Security in Locally Repairable Storage

Abhishek Agarwal and Arya Mazumdar *Senior Member, IEEE*

***Abstract*—In this paper we extend the notion of *locally repairable* codes to *secret sharing* schemes. The main problem that we consider is to find optimal ways to distribute shares of a secret among a set of storage-nodes (participants) such that the content of each node (share) can be recovered by using contents of only few other nodes, and at the same time the secret can be reconstructed by only some allowable subsets of nodes. As a special case, an eavesdropper observing some set of specific nodes (such as less than certain number of nodes) does not get any information. In other words, we propose to study a locally repairable distributed storage system that is secure against a *passive eavesdropper* that can observe some subsets of nodes.**

We provide a number of results related to such systems including upper-bounds and achievability results on the number of bits that can be securely stored with these constraints. In particular, we provide conditions under which a locally repairable code can be turned into a secret sharing scheme and extend the results of secure repairable storage to cooperative repair and storage on networks. Additionally, we consider perfect secret sharing schemes over general access structures under locality constraints and give an example of a perfect secret sharing scheme that can have small locality. Lastly, we provide a lower bound on the size of a share compared to the size of the secret that shows how locality affects the sizes of shares in a perfect scheme.

I. INTRODUCTION

Secret sharing schemes were proposed by Shamir and Blakley [3], [22] to provide security against an eavesdropper with unbounded computational capability. Consider the secret as a realization of a (uniform) random vector $\boldsymbol{S}$ over some support. Define $[n] := \{1, 2, \ldots, n\}$ and let 2^A denote the power set for set A. Suppose that shares of the secret are to be distributed among n participants (storage nodes) such that a set of shares belonging to $\mathcal{A}_s \subseteq 2^{[n]}$, is able to determine the secret. $\mathcal{A}_s$ is called the access structure of the secret sharing scheme. Denote the random variable corresponding to the share of a participant (or node) $i \in [n]$ by C_i and let $\boldsymbol{C} = (C_1 C_2 \ldots C_n)$. Let $\boldsymbol{x}_A$ denote the projection of the vector $\boldsymbol{x} \in \mathbb{F}^n$ to the co-ordinates in $A \subseteq [n]$. For a singleton set $A = \{i\}$ let $\boldsymbol{x}_i := \boldsymbol{x}_{\{i\}}$. A secure scheme has the property that a subset of shares in the block-list $\mathcal{B}_s \subseteq 2^{[n]}$ are unable to determine anything about the secret. Thus, $H(\boldsymbol{S}|\boldsymbol{C}_B) = H(\boldsymbol{S})$ for any $B \in \mathcal{B}_s$ and $H(\boldsymbol{S}|\boldsymbol{C}_A) = 0$ for any $A \in \mathcal{A}_s$, where $H(\cdot)$ denotes the entropy[1]. For a standard *monotone* secret sharing scheme the classes $\mathcal{A}_s$ and $\mathcal{B}_s$ must have the following properties,

$$A' \supseteq A, A \in \mathcal{A}_s \implies A' \in \mathcal{A}_s$$
$$B' \subseteq B, B \in \mathcal{B}_s \implies B' \in \mathcal{B}_s$$
$$\text{and}$$
$$\mathcal{B}_s \subseteq 2^{[n]} \setminus \mathcal{A}_s.$$

For a *perfect* secret sharing scheme we have the above monotone property and $\mathcal{B}_s = 2^{[n]} \setminus \mathcal{A}_s$. Perfect schemes for access structures of the form $\mathcal{A}_s = \{A \subseteq [n] : |A| \geq m\}$ are called *threshold* secret sharing schemes. We refer to [2] for a comprehensive survey of secret sharing schemes.

A convenient property of schemes that need to store data in a distributed storage system is local repairability [8] i.e. any storage node can be repaired by accessing a small subset of other nodes, much smaller than is required for decoding the complete data. Error-correcting codes with the local repair property – locally repairable codes (LRC) – have been the center of a lot of research activities lately [4], [8], [16], [24]. Consider an n length code over a q-ary alphabet, $\mathcal{C} \subseteq \mathbb{F}_q^n$ of size $|\mathcal{C}| = q^k$. The code is said to have *locality* r, if for every i, $1 \leq i \leq n$, there exists a set $\mathcal{R}_i \subseteq [n] \setminus \{i\}$ with $|\mathcal{R}_i| \leq r$ such that for any two codewords $\boldsymbol{u}, \boldsymbol{u}' \in \mathcal{C}$ satisfying $\boldsymbol{u}_i \neq \boldsymbol{u}'_i$, we have $\boldsymbol{u}_{\mathcal{R}_i} \neq \boldsymbol{u}'_{\mathcal{R}_i}$. In a code with locality r, any symbol of a codeword can be deduced by reading only at most r other symbols of the codeword. For application in distributed storage, the code is further required to have a large *minimum distance* d, since that helps recovery in the event of a catastrophic failures (i.e., up to $d-1$ node failures). It is known that [8] for such a code,

$$d \leq n - k - \lceil k/r \rceil + 2, \tag{1}$$

which is also achievable [16], [24]. A q-ary code of length n, size q^k and locality r will be called an $(n, k, r)_q$-optimal LRC if it's minimum distance satisfies (1) with equality.

Security in distributed storage has recently been considered in a number of papers, for example [9], [17], [20], [25] and references therein. In these papers the main objective is to secure stored or downloaded data against an adversary. Threshold secret sharing protocols over a network under some communication constraint has been considered in [21]. Problems most closely related to this paper perhaps appear in [18] where a version of threshold secret sharing scheme with locality has been studied. Motivated by the above applications in distributed storage, we analyze secret sharing schemes with different access structures such that shares of each participant/node can be repaired with locality r.

Abhishek Agarwal is with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455, email: `abhiag@umn.edu`.

Arya Mazumdar is with the College of Information and Computer Science, University of Massachusetts, Amherst, MA 01003, email: `arya@cs.umass.edu`. Part of this work has been done when the author was at University of Minnesota.

This work was supported in part by NSF CCF 1318093, CCF 1453121, CCF 1642658. A preliminary version of this work was presented in the IEEE Information Theory Workshop, Jerusalem, Israel, 2015.

---

[1] The unit of entropy in this paper is q-ary, where q is an integer that will be clear from context.

A. Contributions and organization

Our contributions in this paper are summarized in the following list.

1) *Distributed storage.* We provide bounds and achievability results for a locally repairable scheme for access structure and block-list, $\mathcal{A}_s = \{A \subseteq [n] : |A| \geq m\}$ and $\mathcal{B}_s = \{B \subseteq [n] : |B| \leq \ell\}$, respectively. As evident from definition 1, this access and block structures model a simple distributed storage scenario. We assume that the shares of the secrets are locally recoverable and at the same time an adversary observing up to ℓ shares does not get any information. A more general version of this model that also considers repair bandwidth as a parameter appears in [18]. In section II we also address the conditions under which a locally repairable error-correcting code can be converted into a secret sharing scheme with the above access structure.
   Comparison of this part with results of [18]: In [18], bounds on secrecy capacity for regenerating and locally recoverable codes have been derived using information theoretic inequalities, and achievability of these bounds using schemes that require Gabidulin precoding technique has been shown.
   Our method to prove the converse result is different from that used in [18]. One advantage of our technique for the bound in section II is that it can be easily applied to cooperative repair (section III) and repairable codes on graphs (section IV).
   We provide a random coding argument using network flow graphs to show the existence of an achievability scheme for the bound, and also adapt the method of [18] for more general scenarios mentioned above (i.e., cooperative repair and repairable codes on graphs). For these scenarios, we use lemma 6 and Gabidulin precoding to construct transformations to form secure schemes from existing non-secure locally repairable codes.
2) *Maximal recoverability.* The Gabidulin precoding described above can be used to construct optimal codes but requires an exponentially large (in n) alphabet size. A simple construction of secret sharing schemes from LRCs is provided in eq. (14). We specify in lemma 6 the additional constraints that an optimal LRC would have to satisfy to be able to construct optimal secret sharing schemes in this method. This shows that to construct an optimal secure scheme with small share size we essentially need a *maximally recoverable code* over small alphabet (see theorem 8).
3) *Perfect secret sharing with small locality.* In section V, we consider perfect secret sharing schemes over general access structures under locality constraints. While we show that for threshold secret sharing schemes, there cannot exist any non-trivial local repairability, we give an example of a perfect secret sharing scheme that can have small locality.
4) *Lower bound on the size of shares in terms of the size of the secret.* Furthering the result of [5] to locally repairable schemes we provide an analogous lower-bound on the size of a share compared to the size of the secret. We further show how locality effects the sizes of shares in a perfect scheme as they relate to the size of the secret. These results are presented in section V (see theorem 14).
5) *Extension.* We extend the notion of security to cooperative local repair [19] where a Distributed Storage System can deal with simultaneous multiple node failures. We provide upper-bounds on the secrecy capacity and construct achievable schemes for this scenario in section III.
6) *Extension.* A different and practical generalization for secret sharing scheme is made in which the Distributed Storage System is represented by a graph $\mathcal{G}$ such that a node can only connect to its neighbors in $\mathcal{G}$ for repair. This scenario has been considered in section IV.

II. A SECRET-SHARING SCHEME FOR DISTRIBUTED STORAGE

We start this section by formally defining a secret sharing scheme for a particular, common access structure and block-list: $\mathcal{A}_s = \{A \subseteq [n] : |A| \geq m\}$ and $\mathcal{B}_s = \{B \subseteq [n] : |B| \leq \ell\}$. For a code $\mathcal{C} \subset \mathbb{F}_q^n$ and set $I \subset [n]$ define $\mathcal{C}_I := \{\boldsymbol{x}_I \in \mathbb{F}_q^{|I|} : \boldsymbol{x} \in \mathcal{C}\}$.

Definition 1. *An $(n, k, \ell, m, r)_q$-secret sharing scheme consists of a randomized encoder f that maps a uniform secret $\boldsymbol{S} \in \mathbb{F}_q^k$ randomly to $\boldsymbol{C} = f(\boldsymbol{S}) \in \mathbb{F}_q^n$, and must have the following three properties.*

1) *(Recovery) Given any m symbols of $\boldsymbol{C}$, the secret $\boldsymbol{S}$ is completely determined. This guarantees that the secret is recoverable even with the loss of any $n - m$ shares.*
$$H(\boldsymbol{S}|\boldsymbol{C}_I) = 0, \ \forall I \subseteq [n], |I| = m \tag{2}$$
2) *(Security) Any set of ℓ shares of $\boldsymbol{C}$ does not reveal anything about the secret.*
$$H(\boldsymbol{S}|\boldsymbol{C}_J) = H(\boldsymbol{S}), \ \forall J \subseteq [n], |J| = \ell \tag{3}$$
   A scheme satisfying this condition is called ℓ-secure. An eavesdropper that can observe ℓ nodes is called an ℓ-strength eavesdropper.
3) *(Locality) For any share, there exist at most r other shares that completely determine this. For all i, there exists $\mathcal{R}_i \subseteq [n] \setminus \{i\} : |\mathcal{R}_i| \leq r$, such that*
$$H(\boldsymbol{C}_i|\boldsymbol{C}_{\mathcal{R}_i}) = 0 \tag{4}$$
   $\mathcal{R}_i$ is called the recovery set of share i.

The maximum amount of secret that can be stored as a function of n, ℓ, m and r is called the capacity of the secret sharing scheme and in the following we provide exact characterization of this quantity. We can define the security condition above in a modified way where the eavesdropper is allowed to see any set $J \subseteq [n]$ of shares and we calculate the amount of information revealed, i.e. $I(\boldsymbol{S}; \boldsymbol{C}_J)$, in terms of $n, k, |J|, m$ and r in an optimal scheme. This extension is easy from our result and somewhat summarized in corollary 4.

Note that, for locally repairable schemes with no security requirement i.e. $\ell = 0$ the following lower-bound on m is apparent from (1),
$$m \geq k + \lceil k/r \rceil - 1, \tag{5}$$

This lower bound follows from the definition of the minimum distance of a code $d = n - m + 1$. In the subsequent, we provide the fundamental limit on secrecy capacity and constructions achieving that limit.

As mentioned in the introduction, a generalized version of this type of secret-sharing scheme that include repair-bandwidth and other parameters was studied in [18]. Our theorems 2 and 5 can be obtained as a consequence of results of that paper. We still provide different proofs of these results as the concepts introduced will be useful for later developments.

A. Bounds

Let us first prove an immediate and naive upper bound on the capacity of a locally repairable secret sharing scheme that follows as a consequence of Eq. (5).

Proposition 1. *For any $(n, k, \ell, m, r)_q$-secret sharing scheme,*

$$k \leq m - \ell - \left\lfloor \frac{m - \ell}{r + 1} \right\rfloor$$

Proof: Consider the randomized encoding f of any (n, k, ℓ, m, r)-secret sharing scheme. For any secret $\boldsymbol{s} \in \mathbb{F}_q^k$, define the support of the map $f(\boldsymbol{s})$ to be $\operatorname{supp}(f(\boldsymbol{s})) = \{\boldsymbol{x} \in \mathbb{F}_q^n : \Pr(f(\boldsymbol{s}) = \boldsymbol{x}) \neq 0\}$. Clearly for any pair $\boldsymbol{s}, \boldsymbol{s}' \in \mathbb{F}_q^k$ $\boldsymbol{s} \neq \boldsymbol{s}'$, $\operatorname{supp}(f(\boldsymbol{s})) \cap \operatorname{supp}(f(\boldsymbol{s}')) = \emptyset$.

Suppose, for some $\boldsymbol{s} \in \mathbb{F}_q^k$, $\boldsymbol{x} \in \operatorname{supp}(f(\boldsymbol{s}))$. Let $I \subseteq [n]$ and $|I| = \ell$. Note that, for each $\boldsymbol{s}' \in \mathbb{F}_q^k \setminus \boldsymbol{s}$, there must exist $\boldsymbol{z} \in \operatorname{supp}(f(\boldsymbol{s}'))$ such that $\boldsymbol{z}_I = \boldsymbol{x}_I$ (from the Security property). Let $\mathcal{C} \subseteq \{\boldsymbol{z} \in \operatorname{supp}(f(\boldsymbol{s}')) : \boldsymbol{s}' \in \mathbb{F}_q^k \text{ and } \boldsymbol{z}_I = \boldsymbol{x}_I\}$ such that $|\mathcal{C} \cap \operatorname{supp}(f(\boldsymbol{s}'))| = 1 \forall \boldsymbol{s}' \in \mathbb{F}_q^k$. We have $\mathcal{C} \subseteq \mathbb{F}_q^n$ and $|\mathcal{C}| = q^k$. Moreover, from the Recovery property, any m coordinates of a vector in $\mathcal{C}$ must be unique, which implies $\mathcal{C}$ has minimum distance at least $n - m + 1$.

Since $\{f(\boldsymbol{s}) : \boldsymbol{s} \in \mathbb{F}_q^k\}$ has locality r any set $\mathcal{C} \subset \{f(\boldsymbol{s}) : \boldsymbol{s} \in \mathbb{F}_q^k\}$ must have locality r. Since, all the codewords in $\mathcal{C}$ have fixed value on the co-ordinates I, $\mathcal{C}_{[n]\setminus I} \in \mathbb{F}_q^{n-\ell}$ must be a code of length $n - \ell$ and locality r. Moreover, $\mathcal{C}_{[n]\setminus I}$ has minimum distance at least $n - m + 1$ (same as $\mathcal{C}$). Now from eq. (1) we have,

$$\begin{aligned} & n - m + 1 \leq (n - \ell) - k - \lceil k/r \rceil + 2 \\ \iff \quad & k + \lceil k/r \rceil - 1 \leq m - \ell && \text{(6a)} \\ \iff \quad & k \leq m - \ell - \left\lfloor \frac{m - \ell}{r + 1} \right\rfloor && \text{(6b)} \end{aligned}$$

where eq. (6b) follows by replacing both sides of eq. (6a) by $Incr_0(k + \lceil k/r \rceil - 1)$ and $Incr_0(m - \ell)$ respectively, where $Incr_0(.)$ denotes the increasing function $Incr_0(x) := x - \left\lfloor \frac{x}{r+1} \right\rfloor$. ■

This naive bound in eq. (6a) is not the best possible: it can be further improved to

$$k + \ell + \left\lceil \frac{k + \ell}{r} \right\rceil - 1 \leq m. \tag{7}$$

To prove (7), instead of trying to use eq. (1) as a black-box, we follow its proof method [4], [8].

Theorem 2. *Any $(n, k, \ell, m, r)_q$-secret sharing scheme must satisfy,*

$$k + \ell \leq m - \left\lfloor \frac{m}{r + 1} \right\rfloor. \tag{8}$$

The upper-bound in eq. (8) can also be obtained from [18, Theorem 33] where the authors use a different method. It should be noted that eq. (8) is equivalent to eq. (7). We see that eq. (7) $\implies$ eq. (8) by replacing both sides in eq. (8) by the increasing function $Incr_0(x) := x - \lfloor x/(r+1) \rfloor$. Similarly eq. (8) $\implies$ eq. (7) by replacing each side with the increasing function $Incr_1(x) := x + \lceil x/r \rceil - 1$. This follows because of the following fact,

Claim 3. *For $x, y, r \in \mathbb{Z}^+$,*

$$y = x + \left\lceil \frac{x}{r} \right\rceil - 1 \iff x = y - \left\lfloor \frac{y}{r + 1} \right\rfloor \tag{9}$$

Proof: Let $x = qr + w,\ w < r$. Then, we have.

$$\begin{aligned} & x + \left\lceil \frac{x}{r} \right\rceil - 1 - \left\lfloor \frac{x + \lceil \frac{x}{r} \rceil - 1}{r + 1} \right\rfloor && \text{(10a)} \\ & = x + q + \left\lceil \frac{w}{r} \right\rceil - 1 - \left\lfloor \frac{qr + w + q + \lceil \frac{w}{r} \rceil - 1}{r + 1} \right\rfloor && \text{(10b)} \\ & = x + \left\lceil \frac{w}{r} \right\rceil - \left\lfloor \frac{w + \lceil \frac{w}{r} \rceil - 1}{r + 1} \right\rfloor - 1 && \text{(10c)} \\ & = x && \text{(10d)} \end{aligned}$$

where eq. (10d) follows since $\lceil \frac{w}{r} \rceil - \left\lfloor \frac{w + \lceil \frac{w}{r} \rceil - 1}{r+1} \right\rfloor - 1 = 0$ for $w \in [0, r-1]$. Now, substituting $y = x + \lceil \frac{x}{r} \rceil - 1$ in eq. (10a) we have, eq. (9). ■

Proof of theorem 2: Let $\Lambda_i = \{\mathcal{R}_i \cup \{i\}\}$. Recall that we can recover the secret $\boldsymbol{S}$ from any m symbols in the n length word $f(\boldsymbol{S}) = \boldsymbol{C}$. We construct an m-subset $\mathcal{M} \subseteq [n]$ such that $|\{i : \Lambda_i \subseteq \mathcal{M}\}|$ is maximized. Suppose, $\mathcal{M}' = \bigcup_{i : \Lambda_i \subseteq \mathcal{M}} \mathcal{R}_i$.

We have $H(\boldsymbol{C}_{\mathcal{M}} | \boldsymbol{C}_{\mathcal{M}'}) = 0$. Moreover $H(\boldsymbol{S} | \boldsymbol{C}_{\mathcal{M}}) = 0$. This implies,

$$H(\boldsymbol{S} | \boldsymbol{C}_{\mathcal{M}'}) = 0.$$

Now we can select any ℓ-subset $\mathcal{L}$ of $\mathcal{M}'$ and assume that the eavesdropper observes that set. Therefore, $H(\boldsymbol{S}) = H(\boldsymbol{S} | \boldsymbol{C}_{\mathcal{L}})$ must be less than or equal to the number of symbols in $\mathcal{M}' \setminus \mathcal{L}$. Formally,

$$\begin{aligned} k = H(\boldsymbol{S}) = H(\boldsymbol{S} | \boldsymbol{C}_{\mathcal{L}}) \leq H(\boldsymbol{C}_{\mathcal{M}'} | \boldsymbol{C}_{\mathcal{L}}) & \leq |\mathcal{M}' \setminus \mathcal{L}| \\ & = |\mathcal{M}'| - \ell. \end{aligned} \tag{11}$$

This observation will lead us to eq. (8). We describe below, the only remaining task: the method for constructing the set $\mathcal{M}$ described above, and show that it gives us eq. (8). The construction for $\mathcal{M}$ is given in algorithm 1.

Note that algorithm 1 may not actually give the set containing the maximum number of Λ_i but it would suffice to prove the bound in eq. (8). Let ν denote number of sets Λ_i added to $\mathcal{M}^0$. We have, $|\Lambda_i| \leq r + 1, \forall i$. So the maximum size of the set added in each step is $r + 1$. Since $|\mathcal{M}| = m$

```
Data: $\mathcal{R}_i$ for all $i$
Result: $\mathcal{M} \subseteq [n], |\mathcal{M}| = m$ containing at least
        $\lfloor m/(r+1) \rfloor$ recovery sets
1  $j = 0$; $\mathcal{M}^j = \emptyset$
2  choose any $t \in [n]$
3  while $|\mathcal{M}^j \cup \{\Lambda_t\}| < m$ do
4      $\mathcal{M}^{j+1} = \mathcal{M}^j \cup \Lambda_t$
5      choose $t \notin \mathcal{M}^{j+1}$
6      $j = j + 1$
7  end
8  if $|\mathcal{M}^j \cup \Lambda_t| \leq m$ then
9      $\mathcal{M}^{j+1} = \mathcal{M}^j \cup \Lambda_t$
10 else
11     $\mathcal{I} =$ any $(m - |\mathcal{M}^j|)$-subset of $[n] \setminus \mathcal{M}^j$
12     $\mathcal{M}^{j+1} = \mathcal{M}^j \cup \mathcal{I}$
13 end
14 $j = j + 1$
15 $\mathcal{M} = \mathcal{M}^j$
```

Algorithm 1: Constructing a set $\mathcal{M} \subseteq \{1, 2, \dots, n\}$ to maximize $|\{i : \Lambda_i \subseteq \mathcal{M}\}|$

by construction, when the algorithm ends at line 9 we have $\nu \geq \left\lceil \frac{m}{r+1} \right\rceil$. If the algorithm ends at line 10 we must have, $\nu \geq \left\lfloor \frac{m}{r+1} \right\rfloor$. Evidently we have constructed a set $\mathcal{M}$ such that $|\mathcal{M}'| = |\mathcal{M}| - \nu \leq m - \left\lfloor \frac{m}{r+1} \right\rfloor$. From eq. (11) we have,

$$k \leq m - \left\lfloor \frac{m}{r+1} \right\rfloor - \ell. \tag{12}$$

■

Using eq. (11) we can show the following,

Corollary 4. *There exists a set $J \subseteq [n]$ with $\ell \leq |J| \leq m - \left\lfloor \frac{m}{r+1} \right\rfloor$ such that,*

$$H(\boldsymbol{S}|\boldsymbol{C}_J) \leq m - \left\lfloor \frac{m}{r+1} \right\rfloor - |J|. \tag{13}$$

Equation (13) gives an upper-bound on the maximum ambiguity of the secret of an (n, k, ℓ, m, r)-scheme when the eavesdropper has access to more than ℓ shares.

B. Constructions

It is possible to show matching achievability results to theorem 2 by a number of different methods.

Theorem 5. *There exists a (n, k, ℓ, m, r)-secret sharing scheme such that eq. (7) is satisfied with equality.*

In particular this theorem can be proved by constructing a random linear network code. We delegate that proof to Appendix B.

The achievability result also follows from [18], that gives a construction for optimal secure LRC employing Gabidulin codes to satisfy the security constraint. In the subsequent we describe their method, adapted for our scenario, because this will be useful later in our paper when we consider more general secret sharing schemes.

An intuitive construction of ℓ-secure schemes comes by replacing some inputs to a LRC with uniform random variables. Formally, consider a linear code $\mathcal{C}$ with code-length n and dimension $(k + \ell)$. Let $G = [G^1 \; G^2] \in \mathbb{F}_q^{n \times (k+\ell)}$ be the generator matrix of this code such that $G^1 \in \mathbb{F}_q^{n \times \ell}$ and $G^2 \in \mathbb{F}_q^{n \times k}$. Let $\boldsymbol{a} \in \mathbb{F}_q^{k+\ell}$ be the input to the encoder of $\mathcal{C}$ (i.e., the codeword is generated by multiplying $\boldsymbol{a}$ with the generator matrix of $\mathcal{C}$). Denote by $\boldsymbol{s} \in \mathbb{F}_q^k$ the input we want to store securely. We construct an ℓ-secure secret sharing scheme using $\mathcal{C}$ by taking,

$$\boldsymbol{a} = \begin{bmatrix} \boldsymbol{r} \\ \boldsymbol{s} \end{bmatrix} \tag{14}$$

where $\boldsymbol{r} \in \mathbb{F}_q^\ell$ is an instance of uniformly distributed random vector. This scheme is ℓ-secure if and only if for any ℓ *linearly independent* rows of G the corresponding rows of G^1 are linearly independent.

Lemma 6. *Let $\boldsymbol{g}_i = [g_{i1}g_{i2} \dots g_{i(k+\ell)}], i \in [\ell]$ be any ℓ linearly independent rows of G. The secret sharing scheme constructed in eq. (14) is ℓ-secure if and only if the corresponding row vectors $\boldsymbol{g}_i^1 = [g_{i1}g_{i2} \dots g_{i\ell}], i \in [\ell]$ of G^1 are linearly independent.*

The proof of lemma 6 is given in Appendix A. Note that using lemma 6 we can add the security property to any linear code; we do not assume any locality property for the generator matrix G. But, it is clear that if the generator matrix G has locality r, then so would the scheme constructed in eq. (14). The construction of an optimal $(n, k, \ell, m, r)_q$ scheme is described in the following.

Gabidulin precoding construction: Let N be an integer. The points $\alpha_i \in \mathbb{F}_{q^N}, i \in [n]$ can be represented as vectors in $\mathbb{F}_q^N$ and are said to be $\mathbb{F}_q$*-linearly independent* when the corresponding vectors over $\mathbb{F}_q$ are linearly independent. A Gabidulin code from $\mathbb{F}_{q^N}^k \to \mathbb{F}_{q^N}^n$, for input $(f_1 f_2 \dots f_k), f_i \in \mathbb{F}_{q^N}$, is obtained by evaluating the linearized polynomial $\Theta(y) = \sum_{i=1}^k f_i y^{q^{i-1}}$ at n $\mathbb{F}_q$-linearly independent points $\alpha_i \in \mathbb{F}_{q^N}, i \in [n]$. The linearized polynomial $\Theta(y)$ has the following linearity property,

$$\Theta(ax + by) = a\Theta(x) + b\Theta(y) \tag{15}$$

for all $x, y \in \mathbb{F}_{q^N}$ and $a, b \in \mathbb{F}_q$. Note that, we need $N \geq n$ to obtain n $\mathbb{F}_q$-linearly independent points in $\mathbb{F}_{q^N}$.

Consider the generator matrix, $G = [\boldsymbol{g}_1 \dots \boldsymbol{g}_n]^T$ of a linear $(n, k + \ell, r)_q$-optimal LRC, where $\boldsymbol{g}_i = [g_{i1} \dots g_{i(k+\ell)}]^T$. Consider $\boldsymbol{a} = (\boldsymbol{s} \quad \boldsymbol{r})$, where $\boldsymbol{r}$ is an instance of uniformly distributed random variable in $\mathbb{F}_{q^N}^\ell$ and $\boldsymbol{s} \in \mathbb{F}_{q^N}^k$, $N \geq n$, denotes the secret. First, $\boldsymbol{a}$ is precoded using a Gabidulin code, $\Gamma : \mathbb{F}_{q^N}^{k+\ell} \to \mathbb{F}_{q^N}^{k+\ell}$ which is obtained by evaluating the polynomial,

$$\Psi_{\boldsymbol{a}}(y) = \sum_{i=1}^{k+\ell} a_i y^{q^{i-1}} \tag{16}$$

at the $\mathbb{F}_q$-linearly independent points $\alpha_i \in \mathbb{F}_{q^N}, i \in [k + \ell]$. Now, representing $\Gamma(\boldsymbol{a}) \in \mathbb{F}_{q^N}^{k+\ell}$ as a matrix of size $(k + \ell) \times N$ in $\mathbb{F}_q$, each column of the matrix can be encoded independently using the generator matrix G for the optimal LRC to get $(c_i)_{i=1}^n = \boldsymbol{c} \in \mathbb{F}_{q^N}^n$. It is easy to show that this construction is ℓ-secure. The optimality of the scheme then follows from the optimality of the initial linear LRC. The proof of security of this construction is given below.

Proof of theorem 5 with the Gabidulin construction: Assume without loss of generality (wlog) that the eavesdropper observes $\mathcal{E} = [\ell] \subseteq [n]$ symbols $c_i, i \in \mathcal{E}$. Let $\tilde{G} = [\boldsymbol{g}_1 \dots \boldsymbol{g}_\ell]^T$. Further assume that the $\operatorname{rank}(\tilde{G}) = \ell$, since otherwise the ℓ-strength eavesdropper is equivalent to an $\operatorname{rank}(\tilde{G})$-strength eavesdropper. Let $\tilde{\alpha}_i = \sum_{j=1}^{k+\ell} g_{ij}\alpha_j, i \in \mathcal{E}$. Then since $\tilde{G}$ is full-rank $\{\tilde{\alpha}_i\}_{i\in\mathcal{E}}$ are $\mathbb{F}_q$-linearly independent. Therefore, using eq. (15) we have,

$$\begin{aligned} c_i &= \sum_{j=1}^{k+\ell} g_{ij}\Psi_{\boldsymbol{a}}(\alpha_j) \\ &= \Psi_{\boldsymbol{a}}(\sum_{j=1}^{k+\ell} g_{ij}\alpha_j) = \Psi_{\boldsymbol{a}}(\tilde{\alpha}_i), i \in \mathcal{E}. \end{aligned}$$

Let $\boldsymbol{R}, \boldsymbol{S}, \boldsymbol{C}$ be the random variables corresponding to the vector $\boldsymbol{r}$, the secret $\boldsymbol{s}$, and the node shares $\boldsymbol{C} = (C_i)_i$. To prove security we use the secrecy lemma in [18, Lemma 4], to show that $H(\boldsymbol{C}_\mathcal{E}) \le H(\boldsymbol{R})$ and $H(\boldsymbol{R}|\boldsymbol{S}, \boldsymbol{C}_\mathcal{E}) = 0$ imply $H(\boldsymbol{S}|\boldsymbol{C}_\mathcal{E}) = H(\boldsymbol{S})$. Indeed, $H(\boldsymbol{S}|\boldsymbol{C}_\mathcal{E}) \le H(\boldsymbol{S})$, and

$$\begin{aligned} H(\boldsymbol{S}) + H(\boldsymbol{R}) &= H(\boldsymbol{S}|\boldsymbol{R}) + H(\boldsymbol{R}) \\ &= H(\boldsymbol{S}, \boldsymbol{R}) = H(\boldsymbol{S}, \boldsymbol{C}_\mathcal{E}, \boldsymbol{R}) \\ &= H(\boldsymbol{C}_\mathcal{E}) + H(\boldsymbol{S}, \boldsymbol{R}|\boldsymbol{C}_\mathcal{E}) \\ &= H(\boldsymbol{C}_\mathcal{E}) + H(\boldsymbol{R}|\boldsymbol{S}, \boldsymbol{C}_\mathcal{E}) + H(\boldsymbol{S}|\boldsymbol{C}_\mathcal{E}) \\ &= H(\boldsymbol{C}_\mathcal{E}) + H(\boldsymbol{S}|\boldsymbol{C}_\mathcal{E}) \qquad \text{(17a)} \\ &\le H(\boldsymbol{R}) + H(\boldsymbol{S}|\boldsymbol{C}_\mathcal{E}) \qquad \text{(17b)} \end{aligned}$$

where eqs. (17a) and (17b) follow from the assumptions $H(\boldsymbol{R}|\boldsymbol{S}, \boldsymbol{C}_\mathcal{E}) = 0$ and $H(\boldsymbol{C}_\mathcal{E}) \le H(\boldsymbol{R})$ respectively. On the other hand, assuming that the eavesdropper also knows $\boldsymbol{s}$ (in addition to $\boldsymbol{c}_\mathcal{E}$), she/he has

$$\tilde{c}_i = c_i - \sum_{j=1}^{k} s_j \tilde{\alpha}_i^{q^{\ell+j-1}} = \sum_{j=1}^{\ell} r_j \tilde{\alpha}_i^{q^{j-1}}, i \in \mathcal{E}.$$

Since $B = [\tilde{\alpha}_i^{q^{j-1}}]_{i\in\mathcal{E}, j\in[\ell]}$ is full rank, the eavesdropper can compute $[\tilde{c}_1 \dots \tilde{c}_\ell]B^{-1} = [r_1 \dots r_\ell]$. Thus, $H(\boldsymbol{R}|\boldsymbol{S}, \boldsymbol{C}_\mathcal{E}) = 0$. Now $H(\boldsymbol{C}_\mathcal{E}) \le H(\boldsymbol{R})$, since $|\mathcal{E}| \le \ell$. Therefore, we have an $(n, k, \ell, m, r)_{q^N}$-secret sharing scheme. ■

C. Constructions with small alphabet size: equivalence with maximal recoverability

Note that, the size of the alphabet/shares in the construction of optimal secure scheme using Gabidulin codes is exponential in the number of nodes. In this section, our aim is to show that the construction of an optimal secure scheme with small alphabet size will amount to finding a *maximally recoverable code* over that alphabet. We use the construction in eq. (14) to form a secure scheme from an optimal LRCs with a small alphabet and analyze the conditions for that construction to satisfy lemma 6. We assume $(r+1)|n$ i.e. $r+1$ divides n for simplicity in this subsection.

We will need the following definition of maximally recoverable codes [7].

Definition 2. *Consider an $(n, k, r)_q$-optimal LRC. Let $\mathcal{Q}_j : |\mathcal{Q}_j| = r+1, j \in [n/(r+1)]$ denote a partition of $[n]$ such that the recovery set of ith coordinate is,*

$$\mathcal{R}_i = \mathcal{Q}(i) \setminus \{i\}, \ \forall i \in [n], \tag{18}$$

where $\mathcal{Q}(i) \in \{\mathcal{Q}_j\}_j$ is the partition containing node i. Denote such an LRC by $(n, k, r, \{\mathcal{Q}_j\}_j)_q$. The $(n, k, r, \{\mathcal{Q}_j\}_j)_q$ LRC is called maximally recoverable if the code obtained by puncturing any one symbol from each $\mathcal{Q}_j$ is maximum distance separable (MDS).

Note that, in [8], it was pointed out that an optimal linear LRC must have the recovery structure as in eq. (18).

The main objective of this section is to show that the immediate construction of (n, k, ℓ, m, r)-secret-sharing scheme from an optimal LRC is effective if and and only if the code is maximally recoverable.

Lemma 7. *For any linear $(n, k+\ell, r, \{\mathcal{Q}_j\}_j)_q$ -optimal LRC code with a generator matrix $G \in \mathbb{F}_q^{n\times(k+\ell)}$ consider $\mathcal{S} \subseteq [n] : |\mathcal{S}| = \ell$ and $|\mathcal{S} \cap \mathcal{Q}_j| \le r, j \in [n/(r+1)]$. Then, the rows corresponding to $\mathcal{S}$ in G are linearly independent for any ℓ such that*

$$\ell \le r - 1 + \left(r\left\lfloor\frac{k}{r-1}\right\rfloor - k\right) \tag{19}$$

Proof: Partition $\mathcal{S}$ as follows, $\mathcal{S} = \bigcup_{j\in[n/(r+1)]} \mathcal{S}_j$ with $\mathcal{S}_j = \mathcal{S} \cap \mathcal{Q}_j$ and let $\Lambda := \{j : \mathcal{S}_j \neq 0\}$. Consider a set $\mathcal{S}' \supset \mathcal{S} : |\mathcal{S}'| \le k+\ell$ and define $\mathcal{S}'_j := \mathcal{S}' \cap \mathcal{Q}_j$. Suppose that we can construct $\mathcal{S}'$ with $\mathcal{S}'_j \le r, \forall j \in [n/(r+1)]$ such that the number of partitions $\mathcal{Q}_j$ that contain r co-ordinates of $\mathcal{S}'$ is at least $\lceil (k+\ell)/r \rceil - 1$. Let $\Psi := \{j : \mathcal{S}'_j = r\}$. Thus,

$$|\Psi| \ge \lceil (k+\ell)/r \rceil - 1 \tag{20}$$

Construct a set $\mathcal{S}'' \supseteq \mathcal{S}'$ by adding $k+\ell-|\mathcal{S}'|$ co-ordinates to $\mathcal{S}'$ such that, $|S'' \cap \mathcal{Q}_j| \le r, \forall j \in [n/(r+1)]$. Now at least $|\Psi|$ more co-ordinates are recoverable from $\mathcal{S}''$. Note that the input $\boldsymbol{a}$ for $(n, k+\ell, r, \{\mathcal{Q}_j\}_j)_q$-optimal LRC is recoverable from any $m = (k+\ell) + \lceil (k+\ell)/r \rceil - 1$ co-ordinates and $|\mathcal{S}''| + |\Psi| \ge m$. Thus, $\boldsymbol{a}$ is recoverable from $\boldsymbol{c}_{\mathcal{S}''}$. Now, since $|\mathcal{S}''| = k+\ell$ the rows of G corresponding to $\mathcal{S}''$ (and hence $\mathcal{S}$) must be L.I. We are now left with the task of constructing a set $\mathcal{S}'$ satisfying eq. (20) for the given $\mathcal{S}$ with $|\mathcal{S}| = \ell$ satisfying eq. (19). The construction is given below.

For $|\Lambda| \le k/(r-1)$ we can easily construct $\mathcal{S}'$. Since $|\Lambda| \le k/(r-1) \implies |\Lambda|r \le k+\ell$, we can choose $\Psi (\supseteq \Lambda) : |\Psi| = \lfloor \frac{k+\ell}{r} \rfloor$. Now to each of the partitions $\{\mathcal{S}_j\}_{j\in\Psi}$ add $r - |\mathcal{S}_j|$ co-ordinates from $\mathcal{Q}_j$ to get a set $\mathcal{S}'$ of size $r\lfloor (k+\ell)/r \rfloor \le k+\ell$. It is easy to see that this set satisfies eq. (20).

Now assume that $|\Lambda| > k/(r-1)$. Choose any $\Psi \subseteq \Lambda : |\Psi| = \lfloor k/(r-1) \rfloor$. Select any $r - |\mathcal{S}_j|$ co-ordinates from $\mathcal{Q}_j$ for all $j \in \Psi$. Adding these co-ordinates to $\mathcal{S}$, we get $\mathcal{S}'$ satisfying $|\mathcal{S}'| \le \lfloor k/(r-1) \rfloor (r-1) + \ell \le k+\ell$. Thus, from eq. (19) we have,

$$|\Psi| + 1 - \lceil (k+\ell)/r \rceil \ge \lfloor k/(r-1) \rfloor - \frac{k+\ell}{r}$$

$$\begin{aligned}&\geq \lfloor k/(r-1)\rfloor - \frac{k}{r} - (1+\lfloor k/(r-1)\rfloor - k/r - 1/r)\\&= -(1-1/r)\end{aligned}$$

Since $|\{\mathcal{Q}_j : |\mathcal{Q}_j \cap \mathcal{S}'| = r\}| + 1 - \lceil (k+l)/r\rceil$ is an integer, $m' + 1 - \lceil (k+l)/r\rceil \geq 0$, $\mathcal{S}'$ satisfies eq. (20). ■

For $\ell < r$, the construction (in eq. (14)) using an optimal LRC code is ℓ-secure since any ℓ rows of G_1 form an $\ell \times \ell$ Vandermonde matrix. For $\ell > r$, we have the following result, using definition 2 and lemma 7.

Theorem 8. *Consider a linear $(n, k+\ell, r, \{\mathcal{Q}_j\}_j)_q$ -optimal LRC $\mathcal{C}$. Then the construction in eq. (14) using code $\mathcal{C}$ is ℓ-secure if there exists $\mathcal{C}' \subseteq \mathcal{C}$ of dimension ℓ such that $\mathcal{C}'$ is maximally recoverable. Conversely, if the construction in eq. (14) is ℓ-secure then there must exist a maximally recoverable code $\mathcal{C}' \subseteq \mathcal{C}$ of dimension ℓ, for $\ell \leq r - 1 + (r\lfloor k/(r-1)\rfloor - k)$*

Proof: Let $G = [G^1\ G^2] \in \mathbb{F}_q^{n\times(k+\ell)}$ be the generator matrix of $\mathcal{C}$ where $G^1 \in \mathbb{F}_q^{n\times\ell}$. Let G^1 be the generator matrix of a maximally recoverable code $\mathcal{C}'$. Consider a set $\mathcal{D} \subseteq [n]$ of any ℓ linearly dependent rows of G^1. Since $\mathcal{C}'$ is maximally recoverable, $\mathcal{Q}_j \subseteq \mathcal{D}$ for at least one $j \in [n/(r+1)]$. Hence, the corresponding rows in G must also be linearly dependent. Thus, from lemma 6 the secret sharing construction in eq. (14) must be ℓ-secure.

Now, suppose that $\mathcal{C}$ does not contain any subcode of dimension ℓ which is maximally recoverable. Then, the code generated by G^1 is not maximally recoverable. Thus, there would exist an $\mathcal{S} \subseteq [n] : |\mathcal{S}| = \ell$ and $|\mathcal{S} \cap \mathcal{Q}_j| \leq r, \forall j \in [n/(r+1)]$ such that the rows in G^1 corresponding to $\mathcal{S}$ are linearly dependent. Now from lemma 7 we know that the rows corresponding to $\mathcal{S}$ in G are not linearly dependent for $\ell \leq r - 1 + (r\lfloor k/(r-1)\rfloor - k)$. Hence, from lemma 6 the secret sharing scheme cannot be ℓ secure. ■

Recently an optimal construction of locally repairable codes was proposed in [24] by Tamo and Barg for general values of the parameters $n, k,$ and r and alphabet size of $O(n)$. Our theorem 8 implies that the secret sharing scheme constructed in eq. (14) using such code is ℓ-secure if and only if the Tamo-Barg codes are maximally recoverable. In general these codes are not maximally recoverable. It should be noted that, it is quite a nontrivial open problem to construct maximally recoverable codes with linear or even polynomial (in blocklength) alphabet size [7].

In the next two sections we extend the notions and results of section II to other generalized repair conditions related to distributed storage.

III. Security for Schemes with cooperative repair

Cooperative repair for a locally repairable scheme addresses simultaneous multiple failures in a distributed storage system [19][2]. To this end, we extend the definition in eq. (4) to a (r, δ) scheme where any δ –instead of just one– shares can be recovered from r other shares.

Definition 3. *A set $\mathcal{C} \subseteq \mathbb{F}_q^n$ is said to be (r,δ)-repairable if for every $\Delta \subseteq [n] : |\Delta| \leq \delta$ there exists a set $\mathcal{R}(\Delta) \subseteq [n] \setminus \Delta : |\mathcal{R}(\Delta)| \leq r$ such that for all $\boldsymbol{c}, \boldsymbol{c}' \in \mathcal{C}$,*

$$\boldsymbol{c}_\Delta \neq \boldsymbol{c}'_\Delta \implies \boldsymbol{c}_{\mathcal{R}(\Delta)} \neq \boldsymbol{c}'_{\mathcal{R}(\Delta)} \tag{21}$$

Using definition 3 we can generalize the notion of an $(n, k, \ell, m, r)_q$-secret sharing scheme. For this system we derive an upper bound on the capacity k given $n, m, \ell, r,$ and δ.

Definition 4. *An $(n, k, \ell, m, (r, \delta))_q$-secret sharing scheme consists of a randomized encoder $f(.)$ that stores a file $\boldsymbol{s} \in \mathbb{F}_q^k$ in n separate shares, such that the scheme is (r, δ)-repairable (definition 3), satisfies the recovery condition (cf. eq. (2)) and ℓ-secure (cf. eq. (3)).*

A. *The case of $m = n$*

Error-correcting codes with (r, δ)-repairability were considered in [19] ($\ell = 0$ or no security) and the following upper-bound on the rate of such codes has been proposed, for the case of $m = n$.

$$R = \frac{k}{n} \leq \frac{r}{r+\delta}. \tag{22}$$

For the case of ℓ-secure codes we give an analogous upper bound on the rate of a secret sharing scheme in the following.

Theorem 9. *The rate $R = k/n$ of an $(n, k, \ell, n, (r, \delta))_q$ secret sharing scheme is bounded as,*

$$R \leq \frac{r}{r+\delta} - \frac{\ell}{n}. \tag{23}$$

Proof: For an $(n, k, \ell, (r, \delta))_q$ scheme we construct a set of size $m = n$ similar to algorithm 1 except instead of choosing a set of size 1 in steps 2 and 5, we find a set of size δ. Then using the same arguments we must have at least $\nu = m/(r+\delta)$ number of steps. Hence, subtracting the number recoverable symbols $\delta\nu$ from the m symbols we must have,

$$\begin{aligned}k + \ell &\leq m - \delta\nu = n - \delta\frac{n}{r+\delta}\\ \implies \frac{k+\ell}{n} &\leq \frac{r}{r+\delta}.\end{aligned}$$

■

Construction: Note that, any linear q-ary (r, δ)-repairable error-correcting code of length n and dimension k will give rise to a $(n, k, 0, (r, \delta))$-secret sharing scheme. In [19, Sec. 6], an (r, δ) repairable code has been constructed using bipartite graphs of large girth. In particular, that construction results in parameters such that

$$\frac{k}{n} \geq \frac{r-\delta}{r+\delta}.$$

It can also be seen from the discussion of section II-B that Gabidulin precoding (eq. (16)) would give an ℓ-secure construction with alphabet $\mathbb{F}_{q^N}$, $N \geq n$, from any optimal linear $(n, k+\ell, 0, (r, \delta))_q$-secret sharing scheme. Thus, for any $(n, k+\ell, 0, (r, \delta))_q$ secret sharing scheme achieving the upper-bound in eq. (22) we can achieve the corresponding upper-bound in theorem 9. Hence, using the code of [19, Sec. 6]

---

[2] There is a related notion of cooperative recovery in regenerating codes [23] and security in such systems [12]. In this paper we are concerned with only the local recovery problem, and not the regenerating problem.

in conjunction with the Gabidulin precoding, it is possible to obtain a rate of

$$\frac{k}{n} \geq \frac{r-\delta}{r+\delta} - \frac{\ell}{n},$$

which is an additive term of $\frac{\delta}{r+\delta}$ away from the optimum possible.

B. The case of $m < n$.

The bound for general case of $m < n$ can be deduced from the same arguments as above. In fact, by slightly generalizing algorithm 1, we get the following result: for any $(n, k, \ell, m, (r, \delta))_q$-secret sharing scheme ,

$$k + \ell \leq m - \left\lfloor \frac{m}{r+\delta} \right\rfloor \delta - h \tag{24}$$

where $h = (m \mod (r+\delta) - r)^+$ and $x^+ := \begin{cases} 0 & x \leq 0, \\ x & x > 0. \end{cases}$

Note that, this results in slightly weaker bound for the case of $m = n$ than eq. (23). In general for $m < n$ and arbitrary values of ℓ, we do not have any good construction that will be close to the bound. While the expander-graph based constructions of (r, δ)-locally repairable codes from [19] can be generalized, their performance is very far from the bound of eq. (24).

IV. SECURITY FOR REPAIRABLE CODES ON GRAPHS

Another extension of local repair property for distributed storage has recently been proposed in [13], [14]. Consider a Distributed Storage System as a directed graph $\mathcal{G}$ such that a node of the graph represents a node of the Distributed Storage System and each node can connect to only its out-neighbors for repair. We define an ℓ-secure code in this scenario as follows.

A. Repairable Codes on Graph

Definition 5. *Let $\mathcal{G} = ([n], E)$ be a graph on n nodes. An $(n, k, \ell, m, \mathcal{G})_q$-secret sharing scheme consists of a randomized encoder f that can store a uniformly random secret $\boldsymbol{S} \in \mathbb{F}_q^k$ on n shares/nodes, $\boldsymbol{C} = f(\boldsymbol{S}), \boldsymbol{C} \in \mathbb{F}_q^n$, such that the system is ℓ-secure (cf. eq. (3)) and the data can be recovered from any m shares (cf. eq. (2)). In addition the share of any node can be recovered from its neighbors i.e.*

$$H(\boldsymbol{C}_i | \boldsymbol{C}_{N(i)}) = 0$$

where $N(i) = \{j \in [n] : (i, j) \in E\}$ denotes the neighbors (out-neighbors in the case of a directed graph) of node i in the graph $\mathcal{G} = ([n], E)$.

A bound on the capacity of such a scheme in directed graphs for $\ell = 0$ (no security) was derived in [15],

$$m \geq k + \max_{\substack{U \in \mathcal{I}(\mathcal{G}): \\ |N(U)| \leq k-1}} |U| \tag{25}$$

where $\mathcal{I}(\mathcal{G})$ denotes the set of induced acyclic subgraphs in $\mathcal{G}$, and $N(U) := \cup_{i \in U} N(i) \setminus U$ denotes the neighbors of U. For undirected graphs we have the same bound with $\mathcal{I}(\mathcal{G})$ denoting the collection of all independent sets of the graph. The lower bound on m for an ℓ-secure scheme on a graph $\mathcal{G}$ is given in the following.

Theorem 10. *For any $(n, k, \ell, m, \mathcal{G})_q$-secret sharing scheme on a directed graph $\mathcal{G}$, m satisfies the following lower bound,*

$$m \geq k + \ell + \max_{\substack{U \in \mathcal{I}(\mathcal{G}): \\ |N(U)| \leq \ell+k-1}} |U| \tag{26}$$

where $\mathcal{I}(G)$ denotes the set of induced acyclic graphs in $\mathcal{G}$.

Proof: Since any m co-ordinates in the shares $\boldsymbol{C} = (C_i)_{i \in [n]}$ can recover the secret $\boldsymbol{S}$ we must have,

$$m \geq |W| + 1 \tag{27}$$

for all $W \subseteq [n]$ such that the $H(\boldsymbol{S}|\boldsymbol{C}_W) > 0$. Let U be an acyclic subgraph $U \in \mathcal{I}(\mathcal{G})$, such that $N(U) \leq \ell + k - 1$. Construct a set $V \supseteq \{U \cup N(U)\}$ by adding any $\ell + k - 1 - |N(U)|$ nodes to $U \cup N(U)$. Thus, $|V| = k + \ell + |U| - 1$. We show that $H(\boldsymbol{S}|\boldsymbol{C}_V) > 0$ for any such V.

Note that for any three random X, Y, Z variables we must have,

$$\begin{aligned} H(X|Y, Z) &= H(X, Z|Y) - H(Z|Y) \\ &= H(X|Y) + H(Z|X, Y) - H(Z|Y) \\ &\geq H(X|Y) - H(Z). \end{aligned} \tag{28}$$

Assume that the eavesdropper selects an ℓ-subset $\mathcal{E} \subseteq [n]$ in the set V. Then, since the eavesdropper must not get any information about the secret,

$$H(\boldsymbol{S}|\boldsymbol{C}_{\mathcal{E}}) = H(\boldsymbol{S}) \tag{29}$$

Since the sub-graph U is acyclic the nodes in U must be a function of the leaf nodes and the nodes in $N(U)$. Now, the leaf nodes must also be a function of $N(U)$ since their out-neighbors can only be in $N(U)$. Therefore,

$$\begin{aligned} H(\boldsymbol{S}|\boldsymbol{C}_V) &= H(\boldsymbol{S}|\boldsymbol{C}_{N(U)}) = H(\boldsymbol{S}|\boldsymbol{C}_{\mathcal{E}}, \boldsymbol{C}_{N(U)\setminus\mathcal{E}}) \\ &\overset{(a)}{\geq} H(\boldsymbol{S}|\boldsymbol{C}_{\mathcal{E}}) - H(\boldsymbol{C}_{N(U)\setminus\mathcal{E}}) \\ &\overset{(b)}{=} H(\boldsymbol{S}) - H(\boldsymbol{C}_{N(U)\setminus\mathcal{E}}) \\ &\overset{(c)}{>} 0 \end{aligned}$$

where (a) and (b) follow from eq. (28) and eq. (29) respectively, and (c) is is true since $|N(U) \setminus \mathcal{E}| = k - 1$. ■

When $m = n$, i.e. when the scheme does not need to protect against catastrophic failures, we can formulate a converse bound for repairable codes on graphs that does not follow directly from the above theorem.

Theorem 11. *Consider an $(n, k, \ell, n, \mathcal{G})_q$ secret sharing scheme. The secrecy capacity of the scheme satisfies the following upper-bound.*

$$k \leq n - |U| - |\ell| \tag{30}$$

where U is the largest acyclic induced subgraph in $\mathcal{G}$ when $\mathcal{G}$ is a directed graph, and it is the largest independent set when $\mathcal{G}$ is undirected.

Proof: We will show the proof for $\mathcal{G}$ directed. Consider the shares $\boldsymbol{C}_U$ corresponding to the nodes in $U \subseteq [n]$. The recovery set of any node in U can contain its children in U or co-ordinates in $[n] \setminus U$. Since U is ayclic, all the leaf nodes of U have recovery sets in $[n] \setminus U$. Thus, we can recover all the leaf nodes from the co-ordinates in $[n] \setminus U$. Now, we can recursively recover all the co-ordinates of U from the co-ordinates in $[n] \setminus U$. Thus,

$$H(\boldsymbol{C}_U | \boldsymbol{C}_{[n]\setminus U}) = 0 \tag{31}$$

Equation (31) is true because all the leaf nodes in U must have their recovery sets in $[n] \setminus U$. And by recovering the leaf nodes we can recover all nodes in U. Now, since $H(\boldsymbol{S}|\boldsymbol{C}) = 0$ we must have from eq. (31),

$$H(\boldsymbol{S} | \boldsymbol{C}_{[n]\setminus U}) = 0 \tag{32}$$

Now, suppose that the eavesdropper selects an ℓ-subset $\mathcal{E} \in [n] \setminus U$. Then, we must have,

$$H(\boldsymbol{S}) = H(\boldsymbol{S}|\boldsymbol{C}_{\mathcal{E}}) \tag{33}$$

Therefore, using eqs. (32) and (33) we have,

$$\begin{aligned}
H(\boldsymbol{C}_{[n]\setminus U}|\boldsymbol{C}_{\mathcal{E}}) &= H(\boldsymbol{C}_{[n]\setminus U}|\boldsymbol{C}_{\mathcal{E}}) + H(\boldsymbol{S}|\boldsymbol{C}_{[n]\setminus U}, \boldsymbol{C}_{\mathcal{E}}) \\
&= H(\boldsymbol{S}, \boldsymbol{C}_{[n]\setminus U}|\boldsymbol{C}_{\mathcal{E}}) \\
&= H(\boldsymbol{S}|\boldsymbol{C}_{\mathcal{E}}) + H(\boldsymbol{C}_{[n]\setminus U}|\boldsymbol{S}, \boldsymbol{C}_{\mathcal{E}}) \\
&= H(\boldsymbol{S}) + H(\boldsymbol{C}_{[n]\setminus U}|\boldsymbol{S}, \boldsymbol{C}_{\mathcal{E}}) \\
\implies H(\boldsymbol{S}) &= H(\boldsymbol{C}_{[n]\setminus U}|\boldsymbol{C}_{\mathcal{E}}) - H(\boldsymbol{C}_{[n]\setminus U}|\boldsymbol{S}, \boldsymbol{C}_{\mathcal{E}}) \\
\implies H(\boldsymbol{S}) &\le H(\boldsymbol{C}_{[n]\setminus U}|\boldsymbol{C}_{\mathcal{E}}) \le n - |U| - \ell.
\end{aligned}$$

■

Note that the bound in eq. (30) parallels the feedback vertex set upper-bound in [15, Prop. 11]. Here, a feedback vertex set of a graph is a set of nodes such that every cycle in the graph has a vertex in the set.

B. Achievable Schemes for Secure Repairable Codes on Graphs

In this section we consider construction of $(n, k, \ell, m, \mathcal{G})_q$-secret sharing scheme only when $m = n$. We do not have any nontrivial construction for the case of $m < n$.

Consider a secret sharing scheme for the case of undirected graphs (definition 5). A maximum matching $\mathcal{M}(\mathcal{G})$ of the graph $\mathcal{G}$ is defined as the set of edges of maximum cardinality such that no two edges have a vertex in common. To construct a recoverable scheme for this code, with input $\boldsymbol{x} \in \mathbb{F}^{|\mathcal{M}(\mathcal{G})|}$, we assign a coordinate of $\boldsymbol{x}$ to both vertices for every edge in $\mathcal{M}(\mathcal{G})$. For recoverability, we note that a symbol in vertex v can be recovered from u, where $(v, u) \in \mathcal{M}(\mathcal{G})$.

Suppose $|\mathcal{M}(\mathcal{G})| = k + \ell$. Consider the vector input $\boldsymbol{x} \in \mathbb{F}^{k+\ell}$ to the above scheme. We set $\boldsymbol{x} = G \times [\boldsymbol{s} \quad \boldsymbol{r}], \boldsymbol{s} \in \mathbb{F}^k, \boldsymbol{r} \in \mathbb{F}^\ell$, where $\boldsymbol{s}$ is the secret, $\boldsymbol{r}$ is an instance of a uniform random vector, and G is the $(k+\ell) \times (k+\ell)$ Vandermonde matrix $G = [\alpha_i^{j-1}]_{ij}$ with $\{\alpha_i\}_i$ distinct elements in $\mathbb{F}_q$. Thus, from lemma 6, we see that this scheme is ℓ-secure as well as recoverable.

The capacity of this scheme is $k = |\mathcal{M}(\mathcal{G})| - \ell \ge \frac{n-|U|}{2} - \ell$, where U is the maximum independent set. This is true since if we remove both end-vertices of the edges of the matching then we are left with an independent set. Compared to eq. (30), we are an additive term of at most $\frac{n-|U|}{2}$ away from what is the maximum possible.

For directed graphs $\mathcal{G} = ([n], E)$ we use the repairable codes presented in [15] below to construct a secure scheme. Suppose that the graph has $K := k + \ell$ vertex disjoint cycles. Then it is easy to see that we can form a locally repairable scheme capable of storing $k + \ell$ symbols (one symbol per cycle) by repeating the same symbol on every vertex in a cycle. Hence, it is possible to store as many symbols as the maximum number of vertex disjoint cycles in the graph. In [15], it was shown that we can do better by using vector codes. We describe below the vector linear LRC codes constructed in [15].

Consider the set $\mathcal{P}$ of all cycles in $\mathcal{G}([n], E)$. Suppose, $\Pi : \mathcal{P} \to \mathbb{Q}$ assigns a rational number to every directed cycle. Let $V(C), C \in \mathcal{P}$ denote the vertices of the cycle C. Let K denote the maximum value of $\sum_{C \in \mathcal{P}} \Pi(C)$, over all such mappings Π, under the following constraint,

$$\sum_{C : i \in V(C)} \Pi(C) \le 1, \quad \forall i \in [n].$$

Let the optimal assignment Π on $\mathcal{P}$ be denoted as $\Pi(C) = \frac{n(C)}{p}$, where $n(C), p \in \mathbb{Z}^+$. It is possible to find this optimum by solving a linear program. Then [15] constructs a vector LRC for the graph $\mathcal{G}$ in $\mathbb{F}_q$ with storage capability of pK symbols and per node storage equal to p symbols.

Let $\boldsymbol{s} \in \mathbb{F}_q^{pk}, \boldsymbol{r} \in \mathbb{F}_q^{p\ell}$ represent the secret and an instance of a uniform random vector, respectively. We obtain $\boldsymbol{x} \in \mathbb{F}_q^{pK}, K := k + \ell$, by $\boldsymbol{x} = G \times [\boldsymbol{s} \quad \boldsymbol{r}]$, where G is a $pK \times pK$ Vandermonde matrix $G = [\alpha_i^{j-1}]_{ij}$ with $\{\alpha_i\}_i$ distinct elements in $\mathbb{F}_q$. $\boldsymbol{x}$ is then stored in the graph using the scheme described above. Since an ℓ-strength eavesdropper can only observe at most $p\ell$ co-ordinates in $\boldsymbol{a}$, we can use lemma 6 to see that the scheme is ℓ-secure as well as recoverable.

It is known (cf. [15]) that, $4K \ln 4K \ln \log_2 4K \ge n - |U|$, for U being the maximum acyclic induced subgraph. Hence, we must have,

$$k \ge \frac{n - |U|}{c \log n \log \log n} - \ell.$$

However this achievability result is quite far away from the bound of eq. (30).

V. PERFECT SECRET SHARING AND GENERAL ACCESS STRUCTURES

So far in this paper we were concentrating on a secret sharing scheme that is not perfect, i.e., the access structure and the block-list are not complementary. In this section we provide results regarding existence of locally repairable of perfect secret sharing schemes and the relation between sizes of shares and secret in those schemes.

A. Perfect access structures with locality

To make the (n, k, ℓ, m, r) secret sharing scheme perfect, we must have $m = \ell + 1$. This results in a threshold secret-sharing scheme. Now, from eq. (8) we have,

$$k \le 1 - \left\lfloor \frac{\ell + 1}{r + 1} \right\rfloor.$$

Thus, for storing any secret we must have $r \geq \ell + 1 = m$. Since any secret sharing scheme works when $r \geq m$ (local repair in this case imply full revelation of secret) only trivial locally repairable codes are possible for threshold secret sharing schemes. This implies the following statement.

Proposition 12. *A threshold secret sharing scheme is not locally repairable.*

Note that, perfect secret sharing schemes are a natural generalization of threshold schemes. Although for threshold schemes the locality cannot be small/nontrivial, we show that this is not true for general access structures and perfect schemes. Indeed, the following is true.

Proposition 13. *There exists an access structure $\mathcal{A}_s$, for which a perfect secret sharing scheme is possible with arbitrary non-trivial locality r i.e. $r < \min_{A \in \mathcal{A}_s} |A|$.*

Proof: Let n, κ be such that $r | \kappa$ and $(r+1)|n$. Consider an $(n, \kappa, r, \{\mathcal{Q}_j\}_j)$ maximally recoverable LRC (definition 2). We know that such codes exist from [7]. Now, we use the Gabidulin precoding method described above to construct a $(n, k = 1, \ell = \kappa - 1, m = \kappa(1 + 1/r), r)$ secret sharing scheme from this code.

Define the access structure to be $\mathcal{A}_s = \{A \subseteq [n] : \sum_{j=1}^{n/(r+1)} \min\{|A \cap \mathcal{Q}_j|, r\} \geq \kappa\}$. Now given any $A \in \mathcal{A}_s$, a user accessing the shares corresponding to A can determine the secret s_0 because the set always contains k shares of a punctured $(nr/(r+1), \kappa)$-MDS code.

For a perfect secret sharing scheme the block-list is given by $\mathcal{B}_s = \{B : \sum_{j=1}^{n/(r+1)} \min\{|B \cap \mathcal{Q}_j|, r\} < \kappa\}$. Assume that the eavesdropper has access to a set $B \in \mathcal{B}_s$. Construct the following set of size at most $\kappa - 1$ from B,

$$B' = \cup_{j=1}^{n/(r+1)} N_j', B' \subseteq B$$

where $N_j' \subseteq N_j, N_j = B \cap \mathcal{Q}_j$ is obtained by removing any one co-ordinate if $|N_j| > r$, otherwise $N_j' = N_j$. Note that $|B'| < \kappa$. Since all the shares in B are recoverable from $B' \subseteq B$, an eavesdropper with access to the nodes in B is equivalent to an eavesdropper with access to B'. And since $|B'| \leq \ell = \kappa - 1$, the eavesdropper does not get any information about the secret. ∎

Can the above proposition be made general? Is it possible to characterize the locality for general secret sharing schemes? Shamir's [22] perfect threshold secret sharing scheme for the access structure $\mathcal{A}_s = \{A \subseteq [n] : |A| \geq k\}$ is one of the first general construction of secret sharing protocols. The scheme is defined for a scalar secret $s \in \mathbb{F}$ and a set of n participating nodes P. The scheme uses an (n, k) Reed Solomon code defined using the polynomial $\sigma(x) = s + \sum_{i=1}^{k-1} r_i x^i$, where r_i are instances of uniform random variables in $\mathbb{F}$.

Ito, Shaito, and Nishizeki [11] define a generalization of Shamir's scheme that works for arbitrary monotone access structures. Define a maximal element $B \in \mathcal{B}$ as a set such that $A \supsetneq B \implies A \notin \mathcal{A}$. Similarily, define a minimal set $A \in \mathcal{A}$ as a set such that $B \subsetneq A \implies B \notin \mathcal{A}$. Consider the set of maximal elements of the block-list $\mathcal{B}$, denoted $\mathcal{B}^\dagger$. The scheme uses the generator polynomial $\sigma(x) = s + \sum_{i=1}^{|\mathcal{B}^\dagger|-1} r_i x^i$ to generate $|\mathcal{B}^\dagger|$ shares $\{c_B\}_{B \in \mathcal{B}^\dagger}$ – one share corresponding to each maximal set in $\mathcal{B}$. The shares are distributed such that each user gets the shares corresponding to the subset it does not belong to, i.e. participant node p gets the shares

$$\{c_B : p \notin B, B \in \mathcal{B}^\dagger\} \tag{34}$$

Now, suppose that share of a node p is lost in a secure code with participants P and block-list $\mathcal{B}$. To recover the share of p we access the shares of participants in the set $\mathcal{R}(p)$ where the optimal set $\mathcal{R}(p)$ is

$$\mathcal{R}(p) = \min_{R : \forall B \in \mathcal{B}^\dagger, p \notin B \;\; R \not\subseteq B} |R|. \tag{35}$$

To have non-trivial locality, one must have $\max_p |\mathcal{R}(p)|$ to be strictly less than the maximal sets in the block-list.

B. Size of a share for perfect secret sharing with locality

We know that, for perfect secret sharing schemes, the size of the secret cannot be larger than the size of a share [2, Lemma 2]. Let us see why this statement is true. Let the secret $\boldsymbol{s}$ belong to a domain $\mathcal{K}$ and the share of node j belong to $\mathcal{K}_j$. Assume that there exists a perfect secret sharing scheme which realizes the access structure $\mathcal{A}$ when $|\mathcal{K}| < |\mathcal{K}_j|$. Let $B \subseteq [n]$ be a minimal set in $\mathcal{A}$ such that $j \in B$. Define $B' = B \setminus \{j\}$. Then, since the secret sharing scheme is perfect, for every value of the the shares in B_j all secrets in K must have the same probability. Thus, since the value of the shares of B determine the secret completely there must exist an injective mapping from K to K_j. But since $|K_j| < |K|$ this cannot be possible.

In [5] the minimum node storage required for arbitrary monotone access structures is analyzed. In that paper, an access structure was constructed for which the sizes of the shares has to be $n/log(n)$ times the size of the secret for any perfect scheme. For secret sharing schemes with local repairability and fixed recovery sets, all monotone access structures are not feasible. The minimal sets of the access structure cannot include any recovery set. Here, we extend the result in [5] to the restricted class of monotone access structures.

Assume $(r+1)|n$. Suppose that the secret denoted by the random variable S is stored on n shares as $C_i, i \in [n]$ and the shares have locality r (eq. (4)). Consider a partition of $[n]$, $\mathcal{Q}_j : \mathcal{Q}_j, j \in [n/(r+1)]$ such that the recovery sets are given by eq. (18). For a perfect secret sharing scheme on $[n]$ with monotone access structure $\mathcal{A}_s$, the minimal sets $\mathcal{A}_s^\star$ of $\mathcal{A}_s$, must satisfy,

$$A \in \mathcal{A}_s^\star \implies A \not\supseteq \mathcal{Q}_j \ . \tag{36}$$

Denote this class of monotone access structures with $\mathbb{M}_s$. We have the following result for the minimum size of a share for secret sharing schemes with access structure $\mathcal{A}_s \in \mathbb{M}_s$.

Theorem 14. *Consider distribution of shares of secret S to n nodes with locality r, recovery sets as in eq. (18). Then, there is an access structure $\mathcal{A}_s \in \mathbb{M}_s$ (eq. (36)), such that any perfect scheme for $\mathcal{A}_s$, if exists, must satisfy,*

$$\alpha \geq \frac{(r+1)n}{r \log n} H(S). \tag{37}$$

where α is the average entropy of the shares.

Proof: First, let us define a *polymatroid* $(Q = \{[n], S\}, \phi)$ as follows,

$$\phi(A) = \frac{H(\boldsymbol{c}_A)}{H(S)},\ A \subseteq [n] \tag{38a}$$

$$\phi(A, S) = \frac{H(\boldsymbol{c}_A, S)}{H(S)},\ A \subseteq [n] \tag{38b}$$

A polymatroid function must satisfy the following properties,

- **P1** $\phi(A) \geq 0$ for all $A \subseteq Q$, $\phi(\emptyset) = 0$
- **P2** ϕ is monotone i.e. $A \subseteq B \subseteq Q$, then $\phi(A) \leq \phi(B)$
- **P3** ϕ is submodular i.e. $\phi(A)+\phi(B) \geq \phi(A\cup B)+\phi(A\cap B)$ for any $A, B \subseteq Q$

Note that, the definition in eq. (38) satisfies all the conditions above. In addition, the definition satisfies the following properties,

- **Pa** $\phi(A, S) = \phi(A)$, for every $A \in \mathcal{A}_s$
- **Pb** $\phi(A, S) = \phi(A) + 1$, for every $A \notin \mathcal{A}_s$

which easily follow from the recovery and the security properties i.e. $H(S|\boldsymbol{c}_B) = H(S)$ and $H(S|\boldsymbol{c}_A) = 0$, $A \in \mathcal{A}_s$ and $B \in \mathcal{B}_s = 2^{[n]} - \mathcal{A}_s$ and the definition in eq. (38).

Using properties (P1) to (P3) and properties (Pa) and (Pb) we have the following result, for any $A, B \in \mathcal{A}_s$ such that $A \cap B \notin \mathcal{A}_s$,

$$\begin{aligned} &\phi(A, S) + \phi(B, S) \geq \phi((A\cup B), S) + \phi((A \cap B), S) \\ &\implies \phi(A) + \phi(B) \geq \phi(A \cup B) + \phi(A \cap B) + 1 \end{aligned} \tag{39}$$

Consider the set M of size η such that $(r+1)|\eta$ and it contains $\eta/(r+1)$ partitions $\mathcal{Q}_j$. Another set $N \subseteq [n] \setminus M : |N| = \nu := 2^\eta - (r+2)^{\eta/(r+1)} + 1$ is chosen such that $|N \cap \mathcal{Q}_j| \leq r,\ \ \forall j$. The parameter η for the size of the sets M, N is chosen to be the largest possible, i.e. the maximum η satisfying,

$$\eta - \left\lfloor \frac{\eta}{r+1} \right\rfloor + 2^\eta - (r+2)^{\eta/(r+1)} + 1 \leq n\frac{r}{r+1} \tag{40}$$

Now, construct a sequence $\{M_i\}_{i=0}^{\nu-1}$, for $M_i \in 2^M$ of length ν, such that it satisfies the following conditions for all sets M_i in the sequence,

- **C1** If for any partition $\mathcal{Q}_j$, $\mathcal{Q}_j \cap (M_i - M_{i+1}) \neq \emptyset$ and $|\mathcal{Q}_j \cap M_i| \geq r$, then $|\mathcal{Q}_j \cap M_{i+1}| < r$
- **C2** $M_i \not\subseteq M_{i'}, i < i'$

To construct the sequence $\{M_i\}_i$ of length ν satisfying conditions C1 and C2, we first construct a sequence $\{M_i'\}_{i=0}^{2^\eta-1}$, $M_i' \subseteq M : |M_i'| \leq |M_{i+1}'|$. It is easy to see that all subsequences of $\{A_i'\}$ satisfy condition C2. From this sequence we remove all sets $M_i', i \geq 1$ such that $|(M_0 - M_i') \cap \mathcal{Q}_j| \leq 1$. Note that, the number of the sets removed is,

$$\sum_{1 \leq i \leq \eta/r+1} \binom{\eta/(r+1)}{i} (r+1)^i = (r+2)^{\eta/(r+1)} - 1.$$

The sequence $\{M_i\}_i$ thus constructed has length ν. To see that this sequence satisfies condition C1 note that $|(M_0 - M_i) \cap \mathcal{Q}_j| > 1, \forall i \geq 1$ implies that $\{M_i\}_i$ satisfies condition C1. Thus the constructed sequence satisfies conditions C1 and C2.

Let $N = \{b_1, \ldots, b_{\nu-1}\}$. Define another sequence of sets $N_i = \{b_1, \ldots, b_i\}, i \in [\nu - 1]$ and $N_0 = \emptyset$. Consider a monotone access structure $\mathcal{A}_s$ that contains the sets $U_i := M_i \cup N_i, i \in \{0, \ldots, \nu-2\}$. Let the minimal sets in this access structure be,

$$\mathcal{A}_s^\star = \left\{ A \subseteq U_i : |A \cap \mathcal{Q}_j| = \min\{|A \cap \mathcal{Q}_j|, r\},\ \forall i \in \Big[\frac{n}{r+1} \Big\rfloor\Big] \right\}. \tag{41}$$

Thus, $\mathcal{A}_s \in \mathbb{M}_s$.

Consider the following sets $P = N_i \cup M$ and $Q = M_{i+1} \cup N_{i+1}$. Since $P \supseteq U_i$ and $Q \supseteq U_{i+1}$, $P, Q \in \mathcal{A}_s$. Now, $P \cap Q = N_i \cup M_{i+1}$. From condition C1 and eq. (41), we see that there exists a set $A^\star \in \mathcal{A}_s^\star, A^\star \subseteq U_i$ such that $P \cap Q \subsetneq A^\star$. Therefore, $P \cap Q \notin \mathcal{A}_s$. Applying eq. (39) on P, Q, we have,

$$\begin{aligned} &[\phi(N_i \cup M) - \phi(N_i \cup M_{i+1})] \\ &\quad - [\phi(N_{i+1} \cup M) - \phi(N_{i+1} \cup M_{i+1})] \geq 1. \end{aligned} \tag{42}$$

Using property (P3) we have,

$$\phi(N_{i+1} \cup M_{i+1}) - \phi(N_i \cup M_{i+1}) \geq \phi(N_{i+1}) - \phi(N_i). \tag{43}$$

Thus, combining eqs. (43) and (44) we have,

$$[\phi(N_i \cup M) - \phi(N_i)] - [\phi(N_{i+1} \cup M) - \phi(N_{i+1})] \geq 1. \tag{44}$$

Adding eq. (44) for $i \in \{0, \ldots, \nu - 3\}$ we have,

$$\phi(M) - [\phi(N_{\nu-2} \cup M) - \phi(N_{\nu-2})] \geq \nu - 2. \tag{45}$$

Thus, from the recoverability property we have $\phi(M) \leq \eta r/(r+1)\alpha$. Since, $M \in \mathcal{A}_s$ and $N_{\nu-2} \notin \mathcal{A}_s$, $\phi(N_{\nu-2} \cup M) - \phi(N_{\nu-2}) \geq 1$. Thus, we have from eq. (45),

$$\alpha \geq (r+1)\frac{2^\eta - (r+2)^{\eta/(r+1)}}{\eta r} H(S). \tag{46}$$

Since, $\eta = \Omega(\log n)$ and $(r+2)^{1/(r+1)} < 2$ from eq. (40), eq. (46) asympototically (with n) gives,

$$\alpha \geq \left(\frac{r+1}{r}\right) \frac{n}{\log n} H(S).$$

■

APPENDIX A
PROOF OF LEMMA 6

Consider the submatrix $H_{\ell \times (k+\ell)}$ of G corresponding to ℓ rows, $I_\ell \subseteq [n]$. Assume that the eavesdropper observes I_ℓ. Wlog assume that $\mathrm{rank}(H) = \ell$, since the eavesdropper effectively observes $\mathrm{rank}(H)$ shares.

" $\Longleftarrow$ " Assume that any ℓ rows of G^1 corresponding to ℓ L.I. rows of G are L.I. Thus, $\mathrm{rank}(H_1) = \ell$ by assumption. Let $\boldsymbol{c} = G\boldsymbol{a}$ and $H = [H_1\ \ H_2]$ where H_1 is $\ell \times \ell$ and H_2 is $\ell \times k$. Then,

$$H_1\boldsymbol{r} = \boldsymbol{c}_{I_\ell} - H_2\boldsymbol{s} \tag{47}$$

Now, given c_{I_ℓ}, for every $\boldsymbol{s}$ there is a unique solution to $\boldsymbol{r} = H_1{}^{-1}(\boldsymbol{c}_{I_\ell} - H_2\boldsymbol{s})$. Since, each of those vectors are equally probable the eavesdropper does not get any information about $\boldsymbol{s}$.

" $\Longrightarrow$ " Conversely, suppose that H_1 is not full rank. (but $\operatorname{rank}(H) = \ell$ by assumption). If for a given $\boldsymbol{c}_{I_\ell}$ there does not exist a solution to eq. (47) for some $\boldsymbol{s} \in \mathbb{F}_q^k$ then $H(\boldsymbol{s}|\boldsymbol{c}_{I_\ell}) < H(\boldsymbol{s})$. This happens iff for some $\boldsymbol{a} \in \mathbb{F}_q^{k+\ell}$,

$$H\boldsymbol{a} - \operatorname{colspan}(H_2) \nsubseteq \operatorname{colspan}(H_1) \tag{48}$$

where $\operatorname{colspan}(.)$ denotes the column span of a matrix and $H\boldsymbol{a} - \operatorname{colspan}(H_2) = \{H\boldsymbol{a} - \boldsymbol{v} : \boldsymbol{v} \in \operatorname{colspan}(H_2)\}$. Now, $\operatorname{colspan}(H_2) \nsubseteq \operatorname{colspan}(H_1)$ since $\dim(\operatorname{colspan}(H_1, H_2)) = \ell$ and $\dim(\operatorname{colspan}(H_1)) < \ell$ by assumption. Thus, eq. (48) is satisfied for $\boldsymbol{a} = \boldsymbol{0}$ which implies that in this case the eavesdropper does get some information about $\boldsymbol{s}$.

APPENDIX B
ACHIEVABILITY USING LINEAR NETWORK CODES

In this appendix, we show that the limit derived in theorem 2 is achievable using a random *linear network code* (LNC). The rest of this section is devoted to the proof of theorem 5 via the technique provided in [16]. We assume that k_0 is such that,

$$m = k_0 + k_0/r - 1 \tag{49}$$

For simplicity, further assume that r divides k_0 and $(r+1)$ divides n.

Our roadmap for the proof is the following. We analyze the network flow graph in fig. 1, that has been adapted and modified from [16]. We first show that this graph has multicast capacity k_0. Further there exists an LNC for this graph which corresponds to an $(n, k_0, 0, m, r)$-secret sharing scheme. Then, we impose additional constraints on the LNC for the graph in fig. 1 to get an ℓ-secure scheme, i.e., an $(n, k = k_0 - \ell, \ell, m, r)$-scheme. Clearly this satisfies eq. (7).

We start by describing the graph in fig. 1 (Left). This graph, $\mathcal{G}(n, k_0, m, r)$ consists of a source node X that transmits k_0 q-ary symbols to $T = \binom{n}{m}$ data collectors $DC_\mu, \mu \in [T]$. We assume that X transmit the secret $\boldsymbol{s} \in \mathbb{F}_q^{k_0}$. The unit for the edge capacity is taken to be one q-ary symbol per channel use. The nodes $F_\nu, \nu \in [r]$ connect to the source X through links with capacity k_0/r. The edges that connect $\Gamma_\rho, \rho \in [\frac{n}{r+1}]$ to $Y_i^{\text{in}}, i \in [n]$, has capacity r. All the rest of the edges have unit capacity. Each of $\Gamma_\rho, \rho \in [\frac{n}{r+1}]$ have r incoming edges from $F_\nu, \nu \in [r]$. The edges (X, F_ν) are broken into k_0/r unit capacity edges and labelled $s_1, s_2, \ldots, s_{k_0}$ as shown in the subgraph in fig. 1 (Right). Node F_ν connects to the source X through edges $\{s_{\nu+(\lambda-1)r}\}_{\lambda=1}^{k_0/r}, \nu \in [r]$. Let us denote the subset of nodes $\{\Gamma_\rho, \{Y^{\text{in}}_{(\rho-1)(r+1)+j}\}_{j=1}^{r+1}, \{Y^{\text{out}}_{(\rho-1)(r+1)+j}\}_{j=1}^{r+1}\}$ as the ρ^{th} repair group.

A single network use corresponds to a sequence of single data transmission on every edge. Assume that, data transmitted on the edges $(Y_i^{\text{in}}, Y_i^{\text{out}}), i \in [n]$ in a single network use correspond to the n shares of the secret (i.e., n symbols of $f(\boldsymbol{s})$, where f is the randomized encoding). Note that, the data collectors connect to m nodes (shares) and obtain all of what X transmits: this must be satisfied for all m-subsets (all data collectors). We use the network $\mathcal{G}(n, k_0, m, r)$ to show the existence of a linear $(n, k_0, 0, m, r)$-secret sharing scheme.

Lemma 15. *Given that the network $\mathcal{G}(n, k_0, m, r)$ has multicast capacity k_0, there exists a linear network code with repairability r for this network and the scheme corresponding to the data transmitted on the edges (Y_i^{in}, Y_i^{out}) is an $(n, k_0, 0, m, r)$-secret sharing scheme.*

In the following we show that the network $\mathcal{G}(n, k_0, m, r)$ has multicast capacity k_9.

Definition 6. *A min-cut for any two nodes v, u in $\mathcal{G}(n, k_0, m, r)$, denoted $\operatorname{MinCut}(v, u)$, is defined as a subset of directed edges of minimum aggregate capacity such that if these edges are removed, then there does not exist a path from v to u in the graph $\mathcal{G}(n, k_0, m, r)$. Let $|\operatorname{MinCut}(v, u)|$ denote the aggregate capacity of the edges in $\operatorname{MinCut}(v, u)$.*

It has been shown [1], [10] that the minimum of the min-cuts between a single source and multiple sinks corresponds to the *multicast capacity* of the source. We show that for $\mathcal{G}(n, k_0, m, r)$ this quantity, $\min_{\mu \in [T]} |\operatorname{MinCut}(X, DC_\mu)|$, is equal to k_0.

Lemma 16. *For $\mathcal{G}(n, k_0, m, r)$ the multicast capacity is k_0. That is,*

$$\min_{\mu \in [T]} |\operatorname{MinCut}(X, DC_\mu)| = k_0. \tag{50}$$

Proof: For k_0 satisfying eq. (49) we have,

$$m = k_0 + \frac{k_0}{r} - 1 = (k_0/r - 1)(r+1) + r. \tag{51}$$

Suppose that the minimum in eq. (50) only contains an n_1-subset $\mathcal{E}$ of edges in $\{(X, F_\nu)\}_{\nu \in [r]}$. Assume wlog that $\mathcal{E} = \{(X, F_1), \ldots, (X, F_{n_1})\}$. Consider the data collector DC_μ that connects to $\gamma_\rho, \rho \in [n/(r+1)]$ nodes in each of the repair groups. If $\gamma_\rho \geq r - n_1$ the min-cut should include all the edges $\{(F_{n_1+1}, \Gamma_\rho), \ldots, (F_r, \Gamma_\rho)\}$. Otherwise if $\gamma_\rho < r - n_1$ the min-cut includes all the γ_ρ edges (Y_i^{in}, Y_i^{out}) in the ρ^{th} repair group connected to DC_μ. Therefore, the minimum in eq. (50) would correspond to the data collector that covers entirely as many repair groups as possible. From eq. (51) we see that for a such data collector $\gamma_\rho \geq (r - n_1)$ for all ρ for which $\gamma_\rho > 0$ and for all $0 \leq n_1 \leq r$. Therefore,

$$\min_{\mu} |\operatorname{MinCut}(X, DC_\mu)| = \frac{k_0}{r}(r - n_1) + n_1 \frac{k_0}{r} = k_0$$

■

We know therefore that a random LNC achieves the multicast capacity k_0 for this network. This random LNC corresponds to a secret-sharing scheme with n shares such that the secret in $\mathbb{F}_q^{k_0}$ can be recovered by looking at any m shares. Now to satisfy the local repairability constraint for this LNC, consider the subgraph containing the nodes in the ρ^{th} repair group. Another set of local decoding requirements are imposed on this subgraph. For each r-subset of nodes in any local repair group, a local data collector $LD_i, i \in [n]$ connecting to these nodes should be able to decode the input to Γ_ρ. There are in total n such local decoding requirements. These decoding requirements are similar to the local repairability requirements for the network flow graph considered in [16]. Let $\boldsymbol{z}_\rho \in \mathbb{F}_q^r$ denote the data received by Γ_ρ. Let N_i denote the $r \times r$ local encoding matrix, for the edges $\{(\Gamma_\rho, Y^{in}_{(\rho-1)(r+1)+j})\}_{j \in [r+1] \setminus \{i\}}$ corresponding to i^{th}

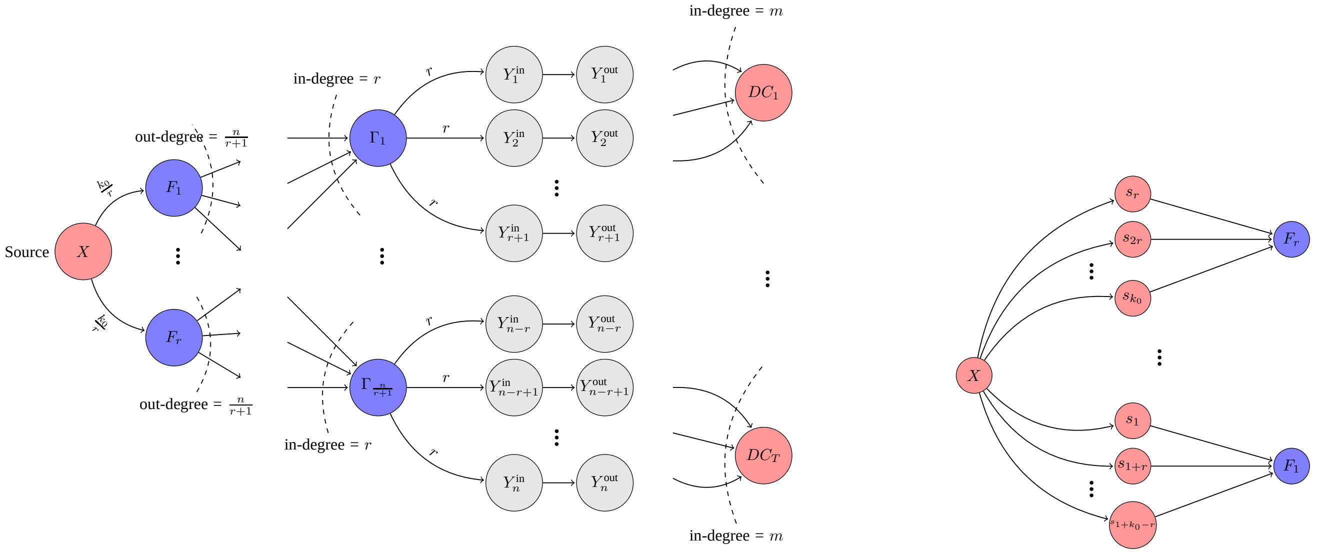

Fig. 1: Left: The information flow-graph $\mathcal{G}(n, k_0, m, r)$ adapted from [16]. The left-most vertex is the source node X. The $T = \binom{n}{m}$ vertices DC_μ are the destination nodes (referred to as the data collectors). Each DC is connected to a different m-tuple of Y_i^{out} nodes. Each of the intermediate nodes $F_\nu, \nu \in [r]$ have out-going edges to all the nodes $\Gamma_\rho, \rho \in \left[\frac{n}{r+1}\right]$. Right: Equivalent representation for the subgraph containing nodes F_ν and the source X.

local data collector. Therefore, the data received by the i^{th} local decoder is,

$$\boldsymbol{z}_\rho N_i, i \in \{(\rho - 1)(r + 1) + 1, \dots, \rho(r + 1)\} \tag{52}$$

We see that, for any local data collector LD_i to recover the data from the node Γ_ρ matrix N_i must be full rank. Since we know that for a large enough alphabet size q we can satisfy these constraints [16, lemma 4], there must exist an LNC that satisfies the local repair requirements. Therefore, we can construct an $(n, k_0, 0, m, r)$-secret-sharing scheme.

Suppose we write the secret as $\boldsymbol{s} = (s_1, \dots, s_{k_0})$, and term $s_1, \dots, s_{k_0}$ as the information symbols. Now, for the random LNC obtained above that satisfy the repairability and recovery requirements, we relabel $k = k_0 - \ell$ information symbols $\{s_{\ell+1}, \dots, s_{k_0}\}$ from the source X as secure information symbols and the choose each of the rest ℓ symbols $\{s_1, \dots, s_\ell\}$ according to a uniformly random distribution in $\mathbb{F}_q$. For such a random LNC to be ℓ-secure any eavesdropper $ED_\tau, \tau \in [\binom{n}{\ell}]$ connecting to any ℓ nodes Y_i^{out} may be able to recover at most the redundant ℓ symbols $\{s_1, \dots, s_\ell\}$ and should have full ambiguity about $\{s_{\ell+1}, \dots, s_{k_0}\}$. We show that these additional security constraints can be satisfied for a random LNC with large enough alphabet and hence we have an (n, k, ℓ, m, r)-secret-sharing scheme satisfying eq. (7).

Note that if a code is secure against an eavesdropper who can observe any of the ℓ shares, it must be secure against any adversary who can only observe less than ℓ shares. Therefore, for $\ell > r$ we can ignore all eavesdroppers who choose all the $(r + 1)$ shares of the same repair group. Since one of the shares in a repair group can be recovered from the other r shares, an eavesdropper who reads t entire repair groups is observing effectively only $\ell - t$ shares. Therefore, we only need to consider the eavesdroppers that observe a maximum of r shares in a repair group. Let us denote this sub-set of eavesdropper as $ED_\tau, \tau \in \mathcal{W}', \mathcal{W}' \subseteq [\binom{n}{\ell}]$.

If $(c_1, \dots, c_n)$ are the n shares for the secret $\boldsymbol{s}$, we must have the data transmitted on the edges (Y_i^{in}, Y_i^{out}) with the following linear form,

$$\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,k_0} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,k_0} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,k_0} \end{pmatrix} \begin{pmatrix} s_1 \\ \vdots \\ s_{k_0} \end{pmatrix} = A\text{s}. \tag{53}$$

We claim that the security against an eavesdropper $ED_\tau, \tau \in \mathcal{W}'$ is equivalent to a full-rank requirement on a $\ell \times \ell$ sub-matrix of A.

Lemma 17. *Let $\mathcal{E}^\tau = \{e_1^\tau, e_2^\tau, \dots, e_\ell^\tau\} \subseteq [n]$ denotes the shares an eavesdropper ED_τ can observe. We have,*

$$\boldsymbol{c}_{\mathcal{E}^\tau} = A_1^\tau \boldsymbol{s}_{[\ell]} + A_2^\tau \boldsymbol{s}_{[k_0] \setminus [\ell]}. \tag{54}$$

If for all eavesdroppers $ED_\tau, \tau \in \mathcal{W}'$ the $\ell \times \ell$ matrix A_1^τ is full-rank then the LNC is ℓ-secure.

Proof: Suppose for some specific $\tau \in \mathcal{W}'$,

$$A_1^\tau = \begin{pmatrix} a_{e_1,1} & a_{e_1,2} & \cdots & a_{e_1,\ell} \\ a_{e_2,1} & a_{e_2,2} & \cdots & a_{e_2,\ell} \\ \vdots & \vdots & \ddots & \vdots \\ a_{e_\ell,1} & a_{e_\ell,2} & \cdots & a_{e_\ell,\ell} \end{pmatrix}; A_2^\tau = \begin{pmatrix} a_{e_1,\ell+1} & \cdots & a_{e_1,k_0} \\ a_{e_2,\ell+1} & \cdots & a_{e_2,k_0} \\ \vdots & \vdots & \ddots \\ a_{e_\ell,\ell+1} & \cdots & a_{e_\ell,k_0} \end{pmatrix}.$$

Since A_1^τ is full rank, there must be a unique solution to $s_1, s_2, \dots, s_\ell$ for every value of $\boldsymbol{c}_{\mathcal{E}^\tau}$ and every value of $\{s_{\ell+1}, \dots, s_{k_0}\} \in \mathbb{F}_q^{k_0}$. Hence, we have,

$$H(\boldsymbol{s}_{[\ell]} | \boldsymbol{c}_{\mathcal{E}^\tau}, \boldsymbol{s}_{[k_0] \setminus [\ell]}) = 0$$

We therefore have the following chain of inequalities that establishes that the eavesdropper does not get any information about the secret from his observation.

$I(\boldsymbol{s}_{[k_0]\setminus[\ell]};\boldsymbol{c}_{\mathcal{E}^\tau}) = H(\boldsymbol{c}_{\mathcal{E}^\tau}) - H(\boldsymbol{c}_{\mathcal{E}^\tau}|\boldsymbol{s}_{[k_0]\setminus[\ell]}) \leq \ell - H(\boldsymbol{c}_{\mathcal{E}^\tau}|\boldsymbol{s}_{[k_0]\setminus[\ell]}) + H(\boldsymbol{c}_{\mathcal{E}^\tau}|\boldsymbol{s}_{[\ell]},\boldsymbol{s}_{[k_0]\setminus[\ell]}) = \ell - I(\boldsymbol{c}_{\mathcal{E}^\tau},\boldsymbol{s}_{[\ell]}|\boldsymbol{s}_{[k_0]\setminus[\ell]}) = \ell - H(\boldsymbol{s}_{[\ell]}|\boldsymbol{s}_{[k_0]\setminus[\ell]}) + H(\boldsymbol{s}_{[\ell]}|\boldsymbol{c}_{\mathcal{E}^\tau},\boldsymbol{s}_{[k_0]\setminus[\ell]}) = \ell - H(\boldsymbol{s}_{[\ell]}) = \ell - \ell = 0.$ ■

We also have the following lemma.

Lemma 18. *Consider the subgraph $\mathcal{G}_e$ formed by removing the edges $s_{\ell+1},\dots,s_{k_0}$ from the graph $\mathcal{G}(n,k_0,m,r)$. For this modified network graph the multicast capacity between the source and the eavesdroppers $ED_\tau, \tau \in \mathcal{W}'$ is ℓ i.e.*

$$\min_{\tau\in\mathcal{W}'} |\text{MinCut}(X, ED_\tau)| = \ell.$$

Proof: It is easy to see from the network structure that min-cut for every eavesdropper $ED_\tau, \tau \in \mathcal{W}'$ corresponds to all the edges (Y_i^{in}, Y_i^{out}) to which an eavesdropper connects in each repair group. Since, every eavesdropper in $\mathcal{W}'$ connects to ℓ nodes, the minimum mincut is also ℓ. ■

Consider the eavesdropper $ED_\tau, \tau \in \mathcal{W}'$ which connects to $t_1, t_2, \dots, t_{n/(r+1)}$ nodes in each of the repair groups. Therefore, we have

$$\sum_{\rho=1}^{n/(r+1)} t_\rho = \ell$$

where $0 \leq t_\rho \leq r, \forall \rho \in [n/(r+1)]$. Let $N'_\rho, \rho \in [n/(r+1)]$ denote the $t_\rho \times r$ local encoding sub-matrix of N_ρ (see, eq. (52)) for the edges (Γ_ρ, Y_i^{in}) connecting the eavesdropper to the ρ^{th} repair group. Also, let $D_\rho, \rho \in [n/(r+1)]$ denote the $r \times \ell$ matrix corresponding to the local encoding vectors for $(F_\nu, \Gamma_\rho), \nu \in [r]$, for the induced graph $\mathcal{G}_e$ described above. The matrix A_1^τ from lemma 17 can be written as,

$$A_1^\tau = \begin{pmatrix} N'_1 D_1 \\ N'_2 D_2 \\ \vdots \\ N'_{\frac{n}{r+1}} D_{\frac{n}{r+1}} \end{pmatrix}. \tag{55}$$

We need all of the matrices $A_1^\tau, \tau \in \mathcal{W}'$ to be full-rank simultaneously. Now using lemma 18 we can see that these constraints on the matrices $\mathbf{D}_\rho$s can all be satisfied simultaneously –with the local repairability and multicast capacity– for all $\tau \in \mathcal{W}'$ for a large enough alphabet size [10], [6, Lemma 4]. Therefore, a random LNC satisfies the full rank constraints of lemma 17.

Therefore, for the random LNC obtained above, for any eavesdropper ED_τ observing $\mathcal{E}^\tau \subseteq [n]$, $I(\boldsymbol{s}_{[k_0]\setminus[\ell]};\boldsymbol{c}_{\mathcal{E}^\tau}) = 0$. Since the data collectors can recover $\boldsymbol{s}$ from any m nodes and $H(\boldsymbol{s}_{[k_0]\setminus[\ell]}|\boldsymbol{s}) = 0$, the secret is recoverable from any m shares. Therefore, we have an (n,k,ℓ,m,r)-scheme achieving the upper bound in eq. (7).

REFERENCES

[1] R. Ahlswede, N. Cai, S.-Y. Li, and R. W. Yeung. Network information flow. *Information Theory, IEEE Transactions on*, 46(4):1204–1216, 2000. 11

[2] A. Beimel. Secret-sharing schemes: A survey, 2011. 1, 9

[3] G. R. Blakley. Safeguarding cryptographic keys. In *Managing Requirements Knowledge, International Workshop on*, pages 313–313. IEEE Computer Society, 1899. 1

[4] V. Cadambe and A. Mazumdar. An upper bound on the size of locally recoverable codes. In *Proc. IEEE Int. Symp. Network Coding*, June 2013. 1, 3

[5] L. Csirmaz. The Size of a Share Must Be Large. *Journal of Cryptology*, 10(4):223–231, Nov. 1997. 2, 9

[6] A. G. Dimakis, P. B. Godfrey, Y. Wu, M. J. Wainwright, and K. Ramchandran. Network coding for distributed storage systems. *IEEE Trans. Inform. Theory*, 56(9):4539–4551, Sep. 2010. 13

[7] P. Gopalan, C. Huang, B. Jenkins, and S. Yekhanin. Explicit maximally recoverable codes with locality. *Computing Research Repository*, abs/1307.4150, 2013. 5, 6, 9

[8] P. Gopalan, C. Huang, H. Simitci, and S. Yekhanin. On the locality of codeword symbols. *IEEE Trans. Inform. Theory*, 58(11):6925–6934, Nov. 2012. 1, 3, 5

[9] S. Goparaju, S. El Rouayheb, R. Calderbank, and H. V. Poor. Data secrecy in distributed storage systems under exact repair. In *Network Coding (NetCod), 2013 International Symposium on*, pages 1–6. IEEE, 2013. 1

[10] T. Ho, M. Médard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and B. Leong. A random linear network coding approach to multicast. *Information Theory, IEEE Transactions on*, 52(10):4413–4430, 2006. 11, 13

[11] M. Ito, A. Saito, and T. Nishizeki. Secret Sharing Scheme Realizing General Access Structure. In *Proc. IEEE GLOBECOM*, 1987. 9

[12] O. O. Koyluoglu, A. S. Rawat, and S. Vishwanath. Secure cooperative regenerating codes for distributed storage systems. *IEEE Transactions on Information Theory*, 60(9):5228–5244, 2014. 6

[13] A. Mazumdar. Achievable schemes and limits for local recovery on a graph. In *Proc. Allerton Conf. Commun., Contr., Computing*, 2014. 7

[14] A. Mazumdar. On a duality between recoverable distributed storage and index coding. In *Proc. Int. Symp. Inform. Theory*, pages 1977–1981. IEEE, 2014. 7

[15] A. Mazumdar. Storage capacity of repairable networks. *Information Theory, IEEE Transactions on*, 61(11), 2015. 7, 8

[16] D. S. Papailiopoulos and A. G. Dimakis. Locally repairable codes. In *Proc. Int. Symp. Inform. Theory*, pages 2771–2775, Cambridge, MA, July 2012. 1, 11, 12

[17] S. Pawar, S. El Rouayheb, and K. Ramchandran. Securing dynamic distributed storage systems against eavesdropping and adversarial attacks. *Information Theory, IEEE Transactions on*, 57(10):6734–6753, 2011. 1

[18] A. S. Rawat, O. O. Koyluoglu, N. Silberstein, and S. Vishwanath. Optimal locally repairable and secure codes for distributed storage systems. preprint, arXiv:1210.6954, 2012. 1, 2, 3, 4, 5

[19] A. S. Rawat, A. Mazumdar, and S. Vishwanath. Cooperative local repair in distributed storage. *EURASIP Journal on Advances in Signal Processing*, 2015(107), 2015. 2, 6, 7

[20] N. B. Shah, K. Rashmi, and P. V. Kumar. Information-theoretically secure regenerating codes for distributed storage. In *Proc. IEEE GLOBECOM*. IEEE, 2011. 1

[21] N. B. Shah, K. Rashmi, and K. Ramchandran. Secure network coding for distributed secret sharing with low communication cost. In *Proc. Int. Symp. Inform. Theory*, pages 2404–2408. IEEE, 2013. 1

[22] A. Shamir. How to share a secret. *Communications of the ACM*, 22(11):612–613, 1979. 1, 9

[23] K. W. Shum and Y. Hu. Cooperative regenerating codes. *IEEE Transactions on Information Theory*, 59(11):7229–7258, 2013. 6

[24] I. Tamo and A. Barg. A family of optimal locally recoverable codes. *IEEE Transactions on Information Theory*, 60(8):4661–4676, Aug 2014. 1, 6

[25] R. Tandon and S. Mohajer. New bounds for distributed storage systems with secure repair. In *Proc. Allerton Conf. Commun., Contr., Computing*, 2014. 1