Optimal Piecewise-based Mechanism for Collecting Bounded Numerical Data under Local Differential Privacy

Ye Zheng
Rochester Institute of Technology
ye.zheng@mail.rit.edu

Sumita Mishra
Rochester Institute of Technology
sumita.mishra@rit.edu

Yidan Hu
Rochester Institute of Technology
yidan.hu@rit.edu

Abstract

Numerical data with bounded domains is a common data type in personal devices, such as wearable sensors. While the collection of such data is essential for third-party platforms, it raises significant privacy concerns. Local differential privacy (LDP) has been shown as a framework providing provable individual privacy, even when the third-party platform is untrusted. For numerical data with bounded domains, existing state-of-the-art LDP mechanisms are piecewise-based mechanisms, which are not optimal, leading to reduced data utility.

This paper investigates the optimal design of piecewise-based mechanisms to maximize data utility under LDP. We demonstrate that existing piecewise-based mechanisms are heuristic instances of the 3-piecewise mechanism, which is far from enough to study optimality. We generalize the 3-piecewise mechanism to its most general form, i.e. m-piecewise mechanism with no pre-defined form of each piece. Under this form, we derive the closed-form optimal mechanism by combining analytical proofs and off-the-shelf optimization solvers. Next, we extend the generalized piecewise-based mechanism to the circular domain (along with the classical domain), defined on a cyclic range where the distance between the two endpoints is zero. By incorporating this property, we design the optimal mechanism for the circular domain, achieving significantly improved data utility compared with existing mechanisms.

Our proposed mechanisms guarantee optimal data utility under LDP among all generalized piecewise-based mechanisms. We show that they also achieve optimal data utility in two common applications of LDP: distribution estimation and mean estimation. Theoretical analyses and experimental evaluations prove and validate the data utility advantages of our proposed mechanisms.

Keywords

local differential privacy, numerical data privacy, bounded domain, circular data

1 Introduction

Numerical data with bounded domains is a fundamental data type in personal devices. These bounded domains can be categorized into two types: linear ranges, such as sensor readings in $[0, 1)$, referred to as *classical domain*; and cyclic ranges, such as angular measurements in $[0, 2\pi)$, referred to as *circular domain*. These types of data are crucial for third-party platforms to provide personalized services. However, collecting such data involves particular privacy concerns, as third-party collectors are potentially untrusted and the data often contain sensitive information. Simple anonymization techniques [29, 40] have been proven insufficient to prevent privacy leakage [27, 38, 41]. Therefore, a provable privacy guarantee is necessary when collecting these sensitive data.

Local differential privacy (LDP) serves as a de facto standard, providing an input-independent formal guarantee regarding the difficulty of inferring sensitive data. Through the LDP mechanism, sensitive data are randomly perturbed before being sent to an untrusted collector. The randomization ensures a sufficient level of indistinguishability (indicated by the privacy parameter ε). Consequently, any observation over the randomized data is essentially powerless to infer the sensitive data. Laplace Mechanism [10] is a classical LDP mechanism for numerical data privacy. It adds random noise, drawn from a Laplace distribution determined by ε, to the sensitive data. However, the unbounded noise of the Laplace mechanism makes it unsuitable for bounded domains.

State-of-the-art LDP mechanisms for numerical data with bounded domains are *piecewise-based mechanisms* [21, 23, 34]. They are widely used as building blocks to provide provable privacy guarantees in various scenarios, such as in sensors networks and federated learning. Piecewise-based mechanisms randomize the sensitive data to a value sampled from a carefully designed piecewise probability distribution. Existing instantiations use different pieces and probabilities, but all are designed for classical domains. Their applicability to other types of bounded domains, e.g. circular domain of angular sensors that commonly appear in personal devices, is unexplored.

Data utility is the most crucial metric for LDP mechanisms, typically measured by the distance between the randomized data and the sensitive data. While the privacy level is theoretically guaranteed by the privacy parameter ε, a mechanism with better data utility allows for more accurate analysis. The data utility of a piecewise-based mechanism is determined by the distance metric, pieces, and their respective probabilities. Unfortunately, none of the existing instantiations of piecewise-based mechanisms are shown to be optimal in terms of data utility, indicating potential for improving analysis accuracy without compromising privacy. These situations highlight the need for optimal piecewise-based mechanisms.

The optimality of piecewise-based mechanisms remains a challenging problem. We will see that the existing instantiations are heuristic forms of the 3-piecewise mechanism (TPM). As a special case with 3 pieces and pre-defined forms of pieces, TPM is far from enough to study the optimality of piecewise-based mechanisms. From the evidence of the staircase Laplace mechanism [15] for unbounded numerical data, the asymptotically optimal mechanism has a staircase (multiple pieces) form. For categorical data, a staircase Randomized Response mechanism (SRR) [33] improves data

Proceedings on Privacy Enhancing Technologies YYYY(X), 1–20

https://doi.org/XXXXXXX.XXXXXXX

utility in location collection compared to the general RR mechanism. Numerous indicators suggest that increasing the variety of probabilities in the data domain, i.e. using more pieces, can improve data utility. In light of these examples, a fundamental question for piecewise-based mechanisms is: *what is the optimal instantiation of piecewise-based mechanism?* In the design of piecewise-based mechanisms, the number of pieces, their probabilities and sizes can be arbitrary. Finding the optimal instantiation within such a large design space is challenging, as it requires optimizing the number of pieces, their probabilities and sizes simultaneously.

This paper studies the optimality of piecewise-based mechanism in its most general form.* We extend TPM into a generalized piecewise-based mechanism (GPM) that is m-piece, with each piece having no pre-defined form. Under GPM, we formulate an optimization problem to minimize the distance between the sensitive and randomized data. By combining the solutions of the optimization problem with analytical proofs, we derive the closed-form optimal GPM for classical domains. For circular domains, where distance metrics have a unique property (e.g. the distance between 0 and 2π is zero), we incorporate this property into mechanism design and link the solving of the optimal mechanism to problems in the classical domain. Table 1 summarizes the key features of this paper in comparison with existing works. Particularly, our contributions are as follows:

- *(Solving framework)* To the best of our knowledge, this is the first work to study the closed-form optimal piecewise-based mechanism under its most general form. We propose a framework that integrates analytical proofs with off-the-shelf optimization solvers to derive the closed-form optimal mechanism. This approach establishes a feasible foundation for achieving optimal data utility under LDP for numerical data with bounded domains.
- *(Closed-form instantiations)* We provide closed-form optimal mechanisms for the classical domain and the circular domain. As alternatives to existing mechanisms, they can be directly used as building blocks in sensor networks and federated learning, etc, while guaranteeing optimal data utility among all piecewise-based mechanisms under LDP.
- *(Theoretical and experimental evaluations)* We provide theoretical analyses of data utility and experimental evaluations on two common applications of LDP: distribution and mean estimation. The results prove and validate our mechanisms' advantages over existing mechanisms. The codes are available at https://github.com/ZhengYeah/Optimal-GPM.

Structure. The main part of this paper is organized as follows: After the preliminaries, we present the optimal piecewise-based mechanism for the classical domain in Section 3. Section 4 focuses on the circular domain and derives its optimal mechanism. Following this, Section 5 discusses the optimality when applying to two common tasks: distribution and mean estimation.

2 Preliminaries

This section formulates the problem and the concept of local differential privacy (LDP). We present existing instantiations of TPM and their limitations, which motivate our proposed optimal generalized piecewise-based mechanism (OGPM).

*This means that we consider all possible forms of piecewise distributions on a bounded domain, ensuring the most comprehensive generalization.

Table 1: OGPM vs existing instantiations of TPM.

	Domain	Optimality	Closed form	Estimation
PM [34]	Classical	No	Yes	Mean
SW [21]		No	Yes	Distribution
PTT [23]		Partly*	No	Mean
This paper (OGPM)	Classical & circular	Yes	Yes	Mean & distribution

* Proved the existence of the optimal under TPM, but did not give closed-form instantiations. Appendix B.1 provides detailed discussion.

2.1 Problem Formulation

We consider a typical data collection schema that consists of a set of *users* and one *collector*. Each user has a numerical sensitive data $x_i \in \mathcal{D}$, where $\mathcal{D}$ is a continuous and bounded domain. The collector needs to collect data from users for statistical estimations, such as the mean value and distribution of the data.

However, the collector is untrusted and may attempt to infer users' sensitive data. To protect privacy, each user locally randomizes their sensitive data using a privacy mechanism $\mathcal{M} : \mathcal{D} \to \tilde{\mathcal{D}}$, then sends $y_i = \mathcal{M}(x_i)$ to the collector.

We seek to design an optimal $\mathcal{M}$ that maximizes the data utility by minimizing the distance between the sensitive data x_i and the reported data y_i, while ensuring ε-LDP (Definition 2.1).

2.2 Local Differential Privacy

Definition 2.1 (ε-LDP [9]). A randomization mechanism $\mathcal{M} : \mathcal{D} \to \tilde{\mathcal{D}}$ satisfies ε-LDP, if for two arbitrary inputs x_1 and x_2, the probability ratio of outputting the same y is bounded:

$$\forall x_1, x_2 \in \mathcal{D}, \forall y \in \tilde{\mathcal{D}} : \frac{\Pr[\mathcal{M}(x_1) = y]}{\Pr[\mathcal{M}(x_2) = y]} \leq \exp(\varepsilon).$$

If $\mathcal{M}(x)$ is continuous, the probability $\Pr[\cdot]$ is replaced by probability density function (pdf). Intuitively, ε-LDP represents the difficulty of distinguishing between x_1 and x_2 given y. A lower privacy parameter $\varepsilon \in [0, +\infty)$ means higher privacy. For example, $\varepsilon = 0$ requires that $\mathcal{M}$ maps two arbitrary inputs to any output y with the same probability, thus y contains no distribution information about x, making any hypothesis-testing method to infer the sensitive x powerless even with known $\mathcal{M}$.

2.3 Piecewise-based Mechanisms

When the input domain $\mathcal{D}$ is both continuous and bounded, the state-of-the-art mechanisms to achieve ε-LDP are piecewise-based mechanisms. We summarize these mechanisms as heuristic instances of the following definition.

Definition 2.2. 3-piecewise mechanism (TPM) $\mathcal{M} : \mathcal{D} \to \tilde{\mathcal{D}}_\varepsilon$ is a family of probability density functions that, given input $x \in \mathcal{D}$, outputs $y \in \tilde{\mathcal{D}}$ according to

$$pdf[\mathcal{M}(x) = y] = \begin{cases} p_\varepsilon & \text{if } y \in [l_{x,\varepsilon}, r_{x,\varepsilon}], \\ p_\varepsilon / \exp(\varepsilon) & \text{if } y \in \tilde{\mathcal{D}}_\varepsilon \setminus [l_{x,\varepsilon}, r_{x,\varepsilon}], \end{cases}$$

where p_ε is a variable determined solely by ε,[†] while $l_{x,\varepsilon}$ and $r_{x,\varepsilon}$ depend on x and ε. The output domain $\tilde{\mathcal{D}}_\varepsilon \supset \mathcal{D}$ is an enlarged domain depending on ε.

TPM samples the output y for each x from a piecewise distribution. This sampling is with a higher probability p_ε within $[l_{x,\varepsilon}, r_{x,\varepsilon}]$ and a lower probability $p_\varepsilon/\exp(\varepsilon)$ within the remaining two pieces $\tilde{\mathcal{D}} \setminus [l_{x,\varepsilon}, r_{x,\varepsilon}]$, satisfying the ε-LDP constraint.

Instantiations. In TPM, the parameters are the central interval $[l_{x,\varepsilon}, r_{x,\varepsilon}]$, its probability p_ε, and the output domain $\tilde{\mathcal{D}}_\varepsilon$. Different instantiations of those parameters yield different existing mechanisms [21, 23, 34]. For example, PM [34] is the first instantiation of TPM, defined on $[-1, 1] \to [-C_\varepsilon, C_\varepsilon]$, where C_ε is a variable determined solely by ε and the central interval has a fixed length $r_{x,\varepsilon} - l_{x,\varepsilon} = C_\varepsilon - 1$.

Data utility metric. To quantify the data utility of different instantiations, we consider the general L_p-similar error metric (i.e. $|y - x|^p$) as a loss function $\mathcal{L} : \mathbb{R} \to \mathbb{R}$. Thus, the error is:

$$Err(x) = \int_{\tilde{\mathcal{D}}} \mathcal{L}(y, x) \mathcal{P}_{\mathcal{M}(x)} \mathrm{d}y, \tag{1}$$

where $\mathcal{P}_{\mathcal{M}(x)}$ is the pdf defined by $\mathcal{M}(x)$. $Err(x)$ illustrates the expected error when applying $\mathcal{M}$ on x under the loss function $\mathcal{L}$. For example, $\mathcal{L}(y, x) \coloneqq |y - x|$ is the absolute error, and $\mathcal{L}(y, x) \coloneqq (y - x)^2$ is the square error. Then $Err(x)$ corresponds to the mean absolute error (MAE) and mean square error (MSE) [23, 34], respectively. Lower $Err(x)$ indicates better data utility.

Limitations of TPM. Existing instantiations of TPM have the following limitations.

- *Not optimal in data utility.* None of the existing instantiations provided closed forms for the optimal data utility. Meanwhile, they also assume an invariable length $r_{x,\varepsilon} - l_{x,\varepsilon}$ of the central piece for all x, and symmetric probability $p_\varepsilon/\exp(\varepsilon)$ for the remaining two pieces. However, a general-form piecewise-based mechanism can have more pieces, unfixed piece lengths, and asymmetric probabilities, potentially improving data utility.
- *Limited applicability.* Existing instantiations of TPM have enlarged and unfixed output domains $\tilde{\mathcal{D}}_\varepsilon \supset \mathcal{D}$. Enlarged output domain incurs applicability issues in scenarios where the collector requires the output domain to align with the input domain (i.e. $\tilde{\mathcal{D}} = \mathcal{D}$),[‡] such as in common sensor-based services.

3 Generalized Piecewise-based Mechanism

This section generalizes TPM to its most general form (GPM). We introduce a framework for deriving the closed-form optimal GPM for the classical domain.

Definition 3.1. Generalized m-piecewise mechanism (m-GPM) $\mathcal{M} : \mathcal{D} \to \tilde{\mathcal{D}}$ is a family of probability density functions that, given input $x \in \mathcal{D}$, outputs $y \in \tilde{\mathcal{D}}$ according to

$$pdf[\mathcal{M}(x) = y] = \begin{cases} p_{1,\varepsilon} & \text{if } y \in [l_{1,x,\varepsilon}, r_{1,x,\varepsilon}), \\ \vdots & \vdots \quad \vdots \\ p_{m,\varepsilon} & \text{if } y \in [l_{m,x,\varepsilon}, r_{m,x,\varepsilon}), \end{cases}$$

$$\forall i, j \in [m], \max \frac{p_{i,\varepsilon}}{p_{j,\varepsilon}} \le \exp(\varepsilon),$$

where $[m] \coloneqq \{1, ..., m\}$. Each probability $p_{i,\varepsilon}$ depends solely on privacy parameter ε, while interval boundaries $l_{i,x,\varepsilon}$ and $r_{i,x,\varepsilon}$ depend on both x and ε.

An m-GPM partitions its output domain into m pieces, assigning probability $p_{i,\varepsilon}$ to each piece $[l_{i,x,\varepsilon}, r_{i,x,\varepsilon})$. The probabilities $p_{i,\varepsilon}$ are independent of the input x, and their ratios must be bounded by $\exp(\varepsilon)$ to satisfy ε-LDP. For notational clarity, we omit subscripts x and ε when their context is clear. Additionally, $\mathcal{M}$ must satisfy standard probability requirements: non-negativity ($p_i \ge 0$), continuity ($r_i = l_{i+1}$), and normalization. TPM is a special case of GPM with $m = 3$.

Finding the optimal GPM requires determining both the optimal number of pieces m and the corresponding $p_{i,\varepsilon}, l_{i,x,\varepsilon}, r_{i,x,\varepsilon}$. Due to the infinite possibilities for $m \in \mathbb{N}^+$ and the resulting $3m$ variables, analytical solutions are computationally intractable. We therefore propose a framework that combines analytical proofs with off-the-shelf optimization solvers.

3.1 Framework for Deriving the Optimal GPM

To derive the *closed-form* optimal GPM, we (i) formulate finding the optimal m-GPM as an optimization problem; (ii) determine the optimal m based on the solutions of the optimization problem; (iii) derive the optimal closed-form expression (among all m-GPM).

Optimal m-GPM. To find the optimal GPM instantiation with m pieces, we need to determine the variables p_i, l_i, and r_i. Any feasible assignment of these variables yields a mechanism $\mathcal{M}$ whose utility can be measured by $Err(x)$ from Formula (1). Finding the optimal m-GPM requires solving a min-max optimization problem that minimizes the worst-case error over all possible inputs x:[§]

$$\min_{p_i, l_{i,x}, r_{i,x}} \max_x \int_{\tilde{\mathcal{D}}} \mathcal{L}(y, x) \mathcal{P}_{\mathcal{M}(x)} \mathrm{d}y, \quad \text{s.t. } \mathcal{M} \text{ satisfies Definition 3.1.} \tag{2}$$

This formulation yields the optimal x-independent p_i values and the corresponding $l_{i,x}, r_{i,x}$ for the worst-case input x. However, since these $l_{i,x}, r_{i,x}$ may not be optimal for other inputs, we need a second optimization step using the obtained optimal p_i:

$$\min_{l_{i,x}, r_{i,x}} \int_{\tilde{\mathcal{D}}} \mathcal{L}(y, x) \mathcal{P}_{\mathcal{M}(x)} \mathrm{d}y, \quad \text{s.t. } \mathcal{M} \text{ satisfies Definition 3.1 with } p_i. \tag{3}$$

Together, these two steps determine the optimal instantiation of m-GPM for any given domain mapping $\mathcal{D} \to \tilde{\mathcal{D}}$, distance metric $\mathcal{L}$, piece number m, privacy parameter ε, and input x. Figure 1 illustrates this solving process.

[†] Otherwise, if p_ε varies with x, it violates the ε-LDP constraint because the probability ratio outputting the same y from x_1 and x_2 is not bounded by $\exp(\varepsilon)$.

[‡] While post-processing the output by truncating it to $\mathcal{D}$ is possible, this approach may still result in low data utility. Sections 7.1.2 and 7.1.3 provide comparisons with mechanisms that include truncation.

[§] Worst-case error is the most common utility metric in mechanism design [23, 34]. We can also optimize the error at other specific points, see Section 6.

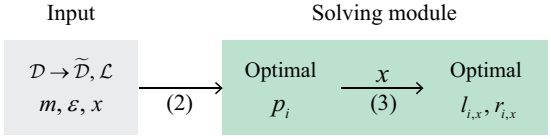


Figure 1: Solving flow for the optimal m-GPM. Two arrows indicate problems in (2) and (3).

Challenges. Nonetheless, solving Formulation (2) has practical difficulties. Even if it can be solved, there is still a gap between the solved optimal m-GPM and the closed-form optimal GPM (among all m-GPM). We detail them as follows.

- *(Solving difficulty)* Formulation (2) is a min-max problem, and the integrand $\mathcal{L}(y, x)\mathcal{P}_{\mathcal{M}(x)}$ is non-linear. It is a non-convex problem whose global optimal hard to solve.
- *(Optimal m)* The solved optimal is only for m-GPM given m. It is necessary to find the optimal piece number m.
- *(Closed form)* For practical usage, we need closed-form p_i, $l_{i,x}$ and $r_{i,x}$ (i.e. their relationships with ε and x), rather than their specific values for every ε and x.

Solutions. We address the challenges and gaps with the following solutions.

- Formulation (2) can be simplified to two *bilinear optimization* problems that can be solved by off-the-shelf solvers. (Section 3.2.1)
- If the optimal $(m+1)$-GPM is identical to the m-GPM, then m is the optimal piece number. (Section 3.2.2)
- Leveraging the above results, the optimal closed-form p_i, $l_{i,x}$ and $r_{i,x}$ can be obtained by analytical deduction (for TPM) or numerical regression (for any m-GPM). (Section 3.2.3)

Following these solutions, we can obtain the closed-form optimal GPM (among all m) for any ε given $\mathcal{D} \to \tilde{\mathcal{D}}$ and $\mathcal{L}$. Before presenting the detailed solutions, we first discuss the conditions under which the obtained GPM is optimal.

Conditions for optimality. When discussing optimality, the following aspects should be specified: (i) the error metric, (ii) the data domain and family of mechanisms, (iii) the strength of the optimality, and (iv) whether post-processing is allowed. In this paper, the optimality of GPM is defined with respect to: (i) the worst-case L_p-similar error metric, (ii) bounded numerical domains $\mathcal{D} \to \tilde{\mathcal{D}}$ and mechanisms based on piecewise distributions, (iii) minimization of error value (not asymptotic or order-of-magnitude optimality), and (iv) without post-processing. These conditions are widely applicable in practice and literature. However, varying any of them may lead to different optimality results. Appendix B.2 provides a detailed discussion of these conditions and related optimalities.

3.2 Detailed Solutions

3.2.1 Solution 1: Simplified Form. The min-max problem in Formulation (2) can be simplified. The key observation is that its inner maximization term, $\max_x$, has a closed form, i.e. the worst-case error is from the endpoints of $\mathcal{D}$. Lemma 3.2 states this observation.

Lemma 3.2. *Assume $\mathcal{D} = [a, b)$, the objective of Formulation (2) can be simplified to*

$$\min_{p_i, l_{i,x}, r_{i,x}} \max_{x \in \{a,b\}} \int_{\tilde{\mathcal{D}}} \mathcal{L}(y, x)\mathcal{P}_{\mathcal{M}(x)} \mathrm{d}y.$$

Proof. (Sketch) The key of the proof is that each integral on $\tilde{\mathcal{D}}$ is convex function w.r.t x. Thus, their non-negative weighted sum is also convex. According to the Bauer maximum principle [37], the maximum is achieved at the endpoints of $\mathcal{D}$, i.e. $x = a$ or b.[¶] Appendix A.1 provides the full proof. □

Complexity. (i) Lemma 3.2 simplifies Formulation (2) to two *bilinear optimization* problems, i.e. when $x = a$ and $x = b$ respectively. The integrand $\mathcal{L}(y, a)\mathcal{P}_{\mathcal{M}(a)}$ includes terms such as $p_i l_i$ and $p_i r_i$, which involve multiplications of two variables. This problem can be solved by off-the-shelf solvers such as Gurobi [2], which employs stochastic *Branch and Bound* [17] method to handle the bilinear terms. (ii) The number of variables is at most $3m$, which can be efficiently solved for small m. For example, we can obtain the exact optimal for $m \leq 7$ and $\mathcal{L} = |y - x|$ within 2 seconds. Furthermore, solvers generally provide over- and under-approximation for bilinear problems [28]. We can obtain solutions with $\leq 1\%$ gap from the optimal for $m \leq 19$ within 1 minute. (iii) Formulation (3) for solving l_i and r_i has at most $2m$ variables and without p_i terms, it is a toy-size problem when m is small.

Encoding details. We need to encode the problem in Lemma 3.2 to simple mathematical expressions that can be handled by the solver. If we instantiate $\mathcal{L} = |y - x|$ and focus on the left endpoint $x = a$, this problem becomes:

$$\begin{aligned} &\min_{p_i, l_i, r_i} \sum_{i=1}^{m} p_i \int_{l_i}^{r_i} (y - a) \mathrm{d}y \\ = &\min_{p_i, l_i, r_i} \sum_{i=1}^{m} \frac{p_i}{2} \left((r_i - a)^2 - (l_i - a)^2 \right). \end{aligned}$$

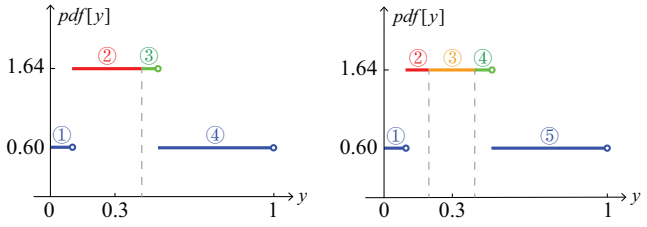
This problem is a bilinear optimization problem. The highest-degree term is $p_i \cdot r_i \cdot r_i$, which can be reformulated as $p_i \cdot t$ with $t = r_i \cdot r_i$, i.e. multiplication of bilinear terms. Such problems can be solved by off-the-shelf bilinear solvers [2]. The other problem in Formulation (3) can be encoded similarly but is much easier due to constant p_i.

3.2.2 Solution 2: Optimal Piece Number. Although we can obtain the optimal m-GPM for any m given sufficient time, solving for each m is unnecessary. The following lemma provides a theoretical basis for capping the optimal number of pieces.

Lemma 3.3. *For all possible ε and x, if the optimal $(m+1)$-GPM is the same as the m-GPM, then the optimal piece number is m.*

Proof. (Sketch) The key insight is that, if the optimal piece number is not m, i.e. an additional piece can lower the error, then this additional piece will be captured by the optimal $(m+1)$-GPM. Therefore, (i) if m is not the optimal piece number, then the optimal $(m+1)$-GPM is different from the optimal m-GPM. (ii) if m is the optimal piece number, then there is no additional piece can lower the error, making the optimal $(m+1)$-GPM is the same as the optimal

[¶] We use $[a, b)$ to denote the domain for consistency with the form of pieces in GPM. This is the same as $[a, b]$ in implementation.



(a) $m = 4$. **(b)** $m = 5$.

Figure 2: Optimal 4-GPM and 5-GPM when $\varepsilon = 1$ and $x = 0.3$. They are identical as $m = 3$ after merging redundant pieces.

m-GPM. Now, no additional piece can be captured, hence m is the optimal piece number. Appendix A.2 provides the full proof. □

Lemma 3.3 requires checking the results for every ε and x, which is challenging as $\varepsilon \in [0, \infty)$ and $x \in \mathcal{D}$ are infinite sets. In practice, we restrict ε within its generally meaningful domain, e.g. $\varepsilon \in [0, 10)$, and employ *Monte Carlo random sampling* to generate a set of random pairs $(\mathcal{E}, \mathcal{X}) = \{(\varepsilon_1, x_1), (\varepsilon_2, x_2), \ldots, (\varepsilon_n, x_n)\}$. If Lemma 3.3 holds for the n-size set $(\mathcal{E}, \mathcal{X})$, the optimality of m-GPM is guaranteed with probability 1 as $n \to \infty$ [31].

Results for Monte Carlo sampling. Fixing $\mathcal{D} = \tilde{\mathcal{D}} = [0, 1)$, for $\mathcal{L} = |y - x|$ and $\mathcal{L} = (y - x)^2$, the optimal 4-GPM are identical as the optimal 3-GPM for random $(\mathcal{E}, \mathcal{X})$ with at least $n = 10^4$. These optimal results align with TPM, but the optimal p, l and r values are different from existing instantiations.

Given the continuity of the objective functions w.r.t ε and x, we posit that it is, in fact, the exact optimal. Therefore, when we say the optimal piece number is $m = 3$, it implies the *statistical optimality* guaranteed by the Monte Carlo asymptotic technique with a strength of at least $n = 10^4$ random samples.

Example 3.4. For $[0, 1) \to [0, 1)$ and $\mathcal{L} = |y - x|$, Figure 2 shows two examples of the optimal 4-GPM and optimal 5-GPM when $\varepsilon = 1$ and $x = 0.3$. After merging redundant pieces, i.e. connected pieces with the same probability, they are the same as the 3-GPM and fall into the TPM category.

3.2.3 Solution 3: Closed-form Instantiation. After determining the optimal m, we can derive the closed-form instantiation. If the optimal results exhibit $m = 3$ and coincide with TPM, it facilitates analytical deduction for the closed-form optimal p, l and r. Otherwise, numerical regression can be employed to obtain the closed-form instantiation. Figure 3 illustrates the workflow.

Analytical deduction. For TPM, the optimization procedure for solving p_i, l_i and r_i can be conducted analytically. The key observation is that there are only three variables: p, l, and r in TPM. Due to the normalization constraint of probability, the central interval length $r - l$ can be replaced by its probability p. This reduces the solving for the optimal p to a univariate optimization problem w.r.t. p, which can be solved by analyzing the first-order derivative. With the solved p, solving l and r also becomes a univariate optimization problem. Appendix B.4 provides the formalized process.

Numerical regression. For any m-GPM, we can obtain the closed-form p_i, l_i, and r_i through numerical regression on their solved optimal values.

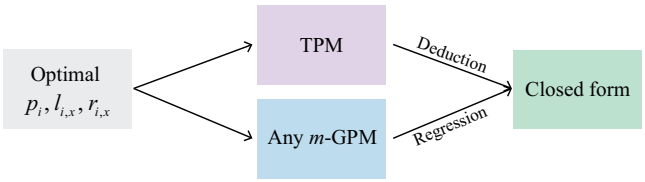


Figure 3: Solving flow for the optimal closed form.

Assume that we want to find the closed-form optimal p_i. Given $(\mathcal{E}, \mathcal{P}_i) = \{(\varepsilon_1, p_{i,1}), \ldots, (\varepsilon_n, p_{i,n})\}$, which contains the solved optimal p_i for random ε, we aim to find the relationship between ε and p_i. This relationship can be approximated by $\hat{p}_i = f(\varepsilon, \beta)$, where f is a designed feature with β as regression parameters.

Ideally, the designed feature f matches the truth form of p_i. In this case, the regression result of f on $(\mathcal{E}, \mathcal{P}_i)$ converges to the optimal closed-form p_i. If not, the regression result may not converge to the optimal. In practice, we can use heuristic forms of f for tighter approximation. Due to the structure of the LDP constraint, we suggest choosing p_i with a form of $\exp(\beta_1 \varepsilon)$, allowing us to design $f = \exp(\beta_1 \varepsilon) + \beta_2$.

Example 3.5. For $\mathcal{L} = |y - x|$, $\mathcal{D} = \tilde{\mathcal{D}} = [0, 1)$ and $m = 3$, assume the probability p has the form $\hat{p} = \exp(\beta_1 \varepsilon) + \beta_2$. We use the `scipy.curve_fit` package to regress p on 50 random ε; then its regression result is $\hat{p} = \exp(\varepsilon/2) - 0.07$ with a maximal error $\leq 10^{-2}$. Note that this result almost coincides with the ground-truth p in Theorem 3.7.

3.3 Closed Form for Classical Domain

The above framework for deriving the closed-form optimal GPM is applicable to any $\mathcal{D} \to \tilde{\mathcal{D}}$. Using this framework, this subsection provides instantiations for a common case: $\mathcal{D} = \tilde{\mathcal{D}}$.

Restricting domain. In real-world applications, the output domain is often required to match the input domain, i.e. $\mathcal{D} = \tilde{\mathcal{D}}$.[‖] Furthermore, a concrete domain (e.g. $\mathcal{D} = [0, 1)$) does not limit the generality, as $\mathcal{D}$ can be transformed to other domains through scaling and shifting operations. The theorem below characterizes the privacy and utility invariants under such transformations.

Theorem 3.6 (Transformation invariants). *Given a GPM $\mathcal{M} : \mathcal{D} \to \tilde{\mathcal{D}}$ satisfying ε-LDP, if domain $\mathcal{D}' = c\mathcal{D} + d$ and $\tilde{\mathcal{D}}' = c\tilde{\mathcal{D}} + d$ with $c > 0$, then transformation*

$$\mathcal{T}(\mathcal{M}) : x' = cx + d,\ p'_i = p_i / c,\ [l'_i, r'_i) = c[l_i, r_i) + d$$

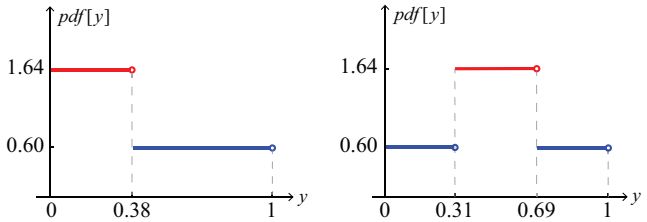
results in GPM $\mathcal{M}' = \mathcal{T}(\mathcal{M}) : \mathcal{D}' \to \tilde{\mathcal{D}}'$ having the same privacy level ε. (privacy invariant)

Meanwhile, if there is another GPM $\mathcal{M}_{bad} : \mathcal{D} \to \tilde{\mathcal{D}}$ with error $Err(x, \mathcal{M}) \leq Err(x, \mathcal{M}_{bad})$ for a given x, then

$$Err(x', \mathcal{M}') \leq Err(x', \mathcal{M}'_{bad}),$$

i.e. $\mathcal{T}$ maintains the data utility ordering. (utility invariant)

[‖] $\tilde{\mathcal{D}}$ generally should be larger than or equal to $\mathcal{D}$ to ensure the mechanism's meaningfulness; otherwise, it means some ranges in $\mathcal{D}$ will disappear after applying the mechanism.



(a) $x = 0$. **(b)** $x = 0.5$.

Figure 4: Optimal GPM (Theorem 3.7) when $\varepsilon = 1$, $x = 0$ and $x = 0.5$.

Proof. (Sketch) The privacy invariant is the same as applying a linear *post-processing* of DP to GPM. The utility invariant is due to $\mathcal{T}$ as a linear function on $\mathcal{D}$. Appendix A.3 provides the full proof of these two invariants. □

Theorem 3.6 allows us to discuss the optimality on a fixed input domain. If a mechanism's input domain differs from $\mathcal{D}$, we can transform it to $\mathcal{D}$, and this transformation maintains the optimality.

Hypothesis 3.1. *For any domain $\mathcal{D} \to \mathcal{D}$, under absolute error and square error metrics, the optimal piecewise-based mechanism falls into* 3*-GPM.*

Support. We validated this hypothesis for $\mathcal{D} = [0, 1)$ by performing Monte Carlo sampling on 10^4 random (ε, x) pairs (detailed in Section 3.2.2). Theorem 3.6 then extends this optimality to any $\mathcal{D}$. Given the continuity of the objective functions w.r.t. ε and x, we posit that $m = 3$ is indeed the exact optimal. To further support this hypothesis, Appendix B.3 outlines two directions for analytical proof and highlights the associated challenges.

Theorem 3.7. *If hypothesis 3.1 holds, then GPM $\mathcal{M} : [0, 1) \to [0, 1)$ with the following closed-form instantiation*

$$pdf[\mathcal{M}(x) = y] = \begin{cases} p_\varepsilon & \text{if } y \in [l_{x,\varepsilon}, r_{x,\varepsilon}), \\ p_\varepsilon / \exp(\varepsilon) & y \in [0, 1) \setminus [l_{x,\varepsilon}, r_{x,\varepsilon}), \end{cases}$$

where $p_\varepsilon = \exp(\varepsilon/2)$,

$$[l_{x,\varepsilon}, r_{x,\varepsilon}) = \begin{cases} [0, 2C) & \text{if } x \in [0, C), \\ x + [-C, C) & \text{if } x \in [C, 1 - C), \\ [1 - 2C, 1) & \text{otherwise,} \end{cases}$$

with $C = (\exp(\varepsilon/2) - 1)/(2\exp(\varepsilon) - 2)$, is optimal for $[0, 1) \to [0, 1)$ under the absolute error and square error metric.

Proof. Provided by analytical deduction with $[a, b) = [\tilde{a}, \tilde{b}) = [0, 1)$, $\mathcal{L} = |y - x|$ and $\mathcal{L} = |y - x|^2$ respectively. Appendix A.4 provides the full proof. □

This optimality on $[0, 1)$ can be transformed to $\mathcal{M}' : [a, b) \to [a, b)$ by applying $\mathcal{T} : x' = (b - a)x + a$, $p'_\varepsilon = p_\varepsilon/(b - a)$, $[l', r') = (b - a) \cdot [l, r) + a$, while maintaining the optimality.

Example 3.8. Figure 4 shows two examples of Theorem 3.7. When $x = 0$ in the left figure, the optimal GPM has $p \approx 1.64$ and $[l, r) \approx [0, 0.38)$. When $x = 0.5$ in the right figure, the optimal GPM has $p \approx 1.64$ and $[l, r) \approx [0.31, 0.69)$.

MSE analysis. We can calculate the mean squared error (MSE) of the optimal GPM in Theorem 3.7 as

$$\begin{aligned} \text{MSE}[\mathcal{M}(x)] &= \int_{\tilde{\mathcal{D}}} (y - x)^2 \cdot pdf[\mathcal{M}(x) = y] \mathrm{d}y \\ &= \int_0^{l_{x,\varepsilon}} (y - x)^2 \frac{p_\varepsilon}{\exp(\varepsilon)} \mathrm{d}y + \int_{l_{x,\varepsilon}}^{r_{x,\varepsilon}} (y - x)^2 p_\varepsilon \mathrm{d}y \\ &\quad + \int_{r_{x,\varepsilon}}^{1} (y - x)^2 \frac{p_\varepsilon}{\exp(\varepsilon)} \mathrm{d}y \,. \end{aligned}$$

which leads to the results in Appendix B.5.

The theoretical MSE allows us to analytically compare the optimal GPM with existing mechanisms. For mechanisms defined on $\mathcal{D} = [0, 1)$, e.g. SW [21], we can compare their MSE with the above result. For mechanisms defined on other domains, e.g. PM [34] on $\mathcal{D} = [-1, 1)$, we can transform the optimal GPM to $[-1, 1)$ and compare their MSE. Detailed comparisons with them are presented in the evaluation section.

4 Optimal GPM for Circular Domain

This section presents the optimal GPM for the circular domain, another type of bounded domain. Circular domains are widely used in cyclic data such as time, angle, and compass direction. However, none of the existing piecewise-based mechanisms consider this type of domain, limiting their applicability.

Different meanings of distance. In the circular domain $[0, 2\pi)$, the distance between two elements differs from that in the classical domain $[0, 2\pi)$. For example, if we denote the distance between x and y in the circular domain as $\mathcal{L}_{\text{mod}}(y, x) = |y - x|$, then it implies $\mathcal{L}_{\text{mod}}(2\pi, 0) = 0$ and $\mathcal{L}_{\text{mod}}(3\pi/2, 0) = \pi/2$, which are different from $\mathcal{L}(y, x) = |y - x|$ in the classical domain. The biggest difference is that there are no endpoints in the circular domain. This unique property makes the mechanisms designed for the classical domain not directly suitable for the circular domain. Although we can "flatten" the circular domain to the classical domain, this conversion changes the distance between elements, leading to data utility loss.

Formally, in the circular domain $[0, 2\pi)$, the distance metric $\mathcal{L}_{\text{mod}}(y, x)$ has the following relationship with $\mathcal{L}(y, x)$ in the classical domain:

$$\mathcal{L}_{\text{mod}}(y, x) = \min\big(\mathcal{L}(y, x), \mathcal{L}(y, 2\pi - x)\big),$$

i.e. the distance between y and x is the smaller one between two arcs from y to x. Under this distance metric, finding the optimal m-GPM for the circular domain is to solve $\mathcal{M} : [0, 2\pi) \to [0, 2\pi)$ such that

$$\min_{p_i, l_{i,x}, r_{i,x}} \max_x \int_0^{2\pi} \mathcal{L}_{\text{mod}}(y, x) \mathcal{P}_{\mathcal{M}(x)} \mathrm{d}y, \quad \text{s.t. } \mathcal{M} \text{ satisfies Definition 3.1,} \tag{4}$$

and use the solved optimal x-independent p_i to determine $l_{i,x}, r_{i,x}$ for any given x:

$$\min_{l_{i,x}, r_{i,x}} \int_0^{2\pi} \mathcal{L}_{\text{mod}}(y, x) \mathcal{P}_{\mathcal{M}(x)} \mathrm{d}y \quad \text{s.t. } \mathcal{M} \text{ satisfies Definition 3.1 with } p_i. \tag{5}$$

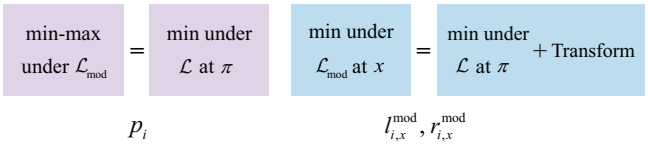


Figure 5: Reduced forms of solving the optimal p_i, $l_{i,x}^{\text{mod}}$ and $r_{i,x}^{\text{mod}}$. Optimizations under circular distance $\mathcal{L}_{\text{mod}}$ can be reduced to those under linear distance $\mathcal{L}$.

We show that these two problems can be reduced to those in the classical domain, thereby enabling the usage of existing results.

4.1 Reduced Forms

Similar to the classical domain, the min-max objective of Formulation (4) also has a closed-form solution, which reduces the problem to the classical domain.

Lemma 4.1. *The objective of Formulation (4) can be reduced to*

$$\min_{p_i, l_{i,x}, r_{i,x}} \int_0^{2\pi} \mathcal{L}(y, \pi) \mathcal{P}_{\mathcal{M}(\pi)} \mathrm{d}y.$$

Proof. (Sketch) We prove it by showing that, for any fixed y, $\max_x \mathcal{L}_{\text{mod}}(y, x) = \mathcal{L}_{\text{mod}}(y, \pi) = \mathcal{L}(y, \pi)$, i.e. the maximum distance between y and x is achieved at a unique $x = \pi$ for any y. Appendix A.5 provides the full proof. □

Following this reduction, the optimal results in the classical domain $[0, 2\pi)$ can be applied to determine the optimal p_i.

For Formulation (5) to solve $l_{i,x}$ and $r_{i,x}$, we shift the domain $[0, 2\pi)$ by $\pi - x$. This is a trick operation as $\mathcal{L}_{\text{mod}}(y, x)$ is transformed to the following form:

$$\begin{aligned}&\mathcal{L}_{\text{mod}}(y + \pi - x, x + \pi - x) = \mathcal{L}_{\text{mod}}(y + \pi - x, \pi)\\ =&\min\left(\mathcal{L}(y + \pi - x, \pi), \mathcal{L}(y + \pi - x, 2\pi - \pi)\right)\\ =&\mathcal{L}(y + \pi - x, \pi),\end{aligned}$$

which transforms $\mathcal{L}_{\text{mod}}$ to $\mathcal{L}$ at $x = \pi$. Then, the optimization problem in the shifted domain becomes

$$\min_{l_{i,x}, r_{i,x}} \int_{\pi - x}^{3\pi - x} \mathcal{L}(y + \pi - x, \pi) \mathcal{P}_{\mathcal{M}(\pi)} \mathrm{d}y.$$

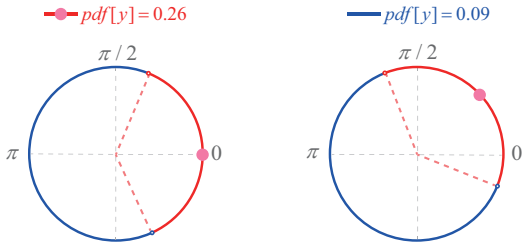
It is a problem in the classical domain $[\pi - x, 3\pi - x)$. We can obtain the closed-form optimal $l_{i,x}$ and $r_{i,x}$ by applying the results of the classical domain. Since the obtained $l_{i,x}$ and $r_{i,x}$ depends on x in the shifted domain, we shift them back to the circular domain using

$$\begin{aligned}l_{i,x}^{\text{mod}} &= l_{i,x} - (\pi - x) \mod 2\pi,\\ r_{i,x}^{\text{mod}} &= r_{i,x} - (\pi - x) \mod 2\pi.\end{aligned}$$

Transformation invariants ensure their optimality in the circular domain. Figure 5 summarizes the above two reductions to solve the optimal p_i, $l_{i,x}^{\text{mod}}$ and $r_{i,x}^{\text{mod}}$ in the circular domain.

4.2 Closed Form for Circular Domain

By applying the above reductions and following the same steps as in the classical domain, we can obtain the optimal GPM for the circular domain.



(a) $x = 0$. **(b)** $x = \pi/4$.

Figure 6: Optimal GPM (Theorem 4.2) for the circular domain with $\varepsilon = 1$. Here $[0, 2\pi)$ is wrapped into a circle. Angle values in the red arc (centered at x) have a higher sampling pdf.

Theorem 4.2. *If Hypothesis 3.1 holds, then GPM $\mathcal{M} : [0, 2\pi) \to [0, 2\pi)$ with the following closed-form instantiation*

$$pdf[\mathcal{M}(x) = y] = \begin{cases} p_\varepsilon & \text{if } y \in [l_{x,\varepsilon}^{\text{mod}}, r_{x,\varepsilon}^{\text{mod}}),\\ p_\varepsilon / \exp(\varepsilon) & y \in [0, 2\pi) \setminus [l_{x,\varepsilon}^{\text{mod}}, r_{x,\varepsilon}^{\text{mod}}),\end{cases}$$

where $p_\varepsilon = \frac{1}{2\pi} \exp(\varepsilon/2)$,

$$\begin{aligned}l_{x,\varepsilon}^{\text{mod}} &= \left(x - \pi \frac{\exp(\varepsilon/2) - 1}{\exp(\varepsilon) - 1}\right) \mod 2\pi,\\ r_{x,\varepsilon}^{\text{mod}} &= \left(x + \pi \frac{\exp(\varepsilon/2) - 1}{\exp(\varepsilon) - 1}\right) \mod 2\pi,\end{aligned}$$

is optimal for the circular domain under the absolute error and square error metric.

Proof. Combination of the reduced forms, the results of the classical domain, Lemma 4.1, and transformation invariants leads to the conclusion. □

Compared to the optimal GPM for the classical domain in Theorem 3.7, the key difference is that the mechanism in the circular domain allows $[l_{x,\varepsilon}^{\text{mod}}, r_{x,\varepsilon}^{\text{mod}})$ to span the 0 or 2π boundary, significantly reducing the error. We also observe that their instantiations of $p_\varepsilon, l_{x,\varepsilon}, r_{x,\varepsilon}$ are connected through transformation invariants. For instance, moving from the classical domain $[0, 1)$ to the circular domain $[0, 2\pi)$, $p_\varepsilon = \exp(\varepsilon/2)$ transforms to $p_\varepsilon = \frac{1}{2\pi} \exp(\varepsilon/2)$, reflecting the ratio of $[0, 1)$ to $[0, 2\pi)$.

Example 4.3. Figure 6 shows two examples of Theorem 4.2 when $\varepsilon = 1$. For $x = 0$ in the left figure, it samples the output y from $[1.62\pi, 2\pi) \cup [0, 0.38\pi)$ with probability density 0.26 and from $[0.38\pi, 1.62\pi)$ with probability density 0.09.

MSE analysis. The MSE of the optimal GPM in Theorem 4.2 needs a separate analysis due to the circular domain. The biggest difference from the classical domain is that there exist no endpoints, i.e. fixed farthest points, in the circular domain. Without loss of generality, assume $\mathcal{L} := |y - x|^2$ and $x > \pi$, the farthest distance from x is always π, i.e. from x to $x - \pi$. If we shift the data domain by $\pi - x$, then point $x - \pi$ is mapped to 0. This domain shift does not change the value of $\mathcal{L}$, but x now locates at π. Therefore, the MSE of the optimal GPM in Theorem 4.2 has an identical value for

all x in the circular domain, which is

$$\begin{aligned}&\mathrm{MSE}[\mathcal{M}(x)] = \mathrm{MSE}[\mathcal{M}(\pi)]\\ =&2\left(\int_0^{l^{\mathrm{mod}}_{\pi,\varepsilon}} (y-\pi)^2 \frac{p_\varepsilon}{\exp(\varepsilon)}\mathrm{d}y + \int_{l^{\mathrm{mod}}_{\pi,\varepsilon}}^{\pi} (y-\pi)^2 p_\varepsilon \mathrm{d}y\right).\end{aligned}$$

Calculating the above integral, we can obtain the MSE of the optimal GPM in the circular domain.

Theorem 4.4. *The optimal GPM in Theorem 4.2 has an identical MSE for all x in the circular domain, which is*

$$\mathrm{MSE}[\mathcal{M}(x)] = \frac{2}{3}\left((\pi^3 - C^3)\frac{p_\varepsilon}{\exp(\varepsilon)} + C^3 p_\varepsilon\right),$$

where $C = \pi(\exp(\varepsilon/2)-1)/(\exp(\varepsilon)-1)$.

Proof. We have shown that the MSE at x is the same as that at π. Then the calculation is straightforward by plugging in the values of p_ε and $l^{\mathrm{mod}}_{\pi,\varepsilon}$ in Theorem 4.2. □

If we do not consider $\mathcal{L}_{\mathrm{mod}}$ and directly apply the optimal GPM in the classical domain (or SW and PM mechanisms) to the circular domain, i.e. flatten the circular domain to the classical domain $[0, 2\pi)$ and consider $\mathcal{L}$, the MSE will varies for different x. In the flattened domain, the worst-case MSE is at $x = 0$ or $x = 2\pi$ and the best-case MSE is at $x = \pi$. Therefore, the optimal GPM in Theorem 4.2 always has a lower MSE than "flattened" mechanisms. The evaluation section will further demonstrate this.

5 Distribution and Mean Estimation

This section applies our mechanisms to support the commonly used distribution and mean estimation. We will show that our mechanisms also provide optimality for these estimation tasks.

Assume a set of users with sensitive data $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$. They apply $\mathcal{M}$ to produce randomized outputs $\mathcal{Y} = \{y_1, y_2, \ldots, y_n\}$. The data collector then estimates the distribution and mean of $\mathcal{X}$ using $\mathcal{Y}$. Specifically, the collector uses the values in $\mathcal{Y}$ as it is, i.e. without knowing and applying any post-processing based on prior knowledge of $\mathcal{X}$.

5.1 Distribution Estimation

To estimate the distribution of $\mathcal{X}$ from $\mathcal{Y}$, the collector must discretize the continuous domain $\mathcal{D}$ into k bins $B_1, B_2, \ldots, B_k$. Each bin B_j's probability is then estimated by the proportion of y_i that falls within B_j. Probabilities of all bins together form the estimated distribution $\hat{\mathcal{F}}_B$ using $\mathcal{Y}$. This estimation's accuracy is measured by the distance between the estimated distribution $\hat{\mathcal{F}}_B$ and the true distribution $\mathcal{F}_B$ from $\mathcal{X}$ [21].

Note that the bin size can impact the estimation accuracy due to the rounding of the bins. However, it is actually a hyperparameter that does not inherently affect the estimation accuracy if we have a sufficiently large number of bins, as the support of $\hat{\mathcal{F}}_B$ (i.e. non-zero bins) converges to $\mathcal{Y}$ and the support of $\mathcal{F}_B$ converges to $\mathcal{X}$, i.e.

$$\lim_{k\to\infty} \mathrm{supp}(\hat{\mathcal{F}}_B) = \mathcal{Y}, \quad \lim_{k\to\infty} \mathrm{supp}(\mathcal{F}_B) = \mathcal{X}.$$

Our mechanisms guarantee the optimal error between $\mathcal{X}$ and $\mathcal{Y}$ for each x_i. This ensures their optimality among GPM when applied to distribution estimation.

Some other statistical estimations are special applications of distribution estimation, such as range and quantiles [21] to estimate a specific part of the distribution. Our mechanisms have optimality for these estimations as well.

5.2 Mean Estimation

To estimate the mean of $\mathcal{X}$ from $\mathcal{Y}$, the collector uses the estimator $\hat{\mu} = \sum_{i=1}^n y_i/n$. The accuracy of this estimator is measured by $|\hat{\mu}-\mu|$, where μ is the true mean of $\mathcal{X}$. Our mechanisms guarantee the optimal error between each x_i and y_i, which in turn leads to the smallest $|\hat{\mu}-\mu|$ among all GPMs under this metric.

Typically, a mean estimator may also need to be unbiased, i.e. $\mathrm{E}[\hat{\mu}] = \mu$. This constraint translates to $\mathrm{E}[y_i] = \mathrm{E}[\mathcal{M}(x_i)] = x_i$.** This is unachievable for same-domain mapping $\mathcal{M}: \mathcal{D} \to \mathcal{D}$ on classical domains, as the endpoints of $\mathcal{D}$ cannot be the mean value (or center) of any distribution over $\mathcal{D}$. For example, for any LDP mechanism $\mathcal{M}: [0,1) \to [0,1)$, distribution of $\mathcal{M}(0)$ can not be unbiased. So the mechanism in Theorem 3.7 is biased.

Unbiased mean estimation. Note that an unbiased mean estimator can be achieved by enlarging the output domain $\mathcal{D} \to \tilde{\mathcal{D}}_\varepsilon$. Mathematically, this involves incorporating the unbiasedness constraint $\mathrm{E}[\mathcal{M}(x)] = x$ into optimization problems for solving $\mathcal{M}$. Following the same optimization process as in the classical domain, we hypothesize that the 3-GPM remains optimal for domain $\tilde{\mathcal{D}}_\varepsilon$.

Hypothesis 5.1. *For any domain $\mathcal{D} \to \tilde{\mathcal{D}}_\varepsilon$, where $\tilde{\mathcal{D}}_\varepsilon$ is a variable w.r.t ε, and under absolute error and square error metrics, the optimal piecewise-based mechanism falls into* 3*-GPM.*

Hypothesis 5.1 is a natural extension of Hypothesis 3.7, as the output domain $\tilde{\mathcal{D}}_\varepsilon$ becomes explicit once ε is specified. Under the 3-GPM, an unbiased mechanism $\mathcal{M}$ with a variable output domain $\tilde{\mathcal{D}}_\varepsilon$ can be analytically derived by incorporating the unbiasedness constraint. As a complement to Theorem 3.7, we provide Theorem 5.1 for mean estimation in the classical domain.

Theorem 5.1. *Denote $\tilde{\mathcal{D}}_\varepsilon = [-C, C+1)$ with $C = (\exp(\varepsilon/2)+1)/(\exp(\varepsilon/2)-1)$. If Hypothesis 5.1 holds, then among the unbiased GPM $\mathcal{M}: \mathcal{D} \to \tilde{\mathcal{D}}_\varepsilon$ (i.e. $\mathrm{E}[\mathcal{M}(x)] = x$), closed form*

$$pdf[\mathcal{M}(x) = y] = \begin{cases} p_\varepsilon & \text{if } y \in [l_{x,\varepsilon}, r_{x,\varepsilon}),\\ p_\varepsilon/\exp(\varepsilon) & y \in \tilde{\mathcal{D}}_\varepsilon \setminus [l_{x,\varepsilon}, r_{x,\varepsilon}),\end{cases}$$

where $p = \exp(\varepsilon/2)/(2C+1)$,

$$l_{x,\varepsilon} = \frac{C+1}{2}\cdot x - \frac{(3C+1)(C-1)}{4C},$$
$$r_{x,\varepsilon} = \frac{C+1}{2}\cdot x + \frac{(C+1)(C-1)}{4C}.$$

is optimal for $[0,1) \to \tilde{\mathcal{D}}_\varepsilon$ and the square error metric.

Proof. Instantiations of C, p, $l_{x,\varepsilon}$, and $r_{x,\varepsilon}$ are derived by analytical deduction. Appendix A.6 proves the unbiasedness. □

In the context of the circular domain, the $\mathcal{M}$ in Theorem 4.2 is unbiased, as the distribution of $\mathcal{M}$ is always centered at x. This property is also illustrated in Figure 6.

**Actually, unbiasedness for numerical data is not as important as for categorical data, as it is for a single data point x_i. When the dataset $\mathcal{X}$ is not concentrated around a single point, an unbiased mechanism may not necessarily provide better performance.

Table 2: Optimal distribution and mean estimation in three domain types under GPM.

	Distribution	Mean
Classical ($\mathcal{D} \to \mathcal{D}$)	Theorem 3.7	Theorem 3.7 (biased)
Circular domain	Theorem 4.2	Theorem 4.2 (unbiased)
Classical ($\mathcal{D} \to \tilde{\mathcal{D}}_\varepsilon$)	–	Theorem 5.1 (unbiased)

Table 2 summarizes the optimal GPM for both distribution and mean estimation in three domain types. Theorems use a concrete domain $\mathcal{D}$ for the sake of clarity. We do not give closed-form optimal GPM for distribution estimation on $\mathcal{D} \to \tilde{\mathcal{D}}_\varepsilon$ because $\tilde{\mathcal{D}}_\varepsilon$ can not be easily concretized by constraints as in mean estimation.

6 Discussion and Extension

Minimize Error at a Specific x. Our proposed optimal GPMs are designed to minimize the worst-case error over the whole domain. However, the framework can be used to minimize the error at any specific x. This is useful when the data distribution is concentrated around a specific data point.

Formally, assume the data distribution is concentrated around x_0, and we want to minimize the error at x_0. The optimization problem in Lemma 3.2 for solving p_i can be modified to

$$\min_{p_i, l_{i,x}, r_{i,x}} \int_{\tilde{\mathcal{D}}} \mathcal{L}(y, x_0) \mathcal{P}_{\mathcal{M}(x_0)} \mathrm{d}y.$$

This optimization problem generally leads to a different p_i from the optimal GPM for the worst-case error.

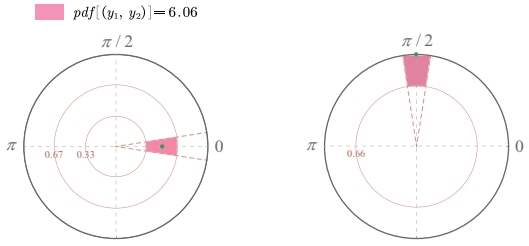
Example 6.1. Assume $\mathcal{D} \to \tilde{\mathcal{D}} = [0, 1) \to [0, 1)$, $\mathcal{L} := |y - x|$, and the data distribution is concentrated around $x_0 = 0.2$. Solving the above optimization problem with $\varepsilon = 1$ gives probability (of the second piece) $p_2 = 1.54$ and $Err(x_0) = 0.241$. In contrast, the optimal GPM for the worst-case error uses $p_2 = 1.64$, as shown in Figure 4, which gives $Err(x_0) = 0.243$.

Importantly, optimizing for a specific x_0 does not leak information about x_0, as the mechanism (i.e. p_i, $l_{i,x}$, and $r_{i,x}$) still contains no information about x_0. Moreover, observing $Err(x)$ at all x does not reveal x_0, as the error at x_0 is not necessarily the smallest. Therefore, the adversary cannot infer x_0 from observing the mechanism.

An Extension: 2D Polar Coordinates. Polar coordinates are widely used in relative location representation, e.g. navigation systems that have the locations of surrounding objects relative to it. Our proposed optimal GPM for the classical and circular domain can be combined and extended to polar coordinates for collecting such data under LDP.

Privacy. A polar coordinate data is represented by a 2D tuple $(x_1, x_2) \in [0, d) \times [0, 2\pi)$, where x_1 is the distance from the pole and x_2 is the angle from the polar axis. The first dimension is linear, while the second is naturally circular, thus we can combine the optimal GPM for both domains to provide LDP for such data.

Utility. Our mechanisms guarantee the optimal error for each dimension. Therefore, if we use $\mathcal{L}_{2\mathrm{D}} := \mathcal{L}(y_1, x_1) + \mathcal{L}_{\mathrm{mod}}(y_2, x_2)$ as the error metric for the polar coordinate data and optimally assign the privacy parameter ε to each dimension, the optimal GPM preserves the optimal error.



(a) $(x_1, x_2) = (0.5, 2\pi)$. **(b)** $(x_1, x_2) = (1, \pi/2)$.

Figure 7: Optimal GPM for 2D polar coordinates when $\varepsilon = 1 + 2\pi$ and $\mathcal{L} = |y - x|^2$ under $\mathcal{L}_{\mathbf{2D}} := \mathcal{L}(y_1, x_1) + \mathcal{L}_{\mathbf{mod}}(y_2, x_2)$.

Optimal assignment of ε. Formally, we want to assign $\varepsilon = \varepsilon_1 + \varepsilon_2$ to x_1 and x_2 respectively to minimize the worst-case error in 2D polar coordinates. This error is the sum of the worst-case error for x_1 and x_2 under $\mathcal{L}$ and $\mathcal{L}_{\mathrm{mod}}$ respectively. Therefore, the optimal assignment of ε can be derived by solving

$$\min_{\varepsilon_1, \varepsilon_2} Err_{1,\mathrm{wor}}(\varepsilon_1) + Err_{2,\mathrm{wor}}(\varepsilon_2),$$

where $Err_{1,\mathrm{wor}}(\varepsilon_1)$ and $Err_{2,\mathrm{wor}}(\varepsilon_2)$ are the worst-case error for the classical domain $[0, d)$ and the circular domain $[0, 2\pi)$ respectively. Closed-form $Err_{1,\mathrm{wor}}(\varepsilon_1)$ and $Err_{2,\mathrm{wor}}(\varepsilon_2)$ can be derived from instantiations of the mechanism and the error metric. Then the above optimization problem gives the optimal ε_1 and ε_2. Appendix B.6 provides details for solving this optimization problem.

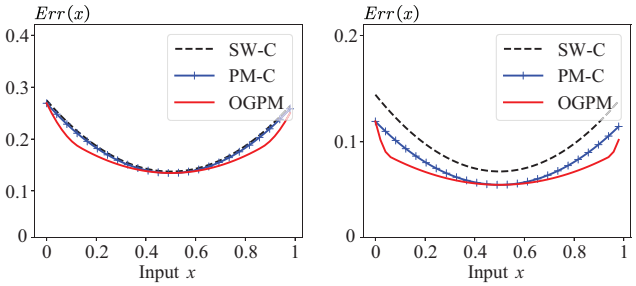
Example 6.2. Figure 7 shows two examples of the optimal GPM for 2D polar coordinates in $[0, 1) \times [0, 2\pi)$ with $\varepsilon = 1 + 2\pi$. The green point represents the sensitive data, the optimal GPM samples the output from the pink area with a higher probability.

7 Evaluations

This section evaluates the theoretical and experimental data utility of our methods by comparing them with existing instantiations and their variants:

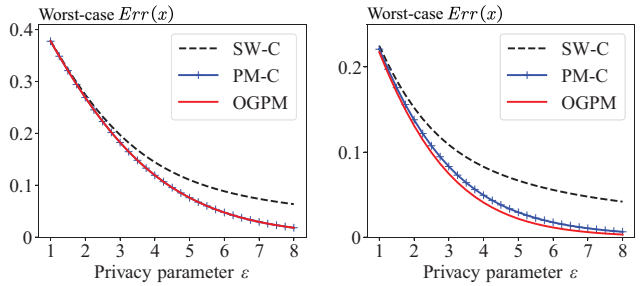
- OGPM: closed-form optimal GPM (Theorem 3.7 and 4.2 for the classical and circular domain, respectively).
- OGPM-U: unbiased closed-form optimal GPM (Theorem 5.1) for mean estimation in the classical domain.
- PM [34], SW [21], and their post-processed versions: PM is the first TPM designed for mean estimation, while SW is designed for distribution estimation. Both mechanisms output enlarged domains but can be post-processed by truncating outputs to the input domain. These post-processed versions are referred to as T-PM and T-SW for convenience.
- PM-C and SW-C: the compressed versions of PM and SW for $\mathcal{D} \to \mathcal{D}$. For the best potential of PM and SW, we adapt them to $\mathcal{D} \to \mathcal{D}$ as PM-C and SW-C by linearly compressing their output domain $\tilde{\mathcal{D}}_\varepsilon$ to $\mathcal{D}$, i.e. by transformation invariants, which maintains the privacy level.

We also compare OGPM's expected error with non-piecewise-based mechanisms that can be applied to bounded numerical domains:



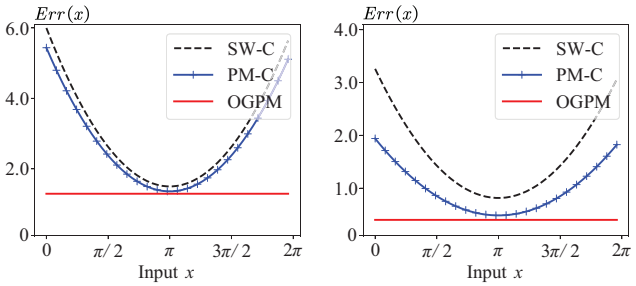
(a) Privacy parameter $\varepsilon = 2$. **(b) Privacy parameter $\varepsilon = 4$.**

Figure 8: Whole-domain error comparison in the classical domain with error metric $\mathcal{L} = |y - x|$.



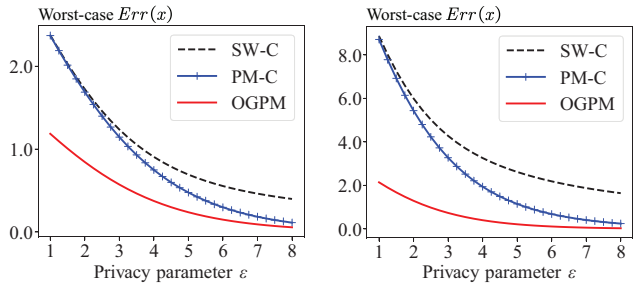
(a) Distance metric $\mathcal{L} = |y - x|$. **(b) Distance metric $\mathcal{L} = |y - x|^2$.**

Figure 9: Worst-case error comparison in the classical domain.



(a) Privacy parameter $\varepsilon = 2$. **(b) Privacy parameter $\varepsilon = 4$.**

Figure 10: Whole-domain error comparison in the circular domain with error metric $\mathcal{L} = |y - x|^2$.



(a) Distance metric $\mathcal{L} = |y - x|$. **(b) Distance metric $\mathcal{L} = |y - x|^2$.**

Figure 11: Worst-case error comparison in the circular domain.

- Variants of the Laplace mechanism: including the staircase mechanism [15], Laplace mechanism with post-processing by truncation (T-Laplace), and the bounded Laplace mechanism (B-Laplace) [19], which redesigns a bounded Laplace-shape distribution.
- Purkayastha mechanism [36]: a mechanism for directional data on spheres $\mathbb{S}^{n-1}$. When $n = 2$, it is a counterpart of OGPM in the circular domain.

We omit comparison with PTT [23] because it does not provide a concrete method to find a closed-form mechanism. We use $\mathcal{D} = [0, 1)$ as the classical domain; this does not change their data utility ordering by the transformation invariant. PM and SW can only be applied to the classical domain, so when evaluating them in the circular domain, we "flatten" the circular domain to the classical domain $[0, 2\pi)$ and apply them to the flattened domain.

The first subsection presents the comparison of expected errors, followed by the distribution and mean estimations on real-world datasets in the second subsection.

7.1 Expected Errors

GPM's data utility under distance metric $\mathcal{L}$ is measured by the expected error $Err(x)$ in Formula (1):

$$Err(x) = \int_{\tilde{\mathcal{D}}} \mathcal{L}(y, x) \mathcal{P}_{\mathcal{M}(x)} \mathrm{d}y,$$

where x is the sensitive input and y is the output of $\mathcal{M}$. Based on $Err(x)$, two types of error need to be considered:

- **whole-domain error:** $Err(x)$ values for the whole domain, i.e. for all $x \in \mathcal{D}$.
- **worst-case error:** the largest $Err(x)$ value among the whole domain. PM [34] and PTT [23] also use this error to evaluate data utility.

We have proved that the worst-case error of the classical domain is from the endpoints, so this error is actually the error at $x = 0$ and $x = 1$ for the classical domain. For the circular domain, the worst-case error is at $x = \pi$.

7.1.1 Comparison with PM-C and SW-C. PM-C and SW-C exhibit the best potential of PM and SW, so we compare OGPM with them in the first place. The comparisons are conducted under the classical domain and the circular domain, respectively.

Classical Domain. Figure 8 shows the comparison of the whole-domain error in the classical domain. We use distance metric $\mathcal{L} = |y - x|$ and set $\varepsilon = 2$ and $\varepsilon = 4$ for the comparison. It can be seen that OGPM consistently has the lowest error across all x values. For all mechanisms, the expected error achieves the maximal value at the endpoints and the minimal value at the midpoint. At small ε values, their errors are similar due to the strong randomness (privacy constraint). At larger ε values, OGPM's error has a significant advantage over PM-C and SW-C, especially between the endpoints and the midpoint. Statistically, under $\mathcal{L} = |y - x|$ with $\varepsilon = 2$,

OGPM's average error is 94.2% of PM-C and 92.3% of SW-C across the whole domain. When $\varepsilon = 4$, OGPM's average error is 90.5% of PM-C and 74.7% of SW-C. More comparisons under smaller ε values are provided in Appendix B.7.

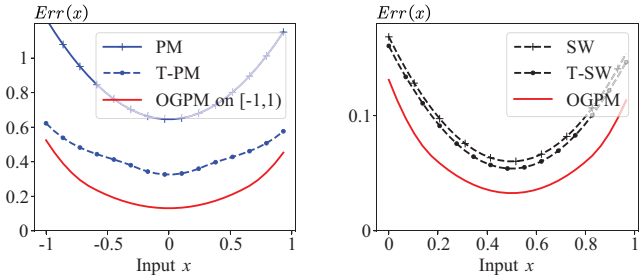
Figure 9 shows the comparison of the worst-case error w.r.t ε in the classical domain. The error is measured by two distance metrics: $\mathcal{L} = |y - x|$ and $\mathcal{L} = |y - x|^2$. It can be seen that the error of all mechanisms decreases with ε, but OGPM and PM-C decreases faster than SW-C. For $\mathcal{L} = |y - x|$, OGPM has almost the same worst-case error as PM-C; for $\mathcal{L} = |y-x|^2$, OGPM's error is slightly smaller than PM-C. For both metrics, OGPM's worst-case error is the lowest across all ε values. Statistically, under $\mathcal{L} = |y - x|^2$, OGPM's average error is 89.9% of PM-C and 61.7% of SW-C.

Circular Domain. Figure 10 shows the comparison of the whole-domain error in the circular domain. We use distance metric $\mathcal{L} = |y - x|^2$ and set $\varepsilon = 2$ and $\varepsilon = 4$ for the comparison. It can be seen that OGPM consistently has the lowest error across all x values, and the error is stable across x, which is consistent with the theoretical analysis in Theorem 4.4. For PM-C and SW-C, which treat the circular domain as the classical domain, their errors vary with x and are higher than OGPM's error, especially near the endpoints. Statistically, under $\mathcal{L} = |y-x|^2$ with $\varepsilon = 2$, OGPM's average error is 47.5% of PM-C and 43.0% of SW-C across the whole domain. When $\varepsilon = 4$, OGPM's average error is 41.3% of PM-C and 24.6% of SW-C.

Figure 11 shows the comparison of the worst-case error w.r.t ε in the circular domain. The error is measured by two distance metrics: $\mathcal{L} = |y - x|$ and $\mathcal{L} = |y - x|^2$. Similar to the classical domain, OGPM has the lowest worst-case error across all ε values, and the advantage is more significant, especially for small ε values. Statistically, under $\mathcal{L} = |y - x|$, OGPM's average error is 50.0% of PM-C and 41.6% of SW-C across the range of ε. Under $\mathcal{L} = |y - x|^2$, OGPM's average error is 22.4% of PM-C and 15.4% of SW-C.

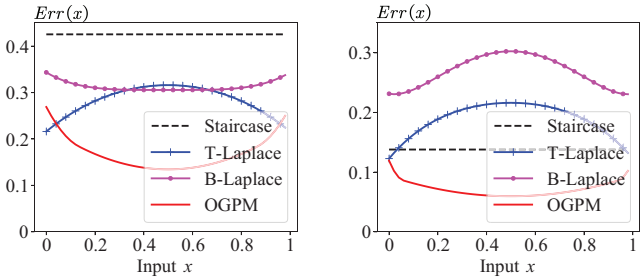
The above comparisons show that OGPM has the lowest expected error in both the classical and circular domains. The advantage of OGPM is more significant in the circular domain, where the error is stable across x.

7.1.2 Comparison with PM and SW. Figure 12 presents the comparison of the whole-domain error in the classical domain for the original PM and SW mechanisms, along with their post-processed versions, T-PM and T-SW. For a fair comparison, OGPM is adapted to the domain $\mathcal{D} = [-1, 1)$ to match PM's design, while SW and OGPM remain consistent with $\mathcal{D} = [0, 1)$. The post-processing of PM and SW involves truncating their outputs in the enlarged domain to the input domain, i.e. applying $\mathcal{I} \circ \mathcal{M}(x)$, where $\mathcal{I} : \tilde{\mathcal{D}} \to \mathcal{D}$ is the truncation operator. We use the distance metric $\mathcal{L} = |y-x|^2$ and set $\varepsilon = 2$ for the comparison among these five mechanisms. It can be observed that OGPM consistently achieves the lowest error across all x values, with a more significant advantage compared to the comparison with PM-C and SW-C. This is because the original PM and SW output larger domains, resulting in higher errors. Meanwhile, T-PM reduces the error of PM more effectively than T-SW reduces the error of SW, as the original PM has a more enlarged output domain than SW, making truncation more impactful. This comparison highlights OGPM's error advantage over the original PM, SW, and their post-processed versions when applied to their respective data domains.



(a) Comparison with PM. **(b) Comparison with SW.**

Figure 12: Whole-domain error comparison with PM and SW on their data domains (i.e. $\mathcal{D} = [-1, 1)$ and $\mathcal{D} = [0, 1)$, respectively) when $\varepsilon = 2$.



(a) Privacy parameter $\varepsilon = 2$. **(b) Privacy parameter $\varepsilon = 4$.**

Figure 13: Whole-domain error comparison with the staircase mechanism [15], T-Laplace and B-Laplace mechanisms [19] in the classical domain with error metric $\mathcal{L} = |y - x|$.

7.1.3 Comparison with the Staircase Mechanism, T-Laplace, and B-Laplace. In addition to piecewise-based mechanisms, the Laplace mechanism and its variants can also be applied to the classical domain to achieve LDP. Among these, the staircase mechanism [15] claims to be optimal under certain assumptions. For the input domain $\mathcal{D} = [0, 1)$ (i.e. sensitivity $\Delta = 1$) and error metric $\mathcal{L} = |y - x|$, its expected error is given by Theorem 3 in [15]: $\exp(\varepsilon/2)/(\exp(\varepsilon) - 1)$. Another approach involves using the Laplace mechanism with truncation [19], referred to here as T-Laplace for convenience. T-Laplace preserves the privacy guarantees of the Laplace mechanism while reducing the expected error, particularly for data points near the endpoints or for small ε values. Additionally, the bounded Laplace mechanism (B-Laplace) [19] introduces a redesigned bounded Laplace-shaped distribution tailored for bounded domains.[††]

Figure 13 compares the whole-domain error in the classical domain $\mathcal{D} = [0, 1)$ for the staircase mechanism, T-Laplace, and B-Laplace. These mechanisms exhibit distinct error patterns across the domain. For the staircase mechanism, the error remains constant, as it is determined by a fixed staircase distribution and is independent of x. For T-Laplace, the error reaches its maximum at the midpoint and its minimum at the endpoints, as truncation favors the endpoints. For instance, when $x = 0$, it is error-free with a probability of $1/2$, due to the symmetry of the Laplace distribution

[††]Appendix B.8 provides details on the expected error of B-Laplace.

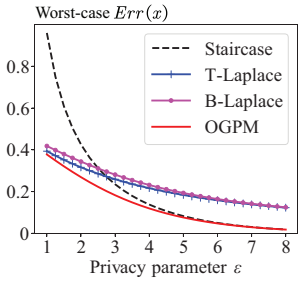


Figure 14: Worst-case error comparison (continued from Figure 13).

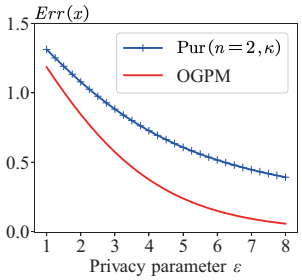


Figure 15: Comparison with the Purkayastha mechanism [36] for sphere $\mathbb{S}^{n-1}$.

around 0. For B-Laplace, the error trend varies with ε. When $\varepsilon = 2$, the error decreases with x and reaches its minimum at the midpoint, whereas for $\varepsilon = 4$, the error increases with x and peaks at the midpoint. Despite these differing error patterns, OGPM generally achieves lower errors than the staircase mechanism, T-Laplace, and B-Laplace across the whole domain.

Figure 14 compares the worst-case error w.r.t. ε in the classical domain. OGPM consistently achieves the lowest worst-case error across all ε values. For small ε, T-Laplace and B-Laplace exhibit a significant advantage over the staircase mechanism; however, this advantage diminishes as ε increases. At larger ε values, the error of the staircase mechanism approaches that of OGPM.

7.1.4 Comparison with the Purkayastha Mechanism. The paper "Differential Privacy for Directional Data" [36] introduces two mechanisms for data on spheres $\mathbb{S}^{n-1}$: the VMF mechanism (ensuring indistinguishability of any two points with distance *through* the sphere) and the Purkayastha mechanism (ensuring indistinguishability of any two points with distance *along* the sphere). When $n = 2$, the sphere $\mathbb{S}^1$ corresponds to a circle, making the Purkayastha mechanism a counterpart of OGPM in the circular domain. Therefore, we compare them in the circular domain.[‡‡]

Figure 15 presents the comparison of the expected error in the circular domain between OGPM and the Purkayastha mechanism. The expected error of the Purkayastha mechanism is derived using the closed-form expressions in Theorem 19 and 22 of [36], with $\kappa = \varepsilon/\Delta_{\angle}$. Since the errors of both mechanisms are x-independent in the circular domain, it suffices to compare their worst-case errors. The results demonstrate that OGPM consistently outperforms the Purkayastha mechanism, achieving significantly lower errors.

7.2 Distribution and Mean Estimations

This section compares the experimental data utility of the mechanisms in distribution and mean estimations.

7.2.1 Setup. We choose the MotionSense dataset [1, 24] for the evaluation. It contains smartphone sensor data recorded during various human activities. Specifically, we use the data from the first three files, encompassing a total of 6 159 data entries. We focus on two types of sensors:

- Accelerometer (linear data): We normalize the dataset to $[0, 1)$ for the classical domain.
- Attitude sensor (angular data): We use this dataset for the circular domain.

Upon applying each LDP mechanism to these datasets, we evaluate the accuracy of distribution and mean estimations on them. For distribution estimation, we divide the domain into $k = 50$ bins and compute the distance between the estimated distribution ($\hat{\mathcal{F}}_B$) and the true distribution ($\mathcal{F}_B$) by summing the absolute difference of each bin's value. Formally,

$$| \hat{\mathcal{F}}_B - \mathcal{F}_B | := \sum_{i=1}^{50} | \hat{\mathcal{F}}_{B_i} - \mathcal{F}_{B_i} |,$$

where $\hat{\mathcal{F}}_{B_i}$ and $\mathcal{F}_{B_i}$ are the i-th bin's value of the estimated and true distributions, respectively. Although this approach cannot capture the property of the circular data, it remains the most viable metric for comparing circular distributions [13, 25]. Under a more relevant approach, the performance of OGPM for the circular domain could be even better.

For mean estimation, we compute the absolute difference between the estimated and true mean, i.e. $|\hat{\mu} - \mu|$, where $\hat{\mu}$ is the estimated mean and μ is the true mean in the classical domain or the circular domain. In the classical domain, the true mean is $\mu = \frac{1}{n}\sum_{i=1}^{n} x_i$. In the circular domain, the mean is computed by the circular mean formula [13, 25]:

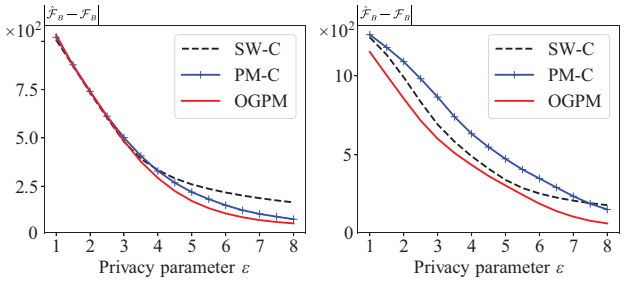
$$\mu = \text{atan2}\left(\frac{1}{n}\sum_{i=1}^{n}\sin x_i, \frac{1}{n}\sum_{i=1}^{n}\cos x_i\right).$$

We repeat the experiments 500 times for stable results and report the average error.

7.2.2 Distribution Estimation. Figure 16 shows the comparison of the errors of distribution estimation in the classical and circular domains. We can see that OGPM outperforms SW and PM with smaller errors in both types of domains. In the classical domain, OGPM's error decreases faster when ε increases above 3, and SW-C decreases slower than PM-C, consistent with the expected error comparison. In the circular domain, OGPM's error is significantly lower than SW-C and PM-C, despite the limitation of the distance metric used for circular distributions. We also observe that SW-C performs better than PM-C in this domain. This is because SW has higher sampling probabilities for both the central piece and other pieces, making it sample the true value more frequently when ε is large in practice in a large-size domain, despite the large expected error theoretically. Statistically, OGPM's distribution error is 93.5% of PM-C and 86.7% of SW-C in the classical domain, and 72.2% of PM-C and 84.0% of SW-C in the circular domain.

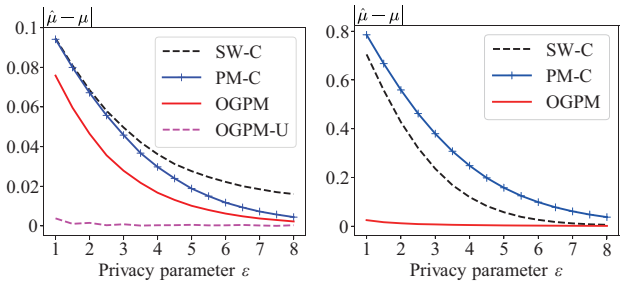
7.2.3 Mean Estimation. Figure 17 shows the comparison of the errors of mean estimation in the classical and circular domains. OGPM consistently outperforms other mechanisms in both types of domains, with significantly lower errors. In the classical domain, we also compare with OGPM-U, which is specifically designed for unbiased mean estimation. Since the Accelerometer dataset is concentrated around zero, it particularly favors unbiased mechanisms,

[‡‡]We omit the comparison with the VMF mechanism also because (i) it has been shown that the Purkayastha mechanism outperforms the VMF mechanism (with the same sensitivity $\Delta_{\angle} = \pi$ for sphere $\mathbb{S}^1$, e.g. Figure 5 and 10 in [36]), and (ii) the expected error of the VMF mechanism lacks a closed-form expression (Theorem 17 in [36]), making it complex to compute.



(a) Classical domain. **(b) Circular domain.**

Figure 16: Comparison of distribution estimation error.



(a) Classical domain. **(b) Circular domain.**

Figure 17: Comparison of mean estimation error.

as their outputs tend to average closely to the true mean. We can see that OGPM-U achieves significantly lower errors than OGPM when ε is small. In the circular domain, OGPM is unbiased, having a negligible mean estimation error. We also observe that SW-C outperforms PM-C in this large-scale domain. Statistically, OGPM's mean estimation error is 66.2% of PM-C and 55.4% of SW-C in the classical domain. In the circular domain, OGPM's mean estimation error is merely 2.3% of PM-C and 3.6% of SW-C.

8 Related Work

This paper focuses on the optimal mechanism for collecting numerical data with bounded domains under LDP, related to numerical data collection and the optimality of DP mechanisms.

Numerical Data Collection under LDP. Classical noise-adding mechanisms such as the Laplace and Gauss mechanisms [10, 11] add randomly sampled noise from a distribution to the data to achieve LDP. However, they generate outputs in an unbounded domain due to the unbounded noise distributions, rendering them unsuitable for applications requiring bounded domains [34].

For bounded data domains, the basic idea is to sample the output from a carefully designed distribution on the bounded domain. Duchi et al. randomize any data in $[-1, 1]$ to two discrete values $y \in \{-C_\varepsilon, C_\varepsilon\}$ [9], where C_ε is a constant depending on the privacy level ε. This binary-output mechanism exhibits a large randomization error as the output space is too coarse. PM [34] extends the binary output to a continuous output in $[-C_\varepsilon, C_\varepsilon]$. It designs a piecewise-based mechanism to sample the output, which uses different sampling intervals for different x, achieving a lower randomization error. Both PM and later SW [21] have shown the data utility advantage of TPM in numerical data collection with bounded domains under LDP. PTT [23] discusses the optimality of TPM. It shows that there exist TPM instantiations that yield optimal data utility, but it does not provide the closed-form mechanism. TPM is a special case of GPM using 3-piece distributions and a specific form for each piece. Meanwhile, existing TPM instantiations are designed for specific error metrics and only classical domains.

Applications of TPM. A natural application of TPM is in high-dimensional numerical data. This includes scenarios where the sensitive data may be a high-dimensional vector [34, 42], an infinite data stream [30], or matrixes [35]. Another application area is federated learning. TPM ensures LDP in the $[0, 1]$ domain, which is commonly used as the normalized domain in model training. TPM avoids the data clipping required by noise-adding mechanisms [3, 14, 22]. Another typical application is in sensors, where the data is bounded by the sensor's physical nature. This paper can replace the existing TPM to achieve better utility.

A recent work aims to design other types of bounded distributions besides TPM to achieve (L)DP [43]. They tailor sin functions and quartic functions on the bounded domain to satisfy the DP constraint. From their experimental results, the piecewise-based design is the best choice among their six instantiations for bounded domains under DP. This also indicates the advantages of piecewise-based mechanisms in the bounded domain.

Optimality of DP Mechanisms. Achieving optimal data utility is a common concern across all DP mechanisms. The staircase mechanism showed that the Laplace mechanism does not generate optimal noise [15, 32]. It samples noise from a staircase distribution, which has been shown to achieve lower error compared to the Laplace mechanism. Besides optimality in concrete error values (which is the focus of this paper), another widely focused concept is asymptotic optimality [6, 23], which studies optimal asymptotic bounds of a mechanism or a statistical estimation. Appendix B.2 discusses more detailed optimalities.

Beyond the optimality of a single DP procedure, the optimality of multifold compositions such as iterative training under DP, is studied by advanced compositions [7, 8, 20, 26]. In high-dimensional settings, the sensitivity set across dimensions also influences the optimality [39], because it affects the choice of the privacy parameter. These works either focus on different privacy constraints or discuss optimality beyond a single DP procedure, making them orthogonal to ours.

9 Conclusions

This paper presents the optimal piecewise-based mechanism for collecting numerical data with bounded domains under LDP. To find the optimal mechanism among all possible piecewise mechanisms, we generalize the existing 3-piece mechanism to an m-piece mechanism with the most general form. We proposed a framework that combines analytical proofs and off-the-shelf optimization solvers to find the optimal mechanism. Our results include the closed-form optimal piecewise mechanisms for both the classical and circular domain. Theoretical and experimental evaluations confirm the advantages of our mechanisms over existing mechanisms.

Acknowledgments

We thank the anonymous reviewers and the revision editor for their valuable feedback and guidance, which significantly improved this paper. We also acknowledge the use of GPT-4 for language refinement in this paper. This research is supported in part by the U.S. National Science Foundation under grants CNS-2245689 (CRII) and DGE-2336252, as well as by the 2022 Meta Research Award for Privacy-Enhancing Technologies.

References

[1] 2018. MotionSense Dataset : Smartphone Sensor Data - HAR. https://www.kaggle.com/datasets/malekzadeh/motionsense-dataset

[2] 2024. Gurobi Help Center. https://support.gurobi.com/hc/en-us

[3] Martín Abadi, Andy Chu, Ian J. Goodfellow, H. Brendan McMahan, Ilya Mironov, Kunal Talwar, and Li Zhang. 2016. Deep Learning with Differential Privacy. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, October 24-28, 2016*, Edgar R. Weippl, Stefan Katzenbeisser, Christopher Kruegel, Andrew C. Myers, and Shai Halevi (Eds.). ACM, 308–318. https://doi.org/10.1145/2976749.2978318

[4] Stephen P Boyd and Lieven Vandenberghe. 2004. *Convex optimization*. Cambridge university press.

[5] Hai Brenner and Kobbi Nissim. 2014. Impossibility of Differentially Private Universally Optimal Mechanisms. *SIAM J. Comput.* 43, 5 (2014), 1513–1540. https://doi.org/10.1137/110846671

[6] T. Tony Cai, Yichen Wang, and Linjun Zhang. 2019. The Cost of Privacy: Optimal Rates of Convergence for Parameter Estimation with Differential Privacy. *CoRR* abs/1902.04495 (2019). arXiv:1902.04495 http://arxiv.org/abs/1902.04495

[7] Jinshuo Dong, David Durfee, and Ryan Rogers. 2020. Optimal Differential Privacy Composition for Exponential Mechanisms. In *Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event (Proceedings of Machine Learning Research, Vol. 119)*. PMLR, 2597–2606. http://proceedings.mlr.press/v119/dong20a.html

[8] Jinshuo Dong, Aaron Roth, and Weijie J. Su. 2019. Gaussian Differential Privacy. *CoRR* abs/1905.02383 (2019). arXiv:1905.02383 http://arxiv.org/abs/1905.02383

[9] John C. Duchi, Martin J. Wainwright, and Michael I. Jordan. 2016. Minimax Optimal Procedures for Locally Private Estimation. *CoRR* abs/1604.02390 (2016). arXiv:1604.02390 http://arxiv.org/abs/1604.02390

[10] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam D. Smith. 2006. Calibrating Noise to Sensitivity in Private Data Analysis. In *Theory of Cryptography, Third Theory of Cryptography Conference, TCC 2006, New York, NY, USA, March 4-7, 2006, Proceedings (Lecture Notes in Computer Science, Vol. 3876)*, Shai Halevi and Tal Rabin (Eds.). Springer, 265–284. https://doi.org/10.1007/11681878_14

[11] Cynthia Dwork and Aaron Roth. 2014. The Algorithmic Foundations of Differential Privacy. *Found. Trends Theor. Comput. Sci.* 9, 3-4 (2014), 211–407. https://doi.org/10.1561/0400000042

[12] Natasha Fernandes, Annabelle McIver, and Carroll Morgan. 2021. The Laplace Mechanism has optimal utility for differential privacy over continuous queries. In *36th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2021, Rome, Italy, June 29 - July 2, 2021*. IEEE, 1–12. https://doi.org/10.1109/LICS52264.2021.9470718

[13] Nicholas I. Fisher. 2000. *Statistical analysis of circular data* (transferred to digital printing ed.). Cambridge University Press, Cambridge, Mass.

[14] Jie Fu, Yuan Hong, Xinpeng Ling, Leixia Wang, Xun Ran, Zhiyu Sun, Wendy Hui Wang, Zhili Chen, and Yang Cao. 2024. Differentially Private Federated Learning: A Systematic Review. *CoRR* abs/2405.08299 (2024), 36. https://doi.org/10.48550/ARXIV.2405.08299 arXiv:2405.08299

[15] Quan Geng and Pramod Viswanath. 2014. The optimal mechanism in differential privacy. In *2014 IEEE International Symposium on Information Theory, Honolulu, HI, USA, June 29 - July 4, 2014*. IEEE, 2371–2375. https://doi.org/10.1109/ISIT.2014.6875258

[16] Arpita Ghosh, Tim Roughgarden, and Mukund Sundararajan. 2009. Universally utility-maximizing privacy mechanisms. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, Bethesda, MD, USA, May 31 - June 2, 2009*, Michael Mitzenmacher (Ed.). ACM, 351–360. https://doi.org/10.1145/1536414.1536464

[17] Akshay Gupte, Shabbir Ahmed, Santanu S. Dey, and Myun-Seok Cheon. 2017. Relaxations and discretizations for the pooling problem. *J. Glob. Optim.* 67, 3 (2017), 631–669. https://doi.org/10.1007/S10898-016-0434-4

[18] Moritz Hardt and Kunal Talwar. 2010. On the geometry of differential privacy. In *Proceedings of the 42nd ACM Symposium on Theory of Computing, STOC 2010, Cambridge, Massachusetts, USA, 5-8 June 2010*, Leonard J. Schulman (Ed.). ACM, 705–714. https://doi.org/10.1145/1806689.1806786

[19] Naoise Holohan, Spiros Antonatos, Stefano Braghin, and Pól Mac Aonghusa. 2020. The Bounded Laplace Mechanism in Differential Privacy. *J. Priv. Confidentiality* 10, 1 (2020). https://doi.org/10.29012/JPC.715

[20] Peter Kairouz, Sewoong Oh, and Pramod Viswanath. 2017. The Composition Theorem for Differential Privacy. *IEEE Trans. Inf. Theory* 63, 6 (2017), 4037–4049. https://doi.org/10.1109/TIT.2017.2685505

[21] Zitao Li, Tianhao Wang, Milan Lopuhaä-Zwakenberg, Ninghui Li, and Boris Skoric. 2020. Estimating Numerical Distributions under Local Differential Privacy. In *Proceedings of the 2020 International Conference on Management of Data, SIGMOD Conference 2020, online conference [Portland, OR, USA], June 14-19, 2020*, David Maier, Rachel Pottinger, AnHai Doan, Wang-Chiew Tan, Abdussalam Alawini, and Hung Q. Ngo (Eds.). ACM, 621–635. https://doi.org/10.1145/3318464.3389700

[22] Chun Liu, Youliang Tian, Jinchuan Tang, Shuping Dang, and Gaojie Chen. 2023. A novel local differential privacy federated learning under multi-privacy regimes. *Expert Syst. Appl.* 227 (2023), 120266. https://doi.org/10.1016/J.ESWA.2023.120266

[23] Fei Ma, Renbo Zhu, and Ping Wang. 2024. PTT: Piecewise Transformation Technique for Analyzing Numerical Data under Local Differential Privacy. *IEEE Transactions on Mobile Computing* (2024), 1–13. https://doi.org/10.1109/TMC.2024.3364496

[24] Mohammad Malekzadeh, Richard G. Clegg, Andrea Cavallaro, and Hamed Haddadi. 2019. Mobile Sensor Data Anonymization. In *Proceedings of the International Conference on Internet of Things Design and Implementation* (Montreal, Quebec, Canada) *(IoTDI '19)*. ACM, New York, NY, USA, 49–58. https://doi.org/10.1145/3302505.3310068

[25] Kanti V Mardia and Peter E Jupp. 2009. *Directional statistics*. John Wiley & Sons.

[26] Jack Murtagh and Salil P. Vadhan. 2018. The Complexity of Computing the Optimal Composition of Differential Privacy. *Theory Comput.* 14, 1 (2018), 1–35. https://doi.org/10.4086/TOC.2018.V014A008

[27] Arvind Narayanan and Vitaly Shmatikov. 2008. Robust De-anonymization of Large Sparse Datasets. In *2008 IEEE Symposium on Security and Privacy (SP) 2008, 18-21 May 2008, Oakland, California, USA*. IEEE Computer Society, 111–125. https://doi.org/10.1109/SP.2008.33

[28] Nekst-Online. 2016. What Is the Pooling Problem? https://nekst-online.nl/what-is-the-pooling-problem/

[29] Ben Niu, Qinghua Li, Xiaoyan Zhu, Guohong Cao, and Hui Li. 2014. Achieving k-anonymity in privacy-aware location-based services. In *2014 IEEE Conference on Computer Communications, INFOCOM 2014, Toronto, Canada, April 27 - May 2, 2014*. IEEE, 754–762. https://doi.org/10.1109/INFOCOM.2014.6848002

[30] Xuebin Ren, Liang Shi, Weiren Yu, Shusen Yang, Cong Zhao, and Zongben Xu. 2022. LDP-IDS: Local Differential Privacy for Infinite Data Streams. In *SIGMOD '22: International Conference on Management of Data, Philadelphia, PA, USA, June 12 - 17, 2022*, Zachary G. Ives, Angela Bonifati, and Amr El Abbadi (Eds.). ACM, 1064–1077. https://doi.org/10.1145/3514221.3526190

[31] Reuven Y. Rubinstein. 1981. *Simulation and the Monte Carlo method*. Wiley. https://www.worldcat.org/oclc/07275104

[32] Jordi Soria-Comas and Josep Domingo-Ferrer. 2013. Optimal data-independent noise for differential privacy. *Inf. Sci.* 250 (2013), 200–214. https://doi.org/10.1016/j.ins.2013.07.004

[33] Han Wang, Hanbin Hong, Li Xiong, Zhan Qin, and Yuan Hong. 2022. L-SRR: Local Differential Privacy for Location-Based Services with Staircase Randomized Response. In *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security, CCS 2022, Los Angeles, CA, USA, November 7-11, 2022*, Heng Yin, Angelos Stavrou, Cas Cremers, and Elaine Shi (Eds.). ACM, 2809–2823. https://doi.org/10.1145/3548606.3560636

[34] 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 *35th IEEE International Conference on Data Engineering, ICDE 2019, Macao, China, April 8-11, 2019*. IEEE, 638–649. https://doi.org/10.1109/ICDE.2019.00063

[35] Yong Wang, Mingxing Gao, Xun Ran, Jun Ma, and Leo Yu Zhang. 2023. An improved matrix factorization with local differential privacy based on piecewise mechanism for recommendation systems. *Expert Syst. Appl.* 216 (2023), 119457. https://doi.org/10.1016/J.ESWA.2022.119457

[36] Benjamin Weggenmann and Florian Kerschbaum. 2021. Differential Privacy for Directional Data. In *CCS '21: 2021 ACM SIGSAC Conference on Computer and Communications Security, Virtual Event, Republic of Korea, November 15 - 19, 2021*, Yongdae Kim, Jong Kim, Giovanni Vigna, and Elaine Shi (Eds.). ACM, 1205–1222. https://doi.org/10.1145/3460120.3484734

[37] Wikipedia. 2024. Bauer maximum principle. https://en.wikipedia.org/w/index.php?title=Bauer_maximum_principle&oldid=1209668728

[38] Wikipedia contributors. 2023. AOL search log release — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=AOL_search_log_release&oldid=1187739761 [Online; accessed 4-January-2024].

[39] Hanshen Xiao, Jun Wan, and Srinivas Devadas. 2023. Geometry of Sensitivity: Twice Sampling and Hybrid Clipping in Differential Privacy with Optimal Gaussian Noise and Application to Deep Learning. In *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security, CCS 2023, Copenhagen, Denmark, November 26-30, 2023*, Weizhi Meng, Christian Damsgaard Jensen, Cas Cremers, and Engin Kirda (Eds.). ACM, 2636–2650. https://doi.org/10.1145/3576915.3623142

[40] Dejun Yang, Xi Fang, and Guoliang Xue. 2013. Truthful incentive mechanisms for k-anonymity location privacy. In *Proceedings of the IEEE INFOCOM 2013, Turin, Italy, April 14-19, 2013*. IEEE, 2994–3002. https://doi.org/10.1109/INFCOM.2013.6567111

[41] Wei Yu, Xinwen Fu, Steve Graham, Dong Xuan, and Wei Zhao. 2007. DSSS-Based Flow Marking Technique for Invisible Traceback. In *2007 IEEE Symposium on Security and Privacy (S&P 2007), 20-23 May 2007, Oakland, California, USA*. IEEE Computer Society, 18–32. https://doi.org/10.1109/SP.2007.14

[42] Dongyue Zhang, Weiwei Ni, Nan Fu, Lihe Hou, and Ruyu Zhang. 2023. Locally differentially private multi-dimensional data collection via haar transform. *Comput. Secur.* 130 (2023), 103291. https://doi.org/10.1016/J.COSE.2023.103291

[43] Kai Zhang, Yanjun Zhang, Ruoxi Sun, Pei-Wei Tsai, Muneeb Ul Hassan, Xin Yuan, Minhui Xue, and Jinjun Chen. 2024. Bounded and Unbiased Composite Differential Privacy. In *IEEE Symposium on Security and Privacy, SP 2024, San Francisco, CA, USA, May 19-23, 2024*. IEEE, 972–990. https://doi.org/10.1109/SP54263.2024.00108

A Proofs

Notations. Table 3 provides the notations used throughout this paper.

A.1 Proof of Lemma 3.2

Proof. We prove the inner $\max_x$ problem has a closed form. According to the definition of $\mathcal{P}_{\mathcal{M}(x)}$, it is

$$\max_x \int_{\tilde{\mathcal{D}}} \mathcal{L}(y,x)\mathcal{P}_{\mathcal{M}(x)}\mathrm{d}y = \max_x \sum_{i=1}^{m} p_i \int_{l_i}^{r_i} \mathcal{L}(y,x)\mathrm{d}y.$$

Denote $f_i(x) \coloneqq \int_{l_i}^{r_i} \mathcal{L}(y,x)\mathrm{d}y$, where $\mathcal{L}(y,x) = |y-x|^p$. First, we prove that $f_i(x)$ is a convex function w.r.t. x. Specifically, based on the relationship between x and $[l_i, r_i)$, the value of x is split into three cases: (i) $x \in [a, l_i)$, (ii) $x \in [l_i, r_i)$, and (iii) $x \in [r_i, b)$. We prove the second derivative of $f_i(x)$ w.r.t. x is non-negative in each case, thus $f_i(x)$ is convex.

Case (i): $x \in [a, l_i)$. The integral is:

$$f_i(x) = \int_{l_i}^{r_i} |y-x|^p \mathrm{d}y = \int_{l_i}^{r_i} (y-x)^p \mathrm{d}y = \frac{(r_i-x)^{p+1} - (l_i-x)^{p+1}}{p+1}.$$

The second derivative w.r.t. x is

$$\frac{\partial^2}{\partial x^2} f_i(x) = p(r_i-x)^{p-1} - p(l_i-x)^{p-1} \geq 0.$$

The inequality holds because $(r_i-x)^{p-1} \geq (l_i-x)^{p-1}$ for $x \in [a, l_i)$.

Case (ii): $x \in [l_i, r_i)$. The integral is

$$f_i(x) = \int_{l_i}^{r_i} |y-x|^p \mathrm{d}y = \int_{l_i}^{x} (x-y)^p \mathrm{d}y + \int_{x}^{r_i} (y-x)^p \mathrm{d}y = \frac{(x-l_i)^{p+1}}{p+1} + \frac{(r_i-x)^{p+1}}{p+1}.$$

The second derivative w.r.t. x is

$$\frac{\partial^2}{\partial x^2} f_i(x) = p(x-l_i)^{p-1} + p(r_i-x)^{p-1} \geq 0.$$

The inequality holds because both $(x-l_i)^{p-1}$ and $(r_i-x)^{p-1}$ are non-negative for $x \in [l_i, r_i)$.

Table 3: Notations

Symbol	Description
x	Sensitive input (from raw data)
y	Randomized output
$\mathcal{D}$	Input domain
$\tilde{\mathcal{D}}$	Output domain
$pdf[\cdot]$	Probability density function
$\mathcal{P}_{\mathcal{M}(x)}$	Probability density function of $\mathcal{M}(x)$
p_ε	Sampling probability w.r.t. ε
$[l_{x,\varepsilon}, r_{x,\varepsilon})$	Sampling interval w.r.t. x and ε

Case (iii): $x \in [r_i, b)$. The integral is

$$f_i(x) = \int_{l_i}^{r_i} |y-x|^p \mathrm{d}y = \int_{l_i}^{r_i} (x-y)^p \mathrm{d}y = \frac{(x-l_i)^{p+1} - (x-r_i)^{p+1}}{p+1}.$$

The second derivative w.r.t. x is

$$\frac{\partial^2}{\partial x^2} f_i(x) = p(x-l_i)^{p-1} - p(x-r_i)^{p-1} \geq 0.$$

The inequality holds because $(x-l_i)^{p-1} \geq (x-r_i)^{p-1}$ for $x \in [r_i, b)$.

The above three cases show that the second derivative of $f_i(x)$ w.r.t. $x \in [a, b)$ is non-negative. Thus, the non-negative weighted sum $\sum_{i=1}^{m} p_i f_i(x)$ is also convex w.r.t. x [4]. According to the Bauer maximum principle [37]: any function that is convex attains its maximum at some extreme points of set. This means that the optimal x is achieved at the endpoints of $x \in \mathcal{D}$, i.e. $x = a$ or $x = b$. Therefore, we have

$$\max_x \int_{\tilde{\mathcal{D}}} \mathcal{L}(y,x)\mathcal{P}_{\mathcal{M}(x)}\mathrm{d}y = \max_{\{a,b\}} \int_{\tilde{\mathcal{D}}} \mathcal{L}(y,x)\mathcal{P}_{\mathcal{M}(x)}\mathrm{d}y,$$

which completes the proof. □

Remark: This lemma can be empirically validated by the whole-domain error plots in Figure 8 and Figure 10, where the maximum of the whole-domain error is achieved at the endpoints.

A.2 Proof of Lemma 3.3

Proof. The optimal $(m+1)$-piecewise mechanism and the optimal m-piecewise mechanism may superficially differ due to the extra piece. Thus, we define a piece-merging operation to merge the redundant pieces. We will show that if the optimal $(m+1)$-piecewise mechanism is the same as the optimal m-piecewise mechanism after merging redundant pieces, then increasing m does not decrease the optimal error, i.e. the optimal piece number is m.

Assume the optimal m-piecewise mechanism is determined by the tuple set

$$S_m = \{(p_i, l_i, r_i) : i \in [m]\}.$$

To merge redundant pieces, we define a piece-merging operation:

$$(p_i, l_i, r_i) \uplus (p_j, l_j, r_j) \coloneqq \begin{cases} (p_i, l_i, r_j) & \text{if } p_i = p_j \text{ and } i+1 = j, \\ \{(p_i, l_i, r_i), (p_j, l_j, r_j)\} & \text{otherwise.} \end{cases}$$

Because the optimal $(m+1)$-piecewise mechanism is the same as the m-piecewise mechanism, it follows that

$$\uplus_{i,j=1}^{m+1} S_{m+1} = \uplus_{i,j=1}^{m} S_m,$$

where $\uplus_{i,j=1}^{m} S_m$ merges all consecutive pieces with the same p. Denote the merged optimal m-piecewise mechanism as $\uplus_{i,j=1}^{m} S_m := S_m^{\uplus}$ and the piece number as $|S_m^{\uplus}| = m^*$. Because both sides of the above equation are optimal, this means that if (p_k, l_k, r_k) is an arbitrary piece in the optimal $(m+1)$-piecewise mechanism, then merging it with $S_m^{\uplus}$ remains $S_m^{\uplus}$, i.e.

$$(p_k, l_k, r_k) \uplus_{i=1}^{m^*} S_m^{\uplus} = S_m^{\uplus}.$$

This premise indicates that there does not exist a piece (p_k, l_k, r_k) besides $S_m^{\uplus}$ lowers the error.

Without loss of generality, we can consider the optimal $(m+2)$-piecewise mechanism, which allows an extra piece besides the optimal $(m+1)$-piecewise mechanism. We claim that the extra piece is still captured by $S_m^{\uplus}$. The key insight is: adding an extra optimal piece to the optimal $(m+1)$-piecewise mechanism is the *same* as adding it to the optimal m-piecewise mechanism, because the optimal $(m+1)$-piecewise mechanism is the same as the optimal m-piecewise mechanism.

Since adding an extra optimal piece to the optimal m-piecewise mechanism remains $S_m^{\uplus}$, then for any $k \in [m+2]$, merging piece (p_k, l_k, r_k) in the optimal $(m+2)$-piecewise mechanism remains an m-piecewise mechanism $S_m^{\uplus}$.

For $m+3$ or more, it follows the same logic. Adding an arbitrary piece is equivalent to adding it to the optimal $(m+2)$-piecewise mechanism, which is the same as adding it to the optimal m-piecewise mechanism. Thus, the optimal m-piecewise mechanism is the same as the optimal $(m+3)$-piecewise mechanism after merging redundant pieces, and so on. □

Remark: Intuitively, the extra pieces (of the optimal $(m+1)$-piecewise mechanism and beyond) is similar to the redundant variables in optimization theory: adding more non-negative variables to a minimization objective does not decrease the *optimal* value. Here the error from each piece is a variable, and it is non-negative, which leads to the same conclusion: adding more pieces (one or more) to the optimal m-piecewise mechanism does not decrease the optimal error.

This lemma means that we can determine the optimal m-piecewise mechanism for $m = 1, 2, \ldots$, until the optimal $(m+1)$-piecewise mechanism is identical to the optimal m-piecewise mechanism for all x and ε. This statement can be empirically validated by attempting to find counterexamples using larger m than the optimal m. The source code of our framework provides scripts and results to empirically validate this lemma.

A.3 Proof of Theorem 3.6

Proof. **Privacy invariant:** For any input $v, v' \in \mathcal{D}'$ and any output $y \in \tilde{\mathcal{D}}'$:

$$\frac{pdf\left[\mathcal{M}'(v) = y\right]}{pdf\left[\mathcal{M}'(v') = y\right]} \le \frac{p}{c} \div \frac{p}{\exp(\varepsilon)c} = \exp(\varepsilon).$$

Utility invariant: For any $x' = cx + d \in \mathcal{D}'$, we can calculate the error difference between $\mathcal{M}'(x')$ and $\mathcal{M}'_{\text{bad}}(x')$ as follows:

$$\begin{aligned} &Err(x', \mathcal{M}') - Err(x', \mathcal{M}'_{\text{bad}}) \\ =&Err(cx+d, \mathcal{M}') - Err(cx+d, \mathcal{M}'_{\text{bad}}) \\ =&\int_{\tilde{\mathcal{D}}'} \mathcal{L}(y, cx+d) \left(\mathcal{P}_{\mathcal{M}'(cx+d)} - \mathcal{P}_{\mathcal{M}'_{\text{bad}}(cx+d)}\right) \mathrm{d}y. \end{aligned}$$

Let $y_t = (y-d)/c$, then $\mathrm{d}y = c\mathrm{d}y_t$ and $y = cy_t + d$, where $y_t \in \tilde{\mathcal{D}}$. The above equation is equivalent to

$$\begin{aligned} &Err(x', \mathcal{M}') - Err(x', \mathcal{M}'_{\text{bad}}) \\ =&\int_{\tilde{\mathcal{D}}} \mathcal{L}(cy_t + d, cx+d) \left(\mathcal{P}_{\mathcal{M}'(cx+d)} - \mathcal{P}_{\mathcal{M}'_{\text{bad}}(cx+d)}\right) c\mathrm{d}y_t \\ =&\int_{\tilde{\mathcal{D}}} \mathcal{L}(cy_t + d, cx+d) \frac{1}{c} \left(\mathcal{P}_{\mathcal{M}(x)} - \mathcal{P}_{\mathcal{M}_{\text{bad}}(x)}\right) c\mathrm{d}y_t. \end{aligned}$$

The last equality holds due to the definition of $\mathcal{T} : \mathcal{M} \to \mathcal{M}'$. For L_p-similar error metric $\mathcal{L}$ (i.e. $\mathcal{L}(y, x) := |y - x|^p$), it follows that

$$\mathcal{L}(cy_t + d, cx + d) = \mathcal{L}(cy_t, cx) = c^p \mathcal{L}(y_t, x).$$

Thus, the above difference of Err is equivalent to

$$\begin{aligned} &Err(x', \mathcal{M}') - Err(x', \mathcal{M}'_{\text{bad}}) \\ =&c^p \int_{\tilde{\mathcal{D}}} \mathcal{L}(y_t, x) \left(\mathcal{M}(x) - \mathcal{M}_{\text{bad}}(x)\right) \mathrm{d}y_t \\ =&c^p \left(Err(x, \mathcal{M}) - Err(x, \mathcal{M}_{\text{bad}})\right) \le 0, \end{aligned}$$

due to the known fact $c > 0$ and $Err(x, \mathcal{M}) - Err(x, \mathcal{M}_{\text{bad}}) \le 0$. □

Remark: Intuitively, this theorem is to prove: if $\mathcal{M}$ is a better mechanism than $\mathcal{M}_{\text{bad}}$ on $\mathcal{D}$, then it is still a better mechanism than $\mathcal{M}_{\text{bad}}$ after linearly mapping their outputs to $\mathcal{D}'$.

A.4 Proof of Theorem 3.7

Proof. Appendix B.4 provides the formalized procedure. Following this procedure, we show the optimal GPM under $\mathcal{D} \to \tilde{\mathcal{D}} = [0, 1) \to [0, 1)$ and $\mathcal{L}(y, x) = |y - x|$.

The variables in TPM are p, l, and r. Since it is a family of probability distributions, the normalization constraint is

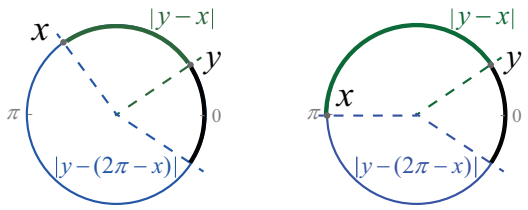
$$(r - l) \cdot p + (1 - (r - l)) \cdot p/\exp(\varepsilon) = 1,$$

which means the length of the central piece is

$$s := r - l = \frac{\exp(\varepsilon) - p}{p(\exp(\varepsilon) - 1)}.$$

Without loss of generality, assume $x = 0$ is the optimal point ($x = 1$ is symmetric). The optimization problem for solving the optimal p is

$$\begin{aligned} &\arg\min_p \left(\int_0^s y \cdot p \, \mathrm{d}y + \int_s^1 y \cdot \frac{p}{\exp(\varepsilon)} \, \mathrm{d}y\right) \\ =&\arg\min_p \left(\frac{s^2}{2}\left(p - \frac{p}{\exp(\varepsilon)}\right) + \frac{1}{2}\frac{p}{\exp(\varepsilon)}\right) \\ =&\arg\min_p \frac{1}{2}\left(\frac{(\exp(\varepsilon) - p)^2}{p(\exp(\varepsilon) - 1)\exp(\varepsilon)} + \frac{p}{\exp(\varepsilon)}\right). \end{aligned}$$



(a) $x = 3\pi/4$. **(b)** $x = \pi$.

Figure 18: Examples of $\mathcal{L}_{\text{mod}}(y, x)$ w.r.t. x. Given a specific y, the shorter between the blue and green arcs is $\mathcal{L}_{\text{mod}}(y, x)$. $\max_x \mathcal{L}_{\text{mod}}(y, x)$ is achieved at $x = \pi$.

To solve the optimal p, we take the first-order derivative w.r.t. p and set it to 0, i.e.

$$\frac{\partial}{\partial p}\left(\frac{(\exp(\varepsilon)-p)^2}{p(\exp(\varepsilon)-1)\exp(\varepsilon)}+\frac{p}{\exp(\varepsilon)}\right)=0,$$

This leads to $p = \exp(\varepsilon/2)$. Then

$$s=\frac{\exp(\varepsilon/2)-1}{\exp(\varepsilon)-1}.$$

Having solved p and s, the optimal r is $r = l + s$. Then the optimal l is determined by

$$\arg\min_l \int_0^l (x-y)\cdot\frac{p}{\exp(\varepsilon)}\,\mathrm{d}y+\int_l^x (x-y)\cdot p\,\mathrm{d}y+ \int_x^{l+s}(y-x)\cdot p\,\mathrm{d}y+\int_{l+s}^1 (y-x)\cdot\frac{p}{\exp(\varepsilon)}\,\mathrm{d}y$$

This is a univariate optimization problem w.r.t. l. Moreover, it is a two-order polynomial w.r.t. l and can be solved by analyzing the first-order and second-order derivatives. The solution is

$$l=\frac{2x(p-pe^{-\varepsilon})-s(p-pe^{-\varepsilon})}{2(p-pe^{-\varepsilon})}=x-\frac{s}{2}.$$

Note that when $x - s/2 < 0$, the above l is outside the domain $[0, 1)$. In this case, the optimal l is $l = 0$.

Relating the above deduction to Theorem 3.7, the term $s/2$ corresponds to C. Then the optimal p is $\exp(\varepsilon/2)$, $l = x - C$, and $r = x + C$ when $x \in [C, 1-C)$, which completes the proof for $\mathcal{L}(y, x) = |y - x|$. The proof for $\mathcal{L}(y, x) = |y - x|^2$ is similar. □

Remark: This proof is the same as finding the optimal 3-piecewise distribution in domain $[0, 1)$. The source code of our framework provides the validation.

A.5 Proof of Lemma 4.1

Proof. Note that $\mathcal{L}_{\text{mod}}(y, x) = \min(\mathcal{L}(y, x), \mathcal{L}(y, 2\pi - x))$. The key observation is that for any y and L_p-similar distance metric $\mathcal{L}$, we have

$$\max_x \mathcal{L}_{\text{mod}}(y, x) = \mathcal{L}_{\text{mod}}(y, \pi) = \mathcal{L}(y, \pi).$$

Figure 18 illustrates the intuition. For any fixed y, it compares the length of $|y - x|$ and $|y - (2\pi - x)|$ w.r.t $x \in [0, 2\pi)$. $\mathcal{L}_{\text{mod}}(y, x)$ is determined by the minimum of the two, and $\max_x \mathcal{L}_{\text{mod}}(y, x)$ will achieved at $x = \pi$.

The remained proof is more intuitive than the proof of Lemma 3.2, as the inner integrand $\mathcal{L}_{\text{mod}}(y, x)$ has a *unique* maximum at $x = \pi$ for any $y \in [l_i, r_i)$, making the swap of $\max_x$ and integration valid. Specifically, we have

$$\max_x \sum_i^m p_i \int_{l_i}^{r_i} \mathcal{L}_{\text{mod}}(y,x)\mathrm{d}y = \sum_i^m p_i \int_{l_i}^{r_i} \max_x \mathcal{L}_{\text{mod}}(y,x)\mathrm{d}y = \sum_i^m p_i \int_{l_i}^{r_i} \mathcal{L}(y,\pi)\mathrm{d}y$$

holds trivially, because all other x values always result in smaller $\mathcal{L}_{\text{mod}}(y, x)$. Therefore, for any values of $p_i \geq 0$, $l_i \leq r_i$, the integration of $\mathcal{L}_{\text{mod}}(y, x)$ is bounded by the integration of $\mathcal{L}(y, \pi)$. □

A.6 Proof of Theorem 5.1

Proof. We prove the unbiasedness of $\mathcal{M}$. The expectation of the given $\mathcal{M}$ is

$$\begin{aligned}\mathrm{E}[\mathcal{M}(x)] &= \int_{-C}^{C+1} y\cdot\mathcal{P}_{\mathcal{M}(x)}\mathrm{d}y\\ &= \int_{-C}^{l} y\frac{p}{\exp(\varepsilon)}\mathrm{d}y+\int_l^r yp\mathrm{d}y+\int_r^{C+1} y\frac{p}{\exp(\varepsilon)}\mathrm{d}y\\ &= \frac{r^2-l^2}{2}\left(p-\frac{p}{\exp(\varepsilon)}\right)+\left(C+\frac{1}{2}\right)\frac{p}{\exp(\varepsilon)}.\end{aligned}$$

Denote $s := (2C + 1)(C - 1)/(2C)$, which rewrites l and r as

$$l=\frac{C+1}{2}x-\frac{C-1}{4}-\frac{s}{2},\ r=\frac{C+1}{2}x-\frac{C-1}{4}+\frac{s}{2}.$$

Then the above $\mathrm{E}[\mathcal{M}(x)]$ is equivalent to

$$\begin{aligned}&\frac{(C+1)sx-(C-1)s/2}{2}\cdot\frac{4C}{(C^2-1)(2C+1)}+\frac{\exp(-\varepsilon/2)}{2}\\ =&x-\frac{C-1}{2(C+1)}+\frac{\exp(-\varepsilon/2)}{2}=x,\end{aligned}$$

leading to $\mathrm{E}[\mathcal{M}] = x$, i.e. $\mathcal{M}$ is unbiased. □

B Complementary Materials

B.1 Detailed Comparison with PTT (Section 2.3)

Piecewise transformation technique (PTT) [23] is a framework for 3-piecewise mechanisms. It shows that (i) many PTT mechanisms are asymptotically optimal when used to obtain an unbiased estimator for mean of numerical data, and (ii) there is a PTT that reaches the theoretical lower bound on variance.

Under the viewpoint of this paper, type-I PTT focuses on TPM that constrains the probabilities of the central interval (p) and the two side intervals (q) as

$$p=\frac{1}{2ak}\frac{\exp(\varepsilon)}{\exp(\varepsilon)-1},\quad q=\frac{\exp(\varepsilon)}{k(\exp(\varepsilon)-1)},$$

where a and k are parameters to be determined. Type-II PTT focuses on TPM that constrains p and q as

$$p=\frac{1}{ak}\frac{\exp(\varepsilon)}{\exp(\varepsilon)-1},\quad q=\frac{\exp(\varepsilon)}{k(\exp(\varepsilon)-1)}.$$

Therefore, PTT is still a specific case of the TPM framework. Additionally, when discussing optimality of PTT, it gives a value of a for type-I PTT but do not provide the optimal k.

B.2 Related Optimality (Section 3.1)

In this paper, the optimality of GPM is defined with respect to: (i) the worst-case L_p-similar error metric, (ii) bounded numerical domains $\mathcal{D} \to \tilde{\mathcal{D}}$ and mechanisms based on piecewise distributions, (iii) minimization of error value (not asymptotic or order-of-magnitude optimality), and (iv) without post-processing. L_p-similar error metrics are natural choices for evaluating data utility [15, 18, 34]. Bounded numerical domains are common in real-world applications. Focusing on error values allows for more precise comparisons between different mechanisms. By excluding post-processing, we can analyze the optimality of the mechanism itself, which provides a more fundamental understanding than considering the mechanism combined with a specific post-processing.

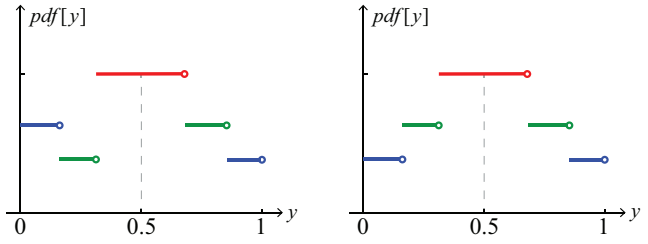
Other types of optimality have been explored in the literature, particularly for variants of Laplace mechanisms. The staircase mechanism [15] adopts the same utility model without prior knowledge or post-processing as this paper. It claims optimality under specific assumptions, one of which is that a staircase (piecewise) distribution *can* achieve the optimal error. The mechanism demonstrates better L_1-error performance than the Laplace mechanism on $\tilde{\mathcal{D}} = (-\infty, \infty)$, and its asymptotic optimality has been formally proven. *Universal optimality* is another type of optimality, defined from the perspective of a user's prior knowledge and post-processing ability [16]. In this utility model, the user observes the output of the mechanism and selects another value based on the output and their prior knowledge, i.e. under a Bayesian utility framework. Formally, if the user's prior is denoted as p_i on the data domain $i \in N$ (i.e. a discrete domain) and the user's post-processing is represented as a remap $z_{i,j}$ that reinterprets the output of the mechanism (on the sensitive value i) to j, then the utility model is defined as

$$Err(i) = \sum_{i \in N} p_i \sum_{j \in N} z_{i,j} \cdot \mathcal{L}(i, j).$$

This utility model incorporates the user's prior knowledge and post-processing ability. A mechanism is called universally optimal if, for any prior p_i, there exists an optimal remap $z_{i,j}$. Under this utility model, it was proven that the truncated geometric mechanism (a discretized version of the Laplace mechanism) can achieve universal optimality for count queries[§§] and a legal error metric $\mathcal{L}(i, j)$. Such universal optimality was shown to be unachievable for more complex queries [5]. Under the same utility model, the universal optimality was extended to the truncated Laplace mechanism for a bounded numerical domain $\mathcal{D} = [0, 1]$ by approximating the geometric mechanism with the Laplace mechanism and post-processing [12].

These results do not hold in our utility model, i.e. utility model without prior and post-processing. Figure 13 has shown that OGPM generally has a smaller error than the truncated Laplace mechanism, especially when the privacy parameter ε is not small, indicating the sub-optimality of the truncated Laplace mechanism in the absence of using prior and post-processing.

[§§]This is in the centralized DP setting, where the data curator holds the dataset and uses *one* mechanism.



(a) Non-staircase distribution. **(b) Staircase distribution.**

Figure 19: A non-staircase distribution (left) can always be shifted into a staircase distribution (right) by moving some pieces closer to x, which reduces the error.

B.3 Directions for Analytically Proving Optimal $m = 3$ (Section 3.2.3)

This appendix outlines two potential directions for analytically proving that the optimal m is 3, along with the challenges associated with each approach.

Mathematically, finding the optimal m-piecewise mechanism is equivalent to identifying the optimal m-piecewise distribution under an L_p-similar error metric. It is seemingly true that the optimal m is 3: if the optimal m-piecewise distribution is not 3 but 4 or more, we can always shift the probability mass from the two side intervals (i.e. other pieces) to the central interval, thereby reducing the error. At the very least, the following fact holds:

Fact B.1. *The optimal m-piecewise distribution has a strict staircase shape, i.e. the probability density of the central interval is greater than that of the two side intervals.*

Figure 19 illustrates this fact. Moving pieces while keeping their probabilities unchanged clearly maintains both the ε-LDP constraint and the probability normalization constraint. This observation reduces the problem to proving that a 3-staircase distribution can achieve the same optimal error as a 4-staircase distribution under the ε-LDP and probability normalization constraints.

Direction 1: If we can further move the green piece in Figure 19b "into" the red central piece while keeping the probabilities of the red and blue pieces unchanged, i.e. transform it into a 3-staircase distribution while ensuring a decrease in the error, then we can prove that the optimal m is 3. However, this is challenging, as it breaks the probability normalization constraint (i.e. the sum of probabilities is no longer 1), requiring adjustments to the probabilities of each piece to satisfy the ε-LDP constraint. The difficulty lies in ensuring that these adjustments will indeed decrease the error.

Direction 2: Another approach is to formulate the problems for 3-staircase and 4-staircase distributions as two constrained optimization problems. The goal would be to prove that the optimal error of the 3-staircase distribution is equivalent to that of the 4-staircase distribution. Ideally, these two multi-variable optimization problems could be solved analytically, resulting in two closed-form error expressions w.r.t. x and ε, thereby completing the proof for any x and ε by showing that the two expressions are equal. This direction aligns with our framework. However, the challenge lies in the complexity of solving such multi-variable optimization problems

analytically. This is why we rely on an off-the-shelf optimization solver, which, while effective, only provides numerical solutions for specific x and ε values.

B.4 Analytical Deduction for the Optimal GPM (Section 3.2.3)

If the optimal GPM under $\mathcal{L}$ is proved to be TPM, then the closed-form optimal can be derived by deduction.

Denote $\mathcal{D} = [a, b)$ and $\tilde{\mathcal{D}} = [\tilde{a}, \tilde{b})$. Notice that the normalization constraint of probability is

$$(r - l) \cdot p + [(b - a) - (r - l)] \cdot p/\exp(\varepsilon) = 1.$$

This means the central interval length is

$$s \coloneqq r - l = \frac{1}{p(1 - \exp(-\varepsilon))} - \frac{\tilde{b} - \tilde{a}}{\exp(\varepsilon) - 1}.$$

If the minimal worst-case error is achieved at $x = a$ in Lemma 3.2, then solving for the optimal p is reduced to

$$\arg\min_{p} \left(\int_{\tilde{a}}^{s} \mathcal{L}(y, a) p \, \mathrm{d}y + \int_{s}^{\tilde{b}} \mathcal{L}(y, a) \frac{p}{\exp(\varepsilon)} \mathrm{d}y \right),$$

which is a univariate optimization problem w.r.t. p and can be solved analytically. With the solved p, Formulation (3) is also reduced to a univariate optimization problem w.r.t. l:

$$\arg\min_{l} \left(\int_{\tilde{a}}^{l} P_1 \, \mathrm{d}y + \int_{l}^{l+s} P_2 \, \mathrm{d}y + \int_{l+s}^{\tilde{b}} P_1 \, \mathrm{d}y \right),$$

where $P_1 = \mathcal{L}(y, x)p$ and $P_2 = \mathcal{L}(y, x)p/\exp(\varepsilon)$. This univariate optimization problem solves the optimal l, and the optimal r is $r = l + s$. Note that l and r should be restricted in $[\tilde{a}, \tilde{b})$ when analyzing the first-order derivative.

B.5 MSE of the Optimal GPM (Section 3.3)

Denote p_ε and C as the same as the instantiations of $\mathcal{M}$ in Theorem 3.7. The MSE of $\mathcal{M}$ is

(1) If $x \in [0, C)$:

$$\frac{p_\varepsilon}{3}\left((2C - x)^3 + x^3\right) + \frac{p_\varepsilon}{3\exp(\varepsilon)}\left((1 - x)^3 - (2C - x)^3\right).$$

(2) If $x \in [C, 1 - C)$:

$$\frac{p_\varepsilon}{3\exp(\varepsilon)}\left(-2C^3 + 3x^2 - 3x + 1\right) + \frac{p_\varepsilon}{3}\left(2C^3\right).$$

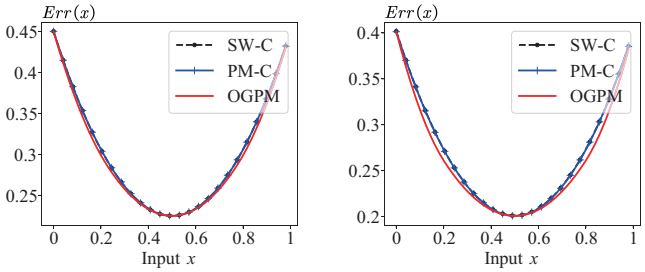
(3) If $x \in [1 - C, 1)$:

$$\frac{p_\varepsilon}{3\exp(\varepsilon)}\left((1 - 2C - x)^3 + x^3\right) + \frac{p_\varepsilon}{3}\left((1 - x)^3 - (1 - 2C - x)^3\right).$$

For example, when $x = 0$, the MSE of $\mathcal{M}$ is

$$\mathrm{MSE}[\mathcal{M}(0)] = \frac{p_\varepsilon}{3}\left(8C^3\right) + \frac{p_\varepsilon}{3\exp(\varepsilon)}\left(1 - 8C^3\right).$$

Setting $\varepsilon = 1$ results in $\mathrm{MSE}[\mathcal{M}(0)] = 0.22$ of OGPM. As a comparison, SW [21], which also designed for $\mathcal{D} = [0, 1)$, has an MSE of 0.29 at $x = 0$.



(a) Privacy parameter $\varepsilon = 0.4$. **(b) Privacy parameter $\varepsilon = 0.8$.**

Figure 20: Whole-domain error comparison in the classical domain with error metric $\mathcal{L} = |y - x|$.

B.6 Optimal Assignment of Privacy Parameter and an Example (Section 6)

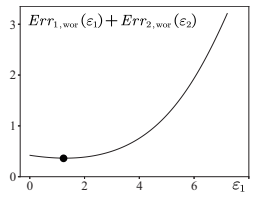
The objective function to minimize the given total error in 2D polar coordinates is

$$\min_{\varepsilon_1, \varepsilon_2} Err_{1,\mathrm{wor}}(\varepsilon_1) + Err_{2,\mathrm{wor}}(\varepsilon_2),$$

where $Err_{1,\mathrm{wor}}(\varepsilon_1)$ and $Err_{2,\mathrm{wor}}(\varepsilon_2)$ are the worst-case errors of the classical domain and the circular domain, respectively.

Without loss of generality, we can assume the polar coordinate data is in $[0, 1) \times [0, 2\pi)$ and $\mathcal{L} = |y_1 - x_1|^2$, then $Err_{1,\mathrm{wor}}(\varepsilon_1)$ and $Err_{2,\mathrm{wor}}(\varepsilon_2)$ are already given by our MSE analysis. However, the above optimization problem as a function of ε_1 and ε_2 is generally non-linear, thus hard to be analytically solved. Therefore, a simple and practical way to find the optimal ε_1 and ε_2 is numerical testing.

Under distance metric $\mathcal{L}(y, x) = |y - x|^2$, the worst-case error of the classical domain $[0, 1)$ is achieved at $x = 0$. Therefore, $Err_{1,\mathrm{wor}}$ equals to $\mathrm{MSE}[\mathcal{M}(0)]$ calculated before. For the circular domain $[0, 2\pi)$, the worst-case error $Err_{2,\mathrm{wor}}$ is stated in Theorem 4.4. If $\varepsilon = 1 + 2\pi$ and we assign ε_1 to the classical domain and $\varepsilon_2 = \varepsilon - \varepsilon_1$ to the circular domain, then the total error is plotted in the right figure. In this figure, the optimal assignment is $\varepsilon_1 = 1.32$ and $\varepsilon_2 = 5.69$.



From the curve of the total error, we can see that ε_2 affects the total error more than ε_1. Even if ε_1 is set to 0, the total error is not significantly affected, and it is still is significantly smaller than the case of $\varepsilon_2 = 0$. This is because the circular domain has a larger range than the classical domain, thus the error of the circular domain is more sensitive to the privacy parameter.

Note that the optimality for the polar coordinate data is under the specific error metric $\mathcal{L}_{2\mathrm{D}} \coloneqq \mathcal{L}(y_1, x_1) + \mathcal{L}_{\mathrm{mod}}(y_2, x_2)$. If the error metric differs, the optimal error might not be preserved.

B.7 Comparison under Small ε (Section 7.1.1)

Figure 20 presents the whole-domain error comparison of OGPM, PM-C, and SW-C under smaller ε values, specifically $\varepsilon = 0.4$ and $\varepsilon = 0.8$. In these scenarios, all three mechanisms approach the uniform distribution more closely compared to cases with larger

ε. Consequently, their errors are also more similar to each other. Statistically, when $\varepsilon = 0.4$, the error of OGPM is at most 0.008 smaller than that of PM-C and SW-C. For $\varepsilon = 0.8$, the error of OGPM is at most 0.015 smaller than that of PM-C and SW-C.

B.8 Expected Error of the B-Laplace Mechanism (Section 7.1.3)

The B-Laplace mechanism redefines a Laplace-shaped distribution on a bounded domain as the perturbation mechanism. For the data domain $\mathcal{D} \to \tilde{\mathcal{D}} = [0, 1) \to [0, 1)$, the B-Laplace mechanism is defined as follows:

Definition B.1 (Bounded Laplace Mechanism, adapted from [19]). The B-Laplace mechanism $\mathcal{M}(x) : [0, 1) \to [0, 1)$ is given by the probability density function (PDF) as follows:

$$pdf[\mathcal{M}(x) = y] = \frac{1}{C_y} \cdot \frac{1}{2b} \exp\left(-\frac{|y - x|}{b}\right) \quad \forall y \in [0, 1),$$

where b is the scale parameter, and $C_y = \int_0^1 \frac{1}{2b} \exp\left(-\frac{|y-x|}{b}\right) \mathrm{d}x$ is the normalization constant.

According to Theorem 3.5 and Corollary 4.5 in [19], the B-Laplace mechanism satisfies ε-LDP whenever $b \geq 1/\varepsilon$. Using the best scale parameter $b = 1/\varepsilon$, the normalization constant becomes $C_y = (1 - \exp(-\varepsilon))/2$. We can compute the expected L_1 error of the B-Laplace mechanism as follows (this computation is not included in [19]):

$$\begin{aligned} Err(x, \mathcal{M}) &= \int_0^1 |y - x| \cdot pdf[\mathcal{M}(x) = y] \mathrm{d}y \\ &= \int_0^1 |y - x| \cdot \frac{1}{C_y} \cdot \frac{1}{2b} \exp\left(-\frac{|y - x|}{b}\right) \mathrm{d}y \\ &= \frac{\varepsilon}{1 - \exp(-\varepsilon)} \int_0^1 |y - x| \exp\left(-\varepsilon |y - x|\right) \mathrm{d}y. \end{aligned}$$

The above integral can be numerically computed using the Python library function `scipy.stats.laplace.expect()` or analytically solved. The final result for the expected error is

$$\frac{2 - (1 + \varepsilon x)e^{-\varepsilon x} - (1 + \varepsilon(1 - x))e^{-\varepsilon(1-x)}}{\varepsilon(1 - e^{-\varepsilon})},$$

which is a closed-form expression w.r.t. x and ε.