Private Information Delivery

Hua Sun, *Member, IEEE*

***Abstract*— We introduce the problem of private information delivery (PID), comprised of K messages, a user, and N servers (each holds $M \leq K$ messages) that wish to deliver one out of K messages to the user privately, i.e., without revealing the delivered message index to the user. The information theoretic capacity of PID, C, is defined as the maximum number of bits of the desired message that can be privately delivered per bit of total communication to the user. For the PID problem with K messages, N servers, M messages stored per server, and $N \geq \lceil \frac{K}{M} \rceil$, we provide an achievable scheme of rate $1/\lceil \frac{K}{M} \rceil$ and an information theoretic converse of rate M/K, i.e., the PID capacity satisfies $1/\lceil \frac{K}{M} \rceil \leq C \leq M/K$. This settles the capacity of PID when $\frac{K}{M}$ is an integer. When $\frac{K}{M}$ is not an integer, we show that the converse rate of M/K is achievable if $N \geq \frac{K}{\gcd(K,M)} - (\frac{M}{\gcd(K,M)} - 1)(\lfloor \frac{K}{M} \rfloor - 1)$, and the achievable rate of $1/\lceil \frac{K}{M} \rceil$ is optimal if $N = \lceil \frac{K}{M} \rceil$. Otherwise if $\lceil \frac{K}{M} \rceil < N < \frac{K}{\gcd(K,M)} - (\frac{M}{\gcd(K,M)} - 1)(\lfloor \frac{K}{M} \rfloor - 1)$, we give an improved achievable scheme and prove its optimality for several small settings.**

***Index Terms*— Capacity, private information delivery, privacy.**

I. Introduction

CONSIDER a dataset comprised of K identically distributed messages and stored over N servers. The servers wish to deliver one of the messages to a user without revealing the identity of the message delivered, i.e., the user does not know which message is delivered to him. For example, the dataset may be medical records from a hospital and each message represents the medical record of a patient. The hospital would like to send the medical record of a patient externally (e.g., for analysis of certain disease that goes beyond the capability of the current hospital), and it is desirable that the name of the patient is not revealed (i.e., the privacy of the patient is preserved). Note that we consider sensitive medical records that are not publicly available (e.g., data-anonymization or privacy-preserving techniques are not applicable). For another example, suppose a company outsources some of its user activity log data externally for statistical analysis, while it does not wish to reveal sensitive information about the user identities (e.g., names, addresses, groups). We call this problem private[1] information delivery (PID).

Manuscript received September 6, 2019; revised April 18, 2020; accepted July 26, 2020. Date of publication August 11, 2020; date of current version November 20, 2020. This article was presented in part at the 2019 IEEE Information Theory Workshop.

The author is with the Department of Electrical Engineering, University of North Texas, Denton, TX 76203 USA (e-mail: hua.sun@unt.edu).

Communicated by A. Thangaraj, Associate Editor for Coding Techniques.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TIT.2020.3015812.

Digital Object Identifier 10.1109/TIT.2020.3015812

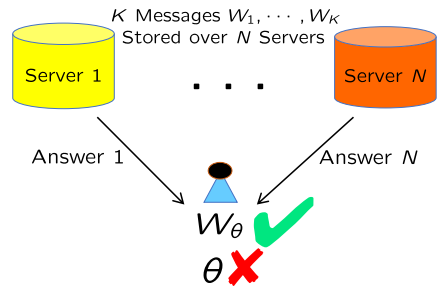

Fig. 1. The private information delivery problem.

This PID problem is trivial for a centralized system, i.e., there is a single server that stores all K messages. In this case, no matter which message the server wishes to deliver, the server simply sends the message to the user and all K choices are indistinguishable from the user. Recently, a fully distributed system is studied in [3], where there are K messages and $N = K$ servers, each stores one message. An example with $K = 3$ and an optimal private coding strategy are shown as follows:

	Server 1	Server 2	Server 3
Storage	W_1, z_1	W_2, z_2	$W_3, z_1 + z_2$
Answer for W_1	$W_1 + z_1$	z_2	$z_1 + z_2$
Answer for W_2	z_1	$W_2 + z_2$	$z_1 + z_2$
Answer for W_3	z_1	z_2	$W_3 + z_1 + z_2$

(1)

Here we have 3 independent messages W_1, W_2, W_3 (one bit each). The servers are equipped with some correlated random variables $z_1, z_2, z_1 + z_2$ that are independent of the messages and z_1, z_2 are two i.i.d. fair coin tosses.

To ensure information theoretic privacy, we need to guarantee that regardless of the message index delivered, the answers seen by the user are identically distributed and the decoding rule remains the same (otherwise, the decoding rule reveals information about the message delivered). For the scheme above, no matter W_1, W_2, or W_3 is to be delivered, the user sees 3 i.i.d. random bits and to decode the desired message, he always adds up the 3 answering strings. In [3], it is proved that the communication rate of $1/3$ is optimal, where the rate is defined as the number of bits privately delivered per bit of total answers sent to the user. For the above $N = K$ and each

[1] In a previous version of this work [1], the problem is called *anonymous* information delivery. We make a clear distinction of privacy and anonymity here, where privacy refers to the behavior or interest of an entity (e.g., which message is delivered) and anonymity refers to the entities of certain activity (e.g., who pays the bill [2]).

server stores $M = 1$ message case, the maximum rate (termed the capacity, C) is $1/K$. Further, it is necessary for each server to hold 1 bit of correlated randomness and for all servers to hold $K - 1$ bits of correlated randomness, per message bit.

As the fully distributed and centralized cases are well understood, our goal in this article is to study the intermediate partially distributed case - each server stores M out of K messages ($1 \leq M \leq K$). We are restricted to replicated systems (i.e., we do not allow coded messages or splitting one message to several servers) in this work,[2] as a first step towards more complex scenarios and a practical set-up for distributed storage systems. Note that we allow the design of the M messages stored. That is, we wish to find the best replication strategy and the corresponding private delivery scheme. The main motivation of this work is to characterize the capacity of PID for replicated systems, as a function of the number of messages, K, the number of servers, N, and the number of messages stored per server, M.

As an example, consider the setting where we have $K = 3$ messages, $N = 3$ servers and $M = 2$ messages are stored per server. The storage and correlated randomness design and the private coding scheme are shown in (2), at the bottom of the page. Here each message is made up of two symbols from $\mathbb{F}_5$, $W_1 = (a_1, a_2)$, $W_2 = (b_1, b_2)$ and $W_3 = (c_1, c_2)$. z is a common random variable shared by the servers and z is uniformly distributed over $\mathbb{F}_5$ (independent of the messages).

We denote the answer from Server $n, n \in \{1, 2, 3\}$ by A_n. Note that A_n is a function of the storage at Server n. To decode the desired message, in all 3 cases where W_1, W_2 or W_3 is delivered, the user employs the same decoding strategy, as follows.

$$\text{Desired Symbol 1} = A_1 + A_2 + A_3 \tag{3}$$
$$\text{Desired Symbol 2} = A_1 + 2A_2 + 3A_3. \tag{4}$$

Further, in all 3 cases, the user receives 3 uniformly random symbols over $\mathbb{F}_5$, thus perfect privacy is achieved. The rate achieved is $2/3$ as 2 symbols are delivered over 3 answering symbols. As we will show later by an information theoretic converse, the rate of $2/3$ is also the maximum possible. Thus the capacity of PID is $2/3$ in this case.

The main contribution of this work is summarized next. We first show that $1/\lceil \frac{K}{M} \rceil \leq C \leq M/K$ by an achievable scheme of rate $1/\lceil \frac{K}{M} \rceil$ and a converse of rate M/K. As a result, we have $C = M/K$ when $\frac{K}{M} \in \mathbb{Z}$. Otherwise, if $\frac{K}{M} \notin \mathbb{Z}$, we prove that when $N \geq \frac{K}{\gcd(K,M)} - (\frac{M}{\gcd(K,M)} - 1)(\lfloor \frac{K}{M} \rfloor - 1)$, the converse rate of M/K is achievable, and when $N = \lceil \frac{K}{M} \rceil$, the achievable rate of $1/\lceil \frac{K}{M} \rceil$ is optimal. For the uncovered regime where $\lceil \frac{K}{M} \rceil < N < \frac{K}{\gcd(K,M)} - (\frac{M}{\gcd(K,M)} - 1)(\lfloor \frac{K}{M} \rfloor - 1)$, we provide an improved achievable scheme and show that it is optimal for certain small cases. Therefore, we have characterized the capacity of PID for most cases, and provided close upper and lower bounds for remaining cases.

Notation: For integers $N_1, N_2, N_1 \leq N_2$, define the notation $[N_1 : N_2]$ as the set $\{N_1, N_1 + 1, \cdots, N_2\}$ and $(N_1 : N_2)$ as the vector $(N_1, N_1 + 1, \cdots, N_2)$. For an index set $\mathcal{I} = \{i_1, i_2, \cdots, i_n\}$, the notation $A_{\mathcal{I}}$ represents the set $\{A_i : i \in \mathcal{I}\}$. For an index vector $\overrightarrow{\mathcal{I}} = (i_1, i_2, \cdots, i_n)$, the notation $A_{\overrightarrow{\mathcal{I}}}$ represents the vector $(A_{i_1}, A_{i_2}, \cdots, A_{i_n})$. For sets (vectors) $\mathcal{I}_1, \mathcal{I}_2$, we define $\mathcal{I}_1 \backslash \mathcal{I}_2$ as the set (vector) of elements that are in $\mathcal{I}_1$ and not in $\mathcal{I}_2$ (in original order). The notation $X \sim Y$ is used to indicate that random variables X and Y are identically distributed. For a matrix $\mathbf{F}$ with i rows and j columns, if we wish to highlight its dimension, we will write $\mathbf{F}_{i \times j}$. For an index vector $\overrightarrow{\mathcal{I}} = (i_1, i_2, \cdots, i_n)$, the notation $\mathbf{F}_{[\overrightarrow{\mathcal{I}},:]}$ represents the submatrix of $\mathbf{F}$ formed by retaining only the rows corresponding to the elements of the vector $\overrightarrow{\mathcal{I}}$. The notation $\mathbf{F}_{[:,\overrightarrow{\mathcal{I}}]}$ is defined similarly (with respect to the columns). The notation $\mathbf{I}_j$ represents the identity matrix with dimension $j \times j$ and the notation $\mathbf{0}$ represents a matrix where each element is 0.

II. Problem Statement

Consider K independent messages $W_1, \cdots, W_K$. Each message is comprised of L i.i.d. uniform symbols from a finite field $\mathbb{F}_p$. In p-ary units,

$$H(W_1) = \cdots = H(W_K) = L, \tag{5}$$
$$H(W_1, \cdots, W_K) = H(W_1) + \cdots + H(W_K). \tag{6}$$

There are N servers, and each server stores M out of the K messages. We denote the storage variable at Server n as S_n.

$$S_n = W_{\mathcal{S}_n},\ \mathcal{S}_n \subset [1 : K], |\mathcal{S}_n| = M. \tag{7}$$

The servers share a common random variable Z, and Z is independent of the messages.

$$H(Z, W_1, \cdots, W_K) = H(Z) + H(W_1) + \cdots + H(W_K). \tag{8}$$

The servers privately generate $\theta \in [1 : K]$ and wish to deliver W_θ to a user while keeping θ a secret from the user.

[2]It turns out that the PID problem is trivial when we may distribute (a distinct part of) *each* message to *each* server as in this case, rate 1 can be achieved easily and the system is essentially centralized in the sense of PID. Therefore for the PID problem, the more interesting case of distributed systems refers to that *some* message is not available at all at *some* server, and we wish to confuse the user about which message is delivered.

	Server 1	Server 2	Server 3
Storage	W_1, W_2, z	W_2, W_3, z	W_3, W_1, z
Answer for W_1	$\frac{3}{2}a_1 - \frac{1}{2}a_2 + z$	$-2z$	$-\frac{1}{2}a_1 + \frac{1}{2}a_2 + z$
Answer for W_2	$2b_1 - b_2 + z$	$-b_1 + b_2 - 2z$	z
Answer for W_3	z	$3c_1 - c_2 - 2z$	$-2c_1 + c_2 + z$

(2)

Depending on θ, there are K strategies that the servers could employ to privately deliver the desired message. For example, if $\theta = k$, then in order to deliver W_k, Server $n \in [1:N]$ sends an answer $A_n^{[k]}$ to the user. The answer $A_n^{[k]}$ is a function of S_n, Z,

$$\forall k \in [1:K], n \in [1:N], \quad H(A_n^{[k]}|S_n, Z) = 0. \tag{9}$$

From all N answers, the user decodes the desired message with zero error.

$$H(W_k|A_1^{[k]}, A_2^{[k]}, \cdots, A_N^{[k]}) = 0. \tag{10}$$

To ensure privacy, the communication strategies must be indistinguishable (identically distributed) from the perspective of the user, i.e., the following privacy constraint must be satisfied, $\forall k \in [1:K]$,

[Privacy]

$$(A_1^{[1]}, A_2^{[1]}, \cdots, A_N^{[1]}, W_1) \sim (A_1^{[k]}, A_2^{[k]}, \cdots, A_N^{[k]}, W_k). \tag{11}$$

The privacy constraint (11) is equivalent to the condition that the answers are i.i.d. and the (deterministic) decoding mappings from the answers to the desired message are the same for all k.

The PID *rate* characterizes how many symbols of desired information are delivered per symbol of total delivery, and is defined as

$$R \triangleq \frac{L}{\sum_{n=1}^{N} D_n} \tag{12}$$

where D_n is the expected number of symbols sent from Server n to the user.

A rate R is said to be achievable if there exists a PID scheme of rate greater than or equal to R, for which zero error decoding is guaranteed. The supremum of achievable rates (over all storage design $S_1, S_2, \cdots, S_N$ and all PID schemes) is called the capacity C.

The randomness size η measures the amount of common randomness at the servers relative to the message size.

$$\eta = \frac{H(Z)}{L}. \tag{13}$$

In this work, we focus on the capacity C and allow as much common randomness as needed.

III. Results

In this section, we state the main results of this work. We start with a upper bound and a lower bound on the PID capacity, stated in the following theorem.

Theorem 1: For the private information delivery problem with K messages, $N \geq \lceil \frac{K}{M} \rceil$ servers and M messages per server, the capacity satisfies

$$1/\lceil \frac{K}{M} \rceil \leq C \leq M/K. \tag{14}$$

We need $N \geq \lceil \frac{K}{M} \rceil$ because otherwise the total storage available at all servers, NM, is smaller than the number of messages, K, and we cannot guarantee that all messages can be delivered correctly. To prove Theorem 1, we provide an achievable scheme of rate $1/\lceil \frac{K}{M} \rceil$ and a converse of rate M/K. The details are presented in Section IV.

The bounds in Theorem 1 match when $\frac{K}{M}$ is an integer. Therefore, in this case, we obtain the exact capacity of PID, stated in the following corollary.

Corollary 1: For the private information delivery problem with K messages, N servers and M messages per server, if $\frac{K}{M} \in \mathbb{Z}$, $N \geq \frac{K}{M}$, the capacity is $C = M/K$.

Remark: When $N = K, M = 1$ (the fully distributed system), $C = 1/K$ and this recovers the capacity result of Theorem 1 in [3]. When $M = K$, we have the fully centralized system and $C = 1$.

When $\frac{K}{M}$ is not an integer, the upper bound and the lower bound in Theorem 1 on the PID capacity are close. The inverse of the capacity ($\frac{1}{C}$, referred to as the optimal download cost) is characterized to within a 1 symbol gap ($= \lceil K/M \rceil - K/M$).

Next, we proceed to consider the conditions on the number of servers N such that the bounds in Theorem 1 are tight.

Theorem 2: For the private information delivery problem with K messages, $N \geq \lceil \frac{K}{M} \rceil$ servers, M messages per server, and $\frac{K}{M} \notin \mathbb{Z}$, the converse rate of M/K is achievable if $N \geq \frac{K}{\gcd(K,M)} - (\frac{M}{\gcd(K,M)} - 1)(\lfloor \frac{K}{M} \rfloor - 1)$, and the achievable rate of $1/\lceil \frac{K}{M} \rceil$ is optimal if $N = \lceil \frac{K}{M} \rceil$.

The proof of Theorem 2 is presented in Section V.

Combining Theorem 1 and Theorem 2, we have characterized the PID capacity when the number of servers is either small or large. In particular, the full regime is characterized when $M = 2$. This result is stated in the following corollary.

Corollary 2: For the private information delivery problem with K messages, N servers and $M = 2$ messages per server, the capacity is

$$C = \begin{cases} 2/K & \text{when } N \geq \lceil \frac{K}{2} \rceil + 1 \\ 1/\lceil \frac{K}{2} \rceil & \text{when } N = \lceil \frac{K}{2} \rceil \\ 0 & \text{when } N < \lceil \frac{K}{2} \rceil \end{cases}. \tag{15}$$

Proof: The case for even K is obvious (covered in Corollary 1) and we only need to consider the case when K is odd. We prove that $\lceil \frac{K}{2} \rceil + 1 = \frac{K}{\gcd(K,M)} - (\frac{M}{\gcd(K,M)} - 1)(\lfloor \frac{K}{M} \rfloor - 1)$, when $M = 2$. Plugging in $M = 2$ to the RHS, we have $K - \lfloor \frac{K}{2} \rfloor + 1 = \text{LHS}$ and the proof is complete. ■

Note that when $M = 2$, the difference between the two thresholds $\left(\lceil \frac{K}{M} \rceil \text{ and } \frac{K}{\gcd(K,M)} - (\frac{M}{\gcd(K,M)} - 1)(\lfloor \frac{K}{M} \rfloor - 1)\right)$ on N is exactly 1. Then we know that the thresholds (conditions for the bounds in Theorem 1 to be tight) cannot be improved in general.

The results obtained so far are summarized in Figure 2. Beyond the intermediate regime $\lceil \frac{K}{M} \rceil < N < \frac{K}{\gcd(K,M)} - (\frac{M}{\gcd(K,M)} - 1)(\lfloor \frac{K}{M} \rfloor - 1)$, we have characterized the PID capacity. The range of N in this regime is at most $\frac{2M}{\gcd(K,M)} - 1$, i.e.,

$$\begin{aligned} & \frac{K}{\gcd(K,M)} - (\frac{M}{\gcd(K,M)} - 1)(\lfloor \frac{K}{M} \rfloor - 1) - \lceil \frac{K}{M} \rceil \\ = \; & \frac{M}{\gcd(K,M)} \frac{K}{M} - \frac{M}{\gcd(K,M)} (\lfloor \frac{K}{M} \rfloor - 1) \\ & + \lfloor \frac{K}{M} \rfloor - \lceil \frac{K}{M} \rceil - 1 \end{aligned} \tag{16}$$

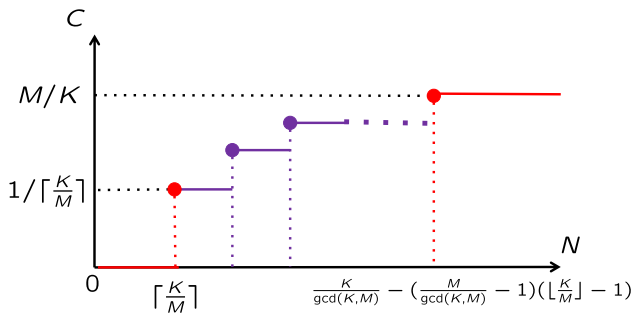

Fig. 2. The PID capacity. When $N \leq \lceil \frac{K}{M} \rceil$ and $N \geq \frac{K}{\gcd(K,M)} - (\frac{M}{\gcd(K,M)} - 1)(\lfloor \frac{K}{M} \rfloor - 1)$, the capacity is fully characterized (colored in red) and otherwise, the capacity is open in general (colored in purple).

$$\begin{aligned} &= \frac{M}{\gcd(K,M)}(\frac{K}{M} - \lfloor \frac{K}{M} \rfloor + 1) + \lfloor \frac{K}{M} \rfloor - \lceil \frac{K}{M} \rceil - 1 \\ &< \frac{2M}{\gcd(K,M)} - 1. \end{aligned} \tag{17}$$

Finally, we consider this intermediate regime and present an improved achievable scheme, in the following theorem.

Theorem 3: For the private information delivery problem with K messages, N servers, and M messages per server, when $\frac{K}{M} \notin \mathbb{Z}$ and $\lceil \frac{K}{M} \rceil < N < \frac{K}{\gcd(K,M)} - (\frac{M}{\gcd(K,M)} - 1)(\lfloor \frac{K}{M} \rfloor - 1)$, the capacity satisfies

$$\begin{aligned} C \geq \frac{l}{N + (l-1)(\lfloor \frac{K}{M} \rfloor - 1)}, \\ \text{where } l = \lfloor \frac{(N - \lfloor \frac{K}{M} \rfloor + 1)M}{K - (\lfloor \frac{K}{M} \rfloor - 1)M} \rfloor. \end{aligned} \tag{18}$$

The proof of Theorem 3 is presented in Section VI. To illustrate Theorem 3, we give two examples.

Example 1: Suppose $M = 3, K = 7$. The only N value that is covered in Theorem 3 is $N = 4$. The achievable rate in Theorem 3 is $2/5$. It turns out that this achievable rate is also optimal (proof deferred to Section VII-B). Therefore, we have characterized the capacity of PID for all possible values of N when $M = 3, K = 7$. This result is plotted in Figure 3(a).

Example 2: Suppose $M = 4, K = 5$. The only N values that are covered in Theorem 3 are $N = 3, 4$. The achievable rate in Theorem 3 is $2/3$ (when $N = 3$), and $3/4$ (when $N = 4$). It turns out that the achievable rates are also optimal (proof deferred to Section VII-A). Therefore, we have characterized the capacity of PID for all possible values of N when $M = 4, K = 5$. This result is plotted in Figure 3(b).

Remark: The achievable rate in Theorem 3 may not be monotonically increasing in N. So for a given N, if we want to find the highest achievable rate, we may search over all $N' \in [\lceil \frac{K}{M} \rceil + 1 : N]$.

Remark: The achievable scheme in Theorem 3 includes those in Theorem 1 and Theorem 2 as special cases. That is, if we set $N = \lceil \frac{K}{M} \rceil$, the rate achieved in Theorem 3 is $R = 1/\lceil \frac{K}{M} \rceil$ (same as that in Theorem 1), and if we set $N = \frac{K}{\gcd(K,M)} - (\frac{M}{\gcd(K,M)} - 1)(\lfloor \frac{K}{M} \rfloor - 1)$, the rate achieved in Theorem 3 is $R = M/K$ (same as that in Theorem 2).

IV. Proof of Theorem 1

A. Converse: $R \leq M/K$

Let us start with two useful lemmas. The first lemma states that if a message is available at a set of servers, then the size of the answers from these servers must be no less than the message size.

Lemma 1: Consider any storage strategy where W_k is only available at servers in the set $\mathcal{N}_k = \{n_{k_1}, n_{k_2}, \cdots, n_{k_i}\}$, i.e., $W_k \in S_j, \forall j \in \mathcal{N}_k$, and $W_k \notin S_l, \forall l \notin \mathcal{N}_k$. We have

$$D_{\mathcal{N}_k^\Sigma} \triangleq D_{n_{k_1}} + D_{n_{k_2}} + \cdots + D_{n_{k_i}} \geq L. \tag{19}$$

Proof:

$$\begin{aligned} L &\overset{(5)}{=} H(W_k) && (20)\\ &\overset{(10)}{=} I(W_k; A_1^{[k]}, A_2^{[k]}, \cdots, A_N^{[k]}) && (21)\\ &= I(W_k; A_{[1:N]\setminus\mathcal{N}_k}^{[k]}) + I(W_k; A_{\mathcal{N}_k}^{[k]} | A_{[1:N]\setminus\mathcal{N}_k}^{[k]}) && (22)\\ &\leq I(W_k; A_{[1:N]\setminus\mathcal{N}_k}^{[k]}, W_{[1:K]\setminus\{k\}}, Z) + H(A_{\mathcal{N}_k}^{[k]}) && (23)\\ &\overset{(8)(11)}{\leq} I(W_k; A_{[1:N]\setminus\mathcal{N}_k}^{[k]} | W_{[1:K]\setminus\{k\}}, Z) + D_{\mathcal{N}_k^\Sigma} && (24)\\ &\overset{(7)}{=} I(W_k; A_{[1:N]\setminus\mathcal{N}_k}^{[k]} | W_{[1:K]\setminus\{k\}}, Z, S_{[1:N]\setminus\mathcal{N}_k}) \\ &\quad + D_{\mathcal{N}_k^\Sigma} && (25)\\ &\overset{(9)}{=} D_{\mathcal{N}_k^\Sigma} && (26) \end{aligned}$$

where (25) follows from the constraint that W_k is not available at Server $l, \forall l \in [1 : N] \setminus \mathcal{N}_k$ so that $S_{[1:N]\setminus\mathcal{N}_k} \subset W_{[1:K]\setminus\{k\}}$. ■

The second lemma states that having multiple servers storing the same set of messages does not help to reduce the private delivery rate.

Lemma 2: Consider any storage strategy $S_1, S_2, \cdots, S_N$ with $N' \leq N$ distinct S_i storage variables. Without loss of generality, assume $S_i \neq S_j, \forall i \neq j, i, j \in [1 : N']$. Then any rate R that is achievable with N servers and the storage strategy $S_1, S_2, \cdots, S_N$ is also achievable with N' servers and the storage strategy $S_1, S_2, \cdots, S_{N'}$.

Proof: Suppose we are given a PID scheme (described by NK answers $A_n^{[k]}, n \in [1 : N], k \in [1 : K]$) that operates over N servers with the storage strategy $S_1, S_2, \cdots, S_N$, where $S_1, \cdots, S_{N'}$ are distinct. Denote the set of server indices for which the storage variables are equal to $S_i, i \in [1 : N']$ by $\mathcal{M}_i$, i.e., if $j \in \mathcal{M}_i$, then $S_j = S_i$. Then we have N' disjoint $\mathcal{M}_i$ sets that form a partition of the N servers, i.e., $\mathcal{M}_1 \cup \mathcal{M}_2 \cup \cdots \mathcal{M}_{N'} = [1 : N]$, and $\mathcal{M}_{i_1} \cap \mathcal{M}_{i_2} = \emptyset, \forall i_1 \neq i_2, i_1, i_2 \in [1 : N']$.

Next we will construct a PID scheme that operates over N' servers with the storage strategy $S_1, S_2, \cdots, S_{N'}$ and achieves the same rate as the N-server scheme above. We will use notations with a tilde symbol to describe the N'-server scheme. The common random variable remains the same, $\widetilde{Z} = Z$. The answer from Server i to deliver $\widetilde{W}_k$ is denoted by $\widetilde{A}_i^{[k]}$. We set

$$\widetilde{A}_i^{[k]} = A_{\overrightarrow{\mathcal{M}_i}}^{[k]}, \ \forall n \in [1 : N'], k \in [1 : K] \tag{27}$$

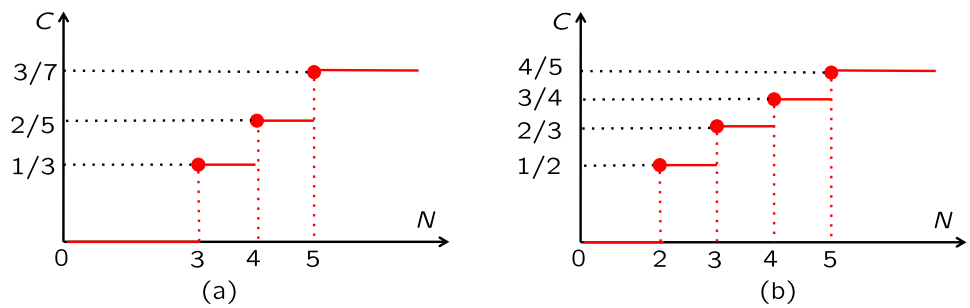

Fig. 3. The PID capacity, (a) when $M = 3, K = 7$, and (b), when $M = 4, K = 5$.

where $\overrightarrow{\mathcal{M}_i}$ is a vector that is in increasing order of the elements in the set $\mathcal{M}_i$. Note that the storage variable of all servers in the set $\mathcal{M}_i$ is the same as that of Server i, so that we may set the answers as above (refer to (9)). After collecting all N' answers, we have that

$$\text{the } A_n^{[k]} \text{ variables in } (\widetilde{A}_1^{[k]}, \cdots, \widetilde{A}_{N'}^{[k]}) \text{ are a permutation of } (A_1^{[k]}, \cdots, A_N^{[k]}) \tag{28}$$

so that we may use the decoding mapping (the order of the arguments in the mapping is correspondingly permuted) from the N-server scheme to decode $\widetilde{W}_k$. From (27) and (28), it is easy to see that the privacy constraint inherits and the same rate is preserved. The proof is therefore complete. ■

We are now ready to show that $R \leq M/K$. From Lemma 2, we may assume without loss of generality that the storage variables $S_n, n \in [1 : N]$ are distinct. Note that S_n is comprised of M out of K messages, so we have at most $\binom{K}{M}$ distinct S_n variables. In other words, we may assume $N = \binom{K}{M}$ (note that having more servers cannot hurt). Further, suppose the sets of stored messages $\mathcal{S}_1, \cdots, \mathcal{S}_N$ are ordered lexicographically. Consider any message $W_k, k \in [1 : K]$, and W_k is available at $\binom{K-1}{M-1}$ servers and this set of servers is denoted by $\mathcal{N}_k$, where $|\mathcal{N}_k| = \binom{K-1}{M-1}$. From Lemma 1, we have

$$L \leq D_{\mathcal{N}_k^\Sigma}. \tag{29}$$

Adding (29) for all $k \in [1 : K]$, we have

$$KL \leq \sum_{k=1}^{K} D_{\mathcal{N}_k^\Sigma} = M \sum_{n=1}^{N} D_n \tag{30}$$

where the last step follows from symmetry and any $D_n, n \in [1 : N]$ appears M times (Server n contains M messages and D_n appears once for each message). Rearranging terms gives us the rate bound and completes the proof:

$$R = \frac{L}{\sum_{n=1}^{N} D_n} \leq M/K. \tag{31}$$

B. Achievability: $R \geq 1/\lceil \frac{K}{M} \rceil$

We provide a scheme with $N = \lceil \frac{K}{M} \rceil$ servers. Suppose each message is comprised of $L = 1$ symbol from $\mathbb{F}_2$ (in fact, any field will work). The common random variable Z consists of $N - 1$ i.i.d. symbols, each from the same field $\mathbb{F}_2$. We denote $Z = (z_1, \cdots, z_{N-1})$.

The storage design is trivial, where the messages are stored sequentially over the servers.

$$S_1 = \{W_1, W_2, \cdots, W_M\} \tag{32}$$
$$S_2 = \{W_{M+1}, W_{M+2}, \cdots, W_{2M}\} \tag{33}$$
$$\vdots \tag{34}$$
$$S_N = \{W_{(N-1)M+1}, \cdots, W_K\}. \tag{35}$$

Suppose $W_k, k \in [1 : K]$ is desired. The delivery scheme is linear, and each answer has $D_i = 1, \forall i \in [1 : N]$ symbol. Then the rate achieved is $R = L/\sum_i D_i = 1/N = 1/\lceil \frac{K}{M} \rceil$, as desired. The answers are shown below.

$$A_i^{[k]} = z_i + 1(k \in [(i-1)M + 1 : iM])W_k, \quad i \in [1 : N-1], \tag{36}$$
$$A_N^{[k]} = -z_1 - \cdots - z_{N-1} + 1(k \in [(N-1)M + 1 : K])W_k \tag{37}$$

where $1(x)$ denotes the indicator function that is equal to 1 if x is true and 0 otherwise. Note that the answering symbol from Server i contains W_k only if W_k is available at Server i.

To decode the desired message symbols, we add the N answering symbols.

$$W_k = A_1^{[k]} + A_2^{[k]} + \cdots + A_N^{[k]} \tag{38}$$

where all common randomness cancels and the desired message retains as it only appears once (in the answer from the server where it is stored). Note that the same decoding mapping is used for all k.

We next show that the privacy constraint (11) is satisfied. To this end, note that regardless of the vale of the desired message index k, the answers are independent uniform random bits, i.e.,

$$H(A_1^{[k]}, \cdots, A_N^{[k]}) = N. \tag{39}$$

Therefore, the scheme is both correct and private.

Finally, we note that $N - 1$ randomness symbols are used to send $L = 1$ message symbol. The randomness size is then $\eta = H(Z)/L = N - 1 = 1/R - 1$.

V. Proof of Theorem 2

A. $N = \frac{K}{\gcd(K,M)} - (\frac{M}{\gcd(K,M)} - 1)(\lfloor \frac{K}{M} \rfloor - 1)$: *Achievability of* $R = M/K$

1) Example With $K = 8, M = 3$*:* To illustrate the main idea in a simpler setting, we first consider an example with $K = 8, M = 3$ so that $N = 8 - (3-1)(2-1) = 6$ and we show that $R = M/K = 3/8$ is achievable.

Suppose the message size $L = 3$ symbols, and each symbol is from $\mathbb{F}_p$, where $p \geq 8$. Then $W_k, k \in [1:8]$ is a 3×1 vector. The common random variable consists of 5 i.i.d. symbols from the same field $\mathbb{F}_p$, i.e., $Z \in \mathbb{F}_p^{5\times 1}$. The storage is designed as follows, where the first 5 servers store $M = 3$ messages out of W_1, W_2, W_3, W_4, W_5 in a cyclic manner and the last server stores the remaining 3 messages W_6, W_7, W_8.

$$S_1 = \{W_1, W_2, W_3\} \tag{40}$$
$$S_2 = \{W_2, W_3, W_4\} \tag{41}$$
$$S_3 = \{W_3, W_4, W_5\} \tag{42}$$
$$S_4 = \{W_4, W_5, W_1\} \tag{43}$$
$$S_5 = \{W_5, W_1, W_2\} \tag{44}$$
$$S_6 = \{W_6, W_7, W_8\}. \tag{45}$$

Let us start with the case where W_1 is desired. The delivery scheme is linear, and the first 5 answer has $D_i = 1, \forall i \in [1:5]$ symbol each while the last answer has $D_6 = 3$ symbols. Then the rate achieved is $R = L/\sum_i D_i = 3/8$, as desired. The collection of the answers is shown below. Define $W_1 = (a_1, a_2, a_3), Z = (z_1, z_2, z_3, z_4, z_5)$.

$$\mathbf{A}^{[1]} \triangleq \begin{bmatrix} A_1^{[1]} \\ A_2^{[1]} \\ A_3^{[1]} \\ A_4^{[1]} \\ A_5^{[1]} \\ A_6^{[1]} \end{bmatrix} = \begin{bmatrix} \mathbf{f}_1^{[1]} & \mathbf{h}_1 \\ \mathbf{0}_{1\times 3} & \mathbf{h}_2 \\ \mathbf{0}_{1\times 3} & \mathbf{h}_3 \\ \mathbf{f}_4^{[1]} & \mathbf{h}_4 \\ \mathbf{f}_5^{[1]} & \mathbf{h}_5 \\ \mathbf{0}_{3\times 3} & \mathbf{H}_6 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \end{bmatrix} \tag{46}$$

where in answer $A_i^{[1]}, i \in [1:5]$ from Server i, $\mathbf{f}_i^{[1]}$ is a 1×3 precoding vector for the message symbols $W_1 \in \mathbb{F}_p^{3\times 1}$ and $\mathbf{h}_i$ is a 1×5 precoding vector for the common randomness symbols $Z \in \mathbb{F}_p^{5\times 1}$. In answer $A_6^{[1]}$, the 3×3 precoding matrix $\mathbf{F}_6^{[1]}$ for W_1 is set as the zero matrix and $\mathbf{H}_6$ is the 3×5 precoding matrix for Z. Note that as W_1 is not stored at Servers $2, 3, 6$, $\mathbf{f}_2^{[1]}, \mathbf{f}_3^{[1]}, \mathbf{F}_6^{[1]}$ must be zero. It turns out that in our scheme, the precoding vectors for the common randomness do not depend on the desired message index. Define

$$\mathbf{F}_{8\times 3}^{[1]} \triangleq \begin{bmatrix} \mathbf{f}_1^{[1]} \\ \mathbf{0}_{3\times 1} \\ \mathbf{0}_{3\times 1} \\ \mathbf{f}_4^{[1]} \\ \mathbf{f}_5^{[1]} \\ \mathbf{0}_{3\times 3} \end{bmatrix}, \quad \mathbf{H}_{8\times 5} \triangleq \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \mathbf{h}_3 \\ \mathbf{h}_4 \\ \mathbf{h}_5 \\ \mathbf{H}_6 \end{bmatrix} \tag{47}$$

and (46) may be re-written as

$$\mathbf{A}^{[1]} = [\mathbf{F}^{[1]} \quad \mathbf{H}] \begin{bmatrix} W_1 \\ Z \end{bmatrix} = \mathbf{F}^{[1]} W_1 + \mathbf{H} Z. \tag{48}$$

To decode the 3 desired message symbols from the 8 answering symbols, we apply a 3×8 linear filtering matrix $\mathbf{G}_{3\times 8}$ to $\mathbf{A}^{[1]}$. We have

$$W_1 = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \mathbf{G}\mathbf{A}^{[1]} = \mathbf{G}\mathbf{F}^{[1]} W_1 + \mathbf{G}\mathbf{H} Z, \tag{49}$$

and to satisfy (49), we set

$$\mathbf{G}\mathbf{F}^{[1]} = \mathbf{I}_3 \Rightarrow \mathbf{G}_{[:,(1,4,5)]} \mathbf{F}_{[(1,4,5),:]}^{[1]} = \mathbf{I}_3, \tag{50}$$
$$\mathbf{G}\mathbf{H} = \mathbf{0}_{3\times 5}. \tag{51}$$

Note that $\mathbf{G}_{[:,(1,4,5)]}, \mathbf{F}_{[(1,4,5),:]}^{[1]}$ are both square matrices.

The situation where $W_k, k \in [2:8]$ is desired is similar. The answers are

$$\mathbf{A}^{[k]} = \mathbf{F}^{[k]} W_k + \mathbf{H} Z \tag{52}$$

and the decoding constraints are (the answers are projected onto $\mathbf{G}$ to decode the desired message)

$$\mathbf{G}_{[:,(1,2,5)]} \mathbf{F}_{[(1,2,5),:]}^{[2]} = \mathbf{I}_3, \tag{53}$$
$$\mathbf{G}_{[:,(1,2,3)]} \mathbf{F}_{[(1,2,3),:]}^{[3]} = \mathbf{I}_3, \tag{54}$$
$$\mathbf{G}_{[:,(2,3,4)]} \mathbf{F}_{[(2,3,4),:]}^{[4]} = \mathbf{I}_3, \tag{55}$$
$$\mathbf{G}_{[:,(3,4,5)]} \mathbf{F}_{[(3,4,5),:]}^{[5]} = \mathbf{I}_3, \tag{56}$$
$$\mathbf{G}_{[:,(6,7,8)]} \mathbf{F}_{[(6,7,8),:]}^{[j]} = \mathbf{I}_3, j \in [6:8] \tag{57}$$
$$\mathbf{G}\mathbf{H} = \mathbf{0}_{3\times 5}. \tag{58}$$

Note that the same decoding mapping $\mathbf{G}$ must be used for each desired message. So the delivery design reduces to find a realization of the matrices $\mathbf{G}, \mathbf{F}^{[1]}, \mathbf{F}^{[2]}, \cdots, \mathbf{F}^{[8]}, \mathbf{H}$ such that (50), (51), (53) - (58) are satisfied.

These matrices are chosen as follows. We first set

$$\mathbf{G} = \begin{bmatrix} \mathbf{I}_3 & \mathbf{V}_{3\times 5} \end{bmatrix} \tag{59}$$

where

$$\mathbf{V}_{3\times 5} = \begin{bmatrix} \frac{1}{\alpha_1-\beta_1} & \frac{1}{\alpha_1-\beta_2} & \frac{1}{\alpha_1-\beta_3} & \frac{1}{\alpha_1-\beta_4} & \frac{1}{\alpha_1-\beta_5} \\ \frac{1}{\alpha_2-\beta_1} & \frac{1}{\alpha_2-\beta_2} & \frac{1}{\alpha_2-\beta_3} & \frac{1}{\alpha_2-\beta_4} & \frac{1}{\alpha_2-\beta_5} \\ \frac{1}{\alpha_3-\beta_1} & \frac{1}{\alpha_3-\beta_2} & \frac{1}{\alpha_3-\beta_3} & \frac{1}{\alpha_3-\beta_4} & \frac{1}{\alpha_3-\beta_5} \end{bmatrix},$$
$$\alpha_i, \beta_j, i \in [1:5], j \in [1:3] \text{ are all distinct.} \tag{60}$$

It is guaranteed that we can find such α_i, β_j because the field size $p \geq 8$. In other words, $\mathbf{V}$ is a Cauchy matrix where each square sub-matrix is invertible. Then $\mathbf{H}$ is solved from (51), as the right null space of $\mathbf{G}$.

$$\mathbf{H} = \begin{bmatrix} \mathbf{V}_{3\times 5} \\ -\mathbf{I}_5 \end{bmatrix}. \tag{61}$$

Next, the submatrices of $\mathbf{F}^{[k]}, k \in [1:8]$ are set as the inverse matrices of corresponding submatrices of $\mathbf{G}$, from (50), (53) - (57). Note that it is easy to see the corresponding submatrices of $\mathbf{G}$ have full rank such that their inverse matrices exist. Then $\mathbf{F}^{[k]}$ are fully determined as the rows that have not appeared are zero vectors, due to the storage constraint. Now all correctness constraints are satisfied. We are left to show that the privacy constraint (11) is satisfied. To this end, we show that regardless of the value of the desired message

index k, the answers are uniformly random, which translates to that the following matrices have full rank.

(Equivalent Privacy Condition):

$$\mathbf{B}_{8\times 8}^{[k]} \triangleq [\mathbf{F}^{[k]} \quad \mathbf{H}],\ \forall k \in [1:8] \text{ have full rank.} \quad (62)$$

As each $\mathbf{F}^{[k]}$ contains 5 zero rows, it suffices to show that any 5 rows of $\mathbf{H}$ are linearly independent (holds trivially by the determinant formula of Cauchy matrices). A more detailed proof will be presented in the general proof.

The construction of the matrices is not unique. In fact, it is not hard to show that if we choose each element of $\mathbf{G}$ i.i.d. and uniformly from a sufficiently large field and follow the above procedure to determine $\mathbf{F}^{[k]}$ and $\mathbf{H}$, then the solution will work with a high probability.

Finally, we note that 5 randomness symbols are used to send 3 message symbols. The randomness size is then $\eta = H(Z)/L = 5/3 = 1/R - 1$.

2) General Proof With Arbitrary K, M: We show that for K messages, M messages per server, and $N = \frac{K}{\gcd(K,M)} - (\frac{M}{\gcd(K,M)} - 1)(\lfloor \frac{K}{M} \rfloor - 1)$ servers, the rate $R = M/K$ is achievable. Note that the proof for larger N values is directly implied, as we only need a fewer number of servers in the achievable scheme.

We treat every $\gcd(K, M)$ messages as a block so that we have $\overline{K} \triangleq \frac{K}{\gcd(K,M)}$ message blocks. Define

$$\begin{aligned}\overline{W}_b &= \{W_{(b-1)\gcd(K,M)+1}, W_{(b-1)\gcd(K,M)+2}, \\ &\quad \cdots, W_{b\gcd(K,M)}\}, b \in [1:\overline{K}]. \end{aligned} \quad (63)$$

Each server now is able to store $\overline{M} \triangleq \frac{M}{\gcd(K,M)}$ message blocks.

Suppose the message size $L = \overline{M}$ symbols. Each symbol is from $\mathbb{F}_p$, where $p \geq \overline{K}$. The common random variable Z consists of $\overline{K} - L$ i.i.d. symbols, each from the same field $\mathbb{F}_p$.

We divide the N servers into 2 sets. The first set is made up of the first N_1 servers and the second set is made up of the last N_2 servers, where

$$N_2 = \lfloor \frac{K}{M} \rfloor - 1 \quad (64)$$

$$N_1 = N - N_2. \quad (65)$$

The message blocks also are divided into 2 sets, where the first set is comprised of the first N_1 message blocks and the second set is comprised of the remaining $\overline{K} - N_1$ message blocks.

The storage is designed as follows. In the first server set, each server stores L $(= \overline{M})$ message blocks out of the first message set in a cyclic manner. In the second server set, each server stores L distinct message blocks from the second message set sequentially.

$$S_1 = \{\overline{W}_1, \overline{W}_2, \cdots, \overline{W}_L\} \quad (66)$$

$$S_2 = \{\overline{W}_2, \overline{W}_3, \cdots, \overline{W}_{L+1}\} \quad (67)$$

$$\vdots \quad (68)$$

$$S_{N_1} = \{\overline{W}_{N_1}, \overline{W}_1, \cdots, \overline{W}_{L-1}\} \quad (69)$$

$$S_{N_1+1} = \{\overline{W}_{N_1+1}, \overline{W}_{N_1+2}, \cdots, \overline{W}_{N_1+L}\} \quad (70)$$

$$\vdots \quad (71)$$

$$S_N = \{\overline{W}_{N_1+(N_2-1)L+1}, \cdots, \overline{W}_{N_1+N_2L}\}. \quad (72)$$

To see that all messages are stored, we show that the last message block $\overline{W}_{N_1+N_2\ L}$ is indeed $\overline{W}_{\overline{K}}$,

$$\begin{aligned} N_1 + N_2\ L &= (N - N_2) + N_2\overline{M} \quad (73)\\ &= N + (\lfloor \frac{K}{M} \rfloor - 1)(\overline{M} - 1) \\ &\quad \text{(Using the definition of } N_2) \quad (74)\\ &= \overline{K} \\ &\quad \text{(Using the definition of } N). \quad (75)\end{aligned}$$

Suppose $W_k, k \in [1:K]$ is desired. The delivery scheme is linear, where each answer from the first server set has $D_i = 1, \forall i \in [1:N_1]$ symbol, and each answer from the second server set has $D_i = L, \forall i \in [N_1+1:N]$ symbols. Then the rate achieved is

$$R = \frac{L}{\sum_i D_i} = \frac{L}{N_1 + LN_2} \overset{(75)}{=} \overline{M}/\overline{K} = M/K \quad (76)$$

and it matches the desired rate expression. The answers are shown below.

$$A_i^{[k]} = \mathbf{F}_i^{[k]} W_k + \mathbf{H}_i Z \quad (77)$$

where if $i \in [1:N_1]$, $\mathbf{F}_i^{[k]}$ has dimension $1 \times L$, $\mathbf{H}_i$ has dimension $1 \times (\overline{K} - L)$, and otherwise if $i \in [N_1+1:N]$, $\mathbf{F}_i^{[k]}$ has dimension $L \times L$, $\mathbf{H}_i$ has dimension $L \times (\overline{K} - L)$. Define

$$\mathbf{F}_{\overline{K}\times L}^{[k]} = [\mathbf{F}_1^{[k]}; \mathbf{F}_2^{[k]}; \cdots; \mathbf{F}_N^{[k]}], \quad (78)$$

$$\mathbf{H}_{\overline{K}\times(\overline{K}-L)} = [\mathbf{H}_1; \mathbf{H}_2; \cdots; \mathbf{H}_N] \quad (79)$$

and we have the collection of all answers,

$$\mathbf{A}^{[k]} = [\mathbf{F}^{[k]} \quad \mathbf{H}] \begin{bmatrix} W_k \\ Z \end{bmatrix} = \mathbf{F}^{[k]} W_k + \mathbf{H} Z. \quad (80)$$

We next specify the availability set $\mathcal{N}_k$ of W_k, i.e., W_k is only available at Server n where $n \in \mathcal{N}_k$. Note that W_k belongs to message block $\overline{W}_{\overline{k}}$, where $\overline{k} \triangleq \lceil \frac{k}{\gcd(K,M)} \rceil$.

$$\mathcal{N}_k = \begin{cases} [\overline{k} - L + 1 : \overline{k}] \mod N_1 & \text{if } \overline{k} \in [1:N_1], \\ \lceil (\overline{k} - N_1)/\overline{M} \rceil & \text{else } \overline{k} \in [N_1+1:\overline{K}]. \end{cases} \quad (81)$$

Due to the above storage constraints, we have the following corresponding constraints on the precoding matrices.

$$\begin{aligned} &\text{If } \overline{k} \in [1:N_1],\ \mathbf{F}_{n_1}^{[k]} = \mathbf{0}_{1\times L}, \forall n_1 \notin \mathcal{N}_k, n_1 \in [1:N_1], \\ &\qquad \mathbf{F}_{n_2}^{[k]} = \mathbf{0}_{L\times L}, \forall n_2 \in [N_1+1:N], \\ &\text{else } \overline{k} \in [N_1+1:\overline{K}], \mathbf{F}_{n_1}^{[k]} = \mathbf{0}_{1\times L}, \forall n_1 \in [1:N_1], \\ &\qquad \mathbf{F}_{n_2}^{[k]} = \mathbf{0}_{L\times L}, \forall n_2 \notin \mathcal{N}_k, n_2 \in [N_1+1:N]. \end{aligned} \quad (82)$$

To decode the L desired message symbols from the $\overline{K}$ answering symbols, we apply a linear filtering matrix $\mathbf{G}_{L\times\overline{K}}$ to $\mathbf{A}^{[k]}$. We have

$$W_k = \mathbf{G}\mathbf{A}^{[k]} = \mathbf{G}\mathbf{F}^{[k]} W_k + \mathbf{G}\mathbf{H} Z, \quad (83)$$

and to satisfy (83), we set

$$\mathbf{G}\mathbf{F}^{[k]} = \mathbf{I}_L \Rightarrow$$
$$\text{If } \overline{k} \in [1:N_1], \mathbf{G}_{[:,\overrightarrow{\mathcal{N}}_k]}\mathbf{F}^{[k]}_{[\overrightarrow{\mathcal{N}}_k,:]} = \mathbf{I}_L,$$
$$\text{else } \overline{k} \in [N_1+1:\overline{K}],$$
$$\mathbf{G}_{[:,N_1+(\mathcal{N}_k-N_1-1)L+1:N_1+(\mathcal{N}_k-N_1)L]}\mathbf{F}^{[k]}_{\mathcal{N}_k} = \mathbf{I}_L; \quad (84)$$
$$\mathbf{G}\mathbf{H} = \mathbf{0}_{L\times(\overline{K}-L)} \quad (85)$$

where the vector $\overrightarrow{\mathcal{N}}_k$ is in increasing order of elements in the set $\mathcal{N}_k$ (the available set for message W_k). For example, suppose $M=6, K=20, k=2$. Then $\gcd(M,K)=2, N=6, N_1=4, N_2=2, L=3, \overline{k}=1$, $\mathcal{N}_2=\{3,4,1\}$, and $\overrightarrow{\mathcal{N}}_2=(1,3,4)$.

We next find matrices $\mathbf{G}, \mathbf{F}^{[1]}, \mathbf{F}^{[2]}, \cdots, \mathbf{F}^{[K]}, \mathbf{H}$ such that (84), (85) are satisfied for all $k \in [1:K]$. We first set

$$\mathbf{G} = \left[\begin{array}{cc} \mathbf{I}_L & \mathbf{V}_{L\times(\overline{K}-L)} \end{array}\right] \quad (86)$$

where $\mathbf{V}_{L\times(\overline{K}-L)}$ is a Cauchy matrix such that the element in i-th row and j-th column is given by

$$V_{ij} = \frac{1}{\alpha_i - \beta_j} \quad (87)$$

and α_i, β_j are distinct elements over $\mathbb{F}_p$ where $p \geq \overline{K}$. Then $\mathbf{H}$ is solved from (85) as the right null space of $\mathbf{G}$. The non-zero rows of $\mathbf{F}^{[k]}$ are solved from (84), as the inverse of some sub-matrices of $\mathbf{G}$.

$$\mathbf{H} = \left[\begin{array}{c} \mathbf{V}_{L\times(\overline{K}-L)} \\ -\mathbf{I}_{\overline{K}-L} \end{array}\right], \quad (88)$$
$$\mathbf{F}^{[k]}_{[\overrightarrow{\mathcal{N}}_k,:]} = \mathbf{G}^{-1}_{[:,\overrightarrow{\mathcal{N}}_k]}, \quad \text{if } \overline{k} \in [1:N_1],$$
$$\mathbf{F}^{[k]}_{\mathcal{N}_k} = \mathbf{G}^{-1}_{[:,N_1+(\mathcal{N}_k-N_1-1)L+1:N_1+(\mathcal{N}_k-N_1)L]},$$
$$\text{else } \overline{k} \in [N_1+1:\overline{K}]. \quad (89)$$

Note that if $\overline{k} \in [1:N_1]$, $\overrightarrow{\mathcal{N}}_k$ consists of L cyclicly consecutive elements in $[1:N_1]$ such that $\mathbf{G}_{[:,\overrightarrow{\mathcal{N}}_k]}$ is non-singular (its determinant is equal to the determinant of a square Cauchy matrix), and otherwise if $\overline{k} \in [N_1+1:\overline{K}]$, $\mathbf{G}_{[:,N_1+(\mathcal{N}_k-N_1-1)L+1:N_1+(\mathcal{N}_k-N_1)L]}$ consists of L consecutive columns from $\mathbf{G}$ and is non-singular as well.

Now all correctness constraints are satisfied. We are left to show that the privacy constraint (11) is satisfied. To this end, we show that regardless of the vale of the desired message index k, the answers are uniformly random, i.e.,

$$H(\mathbf{A}^{[k]}) = \overline{K} = H(W_k, Z). \quad (90)$$

From (80), it is equivalent to show that

(Equivalent Privacy Condition):

$$\mathbf{B}^{[k]}_{\overline{K}\times\overline{K}} = [\mathbf{F}^{[k]} \quad \mathbf{H}], \ \forall k \in [1:K] \text{ have full rank.} \quad (91)$$

First, consider the case where $\overline{k} \in [1:N_1]$. From (82), we know that $N-L$ cyclicly consecutive rows (where the row index does not belong to the set $\mathcal{N}_k$) of $\mathbf{F}^{[k]}$ are the zero vectors. It follows from the determinant formula of a 2×2 block matrix with a zero sub-block that

$$\det(\mathbf{B}^{[k]}) = \det(\mathbf{F}^{[k]}_{[\overrightarrow{\mathcal{N}}_k,:]})\det(\mathbf{H}_{[(1:\overline{K})\setminus\overrightarrow{\mathcal{N}}_k,:]}). \quad (92)$$

We have shown that $\mathbf{F}^{[k]}_{[\overrightarrow{\mathcal{N}}_k,:]}$ is non-singular. Further $|\mathbf{H}_{[(1:\overline{K})\setminus\overrightarrow{\mathcal{N}}_k,:]}|$ is equal to the determinant of a square sub-matrix of a Cauchy matrix (and is another Cauchy matrix) so that $\mathbf{H}_{[(1:\overline{K})\setminus\overrightarrow{\mathcal{N}}_k,:]}$ is non-singular as well.

Second, consider the case where $\overline{k} \in [N_1+1:\overline{K}]$. The proof is similar to that above, where the non-zero part of the $\mathbf{F}^{[k]}$ component is a non-singular square matrix (refer to (89)) and the corresponding sub-matrix of the $\mathbf{H}$ component in the determinant formula (refer to (92)) has a determinant that is given by a square sub-matrix of a Cauchy matrix (thus non-singular as well). Therefore, $\mathbf{B}^{[k]}$ always have full rank and the scheme is private.

Finally, we note that $\overline{K}-L$ randomness symbols are used to send L message symbols. The randomness size is then $\eta = H(Z)/L = (\overline{K}-L)/L = (K-M)/M = 1/R-1$.

Remark: An interesting observation of our scheme is that it is automatically secure, i.e., from the answers for W_k, the user learns absolutely no information about other messages. This indicates that the undelivered messages do not play a role in keeping privacy (the common randomness is responsible for privacy).

B. $N = \lceil \frac{K}{M} \rceil$: Optimality of $R = 1/\lceil \frac{K}{M} \rceil$

We first show that when $N = \lceil \frac{K}{M} \rceil$, each server must contain a message that appears only in that server (not available in any other servers). This result is stated in the following lemma.

Lemma 3: When $N = \lceil \frac{K}{M} \rceil$, the following property holds.

(Property 1) For any $i \in [1:N]$, there exists a message $W_{k_i} \in S_i$, and $W_{k_i} \notin S_j, \forall j \neq i, j \in [1:N]$.

Proof: To set up the proof by contradiction, suppose there exists 1 server (say Server n) where every stored message appears at some other server. As each server stores M messages, we know that the M messages stored at Server n are replicated at least twice. As a result, the total storage required at all N servers is at least $K+M$. However, this is not possible because $K+M$ exceeds the total storage capability of the servers, MN.

$$MN = M \times \lceil \frac{K}{M} \rceil \quad (93)$$
$$< M \times \left(\frac{K}{M}+1\right) = K+M. \quad (94)$$

So we have proved that Property 1 is satisfied. ■

According to Lemma 3, consider the N messages $W_{k_1}, \cdots, W_{k_N}$, where each of them is available at only 1 (distinct) server. Using Lemma 1 for W_{k_i}, we have

$$\mathcal{N}_{k_i} = \{i\}: \quad D_i \geq L. \quad (95)$$

Adding (95) for all $i \in [1:N]$, we have

$$D_1 + \cdots + D_N \geq NL \quad (96)$$
$$\Rightarrow R = \frac{L}{\sum_{n=1}^{N} D_n} \leq 1/N = 1/\lceil \frac{K}{M} \rceil \quad (97)$$

and the proof is complete.

VI. Proof of Theorem 3

The achievable scheme is similar to that presented in Section V-A2 (albeit with a different set of parameters). Here we present the code construction succinctly and only highlight the differences.

We show that for K messages, M messages per server, and N servers, the following rate is achievable.

$$R = \frac{l}{N + (l-1)(\lfloor \frac{K}{M} \rfloor - 1)},\ l = \lfloor \frac{(N - \lfloor \frac{K}{M} \rfloor + 1)M}{K - (\lfloor \frac{K}{M} \rfloor - 1)M} \rfloor. \tag{98}$$

The N servers and K messages are similarly divided into 2 sets. The first server set is made up of the first $N_1 = N - N_2$ servers and the second server set is made up of the last $N_2 = \lfloor \frac{K}{M} \rfloor - 1$ servers. The first message set is comprised of the first $K_1 = K - K_2$ messages and the second message set is comprised of the last $K_2 = N_2\ M$ messages.

Suppose the message size $L = l$ symbols. Define $D_\Sigma \triangleq N_1 + LN_2$. Each message symbol is from $\mathbb{F}_p$, where $p \geq D_\Sigma$. The common random variable Z consists of $D_\Sigma - L$ i.i.d. symbols, each from the same field $\mathbb{F}_p$.

The storage is designed as follows. The first (second) message set is stored over the first (second) server set. Consider the first server set, where the N_1 servers can store $N_1\ M$ messages. Note that there are K_1 messages in the first message set so that at least, each of these K_1 messages can be stored $l = \lfloor N_1 M / K_1 \rfloor$ times (refer to (98)). Imagine these $N_1\ M$ locations as an $N_1 \times M$ table with N_1 rows and M columns. Consider the $N_1 M$ locations of the table in a greedily manner, first from the first row to the last row and then from the first column to the last column, and we throw the $K_1\ l$ messages (from W_1 to W_{K_1}, each message replicated l times) into the locations in the order specified. The desired property of this storage strategy is that each message $W_i, i \in [1 : K-1]$ is available at l cyclicly consecutive servers in the first server set. Denote the availability set of W_k as $\mathcal{N}_k$. In the second server set, each server stores L distinct messages from the second message set sequentially.

For instance, consider the setting in Example 1, where $M = 3, K = 7, N = 4$. Then $N_2 = 1, N_1 = 3, K_2 = 3, K_1 = 4$ and the storage design is as follows.

$$S_1 = \{W_1, W_2, W_4\} \tag{99}$$
$$S_2 = \{W_1, W_3, W_4\} \tag{100}$$
$$S_3 = \{W_2, W_3\} \tag{101}$$
$$S_4 = \{W_5, W_6, W_7\}. \tag{102}$$

Suppose $W_k, k \in [1 : K]$ is desired. In the linear delivery scheme, each answer from the first server set has $D_i = 1, \forall i \in [1 : N_1]$ symbol, and each answer from the second server set has $D_i = L, \forall i \in [N_1 + 1 : N]$ symbols. Then the rate achieved is $R = \frac{L}{N_1 + LN_2} = \frac{l}{N + (l-1)N_2}$, as desired (refer to (98)).

The answers are shown below.

$$A_i^{[k]} = \mathbf{F}_i^{[k]} W_k + \mathbf{H}_i Z \tag{103}$$

where

$$\begin{aligned} &\text{if } i \in [1 : N_1],\ \mathbf{F}_i^{[k]} \text{ is a } 1 \times L \text{ vector,} \\ &\qquad \text{and } \mathbf{H}_i \text{ is a } 1 \times (D_\Sigma - L) \text{ vector,} \\ &\text{else } i \in [N_1 + 1 : N],\ \mathbf{F}_i^{[k]} \text{ is an } L \times L \text{ matrix,} \\ &\qquad \text{and } \mathbf{H}_i \text{ is an } L \times (D_\Sigma - L) \text{ matrix.} \end{aligned} \tag{104}$$

Then the collection of all answers are as follows.

$$\mathbf{A}^{[k]} = \mathbf{F}_{D_\Sigma \times L}^{[k]} W_k + \mathbf{H}_{D_\Sigma \times (D_\Sigma - L)} Z, \tag{105}$$

$$\text{where} \quad \mathbf{F}^{[k]} = [\mathbf{F}_1^{[k]}; \mathbf{F}_2^{[k]}; \cdots; \mathbf{F}_N^{[k]}],$$
$$\mathbf{H} = [\mathbf{H}_1; \mathbf{H}_2; \cdots; \mathbf{H}_N]. \tag{106}$$

The decoding filtering matrix is denoted by $\mathbf{G}_{L \times D_\Sigma}$. Then we have

$$W_k = \mathbf{G}\mathbf{A}^{[k]} = \mathbf{G}\mathbf{F}^{[k]} W_k + \mathbf{G}\mathbf{H} Z, \tag{107}$$

$$\begin{aligned} &\Rightarrow \mathbf{G}\mathbf{F}^{[k]} = \mathbf{I}_L \Rightarrow \\ &\quad \text{If } k \in [1 : K_1],\ \mathbf{G}_{[:, \overrightarrow{\mathcal{N}}_k]} \mathbf{F}_{[\overrightarrow{\mathcal{N}}_k, :]}^{[k]} = \mathbf{I}_L, \\ &\quad \text{else } k \in [K_1 + 1 : K], \\ &\qquad \mathbf{G}_{[:, N_1 + (\mathcal{N}_k - N_1 - 1)L + 1 : N_1 + (\mathcal{N}_k - N_1)L]} \mathbf{F}_{\mathcal{N}_k}^{[k]} = \mathbf{I}_L, \end{aligned} \tag{108}$$

$$\mathbf{G}\mathbf{H} = \mathbf{0}_{L \times (D_\Sigma - L)} \tag{109}$$

and all other unspecified sub-matrices of $\mathbf{F}^{[k]}$ are zero matrices, due to the storage constraint.

To satisfy (108), (109), we set $\mathbf{G}, \mathbf{F}^{[1]}, \mathbf{F}^{[2]}, \cdots, \mathbf{F}^{[K]}, \mathbf{H}$ as follows.

$$\begin{aligned} &\mathbf{G} = \begin{bmatrix} \mathbf{I}_L & \mathbf{V}_{L \times (D_\Sigma - L)} \end{bmatrix}, \text{ where } \mathbf{V} \text{ is a Cauchy} \\ &\quad \text{matrix such that } V_{ij} = \tfrac{1}{\alpha_i - \beta_j}, \alpha_i \neq \beta_j, \\ &\mathbf{H} = \begin{bmatrix} \mathbf{V}_{L \times (D_\Sigma - L)}; -\mathbf{I}_{D_\Sigma - L} \end{bmatrix}, \end{aligned} \tag{110}$$

$$\begin{aligned} &\mathbf{F}_{[\overrightarrow{\mathcal{N}}_k, :]}^{[k]} = \mathbf{G}_{[:, \overrightarrow{\mathcal{N}}_k]}^{-1}, \text{ if } \overline{k} \in [1 : K_1], \\ &\mathbf{F}_{\mathcal{N}_k}^{[k]} = \mathbf{G}_{[:, N_1 + (\mathcal{N}_k - N_1 - 1)L + 1 : N_1 + (\mathcal{N}_k - N_1)L]}^{-1}, \\ &\quad \text{else } \overline{k} \in [K_1 + 1 : K]. \end{aligned} \tag{111}$$

By the same reasoning as that in Section V-A2, the matrices in (111) have full rank so that their inverse matrices are well defined. Now correctness constraints are satisfied, and privacy is guaranteed by the observation that

$$H(\mathbf{A}^{[k]}) = D_\Sigma = H(W_k, Z) \Longleftrightarrow \tag{112}$$

$$\mathbf{B}_{D_\Sigma \times D_\Sigma}^{[k]} = [\mathbf{F}^{[k]} \quad \mathbf{H}],\ \forall k \in [1 : K] \text{ have full rank.} \tag{113}$$

The proof for $\mathbf{B}^{[k]}$ being full rank follows similarly from that in Section V-A2 and the details are thus omitted.

VII. Optimality of Achievable Schemes for Examples 1 and 2

We present the proof for Example 2 first because it is simpler. The proof idea for both examples is the same - we consider all possible storage strategies and show that none of them may outperform the achieved rate. For each storage strategy, we argue that certain combinatoric structure must

exist and the structure leads to a rate upper bound (using Lemma 1).[3]

A. Example 2. $M = 4, K = 5$

We have two settings to consider, i.e., $N = 3$ and $N = 4$.

1) $N = 3$: Proof of $R \leq 2/3$: We assume without loss of generality that each server stores $M = 4$ distinct messages (because storing more messages does not hurt). Then we have $K = 5$ messages and $MN = 12$ messages are stored across all servers. Denote $\mathcal{N}_k, k \in [1:5]$ as the set of servers where W_k is stored, so that $|\mathcal{N}_k|$ represents the number of servers where W_k is stored. We assume that $|\mathcal{N}_1| \geq |\mathcal{N}_2| \geq \cdots \geq |\mathcal{N}_5|$. Therefore, we have a partition of the total storage of 12 messages.

$$12 = |\mathcal{N}_1| + |\mathcal{N}_2| + |\mathcal{N}_3| + |\mathcal{N}_4| + |\mathcal{N}_5|, |\mathcal{N}_k| \in [1:3]. \quad (114)$$

Note that $|\mathcal{N}_k| \geq 1$ because all messages must be stored somewhere and $|\mathcal{N}_k| \leq 3$ because we only have $N = 3$ servers. Because of the range of $|\mathcal{N}_k|$ and the assumption of the monotonic non-increasing property on the $\mathcal{N}_k$ sequence, we only have the following 2 cases.

$$\text{Case 1: } (|\mathcal{N}_1|, |\mathcal{N}_2|, |\mathcal{N}_3|, |\mathcal{N}_4|, |\mathcal{N}_5|) = (3, 3, 3, 2, 1), \quad (115)$$

$$\text{Case 2: } (|\mathcal{N}_1|, |\mathcal{N}_2|, |\mathcal{N}_3|, |\mathcal{N}_4|, |\mathcal{N}_5|) = (3, 3, 2, 2, 2). \quad (116)$$

For both cases, the storage design is deterministic (up to permutations of the servers). Denote (π_1, π_2, π_3) as a permutation of the 3 servers $(1, 2, 3)$. For Case 1, we have

$$\begin{aligned} S_{\pi_1} &= \{W_1, W_2, W_3, W_4\} && (117)\\ S_{\pi_2} &= \{W_1, W_2, W_3, W_4\} && (118)\\ S_{\pi_3} &= \{W_1, W_2, W_3, W_5\}. && (119) \end{aligned}$$

Using Lemma 1 for W_4 and W_5, we have

$$\begin{aligned} D_{\pi_1} + D_{\pi_2} &\geq L && (120)\\ D_{\pi_3} &\geq L && (121)\\ (120) + (121) \Rightarrow D_1 + D_2 + D_3 &\geq 2L && (122)\\ \Rightarrow R = \frac{L}{D_1 + D_2 + D_3} &\leq 1/2 < 2/3. && (123) \end{aligned}$$

For Case 2, we have

$$\begin{aligned} S_{\pi_1} &= \{W_1, W_2, W_3, W_4\} && (124)\\ S_{\pi_2} &= \{W_1, W_2, W_3, W_5\} && (125)\\ S_{\pi_3} &= \{W_1, W_2, W_4, W_5\}. && (126) \end{aligned}$$

Using Lemma 1 for W_3, W_4 and W_5, we have

$$\begin{aligned} D_{\pi_1} + D_{\pi_2} &\geq L && (127)\\ D_{\pi_3} + D_{\pi_1} &\geq L && (128)\\ D_{\pi_2} + D_{\pi_3} &\geq L && (129)\\ (127) + (128) + (129) &\Rightarrow \\ 2(D_1 + D_2 + D_3) &\geq 3L && (130)\\ \Rightarrow R = \frac{L}{D_1 + D_2 + D_3} &\leq 2/3. && (131) \end{aligned}$$

[3]Our proof is brute-force based in essence. This is the reason that we are not able to generalize this converse proof (for which a more algorithmic approach, e.g., linear programming, on using Lemma 1 might be helpful). However, we are not aware of any setting where the best achievable rate given by Theorem 3 is not optimal.

Therefore, for both cases, the rate cannot be higher than $2/3$ so that the achieved rate of $R = 2/3$ is optimal.

2) $N = 4$: Proof of $R \leq 3/4$: The proof idea is similar. We consider a partition of the total storage of $MN = 16$ messages to the $K = 5$ messages.

$$\begin{aligned} 16 &= |\mathcal{N}_1| + |\mathcal{N}_2| + |\mathcal{N}_3| + |\mathcal{N}_4| + |\mathcal{N}_5|, \\ |\mathcal{N}_1| &\geq \cdots \geq |\mathcal{N}_5|, |\mathcal{N}_k| \in [1:4], \forall k \in [1:5]. \quad (132) \end{aligned}$$

For the partition, we have the following 4 cases.

1) $(|\mathcal{N}_1|, |\mathcal{N}_2|, |\mathcal{N}_3|, |\mathcal{N}_4|, |\mathcal{N}_5|) = (4, 4, 4, 3, 1)$.
In this case, W_4 and W_5 are stored over 2 disjoint sets of servers. Using Lemma 1, we have $\sum_{i=1}^{4} D_i \geq 2L$ so that $R \leq 1/2 < 3/4$.
2) $(|\mathcal{N}_1|, |\mathcal{N}_2|, |\mathcal{N}_3|, |\mathcal{N}_4|, |\mathcal{N}_5|) = (4, 4, 4, 2, 2)$.
This case is similar to that above, where W_4 and W_5 are stored over 2 disjoint sets of servers. Then $R \leq 1/2$ follows.
3) $(|\mathcal{N}_1|, |\mathcal{N}_2|, |\mathcal{N}_3|, |\mathcal{N}_4|, |\mathcal{N}_5|) = (4, 4, 3, 3, 2)$.
The storage design is deterministic (up to permutation of the servers). Denote $(\pi_1, \pi_2, \pi_3, \pi_4)$ as a permutation of the 4 servers $(1, 2, 3, 4)$. We have

$$\begin{aligned} S_{\pi_1} &= \{W_1, W_2, W_3, W_4\} && (133)\\ S_{\pi_2} &= \{W_1, W_2, W_3, W_4\} && (134)\\ S_{\pi_3} &= \{W_1, W_2, W_3, W_5\} && (135)\\ S_{\pi_4} &= \{W_1, W_2, W_4, W_5\}. && (136) \end{aligned}$$

Using Lemma 1 for W_3, W_4 and W_5, we have

$$\begin{aligned} D_{\pi_1} + D_{\pi_2} + D_{\pi_3} &\geq L && (137)\\ D_{\pi_4} + D_{\pi_1} + D_{\pi_2} &\geq L && (138)\\ D_{\pi_3} + D_{\pi_4} &\geq L && (139)\\ \Rightarrow 2(D_1 + D_2 + D_3 + D_4) &\geq 3L && (140)\\ \Rightarrow R = \frac{L}{D_1 + D_2 + D_3 + D_4} &\leq 2/3 < 3/4. && (141) \end{aligned}$$

4) $(|\mathcal{N}_1|, |\mathcal{N}_2|, |\mathcal{N}_3|, |\mathcal{N}_4|, |\mathcal{N}_5|) = (4, 3, 3, 3, 3)$. The storage is also deterministic. We have

$$\begin{aligned} S_{\pi_1} &= \{W_1, W_2, W_3, W_4\} && (142)\\ S_{\pi_2} &= \{W_1, W_2, W_3, W_5\} && (143)\\ S_{\pi_3} &= \{W_1, W_2, W_4, W_5\} && (144)\\ S_{\pi_4} &= \{W_1, W_3, W_4, W_5\}. && (145) \end{aligned}$$

Using Lemma 1 for W_2, W_3, W_4 and W_5, we have

$$\begin{aligned} & D_{\pi_{i_1}} + D_{\pi_{i_2}} + D_{\pi_{i_3}} \geq L, \\ & \forall \text{ distinct } i_1, i_2, i_3 \in [1:4] && (146)\\ \Rightarrow\ & D_1 + D_2 + D_3 + D_4 \geq 4L/3 && (147)\\ \Rightarrow\ & R = \frac{L}{D_1 + D_2 + D_3 + D_4} \leq 3/4. && (148) \end{aligned}$$

All cases are covered and we always have $R \leq 3/4$. The proof is thus complete.

B. Example 1. $M = 3, K = 7, N = 4$ and Proof of $R \leq 2/5$

We follow the same proof idea presented in the previous section for Example 2.

Consider a partition of the total storage of $MN = 12$ messages to the $K = 7$ messages.

$$16 = |\mathcal{N}_1| + |\mathcal{N}_2| + \cdots + |\mathcal{N}_7|,$$
$$|\mathcal{N}_1| \geq \cdots \geq |\mathcal{N}_7|, |\mathcal{N}_k| \in [1:4], \forall k \in [1:7]. \quad (149)$$

For the partition, we have the following 5 cases.

1) $(|\mathcal{N}_1|, |\mathcal{N}_2|, \cdots, |\mathcal{N}_7|) = (4, 3, 1, 1, 1, 1, 1)$.
In this case, 4 out of the 5 messages W_3, W_4, W_5, W_6, W_7 (each appeares once) are stored over 4 disjoint sets of servers. Using Lemma 1, we have $\sum_{i=1}^{4} D_i \geq 4L$ so that $R \leq 1/4 < 2/5$.
2) $(|\mathcal{N}_1|, |\mathcal{N}_2|, \cdots, |\mathcal{N}_7|) = (4, 2, 2, 1, 1, 1, 1)$.
In this case, 3 out of the 6 messages $W_2, W_3, W_4, W_5, W_6, W_7$ are stored over 3 disjoint sets of servers. Using Lemma 1, we have $\sum_{i=1}^{4} D_i \geq 3L$ so that $R \leq 1/3 < 2/5$.
3) $(|\mathcal{N}_1|, |\mathcal{N}_2|, \cdots, |\mathcal{N}_7|) = (3, 3, 2, 1, 1, 1, 1)$.
Consider the last 4 messages W_4, W_5, W_6, W_7 (each appears once). If these 4 messages appear in 3 servers, then similar as the case above, we have $R \leq 1/3$. Henceforth, we focus on the setting where these 4 messages appear in 2 servers. The allocation of these 4 messages to the 2 servers might be $3 + 1$ or $2 + 2$. It is easy to see that for both settings, W_3 must appear in the other 2 remaining servers so that we have 3 messages that appear in 3 disjoint sets of servers, i.e., $R \leq 1/3$.
4) $(|\mathcal{N}_1|, |\mathcal{N}_2|, \cdots, |\mathcal{N}_7|) = (3, 2, 2, 2, 1, 1, 1)$.
Consider the last 3 messages W_5, W_6, W_7 (each appears once). If these 3 messages appear in 3 servers, then similar as the case above, we have $R \leq 1/3$. Henceforth, we focus on the setting where these 3 messages appear in 1 server or 2 servers.
When W_5, W_6, W_7 appear in 1 server, the storage is deterministic. We have

$$S_{\pi_1} = \{W_1, W_2, W_3\} \quad (150)$$
$$S_{\pi_2} = \{W_1, W_2, W_4\} \quad (151)$$
$$S_{\pi_3} = \{W_1, W_3, W_4\} \quad (152)$$
$$S_{\pi_4} = \{W_5, W_6, W_7\}. \quad (153)$$

Using Lemma 1 for W_2, W_3, W_4 and W_5, we have

$$D_{\pi_{i_1}} + D_{\pi_{i_2}} \geq L, \ \forall \text{distinct } i_1, i_2 \in [1:3] \quad (154)$$
$$D_{\pi_4} \geq L, \quad (155)$$
$$\Rightarrow \sum_{i=1}^{4} D_i \geq 5L/2, \ R = \frac{L}{\sum_{i=1}^{4} D_i} \leq 2/5. \quad (156)$$

When W_5, W_6, W_7 appear in 2 servers, we have

$$S_{\pi_1} = \{\times, \times, \times\} \quad (157)$$
$$S_{\pi_2} = \{\times, \times, \times\} \quad (158)$$
$$S_{\pi_3} = \{\times, \times, W_5\} \quad (159)$$
$$S_{\pi_4} = \{\times, W_6, W_7\} \quad (160)$$

where $\times$ represents place-holders for the remaining messages, W_1 (will appear 3 times), W_2, W_3, W_4 (will appear 2 times each). By enumerating all possibilities, it is easy to see that there exists 1 message out of W_2, W_3, W_4 that appears only in S_{π_1}, S_{π_2}. Combining this message with W_5, W_6, we have 3 messages that appear in 3 disjoint sets of servers and it follows that $R \leq 1/3$.
5) $(|\mathcal{N}_1|, |\mathcal{N}_2|, \cdots, |\mathcal{N}_7|) = (2, 2, 2, 2, 2, 1, 1)$.
We have 2 possibilities here, depending on how many servers will be occupied by the last 2 messages.
When W_6, W_7 (each appears once) appear in 2 servers, there must exist 1 message out of W_1, W_2, W_3, W_4, W_5 that appears only in the 2 remaining servers. Similarly, we have 3 messages that appear in 3 disjoint sets of servers and $R \leq 1/3$.
When W_6, W_7 (each appears once) appear in 1 server, we have (denote $(\gamma_1, \cdots, \gamma_5)$ as a permutation of $(1, \cdots, 5)$)

$$S_{\pi_1} = \{\times, \times, \times\} \quad (161)$$
$$S_{\pi_2} = \{\times, \times, \times\} \quad (162)$$
$$S_{\pi_3} = \{\times, \times, W_{\gamma_5}\} \quad (163)$$
$$S_{\pi_4} = \{W_{\gamma_5}, W_6, W_7\} \quad (164)$$

where $\times$ represents place-holders for $W_{\gamma_1}, \cdots, W_{\gamma_4}$ (each appears twice). By a similar reasoning as that in the above case, we must have 1 message out of $W_{\gamma_1}, \cdots, W_{\gamma_4}$ that appears only in S_{π_1}, S_{π_2}. Therefore, 3 messages appear in 3 disjoint sets of servers and $R \leq 1/3$.

To summarize, no matter how we design the storage, the rate is always bounded above by $2/5$ so that the proof is complete.

VIII. Discussion

Motivated by dataset privacy, we introduce the problem of private information delivery, where one out of K messages is sent from a set of servers to a user while the delivered message index remains a secret. We take an information theoretic approach to this problem and adopt the capacity as the performance metric (parallel to the recent line of private information retrieval [4]–[7], where the privacy of the user is considered). We propose information theoretic converses that capture this privacy constraint and vector linear coding schemes that satisfy perfect privacy. The rate upper and lower bounds are tight for a wide range of system parameters. We consider the elemental model where the messages are replicated, the user behaves nicely, and a single message is delivered, leaving much room for generalizations.

We have focused exclusively on the metric of rate while the amount of randomness is ignored. The interplay between the communicate rate and the randomness size is an interesting future direction. Further, we are taking a coarse look at the randomness as we assume the same random variable is shared by all servers. It is not hard to see that this is not necessary and we only need the randomness variables to be correlated. A finer view on the rate region of the correlated randomness

variables (instead of the sum randomness rate) will shed light on the consumption of randomness.

In addition, we mention the connection of the private information delivery problem to the anonymous communications problem [3], [8]–[10], where the identity that needs to be hidden is the transmitters, receivers and their associations. Under many circumstances (e.g., [3]), the identity of the delivered message in private information delivery is intimately related to the identity of the nodes in an anonymous communication network. As a result, it is interesting to explore the implications and extensions of the techniques in this work to anonymous communication networks.

Finally, we put our work in the broader context of using information theory tools to model security and privacy primitives and to analyze their fundamental limits. This work presents a model that captures the privacy of dataset delivered for one user and other recent efforts include privacy from the user against the databases [4]–[7] (along the line of private information retrieval), privacy among multiple users [11]–[14] (along the line of caching systems with demand privacy), and new data related models that are increasingly important in the modern information era (see e.g., [15]–[17] and references therein).

REFERENCES

[1] H. Sun, "Private information delivery," 2018, *arXiv:1806.05601*. [Online]. Available: http://arxiv.org/abs/1806.05601

[2] D. Chaum, "The dining cryptographers problem: Unconditional sender and recipient untraceability," *J. Cryptol.*, vol. 1, no. 1, pp. 65–75, Jan. 1988.

[3] H. Sun, "The capacity of anonymous communications," *IEEE Trans. Inf. Theory*, vol. 65, no. 6, pp. 3871–3879, Jun. 2019.

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

[5] 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.

[6] 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.

[7] 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.

[8] K. Peng, *Anonymous Communication Networks: Protecting Privacy on the Web*. Boca Raton, FL, USA: CRC Press, 2014.

[9] J. Ren and J. Wu, "Survey on anonymous communications in computer networks," *Comput. Commun.*, vol. 33, no. 4, pp. 420–431, Mar. 2010.

[10] G. Danezis and C. Diaz, "A survey of anonymous communication channels," Microsoft Res., Cambridge, U.K., Tech. Rep. MSR-TR-2008-35, 2008.

[11] K. Wan and G. Caire, "On coded caching with private demands," 2019, *arXiv:1908.10821*. [Online]. Available: http://arxiv.org/abs/1908.10821

[12] V. R. Aravind, P. K. Sarvepalli, and A. Thangaraj, "Subpacketization in coded caching with demand privacy," in *Proc. Nat. Conf. Commun. (NCC)*, Feb. 2020, pp. 1–6.

[13] S. Kamath, J. Ravi, and B. K. Dey, "Demand-private coded caching and the exact trade-off for N=K=2," in *Proc. Nat. Conf. Commun. (NCC)*, Feb. 2020, pp. 1–6.

[14] K. Wan, H. Sun, M. Ji, D. Tuninetti, and G. Caire, "Fundamental limits of device-to-device private caching with trusted server," 2019, *arXiv:1912.09985*. [Online]. Available: http://arxiv.org/abs/1912.09985

[15] Z. Wang, K. Banawan, and S. Ulukus, "Private set intersection: A multi-message symmetric private information retrieval perspective," 2020, *arXiv:1912.13501*. [Online]. Available: http://arxiv.org/abs/1912.13501

[16] B. Tahmasebi and M. Ali Maddah-Ali, "Private sequential function computation," 2019, *arXiv:1908.01204*. [Online]. Available: http://arxiv.org/abs/1908.01204

[17] H. Sun, "Secure groupcast with shared keys," 2020, *arXiv:2003.11995*. [Online]. Available: http://arxiv.org/abs/2003.11995

Hua Sun (Member, IEEE) received the B.E. degree in communications engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 2011, and the M.S. degree in electrical and computer engineering and the Ph.D. degree in electrical engineering from the University of California at Irvine, USA, in 2013 and 2017, respectively.

He is currently an Assistant Professor with the Department of Electrical Engineering, University of North Texas, USA. His research interest includes information theory and its applications to communications, privacy, security, and storage. He received the IEEE Jack Keil Wolf ISIT Student Paper Award in 2016, the IEEE GLOBECOM Best Paper Award in 2016, and the University of California at Irvine CPCC Fellowship for the year 2011–2012.