On Basing Cryptography on NP-hardness: The Case of Single-Server Private Information Retrieval

Tianren Liu[1⋆] and Vinod Vaikuntanathan[1⋆⋆]

MIT CSAIL

Abstract. The possibility of basing the security of cryptographic objects on the (minimal) assumption that $\mathbf{NP} \not\subseteq \mathbf{BPP}$ is at the very heart of complexity-theoretic cryptography. Unfortunately, most known results along these lines are negative, showing that, assuming widely believed complexity-theoretic conjectures, there are no reductions from an **NP**-hard problem to the task of breaking certain cryptographic schemes. For example, the work of Brassard (FOCS 1979) showed that one-way permutations cannot be based on **NP**-hardness; Akavia, Goldreich, Goldwasser and Moshkovitz (STOC 2006) and Bogdanov and Brzuska (TCC 2015) showed that size-verifiable one-way functions cannot be based on **NP**-hardness; and Bogdanov and Lee (CRYPTO 2013) showed that even simple homomorphic encryption schemes cannot be based on **NP**-hardness. We make progress along this line of inquiry by showing that single-server private information retrieval schemes cannot be based on **NP**-hardness, unless the polynomial hierarchy collapses. Our result is tight in terms of both the correctness and the privacy parameter of the PIR scheme.

1 Introduction

The possibility of basing the security of cryptographic objects on the (minimal) assumption that $\mathbf{NP} \not\subseteq \mathbf{BPP}$ is at the very heart of complexity-theoretic cryptography. To be somewhat more precise, "basing primitive X on **NP**-hardness" means that there is a construction of primitive X and a probabilistic polynomial-time oracle algorithm (a reduction) R such that for every oracle A that "breaks the security of X",

$$\Pr[R^A(\phi) = 1] \geq 2/3 \qquad \text{if } \phi \in \mathsf{SAT}$$
$$\Pr[R^A(\phi) = 1] \leq 1/3 \qquad \text{if } \phi \notin \mathsf{SAT}$$

⋆ liutr@mit.edu.

⋆⋆ vinodv@csail.mit.edu. Research supported in part by DARPA Grant number FA8750-11-2-0225, an Alfred P. Sloan Research Fellowship, the Northrop Grumman Cybersecurity Research Consortium (CRC), the Qatar Computing Research Institute, Microsoft Faculty Fellowship, and a Steven and Renee Finn Career Development Chair from MIT.

Most results on the problem of basing cryptography on **NP**-hardness have been negative. That is, the results show that certain cryptographic primitives cannot be based on **NP**-hardness, under widely believed complexity-theoretic conjectures (for example, that $\mathbf{NP} \neq \mathsf{coNP}$, or that the polynomial hierarchy does not collapse).

- Brassard [Bra79] showed that one-way permutations cannot be based on **NP**-hardness. Subsequently, Goldreich and Goldwasser [GG98], in the process of clarifying Brassard's work, showed that public-key encryption schemes that satisfy certain very special properties cannot be based on **NP**-hardness.
- Akavia, Goldreich, Goldwasser and Moshkovitz [AGGM06], and later Bogdanov and Brzuska [BB15], showed that a special class of one-way functions called *size-verifiable one-way functions* cannot be based on **NP**-hardness. Roughly speaking, a size-verifiable one-way function is one in which the size of the set of pre-images can be efficiently approximated via an **AM** protocol.
- Most recently, Bogdanov and Lee [BL13a] showed that (even simple) homomorphic encryption schemes cannot be based on **NP**-hardness. This includes additively homomorphic encryption as well as homomorphic encryption schemes that only support the majority function, as special cases.

Several works have also explored the problem of basing average-case hardness on (worst case) **NP**-hardness, via restricted types of reductions, most notably non-adaptive reductions that make all its queries to the oracle simultaneously. The work of Feigenbaum and Fortnow, subsequently strengthened by Bogdanov and Trevisan [BT06], show that there cannot be a *non-adaptive* reduction from (worst-case) SAT to the average-case hardness of any problem in **NP**, unless $\mathbf{PH} \subseteq \mathbf{\Sigma_2^p}$.

In a nutshell, while there are a handful of impossibility results for basing various types of cryptographic primitives on **NP**-hardness, our understanding of this question at this point is rather minuscule. In this work, we make progress along these lines of inquiry by showing that (single server) private information retrieval (PIR) schemes cannot be based on **NP**-hardness, unless the polynomial hierarchy collapses.

Main Theorem 1 (Informal) *If there is a probabilistic polynomial time reduction from solving* SAT *to breaking a single-server, one round, private information retrieval scheme, then* $\mathbf{PH} \subseteq \mathbf{\Sigma_2^p}$.

Our result is tight in terms of both the correctness and the privacy parameter of the PIR scheme. Namely, information-theoretically secure PIR schemes exist for the choice of these parameters that are not ruled out by our result. We refer the reader to Section 3 for a formal statement of our result.

Private information retrieval (PIR) is a protocol between a database D holding a string $x \in \{0,1\}^n$, and a user holding an index $i \in [n]$. The user wishes to retrieve the i-th bit x_i from the database, without revealing any information about i. Clearly, the database can rather inefficiently accomplish this by sending the entire string x to the user. The objective of PIR, then, is to achieve

this goal while communicating (significantly) less than n bits. Chor, Goldreich, Kushilevitz and Sudan [CKGS98], who first defined PIR, also showed that non-trivial PIR schemes (with communication less than n bits) require computational assumptions. Subsequently, PIR has been shown to imply one-way functions [BIKM99], oblivious transfer [CMO00] and collision-resistant hashing [IKO05], placing it in cryptomania proper.

On the other hand, there have been several constructions of PIR with decreasing communication complexity under various cryptographic assumptions [KO97,CMS99,Lip05,BGN05,GR05,Gen09,BV11], starting from the work of Kushilevitz and Ostrovsky [KO97] who showed a construction of PIR with $O(n^\epsilon)$ communication (for any constant $\epsilon > 0$) assuming the existence of additively homomorphic encryption schemes.

Notably, some of the later constructions of PIR [CMS99,Lip05,GR05] achieve better efficiency under number-theoretic assumptions such as the Phi-hiding assumption that are *not* known to imply (even additive) homomorphic encryption. In short, while additively homomorphic encryption gives us a PIR scheme, no implication in the other direction is known.

2 Definitions

2.1 Information Theory Notations

A *random variable* X over a finite set $\mathcal{S}$ is defined by its probability mass function $p_X : \mathcal{S} \to [0, +\infty)$ that $\sum_{x \in S} p_X(x) = 1$. We use uppercase letter to denote a random variable.

The *Shannon entropy* of a random variable X, denoted $H(X)$, is defined as $H(X) = \sum_x p_X(x) \log_2 \frac{1}{p_X(x)}$, which measures the uncertainty of X. Let $h(p) = H(\mathrm{Bern}(p)) = p \log_2 \frac{1}{p} + (1-p) \log_2 \frac{1}{1-p}$ denote the Shannon entropy of Bernoulli distribution $\mathrm{Bern}(p)$.

Assume random variables X, Y satisfies a joint distribution. The *conditional entropy* of random variable Y given X, denoted $H(Y|X)$, is defined as $H(Y|X) = H(XY) - H(X)$, which measures the uncertainty of Y given X is known.

The *mutual information* between random variables X and Y is $I(X;Y) = H(X) + H(Y) - H(XY)$, which measures the information shared by X and Y. In particular $H(Y) = I(X;Y) + H(Y|X)$.

Let $X \sim \mathrm{Bern}(\frac{1}{2})$ be a random variable uniformly distributed in $\{0,1\}$, let $N \sim \mathrm{Bern}(\varepsilon)$ be a noise that is independent from X, and let $\hat{X} = X \oplus N$ be the noisy version of X. Then $I(\hat{X}; X) = 1 - h(\varepsilon)$. Moreover, for any random variable X' satisfying $\Pr[X' = X] \geq 1 - \varepsilon$,

$$I(X'; X) = H(X) - H(X|X') = 1 - H(X \oplus X'|X') \geq 1 - H(X \oplus X') \geq 1 - h(\varepsilon)$$

So random variable $\hat{X} = X \oplus N$ minimize the mutual information $I(\hat{X}; X)$ under constraint $\Pr[\hat{X} = X] \geq 1 - \varepsilon$.

2.2 Single-server One-round Private Information Retrieval

In a single-server private information retrieval (PIR) protocol, the database holds n bits of data $x \in \{0,1\}^n$. The user, given an index $i \in [n]$, would like to retrieve the i-th bit from the server. In the meanwhile, the user want to prevent the server from learning any information about index i. The user does so by generating a query based on i using a randomized algorithm; the server responds to the query with an answer. The user, given the answer, should be able to learn the i-th bit x_i.

Definition 2.1 (Private information retrieval). *A single-server single-round private information retrieval (PIR) scheme is a tuple* $(\mathbf{Qry}, \mathbf{Ans}, \mathbf{Rec})$ *of algorithms such that*

- *Query algorithm* $\mathbf{Qry}$*: A probabilistic polynomial-time algorithm* $\mathbf{Qry}$ *that* $\mathbf{Qry}(1^n, i) \to (q, \sigma)$*, where* $i \in [n]$*.*
- *Answer algorithm* $\mathbf{Ans}$ *: A probabilistic polynomial-time algorithm* $\mathbf{Ans}$ *that* $\mathbf{Ans}(x, q) \to a$*, where* $x \in \{0,1\}^n$*. Let* ℓ *denote the length of the answer, i.e.* $a \in \{0,1\}^\ell$*.*
- *Reconstruction algorithm* $\mathbf{Rec}$*: A probabilistic polynomial-time algorithm* $\mathbf{Rec}$ *that* $\mathbf{Rec}(a, \sigma) \to b$ *s.t.* $b \in \{0,1\}$*.*

Correctness A PIR scheme $(\mathbf{Qry}, \mathbf{Ans}, \mathbf{Rec})$ is $(1-\varepsilon)$-correct if for any $x \in \{0,1\}^n$, for any i,

$$\Pr\Big[\mathbf{Qry}(1^n, i) \to (q, \sigma), \mathbf{Ans}(x, q) \to a : \mathbf{Rec}(a, \sigma) = x_i\Big] \geq 1 - \varepsilon(n)$$

where the randomness is over the random tapes of $\mathbf{Qry}, \mathbf{Ans}, \mathbf{Rec}$.

Privacy Computational privacy requires that the database cannot efficiently distinguish between queries for different indexes. Formally, a PIR scheme is δ-privacy (for some $\delta = \delta(n)$) if for any probabilistic polynomial-time algorithm $\mathcal{A}_1, \mathcal{A}_2$, there exists a negligible function δ such that

$$\Pr\left[\begin{array}{c} \mathcal{A}_1(1^n) \to (i_0, i_1, \tau) \\ b \xleftarrow{\$} \{0,1\} \\ \mathbf{Qry}(1^n, i_b) \to (q, \sigma) \\ \mathcal{A}_2(1^n, q, \tau) \to b' \end{array} : b' = b\right] < \frac{1}{2} + \delta(n) \tag{1}$$

The adversary in this privacy definition is interactive, which introduces difficulties in defining an oracle that breaks PIR. To make our task easier, we consider an alternative, non-interactive definition which is equivalent to (1).

Our non-interactive definition postulates that given $k = k(n)$, for any probabilistic polynomial-time algorithm $\mathcal{A}$, there exists a negligible function δ such that

$$\Pr\left[\begin{array}{c} j \xleftarrow{\$} [n] \\ \mathbf{Qry}(1^n, j) \to (q, \sigma) : j' = j \\ \mathcal{A}(1^n, q) \to j' \end{array}\right] < \frac{1}{n}\Big(1 + \delta(n)\Big) \tag{2}$$

These two definitions of privacy are equivalent up to a polynomial factor on δ.

Proposition 2.2. *If a PIR scheme is δ_1-private according to definition* (1)*, then it is δ_2-private according to Definition* (2)*, where $\delta_2 = n\delta_1$. Similarly, if a PIR scheme is δ_2-private according to definition* (2)*, then it is δ_1-private according to Definition* (2)*, where $\delta_1 = \delta_2/2$.*

Proof. Assume that a probabilistic polynomial-time adversary algorithm $\mathcal{A}$ disproves δ_2-private according to definition (2). Construct an adversary $(\mathcal{A}_1, \mathcal{A}_2)$ such that algorithm $\mathcal{A}_1(1^n)$ pick random indexes i_0, i_1 and output $i_0, i_1, \tau = (i_0, i_1)$, algorithm $\mathcal{A}_2(1^n, q, \tau = (i_0, i_1))$ call $\mathcal{A}(1^n, q) \to i$ and output 0 if and only if $i = i_0$. Then $(\mathcal{A}_1, \mathcal{A}_2)$ disprove $\frac{\delta_2}{n}$-private according to definition (1) as

$$\Pr\left[\begin{array}{c}\mathcal{A}_1(1^n) \to (i_0, i_1, \tau)\\ b \xleftarrow{\$} \{0,1\}\\ \mathbf{Qry}(1^n, i_b) \to (q, \sigma)\\ \mathcal{A}_2(1^n, q, \tau) \to b'\end{array} : b' = b\right] = \Pr\left[\begin{array}{c}i_0, i_1 \xleftarrow{\$} [n]\\ b \xleftarrow{\$} \{0,1\}\\ \mathbf{Qry}(1^n, i_b) \to (q, \sigma)\\ \mathcal{A}(1^n, q) \to i\end{array} : \begin{array}{c} i = i_0, b = 0\\ \text{or}\\ i \neq i_0, b \neq 0\end{array}\right]$$

$$= \frac{1}{2}\Pr\left[\begin{array}{c}i_0, i_1 \xleftarrow{\$} [n]\\ \mathbf{Qry}(1^n, i_0) \to (q, \sigma) : i = i_0\\ \mathcal{A}(1^n, q) \to i\end{array}\right] + \frac{1}{2}\Pr\left[\begin{array}{c}i_0, i_1 \xleftarrow{\$} [n]\\ \mathbf{Qry}(1^n, i_1) \to (q, \sigma) : i \neq i_0\\ \mathcal{A}(1^n, q) \to i\end{array}\right]$$

$$\geq \frac{1}{2}\frac{1}{n}\big(1 + \delta_2(n)\big) + \frac{1}{2}\Big(1 - \frac{1}{n}\Big) = \frac{1}{2}\Big(1 + \frac{1}{n}\delta_2(n)\Big)$$

Assume probabilistic polynomial-time adversary algorithm $(\mathcal{A}_1, \mathcal{A}_2)$ disprove δ_1-private according to definition (1), construct an adversary $\mathcal{A}$ that call $(i_0, i_1, \tau) \leftarrow \mathcal{A}_1(1^n)$, $b \leftarrow \mathcal{A}_2(1^n, q, \tau)$ and output i_b. Then $\mathcal{A}$ disprove $2\delta_1$-private according to definition (2) as

$$\Pr\left[\begin{array}{c}j \xleftarrow{\$} [n]\\ \mathbf{Qry}(1^n, j) \to (q, \sigma) : j' = j\\ \mathcal{A}(1^n, q) \to j'\end{array}\right] = \Pr\left[\begin{array}{c}\mathcal{A}_1(1^n) \to (i_0, i_1, \tau)\\ j \xleftarrow{\$} [n]\\ \mathbf{Qry}(1^n, j) \to (q, \sigma)\\ \mathcal{A}_2(1^n, q, \tau) \to b\end{array} : j = i_b\right]$$

$$= \frac{2}{n}\Pr\left[\begin{array}{c}\mathcal{A}_1(1^n) \to (i_0, i_1, \tau)\\ j \xleftarrow{\$} \{i_0, i_1\}\\ \mathbf{Qry}(1^n, j) \to (q, \sigma)\\ \mathcal{A}_2(1^n, q, \tau) \to b\end{array} : j = i_b\right] \geq \frac{2}{n}\Big(\frac{1}{2} + \delta_1(n)\Big) = \frac{1}{n}\Big(1 + 2\delta_1(n)\Big)$$

We consider a weaker privacy definition: for any probabilistic polynomial-time algorithm $\mathcal{A}$, there exists a negligible function δ such that

$$\Pr\left[\begin{array}{c}j \xleftarrow{\$} [n]\\ \mathbf{Qry}(1^n, j) \to (q, \sigma) :\\ \mathcal{A}(1^n, q) \to \mathcal{J}\end{array} \begin{array}{c} j \in \mathcal{J}\\ |\mathcal{J}| \leq k\end{array}\right] < \frac{k}{n}\Big(1 + \delta(n)\Big) = \alpha(n) \tag{3}$$

Define a PIR scheme $(\mathbf{Qry}, \mathbf{Ans}, \mathbf{Rec})$ to be (p, k, α)-private if (3) holds for any adversary $\mathcal{A}$ of running time $p(n)$. A PIR scheme is said to be (k, α)-private if (3) holds for any (computationally unbounded) adversary $\mathcal{A}$.

Proposition 2.3. *For $k' \leq k$, (p, k', α)-privacy implies $(p, k, \frac{k}{k'}\alpha)$-privacy*

Proof. Assume there exists an adversary $\mathcal{A}$ disprove the $(p, k, \frac{k}{k'}\alpha)$-privacy of the PIR scheme; construct an adversary $\mathcal{A}'$ that first executes $\mathcal{J} \leftarrow \mathcal{A}(1^n, q)$, then outputs a random size k' subset of $\mathcal{J}$; then $\mathcal{A}'$ disprove the (p, k', α)-privacy of the PIR scheme.

Answer Communication Complexity In Definition 2.1, ℓ is defined as the length of server's answer. We say ℓ is the *answer communication complexity*. Similarly, the length of the query can be named as the query communication complexity, and their sum is the communication complexity. In this work, we are only interested the answer communication complexity, which is an upper bound of the information flowing from the server to the user.

Typically, people are only interested in PIR scheme such that answer communication complexity $\ell = o(n)$. Otherwise, e.g. when $\ell = n$, there exists a trivial PIR protocol with perfect privacy, where the user sends a meaningless query and the server sends the whole database x. The following proposition shows that the PIR scheme can trivially achieve good privacy if the answer communication complexity ℓ is large.

Proposition 2.4. *There exists a PIR scheme satisfying perfect information-theoretical privacy with answer communication complexity $\ell = n \cdot (1 - h(\varepsilon) + O(n^{-1/4}))$. Similarly, for any $\alpha \geq k/n$, there exists a PIR scheme satisfying (k, α)-privacy with answer communication complexity $\ell = \frac{k}{\alpha}(1 - h(\varepsilon) + O((\frac{k}{\alpha})^{-1/4}))$.*

Proof. Consider a PIR scheme such that the query contains no information about the index i, therefore perfect privacy is achieved. If the server sends the whole database to the user, n bits answer communication complexity is necessary. While the server could somehow compress the database into an ℓ-bit answer, such that the user could still recover the database with at most ε fraction of distortion.

Such task is called "lossy source coding" [Sha59], which is one of the fundamental problems of information theory. Let X be a uniform random Bernoulli variable, then $\Pr[\hat{X} = X] \geq 1 - \varepsilon$ implies $I(\hat{X}, X) \geq 1 - h(\varepsilon)$. Therefore, if you want to compress a random binary string, and are able to recover the string from the lossy compression with $(1 - \varepsilon)$ accuracy, then the compression ratio need to be at least $1 - h(\varepsilon)$.

There exists a lossy source coding scheme almost achieves the information theoretical bound [Ari09,KU10], i.e. when $\ell = n \cdot (1 - h(\varepsilon) + O(n^{-1/4}))$, there exists efficient algorithms $E : \{0,1\}^n \to \{0,1\}^\ell$ and $D : \{0,1\}^\ell \to \{0,1\}^n$, such that for randomly chosen $X \in \{0,1\}^n$ and for any index $i \in [n]$,

$$\Pr_X[\hat{X} = D(E(X)) : \hat{X}_i = X_i] \geq 1 - \varepsilon.$$

Therefore, if the server sends $E(x)$ as the answer, then the PIR scheme could achieve $(1-\varepsilon)$ correctness on random database. Moreover, we could consider the following scheme,

- User sends a query m, which is a random string in $\{0,1\}^n$;
- Server answers by $a = E(m \oplus x)$;
- User retrieve the whole database by $\hat{x} = D(a) \oplus m$.

Then for any index $i \in [n]$, $\Pr[\hat{x}_i = x_i] \geq 1 - \varepsilon$.

Similarly, if $\ell = \frac{k}{\alpha}(1 - h(\varepsilon) + O((\frac{k}{\alpha})^{-1/4}))$, the user could pick a random subset $\mathcal{J} \ni i$ of $\frac{k}{\alpha}$ indexes, the server uses lossy source code to encode $\{x_j\}_{j\in\mathcal{J}}$ in ℓ bits. Based on the query, we know nothing about the index i besides $i \in \mathcal{J}$, so (k, α)-privacy is achieved.

Reduction to breaking PIR. What does it mean for a reduction to decide a language L assuming that there is a p.p.t. adversary that breaks PIR? For any language L, we say L can be reduced to breaking the (k, α)-privacy of PIR scheme $(\mathbf{Qry}, \mathbf{Ans}, \mathbf{Rec})$ if there exists a probabilistic polynomial-time oracle Turing machine (OTM) M such that

$$\begin{aligned} \Pr[M^{\mathcal{O}}(x) = 1] &\geq 2/3 \quad \text{if } x \in L \\ \Pr[M^{\mathcal{O}}(x) = 1] &\leq 1/3 \quad \text{if } x \notin L \end{aligned}$$

for all x, for all oracle $\mathcal{O}$ such that

$$\Pr\left[\begin{array}{c} j \leftarrow [n] \\ \mathbf{Qry}(1^n, j) \to (q, \sigma) : \\ \mathcal{O}(q) \to \mathcal{J} \end{array} \begin{array}{c} \\ j \in \mathcal{J} \\ |\mathcal{J}| \leq k \end{array}\right] \geq \alpha \tag{4}$$

2.3 Entropy Difference Oracle

Entropy Difference (ED) is a complete problem of SZK [GV99]. Entropy Difference is a promise problem defined as

- YES instances: (X, Y) such that $H(X) \geq H(Y) + 1$
- NO instances: (X, Y) such that $H(Y) \geq H(X) + 1$

where X and Y are distributions encoded as circuits which sample from them.

Given an entropy difference oracle, a polynomial-time algorithm would be able the differ the entropy of two circuits up to any inverse-polynomial precision. For example, for distributions X, Y, we query the Entropy Difference oracle with $(X_1 \dots X_s, Y_1 \dots Y_s)$, where $X_i \sim X, Y_i \sim Y$ and $X_1, \dots, X_s$ are i.i.d. and $Y_1, \dots, Y_s$ are i.i.d. Then we would be able to distinguish between $H(X) \geq H(Y) + \frac{1}{s}$ and $H(Y) \geq H(X) + \frac{1}{s}$.

Similarly, polynomial-time algorithm could use Entropy Difference oracle to distinguish between $H(X) \geq \hat{h} + \frac{1}{s}$ and $H(X) \geq \hat{h} + \frac{1}{s}$. Construct a distribution

Y that $2s\hat{h} - 1 < H(Y) < 2s\hat{h} + 1$ and query Entropy Difference oracle with $(X_1 \dots X_{2s}, Y)$, where $X_1, \dots, X_{2s}$ are independent copies of X.

Therefore, a polynomial-time algorithm given Entropy Difference oracle could estimate $H(X)$ to any inverse-polynomial precision by binary search. Assume X, Y satisfies a joint distribution encoded as a circuit which samples from it, then a polynomial-time algorithm given Entropy Difference oracle could estimate conditional entropy $H(X|Y)$, mutual information $I(X;Y)$ to any inverse-polynomial precision. Here the precision is measured by absolute error.

3 PIR and NP-hardness

Theorem 3.1 (Main Theorem). *For any $(1-\epsilon)$-correct PIR scheme $\Pi = (\mathbf{Qry}, \mathbf{Ans}, \mathbf{Rec})$ with n-bit databases and answer communication complexity ℓ, for any language L, if:*

1. *there exists a reduction from L to breaking the PIR δ-privacy (in the sense of Equation (2)); and*
2. *there is a polynomial $p(n)$ such that*

$$\ell \cdot (1+\delta) \leq n \cdot (1 - h(\varepsilon)) - 1/p(n)$$

then $L \in \mathsf{AM} \cap \mathsf{coAM}$.

We prove our main theorem by combining the following two lemmas. The first lemma is our main ingredient, and says that if there is a reduction from breaking a PIR scheme to deciding a language L, and the PIR scheme has a low answer communication complexity, then L can be reduced to the entropy difference problem (see Section 2.3). For the proof of this lemma, we find it somewhat convenient to go through the intermediate and weaker notion of (k, α)-privacy (see Equation (4)), which makes the lemma statement somewhat stronger. Indeed, as Proposition 2.4 shows, the bound in the lemma is tight, since there is in fact an information-theoretic PIR protocol with a matching answer communication complexity.

Lemma 3.2 ($\mathsf{BPP}^{(k,\alpha)\text{-breaking PIR}} \subseteq \mathsf{BPP}^{\mathsf{ED}}$). *For any $(1-\epsilon)$-correct PIR scheme, for any language L, if there exists a reduction from L to (k, α)-breaking the PIR privacy such that*

- *$k \leq \ell$; and*
- *there is a polynomial $p(n)$ such that*

$$\alpha + \frac{1}{p(n)} \leq \frac{k(1 - h(\varepsilon))}{\ell}$$

then there exists a reduction from L to ED.

As noted above, this condition is quite tight when $k \leq \ell$, as there exists a PIR scheme achieving (k, α)-privacy if $\ell \approx \frac{k(1-h(\varepsilon))}{\alpha}$.

Our next lemma shows that any language decidable by a randomized oracle machine with access to an entropy difference oracle, is in $\mathsf{AM} \cap \mathsf{coAM}$. Similar statements have been shown in several previous works (e.g.,[MX10,BL13b]), and we include a proof here for completeness.

Lemma 3.3 ($\mathsf{BPP}^{\mathsf{ED}} \subseteq \mathsf{AM} \cap \mathsf{coAM}$). *For any language L, if there exists an OTM M such that for any oracle $\mathcal{O}$ solving entropy difference*

$$\begin{aligned} \Pr[M^{\mathcal{O}}(x) = 1] &\geq 2/3 \quad \text{if } x \in L \\ \Pr[M^{\mathcal{O}}(x) = 1] &\leq 1/3 \quad \text{if } x \notin L, \end{aligned}$$

then $L \in \mathsf{AM} \cap \mathsf{coAM}$.

3.1 Proof of the Main Theorem

Assume there exists a reduction from L to breaking PIR with parameter as stated in Theorem 3.1. This is a reduction from L to $(1, \alpha)$-breaking PIR where

$$\alpha = \frac{1}{n}(1+\delta) \leq \frac{1-h(\varepsilon)}{\ell} - \frac{n}{\ell \cdot p(n)}.$$

Then by Lemma 3.2, there exists a reduction from L to ED. Combining Lemma 3.3, we can deduce that $L \in \mathsf{AM} \cap \mathsf{coAM}$.

3.2 Proof of the Lemma 3.2

Consider a PIR scheme $(\mathbf{Qry}, \mathbf{Ans}, \mathbf{Rec})$ which is $(1-\varepsilon)$-correct. For any $x \in \{0,1\}^n$, for any index i,

$$\Pr\Big[\mathbf{Qry}(1^n, i) \to (q, \sigma), \mathbf{Ans}(x, q) \to a : \mathbf{Rec}(a, \sigma) = x_i\Big] \geq 1 - \varepsilon(n),$$

the randomness is over the random tapes of $\mathbf{Qry}, \mathbf{Ans}, \mathbf{Rec}$.

Consider a random database, that is, the data X is a uniform random variable over $\{0,1\}^n$. For any fixed index i, the pair $(Q, \Sigma) \leftarrow \mathbf{Qry}(1^n, i)$ satisfies a joint distribution independent from X. Similarly, define random variable $A \leftarrow \mathbf{Ans}(X, Q), \hat{X}_i \leftarrow \mathbf{Rec}(A, \Sigma)$.

As the PIR scheme is $(1-\varepsilon)$-correct, $\Pr[\hat{X}_i = X_i] \geq 1-\varepsilon$. As X_i is a uniform Bernoulli, $I(\hat{X}_i; X_i) \geq 1 - h(\varepsilon)$. As Σ is independent from X_i,

$$I(A; X_i) \geq I(\hat{X}_i; X_i) \geq 1 - h(\varepsilon).$$

As Q is independent from X_i,

$$I(A; X_i|Q) = H(X_i|Q) - H(X_i|AQ) \geq H(X_i) - H(X_i|A) = I(A; X_i) \geq 1 - h(\varepsilon). \tag{5}$$

Remind that, by definition,

$$I(A;X_i|Q) = \mathop{\mathbb{E}}_{Q}\Big[I(A;X_i|Q=Q)\Big] = \sum_q I(A;X_i|Q=q)\Pr[Q=q]$$

For any potential query q, the event $Q=q$ is independent from X. In particular, for any index j, random variable X_j is independent from $X_1 \dots X_{j-1}$ given $Q=q$. So for any q,

$$\begin{aligned}\sum_{j=1}^{n} I(A;X_j|Q=q) &\le \sum_{j=1}^{n} I(A;X_j|X_1\dots X_{j-1},Q=q)\\ &= I(A;X_1\dots X_n|Q=q) \le H(A|Q=q) \le \ell \quad (6)\end{aligned}$$

The first inequality is implied by the fact that $I(Z;X|Y)-I(Z;X) = I(X;Y|Z)-I(X;Y)$ for arbitrary random variables X,Y,Z, and is non-negative if X,Y are independent. So

$$I(A;X_j|X_1\dots X_{j-1},Q=q)-I(A;X_j|Q=q) = I(X_1\dots X_{j-1},X_j|A,Q=q) \ge 0.$$

Equations (5) and (6) are the core in the proof of Lemma 3.2. Equation (5) shows that, when retrieve the i-th bit, the mutual information between X_i and server's answer A is large. Equation (6) shows that, the sum of mutual information between each bit X_j and server's answer A is bounded. Therefore, if we could measure the mutual information by an Entropy Difference oracle, we would have a good knowledge of i.

$(\mathbf{Qry},\mathbf{Ans},\mathbf{Rec})$ is an $(1-\varepsilon)$-correct PIR scheme. Assume language L can be solved by a probabilistic polynomial-time oracle Turing machine $\mathcal{M}$ given any oracle (k,α)-breaking PIR scheme $(\mathbf{Qry},\mathbf{Ans},\mathbf{Rec})$ such that k,α satisfies $k \le \ell$ and

$$\alpha + \frac{1}{p(n)} \le \frac{k(1-h(\varepsilon))}{\ell} \quad (7)$$

where $p(\cdot)$ is a fixed polynomial. We construct an efficient oracle algorithm solving L given an Entropy Difference oracle (Algorithm 1).

Whenever $\mathcal{O}(q)$ is simulated,

$$|\mathcal{J}| = \Big\lceil \frac{s_n-\rho}{B} \Big\rceil \le \frac{s_n}{B} = \frac{1}{B}\sum_{j=1}^{n}\hat{\mu}_j \le \frac{1}{B}\sum_{j=1}^{n}\Big(\mu_j + \frac{1}{2n\cdot p(n)}\Big) \le \frac{1}{B}\Big(\ell + \frac{1}{2\cdot p(n)}\Big) = k$$

For any query q and index i, when $\mathcal{O}(q)$ is simulated,

$$\Pr\big[\mathcal{J} \leftarrow \mathcal{O}(q) : i \in \mathcal{J}\big] = \frac{\hat{\mu}_i}{B} \ge \frac{\mu_i - \frac{1}{2n\cdot p(n)}}{B}$$

Algorithm 1 Solving L given ED oracle on input x

1. Simulate $\mathcal{M}^O(x)$
2. Whenever $\mathcal{M}$ query $\mathcal{O}(q)$
   (a) For each index $j = 1, \ldots, n$, use entropy difference oracle to estimate
   $$\mu_j = I(A; X_j | Q = q)$$
   to $\frac{1}{2n \cdot p(n)}$ precision. More precisely, construct circuit C that $C(x, r) \mapsto (x_j, \mathbf{Ans}(x, q, r))$, and estimate the mutual information between the two components of C's output. Let $\hat{\mu}_j \in [0, 1]$ denote the estimation.
   (b) Define $s_k = \sum_{j=1}^{k} \hat{\mu}_j$, $B = (\ell + \frac{1}{2p(n)})/k$
   (c) Choose a random value $\rho \in [0, B)$
   (d) Let $\mathcal{J} = \{j | \exists z \in \mathbb{Z}, s_{j-1} \leq zB + \rho < s_j\}$
   (e) Answer $\mathcal{M}$'s query by $\mathcal{J}$
3. Output what $\mathcal{M}$ output

Assuming q is generated from $q \leftarrow \mathbf{Qry}(1^n, i)$, then $\mathbb{E}[\mu_i] \geq 1 - h(\varepsilon)$. So

$$\begin{aligned}
&\Pr\big[q \leftarrow \mathbf{Qry}(1^n, i), \mathcal{J} \leftarrow \mathcal{O}(q) : i \in \mathcal{J}\big] \\
&\quad = \mathop{\mathbb{E}}_{q \leftarrow \mathbf{Qry}(1^n, i)} \big[\Pr[i \in J | Q = q]\big] \geq \mathop{\mathbb{E}}_{q \leftarrow \mathbf{Qry}(1^n, i)} \Big[\frac{\mu_i - \frac{1}{2n \cdot p(n)}}{B}\Big] \\
&\quad = \frac{\mathbb{E}_{q \leftarrow \mathbf{Qry}(1^n, i)}[\mu_i] - \frac{1}{2n \cdot p(n)}}{B} \geq \frac{1 - h(\varepsilon) - \frac{1}{2n \cdot p(n)}}{(\ell + \frac{1}{2p(n)})/k} \geq \alpha \qquad (8)
\end{aligned}$$

The last inequality is a consequence of constraint (7),

$$\alpha\Big(1 + \frac{1}{2p(n)}\Big) \leq \alpha + \frac{1}{2p(n)} \leq \frac{1 - h(\varepsilon)}{\ell/k} - \frac{1}{2n \cdot p(n)} \leq \frac{1 - h(\varepsilon) - \frac{1}{2n \cdot p(n)}}{\ell/k}$$

$$\alpha \leq \frac{1 - h(\varepsilon) - \frac{1}{2n \cdot p(n)}}{\ell(1 + \frac{1}{2p(n)})/k} \leq \frac{1 - h(\varepsilon) - \frac{1}{2n \cdot p(n)}}{(\ell + \frac{1}{2p(n)})/k}$$

4 Conclusions

We show that for any non-trivial single-server single-round PIR scheme $\Pi = (\mathbf{Qry}, \mathbf{Ans}, \mathbf{Rec})$, its privacy can be broken in SZK. Therefore, assume there is an polynomial time, adaptive reduction from L to breaking the privacy of Π, then $L \in \mathsf{AM} \cap \mathsf{coAM}$. Such results rule out the possibility that there is a reduction from NP-complete problems to breaking single-server single-round PIR.

We observes that the ability of "rerandomization" is critical in our proof. Let Π be a single-server single-round PIR scheme. The adversary is given the transcript of an execution where the user wants to query the i-th bit. Based on

that, the adversary could generate a distribution over the transcripts where the user query that same index and the database is uniformly random. The ability to generate a transcript distribution of the same index and random database allows the adversary to break PIR scheme Π with a SZK oracle. In particular, the circuit generating the transcript distribution will be fed to SZK oracle as input.

Similar technique shows that breaking homomorphic encryption is not NP-hard [BL13b]. Assume an encryption scheme Π allows rerandomization, i.e. given a ciphertext, you are able to generate a ciphertext distribution of the same message that is statically close to a fresh encryption. Then adversary could break Π with a SZK oracle. ElGamal encryption is an example that allows rerandomization. [BL13b] shows that homomorphic encryption implies rerandomization, i.e. if a encryption scheme supports homomorphic evaluation (and the homomorphic evaluation produces a ciphertext that is statically close to a fresh encryption), then the encryption scheme also allows rerandomization.

We have explored whether same technique applies to single-server multiple-round PIR. In the multiple-round case, assume you are given several transcripts where the user queries the i-bit, and index i is unknown. It's not clear at all how to generate a valid transcript other than the given ones where the user queries the same index.

One open problem is to show there exists no reduction from NP-hard problems to breaking single-server multiple-round PIR. As multiple-round introduces great difficulty, it's worth exploring similar statements for any multiple-round cryptographic primitives.

Acknowledgments: We would like to thank Jayadev Acharya for valuable comments about lossy source coding and polar codes.

References

[AGGM06] Adi Akavia, Oded Goldreich, Shafi Goldwasser, and Dana Moshkovitz. On basing one-way functions on np-hardness. In Jon M. Kleinberg, editor, *Proceedings of the 38th Annual ACM Symposium on Theory of Computing, Seattle, WA, USA, May 21-23, 2006*, pages 701–710. ACM, 2006.

[Ari09] Erdal Arikan. Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels. *Information Theory, IEEE Transactions on*, 55(7):3051–3073, 2009.

[BB15] Andrej Bogdanov and Christina Brzuska. On basing size-verifiable one-way functions on np-hardness. In Yevgeniy Dodis and Jesper Buus Nielsen, editors, *Theory of Cryptography - 12th Theory of Cryptography Conference, TCC 2015, Warsaw, Poland, March 23-25, 2015, Proceedings, Part I*, volume 9014 of *Lecture Notes in Computer Science*, pages 1–6. Springer, 2015.

[BGN05] Dan Boneh, Eu-Jin Goh, and Kobbi Nissim. Evaluating 2-dnf formulas on ciphertexts. In Kilian [Kil05], pages 325–341.

[BIKM99] Amos Beimel, Yuval Ishai, Eyal Kushilevitz, and Tal Malkin. One-way functions are essential for single-server private information retrieval. In Jeffrey Scott Vitter, Lawrence L. Larmore, and Frank Thomson Leighton,

editors, *Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing, May 1-4, 1999, Atlanta, Georgia, USA*, pages 89–98. ACM, 1999.

[BL13a] Andrej Bogdanov and Chin Ho Lee. Limits of provable security for homomorphic encryption. In Ran Canetti and Juan A. Garay, editors, *Advances in Cryptology - CRYPTO 2013 - 33rd Annual Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2013. Proceedings, Part I*, volume 8042 of *Lecture Notes in Computer Science*, pages 111–128. Springer, 2013.

[BL13b] Andrej Bogdanov and Chin Ho Lee. Limits of provable security for homomorphic encryption. In *Advances in Cryptology–CRYPTO 2013*, pages 111–128. Springer, 2013.

[Bra79] Gilles Brassard. Relativized cryptography. In *20th Annual Symposium on Foundations of Computer Science, San Juan, Puerto Rico, 29-31 October 1979*, pages 383–391. IEEE Computer Society, 1979.

[BT06] Andrej Bogdanov and Luca Trevisan. On worst-case to average-case reductions for NP problems. *SIAM J. Comput.*, 36(4):1119–1159, 2006.

[BV11] Zvika Brakerski and Vinod Vaikuntanathan. Efficient fully homomorphic encryption from (standard) LWE. In Rafail Ostrovsky, editor, *IEEE 52nd Annual Symposium on Foundations of Computer Science, FOCS 2011, Palm Springs, CA, USA, October 22-25, 2011*, pages 97–106. IEEE Computer Society, 2011.

[CKGS98] Benny Chor, Eyal Kushilevitz, Oded Goldreich, and Madhu Sudan. Private information retrieval. *J. ACM*, 45(6):965–981, 1998.

[CMO00] Giovanni Di Crescenzo, Tal Malkin, and Rafail Ostrovsky. Single database private information retrieval implies oblivious transfer. In Bart Preneel, editor, *Advances in Cryptology - EUROCRYPT 2000, International Conference on the Theory and Application of Cryptographic Techniques, Bruges, Belgium, May 14-18, 2000, Proceeding*, volume 1807 of *Lecture Notes in Computer Science*, pages 122–138. Springer, 2000.

[CMS99] Christian Cachin, Silvio Micali, and Markus Stadler. Computationally private information retrieval with polylogarithmic communication. In Jacques Stern, editor, *Advances in Cryptology - EUROCRYPT '99, International Conference on the Theory and Application of Cryptographic Techniques, Prague, Czech Republic, May 2-6, 1999, Proceeding*, volume 1592 of *Lecture Notes in Computer Science*, pages 402–414. Springer, 1999.

[Gen09] Craig Gentry. Fully homomorphic encryption using ideal lattices. In Michael Mitzenmacher, editor, *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, Bethesda, MD, USA, May 31 - June 2, 2009*, pages 169–178. ACM, 2009.

[GG98] Oded Goldreich and Shafi Goldwasser. On the possibility of basing cryptography on the assumption that $p \neq np$. *IACR Cryptology ePrint Archive*, 1998:5, 1998.

[GR05] Craig Gentry and Zulfikar Ramzan. Single-database private information retrieval with constant communication rate. In Luís Caires, Giuseppe F. Italiano, Luís Monteiro, Catuscia Palamidessi, and Moti Yung, editors, *Automata, Languages and Programming, 32nd International Colloquium, ICALP 2005, Lisbon, Portugal, July 11-15, 2005, Proceedings*, volume 3580 of *Lecture Notes in Computer Science*, pages 803–815. Springer, 2005.

[GV99] Oded Goldreich and Salil Vadhan. Comparing entropies in statistical zero knowledge with applications to the structure of szk. In *Computational*

Complexity, 1999. Proceedings. Fourteenth Annual IEEE Conference on, pages 54–73. IEEE, 1999.

[IKO05] Yuval Ishai, Eyal Kushilevitz, and Rafail Ostrovsky. Sufficient conditions for collision-resistant hashing. In Kilian [Kil05], pages 445–456.

[Kil05] Joe Kilian, editor. *Theory of Cryptography, Second Theory of Cryptography Conference, TCC 2005, Cambridge, MA, USA, February 10-12, 2005, Proceedings*, volume 3378 of *Lecture Notes in Computer Science*. Springer, 2005.

[KO97] Eyal Kushilevitz and Rafail Ostrovsky. Replication is NOT needed: SINGLE database, computationally-private information retrieval. In *38th Annual Symposium on Foundations of Computer Science, FOCS '97, Miami Beach, Florida, USA, October 19-22, 1997*, pages 364–373. IEEE Computer Society, 1997.

[KU10] Satish Babu Korada and Rüdiger L Urbanke. Polar codes are optimal for lossy source coding. *Information Theory, IEEE Transactions on*, 56(4):1751–1768, 2010.

[Lip05] Helger Lipmaa. An oblivious transfer protocol with log-squared communication. In Jianying Zhou, Javier Lopez, Robert H. Deng, and Feng Bao, editors, *Information Security, 8th International Conference, ISC 2005, Singapore, September 20-23, 2005, Proceedings*, volume 3650 of *Lecture Notes in Computer Science*, pages 314–328. Springer, 2005.

[MX10] Mohammad Mahmoody and David Xiao. On the power of randomized reductions and the checkability of sat. In *Computational Complexity (CCC), 2010 IEEE 25th Annual Conference on*, pages 64–75. IEEE, 2010.

[Sha59] Claude E Shannon. Coding theorems for a discrete source with a fidelity criterion. *IRE Nat. Conv. Rec*, 4(142-163):1, 1959.

A $\mathsf{BPP}^{\mathsf{ED}} \subseteq \mathsf{AM} \cap \mathsf{coAM}$

In this section we'll prove Lemma 3.3. Lemma 3.3 is a direct corollary of $\mathsf{BPP}^{\mathsf{SZK}} \subseteq \mathsf{AM} \cap \mathsf{coAM}$ [MX10]. We include the proof to be self-contained.

Let $\mathcal{M}$ be an probabilistic polynomial-time oracle Turing machine that solves promise problem $L = (L_Y, L_N)$ with bounded error

$$\begin{aligned} x \in L_Y &\implies \Pr\left[\mathcal{M}^{\mathcal{O}}(x) = 1\right] \geq \frac{8}{9} \\ x \in L_N &\implies \Pr\left[\mathcal{M}^{\mathcal{O}}(x) = 1\right] \leq \frac{1}{9} \end{aligned} \tag{9}$$

given any Entropy Difference oracle $\mathcal{O}$. It's sufficient to prove $L \in \mathsf{AM}$. By applying the same argument to $\overline{L} = (L_N, L_Y)$ we also get $L \in \mathsf{coAM}$.

Let $T(n)$ be a polynomial upper bound the running time of $\mathcal{M}$. W.l.o.g. assume $\mathcal{M}$ never make same query twice.

We'll construct an AM protocol solving L by simulating the execution of $\mathcal{M}^{\mathcal{O}}(x)$. First we should specify the oracle. When simulate execution $\mathcal{M}^{\mathcal{O}}(x)$ that $|x| = n$, consider the following randomized ED oracle $\mathcal{D}$.

When no duplicated query is make, $\mathcal{D}$ is equivalent to $\mathcal{O}_f$ when $f : \{0,1\}^* \to \{1, \ldots, 9 \cdot T(n)\}$ is a randomly picked function. As (9) is satisfied when $\mathcal{O} =$

Entropy Difference oracle $\mathcal{D}$ on input (X, Y)

1. Choose a random integer $k \in \{1, \ldots, 9 \cdot T(n)\}$
2. Accept if and only if $H(X) > H(Y) - 1 + \frac{2k-1}{9 \cdot T(n)}$

Entropy Difference oracle $\mathcal{O}_f$ on input (X, Y)

1. Let $k = f(X, Y)$
2. Accept if and only if $H(X) > H(Y) - 1 + \frac{2k-1}{9 \cdot T(n)}$

$\mathcal{O}_f$ for any f, condition (9) is also satisfied when $\mathcal{O} = \mathcal{D}$, in which case the randomness is over the random tape of $\mathcal{M}$ and $\mathcal{D}$.

We construct a decision protocol SP that simulate $\mathcal{M}^{\mathcal{D}}(x)$ to solve L in AM.

Decision protocol SP **solving** $L = (L_Y, L_N)$ on input $x \in \{0,1\}^n$

V: Sample and send random tape $r \in \{0,1\}^{T(n)}$ and random integer $k_1, \ldots, k_{T(n)} \in \{1, \ldots, 9 \cdot T(n)\}^{T(n)}$

P: Send query-answer sequence $((X_1, Y_1), a_1, (X_2, Y_2), a_2, \ldots, (X_q, Y_q), a_q)$ where $a_i \in \{0,1\}$

V: Check query-answer sequence satisfies

1. For every $i \in \{1, \ldots, q\}$, if $\mathcal{M}^{\mathcal{O}}(x)$ is run with random tape r and the oracle answer the first $i-1$ queries by $a_1, \ldots, a_{i-1}$, then the next query is (X_i, Y_i)
2. If $\mathcal{M}^{\mathcal{O}}(x)$ is run with random tape r and the oracle answer the q queries by $a_1, \ldots, a_q$, then $\mathcal{M}$ should accept x

Reject if not

P,V: For each $i \in \{1, \ldots, q\}$, using SZK protocol to distinguish $H(X_i) - H(Y_i) \geq \frac{2k_i}{9 \cdot T(n)} - 1$ and $H(X_i) - H(Y_i) \leq \frac{2(k_i - 1)}{9 \cdot T(n)} - 1$. More precisely, use Entropy Difference protocol for

$$\Big(\underbrace{X_1^{(i)} \ldots X_{9 \cdot T(n)}^{(i)}}_{\text{i.i.d. duplicates of } X_i} U_{9 \cdot T(n)}, \underbrace{Y_1^{(i)} \ldots Y_{9 \cdot T(n)}^{(i)}}_{\text{i.i.d. duplicates of } Y_i} U_{2k-1} \Big)$$

if $a_i = 1$, for its inverse if $a_i = 0$, where U_m is uniform distribution over $\{0,1\}^m$. Repeat $O(n)$ times in parallel and take majority so that the error probability at most $\frac{1}{9 \cdot T(n)}$. Reject if the Entropy Difference protocol rejects.

Now fix any $x \in \{0,1\}^n$. For $r \in \{0,1\}^{T(n)}$, $k_1, \ldots, k_{T(n)} \in \{1, \ldots, 9 \cdot T(n)\}^{T(n)}$, let $((X_1, Y_1), a_1, (X_2, Y_2), a_2, \ldots, (X_q, Y_q), a_q)$ be the query-answer sequence of execution $\mathcal{M}^{\mathcal{D}}(x)$ when $\mathcal{M}$ use r as its random tape and $\mathcal{D}$ use k_i as the randomness to answer the i-th query. Notice that (X_i, Y_i) is determined by

$x, r, k_1, \ldots, k_{i-1}$ and a_i is determined by $(X_i, Y_i), k_i$.

k_i is randomly chosen from $\{1, \ldots, 9 \cdot T(n)\}$. It specifies one of the $9 \cdot T(n)$ disjointed intervales:

$$\Big(-1, \frac{2}{9 \cdot T(n)} - 1\Big), \Big(\frac{2}{9 \cdot T(n)} - 1, \frac{4}{9 \cdot T(n)} - 1\Big), \\ \ldots, \Big(\frac{2(k-1)}{9 \cdot T(n)} - 1, \frac{2k}{9 \cdot T(n)} - 1\Big), \ldots, \Big(1 - \frac{2}{9 \cdot T(n)}, 1\Big)$$

Query (X_i, Y_i) is indenpdent from k_i, and $H(X_i) - H(Y_i)$ lays in at most one of these intervals. Due to k_i's randomness, the probability that $\frac{2(k_i-1)}{9 \cdot T(n)} - 1 < H(X_i) - H(Y_i) < \frac{2(k_i+1)}{9 \cdot T(n)} - 1$ is at most $\frac{1}{9 \cdot T(n)}$. Therefore, the probability that there exists an i that

$$\frac{2(k_i - 1)}{9 \cdot T(n)} - 1 < H(X_i) - H(Y_i) < \frac{2(k_i + 1)}{9 \cdot T(n)} - 1$$

is at most $\frac{1}{9}$.

Completeness of SP For any n-bit $x \in L_Y$, $\Pr[\mathcal{M}^{\mathcal{D}}(x) = 1] \geq \frac{8}{9}$. So with probability at least $\frac{7}{9}$ over the randomness of r and $\{k_i\}$, $\mathcal{M}^{\mathcal{D}}(x) = 1$ and the query-answer sequence satisfies

$$H(X_i) - H(Y_i) \notin \Big(\frac{2(k_i - 1)}{9 \cdot T(n)} - 1, \frac{2(k_i + 1)}{9 \cdot T(n)} - 1\Big)$$

for all i. In such case, and honest prover in SP could simply send query-answer sequence to verifier. The query-answer sequence, together with r, determined a valid execution in which $\mathcal{M}^{\mathcal{D}}(x) = 1$. And in the last step, the Entropy Difference protocol is always run on YES instances, so the verifier would reject, which is an error, with probability at most $\frac{1}{9}$.

Therefore, when $x \in L_Y$, protocol SP would accept x with probability $2/3$ for honest prover.

Soundness of SP For any n-bit $x \in L_N$, $\Pr[\mathcal{M}^{\mathcal{D}}(x) = 1] \leq \frac{1}{9}$. So with probability at least $\frac{7}{9}$ over the randomness of r and $\{k_i\}$, $\mathcal{M}^{\mathcal{D}}(x) \neq 1$ and the query-answer sequence of it satisfies

$$H(X_i) - H(Y_i) \notin \Big(\frac{2(k_i - 1)}{9 \cdot T(n)} - 1, \frac{2(k_i + 1)}{9 \cdot T(n)} - 1\Big)$$

for all i. Conditional on such case, let $((\hat{X}_1, \hat{Y}_1), \hat{a}_1, (\hat{X}_2, \hat{Y}_2), \hat{a}_2, \ldots, (\hat{X}_{q'}, \hat{Y}_{q'}), \hat{a}_{q'})$ be the query-answer sequence sent by the prover. There are a few possibility

- *If this sequence is identical with* $((X_1, Y_1), a_1, \ldots, (X_q, Y_q), a_q)$*,* then this sequence, together with r, determine a execution where $\mathcal{M}^{\mathcal{O}}(x) \neq 1$. So SP protocol would reject.

- *If this sequence is a leading substring of* $((X_1, Y_1), a_1, \ldots, (X_q, Y_q), a_q)$ *or the inverse,* then this sequence doesn't determine a valid execution of $\mathcal{M}^{\mathcal{O}}(x)$. So SP protocol would reject.
- *If there exists* i *such that* $(\hat{X}_j, \hat{Y}_j, \hat{a}_j) = (X_j, Y_j, a_j)$ *for* $j < i$ *and* $(\hat{X}_i, \hat{Y}_i) \neq (X_i, Y_i)$*,* then this sequence doesn't determine a valid execution of $\mathcal{M}^{\mathcal{O}}(x)$, as the next query is uniquely determined by input x, random tape r and previous answers. So SP protocol would reject.
- *If there exists* i *such that* $(\hat{X}_j, \hat{Y}_j, \hat{a}_j) = (X_j, Y_j, a_j)$ *for* $j < i$*,* $(\hat{X}_i, \hat{Y}_i) = (X_i, Y_i)$ *and* $\hat{a}_i \neq a_i$*,* then verifier and prover would run Entropy Difference protocol on a NO instance, which would lead to fail with high probability. So SP protocol would reject with probability at least $1 - \frac{1}{9 \cdot T(n)}$.

In either case, SP protocol would reject with probability at least $1 - \frac{1}{9 \cdot T(n)}$.

Therefore, when $x \in L_N$, protocol SP would accept x with probability at most $\frac{2}{9} + \frac{1}{9 \cdot T(n)}$, for any prover.