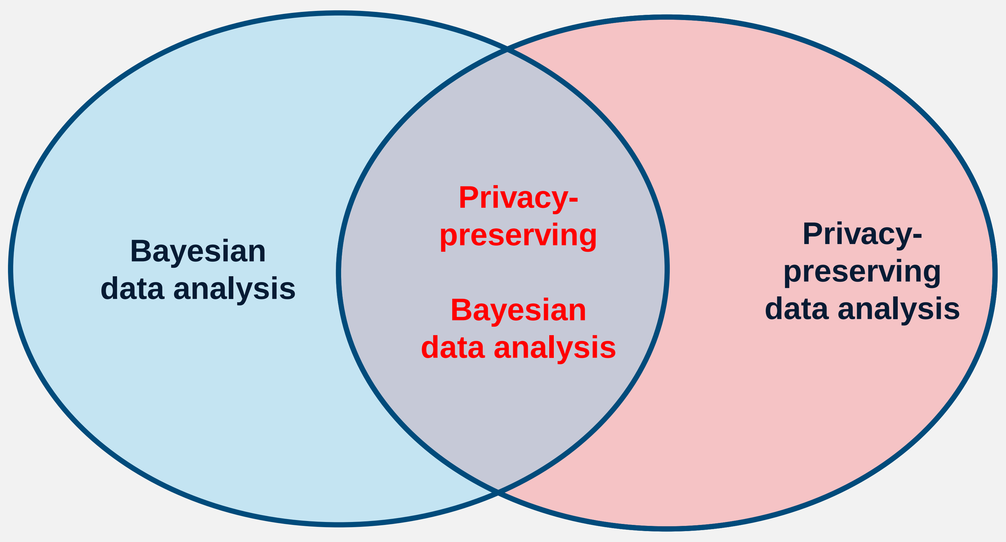
On the Theory and Practice of Privacy-Preserving Bayesian Data Analysis

James Foulds,* Joseph Geumlek,* Max Welling,+ Kamalika Chaudhuri*

*University of California, San Diego +University of Amsterdam

Overview

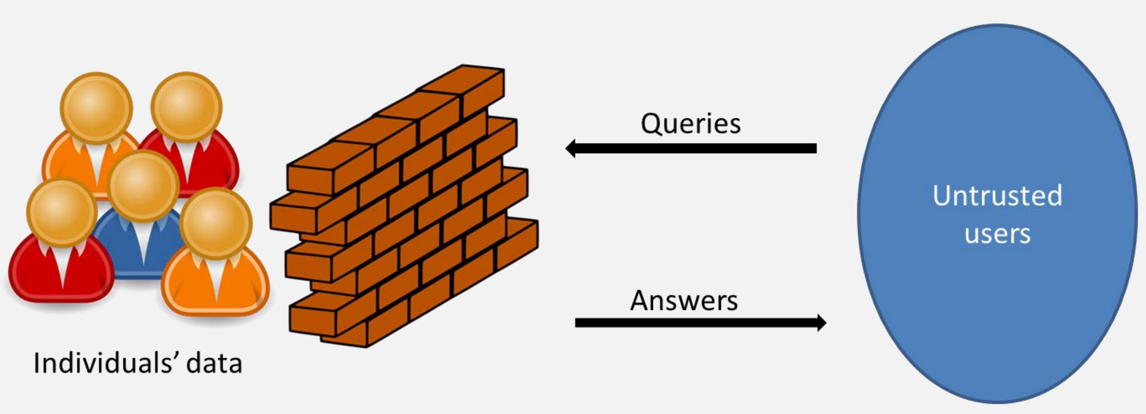


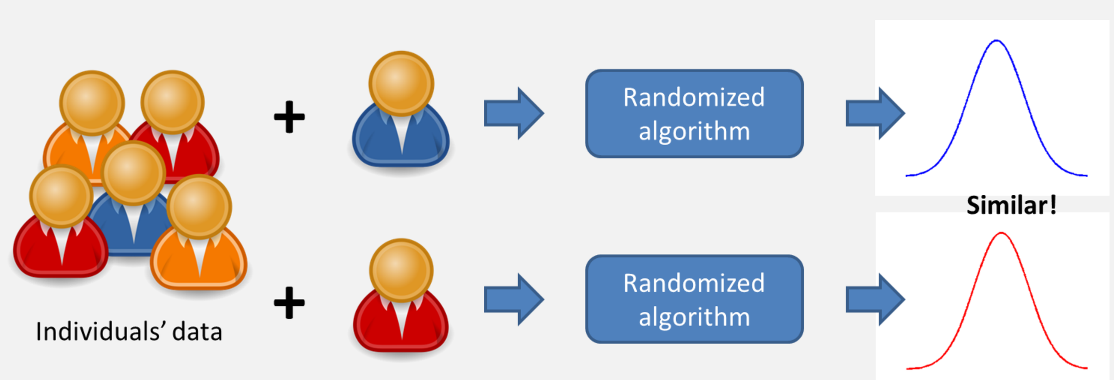
- It was recently shown that Bayesian posterior sampling can provide **privacy "for free"** (Dimitrakakis et al., 2014; Wang et al., 2015)
- This beautiful result has practical limitations: **data inefficiency, approximate inference**
- We develop a very **simple alternative** technique to resolve these limitations, and study it both **theoretically** and **empirically**

Motivation

- As individuals and consumers we benefit daily from ML systems trained on **our** data. The cost is our privacy
- Bayesian inference is widely used for modeling data where privacy is invaluable, including MOOCs, text data, recommendations,…
- Need privacy-preserving, Bayesian data analysis techniques
  - Balance utility and privacy
  - Trade-off should improve with more data

Background: Differential Privacy





Definition of differential privacy (Dwork et al, 2006): A randomized algorithm $\mathcal{M}(\mathbf{X})$ is ϵ-differentially private if

$$\frac{Pr(\mathcal{M}(\mathbf{X}) \in \mathcal{S})}{Pr(\mathcal{M}(\mathbf{X}') \in \mathcal{S})} \leq e^{\epsilon}$$

for all outcomes $\mathcal{S}$, and pairs of databases $\mathbf{X}$, $\mathbf{X}'$ differing in a single element.

Laplace and exponential mechanisms

Laplace mechanism
Add Laplace noise to results of query. Amount of noise depends on the **L1-sensititivity** of the query:

$$\triangle h = \max_{\mathbf{X},\mathbf{X}'} \|h(\mathbf{X}) - h(\mathbf{X}')\|_1$$

Exponential mechanism
Given a utility function, select outputs with high utility more often:

$$Pr(\mathcal{M}_E(\mathbf{X}, u, \epsilon) = \mathbf{r}) \propto \exp\left(\frac{u(\mathbf{X}, \mathbf{r})}{T}\right), \quad T = \frac{2\triangle u}{\epsilon}$$

Sensitivity: $\triangle u \triangleq \max_{r, (\mathbf{X}, \mathbf{X}')} \|u(\mathbf{X}, r) - u(\mathbf{X}', r)\|_1$

Temperature depends on sensitivity, epsilon

Posterior sampling via exponential mechanism
(Dimitrakakis et al., 2014; Wang et al., 2015)

Use utility function $u(\mathbf{X}, \theta) = \log Pr(\theta, \mathbf{X})$

$$\triangle \log Pr(\theta, \mathbf{X}) \triangleq \max_{\theta, (\mathbf{X}^{(1)}, \mathbf{X}^{(2)})} \| \log Pr(\theta, \mathbf{X}^{(1)}) - \log Pr(\theta, \mathbf{X}^{(2)})\|_1$$

Posterior sampling is $\epsilon = 2\triangle \log Pr(\theta, \mathbf{X})$-DP

For smaller ϵ, flatten posterior by increasing the temperature

Privacy for exponential family posteriors

For exponential family posteriors w\ conjugate priors

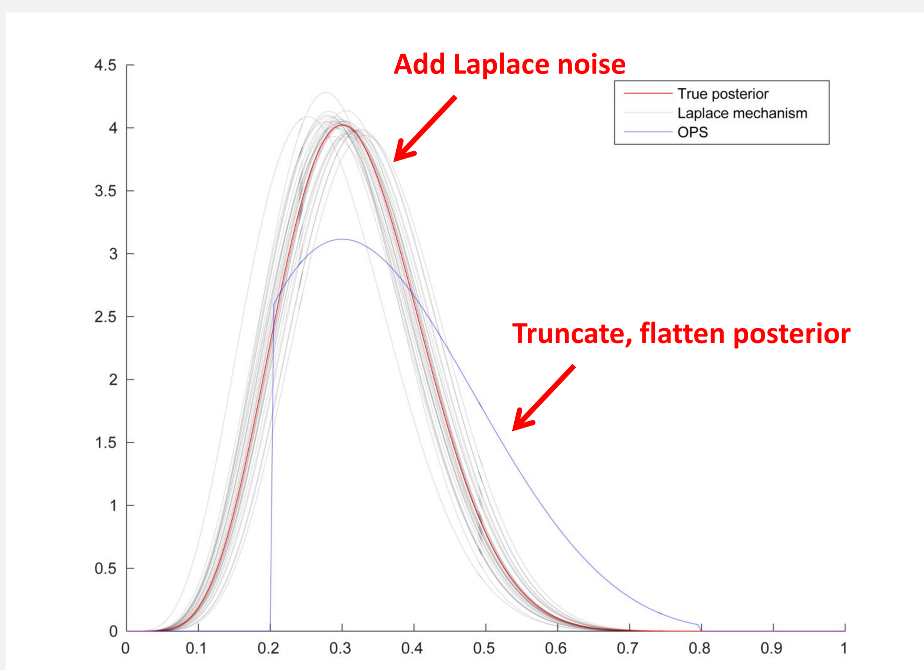
$$Pr(\theta | \mathbf{X}, \chi, \alpha) \propto g(\theta)^{N+\alpha} \exp\left(\theta^{\intercal}\left(\sum_{i=1}^{N} S(\mathbf{x}^{(i)}) + \alpha\chi\right)\right)$$

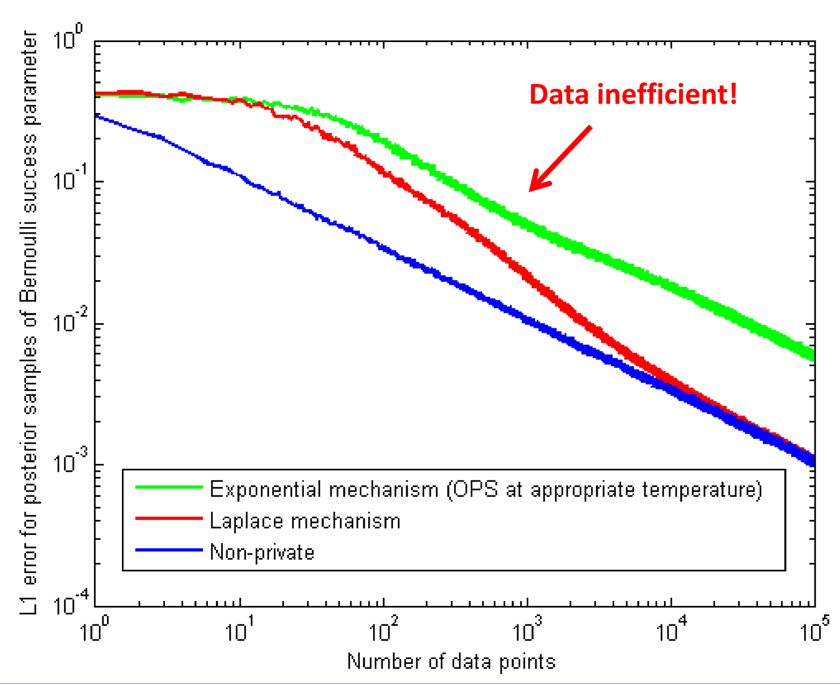
- We propose to use the **Laplace mechanism** to privatize likelihood model's **sufficient statistics**

Mechanism	Sufficient statistics $S(\mathbf{X})$ are:	Release	Sensitivity
Laplace	Noised additively	Statistics	$\sup_{\mathbf{x},\mathbf{x}'} \|S(\mathbf{x}') - S(\mathbf{x})\|_1$
Exponential	Rescaled multiplicatively	One sample	$\sup_{\mathbf{x},\mathbf{x}' \in \chi, \theta \in \Theta} \lvert\theta^{\intercal}\big(S(\mathbf{x}') - S(\mathbf{x})\big) + \log h(\mathbf{x}') - \log h(\mathbf{x})\rvert$

Worst case over parameters as well as data

Example: Beta-Bernoulli model





Asymptotic relative efficiency (ARE) results

- **ARE** = ratio between variance of estimator and optimal variance achieved by posterior mean in the limit, $\mathbb{I}^{-1}/N$
- **Exponential mechanism:** **ARE = 1 + T**
  Temperature T >= 1 *(Wang et al., 2015)*

Our results: under general conditions,

- **Laplace mech. (one sample):** **ARE = 2**
- **Laplace mech. (posterior mean):** **ARE = 1**

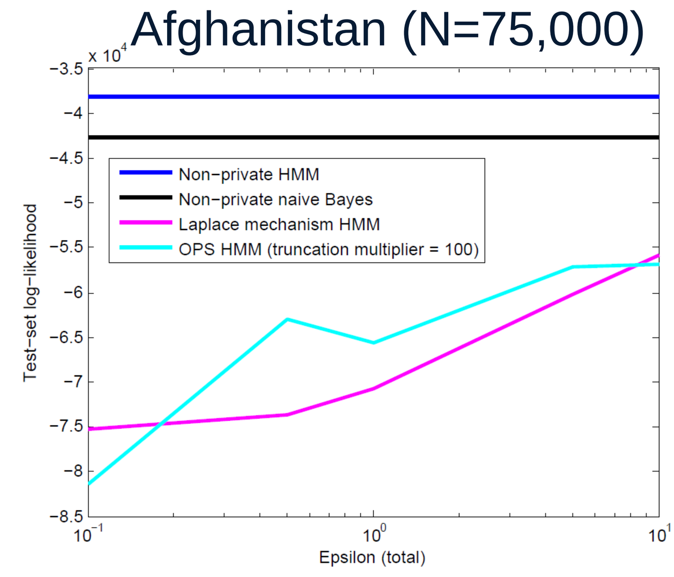
Private Gibbs sampling

- For exponential mechanism, privacy not guaranteed if MCMC sampler not converged
- *Interpret Gibbs update as exponential mechanism*
  - Privacy cost per Gibbs update at temperature T <= privacy cost of posterior sample
- Instead, can use **Laplace mechanism** to protect sufficient statistics needed for Gibbs updates, just **ONCE at beginning of sampling algorithm!**

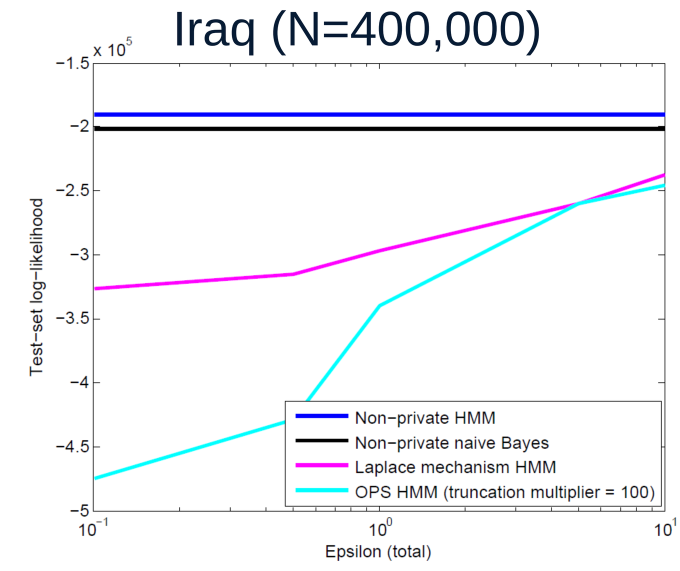
Case study: Wikileaks War Logs

- Privacy-preserving HMM on US military logs from Iraq/Afghanistan wars leaked by Wikileaks

Afghanistan (N=75,000)



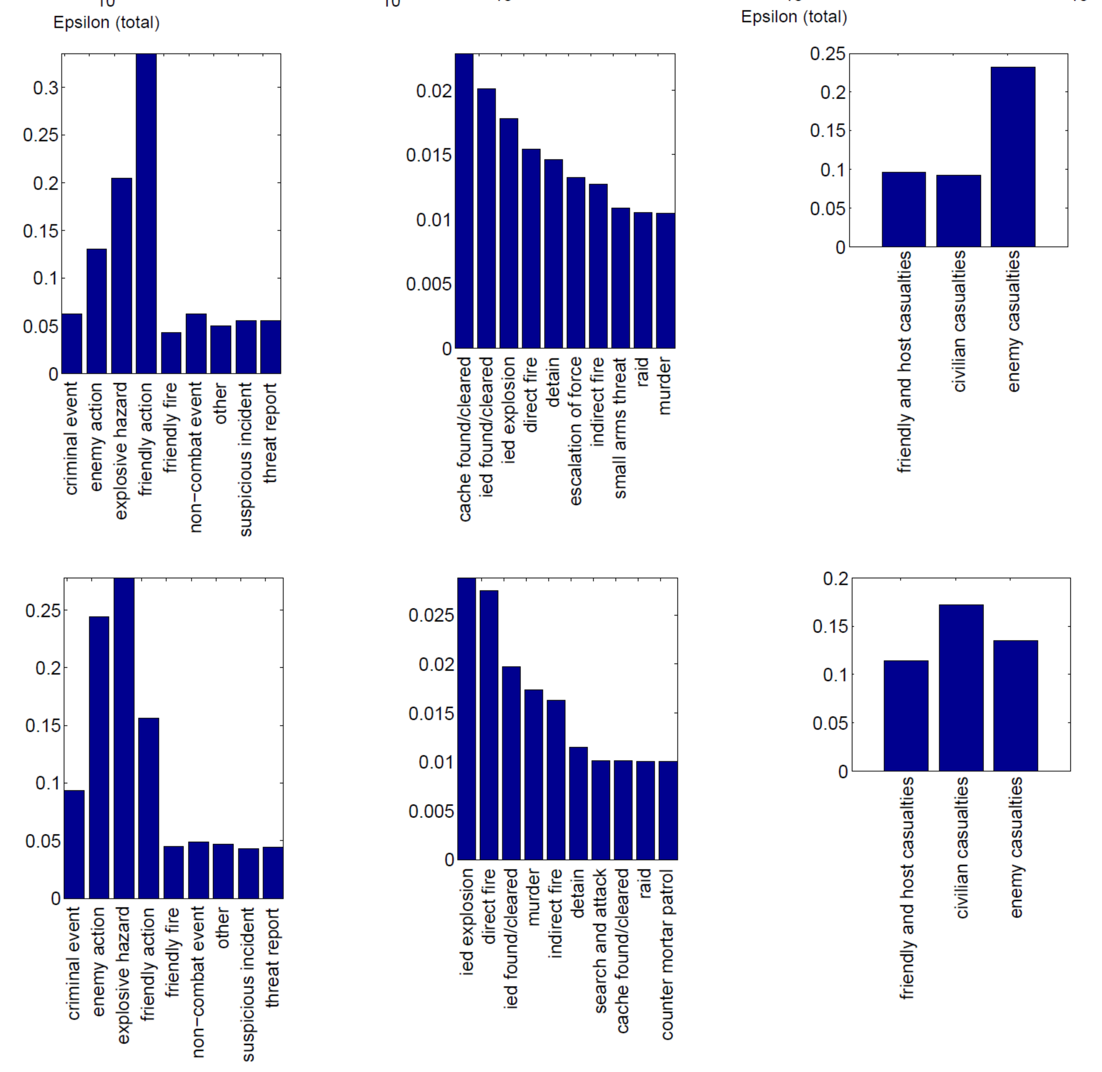
Iraq (N=400,000)



Iraq HMM (Laplace)

State 1 emissions

State 2 emissions



State assignments