The Capacity of Symmetric Private Information Retrieval

Hua Sun, Syed A. Jafar

Abstract

Private information retrieval (PIR) is the problem of retrieving as efficiently as possible, one out of K messages from N non-communicating replicated databases (each holds all K messages) while keeping the identity of the desired message index a secret from each individual database. Symmetric PIR (SPIR) is a generalization of PIR to include the requirement that beyond the desired message, the user learns nothing about the other $K-1$ messages. The information theoretic capacity of SPIR (equivalently, the reciprocal of minimum download cost) is the maximum number of bits of desired information that can be privately retrieved per bit of downloaded information. We show that the capacity of SPIR is $1-1/N$ regardless of the number of messages K, if the databases have access to common randomness (not available to the user) that is independent of the messages, in the amount that is at least $1/(N-1)$ bits per desired message bit, and zero otherwise. Extensions to the capacity region of SPIR and the capacity of finite length SPIR are provided.

This paper was presented in part at NETCOD 2016. Hua Sun (email: huas2@uci.edu) and Syed A. Jafar (email: syed@uci.edu) are with the Center of Pervasive Communications and Computing (CPCC) in the Department of Electrical Engineering and Computer Science (EECS) at the University of California Irvine.

1 Introduction

The private information retrieval (PIR) problem [1, 2] seeks the most efficient way for a user to retrieve a desired message from a set of distributed databases, each of which stores all the messages, without revealing any information about which message is being retrieved to any individual database. This seemingly impossible mission has a trivial (expensive) solution, i.e., the user can request all the messages to hide his interest. The goal of the PIR problem is to find the most efficient solution. The capacity of PIR is defined as the maximum number of bits of desired message that can be privately downloaded per bit of downloaded information. In our recent work [3], the capacity of PIR with K messages and N databases was shown to be $C_{\text{PIR}} = (1 + 1/N + \cdots + 1/N^{K-1})^{-1}$.

The original formulation of PIR only considers the privacy of the user. The privacy of the undesired messages is ignored. However, it is often desirable to restrict the user to retrieve nothing beyond his chosen message. This new constraint is called database privacy, and with this constraint, the problem is called symmetric[1] PIR (SPIR) [4]. Symmetric PIR is especially challenging because the databases must individually learn nothing about the identity of the desired message, but must still collectively allow the user to retrieve his desired message in such a way that the user learns nothing about any other message besides his desired message. For example, the trivial solution of downloading everything, is no longer acceptable. The main contribution of this work is the characterization of the capacity of SPIR, i.e., the maximum number of bits of desired message that can be privately retrieved by a user per bit of downloaded information, without leaking any information about undesired messages to the user. For K messages and N databases, we show that the capacity is $1 - 1/N$. Extensions of the main result, from capacity to capacity region and from infinite message length to arbitrary message length, are also provided.

Besides its direct applications, PIR is especially significant as a fundamental problem that lies at the intersection of several open problems in cryptography [5, 6], coding theory [7, 8, 9] and complexity theory [10]. SPIR inherits many of these connections from PIR. For example, SPIR is essentially a (distributed) form of oblivious transfer [11, 12], where the typical objective is that the transmitter(s) should not know which message is received by the receiver and the receiver should obtain nothing more than the desired message. Oblivious transfer is an important building block (primitive) in cryptography, whose feasibility leads to many other cryptographic protocols [13, 14]. Fundamental limits on the communication efficiency of various forms of oblivious transfer therefore represent an important class of open problems [15, 16]. The capacity characterization of SPIR is a promising step in this direction.

Notation: For $n_1, n_2 \in \mathbb{Z}, n_1 \leq n_2$, define the notation $[n_1 : n_2]$ as the set $\{n_1, n_1 + 1, \cdots, n_2\}$. For an index set $\mathcal{I} = \{i_1, i_2, \cdots, i_n\}$, with $i_1 < i_2 < \cdots < i_n$, the notation $A_{\mathcal{I}}$ represents the vector $(A_{i_1}, A_{i_2}, \cdots, A_{i_n})$. For an element i_θ in the set $\mathcal{I} = \{i_1, i_2, \cdots, i_n\}$, i.e., $i_\theta \in \mathcal{I}$, the notation $\overline{i_\theta}$ represents the complement of $\{i_\theta\}$, i.e., $\overline{i_\theta} \overset{\triangle}{=} \{i_1, \cdots, i_{\theta-1}, i_{\theta+1}, \cdots, i_n\}$.

2 Problem Statement

Consider K independent messages $W_1, \cdots, W_K, W_k \in \mathbb{F}_p^{l_k L \times 1}, k \in [1 : K], l_k \in \mathbb{Z}_+, L \in \mathbb{Z}_+$, where W_k is represented as an $l_k L \times 1$ vector comprised of $l_k L$ i.i.d. uniform symbols from a finite field $\mathbb{F}_p$ for a prime p. In p-ary units,

$$H(W_1, \cdots, W_K) = H(W_1) + \cdots + H(W_K), \tag{1}$$

[1] Symmetry means that the privacy of both the user and the database is considered.

$$H(W_k) = l_k L, \forall k \in [1:K]. \tag{2}$$

There are N databases. Each database stores all the messages $W_1, \cdots, W_K$.

Let us use $\mathcal{F}$ to denote a random variable privately generated by the user, whose realization is not available to the servers. $\mathcal{F}$ represents the randomness in the strategies followed by the user. The user privately generates θ uniformly from $[1:K]$ and wishes to retrieve W_θ privately. The databases do not want to give out any information beyond the one message of the user's choosing (W_θ). In order to achieve database-privacy, we assume that the databases share a common random variable S that is not known to the user. It has been shown that without such common randomness, SPIR is not feasible [4]. For a pictorial illustration of an example of the SPIR problem with K messages and 2 databases, see Figure 1. $\mathcal{F}$ is generated independently and before the realizations of the messages, the common randomness or the desired message index are known, so that

$$H(\theta, \mathcal{F}, W_1, \cdots, W_K, S) = H(\theta) + H(\mathcal{F}) + H(W_1) + \cdots + H(W_K) + H(S). \tag{3}$$

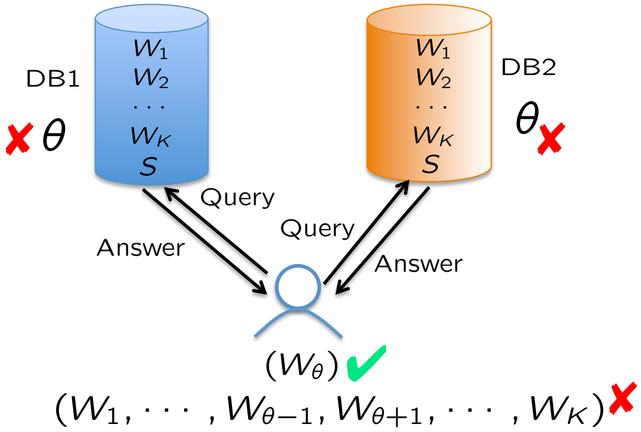

Figure 1: The SPIR problem with K messages and 2 databases.

Suppose $\theta = k$. In order to retrieve message $W_k, k \in [1:K]$ privately, the user privately generates N queries $Q_1^{[k]}, \cdots, Q_N^{[k]}$.

$$H(Q_1^{[k]}, \cdots, Q_N^{[k]} | \mathcal{F}) = 0, \forall k \in [1:K]. \tag{4}$$

The user sends query $Q_n^{[k]}$ to the n-th database, $n \in [1:N]$. Upon receiving $Q_n^{[k]}$, the n-th database generates an answering string $A_n^{[k]}$, which is a function of $Q_n^{[k]}$, all messages $W_1, \cdots, W_K$, and the common randomness S,

$$H(A_n^{[k]} | Q_n^{[k]}, W_1, \cdots, W_K, S) = 0. \tag{5}$$

Each database returns to the user its answer $A_n^{[k]}$.

From all the information that is now available to the user $(Q_{1:N}^{[k]}, A_{1:N}^{[k]}, \mathcal{F})$, the user decodes the desired message W_k according to a decoding rule that is specified by the SPIR scheme. Let P_e denote the probability of error achieved with the specified decoding rule.

To protect the user's privacy, the K strategies must be indistinguishable (identically distributed) from the perspective of any individual database, i.e., the following user-privacy constraint must be satisfied[2],

$$\text{[User-Privacy]} \quad (Q_n^{[k]}, A_n^{[k]}, W_{1:K}, S) \sim (Q_n^{[k']}, A_n^{[k']}, W_{1:K}, S), \\ \forall k, k' \in [1:K], \forall n \in [1:N]. \tag{6}$$

Symmetric PIR also requires protecting the privacy of the database, i.e., it must be ensured that the user learns nothing more than the desired message W_k. So the vector $W_{\overline{k}} = (W_1, \cdots, W_{k-1}, W_{k+1}, \cdots, W_K)$, must be independent of all the information available to the user. Thus, the following database-privacy constraint must be satisfied:

$$\text{[DB-Privacy]} \quad I(W_{\overline{k}}; Q_{1:N}^{[k]}, A_{1:N}^{[k]}, \mathcal{F}) = 0, \forall k \in [1:K]. \tag{7}$$

The SPIR rate of W_k characterizes the amount of desired information retrieved per downloaded symbol, and is defined as follows.

$$R_k \triangleq \frac{l_k L}{D}. \tag{8}$$

where D is the maximum value of the total number of symbols downloaded by the user from all the databases.

A rate tuple $(R_1, \cdots, R_K)$ is said to be ϵ-error achievable if $\forall k \in [1:K]$, there exists a sequence of PIR schemes, indexed by L, where the rate of W_k is greater than or equal to R_k and $P_e \to 0$ as $L \to \infty$. Note that for such a sequence of SPIR schemes, from Fano's inequality, we must have

$$\text{[Correctness]} \; o(L) = \frac{1}{L} H(W_k | Q_{1:N}^{[k]}, A_{1:N}^{[k]}, \mathcal{F}) \tag{9}$$

$$\overset{(4)}{=} \frac{1}{L} H(W_k | A_{1:N}^{[k]}, \mathcal{F}) \tag{10}$$

where $o(L)$ represents a term whose value approaches zero as L approaches infinity. The closure of the set of all ϵ-error achievable rate tuples is called the capacity region $\mathcal{C}$.

3 Results

3.1 Capacity of SPIR

In the typical setting of SPIR, the sizes of the messages are the same, i.e., $l_k = 1, \forall k \in [1:K]$ and the rate (refer to (8)) of each message is the same. Then the capacity region is characterized by one single parameter, i.e., the supremum of the achievable rate, named the capacity. We denote the capacity as C.

When there is only $K = 1$ message, note that the database-privacy constraint is satisfied trivially, so that SPIR reduces to the PIR setting and the capacity is 1. For $K \geq 2$, it is known that some common randomness S is necessary for the feasibility of SPIR. Let us define ρ as the amount of common randomness relative to the message size

$$\rho = \frac{H(S)}{H(W)} = \frac{H(S)}{L} \tag{11}$$

[2] The User-Privacy constraint is equivalently expressed as $I(\theta; Q_n^{[\theta]}, A_n^{[\theta]}, W_{1:K}, S) = 0$.

The capacity should depend on ρ, and because availability of common randomness at the databases is a non-trivial requirement, this dependence is of some interest.

When there is only $N = 1$ database, it is easy to see that the database-privacy constraint, the user-privacy constraint and correctness constraint conflict with each other such that SPIR is not feasible and the capacity is zero. The reason is as follows. First, because of the user-privacy constraint (6), the answer from the only database $A_1^{[k]}$ is identically distributed for all $k \in [1 : K]$. Second, from the correctness constraint (10), from $A_1^{[k]}, \mathcal{F}$, one can decode W_k. Combining these two facts, we have that from $A_1^{[k]}, \mathcal{F}$, one can decode all messages $W_1, \cdots, W_K$. This contradicts the database-privacy constraint (7). Therefore, when $N = 1$ and $K \geq 2$, SPIR is not feasible.

The following theorem states the capacity of SPIR, when we have $N \geq 2$ databases and $K \geq 2$ messages.

Theorem 1 *For SPIR with $K \geq 2$ messages and $N \geq 2$ databases, the capacity is*

$$C_{SPIR} = \begin{cases} 1 - 1/N & \text{if } \rho \geq \frac{1}{N-1} \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

The following observations place Theorem 1 in perspective.

1. We notice a surprising threshold phenomenon in the dependence of SPIR capacity, C_{SPIR}, on the amount of common randomness ρ. When $\rho < \frac{1}{N-1}$, SPIR is not feasible and $C_{\text{SPIR}} = 0$. However, when $\rho \geq \frac{1}{N-1}$, SPIR is not only possible, but the rate can immediately be increased to the maximum possible, i.e., the capacity. Therefore, the minimum common randomness required to achieve any positive rate is already sufficient to achieve the capacity of SPIR. A pictorial illustration of the SPIR capacity and its dependency on the amount of common randomness appears in Figure 2.

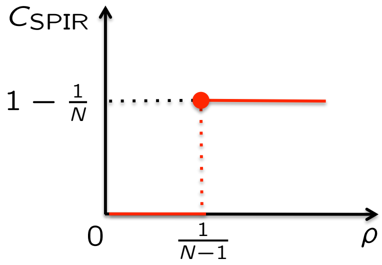

Figure 2: SPIR capacity.

2. The capacity of SPIR is independent of the number of messages, K.

3. When the capacity is non-zero, the capacity is strictly increasing in the number of databases, N, and when N approaches infinity, the capacity approaches 1.

4. It is interesting to compare the capacity of SPIR and the capacity of PIR [3],

$$C_{\text{PIR}} = \left(1 + 1/N + 1/N^2 + \cdots + 1/N^{K-1}\right)^{-1}. \tag{13}$$

We see that the capacity of SPIR is strictly smaller than the capacity of PIR (the additional requirement of preserving database-privacy strictly hurts) and the capacity of PIR approaches the capacity of SPIR when the number of messages, K, approaches infinity (in the large number of messages regime, the penalty vanishes), i.e., $C_{\text{PIR}} > C_{\text{SPIR}}$ for any finite K and $C_{\text{PIR}} \to C_{\text{SPIR}}$ when $K \to \infty$.

5. The achievable scheme presented in Section 4.1.1 has exactly zero error. Further, In the achievability proof for Theorem 1, the message size is $N-1$ bits per message. Therefore, to achieve capacity, message size is not required to approach infinity. By employing the scheme multiple times, we know that when message size is equal to an integer multiple of $N-1$ bits, the capacity is achieved as well. When the message size is not equal to an integer multiple of $N-1$ bits, it turns out that there is a penalty in the form of a ceiling operation. This extension of SPIR to finite length messages is considered in Theorem 3, to be presented in Section 3.3.

6. We note that the converse (upper bound, presented in Section 4.1.2) holds for arbitrary message size L when we require exactly zero error, by replacing the $o(L)$ terms with zero.

7. The extension to unequal message sizes is considered in Section 3.2.

In the following sections, we relax each one of the two assumptions by itself, i.e., equal message sizes and message length L going to infinity.

3.2 Capacity Region of SPIR

In this section, we relax the assumption of equal message sizes, i.e., $l_k, \forall k \in [1:K]$ are arbitrary. Therefore, going beyond the (symmetric) capacity, we wish to characterize the capacity region of SPIR.

When we only have $K=1$ message, similar to the previous section, the capacity region is characterized by the capacity of one message, which is 1. When we only have $N=1$ database and $K \geq 2$ messages, similar to the previous section, SPIR is not feasible and the capacity region is the zero vector. Therefore, we consider $K \geq 2$ messages and $N \geq 2$ databases, where the capacity region of SPIR is characterized in the following theorem. Here the amount of common randomness is normalized with respective to the largest message size.

$$\rho = \frac{H(S)}{\max_{i:i\in[1:K]} H(W_i)} = \frac{H(S)}{\max_{i:i\in[1:K]} l_i L}, \tag{14}$$

Theorem 2 *For SPIR with $K \geq 2$ messages and $N \geq 2$ databases, the capacity region $\mathcal{C}$ is*

$$\mathcal{C} = \left\{(R_1, \cdots, R_K) : R_k \leq \frac{l_k}{\max_i l_i}(1-\frac{1}{N}), \forall k \in [1:K]\right\}, \textit{if} \;\; \rho \geq \frac{1}{N-1} \tag{15}$$

and the zero vector otherwise.

Remark: The optimal (minimum) normalized download cost $D/L = l_k/R_k = \max_i l_i \frac{N}{N-1}$ is the same for each message.

3.3 Capacity of Finite Length SPIR

In this section, we again assume that all messages have the same length, but relax the assumption that L approaches infinity. Instead, we assume that L is an arbitrary finite value. As L is finite, we consider zero error achievable rates and define its supremum as zero error capacity, denoted as C_o. This setting can be obtained from the general problem statement by setting $l_k = 1, \forall k \in [1 : K]$, and L finite.

Similar to Section 3.1, we restrict to $K \geq 2$ and $N \geq 2$ cases as the problem is trivial when $K = 1$ or $N = 1$. The capacity of finite length SPIR is characterized in the following theorem. The relative size of the common randomness, ρ, is defined as in (11).

Theorem 3 *For SPIR with $K \geq 2$ messages, $N \geq 2$ databases, where each message is of size $L \in \mathbb{Z}_+$ symbols, the zero error capacity is*

$$C_{o,LSPIR} = \begin{cases} L/\lceil \frac{L}{C_{SPIR}} \rceil = L/\lceil \frac{L}{1-1/N} \rceil & \textit{if } \rho \geq \frac{\lceil L/(N-1) \rceil}{L} \\ 0 & \textit{otherwise} \end{cases} \tag{16}$$

4 Proofs

4.1 Proof of Theorem 1

4.1.1 Achievability

In this section, we present the scheme that achieves rate $1 - 1/N$, when $\rho = 1/(N-1)$. To this end, we assume each message consists of $N - 1$ bits and each answering string is 1 bit. Specifically, we assume $W_k = (x_{k,1}, \cdots, x_{k,N-1}), \forall k \in [1 : K]$ where each $x_{k,i}, i \in [1 : N-1]$ is one bit. We further assume the entropy of the common random variable S is 1 bit, i.e., S is uniformly distributed over $\{0, 1\}$. Note that S is independent of the messages.

Next we specify the queries. To retrieve W_k privately, the user first generates a random vector of length $(N-1)K$, $[h_{1,1}, \cdots, h_{1,N-1}, \cdots, h_{k,1}, \cdots, h_{K,N-1}]$, where each element is uniformly distributed over $\{0, 1\}$. Then the queries are set as follows.

$$\begin{aligned} Q_1^{[k]} &= [h_{1,1}, \cdots, h_{k,1}, \cdots, h_{k,N-1}, \cdots, h_{K,N-1}] \\ Q_2^{[k]} &= [h_{1,1}, \cdots, h_{k,1} + 1, \cdots, h_{k,N-1}, \cdots, h_{K,N-1}] \\ &\cdots \\ Q_N^{[k]} &= [h_{1,1}, \cdots, h_{k,1}, \cdots, h_{k,N-1} + 1, \cdots, h_{K,N-1}] \end{aligned} \tag{17}$$

The answering strings are generated by using the query vector as the combining coefficients and producing the corresponding linear combination of message bits. We further add the common random variable to each answer.

$$\begin{aligned} A_1^{[k]} &= \sum_{j=1}^{K} \sum_{i=1}^{N-1} h_{j,i} x_{j,i} + S \\ A_2^{[k]} &= \sum_{j=1}^{K} \sum_{i=1}^{N-1} h_{j,i} x_{j,i} + x_{k,1} + S \\ &\cdots \end{aligned}$$

$$A_N^{[k]} = \sum_{j=1}^{K} \sum_{i=1}^{N-1} h_{j,i} x_{j,i} + x_{k,N-1} + S \tag{18}$$

The user obtains $x_{k,i}, i \in [1 : N-1]$ by subtracting $A_1^{[k]}$ from $A_{i+1}^{[k]}$. Therefore, the correctness condition is satisfied.

Privacy of the user is guaranteed because each query is independent of the desired message index k. This is because regardless of the desired message index k, each of the query vectors $Q_n^{[k]}, \forall n$ is individually comprised of elements that are i.i.d. uniform over $\{0, 1\}$. Thus, each database learns nothing about which message is requested.

We now show that database-privacy is preserved as well.

$$\begin{aligned}
& I(W_{\bar{k}}\, ; A_1^{[k]}, A_2^{[k]}, \cdots, A_N^{[k]}, Q_{1:N}^{[k]}, \mathcal{F}) && (19)\\
= \; & I(W_{\bar{k}}\, ; A_1^{[k]}, A_1^{[k]} + x_{k,1}, \cdots, A_1^{[k]} + x_{k,N-1}, Q_{1:N}^{[k]}, \mathcal{F}) && (20)\\
= \; & I(W_{\bar{k}}\, ; A_1^{[k]}, x_{k,1}, \cdots, x_{k,N-1}, Q_{1:N}^{[k]}, \mathcal{F}) && (21)\\
= \; & I(W_{\bar{k}}\, ; A_1^{[k]}, W_k, Q_{1:N}^{[k]}, \mathcal{F}) && (22)\\
\overset{(3)(4)}{=} \; & I(W_{\bar{k}}\, ; A_1^{[k]} | W_k, Q_{1:N}^{[k]}, \mathcal{F}) && (23)\\
= \; & 0 && (24)
\end{aligned}$$

where in each step, the transformation on the variables is invertible such that mutual information remains the same. The last step follows from the independence of the messages and the common randomness (refer to (3)).

Note that because each answering string is 1 bit and the message is $L = N - 1$ bits, the rate achieved is $(N-1)/N = 1 - 1/N$ which matches the capacity. Also note that only the minimum threshold amount of common randomness is utilized, i.e., $\rho = 1/(N-1)$. ■

4.1.2 Converse

Although Theorem 1 restricts to the setting where $l_k = 1, \forall k \in [1 : K]$, we do not assume this at the beginning in the proof of converse. This will make the converse general such that some of the intermediate steps can be used in the converse proofs of Theorem 2 and Theorem 3 as well.

For the converse we allow any feasible SPIR scheme, and prove that its rate cannot be larger than C_{SPIR}. Let us start with two lemmas that will be used later in the proof.

Lemma 1

$$\begin{aligned}
H(A_n^{[k]} | W_k, Q_n^{[k]}) &= H(A_n^{[k']} | W_k, Q_n^{[k']}) && (25)\\
H(A_n^{[k]} | Q_n^{[k]}) &= H(A_n^{[k']} | Q_n^{[k']}), \quad \forall n \in [1 : N] && (26)
\end{aligned}$$

Proof: Since the proofs of (25) and (26) follow from the same arguments, here we will present only the proof of (25). From the User-Privacy constraint (6) we know that $\forall k \in [1 : K], \forall n \in [1 : N]$, $I(\theta; A_n^{[\theta]}, W_k, Q_n^{[\theta]}) = 0$. Therefore, we must have $\forall k' \in [1 : K]$,

$$\begin{aligned}
H(A_n^{[k]}, W_k, Q_n^{[k]}) &= H(A_n^{[k']}, W_k, Q_n^{[k']}) && (27)\\
H(W_k, Q_n^{[k]}) &= H(W_k, Q_n^{[k']}) && (28)
\end{aligned}$$

Combining (27) and (28), we obtain $H(A_n^{[k]} | W_k, Q_n^{[k]}) = H(A_n^{[k']} | W_k, Q_n^{[k']})$. ■

Lemma 2

$$H(A_n^{[k]}|W_k, \mathcal{F}, Q_n^{[k]}) = H(A_n^{[k]}|W_k, Q_n^{[k]}), \quad \forall n \in [1:N] \tag{29}$$

Proof: Since

$$H(A_n^{[k]}|W_k, Q_n^{[k]}) - H(A_n^{[k]}|W_k, \mathcal{F}, Q_n^{[k]}) = I(A_n^{[k]}; \mathcal{F}|W_k, Q_n^{[k]}) \geq 0, \tag{30}$$

we only need to prove $I(A_n^{[k]}; \mathcal{F}|W_k, Q_n^{[k]}) \leq 0$.

$$\begin{aligned}
& I(A_n^{[k]}; \mathcal{F}|W_k, Q_n^{[k]}) && (31)\\
\leq\; & I(A_n^{[k]}, W_1, \cdots, W_K, S; \mathcal{F}|W_k, Q_n^{[k]}) && (32)\\
=\; & I(W_1, \cdots, W_K, S; \mathcal{F}|W_k, Q_n^{[k]}) + \underbrace{I(A_n^{[k]}; \mathcal{F}|W_1, \cdots, W_K, S, W_k, Q_n^{[k]})}_{=0} && (33)\\
\leq\; & I(W_1, \cdots, W_K, S; \mathcal{F}, Q_n^{[k]}) && (34)\\
=\; & 0 && (35)
\end{aligned}$$

where the second term in (33) is zero because of (5) and (35) follows from (3), (4). ■

The proof for $R \leq C_{\textbf{SPIR}}$. For every feasible SPIR scheme, we must satisfy the database-privacy constraint (7),

$$0 = I(W_{\overline{k'}}; A_1^{[k']}, \cdots, A_N^{[k']}, Q_1^{[k']}, \cdots, Q_N^{[k']}, \mathcal{F}) \tag{36}$$

such that $\forall n \in [1:N], \forall k \in [1:K], k \neq k'$,

$$\begin{aligned}
0 &= I(W_k; A_n^{[k']}, Q_n^{[k']}) && (37)\\
&= H(A_n^{[k']}|Q_n^{[k']}) - H(A_n^{[k']}|W_k, Q_n^{[k']}) && (38)\\
&\overset{(25)}{=} H(A_n^{[k']}|Q_n^{[k']}) - H(A_n^{[k]}|W_k, Q_n^{[k]}) && (39)
\end{aligned}$$

Now, consider the answering strings $A_1^{[k]}, \cdots, A_N^{[k]}$, from which we can decode W_k.

$$\begin{aligned}
l_k L = H(W_k) &\overset{(3)}{=} H(W_k|\mathcal{F}) && (40)\\
&\overset{(10)}{=} I(W_k; A_1^{[k]}, \cdots, A_N^{[k]}|\mathcal{F}) + o(L)L && (41)\\
&= H(A_1^{[k]}, \cdots, A_N^{[k]}|\mathcal{F}) - H(A_1^{[k]}, \cdots, A_N^{[k]}|W_k, \mathcal{F}) + o(L)L && (42)\\
&\overset{(4)}{\leq} H(A_1^{[k]}, \cdots, A_N^{[k]}|\mathcal{F}) - H(A_n^{[k]}|W_k, \mathcal{F}, Q_n^{[k]}) + o(L)L && (43)\\
&\overset{(29)}{=} H(A_1^{[k]}, \cdots, A_N^{[k]}|\mathcal{F}) - H(A_n^{[k]}|W_k, Q_n^{[k]}) + o(L)L && (44)\\
&\overset{(39)}{=} H(A_1^{[k]}, \cdots, A_N^{[k]}|\mathcal{F}) - H(A_n^{[k']}|Q_n^{[k']}) + o(L)L && (45)\\
&\overset{(26)}{=} H(A_1^{[k]}, \cdots, A_N^{[k]}|\mathcal{F}) - H(A_n^{[k]}|Q_n^{[k]}) + o(L)L && (46)\\
&\overset{(4)}{\leq} H(A_1^{[k]}, \cdots, A_N^{[k]}|\mathcal{F}) - H(A_n^{[k]}|\mathcal{F}) + o(L)L && (47)
\end{aligned}$$

Adding (47) for all $n \in [1:N]$, we have

$$Nl_kL \leq NH(A_1^{[k]}, \cdots, A_N^{[k]}|\mathcal{F}) - \sum_{n\in[1:N]} H(A_n^{[k]}|\mathcal{F}) + o(L)L \tag{48}$$

$$\leq (N-1)\, H(A_1^{[k]}, \cdots, A_N^{[k]}|\mathcal{F}) + o(L)L \tag{49}$$

$$\leq (N-1) \sum_{n=1}^{N} H(A_n^{[k]}) + o(L)L \tag{50}$$

$$\leq (N-1)\, D + o(L)L \tag{51}$$

$$R_k = \frac{l_kL}{D} \leq 1 - \frac{1}{N} \quad \text{(Letting } L \to \infty\text{)} \tag{52}$$

Thus, the rate of any feasible SPIR scheme cannot be more than C_{SPIR}.

The proof for $\rho \geq 1/(N-1)$. Suppose a feasible SPIR scheme exists that achieves a non-zero SPIR rate. Then we will show that it must have $\rho \geq 1/(N-1)$. Consider the answering strings $A_1^{[k]}, \cdots, A_N^{[k]}$, from which we can decode W_k. From the database-privacy constraint, we have

$$0 = I(W_{\overline{k}}\,; A_1^{[k]}, \cdots, A_N^{[k]}, \mathcal{F}) \tag{53}$$

$$\overset{(3)}{=} I(W_{\overline{k}}\,; A_1^{[k]}, \cdots, A_N^{[k]}|\mathcal{F}) \tag{54}$$

$$\overset{(10)}{=} I(W_{\overline{k}}\,; A_1^{[k]}, \cdots, A_N^{[k]}, W_k|\mathcal{F}) + o(L)L \tag{55}$$

$$\overset{(3)}{=} I(W_{\overline{k}}\,; A_1^{[k]}, \cdots, A_N^{[k]}|W_k, \mathcal{F}) + o(L)L \tag{56}$$

$$\geq I(W_{\overline{k}}\,; A_n^{[k]}|W_k, \mathcal{F}) + o(L)L \tag{57}$$

$$= H(A_n^{[k]}|W_k, \mathcal{F}) - H(A_n^{[k]}|W_1, \cdots, W_K, \mathcal{F}) + o(L)L \tag{58}$$

$$\overset{(4)(5)}{=} H(A_n^{[k]}|W_k, \mathcal{F}) - H(A_n^{[k]}|W_1, \cdots, W_K, \mathcal{F}) + H(A_n^{[k]}|W_1, \cdots, W_K, \mathcal{F}, S) + o(L)L \tag{59}$$

$$= H(A_n^{[k]}|W_k, \mathcal{F}) - I(S; A_n^{[k]}|W_1, \cdots, W_K, \mathcal{F}) + o(L)L \tag{60}$$

$$\overset{(4)}{\geq} H(A_n^{[k]}|W_k, \mathcal{F}, Q_n^{[k]}) - H(S) + o(L)L \tag{61}$$

$$\overset{(29)}{=} H(A_n^{[k]}|W_k, Q_n^{[k]}) - H(S) + o(L)L \tag{62}$$

$$\overset{(39)}{=} H(A_n^{[k']}|Q_n^{[k']}) - H(S) + o(L)L \tag{63}$$

$$\overset{(26)}{=} H(A_n^{[k]}|Q_n^{[k]}) - H(S) + o(L)L \tag{64}$$

Adding (64) for $n \in [1:N]$, we have

$$0 \geq \sum_{n\in[1:N]} H(A_n^{[k]}|Q_n^{[k]}) - NH(S) + o(L)L \tag{65}$$

$$\geq H(A_1^{[k]}, \cdots, A_N^{[k]}|\mathcal{F}) - NH(S) + o(L)L \tag{66}$$

$$\overset{(49)}{\geq} \frac{N}{N-1} l_kL - NH(S) + o(L)L \tag{67}$$

$$\Rightarrow H(S) \quad \geq \quad \frac{1}{N-1} l_k L + o(L)L \tag{68}$$

$$\Rightarrow \rho \quad = \quad \frac{H(S)}{l_k L} \geq \frac{1}{N-1} \quad \text{(Letting } L \to \infty\text{)} \tag{69}$$

Thus, the amount of common randomness relative to the message size of any feasible SPIR scheme cannot be less than $1/(N-1)$.

4.2 Proof for Theorem 2

4.2.1 Achievability

Without loss of generality, we assume that $l_1 \leq l_2 \leq \cdots \leq l_K$. As a result, we need to prove that for W_k, rate $(1-1/N)l_k/l_K$ is achievable, when $\rho = 1/(N-1)$. Here is such a scheme. We set $L = N-1$. We divide each message into K sub-messages and for each sub-message, we use the SPIR scheme (17), (18). That is

$$W_k \quad = \quad (W_k(1), W_k(2), \cdots, W_k(K)) \tag{70}$$

$$W_k(i) \quad \in \quad \mathbb{F}_2^{(l_i - l_{i-1})L \times 1} \quad \text{(Define } l_0 = 0\text{)} \tag{71}$$

Note that $H(W_k) = l_k L$, so we set $W_k(i)$ to zero vectors, when $i \in [k+1:K]$. For sub-messages $(W_1(i), W_2(i), \cdots, W_K(i))$, we employ the SPIR scheme (17), (18) $l_k - l_{k-1}$ times independently. Therefore the number of bits downloaded, denoted as $D_k(i)$, and the amount of common randomness used, $H(S(i))$, are obtained as follows. $\forall i \in [1:K]$,

$$D_k(i) \quad = \quad (l_k - l_{k-1})L/(1-1/N) = (l_k - l_{k-1})N \tag{72}$$

$$H(S(i)) \quad = \quad (l_k - l_{k-1})L/(N-1) = l_k - l_{k-1} \tag{73}$$

The schemes for each sub-message are also independent. As our scheme is a concatenation of multiple independent correct and private SPIR schemes, the overall concatenated scheme is also correct and private. The proof is similar to Theorem 4 of [17] and is thus omitted.

Finally, the rate and the amount of common randomness are as follows.

$$R_k \quad = \quad \frac{H(W_k)}{D} = \frac{\sum_{i=1}^K H(W_k(i))}{\sum_{i=1}^K D_k(i)} \overset{(72)}{=} \frac{l_k L}{l_K N} = \frac{l_k}{l_K}\left(1 - \frac{1}{N}\right) \tag{74}$$

$$\rho \quad = \quad \frac{H(S)}{H(W_K)} = \frac{\sum_{i=1}^K H(S(i))}{l_K L} \overset{(73)}{=} \frac{l_K}{l_k L} = \frac{1}{N-1} \tag{75}$$

Therefore, the achievability of Theorem 2 is proved.

4.2.2 Converse

The converse proof is almost identical to that of Theorem 1. Note that (51) holds for all l_k and all k, we have

$$N \max_i l_i \quad \leq \quad (N-1)D/L \tag{76}$$

$$\Rightarrow R_k \quad = \quad l_k L/D \leq \frac{l_k}{\max_i l_i}\left(1 - \frac{1}{N}\right) \tag{77}$$

Therefore the rate bound is proved. The common randomness bound is identical to (69). Note that (69) holds for all $k \in [1:K]$.

4.3 Proof for Theorem 3

4.3.1 Achievability

Suppose $L = G_1(N-1)+L_1$, where $G_1 = \lfloor L/(N-1) \rfloor$ and $L_1 \in [0 : N-2]$. Note that the capacity achieving scheme for SPIR when L is not restricted is based on dividing the messages to blocks of length $N-1$ (refer to Theorem 1). The optimal scheme for finite L setting is constructed by first using the capacity achieving SPIR scheme G_1 times to retrieve $G_1(N-1)$ bits, and then for the remaining L_1 bits, we use the capacity achieving SPIR schemes with only $L_1+1 \leq N$ databases (say, the first L_1+1 databases), if $L_1 \geq 1$. Otherwise if $L_1 = 0$, then we are done. Note that for the SPIR scheme that uses only L_1+1 databases, the rate is $1-1/(L_1+1)$, the message size is $L_1+1-1 = L_1$ bits, and the common randomness ratio is $\rho = 1/(L_1+1-1) = 1/L_1$. Therefore, overall, the rate and the amount of common randomness are as follows.

$$R = \begin{cases} \frac{G_1(N-1)}{G_1N} = 1 - \frac{1}{N}, & \text{if } L_1 = 0, \\ \frac{G_1(N-1)+L_1}{G_1N+L_1+1}, & \text{otherwise.} \end{cases} \tag{78}$$

$$\rho = \begin{cases} \frac{G_1}{G_1(N-1)} = \frac{1}{N-1}, & \text{if } L_1 = 0, \\ \frac{G_1+1}{G_1(N-1)+L_1} = \frac{\lfloor L/(N-1) \rfloor + 1}{L}, & \text{otherwise.} \end{cases} \tag{79}$$

$$= \frac{\lceil L/(N-1) \rceil}{L} \tag{80}$$

Next, we prove that the rate achieved in (78) matches that in Theorem 3, i.e., $L/\lceil \frac{L}{1-1/N} \rceil$. When $L_1 = 0$ (L is an integer multiple of $N-1$), the claim follows trivially. Hereafter, we consider $L_1 > 0$. It suffices to show that the download cost, $D = L/R = G_1N + L_1 + 1$, satisfies $D \in [\frac{L}{1-1/N}, \frac{L}{1-1/N} + 1)$. In the converse of Theorem 1, we have showed that for arbitrary L and all SPIR schemes, $D \geq \frac{L}{1-1/N}$ holds. So we are left to show that $D < \frac{L}{1-1/N} + 1$.

$$D = G_1N + L_1 + 1 \tag{81}$$

$$< \frac{G_1(N-1)+L_1}{1-1/N} + 1 \quad (N \geq 2) \tag{82}$$

$$= \frac{L}{1-1/N} + 1 \tag{83}$$

Therefore the achievability proof is complete.

4.3.2 Converse

We show for fixed finite L, the achievable rate $R \leq L/\lceil \frac{L}{1-1/N} \rceil$. Equivalently, it suffices to prove that the download cost $D = L/R \geq \lceil \frac{L}{1-1/N} \rceil \geq \frac{L}{1-1/N}$, which follows from Theorem 1. Note that Theorem 1 holds when we require exactly zero error. Also note that since the downloads are assumed to be in terms of symbols from the same field as the message symbols, the download cost must be an integer value.

We are left to prove the common randomness bound. Similar to the download cost, which is restricted to be integers, in the finite length regime the amount of common randomness, ρL is restricted to take integer values as well. Therefore, from (69), we have $\rho L \geq \lceil L/(N-1) \rceil$ (for this setting, $l_k = 1$).

5 Conclusion

For K messages and N databases, the capacity of SPIR was shown to be $C = 1 - 1/N$. In order to achieve any positive rate for SPIR, the minimum amount of common randomness needed among the databases was shown to be $1/(N-1)$ bits per message bit. Remarkably, this is also sufficient to achieve the capacity of SPIR. The insights extend to settings with unequal message sizes and finite length messages.

References

[1] B. Chor, O. Goldreich, E. Kushilevitz, and M. Sudan, "Private information retrieval," in *Proceedings of the 36th Annual Symposium on Foundations of Computer Science*, 1995, pp. 41–50.

[2] B. Chor, E. Kushilevitz, O. Goldreich, and M. Sudan, "Private Information Retrieval," *Journal of the ACM (JACM)*, vol. 45, no. 6, pp. 965–981, 1998.

[3] H. Sun and S. A. Jafar, "The Capacity of Private Information Retrieval," *arXiv preprint arXiv:1602.09134*, 2016.

[4] Y. Gertner, Y. Ishai, E. Kushilevitz, and T. Malkin, "Protecting data privacy in private information retrieval schemes," in *Proceedings of the thirtieth annual ACM symposium on Theory of computing.* ACM, 1998, pp. 151–160.

[5] W. Gasarch, "A Survey on Private Information Retrieval," in *Bulletin of the EATCS*, 2004.

[6] S. Yekhanin, "Private Information Retrieval," *Communications of the ACM*, vol. 53, no. 4, pp. 68–73, 2010.

[7] J. Katz and L. Trevisan, "On the efficiency of local decoding procedures for error-correcting codes," in *Proceedings of the thirty-second annual ACM symposium on Theory of computing.* ACM, 2000, pp. 80–86.

[8] S. Yekhanin, "Locally Decodable Codes and Private Information Retrieval Schemes," Ph.D. dissertation, Massachusetts Institute of Technology, 2007.

[9] Y. Ishai, E. Kushilevitz, R. Ostrovsky, and A. Sahai, "Batch codes and their applications," in *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing.* ACM, 2004, pp. 262–271.

[10] Y. Ishai and E. Kushilevitz, "On the hardness of information-theoretic multiparty computation," in *Advances in Cryptology-EUROCRYPT 2004.* Springer, 2004, pp. 439–455.

[11] M. O. Rabin, "How to exchange secrets with oblivious transfer." 1981.

[12] S. Even, O. Goldreich, and A. Lempel, "A randomized protocol for signing contracts," *Communications of the ACM*, vol. 28, no. 6, pp. 637–647, 1985.

[13] J. Kilian, "Founding crytpography on oblivious transfer," in *Proceedings of the twentieth annual ACM symposium on Theory of computing.* ACM, 1988, pp. 20–31.

[14] Y. Ishai, M. Prabhakaran, and A. Sahai, "Founding cryptography on oblivious transfer–efficiently," in *Annual International Cryptology Conference.* Springer, 2008, pp. 572–591.

[15] R. Ahlswede and I. Csiszár, "On oblivious transfer capacity," in *Information Theory, Combinatorics, and Search Theory.* Springer, 2013, pp. 145–166.

[16] A. C. Nascimento and A. Winter, "On the oblivious-transfer capacity of noisy resources," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2572–2581, 2008.

[17] H. Sun and S. A. Jafar, "Optimal Download Cost of Private Information Retrieval for Arbitrary Message Length," *arXiv preprint arXiv:1610.03048*, 2016.