CS2080: Applied Privacy for Data Science
Composition of Differential Privacy

School of Engineering & Applied Sciences
Harvard University

February 12, 2025

Discussion

Imagine a study asks a random sample of voters in a town about their race and whether they are registered as Republican, Democrat, or Independent. The investigators *claim* to use differentially private mechanisms to compute proportions of party registration by race, along with confidence intervals around these proportions.

Their results indicate that Asian people in the town are almost always registered as Democrats, however people of other racial backgrounds in the town are almost always registered as Republicans.

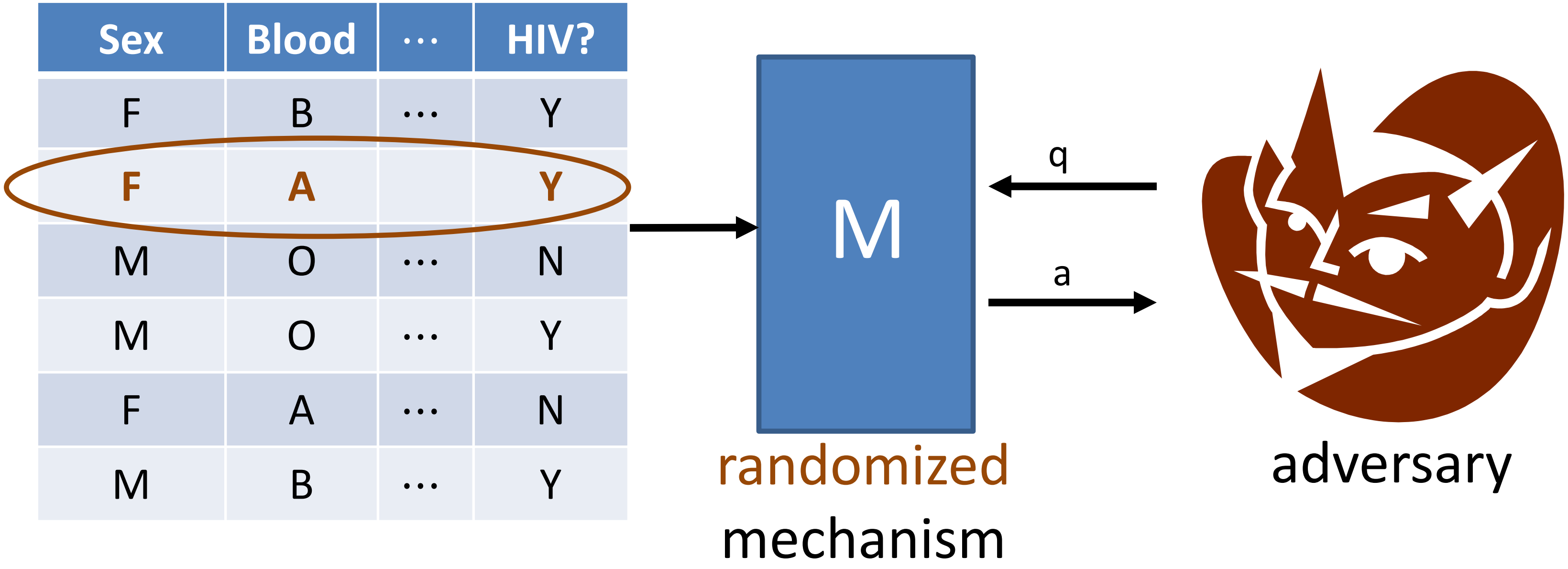
Mr. E is an Asian person living in the town. His boss, a registered Republican, knows Mr. E's race and uses the results of the study to conclude that Mr. E is likely a Democrat.

Mr. E finds out what his boss has learned and is upset, feeling that his privacy was violated. He contacts the study investigators letting them know what happened and questioning whether their analysis actually satisfied differential privacy.

Based on Mr. E's experience, is there reason to doubt the investigators' claim of using differentially private mechanisms in their analysis? Why or why not?

DP for one query/release

[Dwork-McSherry-Nissim-Smith ’06]



Sex	Blood	⋯	HIV?
F	B	⋯	Y
F	A		Y
M	O	⋯	N
M	O	⋯	Y
F	A	⋯	N
M	B	⋯	Y

randomized mechanism

adversary

Def: M is ε-DP if for all x, x' differing on one row, and all q

$\forall$ sets T, $$\Pr[M(x,q) \in T] \leq e^{\varepsilon} \cdot \Pr[M(x',q) \in T]$$

(Probabilities are (only) over the randomness of M.)

Properties of the Definition

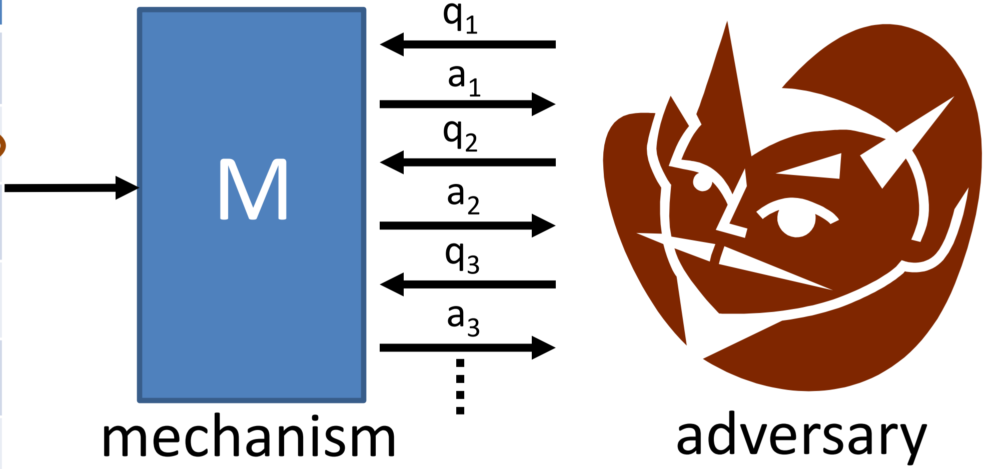
- Suffices to check pointwise: M is ε-DP if and only if $\forall x \sim x' \; \forall q \; \forall y \; \Pr[M(x,q)=y] \le e^{\varepsilon} \cdot \Pr[M(x',q)=y]$.

- Preserved under post-processing: If M is ε-DP and f is any function, then $M'(x,q)=f(M(x,q))$ is ε-DP.

- (Basic) composition: If M_i is ε_i-DP for $i=1,\ldots,k$, then
$$M'\big(x,(q_1,\ldots,q_k)\big)=(M_1(x,q_1),\ldots,M_k(x,q_k))$$
is $(\varepsilon_1+\cdots+\varepsilon_k)$-DP
  - Use independent randomness for the k queries
  - Holds even if q_i's are chosen adaptively

DP for Interactive Mechanisms

Sex	Blood	⋯	HIV?
F	B	⋯	Y
F	A		Y
M	O	⋯	N
M	O	⋯	Y
F	A	⋯	N
M	B	⋯	Y

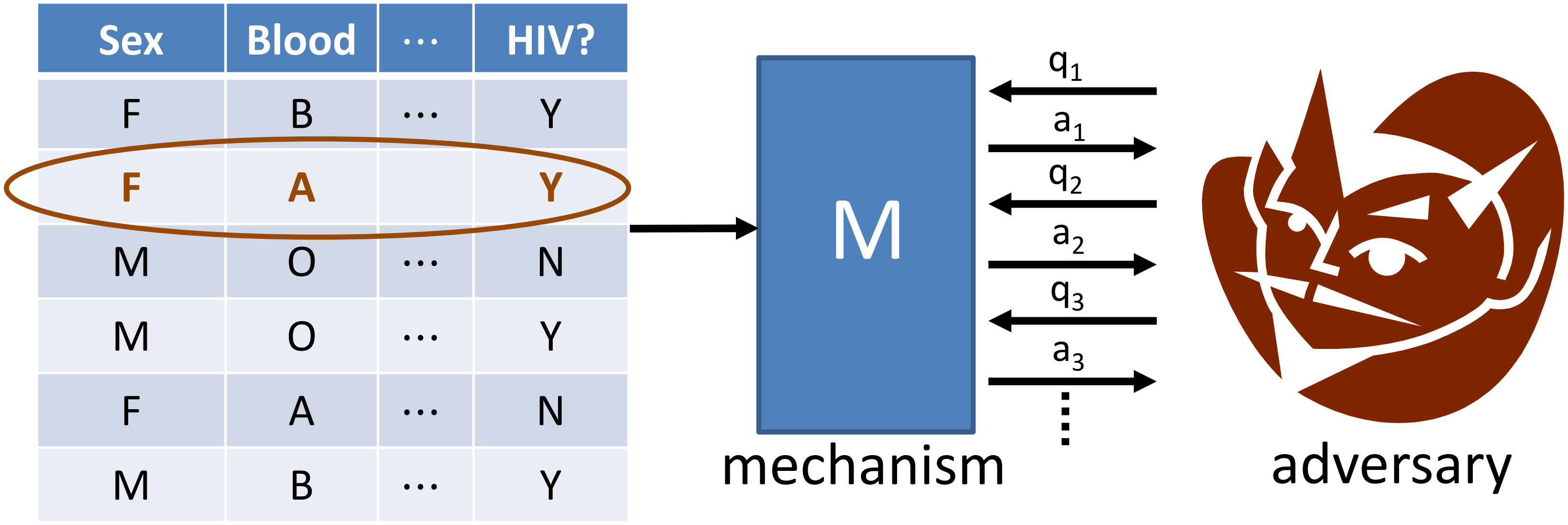


1st Attempt: for all $x \sim x'$, all $q_1, \dots, q_t$, all T

$$\Pr[\underbrace{M(x, q_1, \dots, q_t)} \in T] \leq e^{\varepsilon} \cdot \Pr[\underbrace{M(x', q_1, \dots, q_t)} \in T]$$

vectors of answers $a_1, \dots, a_t$

DP for Interactive Mechanisms



Sex	Blood	⋯	HIV?
F	B	⋯	Y
F	A		Y
M	O	⋯	N
M	O	⋯	Y
F	A	⋯	N
M	B	⋯	Y

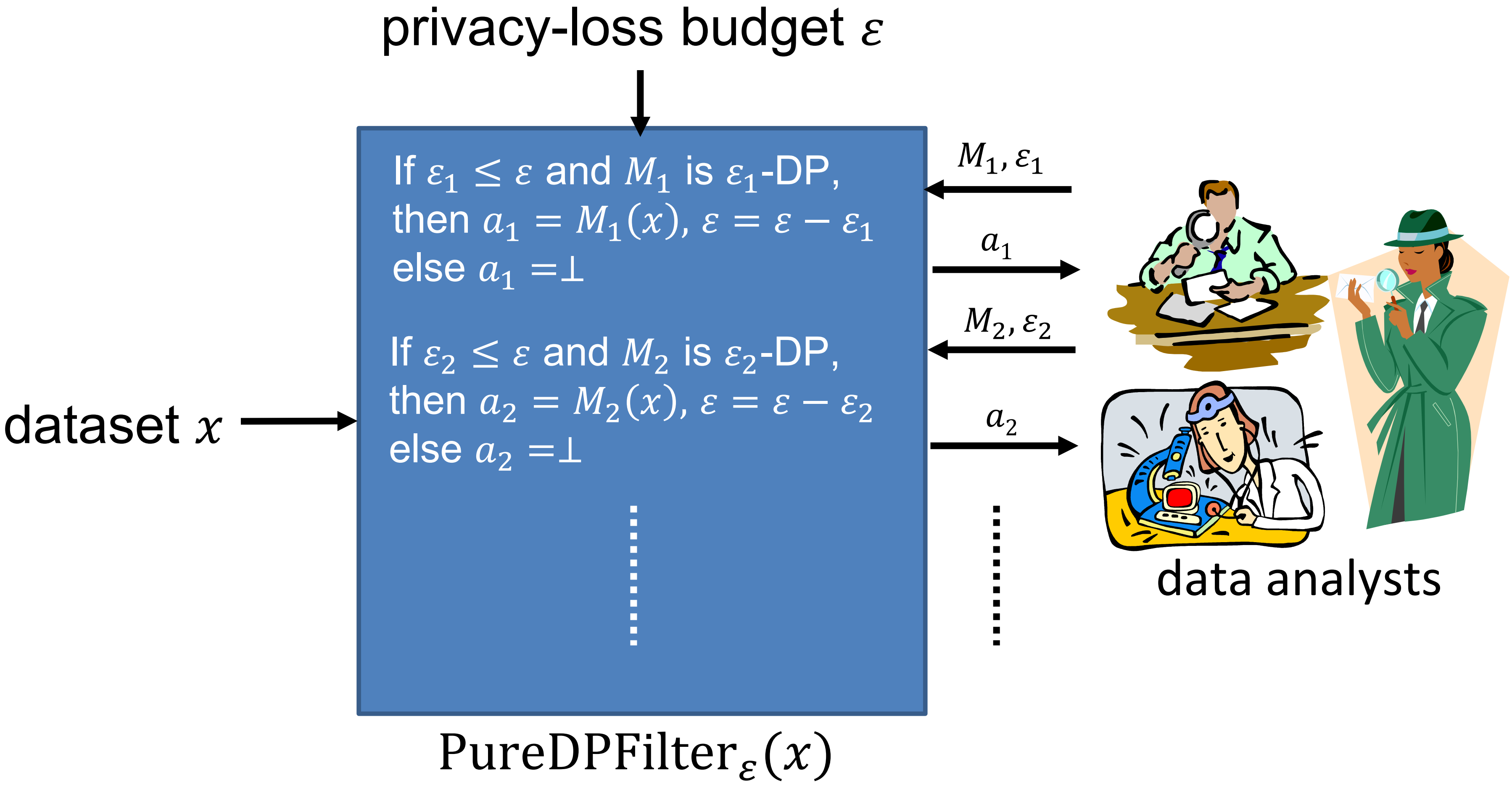
Better: for all $x \sim x'$, all adversarial strategies A

$$\underbrace{\mathrm{View}_A\big(A \leftrightarrow M(x)\big)} \approx_\varepsilon \underbrace{\mathrm{View}_A(A \leftrightarrow M(x'))}$$

Everything A sees (its internal randomness & query answers)

Equivalently: $\forall\ A\ \ \Pr[A \text{ outputs "In" after interacting w}/M(x)]$
$\leq e^{\varepsilon} \cdot \Pr[A \text{ outputs "In" after interacting w}/M(x')]$

Composition as an Interactive Mechanism



privacy-loss budget ε

If $\varepsilon_1 \leq \varepsilon$ and M_1 is ε_1-DP,
then $a_1 = M_1(x)$, $\varepsilon = \varepsilon - \varepsilon_1$
else $a_1 = \perp$

If $\varepsilon_2 \leq \varepsilon$ and M_2 is ε_2-DP,
then $a_2 = M_2(x)$, $\varepsilon = \varepsilon - \varepsilon_2$
else $a_2 = \perp$

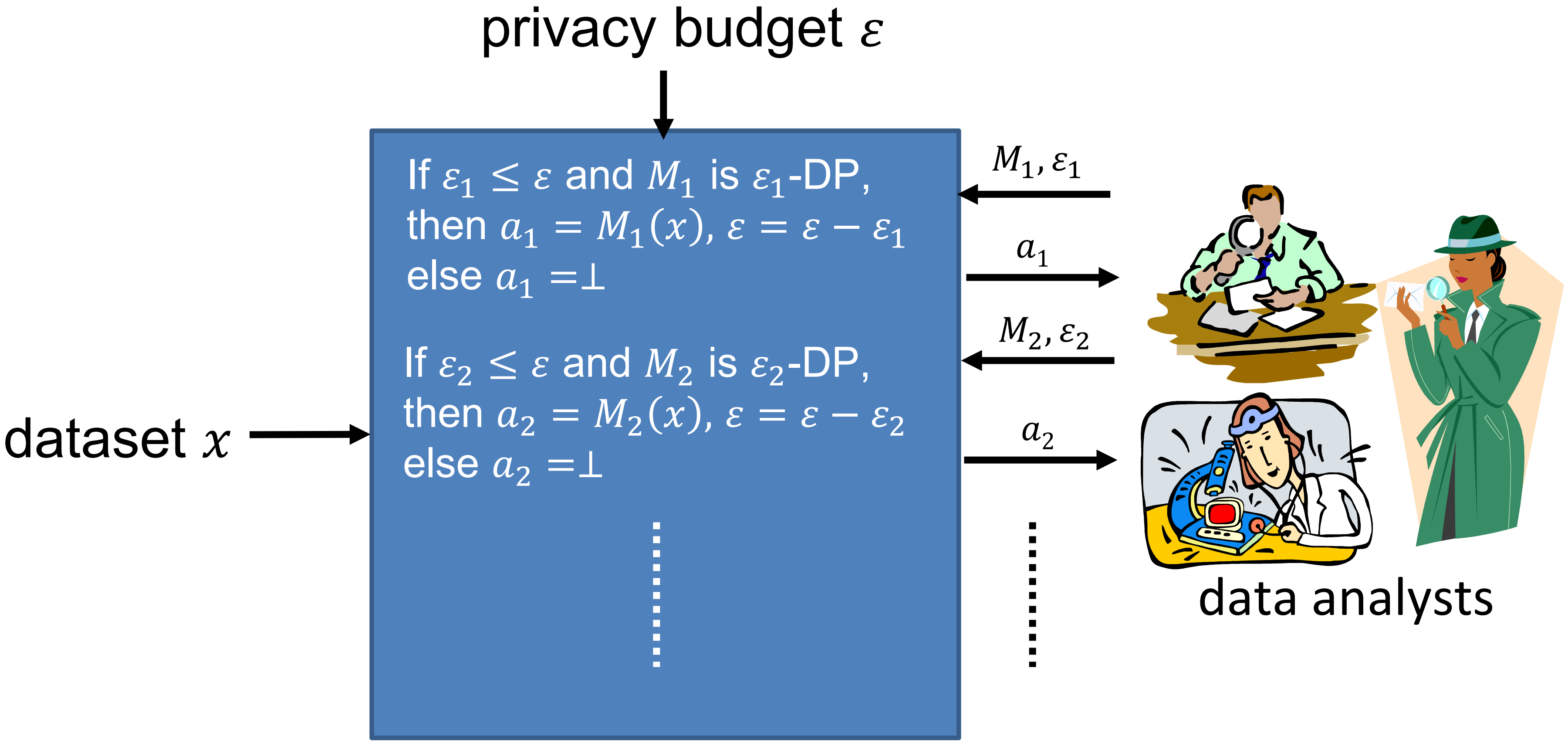
dataset x

data analysts

$\mathrm{PureDPFilter}_{\varepsilon}(x)$

Theorem: $\mathrm{PureDPFilter}_{\varepsilon}$ is an ε-DP interactive mechanism.

Privacy Budgeting



If $\varepsilon_1 \le \varepsilon$ and M_1 is ε_1-DP,
then $a_1 = M_1(x)$, $\varepsilon = \varepsilon - \varepsilon_1$
else $a_1 = \perp$

If $\varepsilon_2 \le \varepsilon$ and M_2 is ε_2-DP,
then $a_2 = M_2(x)$, $\varepsilon = \varepsilon - \varepsilon_2$
else $a_2 = \perp$

- To answer k queries, can set each $\varepsilon_i = \varepsilon/k$.
- More queries ⇒ smaller ε_i ⇒ less accuracy per query.
- Some tradeoff #queries vs. accuracy necessary. (Q: why?)

Composition for Algorithm Design

Composition and post-processing allow designing more complex differentially private algorithms from simpler ones.

Example: The “Statistical Query Model” for ML

- Many ML algorithms (e.g. stochastic gradient descent) can be described as sequence of low-sensitivity queries (e.g. averages) over the dataset, and can tolerate noisy answers to the queries.
- Can answer each query by adding Laplace noise.
- By composition and post-processing, trained model is DP and safe to output.

Group Privacy & Setting ε

- Proposition: If M is ε-DP for individuals, then it is $k\varepsilon$-DP for groups of k individuals. That is, if x and x' differ on at most k individuals, then
$$\forall T\ \ \Pr[M(x) \in T] \leq e^{k\varepsilon} \cdot \Pr[M(x') \in T]$$

- Q: what are examples of “groups” for which this is useful?

- Consequence: need $n \gg 1/\varepsilon$ for any reasonable utility.

- Typical recommendation for a “good” privacy guarantee: $.01 \leq \varepsilon \leq 1$.

DP Histograms

Not all sequences of k queries require noise growing with k.

Thm: Let $B_1, \ldots, B_k$ be disjoint subsets of the row-space $\mathcal{X}$.

- Define $q_j(x) = \left|\{i : x_i \in B_j\right|$.
- Then $M(x) = (q_1(x) + Z_1, \ldots, q_k(x) + Z_k)$ is ε-DP for
- $Z_k \sim \mathrm{Lap}(1/\varepsilon)$ if dataset adjacency is wrt d_{Sym} (add/remove a row)
- $Z_k \sim \mathrm{Lap}(2/\varepsilon)$ if dataset adjacency is wrt d_{Ham} (change a row)

DP Histograms

Thm: Let $B_1, \ldots, B_k$ be disjoint subsets of the row-space $\mathcal{X}$.

- Define $q_j(x) = \left|\{i : x_i \in B_j\right|$.
- Then $M(x) = (q_1(x) + Z_1, \ldots, q_k(x) + Z_k)$ is ε-DP for
- $Z_k \sim \mathrm{Lap}(1/\varepsilon)$ if dataset adjacency is wrt d_{Sym} (add/remove a row)

Proof 1:

DP Histograms

Thm: Let $B_1, \ldots, B_k$ be disjoint subsets of the row-space $\mathcal{X}$.

- Define $q_j(x) = \left|\{i : x_i \in B_j\}\right|$.
- Then $M(x) = (q_1(x) + Z_1, \ldots, q_k(x) + Z_k)$ is ε-DP for
- $Z_k \sim \mathrm{Lap}(1/\varepsilon)$ if dataset adjacency is wrt d_{Sym} (add/remove a row)

Proof 2:

- Transformation $x \mapsto (q_1(x), \ldots, q_k(x))$ is a 1-stable map from d_{Sym} to the ℓ_1 metric on $\mathbb{R}^k$: $\|y - z\|_1 = \sum_{j=1}^{k} \left|y_j - z_j\right|$.
- Chain this transformation with the Vector Laplace mechanism.

Approximate Differential Privacy

Def: M is (ε, δ)-DP if for all $x \sim x'$, we have

$$\forall\, T \quad \Pr[M(x) \in T] \leq \mathrm{e}^{\varepsilon} \cdot \Pr[M(x') \in T] + \delta$$

- Intuitively: ε-DP with probability at least $1 - \delta$.
- Picking a random row & publishing is $(0, 1/n)$-DP, so want $\delta \ll 1/n$.
- Ideally δ is "cryptographically small," e.g. $\delta = 2^{-50}$.
- MIA interpretation: $\mathrm{TPR} \leq e^{\varepsilon} \cdot \mathrm{FPR} + \delta$
- Satisfies post-processing, basic composition (add δ_i's).
- Group privacy for groups up to size $O(1/\varepsilon)$.
- Does not suffice to check pointwise (need to consider sets T).

Benefits of Approximate DP

- More mechanisms, e.g. Gaussian Mechanism:

$$M(x,q) = q(x) + \mathcal{N}\left(0, \sigma^2\right),$$

$$\text{for } \sigma = \frac{\mathrm{GS}_q}{\varepsilon} \cdot \sqrt{2\ln(2/\delta)}$$

- Advanced Composition Thm: If M_i is (ε, δ)-DP for $i = 1, \ldots, k$ and $k < 1/\varepsilon^2$, then $\forall \delta > 0$

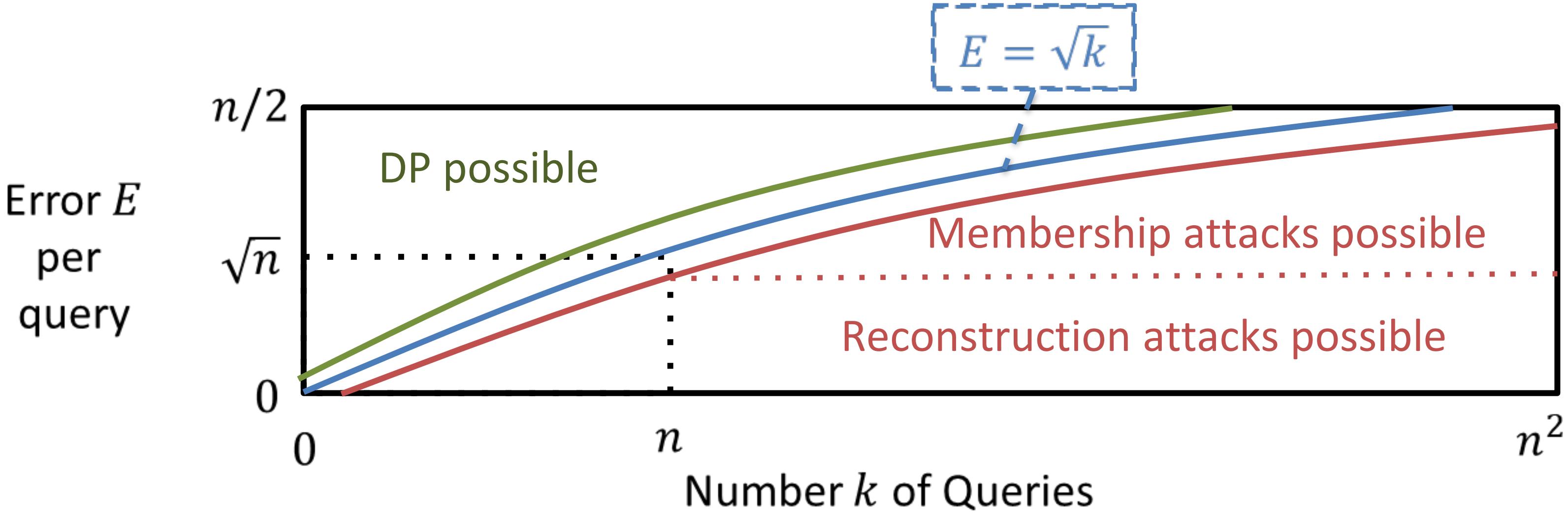
$$M\left(x, (q_1, \ldots, q_k)\right) = \left(M_1(x, q_1), \ldots, M_k(x, q_k)\right)$$

is $(\varepsilon', k \cdot \delta + \delta')$ −DP, for

$$\varepsilon' = O\left(\varepsilon \cdot \sqrt{k \cdot \log(1/\delta')}\right).$$

# Queries vs. Accuracy Tradeoff

- Using Laplace Mechanism to answer k queries, each with global sensitivity 1 (e.g. counts), under fixed privacy budget ε':
  - Set $\varepsilon = 1/\tilde{O}\left(\sqrt{k}\right)$ for each query (via Advanced Comp, hiding δ').
  - Add noise of scale $E = 1/\varepsilon \approx \tilde{O}\left(\sqrt{k}\right)$ per query.



Privacy Measures

- Pure DP:
  - privacy loss $\ln(\Pr[M(x) = y]/\Pr[M(x') = y])$ always $\leq \varepsilon$
  - $\mathrm{TPR} \leq e^{\varepsilon} \cdot \mathrm{FPR}$ at all FPR

- Approx DP:
  - privacy loss $\leq \varepsilon$, except with probability δ over $y \leftarrow M(x)$
  - $\mathrm{TPR} \leq e^{\varepsilon} \cdot \mathrm{FPR} + \delta$ at all FPR

- Other Measures:
  - bound distribution of the privacy loss random variable
  - or full FPR-TPR tradeoff
  - reason more cleanly or tightly about composition

zero-Concentrated DP (zCDP)

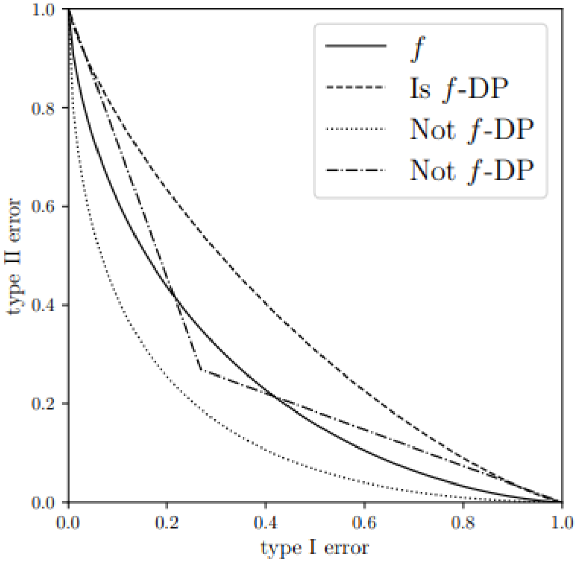
ρ-zCDP: privacy loss is “subGaussian” – dominated by a Gaussian r.v. with mean ρ and variance 2ρ

- ε-DP implies $(\varepsilon^2/2)$-zCDP
- ρ-zCDP implies $\left(\rho + 2\sqrt{\rho \log(1/\delta)}\,, \delta\right)$-DP for all δ
- Composition: $\rho_i's$ add up
- Gaussian mechanism: $M(x,q) = q(x) + \mathcal{N}(0, (\Delta q)^2/2\rho\ )$ is ρ-zCDP

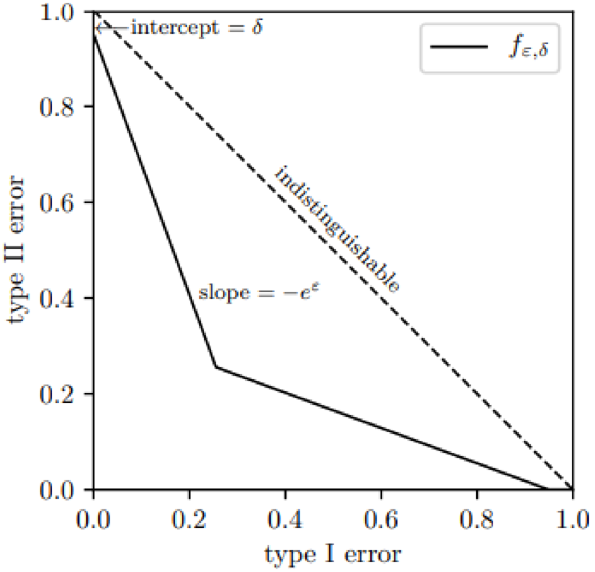
Benefits of approx. DP, advanced composition, Gaussian mechanism with only a single parameter to track

f-DP

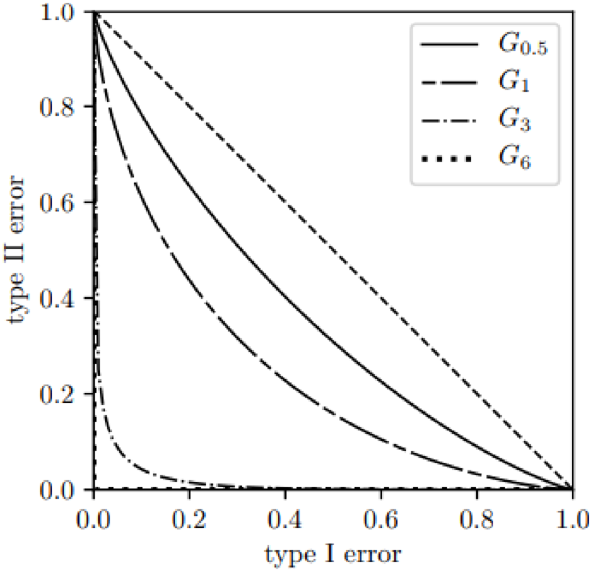
Privacy guarantees specified by a $f : [0,1] \to [0,1]$ s.t.
$\mathrm{FNR} \geq f(\mathrm{FPR})$ at all $\mathrm{FPR} \in [0,1]$
in distinguishing $H_0 = M(x)$ from $H_1 = M(x')$ for $x \sim x'$



Illustrating the def



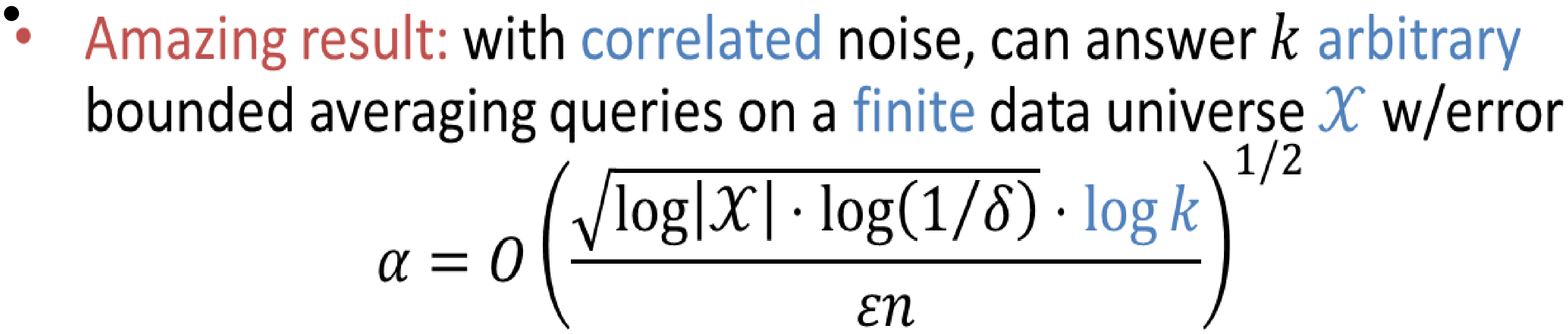
(ε, δ)-DP as f-DP



Gaussian mechanism

f-DP is equivalent to giving a full ε vs. δ curve (rather than a single pair).

An Amazing Possibility



- Amazing result: with correlated noise, can answer k arbitrary bounded averaging queries on a finite data universe $\mathcal{X}$ w/error

$$\alpha = O\left(\frac{\sqrt{\log|\mathcal{X}| \cdot \log(1/\delta)} \cdot \log k}{\varepsilon n}\right)^{1/2}$$

- The row domain $\mathcal{X}$ being bounded is enough to have error that grows much more slowly than $\sqrt{k}/n$.

- But many challenges in making this result practical (including computational complexity).