AAA: an Adaptive Mechanism for Locally Differential Private Mean Estimation (Technical Report)

Fei Wei
National University of Singapore
feiwei@nus.edu.sg

Ergute Bao
National University of Singapore
ergute@comp.nus.edu.sg

Xiaokui Xiao
National University of Singapore
xkxiao@nus.edu.sg

Yin Yang
Hamad Bin Khalifa University
yyang@hbku.edu.qa

Bolin Ding
Alibaba Group
bolin.ding@alibaba-inc.com

ABSTRACT

Local differential privacy (*LDP*) is a strong privacy standard that has been adopted by popular software systems, including Chrome, iOS, MacOS, and Windows. The main idea is that each individual perturbs their own data locally, and only submits the result noisy version to a data aggregator. Although much effort has been devoted to computing various types of aggregates and building machine learning applications under LDP, research on fundamental perturbation mechanisms has not achieved significant improvement in recent years. Towards a finer result utility, existing works in the literature mainly focus on improving the *worst-case* guarantee. However, this approach does not necessarily promise a better *average* performance given the fact that the data in practice obey various distributions.

In this paper, we propose the *advanced adaptive additive* (AAA) mechanism, which is a distribution-aware approach that addresses the average utility and tackles the classical *mean estimation* problem. AAA is carried out in a two-step approach: First, as the global data distribution is not available beforehand, the data aggregator selects a random subset of individuals to compute a (noisy) quantized data descriptor; then, in the second step, the data aggregator collects data from the remaining individuals, which are perturbed in a distribution-aware fashion. The perturbation involved in the latter step is obtained by solving an optimization problem, which is formulated with the data descriptor obtained in the former step and the desired properties of task-determined utilities. We provide rigorous privacy proofs and utility analyses, as well as extensive experiments comparing AAA with state-of-the-art mechanisms. The evaluation results demonstrate that the AAA mechanism consistently outperforms existing solutions with a clear margin in terms of result utility, on a wide range of privacy constraints and real-world and synthetic datasets.

PVLDB Reference Format:
Fei Wei, Ergute Bao, Xiaokui Xiao, Yin Yang, and Bolin Ding. AAA: an Adaptive Mechanism for Locally Differential Private Mean Estimation (Technical Report). PVLDB, XX(X): XXX-XXX, XXXX.
doi:XX.XX/XXX.XX

Proceedings of the VLDB Endowment, Vol. XX, No. X ISSN 2150-8097.
doi:XX.XX/XXX.XX

PVLDB Artifact Availability:
The source code, data, and/or other artifacts have been made available at https://github.com/adaptiveldpmechanism/vldb_2024_submission.

1 INTRODUCTION

Differential privacy (*DP*) [15] is a strong and mathematically rigorous metric that evaluates the privacy guarantee provided by randomization-based data-releasing mechanisms. The concept of DP was first proposed for a centralized setting involving a trusted data curator, which releases information (e.g., mean and heavy hitters) derived from an underlying dataset consisting of sensitive records. In such a setup, DP protects individuals' privacy by requiring that the data curator randomly perturbs the released information. In particular, the perturbation is carefully calibrated to ensure that using the released noisy results, an adversary equipped with arbitrary background knowledge can only have limited confidence when inferring about the underlying sensitive individual data records. Note that this centralized setting requires a trusted data curator, who has direct access to all sensitive records in the dataset. This might not be suitable for scenarios in which individuals do not want to expose their sensitive data to any party, including such a data curator. Moreover, even if the data curator itself can be trusted, it still faces the burden of protecting sensitive data against malicious intruders. Failure to do so may lead to data breaches (e.g., in the recent incidents involving Optus [1] and Samsung [2]), which violate the individuals' privacy and damage the reputation of the data curator.

Local differential privacy (*LDP*) [13, 17, 29] applies the notion of DP to a different setting, in which each individual (i.e., data owner) perturbs her data *locally*, and only submits the perturbed version of her data to an untrusted *data aggregator*. Therefore, the data aggregator only collects data that is already perturbed to satisfy the rigorous LDP requirements, and malicious parties might be less incentivized to intrude into such a data aggregator and steal the (less sensitive) randomized dataset. Perhaps for these reasons, LDP has gained adoption rapidly ever since its proposal. In particular, LDP has been applied in common software systems that collect usage information, which include Apple iOS and MacOS [3, 37], Microsoft Windows [11], and Google Chrome [16]. LDP has also found applications in online services such as Facebook (for gathering users' behavioral data for advertisement placements) [31], and Amazon (for collecting users' shopping preferences) [35].

[1] https://www.acma.gov.au/optus-data-breach
[2] https://www.securityweek.com/samsung-sued-over-recent-data-breaches

Driven by its successful applications, LDP has attracted much research attention in recent years. The majority of LDP-related papers focus on computing various types of statistics from the collected perturbed data, with the goal of maximizing result utility while satisfying LDP for each participating individual. These include range query [7, 9, 41], joint distribution [18, 47, 48] and marginal distribution estimation [8, 50], frequency/histogram [1, 5, 10, 30, 39], heavy hitters [4, 6, 32, 42], etc. We note that there is no single solution that consistently achieves high utility for all the tasks, as the utility metric varies according to different tasks.

In this paper, we focus on the fundamental LDP task of *mean estimation*, for which unbiasedness and minimal variance are desired utilities (see the detailed problem formulations in Section 3). Note that within this scope (i.e. mean estimation), there is only a narrow selection of works, including Duchi's mechanism [14], piecewise and hybrid mechanisms [38], that achieve consistently higher result utility than the classic LDP mechanisms, such as Laplace [15] and randomized response [24, 44]. For achieving better result utility, these works mainly focus on improving the *worst-case* utility guarantee. However, this approach does not necessarily provide a better *average* utility. For example, given an input data distribution dense in the *sub-optimal* regime of the mechanism, the average utility would be less ideal, even with a better worst-case guarantee. When this is true, it is natural to consider a distribution-aware approach that improves the *average* performance, instead of the overly pessimistic worst-case guarantee. This approach has not drawn sufficient attention in the literature, and motivates this work.

Our contributions: We propose the *advanced adaptive additive* (*AAA*) mechanism that is *adaptive* to the global data distribution while aiming to obtain high average result utility under strict LDP constraints. Achieving such a goal, however, is challenging since (i) the metric of utility varies according to the task, (ii) in the LDP setting, the global data distribution is not known beforehand, (iii) even when the data distribution is available, formulating and solving the *distribution-aware data perturbation* problem is still highly non-trivial, as elaborated later in Sections 3 and 4.

The proposed AAA mechanism addresses these challenges through a two-phase approach. In the first phase, the data aggregator randomly selects a subset of all participating individuals to estimate (a quantized representation of) the global data distribution, while LDP is preserved for every individual. In the second phase, each of the remaining individuals submits data perturbed by a calibrated mechanism. In particular, the perturbation in the second phase is carefully calibrated with the estimated global data distribution, which involves solving a convex optimization problem formulated for maximizing utility while enforcing LDP.

The most notable property of the perturbation noise in the second phase of AAA is that the distribution of the noise applied on the individual side depends on the value of her sensitive data. As a consequence, the utility with respect to a single individual varies as the input value varies. In particular, the more frequent data points are perturbed with smaller noises whereas the others are perturbed with larger noises, which is actually aligned with our goal of optimizing the average-case utility with respect to a data distribution. Figure 1 shows an example of conditional injected noise used in AAA that is tailored for a specific input distribution. This differs from the classic approaches such as the Laplace mechanism, where the added noise is independent of the input. More detailed results and explanations on AAA will be given in Section 4.

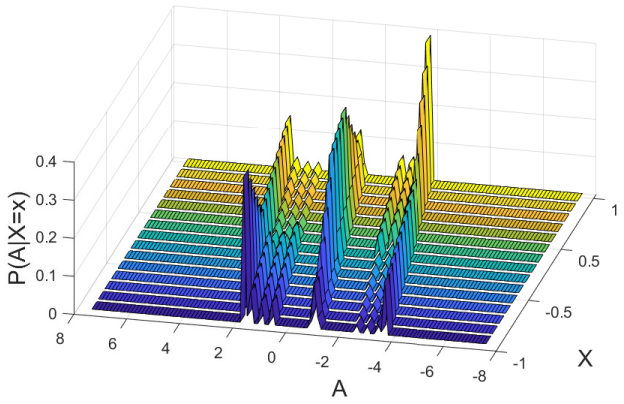


Figure 1: Conditional probability distribution for the random noise injected by the proposed AAA mechanism, assuming that the global data distribution is a Gaussian $N(0, 0.2^2)$ truncated to $\mathcal{X} = [-1, 1]$, and the privacy level of LDP is $\epsilon = 2$. Each horizontal banner in the plot represents a probability distribution function of the conditional additive noise random variable $A|X = x$ for $x \in [-1, 1]$.

We also provide rigorous privacy and utility analysis for the AAA mechanism. To complement our theoretical analysis, we also conduct extensive experiments using both real and synthetic data, demonstrating that AAA consistently outperforms existing solutions by a wide margin. The rest of this paper is organized as follows. Section 2 provides the necessary background on LDP definitions as well as existing LDP mechanisms. Section 3 formally defines the problem of distribution-aware data perturbation under LDP, and Section 4 presents the proposed AAA mechanism for this purpose. Section 5 establishes the privacy guarantees and the utility analysis of AAA. Section 6 discuss the related works and Section 7 presents the experimental evaluation results. Section 8 concludes the paper with directions for future work.

2 BACKGROUND

2.1 Preliminaries

We consider the following scenario: there is an untrusted data aggregator and a set of individual clients $\mathcal{U}_{\text{client}}$ (i.e., data owners), where the aggregator wishes to collect data from all clients. For example, the Census Bureau, as a data aggregator, may want to survey the average annual income of the entire adult population, who are the clients. Following common practice in the LDP literature, we assume that all parties are honest but curious, i.e., each party strictly follows the protocol, and at the same time tries to infer sensitive information from other parties. For simplicity, we assume that each client $u_i \in \mathcal{U}_{\text{client}}$ holds a single numeric data item x_i that follows a ground truth distribution P_X supported by $\mathcal{X}$, which is a finite interval $\mathcal{X} = [-\beta, \beta]$, where $\beta \in \mathbb{R}_+$. Note that here without loss of generality, we assume interval $\mathcal{X}$ is symmetric

to the origin since we can always translate and scale the numeric data into the designated interval.

In the case that each client holds multi-dimensional data, i.e. a tuple of attributes $\mathbf{x}_i = \left(x_i^{(1)}, \ldots, x_i^{(d)}\right) \in \mathcal{X}^d$, one could apply the standard trick in [38] that each client proceeds with a randomly chosen data entry, which reduces the problem to collecting one-dimensional data. Intuitively, this is (roughly) equivalent to randomly partitioning all clients in $\mathcal{U}_{\text{client}}$ into d non-overlapping subsets, each of which reports one of the d attributes. A detailed discussion, including algorithms and experiments, is provided in the full version [45]. In the rest of the paper, we focus on the one-dimensional case where each client holds a single numeric data item.

Following existing work [14, 38], we focus on the fundamental problem of *mean estimation*, and the goal is to minimize the error of the estimated mean computed at the data aggregator with respect to the ground truth. To protect sensitive information, each individual client i perturbs her local data (e.g., by injecting additive noise) before uploading it to the aggregator. This perturbation can be represented by a stochastic mapping $\mathcal{M} : \mathcal{X} \to \mathcal{Y}$. In the context of this work, $\mathcal{M}$ can also be seen as a *conditional distribution* $P_{Y|X}$.

In terms of the utility requirement for the perturbation mechanism, we want (i) the output to remain unbiased, that is, the expectation of the output and input needs to remain identical, and (ii) a smaller output variance to minimize the error in practice. These requirements form the utility metric of mean estimation tasks, which will play an important role in subsequent discussions.

In terms of the privacy constraint, we focus on the standard definition of pure *ϵ-local differential privacy* (LDP) [13].

Definition 1. *For any non-negative ϵ, a privatization mechanism $\mathcal{M}$ is said to satisfy ϵ-*local differential privacy *(LDP) if for any inputs $x \neq x' \in \mathcal{X}$ and subset $\mathcal{U} \subseteq \mathcal{Y}$, it holds that*

$$\Pr[\mathcal{M}(x) \in \mathcal{U}] \leq e^{\epsilon} \cdot \Pr[\mathcal{M}(x') \in \mathcal{U}].$$

The level of privacy protection is quantified by the parameter ϵ. Roughly speaking, a smaller ϵ makes it more difficult for the adversary (including the untrusted data aggregator) to infer the exact input value, given the perturbed version, and vice versa.

2.2 Data Perturbation Mechanisms under LDP

Classic Approaches. There are several classical data perturbation mechanisms in the literature that can be used to enforce LDP. First, we present the Laplace mechanism, assuming that each client's data item is within the interval $[-\beta, \beta]$ as described in the previous subsection.

Lemma 1 (Laplace Mechanism [15]). *The Laplace mechanism that injects Laplace noise of location parameter 0 and scale parameter $\frac{2\beta}{\epsilon}$ into x satisfies $(\epsilon, 0)$-local differential privacy (i.e., pure-LDP), where the probability density function of the Laplace noise of parameters 0 and $\frac{2\beta}{\epsilon}$ is defined as*

$$f(x) = \frac{\epsilon}{4\beta} \exp\left(-\frac{\epsilon|x|}{2\beta}\right).$$

In case X is not continuous but binary, i.e., $\mathcal{X} = \{-\beta, \beta\}$, one can apply the classical *randomized response* mechanism to enforce LDP.

Algorithm 1: Duchi et al.'s Solution for One-Dimensional Numeric Data [14]

Input : Client data $x \in [-\beta, \beta]$, Local DP parameter ϵ.

Output : Perturbed data $y \in \left\{-\frac{e^\epsilon+1}{e^\epsilon-1} \cdot \beta, \frac{e^\epsilon+1}{e^\epsilon-1} \cdot \beta\right\}$.

Sample a Bernoulli variable b such that $\Pr[b = 1] = \frac{e^\epsilon-1}{2e^\epsilon+2} \cdot \frac{x}{\beta} + \frac{1}{2}$.

if $b = 1$ **then**

  $y = \frac{e^\epsilon+1}{e^\epsilon-1} \cdot \beta$

else

  $y = -\frac{e^\epsilon+1}{e^\epsilon-1} \cdot \beta$

return y.

Lemma 2 (Randomized Response [24, 44]). *Let $x \in \{-\beta, \beta\}$ be the sensitive attribute to protect. The randomized response mechanism that randomly flips the sign of x satisfies $(\epsilon, 0)$-local differential privacy, if the probability of flipping is as follows*

$$p_{Y|X}(y|x) = \begin{cases} \frac{e^\epsilon}{1+e^\epsilon}, & \text{if } y = x, \\ \frac{1}{1+e^\epsilon}, & \text{otherwise.} \end{cases}$$

The original randomized response mechanism is only discussed for the case where the client's data is binary. For mean estimation over $\mathcal{X} = [-\beta, \beta]$, an extension of the randomized response mechanism proposed by Duchi et al. [14], presented next.

Duchi et al.'s Mechanism. Duchi et al. propose a mechanism to perturb data under LDP [14], shown in Algorithm 1. Specifically, the algorithm takes as input $X \in [-\beta, \beta]$ and outputs a binary $Y \in \left\{-\frac{e^\epsilon+1}{e^\epsilon-1} \cdot \beta, \frac{e^\epsilon+1}{e^\epsilon-1} \cdot \beta\right\}$ with probabilities

$$p_{Y|X}(y|x) = \begin{cases} \frac{e^\epsilon-1}{2e^\epsilon+2} \cdot \frac{x}{\beta} + \frac{1}{2}, & \text{if } y = \frac{e^\epsilon+1}{e^\epsilon-1} \cdot \beta, \\ -\frac{e^\epsilon-1}{2e^\epsilon+2} \cdot \frac{x}{\beta} + \frac{1}{2}, & \text{if } y = -\frac{e^\epsilon+1}{e^\epsilon-1} \cdot \beta. \end{cases}$$

It is easy to verify that Y is an unbiased estimator of X with variance

$$\mathrm{Var}[Y|X = x] = \beta^2 \cdot \left(\frac{e^\epsilon + 1}{e^\epsilon - 1}\right)^2 - x^2 \leq \beta^2 \cdot \left(\frac{e^\epsilon + 1}{e^\epsilon - 1}\right)^2. \quad (1)$$

Piecewise Mechanism and Hybrid Mechanism. Wang et al. propose the piecewise mechanism (PM) [38], shown in Algorithm 2. PM takes as input some data $X \in [-1, 1]$. For a more general case that $X' \in [-\beta, \beta]$, Ref. [38] shows that we can scale it to $X \in [-1, 1]$, apply PM, and return the scaled result $\beta \cdot y$. Unlike Duchi's mechanism, in PM the support of the output is not discrete but a finite interval $[-C, C]$ where

$$C = \frac{e^{\epsilon/2} + 1}{e^{\epsilon/2} - 1}.$$

The probability density function of the piecewise mechanism is a piecewise constant function

$$f_{Y|X}(y|x) := \begin{cases} p, & \text{if } y \in [l(x), r(x)], \\ \frac{p}{e^\epsilon}, & \text{if } y \in [-C, l(x)) \cup (r(x), C]. \end{cases}$$

Here, $p = \cdot\frac{e^\epsilon - e^{\epsilon/2}}{2e^{\epsilon/2}+2}$, $l(x) = \frac{C+1}{2} \cdot x - \frac{C-1}{2}$, and $r(x) = l(x) + C - 1$. Accordingly, given any numeric data $x \in [-1, 1]$, Ref. [38] proves

Algorithm 2: Piecewise Mechanism for One-Dimensional Numeric Data [38]

Input : Client data $x \in [-1, 1]$, Local DP parameter ϵ.
Output: Perturbed data $y \in [-C, C]$.

Sample α uniformly at random from [0,1]
if $\alpha < \frac{e^{\epsilon/2}}{e^{\epsilon/2}+1}$ **then**
  Sample y uniformly at random from $[l(x), r(x)]$
else
  Sample y uniformly at random from $[-C, l(x)) \cup (r(x), C]$
return y.

that that Y is an unbiased estimate of X, and the conditional variance of Y is

$$\mathrm{Var}[Y|X = x] = \frac{x^2}{e^{\epsilon/2} - 1} + \frac{e^{\epsilon/2} + 3}{3(e^{\epsilon/2} - 1)^2} \leq \frac{4e^{\epsilon/2}}{3(e^{\epsilon/2} - 1)^2}. \quad (2)$$

In the general case that $X \in [-\beta, \beta]$, the estimate remains unbiased, and the variance of the output would be multiplied by β^2. In the same paper [38], Wang et al. also propose a hybrid mechanism (HM) that combines Duchi's mechanism and PM above. According to the analysis in [38], HM achieves better *worst-case performance* compared to both of its underlying components, i.e., Duchi's mechanism and PM. However, as our experiments in Section 7 demonstrate, the result utility of HM is often worse than PM, on several real and synthetic datasets. This highlights the fact that high worst-case performance does not necessarily indicates high *practical performance*, which is the focus of this paper. We refer interested readers to the original paper for further details of HM.

3 PROBLEM FORMULATION

3.1 Rationale

Recall from Section 2.2 that the existing solutions for our problem setting, perturb the output regardless of the input value (or distribution). This misses the opportunity of further improving result utility by performing input-dependent perturbation, as is done in Laplace mechanism, Randomized Response, Duchi's mechanism (which generalizes Randomized Response), PM, and HM. On the other hand, observe that when the perturbation is input-dependent, the result utility can also depend on the input value. For instance, the variance of the output in Duchi's mechanism reaches its worst case when the input $x = 0$, according to Eq. (1). This means that in the unfortunate case that most clients hold a zero (or close to zero) value, Duchi's mechanism would perform poorly. Similarly, according to Eq. (2), the worst case of PM occurs when the absolute value of x is large (bounded by β in our problem setting), meaning that PM would perform poorly in the extreme case that most clients hold a value close to either β or $-\beta$. This issue is recognized in [38], and the authors attempt to address it with HM. However, HM focuses on optimizing worst-case performance, and its result utility is lower than PM on real data and synthetic data following common probability distributions, as shown in our experiments. This leads to the question: can we design an *adaptive* LDP mechanism that works well for common practical data?

The answer to the above question is clearly positive, *if each client already knows the global data distribution before she perturbs her data*. For instance, if all clients know that most data items are close to zero, then it is probably best to avoid Duchi's mechanism, whose worst-case utility is reached when $x = 0$, and instead apply a mechanism that promises good utility for close-to-zero inputs such as PM, according to Eq. (2). Meanwhile, observe that Duchi's mechanism, PM, and their combination HM are just three specific instances in the vast design space of LDP mechanism that performs input-dependent perturbation. In fact, even if all three methods have rather poor performance, e.g., when half of the clients have close-to-zero data values (worst case for Duchi's method) and the other half have close-to-β data values (worst case for PM), it is still possible to design an effective data perturbation scheme for such a scenario that *exploits the knowledge of the global data distribution.*

This above reasoning motivates a *distribution-aware data perturbation* mechanism for enforcing LDP, in which the perturbation depends on not only the input value, but also (an estimation of) the input value distribution. In the following subsections, we formalize this idea and define the optimization objective and constraints.

3.2 Optimization Problem Formulation

Motivated by the discussion in Section 2.1 and above, we start with the following abstract problem definition, and gradually fill out the mathematical details.

Problem 1 (Optimal Mean Estimation Mechanism under LDP). *Given a set of K individual clients, each holding a numeric data item x_i that is a sample of random variable $X \sim f_X$ with support $\mathcal{X} = [-\beta, \beta]$, design a privacy-preserving mechanism $\mathcal{M} : \mathcal{X} \to \mathcal{Y}$ that satisfies the following requirements:*

1. *Privacy-preserving: The mechanism $\mathcal{M}$ satisfies ϵ-*local differential privacy *for $\epsilon > 0$.*
2. *Bias-free: The perturbed estimate Y output by mechanism $\mathcal{M}$ is unbiased with respect to the private data X.*
3. *Minimized variance: The total variance of the perturbed output Y is minimized.*

Privacy requirement. Recall from Section 2.1 that in LDP, each client i applies the mechanism $\mathcal{M}$ to perturb her private data x_i independently, and $\mathcal{M} : \mathcal{X} \mapsto \mathcal{Y}$ can be represented as a conditional probability density function $f_{Y|X}$, i.e., $Y \sim f_{Y|X}$. By Definition 1, for any valid data items $x, x' \in \mathcal{X}$, and any outputs y, mechanism $\mathcal{M}$ should satisfy

$$f_{Y|X}(y|x) \leq e^{\epsilon} \cdot f_{Y|X}(y|x').$$

Bias-free requirement. Recall our assumption that the clients' data are samples of the random variable X obeying a ground truth probability density function f_X. Then, the true expectation of X is given by

$$\mathbb{E}[X] = \int_{\mathcal{X}} x \cdot f_X(x) dx.$$

Meanwhile, with the conditional $f_{Y|X}$, the expectation of Y is

$$\begin{aligned}\mathbb{E}[Y] &= \int_{\mathcal{Y}} y \cdot f_Y(y) dy \\ &= \int_{\mathcal{Y}} y \cdot \left(\int_{\mathcal{X}} f_{Y|X}(y|x) \cdot f_X(x) dx \right) dy.\end{aligned}$$

For the output Y to be unbiased, i.e., $\mathbb{E}[Y] = \mathbb{E}[X]$, it suffices to require:

$$\mathbb{E}[Y|X = x] = x \text{ for all } x \in \mathcal{X}.$$

Minimized variance requirement. Problem 1 states that we aim to minimize the total variance of Y, i.e. $\mathrm{Var}(Y)$. Next, we show that minimizing the total variance of Y is equivalent to minimizing the expected conditional variance $\mathbb{E}[\mathrm{Var}(Y|X)]$. By the law of total variance [46], we have

$$\mathrm{Var}(Y) = \mathbb{E}[\mathrm{Var}(Y|X)] + \mathrm{Var}(\mathbb{E}[Y|X]),$$

where the first term is the expected variance, and the second term

$$\begin{aligned}\mathrm{Var}(\mathbb{E}[Y|X]) &= \mathbb{E}[\mathbb{E}[Y|X]^2] - \mathbb{E}[\mathbb{E}[Y|X]]^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\ &= \mathrm{Var}(X).\end{aligned}$$

The equalities hold as we require Y to be bias-free, i.e., $\mathbb{E}[Y|X = x] = x$ for all $x \in \mathcal{X}$. Note that $\mathrm{Var}(X)$ is determined by the given data distribution and the data domain. Thus, it can be considered as a constant in the optimization problem. Accordingly, minimizing the expected conditional variance $\mathbb{E}[\mathrm{Var}(Y|X)]$ is equivalent to minimizing the total variance $\mathrm{Var}(Y)$.

Optimization program. The above discussions lead to the following optimization program:

$$\begin{aligned}\underset{f_{Y|X}}{\text{minimize}} \quad & \mathbb{E}[\mathrm{Var}(Y|X)] \\ \text{subject to} \quad & \mathbb{E}[Y|X = x] = x \text{ for all } x \in \mathcal{X}, \\ & f_{Y|X}(y|x) \le e^{\epsilon} \cdot f_{Y|X}(y|x') \\ & \text{for all } x, x' \in \mathcal{X}, y \in \mathcal{Y}.\end{aligned} \tag{3}$$

While the above constrained optimization program is a valid problem formulation, it is difficult to solve since (i) the global data distribution (i.e., f_X) is not known beforehand, and (ii) the problem is defined in a continuous domain, and it is unclear how to solve for the optimal $f_{Y|X}$, which has a vast search space. To make the problem tractable, we first quantize the representation for the distribution of variable X, explained in detail in the next subsection.

3.3 Quantized Input Data Distribution

Recall from Section 2.1 that each client $i \in [K]$ holds a private data item $x_i \in \mathcal{X} = [-\beta, \beta]$ which can be viewed as a sample drawn from the ground truth probability distribution f_X. To make it feasible to estimate f_X under LDP and to solve the optimization problem in Eq. (3), we discretize the continuous distribution f_X as follows. First, we define a quantization parameter, denoted by σ, as follows.

$$\sigma = \frac{|\beta - (-\beta)|}{N} = \frac{2\beta}{N}, \tag{4}$$

where N is a system parameter that controls the quantization granularity of f_X, which can be set, e.g., to a relatively large positive integer. Note that it is not necessary to normalize the sample space to an interval symmetric to the origin.

Using the quantization parameter σ defined above, we partition the support of the private data $\mathcal{X} = [-\beta, \beta]$ into non-overlapping bins $\mathcal{X}_1, \ldots, \mathcal{X}_N$ where

$$\mathcal{X}_j = \begin{cases} [-\beta, -\beta + \sigma] & \text{if } j = 1 \\ (-\beta + (j-1)\,\sigma, -\beta + j\sigma], & \text{if } 2 \le j \le N. \end{cases}$$

It is easy to verify that the non-overlapping bins span the support $\mathcal{X}$, i.e., $\bigcup_{j=1}^{N} \mathcal{X}_j = [-\beta, \beta] = \mathcal{X}$

The center of each $\mathcal{X}_j$ (which is also used as the index of each cell later) is

$$x_j = -\beta + \left(j - \frac{1}{2}\right)\sigma. \tag{5}$$

We are now ready to quantize the continuous distribution f_X on $\mathcal{X}$. We define an associated ground truth discrete distribution P_X on the index space $\{x_j : j = 1, \ldots, N\}$. Then, the mass function for P_X is defined as follows:

$$p_X(x_i) = \int_{\mathcal{X}_i} f_X(x)dx. \tag{6}$$

Here, p_X in the above equation represents the probability that a data point is sampled from bin $\mathcal{X}_i$ with respect to distribution f_X. The validity of the mass function as a probability measure is ensured by the fact that $\mathcal{X} = \bigcup_{i=1}^{N} \mathcal{X}_i$.

3.4 Global Data Distribution Estimation

Next we tackle the issue that the global data distribution f_X is not known beforehand. We approach the problem by estimating the quantized distribution p_x, defined in Eq. (6), under LDP. Clearly, this maps to an LDP-compliant histogram estimation problem, which has been studied extensively in the LDP literature, e.g., in [1, 5, 10, 30, 39]. The main difference here is that in our solution, we use a random subsample of the clients for this computation, as elaborated later in Section 4.1.

Let $\bar{P}_X$ be the estimated quantized distribution of X with mass function $\bar{p}_X$ under LDP. We require that $\bar{p}_X$ satisfy the following properties:

(1) Validity: $\bar{p}_X$ is a valid probability mass function, i.e.,

$$0 \le \bar{p}_X(x_j) \le 1 \text{ and } \sum_{j=1}^{N} \bar{p}_X(x_j) = 1.$$

(2) Bounded relative error: The relative error of the estimated probability mass associated with any bin is no larger than $\psi \ge 0$, i.e.,

$$\sup_{j=1,\ldots,N} \left| \frac{p_X(x_j) - \bar{p}_X(x_j)}{p_X(x_j)} \right| \le \psi.$$

Accordingly, our target problem is to solve the optimization program defined in Eq. (3), given an estimated quantized global distribution $\bar{P}_X$. We solve this problem in the next section.

4 THE AAA MECHANISM

4.1 Solution Overview

Algorithm 3 outlines the proposed solution for mean estimation under ϵ-LDP. Specifically, the solution contains two phases: in the first phase (Step 1), a random sample set of the clients jointly participate in a histogram estimation protocol that satisfies ϵ-LDP, as explained in Section 3.4. For instance, one way to perform such sampling is to let each client perform a Bernoulli test with probability s; if the test result is positive, then the client participates in the LDP-compliant histogram estimation protocol, and vice versa. Clearly, the size of the sample set is decided by s, which is a system

Algorithm 3: Outline of the Proposed Solution

Input : Collection of client data $\mathcal{D} = \{x_i : i \in [K]\}$, where $x_i \sim f_X$; LDP parameter ϵ.

Output : A mechanism $\mathcal{M}$ satisfying ϵ-LDP.

Step 1. Estimate a quantized data distribution $\bar{P}_X$ with a random subset of the clients under ϵ-LDP, and remove these clients from $\mathcal{D}$.

Step 2. Solve the optimization problem in Eq. (15) with respect to the LDP parameter ϵ using $\bar{P}_X$, to obtain a solution $\mathbf{Q}$.

Step 3. Use the solution $\mathbf{Q}$ to compute a discrete conditional distribution $\mathbf{P}_{A|X}$ with Eq. (12), and then obtain a piece-wise constant continuous conditional probability density function $f_{A|X}$ according to Eq. (16).

Step 4. Return mechanism $\mathcal{M} : \mathcal{X} \to \mathcal{Y}$ such that $y = \mathcal{M}(x) = a + x$, where a is sampled from the the conditional distribution $f_{A|X=x}$.

parameter. Accordingly to our experiments in Section 7, a reasonable value for s is around 10%, and the performance of the proposed solution is not sensitive to this hyperparameter. Note that clients in the sample set have depleted their privacy budget, and, thus, need to be excluded from the remaining part of the solution as stated in Step 1 of Algorithm 3. The histogram estimation process used in our implementation is detailed in Appendix A for completeness.

Then, in the second phase (Steps 2-4), the remaining clients participate in the sub-protocol for mean estimation, using the proposed AAA mechanism explained in this section that solves the optimization program defined in Eq. (3) under ϵ-LDP, using the estimated global data distribution $\bar{P}_X$ obtained in the first phase.

Given $\bar{P}_X$ from the first phase, it is a challenging task to perform the second phase, which involves solving the difficult optimization problem defined in Eq. (3). In the following, we first discretize the optimization problem in Section 4.2, and then transform the problem to a solvable form in Section 4.3.

4.2 Discretizing the Optimization Problem

One may notice that the ground truth data distribution P_X is not directly available to us, and we only have access to an estimate $\bar{P}_X$. Thus, the optimization problem we can solve is slightly different from the problem in (3). In this section, abusing the notation a bit, we reformulate the problem with respect to the discrete index set $\mathcal{X} = \{x_1, \ldots, x_N\}$, where x_i is defined in Eq. (5).

Similar to the classic solutions, the proposed AAA mechanism injects random noise into each client's data value to obtain the perturbed version. Let A be the added noise, which is a random variable obeying the conditional distribution $P_{A|X}$ with probability mass function $p_{A|X}$. For any $x \in \mathcal{X}$, we denote the random variable obtained by conditioning A on $X = x$ as $(A|X = x)$ or simply A_x. A sample from A_x is later added to the output of the query x for perturbation. Accordingly, the randomized output is $Y_x = x + A_x$. Here, let A_x be supported with alphabet $\mathcal{A}$ that

$$\mathcal{A} \triangleq \{j \cdot \sigma : j \in \mathbb{Z}\}, \tag{7}$$

where σ is the quantization constant defined previously in (4). For ensuring the estimate remains unbiased, it suffices to require that

$$\mathbb{E}[A|X = x] = \sum_{a \in \mathcal{A}} a \cdot p_{A|X}(a|x) = 0 \tag{8}$$

hold for any $x \in \mathcal{X}$, which implies $\mathbb{E}[Y|X = x] = x$ and, thus, $\mathbb{E}[X] = \mathbb{E}[Y]$. While maintaining unbiasedness, our goal is to minimize the conditional variance:

$$\begin{aligned}
\mathbb{E}[\mathrm{Var}(Y|X)] =& \mathbb{E}\left[(Y - \mathbb{E}[Y|X])^2 \,|X\right] \\
=& \sum_{x \in \mathcal{X}} \mathbb{E}\left(Y - \mathbb{E}[Y|X = x]\right)^2 \cdot p_X(x) \\
=& \sum_{x \in \mathcal{X}} \left(\mathbb{E}[Y^2|X = x] - \mathbb{E}^2[Y|X = x]\right) \cdot p_X(x) \\
=& \sum_{x \in \mathcal{X}} \left(\sum_{a \in \mathcal{A}} (x + a)^2 \cdot p_{A|X}(a|x) da - x^2\right) \cdot p_X(x) \\
=& \sum_{x \in \mathcal{X}} \left(\sum_{a \in \mathcal{A}} (x^2 + 2xa + a^2) \cdot p_{A|X}(a|x) - x^2\right) \cdot p_X(x) \\
\overset{(*)}{=}& \sum_{x \in \mathcal{X}} \left(\sum_{a \in \mathcal{A}} a^2 \cdot p_{A|X}(a|x)\right) \cdot p_X(x).
\end{aligned} \tag{9}$$

Here $(*)$ holds as by (8) we have

$$\begin{aligned}
&\sum_{x \in \mathcal{X}} \left(\sum_{a \in \mathcal{A}} 2xa \cdot p_{A|X}(a|x)\right) \cdot p_X(x) \\
=& \sum_{x \in \mathcal{X}} 2x \cdot \left(\sum_{a \in \mathcal{A}} a \cdot p_{A|X}(a|x)\right) \cdot p_X(x) \\
=& 0,
\end{aligned}$$

and $\sum_{a \in \mathcal{A}} p_{A|X}(a|x) = 1$ such that

$$\begin{aligned}
&\sum_{x \in \mathcal{X}} \left(\sum_{a \in \mathcal{A}} x^2 \cdot p_{A|X}(a|x) - x^2\right) \cdot p_X(x) \\
=& \sum_{x \in \mathcal{X}} \left(x^2 \cdot \sum_{a \in \mathcal{A}} p_{A|X}(a|x) - x^2\right) \cdot p_X(x) \\
=& \sum_{x \in \mathcal{X}} \left(x^2 - x^2\right) \cdot p_X(x) = 0.
\end{aligned}$$

By Eq. (9), Problem (3) reduces to the following optimization problem.

$$\begin{aligned}
\underset{\mathbf{P}_{A|X}}{\text{minimize}} \quad & \sum_{x \in \mathcal{X}} \left(\sum_{a \in \mathcal{A}} a^2 \cdot p_{A|X}(a|x)\right) \cdot p_X(x) \\
\text{subject to} \quad & p_{A|X}(y - x|x) \le e^{\epsilon} \cdot p_{A|X}(y - x'|x') \\
& \text{for all } x \ne x' \in \mathcal{X}, y \in \mathcal{Y}, \\
& \sum_{a \in \mathcal{A}} a \cdot p_{A|X}(a|x) = 0 \\
& \text{for all } x \in \mathcal{X}.
\end{aligned} \tag{10}$$

The above discretized version of the optimization program is still rather difficult to solve, since the alphabet $\mathcal{A}$ is unbounded as defined in (7), leading to an infinite-dimensional search space. In the next subsection, we deal with this tricky problem through a

novel definition of the conditional distribution $\mathbf{P}_{A|X}$, as well as a series of mathematical transformations of the optimization problem, arriving at a solvable form at the end of the section.

4.3 Solving the Optimization Problem

To address the problem that the alphabet $\mathcal{A}$ in Problem (10) is unbounded, we restrict the conditional distribution $\mathbf{P}_{A|X}$ to a special form that involves a finite number of unknowns. This is a main insight of the proposed AAA mechanism, and it is the key step towards transforming the optimization problem into a solvable form.

Specifically, given the quantization constant N defined in Section 3, for each $x_i \in \mathcal{X}$, we choose an integer $M \geq N$, and define a tunable vector

$$\mathbf{Q} = (\mathbf{q}^{(1)}, \ldots, \mathbf{q}^{(N)})$$
$$\mathbf{q}^{(i)} = (q_j^{(i)} \in \mathbb{R}^+ : j = -M, \ldots, M)$$

where it needs to be satisfied that the induced vector defines a valid discrete distribution, i.e.,

$$\sum_{j=-M+1}^{M-1} q_j^{(i)} + \frac{1}{1-r} \cdot \left(q_{-M}^{(i)} + q_M^{(i)}\right) = 1 \text{ for all } i = 1 \ldots, N. \quad (11)$$

Here $r \leq 1$ is a positive geometric constant. Then, we define

$$p_{A|X}(a_j|x_i) = \begin{cases} q_{-M}^{(i)} \cdot r^{-M-j}, & \text{if } j \leq -M, \\ q_j^{(i)}, & \text{if } |j| < M, \\ q_M^{(i)} \cdot r^{j-M}, & \text{if } j \geq M. \end{cases} \quad (12)$$

Intuitively, in the above definition of the conditional distribution $\mathbf{P}_{A|X}$, only the middle section involves a finite number of unknowns (i.e., elements of $\mathbf{Q}$), and the two edge sections are simply modeled by geometric series. Based on this definition, next we transform the infinite-dimensional problem defined in Eq. (10) to a finite-dimensional one.

In what follows, for simplicity, we denote $p_X(x_i)$ as p_i. The objective function in Eq. (10) under our setting can be restated as

$$\underset{\mathbf{Q}}{\text{minimize}} \quad \sum_{i=1}^{N} p_i \cdot \left(\sum_{j=-\infty}^{-M} a_j^2 \cdot q_{-M}^{(i)} r^{|j|-M} + \sum_{j=-M+1}^{M} a_j^2 \cdot q_j^{(i)} + \sum_{j=M}^{\infty} a_j^2 \cdot q_M^{(i)} r^{|j|-M} \right).$$

Recall that $a_j = \sigma \cdot j$, so

$$\sum_{j=M}^{\infty} a_j^2 \cdot q_M^{(i)} r^{j-M} = \sigma^2 \cdot q_M^{(i)} \sum_{j=0}^{\infty} (j+M)^2 r^j \overset{(*)}{=} \sigma^2 \cdot q_M^{(i)} \cdot \left(\frac{M^2}{1-r} + \frac{(2M-1)r}{(1-r)^2} + \frac{2r}{(1-r)^3} \right).$$

Note that equality $(*)$ is obtained as follows. Let

$$S = \sum_{j=0}^{\infty} (j+M)^2 r^j = M^2 + \sum_{j=1}^{\infty} (j+M)^2 r^j,$$

then

$$rS = \sum_{j=0}^{\infty} (j+M)^2 r^{j+1} = \sum_{j=1}^{\infty} (j+M-1)^2 r^j,$$

and

$$(1-r)S = M^2 + \sum_{j=1}^{\infty} (2j+2M-1) r^j = (2M-1) \sum_{j=1}^{\infty} r^j + 2 \sum_{j=1}^{\infty} j \cdot r^j. \quad (13)$$

Similarly, let $S' = \sum_{j=1}^{\infty} j r^j = r + \sum_{j=2}^{\infty} j r^j$. We have

$$(1-r)S' = r + \sum_{j=2}^{\infty} r^j = r + \frac{r^2}{1-r} = \frac{r}{1-r},$$

thus

$$S' = \frac{r}{(1-r)^2}. \quad (14)$$

Now, plugging (14) into (13), we obtain

$$S = \frac{M^2}{1-r} + \frac{(2M-1)r}{(1-r)^2} + \frac{2r}{(1-r)^3}.$$

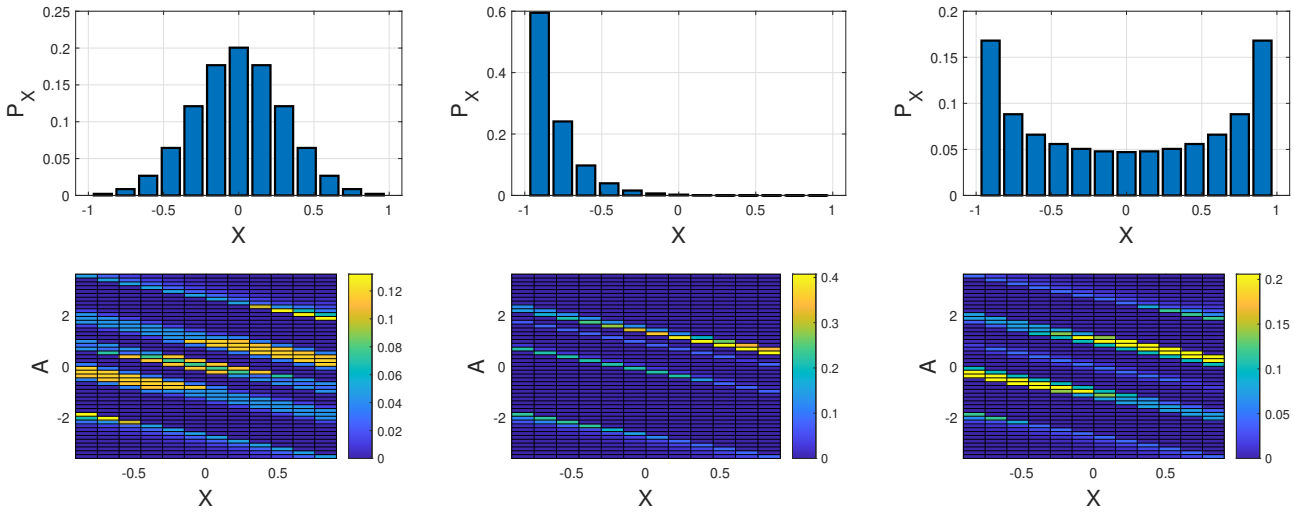
Finally, we obtain an optimization problem as follows:

$$\begin{aligned} \underset{\mathbf{Q}}{\text{minimize}} \quad & \sum_{i=1}^{N} p_i \cdot \left[\sum_{j=-M+1}^{M-1} a_j^2 \cdot q_j^{(i)} + \left(\frac{M^2}{1-r} + \frac{(2M-1)r}{(1-r)^2} + \frac{2r}{(1-r)^3} \right) \cdot \sigma^2 \left(q_{-M}^{(i)} + q_M^{(i)} \right) \right] \\ \text{subject to} \quad & q_j^{(i)} \leq e^{\epsilon} \cdot q_{j'}^{(i')} \\ & \text{for all } (i,j) \neq (i',j') \text{ that } a_j + x_i = a_{j'} + x_{i'}, \\ & \sum_{j=-M+1}^{M-1} q_j^{(i)} + \frac{1}{1-r} \cdot \left(q_{-M}^{(i)} + q_M^{(i)} \right) = 1 \\ & \text{for all } i \in \{1, \ldots, N\}, \\ & q_j^{(i)} \geq 0, \text{ for all } j \in \{-M, \ldots, M\}. \end{aligned} \quad (15)$$

The above optimization problem can then be solved by a standard numerical solver. Let $\mathbf{Q}^*$ denote the solution to the above problem. It is then fed to Step 3 of the algorithm to compute the conditional distribution $\mathbf{P}_{A|X}$ for the added noise. Recall that each bin $\mathcal{X}_i$ is associated with an index $x_i \in \mathcal{X}$, and the ground truth distribution fX is assumed to be continuous. Therefore, based on $\mathbf{P}_{A|X}$, we define a conditional probability density function $f_{A|X}$ as follows:

$$f_{A|X}(a|x) = p_{A|X}(a|x_i) \text{ where } x \in \mathcal{X}_i. \quad (16)$$

This completes the proposed AAA mechanism, which is returned in Step 4 of the algorithm. We provide formal proofs and utility analyses of AAA in the next section. Figure 2 visualizes the optimized conditional noise distribution with respect to several common types of data distributions, in which the proposed AAA mechanism satisfies 2-LDP. As one can observe from Figure 2, in spite of the non-smoothness caused by the numerical solver, not surprisingly, given symmetric data distributions, the resulting noises are also symmetric, and vice versa. Moreover, to minimize the variance, the solver tends to concentrate the resulting noise around zero, with negative noise skew on positive valued inputs and vice versa. Meanwhile, there are probability masses on the tail ends to enforce unbiasedness, which results in an imbalanced multi-peak shape.



(a) Gaussian Distribution (b) Exponential Distribution (c) Beta Distribution

Figure 2: The optimized noise distribution of the AAA mechanism for several common data distributions (all truncated within the interval $[-1, 1]$). The bar plots on the top show the (quantized) data distribution for the data evaluations X, while the depth maps below represent the conditional distributions. Each column in the depth maps represents the discrete distribution conditioning on X. The brightness of the color is associated with the weight of the probability mass, as indicated by the color bar to the right of the figure. As we can see from the plots, the optimized noise distribution varies according to the data distribution.

5 THEORETICAL ANALYSES

5.1 Privacy Analysis

We now demonstrate that the AAA mechanism $f_{Y|X}(X) = X + A_X$ where $A_X \sim f_{A|X}$ satisfies ϵ-LDP. Before we proceed, we present a lemma that will be useful later on.

Lemma 3. *Given $y \in \mathbb{R}$, for any two pairs $(a_j, x) \neq (a_{j'}, x')$ such that $a_j + x = a_{j'} + x' = y$, $x \in \mathcal{X}_i$ and $x' \in \mathcal{X}_{i'}$, it always holds that $a_j + x_i = a_{j'} + x_{i'}$.*

Proof. For any $x \in \mathcal{X}$ that belongs to the bin $\mathcal{X}_i$ centered at x_i with a radius of $\sigma/2$, we can write

$$x = x_i + \gamma \text{ for some } \gamma \in \left(-\frac{\sigma}{2}, \frac{\sigma}{2}\right]. \tag{17}$$

Note that the interval is closed on the left end as well for $i = 1$.

Suppose we have observed a value $y \in \mathbb{R}$, which represents the output of the mechanism in practice. Let (a_j, x) and $(a_{j'}, x')$ be any pair of values such that $a_j + x = a_{j'} + x' = y$. We assume that x and x' fall into the intervals $\mathcal{X}_i$ and $\mathcal{X}_{i'}$, respectively, with centers at x_i and $x_{i'}$. Using Eq. (17), let $\bar{x} = x' - \gamma$, we can write

$$a_j + x - \gamma = a_j + x_i = a_{j'} + \bar{x}. \tag{18}$$

Note that x' belongs to the set $\mathcal{X}$ and is uniquely determined by our definitions of $\mathcal{A}$ and $\mathcal{X}$. We will now show that $\bar{x}$ is equal to x'_i.

Recall that $\bar{x} = x' - \gamma$, and using Eq. (17), we have $|\gamma| \leq \sigma/2$. By our definition of the intervals $\mathcal{X}_i$, if there exists a value $\bar{x} \in \mathcal{X}$ such that $|x' - \bar{x}| \leq \sigma/2$, where $x' \in \mathcal{X}_{i'}$, then we know that $\bar{x} = x_{i'}$. Therefore, the claim holds immediately. □

We are now ready to present our main claim, as follows.

Claim 1. *Given the solution $\mathbf{Q}^*$ of Problem (15) with respect to privacy parameters ϵ, and a ground truth distribution f_X, we define a discrete distribution $\mathbf{P}^*_{A|X}$ by Eq. (12), and induce a piece-wise function:*

$$f_{A|X}(a|x) = p^*_{A|X}(a|x_i)/\sigma \text{ where } x \in \mathcal{X}_i. \tag{19}$$

And we define a mechanism $\mathcal{M} : \mathcal{X} \to \mathcal{Y}$ as $\mathcal{M}(X) = X + A_X$, where $A_X \sim f_{A|X}$. Then, $\mathcal{M}$ satisfies ϵ-local differential privacy.

Proof. To prove that $\mathcal{M}$ satisfies ϵ-local DP, it suffices to show that for any y, it holds that

$$f_{A|X}(a|x) \leq e^{\epsilon} \cdot f_{A|X}(a'|x') \tag{20}$$

where $(a, x) \neq (a', x')$ and $a + x = a' + x' = y$. To prove this, we can revisit the solution $\mathbf{P}_{A|X}$ of Problem (15) and use Lemma 3, which tells us that for any $(a, x) \neq (a', x')$ that $a + x = a' + x'$, where $a, a' \in \mathcal{A}$, $x \in \mathcal{X}_i$, and $x' \in \mathcal{X}i'$, it holds that $a + x_i = a' + x_{i'}$. Therefore,

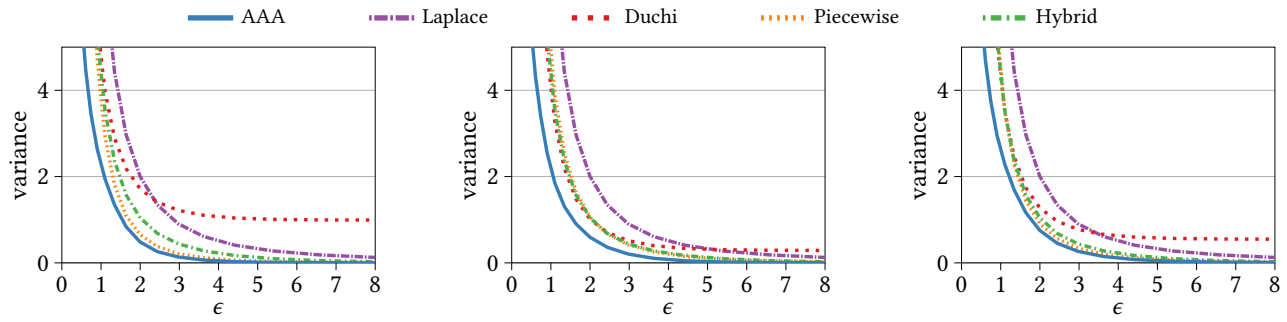
$$p^*_{A|X}(a|x) \leq e^{\epsilon} \cdot p^*_{A|X}(a'|x'). \tag{21}$$

Using Eq. (21) and the definition of $f_{A|X}$ in Eq. (19), we can immediately see that Eq. (20) holds, which completes the proof. □

5.2 Utility Analysis

Recalling the assumption that the error for the estimated distribution of the data is bounded as stated in Section 3.4, we make the following claim.

Claim 2. *Given a ground truth (continuous) distribution f_X over $\mathcal{X}$, let p_X be the quantization of f_X, and $\tilde{p}_X$ be an estimation of p_X.*



(a) Gaussian Distribution. **(b) Exponential Distribution.** **(c) Beta Distribution.**

Figure 3: Comparisons of the expected variance on different input data distributions.

Assume we have

$$\left|\frac{p_X(x) - \tilde{p}_X(x)}{p_X(x)}\right| \leq \psi \text{ for any } x \in \mathcal{X}, \tag{22}$$

and obtain an optimization solution given $\tilde{p}_X$ as $\mathbf{p}_{A|X}$. With respect to the solution $\mathbf{p}_{A|X}$, we define the expected output variance under f_X, $\tilde{p}_X$ as V (the truth variance), $\tilde{V}$ (the estimated variance), respectively. Define the relative error $\phi \triangleq \frac{\tilde{V}-V}{\tilde{V}}$. Then, it holds that $|\phi| \leq \min\left\{\frac{\psi}{1+\psi}, \frac{\psi}{1-\psi}\right\}$.

Proof. By the assumptions in the claim, the estimated (discrete) distribution $\tilde{P}$ has a relative error no larger than ψ with respect to the true (quantized) distribution p_X. Given the estimate $\tilde{p}_X$, we obtain the optimal conditional noise distributions $\mathbf{p}_{A|X}$ which minimize the total variance of the output, which is defined as

$$\tilde{V} = \sum_{i=1}^{N} \tilde{p}_X(x_i) \cdot \left(\sum_{a\in\mathcal{A}} a^2 \cdot p_{A|X}(a|x_i)\right).$$

We apply the converted $f_{A|X}$ as of the previous section, which is defined in Eq. (19). Given the ground truth distribution on $\mathcal{X}$ is f_X and the variance is accordingly

$$\begin{aligned} V &= \int_{\mathcal{X}} f_X(x) \left(\sum_{a\in\mathcal{A}} a^2 \cdot f_{A|X}(a|x)\right) dx \\ &= \sum_{i=1}^{N} \int_{\mathcal{X}_i} f_X(x) \left(\sum_{a\in\mathcal{A}} a^2 \cdot f_{A|X}(a|x)\right) dx \\ &= \sum_{i=1}^{N} \int_{\mathcal{X}_i} f_X(x)dx \cdot \left(\sum_{a\in\mathcal{A}} a^2 \cdot p_{A|X}(a|x_i)\right) \\ &\overset{(*)}{=} \sum_{i=1}^{N} p_X(x_i) \cdot \left(\sum_{a\in\mathcal{A}} a^2 \cdot p_{A|X}(a|x_i)\right). \end{aligned}$$

Here $(*)$ holds by (6). As $\tilde{V} > 0$, we have

$$\phi = \frac{\sum_{i=1}^{N} (p_X(x_i) - \tilde{p}_X(x_i)) \cdot \sum_{a\in\mathcal{A}} a^2 \cdot p_{A|X}(a|x_i)}{\sum_{i=1}^{N} \tilde{p}_X(x_i) \cdot \sum_{a\in\mathcal{A}} a^2 \cdot p_{A|X}(a|x_i)}. \tag{23}$$

By (22), it always holds that

$$-\frac{\psi}{1+\psi} \leq \frac{p_X(x) - \tilde{p}_X(x)}{\tilde{p}_X(x)} \leq \frac{\psi}{1-\psi}. \tag{24}$$

By (23) and RHS of (24),

$$\phi \leq \frac{\sum_{i=1}^{N} \frac{\psi}{1-\psi} \cdot \tilde{p}_X(x) \cdot \sum_{a\in\mathcal{A}} a^2 \cdot p_{A|X}(a|x_i)}{\sum_{i=1}^{N} \tilde{p}_X(x_i) \cdot \sum_{a\in\mathcal{A}} a^2 \cdot p_{A|X}(a|x_i)} = \frac{\psi}{1-\psi} \tag{25}$$

Similarly, we can show $\phi \geq \psi/(1+\psi)$, completing the proof. □

6 RELATED WORKS

Data perturbation under central DP. The Laplace mechanism described in Section 2 was originally designed for the centralized setting of differential privacy with a trusted data curator. In the following, we review several notable fundamental data perturbation mechanisms for enforcing DP in the centralized setting that (often) outperforms the Laplace mechanism and the Gaussian mechanism. Note that assumptions for the centralized setting are radically different from those for LDP, as mentioned in Sections 1 and 2. Most importantly, in the centralized model, the trusted data curator can directly access the exact values of *all* records in the sensitive dataset, which is not possible for any client or the data aggregator in LDP. Accordingly, DP mechanisms designed for the centralized setting that relies upon a trusted data curator are generally inapplicable to LDP. Unless otherwise stated, the solutions reviewed next do not apply to our target problem setting, and, thus, are orthogonal to this work.

Towards better utility, Ghosh et al. [23] propose the geometric mechanism for enforcing ϵ-DP, for which the distribution of the injected noise is formulated as a two-sided geometric series with tunable parameters and can be viewed as an optimized variant of the Laplace mechanism. In particular, Ref. [23] obtains improved result utility under the Bayesian optimization model, in which the privacy practitioner has prior knowledge of the input distribution to calibrate the randomization mechanism. This assumption is fundamentally different from our setting, where no prior knowledge about the input distribution is given.

Another notable series of works that achieve a certain notion of optimality is the staircase mechanisms [20–22]. Specifically, these methods assume that the query (i.e., statistics to be released under DP) is a real-valued function, and the mechanism design aims to minimize the worst-case utility loss for the perturbed output subject to the privacy constraints, which can be formulated as an optimization problem. Specifically, the optimization objective is formulated as a piecewise constant probability density function,

which is symmetric to the origin and decreases geometrically, leading to noise distribution with a symmetric staircase shape. It is worth mentioning that the idea of this work is remotely related to the Cactus mechanism [2], which works towards the optimal noise injection mechanism for large-composition, i.e., answering a large number of queries, under the central DP setting through numerical optimization.

The interactive LDP model. The notion of interactive DP model can be traced back to the classical Ref. [15]. There are a few works in the scope of LDP setting [19, 25–27, 34]. Regarding an interactive model, clients respond to the data aggregator's queries in a sequential manner, and clients are permitted to observe the preceding responses from other clients which affect the perturbation applied to their own data. AAA mechanism falls into this category as we apply a two-step approach, where the responses from a randomly selected portion of clients directly affect the perturbation mechanism applied to the rest of the clients. It has been shown that interactive LDP mechanisms are stronger than the non-interactive ones for tasks such as population quantiles, logistic regression, supporting vector machine, and estimating an unknown Gaussian distribution [19, 25, 26, 34]. On the contrary, the problem of empirical mean estimation for a given set of data points, which is the focus of this paper, has not been investigated under the interactive LDP framework. Our work fills this gap by showing that AAA, which is interactive, achieves better utility than the existing (non-interactive) solutions. Further showing a theoretical separation between interactive LDP and non-interactive LDP mechanisms for mean estimation (and other tasks) is a promising future work direction.

Other LDP mechanisms. There are LDP mechanisms originally designed for tasks other than mean estimation, such as [5, 30, 43, 49]. They are sub-optimal for mean estimation, since their desired utility metrics are fundamentally different from ours (as mentioned in Section 1). For completeness, we provide empirical evaluations of the AAA mechanism against the representative mechanisms in the LDP literature in the full version [45]. We also include the performance of the state-of-the-art solution for frequency estimation [30] in Section 7 as a reference. Overall, our results show that the AAA mechanism consistently outperforms these mechanisms by a large margin for mean estimation.

7 EXPERIMENTS

Our experimental evaluation consists of two parts. In the first part (Section 7.1), we compare the proposed AAA mechanism against the previous state-of-the-art fundamental LDP mechanisms in terms of theoretical performance. In the second part (Sections 7.2-7.4), we evaluate our mechanism's performance on various real and synthetic datasets for the targeted mean estimation task.

7.1 Evaluations of Theoretical Performance

We first evaluate the theoretical performance of the proposed AAA mechanism through numeric simulations. Specifically, we compare the expected variance of the AAA mechanism against that of the previous state-of-the-art mechanisms under the following three different types of data distributions[3]: (1) Gaussian distribution $N(x; 0, 0.1^2)$, which is symmetric to the center of the sample domain (origin). (2) Exponential distribution $\text{Exp}(x + \beta; 6)$, which has probability mass accumulating on one side of the sample domain. (3) Beta distribution $\text{Beta}((x + \beta)/2; 0.5, 0.5)$, whose probability mass concentrates on both ends of the domain.

The proposed AAA mechanism is compared against the following competitors: the Laplace mechanism, Duchi et al.'s Mechanism, the piece-wise mechanism, and the hybrid mechanism, presented in Section 2.2. Since these competitors are described under pure-LDP, we vary ϵ in the range $(0, 8]$.

When solving the discrete optimization problem of Eq. (15), We set the sample space to $[-1, 1]$ (i.e., $\beta = 1$), and the quantization parameter $\sigma = 0.02$ (i.e., $N = 100$). We set $M = k \cdot N$ for $k = 3$ (for the validity of our model, we require $k \geq 1$, by our experiment $k \geq 3$ is sufficient as the tails of the optimized probability mass function attenuate quickly when the noise magnitude $a > \beta$), such that $q_j^{(i)}$ spans between $[-3, 3]$. The geometry series parameter $r = 0.5$; through the experiments, we find that the choice of r has a negligible impact as long as $r < 0.6$.To initiate the numeric optimization solver, we assign a uniform discrete distribution to the sequence $(q_j^{(i)})$.

We plot the expected variance as a function of the privacy parameter ϵ for each of the mechanisms given the ground truth data distributions in Figure 3. Clearly, the proposed AAA mechanism significantly outperforms all of its competitors in all cases. For example, as depicted in Figure 3(a), at $\epsilon = 1$, AAA mechanism's expected variance is approximately 2 and the best result from its competitors is approximately 4. When $\epsilon \to 0$, the performance between AAA and other solutions expands rapidly for all three data distributions.

Another interesting observation from the evaluation results is that although the hybrid mechanism is designed to be an optimized combination of Duchi's mechanism and the piecewise mechanism, its performance under the three data distributions evaluated is sometimes lower than that of the piecewise mechanism, e.g., in Figure 3(a). This confirms our observation that high worst-case performance does not necessarily lead to high average-case performance. Next, we provide empirical results on synthetic and real-world datasets.

7.2 Evaluation Results on Synthetic Data

Next, we present the experimental evaluation results, using synthetic data drawn from the following probability distributions: the Gaussian distribution $N(0, 1)$, Exponential distribution $\text{Exp}(1)$, and Beta distribution $\text{Beta}(0.5, 0.5)$. As these distributions are unbounded, we clip each sample to a range of $[low, high]$, where we set $high = 5$ for all distributions, $low = -5$ for Gaussian, and $low = 0$ for Exponential and Beta distributions. For values outside of this range, we apply truncation, such that if $d > high$, then $d \leftarrow high$; if $d < low$, then $d \leftarrow low$; and d remains unchanged otherwise. For the task of density estimation for our AAA mechanism, we use the randomized response LDP-compliant histogram estimation algorithm, detailed in Appendix A.2. As mentioned in Section 3.4, the histogram estimation process is not the focus of this paper, and it can be performed with any known solution.

[3] All three distributions are truncated to the range $[-1, 1]$.

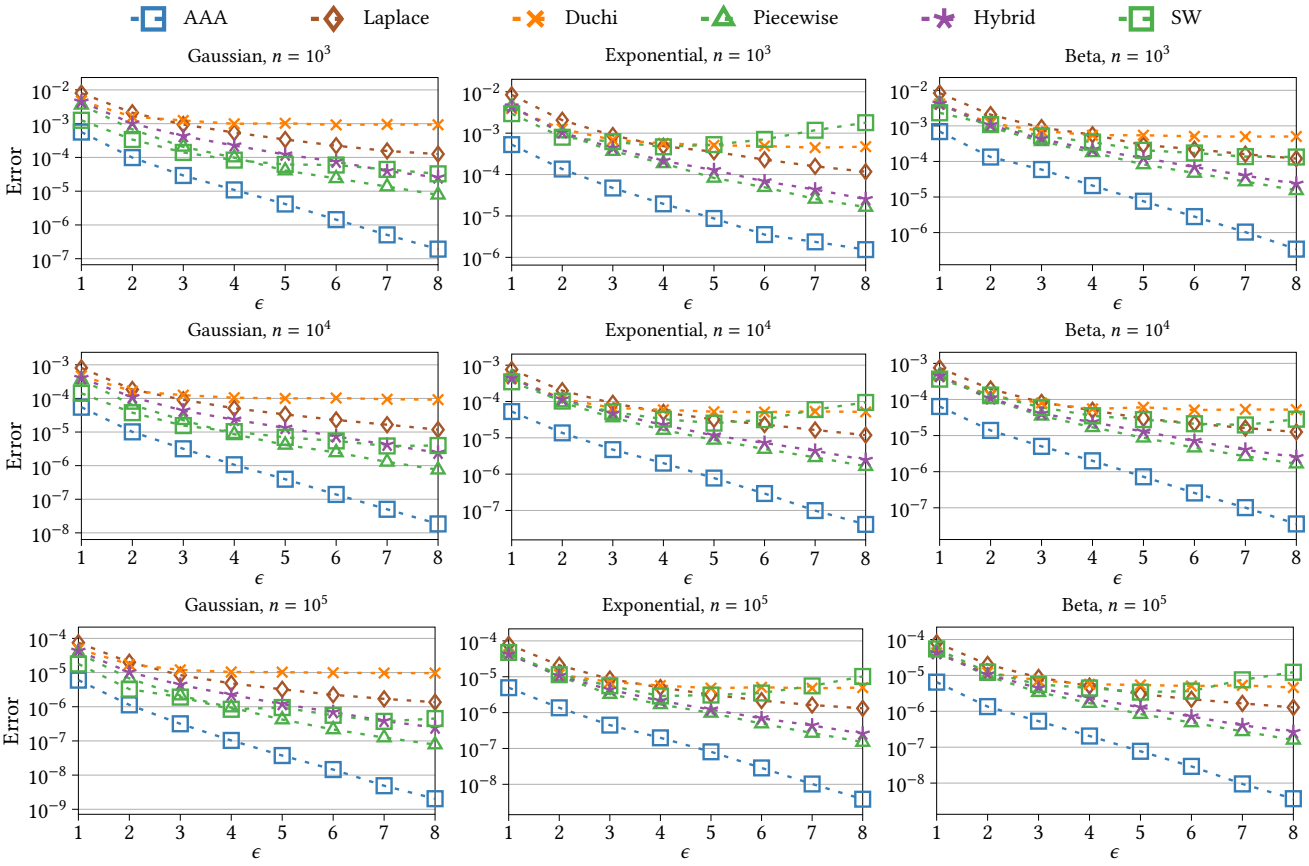


Figure 4: Performance on synthetic data following different distributions with varying privacy parameters.

We further linearly transform each clipped data point to the range $[-1, 1]$ for all distributions. We vary the number of data samples $n \in \{10^3, 10^4, 10^5\}$. Results on higher values of n lead to similar conclusions, are are omitted for brevity.

We consider five competitors for AAA mechanism: Laplace mechanism, Duchi's mechanism [14], piecewise and hybrid mechanisms [38] and the SW mechanism [30] with the suggested hyperparameters. Note that, unlike the other four competitors, SW [30] is designed for distribution (histogram) estimation rather than mean estimation.

We report the average error under varying privacy parameters with ϵ taken from $\{1, 2, 3, 4, 5, 6, 7, 8\}$. For estimating the input distribution, we use randomized response with a portion $s = 0.1$ of the private input data to obtain a noisy histogram. Then, we use the estimated input distribution to instantiate AAA mechanism, and inject the additive noise to the remaining $(1 - s)$ portion of the private data. For the AAA mechanism, we set $N = |\mathcal{X}|$ to 4, and $q = \frac{|\mathcal{A}|}{|\mathcal{X}|}$ to 4 for all experiments.

We present the experimental results in Figure 4, which reports the average squared error between the estimated mean and the true mean over 1000 independent tests. Observe that the proposed AAA mechanism (both pure-DP and approximate-DP versions) outperforms its competitors by a wide margin across all data distributions, dataset sizes, and privacy constraints. The performance gap expands with higher values of the privacy parameter ϵ, and can reach more than an order of magnitude when $\epsilon \geq 7$. Note that a high $\epsilon = 8$ has been used in practice, e.g., in Apple QuickTime [3].

SW [30] has lower performance compared to other methods such as the piecewise mechanism [38] in some settings, since SW is designed for histogram estimation rather than mean estimation.

7.3 Evaluation Results on Real-World Data

Taxi [36]. This dataset contains the duration of trips (in seconds) from 2018 January New York green taxi data. We extract the trips that are shorter than 24 hours (i.e., 86400 seconds), resulting in $n = 792,743$ samples.

Income [33]. This dataset contains income information from the 2018 American Community Survey. The range of the original income attribute is $[-11200, 1423000]$. Following Ref. [30], we extract the values that are non-negative and smaller than $524,288$, resulting in $n = 2,689,888$ samples.

Retirement [28]. This dataset contains the employee compensation from San Francisco. The range is $[-30621.43, 121952.52]$. Following [30], we extract values that are non-negative and smaller than $60,000$, resulting in $n = 682,410$ samples.

For all three datasets, we normalize the data to $[-1, 1]$, and report the average error over 1000 independent runs. We use the same setting of hyperparameters for AAA mechanism as in the

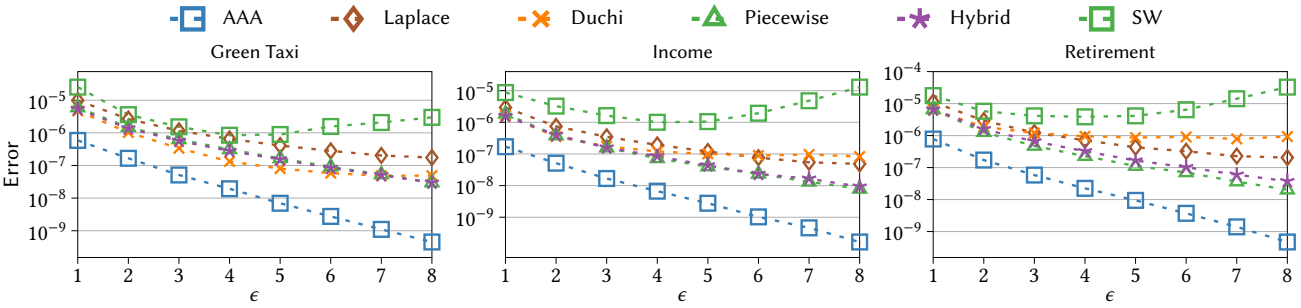
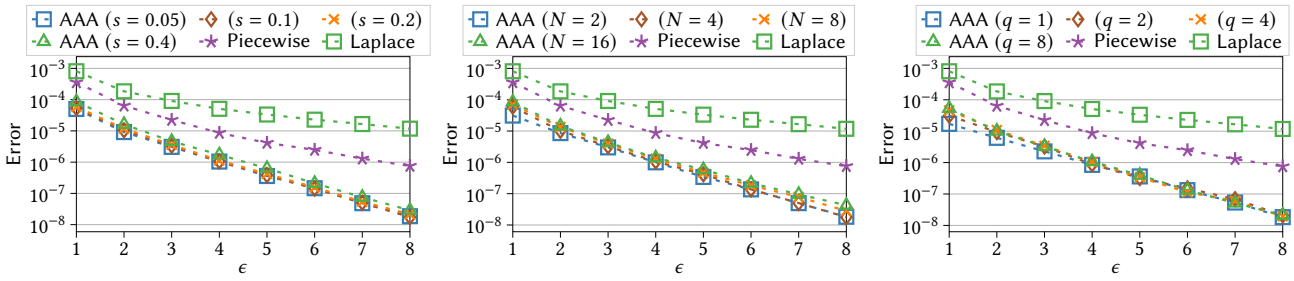


Figure 5: Performance on real-world datasets under varying privacy parameters.



(a) $N = 4$, $q = 4$. **Vary split ratio** s. **(b)** $s = 0.1$, $q = 4$. **Vary number of bins** N. **(c)** $s = 0.1$, $N = 4$. **Vary** $axRatio$ q.

Figure 6: Performance of pure-DP AAA mechanism **on Gaussian data of size** $n = 10^4$.

experiments for synthetic data. As we can see from Figure 5, the results on real-world datasets lead to the same conclusion that AAA mechanism significantly outperforms its competitors on all datasets under different privacy constraints.

7.4 Effect of Hyperparameters

Recall that there are three hyperparameters in our mechanism AAA mechanism: split ratio s, which quantifies the portion of data used for distribution estimation; N, the number of bins to split the input data range; and the ratio $|\mathcal{A}|/|\mathcal{X}|$, denoted as q, which quantifies the ratio between the range of the tunable vector (which is later used for defining the additive noise) and the range of the input private data. We conduct additional experiments to study the effect of varying the aforementioned hyperparameters on AAA mechanism, using $n = 10,000$ data samples generated from the standard Gaussian distribution truncated to $[-1, 1]$. We focus on the pure-DP case and omit the approximate-DP one as the results are similar. We use the representative Laplace mechanism [15] and Piecewise mechanism [38] as the baselines for comparison.

First, we vary the value of split ratio $s = 0.005, 0.1, 0.2, 0.4$ while fixing $N = 4$ and $q = 4$ (see Figure 6a). Next, we vary the value of $N = 2, 4, 8, 16$ while fixing $q = 4$ (see Figure 6b). Finally, we vary the value of $q = 1, 2, 4, 8$ while fixing N to 4 and s to 0.1 (see Figure 6c). The results in Figures 6a, 6b suggest that both hyperparameters s and N have a negligible impact on the performance of AAA mechanism. Figure 6c also implies that hyperparameter q has a negligible impact on the performance of AAA mechanism, except when ϵ is small (i.e., $\epsilon = 1$). With this set of experiments, we conclude that AAA mechanism is highly robust to hyperparameter settings, and it outperforms existing solutions under a wide range of data distribution, privacy constraints, and hyperparameter settings. We defer experiments on high-dimensional data to the full version [45] [45].

8 CONCLUSION

In this work, we present a local differential privacy-preserving mechanism that adapts to the underlying data distribution, whose main component is an algorithm that generates additive noise that is carefully calibrated to achieve an optimal trade-off between privacy guarantee and utility. The proposed AAA mechanism outperforms previous methods in terms of practical performance on common data distributions according to our extensive experiments.

Regarding future work, an interesting direction is to investigate the optimization for utility metrics beyond variance in mean estimation. Further, we intend to apply the proposed AAA mechanism to more complex LDP applications to obtain improved result utility.

REFERENCES

[1] Jayadev Acharya, Ziteng Sun, and Huanyu Zhang. 2018. Hadamard Response: Estimating Distributions Privately, Efficiently, and with Little Communication. In *AISTATS*.

[2] Wael Alghamdi, Shahab Asoodeh, Flavio P Calmon, Oliver Kosut, Lalitha Sankar, and Fei Wei. 2022. Cactus mechanisms: Optimal differential privacy mechanisms in the large-composition regime. In *2022 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 1838–1843.

[3] Apple. 2016. Differential Privacy Overview. (2016). Retrieved December 21, 2020 from https://www.apple.com/privacy/docs/Differential_Privacy_Overview.pdf

[4] Raef Bassily, Kobbi Nissim, Uri Stemmer, and Abhradeep Guha Thakurta. 2017. Practical Locally Private Heavy Hitters. In *NeurIPS*.

[5] Raef Bassily and Adam Smith. 2015. Local, Private, Efficient Protocols for Succinct Histograms. In *STOC*. 127–135.

[6] Mark Bun, Jelani Nelson, and Uri Stemmer. 2019. Heavy Hitters and the Structure of Local Privacy. *ACM Trans. Algorithms* 15, 4, Article 51 (oct 2019), 40 pages.

[7] Rui Chen, Haoran Li, A. Kai Qin, Shiva Prasad Kasiviswanathan, and Hongxia Jin. 2016. Private spatial data aggregation in the local setting. In *ICDE*. 289–300.

[8] Graham Cormode, Tejas Kulkarni, and Divesh Srivastava. 2018. Marginal Release Under Local Differential Privacy. In *SIGMOD*. 131–146.

[9] Graham Cormode, Tejas Kulkarni, and Divesh Srivastava. 2019. Answering Range Queries under Local Differential Privacy. *PVLDB* 12, 10 (jun 2019), 1126–1138.

[10] Graham Cormode, Samuel Maddock, and Carsten Maple. 2021. Frequency Estimation under Local Differential Privacy. *Proc. VLDB Endow.* 14, 11 (oct 2021), 2046–2058.

[11] Bolin Ding, Janardhan Kulkarni, and Sergey Yekhanin. 2017. Collecting Telemetry Data Privately. In *NeurIPS*. 3574–3583.

[12] Frances Ding, Moritz Hardt, John Miller, and Ludwig Schmidt. 2021. Retiring Adult: New Datasets for Fair Machine Learning. In *NeurIPS*. 6478–6490.

[13] John C. Duchi, Michael I. Jordan, and Martin J. Wainwright. 2013. Local Privacy and Minimax Bounds: Sharp Rates for Probability Estimation. In *NeurIPS*. 1529–1537.

[14] John C Duchi, Michael I Jordan, and Martin J Wainwright. 2018. Minimax optimal procedures for locally private estimation. *J. Amer. Statist. Assoc.* 113, 521 (2018), 182–201.

[15] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. 2006. Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography conference*. Springer, 265–284.

[16] Úlfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. 2014. RAPPOR: Randomized Aggregatable Privacy-Preserving Ordinal Response. In *CCS*. 1054–1067.

[17] Alexandre V. Evfimievski, Ramakrishnan Srikant, Rakesh Agrawal, and Johannes Gehrke. 2002. Privacy preserving mining of association rules. In *KDD*. 217–228.

[18] Giulia Fanti, Vasyl Pihur, and Ùlfar Erlingsson. 2015. Building a RAPPOR with the Unknown: Privacy-Preserving Learning of Associations and Data Dictionaries. *Proceedings on Privacy Enhancing Technologies* 2016 (03 2015). https://doi.org/10.1515/popets-2016-0015

[19] Marco Gaboardi, Ryan Rogers, and Or Sheffet. 2019. Locally Private Mean Estimation: Z-test and Tight Confidence Intervals. In *The 22nd International Conference on Artificial Intelligence and Statistics, AISTATS 2019, 16-18 April 2019, Naha, Okinawa, Japan*. 2545–2554.

[20] Quan Geng, Peter Kairouz, Sewoong Oh, and Pramod Viswanath. 2015. The staircase mechanism in differential privacy. *IEEE Journal of Selected Topics in Signal Processing* 9, 7 (2015), 1176–1184.

[21] Quan Geng and Pramod Viswanath. 2014. The optimal mechanism in differential privacy. In *2014 IEEE International Symposium on Information Theory*. IEEE, 2371–2375.

[22] Quan Geng and Pramod Viswanath. 2015. The optimal noise-adding mechanism in differential privacy. *IEEE Transactions on Information Theory* 62, 2 (2015), 925–951.

[23] Arpita Ghosh, Tim Roughgarden, and Mukund Sundararajan. 2009. Universally utility-maximizing privacy mechanisms. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*. 351–360.

[24] Bernard G. Greenberg, Abdel-Latif A. Abul-Ela, Walt R. Simmons, and Daniel G. Horvitz. 1969. The Unrelated Question Randomized Response Model: Theoretical Framework. *J. Amer. Statist. Assoc.* 64, 326 (1969), 520–539.

[25] Matthew Joseph, Janardhan Kulkarni, Jieming Mao, and Zhiwei Steven Wu. 2019. *Locally Private Gaussian Estimation*.

[26] Matthew Joseph, Jieming Mao, Seth Neel, and Aaron Roth. 2019. The Role of Interactivity in Local Differential Privacy. *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)* (2019), 94–105.

[27] Matthew Joseph, Jieming Mao, and Aaron Roth. 2022. Exponential Separations in Local Privacy. *ACM Trans. Algorithms* 18, 4, Article 32 (oct 2022), 17 pages.

[28] Kaggle. 2020. San Francisco Employee Compensation. (2020). Retrieved Feb 21, 2023 from https://www.kaggle.com/datasets/san-francisco/sf-employee-compensation

[29] S. P. Kasiviswanathan, H. K. Lee, K. Nissim, S. Raskhodnikova, and A. Smith. 2008. What Can We Learn Privately?. In *FOCS*. 531–540.

[30] Zitao Li, Tianhao Wang, Milan Lopuhaä-Zwakenberg, Ninghui Li, and Boris Škoric. 2020. Estimating Numerical Distributions under Local Differential Privacy. In *SIGMOD*. 621–635.

[31] Yehuda Lindell and Eran Omri. 2011. A practical application of differential privacy to personalized online advertising. *Cryptology ePrint Archive* (2011).

[32] Zhan Qin, Yin Yang, Ting Yu, Issa Khalil, Xiaokui Xiao, and Kui Ren. 2016. Heavy Hitter Estimation over Set-Valued Data with Local Differential Privacy. In *CCS*. 192–203.

[33] Steven Ruggles, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2019. Integrated Public Use Microdata Series: Version 9.0 [dataset]. (2019). Retrieved Feb 21, 2023 from http://doi.org/10.18128/D010.V9.0

[34] Adam Smith, Abhradeep Thakurta, and Jalaj Upadhyay. 2017. Is Interaction Necessary for Distributed Private Learning?. In *2017 IEEE Symposium on Security and Privacy (SP)*. 58–77. https://doi.org/10.1109/SP.2017.35

[35] Amazon Staff. 2018. Protecting data privacy. https://www.aboutamazon.com/news/amazon-ai/protecting-data-privacy. (2018).

[36] Taxi and Limousine Commission (TLC). 2022. New York Taxi Dataset. (2022). Retrieved Feb 21, 2023 from https://www.nyc.gov/site/tlc/about/tlc-trip-record-data.page

[37] Differential Privacy Team. 2018. Learning with privacy at scale. https://machinelearning.apple.com/research/learning-with-privacy-at-scale. (2018).

[38] Ning Wang, Xiaokui Xiao, Yin Yang, Jun Zhao, Siu Cheung Hui, Hyejin Shin, Junbum Shin, and Ge Yu. 2019. Collecting and analyzing multidimensional data with local differential privacy. In *2019 IEEE 35th International Conference on Data Engineering (ICDE)*. IEEE, 638–649.

[39] Tianhao Wang, Jeremiah Blocki, Ninghui Li, and Somesh Jha. 2017. Locally Differentially Private Protocols for Frequency Estimation. In *USENIX Security*. 729–745.

[40] Tianhao Wang, Jeremiah Blocki, Ninghui Li, and Somesh Jha. 2017. Locally Differentially Private Protocols for Frequency Estimation. In *USENIX Security*, Engin Kirda and Thomas Ristenpart (Eds.). 729–745.

[41] Tianhao Wang, Bolin Ding, Jingren Zhou, Cheng Hong, Zhicong Huang, Ninghui Li, and Somesh Jha. 2019. Answering Multi-Dimensional Analytical Queries under Local Differential Privacy. In *SIGMOD*. 159–176.

[42] Tianhao Wang, Ninghui Li, and Somesh Jha. 2018. Locally Differentially Private Frequent Itemset Mining. In *2018 IEEE Symposium on Security and Privacy (SP)*. 127–143. https://doi.org/10.1109/SP.2018.00035

[43] Tianhao Wang, Milan Lopuhaä-Zwakenberg, Zitao Li, Boris Skoric, and Ninghui Li. 2020. Locally Differentially Private Frequency Estimation with Consistency. In *NDSS*.

[44] Stanley L. Warner. 1965. Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias. *J. Amer. Statist. Assoc.* 60, 309 (1965), 63–69.

[45] Fei Wei, Ergute Bao, Xiaokui Xiao, Yin Yang, and Bolin Ding. 2023. AAA: an Adaptive Mechanism for Locally Differential Private Mean Estimation (Technical report). (2023). Retrieved August 1, 2023 from https://github.com/adaptiveldpmechanism/vldb_2024_submission/blob/main/technical_report.pdf

[46] N.A. Weiss, P.T. Holmes, and M. Hardy. 2006. *A Course in Probability*. Pearson Addison Wesley.

[47] Min Xu, Bolin Ding, Tianhao Wang, and Jingren Zhou. 2020. Collecting and Analyzing Data Jointly from Multiple Services under Local Differential Privacy. *PVLDB* 13, 11 (2020), 2760–2772.

[48] Min Xu, Tianhao Wang, Bolin Ding, Jingren Zhou, Cheng Hong, and Zhicong Huang. 2019. DPSAaS: Multi-Dimensional Data Sharing and Analytics as Services under Local Differential Privacy. *PVLDB* 12, 12 (2019), 1862–1865.

[49] Min Ye and Alexander Barg. 2018. Optimal Schemes for Discrete Distribution Estimation Under Locally Differential Privacy. *IEEE Trans. Inf. Theory* 64, 8 (2018), 5662–5676.

[50] Zhikun Zhang, Tianhao Wang, Ninghui Li, Shibo He, and Jiming Chen. 2018. CALM: Consistent Adaptive Local Marginal for Marginal Release under Local Differential Privacy. In *CCS*. 212–229.

A ESTIMATING THE QUANTIZED GLOBAL DATA DISTRIBUTION

As mentioned in Section 3.4, the computation of $\bar{P_X}$ can be done using any existing LDP-compliant histogram estimation algorithm. For completeness, in the following we provide two simple solutions for this purpose, using additive noise and randomized response respectively. We emphasize that these algorithms are *not* our main contribution as the problem has already been well-studied.

A.1 Using Additive Noise

Algorithm 4 describes a simple LDP-compliant data collection process using additive noise, for the purpose of data distribution estimation. Using Algorithm 4, the output $y = x + a$ can be viewed as a sample of a random variable $Y = X + A$, where X is the random variable following distribution f_X, and A is the injected noise distributed with its own probability density function f_A. The probability density function f_Y of Y is then the convolution of f_X and f_A, i.e.,

$$f_Y(y) = (f_X * f_A)(y) = \int_{y-\beta}^{y+\beta} f_X(y-\tau) \cdot f_A(\tau) d\tau.$$

The interval of the integration is $[y-\beta, y+\beta]$ as f_X is defined on $\mathcal{X} = [-\beta, \beta]$; thus, $-\beta \le y - \tau \le \beta$, and $\tau \in [y-\beta, y+\beta]$. Therefore, regarding each bin $\mathcal{X}_i$, the quantized estimate $\bar{P}_X$ is $\bar{p}_X(x_i) = \int_{\mathcal{X}_i} f_Y(y) dy$.

A.2 Using Randomized Response

In this section, we review the classic algorithm for density estimation based on randomized response [24, 44]. Recall from the notations in Section 3.3 that we are given N non-overlapping bins that partition the support of the private data $\mathcal{X} = [-\beta, \beta]$, denoted as $\mathcal{X}_1, \ldots, \mathcal{X}_N$. The mass function for P_X on these bins is defined as in Eq. (6):

$$p_X(x_i) = \int_{\mathcal{X}_i} f_X(x) dx,$$

which represents the probability that a data point is sampled from the bin $\mathcal{X}_i$ with respect to distribution f_X. The validity of the mass function as a probability measure is ensured by the fact that $\mathcal{X} = \bigcup_{i=1}^{N} \mathcal{X}_i$, as we have mentioned in Section 3.3.

We outline the algorithm that computes a private and discrete estimate $\bar{P}_X$ of f_X while satisfying ϵ-LDP, as in Algorithm 5. The high-level idea of Algorithm 5 is to let each client randomly perturb her bin index before sending it to the data aggregator, who then reconstructs a histogram based on the perturbed indices.

Here we can use an N-by-N matrix A that corresponds to the randomized response Algorithm 5 for histogram estimate. We use the h-th row of A (i.e., $A(h)$) to represent the probability distribution that perturbs bin h, and the j-th column of the h−th row (i.e., $A(h, j)$) to represent the probability that the input bin h is mapped to output bin j. In particular,

$$A(h,h) = \frac{\exp(\epsilon)}{N - 1 + \exp(\epsilon)} \text{ for } h \in [N]; \tag{26}$$

$$A(h,j) = \frac{1}{N - 1 + \exp(\epsilon)} \text{ for } j \neq h. \tag{27}$$

It is easy to verify that Algorithm 5 satisfies ϵ-LDP since $A(h, j) \le \exp(\epsilon) \cdot A(h, j)$, for any $h, h', j \in [N]$. It is also easy to see that matrix A is invertible, and we denote its inverse as A^{-1}.

Now we can think of the input private data as a histogram consisting of N bins, written as an N−dimensional vector $X \in \mathbb{N}^N$. We denote the perturbed histogram collected by the data aggregator as $\tilde{X}$, where each tuple in the bin is perturbed according to the matrix A to preserve LDP. We have that $\mathbb{E}[\tilde{X}] = AX$. To obtain an estimate of X, the data aggregator can simply compute

$$\bar{X} = A^{-1}\tilde{X}.$$

It is easy to check that $\mathbb{E}[\bar{X}] = X$ by the linearity of expectation. In practice, we may encounter negative entries in $\bar{X}$, due to the variance introduced by DP noises. If that happens, one can round the negative entries of $\bar{X}$ to 0, and then normalize $\bar{X}$ to obtain a density estimate for the distribution of P_X.

B HANDLING MULTI-DIMENSIONAL DATA

B.1 Algorithm

Algorithm 4: Data Collection under LDP with Additive Noise

Input : Client data x, LDP parameters (ϵ, δ).
Output : Perturbed data y.

1. Specify a noise $A \sim f_A$ where the noise distribution f_A is determined by choosing a designated noise type (e.g. Gaussian, Laplace), and the associated noise parameters according to the LDP parameters (ϵ, δ).
2. Evaluate A to obtain a noise sample a.
3. Return $y = x + a$.

Algorithm 5: Density Estimation under LDP with Randomized Response

Input : Collection of client data $\mathcal{D} = \{x_i : i \in [n]\}$, where $\mathbf{x}_i \in [-\beta, \beta]$, LDP parameter ϵ, N non-overlapping bins $\mathcal{X}_1, \ldots, \mathcal{X}_N$.
Output : Density estimation $\bar{P}_X$

1. **for** *each client* i **do**
2. Client i quantizes her data x_i to one of the N bins, say b_i.
3. Client i randomly sample $u_i \sim [0, 1]$.
4. **if** $u_i \ge [0, \frac{\exp(\epsilon)-1}{N-1+\exp(\epsilon)}]$ **then**
5. Assign b_i to a uniform sample from $[N]$.
6. Client i sends b_i to the data aggregator.
7. Based on the collected responses $\{b_i\}_i$, the data aggregator constructs a histogram of N bins, denoted as $\tilde{X}$.
8. The data aggregator computes $\bar{X} = A^{-1}\tilde{X}$, where A^{-1} is the inverse of matrix A defined as in Eq. (26) and 27.
9. The data aggregator normalizes $\bar{X}$ to obtain a density estimation $\bar{P}_X$.

Regarding the task of collecting multidimensional data for mean estimation under LDP, we apply the methodology mentioned in

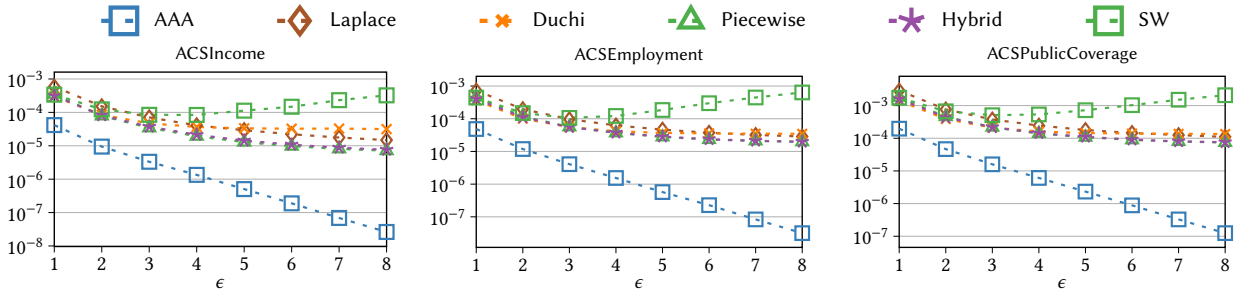


Figure 7: Performance on multidimensional data under varying privacy parameters.

Algorithm 6: Overview of LDP-compliant multi-dimensional data collection

Input : Collection of client data $\mathcal{D} = \{\mathbf{x}_i : i \in [n]\}$, where $\mathbf{x}_i = \left(x_i^{(1)}, \ldots, x_i^{(d)}\right) \in [-\beta, \beta]^d$, LDP parameters (ϵ, δ), client portion parameter s, attribute portion parameter k.

Output : Mean estimation $\bar{\mathbf{X}} = \left(\bar{X}^{(1)}, \ldots, \bar{X}^{(d)}\right)$

for *Each client* i **do**
  Client i uniformly samples k dimensions from $[d]$ without replacement, obtaining a set of indices denoted as $\mathcal{K}_i$.
for *Each attribute* $X^{(j)}$ **do**
  $n^{(j)} = \sum_{i=1}^{n} \mathbb{1}(j \in \mathcal{K}_i)$ represents the number of clients that have selected dimension j, where the indicator function $\mathbb{1}(j \in \mathcal{K}_i)$ evaluates to 1 if client i has selected dimension j .
  The data aggregator randomly picks a s portion of the $n^{(j)}$ clients that have selected dimension j.
  The sampled clients use classic mechanisms satisfying (ϵ, δ)-local DP to perturb the private attribute $x^{(j)}$.
  The clients return the perturbed attribute $\bar{x}^{(j)}$ to the data aggregator.
  The data aggregator obtains a distribution estimation $\hat{P}_j$ for the attribute.
The data aggregator obtains AAA mechanisms satisfying (ϵ, δ)-local DP for each attribute with respect to estimated distribution $\hat{P}_j$.
Broadcast the optimized AAA mechanisms to all clients.
for *Each attribute* $X^{(j)}$*, the rest* $1 - s$ *portion of clients that have selected dimension* j **do**
  The clients use broadcasted AAA mechanisms to perturb the private attribute $x^{(j)}$.
  The clients return the perturbed attribute $\bar{x}^{(j)}$ to the data aggregator.
  The data aggregator obtains $\bar{X}^{(j)} = \sum \frac{1}{(1-s)n^{(j)}} \cdot \bar{x}_i^{(j)}$ as a mean estimation of the attribute.

Section 2.1 that each client proceeds with a random collection of data entries (attributes) and returns the perturbed version. More specifically, the perturbation is executed in an attribute-wise manner on the individual level, i.e., each client applies the broadcast perturbation mechanism, which is optimized with respect to the estimated distributions, on the assigned attributes. We outline the procedure for applying AAA mechanism to multidimensional data in Algorithm 6. The implementations of comparison works are conducted with similar procedures, and are omitted for brevity.

B.2 Experiments

We compare the performance of our solution with the same set of competitors in Section 7 under three real-world datasets, ACSIncome, ACSEmployment, and ACSPublicCoverage [12]. All three datasets are derived from US Census data in year 2018. The number of samples and dimensions (n, d) for ACSIncome, ACSEmployment, and ACSPublicCoverage are $(1664500, 11)$, $(3236107, 17)$, and $(1138289, 20)$, respectively. Similar to what we did in Section 7, we normalize each dimension of the datasets to $[-\beta, \beta]$ with $\beta = 1$.

For this set of experiments, each individual client randomly selects k out of d attributes, and privatizes her data on the selected attributes using the perturbation mechanisms, before sending them to the data aggregator, as explained in Section B.1. Without loss of generality, here we set $k = 1$, which does not affect the relative performance of the algorithms. The error is computed as the sum of mean squared errors over all d dimensions. We report the average error over $1,000$ independent runs in Figure 7.

Once again, the AAA mechanism outperforms its competitors by a large margin, similar to what we observe in the one-dimensional experiments of Section 7. This suggests that the AAA mechanism can be used as a plug-in replacement for any LDP mechanism for mean estimation, and its improvement is independent of the dimensionality of the data. We do not consider more advanced dimensionality reduction techniques such as [1, 5], as they are orthogonal to the underlying perturbation mechanism, and the argument that AAA outperforms other existing solutions would still hold when applying these techniques. Similarly, we also omit further experiments on hyperparameters and additional datasets, since the argument still holds as we have shown for one-dimensional data in Section 7.

C ADDITIONAL EXPERIMENTS

C.1 The Effect of Split Ratio

We also conduct additional experiments investigating the impact of split ratio s. We vary s from 0.001 to 0.4, using a synthetic dataset ($n = 10000$) following Gaussian distribution. From the results in Figure 8, we observe that the error of the proposed AAA mechanism

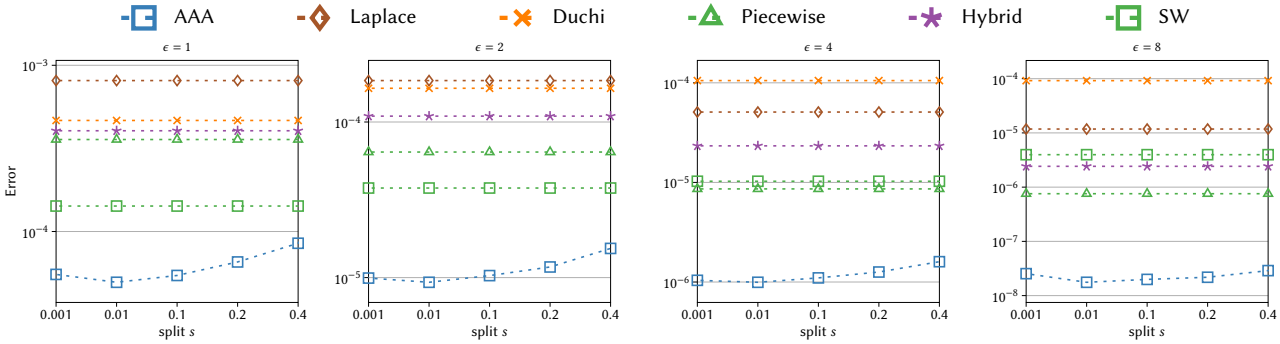


Figure 8: Performance of AAA on $10,000$ Gaussian data points with split ratio s varying from 0.001 to 0.4.

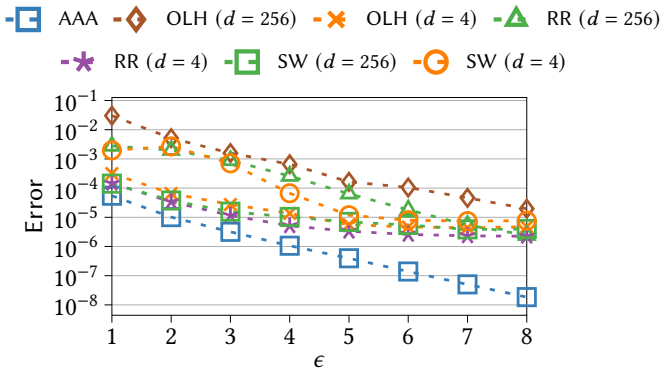


Figure 9: Evaluation of AAA on $10,000$ Gaussian data points against LDP mechanisms designed for frequency estimation.

is consistently and significantly smaller than all existing solutions, for a wide range of values for s, which demonstrates the robustness of AAA with respect to s. Experiments on other datasets lead to similar conclusions and are omitted.

C.2 Other Baselines in LDP

As we have mentioned in Section 6, the LDP mechanisms that are originally designed for tasks other than mean estimation, such as the generalized randomized response (generalized RR), OLH [43], and SW [30], do not perform well under our setting. Nevertheless, we compare the AAA mechanism against the aforementioned LDP protocols on the task population mean. We consider the mean squared error for the population mean as the utility metric, as we did in the previous experiments. From Figure 9, it is clear that AAA significantly outperforms the frequency estimation baselines. Here the parameter d stands for the domain size (resp., number of bins) in frequency estimation (resp., histogram estimation) for generalized RR, OLH [43], and SW [30]. We report the results on $d = 4$ and $d = 256$. Other values of d lead to similar results and hence are omitted from the figure. Results of OUE [40], Hadamard Transform [5], and Subset Selection [49] are also omitted since they are identical to OLH for the resulting error, as pointed out in [43]. We refer interested readers to [43] for a detailed comparison of these LDP mechanisms. We conclude that the AAA mechanism is the state-of-the-art LDP mechanism for the task of mean estimation.