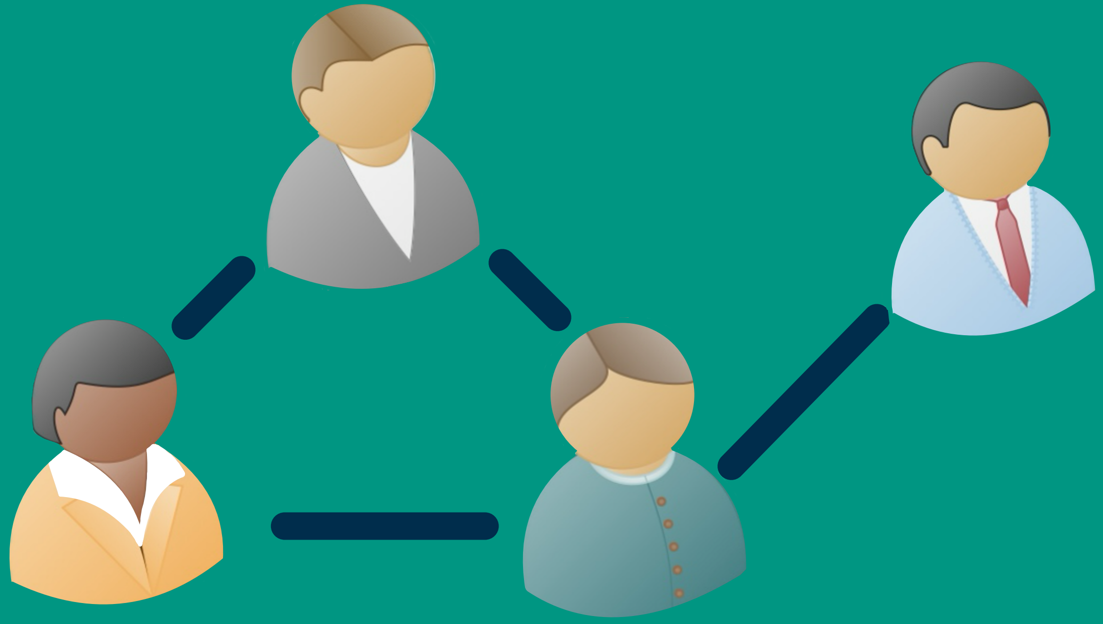


Balancing Privacy and Utility in Correlated Data

INRIA Montpellier, November 12th, 2025
Patricia Guerra-Balboa

PRIVACY AND SECURITY

KASTEL

Our General Goal

Learn population-level information without harming individual’s privacy

Our General Goal

Learn population-level information without harming individual's privacy

Privacy Goal: Protect Alice's location

Utility Goal: Number of cars per street

Our General Goal

Learn population-level information without harming individual's privacy

Privacy Goal: Protect Alice's activity data

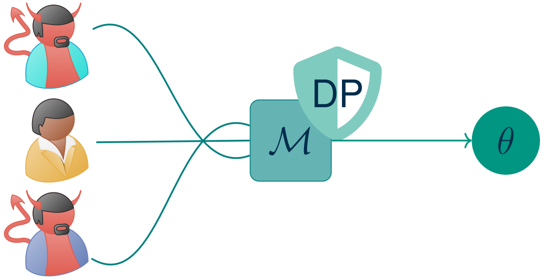
Utility Goal: Correlation between height and health

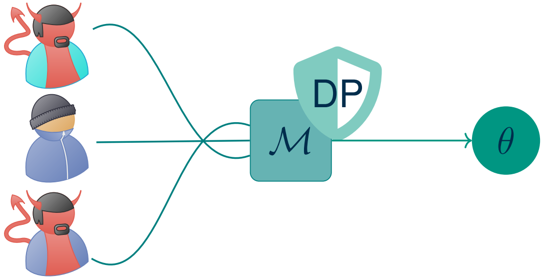
The Best Tool Until Now: Differential Privacy

Idea: We want to bound participation risk.

The Best Tool Until Now: Differential Privacy

Idea: We want to bound participation risk.

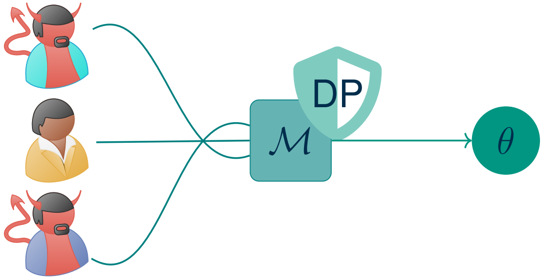


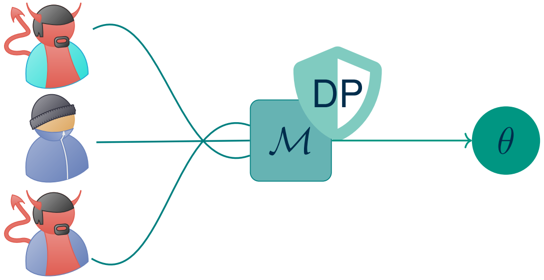


- "Strongest" assumption: everybody's record is known but the target.

The Best Tool Until Now: Differential Privacy

Idea: We want to bound participation risk.





- "Strongest" assumption: everybody's record is known but the target.

The Best Tool Until Now: Differential Privacy

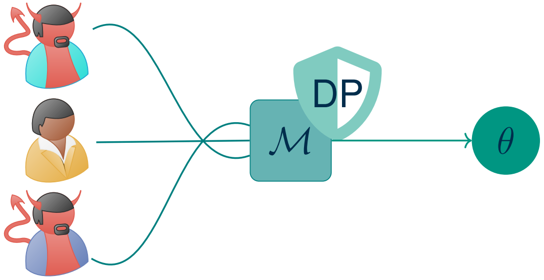
Idea: We want to bound participation risk.

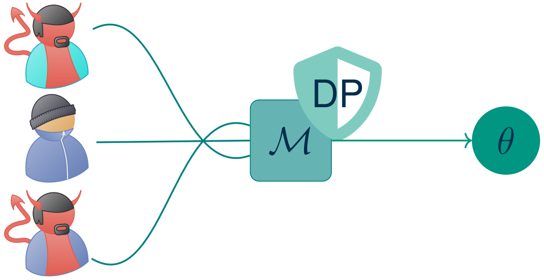
$$ln \frac{p_{\mathcal{M}}(\theta \mid x_1, \ldots, x_{n-1}, \mathbf{x_n})}{p_{\mathcal{M}}(\theta \mid x_1, \ldots, x_{n-1}, \mathbf{y_n})} \leq \mathcal{E}$$

- "Strongest" assumption: everybody's record is known but the target.
- The privacy leakage ε controls the indistinguishability level between $\mathbf{x_n}, \mathbf{y_n}$.

The Best Tool Until Now: Differential Privacy

Idea: We want to bound participation risk.



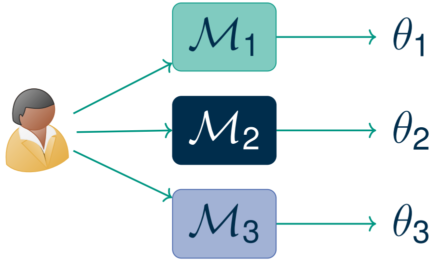


$$ln \frac{p_{\mathcal{M}}(\theta \mid x_1, \ldots, x_{n-1}, \mathbf{x_n})}{p_{\mathcal{M}}(\theta \mid x_1, \ldots, x_{n-1}, \mathbf{y_n})} \leq \mathcal{E}$$

- "Strongest" assumption: everybody's record is known but the target.
- The privacy leakage ε controls the indistinguishability level between $\mathbf{x_n}, \mathbf{y_n}$.
- But at some cost! The smaller the ε the less utility.

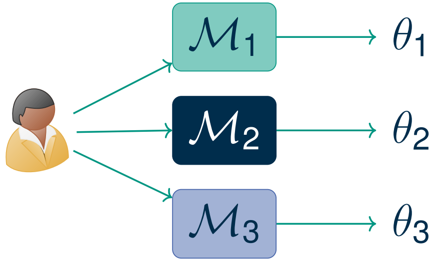
Why DP Is The Best So Far?

Composition

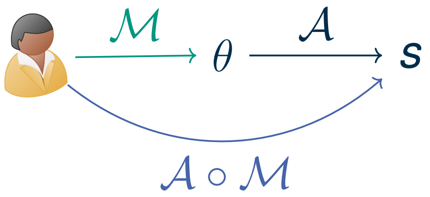


Why DP Is The Best So Far?

Composition

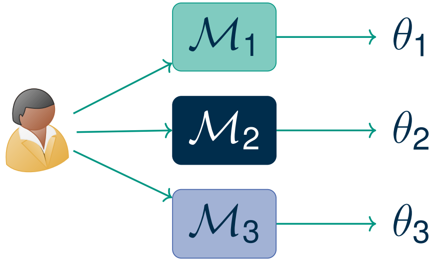


Post-processing

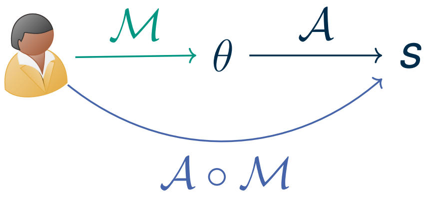


Why DP Is The Best So Far?

Composition



Post-processing



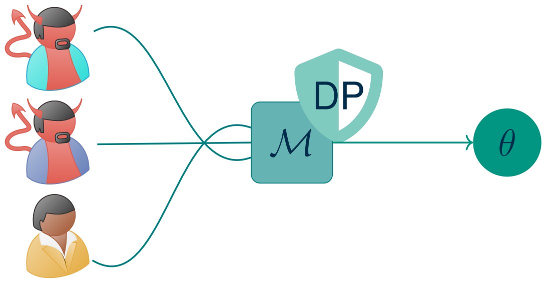
Attack Mitigation

$$\mathcal{M} \quad \varepsilon\text{-DP} \Rightarrow \mathrm{Adv} \leq f(\varepsilon)$$

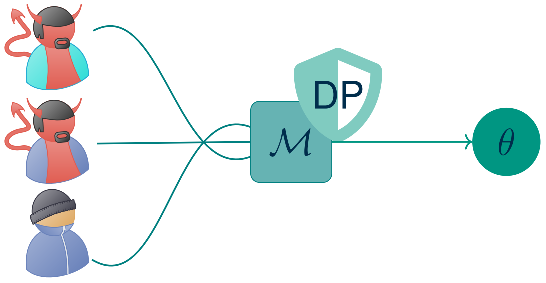
Membership Inference Attack Knowing $D_{\_}$

The attacker receives θ and aims to distinguish between:

$H_0 : x_n$



$H_1 : y_n$



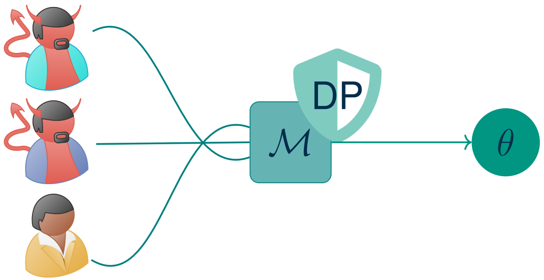
$D_{\_}$ is known:

$H_0 = D_{x_n}$ Vs. $H_1 = D_{y_n}$

Membership Inference Attack Knowing D_{-}

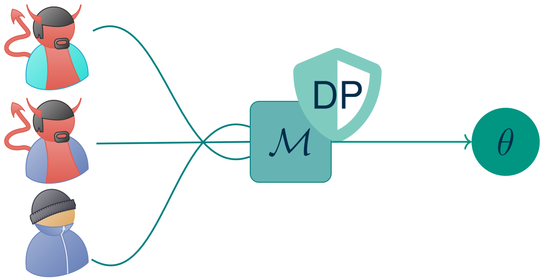
The attacker receives θ and aims to distinguish between:

$H_0 : x_n$



Type I error:
$$\alpha = \Pr_{A \circ \mathcal{M}}(y_n \mid D_{x_n})$$

$H_1 : y_n$



Type II error:
$$\beta = 1 - \Pr_{A \circ \mathcal{M}}(y_n \mid D_{y_n})$$

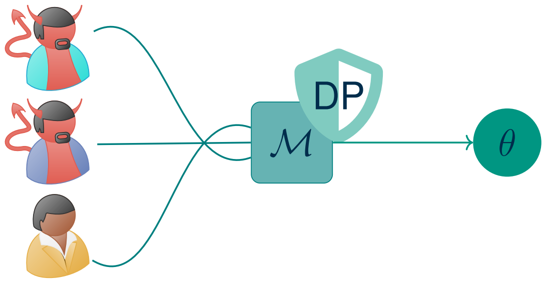
D_{-} is known:

$$H_0 = D_{x_n} \text{ Vs. } H_1 = D_{y_n}$$

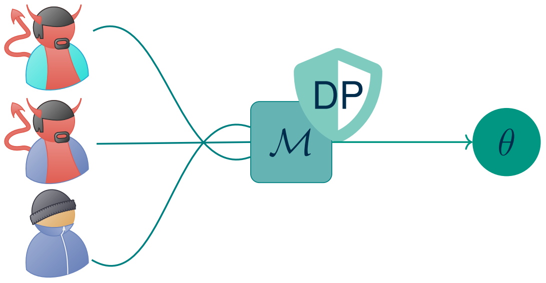
Membership Inference Attack Knowing D_{-}

The attacker receives θ and aims to distinguish between:

$H_0 : x_n$



$H_1 : y_n$



D_{-} is known:

$H_0 = D_{x_n}$ Vs. $H_1 = D_{y_n}$

Type I error:
$\alpha = \Pr_{A \circ \mathcal{M}}(y_n \mid D_{x_n})$

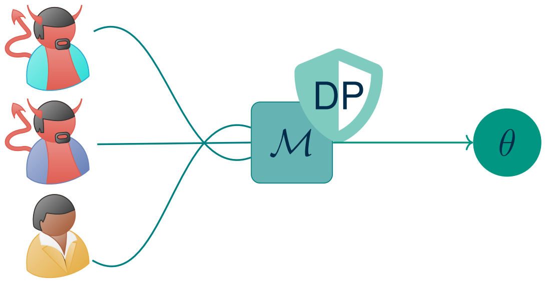
Type II error:
$\beta = 1 - \Pr_{A \circ \mathcal{M}}(y_n \mid D_{y_n})$

$A \circ \mathcal{M}$ is ε-DP $\implies$

Membership Inference Attack Knowing D_{-}

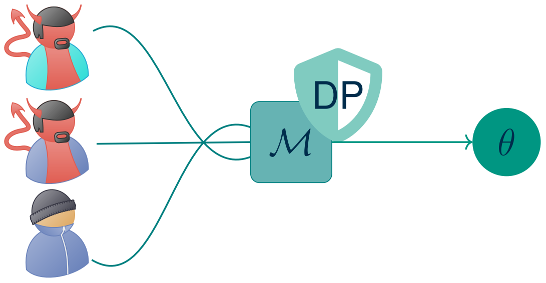
The attacker receives θ and aims to distinguish between:

$H_0 : x_n$



Type I error:
$\alpha = \Pr_{A \circ \mathcal{M}}(y_n \mid D_{x_n})$

$H_1 : y_n$



Type II error:
$\beta = 1 - \Pr_{A \circ \mathcal{M}}(y_n \mid D_{y_n})$

D_{-} is known:

$H_0 = D_{x_n}$ Vs. $H_1 = D_{y_n}$

$A \circ \mathcal{M}$ is ε-DP $\implies$

$$1 - \beta \leq e^{\varepsilon}\alpha$$

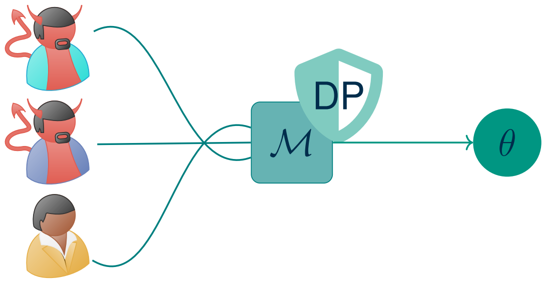
$$\alpha \leq e^{\varepsilon}(1 - \beta)$$

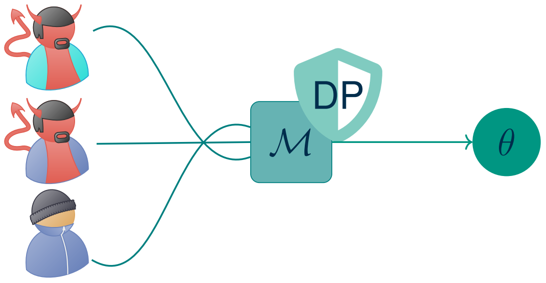
Membership Inference Attack Knowing $D\_$

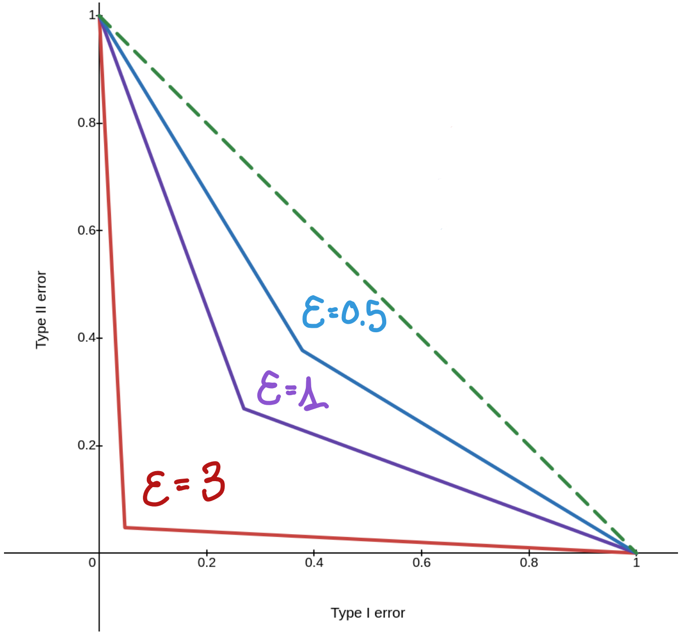
The attacker receives θ and aims to distinguish between:

$H_0 : x_n$

$H_1 : y_n$







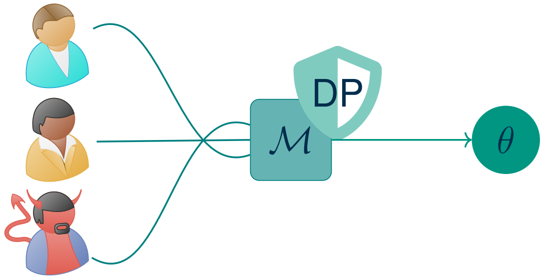
Type I error:
$\alpha = \Pr_{A\circ\mathcal{M}}(y_n \mid D_{x_n})$

Type II error:
$\beta = 1 - \Pr_{A\circ\mathcal{M}}(y_n \mid D_{y_n})$

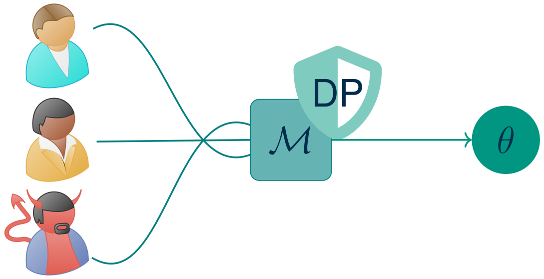
$A \circ \mathcal{M}$ is ε-DP $\Longrightarrow$ $1-\beta \le e^{\varepsilon}\alpha$, $\alpha \le e^{\varepsilon}(1-\beta)$ $\Longrightarrow$

$$\beta \ge \max\{1 - e^{\varepsilon}\alpha, e^{\varepsilon}(1-\alpha)\}$$

What about other attacker models?



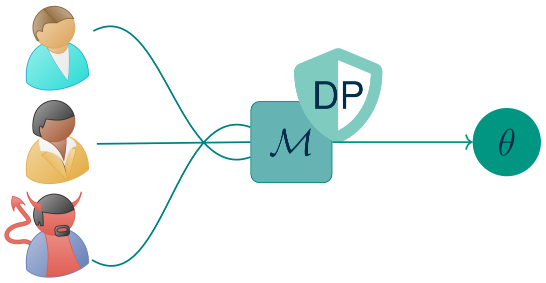
What about other attacker models?

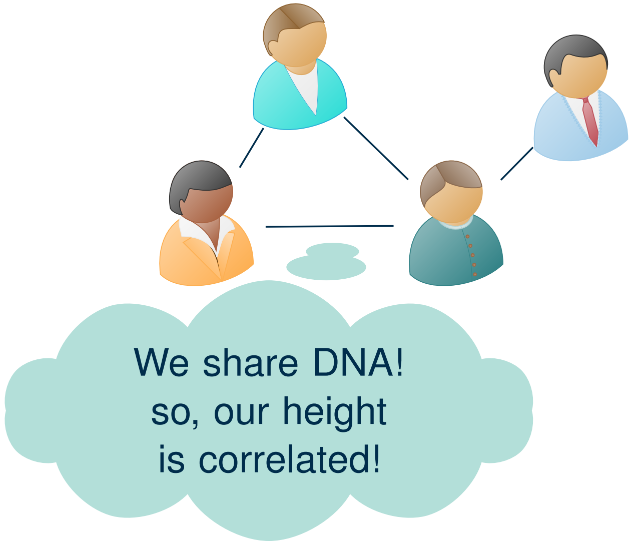


Statistical Independence

The strongest attacker is the worst-case one, and we have at least the same protection.

What about other attacker models?





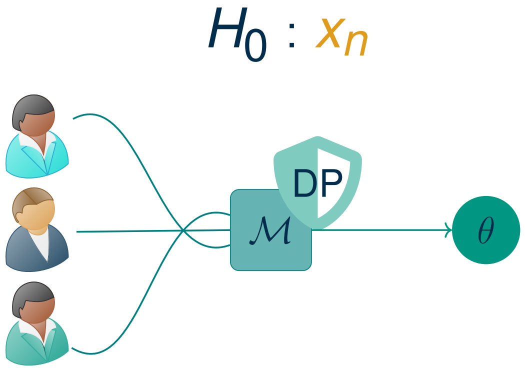
Statistical Independence
Dependencies between Records

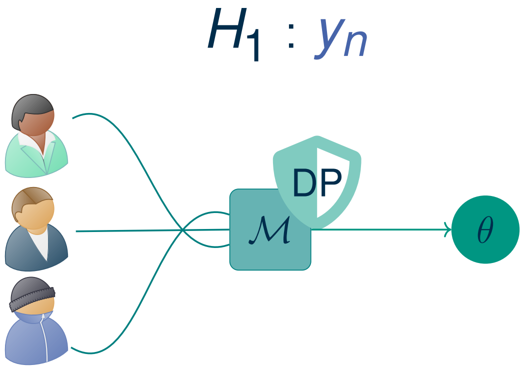
The strongest attacker is the worst-case one, and we have at least the same protection.

DP interpretation does not hold anymore.

Membership Inference Attack With Dependencies

The attacker receives θ and aims to distinguish between:





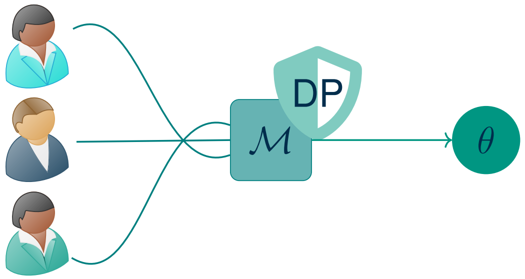
D_- is **unknown**:

$$H_0 = \{D \colon x_n \in D\} \text{ Vs. } H_1 = \{D \colon y_n \in D\}$$

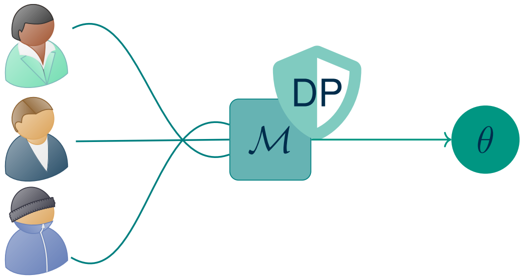
Membership Inference Attack With Dependencies

The attacker receives θ and aims to distinguish between:

$H_0 : x_n$



$H_1 : y_n$



D_- is **unknown**:

$$H_0 = \{D \colon x_n \in D\} \text{ Vs. } H_1 = \{D \colon y_n \in D\}$$

Type I error:

$$\alpha = \Pr_{A \circ \mathcal{M}}(y_n \mid x_n)$$

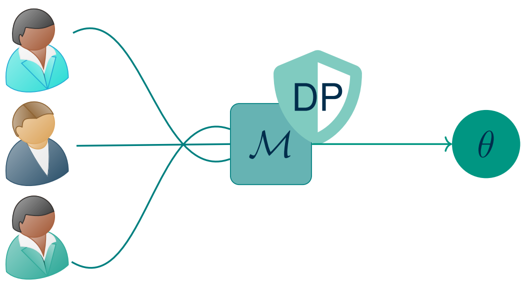
Type II error:

$$\beta = 1 - \Pr_{A \circ \mathcal{M}}(y_n \mid y_n)$$

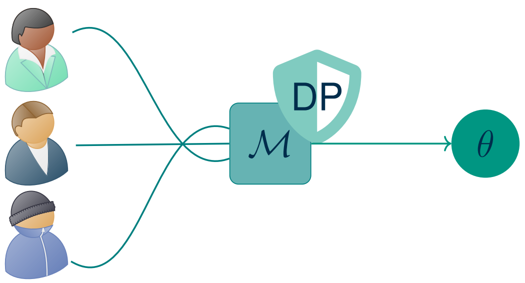
Membership Inference Attack With Dependencies

The attacker receives θ and aims to distinguish between:

$H_0 : x_n$



$H_1 : y_n$



D_- is **unknown**:

$H_0 = \{D\colon x_n \in D\}$ Vs. $H_1 = \{D\colon y_n \in D\}$

Type I error:

$$\alpha = \Pr_{A\circ\mathcal{M}}(y_n \mid x_n) = \sum_{D_-} \Pr_{A\circ\mathcal{M}}(y_n \mid D_{x_n})\,\pi(D_- \mid x_n)$$

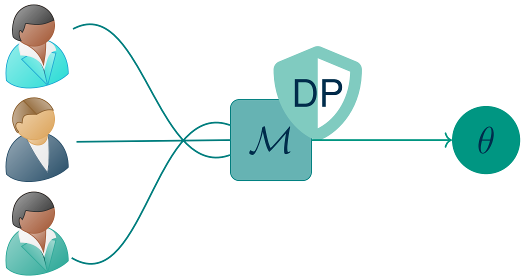
Type II error:

$$\beta = 1 - \Pr_{A\circ\mathcal{M}}(y_n \mid y_n) = 1 - \sum_{D_-} \Pr_{A\circ\mathcal{M}}(y_n \mid D_{y_n})\,\pi(D_- \mid y_n)$$

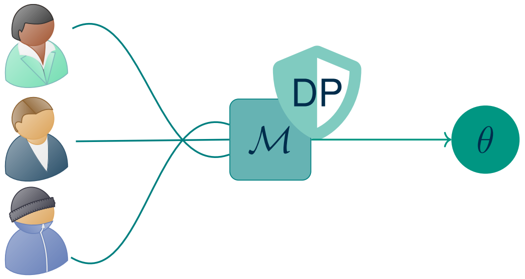
Membership Inference Attack With Dependencies

The attacker receives θ and aims to distinguish between:

$H_0 : x_n$



$H_1 : y_n$



D_{-} is **unknown**:

$H_0 = \{D \colon x_n \in D\}$ Vs. $H_1 = \{D \colon y_n \in D\}$

Type I error:

$$\alpha = \Pr_{A\circ\mathcal{M}}(y_n \mid x_n) = \sum_{D_-} \Pr_{A\circ\mathcal{M}}(y_n \mid D_{x_n})\, \pi(D_- \mid x_n)$$

Type II error:

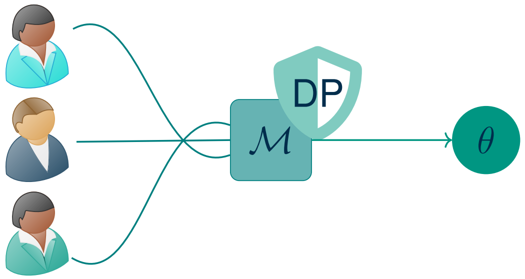
$$\beta = 1 - \Pr_{A\circ\mathcal{M}}(y_n \mid y_n) = 1 - \sum_{D_-} \Pr_{A\circ\mathcal{M}}(y_n \mid D_{y_n})\, \pi(D_- \mid y_n)$$

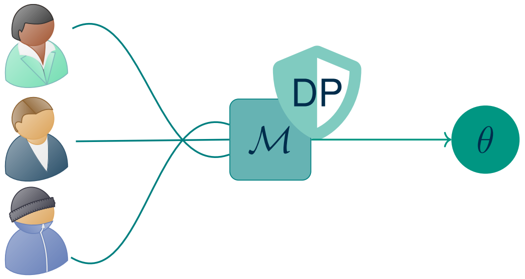
$A \circ \mathcal{M}$ is ε-DP $\implies$ $1-\beta \leq \sum_{D_-} e^{\varepsilon} \Pr_{A\circ\mathcal{M}}(y_n \mid D_{x_n})\pi(D_- \mid y_n)$

Membership Inference Attack With Dependencies

The attacker receives θ and aims to distinguish between:

$H_0 : x_n$ $\quad\quad$ $H_1 : y_n$





D_- is **unknown**:

$$H_0 = \{D \colon x_n \in D\} \text{ Vs. } H_1 = \{D \colon y_n \in D\}$$

Type I error:

$$\begin{aligned}\alpha &= \Pr_{A\circ\mathcal{M}}(y_n \mid x_n) \\ &= \sum_{D_-} \Pr_{A\circ\mathcal{M}}(y_n \mid D_{x_n})\,\pi(D_- \mid x_n)\end{aligned}$$

Type II error:

$$\begin{aligned}\beta &= 1 - \Pr_{A\circ\mathcal{M}}(y_n \mid y_n) \\ &= 1 - \sum_{D_-} \Pr_{A\circ\mathcal{M}}(y_n \mid D_{y_n})\,\pi(D_- \mid y_n)\end{aligned}$$

$A \circ \mathcal{M}$ is ε-DP $\implies$ $1-\beta \leq \sum_{D_-} e^{\varepsilon} \Pr_{A\circ\mathcal{M}}(y_n \mid D_{x_n})\pi(D_- \mid y_n) = e^{\varepsilon} \sum_{D_-} \Pr_{A\circ\mathcal{M}}(y_n \mid D_{x_n})\pi(D_- \mid y_n) \neq e^{\varepsilon}\alpha$

Standard DP Underestimates Participation Risk

Differential Privacy fails to measure privacy leakage under correlation

Empirically confirmed

Theoretically exposed

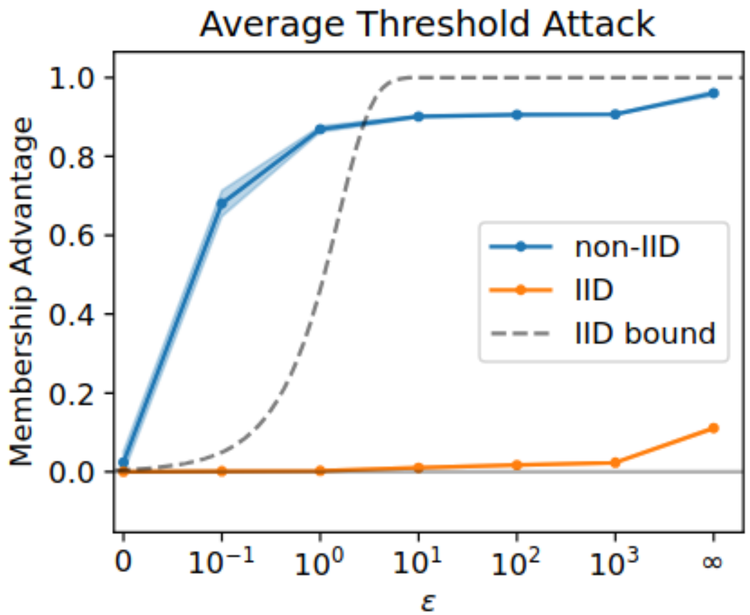


Figure: Humphries et al. 2023 MIA breaks DP guarantees.

Standard DP Underestimates Participation Risk

Differential Privacy fails to measure privacy leakage under correlation

Empirically confirmed

Theoretically exposed

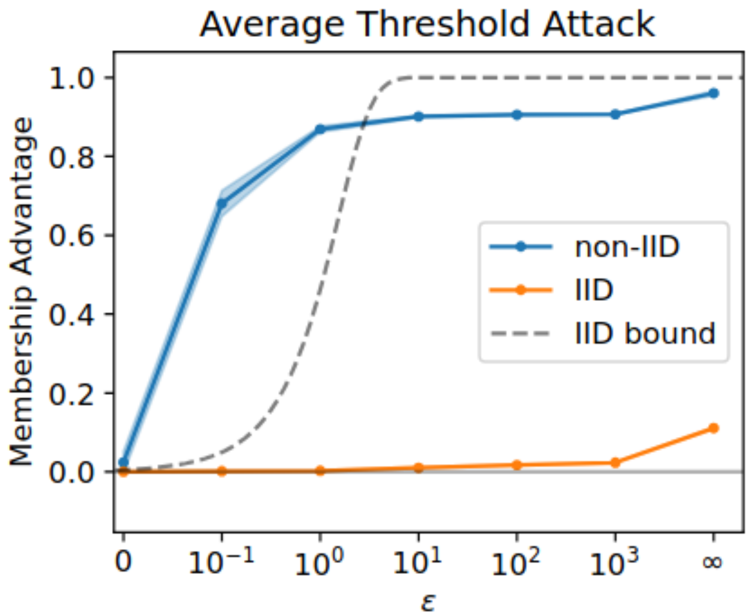
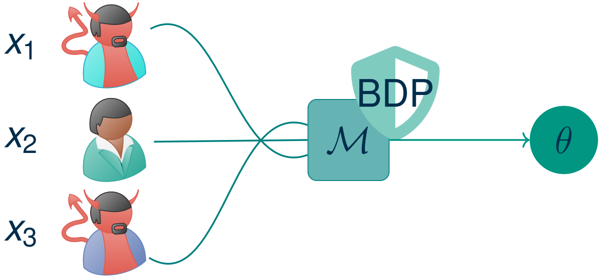


Figure: Humphries et al. 2023 MIA breaks DP guarantees.

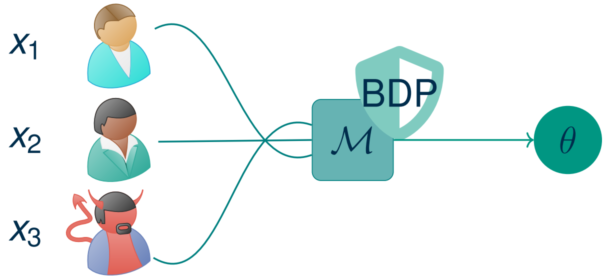
New enhanced notion: Bayesian Differential Privacy

Proposed Solution: Bayesian Differential Privacy



$K = \{1, 3\}$, $i = 2$

Proposed Solution: Bayesian Differential Privacy

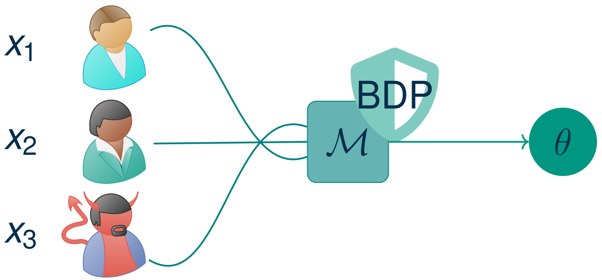


$K = \{3\}, i = 2$

Proposed Solution: Bayesian Differential Privacy

Bayesian DP leakage (Yang et al. 2017)

$$\mathrm{BDPL}_{(K,i)} = \sup_{x_i, x_i', \mathbf{x}_K, S} \ln \frac{\Pr_{\mathcal{M}}[Y \in S \mid \mathbf{X}_K = \mathbf{x}_K, X_i = x_i]}{\Pr_{\mathcal{M}}[Y \in S \mid \mathbf{X}_K = \mathbf{x}_K, X_i = x_i']}, \text{ then } \varepsilon = \sup_{K,i} \mathrm{BDPL}_{(K,i)}.$$



Proposed Solution: Bayesian Differential Privacy

Bayesian DP leakage (Yang et al. 2017)

$$\mathrm{BDPL}_{(K,i)} = \sup_{x_i, x_i', \mathbf{x}_K, S} \ln \frac{\Pr_{\mathcal{M}}[Y \in S \mid \mathbf{X}_K = \mathbf{x}_K, X_i = x_i]}{\Pr_{\mathcal{M}}[Y \in S \mid \mathbf{X}_K = \mathbf{x}_K, X_i = x_i']}, \text{ then } \varepsilon = \sup_{K,i} \mathrm{BDPL}_{(K,i)}.$$

We check our MIA again, with $H_0 : x_n \in D$, $H_1 : y_n \in D$ $(K = \varnothing)$

Proposed Solution: Bayesian Differential Privacy

Bayesian DP leakage (Yang et al. 2017)

$$\mathrm{BDPL}_{(K,i)} = \sup_{x_i, x_i', \mathbf{x}_K, S} \ln \frac{\Pr_{\mathcal{M}}[Y \in S \mid \mathbf{X}_K = \mathbf{x}_K, X_i = x_i]}{\Pr_{\mathcal{M}}[Y \in S \mid \mathbf{X}_K = \mathbf{x}_K, X_i = x_i']}, \text{ then } \varepsilon = \sup_{K,i} \mathrm{BDPL}_{(K,i)}.$$

We check our MIA again, with $H_0 : x_n \in D$, $H_1 : y_n \in D$ $(K = \varnothing)$

Type I error:

$$\alpha = \Pr_{A \circ \mathcal{M}}(y_n \mid x_n)$$

Type II error:

$$\beta = 1 - \Pr_{A \circ \mathcal{M}}(y_n \mid y_n)$$

Proposed Solution: Bayesian Differential Privacy

Bayesian DP leakage (Yang et al. 2017)

$$\mathrm{BDPL}_{(K,i)} = \sup_{x_i, x_i', \mathbf{x}_K, S} \ln \frac{\Pr_{\mathcal{M}}[Y \in S \mid \mathbf{X}_K = \mathbf{x}_K, X_i = x_i]}{\Pr_{\mathcal{M}}[Y \in S \mid \mathbf{X}_K = \mathbf{x}_K, X_i = x_i']}, \text{ then } \varepsilon = \sup_{K,i} \mathrm{BDPL}_{(K,i)}.$$

We check our MIA again, with $H_0 : x_n \in D$, $H_1 : y_n \in D$ $(K = \varnothing)$

Type I error:

$$\alpha = \Pr_{A \circ \mathcal{M}}(y_n \mid x_n)$$

Type II error:

$$\beta = 1 - \Pr_{A \circ \mathcal{M}}(y_n \mid y_n)$$

$A \circ \mathcal{M}$ is ε-**B**DP $\Longrightarrow$

$$1 - \beta \leq e^{\varepsilon} \alpha$$

$$\alpha \leq e^{\varepsilon}(1 - \beta)$$

Proposed Solution: Bayesian Differential Privacy

Bayesian DP leakage (Yang et al. 2017)

$$\mathrm{BDPL}_{(K,i)} = \sup_{x_i, x_i', \mathbf{x}_K, S} \ln \frac{\Pr_{\mathcal{M}}[Y \in S \mid \mathbf{X}_K = \mathbf{x}_K, X_i = x_i]}{\Pr_{\mathcal{M}}[Y \in S \mid \mathbf{X}_K = \mathbf{x}_K, X_i = x_i']}, \text{ then } \varepsilon = \sup_{K,i} \mathrm{BDPL}_{(K,i)}.$$

We check our MIA again, with $H_0 : x_n \in D$, $H_1 : y_n \in D$ $(K = \varnothing)$

Type I error:

$$\alpha = \Pr_{A \circ \mathcal{M}} (y_n \mid x_n)$$

Type II error:

$$\beta = 1 - \Pr_{A \circ \mathcal{M}} (y_n \mid y_n)$$

$A \circ \mathcal{M}$ is ε-**B**DP $\implies$ $1 - \beta \leq e^{\varepsilon}\alpha$, $\alpha \leq e^{\varepsilon}(1 - \beta)$ $\implies$ $\beta \geq \max\{1 - e^{\varepsilon}\alpha, e^{\varepsilon}(1 - \alpha)\}$

Proposed Solution: Bayesian Differential Privacy

Bayesian DP leakage (Yang et al. 2017)

$$\mathrm{BDPL}_{(K,i)} = \sup_{x_i, x_i', \mathbf{x}_K, S} \ln \frac{\Pr_{\mathcal{M}}[Y \in S \mid \mathbf{X}_K = \mathbf{x}_K, X_i = x_i]}{\Pr_{\mathcal{M}}[Y \in S \mid \mathbf{X}_K = \mathbf{x}_K, X_i = x_i']}, \text{ then } \varepsilon = \sup_{K,i} \mathrm{BDPL}_{(K,i)}.$$

Privacy

- ✔ Effective measure and resistance to correlation-based attacks.

Proposed Solution: Bayesian Differential Privacy

Bayesian DP leakage (Yang et al. 2017)

$$\mathrm{BDPL}_{(K,i)} = \sup_{x_i, x_i', \mathbf{x}_K, S} \ln \frac{\Pr_{\mathcal{M}}[Y \in S \mid \mathbf{X}_K = \mathbf{x}_K, X_i = x_i]}{\Pr_{\mathcal{M}}[Y \in S \mid \mathbf{X}_K = \mathbf{x}_K, X_i = x_i']}, \text{ then } \varepsilon = \sup_{K,i} \mathrm{BDPL}_{(K,i)}.$$

Privacy

- ✓ Effective measure and resistance to correlation-based attacks.
- ✓ Good properties: post-processing & composition.
  - While other correlation-aware notions (General Pufferfish framework) don't!

Proposed Solution: Bayesian Differential Privacy

Bayesian DP leakage (Yang et al. 2017)

$$\mathrm{BDPL}_{(K,i)} = \sup_{x_i, x_i', \mathbf{x}_K, S} \ln \frac{\Pr_{\mathcal{M}}[Y \in S \mid \mathbf{X}_K = \mathbf{x}_K, X_i = x_i]}{\Pr_{\mathcal{M}}[Y \in S \mid \mathbf{X}_K = \mathbf{x}_K, X_i = x_i']}, \text{ then } \varepsilon = \sup_{K,i} \mathrm{BDPL}_{(K,i)}.$$

Privacy

- ✔ Effective measure and resistance to correlation-based attacks.
- ✔ Good properties: post-processing & composition.
  - While other correlation-aware notions (General Pufferfish framework) don't!

Utility

- ✘ Poor utility (methods based on group privacy).

Proposed Solution: Bayesian Differential Privacy

Bayesian DP leakage (Yang et al. 2017)

$$\mathrm{BDPL}_{(K,i)} = \sup_{x_i, x_i', \mathbf{x}_K, S} \ln \frac{\Pr_{\mathcal{M}}[Y \in S \mid \mathbf{X}_K = \mathbf{x}_K, X_i = x_i]}{\Pr_{\mathcal{M}}[Y \in S \mid \mathbf{X}_K = \mathbf{x}_K, X_i = x_i']}, \text{ then } \varepsilon = \sup_{K,i} \mathrm{BDPL}_{(K,i)}.$$

Privacy

- ✔ Effective measure and resistance to correlation-based attacks.
- ✔ Good properties: post-processing & composition.
  - While other correlation-aware notions (General Pufferfish framework) don't!

Utility

- ✘ Poor utility (methods based on group privacy).
- ✘ Computationally intractable methods (computing the Wasserstein distance).

Proposed Solution: Bayesian Differential Privacy

Bayesian DP leakage (Yang et al. 2017)

$$\mathrm{BDPL}_{(K,i)} = \sup_{x_i, x_i', \mathbf{x}_K, S} \ln \frac{\Pr_{\mathcal{M}}[Y \in S \mid \mathbf{X}_K = \mathbf{x}_K, X_i = x_i]}{\Pr_{\mathcal{M}}[Y \in S \mid \mathbf{X}_K = \mathbf{x}_K, X_i = x_i']}, \text{ then } \varepsilon = \sup_{K,i} \mathrm{BDPL}_{(K,i)}.$$

Privacy

- ✔ Effective measure and resistance to correlation-based attacks.
- ✔ Good properties: post-processing & composition.
  - While other correlation-aware notions (General Pufferfish framework) don't!

Utility

- ✘ Poor utility (methods based on group privacy).
- ✘ Computationally intractable methods (computing the Wasserstein distance).
- ✘ Limited applicability (lazy, binary, stationary Markov chains).

Our Research Question

Can we reduce utility loss while still retaining the privacy guarantees of BDP?

Our methodology: Understanding how DP leakage relates to BDP leakage:

$$\varepsilon\text{-DP} \Rightarrow ??\text{-BDP}.$$

Against Arbitrary Correlations It Is Impossible

Kifer and Machanavajjhala 2014:

Pufferfish (including BDP) & arbitrary correlation $\Rightarrow$ Free-lunch Privacy $\Rightarrow$ No utility.

Against Arbitrary Correlations It Is Impossible

Kifer and Machanavajjhala 2014:

Pufferfish (including BDP) & arbitrary correlation $\Rightarrow$ Free-lunch Privacy $\Rightarrow$ No utility.

We express this in term of (α, β)**-accuracy** for any numerical target query f:

(α, β)-accuracy

$$\Pr(|f(D) - \mathcal{M}(D)| \geq \alpha) \leq \beta$$

$1 - \beta =$ confidence

$\alpha =$ error interval

Against Arbitrary Correlations It Is Impossible

Kifer and Machanavajjhala 2014:

Pufferfish (including BDP) & arbitrary correlation $\Rightarrow$ Free-lunch Privacy $\Rightarrow$ No utility.

We express this in term of (α, β)**-accuracy** for any numerical target query f:

(α, β)-accuracy

$$\Pr(|f(D) - \mathcal{M}(D)| \geq \alpha) \leq \beta$$

$1 - \beta =$ confidence

$\alpha =$ error interval

Our result (informal):

$$\beta < \frac{1}{e^{\varepsilon}+1} \Rightarrow \alpha > \frac{1}{2}\mathrm{Range}(f).$$

Against Arbitrary Correlations It Is Impossible

Kifer and Machanavajjhala 2014:

Pufferfish (including BDP)

& $\Rightarrow$ Free-lunch Privacy $\Rightarrow$ No utility.

arbitrary correlation

We express this in term of (α, β)**-accuracy** for any numerical target query f:

(α, β)-accuracy

$$\Pr(|f(D) - \mathcal{M}(D)| \geq \alpha) \leq \beta$$

$1 - \beta =$ confidence

$\alpha =$ error interval

Our result (informal):

$$\beta < \frac{1}{e^{\varepsilon}+1} \Rightarrow \alpha > \frac{1}{2}\mathrm{Range}(f).$$

Against Arbitrary Correlations It Is Impossible

Kifer and Machanavajjhala 2014:

Pufferfish (including BDP) & arbitrary correlation $\Rightarrow$ Free-lunch Privacy $\Rightarrow$ No utility.

We express this in term of (α, β)**-accuracy** for any numerical target query f:

(α, β)-accuracy

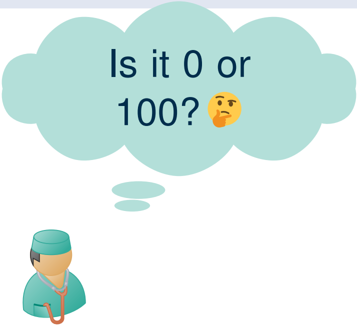
$$\Pr(|f(D) - \mathcal{M}(D)| \geq \alpha) \leq \beta$$

$1 - \beta =$ confidence

$\alpha =$ error interval

Our result (informal):

$\beta < \frac{1}{e^{\varepsilon}+1} \Rightarrow \alpha > \frac{1}{2}\mathrm{Range}(f).$



Few Correlated Records, Same Disaster

Our result (informal)

Privacy decreases linearly proportional to number of correlated records:

$$\varepsilon\text{-DP} \Rightarrow m\varepsilon\text{-BDP}$$

How does it impact utility?

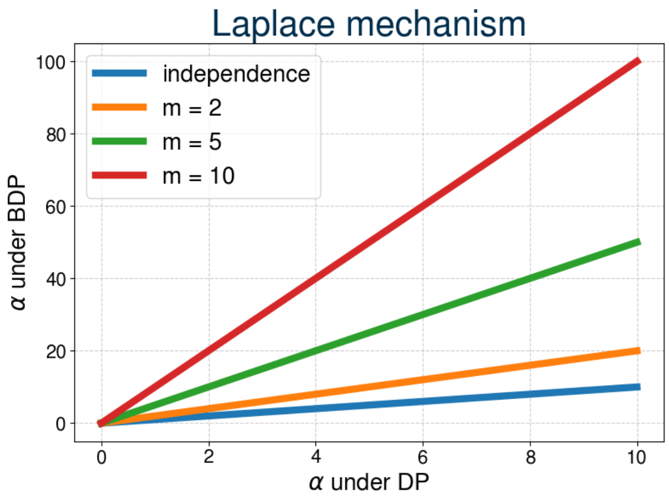


Figure: For the same confidence level, the upper bound on the query error α increases sharply.

Few Correlated Records, Same Disaster

Our result (informal)

Privacy decreases linearly proportional to number of correlated records:

$$\varepsilon\text{-DP} \Rightarrow m\varepsilon\text{-BDP}$$

This result is tight! Even if $\rho \to 0$.

How does it impact utility?

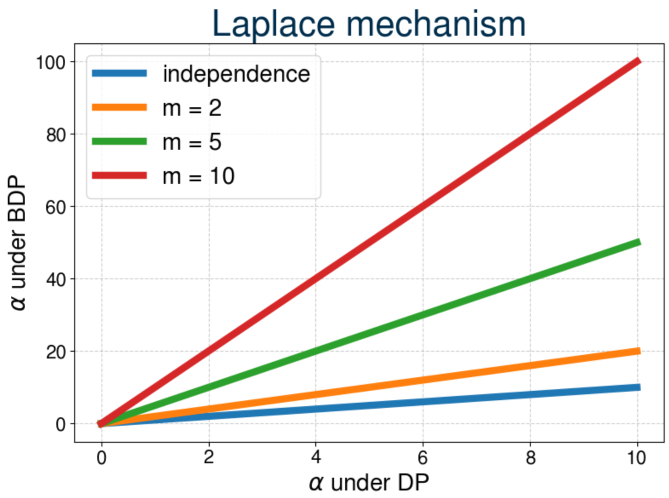


Figure: For the same confidence level, the upper bound on the query error α increases sharply.

Few Correlated Records, Same Disaster

Our result (informal)

Privacy decreases linearly proportional to number of correlated records:

$$\varepsilon\text{-DP} \Rightarrow m\varepsilon\text{-BDP}$$

This result is tight! Even if $\rho \to 0$.

Conclusion:

We need to target specific correlation models π to obtain utility

How does it impact utility?

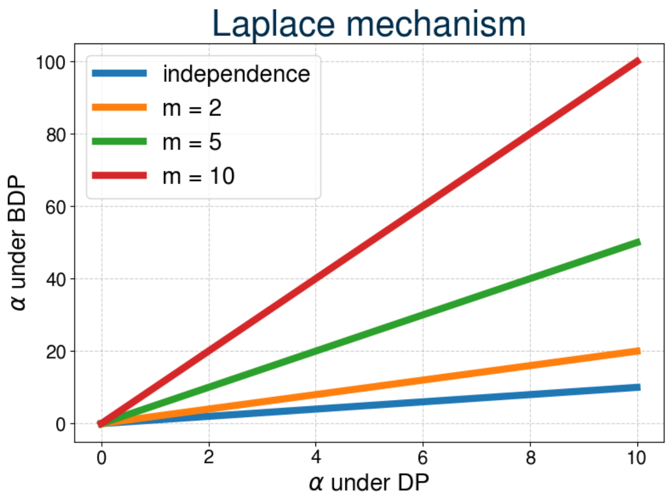


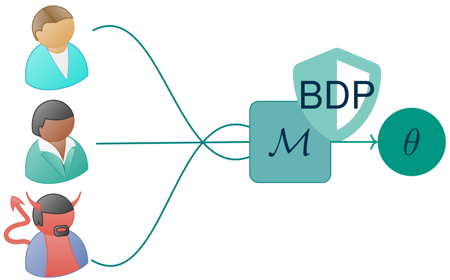
Figure: For the same confidence level, the upper bound on the query error α increases sharply.

New Strategy

Adjust the noise of DP mechanisms to obtain useful BDP mechanisms targeting specific priors π.

Assumptions:

- Global setting: All data is collected by a trusted data curator that applies the mechanism.
- The attacker does not have more knowledge about π than the data curator.

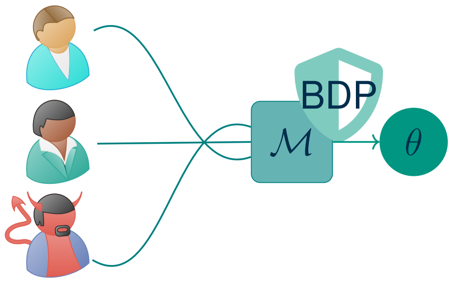


New Strategy

Adjust the noise of DP mechanisms to obtain useful BDP mechanisms targeting specific priors π.

Assumptions:

- Global setting: All data is collected by a trusted data curator that applies the mechanism.
- The attacker does not have more knowledge about π than the data curator.



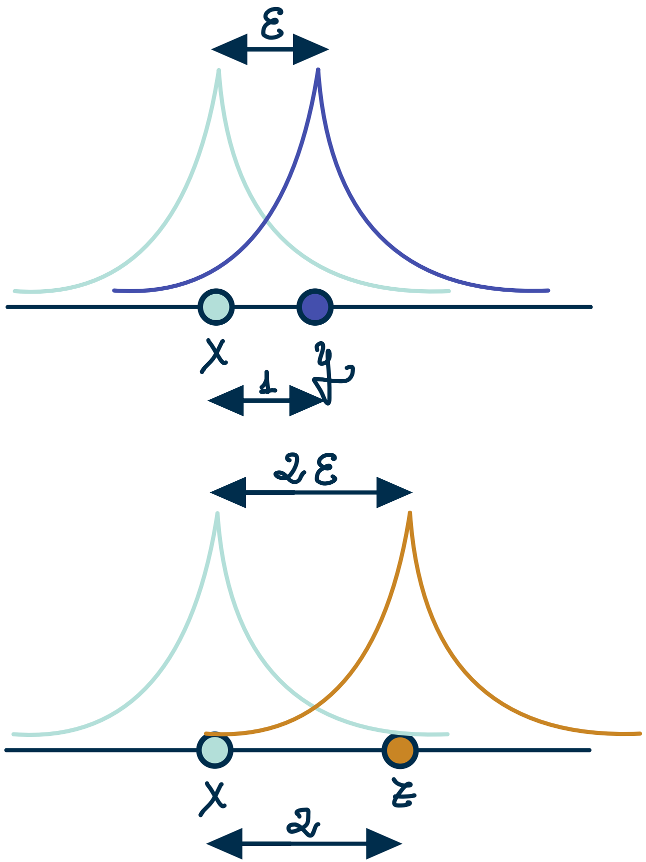
Multivariate Gaussian

Markov Chains

Multivariate Gaussian Correlation

Main Result (Informal)

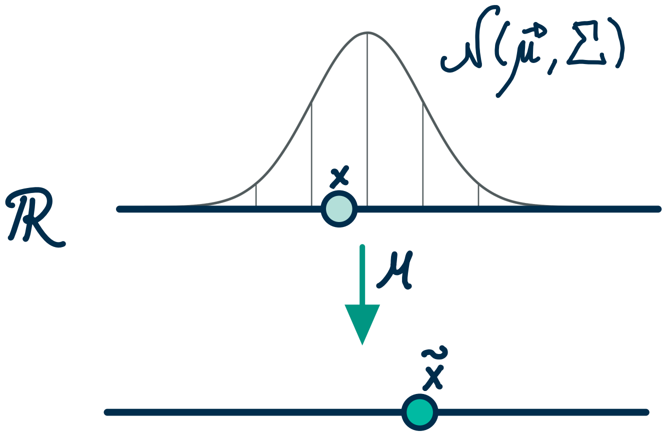
- Let $\mathcal{M}$ be an $\varepsilon\ell_1$-private mechanism,



Multivariate Gaussian Correlation

Main Result (Informal)

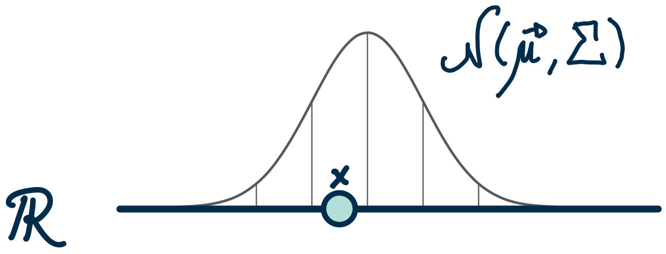
- Let $\mathcal{M}$ be an $\varepsilon\ell_1$-private mechanism,
- input data drawn from a multivariate Gaussian distribution



Multivariate Gaussian Correlation

Main Result (Informal)

- Let $\mathcal{M}$ be an $\varepsilon\ell_1$-private mechanism,
- input data drawn from a multivariate Gaussian distribution
- $\rho(m-2) < 1$ is the maximum correlation coefficient.



$$\vec{\mu} = \begin{pmatrix} \mu_{x_1} \\ \vdots \\ \mu_{x_n} \end{pmatrix} \qquad \Sigma = \begin{pmatrix} \sigma^2 & & \\ & \ddots & \\ & & \sigma^2 \end{pmatrix}$$

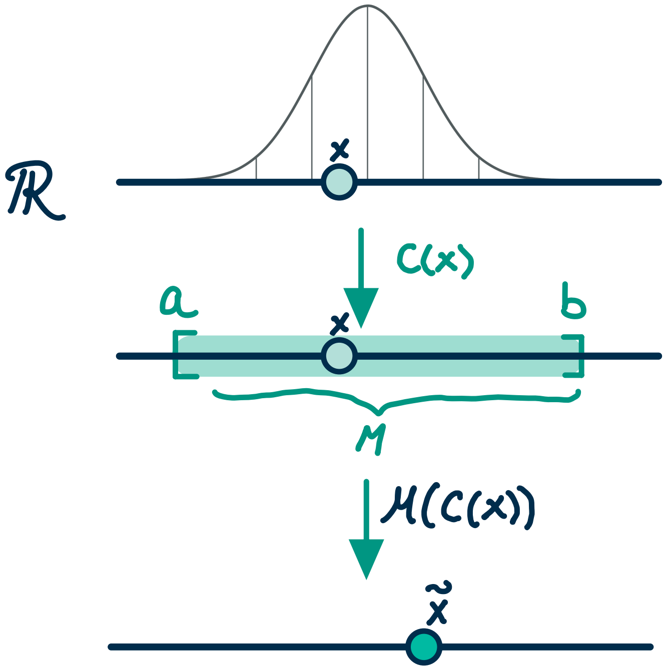
$$|Cov(X_i, X_j)| \leq \rho\sigma^2$$

Multivariate Gaussian Correlation

Main Result (Informal)

- Let $\mathcal{M}$ be an $\varepsilon\ell_1$-private mechanism,
- input data drawn from a multivariate Gaussian distribution
- $\rho(m-2) < 1$ is the maximum correlation coefficient.

Then, using clipping as preprocessing step, $c_I(D)_i = \max(a, \min(b, D_i))$, we obtain $\mathcal{M}_I$ satisfying



Multivariate Gaussian Correlation

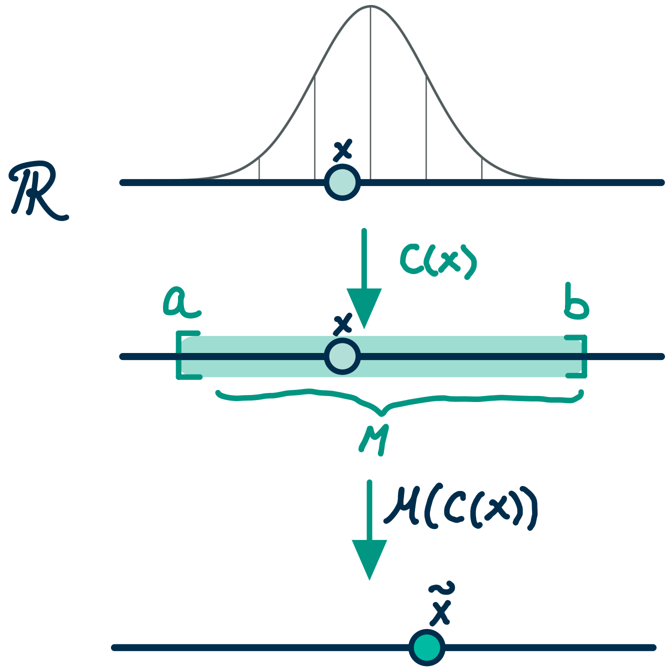
Main Result (Informal)

- Let $\mathcal{M}$ be an $\varepsilon\ell_1$-private mechanism,
- input data drawn from a multivariate Gaussian distribution
- $\rho(m-2) < 1$ is the maximum correlation coefficient.

Then, using clipping as preprocessing step, $c_I(D)_i = \max(a, \min(b, D_i))$, we obtain $\mathcal{M}_I$ satisfying

$$\mathrm{BDPL}(\mathcal{M}_I) \leq \left(\frac{m^2}{4(\frac{1}{\rho} - m + 2)} + 1 \right) M\varepsilon.$$

where M is the diameter of the interval $I = [a, b]$



Multivariate Gaussian Correlation (Impact on Real Databases)

- **Use-case:** Sum queries with Laplace mechanism. $\theta = f(D) + Z$ with $Z \sim \mathrm{Lap}(\mathrm{b})$.

Multivariate Gaussian Correlation (Impact on Real Databases)

- **Use-case:** Sum queries with Laplace mechanism. $\theta = f(D) + Z$ with $Z \sim \mathrm{Lap}(\mathrm{b})$.
- **Strategy:** We calibrate b to obtain BDP using our theorem.

Multivariate Gaussian Correlation (Impact on Real Databases)

- **Use-case:** Sum queries with Laplace mechanism. $\theta = f(D) + Z$ with $Z \sim \mathrm{Lap(b)}$.
- **Strategy:** We calibrate b to obtain BDP using our theorem.
- **Utility metric:** We set $\beta = 0.05$ (i.e., 95% confidence) and measure (α, β)-accuracy, both theoretically $(-)$ and empirically $(\times)$.

Multivariate Gaussian Correlation (Impact on Real Databases)

- **Use-case:** Sum queries with Laplace mechanism. $\theta = f(D) + Z$ with $Z \sim \mathrm{Lap(b)}$.
- **Strategy:** We calibrate b to obtain BDP using our theorem.
- **Utility metric:** We set $\beta = 0.05$ (i.e., 95% confidence) and measure (α, β)-accuracy, both theoretically ($-$) and empirically ($\times$).

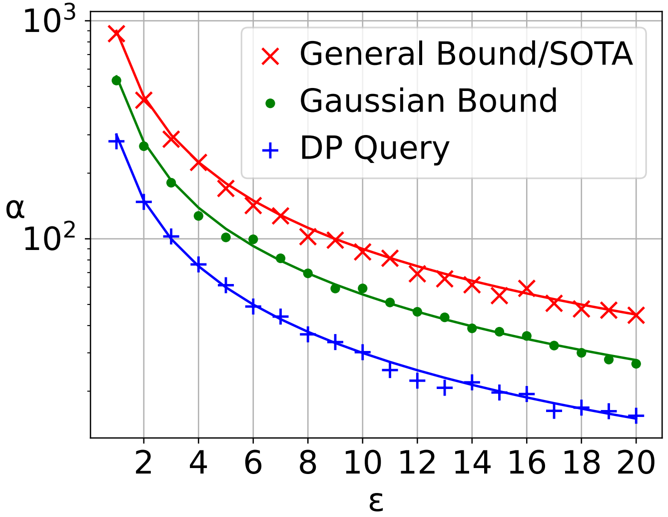


Figure: Galton, $n = 897$ $m = 3$

Key takeway:

Substantial utility gains compared to the general bound!

- More experiments with different real and synthetic datasets in our paper show similar results.

Markov Chain Correlation Model

Main result (Informal)

- Let $\mathcal{M}$ be an ε-DP mechanism,
- input data sampled form Markov chain with transition matrix $P \in \mathbb{R}^{s \times s}$ and initial distribution $w \in \mathbb{R}^s$ with the following properties:

$$\text{(H1) For all } x, y \in \mathcal{S} \text{ we have } P_{x,y} > 0 \text{ and,} \qquad \text{(H2) } wP = w.$$

Then, $\mathcal{M}$ is an $(\varepsilon + 4 \ln \gamma)$-BDP mechanism where $\gamma = \frac{\max_{x,y \in \mathcal{S}} P_{xy}}{\min_{x,y \in \mathcal{S}} P_{xy}}$.

Markov Chain Correlation Model

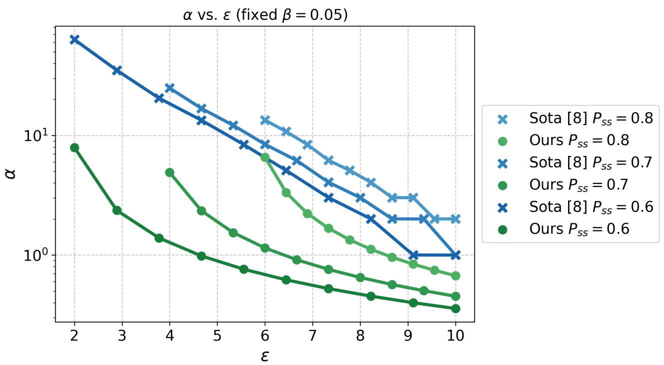
Main result (Informal)

- Let $\mathcal{M}$ be an ε-DP mechanism,
- input data sampled form Markov chain with transition matrix $P \in \mathbb{R}^{s \times s}$ and initial distribution $w \in \mathbb{R}^s$ with the following properties:

$$\text{(H1) For all } x, y \in \mathcal{S} \text{ we have } P_{x,y} > 0 \text{ and,} \qquad \text{(H2) } wP = w.$$

Then, $\mathcal{M}$ is an $(\varepsilon + 4\ln\gamma)$-BDP mechanism where $\gamma = \frac{\max_{x,y\in\mathcal{S}} P_{xy}}{\min_{x,y\in\mathcal{S}} P_{xy}}$.

Previous mechanism	Ours
$P_{xy} > 0$	$P_{xy} > 0$
stationary	stationary
lazy	
binary	
symmetric	
$\varepsilon' > 0$	$\varepsilon' > 4\ln(\gamma)$



Markov Chain Correlation Model (Impact on Real Databases)

- **Use-case:** Counting queries with Laplace mechanism. $\theta = f(D) + Z$ with $Z \sim \mathrm{Lap}(\mathrm{b})$.

Markov Chain Correlation Model (Impact on Real Databases)

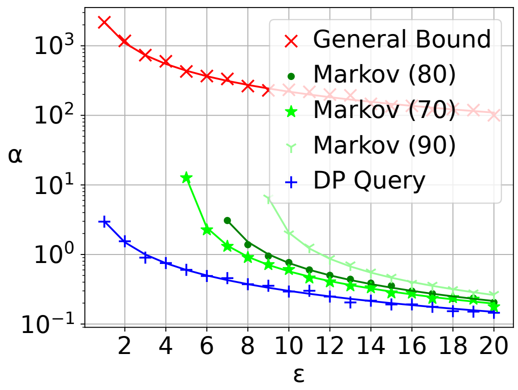
- **Use-case:** Counting queries with Laplace mechanism. $\theta = f(D) + Z$ with $Z \sim \mathrm{Lap(b)}$.
- **Strategy:** We calibrate b to obtain BDP using our theorem.

Markov Chain Correlation Model (Impact on Real Databases)

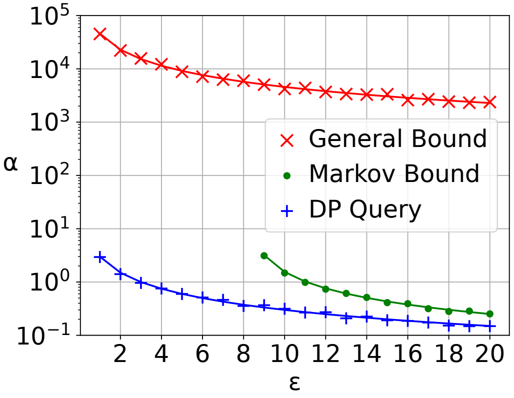
- **Use-case:** Counting queries with Laplace mechanism. $\theta = f(D) + Z$ with $Z \sim \mathrm{Lap(b)}$.
- **Strategy:** We calibrate b to obtain BDP using our theorem.
- **Utility metric:** We set $\beta = 0.05$ (i.e., 95% confidence) and measure $-(\alpha, \beta)$-accuracy, $\times$ upper bound of a $(1 - \beta)$ confidence interval for the absolute query error.

Markov Chain Correlation Model (Impact on Real Databases)

- **Use-case:** Counting queries with Laplace mechanism. $\theta = f(D) + Z$ with $Z \sim \mathrm{Lap(b)}$.
- **Strategy:** We calibrate b to obtain BDP using our theorem.
- **Utility metric:** We set $\beta = 0.05$ (i.e., 95% confidence) and measure $-(\alpha, \beta)$-accuracy, $\times$ upper bound of a $(1 - \beta)$ confidence interval for the absolute query error.



(g) Electricity, $n = 731$.



(h) Activity, $n = 17\,568$.

Key takeway:

- **Substantial utility gains** compared to the general bound!
- Markov bound independent of n $\Rightarrow$ **huge improvement for large datasets.**

Conclusion

- ✓ We provide a **feasible method** to generate a **BDP mechanism** by **recalibrating** existing DP methods, tailored to **Gaussian** and **Markov** models.

Conclusion

- ✓ We provide a **feasible method** to generate a **BDP mechanism** by **recalibrating** existing DP methods, tailored to **Gaussian** and **Markov** models.
- ✓ We offer **significantly better utility than prior results**.

Conclusion

- ✓ We provide a **feasible method** to generate a **BDP mechanism** by **recalibrating** existing DP methods, tailored to **Gaussian** and **Markov** models.
- ✓ We offer **significantly better utility than prior results**.

Key takeway:

BDP becomes usable when correlations are structured.

Conclusion

- ✓ We provide a **feasible method** to generate a **BDP mechanism** by **recalibrating** existing DP methods, tailored to **Gaussian** and **Markov** models.
- ✓ We offer **significantly better utility than prior results**.

Key takeway:

BDP becomes usable when correlations are structured.

Future Work:

- Other distributions **?**
- Can we build methods from scratch instead or recycling **?**
- What if we calibrate directly to the attack advantage **?**

Conclusion

- ✓ We provide a **feasible method** to generate a **BDP mechanism** by **recalibrating** existing DP methods, tailored to **Gaussian** and **Markov** models.
- ✓ We offer **significantly better utility than prior results**.

Key takeway:

BDP becomes usable when correlations are structured.

Future Work:

- Other distributions **?**
- Can we build methods from scratch instead or recycling **?**
- What if we calibrate directly to the attack advantage **?**

Paper

Code

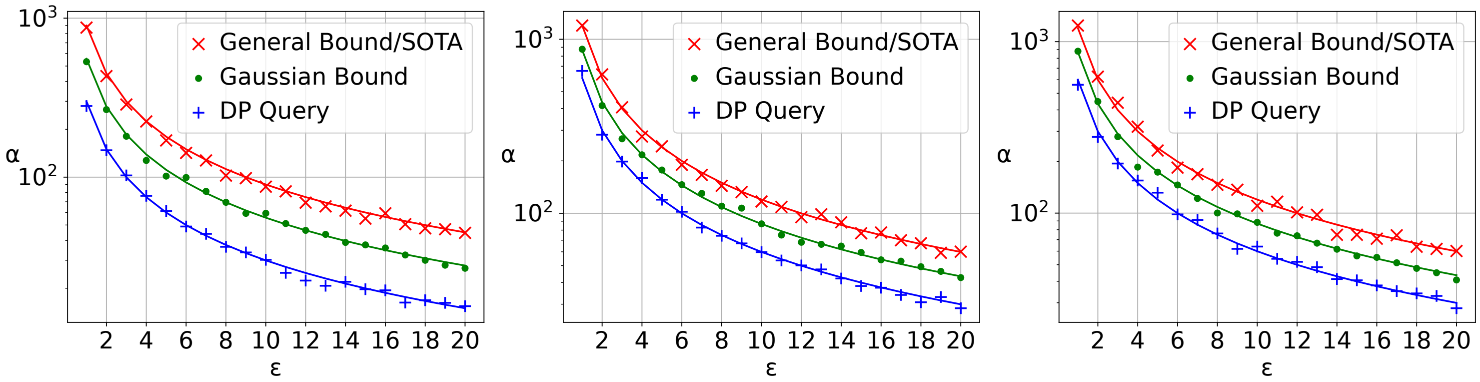
Backup Slides

Experiment Details

Database	n	m	Parameters	Sensitivity
Galton	897	3	$\rho = 0.275$	$\Delta q = 254cm$
FamilyIQ	868	2	$\rho = 0.4483$	$\Delta q = 120$
SyntheticIQ	20000	2	$\rho = 0.45$	$\Delta q = 120$
Activity	17568	n	$\gamma = 7.54$	$\Delta q = 1$
Activity Single Day	288	n	$\gamma = 7.54$	$\Delta q = 1$
Electricity	731	n	70 kWh, $\gamma = 3.29$ 80 kWh, $\gamma = 4.49$ 90 kWh, $\gamma = 8.43$	$\Delta q = 1$

Table: Data description. m is the max number of correlated records and n the total amount.

Multivariate Gaussian More Results



(i) Galton, $n = 897$ $m = 3$ (j) FamilyIQ, $n = 868$, $m = 2$. (k) SyntheticIQ, $n = 20000$, $m = 2$.

Figure: Gaussian data results. Lines show theoretical error at $\beta = 5\%$ and markers indicate empirical 95% upper bounds.