Minimax density estimation in the adversarial framework under local differential privacy

Mélisande Albert, Juliette Chevallier, Béatrice Laurent, Ousmane Sacko

Minimax density estimation in the adversarial framework under local differential privacy

Mélisande Albert[1], Juliette Chevallier[1], Béatrice Laurent[1], and Ousmane Sacko[2]

[1]Institut de Mathématiques de Toulouse; UMR 5219, Université de Toulouse; CNRS, INSA
[2]MODAL'X; UMR 9023, Université Paris Nanterre

March 26, 2024

We consider the problem of nonparametric density estimation under privacy constraints in an adversarial framework. To this end, we study minimax rates under local differential privacy over Sobolev spaces. We first obtain a lower bound which allows us to quantify the impact of privacy compared with the classical framework. Next, we introduce a new Coordinate block privacy mechanism that guarantees local differential privacy, which, coupled with a projection estimator, achieves the minimax optimal rates.

Keywords: Nonparametric density estimation, local differential privacy, minimax optimality, adversarial loss, Sobolev balls

1 Introduction

In this paper, we study minimax nonparametric density estimation under local differential privacy constraints, in an adversarial framework. Let us first present each context.

Minimax nonparametric density estimation. Let $X_1, \dots, X_n$ be n independent and identically distributed (i.i.d.) random variables with common density f on $[0,1]^d$ with respect to the Lebesgue measure. We aim at estimating the underlying density f, a classical challenge in statistics extensively studied in the literature. Early attempts to address this problem primarily rely on kernel methods, as in [Rosenblatt, 1956, Parzen, 1962, Silverman, 1978]. Projection methods also emerge as a viable approach, built upon the decomposition of f into orthonormal bases. Notable contributions include [Schwartz, 1967, Kronmal and Tarter, 1968, Walter, 1977, Efromovich, 1999]. We want to construct optimal estimators in the minimax sense. More precisely, consider a loss function ℓ and a regularity space $\mathcal{F}$. We say that an estimator $\hat{f}$ is optimal in the minimax sense over the class $\mathcal{F}$ if there exists a positive sequence $(\rho_n)_n$ and positive constants c and C such that

$$\sup_{f\in\mathcal{F}} \mathbb{E}\left[\ell(\hat{f}, f)\right] \leq C\rho_n, \qquad \text{and} \qquad \inf_{\tilde{f}} \sup_{f\in\mathcal{F}} \mathbb{E}\left[\ell(\tilde{f}, f)\right] \geq c\rho_n,$$

where the infimum is taken over all estimators $\tilde{f}$ of f based on $(X_1, \dots, X_n)$. Such sequence $(\rho_n)_{n\geq 1}$ is called the minimax rate over the class $\mathcal{F}$. The most commonly used loss functions

are defined from $\mathbb{L}_p$ norms. The literature on nonparametric density estimation is abundant. We can refer to [Tsybakov, 2009] for a review of optimal results in different contexts. This topic remains widely studied in more intricate contexts such as deconvolution, non independent observations, and, more recently, data privacy constraints as considered in this paper.

Differential privacy. The collection of a large amount of personal data requires new tools to protect sensitive information concerning individuals. To guarantee the privacy of each individual, the development of mechanisms to be applied to daba bases has become crucial. Differential privacy has been widely adopted, since the seminal papers by [Dwork et al., 2006b, Dwork et al., 2006a], as it provides rigorous privacy guarantees. The original definition corresponds to the notion of *global differential privacy*, where the entire original data set $(X_1, \dots, X_n)$ is privatized into an output that preserves the privacy of the n individuals and is used for further statistical analyses. This means that the n data holders share confidence with a common curator who has access to the whole sample $(X_1, \dots, X_n)$. In this paper, we consider a stronger notion of privacy, that is called *local differential privacy*, where each individual generates a private view Z_i of its original data X_i independently of the other individuals. More precisely, Z_i is a stochastic transformation of X_i by the channel Q_i : given $X_i = x_i$, $Z_i \sim Q_i(.|X_i = x_i)$. For a given positive number α, we say that the sequence of channels $(Q_i)_{i=1,\dots,n}$ provides α-local differentially private (α-LDP) views of $(X_1, \dots, X_n)$ if

$$\forall\, 1 \le i \le n, \qquad \sup_{B \in \mathcal{B}, (x_i, x_i') \in \mathcal{X}} \frac{Q_i(B|X_i = x_i)}{Q_i(B|X_i = x_i')} \le e^{\alpha}, \tag{1}$$

where $(\mathcal{X}, \mathcal{A})$ and $(\mathcal{Z}, \mathcal{B})$ respectively denote the measure spaces of the original data and of the privatized data. Let us briefly mention the interactive local privacy mechanism where Z_i may depend on the preceding privatized values $Z_1, \dots, Z_{i-1}$, namely $Z_i \sim Q_i(.|x_i, Z_1, \dots, Z_{i-1})$. We focus on non interactive local privacy mechanisms satisfying Equation (1).
In statistical inference under privacy constraints, we do not have access to the original data. Therefore, the statistical performances are necessarily deteriorated. In order to quantify this gap, the literature is prolific since the early work by [Wasserman and Zhou, 2010, Smith, 2011]. The first minimax rates for estimation problems under differential privacy conditions where established in [Duchi et al., 2013b, Duchi et al., 2013a, Duchi et al., 2018] and more recently [Cai et al., 2021]. For instance, [Butucea et al., 2020] estimate density over Besov ellipsoids with respect to the $\mathbb{L}_p$ norm, [Rohde and Steinberger, 2020] deal with the estimation of functionals. [Butucea et al., 2023b] study the effect of interactive versus non interactive privacy mechanisms on the estimation of quadratic functionals, while [Lam-Weil et al., 2022, Dubois et al., 2023] consider goodness-of-fit testing problems.
In this paper, we are interested in minimax density estimation under local differential privacy defined by Equation (1), with respect to adversarial losses. Let us now introduce the adversarial framework.

Adversarial framework. Beyond density estimation, generative models have seen considerable growth in recent years. Generative models aim to reproduce the sampling behavior of a target distribution, rather than explicitly fitting a density function. Specifically, machine learning has made significant empirical progress in generative modeling, using such tools as generative adversarial networks (GANs) [Goodfellow et al., 2014]. GANs provide a flexible framework for sampling from unknown distributions, and have become a standard

tool among practitioners. From a practical point of view, numerous improvements have been made to GANs, enabling them to achieve state-of-the-art performance in various data generation tasks. We refer to [Gui et al., 2020] for a review and all references therein. The empirical success of GANs motivated many researchers to analyze their theoretical properties; among others, we mention the work of [Biau et al., 2020, Liang, 2021, Schreuder et al., 2021, Asatryan et al., 2023, Puchkin et al., 2024].

The rise of GANs has brought with it a focus on losses that can drive generative models, such as integral probability metrics (IPMs) [Müller, 1997, Dziugaite et al., 2015, Liu et al., 2017, Bottou et al., 2018]. IPMs are types of distance between probability distributions, defined by the ability of a class of functions to distinguish between two distributions. More precisely, given a class $\mathcal{G}$ of real-valued functions defined on $[0,1]^d$, we define the distance between two probability density functions f_1 and f_2 by

$$d_{\mathcal{G}}(f_1, f_2) = \sup_{g \in \mathcal{G}} \int_{[0,1]^d} (f_1 - f_2) g. \tag{2}$$

By choosing different sets $\mathcal{G}$, this framework can express a multitude of commonly used measurements. We refer to it as the *adversarial framework*. The name adversarial comes from the idea that $\mathcal{G}$ can be viewed as a *discriminator* able to distinguish between the two probability measures carried by f_1 and f_2. For example, choosing the 1-Lipschitz functions equipped with the Wasserstein metric leads to the Wasserstein GANs [Arjovsky et al., 2017], whereas choosing $\mathcal{G}$ as a Sobolev space leads to Sobolev GANs [Mroueh et al., 2018].

In a pre-published version from 2017, [Liang, 2021] was the first to formalize nonparametric estimation under the adversarial framework, and to prove upper bounds for Sobolev GANs. Following this work, [Singh et al., 2018] succeeded in improving this upper bound. Finally, [Liang, 2021] obtained an optimal minimax rate for Sobolev GANs. The study of the adversarial framework is today a very active field of research, with variations on the choice of regularity classes, or by generalizing the choice of metric. To name a few, [Singh and Póczos, 2018, Bai et al., 2019, Weed and Berthet, 2019, Chen et al., 2022] focused on Wasserstein distance; [Uppal et al., 2019] on Besov spaces; [Luise et al., 2020] on optimal transport-based loss functions. In this work, we focus on private density estimation using adversarial losses as defined in Equation (2), where the discriminator $\mathcal{G}$ is a Sobolev ball. This can be seen as a first step towards understanding the performances of Sobolev GANs under privacy.

Minimax rate in the adversarial framework under privacy constraints In this paper, our aim is to estimate the density of $X_1, \dots, X_n$ under privacy constraints. Hence, the definition of the minimax rate needs to take into account the choice of the privacy mechanism. More precisely, let $\mathcal{Q}_\alpha$ denote the set of all α-local differential private mechanisms, that is the set of mechanisms $Q = (Q_1, \dots, Q_n)$ satisfying Equation (1). The minimax rate over the regularity set $\mathcal{F}$ with respect to the adversarial loss $d_{\mathcal{G}}$ under α-local differential privacy is defined by

$$\rho_n(\mathcal{F}, \mathcal{G}, \mathcal{Q}_\alpha) = \inf_{Q \in \mathcal{Q}_\alpha} \inf_{\tilde{f}} \sup_{f \in \mathcal{F}} \mathbb{E}\left[d_{\mathcal{G}}\left(\tilde{f}, f\right)\right], \tag{3}$$

where the infimum over $\tilde{f}$ is taken among all estimators of f based on privatized data $(Z_1, \dots, Z_n)$ obtained from Q. Consider an α-LDP mechanism Q and an estimator $\hat{f}$ based

on data privatized from Q. We say that the couple $(Q, \hat{f})$ is *minimax optimal* if there exists a constant C (w.r.t. n) such that

$$\sup_{f\in\mathcal{F}} \mathbb{E}\left[d_{\mathcal{G}}(\hat{f}, f)\right] \leq C\rho_n(\mathcal{F}, \mathcal{G}, \mathcal{Q}_\alpha).$$

In this paper, we consider Sobolev spaces as regularity sets $\mathcal{F}$ for the density and $\mathcal{G}$ for the discriminator in the adversarial loss. One of the main contributions of this work is the introduction of the *Coordinate block privacy mechanism* which allows to achieve the minimax optimal rates with projection estimators on the Fourier basis of $\mathbb{L}_2([0,1]^d)$.

The structure of this paper is as follows. Section 2 consists of all the results in the isotropic case. The definition of the Sobolev balls is detailed in Section 2.1. Section 2.2 is devoted to the lower bound of the minimax rate, whereas Section 2.3 presents the Coordinate block privacy mechanism together with our private projection estimator which are minimax optimal. In Section 3, we present results in an anisotropic framework. In Section 4, we demonstrate the main results. The proofs are detailed in the isotropic case to avoid making the notations too cumbersome. Complementary proofs can be found in Appendix A.

In the following, $c_{a,b,\ldots}$ and $C_{a,b,\ldots}$ denote constants depending only on $a, b, \ldots$ that may vary from line to line. Moreover, the notation $u_n \asymp v_n$ indicates that there exist absolute constants c and C such that $c \leq u_n/v_n \leq C$.

2 Results over isotropic Sobolev balls

Before stating our main results, let us define the regularity spaces considered in this section, namely isotropic Sobolev balls.

2.1 Fourier basis and Sobolev spaces

Consider the d-dimensional Fourier basis of $\mathbb{L}_2([0,1]^d)$ obtained by taking the tensor product of the one-dimensional Fourier basis. More precisely, recall the one-dimensional Fourier basis defined for all t in $[0,1]$ by

$$\varphi_1(t) = 1 \qquad \text{and} \qquad \forall j \in \mathbb{N}^*, \quad \begin{cases} \varphi_{2j}(t) &= \sqrt{2}\cos(2\pi jt) \\ \varphi_{2j+1}(t) &= \sqrt{2}\sin(2\pi jt). \end{cases}$$

The d-dimensional Fourier basis is defined for all $\underline{\boldsymbol{j}} = (j_1, \ldots, j_d)$ in $(\mathbb{N}^*)^d$ and all $x = (x_1, \ldots, x_d)$ in $[0,1]^d$ by

$$\varphi_{\underline{\boldsymbol{j}}}(x) = \prod_{m=1}^{d} \varphi_{j_m}(x_m).$$

Then, the family $\left\{\varphi_{\underline{\boldsymbol{j}}}, \underline{\boldsymbol{j}} \in (\mathbb{N}^*)^d\right\}$ is an orthonormal basis of $\mathbb{L}_2([0,1]^d)$. In particular, any function f in $\mathbb{L}_2([0,1]^d)$ can be uniquely decomposed as

$$f = \sum_{\underline{\boldsymbol{j}}\in(\mathbb{N}^*)^d} \theta_{\underline{\boldsymbol{j}}}(f)\varphi_{\underline{\boldsymbol{j}}}, \qquad \text{where} \qquad \theta_{\underline{\boldsymbol{j}}}(f) = \int_{[0,1]^d} f(x)\varphi_{\underline{\boldsymbol{j}}}(x)dx.$$

A natural choice for regularity spaces are Sobolev balls since they can be simply characterized in terms of the Fourier decomposition. The isotropic Sobolev balls are defined for all smoothness parameter $\beta > 0$ and radius $R > 0$ by

$$\mathcal{W}^\beta(R) = \left\{ f \in \mathbb{L}_2([0,1]^d) \, ; \sum_{\underline{\boldsymbol{j}} \in (\mathbb{N}^*)^d} \left(j_1^{2\beta} + \ldots + j_d^{2\beta} \right) \theta^2_{\underline{\boldsymbol{j}}}(f) \leq R^2 \right\}. \tag{4}$$

Note that if f is a density, then

$$\sum_{\underline{\boldsymbol{j}} \in (\mathbb{N}^*)^d} \left(j_1^{2\beta} + \ldots + j_d^{2\beta} \right) \theta^2_{\underline{\boldsymbol{j}}}(f) \geq d \; \theta^2_{(1,1,\ldots,1)}(f) = d.$$

Hence the Sobolev ball $\mathcal{W}^\beta(R)$ may contain densities only if $R^2 \geq d$. In this section, both densities and discriminant functions will belong to such isotropic Sobolev balls.

2.2 Lower bound over Sobolev balls

In this section, we bound from below the minimax rate over Sobolev balls in the adversarial framework under privacy constraints.

Theorem 2.1. *Consider a sample size n, a privacy level α in $(0, A]$ such that $n\alpha^2 \geq 1$, integer smoothness parameters β and δ in $\mathbb{N}^*$, and a positive radius R such that $R^2 > d$. Then, the minimax rate defined in (3) is lower bounded as follows:*

$$\rho_n(\mathcal{W}^\beta(R), \mathcal{W}^\delta(1), \mathcal{Q}_\alpha) \geq c \max\left\{ (n\alpha^2)^{-\frac{\beta+\delta}{2\beta+2d}}, (n\alpha^2)^{-1/2} \right\},$$

where $c = c_{\beta,\delta,R,d,A}$ is a positive constant depending on β, δ, R, d, and A.

The proof of Theorem 2.1 is detailed in Section 4.1. Depending on the value of the smoothness parameter δ for the adversarial distance, the rate is not the same. In particular, if $\delta > d$, we obtain a near-parametric rate, while for $\delta < d$, we see the influence of the regularity of the discriminator class $\mathcal{G} = \mathcal{W}^\delta(1)$. Note that these rates are coherent with the ones achieved by [Liang, 2021] over slightly different Sobolev spaces, with respect to adversarial losses in dimension d, without privacy constraints, that are $n^{-\frac{\beta+\delta}{2\beta+d}} \vee n^{-1/2}$. One may notice the effect of privatization, that replaces n by $n\alpha^2$ and transforms the dimension d in $2d$. These effects can also be observed in [Duchi et al., 2013a], over Sobolev spaces in dimension 1, with respect to the $\mathbb{L}_2$ loss, where the rate $n^{-2\beta/(2\beta+1)}$ without confidentiality becomes $(n\alpha^2)^{-2\beta/(2\beta+2)}$ under privacy constraints.
The lower bound is proved only for integer regularity parameters β and δ which is standard in the literature. Indeed, the proof is based on the construction of parametric spaces, namely $\mathcal{F}^\beta$ and $\mathcal{D}^\delta$, which relies on the characterisation of Sobolev spaces based on the $\mathbb{L}_2$ norm of the partial derivatives (see Lemma A.1).

2.3 Optimal local differential private projection estimator

In the usual *non private* setting, projection estimators have been widely studied. Over isotropic Sobolev balls $\mathcal{W}^\beta(R)$, as considered in this section, where the smoothness parameter β is the same for each dimension, the density is approximated by a truncated sum

$\sum_{\underline{\boldsymbol{j}}\in\{1,2,\dots,J\}^d}\theta_{\underline{\boldsymbol{j}}}(f)\varphi_{\underline{\boldsymbol{j}}}$, where each index j_m in $\underline{\boldsymbol{j}}$ is considered up to the same integer J that needs to be calibrated. If we had access to the original data set $(X_1,\dots,X_n)$, we could then estimate the density f by

$$\hat{f}_J = \sum_{\underline{\boldsymbol{j}}\in\{1,2,\dots,J\}^d} \hat{\theta}_{\underline{\boldsymbol{j}}}\varphi_{\underline{\boldsymbol{j}}}, \qquad \text{where} \qquad \forall \underline{\boldsymbol{j}} \in \{1,2,\dots,J\}^d,\ \hat{\theta}_{\underline{\boldsymbol{j}}} = \frac{1}{n}\sum_{i=1}^n \varphi_{\underline{\boldsymbol{j}}}(X_i),$$

for some J well chosen. Yet, under privacy constraints, we are not allowed to use the original dataset $(X_i)_{1\le i\le n}$, only a privatized version of the data. Let us first define a mechanism that provides α-local differential private data and second, introduce the private projection estimator and prove that it achieves optimal rates.

Privacy mechanism. This mechanism is strongly inspired by the one proposed by [Duchi et al., 2018] and adapted by [Butucea et al., 2023a]. The main difference is that, instead of privatizing all the coefficients at once, as done in the *Coordinate global privacy mechanism* in [Butucea et al., 2023a], we privatize independently blocks of coefficients, which we refer to as the *Coordinate block privacy mechanism*.

Let us first introduce dyadic coordinate blocks. Define for all $\underline{\boldsymbol{\ell}} = (\ell_1,\dots,\ell_d) \in \mathbb{N}^d$,

$$\mathcal{J}_{\underline{\boldsymbol{\ell}}} = \prod_{m=1}^d \left\{2^{\ell_m},\dots,2^{\ell_m+1}-1\right\},$$

such that $\left\{\mathcal{J}_{\underline{\boldsymbol{\ell}}}\right\}_{\underline{\boldsymbol{\ell}}\in\{0,\dots,L\}^d}$ form a partition of $\{1,2,\dots,J\}^d$, where $J = 2^{L+1}-1$. Denote $d_{\underline{\boldsymbol{\ell}}}$ the cardinal of $\mathcal{J}_{\underline{\boldsymbol{\ell}}}$, that is

$$d_{\underline{\boldsymbol{\ell}}} = |\mathcal{J}_{\underline{\boldsymbol{\ell}}}| = \prod_{m=1}^d 2^{\ell_m}.$$

To each block $\mathcal{J}_{\underline{\boldsymbol{\ell}}}$, associate a privacy level $\alpha_{\underline{\boldsymbol{\ell}}}$ that will be calibrated later, and such that $\sum_{\underline{\boldsymbol{\ell}}\in\{0,1,\dots,L\}^d}\alpha_{\underline{\boldsymbol{\ell}}} = \alpha$.
For all i in $\{1,2,\dots,n\}$, we define the private version Z_i of X_i using the following steps.

- Compute for all $\underline{\boldsymbol{\ell}}$ in $\{0,1,\dots,L\}^d$, the coefficients in the block $\mathcal{J}_{\underline{\boldsymbol{\ell}}}$,

$$v_{i,[\underline{\boldsymbol{\ell}}]} = \left(\varphi_{\underline{\boldsymbol{j}}}(X_i)\right)_{\underline{\boldsymbol{j}}\in\mathcal{J}_{\underline{\boldsymbol{\ell}}}} \quad \in [-B_0,B_0]^{d_{\underline{\boldsymbol{\ell}}}},$$

  where $B_0 = \max_{\underline{\boldsymbol{j}}\in(\mathbb{N}^*)^d}\left\|\varphi_{\underline{\boldsymbol{j}}}\right\|_\infty = \sqrt{2}$.

- Independently for all $\underline{\boldsymbol{\ell}}$, draw random vectors $\tilde{v}_{i,[\underline{\boldsymbol{\ell}}]}$ in $\{-B_0,B_0\}^{d_{\underline{\boldsymbol{\ell}}}}$ with coordinates

$$\forall \underline{\boldsymbol{j}} \in \mathcal{J}_{\underline{\boldsymbol{\ell}}}, \quad \tilde{v}_{i,\underline{\boldsymbol{j}}} = \begin{cases} B_0 & \text{with probability } \frac{1}{2} + \frac{\varphi_{\underline{\boldsymbol{j}}}(X_i)}{2B_0} \\ -B_0 & \text{otherwise,} \end{cases}$$

  and set

$$\begin{cases} \mathcal{D}_+\left(\tilde{v}_{i,[\underline{\boldsymbol{\ell}}]}\right) = \left\{z\in\{\pm B_{d_{\underline{\boldsymbol{\ell}}}}(\alpha_{\underline{\boldsymbol{\ell}}})\}^{d_{\underline{\boldsymbol{\ell}}}}\ ;\ \langle z,\tilde{v}_{i,[\underline{\boldsymbol{\ell}}]}\rangle > 0 \text{ or } \left(\langle z,\tilde{v}_{i,[\underline{\boldsymbol{\ell}}]}\rangle = 0 \text{ and } z_1 = \frac{B_{d_{\underline{\boldsymbol{\ell}}}}(\alpha_{\underline{\boldsymbol{\ell}}})}{B_0}\tilde{v}_{i,1}\right)\right\} \\ \mathcal{D}_-\left(\tilde{v}_{i,[\underline{\boldsymbol{\ell}}]}\right) = \left\{z\in\{\pm B_{d_{\underline{\boldsymbol{\ell}}}}(\alpha_{\underline{\boldsymbol{\ell}}})\}^{d_{\underline{\boldsymbol{\ell}}}}\ ;\ \langle z,\tilde{v}_{i,[\underline{\boldsymbol{\ell}}]}\rangle < 0 \text{ or } \left(\langle z,\tilde{v}_{i,[\underline{\boldsymbol{\ell}}]}\rangle = 0 \text{ and } z_1 = -\frac{B_{d_{\underline{\boldsymbol{\ell}}}}(\alpha_{\underline{\boldsymbol{\ell}}})}{B_0}\tilde{v}_{i,1}\right)\right\}, \end{cases}$$

where for all k in $\mathbb{N}^*$ and positive a,

$$B_k(a) = B_0 \frac{e^a+1}{e^a-1} C_k, \qquad \text{with} \qquad \frac{1}{C_k} = \begin{cases} \dfrac{1}{2^{k-1}} \binom{k-1}{\frac{k-1}{2}} & \text{if } k \text{ is odd,} \\ \dfrac{(k-2)!(k-2)}{2^{k-1}\left(\frac{k}{2}-1\right)!\frac{k}{2}!} & \text{if } k \text{ is even.} \end{cases} \tag{5}$$

- Independently for all $\underline{\boldsymbol{\ell}}$ in $\{0,1,\dots,L\}^d$, draw at random

$$T_{i,\underline{\boldsymbol{\ell}}} \sim \mathcal{B}(\pi_{\alpha_{\underline{\boldsymbol{\ell}}}}), \quad \text{where} \quad \pi_{\alpha_{\underline{\boldsymbol{\ell}}}} = \frac{e^{\alpha_{\underline{\boldsymbol{\ell}}}}}{1+e^{\alpha_{\underline{\boldsymbol{\ell}}}}},$$

and generate

$$\tilde{Z}_{i,[\underline{\boldsymbol{\ell}}]} \sim \begin{cases} \mathcal{U}\left(\mathcal{D}_+\left(\tilde{v}_{i,[\underline{\boldsymbol{\ell}}]}\right)\right) & \text{if } T_{i,\underline{\boldsymbol{\ell}}=1} \\ \mathcal{U}\left(\mathcal{D}_-\left(\tilde{v}_{i,[\underline{\boldsymbol{\ell}}]}\right)\right) & \text{if } T_{i,\underline{\boldsymbol{\ell}}=0}. \end{cases}$$

- Define the vector $Z_i = \left(Z_{i,[\underline{\boldsymbol{\ell}}]}\right)_{\underline{\boldsymbol{\ell}} \in \{0,1,\dots,L\}^d}$, where
  - if $d_{\underline{\boldsymbol{\ell}}}$ is odd, $Z_{i,[\underline{\boldsymbol{\ell}}]} = \tilde{Z}_{i,[\underline{\boldsymbol{\ell}}]}$,
  - if $d_{\underline{\boldsymbol{\ell}}}$ is even, the coordinates of $Z_{i,[\underline{\boldsymbol{\ell}}]}$ are

$$Z_{i,\underline{\boldsymbol{j}}} = \begin{cases} \frac{d-2}{2(d-1)!} \times \tilde{Z}_{i,\underline{\boldsymbol{j}}} & \text{if } \underline{\boldsymbol{j}} = (2^{\ell_m})_{1 \le m \le d} \\ \tilde{Z}_{i,\underline{\boldsymbol{j}}} & \text{otherwise .} \end{cases}$$

Note that in $\mathcal{D}_+\left(\tilde{v}_{i,[\underline{\boldsymbol{\ell}}]}\right)$ and $\mathcal{D}_-\left(\tilde{v}_{i,[\underline{\boldsymbol{\ell}}]}\right)$, the second condition is impossible if $d_{\underline{\boldsymbol{\ell}}}$ is odd, i.e., for any vectors $\tilde{z}$ in $\{\pm B_{d_{\underline{\boldsymbol{\ell}}}}(\alpha_{\underline{\boldsymbol{\ell}}})\}^{d_{\underline{\boldsymbol{\ell}}}}$ and $\tilde{v}$ in $\{\pm B_0\}^{d_{\underline{\boldsymbol{\ell}}}}$, $\langle z, \tilde{v}_{i,[\underline{\boldsymbol{\ell}}]}\rangle = 0$ is impossible. Indeed, since the scalar product is proportional to the difference between the number of coordinates of same sign and the number of coordinates of opposite sign, it can be equal to zero only if the number of coordinates is even. In our case, the only $\underline{\boldsymbol{\ell}}$ for which $\mathcal{J}_{\underline{\boldsymbol{\ell}}}$ is odd is the case $\underline{\boldsymbol{\ell}} = (0,0,\dots,0)$.

Proposition 2.1. *Consider a positive integer $J = 2^{L+1}-1$ for some $L \ge 1$, and a privacy level α in $(0,A]$. The mechanism described above satisfies the following properties.*

1. *$Z_1,\dots,Z_n$ provide α-local differential private views of $X_1,\dots,X_n$.*
2. *For all i in $\{1,2,\dots,n\}$, $\underline{\boldsymbol{\ell}}$ in $\{0,\dots,L\}^d$, and $\underline{\boldsymbol{j}}$ in $\mathcal{J}_{\underline{\boldsymbol{\ell}}}$,*

$$\mathbb{E}\left[Z_{i,\underline{\boldsymbol{j}}} \middle| X_i\right] = \varphi_{\underline{\boldsymbol{j}}}(X_i), \quad \text{and} \quad \mathrm{Var}\left(Z_{i,\underline{\boldsymbol{j}}}\right) \le C_A B_0^2 \frac{d_{\underline{\boldsymbol{\ell}}}}{\alpha_{\underline{\boldsymbol{\ell}}}^2},$$

where C_A is a constant depending on A.

The proof is very similar to the one of [Butucea et al., 2023a] and may be found in Appendix A.3. A first natural idea was to consider the *Coordinate global privacy mechanism* in [Duchi et al., 2018, Butucea et al., 2023a]. Yet, one may prove that the variance of each $Z_{i,\underline{\boldsymbol{j}}}$ is then of order J^d/α^2 and this allows to achieve optimal rates only for $\delta < d/2$.
With this new mechanism, the variance of $Z_{i,\underline{\boldsymbol{j}}}$ is controlled by the size $d_{\underline{\boldsymbol{\ell}}}$ of the corresponding block, and it turns out that it makes it possible to achieve the optimal minimax rate.

Private projection estimator Given the observation of the private views $Z_1, \dots, Z_n$ of $X_1, \dots, X_n$, we aim at estimating the density f of the X_i, $1 \le i \le n$. As done by [Duchi et al., 2013a] in dimension one, we estimate the coefficients $\theta_{\underline{\boldsymbol{j}}}(f)$ in the Fourier expansion of f for all $\underline{\boldsymbol{j}} \in \{1, 2, \dots, J\}^d$ by

$$\hat{\theta}_{\underline{\boldsymbol{j}}} = \frac{1}{n} \sum_{i=1}^{n} Z_{i,\underline{\boldsymbol{j}}}.$$

Proposition 2.1 implies that for all $\underline{\boldsymbol{j}}$, $\hat{\theta}_{\underline{\boldsymbol{j}}}$ is an unbiased estimator of $\theta_{\underline{\boldsymbol{j}}}(f)$. Then, we estimate the density f by

$$\hat{f}_J = \sum_{\underline{\boldsymbol{j}} \in \{1,2,\dots,J\}^d} \hat{\theta}_{\underline{\boldsymbol{j}}} \varphi_{\underline{\boldsymbol{j}}}, \tag{6}$$

for a well calibrated J.

Theorem 2.2. *Consider a sample size n, a privacy level α in $(0, A]$ such that $n\alpha^2 > 1$, smoothness parameters β and δ in $\mathbb{R}_+^*$, and a positive radius R. Consider the Coordinate block privacy mechanism described in Section 2.3, where*

$$\alpha_{\underline{\boldsymbol{\ell}}} = \frac{\alpha}{S_d} \prod_{m=1}^{d} 2^{\ell_m(1-\delta/d)/2} \qquad \text{and} \qquad S_d = \sum_{\underline{\boldsymbol{\ell}'} \in \{0,1,\dots,L\}^d} \prod_{m=1}^{d} 2^{\ell'_m(1-\delta/d)/2}.$$

Then, the projection estimator $\hat{f}_J$ defined in Equation (6) satisfies

$$\sup_{f \in \mathcal{W}^\beta(R)} \mathbb{E}\left[d_{\mathcal{W}^\delta(1)}\left(\hat{f}_J, f\right)\right] \le \begin{cases} C(n\alpha^2)^{-\frac{\beta+\delta}{2\beta+2d}} & \text{if } \delta < d \text{ and } J \asymp (n\alpha^2)^{\frac{1}{2\beta+2d}} \\ C\left(\frac{n\alpha^2}{[\ln(n\alpha^2)]^{4d}}\right)^{-1/2} & \text{if } \delta = d \text{ and } J \asymp \left(\frac{n\alpha^2}{[\ln(n\alpha^2)]^{4d}}\right)^{\frac{1}{2\beta+2\delta}} \\ C(n\alpha^2)^{-1/2} & \text{if } \delta > d \text{ and } J \asymp (n\alpha^2)^{\frac{1}{2\beta+2\delta}}, \end{cases}$$

where $C = C_{d,A,\beta,\delta,R}$ denotes a constant depending on d, A, β, δ and R.

Unlike the case of the lower bound, the result holds for non-integers parameters β and δ. One may see with Theorem 2.1 that the private projection estimator with the Coordinate block privacy mechanism are minimax optimal for all $\delta \neq d$. Moreover, if the discriminator regularity $\delta < d$, then one needs to consider larger privacy levels (and thus less confidentiality) for larger blocks, whereas if $\delta > d$, smaller privacy levels (and thus more confidentiality) need to be taken for larger blocks. One may note that a logarithmic term appears in the limiting case where $\delta = d$. In that case, $\alpha_{\underline{\boldsymbol{\ell}}} = \alpha/(L+1)^d$, and all blocs have the same privacy level. In this case, we do not know if the lower bound or the upper bound are suboptimal.
Finally, as the choice of the number of considered coefficients J depends on the unknown regularity β of the density, this procedure is not adaptive in the minimax sense. This is a topic for future work.

3 Results over anisotropic Sobolev balls

Assume now that the regularity is not the same in each direction. More precisely, consider the multidimensional parameter $\underline{\boldsymbol{\beta}} = (\beta_1, \dots, \beta_d)$ in $(\mathbb{R}_+^*)^d$ and define the *anisotropic* Sobolev ball

$$\mathcal{W}^{\underline{\boldsymbol{\beta}}}(R) = \left\{ f \in \mathbb{L}_2([0,1]^d) \, ; \sum_{\underline{\boldsymbol{j}} \in (\mathbb{N}^*)^d} \left(j_1^{2\beta_1} + \ldots + j_d^{2\beta_d} \right) \theta^2_{\underline{\boldsymbol{j}}}(f) \leq R^2 \right\}. \tag{7}$$

All the results may be easily adapted as soon as one considers discriminators with the same "anisotropy" as the density regularity, that is anisotropic Sobolev balls $\mathcal{W}_{\underline{\boldsymbol{\delta}}}(1)$, where $\underline{\boldsymbol{\delta}} = (\delta_1, \ldots, \delta_d)$ is such that the ratio β_m/δ_m is constant.
More precisely, denote

$$\frac{1}{\beta} = \frac{1}{d} \sum_{m=1}^{d} \frac{1}{\beta_m} \quad \text{and} \quad \frac{1}{\delta} = \frac{1}{d} \sum_{m=1}^{d} \frac{1}{\delta_m}. \tag{8}$$

Then, $\mathcal{W}_{\underline{\boldsymbol{\beta}}}(R)$ and $\mathcal{W}_{\underline{\boldsymbol{\delta}}}(1)$ have the same anisotropy if and only if for all $1 \leq m \leq d$,

$$\frac{\beta}{\beta_m} = \frac{\delta}{\delta_m}. \tag{9}$$

In the anisotropic case, the choice of the number of coordinates considered in each dimension should depend on the corresponding regularity. Hence, for an integer J to be calibrated later, denote the set of multi-indices that are considered

$$\mathcal{J} = \prod_{m=1}^{d} \{1, \ldots, J_m\} \quad \text{where} \quad \forall\, 1 \leq m \leq d,\ J_m = J^{\beta/\beta_m} = J^{\delta/\delta_m}.$$

In particular, the number of elements in $\mathcal{J}$ equals

$$\#\mathcal{J} = \prod_{m=1}^{d} J_m = J^d.$$

Note that in the isotropic case, for all $1 \leq m \leq d$, $\beta_m = \beta$ thus $J_m = J$. Then, one recovers the set of multi-indices $\mathcal{J} = \{1, 2, \ldots, J\}^d$ to which the sum relates in the estimator defined in Equation (6).

Theorem 3.1. *Consider a sample size n, a privacy level α in $(0, A]$ such that $n\alpha^2 \geq 1$, integer smoothness parameters $\underline{\boldsymbol{\beta}} \in (\mathbb{N}^*)^d$ and $\underline{\boldsymbol{\delta}} \in (\mathbb{N}^*)^d$ with same anisotropy as defined in Equation (9), and a positive radius R such that $R^2 > d$. Then, there exists a positive constant $c = c_{\underline{\boldsymbol{\beta}},\underline{\boldsymbol{\delta}},R,d,A}$ such that*

$$\rho_n(\mathcal{W}^{\underline{\boldsymbol{\beta}}}(R), \mathcal{W}^{\underline{\boldsymbol{\delta}}}(1), \mathcal{Q}_\alpha) \geq c \max\left\{ (n\alpha^2)^{-\frac{\beta+\delta}{2\beta+2d}}, (n\alpha^2)^{-1/2} \right\},$$

where β and δ are defined by Equation (8).

For the upper bound, we also need to adapt the privacy mechanism. The only difference comes from the fact that we do not consider the same number of dyadic blocks in each dimension. In this case, define

$$L_m = \lfloor \ln_2 \left(J^{\beta/\beta_m} + 1 \right) \rfloor, \quad \text{and} \quad J_m = 2^{L_m+1} - 1,$$

such that $J^{\beta/\beta_m} \leq J_m \leq 3J^{\beta/\beta_m}$. We then consider the same dyadic blocs $\mathcal{J}_{\underline{\ell}}$ for $\underline{\ell}$ in $\mathcal{L} = \prod_{m=1}^{d} \{0, 1, \dots, L_m\}$ as in the isotropic case such that

$$\mathcal{J} = \bigcup_{\underline{\ell} \in \mathcal{L}} \mathcal{J}_{\underline{\ell}}.$$

The privacy mechanism remains unchanged. Finally, we define the private projection estimator by

$$\hat{f}_J = \sum_{\underline{j} \in \mathcal{J}} \hat{\theta}_{\underline{j}} \varphi_{\underline{j}}, \tag{10}$$

for a well calibrated J.

Theorem 3.2. *Consider a privacy level α in $(0, A]$, smoothness parameters $\underline{\boldsymbol{\beta}} \in (\mathbb{R}_+^*)^d$ and $\underline{\boldsymbol{\delta}} \in (\mathbb{R}_+^*)^d$ with same anisotropy as defined in Equation (9), and a positive radius R. Consider the Coordinate block privacy mechanism described in Section 2.3, where*

$$\alpha_{\underline{\ell}} = \frac{\alpha}{S_d} \prod_{m=1}^{d} 2^{\ell_m(1-\delta_m/d)/2} \qquad \text{and} \qquad S_d = \sum_{\underline{\ell'} \in \mathcal{L}} \prod_{m=1}^{d} 2^{\ell'_m(1-\delta_m/d)/2}.$$

Then, the projection estimator $\hat{f}_J$ defined in Equation (10) satisfies

$$\sup_{f \in \mathcal{W}^{\underline{\boldsymbol{\beta}}}(R)} \mathbb{E}\left[d_{\mathcal{W}^{\underline{\boldsymbol{\delta}}}(1)}\left(\hat{f}_J, f\right)\right] \leq \begin{cases} C(n\alpha^2)^{-\frac{\beta+\delta}{2\beta+2d}} & \text{if } \forall\ m,\ \delta_m < d \text{ and } J \asymp (n\alpha^2)^{\frac{1}{2\beta+2d}} \\ C\left(\frac{n\alpha^2}{\ln(n\alpha^2)^{4d}}\right)^{-1/2} & \text{if } \forall\ m,\ \delta_m = d \text{ and } J \asymp \left(\frac{n\alpha^2}{\ln(n\alpha^2)^{4d}}\right)^{\frac{1}{2\beta+2\delta}} \\ C(n\alpha^2)^{-1/2} & \text{if } \forall\ m,\ \delta_m > d \text{ and } J \asymp (n\alpha^2)^{\frac{1}{2\beta+2\delta}}, \end{cases}$$

where β and δ are defined in Equation (8), and where $C = C_{d,A,\underline{\boldsymbol{\beta}},\underline{\boldsymbol{\delta}},R}$ denotes a constant depending on d, A, $\underline{\boldsymbol{\beta}}$, $\underline{\boldsymbol{\delta}}$ and R.

One may notice that the upper bounds are obtained only for all δ_m on the same regime (that is less than, equal to or greater than d). This condition is due to technical reasons, and appears naturally in the proof. Without this condition, the minimax optimality remains an open question. However, given the density smoothness $\underline{\boldsymbol{\beta}}$, one can choose $\underline{\boldsymbol{\delta}}$ with same anisotropy as $\underline{\boldsymbol{\beta}}$ and such that this condition holds. Moreover, as in the isotropic case, the private projection estimator with the Coordinate block privacy mechanism are minimax optimal for all δ_m less than d, or all δ_m greater than d.

The adaptation of the proofs of Theorems 2.1 and 2.2 to the anisotropic case with the adjustments defined above is straightforward, and leads to Theorems 3.1 and 3.2. For the sake of clarity, we decide to present the proofs only in the isotropic case.

4 Main proofs

4.1 Proof of Theorem 2.1

Let $n \geq 1$ and α in $[1/\sqrt{n}, A]$. Let β and δ be two positive integer smoothness parameters, and consider a radius such that $R^2 > d$. Let $J \geq 1$ be a positive integer. We define ψ on $[0, 1]$ by

$$\psi(t) = \exp\left(\frac{-1}{1-(4t-1)^2}\right) \mathbb{1}_{t \in (0,1/2)} - \exp\left(\frac{-1}{1-(4t-3)^2}\right) \mathbb{1}_{t \in (1/2,1)}.$$

Note that the function ψ is based on bump functions and often used to prove lower bounds in nonparametric statistics. In particular, ψ is a periodic function on $[0,1]$ with continuous derivatives of all orders, such that all derivatives are uniformly bounded on $[0,1]$ and periodic, and satisfies $\int_{[0,1]} \psi(x)dx = 0$.
Denote for all $\underline{\boldsymbol{j}} = (j_1, \ldots, j_d)$ in $\{1, 2, \ldots, J\}^d$, and all $x = (x_1, \ldots, x_d)$ in $[0,1]^d$:

$$G_{\underline{\boldsymbol{j}}}(x) = \prod_{m=1}^{d} \psi\left(J\left(x_m - \frac{j_m - 1}{J}\right)\right).$$

Note that the support of $G_{\underline{\boldsymbol{j}}}$ is $\prod_{m=1}^{d}\left[\frac{j_m-1}{J}, \frac{j_m}{J}\right]$, and for all $p > 0$,

$$\int_{[0,1]^d} G_{\underline{\boldsymbol{j}}} = 0, \qquad \text{and} \qquad \left\|G_{\underline{\boldsymbol{j}}}\right\|_p^p = \frac{\left(\|\psi\|_p^p\right)^d}{J^d}. \tag{11}$$

Indeed,

$$\begin{aligned}\left\|G_{\underline{\boldsymbol{j}}}\right\|_p^p &= \int_{[0,1]^d} \left|G_{\underline{\boldsymbol{j}}}(x)\right|^p dx &&= \prod_{m=1}^{d}\left[\int_{\frac{j_m-1}{J}}^{\frac{j_m}{J}} \left|\psi\left(J\left(x_m - \frac{j_m-1}{J}\right)\right)\right|^p dx_m\right] \\ &&&= \prod_{m=1}^{d}\left[\frac{1}{J}\int_0^1 |\psi(y_m)|^p\, dy_m\right] = \frac{(\|\psi\|_p^p)^d}{J^d}.\end{aligned}$$

Let $\gamma_n, \eta > 0$ and define two parametrized families of functions on $[0,1]^d$ by

$$\mathcal{F}^{\beta}(\gamma_n) = \left\{ f_\nu = \mathbb{1}_{[0,1]^d} + \frac{\gamma_n}{J^\beta} \sum_{\underline{\boldsymbol{j}} \in \{1,2,\ldots,J\}^d} \nu_{\underline{\boldsymbol{j}}} G_{\underline{\boldsymbol{j}}};\ \nu \in \{0,1\}^{J^d} \right\},$$

and

$$\mathcal{D}^{\delta}(\eta) = \left\{ g_\lambda = \frac{\eta}{J^\delta} \sum_{\underline{\boldsymbol{j}} \in \{1,2,\ldots,J\}^d} \lambda_{\underline{\boldsymbol{j}}} G_{\underline{\boldsymbol{j}}};\ \lambda \in \{-1,1\}^{J^d} \right\}.$$

Lemma 4.1. *Let β and δ be two positive integer parameters. Assume $R^2 > d$.*

1. *All functions f_ν in $\mathcal{F}^{\beta}(\gamma_n)$ are densities as soon as $\gamma_n \leq \|\psi\|_\infty^{-d}$.*
2. *One has the inclusions $\mathcal{F}^{\beta}(\gamma_n) \subset \mathcal{W}^{\beta}(R)$ and $\mathcal{D}^{\delta} \subset \mathcal{W}^{\delta}(1)$, as soon as*

$$\gamma_n^2 \leq \frac{R^2 - d}{d\,\|\psi\|_2^{2(d-1)}\left[\|\psi\|_2^2 + \left\|\psi^{(\beta)}\right\|_2^2\right]}, \qquad \text{and} \qquad \eta^2 \leq \frac{1}{d\,\|\psi\|_2^{2(d-1)}\left[\|\psi\|_2^2 + \left\|\psi^{(\delta)}\right\|_2^2\right]}. \tag{12}$$

The proof of Lemma 4.1 is detailed in Section A.2. In the following, consider γ_n and η such that $\gamma_n \leq \|\psi\|_\infty^{-d}$, and that (12) is satisfied.

In particular, for all α-LDP privacy mechanisms Q and all estimators $\tilde{f}$ of f based on data privatized from Q,

$$\begin{aligned}\sup_{f\in\mathcal{W}^{\beta}(R)}\mathbb{E}_{f,Q}\left[d_{\mathcal{W}^{\delta}(1)}\left(\tilde{f},f\right)\right] &\geq \sup_{f_\nu\in\mathcal{F}^{\beta}(\gamma_n)}\mathbb{E}_{f_\nu,Q}\left[d_{\mathcal{W}^{\delta}(1)}\left(\tilde{f},f_\nu\right)\right]\\ &= \max_{\nu\in\{0,1\}^{J^d}}\mathbb{E}_{f_\nu,Q}\left[d_{\mathcal{W}^{\delta}(1)}\left(\tilde{f},f_\nu\right)\right].\end{aligned}$$

Moreover, for all ν in $\{0,1\}^{J^d}$, Then,

$$\begin{aligned}d_{\mathcal{W}^{\delta}(1)}\left(\tilde{f},f_\nu\right) &= \sup_{g\in\mathcal{W}^{\delta}(1)}\int_{[0,1]^d}\left[\left(\tilde{f}-f_\nu\right)g\right]\\ &\geq \sup_{g_\lambda\in\mathcal{D}^{\delta}(\eta)}\int_{[0,1]^d}\left[\left(\tilde{f}-f_\nu\right)g_\lambda\right]\\ &= \max_{\lambda\in\{-1,1\}^{J^d}}\frac{\eta}{J^{\delta}}\sum_{\underline{\boldsymbol{j}}\in\{1,\dots,J\}^d}\lambda_{\underline{\boldsymbol{j}}}\int_{[0,1]^d}\left[\left(\tilde{f}-f_\nu\right)G_{\underline{\boldsymbol{j}}}\right]\\ &\geq \frac{\eta}{J^{\delta}}\sum_{\underline{\boldsymbol{j}}\in\{1,\dots,J\}^d}\left|\int_{[0,1]^d}\left[\left(\tilde{f}-f_\nu\right)G_{\underline{\boldsymbol{j}}}\right]\right|, \qquad (13)\end{aligned}$$

with the particular choice of $\tilde{\lambda}$ such that for all $\underline{\boldsymbol{j}}$, $\tilde{\lambda}_{\underline{\boldsymbol{j}}}$ is the sign of the integral $\int_{[0,1]^d}[(\tilde{f}-f_\nu)G_{\underline{\boldsymbol{j}}}]$. Yet, by definition of f_ν in $\mathcal{F}^{\beta}(\gamma_n)$, and since the supports of the functions $G_{\underline{\boldsymbol{j}}}$ are disjoint,

$$\left|\int_{[0,1]^d}\left[\left(\tilde{f}-f_\nu\right)G_{\underline{\boldsymbol{j}}}\right]\right| = \left|\int_{[0,1]^d}\left[\left(\tilde{f}-h_{\underline{\boldsymbol{j}}}(\nu_{\underline{\boldsymbol{j}}})\right)G_{\underline{\boldsymbol{j}}}\right]\right|,$$

where $h_{\underline{\boldsymbol{j}}}(\nu_{\underline{\boldsymbol{j}}}) = 1+\frac{\gamma_n}{J^{\beta}}\nu_{\underline{\boldsymbol{j}}}G_{\underline{\boldsymbol{j}}}$. Introduce

$$\tilde{\nu}_{\underline{\boldsymbol{j}}}\in\operatorname*{argmin}_{\nu_{\underline{\boldsymbol{j}}}\in\{0,1\}}\left|\int_{[0,1]^d}\left[\left(\tilde{f}-h_{\underline{\boldsymbol{j}}}(\nu_{\underline{\boldsymbol{j}}})\right)G_{\underline{\boldsymbol{j}}}\right]\right|.$$

Then by the triangular inequality,

$$\begin{aligned}\left|\int_{[0,1]^d}\left[\left(\tilde{f}-f_\nu\right)G_{\underline{\boldsymbol{j}}}\right]\right| &\geq \frac{1}{2}\left\{\left|\int_{[0,1]^d}\left[\left(\tilde{f}-h_{\underline{\boldsymbol{j}}}(\tilde{\nu}_{\underline{\boldsymbol{j}}})\right)G_{\underline{\boldsymbol{j}}}\right]\right|+\left|\int_{[0,1]^d}\left[\left(\tilde{f}-h_{\underline{\boldsymbol{j}}}(\nu_{\underline{\boldsymbol{j}}})\right)G_{\underline{\boldsymbol{j}}}\right]\right|\right\}\\ &\geq \frac{1}{2}\left|\int_{[0,1]^d}\left[\left(h_{\underline{\boldsymbol{j}}}(\tilde{\nu}_{\underline{\boldsymbol{j}}})-h_{\underline{\boldsymbol{j}}}(\nu_{\underline{\boldsymbol{j}}})\right)G_{\underline{\boldsymbol{j}}}\right]\right|\\ &= \frac{\gamma_n}{2J^{\beta}}\left|\tilde{\nu}_{\underline{\boldsymbol{j}}}-\nu_{\underline{\boldsymbol{j}}}\right|\int_{[0,1]^d}G_{\underline{\boldsymbol{j}}}^2\\ &= \frac{\|\psi\|_2^{2d}}{2}\frac{\gamma_n}{J^{\beta+d}}\left|\tilde{\nu}_{\underline{\boldsymbol{j}}}-\nu_{\underline{\boldsymbol{j}}}\right|. \qquad (14)\end{aligned}$$

by Equation (11) for $p=2$. Finally, combining (13) and (14) leads to

$$d_{\mathcal{W}^{\delta}(1)}\left(\tilde{f},f_\nu\right)\geq\frac{\eta\,\|\psi\|_2^{2d}}{2}\frac{\gamma_n}{J^{\beta+\delta+d}}\rho_H(\tilde{\nu},\nu),$$

where $\rho_H(\tilde{\nu},\nu) = \sum_{\underline{\boldsymbol{j}}} |\tilde{\nu}_{\underline{\boldsymbol{j}}} - \nu_{\underline{\boldsymbol{j}}}| = \sum_{\underline{\boldsymbol{j}}} \mathbb{1}_{\tilde{\nu}_{\underline{\boldsymbol{j}}} \neq \nu_{\underline{\boldsymbol{j}}}}$ denotes the Hamming distance between $\tilde{\nu}$ and ν in $\{0,1\}^{J^d}$. Thus,

$$\begin{aligned}\sup_{f\in\mathcal{W}^\beta(R)} \mathbb{E}_{f,Q}\left[d_{\mathcal{W}^\delta(1)}\left(\tilde{f},f\right)\right] &\geq \frac{\eta\,\|\psi\|_2^{2d}}{2}\frac{\gamma_n}{J^{\beta+\delta+d}} \max_{\nu\in\{0,1\}^{J^d}} \mathbb{E}_{f_\nu,Q}[\rho_H(\tilde{\nu},\nu)] \\ &\geq \frac{\eta\,\|\psi\|_2^{2d}}{2}\frac{\gamma_n}{J^{\beta+\delta+d}} \times \inf_{\hat{\nu}} \max_{\nu\in\{0,1\}^{J^d}} \mathbb{E}_{f_\nu,Q}[\rho_H(\hat{\nu},\nu)]\,, \quad (15)\end{aligned}$$

where the infimum is taken over all estimators $\hat{\nu}$ based on data from the distribution P_ν with density f_ν, that have been privatized using a privacy mechanism Q. Denote M_ν^n the distribution of such privatized data. In order to lower bound this infimum, let us use Theorem 2.12 [Tsybakov, 2009, p. 118] recalled below.

Theorem 4.1 ([Tsybakov, 2009] Theorem 2.12)**.** *Let $\Theta = \{0,1\}^N$, with $N \geq 1$ and $\{P_\theta, \theta\in\Theta\}$ be a set of 2^N probability measures on a measurable space $(\mathcal{X},\mathcal{A})$. Denote $\mathbb{E}_\theta$ the corresponding expectations. Assume that there exists $\xi > 0$ such that for all $\theta,\theta'\in\Theta$ satisfying $\rho_H(\theta,\theta') = 1$, the Kullback-Leibler divergence between P_θ and $P_{\theta'}$ satisfies $KL(P_\theta, P_{\theta'}) \leq \xi$. Then, it yields*

$$\inf_{\hat{\theta}} \max_{\theta\in\Theta} \mathbb{E}_\theta\left[\rho_H\left(\hat{\theta},\theta\right)\right] \geq \frac{N}{2}\max\left\{\frac{e^{-\xi}}{2}, 1-\sqrt{\frac{\xi}{2}}\right\}.$$

We thus need to upper bound the Kullback-Leibler divergence of the privatized distributions M_ν^n and $M_{\nu'}^n$, where ν and ν' belong to $\{0,1\}^{J^d}$ such that $\rho_H(\nu,\nu') = 1$. Yet, according to [Duchi et al., 2013b, Theorem1],

$$\mathrm{KL}(M_\nu^n, M_{\nu'}^n) \leq 4n\left(e^\alpha - 1\right)^2 \mathrm{TV}^2(P_\nu, P'_\nu).$$

Moreover, if $\rho_H(\nu,\nu') = 1$, then there exists a unique $\underline{\boldsymbol{j_0}}$ such that $\nu_{\underline{\boldsymbol{j_0}}} \neq \nu'_{\underline{\boldsymbol{j_0}}}$, and thus

$$\mathrm{TV}(P_\nu, P'_\nu) = \frac{1}{2}\int_{[0,1]^d} |f_\nu - f_{\nu'}| = \frac{\gamma_n}{2J^\beta}\underbrace{\left|\nu_{\underline{\boldsymbol{j_0}}} - \nu'_{\underline{\boldsymbol{j_0}}}\right|}_{=1}\int_{[0,1]^d}\left|G_{\underline{\boldsymbol{j_0}}}\right| = \frac{\|\psi\|_1^d}{2}\frac{\gamma_n}{J^{\beta+d}},$$

by Equation (11) for $p = 1$. We deduce that

$$\mathrm{KL}(M_\nu^n, M_{\nu'}^n) \leq \|\psi\|_1^{2d}\frac{\gamma_n^2}{J^{2\beta+2d}} n\left(e^\alpha - 1\right)^2 \leq e^{2A}\|\psi\|_1^{2d}\frac{\gamma_n^2}{J^{2\beta+2d}} n\alpha^2,$$

as one can prove by Taylor Theorem with the Lagrange form of remainder that for all α in $(0,A]$, $e^\alpha - 1 \leq \alpha e^A$. In particular, if $\frac{\gamma_n^2}{J^{2\beta+2d}} n\alpha^2 =: \gamma^2$ is constant, we obtain that

$$\mathrm{KL}(M_\nu^n, M_{\nu'}^n) \leq \xi, \qquad \text{where} \qquad \xi = \gamma^2 e^{2A}\|\psi\|_1^{2d}.$$

Applying Theorem 4.1 with $N = J^d$ to the privatized distributions leads to

$$\inf_{\hat{\nu}} \max_{\nu\in\{0,1\}^{J^d}} \mathbb{E}_{f_\nu,Q}[\rho_H(\hat{\nu},\nu)] \geq \frac{J^d}{2}\max\left\{\frac{e^{-\xi}}{2}, 1-\sqrt{\frac{\xi}{2}}\right\}.$$

Finally, from (15), we obtain

$$\sup_{f\in\mathcal{W}^\beta(R)} \mathbb{E}_{f,Q}\left[d_{\mathcal{W}^\delta(1)}\left(\tilde{f},f\right)\right] \geq c_0\frac{\gamma_n}{J^{\beta+\delta}},$$

where $c_0 = \left(\eta\left\|\psi\right\|_2^{2d}/4\right)\max\left\{e^{-\xi}/2, 1-\sqrt{\xi/2}\right\}$ is a constant depending on β,δ,R,d,A. This being true for all $Q\in\mathcal{Q}_\alpha$ and $\tilde{f}$, we deduce that

$$\rho_n(\mathcal{W}^\beta(R),\mathcal{W}^\delta(1),\mathcal{Q}_\alpha) = \inf_{Q\in\mathcal{Q}_\alpha}\inf_{\tilde{f}}\sup_{f\in\mathcal{W}^\beta(R)}\mathbb{E}\left[d_{\mathcal{W}^\delta(1)}\left(\tilde{f},f\right)\right] \geq c_0\frac{\gamma_n}{J^{\beta+\delta}}.$$

In order to obtain both rates in Theorem 2.1, consider

$$\gamma = \min\left\{\left\|\psi\right\|_\infty^{-d}, \frac{R^2-d}{d\left\|\psi\right\|_2^{2(d-1)}\left[\left\|\psi\right\|_2^2+\left\|\psi^{(\beta)}\right\|_2^2\right]}\right\},$$

so that we may apply Lemma 4.1 for the different choices of γ_n below.
First, choosing $J = (n\alpha^2)^{\frac{1}{2\beta+2d}}$ and $\gamma_n = \gamma$ implies that $\gamma_n^2\frac{n\alpha^2}{J^{2\beta+2d}} = \gamma^2$ is a constant and thus,

$$\rho_n(\mathcal{W}^\beta(R),\mathcal{W}^\delta(1),\mathcal{Q}_\alpha) \geq \frac{\gamma c_0}{J^{\beta+\delta}} = c\,(n\alpha^2)^{-\frac{\beta+\delta}{2\beta+2d}}.$$

Second, choosing $J=1$ and $\gamma_n = \gamma/(\alpha\sqrt{n}) \leq \gamma$ (as $\alpha\geq 1/\sqrt{n}$) also implies that $\gamma_n^2\frac{n\alpha^2}{J^{2\beta+2d}} = \gamma^2$ is a constant and thus,

$$\rho_n(\mathcal{W}^\beta(R),\mathcal{W}^\delta(1),\mathcal{Q}_\alpha) \geq c_0\frac{\gamma}{\alpha\sqrt{n}} = c\,(n\alpha^2)^{-1/2}.$$

This concludes the proof of Theorem 2.1.

4.2 Proof of Theorem 2.2

In this proof, denote for all $\underline{\boldsymbol{j}} = (j_1,\ldots,j_d)$ in $(\mathbb{N}^*)^d$, $\left\|\underline{\boldsymbol{j}}^\beta\right\|^2 = j_1^{2\beta}+j_2^{2\beta}+\ldots+j_d^{2\beta}$, and let f in $\mathcal{W}^\beta(R)$, such that

$$\sum_{\underline{\boldsymbol{j}}\in(\mathbb{N}^*)^d}\left\|\underline{\boldsymbol{j}}^\beta\right\|^2\theta_{\underline{\boldsymbol{j}}}(f)^2 \leq R^2.$$

We aim at upper bounding

$$\mathbb{E}\left[d_{\mathcal{W}^\delta(1)}\left(\hat{f}_J,f\right)\right] = \mathbb{E}\left[\sup_{g\in\mathcal{W}^\delta(1)}\int_{[0,1]^d}(\hat{f}_J-f)g\right].$$

Let g in $\mathcal{W}^\delta(1)$ and consider its decomposition in the Fourier basis: $g = \sum_{\underline{\boldsymbol{j}}\in(\mathbb{N}^*)^d}\theta_{\underline{\boldsymbol{j}}}(g)\varphi_{\underline{\boldsymbol{j}}}$. Then,

$$\begin{aligned}\int_{[0,1]^d}(\hat{f}_J-f)g &= \sum_{\underline{\boldsymbol{j}}\in(\mathbb{N}^*)^d}\left[\hat{\theta}_{\underline{\boldsymbol{j}}}-\theta_{\underline{\boldsymbol{j}}}(f)\right]\theta_{\underline{\boldsymbol{j}}}(g)\\ &= \sum_{\underline{\boldsymbol{j}}\in\{1,\ldots,J\}^d}\left[\hat{\theta}_{\underline{\boldsymbol{j}}}-\theta_{\underline{\boldsymbol{j}}}(f)\right]\theta_{\underline{\boldsymbol{j}}}(g) + \sum_{\underline{\boldsymbol{j}}\notin\{1,\ldots,J\}^d}\left[-\theta_{\underline{\boldsymbol{j}}}(f)\right]\theta_{\underline{\boldsymbol{j}}}(g).\end{aligned}$$

In particular,

$$\mathbb{E}\left[d_{\mathcal{W}^\delta(1)}\left(\hat{f}_J, f\right)\right] \leq \underbrace{\mathbb{E}\left[\sup_{g\in\mathcal{W}^\delta(1)} \sum_{\underline{\boldsymbol{j}}\in\{1,\dots,J\}^d} \left[\hat{\theta}_{\underline{\boldsymbol{j}}} - \theta_{\underline{\boldsymbol{j}}}(f)\right]\theta_{\underline{\boldsymbol{j}}}(g)\right]}_{E_1} + \underbrace{\sup_{g\in\mathcal{W}^\delta(1)} \sum_{\underline{\boldsymbol{j}}\notin\{1,\dots,J\}^d} \left[-\theta_{\underline{\boldsymbol{j}}}(f)\right]\theta_{\underline{\boldsymbol{j}}}(g)}_{E_2}. \quad (16)$$

• Let us first upper bound E_2.
By the Cauchy-Schwarz inequality,

$$\sum_{\underline{\boldsymbol{j}}\notin\{1,\dots,J\}^d} \left[-\theta_{\underline{\boldsymbol{j}}}(f)\right]\theta_{\underline{\boldsymbol{j}}}(g) \leq \sqrt{\left(\sum_{\underline{\boldsymbol{j}}\notin\{1,\dots,J\}^d} \theta_{\underline{\boldsymbol{j}}}(f)^2\right)\left(\sum_{\underline{\boldsymbol{j}}\notin\{1,\dots,J\}^d} \theta_{\underline{\boldsymbol{j}}}(g)^2\right)}.$$

Yet, since f belongs to $\mathcal{W}^\beta(R)$,

$$\sum_{\underline{\boldsymbol{j}}\notin\{1,\dots,J\}^d} \theta_{\underline{\boldsymbol{j}}}(f)^2 = \sum_{\underline{\boldsymbol{j}}\notin\{1,\dots,J\}^d} \left\|\underline{\boldsymbol{j}}^\beta\right\|^2 \theta_{\underline{\boldsymbol{j}}}(f)^2 \frac{1}{\left\|\underline{\boldsymbol{j}}^\beta\right\|^2} \leq R^2 J^{-2\beta},$$

as for all $\underline{\boldsymbol{j}} \notin \{1,2,\dots,J\}^d$, there exists $m \in \{1,\dots,d\}$ such that $j_m > J$. Then,

$$\left\|\underline{\boldsymbol{j}}^\beta\right\|^2 \geq j_m^{2\beta} > J^{2\beta}.$$

In the same way, since g belongs to $\mathcal{W}^\delta(1)$, $\sum_{\underline{\boldsymbol{j}}\notin\{1,\dots,J\}^d} \theta_{\underline{\boldsymbol{j}}}(g)^2 \leq J^{-2\delta}$. Hence,

$$\sum_{\underline{\boldsymbol{j}}\notin\{1,\dots,J\}^d} \left[-\theta_{\underline{\boldsymbol{j}}}(f)\right]\theta_{\underline{\boldsymbol{j}}}(g) \leq \sqrt{R^2 J^{-2\beta} J^{-2\delta}}.$$

This being true for any g in $\mathcal{W}^\delta(1)$, we deduce that

$$E_2 \leq R J^{-(\beta+\delta)}. \quad (17)$$

• Let us now upper bound E_1.
As the blocks $\mathcal{J}_{\underline{\boldsymbol{\ell}}}$ form a partition of $\{1,\dots,J\}^d$, and by the Cauchy-Schwarz inequality for all $\underline{\boldsymbol{\ell}}$,

$$\begin{aligned}
\sum_{\underline{\boldsymbol{j}}\in\{1,\dots,J\}^d} \left[\hat{\theta}_{\underline{\boldsymbol{j}}} - \theta_{\underline{\boldsymbol{j}}}(f)\right]\theta_{\underline{\boldsymbol{j}}}(g) &= \sum_{\underline{\boldsymbol{\ell}}\in\{0,1,\dots,L\}^d} \sum_{\underline{\boldsymbol{j}}\in\mathcal{J}_{\underline{\boldsymbol{\ell}}}} \left[\hat{\theta}_{\underline{\boldsymbol{j}}} - \theta_{\underline{\boldsymbol{j}}}(f)\right] \frac{1}{\left\|\underline{\boldsymbol{j}}^\delta\right\|} \times \left\|\underline{\boldsymbol{j}}^\delta\right\| \theta_{\underline{\boldsymbol{j}}}(g) \\
&\leq \sum_{\underline{\boldsymbol{\ell}}\in\{0,1,\dots,L\}^d} \sqrt{\left(\sum_{\underline{\boldsymbol{j}}\in\mathcal{J}_{\underline{\boldsymbol{\ell}}}} \left[\hat{\theta}_{\underline{\boldsymbol{j}}} - \theta_{\underline{\boldsymbol{j}}}(f)\right]^2 \frac{1}{\left\|\underline{\boldsymbol{j}}^\delta\right\|^2}\right)\left(\sum_{\underline{\boldsymbol{j}}\in\mathcal{J}_{\underline{\boldsymbol{\ell}}}} \left\|\underline{\boldsymbol{j}}^\delta\right\|^2 \theta_{\underline{\boldsymbol{j}}}(g)^2\right)} \\
&\leq \sum_{\underline{\boldsymbol{\ell}}\in\{0,1,\dots,L\}^d} \sqrt{\sum_{\underline{\boldsymbol{j}}\in\mathcal{J}_{\underline{\boldsymbol{\ell}}}} \left[\hat{\theta}_{\underline{\boldsymbol{j}}} - \theta_{\underline{\boldsymbol{j}}}(f)\right]^2 \frac{1}{\left\|\underline{\boldsymbol{j}}^\delta\right\|^2}},
\end{aligned}$$

and this for all g in $\mathcal{W}^\delta(1)$. Since the upper bound does not depend on g, we deduce that

$$\begin{aligned} E_1 &\leq \sum_{\underline{\boldsymbol{\ell}}\in\{0,1,\dots,L\}^d} \mathbb{E}\left[\sqrt{\sum_{\underline{\boldsymbol{j}}\in\mathcal{J}_{\underline{\boldsymbol{\ell}}}} \left[\hat{\theta}_{\underline{\boldsymbol{j}}} - \theta_{\underline{\boldsymbol{j}}}(f)\right]^2 \frac{1}{\left\|\underline{\boldsymbol{j}}^\delta\right\|^2}}\right] \\ &\leq \sum_{\underline{\boldsymbol{\ell}}\in\{0,1,\dots,L\}^d} \sqrt{\sum_{\underline{\boldsymbol{j}}\in\mathcal{J}_{\underline{\boldsymbol{\ell}}}} \mathbb{E}\left[\left(\hat{\theta}_{\underline{\boldsymbol{j}}} - \theta_{\underline{\boldsymbol{j}}}(f)\right)^2\right] \frac{1}{\left\|\underline{\boldsymbol{j}}^\delta\right\|^2}}, \end{aligned}$$

by Jensen's inequality.
Yet, for any $\underline{\boldsymbol{\ell}}$ in $\{0,1,\dots,L\}^d$ and any $\underline{\boldsymbol{j}} \in \mathcal{J}_{\underline{\boldsymbol{\ell}}}$, by Proposition 2.1, for all $1 \leq i \leq n$, $\mathbb{E}\left[Z_{i,\underline{\boldsymbol{j}}}\middle|X_i\right] = \varphi_{\underline{\boldsymbol{j}}}(X_i)$, thus

$$\mathbb{E}\left[\hat{\theta}_{\underline{\boldsymbol{j}}}\right] = \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[Z_{i,\underline{\boldsymbol{j}}}\right] = \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[\varphi_{\underline{\boldsymbol{j}}}(X_i)\right] = \theta_{\underline{\boldsymbol{j}}}(f).$$

Therefore, using once again Proposition 2.1, as the dimension of the block $\mathcal{J}_{\underline{\boldsymbol{\ell}}}$ equals $d_{\underline{\boldsymbol{\ell}}} = \prod_{m=1}^d 2^{\ell_m}$, and $B_0 = \sqrt{2}$,

$$\mathbb{E}\left[\left(\hat{\theta}_{\underline{\boldsymbol{j}}} - \theta_{\underline{\boldsymbol{j}}}(f)\right)^2\right] = \mathrm{Var}\left(\hat{\theta}_{\underline{\boldsymbol{j}}}\right) = \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}\left(Z_{i,\underline{\boldsymbol{j}}}\right) \leq C_A \frac{\prod_{m=1}^d 2^{\ell_m}}{n\alpha_{\underline{\boldsymbol{\ell}}}^2},$$

where C_A only depends on A. This upper bound no longer depends on $\underline{\boldsymbol{j}}$.
Moreover, recall the inequality of arithmetic and geometric means that is, for all $a_1,\dots,a_d$ positive numbers,

$$\frac{1}{d}\sum_{m=1}^d a_m \geq \left(\prod_{m=1}^d a_m\right)^{1/d}.$$

It is a direct application of Jensen's inequality for the logarithmic function which is concave.

Then, we obtain that for all $\underline{\boldsymbol{j}}$ in $\mathcal{J}_{\underline{\boldsymbol{\ell}}}$,

$$\left\|\underline{\boldsymbol{j}}^\delta\right\|^2 = \sum_{m=1}^d j_m^{2\delta} \;\geq\; d\prod_{m=1}^d j_m^{2\delta/d} \;\geq\; d\prod_{m=1}^d 2^{\ell_m 2\delta/d},$$

where the last inequality comes from the definition of $\mathcal{J}_{\underline{\boldsymbol{\ell}}}$.
Therefore,

$$\sum_{\underline{\boldsymbol{j}}\in\mathcal{J}_{\underline{\boldsymbol{\ell}}}} \mathbb{E}\left[\left(\hat{\theta}_{\underline{\boldsymbol{j}}} - \theta_{\underline{\boldsymbol{j}}}(f)\right)^2\right] \frac{1}{\left\|\underline{\boldsymbol{j}}^\delta\right\|^2} \leq \frac{C_{d,A}}{n\alpha_{\underline{\boldsymbol{\ell}}}^2} \sum_{\underline{\boldsymbol{j}}\in\mathcal{J}_{\underline{\boldsymbol{\ell}}}} \left[\prod_{m=1}^d 2^{\ell_m(1-2\delta/d)}\right] = \frac{C_{d,A}}{n\alpha_{\underline{\boldsymbol{\ell}}}^2} \prod_{m=1}^d 2^{2\ell_m(1-\delta/d)}$$

as the cardinal of $\mathcal{J}_{\underline{\boldsymbol{\ell}}}$ equals $\prod_{m=1}^d 2^{\ell_m}$.
We deduce that

$$E_1 \leq \frac{C_{d,A}}{\sqrt{n}} \sum_{\underline{\boldsymbol{\ell}}\in\{0,1,\dots,L\}^d} \left[\frac{1}{\alpha_{\underline{\boldsymbol{\ell}}}} \prod_{m=1}^d 2^{\ell_m(1-\delta/d)}\right].$$

Then, solving the constrained-optimization problem which consists in

$$\min_{(\alpha_{\underline{\ell}})_{\underline{\ell}}} \left\{ \sum_{\underline{\ell}\in\{0,1,\dots,L\}^d} \left[\frac{1}{\alpha_{\underline{\ell}}} \prod_{m=1}^{d} 2^{\ell_m(1-\delta/d)} \right] \right\} \qquad \text{subject to} \qquad \sum_{\underline{\ell}\in\{0,1,\dots,L\}^d} \alpha_{\underline{\ell}} = \alpha,$$

by the method of Lagrangian multipliers leads to the choice of the privacy levels

$$\alpha_{\underline{\ell}} = \frac{\alpha}{S_d} \prod_{m=1}^{d} 2^{\ell_m(1-\delta/d)/2}, \qquad \text{where} \qquad S_d = \sum_{\underline{\ell'}\in\{0,1,\dots,L\}^d} \prod_{m=1}^{d} 2^{\ell'_m(1-\delta/d)/2},$$

as defined in Theorem 2.2.
For this particular choice of privacy levels, we obtain

$$E_1 \leq \frac{C_{d,A}}{\sqrt{n}} \frac{S_d}{\alpha} \sum_{\underline{\ell}\in\{0,1,\dots,L\}^d} \left[\prod_{m=1}^{d} 2^{\ell_m(1-\delta/d)/2} \right] = \frac{C_{d,A}}{\alpha\sqrt{n}} S_d^2.$$

Note that if $\delta \neq d$, then,

$$S_d^2 = \left(\sum_{l=0}^{L} \left[2^{(1-\delta/d)/2} \right]^l \right)^{2d} = \left(\frac{2^{(1-\delta/d)(L+1)/2} - 1}{2^{(1-\delta/d)/2} - 1} \right)^{2d} = \frac{1}{\left[2^{(1-\delta/d)/2} - 1 \right]^{2d}} \left((J+1)^{(1-\delta/d)/2} - 1 \right)^{2d}.$$

From here, we distinguish three cases.

Case 1: If $\delta < d$, then $S_d^2 \leq C_{d,A,\delta} J^{d-\delta}$ and $E_1 \leq \frac{C_{d,A,\delta}}{\alpha\sqrt{n}} J^{d-\delta}$. Thus, E_1 and E_2 are of the same order if $J^{-(\beta+\delta)} \asymp \frac{1}{\alpha\sqrt{n}} J^{d-\delta}$, that is if $J \asymp \left(n\alpha^2\right)^{\frac{1}{2\beta+2d}}$. In that case, we obtain

$$\mathbb{E}\left[d_{\mathcal{W}^\delta(1)}\left(\hat{f}_J, f \right) \right] \leq C_{d,A,\delta,R} \left(n\alpha^2\right)^{\frac{-(\beta+\delta)}{2\beta+2d}}.$$

Case 2: If $\delta > d$, then $S_d^2 \leq C_{d,A,\delta}$ and $E_1 \leq \frac{C_{d,A,\delta}}{\alpha\sqrt{n}}$. Thus, E_1 and E_2 are of the same order if $J^{-(\beta+\delta)} \asymp \frac{1}{\alpha\sqrt{n}}$, that is if $J \asymp \left(n\alpha^2\right)^{\frac{1}{2\beta+2\delta}}$. In that case, we obtain

$$\mathbb{E}\left[d_{\mathcal{W}^\delta(1)}\left(\hat{f}_J, f \right) \right] \leq C_{d,A,\delta,R} \left(n\alpha^2\right)^{-1/2}.$$

Case 3: If $\delta = d$, then $S_d^2 = (L+1)^{2d} \leq C_{d,A} \ln(J)^{2d}$, as $J = 2^{L+1} - 1$. Then, setting

$$J \asymp \left(\frac{n\alpha^2}{\ln(n\alpha^2)^{4d}} \right)^{\frac{1}{2\beta+2d}},$$

leads to

$$E_1 \leq \frac{C_{d,A}}{\alpha\sqrt{n}} \ln(J)^{2d} \leq \frac{C_{d,A,\beta}}{\alpha\sqrt{n}} \ln(n\alpha^2)^{2d},$$

since $\ln(J) = \frac{1}{2\beta+2d} \left[\ln(n\alpha^2) - 4d \ln(n\alpha^2) \right] \leq C_{d,A,\beta} \ln(n\alpha^2)$. In that case, we obtain

$$\mathbb{E}\left[d_{\mathcal{W}^\delta(1)}\left(\hat{f}_J, f \right) \right] \leq C_{d,A,\beta,R} \left(\frac{n\alpha^2}{\ln(n\alpha^2)^{4d}} \right)^{-1/2}.$$

A Complementary proofs

A.1 Characterization of Sobolev balls using partial derivatives

Lemma A.1. *Let $\underline{\boldsymbol{\beta}} = (\beta_1, \ldots, \beta_d)$ be a regularity parameter with integer coordinates. Consider f a function in $\mathbb{L}_2([0,1]^d)$ such that $\int_{[0,1]^d} f^2(x)dx \leq C_1^2$ and for all $1 \leq m \leq d$, f is β_m-differentiable w.r.t. the m-th variable, and satisfies,*

$$\int_{[0,1]^d} \left[\frac{\partial^{\beta_m} f}{\partial x_m^{\beta_m}}(x)\right]^2 dx \leq C_2^2.$$

Assume also that f and all its partial derivatives are periodic, namely for all $m = 1, \ldots, d$, for all $0 \leq l_m \leq \beta_m$, for all $(x_1, \ldots, x_{m-1}, x_{m+1}, \ldots, x_d)$ in $[0,1]^{d-1}$,

$$\frac{\partial^{l_m} f}{\partial x_m^{l_m}}(x_1, \ldots, x_{m-1}, 0, \ldots, x_d) = \frac{\partial^{l_m} f}{\partial x_m^{l_m}}(x_1, \ldots, x_{m-1}, 1, \ldots, x_d).$$

Then,

$$\sum_{(j_1,\ldots,j_d)\in(\mathbb{N}^*)^d} \left(j_1^{2\beta_1} + \ldots + j_d^{2\beta_d}\right) \theta^2_{(j_1,\ldots,j_d)}(f) \leq L^2,$$

where $L^2 = d\left(C_1^2 + C_2^2\right)$. In particular, f belongs to the Sobolev ball $\mathcal{W}^{\underline{\boldsymbol{\beta}}}(R)$ as defined in (4), as soon as

$$d\left(C_1^2 + C_2^2\right) \leq R^2.$$

First note that, by Parseval's equality,

$$\sum_{\underline{\boldsymbol{j}}\in(\mathbb{N}^*)^d} \theta^2_{\underline{\boldsymbol{j}}}(f) = \int_{[0,1]^d} f^2(x)dx \leq C_1^2, \tag{18}$$

and

$$\sum_{\underline{\boldsymbol{j}}\in(\mathbb{N}^*)^d} \theta^2_{\underline{\boldsymbol{j}}}\left(\frac{\partial^{\beta_m} f}{\partial x_m^{\beta_m}}\right) = \int_{[0,1]^d} \left[\frac{\partial^{\beta_m} f}{\partial x_m^{\beta_m}}(x)\right]^2 dx \leq C_2^2. \tag{19}$$

By integrations by parts, and using the periodicity of f and all its partial derivatives, we have

$$\begin{aligned}\theta_{\underline{\boldsymbol{j}}}\left(\frac{\partial^{\beta_m} f}{\partial x_m^{\beta_m}}\right) &= \int_{[0,1]^d} \frac{\partial^{\beta_m} f}{\partial x_m^{\beta_m}} \varphi_{\underline{\boldsymbol{j}}} \\ &= \pm \int_{[0,1]^d} f \frac{\partial^{\beta_m} \varphi_{\underline{\boldsymbol{j}}}}{\partial x_m^{\beta_m}}.\end{aligned}$$

Recalling that for all $\underline{\boldsymbol{j}} = (j_1, \ldots, j_d)$ in $(\mathbb{N}^*)^d$ and all $x = (x_1, \ldots, x_d)$ in $[0,1]^d$ $\varphi_{\underline{\boldsymbol{j}}}(x) = \prod_{l=1}^d \varphi_{j_l}(x_l)$, we have

$$\frac{\partial^{\beta_m} \varphi_{\underline{\boldsymbol{j}}}}{\partial x_m^{\beta_m}}(x) = \prod_{l\neq m=1}^{d} \varphi_{j_l}(x_l) \varphi_{j_m}^{(\beta_m)}(x_m).$$

Moreover, by definition of the Fourier basis, if β is even, for all $j \in \mathbb{N}^*$, $\varphi_{2j}^{(\beta)} = \pm(2\pi j)^\beta \varphi_{2j}$ and $\varphi_{2j+1}^{(\beta)} = \pm(2\pi j)^\beta \varphi_{2j+1}$, if β is odd, for all $j \in \mathbb{N}^*$, $\varphi_{2j}^{(\beta)} = \pm(2\pi j)^\beta \varphi_{2j+1}$ and $\varphi_{2j+1}^{(\beta)} =$

$\pm(2\pi j)^\beta \varphi_{2j}$.
Hence, if β_m is even and $j_m \geq 2$,

$$\theta_{\underline{\boldsymbol{j}}}\left(\frac{\partial^{\beta_m} f}{\partial x_m^{\beta_m}}\right) = \begin{cases} \pm\left(\frac{2\pi j_m}{2}\right)^{\beta_m} \theta_{\underline{\boldsymbol{j}}}(f) & \text{if } j_m \text{ is even} \\ \pm\left(\frac{2\pi(j_m-1)}{2}\right)^{\beta_m} \theta_{\underline{\boldsymbol{j}}}(f) & \text{if } j_m \text{ is odd.} \end{cases}$$

Since for all $j_m \geq 2$, $j_m \leq 2(j_m - 1)$, we deduce from (19) that

$$\sum_{\underline{\boldsymbol{j}}\in(\mathbb{N}^*)^d,\ j_m\geq 2} j_m^{2\beta_m}\theta^2_{\underline{\boldsymbol{j}}}(f) \leq \sum_{\underline{\boldsymbol{j}}\in(\mathbb{N}^*)^d,\ j_m\geq 2} \left(\frac{\pi j_m}{2}\right)^{2\beta_m} \theta^2_{\underline{\boldsymbol{j}}}(f) \leq C_2^2. \tag{20}$$

If β_m is odd and $j_m \geq 2$,

$$\theta_{\underline{\boldsymbol{j}}}\left(\frac{\partial^{\beta_m} f}{\partial x_m^{\beta_m}}\right) = \begin{cases} \pm\left(\frac{2\pi j_m}{2}\right)^{\beta_m} \theta_{j_1\ldots j_{m+1}\ldots j_d}(f) & \text{if } j_m \text{ is even} \\ \pm\left(\frac{2\pi(j_m-1)}{2}\right)^{\beta_m} \theta_{j_1\ldots j_{m-1}\ldots j_d}(f) & \text{if } j_m \text{ is odd,} \end{cases}$$

and we also obtain (20). Finally, using (18), we get

$$\sum_{\underline{\boldsymbol{j}}\in(\mathbb{N}^*)^d} {j_m}^{2\beta_m}\theta^2_{\underline{\boldsymbol{j}}}(f) \leq C_1^2 + C_2^2,$$

which concludes the proof of Lemma A.1.

A.2 Proof of Lemma 4.1

1. Let f_ν belong to $\mathcal{F}^\beta(\gamma_n)$, and let x in $[0,1]^d$. Then, there exists a unique $\underline{\boldsymbol{j_0}}$ such that x belongs to the support of $G_{\underline{\boldsymbol{j_0}}}$. In particular, if $\gamma_n \leq \|\psi\|_\infty^{-d} \leq J^\beta \|\psi\|_\infty^{-d}$, then

$$f_\nu(x) = 1 + \frac{\gamma_n}{J^\beta}\nu_{\underline{\boldsymbol{j_0}}} G_{\underline{\boldsymbol{j_0}}}(x) \geq 1 - \frac{\gamma_n}{J^\beta}\left|G_{\underline{\boldsymbol{j_0}}}(x)\right| \geq 1 - \frac{\gamma_n}{J^\beta}\|\psi\|_\infty^d \geq 0.$$

   Moreover, using the left hand side equation in (11) directly leads to

$$\int_{[0,1]^d} f_\nu(x)dx \quad = \quad 1 + \frac{\gamma_n}{J^\beta} \sum_{\underline{\boldsymbol{j}}\in\{1,\ldots,J\}^d} \nu_{\underline{\boldsymbol{j}}} \underbrace{\int_{[0,1]^d} G_{\underline{\boldsymbol{j}}}(x)dx}_{=0} = 1.$$

2. To prove this point, we shall use Lemma A.1.
   Consider $\nu = (\nu_{\underline{\boldsymbol{j}}})_{\underline{\boldsymbol{j}}\in\{1,\ldots,J\}^d}$ in $\{0,1\}^{J^d}$ and consider $f_\nu : [0,1]^d \to \mathbb{R}$ in $\mathcal{F}^\beta(\gamma_n)$, defined for all x in $[0,1]^d$ by

$$f_\nu(x) = 1 + \frac{\gamma_n}{J^\beta} \sum_{\underline{\boldsymbol{j}}\in\{1,\ldots,J\}^d} \nu_{\underline{\boldsymbol{j}}} G_{\underline{\boldsymbol{j}}}(x).$$

Let us first upper bound the $\mathbb{L}_2$ norm of f_ν. As the supports of the $G_{\underline{\boldsymbol{j}}}$ are disjoint, by both Equations in (11)

$$\begin{aligned}\int_{[0,1]^d} f_\nu(x)^2 dx &= 1 + \frac{\gamma_n^2}{J^{2\beta}} \sum_{\underline{\boldsymbol{j}} \in \{1,\dots,J\}^d} \nu_{\underline{\boldsymbol{j}}}^2 \int_{[0,1]^d} G_{\underline{\boldsymbol{j}}}(x)^2 dx. \\ &\leq 1 + \frac{\gamma_n^2}{J^{2\beta}} \|\psi\|_2^{2d} \\ &\leq 1 + \gamma_n^2 \|\psi\|_2^{2d}.\end{aligned}$$

Let m in $\{1,\dots,d\}$. Then for all $x \in [0,1]^d$,

$$\frac{\partial^\beta}{\partial x_m^\beta} f_\nu(x) = \frac{\gamma_n}{J^\beta} \sum_{\underline{\boldsymbol{j}} \in \{1,\dots,J\}^d} \nu_{\underline{\boldsymbol{j}}} \left[\frac{\partial^\beta}{\partial x_m^\beta} G_{\underline{\boldsymbol{j}}}(x) \right].$$

Since the supports of the $G_{\underline{\boldsymbol{j}}}$ for $\underline{\boldsymbol{j}} \in \{1,\dots,J\}^d$ are disjoint,

$$\left[\frac{\partial^\beta}{\partial x_m^\beta} f_\nu(x) \right]^2 = \frac{\gamma_n^2}{J^{2\beta}} \sum_{\underline{\boldsymbol{j}} \in \{1,\dots,J\}^d} \nu_{\underline{\boldsymbol{j}}}^2 \left[\frac{\partial^\beta}{\partial x_m^\beta} G_{\underline{\boldsymbol{j}}}(x) \right]^2,$$

and in particular,

$$\int_{[0,1]^d} \left[\frac{\partial^\beta}{\partial x_m^\beta} f_\nu(x) \right]^2 dx = \frac{\gamma_n^2}{J^{2\beta}} \sum_{\underline{\boldsymbol{j}} \in \{1,\dots,J\}^d} \nu_{\underline{\boldsymbol{j}}}^2 \int_{[0,1]^d} \left[\frac{\partial^\beta}{\partial x_m^\beta} G_{\underline{\boldsymbol{j}}}(x) \right]^2 dx.$$

Yet, by definition of $G_{\underline{\boldsymbol{j}}}$,

$$\left| \frac{\partial^\beta}{\partial x_m^\beta} G_{\underline{\boldsymbol{j}}}(x) \right| = \left| J^\beta \psi^{(\beta)} \left(J \left(x_m - \frac{j_m - 1}{J} \right) \right) \times \left[\prod_{m' \neq m} \psi \left(J \left(x_{m'} - \frac{j_{m'} - 1}{J} \right) \right) \right] \right|$$

Therefore,

$$\begin{aligned}\int_{[0,1]^d} \left[\frac{\partial^\beta}{\partial x_m^\beta} G_{\underline{\boldsymbol{j}}}(x) \right]^2 dx &\leq J^{2\beta} \int_{\frac{j_m - 1}{J}}^{\frac{j_m}{J}} \left[\psi^{(\beta)} \left(J \left(x_m - \frac{j_m - 1}{J} \right) \right) \right]^2 dx_m \\ &\quad \times \prod_{m' \neq m} \int_{\frac{j_{m'} - 1}{J}}^{\frac{j_{m'}}{J}} \left[\psi \left(J \left(x_{m'} - \frac{j_{m'} - 1}{J} \right) \right) \right]^2 dx_{m'} \\ &\leq \frac{J^{2\beta} \left(\left\| \psi^{(\beta)} \right\|_2 \|\psi\|_2^{(d-1)} \right)^2}{J^d}\end{aligned}$$

We deduce that, as $\nu_{\underline{\boldsymbol{j}}}^2 \leq 1$ for all $\underline{\boldsymbol{j}}$,

$$\int_{[0,1]^d} \left[\frac{\partial^\beta}{\partial x_m^\beta} f_\nu(x) \right]^2 dx \leq \frac{\gamma_n^2}{J^{2\beta}} \times J^d \times \frac{J^{2\beta} \left(\left\| \psi^{(\beta)} \right\|_2 \|\psi\|_2^{d-1} \right)^2}{J^d} = \gamma_n^2 \left(\left\| \psi^{(\beta)} \right\|_2 \|\psi\|_2^{d-1} \right)^2.$$

This result being true for all $1 \leq m \leq d$, we deduce from Lemma A.1 that f_ν belongs to the isotropic Sobolev ball $\mathcal{W}^\beta(R)$ as soon as

$$d \times \left\{ 1 + \gamma_n^2 \left\|\psi\right\|_2^{2(d-1)} \left[\left\|\psi\right\|_2^2 + \left\|\psi^{(\beta)}\right\|_2^2 \right] \right\} \leq R^2,$$

i.e.

$$\gamma_n^2 \leq \frac{R^2 - d}{d \left\|\psi\right\|_2^{2(d-1)} \left[\left\|\psi\right\|_2^2 + \left\|\psi^{(\beta)}\right\|_2^2 \right]}.$$

The proof is similar for the inclusion $\mathcal{D}^\delta(\eta) \subset \mathcal{W}^\delta(1)$, as soon as

$$d \times \eta^2 \left\|\psi\right\|_2^{2(d-1)} \left[\left\|\psi\right\|_2^2 + \left\|\psi^{(\delta)}\right\|_2^2 \right] \leq 1,$$

which ends the proof.

A.3 Proof of Proposition 2.1

1. To prove the α-local differential privacy, we use the independence between the blocks. Indeed, according to [Butucea et al., 2023a, Proposition 3.1], on each block $\mathcal{J}_{\underline{\ell}}$, we obtain that for all x_i and x_i', for all $z_{[\underline{\ell}]}$ in $\left\{ -B_{d_{\underline{\ell}}}(\alpha_{\underline{\ell}}), B_{d_{\underline{\ell}}}(\alpha_{\underline{\ell}}) \right\}^{d_{\underline{\ell}}}$,

$$\frac{\mathbb{P}\left(Z_{i,[\underline{\ell}]} = z_{[\underline{\ell}]} \middle| X_i = x\right)}{\mathbb{P}\left(Z_{i,[\underline{\ell}]} = z_{[\underline{\ell}]} \middle| X_i = x'\right)} \leq e^{\alpha_{\underline{\ell}}}.$$

Then, by independence between the blocks, for all x, x' in $[0,1]^d$ and all

$$z = (z_{[\underline{\ell}]})_{\underline{\ell} \in \{0,1,\ldots,L\}^d} \in \prod_{\underline{\ell} \in \{0,1,\ldots,L\}^d} \left\{ -B_{d_{\underline{\ell}}}(\alpha_{\underline{\ell}}), B_{d_{\underline{\ell}}}(\alpha_{\underline{\ell}}) \right\}^{d_{\underline{\ell}}},$$

we obtain that

$$\begin{aligned} \frac{\mathbb{P}(Z_i = z | X_i = x)}{\mathbb{P}(Z_i = z | X_i = x')} &= \prod_{\underline{\ell} \in \{0,1,\ldots,L\}^d} \frac{\mathbb{P}\left(Z_{i,[\underline{\ell}]} = z_{[\underline{\ell}]} \middle| X_i = x\right)}{\mathbb{P}\left(Z_{i,[\underline{\ell}]} = z_{[\underline{\ell}]} \middle| X_i = x'\right)} \\ &\leq \prod_{\underline{\ell} \in \{0,1,\ldots,L\}^d} e^{\alpha_{\underline{\ell}}} = e^{\alpha}, \end{aligned}$$

since $\sum_{\underline{\ell} \in \{0,1,\ldots,L\}^d} \alpha_{\underline{\ell}} = \alpha$.

2. Fix i in $\{1, 2, \ldots, n\}$ and, $\underline{\boldsymbol{j}}$ in $\mathcal{J}_{\underline{\ell}}$. As the conditional expectation of the privatized coefficients does not depend on the blocks, one directly obtains from [Butucea et al., 2023a, Proposition 3.2] that $\mathbb{E}\left[Z_{i,\underline{\boldsymbol{j}}} \middle| X_i \right] = \varphi_{\underline{\boldsymbol{j}}}(X_i)$.

For the variance, by definition of each privatized coordinate,

$$\mathrm{Var}\left(Z_{i,\underline{\boldsymbol{j}}}\right) \leq \mathbb{E}\left[Z_{i,\underline{\boldsymbol{j}}}^2 \right] \leq \mathbb{E}\left[\tilde{Z}_{i,\underline{\boldsymbol{j}}}^2 \right] \leq B_{d_{\underline{\ell}}}(\alpha_{\underline{\ell}})^2, \tag{21}$$

where for all $a \in (0, A]$ and all integer k,

$$B_k(a) = B_0 \frac{e^a + 1}{e^a - 1} C_k,$$

with C_k as defined in Equation (5). In order to control $B_k(a)$, one may first note that

$$\frac{e^a + 1}{e^a - 1} \leq \frac{e^A + 1}{a}.$$

Then, using Stirling's approximation, that is $p! \sim \sqrt{2\pi p}(p/e)^p$, one may recover that $C_k \sim \sqrt{\pi p}$, where $k = 2p - 1$ if k si odd, and $k = 2p$ if k is even. Indeed, on the one hand, if $k = 2p + 1$ is odd,

$$C_k = 2^{2p} \frac{(p!)^2}{(2p)!} \sim 2^{2p} \frac{2\pi p (p/e)^{2p}}{\sqrt{4\pi p}(2p/e)^{2p}} = \sqrt{\pi p}.$$

On the other hand, if $k = 2p$ is even,

$$\begin{aligned} C_k &= 2^{2p-1} \frac{(p-1)!p!}{[2(p-1)]!2(p-1)} \\ &\sim 2^{2p-1} \frac{\sqrt{2\pi(p-1)}[(p-1)/e]^{p-1} \times \sqrt{2\pi p}(p/e)^p}{\sqrt{4\pi(p-1)}[2(p-1)/e]^{2p-2} \times 2(p-1)} = \sqrt{\pi p} \left(\frac{p}{p-1}\right)^p \frac{1}{e} \sim \sqrt{\pi p}. \end{aligned}$$

In particular, there exists a constant c such that $C_k \leq c\sqrt{k}$. We deduce that

$$B_k(a) \leq B_0 C_A \sqrt{k},$$

where C_A is a constant only depending on A. Finally, from Equation (21) one deduces that

$$\mathrm{Var}\left(Z_{i,\underline{j}}\right) \leq C_A B_0^2 \frac{d_{\underline{\ell}}}{\alpha_{\underline{\ell}}^2},$$

which ends the proof of Proposition 2.1.

References

[Arjovsky et al., 2017] Arjovsky, M., Chintala, S., and Bottou, L. (2017). Wasserstein generative adversarial networks. In *International conference on machine learning*, pages 214–223. PMLR.

[Asatryan et al., 2023] Asatryan, H., Gottschalk, H., Lippert, M., and Rottmann, M. (2023). A convenient infinite dimensional framework for generative adversarial learning. *Electronic Journal of Statistics*, 17(1):391–428.

[Bai et al., 2019] Bai, Y., Ma, T., and Risteski, A. (2019). Approximability of discriminators implies diversity in GANs. In *International Conference on Learning Representations*.

[Biau et al., 2020] Biau, G., Cadre, B., Sangnier, M., and Tanielian, U. (2020). Some theoretical properties of GANs. *The Annals of Statistics*, 48(3):1539 – 1566.

[Bottou et al., 2018] Bottou, L., Arjovsky, M., Lopez-Paz, D., and Oquab, M. (2018). Geometrical insights for implicit generative modeling. In Lev Rozonoer, Boris Mirkin, I. M., editor, *Braverman Readings in Machine Learning: Key Ideas from Inception to Current State*, LNAI Vol. 11100, pages 229–268. Springer.

[Butucea et al., 2020] Butucea, C., Dubois, A., Kroll, M., and Saumard, A. (2020). Local differential privacy: Elbow effect in optimal density estimation and adaptation over besov ellipsoids. *Bernoulli*, 26(3):1727–1764.

[Butucea et al., 2023a] Butucea, C., Dubois, A., and Saumard, A. (2023a). Phase transitions for support recovery under local differential privacy. *Mathematical Statistics and Learning*, 6(1/2):1–50.

[Butucea et al., 2023b] Butucea, C., Rohde, A., and Steinberger, L. (2023b). Interactive versus noninteractive locally differentially private estimation: Two elbows for the quadratic functional. *The Annals of Statistics*, 51(2):464–486.

[Cai et al., 2021] Cai, T. T., Wang, Y., and Zhang, L. (2021). The cost of privacy: Optimal rates of convergence for parameter estimation with differential privacy. *The Annals of Statistics*, 49(5):2825–2850.

[Chen et al., 2022] Chen, M., Liao, W., Zha, H., and Zhao, T. (2022). Distribution approximation and statistical estimation guarantees of generative adversarial networks. *arXiv preprint arXiv:2002.03938*.

[Dubois et al., 2023] Dubois, A., Berrett, T. B., and Butucea, C. ([2023] ©2023). Goodness-of-fit testing for Hölder continuous densities under local differential privacy. In *Foundations of modern statistics*, volume 425 of *Springer Proc. Math. Stat.*, pages 53–119. Springer, Cham.

[Duchi et al., 2013a] Duchi, J., Wainwright, M. J., and Jordan, M. I. (2013a). Local privacy and minimax bounds: Sharp rates for probability estimation. *Advances in Neural Information Processing Systems*, 26.

[Duchi et al., 2013b] Duchi, J. C., Jordan, M. I., and Wainwright, M. J. (2013b). Local privacy and statistical minimax rates. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, pages 429–438. IEEE.

[Duchi et al., 2018] Duchi, J. C., Jordan, M. I., and Wainwright, M. J. (2018). Minimax optimal procedures for locally private estimation. *Journal of the American Statistical Association*, 113(521):182–201.

[Dwork et al., 2006a] Dwork, C., Kenthapadi, K., McSherry, F., Mironov, I., and Naor, M. (2006a). Our data, ourselves: Privacy via distributed noise generation. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 486–503. Springer.

[Dwork et al., 2006b] Dwork, C., McSherry, F., Nissim, K., and Smith, A. (2006b). Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography conference*, pages 265–284. Springer.

[Dziugaite et al., 2015] Dziugaite, G. K., Roy, D. M., and Ghahramani, Z. (2015). Training generative neural networks via maximum mean discrepancy optimization. In *Conference on Uncertainty in Artificial Intelligence*.

[Efromovich, 1999] Efromovich, S. (1999). *Nonparametric Curve Estimation: Methods, Theory and Applications*. Springer Science & Business Media.

[Goodfellow et al., 2014] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, 27.

[Gui et al., 2020] Gui, J., Sun, Z., Wen, Y., Tao, D., and Ye, J. (2020). A review on generative adversarial networks: Algorithms, theory, and applications. *IEEE Transactions on Knowledge and Data Engineering*, 35:3313–3332.

[Isola et al., 2017] Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134.

[Karras et al., 2021] Karras, T., Aittala, M., Laine, S., Härkönen, E., Hellsten, J., Lehtinen, J., and Aila, T. (2021). Alias-free generative adversarial networks. *Advances in Neural Information Processing Systems*, 34:852–863.

[Kronmal and Tarter, 1968] Kronmal, R. and Tarter, M. (1968). The estimation of probability densities and cumulatives by fourier series methods. *Journal of the American Statistical Association*, 63(323):925–952.

[Lam-Weil et al., 2022] Lam-Weil, J., Laurent, B., and Loubes, J.-M. (2022). Minimax optimal goodness-of-fit testing for densities and multinomials under a local differential privacy constraint. *Bernoulli*, 28(1):579–600.

[Ledig et al., 2017] Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., et al. (2017). Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4681–4690.

[Liang, 2021] Liang, T. (2021). How well generative adversarial networks learn distributions. *The Journal of Machine Learning Research*, 22(1):10366–10406.

[Liu et al., 2017] Liu, S., Bousquet, O., and Chaudhuri, K. (2017). Approximation and convergence properties of generative adversarial learning. *Advances in Neural Information Processing Systems*, 30.

[Luise et al., 2020] Luise, G., Pontil, M., and Ciliberto, C. (2020). Generalization properties of optimal transport gans with latent distribution learning. *arXiv preprint arXiv:2007.14641*.

[Mroueh et al., 2018] Mroueh, Y., Li, C.-L., Sercu, T., Raj, A., and Cheng, Y. (2018). Sobolev GAN. In *International Conference on Learning Representations*.

[Müller, 1997] Müller, A. (1997). Integral probability metrics and their generating classes of functions. *Advances in applied probability*, 29(2):429–443.

[Parzen, 1962] Parzen, E. (1962). On estimation of a probability density function and mode. *The annals of mathematical statistics*, 33(3):1065–1076.

[Puchkin et al., 2024] Puchkin, N., Samsonov, S., Belomestny, D., Moulines, E., and Naumov, A. (2024). Rates of convergence for density estimation with generative adversarial networks. *Preprint, arXiv:2102.00199*.

[Reed et al., 2016] Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B., and Lee, H. (2016). Generative adversarial text to image synthesis. In *International conference on machine learning*, pages 1060–1069. PMLR.

[Rohde and Steinberger, 2020] Rohde, A. and Steinberger, L. (2020). Geometrizing rates of convergence under local differential privacy constraints. *Ann. Statist.*, 48(5):2646–2670.

[Rosenblatt, 1956] Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *The annals of mathematical statistics*, pages 832–837.

[Schreuder et al., 2021] Schreuder, N., Brunel, V.-E., and Dalalyan, A. (2021). Statistical guarantees for generative models without domination. In *Algorithmic Learning Theory*, pages 1051–1071. PMLR.

[Schwartz, 1967] Schwartz, S. C. (1967). Estimation of probability density by an orthogonal series. *The Annals of Mathematical Statistics*, pages 1261–1265.

[Silverman, 1978] Silverman, B. W. (1978). Weak and strong uniform consistency of the kernel estimate of a density and its derivatives. *The Annals of Statistics*, pages 177–184.

[Singh and Póczos, 2018] Singh, S. and Póczos, B. (2018). Minimax distribution estimation in Wasserstein distance. *arXiv preprint arXiv:1802.08855*.

[Singh et al., 2018] Singh, S., Uppal, A., Li, B., Li, C.-L., Zaheer, M., and Póczos, B. (2018). Nonparametric density estimation under adversarial losses. *Advances in Neural Information Processing Systems*, 31.

[Smith, 2008] Smith, A. (2008). Efficient, differentially private point estimators. *arXiv preprint arXiv:0809.4794*.

[Smith, 2011] Smith, A. (2011). Privacy-preserving statistical estimation with optimal convergence rates. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 813–822.

[Stéphanovitch et al., 2023] Stéphanovitch, A., Aamari, E., and Levrard, C. (2023). Wasserstein generative adversarial networks are minimax optimal distribution estimators. *Preprint, hal-04315916*.

[Tsybakov, 2009] Tsybakov, A. B. (2009). *Introduction to Nonparametric Estimation*. Springer Series in Statistics. Springer New York.

[Uppal et al., 2019] Uppal, A., Singh, S., and Póczos, B. (2019). Nonparametric density estimation & convergence rates for GANs under Besov IPM losses. *Advances in neural information processing systems*, 32.

[Vondrick et al., 2016] Vondrick, C., Pirsiavash, H., and Torralba, A. (2016). Generating videos with scene dynamics. *Advances in neural information processing systems*, 29.

[Walter, 1977] Walter, G. G. (1977). Properties of hermite series estimation of probability density. *The Annals of Statistics*, pages 1258–1264.

[Wasserman and Zhou, 2010] Wasserman, L. and Zhou, S. (2010). A statistical framework for differential privacy. *Journal of the American Statistical Association*, 105(489):375–389.

[Weed and Berthet, 2019] Weed, J. and Berthet, Q. (2019). Estimation of smooth densities in Wasserstein distance. In *Conference on Learning Theory*, pages 3118–3119. PMLR.

[Ye and Barg, 2017] Ye, M. and Barg, A. (2017). Asymptotically optimal private estimation under mean square loss. *arXiv preprint arXiv:1708.00059*.

[Ye and Barg, 2018] Ye, M. and Barg, A. (2018). Optimal schemes for discrete distribution estimation under locally differential privacy. *IEEE Transactions on Information Theory*, 64(8):5662–5676.