Exposé de candidature MCF pour le poste N°4887

..en Statistique et Thèmes Connexes (Section 26) à l'UT3 Paul Sabatier / l'IMT

Clément Lalanne

Chercheur postdoctoral à Toulouse School of Economics

clement.lalanne@tse-fr.eu

2024-04-29

Table of Contents

3 Curriculum

2016-2020 : Élève fonctionnaire stagiaire à l'ENS en Informatique

- **Master 1 MPRI Master 2 MVA**
- **Stage de L3** : **De-anonymization and privacy : Study of a random graph model** encadré par **Florian Simatos**

- **Stage de M1** : **Storage-optimal continuous optimization encadré par Volkan Cevher**

- **Stage de M2** : **Nystagmus waveform extraction using convolutional dictionary learning with detrending** encadré par **Laurent Oudre**, **Nicolas Vayatis** et **Thomas Moreau**

4 Curriculum

2020-2023 : Doctorat à l'ENS de Lyon en Informatique

- Sujet : **Sur les compromis liés à l'apprentissage statistique sous contraintes de confidentialité**
- Encadré par **Aurélien Garivier** et **Rémi Gribonval**
- **Chargé de TDs (3x64h)** :
  - Apprentissage Statistique (M1 ENS Lyon)
  - Bases de données et Exploration de données (M1 ENS Lyon)
  - Entrainement à la programmation sportive (L3 ENS Lyon)
  - Optimisation (M1 ENS Lyon)
  - Programmation Système (IUT 2A UCBL)
  - Évaluation de stages, ...

5 Curriculum

2023-Ajd : Postdoc en Mathématiques à TSE

- Sujet : **Optimisation des réseaux de neurones profonds**
- Encadré par **Jérôme Bolte** et **Sébastien Gadat**

Table of Contents

7 Thèmes de recherche et Publications

Thèmes de recherche : Statistique confidentielle, Optimisation

Preprints :

- ⋆ **Privately Learning Smooth Distributions on the Hypercube by Projections**, Clément Lalanne et Sébastien Gadat, 2024

Journaux :

- ⋆ **About the Cost of Central Privacy in Density Estimation**, Clément Lalanne, Aurélien Garivier et Rémi Gribonval, dans **Transactions on Machine Learning Research**, 2023
- **Private Quantiles Estimation in the Presence of Atoms**, Clément Lalanne, Clément Gastaud, Nicolas Grislain, Aurélien Garivier et Rémi Gribonval, dans **Information and Inference**, 2023
- ⋆ **On the Statistical Complexity of Estimation and Testing under Privacy Constraints**, Clément Lalanne, Aurélien Garivier et Rémi Gribonval, dans **Transactions on Machine Learning Research**, 2023

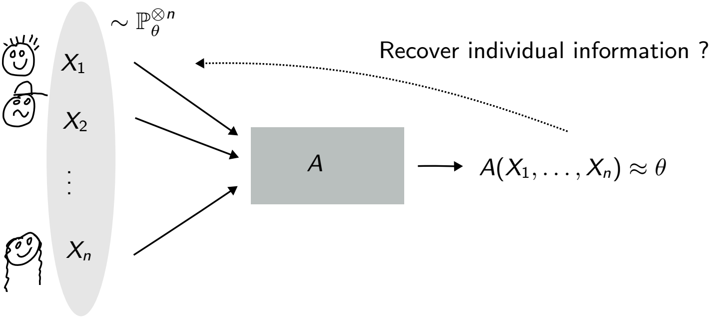
8 Publications

Conférences :

- **Private Statistical Estimation of Many Quantiles**, Clément Lalanne, Aurélien Garivier et Rémi Gribonval, dans **International Conference on Machine Learning**, 2023
- **Can sparsity improve the privacy of neural networks?**, Antoine Gonon, Léon Zheng, Clément Lalanne, Quoc-Tung Le, Guillaume Lauga et Can Pouliquen, dans **GRETSI** (francophone), 2023
- **Extraction of Nystagmus Patterns from Eye-Tracker Data with Convolutional Sparse Coding**, Clément Lalanne, Maxence Rateaux, Laurent Oudre, Matthieu Robert et Thomas Moreau, dans **IEEE Engineering in Medicine and Biology Society**, 2020





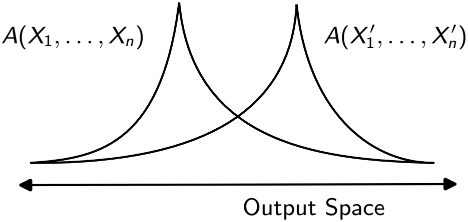
Question :

Peut-on se défendre contre ***toute*** *attaque ?*

Voisinage : $\mathbf{X} \sim \mathbf{X}'$ ssi $\mathbf{X}$ peut être obtenu à partir de $\mathbf{X}'$ en changeant les données d'un individu.

Définition : A est $\epsilon > 0$-DP si $A(\mathbf{X})$ est ϵ-proche de $A(\mathbf{X}')$ pour tous $\mathbf{X} \sim \mathbf{X}'$.[1]

Rôle de ϵ : Plus ϵ est petit, plus A est confidentiel.



Question :

Quel est le coût de la confidentialité sur l'estimation statistique ?

[1] Cynthia Dwork et al. "Calibrating Noise to Sensitivity in Private Data Analysis". In: 2006.

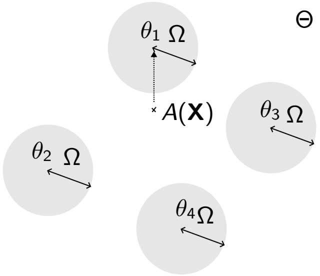
11 Un schéma de preuve pour les bornes inférieures

Setup : $\theta \in \Theta$, $\mathbf{X} = (X_1, \ldots, X_n) \sim \mathbb{P}_\theta^{\otimes n}$.

Risque minimax :[2,3]

$$\boxed{\mathfrak{M}_n := \inf_A \sup_{\theta \in \Theta} \mathbb{E}_{A,\mathbf{X}\sim\theta}\left(\mathrm{Error}(A(\mathbf{X}), \theta)\right)} \geq \inf_A \boxed{\sup_{i=1,\ldots,N} \mathbb{E}_{A,\mathbf{X}\sim\theta_i}\left(\mathrm{Error}(A(\mathbf{X}), \theta_i)\right)}$$

$$\geq \Phi(\Omega) \boxed{\sup_{i=1,\ldots,N} \mathbb{P}_{A,\mathbf{X}\sim\theta_i}\left(\hat{\imath}\,(A(\mathbf{X})) \neq i\right)}$$



[2] Alexandre B. Tsybakov. *Introduction to Nonparametric Estimation*. 2009.

[3] $\mathrm{Error}(\cdot,\cdot) = \Phi(d(\cdot,\cdot))$

Risque minimax :

$$\mathfrak{M}_n \geq \Phi(\Omega) \sup_{i=1,\ldots,N} \mathbb{P}_{A,\mathbf{X}\sim\theta_i}\left(\hat{\imath}\,(A(\mathbf{X})) \neq i\right)$$

Problème de transport :

$$\sup_{i=1,\ldots,N} \mathbb{P}_{A,\mathbf{X}\sim\theta_i}\left(\hat{\imath}\,(A(\mathbf{X})) \neq i\right) \geq \boxed{\sup_{\mathbb{Q}\in\Pi\left(\mathbb{P}_1^{\otimes n},\ldots,\mathbb{P}_N^{\otimes n}\right)} \int s\left(\mathbf{X}_1,\ldots,\mathbf{X}_N\right) d\mathbb{Q}\left(\mathbf{X}_1,\ldots,\mathbf{X}_N\right)},$$

where s is a *similarity function* satisfying, for any $\mathbf{X}_1,\ldots,\mathbf{X}_N$,

$$\boxed{\frac{1}{N}\sum_{i=1}^{N} \mathbb{P}_A\left(\hat{\imath}\,(A(\mathbf{X}_i)) \neq i\right) \geq s\left(\mathbf{X}_1,\ldots,\mathbf{X}_N\right)}.$$

13 Un schéma de preuve pour les bornes inférieures

Risque minimax :

$$\mathfrak{M}_n \geq \Phi(\Omega) \sup_{i=1,\dots,N} \mathbb{P}_{A,\mathbf{X}\sim\theta_i}\left(\hat{i}\,(A(\mathbf{X})) \neq i\right)$$

Le Cam :[4]

$$\boxed{\sup_{i=1,2} \mathbb{P}_{A,\mathbf{X}\sim\theta_i}\left(\hat{i}\,(A(\mathbf{X})) \neq i\right) \geq \frac{1}{2}\left(1 - \left(1 - e^{-\epsilon}\right)\mathrm{TV}\left(\mathbb{P}_{\theta_1}, \mathbb{P}_{\theta_2}\right)\right)^n} \, .$$

Fano :[5]

$$\boxed{\sup_{i=1,\dots,N} \mathbb{P}_{A,\mathbf{X}\sim\theta_i}\left(\hat{i}\,(A(\mathbf{X})) \neq i\right) \geq 1 - \frac{1 + \frac{n\epsilon}{N^2}\sum_{i,j=1}^{N} \frac{2\mathrm{TV}\left(\mathbb{P}_{\theta_i},\mathbb{P}_{\theta_j}\right)}{1+\mathrm{TV}\left(\mathbb{P}_{\theta_i},\mathbb{P}_{\theta_j}\right)}}{\ln(N)}} \, .$$

[4] $\mathrm{TV}\left(\mathbb{P}_1, \mathbb{P}_2\right) := \sup_S \mathbb{P}_1(S) - \mathbb{P}_2(S)$

[5] $\mathrm{KL}\left(\mathbb{P}_1 \| \mathbb{P}_2\right) := \int \ln\left(\frac{d\mathbb{P}_1}{d\mathbb{P}_2}\right) d\mathbb{P}_1$

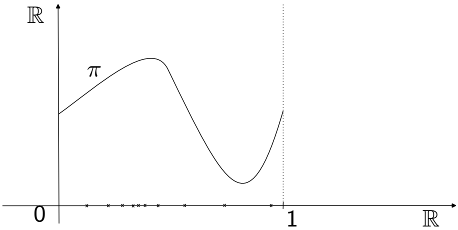
14 Application à l'estimation de densités

Vitesse d'estimation : $\boxed{\Theta\left(n^{-\frac{2\beta}{2\beta+d}} + (n\epsilon)^{-\frac{2\beta}{\beta+d}}\right)}$

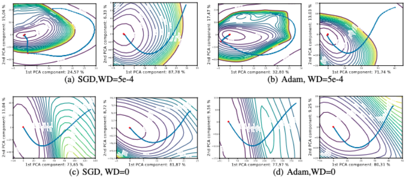
Difficulté technique : Pour les bornes inférieures, espaces de Sobolev-Hölder

$$\sum_{|\alpha|=\lfloor\beta\rfloor} \|\partial^\alpha \pi\|_2^2 + \mathbf{1}_{\beta-\lfloor\beta\rfloor>0} \sum_{|\alpha|=\lfloor\beta\rfloor} \|\partial^\alpha \pi\|_{\mathcal{H}_{\beta-\lfloor\beta\rfloor}}^2 \leq L^2$$

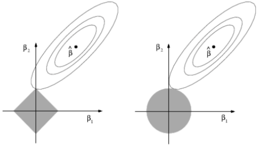
Adaptivité : Possible en adaptant la méthode de Lepskii + Polylog



Optimisation des réseaux de neurones profonds par des méthodes du second ordre Travaux actuels de postdoc

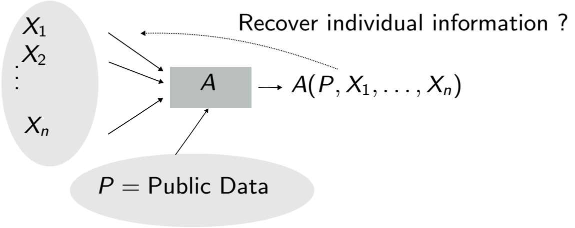


Réduire le fléau de la dimension avec confidentialité

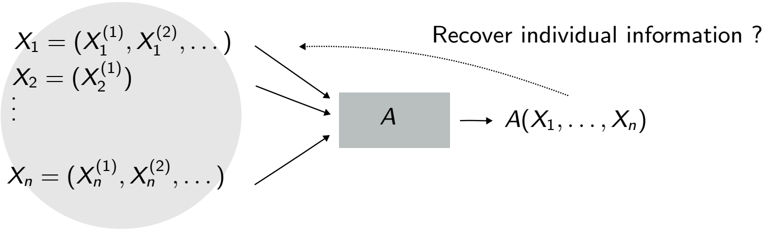


16 Projet de recherche

Confidentialité mixte

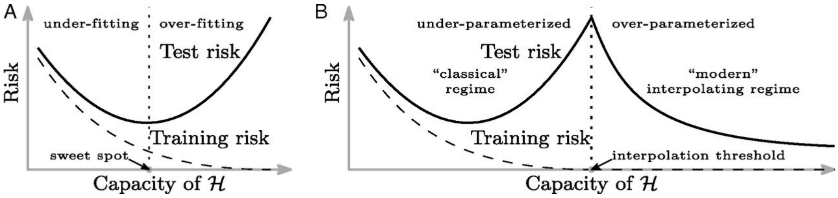


Confidentialité avec contributions asymétriques



17 Projet de recherche

Confidentialité et sur-paramétrisation



Temps d'entrainement des réseaux de neurones et confidentialité

DeepMind

Unlocking High-Accuracy Differentially Private Image Classification through Scale

Soham De[*,1], Leonard Berrada[*,1], Jamie Hayes[1], Samuel L Smith[1] and Borja Balle[1]
[*]Equal contributions, [1]DeepMind

Projet de recherche proche des thématiques de l'équipe Statistiques et Optimisation de l'IMT

- Optimisation pour le machine learning
- Statistiques générales et statistiques non paramétriques
- Machine learning responsable (ex confidentialité, équité, robustesse)
- Possibles applications à des données médicales, aviation, ...

Recherches d'actualité

Figure: Google trend of "Machine Learning"

Figure: Google trend of "Differential Privacy"

Figure: Google trend of "GPU"

Table of Contents

20 Expériences & qualifications

Expérience : Chargé de TDs (3x64h)

- Apprentissage Statistique (M1 ENS Lyon)
- Bases de données et Exploration de données (M1 ENS Lyon)
- Entrainement à la programmation sportive (L3 ENS Lyon)
- Optimisation (M1 ENS Lyon)
- Programmation Système (IUT 2A UCBL)
- Évaluation de stages, ...

Agrégé de mathématiques (2023)

Maîtrise des outils techniques

Maîtrise de l'anglais

21 Intégration

Offre de formation en Mathématiques 2022/2026

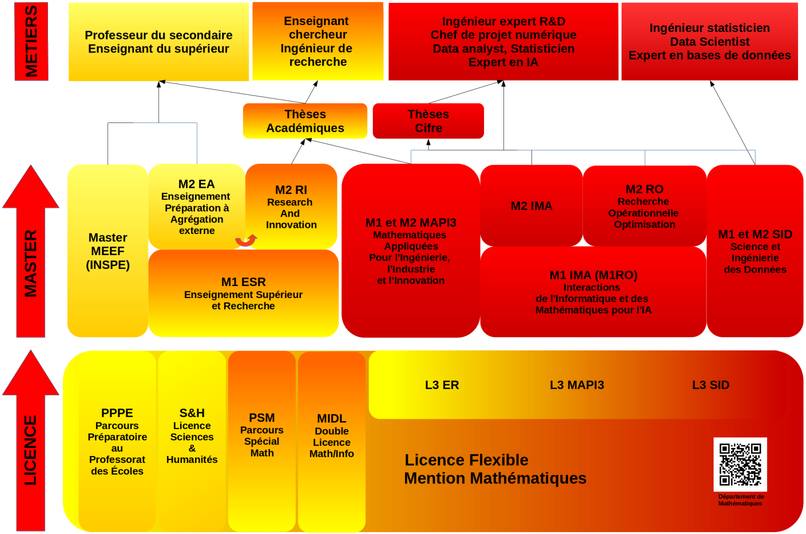


Table of Contents

23 Pourquoi l'UT3 / l'IMT ?

Excellence

Écosystème de recherche

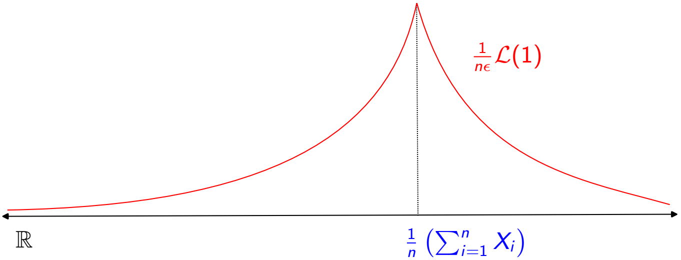
Moyens

Merci de votre attention

25 Private estimator

Laplace mechanism :

$$A(\mathbf{X}) = \frac{1}{n}\left(\sum_{i=1}^{n} X_i\right) + \frac{1}{n\epsilon}\mathcal{L}(1)$$



Laplace mechanism :

$$A(\mathbf{X}) = \frac{1}{n}\left(\sum_{i=1}^{n} X_i\right) + \frac{1}{n\epsilon}\mathcal{L}(1)$$

is ϵ-DP.

Error :

$$\mathbb{E}_{\mathbb{P}_A,\mathcal{B}(\theta)^{\otimes n}}\left((A(\mathbf{X}) - \theta))^2\right) \leq \frac{1/4}{n} + \frac{2}{n^2\epsilon^2}$$

Two regimes :

- **Low privacy regime :** $\epsilon = \Omega(1/\sqrt{n})$, no significant effect on estimation.
- **High privacy regime :** $\epsilon \ll 1/\sqrt{n}$, the precision can be arbitrarily degraded.

Question :

Is it possible to do better ?

27 Building similarity functions

$$\frac{1}{N}\sum_{i=1}^{N}\mathbb{P}_A\left(\hat{i}\left(A\left(\mathbf{X}_i\right)\right)\neq i\right)\geq s\left(\mathbf{X}_1,\ldots,\mathbf{X}_N\right).$$

Definition : A is ϵ-DP if $\mathbf{X}\sim\mathbf{X}' \implies \mathbb{P}\big(A(\mathbf{X})\in S\big)\leq e^{\epsilon}\times\mathbb{P}\big(A(\mathbf{X}')\in S\big)$.[6]

Two marginals : $\boxed{\frac{1}{2}\sum_{i=1}^{2}\mathbb{P}_A\left(\hat{i}\left(A\left(\mathbf{X}_i\right)\right)\neq i\right)\geq\frac{1}{2}e^{-\epsilon d_{\mathrm{ham}}(\mathbf{X}_1,\mathbf{X}_2)}\,^{7}}$

Many marginals : $\boxed{\frac{1}{N}\sum_{i=1}^{N}\mathbb{P}_A\left(\hat{i}\left(A\left(\mathbf{X}_i\right)\right)\neq i\right)\geq 1-\frac{1+\frac{\epsilon}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}d_{\mathrm{ham}}\left(\mathbf{X}_i,\mathbf{X}_j\right)}{\ln(N)}}$

[6] Dwork et al., "Calibrating Noise to Sensitivity in Private Data Analysis".
[7] $d_{\mathrm{ham}}\left((X_1,\ldots,X_n),(X_1',\ldots,X_n')\right):=\sum_{i=1}^{n}\mathbb{1}_{X_i=X_i'}$

28 Back to the transport problem

$$\boxed{\sup_{\mathbb{Q}\in\Pi\left(\mathbb{P}_1^{\otimes n},\ldots,\mathbb{P}_N^{\otimes n}\right)}\int s\left(\mathbf{X}_1,\ldots,\mathbf{X}_N\right)d\mathbb{Q}\left(\mathbf{X}_1,\ldots,\mathbf{X}_N\right)},$$

where s is **non-increasing** *in* $d_{\mathrm{ham}}\left(\mathbf{X}_i,\mathbf{X}_j\right)$ *for any* i,j.

Question :

How to construct a coupling that makes those quantities big ?

29 A good enough coupling

$$\sup_{\mathbb{Q}\in\Pi\left(\mathbb{P}_1^{\otimes n},\dots,\mathbb{P}_N^{\otimes n}\right)} \int s\left(\mathbf{X}_1,\dots,\mathbf{X}_N\right) d\mathbb{Q}\left(\mathbf{X}_1,\dots,\mathbf{X}_N\right),$$

where s is **non-increasing** *in* $d_{\mathrm{ham}}\left(\mathbf{X}_i,\mathbf{X}_j\right)$ *for any* i,j.

Near optimal coupling for equalities : There exists $(X_i)_{i=1,\dots,N}$ of distribution χ, a coupling between $(\mathbb{P}_i)_{i=1,\dots,N}$ such that[8]

$$\forall i,j,\quad \mathrm{TV}\left(\mathbb{P}_i,\mathbb{P}_j\right) \le \boxed{\mathbb{P}(X_i\neq X_j) \le \frac{2\mathrm{TV}\left(\mathbb{P}_i,\mathbb{P}_j\right)}{1+\mathrm{TV}\left(\mathbb{P}_i,\mathbb{P}_j\right)}}\,.$$

Final coupling : $\boxed{\mathbb{Q}^* = \chi^{\otimes n}}$

[8] Omer Angel and Yinon Spinka. *Pairwise optimal coupling of multiple random variables*. 2021.

30 L_2 approximations and projection estimators

Reference Fourier basis :

$$\begin{aligned}\phi_1(x) &= 1 \\ \phi_{2k}(x) &= \sqrt{2}\sin(2\pi kx) \quad k \geq 1 \\ \phi_{2k+1}(x) &= \sqrt{2}\cos(2\pi kx) \quad k \geq 1 \,.\end{aligned}$$

L_2 approximation :

$$\sum_{i=1}^{N} \theta_i \phi_i \xrightarrow[N\to+\infty]{L^2} \pi \quad \text{where} \quad \theta_i := \int_{[0,1]} \pi \, \phi_i \,.$$

Projection estimator :[9]

$$\hat{\pi}^{\text{proj}}(\mathbf{X}) = \sum_{i=1}^{N} \hat{\theta}_i \phi_i \quad \text{where} \quad \hat{\theta}_i := \frac{1}{n}\sum_{j=1}^{n} \phi_i(X_j) \,.$$

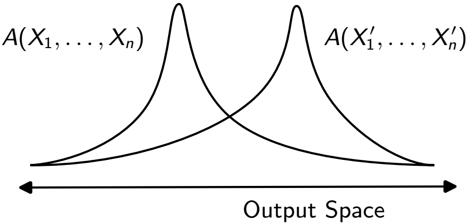
Question : *How do we add privacy ?*

[9] Tsybakov, *Introduction to Nonparametric Estimation*.

31 Concentrated Differential Privacy

Definition :[10][11] A is ρ-zCDP if $\mathbf{X} \sim \mathbf{X}' \implies \forall \alpha > 0, \quad \mathrm{D}_\alpha\left(A(\mathbf{X}) \| A(\mathbf{X}') \right) \leq \alpha\rho$, where

$$\mathrm{D}_\alpha\left(\mathbb{P} \| \mathbb{Q} \right) := \frac{1}{\alpha - 1} \ln \int \left(\frac{d\mathbb{P}}{d\mathbb{Q}} \right)^{\alpha-1} d\mathbb{Q} \,.$$



[10] Cynthia Dwork and Guy N Rothblum. "Concentrated differential privacy". In: (2016).

[11] Mark Bun and Thomas Steinke. "Concentrated Differential Privacy: Simplifications, Extensions, and Lower Bounds". In: 2016.

32 Lower-bounds and upper-bounds with CDP

Private projection estimator :

$$\boxed{\hat{\pi}^{\text{proj}}(\mathbf{X}) = \sum_{i=1}^{N} \left(\hat{\theta}_i + C_{\rho,N}\mathcal{N}(0,1)\right)\phi_i \quad \text{where} \quad \hat{\theta}_i := \frac{1}{n}\sum_{j=1}^{n} \phi_i(X_j)}.$$

$\hat{\pi}^{\text{proj}}$ is ρ-zCDP.

Resulting upper-bound :

$$\boxed{\mathbb{E}\left(\|\hat{\pi}^{\text{proj}}(\mathbf{X}) - \pi\|_{L_2}^2\right) \leq C_{L,\beta} \max\left\{n^{-\frac{2\beta}{2\beta+1}}, (n\sqrt{\rho})^{-\frac{2\beta}{\beta+1}}\right\}}$$

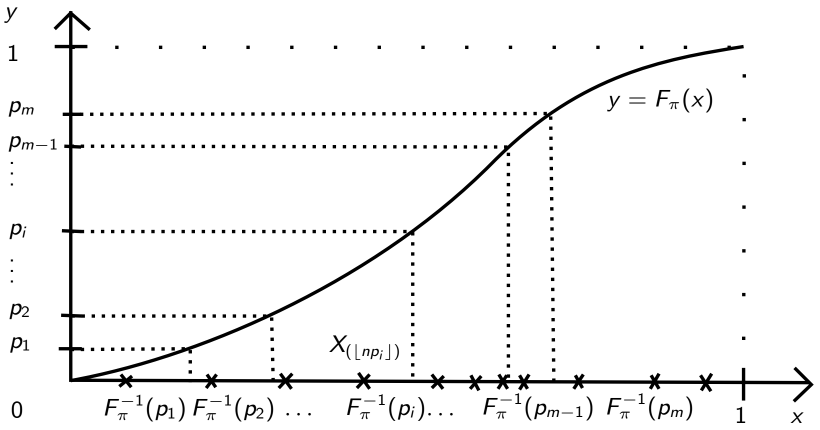
Lower-bound against ρ-zCDP estimators :

$$\boxed{\inf_{A}\sup_{\pi} \mathbb{E}\left(\|A(\mathbf{X}) - \pi\|_{L_2}^2\right) \geq C_{L,\beta} \max\left\{n^{-\frac{2\beta}{2\beta+1}}, (n\sqrt{\rho})^{-\frac{2\beta}{\beta+1}}\right\}}$$

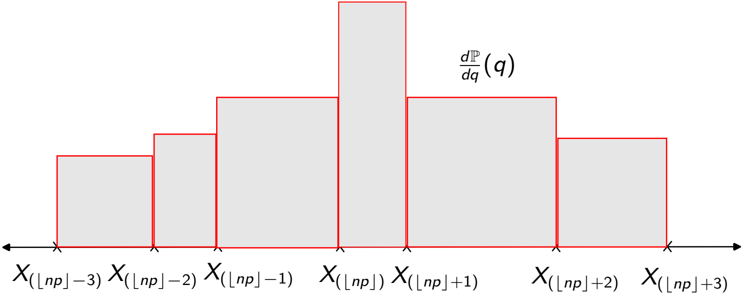
33 Quantiles Estimation Problem

Inputs : samples $\mathbf{X} = (X_1, \ldots, X_n) \overset{\text{i.i.d.}}{\sim} \mathbb{P}_\pi$ $\mathbf{p} = (p_1, \ldots, p_m) \in (0,1)^m$ sorted.

Desired output : Quantile estimator $\mathbf{q} \in [0,1]^m$ of $(F_\pi^{-1}(p_1), \ldots, F_\pi^{-1}(p_m))$.



Mechanism :[12] For a single quantile q (associated with p),



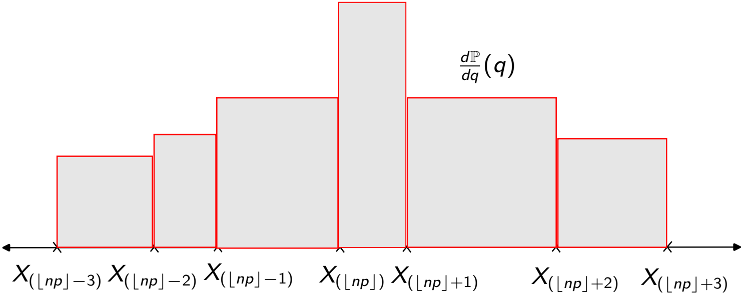
Concentration result :[14] When π is away from 0 on a neighborhood of $F_\pi^{-1}(p)$,

$$\mathbb{P}\left(|q - F_\pi^{-1}(p)| > \gamma\right) \leq P(n) \max\left(e^{-C_1 \epsilon n \gamma}, e^{-C_2 \gamma^2 n}\right).$$

[12] Adam D. Smith. "Privacy-preserving statistical estimation with optimal convergence rates". In: 2011.
[13] $d\mathbb{P}(q) \propto e^{-\frac{\epsilon}{2}\left||\{i|X_i<q\}|-\lfloor np \rfloor\right|} dq$
[14] Clément Lalanne, Aurélien Garivier, and Rémi Gribonval. "Private Statistical Estimation of Many Quantiles". In: 2023.

Independent Private Quantiles

Idea : Use QExp independently on **p** with simple composition.



Concentration result :[16] When π is away from 0 on a neighborhood of $F_\pi^{-1}(\mathbf{p})$,

$$\mathbb{P}\left(\|\mathbf{q} - F_\pi^{-1}(\mathbf{p})\|_\infty > \gamma\right) \leq P(n, m) \max\left(e^{-C_1 \frac{\epsilon n \gamma}{\mathbf{m}}}, e^{-C_2 \gamma^2 n}\right).$$

[15] $d\mathbb{P}(q_i) \propto e^{-\frac{\epsilon}{2m}\left| |\{i | X_i < q_i\}| - \lfloor np_i \rfloor \right|} dq_i$

[16] Lalanne, Garivier, and Gribonval, "Private Statistical Estimation of Many Quantiles".

36 Joint Exponential Private Quantiles

Idea :[17] Leverage structural dependencies. Quantiles are non-decreasing, between q_i and q_j should fall approximately $n|p_i - p_j|$ points. [18]

"Fun" discovery :[19] JointExp $\approx$ Inverse Sensitivity Mechanism[20].

Consistency result : When π is away from 0 on a neighborhood of $F_\pi^{-1}(\mathbf{p})$, JointExp is consistent.

[17] Jennifer Gillenwater, Matthew Joseph, and Alex Kulesza. "Differentially Private Quantiles". In: 2021.

[18] $d\mathbb{P}(\mathbf{q}) \propto e^{-\frac{\epsilon}{2}\sum_{i=1}^{m+1}\left|\delta^{\mathrm{JE}}(i,\mathbf{X},\mathbf{q})\right|} d\mathbf{q}$ where $\delta^{\mathrm{JE}}(i,\mathbf{X},\mathbf{q}) := n(p_i - p_{i-1}) - \#\left(\mathbf{X} \cap (q_{i-1}, q_i]\right)$

[19] Clément Lalanne et al. "Private quantiles estimation in the presence of atoms". In: (2023). ISSN: 2049-8772.

[20] Hilal Asi and John C. Duchi. "Near Instance-Optimality in Differential Privacy". In: *CoRR* (2020). arXiv: 2005.10630.

Idea :[21] Use QExp recursively with a dichotomy on $\mathbf{p}$.

Concentration result :[22] When π is away from 0 on a neighborhood of $F_\pi^{-1}(\mathbf{p})$,

$$\mathbb{P}\left(\|\mathbf{q} - F_\pi^{-1}(\mathbf{p})\|_\infty > \gamma\right) \leq P(n, m) \max\left(e^{-C_1 \frac{\epsilon n \gamma}{(\log_2(m))^2}}, e^{-C_2 \gamma^2 n}\right).$$

Remark : Almost polylogarithmic degradation in m !

[21] Haim Kaplan, Shachar Schnapp, and Uri Stemmer. "Differentially Private Approximate Quantiles". In: ed. by Kamalika Chaudhuri et al. PMLR, 2022.

[22] Lalanne, Garivier, and Gribonval, "Private Statistical Estimation of Many Quantiles".

Idea :[23]

$$\hat{\pi}^{\text{hist}}(t) := \sum_{b \in \text{bins}} \mathbb{1}_b(t) \frac{1}{nh} \left(\sum_{i=1}^{n} \mathbb{1}_b(X_i) + \frac{2}{\epsilon} \mathcal{L}_b \right) .$$

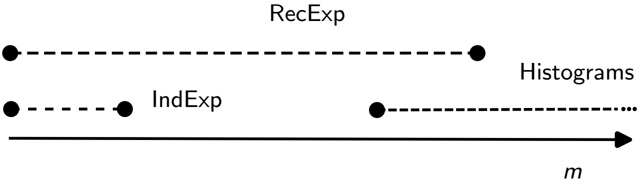
Concentration result :[24] Bins of size h, $\gamma > C_4 h$, π is L-Lipschitz, I is a strict sub-interval

$$\mathbb{P}\left(\| F_{\hat{\pi}^{\text{hist}}}^{-1} - F_{\pi}^{-1} \|_{\infty, I} > \gamma \right) \leq \frac{1}{h} e^{-C_1 \gamma h n \epsilon} + \frac{2}{h} e^{-C_2 h^2 (C_3 \gamma - L h)^2 n} .$$

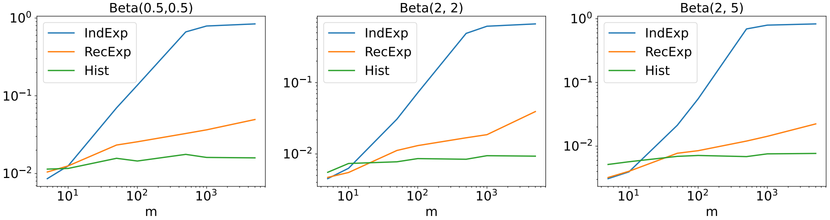
Remark : No degradation in m, but high entry cost.

[23] Larry A. Wasserman and Shuheng Zhou. "A Statistical Framework for Differential Privacy". In: (2010).
[24] Lalanne, Garivier, and Gribonval, "Private Statistical Estimation of Many Quantiles".

 Theoretical choice of algorithm



40 Numerical validation

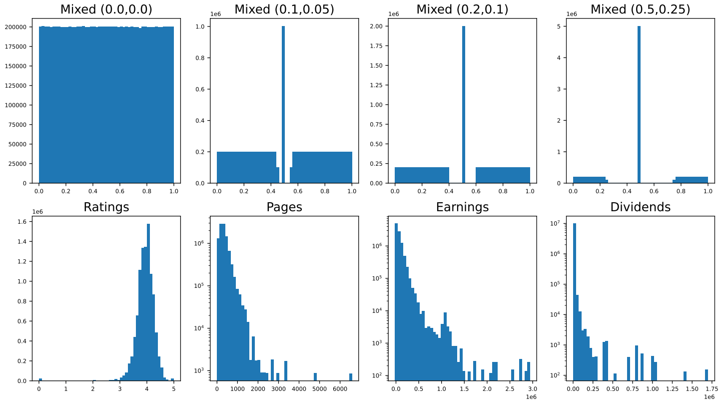


The vertical axis reads the error $\mathbb{E}\left(\|\mathbf{q} - F^{-1}(\mathbf{p})\|_\infty\right)$ where $\mathbf{p} = \left(\frac{1}{4} + \frac{1}{2(m+1)}, \ldots, \frac{1}{4} + \frac{m}{2(m+1)}\right)$ for different values of m, $n = 10000$, $\epsilon = 0.1$, and $\mathbb{E}$ is estimated by Monte-Carlo averaging over 50 runs. The histogram is computed on 200 bins.

41 Dealing with atomic distributions

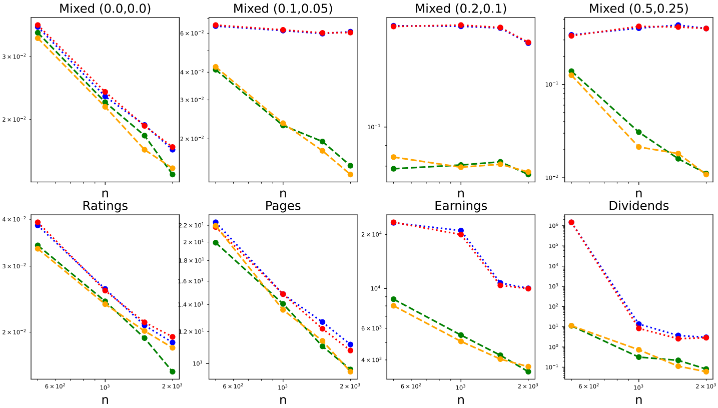
Inconsistency result :[25] When dealing with atomic distributions, all the *Exp mechanisms are inconsistent or have poor performances.

Proposed solution : Smoothing the distribution with noise addition can make those mechanisms consistent and helps the performances.



[25] Lalanne et al., "Private quantiles estimation in the presence of atoms".

42 Dealing with atomic distributions



The vertical axis reads the error $\mathbb{E}\left(\|\hat{\mathbf{q}} - F^{-1}(\mathbf{p})\|_\infty\right)$ where $\mathbf{p} = \left(\frac{1}{m+1}, \ldots, \frac{m}{m+1}\right)$ for $m = 8$, $\epsilon = 1$, $\hat{\mathbf{q}}$ is the private estimator, and $\mathbb{E}$ is estimated by Monte-Carlo averaging over 50 runs.

43 EyeFant Tracker

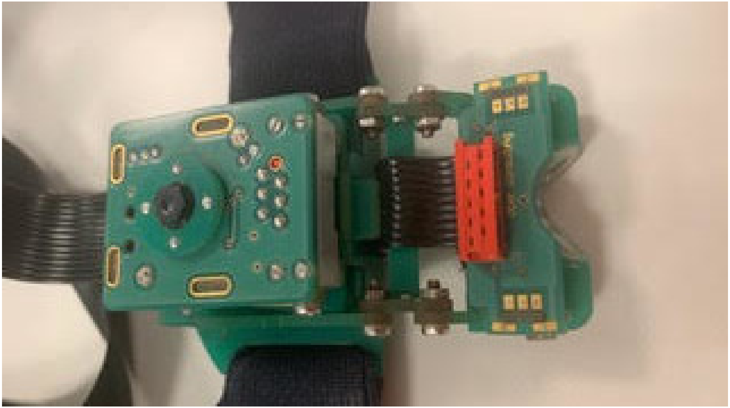


Figure: EyeFant Tracker

44 Nystagmus

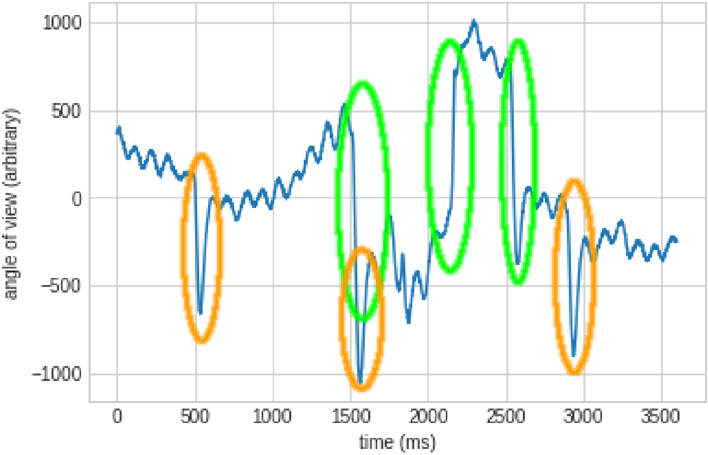
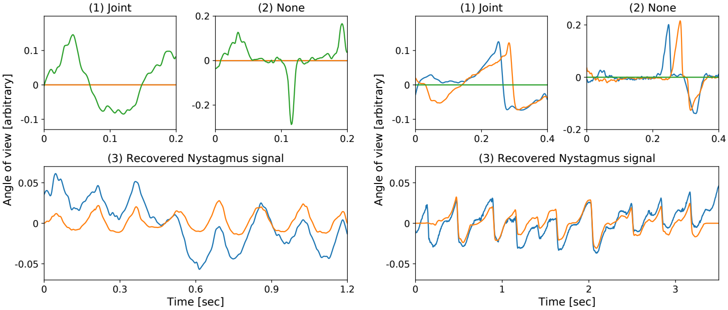


Figure: Example of eye tracker recording of a patient with a pendular Nystagmus (left-eye, horizontal). The signal includes saccades (*green*) as well as eye blinks artefacts (*orange*) hindering the automatic extraction of Nystagmus waveforms.

45 Model

$$\min_{\substack{\mathbf{D},\mathbf{Z},\mathbf{y} \\ \forall k, ||D_k||_2^2 \leq 1}} \frac{1}{2}||\mathbf{D} * \mathbf{Z} + \mathbf{y} - \mathbf{x}||_2^2 + \lambda||\mathbf{Z}||_1 + \lambda_{TV}||\nabla \mathbf{y}||_1,$$

46 Results



Results on Patient 1 Results on Patient 2

Figure: Results for recordings on two patients presenting (*a*) a SN syndrome with a pendular pattern and (*b*) a INS syndrome with characteristic jerky pattern. The patterns learned with JOINT (1) are better at capturing the characteristic waveform of the Nystagmus than the one learned with NONE (2). The bottom part (3) displays a selected part of the original signals (*blue*) as well as the estimated nystagmic component (*orange*). The signal parts are selected to exclude artefacts and saccades, as it is easier to see the Nystagmus patterns in this configuration.