Rényi Differential Privacy

Ilya Mironov
Google Brain

DIMACS, October 24, 2017

Relaxation of Differential Privacy

(ε, δ)-Differential Privacy: The distribution of the output $M(D)$ on database D is (nearly) the same as $M(D')$:

$$\forall S: \quad \Pr[M(D) \in S] \leq \exp(\varepsilon) \cdot \Pr[M(D') \in S] + \delta.$$

privacy loss

allows for a small probability of failure

Dwork, Kenthapadi, McSherry, Mironov, Naor “Our data, Ourselves”, EUROCRYPT 2006

(ε, δ)-Differential Privacy

- Robust to auxiliary data
- Post-processing:

  If $M(D)$ is (ε, δ)-differentially private, so is $f(M(D))$.
- Composability:

  (ε_1,δ_1)-DP and (ε_2,δ_2)-DP is $(\varepsilon_1+\varepsilon_2,\delta_1+\delta_2)$-DP.
- Group privacy:

  Graceful degradation in the presence of correlated inputs.

Why (ε, δ)-Differential Privacy?

Three common use cases:

1. δ probability of failure
2. Gaussian noise
3. Advanced composition theorems

Why (ε, δ)-Differential Privacy?
I. δ probability of failure

Mechanism

Let $c \leftarrow \frac{4 \log |\mathcal{X}|}{\alpha^2}$,
if $\delta = 0$ **then**
  Let $T \leftarrow \frac{18c(\log(2|\mathcal{Q}|)+\log(4c/\beta))}{\epsilon \|x\|_1}$
else
  Let $T \leftarrow \frac{(2+}{}$
end if

Initialize Nu
$\{f_i'\}$, outputti
Let $t \leftarrow 0$, and
for each query f_i **do**
  Let $f'_{2i-1}(\cdot) = f_i(\cdot) - f_i(x^t$
  Let $f'_{2i}(\cdot) = f_i(x^t) - f_i$
  if $E_{2i-1} = \perp$ and $E_{2i} = \perp$ **then**
    Let $a_i = f_i(x^t)$
  else
    if $E_{2i-1} \in \mathbb{R}$ **then**
      Let $a_i = f_i(x^t) + E_{2i-1}$
    else
      Let $a_i = f_i(x^t) - E_{2i}$
    end if
    Let $x^{t+1} = MW(x^t, f_i, a_i)$
    Let $t \leftarrow t + 1$.
  end if
end for

if BAD then GIVE UP

Analysis

Proof. Unlike for multiplicative weights, it will be more convenient to analyze the Perceptron algorithm without normalizing the database to be a probability distribution, and then prove that it is a $T(\alpha')$ database update algorithm for T complete the proof. I view $f_t \in [0,1]^{|\mathcal{X}|}$ as to $\langle f_t, x \rangle$.

We must show tha erty that $|f_t(x^t) - f$ $L > \frac{\|x\|_2^2 |\mathcal{X}|}{\alpha'^2}$.

We use a potential argume to show that for every $t = 1, 2, \ldots, L$, x^{t+1} is significantly closer to x than x^t. Specifically, our potential function is the L_2^2 norm of the database $x - x^t$, defined as

$$\|x\|_2^2 = \sum_{i \in \mathcal{X}} x(i)^2.$$

Observe that $\|x - x^1\|_2^2 = \|x\|_2^2$ since $x^1 = 0$, and $\|x\|_2^2 \geq 0$. Thus it suffices to show that in every step, the potential decreases by $\alpha'^2/|\mathcal{X}|$. We analyze the case where $f_t(x^t) > v_t$, the analysis for the opposite case will be similar. Let $R^t = x^t - x$. Observe that in this case we have

$$f_t(R^t) = f_t(x^t) - f_t(x) \geq \alpha'.$$

$\Pr[\text{BAD}] < \delta$

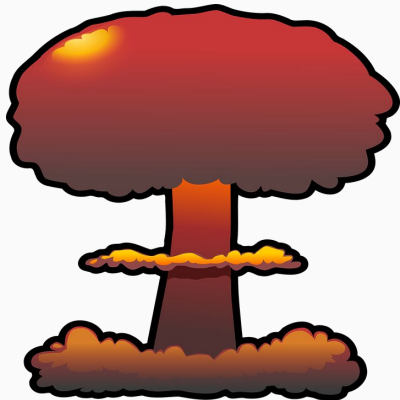
What Can δ Hide?

"Nuclear Option":

- With probability δ publish everything
- With probability 1 publish δ fraction of inputs

Recommendations

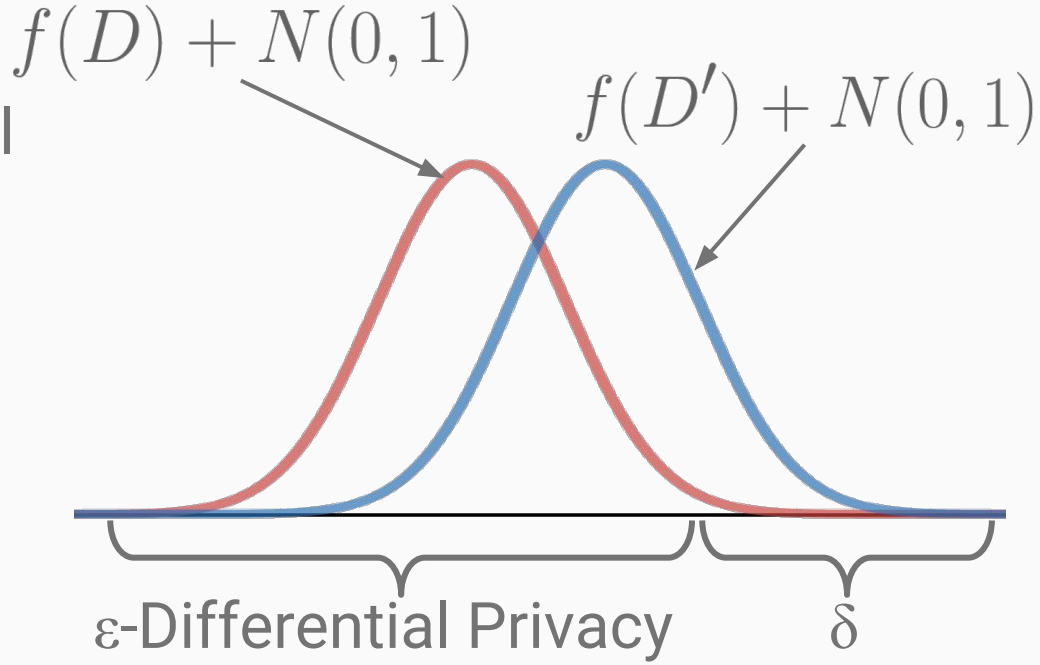
$\delta \ll 1/N$ or $\delta = \text{negl}(1/N)$



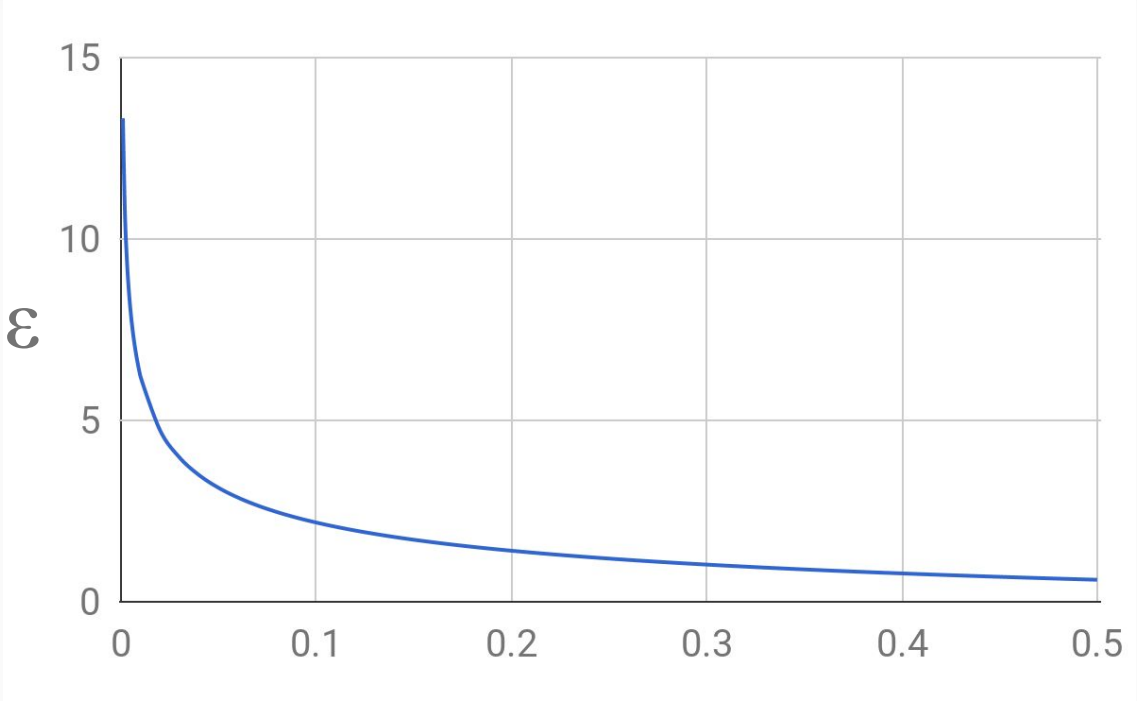
Why (ε, δ)-Differential Privacy?
II. Gaussian Mechanism

Normal distribution:

- Spherically symmetrical
- Closed under addition
- “Noise-like”
- Tightly concentrated



(ε, δ)-Differential Privacy of Gaussian



Why (ε, δ)-Differential Privacy?
III. Advanced Composition Theorem

Basic Composition:

Composition of ε_1-DP and ε_2-DP is $(\varepsilon_1+\varepsilon_2)$-DP

n-fold composition of ε-DP is $n\varepsilon$-DP

Advanced Composition:

n-fold composition of ε-DP is $(\sqrt{2n\ln(1/\delta)}\,\varepsilon, \delta)$-DP for $\delta < 1$

Proof of Advanced Composition

1. Privacy loss variable

$$R \sim \ln \frac{\Pr[x=f(D)]}{\Pr[x=f(D')]} \text{ where } x \sim f(D).$$

2. Azuma inequality for (α, β)-martingales:

$R_1,\ldots, R_n$ such that $|R_i| \leq \alpha$ and $\mathrm{E}[R_i|R_1,\ldots,R_{i-1}] \leq \beta$ then

$$\Pr\left[\sum_{i=1}^{n} R_i > n\beta + z\sqrt{n}\alpha\right] \leq e^{-z^2/2}$$

3. ε-DP ⇒ privacy loss variable is a $(\varepsilon, \varepsilon^2)$-martingale

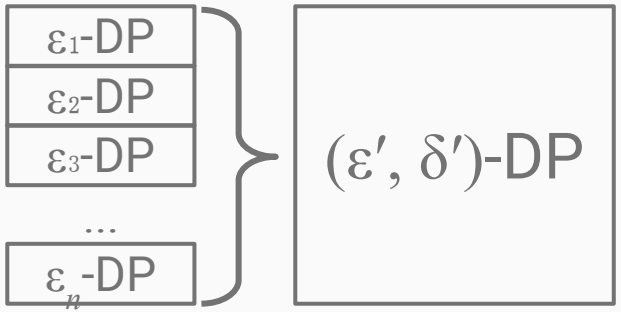
Dwork, Rothblum, Vadhan “Boosting and Differential Privacy”, FOCS 2010

Trouble with (ε, δ)-Differential Privacy

- Gaussian mechanism does not have catastrophic failure!

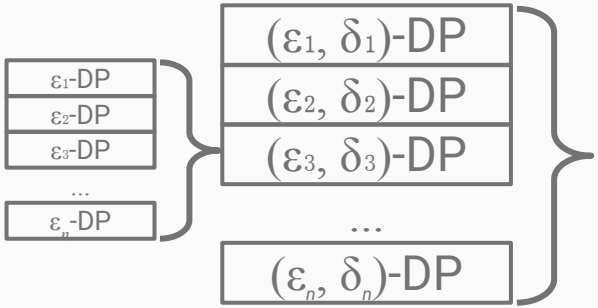
Trouble with (ε, δ)-Differential Privacy

- Gaussian mechanism does not have catastrophic failure!
- Composing advanced composition



Trouble with (ε, δ)-Differential Privacy

- Gaussian mechanism does not have catastrophic failure!
- Composing advanced composition

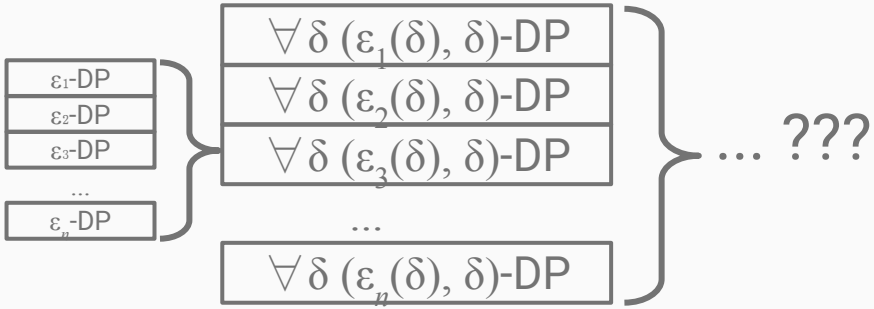


... #P hard!

Murtagh, Vadhan, ``The complexity of computing the optimal composition of differential privacy", TCC 2016-A.

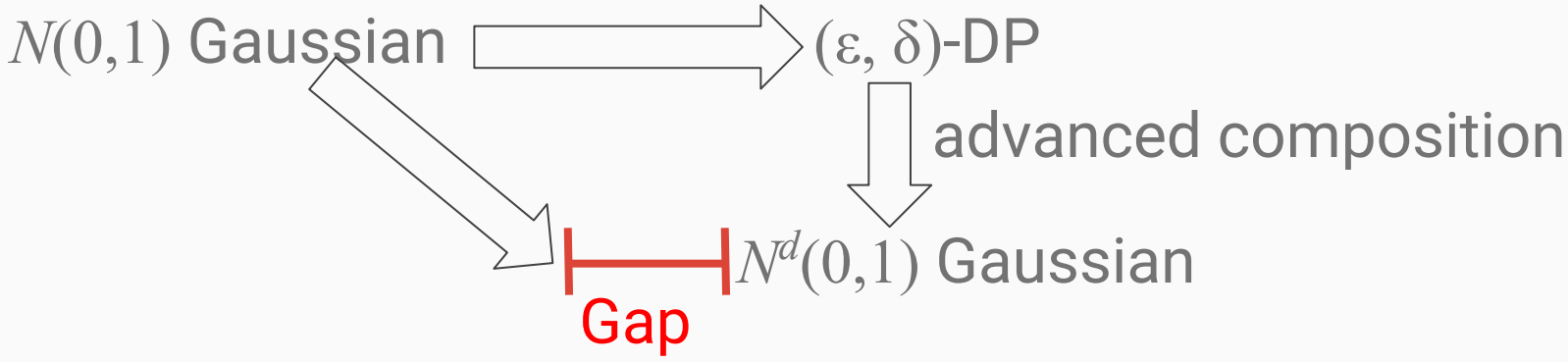
Trouble with (ε, δ)-Differential Privacy

- Gaussian mechanism does not have catastrophic failure!
- Composing advanced composition



Trouble with (ε, δ)-Differential Privacy

- Gaussian mechanism does not have catastrophic failure!
- Composing advanced composition
- Gaussian + Advanced composition is not tight



Why (ε, δ)-Differential Privacy?

Three common use cases:

1. δ probability of failure
2. Gaussian noise
3. Advanced composition theorems

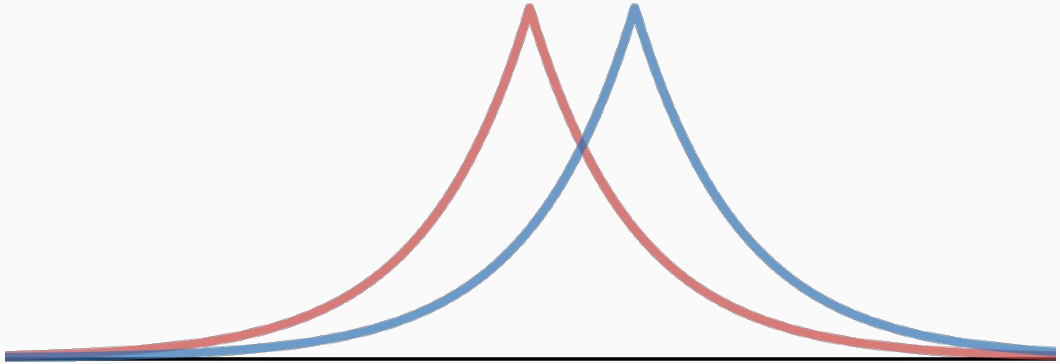
Better Notion of Closeness

ε-Differential Privacy:

$$\max_x P(x)/Q(x) < e^{\varepsilon}$$

Rényi Divergence at ∞:

$$D_\infty(P\|Q) < \varepsilon$$



Rényi Divergence

$$D_1(P\|Q) = \lim_{\alpha\to 1} D_\alpha(P\|Q) = E_P\left[\log \frac{P(x)}{Q(x)}\right]$$

$$D_\alpha(P\|Q) = \frac{1}{\alpha - 1}\log E_Q\left[\left(\frac{P(x)}{Q(x)}\right)^\alpha\right]$$

$$D_\infty(P\|Q) = \lim_{\alpha\to\infty} D_\alpha(P\|Q) = \log \max_x \frac{P(x)}{Q(x)}$$

Rényi Differential Privacy

(α, ε)-Rényi Differential Privacy (RDP):

$$\forall\ D, D' \colon D_{\alpha}(M(D) \,\|\, M(D')) \leq \varepsilon.$$

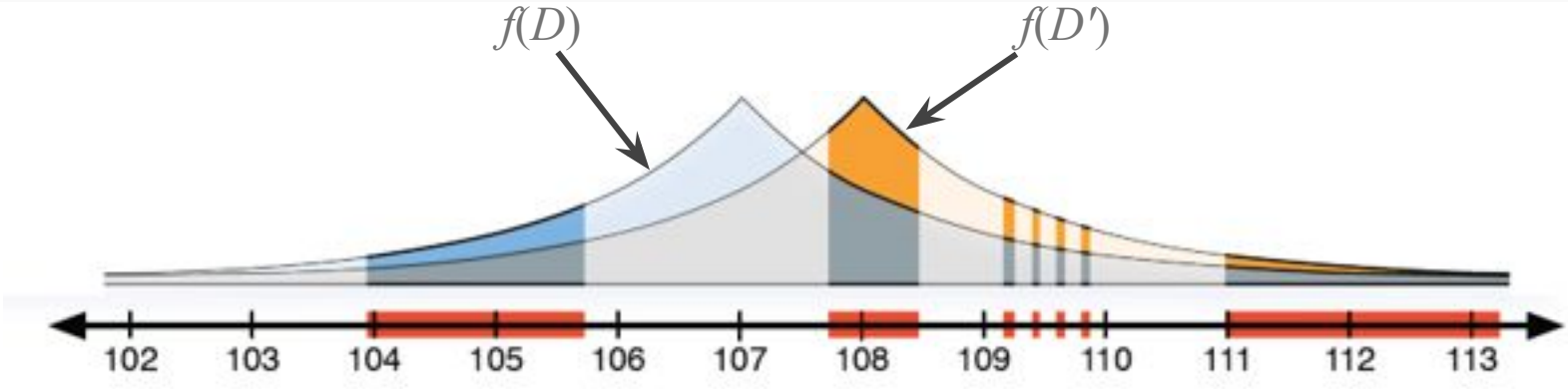
Relaxation of Differential Privacy

(∞, ε)-RDP is ε-DP

Implies (ε, δ)-Differential Privacy

(α, ε)-RDP $\Rightarrow$ $\left(\varepsilon + \frac{\log 1/\delta}{\alpha - 1}, \delta\right)$-DP for any δ

“Bad Outcomes” Interpretation



— bad outcomes

— probability with record x

— probability without record x

“Bad Outcomes” Interpretation

ε-Differential Privacy: $\forall S \;\; \Pr[M(D) \in S] \leq e^{\varepsilon} \cdot \Pr[M(D') \in S]$

(α, ε)-Rényi Diff Privacy: $\forall S \quad \Pr[M(D) \in S] \leq (e^{\varepsilon} \cdot \Pr[M(D') \in S])^{1-1/\alpha}$

(ε, δ)-Differential Privacy: $\forall S \;\; \Pr[M(D) \in S] \leq e^{\varepsilon} \cdot \Pr[M(D') \in S] + \delta$

No Catastrophic Failure Mode!

$$\Pr[M(D)\in S] \leq (e^{\varepsilon} \cdot \Pr[M(D')\in S])^{1-1/\alpha}$$

Monotonicity

For $\alpha_1 \geq \alpha_2$:

(α_1, ε)-RDP $\Rightarrow$ (α_2, ε)-RDP

Composable!

Simultaneous release of (α, ε_1)-RDP and (α, ε_2)-RDP is

$(\alpha, \varepsilon_1+\varepsilon_2)$-RDP

Proof of Advanced Composition

1. Privacy loss variable

$$R \sim \ln \frac{\Pr[x=f(D)]}{\Pr[x=f(D')]}$$ where

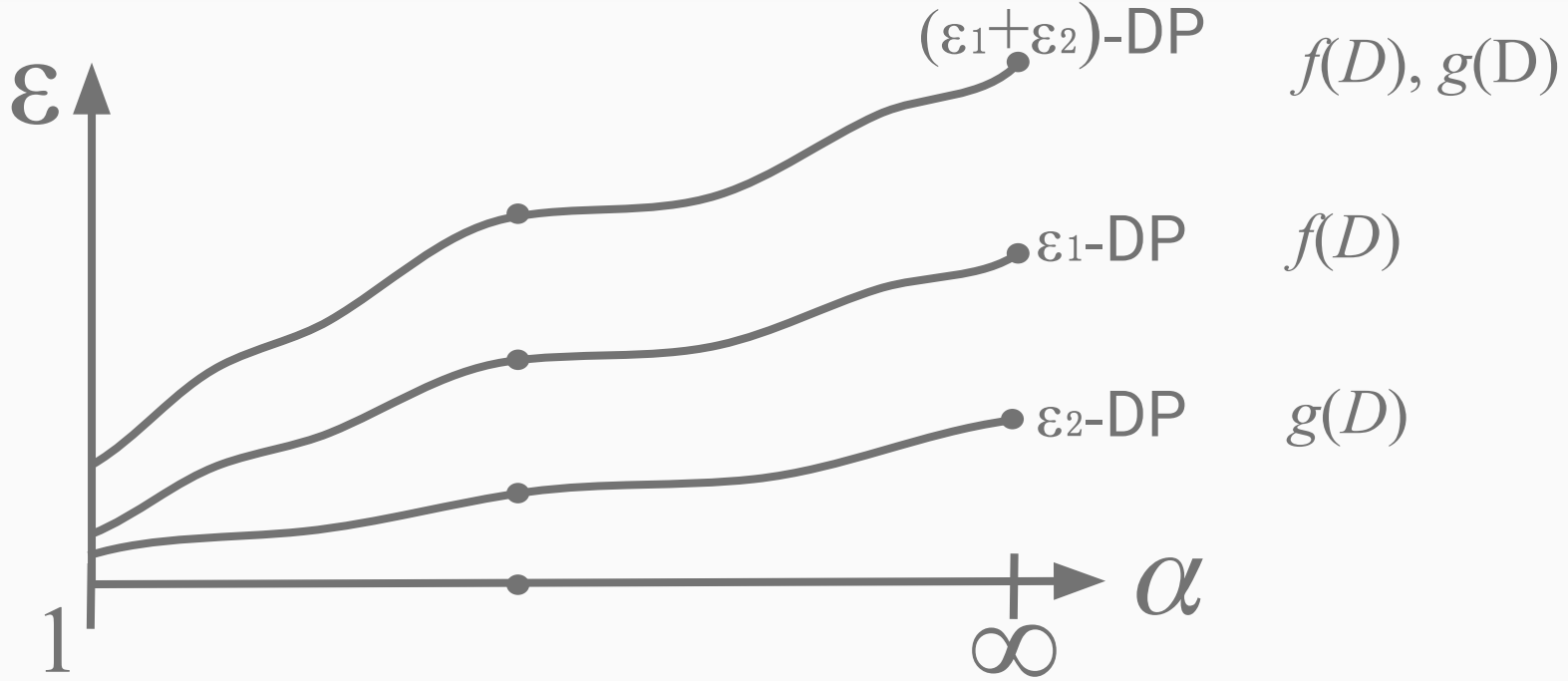
2. Azuma inequality for (α,
   $R_1,\ldots, R_n$ such that |
   $\Pr\left[\sum_{i=1}^{n} R_i > \right.$ hen

3. ε-DP ⇒ privacy loss variable is ale

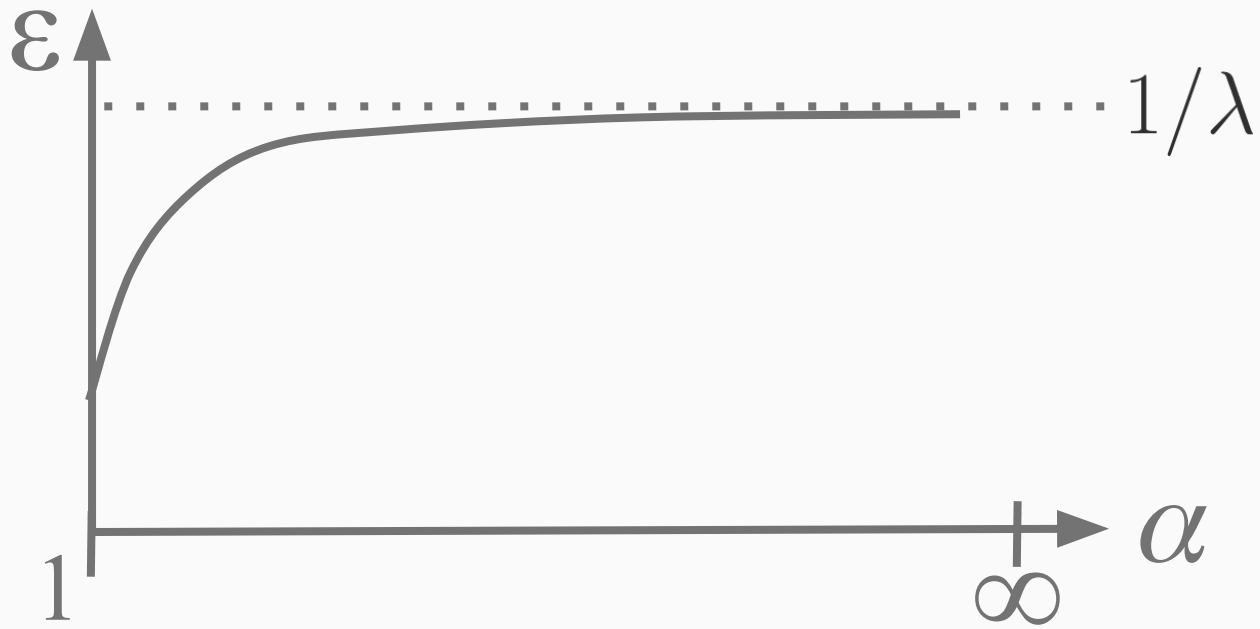
1. Switch to $e^{\lambda R}$
2. Apply Markov's to E
3. Optimize λ

Dwork, Rothblum, Vadhan "Boosting and Differential Privacy", FOCS 2010

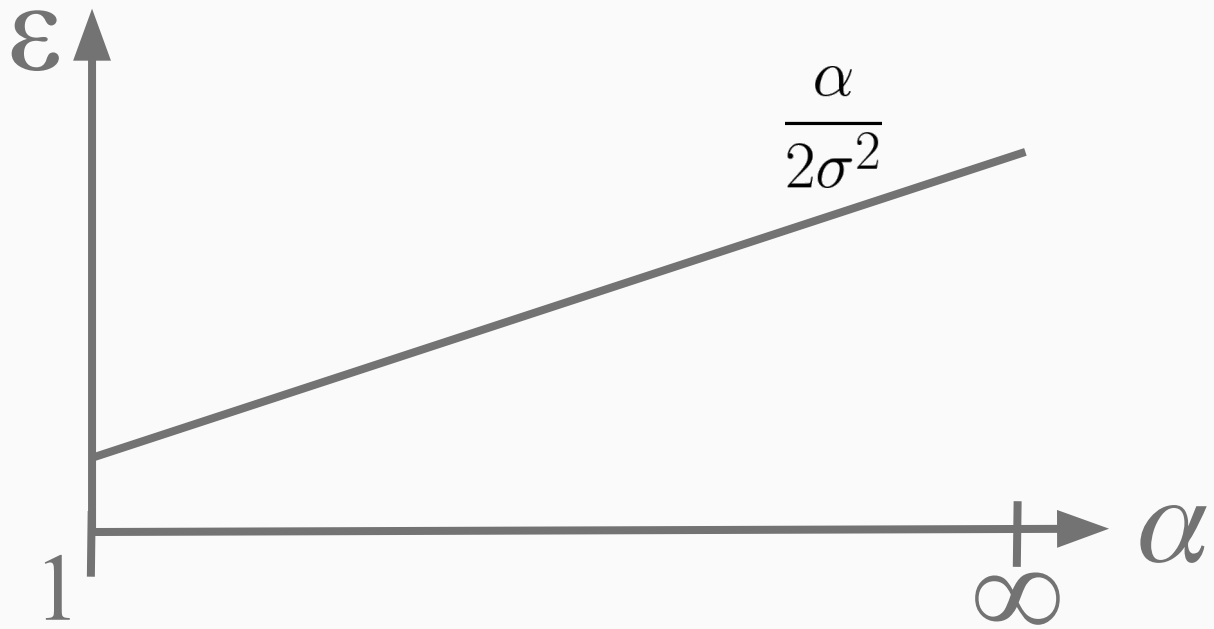
Rényi Budget Curve



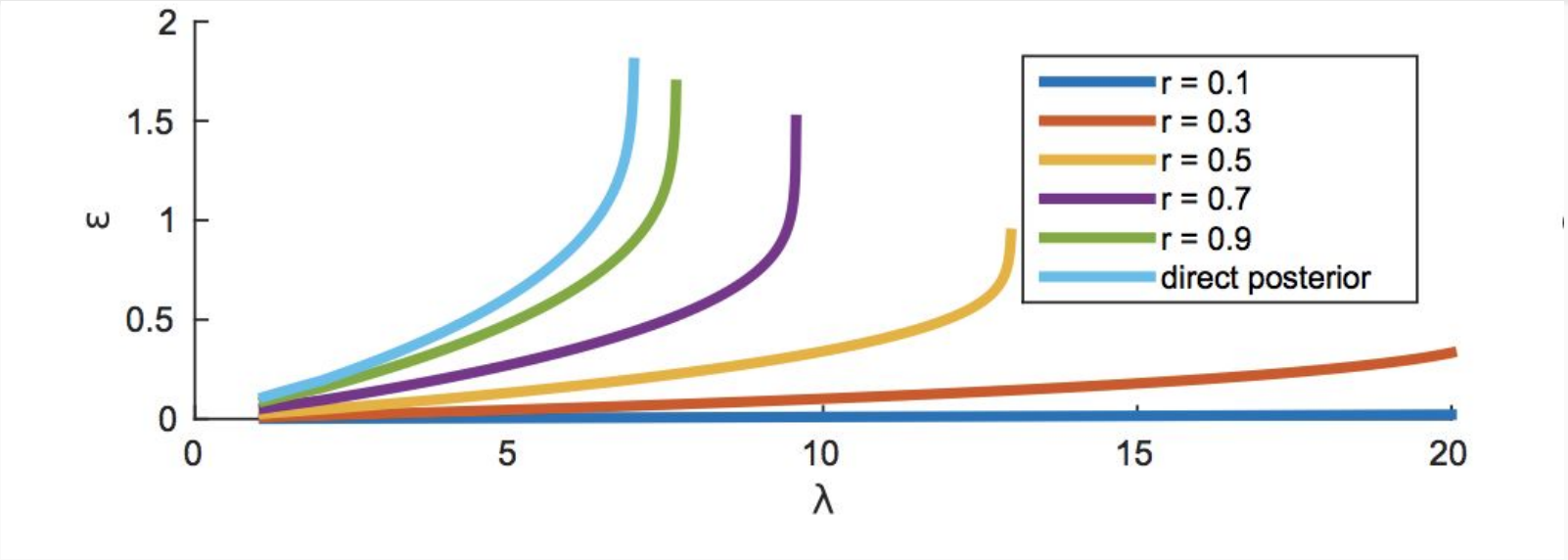
Rényi Budget Curve: Laplace Mechanism



Rényi Budget Curve: Gaussian Mechanism

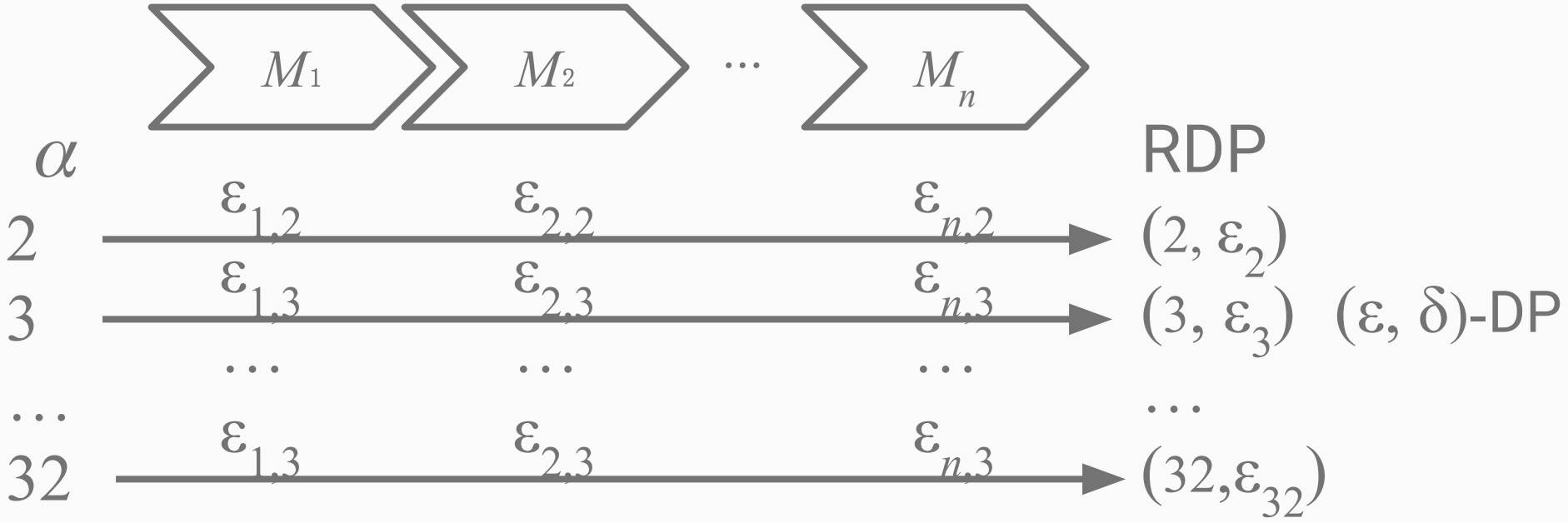


More Complex Mechanisms



Geumlek, Song, Chaudhuri, “Renyi Differential Privacy Mechanisms for Posterior Sampling”, (NIPS 2017)

Rényi Differential Privacy as a Privacy Accountant



Why 2 to 32?

(α, ε)-Rényi Diff Privacy: $\forall S \quad \Pr[M(D) \in S] \leq (e^{\varepsilon} \cdot \Pr[M(D') \in S])^{1-1/\alpha}$

Convergence of Ideas

Prior work:	Dwork, Rothblum, Vadhan (FOCS 2010)
	Dwork, Rothblum (2016)
	Lattice-based cryptography
This work:	“Rényi Differential Privacy” (CSF 2016)
Concurrent work:	Bun, Steinke (TCC 2016-B)
	Abadi et al. (ACM CCS 2016)
	Papernot et al. (ICLR 2017)
Applications:	Geumlek, Song, Chaudhuri (NIPS 2017)

Summary

- Rényi Differential Privacy: generalization and relaxation of differential privacy
- RDP fixes problems of (ε, δ)-DP:
  - No catastrophic failure mode
  - Tight analysis of Gaussian noise
  - Easy composition of heterogeneous mechanisms