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

Rémi A. Chou, Jörg Kliewer, Fellow, IEEE

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, 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

Centralized data storage of sensitive information means compromising the entirety of the data in the case of a data breach. By contrast, well-known distributed storage strategies, where data are stored in multiple servers, can offer resilience against data breaches at a subset of servers and avoid having a single point of failure. Secure distributed storage schemes, e.g., [2]-[5], often rely on the idea of secret sharing as introduced in [6], [7] – we refer to [8] 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., [9], 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 Fneeds to be stored at T servers. The file must be recoverable from τ servers, and needs to remain private from any zcolluding 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 system designer, and require that no more than a fraction α of the file can be learned by a set of z colluding servers. As a function of the parameters (τ, z, α) , we establish the optimal sum of the share sizes and the optimal amount of local randomness needed at the encoder. 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 satisfy such a leakage symmetry are also referred to as uniform secret sharing schemes, e.g., [10], [11].

A. Previous work

Secret sharing was first introduced in [6], [7] and provides perfect security in that any set of colluding participants that is not allowed to reconstruct the secret cannot learn, in an information-theoretic sense, any information about the secret. With the objective to reduce the size of the participants' shares, ramp secret sharing has then been introduced in [12], [13] to relax the security guarantees of secret sharing schemes. Specifically, in [12], [13], any set of colluding users with size smaller than some parameter z cannot learn any information about the secret, any set of colluding users with size larger than or equal to some parameter τ can reconstruct the secret, and any set of colluding users with size stricity larger than z but strictly smaller than τ can learn part of the secret. Additionally, for this last type of set of colluding users, the information leakage about the secret grows linearly with its size. Later, this idea to reduce the size of the shares by allowing information leakage was generalized to non-linear access function, e.g., [14], [15], and further studied under the term non-perfect secret sharing, e.g., [10], [11], [16].

Performance metrics of interest for any secret sharing include the characterization of the optimal size of the shares. More specifically, characterizing the optimal sum of the share sizes, the optimal maximal share size, or the optimal individual share sizes for various secret sharing settings has been the subject of intense research, e.g., [9]-[11], [17]-[22], we also refer to [8] for a survey of known results. For instance, optimal individual share sizes have been characterized for the original secret sharing setting [6], [7], e.g., [9], and the optimal sum of the share sizes for ramp secret sharing has also been fully characterized, e.g., [23]. Unfortunately, even for the relatively simple case of ramp secret sharing the full characterization of optimal individual share sizes is unknown and a challenging problem. A partial solution is proposed in [23] by restricting the class of ramp secret sharing scheme to the class of linear ramp secret schemes, which adds symmetry to the problem. More recently, a more general and

R. Chou is with the Department of Computer Science and Engineering, The University of Texas at Arlington, Arlington, TX. J. Kliewer is with the Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ. This work was supported in part by NSF grants CCF-2201824 and CCF-2201825. E-mails: remi.chou@uta.edu, jkliewer@njit.edu. A preliminary version of this work has been presented at the 2020 IEEE International Symposium on Information Theory (ISIT) [1].

natural definition of symmetry is introduced in the form of uniform secret sharing [10], [11], [24]. In these works, the authors characterize the optimal individual share sizes obtained under such a symmetry condition, which requires that the information leakage only depends on the size of the set of colluding participants, and not on the specific identities of the participants in this set.

B. Comparison to previous work

Our setting not only considers generalizations of previous secret sharing settings but also introduces a fundamentally different view to study the trade-off between storage needs and privacy leakage. Specifically, a major difference between the study of uniform secret sharing in [10], [11] and our work is that, in [10], [11], 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 , and the other point indicates a maximum number of colluding servers z. All the other points of the access function are optimized to minimize the share sizes. This difference introduces a non-trivial optimization problem over a set of access functions to determine optimal individual share sizes. We show that this optimization reduces to maximizing the sum of consecutive gradients of an access function over the set of all possible access functions that satisfy our problem constraints, i.e., it must be less than or equal to α in point z and equal to one in point τ . The crux to solve this optimization is to introduce the concave envelopes of the access functions to show that piecewise linear access functions are solutions to the optimization.

We note that the idea of trading storage space against information leakage is also closely related to non-perfect secret sharing [10], [11], [25], 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.

While the above remark on non-perfect secret sharing applies ramp secret sharing schemes, e.g., [12], [13], we highlight a new result for ramp secret sharing that follows from our main results. Specifically, when the privacy parameter is $\alpha = 0$, i.e., perfect privacy is required against z colluding servers, our results prove that among all uniform ramp secret sharing schemes, which represents a more general class of secret sharing schemes than linear ramp secret sharing schemes, the ones that have a piecewise linear access function have the minimum individual share sizes. This result had also not been previously proved in the literature.

C. Paper organization

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 Sections IV, V to streamline presentation. Finally, we provide concluding remarks in Section VI.

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 $[\![a, b]\!] \triangleq$ $[\lfloor a \rfloor, [b]\!] \cap \mathbb{N}$ and $[a]^+ \triangleq \max(0, a)$. For two arbitrary sets Sand \mathcal{T} , a sequence of elements $x_t \in S$, $t \in \mathcal{T}$, indexed by the set \mathcal{T} is written as $(x_t)_{t\in\mathcal{T}}$.

A. Problem statement

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

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, which is represented by a random binary sequence with finite length;
- Local randomness in the form of a sequence R of ρ bits uniformly distributed over {0,1}^ρ 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 $S \subseteq \mathcal{T}$ of servers, we use the notation $M_S \triangleq (M_t)_{t \in S}$.
- For any subset $S \subseteq T$ such that $|S| \ge \tau$, a decoder

$$d_{\mathcal{S}}: \underset{t \in \mathcal{S}}{\times} \{0, 1\}^{\lambda_t} \to \{0, 1\}^{|F|}, M_{\mathcal{S}} \mapsto \hat{F}(\mathcal{S}),$$

which takes as input M_{S} and outputs $\hat{F}(S)$, an estimate of F.

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

$$\max_{\mathcal{S}\in[\mathcal{T}]^{\geq\tau}} H(F|\hat{F}(\mathcal{S})) = 0, \ (Recoverability)$$
(1)

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

Requirement (1) means that any subset of τ or more servers can reconstruct the file F. Note that Requirement (1) implies $\max_{S \in [\mathcal{T}] \geq \tau} H(F|M_S) = 0$. 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 $(T, \tau, z) = (3, 3, 2)$.









(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.

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 $|\alpha - \beta| \leq \epsilon$.

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

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

 $\triangleq \min\{\rho \in \mathbb{N} : there \ exists \ an \ (\alpha, z)-private$

$$(\tau, (\lambda_t)_{t \in \mathcal{T}}, \rho)$$
 coding scheme for some $(\lambda_t)_{t \in \mathcal{T}} \in \mathbb{N}^T$.

For fixed T, α , τ , and z as in Definition 3, our objective in this paper is to characterize the optimal storage size $\lambda_t^*(\alpha, z, \tau)$ at Server $t \in \mathcal{T}$, the optimal total storage size $\lambda_{sum}^*(\alpha, z, \tau)$, and the optimal amount of local randomness needed at the encoder $\rho^*(\alpha, z, \tau)$. Note that it is a priori unclear whether there exists a coding scheme that can simultaneously achieve $\lambda_t^*(\alpha, z, \tau), t \in \mathcal{T}, \lambda_{sum}^*(\alpha, z, \tau)$, and $\rho^*(\alpha, z, \tau)$. However, our results will prove that such a coding scheme exists.

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., [23], [26],

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

As remarked in [23], in general, one does not have $\lambda_t^*(\alpha = 0, z = \tau - L, \tau) = \frac{1}{L}H(F), \forall t \in \mathcal{T}$, as for some $t \in \mathcal{T}$, the share size could be zero. For this reason, [23] considers linear ramp secret sharing schemes, where the leakage on the file F for a set S of colluding servers scales linearly with the size of S between $\tau - L$ to τ . In other words, a linear ramp secret sharing satisfies the condition

$$\forall \mathcal{S} \in [\mathcal{T}]^{\geqslant \tau - L + 1} \cap [\mathcal{T}]^{\leqslant \tau - 1}, H(F|M_{\mathcal{S}}) = \frac{\tau - |\mathcal{S}|}{L} H(F).$$
(A1)

For such linear ramp secret sharing schemes, [23, Th. 3.3] establishes the following optimal individual share size:

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

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

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

C. Discussion of leakage symmetry conditions used in previous work

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 $\lambda_t^*(\alpha, z, \tau)$ for any $t \in \mathcal{T}$ under the following leakage symmetry condition (A_2)

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

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

In 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.

Note that the leakage symmetry condition (A_2) is introduced under the term *uniform secret sharing* in [10], where the adjective uniform is used in [10] to reflect that (A_2) holds. In [10], 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) . Another difference between our study and [10] is that our study extends [10] in the following aspects: for $\alpha \neq 0$, we study the optimal sum of the share sizes at all the servers and the optimal amount of local randomness required at the encoder in the absence of any leakage symmetry condition.

III. MAIN 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

$$\frac{\lambda_t^{\star}(\alpha, z, \tau)}{H(F)} = \max\left(\frac{1-\alpha}{\tau-z}, \frac{1}{\tau}\right) = \begin{cases} \frac{1-\alpha}{\tau-z} & \text{if } \alpha < \frac{z}{\tau} \\ \frac{1}{\tau} & \text{if } \alpha \geqslant \frac{z}{\tau} \end{cases},\\ \frac{\rho^{\star}(\alpha, z, \tau)}{H(F)} = \frac{[z-\tau\alpha]^+}{\tau-z} = \begin{cases} \frac{z-\tau\alpha}{\tau-z} & \text{if } \alpha < \frac{z}{\tau} \\ 0 & \text{if } \alpha \geqslant \frac{z}{\tau} \end{cases}.$$

Moreover, there exists an (α, z) -private $(\tau, (\lambda_t^*(\alpha, z, \tau))_{t \in \mathcal{T}}, \rho^*(\alpha, z, \tau))$ coding scheme, i.e., $\lambda_t^*(\alpha, z, \tau)$, $t \in \mathcal{T}$, and $\rho^*(\alpha, z, \tau)$ can simultaneously be achieved by a single coding scheme.

Proof. The achievability proof of Theorem 1 is detailed in Section IV. The converse proof of Theorem 1 is presented in Section V.

As expected, Theorem 1 shows that allowing information leakage, controlled by the parameter α , enables a reduction of the individual share size and amount of local randomness needed at the encoder. Theorem 1 also shows the existence of a threshold with respect to α . Specifically, when $\alpha \ge z/\tau$, then a (τ, τ, T) ramp secret sharing is sufficient to achieve the optimal share size $1/\tau$, since, in that case, the share size of any z colluding users is z/τ and the privacy condition is immediately satisfied. **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 recover the well-known fact, e.g., [9, Th. 1], that the optimal share size is the entropy of F for perfect threshold secret sharing, first introduced in [6], [7].

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, Theorem 1 recovers the result in [23], [26], 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, Theorem 1 recovers the results found in [1] and generalizes them to the case where the shares are not assumed to be of equal size in the problem statement.

We numerically illustrate Theorem 1 in the following example.

Example. For the case of T = 12 servers and a reconstruction threshold $\tau = 7$, we depict in Figures 2 and 3, $\frac{\lambda_t^*(\alpha, z, \tau)}{H(F)}$, $t \in \mathcal{T}$, and $\frac{\rho^*(\alpha, z, \tau)}{H(F)}$ obtained in Theorem 1, respectively, as functions of the privacy leakage parameter α and the privacy threshold z.

Note that, in the absence of any leakage symmetry condition, the minimum size of an individual share could be zero. For instance, using the notation of Section II-B, one can construct a $(\tau, \tau - z, T)$ ramp secret sharing scheme with Tparticipants, where the share size of $\tau - z - 1$ participants is zero as follows: Consider a $(z + 1, 1, T - (\tau - z - 1))$ ramp secret sharing scheme with $T - (\tau - z - 1)$ participants and consider $\tau - z - 1$ additional participants to whom we do not give any share. This shows that Theorem 1 does not hold in the absence of Condition (A_2) .

Then, beyond the optimal individual share sizes, we also study the optimal sum of the share sizes in the absence of any leakage symmetry condition. Specifically, in this case, we establish the optimal sum of the share sizes $\lambda_{sum}^{\star}(\alpha, z, \tau)$ and the optimal amount of local randomness $\rho^{\star}(\alpha, z, \tau)$, both defined in Definition 3.

Theorem 2. Let $\tau \in \mathcal{T}$, $\alpha \in \mathbb{Q} \cap [0,1]$, and $z \in \llbracket 1, \tau - 1 \rrbracket$. We have

$$\frac{\lambda_{\text{sum}}^{\star}(\alpha, z, \tau)}{H(F)} = T \max\left(\frac{1-\alpha}{\tau-z}, \frac{1}{\tau}\right) = \begin{cases} T\frac{1-\alpha}{\tau-z} & \text{if } \alpha < \frac{z}{\tau} \\ \frac{T}{\tau} & \text{if } \alpha \geqslant \frac{z}{\tau} \end{cases},\\ \frac{\rho^{\star}(\alpha, z, \tau)}{H(F)} = \frac{[z-\tau\alpha]^{+}}{\tau-z} = \begin{cases} \frac{z-\tau\alpha}{\tau-z} & \text{if } \alpha < \frac{z}{\tau} \\ 0 & \text{if } \alpha \geqslant \frac{z}{\tau} \end{cases}.$$



Fig. 2: $\frac{\lambda_t^*(\alpha, z, \tau)}{H(F)}$, $t \in \mathcal{T}$, when T = 12, $\tau = 7$, and the privacy threshold belongs to $[\![1, 6]\!]$. The bold blue circle corresponds to the optimal share size for Shamir's secret sharing, as reviewed in Corollary 1.



Fig. 3: $\frac{\rho^*(\alpha, z, \tau)}{H(F)}$ when T = 12, $\tau = 7$, and z belongs to [1, 6]. The bold blue circle corresponds to the optimal amount of necessary randomness at the encoder for Shamir's secret sharing, as reviewed in Corollary 1.

Moreover, there exists an (α, z) -private $(\tau, (\lambda_t)_{t \in T}, \rho^*(\alpha, z, \tau))$ coding scheme such that $\sum_{t \in T} \lambda_t = \lambda^*_{sum}(\alpha, z, \tau)$, i.e., $\lambda^*_{sum}(\alpha, z, \tau)$ and $\rho^*(\alpha, z, \tau)$ can simultaneously be achieved by a single coding scheme.

Proof. The achievability part follows from the achievability part of Theorem 1 by summing the sizes of the T shares and the converse part is proved in Appendix B.

Note that the converse proof of Theorem 2 is of combinatorial nature and different from the proof of Theorem 1, which involves an optimization problem. Note also that Theorems 1 and 2 indicate that there is no gain, in terms of necessary local randomness at the encoder, between an optimization over secret sharing schemes that satisfy Condition (A_2) and any secret sharing schemes that do not.

IV. ACHIEVABILITY PROOF OF THEOREM 1

Let $\alpha \in \mathbb{Q} \cap [0, 1]$. We consider the two cases $\alpha \ge z/\tau$ and $\alpha < z/\tau$ in Sections IV-B1 and IV-B2, respectively. We first review some definitions in Section IV-A.

A. Preliminary definitions

1) Access function: For a coding scheme as in Definition 2 that satisfies the leakage symmetry Condition (A_2) , one can define an access function, e.g., [11], which fully describes the leakage of any set of shares. More specifically, with the notation of Definition 2, define the access function of a coding scheme as

$$g: \llbracket 0, T \rrbracket \to [0, 1], t \mapsto C_t.$$

Note that, by (A_2) , the leakage of any set S of shares only depends on the cardinality of S. Hence, it is sufficient to consider g to fully describe the leakage of any set of shares. Note also that, by (1), the reconstruction threshold τ implies that g(t) = 1 for any $t \in [[\tau, T]]$, and, by (2), the privacy threshold z implies that $g(z) \leq \alpha$. Finally, note that by definition of C_t , $t \in \mathcal{T}$, in (A_2) , g is non-decreasing.

2) Ramp secret sharing: Consider $\tau, T, L \in \mathbb{N}$ such that $1 \leq L < \tau \leq T$. A (τ, L, T) linear ramp secret sharing scheme, e.g., [12], [13], is a coding scheme as in Definition 2 with $\alpha = 0, z = \tau - L$, and with access function

$$g: \llbracket 0, T \rrbracket \to [0, 1], t \mapsto \begin{cases} 0 & \text{if } t \in \llbracket 0, \tau - L \rrbracket \\ \frac{t - \tau + L}{L} & \text{if } t \in \llbracket \tau - L + 1, \tau \rrbracket \\ 1 & \text{if } t \in \llbracket \tau + 1, T \rrbracket \end{cases}.$$

In particular, any τ shares can reconstruct F, any set of shares less than or equal to $\tau - L$ does not leak any information about F, and for sets of shares with cardinality in $[[\tau - L + 1, \tau]]$, the leakage increases linearly with the set cardinality.

B. Achievability proof

The first step of the achievability scheme is to characterize the access function g of our desired secret sharing scheme. Note that beyond being an access function as defined in Section IV-A, the only constraint that our setting imposes on g is $g(z) \leq \alpha$. From our converse results, we know that a piecewise linear access function g would provide the lowest possible individual share sizes. We then consider two cases: $\alpha \geq z/\tau$ and $\alpha < z/\tau$. While we handle the first case with a simple ramp secret sharing scheme, we follow the idea from [10] to handle the second case. Specifically, we remark that g can be written as the sum of two other access functions g_1 and g_2 , i.e., $g = g_1 + g_2$, that correspond to two ramp secret sharing schemes. Finally, we construct a secret sharing scheme with access function g by a combination of two ramp secret sharing schemes with access functions g_1 and g_2 .

1) Case 1: Assume that $\alpha \ge z/\tau$. Let g be the access function of a (τ, τ, T) ramp secret sharing scheme, which can be done as in [12], [13] with, for any $t \in \mathcal{T}$, $\frac{\lambda_t(\alpha, z, \tau)}{H(F)} = \frac{1}{\tau}$, and $\frac{\rho(\alpha, z, \tau)}{H(F)} = 0$.

Note that this scheme satisfies (1) because g(t) = 1 for $t \in [\tau, T]$, and also satisfies (2) because for any $t \in [[0, z]]$, $g(t) \leq z/\tau \leq \alpha$.

(6)

2) Case 2: Assume that $\alpha < z/\tau$. Consider the following access function

$$g: t \mapsto \begin{cases} \frac{\alpha}{z}t & \text{if } t \in \llbracket 0, z \rrbracket \\ \frac{1-\alpha}{\tau-z}(t-z) + \alpha & \text{if } t \in \llbracket z+1, \tau \rrbracket \\ 1 & \text{if } t \in \llbracket \tau+1, T \rrbracket \end{cases}$$

Note that if one can construct a coding scheme with access function g, then this coding scheme satisfies (1) because g(t) = 1 for $t \in [\![\tau, T]\!]$, and also satisfies (2) because for any $t \in [\![0, z]\!]$, $g(t) \leq \alpha$. We construct such a coding scheme using the method in [10]. First, note that $g = g_1 + g_2$, where we have defined

$$g_1: t \mapsto \begin{cases} \frac{\alpha}{z}t & \text{if } t \in [\![0,\tau]\!] \\ \frac{\alpha}{z}\tau & \text{if } t \in [\![\tau+1,T]\!] \end{cases},$$

$$g_2: t \mapsto \begin{cases} 0 & \text{if } t \in [\![0,z]\!] \\ \frac{1-\alpha}{\tau-z}(t-z) + \alpha - \frac{\alpha}{z}t & \text{if } t \in [\![z+1,\tau]\!] \\ 1 - \frac{\alpha}{z}\tau & \text{if } t \in [\![\tau+1,T]\!] \end{cases}.$$

Next, we construct a coding scheme with access function g from two ramp secret sharing schemes with the normalized access functions $\left(\frac{\alpha}{z}\tau\right)^{-1}g_1$ and $\left(1-\frac{\alpha}{z}\tau\right)^{-1}g_2$. By [12], [13], there exist a prime q and $n' \in \mathbb{N}$ such that one can construct an optimal (τ, τ, T) ramp secret sharing (with access function $\left(\frac{\alpha}{z}\tau\right)^{-1}g_1$) that uses $\rho^{(1)}$ random symbols at the encoder and yields the shares $(M_t^{(1)})_{t\in\mathcal{T}}$ for a secret $F_1 \in \mathrm{GF}(q^{n_1})$ with $n_1 = \frac{\alpha}{z}\tau n'$ and $\rho^{(1)} = 0$, and a $(\tau, \tau - z, T)$ ramp secret sharing (with access function $\left(1-\frac{\alpha}{z}\tau\right)^{-1}g_2$) that uses $\rho^{(2)}$ random symbols at the encoder and yields the shares $(M_t^{(2)})_{t\in\mathcal{T}}$ for a secret $F_2 \in \mathrm{GF}(q^{n_2})$, independent of F_1 , with $n_2 = \left(1-\frac{\alpha}{z}\tau\right)n'$ and $\rho^{(2)} = H(F_2)\frac{z}{\tau-z}$. Then, define $F \triangleq (F_1, F_2)$ and for any $t \in \mathcal{T}$, $M_t \triangleq (M_t^{(1)}, M_t^{(2)})$. By [10, Th. 3], this defines a coding scheme with access function g such that for any $t \in \mathcal{T}$, $\frac{\lambda_t(\alpha, z, \tau)}{H(F)} = \Delta_{g_1} + \Delta_{g_2}$, where for $i \in \{1, 2\}, \Delta_{g_i} \triangleq \max_{t \in [0, T-1]} (g_i(t+1) - g_i(t))$. Hence, by remarking that $\Delta_{g_1} = \frac{\alpha}{\tau-z}$. Moreover, $\frac{\rho(\alpha, z, \tau)}{H(F)} = \frac{\rho^{(1)} + \rho^{(2)}}{H(F)} = \frac{H(F_2)}{\tau-z} = \frac{n_2}{n_1+n_2} \frac{z}{\tau-z} = \left(1-\frac{\alpha}{z}\tau\right)\frac{z}{\tau-z} = \frac{z-\tau\alpha}{\tau-z}$.

V. 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 V-A and V-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 V-A1 and V-A2, we prove that for any $t \in \mathcal{T}$,

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

and
$$\frac{\lambda_t}{H(F)} \ge \frac{1}{\tau}$$
, (4)

respectively. We will thus deduce from (3) and (4) that for any

$$t \in \mathcal{T}, \frac{\lambda_t}{H(F)} \ge \max\left(\frac{1-\alpha}{\tau-z}, \frac{1}{\tau}\right).$$

1) Proof of the first lower bound (3) on λ_t : Fix $t \in \mathcal{T}$. For
 $i \in [\![z, \tau - 1]\!]$, define $S_i \triangleq \begin{cases} [\![1, i]\!] & \text{if } t > i \\ [\![1, i + 1]\!] \setminus \{t\} & \text{if } t \leqslant i \end{cases}$ and
 $S_{\tau} \triangleq S_{\tau-1} \cup \{t\}$. Then, for $i \in [\![z + 1, \tau - 1]\!]$, we have
 $H(M_t | M_{S_i})$
 $\stackrel{(a)}{=} H(M_t F | M_{S_i}) - H(F | M_{S_i} M_t)$
 $\stackrel{(b)}{=} H(F | M_{S_i}) + H(M_t | F M_{S_i}) - H(F | M_{S_i} M_t)$ (5)
 $\stackrel{(c)}{=} (1 - C_i)H(F) + H(M_t | F M_{S_i}) - (1 - C_{i+1})H(F)$

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

 $= (C_{i+1} - C_i)H(F) + H(M_t|FM_{\mathcal{S}_i}),$

$$H(M_t|M_{\mathcal{S}_{\tau}})$$

$$\stackrel{(a)}{=} H(F|M_{\mathcal{S}_{\tau}}) + H(M_t|FM_{\mathcal{S}_{\tau}}) - H(F|M_{\mathcal{S}_{\tau}}M_t)$$

$$\stackrel{(b)}{=} H(M_t|FM_{\mathcal{S}_{\tau}})$$

$$\stackrel{(c)}{=} (C_{\tau+1} - C_{\tau})H(F) + H(M_t|FM_{\mathcal{S}_{\tau}}), \quad (7)$$

where

(a) holds as in (5);
(b) holds by (1);
(c) holds by defining C_{τ+1} ≜ C_τ = 1.
We also have

$$H(M_t|M_{\mathcal{S}_z})$$

$$\stackrel{(a)}{=} H(F|M_{\mathcal{S}_z}) + H(M_t|FM_{\mathcal{S}_z}) - H(F|M_{\mathcal{S}_z}M_t)$$

$$\stackrel{(b)}{\geq} (1-\alpha)H(F) + H(M_t|FM_{\mathcal{S}_z}) - (1-C_{z+1})H(F)$$

$$= (C_{z+1}-\alpha)H(F) + H(M_t|FM_{\mathcal{S}_z}), \qquad (8)$$

where

(a) holds as in (5);

(b) holds by (2) and (A_2) .

In the following, for convenience, we define $C_z \triangleq \alpha$. Next, we have

$$\begin{split} H(M_t) \\ \stackrel{(a)}{\geq} H(M_t | M_{\mathcal{S}_z}) & (9) \\ \stackrel{(b)}{=} H(M_t | M_{\mathcal{S}_z}) - H(M_t | M_{\mathcal{S}_\tau}) \\ = \sum_{i=z}^{\tau-1} \left(H(M_t | M_{\mathcal{S}_i}) - H(M_t | M_{\mathcal{S}_{i+1}}) \right) \\ \stackrel{(c)}{\geq} \sum_{i=z}^{\tau-1} \left[(C_{i+1} - C_i) H(F) + H(M_t | FM_{\mathcal{S}_i}) \\ - (C_{i+2} - C_{i+1}) H(F) - H(M_t | FM_{\mathcal{S}_{i+1}}) \right]^+ \\ \stackrel{(d)}{\geq} H(F) \sum_{i=z}^{\tau-1} \left[(C_{i+1} - C_i) - (C_{i+2} - C_{i+1}) \right]^+ \end{split}$$

$$\stackrel{(e)}{=} H(F) \sum_{i=z+1}^{\tau} [(\phi(i) - \phi(i-1)) - (\phi(i+1) - \phi(i))]^{+} \\ \stackrel{(f)}{\geq} H(F) \min_{\phi \in \mathcal{F}} \sum_{i=z+1}^{\tau} [(\phi(i) - \phi(i-1)) - (\phi(i+1) - \phi(i))]^{+},$$
(10)

where

- (a) holds because conditioning does not increase entropy;
- (b) holds because $t \in S_{\tau}$;
- (c) holds because H(M_t|M_{S_i}) H(M_t|M<sub>S_{i+1}) ≥ 0 (conditioning does not increase entropy and S_i ⊂ S_{i+1}) and by (6), (7), (8);
 </sub>
- (d) holds because conditioning does not increase 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 nondecreasing (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 (10) by an expression that only depends on the concave envelopes of the access functions that appear in the objective function. This allows us to conclude that a piecewise linear access function is solution to the optimization. Specifically, let $\phi \in \mathcal{F}$ and let ϕ^+ be the concave envelope of ϕ over $[\![z, \tau + 1]\!]$, i.e., for $i \in [\![z, \tau + 1]\!]$, $\phi^+(i) \triangleq \min\{\psi(i) : \psi \ge \phi, \psi \text{ is concave}\}$. Note that $\phi^+(z) = \phi(z)$ and $\phi^+(\tau + 1) = \phi(\tau + 1)$. Then, for any $i \in [\![z+1,\tau]\!]$ such that $\phi(i) = \phi^+(i)$, we have

$$[(\phi(i) - \phi(i - 1)) - (\phi(i + 1) - \phi(i))]^{+} \ge (\phi(i) - \phi(i - 1)) - (\phi(i + 1) - \phi(i)))$$

$$\stackrel{(a)}{\ge} (\phi(i) - \phi^{+}(i - 1)) - (\phi^{+}(i + 1) - \phi(i)))$$

$$\stackrel{(b)}{=} (\phi^{+}(i) - \phi^{+}(i - 1)) - (\phi^{+}(i + 1) - \phi^{+}(i)), \quad (11)$$

where

(a) holds because $\phi^+ \ge \phi$;

(b) holds because $\phi(i) = \phi^+(i)$.

Moreover, for any $i \in [\![z+1,\tau]\!]$ such that $\phi(i) \neq \phi^+(i)$, we have

$$[(\phi(i) - \phi(i-1)) - (\phi(i+1) - \phi(i))]^+ \ge 0$$

= $(\phi^+(i) - \phi^+(i-1)) - (\phi^+(i+1) - \phi^+(i)),$ (12)

where the last equality holds because ϕ^+ is linear between i-1 and i+1, i.e., $\phi^+(i) - \phi^+(i-1) = \phi^+(i+1) - \phi^+(i)$

- details are provided in Appendix A.

Next, we have

 τ

$$\sum_{i=z+1}^{\tau} [(\phi(i) - \phi(i-1)) - (\phi(i+1) - \phi(i))]^{+}$$

$$\stackrel{(a)}{\geq} \sum_{i=z+1}^{\tau} (\phi^{+}(i) - \phi^{+}(i-1)) - (\phi^{+}(i+1) - \phi^{+}(i)))$$

$$= \phi^{+}(z+1) - \phi^{+}(z) - \phi^{+}(\tau+1) + \phi^{+}(\tau)$$

$$\stackrel{(b)}{=} \phi^{+}(z+1) - \phi^{+}(z)$$

$$\stackrel{(c)}{\geq} \frac{1-\alpha}{\tau-z},$$
(13)

where

- (a) holds by (11) and (12);
- (b) holds because $\phi^{+}(\tau + 1) = \phi^{+}(\tau) = 1$;
- (c) holds because $\phi^+(z+1) \phi^+(z) \ge (\phi^+(\tau) \phi^+(z))/(\tau z)$ by concavity of ϕ^+ and where we have used that $\phi^+(\tau) = 1$ and $\phi^+(z) = \phi(z) = \alpha$.

Finally, we have

$$\lambda_t \ge H(M_t)$$
$$\ge H(F)\frac{1-\alpha}{\tau-z}$$

where the last inequality holds by (10) 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

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

where the last inequality holds because $\phi^+(1) - \phi^+(0) \ge (\phi^+(\tau) - \phi^+(0))/\tau$ 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

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

$$\overset{(j)}{\geqslant} z \frac{1-\alpha}{\tau-z} H(F) + (1-\alpha)H(F)$$

$$= \tau \frac{1-\alpha}{\tau-z} H(F),$$
(14)

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 does not increase entropy and we have defined $S_{z,t} \triangleq [\![1, z+1]\!] \setminus \{t\}$ for $t \in [\![1, z]\!]$;
- (g) holds because for any $t \in [\![1, z]\!]$, $H(M_t|M_{S_{z,t}}) \ge \frac{1-\alpha}{\tau-z}H(F)$, by the converse proof of Theorem 1 starting from (9);
- (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 $H(M_{\mathcal{T} \setminus [\![1,z]\!]}F|M_{[\![1,z]\!]}) \ge H(F|M_{[\![1,z]\!]})$ by the chain rule and positivity of conditional entropy;

(j) holds by (2).

Finally, from (14), we have

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

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

$$\frac{\rho}{H(F)} \ge \frac{[z - \tau\alpha]^+}{\tau - z}.$$

VI. 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 system designer. 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. In the absence of any leakage symmetry, we also derived the optimal sum of the share sizes at all the servers and the optimal amount of local randomness needed at the encoder. 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.

APPENDIX A PROOF OF (12)

By contradiction, assume that ϕ^+ is not linear between i-1 and i+1, then we must have

$$\phi^+(i) > \frac{\phi^+(i+1) + \phi^+(i-1)}{2}$$
 (15)

since ϕ^+ is concave. Next, we have a contradiction by constructing ψ_i , a concave function such that $\phi \leq \psi_i < \phi^+$, as follows:

$$\psi_i : \llbracket z, \tau + 1 \rrbracket \to \mathbb{R}$$

$$j \mapsto \begin{cases} \phi^+(j) & \text{if } j \neq i \\ \max\left(\frac{\phi^+(i+1) + \phi^+(i-1)}{2}, \phi(i)\right) & \text{if } j = i \end{cases}$$

We have $\phi \leq \psi_i$ (since $\phi \leq \phi^+$), and $\psi_i < \phi^+$ by (15) and because $\phi^+(i) > \phi(i)$ (since $\phi^+ \geq \phi$ and $\phi^+(i) \neq \phi(i)$). Then, to show concavity of ψ_i , it is sufficient to show that ψ_i^{Δ} is non-increasing, where ψ_i^{Δ} is defined as

$$\begin{split} \psi_i^{\Delta} : \llbracket z, \tau \rrbracket \to \mathbb{R} \\ j \mapsto \psi_i(j+1) - \psi_i(j). \end{split}$$
 For $j \in \llbracket z, i-3 \rrbracket \cup \llbracket i+1, \tau \rrbracket$, we have

$$\psi_i^{\Delta}(j+1) \leqslant \psi_i^{\Delta}(j) \tag{16}$$

by definition of ψ_i^{Δ} and concavity of ϕ^+ . Then, we have

$$\begin{split} \psi_{i}^{\Delta}(i-1) &\stackrel{(a)}{=} \psi_{i}(i) - \psi_{i}(i-1) \\ &\stackrel{(b)}{=} \psi_{i}(i) - \phi^{+}(i-1) \\ &\stackrel{(c)}{\leqslant} \phi^{+}(i) - \phi^{+}(i-1) \\ &\stackrel{(d)}{\leqslant} \phi^{+}(i-1) - \phi^{+}(i-2) \\ &\stackrel{(e)}{=} \psi_{i}(i-1) - \psi_{i}(i-2) \\ &\stackrel{(f)}{=} \psi_{i}^{\Delta}(i-2), \end{split}$$

where

- (a) and (f) hold by definition of ψ_i^{Δ} ;
- (b) and (e) hold by definition of ψ_i ;
- (c) holds because $\psi_i < \phi^+$;
- (d) holds by concavity of ϕ^+ .

Then, we have

$$\psi_i^{\Delta}(i) \stackrel{(a)}{=} \psi_i(i+1) - \psi_i(i)$$

$$\stackrel{(b)}{=} \phi^+(i+1) - \psi_i(i)$$

$$\stackrel{(c)}{\leqslant} \psi_i(i) - \phi^+(i-1)$$

$$\stackrel{(d)}{=} \psi_i(i) - \psi_i(i-1)$$

$$\stackrel{(e)}{=} \psi_i^{\Delta}(i-1), \qquad (17)$$

where

(a) and (e) hold by definition of ψ^Δ_i;
(b) and (d) hold by definition of ψ_i;
(c) holds because φ^{+(i+1)+φ⁺(i-1)}/2 ≤ ψ_i(i).

Then, we also have

$$\psi_i^{\Delta}(i+1) \stackrel{(a)}{=} \psi_i(i+2) - \psi_i(i+1)$$
$$\stackrel{(b)}{=} \phi^+(i+2) - \phi^+(i+1)$$
$$\stackrel{(c)}{\leqslant} \phi^+(i+1) - \phi^+(i)$$
$$\stackrel{(d)}{\leqslant} \phi^+(i+1) - \psi_i(i)$$

$$\stackrel{(e)}{=} \psi_i(i+1) - \psi_i(i)$$

$$\stackrel{(f)}{=} \psi_i^{\Delta}(i),$$
(18)

where

(a) and (f) hold by definition of ψ_i^{Δ} ;

- (b) and (e) hold by definition of ψ_i ;
- (c) holds by concavity of ϕ^+ ;
- (d) holds because $\psi_i < \phi^+$.

Hence, by (16), (17), and (18), ψ_i^{Δ} is non-increasing and we have thus proved (12) by contradiction.

APPENDIX B

CONVERSE PROOF OF THEOREM 2

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. We prove the lower bounds

$$\frac{\sum_{t \in \mathcal{T}} \lambda_t}{H(F)} \ge T \max\left(\frac{1-\alpha}{\tau-z}, \frac{1}{\tau}\right),\tag{19}$$

$$\frac{\rho}{H(F)} \ge \frac{[z - \tau \alpha]^+}{\tau - z},\tag{20}$$

in Appendices B-A and B-B, respectively.

A. Lower bound on the sum of the share sizes

For $\mathcal{W} \subseteq \mathcal{T}$ and $\mathcal{S} \subseteq \mathcal{T} \setminus \mathcal{W}$ such that $|\mathcal{W}| = z$ and $|\mathcal{S}| = z$ $\tau - z$, we have

$$\sum_{l \in S} H(M_l) \stackrel{(a)}{\geq} H(M_S)$$

$$\stackrel{(b)}{\geq} H(M_S | M_W)$$

$$\stackrel{\geq}{\geq} I(M_S; F | M_W)$$

$$= H(F | M_W) - H(F | M_{W \cup S})$$

$$\stackrel{(c)}{=} H(F | M_W)$$

$$= H(F) - I(F; M_W)$$

$$\stackrel{(d)}{\geq} (1 - \alpha)H(F), \qquad (21)$$

where

- (a) and (b) hold by the chain rule and because conditioning does not increase entropy;
- (c) holds by (1) because $|S \cup W| = \tau$;
- (d) holds by (2) because $|\mathcal{W}| = z$.

Then, by defining $\Theta \triangleq \frac{T}{\tau-z} {T \choose z}^{-1} {T-z \choose \tau-z}^{-1}$, we have

$$T\frac{1-\alpha}{\tau-z}H(F)$$

$$\stackrel{(a)}{=}\Theta\sum_{\substack{\mathcal{W}\subseteq\mathcal{T}\\|\mathcal{W}|=z}}\sum_{\substack{S\subseteq\mathcal{W}^c\\|\mathcal{S}|=\tau-z}}(1-\alpha)H(F)$$

$$\stackrel{(b)}{\leqslant}\Theta\sum_{\substack{\mathcal{W}\subseteq\mathcal{T}\\|\mathcal{W}|=z}}\sum_{\substack{S\subseteq\mathcal{W}^c\\|\mathcal{S}|=\tau-z}}EH(M_l)$$

$$\stackrel{(c)}{=}\Theta\sum_{\substack{\mathcal{W}\subseteq\mathcal{T}\\|\mathcal{W}|=z}}\binom{T-z-1}{\tau-z-1}\sum_{l\in\mathcal{W}^c}H(M_l)$$

$$\stackrel{(d)}{=} \Theta \begin{pmatrix} T-z-1\\ \tau-z-1 \end{pmatrix} \sum_{\substack{\mathcal{W} \subseteq \mathcal{T} \\ |\mathcal{W}|=T-z}} \sum_{l \in \mathcal{W}} H(M_l)$$

$$\stackrel{(e)}{=} \Theta \begin{pmatrix} T-z-1\\ \tau-z-1 \end{pmatrix} \begin{pmatrix} T-1\\ T-z-1 \end{pmatrix} \sum_{l \in \mathcal{T}} H(M_l)$$

$$= \sum_{l \in \mathcal{T}} H(M_l)$$

$$\stackrel{(f)}{\leq} \sum_{l \in \mathcal{T}} \lambda_l,$$

$$(22)$$

where

(a) holds because
$$\binom{T}{z}^{-1} \binom{T-z}{\tau-z}^{-1} \sum_{\substack{W \subseteq T \\ |W|=z}} \sum_{\substack{S \subseteq W^c \\ |S|=\tau-z}} 1 = 1;$$

- (b) holds by (21);
- (c) holds because for any $l \in \mathcal{W}^c$, $H(M_l)$ appears exactly $\binom{T-z-1}{\tau-z-1}$ times in the term $\sum_{\substack{\mathcal{S}\subseteq\mathcal{W}^c}}\sum_{l\in\mathcal{S}}H(M_l)$, note that this observation is similar to [27, Lemma 3.2];
- (d) holds by a change of variables in the sums;
- (e) holds because for any $l \in \mathcal{T}$, $H(M_l)$ appears exactly (e) holds because for any $l \in T$, $H(M_l)$ appears exact $\binom{T-1}{T-z-1}$ times in the sum $\sum_{\substack{|W|=T-z \\ |W|=T-z}} \sum_{l \in W} H(M_l)$; (f) holds by definition of M_l , $l \in T$.

Then, for any $S \subseteq T$ such that $|S| = \tau$, we have

$$\sum_{l \in \mathcal{S}} H(M_l) \ge H(M_{\mathcal{S}})$$
$$\ge H(F), \tag{23}$$

where the last inequality holds because if $H(M_S) < H(F)$, then by (1), we have $H(F|M_{\mathcal{S}}) = 0$ and then the contradiction $H(M_{\mathcal{S}}|F) = H(M_{\mathcal{S}}) - H(F) < 0$. We also have

$$\frac{T}{\tau}H(F) = \frac{T}{\tau} {\binom{T}{\tau}}^{-1} \sum_{\substack{\mathcal{S} \subseteq \mathcal{T} \\ |\mathcal{S}| = \tau}} H(F)$$

$$\stackrel{(a)}{\leqslant} \frac{T}{\tau} {\binom{T}{\tau}}^{-1} \sum_{\substack{\mathcal{S} \subseteq \mathcal{T} \\ |\mathcal{S}| = \tau}} \sum_{l \in \mathcal{S}} H(M_l)$$

$$= \frac{T}{\tau} {\binom{T}{\tau}}^{-1} {\binom{T-1}{\tau-1}} \sum_{l \in \mathcal{T}} H(M_l)$$

$$= \sum_{l \in \mathcal{T}} H(M_l)$$

$$\stackrel{(b)}{\leqslant} \sum_{l \in \mathcal{T}} \lambda_l,$$
(24)

where

(a) holds by (23);

(b) holds by definition of M_l , $l \in \mathcal{T}$.

Finally, we conclude from (22) and (24) that (19) holds.

B. Lower bound on the amount of local randomness

Let $\mathcal{V} \subseteq \mathcal{T}$ such that $v \triangleq |\mathcal{V}| < z$. For $\mathcal{W} \subseteq \mathcal{T} \setminus \mathcal{V}$ and $S \subseteq \mathcal{T} \setminus (\mathcal{W} \cup \mathcal{V})$ such that $|\mathcal{W}| = z - v$ and $|S| = \tau - z$, we have

$$\sum_{l \in S} H(M_l | M_{\mathcal{V}}) \stackrel{(a)}{\geq} H(M_{\mathcal{S}} | M_{\mathcal{V}})$$

$$\stackrel{(b)}{\geq} H(M_{\mathcal{S}} | M_{\mathcal{V} \cup \mathcal{W}})$$

$$\stackrel{(c)}{\geq} (1 - \alpha) H(F), \qquad (25)$$

where

- (a) and (b) hold by the chain rule and because conditioning does not increase entropy;
- (c) holds similar to (21) with the substitution $\mathcal{W} \leftarrow \mathcal{V} \cup \mathcal{W}$, which is possible because $|\mathcal{V} \cup \mathcal{W} \cup \mathcal{S}| = \tau$ and $|\mathcal{V} \cup \mathcal{W}| = z$.

Then, by defining $\Lambda \triangleq \frac{1}{\tau-z} {\binom{T-v}{z-v}}^{-1} {\binom{T-z}{\tau-z}}^{-1}$, we have

$$\frac{1-\alpha}{\tau-z}H(F) = \Lambda \sum_{\substack{W \subseteq \mathcal{T} \setminus \mathcal{V} \\ |W| = z-v}} \sum_{\substack{S \subseteq \mathcal{T} \setminus (W \cup \mathcal{V}) \\ |S| = \tau-z}} (1-\alpha)H(F)$$

$$\stackrel{(a)}{\leq} \Lambda \sum_{\substack{W \subseteq \mathcal{T} \setminus \mathcal{V} \\ |W| = z-v}} \sum_{\substack{S \subseteq \mathcal{T} \setminus (W \cup \mathcal{V}) \\ |S| = \tau-z}} H(M_l|M_{\mathcal{V}})$$

$$\stackrel{(b)}{=} \Lambda \sum_{\substack{W \subseteq \mathcal{T} \setminus \mathcal{V} \\ |W| = z-v}} \left(\frac{T-z-1}{\tau-z-1} \right) \sum_{l \in \mathcal{T} \setminus (W \cup \mathcal{V})} H(M_l|M_{\mathcal{V}})$$

$$\stackrel{(c)}{=} \Lambda \left(\frac{T-z-1}{\tau-z-1} \right) \sum_{\substack{W \subseteq \mathcal{T} \setminus \mathcal{V} \\ |W| = T-z}} H(M_l|M_{\mathcal{V}})$$

$$\stackrel{(d)}{=} \Lambda \left(\frac{T-z-1}{\tau-z-1} \right) \left(\frac{T-v-1}{T-z-1} \right) \sum_{l \in \mathcal{T} \setminus \mathcal{V}} H(M_l|M_{\mathcal{V}})$$

$$= \frac{1}{T-v} \sum_{l \in \mathcal{T} \setminus \mathcal{V}} H(M_l|M_{\mathcal{V}})$$

$$\stackrel{(e)}{=} \frac{1}{T-v} \sum_{l \in \mathcal{T} \setminus \mathcal{V}} H(M_l|M_{\mathcal{V}})$$

$$= H(M_{l^*(\mathcal{V})}|M_{\mathcal{V}})$$

$$= H(M_{\mathcal{T}}|M_{\mathcal{V}}) - H(M_{\mathcal{T}}|M_{\mathcal{V} \cup \{l^*(\mathcal{V})\}}), \quad (26)$$

where

- (a) holds by (25);
- (a) holds by (10),
 (b) holds because for any l ∈ T\(W ∪ V), the term H(M_l|M_V) appears exactly (^{T-z-1}/_{τ-z-1}) times in the term ∑S⊆T\(W∪V) ∑_{l∈S} H(M_l|M_V), this argument is similar |S|=τ-z to [27, Lemma 3.2];
- (c) holds by a change of variables in the sums;
- (d) holds because for any $l \in \mathcal{T} \setminus \mathcal{V}$, $H(M_l|M_{\mathcal{V}})$ appears exactly $\binom{T-v-1}{T-z-1}$ times in the term $\sum_{\substack{\mathcal{W} \subseteq \mathcal{T} \setminus \mathcal{V} \\ |\mathcal{W}| = T-z}} \sum_{l \in \mathcal{W}} H(M_l|M_{\mathcal{V}});$
- (e) holds with $l^{\star}(\mathcal{V}) \in \arg \max_{l \in \mathcal{T} \setminus \mathcal{V}} H(M_l | M_{\mathcal{V}}).$

Next, define $\mathcal{V}_0 \triangleq \emptyset$ and for $i \in [[1, z]], \mathcal{V}_i \triangleq \mathcal{V}_{i-1} \cup \{l^*(\mathcal{V}_{i-1})\}$. Then, we have

$$= \left(z\frac{1-\alpha}{\tau-z} - \alpha\right)H(F)$$

$$= -\alpha H(F) + z\frac{1-\alpha}{\tau-z}H(F)$$

$$= -\alpha H(F) + \sum_{i=0}^{z-1}\frac{1-\alpha}{\tau-z}H(F)$$

$$\stackrel{(a)}{\leqslant} -\alpha H(F) + \sum_{i=0}^{z-1}[H(M_{\mathcal{T}}|M_{\mathcal{V}_i}) - H(M_{\mathcal{T}}|M_{\mathcal{V}_{i+1}})]$$

$$= -\alpha H(F) + H(M_{\mathcal{T}}) - H(M_{\mathcal{T}}|M_{\mathcal{V}_z})$$

$$\stackrel{(b)}{\leqslant} -\alpha H(F) + H(F,R) - H(M_{\mathcal{T}}|M_{\mathcal{V}_z})$$

$$\stackrel{(c)}{\equiv} (1-\alpha)H(F) + H(R) - H(M_{\mathcal{T}}|M_{\mathcal{V}_z})$$

$$\stackrel{(d)}{\equiv} (1-\alpha)H(F) + H(R) - H(FM_{\mathcal{T}}|M_{\mathcal{V}_z})$$

$$\stackrel{(e)}{\leqslant} H(R)$$

$$= \rho, \qquad (27)$$

where

- (a) holds by applying z times (26) and the definition of \mathcal{V}_i , $i \in [\![0, z]\!];$
- (b) holds because $M_{\mathcal{T}}$ is a deterministic function of (F, R)
- (c) holds by independence between F and R;
- (d) holds by the chain rule and because $H(F|M_{\mathcal{T}}) = 0$ by (1);
- (e) holds because $-H(F|M_{\mathcal{V}_z}) \leq -(1-\alpha)H(F)$ by (2).

Finally, since we also have $\rho \ge 0$, we conclude from (27) that (20) holds.

REFERENCES

- 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.
- [2] 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.
- [3] 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.
- [4] 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.
- [5] 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.
- [6] A. Shamir, "How to share a secret," *Communications of the ACM*, vol. 22, no. 11, pp. 612–613, 1979.
- [7] G. R. Blakley, "Safeguarding cryptographic keys," in *Proceedings of the National Computer Conference*, vol. 48, 1979, pp. 313–317.
- [8] A. Beimel, "Secret-sharing schemes: A survey," in International Conference on Coding and Cryptology. Springer, 2011, pp. 11–46.
- [9] E. Karnin, J. Greene, and M. Hellman, "On secret sharing systems," *IEEE Transactions on Information Theory*, vol. 29, no. 1, pp. 35–41, 1983.
- [10] 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.
- [11] O. Farràs, T. Hansen, T. Kaced, and C. Padró, "Optimal non-perfect uniform secret sharing schemes," in *Annual Cryptology Conference CRYPTO*. Springer, 2014, pp. 217–234.

- [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] K. Kurosawa, K. Okada, K. Sakano, W. Ogata, and S. Tsujii, "Nonperfect secret sharing schemes and matroids," in Advances in Cryptology—EUROCRYPT'93: Workshop on the Theory and Application of Cryptographic Techniques Lofthus, Norway, May 23–27, 1993 Proceedings 12. Springer, 1994, pp. 126–141.
- [17] R. M. Capocelli, A. De Santis, L. Gargano, and U. Vaccaro, "On the size of shares for secret sharing schemes," *Journal of Cryptology*, vol. 6, no. 3, pp. 157–167, 1993.
- [18] C. Blundo, A. De Santis, L. Gargano, and U. Vaccaro, "On the information rate of secret sharing schemes," in Advances in Cryptology—CRYPTO'92. Springer, 1992, pp. 148–167.
- [19] M. Van Dijk, "On the information rate of perfect secret sharing schemes," *Designs, codes and cryptography*, vol. 6, no. 2, pp. 143–169, 1995.
- [20] L. Csirmaz, "The size of a share must be large," in Advances in Cryptology—EUROCRYPT'94. Springer, 1995, pp. 13–22.
- [21] C. Blundo, A. D. Santis, R. D. Simone, and U. Vaccaro, "Tight bounds on the information rate of secret sharing schemes," *Designs, Codes and Cryptography*, vol. 11, pp. 107–110, 1997.
- [22] O. Farràs, T. Kaced, S. Martín, and C. Padró, "Improving the linear programming technique in the search for lower bounds in secret sharing," *IEEE Transactions on Information Theory*, vol. 66, no. 11, pp. 7088– 7100, 2020.
- [23] 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.
- [24] M. Yoshida, T. Fujiwara, and M. Fossorier, "Optimum general threshold secret sharing," in *International Conference on Information Theoretic Security, ICITS.* Springer, 2012, pp. 187–204.
- [25] O. Farràs, T. B. Hansen, T. Kaced, and C. Padró, "On the information ratio of non-perfect secret sharing schemes," in *Algorithmica*, vol. 79, no. 4. Springer, 2017, pp. 987–1013.
- [26] C. Blundo, A. De Santis, and U. Vaccaro, "Randomness in distribution protocols," *Information and Computation*, vol. 131, no. 2, pp. 111–139, 1996.
- [27] A. De Santis and B. Masucci, "Multiple ramp schemes," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1720–1728, 1999.