Robust Local Differential Privacy

Milan Lopuhaä-Zwakenberg
Eindhoven University of Technology,
the Netherlands

Jasper Goseling
University of Twente, and
CWI, the Netherlands

***Abstract*—We consider data release protocols for data $X = (S, U)$, where S is sensitive; the released data Y contains as much information about X as possible, measured as $\mathrm{I}(X;Y)$, without leaking too much about S. We introduce the Robust Local Differential Privacy (RLDP) framework to measure privacy. This framework relies on the underlying distribution of the data, which needs to be estimated from available data. Robust privacy guarantees ensure privacy for all distributions in a confidence set based on this estimate. We also present three algorithms that construct RLDP protocols from a given dataset. One of these approximates the confidence set by a polytope and uses results from robust optimisation to yield high utility release protocols. However, it relies on vertex enumeration and becomes computationally infeasible for large input alphabets. The other two algorithms are low-complexity and build on randomised response. Experiments verify that all three algorithms offer significantly improved utility over regular LDP.**

I. Introduction

We consider the setting in which users have data $X = (S, U)$ that a data aggregator is interested in, but users do not wish to disclose information about sensitive data S. Therefore, users release an obfuscated version Y of X, such that Y contains as much information about X as possible, measured as $\mathrm{I}(X;Y)$, without leaking too much about S. This scenario and related ones have been studied in, for instance, [1]–[8].

This paper uses a form of local differential privacy (LDP) [9] to measure the amount of information that Y leaks on S. The following version of ε-LDP was introduced in [10]:

$$\mathbb{P}(Y = y|S = s) \leq \mathrm{e}^{\varepsilon}\mathbb{P}(Y = y|S = s'), \tag{1}$$

for all y, s and s'. Note, that this condition is less strict than $\mathbb{P}(Y = y|X = x) \leq \mathrm{e}^{\varepsilon}\mathbb{P}(Y = y|X = x')$ as would be used in ordinary LDP. Also note that (1) relies on the distribution $P_X = P_{S,U}$. From these observations it follows that this privacy definition enables higher utility of the released data Y at the expense of not being completely 'distribution free' as would be the case for ordinary LDP.

In [10] condition (1) is studied for the case of known P_X. This is a strong assumption, since users will need to estimate P_X. When an attacker has better knowledge of P_X than the user, it follows from the odds-ratio interpretation of differential privacy [11] that privacy is not guaranteed in such a scenario.

In this paper we, therefore, provide stronger privacy guarantees. In particular, we introduce robustness constraints, which say that privacy should not just hold for one P_X, but for a set $\mathcal{F}$ of these. In particular, we assume that there is publicly available data from n users, which allows the users as well as an attacker to estimate $\hat{P}_X$. The set $\mathcal{F}$ consists of those P that are close enough to $\hat{P}_X$ so that the difference is not statistically significant for a chosen significance level α. As a result we guarantee privacy against attackers with (at least) reasonable estimates of P_X, without sacrificing utility to protect against attackers with no or unreliable information on P_X. We refer to the resulting privacy framework as Robust Local Differential Privacy (RLDP).

In addition to introducing RLDP, the *main contributions* of this paper are as follows: We characterize $\mathcal{F}$ by an enveloping polytope. We then use techniques from robust optimisation [12]–[14] to characterize the protocol that is optimal over this polytope. This gives high utility and demonstrates the advantage of RLDP over ordinary LDP. A drawback of this algorithm is that it relies on vertex enumeration and is, therefore, computationally unfeasible for large alphabets. Therefore, we introduce two low-complexity algorithms to find release protocols: i) Independent Reporting (IR), in which S and U are reported through separate LDP protocols, and ii) Conditional Reporting (CR), in which first S is obfuscated, and either a slightly obfuscated U or a randomly drawn U' is returned, depending on whether the obfuscated S is 'correct'. For both mechanisms we characterize the conditions that underlying LDP protocols have to satisfy in order to ensure RLDP. Furthermore, while both algorithms can incorporate any LDP protocol, we show that it is optimal to use Randomised Response [15]. This drastically reduces the search space and allows us to find the optimal IR and CR mechanisms using one-dimensional optimisation.

Related work on addressing robustness with respect to estimating the distribution P_X in data release protocols consider robustness under α-leakage [16] and the impact of incorporating robustness on utility [17]. An important difference with the current work is that [16], [17] use α-leakage [18] and related measures instead of local differential privacy. The case in which no information about P_X is known is studied in [19].

The structure of this paper is as follows. In Section II we describe the model in detail. In Section III we present the mechanism based on polyhedral approximation. In Section IV we present the IR and CR protocols and show that the optimal IR and CR protocols can be found using one-dimensional optimisation. In Section V we evaluate the discussed methods experimentally. A discussion of the results and an outlook on future work is given in Section VI.

Due to space constraints all proofs are omitted from this paper. Proofs and additional results are provided in [20].

2021 IEEE International Symposium on Information Theory (ISIT) | 978-1-5386-8209-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ISIT45174.2021.9518151

II. Model

An overview of our model, the details of which are given in this section, is given in Figure 1. There is a publicly accessible dataset $\vec{x}$, in which each entry $x_i = (s_i, u_i)$ is drawn independently from a probability distribution P^* on a set $\mathcal{X} = \mathcal{S} \times \mathcal{U}$, where $\mathcal{S}$ and $\mathcal{U}$ are finite alphabets. New data items $X = (S, U)$ are also drawn from P^*. The user's aim is to create a release protocol $\mathcal{Q}$ such that $Y = \mathcal{Q}(X)$ contains as much information about X as possible, while not leaking too much information about S.

The distribution P^* is not known exactly. The uncertainty set $\mathcal{F} \subset \mathcal{P}_\mathcal{X}$, where $\mathcal{P}_\mathcal{X}$ denotes the probability simplex over $\mathcal{X}$, captures the user's uncertainty about P^*. It is constructed from the dataset $\vec{x}$. More specifically, we let $\mathcal{F}$ be the $(1-\alpha)$-confidence set for P in a χ^2-test, i.e.

$$\mathcal{F} = \left\{ P \middle| \sum_{x \in \mathcal{X}} \frac{(\hat{P}_x - P_x)^2}{P_x} \leq B := \frac{F^{-1}_{|\mathcal{X}|-1}(1-\alpha)}{n} \right\}, \quad (2)$$

where $\hat{P}$ is the empirical probability distribution of X and F_d is the CDF of the χ^2-distribution with d degrees of freedom. By definition $P^* \in \mathcal{F}$ with probability $1 - \alpha$.

From $\hat{P}$ and $\mathcal{F}$ the user creates $\mathcal{Q}$ in such a way that a Local Differential Privacy-like privacy standard is guaranteed when $P^* \in \mathcal{F}$. We will denote this as robust local differential privacy (RLDP). It is defined as follows.

Definition 1. *Let $\varepsilon \geq 0$ and $\mathcal{F} \subset \mathcal{P}_\mathcal{X}$. We say that $\mathcal{Q}$ satisfies $(\varepsilon, \mathcal{F})$-RLDP if for all $s, s' \in \mathcal{S}$, all $y \in \mathcal{Y}$, and all $P \in \mathcal{F}$ we have*

$$\mathbb{P}_{X \sim P}(Y = y | S = s) \leq \mathrm{e}^\varepsilon \mathbb{P}_{X \sim P}(Y = y | S = s'). \quad (3)$$

Following [3], [10] we take $\mathrm{I}_{X \sim \hat{P}}(X; Y)$ as the utility measure; we mean this when we write $\mathrm{I}(X; Y)$ without further comment. Thus the user's goal is, for a given ε, $\hat{P}$, and α, to find the $\mathcal{Q}$ that maximises $\mathrm{I}(X; Y)$ while satisfying $(\varepsilon, \mathcal{F})$-RLDP.

III. Polyhedral approximation: PolyOpt

Our first method to find RLDP (release) protocols relies on optimising $\mathrm{I}(X; Y)$ over protocols that satisfy a more stringent privacy constraint; this yields a lower bound on the maximal $\mathrm{I}(X; Y)$. More concretely, let $P_{\mathcal{U}|s} = (P_{u|s})_{u \in \mathcal{U}} \in \mathcal{P}_\mathcal{U}$ (the probability simplex over $\mathcal{U}$). We consider protocols that satisfy (3) for all P for which $P_{\mathcal{U}|s} \in \mathcal{D}_{\mathcal{U}|s}$, where each $\mathcal{D}_{\mathcal{U}|s}$ is a polytope containing $\mathcal{F}_{\mathcal{U}|s} := \{P_{\mathcal{U}|s} | P \in \mathcal{F}\}$. Such a protocol certainly satisfies $(\varepsilon, \mathcal{F})$-RLDP. Robust optimisation for polytopes [12] allows us to view the set of such protocols as a polytope itself. More concretely, for $(s, u) \in \mathcal{X}$ define $P^{\min}_{u|s} = \min_{P \in \mathcal{F}} P_{u|s}$, and define

$$\mathcal{D}_{\mathcal{U}|s} = \left\{ R \in \mathcal{P}_\mathcal{U} \middle| \forall u \colon R_u \geq P^{\min}_{u|s} \right\}. \quad (4)$$

Then $\mathcal{F}_{\mathcal{U}|s} \subset \mathcal{D}_{\mathcal{U}|s}$. If $a_s = (\sqrt{B+1} + \hat{P}_s - 1)^2 \hat{P}_s^{-2} - 1$, one can calculate that

$$P^{\min}_{u|s} = \frac{a_s + 2\hat{P}_{u|s} - \sqrt{a_s^2 + 4a_s\hat{P}_{u|s} - 4a_s\hat{P}^2_{u|s}}}{2a_s + 2}. \quad (5)$$

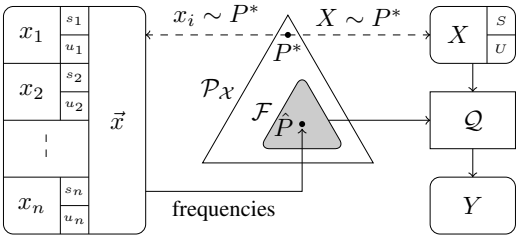

Fig. 1: Overview.

Let $\Gamma \subset \mathbb{R}^\mathcal{X}$ be the convex cone consisting of all T satisfying $\forall x \colon T_x \geq 0$ and

$$\begin{aligned} \forall s_1, s_2, u_1, u_2 \colon \\ T_{s_1,u_1} - \mathrm{e}^\varepsilon T_{s_2,u_2} + \sum_u P^{\min}_{u|s_1} (T_{s_1,u} - T_{s_1,u_1}) \\ - \mathrm{e}^\varepsilon \sum_u P^{\min}_{u|s_2} (T_{s_2,u} - T_{s_2,u_2}) \leq 0. \end{aligned} \quad (6)$$

We identify a protocol $\mathcal{Q}$ with the matrix $Q \in \mathbb{R}^{\mathcal{Y} \times \mathcal{X}}$ given by $Q_{y|x} = \mathbb{P}(\mathcal{Q}(x) = y)$. Robust optimisation for polytopes now yields the following result:

Theorem 1. *Let $\mathcal{Q}$ be a privacy protocol such that for all y we have $(Q_{y|x})_{x \in \mathcal{X}} \in \Gamma$. Then $\mathcal{Q}$ satisfies ε-RLDP.*

From the cone Γ we can employ the techniques of [9] to reduce finding the optimal $\mathcal{Q}$ to a vertex enumeration problem and a linear optimisation problem:

Theorem 2. *Let $\hat{\Gamma}$ be polytope given by $\{T \in \Gamma | \sum_x T_x = 1\}$. Let $\mathcal{V}$ be the set of vertices of $\hat{\Gamma}$. For $v \in \mathcal{V}$, define*

$$\mu_v = \sum_x v_x \hat{P}_x \log \frac{v_x}{\sum_{x'} v_{x'} \hat{P}_{x'}}, \quad (7)$$

Let $\hat{\theta}$ be the solution to the optimisation problem

$$\begin{aligned} &\text{maximise}_\theta \sum_{v \in \mathcal{V}} \theta_v \mu_v \\ &\text{satisfying } \theta \in \mathbb{R}^\mathcal{V}_{\geq 0}, \\ &\sum_v \theta_v v = 1_\mathcal{X}. \end{aligned} \quad (8)$$

Let the protocol $\mathcal{Q}$ be given by $\mathcal{Y} = \{v \in \mathcal{V} : \hat{\theta}_v > 0\}$ and $Q_{v|x} = \hat{\theta}_v v_x$. Then $\mathcal{Q}$ maximises $\mathrm{I}(X; Y)$ among all protocols satisfying the condition of Theorem 1. One has $|\mathcal{Y}| \leq |\mathcal{X}|$.

The method in [9] does not need the vertex enumeration step as their $\hat{\Gamma}$ is always a hypercube.

Together, Theorems 1 and 2 show, if we can solve a vertex enumeration problem, that we can find a protocol $\mathcal{Q}$ that maximises $\mathrm{I}(X; Y)$ among a subset of all $(\varepsilon, \mathcal{F})$-RLDP $\mathcal{Q}$. The algorithm that produces $\mathcal{Q}$ from $\hat{P}$ and ε will be refered to as *PolyOpt* in the remainder of this paper.

Remark 1. A simplex is not the only possible choice for $\mathcal{D}_{\mathcal{U}|s}$. In general, we can make $\mathcal{D}_{\mathcal{U}|s}$ closer to $\mathcal{F}_{\mathcal{U}|s}$ by adding more defining hyperplanes. Doing this allows more $\mathcal{Q}$ to satisfy Theorem 1, and in turn increases the utility of the $\mathcal{Q}$ we find via Theorem 2. However, since Γ is related to the $\mathcal{D}_{\mathcal{U}|s}$ via duality, adding extra constraints to the $\mathcal{D}_{\mathcal{U}|s}$ will increase the dimension of Γ through the addition of auxiliary variables. This makes the vertex enumeration problem of Theorem 2 more computationally involved. Thus we have a tradeoff between utility and computability.

It should be noted that in general the increasing utility found in this way does not approach the optimal utility over all $(\varepsilon, \mathcal{F})$-RLDP protocols. This is because, as we take increasingly finer $\mathcal{D}_{\mathcal{U}|s}$, we approach the set of $\mathcal{Q}$ that satisfy (3) for all P in $\mathcal{F}' := \{P \in \mathcal{P}_{\mathcal{X}} | \forall s \colon P_{\mathcal{U}|s} \in \mathcal{F}_{\mathcal{U}|s}\}$. Since in general $\mathcal{F} \subsetneq \mathcal{F}'$, the set of $(\varepsilon, \mathcal{F}')$-RLDP protocols is strictly smaller than the set of $(\varepsilon, \mathcal{F})$-RLDP protocols.

IV. Low-complexity Mechanisms

While PolyOpt is expected to give good utility results, its downside is that it relies on vertex enumeration in a space with dimension $|\mathcal{X}|$. As such, it becomes computationally infeasible for larger input domains. In this section we introduce two algorithms that build release protocols from existing local differential privacy (LDP) protocols. Recall that a random function $\mathcal{R} \colon \mathcal{A} \to \mathcal{B}$ satisfies ε-LDP if

$$\mathbb{P}(\mathcal{R}(a) = b) \leq \mathrm{e}^{\varepsilon} \mathbb{P}(\mathcal{R}(a') = b) \tag{9}$$

for all $a, a' \in \mathcal{A}$ and $b \in \mathcal{B}$. An important building block for us will be the ε-LDP protocol *Generalised Randomised Response* (GRR) [15], which is, for a given ε, the protocol $\mathcal{G}^{\varepsilon} \colon \mathcal{A} \to \mathcal{A}$ given by

$$\mathbb{P}(\mathcal{G}^{\varepsilon}(a) = a') = \frac{1 + (\mathrm{e}^{\varepsilon} - 1)\delta_{a=a'}}{\mathrm{e}^{\varepsilon} + |\mathcal{A}| - 1}. \tag{10}$$

An important property of GRR is that for every random variable A on $\mathcal{A}$ there is a ε_0 such that GRR is the ε-LDP protocol that maximises $\mathrm{I}(A, \mathcal{R}(A))$ for $\varepsilon \geq \varepsilon_0$ [9].

A. Pure LDP application

One way to ensure $(\varepsilon, \mathcal{F})$-RLDP is to apply an ε-LDP protocol $\mathcal{R}$ to X. Any such protocol certainly satisfies $(\varepsilon, \mathcal{F})$-RLDP for any $\mathcal{F}$. However, such a protocol works by obfuscating the entirety of X, rather than just S; this typically comes at a utility cost. In the experiments, we apply GRR to X to provide a baseline utility.

B. Independent Reporting

The first new algorithm we introduce is Independent Reporting (IR). The basis of IR is to apply two separate LDP protocols $\mathcal{R}^1$ and $\mathcal{R}^2$ to S and U, respectively, and output $\mathcal{Q}(X) := (\mathcal{R}^1(S), \mathcal{R}^2(U))$. This is described in Protocol 1.

While only S needs to be protected, we also need to apply an LDP protocol to U because of the possible correlation between the two. However, since U only indirectly leaks information about S, we can get away with less strict privacy requirements. The degree to which U and S might be correlated depends on two factors: the correlation between S and U under $\hat{P}$, and the size of $\mathcal{F}$. By finding bounds for the latter we get the following result:

Protocol 1: $\mathrm{IR}_{\mathcal{R}^1, \mathcal{R}^2}$ (Independent reporting)

Input : Privacy protocols $\mathcal{R}^1 \colon \mathcal{S} \to \mathcal{Y}^1$ and $\mathcal{R}^2 \colon \mathcal{U} \to \mathcal{Y}^2$; $x = (s, u) \in \mathcal{X}$.
Output: Output datum $y \in \mathcal{Y} := \mathcal{Y}^1 \times \mathcal{Y}^2$

Compute $y_1 \leftarrow \mathcal{R}^1(s)$;
Compute $y_2 \leftarrow \mathcal{R}^2(u)$;
$y \leftarrow (y_1, y_2)$;

Theorem 3. *Let $\varepsilon_1, \varepsilon_2 \in \mathbb{R}_{\geq 0}$. For each s, define a_s as above (5), and let $u_s \in \mathcal{U}$ be such that $\hat{P}_{u_s|s}$ is minimal. Define*

$$d_s := \begin{cases} \frac{a_s(1-2\hat{P}_{u_s|s}) + \sqrt{a_s^2 + 4a_s\hat{P}_{u_s|s} - 4a_s\hat{P}_{u_s|s}^2}}{a_s + 1}, & \text{if } a_s \geq 1; \\ \sqrt{a_s}, & \text{if } a_s \leq 1. \end{cases} \tag{11}$$

Furthermore, define

$$d := \min\left\{2, \max_s(2d_s) + \max_{s,s'} ||\hat{P}_{\mathcal{U}|s} - \hat{P}_{\mathcal{U}|s'}||_1\right\}. \tag{12}$$

Let $\delta_2 = \log\left(1 + \frac{2(\mathrm{e}^{\varepsilon_2} - 1)}{d}\right)$. Suppose that $\mathcal{R}^1$ is ε_1-LDP and that $\mathcal{R}^2$ is δ_2-LDP. Then $\mathrm{IR}_{\mathcal{R}^1, \mathcal{R}^2}$ is $(\varepsilon_1 + \varepsilon_2, \mathcal{F})$-RLDP.

The more independent S and U are, the smaller $\max_{s,s'} ||\hat{P}_{\mathcal{U}|s} - \hat{P}_{\mathcal{U}|s'}||_1$ will be. Theorem 3 then tells us that for more independent S and U, the privacy requirements on $\mathcal{R}^2$ will be less strict, resulting in better utility. The utility of IR is described by the following Theorem:

Theorem 4. *One has*

$$\mathrm{I}(\mathrm{IR}_{\mathcal{R}^1, \mathcal{R}^2}(X); X) = \mathrm{I}(\mathcal{R}^1(S); S) + \mathrm{I}(\mathcal{R}^2(U); U | \mathcal{R}^1(S)). \tag{13}$$

Given an $\varepsilon \geq 0$, we can use these theorems to find $(\varepsilon, \mathcal{F})$-RLDP protocols. Per Theorem 3, it suffices to take a ε_2, and use a ε_1-LDP protocol $\mathcal{R}^1$ and a δ_2-LDP protocol $\mathcal{R}^2$, where $\varepsilon_1 = \varepsilon - \varepsilon_2$ and δ_2 is as in Theorem 3. We want to choose ε_2, $\mathcal{R}^1$ and $\mathcal{R}^2$ in such a way that we maximise the expression in Theorem 4. For ε large enough, the $\mathcal{R}^1$ that maximises $\mathrm{I}(\mathcal{R}^1(S); S)$ is GRR. Furthermore, since

$$\mathrm{I}(\mathcal{R}^2(U); U | \mathcal{R}^1(S)) = \mathbb{E}_r\left[\mathrm{I}_{U \sim \hat{P}_U | \mathcal{R}^1(S) = r}(\mathcal{R}^2(U); U)\right], \tag{14}$$

and GRR maximises $\mathrm{I}(\mathcal{R}^2(U); U)$ for large enough ε for any distribution of U, we should take $\mathcal{R}^2$ to be GRR as well. We are left with only the unknown ε_2, hence to maximise the mutual information of IR we have to solve a onedimensional optimisation problem.

C. Conditional reporting

From Theorem 3 it is clear that in IR, we can afford a larger privacy budget to $\mathcal{R}^2$ if S and U are only weakly correlated. When S and U are closer related, however, the

difference between δ_2 and ε_2 will be small, and IR cannot offer any advantage over general LDP protocols. To this end, we introduce our second new algorithm, *Conditional Reporting*. We first specify a parameter ε_1 and, for each $s \in \mathcal{S}$, a privacy protocol $\mathcal{R}^s : \mathcal{U} \to \mathcal{Y}_s$, where each $\mathcal{Y}_s$ is a finite set. To apply CR to an input datum $(s, u) \in \mathcal{X}$, we first apply GRR with parameter ε_1 to s; call the outcome $\tilde{S}$. If $\tilde{S} = s$, we apply $\mathcal{R}^s$ to u, and we output $(s, \mathcal{R}^s(u))$. If $\tilde{S} \neq s$, we draw a random $\tilde{U} \in \mathcal{U}$ from the probability distribution $\hat{P}_{\mathcal{U}|\tilde{S}}$, and we output $(\tilde{S}, \mathcal{R}^{\tilde{S}}(\tilde{U}))$. This is described in Protocol 2.

Protocol 2: Conditional Reporting (CR)

Input : Privacy parameter ε_1; For every $s \in \mathcal{S}$, a privacy protocol $\mathcal{R}^s : \mathcal{U} \to \mathcal{Y}_s$; input datum $x = (s, u) \in \mathcal{X}$

Output: Output datum $Y \in \mathcal{S} \times \bigcup_{s \in \mathcal{S}} \mathcal{Y}_s$

Take $\tilde{S} \leftarrow \mathcal{G}^{\varepsilon_1}(s) \in \mathcal{S}$;
if $\tilde{S} = s$ **then**
  Compute $Y \leftarrow (s, \mathcal{R}^s(u))$;
else
  Sample $\tilde{U} \in \mathcal{U}$ with $\mathbb{P}(\tilde{U} = u') = \hat{P}_{u'|\tilde{S}}$;
  Compute $Y \leftarrow (\tilde{S}, \mathcal{R}^{\tilde{S}}(\tilde{U}))$;
end
Output Y;

Although we have already obfuscated S via GRR, we still need to obfuscate u and $\tilde{U}$ via $\mathcal{R}^{\tilde{S}}$ for the following reason. Suppose we omit this last step, and instead return $(\tilde{S}, \tilde{U})$, with $\tilde{U} = u$ if $\tilde{S} = s$. From the viewpoint of an attacker, given $\tilde{S}$, the random variable $\tilde{U}$ is drawn from the distribution $P^*_{\mathcal{U}|\tilde{S}}$ if $\tilde{S} = s$, and from the distribution $\hat{P}_{\mathcal{U}|\tilde{S}}$ otherwise. In the LDP model the attacker may collude with an arbitrary amount of users, and as such we may assume that they have access to the real distribution P^*. Under this assumption, the output $\tilde{U}$ contains information about whether it was drawn from $P^*_{\mathcal{U}|\tilde{S}}$ or $\hat{P}_{\mathcal{U}|\tilde{S}}$, and hence whether $S = \tilde{S}$ or not. To prevent this leakage, we have to mask $\tilde{U}$ with the privacy protocol $\mathcal{R}^{\tilde{S}}$. As the following theorem shows, the privacy level that is needed for $\mathcal{R}^s$ depends on d_s, which explains why we need a different protocol $\mathcal{R}^s$ for every s.

Theorem 5. *Let $\varepsilon_1, \varepsilon_2 \in \mathbb{R}_{\geq 0}$. For $s \in \mathcal{S}$ define d_s as in (11). For every s, define $\delta_s := \log\left(1 + \frac{2(\mathrm{e}^{\varepsilon_2}-1)}{d_s}\right)$, and assume each $\mathcal{R}^s$ is δ_s-LDP. Then Algorithm 2 satisfies $(\varepsilon_1 + \varepsilon_2, \mathcal{F})$-RLDP.*

As we can see, the privacy level of $\mathcal{R}^s$ only depends on d_s, which goes to 0 as B grows smaller. This makes CR an attractive protocol if the number of known data points n is large.

On the side of utility, we have the following:

Theorem 6. *One has*

$$\mathrm{I}(\mathrm{CR}(X); X) = \mathrm{I}(\mathcal{G}^{\varepsilon}(S); S) + \frac{\mathrm{e}^{\varepsilon_0}}{\mathrm{e}^{\varepsilon_0}+|\mathcal{S}|-1}\,\mathrm{I}(\mathcal{R}^S(U); U|S). \tag{15}$$

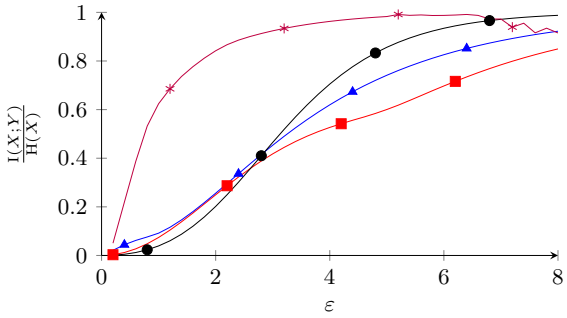

Fig. 2: Experiments on synthetic data for $|\mathcal{S}| = |\mathcal{U}| = 3$. (PolyOpt, GRR, IR, CR)

Compared to Theorem 4, we see that on one hand, CR's utility has a factor $\frac{\mathrm{e}^{\varepsilon_0}}{\mathrm{e}^{\varepsilon_0}+|\mathcal{S}|-1}$ which IR lacks, indicating lower utility. However, the advantage of CR is that $\mathcal{R}^s$ can typically chosen with more relaxed privacy conditions than the $\mathcal{R}^2$ of Theorem 4, which will increase utility again. We investigate the utility difference between IR and CR experimentally in the next section.

Since

$$\mathrm{I}(\mathcal{R}^S(U); U|S) = \sum_s \mathbb{P}(S = s)\,\mathrm{I}_{U \sim \hat{P}_{\mathcal{U}|s}}(\mathcal{R}^s(U); U), \tag{16}$$

Theorem 6 tells us that we want to choose each $\mathcal{Q}^s$ to be the δ_s-LDP protocol for which $\mathrm{I}_{U \sim \hat{P}_{\mathcal{U}|s}}(\mathcal{R}^s(U); U)$ is maximised. For big enough δ_s, this is GRR. As such, the only unknown is again ε_2, reducing finding the optimal CR again to a one-dimensional optimisation problem.

V. Experiments

Throughout the experiments we take $\alpha = 0.05$. We first perform experiments to test the utility of the PolyOpt method introduced in Section III. We perform numerical experiments on synthetic data. For $|\mathcal{S}| = |\mathcal{U}| = 3$, we draw 200 distributions from the Jeffreys prior on the space of probability distributions on $\mathcal{X}$. For each distribution, we draw $n = 1000$ items from this distribution, and we demand robustness w.r.t. this observed distribution. For each observed distribution, for $\varepsilon \in [0.2, 8]$, and for each protocol of PolyOpt, GRR, IR, and CR, we calculate the normalised utility $\frac{\mathrm{I}_{\hat{P}}(X;Y)}{\mathrm{H}(X)}$, which we average over all distributions.

The results are in Figure 2. As we can see, PolyOpt significantly outperforms the other algorithms, although the optimisation method we used (we used Matlab, specifically the MPT3 toolbox) becomes more inaccurate at larger ε. However, the downside of Polyopt lies in its computation time, which is significantly higher than that of other methods. In general, if the user has enough computation power to use the PolyOpt method, then this is recommended, because it clearly outperforms all other protocols. However, it is possible that this is computationally unfeasible.

To investigate the role of n, we perform the same procedure as before, but for $|\mathcal{S}| = |\mathcal{U}| = 5$ and changing n, omitting

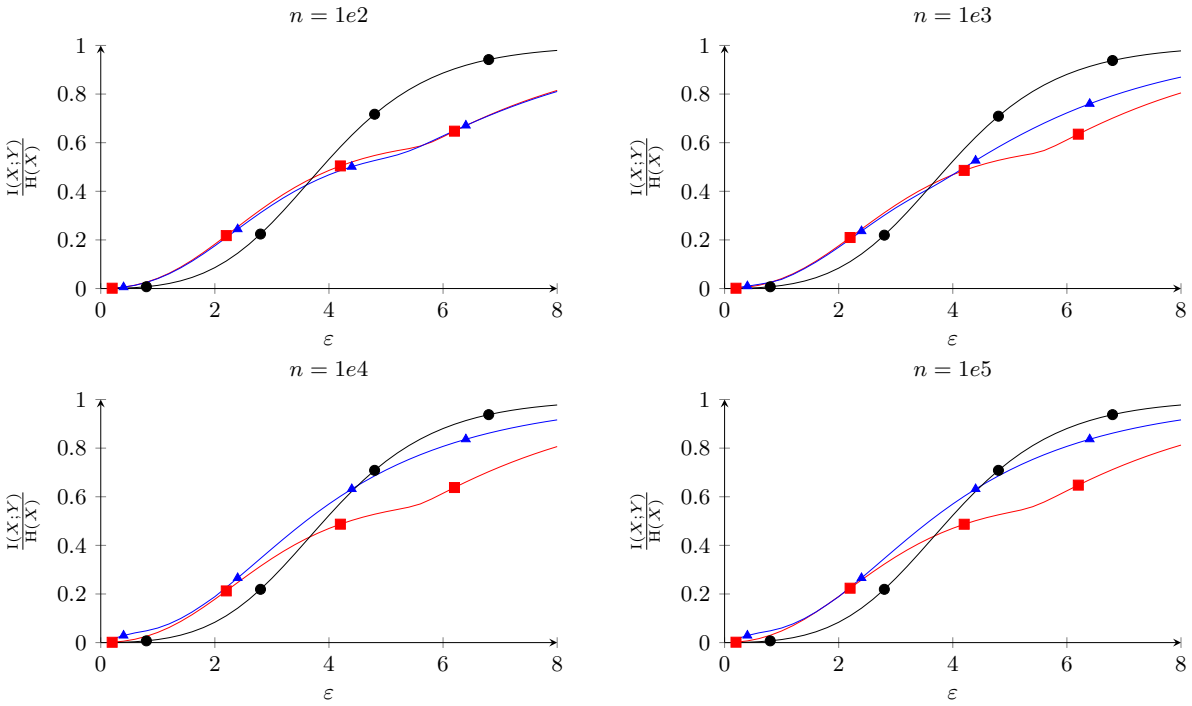

Fig. 3: Experiments on synthetic data with n changing, $|\mathcal{S}| = |\mathcal{U}| = 5$. (—●— GRR, —■— IR, —▲— CR)

PolyOpt for computability reasons. The results are in Figure 3. As is confirmed by Figure 2, straightforward GRR outperforms IR and CR for large enough ε, typically around $\varepsilon = 4$. However, DP-like metrics are often considered for $\varepsilon \approx 1$ [21], at which IR and CR clearly outperform GRR. Furthermore, CR performs better for larger n. This reflects the fact that CR's utility is greatly affected by the size of $\mathcal{F}$.

VI. Conclusion and Future Work

In this paper, we presented a number of algorithms that, given a desired privacy level ε, an estimated distribution $\hat{P}$, and a significance level α, return release protocols that aim to maximise the mutual information between input and output, while satisfying privacy w.r.t. a given sensitive part of the data, for all distributions in a $(1-\alpha)$-confidence set. One of these algorithms, PolyOpt, offers significantly higher utility, especially in the high privacy (low ε) regime. However, it relies on vertex enumeration, making it computationally infeasible for larger input spaces. Two of the other algorithms, IR and CR, rely on processing the sensitive and non-sensitive data separately. These algorithms rely on low-dimensional optimisation, independent of the size of the input space, allowing these to be used when PolyOpt is outside of the user's computational capabilities. IR and CR perform similar in the high privacy regime, with IR performing better for larger n. Furthermore, both of these perform significantly better than GRR. This shows the strength of the RLDP framework, as it shows that tailoring our algorithms to the uncertainty set $\mathcal{F}$ improves utility significantly.

Our results suggest several avenues for future research. First, one may want to incorporate not only robustness in privacy, but also in utility, i.e. to find the protocol $\mathcal{Q}$ that maximises $\min_{P\in\mathcal{F}} \mathrm{I}_P(X;Y)$. An obstacle for this is that $\mathrm{I}_P(X;Y)$ is concave in P, which makes finding its minimum over $\mathcal{F}$ difficult. Second, instead of looking at the situation where X splits into a sensitive part S and a non-sensitive part U, one can consider the more general case that X is correlated with the sensitive data S. This is already done in work on the privacy funnel [3][10], but this generally does not incorporate robustness. Furthermore, the utility of IR and CR might be improved in the high privacy regime by incorporating other LDP protocols than GRR. It is shown in [2] that GRR is the optimal LDP protocol for high ε, but for low ε the optimum typically takes a different form. One obstacle in incorporating this is that these optima depend on P^*, which is inaccessible in the RLDP framework.

Acknowledgements

This work was supported by NWO grant 628.001.026.

References

[1] D. Rebollo-Monedero, J. Forne, and J. Domingo-Ferrer, "From t-closeness-like privacy to postrandomization via information theory," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 11, pp. 1623–1636, 2010.

[2] P. Kairouz, S. Oh, and P. Viswanath, "Extremal mechanisms for local differential privacy," *arXiv:1407.1338*, 2014.

[3] A. Makhdoumi, S. Salamatian, N. Fawaz, and M. Médard, "From the information bottleneck to the privacy funnel," in *2014 IEEE Information Theory Workshop (ITW 2014)*, IEEE, 2014, pp. 501–505.

[4] S. Salamatian, A. Zhang, F. du Pin Calmon, S. Bhamidipati, N. Fawaz, B. Kveton, P. Oliveira, and N. Taft, "Managing your private and public data: Bringing down inference attacks against your privacy.," *J. Sel. Topics Signal Processing*, vol. 9, no. 7, pp. 1240–1255, 2015.

[5] S. Asoodeh, M. Diaz, F. Alajaji, and T. Linder, "Information extraction under privacy constraints," *Information*, vol. 7, no. 1, p. 15, 2016.

[6] S. Kung, "A compressive privacy approach to generalized information bottleneck and privacy funnel problems," *Journal of the Franklin Institute*, vol. 355, no. 4, pp. 1846–1872, 2018.

[7] N. Ding and P. Sadeghi, "A submodularity-based agglomerative clustering algorithm for the privacy funnel," *arXiv preprint arXiv:1901.06629*, 2019.

[8] S. Salamatian, F. P. Calmon, N. Fawaz, A. Makhdoumi, and M. Médard, "Privacy-utility tradeoff and privacy funnel," 2020, Preprint.

[9] J. C. Duchi, M. I. Jordan, and M. J. Wainwright, "Local privacy and statistical minimax rates," in *Foundations of Computer Science (FOCS), 2013 IEEE 54th Annual Symposium on*, IEEE, 2013, pp. 429–438.

[10] M. Lopuhaä-Zwakenberg, H. Tong, and B. Škorić, "Data sanitisation for the privacy funnel with differential privacy guarantees," *arXiv:2008.13151*, 2020.

[11] D. Kifer and A. Machanavajjhala, "Pufferfish: A framework for mathematical privacy definitions," *ACM Transactions on Database Systems (TODS)*, vol. 39, no. 1, pp. 1–36, 2014.

[12] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust optimization*. Princeton University Press, 2009, vol. 28.

[13] A. Ben-Tal, D. Den Hertog, and J.-P. Vial, "Deriving robust counterparts of nonlinear uncertain inequalities," *Mathematical programming*, vol. 149, no. 1-2, pp. 265–299, 2015.

[14] D. Bertsimas, V. Gupta, and N. Kallus, "Data-driven robust optimization," *Mathematical Programming*, vol. 167, no. 2, pp. 235–292, 2018.

[15] S. L. Warner, "Randomized response: A survey technique for eliminating evasive answer bias," *Journal of the American Statistical Association*, vol. 60, no. 309, pp. 63–69, 1965.

[16] M. Diaz, H. Wang, F. P. Calmon, and L. Sankar, "On the robustness of information-theoretic privacy measures and mechanisms," *IEEE Transactions on Information Theory*, vol. 66, no. 4, pp. 1949–1978, 2019.

[17] H. Wang, M. Diaz, F. P. Calmon, and L. Sankar, "The utility cost of robust privacy guarantees," in *2018 IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2018, pp. 706–710.

[18] J. Liao, O. Kosut, L. Sankar, and F. du Pin Calmon, "Tunable measures for information leakage and applications to privacy-utility tradeoffs," *IEEE Transactions on Information Theory*, vol. 65, no. 12, pp. 8043–8066, 2019.

[19] A. Makhdoumi and N. Fawaz, "Privacy-utility tradeoff under statistical uncertainty," in *2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, IEEE, 2013, pp. 1627–1634.

[20] M. Lopuhaä-Zwakenberg and J. Goseling, "The privacy-utility tradeoff of robust local differential privacy," *arXiv:2101.09139*, 2021.

[21] C. Dwork, "A firm foundation for private data analysis," *Communications of the ACM*, vol. 54, no. 1, pp. 86–95, 2011.