The Capacity of Private Information Retrieval From Uncoded Storage Constrained Databases

Mohamed Adel Attia, *Member, IEEE*, Deepak Kumar, and Ravi Tandon, *Senior Member, IEEE*

***Abstract*— Private information retrieval (PIR) allows a user to retrieve a desired message from a set of databases without revealing the identity of the desired message. The replicated database scenario, where N databases store each of the K messages was considered by Sun and Jafar, and the optimal download cost was characterized as $\left(1+\frac{1}{N}+\frac{1}{N^2}+\cdots+\frac{1}{N^{K-1}}\right)$. In this work, we consider the problem of PIR from *uncoded storage constrained* databases. Each database has a storage capacity of μKL bits, where L is the size of each message in bits, and $\mu \in [1/N, 1]$ is the normalized storage. The novel aspect of this work is to characterize the optimum download cost of PIR from uncoded storage constrained databases for any "normalized storage" value in the range $\mu \in [1/N, 1]$. In particular, for any (N, K), we show that the optimal trade-off between normalized storage, μ, and the download cost, $D(\mu)$, is a piece-wise linear function given by the lower convex hull of the N pairs $\left(\frac{t}{N}, \left(1+\frac{1}{t}+\frac{1}{t^2}+\cdots+\frac{1}{t^{K-1}}\right)\right)$ for $t = 1, 2, \ldots, N$. To prove this result, we first present a storage constrained PIR scheme for any (N, K). Next, we obtain a general lower bound on the download cost for PIR, which is valid for any arbitrary storage architecture. The uncoded storage assumption is then applied which allows us to express the lower bound as a linear program (LP). Finally, we solve the LP to obtain tight lower bounds on the download cost for different regimes of storage, which match the proposed storage constrained PIR scheme.**

***Index Terms*— Private information retrieval, distributed storage, capacity, uncoded storage, storage constrained databases.**

I. Introduction

With the paradigm-shifting developments towards distributed storage systems (DSS), assuring privacy while retrieving information from public databases has become a crucial need for users. This problem, also referred to as private information retrieval (PIR) has direct practical applications in cloud storage, social networking, privately accessing stock market records or bank loans, or even activists seeking files that might be considered anti-regime. The original formulation of the PIR problem involves N non-colluding and replicated databases, each storing K messages. Upon receiving queries from the legitimate user, the databases answer truthfully with the required information, which means they are curious but honest. Successful PIR must satisfy two properties: first, each of the N queries sent from the user to the N databases must reveal nothing about the identity of the message being requested; and second, the user must be able to correctly decode the message of interest from the answers received from the N databases.

A trivial solution to PIR is to download all the messages from the databases, but it is highly impractical especially when the number of messages K is too large. The goal is to design an efficient protocol, which is characterized by the total upload/download cost the user has to pay in order to download a message privately. The PIR problem has been studied extensively within the computer science community [1]–[5]. In the pioneering work by Chor *et al.* [1], the authors considered PIR with one bit length messages, where the databases are assumed to be computationally unbounded. The privacy cost is calculated based on the total amount of communication between the user and the databases, i.e., the upload cost which is the size of the N queries, and the download cost which is the size of the N answers. It was shown in [1] that achieving perfect privacy while retrieving from a single computationally unbounded database requires downloading all the messages. Single database PIR was studied in [3]–[5] where the database is assumed to be computationally bounded.

The Shannon theoretic approach for this problem is to allow the size of the messages to be arbitrary large, and therefore the upload cost is considered negligible with respect to the download cost [6]–[8]. This case is more suitable compared to the original formulation in [1] when the size of the files/messages to be retrieved (download cost) is large. Based on the Shannon theoretic formulation, the rate of a PIR scheme is the ratio between the number of desired vs downloaded bits, and PIR capacity is then defined as the maximum achievable rate. In a recent interesting work by Sun and Jafar [7], the exact capacity of PIR (or the inverse of optimum download cost) for any (N, K) was characterized as $(1+\frac{1}{N}+\frac{1}{N^2}+\cdots+\frac{1}{N^{K-1}})^{-1}$.

Since the appearance of [7], significant recent progress has been made on a variety of variations of the basic PIR problem. The case of T-colluding PIR (or TPIR in short) was investigated by Sun and Jafar in [9], where any T

Manuscript received October 8, 2018; revised July 13, 2020; accepted July 19, 2020. Date of publication September 9, 2020; date of current version October 21, 2020. This work was supported in part by NSF under Grant CAREER- 1651492 and Grant CNS-1715947 and in part by the 2018 Keysight Early Career Professor Award. This article was presented in part at the 2018 IEEE International Conference on Communications (ICC) and in part at the 2018 IEEE International Symposium on Information Theory (ISIT). *(Corresponding author: Ravi Tandon.)*

Mohamed Adel Attia and Ravi Tandon are with the Department of Electrical and Computer Engineering, The University of Arizona, Tucson, AZ 85719 USA (e-mail: madel@email.arizona.edu; tandonr@email.arizona.edu).

Deepak Kumar is with Citi Bank, Irving, TX 75039 USA (e-mail: deepakkumar@email.arizona.edu).

Communicated by M. Wigger, Associate Editor for Shannon Theory.

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2020.3023016

0018-9448 © 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: The University of Arizona. Downloaded on January 30,2021 at 20:39:51 UTC from IEEE Xplore. Restrictions apply.

databases out of N are able to collude by sharing their received queries. The problem of PIR with databases storing coded messages, using (N, M) MDS codes, was considered by Tajeddine and El-Rouayheb in [10] and the capacity was subsequently characterized by Banawan and Ulukus in [11] to take the value $(1 + \frac{M}{N} + \frac{M^2}{N^2} + \cdots + \frac{M^{K-1}}{N^{K-1}})^{-1}$. This setting was further investigated for the scenario where any T out of N databases can collude while any subset of M databases out of N fail to respond, also referred to as MDS-TPIR [12], [13]. The exact capacity of robust MDS-TPIR for any (T, N, M, K) and $N \geq M \geq T$ was characterized in [9] as $(1 + \frac{T}{N} + \frac{T^2}{N^2} + \cdots + \frac{T^{K-1}}{N^{K-1}})^{-1}$.

The capacity of cache aided PIR (PIR with side information) was recently characterized in [14], where the user has a local cache of limited storage $0 \leq S \leq K$ and contents known to the databases. It was shown that memory sharing between full storage and no-cache-aided PIR schemes is information-theoretically optimal. The capacity of PIR with private side information, or PIR-PSI [15]–[17], was characterized in [17] to take the value $(1 + \frac{1}{N} + \frac{1}{N^2} + \cdots + \frac{1}{N^{K-M-1}})^{-1}$, where M is the number of messages known as side information to the user. The capacity of multi-round PIR, where the queries in each round are allowed to depend on the answers received in previous rounds, was characterized by Sun and Jafar in [18]. Although no advantage in terms of capacity of having multi-rounds as opposed to the single round case considered in [7], it was shown that the multi-round queries can help in reducing the storage overhead at the databases.

Many other variations of PIR have been considered recently including: Symmetric PIR (SPIR) [19], [20] where the user must be able to retrieve the message of interest privately, while at the same time the databases must avoid any information leakage to the user about the remaining messages; The case of multi-message PIR (or MPIR) [21], [22] in which the user wants to privately retrieve $P \geq 1$ out of K messages in one communication round; PIR with Byzantine databases (or BPIR) [23], [24] where any subset of databases are adversarial and they can respond with incorrect answers; PIR through wiretap channel (or PIR-WTC) [25] where the user wants to retrieve a single message in the presence of an external eavesdropper; and PIR with asymmetric traffic constraint for the databases in [26].

A. PIR From Uncoded Storage Constrained Databases

The replicated databases assumption, where each database stores all the K messages, incurs substantial storage overhead especially for significant large number of messages K. Moreover, in practice, different databases obtained by various servers might not necessarily be replicated, i.e, they may not store the same set of messages/files. In this work, we consider the problem of PIR from *uncoded storage constrained* databases. Each database has a storage capacity of μKL bits, where K is the number of messages, L is the size of each message in bits, and $\mu \in [1/N, 1]$ is the normalized storage. On one extreme, $\mu = 1/N$ is the minimum storage at databases so that the user can retrieve any required message. If the user is interested in retrieving a message, then all the K messages from databases must be downloaded to achieve privacy. This can be viewed through the independence of the storage contents across databases, i.e., each database stores distinct information. In this case, each database is treated independently as a small PIR problem with one database, where all data in that database has be downloaded to achieve privacy. On the other hand, $\mu = 1$ is the replicated databases case settled in [7], where the download cost is minimal. Thus, we aim to characterize this trade-off for any value of μ between these two extremes.

For the storage placement strategy at the databases, we focus on the following special settings:

- As a first step towards solving the problem for any arbitraty storage at the databases, we assume the special case where the storage placement at the databases is centrally optimized. It is important to notice here that while the storage at the databases can be centrally designed in the placement phase, the databases cannot share the received queries from the users in the data delivery phase, i.e., non-colluding databases. The ultimate goal is to extend this study to the problem where the storage at the databases can be made arbitrary.
- We also assume the case where the databases can only store uncoded functions of the messages, i.e., uncoded storage. The works in [10]–[13] focused on the use of MDS coded databases to satisfy the M-out-of-N recoverability, which is not within the scope of this article. However, the results in [11] suggests the benefits of using codes in DB storage to reduce the download cost. Later in Section 3, we conduct a comparison between the results in [11] and our results for storage constrained PIR to show the significance of using coded storage at the databases.

Majority of classical works on PIR assume the presence of replicated databases. Indeed, exceptions to this statement include [18] which investigated the problem of limited storage PIR for the special case of $K = 2$ messages and $N = 2$ databases. The authors presented interesting lower and upper bounds on the capacity for this special case, and show the optimality of the proposed scheme for the case of linear schemes. In [27], the authors proposed a non-linear scheme for the canonical case $K = 2$ and $N = 2$, showing that the proposed non-linear scheme uses less storage than the optimal linear scheme when the retrieval rate is kept optimal.

It is worth mentioning here that our work was extended to the case of decentralized storage constrained databases [28], where the storage placement is done independently. The authors of [28] show that uniform and random caching scheme originally proposed for decentralized coded caching in [29], along with the retrieval scheme which is originally proposed for PIR from replicated databases [7], surprisingly, result in the optimal download cost. Moreover, the extension to the case of heterogeneous storage constrained databases was considered in [30], where databases have limited and heterogeneous sizes. Surprisingly, the authors in [30] show that the optimal download cost matches our results for homogeneous databases having *same average storage constraint.* Both the works [28] and [30] apply the general lower bound derived in our work to prove the optimality of their proposed schemes.

Authorized licensed use limited to: The University of Arizona. Downloaded on January 30,2021 at 20:39:51 UTC from IEEE Xplore. Restrictions apply.

Summary of Contributions– In this work, we characterize the optimal PIR download cost from uncoded storage constrained databases as a piece-wise linear function given by the lower convex hull of N storage-cost pairs $\left(\frac{t}{N}, 1+\frac{t}{N}+\frac{t^2}{N^2}+\cdots+\frac{t^{K-1}}{N^{K-1}})^{-1}\right)$, where $t \in [1:N]$. The achievablility proof, which was presented in parts in [31], works as follows: For the discrete storage values $\mu = \frac{t}{N}$ where $t \in [1:N]$, the storage design at the databases is inspired by storage design schemes in the caching literature [32], where the users prefetch popular content into memories to reduce peak traffic rates when downloading from a single server. As opposed to caching schemes, the storage placement for storage constrained PIR scheme occurs at the databases which should span the whole set of files. Our storage design allows dividing the PIR scheme into blocks of smaller PIR systems where only a subset of databases of size t is involved. The storage points in between the discrete storage values can be achieved by a memory sharing argument, which is given by the lower convex hull of the achieved rate-storage pairs.

As a first step in understanding the fundamental limits, we proved in [33] the optimality of our storage constrained scheme for the special case of $K = 3$ messages, $N = 3$ databases, and any storage value at the databases, under *uncoded and symmetric assumptions* for the storage placement. Our second main contribution of this article is to show that the proposed scheme is information-theoretically optimal for any (N, K, μ), under *uncoded storage placement* assumption. The key technical challenge in proving the lower bounds is dealing with all possible components of storage at the databases limited by the storage and the message size constraints, which significantly go beyond the techniques introduced in [7]. To this end, we tailor the mutual information of the key components used in [7] for the full storage setting to the case of limited storage. We factorize these terms to arrive to the first universal lower bound on the download cost, which is valid *for any arbitrary storage*. Next, we specialize the obtained lower bound to uncoded placement strategies with homogeneous storage constrained databases. This bounding technique is inspired by the methodology recently proposed in [34] for uncoded caching systems, and later applied in our previous work on coded data shuffling [35], [36], and uncoded caching systems with secure delivery [37]. Applying these ideas helps in obtaining a linear program (LP) subject to the storage and message size constraints, which can be solved for different regimes of storage to provide a set of lower bounds on the download cost, and to show that these bounds exactly match the download cost of the proposed storage-constrained PIR scheme.

B. Notation

The notation $[n_1 : n_2]$ for $n_1 < n_2$, and $n_1, n_2 \in \mathbb{N}$ represents the set of all integers between n_1, and n_2, i.e., $\{n_1, n_1+1, \ldots, n_2\}$. The combination coefficient $\binom{n}{k} = \frac{n!}{(n-k)!k!}$ equals zero for $k > n$, or $k < 0$. Elements of ordered sets are placed between round brackets $()$, while for regular sets we use curly brackets $\{\}$. We use bold letters to represent ordered sets, and calligraphy letters for regular sets. In order to describe subsets of ordered sets, we use the subscript to give the indexes of the elements being chosen from the set, e.g., for the ordered set $\boldsymbol{\pi} = (\pi_1 \ldots, \pi_n)$, $\boldsymbol{\pi}_{[1:4]} = (\pi_1, \pi_2, \pi_3, \pi_4)$. We denote Random Variables (RVs) by capital letters, and ordered sets of RVs by capital bold letters. The set in the subscript of a set of ordered RVs is used for short notation of a subset of the set of RVs, e.g., for an ordered set of RVs $\mathbf{Z} = (Z_1, \ldots, Z_n)$, we use $\mathbf{Z}_{\mathcal{W}}$ to denote the all the random variables Z_i where $i \in \mathcal{W}$.

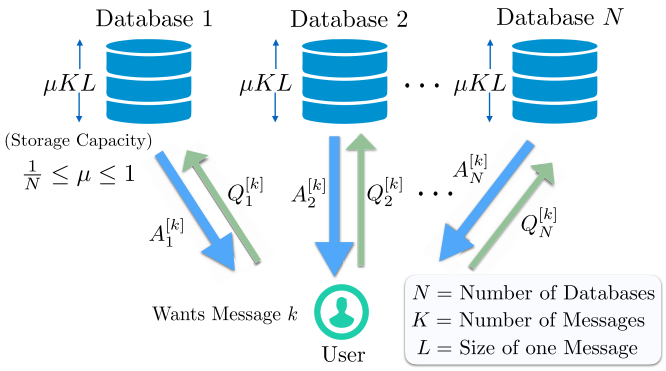

Fig. 1. Storage constrained private information retrieval.

II. Storage Constrained PIR: Problem Statement

We consider the PIR problem with N non-colluding databases, labeled as $\text{DB}_1, \text{DB}_2, \ldots, \text{DB}_N$, and K independent messages, labeled as $W_1, W_2, \ldots, W_K$, where each message is of size L bits, i.e.,

$$H(W_1) = H(W_2) = \cdots = H(W_K) = L. \tag{1}$$

We assume that each database has a storage capacity of μKL bits. If we denote $Z_1, Z_2, \ldots, Z_N$ as the contents stored across the databases, where Z_n is the storage content of DB_n, then we have the following *storage constraint* for each database:

$$H(Z_1) = H(Z_2) = \cdots = H(Z_N) \leq \mu KL. \tag{2}$$

We assume that the storage strategy employed by the user is completely public, i.e., each database knows which contents are stored at all the other databases. The normalized storage μ can take values in the range $1/N \leq \mu \leq 1$. The case when $\mu = 1$ is the setting of replicated databases, with each database storing all the K messages. The lower bound $\mu \geq 1/N$ is in fact a necessary condition for reliable decoding.

The storage-constrained PIR model is shown in Figure 1. To request a message, a user privately selects a number k between 1 and K corresponding to the desired message W_k. Then the user generates N queries $Q_1^{[k]}, Q_2^{[k]}, \ldots, Q_N^{[k]}$, where $Q_n^{[k]}$ is sent to the n^{th} database (DB_n), and the queries are independent of the messages, i.e.,

$$I(W_1, \ldots, W_K; Q_1^{[k]}, \ldots, Q_N^{[k]}) = 0, \quad \forall k \in [1:K]. \tag{3}$$

Upon receiving the query $Q_n^{[k]}$, DB_n returns an answer $A_n^{[k]}$ to the user, which is a function of the corresponding query and the data stored in the DB_n, i.e.,

$$H(A_n^{[k]} | Q_n^{[k]}, Z_n) = 0, \quad \forall k \in [1:K], \ \forall n \in [1:N]. \tag{4}$$

Authorized licensed use limited to: The University of Arizona. Downloaded on January 30,2021 at 20:39:51 UTC from IEEE Xplore. Restrictions apply.

From all the answers from databases, the user must be able to correctly decode the desired message W_k with a small probability of error P_e as the message size L approaches infinity, i.e., the following *correctness constraint* must be satisfied

$$H(W_k|A_1^{[k]},\dots,A_N^{[k]},Q_1^{[k]},\dots,Q_N^{[k]}) = o(L),\ \forall k \in [1:K], \tag{5}$$

where $o(L)$ represents a function of L such that $o(L)/L$ approaches 0 as L approaches infinity. In order to prevent the database DB_n from learning the identity of requested message, privacy must be guaranteed through the following statistical equivalence constraints for all $k_1 \neq k_2 \in [1:K]$:

$$\begin{aligned}&(Q_n^{[k_1]},A_n^{[k_1]},W_1,\dots,W_K,Z_1,\dots,Z_N)\\ &\quad\sim\ (Q_n^{[k_2]},A_n^{[k_2]},W_1,\dots,W_K,Z_1,\dots,Z_N).\end{aligned} \tag{6}$$

For a normalized storage μ, let $\phi_\mu : \{W_{[1:K]} \to Z_{[1:N]}\}$ be the storage placement function mapping the message bits to the database storage. Let us denote the average number of download bits (over all random queries) needed to retrieve a file W_k for $k \in [1:K]$ privately by $D_\mu^{[k]}$, i.e., $D_\mu^{[k]} \geq H(\mathbf{A}_{[1:N]}^{[k]}|\mathbf{Q}_{[1:N]}^{[k]})$. Finally, we denote the average number of download bits (over all possible messages) as $D_\mu = \frac{1}{K}\sum_{k\in[1:K]} D_\mu^{[k]}$. For a placement function ϕ_μ, we say that a pair (D_μ, L) is achievable if there exists a PIR scheme with storage, querying, and decoding functions, which satisfy the storage, correctness and privacy constraints in (2), (5) and (6), respectively. The performance of a PIR scheme is characterized by the number of bits of desired information per one downloaded bit. In particular, if D_μ is the total number of downloaded bits, and L is the size of the desired message, then the download cost is D_μ/L. In other words, the PIR rate is L/D_μ. The goal is to characterize the optimal download cost as a function of the database normalized storage μ:

$$D^*(\mu) = \min\{D_\mu/L : (D_\mu, L) \text{ is achievable}\}. \tag{7}$$

The storage-constrained capacity of PIR is the inverse of the download cost,

$$C^*(\mu) = \max\{L/D_\mu : (D_\mu, L) \text{ is achievable}\}. \tag{8}$$

We next present Claim 1 which shows that the optimal download cost $D^*(\mu)$ (or the inverse of capacity $1/C^*(\mu)$) is a convex function of the normalized storage μ. The proof of Claim 1 is in Appendix A.

Claim 1: The optimal download cost $D^*(\mu)$ is a convex function of μ. In other words, for any (μ_1, μ_2), and $\alpha \in [0,1]$, the following inequality is true:

$$D^*(\alpha\mu_1 + (1-\alpha)\mu_2) \leq \alpha D^*(\mu_1) + (1-\alpha)D^*(\mu_2). \tag{9}$$

A. Storage Constrained PIR: Uncoded Storage Assumption

Now, we specialize the above system model for the storage constrained PIR using uncoded storage assumption, where the databases only store uncoded functions of the K messages subject to the storage constraint. We consider a generic uncoded placement strategy such that if we consider a message W_k, we denote $W_{k,\mathcal{S}}$ as the set of bits of W_k that are fully stored at the databases in the set $\mathcal{S} \subseteq [1:N]$, where $|\mathcal{S}| \geq 1$, and are not stored at any of the other databases in the set $[1:N] \setminus \mathcal{S}$. That is:

$$W_{k,\mathcal{S}} = W_k \cap \mathbf{Z}_{\mathcal{S}} \cap \left(W_k \setminus \mathbf{Z}_{[1:K]\setminus\mathcal{S}}\right), \tag{10}$$

where $W_k \setminus \mathbf{Z}_{[1:K]\setminus\mathcal{S}}$ denotes the parts of W_k that is not available in the storage of databases in the set $[1:K]\setminus\mathcal{S}$. As a result, we can write the content of DB_n, Z_n as

$$Z_n = \bigcup_{k\in[1:K]} \bigcup_{\substack{\mathcal{S}\subseteq[1:N]\\ n\in\mathcal{S}}} W_{k,\mathcal{S}}. \tag{11}$$

Furthermore, the message W_k consists of $2^N - 1$ partitions, $W_{k,\mathcal{S}}$, for $\mathcal{S} \in \mathcal{P}([1:N])$, where $\mathcal{P}([1:N])$ is the power set of all possible subsets of the set $[1:N]$ not including the empty set. Therefore, the message W_k can also be equivalently expressed as

$$W_k = \bigcup_{\substack{\mathcal{S}\subseteq[1:N]\\ |\mathcal{S}|\geq 1}} W_{k,\mathcal{S}}. \tag{12}$$

Now, let us consider $W_{k,\mathcal{S}}$ as a random variable with entropy

$$H(W_{k,\mathcal{S}}) = |W_{k,\mathcal{S}}|L, \tag{13}$$

where $|W_{k,\mathcal{S}}|$ is the size of $W_{k,\mathcal{S}}$ normalized by the message size L. Therefore, the following two constraints are obtained:

- **Message size constraint:** The first constraint is related to the total size of all the messages, W_k and $k \in [1:K]$, given by KL bits,

$$\begin{aligned}1 &= \frac{1}{KL}H(\mathbf{W}_{[1:K]}) = \frac{1}{KL}H(W_1, W_2, \cdots, W_K)\\ &\overset{(a)}{=} \frac{1}{KL}\sum_{k=1}^{K} H(W_k) \overset{(b)}{=} \frac{1}{KL}\sum_{k=1}^{K}\sum_{\substack{\mathcal{S}\subseteq[1:N]\\ |\mathcal{S}|\geq 1}} H(W_{k,\mathcal{S}})\\ &= \sum_{\ell=1}^{N}\frac{1}{K}\sum_{k=1}^{K}\sum_{\substack{\mathcal{S}\subseteq[1:N]\\ |\mathcal{S}|=\ell}} |W_{k,\mathcal{S}}|\\ &= \sum_{\ell=1}^{N}\binom{N}{\ell}\underbrace{\frac{1}{K\binom{N}{\ell}}\sum_{k=1}^{K}\sum_{\substack{\mathcal{S}\subseteq[1:N]\\ |\mathcal{S}|=\ell}} |W_{k,\mathcal{S}}|}_{\triangleq x_\ell} = \sum_{\ell=1}^{N}\binom{N}{\ell}x_\ell,\end{aligned} \tag{14}$$

where (a) follows since the messages are independent, (b) follows from (12), and $x_\ell \geq 0$ is defined as

$$x_\ell \triangleq \frac{1}{K\binom{N}{\ell}}\sum_{k=1}^{K}\sum_{\substack{\mathcal{S}\subseteq[1:N]\\ |\mathcal{S}|=\ell}} |W_{k,\mathcal{S}}|, \quad \ell \in [1:N]. \tag{15}$$

- **Storage constraint:** The second constraint is related to the total storage of all the N databases, which cannot exceed

Authorized licensed use limited to: The University of Arizona. Downloaded on January 30,2021 at 20:39:51 UTC from IEEE Xplore. Restrictions apply.

μNKL bits for $\mu \in [\frac{1}{K}, 1]$,

$$\mu N \geq \frac{1}{KL}\sum_{n=1}^{N} H(Z_n) \overset{(a)}{=} \frac{1}{KL}\sum_{n=1}^{N}\sum_{k=1}^{K}\sum_{\substack{\mathcal{S}\subseteq[1:N]\\ n\in\mathcal{S}}} H(W_{k,\mathcal{S}})$$
$$\overset{(b)}{=} \frac{1}{K}\sum_{k=1}^{K}\sum_{\substack{\mathcal{S}\subseteq[1:N]\\ |\mathcal{S}|\geq 1}} |\mathcal{S}|\,|W_{k,\mathcal{S}}|$$
$$= \sum_{\ell=1}^{N}\frac{\ell}{K}\sum_{k=1}^{K}\sum_{\substack{\mathcal{S}\subseteq[1:N]\\ |\mathcal{S}|=\ell}} |W_{k,\mathcal{S}}| \overset{(c)}{=} \sum_{\ell=1}^{N}\ell\binom{N}{\ell}x_\ell, \tag{16}$$

where (a) follows from (11), (b) is true because when we sum up the contents of the storage at all the databases, the chunk $W_{k,\mathcal{S}}$ is counted $|\mathcal{S}|$ number of times, which is the number of databases storing this chunk, and (c) follows from the definition of x_ℓ in (15). The message size and storage constraints defined in this sub-section will be used in the converse proofs for PIR from uncoded storage constrained databases.

Remark 1 (Implication of Uncoded Storage Assumption): The uncoded storage assumption of the storage in (11) implies that for any subset $\mathcal{K} \in [1:K]$ of messages, $\mathbf{W}_\mathcal{K}$, and subset $\mathcal{N} \in [1:N]$ of databases, $\mathbf{Z}_\mathcal{N}$, there is a Markov chain $\mathbf{W}_\mathcal{K} - \mathbf{Z}_\mathcal{N} - \mathbf{W}_{[1:K]\setminus\mathcal{K}}$. In other words the messages $\mathbf{W}_\mathcal{K}$ and the messages $\mathbf{W}_{[1:K]\setminus\mathcal{K}}$ are independent given the storage contents of the databases, i.e.,

$$H(\mathbf{W}_\mathcal{K}|\mathbf{Z}_\mathcal{N}, \mathbf{W}_{[1:K]\setminus\mathcal{K}}) = H(\mathbf{W}_\mathcal{K}|\mathbf{Z}_\mathcal{N}). \tag{17}$$

That is due to the fact that messages are i.i.d., and due to the nature of uncoded storage placement, where $\mathbf{W}_{[1:K]\setminus\mathcal{K}}$ cannot be used to decode any information of $\mathbf{W}_\mathcal{K}$ from $\mathbf{Z}_\mathcal{N}$.

III. Main Result and Discussions

Our first result is a general information theoretic lower bound on the download cost of the PIR problem with any arbitrary storage at the databases.

Theorem 1: For the storage constrained PIR problem with N databases, K messages (of size L bits each), and arbitrary storage $Z_1, Z_2, \ldots, Z_N$ at the N databases, the optimal download cost is lower bounded as follows,

$$D^*(\mu) \geq 1 + \sum_{n_1=1}^{N}\frac{\lambda_{(N-n_1,1)}}{n_1} + \sum_{n_1=1}^{N}\sum_{n_2=n_1}^{N}\frac{\lambda_{(N-n_1,2)}}{n_1 n_2}$$
$$+\cdots+\sum_{n_1=1}^{N}\cdots\sum_{n_{K-1}=n_{K-2}}^{N}\frac{\lambda_{(N-n_1,K-1)}}{n_1\times\cdots\times n_{K-1}}, \tag{18}$$

where $\lambda_{(n,k)}$ is defined as follows,

$$\lambda_{(n,k)} \triangleq \frac{1}{KL\binom{K-1}{k}\binom{N}{n}}\sum_{\substack{\mathcal{K}\subseteq[1:K]\\ |\mathcal{K}|=k}}\sum_{\substack{\mathcal{N}\subseteq[1:N]\\ |\mathcal{N}|=n}}\sum_{j\in[1:K]\setminus\mathcal{K}} H(W_j|\mathbf{Z}_\mathcal{N}, \mathbf{W}_\mathcal{K}), \tag{19}$$

for $n \in [0:N]$ and $k \in [0:K]$.

The complete proof of Theorem 1 is presented in Section IV.

Boundary Conditions on the function $\lambda_{(n,k)}$:

- We notice that when $n = N$ or $k = K$, then we get all the messages in the conditioning of the entropy terms of $\lambda_{(n,k)}$ in (19), and therefore we get the following boundary conditions on $\lambda_{(n,k)}$:

$$\lambda_{(n=N,k)} = 0,\ \forall k \in [0:K],$$
$$\lambda_{(n,k=K)} = 0,\ \forall n \in [0:N]. \tag{20}$$

- We further notice that for $n = 0$ and all $k \in [0:K]$, then we only have messages in the conditioning of the entropy terms in (19) which are i.i.d., therefore, we get another set of boundary conditions on $\lambda_{(n,k)}$ for all $k \in [0:K]$:

$$\lambda_{(n=0,k)} = \frac{1}{KL\binom{K-1}{k}\binom{N}{n}}\sum_{\substack{\mathcal{K}\subseteq[1:K]\\ |\mathcal{K}|=k}}\sum_{\substack{\mathcal{N}\subseteq[1:N]\\ |\mathcal{N}|=n}}\sum_{j\in[1:K]\setminus\mathcal{K}} H(W_j)$$
$$= \frac{1}{KL\binom{K-1}{k}\binom{N}{n}}\sum_{j=1}^{K}\sum_{\substack{\mathcal{K}\subseteq[1:K]\setminus j\\ |\mathcal{K}|=k}}\sum_{\substack{\mathcal{N}\subseteq[1:N]\\ |\mathcal{N}|=n}} L = 1. \tag{21}$$

- For the replicated databases case, considered in [7] where every database stores all the files, then for the function $\lambda_{(n,k)}$ where $n \in [1:N]$, we retain all the messages in the conditioning of the entropy terms in (19), which gives the following conditions over $\lambda_{(n,k)}$:

$$\lambda_{(n,k)} = 0,\ \forall n \in [1:N],\ \forall k \in [0:K]. \tag{22}$$

Remark 2 (Replicated Databases as a Special Case of Theorem 1): We notice that for the replicated databases case considered in [7], by applying the boundary conditions in (21) and (22) to the general lower bound in Theorem 1, we get the lower bound previously obtained in [7] as follows,

$$D^*(\mu=1) \geq 1 + \sum_{n_1=1}^{N}\frac{\lambda_{(N-n_1,1)}}{n_1} + \sum_{n_1=1}^{N}\sum_{n_2=n_1}^{N}\frac{\lambda_{(N-n_1,2)}}{n_1 n_2}$$
$$+\cdots+\sum_{n_1=1}^{N}\cdots\sum_{n_{K-1}=n_{K-2}}^{N}\frac{\lambda_{(N-n_1,K-1)}}{n_1\times\cdots\times n_{K-1}}$$
$$\overset{(a)}{=} 1 + \sum_{n_1=N}^{N}\frac{\lambda_{(N-n_1,1)}}{n_1} + \sum_{n_1=N}^{N}\sum_{n_2=n_1}^{N}\frac{\lambda_{(N-n_1,2)}}{n_1 n_2}$$
$$+\cdots+\sum_{n_1=N}^{N}\cdots\sum_{n_{K-1}=n_{K-2}}^{N}\frac{\lambda_{(N-n_1,K-1)}}{n_1\times\cdots\times n_{K-1}}$$
$$= 1 + \frac{\lambda_{(0,1)}}{N} + \frac{\lambda_{(0,2)}}{N^2} + \cdots + \frac{\lambda_{(0,K-1)}}{N^{K-1}}$$
$$\overset{(b)}{=} 1 + \frac{1}{N} + \frac{1}{N^2} + \cdots + \frac{1}{N^{K-1}}, \tag{23}$$

where (a) follows from the boundary conditions on $\lambda_{(n,k)}$ from replicated databases in (22), and (b) follows from the boundary condition in (21).

The following Theorem summarizes the second main result of this article, which characterizes the information theoretically optimal download cost of the PIR problem from uncoded storage constrained databases as a function of the available storage.

Authorized licensed use limited to: The University of Arizona. Downloaded on January 30,2021 at 20:39:51 UTC from IEEE Xplore. Restrictions apply.

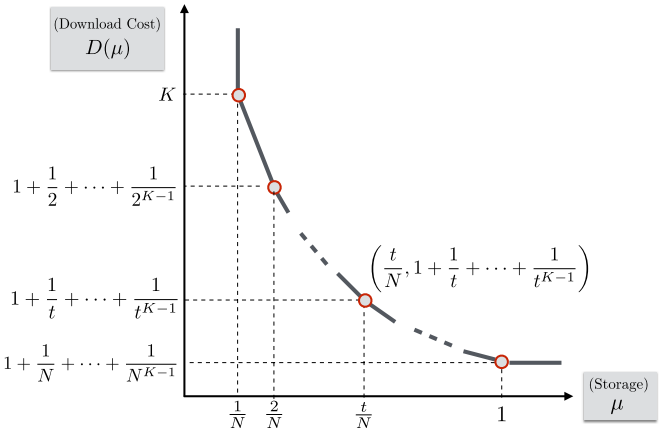

Fig. 2. The optimal trade-off between storage and download cost for uncoded storage constrained PIR.

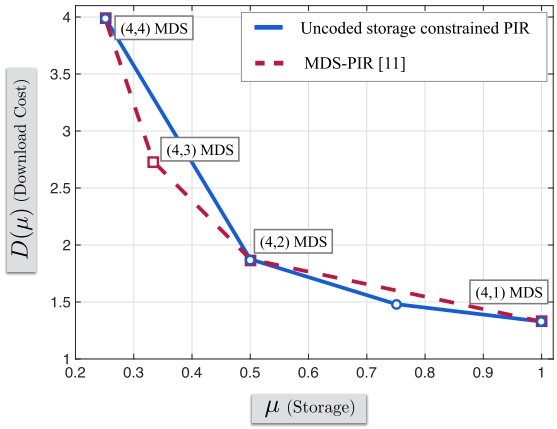

Fig. 3. Optimal trade-off between download and storage for uncoded storage constrained PIR compared to MDS-PIR where $(N, K) = (4, 4)$.

Theorem 2: For the uncoded storage constrained PIR problem with K messages (of size L bits each), and N homogeneous storage constrained databases of μKL bits, the information-theoretically optimal trade-off between storage and download cost is a piece-wise linear function given by the lower convex hull of the following $(\mu, D^*(\mu))$ pairs, for $t = 1, 2, \ldots, N$:

$$\left(\mu = \frac{t}{N},\ D^*(\mu) = \tilde{D}(t)\right), \tag{24}$$

where $\tilde{D}(t)$ is defined as follows for $t \in [1 : N]$:

$$\tilde{D}(t) \triangleq \sum_{k=0}^{K-1} \frac{1}{t^k}. \tag{25}$$

The general achievable scheme for any (N, K) and any μ is described in Section VI. The converse proof of Theorem 2 is an application to Theorem 1 for the special case of homogeneous uncoded storage constrained databases with the proof in details presented in Section V.

The optimal trade-off resulting from Theorem 2 is illustrated in Figure 2. The smallest value of $\mu = 1/N$ corresponds to the parameter $t = 1$, for which the optimal download cost is maximal and is equal to K, corresponding to download all the messages from the databases. The other extreme value of storage is $\mu = 1$, corresponding to $t = N$, i.e., the setting of full storage in which every database can store all the messages. For this case, the optimal download cost was characterized in [7] as $(1 + \frac{1}{N} + \frac{1}{N^2} + \cdots + \frac{1}{N^{K-1}})$. The PIR download cost for the storage values in between outperforms memory sharing between the two extremes, i.e., lower than the line joining between them.

Remark 3 (Applications of Theorem 1 for Other Variants of PIR With Uncoded Storage): The converse proof of Theorem 2 for homogeneous uncoded storage constrained databases is a direct application of Theorem 1. The term $\lambda_{(n,k)}$ in Theorem 1 signifies the normalized average remaining entropy in a message after conditioning on k other messages and the storage from n databases. We note that the result in Theorem 1 can be applied to other models beyond homogeneous uncoded storage databases. Recently, the lower bound in Theorem 1 was applied in [28] to characterize the capacity of decentralized uncoded storage constrained databases where the databases can design their own storage independently at random. The lower bound in Theorem 1 was also proven tight for the case of heterogeneous uncoded storage constrained databases [30], where databases have heterogeneous sizes.

Remark 4 (Significance of Coded Storage to Further Reduce Download Cost): The problem of PIR with databases storing (N, M) MDS coded (or MDS-PIR) messages was considered in [10] to satisfy the M out of N recoverability constraint. In the MDS-PIR scheme, each database stores $\mu_{\text{MDS}} KL$ bits, where $\mu_{\text{MDS}} = 1/M$ and the PIR scheme of [10] achieves the optimal download cost given as $D^*_{\text{MDS}}(\mu_{\text{MDS}}) = 1 + \frac{M}{N} + \frac{M^2}{N^2} + \cdots + \frac{M^{K-1}}{N^{K-1}}$. In Figure 3, we plot the tradeoff between storage and download cost for the MDS-PIR by varying the value of M in the range $M \in [1 : N]$ and compare to the optimal tradeoff for uncoded storage constrained PIR given in Theorem 2 for $N = K = 4$. We notice that our scheme achieves better download cost for higher storage values $\mu \geq 0.5$. At a first glance, it might appear that optimal MDS coded scheme should be better than (or at least same as) our uncoded storage scheme in terms of download cost. However, due to the additional recoverability constraint, MDS-PIR capacity is lower than SC-PIR for some storage values. For storage values $\mu < 0.5$, MDS-PIR scheme performs better in terms of download cost and has the added benefit of data recoverability. While the two schemes are optimal under different assumptions, this comparison shows that coded storage can be beneficial in general to further reduce the download cost.

We briefly describe here the main elements of the proofs of Theorems 1 and 2 through an example of $N = 3$ databases and $K = 3$ messages.

A. Sketch Proof of Theorem 1 for $N = K = 3$

We start by using following bound on $D^{[1]}_{\mu}$ which was first found in [7, Lemma 1] for $N = K = 3$:

$$D^{[1]}_{\mu} - L + o(L) \geq I(\mathbf{W}_{[2:3]}; \mathbf{Q}^{[1]}_{[1:3]}, \mathbf{A}^{[1]}_{[1:3]} | W_1). \tag{26}$$

Authorized licensed use limited to: The University of Arizona. Downloaded on January 30,2021 at 20:39:51 UTC from IEEE Xplore. Restrictions apply.

The above bound can be interpreted as follows: given message W_1 is requested, then the privacy penalty $D_\mu - L$ is bounded by the amount of information the queries and answers tell about the remaining messages $\mathbf{W}_{[2:3]}$ after successfully decoding message W_1. Later, we prove a more general form of this bound in Lemma 2 (Section IV).

Using the chain rule for mutual information in all possible orders for a permutation $\boldsymbol{\sigma} : (1,2,3) \to (\sigma_1,\sigma_2,\sigma_3)$, we expand the RHS of the bound in (26) as,

$$\begin{aligned}
&I(\mathbf{W}_{[2:3]}; \mathbf{Q}^{[1]}_{[1:3]}, \mathbf{A}^{[1]}_{[1:3]}|W_1)\\
&= I(\mathbf{W}_{[2:3]}; Q^{[1]}_{\sigma_1}, A^{[1]}_{\sigma_1}|W_1)\\
&\quad + I(\mathbf{W}_{[2:3]}; Q^{[1]}_{\sigma_2}, A^{[1]}_{\sigma_2}|W_1, Q^{[1]}_{\sigma_1}, A^{[1]}_{\sigma_1})\\
&\quad + I(\mathbf{W}_{[2:3]}; Q^{[1]}_{\sigma_3}, A^{[1]}_{\sigma_3}|W_1, \mathbf{Q}^{[1]}_{\sigma_{[1:2]}}, \mathbf{A}^{[1]}_{\sigma_{[1:2]}}) \qquad (27)\\
&\overset{(a)}{\geq} I(\mathbf{W}_{[2:3]}; Q^{[1]}_{\sigma_1}, A^{[1]}_{\sigma_1}|W_1) + I(\mathbf{W}_{[2:3]}; Q^{[1]}_{\sigma_2}, A^{[1]}_{\sigma_2}|W_1, Z_{\sigma_1})\\
&\quad + I(\mathbf{W}_{[2:3]}; Q^{[1]}_{\sigma_3}, A^{[1]}_{\sigma_3}|W_1, Z_{\sigma_1}, Z_{\sigma_2})\\
&\overset{(b)}{=} I(\mathbf{W}_{[2:3]}; Q^{[2]}_{\sigma_1}, A^{[2]}_{\sigma_1}|W_1) + I(\mathbf{W}_{[2:3]}; Q^{[2]}_{\sigma_2}, A^{[2]}_{\sigma_2}|W_1, Z_{\sigma_1})\\
&\quad + I(\mathbf{W}_{[2:3]}; Q^{[2]}_{\sigma_3}, A^{[2]}_{\sigma_3}|W_1, Z_{\sigma_1}, Z_{\sigma_2}), \qquad (28)
\end{aligned}$$

where (a) follows by bounding the second and the third terms in (27) separately where later in Lemma 1 (Section IV), we generally prove that the mutual information terms with queries and answers of some databases in the conditioning can be lower bounded by replacing the queries and the answers with the corresponding databases storage random variables; and (b) follows from the privacy constraint in (6) where the individual queries and answers are invariant with respect to the requested message index. We would like to point out here that in the original PIR model (replicated databases), the second and third terms in (28) were lower bounded by zero. This was tight in that setting with any DB storage in the conditioning having all the messages, and hence the mutual information terms will be zero. However, for the SC-PIR model in this article, these terms will not be zero as the storage constrained DB may not contain all the messages.

Next, we sum (28) over all possible permutations $\boldsymbol{\sigma} \in [3!]$ to get the following bound:

$$\begin{aligned}
&I(\mathbf{W}_{[2:3]}; \mathbf{Q}^{[1]}_{[1:3]}, \mathbf{A}^{[1]}_{[1:3]}|W_1)\\
&\geq \frac{1}{3}\sum_{i=1}^{3} I(\mathbf{W}_{[2:3]}; Q^{[2]}_i, A^{[2]}_i|W_1)\\
&\quad + \frac{1}{6}\sum_{i=1}^{3}\sum_{j\in[1:3]\setminus i} I(\mathbf{W}_{[2:3]}; Q^{[2]}_j, A^{[2]}_j|W_1, Z_i)\\
&\quad + \frac{1}{3}\sum_{i=1}^{3} I(\mathbf{W}_{[2:3]}; Q^{[2]}_i, A^{[2]}_i|W_1, \mathbf{Z}_{[1:3]\setminus i})\\
&\overset{(a)}{\geq} \frac{1}{3}\sum_{i=1}^{3} H(A^{[2]}_i|W_1, Q^{[2]}_i)\\
&\quad + \frac{1}{6}\sum_{i=1}^{3}\sum_{j\in[1:3]\setminus i} H(A^{[2]}_j|W_1, Z_i, Q^{[2]}_j)\\
&\quad + \frac{1}{3}\sum_{i=1}^{3} H(A^{[2]}_i|W_1, \mathbf{Z}_{[1:3]\setminus i}, Q^{[2]}_i)\\
&\overset{(b)}{\geq} \frac{1}{3}H(\mathbf{A}^{[2]}_{[1:3]}|W_1, \mathbf{Q}^{[2]}_{[1:3]})\\
&\quad + \frac{1}{6}\sum_{i=1}^{3} H(\mathbf{A}^{[2]}_{[1:3]\setminus i}|W_1, Z_i, \mathbf{Q}^{[2]}_{[1:3]})\\
&\quad + \frac{1}{3}\sum_{i=1}^{3} H(A^{[2]}_i|W_1, \mathbf{Z}_{[1:3]\setminus i}, \mathbf{Q}^{[2]}_{[1:3]})\\
&\overset{(c)}{=} \frac{1}{3}I(\mathbf{W}_{[2:3]}; \mathbf{A}^{[2]}_{[1:3]}|W_1, \mathbf{Q}^{[2]}_{[1:3]})\\
&\quad + \frac{1}{6}\sum_{i=1}^{3} I(\mathbf{W}_{[2:3]}; \mathbf{A}^{[2]}_{[1:3]\setminus i}|W_1, Z_i, \mathbf{Q}^{[2]}_{[1:3]})\\
&\quad + \frac{1}{3}\sum_{i=1}^{3} I(\mathbf{W}_{[2:3]}; A^{[2]}_i|W_1, \mathbf{Z}_{[1:3]\setminus i}, \mathbf{Q}^{[2]}_{[1:3]})\\
&\overset{(d)}{=} \frac{1}{3}I(\mathbf{W}_{[2:3]}; W_2, \mathbf{Q}^{[2]}_{[1:3]}, \mathbf{A}^{[2]}_{[1:3]}|W_1)\\
&\quad + \frac{1}{6}\sum_{i=1}^{3} I(\mathbf{W}_{[2:3]}; W_2, \mathbf{Q}^{[2]}_{[1:3]}, \mathbf{A}^{[2]}_{[1:3]\setminus i}|W_1, Z_i)\\
&\quad + \frac{1}{3}\sum_{i=1}^{3} I(\mathbf{W}_{[2:3]}; W_2, \mathbf{Q}^{[2]}_{[1:3]}, A^{[2]}_i|W_1, \mathbf{Z}_{[1:3]\setminus i}) + o(L)\\
&= \frac{1}{3}H(W_2|W_1) + \frac{1}{6}\sum_{i=1}^{3} H(W_2|W_1, Z_i)\\
&\quad + \frac{1}{3}\sum_{i=1}^{3} H(W_2|W_1, \mathbf{Z}_{[1:3]\setminus i})\\
&\quad + \frac{1}{3}\underbrace{I(W_3; \mathbf{Q}^{[2]}_{[1:3]}, \mathbf{A}^{[2]}_{[1:3]}|\mathbf{W}_{[1:2]})}_{\triangleq \text{Term}_1}\\
&\quad + \frac{1}{6}\sum_{i=1}^{3}\underbrace{I(W_3; \mathbf{Q}^{[2]}_{[1:3]\setminus i}, \mathbf{A}^{[2]}_{[1:3]\setminus i}|\mathbf{W}_{[1:2]}, Z_i)}_{\triangleq \text{Term}_2}\\
&\quad + \frac{1}{3}\sum_{i=1}^{3}\underbrace{I(W_3; Q^{[2]}_i, A^{[2]}_i|\mathbf{W}_{[1:2]}, \mathbf{Z}_{[1:3]\setminus i})}_{\triangleq \text{Term}_3} + o(L), \qquad (29)
\end{aligned}$$

where (a) and (c) follow since any answer $A^{[2]}_i$ is a function of the messages $\mathbf{W}_{[1:K]}$ and the query $Q^{[2]}_i$; (b) follows from the union bound and since conditioning reduces entropy; and (d) follows from the fact that queries are independent from the messages, then from the decoding constraint in (7) where W_2 is decodable from $\mathbf{Q}^{[2]}_{[1:3]}$, $\mathbf{A}^{[2]}_{[1:3]\setminus\mathcal{N}}$ and $\mathbf{Z}_{\mathcal{N}}$ for any $\mathcal{N} \subseteq [1:3]$.

Next, we can lower bound the three terms, Term_1, Term_2 and Term_3, in (29) in a similar manner to get the following bounds:

$$\begin{aligned}
\text{Term}_1 &\geq \frac{1}{3}H(W_3|\mathbf{W}_{[1:2]}) + \frac{1}{6}\sum_{i=1}^{3} H(W_3|\mathbf{W}_{[1:2]}, Z_i)\\
&\quad + \frac{1}{3}\sum_{i=1}^{3} H(W_3|\mathbf{W}_{[1:2]}, \mathbf{Z}_{[1:3]\setminus i}) + o(L),\\
\text{Term}_2 &\geq \frac{1}{2}H(W_3|\mathbf{W}_{[1:2]}, Z_i)\\
&\quad + \frac{1}{2}\sum_{j\in[1:3]\setminus i} H(W_3|\mathbf{W}_{[1:2]}, \mathbf{Z}_{[1:3]\setminus j}) + o(L),\\
\text{Term}_3 &\geq H(W_3|\mathbf{W}_{[1:2]}, \mathbf{Z}_{[1:3]\setminus i}) + o(L),
\end{aligned}$$

Authorized licensed use limited to: The University of Arizona. Downloaded on January 30,2021 at 20:39:51 UTC from IEEE Xplore. Restrictions apply.

and plug these back in (29) to arrive at the following bound:

$$\begin{aligned} D_\mu^{[1]} \geq & L + \frac{1}{3}H(W_2|W_1) + \frac{1}{9}H(W_3|\mathbf{W}_{[1:2]}) \\ & + \frac{1}{6}\sum_{i=1}^{3} H(W_2|W_1, Z_i) + \frac{1}{3}\sum_{i=1}^{3} H(W_2|W_1, \mathbf{Z}_{[1:3]\setminus i}) \\ & + \frac{5}{36}\sum_{i=1}^{3} H(W_3|\mathbf{W}_{[1:2]}, Z_i) \\ & + \frac{11}{18}\sum_{i=1}^{3} H(W_3|\mathbf{W}_{[1:2]}, \mathbf{Z}_{[1:3]\setminus i}) + o(L). \end{aligned} \tag{30}$$

By repeating the bounding procedure in (26) with any permutation of the messages indexes $\boldsymbol{\pi} : (1, 2, 3) \to (\pi_1, \pi_2, \pi_3)$, and then sum up over all permutations $\boldsymbol{\pi} \in [3!]$, we obtain the following bound on the average number of downloaded bits D_μ,

$$\begin{aligned} D_\mu \geq & L + o(L) + \frac{1}{6}\sum_{\substack{\mathcal{K}\subseteq[1:3]\\|\mathcal{K}|=1}}\sum_{k\in[1:3]\setminus\mathcal{K}} \left(\frac{1}{3}H(W_k|\mathbf{W}_\mathcal{K}) \right. \\ & \left. + \frac{1}{6}\sum_{i=1}^{3} H(W_k|\mathbf{W}_\mathcal{K}, Z_i) + \frac{1}{3}\sum_{i=1}^{3} H(W_k|\mathbf{W}_\mathcal{K}, \mathbf{Z}_{[1:3]\setminus i})\right) \\ & + \frac{1}{6}\sum_{\substack{\mathcal{K}\subseteq[1:3]\\|\mathcal{K}|=2}}\sum_{k\in[1:3]\setminus\mathcal{K}} \left(\frac{2}{9}H(W_k|\mathbf{W}_\mathcal{K}) \right. \\ & \left. + \frac{5}{18}\sum_{i=1}^{3} H(W_k|\mathbf{W}_\mathcal{K}, Z_i) + \frac{11}{9}\sum_{i=1}^{3} H(W_k|\mathbf{W}_\mathcal{K}, \mathbf{Z}_{[1:3]\setminus i})\right) \\ = & L + \frac{1}{3}\lambda_{(0,1)}L + \frac{1}{2}\lambda_{(1,1)}L + \lambda_{(2,1)}L + \frac{2}{18}\lambda_{(0,2)}L \\ & + \frac{5}{12}\lambda_{(1,2)}L + \frac{11}{6}\lambda_{(2,2)}L + o(L), \end{aligned} \tag{31}$$

where $\lambda_{(n,k)}$ is defined in (19).

Since the bound in (31) is valid for any achievable pair (D_μ, L), it is also a valid bound on the optimal download cost, $D^*(\mu)$, as defined in (7), where $\mu \in [\frac{1}{3}, 1]$. Therefore, by taking the limit $L \to \infty$, we obtain the following bound on $D^*(\mu)$:

$$\begin{aligned} D^*(\mu) \geq & 1 + \frac{1}{3}\lambda_{(0,1)} + \frac{1}{2}\lambda_{(1,1)} + \lambda_{(2,1)} + \frac{2}{18}\lambda_{(0,2)} \\ & + \frac{5}{12}\lambda_{(1,2)} + \frac{11}{6}\lambda_{(2,2)}, \end{aligned} \tag{32}$$

which satisfies the bound in Theorem 1 for $N = K = 3$.

B. Sketch Proof of Theorem 2 for $N = K = 3$ - Converse Proof

Following Theorem 2, the optimal trade-off for the case $N = K = 3$ has three corner points as shown in Figure 4: The corner point P_1 ($\mu = 1/3$) where the optimal scheme is to download all messages to ensure privacy; the corner point P_3 ($\mu = 1$) which corresponds to the replicated databases case considered in [7]; and the middle corner point P_2 ($\mu = 1/2$) where the optimal trade-off outperforms memory sharing between P_1 and P_3.

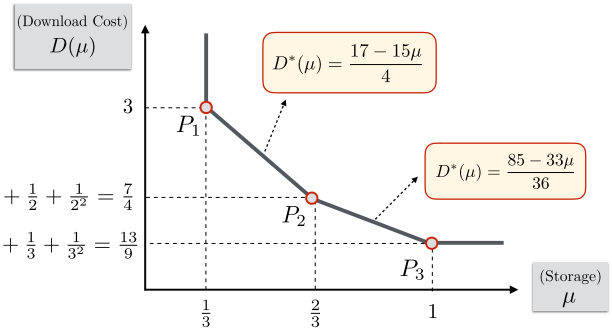

Fig. 4. Optimal trade-off between download and storage for $(N, K) = (3, 3)$. Following Theorem 2, the trade-off has three corner points, labeled as P_1, P_2 and P_3.

From Figure 4, it is clear that we need to need to prove the following two lower bounds on the download cost:

$$D^*(\mu) \geq \frac{17 - 15\mu}{4}, \qquad D^*(\mu) \geq \frac{85 - 33\mu}{36}. \tag{33}$$

To this end, we now specialize the lower bound in (31) for the case of uncoded storage placement using Remark 1 as follows:

$$\begin{aligned} D_\mu \geq & L + \frac{1}{6}\sum_{\substack{\mathcal{K}\subseteq[1:3]\\|\mathcal{K}|=1}}\sum_{k\in[1:3]\setminus\mathcal{K}} \left(\frac{1}{3}H(W_k) + \frac{1}{6}\sum_{i=1}^{3} H(W_k|Z_i) \right. \\ & \left. + \frac{1}{3}\sum_{i=1}^{3} H(W_k|\mathbf{Z}_{[1:3]\setminus i})\right) \\ & + \frac{1}{6}\sum_{\substack{\mathcal{K}\subseteq[1:3]\\|\mathcal{K}|=2}}\sum_{k\in[1:3]\setminus\mathcal{K}} \left(\frac{2}{9}H(W_k) + \frac{5}{18}\sum_{i=1}^{3} H(W_k|Z_i) \right. \\ & \left. + \frac{11}{9}\sum_{i=1}^{3} H(W_k|\mathbf{Z}_{[1:3]\setminus i})\right) + o(L) \\ = & L + \frac{4}{27}\sum_{k=1}^{3} H(W_k) + \frac{11}{108}\sum_{i=1}^{3}\sum_{k=1}^{3} H(W_k|Z_i) \\ & + \frac{17}{54}\sum_{i=1}^{3}\sum_{k=1}^{3} H(W_k|\mathbf{Z}_{[1:3]\setminus i}) + o(L). \end{aligned} \tag{34}$$

Next, we use the representation of the messages for uncoded storage given in (12), where W_k can be partitioned into disjoint parts, to write the lower bound in (34) in terms of the variable x_ℓ defined in (15) as follows,

$$\begin{aligned} D_\mu \overset{(a)}{\geq} & L + o(L) + \frac{4}{27}\sum_{k=1}^{3}\sum_{\substack{\mathcal{S}\subseteq[1:3]\\|\mathcal{S}|\geq 1}} |W_{k,\mathcal{S}}|L \\ & + \frac{11}{108}\sum_{i=1}^{3}\sum_{k=1}^{3}\sum_{\substack{\mathcal{S}\subseteq[1:3]\setminus i\\|\mathcal{S}|\geq 1}} |W_{k,\mathcal{S}}|L + \frac{17}{54}\sum_{i=1}^{3}\sum_{k=1}^{3}|W_{k,\{i\}}|L \\ = & L + \frac{2}{3}\sum_{k=1}^{3}\sum_{\substack{\mathcal{S}\subseteq[1:3]\\|\mathcal{S}|=1}} |W_{k,\mathcal{S}}|L + \frac{1}{4}\sum_{k=1}^{3}\sum_{\substack{\mathcal{S}\subseteq[1:3]\\|\mathcal{S}|=2}} |W_{k,\mathcal{S}}|L \\ & + \frac{4}{27}\sum_{k=1}^{3}\sum_{\substack{\mathcal{S}\subseteq[1:3]\\|\mathcal{S}|=3}} |W_{k,\mathcal{S}}|L + o(L) \\ \overset{(b)}{=} & L + 6x_1L + \frac{9}{4}x_2L + \frac{4}{9}x_3L + o(L), \end{aligned} \tag{35}$$

Authorized licensed use limited to: The University of Arizona. Downloaded on January 30,2021 at 20:39:51 UTC from IEEE Xplore. Restrictions apply.

where (a) follows from (12); and (b) follows from the definition of x_ℓ in (15).

Since the bound in (35) is valid for any achievable pair (D_μ, L), it is also a valid bound on the optimal download cost, $D^*(\mu)$, as defined in (7), where $\mu \in [\frac{1}{3}, 1]$. Therefore, by dividing (35) and taking the limit $L \to \infty$, we obtain the following bound on $D^*(\mu)$:

$$D^*(\mu) \geq 1 + 6x_1 + \frac{9}{4}x_2 + \frac{4}{9}x_3. \tag{36}$$

Moreover, the message size, and the storage constraints for uncoded storage placement for this example $N = K = 3$ follow from (14), and (16), respectively. Hence, we obtain the following constraints:

$$3x_1 + 3x_2 + x_3 = 1, \tag{37}$$

$$3x_1 + 6x_2 + 3x_3 \leq 3\mu. \tag{38}$$

We solve the system of linear inequalities in (36), (37), and (38) using Gaussian elimination to obtain the following two lower bounds on $D^*(\mu)$ which completes the converse proof for $N = K = 3$,

$$D^*(\mu) \geq \frac{17 - 15\mu}{4} + \frac{17}{18}x_3 \overset{(a)}{\geq} \frac{17 - 15\mu}{4}, \tag{39}$$

$$D^*(\mu) \geq \frac{85 - 33\mu}{36} + \frac{17}{6}x_1 \overset{(b)}{\geq} \frac{85 - 33\mu}{36}, \tag{40}$$

where steps (a) and (b) follow since $x_1, x_3 \geq 0$ by definition.

C. Sketch Proof of Theorem 2 for $N = K = 3$ - Achievable Scheme

The optimal trade-off is achieved by memory sharing between different PIR schemes (see Claim 1), which are designed for three values of storage $\mu \in \{1/3, 2/3, 1\}$. The storage placement of our SC-PIR scheme is inspired by the storage placement strategy for caching systems in [32]. Without loss of generality, consider the case where the user wants to retrieve message W_1. Same scheme can be applied for any other requested message due to the symmetry of the setting.

- **Case $\mathbf{P_1}$ ($\mathbf{t = 1}$ or $\mu = \mathbf{1/3}$):**

Storage Placement: For storage placement, we split each message into $\binom{3}{1} = 3$ sub-messages and label each by a unique subset of $[1 : 3]$ of size $t = 1$, i.e., $W_k = \{W_{k,\{1\}}, W_{k,\{2\}}, W_{k,\{3\}}\}$ for $k \in [1 : 3]$, and each sub-message is of size $L/3$ bits. Subsequently, DB_n stores those sub-messages (of each message) whose indexes contains n. For instance, DB_1 stores $\{W_{1,\{1\}}, W_{2,\{1\}}, W_{3,\{1\}}\}$, which satisfies the storage constraint of $\mu KL = L$ bits.

PIR Scheme: The PIR scheme is trivial for this storage point, where in order to maintain privacy all the messages should be downloaded from the databases for any message request. Hence, the download cost for this scheme is given as $D(\mu = \frac{1}{3}) = 9 \times \frac{1}{3} = 1/3$, and point P_1 is achieved.

- **Case $\mathbf{P_2}$ ($\mathbf{t = 2}$ or $\mu = \mathbf{2/3}$):**

Storage Placement: Here, we split each message into $\binom{3}{2} = 3$ sub-messages and label each by a unique subset of $[1 : 3]$ of size $t = 2$, i.e., $W_k = \{W_{k,\{1,2\}}, W_{k,\{1,3\}}, W_{k,\{2,3\}}\}$ for $k \in [1 : 3]$. Each sub-message is of size $L/3$ bits. Subsequently, DB_n stores those sub-messages (of each message) whose index contains n. For instance, DB_1 stores $\{W_{1,\{1,2\}}, W_{1,\{1,3\}}, W_{2,\{1,2\}}, W_{2,\{1,3\}}, W_{3,\{1,2\}}, W_{3,\{1,3\}}\}$, which satisfies the storage constraint of $\mu KL = 2L$ bits.

PIR Scheme: The storage constrained PIR scheme in this case works in 3 blocks, where in every block, only 2 databases are involved whose indexes are in the set $\mathcal{S} \subset [1 : 3]$ where $|\mathcal{S}| = 2$. In each block labeled with $\mathcal{S}$, we apply the original PIR scheme proposed in [7] with $N' = t = 2$ and $K = 3$ only involving the sub-messages $W_{k,\mathcal{S}}$ for $k \in [1 : 3]$. This is enabled by our storage placement scheme for storage constrained databases. Hence, the average download cost equals to that of each block and is given by $D(\mu = 2/3) = 1 + \frac{1}{2} + \frac{1}{2^2} = 7/4$, and point P_2 is achieved.

- **Case $\mathbf{P_3}$ ($\mathbf{t = 3}$ or $\mu = \mathbf{1}$):**

Storage Placement: The storage placement is trivial in this case, where all the databases can store the 3 messages completely, i.e., $\mu KL = 3L$ bits.

PIR Scheme: This storage point is the replicated databases case considered in [7]. Applying the scheme in [7], we achieve a download cost $D(\mu = 1) = (1 + \frac{1}{3} + \frac{1}{3^2}) = 13/9$, and point P_3 is achieved.

Finally, the intermediate values of μ, between the points P_1, P_2, and P_3, can be achieved by memory-sharing (see Claim 1), showing that the lower convex hull given in Figure 4 is achievable. Therefore, the scheme is information-theoretically optimal for $N = 3$, and $K = 3$.

IV. Proof of Theorem 1: General Lower Bound on $D^*(\mu)$

We start by proving the following Lemma, which provides an information theoretic bound useful in many steps of the general converse proof.

Lemma 1: For any $\mathcal{N} \subseteq [1 : N]$, $\mathcal{K} \subseteq [1 : K]$, $i \in [1 : N]$, and $j \in [1 : K]$ we can write the following lower bound:

$$I(\mathbf{W}_{[1:K]\setminus\mathcal{K}}; Q_i^{[j]}, A_i^{[j]} | \mathbf{W}_\mathcal{K}, \mathbf{Q}_\mathcal{N}^{[j]}, \mathbf{A}_\mathcal{N}^{[j]}) \geq I(\mathbf{W}_{[1:K]\setminus\mathcal{K}}; Q_i^{[j]}, A_i^{[j]} | \mathbf{W}_\mathcal{K}, \mathbf{Z}_\mathcal{N}). \tag{41}$$

The proof of Lemma 1 can be found in Appendix B.

Remark 5: Lemma 1 lower bounds the mutual information $I(\mathbf{W}_{[1:K]\setminus\mathcal{K}}; Q_i^{[j]}, A_i^{[j]} | \mathbf{W}_\mathcal{K}, \mathbf{Q}_\mathcal{N}^{[j]}, \mathbf{A}_\mathcal{N}^{[j]})$ by replacing the queries and answers $\mathbf{Q}_\mathcal{N}^{[j]}, \mathbf{A}_\mathcal{N}^{[j]}$ in the conditioning, with the storage contents of the corresponding subset of databases $\mathbf{Z}_\mathcal{N}$. This Lemma is repeatedly used in our converse proof for the download cost.

The following Lemma gives a lower bound on the number of downloaded bits D_μ in terms of a summation of mutual information functions in the form $I(\mathbf{W}_{[1:K]\setminus\mathcal{K}}; Q_i^{[j]}, A_i^{[j]} | \mathbf{W}_\mathcal{K}, \mathbf{Z}_\mathcal{N})$. Notice that this mutual information form appears in right side of the bound in Lemma 1. The proof of Lemma 2 can be found in Appendix C.

Lemma 2: The average number of downloaded bits D_μ of the storage-constrained PIR is lower bounded as follows:

$$D_\mu \geq L + \sum_{n=0}^{N-1} T(n, 1) + o(L), \tag{42}$$

Authorized licensed use limited to: The University of Arizona. Downloaded on January 30,2021 at 20:39:51 UTC from IEEE Xplore. Restrictions apply.

where $T(n,k)$ for $n \in [0:N]$ and $k \in [0:K]$ is defined as follows:

$$T(n,k) \triangleq \frac{1}{NK\binom{K-1}{k}\binom{N-1}{n}} \sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=k}} \sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n}} \sum_{j\in[1:K]\setminus\mathcal{K}} \times \sum_{i\in[1:N]\setminus\mathcal{N}} I(\mathbf{W}_{[1:K]\setminus\mathcal{K}}; Q_i^{[j]}, A_i^{[j]} | \mathbf{W}_{\mathcal{K}}, \mathbf{Z}_{\mathcal{N}}). \tag{43}$$

We notice that when $n = N$ or $k = K$, then we get all the messages in the conditioning of the mutual information term above, and therefore we get the following boundary conditions on $T(n,k)$:

$$\begin{aligned} &T(n=N,k)=0,\ \forall k\in[0:K],\\ &T(n,k=K)=0,\ \forall n\in[0:N]. \end{aligned} \tag{44}$$

In order to utilize the bound developed in Lemma 2, we further lower bound the function $T(n,k)$ in the following Lemma. This lower bound on $T(n,k)$ has an interesting recursive structure, which in turns allows us to leverage the boundary conditions (44) of the function $T(n,k)$ and thus obtain a closed-form lower bound on the download cost.

Lemma 3: The function $T(n,k)$ is lower bounded as follows:

$$T(n,k) \geq \frac{1}{N-n}\left[\sum_{n'=n}^{N-1} T(n',k+1) + \lambda_{(n,k)}L\right] + o(L), \tag{45}$$

where $\lambda_{(n,k)}$ as defined in (19),

$$\lambda_{(n,k)} \triangleq \frac{1}{KL\binom{K-1}{k}\binom{N}{n}} \sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=k}} \sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n}} \sum_{j\in[1:K]\setminus\mathcal{K}} H(W_j|\mathbf{Z}_{\mathcal{N}}, \mathbf{W}_{\mathcal{K}}), \tag{46}$$

for $n \in [0:N]$ and $k \in [0:K]$.

The proof of Lemma 3 is in Appendix D. Now, we use the recursive lower bound on $T(n,k)$ given in Lemma 3 to further lower bound the average number of downloaded bits D_μ obtained in Lemma 2 as follows:

$$\begin{aligned} D_\mu &\geq L + o(L) + \sum_{n_1=0}^{N-1} T(n_1,1) + o(L)\\ &\geq L + \sum_{n_1=0}^{N-1} \frac{1}{N-n_1}\left(\lambda_{(n_1,1)}L + \sum_{n_2=n_1}^{N-1} T(n_2,2)\right)\\ &\geq L + o(L) \sum_{n_1=0}^{N-1} \frac{\lambda_{(n_1,1)}L}{N-n_1}\\ &\quad + \sum_{n_1=0}^{N-1}\sum_{n_2=n_1}^{N-1} \frac{\left(\lambda_{(n_2,2)}L + o(L)\sum_{n_3=n_2}^{N-1} T(n_3,3)\right)}{(N-n_1)(N-n_2)}\\ &\quad \vdots\\ &\geq L + o(L) + \sum_{n_1=0}^{N-1} \frac{\lambda_{(n_1,1)}L}{N-n_1} + \sum_{n_1=0}^{N-1}\sum_{n_2=n_1}^{N-1} \frac{\lambda_{(n_2,2)}L}{(N-n_1)(N-n_2)}\\ &\quad + \sum_{n_1=0}^{N-1}\sum_{n_2=n_1}^{N-1}\sum_{n_3=n_2}^{N-1} \frac{\lambda_{(n_3,3)}L}{(N-n_1)(N-n_2)(N-n_3)} + \cdots\\ &\quad + \sum_{n_1=0}^{N-1}\sum_{n_2=n_1}^{N-1}\cdots\sum_{n_{K-1}=n_{K-2}}^{N-1} \frac{\lambda_{(n_{K-1},K-1)}L + \sum_{n_K=n_{K-1}}^{N-1} T(n_K,K)}{(N-n_1)\times\cdots\times(N-n_{K-1})}\\ &\overset{(a)}{=} L + o(L) + \sum_{n_1=0}^{N-1} \frac{\lambda_{(n_1,1)}L}{N-n_1} + \sum_{n_1=0}^{N-1}\sum_{n_2=n_1}^{N-1} \frac{\lambda_{(n_2,2)}L}{(N-n_1)(N-n_2)}\\ &\quad + \sum_{n_1=0}^{N-1}\sum_{n_2=n_1}^{N-1}\sum_{n_3=n_2}^{N-1} \frac{\lambda_{(n_3,3)}L}{(N-n_1)(N-n_2)(N-n_3)} + \cdots\\ &\quad + \sum_{n_1=0}^{N-1}\cdots\sum_{n_{K-1}=n_{K-2}}^{N-1} \frac{\lambda_{(n_{K-1},K-1)}L}{(N-n_1)\times\cdots\times(N-n_{K-1})}\\ &\overset{(b)}{=} L + o(L) + \sum_{n_1=1}^{N} \frac{\lambda_{(N-n_1,1)}L}{n_1} + \sum_{n_1=1}^{N}\sum_{n_2=1}^{n_1} \frac{\lambda_{(N-n_2,2)}L}{n_1 n_2}\\ &\quad + \cdots + \sum_{n_1=1}^{N}\cdots\sum_{n_{K-1}=1}^{n_{K-2}} \frac{\lambda_{(N-n_{K-1},K-1)}L}{n_1\times\cdots\times n_{K-1}}\\ &\overset{(c)}{=} L + o(L) + \sum_{n_1=1}^{N} \frac{\lambda_{(N-n_1,1)}L}{n_1} + \sum_{n_1=1}^{N}\sum_{n_2=n_1}^{N} \frac{\lambda_{(N-n_1,2)}L}{n_1 n_2}\\ &\quad + \cdots + \sum_{n_1=1}^{N}\cdots\sum_{n_{K-1}=n_{K-2}}^{N} \frac{\lambda_{(N-n_1,K-1)}L}{n_1\times\cdots\times n_{K-1}}, \end{aligned} \tag{47}$$

where (a) follows by applying the boundary condition on $T(n,k)$ as given in (44), where $T(n,k=K)=0$ for $n \in [0:N-1]$; and (b), (c) follow by changing the summation indexes. Taking the limit $L \to \infty$, we obtain the bound on $\frac{D_\mu}{L}$, which is also a valid bound on the optimal download cost, $D^*(\mu)$, as defined in (7) for $\mu \in [\frac{1}{N}, 1]$, since the bound in (54) is valid for any achievable pair (D_μ, L). Therefore, we obtain the following bound on $D^*(\mu)$,

$$\begin{aligned} D^*(\mu) \geq 1 &+ \sum_{n_1=1}^{N} \frac{\lambda_{(N-n_1,1)}}{n_1} + \sum_{n_1=1}^{N}\sum_{n_2=n_1}^{N} \frac{\lambda_{(N-n_1,2)}}{n_1 n_2}\\ &+ \cdots + \sum_{n_1=1}^{N}\cdots\sum_{n_{K-1}=n_{K-2}}^{N} \frac{\lambda_{(N-n_1,K-1)}}{n_1\times\cdots\times n_{K-1}}, \end{aligned} \tag{48}$$

which completes the proof of Theorem 1.

V. Proof of Theorem 2: Lower Bounds for Uncoded Storage Constrained Databases and General (N,K,μ)

We now specialize the lower bound in (47) for the case of uncoded storage placement as defined in Section II-A. Using Remark 1, the term $\lambda_{(n,k)}$ as defined in (19) can be expressed as,

$$\begin{aligned} \lambda_{(n,k)} &= \frac{1}{KL\binom{K-1}{k}\binom{N}{n}} \sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=k}} \sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n}} \sum_{j\in[1:K]\setminus\mathcal{K}} H(W_j|\mathbf{Z}_{\mathcal{N}}, \mathbf{W}_{\mathcal{K}})\\ &= \frac{1}{K\binom{K-1}{k}\binom{N}{n}} \sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=k}} \sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n}} \sum_{j\in[1:K]\setminus\mathcal{K}} H(W_j|\mathbf{Z}_{\mathcal{N}}) \end{aligned}$$

Authorized licensed use limited to: The University of Arizona. Downloaded on January 30,2021 at 20:39:51 UTC from IEEE Xplore. Restrictions apply.

$$= \frac{1}{K\binom{K-1}{k}\binom{N}{n}} \sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n}} \sum_{j=1}^{K} \sum_{\substack{\mathcal{K}\subseteq[1:K]\setminus\{j\}\\|\mathcal{K}|=k}} H(W_j|\mathbf{Z}_{\mathcal{N}})$$

$$= \frac{1}{K\binom{N}{n}} \sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n}} \sum_{j=1}^{K} H(W_j|\mathbf{Z}_{\mathcal{N}}). \tag{49}$$

We notice that $\lambda_{(n,k)}$ is independent of k, and hence we can define $\lambda_n \triangleq \lambda_{(n,k)}$, for all $k \in [1:K]$. Therefore, we can write the bound in (47) as follows,

$$D_\mu \geq L + o(L) + \sum_{n_1=1}^{N} \left(\frac{1}{n_1} + \sum_{n_2=n_1}^{N} \frac{1}{n_1 n_2} + \cdots + \sum_{n_2=n_1}^{N} \cdots \sum_{n_{K-1}=n_{K-2}}^{N} \frac{1}{n_1 \times \cdots \times n_{K-1}} \right) \lambda_{N-n_1}$$

$$= L + \sum_{n_1=1}^{N} S(n_1, K)\ \lambda_{N-n_1} + o(L), \tag{50}$$

where $S(n,k)$, for $n \in [1:N]$ and $k \in [1:K]$, is defined as follows,

$$S(n,k) \triangleq \frac{1}{n} + \sum_{n_2=n}^{N} \frac{1}{n n_2} + \cdots + \sum_{n_2=n}^{N} \cdots \sum_{n_{k-1}=n_{k-2}}^{N} \frac{1}{n n_2 \times \cdots \times n_{k-1}}. \tag{51}$$

It is important to notice the following boundary conditions and properties of $S(n,k)$:

Property 1: $S(n, k=1) = 0$,

Property 2: $S(n, k=2) = \frac{1}{n}$,

Property 3: $NS(n=N,k) = S(n=N, k-1) + 1$,

Property 4: $nS(n,k) - (n+1)S(n+1,k) = S(n,k-1)$. (52)

The first 3 properties are straight forward to prove from the definition of $S(n,k)$ in (51). The fourth property of $S(n,k)$ provides a useful recursive relation and can be proven as follows:

$$nS(n,k) - (n+1)S(n+1,k)$$

$$\overset{(a)}{=} \left(1 + \sum_{n_2=n}^{N} \frac{1}{n_2} + \sum_{n_2=n}^{N} \sum_{n_3=n_2}^{N} \frac{1}{n_2 n_3} + \cdots + \sum_{n_2=n}^{N} \cdots \sum_{n_{k-1}=n_{k-2}} \frac{1}{n_2 \times \cdots \times n_{k-1}}\right)$$

$$- \left(1 + \sum_{n_2=n+1}^{N} \frac{1}{n_2} + \sum_{n_2=n+1}^{N} \sum_{n_3=n_2}^{N} \frac{1}{n_2 n_3} + \cdots + \sum_{n_2=n+1}^{N} \cdots \sum_{n_{k-1}=n_{k-2}} \frac{1}{n_2 \times \cdots \times n_{k-1}}\right)$$

$$= \frac{1}{n} + \sum_{n_3=n}^{N} \frac{1}{n n_3} + \sum_{n_3=n}^{N} \sum_{n_4=n_3}^{N} \frac{1}{n n_3 n_4} + \cdots + \sum_{n_3=n}^{N} \cdots \sum_{n_{k-1}=n_{k-2}}^{N} \frac{1}{n n_3 \times \cdots \times n_{k-1}}$$

$$\overset{(b)}{=} \frac{1}{n} + \sum_{n_2=n}^{N} \frac{1}{n n_2} + \sum_{n_2=n}^{N} \sum_{n_3=n_2}^{N} \frac{1}{n n_2 n_3} + \cdots + \sum_{n_2=n}^{N} \cdots \sum_{n_{k-2}=n_{k-3}}^{N} \frac{1}{n n_2 \times \cdots \times n_{k-2}}$$

$$\overset{(c)}{=} S(n, k-1), \tag{53}$$

where (a) and (c) follow from the definition of $S(n,k)$ in (51); and (b) follows by relabeling the summation indexes.

Next, we express the λ_n term that appears in (50) in terms of x_ℓ as defined in (15) as follows,

$$\lambda_n = \frac{1}{K\binom{N}{n}} \sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n}} \sum_{k=1}^{K} H(W_k/\mathbf{Z}_{\mathcal{N}})$$

$$= \frac{1}{K\binom{N}{n}} \sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n}} \sum_{k=1}^{K} \sum_{\substack{\mathcal{S}\subseteq[1:N]\setminus\mathcal{N}\\|\mathcal{S}|\geq 1}} |W_{k,\mathcal{S}}|L$$

$$= \frac{1}{K\binom{N}{n}} \sum_{\ell=1}^{N-n} \sum_{k=1}^{K} \sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n}} \sum_{\substack{\mathcal{S}\subseteq[1:N]\setminus\mathcal{N}\\|\mathcal{S}|=\ell}} |W_{k,\mathcal{S}}|L$$

$$= \frac{1}{K\binom{N}{n}} \sum_{\ell=1}^{N-n} \sum_{k=1}^{K} \sum_{\substack{\mathcal{S}\subseteq[1:N]\\|\mathcal{S}|=\ell}} \sum_{\substack{\mathcal{N}\subseteq[1:N]\setminus\mathcal{S}\\|\mathcal{N}|=n}} |W_{k,\mathcal{S}}|L$$

$$= \frac{1}{K\binom{N}{n}} \sum_{\ell=1}^{N-n} \sum_{k=1}^{K} \sum_{\substack{\mathcal{S}\subseteq[1:N]\\|\mathcal{S}|=\ell}} \binom{N-\ell}{n} |W_{k,\mathcal{S}}|L$$

$$= \sum_{\ell=1}^{N-n} \binom{N-n}{\ell} \frac{1}{K\binom{N}{\ell}} \sum_{k=1}^{K} \sum_{\substack{\mathcal{S}\subseteq[1:N]\\|\mathcal{S}|=\ell}} |W_{k,\mathcal{S}}|L$$

$$= \sum_{\ell=1}^{N-n} \binom{N-n}{\ell} x_\ell L. \tag{54}$$

Substituting (54) in (50) and taking the limit $L \to \infty$, we obtain the bound on $\frac{D_\mu}{L}$ in terms of x_ℓ as follows,

$$\frac{D_\mu}{L} \geq 1 + \sum_{n_1=1}^{N} \sum_{\ell=1}^{n_1} \binom{n_1}{\ell} S(n_1, K)\ x_\ell$$

$$= 1 + \sum_{\ell=1}^{N} \sum_{n_1=\ell}^{N} \binom{n_1}{\ell} S(n_1, K)\ x_\ell = 1 + \sum_{\ell=1}^{N} \alpha(\ell, K)\ x_\ell, \tag{55}$$

where $\alpha(\ell,k)$ for $\ell \in [1:N]$ and $k \in [1:K]$ is defined as follows,

$$\alpha(\ell,k) \triangleq \sum_{n=\ell}^{N} \binom{n}{\ell} S(n,k). \tag{56}$$

Since the bound in (55) is valid for any achievable pair (D_μ, L), it is also a valid bound on the optimal download cost,

Authorized licensed use limited to: The University of Arizona. Downloaded on January 30,2021 at 20:39:51 UTC from IEEE Xplore. Restrictions apply.

$D^*(\mu)$, as defined in (7), where $\mu \in [\frac{1}{N}, 1]$. Therefore, we obtain the following bound on $D^*(\mu)$,

$$D^*(\mu) \geq 1 + \sum_{\ell=1}^{N} \alpha(\ell, K)\ x_\ell. \tag{57}$$

Next, we use the properties of $S(n,k)$ in (52) to obtain a recursion relation for $\alpha(\ell,k)$ as introduced in the following Lemma:

Lemma 4: The function $\alpha(\ell,k)$ satisfies the following recursion relation:

$$\alpha(\ell,k) = \frac{1}{\ell}\left[\alpha(\ell,k-1) + \binom{N}{\ell}\right]. \tag{58}$$

Proof:

$$\begin{aligned}
\alpha(\ell,k) &\overset{(a)}{=} \sum_{n=\ell}^{N}\binom{n}{\ell}S(n,k) = \frac{1}{\ell}\sum_{n=\ell}^{N}\binom{n-1}{\ell-1}nS(n,k)\\
&= \frac{1}{\ell}\left[\sum_{n=\ell}^{N}\binom{n-1}{\ell-1}nS(n,k) + \sum_{n=\ell+1}^{N}\binom{n-1}{\ell}nS(n,k)\right.\\
&\qquad\left. - \sum_{n=\ell+1}^{N}\binom{n-1}{\ell}nS(n,k)\right]\\
&= \frac{1}{\ell}\left[\sum_{n=\ell}^{N}\binom{n}{\ell}nS(n,k) - \sum_{n=\ell}^{N-1}\binom{n}{\ell}(n+1)S(n+1,k)\right]\\
&= \frac{1}{\ell}\left[\sum_{n=\ell}^{N-1}\binom{n}{\ell}[nS(n,k) - (n+1)S(n,k)]\right.\\
&\qquad\left. + \binom{N}{\ell}NS(N,k)\right]\\
&\overset{(b)}{=} \frac{1}{\ell}\left[\sum_{n=\ell}^{N-1}\binom{n}{\ell}S(n,k-1) + \binom{N}{\ell}S(N,k-1) + \binom{N}{\ell}\right]\\
&= \frac{1}{\ell}\left[\sum_{n=\ell}^{N}\binom{n}{\ell}S(n,k-1) + \binom{N}{\ell}\right]\\
&\overset{(c)}{=} \frac{1}{\ell}\left[\alpha(\ell,k-1) + \binom{N}{\ell}\right],
\end{aligned} \tag{59}$$

where (a) and (c) follow from the definition of $\alpha(\ell,k)$ in (56); and (b) follows from properties 3 and 4 in (52). ∎

Next, we use the recursion relation for $\alpha(\ell,k)$ given in Lemma 4 to obtain a closed form expression of the coefficients $\alpha(\ell,K)$, for $\ell \in [1:N]$, in terms of the system parameters as follows:

$$\begin{aligned}
\alpha(\ell,K) &= \frac{1}{\ell}\left(\alpha(\ell,K-1) + \binom{N}{\ell}\right)\\
&= \frac{1}{\ell}\binom{N}{\ell} + \frac{1}{\ell^2}\left(\alpha(\ell,K-2) + \binom{N}{\ell}\right)\\
&\ \ \vdots\\
&= \binom{N}{\ell}\left(\frac{1}{\ell} + \frac{1}{\ell^2} + \cdots + \frac{1}{\ell^{K-1}}\right)\\
&= \binom{N}{\ell}\left(\tilde{D}(\ell) - 1\right),
\end{aligned} \tag{60}$$

which follows by applying the boundary condition on $\alpha(\ell,k)$ where $\alpha(\ell,k=1) = 0$, and $\tilde{D}(\ell) = \sum_{k=0}^{K-1}\frac{1}{\ell^k}$ as defined in (25). Therefore, the bound in (57) can be written as

$$D^*(\mu) \geq 1 + \sum_{\ell=1}^{N}\binom{N}{\ell}\left(\tilde{D}(\ell) - 1\right)\ x_\ell. \tag{61}$$

Next, we obtain $N-1$ different lower bounds on $D^*(\mu)$, by eliminating the pairs (x_j, x_{j+1}), for each $j \in [1:N-1]$, in the equation (61) using the message size, and the storage constraints for uncoded storage placement given in (14), and (16), respectively. We use (14) to write x_j as follows:

$$x_j = \frac{1}{\binom{N}{j}}\left(1 - \sum_{\ell\in[1:N]\setminus j}\binom{N}{\ell}x_\ell\right). \tag{62}$$

We first apply (62) in (61) to obtain

$$\begin{aligned}
D^*(\mu) &\geq 1 + \sum_{\ell\in[1:N]\setminus j}\binom{N}{\ell}\left(\tilde{D}(\ell) - 1\right)\ x_\ell\\
&\qquad + \left(1 - \sum_{\ell\in[1:N]\setminus j}\binom{N}{\ell}x_\ell\right)\left(\tilde{D}(j) - 1\right)\\
&= \tilde{D}(j) + \sum_{\ell\in[1:N]\setminus j}\binom{N}{\ell}\left(\tilde{D}(\ell) - \tilde{D}(j)\right)\ x_\ell.
\end{aligned} \tag{63}$$

We next apply (62) in the storage constraint (16) to obtain

$$\begin{aligned}
\mu N &\geq \sum_{\ell\in[1:N]\setminus j}\ell\binom{N}{\ell}x_\ell + j\left(1 - \sum_{\ell\in[1:N]\setminus j}\binom{N}{\ell}x_\ell\right)\\
&= j + \sum_{\ell\in[1:N]\setminus j}\binom{N}{\ell}(\ell - j)x_\ell.
\end{aligned} \tag{64}$$

In order to eliminate x_{j+1} from (63), we first use (64) to bound x_{j+1} as

$$x_{j+1} \leq \frac{1}{\binom{N}{j+1}}\left(\mu N - j - \sum_{\ell\in[1:N]\setminus\{j,j+1\}}\binom{N}{\ell}(\ell - j)x_\ell\right), \tag{65}$$

which can be applied in (63) to obtain the following bound on $D^*(\mu)$,

$$\begin{aligned}
D^*(\mu) &\geq \tilde{D}(j) + \sum_{\ell\in[1:N]\setminus j}\binom{N}{\ell}\left(\tilde{D}(\ell) - \tilde{D}(j)\right)\ x_\ell\\
&\overset{(a)}{\geq} \tilde{D}(j) + \sum_{\ell\in[1:N]\setminus\{j,j+1\}}\binom{N}{\ell}\left(\tilde{D}(\ell) - \tilde{D}(j)\right)x_\ell\\
&\qquad + \left(\tilde{D}(j+1) - \tilde{D}(j)\right)\left(\mu N - j\right.\\
&\qquad\left. - \sum_{\ell\in[1:N]\setminus\{j,j+1\}}\binom{N}{\ell}(\ell - j)x_\ell\right)\\
&\overset{(b)}{=} (\mu N - j)\tilde{D}(j+1) - (\mu N - j - 1)\tilde{D}(j)\\
&\qquad + \sum_{\ell\in[1:N]\setminus\{j,j+1\}}\binom{N}{\ell}\Gamma_\ell^{(j)}\ x_\ell\\
&\overset{(c)}{\geq} (\mu N - j)\tilde{D}(j+1) - (\mu N - j - 1)\tilde{D}(j),
\end{aligned} \tag{66}$$

Authorized licensed use limited to: The University of Arizona. Downloaded on January 30,2021 at 20:39:51 UTC from IEEE Xplore. Restrictions apply.

where (a) follows from (65) where the coefficient $\tilde{D}(j+1)-\tilde{D}(j)$ is negative for all $j \in [1:N-1]$; $\Gamma_\ell^{(j)}$ for $\ell \in [1:N] \setminus \{j, j+1\}$ in (b) is defined as

$$\Gamma_\ell^{(j)} \triangleq \tilde{D}(\ell) + (\ell - j - 1)\tilde{D}(j) - (\ell - j)\tilde{D}(j+1); \quad (67)$$

and (c) follows since $x_\ell \geq 0$ and from the following Lemma.

Lemma 5: $\Gamma_\ell^{(j)}$ is non-negative for $\ell \in [1:N] \setminus \{j, j+1\}$.

The proof of Lemma 5 is in Appendix E. From (66) we arrive at the following lower bound on $D^*(\mu)$:

$$D^*(\mu) \geq (\mu N - j)\tilde{D}(j+1) - (\mu N - j - 1)\tilde{D}(j), \quad (68)$$

which is a linear function of μ for a fixed value of $j \in [1:N-1]$ passing through the two points: $(\mu_1 = \frac{j}{N}, \frac{D}{L} = \tilde{D}(j))$ and $(\mu_2 = \frac{j+1}{N}, \frac{D}{L} = \tilde{D}(j+1))$. We obtain $N-1$ such lower bounds for every $j \in [1:N-1]$, which eventually yield the lower bound on the optimal download cost $D^*(\mu)$ as the lower convex envelope of the following N points for $t \in [1:N]$:

$$\left(\mu = \frac{t}{N},\ D(\mu) = \tilde{D}(t) = 1 + \frac{1}{t} + \frac{1}{t^2} + \cdots + \frac{1}{t^{K-1}}\right), \quad (69)$$

which completes the converse proof of Theorem 2.

VI. Proof of Theorem 2: Achievability for General (N, K, μ)

The achievability proof of the optimal download cost given in Theorem 2 has two main parts: a) the storage design (i.e., what to store across N databases) subject to the storage constraint; and b) the design of the PIR scheme from storage constrained databases. We next describe our placement scheme while satisfying the storage constraint at each database. In particular, we focus on the storage points μKL for $\mu = t/N$ and $t \in [1:N]$. Once we achieve a scheme for these storage points, the lower convex envelope is also achieved using memory sharing as discussed in Claim 1.

A. Storage Placement Scheme for $\mu = t/N$ and $t \in [1:N]$

The storage placement scheme is inspired by the placement strategy proposed in the work on coded caching [32]. For a fixed parameter $t \in [1:N]$, we take each message W_k and sub-divide it into $\binom{N}{t}$ equal sized sub-messages of size $L/\binom{N}{t}$ bits each. We then label each sub-message with a unique subset $\mathcal{S} \subseteq [1:N]$ of size t. Therefore, the message W_k can be expressed as:

$$W_k = \bigcup_{\substack{\mathcal{S} \subseteq [1:N] \\ |\mathcal{S}| = t}} W_{k,\mathcal{S}}. \quad (70)$$

Using this message splitting scheme, we propose the databases storage placement scheme as follows: a sub-message $W_{k,\mathcal{S}}$ is stored in exactly t databases whose labels are in the set $\mathcal{S}$. In other words, for every message, each database stores all the sub-messages which contain its index. Therefore, the storage at DB_n is given as

$$Z_n = \bigcup_{k \in [1:K]} \bigcup_{\substack{\mathcal{S} \subseteq [1:N] \\ |\mathcal{S}| = t,\ n \in \mathcal{S}}} W_{k,\mathcal{S}}. \quad (71)$$

Assume that each sub-message is of size t^K bits. Hence the total size of each message L is given as $L = \binom{N}{t} t^K$. We next verify that the above scheme satisfies the storage constraint. To this end, we note that for every message, each database stores $\binom{N-1}{t-1}$ sub-messages (this corresponds to the number of sub-sets of databases of size t in which the the index of the database is present). Hence, the total storage necessary for any database is given as:

$$\underbrace{K}_{\substack{\text{Total number} \\ \text{of messages}}} \times \underbrace{\binom{N-1}{t-1}}_{\substack{\text{Number of submessages} \\ \text{per message per database}}} \times \underbrace{t^K}_{\substack{\text{Size of each} \\ \text{submessage}}}$$
$$= \frac{t}{N} \times K \times \underbrace{\binom{N}{t} t^K}_{\text{Size of a message}} = \frac{t}{N} \times K \times L = \mu K L.$$

This shows that the proposed scheme satisfies the storage constraints for every database.

B. Storage Constrained PIR Scheme for $\mu = t/N$ and $t \in [1:N]$

We now present the storage constrained PIR scheme for any (N, K) originally introduced in our previous work [31]. We focus on the storage parameter $\mu = t/N$ for any $t \in [1:N]$. We assume a fixed query structure at each database independent of the desired message W_k, where the query set $Q_n^{[k]}$ to each DB_n is structured as follows: The query $Q_n^{[k]}$ is composed of $\binom{N-1}{t-1}$ blocks, where every block is labeled by a set $\mathcal{S} \in [1:N]$ of size t, where $n \in \mathcal{S}$. The query block labeled with $\mathcal{S}$ only involves the sub-messages stored at the databases DB_n where $n \in \mathcal{S}$, i.e., $W_{i,\mathcal{S}}$ for $i \in [1:K]$. Furthermore, these sub-messages are stored in $|\mathcal{S}| = t$ databases labeled with indexes in the set $\mathcal{S}$.

For a query block $\mathcal{S}$ and a desired message W_k, we apply the original PIR scheme proposed in [7] with $N' = t$ databases with labels in the set $\mathcal{S}$, and K sub-messages $W_{1,\mathcal{S}}, \ldots, W_{K,\mathcal{S}}$ of size t^K bits each to privately retrieve the desired sub-message $W_{k,\mathcal{S}}$. The download cost to retrieve $W_{i,\mathcal{S}}$ is given as $D^{\mathcal{S}}(\mu) = 1 + \frac{1}{N'} + \ldots + \frac{1}{N'^{K-1}} = 1 + \frac{1}{t} + \ldots + \frac{1}{t^{K-1}}$. By repeating this procedure for all possible $\binom{N}{t}$ query blocks the user can privately retrieve all the sub-messages of W_k as defined in (70). Therefore, the download cost $D(\mu)$ of the proposed storage constrained PIR scheme when $\mu = t/N$ for $t \in [1:N]$ is given as

$$D(\mu) = \frac{\binom{N}{t} \times \text{Total Downloaded bits (per block } \mathcal{S})}{\binom{N}{t} \times \text{Desired bits (per block } \mathcal{S})}$$
$$= D^{\mathcal{S}}(\mu) = 1 + \frac{1}{t} + \frac{1}{t^2} + \cdots + \frac{1}{t^{K-1}} = \tilde{D}(t). \quad (72)$$

Using the memory sharing concept in Claim 1, we can achieve the lower convex envelope of the following N achievable points for $t \in [1:N]$:

$$\left(\mu = \frac{t}{N},\ D(\mu) = \tilde{D}(t) = 1 + \frac{1}{t} + \frac{1}{t^2} + \cdots + \frac{1}{t^{K-1}}\right), \quad (73)$$

which matches the trade-off given in Theorem 2.

Authorized licensed use limited to: The University of Arizona. Downloaded on January 30,2021 at 20:39:51 UTC from IEEE Xplore. Restrictions apply.

VII. Conclusion

In this article, we characterize the optimal download cost of PIR for uncoded storage constrained databases. In particular, for any (N, K), we show that the optimal trade-off between the storage parameter, $\mu \in [1/N, 1]$, and the download cost, $D(\mu)$, is given by the lower convex hull of the pairs $(\frac{t}{N}, \left(1+\frac{1}{t}+\frac{1}{t^2}+\cdots+\frac{1}{t^{K-1}}\right))$ for $t \in [1:N]$. The main technical contribution of this article is obtaining lower bounds on the download cost for PIR as a function of storage, which match the achievable scheme in [31], and hence characterize the optimal trade-off. We first arrived to a lower bound on the download cost, which is valid for any arbitrary storage at the databases. We then specialized the obtained bound for uncoded placement strategies, which helps in obtaining a linear program subject to message size and storage constraints. Solving this linear program, we arrive at a set of $N-1$ lower bounds, where each bound is tight in a certain range of storage. There are several interesting future directions on this important variation of storage-constrained PIR such as a) settling the tradeoff with coded storage allowed at databases, b) colluding databases and c) introducing additional reliability constraints on storage, such that data must be recoverable from any N out of M databases.

Appendix A
Proof of Claim 1

Claim 1 follows from a simple memory sharing argument. Consider any two storage parameters μ_1, and μ_2, with optimal download costs $D^*(\mu_1)$, and $D^*(\mu_2)$, respectively, then for any storage parameter $\bar{\mu} = \alpha\mu_1 + (1-\alpha)\mu_2$, $\alpha \in [0,1]$, there exists a PIR scheme which achieves a download cost of $\bar{D}(\bar{\mu}) = \alpha D^*(\mu_1) + (1-\alpha)D^*(\mu_2)$. This is done as follows: first, we divide each message W_k into two partitions $W_k = \left\{W_k^{(1)}, W_k^{(2)}\right\}$, where $W_k^{(1)}$ and $W_k^{(2)}$ are of size αL and $(1-\alpha)L$, respectively. Likewise, the storage of each database Z_n is divided into two partitions $Z_n = \left\{Z_n^{(1)}, Z_n^{(2)}\right\}$, where $Z_n^{(1)}$ and $Z_n^{(2)}$ are of size $\alpha\mu_1 KL$ and $(1-\alpha)\mu_2 KL$, respectively. Now, for messages partitions denoted by $W_k^{(1)}$ for $k \in [1:K]$ and databases partitions denoted by $Z_n^{(1)}$ for $n \in [1:N]$, we can apply the PIR scheme which achieves a download cost of $\alpha D^*(\mu_1)$, while for messages partitions denoted by $W_k^{(2)}$ for $k \in [1:K]$ and databases partitions denoted by $Z_n^{(2)}$ for $n \in [1:N]$, we can achieve a download cost of $(1-\alpha)D^*(\mu_2)$, which gives a total download cost of $\bar{D}(\bar{\mu}) = \alpha D^*(\mu_1) + (1-\alpha)D^*(\mu_2)$. Since $D^*(\alpha\mu_1 + (1-\alpha)\mu_2)$ by definition is optimal download cost for the storage parameter $\bar{\mu}$, it cannot be larger than the download cost of the memory sharing scheme, i.e.,

$$D^*(\alpha\mu_1 + (1-\alpha)\mu_2) \leq \bar{D}(\bar{\mu}) = \alpha D^*(\mu_1) + (1-\alpha)D^*(\mu_2), \tag{74}$$

which completes the proof of Claim 1.

Appendix B
Proof of Lemma 1

For any $\mathcal{N} \subseteq [1:N]$, $\mathcal{K} \subseteq [1:K]$, $i \in [1:N]$, and $j \in [1:K]$, we can bound the mutual information term $I(\mathbf{W}_{[1:K]\setminus\mathcal{K}}; Q_i^{[j]}, A_i^{[j]} | \mathbf{W}_{\mathcal{K}}, \mathbf{Q}_{\mathcal{N}}^{[j]}, \mathbf{A}_{\mathcal{N}}^{[j]})$ as follows:

$$\begin{aligned}
&I(\mathbf{W}_{[1:K]\setminus\mathcal{K}}; Q_i^{[j]}, A_i^{[j]} | \mathbf{W}_{\mathcal{K}}, \mathbf{Q}_{\mathcal{N}}^{[j]}, \mathbf{A}_{\mathcal{N}}^{[j]}) \\
&\quad = H(Q_i^{[j]}, A_i^{[j]} | \mathbf{W}_{\mathcal{K}}, \mathbf{Q}_{\mathcal{N}}^{[j]}, \mathbf{A}_{\mathcal{N}}^{[j]}) \\
&\qquad\qquad - H(Q_i^{[j]}, A_i^{[j]} | \mathbf{W}_{[1:K]}, \mathbf{Q}_{\mathcal{N}}^{[j]}, \mathbf{A}_{\mathcal{N}}^{[j]}) \\
&\quad \overset{(a)}{=} H(Q_i^{[j]}, A_i^{[j]} | \mathbf{W}_{\mathcal{K}}, \mathbf{Q}_{\mathcal{N}}^{[j]}, \mathbf{A}_{\mathcal{N}}^{[j]}) \\
&\qquad\qquad - H(Q_i^{[j]}, A_i^{[j]} | \mathbf{W}_{[1:K]}, \mathbf{Q}_{\mathcal{N}}^{[j]}, \mathbf{A}_{\mathcal{N}}^{[j]}, \mathbf{Z}_{\mathcal{N}}) \\
&\quad \overset{(b)}{\geq} H(Q_i^{[j]}, A_i^{[j]} | \mathbf{W}_{\mathcal{K}}, \mathbf{Q}_{\mathcal{N}}^{[j]}, \mathbf{A}_{\mathcal{N}}^{[j]}, \mathbf{Z}_{\mathcal{N}}) \\
&\qquad\qquad - H(Q_i^{[j]}, A_i^{[j]} | \mathbf{W}_{[1:K]}, \mathbf{Q}_{\mathcal{N}}^{[j]}, \mathbf{A}_{\mathcal{N}}^{[j]}, \mathbf{Z}_{\mathcal{N}}) \\
&\quad \overset{(c)}{=} H(Q_i^{[j]}, A_i^{[j]} | \mathbf{W}_{\mathcal{K}}, \mathbf{Q}_{\mathcal{N}}^{[j]}, \mathbf{Z}_{\mathcal{N}}) \\
&\qquad\qquad - H(Q_i^{[j]}, A_i^{[j]} | \mathbf{W}_{[1:K]}, \mathbf{Q}_{\mathcal{N}}^{[j]}, \mathbf{Z}_{\mathcal{N}}) \\
&\quad = I(Q_i^{[j]}, A_i^{[j]}; \mathbf{W}_{[1:K]\setminus\mathcal{K}} | \mathbf{W}_{\mathcal{K}}, \mathbf{Q}_{\mathcal{N}}^{[j]}, \mathbf{Z}_{\mathcal{N}}) \\
&\quad \overset{(d)}{=} I(\mathbf{Q}_{\mathcal{N}}^{[j]}, Q_i^{[j]}, A_i^{[j]}; \mathbf{W}_{[1:K]\setminus\mathcal{K}} | \mathbf{W}_{\mathcal{K}}, \mathbf{Z}_{\mathcal{N}}) \\
&\quad \geq I(\mathbf{W}_{[1:K]\setminus\mathcal{K}}; Q_i^{[j]}, A_i^{[j]} | \mathbf{W}_{\mathcal{K}}, \mathbf{Z}_{\mathcal{N}}),
\end{aligned} \tag{75}$$

where (a) follows from the fact that the random variables $\mathbf{Z}_{\mathcal{N}}$ are functions of all the messages $\mathbf{W}_{[1:K]}$; (b) follows since conditioning reduces entropy; (c) follows since the answers $\mathbf{A}_{\mathcal{N}}^{[j]}$ are functions of the storage random variables $\mathbf{Z}_{\mathcal{N}}$ and the queries $\mathbf{Q}_{\mathcal{N}}^{[j]}$; and (d) follows from fact that queries $\mathbf{Q}_{\mathcal{N}}^{[j]}$ are independent from the messages stored at the databases.

Appendix C
Proof of Lemma 2

We start by obtaining the following bound for all $k \in [1:K]$:

$$\begin{aligned}
&I(\mathbf{W}_{[1:K]\setminus k}; \mathbf{Q}_{[1:N]}^{[k]}, \mathbf{A}_{[1:N]}^{[k]} | W_k) \\
&\quad = I(\mathbf{W}_{[1:K]\setminus k}; \mathbf{Q}_{[1:N]}^{[k]}, \mathbf{A}_{[1:N]}^{[k]}, W_k) - I(\mathbf{W}_{[1:K]\setminus k}; W_k) \\
&\quad \overset{(a)}{=} I(\mathbf{W}_{[1:K]\setminus k}; \mathbf{Q}_{[1:N]}^{[k]}) + I(\mathbf{W}_{[1:K]\setminus k}; \mathbf{A}_{[1:N]}^{[k]} | \mathbf{Q}_{[1:N]}^{[k]}) \\
&\qquad\qquad + I(\mathbf{W}_{[1:K]\setminus k}; W_k | \mathbf{Q}_{[1:N]}^{[k]}, \mathbf{A}_{[1:N]}^{[k]}) \\
&\quad \overset{(b)}{=} H(\mathbf{A}_{[1:N]}^{[k]} | \mathbf{Q}_{[1:N]}^{[k]}) - H(\mathbf{A}_{[1:N]}^{[k]} | \mathbf{W}_{[1:K]\setminus k}, \mathbf{Q}_{[1:N]}^{[k]}) \\
&\qquad\qquad + I(\mathbf{W}_{[1:K]\setminus k}; W_k | \mathbf{Q}_{[1:N]}^{[k]}, \mathbf{A}_{[1:N]}^{[k]}) \\
&\quad = H(\mathbf{A}_{[1:N]}^{[k]} | \mathbf{Q}_{[1:N]}^{[k]}) - H(\mathbf{A}_{[1:N]}^{[k]}, W_k | \mathbf{W}_{[1:K]\setminus k}, \mathbf{Q}_{[1:N]}^{[k]}) \\
&\qquad\qquad + H(W_k | \mathbf{W}_{[1:K]\setminus k}, \mathbf{Q}_{[1:N]}^{[k]}, \mathbf{A}_{[1:N]}^{[k]}) \\
&\qquad\qquad + I(\mathbf{W}_{[1:K]\setminus k}; W_k | \mathbf{Q}_{[1:N]}^{[k]}, \mathbf{A}_{[1:N]}^{[k]}) \\
&\quad \leq D_\mu^{[k]} - H(W_k | \mathbf{W}_{[1:K]\setminus k}, \mathbf{Q}_{[1:N]}^{[k]}) \\
&\qquad - H(\mathbf{A}_{[1:N]}^{[k]} | \mathbf{W}_{[1:K]}, \mathbf{Q}_{[1:N]}^{[k]}) + H(W_k | \mathbf{Q}_{[1:N]}^{[k]}, \mathbf{A}_{[1:N]}^{[k]}) \\
&\quad \overset{(c)}{=} D_\mu^{[k]} - L + o(L),
\end{aligned} \tag{76}$$

where (a) follows from the chain rule of mutual information and from the fact that the messages are i.i.d., (b) follows from (3) where queries are not functions of the messages; (c) follows from (4) where answers are functions of the messages and the corresponding queries and also from the

Authorized licensed use limited to: The University of Arizona. Downloaded on January 30,2021 at 20:39:51 UTC from IEEE Xplore. Restrictions apply.

decodability constraint in (5), where W_k is decodable from $\mathbf{Q}^{[k]}_{[1:N]}$ and $\mathbf{A}^{[k]}_{[1:N]}$. Summing up the obtained bound in (76) over $k \in [1:K]$, we get the following bound over the average number of download bits D_μ:

$$
\begin{aligned}
&D_\mu - L + o(L)\\
&\geq \frac{1}{K}\sum_{k=1}^{K} I(\mathbf{W}_{[1:K]\backslash k}; \mathbf{Q}^{[k]}_{[1:N]}, \mathbf{A}^{[k]}_{[1:N]}|W_k)\\
&= \frac{1}{K}\frac{1}{N!}\sum_{k=1}^{K}\sum_{\boldsymbol{\sigma}\in[N!]} I(\mathbf{W}_{[1:K]\backslash k}; \mathbf{Q}^{[k]}_{\boldsymbol{\sigma}_{[1:N]}}, \mathbf{A}^{[k]}_{\boldsymbol{\sigma}_{[1:N]}}|W_k)\\
&\overset{(a)}{=} \sum_{n=1}^{N}\sum_{k=1}^{K}\sum_{\boldsymbol{\sigma}\in[N!]} \frac{I(\mathbf{W}_{[1:K]\backslash k}; Q^{[k]}_{\sigma_n}, A^{[k]}_{\sigma_n}|\mathbf{W}_k, \mathbf{Q}^{[k]}_{\boldsymbol{\sigma}_{[1:n-1]}}, \mathbf{A}^{[k]}_{\boldsymbol{\sigma}_{[1:n-1]}})}{K\times N!}\\
&\overset{(b)}{\geq} \sum_{n=1}^{N}\sum_{k=1}^{K}\sum_{\boldsymbol{\sigma}\in[N!]} \frac{I(\mathbf{W}_{[1:K]\backslash k}; Q^{[k]}_{\sigma_n}, A^{[k]}_{\sigma_n}|\mathbf{W}_k, \mathbf{Z}_{\boldsymbol{\sigma}_{[1:n-1]}})}{K\times N!}\\
&\overset{(c)}{=} \frac{1}{K(K-1)}\frac{1}{N!}\sum_{n=1}^{N}\sum_{k=1}^{K}\sum_{j\in[1:K]\backslash k}\\
&\qquad\times\sum_{\boldsymbol{\sigma}\in[N!]} I(\mathbf{W}_{[1:K]\backslash k}; Q^{[j]}_{\sigma_n}, A^{[j]}_{\sigma_n}|\mathbf{W}_k, \mathbf{Z}_{\boldsymbol{\sigma}_{[1:n-1]}})\\
&= \frac{1}{K\binom{K-1}{1}}\frac{1}{N!}\sum_{n=1}^{N}\sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=1}}\sum_{j\in[1:K]\backslash\mathcal{K}}\\
&\qquad\times\sum_{\boldsymbol{\sigma}\in[N!]} I(\mathbf{W}_{[1:K]\backslash\mathcal{K}}; Q^{[j]}_{\sigma_n}, A^{[j]}_{\sigma_n}|\mathbf{W}_\mathcal{K}, \mathbf{Z}_{\boldsymbol{\sigma}_{[1:n-1]}})\\
&\overset{(d)}{=} \sum_{n=1}^{N}\sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=1}}\sum_{j\in[1:K]\backslash\mathcal{K}} \frac{(N-n)!(n-1)!}{K\binom{K-1}{1}N!}\\
&\qquad\times\sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n-1}}\sum_{i\in[1:N]\backslash\mathcal{N}} I(\mathbf{W}_{[1:K]\backslash\mathcal{K}}; Q^{[j]}_i, A^{[j]}_i|\mathbf{W}_\mathcal{K}, \mathbf{Z}_\mathcal{N})\\
&= \sum_{n=0}^{N-1}\frac{1}{KN\binom{N-1}{n}\binom{K-1}{1}}\sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=1}}\sum_{j\in[1:K]\backslash\mathcal{K}}\sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n}}\\
&\qquad\times\sum_{i\in[1:N]\backslash\mathcal{N}} I(\mathbf{W}_{[1:K]\backslash\mathcal{K}}; Q^{[j]}_i, A^{[j]}_i|\mathbf{W}_\mathcal{K}, \mathbf{Z}_\mathcal{N}) + o(L)\\
&\overset{(e)}{=} \sum_{n=0}^{N-1} T(n,1),
\end{aligned}
\tag{77}
$$

where (a) follows from chain rule of mutual information; (b) follows from Lemma 1; (c) follows from the privacy constraint in (6) where the individual queries and answers are invariant with respect to the requested message index; (d) follows from the symmetry with respect to the summation indexes, where for every set $\mathcal{N}\subseteq[1:N]$ of size $(n-1)$ and every $i\in[1:N]\backslash\mathcal{N}$, the number of permutations σ that lead to the mutual information $I(\mathbf{W}_{[1:K]\backslash\mathcal{K}}; Q^{[j]}_i, A^{[j]}_i|\mathbf{W}_\mathcal{K}, \mathbf{Z}_\mathcal{N})$ is $(N-n)!(n-1)!$; and (e) follows from the definition of $T(n,k)$ in (43).

APPENDIX D
PROOF OF LEMMA 3

We start by bounding $T(n,k)$ defined in (43) as follows,

$$
\begin{aligned}
T(n,k) &= \frac{1}{NK\binom{K-1}{k}\binom{N-1}{n}}\sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=k}}\sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n}}\sum_{j\in[1:K]\backslash\mathcal{K}}\\
&\qquad\times\sum_{i\in[1:N]\backslash\mathcal{N}} I(\mathbf{W}_{[1:K]\backslash\mathcal{K}}; Q^{[j]}_i, A^{[j]}_i|\mathbf{W}_\mathcal{K}, \mathbf{Z}_\mathcal{N})\\
&\overset{(a)}{=} \frac{1}{NK\binom{K-1}{k}\binom{N-1}{n}}\sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=k}}\sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n}}\sum_{j\in[1:K]\backslash\mathcal{K}}\\
&\qquad\times\sum_{i\in[1:N]\backslash\mathcal{N}} I(\mathbf{W}_{[1:K]\backslash\mathcal{K}}; A^{[j]}_i|\mathbf{W}_\mathcal{K}, \mathbf{Z}_\mathcal{N}, Q^{[j]}_i)\\
&\overset{(b)}{=} \frac{1}{NK\binom{K-1}{k}\binom{N-1}{n}}\sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=k}}\sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n}}\sum_{j\in[1:K]\backslash\mathcal{K}}\\
&\qquad\times\sum_{i\in[1:N]\backslash\mathcal{N}} H(A^{[j]}_i|\mathbf{W}_\mathcal{K}, \mathbf{Z}_\mathcal{N}, Q^{[j]}_i)\\
&\overset{(c)}{\geq} \frac{1}{NK\binom{K-1}{k}\binom{N-1}{n}}\sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=k}}\sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n}}\sum_{j\in[1:K]\backslash\mathcal{K}}\\
&\qquad\times\sum_{i\in[1:N]\backslash\mathcal{N}} H(A^{[j]}_i|\mathbf{W}_\mathcal{K}, \mathbf{Z}_\mathcal{N}, \mathbf{Q}^{[j]}_{[1:N]})\\
&\geq \frac{1}{NK\binom{K-1}{k}\binom{N-1}{n}}\sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=k}}\sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n}}\\
&\qquad\times\sum_{j\in[1:K]\backslash\mathcal{K}} H(\mathbf{A}^{[j]}_{[1:N]\backslash\mathcal{N}}|\mathbf{W}_\mathcal{K}, \mathbf{Z}_\mathcal{N}, \mathbf{Q}^{[j]}_{[1:N]})\\
&\overset{(d)}{=} \frac{1}{NK\binom{K-1}{k}\binom{N-1}{n}}\sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=k}}\sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n}}\\
&\qquad\times\sum_{j\in[1:K]\backslash\mathcal{K}} I(\mathbf{W}_{[1:K]\backslash\mathcal{K}}; \mathbf{A}^{[j]}_{[1:N]\backslash\mathcal{N}}|\mathbf{W}_\mathcal{K}, \mathbf{Z}_\mathcal{N}, \mathbf{Q}^{[j]}_{[1:N]})\\
&\overset{(e)}{=} \frac{1}{NK\binom{K-1}{k}\binom{N-1}{n}}\sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=k}}\sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n}}\\
&\qquad\times\sum_{j\in[1:K]\backslash\mathcal{K}} I(\mathbf{W}_{[1:K]\backslash\mathcal{K}}; \mathbf{Q}^{[j]}_{[1:N]}, \mathbf{A}^{[j]}_{[1:N]\backslash\mathcal{N}}|\mathbf{W}_\mathcal{K}, \mathbf{Z}_\mathcal{N})\\
&\overset{(f)}{\geq} o(L) + \frac{1}{NK\binom{K-1}{k}\binom{N-1}{n}}\sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=k}}\sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n}}\\
&\quad\times\sum_{j\in[1:K]\backslash\mathcal{K}} I(\mathbf{W}_{[1:K]\backslash\mathcal{K}}; W_j, \mathbf{Q}^{[j]}_{[1:N]}, \mathbf{A}^{[j]}_{[1:N]\backslash\mathcal{N}}|\mathbf{W}_\mathcal{K}, \mathbf{Z}_\mathcal{N})\\
&= \frac{1}{NK\binom{K-1}{k}\binom{N-1}{n}}\sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=k}}\sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n}}\\
&\qquad\times\sum_{j\in[1:K]\backslash\mathcal{K}} I(\mathbf{W}_{[1:K]\backslash\mathcal{K}}; \mathbf{Q}^{[j]}_{[1:N]}, \mathbf{A}^{[j]}_{[1:N]\backslash\mathcal{N}}|\mathbf{W}_\mathcal{K}, \mathbf{Z}_\mathcal{N}, W_j)\\
&\quad+\frac{1}{NK\binom{K-1}{k}\binom{N-1}{n}}\sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=k}}\sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n}}\\
&\qquad\times\sum_{j\in[1:K]\backslash\mathcal{K}} H(W_j|\mathbf{W}_\mathcal{K}, \mathbf{Z}_\mathcal{N}) + o(L)
\end{aligned}
$$

Authorized licensed use limited to: The University of Arizona. Downloaded on January 30,2021 at 20:39:51 UTC from IEEE Xplore. Restrictions apply.

$$\begin{aligned}
&\overset{(g)}{=} \frac{1}{NK\binom{K-1}{k}\binom{N-1}{n}} \sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=k}} \sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n}} \\
&\times \sum_{j\in[1:K]\setminus\mathcal{K}} I(\mathbf{W}_{[1:K]\setminus(\mathcal{K}\cup j)}; \mathbf{Q}^{[j]}_{[1:N]\setminus\mathcal{N}}, \mathbf{A}^{[j]}_{[1:N]\setminus\mathcal{N}} | \mathbf{W}_{(\mathcal{K}\cup j)}, \mathbf{Z}_{\mathcal{N}}) \\
&\quad + \frac{1}{N-n}\lambda_{(n,k)}L + o(L) \\
&= \frac{1}{NK\binom{K-1}{k}\binom{N-1}{n}} \sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=k+1}} \sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n}} \\
&\quad \times \sum_{j\in\mathcal{K}} I(\mathbf{W}_{[1:K]\setminus\mathcal{K}}; \mathbf{Q}^{[j]}_{[1:N]\setminus\mathcal{N}}, \mathbf{A}^{[j]}_{[1:N]\setminus\mathcal{N}} | \mathbf{W}_{\mathcal{K}}, \mathbf{Z}_{\mathcal{N}}) \\
&\quad + \frac{1}{N-n}\lambda_{(n,k)}L + o(L) \\
&= \tilde{T}(n,k) + \frac{1}{N-n}\lambda_{(n,k)}L + o(L), \qquad (78)
\end{aligned}$$

where (a) and (e) follow from the fact that queries are independent from the messages; (b) and (d) follow from the fact that answers are functions of all the messages; (c) is because conditioning reduces entropy; (f) follows from the decoding constraint in (5) where W_j is decodable from $\mathbf{Q}^{[j]}_{[1:N]}$, $\mathbf{A}^{[j]}_{[1:N]\setminus\mathcal{N}}$ and $\mathbf{Z}_{\mathcal{N}}$; and (g) follows from the definition of $\lambda_{(n,k)}$ in (19) and since the queries $\mathbf{Q}^{[j]}_{\mathcal{N}}$ are independent from the messages. We further lower bound $T(n,k)$ by bounding the term $\tilde{T}(n,k)$ in (78) as follows,

$$\begin{aligned}
&\tilde{T}(n,k) \\
&= \frac{1}{NK\binom{K-1}{k}\binom{N-1}{n}} \sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=k+1}} \sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n}} \\
&\quad \times \sum_{j\in\mathcal{K}} I(\mathbf{W}_{[1:K]\setminus\mathcal{K}}; \mathbf{Q}^{[j]}_{[1:N]\setminus\mathcal{N}}, \mathbf{A}^{[j]}_{[1:N]\setminus\mathcal{N}} | \mathbf{W}_{\mathcal{K}}, \mathbf{Z}_{\mathcal{N}}) \\
&\overset{(a)}{=} \frac{1}{NK\binom{K-1}{k}\binom{N-1}{n}} \sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=k+1}} \sum_{j\in\mathcal{K}} \frac{1}{n!(N-n)!} \\
&\quad \times \sum_{\sigma\in[N!]} I(\mathbf{W}_{[1:K]\setminus\mathcal{K}}; \mathbf{Q}^{[j]}_{\boldsymbol{\sigma}_{[n+1:N]}}, \mathbf{A}^{[j]}_{\boldsymbol{\sigma}_{[n+1:N]}} | \mathbf{W}_{\mathcal{K}}, \mathbf{Z}_{\boldsymbol{\sigma}_{[1:n]}}) \\
&\overset{(b)}{=} \frac{1}{N!K\binom{K-1}{k}(N-n)} \sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=k+1}} \sum_{j\in\mathcal{K}} \sum_{\sigma\in[N!]} \\
&\quad \times \sum_{n'=n}^{N-1} I(\mathbf{W}_{[1:K]\setminus\mathcal{K}}; Q^{[j]}_{\sigma_{n'+1}}, A^{[j]}_{\sigma_{n'+1}} | \mathbf{W}_{\mathcal{K}}, \mathbf{Z}_{\boldsymbol{\sigma}_{[1:n]}}, \\
&\qquad\qquad \mathbf{Q}^{[j]}_{\boldsymbol{\sigma}_{[n+1:n']}}, \mathbf{A}^{[j]}_{\boldsymbol{\sigma}_{[n+1:n']}}) \\
&\overset{(c)}{\geq} \frac{1}{N!K\binom{K-1}{k}(N-n)} \sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=k+1}} \sum_{j\in\mathcal{K}} \sum_{\sigma\in[N!]} \\
&\quad \times \sum_{n'=n}^{N-1} I(\mathbf{W}_{[1:K]\setminus\mathcal{K}}; Q^{[j]}_{\sigma_{n'+1}}, A^{[j]}_{\sigma_{n'+1}} | \mathbf{W}_{\mathcal{K}}, \mathbf{Z}_{\boldsymbol{\sigma}_{[1:n']}}) \\
&\overset{(d)}{=} \sum_{n'=n}^{N-1} \frac{1}{N!K\binom{K-1}{k}(N-n)} \sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=k+1}} \sum_{j\in\mathcal{K}} n'!(N-n'-1)! \\
&\quad \times \sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n'}} \sum_{i\in[1:N]\setminus\mathcal{N}} I(\mathbf{W}_{[1:K]\setminus\mathcal{K}}; Q^{[j]}_i, A^{[j]}_i | \mathbf{W}_{\mathcal{K}}, \mathbf{Z}_{\mathcal{N}}) \\
&\overset{(e)}{=} \frac{1}{N-n} \sum_{n'=n}^{N-1} \frac{1}{NK\binom{K-1}{k}\binom{N-1}{n'}} \sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n'}} \sum_{i\in[1:N]\setminus\mathcal{N}} \sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=k+1}} \\
&\quad \times \sum_{j\in\mathcal{K}} \sum_{j'\in[1:K]\setminus\mathcal{K}} \frac{I(\mathbf{W}_{[1:K]\setminus\mathcal{K}}; Q^{[j']}_i, A^{[j']}_i | \mathbf{W}_{\mathcal{K}}, \mathbf{Z}_{\mathcal{N}})}{K-k-1} \\
&= \frac{1}{N-n} \sum_{n'=n}^{N-1} \frac{1}{NK\binom{K-1}{k+1}\binom{N-1}{n'}} \sum_{\substack{\mathcal{N}\subseteq[1:N]\\|\mathcal{N}|=n'}} \sum_{i\in[1:N]\setminus\mathcal{N}} \sum_{\substack{\mathcal{K}\subseteq[1:K]\\|\mathcal{K}|=k+1}} \\
&\quad \times \sum_{j'\in[1:K]\setminus\mathcal{K}} I(\mathbf{W}_{[1:K]\setminus\mathcal{K}}; Q^{[j']}_i, A^{[j']}_i | \mathbf{W}_{\mathcal{K}}, \mathbf{Z}_{\mathcal{N}}) \\
&\overset{(f)}{=} \frac{1}{N-n} \sum_{n'=n}^{N-1} T(n', k+1), \qquad (79)
\end{aligned}$$

where (a) and (d) follow from a similar argument to step (d) in (77); (b) follows by applying the chain rule of entropy; (c) follows by applying Lemma 1; (e) follows from the privacy constraint in (6) where the individual queries and answers are invariant with respect to the requested message index; and (f) follows from the definition of $T(n,k)$ in (43). By applying the lower bound on the term $\tilde{T}(n,k)$ in (79) to the lower bound on $T(n,k)$ in (78) we conclude the proof of Lemma 3.

Appendix E
Proof of Lemma 5

In order to prove that $\Gamma^{(j)}_{\ell}$ is non-negative for $\ell \in [1:N]\setminus\{j, j+1\}$, we first need to prove an important property for $\tilde{D}(\ell)$ in the following Lemma:

Lemma 6: $\tilde{D}(\ell) - \tilde{D}(\ell+1)$ is non increasing with respect to ℓ, i.e., $\tilde{D}(\ell') - \tilde{D}(\ell'+1) \geq \tilde{D}(\ell) - \tilde{D}(\ell+1)$, for any $\ell' \leq \ell$.

Proof: In order to prove Lemma 6, it is sufficient to prove that $\tilde{D}(\ell') - \tilde{D}(\ell'+1) \geq \tilde{D}(\ell) - \tilde{D}(\ell+1)$ for $\ell' = \ell - 1$, or $\tilde{D}(\ell-1) - 2\tilde{D}(\ell) + \tilde{D}(\ell+1) \geq 0$. The proof for any $\ell' \leq \ell$ follows by induction.

$$\begin{aligned}
&\tilde{D}(\ell-1) - 2\tilde{D}(\ell) + \tilde{D}(\ell+1) \\
&\qquad = \sum_{j=0}^{K-1} \frac{1}{(\ell-1)^j} + \frac{1}{(\ell+1)^j} - \frac{2}{(\ell)^j} \\
&\qquad \overset{(a)}{\geq} 2\sum_{j=0}^{K-1} \frac{1}{(\ell^2-1)^{j/2}} - \frac{1}{(\ell)^j} \\
&\qquad = 2\sum_{j=0}^{K-1} \frac{(\ell^2)^{j/2} - (\ell^2-1)^{j/2}}{(\ell)^j(\ell^2-1)^{j/2}} \geq 0, \qquad (80)
\end{aligned}$$

where, (a) follows from the AM-GM inequality, i.e., arithmetic mean is larger than geometric mean, that is $\frac{x_1+x_2}{2} \geq (x_1x_2)^{1/2}$, $\forall x_1, x_2 \geq 0$. Therefore, we obtain $\tilde{D}(\ell-1) - \tilde{D}(\ell) \geq \tilde{D}(\ell) - \tilde{D}(\ell+1)$, which completes the proof of the Lemma. ∎

Authorized licensed use limited to: The University of Arizona. Downloaded on January 30,2021 at 20:39:51 UTC from IEEE Xplore. Restrictions apply.

• Case $\ell < j$: We prove that $\Gamma_\ell^{(j)} \geq 0$ for $\ell < j$ as follows,

$$\begin{aligned}\Gamma_\ell^{(j)} &= \tilde{D}(\ell) + (\ell - j - 1)\tilde{D}(j) - (\ell - j)\tilde{D}(j+1) \\ &= \left[\tilde{D}(\ell) - \tilde{D}(j)\right] - (j - \ell)\left[\tilde{D}(j) - \tilde{D}(j+1)\right] \\ &= \sum_{i=\ell}^{j-1}\left[\tilde{D}(i) - \tilde{D}(i+1)\right] - (j - \ell)\left[\tilde{D}(j) - \tilde{D}(j+1)\right] \\ &= \sum_{i=\ell}^{j-1}\left(\left[\tilde{D}(i) - \tilde{D}(i+1)\right] - \left[\tilde{D}(j) - \tilde{D}(j+1)\right]\right) \\ &\overset{(a)}{\geq} \sum_{i=\ell}^{j-1}\left(\left[\tilde{D}(j) - \tilde{D}(j+1)\right] - \left[\tilde{D}(j) - \tilde{D}(j+1)\right]\right) = 0,\end{aligned} \tag{81}$$

where (a) follows from Lemma 6.

• Case $\ell > j+1$: Similar to the case $\ell < j$, we prove $\Gamma_\ell^{(j)} \geq 0$ for $\ell > j+1$ as follows,

$$\begin{aligned}\Gamma_\ell^{(j)} &= \tilde{D}(\ell) + (\ell - j - 1)\tilde{D}(j) - (\ell - j)\tilde{D}(j+1) \\ &= (\ell - j - 1)\left[\tilde{D}(j) - \tilde{D}(j+1)\right] - \left[\tilde{D}(j+1) - \tilde{D}(\ell)\right] \\ &= \sum_{i=j+2}^{\ell}\left(\left[\tilde{D}(j) - \tilde{D}(j+1)\right] - \left[\tilde{D}(i-1) - \tilde{D}(i)\right]\right) \\ &\overset{(a)}{\geq} \sum_{i=j+2}^{\ell}\left(\left[\tilde{D}(j) - \tilde{D}(j+1)\right] - \left[\tilde{D}(j) - \tilde{D}(j+1)\right]\right) = 0,\end{aligned} \tag{82}$$

where (a) follows from Lemma 6.

REFERENCES

[1] B. Chor, O. Goldreich, E. Kushilevitz, and M. Sudan, "Private information retrieval," in *Proc. 36th Annu. Symp. Found. Comput. Sci.*, Oct. 1995, pp. 41–50.

[2] W. Gasarch, "A survey on private information retrieval," in *Proc. Bull. Eur. Assoc. Theor. Comput. Sci. (EATCS)*, 2004, p. 113.

[3] S. Yekhanin, "Locally decodable codes," *Found. Trends Theor. Comput. Sci.*, vol. 6, no. 3, pp. 139–255, 2012.

[4] D. Xiaoding Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in *Proc. IEEE Symp. Secur. Privacy (S&P)*, May 2000, pp. 44–55.

[5] R. Ostrovsky and W. E. Skeith, III, "A survey of single-database private information retrieval: Techniques and applications," in *Proc. Int. Workshop Public Key Cryptogr.* Cham, Switzerland: Springer, 2007, pp. 393–411.

[6] T. H. Chan, S.-W. Ho, and H. Yamamoto, "Private information retrieval for coded storage," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2015, pp. 2842–2846.

[7] H. Sun and S. A. Jafar, "The capacity of private information retrieval," *IEEE Trans. Inf. Theory*, vol. 63, no. 7, pp. 4075–4088, Jul. 2017.

[8] N. B. Shah, K. V. Rashmi, and K. Ramchandran, "One extra bit of download ensures perfectly private information retrieval," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2014, pp. 856–860.

[9] H. Sun and S. A. Jafar, "The capacity of robust private information retrieval with colluding databases," *IEEE Trans. Inf. Theory*, vol. 64, no. 4, pp. 2361–2370, Apr. 2018.

[10] R. Tajeddine, O. W. Gnilke, and S. El Rouayheb, "Private information retrieval from MDS coded data in distributed storage systems," *IEEE Trans. Inf. Theory*, vol. 64, no. 11, pp. 7081–7093, Nov. 2018.

[11] K. Banawan and S. Ulukus, "The capacity of private information retrieval from coded databases," *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 1945–1956, Mar. 2018.

[12] R. Freij-Hollanti, O. W. Gnilke, C. Hollanti, and D. A. Karpuk, "Private information retrieval from coded databases with colluding servers," *SIAM J. Appl. Algebra Geometry*, vol. 1, no. 1, pp. 647–664, Jan. 2017.

[13] H. Sun and S. A. Jafar, "Private information retrieval from MDS coded data with colluding servers: Settling a conjecture by Freij-hollanti *et al.*," *IEEE Trans. Inf. Theory*, vol. 64, no. 2, pp. 1000–1022, Feb. 2018.

[14] R. Tandon, "The capacity of cache aided private information retrieval," in *Proc. 55th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Monticello, IL, USA, Oct. 2017, pp. 1078–1082.

[15] Y.-P. Wei, K. Banawan, and S. Ulukus, "Fundamental limits of cache-aided private information retrieval with unknown and uncoded prefetching," *IEEE Trans. Inf. Theory*, vol. 65, no. 5, pp. 3215–3232, May 2019.

[16] S. Kadhe, B. Garcia, A. Heidarzadeh, S. El Rouayheb, and A. Sprintson, "Private information retrieval with side information," *IEEE Trans. Inf. Theory*, vol. 66, no. 4, pp. 2032–2043, Apr. 2020.

[17] Z. Chen, Z. Wang, and S. A. Jafar, "The capacity of T-private information retrieval with private side information," *IEEE Trans. Inf. Theory*, vol. 66, no. 8, pp. 4761–4773, Aug. 2020.

[18] H. Sun and S. A. Jafar, "Multiround private information retrieval: Capacity and storage overhead," *IEEE Trans. Inf. Theory*, vol. 64, no. 8, pp. 5743–5754, Aug. 2018.

[19] H. Sun and S. A. Jafar, "The capacity of symmetric private information retrieval," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2016, pp. 1–5.

[20] Q. Wang and M. Skoglund, "Symmetric private information retrieval for MDS coded distributed storage," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.

[21] K. Banawan and S. Ulukus, "Multi-message private information retrieval: Capacity results and near-optimal schemes," *IEEE Trans. Inf. Theory*, vol. 64, no. 10, pp. 6842–6862, Oct. 2018.

[22] Y. Zhang and G. Ge, "Private information retrieval from MDS coded databases with colluding servers under several variant models," 2017, *arXiv:1705.03186*. [Online]. Available: http://arxiv.org/abs/1705.03186

[23] K. Banawan and S. Ulukus, "The capacity of private information retrieval from Byzantine and colluding databases," *IEEE Trans. Inf. Theory*, vol. 65, no. 2, pp. 1206–1219, Feb. 2019.

[24] R. Tajeddine, O. W. Gnilke, D. Karpuk, R. Freij-Hollanti, and C. Hollanti, "Robust private information retrieval from coded systems with Byzantine and colluding servers," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2018, pp. 2451–2455.

[25] K. Banawan and S. Ulukus, "Private information retrieval through wiretap channel II: Privacy meets security," *IEEE Trans. Inf. Theory*, vol. 66, no. 7, pp. 4129–4149, Jul. 2020.

[26] K. Banawan and S. Ulukus, "Asymmetry hurts: Private information retrieval under asymmetric traffic constraints," *IEEE Trans. Inf. Theory*, vol. 65, no. 11, pp. 7628–7645, Nov. 2019.

[27] C. Tian, H. Sun, and J. Chen, "A Shannon-theoretic approach to the storage-retrieval tradeoff in PIR systems," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2018, pp. 1904–1908.

[28] Y.-P. Wei, B. Arasli, K. Banawan, and S. Ulukus, "The capacity of private information retrieval from decentralized uncoded caching databases," *Information*, vol. 10, no. 12, p. 372, Nov. 2019.

[29] M. A. Maddah-Ali and U. Niesen, "Decentralized coded caching attains order-optimal memory-rate tradeoff," *IEEE/ACM Trans. Netw.*, vol. 23, no. 4, pp. 1029–1040, Aug. 2015.

[30] K. Banawan, B. Arasli, Y.-P. Wei, and S. Ulukus, "The capacity of private information retrieval from heterogeneous uncoded caching databases," *IEEE Trans. Inf. Theory*, vol. 66, no. 6, pp. 3407–3416, Jun. 2020.

[31] R. Tandon, M. Abdul-Wahid, F. Almoualem, and D. Kumar, "PIR from storage constrained databases-coded caching meets PIR," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–6.

[32] M. A. Maddah-Ali and U. Niesen, "Fundamental limits of caching," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2856–2867, May 2014.

[33] M. A. Attia, D. Kumar, and R. Tandon, "The capacity of uncoded storage constrained PIR," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2018, pp. 1959–1963.

[34] K. Wan, D. Tuninetti, and P. Piantanida, "On the optimality of uncoded cache placement," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Sep. 2016, pp. 161–165.

Authorized licensed use limited to: The University of Arizona. Downloaded on January 30,2021 at 20:39:51 UTC from IEEE Xplore. Restrictions apply.

[35] M. Adel Attia and R. Tandon, "Near optimal coded data shuffling for distributed learning," *IEEE Trans. Inf. Theory*, vol. 65, no. 11, pp. 7325–7349, Nov. 2019.
[36] M. A. Attia and R. Tandon, "Approximately optimal distributed data shuffling," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2018, pp. 721–725.
[37] M. Bahrami, M. A. Attia, R. Tandon, and B. Vasic, "Towards the exact rate-memory trade-off for uncoded caching with secure delivery," in *Proc. 55th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Oct. 2017, pp. 878–885.

Mohamed Adel Attia (Member, IEEE) received the B.Sc. degree in electrical engineering from Alexandria University, Alexandria, Egypt, in 2012, and the M.Sc. degree in electronics and communications engineering from The American University, Cairo, Egypt, in 2015. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, The University of Arizona, Tucson, AZ, USA. He joined The University of Arizona as a Graduate Research Assistant in 2015. His current research interests include information theory, machine learning, wireless communications, private information retrieval, and distributed data analysis.

Deepak Kumar received the B.E. degree in electrical and electronics engineering from the M. S. Ramaiah Institute of Technology, Bengaluru, India, in 2012, and the M.S. degree in electrical and computer engineering from The University of Arizona in 2018. He is currently working as a Data Analyst at Citi Bank. His research interests include information theory, machine learning, and data analytics.

Ravi Tandon (Senior Member, IEEE) received the B.Tech. degree in electrical engineering from IIT Kanpur in 2004 and the Ph.D. degree in electrical and computer engineering from the University of Maryland, College Park (UMCP), in 2010. From 2010 to 2012, he was a Post-Doctoral Research Associate with Princeton University. He is currently an Associate Professor with the Department of Electrical and Computer Engineering (ECE), The University of Arizona. Prior to joining The University of Arizona in Fall 2015, he was a Research Assistant Professor at Virginia Tech with positions at the Bradley Department of ECE, Hume Center for National Security and Technology, and the Discovery Analytics Center, Department of Computer Science. His current research interests include information theory and its applications to wireless networks, communications, security and privacy, machine learning, and data mining. He was a recipient of the 2018 Keysight Early Career Professor Award, the NSF CAREER Award in 2017, and the Best Paper Award at the IEEE GLOBECOM 2011. He serves as an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and the IEEE TRANSACTIONS ON COMMUNICATIONS.

Authorized licensed use limited to: The University of Arizona. Downloaded on January 30,2021 at 20:39:51 UTC from IEEE Xplore. Restrictions apply.