

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

Authors: [Ye Zheng](#), Sumita Mishra, Yidan Hu



LDP Mechanisms

- Randomization algorithm $\mathcal{M}: \mathcal{D} \rightarrow \tilde{\mathcal{D}}$
 - provide quantifiable privacy for data $x \in \mathcal{D}$

$$\forall x_1, x_2 \in \mathcal{D}, \forall y \in \tilde{\mathcal{D}} \quad \max \frac{\Pr[\mathcal{M}(x_1) = y]}{\Pr[\mathcal{M}(x_2) = y]} \leq e^\epsilon$$

Distinguishability of x_1 and x_2 (sensitive data)
from y (randomized data)

LDP Mechanisms

- Randomization algorithm $\mathcal{M}: \mathcal{D} \rightarrow \tilde{\mathcal{D}}$
 - provide quantifiable privacy for data $x \in \mathcal{D}$

$$\forall x_1, x_2 \in \mathcal{D}, \forall y \in \tilde{\mathcal{D}} \quad \max \frac{\Pr[\mathcal{M}(x_1) = y]}{\Pr[\mathcal{M}(x_2) = y]} \leq e^\epsilon$$

Distinguishability of x_1 and x_2 (sensitive data)
from y (randomized data)

Privacy

$x_1 \rightarrow \mathcal{M} \rightarrow y$



Provable defense against
data inference attack

LDP Mechanisms

- Randomization algorithm $\mathcal{M}: \mathcal{D} \rightarrow \tilde{\mathcal{D}}$
 - provide quantifiable privacy for data $x \in \mathcal{D}$

$$\forall x_1, x_2 \in \mathcal{D}, \forall y \in \tilde{\mathcal{D}} \quad \max \frac{\Pr[\mathcal{M}(x_1) = y]}{\Pr[\mathcal{M}(x_2) = y]} \leq e^\epsilon$$

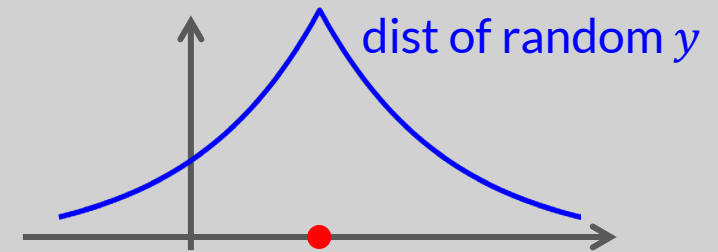
Privacy

$x_1 \rightarrow \mathcal{M} \rightarrow y$



Provable defense against
data inference attack

Data utility



LDP Mechanisms

- Randomization algorithm $\mathcal{M}: \mathcal{D} \rightarrow \tilde{\mathcal{D}}$
 - provide quantifiable privacy for data $x \in \mathcal{D}$

$$\forall x_1, x_2 \in \mathcal{D}, \forall y \in \tilde{\mathcal{D}} \quad \max \frac{\Pr[\mathcal{M}(x_1) = y]}{\Pr[\mathcal{M}(x_2) = y]} \leq e^\epsilon$$

Privacy

quantified by ϵ

$x_1 \rightarrow \mathcal{M} \rightarrow y$

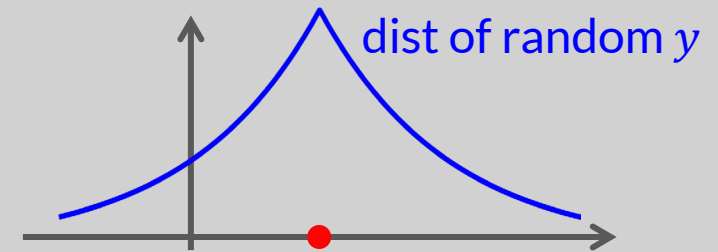


Provable defense against
data inference attack



Data utility

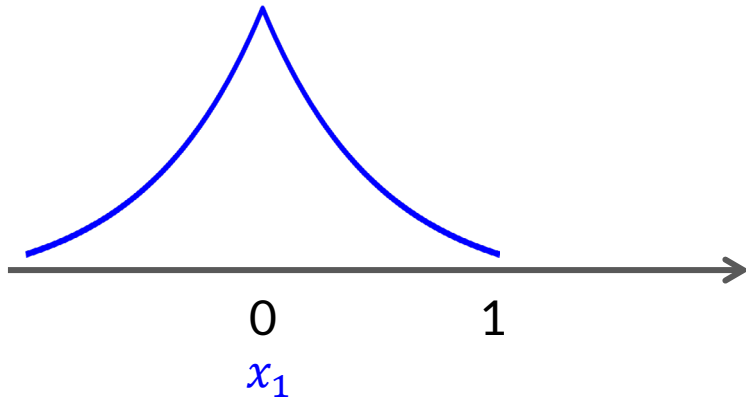
by expected errors



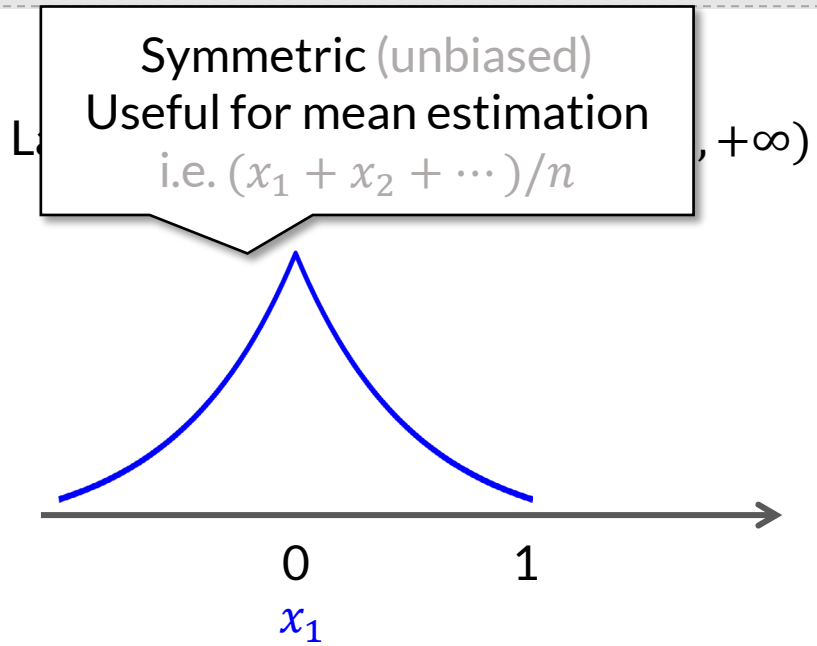
Sensitive $x_1 = 0.5$

LDP Mechanisms for $\mathcal{D} = [0,1]$

Laplace mechanism: $[0,1] \rightarrow (-\infty, +\infty)$

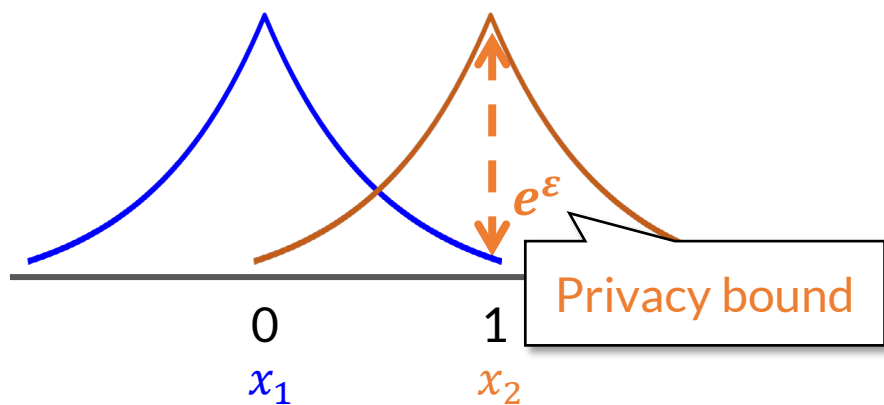


LDP Mechanisms for $\mathcal{D} = [0,1]$

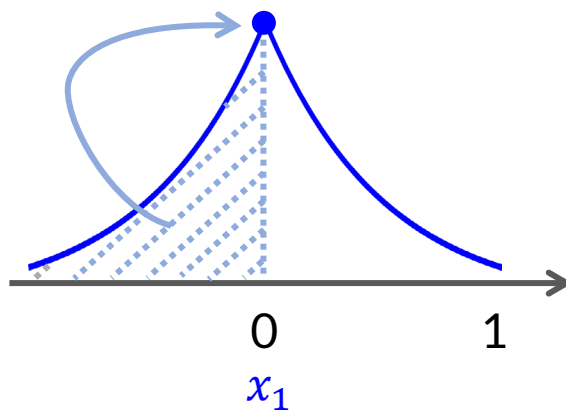


LDP Mechanisms for $\mathcal{D} = [0,1]$

Laplace mechanism: $[0,1] \rightarrow (-\infty, +\infty)$

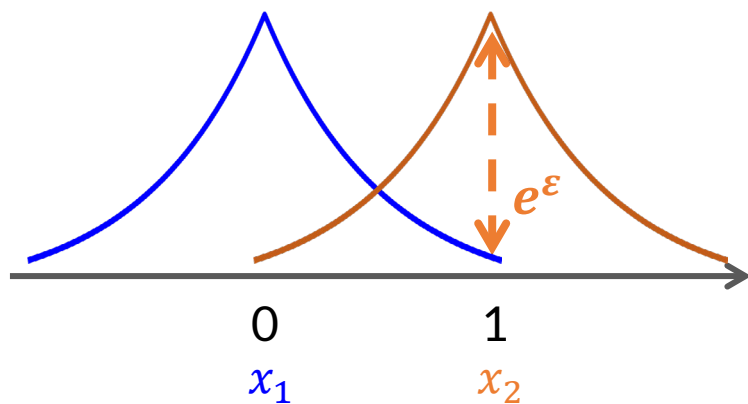


Laplace + **truncation**: $[0,1] \rightarrow [0,1]$

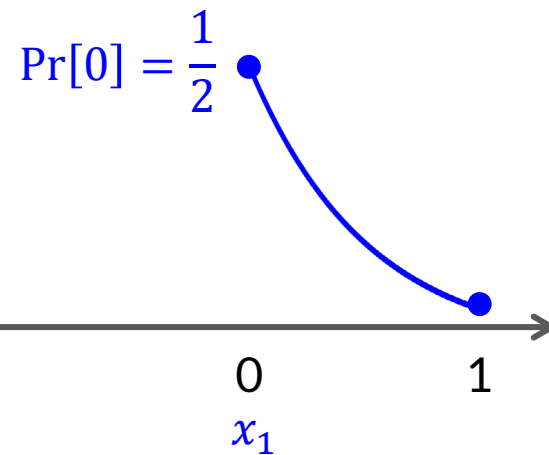


LDP Mechanisms for $\mathcal{D} = [0,1]$

Laplace mechanism: $[0,1] \rightarrow (-\infty, +\infty)$

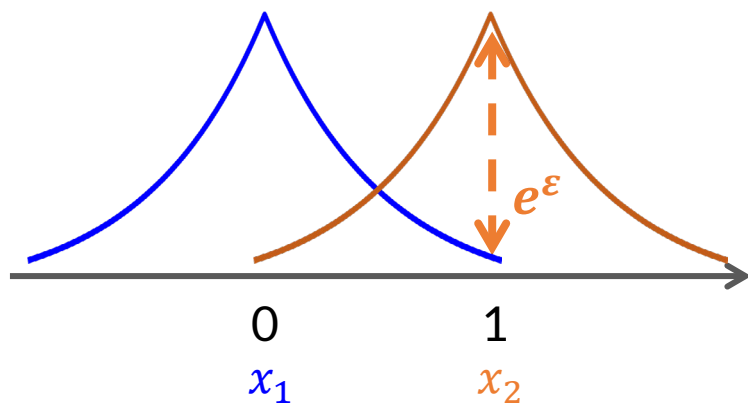


Laplace + **truncation**: $[0,1] \rightarrow [0,1]$

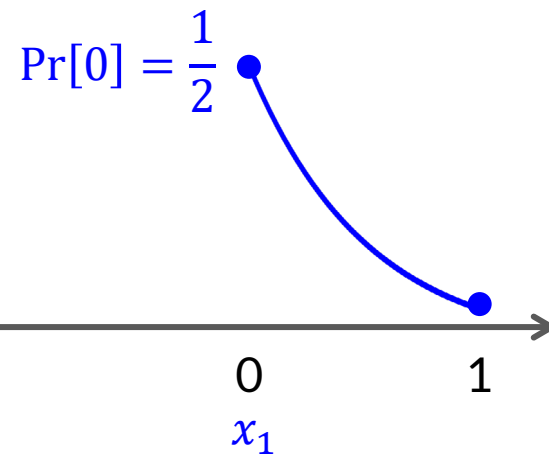


LDP Mechanisms for $\mathcal{D} = [0,1]$

Laplace mechanism: $[0,1] \rightarrow (-\infty, +\infty)$



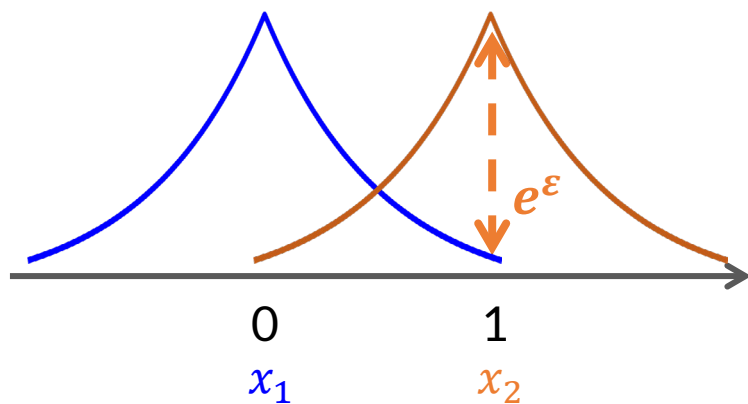
Laplace + **truncation**: $[0,1] \rightarrow [0,1]$



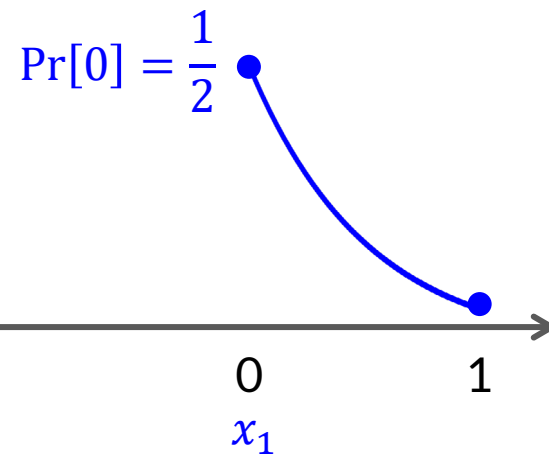
Smaller output
Useful for distribution estimation
i.e. $\{x_1, x_2, \dots\}$

LDP Mechanisms for $\mathcal{D} = [0,1]$

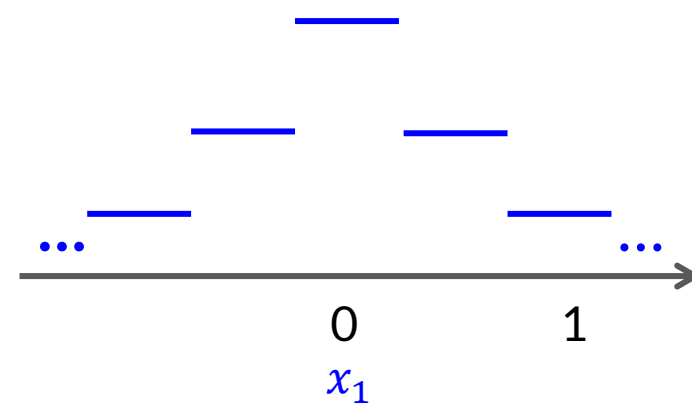
Laplace mechanism: $[0,1] \rightarrow (-\infty, +\infty)$



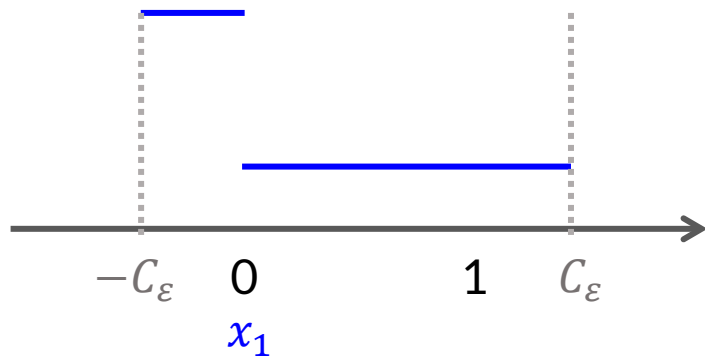
Laplace + **truncation**: $[0,1] \rightarrow [0,1]$



Staircase: $[0,1] \rightarrow (-\infty, +\infty)$

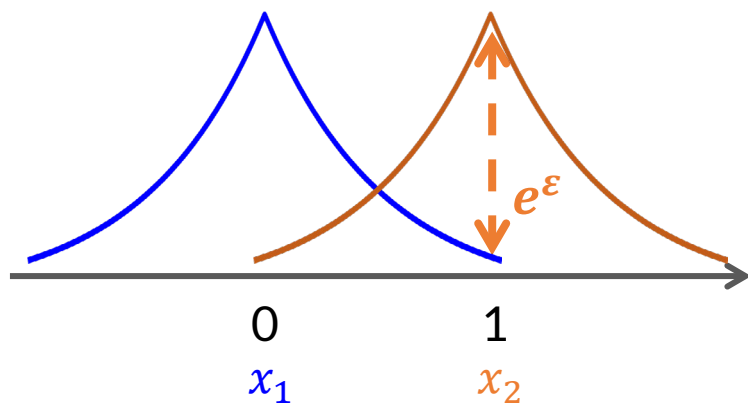


Piecewise mechanism: $[0,1] \rightarrow [-C_\epsilon, C_\epsilon]$

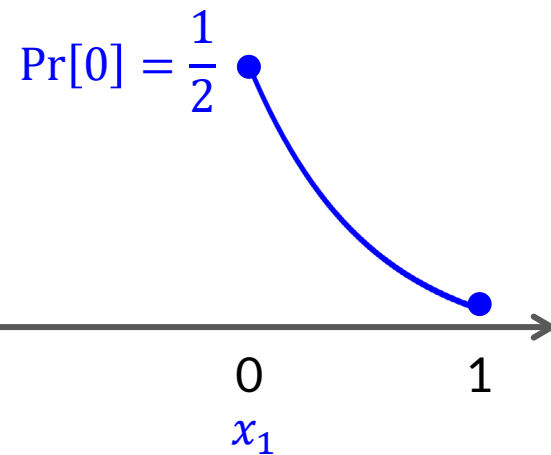


LDP Mechanisms for $\mathcal{D} = [0,1]$

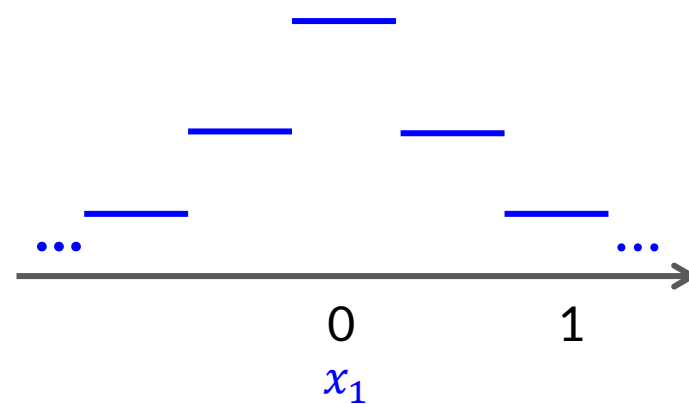
Laplace mechanism: $[0,1] \rightarrow (-\infty, +\infty)$



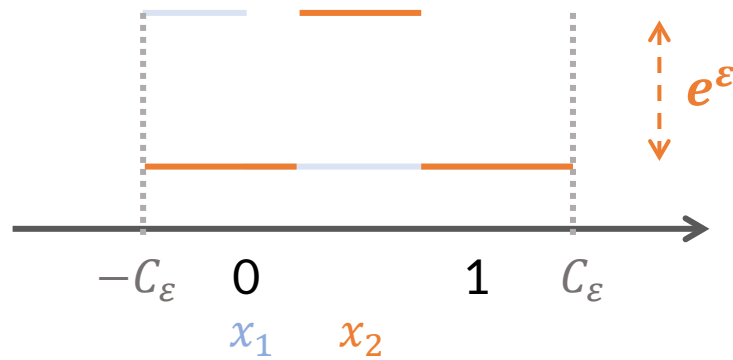
Laplace + **truncation**: $[0,1] \rightarrow [0,1]$



Staircase: $[0,1] \rightarrow (-\infty, +\infty)$

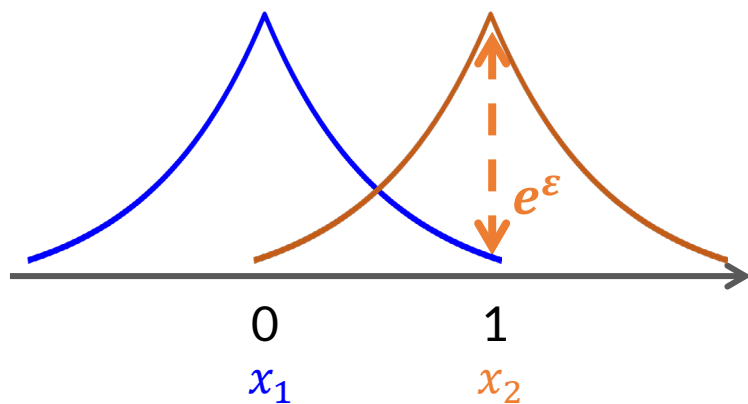


Piecewise mechanism: $[0,1] \rightarrow [-C_\epsilon, C_\epsilon]$

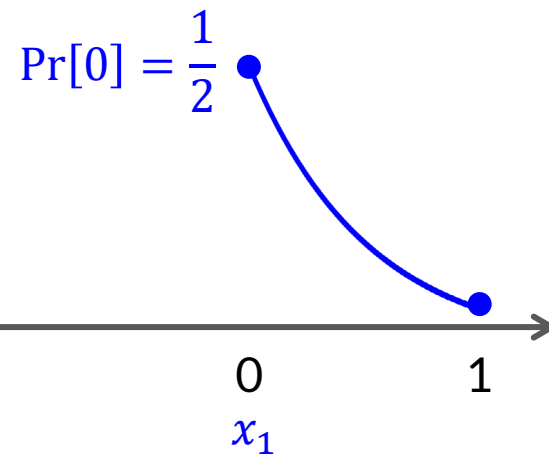


LDP Mechanisms for $\mathcal{D} = [0,1]$

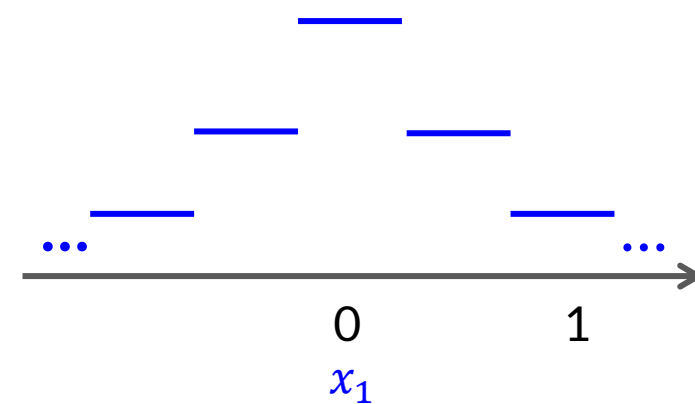
Laplace mechanism: $[0,1] \rightarrow (-\infty, +\infty)$



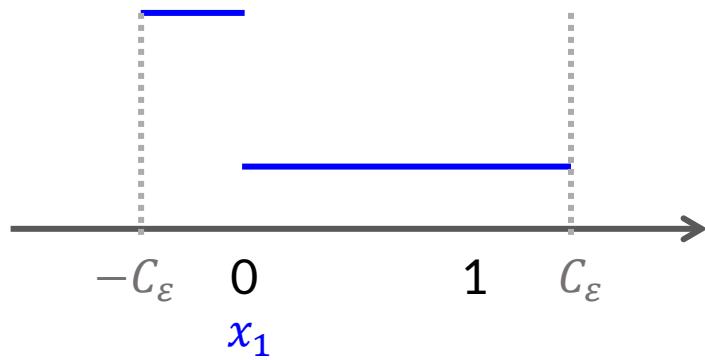
Laplace + **truncation**: $[0,1] \rightarrow [0,1]$



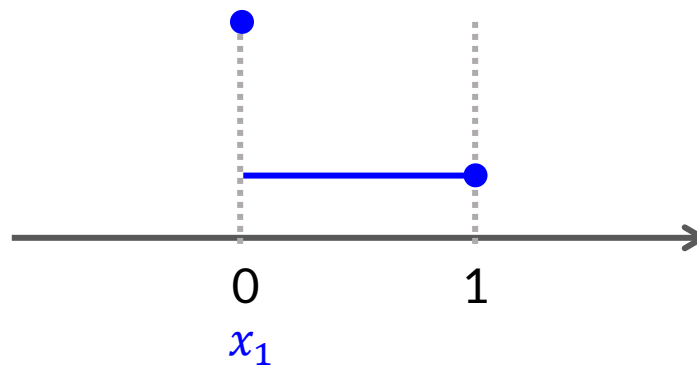
Staircase: $[0,1] \rightarrow (-\infty, +\infty)$



Piecewise mechanism: $[0,1] \rightarrow [-C_\epsilon, C_\epsilon]$



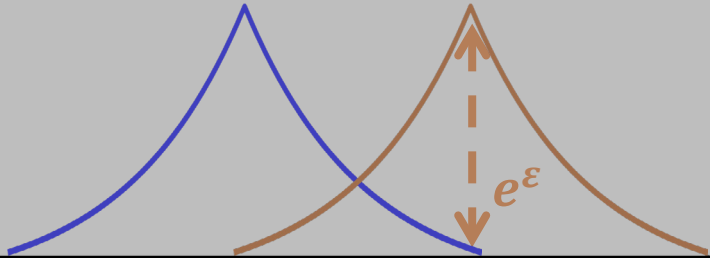
Piecewise + **truncation**: $[0,1] \rightarrow [0,1]$



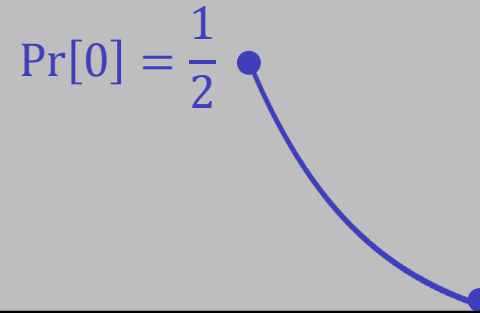
...

LDP Mechanisms for $\mathcal{D} = [0,1]$

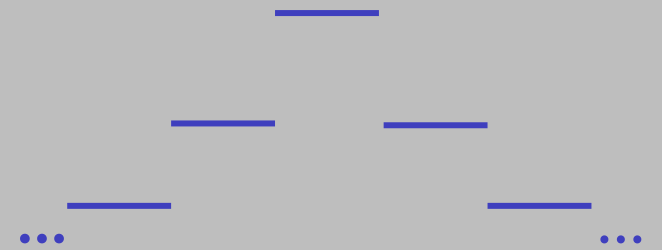
Laplace mechanism: $[0,1] \rightarrow (-\infty, +\infty)$



Laplace + **truncation**: $[0,1] \rightarrow [0,1]$

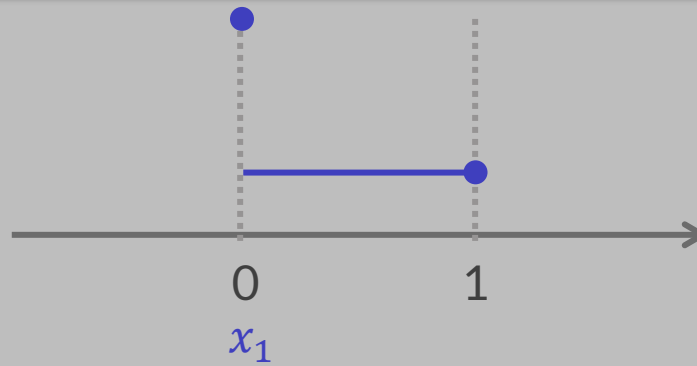
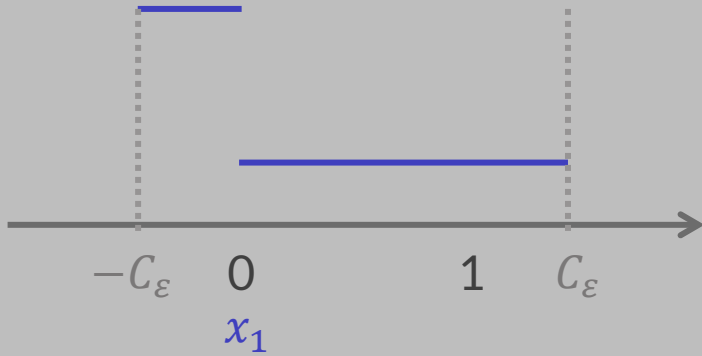


Staircase: $[0,1] \rightarrow (-\infty, +\infty)$



Privacy: LDP with the same ϵ

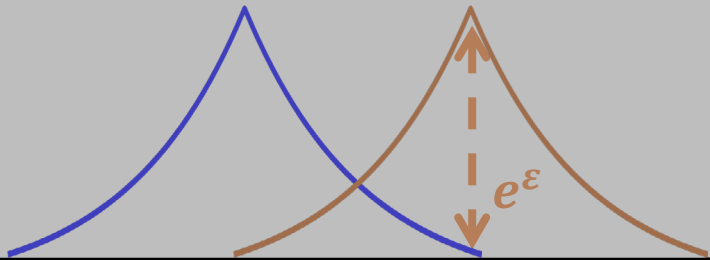
Utility: Different errors



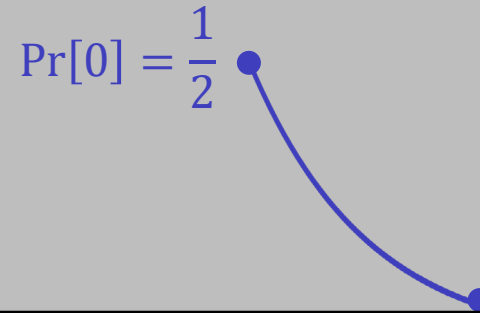
...

LDP Mechanisms for $\mathcal{D} = [0,1]$

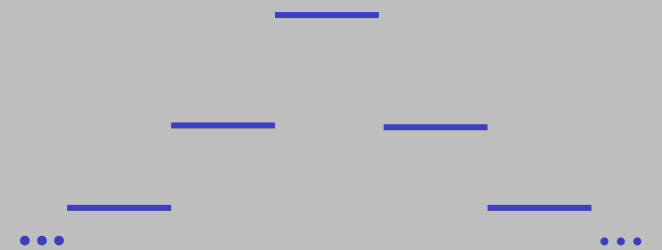
Laplace mechanism: $[0,1] \rightarrow (-\infty, +\infty)$



Laplace + **truncation**: $[0,1] \rightarrow [0,1]$



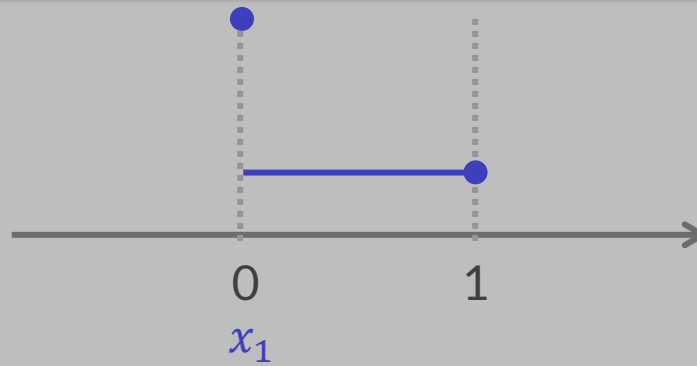
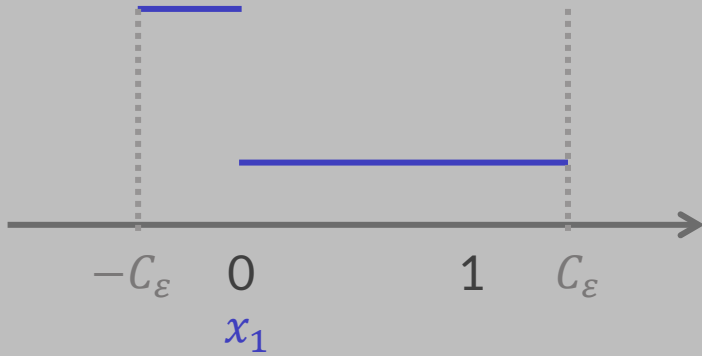
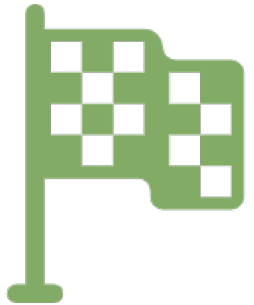
Staircase: $[0,1] \rightarrow (-\infty, +\infty)$



Privacy: LDP with the same ϵ

Utility: Different errors

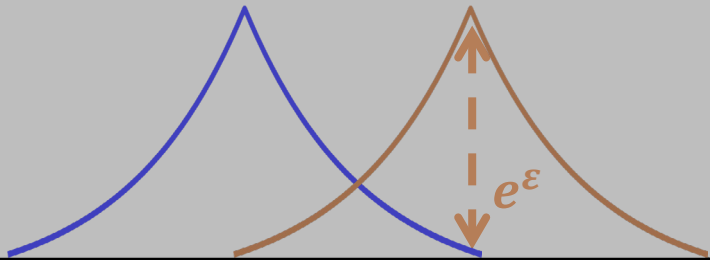
Q: What is the **optimal** LDP mechanism?



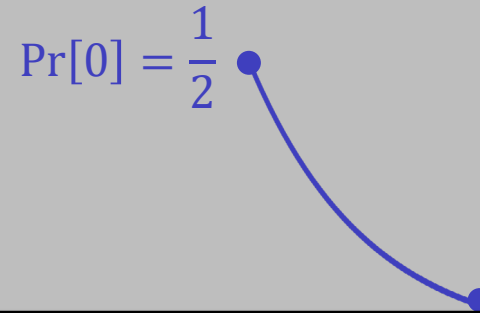
\dots

LDP Mechanisms for $\mathcal{D} = [0,1]$

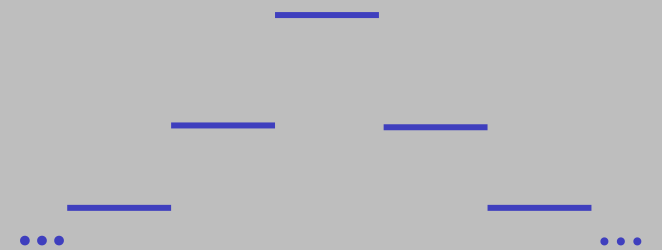
Laplace mechanism: $[0,1] \rightarrow (-\infty, +\infty)$



Laplace + **truncation**: $[0,1] \rightarrow [0,1]$



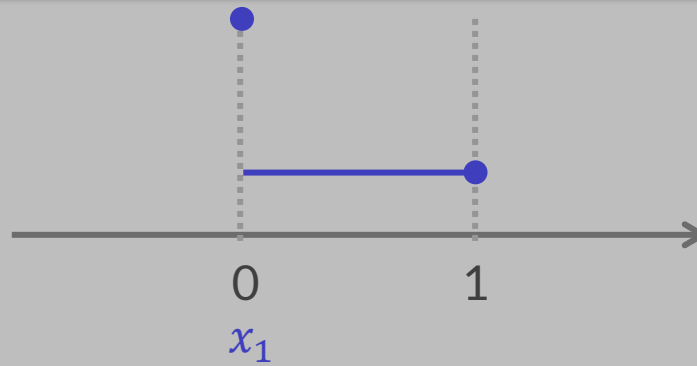
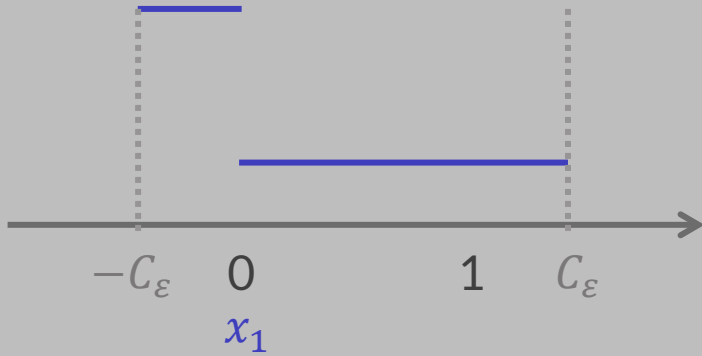
Staircase: $[0,1] \rightarrow (-\infty, +\infty)$



Privacy: LDP with the same ϵ

Utility: Different errors

Q: What is the **optimal piecewise-based** mechanism?

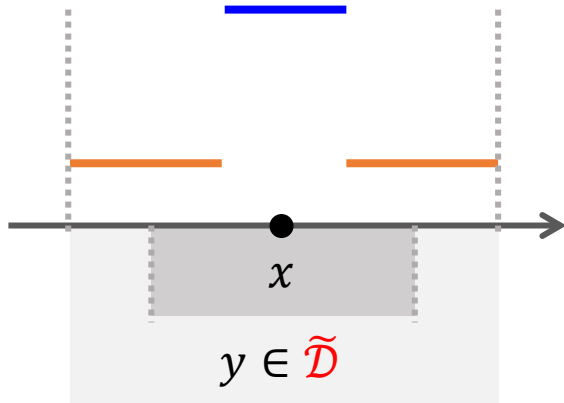


...

3-Piecewise Mechanism

- 3-piecewise distributions on **bounded numerical domain** $\mathcal{D} \rightarrow \tilde{\mathcal{D}}$

- given input x , sample output y from a distribution

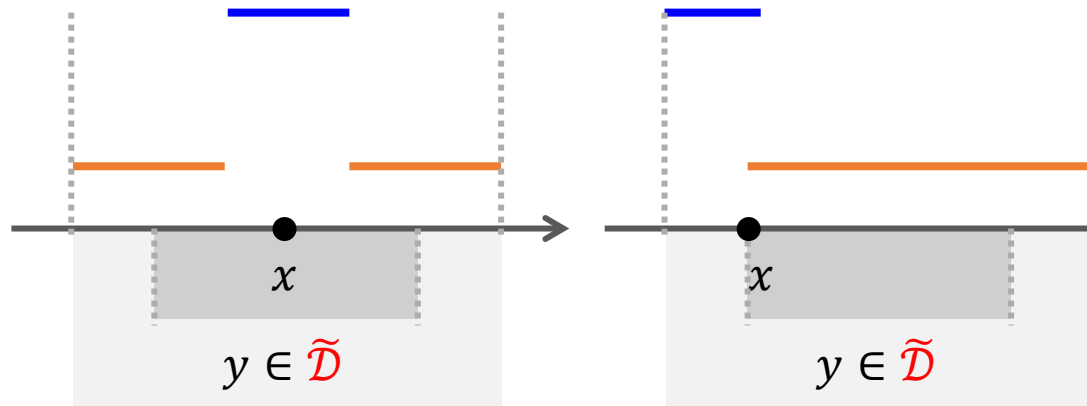


$$pdf[\mathcal{M}(x) = y] = \begin{cases} p_\epsilon & \text{if } y \in [l_{x,\epsilon}, r_{x,\epsilon}] \\ \frac{p_\epsilon}{\exp(\epsilon)} & \text{if } y \in \tilde{\mathcal{D}} \setminus [l_{x,\epsilon}, r_{x,\epsilon}] \end{cases}$$

3-Piecewise Mechanism

- 3-piecewise distributions on **bounded numerical domain $\mathcal{D} \rightarrow \tilde{\mathcal{D}}$**

- given input x , sample output y from a distribution



Sampling probability
depends on ε

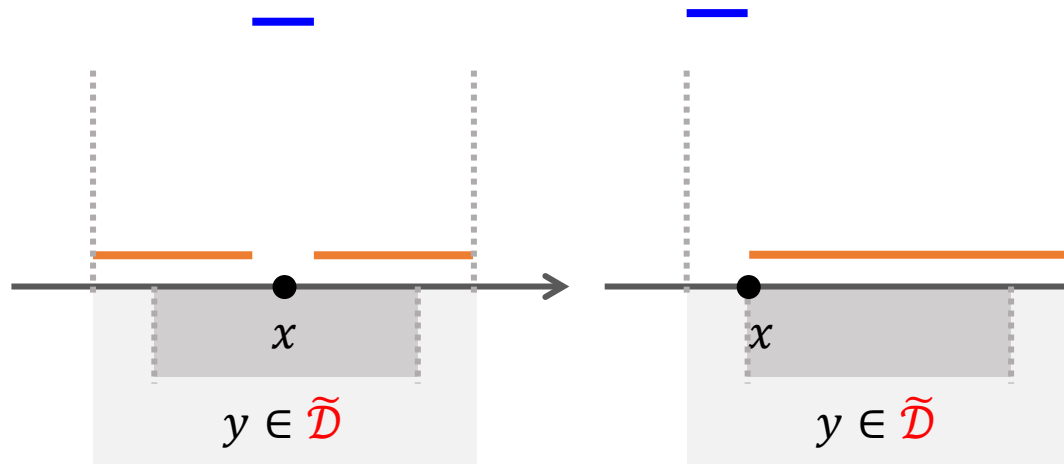
Sampling pieces
depends on x and ε

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

3-Piecewise Mechanism

- 3-piecewise distributions on **bounded numerical domain $\mathcal{D} \rightarrow \tilde{\mathcal{D}}$**

- given input x , sample output y from a distribution



Sampling probability depends on ε

Sampling pieces depends on x and ε

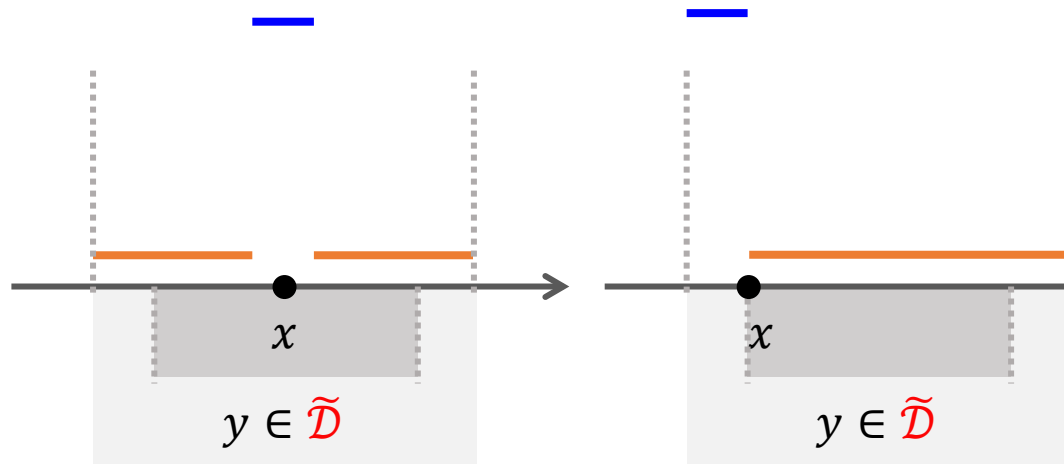
$$pdf[\mathcal{M}(x) = y] = \begin{cases} p_\varepsilon & \text{if } y \in [l_{x,\varepsilon}, r_{x,\varepsilon}] \\ \frac{p_\varepsilon}{\exp(\varepsilon)} & \text{if } y \in \tilde{\mathcal{D}} \setminus [l_{x,\varepsilon}, r_{x,\varepsilon}] \end{cases}$$

- Instantiations: PM [2019], SW [2020], PTT [2024] (design different $p_\varepsilon, l_{x,\varepsilon}, r_{x,\varepsilon}$)

3-Piecewise Mechanism

- 3-piecewise distributions on **bounded numerical domain $\mathcal{D} \rightarrow \tilde{\mathcal{D}}$**

- given input x , sample output y from a distribution



Sampling probability depends on ε

Sampling pieces depends on x and ε

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

- Instantiations: PM [2019], SW [2020], PTT [2024] (design different $p_\varepsilon, l_{x,\varepsilon}, r_{x,\varepsilon}$)
- different errors, but **without optimality**

3-Piecewise Mechanism

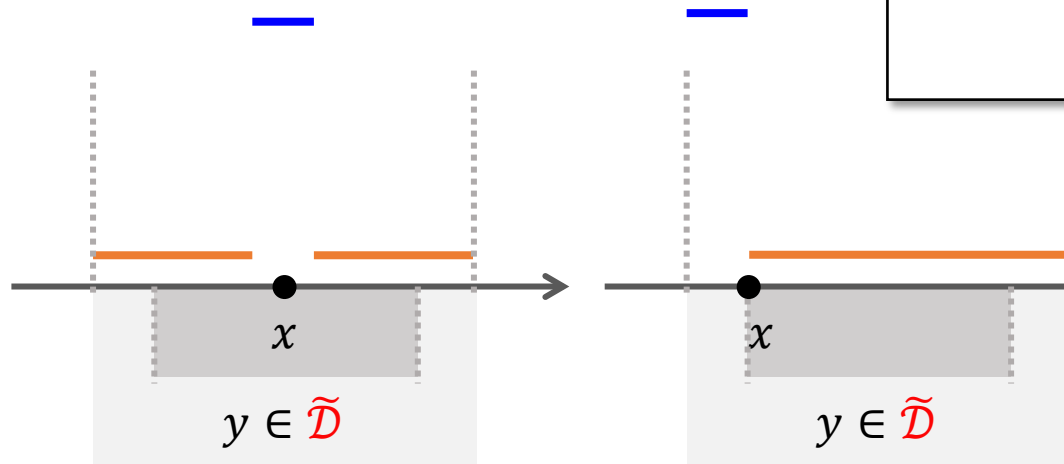


NOT enough to study optimality of piecewise-based mechanism

- 3-piecewise distributions on **bound**

- given input x , sample output y from

- only 3 pieces, two probabilities



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

- Instantiations: PM [2019], SW [2020], PTT [2024] (design different $p_\varepsilon, l_{x,\varepsilon}, r_{x,\varepsilon}$)
- different errors, but **without optimality**

3-Piecewise Mechanism



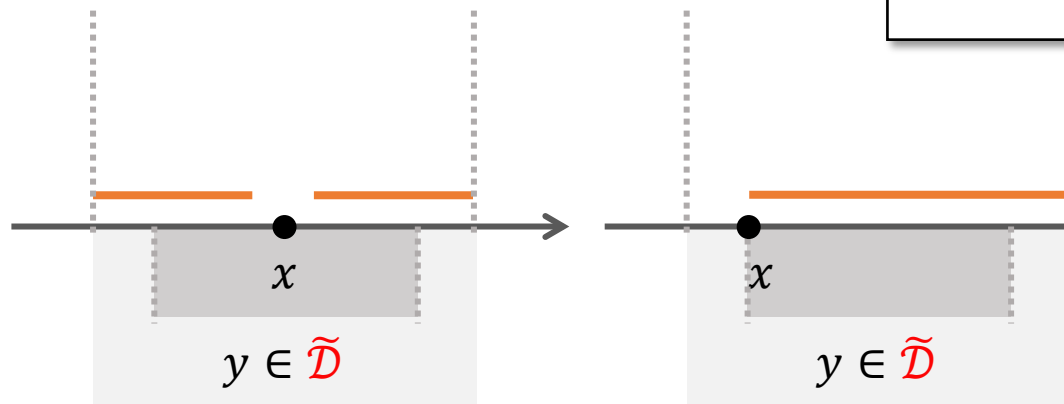
NOT enough to study optimality of piecewise-based mechanism

- 3-piecewise distributions on **bound**

- given input x , sample output y from

- only 3 pieces, two probabilities

- **piecewise distribution can have more pieces and probabilities**

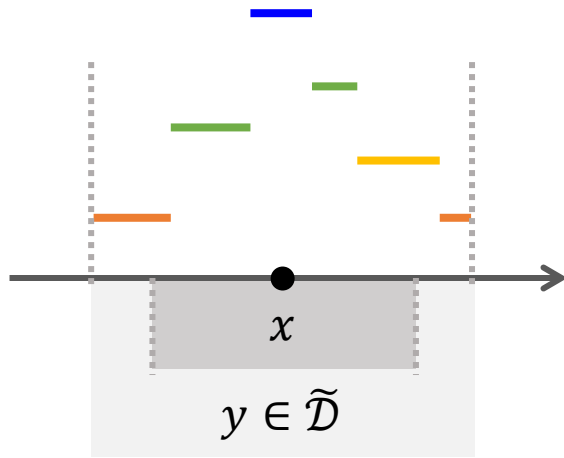


$$pdf[\mathcal{M}(x) = y] = \begin{cases} p_\varepsilon & \text{if } y \in [l_{x,\varepsilon}, r_{x,\varepsilon}] \\ \frac{p_\varepsilon}{\exp(\varepsilon)} & \text{if } y \in \tilde{\mathcal{D}} \setminus [l_{x,\varepsilon}, r_{x,\varepsilon}] \end{cases}$$

- Instantiations: PM [2019], SW [2020], PTT [2024] (design different $p_\varepsilon, l_{x,\varepsilon}, r_{x,\varepsilon}$)
 - different errors, but **without optimality**

Generalized Piecewise-based Mechanism

- Most generalized version: m -piecewise distributions

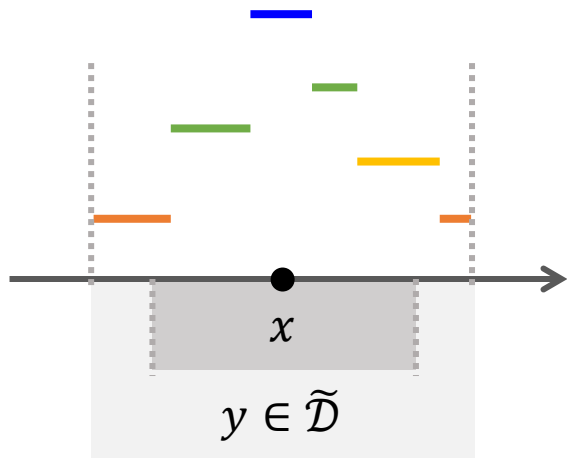


$$pdf[\mathcal{M}(x) = y] = \begin{cases} p_{1,\varepsilon} & \text{if } y \in [l_{1,x,\varepsilon}, r_{1,x,\varepsilon}] \\ p_{2,\varepsilon} & \text{if } y \in [l_{2,x,\varepsilon}, r_{2,x,\varepsilon}] \\ \dots & \dots \\ p_{m,\varepsilon} & \text{if } y \in [l_{m,x,\varepsilon}, r_{m,x,\varepsilon}] \end{cases}$$

$$\max \frac{p_{i,\varepsilon}}{p_{j,\varepsilon}} \leq e^\varepsilon \text{ (LDP constraint)}$$

Generalized Piecewise-based Mechanism

- Most generalized version: m -piecewise distributions



$$pdf[\mathcal{M}(x) = y] = \begin{cases} p_{1,\varepsilon} & \text{if } y \in [l_{1,x,\varepsilon}, r_{1,x,\varepsilon}] \\ p_{2,\varepsilon} & \text{if } y \in [l_{2,x,\varepsilon}, r_{2,x,\varepsilon}] \\ \dots & \dots \\ p_{m,\varepsilon} & \text{if } y \in [l_{m,x,\varepsilon}, r_{m,x,\varepsilon}] \end{cases}$$

$$\frac{p_{i,\varepsilon}}{p_{j,\varepsilon}} \leq e^\varepsilon \text{ (LDP constraint)}$$

- Error (data utility):

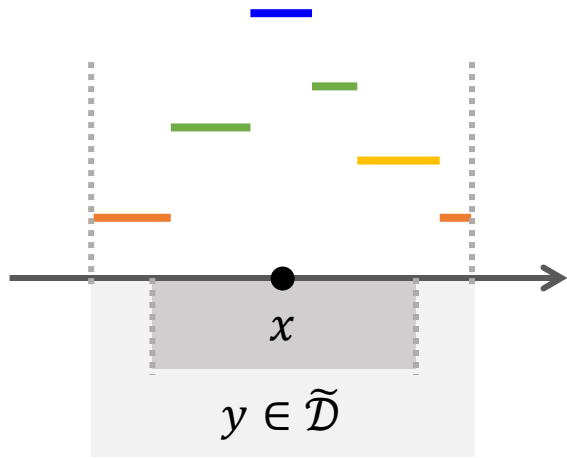
$$\mathcal{L}(y, x)$$

↑

$$\mathcal{L}(y, x) := |y - x|^p$$

Generalized Piecewise-based Mechanism

- Most generalized version: m -piecewise distributions



$$pdf[\mathcal{M}(x) = y] = \begin{cases} p_{1,\varepsilon} & \text{if } y \in [l_{1,x,\varepsilon}, r_{1,x,\varepsilon}] \\ p_{2,\varepsilon} & \text{if } y \in [l_{2,x,\varepsilon}, r_{2,x,\varepsilon}] \\ \dots & \dots \\ p_{m,\varepsilon} & \text{if } y \in [l_{m,x,\varepsilon}, r_{m,x,\varepsilon}] \end{cases}$$

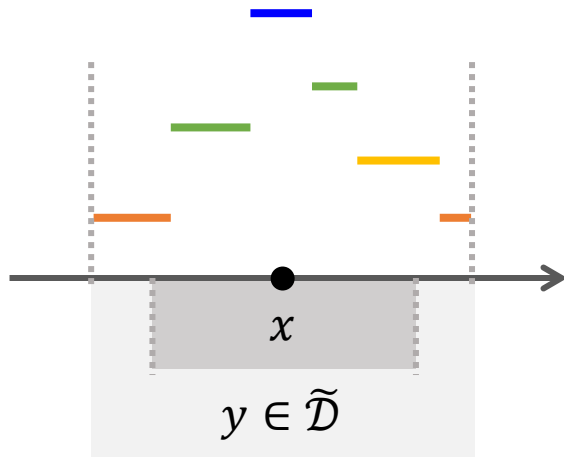
$$\frac{p_{i,\varepsilon}}{p_{j,\varepsilon}} \leq e^\varepsilon \text{ (LDP constraint)}$$

- Expected error:

$$\int_{\tilde{\mathcal{D}}} \mathcal{L}(y, x) \cdot pdf[\mathcal{M}(x) = y] dy$$

Optimal Piecewise-based Mechanism

- Most generalized version: m -piecewise distributions



$$pdf[\mathcal{M}(x) = y] = \begin{cases} p_{1,\varepsilon} & \text{if } y \in [l_{1,x,\varepsilon}, r_{1,x,\varepsilon}] \\ p_{2,\varepsilon} & \text{if } y \in [l_{2,x,\varepsilon}, r_{2,x,\varepsilon}] \\ \dots & \dots \\ p_{m,\varepsilon} & \text{if } y \in [l_{m,x,\varepsilon}, r_{m,x,\varepsilon}] \end{cases}$$

$$\frac{p_{i,\varepsilon}}{p_{j,\varepsilon}} \leq e^\varepsilon \text{ (LDP constraint)}$$

- Expected error:

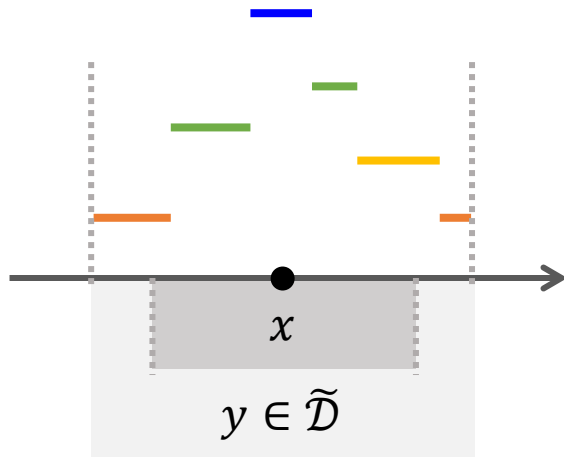
$$\min_{\mathcal{M}: p_i, l_i, r_i} \int_{\tilde{\mathcal{D}}} \mathcal{L}(y, x) \cdot pdf[\mathcal{M}(x) = y] dy$$



Find \mathcal{M} to minimize the error at x

Optimal Piecewise-based Mechanism

- Most generalized version: m -piecewise distributions



$$pdf[\mathcal{M}(x) = y] = \begin{cases} p_{1,\varepsilon} & \text{if } y \in [l_{1,x,\varepsilon}, r_{1,x,\varepsilon}] \\ p_{2,\varepsilon} & \text{if } y \in [l_{2,x,\varepsilon}, r_{2,x,\varepsilon}] \\ \dots & \dots \\ p_{m,\varepsilon} & \text{if } y \in [l_{m,x,\varepsilon}, r_{m,x,\varepsilon}] \end{cases}$$

$$\frac{p_{i,\varepsilon}}{p_{j,\varepsilon}} \leq e^\varepsilon \text{ (LDP constraint)}$$

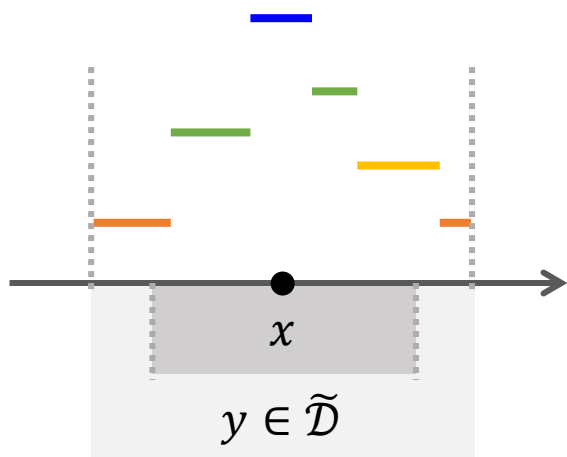
- Expected error:

$$\min_{\mathcal{M}: p_i, l_i, r_i} \max_{x \in \mathcal{D}} \int_{\tilde{\mathcal{D}}} \mathcal{L}(y, x) \cdot pdf[\mathcal{M}(x) = y] dy$$

Find \mathcal{M} to minimize the worst-case error

Optimal Piecewise-based Mechanism

- Most generalized version: m -piecewise distributions



$$pdf[\mathcal{M}(x) = y] = \begin{cases} p_{1,\varepsilon} & \text{if } y \in [l_{1,x,\varepsilon}, r_{1,x,\varepsilon}] \\ p_{2,\varepsilon} & \text{if } y \in [l_{2,x,\varepsilon}, r_{2,x,\varepsilon}] \\ \dots & \dots \\ p_{m,\varepsilon} & \text{if } y \in [l_{m,x,\varepsilon}, r_{m,x,\varepsilon}] \end{cases}$$

$$\frac{p_{i,\varepsilon}}{p_{j,\varepsilon}} \leq e^\varepsilon \text{ (LDP constraint)}$$

Solved \mathcal{M} is
the optimal piecewise-based mechanism

Mathematically \equiv to find the optimal
piecewise distribution under the LDP constraint

$$\min_{\mathcal{M}: p_i, l_i, r_i} \max_{x \in \mathcal{D}} \int_{\tilde{\mathcal{D}}} \mathcal{L}(y, x) \cdot pdf[\mathcal{M}(x) = y] dy$$

Find \mathcal{M} to minimize the worst-case error

Challenges & Proofs

- Challenges

1. min-max problem & multiple variables
2. optimal results only for a specific m

$$\min_{\mathcal{M}: p_i, l_i, r_i} \max_{x \in \mathcal{D}} \int_{\tilde{\mathcal{D}}} \mathcal{L}(y, x) \cdot pdf[\mathcal{M}(x) = y] dy$$

2. $i \in [m]$ 1. variables p_i, l_i, r_i

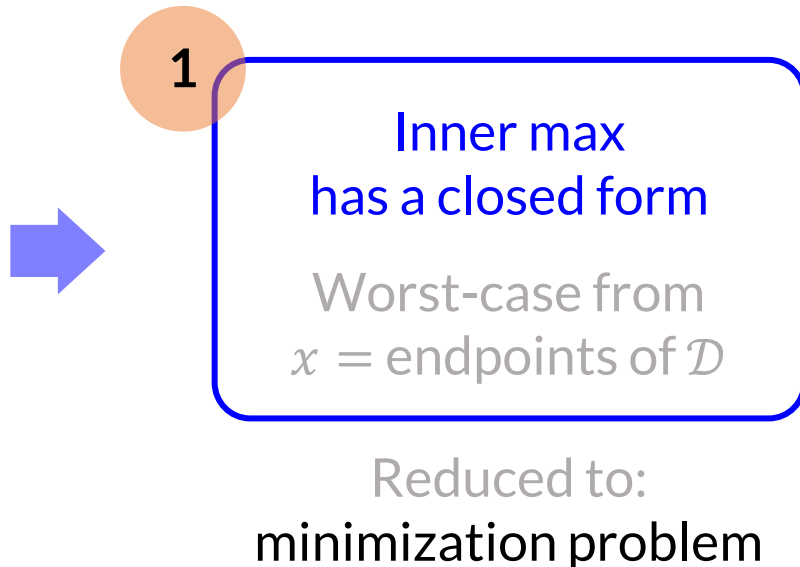
Challenges & Proofs

- Challenges

1. min-max problem & multiple variables
2. optimal results only for a specific m

$$\min_{\mathcal{M}: p_i, l_i, r_i} \max_{x \in \mathcal{D}} \int_{\tilde{\mathcal{D}}} \mathcal{L}(y, x) \cdot pdf[\mathcal{M}(x) = y] dy$$

2. $i \in [m]$ 1. variables p_i, l_i, r_i



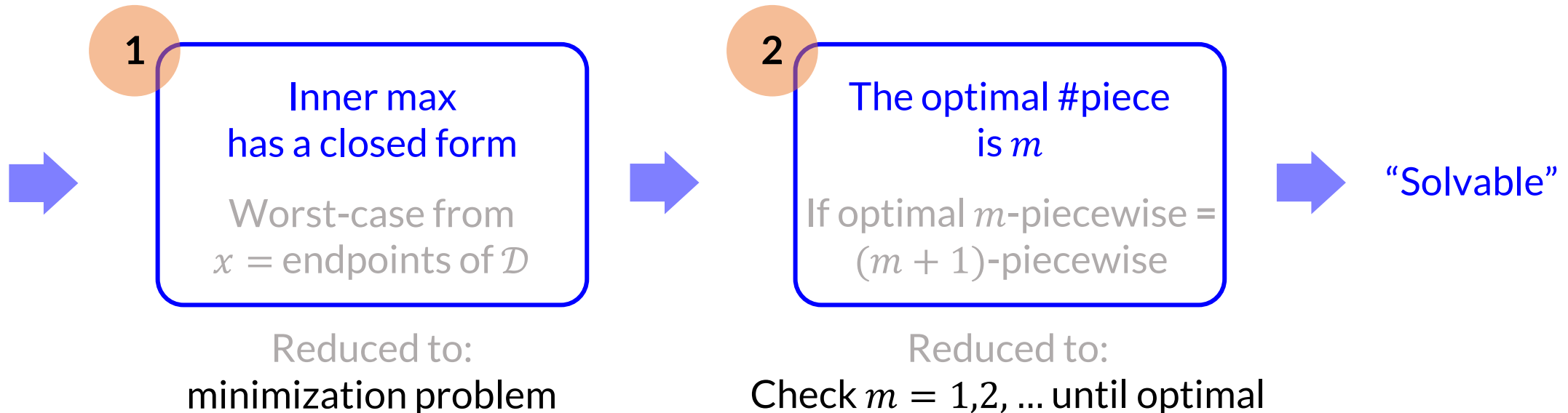
Challenges & Proofs

■ Challenges

1. min-max problem & multiple variables
2. optimal results only for a specific m

$$\min_{\mathcal{M}: p_i, l_i, r_i} \max_{x \in \mathcal{D}} \int_{\tilde{\mathcal{D}}} \mathcal{L}(y, x) \cdot pdf[\mathcal{M}(x) = y] dy$$

2. $i \in [m]$ 1. variables p_i, l_i, r_i



Proof Intuitions

- Worst-case error is achieved at endpoints

$$\max_{x \in \mathcal{D}} \int_{\tilde{\mathcal{D}}} \mathcal{L}(y, x) \cdot pdf[\mathcal{M}(x) = y] dy = \max_{x \in \mathcal{D}} \sum_{i=1}^m p_i \int_{l_i}^{r_i} \mathcal{L}(y, x) dy \quad (m\text{-piecewise distribution})$$

Proof Intuitions

- Worst-case error is achieved at endpoints

$$\max_{x \in \mathcal{D}} \int_{\tilde{\mathcal{D}}} \mathcal{L}(y, x) \cdot pdf[\mathcal{M}(x) = y] dy = \max_{x \in \mathcal{D}} \sum_{i=1}^m p_i \int_{l_i}^{r_i} \mathcal{L}(y, x) dy \quad (m\text{-piecewise distribution})$$

convex function w.r.t x

Proof Intuitions

- Worst-case error is achieved at endpoints

$$\begin{aligned} \max_{x \in \mathcal{D}} \int_{\tilde{\mathcal{D}}} \mathcal{L}(y, x) \cdot pdf[\mathcal{M}(x) = y] dy &= \max_{x \in \mathcal{D}} \sum_{i=1}^m p_i \int_{l_i}^{r_i} \mathcal{L}(y, x) dy && \text{(m-piecewise distribution)} \\ &= \max_{x \in \{a, b\}} \sum_{i=1}^m p_i \int_{l_i}^{r_i} \mathcal{L}(y, x) dy && \text{(maximum principle)} \end{aligned}$$

Proof Intuitions

- Worst-case error is achieved at endpoints

$$\max_{x \in \mathcal{D}} \int_{\tilde{\mathcal{D}}} \mathcal{L}(y, x) \cdot pdf[\mathcal{M}(x)] = \overbrace{\quad}^m \quad \text{c}r_i$$

After merging redundant pieces

- Optimal #piece is m if optimal m -piecewise = $(m+1)$ -piecewise

if: $\min_{e_1, e_2, e_3} e_1 + e_2 + e_3 = \min_{e_1, e_2, e_3, e_4} e_1 + e_2 + e_3 + e_4$

Error from an arbitrary piece
(≥ 0 variable)

i.e. the error can't be lowered by arbitrary ≥ 0 variable

Proof Intuitions

- Worst-case error is achieved at endpoints

$$\max_{x \in \mathcal{D}} \int_{\tilde{\mathcal{D}}} \mathcal{L}(y, x) \cdot pdf[\mathcal{M}(x)] = \overbrace{\quad}^m \quad \overbrace{\quad}^{c r_i}$$

After merging redundant pieces

- Optimal #piece is m if optimal m -piecewise = $(m+1)$ -piecewise

if: $\min_{e_1, e_2, e_3} e_1 + e_2 + e_3 = \min_{e_1, e_2, e_3, e_4} e_1 + e_2 + e_3 + e_4$ Error from an arbitrary piece
(≥ 0 variable)

i.e. the error can't be lowered by arbitrary ≥ 0 variable

then: $= \min_{e_1, e_2, e_3, e_4, e_5} e_1 + e_2 + e_3 + e_4 + e_5$

otherwise, $e_4 \leftarrow e_4 + e_5$ can further lower the error

NOT Manually “Solvable” When $m \geq 4$

- Too many variables & non-linear

$$\max_{x \in \{a, b\}} \sum_{i=1}^m p_i \int_{l_i}^{r_i} \mathcal{L}(y, x) dy$$

NOT Manually “Solvable” When $m \geq 4$

- Too many variables & non-linear
- Efficiently solved by off-the-shelf solvers, e.g. Gurobi
 - limitation: needs given ε
 - limitation: cannot provide closed-form $\mathcal{M}: p_i, l_i, r_i$
 - can be used to analyze optimality

$$\max_{x \in \{a, b\}} \sum_{i=1}^m p_i \int_{l_i}^{r_i} \mathcal{L}(y, x) dy$$

NOT Manually “Solvable” When $m \geq 4$

- Too many variables & non-linear
- Efficiently solved by off-the-shelf solvers, e.g. Gurobi
 - limitation: needs given ε
 - limitation: cannot provide closed-form $\mathcal{M}: p_i, l_i, r_i$
 - can be used to analyze optimality

$$\max_{x \in \{a, b\}} \sum_{i=1}^m p_i \int_{l_i}^{r_i} \mathcal{L}(y, x) dy$$

Monte Carlo testing
with 10^4 random samples

Hypothesis. For any domain $\mathcal{D} \rightarrow \mathcal{D}$, under error metrics $\mathcal{L}(y, x) := |y - x|$ and $\mathcal{L}(y, x) := (y - x)^2$, the optimal piecewise-based mechanism falls into 3-piecewise mechanism.

NOT Manually “Solvable” When $m \geq 4$

- Too many variables & non-linear
- Efficiently solved by off-the-shelf solvers, e.g. Gurobi
 - limitation: needs **given ε**
 - limitation: cannot provide **closed-form \mathcal{M}** : p_i, l_i, r_i
 - can be used to analyze optimality

$$\max_{x \in \{a, b\}} \sum_{i=1}^m p_i \int_{l_i}^{r_i} \mathcal{L}(y, x) dy$$

Monte Carlo testing
with 10^4 random samples

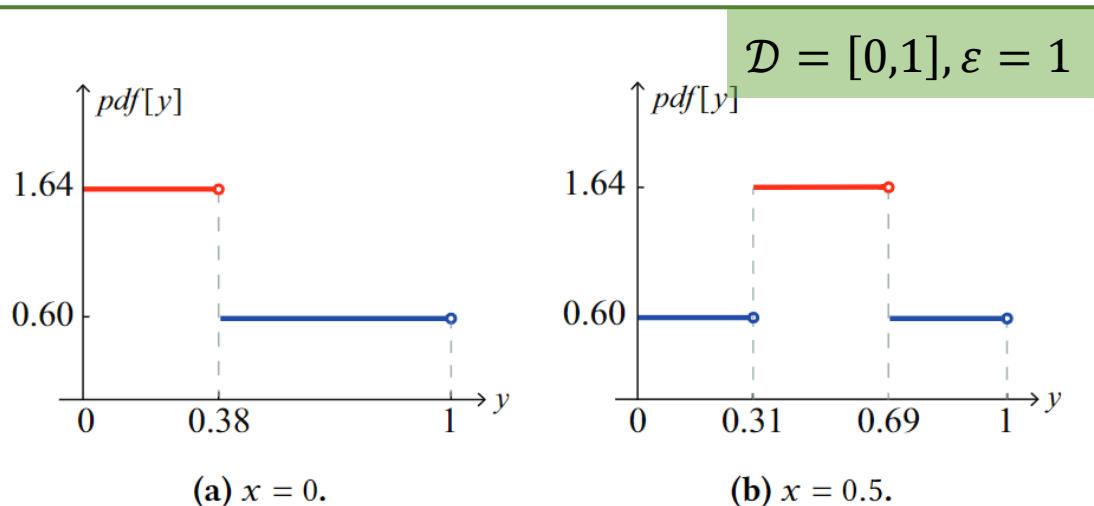
Hypothesis. For any domain $\mathcal{D} \rightarrow \mathcal{D}$, under error metrics $\mathcal{L}(y, x) := |y - x|$ and $\mathcal{L}(y, x) := (y - x)^2$, the optimal piecewise-based mechanism falls into **3-piecewise mechanism**.

closed-form \mathcal{M} can be manually solved
(different from existing instantiations)

NOT Manually “Solvable” When $m \geq 4$

- Too many variables & non-linear
- Efficiently solved by off-the-shelf solvers, e.g. Gurobi
 - limitation: needs given ε
 - limitation: cannot provide closed-form \mathcal{M} : p_i, l_i, r_i

$$\max_{x \in \{a,b\}} \sum_{i=1}^m p_i \int_{l_i}^{r_i} \mathcal{L}(y, x) dy$$



Monte Carlo testing
with 10^4 random samples

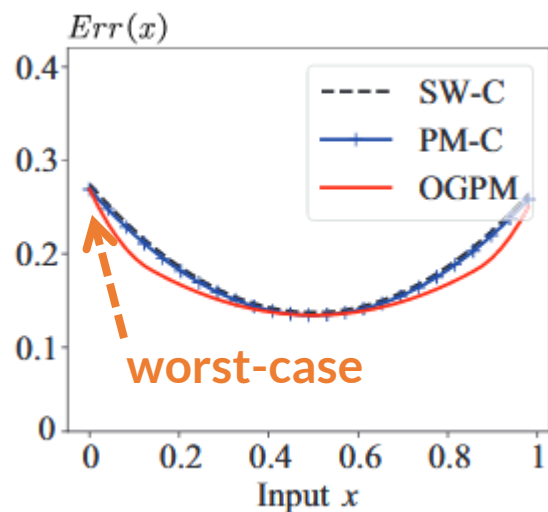
Under error metrics $\mathcal{L}(y, x) := |y - x|$ and $\mathcal{L}(y, x) :=$
mechanism falls into 3-piecewise mechanism.

closed-form \mathcal{M} can be manually solved
(different from existing instantiations)

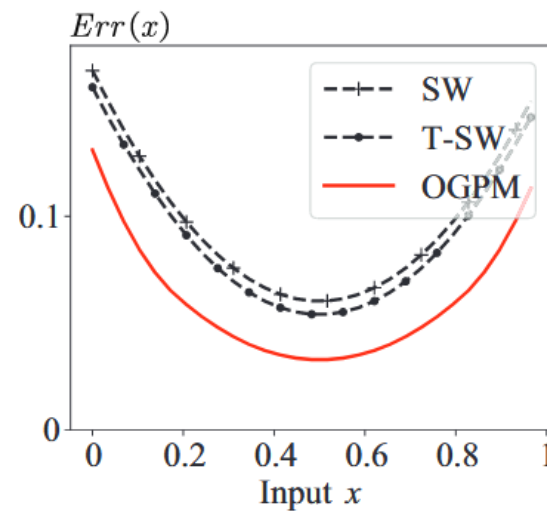
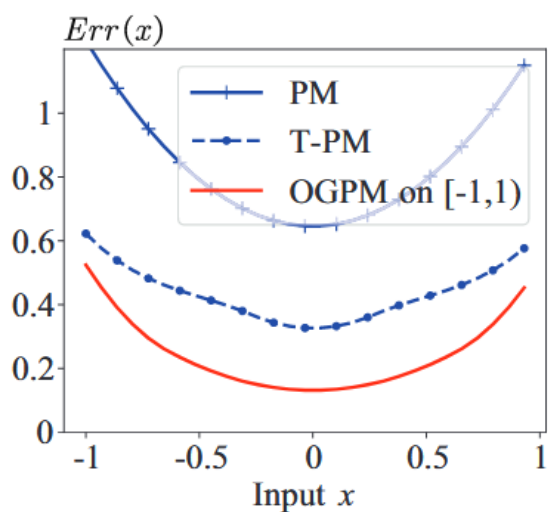
Comparison of Expected Errors

Whole-domain error (i.e. each point in \mathcal{D}) ($\epsilon = 2$)

Compressed PM, SW

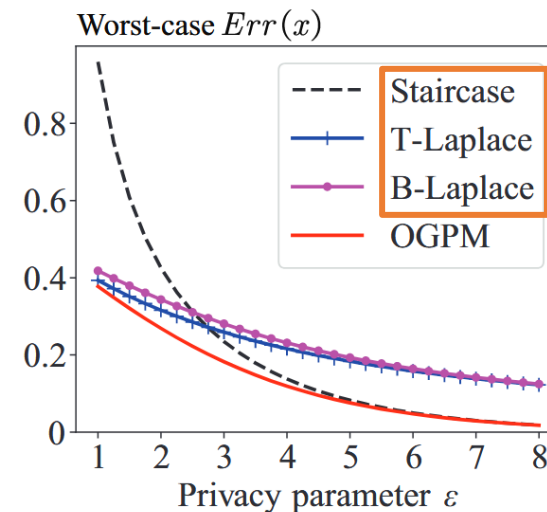


Original and truncated PM, SW



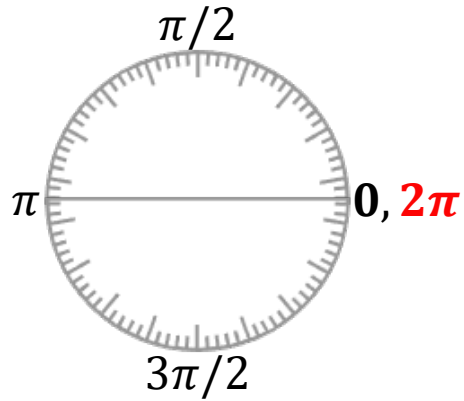
Worst-case error

Non-piecewise-based



Circular Domain

- Different meaning of distance, e.g. distance $(0, 2\pi) = 0$

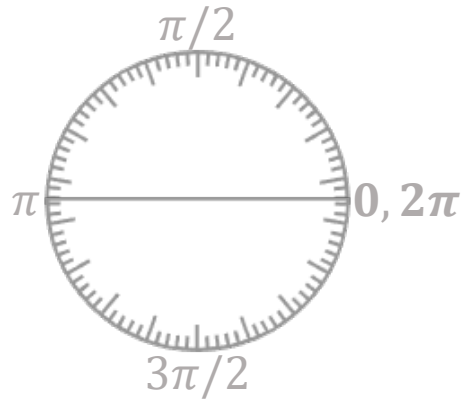


$$\mathcal{L} \rightarrow \mathcal{L}_{\text{mod}}$$

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

Circular Domain

- Different meaning of distance, e.g. $\text{distance}(0, 2\pi) = 0$



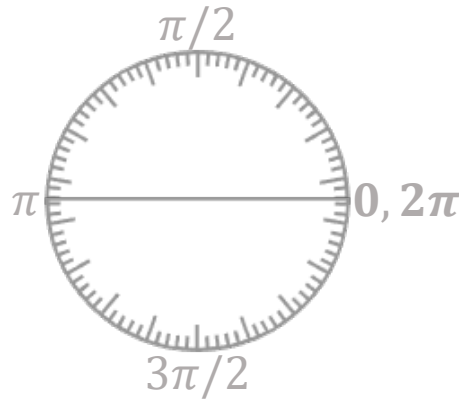
$$\mathcal{L} \rightarrow \mathcal{L}_{\text{mod}}$$

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

$$\Rightarrow \min_{\mathcal{M}: p_i, l_i, r_i} \max_{x \in [0, 2\pi]} \int_{\tilde{\mathcal{D}}} \mathcal{L}_{\text{mod}}(y, x) \cdot \text{pdf}[\mathcal{M}(x) = y] dy$$

Circular Domain

- Different meaning of distance, e.g. distance $(0, 2\pi) = 0$

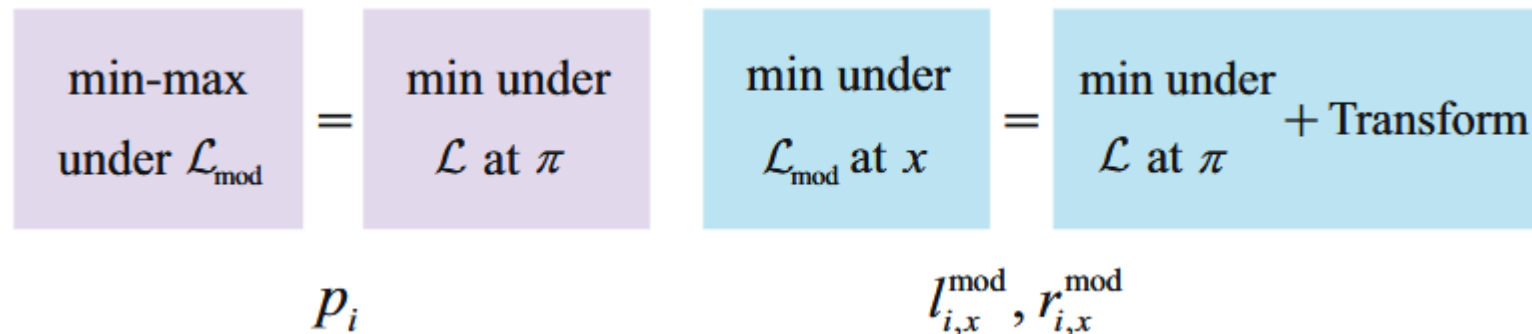


$$\mathcal{L} \rightarrow \mathcal{L}_{\text{mod}}$$

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

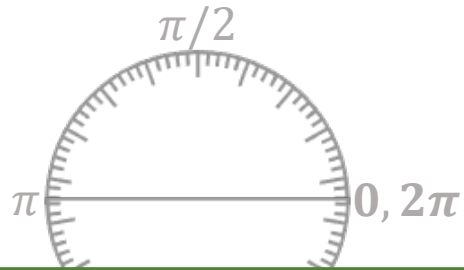
$$\Rightarrow \min_{\mathcal{M}: p_i, l_i, r_i} \max_{x \in [0, 2\pi]} \int_{\tilde{\mathcal{D}}} \mathcal{L}_{\text{mod}}(y, x) \cdot pdf[\mathcal{M}(x) = y] dy$$

- Linking** to problems in the classical domain



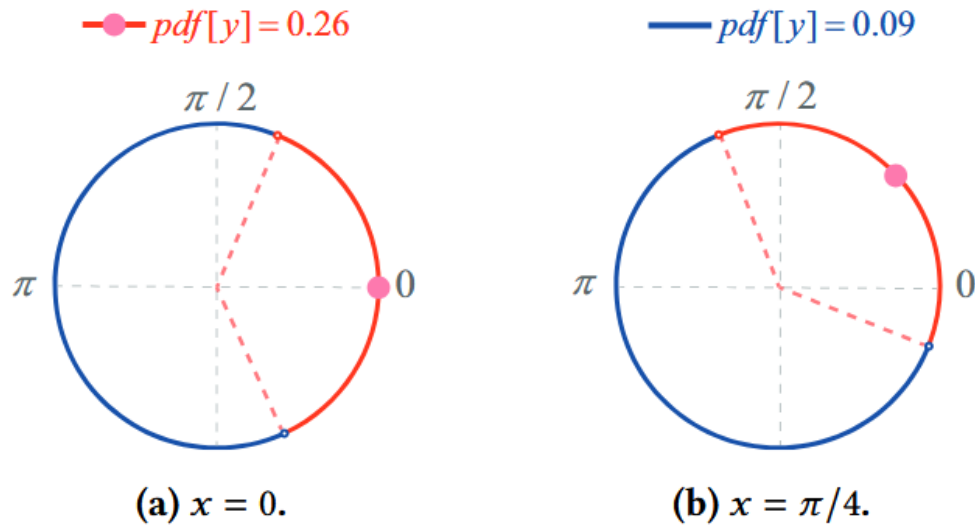
Circular Domain

- Different meaning of distance, e.g. distance $(0, 2\pi) = 0$



$$\mathcal{L} \rightarrow \mathcal{L}_{\text{mod}}$$

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



$$\min_{x: p_i, l_i, r_i} \max_{x \in [0, 2\pi]} \int_{\tilde{\mathcal{D}}} \mathcal{L}_{\text{mod}}(y, x) \cdot pdf[\mathcal{M}(x) = y] dy$$

min under
 \mathcal{L}_{mod} at x

=

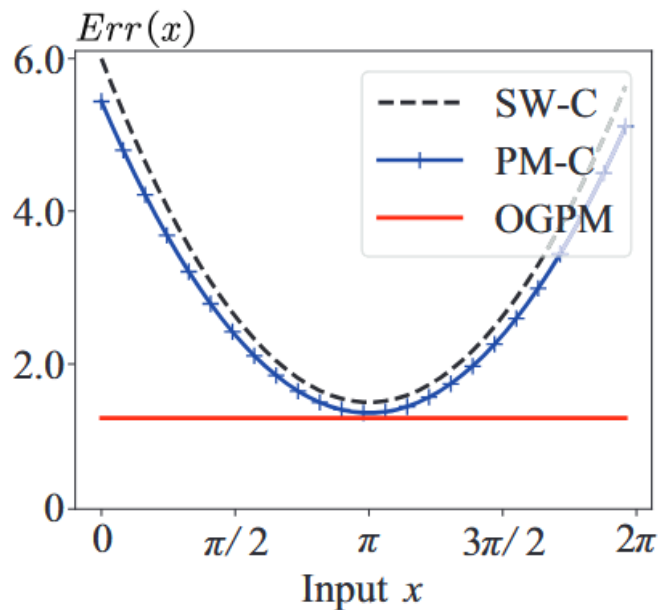
min under
 \mathcal{L} at π + Transform

$$l_{i,x}^{\text{mod}}, r_{i,x}^{\text{mod}}$$

Comparison of Expected Errors

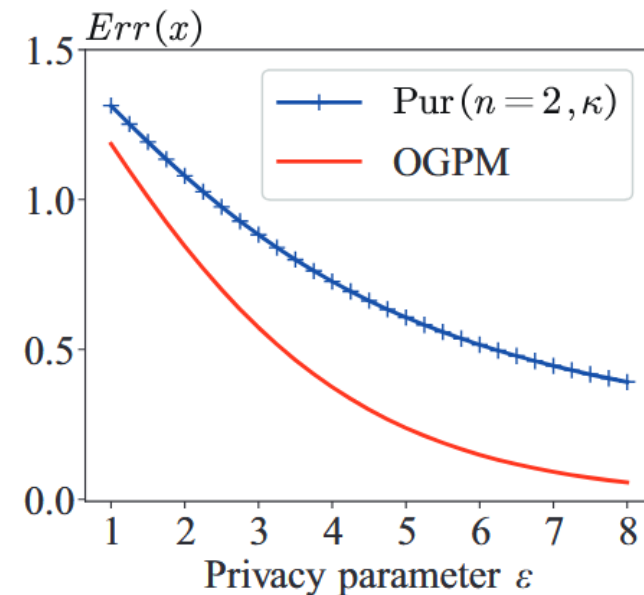
Whole-domain error ($\varepsilon = 2$)

PM-C, SW-C on the **flatten** domain



Worst-case error

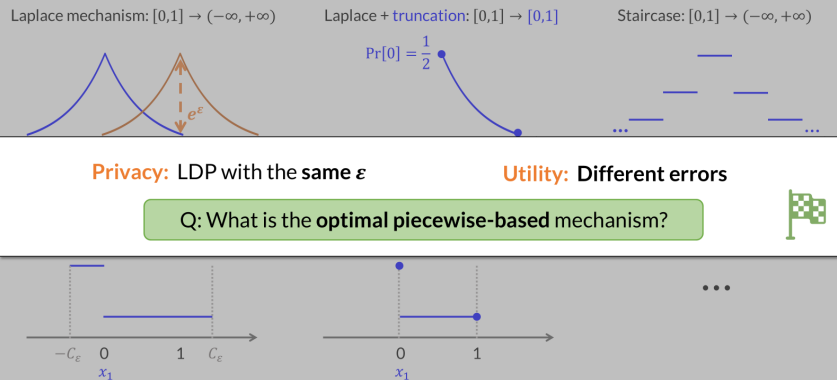
Purkayastha mechanism*



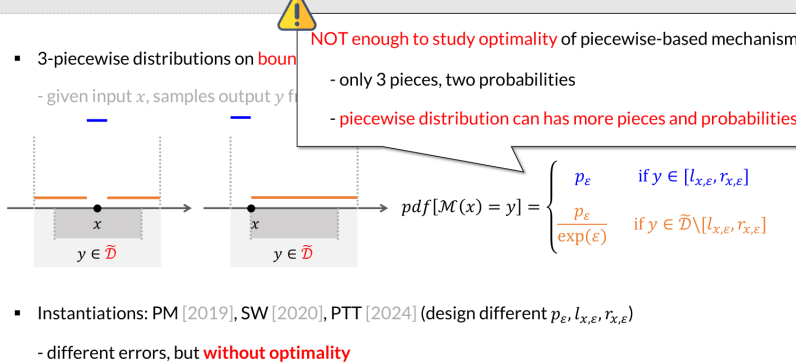
* Differential Privacy for Directional Data, [CCS'21]

Optimal Piecewise-based Mechanism for Collecting Bounded Numerical Data under LDP

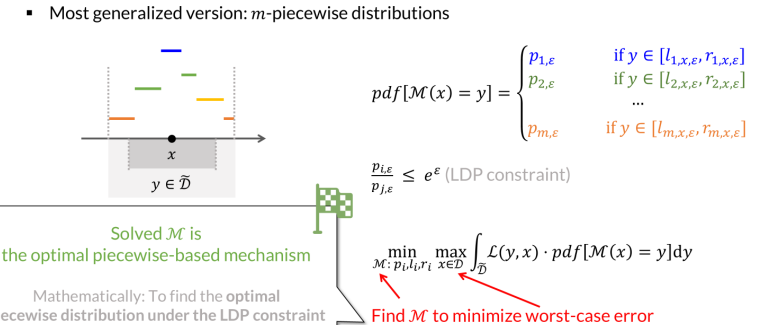
LDP Mechanisms for $\mathcal{D} = [0,1]$



3-Piecewise Mechanism



Optimal Piecewise-based Mechanism



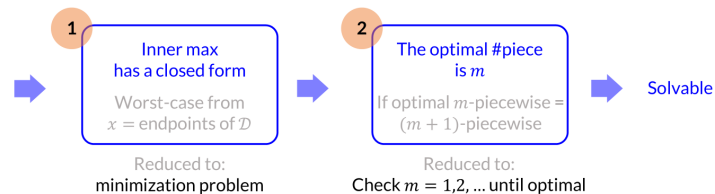
Challenges & Proofs

Challenges

- min-max problem & multiple variables
- optimal results only for a specific m

$$\min_{\mathcal{M}: p_i, l_i, r_i} \max_{x \in \mathcal{D}} \int_{\mathcal{D}} \mathcal{L}(y, x) \cdot pdf[\mathcal{M}(x) = y] dy$$

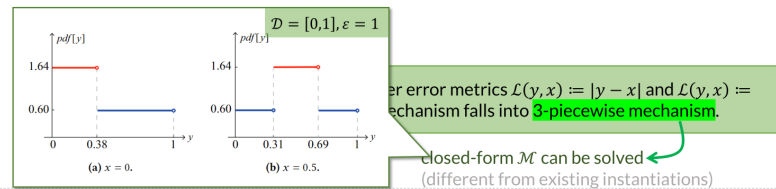
2. $i \in [m]$ 1. variables p_i, l_i, r_i



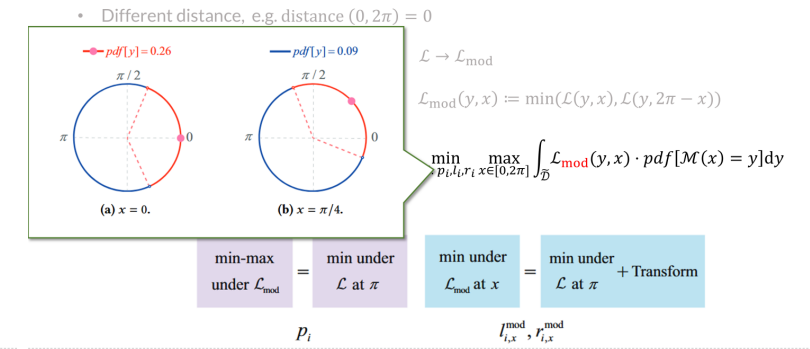
NOT Manually "Solvable" When $m \geq 4$

- Too many variables & non-linear
- Efficiently solved by off-the-shelf solvers, e.g. Gurobi
- limitation: needs given ϵ
- limitation: cannot provide closed-form $\mathcal{M}: p_i, l_i, r_i$

$$\max_{x \in [a,b]} \sum_{i=1}^m p_i \int_{l_i}^{r_i} \mathcal{L}(y, x) dy$$



Circular Domain



Thank you!



Optimality of LDP Mechanisms

- Optimal error (utility) under privacy level ε
 - many mechanisms are optimal in **order-of-magnitude**, e.g. $\Omega(\frac{1}{\sqrt{n}})$ for the counting query*
 - the staircase mechanism is optimal for **domain** $[0,1] \rightarrow (-\infty, +\infty)^\dagger$
 - the geometric mechanism is universally optimal if any **post-processing** is allowed, e.g. truncation^{††}

* The Complexity of Differential Privacy, 2017

† The Staircase Mechanism in Differential Privacy, 2016

†† Universally Utility-maximizing Privacy Mechanisms, 2009

Optimality of LDP Mechanisms

- Optimal error (utility) under privacy level ε
 - many mechanisms are optimal in **order-of-magnitude**, e.g. $\Omega(\frac{1}{\sqrt{n}})$ for the counting query*
 - the staircase mechanism is optimal for **domain** $[0,1] \rightarrow (-\infty, +\infty)^\dagger$
 - the geometric mechanism is universally optimal if any **post-processing** is allowed, e.g. truncation^{††}
- **Specify the utility model** (conditions for optimality)

1

Error metric

$Err(\text{truth}, \text{rand})$

Err or $\Omega(Err)$

2

Data domain &
type of mechanisms

Discrete / cont. $\mathcal{D} \rightarrow \tilde{\mathcal{D}}$

Laplace-shape / piecewise

3

Post-processing

Laplace + truncation: $[0,1] \rightarrow [0,1]$

$\Pr[0] = \frac{1}{2}$

