Math. Stat. Learn. 6 (2023), 51–85
DOI 10.4171/MSL/36

Deconvolution for some singular density errors via a combinatorial median of means approach

Clément Marteau and Mathieu Sart

Abstract. We present a versatile and model based procedure for estimating a density in a deconvolution setting where the error density is assumed to be singular enough. We assess the quality of our estimator by establishing non-asymptotic risk bounds for the $\mathbb{L}^1$ loss. We specify them when the density is piecewise constant on a finite number of (unknown) pieces, when it is unimodal, and when it is concave/convex.

1. Introduction

Let $Y_1, \ldots, Y_n$ be n independent and identically distributed real valued random variables defined on an abstract probability space $(\Omega, \mathcal{E}, \mathbb{P})$. We assume that each random variable Y_i can be written as

$$Y_i = X_i + \varepsilon_i, \quad i \in \{1, \ldots, n\}, \tag{1}$$

where for all $i \in \{1, \ldots, n\}$, X_i denotes a random variable admitting a density f_0 with respect to the Lebesgue measure. The variable ε_i denotes measurement errors, independent of X_i, with known density q. Our aim is to estimate f_0 on a compact interval, say [0, 1], from the indirect observations $Y_1, \ldots, Y_n$.

Deconvolution in a statistical context has been at the core of several investigations in the literature. We refer to, e.g., [26] for a comprehensive introduction to this topic. Over the last decades, several approaches have been proposed. For instance, kernel procedures have been designed, taking advantage of the convolution structure in the Fourier domain. We mention a seminal contribution by [20] who establishes rates of convergence under smoothness assumptions on the target density f_0. Several extensions and improvements have then been obtained, including in particular discussions regarding an adaptive choice of the bandwidth. We refer, e.g., to [17] for bootstrap bandwidth selection, [9] or [12] when the error density q is unknown, [16] for rates of

2020 Mathematics Subject Classification. 62G05.
Keywords. Density deconvolution, combinatorial method, shape constraints.

convergence using the Wasserstein metric, or [23, 29] for adaptation under $\mathbb{L}_p$-losses. Alternatively, several methods based on the minimization of a (penalized) criterion over a given family have been carried out. The penalty allows to include constraints on the target, often in an adaptive way. We refer for instance to [14] for projection methods and to [28] where the performance of a dictionary based method is investigated in a spherical context. Other authors have extended wavelet approaches to the deconvolution model, taking advantage of the density representation property in a wavelet basis. The coefficients of this decomposition must then be estimated and sometimes thresholded in an appropriate way, see [11, 21, 24, 27].

The aforementioned references assess the quality of their estimators by means of $\mathbb{L}^p$-losses, with p larger than 1 (except [20] and [23]). In the present paper, we focus to the case $p = 1$. The loss then does not only measure the distance between the estimator and the density f_0, but also the distance between the two underlying probability measures.

We restrict our study to a moderately ill-posed deconvolution problem with small degree of ill-posedness. More precisely, we assume that the Fourier transform $q^\star$ of q does not vanish and that $|q^\star(t)|^{-1}$ is of the order of $|t|^\beta$ for large values of t. The parameter β should be smaller than $1/2$, which indicates in some sense that q is singular enough, or in other terms, that the framework is not so far away from the direct setting (which would correspond to $\beta = 0$). As noticed by [15, 26], this condition makes it possible the consistent estimation of the distribution function without further assumptions on f_0. This last problem is indeed closely related to the estimation of a density using the $\mathbb{L}^1$ loss. We refer to [19] for more information regarding this issue.

In direct estimation, there exist general procedures that lead to both optimal and robust estimators. We are thinking in particular to methods based on robust tests or like-minded approaches such as those described in [2, 3, 6, 19]. There does not seem to be, in the literature of deconvolution, a *general-purpose* estimation procedure. In this paper, we present a first attempt by proposing a new estimation procedure in line with [19]. It does not lead to risk bounds as general as in the direct case, but at least, it already yields new results in the deconvolution setting under the $\mathbb{L}^1$ loss as described below.

We investigate the case where f_0 is piecewise constant on a finite number r of pieces. The parameter r is assumed to be known but the pieces may be fully unknown to the statistician. We use our method to derive an estimator and get a risk bound of the order of $\|f_0 * q\|_\infty{}^{1/2} r^{1/2+\beta} (\log n/n)^{1/2}$. Moreover, we show that such a result cannot be true in general, no matter the estimator, when β is higher.

Our method also allows us to deal with shape constraints on f_0. The problem of estimation under shape constraints has been widely studied in different statistical frameworks, such as, for example, that of density without noise, or that of regression. This does not seem to be true in density deconvolution, where the only attempts we are aware of are those of [10, 30]. Here, we propose to consider the case where f_0 is unimodal on [0, 1] (with unknown mode) and the case where f_0 is concave on [0, 1]. In each case, our procedure can be applied to get a consistent estimator. The rate of convergence is $(\log n/n)^{1/(2\beta+3)}$ under the unimodal assumption, and $(\log n/n)^{2/(2\beta+5)}$ under the concavity assumption. We prove that these rates are the minimax ones, within a logarithmic factor.

Throughout the paper, we will use the following notation. The terms X, Y, ε denote generic random variables having same law than respectively X_i, Y_i and ε_i. We will sometimes write $\int_A f$ instead of $\int_A f(x)\,\mathrm{d}x$ to shorten the formulas and omit the set A when $A = \mathbb{R}$. We denote the Fourier transform of a integrable, or square integrable, function f by $f^\star$. The following definition is used when f is integrable:

$$f^\star(t) = \int e^{itx} f(x)\,\mathrm{d}x \quad \text{for all } t \in \mathbb{R}.$$

Since we estimate f_0 on [0, 1] only, it is convenient to modify the definition of the $\mathbb{L}^1$ distance by setting

$$d_1(f_1, f_2) = \int_{[0,1]} |f_1(x) - f_2(x)|\,\mathrm{d}x.$$

The domain of integration is thus reduced to [0, 1] even when the functions are defined on a much larger domain such as $\mathbb{R}$. For all $\mathcal{F} \subset \mathbb{L}^1(\mathbb{R})$, and $f \in \mathbb{L}^1(\mathbb{R})$, we set

$$d_1(f, \mathcal{F}) = \inf_{g \in \mathcal{F}} d_1(f, g).$$

The complex conjugate of a number x is $\overline{x}$. The notation $|I|$ may refer either to the Lebesgue measure or the cardinality of I. The supremum norm of a function f is

$$\|f\|_\infty = \sup_{x \in \mathbb{R}} |f(x)|$$

and its derivative (provided it exists) is denoted by f'. The notations $c, c', \boldsymbol{C}, \boldsymbol{C}', \ldots$ are used for quantities that may change from line to line. They are usually constants (that is numbers), but may sometimes depend on some parameters. In that case, the dependency will be specified in the text.

This paper is organized as follows. In Section 2, we present an elementary approach based on an histogram-type estimator. We carry out our main results in Section 3. The description of our general procedure is deferred to Section 4. Section 5 gathers the proofs.

2. Preliminaries

This section gathers some preliminary results. We introduce and comment our main assumption on the error density q. Then, we discuss the estimation of the probability that the variable X belongs to a given set and use the obtained results to design a first histogram estimator.

2.1. Assumptions on the error density

All along the paper, we will assume that the variables ε_i involved in the model (1) admits a density q satisfying the following requirement. This assumption will always assumed to be met in the next sections (except in Proposition 4 below, but this will be specified).

Assumption 1. The Fourier transform $q^\star$ of q does not vanish. Moreover, there exist three constants κ_1, κ_2 and $\beta \in (0, 1/2)$ such that for all $t \in \mathbb{R}$,

$$|q^\star(t)|^{-2} \leq \kappa_1 + \kappa_2 |t|^{2\beta}. \tag{2}$$

Such kind of assumption is quite classical in the statistical literature (see, e.g., [26, Chapter 2]). The fact that $q^\star$ is not allowed to vanish entails in particular that $f_0^\star$ can be recovered from $(f_0 * q)^\star$ at any frequency and hence ensures that the problem is identifiable. Such a constraint is for instance satisfied when q is even and its derivative q' on $(0, +\infty)$ exists and is strictly increasing (see [32]). The difficulty (expressed for instance in terms of convergence rates) of the deconvolution problem is often measured through the behavior of $q^\star(t)$ for large value of t. The condition (2) indicates that $q^\star$ has a polynomial behavior: the deconvolution problem is said in this case to be mildly ill-posed. We restrict our attention to the specific case where the polynomial exponent β is smaller than $1/2$.

The following proposition provides examples of densities satisfying such a requirement. In particular, such examples correspond to densities having a singularity on a given point which appears to be of first mathematical interest. The proof of Proposition 1 is postponed to Section 5.1.

Proposition 1. *Let $\varphi: \mathbb{R} \to \mathbb{R}$ be a map such that $\varphi(0)$ is positive and whose derivative exists, is bounded and integrable on $\mathbb{R}$. Let q be a density of the form $q(x) = |x|^{\beta-1}\varphi(x)$ with $\beta \in (0, 1/2)$. Then if $q^\star$ does not vanish, q satisfies Assumption 1.*

2.2. Estimation of a probability

Our procedures are heavily based on probability estimations of the form

$$\mathbb{P}(X \in I) = \int_I f_0$$

for some $I \subset \mathbb{R}$. The estimation of these terms seems to be tricky when I is only assumed to be a Borel set, but is more easy when I is an interval, or more generally a union of intervals. This has been studied for instance in [15, 26], and we will adopt here the estimation strategy described in [26, Section 2.7.2].

In the following, for any $d \in \mathbb{N}^\star$, we denote by $\mathcal{I}_d$ the collection of unions of at most d intervals included in [0, 1]. For any $I \in \mathcal{I}_d$, the function

$$\phi_I : t \to \overline{\mathbb{1}_I^\star(t)} \times [q^\star(t)]^{-1}$$

belongs to $\mathbb{L}^2(\mathbb{R})$ according to Assumption 1 (this uses $\beta < 1/2$). Its Fourier transform $\phi_I^\star$ is therefore a square integrable function almost everywhere finite. It is also real-valued as ϕ_I is Hermitian. Since Y is absolutely continuous with respect to the Lebesgue measure, the random variable $Z(I) = (1/(2\pi))\phi_I^\star(Y)$ is real-valued (and finite) on an event of probability 1. In the remaining part of the paper, the random variables can be defined almost surely without specifying it again. They can always be modified on zero probability events without changing our results.

When $\| f_0 * q\|_\infty$ is finite, $Z(I)$ has moments up to second order, see Proposition 2 below. We may moreover apply Plancherel isometry to get

$$\begin{aligned}
\mathbb{E}[Z(I)] &= \int \phi_I^\star(y)(f_0 * q)(y)\,\mathrm{d}y \\
&= \frac{1}{2\pi}\int \phi_I(t) f_0^\star(t) q^\star(t)\,\mathrm{d}t \\
&= \frac{1}{2\pi}\int \overline{\mathbb{1}_I^\star(t)} f_0^\star(t)\,\mathrm{d}t \\
&= \mathbb{P}(X \in I),
\end{aligned} \tag{3}$$

where the last equality comes from Plancherel isometry when $f_0 \in \mathbb{L}^2(\mathbb{R})$ and from an additional density argument otherwise, see Section 5.2. We deduce that

$$Z_n(I) = \frac{1}{2\pi n}\sum_{j=1}^{n} \phi_I^\star(Y_j) \tag{4}$$

is an unbiased estimator for the term $\mathbb{P}(X \in I)$. A control on its variance is given by the proposition below. Its proof is postponed to Section 5.3.

Proposition 2. *For all union of at most d intervals $I \in \mathcal{I}_d$ of* $[0,1]$, *we have* $Z(I) \in \mathbb{R}$ *and*

$$\mathbb{E}\big[(Z(I))^2\big] \leq \|f_0 * q\|_\infty \big[\kappa_1 d|I| + c_\beta \kappa_2 d^{1+2\beta}|I|^{1-2\beta}\big], \tag{5}$$

where c_β only depends on β.

Let us mention that the condition $\beta < 1/2$ plays an important role in the estimation of probabilities in deconvolution. The parametric rate of convergence obtained above is impossible to get when $\beta > 1/2$ without additional assumptions on the density. We refer to [15] and [26, Section 2.7.2] for results when f_0 is smooth and $\beta > 1/2$. Note that the estimation rate they obtain depends on the smoothness index of the density (the more irregular the density, the slower the rate).

2.3. Estimation of f_0 using histograms

In this paper, we are interested in the estimation of the density f_0 of the X_i rather than a probability in itself. For pedagogical reasons, we present below an histogram estimator and give a $\mathbb{L}^1$ risk bound. The following results should be seen as a foretaste of our main contributions that will be presented in Sections 3 and 4.

Let m be a (finite) partition of [0, 1] into intervals of positive lengths. The integral of f_0 over each interval I included in m can be estimated thanks to the expression (4). Gathering all these estimations, we can define an histogram estimator $\widehat{f}_m$ of f_0 as follows:

$$\widehat{f}_m = \sum_{I \in m} \frac{Z_n(I)}{|I|} \mathbb{1}_I. \tag{6}$$

By using Proposition 2, we get a control (in expectation) of its $\mathbb{L}^1$-risk as displayed in the following proposition. The proof is deferred to Section 5.4.

Proposition 3. *Let m be a (finite) partition of* $[0,1]$ *into intervals of positive lengths, and*

$$\mathcal{F}_m = \left\{ \sum_{I \in m} \alpha_I 1_I, \ (\alpha_I)_{I \in m} \ \in [0, +\infty)^{|m|}, \ \sum_{I \in m} \alpha_I |I| = 1 \right\} \tag{7}$$

be the collection of piecewise constant densities on m. Then, the above histogram estimator satisfies

$$\mathbb{E}\big[d_1(f_0, \widehat{f}_m)\big] \leq 2 d_1(f_0, \mathcal{F}_m) + \sqrt{\|f_0 * q\|_\infty \frac{\kappa_1 |m| + c_\beta \kappa_2 |m|^{1+2\beta}}{n}}, \tag{8}$$

where c_β only depends on β.

This upper bound is non-asymptotic and shares similarities with results obtained in the direct case (which would roughly correspond to $\beta = 0$). In particular, our bound is composed of a bias term $d_1(f_0, \mathcal{F}_m)$ and an estimation component. The supremum norm of $f_0 * q$ that appears in this inequality can always be related to a suitable $\mathbb{L}^s$ norm of f_0. Indeed, by Young's inequality,

$$\|f_0 * q\|_\infty \leq \inf_{s\geq 1}\left\{\|f_0\|_s \|q\|_{s/(s-1)}\right\}.$$

Since q is not bounded, we cannot get an upper bound of $\|f_0 * q\|_\infty$ independent of f_0. Nevertheless, we always have the rough upper-bound

$$\|f_0 * q\|_\infty \leq \|f_0\|_\infty$$

as q is a density. More generally, if q satisfies our conditions of Proposition 1 with a function φ that decreases fast enough at infinity, $\|q\|_{s/(s-1)}$ is finite as soon as $s > 1/\beta$.

Let us mention that there are many solutions in the literature to estimate a density, whether β is smaller or larger than $1/2$ (the usual solution to deal with larger values of β is to impose a certain regularity on the estimator; see [13, Assumption Kvar(β)], for instance). However, we will see in the next section that (8) is specific to small values of β.

2.4. Piecewise constant functions

The previous result applies in particular to densities f_0 that are piecewise constant on a known partition m of $[0, 1]$. We may indeed consider $M > 0$ and get an upper-bound on the minimax risk over

$$\mathcal{F}_m(M) = \left\{f \in \mathcal{F}_m,\ \|f * q\|_\infty \leq M\right\}.$$

More precisely, the bias term in (8) vanishes when $f_0 \in \mathcal{F}_m(M)$, which leads to

$$\sup_{f_0 \in \mathcal{F}_m(M)} \mathbb{E}\left[d_1(f_0, \hat{f}_m)\right] \leq c M^{1/2} \frac{|m|^{\beta+1/2}}{\sqrt{n}}, \tag{9}$$

where the multiplicative factor c depends on $\kappa_1, \kappa_2, \beta$ only. As in the direct case, we observe here that the optimal estimation rate of a piecewise constant density is $n^{-1/2}$.

It is noteworthy that this upper-bound depends on the partition only through its cardinal and not on the thinness of its elements. It turns out that such a result may be incorrect for larger values of β. The following proposition is the only one in this paper that does not suppose Assumption 1. It is proved in Section 5.5.

Proposition 4. *Suppose that q is bounded and there exists $\tau > 0$ such that for all $x > 0$, then*

$$\int \left(\sqrt{q(t-x)} - \sqrt{q(t)}\right)^2 \mathrm{d}t \leq \tau^2 x^2. \tag{10}$$

Consider $n > 2/(25\tau^2)$ *and the partition* m *of size* 3 *of* $[0,1]$ *defined by*

$$m = \left\{\left[0, 1/(\tau\sqrt{50n})\right), \left[1/(\tau\sqrt{50n}), 2/(\tau\sqrt{50n})\right), \left[2/(\tau\sqrt{50n}), 1\right]\right\}.$$

Then,

$$\inf_{\hat{f}} \sup_{f_0 \in \mathcal{F}_m(2)} \mathbb{E}\left[d_1(f_0, \hat{f})\right] \geq 0.14 \min\left\{1, 1/\|q\|_\infty\right\},$$

where the infimum is taken over all possible estimators of f_0.

The assumption in this proposition requires that $\sqrt{q}$ is sufficiently regular, see [22, Chapter 1, Theorem 7.6]. For example, the Laplace distribution satisfies (10) with $\tau^2 = 1/4$. This result entails that consistent estimation is not possible uniformly over $\mathcal{F}_m(2)$ provided the partition contains small elements. Note that it does not apply when $\beta < 1/2$ as q must be bounded. Thereby, what happens when $\beta < 1/2$ can be quite different from the general case.

2.5. Estimation of a non-increasing density

The previous approach can also be used to establish the minimax rate of a monotone density. Consider the class $\mathcal{D}$ composed of densities bounded from above by 1 and whose restriction to [0, 1] is non-increasing. It follows from [5] that there exists for all $N \geq 1$ a partition m of size N such that

$$\sup_{f_0 \in \mathcal{D}} d_1(f_0, \mathcal{F}_m) \leq C/N,$$

where C is a universal constant. We therefore deduce from (8) and an adequate choice of N,

$$\sup_{f_0 \in \mathcal{D}} \mathbb{E}\left[d_1(f_0, \hat{f}_m)\right] \leq c n^{-1/(2\beta+3)},$$

where c depends on $\kappa_1, \kappa_2, \beta$ only. The proposition below, to be proved in Section 5.6, shows that this rate is optimal:

Proposition 5. *Suppose Assumption 1 is met and that there exist* κ_1', κ_2', κ_3' *such that for all* $t \in \mathbb{R}$,

$$|q^\star(t)|^{-2} \geq \kappa_1' + \kappa_2'|t|^{2\beta} \quad \text{and} \quad |(q^\star)'(t)|^2 \leq \kappa_3'|t|^{-2\beta-2},$$

where $(q^\star)'$ *denotes the derivative of* $q^\star$. *Then, there exists* c *depending only on* q *such that*

$$\inf_{\hat{f}} \sup_{f_0 \in \mathcal{D}} \mathbb{E}\left[d_1(f_0, \hat{f})\right] \geq c n^{-1/(2\beta+3)},$$

where the infimum is taken over all possible estimators $\hat{f}$ *of* f_0.

3. Non-asymptotic risk bounds

We propose below a multi-purpose estimation procedure that leads to more general results than those displayed in Section 2. For the ease of exposition, we concentrate our attention on the results and defer the presentation of the procedure to Section 4.

3.1. Assumption and main result

In the following, $\mathcal{F}$ denotes a collection of non-negative and integrable maps vanishing outside [0, 1] and given by the statistician. This collection can be seen as a translation of the a priori available knowledge on f_0. It is assumed to satisfy the condition below:

Assumption 2. Every $f \in \mathcal{F}$ may be associated to an integer d_f such that the following assertion holds true: for all $g \in \mathcal{F}$, the set

$$[f > g] = \{x \in [0,1],\ f(x) > g(x)\}$$

is a union of at most $\min\{d_f, d_g\}$ intervals. Moreover, there exists an at most countable collection $\overline{\mathcal{F}} \subset \mathcal{F}$ such that: for all $f \in \mathcal{F}$ and $\varepsilon > 0$, there exists $f' \in \overline{\mathcal{F}}$ satisfying $d_1(f, f') \leq \varepsilon$ and $d_{f'} \leq d_f$.

Under this assumption, it is possible to design an estimator $\hat{f}$ (defined in Section 4) whose performances are given by the following theorem.

Theorem 6. *Let $\mathcal{F}$ be a model fulfilling Assumption 2. Then, there exists an estimator $\hat{f}$ satisfying*

$$\mathbb{E}\left[d_1(f_0, \hat{f})\right] \leq 5 \inf_{f \in \mathcal{F}} \left\{ d_1(f_0, f) + c\sqrt{\|f_0 * q\|_\infty \frac{(\kappa_1 d_f + c_\beta \kappa_2 d_f^{1+2\beta}) \log(n d_f)}{n}} + \frac{1}{n} \right\}. \quad (11)$$

In the above inequality, c is a universal constant and c_β only depends on β.

In order to keep this paper to a reasonable size, we illustrate this risk bound by giving only three examples. Other examples are covered in the preprint version of this manuscript [25].

3.2. Estimation of a piecewise constant density

We start by revisiting the problem of estimating a piecewise constant function. In Section 2.4, we assumed that the density was piecewise constant on a known partition m, which led to a minimax bound on $\mathcal{F}_m(M)$. The above theorem allows us to generalize this result by no longer assuming that the partition is known.

To be a little more precise, define the collection $\mathcal{M}_r$ gathering all the partitions m of the form

$$m = \big\{[0, x_1], (x_1, x_2], (x_2, x_3], \dots, (x_{r-1}, 1]\big\}, \tag{12}$$

where $x_1 < x_2 < \cdots < x_{r-1}$ are $r-1$ real numbers (with the convention that $m = \{[0,1]\}$ when $r = 1$). Set,

$$\mathcal{F}_r(M) = \bigcup_{m \in \mathcal{M}_r} \mathcal{F}_m(M), \quad \mathcal{F}_r = \bigcup_{M>0} \mathcal{F}_r(M).$$

We may check that Assumption 2 is fulfilled with $\mathscr{F} = \mathcal{F}_r$, $d_f = 2r$. Theorem 6 hence gives an estimator $\hat{f}$ such that: for all $M > 0$ and n large enough,

$$\sup_{f_0 \in \mathcal{F}_r(M)} \mathbb{E}\big[d_1(f_0, \hat{f})\big] \le c M^{1/2} \frac{r^{1/2+\beta} \log^{1/2}(rn)}{n^{1/2}}.$$

The term c depends on $\kappa_1, \kappa_2, \beta$ only. If we put aside the logarithmic factor, we observe that the upper-bound coincides with the one obtained for the histogram estimator $\hat{f}_m$ on the much smaller class $\mathcal{F}_m(M)$. The estimator $\hat{f}$, however, does not require the knowledge of m and is therefore adaptive. The logarithmic factor is most likely superfluous even though a logarithmic term is involved in the minimax bounds on $\mathcal{F}_r$ in direct estimation under the Hellinger and $\mathbb{L}^2$ losses, see [8].

3.3. Estimation of a unimodal density

The histogram approach leads to optimal rates when the density is non-increasing and bounded. It can easily be adapted to estimate a unimodal density when the mode is known. When the mode is unknown, Theorem 6 can be used instead.

We introduce the collection

$$\begin{aligned}\mathcal{U} = \big\{f,\ f\mathbb{1}_{[0,1]} = f_1\mathbb{1}_{[0,x]} + f_2\mathbb{1}_{(x,1]}, &\text{ where } f_1 \text{ is non-decreasing on } [0,x],\\ &\text{where } f_2 \text{ is non-increasing on } (x,1], \text{ and where } x \in (0,1)\big\}\end{aligned} \tag{13}$$

composed of maps whose restriction to $[0,1]$ is unimodal.

We define for all functions f and all intervals I,

$$V_I(f) = \sup_{a,b \in I} |f(b) - f(a)|. \tag{14}$$

We then set for all $f \in \mathcal{U}$,

$$L_1(f) = \inf_{\substack{x \in (0,1),\\ f_1, f_2 \text{ such that}\\ f\mathbb{1}_{[0,1]} = f_1\mathbb{1}_{[0,x]} + f_2\mathbb{1}_{(x,1]}}} \big\{\log^{1/2}\big(1 + xV_{[0,x]}(f_1)\big) + \log^{1/2}\big(1 + (1-x)V_{(x,1]}(f_2)\big)\big\}^2.$$

Proposition 3 of [1] shows that Assumption 2 is fulfilled for the set

$$\mathcal{F} = \bigcup_{r=1}^{\infty} \mathcal{U} \cap \mathcal{F}_r \tag{15}$$

with d_f proportional to r when $f \in \mathcal{U} \cap \mathcal{F}_r$. Theorem 6 can therefore be applied to define an estimator $\hat{f}$. The main properties of this estimator are stated in the corollary below whose proof ensues from elementary approximation results (see [25, Appendix C], for instance).

Corollary 1. *There exists an estimator $\hat{f}$ satisfying for all $M, L > 0$, and n large enough,*

$$\sup_{\substack{f_0 \in \mathcal{U}, \\ L_1(f_0) \le L, \|f_0 * q\|_\infty \le M}} \mathbb{E}\left[d_1(f_0, \hat{f})\right] \le c L^{(2\beta+1)/(2\beta+3)} \left(M \frac{\log n}{n}\right)^{1/(2\beta+3)}.$$

Moreover, the same estimator satisfies for all $M, r > 0$, and n large enough,

$$\sup_{\substack{f_0 \mathbb{1}_{[0,1]} \in \mathcal{U} \cap \mathcal{F}_r, \\ \|f_0 * q\|_\infty \le M}} \mathbb{E}\left[d_1(f_0, \hat{f})\right] \le c \sqrt{M \frac{r^{1+2\beta} \log(rn)}{n}}.$$

In the above inequalities, c depends on $\kappa_1, \kappa_2, \beta$ only.

We thus get the same rate of convergence (up to log factors) as the one which could be reached by a histogram estimator (which however requires the a priori knowledge of the mode). The second inequality shows that the rate of the estimator is much better when the density is, in reality, piecewise constant. This phenomenon is fully automatic and also appears for some other estimators, such as, for instance the Grenander estimator in the direct setting, or the least squares estimator in isotonic regression.

3.4. Estimation of a concave density

We now consider stronger assumptions on the shape of f_0 in order to get faster rates of convergence.

We define the collection $\mathcal{C}$ gathering all the densities whose restriction to [0, 1] is concave. A function f in $\mathcal{C}$ is derivable on both the left and right-hand sides. Such a derivative (on the right or the left, according to the reader's choice) is denoted as f'. Then, we set

$$L_2(f) = \log^2\left(1 + \sqrt{V_{(0,1)}(f')}\right).$$

We define for $r \ge 1$, the collection

$$\mathcal{A}_r = \left\{ \sum_{I \in m} f_I \mathbb{1}_I, \text{ where } m \in \mathcal{M}_r, \text{ and where } f_I \text{ is a non-negative affine function} \right\}$$

composed of piecewise affine functions. It follows from Proposition 5 of [1] that the set $\mathcal{F}$ defined as

$$\mathcal{F} = \bigcup_{r=1}^{\infty} (\mathcal{C} \cap \mathcal{A}_r), \tag{16}$$

satisfies Assumption 2 with d_f proportional to r when $f \in \mathcal{C} \cap \mathcal{A}_r$.

We may thus apply Theorem 6 to this model. By using suitable approximation results, see [25, Appendix C], we deduce:

Corollary 2. *There exists an estimator $\hat{f}$ satisfying for all $M, L > 0$, and n large enough,*

$$\sup_{\substack{f_0 \in \mathcal{C}, \\ L_2(f_0) \le L,\ \|f_0 * q\|_\infty \le M}} \mathbb{E}\left[d_1(f_0, \hat{f})\right] \le c L^{(2\beta+1)/(2\beta+5)} \left(M \frac{\log n}{n}\right)^{2/(2\beta+5)},$$

where c depends on $\kappa_1, \kappa_2, \beta$ only.

The convergence rate of our estimator is therefore faster under a concavity constraint than under a monotonicity constraint. Although the corollary assumes that f_0 is concave, the same result can be established when f_0 is convex. We now claim that the above rate is optimal, within a logarithmic factor. We indeed show in Section 5.7:

Proposition 7. *Suppose Assumption 1 is met and that there exist κ_1', κ_2', κ_3' such that for all $t \in \mathbb{R}$,*

$$|q^\star(t)|^{-2} \ge \kappa_1' + \kappa_2' |t|^{2\beta} \quad \text{and} \quad |(q^\star)'(t)|^2 \le \kappa_3' |t|^{-2\beta-2},$$

where $(q^\star)'$ is the derivative of $q^\star$. Then, there exists c depending only on q such that for n large enough,

$$\inf_{\hat{f}} \sup_{\substack{f_0 \in \mathcal{C} \\ L_2(f_0) \le 2,\ \|f_0\|_\infty \le 2}} \mathbb{E}\left[d_1(f_0, \hat{f})\right] \ge c(1/n)^{2/(2\beta+5)},$$

where the infimum is taken over all possible estimators $\hat{f}$ of f_0.

4. Estimation procedure

This section is devoted to the construction of the estimator $\hat{f} \in \mathcal{F}$ leading to the results displayed in Theorem 6.

4.1. Probability estimators

In Section 2.2, we proposed simple estimators $Z_n(I)$ of $\mathbb{P}(X \in I)$. Unfortunately, the only deviations bounds that can be established for these quantities come from Chebyshev inequality (as $Z(I)$ may not admit moments of order larger than 2). Such bounds are too rough to be used in the construction of our estimator $\widehat{f}$ of f_0. This is why we propose in the following new probability estimators for which uniform bounds in deviation can be proved. Their constructions involve different steps.

Step 1. We consider an interval I of [0, 1] and begin by defining a median of means estimator of the probability $\mathbb{P}(X \in I)$. We refer, e.g., to [18] for more details regarding this approach.

Let $\delta > 0$ be a parameter whose value will be specified later on. When either $\delta \geq n-1$ or $|I| = 0$, we set $\widehat{Z}_\delta(I) = 0$. Otherwise, we split the data $(Y_1, \dots, Y_n)$ into $r \in (\delta, \delta+1]$ parts $b_1, \dots, b_r$. Each part should be approximately of the same size and more precisely such that $|b_k| \in (n/r, n/r+1]$. For each $k \in \{1, \dots, r\}$, we define the empirical mean based on the observations in the k-th block only:

$$Z_{b_k}(I) = \frac{1}{2\pi |b_k|} \sum_{Y_j \in b_k} \phi_I^\star(Y_j).$$

We then define $\widehat{Z}_\delta(I)$ as any empirical median of $\{Z_{b_1}(I), \dots, Z_{b_r}(I)\}$, that is as any real number such that

$$\left|\left\{k \in \{1, \dots, r\},\ Z_{b_k}(I) \leq \widehat{Z}_\delta(I)\right\}\right| = \left|\left\{k \in \{1, \dots, r\},\ Z_{b_k}(I) \geq \widehat{Z}_\delta(I)\right\}\right|.$$

Step 2. We broaden the previous definition of $\widehat{Z}_\delta(I)$ to unions I of intervals. In the sequel, we denote $\mathcal{I}_\infty = \cup_{d=1}^\infty \mathcal{I}_d$ and the closure of a set I by $\overline{I}$.

Every union of intervals $I \in \mathcal{I}_\infty$ can be written as $\overline{I} = \bigcup_{j=1}^{d_I} \overline{I}_j$ where $d_I \in \mathbb{N}^\star$ and the I_j are intervals such that $\overline{I}_j \cap \overline{I}_{j'} = \emptyset$ for all $j \neq j'$. We stress that d_I and the intervals $\overline{I}_j$ are defined in a unique way. Due to the additivity of measures, it is then natural to set

$$\widehat{Z}_\delta(I) = \sum_{j=1}^{d_I} \widehat{Z}_\delta(\overline{I}_j).$$

Note that this definition is coherent with the first step as $\widehat{Z}_\delta(I) = \widehat{Z}_\delta(\overline{I})$ when I is an interval (in which case $d_I = 1$).

Step 3. One may already show that each estimator $\widehat{Z}_\delta(I)$ is close to $\mathbb{P}(X \in I)$ with high probability. However, we need this result to be true uniformly for all $I \in \mathcal{I}_\infty$. Hence, we add a technical discretization step.

For any $d, j \in \mathbb{N}^\star$, we introduce the grid

$$\mathcal{G}_{d,j} = \left\{k2^{-j}/d,\ k \in \{0, \dots, d2^j\}\right\} \subset [0,1],$$

and define the set $\mathcal{I}_{d,j} \subset \mathcal{I}_d$ of unions of at most d intervals whose endpoints lie in $\mathcal{G}_{d,j}$. For all union of intervals $I \in \mathcal{I}_\infty$, let $\pi_j(I)$ be the largest set of $\mathcal{I}_{d_I,j}$ included in I. If such set does not exist, $\pi_j(I) = \emptyset$.

Consider, $\xi \geq 1$ and $\delta_j(I) = 2\log(2jd_I) + 2\log(1 + d_I 2^{j+1})$ for all $(I, j) \in \mathcal{I}_\infty \times \mathbb{N}^\star$. We define for all $I \in \mathcal{I}_\infty$, our final estimator $\widehat{Z}_{n,\xi}(I)$ of $\mathbb{P}(X \in I)$ by

$$\widehat{Z}_{n,\xi}(I) = \widehat{Z}_{\xi+\delta_1(I)}(\pi_1(I)) + \sum_{j=1}^{\infty} \widehat{Z}_{\xi+\delta_j(I)}(\pi_{j+1}(I) \setminus \pi_j(I)).$$

This sum is composed of a finite number of terms as $\xi + \delta_j(I)$ exceeds $n - 1$ for j large enough. We prove in Section 5.8 the following result.

Proposition 8. *Consider $\xi \geq 1$. Then, there is an event of probability lower bounded by $1 - e^{-\xi}$ on which: for all $I \in \mathcal{I}_\infty$,*

$$\left|\widehat{Z}_{n,\xi}(I) - \mathbb{P}(X \in I)\right| \leq C\sqrt{\frac{\|f_0 * q\|_\infty(\kappa_1 d_I + c_\beta \kappa_2 d_I^{2\beta+1})(\xi + \log d_I)}{n}}.$$

In the above inequality, C denotes a universal constant and c_β only depends on β.

4.2. Estimation procedure

We consider a collection $\mathcal{F}$ of non-negative and integrable maps compactly supported on $[0, 1]$ satisfying Assumption 2 and propose a theoretical procedure in line with the combinatorial method [19] to define an estimator $\hat{f}$ satisfying (11). Examples of such collections are given in Sections 3.2, 3.3 and 3.4.

We define for $d \geq 1$, the set $\overline{\mathcal{I}}_d \subset \mathcal{I}_d$ of unions of at most d intervals of $[0, 1]$ with endpoints in $\mathbb{Q}$. For any $\xi \geq 1$, we introduce the criterion $\gamma_\xi(\cdot)$ defined for $f \in \mathcal{F}$ by

$$\gamma_\xi(f) = \sup_{I \in \overline{\mathcal{I}}_{d_f}} \left|\int_I f - \widehat{Z}_{n,\xi}(I)\right|.$$

Then, we consider an estimator $\tilde{f} \in \mathcal{F}$ satisfying

$$\gamma_\xi(\tilde{f}) \leq \inf_{f \in \mathcal{F}} \gamma_\xi(f) + 1/n, \tag{17}$$

and set

$$\hat{f} = \min\left\{1, \left(\int \tilde{f}\right)^{-1}\right\} \tilde{f}. \tag{18}$$

Assumption 2 asserts that there is a dense and at most countable suitable subset $\overline{\mathcal{F}}$ of $\mathcal{F}$. The lemma below ensures that we may always define $\tilde{f}$ as an element of this set $\overline{\mathcal{F}}$ to avoid possible measurability issues.

Lemma 1. *There exists an estimator $\tilde{f} \in \overline{\mathcal{F}}$ satisfying* (17).

Note that there is no uniqueness concerning the way where our estimator is defined. However, Theorem 6 applies for any of the estimators $\hat{f}$ defined above with $\xi = \log n$. It is proved in Section 5.10.

Let us mention that there are algorithmic issues concerning the computation of this estimator that are far beyond the scope of this paper. This is the counterpart to get a versatile procedure. This weakness also appears for instance in the original combinatorial method of [19], in the procedures based on robust tests, see [6], and in the ρ-estimation procedure [2]. As it is said in the introduction of [7], we cannot have our cake and eat it.

5. Proofs

5.1. Sketch of the proof of Proposition 1

Define for $t \in \mathbb{R}$,

$$I(t) = \int_{-1}^{1} |x|^{\beta-1}\, e^{itx}\, \mathrm{d}x.$$

Then elementary computations give

$$|t|^{\beta} I(t) = 2 \int_{0}^{|t|} y^{\beta-1} \cos(y)\, \mathrm{d}y.$$

We use classical results on Fourier cosine transforms to get

$$|t|^{\beta} I(t) \to 2\cos(\pi\beta/2)\Gamma(\beta) \quad \text{as } |t| \to +\infty.$$

In the above equality, $\Gamma(\cdot)$ denotes the Gamma function. In particular, the limit is finite and non-zero.

Setting $\psi(x) = |x|^{\beta-1}(\varphi(x) - \varphi(0))$ for all $x \in \mathbb{R}$, we can write

$$\int_{-1}^{1} q(x)e^{itx}\, \mathrm{d}x = \varphi(0)I(t) + \int_{-1}^{1} \psi(x)e^{itx}\, \mathrm{d}x.$$

Note that ψ is bounded on $[-1, 1]$. Moreover, ψ admits a derivative ψ' on $\mathbb{R} \setminus \{0\}$ such that

$$\begin{aligned} |\psi'(x)| &\leq (1-\beta)|x|^{\beta-2} \int_{[0,|x|]} |\varphi'(t)|\, \mathrm{d}t + |x|^{\beta-1}\, |\varphi'(x)| \\ &\leq 2|x|^{\beta-1}\|\varphi'\|_{\infty}, \end{aligned}$$

where φ' is the derivative of φ. Therefore, ψ' is integrable on $[-1, 1]$. Integration by parts then shows

$$\int_{-1}^{1} \psi(x)e^{itx} \, \mathrm{d}x = \mathcal{O}(1/t) \quad \text{as } |t| \to +\infty.$$

Since φ' is integrable, φ is bounded and $q(x)$ tends to 0 when $|x| \to +\infty$. We also have,

$$\begin{aligned}\int_{(-\infty,-1]\cup[1,+\infty)} |q'(x)| &\leq (1-\beta)\int_{(-\infty,-1]\cup[1,+\infty)} |x|^{\beta-2}\,|\varphi(x)|\,\mathrm{d}x \\ &\quad + \int_{(-\infty,-1]\cup[1,+\infty)} |x|^{\beta-1}\,|\varphi'(x)|\,\mathrm{d}x \\ &\leq (1-\beta)\int_{(-\infty,-1]\cup[1,+\infty)} |x|^{\beta-2}\,|\varphi(x)|\,\mathrm{d}x \\ &\quad + \sqrt{\int_{(-\infty,-1]\cup[1,+\infty)} |x|^{2\beta-2}\,\mathrm{d}x \int |\varphi'(x)|^2\,\mathrm{d}x}.\end{aligned}$$

Since φ' is bounded and integrable, it belongs to $\mathbb{L}^2$. The above integral is hence finite. We deduce as before,

$$\int_{(-\infty,-1]\cup[1,+\infty)} q(x)e^{itx} \, \mathrm{d}x = \mathcal{O}(1/t).$$

Hence,

$$q^\star(t) = \varphi(0)I(t) + \mathcal{O}(1/t) \quad \text{as } |t| \to +\infty,$$

from which we deduce that $(|t|^\beta |q^\star(t)|)^{-1}$ admits a finite limit when $|t| \to +\infty$. As $q^\star$ does not vanish and is continuous, $|q^\star(t)|^{-1}$ is bounded above on $[-1, 1]$ and $(|t|^\beta |q^\star(t)|)^{-1}$ is bounded above on $(-\infty, -1] \cup [1, +\infty)$. This ends the proof. ■

5.2. Proof of (3) when $f_0 \notin \mathbb{L}^2(\mathbb{R})$

Without loss of generality, we may suppose that I is an interval say $I = [a, b]$. Let K be a density belonging to $\mathbb{L}^2(\mathbb{R})$, $h > 0$ and $K_h(\cdot) = (1/h)K(\cdot/h)$. Then, if F_0 denotes the cumulative distribution function of X,

$$\begin{aligned}&\int \mathbb{1}_I(t)(f_0 * K_h)(t)\,\mathrm{d}t - \int \mathbb{1}_I(t) f_0(t)\,\mathrm{d}t \\ &\qquad = \int K(x)\big(F_0(b - xh) - F_0(b) + F_0(a) - F_0(a - xh)\big)\,\mathrm{d}x.\end{aligned}$$

By using the dominated convergence theorem, we deduce

$$\lim_{h\to 0} \int \mathbb{1}_I(t)(f_0 * K_h)(t)\,\mathrm{d}t - \int \mathbb{1}_I(t) f_0(t)\,\mathrm{d}t = 0.$$

Yet, $f_0 * K_h \in \mathbb{L}^2(\mathbb{R})$ as it is the case for K_h, and by Plancherel isometry,

$$\lim_{h\to 0} \frac{1}{2\pi} \int \overline{\mathbb{1}_I^\star(t)} f_0^\star(t) K_h^\star(t)\,\mathrm{d}t = \int \mathbb{1}_I(t) f_0(t)\,\mathrm{d}t.$$

Note that

$$\overline{\mathbb{1}_I^\star(t)} f_0^\star(t) = \left(\overline{\mathbb{1}_I^\star(t)}[q^\star(t)]^{-1}\right) \times \left(q^\star(t) f_0^\star(t)\right).$$

This map is a product of two square integrable functions ($q^\star f_0^\star$ belongs to $\mathbb{L}^2(\mathbb{R})$ as $q * f_0$ does since this density is bounded) and is therefore integrable. We may thus use $\lim_{h\to 0} K_h^\star(t) = K^\star(0) = 1$ and the dominated convergence theorem to get

$$\int \overline{\mathbb{1}_I^\star(t)} f_0^\star(t)\,\mathrm{d}t = \int \mathbb{1}_I(t) f_0(t)\,\mathrm{d}t$$

as wished. ∎

5.3. Proof of Proposition 2

For any $I \in \mathcal{I}_d$, we can always write $I = \cup_{j=1}^k I_j$, where $k \le d$ and where the I_j are disjoint intervals of [0, 1]. Then,

$$\begin{aligned}\mathbb{E}\left[(Z(I))^2\right] &= \frac{1}{4\pi^2} \int (\phi_I^\star(y))^2 (f_0 * q)(y)\,\mathrm{d}y \\ &\le \frac{\|f_0 * q\|_\infty}{4\pi^2} \int (\phi_I^\star(y))^2\,\mathrm{d}y.\end{aligned}$$

We successively use the Fourier–Plancherel theorem and Cauchy–Schwarz inequality to get

$$\begin{aligned}\mathbb{E}\left[(Z(I))^2\right] &\le \frac{\|f_0 * q\|_\infty}{2\pi} \int \left|\frac{\overline{\mathbb{1}_I^\star(t)}}{q^\star(t)}\right|^2 \mathrm{d}t \\ &= \frac{\|f_0 * q\|_\infty}{2\pi} \int \left|\sum_{j=1}^k \frac{\overline{\mathbb{1}_{I_j}^\star(t)}}{q^\star(t)}\right|^2 \mathrm{d}t \\ &\le \frac{\|f_0 * q\|_\infty}{2\pi} \times k \times \sum_{j=1}^k \int \left|\frac{\mathbb{1}_{I_j}^\star(t)}{q^\star(t)}\right|^2 \mathrm{d}t.\end{aligned}$$

For any $j \in \{1, \ldots, k\}$ and $t \in \mathbb{R}$, we have

$$|\overline{\mathbb{1}_{I_j}^\star(t)}| = |\mathbb{1}_{I_j}^\star(t)| = 2\frac{|\sin(t|I_j|/2)|}{|t|}$$

since I_j is an interval of finite length $|I_j|$. Hence, thanks to Assumption 1,

$$\begin{aligned}\int \left|\frac{\overline{\mathbb{1}^\star_{I_j}(t)}}{q^\star(t)}\right|^2 \mathrm{d}t &= 4\int \frac{\sin^2(t|I_j|/2)}{t^2|q^\star(t)|^2}\,\mathrm{d}t \\ &\leq 4\int \frac{\sin^2(t|I_j|/2)}{t^2}\big(\kappa_1+\kappa_2|t|^{2\beta}\big)\,\mathrm{d}t \\ &\leq 4\kappa_1\big(|I_j|/2\big)\int \frac{\sin^2(t)}{t^2}\,\mathrm{d}t + 4\kappa_2\big(|I_j|/2\big)^{1-2\beta}\int \sin^2(t)|t|^{2\beta-2}\,\mathrm{d}t.\end{aligned}$$

By computing these integrals,

$$\int \left|\frac{\mathbb{1}^\star_{I_j}(t)}{q^\star(t)}\right|^2 \mathrm{d}t \leq 2\kappa_1\pi|I_j| + 4\kappa_2\frac{\sin(\pi\beta)\Gamma(2\beta)}{1-2\beta}|I_j|^{1-2\beta},$$

where $\Gamma(\cdot)$ denotes the Gamma function. Finally, we get

$$\begin{aligned}\mathbb{E}[(Z(I))^2] &\leq \frac{\|f_0 * q\|_\infty}{2\pi}\times k\times\left[2\kappa_1\pi\sum_{j=1}^k|I_j| + \frac{4\kappa_2\sin(\pi\beta)\Gamma(2\beta)}{1-2\beta}\sum_{j=1}^k|I_j|^{1-2\beta}\right], \\ &\leq \frac{\|f_0 * q\|_\infty}{2\pi}\times d\times\left[2\kappa_1\pi|I| + \frac{4\kappa_2\sin(\pi\beta)\Gamma(2\beta)}{1-2\beta}d^{2\beta}|I|^{1-2\beta}\right],\end{aligned}$$

where we have used $k \leq d$, $\sum_{j=1}^k |I_j| = |I|$ and the Hölder inequality (as $\beta < 1/2$). ■

5.4. Proof of Proposition 3

Define

$$\bar{f}_m = \sum_{I\in m}\frac{\mathbb{P}(X\in I)}{|I|}\mathbb{1}_I.$$

For all $I \in m$, and $\alpha \in \mathbb{R}$,

$$\begin{aligned}\int_I\left|f_0 - \frac{1}{|I|}\int_I f_0\right| &\leq \int_I |f_0-\alpha| + \left|\alpha|I| - \int_I f_0\right| \\ &= \int_I |f_0-\alpha| + \left|\int_I(\alpha - f_0)\right| \\ &\leq 2\int_I |f_0-\alpha|.\end{aligned}$$

Therefore,

$$d_1(f_0,\bar{f}_m) = \sum_{I\in m}\int_I\left|f_0 - \frac{1}{|I|}\int_I f_0\right| \leq 2d_1(f_0,\mathcal{F}_m).$$

Moreover,

$$\begin{aligned}\mathbb{E}\big[d_1\big(\widehat{f}_m,\bar{f}_m\big)\big] &= \sum_{I\in m}\mathbb{E}\big[|Z_n(I)-\mathbb{P}(X\in I)|\big]\\ &\le \sum_{I\in m}\sqrt{\mathbb{E}\big[|Z_n(I)-\mathbb{P}(X\in I)|^2\big]}\\ &\le \frac{1}{n^{1/2}}\sum_{I\in m}\sqrt{\mathbb{E}\big[(Z(I))^2\big]}.\end{aligned}$$

By using Proposition 2 and the Hölder inequality (as $\beta<1/2$),

$$\begin{aligned}\mathbb{E}\big[d_1\big(\widehat{f}_m,\bar{f}_m\big)\big] &\le \frac{\|f_0*q\|_\infty^{1/2}}{n^{1/2}}\Bigg[\kappa_1^{1/2}\sum_{I\in m}|I|^{1/2}+c_\beta^{1/2}\kappa_2^{1/2}\sum_{I\in m}|I|^{1/2-\beta}\Bigg]\\ &\le \frac{\|f_0*q\|_\infty^{1/2}}{n^{1/2}}\big[\kappa_1^{1/2}|m|^{1/2}+c_\beta^{1/2}\kappa_2^{1/2}|m|^{1/2+\beta}\big].\end{aligned}$$

The proof then follows from the triangle inequality

$$\mathbb{E}\big[d_1\big(f_0,\widehat{f}_m\big)\big]\le d_1\big(f_0,\bar{f}_m\big)+\mathbb{E}\big[d_1\big(\widehat{f}_m,\bar{f}_m\big)\big].$$ ■

5.5. Proof of Proposition 4

Let h be the Hellinger distance defined for two densities f_1, f_2 by

$$h^2(f_1,f_2)=\frac{1}{2}\int_{\mathbb{R}}\big(\sqrt{f_1(x)}-\sqrt{f_2(x)}\big)^2\,\mathrm{d}x.$$

It follows from standard arguments (see, e.g., [31, Chapter 2]) that

$$\inf_{\widehat{f}}\sup_{f_0\in\mathcal{F}_m(2)}\mathbb{E}\big[d_1(f_0,\widehat{f})\big]\ge\frac{1}{2}\sup_{f_1,f_2\in\mathcal{F}_m(2)}d_1(f_1,f_2)\big(1-\sqrt{2n}h(f_1*q,f_2*q)\big).\tag{19}$$

Consider $p=\min\{1,1/\|q\|_\infty\}$, $a=50\tau^2$, and set

$$\begin{aligned}f_1&=p\sqrt{an}\mathbb{1}_{[0,1/\sqrt{an})}+(1-p)\mathbb{1}_{[0,1]},\\ f_2&=\frac{p\sqrt{an}}{2}\mathbb{1}_{[0,2/\sqrt{an})}+(1-p)\mathbb{1}_{[0,1]}.\end{aligned}$$

Young's Inequality implies $\|f_1*q\|_\infty\le p\|q\|_\infty+1-p\le 2$. The same result is true for f_2*q. Therefore, f_1 and f_2 both belong to $\mathcal{F}_m(2)$.

We now bound $h(f_1*q,f_2*q)$. We have,

$$f_2*q(x)-f_1*q(x)=\frac{p}{2}\int_0^1\big(q\big(x-(u+1)/\sqrt{an}\big)-q\big(x-u/\sqrt{an}\big)\big)\,\mathrm{d}u.$$

We introduce for $x \in \mathbb{R}$,

$$\varphi^2(x) = \int_0^1 \left(\sqrt{q\big(x - (u+1)/\sqrt{an}\,\big)} - \sqrt{q\big(x - u/\sqrt{an}\,\big)} \right)^2 \mathrm{d}u,$$

and decompose $f_2 * q(x) - f_1 * q(x)$ as

$$\begin{aligned} f_2 * q(x) - f_1 * q(x) &= \frac{p}{2}\varphi^2(x) \\ &+ p \int_0^1 \left(\sqrt{q\big(x - (u+1)/\sqrt{an}\,\big)} - \sqrt{q\big(x - u/\sqrt{an}\,\big)} \right) \sqrt{q\big(x - u/\sqrt{an}\,\big)}\, \mathrm{d}u. \end{aligned}$$

We apply the Cauchy–Schwarz inequality:

$$\begin{aligned} |f_2 * q(x) - f_1 * q(x)| &\le \frac{p}{2}\varphi^2(x) + p\varphi(x) \sqrt{\int_0^1 q\big(x - u/\sqrt{an}\,\big)\, \mathrm{d}u} \\ &\le \frac{p}{2}\varphi^2(x) + \varphi(x)\sqrt{p}\sqrt{f_1 * q(x)}. \end{aligned}$$

Let

$$\mathcal{X} = \left\{ x \in \mathbb{R},\ f_1 * q(x) + f_2 * q(x) \le \varphi^2(x) \right\}.$$

We deduce

$$\begin{aligned} 2h^2(f_1 * q, f_2 * q) &\le \int_{\mathcal{X}} |f_1 * q - f_2 * q| + \int_{\mathcal{X}^c} \frac{(f_1 * q - f_2 * q)^2}{f_1 * q + f_2 * q} \\ &\le \int_{\mathcal{X}} \varphi^2 + \frac{1}{2}\int_{\mathcal{X}^c} \frac{p^2\varphi^4 + 4p\varphi^2 f_1 * q}{f_1 * q + f_2 * q} \\ &\le \int_{\mathcal{X}} \varphi^2 + \frac{5}{2}\int_{\mathcal{X}^c} \varphi^2 \\ &\le \frac{5}{2}\int \varphi^2. \end{aligned}$$

Yet,

$$\begin{aligned} \int \varphi^2 &\le \sup_{u \in [0,1]} \int \left(\sqrt{q\big(x - (u+1)/\sqrt{an}\,\big)} - \sqrt{q\big(x - u/\sqrt{an}\,\big)} \right)^2 \mathrm{d}x \\ &\le 2 \sup_{u \in [0,1]} \left[\int \left(\sqrt{q\big(x - (u+1)/\sqrt{an}\,\big)} - \sqrt{q(x)} \right)^2 \mathrm{d}x \right. \\ &\qquad \left. + \int \left(\sqrt{q\big(x - u/\sqrt{an}\,\big)} - \sqrt{q(x)} \right)^2 \mathrm{d}x \right] \\ &\le 10\tau^2/(an) \end{aligned}$$

thanks to (10). We use the definition of a to get $2nh^2(f_1, f_2) \le 1/2$. We conclude by using (19) and by remarking that $d_1(f_1, f_2) = p$. ■

5.6. Proof of Proposition 5

The arguments we propose are inspired by the lower minimax bound of [4] for non-increasing functions in the noise free case. Let $p \geq 1$ whose value will be made precise at the end of the proof. We define

$$\varepsilon = \frac{\log(3/2)}{p} \leq 1, \quad u = [(1+\varepsilon)^p - 1]^{-1}, \quad \text{and} \quad \lambda = \frac{1+\varepsilon}{u(1+\varepsilon/2)}.$$

For any $j \in \{1, \dots, p\}$, we introduce $x_j = u[(1+\varepsilon)^j - 1]$, $I_j = [x_{j-1}, x_j)$ and $\ell_j = x_j - x_{j-1}$. Remark that the intervals I_j are disjoint and included in $[0, 1]$.

Now, for any $j \in \{1, \dots, p\}$, define the functions f_j and g_j as

$$f_j = \lambda(1+\varepsilon)^{-j}(1+\varepsilon/2)\mathbb{1}_{[x_{j-1}, x_j)},$$
$$g_j = \lambda(1+\varepsilon)^{-j+1}\mathbb{1}_{[x_{j-1}, (x_{j-1}+x_j)/2]} + \lambda(1+\varepsilon)^{-j}\mathbb{1}_{((x_{j-1}+x_j)/2, x_j)}.$$

We denote by φ the Cauchy density, introduce for all $\nu \in \{0, 1\}^p$,

$$\phi_\nu = (1 - \log(3/2))\varphi + \sum_{j=1}^{p} \big(\nu_j f_j + (1-\nu_j) g_j\big),$$

and gather all these functions into a set $\mathcal{F} = \{\phi_\nu,\ \nu \in \{0, 1\}^p\}$. We begin by establishing the following inequality:

Claim 1. We have $\mathcal{F} \subset \mathcal{D}$. Moreover, for any $j \in \{1, \dots, p\}$,

$$d_1(f_j, g_j) \geq \varepsilon^2/3.$$

Proof. For any $j \in \{1, \dots, p\}$, we observe that the functions f_j and g_j share the following properties:

- f_j and g_j are piecewise constant functions, compactly supported on I_j;
- for all $x \in I_j$ and $y \in I_{j-1}$,

$$\max\{f_j(x), g_j(x)\} \leq \min\{f_{j-1}(y), g_{j-1}(y)\}; \tag{20}$$

- $\int f_j = \int g_j = \varepsilon$;
- $\|f_1\|_\infty \leq 1/2$ and $\|g_1\|_\infty \leq 1/2$.

We deduce that ϕ_ν is a density, non-increasing on $[0, 1]$, and bounded by 1. Moreover,

$$d_1(f_j, g_j) = \ell_j \lambda (1+\varepsilon)^{-j} \frac{\varepsilon}{2} = \varepsilon^2 \times \frac{1}{2+\varepsilon} \geq \frac{1}{3}\varepsilon^2. \qquad \blacksquare$$

In the sequel, to underline the dependency of the density f_0 in the results, we add a subscript to the expectation. More precisely, $\mathbb{E}_f$ corresponds to the expectation with respect to the variables X_i which are assumed to be of density f. We have for all estimators $\widehat{f}$,

$$\begin{aligned}\sup_{f_0\in\mathcal{D}} \mathbb{E}_{f_0}\big[d_1\big(f_0,\widehat{f}\big)\big] &\geq \frac{1}{2^p}\sum_{\nu\in\{0,1\}^p} \mathbb{E}_{\phi_\nu}\big[d_1\big(f_0,\widehat{f}\big)\big] \\ &\geq \frac{1}{2^p}\sum_{\nu\in\{0,1\}^p}\sum_{j=1}^p \mathbb{E}_{\phi_\nu}\big[d_1\big(\widehat{f}\mathbb{1}_{I_j},\phi_\nu\mathbb{1}_{I_j}\big)\big] \\ &\geq \frac{1}{2}\inf_{\nu\in\{0,1\}^p}\sum_{j=1}^p \Big\{\mathbb{E}_{\phi_{\nu_{j,1}}}\big[d_1\big(\widehat{f}\mathbb{1}_{I_j},\big((1-\log(3/2))\varphi+f_j\big)\mathbb{1}_{I_j}\big)\big] \\ &\qquad\qquad + \mathbb{E}_{\phi_{\nu_{j,0}}}\big[d_1\big(\widehat{f}\mathbb{1}_{I_j},((1-\log(3/2))\varphi+g_j)\mathbb{1}_{I_j}\big)\big]\Big\},\end{aligned}$$

where for any $\nu\in\{0,1\}^p$ and $k\in\{0,1\}$,

$$\nu_{j,k} = (\nu_1,\ldots,\nu_{j-1},k,\nu_{j+1},\ldots,\nu_p).$$

Then, using the triangle inequality,

$$\begin{aligned}&\sup_{f_0\in\mathcal{D}} \mathbb{E}_{f_0}\big[d_1\big(f_0,\widehat{f}\big)\big] \\ &\geq \frac{1}{2}\inf_{\nu\in\{0,1\}^p}\sum_{j=1}^p d_1(f_j,g_j)\int_{\mathbb{R}^n}\min\left(\prod_{k=1}^n \phi_{\nu_{j,1}} * q(x_k), \prod_{k=1}^n \phi_{\nu_{j,0}} * q(x_k)\right)\mathrm{d}x.\end{aligned}$$

Using Claim 1, we get

$$\begin{aligned}&\sup_{f_0\in\mathcal{D}} \mathbb{E}_{f_0}\big[d_1\big(f_0,\widehat{f}\big)\big] \\ &\geq \frac{\varepsilon^2}{6}\inf_{\nu\in\{0,1\}^p}\sum_{j=1}^p \int_{\mathbb{R}^n}\min\left(\prod_{k=1}^n \phi_{\nu_{j,1}} * q(x_k), \prod_{k=1}^n \phi_{\nu_{j,0}} * q(x_k)\right)\mathrm{d}x. \qquad (21)\end{aligned}$$

For any $j\in\{1,\ldots,p\}$, we recall that the χ^2 divergence between $\phi_{\nu_{j,1}} * q$ and $\phi_{\nu_{j,0}} * q$ is defined as

$$\chi^2_\nu(j) = \int \frac{(\phi_{\nu_{j,1}} * q - \phi_{\nu_{j,0}} * q)^2}{\phi_{\nu_{j,0}} * q}.$$

It follows from standard results about distance comparisons that

$$\int_{\mathbb{R}^n}\min\left(\prod_{k=1}^n \phi_{\nu_{j,1}} * q(x_k), \prod_{k=1}^n \phi_{\nu_{j,0}} * q(x_k)\right)\mathrm{d}x \geq 1 - \big((1+\chi^2_\nu(j))^n - 1\big)^{1/2}. \qquad (22)$$

We refer to Chapter 2 of [31] for the proof of this inequality. We use the claim below whose proof is delayed after the present proof.

Claim 2. There exists c depending only on q such that

$$\sup_{1\le j\le p} \chi_\nu^2(j) \le c(1/p)^{2\beta+3}.$$

By choosing p as the smallest integer larger than $(2cn)^{1/(2\beta+3)}$, $\chi_\nu^2(j) \le 1/(2n)$. We now put (21), (22) and the elementary inequality $(1+1/(2n))^n \le e^{1/2}$ together to get

$$\sup_{f_0\in\mathcal{D}} \mathbb{E}_{f_0}\big[d_1\big(f_0,\hat{f}\big)\big] \ge \frac{1}{6}\big(1-\sqrt{e^{1/2}-1}\big)\varepsilon^2 p,$$

which gives the desired result. ∎

Proof of Claim 2. Consider some a such that $\int_{-a}^{a} q(y)\,\mathrm{d}y > 0$. For all $x \in \mathbb{R}$, $j \in \{1,\dots,p\}$ and $\nu \in \{0,1\}^p$,

$$\begin{aligned}\phi_{\nu_{j,0}} * q(x) &\ge (1-\log(3/2))\varphi * q(x),\\ &\ge \frac{1-\log(3/2)}{\pi}\int_{-a}^{a}\frac{q(y)}{1+(x-y)^2}\,\mathrm{d}y\\ &\ge \frac{1-\log(3/2)}{\pi}\int_{-a}^{a}\frac{q(y)}{1+2x^2+2a^2}\,\mathrm{d}y.\end{aligned}$$

There exists therefore c_0 (depending only on a and $\int_{-a}^{a} q(y)\,\mathrm{d}y$ and hence of q) such that

$$\phi_{\nu_{j,0}} * q(x) \ge \frac{1}{c_0(1+x^2)}.$$

Therefore, for all $j \in \{1,\dots,p\}$, $\nu \in \{0,1\}^p$

$$\begin{aligned}\chi_\nu^2(j) \le c_0\bigg[&\int \big(\phi_{\nu_{j,1}} * q(x) - \phi_{\nu_{j,0}} * q(x)\big)^2\,\mathrm{d}x\\ &+ \int x^2\big(\phi_{\nu_{j,1}} * q(x) - \phi_{\nu_{j,0}} * q(x)\big)^2\,\mathrm{d}x\bigg].\end{aligned}$$

For any $j \in \{1,\dots,p\}$, let ψ_j be the map defined as

$$\begin{aligned}\psi_j(t) &= \phi^\star_{\nu_{j,1}}(t) - \phi^\star_{\nu_{j,0}}(t)\\ &= f_j^\star(t) - g_j^\star(t)\\ &= \lambda\varepsilon\frac{\sin(t\ell_j/4)}{t(1+\varepsilon)^j}\big[e^{it(x_j+m_j)/2} - e^{it(x_{j-1}+m_j)/2}\big],\end{aligned}$$

where $m_j = (x_{j-1} + x_j)/2$ and $\ell_j = x_j - x_{j-1} = u(1+\varepsilon)^{j-1}\,\varepsilon$. Then, by using Plancherel isometry,

$$\chi_\nu^2(j) \le \frac{c_0}{2\pi}\left[\int |\psi_j q^\star|^2 + \int |\psi_j' q^\star + \psi_j (q^\star)'|^2\right]$$
$$\le \frac{c_0}{2\pi}\left[\int |\psi_j q^\star|^2 + 2\int |\psi_j (q^\star)'|^2 + 2\int |\psi_j' q^\star|^2\right]. \quad (23)$$

Moreover,

$$\int |\psi_j q^\star|^2 \le \frac{\lambda^2\varepsilon^2}{(1+\varepsilon)^{2j}} \int \left|\frac{\sin(t\ell_j/4)}{t}\right|^2 |q^\star(t)|^2\,\mathrm{d}t.$$

Elementary computations entail

$$\int |\psi_j q^\star|^2 \le c_1\lambda^2(1+\varepsilon)^{-2j}\varepsilon^2\left(\ell_j^2 + \ell_j^{2\beta+1}\right)$$
$$\le c_1'\lambda^2(1+\varepsilon)^{-2j}\varepsilon^2\ell_j^{2\beta+1}$$

as $\beta < 1/2$. Here, c_1, c_1' depend on κ_1', κ_2' and β only. Likewise,

$$\int |\psi_j (q^\star)'|^2 \le c_2\lambda^2(1+\varepsilon)^{-2j}\varepsilon^2\ell_j^{2\beta+3},$$
$$\int |\psi_j' q^\star|^2 \le c_3\lambda^2(1+\varepsilon)^{-2j}\varepsilon^2\ell_j^{2\beta+1},$$

where c_2 depends on κ_3', β only and where c_3 depends on $\kappa_1', \kappa_2', \beta$ only.

We finally deduce from (23) that

$$\chi_\nu^2(j) \le c_4\lambda^2(1+\varepsilon)^{-2j}\varepsilon^2\ell_j^{2\beta+1} \le c_5(1/p)^{2\beta+3},$$

where c_5 depends on q only. ∎

5.7. Proof of Proposition 7

We use here the same notation as those introduced in the proof of Proposition 5, except for p (which is an integer whose value, defined later on, will be different from the one in the previous proof). We set

$$\overline{\lambda}^{-1} = \left(1 + \frac{\varepsilon}{2}\right) \times \frac{1}{p}\sum_{j=1}^{p}\left[\frac{u^2}{2}\varepsilon(1+\varepsilon)^{j-2} + (1 - x_j)\frac{u}{1+\varepsilon}\right].$$

We also define for $x \in \mathbb{R}$ and $j \in \{1, \dots, p\}$:

$$\overline{f_j} = \overline{\lambda}(1+\varepsilon)^{-j}(1+\varepsilon/2)\mathbb{1}_{[x_{j-1}, x_j)},$$
$$\overline{g}_j = \overline{\lambda}(1+\varepsilon)^{-j+1}\mathbb{1}_{[x_{j-1},(x_{j-1}+x_j)/2]} + \overline{\lambda}(1+\varepsilon)^{-j}\mathbb{1}_{((x_{j-1}+x_j)/2, x_j)},$$

$$F_j(x) = \int_{-\infty}^{x} \overline{f}_j(y)\,\mathrm{d}y \quad \text{and} \quad G_j(x) = \int_{-\infty}^{x} \overline{g}_j(y)\,\mathrm{d}y.$$

For $\nu \in \{0,1\}^p$, we set

$$\overline{\varphi}_\nu(x) = \left(1 - \log(3/2) - \frac{\overline{\lambda} u^2 \varepsilon^3}{8(1+\varepsilon)^2} \sum_{j=1}^{p} (1+\varepsilon)^j (1-\nu_j)\right)(1/\sqrt{3})\varphi\big(x/\sqrt{3}\big) + \sum_{j=1}^{p} \big(\nu_j F_j(x) + (1-\nu_j) G_j(x)\big) \mathbb{1}_{[0,1]}(x).$$

We gather these functions $\overline{\varphi}_\nu$ into the set $\mathcal{H} = \{\overline{\varphi}_\nu,\ \nu \in \{0,1\}^p\}$ and prove the following claim.

Claim 3. We have $\mathcal{H} \subset \mathcal{C}$. Moreover, any function $f \in \mathcal{H}$ satisfies $\|f\|_\infty \le 2$, $\|f' \mathbb{1}_{(0,1)}\|_\infty \le 2$, where f' is the derivative of f. For all $j \in \{1, \dots, p\}$,

$$d_1(F_j, G_j) \ge 0.3\varepsilon^3.$$

Proof of Claim 3. Elementary but tedious computations give

$$\overline{\lambda}^{-1} = \frac{(1+\varepsilon/2)u}{2(1+\varepsilon)\log(3/2)} \big[2\log(3/2)(1+u) - 2 - \varepsilon\big].$$

Moreover,

$$0.3 \le \frac{\overline{\lambda} u^2}{8(1+\varepsilon)^2} \le \frac{\overline{\lambda} u^2}{8(1+\varepsilon)^2} (1+\varepsilon)^p \le 0.8. \tag{24}$$

Note now that $\varphi(\cdot/\sqrt{3})$ is concave on $[0,1]$ and

$$1 - \log(3/2) - \frac{\overline{\lambda} u^2 \varepsilon^3}{8(1+\varepsilon)^2} \sum_{j=1}^{p} (1+\varepsilon)^j (1-\nu_j) \ge 1 - \log(3/2) - 0.8\varepsilon^3 p > 0.$$

Moreover, for all $x \in [0,1]$,

$$\overline{\varphi}_\nu(x) = \left(1 - \log(3/2) - \frac{\overline{\lambda} u^2 \varepsilon^3}{8(1+\varepsilon)^2} \sum_{j=1}^{p} (1+\varepsilon)^j (1-\nu_j)\right)(1/\sqrt{3})\varphi\big(x/\sqrt{3}\big) + \int_{-\infty}^{x} \sum_{j=1}^{p} \big(\nu_j \overline{f}_j(t) + (1-\nu_j)\overline{g}_j(t)\big)\,\mathrm{d}t.$$

As in (20),

$$\max\{\overline{f}_j(x), \overline{g}_j(x)\} \le \min\{\overline{f}_{j-1}(y), \overline{g}_{j-1}(y)\},$$

for all $x \in I_j$ and $y \in I_{j-1}$. This implies that the sum in the integrate above is non-increasing and that $\overline{\varphi}_\nu$ is concave on $[0,1]$. Besides, the map in the integrate is bounded from above by $\overline{\lambda}$. Some computations show that this quantity is not greater than 0.95. In particular, $\overline{\varphi}_\nu$ is differentiable on (0, 1) and $|\overline{\varphi}_\nu'|$ is bounded from above by a constant smaller than 1.1. The mean value theorem implies that $|\overline{\varphi}_\nu(x)| \leq 1.3 < 2$ for all $x \in [0, 1]$ and the same result is true for $x \notin [0, 1]$. Now,

$$\int_{I_j} F_j(x)\,\mathrm{d}x = \overline{\lambda}(1+\varepsilon)^{-j}(1+\varepsilon/2)\frac{\ell_j^2}{2},$$

and

$$\begin{aligned}\int_0^1 F_j(x)\,\mathrm{d}x &= \int_{I_j} F_j(x)\,\mathrm{d}x + (1-x_j)F_j(x_j)\\ &= \overline{\lambda}\varepsilon(1+\varepsilon/2)\left[\frac{u^2}{2}\varepsilon(1+\varepsilon)^{j-2} + (1-x_j)\frac{u}{1+\varepsilon}\right].\end{aligned}$$

Therefore,

$$\begin{aligned}\sum_{j=1}^p \int_0^1 F_j(x)\,\mathrm{d}x &= \overline{\lambda}\varepsilon(1+\varepsilon/2)\sum_{j=1}^p\left[\frac{u^2}{2}\varepsilon(1+\varepsilon)^{j-2} + (1-x_j)\frac{u}{1+\varepsilon}\right]\\ &= \varepsilon p = \log(3/2).\end{aligned} \tag{25}$$

We define the triangle function $\zeta(x) = (1-|x|)\mathbb{1}_{[-1,1]}(x)$ and remark that

$$H_j(x) = G_j(x) - F_j(x) = \frac{\overline{\lambda}\varepsilon\ell_j}{4(1+\varepsilon)^j}\zeta\left(\frac{2(x-m_j)}{\ell_j}\right). \tag{26}$$

In particular, since $G_j(x) = F_j(x)$ for all $x \notin [0, 1]$,

$$\begin{aligned}\int_{[0,1]}(G_j - F_j) &= \int_{\mathbb{R}}(G_j - F_j)\\ &= \frac{\overline{\lambda}\varepsilon\ell_j}{4(1+\varepsilon)^j}\int_{\mathbb{R}}\zeta\left(\frac{2(x-m_j)}{\ell_j}\right)\mathrm{d}x,\\ &= \frac{\overline{\lambda}\varepsilon\ell_j^2}{8(1+\varepsilon)^j},\\ &= \frac{\overline{\lambda}u^2}{8}\varepsilon^3(1+\varepsilon)^{j-2}.\end{aligned}$$

This result, combined with (25), entails that $\overline{\varphi}_\nu$ is a density. It also gives as $\zeta \geq 0$,

$$d_1(F_j, G_j) = \frac{\overline{\lambda}u^2}{8(1+\varepsilon)^2}(1+\varepsilon)^j\varepsilon^3 \geq \frac{\overline{\lambda}u^2}{8(1+\varepsilon)^2}\varepsilon^3 \geq 0.3\varepsilon^3. \qquad \blacksquare$$

By using the usual techniques (see Section 5.6, for instance) we get for all estimators $\widehat{f}$,

$$\sup_{\substack{f_0 \in \mathcal{C} \\ L_2(f_0)\leq 2,\, \|f_0\|_\infty \leq 2}} \mathbb{E}_{f_0}\big[d_1\big(f_0, \widehat{f}\big)\big]$$
$$\geq \frac{1}{2} \inf_{\nu \in \{0,1\}^p} \sum_{j=1}^{p} \Big[\mathbb{E}_{\overline{\varphi}_{\nu_{j,1}}} \int_{I_j} |\widehat{f} - \overline{\varphi}_{\nu_{j,1}}| + \mathbb{E}_{\overline{\varphi}_{\nu_{j,0}}} \int_{I_j} |\widehat{f} - \overline{\varphi}_{\nu_{j,0}}|\Big]$$
$$\geq \frac{1}{2} \inf_{\nu \in \{0,1\}^p} \sum_{j=1}^{p} \int_{I_j} |\overline{\varphi}_{\nu_{j,1}} - \overline{\varphi}_{\nu_{j,0}}| \times \big[1 - \big((1 + \overline{\chi}^2_\nu(j))^n - 1\big)^{1/2}\big], \qquad (27)$$

where

$$\overline{\chi}^2_\nu(j) = \int \frac{(\overline{\varphi}_{\nu_{j,1}} * q - \overline{\varphi}_{\nu_{j,0}} * q)^2}{\overline{\varphi}_{\nu_{j,0}} * q}.$$

By using the triangle inequality,

$$\int_{I_j} |\overline{\varphi}_{\nu_{j,1}} - \overline{\varphi}_{\nu_{j,0}}| \geq d_1(F_j, G_j) - \frac{\overline{\lambda} u^2 \varepsilon^3}{8(1+\varepsilon)^2}(1+\varepsilon)^j \int_{I_j} \varphi$$
$$\geq 0.3\varepsilon^3 - 0.13 \frac{\overline{\lambda} u^2 \varepsilon^3}{8(1+\varepsilon)^2}(1+\varepsilon)^j \ell_j$$
$$\geq 0.3\varepsilon^3 - 0.13 \times 0.8\varepsilon^4 u (1+\varepsilon)^{j-1}.$$

By using $\varepsilon = \log(3/2)/p$, the definition of u, and elementary computations, we get

$$\int_{I_j} |\overline{\varphi}_{\nu_{j,1}} - \overline{\varphi}_{\nu_{j,0}}| \geq 0.15\varepsilon^3.$$

We deduce from (27),

$$\sup_{\substack{f_0 \in \mathcal{C} \\ L_2(f_0)\leq 2,\, \|f_0\|_\infty \leq 2}} \mathbb{E}_{f_0}\big[d_1\big(f_0, \widehat{f}\big)\big] \geq 0.07\varepsilon^3 \inf_{\nu \in \{0,1\}^p} \sum_{j=1}^{p} \big[1 - \big((1 + \overline{\chi}^2_\nu(j))^n - 1\big)^{1/2}\big].$$

We need the following result:

Claim 4. There exists c depending on q only such that: for all $\nu \in \{0,1\}^p$,

$$\sup_{1 \leq j \leq p} \overline{\chi}^2_\nu(j) \leq c(1/p)^{2\beta+5}.$$

By choosing p as the smallest integer than $(2cn)^{1/(2\beta+5)}$, we get that $\overline{\chi}^2_\nu(j)$ becomes not larger than $1/(2n)$. In particular,

$$\sup_{\substack{f_0 \in \mathcal{C} \\ L_2(f_0)\leq 2,\, \|f_0\|_\infty \leq 2}} \mathbb{E}_{f_0}\big[d_1\big(f_0, \widehat{f}\big)\big] \geq 0.07\big(1 - \sqrt{e^{1/2} - 1}\big)\varepsilon^3 p,$$

which concludes the proof of Proposition 7. ■

Sketch of the proof of Claim 4. It follows from the proof of Claim 2 that there is some c_0 depending only on q such that

$$\overline{\chi}_\nu^2(j) \le c_0\Big[\int \big(\overline{\varphi}_{\nu_{j,1}} * q(x) - \overline{\varphi}_{\nu_{j,0}} * q(x)\big)^2 \,\mathrm{d}x + \int x^2\big(\overline{\varphi}_{\nu_{j,1}} * q(x) - \overline{\varphi}_{\nu_{j,0}} * q(x)\big)^2 \,\mathrm{d}x\Big].$$

Let $\overline{\psi}_j$ be defined for $j \in \{1, \dots, p\}$ and $t \in \mathbb{R}$ by

$$\begin{aligned}\overline{\psi}_j(t) &= \overline{\varphi}^\star_{\nu_{j,1}}(t) - \overline{\varphi}^\star_{\nu_{j,0}}(t),\\ &= F_j^\star(t) - G_j^\star(t) + \frac{\overline{\lambda} u^2 \varepsilon^3}{8}\varphi^\star\big(t\sqrt{3}\big)(1+\varepsilon)^{j-2}.\end{aligned}$$

Let ζ be the triangle function introduced in the proof of the preceding claim. Since $\zeta^\star(t) = 4\sin^2(t/2)/t^2$ for all $t \in \mathbb{R}$, we deduce from (26)

$$\begin{aligned}\overline{\psi}_j(t) &= -\frac{\overline{\lambda}\varepsilon\ell_j}{4(1+\varepsilon)^j}\left(\zeta\left(\frac{2(.-m_j)}{\ell_j}\right)\right)^\star(t) + \frac{\overline{\lambda} u^2 \varepsilon^3}{8}\varphi^\star\big(t\sqrt{3}\big)(1+\varepsilon)^{j-2},\\ &= -\frac{2\overline{\lambda}\varepsilon}{(1+\varepsilon)^j}\frac{\sin^2(t\ell_j/4)}{t^2}e^{itm_j} + \frac{\overline{\lambda} u^2 \varepsilon^3}{8}e^{-\sqrt{3}|t|}(1+\varepsilon)^{j-2}\\ &= \frac{2\overline{\lambda}\varepsilon}{(1+\varepsilon)^j}\left[-\frac{\sin^2(t\ell_j/4)}{t^2}e^{itm_j} + \frac{\ell_j^2}{16}e^{-\sqrt{3}|t|}\right].\end{aligned}$$

As for (23),

$$\overline{\chi}_\nu^2(j) \le \frac{c_0}{2\pi}\left[\int |\overline{\psi}_j q^\star|^2 + 2\int |\overline{\psi}_j (q^\star)'|^2 + 2\int |\overline{\psi}_j' q^\star|^2\right].$$

There exists C_1 depending only on κ_1', κ_2' such that

$$\int |\overline{\psi}_j q^\star|^2 \le C_1 \frac{\overline{\lambda}^2\varepsilon^2}{(1+\varepsilon)^{2j}}\left[\int \left(\frac{\sin(t\ell_j/4)}{t}\right)^4 \frac{1}{1+|t|^{2\beta}}\,\mathrm{d}t + \ell_j^4\right].$$

By noticing that $\overline{\lambda} \le 0.95$, $\ell_j \le 3\varepsilon$, $\beta < 1/2$, we get

$$\int |\overline{\psi}_j q^\star|^2 \le C_1' \varepsilon^{2\beta+5},$$

where C_1' only depends on $\beta, \kappa_1', \kappa_2'$. By doing quite similar computations, we obtain

$$\begin{aligned}\int |\overline{\psi}_j (q^\star)'|^2 &\le C_2 \varepsilon^{2\beta+5},\\ \int |\overline{\psi}_j' q^\star|^2 &\le C_3 \varepsilon^{2\beta+5},\end{aligned}$$

where C_2, C_3 only depend on $\beta, \kappa_1', \kappa_2', \kappa_3'$. By gathering these results, we deduce

$$\overline{\chi}_\nu^2(j) \le C' \varepsilon^{2\beta+5},$$

and it remains to say that $\varepsilon = \log(3/2)/p$. ■

5.8. Proof of Proposition 8

Claim 5. For all $I \in \mathcal{I}_d$,

$$\mathbb{P}(X \in I) \le \sqrt{\|f_0 * q\|_\infty \left[\kappa_1 d|I| + c_\beta \kappa_2 d^{1+2\beta}|I|^{1-2\beta}\right]}. \tag{28}$$

Proof. We use Proposition 2 and Jensen's inequality $\mathbb{E}[Z(I)] \le \sqrt{\mathbb{E}[(Z(I))^2]}$. ■

Claim 6. For all $\delta > 0$ and interval $I \in \mathcal{I}_1$, the following assertion holds true on an event of probability lower bounded by $1 - e^{-\delta}$:

$$|\widehat{Z}_\delta(I) - \mathbb{P}(X \in I)| \le C\sqrt{\frac{\|f_0 * q\|_\infty[\kappa_1|I| + c_\beta \kappa_2 |I|^{1-2\beta}](1+\delta)}{n}}. \tag{29}$$

In this inequality, C is a universal constant and c_β only depends on β.

Proof. The proof is straightforward when $|I| = 0$. Suppose now that $|I| > 0$. When $\delta < n - 1$, this inequality comes from standard results about median of means estimators, see [18, Section 4.1]. When δ is larger, $\widehat{Z}_\delta(I) = 0$, and (29) ensues from (28). ■

Claim 7. For all $\xi > 0$, the following assertion holds true on an event of probability lower bounded by $1 - e^{-\xi}$: for all d, j and closed interval $I \in \mathcal{I}_1$ with endpoints lying in $\mathcal{G}_{d,j}$,

$$|\widehat{Z}_{\xi+\delta_{d,j}}(I) - \mathbb{P}(X \in I)| \le C\sqrt{\frac{\|f_0 * q\|_\infty[\kappa_1|I| + c_\beta \kappa_2 |I|^{1-2\beta}](\xi + 1 + \delta_{d,j})}{n}}, \tag{30}$$

where $\delta_{d,j} = 2\log(jd) + 2\log(1 + d2^j)$. In this inequality, C is a universal constant and c_β only depends on β.

Proof. Let for all $\delta > 0$ and interval I, $\mathcal{E}(\delta, I)$ be the event on which (29) is true. Let $\mathcal{I}_{d,j}'$ be the collection of closed intervals with endpoints in $\mathcal{G}_{d,j}$. Then, (30) holds true on

$$\mathcal{E} = \bigcap_{(d,j)\in(\mathbb{N}^\star)^2,\, I\in\mathcal{I}_{d,j}} \mathcal{E}(\xi + \delta_{d,j}, I).$$

Moreover, Bonferroni's inequality asserts

$$\begin{aligned}\mathbb{P}\left(\mathcal{E}^c\right) &\le \sum_{(d,j)\in(\mathbb{N}^\star)^2} |\mathcal{I}'_{d,j}| \mathbb{P}\left(\mathcal{E}^c(\xi+\delta_{d,j}, I)\right)\\ &\le e^{-\xi} \sum_{(d,j)\in(\mathbb{N}^\star)^2} (1+d2^j)^2\, e^{-2\log(jd)-2\log(1+d2^j)}\\ &\le e^{-\xi}.\end{aligned}$$

■

Claim 8. For all $\xi > 0$, the following assertion holds true on an event of probability lower bounded by $1-e^{-\xi}$: for all d, j and $I \in \mathcal{I}_{d,j}$,

$$\begin{aligned}&|\widehat{Z}_{\xi+\delta_{d,j}}(I) - \mathbb{P}(X \in I)|\\ &\quad\le C\sqrt{\frac{\|f_0 * q\|_\infty[\kappa_1 d|I| + c_\beta \kappa_2 d^{1+2\beta}|I|^{1-2\beta}](\xi+1+\delta_{d,j})}{n}}.\end{aligned}$$

In this inequality, C is a universal constant and c_β only depends on β.

Proof. Let $I \in \mathcal{I}_{d,j}$ written as $\overline{I} = \cup_{k=1}^d \overline{I}_k$ such that $\overline{I}_{k_1} \cap \overline{I}_{k_2} = \emptyset$ for all $k_1 \neq k_2$. Then,

$$|\widehat{Z}_{\xi+\delta_{d,j}}(I) - \mathbb{P}(X \in I)| \le \sum_{k=1}^d |\widehat{Z}_{\xi+\delta_{d,j}}(\overline{I}_k) - \mathbb{P}(X \in \overline{I}_k)|.$$

Note that each $\overline{I}_k$ belongs to the set $\mathcal{I}'_{d,j}$ defined in the preceding proof. Therefore, on the event defined in the preceding claim,

$$\begin{aligned}&|\widehat{Z}_{\xi+\delta_{d,j}}(I) - \mathbb{P}(X \in I)|\\ &\quad\le C\sqrt{\frac{\|f_0 * q\|_\infty(\xi+1+\delta_{d,j})}{n}} \sum_{k=1}^d \sqrt{\kappa_1|I_k| + c_\beta\kappa_2|I_k|^{1-2\beta}}.\end{aligned}$$

We finally use Cauchy–Schwarz inequality and the Hölder inequality (as $\beta < 1/2$). ■

Claim 9. For all $\xi > 0$, the following assertion holds true on an event of probability lower bounded by $1-e^{-\xi}$: for all j and $I \in \mathcal{I}_\infty$ with endpoints lying in $\mathcal{G}_{2d_I,j}$,

$$\begin{aligned}&|\widehat{Z}_{\xi+\delta_j(I)}(I) - \mathbb{P}(X \in I)|\\ &\quad\le C\sqrt{\frac{\|f_0 * q\|_\infty[\kappa_1 d_I|I| + c_\beta \kappa_2 d_I^{1+2\beta}|I|^{1-2\beta}](\xi+1+\delta_j(I))}{n}}.\end{aligned}$$

In this inequality, C is a universal constant and c_β only depends on β.

Proof. We apply the preceding claim, use that $\delta_j(I) = \delta_{2d_I,j}$ and increase C, c_β. ■

Proof of Proposition 8. For all j and $I \in \mathcal{I}_\infty$,

$$|\pi_{j+1}(I) \setminus \pi_j(I)| \leq 2^{-j+2}$$

and $\pi_{j+1}(I) \setminus \pi_j(I)$ belongs to $\mathcal{I}_{2d_I,j}$. The preceding claim then ensures that

$$\begin{aligned} &\left|\hat{Z}_{\xi+\delta_j(I)}\big(\pi_{j+1}(I) \setminus \pi_j(I)\big) - \mathbb{P}\left(X \in \pi_{j+1}(I) \setminus \pi_j(I)\right)\right| \\ &\quad \leq C\sqrt{\frac{\|f_0 * q\|_\infty(\kappa_1 d_I 2^{-j} + c_\beta \kappa_2 d_I^{1+2\beta} 2^{-(1-2\beta)j})}{n}} \\ &\qquad \times \sqrt{\xi + 1 + 2\log(2jd_I) + 2\log(1 + d_I 2^{j+1})} \end{aligned}$$

Therefore,

$$\begin{aligned} &\sum_{j=1}^{\infty}\left|\hat{Z}_{\xi+\delta_j(I)}\big(\pi_{j+1}(I) \setminus \pi_j(I)\big) - \mathbb{P}\left(X \in \pi_{j+1}(I) \setminus \pi_j(I)\right)\right| \\ &\quad \leq C'\sqrt{\frac{\|f_0 * q\|_\infty(\kappa_1 d_I + c'_\beta \kappa_2 d_I^{2\beta+1})(\xi + \log d_I)}{n}}, \end{aligned}$$

where C' is a universal constant and where c'_β depends on β only. Since

$$\mathbb{P}(X \in I) = \mathbb{P}\left(X \in \pi_1(I)\right) + \sum_{j=1}^{\infty} \mathbb{P}\left(X \in \pi_{j+1}(I) \setminus \pi_j(I)\right),$$

we obtain

$$\begin{aligned} |\hat{Z}_{n,\xi}(I) - \mathbb{P}(X \in I)| &\leq C'\sqrt{\frac{\|f_0 * q\|_\infty(\kappa_1 d_I + c'_\beta \kappa_2 d_I^{2\beta+1})(\xi + \log d_I)}{n}} \\ &\quad + \left|\hat{Z}_{\xi+\delta_1(I)}(\pi_1(I)) - \mathbb{P}\left(X \in \pi_1(I)\right)\right|. \end{aligned}$$

We use Claim 9 again to bound the last term. ■

5.9. Proof of Lemma 1

Let f be an arbitrary function of $\mathcal{F}$. Thanks to Assumption 2, there exists $f' \in \overline{\mathcal{F}}$ such that $d_{f'} \leq d_f$, $d_1(f, f') \leq 1/(2n)$. Then,

$$\begin{aligned} \gamma_\xi(f') &\leq \sup_{I \in \overline{\mathcal{I}}_{d_{f'}}} \left\{\left|\int_I f' - \int_I f\right| + \left|\int_I f - \hat{Z}_{n,\xi}(I)\right|\right\} \\ &\leq d_1(f, f') + \gamma_\xi(f) \\ &\leq 1/(2n) + \gamma_\xi(f). \end{aligned}$$

We thus deduce that

$$\inf_{f\in\overline{\mathcal{F}}} \gamma_\xi(f) \leq 1/(2n) + \inf_{f\in\mathcal{F}} \gamma_\xi(f).$$

Therefore, any $\tilde{f} \in \overline{\mathcal{F}}$ such that

$$\gamma_\xi(\tilde{f}) \leq \inf_{f\in\overline{\mathcal{F}}} \gamma_\xi(f) + 1/(2n)$$

satisfies condition (17). ∎

5.10. Proof of Theorem 6

Lemma 2. *Let $\tilde{f} \in \mathcal{F}$ be an estimator satisfying* (17). *Then,*

$$d_1(f_0, \tilde{f}) \leq \inf_{f\in\mathcal{F}} \left\{ 5 d_1(f_0, f) + 4 \sup_{I\in\overline{\mathcal{I}}_{d_f}} \left| \hat{Z}_{n,\xi}(I) - \int_I f_0 \right| + 2/n \right\}.$$

Proof. For all $f \in \mathcal{F}$, we define

$$\begin{aligned} I &= [f > \tilde{f}] \quad \text{if} \int_{[0,1]} f \geq \int_{[0,1]} \tilde{f}, \text{ and} \\ I &= [\tilde{f} > f] \quad \text{otherwise.} \end{aligned}$$

Note that $I \in \mathcal{I}_{d_f} \cap \mathcal{I}_{d_{\tilde{f}}}$ and

$$d_1(f, \tilde{f}) = 2 \left| \int_I f - \int_I \tilde{f} \right| - \left| \int_{[0,1]} f - \int_{[0,1]} \tilde{f} \right|$$

as f and $\tilde{f}$ vanish outside [0, 1]. There exists $\overline{I} \in \overline{\mathcal{I}}_{d_f} \cap \overline{\mathcal{I}}_{d_{\tilde{f}}}$ included in I such that

$$\begin{aligned} d_1(f, \tilde{f}) &\leq 2 \left| \int_{\overline{I}} f - \int_{\overline{I}} \tilde{f} \right| + 1/n \\ &\leq 2 \left| \int_{\overline{I}} f - \hat{Z}_{n,\xi}(\overline{I}) \right| + 2 \left| \hat{Z}_{n,\xi}(\overline{I}) - \int_{\overline{I}} \tilde{f} \right| + 1/n. \end{aligned}$$

Therefore,

$$\begin{aligned} d_1(f, \tilde{f}) &\leq 2\gamma_\xi(f) + 2\gamma_\xi(\tilde{f}) + 1/n \\ &\leq 4\gamma_\xi(f) + 2/n \end{aligned} \tag{31}$$

by using (17). Now,

$$\begin{aligned}\gamma_\xi(f) &\leq \sup_{I \in \bar{\mathcal{I}}_{d_f}} \left| \int_I f_0 - \int_I f \right| + \sup_{I \in \bar{\mathcal{I}}_{d_f}} \left| \int_I f_0 - \widehat{Z}_{n,\xi}(I) \right| \\ &\leq d_1(f_0, f) + \sup_{I \in \bar{\mathcal{I}}_{d_f}} \left| \int_I f_0 - \widehat{Z}_{n,\xi}(I) \right|.\end{aligned}$$

This inequality, together with (31) entails that

$$d_1(f, \tilde{f}) \leq 4 d_1(f_0, f) + 4 \sup_{I \in \bar{\mathcal{I}}_{d_f}} \left| \int_I f_0 - \widehat{Z}_{n,\xi}(I) \right| + 2/n.$$

We conclude using the triangle inequality

$$d_1(f_0, \tilde{f}) \leq d_1(f_0, f) + d_1(f, \tilde{f}). \qquad \blacksquare$$

Proof of Theorem 6. Elementary computations give $d_1(f_0, \hat{f}) \leq d_1(f_0, \tilde{f})$. We may now combine Proposition 8 and Lemma 2 with $\xi = \log n$ to get an event $\mathcal{E}_n$ of probability $1 - 1/n$ such that

$$\begin{aligned}&d_1(f_0, \hat{f}) \mathbb{1}_{\mathcal{E}_n} \\ &\quad \leq \inf_{f \in \mathcal{F}} \left\{ 5 d_1(f_0, f) + c \sqrt{\|f_0 * q\|_\infty \frac{(\kappa_1 d_f + c_\beta \kappa_2 d_f^{1+2\beta}) \log(n d_f)}{n}} \right\} + 2/n.\end{aligned}$$

Yet,

$$\begin{aligned}\mathbb{E}\big[d_1(f_0, \hat{f})\big] &\leq \mathbb{E}\big[d_1(f_0, \hat{f}) \mathbb{1}_{\mathcal{E}_n}\big] + \mathbb{E}\big[d_1(f_0, \hat{f}) \mathbb{1}_{\mathcal{E}_n^c}\big] \\ &\leq \mathbb{E}\big[d_1(f_0, \hat{f}) \mathbb{1}_{\mathcal{E}_n}\big] + 2/n,\end{aligned}$$

thanks to the inequality $d_1(f_0, \hat{f}) \leq \int_{[0,1]} f_0 + \int_{[0,1]} \hat{f} \leq 2$. ∎

References

[1] Y. Baraud and L. Birgé, Rho-estimators for shape restricted density estimation. *Stochastic Process. Appl.* **126** (2016), no. 12, 3888–3912 Zbl 1419.62070 MR 3565484

[2] Y. Baraud, L. Birgé, and M. Sart, A new method for estimation and model selection: ρ-estimation. *Invent. Math.* **207** (2017), no. 2, 425–517 Zbl 1373.62141 MR 3595933

[3] L. Birgé, Approximation dans les espaces métriques et théorie de l'estimation. *Z. Wahrsch. Verw. Gebiete* **65** (1983), no. 2, 181–237 Zbl 0506.62026 MR 722129

[4] L. Birgé, Estimating a density under order restrictions: nonasymptotic minimax risk. *Ann. Statist.* **15** (1987), no. 3, 995–1012 Zbl 0631.62037 MR 902241

[5] L. Birgé, On the risk of histograms for estimating decreasing densities. *Ann. Statist.* **15** (1987), no. 3, 1013–1022 Zbl 0646.62033 MR 902242

[6] L. Birgé, Model selection via testing: an alternative to (penalized) maximum likelihood estimators. *Ann. Inst. H. Poincaré Probab. Statist.* **42** (2006), no. 3, 273–325 Zbl 1333.62094 MR 2219712

[7] L. Birgé, Robust tests for model selection. In *From probability to statistics and back: high-dimensional models and processes*, pp. 47–64, Inst. Math. Stat. (IMS) Collect. 9, Inst. Math. Statist., Beachwood, OH, 2013 Zbl 1327.62279 MR 3186748

[8] L. Birgé and P. Massart, Minimum contrast estimators on sieves: exponential bounds and rates of convergence. *Bernoulli* **4** (1998), no. 3, 329–375 Zbl 0954.62033 MR 1653272

[9] C. Butucea, C. Matias, and C. Pouet, Adaptivity in convolution models with partially known noise distribution. *Electron. J. Stat.* **2** (2008), 897–915 Zbl 1320.62066 MR 2447344

[10] R. J. Carroll, A. Delaigle, and P. Hall, Testing and estimating shape-constrained nonparametric density and regression in the presence of measurement error. *J. Amer. Statist. Assoc.* **106** (2011), no. 493, 191–202 Zbl 1396.62084 MR 2816713

[11] L. Cavalier and M. Raimondo, Multiscale density estimation with errors in variables. *J. Korean Statist. Soc.* **39** (2010), no. 4, 417–429 Zbl 1294.62074 MR 2780213

[12] F. Comte and C. Lacour, Data-driven density estimation in the presence of additive noise with unknown distribution. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **73** (2011), no. 4, 601–627 Zbl 1226.62034 MR 2853732

[13] F. Comte and C. Lacour, Anisotropic adaptive kernel deconvolution. *Ann. Inst. Henri Poincaré Probab. Stat.* **49** (2013), no. 2, 569–609 Zbl 1348.62121 MR 3088382

[14] F. Comte, Y. Rozenholc, and M.-L. Taupin, Penalized contrast estimator for adaptive density deconvolution. *Canad. J. Statist.* **34** (2006), no. 3, 431–452 Zbl 1104.62033 MR 2328553

[15] I. Dattner, A. Goldenshluger, and A. Juditsky, On deconvolution of distribution functions. *Ann. Statist.* **39** (2011), no. 5, 2477–2501 Zbl 1232.62056 MR 2906875

[16] J. Dedecker and B. Michel, Minimax rates of convergence for Wasserstein deconvolution with supersmooth errors in any dimension. *J. Multivariate Anal.* **122** (2013), 278–291 Zbl 1280.62043 MR 3189324

[17] A. Delaigle and I. Gijbels, Bootstrap bandwidth selection in kernel density estimation from a contaminated sample. *Ann. Inst. Statist. Math.* **56** (2004), no. 1, 19–47 Zbl 1050.62038 MR 2053727

[18] L. Devroye, M. Lerasle, G. Lugosi, and R. I. Oliveira, Sub-Gaussian mean estimators. *Ann. Statist.* **44** (2016), no. 6, 2695–2725 Zbl 1360.62115 MR 3576558

[19] L. Devroye and G. Lugosi, *Combinatorial methods in density estimation*. Springer Ser. Statist., Springer, New York, 2001 Zbl 0964.62025 MR 1843146

[20] J. Fan, On the optimal rates of convergence for nonparametric deconvolution problems. *Ann. Statist.* **19** (1991), no. 3, 1257–1272 Zbl 0729.62033 MR 1126324

[21] J. Fan and J.-Y. Koo, Wavelet deconvolution. *IEEE Trans. Inform. Theory* **48** (2002), no. 3, 734–747 Zbl 1071.94511 MR 1889978

[22] I. A. Ibragimov and R. Z. Has'minskiĭ, *Statistical estimation*. Appl. Math.16, Springer, New York-Berlin, 1981 MR 620321

[23] O. V. Lepski and T. Willer, Oracle inequalities and adaptive estimation in the convolution structure density model. *Ann. Statist.* **47** (2019), no. 1, 233–287 Zbl 1419.62075 MR 3909933

[24] K. Lounici and R. Nickl, Global uniform risk bounds for wavelet deconvolution estimators. *Ann. Statist.* **39** (2011), no. 1, 201–231 Zbl 1209.62060 MR 2797844

[25] C. Marteau and M. Sart, Deconvolution for some singular density errors with respect to the $\mathbb{L}^1$ loss. 2021, https://hal.science/hal-02511859

[26] A. Meister, *Deconvolution problems in nonparametric statistics*. Lect. Notes Stat. 193, Springer, Berlin, 2009 Zbl 1178.62028 MR 2768576

[27] M. Pensky and B. Vidakovic, Adaptive wavelet estimator for nonparametric density deconvolution. *Ann. Statist.* **27** (1999), no. 6, 2033–2053 Zbl 0962.62030 MR 1765627

[28] T. M. Pham Ngoc and V. Rivoirard, The dictionary approach for spherical deconvolution. *J. Multivariate Anal.* **115** (2013), 138–156 Zbl 1258.62042 MR 3004550

[29] G. Rebelles, Structural adaptive deconvolution under $\mathbb{L}_p$-losses. *Math. Methods Statist.* **25** (2016), no. 1, 26–53 Zbl 1342.62054 MR 3480609

[30] J. Schmidt-Hieber, A. Munk, and L. Dümbgen, Multiscale methods for shape constraints in deconvolution: confidence statements for qualitative features. *Ann. Statist.* **41** (2013), no. 3, 1299–1328 Zbl 1293.62104 MR 3113812

[31] A. B. Tsybakov, *Introduction to nonparametric estimation*. Springer Ser. Statist., Springer, New York, 2009 Zbl 1176.62032 MR 2724359

[32] E. O. Tuck, On positivity of Fourier transforms. *Bull. Austral. Math. Soc.* **74** (2006), no. 1, 133–138 Zbl 1097.42008 MR 2248812

Received 16 June 2021; revised 12 July 2022.

Clément Marteau
Univ Lyon, Université Claude Bernard Lyon 1, CNRS UMR 5208, Institut Camille Jordan, 69622, Villeurbanne, France; clement.marteau@univ-lyon1.fr

Mathieu Sart
Université Jean Monnet Saint-Étienne, CNRS, Institut Camille Jordan UMR 5208, 42023, Saint-Étienne, France; mathieu.sart@univ-st-etienne.fr