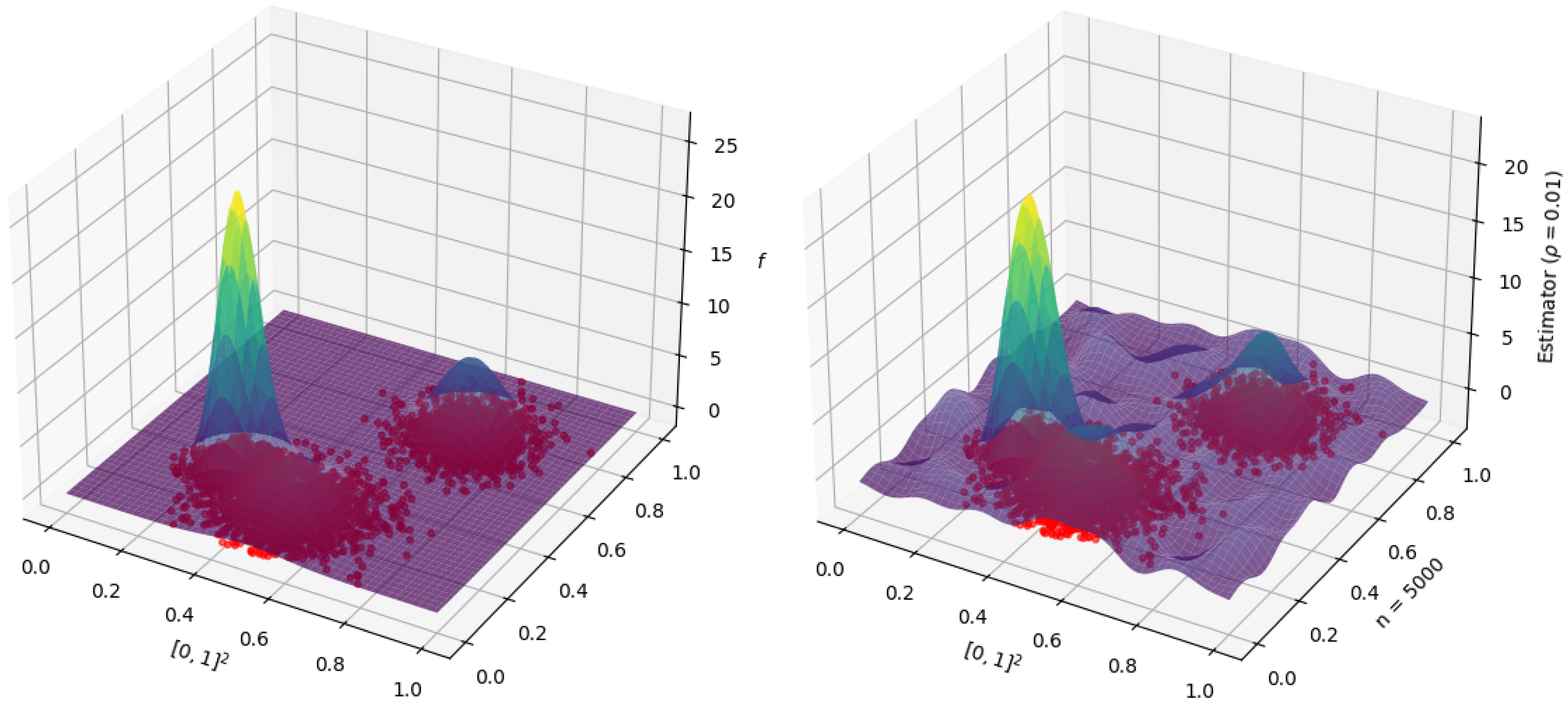
Privately Learning Smooth Distributions on the Hypercube by Projections

Clément Lalanne, Sébastien Gadat
Toulouse School of Economics, Toulouse 1 University, France

I - Problem

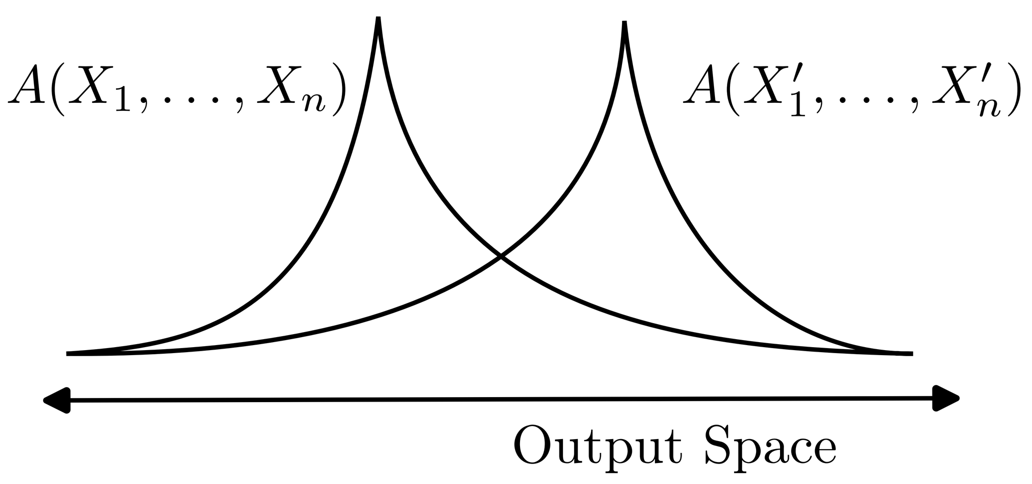
- f : probability density on $[0,1]^d$.
- **Inputs :** samples $X_1, \dots, X_n \overset{\text{i.i.d.}}{\sim} \mathbb{P}_f$
- **Desired output :** Private estimator $\hat{f}$ of f.

I - Problem (Visual)



II - Concentrated DP [1, 2, 3]

- **Neighboring relation :** $(X_1, \dots, X_n) \sim (X'_1, \dots, X'_{n'})$ iff $(X_1, \dots, X_n)$ can be obtained from $(X'_1, \dots, X'_{n'})$ with a permutation and a replacement.
- **Definition :** $\mathbf{X} = (X_1, \dots, X_n) \sim \mathbf{Y} = (X_1, \dots, X_{n-1}, X'_n) \implies \forall 1 < \alpha < +\infty : \mathrm{D}_\alpha\left(A(\mathbf{X}) \| A(\mathbf{Y})\right) \leq \rho\alpha$, $\mathrm{D}_\alpha\left(\mathbb{P} \| \mathbb{Q}\right) := \frac{1}{\alpha - 1} \log \int \left(\frac{d\mathbb{P}}{d\mathbb{Q}}\right)^{\alpha - 1} d\mathbb{Q}$
- **Composition :** If $A_1, \dots, A_k$ are ρ-zCDP, $(A_1, \dots, A_k)$ is $k\rho$-zCDP.
- **Gaussian mechanism :** It is possible to privatize queries by adding Gaussian noise.



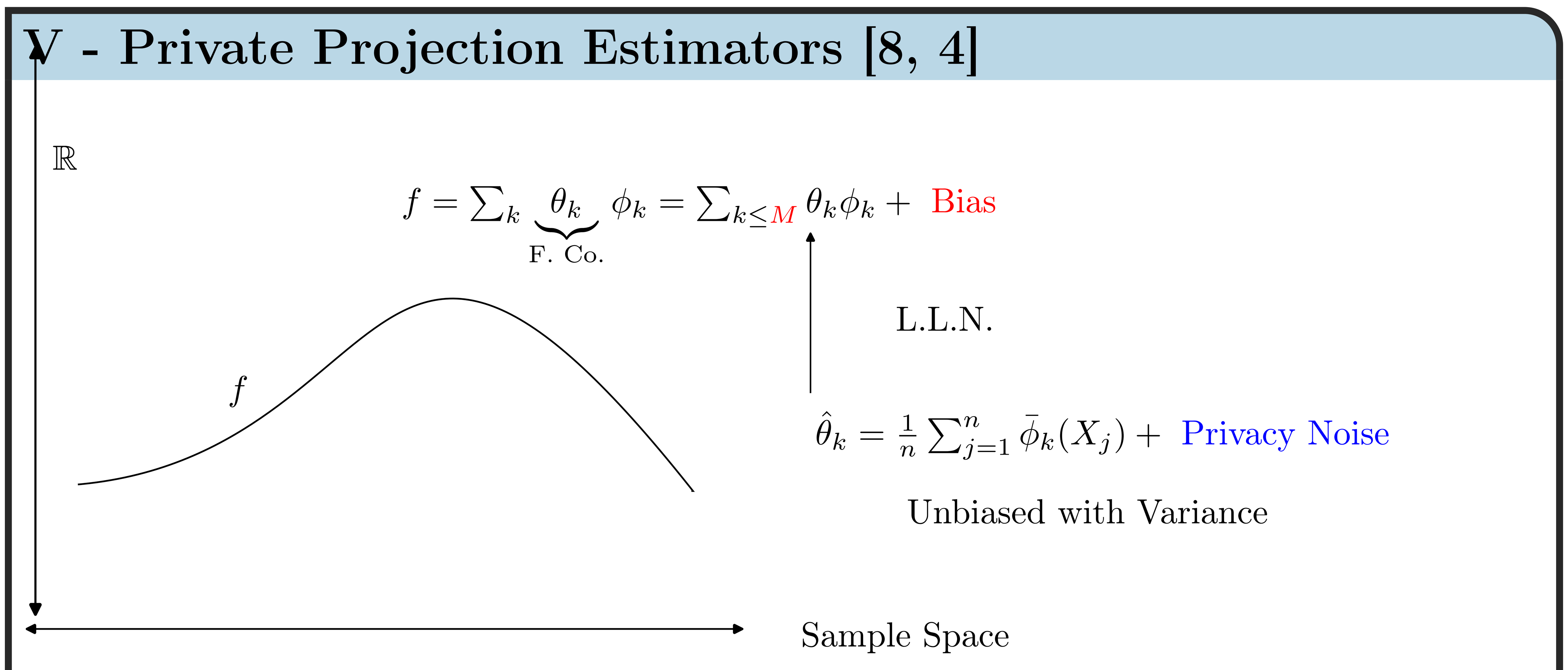
III - Contributions [4, 5, 6]

- **Adaptive :** Optimal even without prior knowledge on the target smoothness.
- **Non-integer smoothness :** Allows for finer-grained modeling than previous work.
- **High(ish) dimension :** Generalizes to arbitrary dimension (but suffers from the curse of dimensionality).

IV - β-Smoothness

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

V - Private Projection Estimators [8, 4]



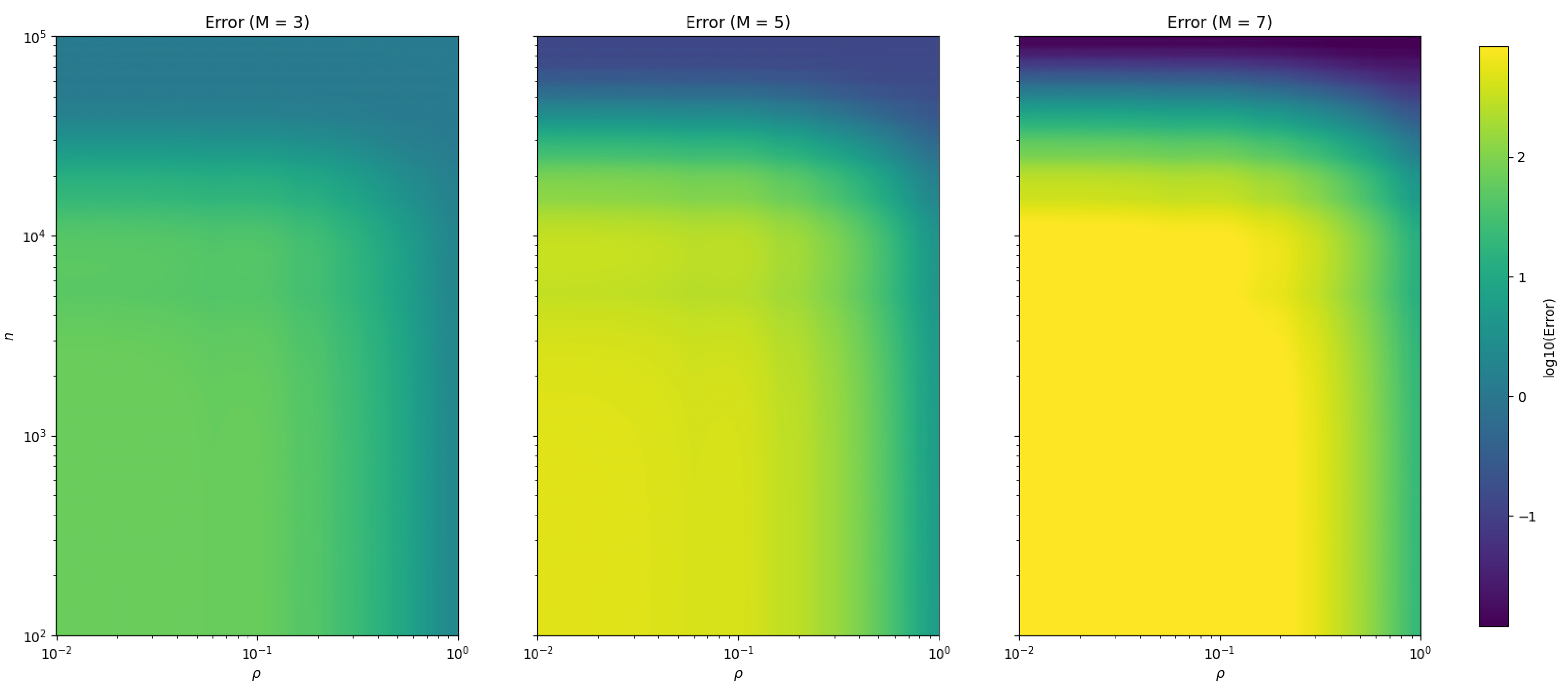
VI - Minimax Rate & Adaptivity

The minimax rate of extimation for the problem is

$$\Theta\left(\max\left\{n^{-\frac{2\beta}{2\beta + d}}, (n\sqrt{\rho})^{-\frac{2\beta}{\beta + d}}\right\}\right) .$$

- **The privacy parameter** ρ : $\rho \gtrsim n^{-\frac{2\beta}{2\beta + d}}$, privacy comes at a negligible cost on the estimation. $\rho \ll n^{-\frac{2\beta}{2\beta + d}}$ the utility can be arbitrarily degraded by making ρ arbitrarily small.
- **The smoothness** β : The higher β, the smaller the cut-off rate $n^{-\frac{2\beta}{2\beta + d}}$.
- **The dimensionality** d : Relative curse of dimensionality (can be balanced by smoothness).
- **Adaptivity** : It is possible to make the estimation adaptive (i.e. without prior knowledge of β) by adapting Lepskii's method [7].

VII - Experimental Results



VIII - References

[1] Cynthia Dwork and Guy Rothblum : *Concentrated differential privacy* (2016)

[2] Cynthia Dwork, Frank McSherry, Kobbi Nissim and Adam D. Smith : *Calibrating Noise to Sensitivity in Private Data Analysis*, TCC (2006)

[3] Mark Bun and Thomas Steinke : *Concentrated differential privacy: Simplifications, extensions, and lower bounds*, Theory of Cryptography (2016)

[4] Larry Wasserman and Shuheng Zhou : *A statistical framework for differential privacy*, Journal of the Ameri- can Statistical Association (2010)

[5] Rina Foygel Barber and John C. Duchi : *Privacy and statistical risk: Formalisms and minimax bounds* (2014)

[6] Lalanne Clément, Aurélien Garivier and Rémi Gribonval : *About the cost of central privacy in density estimation*, Transactions on Machine Learning Research (2023)

[7] Oleg V. Lepskii : *On a problem of adaptive estimation in gaussian white noise*, Theory of Probability and Its Applications (1991)

[8] Alexandre B. Tsybakov : *Introduction to Nonparametric Estimation*, Springer series in statistic (2009)