Private Information Retrieval from Byzantine and Colluding Databases

Karim Banawan Sennur Ulukus
Department of Electrical and Computer Engineering
University of Maryland, College Park, MD 20742
kbanawan@umd.edu *ulukus@umd.edu*

***Abstract*—We consider the problem of single-round private information retrieval (PIR) from N replicated databases. We consider the case when B databases are outdated (unsynchronized), or even worse, adversarial (Byzantine), and therefore, can return incorrect answers. In the PIR problem with Byzantine databases (BPIR), a user wishes to retrieve a specific message from a set of M messages with zero-error, irrespective of the actions performed by the Byzantine databases. We consider the T-privacy constraint in this paper, where any T databases can collude, and exchange the queries submitted by the user. We determine the information-theoretic capacity of this problem, which is the maximum number of *correct symbols* that can be retrieved privately for every symbol of the downloaded data to be $C = \frac{N-2B}{N} \cdot \frac{1-\frac{T}{N-2B}}{1-(\frac{T}{N-2B})^M}$, if $2B+T<N$. Our achievable scheme extends the optimal achievable scheme for the robust PIR (RPIR) problem to correct the *errors* introduced by the Byzantine databases. Our converse proof uses the idea of the cut-set bound in the network coding problem against adversarial nodes.**

I. Introduction

Preserving the privacy of the contents downloaded from open-access databases has been a major area of research within the computer science community [1]–[3]. Many practical applications are related to the private retrieval problem, such as: protecting the identity of stock market records reviewed by an investor, and protecting the nature of restricted content browsed by activists on the internet in oppressive regimes. In the seminal paper [1], Chor et. al. introduced the problem of private information retrieval (PIR). In the classical PIR setting, a user wishes to retrieve a certain message (or file) out of M distinct messages from N non-colluding and replicated databases without leaking any information about the identity of the desired message. The user prepares N queries, one for each database, in a single round, such that the queries do not reveal the user's interest in the desired message. Each database responds *truthfully* with an answering string. The user needs to be able to reconstruct the entire message by decoding the answer strings from all databases. A straightforward solution for the PIR problem is for the user to download the entire database. This solution, however, is highly inefficient with respect to the PIR rate, which is the ratio between the desired message size and the total downloaded symbols.

The computer science formulation of the PIR problem assumes that the message is of length 1. The formulation considers optimizing the download cost, which is the sum of the lengths of the answer strings, in addition to the upload cost, which is the sum of the lengths of the queries. Most of this work adopts computational guarantees as a privacy constraint [3]. Recently, the PIR problem is revisited by information theorists [4]–[7]. The problem is re-formulated such that: the size of the message can be arbitrarily large, the upload cost is ignored, and privacy is guaranteed in the information-theoretic sense. This formulation gives rise to the PIR capacity notion, which is the supremum of PIR rates over all achievable retrieval schemes. In [7], Sun and Jafar determine the capacity of the classical PIR model, and propose a scheme which is based on three principles: message symmetry, database symmetry, and exploitation of side information through interference alignment as observed in [8]. Interesting extensions for the classical PIR problem are investigated following [7], such as: PIR with T colluding databases (TPIR) [9], [10], robust PIR (RPIR) [9], symmetric PIR (SPIR) [11], MDS-coded PIR (CPIR) [6], [12], multi-message PIR (MPIR) [13], PIR under message size constraint L (LPIR) [14], multi-round PIR [15], MDS-coded symmetric PIR [16], MDS-coded PIR with colluding databases [17]–[20].

A common assumption in these works is that the databases respond truthfully with the correct answer strings. This enables the user to exploit the undesired symbol as side information at other databases, and distribute the requests for the desired message among the N databases. An interesting question arises, how can we reconstruct the desired message with no errors even if B databases respond incorrectly? Returning to the practical examples presented earlier: The databases storing the stock market records may not be updated simultaneously, therefore some of the databases may store outdated versions of the messages and can introduce errors to the answering strings. This scenario is referred to as the *unsynchronized PIR* problem [21]. For the oppressive regime example, some databases can be controlled by the regime, and these databases may return incorrect answer strings on purpose to confuse the user. This scenario is referred to as the *PIR with adversarial databases* problem [22], [23]. This motivates our interest in characterizing the exact capacity of the PIR problem with Byzantine databases (BPIR). In BPIR, there exist B databases, which are called Byzantine databases, that respond with erroneous answer strings. The errors introduced by the Byzantine databases can be unintentional (as in the case of databases storing a different copy of the message set), or even worse, can be intentional (as in the case of maliciously

This work was supported by NSF Grants CNS 13-14733, CCF 14-22111, and CNS 15-26608.

controlled databases). In both cases, the user needs to be able to reconstruct the desired message with no error, irrespective of the actions performed by the Byzantine databases. The BPIR problem was introduced in [22]. They propose a generic transformation from RPIR schemes to protocols that tolerate Byzantine servers, and give an explicit Byzantine robust scheme when $B \leq T \leq \frac{N}{3}$. [24] presents a fault-tolerant PIR scheme that can cope with malicious failures for $B \leq T \leq \frac{N}{2}$. [23] observes that allowing for list decoding instead of unique decoding enlarges the feasible set up to $B < N-T-1$. Their achievable scheme allows for a small failure probability. The unsynchronized PIR problem is investigated in [21], where they propose a two-round retrieval scheme. The scheme returns the desired record by first identifying which records are mis-synchronized, and then by constructing a PIR scheme that avoids these problematic records.

In this paper, we consider the *single-round* BPIR problem from N replicated databases in the presence of B Byzantine databases that can introduce errors to the answer strings. The remaining databases store the exact copy of the message set which contains M different messages, and respond truthfully with the correct answer strings. Our goal is to characterize the single-round capacity of the BPIR problem under the zero-error reliability constraint and the T-privacy constraint, which permits colluding between any T databases. To that end, we propose an achievable scheme that is resilient to the worst-case errors resulted from the Byzantine databases. Our achievable scheme extends the optimal scheme for the RPIR problem to correct the *errors* resulted from the Byzantine databases, in contrast to the *erasures* introduced by the unresponsive databases in RPIR. The new ingredients to the achievable scheme are: encoding the undesired symbols via a punctured MDS code, successive interference cancellation of the side information, and encoding the desired symbols by an outer-layer MDS code. For the converse, we extend the arguments developed for the network coding problem in [25] and distributed storage systems in [26] to the PIR problem. This cut-set upper bound can be thought of as a network version of the Singleton bound [27], and intuitively implies that a redundancy of $2B$ nodes is needed in order to mitigate the errors introduced by the B Byzantine databases.

We determine the exact capacity of the BPIR problem to be $C = \frac{N-2B}{N} \cdot \frac{1-\frac{T}{N-2B}}{1-\left(\frac{T}{N-2B}\right)^M}$, if $2B+T < N$. The capacity expression is equivalent the TPIR capacity with $N-2B$ databases with a multiplicative factor of $\frac{N-2B}{N}$, which signifies the ignorance of the user as to which $N-2B$ databases are honest. Our Byzantine formulation includes the special case of the single-round unsynchronized PIR problem, if the user has no knowledge about the number of mis-synchronized messages, and only knows that the entirety of some B databases may be unsynchronized. Under our assumptions, the single-round capacity of the unsynchronized PIR problem and the BPIR problem are the same. Due to space limitations here, proof details, extra examples and some figures can be found in the longer version of this paper [28].

II. Problem Formulation

Consider a single-round PIR setting with N replicated databases storing M messages (or files). The messages $\mathcal{W} = \{W_1, \cdots, W_M\}$ are independent and uniformly distributed over a large enough finite field $\mathbb{F}_q$. Each message $W_i \in \mathbb{F}_q^L$ is a vector of length L (q-ary symbols),

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

$$H(\mathcal{W}) = H(W_1, \cdots, W_M) = ML \tag{2}$$

Each database stores a copy of $\mathcal{W}$. Denote the contents of the nth database by Ω_n. Ideally, $\Omega_n = \mathcal{W}$ for all $n \in \{1, \cdots, N\}$.

In PIR, a user wishes to retrieve a message $W_i \in \mathcal{W}$ without revealing any information about the message index i. The user submits a single-round query $Q_n^{[i]}$ to the nth database. The user does not know the stored messages in advance, therefore, the message set $\mathcal{W}$ and the queries are statistically independent,

$$I\left(W_1, \cdots, W_M; Q_1^{[i]}, \cdots, Q_N^{[i]}\right) = I\left(\mathcal{W}; Q_{1:N}^{[i]}\right) = 0 \tag{3}$$

Ideally, the classical PIR formulation assumes that all databases store the correct database contents, and respond truthfully with the correct answering strings $A_{1:N}^{[i]} = \{A_1^{[i]}, \cdots, A_N^{[i]}\}$. In the BPIR setting, on the other hand, there exists a set $\mathcal{B}$ of databases, that is unknown to the user, such that $|\mathcal{B}| = B$, which are called Byzantine databases. These databases can respond arbitrarily to the user by introducing errors to the answer strings $A_{\mathcal{B}}^{[i]} = \{A_j^{[i]} : j \in \mathcal{B}\}$, i.e.,

$$H\left(A_n^{[i]} | Q_n^{[i]}, \mathcal{W}\right) > 0, \quad n \in \mathcal{B}, \; |\mathcal{B}| = B \tag{4}$$

We assume that these Byzantine databases can coordinate upon submitting the answers. In this paper, we do not assume a specific pattern to the errors. The remaining set of databases $\bar{\mathcal{B}} = \{1, \cdots, N\} \setminus \mathcal{B}$ store $\mathcal{W}$ and respond truthfully, i.e.,

$$H\left(A_n^{[i]} | Q_n^{[i]}, \mathcal{W}\right) = 0, \quad n \in \bar{\mathcal{B}}, \; |\bar{\mathcal{B}}| = N - B \tag{5}$$

We consider a T-privacy constraint as in the TPIR problem in [9], where any T databases can communicate and exchange the queries submitted by the user. To ensure the T-privacy constraint, the queries to any set $\mathcal{T} \subset \{1, \cdots, N\}$ of databases, such that $|\mathcal{T}| = T$, need to be statistically independent of the desired message index i, i.e.,

$$I\left(i; Q_{\mathcal{T}}^{[i]}\right) = 0, \quad \text{for all } \mathcal{T} \subset \{1, \cdots, N\}, \; |\mathcal{T}| = T \tag{6}$$

where $Q_{\mathcal{T}}^{[i]}$ are the queries submitted to the set $\mathcal{T}$ of databases. We do not assume any specific relation between the T colluding databases and the B Byzantine databases.

The user should be able to reconstruct the desired message W_i, no matter what the Byzantine databases do, i.e., if there exists a set of databases $\bar{\mathcal{B}}$, that is unknown to the user, such that (5) holds, then the reliability constraint is given by,

$$H(W_i | A_{1:N}^{[i]}, Q_{1:N}^{[i]}) = 0, \quad \text{such that (5) holds} \tag{7}$$

We define the *resilient* PIR rate R for BPIR as the ratio between the message size L and the total download cost under the reliability constraint in (7) for any possible action of the Byzantine databases, and the T-privacy constraint in (6), i.e.,

$$R = \frac{L}{\sum_{n=1}^{N} H(A_n^{[i]})} \tag{8}$$

The capacity of BPIR is $C = \sup R$ over all possible single-round retrieval schemes. We follow the information-theoretic assumptions of large enough message size, large enough field size, and ignore the upload cost as in [5], [9], [12]. A formal treatment of the capacity under message size constraints can be found in [14]. The BPIR with colluding databases reduces to the TPIR problem in [9] if $B = 0$.

III. Main Result and Discussions

The main result of this paper is to characterize the capacity of the BPIR problem under T-privacy constraint.

Theorem 1 *For the single-round BPIR problem with B Byzantine databases, and T colluding databases, such that $2B + T < N$, the capacity is given by,*

$$C = \frac{N-2B}{N} \cdot \frac{1 - \frac{T}{N-2B}}{1 - \left(\frac{T}{N-2B}\right)^M} \tag{9}$$

$$= \frac{N-2B}{N} \cdot \left(1 + \frac{T}{N-2B} + \cdots + \frac{T^{M-1}}{(N-2B)^{M-1}}\right)^{-1} \tag{10}$$

Furthermore, if $2B+1 \leq N \leq 2B+T$, $C = \frac{1}{(2B+1)M}$, which is the trivial rate in the BPIR problem. Otherwise $C = 0$.

The achievability proof for is given in Section IV, and the converse proof is given in Section V. We have a few remarks.

Remark 1 *The BPIR capacity in (9) is the same as the capacity of PIR with T colluding databases if the number of databases is $N - 2B$ with a penalty factor of $\frac{N-2B}{N}$. This means that the harm introduced by the B Byzantine databases is equivalent to removing $2B$ databases from the system, but the user still needs to download from all N databases, as it does not know which $N - 2B$ databases are honest. This results in the penalty term $\frac{N-2B}{N}$. If $B = 0$, (9) reduces to*

$$C_{\textit{colluded}} = \frac{1 - \frac{T}{N}}{1 - \left(\frac{T}{N}\right)^M} \tag{11}$$

which is the capacity expression in [9] as expected.

Remark 2 *Comparing the BPIR capacity with the robust capacity $C_{\textit{robust}}$ in [9], where U databases are unresponsive,*

$$C_{\textit{robust}} = \frac{1 - \frac{T}{N-U}}{1 - \left(\frac{T}{N-U}\right)^M} \tag{12}$$

we note that the number of redundant databases, which are needed to correct the errors *introduced by the Byzantine databases, is* twice *the number of redundant databases needed to correct the* erasures *introduced in the case of unresponsive databases. We also note that the penalty factor is missing in the RPIR problem, since in the RPIR problem, the user does not get the chance to download from the unresponsive databases, in contrast to the BPIR problem, in which the user downloads answer strings from all databases due to his ignorance.*

Remark 3 *The trivial rate for the BPIR problem is $\frac{1}{(2B+1)M}$, which is less than the trivial rate without the Byzantine databases, $\frac{1}{M}$. The reason for this is that the user must download $(2B+1)$ different copies of the database to decode the desired message via majority decoding. If $N < 2B+1$, the capacity is $C = 0$, as the Byzantine databases can always confuse the user to decode the desired message incorrectly.*

Remark 4 *When the number of messages is large, i.e., as $M \to \infty$, the BPIR capacity $C \to 1 - \frac{2B+T}{N}$, i.e., for large enough number of messages, the capacity acts as if there are no Byzantine databases and $2B + T$ databases are colluding.*

Remark 5 *If T and B are fixed and do not scale with N, i.e., $T = B = o(N)$, then the capacity is a strictly increasing function in N and $C \to 1$ as $N \to \infty$. If the number of the Byzantine databases scales with N, i.e., $B = \gamma N$, where $\gamma \in \left[0, \frac{1}{2}(1 - \frac{T}{N})\right)$, then $C \to 1 - 2\gamma$ as $N \to \infty$. If $2\gamma + \frac{1}{N} \leq 1 \leq 2\gamma + \frac{T}{N}$, then the only possible rate is the trivial rate $\frac{1}{(2B+1)M}$. As $N \to \infty$, then $\gamma \to \frac{1}{2}$, and $C \to 0$. This entails that the asymptotic behaviour of the BPIR capacity is a linear function with a slope of -2, i.e., the asymptotic rate as $N \to \infty$ is decreased by* twice *the ratio of the Byzantine databases. A similar behaviour is observed for secure distributed storage systems against Byzantine attacks in [26]. The problem is infeasible if $\gamma > \frac{1}{2}$, i.e., $C = 0$. This feasibility result conforms with the best result of a* uniquely decodable *BPIR scheme in [24] which needs $B < \frac{N}{2}$.*

Remark 6 *Surprisingly, our retrieval scheme in Section IV is a linear scheme in contrast to the network coding problem in [25] that states that linear coding schemes are not sufficient. We note that although the retrieval process is itself linear, the decoding process employs a successive interference cancellation decoder, which is non-linear.*

Remark 7 *The capacity expression in Theorem 1 is also the capacity result for the unsychronized PIR problem [21]. This occurs under the restriction to single-round schemes and the assumption that the user only knows that there exist B databases that are unsynchronized, but does not know the fraction of messages that are mis-synchronized. The achievability scheme in Section IV is a valid achievable scheme for the unsynchronized PIR problem, since the adversary in the Byzantine setting is stronger. For the converse proof, we restricted the actions of the adversarial databases to changing the contents of the stored messages, i.e., altering Ω_n from $\mathcal{W}$ to $\tilde{\mathcal{W}}$, which is the same setting as the unsynchronized PIR with no restriction on the fraction of messages that can be mis-synchronized.*

IV. Achievability Proof

In this section, we present an achievable scheme that is resilient to the errors introduced by the Byzantine databases. The achievable scheme does not assume any specific error pattern, and enables *correct decoding* of any desired message if any B databases become outdated, or commit an adversarial attack. The scheme generalizes the RPIR scheme in [9].

A. Preliminaries

Lemma 1 (MDS code puncturing [29]) *If $\mathcal{C}$ is an (n, k) MDS code, then by puncturing the code by a sequence of length z, i.e., deleting a sequence of size z from output codewords of $\mathcal{C}$, such that $z < n - k$, the resulting punctured code $\mathcal{C}_z$ is an $(n - z, k)$ MDS code.*

Lemma 2 (Statistical effect of full-rank matrices [9]) *Let $\mathbf{S}_1, \mathbf{S}_2, \cdots, \mathbf{S}_M \in \mathbb{F}_q^{\alpha\times\alpha}$ be M random matrices, drawn independently and uniformly from all $\alpha \times \alpha$ full-rank matrices over $\mathbb{F}_q$. Let $\mathbf{G}_1, \mathbf{G}_2, \cdots, \mathbf{G}_M \in \mathbb{F}_q^{\beta\times\beta}$ be M invertible square matrices of dimension $\beta \times \beta$ over $\mathbb{F}_q$. Let $\mathcal{I}_1, \cdots, \mathcal{I}_M \in \mathbb{N}^\beta$ be M index vectors, each containing β distinct indices from $\{1, \cdots, \alpha\}$, then $\{\mathbf{G}_1\mathbf{S}_1(\mathcal{I}_1, :), \cdots, \mathbf{G}_M\mathbf{S}_M(\mathcal{I}_M, :)\} \sim \{(\mathbf{S}_1([1 : \beta], :), \cdots, \mathbf{S}_M([1 : \beta], :)\}$. where $\sim$ denotes statistical equivalence, $\mathbf{S}_i(\mathcal{I}_i, :)$, $\mathbf{S}_i([1 : \beta], :)$ denote $\beta \times \alpha$ matrices with rows indexed by $\mathcal{I}_i$ and $\{1, \cdots, \beta\}$, respectively.*

Lemma 3 (Code capabilities [30]) *Let $\mathcal{C}$ be an $[n, k, d]$ linear block code over $\mathbb{F}_q$. Let ρ be the number of erasures introduced by the channel. Let $\tau \in \mathbb{N}$, such that $2\tau + \rho \leq d - 1$, then there exists a nearest-codeword decoder that recovers all errors and erasures if the number or errors (excluding erasures) is τ or less.*

B. Motivating Example: $M = 2$, $N = 5$, $T = 2$, $B = 1$

Assume without loss of generality that W_1 is the desired message. Let a_i and b_i be the ith symbol mixture of messages W_1 and W_2, respectively. The specific construction of these mixtures will be presented shortly. We begin the retrieval process by downloading $T^{M-1} = 2$ symbols from W_1, which are a_1, a_2 as in [9]. By *message symmetry*, we download b_1, b_2 from W_2. By *database symmetry*, we download 2 symbols from W_1 and 2 symbols from W_2 from all other databases.

The number of errors that can be corrected increases with d according to Lemma 3. MDS codes meet the Singleton bound [27] with equality, hence encoding both desired and undesired messages by MDS codes is desirable. In addition, Lemma 3 implies a *doubling effect*, which suggests that in order to correct the errors introduced by the Byzantine database, we should consider $N - 2B = 3$ honest databases. We note that any $T = 2$ of them can collude, therefore, we are left with 2 undesired symbols that can be used to generate side information among the 2 colluding databases. Hence, each database should get 1 side information equation $b_{[11:15]}$. These side-information symbols can be added to new desired symbols $a_{[11:15]}$. The query structure is shown in Table I.

TABLE I

The query table for the case $M = 2$, $N = 5$, $T = 2$, $B = 1$.

DB 1	DB 2	DB 3	DB 4	DB 5
a_1 a_2 b_1 b_2	a_3 a_4 b_3 b_4	a_5 a_6 b_5 b_6	a_7 a_8 b_7 b_8	a_9 a_{10} b_9 b_{10}
$a_{11} + b_{11}$	$a_{12} + b_{12}$	$a_{13} + b_{13}$	$a_{14} + b_{14}$	$a_{15} + b_{15}$

Now, we identify the specific construction of the mixtures $a_{[1:15]}$ and $b_{[1:15]}$ in Table I. For the desired message W_1, considering any $N - 2B = 3$ honest databases, we see 9 distinct symbols. Therefore, the length of W_1 is $L = 9$, and we use $\mathbf{S}_1$, which is a 9×9 random mixing matrix picked uniformly from the full-rank matrices over $\mathbb{F}_q^{9\times9}$. These 9 mixed symbols are further mapped to $a_{[1:15]}$ by a $(15, 9)$ MDS code generator matrix $\mathbf{MDS}_{15\times9}$, therefore,

$$a_{[1:15]} = \mathbf{MDS}_{15\times9}\mathbf{S}_1W_1 \tag{13}$$

For the undesired message W_2, considering again 3 honest databases, we have 6 individual symbols in round 1. We should be able to reconstruct the side information equations $b_{[11:15]}$ in round 2 from any 6 individual symbols, hence we get 6 random symbols from W_2. This can be done by considering the first 6 rows of the random mixing matrix $\mathbf{S}_2 \in \mathbb{F}_q^{9\times9}$. These randomly mixed symbols are further mapped to $b_{[1:15]}$ via and MDS code with generator matrix $\mathbf{MDS}_{15\times6}$, i.e.,

$$b_{[1:15]} = \mathbf{MDS}_{15\times6}\mathbf{S}_2([1 : 6], :)W_2 \tag{14}$$

To see the decodability: the worst-case scenario is that the Byzantine database commits errors in all the symbols returned to the user. This means that the database commits 2 errors in the individual symbols from W_1, 2 errors in the individual symbols from W_2, and 1 extra error in the sum of $a + b$.

Consider the codeword $b_{[1:10]}$: this codeword belongs to $(15, 6)$ MDS code with a sequence of length $z = 5$ removed. Hence, this codeword belongs to $(10, 6)$ punctured MDS code. Since $z = 5 < 15 - 6 = 9$, it is still an MDS code. Denote the minimum distance of the $(10, 6)$ punctured MDS code that results in $b_{[1:10]}$ by d_p^b. Then, $d_p^b = 10 - 6 + 1 = 5$. Consequently, from Lemma 3, the $(10, 6)$ punctured MDS code can tolerate errors up to τ_b, such that $\tau_b \leq \left\lfloor \frac{d_p^b - 1}{2} \right\rfloor = 2$. Therefore, this code can correct all errors that can be introduced to the individual undesired symbols $b_{[1:10]}$. Let $b^*_{[1:10]}$ be the correct codeword of $b_{[1:10]}$. Choose any 6 symbols from $b^*_{[1:10]}$. Now, since $\mathbf{MDS}_{15\times6}$ matrix has the property that any 6×6 matrix is an invertible matrix, then from any 6 symbols from $b^*_{[1:10]}$, the *correct side information* equations $b^*_{[11:15]}$ are determined and canceled from the sums of a and b in round 2.

For the desired message W_1: after removing the interference from W_2, we are left with $\tilde{a}_{[1:15]}$. Note that this is not exactly $a_{[1:15]}$, because we canceled $b^*_{[1:10]}$ and not $b_{[1:15]}$. However, the total errors in $\tilde{a}_{[1:15]}$ still is upper bounded by 3, since $\tilde{a}_{[1:15]}$ can differ from $a_{[1:15]}$ only in the positions that correspond to Byzantine databases. The desired message W_1 is coded via $(15,9)$ MDS code. Then, the minimum distance for this code is $d^a = 15 - 9 + 1 = 7$. Consequently, this code can tolerate errors up to τ_a, such that $\tau_a \leq \left\lfloor \frac{d^a-1}{2} \right\rfloor = 3$. Hence, all the errors in $\tilde{a}_{[1:15]}$ can be corrected, and we can obtain true $a^*_{[1:15]}$. Consider the first 9 symbols from $a^*_{[1:15]}$, without loss of generality, then

$$W_1 = (\mathbf{MDS}_{15\times 9}([1:9],:)\mathbf{S}_1)^{-1} a^*_{[1:9]} \tag{15}$$

since $\mathbf{MDS}_{15\times 9}([1:9],:)\mathbf{S}_1$ is a 9×9 invertible matrix.

Therefore, despite Byzantine behaviour of $B = 1$ database, we decode the desired message correctly. In addition, our achievable scheme can identify the Byzantine database as does the scheme in [21] by comparing $a^*_{[1:10]}$ with $a_{[1:10]}$, and $b^*_{[1:10]}$ with $b_{[1:10]}$ and see which database has introduced errors.

To see the privacy: we note that from any $T = 2$ databases, our achievable scheme collects 6 symbols from $a_{[1:15]}$ and 6 symbols from $b_{[1:15]}$ indexed by $\mathcal{I}$ such that $|\mathcal{I}| = 6$. For the undesired message, we collect $b_{\mathcal{I}}$,

$$b_{\mathcal{I}} = \mathbf{MDS}_{15\times 6}(\mathcal{I},:)\mathbf{S}_2([1:6],:)W_2 \sim \mathbf{S}_2([1:6],:)W_2 \tag{16}$$

where (16) follows from Lemma 2 as any 6×6 matrix in $\mathbf{MDS}_{15\times 6}$ matrix is full-rank. Therefore, the symbols $b_{\mathcal{I}}$ are independent and uniformly distributed. For $a_{\mathcal{I}}$, we have

$$a_{\mathcal{I}} = \mathbf{MDS}_{15\times 9}(\mathcal{I},:)\mathbf{S}_1 W_1 = \Psi_{6\times 9} W_1 \tag{17}$$

where $\Psi = \mathbf{MDS}_{15\times 9}(\mathcal{I},:)\mathbf{S}_1$ is a full row-rank matrix as any 6 rows in $\mathbf{MDS}_{15\times 9}$ are linearly independent. Consequently, the symbols $a_{\mathcal{I}}$ are also independent and uniformly distributed, and $a_{\mathcal{I}} \sim b_{\mathcal{I}}$ for every 2 databases, where $\sim$ means that the involved random vectors are statistically identical. Thus, the proposed scheme is 2-private; that is, despite colluding behaviour of $T = 2$ databases, we have privacy.

Finally, the achievable resilient retrieval rate is $R = \frac{9}{25} = \frac{N-2B}{N} \cdot \frac{1-\frac{T}{N-2B}}{1-\left(\frac{T}{N-2B}\right)^M} = C$. In comparison, the trivial rate for this system is $\frac{1}{(2B+1)M} = \frac{1}{6}$, as the user must download the entire database from 3 different databases for correct decoding.

C. General Achievable Scheme

The general achievable scheme is performed in M rounds. The ith round includes all the $\binom{M}{i}$ combinations of the sums of any i messages. The scheme requires $L = (N-2B)^M$. The construction resembles the optimal scheme for RPIR in [9]. The new key ingredient in our achievable scheme is the decoding procedure, which includes correcting the undesired symbols by punctured MDS codes, successive interference cancellation to cancel the interfering messages, and correcting the errors in the desired message by an outer layer MDS code.

1) General Description for the Scheme:

1) *Initialization:* The user downloads T^{M-1} mixed symbols from the desired message from the first database. The scheme sets the round index $i = 1$.
2) *Message symmetry:* To satisfy the privacy constraint, the user downloads the same number of mixed symbols from the undesired messages with all the possible combinations, i.e., in the ith round, the user downloads $\binom{M-1}{i}(N-2B-T)^{i-1}T^{M-i}$ mixed symbols from the remaining $M-1$ messages.
3) *Database symmetry:* The user repeats the same steps at all the databases. Specifically, the user downloads $\binom{M-1}{i-1}(N-2B-T)^{i-1}T^{M-i}$ equations in the form of a desired message mixture symbol and $i-1$ mixed symbols from the undesired messages, and $\binom{M-1}{i}(N-2B-T)^{i-1}T^{M-i}$ mixed symbols from the undesired messages only, from each database.
4) *Exploiting side information:* In the $(i+1)$th round, the user should be able to generate $\frac{N-2B-T}{T}$ side information equations for each undesired symbol in the ith round. The side information generated is added to a new mixed symbol from the desired message.
5) Repeat steps 2, 3, 4 after setting $i = i+1$ until $i = M-1$. We give the specific construction of the desired/undesired mixture in the following subsection.

2) Specific Construction of the Symbol Mixtures: Let $W_m \in \mathbb{F}_q^{(N-2B)^M}$, $m \in \{1, \cdots, M\}$ be the message vectors, and $\mathbf{S}_m$, $m \in \{1, \cdots, M\}$ be random mixing matrices picked independently and uniformly from the full-rank matrices in $\mathbb{F}_q^{(N-2B)^M \times (N-2B)^M}$. At the ith round, the user downloads all possible combinations of the sums of any i messages. In the following specific construction, we enumerate all the sets that contain a symbol from the desired message and assign them labels $\mathcal{L}_1, \cdots, \mathcal{L}_\delta$. For each undesired message, we enumerate also all the sets that contain symbols from this undesired message and do not include any desired symbols and assign them labels $\mathcal{K}_1, \cdots, \mathcal{K}_\Delta$. These sets construct the undesired symbol mixtures and the corresponding side information.

For the desired message: Assume that the desired message is W_ℓ. Let δ be the number of the distinct subsets of $\{1, \cdots, M\}$ that contain ℓ, then $\delta = 2^{M-1}$. Let $\mathcal{L}_i$, $i \in \{1, \cdots, \delta\}$ be the ith subset that contains ℓ. Let $X^{[\ell]} \in \mathbb{F}_q^{N(N-2B)^M}$ be the vector of mixtures that should be obtained from the desired message W_ℓ. Divide $X^{[\ell]}$ into δ partitions denoted by $x^{[\ell]}_{\mathcal{L}_i}$, each corresponds to a distinct set $\mathcal{L}_i$, i.e., $X^{[\ell]} = \begin{bmatrix} x^{[\ell]}_{\mathcal{L}_1} & \cdots & x^{[\ell]}_{\mathcal{L}_\delta} \end{bmatrix}^T$. Now, encode the desired message by a $\left(N(N-2B)^{M-1}, (N-2B)^M\right)$ MDS code as,

$$X^{[\ell]} = \mathbf{MDS}_{N(N-2B)^{M-1} \times (N-2B)^M} \mathbf{S}_\ell W_l \tag{18}$$

where $x^{[\ell]}_{\mathcal{L}_i}$ is a vector of length $N(N-2B-T)^{|\mathcal{L}_i|-1} T^{M-|\mathcal{L}_i|}$.

For any undesired message: Consider the undesired message W_k, $k \in \{1, \cdots, M\} \setminus \{\ell\}$. Let $\Delta = 2^{M-2}$ be the number of distinct subsets that contain k and do not contain ℓ. Let $\mathcal{K}_i$, $i \in \{1, \cdots, \Delta\}$ be the ith subset that contains k and does not contain ℓ. Define $u^{[k]}_{\mathcal{K}_i}$ to be the undesired symbol mixtures in the $|\mathcal{K}_i|$th round corresponding to message k among the $\mathcal{K}_i$ set. Define $\sigma^{[k]}_{\mathcal{K}_i}$ to be the side information symbols from message k among the $\mathcal{K}_i$ subset of undesired messages. These side information equations are added to a desired message symbol in the $(|\mathcal{K}_i|+1)$th round. For each subset $\mathcal{K}_i$,

$$\begin{bmatrix} u_{\mathcal{K}_i}^{[k]} \\ \sigma_{\mathcal{K}_i}^{[k]} \end{bmatrix} = \mathbf{MDS}_{\frac{N}{T}\alpha_i \times \alpha_i} \mathbf{S}_k \left(\mathcal{J}_i, :\right) W_k \tag{19}$$

where $\mathcal{J}_i = \left[\sum_{j=1}^{i-1} \alpha_j + 1 : \sum_{j=1}^{i} \alpha_j\right]$, $\alpha_i = (N-2B)(N-2B-T)^{|\mathcal{K}_i|-1} T^{M-|\mathcal{K}_i|}$, $u_{\mathcal{K}_i}^{[k]}$ is a vector of length $\frac{N}{N-2B}\alpha_i$, and $\sigma_{\mathcal{K}_i}^{[k]}$ is a vector of length $\frac{N-2B-T}{T} \cdot \frac{N}{N-2B}\alpha_i$. This implies that the side information $\sigma_{\mathcal{K}_i}^{[k]}$ in the $(|\mathcal{K}_i|+1)$th round is completely determined by $u_{\mathcal{K}_i}^{[k]}$ in the $|\mathcal{K}_i|$th round. The construction implies that we generate $\frac{N-2B-T}{T}$ side information symbols for each undesired symbol. We note that the same MDS matrix is used for all messages $k \neq \ell$ that belong to the same subset $\mathcal{K}_i$. This is critical to enable *interference alignment*, and *joint error correction*. Let $X^{[k]} \in \mathbb{F}_q^{N(N-2B)^{M-1}}$ be the vector of mixtures corresponding to message $k \neq \ell$ such that $X^{[k]} = \begin{bmatrix} u_{\mathcal{K}_1}^{[k]} & \sigma_{\mathcal{K}_1}^{[k]} & \cdots & u_{\mathcal{K}_\Delta}^{[k]} & \sigma_{\mathcal{K}_\Delta}^{[k]} \end{bmatrix}^T$. Then,

$$X^{[k]} = \mathbf{P}\mathbf{S}_k([1 : T(N-2B)^{M-1}], :)W_k \tag{20}$$

where $\mathbf{P} = \text{diag}\left(\mathbf{MDS}_{\frac{N}{T}\alpha_1 \times \alpha_1}, \cdots, \mathbf{MDS}_{\frac{N}{T}\alpha_\Delta \times \alpha_\Delta}\right)$ is a $\mathbb{F}_q^{N(N-2B)^{M-1} \times T(N-2B)^{M-1}}$ matrix. Now, we are ready to specify the queries. For every non-empty set $\mathcal{M} \subseteq \{1, \cdots, M\}$, define $\mathcal{Q}_{\mathcal{M}}^{[\ell]}$ to be all queries related to set $\mathcal{M}$,

$$\mathcal{Q}_{\mathcal{M}}^{[\ell]} = \begin{cases} x_{\mathcal{L}_1}^{[\ell]}, & \mathcal{M} = \mathcal{L}_1 = \{\ell\} \\ x_{\mathcal{L}_j}^{[\ell]} + \sum_{k \in \mathcal{K}_i} \sigma_{\mathcal{K}_i}^{[k]} & \exists i, j : \mathcal{M} = \mathcal{K}_i \cup \{\ell\} = \mathcal{L}_j \\ \sum_{k \in \mathcal{K}_i} u_{\mathcal{K}_i}^{[k]} & \exists i : \mathcal{M} = \mathcal{K}_i \end{cases} \tag{21}$$

We distribute the queries randomly and evenly among the N databases for each subset $\mathcal{M}$. This completes the construction.

D. Decodability, Privacy, and the Achievable Rate

For the decoding, the first step is to correct the errors in the undesired symbols in the $\mathcal{K}_i$ set in the $|\mathcal{K}_i|$th round, so that we can generate the correct side information in the $(|\mathcal{K}_i|+1)$th round. Since the sum of linear codes is also a linear code, for the every set $\mathcal{K}_i$, $i \in \{1, \cdots, \Delta\}$, we have

$$\begin{bmatrix} \sum_{k \in \mathcal{K}_i} u_{\mathcal{K}_i}^{[k]} \\ \sum_{k \in \mathcal{K}_i} \sigma_{\mathcal{K}_i}^{[k]} \end{bmatrix} = \mathbf{MDS}_{\frac{N}{T}\alpha_i \times \alpha_i} \sum_{k \in \mathcal{K}_i} \mathbf{S}_k \left(\mathcal{J}_i, :\right) W_k \tag{22}$$

This enables *joint error correction* on the aligned sum. The minimum distance of this MDS code is $d^{\mathcal{K}_i} = \frac{N-T}{T}\alpha_i + 1$.

Now, in the $|\mathcal{K}_i|$th round, the user downloads $\sum_{k \in \mathcal{K}_i} u_{\mathcal{K}_i}^{[k]}$ which is a vector of length $\frac{N}{N-2B}\alpha_i$ from all databases. The vector $\sum_{k \in \mathcal{K}_i} u_{\mathcal{K}_i}^{[k]}$ belongs to $\left(\frac{N}{N-2B}\alpha_i, \alpha_i\right)$ punctured MDS code with a puncturing sequence corresponding to the side information symbols, i.e., with a puncturing sequence of length $z = |\sigma_{\mathcal{K}_i}^{[k]}| = \frac{N-2B-T}{T} \cdot \frac{N}{N-2B}\alpha_i$. Therefore,

$$d^{\mathcal{K}_i} - z - 1 = \frac{N-T}{T}\alpha_i - \frac{N-2B-T}{T} \cdot \frac{N}{N-2B}\alpha_i \tag{23}$$

$$= \frac{2B}{N-2B}\alpha_i > 0 \tag{24}$$

Thus, the $\left(\frac{N}{N-2B}\alpha_i, \alpha_i\right)$ punctured MDS code remains an MDS code with a minimum distance d^{u_i}, such that

$$d^{u_i} = \frac{N}{N-2B}\alpha_i - \alpha_i + 1 = \frac{2B}{N-2B}\alpha_i + 1 \tag{25}$$

Hence, the punctured code can correct upto τ_{u_i} errors, such that

$$\tau_{u_i} \leq \left\lfloor \frac{d^{u_i} - 1}{2} \right\rfloor = \frac{B}{N-2B}\alpha_i \tag{26}$$

Each database contributes $\frac{1}{N-2B}\alpha_i$ symbols from $\sum_{k \in \mathcal{K}_i} u_{\mathcal{K}_i}^{[k]}$, hence the Byzantine databases can introduce at most $\frac{B}{N-2B}\alpha_i$ errors. Consequently, the punctured MDS code can correct all errors in $\sum_{k \in \mathcal{K}_i} u_{\mathcal{K}_i}^{[k]}$. This results in a corrected undesired message vector $\left(\sum_{k \in \mathcal{K}_i} u_{\mathcal{K}_i}^{[k]}\right)^*$. Choose any α_i symbols from $\left(\sum_{k \in \mathcal{K}_i} u_{\mathcal{K}_i}^{[k]}\right)^*$. By the MDS property of the $(\frac{N}{T}\alpha_i, \alpha_i)$ MDS code, any $\alpha_i \times \alpha_i$ submatrix is invertible, hence a correct version of the side information vector, which is used in the $(|\mathcal{K}_i|+1)$th round, can be generated. Denote this correct version by $\left(\sum_{k \in \mathcal{K}_i} \sigma_{\mathcal{K}_i}^{[k]}\right)^*$. Now, we cancel the correct side information successively from each set $\mathcal{K}_i$. Note that the successive correction of side information gives rise to non-linearity in the decoding. After interference cancellation, we are left with $\tilde{X}^{[\ell]}$, which is not exactly $X^{[\ell]}$, as we cancelled the correct side information from the sum and not the side information provided by the Byzantine databases. This is not a problem, because $\tilde{X}^{[\ell]}$ and $X^{[\ell]}$ differ in codeword positions if and only if these positions belong to the Byzantine databases, hence the worst-case number of errors in $\tilde{X}^{[\ell]}$ cannot increase. The desired message is encoded by $(N(N-2B)^{M-1}, (N-2B)^M)$ MDS code with minimum distance d^x, such that

$$d^x = N(N-2B)^{M-1} - (N-2B)^M + 1 \tag{27}$$

$$= 2B(N-2B)^{M-1} + 1 \tag{28}$$

Each database returns $(N-2B)^{M-1}$ symbols from the desired message. The B Byzantine databases can at most introduce $B(N-2B)^{M-1}$ errors. The outer MDS code can correct up to τ_x errors, such that

$$\tau_x \leq \left\lfloor \frac{d^x - 1}{2} \right\rfloor = B(N-2B)^{M-1} \tag{29}$$

Thus, the user can correct all the errors introduced by the Byzantine databases to get a correct vector $\left(X^{[\ell]}\right)^* \in \mathbb{F}_q^{N(N-2B)^{M-1}}$. Consider any $(N-2B)^M$ symbols from $\left(X^{[\ell]}\right)^*$. Denote these symbols by x_ℓ^*, and index them by $\mathcal{I}_x$. Then, the user can decode W_ℓ with zero error via

$$W_\ell = (\mathbf{MDS}_{N(N-2B)^{M-1} \times (N-2B)^M}(\mathcal{I}_x, :)\mathbf{S}_1)^{-1} x_\ell^* \tag{30}$$

This is true as matrix $\mathbf{MDS}_{N(N-2B)^{M-1} \times (N-2B)^M}(\mathcal{I}_x, :)\mathbf{S}_1$ is invertible by the MDS property. In addition, the user can identify the Byzantine databases by comparing the correct versions of the undesired symbols at each cancellation step $(\sum_{k\in\mathcal{K}_i} u_{\mathcal{K}_i}^{[k]})^*$, and the desired symbols $\left(X^{[\ell]}\right)^*$ by their counterparts from the retrieval process. Any change between the correct vector and the retrieved vector implies that this database is a Byzantine database (or unsynchronized). The user can expurgate the malicious nodes in this case as in [21], [26].

Next, we show how the privacy is achieved. The queries for any T colluding databases are comprised of $T(N-2B)^{M-1}$ mixed symbols from each message W_i, $i \in \{1, \cdots, M\}$. Let these symbols be indexed by $\mathcal{I}$. Denote the kth message symbols by $x_{\mathcal{I}}^{[k]}$. For the desired symbols, we have

$$x_{\mathcal{I}}^{[\ell]} = \mathbf{MDS}_{N(N-2B)^{M-1}\times(N-2B)^M}(\mathcal{I},:)\mathbf{S}_\ell W_l \tag{31}$$

Since $|\mathcal{I}| = T(N-2B)^{M-1} < (N-2B)^M$ as $2B+T<N$ by construction, and due to the MDS property, the symbols $x_{\mathcal{I}}^{[\ell]}$ have full-rank. Hence, they are independent and uniformly distributed. Furthermore, for any undesired message W_k, $k \neq \ell$, we have,

$$x_{\mathcal{I}}^{[k]} = \mathbf{\Phi}\mathbf{S}_k([1:T(N-2B)^{M-1}],:)W_k \tag{32}$$

where $\mathcal{I} = \bigcup_{j=1}^{\Delta} \mathcal{I}_j$, and $|\mathcal{I}_j| = \alpha_j$, and $\mathbf{\Phi} = \text{diag}\left(\mathbf{MDS}_{\frac{N}{T}\alpha_1\times\alpha_1}(\mathcal{I}_1,:),\cdots,\mathbf{MDS}_{\frac{N}{T}\alpha_\Delta\times\alpha_\Delta}(\mathcal{I}_\Delta,:)\right)$. Therefore, each submatrix in Φ is an $\alpha_i \times \alpha_i$ invertible matrix by the MDS property. Hence, Φ is also an invertible matrix because it is a block-diagonal matrix. By Lemma 2, we have

$$x_{\mathcal{I}}^{[k]} \sim \mathbf{S}_k([1:T(N-2B)^{M-1}],:)W_k \tag{33}$$

Thus, symbols $x_{\mathcal{I}}^{[k]}$ are independent and uniformly distributed, and the privacy is guaranteed. We next calculate the achievable resilient rate. We note that the scheme operates in M rounds. The total download in the ith round is $\binom{M}{i}(N-2B-T)^{i-1}T^{M-i}$ from each database, i.e., the total download of the scheme, D, is $D = N\sum_{i=1}^{M}\binom{M}{i}(N-2B-T)^{i-1}T^{M-i}$. The scheme decodes correctly the desired message, which has length $L=(N-2B)^M$. Thus, the resilient retrieval rate is,

$$\begin{aligned}
R &= \frac{L}{D} && (34)\\
&= \frac{(N-2B)^M}{N\sum_{i=1}^{M}\binom{M}{i}(N-2B-T)^{i-1}T^{M-i}} && (35)\\
&= \frac{N-2B}{N}\cdot\frac{(N-2B)^{M-1}}{\sum_{i=1}^{M}\binom{M}{i}(N-2B-T)^{i-1}T^{M-i}} && (36)\\
&= \frac{N-2B}{N}\cdot\frac{(N-2B)^{M-1}}{\frac{1}{N-2B-T}\left((N-2B)^M-T^M\right)} && (37)\\
&= \frac{N-2B}{N}\cdot\frac{1-\frac{T}{N-2B}}{1-\left(\frac{T}{N-2B}\right)^M} && (38)
\end{aligned}$$

which is the expression in Theorem 1.

V. Converse Proof

In this section, we develop an upper bound for the BPIR problem. We adapt the cut-set upper bound proof in [25], [26] to the PIR setting. The upper bound can be thought of as a network version of the Singleton bound [27]. The upper bound intuitively asserts that the effect of the Byzantine databases on the retrieval rate is harmful as if $2B$ databases are removed from the retrieval process, but the user still needs to access them. Some technical differences from [25] arise in PIR:

1) The Byzantine databases in PIR are not fully omniscient, since they do not know which message the user wishes to retrieve. In the following we assume that the Byzantine databases alter the contents of the entire database.
2) The user does not know the entire codebook in advance.

For sake of deriving an upper bound, we make the following simplifications:

1) We assume that the actions of the Byzantine databases are restricted to altering the contents of the entire database, i.e., the nth Byzantine database changes its contents Ω_n from $\mathcal{W}$ to $\tilde{\mathcal{W}}$, where $\tilde{\mathcal{W}} \neq \mathcal{W}$. This restriction is valid from the converse point of view, since it potentially results in a weaker adversary, which in turn results in a higher rate. Note that, in this sense the Byzantine databases are reduced to being unsynchronized databases (with unknown number of mis-synchronized messages).
2) We further restrict the answering string from the nth database to be a deterministic function $f_n(\cdot)$, i.e., $A_n^{[i]} = f_n(\Omega_n, Q_n^{[i]})$, of the altered database Ω_n. This restriction also limits the capabilities of the Byzantine databases. This results in a further upper bound on rate. Since we restrict the actions of the Byzantine databases to altering Ω_n only, we signify this dependence on Ω_n by writing the answering string $A_n^{[i]}$ as $A_n^{[i]}(\Omega_n)$.
3) We can assume that the retrieval scheme is symmetric. This is without loss of generality as in [7], i.e.,

$$H(A_1^{[i]}|\mathcal{Q}) = H(A_2^{[i]}|\mathcal{Q}) = \cdots = H(A_N^{[i]}|\mathcal{Q}) \tag{39}$$

This assumption remains true in the BPIR problem, because assumptions 1 and 2 above imply that the Byzantine databases answer truthfully to the queries based on their own (altered) Ω_n. Therefore, the lengths of the answer strings will be symmetric in response to a symmetric scheme.

The main argument of the converse proof is summarized in the following lemma. The proof of the lemma can be found in [28]. The main argument in the lemma resembles the converse proofs of [25, Theorem 1] and [26, Theorem 6].

Lemma 4 *Fix a set of honest databases $\mathcal{U} \subset \{1,\cdots,N\}$ such that $|\mathcal{U}| = N-2B$, and $\Omega_n = \mathcal{W}$, for every $n \in \mathcal{U}$. Then, for correct decoding of W_i, the answer strings $A_{\mathcal{U}}^{[i]}(\mathcal{W})$ is unique for every realization of $\mathcal{W}$, i.e., there cannot exist two realizations of the message set $\mathcal{W}, \tilde{\mathcal{W}}$, such that $\mathcal{W} \neq \tilde{\mathcal{W}}$, and $A_{\mathcal{U}}^{[i]}(\mathcal{W}) = A_{\mathcal{U}}^{[i]}(\tilde{\mathcal{W}})$.*

Lemma 4 implies that the answer strings from any $N-2B$ honest databases are enough to reconstruct the desired message, since every realization of the message set produces different answering strings from any $N-2B$ databases. Now, we continue with the main body of the converse proof. From Lemma 4, the answers $A_{\mathcal{U}}^{[i]}(\mathcal{W})$ are unique for every $\mathcal{W}$, hence restricting the decoding function to these answers uniquely determine W_i, i.e., there exists no further confusion about the correct database contents $\mathcal{W}$, and the answering strings are designed to retrieve W_i from this $\mathcal{W}$. Consequently, if the true realization of the database is $\mathcal{W}$, we can write

$$R = \frac{L}{\sum_{n=1}^{N} H(A_n^{[i]})} \tag{40}$$

$$\leq \frac{L}{\sum_{n=1}^{N} H(A_n^{[i]}|\mathcal{Q})} \tag{41}$$

$$= \frac{N-2B}{N} \cdot \frac{L}{\sum_{n\in\mathcal{U}} H(A_n^{[i]}(\mathcal{W})|\mathcal{Q})} \tag{42}$$

$$\leq \frac{N-2B}{N} \cdot C_T(|\mathcal{U}|) \tag{43}$$

$$= \frac{N-2B}{N} \cdot C_T(N-2B) \tag{44}$$

$$= \frac{N-2B}{N} \cdot \frac{1-\frac{T}{N-2B}}{1-\left(\frac{T}{N-2B}\right)^M} \tag{45}$$

where $C_T(\cdot)$ is the capacity of the PIR problem with T colluding databases as a function of the number of databases. Here, (42) follows from the symmetry assumption and the fact that $A_{\mathcal{U}}^{[i]}(\mathcal{W})$ can decode W_i correctly and then $\frac{L}{\sum_{n\in\mathcal{U}} H(A_n^{[i]}(\mathcal{W})|\mathcal{Q})}$ is a valid upper bound on the retrieval rate under the T-privacy constraint if the accessed databases are restricted to $\mathcal{U}$, which is further upper bounded by the TPIR capacity $C_T(|\mathcal{U}|)$ in (43) as $C_T(|\mathcal{U}|)$ is the supremum of all rates that can be achieved using the set of databases $\mathcal{U}$ under the T-privacy constraint, and (45) follows from the capacity expression in [9].

VI. Conclusions

We investigated the PIR problem from N replicated databases in the presence of B Byzantine databases, and T-colluding databases from an information-theoretic perspective. We determined the exact capacity of the BPIR problem to be $C = \frac{N-2B}{N} \cdot \frac{1-\frac{T}{N-2B}}{1-(\frac{T}{N-2B})^M}$. The expression shows that in order to correct the errors introduced by the adversarial databases, the system needs to have $2B$ redundant storage nodes. The retrieval rate is further penalized by the factor $\frac{N-2B}{N}$, which reflects the ignorance of the user to honest databases. The BPIR capacity converges to $C \rightarrow 1-2\gamma$ as $B, N \rightarrow \infty$, $B = \gamma N$, where γ is the fraction of Byzantine databases. For large enough number of messages $C \rightarrow 1 - \frac{2B+T}{N}$. We extended the optimal scheme for the RPIR problem to permit *error correction* of any error pattern introduced by the Byzantine databases. For the converse, we adapted the cut-set bound, which was derived for the network coding problem against adversarial nodes, for the PIR setting.

References

[1] B. Chor, E. Kushilevitz, O. Goldreich, and M. Sudan. Private information retrieval. *Journal of the ACM*, 45(6):965–981, 1998.

[2] R. Ostrovsky and W. Skeith III. A survey of single-database private information retrieval: Techniques and applications. In *International Workshop on Public Key Cryptography*, pages 393–411. Springer, 2007.

[3] S. Yekhanin. Private information retrieval. *Communications of the ACM*, 53(4):68–73, 2010.

[4] N. B. Shah, K. V. Rashmi, and K. Ramchandran. One extra bit of download ensures perfectly private information retrieval. In *IEEE ISIT*, June 2014.

[5] T. Chan, S. Ho, and H. Yamamoto. Private information retrieval for coded storage. In *IEEE ISIT*, June 2015.

[6] R. Tajeddine and S. El Rouayheb. Private information retrieval from MDS coded data in distributed storage systems. In *IEEE ISIT*, July 2016.

[7] H. Sun and S. Jafar. The capacity of private information retrieval. 2016. Available at arXiv:1602.09134.

[8] H. Sun and S. Jafar. Blind interference alignment for private information retrieval. 2016. Available at arXiv:1601.07885.

[9] H. Sun and S. Jafar. The capacity of robust private information retrieval with colluding databases. 2016. Available at arXiv:1605.00635.

[10] 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. 2017. Available at arXiv:1701.07636.

[11] H. Sun and S. Jafar. The capacity of symmetric private information retrieval. 2016. Available at arXiv:1606.08828.

[12] K. Banawan and S. Ulukus. The capacity of private information retrieval from coded databases. *IEEE Trans. on Info. Theory*. Submitted September 2016. Also available at arXiv:1609.08138.

[13] K. Banawan and S. Ulukus. Multi-message private information retrieval: Capacity results and near-optimal schemes. *IEEE Trans. on Info. Theory*. Submitted February 2017. Also available at arXiv:1702.01739.

[14] H. Sun and S. Jafar. Optimal download cost of private information retrieval for arbitrary message length. 2016. Available at arXiv:1610.03048.

[15] H. Sun and S. Jafar. Multiround private information retrieval: Capacity and storage overhead. 2016. Available at arXiv:1611.02257.

[16] Q. Wang and M. Skoglund. Symmetric private information retrieval for MDS coded distributed storage. 2016. Available at arXiv:1610.04530.

[17] R. Freij-Hollanti, O. Gnilke, C. Hollanti, and D. Karpuk. Private information retrieval from coded databases with colluding servers. 2016. Available at arXiv:1611.02062.

[18] H. Sun and S. Jafar. Private information retrieval from MDS coded data with colluding servers: Settling a conjecture by Freij-Hollanti et al. 2017. Available at arXiv: 1701.07807.

[19] Y. Zhang and G. Ge. A general private information retrieval scheme for MDS coded databases with colluding servers. 2017. Available at arXiv: 1704.06785.

[20] Y. Zhang and G. Ge. Multi-file private information retrieval from MDS coded databases with colluding servers. 2017. Available at arXiv: 1705.03186.

[21] G. Fanti and K. Ramchandran. Efficient private information retrieval over unsynchronized databases. *IEEE Journal of Selected Topics in Signal Processing*, 9(7):1229–1239, October 2015.

[22] A. Beimel and Y. Stahl. Robust information-theoretic private information retrieval. In *International Conference on Security in Communication Networks*, pages 326–341. Springer, 2002.

[23] C. Devet, I. Goldberg, and N. Heninger. Optimally robust private information retrieval. In *USENIX Security Symposium*, 2012.

[24] E. Y. Yang, J. Xu, and K. H. Bennett. Private information retrieval in the presence of malicious failures. In *Proceedings 26th Annual International Computer Software and Applications*, August 2002.

[25] O. Kosut, L. Tong, and D. N. C. Tse. Polytope codes against adversaries in networks. *IEEE Trans. on Info. Theory*, 60:3308–3344, June 2014.

[26] S. Pawar, S. El Rouayheb, and K. Ramchandran. Securing dynamic distributed storage systems against eavesdropping and adversarial attacks. *IEEE Trans. on Info.Theory*, 57(10):6734–6753, October 2011.

[27] R. Singleton. Maximum distance Q-nary codes. *IEEE Trans. on Info. Theory*, 10(2):116–118, April 1964.

[28] K. Banawan and S. Ulukus. The capacity of private information retrieval from Byzantine and colluding databases. *IEEE Trans. on Info. Theory*. Submitted June 2017. Also available at arXiv:1706.01442.

[29] C. Feyling. Punctured maximum distance separable codes. *Electronics Letters*, 29(5):470–471, March 1993.

[30] R. Roth. *Introduction to Coding Theory*. Cambridge University Press, 2006.