Lecture 6 Advanced Composition (Part II), Gaussian mechanism

Yu-Xiang Wang

COMPUTER SCIENCE
UC SANTA BARBARA
Computing. ReInvented.

Recap: last lecture

- Private selection
  - Exponential mechanism
  - Report-Noisy-Max
- Application of Exponential mechanism
  - SmallDB algorithm
- Advanced Composition
  - Apply to linear query release
  - Privacy loss random variable

Recap: Utility of Exponential Mechanism and small DB

- Utility of Exp-Mech
- Approximation error of a SmallDB
- Guarantee of SmallDB

Recap: Advanced Composition

Theorem: The adaptive composition of k (ε,δ)-DP mechanisms satisfies $(\tilde{\varepsilon}, \tilde{\delta}) - DP$ where

$$\tilde{\varepsilon} = \varepsilon\sqrt{2k\log(1/\delta')} + \boxed{2k\varepsilon^2}, \qquad \tilde{\delta} = k\delta + \delta'$$

Slightly different, also not the tightest; but among the cleanest with a simple proof.

for any $\varepsilon, \delta \geq 0, \delta' > 0$

TLDR: For reasonable ranges of privacy parameter, total privacy loss scales as sqrt(k).

We will prove the version of this theorem for δ = 0 today.

Recap: Application of Advanced composition to linear query release.

- Advanced composition of Laplace mechanisms
  - For k times.

- Advanced composition of AboveThresh and Laplace Mechanism
  - For N times where N is the number of times to update the synthetic data

Summary of the problem of private query release

	Laplace (release query)	**Laplace (release data)**	**Private Multiplicative Weights** ***(Adaptive queries)***	**SmallDB** ***(Fixed queries)***
Error under Pure-DP	$\frac{k\log(k/\beta)}{n\epsilon}$	$\frac{\sqrt{\lvert\mathcal{X}\rvert}\log\frac{k}{\beta})}{n\epsilon}$	$\left(\frac{\log\lvert\mathcal{X}\rvert\log(k/\beta)}{n\epsilon}\right)^{1/3}$	$\left(\frac{\log\lvert\mathcal{X}\rvert\log\lvert\mathcal{Q}\rvert}{n\epsilon}\right)^{1/3} + \frac{\log\frac{1}{\beta}}{n\epsilon}$
Error under approx-DP	$\frac{\sqrt{k\log\frac{1}{\delta}\log\frac{k}{\beta}}}{n\epsilon}$	Same as above	$\left(\frac{\log\frac{k}{\beta}\sqrt{\log\lvert\mathcal{X}\rvert\log\frac{1}{\delta}}}{n\epsilon}\right)^{1/2}$	Same as above
Computational complexity (per query)	$O(n)$	$O(\lvert\mathcal{X}\rvert)$	$O(\max\{\lvert\mathcal{X}\rvert, n\})$	$O(\lvert\mathcal{X}\rvert)$

Recap: Privacy loss random variable

- PLRV is the log probability ratio as a random variable

$$\boldsymbol{\varepsilon}_{\mathcal{M}}^{x,x'} = \log\left(\frac{p(\mathbf{y})}{p'(\mathbf{y})}\right) \text{ where random variable } \mathbf{y} \sim \mathcal{M}(x).$$

- Tail bound of privacy loss r.v. implies DP

Lemma 1 (Tail bound to (ϵ, δ)-DP conversion). *Let* $\boldsymbol{\varepsilon}_{\mathcal{M}}^{x,x'}$ *be the privacy loss RV defined above. If*

$$\mathbb{P}(\boldsymbol{\varepsilon}_{\mathcal{M}}^{x,x'} > \epsilon) \leq \delta$$

for all *pair of neighboring* x, x' *then* $\mathcal{M}$ *satisfies* (ϵ, δ)*-DP.*

(You are to prove this in HW1.)

This lecture

- Advanced composition (Part II)
    - Proof of advanced composition for pure DP mechanisms
- Gaussian mechanism
    - PLRV of the Gaussian mechanism
    - Composition of Gaussian mechanisms via Adv. composition
- Renyi Differential Privacy
    - Deriving RDP from PLRVs
    - Improved composition of Gaussian mechanism

Readings:

- Advanced Composition for pure-DP
  - Lecture notes
- Gaussian mechanism
  - Balle and W., 2018
- Probability inequalities and subgaussian tail bounds
  - Larry's notes: https://www.stat.cmu.edu/~larry/=stat705/Lecture2.pdf
- Renyi Differential Privacy
  - Bun and Steinke, 2017
  - Mironov, 2017

Adaptive Composition = Sum of privacy loss random variables

- Fix two neighboring datasets, consider a sequence of adaptively chosen pure-DP mechanisms

Proof Idea of Advanced Composition

- Observation 1: sometimes PLRV is positive, other times negative. They cancel with each other.

- Observation 2: as k gets larger, the sum of PLRV concentrates around its mean.
  - Calculate their mean
  - Bound the deviation from the mean

- Observation 3: the adaptivity means that the PLRV will depend on the past

Martingale

- We say that a sequence of r.v. $X_1,\ldots,X_n,\ldots$ is a Martingale if for any n

$$\mathbf{E}(|X_n|) < \infty$$

$$\mathbf{E}(X_{n+1} \mid X_1, \ldots, X_n) = X_n.$$

- Example:
  - Random-walk: Total number of heads minus tails in n coin tosses

Azuma-Hoeffding's inequality

- **Azuma-Hoeffding's inequality**: Assume $X_1, \ldots, X_n$ are Martingale differences

$$S_n = X_1 + \ldots + X_n$$

$$\mathbb{P}\left[S_n \geq \epsilon\right] \leq e^{-\frac{2\epsilon^2}{\sum_{i=1}^{n}(b_i - a_i)^2}}$$

- Apply Azuma-Hoeffding's inequality to our problem
  - What are these martingale differences?
  - What are these bounds?

Proof for the advanced composition for pure DP mechanisms

- Fix x, x', apply Azuma-Hoeffding's inequality

Bounding the KL-divergence

- **Lemma** (Pinsker's inequality)

$$\|P-Q\|_1 \leq \sqrt{2D_{KL}(P\|Q)}$$

- **Corollary**: KL-divergence is nonnegative.
- Now's let's bound the expected value of PLRV:

Improved bounds of the KL-divergence and tighter version of Advanced composition.

Condition: P,Q satisfies that the log-odds ratio ≤ ε.

1. Bound from Dwork and Roth book

$$D_{KL}(P\|Q) \leq \epsilon(e^\epsilon - 1)$$

2. Bound from Bun and Steinke

$$D_{KL}(P\|Q) \leq \frac{\epsilon^2}{2}$$

Use this for theory / get asymptotic rate

3. Tight bound from Adam Smith (also in the proof of Bun and Steinke):

$$D_{KL}(P\|Q) \leq \epsilon \cdot \frac{e^\epsilon - 1}{e^\epsilon + 1} = \epsilon \cdot \tanh(\epsilon/2)$$

Implement this one (already in autodp)

Proofs are left as an exercise (maybe one question in HW2).

Checkpoint

- A simple proof of advanced composition for pure-DP mechanisms via PLRV

Theorem: The adaptive composition of k (ε,δ)-DP mechanisms satisfies $(\tilde{\varepsilon}, \tilde{\delta}) - DP$ where

$$\tilde{\varepsilon} = \varepsilon\sqrt{2k\log(1/\delta')} + \boxed{2}k\varepsilon^2, \qquad \tilde{\delta} = k\delta + \delta'$$

Can be improved to 0.5.

for any $\varepsilon, \delta \geq 0, \delta' > 0$

- Proof of the approx DP mechanism is similar but requires providing a PLRV that works for all approx. DP mechanisms.
- We will (hopefully) cover that in the next lecture as a natural by product.

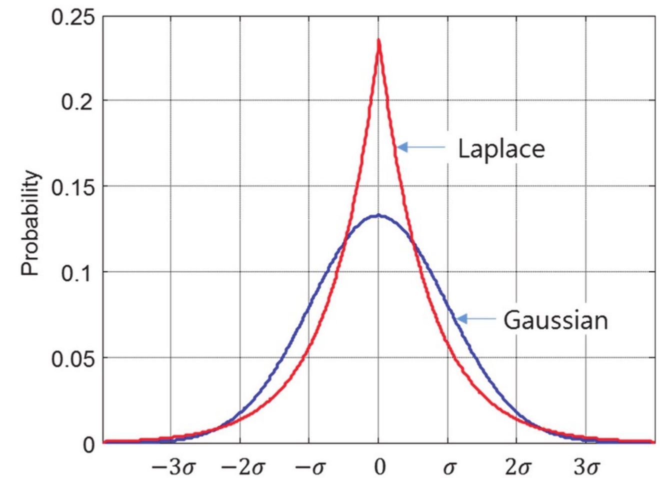
Gaussian mechanism

- Releasing low-sensitivity query f

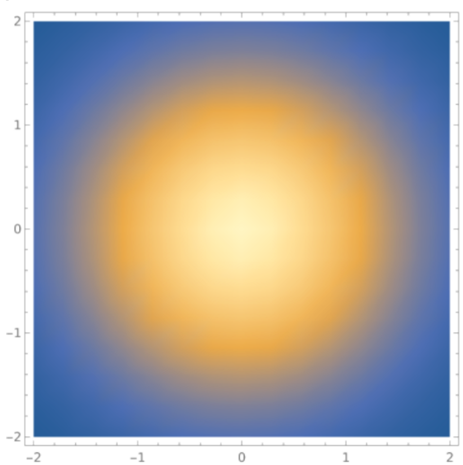
- L2-Sensitivity of f

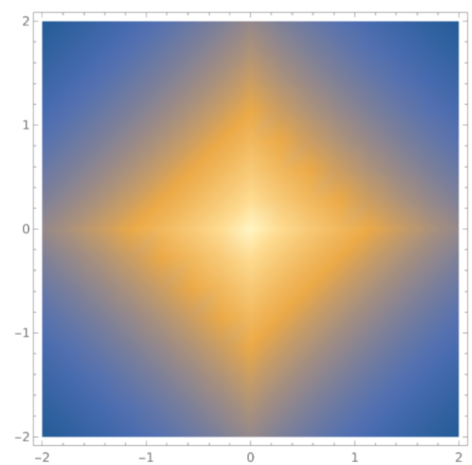
Gaussian noise is more concentrated than Laplace noise

- N(0, σ) vs Lap(0,b)



- In multiple dimensions





Credit: Damien Desfontaines https://desfontain.es/privacy/gaussian-noise.html

Examples of queries and their sensitivities

1. Histogram under “add/remove”

2. Histogram under “replace”

3. Voting when each individual has k-ballots

4. Uncentered sample covariance (“Gram”) matrix

Privacy Loss Random Variable of the Gaussian mechanism is Gaussian

- Recall: $\varepsilon_{\mathcal{M}}^{x,x'} = \log(\frac{p(\mathbf{y})}{p'(\mathbf{y})})$ where random variable $\mathbf{y} \sim \mathcal{M}(x)$.
- The privacy loss RV of a Gaussian mechanism is $\mathcal{N}(\eta, 2\eta) \ with \ \eta = D^2/2\sigma^2, \ where \ D = \|f(x) - f(x')\|.$

The privacy analysis of Gaussian mechanism

Lemma 1 (Tail bound to (ϵ, δ)-DP conversion)**.** *Let $\boldsymbol{\varepsilon}_{\mathcal{M}}^{x,x'}$ be the privacy loss RV defined above. If*

$$\mathbb{P}(\boldsymbol{\varepsilon}_{\mathcal{M}}^{x,x'} > \epsilon) \leq \delta$$

for all *pair of neighboring x, x' then $\mathcal{M}$ satisfies (ϵ, δ)-DP.*

- Useful lemma: Gaussian tail bound

$$\text{Let } X \sim \mathcal{N}(\mu, \sigma^2), \text{ we have:}$$

$$\mathbb{P}(X - \mu \geq u) \leq \exp(-u^2/(2\sigma^2)).$$

$$\mathcal{N}(\eta, 2\eta) \ \textit{with}\ \eta = D^2/2\sigma^2,\ \textit{where}\ D = \|f(x) - f(x')\|.$$

The privacy analysis of Gaussian mechanism

- The Gaussian mechanism with variance σ^2 for a query with L2-sensitivity Δ satisfies (ε,δ)-DP with
  - ε =
  - δ=

Classical Gaussian mechanism: For all $0 < \varepsilon, \delta \leq 1$,
The mechanism obeys (ε,δ)-DP if we choose

$$\sigma = \frac{\Delta}{\epsilon}\sqrt{2\log(1.25/\delta)}$$

Remainder of the lecture

- Detour on centration inequalities
- Concentrated Differential Privacy
- The composition of Gaussian Mechanism

Detour: Concentration inequality

- Markov's inequality
  - For any non-negative r.v. X: $\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}[X]}{t}$

- Chebychev's inequality
  - For any r.v. X with variance σ^2 : $\mathbb{P}\Big(|X - \mathbb{E}[X]| \geq t\sigma\Big) \leq \frac{1}{t^2}$

- Generalizing Chebychev:

Detour: Chernoff's method

Define, $\mu = \mathbb{E}[X]$. For any $t > 0$, we have that,

$$\mathbb{P}((X-\mu) \geq u) = \mathbb{P}(\exp(t(X-\mu)) \geq \exp(tu)) \leq \frac{\mathbb{E}[\exp(t(X-\mu))]}{\exp(tu)}$$

- Chernoff's bound:

$$\mathbb{P}((X-\mu) \geq u) \leq \inf_{0 \leq t \leq b} \exp(-t(u+\mu))\mathbb{E}[\exp(tX)].$$

Subgaussian random variables

- We say a r.v. X with mean μ is σ-subgaussian if

$$\mathbb{E}[\exp(t(X-\mu))] \leq \exp(\sigma^2 t^2/2), \text{ for all } t \in \mathbb{R}.$$

  - We say that X is subgaussian if there exists constant σ.

- Example 1: N(μ, σ^2) is subgaussian.

- Example 2: Bounded random variables are subgaussian.
  - Exercise: what is σ parameter if the range is [a,b]?

Tail bound of subgaussian random variables

- By the Chernoff's bound we get that

$$\mathbb{P}(X - \mu \geq u) \leq \exp(-u^2/(2\sigma^2)).$$

- Proof: By the Chernoff's method

Average of n independent $\boldsymbol{\sigma}$-subgaussian RVs is $\frac{\boldsymbol{\sigma}}{\sqrt{\boldsymbol{n}}}$-subgaussian.

- Why?

$$\begin{aligned}
\mathbb{E}[\exp(t(\widehat{\mu} - \mu))] &= \mathbb{E}[\exp(t/n \sum_{i=1}^{n} (X_i - \mu)] \\
&= \prod_{i=1}^{n} \mathbb{E}[\exp(t(X_i - \mu)/n)] \\
&\leq \exp(t^2 \sigma^2 / (2n)).
\end{aligned}$$

- which implies

$$\mathbb{P}(|\widehat{\mu} - \mu| \geq k\sigma/\sqrt{n}) \leq 2\exp(-k^2/2).$$

Idea: let’s handle mechanisms that are Gaussian-mechanism-like, in a sense that

1. For any neighboring datasets, the PLRV is σ-subgaussian

2. Then composition is straightforward
   - The sum of k PLRVs is $\sigma\sqrt{k}$ - subgaussian

From Moment Generating Functions to Renyi divergence

- Moment Generating function of PLRV

- Renyi Divergence $\alpha \in (0,1) \cup (1,\infty)$

$$D_\alpha(P\|Q) = \frac{1}{\alpha-1} \ln \int p^\alpha q^{1-\alpha} \, \mathrm{d}\mu.$$

$$D_1(P\|Q) = D(P\|Q) = \text{ Kullback-Leibler divergence}$$

$$D_\infty(P\|Q) = \ln \left(\operatorname*{ess\,sup}_P \frac{p}{q} \right)$$

Van Erven, T., & Harremos, P. (2014). Rényi divergence and Kullback-Leibler divergence. *IEEE Transactions on Information Theory*, *60*(7), 3797-3820.

Example of Renyi Divergence

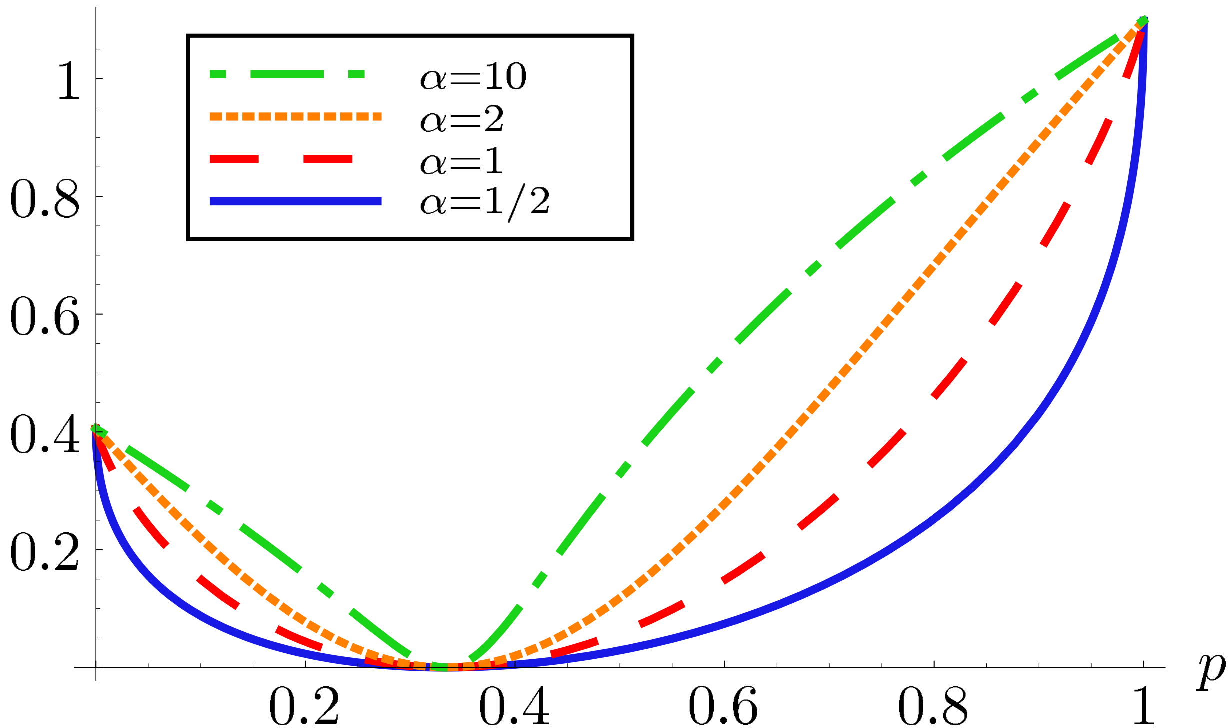


Fig. 2. Rényi divergence as a function of $P = (p, 1 - p)$ for $Q = (1/3, 2/3)$

Renyi Differential Privacy and Concentrated Differential Privacy

- (Mironov, 2017) We say that a mechanism satisfies (α, ϵ)-Renyi DP, if

$$D_\alpha(\mathcal{M}(x) \| \mathcal{M}(x')) \leq \epsilon$$

- (Dwork and Rothblum / Bun and Steinke, 2016) We say a mechanism satisfies ρ-zCDP, if

$$D_\alpha(\mathcal{M}(x) \| \mathcal{M}(x')) \leq \rho\alpha, \forall \alpha > 1$$

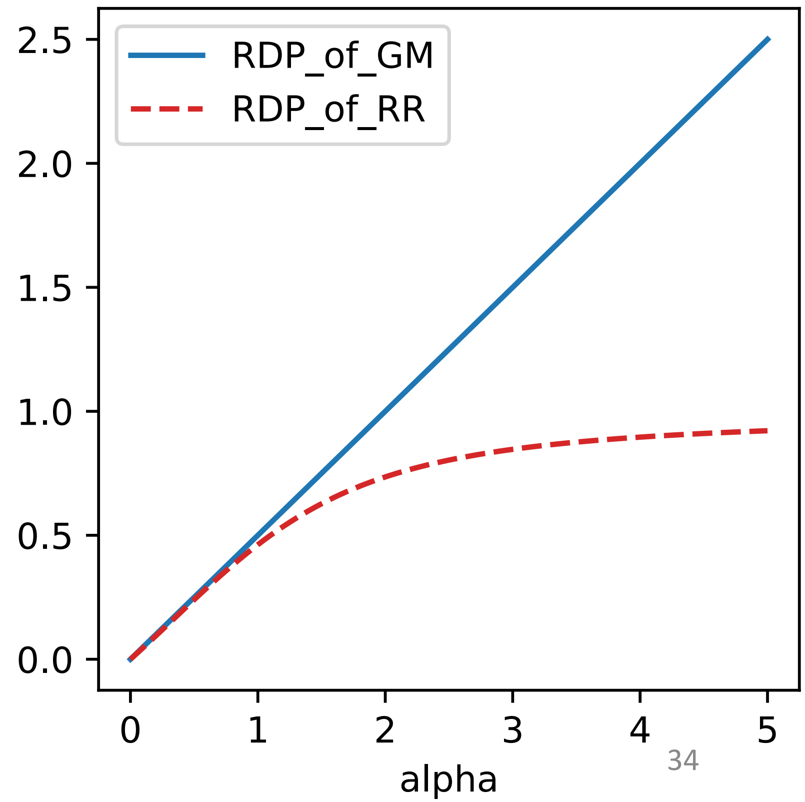
- Connection to PLRV:

Examples of Renyi-DP/ CDP mechanisms

Mechanism	Differential Privacy	Rényi Differential Privacy for α
Randomized Response	$\left\lvert \log \frac{p}{1-p} \right\rvert$	$\alpha > 1$: $\frac{1}{\alpha-1} \log \left(p^\alpha (1-p)^{1-\alpha} + (1-p)^\alpha p^{1-\alpha} \right)$ $\alpha = 1$: $(2p-1) \log \frac{p}{1-p}$
Laplace Mechanism	$1/\lambda$	$\alpha > 1$: $\frac{1}{\alpha-1} \log \left\{ \frac{\alpha}{2\alpha-1} \exp(\frac{\alpha-1}{\lambda}) + \frac{\alpha-1}{2\alpha-1} \exp(-\frac{\alpha}{\lambda}) \right\}$ $\alpha = 1$: $1/\lambda + \exp(-1/\lambda) - 1 = .5/\lambda^2 + O(1/\lambda^3)$
Gaussian Mechanism	∞	$\alpha/(2\sigma^2)$

- Gaussian mechanism

- Pure-DP mechanism



Properties of Renyi DP / CDP

- Adaptive Composition

- Conversion to approximate DP

$$(\epsilon, \alpha)\text{-RDP implies } (\epsilon(\alpha) + \tfrac{\log(1/\delta)}{\alpha-1}, \delta)\text{-DP}$$

If M provides ρ-zCDP, then M is $(\rho + 2\sqrt{\rho \log(1/\delta)}, \delta)$-DP.

- Other properties: Postprocessing, risk multiplier, group privacy (see Mironov, 2017)

Composition, revisited

- Composition of pure-DP mechanism via zCDP
- Composition of Gaussian mechanism via zCDP
- Baseline: Composition of Gaussian mechanism via Advanced Composition

Next Lecture

- More on Renyi Differential Privacy

- Alternative characterization of DP
  - Privacy-profiles
  - Tradeoff functions

- Modern tool: autodp