The Optimal Mechanism in Differential Privacy

Quan Geng

Advisor: Prof. Pramod Viswanath

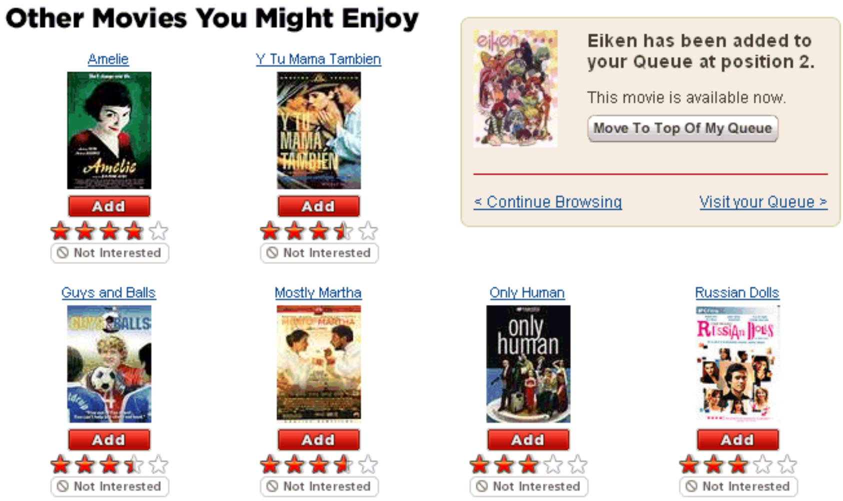
Outline

- Background on Differential Privacy
- ϵ-Differential Privacy: Single Dimensional Setting
- ϵ-Differential Privacy: Multiple Dimensional Setting
- (ϵ, δ)-Differential Privacy
- Conclusion

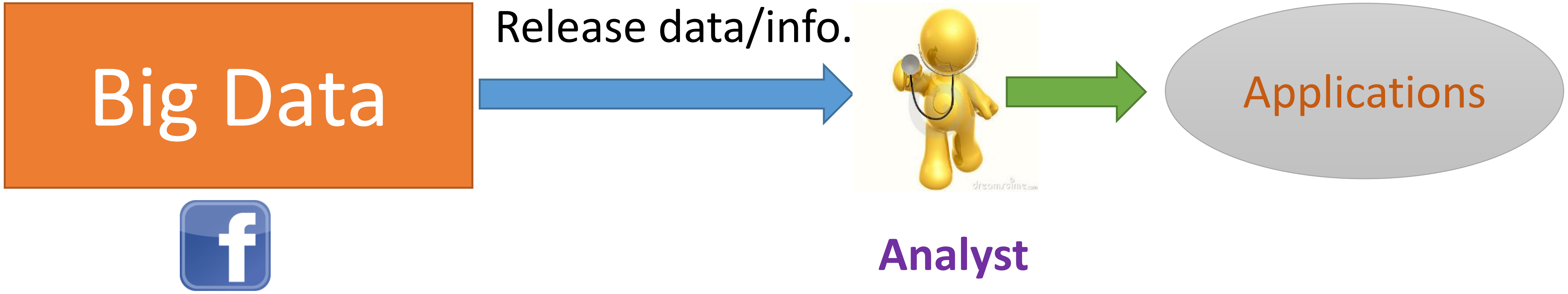
Motivation

- Vast amounts of personal information are collected

- Data analysis produces a lot of useful applications



Motivation



- How to release the data while protecting individual's privacy?

Motivation

- How to protect PRIVACY resilient to attacks with arbitrary side information?

Motivation

- How to protect PRIVACY resilient to attacks with arbitrary side information?
    - **Answer: randomized releasing mechanism**

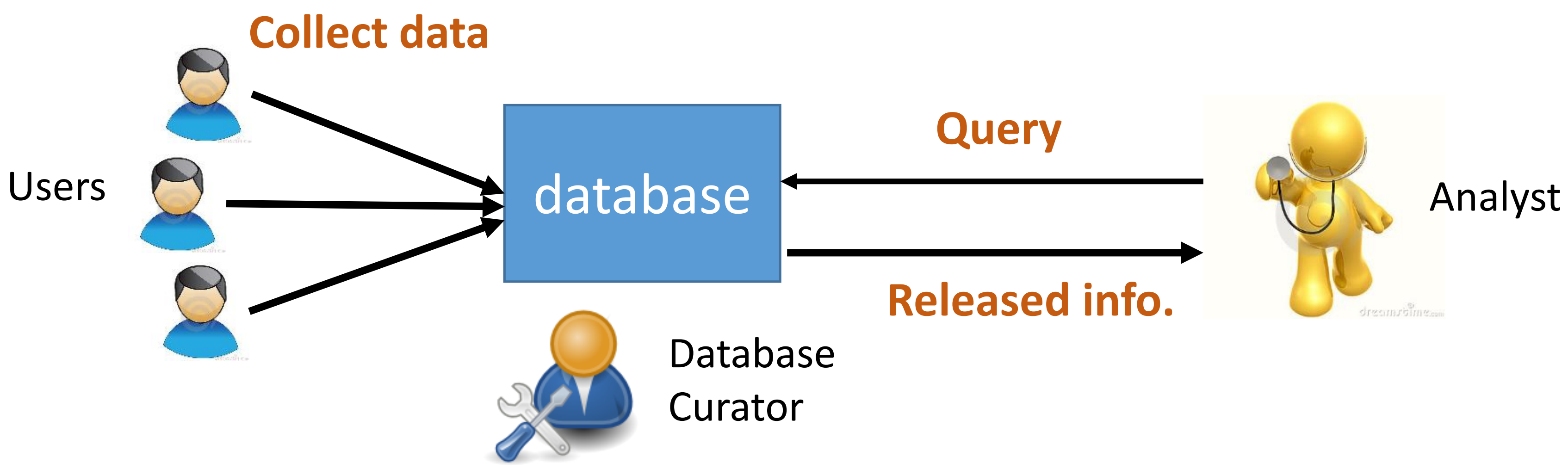
Motivation

- How to protect PRIVACY resilient to attacks with arbitrary side information?
    - **Answer: randomized releasing mechanism**
- How much randomness is needed?
    - completely random (no utility)
    - deterministic (no privacy)

Motivation

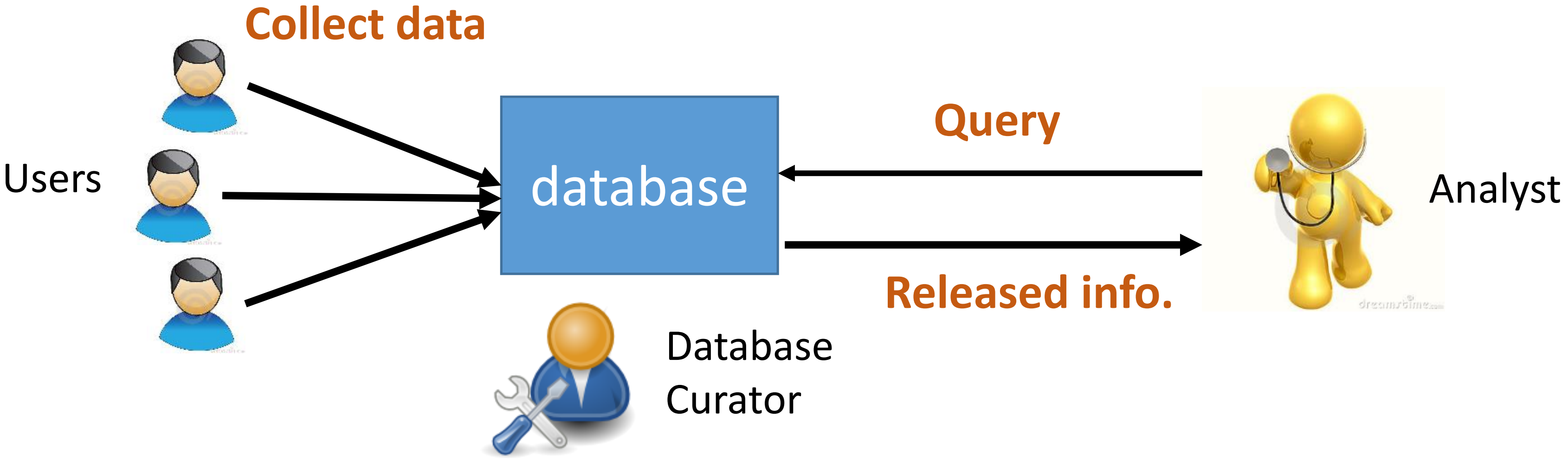
- How to protect PRIVACY resilient to attacks with arbitrary side information?
    - **Answer: randomized releasing mechanism**
- How much randomness is needed?
    - completely random (no utility)
    - deterministic (no privacy)
- **Differential Privacy** [Dwork et. al. 06]: one way to quantify the level of randomness and privacy

Background on Differential Privacy



- **Database Curator**: How to **answer** a query, providing **useful** data information to the analyst, while still protecting the **privacy** of each user.

Background on Differential Privacy



dataset: D

Age
A: 20
B: 35
C: 43
D: 30

query function: q (D)

q: How many people are older than 32?

q(D) = 2

randomized released mechanism: $\mathrm{K(D)}$

K(D) = 1 w.p. 1/8
K(D) = 2 w.p. 1/2
K(D) = 3 w.p. 1/4
K(D) = 4 w.p. 1/8

Background on Differential Privacy

- **Two neighboring datasets**: differ only at one element

Age
A: 20
B: 35
C: 43
D: 30

D_1

Age
A: 20
B: 35
D: 30

D_2

Background on Differential Privacy

- **Two neighboring datasets**: differ only at one element

Age
A: 20
B: 35
C: 43
D: 30

D_1

Age
A: 20
B: 35
D: 30

D_2

$$K(D_1) \approx K(D_2)$$

Differential Privacy: **presence** or **absence** of any individual record in the dataset should not affect the released information significantly.

Background on Differential Privacy

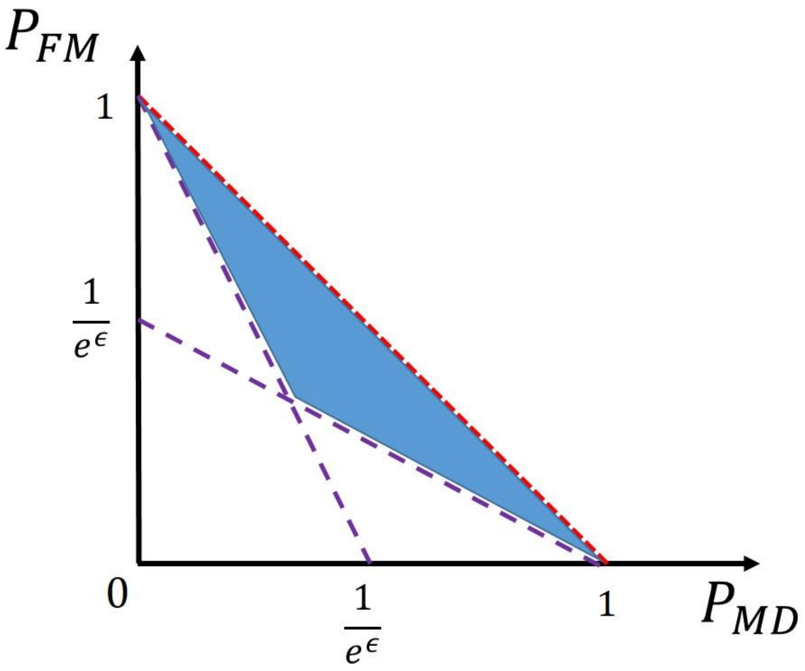
- A randomized mechanism K gives **ϵ-differential privacy**, if for any two neighboring datasets D_1, D_2 , and all S $\subset$Range(K),

$$\mathbf{Pr(K(D_1) \in S) \leq e^{\epsilon}\, Pr(K(D_2) \in S)}$$

Background on Differential Privacy

- A randomized mechanism K gives **ϵ-differential privacy**, if for any two neighboring datasets D_1, D_2 , and all S ⊂Range(K),

$$\mathbf{Pr(K(D_1) \in S) \leq e^{\epsilon}\, Pr(K(D_2) \in S)}$$



Make hypothesis testing hard:

$$e^{\epsilon} P_{FA} + P_{MD} \geq 1$$
$$P_{FA} + e^{\epsilon} P_{MD} \geq 1$$

Background on Differential Privacy

- A randomized mechanism K gives **ϵ-differential privacy**, if for any two neighboring datasets D_1, D_2 , and all $S \subset$Range(K),

$$\mathbf{Pr(K(D_1) \in S) \leq e^{\epsilon}\, Pr(K(D_2) \in S)}$$

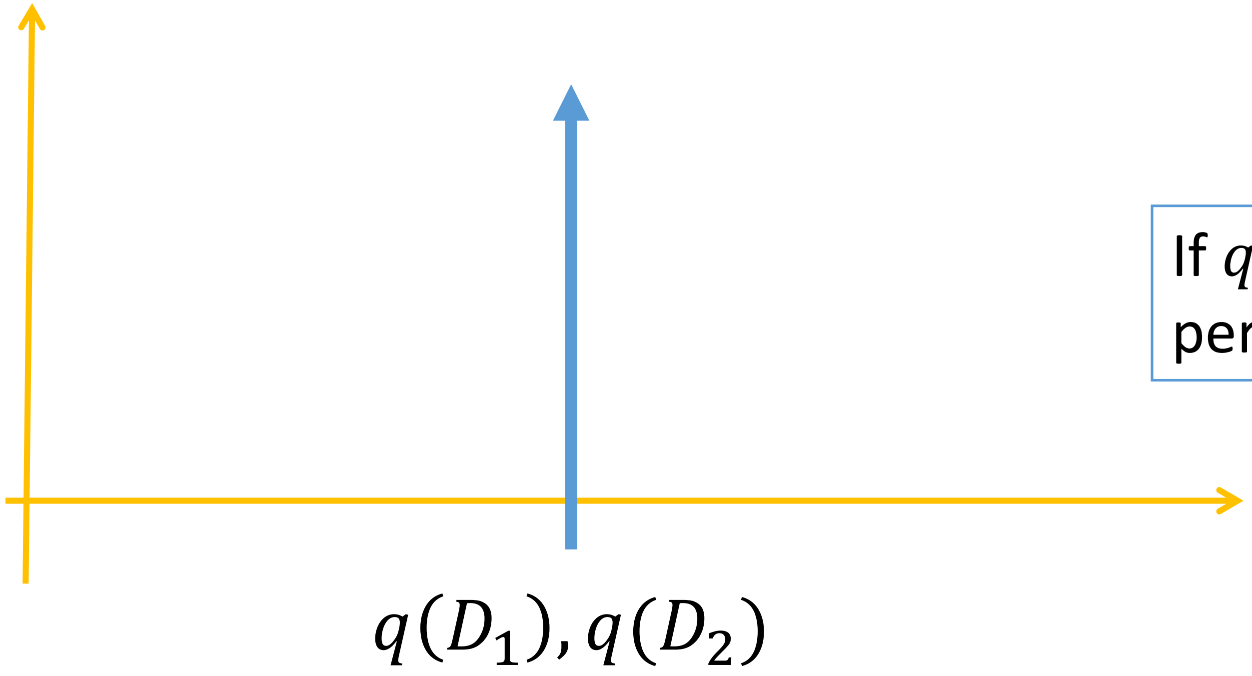
- $\boldsymbol{\epsilon}$ quantifies the level of privacy
  - $\epsilon \rightarrow 0$, high privacy
  - $\epsilon \rightarrow +\infty$, low privacy
  - a **social** question to choose $\boldsymbol{\epsilon}$ (can be 0.01, 0.1, 1, 10 …)

Background on Differential Privacy

- Q: How much perturbation needed to achieve ϵ-DP?
- A: depends on how different $q(D_1), q(D_2)$ are

Background on Differential Privacy

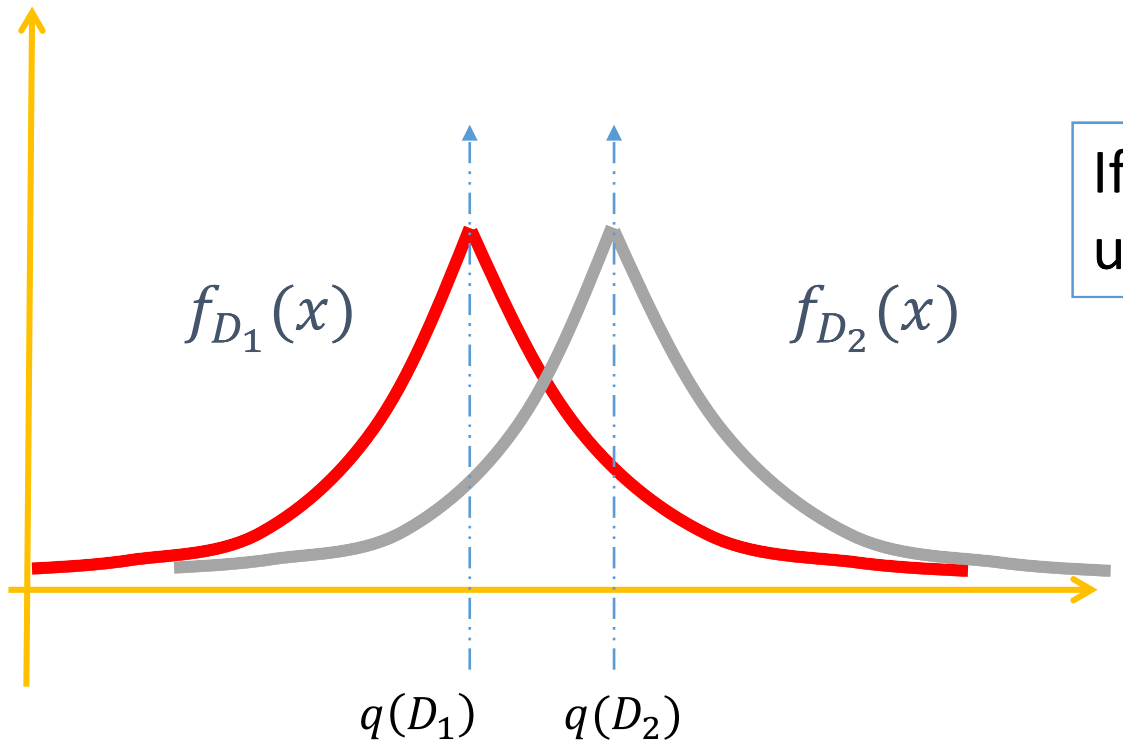
- Q: How much perturbation needed to achieve ϵ-DP?
- A: depends on how different $q(D_1), q(D_2)$ are



If $q(D_1) = q(D_2)$, no perturbation is needed.

Background on Differential Privacy

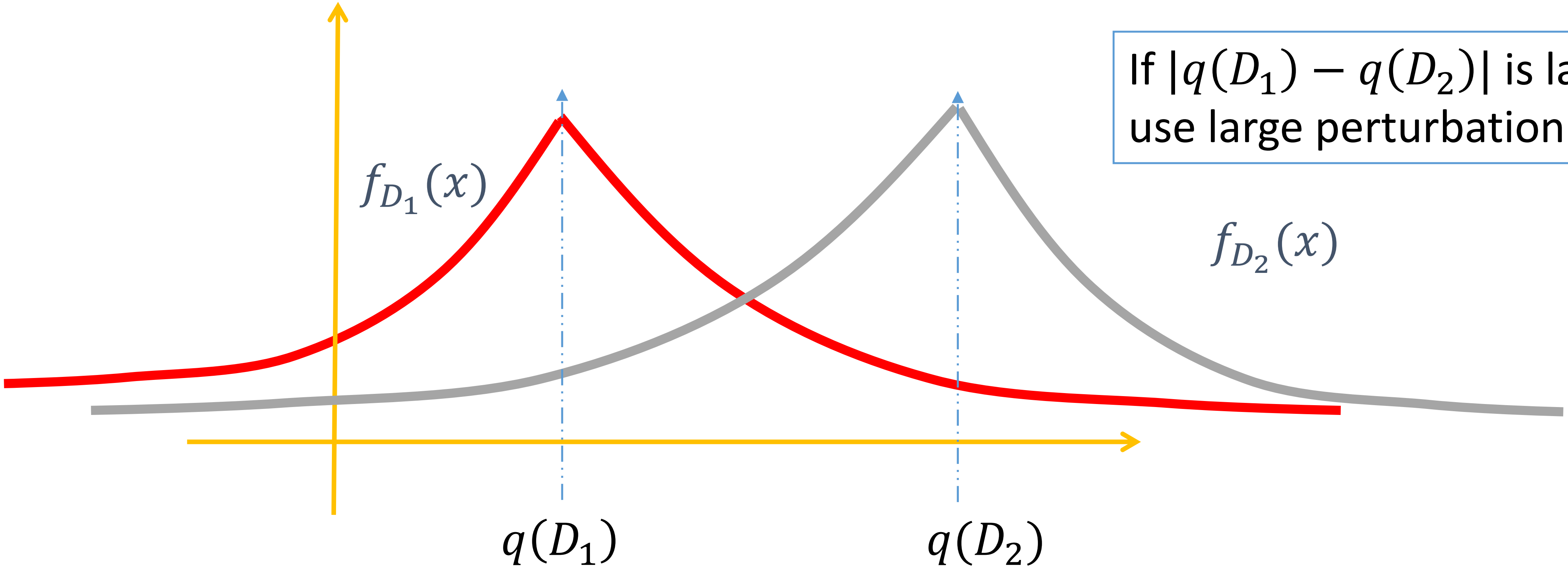
- Q: How much perturbation needed to achieve ϵ-DP?
- A: depends on how different $q(D_1), q(D_2)$ are



If $|q(D_1) - q(D_2)|$ is small, use small perturbation

Background on Differential Privacy

- Q: How much perturbation needed to achieve ϵ-DP?
- A: depends on how different $q(D_1), q(D_2)$ are



If $|q(D_1) - q(D_2)|$ is large, use large perturbation

Background on Differential Privacy

- **Global Sensitivity** Δ: how different when q is applied to neighboring datasets

$$\boldsymbol{\Delta} := \max_{(D_1, D_2)\ are\ neighbors} |q(D_1) - q(D_2)|$$

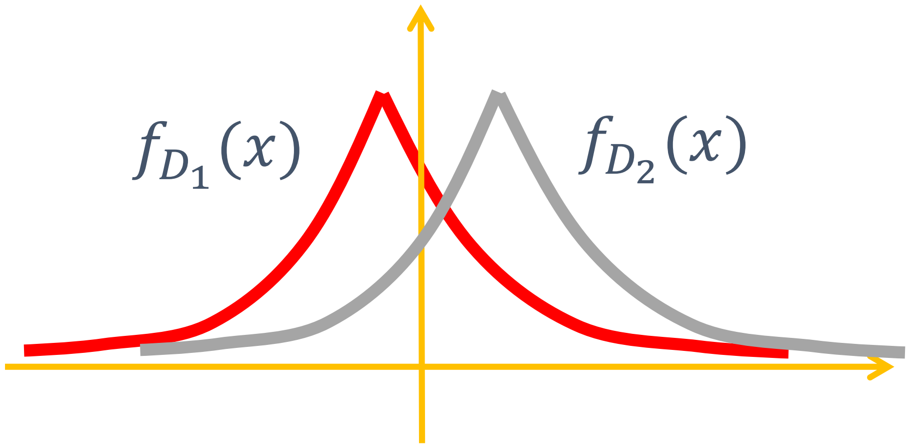
- Example: for a count query, $\Delta = 1$

Background on Differential Privacy

- Laplace Mechanism:

$$K(D) = q(D) + Lap(\frac{\Delta}{\epsilon}),$$

$Lap\left(\frac{\Delta}{\epsilon}\right)$ is a r.v. with p.d.f $f(x) = \frac{1}{2\lambda} e^{-\frac{|x|}{\lambda}}, \ \lambda = \frac{\Delta}{\epsilon}$



- Basic tool in DP

Optimality of Existing work?

- Laplace Mechanism:

$$\boldsymbol{K(D) = q(D) + Lap(\frac{\Delta}{\epsilon})},$$

- Two Questions:
  - Is **data-independent** perturbation optimal?
  - Assume data-independent perturbation, is **Laplacian** distribution optimal?

Optimality of Existing work?

- Laplace Mechanism:

$$K(D) = q(D) + Lap(\frac{\Delta}{\epsilon}),$$

- Two questions:
  - Is **data-independent** perturbation optimal?
  - Assume data-independent perturbation, is **Laplacian** distribution optimal?

- Our results:
  - **data-independent** perturbation is optimal
  - **Laplacian** distribution is not optimal: the optimal is **staircase** distribution

Part One:

The Optimal Mechanism in ϵ-Differential Privacy:
Single Dimensional Setting

presented in my preliminary exam

Recap of Part One: Problem Formulation

$$\textit{minimize} \sup_{t \in R} \int_{x \in R} L(x) \nu_t(dx)$$

$$\text{s.t.} \quad \boldsymbol{\nu}_{t_1}(S) \leq e^{\epsilon} \boldsymbol{\nu}_{t_2}(S + t_1 - t_2),$$

$$\forall \text{measurable set } S, \forall t_1, t_2 \in R, \text{s.t.} |t_1 - t_2| \leq \Delta,$$

$\boldsymbol{L}: R \to R$ cost function on the noise

$\boldsymbol{\nu}_t$ noise probability distribution given query output **t**

Δ global sensitivity

Recap of Part One: Main Results

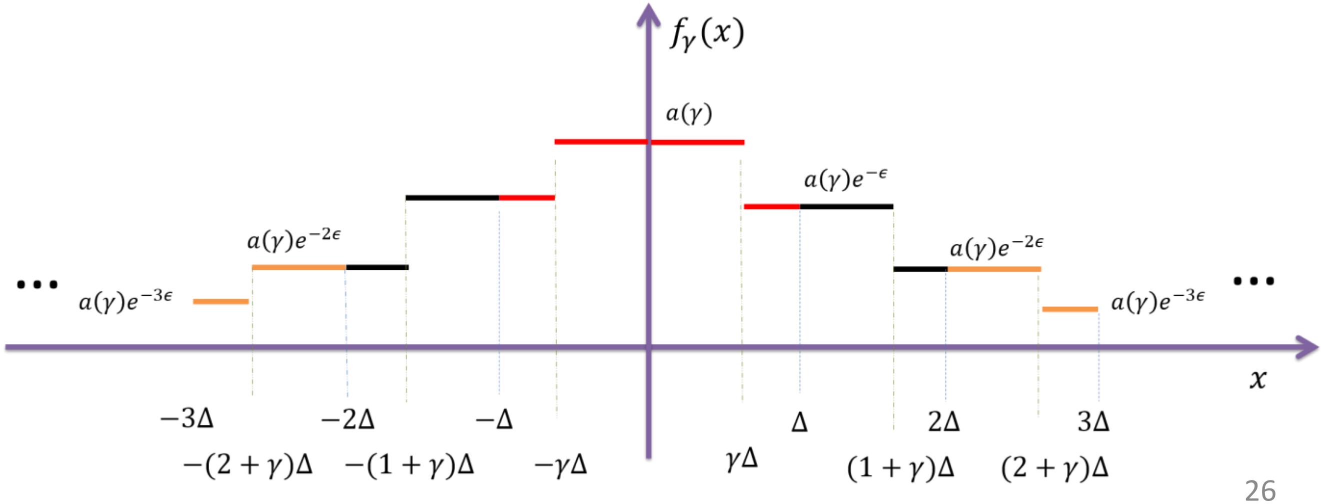
- **Optimality of query-output independent perturbation**

  In the optimal mechanism, ν_t is **independent** of t (under a technical condition).

- **Staircase Mechanism**

  Optimal noise probability distribution has staircase-shaped p.d.f.

$$\begin{aligned} &\text{Minimize} \quad \int L(x)P(dx) \\ \text{s.t.} \quad & \Pr(X \in S) \leq e^{\epsilon} \Pr(X \in S + d), \\ & \forall |d| \leq \Delta, \text{measurable set } S \end{aligned}$$



Recap of Part One: Applications

- ℓ^1: $\mathbf{L(x) = |x|}$

$$V(P_{\gamma^*}) = \Delta \frac{e^{\frac{\epsilon}{2}}}{e^{\epsilon} - 1}$$

$$V_{Lap} = \frac{\Delta}{\epsilon}$$

$\epsilon \to 0$, the additive gap $\to 0$

$\epsilon \to +\infty$, $V(P_{\gamma^*}) = \Theta(\Delta e^{-\frac{\epsilon}{2}})$

- ℓ^2: $\mathbf{L}(x) = \mathbf{x}^2$

$$V(\mathcal{P}_{\gamma^*}) = \Delta^2 \frac{2^{-2/3} b^{2/3} (1+b)^{2/3} + b}{(1-b)^2}.$$

$$V_{Lap} = \frac{2\Delta^2}{\epsilon^2}$$

$\epsilon \to 0$, the additive gap $\leq c\Delta^2$

$\epsilon \to +\infty$, $V(P_{\gamma^*}) = \Theta(\Delta^2\, e^{-\frac{2\epsilon}{3}})$

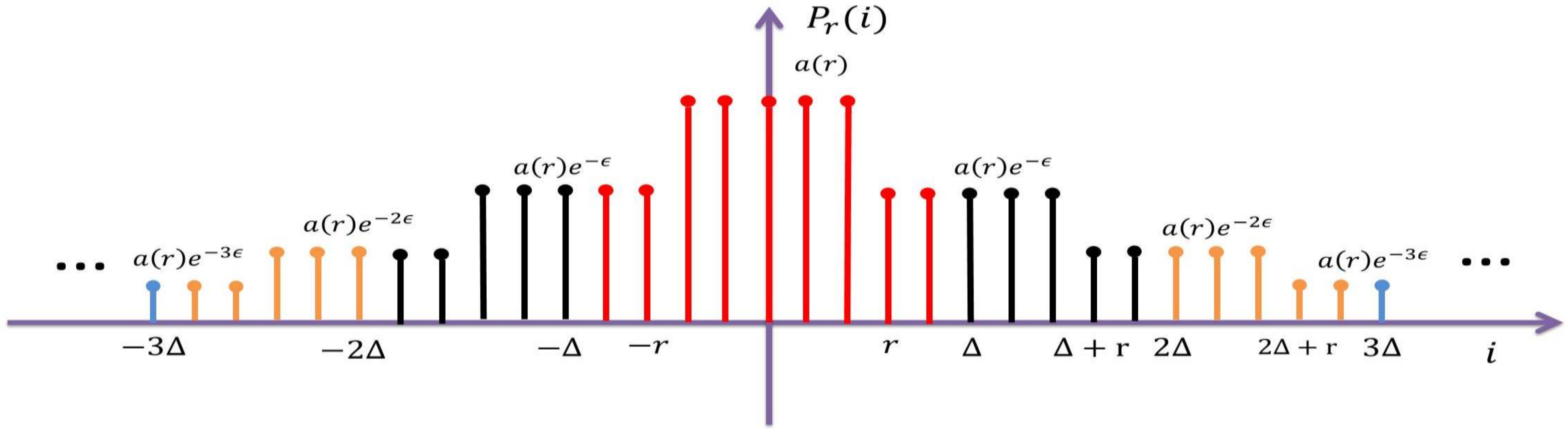
Huge improvement in **low** privacy regime

Recap of Part One: Applications

- Properties of γ^* (for $L(x) = |x|^m$)

$$\gamma^* \to \frac{1}{2},\ as\ \epsilon \to 0,$$
$$\gamma^* \to 0,\ as\ \epsilon \to +\infty.$$

- Extension to Discrete Setting



- Extension to Abstract Setting

Conclusion of Part One:

- Fundamental tradeoff between privacy and utility in ϵ-Differential Privacy
- Staircase Mechanism, optimal mechanism for single real-valued query
  - Huge improvement in low privacy regime
- Extension to discrete setting and abstract setting

Part Two:

The Optimal Mechanism in ϵ-Differential Privacy: Multiple Dimensional Setting

Multiple Dimensional Setting

- Query output can have multiple components

$$\boldsymbol{q}(\boldsymbol{D}) = \left(\boldsymbol{q}_1(\boldsymbol{D}), \boldsymbol{q}_2(\boldsymbol{D}), \dots, \boldsymbol{q}_{\boldsymbol{d}}(\boldsymbol{D})\right) \in \boldsymbol{R}^{\boldsymbol{d}}$$

Multiple Dimensional Setting

- Query output can have multiple components

$$\boldsymbol{q}(\boldsymbol{D}) = \left(\boldsymbol{q}_{\mathbf{1}}(\boldsymbol{D}), \boldsymbol{q}_{\mathbf{2}}(\boldsymbol{D}), \ldots, \boldsymbol{q}_{\boldsymbol{d}}(\boldsymbol{D})\right) \in \boldsymbol{R}^{\boldsymbol{d}}$$

- If all components are uncorrelated, perturb each component independently to preserve DP.

Multiple Dimensional Setting

- Query output can have multiple components

$$\boldsymbol{q}(\boldsymbol{D}) = \left(\boldsymbol{q_1}(\boldsymbol{D}), \boldsymbol{q_2}(\boldsymbol{D}), \dots, \boldsymbol{q_d}(\boldsymbol{D})\right) \in \boldsymbol{R^d}$$

- If all components are uncorrelated, perturb each component independently to preserve DP.

- Composition Theorem:

For each component $\boldsymbol{q_i}\ (\boldsymbol{D})$ preserving ϵ-DP, end-to-end achieves $\boldsymbol{d\epsilon}$-DP.

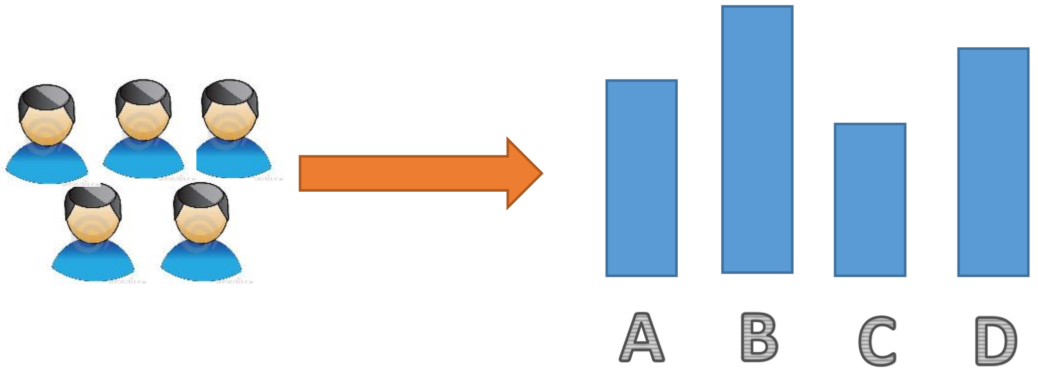
Multiple Dimensional Setting

- Query output can have multiple components

$$q(D) = \left(q_1(D), q_2(D), \dots, q_d(D)\right) \in R^d$$

- For some query, all components are coupled together.

Histogram query:
one user can affect **only one** component



Multiple Dimensional Setting

- Query output can have multiple components

$$\boldsymbol{q(D) = \left(q_1(D), q_2(D), \dots, q_d(D)\right) \in R^d}$$

- For some query, all components are coupled together.
- **Global sensitivity**

$$\boldsymbol{\Delta} := \max_{\boldsymbol{D_1, D_2\ are\ neighbors}} \boldsymbol{||q(D_1) - q(D_2)||_1}$$

Histogram query: **Δ** = 1

Problem Formulation:

Query-output independent perturbation:

$$\boldsymbol{K(D) = q(D) + \ \ X}$$
$$\boldsymbol{= (q_1\,(D) + x_1, q_2\,(D) + x_2, \ldots, q_d\,(D) + x_d)}$$

Problem Formulation:

Query-output independent perturbation:

$$\boldsymbol{K(D)} = \boldsymbol{q(D)} + \ \boldsymbol{X}$$
$$= (\boldsymbol{q_1}\,(\boldsymbol{D}) + \boldsymbol{x_1}, \boldsymbol{q_2}\,(\boldsymbol{D}) + \boldsymbol{x_2}, \dots, \boldsymbol{q_d}\,(\boldsymbol{D}) + \boldsymbol{x_d})$$

Cost function on the noise:

$$L: R^d \rightarrow R$$

Problem Formulation:

Query-output independent perturbation:

$$\boldsymbol{K(D) = q(D) + \ X}$$
$$\boldsymbol{= (q_1\,(D) + x_1, q_2\,(D) + x_2, \ldots, q_d\,(D) + x_d)}$$

Cost function on the noise:

$$L: R^d \rightarrow R$$

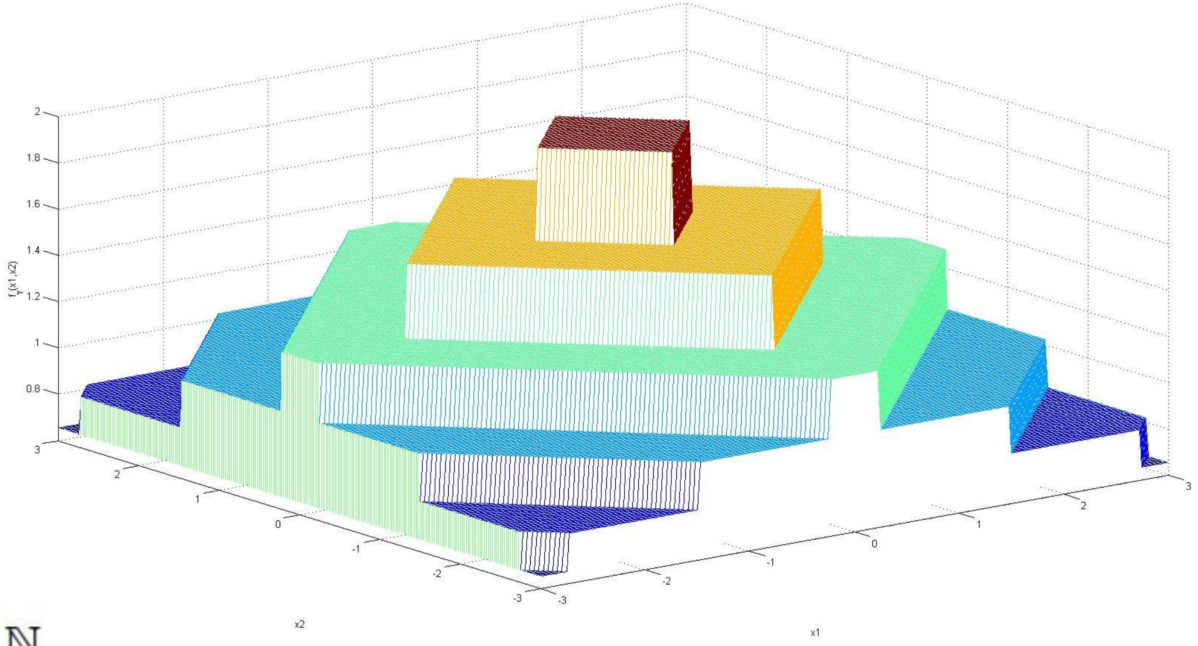
Optimization problem:

$$minimize \iiint_{R^d} L(x_1, \ldots, x_d) P(dx_1 \cdots dx_d)$$

s.t. $P(S) \leq e^{\epsilon} P(S + \boldsymbol{t}), \forall S \subset R^d, \|\boldsymbol{t}\|_1 \leq \Delta$

Our Result: Optimality of Staircase Mechanism

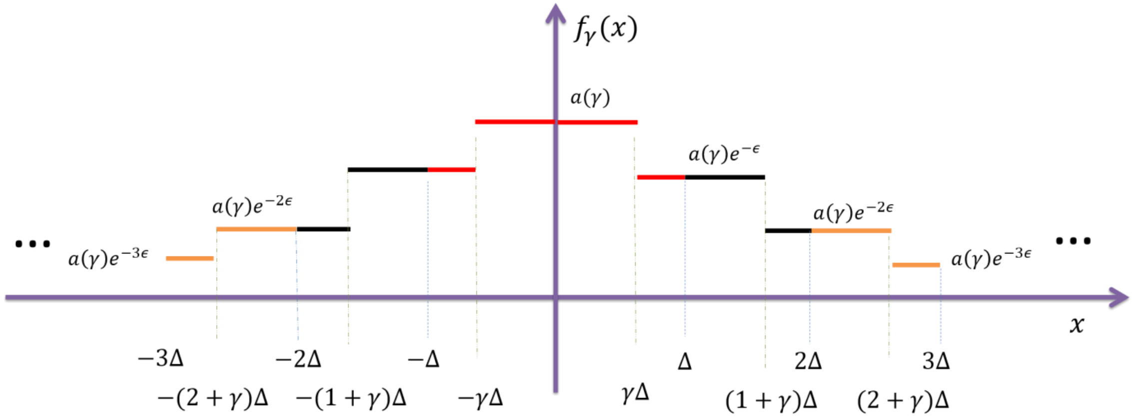
Theorem: If $L(x_1, \dots, x_d) = |x_1| + \cdots + |x_d|, d = 2$, then optimal probability distribution has multidimensional staircase-shaped p.d.f.

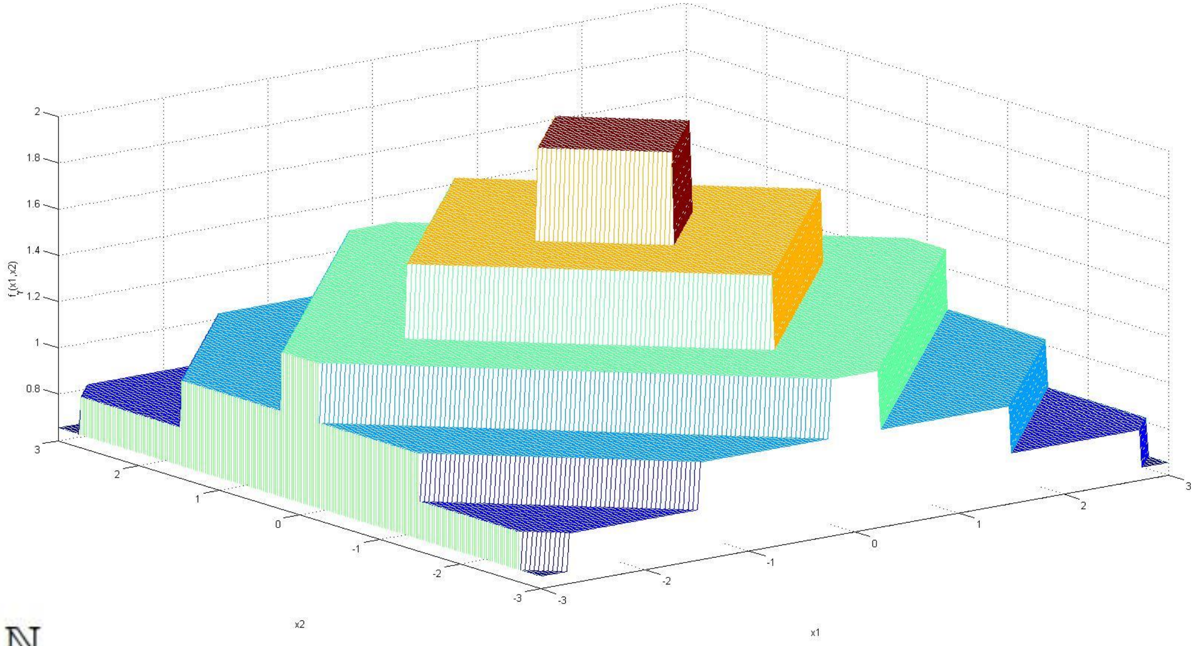


$$f_\gamma(\mathbf{x}) = \begin{cases} e^{-k\epsilon}a(\gamma) & \|\mathbf{x}\|_1 \in [k\Delta, (k+\gamma)\Delta) \text{ for } k \in \mathbb{N} \\ e^{-(k+1)\epsilon}a(\gamma) & \|\mathbf{x}\|_1 \in [(k+\gamma)\Delta, (k+1)\Delta) \text{ for } k \in \mathbb{N} \end{cases}$$

Our Result: Optimality of Staircase Mechanism

Theorem: If $L(x_1, \dots, x_d) = |x_1| + \cdots + |x_d|, d = 2$, then optimal probability distribution has multidimensional staircase-shaped p.d.f.

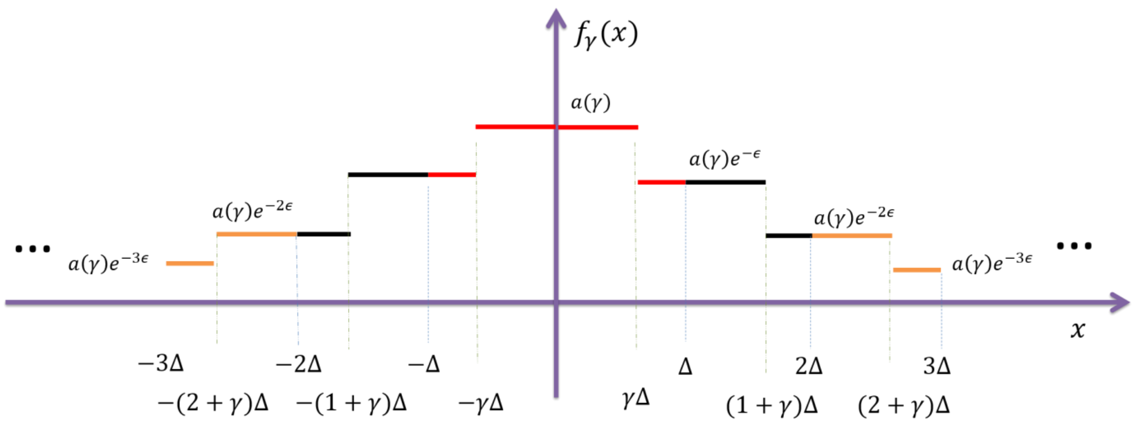


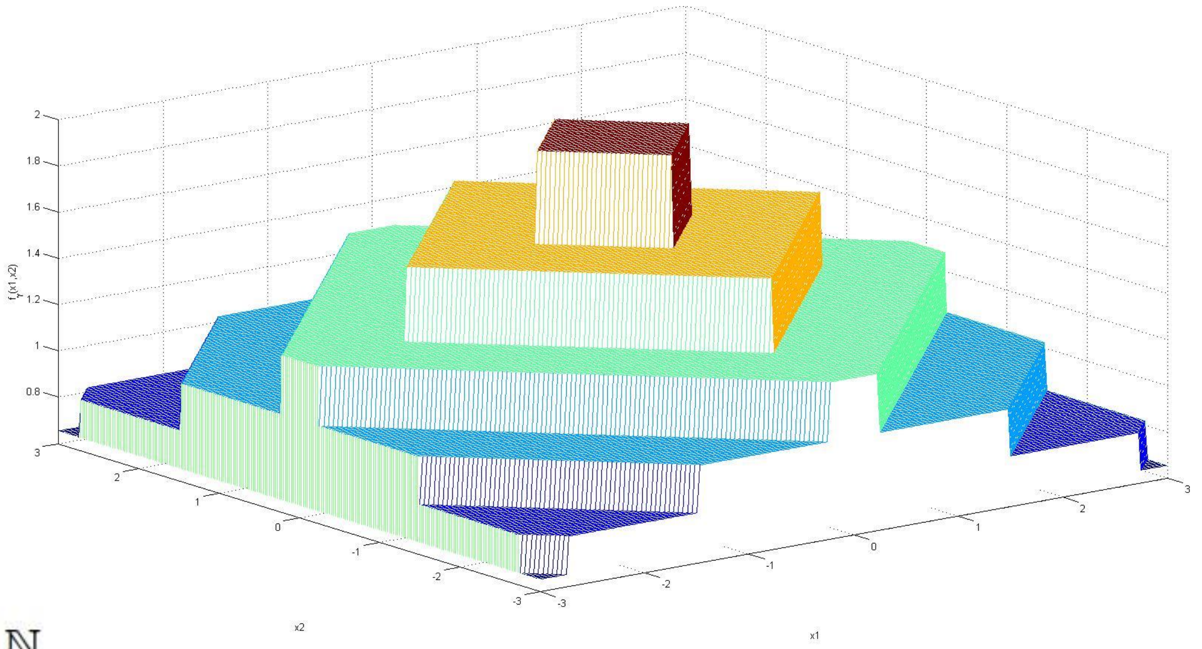


$$f_\gamma(\mathbf{x}) = \begin{cases} e^{-k\epsilon} a(\gamma) & \|\mathbf{x}\|_1 \in [k\Delta, (k+\gamma)\Delta) \text{ for } k \in \mathbb{N} \\ e^{-(k+1)\epsilon} a(\gamma) & \|\mathbf{x}\|_1 \in [(k+\gamma)\Delta, (k+1)\Delta) \text{ for } k \in \mathbb{N} \end{cases}$$

Our Result: Optimality of Staircase Mechanism

Theorem: If $L(x_1, \dots, x_d) = |x_1| + \cdots + |x_d|$, $d = 2$, then optimal probability distribution has multidimensional staircase-shaped p.d.f.



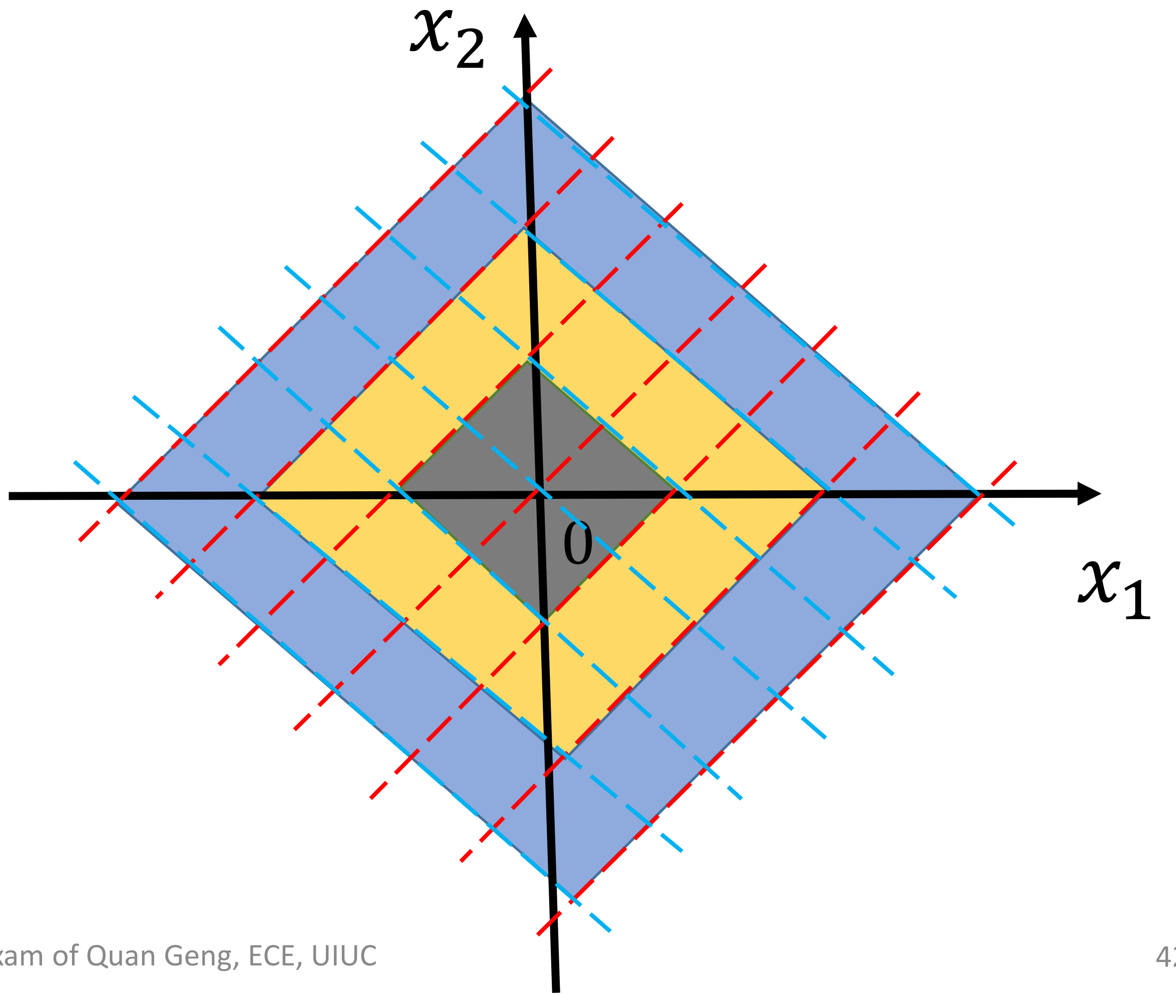


$$f_\gamma(\mathbf{x}) = \begin{cases} e^{-k\epsilon} a(\gamma) & \|\mathbf{x}\|_1 \in [k\Delta, (k+\gamma)\Delta) \text{ for } k \in \mathbb{N} \\ e^{-(k+1)\epsilon} a(\gamma) & \|\mathbf{x}\|_1 \in [(k+\gamma)\Delta, (k+1)\Delta) \text{ for } k \in \mathbb{N} \end{cases}$$

Conjecture: the result holds for any ***d***

Proof Ideas

Step 1: Argue only need to consider multi-dimensional piecewise constant p.d.f. by averaging



Averaging preserves DP

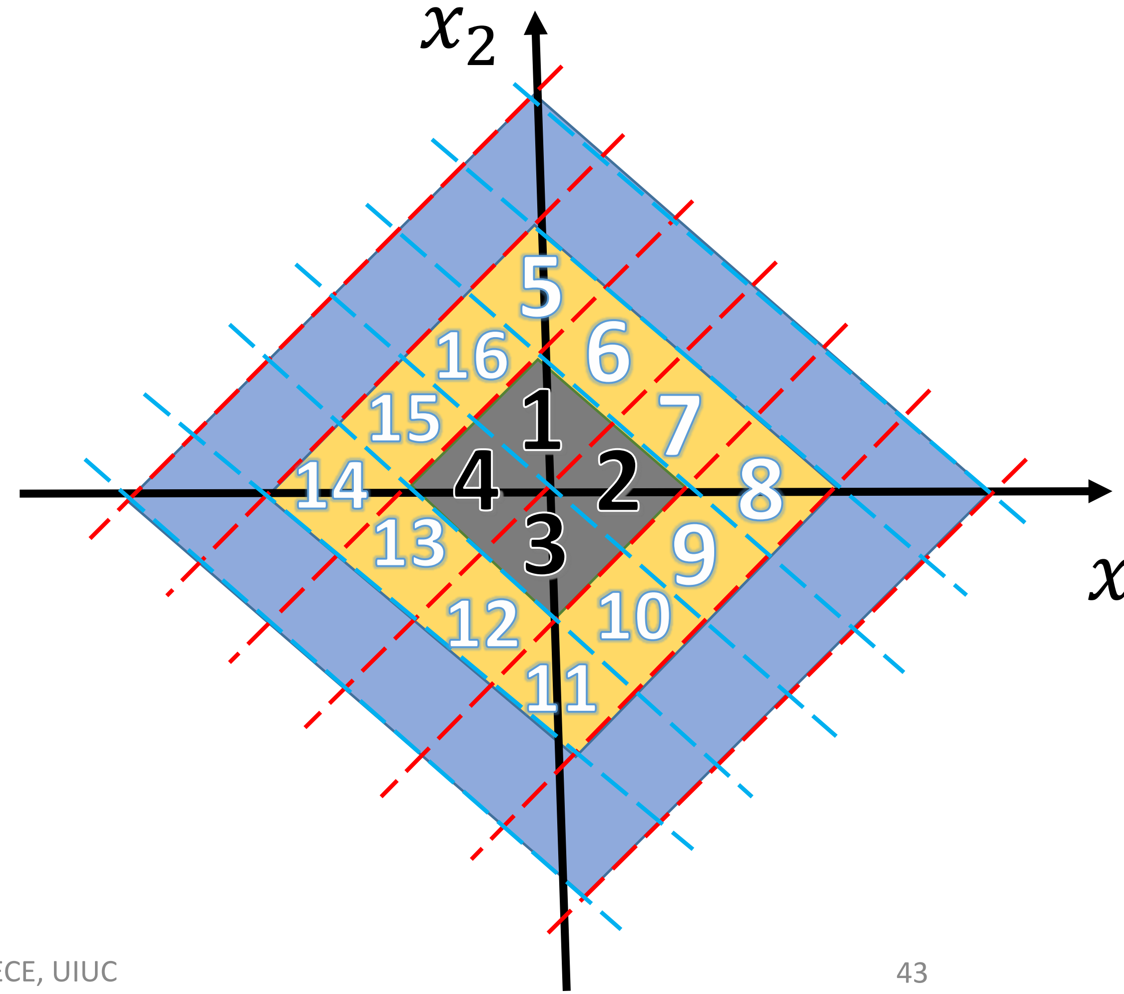
Proof Ideas

Step 1: Argue only need to consider multi-dimensional piecewise constant p.d.f. by averaging

Averaging preserves DP

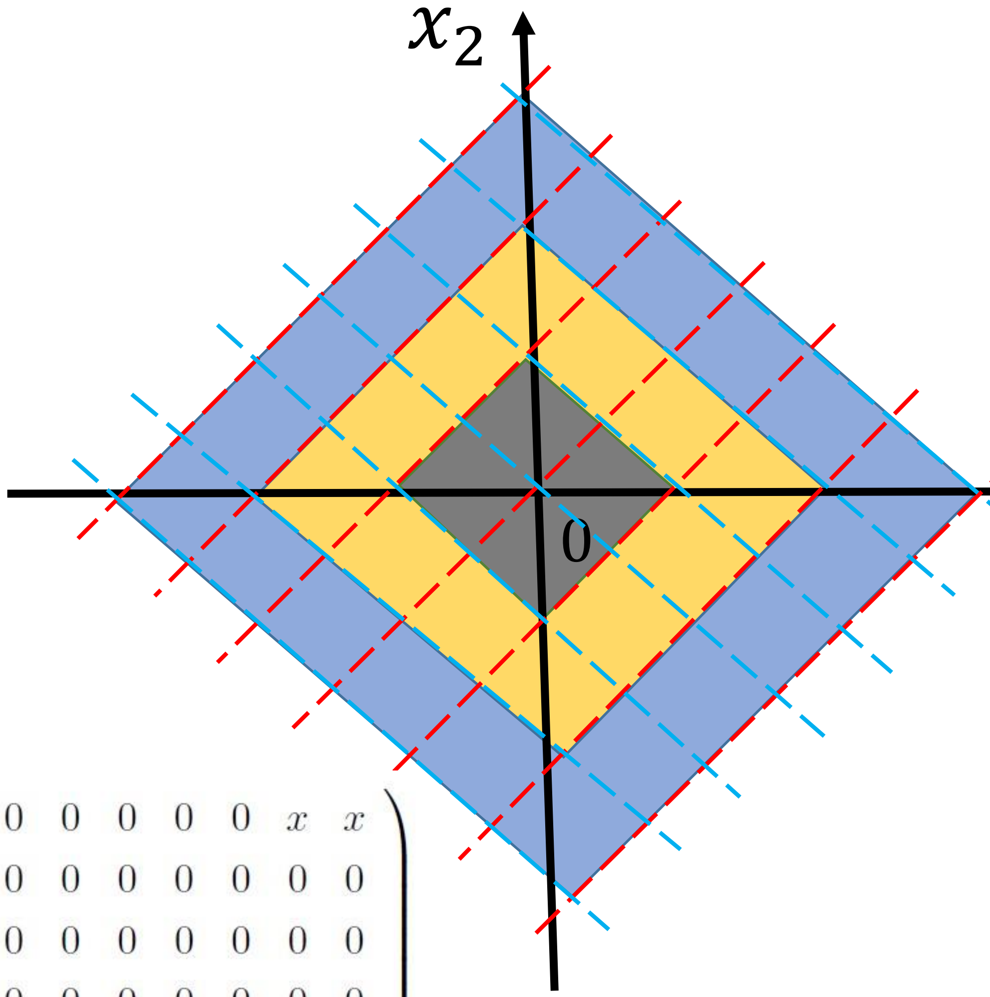
$$P1 \le e^{\epsilon}\, P5$$
$$P1 \le e^{\epsilon}\, P6$$
$$P1 \le e^{\epsilon}\, P16$$
$$P1 \le e^{\epsilon}\, P7$$
$$P1 \le e^{\epsilon}\, P15$$
$$\dots$$
$$\dots$$
$$P4 \le e^{\epsilon}\, P12$$
$$P4 \le e^{\epsilon}\, P13$$
$$P4 \le e^{\epsilon}\, P14$$
$$P4 \le e^{\epsilon}\, P15$$
$$P4 \le e^{\epsilon}\, P16$$

$$\Rightarrow \quad \frac{\sum_{i=1}^{4} Pi}{4} \le e^{\epsilon} \frac{\sum_{i=5}^{16} Pi}{12}$$



Proof Ideas

Averaging preserves DP



Formulate a matrix filling-in problem

- row sum is a constant
- column sum is a constant

Explicit formula for $d = 2$, and arbitrary Δ

$$\begin{pmatrix}
x & x & x & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x & x \\
0 & x & x & x & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & x & x & x & x & x & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & x & x & x & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & x & x & x & x & x & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x & x & x & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x & x & x & x & x & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x & x & x
\end{pmatrix}$$

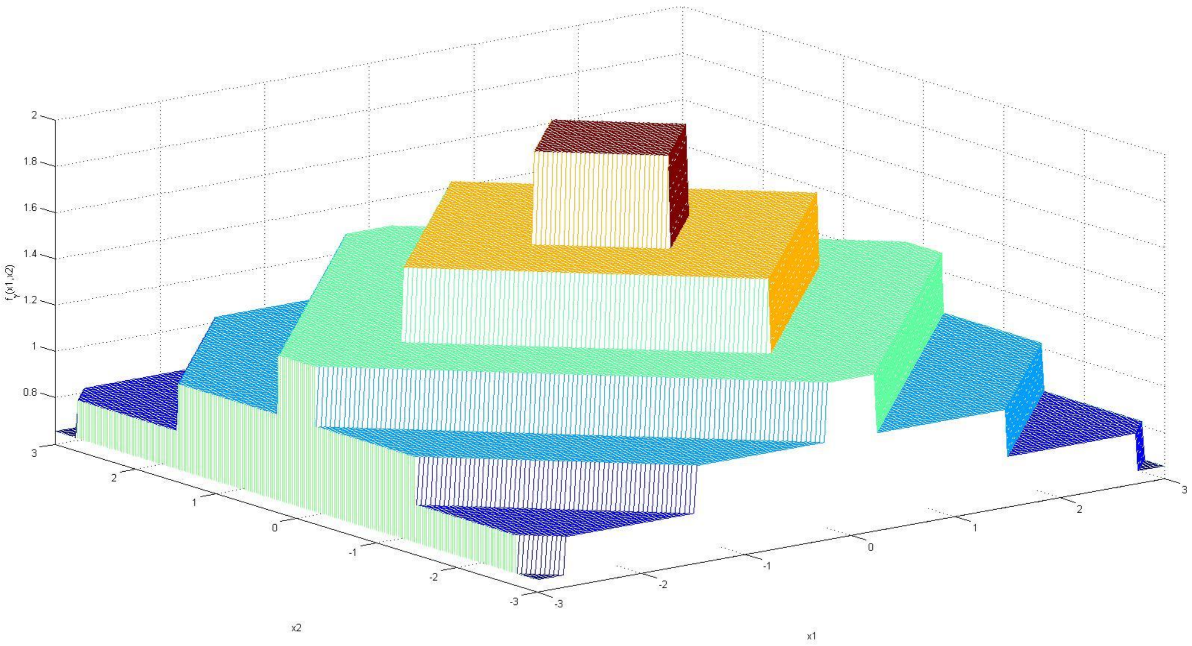
Proof Ideas

Due to Step 1, the p.d.f. is **a function of $\|X\|_1$**

Step 2: Monotonicity

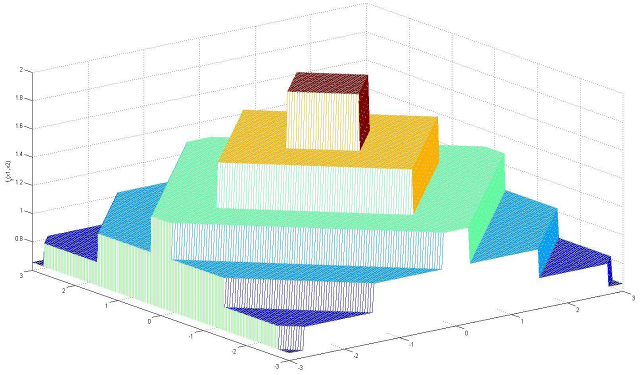
Step 3: Geometrically Decaying

Step 4: Staircase-Shape

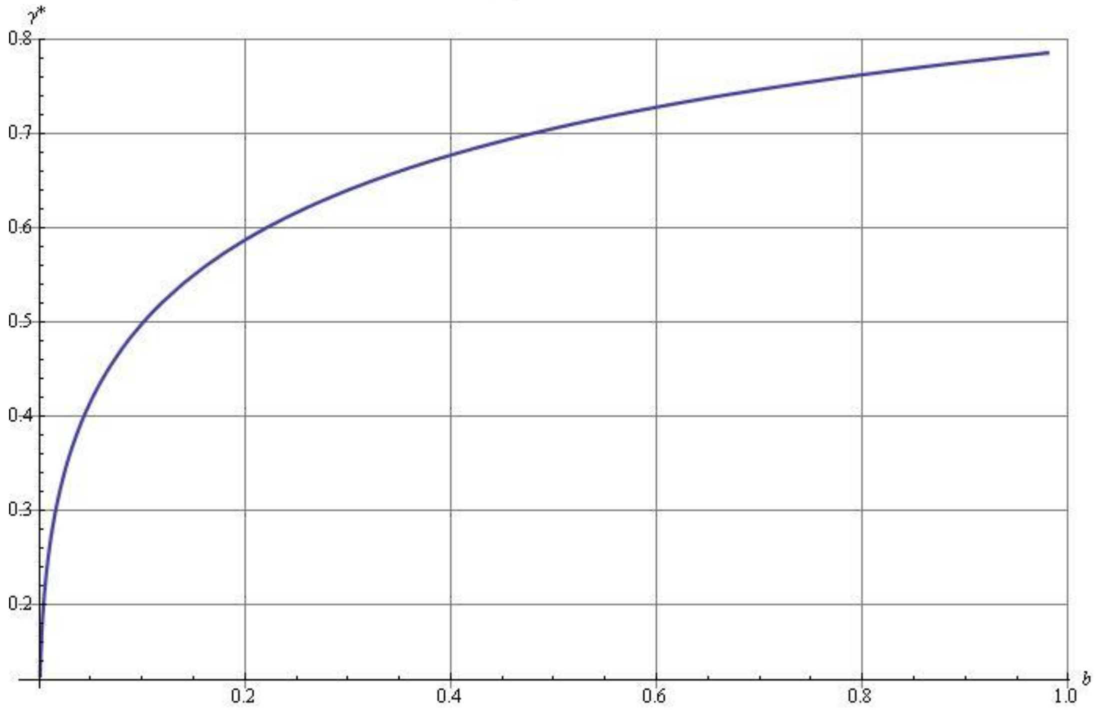


Optimal γ^* for ℓ^1 cost function (d=2)

$$f_\gamma(\mathbf{x}) = \begin{cases} e^{-k\epsilon}a(\gamma) & \|\mathbf{x}\|_1 \in [k\Delta, (k+\gamma)\Delta) \text{ for } k \in \mathbb{N} \\ e^{-(k+1)\epsilon}a(\gamma) & \|\mathbf{x}\|_1 \in [(k+\gamma)\Delta, (k+1)\Delta) \text{ for } k \in \mathbb{N} \end{cases}$$



$$\gamma^* = \underset{\gamma\in[0,1]}{\arg\min} \frac{\gamma^3 + \frac{3b}{1-b}\gamma^2 + \frac{3(b^2+b)}{(1-b)^2}\gamma + b\frac{1+4b+b^2}{(1-b)^3}}{\gamma^2 + \frac{2b}{1-b}\gamma + \frac{b+b^2}{(1-b)^2}}$$



Asymptotic analysis

Laplacian Mechanism: $V = \frac{2\Delta}{\epsilon}$

Multidimensional Staircase Mechanism:

High privacy regime($\epsilon \to 0$):

$$V^* = \frac{2\Delta}{\epsilon} - \frac{\Delta\epsilon^2}{36\sqrt{3}} + O(\epsilon^3)$$

Low privacy regime($\epsilon \to \infty$):

$$V^* = \sqrt[3]{2}\Delta e^{-\frac{\epsilon}{3}} + \frac{\Delta e^{-\frac{2\epsilon}{3}}}{\sqrt[3]{2}} + o(e^{-\frac{2\epsilon}{3}})$$

Asymptotic analysis

Laplacian Mechanism: $\boldsymbol{V = \frac{2\Delta}{\epsilon}}$

Multidimensional Staircase Mechanism:

High privacy regime($\epsilon \to 0$):

$$V^* = \frac{2\Delta}{\epsilon} - \frac{\Delta\epsilon^2}{36\sqrt{3}} + O(\epsilon^3)$$

Low privacy regime($\epsilon \to \infty$):

$$V^* = \sqrt[3]{2}\Delta e^{-\frac{\epsilon}{3}} + \frac{\Delta e^{-\frac{2\epsilon}{3}}}{\sqrt[3]{2}} + o(e^{-\frac{2\epsilon}{3}})$$

Independent Staircase Mechanism:

$$V = \Theta(2\Delta e^{-\frac{\epsilon}{4}})$$

Conclusion of Part Two:

- Extension of Optimality of Staircase Mechanism to Multi-dimensional Setting ($d = 2$)
  - Conjecture: holds for any dimension
- Approximate Optimality of Laplacian Mechanism in High Privacy Regime
- Huge Improvement in Low Privacy Regime

Part Three:

The Optimal Mechanism in (ϵ, δ)-Differential Privacy

Part Three: (ϵ,δ)-differential privacy

- (ϵ, δ)-differential privacy

$$\mathbf{Pr(K(D_1) \in S) \leq e^{\epsilon} \, Pr(K(D_2) \in S) + \delta}$$

- Two special cases:
  - $(\epsilon, 0)$-differential privacy
    - well studied in Part 1 and Part 2
  - $(0, \delta)$-differential privacy

$$\|P_{K(D_1)} - P_{K(D_2)}\|_{\mathrm{TV}} \leq \delta$$

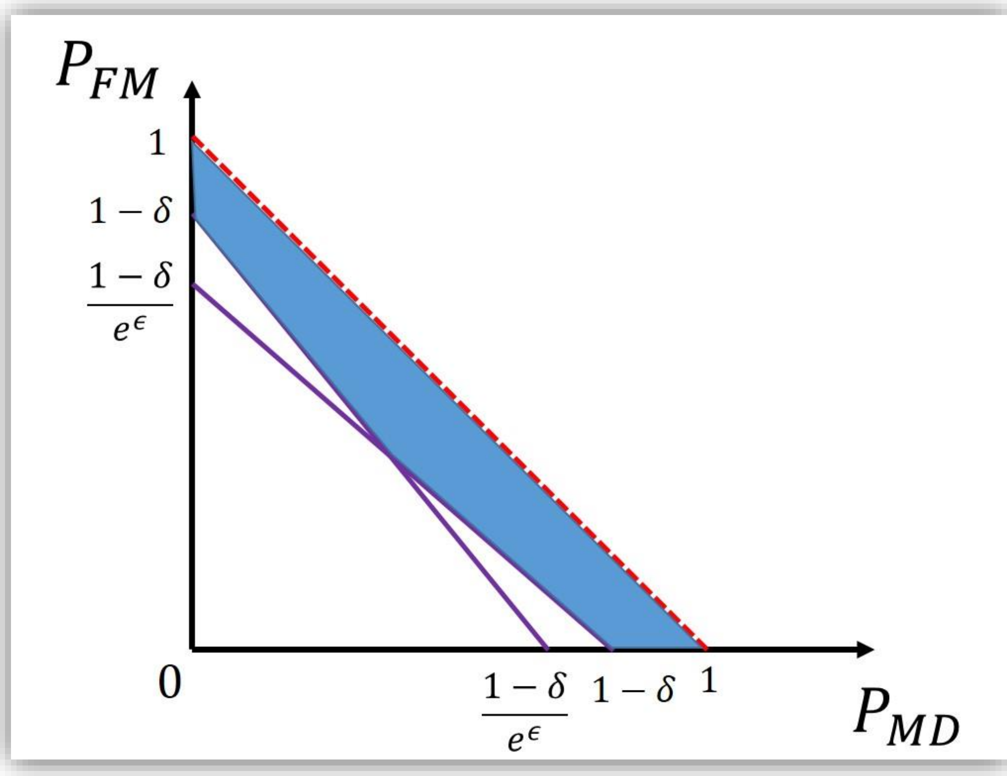
(ϵ,δ)-Differential Privacy

- (ϵ, δ)-differential privacy

$$\mathbf{Pr(K(D_1) \in S) \le e^{\epsilon} Pr(K(D_2) \in S) + \delta}$$

- Make hypothesis testing hard:

$$e^{\epsilon} P_{FA} + P_{MD} \geq 1 - \delta$$
$$P_{FA} + e^{\epsilon} P_{MD} \geq 1 - \delta$$



(ϵ,δ)-Differential Privacy

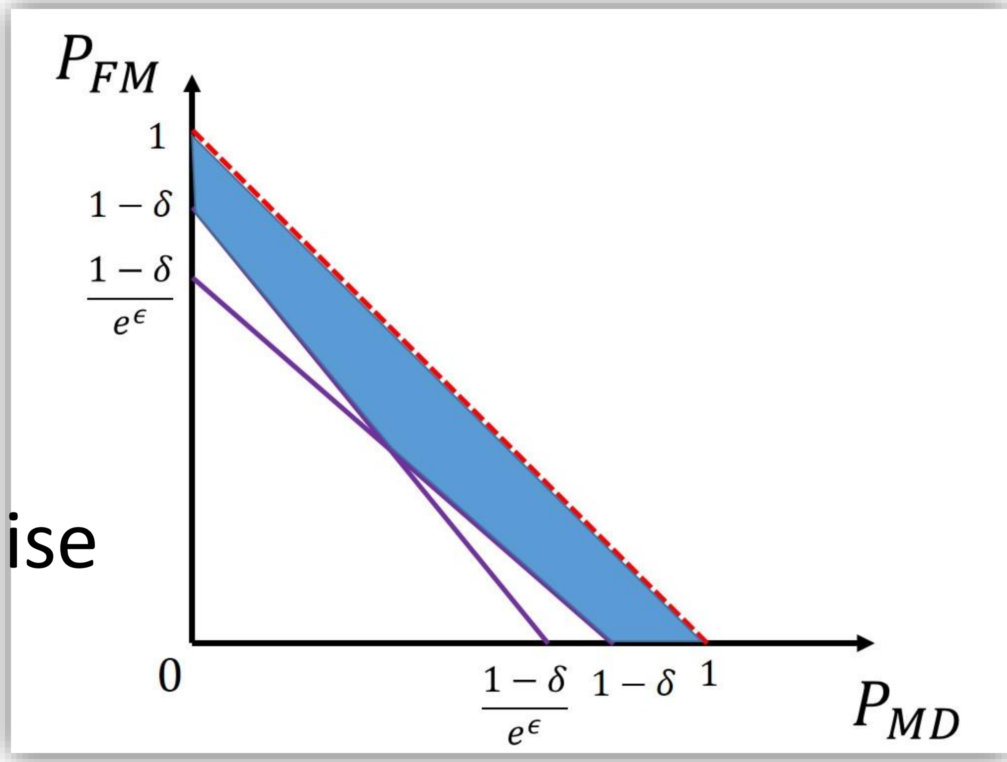
- (ϵ, δ)-differential privacy

$$\mathbf{Pr(K(D_1) \in S) \leq e^{\epsilon} Pr(K(D_2) \in S) + \delta}$$

- Make hypothesis testing hard:

$$e^{\epsilon} P_{FA} + P_{MD} \geq 1 - \delta$$
$$P_{FA} + e^{\epsilon} P_{MD} \geq 1 - \delta$$

- Standard approach: adding Gaussian noise



(ϵ,δ)-Differential Privacy

- (ϵ, δ)-differential privacy

$$\mathbf{Pr}(\mathbf{K}(\mathbf{D_1}) \in \mathbf{S}) \leq \mathbf{e}^{\boldsymbol{\epsilon}}\, \mathbf{Pr}(\mathbf{K}(\mathbf{D_2}) \in \mathbf{S}) + \boldsymbol{\delta}$$

- What is the **optimal** noise probability distribution in this setting?

- **Our Results**:
    - $(0, \delta)$-DP

      Optimality of **Uniform Noise Mechanism**

    - (ϵ, δ)-DP

      Not much more general than $(\epsilon, 0)$- and $(0, \delta)$-DP
      Near-Optimality of **Laplacian** and **Uniform Noise Mechanism** in high privacy regime as $(\epsilon, \delta) \rightarrow (0,0)$

Problem Formulation

$$V^* := \min \sum_{i=-\infty}^{+\infty} L(i)P(i)$$

$$\text{s.t.} \quad P(S) \le e^{\epsilon} P(S + d) + \delta,$$
$$\forall S \subset Z, d \in Z, |d| \le \Delta$$

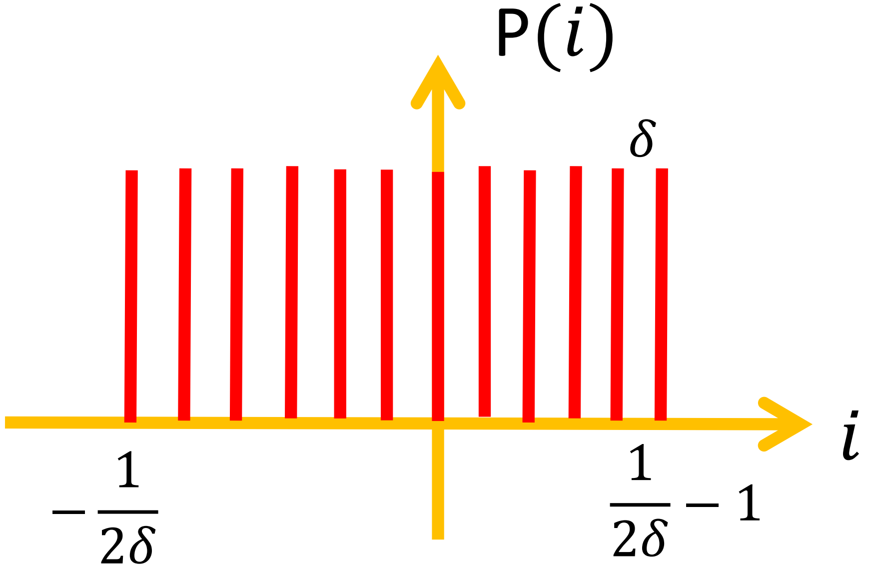
$(0,\delta)$-Differential Privacy

$$V^* := \min \sum_{i=-\infty}^{+\infty} L(i)P(i)$$

$$\text{s.t.} \quad P(S) \leq P(S+d) + \delta, \quad \forall S \subset Z, d \in Z, |d| \leq \Delta$$

- **Theorem:**

If $\Delta = 1$, assuming $\frac{1}{2\delta}$ is an integer, optimal noise P.D. is



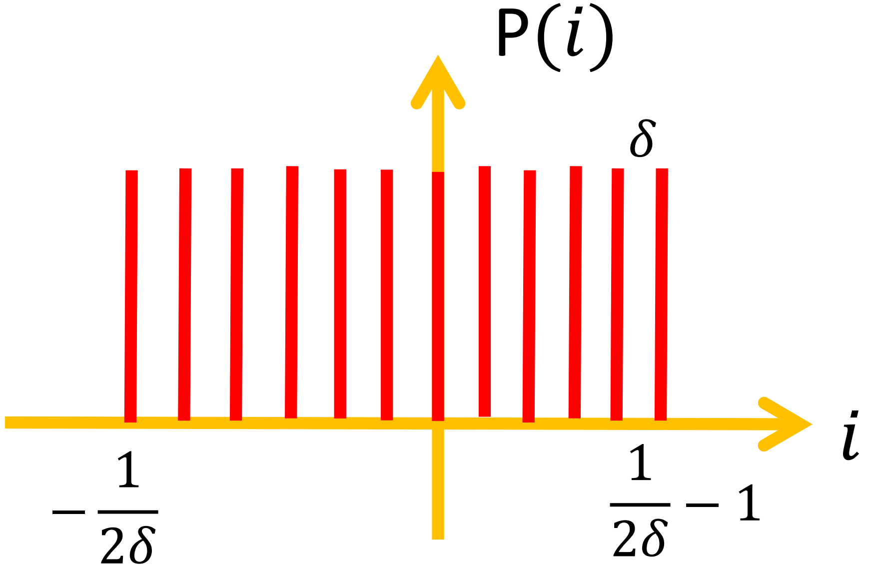
(0,δ)-Differential Privacy

$$V^* := \min \sum_{i=-\infty}^{+\infty} L(i)P(i)$$

$$\text{s.t.} \quad P(S) \le P(S+d) + \delta, \quad \forall S \subset Z, d \in Z, |d| \le \Delta$$

- **Theorem:**

If $\Delta = 1$, assuming $\frac{1}{2\delta}$ is an integer, optimal noise P.D. is



Proof idea:

choose $S = S_k = \{l: l \ge k\}$,
then $P(k) \le \delta, \forall k$

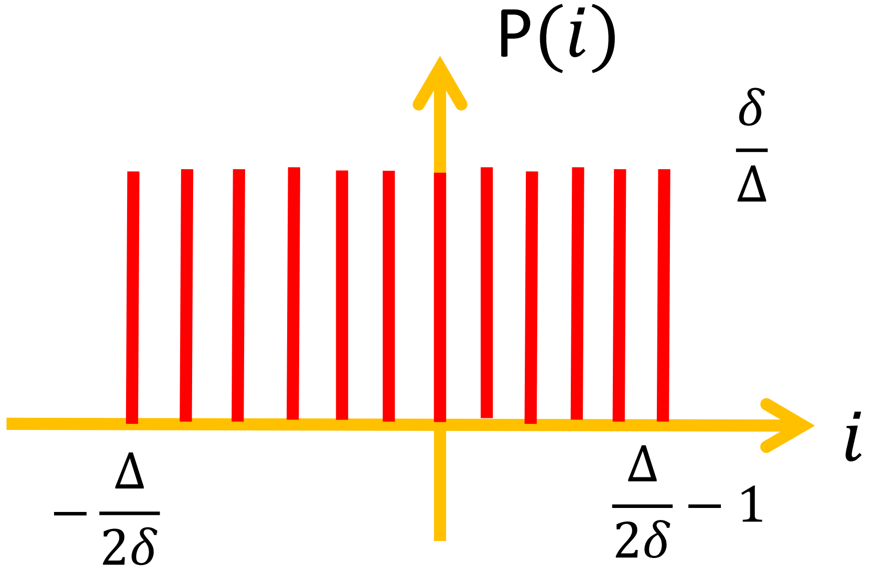
$(0,\delta)$-DP: General Δ

$$V^* := \min \sum_{i=-\infty}^{+\infty} L(i)P(i)$$

s.t. $P(S) \le P(S+d) + \delta,$
$\forall S \subset Z, d \in Z, |d| \le \Delta$

- **Upper Bound:** Uniform Noise Mechanism

$$V^* \le V_{UB} := 2\sum_{i=1}^{\frac{\delta}{2\Delta}-1} \frac{\delta}{\Delta} L(i) + \frac{\delta}{\Delta} L\left(\frac{\Delta}{2\delta}\right)$$



$(0,\delta)$-DP: General Δ

$$V^* := \min \sum_{i=-\infty}^{+\infty} L(i)P(i)$$

$$\text{s.t.} \quad P(S) \leq P(S+d)+\delta, \quad \forall S \subset Z, d \in Z, |d| \leq \Delta$$

- **Lower Bound:** Duality of Linear Programming

choose $S = S_k = \{l: l \geq k\}$, then $\sum_{i=k}^{k+\Delta-1} P(i) \leq \delta, \forall k$

$$V_{LB} := \min \quad 2\sum_{k=1}^{\infty} \mathcal{L}(k)\mathcal{P}_k$$

$$\text{such that} \quad \mathcal{P}_k \geq 0 \quad \forall k \in N$$

$$\frac{\mathcal{P}_0}{2} + \sum_{k=1}^{\infty} \mathcal{P}_k \geq \frac{1}{2}$$

$$-\sum_{\ell=0}^{\Delta-1} \mathcal{P}_{k+\ell} \geq -\delta, \quad \forall k \in \mathbb{N}.$$

$(0,\delta)$-DP: General Δ

$$V^* := \min \sum_{i=-\infty}^{+\infty} L(i)P(i)$$

s.t. $\mathrm{P(S)} \le \mathrm{P}(\mathrm{S}+d)+\delta,$
$\forall \mathrm{S} \subset Z, d \in Z, |d| \le \Delta$

- **Lower Bound:** Duality of Linear Programming

choose $S = S_k = \{l: l \ge k\}$, then $\sum_{i=k}^{k+\Delta-1} P(i) \le \delta, \forall k$

$$V_{LB} := \min \quad 2\sum_{k=1}^{\infty} \mathcal{L}(k)\mathcal{P}_k$$

$$\text{such that} \quad \mathcal{P}_k \ge 0 \quad \forall k \in N$$

$$\frac{\mathcal{P}_0}{2} + \sum_{k=1}^{\infty} \mathcal{P}_k \ge \frac{1}{2}$$

$$-\sum_{\ell=0}^{\Delta-1} \mathcal{P}_{k+\ell} \ge -\delta, \quad \forall k \in \mathbb{N}.$$

$$V_{LB} = 2\delta \sum_{i=0}^{\frac{1}{2\delta}-1} L(1+i\Delta)$$

$(0,\delta)$-DP: Comparison of V_{LB}, V_{UB}

- $L(i) = |i|$

$$V_{LB} = \frac{\Delta}{4\delta} + 1 - \frac{\Delta}{2},$$
$$V_{UB} = \frac{\Delta}{4\delta},$$

Constant additive gap

(0,δ)-DP: Comparison of V_{LB}, V_{UB}

- $L(i) = |i|$

$$V_{LB} = \frac{\Delta}{4\delta} + 1 - \frac{\Delta}{2},$$
$$V_{UB} = \frac{\Delta}{4\delta},$$

Constant additive gap

- $L(i) = |i|^2$

$$V_{LB} = \frac{\Delta^2}{12\delta^2} - \frac{\Delta^2}{4\delta} + \Delta(\frac{1}{2\delta} - 1) + \frac{\Delta^2}{6} + 1,$$
$$V_{UB} = \frac{\Delta^2}{12\delta^2} + \frac{1}{6},$$

Constant multiplicative gap

$(0,\delta)$-DP: Comparison of V_{LB}, V_{UB}

- $L(i) = |i|$

$$V_{LB} = \frac{\Delta}{4\delta} + 1 - \frac{\Delta}{2},$$
$$V_{UB} = \frac{\Delta}{4\delta},$$

Constant additive gap

- $L(i) = |i|^2$

$$V_{LB} = \frac{\Delta^2}{12\delta^2} - \frac{\Delta^2}{4\delta} + \Delta(\frac{1}{2\delta} - 1) + \frac{\Delta^2}{6} + 1,$$
$$V_{UB} = \frac{\Delta^2}{12\delta^2} + \frac{1}{6},$$

Constant multiplicative gap

- $L(i) = |i|^m$

$$\lim_{\delta \to 0} \frac{V_{UB}}{V_{LB}} = 1.$$

(ϵ,δ)-DP: Upper Bound

$$V^* := \min \sum_{i=-\infty}^{+\infty} L(i)P(i)$$

$$\text{s.t.} \quad P(S) \le e^{\epsilon} P(S+d) + \delta, \quad \forall S \subset Z, d \in Z, |d| \le \Delta$$

- Both **(ϵ,0)**-DP and **(0,δ)**-DP imply (ϵ,δ)-DP

$$\boldsymbol{V^* \le \min\left(V_{Lap}, V_{Uniform}\right)}$$

Laplacian Mechanism:

continuous: $f(x) = \frac{1}{2\lambda} e^{-\frac{|x|}{\lambda}}, \ \lambda = \frac{\Delta}{\epsilon}$

discrete: $P(k) = \frac{1-e^{-\frac{\epsilon}{\Delta}}}{1+e^{-\frac{\epsilon}{\Delta}}} e^{-|k|\frac{\epsilon}{\Delta}}$

(ϵ,δ)-DP: Lower Bound

$$V^* := \min \sum_{i=-\infty}^{+\infty} L(i)P(i)$$

s.t. $P(S) \le e^{\epsilon} P(S+d) + \delta,$
$\forall S \subset Z, d \in Z, |d| \le \Delta$

- **Lower Bound:** Duality of Linear Programming

choose $S = S_k = \{l : l \ge k\}$, by symmetry

$$
\begin{aligned}
V_{LB} := \min \quad & 2\sum_{k=1}^{\infty} \mathcal{L}(k)\mathcal{P}_k \\
\text{such that} \quad & \mathcal{P}_k \ge 0 \quad \forall k \in N \\
& \frac{\mathcal{P}_0}{2} + \sum_{k=1}^{\infty} \mathcal{P}_k \ge \frac{1}{2} \\
& \mathcal{P}_0 \frac{1+e^{\epsilon}}{2} + e^{\epsilon} \sum_{k=1}^{\Delta-1} \mathcal{P}_k \le \delta + \frac{e^{\epsilon}-1}{2} \\
& \mathcal{P}_0 \frac{e^{\epsilon}-1}{2} + e^{\epsilon} \sum_{k=1}^{\Delta} \mathcal{P}_k \le \delta + \frac{e^{\epsilon}-1}{2} \\
& \mathcal{P}_0 \frac{e^{\epsilon}-1}{2} + (e^{\epsilon}-1) \sum_{k=1}^{i-1} \mathcal{P}_k + e^{\epsilon} \sum_{k=i}^{i+\Delta-1} \mathcal{P}_k \le \delta + \frac{e^{\epsilon}-1}{2}, \forall i \ge 2.
\end{aligned}
$$

(ϵ,δ)-DP: Comparison of $\boldsymbol{V_{LB}}, \boldsymbol{V_{UB}}$

- $\boldsymbol{L(i) = |i|}$, as $(\epsilon, \delta) \to (0,0)$

$$\frac{min\ (V_{Lap}, V_{Uniform})}{5.29} \leq V_{LB} \leq V^* \leq V_{UB} = min\ (V_{Lap}, V_{Uniform})$$

$$= \Theta(\min\left(\frac{1}{\epsilon}, \frac{1}{\delta}\right))$$

(ϵ,δ)-DP: Comparison of $\boldsymbol{V_{LB}}, \boldsymbol{V_{UB}}$

- $\boldsymbol{L(i)} = |\boldsymbol{i}|$, as $(\epsilon, \delta) \rightarrow (0,0)$

$$\frac{min\,(V_{Lap}, V_{Uniform})}{5.29} \leq V_{LB} \leq V^* \leq V_{UB} = min\,(V_{Lap}, V_{Uniform})$$

$$= \Theta(\min\left(\frac{1}{\epsilon}, \frac{1}{\delta}\right))$$

- $\boldsymbol{L(i)} = |\boldsymbol{i}|^2$, as $(\epsilon, \delta) \rightarrow (0,0)$

$$\frac{min\,(V_{Lap}, V_{Uniform})}{40} \leq V_{LB} \leq V^* \leq V_{UB} = min\,(V_{Lap}, V_{Uniform})$$

$$= \Theta(min\left(\frac{1}{\epsilon^2}, \frac{1}{\delta^2}\right))$$

(ϵ,δ)-DP: Multi-Dimensional Setting

$$\begin{aligned} V^* := \quad & \min \sum_{Z^d} L(i_1, i_2, \dots, i_d) P(i_1, i_2, \dots, i_d) \\ \text{s.t.} \quad & P(S) \le P(S + d) + \delta, \\ & \forall S \subset Z^d, d \in Z^d, \|d\|_1 \le \Delta \end{aligned}$$

- **Upper Bound**

Uniform Noise Mechanism

$$\mathcal{P}(i_1, i_2, \dots, i_d) = \begin{cases} \frac{\delta^d}{\Delta^d} & -\frac{\Delta}{2\delta} \le i_m \le \frac{\Delta}{2\delta} - 1, \forall m \in \{1, 2, \dots, d\} \\ 0 & \text{otherwise} \end{cases}$$

Discrete Laplacian Mechanism

$$P(i_1, \cdots, i_d) = \left(\frac{1-\lambda}{1+\lambda}\right)^d \lambda^{|i_1| + \cdots + |i_d|}, \lambda = e^{-\epsilon/\Delta}$$

(ϵ,δ)-DP: Multi-Dimensional Setting

$$V^* := \min \sum_{Z^d} L(i_1, i_2, \dots, i_d) P(i_1, i_2, \dots, i_d)$$

$$\text{s.t.} \quad P(S) \le P(S + d) + \delta,$$

$$\forall S \subset Z^d, d \in Z^d, \|d\|_1 \le \Delta$$

- **Lower Bound**
  1. Choose certain sets $S = S_k^m := \{(i_1, \cdots, i_d) \in Z^d | i_m \ge k\}$
  2. Write down Linear program, and derive the dual problem
  3. Find proper dual variables to get lower bound

(ϵ,δ)-DP: Multi-Dimensional Setting

Primal Problem

$$V^* \geq V_{LB} := \min \sum_{\mathrm{i}\in\mathbb{Z}^d} \mathcal{P}(\mathrm{i})\mathcal{L}(\mathrm{i}) \qquad (4.30)$$

$$\text{such that} \quad \mathcal{P}(\mathrm{i}) \geq 0 \quad \forall \mathrm{i} \in \mathbb{Z}^d$$

$$\sum_{\mathrm{i}\in\mathbb{Z}^d} \mathcal{P}(\mathrm{i}) \geq 1$$

$$\forall k \in \mathbb{N}, \forall m \in \{1,2,\ldots,d\},$$

$$\sum_{(i_1,i_2,\ldots,i_d)\in\mathbb{Z}^d : k\leq i_m\leq k+\Delta-1} \mathcal{P}(i_1,i_2,\ldots,i_d) - (e^\epsilon - 1) \sum_{(i_1,i_2,\ldots,i_d)\in\mathbb{Z}^d : i_m\geq k+\Delta} \mathcal{P}(i_1,i_2,\ldots,i_d) \leq \delta.$$

Dual Problem

$$V_{LB} := \max \quad \mu - \delta\left(\sum_{i_1\in\mathbb{Z}} y_{i_1}^{(1)} + \sum_{i_2\in\mathbb{Z}} y_{i_2}^{(2)} + \cdots + \sum_{i_d\in\mathbb{Z}} y_{i_d}^{(d)}\right)$$

$$\text{such that} \quad y_{i_1}^{(1)}, y_{i_2}^{(2)}, \ldots, y_{i_d}^{(d)} \geq 0, \forall i_1 \in \mathbb{Z}, i_2 \in \mathbb{Z}, \ldots, i_d \in \mathbb{Z}$$

$$\mu - \sum_{i_1\in[k_1-\Delta+1,k_1]} y_{i_1}^{(1)} + (e^\epsilon - 1) \sum_{i_1\leq k_1-\Delta} y_{i_1}^{(1)}$$

$$- \cdots - \sum_{i_d\in[k_d-\Delta+1,k_d]} y_{i_d}^{(d)} + (e^\epsilon - 1) \sum_{i_d\leq k_d-\Delta} y_{i_d}^{(d)}$$

$$\leq |k_1| + |k_2| + \cdots + |k_d|, \forall (k_1,\ldots,k_d) \in \mathbb{Z}^d.$$

(ϵ,δ)-DP: Multi-Dimensional Setting

- **(0,δ)-DP:**

ℓ^1

$$V_{LB} \geq \frac{d\Delta}{4\delta} - \frac{\Delta - 1}{2}d,$$
$$V_{UB} = \frac{d\Delta}{4\delta},$$

ℓ^2

$$V_{LB} \geq \frac{d\Delta^2}{12\delta^2} + (\frac{1}{\Delta} - 1)\frac{d\Delta^2}{4\delta} + \frac{1-\Delta}{2}d + \frac{d\Delta^2}{6},$$
$$V_{UB} = \frac{d\Delta^2}{12\delta^2} + \frac{d}{6},$$

Uniform Noise Mechanism is Optimal when $\Delta = 1$

- **(ϵ,δ)-DP:**

$$\frac{min\ (V_{Lap}, V_{Uniform})}{C} \leq V_{LB} \leq V^* \leq V_{UB} = min\ (V_{Lap}, V_{Uniform})$$

$C \approx 9.49, for\ \ell^1$
$C \approx 113, for\ \ell^2$

Conclusion of Part Three

- Near-Optimality of Uniform Noise Mechanism in $(0, \delta)$-DP

- (ϵ, δ)-DP is not much more general than $(0, \delta)$-DP and $(\epsilon, 0)$-DP

$$\frac{min\ (V_{Lap}, V_{Uniform})}{C} \leq V_{LB} \leq V^* \leq V_{UB} = min\ (V_{Lap}, V_{Uniform})$$

Conclusion

- Fundamental tradeoff between privacy and utility in Differential Privacy
- Staircase Mechanism, optimal mechanism in ϵ-DP
  - Huge improvement in low privacy regime
  - Extension to Multi-Dimensional Setting
- Uniform Noise Mechanism, near-optimal mechanism in $(0, \delta)$-DP
- (ϵ, δ)-DP is not much more general than $(0, \delta)$-DP and $(\epsilon$,0)-DP

Acknowledgment

Thanks to my advisor, Prof. **Pramod Viswanath**

Thank you!