Upper functions for L_p-norms of Gaussian random fields

Oleg Lepski

Upper functions for $\mathbb{L}_p$-norms of gaussian random fields

Oleg Lepski

Institut de Mathématique de Marseille
Aix-Marseille Université
39, rue F. Joliot-Curie
13453 Marseille, France
e-mail: oleg.lepski@univ-amu.fr

Abstract: In this paper we are interested in finding upper functions for a collection of random variables $\{\|\xi_{\vec{h}}\|_p, \vec{h} \in \mathrm{H}\}, 1 \le p < \infty$. Here $\xi_{\vec{h}}(x), x \in (-b,b)^d, d \ge 1$ is a kernel-type gaussian random field and $\|\cdot\|_p$ stands for $\mathbb{L}_p$-norm on $(-b,b)^d$. The set H consists of d-variate vector-functions defined on $(-b,b)^d$ and taking values in some countable net in $\mathbb{R}^d_+$. We seek a non-random family $\left\{\Psi_\varepsilon(\vec{h}),\ \vec{h}\in\mathrm{H}\right\}$ such that $\mathbb{E}\{\sup_{\vec{h}\in\mathrm{H}}[\|\xi_{\vec{h}}\|_p - \Psi_\varepsilon(\vec{h})]_+\}^q \le \varepsilon^q,\ q\ge 1$, where $\varepsilon > 0$ is prescribed level.

AMS 2000 subject classifications: Primary 60E15; secondary 62G07, 62G08.
Keywords and phrases: upper function, gaussian random field, metric entropy, Dudley's integral.

1. Introduction

Let $\mathbb{R}^d,\ d\ge 1$, be equipped with Borel σ-algebra $\mathfrak{B}(\mathbb{R}^d)$ and Lebesgue measure ν_d. Put $\widetilde{\mathfrak{B}}(\mathbb{R}^d) = \{B\in\mathfrak{B}(\mathbb{R}^d):\ \nu_d(B)<\infty\}$ and let $(W_B,\ B\in\widetilde{\mathfrak{B}}(\mathbb{R}^d))$ be the white noise with intensity ν_d. Throughout of the paper we will use the following notations. For any $u,v\in\mathbb{R}^d$ the operations and relations u/v, uv, $u\vee v$,$u\wedge v$, $u<v$, $au, a\in\mathbb{R}$, are understood in coordinate-wise sense and $|u|$ stands for euclidian norm of u. All integrals are taken over $\mathbb{R}^d$ unless the domain of integration is specified explicitly. For any real a its positive part is denoted by $(a)_+$ and $\lfloor a\rfloor$ is used for its integer part. For any $\mathbf{n}=(n_1,\dots,n_d)\in\mathbb{N}^d,\ d\ge 1$, $|\mathbf{n}|$ stands for $\sum_{j=1}^d n_i$.

1.1. Collection of random variables.

Let $0<\mathfrak{h}\le e^{-2}$ be fixed number and put $\mathfrak{H}=\{\mathfrak{h}_s, s\in\mathbb{N}\}$, where $\mathfrak{h}_s=e^{-s}\mathfrak{h}$. Denote by $\mathfrak{S}(\mathfrak{h})$ the set of all measurable functions defined on $(-b,b)^d,\ b\in(0,\infty)$, and taking values in $\mathfrak{H}$ and introduce

$$\mathfrak{S}_d(\mathfrak{h}) = \left\{\vec{h}:(-b,b)^d\to\mathfrak{H}^d:\quad \vec{h}(x)=\big(h_1(x),\dots,h_d(x)\big),\ x\in(-b,b)^d,\ \ h_i\in\mathfrak{S}(\mathfrak{H}),\ i=\overline{1,d}\right\}.$$

Let $K:\mathbb{R}^d\to\mathbb{R}$ be fixed. With any $\vec{h}\in\mathfrak{S}_d(\mathfrak{h})$ we associate the function

$$K_{\vec{h}}(t,x)=V_{\vec{h}}^{-1}(x)K\left(\frac{t-x}{\vec{h}(x)}\right),\qquad V_{\vec{h}}(x)=\prod_{i=1}^d h_i(x),\ \ t\in\mathbb{R}^d,\ x\in(-b,b)^d.$$

Following the terminology used in the mathematical statistics we call the function K kernel and the vector-function $\vec{h}$ multi-bandwidth. Moreover, if all coordinates of $\vec{h}$ are the same we will say that corresponding collection is *isotropic*. Otherwise it is called *anisotropic*.

Let H be a given subset of $\mathfrak{S}_d(\mathfrak{h})$ and consider the family

$$\left\{\xi_{\vec{h}}(x)=\int K_{\vec{h}}(t,x)W(\mathrm{d}t),\ \ \vec{h}\in\mathrm{H},\ x\in(-b,b)^d\right\}.$$

We note that $\xi_{\vec{h}}$ is centered gaussian random field on $(-b,b)^d$ with the covariance function

$$V_{\vec{h}}^{-1}(x)V_{\vec{h}}^{-1}(y)\int K\left(\frac{t-x}{\vec{h}(x)}\right)K\left(\frac{t-y}{\vec{h}(y)}\right)\nu_d(\mathrm{d}t),\quad x,y\in(-b,b)^d.$$

Throughout the paper $(\xi_{\vec{h}},\vec{h}\in\mathrm{H})$ is supposed to be defined on the probability space $(\mathfrak{X},\mathfrak{A},\mathbb{P})$ and furthermore $\mathbb{E}$ denotes the expectation with respect to $\mathbb{P}$. Moreover, without further mentioning we will assume that $b\geq 1$.

1.2. Objectives.

Our goal is to find an upper function for the collection of random variables

$$\Lambda_p(\mathrm{H})=\left\{\left\|\xi_{\vec{h}}\right\|_p,\ \ \vec{h}\in\mathrm{H}\right\},\quad 1\leq p\leq\infty,$$

where $\|\cdot\|_p$ stands for $\mathbb{L}_p$-norm on $(-b,b)^d$, that is

$$\|g\|_p=\left(\int_{(-b,b)^d}|g|^p\nu_d(\mathrm{d}x)\right)^{\frac{1}{p}},\quad 1\leq p<\infty,\qquad \|g\|_\infty=\sup_{x\in(-b,b)^d}|g(x)|.$$

More precisely we seek for a *non-random collection* $\left\{\Psi_\varepsilon(\vec{h}),\ \ \vec{h}\in\mathrm{H}\right\}$ such that

$$\mathbb{E}\left\{\sup_{\vec{h}\in\mathrm{H}}\left[\left\|\xi_{\vec{h}}\right\|_p-c\Psi_\varepsilon(\vec{h})\right]_+\right\}^q\leq\varepsilon^q,\quad q\geq 1, \tag{1.1}$$

where $\varepsilon>0$ is a prescribed level and $c>0$ is a numerical constant independent of ε.

Some remarks are in order.

1) Although the upper function as well as the inequality (1.1) can be looked for any level $\varepsilon>0$ we will be obviously interested in small values of ε. In this context (1.1) can be replaced by

$$\limsup_{\varepsilon\to 0}\varepsilon^{-q}\mathbb{E}\left\{\sup_{\vec{h}\in\mathrm{H}}\left[\left\|\xi_{\vec{h}}\right\|_p-c\Psi_\varepsilon(\vec{h})\right]_+\right\}^q<\infty,\quad q\geq 1. \tag{1.2}$$

2) We will see that the upper function $\left\{\Psi_\varepsilon(\vec{h}),\ \ \vec{h}\in\mathrm{H}\right\}$ does not necessarily depend on ε, see, in particular Theorems 2 and 3 below. Typically, in such cases the set H depends on ε or reciprocally the level ε depends on assumptions imposed on the set H. In particular, since $\mathrm{H}\subseteq\mathfrak{S}_d(\mathfrak{h})$ we relate later on the level ε with the extra-parameter $\mathfrak{h}$. We will show that in some important cases $\varepsilon=\varepsilon(\mathfrak{h})$ and $\varepsilon(\mathfrak{h})\to 0$ quite rapidly when $\mathfrak{h}\to 0$. This issue is discussed more in detail in the paragraph preceding Corollary 2.

We will say that the upper function $\Psi_\varepsilon(\cdot)$ is *sharp in order* if (1.2) holds and for some $c_0>0$

$$\liminf_{\varepsilon\to 0}\varepsilon^{-q}\mathbb{E}\left\{\sup_{\vec{h}\in\mathrm{H}}\left[\left\|\xi_{\vec{h}}\right\|_p-c_0\Psi_\varepsilon(\vec{h})\right]_+\right\}^q=\infty,\quad q\geq 1. \tag{1.3}$$

It is worth mentioning that uniform probability and moment bounds for $[\sup_{\theta\in\Theta}\Upsilon(\chi_\theta)]$ in the case where χ_θ is empirical or gaussian process and Υ is a positive functional are a subject of vast literature, see, e.g., Alexander (1984), Talagrand (1994, 2005), Lifshits (1995), van der Vaart and

Wellner (1996), van de Geer (2000), Massart (2000), Bousquet (2002), Giné and Koltchinskii (2006) among many others. Such bounds play an important role in establishing the laws of iterative logarithm and central limit theorems [see, e.g., Alexander (1984) and Giné and Zinn (1984)]. However much less attention was paid to the finding of upper functions. Some asymptotical results can be found in Kalinauskaĭte (1966), Qualls and Watanabe (1972), Bobkov (1988), Shiryaev et al. (2002) and references therein. The inequalities similar to (1.1) was obtained by Egishyants and Ostrovskii (1996), Goldenshluger and Lepski (2011a) and Lepski (2013a,b,c).

The upper functions for $\mathbb{L}_p$-norm of "kernel-type" empirical and gaussian processes was studied in recent papers Goldenshluger and Lepski (2011a) and Lepski (2013a). However the results obtained there allow to study only a bandwidth's collection consisting of constant functions, see discussions after Theorems 1–3 below. To the best of our knowledge the problem of constructing upper functions for the collection parameterized by bandwidths being multivariate (univariate) functions was not studied in the literature.

1.3. Relation to the adaptive estimation.

The evaluation of upper functions has become an important technical tool in different areas of mathematical statistics in particular in the minimax and adaptive minimax estimation. Indeed, all known to the author constructions of adaptive estimators e.g. Lepskii (1991), Barron et al. (1999), Cavalier and Golubev (2006), Goldenshluger and Lepski (2009, 2011b) involve the computation of upper functions for stochastic objects of different kinds. We provide below an explicit expression of the functional Ψ_ε that allows, in particular, to use our results for constructing data-driven procedures in multivariate function estimation. It is important to emphasize that the the collection $\big\{\Psi_\varepsilon\big(\vec{h}\big),\ \ \vec{h}\in \mathrm{H}\big\}$ satisfying (1.1) is not unique and obviously we seek for at least sharp in order upper functions. The latter means that some lower bound results (1.3) should be added to the inequality (1.1), see next paragraph and the discussion after Theorem 1. Note however that the theory of adaptive estimation is equipped with very developed criteria of optimality Lepskii (1991), Tsybakov (1998), Kluchnikoff (2005). Hence, we might expect that the corresponding upper function is sharp in order if its use leads to the construction of optimally adaptive estimators.

1.4. Preliminary observations.

This paragraph is devoted to the discussion about what kind of results we expect to obtain. We provide with upper functions and the inequality (1.1) in some simple cases. We present also a universal lower bound for an upper function and discuss its attainability. Although the proofs of all presented results are straightforward and relatively simple for an interested reader we put them in Section 4.4 of Appendix. Moreover, without further mentioning we will consider here only $p<\infty$ and later on γ_p denotes p-th absolute moment of standard gaussian distribution.

Introductive example. Denote by $\mathfrak{S}^{\mathrm{const}}(\mathfrak{h})=\big\{\vec{h}:\ \ \vec{h}(x)=\vec{\mathrm{h}}\in\mathfrak{H}^d,\ \forall x\in(-b,b)^d\big\}$. Thus, $\mathfrak{S}^{\mathrm{const}}(\mathfrak{h})$ consists of bandwidths which are constants. Put also $\mathfrak{S}^{\mathrm{const}}_{\mathrm{isotr}}(\mathfrak{h})=\big\{\vec{\mathrm{h}}\in\mathfrak{S}^{\mathrm{const}}(\mathfrak{h}):\ \vec{\mathrm{h}}=(h,\dots,h),\ h\in\mathfrak{H}\big\}$ (isotropic case).

For any $p\geq 2$, using the results obtained in Lepski (2013a), Theorem 1, we can assert that (1.1) is satisfied with $\mathrm{H}=\mathfrak{S}^{\mathrm{const}}_{\mathrm{isotr}}(\mathfrak{h})$ and

$$\Psi_\varepsilon\big(\vec{\mathrm{h}}\big)=\Psi\big(\vec{\mathrm{h}}\big):=A_1h^{-d/2},\qquad \varepsilon=\varepsilon(\mathfrak{h})=A_2(q)\mathfrak{h}^{\frac{qd(2-p)}{2p}}\exp\Big\{-2^{-3/2}\mathfrak{h}^{-2d/p}\Big\}. \tag{1.4}$$

Here A_1 et $A_2(q)$ are constants completely determined by K, d, b and p. Note also that Theorem 1 in Lepski (2013a) is proved under condition imposed on the kernel K which is similar to Assumption 3 below.

Remark that $h^{-d/2} = (2b)^{-\frac{1}{p}} \big\|V_{\vec{\mathrm{h}}}^{-\frac{1}{2}}\big\|_p$ for any $\vec{\mathrm{h}} \in \mathfrak{S}^{\mathrm{const}}_{\mathrm{isotr}}(\mathfrak{h})$ and $p \in [1,\infty]$. The following question naturally arises in this context.

How is the upper function on an arbitrary subset of $\mathfrak{S}_d(\mathfrak{h})$ related to the functional $\big\|V_{\vec{h}}^{-\frac{1}{2}}\big\|_p$?

Universal lower bound. Our first goal is to show that an upper function on H can not be "better" in order than $\big\|V_{\vec{h}}^{-\frac{1}{2}}\big\|_p$ whenever $\mathrm{H} \subset \mathfrak{S}_d(\mathfrak{h})$ is considered.

Denote $\mathfrak{S}^*_{d,p}(\mathfrak{h}) = \big\{\vec{h} \in \mathfrak{S}_d(\mathfrak{h}) : \ \big\|V_{\vec{h}}^{-\frac{1}{2}}\big\|_p < \infty\big\}$. The following assertion is true: for any $p \geq 1$

$$\mathbb{E}\Big\{\Big[\big\|\xi_{\vec{h}}\big\|_p - 2^{-4}(\gamma_p)^{\frac{1}{p}}\|K\|_2\big\|V_{\vec{h}}^{-\frac{1}{2}}\big\|_p\Big]_+\Big\}^q \geq B_1\mathfrak{h}^{-\frac{dq}{2}} \geq B_1 e^{dq}, \quad \forall \vec{h} \in \mathfrak{S}^*_{d,p}(\mathfrak{h}), \tag{1.5}$$

where B_1 depends only on K, d, b, q and p and its explicit expression can be found in Section 4.4.

Combining (1.4) and (1.5) we can assert that $\Psi\big(\vec{\mathrm{h}}\big) = h^{-d/2}$ is sharp in order on $\mathfrak{S}^{\mathrm{const}}_{\mathrm{isotr}}(\mathfrak{h})$ if $\mathfrak{h} \to 0$ and $p \geq 2$. More generally, we will show that $\prod_{j=1}^d \mathrm{h}_j^{-\frac{1}{2}}$ is a sharp in order upper function on $\mathfrak{S}^{\mathrm{const}}(\mathfrak{h})$, see discussion after Corollary 1.

"Pointwise" upper bound and its trivial consequence. Let $\mathrm{H} = \{\vec{h}\}$, where $\vec{h} \in \mathfrak{S}^*_{d,p}(\mathfrak{h})$ is a given multi-bandwidth. Introduce

$$\sigma^2_p\big(\vec{h}\big) = \sup_{\vartheta \in \mathbb{B}_{s,d}} \int_{\mathbb{R}^d} \bigg(\int_{(-b,b)^d} \vartheta(x) K_{\vec{h}}(t,x)\nu_d(\mathrm{d}x)\bigg)^2 \nu_d(\mathrm{d}t),$$

where, $\mathbb{B}_{\mathrm{s},d} = \{\vartheta : (-b,b)^d \to \mathbb{R} : \ \|\vartheta\|_s \leq 1\}$ and $1/s = 1 - 1/p$.

The following is true: for any $\vec{h} \in \mathfrak{S}^*_{d,p}(\mathfrak{h})$ and any $p \geq 1$

$$\mathbb{E}\Big\{\Big[\big\|\xi_{\vec{h}}\big\|_p - \big((\gamma_p)^{\frac{1}{p}}\|K\|_2 + \sqrt{2}\big)\big\|V_{\vec{h}}^{-\frac{1}{2}}\big\|_p\Big]_+\Big\}^q \leq B_2\sigma^q_p\big(\vec{h}\big)e^{-\sigma_p^{-2}\left(\vec{h}\right)\left\|V_{\vec{h}}^{-\frac{1}{2}}\right\|_p^2}, \quad q \geq 1, \tag{1.6}$$

where B_2 depends only on K, d, b, q and p and its explicit expression can be found in Section 4.4.

Let now $p \in [1,2]$. Using the computations similar to whose led to the bound (3.53) in Section 3.3.4 one can assert that there exists B_3 completely determined by K, d, b and p such that

$$\sigma_p\big(\vec{h}\big) \leq B_3, \quad \forall \vec{h} \in \mathfrak{S}_d(\mathfrak{h}).$$

It yields together with (1.6)

$$\mathbb{E}\Big\{\Big[\big\|\xi_{\vec{h}}\big\|_p - \big((\gamma_p)^{\frac{1}{p}}\|K\|_2 + \sqrt{2}\big)\big\|V_{\vec{h}}^{-\frac{1}{2}}\big\|_p\Big]_+\Big\}^q \leq B_4 e^{-B_5\mathfrak{h}^{-d/2}}, \quad q \geq 1.$$

Let H be a finite set and suppose that $\varepsilon^q(\mathfrak{h}) := \mathrm{card}(\mathrm{H})B_4 e^{-B_5\mathfrak{h}^{-d/2}} \to 0, \ \ \mathfrak{h} \to 0$. Then, in view of (1.5) we assert that $\big\|V_{\vec{h}}^{-\frac{1}{2}}\big\|_p$ is the sharp in order upper function with level $\varepsilon(\mathfrak{h})$ if $p \in [1,2]$.

Concluding remarks. Putting together (1.5) and the statement of Theorem 1 below we can assert that any sharp in order upper function must satisfy

$$\Big\|V_{\vec{h}}^{-\frac{1}{2}}\Big\|_p \lesssim \Psi_\varepsilon\big(\vec{h}\big) \lesssim \Big\|\sqrt{\big|\ln\big(\varepsilon V_{\vec{h}}\big)\big|}V_{\vec{h}}^{-\frac{1}{2}}\Big\|_p, \quad \vec{h} \in \mathrm{H}, \tag{1.7}$$

whenever $\mathrm{H} \in \mathfrak{S}_d(\mathfrak{h})$ is considered.

We present sufficient conditions imposed on H under which $\big\|V_{\vec{h}}^{-\frac{1}{2}}\big\|_p$ is the sharp in order upper function, see Remarks 2 and 3 after Corollary 1 and Theorem 3 respectively. We will see that the latter condition can be checked on rather huge subsets of $\mathfrak{S}_d(\mathfrak{h})$, Section 2.4. However the finding of the necessary condition remains an open problem. The interesting question arising in this context is the right side of the inequality (1.7) tight? The following assertion answers partially on this question.

One can construct $H \subset \mathfrak{S}_d(\mathfrak{h})$ such that

$$\lim_{c\to 0}\lim_{\varepsilon\to 0}\varepsilon^{-q}\mathbb{E}\bigg\{\sup_{\vec{h}\in H}\Big[\big\|\xi_{\vec{h}}\big\|_p - c\Big\|\sqrt{\big|\ln\big(\varepsilon V_{\vec{h}}\big)\big|}V_{\vec{h}}^{-\frac{1}{2}}\Big\|_p\bigg\}^q = \infty, \quad q \geq 1. \tag{1.8}$$

We have no place here in order to prove this result since it takes tens pages. We only mention that the proof of (1.8) is "statistical", cf. Section 1.3. In particular, the description of the set H can be found in the recent paper Lepski (2014), Proposition 2, where it is used in order to prove the optimality of the proposed adaptive procedure. It is important to emphasize that its construction is similar to one of Section 2.4 below. The proof of (1.8) is also based on the lower bound for minimax risks over anisotropic Nikolskii classes established in Kerkyacharian et al. (2008).

1.5. Organization of the paper.

In Section 2 we present three constructions of upper functions and prove for them an inequality of type (1.1), Theorems 1–3. Moreover, in Subsection 2.4 we discuss the example of the bandwidth collection satisfying the assumptions of Theorem 2. Section 3 contains proofs of Theorems 1–3; proofs of auxiliary results are relegated to the Appendix.

2. Main results.

Throughout the paper we will consider the collections $\Lambda(\mathrm{H})$ with K satisfying one of Assumptions 1– 3 indicated below. The parameters $a \geq 1$ and $L > 0$ used there are supposed to be fixed.

2.1. Anisotropic case. First construction.

Assumption 1. *$supp(K) \subset [-a,a]^d$ and*

$$|K(s) - K(t)| \leq L|s-t|, \quad \forall s,t \in \mathbb{R}^d.$$

Introduce $\mathfrak{S}_{d,p}(\mathfrak{h}) = \Big\{\vec{h} \in \mathfrak{S}_d(\mathfrak{h}) : \quad \Big\|\sqrt{\big|\ln\big(V_{\vec{h}}\big)\big|}V_{\vec{h}}^{-\frac{1}{2}}\Big\|_p < \infty\Big\}$. For any $\vec{h} \in \mathfrak{S}_{d,p}(\mathfrak{h})$ and any $0 < \varepsilon \leq e^{-2}$ define

$$\psi_\varepsilon(\vec{h}) = C_1\Big\|\sqrt{\big|\ln\big(\varepsilon V_{\vec{h}}\big)\big|}V_{\vec{h}}^{-\frac{1}{2}}\Big\|_p,$$

where $C_1 = 2\big(q \vee \big[p1\{p<\infty\} + 1\{p=\infty\}\big]\big) + 2\sqrt{2d}\Big[\sqrt{\pi} + \|K\|_2\Big(\sqrt{\big|\ln\big(4bL\|K\|_1\big)\big|} + 1\Big)\Big]$.

Theorem 1. *Let $q \geq 1$, $p \in [1,\infty]$, be fixed and let* H *be an arbitrary countable subset of $\mathfrak{S}_{d,p}(\mathfrak{h})$. Suppose also that Assumption 1 is fulfilled. Then*

$$\mathbb{E}\bigg\{\sup_{\vec{h}\in \mathrm{H}}\Big[\big\|\xi_{\vec{h}}\big\|_p - \psi_\varepsilon(\vec{h})\Big]_+\bigg\}^q \leq \big[C_3\varepsilon\big]^q, \quad \forall \mathfrak{h}, \varepsilon \in \big(0, e^{-2}\big),$$

where $C_3 = C_3(\tilde{q}, p)1\{p < \infty\} + C_3(q, 1)1\{p = \infty\}$, $\tilde{q} = (q/p) \vee 1$ *and*

$$C_3(a, b) = (4b)^{\frac{d}{b}} \left[2a \int_0^\infty z^{a-1} \exp\left(-\frac{z^{\frac{2}{b}}}{8\|K\|_2^2}\right) \mathrm{d}z\right]^{\frac{1}{ab}}, \; a, b \geq 1.$$

Remark 1. *We consider only countable subsets of* $\mathfrak{S}_{d,p}(\mathfrak{h})$ *in order not to discuss the measurability issue. Actually the statement of the theorem remains valid for any subset providing the measurability of the corresponding supremum. It explains why the upper function* ψ_ε *as well as the constants* C_1 *and* C_3 *are independent of the choice of* H.

The advantage of the result presented in Theorem 1 is that it is proved without any condition imposed on the set of bandwidths. Moreover, as it follows from (1.8) this bound can not be improved in order than an arbitrary H is considered. On the other hand for a particular choice of H the obtained result can be essentially improved.

Indeed, let $p \geq 2$ and consider $\mathrm{H} = \mathfrak{S}_{\text{isotr}}^{\text{const}}(\mathfrak{h})$. In this case, the found upper function is given by

$$\sqrt{|\ln(\varepsilon)| + d|\ln(h)|}\; h^{-\frac{d}{2}} > \sqrt{|\ln(\varepsilon)|}\; h^{-\frac{d}{2}}.$$

Choose for instance $\mathfrak{h} = \big(4q|\ln(\varepsilon)|\big)^{-\frac{p}{2d}}$ we deduce from (1.4) and (1.5) that $h^{-\frac{d}{2}}$ is the sharp in order upper function. Thus, the upper function given in Theorem 1 is not optimal.

The problem we address now consists in finding subsets of $\mathfrak{S}_d(\mathfrak{h})$ for which upper functions, more precise than one presented in Theorem 1, can be found.

2.2. Anisotropic case. Functional classes of bandwidths.

Put for any $\vec{h} \in \mathfrak{S}_d(\mathfrak{h})$ and any multi-index $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{N}^d$

$$\Lambda_{\mathbf{s}}\big[\vec{h}\big] = \cap_{j=1}^d \Lambda_{s_j}\big[h_j\big], \qquad \Lambda_{s_j}\big[h_j\big] = \big\{x \in (-b, b)^d : \; h_j(x) = \mathfrak{h}_{s_j}\big\}.$$

Let $\tau \in (0,1)$ and $\mathcal{L} > 0$ be given constants. Define

$$\mathbb{H}_d(\tau, \mathcal{L}) = \bigg\{\vec{h} \in \mathfrak{S}_d(\mathfrak{h}) : \; \sum_{\mathbf{s} \in \mathbb{N}^d} \nu_d^\tau\Big(\Lambda_{\mathbf{s}}\big[\vec{h}\big]\Big) \leq \mathcal{L}\bigg\}.$$

A simple example of the subset of $\mathbb{H}_d(\tau, \mathcal{L})$ is $\mathfrak{S}^{\text{const}}(\mathfrak{h}_\varepsilon)$, since obviously $\mathfrak{S}^{\text{const}}(\mathfrak{h}_\varepsilon) \subset \mathbb{H}_d(\tau, \mathcal{L})$ for any $\tau \in (0,1)$ and $\mathcal{L} \geq (2b)^{d\tau}$. A quite sophisticated construction is postponed to Section 2.4.

Put $\mathbb{N}_p^* = \big\{[p]+1, [p]+2, \dots\big\}$ and introduce for any $\mathcal{A} \geq \mathfrak{h}^{-\frac{d}{2}}$

$$\mathbb{B}(\mathcal{A}) = \bigcup_{r \in \mathbb{N}_p^*} \mathbb{B}_r(\mathcal{A}), \qquad \mathbb{B}_r(\mathcal{A}) = \bigg\{\vec{h} \in \mathfrak{S}_d(\mathfrak{h}) : \; \Big\|V_{\vec{h}}^{-\frac{1}{2}}\Big\|_{\frac{rp}{r-p}} \leq \mathcal{A}\bigg\}.$$

Note that introduced in the previous section $\mathfrak{S}_{d,p}(\mathfrak{h}) \subset \mathbb{B}(\mathcal{A})$ for any $\mathcal{A}$. The following notations related to the functional class $\mathbb{B}(\mathcal{A})$ will be exploited in the sequel. For any $\vec{h} \in \mathbb{B}(\mathcal{A})$ define

$$\mathbb{N}_p^*(\vec{h}, \mathcal{A}) = \mathbb{N}_p^* \cap \big[r_{\mathcal{A}}(\vec{h}), \infty\big), \qquad r_{\mathcal{A}}(\vec{h}) = \inf\big\{r \in \mathbb{N}_p^* : \; \vec{h} \in \mathbb{B}_r(\mathcal{A})\big\}. \tag{2.1}$$

Obviously $r_{\mathcal{A}}\big(\vec{h}\big) < \infty$ for any $\vec{h} \in \mathbb{B}(\mathcal{A})$.

In this section we will be interested in finding an upper function when H is an arbitrary subset of $\mathbb{H}_d(\tau, \mathcal{L}, \mathcal{A}) := \mathbb{H}_d(\tau, \mathcal{L}) \cap \mathbb{B}(\mathcal{A})$.

The following relation between the parameters $\mathfrak{h}, \mathcal{A}$ and τ is supposed to be held throughout of this section.

$$d\ln\ln(\mathcal{A}) \leq 2\sqrt{2(1-\tau)|\ln(\mathfrak{h})|} - d\ln(4). \tag{2.2}$$

For any $\vec{h} \in \mathbb{B}(\mathcal{A})$ define

$$\psi\big(\vec{h}\big) = \inf_{r \in \mathbb{N}^*_p(\vec{h})} C_2(r,\tau,\mathcal{L}) \Big\| V_{\vec{h}}^{-\frac{1}{2}} \Big\|_{\frac{rp}{r-p}},$$

where, $\mathbb{N}^*_p(\vec{h})$ is defined in (2.1) and the quantity $C_2(r,\tau,\mathcal{L}), \tau \in (0,1),\ \mathcal{L} > 0$, is given in Section 3.2.2. Its expression is rather cumbersome and it is why we do not present it right now. Here we only mentioned that $C_2(r,\tau,\mathcal{L})$ is finite for any given r but $\lim_{r\to\infty} C_2(r,\tau,\mathcal{L}) = \infty$.

Note also that the condition $\vec{h} \in \mathbb{B}(\mathcal{A})$ guarantees the $\psi\big(\vec{h}\big) < \infty$ for any $\vec{h}$.

Assumption 2. *There exists $\mathcal{K} : \mathbb{R} \to \mathbb{R}$ such that supp$(\mathcal{K}) \subset [-a, a]$ and*

$$\begin{aligned} &\textbf{(i)} \qquad && |\mathcal{K}(s) - \mathcal{K}(t)| \leq L|s-t|, \quad \forall s,t \in \mathbb{R}; \\ &\textbf{(ii)} \qquad && K(x) = \prod_{i=1}^{d} \mathcal{K}(x_i), \quad \forall x = (x_1,\dots,x_d) \in \mathbb{R}^d. \end{aligned}$$

Theorem 2. *Let $q \geq 1$, $1 \leq p < \infty$, $\tau \in (0,1)$, $\mathcal{L} > 0$ and $\mathcal{A} \geq \mathfrak{h}^{-\frac{d}{2}}$ be fixed and let H be an arbitrary countable subset of $\mathbb{H}_d\big(\tau,\mathcal{L},\mathcal{A}\big)$.*

Then for any $\mathcal{A}$, $\mathfrak{h}$ and τ satisfying (2.2) and K satisfying Assumption 2,

$$\mathbb{E}\bigg\{ \sup_{\vec{h}\in\mathrm{H}} \Big[\big\|\xi_{\vec{h}}\big\|_p - \psi\big(\vec{h}\big) \Big]_+ \bigg\}^q \leq \Big[C_4 \mathcal{A} e^{-e^{2\sqrt{2d|\ln(\mathfrak{h})|}}} \Big]^q, \quad \forall \mathfrak{h} \in \big(0, e^{-2}\big),$$

where C_4 depends on $\mathcal{K}, p, q, b$ and d only and its explicit expression can be found in Section 3.2.2.

The statement of the theorem remains valid for any subset providing the measurability of the corresponding supremum. It explains, in particular, why the upper function $\psi\big(\vec{h}\big)$ is independent of the choice of H and completely determined by the parameters τ, $\mathcal{L}$ and $\mathcal{A}$. It is worth noting that unlike Theorem 1 whose proof is relatively standard the proof of Theorem 2 is rather long and tricky.

Considering classes $\mathbb{H}_d\big(\tau,\mathcal{L},\mathcal{A}\big)$ we are obviously interested in large values of $\mathcal{A}$ since the larger $\mathcal{A}$ is the weaker restriction on the class is imposed. In this context the parameters $\mathfrak{h}$ and $\mathcal{A}$ should be somehow related. Let us discuss one of possible choices of these parameters.

Choose $\mathfrak{h} = \mathfrak{h}_\varepsilon := e^{-\sqrt{|\ln(\varepsilon)|}}$, $\mathcal{A} = \mathcal{A}_\varepsilon := e^{\ln^2(\varepsilon)}$. This yields

$$\lim_{\varepsilon\to 0} \varepsilon^{-a} \mathcal{A}_\varepsilon e^{-e^{2\sqrt{2d|\ln(\mathfrak{h}_\varepsilon)|}}} = 0, \quad \forall a > 0,$$

and moreover, for any $\tau \in (0,1)$ there exist $\varepsilon_0(\tau)$ such that for all $\varepsilon \leq \varepsilon_0(\tau)$ the relation (2.2) is fulfilled. In view of these remarks we come to the following corollary of Theorem 2.

Corollary 1. *Let the assumptions of Theorem 2 hold and let $\mathfrak{h} = \mathfrak{h}_\varepsilon$ and $\mathcal{A} = \mathcal{A}_\varepsilon$ Then for any $\tau \in (0,1)$ and any $q \geq 1$ one can find $\varepsilon(\tau,q)$ such that for any $\varepsilon \leq \varepsilon(\tau,q)$*

$$\mathbb{E}\bigg\{ \sup_{\vec{h}\in\mathrm{H}} \Big[\big\|\xi_{\vec{h}}\big\|_p - \psi\big(\vec{h}\big) \Big]_+ \bigg\}^q \leq \big\{(C_3 + C_4)\varepsilon\big\}^q.$$

The assertion of the corollary can be of course obtained for another choice of the parameters $\mathcal{A}$ and $\mathfrak{h}$. Our choice is dictated by the following reason: $\mathfrak{h}_\varepsilon$ tends to zero rather slowly (slower than polynomial decay) while $\mathcal{A}_\varepsilon$ increases to infinity faster than polynomially in ε. The both restrictions are heavily exploited for the construction of adaptive statistical procedures.

Remark 2. *Let* $\mathrm{H} \subset \mathbb{H}_d\big(\tau, \mathcal{L}, \mathcal{A}_\varepsilon\big)$ *be such that there exists a constant* $\Upsilon > 0$ *independent on* ε *for which*

$$\sup_{\vec{h}\in\mathrm{H}} \psi(\vec{h}) \Big\| V_{\vec{h}}^{-\frac{1}{2}} \Big\|_p^{-1} \leq \Upsilon. \tag{2.3}$$

Taking together the statement of Corollary 1, (1.5) and (2.3) we can assert that $\big\| V_{\vec{h}}^{-\frac{1}{2}} \big\|_p$ *is the sharp in order upper function.*

Let $\mathrm{H} \subset \mathfrak{S}^{\mathrm{const}}(\mathfrak{h}_\varepsilon)$. Since obviously $\mathfrak{S}^{\mathrm{const}}(\mathfrak{h}_\varepsilon) \subset \mathbb{H}_d(\tau, \mathcal{L})$ for any $\tau \in (0,1)$ and $\mathcal{L} \geq (2b)^{d\tau}$ we first assert that $\mathrm{H} \subset \mathbb{H}_d(\tau, \mathcal{L})$. Next, suppose that

$$V_{\vec{\mathrm{h}}} \geq (2b)^{\frac{d}{p}} \mathcal{A}_\varepsilon^{-2}, \quad \forall \vec{\mathrm{h}} \in \mathrm{H}. \tag{2.4}$$

Then, $\mathrm{H} \in \mathbb{B}(\mathcal{A}_\varepsilon)$ and $\mathbb{N}_p^*\big(\vec{h}, \mathcal{A}_\varepsilon\big) = \mathbb{N}_p^*$ for any $\vec{\mathrm{h}} \in \mathrm{H}$. It yields

$$\psi(\vec{\mathrm{h}}) = V_{\vec{\mathrm{h}}}^{-\frac{1}{2}} (2b)^{\frac{d}{p}} \inf_{r\in\mathbb{N}_p^*} C_2(r, \tau, \mathcal{L}), \quad \vec{\mathrm{h}} \in \mathrm{H}.$$

We conclude that (2.3) is fulfilled and, therefore, $V_{\vec{\mathrm{h}}}^{-\frac{1}{2}}$ is the sharp in order upper function for any choice of H satisfying (2.4).

Another interesting question concerns the "sharpness" of the upper function $\psi(\vec{h})$ when H does not satisfy (2.3). The following result, similar (1.8), can be deduced from recent results obtained in Lepski (2014), Proposition 2. One can construct $H \subset \mathbb{H}_d\big(\tau, \mathcal{L}, \mathcal{A}_\varepsilon\big)$ such that

$$\lim_{c\to 0} \lim_{\varepsilon\to 0} \varepsilon^{-q} \mathbb{E}\bigg\{ \sup_{\vec{h}\in H} \Big[\big\|\xi_{\vec{h}}\big\|_p - c\psi(\vec{h}) \Big\} ^q = \infty, \quad q \geq 1. \tag{2.5}$$

It is impossible to compare upper functions found in Theorems 1 and 2 when an arbitrary subset of $\mathbb{H}_d(\tau, \mathcal{L}, \mathcal{A})$ is considered. However they can be easily combined in such a way that the obtained upper function is smaller that both of them. Indeed, set $\Psi_\varepsilon\big(\vec{h}\big) = \psi_\varepsilon\big(\vec{h}\big) \wedge \psi\big(\vec{h}\big)$. We have

$$\bigg\{ \sup_{\vec{h}\in\mathrm{H}} \Big[\big\|\xi_{\vec{h}}\big\|_p - \Psi_\varepsilon(\vec{h}) \Big]_+ \bigg\}^q \leq \bigg\{ \sup_{\vec{h}\in\mathrm{H}} \Big[\big\|\xi_{\vec{h}}\big\|_p - \psi_\varepsilon(\vec{h}) \Big]_+ \bigg\}^q + \bigg\{ \sup_{\vec{h}\in\mathrm{H}} \Big[\big\|\xi_{\vec{h}}\big\|_p - \psi(\vec{h}) \Big]_+ \bigg\}^q.$$

Corollary 2. *Let the assumptions of Theorem 2 hold and let* $\mathfrak{h} = \mathfrak{h}_\varepsilon$ *and* $\mathcal{A} = \mathcal{A}_\varepsilon$ *Then for any* $\tau \in (0,1)$ *and any* $q \geq 1$ *one can find* $\varepsilon(\tau, q)$ *such that for any* $\varepsilon \leq \varepsilon(\tau, q)$

$$\mathbb{E}\bigg\{ \sup_{\vec{h}\in\mathrm{H}} \Big[\big\|\xi_{\vec{h}}\big\|_p - \Psi_\varepsilon\big(\vec{h}\big) \Big]_+ \bigg\}^q \leq \big\{(C_3 + C_4)\varepsilon\big\}^q.$$

2.3. Isotropic case.

In this section we will suppose that $\vec{h}(\cdot) = \big(h(\cdot), \ldots, h(\cdot)\big)$ and consider the case $p \in [1,2]$. We will show that under these restrictions the result similar to the one of in Theorem 2 can be proved without any condition imposed on the set of bandwidths.

Note that in the isotropic case $V_{\vec{h}}(\cdot) = h^d(\cdot)$ and introduce the following notations.

Set $\mathfrak{S}^{\text{isotr}}_{d,p}(\mathfrak{h}) = \cup_{r\in\mathbb{N}^*, r>d}\Big\{\vec{h} \in \mathfrak{S}_d(\mathfrak{h}) : \ \big\|h^{-\frac{d}{2}}\big\|_{p+\frac{1}{r}} < \infty\Big\}$ and define

$$\psi^*\big(\vec{h}\big) = \inf_{r\in\mathbb{N}^*, r>d} C_2^*(r)\Big\|h^{-\frac{d}{2}}\Big\|_{p+\frac{1}{r}}, \quad \vec{h} \in \mathfrak{S}^{\text{isotr}}_{d,p}(\mathfrak{h}),$$

where the explicit expression of $C_2^*(r)$ is given in Section 3.3.1.

Assumption 3. *supp*$(K) \subset [-a,a]^d$ *and for any* $\mathbf{n} \in \mathbb{N}$ *such that* $|\mathbf{n}| \le \lfloor d/2 \rfloor + 1$

$$|D^{\mathbf{n}}K(s) - D^{\mathbf{n}}K(t)| \le L|s-t|, \ \ \forall s,t \in \mathbb{R}^d, \quad D^{\mathbf{n}} = \frac{\partial^{|\mathbf{n}|}}{\partial y_1^{n_1}\cdots \partial y_k^{n_d}}.$$

Theorem 3. *Let $q \ge 1$, $p \in [1,2]$, be fixed and suppose that Assumption 3 is fulfilled.*

Let H *be an arbitrary countable subset of* $\mathfrak{S}^{isotr}_{d,p}(\mathfrak{h})$. *Then,*

$$\mathbb{E}\bigg\{\sup_{\vec{h}\in\mathrm{H}}\Big[\big\|\xi_{\vec{h}}\big\|_p - \psi^*\big(\vec{h}\big)\Big]_+\bigg\}^q \le \Big(C_5 e^{\mathfrak{h}^{-\frac{d}{2}}}\Big)^q, \quad \forall \mathfrak{h} \le e^{-2},$$

where C_5 depends on K,p,q,b and d only and its explicit expression can be found in Section 3.3.1.

Coming back to the example of H consisting of constant functions we conclude that Theorem 3 generalizes the result given by Theorem 2 when $p \in [1,2]$. Indeed, we do not require here the finiteness of the set in which the bandwidth takes its values.

Although the proof of the theorem is based upon the same approach, which is applied for proving Theorem 2, it requires to use quite different arguments. Both assumptions isotropy and $p \in [1,2]$ are crucial for deriving the statement of Theorem 3, see Section 3.3.3 for details.

Remark 3. *In view of (1.5), the condition*

$$\sup_{\vec{h}\in\mathrm{H}} \psi^*(\vec{h})\Big\|h^{-\frac{d}{2}}\Big\|_p^{-1} \le \Upsilon$$

with some $\Upsilon > 0$ independent of $\mathfrak{h}$, guarantees that $\big\|h^{-\frac{d}{2}}\big\|_p$ is the sharp in order upper function on $\mathrm{H} \subset \mathfrak{S}^{isotr}_{d,p}(\mathfrak{h})$ *when $\mathfrak{h} \to 0$.*

Also, combining the results of Theorems 1 and 3 we arrive to the following assertion.

Corollary 3. *Let assumptions of Theorem 3 hold and choose $\mathfrak{h} = \mathfrak{h}_\varepsilon$ Then,*

$$\mathbb{E}\bigg\{\sup_{\vec{h}\in\mathrm{H}}\Big[\big\|\xi_{\vec{h}}\big\|_p - \psi_\varepsilon\big(\vec{h}\big) \wedge \psi^*\big(\vec{h}\big)\Big]_+\bigg\}^q \le \Big([C_3 + C_5]\varepsilon\Big)^q, \quad \forall \varepsilon \in \big(0, e^{-2}\big].$$

2.4. Example of the functional class $\mathbb{H}_d(\boldsymbol{\tau}, \boldsymbol{\mathcal{L}}, \boldsymbol{\mathcal{A}})$.

Le $(\mathbf{e}_1, \dots, \mathbf{e}_d)$ denote the canonical basis of $\mathbb{R}^d$. For function $g : \mathbb{R}^d \to \mathbb{R}^1$ 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

$$\Delta_{u,j}g(x) = g(x + u\mathbf{e}_j) - g(x), \quad j = 1, \dots, d.$$

By induction, the k-th order difference operator with step size u in direction of the variable x_j is defined as

$$\Delta^k_{u,j}g(x) = \Delta_{u,j}\Delta^{k-1}_{u,j}g(x) = \sum_{l=1}^{k}(-1)^{l+k}\binom{k}{l}\Delta_{ul,j}g(x). \tag{2.6}$$

Definition 1. *For given vectors $\vec{r}=(r_1,\dots,r_d)$, $r_j\in[1,\infty]$, $\vec{\beta}=(\beta_1,\dots,\beta_d)$, $\beta_j>0$, and $\vec{L}=(L_1,\dots,L_d)$, $L_j>0$, $j=1,\dots,d$, we say that function $g:\mathbb{R}^d\to\mathbb{R}^1$ belongs to the anisotropic Nikolskii class $\mathbb{N}_d\big(\vec{\beta},\vec{r},\vec{L}\big)$ if*

(i) $\|g\|_{r_j,\mathbb{R}^d}\leq L_j$ *for all* $j=1,\dots,d$;

(ii) *for every $j=1,\dots,d$ there exists natural number $k_j>\beta_j$ such that*

$$\Big\|\Delta^{k_j}_{u,j}g\Big\|_{r_j,\mathbb{R}^d}\leq L_j|u|^{\beta_j},\quad \forall u\in\mathbb{R},\quad \forall j=1,\dots,d. \tag{2.7}$$

Let ℓ be an arbitrary integer number, and let $w:\mathbb{R}\to\mathbb{R}$ be a compactly supported function satisfying $w\in\mathbb{C}^1(\mathbb{R})$. Put

$$w_\ell(y)=\sum_{i=1}^{\ell}\binom{\ell}{i}(-1)^{i+1}\frac{1}{i}w\Big(\frac{y}{i}\Big),\qquad K(t)=\prod_{j=1}^{d}w_\ell(t_j),\quad t=(t_1,\dots,t_d).$$

Although it will not be important for our considerations here we note nevertheless that K satisfies Assumption 2 with $\mathcal{K}=w_\ell$.

Let $\varepsilon,\mathfrak{h}\in\big(0,e^{-2}\big]$ be fixed and set $\frac{1}{\beta}=\sum_{i=1}^d\frac{1}{\beta_i}$, $\frac{1}{\upsilon}=\sum_{i=1}^d\frac{1}{r_i\beta_i}$. For any $j=1,\dots,d$ let $S_\varepsilon(j)\in\mathbb{N}^*$ be defined from the relation

$$e^{-1}\varepsilon^{\frac{2\beta}{(2\beta+1)\beta_j}}<\mathfrak{h}e^{-S_\varepsilon(j)}\leq\varepsilon^{\frac{2\beta}{(2\beta+1)\beta_j}}. \tag{2.8}$$

Without loss of generality we will assume that ε is sufficiently small in order to provide the existence of $S_\varepsilon(j)$ for any j. Put also

$$\mathfrak{H}^{(j)}_\varepsilon=\{\mathfrak{h}_s=\mathfrak{h}e^{-s}, s\in\mathbb{N},\ s\geq S_\varepsilon(j)\},\qquad \mathfrak{H}_\varepsilon=\mathfrak{H}^{(1)}_\varepsilon\times\cdots\times\mathfrak{H}^{(d)}_\varepsilon$$

and introduce for any $x\in(-b,b)^d$ and any $f\in\mathbb{N}_d(\vec{\beta},\vec{r},\vec{L})$

$$\vec{h}_f(x)=\arg\inf_{\vec{h}\in\mathfrak{H}_\varepsilon}\bigg[\Big|\int K_{\vec{h}}(t-x)f(t)\mathrm{d}t-f(x)\Big|+\varepsilon V_{\vec{h}}^{-\frac{1}{2}}\bigg],\quad V_{\vec{h}}=\prod_{i=1}^{d}h_i.$$

Define finally $\mathrm{H}=\Big\{\vec{h}_f,\ f\in\mathbb{N}_d(\vec{\beta},\vec{r},\vec{L})\Big\}$.

Proposition 1. *Let $\vec{\beta}\in(0,\ell]^d$, $\vec{r}\in[1,p]^d$ and $\vec{L}\in(0,\infty)^d$ be given.*
1) For any $\tau\in(0,1)$ there exists $\mathcal{L}>0$ such that

$$\Big\{\vec{h}_f,\ f\in\mathbb{N}_d(\vec{\beta},\vec{r},\vec{L})\Big\}\subset\mathbb{H}_d(\tau,\mathcal{L}).$$

2) If additionally $\upsilon(2+1/\beta)>p$ then there exists $C>0$ such that

$$\Big\{\vec{h}_f,\ f\in\mathbb{N}_d(\vec{\beta},\vec{r},\vec{L})\Big\}\subset\mathbb{B}\big(C\varepsilon^{-\frac{1}{2\beta+1}}\big).$$

The explicit expression for the constants $\mathcal{L}$ and C can be found in the proof of the proposition which is postponed to Appendix.

The condition $\upsilon(2+1/\beta)>p$ appeared in the second assertion of the proposition is known as the dense zone in adaptive minimax estimation over the collection of anisotropic classes of smooth functions on $\mathbb{R}^d$, see Goldenshluger and Lepski (2014).

3. Proof of Theorems 1–3

The proofs of these theorems are based on several auxiliary results, which for the citation convenience are formulated in Lemmas 1 and 2 below.

Furthermore, for any totaly bounded metric space $(\mathfrak{T}, \varrho)$ we denote by $\mathfrak{E}_{\varrho,\mathfrak{T}}(\delta),\ \delta > 0$, the δ-entropy of $\mathfrak{T}$ measured in ϱ, i.e. the logarithm of the minimal number of ϱ-balls of radius $\delta > 0$ needed to cover $\mathfrak{T}$.

$\mathbf{1^0}$. The results formulated in Lemma 1 can be found in Talagrand (1994), Proposition 2.2, and Lifshits (1995), Theorems 14.1 and 15.2.

Lemma 1. *Let $(Z_t,\ t \in \mathbb{T})$ be a centered, bounded on $\mathbb{T}$, gaussian random function.*

I) For any $u > 0$

$$\mathbb{P}\Big\{\sup_{t\in\mathbb{T}} Z_t \geq \mathbb{E}\Big(\sup_{t\in\mathbb{T}} Z_t\Big) + u\Big\} \leq e^{-\frac{u^2}{2\sigma^2}},$$

where $\sigma^2 = \sup_{t\in\mathbb{T}} \mathbb{E}\big(Z_t^2\big)$.

II) Let $\mathbb{T}$ be equipped with intrinsic semi-metric $\rho^2(t,t') := \mathbb{E}\left(Z_t - Z_{t'}\right)^2,\ t,t' \in \mathbb{T}$. Then

$$\mathbb{E}\Big(\sup_{t\in\mathbb{T}} Z_t\Big) \leq D_{\mathbb{T},\rho} := 4\sqrt{2}\int_0^{\sigma/2}\sqrt{\mathfrak{E}_{\rho,\mathbb{T}}(\delta)}\mathrm{d}\delta.$$

III) If $D_{\mathbb{T},\rho} < \infty$ then the $(Z_t,\ t \in \mathbb{T})$ is bounded and uniformly continuous almost surely.

$\mathbf{2^0}$. The result formulated in Lemma 2 below is a particular case of Theorem 5.2 in Birman and Solomjak (1967).

Let $\gamma > 0, \gamma \notin \mathbb{N}^*$, $m \geq 1$ and $R > 0$ be fixed numbers and let $\Delta_k \subset \mathbb{R}^k,\ k \geq 1$, be a given cube with the sides parallel to the axis. Recall that $|y|$ denotes the euclidian norm of $y \in \mathbb{R}^k$ and $\lfloor\gamma\rfloor$ is the integer part of γ. Set also $D^{\mathbf{n}} = \frac{\partial^{|\mathbf{n}|}}{\partial y_1^{n_1}\cdots\partial y_k^{n_k}},\ \mathbf{n} = (n_1,\ldots,n_k) \in \mathbb{N}^k$.

Denote by $\mathbb{S}_m^\gamma\big(\Delta_k\big)$ the Sobolev-Slobodetskii space, i.e. the set of functions $F : \Delta_k \to \mathbb{R}$ equipped with the norm

$$\|F\|_{\gamma,m} = \Big(\int_{\Delta_k}\big|F(y)\big|^m\mathrm{d}y\Big)^{\frac{1}{m}} + \Big(\sum_{|\mathbf{n}|=\lfloor\gamma\rfloor}\int_{\Delta_k}\int_{\Delta_k}\frac{\big|D^{\mathbf{n}}F(y) - D^{\mathbf{n}}F(z)\big|^m}{|y-z|^{k+m(\gamma-\lfloor\gamma\rfloor)}}\mathrm{d}y\mathrm{d}z\Big)^{\frac{1}{m}}.$$

Denote by $\mathbb{S}_m^\gamma\big(\Delta_k, R\big) = \{F : \Delta_k \to \mathbb{R} :\ \|F\|_{\gamma,m} \leq R\}$ the ball of radius R in this space and set

$$\lambda_k\big(\gamma, m, R, \Delta_k\big) = \inf\left\{c :\ \sup_{\delta\in(0,R]}\delta^{k/\gamma}\mathfrak{E}_{\|\cdot\|_2,\mathbb{S}_m^\gamma\left(\Delta_k,R\right)}(\delta) \leq c\right\}.$$

Lemma 2. $\lambda_k\big(\gamma, m, 1, \Delta_k\big) < \infty$ *for any bounded Δ_k and γ, m, k satisfying $\gamma > k/m - k/2$.*

In view of the obvious relation $\mathfrak{E}_{\|\cdot\|_2,\mathbb{S}_m^\gamma\left(\Delta_k,R\right)}(\delta) = \mathfrak{E}_{\|\cdot\|_2,\mathbb{S}_m^\gamma\left(\Delta_k,1\right)}(\delta/R)$ one has for any $R > 0$

$$\lambda_k\big(\gamma, m, R, \Delta_k\big) = R^{k/\gamma}\lambda\big(\gamma, m, 1, \Delta_k\big). \tag{3.1}$$

3.1. Proof of Theorem 1

For any multi-index $\mathbf{s}\in\mathbb{N}^d$ set $\vec{\mathfrak{h}}_{\mathbf{s}}=\big(\mathfrak{h}_{s_1},\ldots,\mathfrak{h}_{s_d}\big)$, $V_{\mathbf{s}}=\prod_{j=1}^d\mathfrak{h}_{s_j}$ and introduce

$$\upsilon_{\mathbf{s}}(x)=\Big(V_{\mathbf{s}}\Big)^{-\frac12}\int K_{\vec{\mathfrak{h}}_{\mathbf{s}}}(t-x)\,W(\mathrm{d}t),\qquad \eta_{\mathbf{s}}=\Big(\big|\ln\left(\varepsilon V_{\mathbf{s}}\right)\big|\Big)^{-\frac12}\sup_{x\in(-b,b)^d}\big|\upsilon_{\mathbf{s}}(x)\big|.$$

Note that for any $\vec{h}\in\mathrm{H}$ and any $\mathbf{s}\in\mathbb{N}^d$ we obviously have

$$\big|\xi_{\vec{h}}(x)\big|\leq\eta_{\mathbf{s}}V_{\mathbf{s}}^{-\frac12}\sqrt{\big|\ln\left(\varepsilon V_{\mathbf{s}}\right)\big|},\quad\forall x\in\Lambda_{\mathbf{s}}\big[\vec{h}\big],\tag{3.2}$$

and consider separately two cases.

Case $p<\infty$. We have in view of (3.2)

$$\|\xi_{\vec{h}}\|_p^p\leq\sum_{\mathbf{s}\in\mathbb{N}^d}\eta_{\mathbf{s}}^p\Big(\big|\ln\left(\varepsilon V_{\mathbf{s}}\right)\big|V_{\mathbf{s}}^{-1}\Big)^{\frac{p}{2}}\nu_d\big(\Lambda_{\mathbf{s}}\big[\vec{h}\big]\big).$$

Since

$$\Big\|V_{\vec{h}}^{-\frac12}\sqrt{\big|\ln\left(\varepsilon V_{\vec{h}}\right)\big|}\Big\|_p^p=\sum_{\mathbf{s}\in\mathbb{N}^d}\Big(\big|\ln\left(\varepsilon V_{\mathbf{s}}\right)\big|V_{\mathbf{s}}^{-1}\Big)^{\frac{p}{2}}\nu_d\big(\Lambda_{\mathbf{s}}\big[\vec{h}\big]\big),$$

using the obvious inequality $\big(y^{1/p}-z^{1/p}\big)_+\leq\big[(y-z)_+\big]^{1/p}$, $y,z\geq0,p\geq1$, we obtain for any $\vec{h}\in\mathrm{H}$

$$\Big(\big\|\xi_{\vec{h}}\big\|_p-\psi_\varepsilon(\vec{h})\Big)_+\leq(2b)^{\frac{d}{p}}\bigg[\sum_{\mathbf{s}\in\mathbb{N}^d}\Big(\big|\ln\left(\varepsilon V_{\mathbf{s}}\right)\big|V_{\mathbf{s}}^{-1}\Big)^{\frac{p}{2}}\big(\eta_{\mathbf{s}}^p-C_1\big)_+\bigg]^{\frac1p}.$$

Noting that the right hand side of the latter inequality is independent of $\vec{h}$ and denoting $\tilde{q}=(q/p)\vee1$ we obtain using Jensen and triangle inequalities

$$\mathbb{E}\bigg\{\sup_{\vec{h}\in\mathrm{H}}\Big[\big\|\xi_{\vec{h}}\big\|_p-\psi_\varepsilon(\vec{h})\Big]_+\bigg\}^q\leq(2b)^{\frac{dq}{p}}\bigg[\sum_{\mathbf{s}\in\mathbb{N}^d}\Big(\big|\ln\left(\varepsilon V_{\mathbf{s}}\right)\big|V_{\mathbf{s}}^{-1}\Big)^{\frac{p}{2}}\Big\{\mathbb{E}\big(\eta_{\mathbf{s}}^p-C_1\big)_+^{\tilde{q}}\Big\}^{\frac{1}{\tilde{q}}}\bigg]^{\frac{q}{p}}.\tag{3.3}$$

Let $\mathbf{s}\in\mathbb{N}^d$ be fixed. We have

$$\begin{aligned}\mathbb{E}\big(\eta_{\mathbf{s}}^p-C_1\big)_+^{\tilde{q}}&=\tilde{q}\int_0^\infty z^{\tilde{q}-1}\mathbb{P}\big\{\eta_{\mathbf{s}}^p\geq C_1+z\big\}\mathrm{d}z\\&=\tilde{q}\int_0^\infty z^{\tilde{q}-1}\mathbb{P}\bigg\{\sup_{x\in(-b,b)^d}\big|\upsilon_{\mathbf{s}}(x)\big|\geq\big[C_1+z\big]^{\frac1p}\sqrt{\big|\ln\left(\varepsilon V_{\mathbf{s}}\right)\big|}\bigg\}\mathrm{d}z.\end{aligned}\tag{3.4}$$

Set $\mathfrak{z}=\big[C_1+z\big]^{\frac1p}\sqrt{\big|\ln\left(\varepsilon V_{\mathbf{s}}\right)\big|}$ and prove that

$$\mathbb{P}\bigg\{\sup_{x\in(-b,b)^d}\big|\upsilon_{\mathbf{s}}(x)\big|\geq\mathfrak{z}\bigg\}\leq2\big(\varepsilon V_{\mathbf{s}}\big)^{2(q\vee p)}\exp\bigg(-\frac{z^{\frac2p}}{8\|\mathcal{K}\|_2^{2d}}\bigg),\quad\forall z\geq0.\tag{3.5}$$

Since $\upsilon_{\mathbf{s}}(\cdot)$ is a zero mean gaussian random field in view of the obvious relation $\sup_x|\upsilon_{\mathbf{s}}(x)| = [\sup_x \upsilon_{\mathbf{s}}(x)]\vee[\sup_x\{-\upsilon_{\mathbf{s}}(x)\}]$ we get

$$\mathbb{P}\Big\{\sup_{x\in(-b,b)^d}\big|\upsilon_{\mathbf{s}}(x)\big|\geq\mathfrak{z}\Big\}\leq 2\mathbb{P}\Big\{\sup_{x\in(-b,b)^d}\upsilon_{\mathbf{s}}(x)\geq\mathfrak{z}\Big\}. \tag{3.6}$$

Let ρ denote the intrinsic semi-metric of $\upsilon_{\mathbf{s}}(\cdot)$ on $(-b,b)^d$.

We have for any $x,x'\in(-b,b)^d$ in view of Assumption 1

$$\begin{aligned}
\rho^2\big(x,x'\big) &\leq \int\Big[K(u)-K\big(\vec{\mathfrak{h}}_{\mathbf{s}}^{-1}(x-x')+u\big)\Big]^2\mathrm{d}u\\
&= 2\|K\|_2^2-2\int_{[-a,a]^d}K(u)K\big(\vec{\mathfrak{h}}_{\mathbf{s}}^{-1}(x-x')+u\big)\mathrm{d}u\\
&= -2\int_{[-a,a]^d}K(u)\Big[K\big(\vec{\mathfrak{h}}_{\mathbf{s}}^{-1}(x-x')+u\big)-K(u)\Big]\mathrm{d}u\\
&\leq 2L\|K\|_1\big|\vec{\mathfrak{h}}_{\mathbf{s}}^{-1}(x-x')\big|\leq 2L\|K\|_1V_{\mathbf{s}}^{-1}\big|x-x'\big|.
\end{aligned} \tag{3.7}$$

Recall that $\mathfrak{E}_{\rho,(-b,b)^d}(\delta)$, $\delta>0$, denotes the δ-entropy of $(-b,b)^d$ measured in ρ.

Putting $c_1=\big|\ln\big(4bL\|K\|_1\big)\big|$, we deduce from (3.7) for any $\delta>0$

$$\mathfrak{E}_{\rho,(-b,b)^d}(\delta)\leq dc_1+d\big|\ln\big(V_{\mathbf{s}}\big)\big|+2d\big[\ln(1/\delta)\big]_+. \tag{3.8}$$

Note that $\sigma^2:=\sup_{x\in(-b,b)^d}\mathbb{E}\left(\eta_{\mathbf{s}}^2(x)\right)=\|K\|_2^2$ and, therefore,

$$D_{(-b,b)^d,\rho}\leq\sqrt{d}\left(c_2+2\sqrt{2}\|K\|_2\sqrt{\big|\ln\big(V_{\mathbf{s}}\big)\big|}\right), \tag{3.9}$$

where $c_2=2\|K\|_2\sqrt{2c_1}+4\sqrt{2}\int_0^{2^{-1}\|K\|_2}\sqrt{[\ln(1/\delta)]_+}\mathrm{d}\delta$.

Thus, using the second assertion of Lemma 1 we have

$$\mathbf{E}:=\mathbb{E}\Big(\sup_{x\in(-b,b)^d}\upsilon_{\mathbf{s}}(x)\Big)\leq 2\sqrt{2d\pi}+2\sqrt{2dc_1}\|K\|_2+2\sqrt{2d}\|K\|_2\sqrt{\big|\ln\big(V_{\mathbf{s}}\big)\big|}.$$

Here we have used that $4\sqrt{2}\int_0^{2^{-1}\|K\|_2}\sqrt{[\ln(1/\delta)]_+}\mathrm{d}\delta\leq 2\sqrt{2\pi}$.

Note that in view of the definition of C_1

$$\mathfrak{z}-\mathbf{E}\geq 2^{-1}C_1^{\frac{1}{p}}\sqrt{\big|\ln\big(\varepsilon V_{\mathbf{s}}\big)\big|}-\mathbf{E}+2^{-1}z^{\frac{1}{p}}\geq 2\sqrt{(q\vee p)}\|K\|_2\sqrt{\big|\ln\big(\varepsilon V_{\mathbf{s}}\big)\big|}+2^{-1}z^{\frac{1}{p}}.$$

Remark that the third assertion of Lemma 1 and (3.9) implies that the first assertion of Lemma 1 is applicable with $\mathbb{T}=(-b,b)^d$ and $Z_t=\upsilon_{\mathbf{s}}(x)$ and we get for any $\mathbf{s}\in\mathbb{N}^d$

$$\mathbb{P}\Big\{\sup_{x\in(-b,b)^d}\upsilon_{\mathbf{s}}(x)\geq\mathfrak{z}\Big\}\leq\big(\varepsilon V_{\mathbf{s}}\big)^{2(q\vee p)}\exp\left(-\frac{z^{\frac{2}{p}}}{8\|K\|_2^2}\right).$$

Thus, the inequality (3.5) follows now from (3.6). We obtain from (3.4) and (3.5)

$$\mathbb{E}\big(\eta_{\mathbf{s}}^p-C_1\big)_+^{\tilde{q}}\leq 2\tilde{q}\big(\varepsilon V_{\mathbf{s}}\big)^{2(q\vee p)}\int_0^\infty z^{\tilde{q}-1}\exp\left(-\frac{z^{\frac{2}{p}}}{8\|K\|_2^2}\right)\mathrm{d}z=:c_3\big(\varepsilon V_{\mathbf{s}}\big)^{2(q\vee p)}. \tag{3.10}$$

Taking into account that $\big|\ln\big(\varepsilon V_{\mathbf{s}}\big)\big| \leq \big|\ln(\varepsilon)\big|V_{\mathbf{s}}^{-1}$, since $\varepsilon, \mathfrak{h} \leq e^{-2}$, we deduce from (3.3) and (3.10) that

$$\mathbb{E}\bigg\{\sup_{\vec{h}\in\mathrm{H}}\Big[\big\|\xi_{\vec{h}}\big\|_p - \psi_\varepsilon(\vec{h})\Big]_+\bigg\}^q \leq (2b)^{\frac{dq}{p}}(c_3)^{\frac{q}{qp}}\varepsilon^q\bigg[\sum_{\mathbf{s}\in\mathbb{N}^d} V_{\mathbf{s}}^p\bigg]^{\frac{q}{p}} \leq (4b)^{\frac{dq}{p}}(c_3)^{\frac{q}{qp}}\varepsilon^q = (C_3\varepsilon)^q.$$

Case $p=\infty$**.** We have in view of (3.2)

$$\big\|\xi_{\vec{h}}\big\|_\infty = \sup_{\mathbf{s}\in\mathbb{N}^d}\sup_{x\in\Lambda_{\mathbf{s}}[\vec{h}]}\big|\xi_{\vec{h}}(x)\big| \leq \sup_{\mathbf{s}\in\mathbb{N}^d}\Big(\eta_{\mathbf{s}}\sqrt{\big|\ln\big(\varepsilon V_{\mathbf{s}}\big)\big|}V_{\mathbf{s}}^{-\frac{1}{2}}\Big).$$

Since, obviously

$$\Big\|V_{\vec{h}}^{-\frac{1}{2}}\sqrt{\big|\ln\big(\varepsilon V_{\vec{h}}\big)\big|}\Big\|_\infty = \sup_{\mathbf{s}\in\mathbb{N}^d}\Big(\sqrt{\big|\ln\big(\varepsilon V_{\mathbf{s}}\big)\big|}V_{\mathbf{s}}^{-\frac{1}{2}}\Big),$$

we obtain for any $\vec{h}\in\mathrm{H}$

$$\Big(\big\|\xi_{\vec{h}}\big\|_\infty - \psi_\varepsilon(\vec{h})\Big)_+ \leq \bigg[\sup_{\mathbf{s}\in\mathbb{N}^d}\Big(\eta_{\mathbf{s}}\sqrt{\big|\ln\big(\varepsilon V_{\mathbf{s}}\big)\big|}V_{\mathbf{s}}^{-\frac{1}{2}}\Big) - C_1\sup_{\mathbf{s}\in\mathbb{N}^d}\Big(\sqrt{\big|\ln\big(\varepsilon V_{\mathbf{s}}\big)\big|}V_{\mathbf{s}}^{-\frac{1}{2}}\Big)\bigg]_+.$$

Since $\big(\sup_\varkappa \mathfrak{a}_\varkappa - \sup_\varkappa \mathfrak{b}_\varkappa\big)_+ \leq \sup_\varkappa\big(\mathfrak{a}_\varkappa - \mathfrak{b}_\varkappa\big)_+$ for arbitrary collections $\{\mathfrak{a}_\varkappa\}_\varkappa$ and $\{\mathfrak{b}_\varkappa\}_\varkappa$ of positives numbers, we obtain for any $q\geq 1$

$$\begin{aligned}\Big(\big\|\xi_{\vec{h}}\big\|_\infty - \psi_\varepsilon(\vec{h})\Big)_+^q &\leq \sup_{\mathbf{s}\in\mathbb{N}^d}\Big(\sqrt{\big|\ln\big(\varepsilon V_{\mathbf{s}}\big)\big|}V_{\mathbf{s}}^{-\frac{1}{2}}\Big)^q\big(\eta_{\mathbf{s}} - C_1\big)_+^q \\ &\leq \sum_{\mathbf{s}\in\mathbb{N}^d}\Big(\sqrt{\big|\ln\big(\varepsilon V_{\mathbf{s}}\big)\big|}V_{\mathbf{s}}^{-\frac{1}{2}}\Big)^q\big(\eta_{\mathbf{s}} - C_1\big)_+^q\end{aligned}$$

Taking into account that the right hand side of the latter inequality is independent of $\vec{h}$ we obtain

$$\mathbb{E}\bigg\{\sup_{\vec{h}\in\mathrm{H}}\Big[\big\|\xi_{\vec{h}}\big\|_\infty - \psi_\varepsilon(\vec{h})\Big]_+\bigg\}^q \leq \sum_{\mathbf{s}\in\mathbb{N}^d}\Big(\sqrt{\big|\ln\big(\varepsilon V_{\mathbf{s}}\big)\big|}V_{\mathbf{s}}^{-\frac{1}{2}}\Big)^q\mathbb{E}\big(\eta_{\mathbf{s}} - C_1\big)_+^q \tag{3.11}$$

Note also that inequality (3.10) is proved for arbitrary $p, \tilde{q} \geq 1$. Applying it formally with $p=1$ and $\tilde{q}=q$ we obtain

$$\mathbb{E}\big(\eta_{\mathbf{s}} - C_1\big)_+^q \leq 2q\big(\varepsilon V_{\mathbf{s}}\big)^{2q}\int_0^\infty z^{q-1}\exp\Big(-\frac{z^2}{8\|K\|_2^2}\Big)\mathrm{d}z \tag{3.12}$$

and the assertion of the theorem for $p=\infty$ follows from (3.11) and (3.12). ∎

3.2. Proof of Theorem 2

3.2.1. Auxiliary lemma

Set $\lambda^*(\gamma,m) = \lambda_1\big(\gamma,m,1,[-a-b,a+b]\big)$, where we recall the number $a>0$ is involved in Assumption 2 and $\lambda_k(\cdot,\cdot,\cdot,\cdot),\ k\in\mathbb{N}^d$ is defined in Lemma 2.

If $d\geq 2$ write $\mathrm{x}=(x_2,\dots,x_d)$ and define for any $\vec{\eta}\in\mathrm{H}$ and any $\mathrm{x}\in(-b,b)^{d-1}$

$$\lambda_{\vec{\eta},\mathbf{s}}(\mathrm{x})=\bigg[\int_{-b}^{b}1_{\Lambda_{\mathbf{s}}[\vec{\eta}]}(x)\nu_1\big(\mathrm{d}x_1\big)\bigg]^{\frac{\tau}{r}}.$$

Later on for any $x\in(-b,b)^d$ we will use the following notation $x=(x_1,\mathrm{x})$. If $d=1$ the dependence of x should be omitted in all formulas. In particular, if $d=1$ then $\lambda_{\eta_1,s_1}=\Big\{\nu_1\big(\Lambda_{s_1}[\eta_1]\big)\Big\}^{\frac{\tau}{r}}$.

For any $\mathrm{x}\in(-b,b)^{d-1}$ and $\mathbf{s}\in\mathbb{N}^d$ introduce the set of functions $Q:\mathbb{R}\to\mathbb{R}$

$$\mathcal{Q}_{\mathrm{x},\mathbf{s}}=\bigg\{Q(\cdot)=\lambda^{-1}_{\vec{\eta},\mathbf{s}}(\mathrm{x})\int_{-b}^{b}\mathfrak{h}_{s_1}^{-1/2}\mathcal{K}\bigg(\frac{\cdot-x_1}{\mathfrak{h}_{s_1}}\bigg)\ell(x_1)1_{\Lambda_{\mathbf{s}}[\vec{\eta}]}(x_1,\mathrm{x})\nu_1(\mathrm{d}x_1),\ \ell\in\mathbb{B}_{\mathfrak{q}},\ \vec{\eta}\in\mathrm{H}\bigg\}.$$

where $\mathbb{B}_{\mathfrak{q}}=\Big\{\ell:(-b,b)\to\mathbb{R}:\ \int_{-b}^{b}|\ell(x_1)|^{\mathfrak{q}}\nu(\mathrm{d}x_1)\leq 1\Big\},\ 1/\mathfrak{q}=1-1/r.$

If $\lambda_{\vec{\eta},\mathbf{s}}(\mathrm{x})=0$ put by continuity $Q\equiv 0$. Let finally $\mu^{-1}=\mathfrak{q}^{-1}+\tau r^{-1}$ and note that $2>\mu>1$ since $\tau<1$ and $r>2$.

Lemma 3. *For any $\mathrm{x}\in(-b,b)^{d-1}$, $\mathbf{s}\in\mathbb{N}^d$ and any $\omega\in\big(1/\mu-1/2,1\big)$ one has*

$$\mathfrak{E}_{\|\cdot\|_2,\mathcal{Q}_{\mathrm{x},\mathbf{s}}}(\epsilon)\leq\lambda^*\big(\omega,\mu\big)R_\mu^{\frac{1}{\omega}}\ \mathfrak{h}_{s_1}^{\frac{1}{2\omega}-1}\epsilon^{-\frac{1}{\omega}}\qquad\forall\epsilon\in\Big(0,R_\mu\ \mathfrak{h}_{s_1}^{\frac{1}{2}-\omega}\Big],$$

where $R_\mu=\Big[\Big\{2^{-1}\|\mathcal{K}\|_{\frac{2\mu}{3\mu-2}}\Big\}\vee\Big\{\|\mathcal{K}\|_1+2\left[5\big\{4L(a+1)\big\}^\mu+4\big\{2\|\mathcal{K}\|_1\big\}^\mu(2-\mu)^{-1}\right]^{\frac{1}{\mu}}\Big\}\Big]$.

3.2.2. Constants and expressions.

Introduce $\Omega=\Big\{\{\omega_1,\omega_2\}:\ \ \omega_1<1/2<\omega_2,\ \ [\omega_1,\omega_2]\subset(1/\mu-1/2,1)\Big\}$ and set

$$C_2(r,\tau,\mathcal{L})=[1\vee(2b)^{d-1}]\left[\mathcal{L}^{\frac{1}{r}}+\mathcal{L}^{\frac{\tau}{r}}\big(1-e^{-\frac{\tau p}{4}}\big)^{\frac{\tau-1}{r}}\right]\big[\widetilde{C}_\mu+\widehat{C}\big]+e^r\sqrt{2(1+q)}\big(r\sqrt{e}\big)^d\|\mathcal{K}\|^d_{\frac{2r}{r+2}};$$

$$\widehat{C}_\mu=\left[\frac{r}{1-\tau}\int_0^\infty\big(u+\widetilde{C}_\mu\big)^{\frac{r+\tau-1}{1-\tau}}\exp\Big\{-u^2\Big[2|\mathcal{K}\|_2^{d-1}\|\mathcal{K}\|_{\frac{2\mu}{3\mu-2}}\Big]^{-1}\Big\}\mathrm{d}u\right]^{\frac{1-\tau}{r}};$$

$$\widetilde{C}_\mu=C_\mu+4^d\big(\sqrt{2e^r}+\sqrt{8\pi}\big)\|\mathcal{K}\|_2^{d-1}\|\mathcal{K}\|_{\frac{2\mu}{3\mu-2}};$$

$$C_\mu=4\sqrt{2}\|\mathcal{K}\|_2^{d-1}\inf_{\{\omega_1,\omega_2\}\in\Omega}\bigg[\sqrt{\lambda^*\big(\omega_2,\mu\big)}\big(1-[2\omega_2]^{-1}\big)R_\mu^{\frac{1}{2\omega_2}}+\sqrt{\lambda^*\big(\omega_1,\mu\big)}\big([2\omega_1]^{-1}-1\big)R_\mu^{\frac{1}{2\omega_1}}\bigg].$$

$$C_4=\bigg(\gamma_{q+1}\sqrt{(\pi/2)}\big[1\vee(2b)^{qd}\big]\sum_{r\in\mathbb{N}_p^*}e^{-e^r}\left[\big(r\sqrt{e}\big)^d\|\mathcal{K}\|^d_{\frac{2r}{r+2}}\right]^{\frac{q}{2}}\bigg)^{\frac{1}{q}}$$

3.2.3. Main steps in the proof of Theorem 2.

The goal of this paragraph is to explain the basic ideas and main ingredients of the proof of Theorem 2 which is rather long and tricky.

Set for any $r \in \mathbb{N}_p^*$ and $\vec{h} \in \mathrm{H}$

$$\zeta_{\vec{h}}(r) = \Big\| V_{\vec{h}}^{\frac{1}{2}} \xi_{\vec{h}} \Big\|_r, \qquad \zeta(r) = \sup_{\vec{h} \in \mathrm{H}} \zeta_{\vec{h}}(r). \tag{3.13}$$

Our basic idea is to prove that for any $r \in \mathbb{N}_p^*$ one can find a constant $U(r)$ being the upper function for $\zeta(r)$ whenever $\mathrm{H} \subset \mathbb{H}_d(\tau, \mathcal{L}, \mathcal{A})$ is considered. Since $U(r)$ is independent of $\vec{h}$ the initial problem is reduced to the study of the deviation of the supremum of $\zeta_{\vec{h}}(r)$ on H.

First part of the proof consists in the aforementioned reduction of the considered problem to the study of the upper function for the $\mathbb{L}_r$-norm of the normalized process $V_{\vec{h}}^{\frac{1}{2}} \xi_{\vec{h}}(\cdot)$. This part is rather short and straightforward and the obtained reduction is given in (3.22).

Our next observation consists in the following. In view of duality arguments

$$\zeta(r) = \sup_{\vec{h} \in \mathrm{H}} \zeta_{\vec{h}}(r) = \sup_{\vec{h} \in \mathrm{H}} \sup_{\vartheta \in \mathbb{B}_{\mathfrak{q},d}} \Upsilon_{\vec{h},\vartheta}, \qquad \Upsilon_{\vec{h},\vartheta} := \int_{(-b,b)^d} V_{\vec{h}}^{\frac{1}{2}}(x) \xi_{\vec{h}}(x) \vartheta(x) \nu_d(\mathrm{d}x),$$

where $\mathbb{B}_{\mathfrak{q},d} = \{\vartheta : (-b,b)^d \to \mathbb{R} : \ \|\vartheta\|_{\mathfrak{q}} \leq 1\}$ and $1/\mathfrak{q} = 1 - 1/r$. Obviously $\Upsilon_{\vec{h},\vartheta}$ is centered gaussian random function on $\mathrm{H} \times \mathbb{B}_{\mathfrak{q},d}$. Hence, if we show that for some $0 < V(r) < \infty$

$$\mathbb{E}\big\{\zeta(r)\big\} \leq V(r), \tag{3.14}$$

then the first assertion of Lemma 1 with

$$\sigma_{\Upsilon}^2 := \sup_{\vec{h} \in \mathrm{H}} \sup_{\theta \in \mathbb{B}_{\mathfrak{q},d}} \mathbb{E}\big\{\Upsilon_{\vec{h},\vartheta}\big\}^2 \tag{3.15}$$

will be applicable to the random variable $\zeta(r)$.

Second part of the proof consists in finding a suitable upper bound for σ_{Υ}. It is also short and straightforward and the obtained bound is presented in (3.25).

Main part of the proof, that deals with establishing (3.14), is divided in several steps. Although the proof is done in an arbitrary dimension some additional difficulties come from the consideration of an anisotropic bandwidth collection. For this reason the explanations below are given in the case $d \geq 2$. Define for any $\mathbf{s} = (s_1, \ldots, s_d) \in \mathbb{N}^d$ and $\mathrm{x} \in (-b,b)^{d-1}$

$$\varsigma_{\mathbf{s}}\big(Q, \mathrm{x}\big) = \int Q(t_1) G_{\mathbf{s}}(\mathrm{t}, \mathrm{x}) W(\mathrm{d}t), \quad Q \in \mathcal{Q}_{\mathrm{x},\mathbf{s}}. \tag{3.16}$$

Here we have put $\mathrm{t} = (t_2, \ldots, t_d)$, denoted $t = (t_1, \mathrm{t})$ for any $t \in \mathbb{R}^d$, and set

$$G_{\mathbf{s}}(\mathrm{t}, \mathrm{x}) = \prod_{i=2}^d \mathfrak{h}_{s_i}^{-\frac{1}{2}} \mathcal{K}\big((\mathrm{t}_i - \mathrm{x}_i)/\mathfrak{h}_{s_i}\big), \ \ \mathrm{t} \in \mathbb{R}^{d-1}, \ \mathrm{x} \in (-b,b)^{d-1}.$$

Remind also that the set $\mathcal{Q}_{\mathrm{x},\mathbf{s}}$ is defined in Lemma 3 and $\mathfrak{h}_s = e^{-s}\mathfrak{h}, \ s \in \mathbb{N}$.

The basic idea used in establishing (3.14) consists in bounding from above $\mathbb{E}\zeta(r)$ by some quantities related to the collection of random variables

$$\Big\{\varsigma_{\mathbf{s}}(\mathrm{x}) := \sup_{Q \in \mathcal{Q}_{\mathbf{s},\mathrm{x}}} \varsigma_{\mathbf{s}}\big(Q, \mathrm{x}\big), \ \ \mathbf{s} \in \mathbb{N}^d, \ \ \mathrm{x} \in (-b,b)^{d-1}\Big\}. \tag{3.17}$$

First step in the proof of (3.14) consists in the realization of the aforementioned idea. The main ingredients for that are: duality arguments, product structure of the kernel (Assumption 2 (**ii**)) and the fact that $\mathrm{H} \in \mathbb{H}_d(\tau, \mathcal{L}, \mathcal{A})$. The required bound is given in (3.36) ($d \geq 2$) and (3.37) ($d = 1$).

Second step in the proof of (3.14). Looking at the inequality (3.36) (or (3.37)) we remark that one has to bound from above the quantities

$$\sup_{\mathrm{x} \in (-b,b)^{d-1}} \mathbb{E}\bigg(\sup_{\mathbf{s} \in \mathcal{S}_d} \varsigma_{\mathbf{s}}^{\frac{r}{1-\tau}}(\mathrm{x}) \bigg), \qquad \sup_{\mathrm{x} \in (-b,b)^{d-1}} \sup_{\mathbf{s} \in \mathbb{N}^d} \mathbb{E}\Big(\varsigma_{\mathbf{s}}^{\frac{r}{1-\tau}}(\mathrm{x}) \Big). \tag{3.18}$$

It is important to note that $\mathcal{S}_d$ is the finite set and its cardinality is completely determined by the parameters $\mathfrak{h}$ and $\mathcal{A}$.

Another important remark is that $\varsigma_{\mathbf{s}}(Q, \mathrm{x})$ is zero-mean gaussian random function on $\mathcal{Q}_{\mathbf{s},\mathrm{x}}$. Hence, in order to compute the quantities given in (3.18) one can use the concentration inequality presented in the first assertion of Lemma 1. The most tricky part of the realization of this program consists in bounding from above $\mathbb{E}\varsigma_{\mathbf{s}}$, which, in its turn, is reduced to the bounding from above the Dudley's integral in view of the second assertion of Lemma 1. The required bound is given in (3.43).

The main technical tool here is Lemma 3 providing very precise estimates for the entropy of the set $\mathcal{Q}_{\mathbf{s},\mathrm{x}}$, which are possible because this set belongs to the intersection of balls in the Sobolev-Slobodetskii space (proof of Lemma 3). The result obtained in Lemma 3 allows to use different bounds for the entropy of $\mathcal{Q}_{\mathbf{s},\mathrm{x}}$ near and outside of the origin in the computation of the Dudley's integral.

Final step in the proof of (3.14) consists of routine computations related to the careful application of the first assertion of Lemma 1.

3.2.4. Proof of Theorem 2.

Put for brevity $C_2(r) = C_2(r, \tau, \mathcal{L})$ and let

$$\psi_r(\vec{h}) = C_2(r) \Big\| V_{\vec{h}}^{-\frac{1}{2}} \Big\|_{\frac{rp}{r-p}}, \; r \in \mathbb{N}_p^*.$$

For any $\vec{h} \in \mathrm{H}$ define $r^*\big(\vec{h}\big) = \arg\inf_{r \in \mathbb{N}_p^*(\vec{h}, \mathcal{A})} \psi_r(\vec{h})$. Note that $C_2(r) < \infty$ for any $r \in \mathbb{N}_p^*$ and

$$\psi_r(\vec{h}) \geq C_2(r)\mathfrak{h}^{-d} \to \infty, \;\; r \to \infty,$$

and, therefore, $r^*\big(\vec{h}\big) < \infty$ for any $\vec{h} \in \mathbb{B}(\mathcal{A})$. The latter fact allows us to assert that

$$\psi(\vec{h}) = \inf_{r \in \mathbb{N}_p^*(\vec{h}, \mathcal{A})} \psi_r(\vec{h}) = \psi_{r^*(\vec{h})}(\vec{h}) =: C_2\Big(r^*\big(\vec{h}\big), \tau, \mathcal{L}\Big) \Big\| V_{\vec{h}}^{-\frac{1}{2}} \Big\|_{\frac{pr^*(\vec{h})}{r^*(\vec{h})-p}}, \tag{3.19}$$

since $\mathbb{N}_p^*(\vec{h}, \mathcal{A})$ is a discrete set.

By definition $r^*\big(\vec{h}\big) \geq r_{\mathcal{A}}\big(\vec{h}\big)$, where recall $r_{\mathcal{A}}\big(\vec{h}\big)$ is defined in (2.1). Hence we get from Hölder inequality and the definition of $r_{\mathcal{A}}\big(\vec{h}\big)$

$$\Big\| V_{\vec{h}}^{-\frac{1}{2}} \Big\|_{\frac{pr^*(\vec{h})}{r^*(\vec{h})-p}} \leq \big[1 \vee (2b)^d\big] \Big\| V_{\vec{h}}^{-\frac{1}{2}} \Big\|_{\frac{pr_{\mathcal{A}}(\vec{h})}{r_{\mathcal{A}}(\vec{h})-p}} \leq \mathcal{A}\big[1 \vee (2b)^d\big]. \tag{3.20}$$

Using the notations given in (3.13), we obtain for any $\vec{h}\in\mathrm{H}$, applying Hölder inequality

$$\left\|\xi_{\vec{h}}\right\|_p \;\leq\; \inf_{r\in\mathbb{N}_p^*}\left\{\zeta(r)\left\|V_{\vec{h}}^{-\frac{1}{2}}\right\|_{\frac{pr}{r-p}}\right\}\leq\zeta\Big(r^*\big(\vec{h}\big)\Big)\left\|V_{\vec{h}}^{-\frac{1}{2}}\right\|_{\frac{pr^*(\vec{h})}{r^*(\vec{h})-p}}. \tag{3.21}$$

We deduce from (3.19), (3.20) and (3.21) that for any $\vec{h}\in\mathrm{H}$

$$\begin{aligned}\Big[\|\xi_{\vec{h}}\|_p-\psi(\vec{h})\Big]_+^q &\leq \left\|V_{\vec{h}}^{-\frac{1}{2}}\right\|^q_{\frac{pr^*(\vec{h})}{r^*(\vec{h})-p}}\Big[\zeta\Big(r^*\big(\vec{h}\big)\Big)-C_2\Big(r^*\big(\vec{h}\big)\Big)\Big]_+^q\\ &\leq \mathcal{A}^q\big[1\vee(2b)^{qd}\big]\Big[\zeta\Big(r^*\big(\vec{h}\big)\Big)-C_2\Big(r^*\big(\vec{h}\big)\Big)\Big]_+^q\leq \mathcal{A}^q\big[1\vee(2b)^{qd}\big]\sum_{r\in\mathbb{N}_p^*}\big[\zeta(r)-C_2(r)\big]_+^q.\end{aligned}$$

To get the last inequality we have used that $r^*\big(\vec{h}\big)\in\mathbb{N}_p^*$ for any $\vec{h}\in\mathrm{H}$.

Taking into account that the right hand side of the latter inequality is independent of $\vec{h}$ we get

$$\mathbb{E}\bigg(\sup_{\vec{h}\in\mathrm{H}}\Big[\|\xi_{\vec{h}}\|_p-\psi(\vec{h})\Big]_+\bigg)^q\leq\big[1\vee(2b)^{qd}\big]\mathcal{A}^q\sum_{r\in\mathbb{N}_p^*}\mathbb{E}\big[\zeta(r)-C_2(r)\big]_+^q. \tag{3.22}$$

Also we have for any $r\in\mathbb{N}_p^*$

$$\mathbb{E}\big[\zeta(r)-C_2(r)\big]_+^q=q\int_0^\infty z^{q-1}\mathbb{P}\Big\{\zeta(r)\geq C_2(r)+z\Big\}\mathrm{d}z. \tag{3.23}$$

1^0. Our goal now is to prove the following inequality: for any $z\geq0$ and $r\in\mathbb{N}_p^*$

$$\mathbb{P}\Big\{\zeta(r)\geq C_2(r)+z\Big\}\leq e^{-e^r}e^{-qe^{2\sqrt{2d|\ln(\mathfrak{h})|}}}\;\exp\Big\{-\Big(2\big(r\sqrt{e}\big)^d\|\mathcal{K}\|^d_{\frac{2r}{r+2}}\Big)^{-1}z^2\Big\}. \tag{3.24}$$

To do that we will realize the program discussed in Section 3.2.3 and consisting in the proof of (3.14) and bounding from above σ_Υ given in (3.15).

1^0**a**. Let us bound from above σ_Υ. By definition

$$\Upsilon_{\vec{h},\vartheta}=\int\bigg[\int_{(-b,b)^d}V_{\vec{h}}^{-\frac{1}{2}}(x)K\bigg(\frac{t-x}{\vec{h}(x)}\bigg)\vartheta(x)\nu_d(\mathrm{d}x)\bigg]W(\mathrm{d}t)$$

and, therefore,

$$\sigma_\Upsilon=\sup_{\vec{h}\in\mathrm{H}}\sup_{\vartheta\in\mathbb{B}_{\mathfrak{q},d}}\bigg[\int\bigg[\int_{(-b,b)^d}V_{\vec{h}}^{-\frac{1}{2}}(x)K\bigg(\frac{t-x}{\vec{h}(x)}\bigg)\vartheta(x)\nu_d(\mathrm{d}x)\bigg]^2\nu_d(\mathrm{d}t)\bigg]^{\frac{1}{2}}.$$

In view of triangle inequality and Assumption 2 (**ii**)

$$\sigma_\Upsilon\leq\sum_{\mathbf{s}\in\mathbb{N}^d}\prod_{j=1}^d\mathfrak{h}_{s_j}^{-\frac{1}{2}}\sup_{\vartheta\in\mathbb{B}_{\mathfrak{q},d}}\bigg(\int\bigg[\int_{(-b,b)^d}\bigg|\prod_{j=1}^d\mathcal{K}\bigg(\frac{t_j-x_j}{\mathfrak{h}_{s_j}}\bigg)\bigg|\big|\vartheta(x)\big|\nu_d(\mathrm{d}x)\bigg]^2\nu_d(\mathrm{d}t)\bigg)^{\frac{1}{2}}.$$

Applying the Young inequality and taking into account that $\vartheta\in\mathbb{B}_{\mathfrak{q},d}$ we obtain

$$\sigma_\Upsilon\leq\|\mathcal{K}\|^d_{\frac{2r}{r+2}}\sum_{\mathbf{s}\in\mathbb{N}^d}\prod_{j=1}^d\mathfrak{h}_{s_j}^{\frac{1}{r}}\leq\big[1-e^{-\frac{1}{r}}\big]^{-d}\|\mathcal{K}\|^d_{\frac{2r}{r+2}}\mathfrak{h}^{\frac{d}{r}}\leq\big(r\sqrt{e}\big)^d\|\mathcal{K}\|^d_{\frac{2r}{r+2}}\mathfrak{h}^{\frac{d}{r}}. \tag{3.25}$$

1^0**b**. Let us prove (3.14). Set for any $\mathbf{s}\in\mathbb{N}^d$, and $\vec{h}\in\mathrm{H}$

$$\xi_{\vec{h},\mathbf{s}}(x)=1_{\Lambda_{\mathbf{s}}\left[\vec{h}\right]}(x)\int\bigg[\prod_{i=1}^d\mathfrak{h}_{s_i}^{-\frac12}\mathcal{K}\big((t_i-x_i)/\mathfrak{h}_{s_i}\big)\bigg]W(\mathrm{d}t),\;\;x\in(-b,b)^d.$$

We obviously have for any $\vec{h}\in\mathrm{H}$

$$\zeta^r_{\vec{h}}(r)=\Big\|V^{\frac12}_{\vec{h}}\xi_{\vec{h}}\Big\|_r^r=\sum_{\mathbf{s}\in\mathbb{N}^d}\big\|\xi_{\vec{h},\mathbf{s}}\big\|_r^r. \tag{3.26}$$

Moreover, note that $\big|\xi_{\vec{h},\mathbf{s}}(x)\big|\leq 1_{\Lambda_{\mathbf{s}}\left[\vec{h}\right]}(x)\big|\ln\big(\varepsilon V_{\mathbf{s}}\big)\big|^{\frac12}\eta_{\mathbf{s}}$ for any $x\in(-b,b)^d$, where, recall, $V_{\mathbf{s}}$ and $\eta_{\mathbf{s}}$ are defined in the beginning of the proof of Theorem 1. Since, we have proved that $\eta_{\mathbf{s}}$ is bounded almost surely, one gets

$$\int_{-b}^b\big|\xi_{\vec{h},\mathbf{s}}(x)\big|^r\nu_1(\mathrm{d}x_1)\leq\lambda^r_{\vec{h},\mathbf{s}}(\mathrm{x})\big|\ln\big(\varepsilon V_{\mathbf{s}}\big)\big|^{\frac r2}\eta^r_{\mathbf{s}}=0,\quad\text{if}\quad\lambda_{\vec{h},\mathbf{s}}(\mathrm{x})=0. \tag{3.27}$$

On the other hand in view of duality arguments

$$\int_{-b}^b\big|\xi_{\vec{h},\mathbf{s}}(x)\big|^r\nu_1(\mathrm{d}x_1)=\bigg[\sup_{\ell\in\mathbb{B}_{\mathfrak{q}}}\int_{-b}^b\xi_{\vec{h},\mathbf{s}}(x)\ell(x_1)\nu_1\big(\mathrm{d}x_1\big)\bigg]^r, \tag{3.28}$$

where, recall, $\mathbb{B}_{\mathfrak{q}}=\Big\{\ell:(-b,b)\to\mathbb{R}:\;\int_{-b}^b|\ell(y)|^{\mathfrak{q}}\nu(\mathrm{d}y)\leq 1\Big\}$, $1/\mathfrak{q}=1-1/r$.

Let $d\geq 2$. The following simple remark is crucial for all further consideration: in view of (3.27) and (3.28) for any $\mathrm{x}\in(-b,b)^{d-1}$, $\mathbf{s}\in\mathbb{N}^d$ and for any $\vec{h}\in\mathrm{H}$

$$\int_{-b}^b\big|\xi_{\vec{h},\mathbf{s}}(x_1,\mathrm{x})\big|^r\nu_1(\mathrm{d}x_1)\leq\lambda^r_{\vec{h},\mathbf{s}}(\mathrm{x})\varsigma^r_{\mathbf{s}}(\mathrm{x}). \tag{3.29}$$

where $\varsigma_{\mathbf{s}}$ is defined in (3.17).

Indeed, if $\lambda_{\vec{h},\mathbf{s}}(\mathrm{x})=0$ (3.29) follows from (3.27). If $\lambda_{\vec{h},\mathbf{s}}(\mathrm{x})>0$ then

$$\int_{-b}^b\xi_{\vec{h},\mathbf{s}}(x)\ell(x_1)\nu_1\big(\mathrm{d}x_1\big)=\lambda_{\vec{h},\mathbf{s}}(\mathrm{x})\int Q(t_1)G_{\mathbf{s}}(\mathrm{t},\mathrm{x})W(\mathrm{d}t),$$

with $Q(\cdot)=\lambda^{-1}_{\vec{h},\mathbf{s}}(\mathrm{x})\int_{-b}^b\mathfrak{h}_{s_1}^{-1/2}\mathcal{K}\Big(\frac{\cdot-x_1}{\mathfrak{h}_{s_1}}\Big)\ell(x_1)1_{\Lambda_{\mathbf{s}}[\vec{h}]}(x_1,\mathrm{x})\nu_1(\mathrm{d}x_1)\in\mathcal{Q}_{\mathrm{x},\mathbf{s}}$, where $\mathcal{Q}_{\mathrm{x},\mathbf{s}}$ is defined in Lemma 3. Then, (3.29) follows from (3.28).

Below we will prove that $\varsigma_{\mathbf{s}}(\mathrm{x}):=\sup_{Q\in\mathcal{Q}_{\mathbf{s},\mathrm{x}}}\varsigma_{\mathbf{s}}\big(Q,\mathrm{x}\big)$ is a random variable. This is important because its definition uses the supremum over $\mathcal{Q}_{\mathbf{s},\mathrm{x}}$ which is not countable.

We get from (3.29) for any $\vec{h}\in\mathrm{H}$ and $\mathbf{s}\in\mathbb{N}^d$ in view of Fubini theorem

$$\begin{aligned}\big\|\xi_{\vec{h},\mathbf{s}}\big\|_r^r&=\int_{(-b,b)^{d-1}}\int_b^b\big|\xi_{\vec{h},\mathbf{s}}(x_1,\mathrm{x})\big|^r\nu_1(\mathrm{d}x_1)\nu_{d-1}(\mathrm{d}\mathrm{x})\leq\int_{(-b,b)^d}\lambda^r_{\vec{h},\mathbf{s}}(\mathrm{x})\varsigma^r_{\mathbf{s}}(\mathrm{x})\nu_{d-1}(\mathrm{d}\mathrm{x})\\&=\int_{(-b,b)^d}\varsigma^r_{\mathbf{s}}(\mathrm{x})\bigg[\int_{-b}^b1_{\Lambda_{\mathbf{s}}\left[\vec{h}\right]}(x)\nu_1\big(\mathrm{d}x_1\big)\bigg]^{\tau}\nu_{d-1}(\mathrm{d}\mathrm{x}).\end{aligned}$$

Taking into account that $\tau<1$ and applying Hölder inequality to the outer integral we get

$$\big\|\xi_{\vec{h},\mathbf{s}}\big\|_r^r \le \nu_d^{\tau}\Big(\Lambda_{\mathbf{s}}\big[\vec{h}\big]\Big)\bigg\{\int_{(-b,b)^d}\varsigma_{\mathbf{s}}^{\frac{r}{1-\tau}}(\mathrm{x})\nu_{d-1}(\mathrm{dx})\bigg\}^{1-\tau},\quad \forall \mathbf{s}\in\mathbb{N}^d. \tag{3.30}$$

If $d=1$ putting $G_{\mathbf{s}}(\mathrm{t},\mathrm{x})\equiv 1$ in (3.16), we obtain using the same arguments

$$\big\|\xi_{h_1,s_1}\big\|_r^r \le \nu_d^{\tau}\Big(\Lambda_{s_1}\big[h_1\big]\Big)\varsigma_{s_1},\qquad \varsigma_{s_1}=\sup_{Q\in\mathcal{Q}_{s_1}}\varsigma_{s_1}(Q). \tag{3.31}$$

1^0**b1**. Let us prove some bounds used in the sequel. Let $S\in\mathbb{N}$ be the number satisfying $e^{-1}<\mathfrak{h}^d e^{-S}\mathcal{A}^4\le 1$, and set $\mathcal{S}_d=\{0,1,\dots S\}^d$ and $\bar{\mathcal{S}}_d=\mathbb{N}^d\setminus\mathcal{S}_d$. If such S does not exist we will assume that $\mathcal{S}_d=\emptyset$ and later on the supremum over empty set is assumed to be 0.

Set also $\mathcal{S}_d^*=\{\mathbf{s}\in\mathbb{N}^d:\ \mathcal{A}^4V_{\mathbf{s}}\le 1\}$, where, recall, $V_{\mathbf{s}}=\prod_{j=1}^d\mathfrak{h}_{\mathbf{s}_j}$. Note that $V_{\mathbf{s}}\le\mathfrak{h}^d e^{-S}\le\mathcal{A}^{-4}$ for any $\mathbf{s}\in\bar{\mathcal{S}}_d$ and, therefore,

$$\bar{\mathcal{S}}_d^*:=\mathbb{N}^d\setminus\mathcal{S}_d^*\subseteq\mathcal{S}_d. \tag{3.32}$$

Putting for brevity $\mathrm{r}=r_{\mathcal{A}}(\vec{h})$, we have for any $\mathbf{s}\in\mathbb{N}^d$ and any $\vec{h}\in\mathbb{B}(\mathcal{A})$

$$\big(V_{\mathbf{s}}\big)^{-\frac{p\mathrm{r}}{2(\mathrm{r}-p)}}\nu_d\Big(\Lambda_{\mathbf{s}}\big[\vec{h}\big]\Big)\le\sum_{\mathbf{k}\in\mathbb{N}^d}\big(V_{\mathbf{k}}\big)^{-\frac{p\mathrm{r}}{2(\mathrm{r}-p)}}\nu_d\Big(\Lambda_{\mathbf{k}}\big[\vec{h}\big]\Big)=\Big\|V_{\vec{h}}^{-\frac{1}{2}}\Big\|_{\frac{p\mathrm{r}}{\mathrm{r}-p}}^{\frac{p\mathrm{r}}{\mathrm{r}-p}}\le\mathcal{A}^{\frac{p\mathrm{r}}{\mathrm{r}-p}}.$$

The last inequality follows from the definition of $r_{\mathcal{A}}(\vec{h})$.

Taking into account that $\frac{p\mathrm{r}}{\mathrm{r}-p}>p$ and that $V_{\mathbf{s}}<1$ we get in view of the definition of $\mathcal{S}_d^*$

$$\nu_d\Big(\Lambda_{\mathbf{s}}\big[\vec{h}\big]\Big)\le V_{\mathbf{s}}^{\frac{p}{4}},\qquad\forall\vec{h}\in\mathbb{B}(\mathcal{A}),\ \ \forall\mathbf{s}\in\mathcal{S}_d^*. \tag{3.33}$$

1^0**b2**. Set $\varsigma(\mathrm{x})=\sup_{\mathbf{s}\in\mathcal{S}_d}\varsigma_{\mathbf{s}}(\mathrm{x})$ and let $d\ge 2$.

We have in view of (3.30) and (3.32) for any $\vec{h}\in\mathrm{H}$

$$\begin{aligned}\sum_{\mathbf{s}\in\bar{\mathcal{S}}_d^*}\big\|\xi_{\vec{h},\mathbf{s}}\big\|_r^r &\le \sum_{\mathbf{s}\in\mathcal{S}_d}\big\|\xi_{\vec{h},\mathbf{s}}\big\|_r^r\le\bigg\{\int_{(-b,b)^{d-1}}\varsigma^{\frac{r}{1-\tau}}(\mathrm{x})\nu_{d-1}(\mathrm{dx})\bigg\}^{1-\tau}\sum_{\mathbf{s}\in\mathbb{N}^d}\nu_d^{\tau}\Big(\Lambda_{\mathbf{s}}\big[\vec{h}\big]\Big)\\ &\le \mathcal{L}\bigg\{\int_{(-b,b)^{d-1}}\varsigma^{\frac{r}{1-\tau}}(\mathrm{x})\nu_{d-1}(\mathrm{dx})\bigg\}^{1-\tau}.\end{aligned}\tag{3.34}$$

To get the last inequality we have used that $\mathrm{H}\subset\mathbb{H}_d(\tau,\mathcal{L})$.

Writing $\tau=\tau^2+\tau(1-\tau)$ and using the bound (3.33) we get in view of (3.30)

$$\sum_{\mathbf{s}\in\mathcal{S}_d^*}\big\|\xi_{\vec{h},\mathbf{s}}\big\|_r^r\ \le\ \sum_{\mathbf{s}\in\mathcal{S}_d^*}\nu_d^{\tau^2}\Big(\Lambda_{\mathbf{s}}\big[\vec{h}\big]\Big)V_{\mathbf{s}}^{\frac{\tau(1-\tau)p}{4}}\bigg\{\int_{(-b,b)^{d-1}}\varsigma_{\mathbf{s}}^{\frac{r}{1-\tau}}(\mathrm{x})\nu_{d-1}(\mathrm{dx})\bigg\}^{1-\tau}.$$

Applying Hölder inequality with exponents $1/\tau$ and $1/(1-\tau)$ we get

$$\begin{aligned}&\sum_{\mathbf{s}\in\mathcal{S}_d^*}\nu_d^{\tau^2}\Big(\Lambda_{\mathbf{s}}\big[\vec{h}\big]\Big)V_{\mathbf{s}}^{\frac{\tau(1-\tau)p}{4}}\bigg\{\int_{(-b,b)^{d-1}}\varsigma_{\mathbf{s}}^{\frac{r}{1-\tau}}(\mathrm{x})\nu_{d-1}(\mathrm{dx})\bigg\}^{1-\tau}\\ &\le\bigg[\sum_{\mathbf{s}\in\mathbb{N}^d}\nu_d^{\tau}\Big(\Lambda_{\mathbf{s}}\big[\vec{h}\big]\Big)\bigg]^{\tau}\bigg[\sum_{\mathbf{s}\in\mathbb{N}^d}V_{\mathbf{s}}^{\frac{\tau p}{4}}\int_{(-b,b)^{d-1}}\varsigma_{\mathbf{s}}^{\frac{r}{1-\tau}}(\mathrm{x})\nu_{d-1}(\mathrm{dx})\bigg]^{1-\tau}\\ &\le\mathcal{L}^{\tau}\bigg[\sum_{\mathbf{s}\in\mathbb{N}^d}V_{\mathbf{s}}^{\frac{\tau p}{4}}\int_{(-b,b)^{d-1}}\varsigma_{\mathbf{s}}^{\frac{r}{1-\tau}}(\mathrm{x})\nu_{d-1}(\mathrm{dx})\bigg]^{1-\tau}.\end{aligned}\tag{3.35}$$

To get the last inequality we have used once again that $\mathrm{H} \subset \mathbb{H}_d(\tau, \mathcal{L})$.

We deduce from (3.26), (3.34) and (3.35) that for any $\vec{h} \in \mathrm{H}$

$$\begin{aligned}\zeta^r_{\vec{h}}(r) \quad &\leq \quad \mathcal{L}\bigg\{\int_{(-b,b)^{d-1}} \varsigma^{\frac{r}{1-\tau}}(\mathrm{x})\nu_{d-1}(\mathrm{dx})\bigg\}^{1-\tau} \\ &+ \quad \mathcal{L}^\tau\bigg[\sum_{\mathbf{s}\in\mathbb{N}^d} V_{\mathbf{s}}^{\frac{\tau p}{4}} \int_{(-b,b)^{d-1}} \varsigma_{\mathbf{s}}^{\frac{r}{1-\tau}}(\mathrm{x})\nu_{d-1}(\mathrm{dx})\bigg]^{1-\tau}.\end{aligned}$$

Noting that the right hand side of the obtained inequality is independent of $\vec{h}$ we get

$$\begin{aligned}\zeta(r) \quad &\leq \quad \mathcal{L}^{\frac{1}{r}}\bigg\{\int_{(-b,b)^{d-1}} \varsigma^{\frac{r}{1-\tau}}(\mathrm{x})\nu_{d-1}(\mathrm{dx})\bigg\}^{\frac{1-\tau}{r}} \\ &+ \quad \mathcal{L}^{\frac{\tau}{r}}\bigg[\sum_{\mathbf{s}\in\mathbb{N}^d} V_{\mathbf{s}}^{\frac{\tau p}{4}} \int_{(-b,b)^{d-1}} \varsigma_{\mathbf{s}}^{\frac{r}{1-\tau}}(\mathrm{x})\nu_{d-1}(\mathrm{dx})\bigg]^{\frac{1-\tau}{r}}.\end{aligned}$$

Hence, applying Jensen inequality and Fubini theorem one has for any $d \geq 2$

$$\begin{aligned}\mathbb{E}\big\{\zeta(r)\big\} \quad &\leq \quad \mathcal{L}^{\frac{1}{r}}\bigg\{\int_{(-b,b)^{d-1}} \mathbb{E}\Big(\varsigma^{\frac{r}{1-\tau}}(\mathrm{x})\Big)\nu_{d-1}(\mathrm{dx})\bigg\}^{\frac{1-\tau}{r}} \\ &+ \quad \mathcal{L}^{\frac{\tau}{r}}\bigg[\sum_{\mathbf{s}\in\mathbb{N}^d} V_{\mathbf{s}}^{\frac{\tau p}{4}} \int_{(-b,b)^{d-1}} \mathbb{E}\Big(\varsigma_{\mathbf{s}}^{\frac{r}{1-\tau}}(\mathrm{x})\Big)\nu_{d-1}(\mathrm{dx})\bigg]^{\frac{1-\tau}{r}} \\ &\leq \quad \mathcal{L}^{\frac{1}{r}}\big[1\vee(2b)^{d-1}\big] \sup_{\mathrm{x}\in(-b,b)^{d-1}} \Big\{\mathbb{E}\Big(\varsigma^{\frac{r}{1-\tau}}(\mathrm{x})\Big)\Big\}^{\frac{1-\tau}{r}} \\ &+ \quad \mathcal{L}^{\frac{\tau}{r}}\big[1\vee(2b)^{d-1}\big]\big(1-e^{-\frac{\tau p}{4}}\big)^{\frac{\tau-1}{r}} \sup_{\mathbf{s}\in\mathbb{N}^d}\sup_{\mathrm{x}\in(-b,b)^{d-1}} \Big\{\mathbb{E}\Big(\varsigma_{\mathbf{s}}^{\frac{r}{1-\tau}}(\mathrm{x})\Big)\Big\}^{\frac{1-\tau}{r}}. \qquad (3.36)\end{aligned}$$

Here we have also used that $V_{\mathbf{s}} \leq \prod_{j=1}^d e^{-s_j-2}$ and that $(1-\tau)/r < 1$.

If $d = 1$ repeating previous computations we obtain from (3.26) and (3.31)

$$\mathbb{E}\big\{\zeta(r)\big\} \leq \mathcal{L}^{\frac{1}{r}}\mathbb{E}\varsigma + \mathcal{L}^{\frac{\tau}{r}}\big(1-e^{-\frac{\tau p}{4}}\big)^{\frac{\tau-1}{r}} \sup_{\mathbf{s}\in\mathbb{N}} \Big[\mathbb{E}\Big(\varsigma_{\mathbf{s}}^{\frac{r}{1-\tau}}\Big)\Big]^{\frac{1-\tau}{r}}. \qquad (3.37)$$

In what follows x is assumed to be fixed that allows us not to separate cases $d = 1$ and $d \geq 2$.

1^0**b3**. Let $\mathrm{x} \in (-b,b)^{d-1}$ be fixed. First let us bound from above

$$\mathbb{E}\varsigma_s(\mathrm{x}) := \mathbb{E}\Big\{\sup_{Q\in\mathcal{Q}_{\mathbf{s},\mathrm{x}}} \varsigma_{\mathbf{s}}\big(Q,\mathrm{x}\big)\Big\}, \;\; \mathbf{s}\in\mathbb{N}^d, \qquad \mathbb{E}\varsigma(\mathrm{x}) := \mathbb{E}\Big\{\sup_{\mathbf{s}\in\mathcal{S}_d}\sup_{Q\in\mathcal{Q}_{\mathbf{s},\mathrm{x}}} \varsigma_{\mathbf{s}}\big(Q,\mathrm{x}\big)\Big\}$$

Note that $\varsigma_{\mathbf{s}}\big(Q,\mathrm{x}\big)$ is zero-mean gaussian random function on $\mathcal{Q}_{\mathbf{s},\mathrm{x}}$. Our objective now is to show that the assertion II of Lemma 1 is applicable with $Z_t = \varsigma_{\mathbf{s}}\big(Q,\mathrm{x}\big), t = Q$, and $\mathbb{T} = \mathcal{Q}_{\mathbf{s},\mathrm{x}}$.

Note that the intrinsic semi-metric of $\varsigma_{\mathbf{s}}\big(Q,\mathrm{x}\big)$ is given by

$$\rho^2\big(Q,\widetilde{Q}\big) = \int G^2_{\mathbf{s}}(\mathrm{t},\mathrm{x})\Big[Q(t_1)-\widetilde{Q}(t_1)\Big]^2\nu_d(\mathrm{dt}), \quad Q,\widetilde{Q}\in\mathcal{Q}_{\mathbf{s},\mathrm{x}}.$$

Noting that $\int_{\mathbb{R}^{d-1}} G^2_{\mathbf{s}}(\mathrm{t},\mathrm{x})\nu_d(\mathrm{dt}) = \|\mathcal{K}\|_2^{2d-2}$ for any $\mathrm{x}\in(-b,b)^{d-1}$, we get

$$\rho\big(Q,\widetilde{Q}\big) = \|\mathcal{K}\|_2^{d-1}\|Q-\widetilde{Q}\|_2, \quad \forall Q,\widetilde{Q}\in\mathcal{Q}_{\mathbf{s},\mathrm{x}}.$$

Below we show that $\big(\mathcal{Q}_{\mathbf{s},\mathrm{x}},\|\cdot\|_2\big)$ is totally bounded metric space and, moreover, the corresponding Dudley's integral is finite. The latter fact allows us to assert that $\varsigma_{\mathbf{s}}(\cdot,\mathrm{x})$ is almost surely continuous on $\mathcal{Q}_{\mathbf{s},\mathrm{x}}$ that implies the measurability of $\varsigma_{\mathbf{s}}(\mathrm{x})$ as well as $\varsigma(\mathrm{x})$. We obviously have

$$\mathfrak{E}_{\rho,\mathcal{Q}_{\mathbf{s},\mathrm{x}}}(\delta) \leq \mathfrak{E}_{\|\cdot\|_2,\mathcal{Q}_{\mathbf{s},\mathrm{x}}}\big(\|\mathcal{K}\|_2^{1-d}\delta\big), \quad \forall\delta>0, \tag{3.38}$$

and, therefore,

$$D_{\mathcal{Q}_{\mathbf{s},\mathrm{x}},\rho} := 4\sqrt{2}\int_0^{2^{-1}\sigma_{\mathbf{s}}}\sqrt{\mathfrak{E}_{\rho,\mathcal{Q}_{\mathbf{s},\mathrm{x}}}(\delta)}\mathrm{d}\delta \leq 4\sqrt{2}\|\mathcal{K}\|_2^{d-1}\int_0^{\tilde{\sigma}_{\mathbf{s}}}\sqrt{\mathfrak{E}_{\|\cdot\|_2,\mathcal{Q}_{\mathbf{s},\mathrm{x}}}(\delta)}\mathrm{d}\delta, \tag{3.39}$$

where $\tilde{\sigma}_{\mathbf{s}} = 2^{-1}\sigma_{\mathbf{s}}\|\mathcal{K}\|_2^{1-d}$ and

$$\sigma_{\mathbf{s}} := \Big[\sup_{Q\in\mathcal{Q}_{\mathbf{s},\mathrm{x}}}\mathbb{E}\big\{\varsigma^2_{\mathbf{s}}(Q,\mathrm{x})\big\}\Big]^{\frac{1}{2}} = \|\mathcal{K}\|_2^{d-1}\sup_{Q\in\mathcal{Q}_{\mathbf{s},\mathrm{x}}}\|Q\|_2.$$

We start with bounding from above the quantity $\sigma_{\mathbf{s}}$.

Recall that $\mu^{-1} = \mathfrak{q}^{-1}+\tau r^{-1}$. Applying Young inequality we have

$$\|Q\|_2 \leq \lambda^{-1}_{\vec{h},\mathbf{s}}(\mathrm{x})\mathfrak{h}_{s_1}^{1-\frac{1}{\mu}}\bigg[\int_{-b}^{b}\big|\ell(x_1)\big|^{\mu}1_{\Lambda_{\mathbf{s}}[\vec{h}]}(x_1,\mathrm{x})\nu_1(\mathrm{d}x_1)\bigg]^{\frac{1}{\mu}}\|\mathcal{K}\|_{\frac{2\mu}{3\mu-2}}.$$

Applying Hölder inequality to the integral in right hand side of the latter inequality and taking into account that $\ell\in\mathbb{B}_{\mathfrak{q}}$ we get

$$\bigg[\int_{-b}^{b}\big|\ell(x_1)\big|^{\mu}1_{\Lambda_{\mathbf{s}}[\vec{h}]}(x_1,\mathrm{x})\nu_1(\mathrm{d}x_1)\bigg]^{\frac{1}{\mu}} \leq \bigg[\int_{-b}^{b}1_{\Lambda_{\mathbf{s}}[\vec{h}]}(x_1,\mathrm{x})\nu_1(\mathrm{d}x_1)\bigg]^{\frac{1}{\mu}-\frac{1}{q}} = \lambda_{\vec{h},\mathbf{s}}(\mathrm{x}). \tag{3.40}$$

Thus we obtain

$$\sigma_{\mathbf{s}} \leq \|\mathcal{K}\|_2^{d-1}\|\mathcal{K}\|_{\frac{2\mu}{3\mu-2}}\mathfrak{h}_{s_1}^{\frac{1-\tau}{r}}. \tag{3.41}$$

Putting $\sigma^*_{\mathbf{s}} = 2^{-1}\|\mathcal{K}\|_{\frac{2\mu}{3\mu-2}}\mathfrak{h}_{s_1}^{\frac{1-\tau}{r}}$ we deduce from (3.39) and (3.41)

$$D_{\mathcal{Q}_{\mathbf{s},\mathrm{x}},\rho} \leq 4\sqrt{2}\|\mathcal{K}\|_2^{d-1}\int_0^{\sigma^*_{\mathbf{s}}}\sqrt{\mathfrak{E}_{\|\cdot\|_2,\mathcal{Q}_{\mathbf{s},\mathrm{x}}}(\delta)}\mathrm{d}\delta. \tag{3.42}$$

Now let us bound from above $\mathbb{E}\big\{\sup_{Q\in\mathcal{Q}_{\mathbf{s},\mathrm{x}}}\varsigma_{\mathbf{s}}\big(Q,\mathrm{x}\big)\big\}$.

Recall that $\Omega = \Big\{\{\omega_1,\omega_2\}: \ \omega_1<1/2<\omega_2, \ [\omega_1,\omega_2]\subset(1/\mu-1/2,1)\Big\}$. Note that the condition $\omega_1>1/\mu-1/2$ implies $1/2-\omega_1<(1-\tau)r^{-1}$ and, therefore

$$\mathfrak{h}_{s_1}^{\frac{1-\tau}{r}} < \mathfrak{h}_{s_1}^{\frac{1}{2}-\omega_1} < \mathfrak{h}_{s_1}^{\frac{1}{2}-\omega_2}$$

since $\mathfrak{h}_{s_1}\leq\mathfrak{h}\leq 1$. It yields that $\big(0,\sigma^*_{\mathbf{s}}\big]\subset\Big(0,R_\mu\,\mathfrak{h}_{s_1}^{\frac{1}{2}-\omega_1}\Big]\subset\Big(0,R_\mu\,\mathfrak{h}_{s_1}^{\frac{1}{2}-\omega_2}\Big]$, since $R_\mu\geq 2^{-1}\|\mathcal{K}\|_{\frac{2\mu}{3\mu-2}}$. Hence Lemma 3 is applicable to the computation of the integral in the right hand side of (3.42).

Recall that $\lambda^*(\cdot,\cdot)$ is defined in Section 3.2.1 and introduce the following notations: $A^2(\omega)=\lambda^*\big(\omega,\mu\big)R_\mu^{\frac{1}{\omega}}\;\mathfrak{h}_{s_1}^{\frac{1}{2\omega}-1}$, $\delta_0=\mathfrak{h}_{s_1}^{\frac{1}{2}}$ and note that $\delta_0<\sigma^*_{\mathbf{s}}$. We get in view of Lemma 3

$$\begin{aligned}\int_0^{\sigma^*_{\mathbf{s}}}\sqrt{\mathfrak{E}_{||\cdot||_2,\mathcal{Q}_{\mathbf{s},\mathrm{x}}}(\delta)}\mathrm{d}\delta &= \int_0^{\delta_0}\sqrt{\mathfrak{E}_{||\cdot||_2,\mathcal{Q}_{\mathbf{s},\mathrm{x}}}(\delta)}\mathrm{d}\delta+\int_{\delta_0}^{\sigma^*_{\mathbf{s}}}\sqrt{\mathfrak{E}_{||\cdot||_2,\mathcal{Q}_{\mathbf{s},\mathrm{x}}}(\delta)}\mathrm{d}\delta\\ &\leq A(\omega_2)\big(1-[2\omega_2]^{-1}\big)\delta_0^{1-\frac{1}{2\omega_2}}+A(\omega_1)\big([2\omega_1]^{-1}-1\big)\delta_0^{1-\frac{1}{2\omega_1}}\\ &= \sqrt{\lambda^*\big(\omega_2,\mu\big)}\big(1-[2\omega_2]^{-1}\big)R_\mu^{\frac{1}{2\omega_2}}+\sqrt{\lambda^*\big(\omega_1,\mu\big)}\big([2\omega_1]^{-1}-1\big)R_\mu^{\frac{1}{2\omega_1}}.\end{aligned}$$

It yields together with (3.42) $D_{\mathcal{Q}_{\mathbf{s},\mathrm{x}},\rho}\leq C_\mu$, where, recall,

$$C_\mu=4\sqrt{2}\|\mathcal{K}\|_2^{d-1}\inf_{\{\omega_1,\omega_2\}\in\Omega}\bigg[\sqrt{\lambda^*\big(\omega_2,\mu\big)}\big(1-[2\omega_2]^{-1}\big)R_\mu^{\frac{1}{2\omega_2}}+\sqrt{\lambda^*\big(\omega_1,\mu\big)}\big([2\omega_1]^{-1}-1\big)R_\mu^{\frac{1}{2\omega_1}}\bigg].$$

Applying the assertion II of Lemma 1 we get

$$\mathbb{E}\varsigma_{\mathbf{s}}(\mathrm{x})=\mathbb{E}\big\{\sup_{Q\in\mathcal{Q}_{\mathbf{s},\mathrm{x}}}\varsigma_{\mathbf{s}}\big(Q,\mathrm{x}\big)\big\}\leq C_\mu. \tag{3.43}$$

We obtain from (3.41) that

$$\sigma_\varsigma:=\sup_{\mathbf{s}\in\mathcal{S}_d}\sup_{Q\in\mathcal{Q}_{\mathbf{s},\mathrm{x}}}\sqrt{\mathbb{E}\varsigma^2_{\mathbf{s}}\big(Q,\mathrm{x}\big)}=:\sup_{\mathbf{s}\in\mathcal{S}_d}\sigma_{\mathbf{s}}\leq\|\mathcal{K}\|_2^{d-1}\|\mathcal{K}\|_{\frac{2\mu}{3\mu-2}}\mathfrak{h}^{\frac{1-\tau}{r}}. \tag{3.44}$$

Applying the assertion I of Lemma 1 we obtain in view of (3.43) for any $z>0$

$$\mathbb{P}\Big\{\varsigma_{\mathbf{s}}(\mathrm{x})\geq C_\mu+z\Big\}\leq e^{-\frac{z^2}{2\sigma^2_{\mathbf{s}}}}\leq e^{-\frac{z^2}{2\sigma^2_\varsigma}}. \tag{3.45}$$

Set $T=C_\mu+\sqrt{2e^r}\|\mathcal{K}\|_2^{d-1}\|\mathcal{K}\|_{\frac{2\mu}{3\mu-2}}$ we obtain using (3.45)

$$\begin{aligned}\mathbb{E}\varsigma(\mathrm{x}) &\leq T+\int_0^\infty\mathbb{P}\big\{\varsigma(\mathrm{x})\geq T+y\big\}\mathrm{d}y\leq T+(S+1)^d\int_0^\infty e^{-\frac{(U-C_\mu+y)^2}{2\sigma^2_\varsigma}}\mathrm{d}y\\ &\leq T+\sqrt{8\pi}\|\mathcal{K}\|_2^{d-1}\|\mathcal{K}\|_{\frac{2\mu}{3\mu-2}}(S+1)^d\exp\Big\{-e^r\mathfrak{h}^{\frac{2(\tau-1)}{r}}\Big\}.\end{aligned}$$

Taking into account that $(S+1)^d\leq\big[4\ln(\mathcal{A})\big]^d$ in view of the definition of S and that

$$\inf_{r>0}e^r\mathfrak{h}^{\frac{2(\tau-1)}{r}}=e^{2\sqrt{2(1-\tau)|\ln(\mathfrak{h})|}},$$

we obtain

$$\begin{aligned}\mathbb{E}\varsigma(\mathrm{x}) &\leq T+\sqrt{8\pi}\|\mathcal{K}\|_2^{d-1}\|\mathcal{K}\|_{\frac{2\mu}{3\mu-2}}\big[4\ln(\mathcal{A})\big]^d e^{e^{-2\sqrt{2(1-\tau)|\ln(\mathfrak{h})|}}}.\\ &\leq C_\mu+4^d\big(\sqrt{2e^r}+\sqrt{8\pi}\big)\|\mathcal{K}\|_2^{d-1}\|\mathcal{K}\|_{\frac{2\mu}{3\mu-2}}=\widetilde{C}_\mu.\end{aligned} \tag{3.46}$$

The last inequality follows from the relation (2.2) and the definition of T.

1^0**b4**. Applying the assertion I of Lemma 1 we obtain in view of (3.44) for any $z>0$

$$\mathbb{P}\Big\{\varsigma(\mathrm{x})\geq \widetilde{C}_\mu+z\Big\}\leq e^{-\frac{z^2}{2\sigma_\varsigma^2}}.$$

It yields together with (3.44)

$$\begin{aligned}\mathbb{E}\Big(\varsigma^{\frac{r}{1-\tau}}(\mathrm{x})\Big) &= \widetilde{C}_\mu^{\frac{r}{1-\tau}}+\frac{r}{1-\tau}\int_0^\infty\big(z+\widetilde{C}_\mu\big)^{\frac{r+\tau-1}{1-\tau}}\mathbb{P}\big\{\varsigma(\mathrm{x})\geq z+\widetilde{C}_\mu\big\}\mathrm{d}z\\ &\leq \widetilde{C}_\mu^{\frac{r}{1-\tau}}+\widehat{C}_\mu^{\frac{r}{1-\tau}}. \end{aligned}\tag{3.47}$$

Here recall

$$\widehat{C}_\mu=\bigg[\frac{r}{1-\tau}\int_0^\infty\big(u+\widetilde{C}_\mu\big)^{\frac{r+\tau-1}{1-\tau}}\exp\Big\{-u^2\Big[2|\mathcal{K}\|_2^{d-1}\|\mathcal{K}\|_{\frac{2\mu}{3\mu-2}}\Big]^{-1}\Big\}\mathrm{d}u\bigg]^{\frac{1-\tau}{r}}.$$

Similarly we deduce from (3.44) and (3.45)

$$\mathbb{E}\Big(\varsigma_{\mathbf{s}}^{\frac{r}{1-\tau}}(\mathrm{x})\Big)\leq C_\mu^{\frac{r}{1-\tau}}+\widehat{C}_\mu^{\frac{r}{1-\tau}}\leq \widetilde{C}_\mu^{\frac{r}{1-\tau}}+\widehat{C}_\mu^{\frac{r}{1-\tau}},\quad \forall\mathbf{s}\in\mathbb{N}^d.\tag{3.48}$$

Noting that the bounds in (3.47) and (3.48) are independent of x and $\mathbf{s}$ we get in view of (3.36)

$$\mathbb{E}\big\{\zeta(r)\big\}\leq\big[1\vee(2b)^{d-1}\big]\big[\mathcal{L}^{\frac{1}{r}}+\mathcal{L}^{\frac{\tau}{r}}\big(1-e^{-\frac{\tau p}{4}}\big)^{\frac{\tau-1}{r}}\big]\big[\widetilde{C}_\mu+\widehat{C}_\mu\big].$$

This proves (3.14) with $V(r)=\big[1\vee(2b)^{d-1}\big]\big[\mathcal{L}^{\frac{1}{r}}+\mathcal{L}^{\frac{\tau}{r}}\big(1-e^{-\frac{\tau p}{4}}\big)^{\frac{\tau-1}{r}}\big]\big[\widetilde{C}_\mu+\widehat{C}_\mu\big]$.

1^0**c**. Remembering that $C_2(r)=T+e^r\sqrt{2(1+q)}\big(r\sqrt{e}\big)^d\|\mathcal{K}\|^d_{\frac{2r}{r+2}}$ we obtain, applying the assertion I of Lemma 1 available in view of (3.14) and (3.25)

$$\mathbb{P}\Big\{\zeta(r)\geq C_2(r)+z\Big\}\leq e^{-e^r}e^{-qe^r\mathfrak{h}^{\frac{2d}{r}}}\exp\Big\{-\big[2\big(r\sqrt{e}\big)^d\|\mathcal{K}\|^d_{\frac{2r}{r+2}}\big]^{-1}z^2\Big\},\quad\forall z\geq 0.$$

Taking into account that $e^r\mathfrak{h}^{\frac{2d}{r}}\leq e^{2\sqrt{2d|\ln(\mathfrak{h})|}}$ for any $r>0$ we come to (3.24).

2^0. We deduce from (3.23) and (3.24) that

$$\mathbb{E}\big[\zeta(r)-C_2(r)\big]_+^q\leq\sqrt{(\pi/2)}e^{-e^r}\big[\big(r\sqrt{e}\big)^d\|\mathcal{K}\|^d_{\frac{2r}{r+2}}\big]^{\frac{q}{2}}e^{-qe^{2\sqrt{2d|\ln(\mathfrak{h})|}}}\boldsymbol{\gamma}_{q+1},$$

where recall $\boldsymbol{\gamma}_{q+1}$ is the $(q+1)$-th moment of the standard normal distribution. This yields together with (3.22)

$$\mathbb{E}\bigg(\sup_{\vec{h}\in\mathrm{H}}\Big[\big\|\xi_{\vec{h}}\big\|_p-\psi_r(\vec{h})\Big]_+\bigg)^q\leq\bigg[C_4\mathcal{A}e^{-e^{2\sqrt{2d|\ln(\mathfrak{h})|}}}\bigg]^q,$$

and the assertion of the theorem follows.

∎

3.3. Proof of Theorem 3

3.3.1. Constants

Let $\mathbf{c}(d)$ be the constant appearing in $(2,2)$-strong maximal inequality, see Folland (1999). Set

$$\begin{aligned}\sigma_* &= \sqrt{2^{d+1}a^d\|K\|_\infty\|K\|_1\mathbf{c}(d)}(2b)^{\frac{d(p-1)}{p}};\\ C_5 &= \Big[\sqrt{8\pi}\sigma_*^{q-1}\boldsymbol{\gamma}_{q+1}\Big]^{\frac{1}{q}}\sum_{r=d+1}^{\infty}\sum_{l=1}^{\infty}e^{-2^l e^r}.\end{aligned}$$

For any $r\in\mathbb{N}^*, r>d$ put $\gamma_r=\frac{d}{2}+\frac{d}{2pr}$ and let $\mathfrak{D}$ denote the unit disc in $\mathbb{R}^d$. Set

$$\begin{aligned}T(r) &= \big[\sigma_*/2\big]\vee\Big[(d/2+1)^dT^*(r)+\|\mathcal{K}\|_1^d(2b)^{1/p}\Big];\\ T^*(r) &= 2^{-d+1}\Big[L(a+2)^d\int\mathfrak{z}^{-d-\gamma_r+\lfloor\gamma_r\rfloor+1}1_{\mathfrak{D}}(\mathfrak{z})\mathrm{d}\mathfrak{z}+C(K)\int\mathfrak{z}^{-d-\gamma_r+\lfloor\gamma_r\rfloor}1_{\overline{\mathfrak{D}}}(\mathfrak{z})\mathrm{d}\mathfrak{z}\Big],\end{aligned}$$

where $C(K)=\sup_{|\mathbf{n}|=\lfloor d/2\rfloor}\big\|D^{\mathbf{n}}K\big\|_1$. Note that $\gamma_r\neq\lfloor\gamma_r\rfloor$ and, therefore, both integrals in the definition of $T^*(r)$ are finite.

Let $\lambda_d^*(r)=\lambda_d\big(\gamma_r,1,1,[-a-b,a+b]^d\big)$, where the quantity $\lambda_k(\cdot,\cdot,\cdot,\cdot), k\in\mathbb{N}^*$, is defined in Lemma 2. Set finally

$$C_2^*(r)=8\sqrt{2\lambda_d^*(r)}\big[T(r)\big]^{d/2\gamma_r}\big(\sigma_*/2\big)^{\frac{1}{2pr}}+4\sqrt{qe^r}\sigma_*.$$

3.3.2. Auxiliary lemma

For any $l\in\mathbb{N}^*$ and any $r\in\mathbb{N}^*$ satisfying $r>d$ put

$$\mathrm{H}_{l,r}=\Big\{\vec{h}\in\mathrm{H}:\ 2^{l-1}\mathfrak{h}^{-\frac{d}{2}}\leq\big\|h^{-\frac{d}{2}}\big\|_{p+\frac{1}{r}}<2^l\mathfrak{h}^{-\frac{d}{2}}\Big\},$$

and introduce

$$\mathfrak{Q}_{l,r}=\Big\{Q:\mathbb{R}^d\to\mathbb{R}:\ Q(\cdot)=\int_{(-b,b)^d}K_{\vec{h}}(\cdot-x)\vartheta(x)\nu_d(\mathrm{d}x),\ \vartheta\in\mathbb{B}_{\mathrm{q},d},\ \vec{h}\in\mathrm{H}_{l,r}\Big\},$$

where, $\mathbb{B}_{\mathrm{q},d}=\{\vartheta:(-b,b)^d\to\mathbb{R}:\ \|\vartheta\|_{\mathrm{q}}\leq1\}$ and $1/\mathrm{q}=1-1/p$.

Lemma 4. *For any $r,l\in\mathbb{N}^*$, $r>d$ and any $\delta\in\Big(0,T(r)\big(2^l\mathfrak{h}^{-\frac{d}{2}}\big)^{\frac{2\gamma_r}{d}}\Big]$ one has*

$$\mathfrak{E}_{\mathfrak{Q}_{l,r},\|\cdot\|_2}(\delta)\leq\lambda_d^*(r)\big[T(r)\big]^{d/\gamma_r}\big(2^l\mathfrak{h}^{-\frac{d}{2}}\big)^2\delta^{-d/\gamma_r}.$$

3.3.3. Preliminary remarks on the proof of Theorem 3.

The goal of this paragraph is to discuss the main technical tools involved in the proof of the theorem. In particular we explain the role of the isotropy and the condition $p\in[1,2]$ in our considerations.

We proceed similarly to the proof of Theorem 2. Using duality arguments we have

$$\sup_{\vec{h}\in \mathrm{H}_{l,r}} \big\|\xi_{\vec{h}}\big\|_p = \sup_{\vec{h}\in \mathrm{H}_{l,r}} \sup_{\vartheta\in\mathbb{B}_{\mathrm{q},d}} \int_{(-b,b)^d} \xi_{\vec{h}}(x)\vartheta(x)\nu_d(\mathrm{d}x).$$

Noting that $\int_{(-b,b)^d} \xi_{\vec{h}}(x)\vartheta(x)\nu_d(\mathrm{d}x) = \int \big[\int_{(-b,b)^d} h^{-d}(x)K\big(\frac{t-x}{h(x)}\big)\vartheta(x)\nu_d(\mathrm{d}x)\big]W(\mathrm{d}t)$ we obtain

$$\sup_{\vec{h}\in \mathrm{H}_{l,r}} \big\|\xi_{\vec{h}}\big\|_p = \sup_{Q\in\mathfrak{Q}_{l,r}} \int Q(t)W(\mathrm{d}t) =: \sup_{Q\in\mathfrak{Q}_{l,r}} \zeta(Q).$$

Remind that $\mathrm{H}_{l,r}$ and $\mathfrak{Q}_{l,r}$ are defined in Lemma 4. Using standard slicing device we reduce the initial problem to the investigation of $\sup_{Q\in\mathfrak{Q}_{l,r}}\zeta(Q)$, see (3.52). Obviously $\zeta(\cdot)$ is centered gaussian random function on $\mathfrak{Q}_{l,r}$ and our goal is to apply to it the assertion I of Lemma 1. To do this it suffices to show that

$$\mathbb{E}\Big\{\sup_{Q\in\mathfrak{Q}_{l,r}} \zeta(Q)\Big\} \leq U_{l,r} \tag{3.49}$$

for some $0 < U_{l,r} < \infty$ and to compute

$$\sigma^2_{l,r} := \sup_{Q\in\mathfrak{Q}_{l,r}} \int Q^2(t)\nu_d(\mathrm{d}t). \tag{3.50}$$

We will see that this programm, being similar to those realized in the proof of Theorem 2, requires completely different arguments. It is related to the fact that we consider the random field $\xi_{\vec{h}}$ itself and not its "normalized" version $\sqrt{V_{\vec{h}}}\xi_{\vec{h}}$.

First step consists in the finding an appropriated upper bound for $\sigma_{l,r}$. In distinction from the similar problem related to the quantity σ_Υ appeared in the proof of Theorem 2 the computations here are more involved. The proof of the bound obtained in (3.53) heavily exploits the condition $p\in[1,2]$ and one can easily checked that (3.53) is not true in general if $p>2$.

Second step consists in proving (3.49). As in the proof of Theorem 2 the main problem here is to bound from above corresponding Dudley'y integral and Lemma 4 is the basic technical tool for it. The aforementioned bound is presented in (3.54).

There is however a great difference between Lemmas 3 and 4. One of the main efforts made in the proof of Theorem 2 is to reduce the considered problem to the study of supremum of gaussian random function defined on $\mathcal{Q}_{\mathbf{s},\mathrm{x}}$. The latter set consists of smooth *univariate* functions and this fact is crucial for the proof of Lemma 3. Namely to make the aforementioned reduction possible the original problem "is replaced" by the study of the process $\sqrt{V_{\vec{h}}}\xi_{\vec{h}}$ and functional classes $\mathbb{H}_d(\tau,\mathcal{L},\mathcal{A})$ are introduced. All of this is dictated by the consideration of anisotropic classes of bandwidths. It turns out that it is not necessary when isotropic classes are studied. Although $\mathfrak{Q}_{l,r}$ is the class of *d-variate* functions, its entropy admits very precise bound presented in Lemma 4, that in its turn leads to the correct estimate in (3.49).

3.3.4. Proof of Theorem 3.

For any $r>d, r\in\mathbb{N}^*$, set $\psi^*_r(h) = C^*_2(r)\big\|h^{-\frac{d}{2}}\big\|_{p+\frac{1}{r}}$. We have

$$\mathbb{E}\Big\{\sup_{\vec{h}\in\mathrm{H}}\Big[\big\|\xi_{\vec{h}}\big\|_p - \inf_{r\in\mathbb{N}^*, r>d}\psi^*_r(\vec{h})\Big]_+\Big\}^q \leq \sum_{r=d+1}^{\infty}\mathbb{E}\Big\{\sup_{\vec{h}\in\mathrm{H}}\Big[\big\|\xi_{\vec{h}}\big\|_p - \psi^*_r(\vec{h})\Big]_+\Big\}^q.$$

Moreover, since $\mathrm{H}=\cup_{l\geq 1}\mathrm{H}_{l,r}$ for any $r\in\mathbb{N}^*$, one has

$$\left\{\sup_{\vec{h}\in\mathrm{H}}\left[\big\|\xi_{\vec{h}}\big\|_p-\psi_r^*(\vec{h})\right]_+\right\}^q\leq\sum_{l=1}^{\infty}\bigg(\sup_{\vec{h}\in\mathrm{H}_{l,r}}\big\|\xi_{\vec{h}}\big\|_p-C_2^*(r)2^{l-1}\mathfrak{h}^{-\frac{d}{2}}\bigg)_+^q.$$

Thus,

$$\mathbb{E}\left\{\sup_{\vec{h}\in\mathrm{H}}\left[\big\|\xi_{\vec{h}}\big\|_p-\inf_{r\in\mathbb{N}^*,r>d}\psi_r^*(\vec{h})\right]_+\right\}^q\leq\sum_{r=d+1}^{\infty}\sum_{l=1}^{\infty}\mathbb{E}\bigg(\sup_{\vec{h}\in\mathrm{H}_{l,r}}\big\|\xi_{\vec{h}}\big\|_p-C_2^*(r)2^{l-1}\mathfrak{h}^{-\frac{d}{2}}\bigg)_+^q. \qquad (3.51)$$

Thus, we get from (3.51)

$$\mathbb{E}\left\{\sup_{\vec{h}\in\mathrm{H}}\left[\big\|\xi_{\vec{h}}\big\|_p-\inf_{r\in\mathbb{N}^*,r>d}\psi_r^*(\vec{h})\right]_+\right\}^q\leq\sum_{r=d+1}^{\infty}\sum_{l=1}^{\infty}\mathbb{E}\bigg(\sup_{Q\in\mathfrak{Q}_{l,r}}\zeta(Q)-C_2^*(r)2^{l-1}\mathfrak{h}^{-\frac{d}{2}}\bigg)_+^q. \qquad (3.52)$$

1^0. We start with bounding the quantity $\sigma_{l,r}$ given in (3.50). Putting for any $x,y\in(-b,b)^d$

$$R(x,y)=\int K\bigg(\frac{t-x}{h(x)}\bigg)K\bigg(\frac{t-y}{h(y)}\bigg)\nu_d(\mathrm{d}t),$$

we obtain for any $Q\in\mathfrak{Q}_{l,r}$

$$\begin{aligned}\int Q^2(t)\nu_d(\mathrm{d}t) &= \int\bigg[\int_{(-b,b)^d}h^{-d}(x)K\bigg(\frac{t-x}{h(x)}\bigg)\vartheta(x)\nu_d(\mathrm{d}x)\bigg]^2\nu_d(\mathrm{d}t)\\ &= \int_{(-b,b)^d}\int_{(-b,b)^d}h^{-d}(x)h^{-d}(y)\vartheta(x)\vartheta(y)R(x,y)\nu_d(\mathrm{d}x)\nu_d(\mathrm{d}y).\end{aligned}$$

Taking into account that $\mathrm{supp}(K)\subseteq[-a,a]^d$ in view of Assumption 3 we get

$$|R(x,y)|\leq\big[h(x)\wedge h(y)\big]\|K\|_\infty\|K\|_1 1_{[-2a,2a]^d}\bigg(\frac{x-y}{h(x)\vee h(y)}\bigg).$$

Hence, putting $\Upsilon=\|K\|_\infty\|K\|_1$, we obtain

$$\begin{aligned}&\int Q^2(t)\nu_d(\mathrm{d}t)\\ &\leq\Upsilon\int_{(-b,b)^d}\int_{(-b,b)^d}\big|\vartheta(x)\vartheta(y)\big|\big[h(x)\vee h(y)\big]^{-d}1_{[-2a,2a]^d}\bigg(\frac{x-y}{h(x)\vee h(y)}\bigg)\nu_d(\mathrm{d}x)\nu_d(\mathrm{d}y).\end{aligned}$$

It remains to note

$$\begin{aligned}&\big[h(x)\vee h(y)\big]^{-d}1_{[-2a,2a]^d}\bigg(\frac{x-y}{h(x)\vee h(y)}\bigg)\\ &\leq h^{-d}(x)1_{[-2a,2a]^d}\bigg(\frac{x-y}{h(x)}\bigg)+h^{-d}(y)1_{[-2a,2a]^d}\bigg(\frac{x-y}{h(y)}\bigg)\end{aligned}$$

and, therefore,

$$\begin{aligned}\int Q^2(t)\nu_d(\mathrm{d}t) &\leq 2\Upsilon\int_{(-b,b)^d}|\vartheta(v)|\bigg[\int_{(-b,b)^d}h^{-d}(v)1_{[-2a,2a]^d}\bigg(\frac{u-v}{h(v)}\bigg)|\vartheta(u)|\nu_d(\mathrm{d}u)\bigg]\nu_d(\mathrm{d}v)\\ &\leq 2^{d+1}a^d\Upsilon\int|\vartheta^*(v)|\sup_{\lambda>0}(2\lambda)^{-d}\bigg[\int_{\mathbb{R}^d}1_{[-\lambda,\lambda]^d}\bigg(\frac{u-v}{\lambda}\bigg)|\vartheta^*(u)|\nu_d(\mathrm{d}u)\bigg]\nu_d(\mathrm{d}v)\\ &\leq 2^{d+1}a^d\Upsilon\int|\vartheta^*(v)|M[|\vartheta^*|](v)\nu_d(\mathrm{d}v).\end{aligned}$$

Here we have put $\vartheta^*(\cdot) = 1_{(-b,b)^d}(\cdot)\vartheta(\cdot)$ and $M[|\vartheta^*|]$ denotes the Hardy-Littlewood maximal operator applied to the function $|\vartheta^*|$.

In view of $(2,2)$-strong maximal inequality, Folland (1999), there exists $\mathbf{c}(d)$ such that

$$\int_{\mathbb{R}^d} \big\{M[|\vartheta^*|](v)\big\}^2 \nu_d(\mathrm{d}v) \leq \mathbf{c}^2(d) \int_{\mathbb{R}^d} |\vartheta^*(v)|^2 \nu_d(\mathrm{d}v).$$

Using the latter bound we obtain applying Cauchy-Schwartz inequality

$$\left[\int Q^2(t)\nu_d(\mathrm{d}t)\right]^{\frac{1}{2}} \leq \sqrt{\mathbf{c}(d)}\left[\int_{(-b,b)^d} |\vartheta(v)|^2 \nu_d(\mathrm{d}v)\right]^{\frac{1}{2}} \leq \sqrt{2^{d+1}a^d \Upsilon \mathbf{c}(d)(2b)^{\frac{d(p-1)}{p}}}.$$

To get the last inequality we applied the Hölder inequality and took into account that $\vartheta \in \mathbb{B}_{\mathrm{q},d}$ and $\mathrm{q} \geq 2$ since $p \leq 2$.

Noting that the right hand side of the obtained inequality is independent of Q we get

$$\sigma_{l,r} \leq \sqrt{2^{d+1}a^d \|K\|_\infty \|K\|_1 \mathbf{c}(d)(2b)^{\frac{d(p-1)}{p}}} := \sigma_*. \tag{3.53}$$

We would like to emphasize that the condition $p \leq 2$ is crucial in order to obtain the bound presented in (3.53).

2^0. Let us now establish (3.50). The intrinsic semi-metric ρ_ζ of $\zeta(\cdot)$ is given by

$$\rho_\zeta(Q_1, Q_2) = \|Q_1 - Q_2\|_2, \quad Q_1, Q_2 \in \mathfrak{Q}_{l,r}.$$

Taking into account that $\frac{d}{2\gamma_r} = \frac{2pr}{2pr+1} < 1$ and applying the second assertion of Lemma 1 and Lemma 4 we obtain in view of (3.53)

$$\begin{aligned}
D_{\mathfrak{Q}_{l,r},\rho_\zeta} &= 4\sqrt{2\lambda^*_d(r)}\big[T(r)\big]^{d/2\gamma_r}\big(2^l \mathfrak{h}^{-\frac{d}{2}}\big)\int_0^{\sigma_{l,r}/2} \delta^{-d/2\gamma_r}\mathrm{d}\delta \\
&= 4\sqrt{2\lambda^*_d(r)}\big[T(r)\big]^{d/2\gamma_r}\big(2^l \mathfrak{h}^{-\frac{d}{2}}\big)(\sigma_{l,r}/2)^{\frac{1}{2pr}} \\
&\leq 4\sqrt{2\lambda^*_d(r)}\big[T(r)\big]^{d/2\gamma_r}\big(\sigma_*/2\big)^{\frac{1}{2pr}}\big(2^l \mathfrak{h}^{-\frac{d}{2}}\big).
\end{aligned}$$

We conclude that Dudley integral is finite and as it is proved in Lemma 4 $\mathfrak{Q}_{l,r}$ is a totally bounded space with respect to the intrinsic semi-metric of $\zeta(\cdot)$. It implies that $\zeta(\cdot)$ is almost surely continuous on $\mathfrak{Q}_{l,r}$ and, therefore, $\sup_{Q\in\mathfrak{Q}_{l,r}} \zeta(Q)$ is a random variable.

Thus, in view of the second assertion of Lemma 1

$$\mathbb{E}\Big\{\sup_{Q\in\mathfrak{Q}_{l,r}} \zeta(Q)\Big\} \leq 4\sqrt{2\lambda^*_d(r)}\big[T(r)\big]^{d/2\gamma_r}\big(\sigma_*/2\big)^{\frac{1}{2pr}}\big(2^l \mathfrak{h}^{-\frac{d}{2}}\big) \tag{3.54}$$

and (3.50) is proved with $U_{l,r} = 4\sqrt{2\lambda^*_d(r)}\big[T(r)\big]^{d/2\gamma_r}\big(\sigma_*/2\big)^{\frac{1}{2pr}}\big(2^l \mathfrak{h}^{-\frac{d}{2}}\big)$.

Moreover, $\zeta(\cdot)$ is almost surely bounded on $\mathfrak{Q}_{l,r}$ and, therefore, the first assertion of Lemma 1 is applicable.

3^0. Hence, noting that $C^*_2(r) = 8\sqrt{2\lambda^*_d(r)}\big[T(r)\big]^{d/2\gamma_r}\big(\sigma_*/2\big)^{\frac{1}{2pr}} + 4\sqrt{qe^r}\sigma_*$ we obtain

$$\mathbb{P}\Big\{\sup_{Q\in\mathfrak{Q}_{l,r}} \zeta(Q) \geq 2^{l-1}\mathfrak{h}^{-\frac{d}{2}}C^*_2(r) + z\Big\} \leq \exp\big\{-2^{l+1}q\mathfrak{h}^{-\frac{d}{2}}e^r\big\}e^{-\frac{z^2}{2\sigma^2_*}}, \quad \forall z > 0.$$

It yields for any $q \geq 1$

$$\mathbb{E}\Big(\sup_{Q\in\mathfrak{Q}_{l,r}} \zeta(Q) - C_2^*(r)2^{l-1}\mathfrak{h}^{-\frac{d}{2}}\Big)_+^q \;=\; q\int_0^\infty z^{q-1}\mathbb{P}\Big\{\sup_{Q\in\mathfrak{Q}_{l,r}} \zeta(Q) \geq 2^{l-1}\mathfrak{h}^{-\frac{d}{2}}C_2^*(r)+z\Big\}$$
$$\leq\; \sqrt{8\pi}\sigma_*^{q-1}\gamma_{q+1}\exp\big\{-2^{l+1}q\mathfrak{h}^{-\frac{d}{2}}e^r\big\}. \tag{3.55}$$

We deduce from (3.52) and (3.55)

$$\mathbb{E}\Big\{\sup_{\vec{h}\in\mathrm{H}}\Big[\big\|\xi_{\vec{h}}\big\|_p - \inf_{r\in\mathbb{N}^*, r>d}\psi_r^*(\vec{h})\Big]\Big\}_+^q \leq \Big(C_5 e^{\mathfrak{h}^{-\frac{d}{2}}}\Big)^q,$$

where, recall, $C_5 = \Big[\sqrt{8\pi}\sigma_*^{q-1}\gamma_{q+1}\Big]^{\frac{1}{q}}\sum_{r=d+1}^\infty\sum_{l=1}^\infty e^{-2^l e^r}$.

∎

4. Appendix

4.1. Proof of Lemma 3

Recall that $\mu^{-1} = \mathfrak{q}^{-1} + \tau r^{-1}$ and note that $2 > \mu > 1$ since $\tau < 1$ and $r > 2$. The proof of the lemma is mostly based on the inclusion

$$\mathcal{Q}_{\mathrm{x},\mathbf{s}} \;\in\; \mathbb{S}_\mu^\omega\Big([-a-b, a+b], \tilde{R}_\mu\Big), \qquad \forall \omega \in \big(1/\mu - 1/2, 1\big), \tag{4.1}$$

where $\tilde{R}_\mu = \|\mathcal{K}\|_1 + 2\left[5\big\{4L(a+1)\big\}^\mu + 4\big\{2\|\mathcal{K}\|_1\big\}^\mu(2-\mu)^{-1}\right]^{\frac{1}{\mu}}$.

First, we note that all functions from $\mathcal{Q}_{\mathrm{x},\mathbf{s}}$ vanish outside the interval $\Delta = [-a-b, a+b]$ since $\mathcal{K}$ is compactly supported on $[-a, a]$ and $\mathfrak{h}_{s_1} \leq \mathfrak{h} < 1$.

Next, applying Young inequality we obtain for any $Q \in \mathcal{Q}_{\mathrm{x},\mathbf{s}}$

$$\big\|Q\big\|_{\mathbb{L}_\mu(\Delta)} \;=\; \lambda_{\vec{h},\mathbf{s}}^{-1}(\mathrm{x})\left[\int_\Delta\left|\int_{-b}^b \mathfrak{h}_{s_1}^{-1/2}\mathcal{K}\Big(\frac{y-x_1}{\mathfrak{h}_{s_1}}\Big)\ell(x_1)1_{\Lambda_{\mathbf{s}}[\vec{h}]}(x_1,\mathrm{x})\nu_1(\mathrm{d}x_1)\right|^\mu \nu_1(\mathrm{d}y)\right]^{\frac{1}{\mu}}$$
$$\leq\; \lambda_{\vec{h},\mathbf{s}}^{-1}(\mathrm{x})(\mathfrak{h}_{s_1})^{\frac{1}{2}}\|\mathcal{K}\|_1\left[\int_{-b}^b |\ell(x_1)|^\mu 1_{\Lambda_{\mathbf{s}}[\vec{h}]}(x_1,\mathrm{x})\nu_1(\mathrm{d}x_1)\right]^{\frac{1}{\mu}} \leq (\mathfrak{h}_{s_1})^{\frac{1}{2}}\|\mathcal{K}\|_1. \tag{4.2}$$

To get the last inequality we have used (3.40).

Let $\omega \in \big(1/\mu - 1/2, 1\big)$ be fixed. Let us bound from above the quantity

$$J_\mu := \int_\Delta\int_\Delta \frac{|Q(y)-Q(z)|^\mu}{|y-z|^{1+\mu\omega}}\mathrm{d}y\mathrm{d}z.$$

Putting $y = u + v$ and $z = u - v$ we obtain by changing of variables

$$J_\mu \leq 2^{-\mu\omega}\int_{-\infty}^\infty |v|^{-1-\mu\omega}\left[\int_{-\infty}^\infty |Q_s(u+v) - Q_s(u-v)|^\mu \mathrm{d}u\right]\mathrm{d}v$$

Note also that

$$
\begin{aligned}
&|Q_s(u+v)-Q_s(u-v)|\\
&\leq \lambda^{-1}_{\vec{h},\mathbf{s}}(\mathrm{x})\int_{-b}^{b}\mathfrak{h}_{s_1}^{-1/2}\Big|\mathcal{K}\Big(\frac{u-x_1}{\mathfrak{h}_{s_1}}+\frac{v}{\mathfrak{h}_{s_1}}\Big)-\mathcal{K}\Big(\frac{u-x_1}{\mathfrak{h}_{s_1}}-\frac{v}{\mathfrak{h}_{s_1}}\Big)\Big||\ell(x_1)|1_{\Lambda_{\mathbf{s}}[\vec{h}]}(x_1,\mathrm{x})\nu_1(\mathrm{d}x_1).
\end{aligned}
$$

Hence,

$$
J_\mu\leq 2^{-\mu\omega}\mathfrak{h}_s^{-\mu(\omega+1/2)}\lambda^{-\mu}_{\vec{h},\mathbf{s}}(\mathrm{x})\int_{-\infty}^{\infty}|w|^{-1-\mu\omega}G^{\mu}(w)\mathrm{d}w,
$$

where we have put for any $w\in\mathbb{R}$

$$
G(w)=\bigg[\int_{-\infty}^{\infty}\bigg[\int_{-b}^{b}\Big|\mathcal{K}\Big(\frac{u-x_1}{\mathfrak{h}_{s_1}}+w\Big)-\mathcal{K}\Big(\frac{u-x_1}{\mathfrak{h}_{s_1}}-w\Big)\Big||\ell(x_1)|1_{\Lambda_{\mathbf{s}}[\vec{h}]}(x_1,\mathrm{x})\nu_1(\mathrm{d}x_1)\bigg]^{\mu}\mathrm{d}u\bigg]^{\frac{1}{\mu}}.
$$

Applying Young inequality for any fixed w and we obtain

$$
\begin{aligned}
G(w)&\leq \mathfrak{h}_{s_1}\bigg[\int_{-\infty}^{\infty}\big|\mathcal{K}(u+w)-\mathcal{K}(u-w)\big|\,\mathrm{d}u\bigg]\bigg[\int_{-b}^{b}|\ell(x_1)|^{\mu}1_{\Lambda_{\mathbf{s}}[\vec{h}]}(x_1,\mathrm{x})\nu_1(\mathrm{d}x_1)\bigg]^{\frac{1}{\mu}}\\
&\leq \mathfrak{h}_{s_1}\bigg[\int_{-\infty}^{\infty}\big|\mathcal{K}(u+w)-\mathcal{K}(u-w)\big|\,\mathrm{d}u\bigg]\lambda_{\vec{h},\mathbf{s}}(\mathrm{x}).
\end{aligned}
$$

To get the last inequality we have used (3.40). Note that

$$
\begin{aligned}
\int_{-\infty}^{\infty}\big|\mathcal{K}(u+w)-\mathcal{K}(u-w)\big|\,\mathrm{d}u&\leq 2\|\mathcal{K}\|_1,\quad\forall w\in\mathbb{R};\\
\int_{-\infty}^{\infty}\big|\mathcal{K}(u+w)-\mathcal{K}(u-w)\big|\,\mathrm{d}u&\leq 4L(a+1)|w|,\quad\forall w\in[-1,1].
\end{aligned}
$$

To get the second inequality we have used Assumption 2 (**i**). Thus, we get finally

$$
J_\mu\leq 2^{-\mu\omega}\mathfrak{h}_{s_1}^{\mu(1/2-\omega)}\left[5\big\{4L(a+1)\big\}^{\mu}+4\big\{2\|\mathcal{K}\|_1\big\}^{\mu}(2-\mu)^{-1}\right]. \tag{4.3}
$$

Here we have also used that $\mu<2$ and $\mu\omega>(2-\mu)(2\mu)^{-1}$.

Putting $\tilde{R}_\mu=\|\mathcal{K}\|_1+\left[5\big\{2L(a+2)\big\}^{\mu}+4\big\{2\|\mathcal{K}\|_1\big\}^{\mu}(2-\mu)^{-1}\right]^{\frac{1}{\mu}}$ we get from (4.2) and (4.3) for any $\omega\in\big(1/\mu-1/2,1\big)$

$$
\big\|Q\big\|_{\mathbb{L}_\mu(\Delta)}+\bigg[\int_\Delta\int_\Delta\frac{|Q(y)-Q(z)|^{\mu}}{|y-z|^{1+\mu\omega}}\mathrm{d}y\mathrm{d}z\bigg]^{1/\mu}\leq\tilde{R}_\mu\mathfrak{h}_{s_1}^{\frac{1}{2}-\omega}.
$$

Thus, the inclusion (4.1) is proved since $\tilde{R}_\mu\leq R_\mu$. The assertion of the lemma follows from Lemma 2 with $k=1$ and its consequence (3.1).

∎

4.2. Proof of Lemma 4

Similarly to the proof of Lemma 3 the proof of the present lemma is based on the inclusion

$$\mathfrak{Q}_{l,r} \subset \mathbb{S}_1^{\gamma_r}\Big(\big(-a-b,a+b\big)^d, R\Big), \quad R = T(r)\big(2^l\mathfrak{h}^{-\frac{d}{2}}\big)^{\frac{2\gamma_r}{d}}. \tag{4.4}$$

Indeed, if (4.4) holds then the required assertion follows from the consequence (3.1) of Lemma 2.

Thus, let us prove (4.4). First, we note that all functions from $\mathfrak{Q}_{l,r}$ vanish outside the cube $\mathbf{\Delta} = [-a-b,a+b]^d$ since K is compactly supported on $[-a,a]^d$ and $\mathfrak{h} < 1$.

Next, for any $Q \in \mathfrak{Q}_{l,r}$ we obviously have

$$\|Q\|_1 := \int_{\mathbf{\Delta}} |Q(t)|\nu_d(\mathrm{d}t) \le \|\mathcal{K}\|_1^d \int_{(-b,b)^d} |\vartheta(x)|\nu_d(\mathrm{d}x) \le \|\mathcal{K}\|_1^d(2b)^{1/p}, \tag{4.5}$$

where the last inequality follows from the condition $\vartheta \in \mathbb{B}_{q,d}$ and the Hölder inequality.

Taking into account that $\vec{h}(x) = \big(h(x),\dots,h(x)\big)$ and that $\lfloor\gamma_r\rfloor = \lfloor d/2\rfloor$, we have for any $\mathbf{n} \in \mathbb{N}^d$ satisfying $|\mathbf{n}| = \lfloor\gamma_r\rfloor$ in view of Assumption 3

$$D^{\mathbf{n}}Q(t) = \int_{(-b,b)^d} \big[h(x)\big]^{-|\mathbf{n}|-d}\big[D^{\mathbf{n}}K\big]\left(\frac{t-x}{\vec{h}(x)}\right)\vartheta(x)\nu_d(\mathrm{d}x).$$

Moreover, putting $y = u+v$ and $z = u-v$ we obtain by changing of variables

$$I_{\mathbf{n}} := \int_{\mathbf{\Delta}}\int_{\mathbf{\Delta}} \frac{\big|D^{\mathbf{n}}Q(y) - D^{\mathbf{n}}Q(z)\big|}{|y-z|^{d+\varepsilon}}\mathrm{d}y\mathrm{d}z \le 2^{-d-\alpha}\int_{\mathbb{R}^d} |v|^{-d-\varepsilon}T(v)\mathrm{d}v.$$

Here $\alpha = \gamma_r - \lfloor\gamma_r\rfloor$ and $T(v) = \int_{\mathbb{R}^d}\big|D^{\mathbf{n}}Q(u+v) - D^{\mathbf{n}}Q(u-v)\big|\mathrm{d}u$.

We get using Fubini theorem

$$\begin{aligned} I_{\mathbf{n}} \ &\le\ 2^{-d-\alpha}\int_{(-b,b)^d} \big[h(x)\big]^{-|\mathbf{n}|-d}|\vartheta(x)|\bigg\{\int |v|^{-d-\alpha}\bigg[\int \\ &\bigg|\big[D^{\mathbf{n}}K\big]\left(\frac{u+v-x}{h(x)}\right) - \big[D^{\mathbf{n}}K\big]\left(\frac{u-v-x}{h(x)}\right)\bigg|\mathrm{d}u\bigg]\mathrm{d}v\bigg\}\nu_d(\mathrm{d}x), \end{aligned}$$

By changing variables in inner integrals $w = (u-x)/h(x)$ and $\mathfrak{z} = v/h(x)$ we obtain

$$I_{\mathbf{n}} \ \le\ T\int_{(-b,b)^d} \big[h(x)\big]^{-|\mathbf{n}|-\varepsilon}|\vartheta(x)|\nu_d(\mathrm{d}x), \tag{4.6}$$

where $T = 2^{-d-\alpha}\int |\mathfrak{z}|^{-d-\alpha}\int \big|D^{\mathbf{n}}K(w+\mathfrak{z}) - D^{\mathbf{n}}K(w-\mathfrak{z})\big|\mathrm{d}w\mathrm{d}\mathfrak{z}$.

We obtain in view of Assumption 3 for any $|\mathbf{n}| \le \lfloor d/2\rfloor + 1$

$$\begin{aligned} \int |D^{\mathbf{n}}K(w+\mathfrak{z}) - D^{\mathbf{n}}K(w-\mathfrak{z})|\,\mathrm{d}w \ &\le\ 2C(K), \quad \forall \mathfrak{z} \in \mathbb{R}^d; \\ \int |D^{\mathbf{n}}K(w+\mathfrak{z}) - D^{\mathbf{n}}K(w-\mathfrak{z})|\,\mathrm{d}w \ &\le\ 2L(a+2)^d|\mathfrak{z}|, \quad \forall |\mathfrak{z}| \le 1. \end{aligned}$$

It yields (recall that $\mathfrak{D}$ denotes the unit disc in $\mathbb{R}^d$),

$$T \le 2^{-d+1}\bigg[L(a+2)^d\int \mathfrak{z}^{-d-\alpha+1}1_{\mathfrak{D}}(\mathfrak{z})\mathrm{d}z + C(K)\int \mathfrak{z}^{-d-\alpha}1_{\overline{\mathfrak{D}}}(\mathfrak{z})\mathrm{d}\mathfrak{z}\bigg] = T^*(r).$$

Thus, we deduce from (4.6) for any $\mathbf{n}$ satisfying $|\mathbf{n}| = \lfloor \gamma_r \rfloor$

$$\begin{aligned} I_{\mathbf{n}} &\leq T^*(r) \int_{(-b,b)^d} \big[h(x)\big]^{-\gamma_r} |\vartheta(x)| \nu_d(\mathrm{d}x) \leq T^*(r) \bigg(\int_{(-b,b)^d} \big[h(x)\big]^{-p\gamma} \nu_d(\mathrm{d}x) \bigg)^{\frac{1}{p}} \\ &= T^*(r) \bigg(\Big\| h^{-\frac{d}{2}} \Big\|_{\frac{2p\gamma_r}{d}} \bigg)^{\frac{2\gamma_r}{d}} = T^*(r) \bigg(\Big\| h^{-\frac{d}{2}} \Big\|_{p+\frac{1}{r}} \bigg)^{\frac{2\gamma_r}{d}}. \end{aligned} \tag{4.7}$$

Here we have used Hölder inequality, the condition $\vartheta \in \mathbb{B}_{\mathrm{q},d}$ and the definition of γ_r.

Taking into account that $\vec{h} \in \mathrm{H}_{l,r}$ we obtain from (4.7) that

$$\sum_{|\mathbf{n}|=\lfloor \gamma_r \rfloor} I_{\mathbf{n}} \leq (d/2+1)^d T^*(r) \big(2^l \mathfrak{h}^{-\frac{d}{2}} \big)^{\frac{2\gamma_r}{d}}.$$

It leads together with (4.5) to the assertion of the lemma. ∎

4.3. Proof of Proposition 1

Set

$$B_{\vec{h}}(f,x) = \bigg| \int K_{\vec{h}}(t-x) f(t) \mathrm{d}t - f(x) \bigg|, \quad x \in \mathbb{R}^d.$$

We start the proof with several remarks.

1) Obviously $\Lambda_{\mathbf{s}}\big[\vec{h}_f\big] \in \mathfrak{B}\big(\mathbb{R}^d\big)$ for any $f \in \mathbb{N}_d(\vec{\beta}, \vec{r}, \vec{L})$ and any multi-index $\mathbf{s}$ since $B_{\vec{h}}(f,\cdot)$ is measurable function. Moreover $\vec{h}_f(\cdot)$ takes its values in countable set that implies that $\vec{h}_f(\cdot)$ is measurable function.

2) The definition of the Nikolskii class implies that $\|f\|_{r_j} \leq L_j$ for any $j = 1, \dots d$. It yields, in view of the Young inequality

$$\big\| B_{\vec{h}}(f,\cdot) \big\|_{r_j} \leq (1 + \|K\|_1) L_j, \quad \forall j = 1, \dots d,$$

and therefore,

$$\nu_d\big(x \in (-b,b)^d : \ B_{\vec{h}}(f,x) = \infty \big) = 0, \quad \forall \vec{h} \in \mathfrak{H}_\varepsilon^d.$$

This, in its turn, implies that

$$\nu_d\Big(\cup_{j=1}^d \big\{ x \in (-b,b)^d : \ \ h_j(f,x) = \infty \big\} \Big) = 0. \tag{4.8}$$

3) The following statement was proved in Goldenshluger and Lepski (2014), Lemma 3: there exists a constant $\widetilde{C}$ completely determined by $\vec{\beta}, d$ and the function w such that

$$B_{\vec{h}}(f,x) \leq \sum_{j=1}^d B_{\vec{h},j}(f,x), \quad x \in \mathbb{R}^d, \qquad \big\| B_{\vec{h},j}(f,\cdot) \big\|_{r_j} \leq \tilde{C} L_j h_j^{\beta_j}, \quad \forall j = 1, \dots, d. \tag{4.9}$$

1^0. *Proof of the first assertion.* For any $\mathbf{s} \in \mathbb{N}^*$ recall that $\vec{\mathfrak{h}}_{\mathbf{s}} = (\mathfrak{h}_{s_1}, \dots, \mathfrak{h}_{s_d})$ and $V_{\mathbf{s}} = \prod_{j=1}^d \mathfrak{h}_{s_j}$. Denote by $\mathcal{S}_d$ the set consisting of $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{N}^d$ satisfying $s_j \geq S_\varepsilon(j)$ for any $j = 1, \dots, d$. We will also use the following notation: for any $\mathbf{s} \in \mathcal{S}_d$ let $\hat{\mathbf{s}} \in \mathbb{N}^d$ be such that $\hat{\mathbf{s}} < \mathbf{s}$ and $|\mathbf{s} - \hat{\mathbf{s}}| = 1$.

Putting $\mathcal{X}=\cap_{j=1}^d\{x\in(-b,b)^d:\ h_j(f,x)<\infty\}$ we have in view of the definition $\vec{h}(f,\cdot)$ for any $\mathbf{s}\in\mathcal{S}_d$ such that $\mathbf{s}\neq\big(S_\varepsilon(1),\ldots,S_\varepsilon(d)\big)$.

$$\begin{aligned}
\Lambda_{\mathbf{s}}\big[\vec{h}_f\big]\cap\mathcal{X} &\subseteq \Big\{x\in(-b,b)^d:\ B_{\vec{\mathfrak{h}}_{\mathbf{s}}}(f,x)+\varepsilon V_{\mathbf{s}}^{-\frac12}\le B_{\vec{\mathfrak{h}}_{\hat{\mathbf{s}}}}(f,x)+\varepsilon V_{\hat{\mathbf{s}}}^{-\frac12}\Big\}\\
&\subseteq \Big\{x\in(-b,b)^d:\ B_{\vec{\mathfrak{h}}_{\hat{\mathbf{s}}}}(f,x)\ge\varepsilon V_{\mathbf{s}}^{-\frac12}(1-e^{-1/2})\Big\}\\
&\subseteq \bigcup_{j=1}^d\Big\{x\in(-b,b)^d:\ B_{\vec{\mathfrak{h}}_{\hat{\mathbf{s}}},j}(f,x)\ge\varepsilon V_{\mathbf{s}}^{-\frac12}(1-e^{-1/2})d^{-1}\Big\}
\end{aligned}$$

The last inclusion follows from the first inequality in (4.9) and the definition of $\hat{\mathbf{s}}$.

We get from (4.8), the second inequality in (4.9) and the Markov inequality

$$\begin{aligned}
\nu_d\Big(\Lambda_{\mathbf{s}}\big[\vec{h}_f\big]\Big)=\nu_d\Big(\Lambda_{\mathbf{s}}\big[\vec{h}_f\big]\cap\mathcal{X}\Big) &\le \sum_{j=1}^d d^{r_j}V_{\mathbf{s}}^{\frac{r_j}{2}}\big[\varepsilon(1-e^{-1/2})\big]^{-r_j}\big\|B_{\vec{\mathfrak{h}}_{\hat{\mathbf{s}}},j}(f,\cdot)\big\|_{r_j}^{r_j}\\
&\le \sum_{j=1}^d\varkappa_j\big[\varepsilon^{-1}V_{\mathbf{s}}^{\frac12}\mathfrak{h}_{s_j}^{\beta_j}\big]^{r_j},
\end{aligned}$$

where we have put $\varkappa_j=\big\{d(e^{\beta_j}-e^{\beta_j-1/2})\widetilde{C}L_j\big\}^{r_j}$ and used once again the definition of $\hat{\mathbf{s}}$.

Since $\nu_d\Big(\Lambda_{\mathbf{s}}\big[\vec{h}_f\big]\Big)=0$ for any $\mathbf{s}\notin\mathcal{S}_d$ by the definition of $\vec{h}_f$ and $\nu_d\Big(\Lambda_{\mathbf{s_0}}\big[\vec{h}_f\big]\Big)\le(2b)^d$, $\mathbf{s_0}=\big(S_\varepsilon(1),\ldots,S_\varepsilon(d)\big)$, we obtain for any $\tau\in(0,1)$

$$\sum_{\mathbf{s}\in\mathbb{N}^d}\nu_d^\tau\Big(\Lambda_{\mathbf{s}}\big[\vec{h}_f\big]\Big)\le\sum_{j=1}^d\varkappa_j^\tau\sum_{\mathbf{s}\in\mathcal{S}_d,\mathbf{s}\neq\mathbf{s_0}}\big[\varepsilon^{-1}V_{\mathbf{s}}^{\frac12}\mathfrak{h}_{s_j}^{\beta_j}\big]^{\tau r_j}+(2b)^{\frac{d}{\tau}}.$$

In view of (2.8) (the definition of $S_\varepsilon(j), j=1,\ldots,d$) we get

$$\begin{aligned}
V_{\mathbf{s}}^{\frac12} &= \Big[\mathfrak{h}^d e^{-\sum_{l=1}^d S_\varepsilon(l)}e^{\sum_{l=1}^d(S_\varepsilon(l)-s_l)}\Big]^{\frac12}\le\varepsilon^{\frac{1}{2\beta+1}}e^{\frac12\sum_{l=1}^d(S_\varepsilon(l)-s_l)};\\
\mathfrak{h}_{s_j}^{\beta_j} &= \mathfrak{h}^{\beta_j}e^{-\beta_jS_\varepsilon(j)}e^{\beta_j(S_\varepsilon(j)-s_j)}\le\varepsilon^{\frac{2\beta}{2\beta+1}}e^{\beta_j(S_\varepsilon(j)-s_j)}\le\varepsilon^{\frac{2\beta}{2\beta+1}}.
\end{aligned}$$

It yields $\varepsilon^{-1}V_{\mathbf{s}}^{\frac12}\mathfrak{h}_{s_j}^{\beta_j}\le e^{\frac12\sum_{k=1}^d(S_\varepsilon(k)-s_k)}$ and, therefore,

$$\sum_{\mathbf{s}\in\mathbb{N}^d}\nu_d^\tau\Big(\Lambda_{\mathbf{s}}\big[\vec{h}_f\big]\Big)\le\sum_{j=1}^d\varkappa_j^\tau\Big(1-e^{-\frac{\tau r_j}{2}}\Big)^{-d}+(2b)^{\frac{d}{\tau}}=:\mathcal{L}.$$

The first assertion is proved.

2^0. *Proof of the second assertion.* The condition of the proposition allows us to assert that there exists $\mathfrak{p}>p$ such that $\upsilon(2+1/\beta)>\mathfrak{p}$. Putting $\phi_\varepsilon=e^{d/2}\varepsilon^{\frac{2\beta}{2\beta+1}}$ we obtain using the definition of $\vec{h}_f$

$$\begin{aligned}
\Big\|V_{\vec{h}_f}^{-\frac12}\Big\|_{\mathfrak{p}}^{\mathfrak{p}} &\le \varepsilon^{-\mathfrak{p}}\Big\|B_{\vec{h}_f}(f,\cdot)+\varepsilon V_{\vec{h}_f}^{-\frac12}\Big\|_{\mathfrak{p}}^{\mathfrak{p}}=\varepsilon^{-\mathfrak{p}}\int_{(-b,b)^d}\inf_{\vec{h}\in\mathfrak{H}_\varepsilon}\Big[B_{\vec{h}}(f,x)+\varepsilon V_{\vec{h}}^{-\frac12}\Big]^{\mathfrak{p}}\mathrm{d}x\\
&\le \big(2\phi_\varepsilon\varepsilon^{-1}\big)^{\mathfrak{p}}+\sum_{k=0}^\infty\big(2e^{k+1}\phi_\varepsilon\varepsilon^{-1}\big)^{\mathfrak{p}}\nu_d\Big(x:\ \inf_{\vec{h}\in\mathfrak{H}_\varepsilon}\Big[B_{\vec{h}}(f,x)+\varepsilon V_{\vec{h}}^{-\frac12}\Big]\ge2e^k\phi_\varepsilon\Big)\\
&\le \big(2\phi_\varepsilon\varepsilon^{-1}\big)^{\mathfrak{p}}+\sum_{k=0}^\infty\big(2e^{k+1}\phi_\varepsilon\varepsilon^{-1}\big)^{\mathfrak{p}}\nu_d\Big(x:\ B_{\vec{\mathfrak{h}}[k]}(f,x)+\varepsilon V_{\vec{\mathfrak{h}}[k]}^{-\frac12}\ge2e^k\phi_\varepsilon\Big),
\end{aligned}$$

where we choose $\vec{\mathfrak{h}}[k] \in \mathfrak{H}_\varepsilon$ as follows. Let $\vec{h}[k] = \big(h_1[k], \dots, h_d[k]\big)$ be given by

$$h_j[k] = \big(\phi_\varepsilon\big)^{1/\beta_j} e^{k\left(\frac{1}{\beta_j} - \frac{\upsilon(2+1/\beta)}{\beta_j r_j}\right)}, \quad j = 1, \dots, d,$$

and define $\vec{\mathfrak{h}}[k] \in \mathfrak{H}_\varepsilon$ from the relation $e^{-1}\vec{h}[k] \le \vec{\mathfrak{h}}[k] < \vec{h}[k]$.

First we note that

$$h_j[k] \le \big(\phi_\varepsilon\big)^{1/\beta_j} \le \mathfrak{h} e^{-S_\varepsilon(j)+1},$$

since $\vec{r} \in [1,p]^d$ and $p < \upsilon(2+1/\beta)$. This guarantees the existence of $\vec{\mathfrak{h}}[k]$. Next,

$$\varepsilon V_{\vec{\mathfrak{h}}[k]}^{-\frac{1}{2}} \le \varepsilon V_{e^{-1}\vec{h}[k]}^{-\frac{1}{2}} = e^{k+d/2}\varepsilon^{\frac{2\beta}{2\beta+1}} = e^k\phi_\varepsilon,$$

and, therefore, using the latter bound, (4.9) and Markov inequality we obtain

$$\begin{aligned}
\Big\|V_{\vec{h}_f}^{-\frac{1}{2}}\Big\|_{\mathfrak{p}}^{\mathfrak{p}} &\le \big(2\phi_\varepsilon\varepsilon^{-1}\big)^{\mathfrak{p}} + \sum_{k=0}^{\infty}\big(2e^{k+1}\phi_\varepsilon\varepsilon^{-1}\big)^{\mathfrak{p}}\nu_d\Big(x:\ B_{\vec{\mathfrak{h}}_{\mathbf{s}[k]}}(f,x) \ge e^k\phi_\varepsilon\Big) \\
&\le \big(2\phi_\varepsilon\varepsilon^{-1}\big)^{\mathfrak{p}} + \sum_{k=0}^{\infty}\big(2e^{k+1}\phi_\varepsilon\varepsilon^{-1}\big)^{\mathfrak{p}}\sum_{j=1}^{d}(e^k\phi_\varepsilon)^{-r_j}(\tilde{C}L_j)^{r_j}\big(\mathfrak{h}_{s_j}[k]\big)^{\beta_j r_j} \\
&\le \big(2\phi_\varepsilon\varepsilon^{-1}\big)^{\mathfrak{p}} + \sum_{k=0}^{\infty}\big(2e^{k+1}\phi_\varepsilon\varepsilon^{-1}\big)^{\mathfrak{p}} e^{-k\upsilon(2+1/\beta)}\sum_{j=1}^{d}\big(\tilde{C}L_j\big)^{r_j} \\
&= \varepsilon^{-\frac{\mathfrak{p}}{2\beta+1}}\Big\{\big(2e^{d/2}\big)^{\mathfrak{p}} + \big(2e^{d/2+1}\big)^{\mathfrak{p}}\sum_{k=0}^{\infty}e^{-k[\upsilon(2+1/\beta)-\mathfrak{p}]}\sum_{j=1}^{d}\big(\tilde{C}L_j\big)^{r_j}\Big\}.
\end{aligned}$$

As we see the assumption of the proposition $\upsilon(2+1/\beta) > p$ allowing us to choose $\mathfrak{p} > p$ and $\upsilon(2+1/\beta) > \mathfrak{p}$ is crucial. The second assertion is proved.

∎

4.4. Proofs of (1.5) and (1.6).

We start with the following bound obtained by application of the Minkovski inequality for integrals and the Hölder inequality.

$$\sigma_p\big(\vec{h}\big) \le \|K\|_2\big\|V_{\vec{h}}^{-\frac{1}{2}}\big\|_p, \quad \forall \vec{h} \in \mathfrak{S}^*_{d,p}(\mathfrak{h}). \tag{4.10}$$

Set $\mathfrak{S}^y = \Big\{\vec{h} \in \mathfrak{S}^*_{d,p}(\mathfrak{h}) :\quad \sigma_p\big(\vec{h}\big) \le y\|K\|_2\big\|V_{\vec{h}}^{-\frac{1}{2}}\big\|_p\Big\}$, where $y \le 2^{-2-\frac{1}{p}}$ will be chosen later. Our first goal consists in establishing the following inequality.

$$(3/4)(\boldsymbol{\gamma}_p)^{\frac{1}{p}}\|K\|_2\big\|V_{\vec{h}}^{-\frac{1}{2}}\big\|_p \le \mathbb{E}\big(\big\|\xi_{\vec{h}}\big\|_p\big) \le (\boldsymbol{\gamma}_p)^{\frac{1}{p}}\|K\|_2\big\|V_{\vec{h}}^{-\frac{1}{2}}\big\|_p, \quad \forall \vec{h} \in \mathfrak{S}^y. \tag{4.11}$$

where, remind, $\boldsymbol{\gamma}_p$ is the p-th absolute moment of the standard normal distribution

The right hand side of the latter inequality is obvious. Indeed, we have in view of Jensen inequality and Fubini theorem

$$\mathbb{E}\big(\|\xi_{\vec{h}}\|_p\big) \leq \Big[\mathbb{E}\big(\|\xi_{\vec{h}}\|_p^p\big)\Big]^{\frac{1}{p}} = \Big[\int_{(-b,b)^d} \big(\mathbb{E}|\xi_{\vec{h}}(x)|^p\big)\nu_d(\mathrm{d}x)\Big]^{\frac{1}{p}} = (\gamma_p)^{\frac{1}{p}}\|K\|_2\big\|V_{\vec{h}}^{-\frac{1}{2}}\big\|_p.$$

Thus, let us prove the left hand side in (4.11). In view of duality arguments

$$\zeta := \big\|\xi_{\vec{h}}\big\|_p = \sup_{\vartheta\in\mathbb{B}_{s,d}} \int_{(-b,b)^d} \vartheta(x)\xi_{\vec{h}}(x)\nu_d(\mathrm{d}x) =: \sup_{Q\in\mathfrak{Q}} \zeta_Q$$

where we have put $\zeta_Q = \int_{\mathbb{R}^d} Q(t)W(\mathrm{d}t)$ and

$$\mathfrak{Q} = \Big\{Q\in\mathbb{R}^d\to\mathbb{R}:\ \ Q(\cdot) = \int_{(-b,b)^d} \vartheta(x)K_{\vec{h}}(\cdot,x)\nu_d(\mathrm{d}x),\ \vartheta\in\mathbb{B}_{s,d}\Big\}.$$

Let M_ζ be the median of ζ and let $\eta\sim\mathcal{N}\big(0,\sigma_p^2(\vec{h})\big)$. We have in view of triangle inequality

$$(\gamma_p)^{\frac{1}{p}}\|K\|_2\big\|V_{\vec{h}}^{-\frac{1}{2}}\big\|_p = \Big[\mathbb{E}\big(\|\xi_{\vec{h}}\|_p^p\big)\Big]^{\frac{1}{p}} =: \big[\mathbb{E}|\zeta|^p\big]^{\frac{1}{p}} \leq M_\zeta + \big[\mathbb{E}|\zeta-M_\zeta|^p\big]^{\frac{1}{p}}.$$

Note that $\zeta=\sup_{Q\in\mathfrak{Q}}\zeta_Q$ and ζ_Q is zero mean gaussian random function on $\mathfrak{Q}$. Moreover, this function is bounded since $\mathbb{E}\zeta<\infty$ in view of the right hand side of (4.11).

Hence, in view of Theorem 12.2 in Lifshits (1995), $\mathbb{P}\big(|\zeta-M_\zeta|>z\big)\leq 2\mathbb{P}(|\eta|>z)$ for any $z>0$. It yields, $\mathbb{E}|\zeta-M_\zeta|^p\leq 2\mathbb{E}|\eta|^p = 2\gamma_p\sigma_p^p(\vec{h})$ Since $y\leq 2^{-2-\frac{1}{p}}$, we obtain for any $\vec{h}\in\mathfrak{S}^y$

$$\big[\mathbb{E}|\zeta-M_\zeta|^p\big]^{\frac{1}{p}} \leq 4^{-1}(\gamma_p)^{\frac{1}{p}}\|K\|_2\big\|V_{\vec{h}}^{-\frac{1}{2}}\big\|_p.$$

It remains to note that $M_\zeta\leq\mathbb{E}\zeta$, Theorem 14.1 in Lifshits (1995), and the left hand side of (4.11) follows. We easily deduce from (4.11) that

$$4^{-1}(\gamma_p)^{\frac{1}{p}}\|K\|_2\big\|V_{\vec{h}}^{-\frac{1}{2}}\big\|_p \leq M_\zeta \leq (\gamma_p)^{\frac{1}{p}}\|K\|_2\big\|V_{\vec{h}}^{-\frac{1}{2}}\big\|_p,\quad \forall\vec{h}\in\mathfrak{S}^y. \tag{4.12}$$

Indeed, the right hand side follows from $M_\zeta\leq\mathbb{E}\zeta$ and the right hand side of (4.11). Additionally,

$$\mathbb{E}\zeta \leq M_\zeta + \mathbb{E}|\zeta-M_\zeta| \leq M_\zeta + 2\gamma_1\sigma_p(\vec{h}) \leq M_\zeta + 2^{-1}(\gamma_p)^{\frac{1}{p}}\|K\|_2\big\|V_{\vec{h}}^{-\frac{1}{2}}\big\|_p,\quad \forall\vec{h}\in\mathfrak{S}^y.$$

This, together with left hand side of (4.11) completes the proof of (4.12).

Proof of (1.5). 1^0. Suppose first that $\vec{h}\in\mathfrak{S}^y$ and put for brevity $\lambda_p = \big\|V_{\vec{h}}^{-\frac{1}{2}}\big\|_p$. We have

$$\begin{aligned}\mathbb{E}\Big\{\Big[\zeta-2^{-4}(\gamma_p)^{\frac{1}{p}}\|K\|_2\lambda_p\Big]_+\Big\}^q \ &\geq\ \big[2^{-4}(\gamma_p)^{\frac{1}{p}}\|K\|_2\lambda_p\big]^q\mathbb{P}\Big\{|\zeta-M_\zeta|\leq 2^{-3}(\gamma_p)^{\frac{1}{p}}\|K\|_2\lambda_p\Big\}\\ &\geq\ B_1'\mathfrak{h}^{-\frac{dq}{2}}\Big[1-2\mathbb{P}\Big\{|\eta|>2^{-3}(\gamma_p)^{\frac{1}{p}}\|K\|_2\lambda_p\Big\}\Big].\end{aligned} \tag{4.13}$$

To get the first inequality we have used the left hand side of (4.12). Taking into account that $\vec{h}\in\mathfrak{S}^y$ we obtain

$$\mathbb{P}\Big\{|\eta|>2^{-3}(\gamma_p)^{\frac{1}{p}}\|K\|_2\lambda_p\Big\}\Big] \leq 2-2\Phi\big(2^{-3}(\gamma_p)^{\frac{1}{p}}y^{-1}\big)$$

where Φ is the distribution function of the standard normal law. Choosing y_0 from the equality $2-2\Phi\big(2^{-3}(\boldsymbol{\gamma}_p)^{\frac{1}{p}}y^{-1}\big)=4^{-1}$ and setting $y=y_0\wedge 2^{-2-\frac{1}{p}}$ we deduce from (4.13)

$$\mathbb{E}\Big\{\Big[\zeta-2^{-4}(\boldsymbol{\gamma}_p)^{\frac{1}{p}}\|K\|_2\lambda_p\Big]_+\Big\}^q \;\;\geq\;\; 2^{-1}B_1'\mathfrak{h}^{-\frac{d}{2}},\quad \forall\vec{h}\in\mathfrak{S}^y. \tag{4.14}$$

2^0. Suppose now that $\vec{h}\in\mathfrak{S}^*_{d,p}(\mathfrak{h})\setminus\mathfrak{S}^y$ and put for brevity $X=2^{-3}(\boldsymbol{\gamma}_p)^{\frac{1}{p}}\|K\|_2$. One has

$$\mathbb{E}\Big\{\Big[\zeta-2^{-4}(\boldsymbol{\gamma}_p)^{\frac{1}{p}}\|K\|_2\lambda_p\Big]_+\Big\}^q \;\;\geq\;\; B_1'\mathfrak{h}^{-\frac{dq}{2}}\mathbb{P}\{\zeta\geq X\lambda_p\}.$$

Remembering that $\zeta=\sup_{Q\in\mathfrak{Q}}\zeta_Q$ we get

$$\mathbb{E}\Big\{\Big[\zeta-2^{-4}(\boldsymbol{\gamma}_p)^{\frac{1}{p}}\|K\|_2\lambda_p\Big]_+\Big\}^q \;\;\geq\;\; B_1'\mathfrak{h}^{-\frac{dq}{2}}\sup_{Q\in\mathfrak{Q}}\mathbb{P}\{\zeta_Q\geq X\lambda_p\}. \tag{4.15}$$

Taking into account that $\zeta_Q\sim\mathcal{N}\big(0,||Q||_2^2\big)$ we have

$$\sqrt{2\pi}\mathbb{P}\big\{\zeta_Q\geq X\lambda_p\big\}\geq ||Q||_2(X\lambda_p)^{-1}\big[1+||Q||_2^2(X\lambda_p)^{-2}\big]^{-1}e^{-\frac{(X\lambda_p)^2}{2||Q||_2^2}}$$

Since $\sigma_p(\vec{h})=\sup_{Q\in\mathfrak{Q}}||Q||_2$ we obtain from (4.10) $\big[1+||Q||_2^2(X\lambda_p)^{-2}\big]^{-1}\geq\big[1+8(\boldsymbol{\gamma}_p)^{-\frac{1}{p}}\big]^{-1}$.
Therefore,

$$\sup_{Q\in\mathfrak{Q}}\mathbb{P}\big\{\zeta_Q\geq X\lambda_p\big\}\geq B_1''\sigma_p(\vec{h})(X\lambda_p)^{-1}e^{-\frac{(X\lambda_p)^2}{2\sigma_p^2(\vec{h})}}$$

Since $\vec{h}\in\mathfrak{S}^*_{d,p}(\mathfrak{h})\setminus\mathfrak{S}^y$ one has $\sigma_p(\vec{h})(X\lambda_p)^{-1}\geq 8y(\boldsymbol{\gamma}_p)^{-\frac{1}{p}}$ that implies

$$\sup_{Q\in\mathfrak{Q}}\mathbb{P}\big\{\zeta_Q\geq X\lambda_p\big\}\geq 8B_1''y(\boldsymbol{\gamma}_p)^{-\frac{1}{p}}e^{-\frac{(\boldsymbol{\gamma}_p)^{\frac{2}{p}}}{128y^2}}$$

It yields together with (4.15)

$$\mathbb{E}\Big\{\Big[\zeta-2^{-4}(\boldsymbol{\gamma}_p)^{\frac{1}{p}}\|K\|_2\lambda_p\Big]_+\Big\}^q \;\;\geq\;\; B_1'''\mathfrak{h}^{-\frac{dq}{2}},\quad \forall\vec{h}\in\mathfrak{S}^*_{d,p}(\mathfrak{h})\setminus\mathfrak{S}^y. \tag{4.16}$$

The inequality (1.5) follows now from (4.14) and (4.16).

Proof of (1.6). In view of the right hand side of (4.11) and the first assertion of Lemma 1 we have

$$\begin{aligned}&\mathbb{E}\Big\{\Big[\big\|\xi_{\vec{h}}\big\|_p-\big((\boldsymbol{\gamma}_p)^{\frac{1}{p}}\|K\|_2+\sqrt{2}\big)\big\|V_{\vec{h}}^{-\frac{1}{2}}\big\|_p\Big]_+\Big\}^q\\ &\leq q\int_0^\infty z^{q-1}\mathbb{P}\big\{\zeta-\mathbb{E}\zeta>\sqrt{2}\lambda_p+z\big\}\mathrm{d}z\;\leq\; e^{-\sigma_p^{-2}(\vec{h})\lambda_p}\sigma_p^q\big(\vec{h}\big)q\int_0^\infty z^{q-1}e^{-\frac{z^2}{2}}\mathrm{d}z.\end{aligned}$$

∎

References

Alexander, K. S. (1984). Probability inequalities for empirical processes and a law of the iterated logarithm. *Ann. Probab.* **12**, 1041–1067.

Barron, A., Birgé, L. and Massart, P. (1999). Risk bounds for model selection via penalization. *Probab. Theory Related Fields* **113**, 301–413.

Birman, M.S, and Solomjak M.Z. (1967). Piecewise polynomial approximation of functions of the classes W_p^{ε}. *Math. Sbornik.* **73**, 331–355.

Bobkov, S.G. (1988). Upper functions and oscillating Gaussian processes. *J. Soviet Math.* **43, 6**, 2745–2751.

Bousquet, O. (2002). A Bennett concentration inequality and its application to suprema of empirical processes. *C. R. Math. Acad. Sci. Paris* **334**, 495–500.

Cavalier, L. and Golubev, Yu. (2006). Risk hull method and regularization by projections of ill-posed inverse problems. *Ann. Statist.* **34**, 1653–1677.

Folland, G. B. (1999). *Real Analysis. Modern Techniques and Their Applications.* Second edition. John Wiley & Sons, Inc., New York.

Egishyants, S.A. and Ostrovskii, E. I. (1996). Local and global upper function for random fields. *Theory Probab. App.* **41, 4**, 657–665.

Giné, E. and Koltchinskii, V. (2006). Concentration inequalities and asymptotic results for ratio type empirical processes. *Ann. Probab.* **34**, 1143-1216.

Giné, E. and Zinn, J. (1984). Some limit theorems for empirical processes. *Ann. Probab.* **12**, 928–989.

Goldenshluger, A. and Lepski, O. (2009). Structural adaptation via L_p–norm oracle inequalities. *Probability Theory and Related Fields* **143**, 41-71.

Goldenshluger, A. and Lepski, O. (2011a). Uniform bounds for norms of sums of independent random functions. *Ann. Probab.* **39, 6**, 2318–2384.

Goldenshluger, A. and Lepski, O. (2011b). Bandwidth selection in kerrnel density estimation: oracle inequalities and adaptive minimax optimality. *Ann. Statist.* **39**, 1608-1632.

Goldenshluger, A. and Lepski, O. (2014). On adaptive minimax density estimation on $\mathbb{R}^d$. *Probability Theory and Related Fields* **159**, 479–543

Golubev, Y. and Spokoiny, V. (2009). Exponential bounds for minimum contrast estimators. *Electron. J. Stat.* **3**, 712–746.

Kalinauskaĭte, N. (1966). Upper and lower functions for sums of independent random variables with limiting stable distribution. *Litovsk. Mat.Sb.* **6**, 249–256.

Kerkyacharian, G., Lepski, O. and Picard, D. (2008). Nonlinear estimation in anisotropic multiindex denoising. Sparse case. *Theory Probab. Appl.* **52**, 58–77.

Kluchnikoff, N. (2005). On adaptive estimation of anysotropic functions. *Ph.d Thesis. Aix-Marseille 1.*

Lepskii, O. V. (1991). Asymptotically minimax adaptive estimation. I. Upper bounds. Optimally adaptive estimates. *Theory Probab. Appl.* **36**, 682–697.

Lepski, O. (2013a) Upper Functions for Positive Random Functionals. I. General Setting and Gaussian Random Functions. *Math. Methods of Statist., 22, no. 1 (2013), 1-27.*

Lepski, O. (2013b) Upper Functions for Positive Random Functionals. II. Application to the Empirical Processes Theory, Part 1. *Math. Methods of Statist., 22, 2, 83-99.*

Lepski, O. (2013c) Upper Functions for Positive Random Functionals. II. Application to the Empirical Processes Theory, Part 2. *Math. Methods of Statist., 22, no. 3 (2013), 193-212.*

Lepski, O. (2014) Adaptive estimation over anisotropic functional classes via oracle approach.

`arXiv:1405.4504v1, http://arxiv.org`

LIFSHITS, M.A. (1995). *Gaussian Random Functions.* Kluwer, Dordrecht.

MASSART, P. (2000). About the constants in Talagrand's concentration inequalities for empirical processes. *Ann. Probab.* **28**, 863–884.

NIKOL'SKII, S. M. (1977). *Priblizhenie Funktsii Mnogikh Peremennykh i Teoremy Vlozheniya.* (in Russian). [Approximation of functions of several variables and imbedding theorems.] Second edition, revised and supplemented. Nauka, Moscow.

QUALLS, C. and WATANABE, H. (1972). Asymptotic prpperties of Gaussian processes. *Annals Math. Stat.* **43**, 580–596.

SHIRAYEV, A.N., VALKEILA, E. and VOSTRIKOVA, L. (2002). On lower and upper function for square integrable martingales. *in Proc.Steklov Inst. Math.* **2 (237)**, 281–292.

TALAGRAND, M. (1994). Sharper bounds for Gaussian and empirical processes. *Ann. Probab.* **22**, 28–76.

TALAGRAND, M. (2005). *The Generic Chaining.* Springer, New York.

TSYBAKOV, A. (1998). Pointwise and sup-norm sharp adaptive estimation of functions on the Sobolev classes. *Ann.Statist.* **26**, 2420–2469.

VAN DE GEER, S. (2000). *Applications of Empirical Process Theory.* Cambridge University Press, Cambridge.

VAN DER VAART, A. W. and WELLNER, J. A. (1996). *Weak Convergence and Empirical Processes.* Springer, New York.