PAPER Special Section on Cryptography and Information Security

On Lower Bounds for the Communication Complexity of Private Information Retrieval*

Toshiya ITOH†, ***Regular Member***

SUMMARY Private information retrieval for $k \geq 1$ databases (denoted by (k, ℓ)-PIR for short) is a protocol that (1) a user sends an ℓ tuple query to each of k noncommunicating replicated databases; (2) each database responds the user with an answer corresponding to the ℓ tuple query; (3) the user privately retrieve any single bit out of the n bits of data stored in k databases. In this model, "privacy" implies that the user retrieves the bit he is interested in but releases to each database nothing about which bit he wishes to get. In general, the efficiency of (k, ℓ)-PIR is measured by the total amount of bits exchanged between the user and the k databases, but few about its lower bounds are known except for restricted cases. In this paper, we classify (k, ℓ)-PIR into a *linear* type, a *multilinear* type, and an *affine* type with respect to the relationship between queries to each database (made by the user) and answers to the user (made by each database), and show that (1) the lower bound for the communication complexity of any multilinear type (k, ℓ)-PIR is $\Omega(\sqrt[\ell+1]{n})$ (Theorem 3.1); (2) the lower bound for the communication complexity of any linear type (k, ℓ)-PIR is $\Omega(\sqrt{n})$ (Corollary 3.2); (3) the lower bound for the communication complexity of any affine type (k, ℓ)-PIR is $\Omega(\sqrt[\ell+1]{n})$ (Theorem 4.2).

key words: *private information retrieval, communication complexity, linear type, multilinear type, affine type*

1. Introduction

1.1 Background

Private information retrieval for $k \geq 1$ databases (denoted by (k, ℓ)-PIR for short) is initiated by Chor et al. [4] as a useful tool for a user to privately get information through networks. It seems quite natural to ask for the privacy of the user. For example, an investor (or a speculator) that makes access to the stock-market database to get the value of a certain stock may wish to keep private which stock he is interested in. Informally, (k, ℓ)-PIR is an interactive protocol that (1) a user sends an ℓ tuple query to each of k noncommunicating replicated databases; (2) each of the k databases responds the user with an answer corresponding to the ℓ tuple query; (3) the user privately retrieves any single bit out of the n bits of data stored in the databases. In this model, "private" implies that the user is able to retrieve his desired bit but releases to each database nothing about which bit he wishes to get in the information-theoretic sense. In the practical point of view, the communication complexity between the user and each of the k databases would be one of the most essential resources for constructing efficient (k, ℓ)-PIR.

When the user makes access to only a single database, he may ask for a copy of the whole database to privately retrieve the bit that he is interested in. This requires $O(n)$ communication complexity, however, it is proved to be essentially the best he can do. In fact, Chor et al. [4] showed that for any integer $\ell \geq 1$, $(1, \ell)$-PIR requires $\Omega(n)$ communication complexity. To reduce the communication complexity of (k, ℓ)-PIR, Chor et al. [4] proposed (2,1)-PIR with communication complexity $12\sqrt[3]{n}$ by applying the covering codes [13], Ambainis [1] showed $(k, k-1)$-PIR with communication complexity $O(2^{k^2}\sqrt[2k-1]{n})$, and Ishai and Kushilevitz [9] showed $(k, (2k-1)(k-1))$-PIR with communication complexity $O(k^3\sqrt[2k-1]{n})$, where n is the length of the data stored in the k replicated databases.

It is conjectured by Chor et al. [4] that $O(\sqrt[3]{n})$ is the lower bound for the communication complexity of $(2, \ell)$-PIR for any integer $\ell \geq 1$. To overcome this, Chor and Gilboa [3] extended (k, ℓ)-PIR and defined in a natural manner *computationally* private information retrieval for $k \geq 1$ databases (denoted by (k, ℓ)-CPIR) applying cryptographic primitives. This is also an interactive protocol similar to (k, ℓ)-PIR but the user releases to each of the k databases nothing about which bit he really tries to get in the computational sense. In this model, Chor and Gilboa [3] showed that for any $\varepsilon > 0$, there exists $(2, \ell)$-CPIR with communication complexity $O(n^\varepsilon)$ assuming that pseudorandom generators exist [8], [10]. For any integer $\ell \geq 1$, it seems impossible to construct $(1, \ell)$-CPIR with communication complexity $o(n)$. However, Kushilevitz and Ostrovsky [11] showed that for any $\varepsilon > 0$, there exists $(1, \ell)$-CPIR with communication complexity $O(n^\varepsilon)$ under the stronger assumption that the quadratic residuosity is hard [7], and Cachin et al. [5] showed that there exists $(1, \ell)$-CPIR with communication complexity $O(\log^{O(1)} n)$ under the novel assumption that the Φ-hiding is hard. Recently, Kushilevitz and Ostrovsky [12] showed that there exists $(1, \ell)$-CPIR with communication complexity less than n under a general assump-

Manuscript received March 24, 2000.
Manuscript revised June 12, 2000.

†The author is with Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology, Yokohama-shi, 226-8502 Japan.

*This research has been supported by the Kayamori Foundation of Informational Science Advancement.

tion that one-way trapdoor permutations exist.

1.2 Motivation

In general, the privacy of users in any (k,ℓ)-CPIR is asymptotically guaranteed with respect to n (the length of data stored in the database), i.e., the privacy of users in any (k,ℓ)-CPIR is protected for sufficiently large n. This implies that the privacy of users in any (k,ℓ)-CPIR is not necessarily guaranteed for reasonable size of data stored in the database(s). Thus from the privacy point of view, (k,ℓ)-PIR is more advantageous than (k,ℓ)-CPIR, because any (k,ℓ)-PIR protects the privacy of users (in the information-theoretic sense) for any size of data stored in the k replicated databases.

As mentioned in Sect. 1.1, we already have $(1,\ell)$-CPIR with low communication complexity under the reasonable assumptions, but we know few about how small the communication complexity of (k,ℓ)-PIR could be. So we investigate lower bounds for the communication complexity of (k,ℓ)-PIR to precisely understand the inherent natures and properties of (k,ℓ)-PIR.

1.3 Main Results

To investigate the lower bounds for the communication complexity of (k,ℓ)-PIR, we classify (k,ℓ)-PIR into *linear*, *multilinear*, and *affine* types with respect to the relationship between queries to each database and answers to the user (the definitions of linear, multilinear, and affine types will be given later). It is not known whether every (k,ℓ)-PIR necessarily belongs to one of these three types, however, all known (k,ℓ)-PIR (see, e.g., [1], [4], [9]) can be classified into a linear type, an multilinear type, or an affine type.

Then we show that (1) the lower bound for the communication complexity of any multilinear type (k,ℓ)-PIR is $\Omega(\sqrt[\ell+1]{n})$ (Theorem 3.1); (2) the lower bound for the communication complexity of any linear type (k,ℓ)-PIR is $\Omega(\sqrt{n})$ (Corollary 3.2); (3) the lower bound for the communication complexity of any affine type (k,ℓ)-PIR is $\Omega(\sqrt[\ell+1]{n})$ (Theorem 4.2).

2. Preliminaries

A model of (k,ℓ)-PIR is similar to that of multi-prover interactive proofs [2]. We use $\mathcal{U}$ to denote a user that is a probabilistic polynomial time (interactive) Turing machine and $\mathcal{DB}_1, \mathcal{DB}_2, \ldots, \mathcal{DB}_k$ to denote deterministic polynomial time (interactive) Turing machines that are allowed to communicate with $\mathcal{U}$ but are not allowed to communicate with each other, i.e., $\mathcal{DB}_1, \mathcal{DB}_2, \ldots, \mathcal{DB}_k$ are physically separated. Assume that each database $\mathcal{DB}_j$ stores the same content $x \in \{0,1\}^n$ of length $n > 0$ and that for some polynomial s in n, the user $\mathcal{U}$ has access to random sequence $r \in \{0,1\}^{s(n)}$. For simplicity, we use s instead of $s(n)$ in the rest of the paper.

In general, we can define multiple round (k,ℓ)-PIR for any integers $k, \ell \geq 1$. In this paper, however, we consider only single round (k,ℓ)-PIR, because all known (k,ℓ)-PIR (see, e.g., [1], [4], [9]) are single round. For simplicity, we use $[h]$ to denote the set $\{1, 2, \ldots, h\}$ for any integer $h \geq 1$.

Definition 2.1 (Information Retrieval Scheme [4])**:** We say that $\Pi = (\mathcal{U}; \mathcal{DB}_1, \mathcal{DB}_2, \ldots, \mathcal{DB}_k)$ is single round information retrieval for $k \geq 1$ databases (denoted by (k,ℓ)-IR for short) if for any $n > 0$ and any $x \in \{0,1\}^n$, it satisfies the following: For each $i \in [n]$, each $j \in [k]$, and each $r \in \{0,1\}^s$, (1) $\mathcal{U}$ sends $q_j(i,r) = \langle q_{j,1}(i,r), q_{j,2}(i,r), \ldots, q_{j,\ell}(i,r)\rangle = \mathcal{U}(i,j,r;n)$ to $\mathcal{DB}_j$; (2) $\mathcal{DB}_j$ sends $a_j = \mathcal{DB}_j(x, q_j(i,r))$ to $\mathcal{U}$; (3) $\mathcal{U}$ computes $x_i = \mathcal{U}(i,r;a_1,a_2,\ldots,a_k)$.

Since $\mathcal{U}$ is a probabilistic polynomial time Turing machine, $q_j(i,r) : \{0,1\}^s \to \{0,1\}^*$ is a random variable for any $j \in [k]$ and any $i \in [n]$. Informally, we say that $\Pi = (\mathcal{U}; \mathcal{DB}_1, \mathcal{DB}_2, \ldots, \mathcal{DB}_k) \in (k,\ell)$-IR is *private* if $\mathcal{U}$ releases nothing about which bit he tries to retrieve.

Definition 2.2 (User Privacy [4])**:** We say that $\Pi = (\mathcal{U}; \mathcal{DB}_1, \mathcal{DB}_2, \ldots, \mathcal{DB}_k)$ is single round private information retrieval for $k \geq 1$ databases (denoted by (k,ℓ)-PIR for short) if it is (k,ℓ)-IR and satisfies the following: For each $j \in [k]$, any $n > 0$, and any $i_1, i_2 \in [n]$, $q_j(i_1,r)$ and $q_j(i_2,r)$ are identically distributed.

Definition 2.3 (Communication Complexity [4])**:** Let $\Pi = (\mathcal{U}; \mathcal{DB}_1, \mathcal{DB}_2, \ldots, \mathcal{DB}_k)$ be (k,ℓ)-PIR for any integers $k, \ell \geq 1$. Then for all $n > 0$, the communication complexity $\textsc{Comm}_\Pi(n)$ of the protocol Π is the sum of the total bits exchanged between $\mathcal{U}$ and $\mathcal{DB}_j$'s, i.e.,

$$\textsc{Comm}_\Pi(n) = \max_{(x,i,r)\in\{0,1\}^n\times[n]\times\{0,1\}^s} \sum_{j\in[k]} \{|q_j(i,r)| + |a_j|\}.$$

For each $j \in [k]$, each $i \in [n]$, and any $r \in \{0,1\}^s$, we can assume without loss of generality that $a_j \in \{0,1\}^*$ and $q_j(i,r) \in \{0,1\}^*$, and we can regard a_j and $q_j(i,r)$ as vectors over $GF(2)$.

Remark 2.4: For any pair of integers $k, \ell \geq 1$, let $\Pi = \langle \mathcal{U}; \mathcal{DB}_1, \mathcal{DB}_2, \ldots, \mathcal{DB}_k\rangle$ be (k,ℓ)-IR as given in Definition 2.1. For each $j \in [k]$, each $i \in [n]$, and each $h \in [\ell]$, assume that $q_{j,h}(i,r) \in \{0,1\}^{m_{j,h}}$ for any $r \in \{0,1\}^s$, and let $e_{j,h}[t_{j,h}] = 0^{t_{j,h}-1}10^{m_{j,h}-t_{j,h}}$ for each $t_{j,h} \in [m_{j,h}]$ and $e_{j,h}[0] = 0^{m_{j,h}}$.

For any pair of vectors $a, b \in \{0,1\}^m$, we use $a \oplus b$ to denote the *bitwise* Xor of a and b. We classify (k,ℓ)-PIR into a *linear* type, a *multilinear* type, and an *affine* type according to the relationship between queries to each database and answers to the user.

Definition 2.5 (Multilinear)**:** For any pair of integers $k, \ell \geq 1$, we say that $\Pi = (\mathcal{U}; \mathcal{DB}_1, \mathcal{DB}_2, \ldots, \mathcal{DB}_k)$ is multilinear type (k,ℓ)-PIR (denoted by (k,ℓ)-MLPIR for short) if it is (k,ℓ)-PIR and satisfies that

$$\begin{aligned}
&\mathcal{DB}_j(x, \langle q_{j,1}(i,r), \ldots, q_{j,h-1}(i,r), \alpha \oplus \beta, \\
&\qquad\qquad q_{j,h+1}(i,r), \ldots, q_{j,\ell}(i,r)\rangle) \\
&= \mathcal{DB}_j(x, \langle q_{j,1}(i,r), \ldots, q_{j,h-1}(i,r), \alpha, \\
&\qquad\qquad q_{j,h+1}(i,r), \ldots, q_{j,\ell}(i,r)\rangle) \\
&\quad \oplus \mathcal{DB}_j(x, \langle q_{j,1}(i,r), \ldots, q_{j,h-1}(i,r), \beta, \\
&\qquad\qquad q_{j,h+1}(i,r), \ldots, q_{j,\ell}(i,r)\rangle),
\end{aligned}$$

for any $j \in [k]$, any $h \in [\ell]$, any $x \in \{0,1\}^n$, any $i \in [n]$, any $r \in \{0,1\}^s$, and any $\alpha, \beta \in \{0,1\}^{m_{j,h}}$.

For any pair of integers $k, \ell \geq 1$, we say that $\Pi = (\mathcal{U}; \mathcal{DB}_1, \mathcal{DB}_2, \ldots, \mathcal{DB}_k)$ is linear type (k,ℓ)-PIR (denoted by (k,ℓ)-LPIR for short) if it is multilinear type (k,ℓ)-PIR and $\ell = 1$.

Definition 2.6 (Affine)**:** For any pair of integers $k, \ell \geq 1$, we say that $\Pi = (\mathcal{U}; \mathcal{DB}_1, \mathcal{DB}_2, \ldots, \mathcal{DB}_k)$ is affine type (k,ℓ)-PIR (denoted by (k,ℓ)-APIR for short) if it is (k,ℓ)-PIR and satisfies that

$$\begin{aligned}
&\mathcal{DB}_j(x, \langle q_{j,1}(i,r), \ldots, q_{j,h-1}(i,r), \alpha \oplus \beta, \\
&\qquad\qquad q_{j,h+1}(i,r), \ldots, q_{j,\ell}(i,r)\rangle) \\
&= \mathcal{DB}_j(x, \langle q_{j,1}(i,r), \ldots, q_{j,h-1}(i,r), \alpha, \\
&\qquad\qquad q_{j,h+1}(i,r), \ldots, q_{j,\ell}(i,r)\rangle) \\
&\quad \oplus \mathcal{DB}_j(x, \langle q_{j,1}(i,r), \ldots, q_{j,h-1}(i,r), \beta, \\
&\qquad\qquad q_{j,h+1}(i,r), \ldots, q_{j,\ell}(i,r)\rangle) \\
&\quad \oplus \mathcal{DB}_j(x, \langle q_{j,1}(i,r), \ldots, q_{j,h-1}(i,r), 0^{m_{j,h}}, \\
&\qquad\qquad q_{j,h+1}(i,r), \ldots, q_{j,\ell}(i,r)\rangle),
\end{aligned}$$

for any $j \in [k]$, any $h \in [\ell]$, any $x \in \{0,1\}^n$, any $i \in [n]$, any $r \in \{0,1\}^s$, and any $\alpha, \beta \in \{0,1\}^{m_{j,h}}$.

In the rest of this paper, we simply denote k-LPIR instead of (k,ℓ)-LPIR, because $\ell = 1$ for (k,ℓ)-LPIR. For any integers $k, \ell \geq 1$, we use k-$\mathcal{LPIR}$ to denote the collection of k-LPIR, (k,ℓ)-$\mathcal{MLPIR}$ to denote the collection of (k,ℓ)-MLPIR, and (k,ℓ)-$\mathcal{APIR}$ to denote the collection of (k,ℓ)-APIR. We also use (k,ℓ)-$\mathcal{PIR}$ to denote the collection of (k,ℓ)-PIR for any integer $k \geq 1$ and any integer $\ell \geq 1$.

The following fact plays a crucial role to derive lower bounds for the communication complexity of (k,ℓ)-MLPIR, k-LPIR, and (k,ℓ)-APIR for any integer $k \geq 2$ and any integer $\ell \geq 1$.

Fact 2.7 [4, §5.1]**:** For all $n > 0$, any integer $\ell \geq 1$, and any protocol $\Pi \in (1,\ell)$-$\mathcal{PIR}$, $\textsc{Comm}_\Pi(n) \geq n$.

3. Communication Complexity of Multilinear Type PIR

3.1 A Lower Bound for Multilinear Type PIR

In this subsection, we derive a lower bound for the communication complexity of (k,ℓ)-MLPIR.

Theorem 3.1: For any pair of integer $k \geq 2$, any integer $\ell \geq 1$, any protocol $\Pi \in (k,\ell)$-$\mathcal{MLPIR}$, any $\varepsilon > 0$, and all but finitely many $n > 0$, the following holds: $\textsc{Comm}_\Pi(n) \geq (1/\sqrt[\ell+1]{k-1} - \varepsilon)\sqrt[\ell+1]{n}$.

Proof: We show the theorem by contradiction. Then we assume that $\textsc{Comm}_\Pi(n) < (1/\sqrt[\ell+1]{k-1} - \varepsilon)\sqrt[\ell+1]{n}$ for some $k \geq 2$, some $\Pi = (\mathcal{U}; \mathcal{DB}_1, \mathcal{DB}_2, \ldots, \mathcal{DB}_k) \in (k,\ell)$-$\mathcal{MLPIR}$, some $\varepsilon > 0$, and infinitely many $n > 0$. Recall Remark 2.4 and define the single database protocol $\Pi' = (\mathcal{U}'; \mathcal{DB}')$ as follows:

Step 1: $\mathcal{U}'$ simulates $\mathcal{U}$ to generate

$$q_1(i,r), q_2(i,r), \ldots, q_k(i,r)$$

and sends $q_1(i,r)$ to $\mathcal{DB}'$.

Step 2: $\mathcal{DB}'$ simulates $\mathcal{DB}_1$ to generate

$$a_1 = \mathcal{DB}_1(x, q_1(i,r))$$

and sends a_1 to $\mathcal{U}'$.

Step 3: For each $2 \leq j \leq k$, each $h \in [\ell]$, and each $t_{j,h} \in [m_{j,h}]$, $\mathcal{DB}'$ simulates $\mathcal{DB}_j$ to generate

$$\begin{aligned}
&\tilde{a}_j[t_{j,1}, \ldots, t_{j,\ell}] \\
&\quad = \mathcal{DB}_j(x, \langle e_{j,1}[t_{j,1}], \ldots, e_{j,\ell}[t_{j,\ell}]\rangle),
\end{aligned}$$

and sends $\tilde{a}_j[t_{j,1}, \ldots, t_{j,\ell}]$ to $\mathcal{U}'$.

Step 4: For each $2 \leq j \leq k$ and each $h \in [\ell]$, $\mathcal{U}'$ computes $c_{j,h}[1], \ldots, c_{j,h}[m_{j,h}] \in \{0,1\}$ such that

$$q_{j,h}(i,r) = \bigoplus_{t_{j,h}=1}^{m_{j,h}} c_{j,h}[t_{j,h}] e_{j,h}[t_{j,h}].$$

Step 5: For each $2 \leq j \leq k$, $\mathcal{U}'$ computes

$$\begin{aligned}
a'_j = \bigoplus_{t_{j,1}=1}^{m_{j,1}} \cdots \bigoplus_{t_{j,\ell}=1}^{m_{j,\ell}} \{&(c_{j,1}[t_{j,1}] \cdots c_{j,\ell}[t_{j,\ell}]) \\
&\times \tilde{a}_j[t_{j,1}, \ldots, t_{j,\ell}]\}.
\end{aligned}$$

Step 6: $\mathcal{U}'$ outputs $\mathcal{U}(i,r; a_1, a'_2, a'_3, \ldots, a'_k)$.

We show that $\Pi' \in (1,\ell)$-$\mathcal{PIR}$. From the definition of (k,ℓ)-MLPIR, it follows that for each $2 \leq j \leq k$,

$$\begin{aligned}
a_j &= \mathcal{DB}_j\left(x, \langle q_{j,1}(i,r), \ldots, q_{j,\ell}(i,r)\rangle\right) \\
&= \mathcal{DB}_j\left(x, \left\langle \bigoplus_{t_{j,1}=1}^{m_{j,1}} c_{j,1}[t_{j,1}] e_{j,1}[t_{j,1}], \ldots, \right.\right. \\
&\qquad\qquad \left.\left. \bigoplus_{t_{j,\ell}=1}^{m_{j,\ell}} c_{j,\ell}[t_{j,\ell}] e_{j,\ell}[t_{j,\ell}] \right\rangle\right)
\end{aligned}$$

$$
\begin{aligned}
&= \bigoplus_{t_{j,1}=1}^{m_{j,1}} \cdots \bigoplus_{t_{j,\ell}=1}^{m_{j,\ell}} \{(c_{j,1}[t_{j,1}] \cdots c_{j,\ell}[t_{j,\ell}]) \\
&\qquad \times \mathcal{DB}_j\,(x, \langle e_{j,1}[t_{j,1}], \ldots, e_{j,\ell}[t_{j,\ell}]\rangle)\} \\
&= \bigoplus_{t_{j,1}=1}^{m_{j,1}} \cdots \bigoplus_{t_{j,\ell}=1}^{m_{j,\ell}} \{(c_{j,1}[t_{j,1}] \cdots c_{j,\ell}[t_{j,\ell}]) \\
&\qquad\qquad \times \tilde{a}_j[t_{j,1}, \ldots, t_{j,\ell}]\} \\
&= a'_j,
\end{aligned}
$$

where the second equality follows from Step 4, the third equality follows from the multilinearity of the protocol Π (Definition 2.5), the fourth equality follows from Step 3, and the fifth equality follows from Step 5. Thus we have that in Step 6 of the protocol Π',

$$\mathcal{U}(i; a_1, a'_2, \ldots, a'_k) = \mathcal{U}(i; a_1, a_2, \ldots, a_k) = x_i,$$

i.e., $\mathcal{U}'$ retrieves x_i for any $i \in [n]$. Since $\mathcal{U}'$ sends only $q_1(i)$ to $\mathcal{DB}'$ in the protocol Π' and the protocol $\Pi \in (k,\ell)$-$\mathcal{MLPIR}$ satisfies the condition of Definition 2.2, we can conclude that $\Pi' \in (1,\ell)$-$\mathcal{PIR}$.

We estimate the communication complexity of the protocol $\Pi' \in (1,\ell)$-$\mathcal{PIR}$. For the underlying protocol $\Pi \in (k,\ell)$-$\mathcal{MLPIR}$, recall the assumption that

$$\textsc{Comm}_\Pi(n) < (1/\sqrt[\ell+1]{k-1} - \varepsilon)\sqrt[\ell+1]{n}$$

for some $k \geq 2$, some $\varepsilon > 0$, and infinitely many $n > 0$. So we have that

$$
\begin{aligned}
&|q_1(i,r)| < (1/\sqrt[\ell+1]{k-1} - \varepsilon)\sqrt[\ell+1]{n}; \\
&|a_1| < (1/\sqrt[\ell+1]{k-1} - \varepsilon)\sqrt[\ell+1]{n}; \\
&m_{j,h} < (1/\sqrt[\ell+1]{k-1} - \varepsilon)\sqrt[\ell+1]{n}; \\
&|\tilde{a}_j[t_{j,1}, t_{j,2}, \ldots, t_{j,\ell}]| < (1/\sqrt[\ell+1]{k-1} - \varepsilon)\sqrt[\ell+1]{n},
\end{aligned}
$$

for each $2 \leq j \leq k$, each $h \in [\ell]$, and each $t_{j,h} \in [m_{j,h}]$. Then it follows that for infinitely many $n > 0$,

$$
\begin{aligned}
\textsc{Comm}_{\Pi'}(n) &= |q_1(i,r)| + |a_1| \\
&\qquad + \sum_{j=2}^{k} \sum_{t_{j,1}=1}^{m_{j,1}} \cdots \sum_{t_{j,\ell}=1}^{m_{j,\ell}} |\tilde{a}_j[t_{j,1}, \ldots, t_{j,\ell}]| \\
&< 2\left(\frac{1}{\sqrt[\ell+1]{k-1}} - \varepsilon\right)\sqrt[\ell+1]{n} \\
&\qquad + (k-1)\left\{\left(\frac{1}{\sqrt[\ell+1]{k-1}} - \varepsilon\right)\sqrt[\ell+1]{n}\right\}^{\ell+1} \\
&= 2\left(\frac{1}{\sqrt[\ell+1]{k-1}} - \varepsilon\right)\sqrt[\ell+1]{n} \\
&\qquad + (k-1)\left(\frac{1}{\sqrt[\ell+1]{k-1}} - \varepsilon\right)^{\ell+1} n.
\end{aligned}
$$

Since $\varepsilon > 0$, we have that $(k-1)(1/\sqrt[\ell+1]{k-1} - \varepsilon)^{\ell+1} < 1$. So $\textsc{Comm}_{\Pi'}(n) < 2(1/\sqrt[\ell+1]{k-1} - \varepsilon)\sqrt[\ell+1]{n} + (k-1)(1/\sqrt[\ell+1]{k-1} - \varepsilon)^{\ell+1} n < n$ for infinitely many $n > 0$, which contradicts Fact 2.7.

Thus for any integer $k \geq 2$, any integer $\ell \geq 1$, any protocol $\Pi \in (k,\ell)$-$\mathcal{MLPIR}$, and any $\varepsilon > 0$, it follows that $\textsc{Comm}_\Pi(n) \geq (1/\sqrt[\ell+1]{k-1} - \varepsilon)\sqrt[\ell+1]{n}$ for all but finitely many $n > 0$. □

Corollary 3.2 (to Theorem 3.1)**:** For any integer $k \geq 2$, any protocol $\Pi \in k$-$\mathcal{LPIR}$, any $\varepsilon > 0$, and all but finitely many $n > 0$, the following holds: $\textsc{Comm}_\Pi(n) \geq (1/\sqrt{k-1} - \varepsilon)\sqrt{n}$.

Proof: From Definition 2.5, we have that k-$\mathcal{LPIR} = (k,1)$-$\mathcal{MLPIR}$ for any integer $k \geq 2$. Thus the corollary follows from Theorem 3.1 by letting $\ell = 1$. □

3.2 An Upper Bound for Multilinear Type PIR

To evaluate the lower bound for the communication complexity of (k,ℓ)-MLPIR given in Theorem 3.1, we show an upper bound for the communication complexity of (k,ℓ)-MLPIR.

Theorem 3.3 [9, §3.2]**:** For any integer $k \geq 2$, there exist an integer $\ell = (k-1)^2$ and a protocol $\Pi \in (k,\ell)$-$\mathcal{MLPIR}$ such that $\textsc{Comm}_\Pi(n) = O(k^3\sqrt[k]{n})$.

Proof: For any integer $k \geq 2$, Ishai and Kushilevitz [9, §3.2] showed the protocol IK[1] $\in (k,\ell)$-$\mathcal{PIR}$ such that $\ell = (k-1)^2$ and $\textsc{Comm}_{\text{IK}[1]}(n) = O(k^3\sqrt[k]{n})$. We note that the protocol IK[1] is based on the multilinear function $\prod$ (see [9]) and the balancing scheme [3]. So from the definition of (k,ℓ)-MLPIR (Definition 2.5), it is easy to verify that IK[1] $\in (k,\ell)$-$\mathcal{MLPIR}$. □

It follows from Theorem 3.1 that for any integer $k \geq 2$, any protocol $\Pi \in (k,(k-1)^2)$-$\mathcal{MLPIR}$, any $\varepsilon > 0$, and all but finitely many $n > 0$,

$$\textsc{Comm}_\Pi(n) \geq \left(\frac{1}{\sqrt[(k-1)^2+1]{k-1}} - \varepsilon\right)\sqrt[(k-1)^2+1]{n}. \quad (1)$$

From Theorem 3.3 and Eq. (1), it is obvious that we still have a gap between the lower and upper bounds for the communication complexity of (k,ℓ)-MLPIR.

Table 1 exemplifies the upper and lower bounds for the communication complexity of (k,ℓ)-MLPIR for each $k \geq 2$, where $\ell = (k-1)^2$.

To evaluate the lower bound for the communication complexity of k-LPIR given in Corollary 3.2, we also show a upper bound for the communication complexity of k-LPIR (especially for $k = 2$).

Table 1 Communication complexity of (k,ℓ)-MLPIR.

Number of Databases: k	Upper Bound	Lower Bound
2	$O(\sqrt{n})$	$\Omega(\sqrt{n})$
3	$O(\sqrt[3]{n})$	$\Omega(\sqrt[5]{n})$
4	$O(\sqrt[4]{n})$	$\Omega(\sqrt[10]{n})$
5	$O(\sqrt[5]{n})$	$\Omega(\sqrt[17]{n})$
⋮	⋮	⋮

Table 2 Communication complexity of k-LPIR.

Number of Databases: k	Upper Bound	Lower Bound
2	$O(\sqrt{n})$	$\Omega(\sqrt{n})$
3	$O(\sqrt{n})$	$\Omega(\sqrt{n})$
4	$O(\sqrt{n})$	$\Omega(\sqrt{n})$
5	$O(\sqrt{n})$	$\Omega(\sqrt{n})$
$\vdots$	$\vdots$	$\vdots$

Theorem 3.4: For all $n > 0$, there exists a protocol $\Pi \in 2\text{-}\mathcal{LPIR}$ such that $\textsc{Comm}_\Pi(n) = 4\sqrt{n}$.

Proof: For any $S \subseteq [m]$ and any $i \in [m]$, define $S \oplus i = S - \{i\}$ if $i \in S$; $S \oplus i = S \cup \{i\}$ otherwise. Let $s = \sqrt{n}$. For any $x \in \{0,1\}^n$, regard x as an $s \times s$ matrix $X = (x_1, x_2, \ldots, x_s)$ over $GF(2)$, where $x_j \in \{0,1\}^s$ is the column vector over $GF(2)$ for each $j \in [s]$. Let $i \in [n]$ be the bit position of x that $\mathcal{U}$ tries to get and $(i_1, i_2) \in [s] \times [s]$ be the bit position in the matrix X corresponding to $i \in [n]$. Combining the protocol $\Pi' \in (2,\ell)\text{-}\mathcal{PIR}$ given by Chor et al. [4, §3.1] and the balancing scheme [3], we define the two database protocol $\Pi = (\mathcal{U}; \mathcal{DB}_1, \mathcal{DB}_2)$ as follows:

Step 1: $\mathcal{U}$ chooses $S_1 \subseteq [s]$ uniformly at random and computes $S_2 = S_1 \oplus i_2$.

Step 2: $\mathcal{U}$ sends $S_1 \subseteq [s]$ to $\mathcal{DB}_1$ and $S_2 \subseteq [s]$ to $\mathcal{DB}_2$.

Step 3: $\mathcal{DB}_1$ computes $a_1 = \bigoplus_{h \in S_1} x_h$ and sends $a_1 \in \{0,1\}^s$ to $\mathcal{U}$.

Step 4: $\mathcal{DB}_2$ computes $a_2 = \bigoplus_{h \in S_2} x_h$ and sends $a_2 \in \{0,1\}^s$ to $\mathcal{U}$.

Step 5: $\mathcal{U}$ computes $a = a_1 \oplus a_2$ and picks the $i_1 th$ position of the column vector $a \in \{0,1\}^s$.

It is easy to verify that $\Pi = (\mathcal{U}; \mathcal{DB}_1, \mathcal{DB}_2) \in 2\text{-}\mathcal{LPIR}$ and that $\textsc{Comm}_\Pi(n) = 4\sqrt{n}$. □

From Theorem 3.4, it follows that for any integer $k \geq 2$, the lower bound for the communication complexity of k-LPIR given in Corollary 3.2 is *optimal* up to constant factor. Table 2 exemplifies the upper and lower bounds for the communication complexity of k-LPIR for each $k \geq 2$.

4. Communication Complexity of Affine Type PIR

4.1 A Lower Bound for Affine Type PIR

In this subsection, we derive a lower bound for the communication complexity of (k,ℓ)-APIR for any integer $k \geq 2$ and any integer $\ell \geq 1$ in almost the same way as the proof of Theorem 3.1.

Recall Remark 2.4. For each $j \in [k]$ and each $h \in [\ell]$, the coefficients $c_{j,h}[0], c_{j,h}[1], \ldots, c_{j,h}[m_{j,h}] \in \{0,1\}$ are said to be *good* for $q_{j,h}(i,r)$ if

$$\bigoplus_{t_{j,h}=0}^{m_{j,h}} c_{j,h}[t_{j,h}] = 1;$$

$$q_{j,h}(i,r) = \sum_{t_{j,h}=0}^{m_{j,h}} c_{j,h}[t_{j,h}] e_{j,h}[t_{j,h}].$$

It is immediate that there exist the *unique* good coefficients $c_{j,h}[0], c_{j,h}[1], \ldots, c_{j,h}[m_{j,h}] \in \{0,1\}$ for $q_{j,h}(i,r)$, because $e_{j,h}[0] = 0^{m_{j,h}}$ and $\{e_{j,h}[1], \ldots, e_{j,h}[m_{j,h}]\}$ is the basis for the linear space $\{0,1\}^{m_{j,h}}$.

Lemma 4.1: For each $j \in [k]$, each $h \in [\ell]$, and each $i \in [n]$, let $c_{j,h}[0], c_{j,h}[1], \ldots, c_{j,h}[m_{j,h}] \in \{0,1\}$ be the good coefficients for $q_{j,h}(i,r)$. Then for each $j \in [k]$ and each $h \in [\ell]$, the following holds:

$$\mathcal{DB}_j\Bigg(x, \Bigg\langle q_{j,1}(i,r), \ldots, q_{j,h-1}(i,r), \bigoplus_{t_{j,h}=0}^{m_{j,h}} c_{j,h}[t_{j,h}] e_{j,h}[t_{j,h}], q_{j,h+1}(i,r), \ldots, q_{j,\ell}(i,r) \Bigg\rangle\Bigg)$$

$$= \bigoplus_{t_{j,h}=0}^{m_{j,h}} c_{j,h}[t_{j,h}] \mathcal{DB}_j(x, \langle q_{j,1}(i,r), \ldots, q_{j,h-1}(i,r), e_{j,h}[t_{j,h}], q_{j,h+1}(i,r), \ldots, q_{j,\ell}(i,r)\rangle).$$

Proof: The lemma follows from Definition 2.6 and the condition that $\bigoplus_{t_{j,h}=0}^{m_{j,h}} c_{j,h}[t_{j,h}] = 1$. □

Theorem 4.2: For any integer $k \geq 2$, any integer $\ell \geq 1$, any protocol $\Pi \in (k,\ell)\text{-}\mathcal{APIR}$, any $\varepsilon > 0$, and all but finitely many $n > 0$, the following holds: $\textsc{Comm}_\Pi(n) \geq (1/\sqrt[\ell+1]{k-1} - \varepsilon) \cdot \sqrt[\ell+1]{n}$.

Proof: In a way similar to the proof of Theorem 3.1, we show the theorem by contradiction. Assume that $\textsc{Comm}_\Pi(n) < (1/\sqrt[\ell+1]{k-1} - \varepsilon)\sqrt[\ell+1]{n}$ for some $k \geq 2$, some protocol $\Pi = (\mathcal{U}; \mathcal{DB}_1, \mathcal{DB}_2, \ldots, \mathcal{DB}_k) \in (k,\ell)\text{-}\mathcal{APIR}$, some $\varepsilon > 0$, and infinitely many $n > 0$. We notice that (k,ℓ)-APIR is the same as (k,ℓ)-MLPIR except for the third additive term in Definition 2.6. To construct single database protocol $\Pi' = (\mathcal{U}', \mathcal{DB}')$ as in the proof of Theorem 3.1, we need to remove those third additive terms. So we employ $e_{j,h}[0] = 0^{m_{j,h}}$ to apply Lemma 4.1 and define the following single database protocol $\Pi' = (\mathcal{U}'; \mathcal{DB}')$:

Step 1: $\mathcal{U}'$ simulates $\mathcal{U}$ to generate

$$q_1(i,r), q_2(i,r), \ldots, q_k(i,r)$$

and sends $q_1(i,r)$ to $\mathcal{DB}'$.

Step 2: $\mathcal{DB}'$ simulates $\mathcal{DB}_1$ to generate

$$a_1 = \mathcal{DB}_1(x, q_1(i,r))$$

and sends a_1 to $\mathcal{U}'$.

Step 3: For each $2 \le j \le k$, each $h \in [\ell]$, and each $0 \le t_{j,h} \le m_{j,h}$, $\mathcal{DB}'$ simulates $\mathcal{DB}_j$ to generate

$$\tilde{a}_j[t_{j,1}, \ldots, t_{j,\ell}] = \mathcal{DB}_j(x, \langle e_{j,1}[t_{j,1}], \ldots, e_{j,\ell}[t_{j,\ell}]\rangle),$$

and sends $\tilde{a}_j[t_{j,1}, \ldots, t_{j,\ell}]$ to $\mathcal{U}'$.

Step 4: For each $2 \le j \le k$ and each $h \in [\ell]$, $\mathcal{U}'$ computes the good coefficients $c_{j,h}[0], \ldots, c_{j,h}[m_{j,h}]$ for $q_{j,h}(i,r) \in \{0,1\}^{m_{j,h}}$.

Step 5: For each $2 \le j \le k$, $\mathcal{U}'$ computes

$$a'_j = \bigoplus_{t_{j,1}=0}^{m_{j,1}} \cdots \bigoplus_{t_{j,\ell}=0}^{m_{j,\ell}} \{(c_{j,1}[t_{j,1}] \cdots c_{j,\ell}[t_{j,\ell}]) \times \tilde{a}_j[t_{j,1}, \ldots, t_{j,\ell}]\}.$$

Step 6: $\mathcal{U}'$ outputs $\mathcal{U}(i, r; a_1, a'_2, a'_3, \ldots, a'_k)$.

We show that $\Pi' \in (1,\ell)$-$\mathcal{PIR}$. From the definition of (k,ℓ)-APIR, it follows that for each $2 \le j \le k$,

$$\begin{aligned}
a_j &= \mathcal{DB}_j\left(x, \langle q_{j,1}(i,r), \ldots, q_{j,\ell}(i,r)\rangle\right) \\
&= \mathcal{DB}_j\left(x, \left\langle \bigoplus_{t_{j,1}=0}^{m_{j,1}} c_{j,1}[t_{j,1}]e_{j,1}[t_{j,1}], \ldots, \bigoplus_{t_{j,\ell}=0}^{m_{j,\ell}} c_{j,\ell}[t_{j,\ell}]e_{j,\ell}[t_{j,\ell}] \right\rangle\right) \\
&= \bigoplus_{t_{j,1}=0}^{m_{j,1}} \cdots \bigoplus_{t_{j,\ell}=0}^{m_{j,\ell}} \{(c_{j,1}[t_{j,1}] \cdots c_{j,\ell}[t_{j,\ell}]) \times \mathcal{DB}_j\left(x, \langle e_{j,1}[t_{j,1}], \ldots, e_{j,\ell}[t_{j,\ell}]\rangle\right) \\
&= \bigoplus_{t_{j,1}=0}^{m_{j,1}} \cdots \bigoplus_{t_{j,\ell}=0}^{m_{j,\ell}} \{(c_{j,1}[t_{j,1}] \cdots c_{j,\ell}[t_{j,\ell}]) \times \tilde{a}_j[t_{j,1}, t_{j,2}, \ldots, t_{j,\ell}] \\
&= a'_j,
\end{aligned}$$

where the second equality follows from Step 4, the third equality follows from Lemma 4.1, the fourth equality follows from Step 3, and the fifth equality follows from Step 5. Thus in Step 6 of the protocol Π', we have that $\mathcal{U}(i; a_1, a'_2, a'_3, \ldots, a'_k) = \mathcal{U}(i; a_1, a_2, a_3, \ldots, a_k) = x_i$, i.e., $\mathcal{U}'$ retrieves x_i for any $i \in [n]$. Since $\mathcal{U}'$ sends only $q_1(i)$ to $\mathcal{DB}'$ in the protocol Π' and the protocol $\Pi \in (k,\ell)$-$\mathcal{APIR}$ satisfies the condition of Definition 2.2, we can conclude that $\Pi' \in (1,\ell)$-$\mathcal{PIR}$.

We estimate the communication complexity of the protocol $\Pi' \in (1,\ell)$-$\mathcal{PIR}$. For the underlying protocol $\Pi \in (k,\ell)$-$\mathcal{APIR}$, recall the assumption that

$$\textsc{Comm}_\Pi(n) < (1/\sqrt[\ell+1]{k-1} - \varepsilon)\sqrt[\ell+1]{n}$$

for some $k \ge 2$, some $\varepsilon > 0$, and infinitely many $n > 0$. So we have that

$$\begin{aligned}
&|q_1(i,r)| < (1/\sqrt[\ell+1]{k-1} - \varepsilon)\sqrt[\ell+1]{n}; \\
&|a_1| < (1/\sqrt[\ell+1]{k-1} - \varepsilon)\sqrt[\ell+1]{n}; \\
&m_{j,h} < (1/\sqrt[\ell+1]{k-1} - \varepsilon)\sqrt[\ell+1]{n}; \\
&|\tilde{a}_j[t_{j,1}, t_{j,2}, \ldots, t_{j,\ell}]| < (1/\sqrt[\ell+1]{k-1} - \varepsilon)\sqrt[\ell+1]{n},
\end{aligned}$$

for each $2 \le j \le k$, each $h \in [\ell]$, and each $0 \le t_{j,h} \le m_{j,h}$. It follows that for infinitely many $n > 0$,

$$\begin{aligned}
\textsc{Comm}_{\Pi'}(n) &= |q_1(i)| + |a_1| + \sum_{j=2}^{k} \sum_{t_1=0}^{m_{j1}} \cdots \sum_{t_\ell=0}^{m_{j\ell}} |\tilde{a}_j[t_{j,1}, \ldots, t_{j,\ell}]| \\
&< 2\left(\frac{1}{\sqrt[\ell+1]{k-1}} - \varepsilon\right)\sqrt[\ell+1]{n} \\
&\quad + (k-1)\left\{\left(\frac{1}{\sqrt[\ell+1]{k-1}} - \varepsilon\right)\sqrt[\ell+1]{n} + 1\right\}^{\ell} \times \left(\frac{1}{\sqrt[\ell+1]{k-1}} - \varepsilon\right)\sqrt[\ell+1]{n} \\
&< 2\left(\frac{1}{\sqrt[\ell+1]{k-1}} - \varepsilon\right)\sqrt[\ell+1]{n} \\
&\quad + (k-1)\left(\frac{1}{\sqrt[\ell+1]{k-1}} - \frac{\varepsilon}{2}\right)^{\ell+1} n.
\end{aligned}$$

Since $\varepsilon > 0$, it is immediate that $(k-1)(1/\sqrt[\ell+1]{k-1} - \varepsilon/2)^{\ell+1} < 1$. So $\textsc{Comm}_{\Pi'}(n) < 2(1/\sqrt[\ell+1]{k-1} - \varepsilon)\sqrt[\ell+1]{n} + (k-1)(1/\sqrt[\ell+1]{k-1} - \varepsilon/2)^{\ell+1} n < n$ for infinitely many $n > 0$, which contradicts Fact 2.7.

Thus for any integer $k \ge 2$, any integer $\ell \ge 1$, any protocol $\Pi \in (k,\ell)$-$\mathcal{APIR}$, and any $\varepsilon > 0$, it follows that $\textsc{Comm}_\Pi(n) \ge (1/\sqrt[\ell+1]{k-1} - \varepsilon)\sqrt[\ell+1]{n}$ for all but finitely many $n > 0$. □

4.2 An Upper Bound for Affine Type PIR

To evaluate the lower bound for the communication complexity of (k,ℓ)-APIR given in Theorem 4.2, we show an upper bound for the communication complexity of (k,ℓ)-APIR.

Theorem 4.3 [9, §3.3]**:** For any integer $k \ge 2$, there exist an integer $\ell = (2k-1)(k-1)$ and a protocol $\Pi \in (k,\ell)$-$\mathcal{APIR}$ such that $\textsc{Comm}_\Pi(n) = O(k^3\sqrt[2k-1]{n})$.

Proof: For any integer $k \ge 2$, Ishai and Kushilevitz [9, §3.3] showed the protocol IK[2] $\in (k,\ell)$-$\mathcal{PIR}$ such that $\ell = (2k-1)(k-1)$ and $\textsc{Comm}_{\text{IK[2]}}(n) = O(k^3\sqrt[2k-1]{n})$. Note that the protocol IK[2] is based on the multilinear function $\prod$ (see [9] for the definition) and the idea similar to the covering code scheme [4]. From the definition of (k,ℓ)-APIR (Definition 2.6), it is obvious that IK[2] $\in (k,\ell)$-$\mathcal{APIR}$. □

It follows from Theorem 4.2 that for any integer $k \ge 2$, any protocol $\Pi \in (k,(2k-1)(k-1))$-$\mathcal{APIR}$, any $\varepsilon > 0$, and all but finitely many $n > 0$,

Table 3 Communication complexity of (k, ℓ)-APIR.

Number of Databases: k	Upper Bound	Lower Bound
2	$O(\sqrt[3]{n})$	$\Omega(\sqrt[4]{n})$
3	$O(\sqrt[5]{n})$	$\Omega(\sqrt[11]{n})$
4	$O(\sqrt[7]{n})$	$\Omega(\sqrt[22]{n})$
5	$O(\sqrt[9]{n})$	$\Omega(\sqrt[37]{n})$
$\vdots$	$\vdots$	$\vdots$

$$\begin{aligned}&\textsc{Comm}_{\Pi}(n)\\ &\geq \left(\frac{1}{\sqrt[(2k-1)(k-1)+1]{k-1}} - \varepsilon\right)\sqrt[(2k-1)(k-1)+1]{n}. \qquad (2)\end{aligned}$$

From Theorem 4.3 and Eq. (2), it is obvious that we still have a gap between the lower and upper bounds for the communication complexity of (k, ℓ)-APIR.

Table 3 exemplifies the upper and lower bounds for the communication complexity of (k, ℓ)-APIR for each $k \geq 2$, where $\ell = (2k-1)(k-1)$.

5. Concluding Remarks

In this paper, we have classified (k, ℓ)-PIR into a linear type, an multilinear type, and an affine type and then we have shown the lower bounds for the communication complexity of each type, i.e.,

(1) for any protocol $\Pi \in (k, \ell)$-$\mathcal{MLPIR}$, $\textsc{Comm}_{\Pi}(n) \geq \Omega(\sqrt[\ell+1]{n})$ (Theorem 3.1);

(2) for any protocol $\Pi \in k$-$\mathcal{LPIR}$, $\textsc{Comm}_{\Pi}(n) \geq \Omega(\sqrt{n})$ (Corollary 3.2);

(3) for any protocol $\Pi \in (k, \ell)$-$\mathcal{APIR}$, $\textsc{Comm}_{\Pi}(n) \geq \Omega(\sqrt[\ell+1]{n})$ (Theorem 4.2).

To evaluate those lower bounds, we have also shown that for any integer $k \geq 2$,

(4) there exist $\Pi \in (k, (k-1)^2)$-$\mathcal{MLPIR}$ with $\textsc{Comm}_{\Pi}(n) = O(k^3\sqrt[k]{n})$ (Theorem 3.3);

(5) there exists $\Pi \in 2$-$\mathcal{LPIR}$ with $\textsc{Comm}_{\Pi}(n) = 4\sqrt{n}$ (Theorem 3.4);

(6) there exists $\Pi \in (k, (2k-1)(k-1))$-$\mathcal{APIR}$ with $\textsc{Comm}_{\Pi}(n) = O(k^3\sqrt[2k-1]{n})$ (Theorem 4.3).

Note that the protocols given in Theorems 3.3, 3.4, and 4.3 are communication-efficient, but do not satisfy the *data privacy* [6]. To achieve the data privacy, we need to slightly modify the present model into the *shared randomness* model. Gertner et al. [6] proposed the multiparty protocol that conditional discloses secrets (CDS protocol) to transform specific (k, ℓ)-PIR into (k, ℓ')-PIR with the data privacy under the shared randomness model. In a way similar to the transformation by Gertner et al. [6], we can show the theorems below applying the CDS protocol to the protocols in Theorems 3.3, 3.4, and 4.3.

Theorem 5.1: For any integer $k \geq 2$, the protocol Π given in Theorem 3.3 can be transformed into the protocol Π' with user privacy and data privacy such that $\textsc{Comm}_{\Pi'}(n) = O(k^3\sqrt[k]{n}\log n)$.

Theorem 5.2: The protocol Π given in Theorem 3.4 can be transformed into the protocol Π' with user privacy and data privacy such that $\textsc{Comm}_{\Pi'}(n) = O(n \log n)$.

Theorem 5.3: For any integer $k \geq 2$, the protocol Π given in Theorem 4.3 can be transformed into the protocol Π' with user privacy and data privacy such that $\textsc{Comm}_{\Pi'}(n) = O(k^3\sqrt[2k-1]{n}\log n)$.

As we have mentioned at the end of Sect. 3.2, the communication complexity of the protocol $\Pi \in 2$-$\mathcal{LPIR}$ given in Theorem 3.4 is *optimal* up to constant factor, but we still have large gaps between the lower and upper bounds for the communication complexity of (k, ℓ)-MLPIR and (k, ℓ)-APIR. So

(1) Improve the lower bound for the communication complexity of $\Pi \in (k, \ell)$-$\mathcal{MLPIR}$;

(2) Improve the lower bound for the communication complexity of $\Pi \in (k, \ell)$-$\mathcal{APIR}$;

(3) Improve the upper bound for the communication complexity of $\Pi \in (k, \ell)$-$\mathcal{MLPIR}$;

(4) Improve the upper bound for the communication complexity of $\Pi \in (k, \ell)$-$\mathcal{APIR}$.

References

[1] A. Ambainis, "Upper bound on the communication complexity of private information retrieval," Proc. 24th International Colloquium on Automata, Languages, and Programming, Lecture Notes in Computer Science 1256, pp.401–409, 1997.

[2] M. Ben-Or, S. Goldwasser, J. Kilian, and A. Wigderson, "Multi-prover interactive proofs: How to remove intractability assumptions," Proc. 20th ACM Symposium on Theory of Computing, pp.113–131, 1988.

[3] B. Chor and N. Gilboa, "Computationally private information retrieval," Proc. 29th Annual ACM Symposium on Theory of Computing, pp.304–313, 1997.

[4] B. Chor, O. Goldreich, E. Kushilevitz, and M. Sudan, "Private information retrieval," J. Assoc. Comput. Mach., vol.45, pp.965–982, 1998.

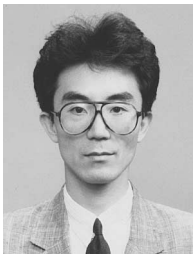
[5] C. Cachin, S. Micali, and M. Stadler, "Computationally private information retrieval with polylogarithmic communication," Proc. Eurocrypt, Lecture Notes in Computer Science 1592, pp.402–414, 1999.

[6] Y. Gertner, Y. Ishai, E. Kushilevitz, and T. Malkin, "Protecting data privacy in private information retrieval schemes," Proc. 30th Annual ACM Symposium on Theory of Computing, pp.151–160, 1998.

[7] S. Goldwasser and S. Micali, "Probabilistic encryption," J. Comput. Sys. Sci., vol.28, pp.270–299, 1984.

[8] J. Håstad, "Pseudo-random generators with uniform assumptions," Proc. 22nd ACM Symposium on Theory of Computing, pp.395–404, 1990.

[9] Y. Ishai and E. Kushilevitz, "Improved upper bounds on information-theoretic private information retrieval," Proc. 31st Annual ACM Symposium on Theory of Computing, pp.79–88, 1999.
[10] R. Impagliazzo, L.A. Levin, and M. Luby, "Pseudo-random generation from one-way functions," Proc. 30th IEEE Symposium on Foundations of Computer Science, pp.12–24, 1989.
[11] E. Kushilevitz and R. Ostrovsky, "Replication is not needed: Single database, computationally-private information retrieval," Proc. 38th Annual IEEE Symposium on Foundations of Computer Science, pp.364–373, 1997.
[12] E. Kushilevitz and R. Ostrovsky, "One-way trapdoor permutations are sufficient for non-trivial single-server private information retrieval," Proc. Eurocrypt, Lecture Notes in Computer Science 1807, pp.104–121, 2000.
[13] F.J. MacWilliams and N.J.A. Sloane, The Theory of Error-Correcting Codes, 9th edition, North-Holland, 1996.



Toshiya Itoh was born in Urawa, Japan, in 1959. He received the B.Eng., the M.S. Eng., and the Dr. Eng. degree in electronic engineering in 1982, 1984, and 1988, respectively from Tokyo Institute of Technology, Tokyo, Japan. From 1985 to 1990, he was an Assistant Professor in the Department of Electrical and Electronic Engineering at Tokyo Institute of Technology, and from 1990 to 1992, he was a Lecturer in the Department of Information Processing at Tokyo Institute of Technology. Since 1992, he has been an Associate Professor in the Department of Information Processing at Tokyo Institute of Technology. His current interests are modern cryptographies, finite field arithmetics, and complexity theory. Dr. Itoh is a member of the Information Processing Society of Japan, the International Association for Cryptologic Research, the Association for Computing Machinery, and LA.