Secure Distributed Storage: Optimal Trade-Off Between Storage Rate and Privacy Leakage

Rémi A. Chou
Department of Electrical Engineering & Computer Science
Wichita State University, Wichita, KS 67260
remi.chou@wichita.edu

Jörg Kliewer
Department of Electrical & Computer Engineering
New Jersey Institute of Technology, Newark, NJ 07102
jkliewer@njit.edu

2023 IEEE International Symposium on Information Theory (ISIT) | 978-1-6654-7554-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISIT54713.2023.10206726

***Abstract*—Consider the problem of storing data in a distributed manner over T servers. Specifically, the data needs to (i) be recoverable from any τ servers, and (ii) remain private from any z colluding servers, where privacy is quantified in terms of mutual information between the data and all the information available at any z colluding servers. For this model and under a leakage symmetry requirement at the servers, our main results are (i) the fundamental trade-off between storage size and the level of desired privacy, and (ii) the optimal amount of local randomness necessary at the encoder. As a byproduct, our results provide an optimal lower bound on the individual share size of ramp secret sharing schemes under a more general leakage symmetry condition than the ones previously considered in the literature.**

I. Introduction

Secure distributed storage schemes, e.g., [1]–[4], often rely on the idea of secret sharing as introduced in [5], [6] – we refer to [7] for a comprehensive literature review on secret sharing. Hence, there is a fundamental lower bound on the required storage space necessary to securely store information in a distributed manner. Specifically, in any threshold secret sharing scheme, the total amount of information that needs to be stored must at least be equal to the entropy of the secret times the number of participants, see e.g., [8], and it is thus impossible to reduce the storage space without any changes to the model assumptions.

In this paper, we propose to determine the optimal cost reduction, in terms of storage space, that can be obtained in exchange of tolerating a *controlled amount* of reduced privacy. Specifically, we focus on a setting where a file F needs to be stored at T servers. The file must be recoverable from τ servers, and needs to remain private from any z colluding servers. Here, privacy is quantified in terms of mutual information between the data and all the information available at any z colluding servers. In particular, we introduce a parameter $\alpha \in [0,1]$, to be chosen by the users, and require that no more than a fraction α of the file can be learned by a set of z colluding users. As a function of the parameters (τ, z, α), under the assumption of leakage symmetry, i.e., when the information leakage about the file at a given set of colluding servers only depends on the cardinality of the set and not on the identities of the servers among this set, we establish the optimal individual share size for each server. Secret sharing schemes that satisfies such a leakage symmetry are sometimes referred to as uniform secret sharing schemes, e.g., [9], [10]. A major difference between [9], [10] and our work is that, in [9], [10], optimal individual share sizes are derived for a fixed access function, i.e., the information leakage about the file tolerated at a given set of colluding servers is a fixed and given value. In contrast, in our setting we derive optimal individual share sizes for secret sharing schemes whose access functions are not fixed but are allowed to belong to a set of access functions, indeed, in our setting, only two points of the access functions are fixed as parameters: one point indicates a reconstruction threshold τ, another point indicates a maximum number of colluding servers z, and all the other points of the access function are optimized to minimize the share sizes. This difference introduces an optimization problem over a set of access functions to determine optimal individual share sizes that we solve in this study. As a byproduct of our results, when the privacy parameter is $\alpha = 0$, i.e., perfect privacy is required, we prove that among all uniform ramp secret sharing schemes, the ones that have a linear access function have the minimum individual share size.

We note that the idea of trading storage space against privacy is also closely related to non-perfect secret sharing [9]–[11], including ramp secret sharing with linear [12], [13] or non-linear access functions [14], [15]. Similar to our previous comment, these settings have been studied for fixed access functions, whereas, in this study, to minimize share sizes, we consider secret sharing schemes with access functions allowed to belong to a set of access functions.

The remainder of the paper is organized as follows. We formulate our problem statement and review known results in Section II. We present our main results in Section III and relegate the proofs to Section IV. Finally, we provide concluding remarks in Section V. Some proofs are omitted due to space constraints.

II. Problem statement and review of known results

Notation: Let $\mathbb{N}$, $\mathbb{R}$, and $\mathbb{Q}$ be the sets of natural, real, and rational numbers, respectively. For $a, b \in \mathbb{R}$, define $[\![a,b]\!] \triangleq [\lfloor a \rfloor, \lceil b \rceil] \cap \mathbb{N}$ and $[a]^+ \triangleq \max(0,a)$. Consider two arbitrary

This work was supported in part by NSF grants CCF-2201824 and CCF-2201825.

sets $\mathcal{S}$ and $\mathcal{T}$, a sequence of elements $x_t \in \mathcal{S}$, $t \in \mathcal{T}$, indexed by the set $\mathcal{T}$ is written as $(x_t)_{t\in\mathcal{T}}$.

A. Problem statement

Consider $T \geqslant 2$ servers indexed by $\mathcal{T} \triangleq [\![1, T]\!]$. For $t \in \mathcal{T}$, define $[\mathcal{T}]^{\geqslant t}$ as the set of all the subsets of $\mathcal{T}$ that have a cardinality larger than or equal to t, i.e., $[\mathcal{T}]^{\geqslant t} \triangleq \{\mathcal{S} \subset \mathcal{T} : |\mathcal{S}| \geqslant t\}$. Similarly, define $[\mathcal{T}]^{\leqslant t} \triangleq \{\mathcal{S} \subset \mathcal{T} : |\mathcal{S}| \leqslant t\}$ and $[\mathcal{T}]^{=t} \triangleq \{\mathcal{S} \subset \mathcal{T} : |\mathcal{S}| = t\}$.

Definition 1. *Let $(\lambda_t)_{t\in\mathcal{T}} \in \mathbb{N}^T$, $\rho \in \mathbb{N}$, and $\tau \in \mathcal{T}$. A $(\tau, (\lambda_t)_{t\in\mathcal{T}}, \rho)$ coding scheme consists of*

- *A file $F \in \{0,1\}^{|F|}$;*
- *Local randomness in the form of a sequence R of ρ bits uniformly distributed over $\{0,1\}^\rho$ and independent of F;*
- *T encoders $(e_t)_{t\in\mathcal{T}}$, where for $t \in \mathcal{T}$,*

$$e_t : \{0,1\}^{|F|} \times \{0,1\}^\rho \to \{0,1\}^{\lambda_t}, \\ (F, R) \mapsto M_t,$$

which takes as input the file F and the local randomness R, and outputs the sequence M_t, referred to as share in the following, of length $\lambda_t \in \mathbb{N}$. λ_t is referred to as share size in the following.

- *T servers, where Server $t \in \mathcal{T}$ stores M_t. In the following, for any subset $\mathcal{S} \subseteq \mathcal{T}$ of servers, we use the notation $M_\mathcal{S} \triangleq (M_t)_{t\in\mathcal{S}}$.*

Definition 2. *For $\tau \in \mathcal{T}$, $\alpha \in \mathbb{Q} \cap [0,1]$, and $z \in [\![1, \tau-1]\!]$, a $(\tau, (\lambda_t)_{t\in\mathcal{T}}, \rho)$ coding scheme is (α, z)-private if*

$$\max_{\mathcal{S}\in[\mathcal{T}]^{\geqslant\tau}} H(F|M_\mathcal{S}) = 0, \ \textit{(Recoverability)} \tag{1}$$

$$\max_{\mathcal{S}\in[\mathcal{T}]^{\leqslant z}} \frac{I(F; M_\mathcal{S})}{H(F)} \leqslant \alpha, \ \textit{(Privacy)}. \tag{2}$$

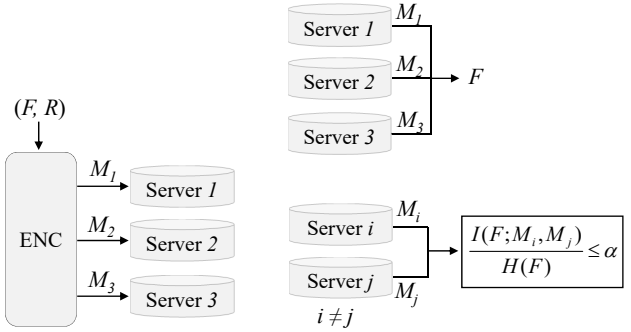
Requirement (1) means that any subset of τ or more servers can reconstruct the file F. Note that since conditioning reduces entropy, it is sufficient to take the maximization over $\mathcal{S} \in [\mathcal{T}]^{=\tau}$ in (1). Requirement (2) means that any subset of servers with size smaller than or equal to z must not learn more than $\alpha H(F)$ bits of information about F. In the following, τ is referred to as reconstruction threshold, α is referred to as privacy leakage parameter, and z is referred to as privacy threshold. The setting is illustrated in Figure 1 when $(T, \tau, z) = (3, 3, 2)$.

Remark. *In Definition 2, α is restricted to be a rational number. However, note that by density of $\mathbb{Q}$ in $\mathbb{R}$, for any $\beta \in [0,1]$, for any $\epsilon > 0$, there exists $\alpha \in \mathbb{Q} \cap [0,1]$ such that $|\alpha - \beta| \leqslant \epsilon$.*

Definition 3. *Let $\tau \in \mathcal{T}$, $\alpha \in \mathbb{Q} \cap [0,1]$, and $z \in [\![1, \tau-1]\!]$. Then, for $t \in \mathcal{T}$, define*

$$\lambda_t^\star(\alpha, z, \tau) \triangleq \min\{\lambda_t \in \mathbb{N} : \textit{there exists an } (\alpha, z)\textit{-private} \\ (\tau, (\lambda_{t'})_{t'\in\mathcal{T}}, \rho) \textit{ coding scheme} \\ \textit{for some } \rho \in \mathbb{N} \textit{ and } (\lambda_{t'})_{t'\in\mathcal{T}\setminus\{t\}} \in \mathbb{N}^{T-1}\},$$

$$\rho^\star(\alpha, z, \tau) \triangleq \min\{\rho \in \mathbb{N} : \textit{there exists an } (\alpha, z)\textit{-private} \\ (\tau, (\lambda_t)_{t\in\mathcal{T}}, \rho) \textit{ coding scheme} \\ \textit{for some } (\lambda_t)_{t\in\mathcal{T}} \in \mathbb{N}^T\}.$$



(a) Storage (b) Retrieval

Fig. 1: Secure distributed storage (a) and retrieval (b) with privacy leakage for $T = 3$ servers, reconstruction threshold $\tau = 3$, privacy threshold $z = 2$, and privacy leakage parameter α. M_i is stored at Server $i \in \{1, 2, 3\}$ and created from the File F and the local randomness R available at the encoder.

For fixed T, α, τ, and z as in Definition 3, our objective in this paper is to characterize the optimal storage size $\lambda_t^\star(\alpha, z, \tau)$ at Server $t \in \mathcal{T}$, and the optimal amount of local randomness needed at the encoder $\rho^\star(\alpha, z, \tau)$.

B. Previous results

The special case $\alpha = 0$ has been studied in the literature and corresponds to ramp secret sharing [12], [13]. Specifically, by choosing $\alpha = 0$ and $z = \tau - L$, for some $L \in [\![1, \tau-1]\!]$, the problem statement of Section II-A describes a so-called (τ, L, T) ramp secret sharing scheme. Additionally, for ramp secret sharing, we have, e.g., [16], [17],

$$\rho^\star(\alpha = 0, z = \tau - L, \tau) = \frac{\tau - L}{L} H(F),$$

and the optimal size of the sum of the T shares is $\frac{T}{L}H(F)$. However, as remarked in [17], in general, one does not have $\lambda_t^\star(\alpha = 0, z = \tau - L, \tau) = \frac{1}{L}H(F), \forall t \in \mathcal{T}$, as for some $t \in \mathcal{T}$, the share size could be zero. For this reason, [17] considers linear ramp secret sharing schemes, where the leakage on the file F for a set $\mathcal{S}$ of colluding servers scales linearly with the size of $\mathcal{S}$ between $\tau - L$ to τ. In other words, a linear ramp secret sharing satisfies the condition

$$\forall \mathcal{S} \in [\mathcal{T}]^{\geqslant \tau-L+1} \cap [\mathcal{T}]^{\leqslant \tau-1}, H(F|M_\mathcal{S}) = \frac{\tau - |\mathcal{S}|}{L} H(F). \tag{A_1}$$

For such linear ramp secret sharing schemes, [17] establishes the following optimal individual share size:

$$\lambda_t^\star(\alpha = 0, z = \tau - L, \tau) = \frac{1}{L} H(F), \forall t \in \mathcal{T}.$$

Remark that the definition of linear secret sharing schemes means that a fixed value is assigned to the information leakage

at a given set of colluding servers, i.e., (A_1) can be rewritten as

$$\forall \mathcal{S} \in [\mathcal{T}]^{\geqslant \tau-L+1} \cap [\mathcal{T}]^{\leqslant \tau-1}, I(F; M_{\mathcal{S}}) = \frac{|\mathcal{S}|-(\tau-L)}{L} H(F).$$

III. Main results

A. Preliminary discussion of leakage symmetry conditions

For any $\tau \in \mathcal{T}$, $\alpha \in \mathbb{Q} \cap [0,1]$, and $z \in [\![1, \tau-1]\!]$, we will establish the optimal individual share size $\lambda_t^\star(\alpha, z, \tau)$ for any $t \in \mathcal{T}$ under the following leakage symmetry condition (A_2)

$$\forall t \in \mathcal{T}, \exists C_t \in \mathbb{R}^+, \forall \mathcal{S} \in [\mathcal{T}]^{=t}, \frac{I(F; M_{\mathcal{S}})}{H(F)} = C_t, \qquad (A_2)$$

where, by convention, we define $C_0 \triangleq 0$. Condition (A_2) means that when considering a subset of servers $\mathcal{S} \subseteq \mathcal{T}$, the privacy leakage about F, i.e., $I(F; M_{\mathcal{S}})$, must only depend on the cardinality of $\mathcal{S}$ and not the specific members in $\mathcal{S}$. Note that after normalization by $H(F)$, $\frac{I(F;M_{\mathcal{S}})}{H(F)} \in [0,1]$ for any $\mathcal{S} \subseteq \mathcal{T}$. Note also that, by (2), we must have $C_t \leqslant \alpha$ for any $t \in [\![1, z]\!]$.

In the the special case $\alpha = 0$, observe that Condition (A_2) is more general than Condition (A_1), which is reviewed in Section II-B and used to derive the optimal size of individual share for linear secret sharing schemes. Indeed, Condition (A_1) is recovered by setting $C_t \triangleq \frac{t-(\tau-L)}{L}$ for $t \in [\![\tau-L+1, \tau-1]\!]$ in Condition (A_2) with $L \triangleq \tau - z$. Hence, when $\alpha = 0$, Condition (A_2) describes a class of ramp secret sharing schemes that contains linear ramp secret sharing schemes.

Note that the leakage symmetry condition (A_2) is introduced under the term *uniform secret sharing* in [9], where the adjective uniform is used in [9] to reflect that (A_2) holds. In [9], the optimal share size is established when the constants $(C_t)_{t\in\mathcal{T}}$ in (A_2) are fixed. By contrast, in this paper, we are interested in finding the constants $(C_t)_{t\in\mathcal{T}}$ that minimize the individual share size and the necessary amount of local randomness at the encoder. To this end, we will carry an optimization over all possible secret sharing schemes that satisfy the leakage symmetry condition (A_2).

B. Results

We first establish in Theorem 1 the optimal individual share size and optimal amount of local randomness under the leakage symmetry condition (A_2). We then derive three corollaries from Theorem 1 that recover or extend known results, as outlined below.

Theorem 1. *Let $\tau \in \mathcal{T}$, $\alpha \in \mathbb{Q} \cap [0,1]$, and $z \in [\![1, \tau-1]\!]$. Suppose that the leakage symmetry condition (A_2) holds. Then, for any $t \in \mathcal{T}$, we have*

$$\frac{\lambda_t^\star(\alpha, z, \tau)}{H(F)} = \max\left(\frac{1-\alpha}{\tau - z}, \frac{1}{\tau}\right) = \begin{cases} \frac{1-\alpha}{\tau-z} & \text{if } \alpha < \frac{z}{\tau} \\ \frac{1}{\tau} & \text{if } \alpha \geqslant \frac{z}{\tau} \end{cases},$$

$$\frac{\rho^\star(\alpha, z, \tau)}{H(F)} = \frac{[z - \tau\alpha]^+}{\tau - z} = \begin{cases} \frac{z-\tau\alpha}{\tau-z} & \text{if } \alpha < \frac{z}{\tau} \\ 0 & \text{if } \alpha \geqslant \frac{z}{\tau} \end{cases}.$$

Proof. The achievability proof of Theorem 1 relies on ramp secret sharing and is omitted due to space constraints. The converse proof of Theorem 1 is presented in Section IV. ∎

Corollary 1. *Assume that the privacy leakage is $\alpha = 0$ and the privacy threshold is $z = \tau - 1$. Observe from (1) and (2) that, in this case, Condition (A_2) is always satisfied, in particular, $C_t = 0$ when $t \in [\![1, \tau-1]\!]$, and $C_t = 1$ when $t \in [\![\tau, T]\!]$. Then, by Theorem 1, we have*

$$\frac{\lambda_t^\star(\alpha, z, \tau)}{H(F)} = 1, \forall t \in \mathcal{T}$$
$$\frac{\rho^\star(\alpha, z, \tau)}{H(F)} = \tau - 1.$$

Hence, we recover the well-known fact, e.g., [8, Th. 1], that the optimal share size is the entropy of F for perfect threshold secret sharing, first introduced in [5], [6]. Note that, in this special case, $\frac{\lambda_t^\star(\alpha,z,\tau)}{H(F)}$, $t \in \mathcal{T}$, is independent of τ, which is not true in general.

Corollary 2. *Suppose that the leakage symmetry condition (A_2) holds. Assume that the privacy leakage is $\alpha = 0$ and the privacy threshold is $z = \tau - L$, for some $L \in [\![1, \tau-1]\!]$. Then, by Theorem 1, we have*

$$\frac{\lambda_t^\star(\alpha, z, \tau)}{H(F)} = \frac{1}{L}, \forall t \in \mathcal{T}$$
$$\frac{\rho^\star(\alpha, z, \tau)}{H(F)} = \frac{\tau - L}{L}.$$

Hence, it recovers the result in [16], [17], for (τ, L, T) linear ramp secret sharing schemes, i.e., secret sharing schemes that satisfy Condition (A_1), and generalizes it to the larger class of uniform secret sharing schemes, i.e., secret sharing schemes that satisfy Condition (A_2). The result can also be interpreted as follows: Among all uniform secret sharing schemes, linear secret sharing schemes are optimal in terms of individual share size and local randomness necessary at the encoder.

Corollary 3. *Assume that the reconstruction threshold is $\tau = T$, the privacy threshold is $z = T - 1$, and the Condition (A_2) holds. Then,*

$$\frac{\lambda_t^\star(\alpha, z, \tau)}{H(F)} = \begin{cases} 1-\alpha & \text{if } \alpha < 1 - \frac{1}{T} \\ \frac{1}{T} & \text{if } \alpha \geqslant 1 - \frac{1}{T} \end{cases}, \forall t \in \mathcal{T}$$

$$\frac{\rho^\star(\alpha, z, \tau)}{H(F)} = \begin{cases} T(1-\alpha) - 1 & \text{if } \alpha < 1 - \frac{1}{T} \\ 0 & \text{if } \alpha \geqslant 1 - \frac{1}{T} \end{cases},$$

which recovers the results found in [18] and generalize them to the case where the shares are not assumed to be of equal size in the problem statement.

IV. Converse proof of Theorem 1

Under the leakage symmetry Condition (A_2), we prove lower bounds on the individual share size and the necessary amount of local randomness at the encoder in Sections IV-A and IV-B, respectively.

A. Lower bound on individual share size

Let $\tau \in \mathcal{T}$, $\alpha \in [0,1]$, $z \in [\![1, \tau-1]\!]$, and consider an (α, z)-private $(\tau, (\lambda_t)_{t\in\mathcal{T}}, \rho)$ coding scheme for some $(\lambda_t)_{t\in\mathcal{T}} \in \mathbb{N}^T$, $\rho \in \mathbb{N}$, as defined in Definition 2 under the leakage symmetry Condition (A_2). In Sections IV-A1 and IV-A2, we prove that for any $t \in \mathcal{T}$,

$$\frac{\lambda_t}{H(F)} \geqslant \frac{1-\alpha}{\tau - z}, \tag{3}$$

and

$$\frac{\lambda_t}{H(F)} \geqslant \frac{1}{\tau}, \tag{4}$$

respectively. We will thus deduce from (3) and (4) that for any $t \in \mathcal{T}$,

$$\frac{\lambda_t}{H(F)} \geqslant \max\left(\frac{1-\alpha}{\tau - z}, \frac{1}{\tau}\right) = \begin{cases} \frac{1-\alpha}{\tau-z} & \text{if } \alpha < \frac{z}{\tau} \\ \frac{1}{\tau} & \text{if } \alpha \geqslant \frac{z}{\tau} \end{cases}. \tag{5}$$

1) Proof of the first lower bound (3) *on* λ_t*:* Fix $t \in \mathcal{T}$. For $i \in [\![z, \tau-1]\!]$, define $\mathcal{S}_i \triangleq \begin{cases} [\![1, i]\!] & \text{if } t > i \\ [\![1, i+1]\!]\backslash\{t\} & \text{if } t \leqslant i \end{cases}$ and $\mathcal{S}_\tau \triangleq \mathcal{S}_{\tau-1} \cup \{t\}$. Then, for $i \in [\![z+1, \tau-1]\!]$, we have

$$\begin{aligned} &H(M_t|M_{\mathcal{S}_i}) \\ &\overset{(a)}{=} H(M_tF|M_{\mathcal{S}_i}) - H(F|M_{\mathcal{S}_i}M_t) \\ &\overset{(b)}{=} H(F|M_{\mathcal{S}_i}) + H(M_t|FM_{\mathcal{S}_i}) - H(F|M_{\mathcal{S}_i}M_t) \qquad (6) \\ &\overset{(c)}{=} (1-C_i)H(F) + H(M_t|FM_{\mathcal{S}_i}) - (1-C_{i+1})H(F) \\ &= (C_{i+1}-C_i)H(F) + H(M_t|FM_{\mathcal{S}_i}), \qquad (7) \end{aligned}$$

where
(a) and (b) hold by the chain rule;
(c) holds for some constants C_i and C_{i+1} by (A_2).
Next, we have

$$\begin{aligned} &H(M_t|M_{\mathcal{S}_\tau}) \\ &\overset{(a)}{=} H(F|M_{\mathcal{S}_\tau}) + H(M_t|FM_{\mathcal{S}_\tau}) - H(F|M_{\mathcal{S}_\tau}M_t) \\ &\overset{(b)}{=} H(M_t|FM_{\mathcal{S}_\tau}) \\ &\overset{(c)}{=} (C_{\tau+1}-C_\tau)H(F) + H(M_t|FM_{\mathcal{S}_\tau}), \qquad (8) \end{aligned}$$

where
(a) holds as in (6);
(b) holds by (1);
(c) holds by defining $C_{\tau+1} \triangleq C_\tau = 1$.
We also have

$$\begin{aligned} &H(M_t|M_{\mathcal{S}_z}) \\ &\overset{(a)}{=} H(F|M_{\mathcal{S}_z}) + H(M_t|FM_{\mathcal{S}_z}) - H(F|M_{\mathcal{S}_z}M_t) \\ &\overset{(b)}{\geqslant} (1-\alpha)H(F) + H(M_t|FM_{\mathcal{S}_z}) - (1-C_{z+1})H(F) \\ &= (C_{z+1}-\alpha)H(F) + H(M_t|FM_{\mathcal{S}_z}), \qquad (9) \end{aligned}$$

where
(a) holds as in (6);
(b) holds by (2) and (A_2).
In the following, for convenience, we define $C_z \triangleq \alpha$. Next, we have

$$\begin{aligned} &H(M_t) \\ &\overset{(a)}{\geqslant} H(M_t|M_{\mathcal{S}_z}) \qquad (10) \\ &\overset{(b)}{=} H(M_t|M_{\mathcal{S}_z}) - H(M_t|M_{\mathcal{S}_\tau}) \\ &= \sum_{i=z}^{\tau-1} \big(H(M_t|M_{\mathcal{S}_i}) - H(M_t|M_{\mathcal{S}_{i+1}})\big) \\ &\overset{(c)}{\geqslant} \sum_{i=z}^{\tau-1} \big[(C_{i+1}-C_i)H(F) + H(M_t|FM_{\mathcal{S}_i}) \\ &\qquad -(C_{i+2}-C_{i+1})H(F) - H(M_t|FM_{\mathcal{S}_{i+1}})\big]^+ \\ &\overset{(d)}{\geqslant} H(F)\sum_{i=z}^{\tau-1}[(C_{i+1}-C_i) - (C_{i+2}-C_{i+1})]^+ \\ &\overset{(e)}{=} H(F)\sum_{i=z+1}^{\tau}[(\phi(i)-\phi(i-1)) - (\phi(i+1)-\phi(i))]^+ \\ &\overset{(f)}{\geqslant} H(F)\min_{\phi\in\mathcal{F}}\sum_{i=z+1}^{\tau}[(\phi(i)-\phi(i-1)) - (\phi(i+1)-\phi(i))]^+ \qquad (11) \end{aligned}$$

where
(a) holds because conditioning reduces entropy;
(b) holds because $t \in \mathcal{S}_\tau$;
(c) holds because $H(M_t|M_{\mathcal{S}_i}) - H(M_t|M_{\mathcal{S}_{i+1}}) \geqslant 0$ (conditioning reduces entropy and $\mathcal{S}_i \subset \mathcal{S}_{i+1}$) and by (7), (8), (9);
(d) holds because conditioning reduces entropy;
(e) holds with the function $\phi : [\![z, \tau+1]\!] \to [0,1]$ defined such that $\phi(i) = C_i$ for $i \in [\![z, \tau+1]\!]$;
(f) holds with the minimum taken over the set $\mathcal{F}$ of all the functions $\phi : [\![z, \tau+1]\!] \to [0,1]$ that are non-decreasing (by (A_2) because for any $\mathcal{S} \subset \mathcal{S}' \subset \mathcal{T}$, $\frac{I(F;M_\mathcal{S})}{H(F)} \leqslant \frac{I(F;M_{\mathcal{S}'})}{H(F)}$) and such that $\phi(z) = \alpha$ (because $C_z = \alpha$), $\phi(\tau+1) = \phi(\tau) = 1$ (because $C_{\tau+1} = C_\tau = 1$).

We now lower bound the minimum in the right-hand side of (11). Let $\phi \in \mathcal{F}$ and let ϕ^+ be the concave envelope of ϕ over $[\![z, \tau+1]\!]$, i.e., for $i \in [\![z, \tau+1]\!]$, $\phi^+(i) \triangleq \min\{\psi(i) : \psi \geqslant \phi, \psi \text{ is concave}\}$. Note that $\phi^+(z) = \phi(z)$ and $\phi^+(\tau+1) = \phi(\tau+1)$.

Next, we have

$$\begin{aligned} &\sum_{i=z+1}^{\tau}[(\phi(i)-\phi(i-1)) - (\phi(i+1)-\phi(i))]^+ \\ &\overset{(a)}{\geqslant} \sum_{i=z+1}^{\tau}(\phi^+(i)-\phi^+(i-1)) - (\phi^+(i+1)-\phi^+(i))) \qquad (12) \\ &= \phi^+(z+1) - \phi^+(z) + \phi^+(\tau+1) - \phi^+(\tau) \\ &\overset{(b)}{=} \phi^+(z+1) - \phi^+(z) \\ &\overset{(c)}{\geqslant} \frac{1-\alpha}{\tau-z}, \qquad (13) \end{aligned}$$

where

(a) can be proved as in [19], details are omitted due to space constraints;
(b) holds because $\phi^+(\tau+1)=\phi^+(\tau)=1$;
(c) holds because $\phi^+(z+1)-\phi^+(z) \geqslant (\phi^+(\tau)-\phi^+(z))/(\tau-z)$ by concavity of ϕ^+ and where we have used that $\phi^+(\tau)=1$ and $\phi^+(z)=\phi(z)=\alpha$.

Finally, we have

$$\begin{aligned}\lambda_t &\geqslant H(M_t)\\ &\geqslant H(F)\frac{1-\alpha}{\tau-z},\end{aligned}$$

where the last inequality holds by (11) and (13), which is valid for any $\phi \in \mathcal{F}$.

2) Proof of the second lower bound (4) *on* λ_t*:* Note that in the proof of (3), one can substitute the variable z by zero such that one can show

$$\begin{aligned}\lambda_t &\geqslant H(M_t)\\ &\geqslant H(F)(\phi^+(1)-\phi^+(0))\\ &\geqslant H(F)\frac{1}{\tau},\end{aligned}$$

where the last inequality holds because $\phi^+(1)-\phi^+(0) \geqslant (\phi^+(\tau)-\phi^+(0))/\tau$ by concavity of ϕ^+ and where we have used that $\phi^+(\tau)=1$ and $\phi^+(0)=0$.

B. Lower bound on the amount of local randomness

Let $\tau \in \mathcal{T}$, $\alpha \in [0,1]$, $z \in [\![1, \tau-1]\!]$, and consider an (α, z)-private $(\tau, (\lambda_t)_{t\in\mathcal{T}}, \rho)$ coding scheme for some $(\lambda_t)_{t\in\mathcal{T}} \in \mathbb{N}^T$, $\rho \in \mathbb{N}$, as defined in Definition 2 under the leakage symmetry Condition (A_2). Then, we have

$$\begin{aligned}
&\rho + H(F)\\
&\overset{(a)}{=} H(R) + H(F)\\
&\overset{(b)}{=} H(RF)\\
&\overset{(c)}{\geqslant} H(M_{\mathcal{T}})\\
&\overset{(d)}{=} H(M_{[\![1,z]\!]}) + H(M_{\mathcal{T}\setminus[\![1,z]\!]}|M_{[\![1,z]\!]})\\
&\overset{(e)}{=} \sum_{t=1}^{z} H(M_t|M_{[\![1,t-1]\!]}) + H(M_{\mathcal{T}\setminus[\![1,z]\!]}|M_{[\![1,z]\!]})\\
&\overset{(f)}{\geqslant} \sum_{t=1}^{z} H(M_t|M_{\mathcal{S}_{z,t}}) + H(M_{\mathcal{T}\setminus[\![1,z]\!]}|M_{[\![1,z]\!]})\\
&\overset{(g)}{\geqslant} z\frac{1-\alpha}{\tau-z}H(F) + H(M_{\mathcal{T}\setminus[\![1,z]\!]}|M_{[\![1,z]\!]})\\
&\overset{(h)}{=} z\frac{1-\alpha}{\tau-z}H(F) + H(M_{\mathcal{T}\setminus[\![1,z]\!]}F|M_{[\![1,z]\!]})\\
&\quad - H(F|M_{\mathcal{T}\setminus[\![1,z]\!]}M_{[\![1,z]\!]})\\
&\overset{(i)}{\geqslant} z\frac{1-\alpha}{\tau-z}H(F) + H(F|M_{[\![1,z]\!]})\\
&\overset{(j)}{\geqslant} z\frac{1-\alpha}{\tau-z}H(F) + (1-\alpha)H(F)\\
&= \tau\frac{1-\alpha}{\tau-z}H(F), \qquad (14)
\end{aligned}$$

where

(a) holds by uniformity of R;
(b) holds by independence between F and R;
(c) holds because $M_{\mathcal{T}}$ is a deterministic function of (R,F);
(d) and (e) hold by the chain rule;
(f) holds because conditioning reduces entropy and we have defined $\mathcal{S}_{z,t} \triangleq [\![1, z+1]\!]\setminus\{t\}$ for $t \in [\![1,z]\!]$;
(g) holds because for any $t \in [\![1,z]\!]$, $H(M_t|M_{\mathcal{S}_{z,t}}) \geqslant \frac{1-\alpha}{\tau-z}H(F)$, by the converse proof of Theorem 1 starting from (10);
(h) holds by the chain rule;
(i) holds because $H(F|M_{\mathcal{T}\setminus[\![1,z]\!]}M_{[\![1,z]\!]}) = H(F|M_{\mathcal{T}}) = 0$ by (1), and $H(M_{\mathcal{T}\setminus[\![1,z]\!]}F|M_{[\![1,z]\!]}) \geqslant H(M_{\mathcal{T}\setminus[\![1,z]\!]}|M_{[\![1,z]\!]})$ by the chain rule and positivity of conditional entropy;
(j) holds by (2).

Finally, from (14), we have

$$\begin{aligned}\rho &\geqslant \left(\tau\frac{1-\alpha}{\tau-z}-1\right)H(F)\\ &= \frac{z-\tau\alpha}{\tau-z}H(F),\end{aligned}$$

and since we also have $\rho \geqslant 0$, we conclude

$$\frac{\rho}{H(F)} \geqslant \frac{[z-\tau\alpha]^+}{\tau-z} = \begin{cases}\frac{z-\tau\alpha}{\tau-z} & \text{if } \alpha < \frac{z}{\tau}\\ 0 & \text{if } \alpha \geqslant \frac{z}{\tau}\end{cases}.$$

V. Concluding remarks

We considered a setting where a file must be stored in L servers such that: (i) any τ servers that pool their information together can reconstruct the file, and (ii) any z servers cannot learn more than a fraction $\alpha \in [0,1]$ of the file, where τ, z, and α are parameters to be chosen by the users. This setting generalizes ramp secret sharing in that information leakage about the file is allowed up to a fraction α, and goes beyond existing works on uniform secret sharing by considering share size optimization over a set of access functions rather than for a fixed access function. Specifically, for given parameters τ, z, α, and under the leakage symmetry assumption that any set of colluding servers must have the same information leakage about the file that any other set of colluding servers of same size, we derived the optimal individual share size at each server. As a byproduct, in the case $\alpha = 0$, our results prove that among all uniform secret sharing schemes for our model, linear ramp secret sharing schemes require the smallest individual share size at the servers.

REFERENCES

[1] J. A. Garay, R. Gennaro, C. Jutla, and T. Rabin, "Secure distributed storage and retrieval," *Theoretical Computer Science*, vol. 243, no. 1-2, pp. 363–389, 2000.

[2] A. S. Rawat, O. O. Koyluoglu, and S. Vishwanath, "Centralized repair of multiple node failures with applications to communication efficient secret sharing," *IEEE Transactions on Information Theory*, vol. 64, no. 12, pp. 7529–7550, 2018.

[3] R. Bitar and S. El Rouayheb, "Staircase codes for secret sharing with optimal communication and read overheads," *IEEE Transactions on Information Theory*, vol. 64, no. 2, pp. 933–943, 2018.

[4] W. Huang, M. Langberg, J. Kliewer, and J. Bruck, "Communication efficient secret sharing," *IEEE Transactions on Information Theory*, vol. 62, no. 12, pp. 7195–7206, 2016.

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

[6] G. R. Blakley, "Safeguarding cryptographic keys," in *Proceedings of the national computer conference*, vol. 48, 1979, pp. 313–317.

[7] A. Beimel, "Secret-sharing schemes: A survey," in *International Conference on Coding and Cryptology*. Springer, 2011, pp. 11–46.

[8] E. Karnin, J. Greene, and M. Hellman, "On secret sharing systems," *IEEE Transactions on Information Theory*, vol. 29, no. 1, pp. 35–41, 1983.

[9] M. Yoshida, T. Fujiwara, and M. P. Fossorier, "Optimal uniform secret sharing," *IEEE Transactions on Information Theory*, vol. 65, no. 1, pp. 436–443, 2018.

[10] O. Farràs, T. Hansen, T. Kaced, and C. Padró, "Optimal non-perfect uniform secret sharing schemes," in *Annual Cryptology Conference*. Springer, 2014, pp. 217–234.

[11] O. Farràs, T. B. Hansen, T. Kaced, and C. Padró, "On the information ratio of non-perfect secret sharing schemes," vol. 79, no. 4. Springer, 2017, pp. 987–1013.

[12] H. Yamamoto, "Secret sharing system using (k, l, n) threshold scheme," *Electronics and Communications in Japan (Part I: Communications)*, vol. 69, no. 9, pp. 46–54, 1986.

[13] G. R. Blakley and C. Meadows, "Security of ramp schemes," in *Workshop on the Theory and Application of Cryptographic Techniques*. Springer, 1984, pp. 242–268.

[14] K. Yoneyama, N. Kunihiro, B. Santoso, and K. Ohta, "Non-linear function ramp scheme," in *Proc. Int. Symp. Inf. Theory Appl.(ISITA)*, 2004, pp. 788–793.

[15] M. Yoshida and T. Fujiwara, "Secure construction for nonlinear function threshold ramp secret sharing," in *Proc. of IEEE Int. Symp. Inf. Theory*, 2007, pp. 1041–1045.

[16] C. Blundo, A. De Santis, and U. Vaccaro, "Randomness in distribution protocols," *Information and Computation*, vol. 131, no. 2, pp. 111–139, 1996.

[17] C. Blundo, A. D. Santis, and U. Vaccaro, "Efficient sharing of many secrets," in *Annual Symposium on Theoretical Aspects of Computer Science*. Springer, 1993, pp. 692–703.

[18] R. A. Chou and J. Kliewer, "Secure distributed storage: Rate-privacy trade-off and XOR-based coding scheme," in *IEEE International Symposium on Information Theory (ISIT)*, 2020, pp. 605–610.

[19] R. A. Chou, "Quantifying the cost of privately storing data in distributed storage systems," *IEEE Transactions on Information Theory*, vol. 68, no. 11, pp. 7485–7499, 2022.