ISIT'2013 1569716769

The Secrecy Capacity of Minimum Bandwidth Cooperative Regenerating Codes

O. Ozan Koyluoglu, Ankit S. Rawat, and Sriram Vishwanath
Department of Electrical and Computer Engineering
The University of Texas at Austin
Email: ozan@mail.utexas.edu, ankitsr@utexas.edu, sriram@ece.utexas.edu

***Abstract*—Regenerating codes enable trading off repair bandwidth for storage in distributed storage systems (DSS). Due to their distributed nature, these systems are intrinsically susceptible to attacks, and they may be susceptible to multiple node failures. This paper analyzes storage systems that employ cooperative regenerating codes that are robust to passive eavesdroppers, and proposes codes achieving the secrecy capacity for the minimum bandwidth cooperative regenerating point. The achievability results correspond to exact repair, and secure file size upper bounds are obtained using mincut analyses over a suitable secrecy graph representation of DSS. The main achievability argument is based on appropriate precoding of the data using MRD (Gabidulin) codes to eliminate any information leakage to the eavesdropper.**

***Index Terms*—Coding for distributed storage systems, minimum bandwidth cooperative regenerating codes, security.**

I. Introduction

Distributed storage systems (DSS) are designed to store data over a distributed network of nodes. Data to be stored is more than doubling every two years, and efficiency in storage and data recovery is particularly critical today. In addition to resilience against node failures, DSS require adequate mechanisms to endure adversarial attacks, such as one from eavesdroppers aiming to gain access to the stored data. Therefore, designing systems that meet security requirements while performing efficient repairs is of definite interest.

In [1], Dimakis et al. present a class of *regenerating codes*, which efficiently trade-off per node storage and repair bandwidth for single node repair in DSS. Explicit codes that achieve one of the two ends of the trade off between storage and repair bandwidth (namely, the points referred to as minimum bandwidth regeneration (MBR) and minimum storage regeneration (MSR)) have been investigated in several works recently (see, e.g., [2], [3] and references therein). In these works, while DSS can exhibit multiple simultaneous node failures, the repair process is sequential, i.e., one by one. However, it is desirable that the multiple failures be repaired simultaneously: As large-scale systems, DSS can have multiple failures, and some administrators (e.g., TotalRecall [4]), in order to render the entire process more efficient and less frequent, may choose to wait to initiate a repair process after a certain threshold (t) on the number of failures is reached. In such multiple failure scenarios, each new node replacing a failed one can still contact d remaining (surviving) nodes to download data for the repair process. In addition, replacement nodes, after downloading data from surviving nodes, can also exchange data within themselves to complete the repair process. This repair process is referred to as *cooperative repair* in [5]. Recent works, [6] and [7], provide a cut-set bound argument and derive the cooperative counterparts of the end points of the trade off region: minimum bandwidth cooperative regenerating (MBCR) point and the minimum storage cooperative regenerating (MSCR) point. Explicit code constructions for exact repair at the MBCR point are presented in [7] for $d = k$, and in [8] for $n = d + t$.

To secure DSS, cryptographic approaches like private-key cryptography are often logistically prohibitive, as the secret key distribution between each pair of nodes and its renewal are highly challenging. Compared to cryptographic approaches, information theoretic security (see, e.g., [9], [10]) offers secrecy guarantees even with infinite computational power at eavesdroppers without requiring the sharing and/or distribution of keys. The design of (information theoretically) secure DSS against eavesdropping attacks has been recently studied in [11], where the authors consider a passive eavesdropper model that observe the data stored on ℓ ($< k$) storage nodes for a DSS employing an MBR code. In another work, Shah et al. [12], utilizing product matrix codes, present coding schemes that achieves the bound on secrecy capacity at the MBR point.

In this paper, we focus on secure and cooperative regenerating codes for DSS at the minimum bandwidth regenerating point. This model generalizes the single node repair setting considered in earlier works to multiple node failures. In terms of security requirements, we consider an attacker having access to data stored on any ℓ number of nodes (a passive and colluding eavesdropper model analyzed in earlier works). Given such a model, we first derive an upper bound on the secrecy capacity for MBCR codes. Compared to earlier works, this bound is obtained from the min-cut analyses over the *secrecy* graph representation of DSS employing *cooperative repair*. Then, we propose a novel coding scheme which precodes the data with MRD (Gabidulin) codes, and then utilize the code proposed in [8] (having download links with minimum bandwidth for $d = n-t$). We show that the proposed scheme achieves secrecy *efficiently* as it achieves the proposed upper bound, characterizing the secrecy capacity for MBCR codes. The proposed code design allows for exact repair. And, the precoding argument can be seen as an extension of the seminal work of Shamir [13] to the context of DSS.

II. System Model and Preliminaries

Consider a DSS with n live nodes and a file $\mathbf{f}$ of size $\mathcal{M}$ over a field $\mathbb{F}$, size of which is to be finalized later. In order to store the file $\mathbf{f}$, it is divided into k blocks $(\mathbf{f}_1, \ldots, \mathbf{f}_k)$ each of size $\mathcal{M}/k$ ($\mathbf{f}_i \in \mathbb{F}_q^{\frac{\mathcal{M}}{k}}$). These k data blocks are encoded into n data blocks, $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$, each of length α over $\mathbb{F}_q$ ($\alpha \geq \frac{\mathcal{M}}{k}$). Given the codewords, node i in an n-node DSS stores encoded block $\mathbf{x}_i$. In this paper, we focus on "any k out of n" property, i.e., the content of any k nodes suffices to recover the file. The symbols stored at node i is represented by the vector $\mathbf{s}_i$, the symbols transmitted from node i to node j is denoted as $\mathbf{d}_{i,j}$, and the set $\mathbf{d}_j$ is used to denote all of the downloaded symbols to node j. DSS is initialized with $\mathbf{s}_i = \mathbf{x}_i$ for $i = [1:n]$. (For $a < b$, $[a:b]$ represents the set of numbers $\{a, a+1, \cdots, b\}$.)

A. Cooperative repair in DSS and information flow graph

In their seminal work [1], Dimakis et al. models the operation of DSS by a multicasting scenario over an information flow graph. In the cooperative setting (see Fig. 1), information flow graph consists of three types of nodes: 1) Source node (S): Contains original file $\mathbf{f}$ of size $\mathcal{M}$ symbols. 2) Storage nodes ($(x_i^{\text{in}}, x_i^{\text{co}}, x_i^{\text{out}})$): x^{in} is the sub-node having the connections from the live nodes, x^{co} is the sub-node having the connections from the nodes under repair in the same repair group, and x^{out} is the storage sub-node, which stores the data and is contacted by a data collector or other nodes under repair. 3) Data collector (DC): Each data collector (DC) contacts x^{out} sub-node of k live nodes (with edges each having ∞-link capacity).

Here, x^{in} is connected to x^{co} with a link of infinite capacity, x^{co} is connected to x^{out} with a link of capacity α. We represent cuts with a notation with bars as in $(x^{\text{in}}, x^{\text{co}}|x^{\text{out}})$, meaning the cut is passing through the link between x^{co} and x^{out}. (See Fig. 1.) The nodes on the right hand side of the cuts belong to DC side, represented by the set $\mathcal{D}$, whereas the nodes belonging to the left hand side of the cuts belong to $\mathcal{D}^c$, the source side. For a newcomer node, x_i^{in} is connected to x^{out} sub-nodes of d live nodes with links of capacity β symbols each, representing the data downloaded during node repair. This newcomer node also connects to x^{in} sub-nodes of $(t-1)$ nodes being repaired in the same group, each having a link capacity of β'. Hence, the total repair cost is given by $\gamma = d\beta + (t-1)\beta'$.

B. MBCR and MSCR points

For a given graph $\mathcal{G}$ and DCs DC_i, the file size can be bounded using the max flow-min cut theorem for multicasting utilized in network coding [1], [14].

Lemma 1 (Max flow-min cut theorem for multicasting)**.**

$$\mathcal{M} \leq \min_{\mathcal{G}} \min_{\text{DC}_\text{i}} \text{maxflow}(\text{S} \rightarrow \text{DC}_\text{i}, \mathcal{G}),$$

where $\text{flow}(\text{S} \rightarrow \text{DC}_\text{i}, \mathcal{G})$ *represents the flow from the source node* S *to* DC_i *over the graph* $\mathcal{G}$.

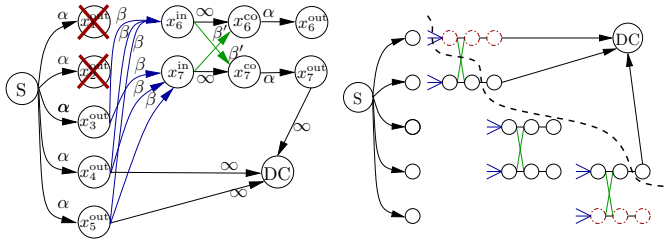

Fig. 1: Information flow graph of DSS with $n = 5$, $d = k = 3$, and $t = 2$. Left: After the failure of node 1 and node 2, the system cooperatively repairs these two nodes as node 6 and node 7 such that $x_6^{\text{out}} = x_1^{\text{out}}$ and $x_7^{\text{out}} = x_2^{\text{out}}$. Right: Multiple repair stages and a cut, represented by dotted line, are shown. The first repaired node has a cut of type $(|x^{\text{in}}, x^{\text{co}}, x^{\text{out}})$ and the second has a cut of type $(x^{\text{in}}, x^{\text{co}}|x^{\text{out}})$. Nodes that are being eavesdropped are indicated with dashed-dotted circles.

Therefore, $\mathcal{M}$ symbol long file can be delivered to a DC, only if the min cut is at least $\mathcal{M}$. Dimakis et al., [1], consider k successive node failures and evaluate the min-cut over possible graphs, and obtain a file size bound for $t = 1$ case. The codes that attain the bound are named as regenerating codes [1]. Using a similar approach, a file size bound in the cooperative setting can be obtained as follows [3], [6], [7]. (In the next section, we derive a secure file size bound using a similar min cut approach.)

$$\mathcal{M} \leq \sum_{i=0}^{\mu-1} u_i \min\left\{\alpha, \left(d - \sum_{j=0}^{i-1} u_j\right)\beta + (t-u_i)\beta'\right\}, \quad (1)$$

where $u_i \in [0:t]$ is the number of repaired nodes in repair group $i \in [0:\mu-1]$ that is connected to DC. Similar to the $t = 1$ case analyzed in [1], the cut of type $(x^{\text{in}}, x^{\text{co}}|x^{\text{out}})$ has a value of α. The cut of type $(|x^{\text{in}}, x^{\text{co}}, x^{\text{out}})$, on the other hand, has a value of $(t-u_i)\beta'$ due to the links coming from the nodes under repair that are not connected to DC and additional value of $(d - \sum_{j=0}^{i-1} u_j)\beta$ is due to the connections to the previously repaired live nodes that are not contacted by DC. (Here, we again subtract the values of the flows from the nodes already belonging to the DC side, $\mathcal{D}$.)

Note that, given a file size $\mathcal{M}$, there is an inherent trade off between storage per node α and *repair bandwidth* $\gamma \triangleq d\beta + (t-1)\beta'$. Two classes of codes that achieve two extreme points of this trade off are named as *minimum bandwidth cooperative regenerating (MBCR)* codes and *minimum storage cooperative regenerating (MSCR)* codes. The former is obtained by first finding the minimum possible γ and then finding the minimum α satisfying (1). MSCR point, on the other hand, is obtained by first choosing a minimum storage per node (i.e., $\alpha = \mathcal{M}/k$), and then minimizing γ satisfying the min cut (1). We depict these two trade off points, which are directly computable from (1), in Fig. 2. (See [6], [7] for a detailed derivation of these two points.) Note that, when $t = 1$, these points correspond to MBR/MSR points characterized in [1].

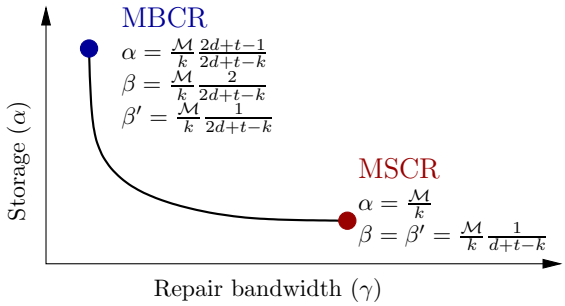

Fig. 2: Storage vs. repair bandwidth trade off for cooperative regenerating codes. The repair bandwidth is given by $\gamma = d\beta + (t-1)\beta'$.

C. Eavesdropper model

In the following, we provide the eavesdropper model together with a definition of achievability of secure file size. (Note that, as $\alpha = \gamma$ for MBCR codes, the eavesdropper does not gain additional knowledge regarding the file by observing the downloaded information after observing the stored content.)

Definition 2 (Security against an ℓ-eavesdropper)**.** *File size of $\mathcal{M}^s$ is secure against an ℓ-eavesdropper, if, for any sets $\mathcal{E}$ of size ℓ, $I(\mathbf{f}^s;\mathbf{e}) = 0$, where $\mathbf{f}^s$ is the secure file of size $\mathcal{M}^s$, which is first encoded to a data $\mathbf{f}$ of size $\mathcal{M}$ before storing on DSS, and $\mathbf{e}$ is the eavesdropper observation vector given by $\mathbf{e} \triangleq \{x_i^{\text{out}} : i \in \mathcal{E}\}$.*

The following lemma will be used in the sequel.

Lemma 3 (Secrecy Lemma)**.** *Consider a system with information symbols $\mathbf{u}$, random symbols $\mathbf{r}$ (independent of $\mathbf{u}$), and an eavesdropper with observations given by $\mathbf{e}$. If $H(\mathbf{e}) \leq H(\mathbf{r})$ and $H(\mathbf{r}|\mathbf{u},\mathbf{e}) = 0$, then $I(\mathbf{u};\mathbf{e})=0$.*

Proof: We have $I(\mathbf{u};\mathbf{e}) = H(\mathbf{e}) - H(\mathbf{e}|\mathbf{u}) \overset{(a)}{\leq} H(\mathbf{e}) - H(\mathbf{e}|\mathbf{u}) + H(\mathbf{e}|\mathbf{u},\mathbf{r}) \overset{(b)}{\leq} H(\mathbf{r}) - I(\mathbf{e};\mathbf{r}|\mathbf{u}) \overset{(c)}{=} H(\mathbf{r}|\mathbf{u},\mathbf{e}) \overset{(d)}{=} 0$, where (a) follows by non-negativity of $H(\mathbf{e}|\mathbf{u},\mathbf{r})$, (b) is due to $H(\mathbf{e}) \leq H(\mathbf{r})$, (c) follows as $\mathbf{r}$ and $\mathbf{u}$ are independent, (d) is due to $H(\mathbf{r}|\mathbf{u},\mathbf{e}) = 0$. (The proof above follows the classical techniques given in [10]. See also [12].) ■

III. Upper bound on secure file size for secure MBCR codes

Analysis of the cut-set bounds for cooperative regenerating codes are provided in [6], [7]. (See also [3], [5]. Here, we follow the notations of [3], [6].) Consider a scenario where groups of nodes are repaired sequentially. Let u_i denote the number of nodes in group i that are repaired in group i and contacted by the DC. We have $u_i \in [1:t], \forall i = 0,1,\cdots,\mu-1$, $\sum_{i=0}^{\mu-1} u_i = k$, where μ is the total number of groups that have been repaired. For deriving secure file size bound, the DC is assumed to contact only these k nodes that belong to one of these μ groups.

We consider two types of cuts: m_i number of nodes have the first cut type $(x^{\text{in}}, x^{\text{co}}|x^{\text{out}})$, and $u_i - m_i$ number of nodes have the second cut type $(|x^{\text{in}}, x^{\text{co}}, x^{\text{out}})$, $0 \leq i \leq \mu-1$. We consider ℓ number of colluding eavesdroppers, each observing the contents of different nodes. We denote the number of eavesdroppers on the nodes in the first cut type as $l^{i,1}$; and denote the number of eavesdroppers on the nodes in the second cut type as $l^{i,2}$ such that

$$l^{i,1} \leq m_i,\ \ l^{i,2} \leq u_i - m_i,\ \ \sum_{i=0}^{\mu-1}(l^{i,1} + l^{i,2}) = \ell.$$

Thus, for group i, due to the eavesdroppers, the nodes that belong to the first type can only add the value of $(m_i - l^{i,1})\alpha$ to the cut. The second type, on the other hand, consists of $u_i - m_i$ nodes, out of which $l^{i,2}$ of them are eavesdropped. As the data downloaded is equal to the data stored at MBCR point, the nodes that are eavesdropped do not add a value to the cut. The remaining $u_i - m_i - l^{i,2}$ number of nodes contact d live nodes, $\sum_{j=0}^{i-1} u_j$ number of these belong to the previous groups being repaired. In addition, these nodes contact $t-1$ nodes from the same repair group, out of which $u_i - m_i - 1$ number of nodes belong to $\mathcal{D}$. Accordingly, this cut-set bound is given by the following.

$$\mathcal{M}^s \leq \sum_{i=0}^{\mu-1} \left(\left(m_i - l^{i,1}\right)\alpha + \left(u_i - m_i - l^{i,2}\right) C_i\right), \qquad (2)$$

where $C_i = \left(d - \sum_{j=0}^{i-1} u_j\right)\beta + (t - u_i + m_i)\,\beta'$.

We consider two scenarios in (2): (i) $m_i = 0, l^{i,2} = l^i$, and (ii) $m_i = u_i, l^{i,1} = l^i$. Hence, we obtain,

$$\mathcal{M}^s \leq \sum_{i=0}^{\mu-1}(u_i - l^i)\min\left\{\alpha, \left(d - \sum_{j=0}^{i-1} u_j\right)\beta + (t-u_i)\beta'\right\}$$

Note that, at MBCR point, the nodes store what they download. Therefore, $\alpha = d\beta + (t-1)\beta'$. Utilizing this, consider having $\mu = k$, $u_i = 1$, $\forall i = 0,\cdots,k-1$ in equation above. Accordingly, we obtain

$$\mathcal{M}^s \leq \sum_{i=0}^{k-1}(1 - l^i)\left((d-i)\beta + (t-1)\beta'\right). \qquad (3)$$

Here, the minimum cut value corresponds to having $l^i = 1$ for $i = 0,1,\cdots,\ell-1$; and $l^i = 0$ otherwise. Hence, we get

$$\begin{aligned}\mathcal{M}^s &\leq \sum_{i=\ell}^{k-1}(d-i)\beta + (t-1)\beta' \\ &= \frac{(k-\ell)(2d-k-\ell+1)}{2}\beta + (k-\ell)(t-1)\beta'.\end{aligned}$$

The normalized values at the MBCR point are given by $\beta' = 1$, $\beta = 2$, $\alpha = \gamma = 2d + t - 1$, $\mathcal{M} = k(2d - k + t)$. Using this in the equation above, we get the following result.

Theorem 4. *Cooperative regenerating codes operating at the MBCR point with a secure file size of $\mathcal{M}^s$ satisfy*

$$\mathcal{M}^s \quad \leq \quad k(2d-k+t) - \ell(2d-\ell+t), \tag{4}$$

and the MBCR point is given by $\beta' = 1$, $\beta = 2$, $\alpha = \gamma = 2d+t-1$ for a file size of $\mathcal{M} = k(2d-k+t)$.

IV. Code construction for secure MBCR codes

We consider secrecy precoding of the data at hand before storing it on DSS nodes using an MBCR code. We establish this precoding with maximum rank distance (MRD) codes. In vector representation, assuming $m \geq N$, the norm of a vector $\mathbf{v} \in \mathbb{F}_{q^m}^N$ is the column rank of $\mathbf{v}$ over the base field $\mathbb{F}_q$, denoted by $Rk(\mathbf{v})$. (This is the maximum number of linearly independent coordinates of $\mathbf{v}$ over the base field $\mathbb{F}_q$, for a given basis of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$.) Rank distance between two vectors is defined by $d(\mathbf{v}_1, \mathbf{v}_2) = Rk(\mathbf{v}_1 - \mathbf{v}_2)$. (In matrix representation, this is equivalent to the rank of the difference of the two corresponding matrices of the vectors.) An $[N, K, D]$ MRD code over the extension field $\mathbb{F}_{q^m}$ achieving the maximum rank distance $D = N-K+1$ (for $m \geq N$) can be constructed with the following linearized polynomial. (This is referred to as the Gabidulin construction of MRD codes, or Gabidulin codes [15].)

$$f(g) = \sum_{i=0}^{K-1} u_i g^{[i]}, \tag{5}$$

where $[i] = q^i$, and $g, u_i \in \mathbb{F}_{q^m}$. Then, given N linearly independent elements over $\mathbb{F}_q$, $\{g_1, \cdots, g_N\}$ with $g_j \in \mathbb{F}_{q^m}$, the codewords for a given set of K message (information) symbols, $u_i \in \mathbb{F}_{q^m}, i = [0 : K-1]$, are obtained by $x_j = f(g_j) = \sum_{i=0}^{K-1} u_i g_j^{[i]}$ for $j = [1 : N]$. (With generator matrix representation, we have $\mathbf{x} = \mathbf{uG}$, where $\mathbf{G} = [g_1, \cdots, g_N; \cdots; g_1^{[K-1]}, \cdots, g_N^{[K-1]}]$.) Note that the linearized polynomial satisfies $f(a_1 g_1 + a_2 g_2) = a_1 f(g_1) + a_2 f(g_2)$, for a given $a_1, a_2 \in \mathbb{F}_q$ and $g_1, g_2 \in \mathbb{F}_{q^m}$, and this will be utilized in the following.

Consider now the MBCR point given by $\mathcal{M} = k(2d-k+t)$, $\beta' = 1, \beta = 2, \alpha = \gamma = 2d+t-1, \mathcal{M}^s = k(2d-k+t) - \ell(2d-\ell+t)$, and $n = d+t$. We use MRD codes with $N = K = \mathcal{M}$; hence, the rank distance bound $D \leq N - K + 1$ is saturated at $D = 1$. Accordingly, we utilize $[\mathcal{M}, \mathcal{M}, 1]$ Gabidulin codes over $\mathbb{F}_{q^m}$, which maps length $\mathcal{M}$ vectors (each element of it being in $\mathbb{F}_{q^m}$) to length $\mathcal{M}$ codewords in $\mathbb{F}_{q^m}^{\mathcal{M}}$ (with $m \geq \mathcal{M}$). The coefficients of the underlying linearized polynomial ($f(g)$) are chosen by $\mathcal{M} - \mathcal{M}^s$ random symbols denoted by $\mathbf{r} \in \mathbb{F}_{q^m}^{\mathcal{M}-\mathcal{M}^s}$ and $\mathcal{M}^s$ secure data symbols denoted by $\mathbf{u} \in \mathbb{F}_{q^m}^{\mathcal{M}^s}$. The corresponding polynomial $f(g)$ is evaluated at $\mathcal{M}$ points $\{g_1, \ldots, g_{\mathcal{M}}\}$, which are linearly independent over $\mathbb{F}_q$. We denote these as $x_j = f(g_j)$ for $j = 1, \cdots, \mathcal{M}$. This finalizes the secrecy precoding step.

The second encoding step is based on the encoding scheme for cooperative repair proposed in [8]. (Here, we summarize file recovery and node repair processes for the case of MRD precoding, and provide the proof of security.) Split the $\mathcal{M}$ symbols into two parts a) x_1 to x_{nk}, and b) x_{nk+1} to $x_{nk+k(d-k)}$. (Note that $n = d+t$ and $\mathcal{M} = nk + k(d-k)$.) The first part is divided into n groups of k symbols, and stored in n nodes. Here, node i stores $x_{(i-1)k+1}$ to x_{ik}. The second part is divided into $d-k$ groups of k symbols. These symbols are encoded with an (n, k) MDS code, and stored on n nodes. In particular, $\{y_{j,1}, \ldots, y_{j,n}\}$ are generated from symbols $\{x_{nk+(j-1)k+1}, \ldots, x_{nk+jk}\}$, and $y_{j,i}$ is stored at node i, for $j = 1, \cdots, d-k$. Node i, having stored $\{x_{(i-1)k+1}, \ldots, x_{ik}, y_{1,i}, \ldots y_{d-k,i}\}$, which is referred to as the primary data of node i, encodes these symbols using an $(n-1, d)$ MDS code that has a generator matrix given by a (generalized) Cauchy matrix Φ of size $d \times (n-1)$. (This choice of Φ ensures that $[\mathbf{I}_d \ \Phi]$ is generator matrix for an $(n+d-1, d)$ MDS code [16].) These $n-1$ symbols are stored in every other node one-by-one. We denote the encoded primary data of node i that is stored in node $j \neq i$ as $z_{j,i}$. We call these as the secondary data. This procedure is repeated for every node, so that each node i stores $\{x_{(i-1)k+1}, \ldots, x_{ik}, y_{1,i}, \ldots, y_{d-k,i}, z_{i,1}, \ldots, z_{i,i-1}, z_{i,i+1}, \ldots, z_{i,n}\}$, and hence total number of symbols stored at each node is $k + (d-k) + (n-1) = d + n - 1 = 2d + t - 1 = \alpha$.

File recovery at DC: DC connects to any k nodes, without loss of generality we assume the first k nodes. From $y_{j,1:k}$, DC can obtain $x_{nk+(j-1)k+1}, \cdots, x_{nk+jk}$, for each $j = [1 : d-k]$. It can re-encode this into $y_{j,1:n}$ using the MDS code, and obtain the other y symbols at the remaining nodes. Then, for each $i \in [k+1 : n]$, DC can use the MDS property of $[\mathbf{I}_d \ \Phi]$, to obtain $x_{(i-1)k+1}, \cdots, x_{ik}$ symbols of node i from the k secondary data symbols of the contacted nodes, i.e., $z_{j,i}$ for $j = [1 : k]$, and additional $d-k$ symbols, $y_{j,i}$ for $j = [1 : d-k]$. Having obtained $x_1, \cdots, x_{\mathcal{M}}$, DC can perform interpolation to solve for both data and random coefficients.

Node repair: Assume that the first t nodes fail. From the secondary data stored in the remaining $d = n - t$ nodes, $z_{t+1,i}, \cdots, z_{n,i}$, one can recover $x_{(i-1)k+1}, \cdots, x_{ik}$ and $y_{1,i}, \cdots, y_{d-k,i}$; for node $i = 1, \cdots, t$. (This corresponds to sending 1 symbol from each of d nodes to each of the t nodes.) Then, to recover the secondary data stored at each node under repair, say for the node $j = 1, \cdots, t$, every other node, i.e., nodes $i \neq j$, including the nodes under repair, computes and sends its corresponding encoded primary data, i.e., $z_{j,i}$, to node j. (This corresponds to sending 1 symbol from each node to each of the t nodes.) This achieves $\beta = 2$ and $\beta' = 1$ symbols for the repair procedure.

Security: Consider that the eavesdropper is observing the first ℓ nodes. Due to the code construction, the symbols in the sets $\mathcal{X} = \{x_1, \ldots, x_{\ell k}\}$, $\mathcal{Y} = \{y_{1,1}, \ldots, y_{d-k,1}, \cdots, y_{1,\ell}, \ldots, y_{d-k,\ell}\}$, $\mathcal{Z} = \{z_{j,i}$ for $j = 1, \ldots, \ell$, and $i = \ell+1, \cdots, n\}$ correspond to linearly independent evaluation points. (Note that, the symbols $\{z_{j,i}\}$ for $j = 1, \cdots, \ell$; $i = 1, \cdots, \ell$; $j \neq i$, are linear combinations of the symbols in $\mathcal{X} \cup \mathcal{Y}$.) Due to the linearized property of the code, the eavesdropper observing $\ell\alpha = \ell(2d+t-1)$ symbols, has evaluation of polynomial $f(\cdot)$ at $\ell(2d+t-\ell)$ linearly

independent points. Using the data symbols, together with interpolation from these $\ell(2d+t-\ell)$ symbols, the eavesdropper can solve for $\ell(2d+t-\ell)$ random symbols. Therefore, denoting the eavesdroppers' observation as $\mathbf{e}$, we have $H(\mathbf{r}|\mathbf{e},\mathbf{u})=0$. As, $H(\mathbf{e})=H(\mathbf{r})$, from Lemma 3, we have $I(\mathbf{u};\mathbf{e})=0$.

From above and Theorem 4, we obtain the following result.

Theorem 5. *The secrecy capacity at MBCR point for a file size of $\mathcal{M}=k(2d-k+t)$ is given by $\mathcal{M}^s=k(2d-k+t)-\ell(2d-\ell+t)$, if $n=d+t$.*

V. Discussion and concluding remarks

A. Secure MBCR code examples

Cooperative regenerating codes has a repair bandwidth given by $\gamma=d\beta+(t-1)\beta'$. Here, we analyze $\frac{\gamma}{\mathcal{M}^s}$, the ratio of repair bandwidth to the secure file size, referred to as the normalized repair bandwidth (NRBW). The parameters of Theorem 5 are given in the following table. ($\ell=0$ case corresponds to the systems without security constraints. $t=1$ case corresponds to non-cooperative case.)

TABLE I: NRBW for $n=4,5$, $d\geq k$, $d+t=n$.

n	k	l	t	d	$\beta/\mathcal{M}^s$	$\beta'/\mathcal{M}^s$	$\gamma/\mathcal{M}^s$	$\mathcal{M}$	$\mathcal{M}^s$
4	2	0	1	3	0.2000	0.1000	0.6000	10	10
4	2	0	2	2	0.2500	0.1250	**0.6250**	8	8
4	2	1	1	3	0.5000	0.2500	1.5000	10	4
4	2	1	2	2	0.6667	0.3333	**1.6667**	8	3
4	3	0	1	3	0.1667	0.0833	0.5000	12	12
4	3	1	1	3	0.3333	0.1667	1.0000	12	6
4	3	2	1	3	1.0000	0.5000	3.0000	12	2
5	2	0	1	4	0.1429	0.0714	0.5714	14	14
5	2	0	2	3	0.1667	0.0833	**0.5833**	12	12
5	2	0	3	2	0.2000	0.1000	**0.6000**	10	10
5	2	1	1	4	0.3333	0.1667	1.3333	14	6
5	2	1	2	3	0.4000	0.2000	**1.4000**	12	5
5	2	1	3	2	0.5000	0.2500	**1.5000**	10	4

A direct calculation shows that NRBW for the case $t>1$ is strictly greater than that of $t=1$ when $n=d+t$ for any $l<k$. (In the table above, bold-red font indicates higher NRBW compared to $(t=1,d=n-1)$ case.) This in turn means that one may not deliberately delay the repairs to achieve a better performance than that of single failure-repair if d is chosen such that $n=d+t$ for a given (n,t). However, if the downloads within the cooperative group are less costly compared to the downloads from the live nodes, then delaying repairs would be beneficial in reducing the total cost. (Note that, the bandwidth for $t>1$ case can be smaller than $t=1$ case for $d<n-t$ [6].)

B. Concluding remarks

In this paper, we considered secure cooperative regenerating codes for DSS. We characterized the secrecy capacity (against a passive eavesdropper observing contents of ℓ storage nodes) at the minimum bandwidth cooperative regeneration (MBCR) point. The code construction proposed in this paper utilizes maximum rank distance (MRD), in particular Gabidulin, codes as a precoding step. Indeed, the properties of linearized polynomials are essential in obtaining the results. The properties of such codes have been utilized to achieve secrecy in different contexts in the literature: [17] for error control in network coding, [18] for resilience against active eavesdroppers, and [19] for security in locally repairable codes. The results in this paper shows that, in addition to previous works, MRD codes have useful properties in constructing secure cooperative DSS.

As a final note, we point out that the code construction presented in this paper has the requirement of $d=n-t$. However, for practical systems, it may not be possible that a failed node connects to all the remaining nodes. In an extended version of this work [20], utilizing the codes presented in [21], we proposed secure MBCR codes for $d<n-t$.

References

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

[2] A. G. Dimakis, K. Ramchandran, Y. Wu, and C. Suh, "A Survey on Network Codes for Distributed Storage," *Proceedings of the IEEE*, vol. 99, no. 3, pp. 476–489, Mar. 2011.

[3] F. Oggier and A. Datta, "Coding techniques for repairability in networked distributed storage systems," NTU, Tech. Rep., Sep. 2012.

[4] R. Bhagwan, K. Tati, Y.-C. Cheng, S. Savage, and G. M. Voelker, "Total recall: system support for automated availability management," in *Proc. NSDI'04*, Berkeley, CA, USA, Mar. 2004.

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

[6] A.-M. Kermarrec, N. Le Scouarnec, and G. Straub, "Repairing multiple failures with coordinated and adaptive regenerating codes," in *Proc. 2011 International Symposium on Network Coding (NetCod 2001)*, Jul. 2011.

[7] K. W. Shum and Y. Hu, "Cooperative regenerating codes," *CoRR*, vol. abs/1207.6762, Jul. 2012.

[8] S. Jiekak and N. Le Scouarnec, "CROSS-MBCR: Exact minimum bandwith coordinated regenerating codes," *CoRR*, vol. abs/1207.0854, Jul. 2012.

[9] C. E. Shannon, "Communication theory of secrecy systems," *The Bell System Technical Journal*, vol. 28, pp. 656–715, 1949.

[10] A. Wyner, "The wire-tap channel," *The Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[11] S. Pawar, S. El Rouayheb, and K. Ramchandran, "Securing dynamic distributed storage systems against eavesdropping and adversarial attacks," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6734 –6753, Oct. 2011.

[12] N. B. Shah, K. V. Rashmi, and P. V. Kumar, "Information-theoretically secure regenerating codes for distributed storage," in *Proc. 2011 IEEE Global Telecom. Conference (GLOBECOM 2011)*, Dec. 2011.

[13] A. Shamir, "How to share a secret," *Communications of the ACM*, vol. 22, no. 11, pp. 612–613, Nov. 1979.

[14] T. Ho, M. Medard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast," *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4413–4430, Oct. 2006.

[15] E. M. Gabidulin, "Theory of codes with maximum rank distance," *Probl. Peredachi Inf.*, vol. 21, no. 1, pp. 3–16, Jul. 1985.

[16] R. M. Roth and G. Seroussi, "On generator matrices of MDS codes," *IEEE Trans. Inf. Theory*, vol. 31, no. 6, pp. 826–830, Nov. 1985.

[17] D. Silva, F. R. Kschischang, and R. Koetter, "A rank-metric approach to error control in random network coding," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 3951–3967, Sep. 2008.

[18] N. Silberstein, A. S. Rawat, and S. Vishwanath, "Error resilience in distributed storage via rank-metric codes," in *Proc. 50th Allerton Conference on Communication, Control, and Computing*, Oct. 2012.

[19] A. S. Rawat, O. O. Koyluoglu, N. Silberstein, and S. Vishwanath, "Optimal locally repairable and secure codes for distributed storage systems," *CoRR*, vol. abs/1210.6954, Oct. 2012.

[20] O. O. Koyluoglu, A. S. Rawat, and S. Vishwanath, "Secure cooperative regenerating codes for distributed storage systems," *CoRR*, vol. abs/1210.3664, Oct. 2012.

[21] A. Wang and Z. Zhang, "Exact cooperative regenerating codes with minimum-repair-bandwidth for distributed storage," *CoRR*, vol. abs/1207.0879, Jul. 2012.