Universal Exact Compression of Differentially Private Mechanisms

Yanxiao Liu[1] Wei-Ning Chen[2] Ayfer Özgür[2] Cheuk Ting Li[1]

[1]The Chinese University of Hong Kong [2]Stanford University

Introduction

Local differential privacy (DP).

A local randomizer $\mathcal{A} : \mathcal{X} \to \mathcal{Z}$ satisfies local DP if for any $x, x' \in \mathcal{X}$ and measurable set $\mathcal{S} \subset \mathcal{Z}$,

$$\Pr\{\mathcal{A}(x) \in \mathcal{S}\} \le e^{\varepsilon} \cdot \Pr\{\mathcal{A}(x') \in \mathcal{S}\} + \delta.$$

Compression with public randomness.

A local DP mechanism $\mathcal{A}$ can be compressed in b bits if

- $\mathcal{A}(x) \overset{\mathrm{d}}{=} \mathcal{M}(K(x,R), R)$ where R is public randomness;
- $K(x,R)$ can be encoded by a prefix-free code with expected length at most b;
- $(K(x,R), R)$ *jointly* satisfies local DP.

Prior works. It is known that every ε-local DP mechanism can be compressed in $O(\varepsilon)$ bits with small distortion [1, 2]; however, the compressed schemes are *approximate*.

Our contributions. We introduce Poisson private representation (PPR) that *exactly* simulates *any* local randomizer within $O(\varepsilon)$ bits while ensuring local DP.

Poisson Functional Representation

Let $(T_i)_i$ be a Poisson process with rate 1 (i.e., $T_1, T_2 - T_1, T_3 - T_2, \ldots \overset{\text{i.i.d.}}{\sim} \mathrm{Exp}(1)$), independent of $Z_i \overset{\text{i.i.d.}}{\sim} Q$. Then $(Z_i, T_i)_i$ is a Poisson process with intensity measure $Q \times \lambda_{[0,\infty)}$.

Fix any distribution P over $\mathcal{Z}$ that is absolutely continuous with respect to Q. Let

$$\tilde{T}_i := T_i \cdot \left(\frac{\mathrm{d}P}{\mathrm{d}Q}(Z_i)\right)^{-1}.$$

Then $(Z_i, \tilde{T}_i)$ is a Poisson process with intensity measure $P \times \lambda_{[0,\infty)}$ [3,4,5].

Theorem (Poisson functional representation [3]).

Let $Z = Z_K$ be with the smallest associated $\tilde{T}_K$, i.e., $K := \arg\min_i \tilde{T}_i$ and $Z := Z_K$. Then $Z \sim P$.

Poisson Private Representation (PPR)

Algorithm 1 (PPR).

Input: private data $x \in \mathcal{X}$, (ε, δ)-local DP mechanism $P(\cdot|x)$, reference distribution $Q(\cdot)$, compression parameter $\alpha > 1$.

(a) Generate public random variables

$$(Z_i)_{i=1,2,\ldots} \overset{\text{i.i.d.}}{\sim} Q(\cdot).$$

(b) The local user knows $(Z_i)_i$, x, $P(\cdot|x)$ and performs:

(1) Generate the Poisson process $(T_i)_i$ with rate 1.

(2) Computes $\tilde{T}_i \triangleq T_i \cdot \left(\frac{dP}{dQ}(Z_i)\right)^{-1}$.

(3) Generates $K \triangleq K(x; (Z_i, T_i)_i) \in \mathbb{Z}_+$ with

$$\Pr(K = k) = \frac{\tilde{T}_k^{-\alpha}}{\sum_{i=1}^{\infty} \tilde{T}_i^{-\alpha}}.$$

(4) Compresses and sends $K \in \mathbb{Z}_+$ (e.g., with Elias delta code).

(c) The server, which knows $(Z_i)_i, K$, outputs $Z = Z_K$.

Proposition 1 (Exactness).

The output Z of PPR follows $P(\cdot|x)$ exactly.

Theorem 2 (Compression size).

For PPR with $\alpha > 1$, message $K \in \mathbb{Z}_+$ satisfies

$$\mathbb{E}[\log_2 K] \le D_{\mathsf{KL}}(P(\cdot|x) \| Q(\cdot)) + \log_2(3.56)/\min((\alpha-1)/2, 1).$$

K can be encoded by a prefix-free code with expected length approximately $D_{\mathsf{KL}}(P(\cdot|x)\|Q(\cdot))$ bits within a logarithmic gap. If X is random, the expected length is approximately $I(X;Z)$ which is almost optimal.

As a result, when $P(\cdot|x)$ satisfies ε-local DP, then the compression size is at most

$$\ell + \log_2(\ell + 1) + 2 \ \text{(bits)},$$

where $\ell \triangleq \varepsilon \log_2 e + \log_2(3.56)/\min((\alpha-1)/2, 1)$.

Remarks:

- The exactness of PPR follows from the Poisson functional representation.
- While the algorithm requires an infinite number of samples, it can be reparameterized and terminates in finite steps.
- PPR can also be used to compress central DP mechanisms and offer (weaker) local DP guarantees.

Privacy Guarantees of PPR

Theorem 3 (ε-DP of PPR).

If the mechanism $P(\cdot|x)$ is ε-DP, then PPR $P_{(Z_i)_i, K|x}$ with $\alpha > 1$ is $2\alpha\varepsilon$-DP.

Theorem 4 ((ε, δ)-DP of PPR).

If the mechanism $P(\cdot|x)$ is (ε, δ)-DP, then PPR $P_{(Z_i)_i, K|x}$ with $\alpha > 1$ is $(2\alpha\varepsilon, 2\delta)$-DP.

Theorem 5 (Tighter (ε, δ)-DP of PPR).

If the mechanism $P(\cdot|x)$ is (ε, δ)-DP, then PPR $P_{(Z_i)_i, K|x}$ with $\alpha > 1$ is $(\alpha\varepsilon + \tilde{\varepsilon}, 2(\delta + \tilde{\delta}))$-DP, for every $\tilde{\varepsilon} \in (0, 1]$ and $\tilde{\delta} \in (0, 1/3]$ satisfying

$$\alpha \le e^{-4.2}\tilde{\delta}\tilde{\varepsilon}^2/(-\ln\tilde{\delta}) + 1.$$

Applications

PPR-compressed Gaussian mechanism.

Consider the Gaussian mechanism

$$P_{Z|X}(\cdot|x) = \mathcal{N}\left(x, \frac{\sigma^2}{n}\mathbb{I}_d\right)$$

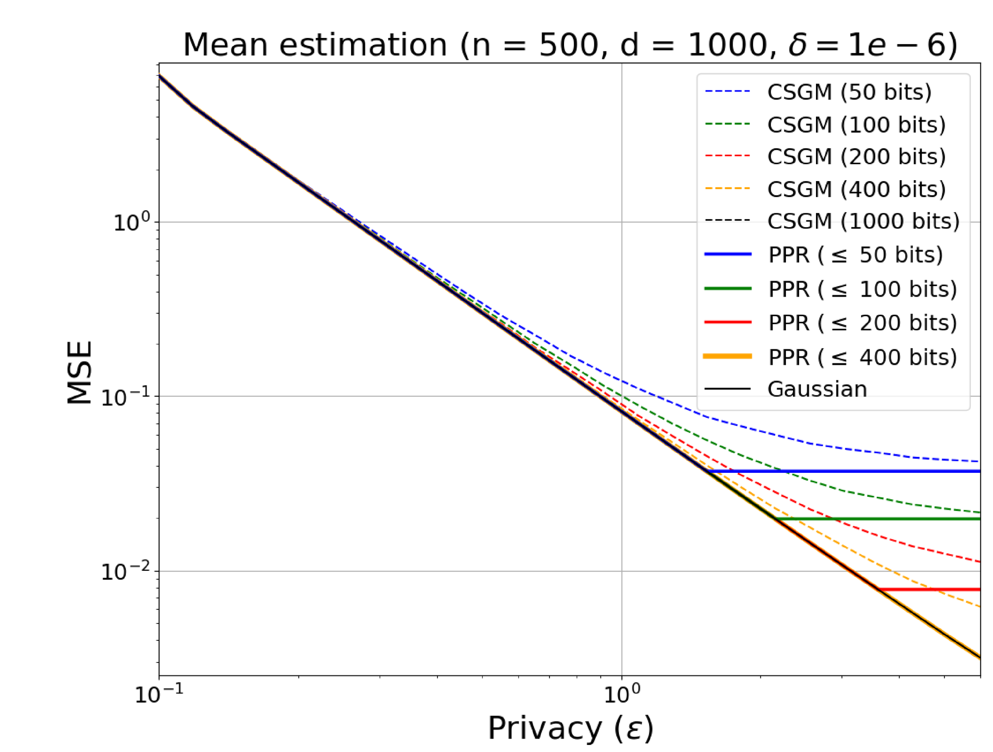
and the reference distribution

$$Q = \mathcal{N}\left(0, \left(\frac{C^2}{d} + \frac{\sigma^2}{n}\right)\mathbb{I}_d\right),$$

where $\sigma \ge C\sqrt{2\ln(1.25/\delta)}/\varepsilon$. Let Z_i be the output of PPR applied on $P_{Z|X}(\cdot|x_i)$. Then it holds that

- $\hat{\mu}(Z^n) := \frac{1}{n}\sum_i Z_i$ is unbiased w.r.t. the true mean.
- $\hat{\mu}(Z^n)$ satisfies (ε, δ)-central DP.
- PPR satisfies $(2\alpha\sqrt{n}\varepsilon, 2\delta)$-local DP.
- The average per-client communication cost is at most

$$O\left(d \log\left(\frac{n\varepsilon^2}{d\log(1/\delta) + 1}\right) + 1\right) \text{ bits.}$$



References

[1] Feldman and Talwar, *"Lossless compression of efficient private local randomizers,"* ICML 2021.

[2] Shah, Chen, Balle, Kairouz and Theis, *"Optimal Compression of Locally Differentially Private Mechanisms,"* AISTATS 2022.

[3] Li and El Gamal, *"Strong Functional Representation Lemma and Applications to Coding Theorems,"* IEEE Trans. Inf. Theory, 2018.

[4] Li and Anantharamm, *"A unified framework for one-shot achievability via the Poisson matching lemma,"* IEEE Trans. Inf. Theory, 2021.

[5] Maddison, *"A Poisson process model for Monte Carlo,"* Perturbation, Optimization, and Statistics, 2016.