Online Local Differential Private Quantile Inference via Self-normalization

Yi Liu [1]   **Qirui Hu** [2]   **Lei Ding** [1]   **Linglong Kong** [1]

Abstract

Based on binary inquiries, we developed an algorithm to estimate population quantiles under Local Differential Privacy (LDP). By self-normalizing, our algorithm provides asymptotically normal estimation with valid inference, resulting in tight confidence intervals without the need for nuisance parameters to be estimated. Our proposed method can be conducted fully online, leading to high computational efficiency and minimal storage requirements with $\mathcal{O}(1)$ space. We also proved an optimality result by an elegant application of one central limit theorem of Gaussian Differential Privacy (GDP) when targeting the frequently encountered median estimation problem. With mathematical proof and extensive numerical testing, we demonstrate the validity of our algorithm both theoretically and experimentally.

1. Introduction

Personal data is currently widely used for various purposes, such as facial recognition, personalized advertising, medical trials, and recommendation systems to name a few. While there are potential benefits, it is important also to consider the risks associated with handling sensitive personal information. For instance, research on diabetes can provide valuable insights that may benefit society as a whole in the long term. However, it is crucial to keep in mind that participants may suffer direct consequences if their data is not properly protected through controlled disclosure, such as a rise in health insurance premiums.

The concept of Differential Privacy (DP; Dwork et al., 2006) has been successful in providing a rigorous condition for controlled disclosure by bounding the change in the distribution of outputs of a query made on a dataset under the alteration of one data point. This has led to a vast amount of literature under the umbrella of DP, resulting in various generalizations, tools, and applications. However, while enjoying the mathematically solid guarantee of DP and its variants, concerns about a weak link in the process, the trusted curator, are beginning to arise.

The use of trusted curators undermines the spirit of the solid cryptographic level of privacy protection that DP provides. This risk is not limited to information security breaches and rogue researchers but also includes legal proceedings where researchers may be compelled to hand over the raw data, breaking the initial promise made to DP at the time of data collection. Two concepts, Local Differential Privacy (LDP) and pan-DPs, are proposed as solutions. The pan-DP directly counters this issue by solidifying the algorithm to withstand multiple announced intrusions (subpoenas) or one unannounced intrusion (hackers). The concept of LDP was first introduced formally by (Kasiviswanathan et al., 2011), but its early form can be traced back to (Evfimievski et al., 2003) and (Warner, 1965), in the name of "amplification" and "randomized response survey," respectively.

In LDP settings, the sensitive information never leaves the control of the users unprotected. The users encode and alter their data locally before sending them to an untrusted central data server for further analysis and computation. Recently, in (Amin et al., 2020) unveiled a connection between pan-DP and LDP by considering variants of pan-DP framework that can defend against multiple unannounced intrusions. Surprisingly, this requirement can only be fulfilled if the data is scrambled before it leaves the owner's control, which goes back to the definition of LDP. For better privacy protection, many big tech companies have already implemented LDP into their products, such as Google (Erlingsson et al., 2014) and Microsoft (Ding et al., 2017).

This discovery rekindled the research interest in LDP. Researchers have begun to consider fundamental statistical problems, such as estimating parameters, modeling, and hypothesis testing under this constraint. The quantiles, including the median, are basic summary statistics that have been widely studied within the framework of differential privacy. Early research in this area includes the estimation of quantiles under the central DP setting, as presented in (Dwork

[1]Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Canada [2]Center for Statistical Science, Department of Industrial Engineering, Tsinghua University, Beijing, China. Correspondence to: Linglong Kong <lkong@ualberta.ca>.

& Lei, 2009) and (Lei, 2011). More recent advancements, such as (Smith, 2011), have proposed a rate-optimal sample quantile estimator that does not rely on the evaluation of histograms. (Gillenwater et al., 2021) further extended this research by estimating multiple quantiles simultaneously. Despite these advances, the quantile estimation under the central DP setting remains an active area of research, with new work in various applications such as (Alabi et al., 2022) and (Ben-Eliezer et al., 2022).

In the central DP setting, a trusted curator can acquire the actual sample quantiles and other summary statistics, with the only limitation being that the release of the output must conform to the DP condition. However, under the local DP setting, the curator does not have access to the true data and can only see proxies generated by the users. This makes it more challenging to design local DP algorithms that can provide valid results leading to greater problems in developing corresponding theoretical properties and providing further statistical inference.

Researchers often propose consistent estimators for the parameters of interest and derive the asymptotic normality. However, these estimators often involve nuisance parameters that are not trivial to obtain or estimate, making them difficult to deploy in real-world scenarios. To address this issue, (Shao, 2010) developed the methodology of self-normalization for constructing confidence intervals. This method involves designing a statistic called the self-normalizer, which is proportional to the nuisance parameters, and making the original estimate a pivotal quantity by placing it and the self-normalizer in the numerator and the denominator, thereby canceling out the nuisance parameters and leading to an asymptotically pivotal distribution. This methodology provides a powerful tool for statistical inference under complex data, particularly in the context of LDP frameworks where obtaining accurate original data or consistently estimating nuisance parameters without an additional privacy budget is challenging.

Efficient computation is essential for the practicality of LDP algorithms, as large sample sizes are necessary to counteract the effects of local perturbations and achieve optimal performance. Meanwhile, online computation is another valuable attribute of LDP algorithms, as it reduces storage requirements and diminishes risks associated with information storage. Early attempts of introduce online computation to DP algorithms can be traced back to (Jain et al., 2012), where additive Gaussian noise was injected into the gradient to provide DP protection. Later, (Agarwal & Singh, 2017) gives an online linear optimization DP algorithm that with optimal regret bounds. The concept of online computation has also been incorporated into federated learning, as discussed by (Wei et al., 2020). More recently, (Lee et al., 2022) has facilitated online computation for a random

scaling quantity using only the trajectory of stochastic optimization, effectively eliminating the need for past state storage and enhancing computational efficiency. In contrast to traditional studies on DP online algorithms, our emphasis is on harnessing online computation for convenience. Our theoretical analysis concentrates on the statistical properties of the proposed estimators, encompassing aspects such as consistency, asymptotic normality, and more.

In this paper, our contributions are listed as follows.

- We propose a new LDP algorithm for population quantile estimation that does not require a trusted curator. Under some mild conditions, we derive the consistency and asymptotic normality of the proposed quantile estimator.

- We construct the confidence interval of the population quantiles via self-normalization, which eliminates the need for estimating the asymptotic variance in the limiting distribution. Furthermore, this procedure can be implemented online without storing all past statuses.

- We also discuss the optimality of the proposed algorithm. By combining it with the central limit theorem of GDP, we demonstrate that our algorithm for median estimation achieves the lower bound of asymptotic variance among all median estimators constructed by a binary random response-based sequential interactive mechanism under LDP.

The structure of this paper is as follows. We begin by providing an overview of the concepts of central DP and LDP. Then present our proposed methodology, detailing the algorithms and their corresponding theoretical results. Finally, we provide experimental results to demonstrate the effectiveness of our approach.

2. Preliminaries

2.1. Central Differential Privacy

Definition 2.1. (Dwork et al., 2006) A randomized algorithm $\mathcal{A}$, taking a dataset consisting of individuals as its input, is (ϵ, δ)-differentially private if, for any pair of datasets S and S' that differ in the record of a single individual and any event E, satisfies the below condition:

$$\mathbb{P}[\mathcal{A}(S) \in E] \leq e^{\epsilon} \mathbb{P}\left[\mathcal{A}\left(S'\right) \in E\right] + \delta.$$

When $\delta = 0$, $\mathcal{A}$ is called ϵ-Differentially Private (ϵ-DP).

The concept of DP only imposes constraints on the output distribution of an algorithm $\mathcal{A}$, rather than placing restrictions on the credibility of the entity running the algorithm or

protecting the internal states of $\mathcal{A}$. The existence of the curator who has access to the raw data set is why this approach is known as "Central" DP. The curator simplifies the algorithm design and often leads to an asymptotically negligible loss of accuracy from privacy protection (Cai et al., 2021).

2.2. Local Differential Privacy

Despite the varying definitions of LDP due to the level of interactions, all of them depend on the following concept called (ϵ, δ)-randomizer.

Definition 2.2. (Joseph et al., 2019) An (ϵ, δ)-randomizer $R : X \rightarrow Y$ is an (ϵ, δ)-differentially private function taking a single data point as input.

The definition of randomizer is mathematically a special case of the central DP. The main difference between the central and local DP is the role of the curator, which is further determined by the level of interactions allowed. In LDP, the curator coordinates interactions between n users, each of whom holds their own private information X_i. In each round of interaction, the curator selects a user and assigns them a randomizer R_t. If the (ϵ, δ) parameters are allowed by the experiment setting, the user will run the randomizer on their private information and release the output to the curator. The level of interactions can vary from full-interactive, where the curator can choose the randomizer and the next user based on all previous interactions, to sequential (also called one-shot) interactive, where the curator is not allowed to pick one user twice but is still able to adaptively picking the next the user-randomizer pairs based on all previous interactions, to non-interactive, where adaptivity is forbidden, and all user-randomizer pairs must be determined before any information is collected. If the curator is further forbidden from varying the randomizer R and tracking back outputs to a specific user, it will lead to another interesting setting called shuffle-DP. In this paper, our algorithm complies with the $(\epsilon, 0)$ sequentially interactive Local LDP conditions. This is because our inquiries and internal states are driven by randomized responses from users, while no user is queried more than once. For a more detailed information, please refer to Definitions 2.3 and 2.6 in (Cheu et al., 2019).

2.3. Notations

In this paper, we employ the following notations. $\mathbf{1}_{\{\cdot\}}$ is the indicator function and $[a]$ denotes the largest integer that does not exceed a. $\mathcal{O}$ (or $\mathpzc{o}$) denotes a sequence of real numbers of a certain order. For instance, $\mathpzc{o}(n^{-1/2})$ means a smaller order than $n^{-1/2}$, and by $\mathcal{O}_{a.s.}$ (or $\mathpzc{o}_{a.s.}$) almost surely $\mathcal{O}$ (or $\mathpzc{o}$). For sequences a_n and b_n, denote $a_n \asymp b_n$ if there exist positive contants c and C such that $cb_n \leq a_n \leq Cb_n$. The symbol $\xrightarrow{d}$ means weak convergence

or converge in distribution.

3. Algorithm and Main Results

3.1. Algorithm

Let $x_1, \ldots, x_n, \ldots$ be independently and identically distributed(i.i.d.) random variables defined on $\mathbb{R}$ representing private information of each user, with target quantile τ and corresponding true value Q, i.e., $\mathbb{P}(x_i \leq Q) = \tau$. To ensure the uniqueness of quantiles, we assume the x_i's are continuous random variables, with positive density on the target quantile. In practice, we can perturb the data by a small amount of additive data-independent noise to remove atoms in the distribution as is in (Gillenwater et al., 2021).

The design of the local randomizer is crucial for LDP mechanisms as it must properly choose the inquiry to the user in order to maximize the gathering of information related to the estimation of the target quantile without violating privacy conditions. The population quantiles can be considered as a minimizer of the check loss function:

$$l_\tau(x, \theta) = \begin{cases} \tau x & \text{if } x \geq \theta \\ (\tau - 1)x, & \text{if } x < \theta \end{cases}.$$

In the non-DP case, a known solution is the use of stochastic gradient descent, as outlined in (Joseph & Bhatnagar, 2015). It is important to note that for each point, the gradient it contributes is purely determined by the binary variable representing whether the value is greater than θ or not. This motivates us to modify the stochastic gradient descent process by adding a local randomization process, resulting in the Algorithm 1 and 2 outlined below:

Algorithm 1 Locally Randomized Compare (LRC)

Input: Inquiry q, response rate r, private data x
$u \sim Bernoulli(r)$
$v \sim Bernoulli(0.5)$
if $u = 1$ **then**
    return $\mathbf{1}_{x>q}$
else
    return v
end if

In Algorithm 1, generating randomness of v before the if-condition fork may seem wasteful, but it prevents side-channel attacks such as inferring the true value based on the timing of response (Coppens et al., 2009; Lawson, 2009). Algorithm 2 collects random responses and generates the next inquiry accordingly. Therefore, Algorithm 2 satisfies the definition of sequential interactive local DP.

The following algorithm can be used when estimations and

Algorithm 2 Main Algorithm

Input: Step sizes d_n, target quantile $\tau \in (0,1)$, truthful response rate r
Initialize: $n \leftarrow 0$, $q_0 \leftarrow 0$, $v_0^a \leftarrow 0$, $v_0^b \leftarrow 0$, $Q_0 \leftarrow 0$
repeat
   $n \leftarrow n + 1$
   Inquire: $s \leftarrow LRC(q_{n-1}, r, x_n)$
   if s is 1 **then**

$$q_n \leftarrow q_{n-1} + \frac{1 - r + 2\tau r}{2} d_n$$

   else

$$q_n \leftarrow q_{n-1} - \frac{1 + r - 2\tau r}{2} d_n$$

   end if
   $Q_n = ((n-1)Q_n + q_n)/n$
   $v_n^a \leftarrow v_{n-1}^a + n^2 Q_n^2$
   $v_n^b \leftarrow v_{n-1}^b + n^2 Q_n$
   Destroy $v_{n-1}^a, v_{n-1}^b, Q_{n-1}, q_{n-1}$
until Forever

---

confidence intervals are required. These values are not calculated at every step to minimize computational expenses.

Algorithm 3 Generate Confidence Interval

Input: Internal states of Algorithm 2: n, Q_n, v_n^a, v_n^b
$N_n \leftarrow n^{-1} \left(v_n^a - 2Q_n v_n^b + Q_n^2 n(n+1)(2n+1)/6 \right)$
$W \leftarrow n^{-1} \mathcal{U}_{1-\alpha/2} \sqrt{N_n}$
Return: Confidence interval $(Q_n - W, Q_n + W)$

The use of dichotomous inquiry in data privacy brings multiple advantages. One benefit is the reduced communication cost, as it only takes one bit to respond. Additionally, the binary response can make full use of the DP budget, as opposed to methods such as the Laplace mechanism, which may provide unnecessary privacy guarantees beyond ϵ-DP, as outlined in Theorem 3 in (Balle et al., 2018) and Theorem 2.1 in (Liu et al., 2022).

Furthermore, people tend to be more comfortable answering dichotomous questions compared to open-ended ones (Brown et al., 1996) as they present a choice between two options and may be perceived as less threatening than open-ended questions, which require more detailed and nuanced responses. In addition, the binary approach is easy to understand for users. With the proper choice of truthful response rate r, the algorithm known as the random response can be easily simulated through coin flips or dice rolls, allowing users to understand it fully and are able to "run" it without the help of electronic devices. This is in contrast to a DP mechanism involving the usage of random distribution on

real numbers. Due to the finite nature of the computer, the imperfection of floating-point arithmetic leads to serious risks with effective exploits. For more information, please refer to (Mironov, 2012; Jin et al., 2021; Haney et al., 2022).

Before discussing the specific characteristics of our estimator, we will first demonstrate its performance through a sample trajectory. The experiment is conducted with a truthful response rate with $r = 0.5$, which means half of the responses are purely random. The objective is to estimate the median from i.i.d. samples. The true underlying distribution is a standard normal distribution.

It can be seen that from Figure 1, the proposed estimator converges to the true value, and both infeasible and proposed confidence intervals, defined later, contain the true value at a slightly larger sample size. Also, the proposed confidence intervals are highly competitive with the infeasible one in width. Refer to Figure 5 and 6 for convergence trajectories under different initialization or target quantiles.
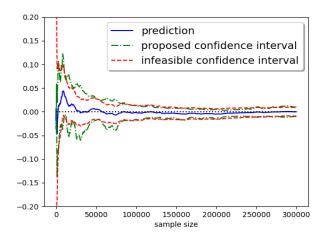

Figure 1. A sample trajectory of estimator Q_n, infeasible confidence interval (2) and proposed confidence interval (3). The horizon dotted line is the true value $Q = 0$.

Next, we show the LDP property of our algorithm:

Theorem 3.1. *Algorithm 1 is an $(\epsilon, 0)$-randomizer with $\epsilon = \log((1+r)/(1-r))$.*

Proof. see Appendix C.1 ☐

The algorithm presented in Algorithm 2 adaptively selects the next randomizer, determined by the parameter q in Algorithm 1, based on its internal state q_n. However, it never revisits previous users. As a result, Algorithm 2 satisfies sequential interactive $(\epsilon, 0)$-LDP, where $\epsilon = \log\left((1+r)/(1-r)\right)$ (equivalently, $r = (e^\epsilon - 1)/(e^\epsilon + 1) = \tanh\left(\epsilon/2\right)$).

Throughout the remainder of this paper, we will use the truthful response rate r to represent the privacy budget, as opposed to the more standard ϵ. This choice is made for the following reasons:

In the context of LDP, it is crucial to ensure understanding and acceptance by end-users who may not possess expertise in the field. The truthful response rate, denoted by r, has a more intuitive interpretation. Additionally, r appears in multiple results presented in this paper, and maintaining this form allows for a more direct presentation. If necessary, the results can be easily converted by replacing all instances of r with $\tanh(\epsilon/2)$. For a conversion table, please refer to Table 5.

3.2. Consistency

To discuss the asymptotic properties of estimator Q_n, we rewrite it as a recursive equation. Let $\{U_n\}$ and $\{V_n\}$ be the i.i.d. Bernoulli sequences with

$$
\mathbb{P}(U_n = 1) = r, \ \mathbb{P}(U_n = 0) = 1 - r,
$$
$$
\mathbb{P}(V_n = 1) = \mathbb{P}(V_n = 0) = 1/2.
$$

For $q_0 \in \mathbb{R}$,

$$
q_{n+1} = q_n + \frac{1 - r + 2r\tau}{2d_n} \left(\mathbf{1}_{x_{n+1} > q_n} U_n + (1 - U_n)V_n \right)
$$
$$
- \frac{1 + r - 2r\tau}{2d_n} \left(\mathbf{1}_{x_{n+1} < q_n} U_n + (1 - U_n)(1 - V_n) \right),
$$
(1)

where the step size $\{d_n\}_{n=1}^{\infty}$, satisfies

$$
\sum_{n=1}^{\infty} d_n = \infty, \qquad \sum_{n=1}^{\infty} d_n^2 < \infty.
$$

The step size d_n is vital for the convergence of q_n, but it has a relatively minor effect on Q_n. The following theorem guarantees consistency:

Theorem 3.2. *For increasing positive number γ_n, satisfied*

$$
\frac{\gamma_n}{\gamma_{n-1}} = 1 + \mathcal{O}(d_n), \qquad \sum_{n=1}^{\infty} d_n^2 \gamma_n^2 < \infty,
$$

the n-step output q_n enjoys that

$$
\gamma_n |q_n - Q| = \mathcal{O}_{a.s.}(1).
$$

Proof. see Appendix C.2. □

In particular, if $d_n \asymp a/n^{\beta}$, for some constant $a > 0$ and $\beta \in (1/2, 1)$, then $\gamma_n \asymp n^{\gamma}$ for some $\gamma < \beta - 1/2$, and for the sake of simplicity, we will set the step sizes as $d_n \asymp a/n^{\beta}$.

3.3. Asymptotic Normality

Next, the asymptotic normality will be discussed.

Theorem 3.3. *If $\beta \in (0, 1)$, then*

$$
\sqrt{n}\,(Q_n - Q) \xrightarrow{d} N\left(0, \frac{1 - r^2(1 - 2(1 - \tau))^2}{4r^2 f_X^2(Q)}\right),
$$

where $f_X^2(Q)$ is the value on Q for density function of X.

Proof. see Appendix C.2. □

Noticed that the conditions on β in Theorem 3.2 and Theorem 3.3 are different. It is possible that q_n fails to converge to Q, but Q_n still enjoys asymptotic normality. Following Theorem 3.3, one constructs the confidence interval of Q, if $f_X(Q)$ can be obtained or estimated by $\widehat{f_X(Q)}$. Denote $z_{1-\alpha}$ as the upper $\alpha-$quantile of standard normal distribution. The infeasible confidence interval with significance level α is:

$$
\Big(Q_n - z_{1-\alpha}\sqrt{n(1 - r^2(1 - 2(1 - \tau))^2)}/(2r\widehat{f_X(Q)}),
$$
$$
Q_n + z_{1-\alpha}\sqrt{n(1 - r^2(1 - 2(1 - \tau))^2)}/(2r\widehat{f_X(Q)})\Big).
$$
(2)

However, obtaining a consistent estimator $\widehat{f_X(Q)}$, such as using non-parametric methods under our differential privacy framework, is not straightforward, since we can only obtain the binary sequence $\mathbf{1}_{x_n > q_{n-1}}$ for protecting privacy, and the original data set $x_1, \ldots, x_n$ cannot be accessed directly.

An alternative approach to estimate the nuisance parameter $f_X(Q)$ is through the use of bootstrap methods to simulate the asymptotic distribution. Traditional bootstrap methods that rely on re-sampling are not suitable for the stochastic gradient descent method because of failing to recover the special dependence structure defined in (1).

Recently, (Fang et al., 2018) proposed online bootstrap confidence intervals for stochastic gradient descent, which involve recursively updating randomly perturbed stochastic estimates. Although this approach performs well when there are no constraints on DP, it requires multiple interactions with the users and will therefore blow up the privacy budget.

3.4. Inference via Self-normalization

To overcome the difficulties above, we propose a novel inference procedure of quantiles under the LDP framework via self-normalization, which will avoid estimating the nuisance parameter $f_X(Q)$. We hope to construct an estimator that is proportional to the nuisance parameters. To approach that, we will first establish further theoretical properties of the proposed estimator Q_n. Define the process $S_{[nt]} = \sum_{i=1}^{[nt]} q_i, t \in [0, 1]$.

Theorem 3.4. *If $\beta \in (0, 1)$, then*

$$n^{-1/2}(S_{[nt]} - nQ) \xrightarrow{d} \frac{\sqrt{(1 - r^2(1 - 2(1 - \tau))^2)}}{2rf_X(Q)}W(t),$$

where $W(t)$ is the Brownian motion in $(C[0, 1], \mathbb{R})$.

Proof. see Appendix C.2. □

Noticed that Theorem 3.3 is the special case in Theorem 3.4 when $t = 1$. Then, following Theorem 3.4, we define the self-normalizer:

$$N_n = \int_0^1 \left(S_{[nt]} - [nt]Q_n\right)^2 dt,$$

By the continuous mapping theorem, we can derive:

$$\frac{n^{-1/2}(S_n - nQ)}{\sqrt{n^{-1}N_n}} \xrightarrow{d} \mathcal{S} := \frac{W(1)}{\sqrt{\int_0^1 (W(t) - tW(1))^2 \, dt}},$$

where the asymptotical distribution $\mathcal{S}$ is not associated with any unknown parameters, and its quantile can be computed by Monte Carlo simulation. Therefore, we have constructed an asymptotical pivotal quantity. Denote $\mathcal{U}_{1-\alpha}$ the $1 - \alpha$ quantile of $\mathcal{S}$, the $1 - \alpha$ self-normalized confidence interval of Q is constructed by:

$$\left(Q_n - n^{-1}\mathcal{U}_{1-\alpha/2}\sqrt{N_n}, Q_n + n^{-1}\mathcal{U}_{1-\alpha/2}\sqrt{N_n}\right). \quad (3)$$

As noted by (Shao, 2015), the distribution of $\mathcal{S}$ has a heavier tail than that of the standard normal distribution, which is analogous to the heavier tail of $t-$distribution compared to the standard normal distribution, resulting in a wider but not conservative corresponding confidence interval. However, the average width of the confidence interval constructed through self-normalization is not excessively large when compared to the infeasible confidence interval, as demonstrated by numerical experiments in Figure 1. Furthermore, the construction of an asymptotic pivotal quantity is not unique. See Appendix B for examples of other possibilities.

Whether there are theoretical advantages between the different constructions of self-normalizer is still open to discussion, but according to (Lee et al., 2022), the proposed self-normalizer can be computed in a fully online fashion and is computationally efficient, as outlined in Algorithm 2 and 3. The algorithm only needs to store a single integer n and four float numbers: v_n^a, v_n^b, q_n, Q_n and conduct only a dozen of arithmetic operations.

3.5. Discussion of Optimality

In this subsection, we will discuss the optimality of the proposed algorithm. To generalize the setting, we consider all

binary random response-based sequential interactive mechanisms. The random response mechanism can be written as the following $K : \{0, 1\} \rightarrow \{0, 1\}$:

$$K(x) = \begin{cases} 0, & \text{w.p. } (1-r)/2, \\ 1, & \text{w.p. } (1-r)/2, \\ x, & \text{w.p. } r. \end{cases}$$

Let $\{T_1, \cdots, T_n\}$ be a collection of binary query functions, which means $T_i(x) = \mathbf{1}_{x \in C_i}$, for some subset $C \subset \mathbb{R}$. In the sequential interactive LDP setting, the curator will generate its output based on the transcript $\{\{K \circ T_1(x_1), \cdots, K \circ T_n(x_n)\}, \{C_1, \cdots, C_n\}\}$ and the choice of C_i may depend on the transcript up to this point:$\{\{K \circ T_1(x_1), \cdots, K \circ T_{i-1}(x_{i-1})\}, \{C_1, \cdots, C_{i-1}\}\}$. Notice that the Algorithm 1 is a special case where $C_i = \{z : z \geq q_{i-1}\}$, and q_{i-1} is given by

$$\sum_{j=1}^{i-1} T_j(x_j)\frac{1 - r + 2\tau r}{2}d_j - (1 - T_j(x_j))\frac{1 + r - 2\tau r}{2}d_j.$$

We aim to determine a lower bound for the estimation variance. Therefore, any lower bounds derived under specific conditions also serve as a general lower bound for the estimation variance. To demonstrate this, we will present a pair of distributions with distinct medians that are, to the best of our knowledge, the most indistinguishable given randomized binary queries.

Define:

$$H_0 : x_i \sim Laplace(1) \text{ vs. } H_1 : x_i \sim Laplace(1) + \epsilon_n \quad (4)$$

Let $\epsilon = \log\left[\left(e^{\frac{1}{\sqrt{n}}}(r+1) + r - 1\right)/\left(e^{\frac{1}{\sqrt{n}}}(r-1) + r + 1\right)\right]$. Simple computation yields that for any $(a, b) \in \{0, 1\}^2$

$$\frac{\mathbb{P}(K \circ T_i(x_i) = a | H_b)}{\mathbb{P}(K \circ T_i(x_i) = a | H_{1-b})} \leq \frac{e^{\epsilon_n}(r+1) - r + 1}{-e^{\epsilon_n}(r-1) + r + 1} = e^{\sqrt{\frac{1}{n}}}. \quad (5)$$

Interestingly, if we consider the truth $H \in \{H_0, H_1\}$ as a data set containing only one data point, (5) shows that $K \circ T_i$ is $1/\sqrt{n}$-DP. Notice that the transcript is a n-fold adaptive composition (Kairouz et al., 2015) of $1/\sqrt{n}$-DP mechanisms. By Theorem 8 (Dong et al., 2021a), the transcript and all post-processing of it (Proposition 4; (Dong et al., 2021a)) asymptotically satisfies the Gaussian Differential Privacy condition with $\mu = 1$ (or briefly 1-GDP).

We will now examine the limit on the best possible variance imposed by the 1-GDP condition. Denote the estimator of median as $\hat{\theta}_n$. First, we will consider asymptotically normal, unbiased, shift-invariant estimators of the median. By

restricting our discussion to unbiased, shift-invariant estimators, we ensure that no estimator has an unfair advantage by favoring specific values. Under the null hypothesis, for the standard deviation σ_n of $\hat{\theta}_n$, one has that

$$\frac{\hat{\theta}_n}{\sigma_n} \xrightarrow{d} N(0,1),$$

and under the alternative hypothesis,

$$\frac{\hat{\theta}_n - \epsilon_n}{\sigma_n} \xrightarrow{d} N(0,1).$$

The 1-GDP condition implies that for sufficiently large n, $\epsilon_n/\sigma_n \leq 1$ (see Appendix C.3). By plugging in the values $\epsilon_n = (r\sqrt{n})^{-1} + \mathcal{O}(n^{-3/2})$ and $1/2 = f(F^{-1}(1/2))$, we deduce that:

$$\sigma_n \geq \frac{1}{2r\sqrt{n}f(F^{-1}(1/2))} + \mathcal{O}\left(n^{-1}\right),$$

which gives us an asymptotic lower bound of the variance: $(4r^2 n f^2(F^{-1}(1/2)))^{-1}$. This lower bound matches the asymptotic variance obtained in Theorem 3.3, showing the optimality of our approach. Although most estimators we are interested in have an asymptotically normal distribution, we wish to generalize the minimal variance result to other families as the theorem below.

Theorem 3.5. *If $\hat{\theta}_n$ is a median estimator based on the random response of binary-based sequential interactive inquiries such that:*

$$\frac{\hat{\theta}_n - F^{-1}(1/2)}{\sigma_n} \xrightarrow{d} G$$

where G has a log-concave density $f_G(x) \propto \mathrm{e}^{-\varphi(x)}$ on $\mathbb{R}$ such that $\varphi(x) = \varphi(-x)$, $\mathbb{E}\left[(\varphi'(G))^2\right] < +\infty$, and $\mathbb{E}\left[G^2\right] = 1$.

Then,

$$\sigma_n \geq \frac{1}{2r\sqrt{n}f(F^{-1}(1/2))} + \mathcal{O}\left(n^{-1}\right).$$

The minimal variance result can be attributed to two factors. In Appendix C.3, we demonstrate that asymptotic GDP imposes a condition on the variance of estimators that follow a normal distribution. This condition serves as a lower bound for 1-GDP estimators, without relying on any specific mechanism assumption. Secondly, the relaxation from the assumption of normality to milder conditions on the function G is a consequence of Theorem 1.2 in (Dong et al., 2021b). This theorem establishes that among all μ-GDP estimators satisfying the aforementioned conditions, the variance is lower bounded by $1/\mu^2$. This lower bound is attainable when the underlying distribution is normal.

4. Experiments

We evaluate the performance of our algorithms using a variety of distributions. The data come from four cases: standard Normal $N(0,1)$, Uniform $U(-1,1)$, standard Cauchy $C(0,1)$, and PERT distribution (Clark, 1962) with probability density function:

$$f(x) = 0.625(1-x)(1+x)^3, \quad x \in (-1,1).$$

These cases represent situations with heavy tails, compact or non-compact support, and asymmetric distributions commonly found in practice, as shown in Figure 2.
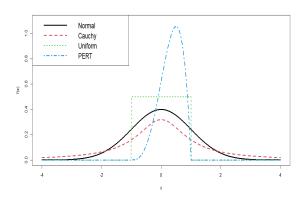

Figure 2. Plot of the density function, where the types of lines represent different distribution, solid: Normal, dashed: Cauchy, dotted: Uniform, dot-dash: PERT.

The target quantiles are $\tau = 0.3, 0.5, 0.8$, and the truthful response rate $r = 0.25, 0.5, 0.9$, which the privacy budget is $\epsilon = \log(1 + 2r/(1-r))$ corresponding to $0.51, 1.09, 2.94$ respectively. We use the step sizes $d_n = 2/(n^{0.51}+100)$ for all experiments, which satisfies the assumptions of Theorem 3.3 and 3.4. The range of sample size n is $(10000, 400000)$, the initial value $q_0 = 0$, and the number of replication is 10000. The results from different sample sizes are independently conducted from scratch to eliminate the correlation among experiments.

To show the consistency of the proposed estimator Q_n, Figure 3 displays the box plots of estimator Q_n under Normal distribution with sample size $n = 10000, \ldots, 50000$. As the sample size increases, the estimation becomes closer to the true values Q, the corresponding standard errors decay across all settings, and the truthful response rate leads to significantly better performance in small finite sample sizes but has diminishing effects afterward. Meanwhile, we can also see that the proximity between the true target value and the initialization 0 is beneficial to early performances. But in an asymptotic view, the proposed algorithm is insensitive to the initial value selection.

We also demonstrate the empirical coverage rate and mean

absolute error of the developed method in Table 1. The empirical coverage rate of the proposed method becomes closer to the nominal confidence level as the sample size increases in most cases and the mean absolute error decreases to zero. The corresponding figures and tables of other distributions can be found in Appendix A, which describes a similar phenomenon.

Figure 4 investigates the performance of the proposed confidence interval in other nominal levels. One can discover that the curves of the empirical coverage rate are getting closer to $y = x$ uniformly, as sample size increases in all privacy budget settings, which reveals the performance of the proposed method is irrelevant to the pre-determined significance level. It is worth noting that when $r = 0.25$, the effective sample size is $1/16$ of the original one, yet the performance of the proposed method remains excellent, which strongly supports the asymptotic theory.
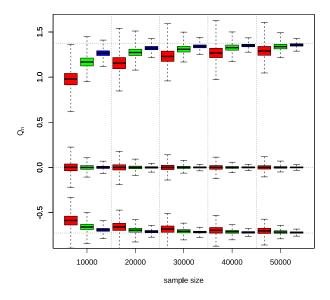

(a) Left: $n = 10000$. Right: $n = 50000$

(b) Left: $n = 100000$. Right: $n = 200000$

Figure 4. The curve of the empirical coverage rate of proposed confidence interval (3) with nominal significance level, when the data are Normal and target quantile $\tau = 0.3$ under different privacy budget (dotted $r = 0.25$, dot-dash $r = 0.5$ and dashed $r = 0.9$).

Table 1. Empirical results of coverage rate(mean absolute error) of proposed confidence interval (3) (estimator Q_n) with data collected from Normal.

n	τ	$r = 0.25$	$r = 0.5$	$r = 0.9$
	0.3	0.926(0.069)	0.965(0.034)	0.982(0.018)
10000	0.5	0.834(0.037)	0.897(0.019)	0.911(0.011)
	0.8	0.962(0.121)	0.992(0.058)	0.999(0.031)
	0.3	0.936(0.041)	0.958(0.020)	0.971(0.011)
20000	0.5	0.888(0.027)	0.915(0.014)	0.936(0.008)
	0.8	0.965(0.063)	0.984(0.030)	0.994(0.016)
	0.3	0.943(0.025)	0.958(0.013)	0.967(0.007)
40000	0.5	0.910(0.020)	0.931(0.010)	0.937(0.006)
	0.8	0.966(0.035)	0.978(0.017)	0.984(0.009)
	0.3	0.946(0.015)	0.954(0.007)	0.958(0.004)
100000	0.5	0.929(0.013)	0.944(0.006)	0.941(0.004)
	0.8	0.954(0.019)	0.965(0.009)	0.973(0.005)
	0.3	0.947(0.010)	0.951(0.005)	0.956(0.003)
200000	0.5	0.942(0.009)	0.949(0.004)	0.947(0.002)
	0.8	0.956(0.013)	0.960(0.006)	0.964(0.003)
	0.3	0.945(0.007)	0.953(0.004)	0.948(0.002)
400000	0.5	0.942(0.006)	0.949(0.003)	0.944(0.002)
	0.8	0.952(0.009)	0.957(0.004)	0.958(0.002)

Figure 3. Box-plot of estimator Q_n for different target quantiles of Cauchy distribution. In each sample size divided by a vertical dotted line, the three boxes establish results with different privacy budgets by left: $r = 0.25$, middle $r = 0.5$, and right: $r = 0.9$. The horizontal dashed lines represent the true value Q in $\tau = 0.3$, 0.5, 0.8 from the bottom to the top.

5. Conclusion and Future Works

In this paper, we proposed a novel algorithm for estimating population quantiles under the settings of LDP. The core design idea of the algorithm is based on using dichotomous inquiry. The proposed estimator enjoys excellent theoretical properties, including consistency, asymptotic normality, and optimality in some special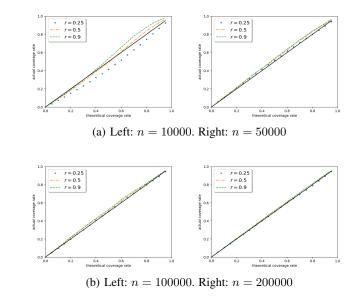 cases. Importantly, by applying the technique of self-normalization to cancel out the nuisance parameters, we can construct confidence intervals of population quantiles for statistical inference. Finally, our algorithm is designed in an online setting, making it suitable for handling large streaming data without the need for data storage. Extensive simulation studies reveal a positive confirmation of the asymptotic theory.

Despite the contributions above, this article still leaves many exciting questions unanswered, which opens many avenues for future research. A general tight lower bound for other quantiles under our setting is still undetermined, and we have yet to consider other variants of LDP (e.g. full-interactive). Other directions include exploring data that is not independently and identically distributed, such as time series or spatial series data. Additionally, the quantile of interest may be influenced by other covariates, leading to the study of LDP quantile regression. This paper focuses on estimating quantiles for a specific sample size n, with the potential for developing consistent bounds, resulting in the transition from quantile confidence intervals to confidence sequences.

Acknowledgements

References

Agarwal, N. and Singh, K. The price of differential privacy for online learning. In Precup, D. and Teh, Y. W. (eds.), *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 32–40. PMLR, 06–11 Aug 2017.

Alabi, D., Ben-Eliezer, O., and Chaturvedi, A. Bounded space differentially private quantiles. *arXiv preprint arXiv:2201.03380*, 2022.

Amin, K., Joseph, M., and Mao, J. Pan-private uniformity testing. In Abernethy, J. and Agarwal, S. (eds.), *Proceedings of Thirty Third Conference on Learning Theory*, volume 125 of *Proceedings of Machine Learning Research*, pp. 183–218. PMLR, 09–12 Jul 2020.

Balle, B., Barthe, G., and Gaboardi, M. Privacy amplification by subsampling: Tight analyses via couplings and divergences. *Advances in Neural Information Processing Systems*, 31, 2018.

Ben-Eliezer, O., Mikulincer, D., and Zadik, I. Archimedes meets privacy: On privately estimating quantiles in high dimensions under minimal assumptions. *arXiv preprint arXiv:2208.07438*, 2022.

Brown, T. C., Champ, P. A., Bishop, R. C., and McCollum, D. W. Which response format reveals the truth about donations to a public good? *Land Economics*, pp. 152–166, 1996.

Cai, T. T., Wang, Y., and Zhang, L. The cost of privacy: Optimal rates of convergence for parameter estimation with differential privacy. *The Annals of Statistics*, 49(5): 2825–2850, 2021.

Cheu, A., Smith, A., Ullman, J., Zeber, D., and Zhilyaev, M. Distributed differential privacy via shuffling. In *Advances in Cryptology–EUROCRYPT 2019: 38th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Darmstadt, Germany, May 19–23, 2019, Proceedings, Part I 38*, pp. 375–403. Springer, 2019.

Clark, C. E. The pert model for the distribution of an activity time. *Operations Research*, 10(3):405–406, 1962.

Coppens, B., Verbauwhede, I., De Bosschere, K., and De Sutter, B. Practical mitigations for timing-based side-channel attacks on modern x86 processors. In *2009 30th IEEE Symposium on Security and Privacy*, pp. 45–60, 2009. doi: 10.1109/SP.2009.19.

Ding, B., Kulkarni, J., and Yekhanin, S. Collecting telemetry data privately. *Advances in Neural Information Processing Systems*, 30, 2017.

Dippon, J. Globally convergent stochastic optimization with optimal asymptotic distribution. *Journal of applied probability*, 35(2):395–406, 1998.

Dong, J., Roth, A., and Su, W. J. Gaussian differential privacy. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 2021a.

Dong, J., Su, W., and Zhang, L. A central limit theorem for differentially private query answering. In Ranzato, M., Beygelzimer, A., Dauphin, Y., Liang, P., and Vaughan, J. W. (eds.), *Advances in Neural Information Processing Systems*, volume 34, pp. 14759–14770. Curran Associates, Inc., 2021b.

Dwork, C. and Lei, J. Differential privacy and robust statistics. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pp. 371–380, 2009.

Dwork, C., Kenthapadi, K., McSherry, F., Mironov, I., and Naor, M. Our data, ourselves: Privacy via distributed noise generation. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 486–503. Springer, 2006.

Erlingsson, Ú., Pihur, V., and Korolova, A. Rappor: Randomized aggregatable privacy-preserving ordinal response. In *Proceedings of the 2014 ACM SIGSAC conference on computer and communications security*, pp. 1054–1067, 2014.

Evfimievski, A., Gehrke, J., and Srikant, R. Limiting privacy breaches in privacy preserving data mining. In *Proceedings of the twenty-second ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pp. 211–222, 2003.

Fang, Y., Xu, J., and Yang, L. Online bootstrap confidence intervals for the stochastic gradient descent estimator. *The Journal of Machine Learning Research*, 19(1):3053–3073, 2018.

Gillenwater, J., Joseph, M., and Kulesza, A. Differentially private quantiles. In Meila, M. and Zhang, T. (eds.), *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 3713–3722. PMLR, 18–24 Jul 2021.

Haney, S., Desfontaines, D., Hartman, L., Shrestha, R., and Hay, M. Precision-based attacks and interval refining: how to break, then fix, differential privacy on finite computers. *arXiv preprint arXiv:2207.13793*, 2022.

Jain, P., Kothari, P., and Thakurta, A. Differentially private online learning. In *Conference on Learning Theory*, pp. 24–1. JMLR Workshop and Conference Proceedings, 2012.

Jin, J., McMurtry, E., Rubinstein, B. I., and Ohrimenko, O. Are we there yet? timing and floating-point attacks on differential privacy systems. *arXiv preprint arXiv:2112.05307*, 2021.

Joseph, A. G. and Bhatnagar, S. A stochastic approximation algorithm for quantile estimation. In *International Conference on Neural Information Processing*, pp. 311–319. Springer, 2015.

Joseph, M., Mao, J., Neel, S., and Roth, A. The role of interactivity in local differential privacy. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 94–105, 2019. doi: 10.1109/FOCS. 2019.00015.

Kairouz, P., Oh, S., and Viswanath, P. The composition theorem for differential privacy. In *International conference on machine learning*, pp. 1376–1385. PMLR, 2015.

Kasiviswanathan, S. P., Lee, H. K., Nissim, K., Raskhodnikova, S., and Smith, A. What can we learn privately? *SIAM Journal on Computing*, 40(3):793–826, 2011.

Langley, P. Crafting papers on machine learning. In Langley, P. (ed.), *Proceedings of the 17th International Conference on Machine Learning (ICML 2000)*, pp. 1207–1216, Stanford, CA, 2000. Morgan Kaufmann.

Lawson, N. Side-channel attacks on cryptographic software. *IEEE Security & Privacy*, 7(6):65–68, 2009. doi: 10. 1109/MSP.2009.165.

Lee, S., Liao, Y., Seo, M. H., and Shin, Y. Fast and robust online inference with stochastic gradient descent via random scaling. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pp. 7381–7389, 2022.

Lei, J. Differentially private m-estimators. *Advances in Neural Information Processing Systems*, 24, 2011.

Liu, Y., Sun, K., Kong, L., and Jiang, B. Identification, amplification and measurement: A bridge to gaussian differential privacy. *Advances in Neural Information Processing Systems*, 2022.

Mironov, I. On significance of the least significant bits for differential privacy. In *Proceedings of the 2012 ACM conference on Computer and communications security*, pp. 650–661, 2012.

Shao, X. A self-normalized approach to confidence interval construction in time series. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(3): 343–366, 2010.

Shao, X. Self-normalization for time series: a review of recent developments. *Journal of the American Statistical Association*, 110(512):1797–1817, 2015.

Smith, A. Privacy-preserving statistical estimation with optimal convergence rates. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pp. 813–822, 2011.

Warner, S. L. Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60(309):63–69, 1965.

Wei, K., Li, J., Ding, M., Ma, C., Yang, H. H., Farokhi, F., Jin, S., Quek, T. Q. S., and Vincent Poor, H. Federated learning with differential privacy: Algorithms and performance analysis. *IEEE Transactions on Information Forensics and Security*, 15:3454–3469, 2020. doi: 10.1109/TIFS.2020.2988575.

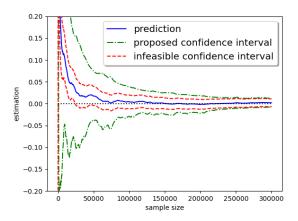# A. Additional figures and tables

Figure 5. An alternative sample trajectory of estimator Q_n using a different initialization $q_0 = 1$.
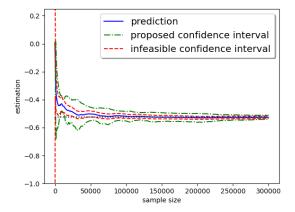

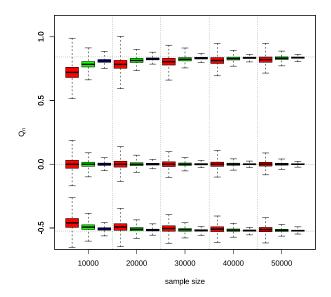*Figure 6.* An alternative sample trajectory of estimator Q_n using a different target quantile $\tau = 0.3$.

11

Figure 7. Box-plot of estimator Q_n for different target quantile of Cauchy distribution. In each sample size divided by a vertical dotted line, the three boxes establish results with different privacy budgets by left: $r = 0.25$, middle $r = 0.5$, and right: $r = 0.9$. The horizontal dashed lines represent the true value Q in $\tau = 0.3, 0.5, 0.8$ from the bottom to the top.

Figure 8. Box-plot of estimator Q_n for different target quantile of Uniform distribution. In each sample size divided by a vertical dotted line, the three boxes establish results with different privacy budgets by left: $r = 0.25$, middle $r = 0.5$, and right: $r = 0.9$. The horizontal dashed lines represent the true value Q in $\tau = 0.3, 0.5, 0.8$ from the bottom to the top.
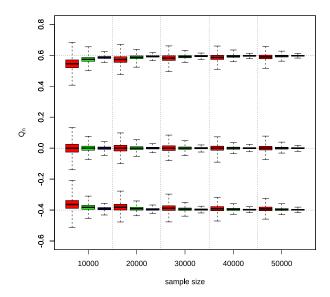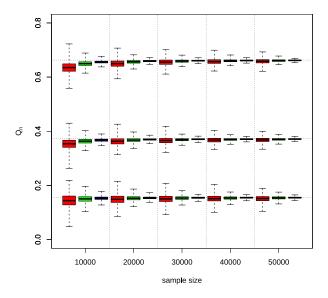
Figure 9. Box-plot of estimator Q_n for different target quantile of PERT distribution. In each sample size divided by a vertical dotted line, the three boxes establish results with different privacy budgets by left: $r = 0.25$, middle $r = 0.5$, and right: $r = 0.9$. The horizontal dashed lines represent the true value Q in $\tau = 0.3, 0.5, 0.8$ from the bottom to the top.

Table 2. Empirical results of coverage rate(mean absolute error) of proposed confidence interval (3) (estimator Q_n) with data collected from Cauchy.

n	τ	$r = 0.25$	$r = 0.5$	$r = 0.9$
	0.3	0.894(0.140)	0.972(0.068)	0.987(0.037)
10000	0.5	0.807(0.045)	0.876(0.024)	0.906(0.014)
	0.8	0.853(0.399)	0.989(0.207)	1.000(0.112)
	0.3	0.928(0.076)	0.966(0.037)	0.982(0.020)
20000	0.5	0.872(0.034)	0.908(0.018)	0.927(0.010)
	0.8	0.950(0.219)	0.991(0.105)	0.998(0.055)
	0.3	0.944(0.044)	0.964(0.022)	0.974(0.012)
40000	0.5	0.900(0.025)	0.926(0.012)	0.939(0.007)
	0.8	0.965(0.114)	0.984(0.053)	0.993(0.028)
	0.3	0.944(0.025)	0.956(0.012)	0.963(0.007)
100000	0.5	0.927(0.016)	0.935(0.008)	0.945(0.004)
	0.8	0.956(0.054)	0.970(0.026)	0.980(0.013)
	0.3	0.948(0.017)	0.954(0.008)	0.958(0.004)
200000	0.5	0.936(0.011)	0.944(0.006)	0.945(0.003)
	0.8	0.952(0.034)	0.966(0.017)	0.971(0.008)
	0.3	0.942(0.012)	0.954(0.006)	0.952(0.003)
400000	0.5	0.944(0.008)	0.949(0.004)	0.946(0.002)
	0.8	0.948(0.023)	0.960(0.011)	0.961(0.005)

13

Table 3. Empirical results of coverage rate(mean absolute error) of proposed confidence interval (3) (estimator Q_n) with data collected from PERT.

n	τ	$r = 0.25$	$r = 0.5$	$r = 0.9$
	0.3	0.900(0.021)	0.927(0.011)	0.938(0.006)
10000	0.5	0.951(0.022)	0.970(0.011)	0.971(0.006)
	0.8	0.990(0.029)	0.997(0.014)	0.998(0.008)
	0.3	0.920(0.015)	0.932(0.007)	0.941(0.004)
20000	0.5	0.950(0.014)	0.957(0.007)	0.962(0.004)
	0.8	0.983(0.016)	0.990(0.008)	0.992(0.004)
	0.3	0.927(0.011)	0.937(0.005)	0.936(0.003)
40000	0.5	0.947(0.009)	0.951(0.004)	0.955(0.002)
	0.8	0.974(0.009)	0.978(0.005)	0.982(0.002)
	0.3	0.934(0.007)	0.936(0.003)	0.942(0.002)
100000	0.5	0.948(0.005)	0.948(0.003)	0.956(0.001)
	0.8	0.967(0.005)	0.969(0.003)	0.972(0.001)
	0.3	0.936(0.005)	0.935(0.002)	0.939(0.001)
200000	0.5	0.943(0.004)	0.952(0.002)	0.949(0.001)
	0.8	0.960(0.004)	0.963(0.002)	0.964(0.001)
	0.3	0.936(0.003)	0.935(0.002)	0.936(0.001)
400000	0.5	0.946(0.003)	0.946(0.001)	0.946(0.001)
	0.8	0.955(0.003)	0.956(0.001)	0.956(0.001)

Table 4. Empirical results of coverage rate(mean absolute error) of proposed confidence interval (3) (estimator Q_n) with data collected from Uniform.

n	τ	$r = 0.25$	$r = 0.5$	$r = 0.9$
	0.3	0.922(0.043)	0.956(0.021)	0.972(0.011)
10000	0.5	0.853(0.030)	0.898(0.016)	0.928(0.009)
	0.8	0.965(0.057)	0.984(0.028)	0.994(0.015)
	0.3	0.930(0.027)	0.950(0.013)	0.963(0.007)
20000	0.5	0.896(0.022)	0.928(0.011)	0.934(0.006)
	0.8	0.960(0.032)	0.977(0.016)	0.984(0.008)
	0.3	0.939(0.017)	0.953(0.009)	0.959(0.004)
40000	0.5	0.921(0.016)	0.934(0.008)	0.943(0.004)
	0.8	0.959(0.019)	0.969(0.009)	0.974(0.005)
	0.3	0.942(0.010)	0.953(0.005)	0.955(0.003)
100000	0.5	0.939(0.010)	0.942(0.005)	0.943(0.003)
	0.8	0.954(0.011)	0.959(0.005)	0.960(0.003)
	0.3	0.944(0.007)	0.950(0.003)	0.950(0.002)
200000	0.5	0.938(0.007)	0.947(0.004)	0.946(0.002)
	0.8	0.950(0.007)	0.956(0.004)	0.957(0.002)
	0.3	0.945(0.005)	0.947(0.002)	0.950(0.001)
400000	0.5	0.944(0.005)	0.951(0.002)	0.950(0.001)
	0.8	0.946(0.005)	0.955(0.002)	0.948(0.001)

Table 5. Conversion table between r and ϵ

r	ϵ	r	ϵ
0	0	0.5	1.10
0.05	0.10	0.55	1.24
0.1	0.20	0.6	1.39
0.15	0.30	0.65	1.55
0.2	0.40	0.7	1.73
0.25	0.51	0.75	1.95
0.3	0.62	0.8	2.20
0.35	0.73	0.85	2.51
0.4	0.85	0.9	2.94
0.45	0.97	0.95	3.66

B. Alternative self-normalizes

The following self-normalizer can also be used to construct the asymptotically pivotal quantity,

$$N'_n = \sup_{t \in [0,1]} \left| S_{[nt]} - [nt]Q_n \right|,$$

$$N''_n = \int_0^1 \left| S_{[nt]} - [nt]Q_n \right| dt,$$

and based on the continuous mapping theorem again, one has that,

$$\frac{n^{-1/2}(S_n - nQ)}{n^{-1/2}N'_n} \xrightarrow{d} \frac{W(1)}{\sup_{t \in [0,1]} |W(t) - tW(1)|},$$

$$\frac{n^{-1/2}(S_n - nQ)}{n^{-1/2}N''_n} \xrightarrow{d} \frac{W(1)}{\int_0^1 |W(t) - tW(1)| \, dt}.$$

C. Proof

C.1. Proof of Theorem 3.1

Exhaustive computation yields that for any $(a, b) \in \{0, 1\}^2$

$$\frac{\mathbb{P}(LRC(q, r, x) = a | \mathbf{1}_{x>q} = b)}{\mathbb{P}(LRC(q, r, x) = a | \mathbf{1}_{x>q} = 1 - b)} \in \{\frac{1+r}{1-r}, \frac{1-r}{1+r}\} \tag{6}$$

C.2. Proof of Theorem 3.2, 3.3 and 3.4

One can verify that the recursive equation (1) is asymptotically equivalent to

$$q_{n+1} = q_n + \frac{1}{d_n}\left(1 - \frac{2}{1 - r + 2r(1-\tau)}\mathbf{1}_{x_n^* > q_n}\right),$$

where $\mathbb{P}(x_n^* = x_n) = r$, $\mathbb{P}(x_n^* = -\infty) = \mathbb{P}(x_n^* = \infty) = (1-r)/2$. Let

$$H(z, X) = 1 - \frac{2}{1 - r + 2r(1-\tau)}\mathbf{1}_{X>z}$$

$$h(z, X) = \mathbb{E}H(z, X) = 1 - \frac{2(1 - F(z))}{1 - r + 2r(1-\tau)}.$$

Hence $F(Q) = \tau$ is equivalent to $h(Q, X^*) = 0$. Then, one will find that the estimation of Q with sample $x_1, \ldots, x_n$ under LDP is equivalent to the estimation of Q^* with sample $x_1, \ldots, x_n$ without LDP constraints. The standard framework of the SGD method, such as Theorem 2 and 3 in (Dippon, 1998), can be applied. Moreover, the statements in Theorems 3.2, 3.3, and 3.4 hold true.

C.3.

We prove this by contradiction. Assuming that for any $n_0 > 1$ there is a $n > n_0$ such that :

$$\epsilon_n/\sigma_n > k > 1.$$

Let

$$w = \Phi\left(-\frac{1}{2}\right) - \Phi\left(\frac{1}{2} - k\right) > 0.$$

We choose a sufficiently large n_0 such that for any $n > n_0$

$$\mathbb{P}(\hat{\theta}_n/\sigma_n < 1/2 | H_0) \geq \Phi(1/2) - w/3$$

and

$$\mathbb{P}(\hat{\theta}_n/\sigma_n < 1/2|H_1) \le \Phi(1/2 - \epsilon_n/\sigma_n) + w/3 \le \Phi(1/2 - k) + w/3.$$

Then,

$$
\begin{aligned}
\mathbb{P}(\hat{\theta}_n/\sigma_n < 1/2|H_0) - \mathbb{P}(\hat{\theta}_n/\sigma_n < 1/2|H_1) &\ge \Phi(1/2) - \Phi(1/2 - k) - 2w/3 \\
&= 2\Phi\left(\frac{1}{2}\right) - 1 + \Phi\left(-\frac{1}{2}\right) - \Phi\left(\frac{1}{2} - k\right) - 2w/3 \\
&= 2\Phi\left(\frac{1}{2}\right) - 1 + w/3 \\
&> 2\Phi\left(\frac{1}{2}\right) - 1 + w/6
\end{aligned}
$$

Then $\hat{\theta}_n$ is not $(0, 2\Phi\left(\frac{1}{2}\right) - 1 + w/6)$-DP and therefore is not asymptotically 1-GDP leading to a contradiction.