A General Framework for Accurate and Private Mean Estimation

Zhouhao Yang, Xingyu Xu, *Graduate Student Member, IEEE*, and Yuantao Gu, *Senior Member, IEEE*

***Abstract*—In this letter, we present a differentially private algorithm which accurately estimates the mean of an underlying population with given cumulative distribution function. Our algorithm outperforms the former algorithms in two aspects. First, our algorithm is capably of handling more general types of probability distributions, possibly with a very heavy tail. Second, for light-tailed distributions, our algorithm achieves a better level of accuracy with fewer samples.**

***Index Terms*—Differential privacy, mean estimation, heavy-tailed distribution.**

I. Introduction

ESTIMATING the mean of a distribution given some independently and identically distributed (*i.i.d.*) samples is definitely a classical and fundamental problem in statistics, signal processing, and many other fields in science and engineering. There have been a sea of research in this vein striving for better *accuracy* of mean estimation. However, in modern days some more concerns other than accuracy are also found to be crucial. In particular, under many circumstances, the samples contain sensitive personal information susceptible to privacy attacks, making the preservation of privacy crucial and indispensable. Therefore, there is an ascending demand for learning algorithms that can ensure the *privacy* of individuals.

There are several different ways to formalize the notion of privacy in a scientific field like signal processing. Among these different ways, the concept of *differential privacy* (DP), proposed in [1], is arguably the most popular way by now [2], [3], [4]. Well-designed differentially private algorithms is ought to assemble a combination of utility and privacy. On one hand, they provide a useful estimation of statistics. On the other hand, they make it computationally impossible to speculate about sensitive information of individuals from data [5].

Informally, we call a statistical analysis differentially private if the likelihood of the (randomized) outcome does not differ much when a single datapoint in the dataset is altered [6]. More precisely, we consider any algorithm that takes n samples from some space $\mathcal{X}$ and outputs the result of it analyzing the samples, e.g. an estimate of the mean of the samples, taking value in some output space $\mathcal{Y}$. The formal definition of differential privacy is as follows.

Definition 1 (Differential Privacy (DP) [1], [7]): A randomized algorithm $M : \mathcal{X}^n \to \mathcal{Y}$ satisfies (ε, δ)-differential privacy ((ε, δ)-DP) if for every pair of neighboring datasets $X, X' \in \mathcal{X}^n$ (i.e., datasets that differ in exactly one entry), $\forall S \subset \mathcal{Y}$,

$$\mathbb{P}[M(X) \in S] \le e^{\varepsilon}\mathbb{P}[M(X') \in S] + \delta. \tag{1}$$

When $\delta = 0$, we say that M satisfies ε-differential privacy or pure differential privacy.

As mentioned above, an important scenario where differential privacy is frequently involved is statistical inference. Statistical inference is a classical statistical problem, which can be described as follows:

Given some samples $X_1, X_2, \ldots, X_n$ from an unknown probabilistic model, how can we estimate certain properties, including mean, variance, moments and other statistics, of the underlying population [8]?

Furthermore, a rule depending on observed data to compute an estimate of a given quantity is called an estimator. Specifically, for some statistical property θ, we call $\widehat{\theta}(X_1, X_2, \ldots, X_n)$ an estimator of θ, if $\widehat{\theta}$ estimates the quantity of θ after taking in the samples. For the sake of convenience, we specify that an estimator $\widehat{\theta}$ is claimed to satisfy α-accuracy if $|\theta - \widehat{\theta}| \le \alpha$.

It is noteworthy that there are various reasons a classical estimator fails to attain accuracy and privacy [9]. On one hand, classical estimator can be non-optimal in accuracy in case that the probability distribution is heavy-tailed. For example, in the seemingly simple task of mean estimation, the most commonly used estimator, the empirical mean $(X_1 + \cdots + X_n)/n$, does not achieve optimal accuracy when the distribution of X_i are heavy-tailed [10]. On the other hand, such estimators also grant no guarantee of privacy [11]. It is therefore a nontrivial task to design private statistical inference algorithms. Much effort has been put in this direction in recent years, including private mean estimation [11], [12], [13], [14], [15], private covariance estimation [16], [17], [18], private hypothesis selection [19], etc.

In this work, we focus on univariate private mean estimation. That is, estimating the mean of a distribution $\mathcal{D}$ on $\mathbb{R}$ given a few *i.i.d.* samples $X_1, \ldots, X_n$ from $\mathcal{D}$. This topic has attracted a lot of interest recently and private mean estimation algorithms have been designed for various distributions. Unsurprisingly, the overall trend is to design algorithms that work for more and more general classes of distributions, while guaranteeing a reasonable sample complexity, i.e. the number of samples required

Manuscript received 1 July 2022; revised 9 October 2022; accepted 26 October 2022. Date of publication 3 November 2022; date of current version 18 November 2022. This work was supported in part by the National Natural Science Foundation of China under Grants NSFC U2230201 and 61971266, in part by Guoqiang Institute, Tsinghua University, and in part by the Clinical Medicine Development Fund of Tsinghua University. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. KokSheik Wong. *(Corresponding author: Yuantao Gu.)*

Zhouhao Yang is with the Zhiyuan College, Shanghai Jiaotong University, Shanghai 200240, China (e-mail: Yang,yangzhouhao@sjtu.edu.cn).

Xingyu Xu and Yuantao Gu are with the Beijing National Research Center for Information Science and Technology (BNRist) and Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: xy-xu19@mails.tsinghua.edu.cn; gyt@tsinghua.edu.cn).

This letter has supplementary downloadable material available at https://doi.org/10.1109/LSP.2022.3219356, provided by the authors.

Digital Object Identifier 10.1109/LSP.2022.3219356

TABLE I
FORMER RESULTS (SORTED IN ASCENDING ORDER OF GENERALITY OF ASSUMPTION)

Assumption	Sample Complexity	Reference
X Gaussian, $\mathrm{Var}(X) \le 1$	$\widetilde{O}\left(\frac{1}{\alpha^2} + \frac{1}{\alpha\varepsilon}\right)$	[12]
$\|X\|_{\psi_2} \le K$, $\mathrm{Var}\,(X) \le 1$	$\widetilde{O}\left(\frac{1}{\alpha^2} + \frac{K}{\alpha\varepsilon}\right)$	[14], [19]
$\mathbb{E}X^2 \le M^2$	$\widetilde{O}\left(\frac{M}{\alpha^2\varepsilon^2}\right)$	[20]
$\mathbb{E}\lvert X\rvert^k \le M_k^k$, $\mathrm{Var}\,(X) \le 1$	$\widetilde{O}\left(\frac{1}{\alpha^2} + \frac{M_k}{\varepsilon\alpha^{\frac{k}{k-1}}}\right)$	[11]

In the table, for different classes of distributions, we present the sample complexity of (ε, δ)-DP algorithms that gives an estimation of the mean within α-accuracy. In addition, we omit the term $O\left(\log(R)/\varepsilon\right)$ in the sample complexities, where R is the bound of constructed histograms. $\|X\|_{\psi_2}$ denotes the sub-Gaussian norm of the random variable X.

to achieve α-accuracy. In Table I, we summarize the former results, namely the sample complexity for different classes of distributions.

There is however a significant gap between the sample complexity results for sub-Gaussian distributions and the results for distributions with a bounded k-th moment. Although the algorithm designed for distributions with a bounded k-th moment can also be applied to Gaussian distributions, which means it can be considered as a more general algorithm, it actually falls short in sample efficiency in that case. In fact, for standard Gaussian distribution, we have $M_k \asymp \sqrt{k}$, thus the corresponding sample complexity will be

$$n = \widetilde{O}\left(\frac{1}{\alpha^2} + \frac{\sqrt{k}}{\varepsilon\alpha^{k/(k-1)}}\right). \tag{2}$$

The second term is much larger than the corresponding term $1/\varepsilon\alpha$ compared to the Gaussian case as $\alpha \to 0$, for most choices of k (for small k the exponent of α is far from optimal, while for large k the factor $\sqrt{k}$ grows too large).

Now, based on the above reasoning, the following question arises naturally: is it possible to establish a more unified private mean estimation framework which enjoys both generality and sample efficiency (i.e. accuracy) for different distributions?

II. INFORMAL STATEMENT OF RESULTS

In our work, we answer this question affirmatively by proposing an ε-differentially private histogram-based algorithm that estimates the mean of a population with α-accuracy. We focus on the univariate case and leave the multivariate case for future discussion.

A more detailed formulation is as follows. Denote by $\mathcal{D}$ the underlying centered distribution. Now we have some *i.i.d.* samples from the translated distribution $\mu + \mathcal{D}$ and would like to estimate μ. Instead of restricting $\mathcal{D}$ to be sub-Gaussian or to be of bounded k-th moments *a priori*, we take a different perspective and assume we have access to a function $T(\alpha)$ such that

$$\mathbb{E}_{X\sim\mathcal{D}}(|X| - T(\alpha))\mathbf{1}_{|X|>T(\alpha)} \le \alpha \tag{3}$$

for every $\alpha \in (0, 1)$. Note that there always exists some function T satisfying (3) (as long as X is integrable, which is of course always assumed since we are concerned with estimating the mean of X). What we assume here is a practical way to compute T which should be interpreted as a form of prior knowledge on $\mathcal{D}$. For example, if we know that $\mathcal{D}$ is sub-Gaussian with sub-Gaussian norm $\le K$, we may use $T(\alpha) = CK\sqrt{\log(C/\alpha)}$ for some sufficiently large constant $C > 0$. If we know that the k-th moment of $\mathcal{D}$ is upper bounded by 1 and know nothing otherwise, we may use $T(\alpha) = C\alpha^{-1/(k-1)}$. If we have a rather comprehensive knowledge on $\mathcal{D}$ that we even know its c.d.f., we may simply compute $G(x) = \mathbb{E}_{X\sim\mathcal{D}}(|X| - x)\mathbf{1}_{|X|>x}$ and set $T = G^{-1}$ to be an inverse of G (which exists since G is monotone).

We also need to assume the existence of a low order moment of the underlying population $\mathcal{D}$, since otherwise even non-private estimation is not possible. For simplicity we assume $\mathbb{E}|X|^2 \le M$, though $(1+\delta)$-th moment $(\delta > 0)$ would suffice.

With the above assumptions, we propose an ε-differentially private algorithm that takes

$$n \ge O\left(\log\left(\frac{1}{\beta}\right)\left(\frac{1}{\alpha^2} + \frac{T(\alpha)}{\varepsilon\alpha} + \frac{\log(R)}{\varepsilon}\right)\right) \tag{4}$$

samples and guarantees that with probability at least $1 - \beta$, the output statistical mean $\hat{\mu}$ is close to the original mean within α-accuracy. Here R denotes a pre-specfied upper bound of μ, which can usually be set as a sufficiently large constant.

III. PRIVATE MEAN ESTIMATION OF GENERAL DISTRIBUTIONS

In this section, we present our algorithm as well as its guarantee on privacy and accuracy. The proofs are in supplementary material.

A. Technical Lemma

To begin with, we present a vital technical lemma necessary for our ultimate algorithm.

In Lemma 1, we claim that there exists an ε-DP algorithm that privately returns an interval containing a large portion of sample points as well as the mean of the underlying population with high probability. We name the algorithm as DP Range Estimation Algorithm. This lemma is inspired from Algorithm 1 in [12] and Algorithm 1 in [11].

Lemma 1 (DP Range Estimation): Let $\mathcal{D}$ be a mean-zero distribution over $\mathbb{R}$. Assume that $\mathbb{E}_{X\sim\mathcal{D}}[X] = \mu \in [-R, R]$ with some sufficiently large number R. Set

$$G(x) = \mathbb{E}_{X\sim\mathcal{D}}[(|X| - x)\,\mathbf{1}_{|X|>x}],$$

and $T = G^{-1}$. Then for every $0 < \alpha < 0.1$, ε, R, there exists an ε-DP Range Estimation Algorithm that requires

$$n \ge O\left(\frac{1}{\alpha} + \frac{\log(R\alpha)}{\varepsilon}\right) \tag{5}$$

samples, and outputs $I = [a, b]$, such that with probability at least 0.9, we have:
1) $b - a \in \Theta(T(\alpha))$,
2) At most αn samples lie outside I,
3) $\mu \in I$ and $\mu - a, b - \mu > 2T(\alpha)$.

Algorithm 1: DP Mean Estimation $\mathrm{PME}_{\alpha,\varepsilon,R}(X)$.

Input: Samples $X_1, X_2, \ldots, X_n, X_1,' X_2,' \ldots, X_n' \in \mathbb{R}$. Parameters ε, α, R
Output: $\widehat{\mu} \in \mathbb{R}$: An ε-DP estimation of the mean of the underlying population

```
Set iteration parameter m = 10 log(2/β)
for i = 1, 2, ..., m do
  Let Y_i = {X_i : (i−1)n/m ≤ i ≤ in/m + 1}
  Let Y'_i = {X'_i : (i−1)n/m ≤ i ≤ in/m + 1}
  Run DP Range Estimation Algorithm for Y'_i and
    outputs an interval I_i, r_i = |I_i|
  for x ∈ Y_i do
    y = arg min_{y∈I_i} |y − x|
    x = y
  end for
end for
μ̂_i = (1/n) Σ_{y∈Y_i} y + Lap(m r_i / (ε n))
μ̂ = Median(μ̂_1, ..., μ̂_m)
return μ̂
```

B. Our Algorithm

In this subsection, we present our main contribution: an ε-DP Private Mean Estimation Algorithm for general distributions. The algorithm is divided into two parts. First, we limit the data into a designated range for the purpose of (1) preserving privacy (2) achieving optimal accuracy, which requires constraining the amount of noise added. Second, we estimate the mean of a given dataset in a differentially private way.

In the following, we concentrate on private mean estimation. Essentially, the algorithm is based on the trade-off between privacy and accuracy. On one hand, the range of our previously constructed interval should be large enough so that if we truncate the distribution within it, then the mean of the truncated distribution will be close to the mean of the original distribution. On the other hand, the range ought to be small enough so that the noise added guarantees the privacy of the algorithm. Based on those considerations, we assert that our algorithm is not only differentially private, but is accurate enough without bringing a large overhead to the sample complexity.

Theorem 1: Let $\mathcal{D}$ be a distribution over $\mathbb{R}$ with its $(1+k)$-th central moment bounded by M for some $k \in (0, 1]$. Assume that $\mathbb{E}_{X\sim\mathcal{D}}[X] = \mu \in [-R, R]$ with some sufficiently large number R. Let

$$G(x) = \mathbb{E}_{X\sim\mathcal{D}}[(|X| - x)\,\mathbf{1}_{|X|>x}],$$

and $T = G^{-1}$. Then for all $\varepsilon, \alpha, R > 0$, there exists an ε-DP algorithm that takes

$$n \geq O\left(\log\left(\frac{1}{\beta}\right)\left(\frac{1}{\alpha^2} + \frac{T(\alpha)}{\varepsilon\alpha} + \frac{\log(R)}{\varepsilon}\right)\right). \quad (6)$$

samples from $\mathcal{D}$ and outputs $\widehat{\mu} \in \mathbb{R}$, such that with probability at least $1-\beta$, we have

$$|\mu - \widehat{\mu}| \leq \alpha.$$

IV. Experiments

In this section, we justify our conclusion both theoretically and experimentally. For the theoretical part, we will compare our algorithm with the former ones in the cases of Gaussian distributions and of distributions with bounded k-th moment. For the experimental part, we compare our algorithm with the one proposed by Kamath et al. in [11]. For distributions with a very heavy tail, Kamath et al.'s algorithm may fail, whereas our algorithm can still be applied. We design an experiment for Gaussian distributions and Levy-stable distributions.

We shall begin with theoretical verification, contrasting our results, which is based on $G(x) = \mathbb{E}[(|X| - x)\mathbf{1}_{|X|>x}]$ and $T = G^{-1}$, with the results proved in the literature before.

A. Theoretical Verification

1) Gaussian Distributions: Since α is close to 0, by pure computation, we have

$$T(\alpha) = O\left(\sqrt{\log(1/\alpha)}\right) \quad (7)$$

under the assumption that $\mathcal{D}$ is a standard Gaussian distribution. Then the sample complexity becomes

$$n \geq O\left(\log\left(\frac{1}{\beta}\right)\left(\frac{1}{\alpha^2} + \frac{\sqrt{\log(\frac{1}{\alpha})}}{\varepsilon\alpha} + \frac{\log(R)}{\varepsilon}\right)\right), \quad (8)$$

which matches the result of [12].

2) Distributions With Bounded K-th Moment: Let X be a distribution over $\mathbb{R}$ with mean $\mu = 0$ and k-th moment bounded by M. By Chebyshev's inequality,

$$\mathbb{P}(|X| > tM^{\frac{1}{k}}) \leq \frac{M}{t^k} \quad (\forall t > 0). \quad (9)$$

Subsequently, since α is small,

$$T(\alpha) = O\left(\frac{M^{\frac{1}{k}}}{\alpha^{\frac{1}{k-1}}}\right). \quad (10)$$

And the sample complexity turns into

$$n \geq O\left(\log\left(\frac{1}{\beta}\right)\left(\frac{1}{\alpha^2} + \frac{1}{\varepsilon\alpha^{\frac{k}{k-1}}} + \frac{\log(R)}{\varepsilon}\right)\right), \quad (11)$$

which matches the result in [11] for $k \geq 2$. However, note that we never assume $k \geq 2$ in our algorithm, thus our algorithm works for more general distributions, in particular, the important class of Levy distributions (c.f. the section on Experiments).

B. Numerical Experiments

We compare our algorithm with the one proposed in [11]. We demonstrate that our algorithm enjoys a better sample efficiency: we entail fewer samples to achieve a similarly accurate mean estimation. We focus on two classes of mean-zero underlying populations. For each distribution, we will repeat the experiments for 30 times and utilize the mean square root of the outputs as the final estimated mean of our algorithm. Since we set the mean of underlying populations to be 0, the output of the algorithms is regarded as the final accuracy.

Furthermore, note that there is a constant factor in the length r of buckets in histogram algorithm as well as the sample complexity n, we should pre-run the algorithms and find appropriate constants in advance. Then, when we run the algorithms as in the above procedure, the constants are fixed for changing α, ε and R.

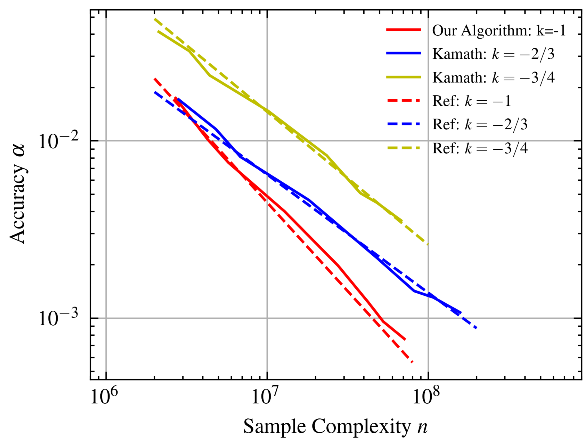

Fig. 1. Log-log plot of $\alpha \sim n$ dependence for Standard Gaussian Distribution.

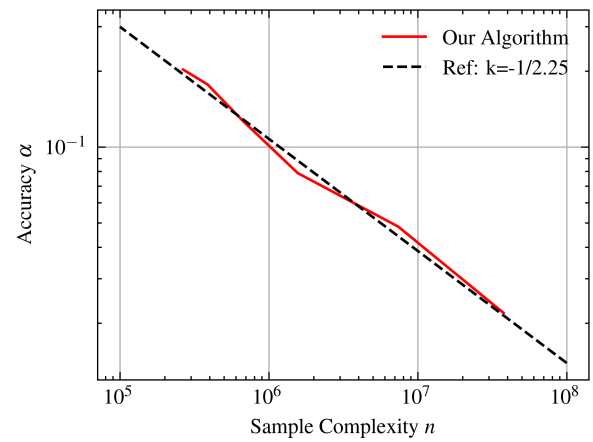

Fig. 3. Log-log plot of $\alpha \sim n$ dependence for Levy-stable distribution with $\alpha^* = 1.8$.

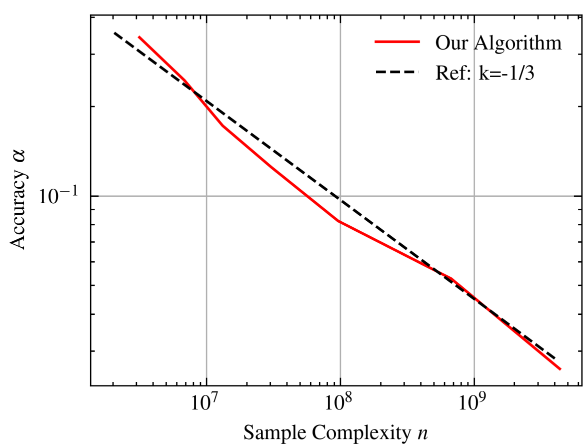

Fig. 2. Log-log plot of $\alpha \sim n$ dependence for Levy-stable distribution with $\alpha^* = 1.5$.

Gaussian Distributions (light tail, better accuracy): Assume that the underlying population $\mathcal{D}$ is a standard normal distribution $\mathcal{N}(0,1)$. Using our algorithm, the theoretical relation between sample complexity and input parameters ought to be

$$n = O\left(\log\left(\frac{1}{\beta}\right)\left(\frac{1}{\alpha^2} + \frac{\sqrt{\log(\frac{1}{\alpha})}}{\varepsilon\alpha} + \frac{\log(R)}{\varepsilon}\right)\right). \quad (12)$$

For Kamath et al.'s algorithm, we use the k-th moment of $\mathcal{D}$ and set e.g. $k = 3, 4$ in experiments. Thus we have

$$n = O\left(\log\left(\frac{1}{\beta}\right)\left(\frac{1}{\alpha^2} + \frac{1}{\varepsilon\alpha^{\frac{k}{k-1}}} + \frac{\log(R)}{\varepsilon}\right)\right). \quad (13)$$

The superiority of our algorithm is that we require less number of samples n to achieve a pre-specified accuracy α. To demonstrate this phenomenon, we consider the case where $\epsilon \ll 1$, corresponding to a strict-privacy scenario, such that the second term in the above equations are dominant.

See Fig. 1 for an illustration of the $\alpha \sim n$ dependence. We can conclude that for our algorithm, $\alpha \propto n^{-1}$. For Kamath et al.'s algorithm, $k = 3, 4$ corresponds resp. to $\alpha \propto n^{-2/3}$ and $\alpha \propto n^{-3/4}$, matching our previous analysis.

Levy-stable distributions (heavier tail, better applicability): Levy-stable distributions are an important class of heavy-tailed distributions that are often used in practice [21], [22], [23], [24], [25]. They are controlled by four parameters $(\alpha^*, \beta^*, \gamma^*, \mu^*)$, which respectively stand for stability, skewness, scale and location. Assume $f(x; \alpha^*, \beta^*, \gamma^*, \mu^*)$ to be the probability density function of a Levy-stable distribution $\mathcal{L}(\alpha^*, \beta^*, \gamma^*, \mu^*)$, then the asymptotic behavior of an Levy-stable distribution suggests that

$$\lim_{x\to\pm\infty} f(x; \alpha^*, 0, 1, 0) = \frac{\alpha^*\Gamma(\alpha^*)\sin(\frac{\pi\alpha^*}{2})}{\pi|x|^{1+\alpha^*}}. \quad (14)$$

Therefore, $G(x) \in \Theta(x^{1-\alpha^*})$ $(x \to +\infty)$ and $T(\alpha) \in \Theta(\alpha^{\frac{1}{1-\alpha^*}})$ $(\alpha \to 0^+)$. Using our algorithm, the sample complexity shall be

$$n = O\left(\log\left(\frac{1}{\beta}\right)\left(\frac{1}{\alpha^{\frac{1+\alpha^*}{\alpha^*}}} + \frac{1}{\varepsilon\alpha^{\frac{\alpha^*}{\alpha^*-1}}} + \frac{\log(R)}{\varepsilon}\right)\right). \quad (15)$$

Meanwhile, since the order of finite moment of a Levy-stable distribution $\mathcal{L}(\alpha^*, \beta^*, \gamma^*, \mu^*)$ can not exceed α^*, it's inappropriate to use Kamath et al.'s algorithm in this case.

In the experiments, we consider $\mathcal{L}(\alpha^*, \beta^*, \gamma^*, \mu^*) = (\alpha^*, 0, 1, 0)$ with $\alpha^* = 1.5, 1.8$ respectively. See Fig. 2 and Fig. 3 for the $\alpha \sim n$ dependence. We can conclude that using our algorithm, $\alpha \propto n^{-(\alpha^*-1)/\alpha^*}$, which is in accordance with our former analysis.

V. Conclusion

In this letter, a differentially private mean estimation algorithm for general distributions is proposed. Our algorithm has an edge in both generality and sample efficiency. We provide an upper bound for the sample complexity of our algorithm. In addition, experiments regarding Gaussian distributions and Levy-stable distributions are also included to justify the practical utility of our result.

REFERENCES

[1] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *Proc. Theory Cryptogr. Conf.*, Springer, 2006, pp. 265–284.
[2] A. D. Sarwate and K. Chaudhuri, "Signal processing and machine learning with differential privacy: Algorithms and challenges for continuous data," *IEEE Signal Process. Mag.*, vol. 30, no. 5, pp. 86–94, Sep. 2013.
[3] M. Kim, O. Günlü, and R. F. Schaefer, "Federated learning with local differential privacy: Trade-offs between privacy, utility, and communication," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2021, pp. 2650–2654.
[4] X. Shen and Y. Liu, "Distributed differential utility/cost analysis for privacy protection," *IEEE Signal Process. Lett.*, vol. 26, no. 10, pp. 1436–1440, Oct. 2019.
[5] C. Tzamos, E.-V. Vlatakis-Gkaragkounis, and I. Zadik, "Optimal private median estimation under minimal distributional assumptions," in *Proc. Adv. Neural Inf. Process. Syst.*, 2020, vol. 33, pp. 3301–3311.
[6] X. Liu, W. Kong, S. Kakade, and S. Oh, "Robust and differentially private mean estimation," *Adv. Neural Inf. Process. Syst.*, vol. 34, pp. 3887–3901, 2021.
[7] C. Dwork et al., "The algorithmic foundations of differential privacy," *Found. Trends Theor. Comput. Sci.*, vol. 9, no. 3-4, pp. 211–407, 2014.
[8] H. Zhang, "Statistical inference in the differential privacy model," Ph.D. dissertation, Dept. Elect. Comput. Eng., Cornell Univ., Ithaca, NY, USA, 2021.
[9] G. Kamath and J. Ullman, "A primer on private statistics," 2020, *arXiv:2005.00010*.
[10] G. Lugosi and S. Mendelson, "Mean estimation and regression under heavy-tailed distributions: A survey," *Found. Comput. Math.*, vol. 19, no. 5, pp. 1145–1190, 2019.
[11] G. Kamath, V. Singhal, and J. Ullman, "Private mean estimation of heavy-tailed distributions," in *Proc. Conf. Learn. Theory* (PMLR), 2020, pp. 2204–2235.
[12] V. Karwa and S. Vadhan, "Finite sample differentially private confidence intervals," in *Proc. 9th Innov. Theor. Comput. Sci. Conf.*, (ser. Leibniz International Proceedings in Informatics (LIPIcs)), A. R. Karlin, Ed., vol. 94, Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2018, pp. 44:1–44:9. [Online]. Available: http://drops.dagstuhl.de/opus/volltexte/2018/8344
[13] V. Balcer and S. Vadhan, "Differential privacy on finite computers," *J. Privacy Confidentiality*, vol. 9, no. 2, Sep. 2019. [Online]. Available: https://journalprivacyconfidentiality.org/index.php/jpc/article/view/679
[14] G. Kamath, J. Li, V. Singhal, and J. Ullman, "Privately learning high-dimensional distributions," in *Proc. Conf. Learn. Theory*, 2019, pp. 1853–1902.
[15] G. Kamath, O. Sheffet, V. Singhal, and J. Ullman, "Differentially private algorithms for learning mixtures of separated Gaussians," in *Proc. 33rd Int. Conf. Adv. Neural Inf. Process. Syst.*, 2019, pp. 168–180.
[16] S. Biswas, Y. Dong, G. Kamath, and J. Ullman, "Coinpress: Practical private mean and covariance estimation," in *Proc. Adv. Neural Inf. Process. Syst.*, 2020, vol. 33, pp. 14475–14485.
[17] K. Amin, T. Dick, A. Kulesza, A. Munoz, and S. Vassilvitskii, "Differentially private covariance estimation," in *Proc. 33rd Int. Conf. Adv. Neural Inf. Process. Syst.*, 2019, pp. 14213–14222.
[18] M. Avella-Medina and V.-E. Brunel, "Differentially private sub-Gaussian location estimators," 2019, *arXiv:1906.11923*.
[19] M. Bun, G. Kamath, T. Steinke, and S. Z. Wu, "Private hypothesis selection," in *Proc. 33rd Int. Conf. Adv. Neural Inf. Process. Syst.*, 2019, pp. 156–167.
[20] M. Bun and T. Steinke, "Average-case averages: Private algorithms for smooth sensitivity and mean estimation," in *Proc. 33rd Int. Conf. Adv. Neural Inf. Process. Syst.*, 2019, pp. 181–191.
[21] J. H. McCulloch, "Simple consistent estimators of stable distribution parameters," *Commun. Statist.- Simul. Computation*, vol. 15, no. 4, pp. 1109–1136, 1986.
[22] C. L. Nikias and M. Shao, *Signal Processing With Alpha-Stable Distributions and Applications*, Hoboken, NJ, USA: Wiley, 1995.
[23] E. E. Kuruoglu, W. J. Fitzgerald, and P. J. Rayner, "Near optimal detection of signals in impulsive noise modeled with a symmetric/spl alpha/-stable distribution," *IEEE Commun. Lett.*, vol. 2, no. 10, pp. 282–284, Oct. 1998.
[24] H. Bergström, "On some expansions of stable distribution functions," *Arkiv Matematik*, vol. 2, no. 4, pp. 375–378, 1952.
[25] J. H. McCulloch, "13 financial applications of stable distributions," *Handbook Statist.*, vol. 14, pp. 393–425, 1996.