On Basing Private Information Retrieval on NP-Hardness

Tianren Liu
MIT*

Vinod Vaikuntanathan
MIT†

November 3, 2015

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. 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. We make progress along this line of inquiry by showing that the security of single-server single-round private information retrieval schemes cannot be based on **NP**-hardness, unless the polynomial hierarchy collapses. Our main technical contribution is in showing how to break the security of a PIR protocol given an **SZK** oracle. Our result is tight in terms of both the correctness and the privacy parameter of the PIR scheme.

*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.

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. Somewhat more precisely, "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", $\Pr[R^A(\phi) = 1] \geq 2/3$ if $\phi \in \mathsf{SAT}$ and $\Pr[R^A(\phi) = 1] \leq 1/3$ otherwise.

There are a handful of impossibility results which show that, assuming widely believed complexity-theoretic conjectures, the security of various cryptographic objects cannot be based on **NP**-hardness. We discuss these results in detail in Section 1.2. 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 (Informal)**.** If there is a probabilistic polynomial time reduction from solving SAT to breaking a single-server, one round, private information retrieval scheme, then $\mathbf{NP} \subseteq \mathbf{coAM}$.

Our result rules out security reductions from SAT that make black-box use of the adversary that breaks a PIR scheme. Other than being black-box in the adversary, the security reduction can be very general, in particular, it is allowed to make polynomially many adaptively chosen calls to the PIR-breaking adversary.

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 those choice of 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. 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].

In particular, Kushilevitz and Ostrovsky [KO97] were the first to show a construction of PIR with $O(n^\epsilon)$ communication (for any constant $\epsilon > 0$) assuming the existence of additively homomorphic encryption schemes. Some of the later constructions of PIR [CMS99, Lip05, GR05, BV11] achieve $\mathsf{polylog}(n)$ communication under number-theoretic assumptions such as the Phi-hiding assumption and the LWE assumption. Notably, all of them are single-round protocols, involving one message from the user to the server and one message back.

1.1 Our Techniques

The core of our proof is an attack against any single-server one-round PIR protocol given access to an **SZK** oracle. In particular, we show that given an oracle to the *entropy difference* (ED) problem, which is complete for **SZK**, one can break any single-server one-round PIR protocol. Once we have this result, the rest follows from a beautiful work of Mahmoody and Xiao [MX10] who show

that $\mathbf{BPP}^{\mathbf{SZK}} \subseteq \mathbf{AM} \cap \mathbf{coAM}$. That is, if there is a reduction from deciding SAT to breaking single-server one-round PIR, then $\mathsf{SAT} \in \mathbf{BPP}^{\mathbf{SZK}}$ and therefore, by [MX10], $\mathsf{SAT} \in \mathbf{AM} \cap \mathbf{coAM}$. In turn, from the work of Boppana, Håstad and Zachos [BHZ87], this means that the polynomial hierarchy collapses to the second level.

The intuition behind the attack against PIR protocols is simple. Assume that the database is uniformly random and the user's query is fixed. Let X be a random variable that denotes the database, and let A be a random variable that denotes the PIR answer (on input a query q from a user trying to retrieve the i-th bit). We have two observations.

1. *The answer enables the user to learn the i-th bit.* In other words, the mutual information between the i-th database bit X_i and the answer A has to be large. Indeed, we show that if the PIR protocol is correct with probability $1 - \varepsilon$, then this mutual information is at least $1 - h(\varepsilon)$, where h is the binary entropy function.

2. *The answer does not contain a large amount of information about all the database entries.* Indeed, the entropy of the answer is limited by its length which is much shorter than the size of the database. We show that for most indices j, the answer contains little information about the j-th bit, that is the mutual information between A and X_j is small.

We then proceed as follows. Given the user's query q, an efficient adversary can construct a circuit sampling from joint distribution $(X; A)$. Armed with the entropy difference ED oracle, the adversary can estimate $I(X_j; A)$ for any index j. Since $I(X_i; A)$ is close to 1 (where i is the index underlying the query q) and $I(X_j; A)$ is small for most indices j, the adversary can predict i much better than random guessing. This breaks the security of PIR.

We refer the reader to Theorem 8 for the formal statement, and to Proposition 7 which shows that the parameters of Theorem 8 are tight.

1.2 Related Work

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. In particular, one of their conditions require that it should be easy to certifying an invalid key as such.

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. A size-verifiable one-way function, roughly speaking, 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. While PIR schemes can be constructed from additively homomorphic encryption, we are not aware of a way to use PIR to obtain any type of non-trivial homomorphic encryption scheme.

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 $\mathbf{PH} \subseteq \mathbf{\Sigma_2}$ (that is, the polynomial hierarchy collapses to the second level). In contrast, our results rule out even adaptive reductions (to much stronger primitives).

2 Definitions

2.1 Information Theory Background

A *random variable* X over a finite set S is defined by its probability mass function $p_X : S \to [0,1]$ such that $\sum_{x \in S} p_X(x) = 1$. We use uppercase letters to denote random variables. The *Shannon entropy* of a random variable X, denoted $H(X)$, is defined as

$$H(X) = \sum_x p_X(x) \log_2 \frac{1}{p_X(x)}.$$

Let $\mathrm{Bern}(p)$ denote the Bernoulli distribution on $\{0,1\}$ which assigns a probability of p to 1 and $1-p$ to 0. We will denote by $h(p) = H(\mathrm{Bern}(p)) = p \log_2 \frac{1}{p} + (1-p) \log_2 \frac{1}{1-p}$ the Shannon entropy of the distribution $\mathrm{Bern}(p)$.

Let X and Y be two (possibly dependent) random variables. The *conditional entropy* of Y given X, denoted $H(Y|X)$, is defined as $H(Y|X) = H(XY) - H(X)$, where XY denotes the joint distribution of X and Y. Informally, $H(Y|X)$ measures the (residual) uncertainty of Y when X is known.

The *mutual information* between random variables X and Y is

$$I(X;Y) = H(X) + H(Y) - H(XY) = H(Y) - H(Y|X) = H(X) - H(X|Y)$$

which measures the information that X reveals about Y (and vice versa). In particular, if two random variables X, Y are independent, their mutual information is zero.

The *conditional mutual information* between random variables X and Y given Z, denoted $I(X;Y|Z)$, is defined as

$$I(X;Y|Z) = H(X|Z) + H(Y|Z) - H(XY|Z).$$

We will use without proof that entropy, conditional entropy, mutual information, conditional mutual information are non-negative.

We will need the following simple propositions.

Proposition 1. *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 $\Pr[X' = X] \geq 1 - \varepsilon$,*

$$I(X';X) \geq 1 - h(\varepsilon).$$

Proof. Clearly, $I(\hat{X};X) = H(X) - H(X|\hat{X}) = 1 - h(\varepsilon)$. Furthermore, the random variable $\hat{X} = X \oplus N$ minimizes the mutual information $I(\hat{X};X)$ under the constraint that $\Pr[\hat{X} = X] \geq 1 - \varepsilon$. In particular, we have

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

for any random variable X' satisfying $\Pr[X' = X] \geq 1 - \varepsilon$. □ □

Proposition 2 (Conditioning decreases entropy)**.** *For any random variables X, Y, Z, it holds that $H(X) \geq H(X|Y) \geq H(X|YZ)$.*

In general, conditioning can increase or decrease mutual information, but when conditioning on an *independent* variable, mutual information increases.

Proposition 3 (Conditioning on independent variables increases mutual information)**.** *For random variables X, Y, Z such that Y and Z are independent, $I(X;Y|Z) \geq I(X;Y)$.*

Proof. As Y, Z are independent, $H(Y|Z) = H(Y)$.

$$I(X;Y|Z) = H(Y|Z) - H(Y|XZ) \geq H(Y) - H(Y|X) = I(X;Y).$$

□

□

Proposition 4 (Data processing for mutual information)**.** *Assume random variables X, Y, Z satisfies $X \to Y \to Z$, i.e. X and Z are independent conditional on Y, then $I(X;Y) \geq I(X;Z)$.*

Proof. Since X and Z are independent conditional on Y (meaning $I(X;Z|Y) = 0$), we have $H(X|YZ) = H(X|Y)$. Thus

$$I(X;Y) = H(X) - H(X|Y) = H(X) - H(X|YZ) \geq H(X) - H(X|Z) = I(X;Z).$$

□

□

Proposition 5 (Chain rule for mutual information)**.** *For random variables $X_1, \ldots, X_n, Y$, it holds that*

$$I(X_1 \ldots X_n; Y) = \sum_{i=1}^{n} I(X_i; Y | X_1 \ldots X_{i-1}).$$

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, without revealing any information about 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 and the randomness used to generate the query, should be able to learn the i-th bit x_i.

We specialize our definitions to the case of single round protocols.

Definition 1 (Private information retrieval)**.** A single-server one round private information retrieval (PIR) scheme is a tuple $(\mathbf{Qry}, \mathbf{Ans}, \mathbf{Rec})$ of algorithms such that

- The query algorithm $\mathbf{Qry}$ is a probabilistic polynomial-time algorithm such that $\mathbf{Qry}(1^n, i) \to (q, \sigma)$, where $i \in [n]$. Here, q is the PIR query and σ is the secret state of the user (which, without loss of generality, is the randomness used by the algorithm).
- The answer algorithm $\mathbf{Ans}$ is a probabilistic polynomial-time algorithm such that $\mathbf{Ans}(x, q) \to a$, where $x \in \{0,1\}^n$. Let ℓ denote the length of the answer, i.e. $a \in \{0,1\}^\ell$.
- The reconstruction algorithm $\mathbf{Rec}$ is a probabilistic polynomial-time algorithm such that $\mathbf{Rec}(a, \sigma) \to b$ where $b \in \{0,1\}$.

Correctness. A PIR scheme $(\mathbf{Qry}, \mathbf{Ans}, \mathbf{Rec})$ is $(1-\varepsilon)$-correct if for any $x \in \{0,1\}^n$ and 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 probability is taken over the random tapes of $\mathbf{Qry}, \mathbf{Ans}, \mathbf{Rec}$. We call ϵ the error probability of the PIR scheme.

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

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

(Here and in the sequel, τ will denote the state that $\mathcal{A}_1$ passes on to $\mathcal{A}_2$).

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).

We call a PIR scheme δ-GUESS-secure if 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 in n, as we show in the next proposition.

Proposition 6. *If a PIR scheme is δ_1-IND-secure (according to Definition (1)), then it is δ_2-GUESS-secure (according to Definition (2)) where $\delta_2 = n\delta_1$. Similarly, if a PIR scheme is δ_2-GUESS-secure, then it is δ_1-IND-secure where $\delta_1 = \delta_2/2$.*

Proof. Assume that a probabilistic polynomial-time (p.p.t.) adversary algorithm $\mathcal{A}$ breaks δ_2-privacy according to definition (2). We construct an adversary $\mathcal{B} = (\mathcal{B}_1, \mathcal{B}_2)$ that breaks definition (1).

The algorithm $\mathcal{B}_1(1^n)$ picks two random indices i_0 and i_1 and outputs i_0, i_1 and $\tau = (i_0, i_1)$, algorithm $\mathcal{B}_2(1^n, q, \tau = (i_0, i_1))$ calls $\mathcal{A}(1^n, q)$ to get an index i, and outputs 0 if and only if $i = i_0$.

Then,

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

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

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

Thus, $(\mathcal{B}_1,\mathcal{B}_2)$ breaks $\frac{\delta_2}{n}$-privacy according to definition (1).

In the other direction, assume that a p.p.t. adversary algorithm $\mathcal{A}=(\mathcal{A}_1,\mathcal{A}_2)$ breaks δ_1-privacy according to definition (1). We construct an adversary $\mathcal{B}$ that works as follows. $\mathcal{B}$ runs $\mathcal{A}_1$ to get $(i_0,i_1,\tau)\leftarrow\mathcal{A}_1(1^n)$, gets a challenge query q and runs $\mathcal{A}_2$ to get $b\leftarrow\mathcal{A}_2(1^n,q,\tau)$. $\mathcal{B}$ simply outputs i_b. Then, we have:

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

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

Thus, $\mathcal{B}$ breaks $2\delta_1$-privacy according to definition (2). □ □

Answer Communication Complexity. We define the answer communication complexity of the PIR scheme to be the number of bits in the server's response to a PIR query. (This is denoted by ℓ in Definition 1). Similarly, we call the bit-length of the query as the query communication complexity, and their sum as the total communication complexity. In this work, we are interested in PIR protocols with a "small" answer communication complexity (regardless of their query communication complexity). Since our main result is a lower bound, this only makes it stronger.

Typically, we are interested in PIR schemes with answer communication complexity $\ell=o(n)$. Otherwise, e.g. when $\ell=n$, there is a trivial PIR protocol with perfect privacy, where the user sends nothing and the server sends the whole database x. The following proposition shows a tradeoff between the correctness error and answer communication complexity of perfectly private PIR schemes.

Proposition 7. *There exists a PIR scheme with perfect information-theoretic privacy, error probability ε, and answer communication complexity $\ell=n\cdot(1-h(\varepsilon)+O(n^{-1/4}))$.*

Consider a PIR scheme where the user sends nothing and the server sends the whole database to the user, incurring an answer communication complexity of n bits. The query contains no information about the index i, and this achieves perfect privacy and correctness. The idea is that given the possibility of a correctness error of ε, the server can compress the database into $\ell < n$ bits, such that the user can still recover the database with at most ε error.

This is a fundamental problem in information theory, called "lossy source coding" [Sha59]. Let X be a uniform random Bernoulli variable. Proposition 1 says that for any random variable $\hat{X}$ such that $\Pr[\hat{X} = X] \geq 1 - \varepsilon$, $I(\hat{X}, X) \geq 1 - h(\varepsilon)$. Therefore, to compress a random binary string and to recover the string from the lossy compression with $(1 - \varepsilon)$ accuracy, 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 achieves $(1 - \varepsilon)$ correctness on a random database. Moreover, we can extend this to work for any database by the following scheme which has a query communication complexity of n bits and an answer communication complexity of ℓ bits.

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

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

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 δ-GUESS-security of PIR scheme $(\mathbf{Qry}, \mathbf{Ans}, \mathbf{Rec})$ if there exists a probabilistic polynomial-time oracle Turing machine (OTM) M such that for all x and for all "legal" oracles $\mathcal{O}_\delta^{\mathsf{PIR}}$,

$$\Pr[M^{\mathcal{O}_\delta^{\mathsf{PIR}}}(x) = 1] \geq 2/3 \quad \text{if } x \in L$$
$$\Pr[M^{\mathcal{O}_\delta^{\mathsf{PIR}}}(x) = 1] \leq 1/3 \quad \text{if } x \notin L$$

where the probability is taken over the coins of the machine M and the oracle $\mathcal{O}_\delta^{\mathsf{PIR}}$. We stress that M is allowed to make adaptive queries to the oracle.

By a legal δ-breaking oracle $\mathcal{O}_\delta^{\mathsf{PIR}}$, we mean one that satifies

$$\Pr\left[\begin{array}{c} j \leftarrow [n] \\ \mathbf{Qry}(1^n, j) \to (q, \sigma) : \ j = j' \\ \mathcal{O}_\delta^{\mathsf{PIR}}(q) \to j' \end{array}\right] \geq \frac{1}{n}(1 + \delta) \tag{3}$$

where the probability is taken over the coins used in the experiment, including those of $\mathbf{Qry}$ and $\mathcal{O}_\delta^{\mathsf{PIR}}$.

2.3 Entropy Difference

Entropy Difference (ED) is a promise problem that is complete for **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.

We list a few elementary observations regarding the power of an oracle that decides the entropy difference problem.

First, given an entropy difference oracle, a polynomial-time algorithm can distinguish between two distributions X and Y such that either $H(X) \geq H(Y) + \frac{1}{s}$ or $H(Y) \geq H(X) + \frac{1}{s}$ for any polynomial function s. That is, one can solve the entropy difference problem up to any inverse-polynomial precision. This can be done as follows: 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, a polynomial-time algorithm can use the Entropy Difference oracle to distinguish between $H(X) \geq \hat{h} + \frac{1}{s}$ and $H(X) \leq \hat{h} - \frac{1}{s}$ for a given $\hat{h}$. This can be done as follows: construct a distribution Y that $2s\hat{h} - 1 < H(Y) < 2s\hat{h} + 1$ and query the Entropy Difference oracle with the distributions $X_1 \dots X_{2s}$ and Y, where $X_1, \dots, X_{2s}$ are independent copies of X. Therefore, a polynomial-time algorithm given Entropy Difference oracle can estimate $H(X)$ to within any additive inverse-polynomial precision by binary search.

Finally, assume that X and Y are random variables encoded as a circuit which samples from their joint distributions. Then, a polynomial-time algorithm given an Entropy Difference oracle can also estimate the conditional entropy $H(X|Y)$, mutual information $I(X;Y)$ to any inverse-polynomial precision. Here the precision is measured by absolute additive error.

3 PIR and NP-hardness

Theorem 8 (Main Theorem)**.** *Let* $\Pi = (\mathbf{Qry}, \mathbf{Ans}, \mathbf{Rec})$ *be any* $(1 - \epsilon)$*-correct PIR scheme with* n*-bit databases and answer communication complexity* ℓ*. Let* L *be any language. If*

1. *there exists a reduction from* L *to breaking the* δ*-privacy of* Π *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 $L \in \mathbf{AM} \cap \mathbf{coAM}$*.*

In particular, using the result of [BHZ87], this tells us that unless the polynomial hierarchy collapses, there is no reduction from SAT to breaking the privacy of a PIR scheme with parameters as above.

We note that the bound in the lemma is tight. As Proposition 7 shows, there is in fact a perfectly (information-theoretically) private PIR protocol with a matching answer communication complexity of $n \cdot (1 - h(\varepsilon)) + o(n)$.

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 deciding a language L to breaking a PIR scheme, and the PIR scheme has a low answer communication complexity, then L can be reduced to the entropy difference problem (defined in Section 2.3).

Lemma 9 ($\mathbf{BPP}^{\mathcal{O}_\delta^{\mathsf{PIR}}} \subseteq \mathbf{BPP}^{\mathsf{ED}}$)**.** *Let $\Pi = (\mathbf{Qry}, \mathbf{Ans}, \mathbf{Rec})$ be any $(1-\epsilon)$-correct PIR scheme with answer communication complexity ℓ and let L be any language. If there exists a reduction from L to δ-breaking the privacy of a PIR protocol such that*

$$\frac{1-h(\varepsilon)}{\ell} - \frac{1+\delta}{n} \geq \frac{1}{p(n)}$$

for some polynomial function $p(n)$, then there exists a probabilistic polynomial time reduction from L to ED.

As noted in Proposition 7, this condition is tight as there exists a PIR scheme achieving perfect privacy ($\delta = 0$) if $\ell \approx n \cdot (1 - h(\varepsilon))$.

The next lemma, originally shown in [MX10] and used in [BL13b], states that any language decidable by a randomized oracle machine with access to an entropy difference oracle is in $\mathbf{AM} \cap \mathbf{coAM}$.

Lemma 10 ($\mathbf{BPP}^{\mathsf{ED}} \subseteq \mathbf{AM} \cap \mathbf{coAM}$ [MX10])**.** *For any language L, if there exists an OTM M such that for any oracle $\mathcal{O}$ solving entropy difference*

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

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

3.1 Proof of the Main Theorem

Assume that there exists a reduction from deciding a language L to breaking PIR with parameters as stated in Theorem 8. In other words, there is a reduction from L to δ-breaking PIR where

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

where the inequality is using the hypothesis in Theorem 8 that $\ell \cdot (1+\delta) \leq n \cdot (1 - h(\varepsilon)) - 1/p(n)$.

Then, by Lemma 9, there is a reduction from deciding L to solving the entropy difference problem ED. Combined with Lemma 10, we deduce that $L \in \mathbf{AM} \cap \mathbf{coAM}$.

3.2 Proof of Lemma 9

We start with two claims that are central to our proof. The first claim says that because of $(1-\varepsilon)$-correctness of the PIR scheme, the PIR answer a on a query $q \leftarrow \mathbf{Qry}(1^n, i)$ has to contain information about the i^{th} bit of the database x_i.

Claim. Let $\Pi = (\mathbf{Qry}, \mathbf{Ans}, \mathbf{Rec})$ be a PIR scheme which is $(1-\varepsilon)$-correct. Fix any index $i \in [n]$. Let X denote a random n-bit database; $(Q, \Sigma) \leftarrow \mathbf{Qry}(1^n, i)$; and $A \leftarrow \mathbf{Ans}(X, Q)$. Then,

$$I(A; X_i | Q) \geq 1 - h(\varepsilon). \tag{4}$$

Proof. Define the random variable $\hat{X}_i \leftarrow \mathbf{Rec}(A, \Sigma)$. Since the PIR scheme is $(1-\varepsilon)$-correct, $\Pr[\hat{X}_i = X_i] \geq 1 - \varepsilon$. Since X_i is a uniform Bernoulli variable, we know from Proposition 1 that $I(\hat{X}_i; X_i) \geq 1 - h(\varepsilon)$.

As X_i is independent from Q, we know from Proposition 3 that

$$I(\hat{X}_i; X_i|Q) \geq I(\hat{X}_i; X_i).$$

Next, we claim that conditioning on Q, we have $X_i \to A \to \hat{X}_i$, in other words, $I(X_i; \hat{X}_i|A, Q) = 0$. This is because when A and Q are given, one can sample a random Σ consistent with Q, then compute $\hat{X}_i$ from Σ and A, with no knowledge of X_i. Now, Proposition 4 (data processing inequality for mutual information) shows that $I(A; X_i|Q) \geq I(\hat{X}_i; X_i|Q)$.

Combining what we have,

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

This completes the proof. □ □

Claim. Let $\Pi = (\mathbf{Qry}, \mathbf{Ans}, \mathbf{Rec})$ be a PIR scheme with an answer communication complexity of ℓ bits. Let X denote a random n-bit database; $(Q, \Sigma) \leftarrow \mathbf{Qry}(1^n, i)$; and $A \leftarrow \mathbf{Ans}(X, Q)$. Then, for any potential query q,

$$\sum_{j=1}^{n} I(A; X_j|Q = q) \leq \ell. \tag{5}$$

Proof. Recall that, by definition,

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

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,

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

where the first inequality is implied by the Proposition 3 and the second equality is Proposition 5 (chain rule for mutual information). □ □

Equations (4) and (5) are the core of the proof of Lemma 9. Equation (4) shows that, when retrieving the i-th bit, the mutual information between X_i and server's answer A is large. Equation (5) shows that, the sum of mutual information between each bit X_j and server's answer A is bounded by the answer communication complexity. Therefore, if we could measure the mutual information by an Entropy Difference oracle, we would have a pretty good knowledge of i.

In particular, we proceed as follows. Assume language L can be solved by a probabilistic polynomial-time oracle Turing machine $\mathcal{M}$ given any oracle $\mathcal{O}_\delta^{\mathsf{PIR}}$ that breaks the δ-GUESS-security of the PIR scheme $(\mathbf{Qry}, \mathbf{Ans}, \mathbf{Rec})$ where

$$\frac{1+\delta}{n} \leq \frac{1 - h(\varepsilon)}{\ell} - \frac{1}{p(n)} \tag{6}$$

Algorithm 1 Solving L given ED oracle on input x

1. Simulate $\mathcal{M}^{\mathcal{O}_\delta^{\mathsf{PIR}}}(x)$
2. Whenever $\mathcal{M}$ queries $\mathcal{O}_\delta^{\mathsf{PIR}}(q)$, do the following:
   (a) For each index $j = 1, \ldots, n$, use the entropy difference oracle to estimate
   $$\mu_j = I(A; X_j | Q = q)$$
   to $\frac{1}{2n \cdot p(n)}$ precision. More precisely, construct a circuit $C = C_{q,j}$ such that
   $$C_{q,j}(x, r) = (x_j, \mathbf{Ans}(x, q, r))$$
   and estimate the mutual information between the two components of C's output. Let $\hat{\mu}_j \in [0, 1]$ denote the estimation.
   (b) Sample a random value $\hat{i} \in [n]$ according to probability distribution $p(\hat{i}) = \hat{\mu}_{\hat{i}} / \sum_j \hat{\mu}_j$
   (c) Answer $\mathcal{M}$'s query by $\hat{i}$
3. Output what $\mathcal{M}$ output

where $p(\cdot)$ is a fixed polynomial. We construct an efficient oracle algorithm (see Algorithm 1) that solves L given an Entropy Difference oracle $\mathcal{O}^{\mathsf{ED}}$.

For any query q and index i, when $\mathcal{O}_\delta^{\mathsf{PIR}}(q)$ is simulated,

$$\begin{aligned}\Pr\big[\hat{i} \leftarrow \mathcal{O}_\delta^{\mathsf{PIR}}(q) : \hat{i} = i\big] = \frac{\hat{\mu}_i}{\sum_j \hat{\mu}_j} &\geq \frac{\mu_i - \frac{1}{2n \cdot p(n)}}{\sum_j \mu_j + \frac{1}{2p(n)}} \\ &\geq \frac{\mu_i - \frac{1}{2p(n)}}{\ell + \frac{1}{2p(n)}} \geq \frac{\mu_i}{\ell} \frac{1 - \frac{1}{2p(n)}}{1 + \frac{1}{2p(n)}} \geq \frac{\mu_i}{\ell}\Big(1 - \frac{1}{p(n)}\Big) \geq \frac{\mu_i}{\ell} - \frac{1}{p(n)}\end{aligned}$$

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

$$\begin{aligned}&\Pr\big[q \leftarrow \mathbf{Qry}(1^n, i), \hat{i} \leftarrow \mathcal{O}_\delta^{\mathsf{PIR}}(q) : \hat{i} = i\big] \\ &\quad= \mathop{\mathbb{E}}_{q \leftarrow \mathbf{Qry}(1^n, i)}\big[\Pr[\hat{i} = i | Q = q]\big] \\ &\quad\geq \mathop{\mathbb{E}}_{q \leftarrow \mathbf{Qry}(1^n, i)}\Big[\frac{\mu_i}{\ell} - \frac{1}{p(n)}\Big] \\ &\quad= \frac{\mathbb{E}_{q \leftarrow \mathbf{Qry}(1^n, i)}[\mu_i]}{\ell} - \frac{1}{p(n)} \\ &\quad\geq \frac{1 - h(\varepsilon)}{\ell} - \frac{1}{p(n)} \\ &\quad\geq \frac{1}{n}(1 + \delta)\end{aligned}$$

4 Discussion and Open Questions

We show that any non-trivial single-server single-round PIR scheme can be broken in **SZK**. Since languages that can be decided with (adaptive) oracle access to **SZK** live in $\mathbf{AM} \cap \mathbf{coAM}$, this shows that there cannot be a reduction from SAT to **SZK**, and therefore also from SAT to breaking single-server single-round PIR.

The crucial underlying feature of single-round PIR schemes that we use is the ability to "re-randomize". By this, we mean that given a user query q for an index i, one can generate not just a single transcript, but the distribution over all transcripts where the database is uniformly random and the prefix of the transcript is q. This ability to generate a transcript distribution of the same index and random database allows the adversary to break a PIR scheme with an **SZK** oracle.

Indeed, this is reminiscent of the work of Bogdanov and Lee who show that breaking homomorphic encryption is not **NP**-hard [BL13b]. Their main contribution is to show that any homomorphic encryption (whose homomorphic evaluation process produces a ciphertext that is statistically close to a fresh encryption) can be turned into a (weakly) re-randomizable encryption scheme. Once this is done, an **SZK** oracle can be used to break the scheme in much the same way as we do.

A natural question arising from our work is to extend our results to multi-round PIR. The key technical difficulty that arises is in sampling a random "continuation" of a partial transcript. We conjecture that our lower bound can nevertheless be extended to the multi-round case, and leave this as an interesting open problem.

Acknowledgments. We would like to thank the anonymous TCC reviewers for their careful reading and excellent suggestions, and 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.

[BHZ87] Ravi B. Boppana, Johan Håstad, and Stathis Zachos. Does co-np have short interactive proofs? *Inf. Process. Lett.*, 25(2):127–132, 1987.

[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.