Asymmetry Hurts: Private Information Retrieval Under Asymmetric Traffic Constraints

Karim Banawan, *Member, IEEE*, and Sennur Ulukus, *Fellow, IEEE*

***Abstract*— We consider the classical setting of private information retrieval (PIR) of a single message (file) out of M messages from N distributed databases under the new constraint of *asymmetric traffic* from databases. In this problem, the *ratios between the traffic* from the databases are constrained, i.e., the ratio of the length of the answer string that the user (retriever) receives from the nth database to the total length of all answer strings from all databases is constrained to be τ_n. This may happen if the user's access to the databases is restricted due to database availability, channel quality to the databases, and other factors. For this problem, for fixed M, N, we develop a general upper bound $\bar{C}(\tau)$, which generalizes the converse proof of Sun-Jafar, where database symmetry was inherently used. Our converse bound is a piece-wise affine function in the traffic ratio vector $\tau = (\tau_1, \cdots, \tau_N)$. For the lower bound, we explicitly show the achievability of $\binom{M+N-1}{M}$ corner points. For the remaining traffic ratio vectors, we perform time-sharing between these corner points. The recursive structure of our achievability scheme is captured via a system of difference equations. The upper and lower bounds exactly match for $M = 2$ and $M = 3$ for any N and any τ. The results show strict loss of PIR capacity due to the asymmetric traffic constraints compared with the symmetric case of Sun-Jafar which implicitly uses $\tau_n = \frac{1}{N}$ for all n.**

***Index Terms*— Private information retrieval, asymmetric traffic constraints, database access constraints, capacity.**

I. Introduction

PROTECTING the privacy of downloaded information from curious publicly accessible databases has been the focus of considerable research within the computer science community [1]–[5]. The problem of privacy has become even more relevant today in the presence of efficient data-mining techniques. Private information retrieval (PIR), introduced by Chor *et al.* in [1], studies the privacy of the downloaded content from public databases. In the classical PIR setting, a user requests to download a certain message (or file) out of M distinct messages from N non-communicating (non-colluding) databases without leaking the identity of the desired message to any individual database. The contents of these databases are identical. The user prepares N queries, one for each database, such that the queries do not reveal the user's interest in the desired message. Upon receiving these queries, each database responds truthfully with an answer string. The user needs to be able to reconstruct the entire message by decoding the answer strings from all databases. PIR schemes are designed to be more efficient than the trivial scheme of downloading all the files stored in the databases. The efficiency of a retrieval scheme is measured by the retrieval rate, which is the ratio of the number of decodable desired message symbols to the number of total downloaded symbols.

Recently, the PIR problem is revisited by information theorists [6]–[11]. The information-theoretic reformulation of the problem assumes that the messages are of arbitrarily large size and hence the upload cost can be neglected with respect to the download cost [8] in contrast to the computer science formulation. This formulation provides an absolute guarantee (as opposed to computational PIR, e.g., [3], [5]). In the leading work [12], Sun and Jafar introduce the PIR capacity notion to characterize the fundamental limits of the PIR problem. The PIR capacity is defined as the supremum of PIR rates over all achievable retrieval schemes. [12] determines the exact capacity of the classical PIR to be $C = (1 + \frac{1}{N} + \frac{1}{N^2} + \cdots + \frac{1}{N^{M-1}})^{-1}$. Following the work of [12], the fundamental limits of many interesting variants of the classical PIR problem have been considered, such as: PIR from colluding databases, robust PIR, symmetric PIR, PIR from MDS-coded databases, PIR for arbitrary message lengths, multi-round PIR, multi-message PIR, PIR from Byzantine databases, secure symmetric PIR with adversaries, cache-aided PIR, PIR with private side information (PSI), PIR for functions, storage constrained PIR, and their several combinations [13]–[36].

A common property of the achievability schemes constructed for these PIR problems is that they exhibit a *symmetric structure* across the databases. In most existing PIR schemes, the user retrieves pieces of the desired message from all databases, and generates and uses side information at all databases in a symmetric manner. This enables the user to balance the load of retrieval of the desired message equally among the databases, and re-use the side information

Manuscript received January 8, 2018; revised February 9, 2019; accepted May 29, 2019. Date of publication August 2, 2019; date of current version October 18, 2019. This work was supported by the NSF under Grant CNS 13-14733, Grant CCF 14-22111, Grant CNS 15-26608, and Grant CCF 17-13977. This article was presented in part at the 2018 IEEE International Symposium on Information Theory.

K. Banawan was with the Department of Electrical and Computer Engineering, University of Maryland at College Park, College Park, MD 20742 USA. He is now with the Electrical Engineering Department, Faculty of Engineering, Alexandria University, Alexandria 21544, Egypt (e-mail: kbanawan@alexu.edu.eg).

S. Ulukus is with the Department of Electrical and Computer Engineering, University of Maryland at College Park, College Park, MD 20742 USA (e-mail: ulukus@umd.edu).

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

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.2019.2933011

generated from one database equally in all the remaining databases. Now, consider the following scenarios that render symmetry assumption unworkable: *Varying database availability:* Certain databases are available only a fraction of the time other databases are available for downloads. *Different capacities:* The capacities of the links (bit pipes) from the databases to the user have different capacities. This may be due to different physical locations of the databases, e.g., the user may be able to access physically closer databases more often than physically distant databases, or it may be due to the quality of the physical layer communication channel, e.g., the bandwidths (rates) of the download channels may be different for different databases. In these cases, the user is forced to deal with each database differently, i.e., the user should utilize the databases which have better quality links more often than the other databases. This breaks the database symmetry assumption, makes load balancing of desired message and side information more challenging, and poses the following interesting questions: Can we perform efficient PIR without applying database symmetry? Is there a fundamental PIR rate loss due to not being able to use symmetric schemes?

Motivated by these practical scenarios, we consider the PIR problem under *asymmetric traffic constraints*. Formally, we consider a classical PIR setting with N replicated and non-communicating databases storing M messages. We assume that the nth database responds with a t_n-length answer string. We constrain the lengths of the answer strings such that $t_n = \lambda_n t_1$ for $n \in \{2, \cdots, N\}$. This, in turn, forces the ratios between the traffic from the databases to be $1 : \lambda_2 : \lambda_3 : \cdots : \lambda_N$. We denote the traffic ratio with respect to the total download by a vector $\boldsymbol{\tau} = (\tau_1, \cdots, \tau_N)$, where $\tau_n = \frac{\lambda_n}{\sum_{j=1}^{N} \lambda_j}$. We aim at characterizing the capacity of this PIR problem, $C(\boldsymbol{\tau})$, as a function of the given traffic ratio vector $\boldsymbol{\tau}$ for arbitrary M and N. We note that in this problem, we do not constrain t_1 itself, but rather constrain the ratios between the responses according to $\boldsymbol{\tau}$; in fact, we assume that t_1 can grow arbitrarily large to conform with the classical information-theoretic formulation. Furthermore, we remark that although our problem seems to be related to the *upload-constrained* PIR problem [12], we note that the upload-constrained problem investigates the *minimum* possible query size if the user and the databases exchange a codebook prior to the retrieval process, while in the asymmetric traffic constrained problem here we do not assume the existence of a codebook, and hence we minimize the number of queries subject to an additional constraint on the traffic ratios.

In this paper, we investigate the fundamental limits of the PIR problem under asymmetric traffic constraints. To that end, we develop a novel upper bound for the capacity $\bar{C}(\boldsymbol{\tau})$. This generalizes the converse proof of [12] to incorporate the asymmetric traffic constraints. Originally, the proof in [12] exploits the database symmetry. The rationale is that even if the optimal scheme is not symmetric, we can transform it into a symmetric scheme without changing the retrieval rate by means of time-sharing [12]. In our case, we cannot use this technique as we must deal with the databases differently. We characterize the upper bound as a piece-wise affine function in $\boldsymbol{\tau}$ (see Theorem 1). The upper bound implies that asymmetry fundamentally hurts the retrieval rate (see Corollary 1 and Remark 4). Then, we propose explicit achievability schemes for $\binom{M+N-1}{M}$ corner points. Each corner point corresponds to a specific partitioning of the databases according to the number of side information symbols that are used simultaneously within the initial round of the download. We describe the achievability scheme via a system of difference equations in the number of stages at each round of the download (which is parallel to [22]). For any other traffic ratio vector $\boldsymbol{\tau}$, we employ time-sharing between the corner points that enclose $\boldsymbol{\tau}$. We provide an explicit rate expression for the case of $N = 2$ for arbitrary M. We show that the upper bound and the lower bound exactly match for the cases of $M = 2$ and $M = 3$ messages for any N and any $\boldsymbol{\tau}$, leading to the exact capacity $C(\boldsymbol{\tau})$ for these cases.

II. System Model

Consider a classical PIR model with N non-communicating and replicated databases storing M messages (or files). Each database stores the same set of messages $W_{1:M} = \{W_1, \cdots, W_M\}$. Messages $W_{1:M}$ are independent and identically distributed over all vectors of size L picked from a finite field $\mathbb{F}_q^L$, i.e.,

$$H(W_i) = L, \quad i \in \{1, \cdots, M\} \tag{1}$$

$$H(W_1, \cdots, W_M) = ML, \quad (q\text{-ary units}) \tag{2}$$

In the PIR problem, a user wants to retrieve a message $W_i \in W_{1:M}$ correctly without revealing any information about the identity of the message i to any individual database. To that end, the user submits a query $Q_n^{[i]}$ to the nth database. The messages and the queries are statistically independent due to the fact that the user does not know the message realizations in advance, i.e.,

$$I(W_{1:M}; Q_{1:N}^{[i]}) = 0 \tag{3}$$

where $Q_{1:N}^{[i]} = \{Q_1^{[i]}, \cdots, Q_N^{[i]}\}$. The nth database responds truthfully by an answer string $A_n^{[i]}$. The answer string $A_n^{[i]}$ is a deterministic function of the query $Q_n^{[i]}$ and all the messages $W_{1:M}$, hence

$$H(A_n^{[i]} | Q_n^{[i]}, W_{1:M}) = 0, \quad n \in \{1, \cdots, N\} \tag{4}$$

In the PIR model with asymmetric traffic constraints, the lengths of the answer strings are different (see Fig. 1). More specifically, we assume that the nth database responds with a t_n-length answer string, such that $t_n = \lambda_n t_1$, where λ_n is the ratio between the traffic from the nth database to the traffic from the first database. Without loss of generality, we assume that the first database has the highest traffic and the remaining databases are ordered descendingly in λ_n. Hence, $\{\lambda_n\}_{n=1}^{N}$ is a *non-increasing monotone* sequence with $\lambda_1 = 1$, and $\lambda_n \in [0, 1]$, i.e.,

$$H(A_n^{[i]}) \leq \lambda_n t_1, \quad i \in \{1, \cdots, M\}, \; n \in \{1, \cdots, N\} \tag{5}$$

where $1 \geq \lambda_2 \geq \cdots \geq \lambda_N$.

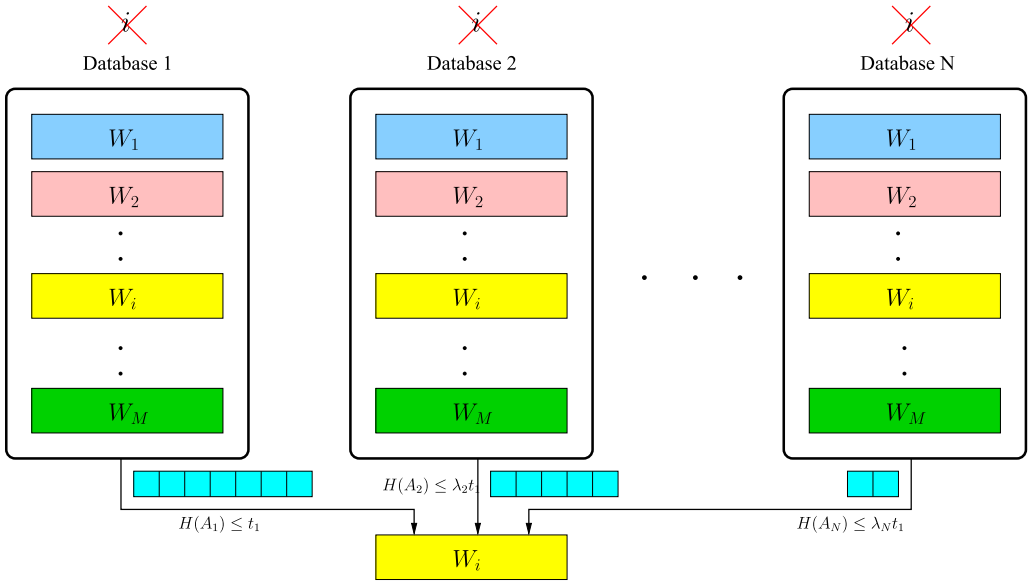

Fig. 1. PIR under asymmetric traffic constraints.

We define the *traffic ratio* of the nth database τ_n as the ratio between the traffic from the nth database and the total traffic from all databases, i.e.,

$$\tau_n = \frac{\lambda_n}{\sum_{j=1}^{N} \lambda_j} \tag{6}$$

We note that there is a one-to-one transformation between the vector $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \cdots, \lambda_N)$ and the vector $\boldsymbol{\tau} = (\tau_1, \tau_2, \cdots, \tau_N)$. Thus, $\boldsymbol{\lambda}$ and $\boldsymbol{\tau}$ are used interchangeably within the context of this paper.

In order to ensure the privacy, at the nth database, the query $Q_n^{[i]}$ designed to retrieve W_i should be indistinguishable from the queries designed to retrieve any other message, i.e.,

$$(Q_n^{[i]}, A_n^{[i]}, W_{1:M}) \sim (Q_n^{[j]}, A_n^{[j]}, W_{1:M}), \quad \forall j \in \{1, \cdots, M\} \tag{7}$$

where $\sim$ denotes statistical equivalence.

In addition, the user should be able to reconstruct W_i from the collected answer strings $A_{1:N}^{[i]}$ with arbitrarily small probability of error. By Fano's inequality, we have the following reliability constraint,

$$H(W_i | Q_{1:N}^{[i]}, A_{1:N}^{[i]}) = o(L) \tag{8}$$

where $\frac{o(L)}{L} \to 0$ as $L \to \infty$.

For a fixed N, M, and a traffic ratio vector $\boldsymbol{\tau}$, a retrieval rate $R(\boldsymbol{\tau})$ is achievable if there exists a PIR scheme which satisfies the privacy constraint (7) and the reliability constraint (8) for some message lengths $L(\boldsymbol{\tau})$ and answer strings of lengths $\{t_n(\boldsymbol{\tau})\}_{n=1}^{N}$ that satisfy the asymmetric traffic constraint (5), such that

$$R(\boldsymbol{\tau}) = \frac{L(\boldsymbol{\tau})}{\sum_{n=1}^{N} t_n(\boldsymbol{\tau})} \tag{9}$$

We note that in this problem, we do not constrain either the message length $L(\boldsymbol{\tau})$ or the lengths of the answer strings $t_n(\boldsymbol{\tau})$, but we rather constrain the ratios of the traffic of each database with respect to the traffic of the first database. The pair $(L(\boldsymbol{\tau}), t_1(\boldsymbol{\tau}))$ can grow arbitrarily large to conform with the information-theoretic framework.

The capacity of the PIR problem under asymmetric traffic constraints $C(\boldsymbol{\tau})$ is defined as the supremum of all achievable retrieval rates, i.e., $C(\boldsymbol{\tau}) = \sup R(\boldsymbol{\tau})$.

III. Main Results and Discussions

Our first result is an upper bound on $C(\boldsymbol{\tau})$ as a function of $\boldsymbol{\tau}$ for any fixed M, N.

Theorem 1 (Upper bound) *For the PIR problem under monotone non-increasing asymmetric traffic constraints $\boldsymbol{\tau} = (\tau_1, \cdots, \tau_N)$, the PIR capacity $C(\boldsymbol{\tau})$ is upper bounded by*

$$C(\boldsymbol{\tau}) \le \bar{C}(\boldsymbol{\tau}) = \min_{n_i \in \{1, \cdots, N\}} \frac{1 + \frac{\gamma(n_1)}{n_1} + \frac{\gamma(n_2)}{n_1 n_2} + \cdots + \frac{\gamma(n_{M-1})}{\prod_{i=1}^{M-1} n_i}}{1 + \frac{1}{n_1} + \frac{1}{n_1 n_2} + \cdots + \frac{1}{\prod_{i=1}^{M-1} n_i}} \tag{10}$$

where $\gamma(\ell) = \frac{\sum_{n=\ell+1}^{N} \lambda_n}{\sum_{n=1}^{N} \lambda_n} = \sum_{n=\ell+1}^{N} \tau_n$ corresponds to the sum of the traffic ratios from databases $[\ell + 1 : N]$.

The proof of this upper bound is given in Section IV. We have the following remarks.

Remark 1 *The minimization in (10) is performed to obtain the tightest bound, i.e., the bound in* (10) *is valid for any sequence of $\{n_i\}_{i=1}^{N} \subset \{1, \cdots, N\}^{M-1}$. In particular, restricting the minimization in the bound in* (10) *to monotone non-decreasing sequences $\{n_i\}_{i=1}^{M-1} \subset \{1, \cdots, N\}^{M-1}$ such that $n_1 \le n_2 \le \cdots \le n_{M-1}$ is still a valid upper bound, as it cannot decrease the upper bound $\bar{C}(\boldsymbol{\tau})$ (the feasible set shrinks, hence the optimal value $\bar{C}(\boldsymbol{\tau})$ would be potentially higher). For fixed M, N, the number of such monotone bounds is $\binom{M+N-2}{M-1}$.*

Remark 2 *The upper bound for the capacity function $\bar{C}(\boldsymbol{\tau})$ in (10) is a piece-wise affine function in the traffic ratio vector $\boldsymbol{\tau}$.*

Remark 3 *The upper bound in* (10) *generalizes the known results about the PIR problem. By picking $n_1 = \cdots = n_{M-1} = N$,* (10) *leads to*

$$C(\boldsymbol{\tau}) \leq \frac{1}{1+\frac{1}{N}+\frac{1}{N^2}+\cdots+\frac{1}{N^{M-1}}} \tag{11}$$

which is the capacity of PIR with symmetric traffic (no traffic constraints) in [12]. On the other hand, if $\boldsymbol{\tau} = (1,0,0,\cdots,0)$, which implies that no traffic is returned from any database except for the first one, by picking $n_1 = \cdots = n_{M-1} = 1$, the upper bound in (10) *leads to $\frac{1}{M}$, which is the capacity of the PIR problem with one database [1].*

The following corollary is a direct consequence of Theorem 1. The corollary asserts that there is a strict capacity loss due to the asymmetric traffic constraints if the traffic ratio of the weakest link falls below a certain threshold.

Corollary 1 (Asymmetry hurts) *For the PIR problem under monotone non-increasing asymmetric traffic constraints $\boldsymbol{\tau} = (\tau_1,\cdots,\tau_N)$, if $\tau_N < \tau^*$, such that*

$$\tau^* = \frac{N^{M-1}-1}{N^M-1}, \quad N > 1 \tag{12}$$

then $C(\boldsymbol{\tau}) < C$, where $C = \frac{1}{1+\frac{1}{N}+\cdots+\frac{1}{N^{M-1}}}$ is the PIR capacity without the asymmetric traffic constraints in[1] *[12].*

Proof: From Theorem 1, the upper bound corresponding to $n_1 = N-1$, and $n_2 = \cdots = n_{M-1} = N$ is strictly tighter than the capacity without asymmetric traffic constraints C if

$$\frac{1+\frac{\tau_N}{N-1}}{1+\frac{1}{N-1}+\frac{1}{(N-1)N}+\cdots+\frac{1}{(N-1)N^{M-2}}} < C \tag{13}$$

which leads to

$$\frac{\tau_N}{N-1}\left(1+\frac{1}{N}+\cdots+\frac{1}{N^{M-1}}\right) < \left(\frac{1}{N-1}-\frac{1}{N}\right)\left(1+\frac{1}{N}+\cdots+\frac{1}{N^{M-2}}\right) \tag{14}$$

which further simplifies to

$$\tau_N < \frac{\frac{1}{N}\left(1+\frac{1}{N}+\cdots+\frac{1}{N^{M-2}}\right)}{\left(1+\frac{1}{N}+\cdots+\frac{1}{N^{M-1}}\right)} = \frac{\frac{1}{N}\sum_{i=0}^{M-2}N^i}{\sum_{i=0}^{M-1}N^i} = \tau^* \tag{15}$$

[1]We note that Corollary 1 can be generalized to cases other than the constraint on the traffic of the lowest link. In fact, there exist $\binom{M+N-2}{M-1}$ inequalities regarding the conditions such that the traffic ratio vector results in hurting the retrieval rate. These conditions result from plugging different monotone non-decreasing sequences in (10). That is why, we present the condition in Corollary 1, as an instance of a condition for which we incur a capacity loss, but this indeed is without loss of generality. As a concrete example, let $n_1 = N-2$, and $n_i = N$ for all $n_i \neq 1$. Plugging these numbers in (10) yields the condition $\tau_{N-1}+\tau_N < 2\tau^*$ by following the same steps as in Corollary 1.

which implies that the upper bound for the capacity under the asymmetric traffic constraint is strictly less than C, which in turn implies that any achievable rate is strictly less than the unconstrained capacity. ■

Remark 4 *As the number of messages M becomes large enough, i.e., as $M \to \infty$, the traffic ratio threshold in (12) $\tau^* \to \frac{1}{N}$. This implies that as $M \to \infty$, any asymmetric traffic constraint incurs strict capacity loss.*

Our second result is a lower bound on $C(\boldsymbol{\tau})$ as a function of $\boldsymbol{\tau}$ for any fixed M, N.

Theorem 2 (Lower bound) *For the PIR problem under asymmetric traffic constraints, for a monotone non-decreasing sequence $\mathbf{n} = \{n_i\}_{i=0}^{M-1} \subset \{1,\cdots,N\}^M$, let $n_{-1} = 0$, and $\mathcal{S} = \{i \geq 0 : n_i - n_{i-1} > 0\}$. Denote $y_\ell[k]$ as the number of stages of the achievable scheme that downloads k-sums from the nth database, such that $n_{\ell-1} \leq n \leq n_\ell$, and $\ell \in \mathcal{S}$. Let $\xi_\ell = \prod_{s\in\mathcal{S}\setminus\{\ell\}}\binom{M-2}{s-1}$. The number of stages $y_\ell[k]$ is characterized by the following system of difference equations:*

$$\begin{aligned}
y_0[k] &= (n_0-1)y_0[k-1] + \sum_{j\in\mathcal{S}\setminus\{0\}}(n_j-n_{j-1})y_j[k-1] \\
y_1[k] &= (n_1-n_0-1)y_1[k-1] + \sum_{j\in\mathcal{S}\setminus\{1\}}(n_j-n_{j-1})y_j[k-1] \\
y_\ell[k] &= n_0\xi_\ell\delta[k-\ell-1] + (n_\ell-n_{\ell-1}-1)y_\ell[k-1] \\
&\quad + \sum_{j\in\mathcal{S}\setminus\{\ell\}}(n_j-n_{j-1})y_j[k-1], \quad \ell \geq 2
\end{aligned} \tag{16}$$

where $\delta[\cdot]$ denotes the Kronecker delta function. The initial conditions of (16) *are $y_0[1] = \prod_{s\in\mathcal{S}}\binom{M-2}{s-1}$, and $y_j[k] = 0$ for $k \leq j$. Consequently, the traffic ratio vector $\boldsymbol{\tau}(\mathbf{n}) = (\tau_1(\mathbf{n}),\cdots,\tau_N(\mathbf{n}))$ corresponding to the sequence $\mathbf{n} = \{n_i\}_{i=0}^{M-1}$ is given by:*

$$\tau_n(\mathbf{n}) = \frac{\sum_{k=1}^{M}\binom{M}{k}y_j[k]}{\sum_{\ell\in\mathcal{S}}\sum_{k=1}^{M}\binom{M}{k}y_\ell[k](n_\ell-n_{\ell-1})} \tag{17}$$

for $n_{j-1}+1 \leq n \leq n_j$, and the achievable rate corresponding to $\boldsymbol{\tau}(\mathbf{n})$ is given by:

$$R(\boldsymbol{\tau}(\mathbf{n})) = \frac{\sum_{\ell\in\mathcal{S}}\sum_{k=1}^{M}\binom{M-1}{k-1}y_\ell[k](n_\ell-n_{\ell-1})}{\sum_{\ell\in\mathcal{S}}\sum_{k=1}^{M}\binom{M}{k}y_\ell[k](n_\ell-n_{\ell-1})} \tag{18}$$

Moreover, for $\boldsymbol{\tau} = \sum_{i=1}^{N}\alpha_i\boldsymbol{\tau}(\mathbf{n}_i)$ for $\alpha_i \geq 0$, for all i, and $\sum_{i=1}^{N}\alpha_i = 1$, the following is a lower bound on $C(\boldsymbol{\tau})$,

$$C(\boldsymbol{\tau}) \geq R(\boldsymbol{\tau}) = \sum_{i=1}^{N}\alpha_i R(\boldsymbol{\tau}(\mathbf{n}_i)) \tag{19}$$

The proof of Theorem 2 can be found in Section V. The theorem characterizes an achievable rate for the corner points $\boldsymbol{\tau}(\mathbf{n})$ corresponding to any monotone non-decreasing sequence $\mathbf{n} = \{n_i\}_{i=0}^{M-1} \subset \{1,\cdots,N\}^M$. For any other traffic ratio vector $\boldsymbol{\tau}$, the achievability scheme is obtained by time-sharing between the nearest corner points. We note that due to the large number of corner points, we do not provide an explicit achievable rate for each corner point but we

rather describe the achievable rate by a system of difference equations. The solution of this system of difference equations specifies the traffic ratio vector $\boldsymbol{\tau}(\mathbf{n})$ and the achievable rate $R(\boldsymbol{\tau}(\mathbf{n}))$ corresponding to the monotone non-decreasing sequence $\{n_i\}_{i=0}^{M-1}$. We have the following remarks.

Remark 5 *If $n_i = N$ for all $i \in \{0, \cdots, M-1\}$, then $\mathcal{S} = \{0\}$ and the number of stages of k-sums is described by the following difference equation for any database*

$$y[k] = (N-1)y[k-1] \tag{20}$$

with initial condition of $y[1] = 1$. In this case $\tau_n = \frac{1}{N}$ for all n, and $R = \frac{1}{1+\frac{1}{N}+\cdots+\frac{1}{N^{M-1}}}$, i.e., the scheme in Theorem 2 reduces to the symmetric scheme in [12] if the sequence $\mathbf{n} = (N, N, \cdots, N)$ is used.

Remark 6 *We note that the sequence $\{n_i\}_{i=0}^{M-1}$ suffices to completely specify the traffic ratio vector $\boldsymbol{\tau}(\mathbf{n})$ for every corner point as a consequence of the monotonicity of the sequence, i.e.,*

$$\{n_i\}_{i=0}^{M-1} \Rightarrow (\underbrace{\tilde{\tau}_0, \cdots, \tilde{\tau}_0}_{n_0 \text{ elements}}, \underbrace{\tilde{\tau}_1, \cdots, \tilde{\tau}_1}_{(n_1-n_0) \text{ elements}}, \cdots, \underbrace{\tilde{\tau}_{M-1}, \cdots, \tilde{\tau}_{M-1}}_{(n_{M-1}-n_{M-2}) \text{ elements}}) \tag{21}$$

where $\tilde{\tau}_j = \frac{\sum_{k=1}^{M} \binom{M}{k} y_j[k]}{\sum_{\ell \in \mathcal{S}} \sum_{k=1}^{M} \binom{M}{k} y_j[k](n_\ell - n_{\ell-1})}$.

Remark 7 *For fixed M, N, the number of corner points in Theorem 2 corresponds to the number of monotone non-decreasing sequences $\mathbf{n} = \{n_i\}_{i=0}^{M-1}$, which is*[2] $\binom{M+N-1}{M}$.

The next corollary asserts that the achievable scheme in Theorem 2 is optimal for $M = 2$ and $M = 3$ messages for any traffic ratio vector $\boldsymbol{\tau}$ and any number of databases N.

Corollary 2 (Capacity for $M = 2$ and $M = 3$ messages) *For the PIR problem with asymmetric traffic constraints $\boldsymbol{\tau}$, the capacity $C(\boldsymbol{\tau})$ for $M = 2$ and $M = 3$, and for any arbitrary N is given by:*

$$C(\boldsymbol{\tau}) = \begin{cases} \min\limits_{n_0 \in \{1,\cdots,N\}} \dfrac{1 + \frac{\sum_{n=n_0+1}^{N} \tau_n}{n_0}}{1 + \frac{1}{n_0}}, & M = 2 \\ \min\limits_{n_0 \le n_1 \in \{1,\cdots,N\}} \dfrac{1 + \frac{\sum_{n=n_0+1}^{N} \tau_n}{n_0} + \frac{\sum_{n=n_1+1}^{N} \tau_n}{n_0 n_1}}{1 + \frac{1}{n_0} + \frac{1}{n_0 n_1}}, & M = 3 \end{cases} \tag{22}$$

The proof of Corollary 2 is given in Section VI.

Fig. 2 shows the PIR capacity under asymmetric constraints $C(\lambda_2)$ as a function of λ_2 (which is bijective to $\boldsymbol{\tau}$) for the case of $M = 3$ messages and $N = 2$ databases. We note that the capacity $C(\lambda_2)$ is a piece-wise monotone curve in λ_2, which consists of $\binom{M+N-2}{M-1} = 3$ regimes. There exist $\binom{M+N-1}{M} = 4$ corner points. Specific achievable schemes for the case of $M = 3$ and $N = 2$ are given in Section V-A. Each corner point shown in Fig. 2 corresponds to an explicit achievable scheme given in Section V-A.1. For any other point, time-sharing between corner points is used to achieve these points as shown in Section V-A.2.

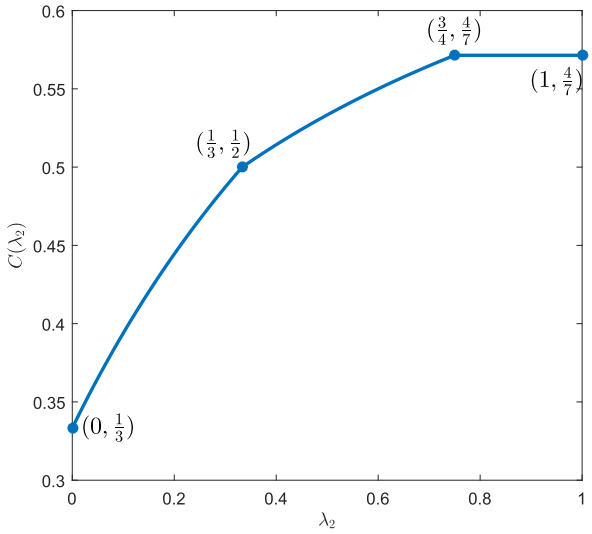

Fig. 2. Capacity function $C(\lambda_2)$ for $M = 3$, $N = 2$.

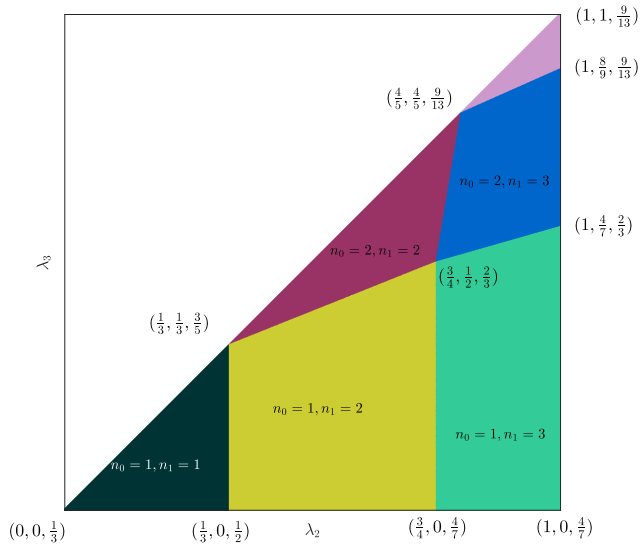

Fig. 3. Illustration of corner points and regions of $C(\lambda_2, \lambda_3)$ for $M = 3$, $N = 3$.

Fig. 3 shows the capacity region $C(\lambda_2, \lambda_3)$ for the case of $M = 3$ messages and $N = 3$ databases as a function of the pair (λ_2, λ_3) (which is bijective to $\boldsymbol{\tau}$). Fig. 3 shows that there exist $\binom{M+N-1}{M} = 10$ corner points, and $\binom{M+N-2}{M-1} = 6$ regions. We show the capacity regions in terms of the triple $(\lambda_2, \lambda_3, C(\lambda_2, \lambda_3))$. Furthermore, for every region we show the corresponding (n_0, n_1) to be plugged in (22). The capacity for any point (λ_2, λ_3) other than the corner points is obtained by time-sharing between the corner points that

[2]We note that the number of *corner points* in the lower and upper bounds are the same as the upper bound corresponds to $\binom{M+N-2}{M-1}$ regions (surfaces) in N-dim space. This induces $\binom{M+N-1}{M}$ corner points at the intersections of the regions (see Fig. 3, Fig. 4, and Fig. 5).

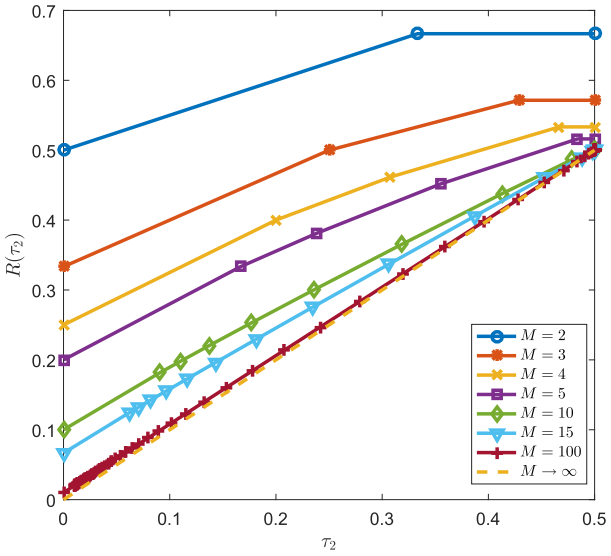

Fig. 4. Achievable rate-traffic ratio tradeoff for $N = 2$.

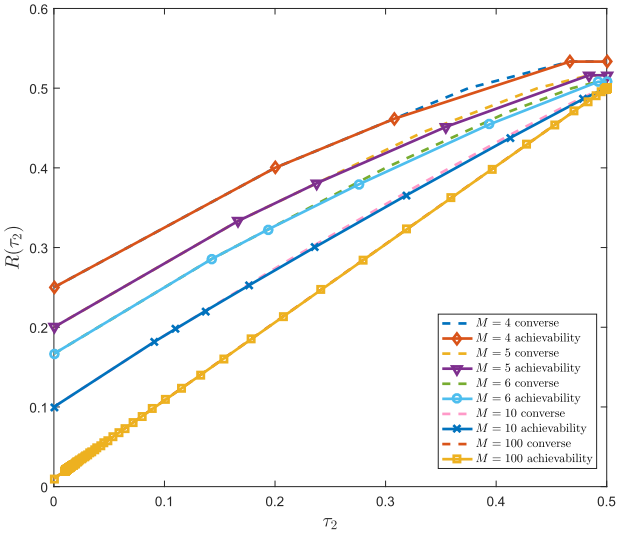

Fig. 5. Lower and upper bounds for $N = 2$.

enclose (λ_2, λ_3). Specific achievable schemes for $M = 3$, $N = 3$ are given in Section VIII-B.

Finally, in the following corollary, we specialize the achievable scheme in Theorem 2 to the case of $N = 2$ for any arbitrary M.

Corollary 3 (Achievable traffic versus retrieval rate tradeoff for $N = 2$ databases) *For the PIR problem with $N = 2$ and an arbitrary M under asymmetric traffic constraints $\boldsymbol{\tau} = (1 - \tau_2, \tau_2)$, $\tau_2 \leq \frac{1}{2}$, let $s_2 \in \{1, \cdots, M-1\}$, for the traffic ratio $\tau_2(s_2)$, where*

$$\tau_2(s_2) = \frac{\sum_{i=0}^{\left\lfloor \frac{M-s_2-1}{2} \right\rfloor} \binom{M}{s_2+2i+1}}{M\binom{M-2}{s_2-1} + \sum_{i=0}^{M-s_2-1} \binom{M}{s_2+1+i}} \tag{23}$$

the PIR capacity $C(\tau_2(s_2))$ is lower bounded by:

$$\begin{aligned} C(\tau_2(s_2)) &\geq R(\tau_2(s_2)) \\ &= \frac{\binom{M-2}{s_2-1} + \sum_{i=0}^{M-s_2-1} \binom{M-1}{s_2+i}}{M\binom{M-2}{s_2-1} + \sum_{i=0}^{M-s_2-1} \binom{M}{s_2+1+i}} \end{aligned} \tag{24}$$

Moreover, if $\tau_2(s_2) < \tau_2 < \tau_2(s_2+1)$, and $\alpha \in (0,1)$, such that $\tau_2 = \alpha\tau_2(s_2) + (1-\alpha)\tau_2(s_2+1)$, then

$$C(\tau_2) \geq R(\tau_2) = \alpha R(\tau_2(s_2)) + (1-\alpha)R(\tau_2(s_2+1)) \tag{25}$$

The proof of Corollary 3 is given in Section VII.

Remark 8 *Fig. 4 shows the tradeoff between the traffic ratio τ_2 and the achievable retrieval rate $R(\tau_2)$. We note that as M increases $R(\tau_2)$ decreases pointwise. We observe that as $M \to \infty$, the rate-traffic tradeoff converges to $R(\tau_2) = \tau_2$. This implies that for large enough M, our achievable scheme reduces to time-sharing between the trivial achievable scheme of downloading all the messages from database 1 which achieves a rate of $\frac{1}{M}$, and the asymptotically-optimal achievable scheme in [6] which achieves $R = 1 - \frac{1}{N}$.*

Remark 9 *Fig. 5 shows the lower and upper bounds for $N = 2$ and $M = 4, 5, 6, 10, 100$. It is clear that the bounds do not match. However, we note that the largest gap monotonically decreases as M increases until both bounds match at the limit $M \to \infty$, where $R(\tau_2) = \tau_2$. The maximum possible gap for $N = 2$ is 0.008, which appears at $M = 4$. Furthermore, we can analytically prove that the upper bound $\bar{C}(\boldsymbol{\tau}) \to \tau_2$ by choosing $n_i = 1$ for all i, hence the upper bound becomes,*

$$\bar{C}(\boldsymbol{\tau}) = \frac{1 + (M-1)\tau_2}{M} \to \tau_2 \tag{26}$$

as $M \to \infty$. This settles the asymptotic PIR capacity to be $C(\tau_2) = \tau_2$ for $N = 2$ and $M \to \infty$.

IV. Converse Proof

In this section, we derive an upper bound for the PIR problem with asymmetric traffic constraints. We extend the converse techniques introduced in [12] to account for the asymmetry of the returned answer strings.

We first need the following lemma, which characterizes a fundamental lower bound on the interference from the undesired messages within the answer strings, i.e., a lower bound on $\sum_{n=1}^{N} t_n - L$, as a consequence of the privacy constraint. The proof of this lemma can be found in [12, Lemma 5]. The proof follows for our case since the privacy constraint does not change in the PIR with asymmetric traffic constraints, and the fact that the proof in [12, Lemma 5] deals with the length of the entire downloaded answer strings $A_{1:N}^{[1]}$ and not the individual lengths of each answer string, see [12, equations (46)-(47)].

Lemma 1 (Interference lower bound) *For the PIR problem under asymmetric traffic constraints $\{t_n\}_{n=1}^{N}$, the interference from undesired messages within the answer strings $\sum_{n=1}^{N} t_n - L$ is lower bounded as,*

$$\sum_{n=1}^{N} t_n - L + o(L) \geq I\left(W_{2:M}; Q_{1:N}^{[1]}, A_{1:N}^{[1]} | W_1\right) \tag{27}$$

In the following lemma, we prove an inductive relation for the mutual information term on the right hand side of (27). In this lemma, the interference lower bound in (27) is expanded into two parts. The first part, which contains the answer strings from the first n_{m-1} databases $A^{[m]}_{1:n_{m-1}}$, is dealt with as in the proof of [12, Lemma 6]. For the second part, which contains the remaining answer strings $A^{[m]}_{n_{m-1}+1:N}$, each answer string $A^{[m]}_n$ is bounded trivially by the length of the answer string t_n.

Lemma 2 (Induction lemma) *For all $m \in \{2, \ldots, M\}$ and for an arbitrary $n_{m-1} \in \{1, \cdots, N\}$, the mutual information term in Lemma 1 can be inductively lower bounded as,*

$$I\left(W_{m:M}; Q^{[m-1]}_{1:N}, A^{[m-1]}_{1:N} | W_{1:m-1}\right) \geq \frac{1}{n_{m-1}} I\left(W_{m+1:M}; Q^{[m]}_{1:N}, A^{[m]}_{1:N} | W_{1:m}\right) + \frac{1}{n_{m-1}} \left(L - t_1 \sum_{n=n_{m-1}+1}^{N} \lambda_n\right) - \frac{o(L)}{n_{m-1}} \tag{28}$$

We note that [12, Lemma 6] can be interpreted as a special case of Lemma 2 with setting $n_{m-1} = N$. Intuitively, n_{m-1} represents the number of databases that can apply the optimal symmetric scheme in [12] if the user wants to retrieve message W_{m-1} from the set of $W_{m-1:M}$ messages (i.e., conditioned on $W_{1:m-1}$).

Proof: We start with the left hand side of (28) after multiplying by n_{m-1},

$$\begin{aligned}
&n_{m-1} I\left(W_{m:M}; Q^{[m-1]}_{1:N}, A^{[m-1]}_{1:N} | W_{1:m-1}\right) \\
&\geq n_{m-1} I\left(W_{m:M}; Q^{[m-1]}_{1:n_{m-1}}, A^{[m-1]}_{1:n_{m-1}} | W_{1:m-1}\right) && (29)\\
&\geq \sum_{n=1}^{n_{m-1}} I\left(W_{m:M}; Q^{[m-1]}_{n}, A^{[m-1]}_{n} | W_{1:m-1}\right) && (30)\\
&\overset{(7)}{=} \sum_{n=1}^{n_{m-1}} I\left(W_{m:M}; Q^{[m]}_{n}, A^{[m]}_{n} | W_{1:m-1}\right) && (31)\\
&\overset{(3)}{=} \sum_{n=1}^{n_{m-1}} I\left(W_{m:M}; A^{[m]}_{n} | Q^{[m]}_{n}, W_{1:m-1}\right) && (32)\\
&\overset{(4)}{=} \sum_{n=1}^{n_{m-1}} H\left(A^{[m]}_{n} | Q^{[m]}_{n}, W_{1:m-1}\right) && (33)\\
&\geq \sum_{n=1}^{n_{m-1}} H\left(A^{[m]}_{n} | A^{[m]}_{1:n-1}, Q^{[m]}_{1:n_{m-1}}, W_{1:m-1}\right) && (34)\\
&\overset{(4)}{=} \sum_{n=1}^{n_{m-1}} I\left(W_{m:M}; A^{[m]}_{n} | A^{[m]}_{1:n-1}, Q^{[m]}_{1:n_{m-1}}, W_{1:m-1}\right) && (35)\\
&= I\left(W_{m:M}; A^{[m]}_{1:n_{m-1}} | Q^{[m]}_{1:n_{m-1}}, W_{1:m-1}\right) && (36)\\
&\overset{(3)}{=} I\left(W_{m:M}; Q^{[m]}_{1:n_{m-1}}, A^{[m]}_{1:n_{m-1}} | W_{1:m-1}\right) && (37)\\
&\overset{(3),(4)}{=} I\left(W_{m:M}; Q^{[m]}_{1:N}, A^{[m]}_{1:N} | W_{1:m-1}\right) \\
&\quad - I\left(W_{m:M}; A^{[m]}_{n_{m-1}+1:N} | Q^{[m]}_{1:N}, A^{[m]}_{1:n_{m-1}}, W_{1:m-1}\right) && (38)\\
&\overset{(4)}{=} I\left(W_{m:M}; Q^{[m]}_{1:N}, A^{[m]}_{1:N} | W_{1:m-1}\right) \\
&\quad - H\left(A^{[m]}_{n_{m-1}+1:N} | Q^{[m]}_{1:N}, A^{[m]}_{1:n_{m-1}}, W_{1:m-1}\right) && (39)\\
&\overset{(5)}{\geq} I\left(W_{m:M}; Q^{[m]}_{1:N}, A^{[m]}_{1:N} | W_{1:m-1}\right) - t_1 \sum_{n=n_{m-1}+1}^{N} \lambda_n && (40)\\
&\overset{(8)}{=} I\left(W_{m:M}; W_m, Q^{[m]}_{1:N}, A^{[m]}_{1:N} | W_{1:m-1}\right) \\
&\quad - t_1 \sum_{n=n_{m-1}+1}^{N} \lambda_n - o(L) && (41)\\
&= I(W_{m:M}; W_m | W_{1:m-1}) + I\left(W_{m:M}; Q^{[m]}_{1:N}, A^{[m]}_{1:N} | W_{1:m}\right) \\
&\quad - t_1 \sum_{n=n_{m-1}+1}^{N} \lambda_n - o(L) && (42)\\
&= I\left(W_{m+1:M}; Q^{[m]}_{1:N}, A^{[m]}_{1:N} | W_{1:m}\right) \\
&\quad + \left(L - t_1 \sum_{n=n_{m-1}+1}^{N} \lambda_n\right) - o(L) && (43)
\end{aligned}$$

where (29), (30) follow from the non-negativity of mutual information, (31) follows from the privacy constraint, (32) follows from the independence of $\left(W_{m:M}, Q^{[m]}_n\right)$, (33), (35) follow from the fact that the answer string $A^{[m]}_n$ is a deterministic function of $\left(Q^{[m]}_n, W_{1:M}\right)$, (34) follows from conditioning reduces entropy, (37) follows from the independence of $\left(W_{m:M}, Q^{[m]}_{1:n_{m-1}}\right)$, (38) follows from the chain rule, the independence of the queries and the messages, and the fact that $Q^{[m]}_{1:N} \to Q^{[m]}_{1:n_{m-1}} \to A^{[m]}_{1:n_{m-1}}$ forms a Markov chain by (4), (39) follows from the fact that the answer strings $A^{[m]}_{1:n_{m-1}}$ are fully determined from $\left(Q^{[m]}_{1:N}, W_{1:M}\right)$, (40) follows from the fact that conditioning reduces entropy and $H(A_{n_{m-1}+1:N}) \leq \sum_{n=n_{m-1}+1}^{N} t_n$ which is equal to $t_1 \sum_{n=n_{m-1}+1}^{N} \lambda_n$ from the asymmetric traffic constraints, (41) follows from the reliability constraint. Finally, dividing both sides by n_{m-1} leads to (28). ■

Now, we are ready to derive an explicit upper bound for the retrieval rate under asymmetric traffic constraints. Applying Lemma 1 and Lemma 2 successively for an arbitrary sequence $\{n_i\}_{i=1}^{M-1} \subset \{1, \cdots, N\}^{M-1}$ and observing that $\sum_{n=1}^{N} t_n = t_1 \sum_{n=1}^{N} \lambda_n$ under the asymmetric traffic constraints, we have the following

$$\begin{aligned}
&t_1 \sum_{n=1}^{N} \lambda_n - L + \tilde{o}(L) \\
&\overset{(27)}{\geq} I\left(W_{2:M}; Q^{[1]}_{1:N}, A^{[1]}_{1:N} | W_1\right) && (44)\\
&\overset{(28)}{\geq} \frac{1}{n_1}\left(L - t_1 \sum_{n=n_1+1}^{N} \lambda_n\right) + \frac{1}{n_1} I\left(W_{3:M}; Q^{[2]}_{1:N}, A^{[2]}_{1:N} | W_{1:2}\right) && (45)
\end{aligned}$$

$$\overset{(28)}{\geq} \frac{1}{n_1}\left(L-t_1\sum_{n=n_1+1}^{N}\lambda_n\right)+\frac{1}{n_1 n_2}\left(L-t_1\sum_{n=n_2+1}^{N}\lambda_n\right) + \frac{1}{n_2} I\left(W_{4:M}; Q_{1:N}^{[3]}, A_{1:N}^{[3]} | W_{1:3}\right) \quad (46)$$

$$\overset{(28)}{\geq} \cdots$$

$$\overset{(28)}{\geq} \frac{1}{n_1}\left(L-t_1\sum_{n=n_1+1}^{N}\lambda_n\right)+\frac{1}{n_1 n_2}\left(L-t_1\sum_{n=n_2+1}^{N}\lambda_n\right) + \cdots + \frac{1}{\prod_{i=1}^{M-1} n_i}\left(L-t_1\sum_{n=n_{M-1}+1}^{N}\lambda_n\right) \quad (47)$$

where $\tilde{o}(L) = \left(1+\frac{1}{n_1}+\frac{1}{n_1 n_2}+\cdots+\frac{1}{\prod_{i=1}^{M-1} n_i}\right) o(L)$, (44) follows from Lemma 1, and the remaining bounding steps follow from successive application of Lemma 2.

Ordering terms, we have,

$$\left(1+\frac{1}{n_1}+\frac{1}{n_1 n_2}+\cdots+\frac{1}{\prod_{i=1}^{M-1} n_i}\right)L-\tilde{o}(L) \leq \left(1+\frac{\gamma(n_1)}{n_1}+\cdots+\frac{\gamma(n_{M-1})}{\prod_{i=1}^{M-1} n_i}\right) t_1 \sum_{n=1}^{N}\lambda_n \quad (48)$$

where $\gamma(\ell) = \frac{\sum_{n=\ell+1}^{N}\lambda_n}{\sum_{n=1}^{N}\lambda_n} = \sum_{n=\ell+1}^{N}\tau_n$ corresponds to the sum of the traffic ratios from databases $[\ell+1:N]$.

We conclude the proof by taking $L \to \infty$. Thus, for an arbitrary sequence $\{n_i\}_{i=1}^{M-1}$, we have

$$R(\boldsymbol{\tau}) = \frac{L}{t_1 \sum_{n=1}^{N}\lambda_n} \leq \frac{1+\frac{\gamma(n_1)}{n_1}+\frac{\gamma(n_2)}{n_1 n_2}+\cdots+\frac{\gamma(n_{M-1})}{\prod_{i=1}^{M-1} n_i}}{1+\frac{1}{n_1}+\frac{1}{n_1 n_2}+\cdots+\frac{1}{\prod_{i=1}^{M-1} n_i}} \quad (49)$$

Finally, we get the tightest bound by minimizing over the sequence $\{n_i\}_{i=1}^{M-1}$ over the set $\{1,\cdots,N\}$, as

$$R(\boldsymbol{\tau}) \leq \min_{n_i\in\{1,\cdots,N\}} \frac{1+\frac{\gamma(n_1)}{n_1}+\frac{\gamma(n_2)}{n_1 n_2}+\cdots+\frac{\gamma(n_{M-1})}{\prod_{i=1}^{M-1} n_i}}{1+\frac{1}{n_1}+\frac{1}{n_1 n_2}+\cdots+\frac{1}{\prod_{i=1}^{M-1} n_i}} \quad (50)$$

$$= \min_{n_i\in\{1,\cdots,N\}} \frac{1+\frac{\sum_{n=n_1+1}^{N}\tau_n}{n_1}+\cdots+\frac{\sum_{n=n_{M-1}+1}^{N}\tau_n}{\prod_{i=1}^{M-1} n_i}}{1+\frac{1}{n_1}+\cdots+\frac{1}{\prod_{i=1}^{M-1} n_i}} \quad (51)$$

Remark 10 *From the converse proof, we note that we can intuitively interpret n_i as the number of databases that can apply the symmetric traffic scheme in [12] if the number of messages is reduced to be $M-i+1$. We point out that in the absence of asymmetric traffic constraints as in [12], all databases can apply symmetric schemes, therefore $n_i = N$ for all $i \in \{1,\cdots,M-1\}$. Now, in Lemma 2, we lower bound the term $I\left(W_{m:M}; Q_{1:N}^{[m-1]}, A_{1:N}^{[m-1]} | W_{1:m-1}\right)$ which refers to a reduced PIR problem with $M-m+1$ messages. The privacy constraint*

$$(Q_n^{[i]}, A_n^{[i]}, W_{1:M}) \sim (Q_n^{[j]}, A_n^{[j]}, W_{1:M}), \quad \forall j \in \{1,\cdots,M\} \quad (52)$$

is less constrained when the number of messages decreases from M to $M-m+1$ (as we have less pair-wise statistical equivalence). Hence, if n_m databases can adopt symmetric scheme when the number of messages is $M-m+1$, then $n_{m+1} \geq n_m$ databases can also adopt the same symmetric scheme as the number of messages is reduced to $M-m$ messages, which leads to more flexibility in terms of satisfying the traffic constraints. Therefore, it suffices to evaluate the bound in (10) *for monotone non-decreasing sequences $\{n_i\}_{i=1}^{M-1} \subset \{1,\cdots,N\}^{M-1}$ such that $n_1 \leq n_2 \leq \cdots \leq n_{M-1}$. Note that from our achievable scheme point of view, the monotone sequences have an operational meaning, in which the number of databases in group ℓ is equal to $n_\ell - n_{\ell-1} \geq 0$. Furthermore, for the cases of $M = 2, 3$, the lower and upper bounds match, which implies that there is no loss of generality in focusing on monotone non-decreasing sequences.*

V. Achievability Proof

The achievability scheme for the PIR problem under asymmetric traffic constraints is inspired by the PIR schemes in [12], [30]. Our achievable scheme applies message symmetry, and side information exploitation as in [12], [30]. However, due to the asymmetric traffic constraints, database symmetry cannot be applied. In an alternative view, we use the side information in an asymmetric fashion among the databases. The most relevant achievable scheme to our achievable scheme here is the scheme in [30], in which the bits stored in the user's cache is exploited differently depending on the caching ratio. We begin the discussion by presenting the $M = 3$, $N = 2$ case as a concrete example to illustrate the main concepts of the scheme.

A. Motivating Example: $M = 3$ Messages, $N = 2$ Databases

In this section, we show the achievability scheme for $M = 3$, $N = 2$. We first carry out the minimization in (10) over $n_1, n_2 \in \{1, 2\}$. In this case, we have 4 upper bounds (or effectively 3 bounds if $n_1 \leq n_2$ restriction is applied). By taking the minimum of these bounds for every $\lambda_2 \in [0, 1]$, we have the following explicit upper bound on the capacity as a function of λ_2 (which is in bijection to τ_2)

$$C(\lambda_2) \leq \begin{cases} \frac{1+3\lambda_2}{3(1+\lambda_2)}, & 0 \leq \lambda_2 \leq \frac{1}{3} \\ \frac{2(1+2\lambda_2)}{5(1+\lambda_2)}, & \frac{1}{3} \leq \lambda_2 \leq \frac{3}{4} \\ \frac{4}{7}, & \frac{3}{4} \leq \lambda_2 \leq 1 \end{cases} \quad (53)$$

To show the achievability of the upper bound in (53), let a_i, b_i, c_i denote randomly and independently permuted symbols of messages W_1, W_2, W_3, respectively. Define $s_2 \in \{0, 1, 2\}$ to be the number of side information symbols that are used simultaneously in database 2 within the initial round of downloads. First, we show the achievability of the corner points, i.e., the achievability of the points corresponding to $\lambda_2 \in \{0, \frac{1}{3}, \frac{3}{4}, 1\}$.

1) Achievability of the Corner Points:

a) The $\lambda_2 = 0$ Corner Point: $\lambda_2 = 0$ means that the second database does not return any answer strings. The optimal achievable scheme is to download all files from the first database (see Table I). This scheme achieves $R = \frac{1}{3} = C(0)$.

TABLE I
THE QUERY TABLE FOR $M = 3$, $N = 2$, $\lambda_2 = 0$

Database 1	Database 2
a_1, b_1, c_1	

TABLE II
THE QUERY TABLE FOR $M = 3$, $N = 2$, $\lambda_2 = 1$

Database 1	Database 2
a_1, b_1, c_1	a_2, b_2, c_2
$a_3 + b_2$ $a_4 + c_2$ $b_3 + c_3$	$a_5 + b_1$ $a_6 + c_1$ $b_4 + c_4$
$a_7 + b_4 + c_4$	$a_8 + b_3 + c_3$

TABLE III
THE QUERY TABLE FOR $M = 3$, $N = 2$, $\lambda_2 = \frac{3}{4}$

Database 1	Database 2
a_1, b_1, c_1	
	$a_2 + b_1$ $a_3 + c_1$ $b_2 + c_2$
$a_4 + b_2 + c_2$	

TABLE IV
THE QUERY TABLE FOR $M = 3$, $N = 2$, $\lambda_2 = \frac{1}{3}$

Database 1	Database 2
a_1, b_1, c_1	
	$a_2 + b_1 + c_1$

b) The $\lambda_2 = 1$ Corner Point: Since $\lambda_1 = 1$ by definition, $\lambda_2 = 1$ means that a symmetric scheme can be applied to both databases. Thus, the optimal achievable scheme is the optimal symmetric scheme in [12] (see Table II). We present the scheme here for completeness. In this scheme, the user starts with downloading the individual symbols a_1, b_1, c_1 from database 1. Since $\lambda_2 = 1$, database symmetry can be applied, hence the user downloads a_2, b_2, c_2 from database 2. Note that in this case, the user does not exploit the side information generated from database 1 in the first round of downloads, but rather downloads new individual symbols, hence $s_2 = 0$ in this case. The undesired symbols $b_i, c_i, i = 1, 2$ can be exploited in the other database. This can be done by downloading $a_3 + b_2$, $a_4 + c_2$ from database 1, and similarly by applying database symmetry, the user downloads $a_5 + b_1$, $a_6 + c_1$ from database 2. In order to satisfy the privacy constraint, the user applies the message symmetry and downloads $b_3 + c_3$ from database 1, and $b_4 + c_4$ from database 2. Finally, the user exploits the newly generated side information by downloading $a_7 + b_4 + c_4$ from database 1, and $a_8 + b_3 + c_3$ from database 2. Consequently, the user downloads $L = 8$ symbols in 14 downloads which results in $R = \frac{8}{14} = \frac{4}{7} = C(1)$.

c) The $\lambda_2 = \frac{3}{4}$ Corner Point: The user can cut the first round of downloads in database 2 and exploit the side information generated from database 1 directly in the form of sums of 2, i.e., the user downloads a_1, b_1, c_1 from database 1 and then exploits the undesired symbols as side information by downloading $a_2 + b_1$, $a_3 + c_1$ from database 2. The user then applies message symmetry and downloads $b_2 + c_2$. Since the user uses 1 bit of side information in the initial download round from database 2, $s_2 = 1$ in this case. Finally, the user exploits the undesired sum $b_2 + c_2$ from database 2 as a side information in database 1 and downloads $a_4 + b_2 + c_2$. Using this scheme the user downloads 4 symbols from database 1 and 3 symbols from database 2, hence $\lambda_2 = \frac{3}{4}$. The user downloads $L = 4$ desired symbols out of 7 downloads, thus $R = \frac{4}{7} = C(\frac{3}{4})$. The privacy is satisfied since W_1, W_2, W_3 are independently and randomly permuted, and since the scheme includes all the possible combinations of the sums in any round. The query table for this scheme is given in Table III. We note that this scheme is exactly the asymmetric achievable scheme presented in [16].

d) The $\lambda_2 = \frac{1}{3}$ Corner Point: In this case, the user downloads a_1, b_1, c_1 from database 1. In database 2, the user exploits the side information b_1, c_1 simultaneously and downloads $a_2 + b_1 + c_1$. Due to the fact that 2 side information symbols are used simultaneously in the initial round of download from database 2, $s_2 = 2$ in this case. Using this scheme the user downloads 3 symbols from database 1 and 1 symbol from database 2, therefore $\lambda_2 = \frac{1}{3}$. The user downloads $L = 2$ desired symbols in 4 downloads, hence $R = \frac{1}{2} = C(\frac{1}{3})$. The privacy follows by the same argument as in the previous case. The query table for this case is given in Table IV.

2) Achievability of Non-Corner Points: In the following, we show that by combining the achievable schemes of the corner points over different symbols, the upper bound in (53) is tight for any λ_2. We note that the privacy constraint is still satisfied after this combination as each scheme operates over different sets of symbols and the fact that each scheme satisfies the privacy constraint individually. A formal argument for proving that combination of private schemes remains private can be found in [16, Theorem 4]. Let ν_{s_2}, where $s_2 = 0, 1, 2$, denote the number of repetitions of the scheme that uses s_2 side information symbols simultaneously in the first round of download in database 2. By convention, let ν_3 denote the number of repetitions of the trivial retrieval scheme, i.e., when the retrieval is solely done from database 1.

a) The $0 \leq \lambda_2 \leq \frac{1}{3}$ Regime: We combine the achievable scheme of $\lambda_2 = 0$ corner point with the achievable scheme of $\lambda_2 = \frac{1}{3}$ corner point. The achievable scheme of $\lambda_2 = 0$ downloads 3 symbols from database 1 and 0 symbols from database 2. We perform this scheme ν_3 repetitions. The achievable scheme of $\lambda_2 = \frac{1}{3}$ downloads 3 symbols from database 1 and 1 symbol from database 2. We perform this scheme ν_2 repetitions. Under the asymmetric traffic constraints, this results in the following system of equations

$$3\nu_3 + 3\nu_2 = t_1 \tag{54}$$

$$\nu_2 = \lambda_2 t_1 \tag{55}$$

This system has a unique solution (parametrized by t_1) of $\nu_2 = \lambda_2 t_1$ and $\nu_3 = \frac{1-3\lambda_2}{3} t_1$. Note that $\nu_3 \geq 0$ in the regime of $0 \leq \lambda_2 \leq \frac{1}{3}$. Since the scheme of $\lambda_2 = 0$ downloads 1 symbol from the desired message and the scheme of $\lambda_2 = \frac{1}{3}$ downloads 2 symbols from the desired message. The achievable rate $R(\lambda_2)$ is given by

$$R(\lambda_2) = \frac{2\nu_2 + \nu_3}{(1+\lambda_2)t_1} = \frac{1+3\lambda_2}{3(1+\lambda_2)} = C(\lambda_2), \quad 0 \leq \lambda_2 \leq \frac{1}{3} \tag{56}$$

b) The $\frac{1}{3} \leq \lambda_2 \leq \frac{3}{4}$ Regime: Similarly, the user combines the achievable schemes of $\lambda_2 = \frac{1}{3}$ and $\lambda_2 = \frac{3}{4}$ corner points. The user applies the scheme of $\lambda_2 = \frac{1}{3}$ for ν_2 repetitions, which downloads 3 symbols from database 1 and 1 symbol from database 2 and has $L = 2$. The user applies the scheme of $\lambda_2 = \frac{3}{4}$ for ν_1 repetitions, which downloads 4 symbols from database 1 and 3 symbols from database 2 and has $L = 4$. This results in the following system of equations

$$4\nu_1 + 3\nu_2 = t_1 \tag{57}$$

$$3\nu_1 + \nu_2 = \lambda_2 t_1 \tag{58}$$

which has the following solution: $\nu_1 = \frac{-1+3\lambda_2}{5} t_1 \geq 0$ and $\nu_2 = \frac{3-4\lambda_2}{5} t_1 \geq 0$ in the regime of $\frac{1}{3} \leq \lambda_2 \leq \frac{3}{4}$. Consequently, the achievable rate is given by

$$R(\lambda_2) = \frac{4\nu_1 + 2\nu_2}{(1+\lambda_2)t_1} = \frac{2(1+2\lambda_2)}{5(1+\lambda_2)} = C(\lambda_2), \quad \frac{1}{3} \leq \lambda_2 \leq \frac{3}{4} \tag{59}$$

c) The $\frac{3}{4} \leq \lambda_2 \leq 1$ Regime: The user combines the achievable schemes of $\lambda_2 = \frac{3}{4}$ and $\lambda_2 = 1$ corner points. The user repeats the scheme of $\lambda_2 = \frac{3}{4}$ for ν_1 repetitions, and the scheme of $\lambda_2 = 1$ for ν_0 repetitions. This results in the following system of equations

$$4\nu_1 + 7\nu_0 = t_1 \tag{60}$$

$$3\nu_1 + 7\nu_0 = \lambda_2 t_1 \tag{61}$$

The solution for this system is given by: $\nu_1 = (1-\lambda_2)t_1 \geq 0$ and $\nu_0 = \frac{-3+4\lambda_2}{7} t_1 \geq 0$ in the regime of $\frac{3}{4} \leq \lambda_2 \leq 1$. The corresponding rate is given by

$$R(\lambda_2) = \frac{4\nu_1 + 8\nu_0}{(1+\lambda_2)t_1} = \frac{4}{7} = C(\lambda_2), \qquad \frac{3}{4} \leq \lambda_2 \leq 1 \tag{62}$$

d) Specific Example for Non-Corner Points, $\lambda_2 = \frac{1}{2}$: The query table for this case is given in Table V. The user applies the scheme of $\lambda_2 = \frac{3}{4}$ for $\nu_1 = \frac{-1+3\lambda_2}{5} t_1 = \frac{1}{10} t_1$ repetitions, and the scheme of $\lambda_2 = \frac{1}{3}$ for $\nu_2 = \frac{3-4\lambda_2}{5} t_1 = \frac{1}{5} t_1$ repetitions. Choosing $t_1 = 10$, we have $\nu_1 = 1$ and $\nu_2 = 2$. The scheme downloads 10 symbols from database 1 and 5 symbols from database 2, thus, $\lambda_2 = \frac{1}{2}$. The scheme downloads 8 symbols in 15 downloads, hence $R(\frac{1}{2}) = \frac{8}{15} = \frac{2(1+2\lambda_2)}{5(1+\lambda_2)} = C(\frac{1}{2})$.

TABLE V
THE QUERY TABLE FOR $M = 3$, $N = 2$, $\lambda_2 = \frac{1}{2}$

Database 1	Database 2
a_1, b_1, c_1	$a_2 + b_1$ $a_3 + c_1$ $b_2 + c_2$
$a_4 + b_2 + c_2$	
a_5, b_3, c_3	$a_6 + b_3 + c_3$
a_7, b_4, c_4	$a_8 + b_4 + c_4$

B. Description of the General Scheme

In this section, we describe the general achievable scheme that achieves the retrieval rates in Theorem 2. We first show explicitly the achievability schemes for corner points, i.e., the achievability scheme for every monotone non-decreasing sequence $\{n_i\}_{i=0}^{M-1} \subset \{1, \cdots, N\}^M$. We note that our achievability scheme is different in two key steps: First regarding the database symmetry, we note that it is not applied over all databases directly as in [12], but rather it is applied over groups of databases, such as, group 0 includes databases 1 through n_0, group 1 includes databases $n_0 + 1$ through n_1, etc. Second, regarding the exploitation of side information step, we note that each group of databases exploits side information differently in the *initial* round of downloading. More specifically, we note that group 0 of databases do not exploit any side information in the initial round of the download, group 1 exploits 1 side information symbol in the initial round of the download, group 2 exploits sums of 2 side information symbols in the initial round of the download, and so on.

Next, we show that for non-corner points, time-sharing between corner points is achievable and this concludes the achievability proof of Theorem 2.

1) Achievability Scheme for the Corner Points: Let $s_n \in \{0, 1, \cdots, M-1\}$ denote the number of side information symbols that are used simultaneously in the initial round of downloads at the nth database. For a given non-decreasing sequence[3] $\{n_i\}_{i=0}^{M-1} \subset \{1, \cdots, N\}^M$, let $s_n = i$ for all $n_{i-1} + 1 \leq n \leq n_i$ with $n_{-1} = 0$ by convention. Denote $\mathcal{S} = \{i : s_n = i \text{ for some } n \in \{1, \cdots, N\}\}$. We follow the round and stage definitions in [22]. The kth round is the download queries that admit a sum of k different messages (k-sum in [12]). A stage of the kth round is a query block of the kth round that exhausts all $\binom{M}{k}$ combinations of the k-sum. Denote $y_\ell[k]$ to be the number of stages in round k downloaded from the nth database, such that $n_{\ell-1} + 1 \leq n \leq n_\ell$. The details of the achievable scheme are as follows:

1) *Initialization:* The user permutes each message independently and uniformly using a random interleaver, i.e.,

$$x_m(i) = W_m(\pi_m(i)), \quad i \in \{1, \cdots, L\} \tag{63}$$

[3]We note that the monotone non-decreasing sequences in the achievability and the converse proofs serve similar but not exactly the same roles. In the achievability proof, the index n_i corresponds to the index of the last database that exploits a sum of i symbols in the initial round of download. For the converse proof, the index n_i corresponds to the number of databases that apply the symmetric PIR scheme if the number of messages is reduced to $M - i + 1$ messages. We use the same notation in both proofs because in both cases we minimize over the non-decreasing sequence to get the lower/upper bound and also to simplify the proof of the capacity for the cases of $M = 2$, $M = 3$. As an example, please see footnote 4.

where $x_m(i)$ is the ith symbol of the permuted W_m, $\pi_m(\cdot)$ is a random interleaver for the mth message that is chosen independently, uniformly, and privately at the user's side. From the nth database where $1 \leq n \leq n_0$, the user downloads $\prod_{s\in\mathcal{S}} \binom{M-2}{s-1}$ symbols from the desired message. The user sets the round index $k = 1$. I.e., the user starts downloading the desired symbols from $y_0[1] = \prod_{s\in\mathcal{S}} \binom{M-2}{s-1}$ different stages.

2) *Message symmetry:* To satisfy the privacy constraint, for each stage initiated in the previous step, the user completes the stage by downloading the remaining $\binom{M-1}{k}$ k-sum combinations that do not include the desired symbols, in particular, if $k = 1$, the user downloads $\prod_{s\in\mathcal{S}} \binom{M-2}{s-1}$ individual symbols from each undesired message.

3) *Database symmetry:* Due to the asymmetric traffic constraints, the original database symmetry step in [12] cannot be applied directly to our problem. Instead, we divide the databases into groups. Group $\ell \in \mathcal{S}$ corresponds to databases $n_{\ell-1}+1$ to n_ℓ. Database symmetry is applied within each group only. Consequently, the user repeats step 2 over each group of databases, in particular, if $k = 1$, the user downloads $\prod_{s\in\mathcal{S}} \binom{M-2}{s-1}$ individual symbols from each message from the first n_0 databases (group 1).

4) *Exploitation of side information:* This step is also different from [12] because of the asymmetric traffic constraints. In order to create different lengths of the answer strings, the initial exploitation of side information is group-dependent as well. More specifically, the undesired symbols downloaded within the kth round (the k-sums that do not include the desired message) are used as side information in the $(k+1)$th round. This exploitation of side information is performed by downloading $(k+1)$-sum consisting of 1 desired symbol and a k-sum of undesired symbols only that were generated in the kth round. However, the main difference from [12] is that for the nth database, if $s_n > k$, then this database does not exploit the side information generated in the kth round. Thus, the nth database belonging to the ℓth group exploits the side information generated in the kth round from all databases except itself if $s_n \leq k$. Moreover, for $s_n = k$, extra side information can be used in the nth database. This is because the user can form $n_0 \prod_{s\in\mathcal{S}\setminus\{s_n\}} \binom{M-2}{s-1}$ extra stages of side information by constructing k-sums of the undesired symbols in round 1 from the databases in group 0.

5) *Repeat* steps 2, 3, 4 after setting $k = k+1$ until $k = M$.

6) *Shuffling the order of the queries:* By shuffling the order of the queries uniformly, all possible queries can be made equally likely regardless of the message index. This guarantees the privacy.

2) Achievability Scheme for Non-Corner Points: In this section, we show that achievability schemes for non-corner points can be derived by time-sharing between the nearest corner points, i.e., the achievable scheme under $\boldsymbol{\tau}$ constraints is performed by time-sharing between the corner points of an N-dimensional polytope that enclose the traffic vector $\boldsymbol{\tau}$. The following lemma formalizes the time-sharing argument. Lemma 3 can be thought of as an adaptation of [16, Theorem 4] and [27, Lemma 1] to the PIR problem under asymmetric traffic constraints.

Lemma 3 (Time-sharing) *For the PIR problem under asymmetric traffic constraints $\boldsymbol{\tau}$, let the retrieval rate $R(\boldsymbol{\tau}_i)$ be achievable for the traffic ratio vector $\boldsymbol{\tau}_i$ for all $i \in \{1, \cdots, N\}$. Moreover, assume that $\boldsymbol{\tau} = \sum_{i=1}^N \alpha_i \boldsymbol{\tau}_i$ for some $\{\alpha_i\}_{i=1}^N$ such that $\alpha_i \geq 0$, for all i, and $\sum_{i=1}^N \alpha_i = 1$. Then, the following retrieval rate $R(\boldsymbol{\tau})$ is achievable,*

$$R(\boldsymbol{\tau}) = \sum_{i=1}^N \alpha_i R(\boldsymbol{\tau}_i) \tag{64}$$

Proof: Let PIR_i denote the PIR scheme that achieves retrieval rate $R(\boldsymbol{\tau}_i)$ for a traffic ratio vector $\boldsymbol{\tau}_i$. Denote the total download of PIR_i by D_i and the corresponding message length by L_i.

Now, construct the following PIR scheme with total download D and message length L. For each database, concatenate the queries from the N PIR schemes with ensuring that each symbol is queried by one PIR scheme only. Hence, $D = \sum_{i=1}^N D_i$, such that $D_i = \alpha_i D$, for $i \in \{1, \cdots, N\}$, and the download from the nth database is $t_n(\boldsymbol{\tau}) = \sum_{i=1}^N t_n(\boldsymbol{\tau}_i)$. This concatenation of the achievable schemes is feasible under asymmetric traffic constraints since $\boldsymbol{\tau} = \sum_{i=1}^N \alpha_i \boldsymbol{\tau}_i$. To see this, we note that the nth element of the traffic ratio vector τ_n is given by

$$\begin{aligned}\tau_n &= \frac{t_n(\boldsymbol{\tau})}{D} \\ &= \frac{\sum_{i=1}^N t_n(\boldsymbol{\tau}_i)}{D} \\ &= \frac{\sum_{i=1}^N \tau_n^{(i)} D_i}{D} = \frac{\sum_{i=1}^N \tau_n^{(i)} \alpha_i D}{D} = \sum_{i=1}^N \alpha_i \tau_n^{(i)}\end{aligned} \tag{65}$$

where $\tau_n^{(i)}$ denotes the nth element in $\boldsymbol{\tau}_i$. Since these implications are true for each element in $\boldsymbol{\tau}$, we have $\boldsymbol{\tau} = \sum_{i=1}^N \alpha_i \boldsymbol{\tau}_i$ as required.

PIR_i scheme downloads L_i symbols from the desired messages, such that

$$L_i = R(\boldsymbol{\tau}_i) D_i = \alpha_i R(\boldsymbol{\tau}_i) D \tag{66}$$

Hence, the total message length by concatenating all the achievable schemes together is

$$L = \sum_{i=1}^N L_i = \sum_{i=1}^N \alpha_i R(\boldsymbol{\tau}_i) D \tag{67}$$

and the corresponding achievable rate is given by

$$R(\boldsymbol{\tau}) = \frac{L}{D} = \sum_{i=1}^N \alpha_i R(\boldsymbol{\tau}_i) \tag{68}$$

The reliability constraint follows from the reliability of each PIR scheme. The privacy constraint is satisfied due to the

fact that each PIR scheme operates on a different portion of the messages and these portions are picked uniformly and independently. Hence, the privacy constraint for the concatenated scheme follows from the privacy of each PIR scheme. A formal treatment of the privacy constraint of concatenated schemes can be found in [16]. ■

Thus, Lemma 3 provides an achievability proof for any traffic ratio vector $\boldsymbol{\tau}$ that is not a corner point. Finally, we have the following remark regarding this time-sharing lemma.

Remark 11 *We note that although the vector* $\boldsymbol{\lambda} = (\lambda_1, \cdots, \lambda_N)$ *is in bijection with* $\boldsymbol{\tau} = (\tau_1, \cdots, \tau_N)$*, the time-sharing argument in Lemma 3 does not hold for* $R(\boldsymbol{\lambda})$*. This is due to the fact that* $R(\boldsymbol{\lambda})$ *is a non-linear function of* $\boldsymbol{\lambda}$ *whereas* $R(\boldsymbol{\tau})$ *is an affine function of* $\boldsymbol{\tau}$*.*

C. Decodability, Privacy, and Calculation of the Achievable Rate

In this section, we prove the decodability, privacy and the achievable rate in Theorem 2. We note that it suffices to consider the corner points only, as Lemma 3 settles the decodability, privacy and achievable rate for non-corner points based on the existence of feasible PIR schemes that achieve the corner points.

a) Decodability: By construction, in the $(k+1)$th round at the nth database, the user exploits the side information generated in the kth round in the remaining active databases by adding 1 symbol of the desired message with $(k-1)$-sum of undesired messages which was downloaded previously in the kth round. Moreover, for the nth database belonging to the ℓth group at the $(\ell+1)$th round, the user adds every ℓ symbols of the undesired symbols downloaded from group 0 to make one side information symbol. Since the user downloads $\prod_{\ell \in \mathcal{S}} \binom{M-2}{\ell-1}$ symbols from every database in the first n_0 databases (group 0), the user can exploit such side information to initiate $n_0 \prod_{\ell \in \mathcal{S}\setminus\{\ell\}} \binom{M-2}{\ell-1}$ stages in the $(\ell+1)$th round from every database in group ℓ. Since all side information symbols used in the $(k+1)$th round are decodable in the kth round or from round 1, the user cancels out these side information symbols and is left with symbols from the desired message.

b) Privacy: For every stage of the kth round initiated in the exploitation of the side information step, the user completes the stage by including all the remaining $\binom{M-1}{k-1}$ undesired symbols. This implies that all $\binom{M}{k}$ combinations of the k-sum are included at each round. Thus, the structure of the queries is the same for any desired message. The privacy constraint in (7) is satisfied by the random and independent permutation of each message and the random shuffling of the order of the queries. This ensures that all queries are equally likely independent of the desired message index.

c) Calculation of the Achievable Rate: For a corner point characterized by the non-decreasing sequence $\{n_i\}_{i=0}^{M-1}$, as mentioned before, we denote $y_\ell[k]$ to be the number of stages that admit k-sums downloaded from any database belonging to the ℓth group, i.e., the nth database such that $n_{\ell-1}+1 \le n \le n_\ell$. By construction, we observe that all databases belonging to the ℓth group are inactive until the $(\ell+1)$th round as the user initiates download in such databases by exploiting ℓ bits of side information simultaneously by definition of the group. Consequently, we have the initial condition $y_\ell[k] = 0$ for $k \le \ell$. Since the user downloads $\prod_{s\in\mathcal{S}} \binom{M-2}{s-1}$ individual symbols (i.e., from round 1) from group 0, we have the following initial condition $y_0[1] = \prod_{s\in\mathcal{S}} \binom{M-2}{s-1}$.

Now, we note from the side information exploitation step that the user initiates new stages in the kth round from the nth database depending on the number of stages of the $(k-1)$th round for group 0 and group 1 (i.e., for $1 \le n \le n_1$). More specifically, for the nth database belonging to group 0, the user considers all the undesired symbols downloaded from all databases (except the nth database) in the $(k-1)$th round as side information. Since database symmetry applies over each group, and from the fact that each stage in the $(k-1)$th round initiates a stage in the kth round, we have

$$y_0[k] = (n_0 - 1) y_0[k-1] + \sum_{\ell \in \mathcal{S}\setminus\{0\}} (n_\ell - n_{\ell-1}) y_\ell[k-1] \tag{69}$$

where the left side is the total number of stages in the $(k-1)$th round from all the $N-1$ databases (i.e., except for the nth database that belongs to group 0). The same argument holds for group 1 as well, hence

$$y_1[k] = (n_1 - n_0 - 1) y_1[k-1] + \sum_{\ell \in \mathcal{S}\setminus\{1\}} (n_\ell - n_{\ell-1}) y_\ell[k-1] \tag{70}$$

where $(n_1 - n_0 - 1)$ denotes the number of databases in group 1 other than the nth database.

For a database belonging to the ℓth group such that $\ell \ge 2$, the user can generate extra stages by exploiting the symbols downloaded in round 1. To initiate one stage in the $(\ell+1)$th round, the user needs to combine symbols from $\frac{\binom{M-1}{\ell}^{\ell}}{M-1} = \binom{M-2}{\ell-1}$ stages. Therefore, the number of stages initiated in the $(\ell+1)$th round as a consequence of the side information in round 1 is $\xi_\ell = \frac{y_0[1]}{\binom{M-2}{\ell-1}} = \prod_{s\in\mathcal{S}\setminus\{\ell\}} \binom{M-2}{s-1}$. Since these extra side information can be used once (at the $(\ell+1)$th round only) and after that for the kth round, the database exploits the side information generated in the $(k-1)$th round only. We represent this one-time exploitation of side information in the $(\ell+1)$th round by the Kronecker delta function $\delta[k-\ell-1]$. Consequently, the number of stages for the ℓth group, $\ell \ge 2$ is related via the following difference equation:

$$y_\ell[k] = n_0 \xi_\ell \delta[k-\ell-1] + (n_\ell - n_{\ell-1} - 1) y_\ell[k-1] + \sum_{j \in \mathcal{S}\setminus\{\ell\}} (n_j - n_{j-1}) y_j[k-1] \tag{71}$$

Now, we are ready to characterize $\boldsymbol{\tau}(\mathbf{n})$ and $R(\boldsymbol{\tau}(\mathbf{n}))$ in terms of $y_\ell[k]$, where $\ell \in \mathcal{S}$ and $k = 1, \cdots, M$. For any stage in the kth round, the user downloads $\binom{M-1}{k-1}$ desired symbols from a total of $\binom{M}{k}$ downloads. Therefore, from a database belonging to the ℓth group, the user

downloads $\sum_{k=1}^{M}\binom{M-1}{k-1}y_\ell[k]$ desired symbols from a total of $\sum_{k=1}^{M}\binom{M}{k}y_\ell[k]$. The number of databases belonging to the ℓth group is given by $n_\ell - n_{\ell-1}$. Therefore, the total download is given by,

$$\sum_{n=1}^{N} t_n(\boldsymbol{\tau}(\mathbf{n})) = \sum_{\ell\in\mathcal{S}}\sum_{k=1}^{M}\binom{M}{k}y_\ell[k](n_\ell - n_{\ell-1}) \tag{72}$$

Thus, the traffic ratio of the nth database belonging to the ℓth group (i.e., $n_{\ell-1}+1 \le n \le n_\ell$) corresponding to $\mathbf{n} = \{n_i\}_{i=0}^{M-1}$ is given by

$$\tau_n(\mathbf{n}) = \tilde{\tau}_\ell = \frac{\sum_{k=1}^{M}\binom{M}{k}y_\ell[k]}{\sum_{\ell\in\mathcal{S}}\sum_{k=1}^{M}\binom{M}{k}y_\ell[k](n_\ell - n_{\ell-1})} \tag{73}$$

where $n_{\ell-1}+1 \le n \le n_\ell$. Furthermore, the total desired symbols from all databases is given by

$$L(\boldsymbol{\tau}(\mathbf{n})) = \sum_{\ell\in\mathcal{S}}\sum_{k=1}^{M}\binom{M-1}{k-1}y_\ell[k](n_\ell - n_{\ell-1}) \tag{74}$$

which further leads to the following achievable rate

$$R(\boldsymbol{\tau}(\mathbf{n})) = \frac{\sum_{\ell\in\mathcal{S}}\sum_{k=1}^{M}\binom{M-1}{k-1}y_\ell[k](n_\ell - n_{\ell-1})}{\sum_{\ell\in\mathcal{S}}\sum_{k=1}^{M}\binom{M}{k}y_\ell[k](n_\ell - n_{\ell-1})} \tag{75}$$

VI. Optimality of $M = 2$ and $M = 3$ Cases

In this section, we prove Corollary 2, i.e., we prove that the capacity of the PIR problem under asymmetric traffic constraints $C(\boldsymbol{\tau})$ for $M = 2, 3$ is given by (22). We note that since the upper bound in Theorem 1 is affine in $\boldsymbol{\tau}$ and time-sharing rates are achievable from Lemma 3, it suffices to prove the optimality of all corner points to settle the PIR capacity $C(\boldsymbol{\tau})$ for $M = 2, 3$. In the following, we use Theorem 1 and Theorem 2 to show the optimality of these corner points.

A. $M = 2$ Messages

We start the proof from the achievability side. From Theorem 2, the corner points are specified by the non-decreasing sequence $\mathbf{n} = (n_0, n_1)$. In this case, the system of difference equations in (16) is reduced to

$$y_0[k] = (n_0 - 1)y_0[k-1] \tag{76}$$

$$y_1[k] = n_0 y_0[k-1] \tag{77}$$

for $k = 1, 2$, where $y_0[1] = 1$, and $y_1[1] = 0$. Hence, $y_0[2] = n_0 - 1$, and $y_1[2] = n_0$. Hence, the total downloads for the corner point $\mathbf{n} = (n_0, n_1)$ is

$$\begin{aligned}\sum_{n=1}^{N} t_n(\boldsymbol{\tau}(\mathbf{n})) &= \sum_{\ell=0}^{1}\sum_{k=1}^{2}\binom{2}{k}y_\ell[k](n_\ell - n_{\ell-1}) \\ &= n_0(n_1+1)\end{aligned} \tag{78}$$

Thus, from Theorem 2, the traffic-ratio vector $\boldsymbol{\tau}(\mathbf{n})$ is given by

$$\tilde{\tau}_0 = \frac{\binom{2}{1}y_0[1] + \binom{2}{2}y_0[2]}{\sum_{n=1}^{N} t_n(\boldsymbol{\tau}(\mathbf{n}))} = \frac{n_0+1}{n_0(n_1+1)} \tag{79}$$

$$\tilde{\tau}_1 = \frac{\binom{2}{1}y_1[1] + \binom{2}{2}y_1[2]}{\sum_{n=1}^{N} t_n(\boldsymbol{\tau}(\mathbf{n}))} = \frac{1}{n_1+1} \tag{80}$$

where $\tau_n = \tilde{\tau}_0$, for $1 \le n \le n_0$, and $\tau_n = \tilde{\tau}_1$, for $n_0 + 1 \le n \le n_1$, and $\tau_n = 0$ otherwise. For the desired symbols, the user downloads $L_0(\boldsymbol{\tau}(\mathbf{n}))$ symbols from the nth database when $1 \le n \le n_0$, and $L_1(\boldsymbol{\tau}(\mathbf{n}))$ symbols from the nth database when $n_0 + 1 \le n \le n_1$

$$L_0(\boldsymbol{\tau}(\mathbf{n})) = y_0[1] + y_0[2] = n_0 \tag{81}$$

$$L_1(\boldsymbol{\tau}(\mathbf{n})) = y_1[1] + y_1[2] = n_0 \tag{82}$$

Consequently, $L = n_0 L_0 + (n_1 - n_0)L_1 = n_0 n_1$, and the achievable retrieval rate $R(\boldsymbol{\tau}(\mathbf{n}))$ is given by

$$R(\boldsymbol{\tau}(\mathbf{n})) = \frac{L(\boldsymbol{\tau}(\mathbf{n}))}{\sum_{n=1}^{N} t_n(\boldsymbol{\tau}(\mathbf{n}))} = \frac{n_1}{n_1+1} \tag{83}$$

For the converse, we evaluate the bound in (10) (without the minimization) for $n_1 = n_0$, i.e., we substitute with n_0 in the argument of the upper bound. Then, we have the following upper bound

$$R(\boldsymbol{\tau}(\mathbf{n})) \le \frac{1 + \frac{\sum_{n=n_0+1}^{N}\tau_n}{n_0}}{1 + \frac{1}{n_0}} \tag{84}$$

$$= \frac{1 + \frac{(n_1-n_0)\tilde{\tau}_1}{n_0}}{1 + \frac{1}{n_0}} \tag{85}$$

$$= \frac{n_1}{n_1+1} \tag{86}$$

This concludes the optimality of our achievable scheme for $M = 2$.

B. $M = 3$ Messages

Similarly, for the corner point specified by the non-decreasing sequence $\mathbf{n} = (n_0, n_1, n_2)$, we have the following system of difference equations for $k = 1, 2, 3$

$$\begin{aligned}y_0[k] = {} & (n_0 - 1)y_0[k-1] + (n_1 - n_0)y_1[k-1] \\ & + (n_2 - n_1)y_2[k-1]\end{aligned} \tag{87}$$

$$\begin{aligned}y_1[k] = {} & n_0 y_0[k-1] + (n_1 - n_0 - 1)y_1[k-1] \\ & + (n_2 - n_1)y_2[k-1]\end{aligned} \tag{88}$$

$$\begin{aligned}y_2[k] = {} & n_0\delta[k-3] + n_0 y_0[k-1] + (n_1 - n_0)y_1[k-1] \\ & + (n_2 - n_1 - 1)y_2[k-1]\end{aligned} \tag{89}$$

with the initial conditions $y_0[1] = 1$, $y_1[1] = 0$, and $y_2[1] = y_2[2] = 0$. Evaluating $y_\ell[k]$, for $\ell = 0, 1, 2$, and $k = 1, 2, 3$ recursively leads to $y_0[2] = n_0 - 1$, $y_1[2] = n_0$, $y_0[3] = n_1 n_0 - 2n_0 + 1$, $y_1[3] = n_1 n_0 - 2n_0$, and $y_2[3] = n_1 n_0$. This leads to the following total download

$$\begin{aligned}\sum_{n=1}^{N} t_n(\boldsymbol{\tau}(\mathbf{n})) &= \sum_{\ell=0}^{2}\sum_{k=1}^{3}\binom{3}{k}y_\ell[k](n_\ell - n_{\ell-1}) \\ &= n_0(n_1 n_2 + n_1 + 1)\end{aligned} \tag{90}$$

The sequence $\mathbf{n} = (n_0, n_1, n_2)$ specifies the traffic ratio vector $\boldsymbol{\tau}(\mathbf{n})$ such that

$$\tilde{\tau}_0 = \frac{n_0 n_1 + n_0 + 1}{n_0(n_2 n_1 + n_1 + 1)} \tag{91}$$

$$\tilde{\tau}_1 = \frac{n_1 + 1}{n_2 n_1 + n_1 + 1} \tag{92}$$

$$\tilde{\tau}_2 = \frac{n_1}{n_2 n_1 + n_1 + 1} \tag{93}$$

where $\tau_n = \tilde{\tau}_0$ for $1 \leq n \leq n_0$, $\tau_n = \tilde{\tau}_1$ for $n_0 + 1 \leq n \leq n_1$, $\tau = \tilde{\tau}_2$ for $n_1 + 1 \leq n \leq n_2$, and $\tau_n = 0$ otherwise.

For the desired symbols, the user downloads $L_0(\boldsymbol{\tau}(\mathbf{n}))$ symbols from the nth database if $1 \leq n \leq n_0$, $L_1(\boldsymbol{\tau}(\mathbf{n}))$ symbols if $n_0 + 1 \leq n \leq n_1$, and $L_2(\boldsymbol{\tau}(\mathbf{n}))$ symbols if $n_1 + 1 \leq n \leq n_2$, hence

$$L_\ell(\boldsymbol{\tau}(\mathbf{n})) = \sum_{k=1}^{3} \binom{2}{k-1} y_\ell[k] = n_0 n_1, \quad \ell = 0, 1, 2 \tag{94}$$

Consequently, the following rate is achievable

$$R(\boldsymbol{\tau}(\mathbf{n})) = \frac{n_1 n_2}{n_1 n_2 + n_1 + 1} \tag{95}$$

For the converse, pick (n_1, n_2) in the converse bound to be (n_0, n_1), which leads to the following bound

$$R(\boldsymbol{\tau}(\mathbf{n})) \leq \frac{1 + \frac{\sum_{n=n_0+1}^{N} \tau_n}{n_0} + \frac{\sum_{n=n_1+1}^{N} \tau_n}{n_0 n_1}}{1 + \frac{1}{n_0} + \frac{1}{n_0 n_1}} \tag{96}$$

$$= \frac{1 + \frac{(n_1 - n_0)\tilde{\tau}_1}{n_0} + \frac{(n_2 - n_1)\tilde{\tau}_2}{n_0} + \frac{(n_2 - n_1)\tilde{\tau}_2}{n_0 n_1}}{1 + \frac{1}{n_0} + \frac{1}{n_0 n_1}} \tag{97}$$

$$= \frac{n_1 n_2}{n_1 n_2 + n_1 + 1} \tag{98}$$

This concludes the optimality of our achievable scheme for $M = 3$.

Remark 12 *We note that, surprisingly, for the corner points of the cases $M = 2$ and $M = 3$, the number of desired symbols downloaded from each active database is the same irrespective to the traffic ratio of the database; see* (81)-(82) *for $M = 2$ and* (94) *for $M = 3$. This suggests that at these corner points, the optimal scheme performs* combinatorial water-filling *for the undesired symbols first, i.e., the nth active database downloads $t_n - n_0$ undesired symbols for $M = 2$ and $t_n - n_0 n_1$ undesired symbols for $M = 3$, and then downloads the same number of desired symbols from all active databases.*

VII. Achievable Tradeoff for $N = 2$ and Arbitrary M

For the special case of $N = 2$, and an arbitrary M, the retrieval rate calculation in Theorem 2 is significantly simplified. Let $s_2 \in \{0, \cdots, M-1\}$ be the number of side information symbols that are used simultaneously in the initial round of download at the second database. Note that there is a bijection between s_2 and the non-decreasing sequence $\mathbf{n}$ as $n_0 = n_1 = \cdots = n_{s_2 - 1} = 1$, and $n_{s_2} = 2$ for any corner point other than the corner point corresponding to the trivial scheme of downloading the contents of the first database.

The user starts with downloading $\binom{M-2}{s_2-1}$ stages of individual symbols (i.e., the user downloads $M\binom{M-2}{s_2-1}$ symbols in round 1 from all messages) from the first database to create 1 stage in the $(s_2 + 1)$th round. After the initial exploitation of side information, the two databases exchange side information. More specifically, from database 1 in the $(s_2 + 2k)$th round, where $k = 1, \cdots, \left\lfloor \frac{M - s_2}{2} \right\rfloor$, the user exploits the side information generated in database 2 in the $(s_2 + 2k - 1)$th round to download $\binom{M-1}{s_2+2k-1}$ desired symbols (by adding one symbol of the desired symbols to the $(s_2 + 2k - 1)$-sum of undesired symbols generated in database 2) from total download in the $(s_2 + 2k)$th round of $\binom{M}{s_2+2k}$. Similarly from database 2, in the $(s_2 + 2k + 1)$th round, where $k = 0, \cdots, \left\lfloor \frac{M - s_2 - 1}{2} \right\rfloor$, the user exploits the side information generated in database 1 in the $(s_2 + 2k)$th round, and downloads $\binom{M-1}{s_2+2k}$ desired symbols from total of $\binom{M}{s_2+2k+1}$ downloads in the $(s_2 + 2k + 1)$th round.

Consequently, we have

$$t_1(s_2) = M\binom{M-2}{s_2-1} + \sum_{k=1}^{\left\lfloor \frac{M-s_2}{2} \right\rfloor} \binom{M}{s_2+2k} \tag{99}$$

$$t_2(s_2) = \sum_{k=0}^{\left\lfloor \frac{M-s_2-1}{2} \right\rfloor} \binom{M}{s_2+2k+1} \tag{100}$$

which further leads to the following total download

$$t_1(s_2) + t_2(s_2) = M\binom{M-2}{s_2-1} + \sum_{k=0}^{M-s_2-1} \binom{M}{s_2+k+1} \tag{101}$$

Thus, the traffic ratio $\tau_2(s_2)$ is given by

$$\tau_2(s_2) = \frac{t_2(s_2)}{t_1(s_2) + t_2(s_2)} = \frac{\sum_{k=0}^{\left\lfloor \frac{M-s_2-1}{2} \right\rfloor} \binom{M}{s_2+2k+1}}{M\binom{M-2}{s_2-1} + \sum_{k=0}^{M-s_2-1} \binom{M}{s_2+k+1}} \tag{102}$$

The total number of desired symbols is given by

$$L(s_2) = \binom{M-2}{s_2-1} + \sum_{k=1}^{\left\lfloor \frac{M-s_2}{2} \right\rfloor} \binom{M-1}{s_2+2k-1} + \sum_{k=0}^{\left\lfloor \frac{M-s_2-1}{2} \right\rfloor} \binom{M-1}{s_2+2k} \tag{103}$$

$$= \binom{M-2}{s_2-1} + \sum_{k=0}^{M-s_2-1} \binom{M-1}{s_2+k} \tag{104}$$

Thus, the following rate is achievable for $N = 2$ and arbitrary M

$$R(s_2) = \frac{L(s_2)}{t_1(s_2) + t_2(s_2)} = \frac{\binom{M-2}{s_2-1} + \sum_{k=0}^{M-s_2-1} \binom{M-1}{s_2+k}}{M\binom{M-2}{s_2-1} + \sum_{k=0}^{M-s_2-1} \binom{M}{s_2+k+1}} \tag{105}$$

VIII. Further Examples

In this section, we present further examples to clarify the achievable scheme for some additional tractable values of M, N.

TABLE VI
THE QUERY TABLE FOR $M = 4$, $N = 2$, $s_2 = 1$ (CORRESPONDING TO $\tau_2 = \frac{7}{15}$)

Database 1	Database 2
a_1, b_1, c_1, d_1	
	$a_2 + b_1$ $a_3 + c_1$ $a_4 + d_1$ $b_2 + c_2$ $b_3 + d_2$ $c_3 + d_3$
$a_5 + b_2 + c_2$ $a_6 + b_3 + d_2$ $a_7 + c_3 + d_3$ $b_4 + c_4 + d_4$	
	$a_8 + b_4 + c_4 + d_4$

TABLE VII
THE QUERY TABLE FOR $M = 4$, $N = 2$, $s_2 = 2$ (CORRESPONDING TO $\tau_2 = \frac{4}{13}$)

Database 1	Database 2
a_1, b_1, c_1, d_1 a_2, b_2, c_2, d_2	
	$a_3 + b_1 + c_1$ $a_4 + b_2 + d_1$ $a_5 + c_2 + d_2$ $b_3 + c_3 + d_3$
$a_6 + b_3 + c_3 + d_3$	

TABLE VIII
THE QUERY TABLE FOR $M = 4$, $N = 2$, $s_2 = 3$ (CORRESPONDING TO $\tau_2 = \frac{1}{5}$).

Database 1	Database 2
a_1, b_1, c_1, d_1	
	$a_2 + b_1 + c_1 + d_1$

A. $M = 4$ Messages, $N = 2$ Databases

In this example, we show that the achievable rate $R(\tau_2)$ does not match the upper bound $\bar{C}(\tau_2)$ for all traffic ratios τ_2. For $M = 4$, we have $M + 1 = 5$ corner points, corresponding to $s_2 = \{0, 1, 2, 3\}$ and another corner point corresponding to the trivial scheme of downloading the contents of database 1. Let a_i, b_i, c_i, d_i denote the randomly permuted symbols of messages W_1, W_2, W_3, W_4, respectively. Then, $R(0) = \frac{1}{4}$ by trivially downloading a_1, b_1, c_1, d_1 from database 1. In addition, $R(\frac{1}{2}) = \frac{1-\frac{1}{2}}{1-(\frac{1}{2})^4} = \frac{8}{15}$ using the symmetric scheme in [12].

1) Corner Point $s_2 = 1$: (See the query table in Table VI.) The user uses 1 bit of side information in database 2, hence the user starts downloading from round 2 (that admits 2-sums). The user exploits the side information generated in round 1 by downloading $a_2 + b_1$, $a_3 + c_1$, and $a_4 + d_1$. The user completes the stage by downloading undesired symbols consisting of 2-sums that do not include a_i, hence the user downloads $b_2 + c_2$, $b_3 + d_2$, $c_3 + d_3$. The undesired symbols are exploited in database 1, thus the user downloads $a_5 + b_2 + c_2$, $a_6 + b_3 + d_2$, and $a_7 + c_3 + d_3$. The user completes the stage by downloading $b_4 + c_4 + d_4$, which can be exploited in database 2 by downloading $a_8 + b_4 + c_4 + d_4$. In this case, the user downloads 8 symbols from database 1 and 7 symbols from database 2, hence we have $\tau_2 = \frac{7}{15}$. Since the user downloads $L = 8$ desired symbols, the achievable rate $R(\frac{7}{15}) = \frac{8}{15}$.

2) Corner Point $s_2 = 2$: (See the query table in Table VII.) The user downloads $\binom{M-2}{s_2-1} = 2$ stages of individual symbols (1-sum) from database 1, so that the user forms 2-sums that can be used in database 2 as side information to start round 3 directly, i.e., by forming 2-sums as side information from the individual symbols, the user effectively skips round 2. More specifically, the user downloads $a_3 + b_1 + c_1$, $a_4 + b_2 + d_1$, $a_5 + c_2 + d_2$ from database 2 taking into considerations that all these undesired symbols are decodable from database 1. The user completes the stage by downloading $b_3 + c_3 + d_3$ that can be further exploited in database 1 by downloading $a_6 + b_3 + c_3 + d_3$. In this case, the user downloads 9 symbols from database 1 and 4 symbols from database 2, therefore $\tau_2 = \frac{4}{13}$. The user downloads $L = 6$ desired symbols, thus, $R(\frac{4}{13}) = \frac{6}{13}$.

3) Corner Point $s_2 = 3$: (See the query table in Table VIII.) In this case, the user skips rounds 2, 3 and jumps directly to round 4 at database 2. Therefore, the user downloads $a_2 + b_1 + c_1 + d_1$ from database 2, which uses $b_1 + c_1 + d_1$ as side information which is decodable from database 1. Thus, we have $\tau_2 = \frac{1}{5}$, and the corresponding rate $R(\frac{1}{5}) = \frac{2}{5}$.

4) Comparison with the Upper Bound: The upper bound in Theorem 1 can be explicitly expressed as:

$$R(\tau_2) \leq \begin{cases} \frac{1}{4} + \frac{3\tau_2}{4}, & 0 \leq \tau_2 \leq \frac{1}{5} \\ \frac{2}{7} + \frac{4\tau_2}{7}, & \frac{1}{5} \leq \tau_2 \leq \frac{3}{8} \\ \frac{4}{11} + \frac{4\tau_2}{11}, & \frac{3}{8} \leq \tau_2 \leq \frac{7}{15} \\ \frac{8}{15}, & \frac{7}{15} \leq \tau_2 \leq \frac{1}{2} \end{cases} \tag{106}$$

We observe that for all the corner points of the achievable scheme, the upper and lower bounds match. However, the upper bound has an extra corner point $(\frac{3}{8}, \frac{1}{2})$ which is not achievable using time-sharing. This is illustrated in Fig. 6

B. $M = 3$ Messages, $N = 3$ Databases

In this example, we show the capacity-achieving scheme for $M = 3$, $N = 3$ (the capacity region is illustrated in Fig. 3 as a function of $C(\lambda_2, \lambda_3)$). Let a_i, b_i, c_i denote the permuted symbols of messages W_1, W_2, W_3, respectively. We show here only the query tables for achieving non-trivial corner points. In this case, we have $\binom{M+N-1}{M} = 10$ corner points corresponding to non-decreasing sequences (n_0, n_1, n_2).

For the pair $(\tau_2, \tau_3) = (0, 0)$, the achievable scheme is the trivial scheme that downloads a_1, b_1, c_1 from the first database only achieving $R(0, 0) = \frac{1}{3}$. For the corner point $(\frac{1}{4}, 0)$, this

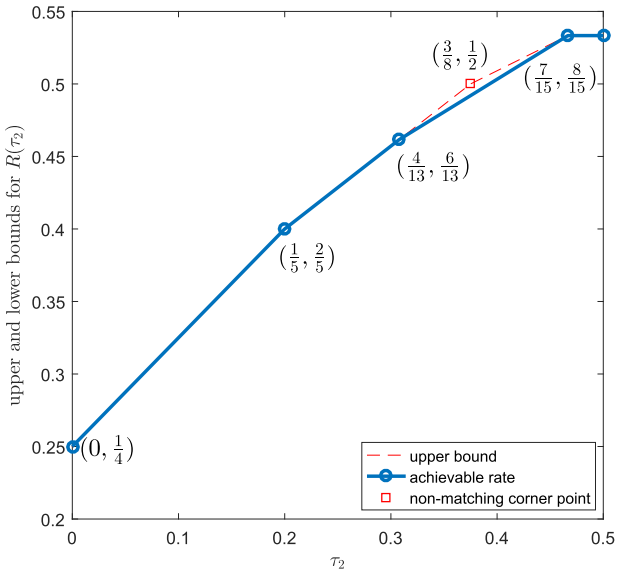

Fig. 6. Upper and lower bounds for $R(\tau_2)$ for $M = 4$, $N = 2$.

is exactly the same corner point presented in Section V-A.1 (for $\lambda_2 = \frac{1}{3}$) as $\tau_3 = 0$, which effectively reduces the problem to $N = 2$ databases. The achievable scheme for this corner point is illustrated in Table IV, hence $R(\frac{1}{4}, 0) = \frac{1}{2}$. For the corner point $(\frac{3}{7}, 0)$, again this point reduces to 2 databases. The achievable scheme is given in Table III, and $R(\frac{3}{7}, 0) = \frac{4}{7}$. For the corner point $(\frac{1}{3}, \frac{1}{3})$, which is the symmetric-traffic point, the achievable scheme is the symmetric scheme in [12], which achieves $R(\frac{1}{3}, \frac{1}{3}) = \frac{9}{13}$. For the corner point $(\frac{1}{2}, 0)$, we can apply the symmetric achievable scheme for $N = 2$ databases only as $\tau_3 = 0$ in this case, hence $R(\frac{1}{2}, 0) = \frac{4}{7}$.

Now, we focus on the non-trivial corner points. As mentioned previously, the pair (s_2, s_3) is in bijection with the sequence (n_0, n_1, n_2). Therefore, we enumerate the remaining cases using the pair (s_2, s_3).

1) Corner Point $(s_2, s_3) = (0, 1)$*:* In this case, the user does not use the side information generated in database 1 within the initial download of database 2 ($s_2 = 0$), hence the user downloads new individual symbols from database 2. The user uses 1 bit of side information in database 3 in its round of download (round 2). These side information symbols come from database 1 and database 2. The query table for this case is shown in Table IX. In this case, we have $(\tau_2, \tau_3) = (\frac{9}{26}, \frac{4}{13})$, and the achievable rate is $R(\frac{9}{26}, \frac{4}{13}) = \frac{9}{13}$.

2) Corner Point $(s_2, s_3) = (0, 2)$*:* The user does not exploit the side information generated from database 1 in the first round of download at database 2. The user uses 2 side information symbols simultaneously in the initial round (round 3) of download at database 3. Note that in round 3 database 3 receives side information from rounds 1 and 2 of databases 1 and 2. The query table for this case is shown in Table X. In this case, we have $(\tau_2, \tau_3) = (\frac{7}{18}, \frac{2}{9})$, and the achievable rate is $R(\frac{7}{18}, \frac{2}{9}) = \frac{2}{3}$.

3) Corner Point $(s_2, s_3) = (1, 1)$*:* In this case, both databases 2 and 3 exploit the side information generated from database 1 in their initial round of download (round 1).

TABLE IX
THE QUERY TABLE FOR $M = 3$, $N = 3$, $(s_2, s_3) = (0, 1)$ (I.E., $(\tau_2, \tau_3) = (\frac{9}{26}, \frac{4}{13})$).

Database 1	Database 2	Database 3
a_1, b_1, c_1	a_2, b_2, c_2	
$a_3 + b_2$ $a_4 + c_2$ $b_3 + c_3$	$a_5 + b_1$ $a_6 + c_1$ $b_4 + c_4$	$a_7 + b_1$ $a_8 + c_1$ $b_5 + c_5$ $a_9 + b_2$ $a_{10} + c_2$ $b_6 + c_6$
$a_{11} + b_4 + c_4$ $a_{12} + b_5 + c_5$ $a_{13} + b_6 + c_6$	$a_{14} + b_3 + c_3$ $a_{15} + b_5 + c_5$ $a_{16} + b_6 + c_6$	$a_{17} + b_3 + c_3$ $a_{18} + b_4 + c_4$

TABLE X
THE QUERY TABLE FOR $M = 3$, $N = 3$, $(s_2, s_3) = (0, 2)$ (I.E., $(\tau_2, \tau_3) = (\frac{7}{18}, \frac{2}{9})$)

Database 1	Database 2	Database 3
a_1, b_1, c_1	a_2, b_2, c_2	
$a_3 + b_2$ $a_4 + c_2$ $b_3 + c_3$	$a_5 + b_1$ $a_6 + c_1$ $b_4 + c_4$	
$a_7 + b_4 + c_4$	$a_8 + b_3 + c_3$	$a_9 + b_1 + c_1$ $a_{10} + b_2 + c_2$ $a_{11} + b_3 + c_3$ $a_{12} + b_4 + c_4$

The query table for this case is shown in Table XI. In this case, we have $(\tau_2, \tau_3) = (\frac{4}{13}, \frac{4}{13})$, and the achievable rate is $R(\frac{4}{13}, \frac{4}{13}) = \frac{9}{13}$.

4) Corner Point $(s_2, s_3) = (1, 2)$*:* In this case[4], database 2 exploits 1 side information in its initial download (round 2), while database 3 skips to round 3 directly. Database 3 receives side information from the round 1 of database 1 and round 2 of database 2. The query table for this case is shown in Table XII. In this case, we have $(\tau_2, \tau_3) = (\frac{1}{3}, \frac{2}{9})$, and the achievable rate is $R(\frac{1}{3}, \frac{2}{9}) = \frac{2}{3}$.

5) Corner Point $(s_2, s_3) = (2, 2)$*:* Both databases 2 and 3 skip round 1 and 2 of downloads and go directly to round 3, in which they exploits 2 side information symbols simultaneously. The query table for this case is shown in

[4]We use this case to clarify the differences between the non-decreasing sequences in the achievability and the converse proofs. In this case $\mathbf{n} = (1, 2, 3)$. From he achievability point of view, this means that database 1 ($n_0 = 1$) is the last database that does not use side information from other databases, database 2 ($n_1 = 2$) is the last database that uses 1 side information symbol, and database 3 ($n_2 = 3$) is the last database that uses a 2-sum of side information symbols in the first round of download. From the converse point of view, if we condition on one message (say W_3), then the user downloads $(a_1, b_1, a_4 + b_2)$ from database 1, $(a_3, b_2, a_2 + b_1)$ from database 2, and $(a_5 + b_1, a_6 + b_2)$ from database 3. That means that there are ($n_1 = 2$) databases that use the symmetric scheme of [12] when the number of messages is reduced to two. If we further condition on W_1, the user downloads (a_1, a_4) from database 1, (a_3, a_2) from database 2, and (a_5, a_6) from database 3, which means that there are ($n_2 = 3$) databases that apply the optimal symmetric scheme if the number of messages is decreased to one.

TABLE XI
THE QUERY TABLE FOR $M = 3$, $N = 3$, $(s_2, s_3) = (1, 1)$ (I.E., $(\tau_2, \tau_3) = (\frac{4}{13}, \frac{4}{13})$)

Database 1	Database 2	Database 3
a_1, b_1, c_1		
	$a_2 + b_1$ $a_3 + c_1$ $b_2 + c_2$	$a_4 + b_1$ $a_5 + c_1$ $b_3 + c_3$
$a_6 + b_2 + c_2$ $a_7 + b_3 + c_3$	$a_8 + b_3 + c_3$	$a_9 + b_2 + c_2$

TABLE XII
THE QUERY TABLE FOR $M = 3$, $N = 3$, $(s_2, s_3) = (1, 2)$ (I.E., $(\tau_2, \tau_3) = (\frac{1}{3}, \frac{2}{9})$)

Database 1	Database 2	Database 3
a_1, b_1, c_1		
	$a_2 + b_1$ $a_3 + c_1$ $b_2 + c_2$	
$a_4 + b_2 + c_2$		$a_5 + b_1 + c_1$ $a_6 + b_2 + c_2$

TABLE XIII
THE QUERY TABLE FOR $M = 3$, $N = 3$, $(s_2, s_3) = (2, 2)$ (I.E., $(\tau_2, \tau_3) = (\frac{1}{5}, \frac{1}{5})$)

Database 1	Database 2	Database 3
a_1, b_1, c_1		
	$a_2 + b_1 + c_1$	$a_3 + b_1 + c_1$

Table XIII. In this case, we have $(\tau_2, \tau_3) = (\frac{1}{5}, \frac{1}{5})$, and the achievable rate is $R(\frac{1}{5}, \frac{1}{5}) = \frac{3}{5}$.

IX. CONCLUSION

In this paper, we introduced the PIR problem under asymmetric traffic constraints $\boldsymbol{\tau}$. We investigated the fundamental limits of this problem by developing the novel upper bound $\bar{C}(\boldsymbol{\tau})$. The upper bound generalizes the converse proof in [12], which inherently utilizes database symmetry. The upper bound is a piece-wise affine function in $\boldsymbol{\tau}$. The upper bound implies a strict capacity loss due to the asymmetric traffic constraints for certain cases. We developed explicit achievable schemes for $\binom{M+N-1}{M}$ corner points, and achieved the remaining points by time-sharing. We described the achievable scheme by means of a system of difference equations. We explicitly derived the achievable rate for $N = 2$ and arbitrary M. We proved that the upper bound and the lower bound exactly match for every $\boldsymbol{\tau}$ for the cases of $M = 2$ and $M = 3$ for any N.

It is worth noting that for general M, which is not equal to $2, 3$, the problem is open from both sides (achievability and converse). However, focusing on the achievability side, one can think about different round skipping techniques other than just skipping the first rounds. It is unclear how to exploit side information in this case though. We see the problem of closing the gap for general M in our setting as a central problem as it relates to the cache-aided PIR problems [30], [36], PIR from wiretap channel II [37], and noisy PIR [38]. Consequently, closing the gap in our problem solves the other problems almost directly.

REFERENCES

[1] B. Chor, E. Kushilevitz, O. Goldreich, and M. Sudan, "Private information retrieval," *J. ACM*, vol. 45, no. 6, pp. 965–981, 1998.

[2] W. Gasarch, "A survey on private information retrieval," *Bull. EATCS*, vol. 82, p. 113, Feb. 2004.

[3] C. Cachin, S. Micali, and M. Stadler, "Computationally private information retrieval with polylogarithmic communication," in *Proc. Int. Conf. Theory Appl. Cryptograph. Techn.* Berlin, Germany: Springer, 1999, pp. 402–414.

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

[5] S. Yekhanin, "Private information retrieval," *Commun. ACM*, vol. 53, no. 4, pp. 68–73, Apr. 2010.

[6] 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 (ISIT)*, Jun./Jul. 2014, pp. 856–860.

[7] G. Fanti and K. Ramchandran, "Efficient private information retrieval over unsynchronized databases," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 7, pp. 1229–1239, Oct. 2015.

[8] 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.

[9] A. Fazeli, A. Vardy, and E. Yaakobi, "Codes for distributed PIR with low storage overhead," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2015, pp. 2852–2856.

[10] R. Tajeddine and S. El Rouayheb, "Private information retrieval from MDS coded data in distributed storage systems," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 1411–1415.

[11] H. Sun and S. A. Jafar, "Blind interference alignment for private information retrieval," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 560–564.

[12] 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.

[13] 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.

[14] H. Sun and S. A. Jafar, "The capacity of symmetric private information retrieval," *IEEE Trans. Inf. Theory*, vol. 65, no. 1, pp. 322–329, Jan. 2019.

[15] 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.

[16] H. Sun and S. A. Jafar, "Optimal download cost of private information retrieval for arbitrary message length," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 12, pp. 2920–2932, Dec. 2017.

[17] 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.

[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] 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, 2017.

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

[21] R. Tajeddine, O. W. Gnilke, D. Karpuk, R. Freij-Hollanti, C. Hollanti, and S. El Rouayheb, "Private information retrieval schemes for coded data with arbitrary collusion patterns," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 1908–1912.

[22] 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.

[23] Y. Zhang and G. Ge, "A general private information retrieval scheme for MDS coded databases with colluding servers," 2017, *arXiv:1704.06785*. [Online]. Available: https://arxiv.org/abs/1704.06785

[24] 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: https://arxiv.org/abs/1705.03186
[25] 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.
[26] Q. Wang and M. Skoglund, "Secure symmetric private information retrieval from colluding databases with adversaries," in *Proc. 55th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Oct. 2017, pp. 1083–1090.
[27] R. Tandon, "The capacity of cache aided private information retrieval," in *Proc. 55th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Oct. 2017, pp. 1078–1082.
[28] Q. Wang and M. Skoglund, "Linear symmetric private information retrieval for MDS coded distributed storage with colluding servers," 2017, *arXiv:1708.05673*. [Online]. Available: https://arxiv.org/abs/1708.05673
[29] S. Kadhe, B. Garcia, A. Heidarzadeh, S. El Rouayheb, and A. Sprintson, "Private information retrieval with side information," 2017, *arXiv:1709.00112*. [Online]. Available: https://arxiv.org/abs/1709.00112
[30] 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.
[31] Z. Chen, Z. Wang, and S. Jafar, "The capacity of T-private information retrieval with private side information," 2017, *arXiv:1709.03022*. [Online]. Available: https://arxiv.org/abs/1709.03022
[32] Y.-P. Wei, K. Banawan, and S. Ulukus, "The capacity of private information retrieval with partially known private side information," *IEEE Trans. Inf. Theory*, to be published. [Online]. Available: https://arxiv.org/abs/1710.00809
[33] H. Sun and S. A. Jafar, "The capacity of private computation," 2017, *arXiv:1710.11098*, [Online]. Available: https://arxiv.org/abs/1710.11098
[34] M. Mirmohseni and M. A. Maddah-Ali, "Private function retrieval," 2017, *arXiv:1711.04677*. [Online]. Available: https://arxiv.org/abs/1711.04677
[35] M. Abdul-Wahid, F. Almoualem, D. Kumar, and R. Tandon, "Private information retrieval from storage constrained databases—Coded caching meets PIR," 2017, *arXiv:1711.05244*, [Online]. Available: https://arxiv.org/abs/1711.05244
[36] Y.-P. Wei, K. Banawan, and S. Ulukus, "Cache-aided private information retrieval with partially known uncoded prefetching: Fundamental limits," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 6, pp. 1126–1139, Jun. 2018.
[37] K. Banawan and S. Ulukus, "Private information retrieval through wiretap channel II: Privacy meets security," *IEEE Trans. Inf. Theory*, to be published.
[38] K. Banawan and S. Ulukus, "Noisy private information retrieval: Separability of channel coding and information retrieval," *IEEE Trans. Inf. Theory*, to be published. [Online]. Available: https://arxiv.org/abs/1807.05997

Karim Banawan (S'13–M'18) received the B.Sc. and M.Sc. degrees, with highest honors, in electrical engineering from Alexandria University, Alexandria, Egypt, in 2008, 2012, respectively, and the M.Sc. and Ph.D. degrees in electrical engineering from the University of Maryland at College Park, MD, USA, in 2017 and 2018, respectively, with his Ph.D. thesis on private information retrieval and security in networks. He was the recipient of the Distinguished Dissertation Fellowship from the Department of Electrical and Computer Engineering, at the University of Maryland College Park, for his Ph.D. thesis work.

In 2019, he joined the department of electrical engineering, Alexandria University, as an assistant professor. His research interests include information theory, wireless communications, physical layer security and private information retrieval.

Sennur Ulukus (S'90–M'98–SM'15–F'16) is the Anthony Ephremides Professor in Information Sciences and Systems in the Department of Electrical and Computer Engineering at the University of Maryland at College Park, where she also holds a joint appointment with the Institute for Systems Research (ISR). Prior to joining UMD, she was a Senior Technical Staff Member at AT&T Labs-Research. She received her Ph.D. degree in Electrical and Computer Engineering from Wireless Information Network Laboratory (WINLAB), Rutgers University, and B.S. and M.S. degrees in Electrical and Electronics Engineering from Bilkent University. Her research interests are in information theory, wireless communications, machine learning, signal processing and networks, with recent focus on private information retrieval, age of information, distributed coded computation, energy harvesting communications, physical layer security, and wireless energy and information transfer.

Dr. Ulukus is a fellow of the IEEE, and a Distinguished Scholar-Teacher of the University of Maryland. She received the 2003 IEEE Marconi Prize Paper Award in Wireless Communications, the 2019 IEEE Communications Society Best Tutorial Paper Award, an 2005 NSF CAREER Award, the 2010–2011 ISR Outstanding Systems Engineering Faculty Award, and the 2012 ECE George Corcoran Outstanding Teaching Award. She is a Distinguished Lecturer of the IEEE Information Theory Society for 2018–2019. She is on the Editorial Board of the IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING since 2016. She was an Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS Series on Green Communications and Networking (2015–2016), IEEE TRANSACTIONS ON INFORMATION THEORY (2007–2010), and IEEE TRANSACTIONS ON COMMUNICATIONS (2003–2007). She was a Guest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (2015 and 2008), *Journal of Communications and Networks* (2012), and IEEE TRANSACTIONS ON INFORMATION THEORY (2011). She is a TPC co-chair of 2019 ITW, 2017 IEEE ISIT, 2016 IEEE Globecom, 2014 IEEE PIMRC, and 2011 IEEE CTW.