Optimal Mechanism for Randomized Responses under Universally Composable Security Measure

Yuuya Yoshida\textsuperscript{a}, Man-Hong Yung\textsuperscript{b,c,d}, and Masahito Hayashi\textsuperscript{a,b}

\textsuperscript{a}Graduate School of Mathematics, Nagoya University
\textsuperscript{b}Shenzhen Institute for Quantum Science and Engineering, Southern University of Science and Technology
\textsuperscript{c}Department of Physics, Southern University of Science and Technology
\textsuperscript{d}Shenzhen Key Laboratory of Quantum Science and Engineering

Abstract—We consider a problem of analyzing a global property of private data through randomized responses subject to a certain rule, where private data are used for another cryptographic protocol, e.g., authentication. For this problem, the security of private data was evaluated by a universally composable security measure, which can be regarded as $(0, \delta)$-differential privacy. Here we focus on the trade-off between the global accuracy and a universally composable security measure, and derive an optimal solution to the trade-off problem. More precisely, we adopt the Fisher information of a certain distribution family as the estimation accuracy of a global property and impose $(0, \delta)$-differential privacy on a randomization mechanism protecting private data. Finally, we maximize the Fisher information under the $(0, \delta)$-differential privacy constraint and obtain an optimal mechanism explicitly.

Index Terms—universally composable security measure, $(0, \delta)$-differential privacy, $l^1$-norm, Fisher information, parameter estimation, sublinear function

I. INTRODUCTION

For many applications, it is of great interest in estimating a global property of an ensemble while protecting individual privacy. In these scenarios, disclosed data are randomized to protect individual privacy, but it causes uncertainty to the global property. This fact implies a trade-off between global accuracy and individual privacy. This paper focuses on a scenario of answering YES/NO question, where the goal is to estimate the number of YES’s (labeled as “1”) or NO’s (labeled as “0”) with high accuracy, given that respondents randomize their responses to protect individual privacy.

More specifically, we assume that an investigator is interested only in the ratio $1 - \theta : \theta$ of the binary private data, where $\theta$ is a real number between zero and one. The investigator randomly chooses $n$ individuals to ask for their private data. If selected individuals directly send their private data $X \in \mathcal{X} := \{0, 1\}$ to the investigator, then their private data could be completely leaked to the investigator. To protect individual privacy, we use the following scheme [1], [2]: when private data is $X = 0$, the individual generates a disclosed data $Y \in \mathcal{Y}$ subject to a distribution $p_0$ on a probability space $\mathcal{Y}$ and sends it to the investigator. In the following, the sets $\mathcal{X}$ and $\mathcal{Y}$ are called the private data set and the disclosed data set, and $|S|$ denotes the number of elements of a set $S$. When private data is $X = 1$, the individual generates a disclosed data $Y$ subject to another distribution $p_1$ on $\mathcal{Y}$ and sends it to the investigator.

To analyze estimation of the ratio $1 - \theta : \theta$, we introduce the parameterized distribution $p_\theta$ defined as $p_\theta = (1 - \theta)p_0 + \theta p_1$. In this way, the estimation of the ratio $1 - \theta : \theta$ is reduced to the estimation of the parameter $\theta \in [0, 1]$ of the distribution family $\{p_\theta\}_{\theta \in [0, 1]}$ when $n$ data are generated from the same unknown distribution $p_\theta$ independently, as we allow duplication in the selection of individuals. In literatures [3]–[5] of statistical parameter estimation, it is well-known that an optimal estimator is given as the maximum likelihood estimator (MLE) with respect to sufficiently large $n$ data. The error of the MLE is asymptotically characterized by the inverse of the Fisher information $I_\theta(Y)$, which is called the Cramér-Rao bound. Thus we adopt the Fisher information $I_\theta(Y)$ as the estimation accuracy.

In addition to the above scenario, it is natural to use private data as resources for another cryptographic protocol like an authentication protocol [6]. In fact, we often use private data, e.g., birthday, to identify an individual. In this case, we need to guarantee the security of the whole protocol. That is, if a part of information of private data is leaked, we need to consider its effect to the cryptographic protocol that uses private data. In the cryptography community, to evaluate the security of the whole protocol, a security measure based on the $l^1$-norm is proposed as a universally composable (UC) security measure [7]. If the UC-security measure of the first protocol equals $\delta$ and that of the second protocol equals $\delta'$, then that of the combined protocol is upper bounded by $\delta + \delta'$. This property is called universal compositability. Thanks to this property, the $l^1$-norm is widely accepted as a security measure in the communities of cryptography and information-theoretic security [9]–[11]. Thus, when using private data for

MHY was supported by National Natural ScienceFoundation of China (Grants No. 11405093), Natural Science Foundation of Guangdong Province (2017B030308003), the Guangdong Innovative and Entrepreneurial Research Team Program (Grant No. 2016ZT06D348), and the Science Technology and Innovation Commission of Shenzhen Municipality (Grants No. ZDSYS20170303165926217 and No. JCYJ201701041215260376). MH was supported in part by Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for Scientific Research (A) No. 17H01280 and (B) No. 16K10017, Fund for the Promotion of Joint International Research Grant No. 15KK0007, and Kayamori Foundation of Informational Science Advancement.
a cryptographic protocol, we need to guarantee that the UC-security measure is upper bounded by a certain threshold. On the other hand, as a privacy measure, Kairouz et al. [12] focused on $(\epsilon, 0)$-differential privacy and maximized the Fisher information $J_\theta$ under the $(\epsilon, 0)$-differential privacy constraint.

**Differential privacy (DP)** is a standard privacy measure that is widely accepted and introduced by [13] and [14]. (For the definition, see (2) in Section II.) However, $(\epsilon, 0)$-differential privacy does not have the universally composable property for private data. Therefore, we need to address the trade-off between the Fisher information and the UC-security measure.

In our setting, the UC-security measure is given as the variational distance $d_1(p_0, p_1) := (1/2)||p_0 - p_1||_1$ between two distributions $p_0$ and $p_1$, where $||\cdot||_1$ denotes the $l^1$-norm. The variational distance also has the following meaning: if an adversary tries to distinguish private date, the minimum value of the average error probability equals $\min_{\mathcal{S} \subseteq \mathcal{Y}} \left( \frac{1}{2} p_0(S) + \frac{1}{2} p_1(S^c) \right) = (1/2)(1 - d_1(p_0, p_1))$, where $S^c$ denotes the complement of $S$. Fortunately, it can be regarded as $(0, \delta)$-differential privacy. Thus we can also say that we maximize the Fisher information $J_\theta$ under the $(0, \delta)$-differential privacy constraint.

Further, to address the above maximization, we encounter a new aspect that never appeared in preceding studies for the trade-off between the Fisher information and $(\epsilon, 0)$-differential privacy. Kairouz et al. [12] maximized (non-explicitly) the Fisher information under the $(\epsilon, 0)$-differential privacy constraint when $|\mathcal{X}|, |\mathcal{Y}| \geq 2$. Then they showed that the maximization under the $(\epsilon, 0)$-differential privacy constraint achieves the maximum value when $|\mathcal{X}| = |\mathcal{Y}| = 2$. Holohan et al. [15] considered the maximization under the $(\epsilon, \delta)$-differential privacy constraint. However, they assumed $|\mathcal{X}| = |\mathcal{Y}| = 2$. Therefore, this kind of maximization has been open for a general disclosed data set $\mathcal{Y}$ even if $|\mathcal{X}| = 2$. To find an optimal mechanism in our framework, we need to maximize the Fisher information for a general disclosed data set $\mathcal{Y}$. In fact, we can show that the maximization under the $(0, \delta)$-differential privacy constraint achieves the maximum value only when $|\mathcal{Y}| > |\mathcal{X}|$. As a result, we obtain a randomized response scheme with $|\mathcal{X}| = 2$ and $|\mathcal{Y}| = 3$ that has been never obtained in preceding studies. Our optimal solution is completely different from those of [12] and [15]. To handle the case with $|\mathcal{Y}| > |\mathcal{X}|$, we need to address an additional case that is more complicated than the case with $|\mathcal{X}| = |\mathcal{Y}|$. Table I summarizes the relation among [12], [15] and this paper.

The remaining part of this paper is organized as follows.

**TABLE I**

| Relation with preceding studies | Constraint | $|\mathcal{X}|$ | $|\mathcal{Y}|$ | Condition for optimality |
|---------------------------------|------------|----------------|----------------|--------------------------|
| Kairouz et al. [12]             | $(\epsilon, 0)$-DP | $\geq 2$ | $\geq 2$ | $|\mathcal{X}| = |\mathcal{Y}|$ |
| Holohan et al. [15]             | $(\epsilon, 0)$-DP | $2$ | $2$ | $|\mathcal{Y}| \geq 3$ |
| This paper                      | $(0, \delta)$-DP | $2$ | $2$ | $|\mathcal{Y}| \geq 3$ |

Section II states the formulation of our problem and the maximum Fisher information under the $l^1$-norm constraint, which describes that the maximum Fisher information depends on the weight $w$ and the parameter $\theta$. Section III proves our theorem on the maximum Fisher information by solving an optimization with a general sublinear function. Section IV explains the relation among preceding studies and this paper. Section V is devoted to concluding remarks. The full version of this paper is available at arXiv:1805.06278.

**II. OPTIMAL ESTIMATION**

According to Fig. 1, we describe a scheme to estimate the parameter $\theta$. In our scheme, we first fix two distributions $p_0$ and $p_1$ on a finite probability space $\mathcal{Y}$, which describes a conversion rule from private data $X_i \in \mathcal{X} = \{0, 1\}$ to disclosed data $Y_i \in \mathcal{Y}$. Assume that private data $X_1, \ldots, X_n \in \mathcal{X}$ are independent and subject to the binary distribution parametrized by the parameter $\theta \in (0, 1)$. That is, the true probability of $X_i = 0$ is $1 - \theta$ and that of $X_i = 1$ is $\theta$, where $X_i$ denotes the $i$-th individual’s private data. The $i$-th individual generates a disclosed data $Y_i$ subject to $p_\theta$, dependent on the value $X_i = x$ and then sends $Y_i$ to the investigator. From the investigator’s viewpoint, the disclosed data $Y_i$ is given by the distribution $p_\theta := (1 - \theta)p_0 + \theta p_1$.

Next, the parameter $\theta$ is estimated by the investigator in the following way. The investigator observes $n$ disclosed data $Y_1, \ldots, Y_n$ and employs the MLE $\hat{\theta}_n := \arg \max_{\theta \in [0, 1]} \sum_{i=1}^n \ln p_\theta(Y_i)$, whose asymptotic optimality is well-known [3]–[5]. That is, if $n$ is sufficiently large, the MLE $\hat{\theta}_n$ behaves approximately as $\theta + (nJ_\theta)^{-1/2} Z$, where $Z$ is a random variable subject to the standard Gaussian distribution and $J_\theta$ is the Fisher information of the distribution family $\{p_\theta\}_{\theta \in [0, 1]}$: $J_\theta = \sum_{\theta \in [0, 1]} \left( (d/d\theta) \ln p_\theta(y) \right)^2 p_\theta(y)$. Therefore, the mean square error behaves as $1/nJ_\theta$. For example, when we require the confidence level to be $\alpha$, the confidence interval is approximately given as $\hat{\theta}_n + (nJ_\theta)^{-1/2} [\Phi^{-1}(\alpha/2), \Phi^{-1}(1 - \alpha/2)]$, where $\Phi(y) := (2\pi)^{-1/2} \int_{-\infty}^{\infty} e^{-x^2/2} dx$. In this sense, we can conclude that the Fisher information $J_\theta$ is the estimation accuracy.

If the investigator only observes the $i$-th disclosed data $Y_i$, the investigator can also infer, at least partially, the $i$-th private...
data $X_i$. Since we assume to use the private data $X_i$ in another cryptographic protocol, the UC-security measure $\|p_0 - p_1\|_1$ is suitable for a privacy criterion. The UC-security measure characterizes the distinguishability as follows. The minimum value of the weighted error probability of the above inference with a weight $w \in (0, 1)$ is $\min_{S \subseteq Y} \{(1 - w)p_0(S) + w p_1(S^c)\}$, where $S$ is a subset of events to infer $X_i = 1$. When treating two error probabilities $p_0(S)$ and $p_1(S^c)$ equally, the weight $w$ is one half. When taking the error probability $p_0(S)$ more seriously than $p_1(S^c)$, the weight $w$ is smaller than one half. Now, the minimum value of the weighted error probability is calculated as

$$\min_{S \subseteq Y} \{(1 - w)p_0(S) + wp_1(S^c)\} = \frac{1}{2}(1 - \|p_0 - p_1\|_1),$$

which quantifies the distinguishability. Thus, to keep individual privacy a given level, we impose a constant constraint on the $l^1$-norm as $\|p_0 - p_1\|_1 \leq \delta$, which is equivalent to the condition

$$\min_{S \subseteq Y} \{(1 - w)p_0(S) + wp_1(S^c)\} \geq a := \frac{1 - \delta}{2}. \quad (1)$$

To protect individual privacy, many preceding studies consider $(\epsilon, \delta)$-differential privacy, which is defined by the following condition: for any $S \subseteq Y$ and any $i, j \in X$,

$$p_i(S) \leq e^{\epsilon}p_j(S) + \delta. \quad (2)$$

When $\epsilon = 0$, the above condition can be simplified to $\|p_0 - p_1\|_1 \leq \delta$. Thus we can also say that the $l^1$-norm constraint is $(0, \delta)$-differential privacy when $w = 1/2$. As explained in the previous paragraph, we also address the minimum value of the weighted error probability as a privacy criterion. Hence we maximize the Fisher information $J_\theta$ under the $l^1$-norm constraint $\|p_0 - p_1\|_1 \leq \delta$. In the following, we denote the Fisher information by $J_\theta(p_0, p_1)$.

From now on, we shall discuss only the case $\delta \in (0, 1)$ because the other cases correspond to trivial domains. Further, since the triangle inequality yields $|1 - 2w| = |(1 - w) - w| \leq ||(1 - w)p_0 - wp_1||_1$, the constraint $||(1 - w)p_0 - wp_1||_1 \leq \delta$ implies $|1 - 2w| \leq \delta$, which is equivalent to the condition $a = (1 - \delta)/2 \leq w \leq (1 + \delta)/2 = 1 - a$. Hence we also impose the condition $w \in [a, 1 - a]$ on $w \in (0, 1)$. Then we have the following theorem:

**Theorem 1.** Assume the conditions $w \in [a, 1 - a]$ and $\delta, \theta \in (0, 1)$, and set the parameters $a := (1 - \delta)/2$ and $\theta_0 := (w - a)/\delta$. Then we have three cases: (i) $|Y| = 2$ and $\theta \leq \theta_0$, (ii) $|Y| = 2$ and $\theta > \theta_0$, and (iii) $|Y| \geq 3$. In these three cases, from top to bottom, we have

$$\max_{w \in [a, 1 - a]} \frac{J_\theta(p_0, p_1)}{\|p_0 - p_1\|_1 \leq \delta} \begin{cases} a & \text{if } |Y| = 2 \text{ and } \theta \leq \theta_0, \\ \frac{a}{1 - \theta} & \text{if } |Y| = 2 \text{ and } \theta > \theta_0, \\ \frac{a}{1 - \theta} & \text{if } |Y| \geq 3 \text{ and } \theta \leq \theta_0, \\ \frac{a}{1 - \theta} & \text{if } |Y| \geq 3 \text{ and } \theta > \theta_0. \end{cases}$$

Pairs of two distributions in the cases (iii) contain all pairs in the case $|Y| = 2$. Hence the maximum Fisher information can be achieved at least when the number of elements of the probability space $Y$ is three. In this case, the optimal choice of two distributions $p_0$ and $p_1$ does not depend on $\theta$. However, it depends on $\delta$ and $w$. Fig. 2 illustrates the relation between the parameter $\theta$ and the maximum value given in Theorem 1 when the threshold $\delta$ and the weight $w$ take the specific values.

![Fig. 2.](image_url)

**III. Optimization with General Sublinear Function**

To prove Theorem 1 when $|Y| = 3$, we maximize a more general objective function, which is the sum of values of a sublinear function. Although this objective function was optimized in [12] under the $(\epsilon, 0)$-differential privacy constraint, we optimize it under the $l^1$-norm constraint. First, we define sublinear functions as follows.

**Definition 1.** A function $\psi : [0, \infty)^2 \to \mathbb{R}$ is sublinear if the following conditions hold:

- $\psi(\alpha x, \alpha y) = \alpha \psi(x, y)$ for all $\alpha > 0$ and $x, y \geq 0$;
- $\psi(x_1 + x_2, y_1 + y_2) \leq \psi(x_1, y_1) + \psi(x_2, y_2)$ for all $x_1, x_2, y_1, y_2 \geq 0$.

Any sublinear function is convex. Thus the sum of values of a sublinear function is also convex. Conversely, if a convex
function $\psi$ satisfies the first condition in Definition 1, then $\psi$ is sublinear. This fact follows from Definition 1 immediately.

Let $\psi$ be a sublinear function and define the function $\Psi$ as

$$
\Psi(p_0, p_1) := \sum_{y \in \mathcal{Y}} \psi(p_0(y), p_1(y))
$$

for any two distributions $p_0$ and $p_1$. Then we can maximize $\Psi$ under the $l^1$-norm constraint as follows.

**Theorem 2.** When the conditions $|\mathcal{Y}| \geq 3$, $w \in [a, 1-a]$, and $\delta \in (0,1)$ hold, the maximization problem

$$
\max_{\| (1-w)p_0 - wp_1 \|_1 \leq \delta} \Psi(p_0, p_1)
$$

achieves the maximum value at the ordered pair $(\tilde{p}_0, \tilde{p}_1)$ of the two distributions

$$
\tilde{p}_0 = \left[ \frac{a}{1-w}, 1 - \frac{a}{1-w}, 0 \right], \quad \tilde{p}_1 = \left[ \frac{a}{w}, 0, 1 - \frac{a}{w} \right].
$$

**Proof.** We regard the probability space $\mathcal{Y}$ as the set $\{1, \ldots, N\}$. To show this theorem, we remark a few facts. First, the domain of the maximization problem (4) is compact and convex. Second, the objective function is convex. Thus the maximum value is achieved at an extreme point of the domain. We assume that $(p_0, p_1)$ is an extreme point of the domain in the following steps.

**Step 1.** We show that there exists an element $y \in \mathcal{Y}$ satisfying $p_0(y)p_1(y) > 0$ by contradiction. Suppose that any element $y \in \mathcal{Y}$ satisfies $p_0(y)p_1(y) = 0$. This assumption implies $\delta = \| (1-w)p_0 - wp_1 \|_1 = |1-w| + w = 1$, which contradicts the assumption $0 < \delta < 1$. Therefore, an element $y \in \mathcal{Y}$ satisfies $p_0(y)p_1(y) > 0$. Without loss of generality, we may assume that the above element $y$ is 1, i.e., $p_0(1)p_1(1) > 0$ by changing elements if necessarily.

**Step 2.** We prove that any element $y \neq 1$ satisfies $p_0(y)p_1(y) = 0$ by contradiction. Suppose $p_0(2)p_1(2) > 0$ by changing elements if necessarily. Taking a sufficiently small positive number $\epsilon$, we define the distributions $q_k$ and $q_k'$ for $k = 0,1$ as

$$
q_0(1) = p_0(1) - \epsilon w, \quad q_0'(1) = p_0(1) + \epsilon w,
$$

$$
q_0(2) = p_0(2) + \epsilon w, \quad q_0'(2) = p_0(2) - \epsilon w,
$$

$$
q_1(1) = p_1(1) - (1-\epsilon)w, \quad q_1'(1) = p_1(1) + (1-\epsilon)w,
$$

$$
q_1(2) = p_1(2) + (1-\epsilon)w, \quad q_1'(2) = p_1(2) - (1-\epsilon)w,
$$

$$
q_1(y) = q_1'(y) = p_1(y) \quad (3 \leq y \leq N).
$$

Then, for any $k = 0,1$, the relations $\| (1-w)q_0 - wq_1 \|_1 = \| (1-w)q_0' - wq_1' \|_1 \leq \delta$, $p_k = (q_k + q_k')/2$, and $q_k \neq q_k'$ hold, which contradicts that the point $(p_0, p_1)$ is an extreme point of the domain. Thus any element $y \neq 1$ satisfies $p_0(y)p_1(y) = 0$.

**Step 3.** We show the relations

$$
\min\{ (1-w)p_0(1), wp_1(1) \} \geq a, \quad \Psi(p_0, p_1) = \psi(p_0(1), p_1(1)) + (1 - p_0(1))\psi(1,0) + (1 - p_1(1))\psi(0,1). \tag{5}
$$

Using the result in Step 2, for each element $y \neq 1$, we can take an element $k(y) \in \mathbb{Z}/2\mathbb{Z}$ satisfying $p_{k(y)+1}(y) = 0$. Then the inequality

$$
d \geq \| (1-w)p_0 - wp_1 \|_1
$$

$$
= \| (1-w)p_0(1) - wp_1(1) \|
$$

$$
+ (1-w) \sum_{2 \leq y \leq 2N} p_0(y) + w \sum_{k(y) = 1} p_1(y)
$$

$$
= \| (1-w)p_0(1) - wp_1(1) \|
$$

$$
+ (1-w)(1 - p_0(1)) + w(1 - p_1(1))
$$

$$
= \| (1-w)p_0(1) - wp_1(1) \| + 1 - (1-w)p_0(1) - wp_1(1)
$$

holds, which means (5). Further, (6) is verified as follows:

$$
\Psi(p_0, p_1) = \psi(p_0(1), p_1(1)) + \sum_{2 \leq y \leq N} p_0(y)\psi(1,0)
$$

$$
+ \sum_{k(y) = 1} p_1(y)\psi(0,1)
$$

$$
= \psi(p_0(1), p_1(1)) + (1 - p_0(1))\psi(1,0) + (1 - p_1(1))\psi(0,1). \tag{6}
$$

**Step 4.** In this step, assuming $N \geq 3$, we show our assertion. The linearity of $\psi$ implies

$$
|\psi(1 - \frac{a}{1-w}, \frac{a}{1-w})| + (1 - \frac{a}{1-w})\psi(1,0) + (1 - \frac{a}{w})\psi(0,1)
$$

$$
- \psi(p_0(1), p_1(1))
$$

$$
= \psi(1 - \frac{a}{1-w}, \frac{a}{1-w}) - \psi(p_0(1), p_1(1))
$$

$$
+ \left( p_0(1) - \frac{a}{1-w} \right)\psi(1,0) + \left( p_1(1) - \frac{a}{w} \right)\psi(0,1)
$$

$$
\geq \psi\left( \frac{a}{1-w}, \frac{a}{1-w} \right) - \psi\left( \frac{a}{1-w}, \frac{a}{1-w} \right)
$$

$$
- \psi\left( p_0(1) - \frac{a}{1-w}, 0 \right) - \psi\left( 0, p_1(1) - \frac{a}{w} \right)
$$

$$
+ \left( p_0(1) - \frac{a}{1-w} \right)\psi(1,0) + \left( p_1(1) - \frac{a}{w} \right)\psi(0,1) \tag{7}
$$

$$
\geq 0,
$$

where (6) and (5) have been used to obtain (a) and (b), respectively. Therefore, $\Psi(p_0, p_1)$ is the maximum value of the maximization problem (4).

**Proof of Theorem 1 when $|\mathcal{Y}| \geq 3$.** Since the Fisher information $I_{\beta}(p_0, p_1) = \sum_{y \in \mathcal{Y}} (p_1(y) - p_0(y))^2/p_0(y)$ is the sum of values of a sublinear function, Theorem 2 implies Theorem 1 when $|\mathcal{Y}| \geq 3$.

**IV. Related work**

Here we compare our result with preceding studies. The earlier studies [1], [2] discussed the estimation error in a similar way, but they did not give any privacy criteria. Warner [1] proposed a scheme to protect individual privacy, in which
TABLE II
COMPARISON WITH EXISTING RESULTS BASED ON UC-SECURITY MEASURE

This table shows optimal pairs of two distributions $p_0$ and $p_1$ in [1], [2], [15], and ours.

| Scheme | Value | Optimal pair of two distributions |
|--------|-------|-----------------------------------|
| Warner [1] | $\theta = 2$ | $p_0 = ([1 + \delta)/2, (1 - \delta)/2]$  
$p_1 = ([1 - \delta)/2, (1 + \delta)/2]$ |
| Unrelated question [2] | $\theta = 0$ | $p_0 = (\delta + (1 - \delta)(1 - \eta), (1 - \delta)\eta]$  
$p_1 = (1 - \delta)(1 - \eta), \delta + (1 - \delta)\eta]$ |
| Holohan et al. [15] | $\theta \leq 1/2$ | $p_0 = [1, 0]$  
$p_1 = [1 - \delta, 0]$ |
| This paper | $\theta > 1/2$ | $p_0 = [2a, 1 - 2a, 0]$  
$p_1 = [2a, 0, 1 - 2a] = [1, \delta, 0, \delta]$ |

Each individual flips each true answer with probability $1 - \pi$ and does not flip it with probability $\pi \in (0, 1)$. That is, he proposed to set $p_0$ and $p_1$ as $p_0 = [\pi, 1 - \pi]$ and $p_1 = [1 - \pi, \pi]$ in our notation. Greenberg et al. [2] proposed another scheme by using a question unrelated to an intended YES/NO question. In their scheme, the investigator asks each individual the intended question and the unrelated one. The asked individual answers the former with probability $\epsilon$ and the latter with probability $1 - \pi$. When the unrelated question has the true ratio $\eta = 1 - \delta$, the distributions $p_0$ and $p_1$ are set to $p_0 = [\pi + (1 - \pi)(1 - \eta), (1 - \pi)\eta]$ and $p_1 = [(1 - \pi)(1 - \eta), \pi + (1 - \pi)\eta]$ in our notation. Maximizing the Fisher information in two sets of the above respective pairs $(p_0, p_1)$, we obtain the optimal pairs in Table II. Our scheme is best of the pairs in Table II because our scheme is to maximize the Fisher information in the set of all pairs of two distributions. Moreover, the studies [1], [2] did not consider the Fisher information and considered only the case $|X| = |Y| = 2$. Hence, even if their schemes are optimized, it is impossible to surpass our optimal performance not only the blue broken curve but also the red solid curve in Fig. 2.

The studies [12], [15], [16] are closely related to this paper. Holohan et al. [15] showed optimal $(\epsilon, \delta)$-differentially private mechanisms explicitly in the case $|X| = 2$. However, they considered only the case $|Y| = 2$. Hence, when $\epsilon = 0$, their result [15, Theorem 3] corresponds to the case (i) in Theorem 1, but the case (iii) in Theorem 1 is not known at all. Their result in the case $\epsilon = 0$ is in Table II and illustrated by the red broken curve in Fig. 2. Kairouz et al. [12] provided a theorem on optimal $(\epsilon, 0)$-differentially private mechanisms in the case $|X| \geq 2$. Their theorem can be applied to many objective functions including the Fisher information and the f-divergence, and turns convex optimization problems to linear programs. As stated in the previous sections, their objective functions are the same as ours. Hence, in the case $|X| = 2$, Theorem 2 can be regarded as the $(0, \delta)$-differential privacy version of the result in [12]. Table I in Section I summarizes the relation among [12], [15] and this paper. Moreover, in the case $X = Z$, Geng and Viswanath [16] discussed minimization of $L^1$ and $L^2$ cost functions under the $(\epsilon, \delta)$-differential privacy constraint and gave lower and upper bounds of the minimum values. As a special case, when $\epsilon = 0$ and $\delta \to 0$, their lower and upper bounds are equal asymptotically. Their scheme to protect individual privacy is different from ours because their scheme is to add uniform noise or discrete Laplacian noise to integers. For other related work, see the full version arXiv:1805.06278.

V. CONCLUDING REMARKS

In conclusion, for the trade-off problem associated with binary private data, we have proposed a randomized mechanism that can maximize the estimation accuracy of a global property while keeping individual privacy a given level. Since we assume that private data are used for another cryptographic protocol like authentication, the UC-security measure is suitable for a privacy criterion. In our setting, the UC-security measure is the $1^2$-norm $(1/2)||p_0 - p_1||_1$ and the constraint $(1/2)||p_0 - p_1||_1 \leq \delta$ can be regarded as $(0, \delta)$-differentially private. Under this constraint, we have maximized the Fisher information that is the estimation accuracy of a global property. In particular, to achieve the maximum value, the disclosed data set must consist of at least three elements: $|X| \geq 3$. This fact is new and different from the $(\epsilon, 0)$-differential privacy case.

REFERENCES

[1] S. L. Warner. Randomized response: a survey technique for eliminating evasive answer bias. J. Amer. Statist. Assoc., 60(309):63–69, 1965.
[2] B. G. Greenberg, A.-L. A. Abul-Ela, W. R. Simmons, and D. G. Horvitz. The unrelated question randomized response model: theoretical framework. J. Amer. Statist. Assoc., 64(326):520–539, 1969.
[3] L. Le Cam. Locally Asymptotically Normal Families of Distributions: Certain approximations to families of distributions and their use in the theory of estimation and testing hypotheses. University of California Press, Berkeley, CA, 1960.
[4] L. Le Cam. Asymptotic Methods in Statistical Decision Theory. Springer-Verlag, New York, NY, 1986.
[5] A. W. Van der Vaart. Asymptotic Statistics. Cambridge University Press, Cambridge, 1998.
[6] H. Krawczyk. LFSR-based hashing and authentication. In Proc. of CRYPTO’94, pages 129–139, Santa Barbara, CA, USA, 1994.
[7] R. Canetti. Universal composable security: A new paradigm for cryptographic protocols. In Proc. of FOCS’01, pages 136–145, Washington, DC, USA, 2001.
[8] H. Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. Ann. Math. Stat., 23(4):493–507, 1952.
[9] M. Hayashi. Tight exponential analysis of universally composable privacy amplification and its applications. IEEE Trans. Inf. Theory, 59(11):7728–7746, 2013.
[10] U. Maurer and S. Wolf. Information-theoretic key agreement: From weak to strong secrecy for free. In Proc. of EUROCRYPT’00, pages 351–368, Bruges, Belgium, 2000.
[11] R. Renner and S. Wolf. Simple and tight bounds for information reconciliation and privacy amplification. In Proc. of ASIACRYPT’03, pages 199–216, Chennai, India, 2005.
[12] P. Kairouz, S. Oh, and P. Viswanath. Extremal mechanisms for local differential privacy. J. Mach. Learn. Res., 17(1):492–542, 2016.
[13] C. Dwork, F. McSherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. In Proc. of TCC’06, pages 265–284, New York, NY, USA, 2006.
[14] C. Dwork. Differential privacy. In Proc. of ICAI’06, pages 1–12, Venice, Italy, 2006.
[15] N. Holohan, D. J. Leith, and O. Mason. Optimal differentially private mechanisms for randomised response. IEEE Trans. Inf. Forensics, 12(11):2726–2735, 2017.
[16] Q. Geng and P. Viswanath. Optimal noise adding mechanisms for approximate differential privacy. IEEE Trans. Inf. Theory, 62(2):952–969, 2016.