

## Lecture 3: Lower Bounds for PIR

Scribe: Diana Vieira Fernandes

February 6, 2024

### 1 Lower Bounds for PIR

In recent lectures, we showed how to construct a 2-server Private Information Retrieval (PIR) schemes with sublinear bandwidth, without relying on any cryptographic assumptions. One question is whether we can achieve sublinear bandwidth in the single-server setting. We know that using cryptography (e.g. fully homomorphic encryption), we can easily get a single-server PIR scheme with constant bandwidth. However, is it possible to get sublinear bandwidth without cryptographic assumptions? In today's lecture, we will prove that

1. *A perfect (or even statistical) PIR scheme without cryptographic assumptions requires linear bandwidth;*
2. We will prove a even stronger version of the lower bound, that is, *a single-server PIR scheme with non-trivial bandwidth implies Oblivious Transfer (OT).*

The high-level idea of the first lower bound is to show that there exists a computationally unbounded extractor, such that given any database  $DB$  and any query  $q$ , the extractor can extract the whole database given the communication script. This shows that the communication script has at least  $n$ -bit of entropy for a randomly sampled  $n$ -bit database, and thus it is at least  $n$ -bit long.

The idea of the second lower bound is construct an OT scheme from a PIR scheme with non-trivial bandwidth. The challenge in the proof arises from the fact that PIR has only one-sided privacy, but OT has two-sided privacy (as explained more later).

#### 1.1 Unconditional PIR Requires Linear Bandwidth

The following theorem has been a folklore lower bound and is formalized by Damgård, Larsen and Nielsen [?].

**Theorem 1** ([?]). *1-server PIR scheme with perfect correctness and perfect privacy must have  $\Omega(n)$  bandwidth where  $n = |DB|$ . Further, this lower bound holds regardless of the number of rounds or client/server computation.*

In our lecture, we will only provide the proof for PIR schemes with perfect privacy and correctness. However, Damgård, Larsen and Nielsen [?] generalized the proof to statistical correctness and privacy as well.

**Notations.** Given any database DB of size  $n$ , any client query  $i \in [n]$ , we denote the PIR protocol's communication script as  $\langle \text{Server}(\text{DB}, r_1) \leftrightarrow \text{Client}(i, r_2) \rangle$  (if it's multi-round, we just concatenate all the exchanged messages). Here,  $r_1$  and  $r_2$  are the random coins consumed by the server and the client, respectively. By the definition of perfect correctness, there exists an algorithm **Coutput**, such that  $\text{Coutput}(\langle \text{Server}(\text{DB}, r_1) \leftrightarrow \text{Client}(i, r_2) \rangle, i, r_2) = \text{DB}[i]$  with probability 1. That is, the **Coutput** is the client-side algorithm to construct the final answer in the PIR protocol.

*Proof.* We want to show that this protocol must have at least  $n$  bandwidth and the way we are going to propose an encoding scheme. We are going to see that by perfect privacy and perfect correctness of the PIR scheme, given any possible communication script  $C$ , a computationally unbounded extractor can extract all original database bits. Then, the communication script  $C$  can actually be seen as the codeword for the database, and the decoder just runs the extractor to reconstruct the database. It gives a uniquely decodable scheme and by Shannon's source coding theorem, the expected length of the codeword has to be at least  $n$ -bit.

We now define the encoding scheme and the decoding scheme. Notice that the encoder and the decoder algorithms do not need to be efficient (polynomially bounded).

- *Encoding.* Given any  $n$ -bit database DB, the encoding for DB is the shortest possible communication script  $C$  that is compatible with the DB. That is, there exist  $r_1, r_2 \in \{0, 1\}^*$ ,  $i \in [n]$  such that  $\langle \text{Server}(\text{DB}, r_1) \leftrightarrow \text{Client}(i, r_2) \rangle$ .
- *Decoding.* Given any code word  $C$ , the decoder algorithm will reconstruct the  $i$ -th bit of DB as follows. The algorithm fixes  $i$  and enumerates  $r_2$  until  $C$  is a possible communication script given the client's inputs are  $i, r_2$ . Then, the  $i$ -th bit of DB is reconstructed by  $\text{Coutput}(C, i, r_2)$ .

**Claim 2.** *The decoding algorithm terminates in finite time.*

This claim is surprisingly implied by the perfect privacy of the PIR scheme – given a fixed DB, the perfect privacy requires that any possible communication script  $C$  should happen with non-zero probability for all client's query indices  $i \in [n]$ . Otherwise, the adversary, who acts as the server, can exclude at least one possible index when it sees  $C$ , resulting in a privacy violation. Since  $C$  is possible for all possible  $i$ , the decoding algorithm is guaranteed to find a possible  $r_2$  in finite time, such that  $C$  is compatible with  $i$  and  $r_2$ .

**Claim 3.** *The decoding algorithm always decodes correctly.*

Since  $C$  is compatible with the original DB and the client's query  $j$ , the perfect correctness of the PIR scheme implies that  $\text{Coutput}(C, i, r_2) = \text{DB}[i]$ .

Based on the claims, we show that this encoding-decoding scheme is a uniquely decodable code, i.e.,  $\forall \text{DB} \in \{0, 1\}^n$ ,  $\Pr[\text{Decode}(\text{Encode}(\text{DB})) = \text{DB}] = 1$ . Then, by Shannon's source coding theorem, we have that

$$\mathbb{E}_{\text{DB} \leftarrow \{0, 1\}^n} [\text{Encode}(\text{DB})] \geq H(\text{DB}) = n.$$

Here,  $H(\text{DB})$  is the entropy of DB. Then, we show that the expected length for the shortest possible communication script  $C$  is at least  $n$  bits.  $\square$

**Remark 1.** [?] extended the proof to statistical-correct and statistical-private PIR schemes.

## 1.2 PIR with non-trivial BW implies Oblivious Transfer

Crescenzo, Malkin, and Ostrovsky [?] showed that a single-server PIR with non-trivial bandwidth implies Oblivious Transfer (OT). Since Oblivious Transfer implies the existence of One-Way Function (OWF), this essentially says that OWF is needed for any single-server PIR with non-trivial BW. In this lecture, we focus on the first step of this proof that shows that a single-server PIR with non-trivial bandwidth implies an honest-receiver OT.

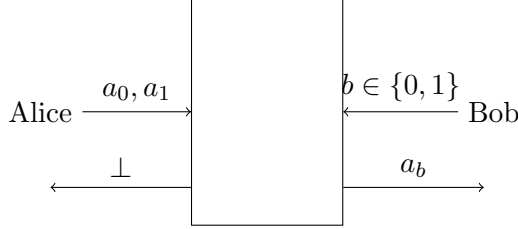


Figure 1: Protocol Illustration

**A  $\binom{2}{1}$ -OT protocol between Alice and Bob is as follows:**

- Alice has two numbers  $a_0$  and  $a_1$ , Bob has a single bit  $b$ .
- Bob wishes to learn  $a_b$  without Alice learning his choice of  $b$  (Bob's privacy).
- Moreover, Bob does not learn any information about  $a_{1-b}$  (Alice's privacy).

**Notations:** The protocol's execution is denoted as

$$\begin{pmatrix} \text{View}_A \\ \text{View}_B \\ \text{out}_B \end{pmatrix} = \langle \text{Alice}(1^\lambda, a_0, a_1) \leftrightarrow \text{Bob}(1^\lambda, b) \rangle,$$

where

- View  $A$  is Alice's view, including the random coins and the observed messages;
- View  $B$  is Bob's view, including the random coins and the observed messages;
- $\text{out}_B$  is Bob's output.

**Definition 4** (Correctness of OT). *Given any  $a_0, a_1, b \in \{0, 1\}^3$ ,*

$$\Pr[(\cdot, \cdot, \text{out}_B) \leftarrow \langle \text{Alice}(1^\lambda, a_0, a_1) \leftrightarrow \text{Bob}(1^\lambda, b) \rangle : \text{out}_B = a_b] = 1 - \text{negl}(\lambda).$$

That is, Bob should output the correct bit he tries to fetch with overwhelming probability.

We now see the privacy definitions. We focus on the case where Alice can be malicious and Bob is honest but curious, which means Alice may deviate from the protocol, but Bob follows the protocol honestly and only tries to break the privacy based on his local view.

**Definition 5** (Sender's privacy against an honest-but-curious receiver). *For every uniform probabilistic polynomial-time (PPT) reconstruction algorithm  $R$ , there exists a negligible function  $\text{negl}$  such that for all inputs  $a_0, a_1, b \in \{0, 1\}$ , the following holds:*

$$\Pr \left[ \begin{array}{c} (a_0, a_1) \xleftarrow{\$} \{0, 1\}^2 \\ (\cdot, \text{View}_B, \cdot) \leftarrow \langle \text{Alice}(1^\lambda, a_0, a_1) \leftrightarrow \text{Bob}(1^\lambda, b) \rangle : \\ R(1^\lambda, \text{View}_B) = a_{1-b} \end{array} \right] < \frac{1}{2} + \text{negl}(\lambda)$$

**Definition 6** (Receiver’s privacy against a malicious sender). *There exists a negligible function  $\text{negl}$  that for any uniform PPT adversary  $A^*$  representing a malicious sender (Alice) and any uniform PPT reconstruction algorithm  $R^*$ , the following holds:*

$$\Pr \left[ \begin{array}{c} b \xleftarrow{\$} \{0, 1\} \\ (\text{View}_A, \cdot, \cdot) \leftarrow \langle A^*(1^\lambda, a_0, a_1) \leftrightarrow \text{Bob}(1^\lambda, b) \rangle : \\ R^*(1^\lambda, \text{View}_A) = b \end{array} \right] \leq \frac{1}{2} + \text{negl}(\lambda)$$

**Implications of OT.** The major implication of OT is shown by Kilian [?], and Impagliazzo and Rudich [?]:

- [?]: OT is complete for constructing MPC for any computational task (in the presence of a dishonest majority).
- [?] : OT  $\xRightarrow{\text{Black-box Reduction}}$  OWF
- [?] : OWF  $\not\xRightarrow{\text{Black-box Reduction}}$  OT

**Theorem 7.** [?] *Constructing OT from one-way functions via black-box reductions is impossible. Such a reduction does not examine the internal structure of the circuit realizing the one-way function.*

**Remark 2.** *In general, PIR is different from OT in the following sense:*

1. *Classical PIR has only one-sided privacy;*
2. *Classical PIR has efficiency requirements on the communication (otherwise it will be trivial to simply download the whole database).*

*OT is sometimes referred to as “symmetric PIR”.*

**Constructing OT from PIR [?].** Crescenzo, Malkin, and Ostrovsk [?] showed that PIR with non-trivial bandwidth implies the possibility of constructing Oblivious Transfer (OT) even with a malicious receiver. In this class, we are just going to do the easier version which shows that PIR implies an honest-but-curious Bob OT.

The construction of OT from PIR, is shown in Figure 2. The correctness and Bob’s privacy are easy to show by the correctness and the privacy of the PIR scheme. It remains to show Alice’s privacy (sender privacy).

The high-level idea is as follows. In the  $j$ -th repetition of the PIR scheme, since we assume the communication of the PIR scheme is sublinear, Bob cannot learn every index of the database with overwhelming certainty (with a simple argument of information entropy). Then, if Bob samples a random index, the corresponding database entry remains some uncertainty to Bob. Moreover, Alice will cover the  $a_{1-b}$  bit with all the random database indices chosen by Bob, essentially amplifying the uncertainty. Then, Bob cannot recover  $a_{1-b}$  with a non-negligible advantage over random guessing. We will use the following lemma to show the “amplification” effect.

**Lemma 8.** *Suppose  $X_1, \dots, X_n$  are binary random variables such that for each  $i$ ,  $\Pr[X_i = 1] = \frac{1}{2} + \delta$ ,  $\delta \in (-\frac{1}{2}, \frac{1}{2})$ . Then, we have that*

$$\Pr[X_i \oplus \dots \oplus X_n = 1] = \frac{1}{2} + \delta(2\delta)^{n-1}.$$

**Phase 1.** Repeat the following process  $m$  times (indexed by  $j$ ):

- Alice samples  $\text{DB}_j \xleftarrow{\$} \{0, 1\}^k$
- Bob samples  $i_j \xleftarrow{\$} [k]$
- Alice and Bob execute the PIR protocol acting as the server and the client, respectively. Let  $x_j$  be Bob's output.

**Phase 2.**

- Bob samples  $i'_1, \dots, i'_m \xleftarrow{\$} [k]^m$
- Bob sends the following tuples to Alice:

$$\begin{cases} (i_1, \dots, i_m), (i'_1, \dots, i'_m), & \text{if } b = 0 \\ (i'_1, \dots, i'_m)(i_1, \dots, i_m), & \text{if } b = 1 \end{cases}$$

- Alice parses the message as  $(t_1^0, \dots, t_m^0), (t_1^1, \dots, t_m^1)$ .
- Alice returns these two bits to Bob:

$$c_0 \leftarrow a_0 \oplus \text{DB}_j[t_1^0] \oplus \dots \oplus \text{DB}_j[t_m^0]$$

$$c_1 \leftarrow a_0 \oplus \text{DB}_j[t_1^1] \oplus \dots \oplus \text{DB}_j[t_m^1]$$

- Bob can reconstruct  $a_b$  from  $c_b$  because it knows  $x_1 = \text{DB}[i_1^b], \dots, x_m = \text{DB}[i_m^b]$ .

Figure 2: Constructing an honest-but-curious Bob OT from PIR.

*Proof.* Base Case ( $n = 1$ ): For  $n = 1$ , the statement simplifies to  $\Pr[X_1 = 1] = \frac{1}{2} + \delta$ , which is true by definition.

Inductive Step: Assume the lemma is true for  $n$ , i.e.,  $\Pr[X_1 \oplus \dots \oplus X_n = 1] = \frac{1}{2} + \delta(2\delta)^{n-1}$ . We need to prove it for  $n + 1$ .

Given:

$$\frac{1}{2} + \gamma_{n+1} = \left(\frac{1}{2} + \gamma_n\right) \left(\frac{1}{2} + \delta\right) + \left(\frac{1}{2} - \gamma_n\right) \left(\frac{1}{2} - \delta\right)$$

Expanding this, we get:

$$\frac{1}{2} + \gamma_{n+1} = \frac{1}{4} + \frac{\gamma_n}{2} + \frac{\delta}{2} + \gamma_n\delta + \frac{1}{4} - \frac{\gamma_n}{2} - \frac{\delta}{2} + \gamma_n\delta$$

$$\gamma_{n+1} = 2\gamma_n\delta$$

By the inductive hypothesis,  $\gamma_n = \delta(2\delta)^{n-1}$ , so:

$$\gamma_{n+1} = 2\delta(2\delta)^{n-1}\delta = \delta(2\delta)^n$$

Thus,

$$\Pr[X_1 \oplus \dots \oplus X_{n+1} = 1] = \frac{1}{2} + \delta(2\delta)^n$$

□

For the rest of the proof, we only need to show that in a single copy of the PIR scheme, Bob can only guess the random index's value with probability no more than  $\frac{1}{2} + \delta$  with some non-negligible  $\delta$ , then using the XOR amplification lemma is sufficient to prove the sender's privacy. Given any  $i \in [k]$ , any PPT reconstruction algorithm  $R$ , consider the following experiment.

- $\text{Expt}(1^\lambda, k, \text{coin}_A, \text{coin}_B, i, R)$ :
  - $\text{DB} \xleftarrow{\$} \{0, 1\}^k$
  - $(\cdot, \text{View}_B) \leftarrow \langle \text{PIR.Alice}(1^\lambda, \text{DB}, \text{coin}_A) \leftrightarrow \text{PIR.Bob}(1^\lambda, i, \text{coin}_B) \rangle$
  - $r \xleftarrow{\$} [k]$ ;
  - output 1 if  $R(1^\lambda, \text{View}_B, r, \text{coin}_B) = \text{DB}[r]$

**Claim 9.** Let  $p = \Pr[\text{Expt}(1^\lambda, k, \text{coin}_A, \text{coin}_B, i, R) = 1]$ . Then,

$$H(p) \geq \frac{k-l}{k},$$

where  $H(p)$  is the binary entropy of  $p$ ,  $k$  is the database size, and  $l$  is the communication script length in the PIR scheme.

*Proof.* Let  $\text{Comm}$  be the communication script in the PIR scheme. By definition of entropy,  $H(\text{Comm}) \leq l$ . Denote  $\text{DB} = (y_1, \dots, y_k)$ . Let  $z_j = R(1^\lambda, \text{View}_B, j, \text{coin}_B)$  for  $j \in [k]$ . Let  $p_j = \Pr[y_j \neq z_j]$  and  $p = \frac{1}{k} \sum_{j \in [k]} p_j$ .

By Fano's inequality <sup>( $\Delta$ )</sup>:

$$H(p_j) \geq H(y_j | \text{Comm})$$

By chain rule<sup>(\*)</sup>:

$$H(\text{DB} | \text{Comm}) = \sum_{j=1}^k H(y_j | \text{Comm}, y_{j-1}, \dots, y_1) \leq \sum_{j=1}^k H(y_j | \text{Comm}).$$

Thus,

$$H(\text{DB}|\text{Comm}) = H(\text{DB}) - H(\text{Comm}) + H(\text{Comm}|\text{DB}) \geq k - H(\text{Comm}) \geq k - l,$$

Hence,

$$H(p) = H\left(\frac{1}{k} \sum_{j=1}^k p_j\right) \geq^{(\triangle)} \frac{1}{k} \sum_{j=1}^k H(p_j) \geq \frac{1}{k} \sum_{j=1}^k \frac{H(yg_j|\text{Comm})}{k} \geq^{(\star)} \frac{k-l}{k}$$

Where  $\triangle$  and  $\star$  denote where each inequality is applied. □

The full proof can also be seen in [?] and for the lemmas 4.5 (absolute difference in entropy between two probability distributions  $p$  and  $q$  on a finite set  $M$  is bounded above by the  $L_1$ -norm of their difference) and 4.6 (bounds the entropy of a function and is maximized by the uniform distribution) [?].

## Notes

### Conditional Entropy

$$H(Y|X) = - \sum_{x \in D_X, y \in D_Y} p(x, y) \log_2 \frac{p(x, y)}{p(x)}$$

### Entropy

Let  $X$  be a random variable taking values over a finite domain  $D_X$ .

$$H(X) = - \sum_{x \in D_X} p(x) \log_2 p(x)$$

Fact:  $0 \leq H(X) \leq \log_2(|D_X|)$

### Chain Rule

$$H(Y|X) = H(X) + H(Y|X)$$

$$H(X_1, \dots, X_n) = \sum_{i=1}^n H(X_i|X_1, \dots, X_{i-1})$$

### Bayes' Theorem

$$H(Y|X) = H(X|Y) - H(X) + H(Y)$$

### Binary Entropy Function

For  $p \in [0, 1]$ :

$$H(p) = -p \log_2(p) - (1-p) \log_2(1-p)$$

### Concavity of Entropy

For  $p, q \in [0, 1]$  and  $0 \leq \lambda \leq 1$ :

$$H(\lambda p + (1-\lambda)q) \geq \lambda H(p) + (1-\lambda)H(q)$$

### Fano's Inequality

Let  $X$  and  $Y$  be random variables with  $X \in D_X$  and  $Y \in D_Y$ . Let  $\hat{X} = f(Y)$  and  $p = P(X \neq \hat{X})$ .

$$H(X|Y) \leq H(p) + p \log_2(|D_X| - 1)$$