

ORACLE INEQUALITIES AND ADAPTIVE ESTIMATION IN THE CONVOLUTION STRUCTURE DENSITY MODEL¹

BY O. V. LEPSKI AND T. WILLER

Aix-Marseille Université

We study the problem of nonparametric estimation under \mathbb{L}_p -loss, $p \in [1, \infty)$, in the framework of the convolution structure density model on \mathbb{R}^d . This observation scheme is a generalization of two classical statistical models, namely density estimation under direct and indirect observations. The original pointwise selection rule from a family of “kernel-type” estimators is proposed. For the selected estimator, we prove an \mathbb{L}_p -norm oracle inequality and several of its consequences. Next, the problem of adaptive minimax estimation under \mathbb{L}_p -loss over the scale of anisotropic Nikol’skii classes is addressed. We fully characterize the behavior of the minimax risk for different relationships between regularity parameters and norm indexes in the definitions of the functional class and of the risk. We prove that the proposed selection rule leads to the construction of an optimally or nearly optimally (up to logarithmic factors) adaptive estimator.

CONTENTS

1. Introduction	234
1.1. Oracle approach via local selection	235
1.2. Adaptive estimation	236
1.3. Assumption on the function g	239
2. Pointwise selection rule and \mathbb{L}_p -norm oracle inequality	239
2.1. Pointwise selection rule from the family of kernel estimators	240
Pointwise selection rule	240
2.2. \mathbb{L}_p -norm oracle inequality	240
3. Abstract upper bound theorem	242
4. Adaptive estimation over the scale of anisotropic classes	245
4.1. Anisotropic Nikol’skii classes	245
4.2. Construction of kernel K	246
4.3. Main results	246
4.3.1. Bounded case	246
4.3.2. Unbounded case, $\alpha = 1$	248
5. Proofs of Theorems 1–2	250
5.1. Proof of Theorem 1	250

Received April 2017; revised November 2017.

¹This work has been carried out in the framework of the Labex Archimède (ANR-11-LABX-0033) and of the A*MIDEX project (ANR-11-IDEX-0001-02), funded by the “Investissements d’Avenir” French Government program managed by the French National Research Agency (ANR).

MSC2010 subject classifications. 62G05, 62G20.

Key words and phrases. Deconvolution model, density estimation, oracle inequality, adaptive estimation, kernel estimators, \mathbb{L}_p -risk, anisotropic Nikol’skii class.

5.1.1. Proof of the theorem 251
 5.1.2. Proof of Proposition 1 253
 5.2. Proof of Theorem 2 257
 5.2.1. Preliminaries 258
 5.2.2. Part I 260
 5.2.3. Part II 265
 6. Proof of Theorems 3 and 4 268
 6.1. Special set of bandwidths 268
 6.2. Auxiliary statements 269
 6.3. Several bounds 271
 6.4. Proof of Theorem 3 272
 6.5. Proof of Theorem 4 272
 7. Proofs of Propositions 2 and 3 273
 7.1. Proof of Proposition 2 273
 7.2. Proof of Proposition 3 277
 Appendix 281
 References 285

1. Introduction. In the present paper, we will investigate the following observation scheme introduced in [Lepski and Willer \(2017\)](#). Suppose that we observe i.i.d. vectors $Z_i \in \mathbb{R}^d, i = 1, \dots, n$, with a common probability density p satisfying the following structural assumption:

$$(1.1) \quad p = (1 - \alpha)f + \alpha[f \star g], \quad f \in \mathbb{F}_g(R), \alpha \in [0, 1],$$

where $\alpha \in [0, 1]$ and $g : \mathbb{R}^d \rightarrow \mathbb{R}$ are supposed to be known and $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is the function to be estimated. We will call the observation scheme (1.1) *convolution structure density model*.

Here and later, for two functions $f, g \in \mathbb{L}_1(\mathbb{R}^d)$

$$[f \star g](x) = \int_{\mathbb{R}^d} f(x - z)g(z)v_d(dz), \quad x \in \mathbb{R}^d,$$

and for any $\alpha \in [0, 1], g \in \mathbb{L}_1(\mathbb{R}^d)$ and $R > 1$,

$$\mathbb{F}_g(R) = \{f \in \mathbb{B}_{1,d}(R) : (1 - \alpha)f + \alpha[f \star g] \in \mathfrak{P}(\mathbb{R}^d)\}.$$

Here, $\mathfrak{P}(\mathbb{R}^d)$ denotes the set of probability densities on $\mathbb{R}^d, \mathbb{B}_{s,d}(R)$ is the ball of radius $R > 0$ in $\mathbb{L}_s(\mathbb{R}^d) := \mathbb{L}_s(\mathbb{R}^d, \nu_d), 1 \leq s \leq \infty$ and ν_d is the Lebesgue measure on \mathbb{R}^d . The convolution structure density model (1.1) will be studied for an arbitrary $g \in \mathbb{L}_1(\mathbb{R}^d)$ and $f \in \mathbb{F}_g(R)$. Then, except in the case $\alpha = 0$, the function f is not necessarily a probability density.

We remark that if one assumes additionally that $f, g \in \mathfrak{P}(\mathbb{R}^d)$, this model can be interpreted as follows. The observations $Z_i \in \mathbb{R}^d, i = 1, \dots, n$, can be written as a sum of two independent random vectors, that is,

$$(1.2) \quad Z_i = X_i + \epsilon_i Y_i, \quad i = 1, \dots, n,$$

where $X_i, i = 1, \dots, n$, are i.i.d. d -dimensional random vectors with a common density f , to be estimated. The noise variables $Y_i, i = 1, \dots, n$, are i.i.d. d -dimensional random vectors with a known common density g . At last $\varepsilon_i \in \{0, 1\}, i = 1, \dots, n$, are i.i.d. Bernoulli random variables with $\mathbb{P}(\varepsilon_1 = 1) = \alpha$, where $\alpha \in [0, 1]$ is supposed to be known. The sequences $\{X_i, i = 1, \dots, n\}, \{Y_i, i = 1, \dots, n\}$ and $\{\varepsilon_i, i = 1, \dots, n\}$ are supposed to be mutually independent.

The observation scheme (1.2) can be viewed as the generalization of two classical statistical models. Indeed, the case $\alpha = 1$ corresponds to the standard deconvolution model $Z_i = X_i + Y_i, i = 1, \dots, n$. Another ‘‘extreme’’ case $\alpha = 0$ corresponds to the direct observation scheme $Z_i = X_i, i = 1, \dots, n$. The ‘‘intermediate’’ case $\alpha \in (0, 1)$, considered for the first time in Hesse (1995), can be treated as the mathematical modeling of the following situation. One part of the data, namely $(1 - \alpha)n$, is observed without noise, while the other part is contaminated by additional noise. If the indexes corresponding to that first part were known, the density f could be estimated using only this part of the data, with the accuracy corresponding to the direct case. The question we address now is: can one obtain the same accuracy if the latter information is not available? We will see that the answer to the aforementioned question is positive, but the construction of optimal estimation procedures is based upon ideas corresponding to the ‘‘pure’’ deconvolution model.

We want to estimate f using the observations $Z^{(n)} = (Z_1, \dots, Z_n)$. By estimator, we mean any $Z^{(n)}$ -measurable map $\hat{f} : \mathbb{R}^n \rightarrow \mathbb{L}_p(\mathbb{R}^d)$. The accuracy of an estimator \hat{f} is measured by the \mathbb{L}_p -risk

$$\mathcal{R}_n^{(p)}[\hat{f}, f] := (\mathbb{E}_f \|\hat{f} - f\|_p^p)^{1/p}, \quad p \in [1, \infty),$$

where \mathbb{E}_f denotes the expectation with respect to the probability measure \mathbb{P}_f of the observations $Z^{(n)} = (Z_1, \dots, Z_n)$. Also, $\|\cdot\|_p, p \in [1, \infty)$, is the \mathbb{L}_p -norm on \mathbb{R}^d and without further mentioning we will assume that $f \in \mathbb{L}_p(\mathbb{R}^d)$.

1.1. *Oracle approach via local selection.* Let $\mathcal{F}(\mathbb{H}) = \{\hat{f}_{\vec{h}}, \vec{h} \in \mathbb{H}\}$ be a family of ‘‘kernel-type’’ estimators (see Section 2.1), parameterized by a collection of multibandwidths \mathbb{H} built from the observation $Z^{(n)}$. We want to construct a $Z^{(n)}$ -measurable random map $\vec{h} : \mathbb{R}^d \rightarrow \mathbb{H}$ and for any $p \in [1, \infty)$ and $n \geq 1$ to bound from above the \mathbb{L}_p -risk of the selected estimator $\hat{f}_{\vec{h}(\cdot)}$. Our selection rule presented in Section 2.1 can be viewed as a generalization and modification of statistical procedures proposed in Kerkycharian, Lepski and Picard (2001) and Goldenshluger and Lepski (2014). In Section 2.2, the following risk bound will be established:

$$(1.3) \quad \mathcal{R}_n^{(p)}[\hat{f}_{\vec{h}(\cdot)}; f] \leq C_1 \left\| \inf_{\vec{h} \in \mathbb{H}} A_n(f, \vec{h}, \cdot) \right\|_p + C_2 n^{-\frac{1}{2}} \quad \forall f \in \mathbb{F}_g(R).$$

Here, C_1 and C_2 are numerical constants which depend on d and p only, and $A_n(\cdot, \cdot, x), x \in \mathbb{R}^d$, is an explicitly known functional. We call (1.3) an \mathbb{L}_p -norm oracle inequality obtained by local selection. Since the selection rule from the

considered family is done pointwisely, that is, for any $x \in \mathbb{R}^d$, this allows to take into account the “local structure” of the function to be estimated. The \mathbb{L}_p -norm oracle inequality is then obtained by the integration of the pointwise risk of the proposed estimator, which is a kernel estimator with the bandwidth being a multivariate random function. This, in its turn, allows us to derive different minimax adaptive results thanks to an unique \mathbb{L}_p -norm oracle inequality. It is worth noting in this context that estimation procedures based on a local selection scheme can be applied to the estimation of functions belonging to much more general functional classes than those based on global selection schemes; see, for instance, Goldenshluger and Lepski (2011) and Goldenshluger and Lepski (2014) for comparison. We will see however that although $A_n(\cdot, \cdot, x)$, $x \in \mathbb{R}^d$, is known explicitly, its computation in particular problems is not a simple task. The main difficulty here is mostly related to the fact that (1.3) is proved without any assumption (except for the model requirements) imposed on the underlying function f . It turns out that under some nonrestrictive assumptions imposed on f , the obtained bound can be considerably simplified; see Section 3.

1.2. *Adaptive estimation.* Let Σ be a given subset of $\mathbb{L}_p(\mathbb{R}^d)$. For any estimator \tilde{f}_n , define its *maximal risk* by $\mathcal{R}_n^{(p)}[\tilde{f}_n; \Sigma] = \sup_{f \in \Sigma} \mathcal{R}_n^{(p)}[\tilde{f}_n; f]$ and its *minimax risk* on Σ is given by

$$\phi_n(\Sigma) := \inf_{\tilde{f}_n} \mathcal{R}_n^{(p)}[\tilde{f}_n; \Sigma].$$

Here, the infimum is taken over all possible estimators. An estimator whose maximal risk is bounded, up to some constant factor, by $\phi_n(\Sigma)$, is called minimax on Σ .

Let $\{\Sigma_\vartheta, \vartheta \in \Theta\}$ be a collection of subsets of $\mathbb{L}_p(\mathbb{R}^d, \nu_d)$, where ϑ is a nuisance parameter which may have a very complicated structure.

The problem of adaptive estimation can be formulated as follows: *is it possible to construct a single estimator \hat{f}_n which would be simultaneously minimax on each class Σ_ϑ , $\vartheta \in \Theta$, that is,*

$$\limsup_{n \rightarrow \infty} \phi_n^{-1}(\Sigma_\vartheta) \mathcal{R}_n^{(p)}[\hat{f}_n; \Sigma_\vartheta] < \infty \quad \forall \vartheta \in \Theta?$$

We refer to this question as *the problem of minimax adaptive estimation over the scale of $\{\Sigma_\vartheta, \vartheta \in \Theta\}$* . If such an estimator exists, we will call it optimally adaptive.

From oracle approach to adaptation. Let the oracle inequality (1.3) be established. Define

$$R_n(\Sigma_\vartheta) = \sup_{f \in \Sigma_\vartheta} \left\| \inf_{\vec{h} \in \mathbb{H}} A_n(f, \vec{h}, \cdot) \right\|_p + n^{-\frac{1}{2}}, \quad \vartheta \in \Theta.$$

We immediately deduce from (1.3) that for any $\vartheta \in \Theta$

$$\limsup_{n \rightarrow \infty} R_n^{-1}(\Sigma_\vartheta) \mathcal{R}_n^{(p)}[\hat{f}_{\vec{h}(\cdot)}; \Sigma_\vartheta] < \infty.$$

Hence, the minimax adaptive optimality of the estimator $\hat{f}_{\mathbf{h}(\cdot)}$ is reduced to the comparison of the normalization $R_n(\Sigma_\vartheta)$ with the minimax risk $\phi_n(\Sigma_\vartheta)$. Indeed, if one proves that for any $\vartheta \in \Theta$

$$\liminf_{n \rightarrow \infty} R_n(\Sigma_\vartheta) \phi_n^{-1}(\Sigma_\vartheta) < \infty,$$

then the estimator $\hat{f}_{\mathbf{h}(\cdot)}$ is *optimally adaptive* over the scale $\{\Sigma_\vartheta, \vartheta \in \Theta\}$. Using the modern statistical language we call the estimator \hat{f}_n *nearly optimally adaptive* if

$$\limsup_{n \rightarrow \infty} \phi_n^{-1}(\Sigma_\vartheta) \mathcal{R}_n^{(p)}[\hat{f}_n; \Sigma_\vartheta] < \infty \quad \forall \vartheta \in \Theta.$$

Objectives. In the framework of the *convolution structure density model*, we will be interested in adaptive estimation over the scale

$$\Sigma_\vartheta = \mathbb{N}_{\vec{r}, d}(\vec{\beta}, \vec{L}) \cap \mathbb{F}_{g, \infty}(R, Q), \quad \vartheta = (\vec{\beta}, \vec{r}, \vec{L}, R, Q),$$

where $\mathbb{F}_{g, \infty}(R, Q) := \{f \in \mathbb{F}_g(R) : (1 - \alpha)f + \alpha[f \star g] \in \mathbb{B}_{\infty, d}(Q)\}$ and $\mathbb{N}_{\vec{r}, d}(\vec{\beta}, \vec{L})$ is the anisotropic Nikol'skii class; see Definition 1 below. Here, we only mention that the adaptive estimation over the scale $\{\mathbb{N}_{\vec{r}, d}(\vec{\beta}, \vec{L}), (\vec{\beta}, \vec{r}, \vec{L}) \in (0, \infty)^d \times [1, \infty]^d \times (0, \infty)^d\}$ is usually viewed as the adaptation to anisotropy and inhomogeneity of the function to be estimated. As to the assumption $f \in \mathbb{F}_{g, \infty}(R, Q)$, it simply means that the common density of observations \mathfrak{p} is uniformly bounded by Q . In particular, this is always the case if $\alpha = 1$ and $\|g\|_\infty < \infty$.

Additionally, we will study the adaptive estimation over the collection

$$\Sigma_\vartheta = \mathbb{N}_{\vec{r}, d}(\vec{\beta}, \vec{L}) \cap \mathbb{F}_g(R) \cap \mathbb{B}_{\infty, d}(Q), \quad \vartheta = (\vec{\beta}, \vec{r}, \vec{L}, R, Q).$$

We will show that the boundedness of the underlying function allows to improve considerably the accuracy of estimation.

Historical notes. The minimax adaptive estimation is a very active area of mathematical statistics, and the interested reader can find a very detailed overview as well as several open problems in adaptive estimation in Lepski (2015). Below we will discuss only the articles whose results are relevant to our consideration, that is, the density setting under \mathbb{L}_p -loss, from a minimax perspective. As already said, the convolution structure density model includes itself the density estimation under direct and indirect observations.

Direct case, $\alpha = 0$. There is a vast literature dealing with minimax and minimax adaptive density estimation; see, for example, Efroimovich (1986), Hasminskii and Ibragimov (1990), Golubev (1992), Donoho et al. (1996), Devroye and Lugosi (1997), Rigollet (2006), Rigollet and Tsybakov (2007), Samarov and Tsybakov (2007), Birgé (2014), Giné and Nickl (2009), Akakpo (2012), Gach, Nickl and Spokoiny (2013), Lepski (2013), among many others. Special attention was paid

to the estimation of densities with unbounded support; see Juditsky and Lambert-Lacroix (2004), Reynaud-Bouret, Rivoirard and Tuleau-Malot (2011). The most general results can be found in Goldenshluger and Lepski (2011, 2014) and in Section 4 we will compare in detail our results with those obtained in these papers.

Intermediate case, $\alpha \in (0, 1)$. To the best of our knowledge, adaptive estimation in the case of partially contaminated observations has not been studied yet. We were able to find only two papers dealing with minimax estimation. The first one is Hesse (1995) (where the discussed model was introduced in dimension 1) in which the author evaluated the \mathbb{L}_∞ -risk of the proposed estimator over a functional class formally corresponding to the Nikol'skii class $\mathbb{N}_{\infty,1}(\vec{2}, 1)$. In Yuan and Chen (2002), the latter result was extended to the multidimensional setting, that is, to the minimax estimation on $\mathbb{N}_{\infty,d}(\vec{2}, 1)$. The most intriguing fact is that the accuracy of estimation in partially contaminated noise is the same as in the direct observation scheme; however, none of these articles studied the optimality of the proposed estimators. We will come back to the aforementioned papers in Section 1.3 in order to compare the assumptions imposed on the noise density g .

Deconvolution case, $\alpha = 1$. First, let us remark that the behavior of the Fourier transform of the function g plays an important role in all the works dealing with deconvolution. Indeed ill-posed problems correspond to Fourier transforms decaying toward zero. Our results will be established for “moderately” ill posed problems, so we detail only results in papers studying that type of operators. This assumption means that there exist $\vec{\mu} = (\mu_1, \dots, \mu_d) \in (0, \infty)^d$ and $\Upsilon_1 > 0, \Upsilon_2 > 0$ such that the Fourier transform \check{g} of g satisfies

$$(1.4) \quad \Upsilon_1 \prod_{j=1}^d (1+t_j^2)^{-\frac{\mu_j}{2}} \leq |\check{g}(t)| \leq \Upsilon_2 \prod_{j=1}^d (1+t_j^2)^{-\frac{\mu_j}{2}} \quad \forall t = (t_1, \dots, t_d) \in \mathbb{R}^d.$$

Some minimax and minimax adaptive results in dimension 1 over different classes of smooth functions can be found in particular in Stefanski (1990), Stefanski and Carroll (1990), Fan (1991, 1993), Pensky and Vidakovic (1999), Fan and Koo (2002), Comte, Rozenholc and Taupin (2006), Butucea and Tsybakov (2008), Hall and Meister (2007), Meister (2009), Lounici and Nickl (2011), Kerkyacharian, Pham Ngoc and Picard (2011).

There are very few results in the multidimensional setting. It seems that Masry (1993) was the first paper where the deconvolution problem was studied for multivariate densities. It is worth noting that Masry (1993) considered more general weakly dependent observations and this paper formally does not deal with the minimax setting. However, the results obtained in this paper could be formally compared with the estimation under \mathbb{L}_∞ -loss over the *isotropic* Hölder class of regularity 2, that is, $\mathbb{N}_{\infty,d}(\vec{2}, 1)$ which is exactly the same setting as in Yuan and Chen (2002) in the case of partially contaminated observations. Let us also remark that there is no lower bound result in Masry (1993). The most general results in the deconvolution model were obtained in Comte and Lacour (2013) and Rebelles

(2016) and in Section 4 we will compare in detail our results with those obtained in these papers.

1.3. *Assumption on the function g.* Later on for any $U \in \mathbb{L}_1(\mathbb{R}^d)$, let \check{U} denote its Fourier transform. All our results will be established under the following condition.

ASSUMPTION 1. (1) If $\alpha \neq 1$, then there exists $\varepsilon > 0$ such that

$$|1 - \alpha + \alpha \check{g}(t)| \geq \varepsilon \quad \forall t \in \mathbb{R}^d.$$

(2) If $\alpha = 1$, then there exists $\vec{\mu} = (\mu_1, \dots, \mu_d) \in (0, \infty)^d$ and $\Upsilon_0 > 0$ s.t.

$$|\check{g}(t)| \geq \Upsilon_0 \prod_{j=1}^d (1 + t_j^2)^{-\frac{\mu_j}{2}} \quad \forall t = (t_1, \dots, t_d) \in \mathbb{R}^d.$$

Note that Assumption 1(1) is very weak and it is verified for many distributions, including centered multivariate Laplace and Gaussian ones. Note also that this assumption always holds with $\varepsilon = 1 - 2\alpha$ if $\alpha < 1/2$. Additionally, it holds with $\varepsilon = 1 - \alpha$ if \check{g} is a real positive function. The latter is true, in particular, for any probability law obtained by an even number of convolutions of a symmetric distribution with itself. At last, our Assumption 1(1) is weaker than the conditions imposed in Hesse (1995) and Yuan and Chen (2002). In these papers, $\check{g} \in \mathcal{C}^{(2)}(\mathbb{R}^d)$, $\check{g}(t) \neq 0$ for any $t \in \mathbb{R}^d$ and $|1 - \alpha + \alpha \check{g}(t)| \geq 1 - \alpha$ for any $t \in \mathbb{R}^d$. As to Assumption 1(2) [cf. (1.4)], it is much more restrictive. The centered multivariate Laplace law is an example in which this condition is satisfied.

2. Pointwise selection rule and \mathbb{L}_p -norm oracle inequality. To present our results in an unified way, let us define $\vec{\mu}(\alpha) = \vec{\mu}$, $\alpha = 1$, $\vec{\mu}(\alpha) = (0, \dots, 0)$, $\alpha \in [0, 1)$. Let $K : \mathbb{R}^d \rightarrow \mathbb{R}$ be a continuous function belonging to $\mathbb{L}_1(\mathbb{R}^d)$, $\int_{\mathbb{R}} K = 1$, and such that its Fourier transform \check{K} satisfies the following condition.

ASSUMPTION 2. There exist $\mathbf{k}_1 > 0$ and $\mathbf{k}_2 > 0$ such that

$$\int_{\mathbb{R}^d} |\check{K}(t)| \prod_{j=1}^d (1 + t_j^2)^{\frac{\mu_j(\alpha)}{2}} dt \leq \mathbf{k}_1, \quad \int_{\mathbb{R}^d} |\check{K}(t)|^2 \prod_{j=1}^d (1 + t_j^2)^{\mu_j(\alpha)} dt \leq \mathbf{k}_2^2.$$

Set $\mathcal{H} = \{e^k, k \in \mathbb{Z}\}$ and let $\vec{\mathcal{H}}^d = \{\vec{h} = (h_1, \dots, h_d) : h_j \in \mathcal{H}, j = 1, \dots, d\}$. Define for any $\vec{h} = (h_1, \dots, h_d) \in \vec{\mathcal{H}}^d$

$$K_{\vec{h}}(t) = V_{\vec{h}}^{-1} K(t_1/h_1, \dots, t_d/h_d), \quad t \in \mathbb{R}^d, V_{\vec{h}} = \prod_{j=1}^d h_j.$$

Later on for any $u, v \in \mathbb{R}^d$ the operations and relations $u/v, uv, u \vee v, u \wedge v, u \geq v, au, a \in \mathbb{R}$, are understood in coordinatewise sense. In particular, $u \geq v$ means that $u_j \geq v_j$ for any $j = 1, \dots, d$.

2.1. *Pointwise selection rule from the family of kernel estimators.* For any $\vec{h} \in (0, \infty)^d$, let $M(\cdot, \vec{h})$ satisfy the operator equation

$$(2.1) \quad K_{\vec{h}}(y) = (1 - \alpha)M(y, \vec{h}) + \alpha \int_{\mathbb{R}^d} g(t - y)M(t, \vec{h}) dt, \quad y \in \mathbb{R}^d.$$

Note that although the explicit expression of $M(\cdot, \vec{h})$ is not available, its Fourier transform can be easily deduced from (2.1); see Section 5.1.2. Define

$$\widehat{f}_{\vec{h}}(x) = n^{-1} \sum_{i=1}^n M(Z_i - x, \vec{h}), \quad x \in \mathbb{R}^d, \vec{h} \in \mathcal{H}^d.$$

Our first goal is to propose for any given $x \in \mathbb{R}^d$ a data driven selection rule from the family of estimators $\mathcal{F}(\mathbb{H}) = \{\widehat{f}_{\vec{h}}(x), \vec{h} \in \mathbb{H}\}$, where \mathbb{H} is an arbitrary subset of \mathcal{H}^d . Set for any $\vec{h} \in \mathcal{H}^d$ and $x \in \mathbb{R}^d$

$$\widehat{U}_n(x, \vec{h}) = \sqrt{2n^{-1}\lambda_n(\vec{h})\widehat{\sigma}^2(x, \vec{h})} + \frac{4}{3}n^{-1}M_\infty\lambda_n(\vec{h}) \prod_{j=1}^d h_j^{-1}(h_j \wedge 1)^{-\mu_j(\alpha)},$$

where we have put $\widehat{\sigma}^2(x, \vec{h}) = \frac{1}{n} \sum_{i=1}^n M^2(Z_i - x, \vec{h})$ and

$$\lambda_n(\vec{h}) = 4 \ln(M_\infty) + 6 \ln(n) + (8p + 26) \sum_{j=1}^d [1 + \mu_j(\alpha)] |\ln(h_j)|,$$

$$M_\infty = [(2\pi)^{-d} \{\varepsilon^{-1} \|\check{K}\|_1 1_{\alpha \neq 1} + \Upsilon_0^{-1} \mathbf{k}_1 1_{\alpha=1}\}] \vee 1.$$

Pointwise selection rule. Let \mathbb{H} be an arbitrary subset of \mathcal{H}^d . For any $\vec{h} \in \mathbb{H}$ and $x \in \mathbb{R}^d$, introduce $\widehat{U}_n^*(x, \vec{h}) = \sup_{\vec{\eta} \in \mathbb{H}: \vec{\eta} \geq \vec{h}} \widehat{U}_n(x, \vec{\eta})$,

$$\widehat{\mathcal{R}}_{\vec{h}}(x) = \sup_{\vec{\eta} \in \mathbb{H}} [|\widehat{f}_{\vec{h} \vee \vec{\eta}}(x) - \widehat{f}_{\vec{\eta}}(x)| - 4\widehat{U}_n(x, \vec{h} \vee \vec{\eta}) - 4\widehat{U}_n(x, \vec{\eta})]_+,$$

$$(2.2) \quad \vec{\mathbf{h}}(x) = \arg \inf_{\vec{h} \in \mathbb{H}} [\widehat{\mathcal{R}}_{\vec{h}}(x) + 8\widehat{U}_n^*(x, \vec{h})].$$

Our final estimator is $\widehat{f}_{\vec{\mathbf{h}}(x)}(x), x \in \mathbb{R}^d$ and we will call (2.2) the *pointwise selection rule*. Note that the estimator $\widehat{f}_{\vec{\mathbf{h}}(\cdot)}(\cdot)$ does not necessarily belong to the collection $\mathcal{F}(\mathbb{H})$ since the multibandwidth $\vec{\mathbf{h}}(\cdot)$ is a d -variate function, which is not necessarily constant on \mathbb{R}^d . The latter fact allows to take into account the “local structure” of the function to be estimated. Moreover, $\vec{\mathbf{h}}(\cdot)$ is chosen with respect to the observations and, therefore, it is a random vector function.

2.2. *\mathbb{L}_p -norm oracle inequality.* Introduce for any $x \in \mathbb{R}^d$ and $\vec{h} \in \mathcal{H}^d$

$$U_n^*(x, \vec{h}) = \sup_{\vec{\eta} \in \mathcal{H}^d: \vec{\eta} \geq \vec{h}} U_n(x, \vec{\eta}), \quad S_{\vec{h}}(x, f) = \int_{\mathbb{R}^d} K_{\vec{h}}(t - x) f(t) \nu_d(dt),$$

where we have put

$$U_n(x, \vec{\eta}) = \sqrt{2n^{-1}\lambda_n(\vec{\eta})\sigma^2(x, \vec{\eta})} + \frac{4}{3}n^{-1}M_\infty\lambda_n(\vec{\eta}) \prod_{j=1}^d \eta_j^{-1}(\eta_j \wedge 1)^{-\mu_j(\alpha)}$$

and $\sigma^2(x, \vec{\eta}) = \int_{\mathbb{R}^d} M^2(t - x, \vec{\eta})\mathfrak{p}(t)\nu_d(dt)$.

For any $\mathbb{H} \subseteq \mathcal{H}^d$, $\vec{h} \in \mathbb{H}$ and $x \in \mathbb{R}^d$ introduce also

$$(2.3) \quad \begin{aligned} B_{\vec{h}}^*(x, f) &= \sup_{\vec{\eta} \in \mathbb{H}} |S_{\vec{h} \vee \vec{\eta}}(x, f) - S_{\vec{\eta}}(x, f)|, \\ B_{\vec{h}}(x, f) &= |S_{\vec{h}}(x, f) - f(x)|. \end{aligned}$$

THEOREM 1. *Let Assumptions 1 and 2 be fulfilled. Then for any $\mathbb{H} \subseteq \mathcal{H}^d$, $n \geq 3$, $p \in [1, \infty)$ and any $f \in \mathbb{F}_g(\mathbb{R})$*

$$\mathcal{R}_n^{(p)}[\widehat{f}_{\mathbb{H}(\cdot)}, f] \leq \left\| \inf_{\vec{h} \in \mathbb{H}} \{2B_{\vec{h}}^*(\cdot, f) + B_{\vec{h}}(\cdot, f) + 49U_n^*(\cdot, \vec{h})\} \right\|_p + \mathbf{C}_p n^{-\frac{1}{2}}.$$

The explicit expression for the constant \mathbf{C}_p (independent of f , n and \mathbb{H}) can be found in the proof of the theorem. Later on we will pay attention to a special choice for the collection of multibandwidths, namely

$$\mathcal{H}_{\text{isotr}}^d := \{\vec{h} \in \mathcal{H}^d : \vec{h} = (h, \dots, h), h \in \mathcal{H}\}.$$

In Section 4, the selection from the corresponding family of kernel estimators will be used for the adaptive estimation over the collection of isotropic Nikol'skii classes. Note also that if $\mathbb{H} = \mathcal{H}_{\text{isotr}}^d$ then obviously

$$B_{\vec{h}}^*(\cdot, f) \leq 2 \sup_{\vec{\eta} \in \mathcal{H}_{\text{isotr}}^d : \eta \leq h} B_{\vec{\eta}}(\cdot, f) \quad \forall \vec{h} = (h, \dots, h) \in \mathcal{H}_{\text{isotr}}^d$$

and we come to the following corollary of Theorem 1.

COROLLARY 1. *Let Assumptions 1 and 2 be fulfilled and the selection rule runs $\mathbb{H} = \mathcal{H}_{\text{isotr}}^d$. Then for any $n \geq 3$, $p \in [1, \infty)$ and any $f \in \mathbb{F}_g(\mathbb{R})$*

$$\mathcal{R}_n^{(p)}[\widehat{f}_{\mathbb{H}(\cdot)}, f] \leq \left\| \inf_{\vec{h} \in \mathcal{H}_{\text{isotr}}^d} \left\{ 5 \sup_{\vec{\eta} \in \mathcal{H}_{\text{isotr}}^d : \eta \leq h} B_{\vec{\eta}}(\cdot, f) + 49U_n^*(\cdot, \vec{h}) \right\} \right\|_p + \mathbf{C}_p n^{-\frac{1}{2}}.$$

The oracle inequality proved in Theorem 1 is particularly useful since it does not require any assumption on the underlying function f (except for the restrictions ensuring the existence of the model and of the risk). However, the quantity appearing in the right-hand side of this inequality, namely

$$\left\| \inf_{\vec{h} \in \mathbb{H}} \{2B_{\vec{h}}^*(\cdot, f) + B_{\vec{h}}(\cdot, f) + 49U_n^*(\cdot, \vec{h})\} \right\|_p$$

is not easy to analyze. In particular, in order to use the result of Theorem 1 for adaptive estimation, one has to be able to compute

$$\sup_{f \in \mathbb{F}} \left\| \inf_{\vec{h} \in \mathbb{H}} \{2B_{\vec{h}}^*(\cdot, f) + B_{\vec{h}}(\cdot, f) + 49U_n^*(\cdot, \vec{h})\} \right\|$$

for a given class $\mathbb{F} \subset \mathbb{L}_p(\mathbb{R}^d) \cap \mathbb{F}_g(R)$ with either $\mathbb{H} = \mathcal{H}^d$ or $\mathbb{H} = \mathcal{H}_{\text{isotr}}^d$. It turns out that under some nonrestrictive assumptions imposed on f , the obtained bounds can be considerably simplified. Moreover, new inequalities will allow us to better understand the way for proving adaptive results.

3. Abstract upper bound theorem. Define $\forall \mathbf{u} \in [1, \infty], D > 0,$

$$\mathbb{F}_{g, \mathbf{u}}(R, D) := \{f \in \mathbb{F}_g(R) : (1 - \alpha)f + \alpha[f \star g] \in \mathbb{B}_{\mathbf{u}, d}^{(\infty)}(D)\},$$

where $\mathbb{B}_{\mathbf{u}, d}^{(\infty)}(D)$ is the ball of radius D in the weak-type space $\mathbb{L}_{\mathbf{u}, \infty}(\mathbb{R}^d)$, that is, $\mathbb{B}_{\mathbf{u}, d}^{(\infty)}(D) = \{\lambda : \mathbb{R}^d \rightarrow \mathbb{R} : \|\lambda\|_{\mathbf{u}, \infty} < D\}$, where

$$\|\lambda\|_{\mathbf{u}, \infty} = \inf\{C : \nu_d(x : |\lambda(x)| > \mathfrak{z}) \leq C^{\mathbf{u}} \mathfrak{z}^{-\mathbf{u}}, \forall \mathfrak{z} > 0\}.$$

As usual $\mathbb{B}_{\infty, d}^{(\infty)}(D) = \mathbb{B}_{\infty, d}(D)$ and obviously $\mathbb{B}_{\mathbf{u}, d}^{(\infty)}(D) \supset \mathbb{B}_{\mathbf{u}, d}(D)$.

It is worth noting that the assumption $f \in \mathbb{F}_{g, \mathbf{u}}(R, D)$ simply means that the common density of the observations \mathbf{p} belongs to $\mathbb{B}_{\mathbf{u}, d}^{(\infty)}(D)$. Our objective is to bound from above $\sup_{f \in \mathbb{F}} \mathcal{R}_n^{(p)}[\widehat{f}_{\vec{h}(\cdot)}, f]$ for any $\mathbb{F} \subset \mathbb{F}_{g, \mathbf{u}}(R, D) \cap \mathbb{B}_{\mathbf{q}, d}(D)$, where $\mathbf{q} \in [1, \infty]$. Since \mathbb{F} is an arbitrary set, this bound can be applied to the adaptation over different scales of functional classes. In particular, the results below form the basis for our consideration in Section 4.

Introduce for any $\vec{h} \in \mathcal{H}^d$

$$F_n(\vec{h}) = \frac{\sqrt{\ln n + \sum_{j=1}^d |\ln h_j|}}{\sqrt{n} \prod_{j=1}^d h_j^{\frac{1}{2}} (h_j \wedge 1)^{\mu_j(\alpha)}}, \quad G_n(\vec{h}) = \frac{\ln n + \sum_{j=1}^d |\ln h_j|}{n \prod_{j=1}^d h_j (h_j \wedge 1)^{\mu_j(\alpha)}}.$$

Furthermore let \mathbb{H} be either \mathcal{H}^d or $\mathcal{H}_{\text{isotr}}^d$ and for any $v, z > 0$ define

$$\mathfrak{H}(v) = \{\vec{h} \in \mathbb{H} : G_n(\vec{h}) \leq av\}, \quad \mathfrak{H}(v, z) = \{\vec{h} \in \mathfrak{H}(v) : F_n(\vec{h}) \leq avz^{-\frac{1}{2}}\}.$$

Here, $a > 0$ is a numerical constant whose explicit expression is given in the beginning of Section 5.2. Put also for any $v > 0, l_{\mathbb{H}}(v) = v^{p-1} (1 + |\ln(v)|)^{t(\mathbb{H})}$, where $t(\mathbb{H}) = d - 1$ if $\mathbb{H} = \mathcal{H}^d$ and $t(\mathbb{H}) = 0$ if $\mathbb{H} = \mathcal{H}_{\text{isotr}}^d$.

REMARK 1. Note that whatever the values of $v > 0$ and $z \geq 2, \mathfrak{H}(v) \neq \emptyset$ and $\mathfrak{H}(v, z) \neq \emptyset$ since one can find $b > 1$ such that

$$(\ln n + d \ln b)(nb^d)^{-1} \leq [a^2 v^2 z^{-1}] \wedge av.$$

The latter means that $\vec{b} = (b, \dots, b) \in \mathfrak{H}(v, z) \cap \mathfrak{H}(v)$.

All the results in this section will be proved under an additional condition imposed on the kernel K .

ASSUMPTION 3. Let $K(x) = \prod_{j=1}^d \mathcal{K}(x_j)$, $x \in \mathbb{R}^d$, where $\mathcal{K} : \mathbb{R} \rightarrow \mathbb{R}$ is a compactly supported, bounded function and $\int \mathcal{K} = 1$.

Without loss of generality, we will assume that $\|\mathcal{K}\|_\infty \geq 1$ and $\text{supp}(\mathcal{K}) \subset [-c_{\mathcal{K}}, c_{\mathcal{K}}]$ with $c_{\mathcal{K}} \geq 1$.

Introduce the following notation. Set for any $h \in \mathcal{H}$ and $j = 1, \dots, d$

$$b_{h,f,j}^*(\cdot) = \left| \int_{\mathbb{R}} \mathcal{K}(u) f(\cdot + u\mathbf{e}_j) v_1(du) - f(\cdot) \right|, \quad b_{h,f,j}(\cdot) = \sup_{\eta \in \mathcal{H}: \eta \leq h} b_{\eta,f,j}^*(\cdot),$$

where $(\mathbf{e}_1, \dots, \mathbf{e}_d)$ denotes the canonical basis of \mathbb{R}^d . Introduce $\forall s \in [1, \infty]$

$$\mathbf{B}_{j,s,\mathbb{F}}^*(\mathbf{h}) = \sup_{f \in \mathbb{F}} \sum_{h \in \mathcal{H}: h \leq \mathbf{h}} \|b_{h,f,j}^*\|_s,$$

$$\mathbf{B}_{j,s,\mathbb{F}}(\mathbf{h}) = \sup_{f \in \mathbb{F}} \|b_{\mathbf{h},f,j}\|_s, \quad j = 1, \dots, d.$$

For any $v > 0$ and $j = 1, \dots, d$, set $\mathbf{V}_j(v) = \{\mathbf{h} \in \mathcal{H} : \mathbf{B}_{j,\infty,\mathbb{F}}(\mathbf{h}) \leq \mathbf{c}v\}$ and

$$J(\vec{h}, v) = \{j \in \{1, \dots, d\} : h_j \in \mathbf{V}_j(v)\}, \quad \vec{h} \in \mathcal{H}^d,$$

where $\mathbf{c} = (20d)^{-1}[\max(2c_{\mathcal{K}}\|\mathcal{K}\|_\infty, \|\mathcal{K}\|_1)]^{-d}$. As usual the complement of $J(\vec{h}, v)$ will be denoted by $\bar{J}(\vec{h}, v)$. Furthermore, the summation over the empty set is supposed to be zero.

For any $\vec{s} = (s_1, \dots, s_d) \in [1, \infty)^d$, $\mathbf{u} \geq 1$ and $v > 0$ introduce

$$(3.1) \quad \Lambda_{\vec{s}}(v, \mathbb{F}, \mathbf{u}) = \inf_{z \geq 2} \inf_{\vec{h} \in \mathfrak{H}(v,z)} \left[\sum_{j \in \bar{J}(\vec{h},v)} v^{-s_j} [\mathbf{B}_{j,s_j,\mathbb{F}}(h_j)]^{s_j} + z^{-\mathbf{u}} \right];$$

$$(3.2) \quad \Lambda_{\vec{s}}(v, \mathbb{F}) = \inf_{\vec{h} \in \mathfrak{H}(v)} \left[\sum_{j \in \bar{J}(\vec{h},v)} v^{-s_j} [\mathbf{B}_{j,s_j,\mathbb{F}}(h_j)]^{s_j} + v^{-2} F_n^2(\vec{h}) \right].$$

Since the sets $\mathfrak{H}(v)$ and $\mathfrak{H}(v, z)$ both depend on n , the quantities above depend on n as well. We omit this dependence to simplify the notation.

THEOREM 2. *Let assumptions of Theorem 1 be fulfilled and suppose additionally that K satisfies Assumption 3. Then for any $n \geq 3$, $p > 1$, $\mathbf{q} > 1$, $R > 1$, $D > 0$, $0 < \underline{v} \leq \bar{v} \leq \infty$, $\mathbf{u} \in (p/2, \infty]$, $\mathbf{u} \geq \mathbf{q}$, $\vec{s} \in (1, \infty)^d$, $\vec{q} \in [p, \infty)^d$ and any $\mathbb{F} \subset \mathbb{B}_{\mathbf{q},d}(D) \cap \mathbb{F}_{\mathbf{g},\mathbf{u}}(R, D)$*

$$\sup_{f \in \mathbb{F}} \mathcal{R}_n^{(p)}[\widehat{f}_{\mathbf{h}(\cdot)}, f] \leq C^{(2)} \left[l_{\mathbb{H}}(\underline{\mathbf{v}}) + \int_{\underline{\mathbf{v}}}^{\bar{\mathbf{v}}} v^{p-1} [\Lambda_{\vec{s}}(v, \mathbb{F}, \mathbf{u}) \wedge \Lambda_{\vec{s}}(v, \mathbb{F})] dv \right. \\ \left. + \bar{\mathbf{v}}^p \Lambda_{\vec{q}}(\bar{\mathbf{v}}, \mathbb{F}, \mathbf{u}) \right]^{\frac{1}{p}} + C_p n^{-\frac{1}{2}}.$$

If additionally $\mathbf{q} \in (p, \infty)$, one has also

$$\sup_{f \in \mathbb{F}} \mathcal{R}_n^{(p)}[\widehat{f}_{\mathbf{h}(\cdot)}, f] \leq C^{(2)} \left[l_{\mathbb{H}}(\underline{\mathbf{v}}) + \int_{\underline{\mathbf{v}}}^{\bar{\mathbf{v}}} v^{p-1} [\mathbf{\Lambda}_{\vec{s}}(v, \mathbb{F}, \mathbf{u}) \wedge \mathbf{\Lambda}_{\vec{s}}(v, \mathbb{F})] dv + \bar{\mathbf{v}}^{p-\mathbf{q}} \right]^{\frac{1}{p}} + \mathbf{C}_p n^{-\frac{1}{2}}.$$

Moreover, if $\mathbf{q} = \infty$ one has

$$\sup_{f \in \mathbb{F}} \mathcal{R}_n^{(p)}[\widehat{f}_{\mathbf{h}(\cdot)}, f] \leq C^{(2)} \left[l_{\mathbb{H}}(\underline{\mathbf{v}}) + \int_{\underline{\mathbf{v}}}^{\bar{\mathbf{v}}} v^{p-1} [\mathbf{\Lambda}_{\vec{s}}(v, \mathbb{F}, \mathbf{u}) \wedge \mathbf{\Lambda}_{\vec{s}}(v, \mathbb{F})] dv + \mathbf{\Lambda}_{\vec{s}}(\bar{\mathbf{v}}, \mathbb{F}, \mathbf{u}) \right]^{\frac{1}{p}} + \mathbf{C}_p n^{-\frac{1}{2}}.$$

Finally, if $\mathbb{H} = \mathcal{H}_{\text{isotr}}^d$ all the assertions above remain true for any $\vec{s} \in [1, \infty)^d$ if one replaces in (3.1)–(3.2) $\mathbf{B}_{j,s_j,\mathbb{F}}(\cdot)$ by $\mathbf{B}_{j,s_j,\mathbb{F}}^*(\cdot)$.

In all the assertions, the third terms are assumed to be zero when $\bar{\mathbf{v}} = \infty$.

1⁰. It is important to emphasize that $C^{(2)}$ depends only on $\vec{s}, \vec{q}, g, \mathcal{K}, d, R, D, \mathbf{u}$ and \mathbf{q} . Note also that the assertions of the theorem remain true if we minimize the right-hand sides of the obtained inequalities w.r.t. \vec{s}, \vec{q} since their left-hand sides are independent of \vec{s} and \vec{q} . In this context, it is important to realize that $C^{(2)} = C^{(2)}(\vec{s}, \dots)$ is bounded for any $\vec{s} \in (1, \infty)^d$ but $C^{(2)}(\vec{s}, \dots)$ is unbounded if there exists $j = 1, \dots, d$ such that $s_j = 1$. Contrary to that $C^{(2)}(\vec{s}, \dots) < \infty$ for any $\vec{s} \in [1, \infty)^d$ if $\mathbb{H} = \mathcal{H}_{\text{isotr}}^d$ and it explains in particular the fourth assertion of the theorem.

2⁰. It is worth noting that all the bounds presented in the theorem are mostly based on the result given in (5.39) of Section 5.2. This is an \mathbb{L}_p -norm oracle inequality on $\mathbb{F}_{g,\mathbf{u}}(R, D) \cap \mathbb{B}_{\mathbf{q},d}(D)$ having independent interest. In particular, it does not require Assumption 3 and it is established for any compactly supported K satisfying Assumption 2.

3⁰. Note also that $D, R, \mathbf{u}, \mathbf{q}$ are not involved in the construction of our pointwise selection rule. That means that one and the same estimator can be actually applied on any $\mathbb{F} \subset \bigcup_{R,D,\mathbf{u},\mathbf{q}} \mathbb{B}_{\mathbf{q},d}(D) \cap \mathbb{F}_{g,\mathbf{u}}(R, D)$. Moreover, the assertion of the theorem has a nonasymptotical nature; we do not suppose that the number of observations n is large.

4⁰. As we see, the application of our results to some functional class is mainly reduced to the computation of the functions $\mathbf{B}_{j,s,\mathbb{F}}(\cdot), j = 1, \dots, d$, for some properly chosen s . Note however that this task is not necessary for many functional classes, at least for the classes defined by the help of kernel approximation. Indeed, a typical description of \mathbb{F} can be summarized as follows. Let $\lambda_j : \mathbb{R}_+ \rightarrow \mathbb{R}_+$,

be such that $\lambda_j(0) = 0, \lambda_j \uparrow$ for any $j = 1, \dots, d$. Then the functional class is defined as a collection of functions satisfying

$$(3.3) \quad \|b_{\mathbf{h},f,j}\|_{r_j} \leq \lambda_j(\mathbf{h}) \quad \forall \mathbf{h} \in \mathcal{H}$$

for some $\vec{r} \in [1, \infty]$. It yields obviously $\mathbf{B}_{j,r_j,\mathbb{F}}(\cdot) \leq \lambda_j(\cdot)$ for any $j = 1, \dots, d$, and the result of Theorem 2 remains valid if we replace formally $\mathbf{B}_{j,r_j,\mathbb{F}}(\cdot)$ by $\lambda_j(\cdot)$ in all the expressions appearing in this theorem.

In the Appendix (proof of Lemma 4), we show that for some particular kernel K^* , the anisotropic Nikol'skii class $\mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L})$ is included into the class defined by (3.3) with $\lambda_j(\mathbf{h}) = L_j \mathbf{h}^{\beta_j}$, whatever the values of $\vec{\beta}, \vec{L}$ and \vec{r} .

4. Adaptive estimation over the scale of anisotropic classes. As we have seen, the estimator $\hat{f}_{\mathbf{h}}$ depends on \mathbb{H} and later on we will consider two choices of the set \mathbb{H} , namely $\mathbb{H} = \mathcal{H}^d$ and $\mathbb{H} = \mathcal{H}_{\text{isotr}}^d$. So, to present our results we will write $\hat{f}_{\mathbf{h},\mathbb{H}}$ to underline the aforementioned dependence.

4.1. *Anisotropic Nikol'skii classes.* Let $(\mathbf{e}_1, \dots, \mathbf{e}_d)$ denote the canonical basis of \mathbb{R}^d . For some function $G : \mathbb{R}^d \rightarrow \mathbb{R}^1$ and real number $u \in \mathbb{R}$ define the first-order difference operator with step size u in direction of the variable x_j by $\Delta_{u,j}G(x) = G(x + u\mathbf{e}_j) - G(x), j = 1, \dots, d$. By induction, the k th order difference operator with step size u in direction of the variable x_j is

$$\Delta_{u,j}^k G(x) = \Delta_{u,j} \Delta_{u,j}^{k-1} G(x) = \sum_{l=1}^k (-1)^{l+k} \binom{k}{l} \Delta_{ul,j} G(x).$$

DEFINITION 1. For given vectors, $\vec{\beta} = (\beta_1, \dots, \beta_d) \in (0, \infty)^d, \vec{r} = (r_1, \dots, r_d) \in [1, \infty]^d$, and $\vec{L} = (L_1, \dots, L_d) \in (0, \infty)^d$ we say that a function $G : \mathbb{R}^d \rightarrow \mathbb{R}^1$ belongs to the anisotropic Nikol'skii class $\mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L})$ if $\|G\|_{r_j} \leq L_j$ for all $j = 1, \dots, d$ and there exists natural number $k_j > \beta_j$ such that

$$\|\Delta_{u,j}^{k_j} G\|_{r_j} \leq L_j |u|^{\beta_j} \quad \forall u \in \mathbb{R}, \forall j = 1, \dots, d.$$

If $\beta_j = \boldsymbol{\beta} \in (0, \infty), r_j = \mathbf{r} \in [1, \infty]$ and $L_j = \mathbf{L} \in (0, \infty)$ for any $j = 1, \dots, d$ the corresponding Nikol'skii class, denoted furthermore $\mathbb{N}_{\mathbf{r},d}(\boldsymbol{\beta}, \mathbf{L})$, is called isotropic. The following quantities related to the parameters of the Nikol'skii class will be very important in the sequel:

$$\frac{1}{\beta(\alpha)} = \sum_{j=1}^d \frac{2\mu_j(\alpha) + 1}{\beta_j}, \quad \frac{1}{\omega(\alpha)} = \sum_{j=1}^d \frac{2\mu_j(\alpha) + 1}{\beta_j r_j},$$

$$L(\alpha) = \prod_{j=1}^d L_j \frac{2\mu_j(\alpha)+1}{\beta_j}.$$

Define also for any $1 \leq s \leq \infty$ and $\alpha \in [0, 1]$

$$\varkappa_\alpha(s) = \omega(\alpha)(2 + 1/\beta(\alpha)) - s, \quad \tau(s) = 1 - 1/\omega(0) + 1/(s\beta(0)).$$

4.2. *Construction of kernel K.* We keep Assumption 2 and enforce Assumption 3 by Assumption 4 below related to the following specific construction of kernel K used in the definition of the family of estimators $\{\widehat{f}_{\vec{h}}(\cdot), \vec{h} \in \mathcal{H}^d\}$ [see, e.g., Kerkyacharian, Lepski and Picard (2001) or Goldenshluger and Lepski (2014)]. Let ℓ be an integer number, $\mathcal{K} : \mathbb{R}^1 \rightarrow \mathbb{R}^1$ be a compactly supported continuous function satisfying $\int_{\mathbb{R}^1} \mathcal{K}(y) dy = 1$. Put

$$\mathcal{K}_\ell(y) = \sum_{i=1}^{\ell} \binom{\ell}{i} (-1)^{i+1} \frac{1}{i} \mathcal{K}\left(\frac{y}{i}\right),$$

and add the following structural condition to Assumption 2.

ASSUMPTION 4. $K(x) = \prod_{j=1}^d \mathcal{K}_\ell(x_j), \forall x \in \mathbb{R}^d$.

4.3. *Main results.* Set $\delta_n = L(\alpha)n^{-1} \ln(n)$, $t(\mathbb{H}) = d - 1$ if $\mathbb{H} = \mathcal{H}^d$ and $t(\mathbb{H}) = 0$ if $\mathbb{H} = \mathcal{H}_{\text{isotr}}^d$ and let

$$\mathfrak{b}_n(\mathbb{H}) = \begin{cases} [\ln(n)]^{t(\mathbb{H})}, & \varkappa_\alpha(p) > p\omega(\alpha); \\ \ln^{\frac{1}{p}}(n) \vee [\ln(n)]^{t(\mathbb{H})}, & \varkappa_\alpha(p) = p\omega(\alpha); \\ \ln^{\frac{1}{p}}(n), & \varkappa_\alpha(p) = 0; \\ 1 & \text{otherwise.} \end{cases}$$

4.3.1. *Bounded case.* The first problem we address is the adaptive estimation over the collection of the functional classes $\{\mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L}) \cap \mathbb{F}_g(R) \cap \mathbb{B}_{\infty,d}(Q)\}_{\vec{\beta}, \vec{r}, \vec{L}, R, Q}$. The results obtained in Theorem 3 together with those from Theorem 2 in Lepski and Willer (2017) show that the boundedness of the function belonging to $\mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L}) \cap \mathbb{F}_g(R)$ is a minimal condition allowing to eliminate the inconsistency zone, as it was conjectured in Lepski and Willer (2017). Define

$$\rho(\alpha) = \begin{cases} \frac{1 - 1/p}{1 - 1/\omega(\alpha) + 1/\beta(\alpha)}, & \varkappa_\alpha(p) > p\omega(\alpha); \\ \frac{\beta(\alpha)}{2\beta(\alpha) + 1}, & 0 < \varkappa_\alpha(p) \leq p\omega(\alpha); \\ \frac{\tau(p)\omega(\alpha)\beta(0)}{z(\alpha)}, & \varkappa_\alpha(p) \leq 0, \tau(\infty) > 0; \\ \frac{\omega(\alpha)}{p}, & \varkappa_\alpha(p) \leq 0, \tau(\infty) \leq 0. \end{cases}$$

THEOREM 3. Let $\alpha \in [0, 1]$, $\ell \in \mathbb{N}^*$ and $g \in \mathbb{L}_1(\mathbb{R}^d)$, satisfying Assumption 1, be fixed. Let K satisfy Assumptions 2 and 4:

(1) Then for any $p \in (1, \infty)$, $Q > 0$, $R > 0$, $L_0 > 0$, $\vec{\beta} \in (0, \ell]^d$, $\vec{r} \in (1, \infty]^d$ and $\vec{L} \in [L_0, \infty)^d$ there exists $C < \infty$, independent of \vec{L} , such that

$$\limsup_{n \rightarrow \infty} \sup_{f \in \mathbb{N}_{\vec{r}, d}(\vec{\beta}, \vec{L}) \cap \mathbb{F}_g(R) \cap \mathbb{B}_{\infty, d}(Q)} \mathfrak{b}_n(\mathcal{H}^d)^{-1} \delta_n^{-\rho(\alpha)} \mathcal{R}_p^{(n)}[\widehat{f}_{\mathbf{h}, \mathcal{H}^d}; f] \leq C.$$

(2) For any $p \in (1, \infty)$, $Q > 0$, $R > 0$, $L_0 > 0$, $\beta \in (0, \ell]$, $\mathbf{r} \in [1, \infty]$ and $\mathbf{L} \in [L_0, \infty)$ there exists $C < \infty$, independent of \mathbf{L} , such that

$$\limsup_{n \rightarrow \infty} \sup_{f \in \mathbb{N}_{\mathbf{r}, d}(\beta, \mathbf{L}) \cap \mathbb{F}_g(R) \cap \mathbb{B}_{\infty, d}(Q)} \mathfrak{b}_n(\mathcal{H}_{\text{isotr}}^d)^{-1} \delta_n^{-\rho(\alpha)} \mathcal{R}_p^{(n)}[\widehat{f}_{\mathbf{h}, \mathcal{H}_{\text{isotr}}^d}; f] \leq C.$$

1⁰. Our estimation procedure is completely data driven, that is, independent of $\vec{\beta}, \vec{r}, \vec{L}, R, Q$ and the assertions of Theorem 3 are completely new if $\alpha \neq 0$. Comparing the results obtained in Theorem 3 with those proved in Theorem 2 in Lepski and Willer (2017), we can assert that our estimator is optimally adaptive if $\varkappa_\alpha(p) < 0$ and nearly optimally adaptive if $0 < \varkappa_\alpha(p) < p\omega(\alpha)$. The construction of an estimation procedure which would be optimally adaptive when $\varkappa_\alpha(p) \geq 0$ is an open problem, and we conjecture that the lower bounds for the asymptotics of the minimax risk found in Theorem 2 in Lepski and Willer (2017) are sharp in order. This conjecture in the case $\alpha = 1$ is partially confirmed by the results obtained in Comte and Lacour (2013) and Rebelles (2016). Since both articles deal with the estimation of unbounded functions we will discuss them in the next section.

2⁰. We note that the asymptotic of the minimax risk under partially contaminated observations, $\alpha \in (0, 1)$, is independent of α and coincides with the asymptotic of the risk in the direct observation model, $\alpha = 0$. This phenomenon was discovered in Hesse (1995) and Yuan and Chen (2002). In the very recent papers, Duval (2017) and Lepski (2018), the particular case $\vec{r} = (r, \dots, r)$, $r \in (1, \infty)$ was studied. In Duval (2017), the isotropic classes, compactly supported densities and $\alpha < 1/2$ were considered and a nearly optimally adaptive estimator was built via wavelet thresholding. In Lepski (2018), the anisotropic classes on \mathbb{R}^d with $r = p$ were studied under the same condition $\alpha < 1/2$, except for the case $p = 2$, in which $\alpha \in (0, 1)$, and an optimally adaptive estimator was built via a plug-in method.

3⁰. As to the direct observation scheme, $\alpha = 0$, our results coincide with those obtained recently in Goldenshluger and Lepski (2014), when $p\omega(0) > \varkappa_0(p)$. However, for the tail zone $p\omega(0) \leq \varkappa_0(p)$, our bound is slightly better since the bound obtained in the latter paper contains an additional factor $\ln^{\frac{d}{p}}(n)$. It is interesting to note that although both estimator constructions are based upon local selections from the family of kernel estimators, the selection rules are different.

4⁰. Let us discuss the results corresponding to the *tail zone*, $\varkappa_\alpha(p) > p\omega(\alpha)$. The lower bound for the minimax risk is given by $[L(\alpha)n^{-1}]^{\rho(\alpha)}$ [see Lepski and Willer (2017)], while the accuracy provided by our estimator is

$$[\ln(n)]^{t(\mathbb{H})} [L(\alpha)n^{-1} \ln(n)]^{\rho(\alpha)}.$$

As it was mentioned, the passage from $[L(\alpha)n^{-1}]^{\rho(\alpha)}$ to $[L(\alpha)n^{-1} \ln(n)]^{\rho(\alpha)}$ seems to be an unavoidable payment for the application of a local selection scheme. It is interesting to note that the additional factor $[\ln(n)]^{t(\mathbb{H})}$ disappears in the dimension $d = 1$. First, note that if $\alpha = 0$ the one-dimensional setting was considered in Juditsky and Lambert-Lacroix (2004) and Reynaud-Bouret, Rivoirard and Tuleau-Malot (2011). The setting of Juditsky and Lambert-Lacroix (2004) corresponds to $r = \infty$, while Reynaud-Bouret, Rivoirard and Tuleau-Malot (2011) deal with the case of $p = 2$ and $\tau(2) > 0$. Both settings rule out the sparse zone. The rates of convergence found in these papers are easily recovered from our results corresponding to the tail and dense zones. Also, we remark that the aforementioned factor appears only when anisotropic functional classes are considered, as can be seen in the second assertion of Theorem 3. A natural question is whether the $[\ln(n)]^{d-1}$ -factor is an unavoidable payment for anisotropy of the underlying function or not?

5⁰. We finish our discussion with the following remark. If $\alpha \neq 1$, the assumption $f \in \mathbb{F}_{g,\infty}(R, Q)$ implies in many cases that f is uniformly bounded and, therefore, Theorem 3 is applicable. In particular, it is always the case if the model (1.2) is considered. Indeed $f, g \in \mathfrak{F}(\mathbb{R}^d)$ in this case, which implies $\|f\|_\infty \leq (1 - \alpha)^{-1} \|p\|_\infty \leq (1 - \alpha)^{-1} Q$. Another case is $\|g\|_\infty < \infty$ and recall that this assumption was used in the proofs of Theorems 1 and 2 in Lepski and Willer (2017). We obviously have that

$$\|f\|_\infty \leq (1 - \alpha)^{-1} [Q + \alpha R \|g\|_\infty].$$

More generally, $\|f\|_\infty \leq (1 - \alpha)^{-1}(Q + \alpha D)$ if $f \in \mathbb{F}_{g,\infty}(R, Q)$ and $\|f \star g\|_\infty \leq D$. Since the definition of the Nikol'skii class implies that $\|f\|_{r^*} \leq L^*$, where $r^* = \sup_{j=1,\dots,d} r_j$ and $L^* = \sup_{j=1,\dots,d} L_j$, the latter condition can be verified in particular if $\|g\|_q < \infty$, $1/q = 1 - 1/r^*$. All these facts explain why, under our general assumption $f \in \mathbb{F}_{g,\infty}(R, Q)$, we study the estimation of unbounded functions only in the case $\alpha = 1$.

4.3.2. *Unbounded case, $\alpha = 1$.* The problem we address now is the adaptive estimation over the collection of functional classes $\{\mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L}) \cap \mathbb{F}_{g,\infty}(R, Q)\}_{\vec{\beta}, \vec{r}, \vec{L}, R, Q}$. Since, if $\|g\|_\infty < \infty$, then $\mathbb{F}_{g,\infty}(R, Q) = \mathbb{F}_g(R)$ for any $Q \geq R \|g\|_\infty$, there is no consistent estimator if either $p = 1$ or $\varkappa_\alpha(p) \leq 0$, $\tau(p) \leq 0$, $\max_{j=1,\dots,d} r_j \leq p$ in view of Theorem 1 in Lepski and Willer (2017). For this reason, later on we will only consider the parameters $\vec{\beta}, \vec{r}$ belonging to the set $\mathcal{P}_{p,\vec{\mu}}$ defined below:

$$\mathcal{P}_{p,\vec{\mu}} = (0, \infty)^d \times [1, \infty]^d \setminus \left\{ \vec{\beta}, \vec{r} : \varkappa_\alpha(p) \leq 0, \tau(p) \leq 0, \max_{j=1,\dots,d} r_j \leq p \right\}.$$

Set $z(\alpha) = \omega(\alpha)(2 + 1/\beta(\alpha))\beta(0)\tau(\infty) + 1$, $p^* = [\max_{l=1,\dots,d} r_l] \vee p$ and let

$$\varrho(\alpha) = \begin{cases} \frac{1 - 1/p}{1 - 1/\omega(\alpha) + 1/\beta(\alpha)}, & \varkappa_\alpha(p) > p\omega(\alpha); \\ \frac{\beta(\alpha)}{2\beta(\alpha) + 1}, & 0 < \varkappa_\alpha(p) \leq p\omega(\alpha); \\ \frac{\tau(p)\omega(\alpha)\beta(0)}{z(\alpha)}, & \varkappa_\alpha(p) \leq 0, \tau(p^*) > 0; \\ \frac{\omega(\alpha)(1 - p^*/p)}{\varkappa_\alpha(p^*)}, & \varkappa_\alpha(p) \leq 0, \tau(p^*) \leq 0. \end{cases}$$

We will assume $0/0 = 0$, which implies in particular $\frac{1-p^*/p}{\varkappa_\alpha(p^*)} = 0$ if $p^* = p$ and $\varkappa_\alpha(p) = 0$. Note also that $\varkappa_\alpha(p^*)/p^* = -1$ if $p^* = \infty$.

THEOREM 4. *Let $\ell \in \mathbb{N}^*$ and $g \in \mathbb{L}_1(\mathbb{R}^d)$, satisfying Assumption 1 be fixed and let K satisfy Assumptions 2 and 4:*

(1) *Then for any $p > [\min_{j=1,\dots,\mu_j}]^{-1}$, $R, Q > 0, 0 < L_0 \leq L_\infty < \infty, (\vec{\beta}, \vec{r}) \in \mathcal{P}_{p,\vec{\mu}} \cap \{(0, \ell]^d \times (1, \infty)^d\}$ and $\vec{L} \in [L_0, L_\infty]^d$ there exists $C < \infty$, independent of \vec{L} , such that*

$$\limsup_{n \rightarrow \infty} \sup_{f \in \mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L}) \cap \mathbb{F}_{g,\infty}(R, Q)} \mathfrak{b}_n(\mathcal{H}^d)^{-1} \delta_n^{-\varrho(1)} \mathcal{R}_p^{(n)}[\widehat{f}_{\mathbf{h}, \mathcal{H}^d}; f] \leq C.$$

(2) *For any $p > [\min_{j=1,\dots,\mu_j}]^{-1}$, $R, Q > 0, 0 < L_0 \leq L_\infty < \infty, (\boldsymbol{\beta}, \mathbf{r}) \in \mathcal{P}_{p,\vec{\mu}} \cap \{(0, \ell] \times [1, \infty)\}$ and $\mathbf{L} \in [L_0, L_\infty]$ there exists $C < \infty$, independent of \mathbf{L} , such that*

$$\limsup_{n \rightarrow \infty} \sup_{f \in \mathbb{N}_{\mathbf{r},d}(\boldsymbol{\beta}, \mathbf{L}) \cap \mathbb{F}_{g,\infty}(R, Q)} \mathfrak{b}_n(\mathcal{H}_{\text{isotr}}^d)^{-1} \delta_n^{-\varrho(1)} \mathcal{R}_p^{(n)}[\widehat{f}_{\mathbf{h}, \mathcal{H}_{\text{isotr}}^d}; f] \leq C.$$

1⁰. Note that $\|g\|_1 < \infty, \|g\|_\infty < \infty$ implies that $\|g\|_2 < \infty$ and, therefore, the Parseval identity together with Assumption 1 allows us to assert that

$$\|g\|_\infty < \infty \quad \Rightarrow \quad \mu_j > 1/2 \quad \forall j = 1, \dots, d.$$

Hence, the condition $p > [\min_{j=1,\dots,\mu_j}]^{-1}$ is automatically checked if $p \geq 2$ and $\|g\|_\infty < \infty$. Also, it is worth noting that considering the adaptation over the collection of isotropic classes, we do not require that the coordinates of $\vec{\mu}$ would be the same. The latter is true for the second assertion of Theorem 3 as well. At last, analyzing the proof of the theorem, we can assert that the second assertion remains true under the slightly weaker assumption $p > d(\mu_1 + \dots + \mu_d)^{-1}$.

2⁰. The assertion of Theorem 4 has no analogue in the existing literature except for the results obtained in Comte and Lacour (2013) and Rebelles (2016). Comte and Lacour (2013) deals with the particular case $p = 2, \vec{r} = (2, \dots, 2)$

while Rebelles (2016) studied the case $\vec{r} = (p, \dots, p)$, $p \in (1, \infty)$. It is easy to check that in both papers whatever the value of $\vec{\beta}$ and $\vec{\mu}$, the corresponding set of parameters belongs to the *dense* zone. Note also that the estimation procedures used in Comte and Lacour (2013) as well as in Rebelles (2016), if $p \geq 2$ (both based on a global version of the Goldenshluger–Lepski method) are optimally adaptive. They attain the asymptotic of the minimax risks corresponding to the dense zone found in Theorem 1 in Lepski and Willer (2017), while our method is only nearly optimally adaptive. However, it is well known that the global selection from the family of standard kernel estimators leads to correct results only if $\vec{r} = (p, \dots, p)$ when the \mathbb{L}_p -risk is considered; see, for instance, Goldenshluger and Lepski (2011, 2014). On the other hand, estimation procedures based on a local selection scheme, which can be applied to the estimation of functions belonging to much more general functional classes, often do not lead to an optimally adaptive method. Fortunately, the loss of accuracy inherent to local procedures is logarithmic w.r.t. the number of observations.

³⁰. Together with Theorems 1 and 2 in Lepski and Willer (2017), Theorems 3 and 4 provide the full classification of the asymptotics of the minimax risks over anisotropic/isotropic Nikol’skii classes for the class parameters belonging to the sparse zone and, up to some logarithmic factor, belonging to the tail and dense zones as well as the boundaries. We mean that the results of these theorems are valid for any fixed $\vec{\beta} \in (0, \infty)^d$, $\vec{r} \in (1, \infty]^d$ and $\vec{L} \in (0, \infty)^d$. Indeed, for given $\vec{\beta}$ and \vec{L} one can choose $L_0 = \min_{j=1, \dots, d} L_j$, $L_\infty = \max_{j=1, \dots, d} L_j$ and the number ℓ , used in the construction of kernel \mathcal{K}_ℓ , as any integer strictly larger than $\max_{j=1, \dots, d} \beta_j$.

5. Proofs of Theorems 1–2.

5.1. *Proof of Theorem 1.* The main ingredients of the proof of the theorem are given in Proposition 1. Their proofs are postponed to Section 5.1.2. Introduce for any $\vec{h} \in \mathcal{H}^d$:

$$\xi_n(x, \vec{h}) = \frac{1}{n} \sum_{i=1}^n [M(Z_i - x, \vec{h}) - \mathbb{E}_f M(Z_i - x, \vec{h})], \quad x \in \mathbb{R}^d.$$

PROPOSITION 1. *Let Assumptions 1 and 2 be fulfilled. Then for any $n \geq 3$ and any $p > 1$:*

- (i) $\int_{\mathbb{R}^d} \mathbb{E}_f \left\{ \sup_{\vec{h} \in \mathcal{H}^d} [|\xi_n(x, \vec{h})| - U_n(x, \vec{h})]_+^p \right\} \nu_d(dx) \leq C_p n^{-\frac{p}{2}},$
- (ii) $\int_{\mathbb{R}^d} \mathbb{E}_f \left\{ \sup_{\vec{h} \in \mathcal{H}^d} [\widehat{U}_n(x, \vec{h}) - 3U_n(x, \vec{h})]_+^p \right\} \nu_d(dx) \leq C'_p n^{-\frac{p}{2}},$
- (iii) $\int_{\mathbb{R}^d} \mathbb{E}_f \left\{ \sup_{\vec{h} \in \mathcal{H}^d} [U_n(x, \vec{h}) - 4\widehat{U}_n(x, \vec{h})]_+^p \right\} \nu_d(dx) \leq C'_p n^{-\frac{p}{2}}.$

The explicit expression of constant C_p and C'_p can be found in the proof.

5.1.1. *Proof of the theorem.* We start by proving the so-called pointwise oracle inequality for losses.

Pointwise oracle inequality for losses. Let $\vec{h} \in \mathbb{H}$ and $x \in \mathbb{R}^d$ be fixed. We have in view of the triangle inequality

$$(5.1) \quad \begin{aligned} |\widehat{f}_{\vec{h}(x)}^-(x) - f(x)| &\leq |\widehat{f}_{\vec{h}(x) \vee \vec{h}}^-(x) - \widehat{f}_{\vec{h}(x)}^-(x)| + |\widehat{f}_{\vec{h}(x) \vee \vec{h}}^-(x) - \widehat{f}_{\vec{h}}^-(x)| \\ &\quad + |\widehat{f}_{\vec{h}}^-(x) - f(x)|. \end{aligned}$$

1⁰. First, note that obviously $\widehat{f}_{\vec{h}(x) \vee \vec{h}}^-(x) = \widehat{f}_{\vec{h} \vee \vec{h}(x)}^-(x)$ and, therefore,

$$\begin{aligned} |\widehat{f}_{\vec{h}(x) \vee \vec{h}}^-(x) - \widehat{f}_{\vec{h}(x)}^-(x)| &= |\widehat{f}_{\vec{h} \vee \vec{h}(x)}^-(x) - \widehat{f}_{\vec{h}(x)}^-(x)| \\ &\leq \widehat{\mathcal{R}}_{\vec{h}}^-(x) + 4\widehat{U}_n(x, \vec{h}(x) \vee \vec{h}) + 4\widehat{U}_n(x, \vec{h}(x)). \end{aligned}$$

Moreover, by definition, $\widehat{U}_n(x, \vec{\eta}) \leq \widehat{U}_n^*(x, \vec{\eta})$ for any $\vec{\eta} \in \mathcal{H}^d$. Next, for any $\vec{h}, \vec{\eta} \in \mathcal{H}^d$ we have obviously $\widehat{U}_n(x, \vec{h} \vee \vec{\eta}) \leq \widehat{U}_n^*(x, \vec{h}) \wedge \widehat{U}_n^*(x, \vec{\eta})$. Thus,

$$(5.2) \quad |\widehat{f}_{\vec{h}(x) \vee \vec{h}}^-(x) - \widehat{f}_{\vec{h}(x)}^-(x)| \leq \widehat{\mathcal{R}}_{\vec{h}}^-(x) + 8\widehat{U}_n^*(x, \vec{h}(x)).$$

Similarly, we have

$$(5.3) \quad |\widehat{f}_{\vec{h}(x) \vee \vec{h}}^-(x) - \widehat{f}_{\vec{h}}^-(x)| \leq \widehat{\mathcal{R}}_{\vec{h}(x)}^-(x) + 8\widehat{U}_n^*(x, \vec{h}).$$

The definition of $\vec{h}(x)$ implies that for any $\vec{h} \in \mathbb{H}$,

$$\widehat{\mathcal{R}}_{\vec{h}(x)}^-(x) + 8\widehat{U}_n^*(x, \vec{h}(x)) + \widehat{\mathcal{R}}_{\vec{h}}^-(x) + 8\widehat{U}_n^*(x, \vec{h}) \leq 2\widehat{\mathcal{R}}_{\vec{h}}^-(x) + 16\widehat{U}_n^*(x, \vec{h})$$

and we get from (5.1), (5.2) and (5.3) for any $\vec{h} \in \mathbb{H}$,

$$(5.4) \quad |\widehat{f}_{\vec{h}(x)}^-(x) - f(x)| \leq 2\widehat{\mathcal{R}}_{\vec{h}}^-(x) + 16\widehat{U}_n^*(x, \vec{h}) + |\widehat{f}_{\vec{h}}^-(x) - f(x)|.$$

2⁰. We obviously have for any $\vec{h}, \vec{\eta} \in \mathcal{H}^d$,

$$\begin{aligned} |\widehat{f}_{\vec{h} \vee \vec{\eta}}^-(x) - \widehat{f}_{\vec{\eta}}^-(x)| &\leq |\mathbb{E}_f M(Z_1 - x, \vec{h} \vee \vec{\eta}) - \mathbb{E}_f M(Z_1 - x, \vec{\eta})| \\ &\quad + |\xi_n(x, \vec{h} \vee \vec{\eta})| + |\xi_n(x, \vec{\eta})|. \end{aligned}$$

Note that, for any $\vec{h} \in \mathcal{H}^d$,

$$\begin{aligned} \mathbb{E}_f M(Z_1 - x, \vec{h}) &:= \int_{\mathbb{R}^d} M(t - x, \vec{h}) \mathbf{p}(t) \nu_d(dt) \\ &= (1 - \alpha) \int_{\mathbb{R}^d} M(t - x, \vec{h}) f(t) \nu_d(dt) \\ &\quad + \alpha \int_{\mathbb{R}^d} M(t - x, \vec{h}) [f \star g](t) \nu_d(dt), \end{aligned}$$

in view of the assumption (1.1) imposed on the density p . Note that

$$\begin{aligned} & (1 - \alpha) \int_{\mathbb{R}^d} M(t - x, \vec{h}) f(t) \nu_d(dt) + \alpha \int_{\mathbb{R}^d} M(t - x, \vec{h}) [f \star g](t) \nu_d(dt) \\ &= \int_{\mathbb{R}^d} f(z) \left[(1 - \alpha) M(z - x, \vec{h}) + \alpha \int_{\mathbb{R}^d} M(u, \vec{h}) g(u - z + x) \nu_d(du) \right] \nu_d(dz) \end{aligned}$$

and, therefore, in view of the definition of $M(\cdot, \vec{h})$ [cf. (2.1)], we obtain

$$(5.5) \quad \mathbb{E}_f M(Z_1 - x, \vec{h}) = \int_{\mathbb{R}^d} K_{\vec{h}}(z - x) f(z) \nu_d(dz) =: S_{\vec{h}}(x, f) \quad \forall \vec{h} \in \mathcal{H}^d.$$

We deduce from (5.5) that

$$|\mathbb{E}_f M(Z_1 - x, \vec{h} \vee \vec{\eta}) - \mathbb{E}_f M(Z_1 - x, \vec{\eta})| = |S_{\vec{h} \vee \vec{\eta}}(x, f) - S_{\vec{\eta}}(x, f)|$$

and, therefore, for any $\vec{h}, \vec{\eta} \in \mathcal{H}^d$,

$$(5.6) \quad \begin{aligned} |\widehat{f}_{\vec{h} \vee \vec{\eta}}(x) - \widehat{f}_{\vec{\eta}}(x)| &\leq |S_{\vec{h} \vee \vec{\eta}}(x, f) - S_{\vec{\eta}}(x, f)| \\ &\quad + |\xi_n(x, \vec{h} \vee \vec{\eta})| + |\xi_n(x, \vec{\eta})|. \end{aligned}$$

3⁰. Set for any $\vec{h} \in \mathcal{H}^d$ and any $x \in \mathbb{R}^d$

$$\begin{aligned} \nu(x) &= \sup_{\vec{\eta} \in \mathcal{H}^d} [|\xi_n(x, \vec{\eta})| - U_n(x, \vec{\eta})]_+, \\ \varpi_1(x) &= \sup_{\vec{h} \in \mathcal{H}^d} [U_n(x, \vec{h}) - 4\widehat{U}_n(x, \vec{h})]_+, \\ \varpi_2(x) &= \sup_{\vec{h} \in \mathcal{H}^d} [\widehat{U}_n(x, \vec{h}) - 3U_n(x, \vec{h})]_+. \end{aligned}$$

We obtain in view of (5.6) that for any $\vec{h} \in \mathbb{H}$ (since obviously $\vec{h} \vee \vec{\eta} \in \mathcal{H}^d$ for any $\vec{h}, \vec{\eta} \in \mathcal{H}^d$)

$$(5.7) \quad \widehat{\mathcal{R}}_{\vec{h}}(x) \leq B_{\vec{h}}^*(x, f) + 2\nu(x) + 2\varpi_1(x).$$

Using the obvious inequality $(\sup_{\alpha} F_{\alpha} - \sup_{\alpha} G_{\alpha})_+ \leq \sup_{\alpha} (F_{\alpha} - G_{\alpha})_+$ get

$$(5.8) \quad [\widehat{U}_n^*(x, \vec{h}) - 3U_n^*(x, \vec{h})]_+ \leq \sup_{\vec{\eta} \in \mathcal{H}^d} [\widehat{U}_n(x, \vec{\eta}) - 3U_n(x, \vec{\eta})]_+ =: \varpi_2(x).$$

We get from (5.4), (5.7) and (5.8)

$$\begin{aligned} |\widehat{f}_{\vec{h}(x)}(x) - f(x)| &\leq 2B_{\vec{h}}^*(x, f) + 4\nu(x) + 4\varpi_1(x) + 48U_n^*(x, \vec{h}) \\ &\quad + 16\varpi_2(x) + |\widehat{f}_{\vec{h}}(x) - f(x)|. \end{aligned}$$

It remains to note that

$$|\widehat{f}_{\vec{h}}(x) - f(x)| \leq B_{\vec{h}}(x, f) + |\xi_n(x, \vec{h})| \leq B_{\vec{h}}(x, f) + U_n(x, \vec{h}) + \nu(x)$$

and we obtain for any $x \in \mathbb{R}^d$, putting $\mathbf{z}(x) = 5\nu(x) + 4\varpi_1(x) + 16\varpi_2(x)$,

$$|\widehat{f}_{\vec{h}(x)}(x) - f(x)| \leq 2B_{\vec{h}}^*(x, f) + B_{\vec{h}}^-(x, f) + 49U_n^*(x, \vec{h}) + \mathbf{z}(x) \quad \forall \vec{h} \in \mathbb{H}.$$

Noting that the left-hand side of the latter inequality is independent of \vec{h} we obtain for any $x \in \mathbb{R}^d$:

$$(5.9) \quad |\widehat{f}_{\vec{h}(x)}(x) - f(x)| \leq \inf_{\vec{h} \in \mathbb{H}} \{2B_{\vec{h}}^*(x, f) + B_{\vec{h}}^-(x, f) + 49U_n^*(x, \vec{h})\} + \mathbf{z}(x).$$

This is the pointwise oracle inequality for losses.

Application of Proposition 1. Set for any $x \in \mathbb{R}^d$

$$R_n(x) = \inf_{\vec{h} \in \mathbb{H}} \{2B_{\vec{h}}^*(x, f) + B_{\vec{h}}^-(x, f) + 49U_n^*(x, \vec{h})\}.$$

Applying Proposition 1, we get from (5.9) and the triangle inequality

$$\begin{aligned} \mathcal{R}_n^{(p)}[\widehat{f}_{\vec{h}(\cdot)}, f] &\leq \|R_n\|_p + 5 \left[\int_{\mathbb{R}^d} \mathbb{E}_f \{ \nu(x) \}^p \right]^{\frac{1}{p}} + 4 \left[\int_{\mathbb{R}^d} \mathbb{E}_f \{ \varpi_1(x) \}^p \right]^{\frac{1}{p}} \\ &\quad + 16 \left[\int_{\mathbb{R}^d} \mathbb{E}_f \{ \varpi_2(x) \}^p \right]^{\frac{1}{p}} \\ &\leq \|R_n\|_p + C_p n^{-\frac{1}{2}}, \end{aligned}$$

where $C_p = 5(C_p)^{\frac{1}{p}} + 20(C'_p)^{\frac{1}{p}}$. The theorem is proved. \square

5.1.2. *Proof of Proposition 1.* Since the proof of the proposition is quite long and technical, we divide it into several steps.

Preliminaries. We start the proof with the following simple remark. Let $\check{M}(t, \vec{h}), t \in \mathbb{R}^d$, denote the Fourier transform of $M(\cdot, \vec{h})$. Then we obtain in view of the definition of $M(\cdot, \vec{h})$:

$$\check{M}(t, \vec{h}) = \check{K}(t\vec{h})[(1 - \alpha) + \alpha\check{g}(-t)]^{-1}, \quad t \in \mathbb{R}^d.$$

1⁰. Note that Assumptions 1 and 2 guarantee that $\check{M}(\cdot, \vec{h}) \in \mathbb{L}_1(\mathbb{R}^d) \cap \mathbb{L}_2(\mathbb{R}^d)$ for any $\vec{h} \in \mathcal{H}^d$ and, therefore,

$$\|M(\cdot, \vec{h})\|_\infty \leq (2\pi)^{-d} \|\check{M}(\cdot, \vec{h})\|_1, \quad \|M(\cdot, \vec{h})\|_2 = (2\pi)^{-d} \|\check{M}(\cdot, \vec{h})\|_2.$$

Thus, putting $\mathcal{M}_\infty(\vec{h}) = M_\infty \prod_{j=1}^d h_j^{-1} (h_j \wedge 1)^{-\mu_j(\alpha)}$, we obtain $\forall \vec{h} \in \mathcal{H}^d$

$$(5.10) \quad \|M(\cdot, \vec{h})\|_\infty \leq \mathcal{M}_\infty(\vec{h}), \quad \|M(\cdot, \vec{h})\|_2 \leq M_2 \prod_{j=1}^d h_j^{-\frac{1}{2}} (h_j \wedge 1)^{-\mu_j(\alpha)},$$

in view of Assumptions 1 and 2. Here, we have put

$$M_2 = [(2\pi)^{-d} \{\varepsilon^{-1} \|\check{K}\|_2 1_{\alpha \neq 1} + \gamma_0^{-1} \mathbf{k}_2 1_{\alpha=1}\}] \vee 1.$$

Additionally we deduce from (5.10) for any $\vec{h} \in \mathcal{H}^d$

$$(5.11) \quad \|M(\cdot, \vec{h})\|_4^4 \leq M_2^2 M_\infty^2 \prod_{j=1}^d h_j^{-3} (h_j \wedge 1)^{-4\mu_j(\alpha)}.$$

Let $\mathcal{L}(\cdot, \vec{h})$ be either $M(\cdot, \vec{h})$ or $M^2(\cdot, \vec{h})$ and let $\mathcal{L}_\infty(\vec{h})$ denote either $\mathcal{M}_\infty(\vec{h})$ or $\mathcal{M}_\infty^2(\vec{h})$.

We have in view of (5.10), denoting $T(\vec{h}) = \sum_{j=1}^d [1 + \mu_j(\alpha)] |\ln(h_j)|$,

$$(5.12) \quad \mathcal{L}_\infty^{-1}(\vec{h}) \vee \mathcal{L}_\infty(\vec{h}) \leq M_\infty^2 e^{2T(\vec{h})} \quad \forall \vec{h} \in \mathcal{H}^d.$$

Additionally, we get from (5.10) and (5.11)

$$(5.13) \quad \|\mathcal{L}(\cdot, \vec{h})\|_2^2 \leq M_2^2 M_\infty^2 e^{4T(\vec{h})} \quad \forall \vec{h} \in \mathcal{H}^d.$$

Set $\sigma^\mathcal{L}(x, \vec{h}) = \sqrt{\int_{\mathbb{R}^d} \mathcal{L}^2(t - x, \vec{h}) \mathfrak{p}(t) \nu_d(dt)}$ and note that in view of (5.13)

$$(5.14) \quad \int_{\mathbb{R}^d} [\sigma^\mathcal{L}(x, \vec{h})]^2 \nu_d(dx) = \|\mathcal{L}(\cdot, \vec{h})\|_2^2 \leq M_2^2 M_\infty^2 e^{4T(\vec{h})} \quad \forall \vec{h} \in \mathcal{H}^d.$$

Next, we have in view of (5.12)

$$(5.15) \quad \|\sigma^\mathcal{L}(\cdot, \vec{h})\|_\infty \leq \mathcal{L}_\infty(\vec{h}) \leq M_\infty^2 e^{2T(\vec{h})}.$$

2⁰. Define for any $x \in \mathbb{R}^d$ and $\vec{h} \in \mathcal{H}^d$

$$\zeta^\mathcal{L}(x, \vec{h}) = n^{-1} \sum_{i=1}^n [\mathcal{L}(Z_i - x, \vec{h}) - \mathbb{E}\mathcal{L}(Z_i - x, \vec{h})],$$

$$z_n(x, \vec{h}) = 3 \ln(n) + (8p + 22)T(\vec{h}) + 2 |\ln(\{\sigma^\mathcal{L}(x, \vec{h})\} \vee \{n^{-3/2} \mathcal{L}_\infty(\vec{h})\})|,$$

$$V^\mathcal{L}(x, \vec{h}) = \sigma^\mathcal{L}(x, \vec{h}) \sqrt{2n^{-1} z_n(x, \vec{h})} + (4/3)n^{-1} z_n(x, \vec{h}) \mathcal{L}_\infty(\vec{h}),$$

$$U^\mathcal{L}(x, \vec{h}) = \sigma^\mathcal{L}(x, \vec{h}) \sqrt{2n^{-1} \lambda_n(\vec{h})} + (4/3)n^{-1} \lambda_n(\vec{h}) \mathcal{L}_\infty(\vec{h}),$$

where remind $\lambda_n(\vec{h}) = 4 \ln(M_\infty) + 6 \ln(n) + (8p + 26)T(\vec{h})$.

Noting that $\sup_{z \in [a, b]} |\ln z| \leq |\ln a| \vee |\ln b|$ for any $0 < a < b < \infty$ we deduce from (5.15) $z_n(x, \vec{h}) \leq \lambda_n(\vec{h})$ for any $x \in \mathbb{R}^d$ and, therefore,

$$(5.16) \quad V^\mathcal{L}(x, \vec{h}) \leq U^\mathcal{L}(x, \vec{h}) \quad \forall \vec{h} \in \mathcal{H}^d.$$

First step. Let $x \in \mathbb{R}^d$ and $\vec{h} \in \mathcal{H}^d$ be fixed. Put $b = 8p + 22$. We obtain for any $z \geq 1$ and $q \geq 1$ by the integration of the Bernstein inequality

$$\begin{aligned} & \mathbb{E}_f \{ |\zeta^\mathcal{L}(x, \vec{h})| - \sqrt{2n^{-1} z} \sigma^\mathcal{L}(x, \vec{h}) - (4/3)n^{-1} z \mathcal{L}_\infty(\vec{h}) \}_+^q \\ & \leq 2\Gamma(q+1) [\sqrt{2n^{-1} z} \sigma^\mathcal{L}(x, \vec{h}) + (4/3)n^{-1} \mathcal{L}_\infty(\vec{h})]^q \exp\{-z\}, \end{aligned}$$

where Γ is the Gamma function.

1⁰. Choose $z = z_n(x, \vec{h})$. Noting that for any $n \in \mathbb{N}^*$ and $x \in \mathbb{R}^d$

$$\sqrt{2n^{-1}z\sigma^{\mathcal{L}}(x, \vec{h})} + (4/3)n^{-1}\mathcal{L}_{\infty}(\vec{h}) \leq 3\mathcal{L}_{\infty}(\vec{h})n^{-\frac{1}{2}}$$

and taking into account that $\exp\{-|\ln(y)|\} \leq y$ for any $y > 0$, we get

$$\begin{aligned} & \mathbb{E}_f\{|\zeta^{\mathcal{L}}(x, \vec{h})| - V^{\mathcal{L}}(x, \vec{h})\}_+^q \\ (5.17) \quad & \leq 2 \times 3^q \Gamma(q + 1)n^{-\frac{q}{2}-3}\mathcal{L}_{\infty}^q(\vec{h})e^{bT(\vec{h})}(\{\sigma^{\mathcal{L}}(x, \vec{h})\} \vee \{n^{-3/2}\mathcal{L}_{\infty}(\vec{h})\})^2 \\ & \leq C_q^{(1)}n^{-\frac{q}{2}-3}e^{(2q-b)T(\vec{h})}(\{\sigma^{\mathcal{L}}(x, \vec{h})\} \vee \{n^{-3/2}\mathcal{L}_{\infty}(\vec{h})\})^2. \end{aligned}$$

Here, to get the second inequality, we have used (5.12) and put $C_q^{(1)} = 2M_{\infty}^{2q}3^q \times \Gamma(q + 1)$.

Set $\mathcal{X}(\vec{h}) = \{x \in \mathbb{R}^d : \sigma^{\mathcal{L}}(x, \vec{h}) \geq n^{-3/2}\mathcal{L}_{\infty}(\vec{h})\}$, $\bar{\mathcal{X}}(\vec{h}) = \mathbb{R}^d \setminus \mathcal{X}(\vec{h})$ and later on the integration over the empty set is supposed to be zero.

Putting $C_p^{(2)} = C_p^{(1)}M_p^2M_{\infty}^2$, we have in view of (5.16), (5.14) and (5.17) applied with $q = p$ that for any $\vec{h} \in \mathcal{H}^d$

$$(5.18) \quad \int_{\mathcal{X}(\vec{h})} \mathbb{E}_f\{|\zeta^{\mathcal{L}}(x, \vec{h})| - U^{\mathcal{L}}(x, \vec{h})\}_+^p \nu_d(dx) \leq C_p^{(2)}n^{-\frac{p}{2}}e^{(2p+4-b)T(\vec{h})}.$$

2⁰. Introduce the following notation. For any $i = 1, \dots, n$, set

$$\Psi_i(x, \vec{h}) = 1_{\{|\mathcal{L}(Z_i-x, \vec{h}) - \mathbb{E}\mathcal{L}(Z_i-x, \vec{h})| \geq n^{-1}\mathcal{L}_{\infty}(\vec{h})\}}$$

and define the random event $D(x, \vec{h}) = \{\sum_{i=1}^n \Psi_i(x, \vec{h}) \geq 2\}$. As usual, the complementary event will be denoted by $\bar{D}(x, \vec{h})$. Set finally $\pi(x, \vec{h}) = \mathbb{P}_f\{\Psi_1(x, \vec{h}) = 1\}$.

We obviously have $|\zeta^{\mathcal{L}}(x, \vec{h})|1_{\bar{D}(x, \vec{h})} \leq 3n^{-1}\mathcal{L}_{\infty}(\vec{h}) < U^{\mathcal{L}}(\vec{h})$ and, therefore,

$$(5.19) \quad 1_{\bar{D}(x, \vec{h})}\{|\zeta^{\mathcal{L}}(x, \vec{h})| - U^{\mathcal{L}}(x, \vec{h})\}_+^p = 0.$$

Applying the Cauchy–Schwarz inequality, we deduce from (5.19) that

$$\mathbb{E}_f\{|\zeta^{\mathcal{L}}(x, \vec{h})| - U^{\mathcal{L}}(x, \vec{h})\}_+^p \leq \mathbb{E}_f^{\frac{1}{2}}\{|\zeta^{\mathcal{L}}(x, \vec{h})| - U^{\mathcal{L}}(x, \vec{h})\}_+^{2p} \mathbb{P}_f^{\frac{1}{2}}\{D(x, \vec{h})\}.$$

Using (5.17) with $q = 2p$ and (5.12), we obtain for any $x \in \bar{\mathcal{X}}(\vec{h})$

$$(5.20) \quad \mathbb{E}_f\{|\zeta^{\mathcal{L}}(x, \vec{h})| - U^{\mathcal{L}}(x, \vec{h})\}_+^p \leq C_p^{(3)}n^{-\frac{p+6}{2}}e^{\mathbf{c}_p T(\vec{h})}[\mathbb{P}_f\{D(x, \vec{h})\}]^{\frac{1}{2}},$$

where we have put $C_p^{(3)} = [C_{2p}^{(1)}]^{\frac{1}{2}}M_{\infty}^2$ and $\mathbf{c}_p = 2p + 2 - b/2$.

For any $\lambda > 0$, we have in view of the exponential Markov inequality

$$\begin{aligned} \mathbb{P}_f\{D(x, \vec{h})\} &= \mathbb{P}_f\left\{\sum_{i=1}^n \Psi_i(x, \vec{h}) \geq 2\right\} \\ &\leq e^{-2\lambda}[e^\lambda \pi(x, \vec{h}) + 1 - \pi(x, \vec{h})]^n \\ &= e^{-2\lambda}[(e^\lambda - 1)\pi(x, \vec{h}) + 1]^n \\ &\leq \exp\{-2\lambda + n(e^\lambda - 1)\pi(x, \vec{h})\}. \end{aligned}$$

Tchebychev inequality yields $\pi(x, \vec{h}) \leq n^2 \mathcal{L}_\infty^{-2}(\vec{h})[\sigma^\mathcal{L}(x, \vec{h})]^2$ and we get

$$\mathbb{P}_f\{D(x, \vec{h})\} \leq \exp\{-2\lambda + n^3 \mathcal{L}_\infty^{-2}(\vec{h})[\sigma^\mathcal{L}(x, \vec{h})]^2(e^\lambda - 1)\} \quad \forall \vec{h} \in \mathcal{H}^d.$$

Note that the definition of $\vec{\mathcal{X}}(\vec{h})$ implies $n^3 \mathcal{L}_\infty^{-2}(\vec{h})[\sigma^\mathcal{L}(x, \vec{h})]^2 < 1$ for any $x \in \vec{\mathcal{X}}(\vec{h})$. Hence, choosing $\lambda = \ln 2 - 2 \ln \{n^{3/2} \mathcal{L}_\infty^{-1}(\vec{h}) \sigma^\mathcal{L}(x, \vec{h})\}$ we have

$$\mathbb{P}_f\{D(x, \vec{h})\} \leq (e^2/4)n^6 \mathcal{L}_\infty^{-4}(\vec{h})[\sigma^\mathcal{L}(x, \vec{h})]^4 \quad \forall x \in \vec{\mathcal{X}}(\vec{h}).$$

It yields, together with (5.12), (5.14) and (5.20) and for any $\vec{h} \in \mathcal{H}^d$

$$(5.21) \quad \int_{\vec{\mathcal{X}}(\vec{h})} \mathbb{E}_f\{|\zeta^\mathcal{L}(x, \vec{h})| - U^\mathcal{L}(x, \vec{h})\}_+^p v_d(dx) \leq C_p^{(4)} n^{-\frac{p}{2}} e^{(2p+10-b/2)T(\vec{h})},$$

where $C_p^{(4)} = C_p^{(3)}(e/2)M_\infty^6 M_2^2$. Putting $C_p^{(5)} = C_p^{(2)} + C_p^{(4)}$ and noting that $2p + 10 - b/2 < 0$ we obtain from (5.18) and (5.21) for any $\vec{h} \in \mathcal{H}^d$

$$(5.22) \quad \int_{\mathbb{R}^d} \mathbb{E}_f\{|\zeta^\mathcal{L}(x, \vec{h})| - U^\mathcal{L}(x, \vec{h})\}_+^p v_d(dx) \leq C_p^{(5)} n^{-\frac{p}{2}} e^{(2p+10-b/2)T(\vec{h})}.$$

³⁰. Choosing $\mathcal{L} = M$ and $\mathcal{L}_\infty = \mathcal{M}_\infty$, we get from (5.22) and the definition of b for any $\vec{h} \in \mathcal{H}^d$:

$$(5.23) \quad \int_{\mathbb{R}^d} \mathbb{E}_f\{|\xi_n(x, \vec{h})| - U_n(x, \vec{h})\}_+^p v_d(dx) \leq C_p^{(5)} n^{-\frac{p}{2}} e^{-T(\vec{h})}.$$

The first assertion follows from (5.23) with $C_p = C_p^{(5)} \sum_{k \in \mathbb{Z}^d} e^{-\sum_{j=1}^d |k_j|}$.

Second step. Denoting $\chi(x, \vec{h}) = \{|\widehat{\sigma}^2(x, \vec{h}) - \sigma^2(x, \vec{h})| - \mathfrak{U}_n(x, \vec{h})\}_+$,

$$\mathfrak{U}_n(x, \vec{h}) = \sigma^{M^2}(x, \vec{h}) \sqrt{2n^{-1} \lambda_n(\vec{h})} + (4/3)n^{-1} \lambda_n(\vec{h}) \mathcal{M}_\infty^2(\vec{h}),$$

and choosing $\mathcal{L} = M^2$ and $\mathcal{L}_\infty = \mathcal{M}_\infty^2$, we get from (5.22) for any $\vec{h} \in \mathcal{H}^d$,

$$(5.24) \quad \int_{\mathbb{R}^d} \mathbb{E}_f\{\chi^p(x, \vec{h})\} v_d(dx) \leq C_p^{(5)} n^{-\frac{p}{2}} e^{(2p+10-b/2)T(\vec{h})}.$$

Note that $\sigma^{M^2}(x, \vec{h}) \leq \mathcal{M}_\infty(\vec{h}) \sigma(x, \vec{h})$ and, therefore, for any $x \in \mathbb{R}^d$,

$$\mathfrak{U}_n(x, \vec{h}) \leq \mathcal{M}_\infty(\vec{h}) U_n(x, \vec{h}) \quad \forall \vec{h} \in \mathcal{H}^d.$$

This implies

$$2n^{-1}\lambda_n(\vec{h})\widehat{\sigma}^2(x, \vec{h}) \leq 2n^{-1}\lambda_n(\vec{h})\sigma^2(x, \vec{h}) + 2n^{-1}\lambda_n(\vec{h})\mathcal{M}_\infty(\vec{h})U_n(x, \vec{h}) + 2n^{-1}\lambda_n(\vec{h})\mathcal{M}_\infty(\vec{h})\chi^*(x, \vec{h}),$$

where we have denoted $\chi^*(x, \vec{h}) = \mathcal{M}_\infty^{-1}(\vec{h})\chi(x, \vec{h})$. Hence

$$(5.25) \quad \widehat{U}_n(x, \vec{h}) \leq U_n(x, \vec{h}) + \sqrt{2n^{-1}\lambda_n(\vec{h})\mathcal{M}_\infty(\vec{h})[U_n(x, \vec{h}) + \chi^*(x, \vec{h})]}.$$

By the same reason,

$$(5.26) \quad U_n(x, \vec{h}) \leq \widehat{U}_n(x, \vec{h}) + \sqrt{2n^{-1}\lambda_n(\vec{h})\mathcal{M}_\infty(\vec{h})[U_n(x, \vec{h}) + \chi^*(x, \vec{h})]}.$$

Note that the definition of $\widehat{U}_n(x, \vec{h})$ and $U_n(x, \vec{h})$ implies that

$$(5.27) \quad 2n^{-1}\lambda_n(\vec{h})\mathcal{M}_\infty(\vec{h}) \leq (3/2) \min[\widehat{U}_n(x, \vec{h}), U_n(x, \vec{h})].$$

Using the inequality $\sqrt{|ab|} \leq 2^{-1}(|a| + |b|)$, $y > 0$ we get from (5.25), (5.26) and (5.27)

$$\widehat{U}_n(x, \vec{h}) \leq (1 + \sqrt{3/2} + (3/4)y)U_n(x, \vec{h}) + (2y)^{-1}\chi^*(x, \vec{h});$$

$$U_n(x, \vec{h}) \leq (1 + (3/4)y)\widehat{U}_n(x, \vec{h}) + (2y)^{-1}U_n(x, \vec{h}) + (2y)^{-1}\chi^*(x, \vec{h}).$$

Choosing $y = 1/2$ in the first inequality and $y = 1$ in the second we get for any $x \in \mathbb{R}^d$ and $\vec{h} \in \mathcal{H}^d$,

$$(5.28) \quad [\widehat{U}_n(x, \vec{h}) - 3U_n(x, \vec{h})]_+ \leq \chi^*(x, \vec{h});$$

$$(5.29) \quad [U_n(x, \vec{h}) - 4\widehat{U}_n(x, \vec{h})]_+ \leq \chi^*(x, \vec{h}).$$

Remembering that $b = 8p + 22$ we obtain from (5.28), (5.29), (5.24) and (5.12) for any $\vec{h} \in \mathcal{H}^d$, denoting $C'_p = M_\infty^{2p}C_p^{(5)}$,

$$(5.30) \quad \int_{\mathbb{R}^d} \mathbb{E}_f[\widehat{U}_n(x, \vec{h}) - 3U_n(x, \vec{h})]_+^p \nu_d(dx) \leq C'_p n^{-\frac{p}{2}} e^{-T(\vec{h})};$$

$$(5.31) \quad \int_{\mathbb{R}^d} \mathbb{E}_f[U_n(x, \vec{h}) - 4\widehat{U}_n(x, \vec{h})]_+^p \nu_d(dx) \leq C'_p n^{-\frac{p}{2}} e^{-T(\vec{h})}.$$

The second and third assertions follow from (5.30) and (5.31). \square

5.2. Proof of Theorem 2. The proof of the theorem is very long and technical and we break it on two parts, which in its turn are divided on several steps. Introduce the following notation: $c_1 = M_2\sqrt{2D}$, $c_2 = \frac{4M_\infty}{3}$,

$$a = \{196[(c_1\sqrt{c_3}) \vee (c_2c_3)]\}^{-1},$$

where $c_3 = 2 \max\{4 \ln(M_\infty), (8p + 26) \max_{j=1, \dots, d}[1 + \mu_j(\alpha)]\}$.

5.2.1. *Preliminaries.* Recall that for any locally integrable function $\lambda : \mathbb{R}^d \rightarrow \mathbb{R}$ its strong maximal function is defined as

$$(5.32) \quad \mathfrak{M}[\lambda](x) := \sup_H \frac{1}{\nu_d(H)} \int_H \lambda(t) dt, \quad x \in \mathbb{R}^d,$$

where the supremum is taken over all possible rectangles H in \mathbb{R}^d with sides parallel to the coordinate axes, containing point x . It is well known that the strong maximal operator $\lambda \mapsto \mathfrak{M}[\lambda]$ is of the strong (\mathbf{t}, \mathbf{t}) -type for all $1 < \mathbf{t} \leq \infty$, that is, if $\lambda \in \mathbb{L}_{\mathbf{t}}(\mathbb{R}^d)$ then $\mathfrak{M}[\lambda] \in \mathbb{L}_{\mathbf{t}}(\mathbb{R}^d)$ and there exists a constant $C_{\mathbf{t}}$ depending on \mathbf{t} only such that

$$(5.33) \quad \|\mathfrak{M}[\lambda]\|_{\mathbf{t}} \leq C_{\mathbf{t}} \|\lambda\|_{\mathbf{t}}, \quad \mathbf{t} \in (1, \infty].$$

Let $\mathfrak{m}[\lambda]$ be defined by (5.32), where, instead of rectangles, the supremum is taken over all possible cubes H in \mathbb{R}^d with sides parallel to the coordinate axes, containing point x . Then it is known that $\lambda \mapsto \mathfrak{m}[\lambda]$ is of the weak $(1, 1)$ -type, that is, there exists C_1 depending on d only s.t. for any $\lambda \in \mathbb{L}_1(\mathbb{R}^d)$,

$$(5.34) \quad \nu_d\{x : |\mathfrak{m}[\lambda](x)| \geq \mathfrak{z}\} \leq C_1 \mathfrak{z}^{-1} \|\lambda\|_1 \quad \forall \mathfrak{z} > 0.$$

The results presented below deal with the weak property of the strong maximal function. The following inequality can be found in [de Guzmán \(1975\)](#). There exists a constant $C > 0$ depending on d only such that

$$\nu_d\{x : |\mathfrak{M}[\lambda](x)| \geq \mathfrak{z}\} \leq C \int_{\mathbb{R}^d} \frac{|\lambda(x)|}{\mathfrak{z}} \left\{ 1 + \left(\ln_+ \frac{|\lambda(x)|}{\mathfrak{z}} \right)^{d-1} \right\} dx, \quad \mathfrak{z} > 0,$$

where for all $z \in \mathbb{R}$, $\ln_+(z) := \max\{\ln(z), 0\}$.

LEMMA 1. *For any given $d \geq 1, R > 0, Q > 0$ and $\mathbf{q} \in (1, \infty]$ there exists $C(d, \mathbf{q}, R, Q)$ such that for any $\lambda \in \mathbb{B}_{1,d}(R) \cap \mathbb{B}_{\mathbf{q},d}(Q)$,*

$$\nu_d\{x : |\mathfrak{M}[\lambda](x)| \geq \mathfrak{z}\} \leq C(d, \mathbf{q}, R, Q) \mathfrak{z}^{-1} (1 + |\ln(\mathfrak{z})|)^{d-1} \quad \forall \mathfrak{z} > 0$$

The proof of the lemma is an elementary consequence of the aforementioned result and can be omitted.

Recall also the particular case of the Young inequality for weak-type spaces; see [Grafakos \(2008\)](#), Theorem 1.2.13. For any $\mathbf{u} \in (1, \infty]$, there exists $C_{\mathbf{u}} > 0$ such that for any $\lambda_1 \in \mathbb{L}_1(\mathbb{R}^d)$ and $\lambda_2 \in \mathbb{L}_{\mathbf{u},\infty}(\mathbb{R}^d)$ one has

$$(5.35) \quad \|\lambda_1 \star \lambda_2\|_{\mathbf{u},\infty} \leq C_{\mathbf{u}} \|\lambda_1\|_1 \|\lambda_2\|_{\mathbf{u},\infty}.$$

Let \mathfrak{J} denote the set of all the subsets of $\{1, \dots, d\}$ endowed with the empty set \emptyset . For any $J \in \mathfrak{J}$ and $y \in \mathbb{R}^d$, set $y_J = \{y_j, j \in J\} \in \mathbb{R}^{|J|}$ and we will write $y = (y_J, y_{\bar{J}})$, where as usual $\bar{J} = \{1, \dots, d\} \setminus J$.

For any $j = 1, \dots, d$ introduce the $d \times d$ matrix $\mathbf{E}_j = (\mathbf{0}, \dots, \mathbf{e}_j, \dots, \mathbf{0})$ where, recall, $(\mathbf{e}_1, \dots, \mathbf{e}_d)$ denotes the canonical basis of \mathbb{R}^d . Set also $\mathbf{E}[J] = \sum_{j \in J} \mathbf{E}_j$. Later on $\mathbf{E}_\emptyset = \mathbf{E}[\emptyset]$ denotes the matrix with zero entries.

To any $J \in \mathfrak{J}$ and any $\lambda : \mathbb{R}^d \rightarrow \mathbb{R}$, associate the function

$$\lambda_J(y_J, z_{\bar{J}}) = \lambda(z + \mathbf{E}[J](y - z)), \quad y, z \in \mathbb{R}^d$$

with the obvious agreement $\lambda_J \equiv \lambda$ if $J = \{1, \dots, d\}$, which is always the case if $d = 1$. For any $\vec{h} \in \mathcal{H}^d$ and $J \subseteq \{1, \dots, d\}$, set $K_{\vec{h}, J}(u_J) = \prod_{j \in J} h_j^{-1} \mathcal{K}(u_j/h_j)$ and define for any $y \in \mathbb{R}^d$,

$$[K_{\vec{h}} \circ \lambda]_J(y) = \int_{\mathbb{R}^{|\bar{J}|}} K_{\vec{h}, \bar{J}}(u_{\bar{J}} - y_{\bar{J}}) \lambda(y_J, u_{\bar{J}}) \nu_{|\bar{J}|}(du_{\bar{J}}),$$

where $\nu_{|\bar{J}|}$ is the Lebesgue measure on $\mathbb{R}^{|\bar{J}|}$. For any $\vec{h}, \vec{\eta} \in \mathcal{H}^d$, set

$$B_{\vec{h}, \vec{\eta}}(x, f) = |S_{\vec{h} \vee \vec{\eta}}(x, f) - S_{\vec{\eta}}(x, f)|.$$

LEMMA 2. *Let Assumption 3 hold. One can find $k \in \{1, \dots, d\}$ and a collection of indexes $\{j_1 < j_2 < \dots < j_k\} \in \{1, \dots, d\}$ such that for any $x \in \mathbb{R}^d$ and any $f : \mathbb{R}^d \rightarrow \mathbb{R}$:*

$$B_{\vec{h}, \vec{\eta}}(x, f) \leq \sum_{l=1}^k ([|K_{\vec{h} \vee \vec{\eta}}| \circ b_{h_{j_l}, f, j_l}]_{J_l}(x) + [|K_{\vec{\eta}}| \circ b_{h_{j_l}, f, j_l}]_{J_l}(x)),$$

$$B_{\vec{h}}(x, f) \leq \sum_{l=1}^k [|K_{\vec{h}}| \circ b_{h_{j_l}, f, j_l}]_{J_l}(x), \quad J_l = \{j_1, \dots, j_l\}.$$

The proof of the lemma can be found in [Lepski \(2015\)](#), Lemma 2. Also, let us mention the following bound which is a trivial consequence of the Young inequality and the Fubini theorem. If $\lambda \in \mathbb{L}_{\mathbf{t}}(\mathbb{R}^d)$, then for any $\mathbf{t} \in [1, \infty]$

$$(5.36) \quad \sup_{J \in \mathfrak{J}} \|[K_{\vec{h}} \circ \lambda]_J\|_{\mathbf{t}} \leq \|\mathcal{K}\|_1^d \|\lambda\|_{\mathbf{t}} \quad \forall \vec{h} \in \mathcal{H}^d.$$

To any $J \in \mathfrak{J}$ and any locally integrable function $\lambda : \mathbb{R}^d \rightarrow \mathbb{R}_+$, we associate the operator

$$(5.37) \quad \mathfrak{M}_J[\lambda](x) = \sup_{H_{|\bar{J}|}} \frac{1}{\nu_{|\bar{J}|}(H_{|\bar{J}|})} \int_{H_{|\bar{J}|}} \lambda(t + \mathbf{E}[J][x - t]) \nu_{|\bar{J}|}(dt_{\bar{J}}),$$

where the supremum is taken over all hyper-rectangles in $\mathbb{R}^{|\bar{J}|}$ containing $x_{\bar{J}} = (x_j, j \in \bar{J})$ and with sides parallel to the axis.

As we see, $\mathfrak{M}_J[\lambda]$ is the strong maximal operator applied to the function obtained from λ by fixing the coordinates whose indices belong to J . It is obvious that $\mathfrak{M}_{\emptyset}[\lambda] \equiv \mathfrak{M}[\lambda]$ and $\mathfrak{M}_{\{1, \dots, d\}}[\lambda] \equiv \lambda$.

The following result is a direct consequence of (5.33) and of the Fubini theorem. For any $\mathbf{t} \in (1, \infty]$, there exists $\mathbf{C}_{\mathbf{t}}$ such that for any $\lambda \in \mathbb{L}_{\mathbf{t}}(\mathbb{R}^d)$

$$(5.38) \quad \sup_{J \in \mathfrak{J}} \|\mathfrak{M}_J[\lambda]\|_{\mathbf{t}} \leq \mathbf{C}_{\mathbf{t}} \|\lambda\|_{\mathbf{t}}.$$

Note that $\mathbf{C}_{\infty} = 1$.

5.2.2. *Part I.* For any $\vec{h} \in \mathcal{H}^d$ and any $v > 0$, let

$$\mathcal{B}_{\vec{h}}^{\pm}(\cdot, f) = 2B_{\vec{h}}^*(\cdot, f) + B_{\vec{h}}^-(\cdot, f), \quad \mathcal{A}(\vec{h}, f, v) = \{x \in \mathbb{R}^d : \mathcal{B}_{\vec{h}}^-(x, f) \geq 2^{-1}v\}.$$

Introduce for any $v > 0$ and $f \in \mathbb{F}_{g, \mathbf{u}}(R, D)$

$$\begin{aligned} \Lambda(v, f) &= \inf_{\vec{h} \in \mathfrak{H}(v)} [v_d(\mathcal{A}(\vec{h}, f, v)) + v^{-2}F_n^2(\vec{h})], \\ \Lambda(v, f, \mathbf{u}) &= \inf_{z \geq 2} \inf_{\vec{h} \in \mathfrak{H}(v, z)} [v_d(\mathcal{A}(\vec{h}, f, v)) + z^{-\mathbf{u}}], \\ \Lambda_p(v, f, \mathbf{u}) &= \inf_{z \geq 2} \inf_{\vec{h} \in \mathfrak{H}(v, z)} \left[\int_{\mathcal{A}(\vec{h}, f, v)} |\mathcal{B}_{\vec{h}}^-(x, f)|^p v_d(dx) + v^p z^{-\mathbf{u}} \right]. \end{aligned}$$

Note that $\Lambda_p(\infty, f, \mathbf{u}) \equiv 0$. Let K be a compactly supported function satisfying Assumption 2. The goal of Part I is to prove the following bound.

For any $n \geq 3$, $p > 1$, $\mathbf{q} > 1$, $R > 1$, $D > 0$, $0 < \underline{v} \leq \bar{v} \leq \infty$, $\mathbf{u} \in (p/2, \infty]$, $\mathbf{u} \geq \mathbf{q}$ and any $f \in \mathbb{F}_{g, \mathbf{u}}(R, D) \cap \mathbb{B}_{\mathbf{q}, d}(D)$,

$$(5.39) \quad \begin{aligned} \mathcal{R}_n^{(p)}[\widehat{f}_{\mathbf{h}(\cdot)}, f] &\leq C^{(1)} \left[l_{\mathbb{H}}(\underline{v}) + \int_{\underline{v}}^{\bar{v}} v^{p-1} \{\Lambda(v, f) \wedge \Lambda(v, f, \mathbf{u})\} dv \right. \\ &\quad \left. + \Lambda_p(\bar{v}, f, \mathbf{u}) \right]^{\frac{1}{p}} + \mathbf{C}_p n^{-\frac{1}{2}}. \end{aligned}$$

Here, $C^{(1)}$ is a constant independent of f and n . Its explicit expression can be found in the proof of the theorem. We remark also that only this constant depends on \mathbf{q} . Since the risk of our estimator is independent of \underline{v} , $\bar{v} > 0$, we can minimize the right-hand side of (5.39) w.r.t. these parameters.

Auxiliary results. Let us prove several simple facts. First, note that for any $n \geq 3$ for any $\vec{h} \in \mathcal{H}^d$

$$(5.40) \quad \lambda_n(\vec{h}) \leq c_3 \left[\ln(n) + \sum_{j=1}^d |\ln(h_j)| \right].$$

Next, it is easy to see that for any any $n \geq 3$ and $\vec{\eta}, \vec{h} \in (0, \infty)^d : \vec{\eta} \geq \vec{h}$

$$F_n(\vec{\eta}) \leq F_n(\vec{h}) \sqrt{l(V_{\vec{\eta}}/V_{\vec{h}})}, \quad G_n(\vec{\eta}) \leq G_n(\vec{h}) l(V_{\vec{\eta}}/V_{\vec{h}}),$$

where $l(v) = v^{-1}(1 + \ln v)$. Since $\vec{\eta} \geq \vec{h}$ implies $V_{\vec{\eta}} \geq V_{\vec{h}}$ and $l(v) \leq 1$ if $v \geq 1$, we have

$$(5.41) \quad F_n(\vec{\eta}) \leq F_n(\vec{h}), \quad G_n(\vec{\eta}) \leq G_n(\vec{h}) \quad \forall \vec{\eta}, \vec{h} \in (0, \infty)^d : \vec{\eta} \geq \vec{h}.$$

Then by (5.40) and the second inequality in (5.41), we have

$$(5.42) \quad \sup_{\vec{\eta} \in \mathcal{H}^d : \vec{\eta} \geq \vec{h}} \frac{4M_\infty \lambda_n(\vec{\eta})}{3n \prod_{j=1}^d \eta_j (\eta_j \wedge 1)^{\mu_j(\alpha)}} \leq c_2 c_3 G_n(\vec{h}).$$

Now let us establish two bounds for $\|U_n^*(\cdot, \vec{h})\|_\infty$.

1⁰a. Let $\mathbf{u} = \infty$. We have, in view of the second inequality in (5.10) for any $\vec{\eta} \in \mathcal{H}^d$,

$$\sigma(x, \vec{\eta}) \leq \sqrt{D} \|M(\cdot, \vec{\eta})\|_2 \leq M_2 \sqrt{D} \prod_{j=1}^d \eta_j^{-\frac{1}{2}} (\eta_j \wedge 1)^{-\mu_j(\alpha)} \quad \forall x \in \mathbb{R}^d.$$

It yields for any $x \in \mathbb{R}^d$ in view of the first inequality in (5.41)

$$(5.43) \quad \sup_{\vec{\eta} \in \mathcal{H}^d : \vec{\eta} \geq \vec{h}} \sqrt{2n^{-1} \lambda_n(\vec{\eta}) \sigma^2(x, \vec{\eta})} \leq \sup_{\vec{\eta} \in \mathcal{H}^d : \vec{\eta} \geq \vec{h}} c_1 \sqrt{c_3} F_n(\vec{\eta}) \leq c_1 \sqrt{c_3} F_n(\vec{h}).$$

Then gathering (5.42), (5.43) and by definition of a , we have

$$(5.44) \quad \|U_n^*(\cdot, \vec{h})\|_\infty \leq (196a)^{-1} [F_n(\vec{h}) + G_n(\vec{h})].$$

1⁰b. Another bound for $\|U_n^*(\cdot, \vec{h})\|_\infty$ is available regardless of the value of \mathbf{u} . Indeed for any $\vec{\eta} \in \mathcal{H}^d$ in view of the first inequality in (5.10)

$$\sigma(x, \vec{\eta}) \leq \|M(\cdot, \vec{\eta})\|_\infty \leq M_\infty \prod_{j=1}^d \eta_j^{-1} (\eta_j \wedge 1)^{-\mu_j(\alpha)} \quad \forall x \in \mathbb{R}^d.$$

It yields for any $x \in \mathbb{R}^d$ and any $n \geq 3$

$$\begin{aligned} & \sup_{\vec{\eta} \in \mathcal{H}^d : \vec{\eta} \geq \vec{h}} \sqrt{2n^{-1} \lambda_n(\vec{\eta}) \sigma^2(x, \vec{\eta})} \\ & \leq \sup_{\vec{\eta} \in \mathcal{H}^d : \vec{\eta} \geq \vec{h}} \sqrt{2c_3} M_\infty \frac{\sqrt{\ln(n) + \sum_{j=1}^d |\ln(\eta_j)|}}{\sqrt{n} \prod_{j=1}^d \eta_j (\eta_j \wedge 1)^{\mu_j(\alpha)}} \leq \sqrt{\frac{2c_3 n}{\ln(n)}} M_\infty G_n(\vec{h}). \end{aligned}$$

Then gathering with (5.42) again, we have

$$\|U_n^*(\cdot, \vec{h})\|_\infty \leq \sqrt{2c_3} M_\infty \vee (c_2 c_3) \sqrt{n / \ln n} G_n(\vec{h}) \quad \forall \vec{h} \in \mathcal{H}^d.$$

Denoting $\vec{\mathbf{b}}[b] = (b, \dots, b)$, we obtain from the previous bound

$$(5.45) \quad \inf_{\vec{h} \in \mathbb{H}} \|U_n^*(\cdot, \vec{h})\|_\infty \leq \inf_{b \geq 1} \|U_n^*(\cdot, \vec{\mathbf{b}}[b])\|_\infty = 0.$$

2⁰. Let now $\mathbf{u} < \infty$. Let us prove that for any $\mathfrak{z} > 0$, $\mathbf{s} \in \{1, \mathbf{u}\}$ and any $f \in \mathbb{F}_{g, \mathbf{u}}(R, D)$

$$(5.46) \quad \nu_d \left(x \in \mathbb{R}^d : \sup_{\vec{\eta} \in \mathcal{H}^d : \vec{\eta} \geq \vec{h}} \mathcal{U}_n(x, \vec{\eta}, f) \geq \mathfrak{z} \right) \leq c_5 [\tilde{D} \mathfrak{z}^{-2} F_n^2(\vec{h})]^{\mathbf{s}},$$

where we have put $\mathcal{U}_n^2(\cdot, \vec{\eta}, f) = 2n^{-1} \lambda_n(\vec{\eta}) \sigma^2(\cdot, \vec{\eta})$ and $\tilde{D} = 1$ if $\mathbf{s} = 1$ and $\tilde{D} = D$ if $\mathbf{s} = \mathbf{u}$. Indeed, if $\mathbf{s} = 1$, applying the Markov inequality, we obtain in view of the second inequality in (5.10) for any $\vec{\eta} \in \mathcal{H}^d$

$$(5.47) \quad \begin{aligned} \nu_d(x \in \mathbb{R}^d : \mathcal{U}_n(x, \vec{\eta}, f) \geq \mathfrak{z}) &\leq 2(n\mathfrak{z}^2)^{-1} \lambda_n(\vec{\eta}) \int_{\mathbb{R}^d} \sigma^2(x, \vec{\eta}) \nu_d(dx) \\ &= 2(n\mathfrak{z}^2)^{-1} \lambda_n(\vec{\eta}) \|M(\cdot, \vec{\eta})\|_2^2 \\ &\leq 2M_2^2(n\mathfrak{z}^2)^{-1} \frac{\lambda_n(\vec{\eta})}{\prod_{j=1}^d \eta_j (\eta_j \wedge 1)^{2\mu_j(\alpha)}} \\ &\leq c_6 \mathfrak{z}^{-2} F_n^2(\vec{\eta}). \end{aligned}$$

Here, we have put $c_6 = 2M_2^2 c_1^2 c_3$ and to get the last inequality we have used (5.40). To get a similar result if $\mathbf{s} = \mathbf{u}$, we remark that $\sigma^2(\cdot, \vec{\eta}) = M^2(\cdot, \vec{\eta}) \star \mathfrak{p}(\cdot)$ and that $M^2(\cdot, \vec{\eta}) \in \mathbb{L}_1(\mathbb{R}^d)$ in view of the second inequality in (5.10). It remains to note that $f \in \mathbb{F}_{g, \mathbf{u}}(R, D)$ implies $\mathfrak{p} \in \mathbb{B}_{\mathbf{u}, d}^{(\infty)}(D)$ and to apply inequality (5.35).

It yields together with the second inequality in (5.10) for any $\vec{\eta} \in \mathcal{H}^d$,

$$(5.48) \quad \nu_d(x \in \mathbb{R}^d : \mathcal{U}_n(x, \vec{\eta}, f) \geq \mathfrak{z}) \leq [c_6 C_{\mathbf{u}} D \mathfrak{z}^{-2} F_n^2(\vec{\eta})]^{\mathbf{u}}.$$

Denoting $\tilde{C} = 1$ if $\mathbf{s} = 1$ and $\tilde{C} = C_{\mathbf{u}}$ if $\mathbf{s} = \mathbf{u}$, we get from (5.47), (5.48)

$$\nu_d \left(x \in \mathbb{R}^d : \sup_{\vec{\eta} \in \mathcal{H}^d : \vec{\eta} \geq \vec{h}} \mathcal{U}_n(x, \vec{\eta}, f) \geq \mathfrak{z} \right) \leq [c_6 \tilde{C} \tilde{D} \mathfrak{z}^{-2}]^{\mathbf{s}} \sum_{\vec{\eta} \in \mathcal{H}^d : \vec{\eta} \geq \vec{h}} F_n^{2\mathbf{s}}(\vec{\eta}).$$

It remains to note that since $\vec{\eta}, \vec{h} \in \mathcal{H}^d$ and $\vec{\eta} \geq \vec{h}$ we can write $\eta_j = e^{m_j} h_j$ with $m_j \geq 0$ for any $j = 1, \dots, d$. Putting $\mathbf{m} = (m_1, \dots, m_d) \in \mathbb{N}^d$, the latter result yields together with the first inequality in (5.41)

$$\sum_{\vec{\eta} \in \mathcal{H}^d : \vec{\eta} \geq \vec{h}} F_n^{2\mathbf{s}}(\vec{\eta}) \leq F_n^{2\mathbf{s}}(\vec{h}) \sum_{\mathbf{m} \in \mathbb{N}^d} \left(1 + \sum_{j=1}^d m_j \right)^{\mathbf{s}} e^{-\mathbf{s} \sum_{j=1}^d m_j} =: c_7 F_n^{2\mathbf{s}}(\vec{h}).$$

Thus, (5.46) with $c_5 = c_7 [c_6 \tilde{C}]^{\mathbf{s}}$ is established.

3⁰. Let $c_K \geq 1$ be s.t. $\text{supp}(K) \subset [-c_K, c_K]^d$. We have $\forall \vec{h} \in (0, \infty)^d$

$$|S_{\vec{h}}(x, f)| = \left| \int_{\mathbb{R}^d} K_{\vec{h}}(t - x) f(t) \nu_d(dt) \right| \leq (2c_K)^d \|K\|_{\infty}^d \mathfrak{M}[|f|](x).$$

If $\vec{h} = (h, \dots, h)$, $h \in (0, \infty)$, the latter inequality holds with $\mathfrak{m}[|f|]$ instead of $\mathfrak{M}[|f|]$. Thus,

$$(5.49) \quad \sup_{\vec{h} \in \mathbb{H}} |\mathcal{B}_{\vec{h}}(x, f)| \leq 3(2c_K)^d \|K\|_\infty \mathfrak{M}_{\mathbb{H}}[|f|](x) + |f(x)| \quad \forall x \in \mathbb{R}^d,$$

where we have denoted $\mathfrak{M}_{\mathbb{H}} = \mathfrak{M}$ if $\mathbb{H} = \mathcal{H}^d$ and $\mathfrak{M}_{\mathbb{H}} = \mathfrak{m}$ if $\mathbb{H} = \mathcal{H}_{\text{isotr}}^d$. Moreover, we deduce from (5.49) and (5.45) putting $T_{\vec{h}}(x, f) = \mathcal{B}_{\vec{h}}(x, f) + 49U_n^*(\cdot, \vec{h})$ that

$$(5.50) \quad \inf_{\vec{h} \in \mathbb{H}} |T_{\vec{h}}(x, f)| \leq 3(2c_K)^d \|K\|_\infty \mathfrak{M}_{\mathbb{H}}[|f|](x) + |f(x)|.$$

Proof of (5.39). Put $\mathbf{T}(x, f) = \inf_{\vec{h} \in \mathbb{H}} |T_{\vec{h}}(x, f)|$ and introduce $\mathcal{C}_v = \{x \in \mathbb{R}^d : \mathbf{T}(x, f) \geq v\}$, $v > 0$. For any given $\bar{v} > 0$, one has

$$(5.51) \quad \|\mathbf{T}(\cdot, f)\|_p^p \leq p \int_0^{\bar{v}} v^{p-1} \nu_d(\mathcal{C}_v) dv + \int_{\mathcal{C}_{\bar{v}}} |\mathbf{T}(x, f)|^p \nu_d(dx).$$

Note that the second term disappears if one chooses $\bar{v} = \infty$. Denoting $\mathcal{W}_v = \{x \in \mathbb{R}^d : 49U_n^*(x, \vec{h}) \geq 2^{-1}v\}$, we have for any $\vec{h} \in \mathbb{H}$ and $v > 0$

$$(5.52) \quad \nu_d(\mathcal{C}_v) \leq \nu_d(\mathcal{A}(\vec{h}, f, v)) + \nu_d(\mathcal{W}_v(\vec{h}, f));$$

$$(5.53) \quad |\mathbf{T}(x, f)|^p 1_{\mathcal{C}_v}(x) \leq 2^p |\mathcal{B}_{\vec{h}}(x, f)|^p 1_{\mathcal{A}(\vec{h}, f, v)} + 98^p |U_n^*(x, \vec{h})|^p 1_{\mathcal{W}_v}(x);$$

$$(5.54) \quad \nu_d(\mathcal{C}_v) \leq \nu_d(x \in \mathbb{R}^d : 3(2c_K)^d \|K\|_\infty \mathfrak{M}_{\mathbb{H}}[|f|](x) + |f(x)| > v).$$

The last inequality follows from (5.50).

1⁰. Set $\mathcal{U}_n^*(x, \vec{h}, f) = \sup_{\vec{\eta} \in \mathcal{H}^d : \vec{\eta} \geq \vec{h}} \mathcal{U}_n(x, \vec{\eta}, f)$. Noting that $U_n^*(x, \vec{h}) \leq \mathcal{U}_n^*(x, \vec{h}, f) + (196a)^{-1} G_n(\vec{h})$ in view of (5.42), we get for any $\vec{h} \in \mathfrak{H}(v)$

$$(5.55) \quad \mathcal{W}_v \subseteq \{x \in \mathbb{R}^d : 49\mathcal{U}_n^*(x, \vec{h}) \geq 4^{-1}v\} := \widetilde{\mathcal{W}}_v.$$

Applying (5.46) with $\mathbf{s} = 1$, we deduce from (5.52) that

$$\nu_d(\mathcal{C}_v) \leq \nu_d(\mathcal{A}(\vec{h}, f, v)) + 196^2 c_5 v^{-2} F_n^2(\vec{h}) \quad \forall \vec{h} \in \mathfrak{H}(v).$$

Since the left-hand side of the latter inequality is independent of \vec{h} , we get

$$(5.56) \quad \nu_d(\mathcal{C}_v) \leq \max[1, 196^2 c_5] \Lambda(v, f).$$

2⁰. Let us establish the following bounds, where c_9 is given in the paragraph 2⁰b. below. For any $\mathbf{u} \in [1, \infty]$ and $v > 0$,

$$(5.57) \quad \nu_d(\mathcal{C}_v) \leq \max[1, c_5 196^2, c_5 196^{2\mathbf{u}} D^{\mathbf{u}} a^{2\mathbf{u}}] \{\Lambda(v, f) \wedge \Lambda(v, f, \mathbf{u})\}$$

and for any $\mathbf{u} \in (p/2, \infty]$,

$$(5.58) \quad \int_{\mathcal{C}_v} |\mathbf{T}(x, f)|^p \nu_d(dx) \leq \max[2^p, 98^p c_9] \Lambda_p(v, f, \mathbf{u}) \quad \forall v > 0.$$

2⁰a. Let $\mathbf{u} = \infty$. Note that minimum over z in the definition of $\Lambda(\cdot, \cdot, \infty)$ and $\Lambda_p(\cdot, \cdot, \infty)$ is obviously attained for $z = 2$. Also, we remark that $\mathcal{W}_v = \emptyset$ for any $\vec{h} \in \mathfrak{H}(v, 2)$ in view of (5.44). Thus, we deduce from (5.52) and (5.53), since the left-hand sides of both inequalities are independent of \vec{h} ,

$$(5.59) \quad \nu_d(\mathcal{C}_v) \leq \Lambda(v, f, \infty), \quad \int_{\mathcal{C}_v} |\mathbf{T}(x, f)|^p \nu_d(dx) \leq \Lambda_p(v, f, \infty).$$

This inequality and (5.56) ensure that (5.57) and (5.58) hold if $\mathbf{u} = \infty$.

2⁰b. Let $\mathbf{u} < \infty$. Applying (5.46) with $\mathbf{s} = \mathbf{u}$, we obtain in view of (5.55)

$$\nu_d(\mathcal{W}_v) \leq c_5 196^{2\mathbf{u}} D^{\mathbf{u}} v^{-2\mathbf{u}} F_n^{2\mathbf{u}}(\vec{h}) \leq c_5 196^{2\mathbf{u}} D^{\mathbf{u}} a^{2\mathbf{u}} z^{-\mathbf{u}} \quad \forall \vec{h} \in \mathfrak{H}(v, z).$$

It yields together with (5.52)

$$(5.60) \quad \nu_d(\mathcal{C}_v) \leq \max[1, c_5 196^{2\mathbf{u}} D^{\mathbf{u}} a^{2\mathbf{u}}] \Lambda(v, f, \mathbf{u}).$$

This inequality and (5.56) ensure that (5.57) holds if $\mathbf{u} < \infty$.

What is more, we have in view of (5.42) and (5.55) for any $\vec{h} \in \mathfrak{H}(v)$,

$$|U_n^*(x, \vec{h})|^p 1_{\mathcal{W}_v} \leq 2^p |U_n^*(x, \vec{h}, f)|^p 1_{\tilde{\mathcal{W}}_v}.$$

Moreover, applying (5.46) with $\mathbf{s} = \mathbf{u}$, we have for any $y > 0$ and $\vec{h} \in \mathfrak{H}(v, z)$

$$\nu_d(\tilde{\mathcal{W}}_y) \leq c_5 196^{2\mathbf{u}} D^{\mathbf{u}} y^{-2\mathbf{u}} F_n^{2\mathbf{u}}(\vec{h}) \leq c_5 196^{2\mathbf{u}} D^{\mathbf{u}} y^{-2\mathbf{u}} (av)^{2\mathbf{u}} z^{-\mathbf{u}}.$$

Hence, if additionally $\mathbf{u} > p/2$, we have for any $\vec{h} \in \mathfrak{H}(v, z)$

$$\begin{aligned} \int_{\mathcal{W}_v} |U_n^*(x, \vec{h})|^p \nu_d(dx) &\leq 2^p p \int_v^\infty y^{p-1} \nu_d(\tilde{\mathcal{W}}_y) dy \\ &\leq c_5 196^{2\mathbf{u}} D^{\mathbf{u}} a^{2\mathbf{u}} 2^p p v^{2\mathbf{u}} z^{-\mathbf{u}} \int_v^\infty y^{p-1-2\mathbf{u}} dy =: c_9 v^p z^{-\mathbf{u}}. \end{aligned}$$

This yields together with (5.53)

$$(5.61) \quad \int_{\mathcal{C}_v} |\mathbf{T}(x, f)|^p \nu_d(dx) \leq \max[2^p, 98^p c_9] \Lambda_p(v, f, \mathbf{u}).$$

This inequality ensures that (5.58) holds if $\mathbf{u} < \infty$.

3⁰. Recall that $f \in \mathbb{F}_g(R)$ implies that $f \in \mathbb{B}_{1,d}(R)$. Since additionally $f \in \mathbb{B}_{\mathbf{q},d}(D)$, $\mathbf{q} > 1$, Lemma 1 as well as (5.34) is applicable and we obtain in view of (5.54) $\nu_d(\mathcal{C}_v) \leq c_{10} v^{-1} (1 + |\ln v|)^{t(\mathbb{H})}$ for any $v > 0$. It yields for any $\underline{v} > 0$ and $p > 1$

$$(5.62) \quad p \int_0^{\underline{v}} v^{p-1} \nu_d(\mathcal{C}_v) dv \leq c_{11} \underline{v}^{p-1} (1 + |\ln \underline{v}|)^{t(\mathbb{H})}.$$

In the case of $t(\mathbb{H}) = 0$, the last inequality is obvious and if $t(\mathbb{H}) = d - 1$ it follows by integration by parts. The bound (5.39) follows now from (5.51), where the bound (5.62) is used when the integration is made over $[0, \underline{v}]$, the estimate (5.57) for integration over $[\underline{v}, \bar{v}]$ and the bound (5.58) with $v = \bar{v}$.

5.2.3. *Part II.* In the subsequent proof c_1, c_2, \dots , stand for constants depending only on $\vec{s}, \vec{q}, g, \mathcal{K}, d, R, D, \mathbf{u}$ and \mathbf{q} .

1⁰. We start with the following obvious observation. For any $\lambda : \mathbb{R}^d \rightarrow \mathbb{R}_+$, $\vec{u} \in \mathbb{R}^d$ and $J \in \mathfrak{J}$,

$$(5.63) \quad [K_{\vec{u}} \circ \lambda]_J(x) \leq (2c_{\mathcal{K}} \|\mathcal{K}\|_{\infty})^d \mathfrak{M}_J[\lambda](x) \quad \forall x \in \mathbb{R}^d.$$

Putting $C_1 = (2c_{\mathcal{K}} \|\mathcal{K}\|_{\infty})^d$, we get for any $\vec{h}, \vec{\eta} \in \mathcal{H}^d$ in view of (5.63) and assertions of Lemma 2 that

$$B_{\vec{h}, \vec{\eta}}^-(\cdot, f) \leq 2C_1 \sum_{j=1}^d \sup_{J \in \mathfrak{J}} \mathfrak{M}_J[b_{h_j, f, j}](\cdot), \quad B_{\vec{h}}^-(\cdot, f) \leq C_1 \sum_{j=1}^d \sup_{J \in \mathfrak{J}} \mathfrak{M}_J[b_{h_j, f, j}](\cdot).$$

Thus noting that the right-hand side of the first inequality above is independent of $\vec{\eta}$, we obtain

$$(5.64) \quad B_{\vec{h}}^-(x, f) \leq 5C_1 \sum_{j=1}^d \sup_{J \in \mathfrak{J}} \mathfrak{M}_J[b_{h_j, f, j}](x) \quad \forall x \in \mathbb{R}^d, \forall \vec{h} \in \mathcal{H}^d.$$

Applying (5.38) with $\mathbf{t} = \infty$, we have for any $v > 0$ in view of the definition of $J(\vec{h}, v)$,

$$(5.65) \quad \begin{aligned} B_{\vec{h}}^-(x, f) &\leq 5C_1 \left[\sum_{j \in \bar{J}(\vec{h}, v)} \sup_{J \in \mathfrak{J}} \mathfrak{M}_J[b_{h_j, f, j}](x) + \sum_{j \in J(\vec{h}, v)} \sup_{J \in \mathfrak{J}} \|\mathfrak{M}_J[b_{h_j, f, j}]\|_{\infty} \right] \\ &\leq 5C_1 \sum_{j \in \bar{J}(\vec{h}, v)} \sup_{J \in \mathfrak{J}} \mathfrak{M}_J[b_{h_j, f, j}](x) + 5C_1 \sum_{j \in J(\vec{h}, v)} \mathbf{B}_{j, \infty, \mathbb{F}}(h_j) \\ &\leq 5C_1 \sum_{j \in \bar{J}(\vec{h}, v)} \sup_{J \in \mathfrak{J}} \mathfrak{M}_J[b_{h_j, f, j}](x) + 4^{-1}v \quad \forall f \in \mathbb{F}. \end{aligned}$$

We obtain for any $f \in \mathbb{F}$, $v > 0$ and $\vec{s} = (s_1, \dots, s_d) \in (1, \infty)^d$, applying consecutively the Markov inequality and (5.38) with $\mathbf{t} = s_j$,

$$(5.66) \quad \begin{aligned} v_d \{ \mathcal{A}(\vec{h}, f, v) \} &\leq v_d \left(\bigcup_{J \in \mathfrak{J}} \bigcup_{j \in \bar{J}(\vec{h}, v)} \{ x : 5C_1 \mathfrak{M}_J[b_{h_j, f, j}](x) \geq (4d)^{-1}v \} \right) \\ &\leq c_1 \sum_{j \in \bar{J}(\vec{h}, v)} v^{-s_j} \|b_{h_j, f, j}\|_{s_j}^{s_j} \\ &\leq c_1 \sum_{j \in \bar{J}(\vec{h}, v)} v^{-s_j} [\mathbf{B}_{j, s_j, \mathbb{F}}(h_j)]^{s_j}. \end{aligned}$$

Noting that the right-hand side of the latter inequality is independent of f and the left-hand side is independent of \vec{s} , we get for any $v > 0$ and $\vec{s} \in (1, \infty)^d$,

$$(5.67) \quad c_1^{-1} \sup_{f \in \mathbb{F}} \{ \Lambda(v, f) \wedge \Lambda(v, f, \mathbf{u}) \} \leq \Lambda_{\vec{s}}(v, \mathbb{F}, \mathbf{u}) \wedge \Lambda_{\vec{s}}(v, \mathbb{F}).$$

²⁰. Note also that in view of (5.65), we have for any $v > 0$:

$$\begin{aligned}
 & \int_{\mathcal{A}(\vec{h}, f, v)} |\mathcal{B}_{\vec{h}}(x, f)|^p \nu_d(dx) \\
 & \leq c_2 \int_{\mathcal{A}(\vec{h}, f, v)} \left| \sum_{j \in \bar{J}(\vec{h}, v)} \sup_{J \in \bar{\mathfrak{J}}} \mathfrak{M}_J[b_{h_j, f, j}](x) \right|^p \nu_d(dx) \\
 (5.68) \quad & + c_3 v^p \nu_d\{\mathcal{A}(\vec{h}, f, v)\} \\
 & \leq c_4 \left[\sum_{j \in \bar{J}(\vec{h}, v)} \int_{\mathcal{A}(\vec{h}, f, v)} \left| \sup_{J \in \bar{\mathfrak{J}}} \mathfrak{M}_J[b_{h_j, f, j}](x) \right|^p \nu_d(dx) \right. \\
 & \quad \left. + v^p \nu_d\{\mathcal{A}(\vec{h}, f, v)\} \right].
 \end{aligned}$$

For any $v > 0$ and $j = 1, \dots, d$, introduce $\mathcal{A}_j(v) = \mathcal{A}(\vec{h}, f, v) \cap \bar{\mathfrak{A}}_j(v)$, where

$$\bar{\mathfrak{A}}_j(v) = \left\{ x \in \mathbb{R}^d : \sup_{J \in \bar{\mathfrak{J}}} \mathfrak{M}_J[b_{h_j, f, j}](x) \geq (40C_1)^{-1}v \right\}.$$

Noting that in view of (5.65) for any $v > 0$ and any $j \in \bar{J}(\vec{h}, v)$,

$$\begin{aligned}
 \mathcal{A}_j(v) & \subseteq \left\{ x \in \mathbb{R}^d : 5C_1 \sum_{k \in \bar{J}(\vec{h}, v), k \neq j} \sup_{J \in \bar{\mathfrak{J}}} \mathfrak{M}_J[b_{h_j, f, k}](x) \geq v/8 \right\} \\
 & \subseteq \left\{ x \in \mathbb{R}^d : 5C_1 \sum_{k \in \bar{J}(\vec{h}, v)} \sup_{J \in \bar{\mathfrak{J}}} \mathfrak{M}_J[b_{h_j, f, k}](x) \geq v/8 \right\} =: \mathcal{A}^*(\vec{h}, f, v),
 \end{aligned}$$

we deduce from (5.68) that for any $\vec{q} \in [p, \infty)^d$,

$$\begin{aligned}
 \int_{\mathcal{A}(\vec{h}, f, v)} |\mathcal{B}_{\vec{h}}(x, f)|^p \nu_d(dx) & \leq c_4 \sum_{j \in \bar{J}(\vec{h}, v)} \int_{\bar{\mathfrak{A}}_j(v)} \left| \sup_{J \in \bar{\mathfrak{J}}} \mathfrak{M}_J[b_{h_j, f, j}](x) \right|^p \nu_d(dx) \\
 (5.69) \quad & + c_5 v^p [\nu_d\{\mathcal{A}^*(\vec{h}, f, v)\} + \nu_d\{\mathcal{A}(\vec{h}, f, v)\}] \\
 & \leq c_6 \sum_{j \in \bar{J}(\vec{h}, v)} v^{p-q_j} \left\| \sup_{J \in \bar{\mathfrak{J}}} \mathfrak{M}_J[b_{h_j, f, j}](x) \right\|_{q_j}^{q_j} \\
 & + c_5 v^p [\nu_d\{\mathcal{A}^*(\vec{h}, f, v)\} + \nu_d\{\mathcal{A}(\vec{h}, f, v)\}].
 \end{aligned}$$

It remains to note that similarly (5.66) for any $\vec{s} \in (1, \infty)^d$

$$\nu_d\{\mathcal{A}^*(\vec{h}, f, v)\} \leq c_7 \sum_{j \in \bar{J}} (\vec{h}, v) v^{-s_j} \|b_{h_j, f, j}\|_{s_j}^{s_j}$$

and to apply (5.38) with $\mathbf{t} = q_j$ to the each term in the sum appeared in (5.69). All of this together with (5.66), applied with $\vec{s} = \vec{q}$ yields for any $v > 0$

and $\vec{q} \in [p, \infty)^d$

$$\int_{\mathcal{A}(\vec{h}, f, v)} |\mathcal{B}_{\vec{h}}(x, f)|^p v_d(dx) \leq c_9 \sum_{j \in \vec{J}(\vec{h}, v)} v^{p-q_j} [\mathbf{B}_{j, q_j, \mathbb{F}}(h_j)]^{q_j}.$$

Noting that the right-hand side of the latter inequality is independent of f and the left-hand side is independent of \vec{q} , the we get

$$(5.70) \quad \sup_{f \in \mathbb{F}} \Lambda_p(v, f, \mathbf{u}) \leq c_9 v^p \Lambda_{\vec{q}}(v, \mathbb{F}, \mathbf{u}) \quad \forall v > 0, \vec{q} \in [p, \infty)^d.$$

The first assertion follows from (5.67), (5.70) and (5.39).

3⁰. Remark that in view of (5.49) and (5.33) $f \in \mathbb{B}_{\mathbf{q}, d}(D)$ implies

$$(5.71) \quad \|\mathcal{B}_{\vec{h}}(\cdot, f)\|_{\mathbf{q}} \leq [3(2c_{\mathcal{K}})^d \|\mathcal{K}\|_{\infty}^d C_{\mathbf{q}} + 1] D \quad \forall \vec{h} \in (0, \infty)^d,$$

where $C_{\mathbf{q}}$ is the constant which appeared in (5.33). Hence for any $v > 0$ and $\mathbf{q} \in [p, \infty)$,

$$(5.72) \quad \int_{\mathcal{A}(\vec{h}, f, v)} |\mathcal{B}_{\vec{h}}(x, f)|^p v_d(dx) \leq 2^{\mathbf{q}-p} v^{p-\mathbf{q}} \|\mathcal{B}_{\vec{h}}(\cdot, f)\|_{\mathbf{q}}^{\mathbf{q}} \leq c_{10} v^{p-\mathbf{q}}.$$

Remind that $\mathfrak{H}(v) \neq \emptyset$, $\mathfrak{H}(v, z) \neq \emptyset$ whatever $v > 0$ and $z \geq 2$; see Remark 1. Hence, in view of (5.72) for any f ,

$$\Lambda_p(v, f, \mathbf{u}) \leq \inf_{z \geq 2} [c_{10} v^{p-\mathbf{q}} + z^{-\mathbf{u}}] = c_{10} v^{p-\mathbf{q}}.$$

It remains to note that the right-hand side of the obtained inequality is independent of f and the second assertion of the theorem follows from this inequality, (5.67) and (5.39).

4⁰. Since $C_{\infty} = 1$, we obtain in view of (5.71) for all $f \in \mathbb{B}_{\infty, d}(D)$:

$$\|\mathcal{B}_{\vec{h}}(\cdot, f)\|_{\infty} \leq [3(2c_{\mathcal{K}})^d \|\mathcal{K}\|_{\infty}^d + 1] D \quad \forall \vec{h} \in (0, \infty)^d.$$

It yields for any $\vec{s} \in (1, \infty)$ in view of (5.66) if $\mathbf{q} = \infty$,

$$\int_{\mathcal{A}(\vec{h}, f, v)} |\mathcal{B}_{\vec{h}}(x, f)|^p v_d(dx) \leq c_{12} \sum_{j \in \vec{J}(\vec{h}, v)} v^{-s_j} [\mathbf{B}_{j, s, \mathbb{F}}(h_j)]^{s_j}.$$

Since the left-hand side of the obtained inequality is independent of f and the left-hand side is independent of \vec{s} we conclude that

$$(5.73) \quad \sup_{f \in \mathbb{F}} \Lambda_p(v, f, \mathbf{u}) \leq c_{12} \Lambda_{\vec{s}}(v, \mathbb{F}, \mathbf{u}) \quad \forall v > 0, \vec{s} \in (1, \infty)^d.$$

The third assertion of the theorem follows now from (5.67), (5.73) and (5.39).

5⁰. We have seen (Corollary 1), that $B_{\vec{h}}^*(\cdot, f) \leq 2 \sup_{\eta \in \mathcal{H}: \eta \leq h} B_{\vec{\eta}}(\cdot, f)$ if $\vec{h} = (h, \dots, h) \in \mathcal{H}_{\text{isotr}}^d$. Therefore, by the definition of $\mathcal{B}_{\vec{h}}(\cdot, f)$,

$$(5.74) \quad \mathcal{B}_{\vec{h}}(\cdot, f) \leq 5 \sup_{\eta \in \mathcal{H}: \eta \leq h} B_{\vec{\eta}}(\cdot, f) \leq 5 \sup_{\eta \in \mathcal{H}: \eta \leq h} \sum_{j=1}^d \sup_{J \in \vec{3}} [|\mathcal{K}_{\vec{\eta}}| \circ b_{\eta, f, j}^*]_J(x),$$

where, remind $\vec{\eta} = (\eta, \dots, \eta) \in \mathcal{H}_{\text{isotr}}^d$. We remark that (5.74) is similar to (5.64) but the maximal operator is not involved in this bound. This, in its turn, allows to consider $\vec{s} \in [1, \infty)^d$. Indeed, similar to (5.65) we have for any $v > 0$, applying (5.36) with $\mathbf{t} = \infty$,

$$(5.75) \quad \mathcal{B}_{\vec{h}}^-(x, f) \leq 5 \sup_{\eta \in \mathcal{H}: \eta \leq h} \sum_{j \in \bar{J}(\vec{h}, v)} \sup_{J \in \bar{\mathcal{J}}} [|K_{\vec{\eta}}| \circ b_{\eta, f, j}^*]_J(x) + 4^{-1}v \quad \forall f \in \mathbb{F}.$$

We obtain for any $f \in \mathbb{F}$, $v > 0$ and $\vec{s} = (s_1, \dots, s_d) \in [1, \infty)^d$ applying consecutively the Markov inequality and (5.36) with $\mathbf{t} = s_j$,

$$v_d(\mathcal{A}(\vec{h}, f, v)) \leq c_{14} \sum_{j \in \bar{J}(\vec{h}, v)} v^{-s_j} [\mathbf{B}_{j, s_j, \mathbb{F}}^*(h)]^{s_j}.$$

We note that the obtained inequality coincides with (5.66) if one replaces $\mathbf{B}_{j, s_j, \mathbb{F}}(\cdot)$ by $\mathbf{B}_{j, s_j, \mathbb{F}}^*(\cdot)$. It remains to note that $\mathbf{B}_{j, s_j, \mathbb{F}}(\cdot) \leq \mathbf{B}_{j, s_j, \mathbb{F}}^*(\cdot)$. Indeed,

$$b_{\mathbf{v}, f, j}(x) = \lim_{k \rightarrow \infty} \sup_{h \in \mathcal{H}: e^{-k} \leq h \leq \mathbf{v}} b_{h, f, j}^*(x).$$

Therefore, by the monotone convergence theorem and the triangle inequality for any $s \in [1, \infty)$,

$$\begin{aligned} \mathbf{B}_{j, s, \mathbb{F}}(\mathbf{h}) &:= \sup_{f \in \mathbb{F}} \|b_{\mathbf{v}, f, j}\|_s = \sup_{f \in \mathbb{F}} \lim_{k \rightarrow \infty} \left\| \sup_{h \in \mathcal{H}: e^{-k} \leq h \leq \mathbf{h}} b_{h, f, j}^* \right\|_s \\ &\leq \sup_{f \in \mathbb{F}} \lim_{k \rightarrow \infty} \sum_{h \in \mathcal{H}: e^{-k} \leq h \leq \mathbf{h}} \|b_{h, f, j}^*\|_s = \sup_{f \in \mathbb{F}} \sum_{h \in \mathcal{H}: h \leq \mathbf{h}} \|b_{h, f, j}^*\|_s =: \mathbf{B}_{j, s, \mathbb{F}}^*(\mathbf{v}). \end{aligned}$$

The fourth statement of the theorem follows now from (5.67), (5.70), (5.72) and (5.39). \square

6. Proof of Theorems 3 and 4. In the subsequent proof $\mathbf{c}, \mathbf{c}_1, C, C_1, \dots$, stand for constants that can depend on $g, L_0, L_\infty, Q, R, \vec{\beta}, \vec{r}, d$ and p , but are independent of \vec{L} and n . These constants can be different on different appearances. The proofs are based on the application of Theorem 3 and on some auxiliary assertions presented below.

The bandwidth's construction presented below as well as auxiliary statements from the next section will be exploited not only for proving Theorems 3 and 4, but also in the consideration forming some future Part II of this work. By this reason, we formulate them in a bit more general form than what is needed for our current purposes.

6.1. *Special set of bandwidths.* Let $\mathcal{J}_\infty = \{j = 1, \dots, d : r_j = \infty\}$ and put $p_\pm = [\sup_{j \in \bar{\mathcal{J}}_\infty} r_j] \vee p$, where $\bar{\mathcal{J}}_\infty$ is complimentary to \mathcal{J}_∞ . Introduce

$$(6.1) \quad q_j = \begin{cases} p_\pm, & j \in \bar{\mathcal{J}}_\infty, \\ \infty, & j \in \mathcal{J}_\infty, \end{cases} \quad \gamma_j = \begin{cases} \frac{\beta_j \tau(p_\pm)}{\tau(r_j)}, & j \in \bar{\mathcal{J}}_\infty, \\ \beta_j, & j \in \mathcal{J}_\infty. \end{cases}$$

Let $\mathbf{c} = (20d)^{-1}[\max(2c_{\mathcal{K}_\ell} \|\mathcal{K}_\ell\|_\infty, \|\mathcal{K}_\ell\|_1)]^{-d}$ and let $L > 0$ be any number satisfying

$$(6.2) \quad L \leq 1 \wedge (C_1^{-1} \mathbf{c}) \wedge L_0.$$

Set for any $r, \mathbf{s} \in [1, \infty]$

$$\varkappa_\alpha(r, \mathbf{s}) = \frac{\mathbf{s}\omega(\alpha)(2 + 1/\beta(\alpha))}{(\mathbf{s} + \omega(\alpha))} - r, \quad \alpha \in [0, 1].$$

Recall that $\delta_n = L(\alpha)n^{-1} \ln n$ and introduce for any $v > 0, \mathbf{s} \in [1, \infty]$,

$$(6.3) \quad \tilde{\eta}_j(v, \mathbf{s}) = (LL_j^{-1})^{\frac{1}{\beta_j}} \{ \mathbf{a}^{-2} \delta_n \}^{\frac{\mathbf{s}\omega(\alpha)}{(\mathbf{s} + \omega(\alpha))\beta_j r_j} v^{\frac{1}{\beta_j} - \frac{\mathbf{s}\omega(\alpha)(2+1/\beta(\alpha))}{(\mathbf{s} + \omega(\alpha))\beta_j r_j}}};$$

$$(6.4) \quad \hat{\eta}_j(v, \mathbf{s}) = (LL_j^{-1})^{\frac{1}{\gamma_j}} \{ \mathbf{a}^{-2} \delta_n \}^{\frac{\mathbf{s}v(\alpha)}{(\mathbf{s} + v(\alpha))\gamma_j q_j} v^{\frac{1}{\gamma_j} - \frac{\mathbf{s}v(\alpha)(2+1/\gamma(\alpha))}{(\mathbf{s} + v(\alpha))\gamma_j q_j}}},$$

where $\frac{1}{\gamma(\alpha)} := \sum_{j=1}^d \frac{2\mu_j(\alpha)+1}{\gamma_j}$, $\frac{1}{v(\alpha)} := \sum_{j=1}^d \frac{2\mu_j(\alpha)+1}{\gamma_j q_j}$.

The constant $\mathbf{a} > 0$ will be chosen differently in accordance with some special relationships between the parameters $\vec{\beta}, \vec{r}, \vec{\mu}, \alpha$ and p .

Determine $\mathbf{h}_j(\cdot, \mathbf{s})$ and $\mathfrak{h}_j(\cdot, \mathbf{s}), j = 1, \dots, d$, from the relations

$$(6.5) \quad \mathbf{h}_j(v, \mathbf{s}) = \max\{h \in \mathcal{H} : h \leq \tilde{\eta}_j(v, \mathbf{s})\}, \quad v > 0;$$

$$(6.6) \quad \mathfrak{h}_j(v, \mathbf{s}) = \max\{h \in \mathcal{H} : h \leq \hat{\eta}_j(v, \mathbf{s})\}, \quad v > 0,$$

and set $\vec{\mathbf{h}}(\cdot, \mathbf{s}) = (\mathbf{h}_1(\cdot, \mathbf{s}), \dots, \mathbf{h}_d(\cdot, \mathbf{s}))$ and $\vec{\mathfrak{h}}(\cdot, \mathbf{s}) = (\mathfrak{h}_1(\cdot, \mathbf{s}), \dots, \mathfrak{h}_d(\cdot, \mathbf{s}))$.

6.2. *Auxiliary statements.* All the results formulated below are proved in Section 7. Let

$$\mathfrak{z}(v) = 2(\mathbf{a}^{-2} \delta_n)^{-\frac{\omega(\alpha)}{\omega(\alpha)+\mathbf{u}} v^{\frac{\omega(\alpha)(2+1/\beta(\alpha))}{\omega(\alpha)+\mathbf{u}}}}, \quad \mathbf{u} \in [1, \infty],$$

and remark that $\mathfrak{z}(\cdot) \equiv 2$ if $\mathbf{u} = \infty$. Note also that

$$(6.7) \quad \mathfrak{z}(v) \geq 2 \quad \forall v \geq (\mathbf{a}^{-2} \delta_n)^{\frac{1}{2+1/\beta(\alpha)}} =: \mathbf{v}.$$

Introduce the following notation: $\mu(\alpha) = \min_{j=1, \dots, d} \mu_j(\alpha)$,

$$X = \frac{1}{2\beta(\alpha)} - \frac{1}{2\beta(0)} = \sum_{j=1}^d \frac{\mu_j(\alpha)}{\beta_j}, \quad Y = \frac{1}{2\omega(\alpha)} - \frac{1}{2\omega(0)} = \sum_{j=1}^d \frac{\mu_j(\alpha)}{\beta_j r_j}.$$

Recall that $z(\alpha) = \omega(\alpha)(2 + 1/\beta(\alpha))\beta(0)\tau(\infty) + 1$ and define

$$(6.8) \quad \underline{\mathbf{y}} = (\mathbf{a}^{-2} \delta_n)^{\frac{1}{1-1/\omega(\alpha)+1/\beta(\alpha)}}, \quad \mathbf{v} = (\mathbf{a}^{-2} \delta_n)^{\frac{\omega(\alpha)\tau(\infty)\beta(0)}{z(\alpha)+\omega(\alpha)/\mathbf{u}}}.$$

Set $\mathbf{u}^* = [-\tau(\infty)\beta(0)]^{-1}$ if $\tau(\infty) < 0$ and let $\mathbf{u}^* = \infty$ if $\tau(\infty) \geq 0$. Put finally $\mathbf{y} = \mathbf{u}^* \vee p^*$ and $\mathbf{Z}_{\mathbf{y}, \mathbf{u}}(\alpha) = Y - [X + 1]\mathbf{y}^{-1} + 1/\mathbf{u}$.

PROPOSITION 2. Let $\vec{\beta}, \vec{r}, L_0, L_\infty, \vec{\mu}, \alpha$ and p be given. Assume that $\vec{L} \in [L_0, L_\infty]^d$. Then:

(1) There exists $\mathbf{a} > 0$ independent of \vec{L} such that for all n large enough

$$\vec{h}(v, \mathbf{1}) \in \mathfrak{H}(v) \quad \forall v \in [\underline{\mathbf{y}}, 1].$$

(2) There exists $\mathbf{a} > 0$ independent of \vec{L} and $\mathbf{u} > 1$ such that $\vec{h}(\mathbf{v}, \mathbf{u}) \in \mathfrak{H}(\mathbf{v}, \mathfrak{z}(\mathbf{v}))$ for all large n if either $\tau(\infty) \geq 0$ or $\mathbf{Z}_{\mathbf{y}, \mathbf{u}}(\alpha) \geq 0, \tau(p^*) \geq 0$.

Recall that $\mathbf{v} \rightarrow 0, n \rightarrow \infty$, is defined in (6.7) and introduce

$$(6.9) \quad \mathbf{v}_1 = (\mathbf{a}^{-2} \delta_n)^{\frac{1}{1-\mathbf{u}/\omega(0)+1/\beta(0)}}, \quad \mathbf{v}_3 = (\mathbf{a}^{-2} \delta_n)^{-\frac{Y+1/\mathbf{u}}{\pi(\mathbf{u})\sqrt{0}}},$$

where $\pi(\mathbf{u}) = [1/\omega(0) - 1/\mathbf{u}][1 + X] - 1/\beta(0)[Y + 1/\mathbf{u}]$. Define also

$$(6.10) \quad \vec{\mathbf{v}} = \mathbf{v}_1 1_{\{\tau(p^*) > 0\}} + \mathbf{v}_2 1_{\{\tau(p^*) \leq 0\}}, \quad \mathbf{v}_2 = (\mathbf{a}^{-2} \delta_n)^{\frac{\mathbf{u}\omega(1)}{\varkappa_1(p^*, \mathbf{u})(\omega(1)+\mathbf{u})}}.$$

Note that $\mathbf{v}_1 \rightarrow \infty, n \rightarrow \infty$, if $\infty > \mathbf{u} \geq \mathbf{u}^* \vee p^*$ (it will be proved in Proposition 3 below). However, $\mathbf{v}_1 = 1$ if $\mathbf{u} = \infty$. As it is shown in the proof of Proposition 2, formulae (7.12), $\mathbf{v} < \mathbf{v}$ for all n large enough. Also $\mathbf{v}_2 \rightarrow \infty, n \rightarrow \infty$, if $\varkappa_1(p^*, \mathbf{u}) < 0$. At last $\mathbf{v}_3 \rightarrow \infty, n \rightarrow \infty$, since $\omega(0) > \omega(1)$. Moreover, $\mathbf{v}_3 = \infty$ if $\pi(\mathbf{u}) \leq 0$. Introduce finally

$$\mathcal{I}_{\mathbf{u}}(\alpha) = \begin{cases} [\mathbf{v}, 1], & p^* = \infty, \\ [\mathbf{v}, \mathbf{v}_1], & \alpha \neq 1, p^* < \infty; \\ [\mathbf{v}, \vec{\mathbf{v}}], & \alpha = 1, p^* < \infty, \mathbf{Z}_{\mathbf{y}, \mathbf{u}}(\alpha) \geq 0; \\ [\mathbf{v}, \mathbf{v}_3], & \alpha = 1, p^* < \infty, \mathbf{Z}_{\mathbf{y}, \mathbf{u}}(\alpha) < 0. \end{cases}$$

PROPOSITION 3. Let $\vec{\beta}, \vec{r}, L_0, L_\infty, \vec{\mu}, \alpha$ and p be given and let $\vec{L} \in [L_0, L_\infty]^d, \mathbf{u} \in [\mathbf{y}, \infty]$. Then there exists $\mathbf{a} > 0$ independent of \vec{L} and \mathbf{u} such that for all n large enough $\vec{h}(v, \mathbf{u}) \in \mathfrak{H}(v, \mathfrak{z}(v)), v \in \mathcal{I}_{\mathbf{u}}(\alpha)$.

In the current paper, we will use the statements of Proposition 2 and 3 only with $\mathbf{u} = \infty$. In this context, we remark that $\varkappa_\alpha(\cdot) \equiv \varkappa_\alpha(\cdot, \infty)$.

We finish this section with the following observations which will be useful in the sequel.

LEMMA 3. For any $\mathbf{u} \in (1, \infty]$ and $\alpha \in [0, 1]$,

$$(6.11) \quad \mathbf{Z}_{\mathbf{y}, \mathbf{u}}(\alpha) \geq 0, \quad \tau(p^*) \geq 0 \quad \Rightarrow \quad z(\alpha)/\omega(\alpha) - 1 + 2/\mathbf{u} \geq 0;$$

$$(6.12) \quad \mathbf{Z}_{\mathbf{y}, \mathbf{u}}(\alpha) \geq 0, \quad \tau(p^*) \leq 0 \quad \Rightarrow \quad \varkappa_\alpha(p^*, \mathbf{u}) < 0;$$

$$(6.13) \quad \varkappa_\alpha(p^*, \mathbf{u}) \leq 0, \quad \tau(p^*) > 0 \quad \Rightarrow \quad z(\alpha) + \omega(\alpha)/\mathbf{u} > 0;$$

$$(6.14) \quad \tau(\infty) \geq 0 \quad \text{or} \quad \mathbf{Z}_{\mathbf{y}, \mathbf{u}}(\alpha) \geq 0, \quad \tau(p^*) \geq 0 \quad \Rightarrow \quad z(\alpha) + \omega(\alpha)/\mathbf{u} > 0;$$

$$(6.15) \quad \mu(\alpha) + 1/\mathbf{u} - 1/\mathbf{y} \geq 0 \quad \Rightarrow \quad \mathbf{Z}_{\mathbf{y}, \mathbf{u}}(\alpha) \geq 0.$$

The proof of the lemma is put into the [Appendix](#).

Moreover, set $r(\alpha) = \frac{1-1/p}{1-1/\omega(\alpha)+1/\beta(\alpha)} \wedge \frac{\beta(\alpha)}{2\beta(\alpha)+1}$. If $\varkappa_\alpha(p^*) \geq 0$, one has

$$(6.16) \quad \varrho(\alpha) = r(\alpha), \quad \rho(\alpha) = r(\alpha) \wedge [\omega(\alpha)/p].$$

If $\varkappa_\alpha(p^*) < 0$, one has

$$(6.17) \quad \varrho(\alpha) = r(\alpha) \wedge \left[\frac{\tau(p)\omega(\alpha)\beta(0)}{z(\alpha)} 1_{\{\tau(p^*) > 0\}} + \frac{\omega(\alpha)(1-p^*/p)}{\varkappa_\alpha(p^*)} 1_{\{\tau(p^*) \leq 0\}} \right];$$

$$(6.18) \quad \rho(\alpha) = r(\alpha) \wedge \left[\frac{\tau(p)\omega(\alpha)\beta(0)}{z(\alpha)} 1_{\{\tau(\infty) > 0\}} + \frac{\omega(\alpha)}{p} 1_{\{\tau(\infty) \leq 0\}} \right].$$

6.3. *Several bounds.* Let us collect some bounds for several terms appearing in Theorem 2 and used in the proofs of Theorems 3 and 4 simultaneously. The proof of the following lemma is put into the [Appendix](#).

LEMMA 4. For any $v \in [\underline{\mathbf{v}}, \mathbf{v}] \cup \mathcal{I}_\infty(\alpha)$,

$$\mathbf{\Lambda}_{\vec{r}}(v, \mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L}), \infty) \leq \mathbf{C}_1 \delta_n^{\omega(\alpha)} v^{-\omega(\alpha)(2+1/\beta(\alpha))} \quad \forall v \in \mathcal{I}_\infty(\alpha);$$

$$\mathbf{\Lambda}_{\vec{r}}(v, \mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L})) \leq \mathbf{C}_1 \delta_n^{\frac{\omega(\alpha)}{\omega(\alpha)+1}} v^{-\frac{\omega(\alpha)(2+1/\beta(\alpha))}{\omega(\alpha)+1}} \quad \forall v \in [\underline{\mathbf{v}}, \mathbf{v}].$$

If additionally $\tau(p^*) > 0$, the following inequality with \vec{q} defined in (6.1) holds:

$$\mathbf{v}^p \mathbf{\Lambda}_{\vec{q}}(\mathbf{v}, \mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L}), \infty) \leq \mathbf{C}_2 \delta_n^{\frac{\omega(\alpha)\tau(p)\beta(0)}{z(\alpha)}}.$$

¹⁰. From now on, we choose $\underline{\mathbf{v}} = \underline{\mathbf{v}}$. In view of the first and the second bounds from Lemma 4 and the definitions of $\underline{\mathbf{v}}$ and \mathbf{v} , we get

$$(6.19) \quad \begin{aligned} & \int_{\underline{\mathbf{v}}}^{\overline{\mathbf{v}}} v^{p-1} [\mathbf{\Lambda}_{\vec{r}}(v, \mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L}), \infty) \wedge \mathbf{\Lambda}_{\vec{r}}(v, \mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L}))] dv \\ & \leq c_4 \left[\delta_n^{\frac{\omega(\alpha)}{\omega(\alpha)+1}} \underline{\mathbf{v}}^{p-\frac{\omega(\alpha)(2+1/\beta(\alpha))}{\omega(\alpha)+1}} 1_{\{\varkappa_\alpha(p) > p\omega(\alpha)\}} \right. \\ & \quad + \delta_n^{\frac{\omega(\alpha)}{\omega(\alpha)+1}} \mathbf{v}^{p-\frac{\omega(\alpha)(2+1/\beta(\alpha))}{\omega(\alpha)+1}} 1_{\{\varkappa_\alpha(p) < p\omega(\alpha)\}} \\ & \quad + \delta_n^{\omega(\alpha)} \mathbf{v}^{p-\omega(\alpha)(2+1/\beta(\alpha))} 1_{\{\varkappa_\alpha(p) > 0\}} \\ & \quad + \delta_n^{\omega(\alpha)} \overline{\mathbf{v}}^{p-\omega(\alpha)(2+1/\beta(\alpha))} 1_{\{\varkappa_\alpha(p) < 0\}} \\ & \quad + \ln(n) \left(\delta_n^{\frac{\omega(\alpha)}{\omega(\alpha)+1}} 1_{\{\varkappa_\alpha(p) = p\omega(\alpha)\}} \right. \\ & \quad \left. + \delta_n^{\omega(\alpha)} 1_{\{\varkappa_\alpha(p) = 0\}} \right) \\ & =: A_n + c_2 \delta_n^{\omega(\alpha)} \overline{\mathbf{v}}^{p-\omega(\alpha)(2+1/\beta(\alpha))} 1_{\{\varkappa_\alpha(p) < 0\}}. \end{aligned}$$

After elementary computations and taking into account (6.16), we obtain

$$(6.20) \quad A_n \leq c_5 \mathfrak{b}_n^p(\mathbb{H}) \delta_n^{p\rho(\alpha)}, \quad A_n \leq c_5 \mathfrak{b}_n^p(\mathbb{H}) \delta_n^{p\varrho(\alpha)}.$$

These bounds are not surprising because $\varrho(\alpha) = \rho(\alpha)$ if $\varkappa_\alpha(p) \geq 0$.

2⁰. We have $\ell_{\mathbb{H}}(\underline{\mathbf{v}}) \leq c_6 \delta_n^{\frac{p-1}{1-\frac{1}{\omega(\alpha)}+1/\beta(\alpha)}} (\ln n)^{t(\mathbb{H})}$, which yields by (6.16), (6.17) and (6.18):

$$(6.21) \quad \ell_{\mathbb{H}}(\underline{\mathbf{v}}) \leq c_1 \mathfrak{b}_n^p(\mathbb{H}) \delta_n^{p\rho(\alpha)}, \quad \ell_{\mathbb{H}}(\underline{\mathbf{v}}) \leq c_1 \mathfrak{b}_n^p(\mathbb{H}) \delta_n^{p\varrho(\alpha)}.$$

6.4. *Proof of Theorem 3.* Furthermore, $\mathcal{F} = \mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L}) \cap \mathbb{F}_g(R) \cap \mathbb{B}_{\infty,d}(Q)$. Since $\mathbb{F}_g(R) \cap \mathbb{B}_{\infty,d}(Q) \subset \mathbb{F}_{g,\infty}(R, D)$ with $D = Q[1 - \alpha + \alpha \|g\|_1]$, Theorem 2 with $\mathbf{u} = \infty$, $\mathbf{q} = \infty$, $D = Q[1 - \alpha + \alpha \|g\|_1] \vee Q$ is applicable with $\mathbb{F} = \mathcal{F}$.

1⁰. Consider the case $\tau(\infty) \leq 0$. Choose $\bar{\mathbf{v}} = 1$ and remark that the statements of Propositions 2 and 3 hold for any $v \in [\underline{\mathbf{v}}, \bar{\mathbf{v}}]$. Indeed, it suffices to note that $\mathcal{I}_\infty(\alpha) \supseteq [\underline{\mathbf{v}}, \bar{\mathbf{v}}] := [\underline{\mathbf{v}}, 1]$, because $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 > 1$ and $\bar{\mathbf{v}} \geq 1$ if $\tau(\infty) \leq 0$, $\tau(p^*) > 0$, $\mathbf{Z}_{\mathbf{y},\infty}(\alpha) \geq 0$ since in this case $\mathbf{v} \geq 1$ by (6.14). Then we can apply all the bounds obtained above, in particular we get from the first inequality of Lemma 4

$$(6.22) \quad \Lambda_{\vec{r}}(1, \mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L}), \infty) \leq C_1 \lambda_1(1) \leq c_6 \delta_n^{\omega(\alpha)} \leq c_6 \mathfrak{b}_n^p(\mathbb{H}) \delta_n^{p\rho(\alpha)},$$

since $\omega(\alpha) \geq p\rho(\alpha)$ in both considered cases in view of the second equality in (6.16) and of (6.18). Applying the third assertion of Theorem 2, we obtain from (6.19), (6.20), (6.22) and (6.21)

$$\sup_{f \in \mathcal{F}} \mathcal{R}_n^{(p)}[\widehat{f}_{\widehat{\mathbf{h}}(\cdot)}, f] \leq C[(c_1 + c_4 + c_5 + c_6) \mathfrak{b}_n^p(\mathbb{H}) \delta_n^{p\rho(\alpha)}]^{1/p} \leq c_7 \mathfrak{b}_n(\mathbb{H}) \delta_n^{\rho(\alpha)}$$

and the assertion of Theorem 3 follows in both considered cases.

2⁰. Consider the case $\tau(\infty) > 0$.

Choose $\bar{\mathbf{v}} = \mathbf{v}$ and remark that the statements of Propositions 2 and 3 hold for any $v \in [\underline{\mathbf{v}}, \bar{\mathbf{v}}]$. Indeed, $\tau(\infty) > 0$ implies $\mathbf{v} < 1$ in view of (6.14) and, therefore, $[\underline{\mathbf{v}}, \mathbf{v}] \subseteq \mathcal{I}_\infty(\alpha)$. We deduce from (6.19), (6.20), the third bound in Lemma 4, (6.21) and applying the first assertion of Theorem 2 that

$$(6.23) \quad \sup_{f \in \mathcal{F}} \mathcal{R}_n^{(p)}[\widehat{f}_{\widehat{\mathbf{h}}(\cdot)}, f] \leq C[c_8 \delta_n^{\frac{\omega(\alpha)\tau(p)\beta(0)}{z(\alpha)}} + (c_3 + c_6) \mathfrak{b}_n^p(\mathbb{H}) \delta_n^{p\rho(\alpha)}]^{1/p}.$$

This completes the proof of Theorem 3 in view of (6.18).

6.5. *Proof of Theorem 4.* In the following, we assume $p^* < \infty$, since $p^* = \infty$ implies by definition of the anisotropic Nikol'skii class that $\mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L}) \subset \mathbb{B}_{\infty,d}(L_\infty)$. Hence, the results in that case follow from Theorem 3 since $\varrho(\alpha) = \rho(\alpha)$ when $p^* = \infty$.

Moreover, we remark that the imposed condition $p > [\min_{j=1,\dots,\mu_j} \mu_j]^{-1}$ implies $\mathbf{Z}_{\mathbf{y},\infty}(\alpha) \geq 0$ in view of (6.15) proved in Lemma 3. This, first, makes the second

assertion of Proposition 2 applicable in the case $\tau(p^*) > 0$. Next, it allows (recall that $p^* < \infty$ and $\alpha = 1$) to rewrite $\mathcal{I}_\infty(1)$ which appeared in Proposition 3 as $\mathcal{I}_\infty(1) = [\mathbf{v}, \bar{\mathbf{v}}]$.

1⁰. Consider the case $\tau(p^*) > 0$.

Taking into account that $\vec{L} \in [L_0, L_\infty]$ we remark that in view of Nikol'skii (1977) [Theorem 6.9.1, Section 6.9] $\mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L}) \subset \mathbb{B}_{p^*,d}(c_9 L_\infty)$, where c_9 is independent of \vec{L} . Thus, Theorem 2 is applicable with $\mathbf{u} = \infty$, $\mathbf{q} = p^*$ and $D = c_9 L_\infty \vee Q$. Choose $\bar{\mathbf{v}} = \mathbf{v}$ and remark that the statements of Propositions 2 and 3 hold since $\bar{\mathbf{v}} = \mathbf{v}$. The assertion of the theorem is obtained from (6.19), (6.20), the third bound in Lemma 4, (6.21), (6.17) and the first assertion of Theorem 2 by the same computations that led to (6.23).

2⁰. Consider the case $\tau(p^*) \leq 0$. Since $\mathbf{Z}_{y,\infty}(\alpha) \geq 0$, we have $\varkappa_1(p^*) < 0$ in view of (6.12) of Lemma 3. This in its turn implies that $p^* > p$ in this case because we consider only class parameters belonging to $\mathcal{P}_{p,\vec{\mu}}$. Since the definition of the anisotropic Nikol'skii class implies that $\mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L}) \subset \mathbb{B}_{p^*,d}(L_\infty)$, we assert that the second assertion of Theorem 2 is applicable with $\mathbf{u} = \infty$, $\mathbf{q} = p^*$ and $D = L_\infty \vee Q$. Choose $\bar{\mathbf{v}} = \mathbf{v}_2$ and note that $\bar{\mathbf{v}} = \mathbf{v}_2$ in the considered case. Thus, we deduce from (6.19), (6.20), (6.21) and (6.17), denoting $\mathcal{F} = \mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L}) \cap \mathbb{F}_{g,\infty}(R, Q)$,

$$\sup_{f \in \mathcal{F}} \mathcal{R}_n^{(p)}[\widehat{f}_{\mathbf{h}(\cdot)}, f] \leq C [c_9 \delta_n^{\omega(1) - \frac{\omega(1)\varkappa_1(p,\infty)}{\varkappa_1(p^*,\infty)}} + (c_3 + c_6) \mathfrak{b}_n^p(\mathbb{H}) \delta_n^{p\varrho(\alpha)} + \delta_n^{\frac{\omega(1)(p-p^*)}{\varkappa_1(p^*,\infty)}}]^{1/p}$$

and the assertion of the theorem follows in this case. Theorem 4 is proved.

7. Proofs of Propositions 2 and 3. Without further mentioning, we will assume that n is large enough to provide $\alpha^{-2} \delta_n \leq 1$. The proof of the following lemma is in the Appendix.

LEMMA 5. For any $\vec{\beta}, \vec{r}, \vec{\mu}, p \geq 1$ and $\alpha \in [0, 1]$, the following is true:

$$1/\gamma(\alpha) - 1/\beta(\alpha) = [\tau(\infty)\beta(0)]^{-1} [1/\omega(\alpha) - 1/\nu(\alpha)].$$

7.1. Proof of Proposition 2. We start the proof with several remarks which will be useful in the sequel. First, obviously there exists $0 < \mathbf{T} := T(\vec{\beta}, \vec{r}, \vec{\mu}, p) < \infty$ independent of \vec{L} such that

$$(7.1) \quad \lim_{n \rightarrow \infty} \sup_{\alpha \in (0, 1]} \sup_{\mathbf{s} \in [1, \infty]} \sup_{v \in [\underline{\mathbf{v}}, 1 \vee \mathbf{v}]} \sup_{j=1}^d \left\{ \frac{|\ln(\mathbf{h}_j(v, \mathbf{1}))| + |\ln(\mathbf{h}_j(v, \mathbf{s}))|}{\ln n} \right\} = \mathbf{T}.$$

Next, for any $\mathbf{s} \in [1, \infty]$ and any $v > 0$,

$$(7.2) \quad \frac{\ln n}{n} \prod_{j=1}^d (\tilde{\eta}_j(v, \mathbf{s}))^{-1-2\mu_j(\alpha)} = 2L^{-\frac{1}{\beta(\alpha)}} \alpha^2 v^2 \mathfrak{z}^{-1}(v).$$

(1) Let us proceed to the proof of the first assertion. Note that for all $n \geq 3$,

$$(7.3) \quad \vec{h}(v, \mathbf{1}) \in (0, 1]^d \quad \forall v \in [\underline{y}, 1].$$

Indeed for any $v > 0$ we have since $\mathbf{L} \leq L_0$,

$$(7.4) \quad \tilde{\eta}_j^{\beta_j r_j}(v, \mathbf{1}) \leq (\mathfrak{a}^{-2} \delta_n)^{\frac{\omega(\alpha)}{1+\omega(\alpha)}} v^{r_j - \frac{\omega(\alpha)(2+1/\beta(\alpha))}{1+\omega(\alpha)}}, \quad j \in \bar{\mathcal{J}}_\infty.$$

Therefore, for any $v \in [\underline{y}, 1]$ one has, in view of the definition of \underline{y} ,

$$\tilde{\eta}_j^{\beta_j r_j}(v, \mathbf{1}) \leq (\mathfrak{a}^{-2} \delta_n)^{\frac{\omega(\alpha)}{1+\omega(\alpha)}} \underline{y}^{1 - \frac{\omega(\alpha)(2+1/\beta(\alpha))}{1+\omega(\alpha)}} = 1, \quad j \in \bar{\mathcal{J}}_\infty.$$

Note that for any $j \in \mathcal{J}_\infty$,

$$\tilde{\eta}_j(v, \mathbf{1}) = (\mathbf{L}L_j^{-1}v)^{\frac{1}{\beta_j}} \leq v^{\frac{1}{\beta_j}} \leq 1 \quad \forall v \leq 1$$

and the proof of (7.3) is completed since $\mathbf{h}_j(\cdot, \mathbf{1}) \leq \tilde{\eta}_j(\cdot, \mathbf{1})$ by construction.

Set $T_0 = [\mathbf{T} + 2]e^{d+2\sum_{j=1}^d \mu_j(\alpha)} \mathbf{L}^{-\frac{1}{\beta(\alpha)}}$ and remark that in view of (7.1), (7.2) and (7.3) for all n large enough and any $v \in [\underline{y}, 1]$,

$$(7.5) \quad \begin{aligned} G_n(\vec{h}(v, \mathbf{1})) &\leq \frac{(\mathbf{T} + 2) \ln n}{n \prod_{j=1}^d (\mathbf{h}_j(v, \mathbf{1}))^{1+\mu_j(\alpha)}} \leq \frac{T_0 \mathbf{L}^{\frac{1}{\beta(\alpha)}} \ln n}{n \prod_{j=1}^d (\tilde{\eta}_j(v, \mathbf{1}))^{1+\mu_j(\alpha)}} \\ &\leq \frac{T_0 \mathbf{L}^{\frac{1}{\beta(\alpha)}} \ln n}{n \prod_{j=1}^d (\tilde{\eta}_j(v, \mathbf{1}))^{1+2\mu_j(\alpha)}} = T_0 \mathfrak{a}^{\frac{2}{1+\omega(\alpha)}} \delta_n^{\frac{\omega(\alpha)}{1+\omega(\alpha)}} v^{\frac{2-\omega(\alpha)/\beta(\alpha)}{1+\omega(\alpha)}}. \end{aligned}$$

Here, we have taken into account that $\mathbf{h}_j(v, \mathbf{s}) \geq e^{-1} \eta_j(v, \mathbf{s})$. Since

$$T_0 \mathfrak{a}^{\frac{2}{1+\omega(\alpha)}} \delta_n^{\frac{\omega(\alpha)}{1+\omega(\alpha)}} v^{\frac{2-\omega(\alpha)/\beta(\alpha)}{1+\omega(\alpha)}} \leq T_0 \mathfrak{a}^2 v \quad \Leftrightarrow \quad v \geq \underline{y}$$

denoting $\mathfrak{a} = \sqrt{a/T_0}$ we assert that

$$G_n(\vec{h}(v, \mathbf{1})) \leq av \quad \forall v \in [\underline{y}, 1].$$

The first assertion is established.

(2) Before proving the second assertion, let us make several remarks.

1⁰. For any $\mathbf{u} \in [1, \infty]$, the following is true:

$$(7.6) \quad \hat{\eta}_j(\mathbf{v}, \mathbf{u}) = (\mathbf{L}L_j^{-1}\mathbf{v})^{\frac{1}{\beta_j}}, \quad j \in \mathcal{J}_\infty;$$

$$(7.7) \quad \hat{\eta}_j(\mathbf{v}, \mathbf{u}) = (\mathbf{L}L_j^{-1}\mathbf{v})^{\frac{1}{\gamma_j}} (\mathfrak{a}^{-2} \delta_n)^{\frac{\omega(\alpha)\tau(p_\pm)\beta(0)}{\gamma_j|\zeta(\alpha)+\omega(\alpha)/\mathbf{u}|}}, \quad j \in \bar{\mathcal{J}}_\infty.$$

The equality (7.6) follows directly from the definition of $\hat{\eta}_j(\mathbf{v}, \mathbf{u})$ since, we remind $\gamma_j = \beta_j$, $q_j = \infty$ if $j \in \mathcal{J}_\infty$. Thus, let us prove the equality (7.7):

$$\hat{\eta}_j^{\gamma_j q_j}(\mathbf{v}, \mathbf{u}) = (\mathbf{L}L_j^{-1})^{p_\pm} (\mathfrak{a}^{-2} \delta_n)^{\frac{\mathbf{u}v(\alpha)}{\mathbf{u}+v(\alpha)}} \mathbf{v}^{p_\pm - \frac{\mathbf{u}v(\alpha)(2+1/\gamma(\alpha))}{\mathbf{u}+v(\alpha)}} \quad \forall j \in \bar{\mathcal{J}}_\infty.$$

Here, we used that $q_j = p_{\pm}$ for any $j \in \bar{\mathcal{J}}_{\infty}$. Using the definition of \mathbf{v} , we get

$$\widehat{\eta}_j^{\gamma_j q_j}(\mathbf{v}, \mathbf{u}) = (\mathbf{L}L_j^{-1})^{p_{\pm}} (\mathfrak{a}^{-2} \delta_n)^{\frac{\mathbf{u}v(\alpha)}{\mathbf{u}+v(\alpha)} + \frac{\omega(\alpha)\tau(\infty)\beta(0)}{z(\alpha)+\omega(\alpha)/\mathbf{u}}} [p_{\pm} - \frac{\mathbf{u}v(\alpha)(2+1/\gamma(\alpha))}{\mathbf{u}+v(\alpha)}]$$

for any $j \in \bar{\mathcal{J}}_{\infty}$. Using the definition of $z(\alpha)$, we obtain

$$\begin{aligned} A &:= \frac{\mathbf{u}v(\alpha)}{\mathbf{u}+v(\alpha)} + \frac{\omega(\alpha)\tau(\infty)\beta(0)}{z(\alpha)+\omega(\alpha)/\mathbf{u}} \left[p_{\pm} - \frac{\mathbf{u}v(\alpha)(2+1/\gamma(\alpha))}{\mathbf{u}+v(\alpha)} \right] \\ &= \frac{\omega(\alpha)\tau(\infty)\beta(0)p_{\pm}}{z(\alpha)+\omega(\alpha)/\mathbf{u}} \\ &\quad + \frac{\mathbf{u}v(\alpha)[1+\omega(\alpha)/\mathbf{u}-\omega(\alpha)\tau(\infty)\beta(0)\{1/\gamma(\alpha)-1/\beta(\alpha)\}]}{(\mathbf{u}+v(\alpha))(z(\alpha)+\omega(\alpha)/\mathbf{u})}. \end{aligned}$$

We obtain applying Lemma 5

$$A = \frac{\omega(\alpha)\tau(\infty)\beta(0)p_{\pm}}{z(\alpha)+\omega(\alpha)/\mathbf{u}} + \frac{\mathbf{u}v(\alpha)\omega(\alpha)[1/s+1/v(\alpha)]}{(\mathbf{u}+v(\alpha))(z(\alpha)+\omega(\alpha)/\mathbf{u})} = \frac{\omega(\alpha)\tau(p_{\pm})\beta(0)p_{\pm}}{z(\alpha)+\omega(\alpha)/\mathbf{u}}.$$

Thus, (7.7) is established.

2⁰. Next, let us prove that

$$(7.8) \quad \vec{h}(\mathbf{v}, \mathbf{u}) \in (0, 1]^d \quad \forall \mathbf{u} \in (1, \infty].$$

If $\mathcal{J}_{\infty} \neq \emptyset$, which is equivalent to $p^* = \infty$, the definition of \mathbf{v} implies that $\mathbf{v} \leq 1$ for all n large enough, since $\tau(p^*) = \tau(\infty) \geq 0$ and in view of (6.14). We deduce from (7.6)

$$h_j(\mathbf{v}, \mathbf{u}) \leq \widehat{\eta}_j(\mathbf{v}, \mathbf{u}) = (\mathbf{L}L_j^{-1}\mathbf{v})^{\frac{1}{\beta_j}} \leq \mathbf{v}^{\frac{1}{\beta_j}} \leq 1 \quad \forall j \in \mathcal{J}_{\infty}$$

and (7.8) is proved for any $j \in \mathcal{J}_{\infty}$.

It remains to note that $\tau(p_{\pm}) \geq \tau(p^*)$ since $p^* \geq p_{\pm}$ and, therefore, if $\tau(p^*) \geq 0$ we have $h_j(\mathbf{v}, \mathbf{u}) \leq \widehat{\eta}_j(\mathbf{v}, \mathbf{u}) \leq 1$, for any $j \in \bar{\mathcal{J}}_{\infty}$ and all n large enough in view of (6.14), (7.7) and since $\mathbf{L}L_j^{-1} \leq 1$. Thus, (7.8) is proved.

3⁰. For any $\mathbf{u} \in (1, \infty]$, one has

$$(7.9) \quad \mathfrak{a}^{-2} \delta_n \prod_{j=1}^d \widehat{\eta}_j^{-1-2\mu_j(\alpha)}(\mathbf{v}, \mathbf{u}) \leq T^{-1}(\alpha) (\mathfrak{a}^{-2} \delta_n)^{1-\frac{\omega(\alpha)\tau(\infty)\beta(0)/\beta(\alpha)+1}{z(\alpha)+\omega(\alpha)/\mathbf{u}}};$$

$$(7.10) \quad \mathfrak{a}^{-2} \delta_n \prod_{j=1}^d \widehat{\eta}_j^{-1}(\mathbf{v}, \mathbf{u}) \leq T^{-1}(0) (\mathfrak{a}^{-2} \delta_n)^{1-\frac{\omega(\alpha)}{z(\alpha)+\omega(\alpha)/\mathbf{u}}},$$

where $T(\alpha) = \inf_{\bar{L} \in [L_0, L_{\infty}]^d} \prod_{j \in \mathcal{J}_{\infty}} (\mathbf{L}L_j^{-1})^{\frac{1+2\mu_j(\alpha)}{\beta_j}} \prod_{j \in \bar{\mathcal{J}}_{\infty}} (\mathbf{L}L_j^{-1})^{\frac{1+2\mu_j(\alpha)}{\gamma_j}}$.

Indeed, we have in view of (7.6), (7.7) and the definition of \mathbf{v}

$$\prod_{j=1}^d \widehat{\eta}_j^{1+2\mu_j(\alpha)}(\mathbf{v}, \mathbf{u}) \geq T^{-1}(\alpha)(\alpha^{-2}\delta_n)^{\frac{\omega(\alpha)\tau(p_{\pm})\beta(0)}{\gamma_{\pm}(\alpha)[z(\alpha)+\omega(\alpha)/\mathbf{u}] + \frac{\omega(\alpha)\tau(\infty)\beta(0)}{\beta_{\infty}(\alpha)[z(\alpha)+\omega(\alpha)/\mathbf{u}]}};$$

$$\prod_{j=1}^d \widehat{\eta}_j(\mathbf{v}, \mathbf{u}) \geq T^{-1}(0)(\alpha^{-2}\delta_n)^{\frac{\omega(\alpha)\tau(p_{\pm})\beta(0)}{\gamma_{\pm}(0)[z(\alpha)+\omega(\alpha)/\mathbf{u}] + \frac{\omega(\alpha)\tau(\infty)\beta(0)}{\beta_{\infty}(0)[z(\alpha)+\omega(\alpha)/\mathbf{u}]}};$$

where we have put $\frac{1}{\beta_{\infty}(\alpha)} = \sum_{j \in \mathcal{J}_{\infty}} \frac{1+2\mu_j(\alpha)}{\beta_j}$, $\frac{1}{\gamma_{\pm}(\alpha)} = \sum_{j \in \bar{\mathcal{J}}_{\infty}} \frac{1+2\mu_j(\alpha)}{\gamma_j}$. Note that for any $\alpha \in [0, 1]$,

$$\frac{\tau(p_{\pm})}{\gamma_{\pm}(\alpha)} + \frac{\tau(\infty)}{\beta_{\infty}(\alpha)} = \sum_{j \in \bar{\mathcal{J}}_{\infty}} \frac{(1+2\mu_j(\alpha))\tau(r_j)}{\beta_j} + \frac{\tau(\infty)}{\beta_{\infty}(\alpha)} = \frac{\tau(\infty)}{\beta(\alpha)} + \frac{1}{\omega(\alpha)\beta(0)}$$

and (7.9) and (7.10) are established.

4⁰. Simple algebra shows that for any $\mathbf{u} \in [1, \infty]$,

$$(\alpha^{-2}\delta_n)^{1-\frac{\omega(\alpha)\tau(\infty)\beta(0)/\beta(\alpha)+1}{z(\alpha)+\omega(\alpha)/\mathbf{u}}} = 2\mathbf{v}^2\mathfrak{z}^{-1}(\mathbf{v}),$$

and we deduce from (7.9) for any $\mathbf{u} \in (1, \infty]$ (recall that $\mathfrak{z} \equiv 2$ if $\mathbf{u} = \infty$)

$$(7.11) \quad \delta_n \prod_{j=1}^d \widehat{\eta}_j^{-1-2\mu_j(\alpha)}(\mathbf{v}, \mathbf{u}) \leq 2T^{-1}(\alpha)\alpha^2\mathbf{v}^2\mathfrak{z}^{-1}(\mathbf{v}).$$

Let us also prove that for any $\mathbf{u} \in [1, \infty]$ and all n large enough

$$(7.12) \quad \mathbf{v} > \mathbf{v} := (\alpha^{-2}\delta_n)^{\frac{1}{2+1/\beta(\alpha)}} \Rightarrow \mathfrak{z}(\mathbf{v}) \geq 2.$$

The latter inclusion follows from (6.7). Indeed, if $\tau(\infty) \leq 0$ then $\mathbf{v} \geq 1 \geq \mathbf{v}$. If $\tau(\infty) > 0$, then in view of (6.14)

$$\frac{\omega(\alpha)\tau(\infty)\beta(0)}{z(\alpha) + \omega(\alpha)/\mathbf{u}} - \frac{1}{2 + 1/\beta(\alpha)} = -\frac{1 + \omega(\alpha)/\mathbf{u}}{[z(\alpha) + \omega(\alpha)/\mathbf{u}][2 + 1/\beta(\alpha)]} < 0$$

so $\mathbf{v} > \mathbf{v}$. Note at last that for any $\mathbf{u} \in (1, \infty]$

$$(7.13) \quad \mathbf{v}\mathfrak{z}^{-1}(\mathbf{v}) = 2(\alpha^{-2}\delta_n)^{\frac{\omega(\alpha)\tau(\mathbf{u})\beta(0)}{z(\alpha)+\omega(\alpha)/\mathbf{u}}}.$$

5⁰. Let us proceed to the proof of the second assertion. Choose $\alpha^2 < a^2T(\alpha)/(4T_0) < 1$. We get from (7.1), (7.9) and (7.11) similar to (7.5)

$$(7.14) \quad F_n^2(\vec{\mathfrak{h}}(\mathbf{v}, \mathbf{u})) \leq \frac{T_0\delta_n}{\prod_{j=1}^d (\widehat{\eta}_j(\mathbf{v}, \mathbf{u}))^{1+2\mu_j(\alpha)}} \leq 2T_0T^{-1}(\alpha)\alpha^2\mathbf{v}^2\mathfrak{z}^{-1}(\mathbf{v})$$

$$\leq 2^{-1}\alpha^2\mathbf{v}^2\mathfrak{z}^{-1}(\mathbf{v}).$$

Thus to prove the assertion all we need to show is that $\vec{\mathfrak{h}}(\mathbf{v}, \mathbf{u}) \in \mathfrak{H}(\mathbf{v})$, that is, $G_n(\vec{\mathfrak{h}}(\mathbf{v}, \mathbf{u})) \leq a\mathbf{v}$. Let us distinguish three cases.

5⁰a. Let $\tau(\infty) \geq 0$. We remark that the definition of \mathbf{v} in this case yields $\mathbf{v} \leq 1$ for all n large enough and we obtain from (7.11) and (7.12) that

$$(7.15) \quad \delta_n \prod_{j=1}^d \widehat{\eta}_j^{-1-2\mu_j(\alpha)}(\mathbf{v}, \mathbf{u}) \leq T^{-1}(\alpha) \mathbf{a}^2 \mathbf{v}.$$

Then we have in view of (7.1), (7.8), (7.9) and (7.15) similar to (7.5)

$$(7.16) \quad G_n(\vec{h}(\mathbf{v}, \mathbf{u})) \leq \frac{T_0 \delta_n}{\prod_{j=1}^d (\widehat{\eta}_j(\mathbf{v}, \mathbf{u}))^{1+2\mu_j(\alpha)}} \leq T_0 T^{-1}(\alpha) \mathbf{a}^2 \mathbf{v} \leq a \mathbf{v}.$$

5⁰b. Let $\tau(\infty) < 0$, $\tau(p^*) \geq 0$ and $\alpha \neq 1$. Then by the imposed assumption $\mathbf{u} \leq \mathbf{u}^*$, and, therefore, $\tau(\mathbf{u}) \geq 0$. We get from (7.13), (6.14) and (7.14)

$$(7.17) \quad G_n(\vec{h}(\mathbf{v}, \mathbf{u})) = F_n^2(\vec{h}(\mathbf{v}, \mathbf{u})) \leq a^2 \mathbf{v} < a \mathbf{v}.$$

5⁰c. Let $\tau(\infty) < 0$, $\tau(p^*) \geq 0$, $\alpha = 1$. We have as previously

$$(7.18) \quad \begin{aligned} G_n^2(\vec{h}(\mathbf{u})) &\leq \frac{(\mathbf{T} + 2) \ln n}{n \prod_{j=1}^d (\mathfrak{h}_j(\mathbf{v}, \mathbf{u}))^{1+2\mu_j(\alpha)}} \frac{(\mathbf{T} + 2) \ln n}{n \prod_{j=1}^d \mathfrak{h}_j(\mathbf{v}, \mathbf{u})} \\ &\leq 2T_0^2 T^{-1}(1) T_1 \mathbf{a}^4 \mathbf{v}^2 \mathfrak{z}^{-1}(\mathbf{v}) \left[\frac{T(0) \mathbf{a}^{-2} \delta_n}{\prod_{j=1}^d \widehat{\eta}_j(\mathbf{v}, \mathbf{u})} \right]. \end{aligned}$$

Here, we have used (7.11) and put $T_1 = T^{-1}(0) L^{-1/\beta(0)}$. Our goal now is to show that for any $\mathbf{u} \in [1, \infty]$ and all n large enough

$$(7.19) \quad T(0) \mathbf{a}^{-2} \delta_n \mathfrak{z}^{-1}(\mathbf{v}) \prod_{j=1}^d \widehat{\eta}_j^{-1}(\mathbf{v}, \mathbf{u}) \leq 1.$$

In view of (7.10) and of the definition of $\mathfrak{z}(\cdot)$ in order to establish (7.19), it suffices to show that $z(1)/\omega(1) - 1 + 2/\mathbf{u} \geq 0$. Since we assumed $\tau(\infty) < 0$ and $\tau(p^*) \geq 0$, the required results follows from (6.11). Thus, (7.19) is proved. Then choosing a such that $T_0(2T^{-1}(1)T_1)^{1/2} \mathbf{a}^2 \leq a$, we obtain from (7.18) and (7.19) that

$$G_n(\vec{h}(\mathbf{v}, \mathbf{u})) \leq T_0(2T^{-1}(\alpha)T_1)^{1/2} \mathbf{a}^2 \mathbf{v} \leq a \mathbf{v}$$

for all all n large enough. The second assertion is proved. \square

7.2. *Proof of Proposition 3.* We start the proof with several remarks which will be useful in the sequel.

1⁰. Let us show that for all n large enough

$$(7.20) \quad \vec{h}(v, \mathbf{u}) \in (0, 1]^d \quad \forall v \in \mathcal{I}_{\mathbf{u}}(\alpha), \forall \mathbf{u} \geq \mathbf{y}.$$

In view of the definition of $\tilde{\eta}_j(\cdot, \mathbf{u})$, $j = 1, \dots, d$,

$$(7.21) \quad \tilde{\eta}_j^{\beta_j r_j}(v, \mathbf{u}) = (L L_j^{-1})^{r_j} \{ \mathbf{a}^{-2} \delta_n \}^{\frac{\mathbf{u}\omega(\alpha)}{\mathbf{u}+\omega(\alpha)}} v^{r_j - \frac{\mathbf{u}\omega(\alpha)(2+1/\beta(\alpha))}{\mathbf{u}+\omega(\alpha)}}, \quad j \in \bar{\mathcal{J}}_{\infty}.$$

Therefore, for any $v \in [\underline{\mathbf{v}}, 1]$ one has, taking into account that $\mathbf{L} \leq L_0$,

$$\tilde{\eta}_j^{\beta_j r_j}(v, \mathbf{u}) \leq \{\mathbf{a}^{-2} \delta_n\}^{\frac{\mathbf{u}\omega(\alpha)}{\mathbf{u}+\omega(\alpha)}} \underline{\mathbf{v}}^{1-\frac{\mathbf{u}\omega(\alpha)(2+1/\beta(\alpha))}{\mathbf{u}+\omega(\alpha)}} = 1, \quad j \in \bar{\mathcal{J}}_\infty.$$

It remains to note that $\mathbf{v} > \underline{\mathbf{v}}$ for all n large enough and, therefore,

$$(7.22) \quad \tilde{\eta}_j(v, \mathbf{u}) \leq 1, \quad j \in \bar{\mathcal{J}}_\infty, \forall v \in [v, 1] \cap \mathcal{I}_\mathbf{u}(\alpha).$$

We also have in view of the definition of $\tilde{\eta}_j(\cdot, \mathbf{u}), j = 1, \dots, d$,

$$\tilde{\eta}_j(v, \mathbf{u}) = (\mathbf{L}L_j^{-1}v)^{\frac{1}{\beta_j}} \leq 1, \quad j \in \mathcal{J}_\infty$$

for any $v \leq 1$. This together with (7.22) proves (7.20) in the cases when $\mathcal{I}_\mathbf{u}(\alpha) = [v, 1]$. Noting that $p^* < \infty$ is equivalent to $\mathcal{J}_\infty = \emptyset$, we deduce from (7.21) for any $v \geq 1$ and any $j = 1, \dots, d$,

$$\tilde{\eta}_j^{\beta_j r_j}(v, \mathbf{u}) \leq \{\mathbf{a}^{-2} \delta_n\}^{\frac{\mathbf{u}\omega(\alpha)}{\mathbf{u}+\omega(\alpha)}} v^{p^* - \frac{\mathbf{u}\omega(\alpha)(2+1/\beta(\alpha))}{\mathbf{u}+\omega(\alpha)}} \leq \{\mathbf{a}^{-2} \delta_n\}^{\frac{\mathbf{u}\omega(\alpha)}{\mathbf{u}+\omega(\alpha)}} v^{-\varkappa_\alpha(p^*, \mathbf{u})}.$$

Thus, for any $v \geq 1, j = 1, \dots, d$ and for all n large enough

$$(7.23) \quad \tilde{\eta}_j^{\beta_j r_j}(v, \mathbf{u}) \leq 1_{\{\varkappa_\alpha(p^*, \mathbf{u}) \geq 0\}} + \{\mathbf{a}^{-2} \delta_n\}^{\frac{\mathbf{u}\omega(\alpha)}{\mathbf{u}+\omega(\alpha)}} \tilde{\mathbf{v}}^{-\varkappa_\alpha(p^*, \mathbf{u})} 1_{\{\varkappa_\alpha(p^*, \mathbf{u}) < 0\}},$$

where we denoted $\tilde{\mathbf{v}} = \mathbf{v}_1$ if $\alpha \neq 1$ and $\tilde{\mathbf{v}} = \bar{\mathbf{v}}$ if $\alpha = 1$.

Let $\alpha = 1, p^* < \infty, \tau(p^*) > 0$. Then $\tilde{\mathbf{v}} = \mathbf{v}$ and we have for any $j = 1, \dots, d$ and $v \in [1, \mathbf{v}]$, using the definition of \mathbf{v} ,

$$\{\mathbf{a}^{-2} \delta_n\}^{\frac{\mathbf{u}\omega(1)}{\mathbf{u}+\omega(1)} - \frac{\varkappa_1(p^*, \mathbf{u})\omega(1)\tau(\infty)\beta(0)}{z(1)+\omega(1)/\mathbf{u}}} = \{\mathbf{a}^{-2} \delta_n\}^{\frac{p^* \tau(p^*)\omega(1)}{z(1)+\omega(1)/\mathbf{u}}} \rightarrow 0, \quad n \rightarrow \infty$$

in view of (6.13). Hence, (7.20) follows from (7.23) in this case.

Let $\alpha = 1, \tau(p^*) \leq 0$. Then $\tilde{\mathbf{v}} = \mathbf{v}_2$ and we have for any $v \in [1, \mathbf{v}_2]$ in view of the definition of \mathbf{v}_2

$$\tilde{\eta}_j^{\beta_j r_j}(v, \mathbf{u}) \leq \tilde{\eta}_j^{\beta_j r_j}(\mathbf{v}_2, \mathbf{u}) = \{\mathbf{a}^{-2} \delta_n\}^{\frac{\mathbf{u}\omega(1)}{\mathbf{u}+\omega(1)}} \mathbf{v}_2^{-\varkappa_1(p^*, \mathbf{u})} = 1, \quad j = 1, \dots, d$$

and, therefore (7.20) follows from (7.23) in this case.

Let $\alpha \neq 1, \mathbf{u} < \infty$. First, we note that $\tau(\infty) < 0$ and $\mathbf{u} \geq \mathbf{y}$ imply

$$1 - \mathbf{u}/\omega(0) + 1/\beta(0) = 1 - \mathbf{u} + \mathbf{u}\tau(\mathbf{u}) \leq 1 - \mathbf{u} + \mathbf{u}\tau(\mathbf{y}) \leq 1 - \mathbf{u} < 0$$

since either $\mathbf{y} = p^*$, which is equivalent to $\tau(p^*) \leq 0$, or $\mathbf{y} = \mathbf{u}^*$ and then $\tau(\mathbf{y}) = 0$. Thus $\mathbf{v}_1 \rightarrow \infty, n \rightarrow \infty$ and, therefore, in view of (7.23) for any $v \in [1, \mathbf{v}_1]$ and $j = 1, \dots, d$ one has

$$\tilde{\eta}_j^{\beta_j r_j}(v, \mathbf{u}) \leq 1_{\{\varkappa_\alpha(p^*, \mathbf{u}) \geq 0\}} + \{\mathbf{a}^{-2} \delta_n\}^{\frac{\mathbf{u}\omega(\alpha)}{\mathbf{u}+\omega(\alpha)}} \mathbf{v}_1^{-\varkappa_\alpha(p^*, \mathbf{u})} 1_{\{\varkappa_\alpha(p^*, \mathbf{u}) < 0\}}.$$

Note that $1 - \mathbf{u}/\omega(0) + 1/\beta(0) = \varkappa_0(p^*, \mathbf{u})[1/\mathbf{u} + 1/\omega(0)] - (\mathbf{u} - p^*)[1/\mathbf{u} + 1/\omega(0)]$ and, therefore,

$$-\frac{\varkappa_0(p^*, \mathbf{u})}{1 - \mathbf{u}/\omega(0) + 1/\beta(0)} \geq -\frac{\mathbf{u}\omega(0)}{\mathbf{u} + \omega(0)} \Rightarrow \mathbf{v}_1^{-\varkappa_0(p^*, \mathbf{u})} \leq \{\mathbf{a}^{-2} \delta_n\}^{-\frac{\mathbf{u}\omega(0)}{\mathbf{u}+\omega(0)}}.$$

It remains to note that if $\tau(\infty) \geq 0$ then $\mathbf{u}^* = \infty$ and, therefore, $\mathbf{u} = \infty$. It implies $\mathbf{v}_1 = 1$ and $\mathcal{I}_{\mathbf{u}}(\alpha) = [v, 1]$ and this case has been already treated. This completes the proof of (7.20).

2^0 . Remark that there obviously exists $0 < \mathbf{S} := S(\vec{\beta}, \vec{r}, \vec{\mu}, p) < \infty$ independent of \vec{L} such that

$$\lim_{n \rightarrow \infty} (\ln n)^{-1} \sup_{\alpha \in \{0,1\}} \sup_{\mathbf{u} \in [1, \infty]} \sup_{v \in \mathcal{I}_{\mathbf{u}}(\alpha)} \sum_{j=1}^d |\ln(\mathbf{h}_j(v, \mathbf{u}))| = \mathbf{S}.$$

Hence, in view of (7.20) one has for all n large enough and $v \in \mathcal{I}_{\mathbf{u}}(\alpha)$,

$$(7.24) \quad F_n(\vec{\mathbf{h}}(v, \mathbf{u})) \leq \sqrt{(\mathbf{S} + 2)n^{-1} \ln n} \prod_{j=1}^d (\mathbf{h}_j(v, \mathbf{u}))^{-\frac{1}{2} - \mu_j(\alpha)};$$

$$(7.25) \quad G_n(\vec{\mathbf{h}}(v, \mathbf{u})) \leq (\mathbf{S} + 2)n^{-1} \ln n \prod_{j=1}^d (\mathbf{h}_j(v, \mathbf{u}))^{-1 - \mu_j(\alpha)}.$$

Taking into account that $\mathbf{h}_j(v, \mathbf{u}) \geq e^{-1} \tilde{\eta}_j(v, \mathbf{u})$ and setting

$$S_0 = [\mathbf{S} + 2] e^{d+2 \sum_{j=1}^d \mu_j} \mathbf{L}^{-\frac{1}{\beta(1)}},$$

we obtain from (7.2) for any $\alpha \in [0, 1]$ and $v \in \mathcal{I}_{\mathbf{u}}(\alpha)$,

$$(7.26) \quad (\mathbf{S} + 2)n^{-1} \ln(n) \prod_{j=1}^d (\mathbf{h}_j(v, \mathbf{s}))^{1+2\mu_j(\alpha)} \leq 2S_0 a^2 v^2 \mathfrak{z}^{-1}(v).$$

From now on, we choose $a \leq a/(2S_0) < 1$. It yields in view of (7.24), (7.26)

$$(7.27) \quad F_n^2(\vec{\mathbf{h}}(v, \mathbf{u})) \leq a^2 v^2 \mathfrak{z}^{-1}(v) \quad \forall v \in \mathcal{I}_{\mathbf{u}}(\alpha).$$

3^0 . Since (7.27) holds, to complete the proof of Proposition 3 all we need to show is that $G_n(\vec{\mathbf{h}}(v, \mathbf{u})) \leq av, \forall v \in \mathcal{I}_{\mathbf{u}}(\alpha)$. Let us distinguish three cases.

3^0 a. Let $p^* = \infty$ or $\alpha \neq 1, \mathbf{u} = \infty$. First, we note that in these cases $\mathcal{I}_{\mathbf{u}}(\alpha) = [v, 1]$. Next, in view of the second inequality in (7.24), (7.20), (7.26) and (7.27) we obtain for any $v \in \mathcal{I}_{\mathbf{u}}(\alpha)$:

$$(7.28) \quad G_n(\vec{\mathbf{h}}(v, \mathbf{u})) \leq \frac{(\mathbf{S} + 2) \ln n}{n \prod_{j=1}^d (\mathbf{h}_j(v, \mathbf{u}))^{1+2\mu_j(\alpha)}} \leq a^2 v^2 \mathfrak{z}^{-1}(v) \leq av.$$

To get the last inequality, we have used that $a < 1, \mathfrak{z}(\cdot) \geq 2$ and $v \leq 1$.

3^0 b. Let $\alpha \neq 1, p^* < \infty, \mathbf{u} < \infty$. We have in view of (7.25) and (7.26)

$$G_n(\vec{\mathbf{h}}(v, \mathbf{u})) \leq a^2 v^2 \mathfrak{z}^{-1}(v) \quad \forall v \in \mathcal{I}_{\mathbf{u}}(0).$$

Here, we have also used that $\mu_j(\alpha) = 0$ for all j . Simple algebra shows that

$$v \mathfrak{z}^{-1}(v) = \{a^{-2} \delta_n\}^{\frac{\mathbf{u}\omega(0)}{\mathbf{u}+\omega(0)}} v^{\frac{\mathbf{u}-\omega(0)-\omega(0)/\beta(0)}{\mathbf{u}+\omega(0)}}, \quad \mathbf{u} \neq \infty,$$

and since $\mathbf{u} - \omega(0) - \omega(0)/\beta(0) > 0$ for any $\mathbf{u} \geq \mathbf{u}^*$, the result follows from

$$\sup_{v \in \mathcal{I}_{\mathbf{u}}(0)} v \mathfrak{z}^{-1}(v) = \mathbf{v}_1 \mathfrak{z}^{-1}(\mathbf{v}_1) = 1.$$

3⁰c. Let $\alpha = 1, p^* < \infty$. For any $v \in \mathcal{I}_{\mathbf{u}}(1)$, we have in view of the second inequality in (7.24) and (7.26), denoting $S_1 = S_0 L^{-1/\beta(0)}$,

$$\begin{aligned} G_n^2(\vec{\mathbf{h}}(v, \mathbf{u})) &\leq \frac{(\mathbf{S} + 2) \ln n}{n \prod_{j=1}^d (\mathbf{h}_j(v, \mathbf{u}))^{1+2\mu_j(\alpha)}} \frac{(\mathbf{S} + 2) \ln n}{n \prod_{j=1}^d (\mathbf{h}_j(v, \mathbf{u}))} \\ (7.29) \qquad \qquad \qquad &\leq S_1 \alpha^2 a^2 v^2 \mathfrak{z}^{-1}(v) \alpha^{-2} \delta_n \prod_{j=1}^d \tilde{\eta}_j^{-1}(v, \mathbf{u}). \end{aligned}$$

Our goal now is to show that for all n large enough

$$(7.30) \qquad \sup_{v \in \mathcal{I}_{\mathbf{u}}(1)} \alpha^{-2} \delta_n \mathfrak{z}^{-1}(v) \prod_{j=1}^d \tilde{\eta}_j^{-1}(v, \mathbf{u}) \leq 1.$$

Denoting $P(v) = \alpha^{-2} \delta_n \mathfrak{z}^{-1}(v) \prod_{j=1}^d \tilde{\eta}_j^{-1}(v, \mathbf{u})$ we easily compute

$$(7.31) \qquad P(v) = 2^{-1} \left\{ \alpha^{-2} \delta_n \right\}^{\frac{2\mathbf{u}\omega(1)(\mathbf{Y}+1/\mathbf{u})}{\omega(1)+\mathbf{u}}} v^{\frac{2\mathbf{u}\omega(1)\pi(\mathbf{u})}{\mathbf{u}+\omega(1)}}, \quad v > 0.$$

It yields obviously

$$(7.32) \qquad \sup_{v \in \mathcal{I}_{\mathbf{u}}(1)} \alpha^{-2} \delta_n \mathfrak{z}^{-1}(v) \prod_{j=1}^d \tilde{\eta}_j^{-1}(v, \mathbf{u}) \leq \max[P(\mathbf{v}), P(\tilde{\mathbf{v}})],$$

where $\tilde{\mathbf{v}} \in \{\bar{\mathbf{v}}, \mathbf{v}_3\}$. We deduce from (7.31) that for any $\mathbf{u} \in [1, \infty]$,

$$(7.33) \qquad P(\mathbf{v}) = 2^{-1} \left\{ \alpha^{-2} \delta_n \right\}^{\frac{2+1/\beta(\alpha)-1/\beta(0)}{2+1/\beta(\alpha)}} \rightarrow 0, \quad n \rightarrow \infty.$$

3⁰c1. Consider the case $\mathbf{Z}_{\mathbf{y}, \mathbf{u}}(\alpha) \geq 0$. Here, $\tilde{\mathbf{v}} = \bar{\mathbf{v}}$.

If $\tau(p^*) > 0$, then $\bar{\mathbf{v}} = \mathbf{v}$. Moreover, $\mathbf{y} = \mathbf{u}^*$ since $\mathbf{u}^* = \infty$ if $\tau(\infty) \geq 0$ and $\tau(\mathbf{u}^*) = 0$ if $\tau(\infty) < 0$. Hence $z(1)/\omega(1) - 1 + 2/\mathbf{u} \geq 0$ in view of (6.11) and we have, in view of the definition of \mathbf{v} ,

$$(7.34) \qquad P(\mathbf{v}) = 2^{-1} \left\{ \alpha^{-2} \delta_n \right\}^{\frac{\mathbf{u}\omega(1)(1/\omega(1)-1/\omega(0)+2/\mathbf{u})}{\omega(1)+\mathbf{u}} + \frac{\mathbf{u}\omega^2(1)\tau(\infty)\beta(0)\pi(\mathbf{u})}{[\mathbf{u}+\omega(1)][z(1)+\omega(1)/\mathbf{u}]}}.$$

Note that

$$\begin{aligned} &\frac{\mathbf{u}\omega(1)(1/\omega(1) - 1/\omega(0) + 2/\mathbf{u})}{\omega(1) + \mathbf{u}} + \frac{\mathbf{u}\omega^2(1)\tau(\infty)\beta(0)\pi(\mathbf{u})}{[\mathbf{u} + \omega(1)][z(1) + \omega(1)/\mathbf{u}]} \\ &= 1 - \frac{\omega(1)[1/\omega(0) - 1/\mathbf{u}]}{z(1) + \omega(1)/\mathbf{u}} - \frac{\omega(1)\tau(\infty)}{z(1) + \omega(1)/\mathbf{u}} = 1 - \frac{\omega(1)[1 - 1/\mathbf{u}]}{z(1) + \omega(1)/\mathbf{u}} > 0. \end{aligned}$$

To get the last inequality, we have used that

$$1 - \frac{\omega(1)[1 - 1/\mathbf{u}]}{z(1) + \omega(1)/\mathbf{u}} > 0 \quad \Leftrightarrow \quad z(1)/\omega(1) - 1 + 2/\mathbf{u} > 0.$$

Thus, we conclude that $P(\mathbf{v}) \leq 1$, for all large n , which together with (7.33) implies (7.30) in the considered case.

If $\tau(p^*) \leq 0$, then $\bar{\mathbf{v}} = \mathbf{v}_2$ and, moreover, $\mathbf{y} = \mathbf{p}^*$. Also, $\varkappa_1(p^*, \mathbf{u}) < 0$ thanks to (6.12) of Lemma 3. We have, in view of the definition of \mathbf{v}_2 ,

$$(7.35) \quad P(\mathbf{v}_2) = 2^{-1} \left\{ \mathfrak{a}^{-2} \delta_n \right\}^{\frac{\mathbf{u}\omega(1)(1/\omega(1) - 1/\omega(0) + 2/\mathbf{u})}{\omega(1) + \mathbf{u}} + \frac{[\mathbf{u}\omega(1)]^2 \pi(\mathbf{u})}{\varkappa_1(p^*, \mathbf{u})[\mathbf{u} + \omega(1)]^2}}.$$

After routine computations, we come to the following equality:

$$\begin{aligned} & \frac{\mathbf{u}\omega(1)(1/\omega(1) - 1/\omega(0) + 2/\mathbf{u})}{\omega(1) + \mathbf{u}} + \frac{[\mathbf{u}\omega(1)]^2 \pi(\mathbf{u})}{\varkappa_1(p^*, \mathbf{u})[\mathbf{u} + \omega(1)]^2} \\ &= - \frac{2\mathbf{u}\omega(1)p^*[Y - (X + 1)(\mathbf{y})^{-1} + 1/\mathbf{u}]}{\varkappa_1(p^*, \mathbf{u})[\mathbf{u} + \omega(1)]} \geq 0. \end{aligned}$$

Hence, $P(\mathbf{v}_2) \leq 1$ for all n large enough, which together with (7.33) allows us to assert (7.30) in the considered case.

3⁰c3. Consider the case $\mathbf{Z}_{\mathbf{y}, \mathbf{u}}(\alpha) < 0$. Here, $\tilde{\mathbf{v}} = \mathbf{v}_3$. If $\pi(\mathbf{u}) \leq 0$, then $\mathbf{v}_3 = \infty$ and obviously $P(\mathbf{v}_3) = 0$. If $\pi(\mathbf{u}) > 0$, then $P(\mathbf{v}_3) = 1$ in view of (7.31) and the definition of \mathbf{v}_3 . This completes the proof (7.30).

Finally, in the case 3⁰c, choosing $\mathfrak{a} \leq \sqrt{1/S_1}$, we deduce from (7.29) and (7.30) that for all n large enough $G_n(\bar{\mathbf{h}}(v, \mathbf{u})) \leq \sqrt{S_1} \mathfrak{a} v \leq av, v \in \mathcal{I}_{\mathbf{u}}(1)$. \square

APPENDIX

PROOF OF LEMMA 3. Note that

$$\begin{aligned} z(\alpha) + \omega(\alpha)/\mathbf{u} &= \omega(\alpha)(2 + 1/\beta(\alpha))\beta(0)\tau(p^*) + 1 - \omega(\alpha)(2 + 1/\beta(\alpha))(p^*)^{-1} \\ &\quad + \omega(\alpha)/\mathbf{u} \\ &= \omega(\alpha)(2 + 1/\beta(\alpha))\beta(0)\tau(p^*) - (p^*)^{-1}(1 + \omega(\alpha)/\mathbf{u})\varkappa_\alpha(p^*, \mathbf{u}) \end{aligned}$$

and (6.13) follows. On the other hand, we have

$$z(\alpha)/\omega(\alpha) - 1 + 2/\mathbf{u} = (2 + 2X)\beta(0)\tau(\infty) + 2Y + 2/\mathbf{u}$$

and (6.11) is checked if $\tau(\infty) \geq 0$ since $X, Y \geq 0$. If $\tau(\infty) < 0$ and $\tau(p^*) \geq 0$, then we note first that necessarily $\mathbf{u}^* \geq p^*$ since $\tau(\mathbf{u}^*) = 0$ and $\tau(\cdot)$ is strictly decreasing. Hence $\mathbf{y} = \mathbf{u}^*$ and we have

$$z(\alpha)/\omega(\alpha) - 1 + 2/\mathbf{u} = 2\{Y - (X + 1)\mathbf{y}^{-1} + 1/\mathbf{u}\} = 2\mathbf{Z}_{\mathbf{y}, \mathbf{u}}(\alpha) \geq 0$$

and (6.11) is established. Let us prove (6.12). We obviously have

$$\frac{\varkappa_\alpha(p^*, \mathbf{u})(\mathbf{u} + \omega(\alpha))}{p^* \mathbf{u} \omega(\alpha)} = -2[Y - (X + 1)/p^* + 1/\mathbf{u}] + \tau(p^*) - 1 + 1/\mathbf{u}.$$

If $\tau(p^*) \leq 0$, then necessarily $\mathbf{y} = p^*$ and, therefore, for any $\mathbf{u} > 1$,

$$\frac{\varkappa_\alpha(p^*, \mathbf{u})(\mathbf{u} + \omega(\alpha))}{p^* \mathbf{u} \omega(\alpha)} = -2\mathbf{Z}_{\mathbf{y}, \mathbf{u}}(\alpha) + \tau(p^*) - 1 + 1/\mathbf{u} < 0,$$

since we have supposed that $\mathbf{Z}_{\mathbf{y}, \mathbf{u}}(\alpha) \geq 0$.

Let us prove (6.14). If $\tau(\infty) \geq 0$, then $z(\alpha) \geq 1$ and (6.14) follows. If $\mathbf{Z}_{\mathbf{y}, \mathbf{u}}(\alpha) \geq 0$, $\tau(p^*) \geq 0$, then (6.14) follows from (6.11) since $\mathbf{u} > 1$.

It remains to prove (6.15). If $\alpha \neq 1$, (6.15) is trivial because $\mu(\alpha) = Y = X = 0$. If $\alpha = 1$, noting that $r_j \leq p^* \leq \mathbf{y}$ for any $j = 1, \dots, d$, we have

$$Y - [X + 1]\mathbf{y}^{-1} + 1/\mathbf{u} \geq \mu(1)[1 - \tau(\mathbf{y})] - 1/\mathbf{y} + 1/\mathbf{u} \geq \mu(1) - 1/\mathbf{y} + 1/\mathbf{u}$$

and (6.15) follows. To get the last inequality, we have used that $\tau(\mathbf{u}^*) = 0$ and that $\tau(\cdot)$ is strictly decreasing, so $\tau(\mathbf{y}) \leq 0$. \square

PROOF OF LEMMA 4. Let us first prove the following assertions.

1⁰. Let $\ell \in \mathbb{N}^*$, $p > 1$ and K satisfying Assumption 4 be fixed. Then for any $\vec{\beta} \in (0, \ell]^d$, $\vec{r} \in [1, \infty)^d$ and $\vec{L} \in (0, \infty)^d$ one can find $C_1 > 0$ independent of \vec{L} such that $\forall \vec{h} \in \mathcal{H}^d$:

$$(A.2) \quad \mathbf{B}_{j, r_j, \mathbb{N}_{\vec{r}, d}(\vec{\beta}, \vec{L})}(h_j) \leq C_1 L_j h_j^{\beta_j}, \quad j = 1, \dots, d.$$

If additionally $\tau(p^*) > 0$, then

$$(A.3) \quad \mathbf{B}_{j, q_j, \mathbb{N}_{\vec{r}, d}(\vec{\beta}, \vec{L})}(h_j) \leq C_1 L_j h_j^{\gamma_j}, \quad j = 1, \dots, d.$$

At last, (A.2) and (A.3) remain true if one replaces the quantity \mathbf{B} by \mathbf{B}^* .

1⁰a. In view of Lemma 5 in Lepski (2015), if $\tau(p^*) > 0$ then

$$(A.4) \quad \mathbb{N}_{\vec{r}, d}(\vec{\beta}, \vec{L}) \subseteq \mathbb{N}_{\vec{q}, d}(\vec{\gamma}, c_2 \vec{L}),$$

where c_2 is independent on \vec{L} . Note also that $\gamma_j \leq \beta_j$ for any $j = 1, \dots, d$.

Let $(\vec{\pi}, \vec{s})$ be either $(\vec{\beta}, \vec{r})$ or $(\vec{\gamma}, \vec{q})$ and without further mentioning the couple $(\vec{\gamma}, \vec{q})$ is used below under the condition $\tau(p^*) > 0$. We obviously have for any $\vec{h} \in \mathcal{H}$

$$\begin{aligned} b_{\vec{h}, f, j}(x) &:= \sup_{h \in \mathcal{H}: h \leq \vec{h}} \left| \int_{\mathbb{R}} \mathcal{K}_\ell(u) [f(x + u\mathbf{h}e_j) - f(x)] \nu_1(du) \right| \\ &= \sup_{h \in \mathcal{H}: h \leq \vec{h}} \left| \int_{\mathbb{R}} \mathcal{K}_\ell(u) [\Delta_{uh, j} f(x)] \nu_1(du) \right|. \end{aligned}$$

For $j = 1, \dots, d$, we have

$$\begin{aligned} \int_{\mathbb{R}} \mathcal{K}_\ell(u) \Delta_{uh,j} f(x) v_1(du) &= \int_{\mathbb{R}} \sum_{i=1}^{\ell} \binom{\ell}{i} (-1)^{i+1} \frac{1}{i} \mathcal{K}_\ell\left(\frac{u}{i}\right) [\Delta_{hu,j} f(x)] v_1(du) \\ &= (-1)^{\ell-1} \int_{\mathbb{R}} \mathcal{K}_\ell(z) \sum_{i=1}^{\ell} \binom{\ell}{i} (-1)^{i+\ell} [\Delta_{izh,j} f(x)] v_1(dz) \\ &= (-1)^{\ell-1} \int_{\mathbb{R}} \mathcal{K}_\ell(z) [\Delta_{zh,j}^\ell f(x)] v_1(dz). \end{aligned}$$

The last equality follows from the definition of the ℓ th order difference operator determined in Definition 1. Hence, for any $j \in \mathcal{J}_\infty$ we have in view of the definition of the Nikol'skii class (remind that $\gamma_j = \beta_j, j \in \mathcal{J}_\infty$)

$$\|b_{\mathbf{h},f,j}\|_\infty \leq L_j \sup_{h \in \mathcal{H}: h \leq \mathbf{h}} h_j^{\pi_j} \int_{\mathbb{R}} |\mathcal{K}_\ell(z)| |z|^{\pi_j} v_1(dz).$$

This yields for any $\mathbf{h} \in \mathcal{H}$

$$(A.5) \quad \mathbf{B}_{j,\infty, \mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L})}(\mathbf{h}) \leq c_1 L_j \mathbf{h}^{\pi_j},$$

and (A.2) and (A.3) are proved for any $j \in \mathcal{J}_\infty$.

Let $j \in \tilde{\mathcal{J}}_\infty$. Choosing \mathbf{k} from the relation $e^{\mathbf{k}} = \mathbf{h}$ (recall that $\mathbf{h} \in \mathcal{H}$), we have for any $x \in \mathbb{R}^d$,

$$\begin{aligned} b_{\mathbf{h},f,j}(x) &= \sup_{k \leq \mathbf{k}} \left| \int_{\mathbb{R}} \mathcal{K}_\ell(z) [\Delta_{ze^k,j}^\ell f(x)] v_1(dz) \right| \\ &=: \lim_{l \rightarrow -\infty} \sup_{l \leq k \leq \mathbf{k}} \left| \int_{\mathbb{R}} \mathcal{K}_\ell(z) [\Delta_{ze^k,j}^\ell f(x)] v_1(dz) \right|. \end{aligned}$$

Using the monotone convergence theorem and the triangle inequality,

$$\begin{aligned} \|b_{\mathbf{h},f,j}\|_{s_j} &= \lim_{l \rightarrow -\infty} \sup_{l \leq k \leq \mathbf{k}} \left\| \int_{\mathbb{R}} \mathcal{K}_\ell(z) [\Delta_{ze^k,j}^\ell f(\cdot)] v_1(dz) \right\|_{s_j} \\ &\leq \sum_{k=-\infty}^{\mathbf{k}} \left\| \int_{\mathbb{R}} \mathcal{K}_\ell(z) [\Delta_{ze^k,j}^\ell f(\cdot)] v_1(dz) \right\|_{s_j}. \end{aligned}$$

By the Minkowski inequality for integrals [see, e.g., Folland (1999), Section 6.3], we obtain

$$\|b_{\mathbf{v},f,j}\|_{s_j} \leq \sum_{k=-\infty}^{\mathbf{k}} \int_{\mathbb{R}} |\mathcal{K}_\ell(z)| \|\Delta_{ze^k,j}^\ell f\|_{s_j} v_1(dz), \quad j = 1, \dots, d$$

Taking into account that $f \in \mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L})$ and (A.4), we have for any $\mathbf{h} \in \mathcal{H}$ and $j = 1, \dots, d$,

$$(A.6) \quad \|b_{\mathbf{h},f,j}\|_{s_j} \leq \left[\int_{\mathbb{R}} |\mathcal{K}_\ell(z)| |z|^{\beta_j} v_1(dz) \right] L_j \sum_{k=-\infty}^{\mathbf{k}} e^{k\pi_j} \leq c_1 L_j \mathbf{h}^{\pi_j}.$$

This proves (A.2) and (A.3) for any $j \in \tilde{\mathcal{J}}_\infty$.

1⁰b. Set $\mathbb{F} = \mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L})$ and recall that

$$\begin{aligned} \mathbf{B}_{j,s_j,\mathbb{F}}^*(\mathbf{h}) &:= \sup_{f \in \mathbb{F}} \sum_{h \in \mathcal{H}: h \leq \mathbf{h}} \left\| \int_{\mathbb{R}} \mathcal{K}_\ell(u) [f(x + u\mathbf{h}e_j) - f(x)] \nu_1(du) \right\|_{s_j} \\ &\leq \sup_{f \in \mathbb{F}} \sum_{h \in \mathcal{H}: h \leq \mathbf{h}} \|b_{h,f,j}\|_{s_j}. \end{aligned}$$

Hence, (A.2) and (A.3) with \mathbf{B}^* instead of \mathbf{B} follows from (A.5) and (A.6).

2⁰. First, we remark that $\mathbf{h}_j(\cdot, \mathbf{1}) \equiv \mathbf{h}_j(\cdot, \infty) \equiv \mathfrak{h}_j(\cdot, \infty) \leq (\mathbf{L}L_j^{-1})^{\frac{1}{\beta_j}}$, $j \in \mathcal{J}_\infty$. Then we get, from (A.2) and (6.2) for any $j \in \mathcal{J}_\infty$,

$$\begin{aligned} \mathbf{B}_{j,\infty,\mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L})}(\mathbf{h}_j(v, \mathbf{1})) &= \mathbf{B}_{j,\infty,\mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L})}(\mathbf{h}_j(v, \infty)) \\ &= \mathbf{B}_{j,\infty,\mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L})}(\mathfrak{h}_j(v, \infty)) \\ &\leq \mathbf{c}v \quad \forall v > 0. \end{aligned}$$

It yields, in particular, that for any $v > 0$,

$$(A.7) \quad J(\vec{\mathbf{h}}(v, \mathbf{1}), v) \supseteq \mathcal{J}_\infty, \quad J(\vec{\mathbf{h}}(v, \infty), v) \supseteq \mathcal{J}_\infty, \quad J(\vec{\mathbf{h}}(v, \infty), v) \supseteq \mathcal{J}_\infty.$$

Thus, putting

$$\lambda_2(v) = \sum_{j \in \mathcal{J}_\infty} v^{-r_j} L_j^{r_j} [\mathbf{h}_j(v, \mathbf{1})]^{r_j \beta_j} + v^{-2} (\ln n/n) \prod_{j=1}^d (\mathbf{h}_j(v, \mathbf{1}))^{-1-2\mu_j(\alpha)}$$

we obtain in view of (A.2), Propositions 2, 3, (A.7) and the definition of $\vec{\mathbf{h}}(\cdot, \mathbf{s})$, $\mathbf{s} \in \{1, \infty\}$ that for any $v \in \mathcal{I}_\infty(\alpha)$ and $v \in [\mathbf{y}, 1]$, respectively,

$$(A.8) \quad \begin{aligned} \mathbf{A}_{\vec{r}}(v, \mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L}), \infty) &\leq C_1 \sum_{j \in \mathcal{J}_\infty} v^{-r_j} L_j^{r_j} [\mathbf{h}_j(v, \infty)]^{r_j \beta_j} \\ &\leq C_3 \delta_n^{\omega(\alpha)} v^{-\omega(\alpha)(2+1/\beta(\alpha))}, \end{aligned}$$

$$(A.9) \quad \mathbf{A}_{\vec{r}}(v, \mathbb{N}_{\vec{r},d}(\vec{\beta}, \vec{L})) \leq C_4 \lambda_2(v) \leq C_5 \delta_n^{\frac{\omega(\alpha)}{\omega(\alpha)+1}} v^{-\frac{\omega(\alpha)(2+1/\beta(\alpha))}{\omega(\alpha)+1}}.$$

To get (A.9), we have used that for all n large enough and all $v \in [\mathbf{y}, 1]$,

$$F_n(\vec{\mathbf{h}}(v, \mathbf{1})) \leq C_2 (\ln n/n) \prod_{j=1}^d (\mathbf{h}_j(v, \mathbf{1}))^{-1-2\mu_j(\alpha)},$$

where C_2 is independent of \vec{L} . This follows from assertions (7.1) and (7.3) established in the proof of Proposition 2. The first and second assertions of the lemma follow now from (A.8) and (A.9), respectively.

Moreover, if $\tau(p^*) > 0$ we get in view of (A.3), Propositions 2 and (A.7)

$$\begin{aligned} \mathbf{v}^p \Lambda_{\vec{q}}(\mathbf{v}, \mathbb{N}_{\vec{r}, d}(\vec{\beta}, \vec{L}), \infty) &\leq C_1 \sum_{j \in \bar{\mathcal{J}}_\infty} \mathbf{v}^{-q_j} L_j^{q_j} [h_j(\mathbf{v}, \infty)]^{q_j \gamma_j} \\ &\leq C_2 \delta_n^{\frac{\omega(\alpha)\tau(p)\beta(0)}{z(\alpha)}} \end{aligned}$$

and the third assertion of the lemma is established. \square

PROOF OF LEMMA 5. Note that

$$\begin{aligned} 1/\gamma(\alpha) - 1/\beta(\alpha) &= 1/\gamma_\pm(\alpha) - 1/\beta_\pm(\alpha) \\ &= \sum_{j \in \mathcal{J}_\pm} \frac{1 + 2\mu_j(\alpha)}{\beta_j} [\tau(r_j)/\tau(p_\pm) - 1] \\ &= [\beta(0)\tau(p_\pm)]^{-1} \sum_{j \in \mathcal{J}_\pm} \frac{1 + 2\mu_j(\alpha)}{\beta_j} (1/r_j - 1/p_\pm) \\ &= [\tau(p_\pm)\beta(0)]^{-1} [1/\omega(\alpha) - 1/(\beta_\pm(\alpha)p_\pm)]. \end{aligned}$$

Moreover, in view of the latter inequality,

$$\begin{aligned} 1/\omega(\alpha) - 1/v(\alpha) &= 1/\omega(\alpha) - 1/(p_\pm\gamma_\pm(\alpha)) \\ &= 1/\omega(\alpha) - 1/(p_\pm\beta_\pm(\alpha)) \\ &\quad - [\tau(p_\pm)\beta(0)p_\pm]^{-1} [1/\omega(\alpha) - 1/(\beta_\pm(\alpha)p_\pm)] \\ &= \{1 - [\tau(p_\pm)\beta(0)p_\pm]^{-1}\} [1/\omega(\alpha) - 1/(\beta_\pm(\alpha)p_\pm)]. \end{aligned}$$

Note that $1 - [\tau(p_\pm)\beta(0)p_\pm]^{-1} = \tau(\infty)/\tau(p_\pm)$ and the lemma follows. \square

REFERENCES

- AKAKPO, N. (2012). Adaptation to anisotropy and inhomogeneity via dyadic piecewise polynomial selection. *Math. Methods Statist.* **21** 1–28. [MR2901269](#)
- BIRGÉ, L. (2014). Model selection for density estimation with \mathbb{L}_2 -loss. *Probab. Theory Related Fields* **158** 533–574. [MR3176358](#)
- BUTUCEA, C. and TSYBAKOV, A. B. (2008). Sharp optimality in density deconvolution with dominating bias. I, II. *Theory Probab. Appl.* **52** 111–128, 237–249. [MR2354572](#)
- COMTE, F. and LACOUR, C. (2013). Anisotropic adaptive kernel deconvolution. *Ann. Inst. Henri Poincaré Probab. Stat.* **49** 569–609. [MR3088382](#)
- COMTE, F., ROZENHOLC, Y. and TAUPIN, M.-L. (2006). Penalized contrast estimator for adaptive density deconvolution. *Canad. J. Statist.* **34** 431–452. [MR2328553](#)
- DE GUZMÁN, M. (1975). *Differentiation of Integrals in R^n* . *Lecture Notes in Mathematics* **481**. Springer, Berlin. [MR0457661](#)
- DEVROYE, L. and LUGOSI, G. (1997). Nonasymptotic universal smoothing factors, kernel complexity and Yatracos classes. *Ann. Statist.* **25** 2626–2637. [MR1604428](#)

- DONOHO, D. L., JOHNSTONE, I. M., KERKYACHARIAN, G. and PICARD, D. (1996). Density estimation by wavelet thresholding. *Ann. Statist.* **24** 508–539. [MR1394974](#)
- DUVAL, C. (2017). A note on a fixed-point method for deconvolution. *Statistics* **51** 347–362. [MR3609324](#)
- EFRŌIMOVICH, S. YU. (1986). Nonparametric estimation of a density of unknown smoothness. *Theory Probab. Appl.* **30** 557–568. [MR0805304](#)
- FAN, J. (1991). On the optimal rates of convergence for nonparametric deconvolution problems. *Ann. Statist.* **19** 1257–1272. [MR1126324](#)
- FAN, J. (1993). Adaptively local one-dimensional subproblems with application to a deconvolution problem. *Ann. Statist.* **21** 600–610. [MR1232507](#)
- FAN, J. and KOO, J.-Y. (2002). Wavelet deconvolution. *IEEE Trans. Inform. Theory* **48** 734–747. [MR1889978](#)
- FOLLAND, G. B. (1999). *Real Analysis: Modern Techniques and Their Applications*, 2nd ed. Wiley, New York. [MR1681462](#)
- GACH, F., NICKL, R. and SPOKOINY, V. (2013). Spatially adaptive density estimation by localised Haar projections. *Ann. Inst. Henri Poincaré Probab. Stat.* **49** 900–914. [MR3112439](#)
- GINÉ, E. and NICKL, R. (2009). An exponential inequality for the distribution function of the kernel density estimator, with applications to adaptive estimation. *Probab. Theory Related Fields* **143** 569–596. [MR2475673](#)
- GOLDENSHLUGER, A. and LEPSKI, O. (2011). Bandwidth selection in kernel density estimation: Oracle inequalities and adaptive minimax optimality. *Ann. Statist.* **39** 1608–1632. [MR2850214](#)
- GOLDENSHLUGER, A. and LEPSKI, O. (2014). On adaptive minimax density estimation on R^d . *Probab. Theory Related Fields* **159** 479–543. [MR3230001](#)
- GOLUBEV, G. K. (1992). Non-parametric estimation of smooth probability densities. *Probl. Inf. Transm.* **1** 52–62.
- GRAFAKOS, L. (2008). *Classical Fourier Analysis*, 2nd ed. *Graduate Texts in Mathematics* **249**. Springer, New York. [MR2445437](#)
- HALL, P. and MEISTER, A. (2007). A ridge-parameter approach to deconvolution. *Ann. Statist.* **35** 1535–1558. [MR2351096](#)
- HASMINSKII, R. and IBRAGIMOV, I. (1990). On density estimation in the view of Kolmogorov’s ideas in approximation theory. *Ann. Statist.* **18** 999–1010. [MR1062695](#)
- HESSE, C. H. (1995). Deconvolving a density from partially contaminated observations. *J. Multivariate Anal.* **55** 246–260. [MR1370403](#)
- JUDITSKY, A. and LAMBERT-LACROIX, S. (2004). On minimax density estimation on \mathbb{R} . *Bernoulli* **10** 187–220. [MR2046772](#)
- KERKYACHARIAN, G., LEPSKI, O. and PICARD, D. (2001). Nonlinear estimation in anisotropic multi-index denoising. *Probab. Theory Related Fields* **121** 137–170. [MR1863916](#)
- KERKYACHARIAN, G., PHAM NGOC, T. M. and PICARD, D. (2011). Localized spherical deconvolution. *Ann. Statist.* **39** 1042–1068. [MR2816347](#)
- LEPSKI, O. (2013). Multivariate density estimation under sup-norm loss: Oracle approach, adaptation and independence structure. *Ann. Statist.* **41** 1005–1034. [MR3099129](#)
- LEPSKI, O. (2015). Adaptive estimation over anisotropic functional classes via oracle approach. *Ann. Statist.* **43** 1178–1242. [MR3346701](#)
- LEPSKI, O. V. (2018). A new approach to estimator selection. *Bernoulli* **24**. To appear. Available at [arXiv:1603.03934v1](#).
- LEPSKI, O. V. and WILLER, T. (2017). Lower bounds in the convolution structure density model. *Bernoulli* **23** 884–926. [MR3606754](#)
- LOUNICI, K. and NICKL, R. (2011). Global uniform risk bounds for wavelet deconvolution estimators. *Ann. Statist.* **39** 201–231. [MR2797844](#)
- MASRY, E. (1993). Strong consistency and rates for deconvolution of multivariate densities of stationary processes. *Stochastic Process. Appl.* **47** 53–74. [MR1232852](#)

- MEISTER, A. (2009). *Deconvolution Problems in Nonparametric Statistics. Lecture Notes in Statistics* **193**. Springer, Berlin. [MR2768576](#)
- NIKOL'SKII, S. M. (1977). *Priblizhenie Funktsii Mnogikh Peremennykh i Teoremy Vlozheniya* [Approximation of Functions of Several Variables and Imbedding Theorems], 2nd ed., revised and supplemented. Nauka, Moscow (in Russian).
- PENSKY, M. and VIDAKOVIC, B. (1999). Adaptive wavelet estimator for nonparametric density deconvolution. *Ann. Statist.* **27** 2033–2053. [MR1765627](#)
- REBELLES, G. (2016). Structural adaptive deconvolution under L_p -losses. *Math. Methods Statist.* **25** 26–53. [MR3480609](#)
- REYNAUD-BOURET, P., RIVOIRARD, V. and TULEAU-MALOT, C. (2011). Adaptive density estimation: A curse of support? *J. Statist. Plann. Inference* **141** 115–139. [MR2719482](#)
- RIGOLLET, P. (2006). Adaptive density estimation using the blockwise Stein method. *Bernoulli* **12** 351–370. [MR2218559](#)
- RIGOLLET, PH. and TSYBAKOV, A. B. (2007). Linear and convex aggregation of density estimators. *Math. Methods Statist.* **16** 260–280. [MR2356821](#)
- SAMAROV, A. and TSYBAKOV, A. (2007). Aggregation of density estimators and dimension reduction. In *Advances in Statistical Modeling and Inference. Ser. Biostat.* **3** 233–251. World Sci. Publ., Hackensack, NJ. [MR2416118](#)
- STEFANSKI, L. A. (1990). Rates of convergence of some estimators in a class of deconvolution problems. *Statist. Probab. Lett.* **9** 229–235. [MR1045189](#)
- STEFANSKI, L. and CARROLL, R. J. (1990). Deconvoluting kernel density estimators. *Statistics* **21** 169–184. [MR1054861](#)
- YUAN, M. and CHEN, J. (2002). Deconvolving multidimensional density from partially contaminated observations. *J. Statist. Plann. Inference* **104** 147–160. [MR1900523](#)

INSTITUT DE MATHÉMATIQUE DE MARSEILLE
AIX-MARSEILLE UNIVERSITÉ
39, RUE F. JOLIOT-CURIE
13453 MARSEILLE
FRANCE
E-MAIL: oleg.lepski@univ-amu.fr
thomas.willer@univ-amu.fr