that each column of the array is an (r,δ)-repair set. Specifically, we consider the scenario where some repair sets have failed, meaning that certain columns contain more than $\delta-1$ erasures. We seek to determine the minimum amount of data that needs to be downloaded from D remaining columns, and how to construct a code that achieves the minimum repair bandwidth for this model.

To begin, we provide some necessary definitions for the partial erasure case in Section III.

Definition 8: Let $\mathcal{C}$ be an $(N,K,k=Kr;L=r+\delta-1,\delta)$-RASL code, where $0<r\leqslant L-1$. Let $\mathcal{I}=\{i_1,i_2,\ldots,i_\tau\}\subseteq[N]$ denote the failed columns, and let $E_{i_t}\subseteq[L]$ with $|E_{i_t}|\geqslant\delta$ for $t\in[\tau]$ denote the corresponding erasures in the i_t-th column. For a D-subset $\mathcal{R}\subseteq[N]\setminus\mathcal{I}$, define $B(\mathcal{C},\mathcal{I},\mathcal{E},\mathcal{R})$ as the *minimum repair bandwidth* for $\{c_{i,j}\ :\ i\in\mathcal{I},j\in E_i\}$, i.e., the smallest number of symbols of $\mathbb{F}_q$ helper racks (columns) need to send in order to recover the erasure pattern $\mathcal{E}=\{E_{i_t}\ :\ t\in[\tau]\}$ (where each helper rack (column) $j\in\mathcal{R}$ may send symbols that depend solely on $c_j\in\mathbb{F}_q^L$).

First, we consider the case where a single column contains erasures, but unlike Theorem 2, the column is not fully erased. Namely, we consider the case where $\mathcal{E}=\{E_i\}$, and $E_i=E$, i.e., the partial erasure case for $\lambda=1$ in Section III. The main idea of the repair process is explained in Figure 4.

Theorem 4: Consider the setting of Construction B. Let q_0 be a prime power, and $\mathbb{F}_q=\mathbb{F}_{q_0}(y_1,y_2,\ldots,y_m)$, $\mathbb{F}_{q_i}=\mathbb{F}_{q_0}(y_1,\ldots,y_{i-1},y_{i+1},\ldots,y_m)$. Define $w_i\triangleq[\mathbb{F}_q:\mathbb{F}_{q_i}]$ for each $i\in[m]$. For any given set $E\subseteq[L]$ with $\delta\leqslant|E|\leqslant L$, consider the case that the elements in $\boldsymbol{A}_i|_E$ are erasures.

I. Define $w_i^*\triangleq\frac{L-\delta+1}{\gcd(w_i,L-\delta+1)}$. If $w_i^*\leqslant w_i$ and $m_2\geqslant\max\{2,w_i^*\}$, then we can recover $\boldsymbol{A}_i|_E$ by downloading $(w_i^*+m_1-1)(|E|-\delta+1)$ elements of $\mathbb{F}_{q_i}$, i.e., $\frac{(w_i^*+m_1-1)(|E|-\delta+1)}{w_i^*}$ symbols in $\mathbb{F}_q$ from any other $w_i^*+m_1-1$ racks (columns).
II. Assume w_i^* is a positive integer. If $w_i^*|w_i$ and $m_2\geqslant\max\{2,w_i^*\}$, then we can recover $\boldsymbol{A}_i|_E$ by downloading $(w_i^*+m_1-1)(|E|-\delta+1)$ elements of $\mathbb{F}_{q_i}$, i.e., $\frac{(w_i^*+m_1-1)(|E|-\delta+1)}{w_i^*}$ symbols in $\mathbb{F}_q$ from any other $w_i^*+m_1-1$ racks (columns).

Proof: By Construction B, for recovering of $\boldsymbol{A}_i|_E$, it is sufficient to recover $f_{\boldsymbol{a},i}(x)\equiv F_{\boldsymbol{a}}(x)\pmod{h(x)-y_i}$. By Lemma 3, for $i\in[m]$,

$$F_{\boldsymbol{a}}(x)\pmod{h(x)-y_i}=f_{\boldsymbol{a},i}(x)=\sum_{j=0}^{r-1}H_{\boldsymbol{a},j}(y_i)x^j,$$

where $\deg(H_{\boldsymbol{a},j}(x))\leqslant m_1-1$. Recall that in this case we have $\delta=L-r+1$, and we know $r+\delta-1-|E|$ elements of $\boldsymbol{A}_i$, i.e., those elements in $\boldsymbol{A}_i|_{[r+\delta-1]\setminus E}$. Thus, one sufficient condition for repairing $f_{\boldsymbol{a},i}(x)$ is to recover $H_{\boldsymbol{a},j}(y_i)$ for $0\leqslant j\leqslant|E|-\delta$. Note that for these j's, $H_{\boldsymbol{a},j}(y_t)$ for $t\in[m]\setminus\{i\}$ can be calculated by the elements in $\boldsymbol{A}_t$.

Define the following array $\mathcal{A}_1$ from the Tamo-Barg code:

$$\mathcal{A}_1=\Big\{A_{\boldsymbol{b}}^*:A_{\boldsymbol{b}}^*=(\boldsymbol{A}_1^*,\boldsymbol{A}_2^*,\ldots,\boldsymbol{A}_m^*)=(F_{\boldsymbol{b}}(\beta_{u,v}))_{(r+\delta-1)\times m},\ \text{for, any }\boldsymbol{b}\in\mathbb{F}_q^{m_1(|E|-\delta+1)}\Big\},$$

where

$$F_{\boldsymbol{b}}(x)\pmod{h(x)-y_i}=f_{\boldsymbol{b},i}^*(x)=\sum_{t=0}^{|E|-\delta}b_{i,t}x^t$$

and $\boldsymbol{b}=(b_{1,0},b_{1,1},\ldots,b_{1,|E|-\delta},b_{2,0},\ldots,b_{m_1,|E|-\delta})$. For any $\boldsymbol{a}\in\mathbb{F}_q^{rm_1}$, let $V(\boldsymbol{a})=(v_{1,0},v_{1,1},\ldots,v_{1,|E|-\delta},v_{2,0},\ldots,v_{m_1,|E|-\delta})$ with $v_{i,j}=H_{\boldsymbol{a},j}(y_i)$ for $i\in[m_1]$ and $0\leqslant j\leqslant|E|-\delta$, then for any $m_1+1\leqslant i\leqslant m$ we have

$$F_{V(\boldsymbol{a})}(x)\equiv\sum_{0\leqslant j\leqslant|E|-\delta}v_{i,j}x^j\pmod{h(x)-y_i},$$

where

$$v_{i,j}=H_{\boldsymbol{a},j}(y_i)\quad\text{for }m_1+1\leqslant i\leqslant m,\ 0\leqslant j\leqslant|E|-\delta,$$

according to Lemma 3. Now consider the repair problem of the i-th column $\boldsymbol{A}_i^*$ of the codeword $A_{V(\boldsymbol{a})}^*\in\mathcal{A}_1$. Since the repair problem is equivalent to recovering the coefficients of $f_{V(\boldsymbol{a}),i}^*(x)$, i.e., $v_{i,j}=H_{\boldsymbol{a},j}(y_i)$ for $0\leqslant j\leqslant|E|-\delta$, it is also a sufficient condition to recover $A_i|_E$. Thus, the bandwidth of repairing $\boldsymbol{A}_i|_E$ is upper bounded by the bandwidth of repairing $\boldsymbol{A}_i^*$ for $\boldsymbol{A}_{V(\boldsymbol{a})}^*$. Now the desired results follows from Theorem 2. ■

We now move on to address the challenge of repairing erasure patterns in scenarios where multiple repair sets (columns) have failed, specifically when $|E_{i_t}|\geqslant\delta$ for $t\in[\tau]$, i.e., the partial erasure case with $\lambda=\tau>1$.

Consider a given erasure pattern $\mathcal{E}=\{E_{i_1},E_{i_2},\ldots,E_{i_\tau}\}$ with $|E_{i_t}|\geqslant\delta$ for all $t\in[\tau]$. Define

$$\begin{aligned}\Lambda_{\mathcal{E}}(x)&\triangleq\prod_{E_j\in\mathcal{E}}(x-y_j)=\prod_{t\in[\tau]}(x-y_{i_t}),\\ \Omega_{t,\mathcal{E},\alpha}(x)&\triangleq\frac{\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_{q_1}}(\alpha\Lambda_{\mathcal{E}}(x))}{(x-y_{i_t})\prod_{E_{i_j}\in\mathcal{E},\,j\neq t}(y_{i_t}-y_{i_j})}.\end{aligned}\tag{6}$$

Theorem 5: Let $\mathcal{C}$ be the code generated by Construction B, and $\mathcal{E}=\{E_{i_1},E_{i_2},\ldots,E_{i_\tau}\}$ be an erasure pattern with $|E_{i_t}|\geqslant\delta$ for all $t\in[\tau]$. For a subfield $\mathbb{F}_{q_1}\subset\mathbb{F}_q$, if $m_2\geqslant\frac{q\tau}{q_1}$ then the erasure pattern can be recovered by downloading M elements of $\mathbb{F}_{q_1}$, i.e., with repair bandwidth $\frac{M}{\ell}$ by contacting all the remaining racks (columns), where $M=\sum_{t\in[\tau]}|E_{i_t}|-\tau(\delta-1)$ and $q=q_1^\ell$.

Proof: Without loss of generality, assume $i_t=t$, namely, we consider the case $\mathcal{E}=\{E_1,E_2,\ldots,E_\tau\}$, $|E_i|\geqslant\delta$ for all $i\in[\tau]$. Let $w=\max\{|E_i|\ :\ i\in[\tau]\}$ and for any $1\leqslant\ell\leqslant w-\delta+1$ define

$$R_\ell\triangleq\{i\in[\tau]\ :\ |E_i|\geqslant\delta-1+\ell\}.$$

By Construction B, for recovering of $\boldsymbol{A}_i|_{E_i}$, $E_i\in\mathcal{E}$, one possible method is to recover $f_{\boldsymbol{a},i}(x)=F_{\boldsymbol{a}}(x)\pmod{h(x)-y_i}$. By Lemma 3, for $i\in[m]$,

$$F_{\boldsymbol{a}}(x)\pmod{h(x)-y_i}=f_{\boldsymbol{a},i}(x)=\sum_{j=0}^{r-1}H_{\boldsymbol{a},j}(y_i)x^j,$$

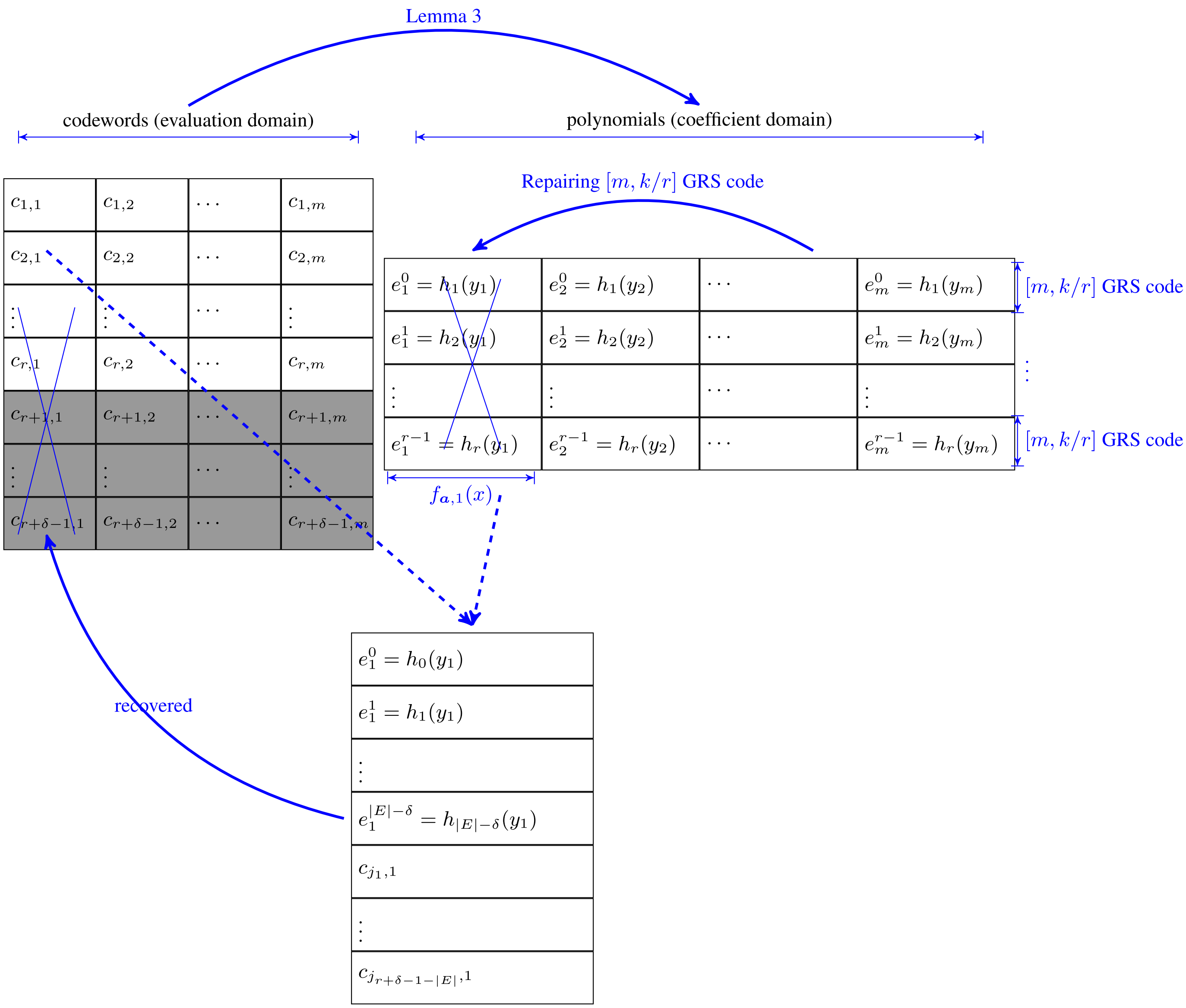

Fig. 4. Partial repairing of Tamo-Barg codes. Herein, we initially reduce the repair problem of failed items in the first column to the repair problem for the coefficients of the corresponding polynomial remainder. Secondly, according to Lemma 3, these coefficients can be recovered by repairing certain codewords within a Reed-Solomon code. Finally, repairing these symbols of the codewords ensures the repair of the desired components.

where $\deg(H_{\boldsymbol{a},j}(x)) \leqslant m_1 - 1$. Note that $(H_{\boldsymbol{a},j}(y_1), H_{\boldsymbol{a},j}(y_2), \cdots, H_{\boldsymbol{a},j}(y_m))$ can be regarded as a codeword of an $[m = m_1 + m_2, m_1, m_2 + 1]_q$ GRS code, i.e., $(H_{\boldsymbol{a},j}(y_1), H_{\boldsymbol{a},j}(y_2), \cdots, H_{\boldsymbol{a},j}(y_m)) \in \mathrm{GRS}_{m_1}(\mathbf{1}, \boldsymbol{Y})$, where $\boldsymbol{Y} = (y_1, y_2, \ldots, y_m)$.

First, we are going to apply the trace function and $\Omega_{i,\mathcal{E},\alpha}(x)$ defined by (6) to recover $H_{\boldsymbol{a},j}(y_i)$ for all $j \in [w-\delta+1]$ and $i \in R_\ell$. For $i \in [m] \setminus R_1$, $\boldsymbol{A}_i$ suffers from at most $\delta - 1$ erasures, i.e., at least r components are accessible. Hence, rack (column) i, $i \in [m] \setminus R_1$, can recover $f_{\boldsymbol{a},i}(x) = \sum_{j=0}^{r-1} H_{\boldsymbol{a},j}(y_i)x^j$ since $\deg(f_{\boldsymbol{a},i}(x)) < r$. For $j \in [w-\delta+1]$ and $i \in R_j$, the rack (column) i downloads

$$\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_{q_1}}\left(\frac{-H_{\boldsymbol{a},j}(y_t)}{y_t - y_i}\right)$$

from the racks (columns) t, for $t \in [m] \setminus R_1$. Since the dual of a GRS code is a GRS code with the same code locators, there exists a vector $\boldsymbol{\Theta} = (\theta_1, \theta_2, \ldots, \theta_m) \in \mathbb{F}_q^m$ such that $\mathrm{GRS}_{m_1}^{\perp}(\mathbf{1}, \boldsymbol{Y}) = \mathrm{GRS}_{m_2}(\boldsymbol{\Theta}, \boldsymbol{Y})$. Since $m_2 \geqslant \frac{q\tau}{q_1}$ and $\deg(\Omega_{i,\mathcal{E},\alpha}(x)) \leqslant \frac{q\tau}{q_1} - 1$ we have that

$$\begin{aligned}&(\theta_1\Omega_{i,\mathcal{E},\alpha}(y_1), \theta_2\Omega_{i,\mathcal{E},\alpha}(y_2), \ldots, \theta_m\Omega_{i,\mathcal{E},\alpha}(y_m))\\ &\quad \in \mathrm{GRS}_{m_2}(\boldsymbol{\Theta}, \boldsymbol{Y}) = \mathrm{GRS}_{m_1}^{\perp}(\mathbf{1}, \boldsymbol{Y}).\end{aligned}$$

Thus, for $\ell \in [w-\delta+1]$ and $i \in R_\ell$, we have

$$0 = \sum_{t\in[m]} \theta_t \Omega_{i,\mathcal{E},\alpha}(y_t) H_{\boldsymbol{a},r-j}(y_t) \text{ for } \alpha \in \mathbb{F}_q.$$

By (6), we have $\Omega_{i,\mathcal{E},\alpha}(y_i) = \alpha$ for $t \in R_1 \setminus \{i\}$ and $\Omega_{i,\mathcal{E},\alpha}(y_t) = 0$ for $i \in R_1$, which means that the preceding equation can be rewritten as

$$0 = \sum_{t\in[m]} \theta_t \Omega_{i,\mathcal{E},\alpha}(y_t) H_{\boldsymbol{a},j}(y_t)$$

$$= \theta_i\Omega_{i,\mathcal{E},\alpha}(y_i)H_{\boldsymbol{a},j}(y_i) + \sum_{t\in[m]\setminus R_1} \theta_t\Omega_{i,\mathcal{E},\alpha}(y_t)H_{\boldsymbol{a},j}(y_t)$$
$$= \theta_i\alpha H_{\boldsymbol{a},j}(y_i) + \sum_{t\in[m]\setminus R_1} \theta_t\Omega_{i,\mathcal{E},\alpha}(y_t)H_{\boldsymbol{a},j}(y_t),$$

i.e.,

$$\theta_i\alpha H_{\boldsymbol{a},j}(y_i) = -\sum_{t\in[m]\setminus R_1} \theta_t\Omega_{i,\mathcal{E},\alpha}(y_t)H_{\boldsymbol{a},j}(y_t) \tag{7}$$

for $\ell \in [w-\delta+1]$, $i \in R_\ell$, and $\alpha \in \mathbb{F}_q$. Let $\{\alpha_1, \alpha_2, \dots, \alpha_T\}$ be a basis of $\mathbb{F}_q$ over $\mathbb{F}_{q_1}$ and $\{\beta_1, \beta_2, \dots, \beta_T\}$ be its dual basis. By (7), for $j \in [w-\delta+1]$, $i \in R_j$, and $v \in [T]$,

$$\begin{aligned}
&\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_{q_1}}(\theta_i\alpha_v H_{\boldsymbol{a},j}(y_i))\\
&= \mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_{q_1}}\left(-\sum_{t\in[m]\setminus R_1} \theta_t\Omega_{i,\mathcal{E},\alpha_t}(y_t)H_{\boldsymbol{a},j}(y_t)\right)\\
&= \sum_{t\in[m]\setminus R_1} \mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_{q_1}}\left(\frac{-\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_{q_1}}(\alpha_v\Lambda_{\mathcal{E}}(y_t))\,H_{\boldsymbol{a},j}(y_t)}{y_t-y_i}\right)\\
&= \sum_{t\in[m]\setminus R_1} \mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_{q_1}}(\alpha_v\Lambda_{\mathcal{E}}(y_t))\,\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_{q_1}}\left(\frac{-H_{\boldsymbol{a},j}(y_t)}{y_t-y_i}\right),
\end{aligned}$$

where the last two equalities hold by (6) and the fact $\mathrm{Tr}_{q/q_1}(\alpha_v\Lambda_{\mathcal{E}}(x)) \in \mathbb{F}_{q_1}$ for $x \in \mathbb{F}_q$. Therefore, for $\ell \in [w-\delta+1]$, $i \in R_\ell$, $H_{\boldsymbol{a},j}(y_i)$ can be recovered as

$$\theta_i H_{\boldsymbol{a},j}(y_i) = \sum_{v\in[T]} \mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_{q_1}}(\theta_i\alpha_v H_{\boldsymbol{a},j}(y_i))\,\beta_v,$$

by downloading $\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_{q_1}}\left(\frac{-H_{\boldsymbol{a},j}(y_t)}{y_t-y_i}\right)$ from the racks (columns) t for $t \in [m]\setminus R_1$. Namely, we can recover $H_{\boldsymbol{a},j}(y_i)$ for $\ell \in [w-\delta+1]$ and $i \in R_\ell$, i.e., $H_{\boldsymbol{a},j}(y_i)$ with $i \in [\tau]$ and $r+\delta-1-|E_i| \leqslant j \leqslant r-1$.

Secondly, for $i \in [\tau]$, we may rewrite $f_{\boldsymbol{a},i}(x)$ as

$$\begin{aligned}
f_{\boldsymbol{a},i}(x) &= \sum_{j=0}^{r-1} H_{\boldsymbol{a},j}(y_i)x^j\\
&= \sum_{j=r+\delta-1-|E_i|}^{r-1} H_{\boldsymbol{a},j}(y_i)x^j + g_{\boldsymbol{a},i}(x),
\end{aligned}$$

where $\deg(g_{\boldsymbol{a},i}(x)) \leqslant r+\delta-|E_i|-2$. Note that the above analysis recovers $H_{\boldsymbol{a},j}(y_i)$ for $i \in [\tau]$ and $r+\delta-|E_i|-1 \leqslant j \leqslant r-1$. Recall that we know $r+\delta-1-|E_i|$ elements of $\boldsymbol{A}_i$, i.e., those elements in $\boldsymbol{A}_i|_{[r+\delta-1]\setminus E_i}$. This is to say that $\boldsymbol{A}_i|_{[r+\delta-1]\setminus E_i}$ can recover $g_{\boldsymbol{a},i}(x)$ with $\deg(g_{\boldsymbol{a},i}(x)) \leqslant r+\delta-|E_i|-2$ for $i \in [\tau]$. Thus, for $i \in [\tau]$, $f_{\boldsymbol{a},i}(x) \equiv F_{\boldsymbol{a}}(x) \pmod{h(x)-y_i}$ can be recovered. Therefore, we can recover $\boldsymbol{A}_i$ for $i \in [\tau]$ according to Construction B, which completes the proof. ∎

When the finite field is sufficiently large, it is possible to further decrease the partial repair bandwidth. The main idea of repairing these erasure patterns is applying Lemma 3 to reduce this problem into the repair problem of a class of Reed-Solomon codes.

Lemma 4: Let $\boldsymbol{Y} = (y_1, y_2, \dots, y_m) \in \mathbb{F}_{q_1}^m$ and $\mathcal{C}_1 = \mathrm{GRS}_{m_1}(\boldsymbol{1}, \boldsymbol{Y})$ be an $[m = m_1+m_2, m_1, m_2+1]_{q_1}$ GRS code. If $\mathcal{C}_1$ has (τ, D) optimal repair property with $1 \leqslant \tau \leqslant m_2$ and $m_1 < D \leqslant m-\tau$, then for any $\mathcal{E} = \{E_{i_1}, E_{i_2}, \dots, E_{i_\tau}\}$ with $E_{i_j} \subseteq [L]$ and $|E_{i_j}| = w \geqslant \delta$ for $j \in [\tau]$, the code $\mathcal{C}$ generated by Construction B satisfies

$$B(\mathcal{C}, \mathcal{I}, \mathcal{E}, \mathcal{R}) \leqslant \frac{\tau D\ell(w-\delta+1)}{D-k+\tau}$$

for any $|\mathcal{I}| = \tau$ and $|\mathcal{R}| = D$ with $\mathcal{R} \subseteq [m] \setminus \mathcal{I}$, where $L = r+\delta-1$ and $q_1 = q^\ell$.

Proof: According to Construction B, in order to recover $\boldsymbol{A}_{i_t}|_{E_{i_t}}$, $i \in [\tau]$, it is sufficient to recover $f_{\boldsymbol{a},i_t}(x) = F_{\boldsymbol{a}}(x) \pmod{h(x)-y_{i_t}}$. Note that by Lemma 3, we have

$$f_{\boldsymbol{a},i_t}(x) = \sum_{j=0}^{r-1} H_{\boldsymbol{a},j}(y_{i_t})x^j,$$

where $\deg(H_{\boldsymbol{a},j}(x)) < m_1$. Since we know $L-w$ evaluations for each $f_{\boldsymbol{a},i_t}$, i.e., the values of $f_{\boldsymbol{a},i_t}(\beta_{i_t,j})$ for $j \in [\ell] \setminus E_{i_t}$, we only need to figure out $H_{\boldsymbol{a},j}(y_{i_t})$ for $0 \leqslant j \leqslant w-\delta$ in order to recover $f_{\boldsymbol{a},i_t}(\beta_{i_t,j})$ for $j \in E_{i_t}$. Recall that for each $i \in \mathcal{R} \subseteq [m]\setminus\mathcal{I}$ we know $r+\delta-1$ evaluations for $f_{\boldsymbol{a},i}$, which means we can figure out $H_{\boldsymbol{a},j}(y_i)$ for $0 \leqslant j \leqslant r-1$.

Now, for any $0 \leqslant j \leqslant w-\delta$, we regard $(H_{\boldsymbol{a},j}(y_1), H_{\boldsymbol{a},j}(y_2), \dots, H_{\boldsymbol{a},j}(y_{m_1+m_2}))$ as a codeword of $\mathcal{C}_1 = \mathrm{GRS}_{m_1}(\boldsymbol{1}, \boldsymbol{Y})$. In this way, the repair problem is reduced to $w-\delta+1$ parallel repair problems for the code $\mathcal{C}_1 = \mathrm{GRS}_{m_1}(\boldsymbol{1}, \boldsymbol{Y})$ when the code symbols with index $i \in \mathcal{R}$ are known (helper racks (columns)) and the code symbols with index $i \in \mathcal{I}$ are erasures. Since $\mathcal{C}_1$ has the (τ, D) optimal repair property, then for any $\mathcal{I} \subseteq [m]$ with $|\mathcal{I}| = \tau$ and $\mathcal{R} \subseteq [m] \setminus \mathcal{I}$ with $|\mathcal{R}| = D$, we can recover $H_{\boldsymbol{a},j}(y_i)$ for $i \in \mathcal{I}$ by downloading $\frac{\tau D}{D-k+\tau}$ elements over $\mathbb{F}_{q_1}$, i.e., $\frac{\tau D\ell}{D-k+\tau}$ elements over $\mathbb{F}_q$. Thus, the desired result follows since we have $w-\delta+1$ layers to recover $H_{\boldsymbol{a},j}(y_i)$ for these j and $i \in \mathcal{I}$. ∎

By the preceding proof of Lemma 4, the repair method achieves bandwidth $\frac{\tau D\ell(w-\delta+1)}{D-k+\tau}$ for the code $\mathcal{C}$ constructed by Construction B only for the specified $\mathcal{I}$ and $\mathcal{R}$. We remark that there are known explicit repair methods for GRS codes such as the one introduced in [44] for multiple erasures.

A. A Lower Bound on the Partial-Repair Bandwidth

A natural question arises regarding the partial-repair problem: what is the theoretical bound for the partial-repair bandwidth? Since the main tool to solve this problem is a kind of direct graphs, in this section, we use vertex to denote a rack (column) of the RASL codes. Let L denote the number of elements stored in each node. Assume that there is a node, say the i-th vertex, suffering from $L-s_i$ erasures in the positions given in E_i. Define $\beta(L, s_i, D)$ as the number of elements the system needs to download from each of the D helper vertices to recover the $L-s_i$ erased elements.

Inspired by the idea of the information-flow graph presented in [11], we propose a solution to this problem by defining a special kind of information-flow graph called a *partial information-flow graph*. The basic idea is to allow each vertex that experiences erasures to have a certain amount of surviving information. When a vertex experiences partial erasure, the system needs to recover the erased portion of information for the goal vertex. Since the recovered vertex inherits the surviving information from the original vertex, it is named as an inheritor.

Definition 9: A directed acyclic graph is said to be a *partial information-flow graph* if it satisfies the following:

- A source vertex S, corresponding to the original data which will be stored into N initial storage vertices.
- Initial storage vertices $X^{(i)}$, each of them consists of an input vertex $X_{in}^{(i)}$ and an output vertex $X_{out}^{(i)}$. $X_{in}^{(i)}$ and $X_{out}^{(i)}$ are connected by a directed edge $(X_{in}^{(i)}, X_{out}^{(i)})$ with capacity equal to the number of elements stored at $X^{(i)}$, i.e., $r \leqslant L$, where r is the number of original elements stored at $X_{in}^{(i)}$. S connects to each $X_{in}^{(i)}$ by a directed edge $(S, X_{in}^{(i)})$ with capacity r.
- To model the dynamic behavior of storage systems such as erasures and repair, the time factor is also considered. At any given time, vertices are either active or inactive. At the initial time step, the storage vertices $X^{(i)}$ are all active and the source vertex S is inactive. Later on, at any given time step, if a vertex suffers from a partial erasure, say the vertex $v = (v_{in}, v_{out})$, then the vertex v is set to be inactive and we create a direct inheritor (I_{in}, I_{out}), which is connected with v_{out} by an edge (v_{out}, I_{in}) with capacity $s(v)$, where $s(v)$ is the number of surviving information symbols of the vertex v. The vertex I also needs to download $\beta(L, s(v), D)$ symbols from each of D other active vertices, i.e., we add D directed edges $(v_{out}^{(j)}, I_{in}^{(i)})$ with a capacity of $\beta = \beta(L, s(v), D)$. Finally, we set the vertex I as active.
- A data collector vertex DC, corresponding to a request to construct the data. DC connects to K active vertices with subscript "out" by directed edges with infinite capacity to recover the original data.

Remark 8: The partial information-flow graph is a direct generalization of the information-flow graph to include the case of repairing partial erasures of a vertex. Specifically, each failed vertex v_i (progenitor) will have a portion of information (denoted by s_i surviving elements) inherited by its reconstructed inheritor vertex. We also extend the information-flow graph to include the possibility that each vertex may contain redundancies, which corresponds to the concept of locality in Definition 4, where each vertex is regarded as a repair set. Therefore, when $s_i = 0$ for all $i \in [N]$ and $L = r$, the partial information-flow graph reduces to the information-flow graph defined in [11].

For partial information-flow graphs, two examples are presented in Fig. 5 and 6 to explain the initiation and dynamic behavior of storage systems, respectively. We now use partial information-flow graphs to analyze the repair problem of distributed storage systems in the following setting:

- Basic recovery assumption: The file contains M symbols, and is stored in N vertices, each of which has L symbols, and any K vertices are capable of recovering all the M symbols of the file.
- $(r, L-r+1)$-locality: For each vertex, any $r \leqslant L$ symbols can recover all the symbols stored in this vertex.

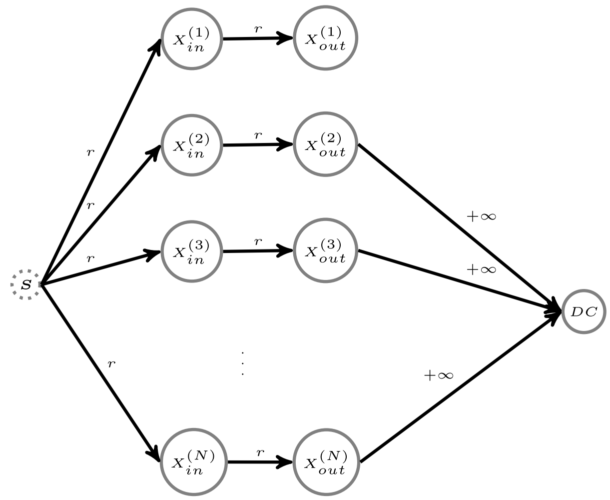

Fig. 5. An example of the the initial status of the information flow graph, where one possible data collector DC is included.

- Erasure model: There are $L-s_i$ symbols in the i-th vertex that are erased.

Given positive integers $N > K$, $D \leqslant N-1$, $L \geqslant r$, $s_i \leqslant L$ for all $i \in [N]$ and a real number $\beta \geqslant 0$, let $G(N, K, r, D, \beta; \boldsymbol{s} = (s_1, s_2, \ldots, s_N))$ denote a family of partial information-flow graphs with all possible evolutions. The parameter tuple $(N, K, r, D, \beta; \boldsymbol{s})$ is said to be *feasible* if there exists a locally repairable array code with repair bandwidth β and sub-packetization L, with $L - s_i$ erasures in the i-th vertex.

Proposition 1: Let $G \in G(N, K, r, D, \beta; \boldsymbol{s} = (s_1, s_2, \ldots, s_N))$ be the partial information-flow graph for a given time step. Then the following hold:

- For any $v \in G$, there exists a unique list of vertices $B(v) = \{v^{(1)} = v, v^{(2)}, \cdots, v^{(t)}\}$ such that $v^{(i+1)}$ is an inheritor of $v^{(i)}$ for $i \in [t-1]$, where t denotes the amount of inheritors in G for vertex v.
- For any pair of vertices $v, v' \in G$, if $v' \notin B(v)$ and $v \notin B(v')$ then there is at most one edge $(v_{out}^{(i)}, v'_{in})$ for $v_{out}^{(i)} \in \widetilde{B}(v)$ and $v'_{in} \in \widetilde{B}(v')$, where $\widetilde{B}(v) = \{v_{in}, v_{out} : v = (v_{in}, v_{out}) \in B(v)\}$.

Proof: For each vertex v in G, if it has a direct inheritor, then that inheritor is unique. We can then define $B(v)$ as the set of all inheritors (direct or indirect) of v in G, which is also unique. The size of $B(v)$, denoted by $|B(v)|$, represents the number of inheritors of v in G, including v itself.

For the second part, we assume that there exists an edge $(v_{out}^{(i)}, v'_{in})$ for $v^{(i)} \in B(v)$. Since $v' \notin B(v)$ and $v \notin B(v')$, we can conclude that $v^{(i)}$ is active when v' is included in G. This also means that $v^{(j)}$ for $j \in [i-1]$ are inactive, i.e., $(v_{out}^{(j)}, v'_{in}) \notin E$. Note that $v^{(i)}$ being active implies that $v^{(j)}$ for $j > i$ have not been included yet, i.e., $(v_{out}^{(j)}, v'_{in}) \notin E$. This completes the proof. ■

Definition 10: A *cut* in the partial information flow graph G between S and DC is a subset of edges W such that each directed path from S to DC contains at least one edge in W. Furthermore, the minimal cut is the cut with the smallest edge

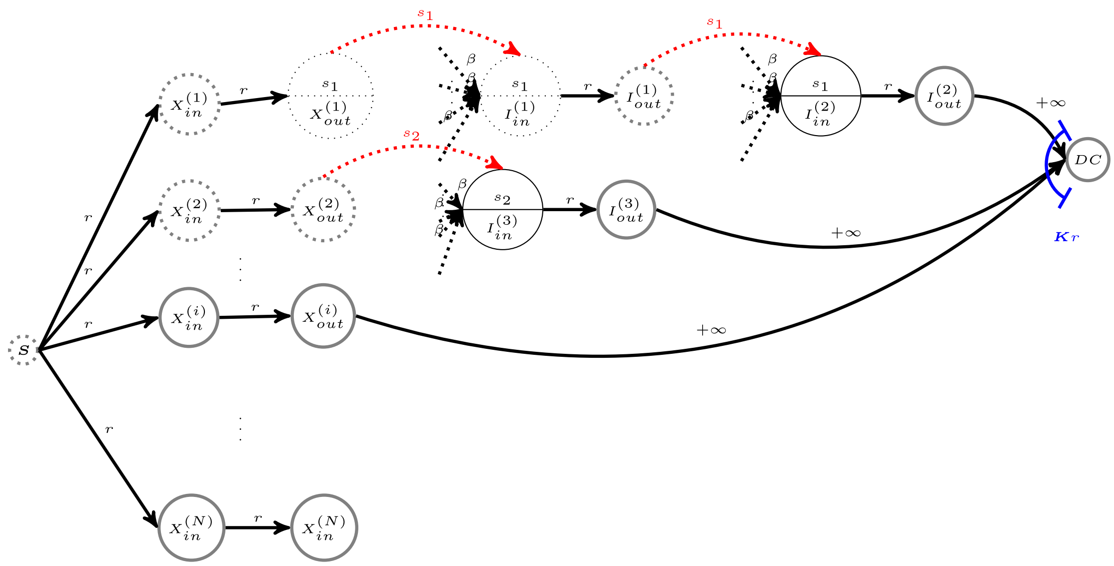

Fig. 6. An example for three inheritors in the partial information flow graph with parent $X^{(i)}$, $i \in \{1, 2\}$, where the first node suffers partial erasures twice and the second node is erased once. Here we use dotted vertices to denote inactive nodes. We also include a possible data collector DC as an example.

capacity sum. We define the capacity of W as

$$C(W) = \sum_{e \in W} C(e),$$

where $C(e)$ denotes the capacity of an edge e.

Theorem 6: For given positive integers $N > K$, $D \leqslant N - 1$, $L \geqslant r$, $s_i \leqslant L$ for $i \in [N]$, the parameter tuple $(N, K, D, L, \beta; \boldsymbol{s})$ is feasible if and only if $Kr \leqslant c(N, K, D, \beta; \boldsymbol{s})$ under a large enough finite field, where $c(N, K, D, \beta; \boldsymbol{s})$ satisfies

$$\begin{aligned} c(N, K, D, \beta; \boldsymbol{s}) = &\sum_{i=0}^{\min\{K,D\}-1} \min\{(D-i)\beta + s_{j_i}, r\} \\ &+ \sum_{i=\min\{K,D\}}^{K-1} \min\{s_{j_i}, r\}, \end{aligned}$$

where $\beta \geqslant 0$ is a real number, and $s_{j_0} \leqslant s_{j_1} \leqslant \ldots \leqslant s_{j_{N-1}}$.

Before going to the proof of Theorem 6, we recall the well known theorem about minimum cut and maximum flow for network coding, and translate it into our partial information-flow graph setting.

Lemma 5: For a partial information-flow graph G, a data collector DC can recover the original file if and only if the minimum cut between the source S and a data collector DC is larger than or equal to the file size under a large enough finite field.

Proof: Regard DC as a terminal, then the partial information-flow graph G can be viewed as a broadcast network with source S and original file with size M. Then, according to the well known cut-set bound [10] for network coding, under a large enough finite field the network has a linear solution [16], i.e., DC can get enough information to recover the original file if and only if the minimum cut between the source S and the terminal DC is larger than or equal to the file size M. ∎

Proof of Theorem 6: According to Lemma 5, proving this theorem is equivalent to proving that the minimum cut between the source vertex S and the data collector DC is larger than or equal to the file size $M = Kr$ if and only if $Kr \leqslant c(N, K, D, \beta; \boldsymbol{s})$, for any $G \in G(N, K, r, D, \beta; \boldsymbol{s})$.

In one direction, we show that there exists a partial information-flow graph $G_1 \in G(N, K, r, D, \beta; \boldsymbol{s})$ such that the minimum cut $W(G_1)$ between the source S and the data collector DC is at most $c(N, K, D, \beta; \boldsymbol{s})$, i.e., $C(W(G_1)) \leqslant c(N, K, D, \beta; \boldsymbol{s})$. Consider a sequence of inheritor vertices denoted as $(I_{in}^{(j_t)}, I_{out}^{(j_t)})$ for $0 \leqslant t \leqslant K-1$ such that there exists a directed edge $(I_{out}^{(j_{t_1})}, I_{in}^{(j_{t_2})})$ with capacity β if and only if $0 < t_2 - t_1 \leqslant D$. Now consider the data collector that connects to the vertices $(I_{in}^{(j_t)}, I_{out}^{(j_t)})$ for $0 \leqslant t \leqslant K-1$. It is easy to check that the minimum cut satisfies

$$C(W(G_1)) \leqslant \begin{cases} \sum_{i=0}^{K-1} \min\{(D-i)\beta + s_{j_i}, r\}, & \text{if } K \leqslant D, \\ \sum_{i=0}^{D-1} \min\{(D-i)\beta + s_{j_i}, r\} \\ \quad + \sum_{i=D}^{K-1} \min\{s_{j_i}, r\}, & \text{if } K > D, \end{cases}$$

as illustrated in Fig. 7. Thus, G_1 is feasible only if

$$\begin{aligned} &\sum_{i=0}^{\min\{K,D\}-1} \min\{(D-i)\beta + s_{j_i}, r\} \\ &\qquad + \sum_{i=\min\{D,K\}}^{K-1} \min\{s_{j_i}, r\} \geqslant Kr. \end{aligned}$$

In the other direction, we claim that for any $G_2 \in G(N, K, r, D, \beta; \boldsymbol{s})$ with D helper vertices, any cut W between the source S and a data collector DC should satisfy

$$C(W) \geqslant \sum_{i=0}^{\min\{K,D\}-1} \min\{(D-i)\beta + s_{j_i}, r\} + \sum_{i=\min\{D,K\}}^{K-1} \min\{s_{j_i}, r\}.$$

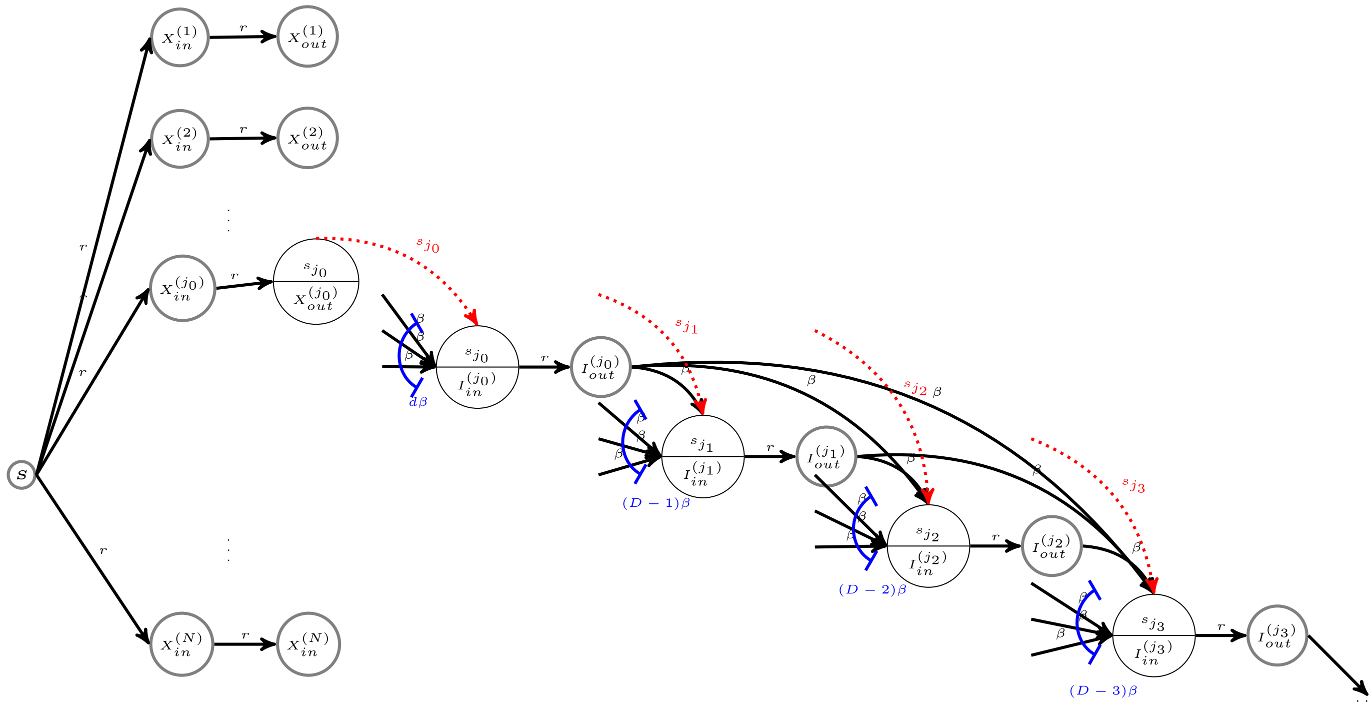

Fig. 7. Each inheritor receives s_j surviving symbols from its parent, which does not cost network traffic, and is denoted with a dotted arrow. We also highlight the inheritance by dividing the node into two parts, where the one labeled with s_j corresponds to the inheritance.

To prove this we assume the K active vertices connected with DC are $\{Y^{(i)}_{out} : i \in [K]\}$ and the cut W partitions the vertices into two subsets, U and $\overline{U}$, such that $S \in U$ and $DC \in \overline{U}$. Let $\overline{U}_{out}$ denote the set of vertices with subscript "out". We only need to consider the case $\{Y^{(i)}_{out} : i \in [K]\} \subseteq \overline{U}_{out}$ since the capacity between $Y^{(i)}_{out}$ and DC is infinite which implies $C(W) = +\infty$ for the case $Y^{(i)}_{out} \in U$. Since the graph G_2 is a directed acyclic graph then it has a topological sorting which means the vertices can be ordered such that the edge (v_i, v_j) implies that $i < j$. We assume the set $\{Y^{(i)}_{out} : i \in [K]\} \cap \overline{U}$ has size τ and is topologically ordered as $(Y^{(i_1)}_{out}, Y^{(i_2)}_{out}, \cdots, Y^{(i_\tau)}_{out})$, where $\tau \leqslant K$. Now, we find a list of vertices $(L^1_{out}, L^2_{out}, \cdots, L^K_{out})$ satisfying the following two conditions:

I L^1_{out} is the first vertex of $\overline{U}_{out}$ under the topological sorting.

II For $i \in [K]$, L^i_{out} is the first vertex of $\overline{U}_{out} \setminus (\bigcup_{j=1}^{i-1} \widetilde{B}(L^j))$ under the topological sorting.

For L^1_{out}, if $L^1_{in} \notin \overline{U}$ then $(L^1_{in}, L^1_{out}) \in W$ with capacity r. If $L^1_{in} \in \overline{U}$ then, by Condition I, W contains all the input edges of L^1_{in}. Those edges have capacity r in total if L^1 does not have a progenitor, or capacity $s_{t_1} + D\beta$ otherwise.

For $1 < i \leqslant K$ and $L^i_{in} \notin \overline{U}$, we have $(L^i_{in}, L^i_{out}) \in W$ with capacity r.

For $1 < i \leqslant K$ and $L^i_{in} \in \overline{U}$, we have the following two subcases:

- L^i has no progenitor, i.e., L^i is one of the original storage vertices connected with the source S. In that case, we have $(S, L^i_{in}) \in W$ with capacity r.
- L^i has a progenitor, say L. In that case, we claim that $L_{out} \in U$. Otherwise, by Proposition 1-(1), the fact $L^i \in \overline{U}_{out} \setminus (\bigcup_{j=1}^{i-1} \widetilde{B}(L^j))$ means $L_{out} \in \overline{U}_{out} \setminus (\bigcup_{j=1}^{i-1} \widetilde{B}(L^j))$, which contradicts the fact that L^i is the first vertex of $\overline{U}_{out} \setminus (\bigcup_{j=1}^{i-1} \widetilde{B}(L^j))$ under the the topological sorting. Similarly, W should contain all the input edges of L^i_{in} except for those edges from vertices belonging to $\widetilde{B}(L^j)$ for $j \in [i-1]$. By Proposition 1-(2), there are at most $i-1$ edges coming from the vertices of $\widetilde{B}(L^j)$ for $j \in [i-1]$ and each of them should have capacity β. Thus, W should contain input edges of L^i_{in} with total capacity at least s_{t_i} if $i > D$ and $s_{t_i} + (D-i+1)\beta$ otherwise. Herein, we highlight that $t_i \notin \{t_1, t_2, \ldots, t_{i-1}\}$ since $L^i \in \overline{U}_{out} \setminus (\bigcup_{j=1}^{i-1} \widetilde{B}(L^j))$.

Therefore, for any cut W between S and DC, the capacity satisfies

$$C(W) \geqslant \sum_{i=0}^{\min\{K,D\}-1} \min\{(D-i)\beta + s_{j_i}, r\} + \sum_{i=\min\{D,K\}}^{K-1} \min\{s_{j_i}, r\},$$

where the last inequality holds by the fact $s_{j_1} \leqslant s_{j_2} \leqslant \ldots \leqslant s_{j_N}$, i.e.,

$$\sum_{i=1}^{K} s_{t_i} \geqslant \sum_{i=1}^{K} s_{j_i}$$

for any possible t_i with $0 \leqslant i \leqslant K-1$. Thus, we have

$$C_{\min} = C(W(G_1)) = \sum_{i=0}^{\min\{K,D\}-1} \min\{(D-i)\beta + s_{j_i}, r\} + \sum_{i=\min\{D,K\}}^{K-1} \min\{s_{j_i}, r\},$$

where $C_{\min}$ denotes the minimum cut capacity of all the graphs $G \in G(N,K,r,D,\beta;\boldsymbol{s})$. For an example of the exact minimum cut, the reader may refer to Fig. 7. Therefore, $C_{\min} \geqslant Kr$ is equivalent with $\sum_{i=0}^{\min\{K,D\}-1} \min\{(D-i)\beta + s_{j_i}, r\} + \sum_{i=\min\{D,K\}}^{K-1} \min\{s_{j_i}, r\} \geqslant Kr$. ■

As the next step, we study the relationship between r and N,K,β,D,M by solving the inequality $Kr \leqslant c(N,K,D,\beta;\boldsymbol{s})$ in Theorem 6.

Theorem 7: Let $\boldsymbol{s} = (s_1, s_2, \cdots, s_N)$ and $s_{j_0} \leqslant s_{j_1} \leqslant \ldots \leqslant s_{j_{N-1}} < s_{j_N} \triangleq +\infty$. For $1 \leqslant t \leqslant K-1$, define

$$s_{j_t}^* \triangleq \frac{\sum_{i=0}^{t-1} s_{j_i} + (K-t)s_{j_t}}{K},$$

$$g(t) \triangleq \sum_{i=0}^{t-1} s_{j_i} + t(D-K)\beta + \frac{t(t+1)\beta}{2},$$

and for $0 \leqslant t \leqslant K-1$,

$$f(t) \triangleq \frac{1}{K(D+1-K) + \frac{t(t+1)}{2} + t(K-t-1)}. \tag{8}$$

When $s_{j_{\tau-1}} < r \leqslant s_{j_\tau}$ for $\tau \in [N]$, the parameter tuple $(N,K,r,D,\beta;\boldsymbol{s})$ is feasible if and only if $r \geqslant r^*(N,K,D,\beta,\boldsymbol{s})$ and the solution can be achieved via linear codes, where

$$r^*(N,K,D,\beta;\boldsymbol{s}) = \begin{cases} \frac{M}{K}, & \beta \in [f(0)K(r-s_{j_0}), \infty), \\ \frac{M-g(t)}{K-t}, & \beta \in (f(t)K(r-s_{j_t}^*), f(t-1)K(r-s_{j_{t-1}}^*)], \end{cases}$$

for $t \in [\tau^*-1]$ with $\tau^* \triangleq \min\{\tau, K\}$. If $r \leqslant s_{j_0}$ then the parameter tuple $(N,K,r,D,\beta;\boldsymbol{s})$ is always feasible for any $\beta \geqslant 0$.

Proof: By Theorem 6, we need to determine the exact threshold by considering fixed values of β, $\boldsymbol{s}$, and $D \geqslant K$, and minimizing $r = r_{\min}(\beta, \boldsymbol{s}, D)$ such that

$$\sum_{i=0}^{K-1} \min\{(D-i)\beta + s_{j_i}, r\} \geqslant M.$$

Let

$$b_i = (D-(K-1-i))\beta + s_{j_i} \quad \text{for } 0 \leqslant i \leqslant K-1 \tag{9}$$

and $b_K = +\infty$. Let τ be the integer such that $s_{j_{\tau-1}} < r \leqslant s_{j_\tau}$ with $1 \leqslant \tau \leqslant N$, we have

$$C(r) \triangleq \sum_{i=0}^{K-1} \min\{(D-i)\beta + s_{j_i}, r\} = \begin{cases} Kr, & 0 \leqslant r < b_0, \\ (K-t)r + \sum\limits_{0 \leqslant i \leqslant t-1} b_i, & b_{t-1} \leqslant r < b_t, \\ & 1 \leqslant t \leqslant \tau^* = \min\{\tau, K\}, \end{cases}$$

where $b_0 \leqslant b_1 \leqslant \ldots \leqslant b_{K-1}$ by (9). The preceding equality means that $C(r)$ is strictly increasing from 0 to $(K-\tau^*)r + \sum_{i=0}^{\tau^*-1} b_i$ as r increasing from 0 to b_{τ^*-1}. Thus, to find a minimum value for r such that $C(r) \geqslant M$, we only need to calculate $C^{-1}(M)$ when $M \leqslant (K-\tau^*)r + \sum_{i=0}^{\tau^*-1} b_i$. That is, for $1 \leqslant t \leqslant \tau^*-1$

$$C^{-1}(M) = \begin{cases} \frac{M}{K}, & M \in [0, Kb_0), \\ \frac{M-g(t)}{K-t}, & M \in [\sum_{i=0}^{t-2} b_i + (K-t+1)b_{t-1}, \\ & \qquad \sum_{i=0}^{t-1} b_i + (K-t)b_t), \end{cases} \tag{10}$$

where

$$\sum_{i=0}^{t-1} b_i = \sum_{i=0}^{t-1} s_{j_i} + t(D-K)\beta + \frac{t(t+1)\beta}{2} = g(t).$$

Recall that for $1 \leqslant t \leqslant K$,

$$\begin{aligned} &\sum_{i=0}^{t-1} b_i + (K-t)b_t \\ &= \sum_{i=0}^{t-1} s_{j_i} + t(D-K)\beta + \frac{t(t+1)\beta}{2} \\ &\quad + (K-t)(s_{j_t} + (D+1-K+t)\beta) \\ &= \sum_{i=0}^{t-1} s_{j_i} + (K-t)s_{j_t} + K(D+1-K)\beta \\ &\quad + \frac{t(t+1)\beta}{2} + t(K-t-1)\beta \\ &= Ks_{j_t}^* + \frac{\beta}{f(t)}, \end{aligned} \tag{11}$$

where $s_{j_t}^* \triangleq \frac{\sum_{i=0}^{t-1} s_{j_i} + (K-t)s_{j_t}}{K}$ and the last equality holds by (8). Note that $M = Kr$. Combining (11) and (10),

$$\sum_{i=0}^{t-2} b_i + (K-t+1)b_{t-1} \leqslant M < \sum_{i=0}^{t-1} b_i + (K-t)b_t$$

is equivalent with

$$K(r-s_{j_t}^*)f(t) < \beta \leqslant K(r-s_{j_{t-1}}^*)f(t-1)$$

for $1 \leqslant t \leqslant K$, which results in the claim for the case $r > s_{j_0}$. Finally, note that for the case $r \leqslant s_{j_0} \leqslant \ldots \leqslant s_{j_{N-1}}$, we have $C(r) \equiv Kr = M$, which means for this case the parameter tuple $(N,K,r,D,\beta;\boldsymbol{s})$ is always feasible. ■

Remark 9: For the case $L > s_{j_0} \geqslant r$, i.e., the case that the number of erased symbols is less than $L-r$, we know that $\beta = 0$ is sufficient to repair those erasures since we have $(r, L-s+1)$-locality inside each vertex.

Remark 10: For the case $s_{j_0} = s_{j_1} = \cdots = s_{j_{N-1}} = 0$ and $L = r$, i.e., the ordinary case without locality, the bounds in Theorems 6 and 7 are exactly the cut-set bound described in [11] and [19] for the rack-aware model.

Considering the regular case, $s_1 = s_2 = \cdots = s_N$, we have the following corollary that is implied directly from Theorem 7.

Corollary 4: Let $\mathcal{C}$ be an $(N,K,k=Kr; L=r+\delta-1,\delta)$-RASL code, where each column is a repair set with $(r, \delta = L-r+1)$-locality. Let D be an integer satisfying $K < D \leqslant N-1$. For any $i \in [N]$, $E_i \subseteq [L]$, and D-subset $\mathcal{R} \subseteq [N] \setminus \{i\}$, we have

$$B(\mathcal{C}, \{i\}, \{E_i\}, \mathcal{R}) \geqslant \begin{cases} \frac{D(|E_i|-\delta+1)}{D-K+1}, & \text{if } |E_i| \geqslant \delta, \\ 0, & \text{otherwise.} \end{cases}$$

Proof: By setting $s_1 = s_2 = \cdots = s_N = L - |E_i|$, if $L - |E_i| < r$, i.e., $|E_i| \geqslant \delta$, then we have

$$\beta \geqslant \frac{K(|E_i| - \delta + 1)}{K(D - K + 1)} = \frac{|E_i| - \delta + 1}{D - K + 1}$$

by Theorem 7, since $M = Kr$. Thus, we have $B(\mathcal{C}, \{i\}, \{E_i\}, \mathcal{R}) = D\beta \geqslant \frac{D(|E_i|-\delta+1)}{D-K+1}$ when $|E_i| \geqslant \delta$. If $|E_i| < \delta$, then by Theorem 7 we have $\beta \geqslant 0$, i.e., $B(\mathcal{C}, \{i\}, \{E_i\}, \mathcal{R}) \geqslant 0$ ■

Thus, similarly to Corollary 3, we have the following conclusion for partial repairing for Tamo-Barg codes, which follows directly from Corollary 4 and Theorem 4.

Corollary 5: Consider the setting of Construction B. Let q_0 be a prime power, and $\mathbb{F}_q = \mathbb{F}_{q_0}(y_1, y_2, \ldots, y_m)$, $\mathbb{F}_{q_i} = \mathbb{F}_{q_0}(y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_m)$. Define $w_i \triangleq [\mathbb{F}_q : \mathbb{F}_{q_i}]$ for each $i \in [m]$. For any given erasure set $E \subseteq [L]$ with $\delta \leqslant |E| \leqslant L$, we can recover $\boldsymbol{A}_i|_E$ by downloading $\frac{(r+m_1-1)(|E|-\delta+1)}{r}$ symbols in $\mathbb{F}_q$ from any other $r + m_1 - 1$ racks (columns), which is exactly the optimal bandwidth with respect to the cut-set bound according to Corollary 4.

VII. Conclusion

In this paper, we explored the rack-aware systems with locality. Under this model, by arranging each repair set as a rack, we considered repairing erasures that extend beyond the locality for locally repairable codes. We presented two repair schemes to minimize the repair bandwidth for Tamo-Barg codes under the rack-aware model by designating each repair set as a rack. One of the schemes achieved optimal repair for a single rack erasure. Additionally, the cut-set bound was established for locally repairable codes under the rack-aware model, and our repair schemes were proven to be optimal with respect to this bound.

Moreover, we also studied the partial-repair problem for locally repairable codes under the rack-aware model with locality, and we introduced repair schemes and bounds for this scenario. However, research on the partial-repair problem for locally repairable codes is still in its early stages. In general, the partial-repair problem for locally repairable codes may be a three-dimensional one where each rack (column) contains L symbols with locality, and each symbol is a vector over a given base field, say $\mathbb{F}_q$. This general problem is still widely open. Furthermore, it remains an open question how to repair known locally repairable codes such as maximally recoverable codes (partial MDS codes), sector-disk codes, and locally repairable codes with super-linear length. We leave these questions for future work.

Acknowledgment

The authors would like to thank the anonymous reviewers and the Associate Editor, Prof. Itzhak Tamo, for their valuable comments that improved the quality and presentation of this article.

References

[1] V. R. Cadambe, S. A. Jafar, H. Maleki, K. Ramchandran, and C. Suh, "Asymptotic interference alignment for optimal repair of MDS codes in distributed storage," *IEEE Trans. Inf. Theory*, vol. 59, no. 5, pp. 2974–2987, May 2013.

[2] V. R. Cadambe and A. Mazumdar, "Bounds on the size of locally recoverable codes," *IEEE Trans. Inf. Theory*, vol. 61, no. 11, pp. 5787–5794, Nov. 2015.

[3] H. Cai, M. Cheng, C. Fan, and X. Tang, "Optimal locally repairable systematic codes based on packings," *IEEE Trans. Commun.*, vol. 67, no. 1, pp. 39–49, Jan. 2019.

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

[5] H. Cai, Y. Miao, M. Schwartz, and X. Tang, "On optimal locally repairable codes with super-linear length," *IEEE Trans. Inf. Theory*, vol. 66, no. 8, pp. 4853–4868, Aug. 2020.

[6] H. Cai and M. Schwartz, "On optimal locally repairable codes and generalized sector-disk codes," *IEEE Trans. Inf. Theory*, vol. 67, no. 2, pp. 686–704, Feb. 2021.

[7] B. Chen, W. Fang, S.-T. Xia, J. Hao, and F.-W. Fu, "Improved bounds and singleton-optimal constructions of locally repairable codes with minimum distance 5 and 6," *IEEE Trans. Inf. Theory*, vol. 67, no. 1, pp. 217–231, Jan. 2021.

[8] Z. Chen and A. Barg, "Explicit constructions of MSR codes for clustered distributed storage: The rack-aware storage model," *IEEE Trans. Inf. Theory*, vol. 66, no. 2, pp. 886–899, Feb. 2020.

[9] Z. Chen, M. Ye, and A. Barg, "Enabling optimal access and error correction for the repair of Reed–Solomon codes," 2020, *arXiv:2001.07189*.

[10] T. M. Cover, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 1999.

[11] 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.

[12] M. Forbes and S. Yekhanin, "On the locality of codeword symbols in non-linear codes," *Discrete Math.*, vol. 324, no. 6, pp. 78–84, 2014.

[13] 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.

[14] V. Guruswami and M. Wootters, "Repairing Reed–Solomon codes," *IEEE Trans. Inf. Theory*, vol. 63, no. 9, pp. 5684–5698, Sep. 2017.

[15] J. Hao, S. Xia, K. W. Shum, B. Chen, F. Fu, and Y. Yang, "Bounds and constructions of locally repairable codes: Parity-check matrix approach," *IEEE Trans. Inf. Theory*, vol. 66, no. 12, pp. 7465–7474, Dec. 2020.

[16] T. Ho et al., "A random linear network coding approach to multicast," *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4413–4430, Oct. 2006.

[17] L. Holzbaur, S. Puchinger, E. Yaakobi, and A. Wachter-Zeh, "Partial MDS codes with regeneration," *IEEE Trans. Inf. Theory*, vol. 67, no. 10, pp. 6425–6441, Oct. 2021.

[18] H. Hou, P. P. C. Lee, K. W. Shum, and Y. Hu, "Rack-aware regenerating codes for data centers," *IEEE Trans. Inf. Theory*, vol. 65, no. 8, pp. 4730–4745, Aug. 2019.

[19] Y. Hu, P. P. C. Lee, and X. Zhang, "Double regenerating codes for hierarchical data centers," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 245–249.

[20] Y. Hu et al., "Optimal repair layering for erasure-coded data centers," *ACM Trans. Storage*, vol. 13, no. 4, pp. 1–24, Nov. 2017.

[21] C. Huang, M. Chen, and J. Li, "Pyramid codes: Flexible schemes to trade space for access efficiency in reliable data storage systems," *ACM Trans. Storage*, vol. 9, no. 1, pp. 1–28, Mar. 2013.

[22] L. Jin, L. Ma, and C. Xing, "Construction of optimal locally repairable codes via automorphism groups of rational function fields," *IEEE Trans. Inf. Theory*, vol. 66, no. 1, pp. 210–221, Jan. 2020.

[23] 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.

[24] G. Kim and J. Lee, "Locally repairable codes with unequal local erasure correction," *IEEE Trans. Inf. Theory*, vol. 64, no. 11, pp. 7137–7152, Nov. 2018.

[25] J. Li and X. Tang, "Optimal exact repair strategy for the parity nodes of the $(k+2, k)$ zigzag code," *IEEE Trans. Inf. Theory*, vol. 62, no. 9, pp. 4848–4856, Sep. 2016.

[26] J. Li, X. Tang, H. Hou, Y. S. Han, B. Bai, and G. Zhang, "PMDS array codes with small sub-packetization, small repair bandwidth/rebuilding access," *IEEE Trans. Inf. Theory*, vol. 69, no. 3, pp. 1551–1566, Mar. 2023.

[27] J. Li, X. Tang, and U. Parampalli, "A framework of constructions of minimal storage regenerating codes with the optimal access/update property," *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1920–1932, Apr. 2015.

[28] J. Li, X. Tang, and C. Tian, "A generic transformation to enable optimal repair in MDS codes for distributed storage systems," *IEEE Trans. Inf. Theory*, vol. 64, no. 9, pp. 6257–6267, Sep. 2018.

[29] X. Li, L. Ma, and C. Xing, "Construction of asymptotically good locally repairable codes via automorphism groups of function fields," *IEEE Trans. Inf. Theory*, vol. 65, no. 11, pp. 7087–7094, Nov. 2019.

[30] J. Liu, S. Mesnager, and L. Chen, "New constructions of optimal locally recoverable codes via good polynomials," *IEEE Trans. Inf. Theory*, vol. 64, no. 2, pp. 889–899, Feb. 2018.

[31] Y. Liu, J. Li, and X. Tang, "A generic transformation to enable optimal repair/access MDS array codes with multiple repair degrees," *IEEE Trans. Inf. Theory*, vol. 69, no. 7, pp. 4407–4428, 2023.

[32] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, vol. 16. Amsterdam, The Netherlands: Elsevier, 1977.

[33] A. Mazumdar, V. Chandar, and G. W. Wornell, "Update-efficiency and local repairability limits for capacity approaching codes," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 5, pp. 976–988, May 2014.

[34] D. Pei, A. Salomaa, and C. Ding, *Chinese Remainder Theorem: Applications in Computing, Coding, Cryptography*. Singapore: World Scientific, 1996.

[35] V. Pless, W. Huffman, and R. Brualdi, *Handbook of Coding Theory*. Amsterdam, The Netherlands: Elsevier, 1998.

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

[37] 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.

[38] A. S. Rawat, D. S. Papailiopoulos, A. G. Dimakis, and S. Vishwanath, "Locality and availability in distributed storage," *IEEE Trans. Inf. Theory*, vol. 62, no. 8, pp. 4481–4493, Aug. 2016.

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

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

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

[42] I. Tamo, Z. Wang, and J. Bruck, "Zigzag codes: MDS array codes with optimal rebuilding," *IEEE Trans. Inf. Theory*, vol. 59, no. 3, pp. 1597–1616, Mar. 2013.

[43] I. Tamo, M. Ye, and A. Barg, "Optimal repair of reed-solomon codes: Achieving the cut-set bound," in *Proc. IEEE 58th Annu. Symp. Found. Comput. Sci. (FOCS)*, Oct. 2017, pp. 216–227.

[44] I. Tamo, M. Ye, and A. Barg, "The repair problem for Reed–Solomon codes: Optimal repair of single and multiple erasures with almost optimal node size," *IEEE Trans. Inf. Theory*, vol. 65, no. 5, pp. 2673–2695, May 2019.

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

[46] J. Wang, D. Zheng, S. Li, and X. Tang, "Rack-aware MSR codes with error correction capability for multiple erasure tolerance," *IEEE Trans. Inf. Theory*, vol. 69, no. 10, pp. 6428–6442, Feb. 2023.

[47] C. Xing and C. Yuan, "Construction of optimal locally recoverable codes and connection with hypergraph," in *Proc. 46th Int. Colloq. Automata, Lang., Program. (ICALP)*, 2019, pp. 98:1–98:13.

[48] M. Ye and A. Barg, "Explicit constructions of high-rate MDS array codes with optimal repair bandwidth," *IEEE Trans. Inf. Theory*, vol. 63, no. 4, pp. 2001–2014, Apr. 2017.

[49] A. Zeh and E. Yaakobi, "Bounds and constructions of codes with multiple localities," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 640–644.

[50] Z. Zhang and L. Zhou, "A vertical-horizontal framework for building rack-aware regenerating codes," *IEEE Trans. Inf. Theory*, vol. 69, no. 5, pp. 2874–2885, May 2023.

Han Cai (Member, IEEE) received the B.S. and M.S. degrees in mathematics from Hubei University, Wuhan, China, in 2009 and 2013, respectively, and the Ph.D. degree from the Department of Communication Engineering, Southwest Jiaotong University, Chengdu, China, in 2017. From October 2015 to October 2017, he was a Visiting Ph.D. Student with the Faculty of Engineering, Information and Systems, University of Tsukuba, Japan. From 2018 to 2021, he was a Post-Doctoral Fellow with the School of Electrical & Computer Engineering, Ben-Gurion University of the Negev, Israel. In 2021, he joined Southwest Jiaotong University, where he currently hold a tenure-track position. His research interests include coding theory and sequence design.

Ying Miao received the D.Sci. degree in mathematics from Hiroshima University, Hiroshima, Japan, in 1997.

From 1989 to 1993, he was with Suzhou Institute of Silk Textile Technology, Suzhou, Jiangsu, China. From 1995 to 1997, he was a Research Fellow of Japan Society for the Promotion of Science. From 1997 to 1998, he was a Post-Doctoral Fellow with the Department of Computer Science, Concordia University, Montreal, QC, Canada. In 1998, he joined the University of Tsukuba, Tsukuba, Ibaraki, Japan, where he is currently a Full Professor with the Faculty of Engineering, Information and Systems. His current research interests include combinatorics, coding theory, and information security.

Dr. Miao received the 2001 Kirkman Medal from the Institute of Combinatorics and its Applications. He has been serving as an Associate Editor for several journals, such as IEEE TRANSACTIONS ON INFORMATION THEORY, *Graphs and Combinatorics*, and *Journal of Combinatorial Designs*.

Moshe Schwartz (Fellow, IEEE) received the B.A. (summa cum laude), M.Sc., and Ph.D. degrees from the Computer Science Department, Technion—Israel Institute of Technology, Haifa, Israel, in 1997, 1998, and 2004, respectively.

He is currently a Professor with the Department of Electrical and Computer Engineering, McMaster University, Canada, and a Professor with the School of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Israel (on a leave of absence). His research interests include algebraic coding, combinatorial structures, and digital sequences. He was a Fulbright Post-Doctoral Researcher with the Department of Electrical and Computer Engineering, University of California at San Diego, and a Post-Doctoral Researcher with the Department of Electrical Engineering, California Institute of Technology. While on sabbatical (2012–2014), he was a Visiting Scientist with Massachusetts Institute of Technology (MIT).

Prof. Schwartz received the 2009 IEEE Communications Society Best Paper Award in Signal Processing and Coding for Data Storage and the 2020 NVMW Persistent Impact Prize. He served as an Associate Editor for Coding Techniques and Coding Theory for IEEE TRANSACTIONS ON INFORMATION THEORY from 2014 to 2021. Since 2021, he has been serving as an Area Editor for Coding and Decoding for IEEE TRANSACTIONS ON INFORMATION THEORY. He has been an Editorial Board Member of the *Journal of Combinatorial Theory, Series A* since 2021.

Xiaohu Tang (Senior Member, IEEE) received the B.S. degree in applied mathematics from Northwest Polytechnic University, Xi'an, China, in 1992, the M.S. degree in applied mathematics from Sichuan University, Chengdu, China, in 1995, and the Ph.D. degree in electronic engineering from Southwest Jiaotong University, Chengdu, in 2001.

From 2003 to 2004, he was a Research Associate with the Department of Electrical and Electronic Engineering, The Hong Kong University of Science and Technology. From 2007 to 2008, he was a Visiting Professor with the University of Ulm, Germany. Since 2001, he has been with the School of Information Science and Technology, Southwest Jiaotong University, where he is currently a Professor. His research interests include coding theory, network security, distributed storage, and information processing for big data.

Dr. Tang was a recipient of the National Excellent Doctoral Dissertation Award in 2003, China, the Humboldt Research Fellowship in 2007, Germany, and the Outstanding Young Scientist Award by NSFC in 2013, China. He served as an Associate Editor for several journals, including IEEE TRANSACTIONS ON INFORMATION THEORY and *IEICE Transactions on Fundamentals* and served on a number of technical program committees of conferences.