



HAL
open science

UNIFORM BOUNDS FOR NORMS OF SUMS OF INDEPENDENT RANDOM FUNCTIONS

Alexander Goldenshluger, Oleg Lepski

► **To cite this version:**

Alexander Goldenshluger, Oleg Lepski. UNIFORM BOUNDS FOR NORMS OF SUMS OF INDEPENDENT RANDOM FUNCTIONS. *Annals of Probability*, 2011, 39 (6), pp.2318-2384. 10.1214/10-AOP595 . hal-01265260

HAL Id: hal-01265260

<https://hal.science/hal-01265260>

Submitted on 2 Feb 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIFORM BOUNDS FOR NORMS OF SUMS OF INDEPENDENT RANDOM FUNCTIONS

BY ALEXANDER GOLDENSHLUGER¹ AND OLEG LEPSKI²

University of Haifa and Université Aix–Marseille I

In this paper, we develop a general machinery for finding explicit uniform probability and moment bounds on sub-additive positive functionals of random processes. Using the developed general technique, we derive uniform bounds on the \mathbb{L}_S -norms of empirical and regression-type processes. Usefulness of the obtained results is illustrated by application to the processes appearing in kernel density estimation and in nonparametric estimation of regression functions.

1. Introduction.

1.1. *General setting.* Let \mathfrak{G} and \mathfrak{H} be linear topological spaces, $(\Omega, \mathfrak{A}, \mathbb{P})$ be a complete probability space, and let $\xi_\theta : \Omega \rightarrow \mathfrak{G}$, $\theta \in \mathfrak{H}$ be a family of random mappings. In the sequel, $\xi_\bullet(\omega)$ is assumed linear and continuous on \mathfrak{H} for any $\omega \in \Omega$. Let $\Psi : \mathfrak{G} \rightarrow \mathbb{R}_+$ be a given *sub-additive* functional. Suppose that there exist functions $A : \mathfrak{H} \rightarrow \mathbb{R}_+$, $B : \mathfrak{H} \rightarrow \mathbb{R}_+$ and $U : \mathfrak{H} \rightarrow \mathbb{R}_+$ such that

$$(1.1) \quad \mathbb{P}\{\Psi(\xi_\theta) - U(\theta) \geq z\} \leq g\left(\frac{z^2}{A^2(\theta) + B(\theta)z}\right) \quad \forall \theta \in \mathfrak{H},$$

where $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is a monotone decreasing to zero function.

Let Θ be a fixed subset of \mathfrak{H} . In this paper, under rather general assumptions on U , A , B and Θ , we establish uniform probability and moment bounds of the following type: for any $\epsilon \in (0, 1)$, $y > 0$ and some $q \geq 1$

$$(1.2) \quad \mathbb{P}\left\{\sup_{\theta \in \Theta} [\Psi(\xi_\theta) - u_\epsilon(1 + \sqrt{y}\lambda_A + y\lambda_B)U(\theta)] \geq 0\right\} \leq P_{\epsilon, g}(y),$$

$$(1.3) \quad \mathbb{E} \sup_{\theta \in \Theta} [\Psi(\xi_\theta) - u_\epsilon(1 + \sqrt{y}\lambda_A + y\lambda_B)U(\theta)]_+^q \leq E_{\epsilon, g}(y).$$

Here λ_A and λ_B are the quantities completely determined by U , A , Θ and U , B , Θ , respectively, and the inequalities (1.2) and (1.3) hold if these quantities are finite;

Received April 2009; revised February 2010.

¹Supported by the ISF Grant 389/07.

²Supported by the Grant ANR-07-BLAN-0234.

MSC2010 subject classifications. Primary 60E15; secondary 62G07, 62G08.

Key words and phrases. Empirical processes, concentration inequalities, kernel density estimation, regression.

$P_{\epsilon,g}(\cdot)$ and $E_{\epsilon,g}(\cdot)$ are continuous decreasing to zero functions completely determined by ϵ and g ; and the factor u_ϵ is such that $u_\epsilon \rightarrow 1, \epsilon \rightarrow 0$. We present explicit expressions for all quantities appearing in (1.2) and (1.3).

In order to derive (1.2) and (1.3) from (1.1), we assume that the set Θ is the image of a totally bounded set in some metric space under a continuous mapping. Namely, if (\mathfrak{Z}, d) is a metric space, and \mathbb{Z} is a totally bounded subset of (\mathfrak{Z}, d) then we assume that there exists a *continuous* mapping ϕ from \mathfrak{Z} to \mathfrak{H} such that

$$(1.4) \quad \Theta = \{\theta \in \mathfrak{H} : \theta = \phi[\zeta], \zeta \in \mathbb{Z}\}.$$

Let $N_{\mathbb{Z},d}(\delta), \delta > 0$ be the minimal number of balls of radius δ in the metric d needed to cover \mathbb{Z} . The inequalities (1.2) and (1.3) are proved under some condition that relates $N_{\mathbb{Z},d}(\cdot)$ and $g(\cdot)$. It is worth mentioning that in particular examples the parametrization $\Theta = \phi[\mathbb{Z}]$ is often natural, while the metric d may have a rather unusual form.

Inequalities (1.2) and (1.3) can be considered as a refinement of usual bounds on the tail distribution of suprema of random functions. In particular, probability and moment bounds for $\sup_{\theta \in \Theta} \Psi(\xi_\theta)$ can be easily derived from (1.2) and (1.3). The well-known concentration results deal with deviation of the supremum of a random process from the expectation of this supremum, and estimation of the expectation is a separate rather difficult problem. In contrast, in this paper we develop *explicit* uniform bounds on the whole trajectory $\{\Psi(\xi_\theta), \theta \in \Theta\}$. The inequality in (1.1) provides the basic step in the development of such uniform probability bounds. The usual technique is based on the *chaining argument* that repeatedly applies inequality in (1.1) to increments of the considered random process [see, e.g., [Ledoux and Talagrand \(1991\)](#) and [van der Vaart and Wellner \(1996\)](#), Section 2.2].

The most interesting phenomena can be observed when a sequence of random mappings $\{\xi_\theta^{(n)}, \theta \in \mathfrak{H}\}, n \in \mathbb{N}^*$ is considered. There exists a class of problems where the quantities λ_A and λ_B depend on n , and $\lambda_A \rightarrow 0, \lambda_B \rightarrow 0$ as $n \rightarrow \infty$. Under these circumstances, one can choose $y = y_n \rightarrow \infty$ and $\epsilon = \epsilon_n \rightarrow 0$ such that

$$(1.5) \quad P_{\epsilon,g}(y_n) \rightarrow 0, \quad E_{\epsilon,g}(y_n) \rightarrow 0, \quad n \rightarrow \infty$$

and, at the same time,

$$(1.6) \quad u_{\epsilon_n}(1 + \sqrt{y_n}\lambda_A + y_n\lambda_B)U(\cdot) \rightarrow U(\cdot), \quad n \rightarrow \infty.$$

The relation in (1.5) means that $u_{\epsilon_n}(1 + \sqrt{y_n}\lambda_A + y_n\lambda_B)U(\cdot)$ is indeed a uniform upper bound for $\Psi(\xi_\theta^{(n)})$ on Θ , while (1.6) indicates that for large n this uniform bound is nearly as good as a nonuniform bound $U(\cdot)$ given in (1.1). Typically for a fixed $y > 0$, we have $P_{\epsilon,g}(y) \rightarrow \infty$ and $E_{\epsilon,g}(y) \rightarrow \infty$ as $\epsilon \rightarrow 0$; therefore, in order to get (1.5) and (1.6), $\epsilon_n \rightarrow 0$ and $y_n \rightarrow \infty$ should be calibrated in an appropriate way.

The general setting outlined above includes important specific problems that are in the focus of the present paper. We consider sequences of random mappings that are sums of real-valued random functions defined on some measurable space (here the parameter $n \in \mathbb{N}^*$ is the number of summands). We are interested in uniform bounds on the norms of such random functions; thus the sub-additive functional of interest Ψ is the \mathbb{L}_s -norm, $s \geq 1$. First, the nonuniform bound (1.1) is established, and then the inequalities of the type (1.2) and (1.3) are derived. It is shown that (1.5) and (1.6) hold under mild assumptions on the parametric set Θ . We also discuss sharpness of the nonuniform inequality in (1.1).

1.2. *Norms of sums of independent random functions.* Let $(\mathcal{T}, \mathfrak{X}, \tau)$ and $(\mathcal{X}, \mathfrak{X}, \nu)$ be σ -finite spaces, and let \mathcal{X} be a separable Banach space. Consider an \mathcal{X} -valued random element X defined on the complete probability space $(\Omega, \mathfrak{A}, \mathbb{P})$ and having the density f with respect to the measure ν . Let ε be a real random variable defined on the same probability space, independent of X and having a symmetric distribution.

For any $(\mathfrak{X} \times \mathfrak{X})$ -measurable function w on $\mathcal{T} \times \mathcal{X}$ and for any $t \in \mathcal{T}$, $n \in \mathbb{N}^*$, define the random functions

$$(1.7) \quad \xi_w(t) := \sum_{i=1}^n [w(t, X_i) - \mathbb{E}w(t, X)], \quad \eta_w(t) := \sum_{i=1}^n w(t, X_i)\varepsilon_i,$$

where (X_i, ε_i) , $i = 1, \dots, n$, are independent copies of (X, ε) . Put for $1 \leq s < \infty$

$$\|\xi_w\|_{s,\tau} = \left[\int |\xi_w(t)|^s \tau(dt) \right]^{1/s}, \quad \|\eta_w\|_{s,\tau} = \left[\int |\eta_w(t)|^s \tau(dt) \right]^{1/s}.$$

We are interested in uniform bounds of the type (1.2) and (1.3) for $\|\xi_w\|_{s,\tau}$ and $\|\eta_w\|_{s,\tau}$ when $w \in \mathcal{W}$, where \mathcal{W} is a given set of $(\mathfrak{X} \times \mathfrak{X})$ -measurable functions. This setup is a specific case of the general framework with $\Psi(\cdot) = \|\cdot\|_{s,\tau}$, $\theta = w$ and $\Theta = \mathcal{W}$. More precisely, if ψ_w denotes either ξ_w or η_w , and if \mathbb{P} is the probability law of X_1, \dots, X_n (when ξ_w is studied) or of $(X_1, \varepsilon_1), \dots, (X_n, \varepsilon_n)$ (when η_w is studied) then we want to find a functional $U(\psi_w) = U_\psi(w, f)$ such that (1.1) holds and

$$(1.8) \quad \mathbb{P} \left\{ \sup_{w \in \mathcal{W}} [\|\psi_w\|_{s,\tau} - u_\epsilon(1 + \sqrt{y}\lambda_A + y\lambda_B)U_\psi(w, f)] \geq 0 \right\} \leq P_{\epsilon,g}(y),$$

$$(1.9) \quad \mathbb{E} \sup_{w \in \mathcal{W}} [\|\psi_w\|_{s,\tau} - u_\epsilon(1 + \sqrt{y}\lambda_A + y\lambda_B)U_\psi(w, f)]_+^q \leq E_{\epsilon,g}(y),$$

$$q \geq 1.$$

Note that $\{\xi_w, w \in \mathcal{W}\}$ is the *empirical process*. In the sequel, we refer to $\{\eta_w, w \in \mathcal{W}\}$ as the *regression-type process* as it naturally appears in nonparametric estimation of regression functions. In the regression context, X_i are the design variables, ε_i are the random noise variables.

Uniform probability and moment bounds for empirical processes are a subject of vast literature; see, for example, Alexander (1984), Talagrand (1994), van der Vaart and Wellner (1996), Massart (2000), Bousquet (2002), Giné and Koltchinskii (2006) among many others. Such bounds play an important role in establishing the laws of iterated logarithm and central limit theorems [see, e.g., Alexander (1984) and Giné and Zinn (1984)]. However, we are not aware of works studying uniform bounds of the type (1.8) and (1.9) satisfying (1.5) and (1.6) for the \mathbb{L}_s -norms of such processes.

Apart from the pure probabilistic interest, development of uniform bounds on the \mathbb{L}_s -norms of processes $\{\xi_w, w \in \mathcal{W}\}$ and $\{\eta_w, w \in \mathcal{W}\}$ is motivated by problems of adaptive estimation arising in nonparametric statistics. In particular, the processes $\{\xi_w, w \in \mathcal{W}\}$ and $\{\eta_w, w \in \mathcal{W}\}$ represent stochastic errors of linear estimators with the weight w in the density estimation and nonparametric regression models, respectively. Uniform bounds on the error process are key technical tools in development of virtually all adaptive estimation procedures [see, e.g., Barron, Birgé and Massart (1999), Devroye and Lugosi (2001) Cavalier and Golubev (2006), Goldenshluger and Lepski (2008) and Golubev and Spokoiny (2009)].

The *kernel density estimator process* is a particular case of the empirical process $\{\xi_w, w \in \mathcal{W}\}$ that was frequently studied in the probabilistic literature. It is associated with the weight function w given by

$$(1.10) \quad w(t, x) = \frac{1}{n \prod_{i=1}^d h_i} K\left(\frac{t-x}{h}\right), \quad x \in \mathcal{X} = \mathbb{R}^d, t \in \mathcal{T} = \mathbb{R}^d,$$

where $K: \mathbb{R}^d \rightarrow \mathbb{R}$ is a kernel, $h = (h_1, \dots, h_d)$ is the bandwidth vector, and u/v denotes the coordinate-wise division for $u, v \in \mathbb{R}^d$. Limit laws for the \mathbb{L}_s -norms of the kernel density estimators were derived in Beirlant and Mason (1995); Dümbgen and Fatalov (2002) study exact asymptotics for the large/moderate deviation probabilities. Giné, Mason and Zaitsev (2003) investigate weak convergence of the \mathbb{L}_1 -norm kernel density estimator process indexed by a class of kernels under entropy conditions. For other closely related work, see Einmahl and Mason (2000), Giné, Koltchinskii and Zinn (2004), Giné and Nickl (2008) and references therein. We remark that the kernel density estimator process is naturally parametrized by $\mathcal{W} = \mathcal{K} \times \mathcal{H}$, where \mathcal{H} is a set of bandwidths and \mathcal{K} is a family of kernels. The *convolution kernel density estimator process* will be also studied in Section 3.4.

The inequalities (1.8) and (1.9) are useful for constructing statistical procedures provided that the following requirements are met.

- (i) *Explicit expression for $U_\psi(w, f)$.* Typically, the bound $U_\psi(w, f)$ is directly involved in the construction of statistical procedures; thus, it should be explicitly given.
- (ii) *Minimal assumptions on \mathcal{W} .* This condition is dictated by a variety of problems where the inequalities (1.8) and (1.9) can be applied. In particular, the sets \mathcal{W} may have a complicated structure (see, e.g., examples in Section 3.4).

- (iii) *Minimal assumptions on f .* The probability measure \mathbb{P} (and the expectation \mathbb{E}) as well as the right-hand sides of (1.8) and (1.9) are determined by the density f . Therefore, we want to establish (1.8) and (1.9) under weak assumptions on f . In particular, we would like to emphasize that all our results are established for the set of all probability densities uniformly bounded by a given constant. No regularity conditions are supposed.
- (iv) *Minimal assumptions on the distribution of ε .* If the process $\{\eta_w, w \in \mathcal{W}\}$ is considered, then the probability measure \mathbb{P} (and the expectation \mathbb{E}) is also determined by the distribution of ε . Therefore, we would like to have (1.8) and (1.9) under mild assumptions on this distribution. We will see that the function g given in (1.1) depends on the distribution tail of ε .

Let us briefly discuss some consequences of requirement (i) for the process $\{\xi_w, w \in \mathcal{W}\}$. Using the Talagrand concentration inequality, we prove that (1.1) holds with $U_\xi(w, f) = \mathbb{E}\|\xi_w\|_{s,\tau}$, on the space of functions $\mathfrak{H} = \{w : \sup_{x \in \mathcal{X}} \|w(\cdot, x)\|_{s,\tau} < \infty\}$. However, this bound cannot be used in statistical problems at least for two reasons.

First, it is implicit and a reasonably sharp explicit upper bound $\overline{U}_\xi(w, f)$ on $U_\xi(w, f)$ should be used instead. Sometimes if the class \mathcal{W} is not so complex (e.g., $\mathcal{W} = \mathcal{K} \times \mathcal{H}$) one can find a constant c independent of w, f and n such that

$$c\overline{U}_\xi(w, f) \leq U_\xi(w, f) \leq \overline{U}_\xi(w, f).$$

In such cases, $\overline{U}_\xi(w, f)$ can be regarded as a sharp bound on $U_\xi(w, f)$. We note, however, that establishing the above inequalities requires additional assumptions on \mathcal{W} and f and nontrivial technical work. It seems that for more complex classes \mathcal{W} the problem of finding an “optimal” upper estimate for $U_\xi(w, f)$ cannot be solved in the framework of probability theory. Contrary to that, theory of adaptive nonparametric estimation is equipped with the optimality criterion, and an upper bound $\overline{U}_\xi(w, f)$ can be regarded as sharp if it leads to the optimal statistical procedure. Thus, sharpness of $\overline{U}_\xi(w, f)$ can be assessed through accuracy analysis of the resulting statistical procedure.

Second, $U_\xi(w, f)$ [and presumably its sharp upper bound $\overline{U}_\xi(w, f)$] depends on f . In the density estimation context where the process $\{\xi_w, w \in \mathcal{W}\}$ appears, f is the parameter to be estimated. Therefore, bounds depending on f cannot be used in construction of estimation procedures. A natural idea is to replace $U_\xi(w, f)$ by its empirical counterpart $\hat{U}_\xi(w)$ whose construction is based only on the observations X_1, \dots, X_n . We adopt this strategy and establish the corresponding inequality

$$(1.11) \quad \mathbb{E} \sup_{w \in \mathcal{W}} [\|\xi_w\|_{s,\tau} - v_\epsilon(1 + \sqrt{y}\lambda_A + y\lambda_B)\hat{U}_\xi(w)]_+^q \leq \tilde{E}_{\epsilon,g}(y), \quad q \geq 1,$$

where $\tilde{E}_{\epsilon,g}(\cdot)$ differs from $E_{\epsilon,g}(\cdot)$ in (1.9) only by some absolute multiplicative factor, and, therefore, satisfies (1.5) if (1.6) holds for $\overline{U}_\xi(w, f)$. Here v_ϵ is bounded by some absolute constant and completely determined by ϵ and \mathcal{W} . We provide an explicit expression for v_ϵ .

Thus, requirement (i) leads to a new type of uniform bounds that are random. A natural question about sharpness of these bounds arises. In order to give an answer to this question, we prove that under mild assumptions on the class of weights \mathcal{W} one can choose $\epsilon = \epsilon_n \rightarrow 0$ and $y_n \rightarrow \infty$ as $n \rightarrow \infty$ so that

$$\lim_{n \rightarrow \infty} v_{\epsilon_n} (1 + \sqrt{y_n \lambda_A} + y_n \lambda_B) = 1, \quad \lim_{n \rightarrow \infty} \tilde{E}_{\epsilon_n, g}(y_n) = 0,$$

and there exists $\tilde{\epsilon}_n \rightarrow 0, n \rightarrow \infty$ such that for any subset $\mathcal{W}_0 \subseteq \mathcal{W}$ and any $q \geq 1$

$$\mathbb{E} \left[\sup_{w \in \mathcal{W}_0} \hat{U}_{\xi}(w) \right]^q \leq \left[(1 + \tilde{\epsilon}_n) \sup_{w \in \mathcal{W}_0} \bar{U}_{\xi}(w, f) \right]^q + R_n(\mathcal{W}_0),$$

where the remainder term $R_n(\mathcal{W}_0)$ is asymptotically negligible in the sense that for any $\ell > 0$ one has $\limsup_{n \rightarrow \infty} \sup_{f \in \mathcal{F}} \sup_{\mathcal{W}_0 \subseteq \mathcal{W}} [n^\ell R_n(\mathcal{W}_0)] = 0$. Here \mathcal{F} denotes the set of all probability densities uniformly bounded by a given constant [see (3.8)]. These results show that in asymptotic terms the random uniform bound is almost as good as the nonrandom one, and there is no loss of sharpness due to the use of the random uniform bound.

1.3. *Summary of results and organization of the paper.* In this paper, we develop a general machinery for finding uniform upper bounds on sub-additive positive functionals of sums of independent random functions. We start with the general setting as outlined in Section 1.1 above, and establish inequalities of the type (1.2) and (1.3) (see Proposition 2). Proofs of these results are based on the chaining and slicing/peeling techniques. The distinctive feature of our approach is that Θ is assumed to be an image of a subset \mathbb{Z} , of a metric space under some continuous mapping ϕ , that is, $\Theta = \phi(\mathbb{Z})$ as in (1.4). Then chaining on Θ is performed according to the distance induced on Θ by the mapping ϕ .

Section 3 is devoted to a systematical study of the \mathbb{L}_s -norm of the empirical process $\{\xi_w, w \in \mathcal{W}\}$. First, we derive an inequality on the tail probability of $\|\xi_w\|_{s, \tau}$ for an individual function $w \in \mathcal{W}$ (see Theorem 1 in Section 3.1). Here we use the Bernstein inequality for empirical processes proved by Bousquet (2002) and inequalities for norms of integral operators. Then in Section 3.2 we proceed with establishing uniform bounds. In Theorem 2 of Section 3.2.1, we derive uniform nonrandom bounds for $\|\xi_w\|_{s, \tau}, w \in \mathcal{W}$ that hold for all $s \geq 1$. In the case $s > 2$, the nonrandom bound depends on the density f ; therefore, for $s > 2$ we construct a random bound and present the corresponding result in Theorem 3. Theorems 2 and 3 hold for classes of weights \mathcal{W} satisfying rather general conditions. In Section 3.3, we specialize results of Theorems 2 and 3 to the classes \mathcal{W} of weights depending on the difference of their arguments. This allows us to derive explicit both nonrandom and random uniform bounds on $\|\xi_w\|_{s, \tau}$ under conditions on the weights which can be easily interpreted. The corresponding results are given in Theorems 4 and 5. We also present some asymptotic corollaries which demonstrate sharpness of the derived uniform bounds. Section 3.4 applies the results of

Theorems 4 and 5 to special examples of the set \mathcal{W} . In particular, we consider the kernel density estimator process given by (1.10), and the convolution kernel density estimator processes. It turns out that corresponding results can be formulated in a unified way, and they are presented in Theorem 7.

In Section 4, we study \mathbb{L}_s -norm of the regression-type processes $\{\eta_w, w \in \mathcal{W}\}$ given in (1.7). First, we derive an inequality on the tail probability of $\|\eta_w\|_{s,\tau}$ for an individual function $w \in \mathcal{W}$ (Theorem 8 in Section 4.1). This theorem is proved under two different types of conditions on the tail probability of the random variable ε . In Section 4.2, we present a nonrandom uniform bound for $\|\eta_w\|_{s,\tau}$ for all $s \geq 1$ over the class of weights depending on the difference of their arguments. The corresponding result is given in Theorem 9, and some asymptotic results that follow from Theorem 9 are formulated in Corollary 7. Sections 5–10 contain proofs of main results of this paper. Proofs of auxiliary lemmas are given in the Appendix.

2. Uniform bounds in general setting. In this section, we establish uniform probability bounds for the supremum of a general sub-additive functional of a random process from the probability inequality for the individual process.

Let \mathfrak{G} and \mathfrak{H} be linear topological spaces, $(\Omega, \mathfrak{A}, \mathbb{P})$ be a complete probability space, and let $\xi_\theta : \Omega \rightarrow \mathfrak{G}$, $\theta \in \mathfrak{H}$ be a family of random mappings such that:

- $\xi_\bullet(\omega)$ is linear and continuous on \mathfrak{H} for any $\omega \in \Omega$;
- $\xi_\theta(\cdot)$ is \mathfrak{A} -measurable for any $\theta \in \mathfrak{H}$.

Let $\Psi : \mathfrak{G} \rightarrow \mathbb{R}_+$ be a given *sub-additive* functional, and Θ be a fixed subset of \mathfrak{H} .

ASSUMPTION 1. There exist functions $A : \mathfrak{H} \rightarrow \mathbb{R}_+$, $B : \mathfrak{H} \rightarrow \mathbb{R}_+$, $U : \mathfrak{H} \rightarrow \mathbb{R}_+$ and $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that:

- (i) for any $z > 0$

$$\mathbb{P}\{\Psi(\xi_\theta) - U(\theta) \geq z\} \leq g\left(\frac{z^2}{A^2(\theta) + B(\theta)z}\right) \quad \forall \theta \in \mathfrak{H};$$

- (ii) the function g is monotonically decreasing to 0;
- (iii) $0 < r := \inf_{\theta \in \Theta} U(\theta) \leq \sup_{\theta \in \Theta} U(\theta) =: R < \infty$.

Condition (i) is a Bernstein-type probability inequality on $\Psi(\xi_\theta)$ for a fixed $\theta \in \mathfrak{H}$. In particular, in examples of Sections 3 and 4 we have $g(x) = ce^{-x^\alpha}$ and $g(x) = cx^{-p}$ for some $c, \alpha, p > 0$. Based on Assumption 1, our goal is to derive uniform probability and moment bounds of the type (1.2) and (1.3). For this purpose, we suppose that the set Θ is parametrized in a special way; this assumption facilitates the use of the standard chaining technique and leads to quite natural conditions on the functions U, A and B .

ASSUMPTION 2. Let (\mathfrak{Z}, d) be a metric space, and let \mathbb{Z} be a totally bounded subset of (\mathfrak{Z}, d) . There exists a *continuous* mapping ϕ from \mathfrak{Z} to \mathfrak{H} such that

$$\Theta = \{\theta \in \mathfrak{H} : \theta = \phi[\zeta], \zeta \in \mathbb{Z}\}.$$

REMARK 1. In statistical applications the set Θ is parametrized in a natural way. For instance, if, as in the introduction section, $\Psi(\cdot) = \|\cdot\|_{s,\tau}$ and $\xi_\theta = \xi_w$ with w given by (1.10), then Θ is parametrized by the kernel and bandwidth $(K, h) \in \mathcal{K} \times \mathcal{H}$. The distance d on $\mathcal{K} \times \mathcal{H}$ may have a rather special form.

Let Z be a subset of \mathbb{Z} . Define the following quantities:

$$(2.1) \quad \varkappa_U(Z) := \sup_{\zeta_1, \zeta_2 \in Z} \frac{U(\phi[\zeta_1] - \phi[\zeta_2])}{d(\zeta_1, \zeta_2)} \vee \sup_{\zeta \in Z} U(\phi[\zeta]),$$

$$(2.2) \quad \Lambda_A(Z) := \sup_{\zeta_1, \zeta_2 \in Z} \frac{A(\phi[\zeta_1] - \phi[\zeta_2])}{d(\zeta_1, \zeta_2)} \vee \sup_{\zeta \in Z} A(\phi[\zeta]),$$

$$(2.3) \quad \Lambda_B(Z) := \sup_{\zeta_1, \zeta_2 \in Z} \frac{B(\phi[\zeta_1] - \phi[\zeta_2])}{d(\zeta_1, \zeta_2)} \vee \sup_{\zeta \in Z} B(\phi[\zeta]).$$

Let $N_{Z,d}(\delta)$ denote the minimal number of balls of radius δ in the metric d needed to cover the set Z , and let $\mathcal{E}_{Z,d}(\delta) = \ln[N_{Z,d}(\delta)]$ be the δ -entropy of Z . For any $y > 0$ and $\epsilon > 0$, put

$$L_{g,Z}^{(\epsilon)}(y) = g(y) + \sum_{k=1}^{\infty} [N_{Z,d}(\epsilon 2^{-k})]^2 g(9y2^{k-3}k^{-2}).$$

Key propositions. The next two statements are the main results of this section. Define

$$(2.4) \quad C^*(y, Z) := \sqrt{y}\Lambda_A(Z) + y\Lambda_B(Z), \quad y > 0,$$

where Λ_A and Λ_B are given in (2.2) and (2.3).

PROPOSITION 1. Suppose that Assumptions 1 and 2 hold, and let Z be a subset of \mathbb{Z} such that $\sup_{\zeta, \zeta' \in Z} d(\zeta, \zeta') \leq \epsilon/4$ and $\varkappa_U(Z) < \infty$. Then for all $y > 0$ and $\epsilon > 0$ one has

$$P\left\{\sup_{\zeta \in Z} \Psi(\xi_{\phi[\zeta]}) \geq (1 + \epsilon)[\varkappa_U(Z) + C^*(y, Z)]\right\} \leq L_{g,Z}^{(\epsilon)}(y).$$

REMARK 2. Inspection of the proof of Proposition 1 shows that continuity of ξ_\bullet on \mathfrak{H} can be replaced by the assumption that $\Psi(\xi_\bullet)$ is continuous P-almost surely on $\phi[\mathbb{Z}]$ in the distance d . The latter assumption is often easier to verify in specific problems.

Define

$$(2.5) \quad \Psi_u^*(y, Z) := \sup_{\zeta \in Z} \{ \Psi(\xi_{\phi[\zeta]}) - u C^*(y) U(\phi[\zeta]) \}, \quad y > 0,$$

where $Z \subseteq \mathbb{Z}$ is a subset of \mathbb{Z} , $u \geq 1$ is a constant, and $C^*(\cdot)$ is the function defined below in (2.9). We derive bounds on the tail probability and q th moment of the random variable $\Psi_u^*(y, \mathbb{Z})$. Note that $\Psi_u^*(y, \mathbb{Z})$ is \mathfrak{A} -measurable for given y and u because the mapping $\zeta \mapsto \xi_{\phi[\zeta]}$ is \mathbb{P} -almost surely continuous, and \mathbb{Z} is a totally bounded set. By the same reason the supremum taken over any subset of \mathbb{Z} will be measurable as well.

With r and R defined in Assumption 1(iii), for any $a \in [r, R]$ consider the following subsets of \mathbb{Z} :

$$(2.6) \quad \mathbb{Z}_a := \{ \zeta \in \mathbb{Z} : a/2 < U(\phi[\zeta]) \leq a \}.$$

In words, for given $a \in [r, R]$, \mathbb{Z}_a is the *slice* of the parameter values $\zeta \in \mathbb{Z}$ for which the function $U(\phi[\zeta])$ takes values between $a/2$ and a .

In what follows, the quantities $\varkappa_U(\mathbb{Z}_a)$, $\Lambda_A(\mathbb{Z}_a)$, $\Lambda_B(\mathbb{Z}_a)$ and $L_{g, \mathbb{Z}_a}^{(\epsilon)}(y)$ will be considered as functions of $a \in [r, R]$. That is why, with slight abuse of notation, we will write

$$(2.7) \quad \varkappa_U(a) := \varkappa_U(\mathbb{Z}_a), \quad L_g^{(\epsilon)}(y, a) := L_{g, \mathbb{Z}_a}^{(\epsilon)}(y).$$

Put also

$$(2.8) \quad \Lambda_A := \sup_{a \in [r, R]} a^{-1} \Lambda_A(\mathbb{Z}_a); \quad \Lambda_B := \sup_{a \in [r, R]} a^{-1} \Lambda_B(\mathbb{Z}_a),$$

and let the function $C^*(\cdot)$ in (2.5) be defined as

$$(2.9) \quad C^*(y) := 1 + 2\sqrt{y} \Lambda_A + 2y \Lambda_B, \quad y > 0.$$

PROPOSITION 2. *Suppose that Assumptions 1 and 2 hold, and let $\varkappa_U(\mathbb{Z}) < \infty$. If*

$$(2.10) \quad \varkappa_U(a) \leq a \quad \forall a \in [r, R],$$

and if $u_\epsilon = 2^\epsilon(1 + \epsilon)$ then for any $\epsilon \in (0, 1]$, $y > 0$ and any $q \geq 1$ one has

$$(2.11) \quad \mathbb{P}\{\Psi_{u_\epsilon}^*(y, \mathbb{Z}) \geq 0\} \leq N_{\mathbb{Z}, d}(\epsilon/8) \sum_{j=0}^{[\epsilon^{-1} \log_2(R/r) - 1]_+} L_g^{(\epsilon)}(y, r2^{\epsilon(j+1)}),$$

$$(2.12) \quad \begin{aligned} \mathbb{E}[\Psi_{u_\epsilon}^*(y, \mathbb{Z})]_+^q &\leq N_{\mathbb{Z}, d}(\epsilon/8) [u_\epsilon C^*(y)]^q \\ &\times \sum_{j=0}^{[\epsilon^{-1} \log_2(R/r) - 1]_+} [r2^{\epsilon(j+1)}]^q J_g^{(\epsilon)}(y, r2^{\epsilon(j+1)}), \end{aligned}$$

where $J_g^{(\epsilon)}(z, a) := q \int_1^\infty (x - 1)^{q-1} L_g^{(\epsilon)}(zx, a) dx$.

REMARK 3.

1. Proposition 1 establishes an upper bound on the tail probability of the supremum of $\Psi(\xi_{\phi[\zeta]})$ over an arbitrary subset of \mathbb{Z} contained in a ball of radius $\epsilon/8$ in the metric d . The proof of Proposition 2 uses this bound for balls Z of the radius $\epsilon/8$ that form a covering of \mathbb{Z} . Each ball Z is divided on *slices* on which the value of $U(\phi[\zeta])$ is roughly the same. Then the supremum over \mathbb{Z} is bounded by the sum of suprema over the *slices*. This simple technique is often used in the literature on empirical processes where it is referred to as *peeling* or *slicing* [see, e.g., van de Geer (2000), Section 5.3, and Giné and Koltchinskii (2006)].
2. Note that Proposition 2 holds for any distance d on \mathfrak{Z} . Therefore, if $\varkappa_U(a)$ is proportional to a , condition (2.10) can be enforced by rescaling the distance d .

We now present a useful bound that can be easily derived from (2.11) and (2.12). Let

$$(2.13) \quad L_g^{(\epsilon)} := \sum_{k=1}^{\infty} [N_{\mathbb{Z},d}(\epsilon 2^{-k})]^2 \sqrt{g(9 \cdot 2^{k-3} k^{-2})}.$$

Note that for all $Z \subseteq \mathbb{Z}$ and $y \geq 1$

$$L_{g,Z}^{(\epsilon)}(y) \leq g(y) + L_g^{(\epsilon)} \sqrt{g(y)},$$

because $\inf_{k \geq 1} 2^k (k)^{-2} = 8/9$ and g is monotone decreasing. Therefore, we arrive to the following corollary to Proposition 2.

COROLLARY 1. *If the assumptions of Proposition 2 hold, and $L_g^{(\epsilon)} < \infty$ then for all $y \geq 1$ and $\epsilon \in (0, 1]$*

$$\begin{aligned} \mathbb{P}\{\Psi_{u_\epsilon}^*(y, \mathbb{Z}) \geq 0\} &\leq N_{\mathbb{Z},d}(\epsilon/8) [1 \vee \epsilon^{-1} \log_2(R/r)] [g(y) + L_g^{(\epsilon)} \sqrt{g(y)}], \\ \mathbb{E}[\Psi_{u_\epsilon}^*(y, \mathbb{Z})]_+^q &\leq N_{\mathbb{Z},d}(\epsilon/8) [2^{2\epsilon} R(1 + \epsilon) C^*(y)]^q [2^{q\epsilon} - 1]^{-1} J_g^{(\epsilon)}(y), \end{aligned}$$

where $J_g^{(\epsilon)}(z) = q \int_1^\infty (x - 1)^{q-1} [g(zx) + L_g^{(\epsilon)} \sqrt{g(zx)}] dx$.

3. Uniform bounds for norms of empirical processes. Based on the results obtained in Propositions 1 and 2, in this section we develop uniform bounds for the family $\{\|\xi_w\|_{s,\tau}, w \in \mathcal{W}\}$, where ξ_w is defined in (1.7). The first step here is to check Assumption 1. For this purpose, we establish an exponential inequality for $\|\xi_w\|_{s,\tau}$ when the function $w \in \mathcal{W}$ is fixed. Next, using Corollary 1 we derive a *nonrandom* uniform bound and establish corresponding inequalities of the type (1.8) and (1.9) satisfying requirements (i)–(iv) of the Introduction. We develop also a *random* uniform bound based on X_1, \dots, X_n and derive an inequality of the type (1.11).

To proceed, we need the following assumption.

ASSUMPTION (A1). Let $\overline{\mathcal{X}}$ be a countable dense subset of \mathcal{X} . For any $\varepsilon > 0$ and any $x \in \mathcal{X}$, there exists $\bar{x} \in \overline{\mathcal{X}}$ such that

$$\|w(\cdot, x) - w(\cdot, \bar{x})\|_{s,\tau} \leq \varepsilon.$$

In the sequel, we consider only the sets \mathcal{W} of $(\mathfrak{T} \times \mathfrak{X})$ -measurable functions satisfying Assumption (A1). Let

$$v'(dx) = f(x)v(dx),$$

and for any $s \in [1, \infty]$ define

$$\begin{aligned} \Sigma_s(w, f) &:= \left[\int \|w(t, \cdot)\|_{2,v'}^s \tau(dt) \right]^{1/s} \\ &= \left[\int \left(\int |w(t, x)|^2 f(x)v(dx) \right)^{s/2} \tau(dt) \right]^{1/s}, \\ M_{s,\tau,v'}(w) &:= \sup_{x \in \mathcal{X}} \|w(\cdot, x)\|_{s,\tau} \vee \sup_{t \in \mathfrak{T}} \|w(t, \cdot)\|_{s,v'}, \\ M_s(w) &:= M_{s,\tau,v}(w). \end{aligned} \tag{3.1}$$

Let $c_1(s) := 15s/\ln s$, $s > 2$, $c_2(s)$ be the constant appearing below in inequality (6.2) of Lemma 3, and define

$$\begin{aligned} c_3(s) &:= c_1(s) \vee c_2(s/(s-1)) \quad \forall s > 2, \\ c_*(s) &:= \begin{cases} 0, & 1 \leq s < 2, \\ 1, & s = 2, \\ c_3(s), & s > 2. \end{cases} \end{aligned} \tag{3.2}$$

It is worth mentioning that $c_1(s)$ is the best known constant in the Rosenthal inequality [see Johnson, Schechtman and Zinn (1985)], and in many particular examples $c_2(s) = 1$ (see Lemma 3 below). Although $c_1(s)$ is defined for $s > 2$ only, it will be convenient to set $c_1(s) = 1$ if $s \in [1, 2]$. We use this convention in what follows without further mention.

3.1. *Probability bounds for fixed weight function.* For any $w \in \mathcal{W}$, we define

$$\begin{aligned} \rho_s(w, f) &:= \begin{cases} [\sqrt{n}\Sigma_s(w, f)] \wedge [4n^{1/s}M_s(w)], & s < 2, \\ \sqrt{n}M_2(w), & s = 2, \\ c_1(s)[\sqrt{n}\Sigma_s(w, f) + 2n^{1/s}M_s(w)], & s > 2, \end{cases} \\ \omega_s^2(w, f) &:= \begin{cases} M_s^2(w)[14n + 96n^{1/s}], & s < 2, \\ 6nM_{1,\tau,v'}^2(w) + 24\sqrt{n}M_2^2(w), & s = 2, \end{cases} \end{aligned} \tag{3.3}$$

and if $s > 2$ then we set

$$\begin{aligned} \omega_s^2(w, f) &:= 6c_3(s)[nM_{2s/(s+2),\tau,v'}^2(w) \\ &\quad + 4\sqrt{n}\Sigma_s(w, f)M_s(w) + 8n^{1/s}M_s^2(w)]. \end{aligned} \tag{3.4}$$

THEOREM 1. *Let $s \in [1, \infty)$ be fixed, and suppose that Assumption (A1) holds. If $M_s(w) < \infty$, then for any $z > 0$*

$$(3.5) \quad \begin{aligned} & \mathbb{P}\{\|\xi_w\|_{s,\tau} \geq \rho_s(w, f) + z\} \\ & \leq \exp\left\{-\frac{z^2}{(1/3)\omega_s^2(w, f) + (4/3)c_*(s)M_s(w)z}\right\}, \end{aligned}$$

where $c_*(s)$ is given in (3.2).

REMARK 4. Because $c_*(s) = 0$ for $s \in [1, 2)$, the distribution of the random variable $\|\xi_w\|_{s,\tau}$ has a sub-Gaussian tail. In this case, similar bounds can be obtained from the inequalities given in [Pinelis \(1990\)](#), Theorem 2.1, [Pinelis \(1994\)](#), Theorems 3.3–3.5, and [Ledoux and Talagrand \(1991\)](#), Section 6.3. In particular, Theorem 1.2 of [Pinelis \(1990\)](#) gives the upper bound $\exp\{-z^2/2nM_s^2(w)\}$ which is better by a constant factor than our upper bound in (3.5) whenever $s \in [1, 2)$. However, if $s \geq 2$ then the cited results are not accurate enough in the sense that the corresponding bounds do not satisfy relations (1.5) and (1.6) of the [Introduction](#). It seems that only concentration principle leads to tight upper bounds; that is why we use this unified method in our derivation.

It is obvious that the upper bound of Theorem 1 remains valid if we replace $\rho_s(w, f)$, $\omega_s^2(w, f)$ and $M_s(w)$ by their upper bounds. The next result can be derived from Theorem 1 in the case $s \in [1, 2)$.

COROLLARY 2. *Let $s \in [1, 2)$ be fixed, and suppose that Assumption (A1) holds. If $M_s(w) < \infty$ then for every $z > 0$ and for all $n \geq 1$*

$$\mathbb{P}\{\|\xi_w\|_{s,\tau} \geq 4n^{1/s}M_s(w) + z\} \leq \exp\left\{-\frac{z^2}{37nM_s^2(w)}\right\}.$$

The result of the corollary is valid without any conditions on the density f . Moreover, neither the bound for $\|\xi_w\|_{s,\tau}$, nor the right-hand side of the inequality depend on f . It is important to realize that the probability inequality of Corollary 2 is sharp in some cases. In particular, it is not too difficult to construct a density f such that $\Sigma_s(w, f) = +\infty$ for any function w satisfying rather general assumptions. In this case, the established inequality seems to be sharp. On the other hand, for any density f satisfying a moment condition $\sqrt{n}\Sigma_s(w, f)$ can be bounded from above, up to a numerical constant, by $\sqrt{n}M_2(w)$ which is typically much smaller than $n^{1/s}M_s(w)$.

Several useful bounds can be derived from Theorem 1. In particular, it is shown at the end of the proof of Theorem 1 that for all $s \geq 2$ and $p \geq 1$

$$(3.6) \quad \Sigma_s(w, f) \leq M_2(w)\|\sqrt{f}\|_{s,\nu}, \quad M_{p,\tau,\nu'}(w) \leq [1 \vee \|f\|_\infty]^{1/p}M_p(w).$$

Using these inequalities, we arrive to the following result.

COROLLARY 3. *Let $s > 2$ be fixed, and suppose that Assumption (A1) holds. If $M_s(w) < \infty$, then for every $z > 0$ and for all $n \geq 1$*

$$\begin{aligned} & \mathbb{P}\{\|\xi_w\|_{s,\tau} \geq \tilde{\rho}_s(w, f) + z\} \\ & \leq \exp\left\{-\frac{z^2}{(1/3)\tilde{\omega}_s^2(w, f) + (4/3)c_3(s)M_s(w)z}\right\}, \end{aligned}$$

where $\tilde{\rho}_s(w, f) := c_1(s)[\sqrt{n}M_2(w)\|\sqrt{f}\|_{s,v} + 2n^{1/s}M_s(w)]$ and

$$\begin{aligned} \tilde{\omega}_s^2(w, f) := & 6c_3(s)\{n[1 \vee \|f\|_\infty]^{(s+2)/s}M_{2s/(s+2)}^2(w) \\ & + 4\sqrt{n}M_2(w)M_s(w)\|\sqrt{f}\|_{s,v} + 8n^{1/s}M_s^2(w)\}. \end{aligned}$$

3.2. Uniform bounds. Theorem 1 together with Corollaries 2 and 3 ensures that Assumption 1 is fulfilled for $\|\xi_w\|_{s,\tau}$. In this section, we use Proposition 2 together with Theorem 1 in order to derive a uniform over \mathcal{W} bounds on $\|\xi_w\|_{s,\tau}$.

Following the general setting of Section 2, we assume that \mathcal{W} is a parametrized set of weights, that is,

$$(3.7) \quad \mathcal{W} = \{w : w = \phi[\zeta], \zeta \in \mathbb{Z}\},$$

where \mathbb{Z} is a totally bounded subset of some metric space (\mathfrak{Z}, d) . Thus, any $w \in \mathcal{W}$ can be represented as $w = \phi[\zeta]$ for some $\zeta \in \mathbb{Z}$. Recall that $N_{\mathbb{Z},d}(\delta)$, $\delta > 0$ stands for the minimal number of balls of radius δ in the metric d needed to cover the set \mathbb{Z} , and $\mathcal{E}_{\mathbb{Z},d}(\delta) = \ln[N_{\mathbb{Z},d}(\delta)]$ is the δ -entropy of \mathbb{Z} .

The next assumption requires that the mapping $\zeta \mapsto \phi[\zeta] = w$ be continuous in the supremum norm.

ASSUMPTION (A2). For every $\varepsilon > 0$, there exists $\gamma > 0$ such that for all $\zeta_1, \zeta_2 \in \mathbb{Z}$ satisfying $d(\zeta_1, \zeta_2) \leq \gamma$ one has

$$\sup_t \sup_x |w_1(t, x) - w_2(t, x)| \leq \varepsilon,$$

where $w_1(t, x) = \phi[\zeta_1](t, x)$ and $w_2(t, x) = \phi[\zeta_2](t, x)$.

Because ξ_w is linear in w , this assumption along with Assumption (A1) guarantees that all the considered objects are measurable.

Let \mathcal{F} be the class of all probability densities uniformly bounded by constant f_∞ ,

$$(3.8) \quad \mathcal{F} := \left\{ p : \mathbb{R}^d \rightarrow \mathbb{R} : p \geq 0, \int p = 1, \|p\|_\infty \leq f_\infty < \infty \right\}.$$

It is easily seen that the inequalities of Theorem 1 and Corollary 3 can be made uniform with respect to the class \mathcal{F} . Indeed, the bound of Corollary 3 remains valid if one replaces $\|f\|_\infty$ and $\|\sqrt{f}\|_{s,\nu}$ by f_∞ and $f_\infty^{1/2-1/s}$, respectively. From now on, we suppose without loss of generality that $f_\infty \geq 1$.

3.2.1. *Uniform nonrandom bound.* Theorem 1 together with Corollaries 2 and 3 show that Assumption 1 is fulfilled for $\|\xi_w\|_{s,\tau}$ with $g(x) = e^{-x}$,

$$\begin{aligned}
 (3.9) \quad U(w) &= U_\xi(w, f) \\
 &:= \begin{cases} 4n^{1/s} M_s(w), & s \in [1, 2), \\ \sqrt{n} M_2(w), & s = 2, \\ c_1(s) [\sqrt{n} \Sigma_s(w, f) + 2n^{1/s} M_s(w)], & s > 2; \end{cases} \\
 A^2(w) &= A_\xi^2(w) \\
 &:= \begin{cases} 37n M_s^2(w), & s < 2, \\ 2f_\infty^2 n M_1^2(w) + 8\sqrt{n} M_2^2(w), & s = 2, \\ 2c_3(s) f_\infty^2 [n M_{2s/(s+2)}^2(w) + 4\sqrt{n} M_2(w) M_s(w) \\ \quad + 8n^{1/s} M_s^2(w)], & s > 2; \end{cases}
 \end{aligned}$$

and $B(w) = B_\xi(w) := \frac{4}{3} c_*(s) M_s(w)$, where $c_*(s)$ is defined in (3.2).

Put

$$(3.10) \quad r_\xi := \inf_{w \in \mathcal{W}} U_\xi(w, f), \quad R_\xi := \sup_{w \in \mathcal{W}} U_\xi(w, f).$$

Let $\varkappa_{U_\xi}(\cdot)$ be given by (2.7) with $U = U_\xi$, and

$$C_\xi^*(y) = 1 + 2\sqrt{y} \Lambda_{A_\xi} + 2y \Lambda_{B_\xi}, \quad y > 0,$$

where Λ_A and Λ_B are defined in (2.8); see also (2.9).

THEOREM 2. *Let $s \geq 1$ be fixed, (3.7) hold, and let $f \in \mathcal{F}$ if $s \geq 2$. Let Assumption (A2) be fulfilled. If $\varkappa_{U_\xi}(a) \leq a$ for all $a \in [r_\xi, R_\xi]$ then for any $y \geq 1$ and $\epsilon \in (0, 1]$ one has*

$$\begin{aligned}
 &\mathbb{P} \left\{ \sup_{w \in \mathcal{W}} [\|\xi_w\|_{s,\tau} - u_\epsilon C_\xi^*(y) U_\xi(w, f)] \geq 0 \right\} \\
 &\leq \frac{1}{\epsilon} N_{\mathbb{Z},d}(\epsilon/8) [1 \vee \log_2(R_\xi/r_\xi)] [1 + L_{\text{exp}}^{(\epsilon)}] e^{-y/2}, \\
 &\mathbb{E} \sup_{w \in \mathcal{W}} [\|\xi_w\|_{s,\tau} - u_\epsilon C_\xi^*(y) U_\xi(w, f)]_+^q \\
 &\leq \frac{2^q(\epsilon+1) u_\epsilon^q}{2^{q\epsilon} - 1} \Gamma(q+1) N_{\mathbb{Z},d}(\epsilon/8) [R_\xi C_\xi^*(1)]^q [1 + L_{\text{exp}}^{(\epsilon)}] e^{-y/2},
 \end{aligned}$$

where $u_\epsilon = 2^\epsilon(1 + \epsilon)$, $\Gamma(\cdot)$ is the gamma-function and

$$(3.11) \quad L_{\text{exp}}^{(\epsilon)} = \sum_{k=1}^{\infty} \exp\{2\mathcal{E}_{\mathbb{Z},d}(\epsilon 2^{-k}) - (9/16)2^k k^{-2}\}.$$

The proof follows immediately by application of Corollary 1, and noting that for $g(x) = e^{-x}$ the quantity $L_g^{(\epsilon)}$ is given by the above formula [cf. (2.13)], while $J_g^{(\epsilon)}(\cdot)$ for $g(x) = e^{-x}$ is bounded as follows

$$\begin{aligned} J_g^{(\epsilon)}(z) &= q \int_1^{\infty} (x - 1)^{q-1} [e^{-zx} + L_g^{(\epsilon)} \sqrt{e^{-zx}}] dx \\ &\leq \Gamma(q + 1)[1 + L_{\text{exp}}^{(\epsilon)}](2/z)^q e^{-z/2}. \end{aligned}$$

REMARK 5. It is instructive to compare the results of Theorem 2 with those of Theorem 1 (and Corollaries 2 and 3). The uniform bound on $\|\xi_w\|_{s,\tau}$ in Theorem 2 is determined by the individual bound $U_\xi(w, f)$ for a fixed weight $w \in \mathcal{W}$, and by the function $C_\xi^*(\cdot)$ which, in its turn, is computed on the basis of $A_\xi(w)$, $B_\xi(w)$ and $U_\xi(w, f)$. The function $C_\xi^*(\cdot)$ depends on the parametrization (3.7) and on the distance d on \mathfrak{Z} via the quantities Λ_A and Λ_B [see (2.8)]. The right-hand sides of the inequalities in Theorem 2 depend on massiveness of the set of weights \mathcal{W} as measured by the entropy $\mathcal{E}_{\mathbb{Z},d}(\cdot)$. Note also that these bounds decrease exponentially in y .

3.2.2. *Uniform random bound.* The uniform nonrandom bounds on $\|\xi_w\|_{s,\tau}$ given in Theorem 2 depend on the density f via $U_\xi(w, f)$. As discussed in the Introduction, this does not allow one to use this bound in statistical problems. Our goal is to recover the statement of Theorem 2 (up to some numerical constants) with the unknown quantity $U_\xi(w, f)$ replaced by its estimator $\hat{U}_\xi(w)$. Note also that $U_\xi(w, f)$ of Theorem 2 depends on f only if $s > 2$; here the quantity depending on f is $\Sigma_s(w, f)$.

Assume that the conditions of Theorem 2 are satisfied, and let $s > 2$. For any $t \in \mathcal{T}$ define

$$(3.12) \quad \hat{\Sigma}_s(w) := \|S_w\|_{s,\tau}, \quad S_w^2(t) := \frac{1}{n} \sum_{i=1}^n w^2(t, X_i),$$

$$(3.13) \quad \hat{U}_\xi(w) := c_1(s)[\sqrt{n} \hat{\Sigma}_s(w) + 2n^{1/s} M_s(w)].$$

It is easily seen that $\hat{U}_\xi(w)$ is a reasonable estimate of $U_\xi(w, f)$ because under mild assumptions for any fixed $t \in \mathcal{T}$ by the law of large numbers

$$S_w^2(t) - \|w(t, \cdot)\|_{2,\nu'}^2 \rightarrow 0, \quad n \rightarrow \infty \quad \text{in probability.}$$

Moreover,

$$\begin{aligned}
 |\hat{\Sigma}_s(w) - \Sigma_s(w, f)|^2 &\leq \|S_w - \|w(\cdot, \cdot)\|_{2, \nu'}\|_{s, \tau}^2 \\
 &\leq \|\sqrt{|S_w^2 - \|w(\cdot, \cdot)\|_{2, \nu'}^2}\|_{s, \tau}^2 \\
 &= \|S_w^2 - \|w(\cdot, \cdot)\|_{2, \nu'}^2\|_{s/2, \tau} \\
 &= \left\| \frac{1}{n} \sum_{i=1}^n [w^2(\cdot, X_i) - \mathbb{E}w^2(\cdot, X)] \right\|_{s/2, \tau}.
 \end{aligned}$$

Thus, for any $s > 2$ we have

$$(3.14) \quad |\hat{\Sigma}_s(w) - \Sigma_s(w, f)| \leq \sqrt{\frac{\|\xi_{w^2}\|_{s/2, \tau}}{n}},$$

that is, the difference between $\hat{\Sigma}_s(w)$ and $\Sigma_s(w, f)$ is controlled in terms of $\|\xi_{w^2}\|_{s/2, \tau}$. The idea now is to use Theorem 2 in order to find a nonrandom upper bound on $\|\xi_{w^2}\|_{s/2, \tau}$. One can expect that this bound will be much smaller than $\Sigma_s(w, f)$ provided that the function w is small enough. If this is true then $\hat{\Sigma}_s(w)$ approximates well $\Sigma_s(w, f)$, and it can be used instead of $\Sigma_s(w, f)$ in the definition of the uniform over \mathcal{W} upper bound on $\|\xi_w\|_{s, \tau}$.

In order to control uniformly $\|\xi_{w^2}\|_{s/2, \tau}$ by applying Theorem 1 and Corollary 1, we need the following definitions. Put

$$\begin{aligned}
 \tilde{U}(w^2) &:= \begin{cases} 4n^{2/s} M_{s/2}(w^2), & s \in (2, 4), \\ c_1(s/2) [f_\infty^{1/2} \sqrt{n} M_2(w^2) + 2n^{2/s} M_{s/2}(w^2)], & s \geq 4; \end{cases} \\
 \tilde{A}^2(w^2) &:= \begin{cases} 37n M_{s/2}^2(w^2), & s \in (2, 4), \\ 2c_3(s/2) f_\infty^2 [n M_{2s/(s+4)}^2(w^2) \\ \quad + 4\sqrt{n} M_2(w^2) M_{s/2}(w^2) \\ \quad + 8n^{2/s} M_{s/2}^2(w^2)], & s \geq 4; \end{cases}
 \end{aligned}$$

and $\tilde{B}(w^2) := \frac{4}{3} c_*(s/2) M_{s/2}(w^2)$, where $c_*(\cdot)$ is given in (3.2).

For any subset $Z \subseteq \mathbb{Z}$, let $\varkappa_{\tilde{U}}(Z)$, $\Lambda_{\tilde{A}}(Z)$, and $\Lambda_{\tilde{B}}(Z)$ be given by (2.1)–(2.3) with $U = \tilde{U}$, $A = \tilde{A}$ and $B = \tilde{B}$. With r_ξ and R_ξ defined in (3.10), let

$$(3.15) \quad \mathbb{Z}_a = \{\zeta \in \mathbb{Z} : a/2 < U_\xi(w, f) \leq a\}, \quad a \in [r_\xi, R_\xi],$$

and we set

$$(3.16) \quad \begin{aligned}
 \varkappa_{\tilde{U}}(a) &:= \varkappa_{\tilde{U}}(\mathbb{Z}_a), \quad \lambda_{\tilde{A}} = \sup_{a \in [r_\xi, R_\xi]} a^{-2} \Lambda_{\tilde{A}}(\mathbb{Z}_a), \\
 \lambda_{\tilde{B}} &= \sup_{a \in [r_\xi, R_\xi]} a^{-2} \Lambda_{\tilde{B}}(\mathbb{Z}_a),
 \end{aligned}$$

[cf. (2.7) and (2.8)]. It is important to emphasize here that in the definition of $\varkappa_{\tilde{U}}, \lambda_{\tilde{A}}$ and $\lambda_{\tilde{B}}$ we use the same set \mathbb{Z}_a as in the definition of $\varkappa_{U_\xi}(\cdot), \Lambda_{A_\xi}(\cdot)$ and $\Lambda_{B_\xi}(\cdot)$.

The next result establishes a random uniform bound on $\|\xi_w\|_{s,\tau}$.

THEOREM 3. *Let $s > 2$ be fixed, (3.7) hold, Assumption (A2) be fulfilled, and*

$$(3.17) \quad \varkappa_{U_\xi}(a) \leq a \quad \forall a \in [r_\xi, R_\xi].$$

Let $\epsilon \in (0, 1]$ be fixed, and suppose that there exists a positive number $\gamma < [4c_1(s)(1 + \epsilon)]^{-1}$ such that

$$(3.18) \quad \varkappa_{\tilde{U}}(a) \leq (\gamma a)^2 \quad \forall a \in [r_\xi, R_\xi].$$

If y_γ denotes the root of the equation

$$(3.19) \quad \sqrt{y}\lambda_{\tilde{A}} + y\lambda_{\tilde{B}} = \gamma^2,$$

and if $y_\gamma > 1$ then:

(i) *For every $y \in [1, y_\gamma]$ one has*

$$\mathbb{E} \sup_{w \in \mathcal{W}} \{ \|\xi_w\|_{s,\tau} - \bar{u}_\epsilon(\gamma) C_\xi^*(y) \hat{U}_\xi(w) \}_+^q \leq T_{1,\epsilon} [C_\xi^*(y)]^q \exp\{-y/2\},$$

where $\bar{u}_\epsilon(\gamma) := u_\epsilon [1 - 4c_1(s)(1 + \epsilon)\gamma]^{-1}$ and $u_\epsilon = 2^\epsilon(1 + \epsilon)$.

(ii) *For any subset $\mathcal{W}_0 \subseteq \mathcal{W}$, one has*

$$(3.20) \quad \mathbb{E} \left[\sup_{w \in \mathcal{W}_0} \hat{U}_\xi(w) \right]^q \leq [1 + 4c_1(s)(1 + \epsilon)\gamma]^q \sup_{w \in \mathcal{W}_0} [U_\xi(w, f)]^q + T_{2,\epsilon} \left[\sqrt{n} \sup_{w \in \mathcal{W}_0} M_s(w) \right]^q \exp\{-y_\gamma/2\}.$$

The explicit expressions for the constants $T_{1,\epsilon}$ and $T_{2,\epsilon}$ are given in the beginning of proof of the theorem.

REMARK 6.

1. Theorem 3 requires two sets of conditions: conditions of Theorem 2, and conditions on behavior of the functions $\varkappa_{\tilde{U}}(\cdot), \Lambda_{\tilde{A}}(\cdot)$ and $\Lambda_{\tilde{B}}(\cdot)$ on the slices \mathbb{Z}_a defined through $U_\xi(w, f)$.
2. The parameter γ controls closeness of $\hat{U}_\xi(\cdot)$ to $U_\xi(\cdot, f)$: the smaller γ , the closer the random bound $\hat{U}_\xi(\cdot)$ to the nonrandom one $U_\xi(\cdot, f)$ [see (3.20)]. In this case, we do not loose much if $U_\xi(\cdot, f)$ is replaced by its empirical counterpart $\hat{U}_\xi(w)$. Clearly, it is possible to choose γ small and simultaneously to keep y_γ large only if $\lambda_{\tilde{A}}$ and $\lambda_{\tilde{B}}$ are small enough. Fortunately, this is the case in many examples.
3. Note also that when γ approaches $[4c_1(s)(1 + \epsilon)]^{-1}$ the parameter $\bar{u}_\epsilon(\gamma)$ increases to infinity [clearly, we want to keep $\bar{u}_\epsilon(\gamma)$ as close to one as possible].

Thus, the assumption $\gamma < [4c_1(s)(1 + \epsilon)]^{-1}$ is important; this poses a restriction on the parameter set \mathcal{W} . We conjecture that the following condition is necessary: for given $s \geq 2$ there exists a universal constant, say, $c(s)$, such that $\gamma < c(s)$.

The next corollary to Theorem 3 will be useful in what follows.

COROLLARY 4. *The statements of Theorem 3 remain valid if $U_\xi(w, f)$ and $\hat{U}_\xi(w)$ are redefined as $\max\{U_\xi(w, f), \sqrt{n}M_2(w)\}$ and $\max\{\hat{U}_\xi(w), \sqrt{n}M_2(w)\}$, respectively.*

3.3. Uniform bounds for classes of weights depending on the difference of arguments. As we have seen, the results and assumptions in Theorems 2 and 3 are stated in terms of the quantities (such as $\lambda_{\tilde{A}}, \lambda_{\tilde{B}}, y_\gamma$) that are given implicitly. In particular, additional computations are still necessary in order to apply Theorems 2 and 3 in specific problems. In this section, we specialize the results of Theorems 2 and 3 for the classes of weights \mathcal{W} depending on the difference of arguments. Under natural and easily interpretable assumptions on the class of such weights, we derive explicit uniform bounds on the norms of empirical processes.

Throughout this section, $\mathcal{X} = \mathcal{T} = \mathbb{R}^d$, $\tau = \nu = \text{mes}$ is the Lebesgue measure and we write $\|\cdot\|_s$ instead of $\|\cdot\|_{s,\tau}$. In this section, the class of weights \mathcal{W} is a set of functions from $\mathbb{R}^d \times \mathbb{R}^d$ to \mathbb{R} of the following form

$$(3.21) \quad \mathcal{W} = \{w(t - x), w \in \mathcal{V}\},$$

where \mathcal{V} is a given set of d -variate functions. For the sake of notational convenience, we will identify the weight $w \in \mathcal{W}$ with the d -variate function $w \in \mathcal{V}$ in the definition of the process ξ_w and the quantities such as U_ξ, A_ξ, B_ξ etc. Thus, when we write $w \in \mathcal{W}$ we mean the weight $w(\cdot - \cdot)$ while $w \in \mathcal{V}$ denotes the corresponding d -variate function; this should not lead to a confusion.

Let (\mathfrak{Z}, d) be a fixed metric space; as before, we suppose that \mathcal{V} is parametrized by the parameter $\zeta \in \mathfrak{Z}$, that is,

$$(3.22) \quad \mathcal{V} = \{w : w = \phi[\zeta], \zeta \in \mathfrak{Z}\},$$

where \mathfrak{Z} is a totally bounded subset of the metric space (\mathfrak{Z}, d) . Recall that $N_{\mathfrak{Z},d}(\delta)$, $\delta > 0$ is the number of the balls of radius δ in the metric d that form a minimal covering of the set \mathfrak{Z} .

We need the following assumptions on the class of weights \mathcal{W} (the functional set \mathcal{V}).

ASSUMPTION (W).

(W1) The Lebesgue measure of support of all functions from \mathcal{V} is finite, that is,

$$(3.23) \quad \mu_* := \sup_{w \in \mathcal{V}} \text{mes}\{\text{supp}(w)\} < \infty.$$

(W2) There exist real numbers $\alpha_1 \in (0, 1)$ and $\alpha_2 \in (0, 1)$ such that

$$\text{mes}\{x \in \mathbb{R}^d : |w(x)| \geq \alpha_1 \|w\|_\infty\} \geq \alpha_2 \text{mes}\{\text{supp}(w)\} \quad \forall w \in \mathcal{V}.$$

(W3) There exists a real number $\mu \geq 1$ such that

$$n \text{mes}\{\text{supp}(w)\} \geq \mu \quad \forall w \in \mathcal{V}.$$

(W4) There exists a real number $\beta \in (0, 1)$ such that

$$\sup_{\delta \in (0,1)} \{\ln[N_{\mathbb{Z},d}(\delta)] - \delta^{-\beta}\} =: C_{\mathbb{Z}}(\beta) < \infty.$$

REMARK 7. We will show that Assumption (W2) is fulfilled if \mathcal{V} is a set of smooth functions. Assumption (W3) together with (W2) allows one to establish relations between \mathbb{L}_p -norms of functions from \mathcal{V} ; this will be extensively used in what follows. Assumption (W4) is a usual entropy condition. In particular, (W4) ensures that the quantity $L_{\text{exp}}^{(\epsilon)}$ in (3.11) is finite.

In addition to Assumption (W), we will need the following assumption on the properties of the mapping ϕ in (3.22). For $p \geq 1$, put

$$(3.24) \quad 0 < \underline{w}_p := n^{1/p} \inf_{w \in \mathcal{V}} \|w\|_p \leq n^{1/p} \sup_{w \in \mathcal{V}} \|w\|_p =: \overline{w}_p < \infty$$

and define

$$(3.25) \quad \mathbb{Z}_p(b) := \{\zeta \in \mathbb{Z} : n^{1/p} \|\phi[\zeta]\|_p \leq b\}, \quad b \in [\underline{w}_p, \overline{w}_p].$$

ASSUMPTION (L). The mapping ϕ in (3.22) satisfies the following conditions:

- if $s \in [1, 2)$ then

$$\sup_{\zeta_1, \zeta_2 \in \mathbb{Z}_s(b)} \frac{n^{1/s} \|\phi[\zeta_1] - \phi[\zeta_2]\|_s}{d(\zeta_1, \zeta_2)} \leq b \quad \forall b \in [\underline{w}_s, \overline{w}_s],$$

- if $s \geq 2$ then

$$(3.26) \quad \sup_{\zeta_1, \zeta_2 \in \mathbb{Z}_2(b)} \frac{\sqrt{n} \|\phi[\zeta_1] - \phi[\zeta_2]\|_2}{d(\zeta_1, \zeta_2)} \leq b \quad \forall b \in [\underline{w}_2, \overline{w}_2].$$

We note that Assumption (L) guarantees continuity of $\|\xi_w\|_s$ on $\phi[\mathbb{Z}]$ for any $s \leq 2$. The same property for $s > 2$ follows from Lemma 7. This, in view of Remark 2, replaces Assumption (A2).

The next statement presents the uniform moment bound on $\|\xi_w\|_s$ when $s \in [1, 2]$, and \mathcal{W} is given by (3.21).

THEOREM 4. *Let the class of weights \mathcal{W} be defined by (3.21), and let (3.22) and Assumptions (W1), (W4) and (L) hold.*

(i) If $s \in [1, 2)$ then for all $n \geq 1, z \geq [\sqrt{37}/2]n^{1/2-1/s}$, and $\epsilon \in (0, 1]$ one has

$$\mathbb{E} \sup_{w \in \mathcal{W}} [\|\xi_w\|_s - 4u_\epsilon(1+z)n^{1/s}\|w\|_s]_+^q \leq T_{3,\epsilon}n^{q/s} \exp\left\{-\frac{2z^2}{37}n^{(2/s)-1}\right\}.$$

(ii) If $f \in \mathcal{F}$ then for all $n \geq 1, z \geq \sqrt{8[\mu_*f_\infty^2 + 4n^{-1/2}]}$, and $\epsilon \in (0, 1]$ one has

$$\begin{aligned} \mathbb{E} \sup_{w \in \mathcal{W}} [\|\xi_w\|_2 - u_\epsilon(1+z+z^2/12)\sqrt{n}\|w\|_2]_+^q \\ \leq T_{4,\epsilon}n^{q/2} \exp\left\{-\frac{z^2}{16[f_\infty^2\mu_* + 4n^{-1/2}]}\right\}. \end{aligned}$$

The explicit expressions for the constants $T_{3,\epsilon}$ and $T_{4,\epsilon}$ are given in the beginning of the proof of the theorem.

The bound of Theorem 4 is nonrandom because $U_\xi(w, f)$ does not depend on f whenever $s \in [1, 2]$. The proof of this statement is based on application of Theorem 2.

Now we proceed with the case $s > 2$. Here we need some further notation. Given $p \geq 2$, let $m_p \in (0, 1]$ be such that

$$(3.27) \quad \sup_{b \in [\underline{w}_2, \bar{w}_2]} b^{-1} \sup_{\zeta_1, \zeta_2 \in \mathbb{Z}_2(b)} \frac{n^{1/p} \|\phi[\zeta_1] - \phi[\zeta_2]\|_p}{[d(\zeta_1, \zeta_2)]^{m_p}} =: C_p < \infty.$$

Existence of $m_p \in (0, 1]$ such that (3.27) holds is ensured by Lemma 7 given in Section 8.2. In particular, it is shown there that if Assumptions (W) and (L) hold then m_p can be taken equal to $2/p$. We note also that $m_2 = 1$ and $C_2 = 1$ by Assumption (L).

Following (3.12), (3.13) and Corollary 4, we set

$$(3.28) \quad \begin{aligned} \hat{U}_\xi(w) &= c_1(s) \left\{ \sqrt{n} \left\| \left[\frac{1}{n} \sum_{i=1}^n w^2(\cdot - X_i) \right]^{1/2} \right\|_s + 2n^{1/s}\|w\|_s \right\}, \\ \check{U}_\xi(w) &:= \max\{\hat{U}_\xi(w), \sqrt{n}\|w\|_2\}, \\ \bar{U}_\xi(w) &:= \max\{U_\xi(w, f), \sqrt{n}\|w\|_2\}. \end{aligned}$$

Put also

$$(3.29) \quad \begin{aligned} C_\xi^*(y) &= 1 + 2\vartheta_0 \{ \sqrt{y} [\mu_*^{1/s} + n^{-1/(2s)}] + yn^{-1/s} \}, \\ m &:= \begin{cases} 1 \wedge m_s, & s \in (2, 4), \\ 1 \wedge m_s \wedge m_{s/2}, & s \geq 4, \end{cases} \end{aligned}$$

where $\vartheta_0 := 5c_1(s)[C_s \vee 1]f_\infty\alpha_1^{-1}\alpha_2^{-1/2}$, α_1 and α_2 are given in Assumption (W2), and m_p and C_p are defined in (3.27).

THEOREM 5. *Let Assumptions (W) and (L) hold, and assume that $f \in \mathcal{F}$. Suppose that (W3) is fulfilled with $\mu > [64c_1^2(s)]^{(s \wedge 4)/(s \wedge 4 - 2)}$, and (W4) is fulfilled with $\beta < m$. Let $\gamma = \mu^{1/(s \wedge 4) - 1/2}$, and*

$$(3.30) \quad y_* := \begin{cases} \vartheta_1 n^{4/s - 1}, & s \in (2, 4), \\ \vartheta_2 \mu^{-1/2} [\mu_*^{2/s} + n^{-1/s}]^{-2}, & s \geq 4, \end{cases}$$

with constants ϑ_1 and ϑ_2 specified explicitly in the proof; then for any $s > 2$ and $y \in [1, y_*]$ one has

$$\mathbb{E} \sup_{w \in \mathcal{W}} \{ \|\xi_w\|_s - \bar{u}_\epsilon(\gamma) C_\xi^*(y) \check{U}_\xi(w) \}_+^q \leq T_{5,\epsilon} n^{q/2} [C_\xi^*(y)]^q \exp\{-y/2\},$$

where $\bar{u}_\epsilon(\cdot)$ is defined in Theorem 3. In addition, if $\mathcal{W}_0 \subseteq \mathcal{W}$ is an arbitrary subset of \mathcal{W} then

$$\begin{aligned} \mathbb{E} \left[\sup_{w \in \mathcal{W}_0} \check{U}_\xi(w) \right]^q &\leq \left\{ [1 + 4c_1(s)(1 + \epsilon)\mu^{1/(s \wedge 4) - 1/2}] \sup_{w \in \mathcal{W}_0} \bar{U}_\xi(w) \right\}^q \\ &\quad + T_{6,\epsilon} n^{q(s-2)/(2s)} \exp\{-y_*/2\}. \end{aligned}$$

The explicit expressions for the constants $T_{5,\epsilon}$ and $T_{6,\epsilon}$ are given in the proof.

Theorem 5 establishes random uniform bounds on the norms of empirical processes in terms of the parameters determining the class \mathcal{W} . In particular, the parameters μ and μ_* play an important role. Theorem 5 leads to a number of powerful asymptotic results that demonstrate sharpness of the proposed random bound.

COROLLARY 5. *Let assumptions of Theorem 5 hold, and let $s > 2$ be fixed. There exist positive constants $k_i = k_i(s)$, $i = 1, 2, 3$ such that if*

$$\begin{aligned} \mu &= \mu_n \asymp [\ln n]^{k_1}, & \mu_* &= \mu_{*,n} \asymp [\ln n]^{-k_2}, \\ \epsilon &= \epsilon_n \asymp [\ln n]^{-k_3}, & n &\rightarrow \infty, \end{aligned}$$

then for all $\ell > 0$ and $q \geq 1$

$$\begin{aligned} \lim_{n \rightarrow \infty} \sup_{f \in \mathcal{F}} n^\ell \mathbb{E} \sup_{w \in \mathcal{W}} [\|\xi_w\|_s - (1 + 3\epsilon_n) \check{U}_\xi(w)]_+^q &= 0, \\ \mathbb{E} \left[\sup_{w \in \mathcal{W}_0} \check{U}_\xi(w) \right]^q &\leq \left[(1 + \epsilon_n) \sup_{w \in \mathcal{W}_0} \bar{U}_\xi(w, f) \right]^q + R_n(\mathcal{W}_0), \end{aligned}$$

where $\limsup_{n \rightarrow \infty} \sup_{f \in \mathcal{F}} \sup_{\mathcal{W}_0 \subseteq \mathcal{W}} [n^\ell R_n(\mathcal{W}_0)] = 0$.

The explicit expressions for the constants $k_i > 0$, $i = 1, 2, 3$ are easily computed from Theorem 5.

REMARK 8. Corollary 5 shows that if the class of weights \mathcal{W} is such that $\mu = \mu_n$ and $\mu_* = \mu_{*,n}$, and if ϵ is set to be $\epsilon = \epsilon_n$, then $(1 + 3\epsilon_n)\check{U}_n(w)$ is a uniform random bound on $\|\xi_w\|_s$ which is asymptotically almost as good as the nonrandom bound $\bar{U}_\xi(w, f)$ depending on f . Thus, in asymptotic terms, there is no loss in sharpness of the random uniform bound in comparison with the nonrandom bound that depends on f .

3.4. *Specific problems.* In this section, we consider process ξ_w corresponding to special classes of weights \mathcal{W} that arise in kernel density estimation. Using results of Theorems 4 and 5, we derive uniform bounds on the norms of these processes. As in Section 3.3, here $\mathcal{X} = \mathcal{T} = \mathbb{R}^d$, and ν and τ are both the Lebesgue measure.

Let \mathcal{K} be a given set of real functions defined on \mathbb{R}^d and suppose that \mathcal{K} is a *totally bounded* set with respect to the \mathbb{L}_∞ -norm. Let $\mathcal{H} := \otimes_{i=1}^d [h_i^{\min}, h_i^{\max}]$, where the vectors $h^{\min} = (h_1^{\min}, \dots, h_d^{\min})$, $h^{\max} = (h_1^{\max}, \dots, h_d^{\max})$, $0 < h_i^{\min} \leq h_i^{\max} \leq 1, \forall i = 1, \dots, d$ are fixed.

For any $h \in \mathcal{H}$ define $V_h := \prod_{i=1}^d h_i$, and endow the set \mathcal{H} with the following distance:

$$(3.31) \quad \Delta_{\mathcal{H}}(h, h') = \max_{i=1, \dots, d} \ln \left(\frac{h_i \vee h'_i}{h_i \wedge h'_i} \right).$$

In order to verify that $\Delta_{\mathcal{H}}$ is indeed a distance on \mathcal{H} it suffices to note that the function $(x, y) \mapsto \ln(x \vee y) - \ln(x \wedge y)$, $x > 0, y > 0$ satisfies all axioms of distance on $\mathbb{R}_+ \setminus \{0\}$.

We will be interested in the following classes of weights \mathcal{W} and the corresponding processes ξ_w .

Kernel density estimator process. With any $K \in \mathcal{K}$ and $h \in \mathcal{H}$, we associate the weight function

$$w(t - x) = n^{-1} K_h(t - x) := (n V_h)^{-1} K[(t - x)/h].$$

As before, $u/v, u, v \in \mathbb{R}^d$, stands for the coordinate-wise division $(u_1/v_1, \dots, u_d/v_d)$.

The weight w is naturally parametrized by K and h so that we put

$$(3.32) \quad \mathbb{Z}^{(1)} := \mathcal{K} \times \mathcal{H}, \quad \zeta = (K, h), \quad w = \phi_1[\zeta] := n^{-1} K_h.$$

We equip $\mathbb{Z}^{(1)}$ with the family of distances $\{d_\vartheta^{(1)}(\cdot, \cdot), \vartheta > 0\}$ defined by

$$d_\vartheta^{(1)}(\zeta, \zeta') = \vartheta \max\{\|K - K'\|_\infty, \Delta_{\mathcal{H}}(h, h')\},$$

$$\zeta = (K, h), \quad \zeta' = (K', h'), \quad \vartheta > 0.$$

Obviously, $\mathbb{Z}^{(1)}$ is a totally bounded set with respect to $d_\vartheta^{(1)}$ for any $\vartheta > 0$.

The corresponding family of random fields is

$$(3.33) \quad \xi_w^{(1)}(t) := \xi_{\phi_1[\zeta]}(t) = \frac{1}{n} \sum_{i=1}^n \{K_h(t - X_i) - \mathbb{E}K_h(t - X)\}, \quad \zeta \in \mathbb{Z}^{(1)},$$

and we are interested in bounds on the \mathbb{L}_s -norm of this process uniform over the class of weights

$$\mathcal{W}^{(1)} := \{w(\cdot - \cdot) = n^{-1}K_h(\cdot - \cdot) : (K, h) \in \mathbb{Z}^{(1)}\}.$$

We note that $\xi_w^{(1)}$ is the stochastic error of the kernel density estimator associated with the kernel $K \in \mathcal{K}$ and bandwidth $h \in \mathcal{H}$. According to Theorems 4 and 5, for the process $\{\xi_w, w \in \mathcal{W}^{(1)}\}$, the uniform bounds on $\|\xi_w\|_s$ should be based on the following functionals. Define

$$U_\xi^{(1)}(w) := \begin{cases} 4(nV_h)^{1/s-1} \|K\|_s, & s \in [1, 2), \\ (nV_h)^{-1/2} \|K\|_2, & s = 2. \end{cases}$$

For $s > 2$, we put

$$U_\xi^{(1)}(w, f) := c_1(s) \left[n^{-1/2} \left(\int \left[\int K_h^2(t-x) f(x) dx \right]^{s/2} dt \right)^{1/s} + 2(nV_h)^{1/s-1} \|K\|_s \right],$$

$$\hat{U}_\xi^{(1)}(w) := c_1(s) \left[n^{-1/2} \left(\int \left[n^{-1} \sum_{i=1}^n K_h^2(t - X_i) \right]^{s/2} dt \right)^{1/s} + 2(nV_h)^{1/s-1} \|K\|_s \right],$$

and finally

$$(3.34) \quad \begin{aligned} \bar{U}_\xi^{(1)}(w, f) &:= \max[U_\xi^{(1)}(w, f), (nV_h)^{-1/2} \|K\|_2] \\ \check{U}_\xi^{(1)}(w) &:= \max[\hat{U}_\xi^{(1)}(w), (nV_h)^{-1/2} \|K\|_2]. \end{aligned}$$

Convolution kernel density estimator process. For any $(K, h) \in \mathbb{Z}^{(1)}$ and $(Q, \mathfrak{h}) \in \mathbb{Z}^{(1)}$, we define

$$(3.35) \quad w(t - x) = n^{-1}[K_h * Q_{\mathfrak{h}}](t - x),$$

where $\mathbb{Z}^{(1)}$ is defined in (3.32), and $*$ stands for the convolution on \mathbb{R}^d . Put

$$\mathbb{Z}^{(2)} := \mathbb{Z}^{(1)} \times \mathbb{Z}^{(1)}, \quad z = [(K, h), (Q, \mathfrak{h})], \quad w = \phi_2[z] = n^{-1}(K_h * Q_{\mathfrak{h}}),$$

and define the family of distances on $\mathbb{Z}^{(2)}$ as

$$d_{\vartheta}^{(2)}(z, z') = \vartheta \max\{\|K - K'\|_{\infty} \vee \|Q - Q'\|_{\infty}, \Delta_{\mathcal{H}}(h, h') \vee \Delta_{\mathcal{H}}(\mathfrak{h}, \mathfrak{h}')\},$$

$\vartheta > 0,$

where $z = [(K, h), (Q, \mathfrak{h})], z' = [(K', h'), (Q', \mathfrak{h}')] , z, z' \in \mathbb{Z}^{(2)}$. Obviously, $\mathbb{Z}^{(2)}$ is a totally bounded set with respect to the distance $d_{\vartheta}^{(2)}$ for any $\vartheta > 0$.

The corresponding family of random fields is

$$\begin{aligned} \xi_w^{(2)}(t) &:= \xi_{\phi_2[z]}(t) \\ (3.36) \quad &= \frac{1}{n} \sum_{i=1}^n \{[K_h * Q_{\mathfrak{h}}](t - X_i) - \mathbb{E}[K_h * Q_{\mathfrak{h}}](t - X)\}, \end{aligned}$$

$\zeta \in \mathbb{Z}^{(2)},$

and we are interested in a uniform bound on $\|\xi_w^{(2)}\|_s$ over

$$\mathcal{W}^{(2)} := \{w(\cdot - \cdot) = n^{-1} K_h * Q_{\mathfrak{h}}(\cdot - \cdot), [(K, h), (Q, \mathfrak{h})] \in \mathbb{Z}^{(2)}\}.$$

The random field ξ_w with w given by (3.35) appears in the context of multivariate density estimation. In particular, the uniform bounds on $\|\xi_w\|_s$ are instrumental in construction of a selection rule for the family of kernel estimators parametrized by $\mathcal{K} \times \mathcal{H}$ [see Goldenshluger and Lepski (2009)]. Theorems 4 and 5 suggest to base the uniform bounds on the following quantities. Define

$$U_{\xi}^{(2)}(w) := \begin{cases} 4n^{1/s-1} \|K_h * Q_{\mathfrak{h}}\|_s, & s \in [1, 2), \\ n^{-1/2} \|K_h * Q_{\mathfrak{h}}\|_2, & s = 2. \end{cases}$$

For $s > 2$, we put

$$\begin{aligned} U_{\xi}^{(2)}(w, f) &:= c_1(s) \left[n^{-1/2} \left(\int \left[\int [K_h * Q_{\mathfrak{h}}]^2(t - x) f(x) dx \right]^{s/2} dt \right)^{1/s} \right. \\ &\quad \left. + 2n^{1/s-1} \|K_h * Q_{\mathfrak{h}}\|_s \right], \\ \hat{U}_{\xi}^{(2)}(w) &:= c_1(s) \left[n^{-1/2} \left(\int \left[n^{-1} \sum_{i=1}^n [K_h * Q_{\mathfrak{h}}]^2(t - X_i) \right]^{s/2} dt \right)^{1/s} \right. \\ &\quad \left. + 2n^{1/s-1} \|K_h * Q_{\mathfrak{h}}\|_s \right]; \end{aligned}$$

and finally

$$\begin{aligned} (3.37) \quad \bar{U}_{\xi}^{(2)}(w, f) &:= \max[U_{\xi}^{(2)}(w, f), n^{-1/2} \|K_h * Q_{\mathfrak{h}}\|_2], \\ \check{U}_{\xi}^{(2)}(w) &:= \max[\hat{U}_{\xi}^{(2)}(w), n^{-1/2} \|K_h * Q_{\mathfrak{h}}\|_2]. \end{aligned}$$

Theorems 4 and 5 can be used in order to establish upper bounds on the norms of the processes $\xi_w^{(i)}$, $i = 1, 2$. For this purpose, Assumptions (W) and (L) should be verified for the classes of weights $\mathcal{W}^{(i)}$, $i = 1, 2$, defined above. To this end, we introduce conditions on the family of kernels \mathcal{K} that imply Assumptions (W) and (L). These conditions are rather natural and easily verifiable; they can be weakened in several ways, but we do not pursue this issue here and try to minimize cumbersome calculations to be done.

ASSUMPTION (K).

(K1) The family \mathcal{K} is a subset of the isotropic Hölder ball of functions $\mathbb{H}_d(1, L_{\mathcal{K}})$ with the exponent 1 and the Lipschitz constant $L_{\mathcal{K}}$, that is,

$$|K(x) - K(y)| \leq L_{\mathcal{K}}|x - y| \quad \forall x, y \in \mathbb{R}^d,$$

where $|\cdot|$ denotes the Euclidean distance. Moreover, any function K from \mathcal{K} is compactly supported and, without loss of generality, $\text{supp}(K) \subseteq [-1/2, 1/2]^d$ for all $K \in \mathcal{K}$.

(K2) There exist real numbers $k_1 > 0$ and $k_{\infty} < \infty$ such that

$$k_1 \leq \left| \int K(t) dt \right| \leq \|K\|_{\infty} \leq k_{\infty} \quad \forall K \in \mathcal{K}.$$

Without loss of generality, we will assume that $k_{\infty} \geq 1$ and $k_1 \leq 1$.

(K3) The set \mathcal{K} is a totally bounded set with respect to the \mathbb{L}_{∞} -norm, and there exists a real number $\beta_{\mathcal{K}} \in (0, 1)$ such that the entropy $\mathcal{E}_{\mathcal{K}}(\cdot)$ of \mathcal{K} satisfies

$$\sup_{\delta \in (0,1)} [\mathcal{E}_{\mathcal{K}}(\delta) - \delta^{-\beta_{\mathcal{K}}}] =: C_{\mathcal{K}} < \infty.$$

Several remarks on the above assumptions are in order. First, we note that Assumptions (K1) and (K3) are not completely independent. In fact, if we suppose that $\mathcal{K} \subset \mathbb{H}_d(\alpha, L_{\mathcal{K}})$ with some $\alpha > d$ then Assumption (K3) is automatically fulfilled with $\beta_{\mathcal{K}} = \alpha/d$. On the other hand, all our results remain valid if $\mathcal{K} \subset \mathbb{H}_d(\alpha, L_{\mathcal{K}})$ with some $\alpha > 0$. Observe also that the condition $|\int K(t) dt| \geq k_1$ of Assumption (K2) is not restrictive at all because for kernel estimators $\int K(t) dt = 1$. Therefore, the first inequality in (K2) is satisfied with $k_1 = 1$.

REMARK 9. It is easy to check that Assumption (K1) implies Assumption (A2) in Section 3.2 and Assumption 2 in Section 2.

Now we apply Theorems 4 and 5 to the families of random fields given by (3.33) and (3.36). We present the results for the processes $\{\xi_{\phi_1[\zeta]}, \zeta \in \mathbb{Z}^{(1)}\}$ and $\{\xi_{\phi_2[z]}, z \in \mathbb{Z}^{(2)}\}$ in a unified way.

3.4.1. *Case $s \in [1, 2]$. Uniform nonrandom bounds.* In order to derive the uniform upper bounds for $s \in [1, 2]$, we use Theorem 4. Obviously, Assumption (K) implies Assumptions (W1) and (W4). Thus, in order to apply Theorem 4, we need to verify Assumption (L). This is done in Lemma 9 given in Section 9. Thus, Theorem 4 is directly applicable, and nonasymptotic bounds can be straightforwardly derived from this theorem; one needs only to recalculate the constants appearing in the statements of the theorem.

We note that the quantity μ_* defined in (3.23) satisfies $\mu_* \leq V_{h^{\max}}$ for the set of weights $\mathcal{W}^{(1)}$ and $\mu_* \leq 2^d V_{h^{\max}}$ for the set of weights $\mathcal{W}^{(2)}$. If we assume that $V_{h^{\max}} \rightarrow 0$ as $n \rightarrow \infty$, then we can establish some asymptotic results, one of which is given in the next theorem.

THEOREM 6. *If Assumption (K) holds, then for all $s \in [1, 2]$, $\ell > 0$ and $\epsilon \in (0, 1)$*

$$\lim_{n \rightarrow \infty} n^\ell \sup_{f \in \mathcal{F}} \mathbb{E} \sup_{w \in \mathcal{W}^{(i)}} [\|\xi_w^{(i)}\|_s - (1 + \epsilon)U_\xi^{(i)}(w)]_+^q = 0, \quad i = 1, 2.$$

If Assumption (K) holds and $V_{h^{\max}} = o(1/\ln n)$ as $n \rightarrow \infty$, then for all $\ell > 0$ and $\epsilon \in (0, 1)$

$$\lim_{n \rightarrow \infty} n^\ell \sup_{f \in \mathcal{F}} \mathbb{E} \sup_{w \in \mathcal{W}^{(i)}} [\|\xi_w^{(i)}\|_2 - (1 + \epsilon)U_\xi^{(i)}(w)]_+^q = 0, \quad i = 1, 2.$$

Proof of the theorem is omitted; it is a straightforward consequence of Theorem 4 and Lemma 9 given below in Section 9.

3.4.2. *Case $s > 2$. Uniform random bounds.* In the case $s > 2$, the uniform bounds are derived from Theorem 5. To state these results, we need the following notation. Define

$$\begin{aligned} \vartheta_0^{(1)} &:= 10c_1(s) f_\infty [L_{\mathcal{K}} \sqrt{d}/k_1]^{d/2}, \\ \vartheta_0^{(2)} &:= 10c_1(s) f_\infty [2^{d+2} \sqrt{d} L_{\mathcal{K}} k_\infty / k_1^2]^{d/2}. \end{aligned} \tag{3.38}$$

The next two quantities, $A_{\mathcal{H}}$ and $B_{\mathcal{H}}$, are completely determined by the bandwidth set \mathcal{H} :

$$\begin{aligned} A_{\mathcal{H}} &:= \prod_{j=1}^d \ln(h_j^{\max} / h_j^{\min}), \\ B_{\mathcal{H}} &:= \log_2(V_{h^{\max}} / V_{h^{\min}}) = \sum_{j=1}^d \log_2(h_j^{\max} / h_j^{\min}). \end{aligned} \tag{3.39}$$

For $y > 0$ put

$$C_{\xi,i}^*(y) := 1 + 2\vartheta_0^{(i)} \{ \sqrt{y} ([2^{d(i-1)} V_{h^{\max}}]^{1/s} + n^{-1/2s}) + yn^{-1/s} \}, \tag{3.40}$$

$i = 1, 2.$

Define also

$$y_*^{(i)} := \begin{cases} \vartheta_1^{(i)} n^{4/s-1}, & s \in (2, 4), \\ \vartheta_2^{(i)} (nV_{h\min})^{-1/2} [(2^{d(i-1)}V_{h\max})^{2/s} + n^{-1/s}]^{-2}, & s \geq 4, \end{cases}$$

where explicit expressions for the constants $\vartheta_1^{(i)}, \vartheta_2^{(i)}, i = 1, 2$ are given in the proof of Theorem 7.

THEOREM 7. *Let Assumption (K) hold, $f \in \mathcal{F}$, and let $\max_{j=1,\dots,d} |h_j^{\max}| \leq 1$. For $i = 1, 2$ assume that*

$$(3.41) \quad nV_{h\min} > [64c_1^2(s)]^{(s \wedge 4)/(s \wedge 4 - 2)} [2^{d+2} \sqrt{d} L_{\mathcal{K}} k_{\infty} / k_1^2]^{d(i-1)}.$$

If $\gamma := (nV_{h\min})^{1/(s \wedge 4) - 1/2}$, then for any $s > 2, y \in [1, y_*^{(i)}]$ and for $i = 1, 2$ one has

$$\begin{aligned} \mathbb{E} \sup_{w \in \mathcal{W}^{(i)}} \{ \|\xi_w^{(i)}\|_s - \bar{u}_\epsilon(\gamma) C_{\xi,i}^*(y) \check{U}_\xi^{(i)}(w) \}_+^q \\ \leq \tilde{T}_{1,\epsilon}^{(i)} (1 + A_{\mathcal{H}})^{2i} (1 + B_{\mathcal{H}}) n^{q/2} [C_{\xi,i}^*(y)]^q e^{-y/2}, \end{aligned}$$

where $\bar{u}_\epsilon(\cdot)$ is defined in Theorem 3, and $\check{U}_\xi^{(i)}(w)$ are defined in (3.34) and (3.37).

In addition, for any subset $\mathcal{W}_0 \subseteq \mathcal{W}^{(i)}$, any $s > 2$ and for $i = 1, 2$ one has

$$\begin{aligned} \mathbb{E} \left[\sup_{w \in \mathcal{W}_0} \check{U}_\xi^{(i)}(w) \right]^q \leq [1 + 4c_1(s)(1 + \epsilon)(nV_{h\min})^{1/(s \wedge 4) - 1/2}]^q \sup_{w \in \mathcal{W}_0} \{ \bar{U}_\xi^{(i)}(w) \}^q \\ + \tilde{T}_{2,\epsilon}^{(i)} (1 + A_{\mathcal{H}})^{2i} (1 + B_{\mathcal{H}}) n^{q(s-2)/(2s)} \exp\{-y_*^{(i)}/2\}. \end{aligned}$$

The explicit expressions for the constants $\tilde{T}_{1,\epsilon}^{(i)}$ and $\tilde{T}_{2,\epsilon}^{(i)}$ are given in the proof.

We emphasize that the upper bounds of Theorem 7 are nonasymptotic. The constants $\vartheta_1^{(i)}, \vartheta_2^{(i)}, \tilde{T}_{1,\epsilon}^{(i)}$ and $\tilde{T}_{2,\epsilon}^{(i)}$ are written down explicitly in the proof of the theorem; they are completely determined through the quantities $L_{\mathcal{K}}, k_1, k_{\infty}, C_{\mathcal{K}}$ and $\beta_{\mathcal{K}}$ appearing in Assumption (K), and the constant $c_1(s)$ in the Rosenthal inequality.

REMARK 10. Condition (3.41) is not restrictive because the standard assumption on the bandwidth set \mathcal{H} in the kernel density estimation is that

$$nV_{h\min} \rightarrow \infty, \quad V_{h\max} \rightarrow 0, \quad n \rightarrow \infty.$$

The bounds established in Theorem 7 can be used in order to derive asymptotic (as $n \rightarrow \infty$) results under general assumptions on the set of bandwidths \mathcal{H} . One of such results is given in the next corollary.

COROLLARY 6. *Let $s > 2$ be fixed, Assumption (K) hold, and $f \in \mathcal{F}$. There exist positive constants $k_{1,i} = k_{1,i}(s)$, $k_{2,i} = k_{2,i}(s)$ and $k_{3,i} = k_{3,i}(s)$, $i = 1, 2$, such that if*

$$V_{h^{\max}} \asymp [\ln n]^{-k_{1,i}}, \quad nV_{h^{\min}} \asymp [\ln n]^{k_{2,i}},$$

$$\epsilon = \epsilon_n \asymp [\ln n]^{-k_{3,i}}, \quad n \rightarrow \infty,$$

then for all $\ell > 0, q \geq 1$

$$\lim_{n \rightarrow \infty} \sup_{f \in \mathcal{F}} n^\ell \mathbb{E} \sup_{w \in \mathcal{W}^{(i)}} [\|\xi_w^{(i)}\|_s - (1 + 3\epsilon_n)\check{U}_\xi^{(i)}(w)]_+^q = 0.$$

In addition, for any subset $\mathcal{W}_0 \subseteq \mathcal{W}^{(i)}$ one has

$$\mathbb{E} \left[\sup_{w \in \mathcal{W}_0} \check{U}_\xi^{(i)}(w) \right]^q \leq \left[(1 + \epsilon_n) \sup_{w \in \mathcal{W}_0} \bar{U}_\xi^{(i)}(w, f) \right]^q + R_n^{(i)}(\mathcal{W}_0),$$

where $\limsup_{n \rightarrow \infty} \sup_{f \in \mathcal{F}} \sup_{\mathcal{W}_0 \subseteq \mathcal{W}^{(i)}} [n^\ell R_n^{(i)}(\mathcal{W}_0)] = 0, i = 1, 2$.

We remark that explicit expressions for the constants $k_{1,i}$ and $k_{2,i}, i = 1, 2$, are easily derived from Theorem 7.

4. Uniform bounds for norms of regression-type processes. In this section, we use Proposition 2 in order to derive uniform bounds for the family $\|\eta_w\|_{s,\tau}, w \in \mathcal{W}$; we recall that

$$\eta_w(t) = \sum_{i=1}^n w(t, X_i) \varepsilon_i,$$

see (1.7). First, we verify Assumption 1 by establishing an analogue of Theorem 1 for a fixed weight function $w \in \mathcal{W}$ [see Theorem 8 below]. It turns out that the corresponding inequality depends heavily on the tail probability of the random variable ε . In other words, we prove that Assumption 1 is fulfilled with function g that is determined by the rate at which the tail probability of ε decreases. Next, under Assumptions (W) and (L), we derive uniform bounds using Corollary 1; this leads to an analogue of Theorem 4 for the regression-type processes.

4.1. *Probability bounds for fixed weight function.* We consider two types of moment conditions on the distribution of ε .

ASSUMPTION (E). The distribution of ε is symmetric, and one of the following two conditions is fulfilled:

(E1) there exist constants $\alpha > 0, v > 0$ and $b > 0$ such that

$$\mathbb{P}\{|\varepsilon| \geq x\} \leq v \exp\{-bx^\alpha\} \quad \forall x > 0,$$

(E2) there exist constants $p \geq [s \vee 2]$ and $P > 0$ such that

$$\mathbb{E}|\varepsilon|^p \leq P.$$

Let $\sigma_\varepsilon^2 := \mathbb{E}\varepsilon^2$ and $e_s := (\mathbb{E}|\varepsilon|^s)^{1/s}$. For any $w \in \mathcal{W}$ define

$$\varrho_s(w, f) := \begin{cases} \sigma_\varepsilon \{ \sqrt{n} \Sigma_s(w, f) \wedge 4n^{1/s} M_s(w) \}, & s < 2, \\ \sigma_\varepsilon \sqrt{n} M_2(w), & s = 2, \\ c_1(s) [\sigma_\varepsilon \sqrt{n} \Sigma_s(w, f) + 2n^{1/s} e_s M_s(w)], & s > 2, \end{cases}$$

$$\varpi_s^2(w, f) := \begin{cases} M_s^2(w) [(6\sigma_\varepsilon^2 + 8)n + 96\sigma_\varepsilon n^{1/s}], & s < 2, \\ 6\sigma_\varepsilon^2 n M_{1,\tau,\nu'}^2(w) + 24\sigma_\varepsilon \sqrt{n} M_2^2(w), & s = 2, \end{cases}$$

and if $s > 2$ then we set

$$\varpi_s^2(w, f) := 6c_3(s) [\sigma_\varepsilon^2 n M_{2s/(s+2),\tau,\nu'}^2(w) + 4\sigma_\varepsilon \sqrt{n} \Sigma_s(w, f) M_s(w) + 8e_s n^{1/s} M_s^2(w)].$$

In the above formulas, we use notation introduced in the beginning of Section 3; the formulas should be compared with (3.3) and (3.4).

The next theorem is the analogue of Theorem 1 for the regression-type processes.

THEOREM 8.

(i) Suppose that Assumption (E1) holds, and for $x > 0$ define the function

$$(4.1) \quad G_1(x) := (1 + nv) g_{\alpha,b}(x),$$

$$g_{\alpha,b}(x) := \begin{cases} \exp\{-|x| \wedge |b^{1/\alpha} x|^{\alpha/(2+\alpha)}\}, & s < 2, \\ \exp\{-|x| \wedge |b^{1/\alpha} x|^{\alpha/(1+\alpha)}\}, & s \geq 2. \end{cases}$$

Then for all $s \in [1, \infty)$ and $z > 0$ one has

$$\mathbb{P}\{\|\eta_w\|_{s,\tau} \geq \varrho_s(w, f) + z\} \leq G_1\left(\frac{z^2}{(1/3)\varpi_s^2(w, f) + (4/3)c_*(s)M_s(w)z}\right),$$

where $c_*(\cdot)$ is given in (3.2).

(ii) Suppose that Assumption (E2) holds and for $x > 0$ define the function

$$G_2(x) := (1 + nP) \times \begin{cases} (x^{-1} p \ln[1 + p^{-1}x])^{p/2}, & s < 2, \\ (x^{-1} p \ln[1 + p^{-1}x])^p, & s \geq 2. \end{cases}$$

Then for all $s \in [1, \infty)$ and $z > 0$ one has

$$\mathbb{P}\{\|\eta_w\|_{s,\tau} \geq \varrho_s(w) + z\} \leq G_2\left(\frac{z^2}{(1/3)\varpi_s^2(w, f) + (4/3)c_*(s)M_s(w)z}\right).$$

4.2. *Uniform bound.* Theorem 8 guarantees that Assumption 1 holds with function g being either G_1 or G_2 . This result is the basis for derivation of uniform bounds, and the general machinery presented in the previous sections can be fully applied here. In this section, we restrict ourselves only with uniform bounds over the classes of weights depending on the difference of arguments. In other words, under Assumptions (W), (L) and (E1) we prove an analogue of Theorem 4 for the regression-type processes.

A natural assumption in the regression model where the process $\{\eta_w, w \in \mathcal{W}\}$ appears is that the design variable X is distributed on a bounded interval of \mathbb{R}^d , that is, the density f is compactly supported. This will be assumed throughout this section.

Let $\mathcal{I} \in \mathbb{R}^d$ be a bounded interval, $\mathcal{T} = \mathcal{X} = \mathcal{I}$, and let $\tau = \nu = \text{mes}$ be the Lebesgue measure. For the sake of brevity, we write $\alpha_* = \alpha_1^{-1} \alpha_2^{-1/2}$ where α_1 and α_2 appear in Assumption (W2). Define

$$\begin{aligned} a &:= \max(\sigma_\varepsilon \sqrt{\text{mes}(\mathcal{I})}, c_1(s) [\sigma_\varepsilon f_\infty^{1/2} + 2e_s \alpha_*]), \\ c_n &:= \frac{4}{3} c_*(s) \alpha_* n^{-1/s}; \\ b_n^2 &:= \begin{cases} [2\sigma_\varepsilon^2 + \frac{8}{3} + 32\sigma_\varepsilon n^{1/s-1}] \mu_*^{2/s-1}, & s < 2, \\ 2f_\infty^2 \mu_* + 8n^{-1/2}, & s = 2, \\ 2c_3(s) f_\infty^2 [\sigma_\varepsilon^2 \mu_*^{2/s} + (4\sigma_\varepsilon \alpha_* + 8e_s \alpha_*^2) n^{-1/s}], & s > 2. \end{cases} \end{aligned}$$

THEOREM 9. *Let Assumptions (W) and (E1) hold. Suppose $f \in \mathcal{F}$, and assume that (3.26) is valid for all $s \geq 1$. Let Assumption (W4) be fulfilled with $\beta < \alpha/(2 + \alpha)$, if $s < 2$, and with $\beta < \alpha/(1 + \alpha)$ if $s \geq 2$. Then for all $s \geq 1$, $q \geq 1$ and $y > 1$ one has*

$$\begin{aligned} \mathbb{E} \sup_{w \in \mathcal{W}} [\|\eta_w\|_s - a u_\varepsilon (1 + 2\sqrt{y} b_n + 2y c_n) \sqrt{n} \|w\|_2]_+^q \\ \leq T_{n,\varepsilon} [1 + 2\sqrt{y} b_n + 2y c_n]^q [g_{\alpha,b}(y)]^{1/4}, \end{aligned}$$

where $u_\varepsilon = 2^\varepsilon (1 + \varepsilon)$, $g_{\alpha,b}(\cdot)$ is defined in (4.1), and the explicit expression of the constant $T_{n,\varepsilon}$ is given in the beginning of the proof of the theorem.

The following asymptotic result is an immediate consequence of Theorem 9.

COROLLARY 7. *Let the assumptions of Theorem 9 hold. For any $\alpha > 0$ there exist a universal constant $c = c(\alpha) > 0$ such that if $\mu_* \asymp [\ln n]^{-c}$ then for all $s \geq 1$, $\varepsilon \in (0, 1)$ and for all $\ell > 0$*

$$\lim_{n \rightarrow \infty} n^\ell \sup_{f \in \mathcal{F}} \mathbb{E} \sup_{w \in \mathcal{W}} [\|\eta_w\|_s - (1 + \varepsilon) a \sqrt{n} \|w\|_2]_+^q = 0.$$

The explicit expression for $c(\alpha)$ is easily derived from Theorem 9.

5. Proofs of Propositions 1 and 2.

5.1. *Proof of Proposition 1.* Let $Z_k, k \in \mathbb{N}$ be an $\epsilon 2^{-k-3}$ -net of Z , and let $z_k(\zeta), \zeta \in Z$ denote the element of Z_k closest to ζ in the metric d .

The continuity of the mapping $\zeta \mapsto \xi_{\phi[\zeta]}$ guarantees that P-almost surely the following relation holds for any $\zeta \in Z$:

$$(5.1) \quad \xi_{\phi[\zeta]} = \xi_{\phi[\zeta^{(0)}]} + \sum_{k=0}^{\infty} [\xi_{\phi[z_{k+1}(\zeta)]} - \xi_{\phi[z_k(\zeta)]}],$$

where $\zeta^{(0)}$ is an arbitrary fixed element of Z and $z_0(\zeta) = \zeta^{(0)}, \forall \zeta \in Z$.

Note also that independently of ζ for all $k \geq 0$

$$(5.2) \quad d(z_{k+1}(\zeta), z_k(\zeta)) \leq \epsilon 2^{-k-2}.$$

We get from sub-additivity of Ψ , (5.1) and (5.2) that for any $\zeta \in Z$

$$(5.3) \quad \begin{aligned} \Psi(\xi_{\phi[\zeta]}) &\leq \Psi(\xi_{\phi[\zeta^{(0)}]}) + \frac{\pi^2}{6} \sum_{k=0}^{\infty} p_k \Psi(\xi_{\phi[z_{k+1}(\phi)]} - \xi_{\phi[z_k(\phi)]})(k+1)^2 \\ &\leq \Psi(\xi_{\phi[\zeta^{(0)}]}) + \frac{\pi^2}{6} \sup_{k \geq 0} \sup_{\substack{(z, z') \in Z_{k+1} \times Z_k: \\ d(z, z') \leq \epsilon 2^{-k-2}}} (k+1)^2 \Psi(\xi_{\phi[z]} - \xi_{\phi[z']}), \end{aligned}$$

where $p_k := 6/(\pi^2(k+1)^2)$ and $\sum_{k=0}^{\infty} p_k = 1$. Since ξ_{\bullet} is linear, $\xi_{\phi[z]} - \xi_{\phi[z']} = \xi_{\phi[z]-\phi[z']}$ for all $z, z' \in Z$, and we obtain from (5.3) and the triangle inequality for probabilities that

$$(5.4) \quad \begin{aligned} &\mathbb{P}\left\{ \sup_{\zeta \in Z} \Psi(\xi_{\phi[\zeta]}) \geq (1 + \epsilon)[\varkappa_U(Z) + C^*(y, Z)] \right\} \\ &\leq \mathbb{P}\left\{ \Psi(\xi_{\phi[\zeta^{(0)}]}) \geq \varkappa_U(Z) + C^*(y, Z) \right\} \\ &\quad + \sum_{k=0}^{\infty} \sum_{\substack{(z, z') \in Z_{k+1} \times Z_k: \\ d(z, z') \leq \epsilon 2^{-k-2}}} \mathbb{P}\left\{ \Psi(\xi_{\phi[z]-\phi[z']}) \geq \frac{6\epsilon[\varkappa_U(Z) + C^*(y, Z)]}{\pi^2(k+1)^2} \right\} \\ &=: I_1 + I_2. \end{aligned}$$

In view of (2.1) and because $\zeta^{(0)} \in Z$, we have that $U(\phi[\zeta^{(0)}]) \leq \varkappa_U(Z)$. Therefore, we get from Assumption 1(i) and monotonicity of the function g that for any $y > 0$

$$(5.5) \quad \begin{aligned} I_1 &\leq \mathbb{P}\left\{ \Psi(\xi_{\phi[\zeta^{(0)}]}) - U(\phi[\zeta^{(0)}]) \geq C^*(y, Z) \right\} \\ &\leq g\left(\frac{[C^*(y, Z)]^2}{A^2(\phi[\zeta^{(0)}]) + B(\phi[\zeta^{(0)}])C^*(y, Z)} \right) \\ &\leq g\left(\frac{[C^*(y, Z)]^2}{\Lambda_A^2(Z) + \Lambda_B(Z)C^*(y, Z)} \right) \leq g(y). \end{aligned}$$

To order to get the last inequality, we have used monotonicity of g and that for any $y > 0$

$$\frac{[C^*(y, Z)]^2}{\Lambda_A^2(Z) + \Lambda_B(Z)C^*(y, Z)} = \frac{[\sqrt{y}\Lambda_A(Z) + y\Lambda_B(Z)]^2}{\Lambda_A^2(Z) + \Lambda_B(Z)[\sqrt{y}\Lambda_A(Z) + y\Lambda_B(Z)]} \geq y.$$

By (2.1), if $z, z' \in Z$ and $d(z, z') \leq \epsilon 2^{-k-2}$ then

$$U(\phi[z] - \phi[z']) \leq \epsilon 2^{-k-2} \varkappa_U(Z),$$

and, therefore, for any $y \geq 0$

$$\begin{aligned} & \mathbb{P}\left\{\Psi(\xi_{\phi[z]-\phi[z']}) \geq \frac{6\epsilon[\varkappa_U(Z) + C^*(y, Z)]}{\pi^2(k+1)^2}\right\} \\ & \leq \mathbb{P}\left\{\Psi(\xi_{\phi[z]-\phi[z']}) - U(\phi[z] - \phi[z']) \right. \\ & \quad \left. \geq \frac{6\epsilon[\varkappa_U(Z) + C^*(y, Z)]}{\pi^2(k+1)^2} - \varkappa_U(Z)\epsilon 2^{-k-2}\right\} \\ & \leq \mathbb{P}\left\{\Psi(\xi_{\phi[z]-\phi[z']}) - U(\phi[z] - \phi[z']) \geq \frac{9\epsilon C^*(y, Z)}{16(k+1)^2}\right\}. \end{aligned}$$

Here we took into account that $\min_{k \geq 0}[6\pi^{-2}(k+1)^{-2} - 2^{-k-2}] > 0$ and $9/16 < (6/\pi^2)$. Putting $C_k = \frac{9\epsilon C^*(y, Z)}{16(k+1)^2}$ and applying Assumption 1(i), we obtain for any $z, z' \in Z_{k+1} \times Z_k$ satisfying $d(z, z') \leq \epsilon 2^{-k-2}$:

$$\begin{aligned} & \mathbb{P}\left\{\Psi(\xi_{\phi[z]-\phi[z']}) \geq \frac{6\epsilon[\varkappa_U(Z) + C^*(y, Z)]}{\pi^2(k+1)^2}\right\} \\ & \leq g\left(\frac{C_k^2}{A^2(\phi[z] - \phi[z']) + B(\phi[z] - \phi[z'])C_k}\right) \\ & \leq g\left(\frac{C_k^2}{[\Lambda_A(Z)\epsilon 2^{-k-2}]^2 + [\Lambda_B(Z)\epsilon 2^{-k-2}]C_k}\right) \\ & \leq g\left(\frac{\tilde{C}_k^2}{\Lambda_A^2(Z) + \Lambda_B(Z)\tilde{C}_k}\right), \end{aligned}$$

where we denoted $\tilde{C}_k = C_k 2^{k+2}$. Taking into account that $9(k+1)^{-2}2^{k-2} \geq 1$ for any $k \geq 0$, and by definition of $C^*(y, Z)$, we obtain for any $y > 0$ that

$$\frac{\tilde{C}_k^2}{\Lambda_A^2(Z) + \Lambda_B(Z)\tilde{C}_k} \geq 9y(k+1)^{-2}2^{k-2}.$$

Hence, for any $z, z' \in Z_{k+1} \times Z_k$ satisfying $d(z, z') \leq \epsilon 2^{-k-2}$ one has

$$(5.6) \quad \mathbb{P} \left\{ \Psi(\xi_{\phi[z]-\phi[z']}) \geq \frac{6\epsilon[\varkappa_U(Z) + C^*(y, Z)]}{\pi^2(k+1)^2} \right\} \leq g(9y2^{k-2}(k+1)^{-2}).$$

Noting that the right-hand side of (5.6) does not depend on z, z' we get

$$(5.7) \quad I_2 \leq \sum_{k=0}^{\infty} \{N_{Z,d}(\epsilon 2^{-k-1})\}^2 g(9y2^{k-2}(k+1)^{-2}).$$

The theorem statement follows now from (5.4), (5.5) and (5.7).

5.2. Proof of Proposition 2. Let $Z_l, l = 1, \dots, N_{\mathbb{Z},d}(\epsilon/8)$ be d -balls of radius $\epsilon/8$ forming a minimal covering of the set \mathbb{Z} . For any $0 \leq j \leq [\epsilon^{-1} \log_2(R/r) - 1]_+$ [without loss of generality, we assume that $\epsilon^{-1} \log_2(R/r)$ is integer], let $\delta_j = r2^{\epsilon^j}$, and put

$$\tilde{\mathbb{Z}}_{\delta_{j+1}} = \{\zeta \in \mathbb{Z} : \delta_j < U(\phi[\zeta]) \leq \delta_{j+1}\}.$$

Note that $\tilde{\mathbb{Z}}_{\delta_j} \subseteq \mathbb{Z}_{\delta_j}$ for all j because $\epsilon \in (0, 1]$; recall that \mathbb{Z}_a is defined in (2.6).

We have $Z_l = \bigcup_{j=0}^{[\epsilon^{-1} \log_2(R/r)-1]_+} \{Z_l \cap \tilde{\mathbb{Z}}_{\delta_{j+1}}\}$ for any $l = 1, \dots, N_{\mathbb{Z},d}(\epsilon/8)$. Therefore, for any $y > 0$,

$$(5.8) \quad \Psi_{u_\epsilon}^*(y, Z_l) \leq \sup_{j=0, \dots, [\epsilon^{-1} \log_2(R/r)-1]_+} \left[\sup_{\zeta \in Z_l \cap \tilde{\mathbb{Z}}_{\delta_{j+1}}} \Psi(\xi_{\phi[\zeta]}) - u_\epsilon C^*(y) \delta_j \right].$$

Let $0 \leq j \leq [\epsilon^{-1} \log_2(R/r) - 1]_+$ be fixed; then using the definition of Λ_A and Λ_B [see (2.2) and (2.3)] and the fact that $\tilde{\mathbb{Z}}_{\delta_{j+1}} \subseteq \mathbb{Z}_{\delta_{j+1}}$ we have that

$$\begin{aligned} C^*(y) &\geq 1 + \delta_{j+1}^{-1} [2\sqrt{y} \Lambda_A(\mathbb{Z}_{\delta_{j+1}}) + 2y \Lambda_B(\mathbb{Z}_{\delta_{j+1}})] \\ &\geq 1 + \delta_j^{-1} [\sqrt{y} \Lambda_A(\mathbb{Z}_{\delta_{j+1}}) + y \Lambda_B(\mathbb{Z}_{\delta_{j+1}})] \\ &\geq 1 + \delta_j^{-1} [\sqrt{y} \Lambda_A(\tilde{\mathbb{Z}}_{\delta_{j+1}}) + y \Lambda_B(\tilde{\mathbb{Z}}_{\delta_{j+1}})]. \end{aligned}$$

Therefore

$$\begin{aligned} C^*(y) \delta_j &\geq \delta_j + [\sqrt{y} \Lambda_A(\tilde{\mathbb{Z}}_{\delta_{j+1}}) + y \Lambda_B(\tilde{\mathbb{Z}}_{\delta_{j+1}})] \\ &\geq 2^{-\epsilon} \varkappa_U(\tilde{\mathbb{Z}}_{\delta_{j+1}}) + C^*(y, \tilde{\mathbb{Z}}_{\delta_{j+1}}), \end{aligned}$$

since by the premise of the proposition $\delta_j = 2^{-\epsilon} \delta_{j+1} \geq 2^{-\epsilon} \varkappa_U(\mathbb{Z}_{\delta_{j+1}}) \geq 2^{-\epsilon} \times \varkappa_U(\tilde{\mathbb{Z}}_{\delta_{j+1}})$. Note also that the definition of $C^*(\cdot, \cdot)$ implies that $C^*(\cdot, Z_1) \leq C^*(\cdot, Z_2)$ whenever $Z_1 \subseteq Z_2$. Thus, we have for any $0 \leq j \leq [\epsilon^{-1} \log_2(R/r) - 1]_+$ and any $l = 1, \dots, N_{\mathbb{Z},d}(\epsilon/8)$

$$(5.9) \quad u_\epsilon C^*(y) \delta_j \geq (1 + \epsilon) [\varkappa_U(Z_l \cap \tilde{\mathbb{Z}}_{\delta_{j+1}}) + C^*(y, Z_l \cap \tilde{\mathbb{Z}}_{\delta_{j+1}})].$$

Taking into account (5.8), we obtain

$$\begin{aligned} & \mathbb{P}\{\Psi_{u_\epsilon}^*(y, Z_l) \geq 0\} \\ & \leq \sum_{j=0}^{[\epsilon^{-1} \log_2(R/r)-1]_+} \mathbb{P}\left\{ \sup_{\zeta \in Z_l \cap \tilde{\mathbb{Z}}_{\delta_{j+1}}} \Psi(\xi_{\phi[\zeta]}) \geq (1 + \epsilon)[\mathcal{X}_U(Z_l \cap \tilde{\mathbb{Z}}_{\delta_{j+1}})] \right. \\ & \qquad \qquad \qquad \left. + C^*(y, Z_l \cap \tilde{\mathbb{Z}}_{\delta_{j+1}}) \right\}. \end{aligned}$$

Applying Proposition 1 for the sets $Z_l \cap \tilde{\mathbb{Z}}_{\delta_{j+1}}$, we get for any $y > 0$

$$\begin{aligned} \mathbb{P}\{\Psi_{u_\epsilon}^*(y, Z_l) \geq 0\} & \leq \sum_{j=0}^{[\epsilon^{-1} \log_2(R/r)-1]_+} L_{g, \mathbb{Z}_{\delta_{j+1}}}^{(\epsilon)}(y) \\ & = \sum_{j=0}^{[\epsilon^{-1} \log_2(R/r)-1]_+} L_g^{(\epsilon)}(y, r2^{\epsilon(j+1)}). \end{aligned}$$

It remains to note that the right-hand side of the last inequality does not depend on l ; thus, we come to the first assertion of the proposition.

Now we derive the bound for the moments of $\Psi_{u_\epsilon}^*(y, \mathbb{Z})$. We have from (5.8) with $y > 0$ that for any $q \geq 1$

$$\begin{aligned} & \mathbb{E}\left(\sup_{\zeta \in \mathbb{Z}} \{\Psi(\xi_{\phi[\zeta]}) - u_\epsilon C^*(y)U(\phi[\zeta])\}_+^q\right) \\ (5.10) \quad & \leq \sum_{l=1}^{N_{\mathbb{Z},d}(\epsilon/8)} \sum_{j=0}^{[\epsilon^{-1} \log_2(R/r)-1]_+} \mathbb{E}\left(\sup_{\zeta \in Z_l \cap \tilde{\mathbb{Z}}_{\delta_{j+1}}} \{\Psi(\xi_{\phi[\zeta]}) - u_\epsilon C^*(y)\delta_j\}_+^q\right) \\ & =: \sum_{l=1}^{N_{\mathbb{Z},d}(\epsilon/8)} \sum_{j=0}^{[\epsilon^{-1} \log_2(R/r)-1]_+} E_j(l). \end{aligned}$$

For $l = 1, \dots, N_{\mathbb{Z},d}(\epsilon/8)$ and $0 \leq j \leq [\epsilon^{-1} \log_2(R/r) - 1]_+$ we have

$$\begin{aligned} E_j(l) & = q \int_{u_\epsilon C^*(y)\delta_j}^\infty [x - u_\epsilon C^*(y)\delta_j]^{q-1} \\ & \qquad \qquad \qquad \times \mathbb{P}\left\{ \sup_{\zeta \in Z_l \cap \tilde{\mathbb{Z}}_{\delta_{j+1}}} \Psi(\xi_{\phi[\zeta]}) \geq x \right\} dx \\ & = [u_\epsilon C^*(y)]^q \delta_j^q q \\ (5.11) \quad & \times \int_1^\infty (z - 1)^{q-1} \mathbb{P}\left\{ \sup_{\zeta \in Z_l \cap \tilde{\mathbb{Z}}_{\delta_{j+1}}} \Psi(\xi_{\phi[\zeta]}) \geq zu_\epsilon C^*(y)\delta_j \right\} dz \end{aligned}$$

$$\begin{aligned} &\leq [u_\epsilon C^*(y)]^q \delta_j^q q \\ &\quad \times \int_1^\infty (z-1)^{q-1} \mathbb{P}\left\{ \sup_{\zeta \in Z_l \cap \tilde{\mathbb{Z}}_{\delta_{j+1}}} \Psi(\xi_{\phi[\zeta]}) \geq u_\epsilon C^*(yz) \delta_j \right\} dz \\ &\leq [u_\epsilon C^*(y)]^q \delta_j^q q \int_1^\infty (z-1)^{q-1} L_g^{(\epsilon)}(yz, r2^{\epsilon(j+1)}) dz. \end{aligned}$$

Here the third line follows from $zC^*(y) \geq C^*(yz)$ for any $z \geq 1$, and the last line is a consequence of (5.9) and the probability bound established above.

The second statement of the theorem follows now from (5.10) and (5.11) since the right-hand side in (5.11) does not depend on l .

6. Proof of Theorem 1.

6.1. *Preliminaries.* For convenience in this section, we present some well-known results that will be repeatedly used in the proofs.

Empirical processes. Let \mathcal{F} be a countable set of functions $f : \mathcal{X} \rightarrow \mathbb{R}$. Suppose that $\mathbb{E}f(X) = 0, \|f\|_\infty \leq b, \forall f \in \mathcal{F}$ and put

$$Y = \sup_{f \in \mathcal{F}} \sum_{i=1}^n f(X_i), \quad \sigma^2 = \sup_{f \in \mathcal{F}} \mathbb{E}[f(X)]^2.$$

LEMMA 1. *For any $x \geq 0$*

$$\mathbb{P}\{Y - \mathbb{E}Y \geq x\} \leq \exp\left\{-\frac{x^2}{2n\sigma^2 + 4b\mathbb{E}Y + (2/3)bx}\right\}.$$

The statement of the lemma is an immediate consequence of the the Bennett inequality for empirical processes [see Bousquet (2002)] and the standard arguments allowing to derive the Bernstein inequality from the Bennett inequality.

Inequalities for sums of independent random variables. We recall the well-known Rosenthal and Bahr–Esseen [see von Bahr and Esseen (1965)] bounds on the moments of sums of independent random variables.

LEMMA 2. *Let Y_1, \dots, Y_n be independent random variables, $\mathbb{E}Y_i = 0, i = 1, \dots, n$. Then*

$$\begin{aligned} \mathbb{E}\left|\sum_{i=1}^n Y_i\right|^p &\leq [c_1(p)]^p \left\{ \sum_{i=1}^n \mathbb{E}|Y_i|^p + \left(\sum_{i=1}^n \mathbb{E}Y_i^2\right)^{p/2} \right\}, \quad p > 2; \\ \mathbb{E}\left|\sum_{i=1}^n Y_i\right|^p &\leq 2 \sum_{i=1}^n \mathbb{E}|Y_i|^p, \quad p \in [1, 2), \end{aligned}$$

where $c_1(p) = 15p/\ln p$.

The constant $c_1(p) = 15p/\ln p$ in the Rosenthal inequality is obtained by symmetrization of the inequality of Theorem 4.1 in Johnson, Schechtman and Zinn (1985).

Norms of integral operators. The next statement presents inequalities for norms of integral operators.

LEMMA 3. Let $(\mathcal{T}, \mathfrak{T}, \tau)$ and $(\mathcal{X}, \mathfrak{X}, \chi)$ be σ -finite spaces, w be a $(\mathfrak{T} \times \mathfrak{X})$ -measurable function on $\mathcal{T} \times \mathcal{X}$, and let

$$M_{p,\tau,\chi}(w) := \sup_{x \in \mathcal{X}} \|w(\cdot, x)\|_{p,\tau} \vee \sup_{t \in \mathcal{T}} \|w(t, \cdot)\|_{p,\chi}.$$

If $R \in \mathbb{L}_p(\mathcal{X}, \chi)$ and $\mathcal{I}_R(t) := \int w(t, x)R(x)\chi(dx)$ then the following statements hold:

(a) For any $p \in [1, \infty]$

$$(6.1) \quad \|\mathcal{I}_R\|_{p,\tau} \leq M_{1,\tau,\chi}(w) \|R\|_{p,\chi}.$$

(b) For any $1 < p < r < \infty$

$$(6.2) \quad \|\mathcal{I}_R\|_{r,\tau} \leq c_2(p) M_{q,\tau,\chi}(w) \|R\|_{p,\chi},$$

where $\frac{1}{q} = 1 + \frac{1}{r} - \frac{1}{p}$, and $c_2(p)$ is a numerical constant independent of w .

The statements of the lemma can be found in Folland (1999), Theorems 6.18 and 6.36.

Note that if $\chi = \nu' := f\nu$ then $M_{p,\tau,\chi}(w) = M_p(w)$, $\forall w$ [see (3.1)]. If $\mathcal{T} = \mathcal{X} = \mathbb{R}^d$, τ and χ are the Lebesgue measures, and if $w(t, x)$ depends on the difference $t - x$ only, then $c_2(p) = 1$, and (6.2) is the well-known Young inequality.

6.2. *Proof of Theorem 1.* We begin with two technical lemmas; their proofs are given in the Appendix.

LEMMA 4. Let $\mathbb{B}_{s/(s-1)}$ be the unit ball in $\mathbb{L}_{s/(s-1)}(\mathcal{T}, \tau)$, and suppose that Assumption (A1) hold. Then, there exists a countable set $\mathfrak{L} \subset \mathbb{B}_{s/(s-1)}$ such that

$$\|\xi_w\|_{s,\tau} = \sup_{l \in \mathfrak{L}} \int l(t)\xi_w(t)\tau(dt).$$

LEMMA 5. Let $\bar{w}(t, x) = w(t, x) - \mathbb{E}w(t, X)$; then for all $p \geq 1$ one has:

- (a) $\|\bar{w}(\cdot, x)\|_{p,\tau} \leq 2 \sup_{x \in \mathcal{X}} \|w(\cdot, x)\|_{p,\tau}$.
- (b) $M_p(\bar{w}) \leq 2M_p(w)$.

We break the proof of Theorem 1 into several steps.

Step 1: Reduction to empirical process. We obtain from Lemma 4

$$\begin{aligned} \|\xi_w\|_{s,\tau} &= \sup_{l \in \mathcal{L}} \int l(t)\xi_w(t)\tau(dt) \\ &= \sup_{l \in \mathcal{L}} \sum_{i=1}^n \int l(t)\bar{w}(t, X_i)\tau(dt) \\ &= \sup_{\lambda \in \Lambda} \sum_{i=1}^n \lambda(X_i), \end{aligned}$$

where

$$\Lambda = \left\{ \lambda : \mathcal{X} \rightarrow \mathbb{R} : \lambda(x) = \int l(t)\bar{w}(t, x)\tau(dt), l \in \mathcal{L} \right\}.$$

Thus,

$$(6.3) \quad \|\xi_w\|_{s,\tau} = \sup_{\lambda \in \Lambda} \sum_{i=1}^n \lambda(X_i) =: Y$$

and, obviously, $\mathbb{E}\lambda(X) = 0$. The idea now is to apply Lemma 1 to the random variable Y .

Step 2: Some upper bounds. In order to apply Lemma 1, we need to bound from above the following quantities: (i) $\mathbb{E}Y$; (ii) $b := \sup_{\lambda \in \Lambda} \|\lambda\|_\infty$; and (iii) $\sigma^2 := \sup_{\lambda \in \Lambda} \mathbb{E}\lambda^2(X)$.

(i) *Upper bound for $\mathbb{E}Y$.* Applying the Hölder inequality, we get from (6.3)

$$\mathbb{E} \left[\sup_{\lambda \in \Lambda} \sum_{i=1}^n \lambda(X_i) \right] = \mathbb{E} \|\xi_w\|_{s,\tau} \leq [\mathbb{E} \|\xi_w\|_{s,\tau}^s]^{1/s} = \left[\int \mathbb{E} |\xi_w(t)|^s \tau(dt) \right]^{1/s}.$$

If $s \in [1, 2]$, then for all $t \in \mathcal{T}$

$$\mathbb{E} |\xi_w(t)|^s \leq [\mathbb{E} |\xi_w(t)|^2]^{s/2} \leq [n \mathbb{E} w^2(t, X)]^{s/2} = \left[n \int w^2(t, x) f(x) \nu(dx) \right]^{s/2}.$$

Thus, we have for all $s \in [1, 2]$

$$(6.4) \quad \mathbb{E}Y = \mathbb{E} \left[\sup_{\lambda \in \Lambda} \sum_{i=1}^n \lambda(X_i) \right] \leq \sqrt{n} \Sigma_s(w, f).$$

Note that the same quantity can be bounded from above in a different way. Indeed, in view of the Barh–Esseen inequality (the second statement of Lemma 6.8)

$$\mathbb{E} |\xi_w(t)|^s \leq 2n \mathbb{E} |\bar{w}(t, X)|^s = 2^{1+s} n \mathbb{E} |w(t, X)|^s$$

and we obtain for all $s \in [1, 2]$

$$(6.5) \quad \mathbb{E}Y = \mathbb{E} \left[\sup_{\lambda \in \Lambda} \sum_{i=1}^n \lambda(X_i) \right] \leq 2^{1+1/s} n^{1/s} M_s(w).$$

We get finally from (6.4) and (6.5)

$$(6.6) \quad \mathbb{E}Y \leq \{\sqrt{n}\Sigma_s(w, f)\} \wedge \{4n^{1/s}M_s(w)\}.$$

If $s = 2$, we obtain a bound independent of f : indeed, in this case

$$(6.7) \quad \begin{aligned} \mathbb{E}Y &= \mathbb{E}\left[\sup_{\lambda \in \Lambda} \sum_{i=1}^n \lambda(X_i)\right] \leq \sqrt{n}\left[\iint w^2(t, x)f(x)v(dx)\tau(dt)\right]^{1/2} \\ &\leq \sqrt{n}M_2(w). \end{aligned}$$

If $s > 2$, then applying the Rosenthal inequality (the first assertion of Lemma 6.8) to $\xi_w(t)$, which is a sum of i.i.d. random variables for any $t \in \mathcal{T}$, we get

$$[\mathbb{E}(|\xi_w(t)|^s)]^{1/s} \leq c_1(s)[(n\mathbb{E}w^2(t, X))^{s/2} + n\mathbb{E}|\bar{w}(t, X)|^s]^{1/s}$$

and, therefore,

$$(6.8) \quad \begin{aligned} &\mathbb{E}\left[\sup_{\lambda \in \Lambda} \sum_{i=1}^n \lambda(X_i)\right] \\ &\leq c_1(s)\left\{\sqrt{n}\left[\iint \left(\int w^2(t, x)f(x)v(dx)\right)^{s/2} \tau(dt)\right]^{1/s}\right. \\ &\quad \left.+ 2n^{1/s}\left[\iint |w(t, x)|^s f(x)v(dx)\tau(dt)\right]^{1/s}\right\}. \end{aligned}$$

To get the last inequality we have used that $\mathbb{E}|\bar{w}(t, X)|^s \leq 2^s\mathbb{E}|w(t, X)|^s$, for all $s \geq 1$.

It is evident that the second integral on the right-hand side of (6.8) does not exceed $M_s(w)$. Moreover, since $(\mathbb{E}w^2(t, X))^{s/2} \leq \mathbb{E}|w(t, X)|^s$, $s \geq 2$, the following bound is true $\Sigma_s(w, f) \leq M_s(w)$. We conclude that $\mathbb{E}Y < \infty$ whenever $M_s(w) < \infty$, and

$$(6.9) \quad \mathbb{E}Y = \mathbb{E}\left[\sup_{\lambda \in \Lambda} \sum_{i=1}^n \lambda(X_i)\right] \leq c_1(s)\{\sqrt{n}\Sigma_s(w, f) + 2n^{1/s}M_s(w)\}.$$

(ii) *Upper bound for $b = \sup_{\lambda \in \Lambda} \|\lambda\|_\infty$.* Taking into account that $l \in \mathcal{L} \subset \mathbb{B}_{s/(s-1)}$ (Lemma 4) and applying the Hölder inequality, we get for any $x \in \mathcal{X}$

$$|\lambda(x)| \leq \left[\int |w(t, x) - \mathbb{E}w(t, X)|^s \tau(dt)\right]^{1/s} = \|\bar{w}(\cdot, x)\|_{s, \tau}.$$

Therefore, in view of Lemma 5(a)

$$(6.10) \quad b = \|\lambda\|_\infty \leq 2 \sup_{x \in \mathcal{X}} \|w(\cdot, x)\|_{s, \tau} \leq 2M_s(w).$$

(iii) Upper bound on the “dual” variance σ^2 . Since $\mathbb{E}\lambda(X) = 0$, we have

$$\begin{aligned} \sigma^2 &= \sup_{\lambda \in \Lambda} \int \lambda^2(x) f(x) \nu(dx) \\ &= \sup_{l \in \mathcal{L}} \int \left[\int \bar{w}(t, x) l(t) \tau(dt) \right]^2 f(x) \nu(dx) \\ &\leq \sup_{l \in \mathbb{B}_{s/(s-1)}} \int \left[\int \bar{w}(t, x) l(t) \tau(dt) \right]^2 f(x) \nu(dx) \\ &\leq \sup_{l \in \mathbb{B}_{s/(s-1)}} \int \left[\int w(t, x) l(t) \tau(dt) \right]^2 f(x) \nu(dx). \end{aligned}$$

The expression on the right-hand side is bounded differently depending on the value of s .

If $s \in [1, 2)$, then applying the Hölder inequality to the inner integral in the previous expression we obtain

$$\begin{aligned} \sigma^2 &\leq \int \left[\int |w(t, x)|^s \tau(dt) \right]^{2/s} f(x) \nu(dx) \\ (6.11) \quad &\leq \sup_{x \in \mathcal{X}} \|w(\cdot, x)\|_{s, \tau}^2 \leq M_s^2(w). \end{aligned}$$

We remark also that the bound given by (6.11) remains true for all $s \geq 1$. This shows, in particular, that σ is always bounded whenever $M_s(w) < \infty$.

If $s = 2$, then we apply inequality (6.1) of Lemma 3 with $p = 2$ and $\chi(dx) = \nu'(dx) = f(x)\nu(dx)$ to the integral operator $\mathcal{I}_l(x) = \int w(t, x)l(t)\tau(dt)$. This leads to the following bound

$$(6.12) \quad \sigma^2 \leq M_{1, \tau, \nu'}^2(w).$$

If $s > 2$, then we apply inequality (6.2) of Lemma 3 with $r = 2$, $p = \frac{s}{s-1}$, $q = \frac{2s}{s+2}$ and $\chi = \nu'$ to the integral operator $\mathcal{I}_l(x) = \int w(t, x)l(t)\tau(dt)$. This yields

$$(6.13) \quad \sigma^2 \leq c_2(s/(s-1))M_{q, \tau, \nu'}^2(w) = c_2(s/(s-1))M_{2s/(s+2), \tau, \nu'}^2(w).$$

Step 3: Application of Lemma 1. 1. Case $s \in [1, 2)$. Here we have from (6.6), (6.10) and (6.11)

$$\begin{aligned} \mathbb{E}Y &\leq \{\sqrt{n}\Sigma_s(w, f)\} \wedge \{4n^{1/s}M_s(w)\} =: \rho_s(w, f), \\ b &\leq 2M_s(w), \quad \sigma^2 \leq M_s^2(w). \end{aligned}$$

Therefore applying Lemma 1, we have for all $z > 0$

$$\begin{aligned} (6.14) \quad &\mathbb{P}\{\|\xi_w\|_{s, \tau} \geq \rho_s(w, f) + z\} \\ &\leq \exp\left\{-\frac{z^2}{2M_s^2(w)[n + 16n^{1/s}] + [4M_s(w)z/3]}\right\}, \end{aligned}$$

where we have used (6.5) in the denominator of the expression inside of the exponent.

To get the result of the theorem, we note that the following trivial upper bound follows from the triangle inequality and the statement (a) of Lemma 5:

$$\|\xi_w\|_{s,\tau} \leq 2nM_s(w) \quad \forall s \geq 1.$$

Thus, the probability in (6.14) is equal to zero if $z > 2nM_s(w)$; hence, we can replace z by $2nM_s(w)$ in the denominator of the expression on the right-hand side. This leads to the statement of the theorem for $s \in [1, 2)$.

2. *Case $s = 2$.* We have from (6.7), (6.10) and (6.12)

$$\mathbb{E}Y \leq \sqrt{n}M_2(w), \quad b \leq 2M_2(w), \quad \sigma^2 \leq M_{1,\tau,v'}^2(w).$$

Thus, for all $z > 0$

$$\begin{aligned} &\mathbb{P}\{\|\xi_w\|_{2,\tau} \geq \sqrt{n}M_2(w) + z\} \\ &\leq \exp\left\{-\frac{z^2}{2[nM_{1,\tau,v'}^2(w) + 4\sqrt{n}M_2^2(w) + (2/3)M_2(w)z]}\right\}, \end{aligned}$$

and the statement of Theorem 1 is established for $s = 2$.

3. *Case $s > 2$.* We have from (6.9), (6.10) and (6.13)

$$\begin{aligned} \mathbb{E}Y &\leq c_1(s)[\sqrt{n}\Sigma_s(w, f) + 2n^{1/s}M_s(w)], \\ b &\leq 2M_s(w); \quad \sigma^2 \leq c_2(s/(s-1))M_{2s/(s+2),\tau,v'}^2(w). \end{aligned}$$

Thus, for any $z > 0$ we get

$$\begin{aligned} &\mathbb{P}\{\|\xi_w\|_{s,\tau} \geq c_1(s)[\sqrt{n}\Sigma_s(w, f) + 2n^{1/s}M_s(w)] + z\} \\ &\leq \exp\{-z^2(2c_3(s)[nM_{2s/(s+2),\tau,v'}^2(w) + 4\sqrt{n}\Sigma_s(w, f)M_s(w) \\ &\quad + 8n^{1/s}M_s^2(w) + \frac{2}{3}M_s(w)z])^{-1}\}, \end{aligned}$$

where $c_3(s)$ is given in (3.2). This completes the proof of the theorem for the case of $s > 2$.

We conclude by establishing the inequalities in (3.6). In order to derive the first inequality, we apply (6.1) of Lemma 3 with $p = s/2 > 1$, $\chi = \nu$ to the integral operator $\mathcal{I}_f(t) := \int w^2(t, x) f(x) \nu(dx)$. This yields

$$\left[\int \left(\int w^2(t, x) f(x) \nu(dx) \right)^{s/2} \tau(dt) \right]^{1/s} \leq M_2(w) \|\sqrt{f}\|_{s,\nu},$$

as claimed. The second inequality in (3.6) follows straightforwardly from the definition of $M_{p,\tau,v'}$ and M_p .

7. Proofs of Theorem 3 and Corollary 4.

7.1. *Proof of Theorem 3.* First, we specify the constants appearing in the statement of the theorem:

$$T_{1,\epsilon} := \left(\frac{2^{q(\epsilon+1)}}{2^{q\epsilon} - 1} \Gamma(q + 1) + 1 \right) N_{\mathbb{Z},d}(\epsilon/8) (2u_\epsilon R_\xi)^q [1 \vee \log_2(R_\xi/r_\xi)] [1 + L_{\text{exp}}^{(\epsilon)}],$$

$$T_{2,\epsilon} := [c_1(s) + 2]^q N_{\mathbb{Z},d}(\epsilon/8) [1 \vee \log_2(R_\xi/r_\xi)] [1 + L_{\text{exp}}^{(\epsilon)}].$$

Recall that in view of (3.7), any $w \in \mathcal{W}$ is represented as $w = \phi[\zeta]$ for some $\zeta \in \mathbb{Z}$. For every $0 \leq j \leq \lceil \log_2(R_\xi/r_\xi) - 1 \rceil_+$ [without loss of generality, we assume that $\log_2(R_\xi/r_\xi)$ is an integer number], put $\delta_j = r2^{j+1}$, and define the random events

$$\mathcal{A} := \bigcap_{j=0}^{\lceil \log_2(R_\xi/r_\xi) - 1 \rceil_+} \mathcal{A}_j, \quad \mathcal{A}_j := \left\{ \sup_{\zeta \in \mathbb{Z}_{\delta_j}} \|\xi_{\phi^2[\zeta]}\|_{s/2,\tau} \leq [2(1 + \epsilon)\gamma\delta_j]^2 \right\}.$$

(i) The following trivial inequality holds:

$$\begin{aligned} & \sup_{\zeta \in \mathbb{Z}} \{ \|\xi_{\phi[\zeta]}\|_{s,\tau} - \bar{u}_\epsilon(\gamma) C_\xi^*(y) \hat{U}_\xi(\phi[\zeta]) \}_+ \\ & \leq \sup_{\zeta \in \mathbb{Z}} \{ \|\xi_{\phi[\zeta]}\|_{s,\tau} - u_\epsilon C_\xi^*(y) U_\xi(\phi[\zeta], f) \}_+ \\ & \quad + u_\epsilon C_\xi^*(y) \sup_{\zeta \in \mathbb{Z}} U_\xi(\phi[\zeta], f). \end{aligned}$$

Therefore,

$$\begin{aligned} & \mathbb{E} \sup_{\zeta \in \mathbb{Z}} \{ \|\xi_{\phi[\zeta]}\|_{s,\tau} - \bar{u}_\epsilon(\gamma) C_\xi^*(y) \hat{U}_\xi(\phi[\zeta]) \}_+^q \\ (7.1) \quad & \leq \mathbb{E} \left[\sup_{\zeta \in \mathbb{Z}} \{ \|\xi_{\phi[\zeta]}\|_{s,\tau} - \bar{u}_\epsilon(\gamma) C_\xi^*(y) \hat{U}_\xi(\phi[\zeta]) \}_+^q \mathbf{1}(\mathcal{A}) \right] \\ & \quad + 2^{q-1} \mathbb{E} \sup_{\zeta \in \mathbb{Z}} \{ \|\xi_{\phi[\zeta]}\|_{s,\tau} - u_\epsilon C_\xi^*(y) U_\xi(\phi[\zeta], f) \}_+^q \\ & \quad + 2^{q-1} [u_\epsilon C_\xi^*(y) R_\xi]^q \sum_{j=0}^{\lceil \log_2(R_\xi/r_\xi) - 1 \rceil_+} \mathbb{P}\{\bar{\mathcal{A}}_j\}, \end{aligned}$$

where $\bar{\mathcal{A}}_j$ denotes the event complementary to \mathcal{A}_j , and $\mathbf{1}(\mathcal{A})$ is the indicator of the event \mathcal{A} . The second term on the right-hand side is bounded using Theorem 2; our current goal is to bound the first and the third terms.

Note that, if the event \mathcal{A} occurs then for every $\zeta \in \mathbb{Z}$

$$\begin{aligned} (7.2) \quad & U_\xi(\phi[\zeta], f) [1 + 4c_1(s)(1 + \epsilon)\gamma] \\ & \geq \hat{U}_\xi(\phi[\zeta]) \geq U_\xi(\phi[\zeta], f) [1 - 4c_1(s)(1 + \epsilon)\gamma]. \end{aligned}$$

Indeed, in view of (3.9), (3.13) and (3.14) we get

$$\begin{aligned}
 \hat{U}_\xi(\phi[\zeta]) &\geq U_\xi(\phi[\zeta], f) - |\hat{U}_\xi(\phi[\zeta]) - U_\xi(\phi[\zeta], f)| \\
 (7.3) \qquad &= U_\xi(\phi[\zeta], f) - c_1(s)\sqrt{n}|\hat{\Sigma}_s(\phi[\zeta]) - \Sigma_s(\phi[\zeta], f)| \\
 &\geq U_\xi(\phi[\zeta], f) - c_1(s)\sqrt{\|\xi\phi^2[\zeta]\|_{s/2,\tau}}.
 \end{aligned}$$

Let $\zeta \in \mathbb{Z}$ be fixed. Since \mathbb{Z}_{δ_j} , $j = 0, \dots, \lceil \log_2(R_\xi/r_\xi) - 1 \rceil_+$, defined in (3.15), form the partition of \mathbb{Z} , there exists j_* such that $\zeta \in \mathbb{Z}_{\delta_{j_*}}$. Because $\zeta \in \mathbb{Z}_{\delta_{j_*}}$ implies $U_\xi(\phi[\zeta], f) \geq \delta_{j_*}/2 = \delta_{j_*-1}$, we obtain from (7.3) on the event \mathcal{A} that

$$\begin{aligned}
 (7.4) \qquad \hat{U}_\xi(\phi[\zeta]) &\geq U_\xi(\phi[\zeta], f) - 2c_1(s)(1 + \epsilon)\gamma\delta_{j_*} \\
 &\geq U_\xi(\phi[\zeta], f)[1 - 4c_1(s)(1 + \epsilon)\gamma].
 \end{aligned}$$

Thus, the right-hand side inequality in (7.2) is proved. Similarly, we have from (7.3) and (7.4) that

$$\begin{aligned}
 \hat{U}_\xi(\phi[\zeta]) &\leq U_\xi(\phi[\zeta], f) + |\hat{U}_\xi(\phi[\zeta]) - U_\xi(\phi[\zeta], f)| \\
 &\leq U(\phi[\zeta]) + c_1(s)\sqrt{\|\xi\phi^2[\zeta]\|_{s/2,\tau}} \\
 &\leq U_\xi(\phi[\zeta], f)[1 + 4c_1(s)(1 + \epsilon)\gamma].
 \end{aligned}$$

Thus, (7.2) is proved.

Using the right-hand side inequality in (7.2) and applying Theorem 2, we obtain

$$\begin{aligned}
 (7.5) \qquad &\mathbb{E}\left[\sup_{\zeta \in \mathbb{Z}}\{\|\xi\phi[\zeta]\|_{s,\tau} - \bar{u}_\epsilon(\gamma)C_\xi^*(y)\hat{U}_\xi(\phi[\zeta])\}_+^q \mathbf{1}(\mathcal{A})\right] \\
 &\leq \mathbb{E}\sup_{\zeta \in \mathbb{Z}}\{\|\xi\phi[\zeta]\|_{s,\tau} - u_\epsilon C_\xi^*(y)U_\xi(\phi[\zeta], f)\}_+^q \\
 &\leq \frac{2^{q(\epsilon+1)}u_\epsilon^q}{2^{q\epsilon} - 1}\Gamma(q + 1)N_{\mathbb{Z},d}(\epsilon/8)[R_\xi C_\xi^*(1)]^q[1 + L_{\text{exp}}^{(\epsilon)}]\exp\{-y/2\}.
 \end{aligned}$$

Now we bound the probability $\mathbb{P}\{\bar{\mathcal{A}}_j\}$. Let Z_l , $l = 1, \dots, N_{\mathbb{Z},d}(\epsilon/8)$ be a minimal covering of \mathbb{Z} by balls of radius $\epsilon/8$ in the metric d . By definition of \mathcal{A}_j , we have

$$(7.6) \qquad \mathbb{P}\{\bar{\mathcal{A}}_j\} \leq \sum_{l=1}^{N_{\mathbb{Z},d}(\epsilon/8)} \mathbb{P}\left\{\sup_{\zeta \in Z_l \cap \mathbb{Z}_{\delta_j}} \|\xi\phi^2[\zeta]\|_{s/2,\tau} \geq [2(1 + \epsilon)\gamma\delta_j]^2\right\}.$$

Note that

$$[2\gamma\delta_j]^2 \geq \varkappa_{\bar{U}}(\mathbb{Z}_{\delta_j}) + \delta_j^2[\sqrt{y_\gamma}\lambda_{\bar{A}} + y_\gamma\lambda_{\bar{B}}] \geq \varkappa_{\bar{U}}(Z_l \cap \mathbb{Z}_{\delta_j}) + \delta_j^2[\sqrt{y}\lambda_{\bar{A}} + y\lambda_{\bar{B}}];$$

here the first inequality follows from the condition $\varkappa_{\bar{U}}(\mathbb{Z}_a) = \varkappa_{\bar{U}}(a) \leq (\gamma a)^2$, $\forall a \in [r_\xi, R_\xi]$ and from definition of y_γ ; the second inequality holds by the inclusion $Z_l \cap \mathbb{Z}_{\delta_j} \subseteq \mathbb{Z}_{\delta_j}$ and because $y \leq y_\gamma$. Furthermore, by (3.16) and by the above

inclusion

$$\begin{aligned} \lambda_{\tilde{A}} &\geq \delta_j^{-2} \Lambda_{\tilde{A}}(\mathbb{Z}_{\delta_j}) \geq \delta_j^{-2} \Lambda_{\tilde{A}}(\mathbb{Z}_{\delta_j} \cap Z_l), \\ \lambda_{\tilde{B}} &\geq \delta_j^{-2} \Lambda_{\tilde{B}}(\mathbb{Z}_{\delta_j}) \geq \delta_j^{-2} \Lambda_{\tilde{B}}(\mathbb{Z}_{\delta_j} \cap Z_l), \end{aligned}$$

which leads to

$$\begin{aligned} [2\gamma\delta_j]^2 &\geq \varkappa_{\tilde{U}}(Z_l \cap \mathbb{Z}_{\delta_j}) + \sqrt{y} \Lambda_{\tilde{A}}(\mathbb{Z}_{\delta_j} \cap Z_l) + y \Lambda_{\tilde{B}}(\mathbb{Z}_{\delta_j} \cap Z_l) \\ &= \varkappa_{\tilde{U}}(Z_l \cap \mathbb{Z}_{\delta_j}) + \tilde{C}_*(y, Z_l \cap \mathbb{Z}_{\delta_j}), \end{aligned}$$

where $\tilde{C}_*(y, \cdot) := \sqrt{y} \Lambda_{\tilde{A}}(\cdot) + y \Lambda_{\tilde{B}}(\cdot)$ [cf. (2.4)].

Hence, applying Proposition 1, we obtain from (7.6) that

$$\begin{aligned} \mathbb{P}\{\bar{\mathcal{A}}_j\} &\leq \sum_{l=1}^{N_{\mathbb{Z},d}(\epsilon/8)} \mathbb{P}\left\{ \sup_{\zeta \in Z_l \cap \mathbb{Z}_{\delta_j}} \|\xi_{\phi^2[\zeta]}\|_{s/2, \tau} \right. \\ &\quad \left. \geq (1 + \epsilon)[\varkappa_{\tilde{U}}(Z_l \cap \mathbb{Z}_{\delta_j}) + \tilde{C}^*(y, Z_l \cap \mathbb{Z}_{\delta_j})] \right\} \\ (7.7) \quad &\leq N_{\mathbb{Z},d}(\epsilon/8) \left[\exp\{-y\} + \sum_{k=0}^{\infty} \exp\{2\mathcal{E}_{\mathbb{Z},d}(\epsilon 2^{-k}) - 9y2^{k-3}k^{-2}\} \right] \\ &\leq N_{\mathbb{Z},d}(\epsilon/8) [1 + L_{\text{exp}}^{(\epsilon)}] \exp\{-y/2\}, \end{aligned}$$

where we have used that $y \geq 1$.

Finally, combining (7.1), (7.5), the bound of Theorem 2, and (7.7) we come to the first assertion of the theorem. Here we also used that $C_{\xi}^*(1) \leq C_{\xi}(y)$ because $y \geq 1$.

(ii) In order to prove the second statement, we note first the following nonrandom bound: since $\hat{\Sigma}_s(w) \leq M_s(w)$ for all $w \in \mathcal{W}$ and $s > 2$,

$$\hat{U}_s(w) \leq M_s(w)[c_1(s)\sqrt{n} + 2n^{1/s}] \leq [c_1(s) + 2]\sqrt{n}M_s(w) \quad \forall w \in \mathcal{W}.$$

Next, the left-hand side inequality in (7.2) implies that for any subset $\mathcal{W}_0 \subseteq \mathcal{W}$

$$\mathcal{A} \subseteq \left\{ \sup_{w \in \mathcal{W}_0} \hat{U}_{\xi}(w) < [1 + 4c_1(s)(1 + \epsilon)\gamma] \sup_{w \in \mathcal{W}_0} U_{\xi}(w, f) \right\} =: \mathcal{A}_0.$$

Therefore $\mathbb{P}(\bar{\mathcal{A}}_0) \leq \mathbb{P}(\bar{\mathcal{A}})$ and

$$\mathbb{E}\{[\hat{U}(w)]^q \mathbf{1}(\bar{\mathcal{A}}_0)\} \leq [c_1(s) + 2]^q [\sqrt{n}M_s(w)]^q \mathbb{P}(\bar{\mathcal{A}}).$$

Using (7.7) with $y = y_{\gamma}$, and definition of the event \mathcal{A} , we complete the proof.

7.2. *Proof of Corollary 4.* First, as in (7.2), we need to bound $\check{U}_\xi(w) := \max\{\hat{U}_\xi(w), \sqrt{n}M_2(w)\}$ from above and from below in terms of $\bar{U}_\xi(w, f) := \max\{U_\xi(w, f), \sqrt{n}M_2(w)\}$. Such bounds are easily derived from the following trivial fact: for any positive A, B, C and any $\delta \in (0, 1)$

$$A(1 + \delta) \geq B \geq A(1 - \delta) \quad \Rightarrow \quad [A \vee C](1 + \delta) \geq [B \vee C] \geq [A \vee C](1 - \delta).$$

Next, (7.5) remains valid because, by construction, $U_\xi(w, f) \leq \bar{U}_\xi(w, f)$ and the assumptions, allowing to apply Theorem 2 are imposed now on $\bar{U}_\xi(w, f)$ instead of $U_\xi(w, f)$. The computations leading to (7.7) remain also unchanged if $U_\xi(w, f)$ is replaced by $\bar{U}_\xi(w, f)$. Note that now $\lambda_{\bar{A}}$ and $\lambda_{\bar{B}}$ are defined via $\bar{U}_\xi(w, f)$.

8. Proofs of Theorems 4, 5.

8.1. *Proof of Theorem 4.* The proof is based on an application of Theorem 2. Put

$$T_{3,\epsilon} := \frac{2^{q(\epsilon+1)}u_\epsilon^q}{2^{q\epsilon} - 1} \Gamma(q + 1)N_{\mathbb{Z},d}(\epsilon/8)[1 + L_{\text{exp}}^{(\epsilon)}][4\bar{w}_s(1 + 4n^{1/2-1/s})]^q;$$

$$T_{4,\epsilon} := \frac{2^{q(\epsilon+1)}u_\epsilon^q}{2^{q\epsilon} - 1} \Gamma(q + 1)N_{\mathbb{Z},d}(\epsilon/8)[1 + L_{\text{exp}}^{(\epsilon)}]$$

$$\times \bar{w}_2^q \{1 + 2\sqrt{2\mu_*f_\infty^2 + 8n^{-1/2} + (8/3)n^{-1/2}}\}^q.$$

We have $M_p(w) = \|w\|_p$ for all $w \in \mathcal{V}$ and $p \geq 1$, and (3.9) yields

$$(8.1) \quad U_\xi(w, f) = \begin{cases} 4n^{1/s}\|w\|_s, & s \in [1, 2), \\ \sqrt{n}\|w\|_2, & s = 2. \end{cases}$$

Therefore, in view of (3.10)

$$(8.2) \quad r_\xi = \begin{cases} 4n^{1/s}\underline{w}_s, & s \in [1, 2), \\ \sqrt{n}\underline{w}_2, & s = 2, \end{cases} \quad R_\xi = \begin{cases} 4n^{1/s}\bar{w}_s, & s \in [1, 2), \\ \sqrt{n}\bar{w}_2, & s = 2. \end{cases}$$

It follows from (3.23), the Hölder inequality and the formulas for $A_\xi^2(w)$ and $B_\xi(w)$ immediately after (3.9) that

$$(8.3) \quad A_\xi^2(w) \leq \begin{cases} 37n\|w\|_s^2, & s \in [1, 2), \\ [2f_\infty^2n\mu_* + 8\sqrt{n}]\|w\|_2^2, & s = 2, \end{cases}$$

$$B_\xi(w) = \begin{cases} 0, & s \in [1, 2), \\ \frac{4}{3}\|w\|_2, & s = 2. \end{cases}$$

In order to apply Theorem 2, we need to check that $\varkappa_{U_\xi}(a) \leq a$ for all $a \in [r_\xi, R_\xi]$.

Let $s \in [1, 2)$; here $\mathbb{Z}_a = \{\zeta : a/2 < 4n^{1/s} \|\phi[\zeta]\|_s = 4n^{1/2} \|w\|_s \leq a\}$; see (3.15). By (8.1), Assumption (L) and because $\mathbb{Z}_a \subseteq \mathbb{Z}_s(a/4)$ we have

$$\sup_{\zeta_1, \zeta_2 \in \mathbb{Z}_a} \frac{U_\xi(\phi[\zeta_1] - \phi[\zeta_2], f)}{d(\zeta_1, \zeta_2)} \leq \sup_{\zeta_1, \zeta_2 \in \mathbb{Z}_s(a/4)} \frac{4n^{1/s} \|\phi[\zeta_1] - \phi[\zeta_2]\|_s}{d(\zeta_1, \zeta_2)} \leq a.$$

If $s = 2$, then $\mathbb{Z}_a = \{\zeta : a/2 < \sqrt{n} \|\phi[\zeta]\|_2 = \sqrt{n} \|w\|_2 \leq a\}$, and again by Assumption (L) $\sup_{\zeta_1, \zeta_2 \in \mathbb{Z}_a} [\sqrt{n} \|\phi[\zeta_1] - \phi[\zeta_2]\|_2 / d(\zeta_1, \zeta_2)] \leq a$. Thus, $\varkappa_{U_\xi}(a) \leq a$ for all $a \in [r_\xi, R_\xi]$, and Theorem 2 can be applied. To this end, we should compute the quantities Λ_{A_ξ} and Λ_{B_ξ} [see (2.2), (2.3) and (2.8)].

For $s \in [1, 2)$, we have by (8.3), definition of \mathbb{Z}_a and Assumption (L) that

$$\begin{aligned} \sup_{\zeta \in \mathbb{Z}_a} A_\xi(\phi[\zeta]) &= \sup_{\zeta \in \mathbb{Z}_a} \sqrt{37n} \|\phi[\zeta]\|_s = \frac{\sqrt{37}}{4} a n^{1/2-1/s}, \\ \sup_{\zeta_1, \zeta_2 \in \mathbb{Z}_a} \frac{A_\xi(\phi[\zeta_1] - \phi[\zeta_2])}{d(\zeta_1, \zeta_2)} &\leq \sup_{\zeta_1, \zeta_2 \in \mathbb{Z}_s(a/4)} \frac{\sqrt{37} n^{1/2} \|\phi[\zeta_1] - \phi[\zeta_2]\|_s}{d(\zeta_1, \zeta_2)} \\ &\leq \frac{\sqrt{37}}{4} a n^{1/2-1/s}. \end{aligned}$$

Similarly, if $s = 2$ then $\sup_{\zeta \in \mathbb{Z}_a} A_\xi(\phi[\zeta]) \leq a(2f_\infty^2 \mu_* + 8n^{-1/2})^{1/2}$ and

$$\begin{aligned} \sup_{\zeta_1, \zeta_2 \in \mathbb{Z}_a} \frac{A_\xi(\phi[\zeta_1] - \phi[\zeta_2])}{d(\zeta_1, \zeta_2)} &\leq \sup_{\zeta_1, \zeta_2 \in \mathbb{Z}_2(a)} [2f_\infty^2 n \mu_* + 8\sqrt{n}]^{1/2} \frac{\|\phi[\zeta_1] - \phi[\zeta_2]\|_2}{d(\zeta_1, \zeta_2)} \\ &\leq a \left(2f_\infty^2 \mu_* + \frac{8}{\sqrt{n}} \right)^{1/2}. \end{aligned}$$

These computations and similar computations for Λ_{B_ξ} yield

$$\begin{aligned} \Lambda_{A_\xi} &\leq \begin{cases} \frac{\sqrt{37}}{4} n^{1/2-1/s}, & s \in [1, 2), \\ [2f_\infty^2 \mu_* + 8n^{-1/2}]^{1/2}, & s = 2, \end{cases} \\ \Lambda_{B_\xi} &= \begin{cases} 0, & s \in [1, 2), \\ \frac{4}{3} n^{-1/2}, & s = 2. \end{cases} \end{aligned}$$

Recall that $C_\xi^*(y) = 1 + 2\sqrt{y} \Lambda_{A_\xi} + 2y \Lambda_{B_\xi}$ [see (2.9)]. Therefore if for arbitrary $z > 0$, we set

$$y = \begin{cases} \frac{4}{37} n^{(2/s)-1} z^2, & s \in [1, 2), \\ \frac{z^2}{8} [f_\infty^2 \mu_* + 4n^{-1/2}]^{-1}, & s = 2, \end{cases}$$

then we get $C_\xi^*(y) = 1 + z$ if $s \in [1, 2)$ and

$$C_\xi^*(y) = 1 + z + \frac{z^2}{3\sqrt{n}[f_\infty^2 \mu_* + 4n^{-1/2}]} \leq 1 + z + \frac{z^2}{12},$$

if $s = 2$. Then the statements (i) and (ii) follow by application of the moment bound of Theorem 2. Observe that $C_\xi(1) = 1 + \frac{\sqrt{37}}{2}n^{-1/2-1/s}$ for $s \in [1, 2)$, and $C_\xi(1) = 1 + 2[2f_\infty^2\mu_* + 8n^{-1/2}]^{1/2} + \frac{4}{3}n^{-1/2}$ for $s = 2$; R_ξ is given in (8.2). These expressions along with the moment bound of Theorem 2 lead to the formulas for $T_{1,\epsilon}$ and $T_{2,\epsilon}$ given in the beginning of the proof.

8.2. *Proof of Theorem 5.* First, we specify the constants appearing in the statement of the theorem. Put $\alpha_* := \alpha_1^{-1}\alpha_2^{-1/2}$ where α_1 and α_2 appear in Assumption (W2); then

$$(8.4) \quad \vartheta_1 := [148\alpha_*^4]^{-1}, \quad \vartheta_2 := 5\sqrt{2}c_1(s/2)f_\infty\alpha_*^2C_{s/2}.$$

Define also

$$(8.5) \quad k_* := 8\alpha_*^2c_1(s)[C_s \vee C_{s/2} \vee 1],$$

$$L_*^{(\epsilon)}(\beta) := \sum_{k=1}^\infty \exp\{2^{1+k\beta/m}(k_*^{-1}\epsilon)^{-\beta/m} - (9/16)2^k k^{-2}\},$$

and note that $L_*^{(\epsilon)}(\beta) < \infty$ because $\beta < m$. If we set $I_\epsilon(q) := 2^{q(\epsilon+1)}[2^{q\epsilon} - 1]^{-1}\Gamma(q + 1) + 1$, then the constants $T_{5,\epsilon}$ and $T_{6,\epsilon}$ appearing in the statement of the theorem are given by

$$(8.6) \quad T_{5,\epsilon} := I_\epsilon(q)(2u_\epsilon k_* \bar{w}_2)^q N_{\mathbb{Z},d}([k_*^{-1}\epsilon/8]^{1/m})$$

$$\times \log_2\left(\frac{k_* \bar{w}_2}{w_2}\right)[1 + L_*^{(\epsilon)}(\beta) \exp\{2C_{\mathbb{Z}}(\beta)\}],$$

$$(8.7) \quad T_{6,\epsilon} := [c_1(s) + 2]^q (\alpha_* \bar{w}_2)^q N_{\mathbb{Z},d}([k_*^{-1}\epsilon/8]^{1/m})$$

$$\times \log_2\left(\frac{k_* \bar{w}_2}{w_2}\right)[1 + L_*^{(\epsilon)}(\beta) \exp\{2C_{\mathbb{Z}}(\beta)\}].$$

The proof is based on application of Theorem 3 and Corollary 4. These results will be utilized with a distance d_* on \mathfrak{Z} which is related to the original distance d , and specified below. In order to apply Theorem 3, we need to verify its conditions and to compute the quantities $\Lambda_{A_\xi}, \Lambda_{B_\xi}, \lambda_{\bar{A}}, \lambda_{\bar{B}}$ and γ_γ . These computations are routine and tedious.

We break the proof into steps.

0⁰. *Auxiliary results.* We begin with preliminary results that will be used in the subsequent proof.

LEMMA 6. *Let (3.21) hold and Assumptions (W2) and (W3) be satisfied; then for all $w \in \mathcal{W}$ and $1 \leq p < q \leq \infty$ one has*

$$[n^{1/q} M_q(w)] \leq \alpha_1^{-1} \alpha_2^{-1/p} \mu^{1/q-1/p} [n^{1/p} M_p(w)].$$

PROOF. Recall that under (3.21), $M_p(w) = \|w\|_p$ for all $p \geq 1$. In view of Assumption (W2) for any $w \in \mathcal{V}$, we have

$$\alpha_1 \alpha_2^{1/p} \|w\|_\infty [\text{mes}\{\text{supp}(w)\}]^{1/p} \leq \|w\|_p \leq \|w\|_\infty [\text{mes}\{\text{supp}(w)\}]^{1/p}.$$

This inequality, together with Assumption (W3), yields

$$\begin{aligned} n^{1/q} \|w\|_q &\leq n^{1/q} \|w\|_\infty [\text{mes}\{\text{supp}(w)\}]^{1/q} \\ &= \frac{n^{1/p} \|w\|_\infty [\text{mes}\{\text{supp}(w)\}]^{1/p}}{[n \text{mes}\{\text{supp}(w)\}]^{1/p-1/q}} \\ &\leq \alpha_1^{-1} \alpha_2^{-1/p} \mu^{1/q-1/p} [n^{1/p} \|w\|_p]. \quad \square \end{aligned}$$

Our next lemma demonstrates that there exists a real number $m_p \in (0, 1]$ such that (3.27) holds.

LEMMA 7. *Let Assumptions (W) and (L) hold; then for any $p \geq 2$, the inequality (3.27) is valid with $m_p = 2/p$ and $C_p = (2\alpha_*)^{1-2/p} \mu^{1/p-1/2}$, that is,*

$$\sup_{b \in [\underline{w}_2, \bar{w}_2]} b^{-1} \sup_{\zeta_1, \zeta_2 \in \mathbb{Z}_2(b)} \frac{n^{1/p} \|\phi[\zeta_1] - \phi[\zeta_2]\|_p}{[d(\zeta_1, \zeta_2)]^{2/p}} \leq (2\alpha_*)^{1-2/p} \mu^{1/p-1/2}.$$

PROOF. We obviously have for any $p > 2$

$$\|\phi[\zeta_1] - \phi[\zeta_2]\|_p \leq (\|\phi[\zeta_1]\|_\infty + \|\phi[\zeta_2]\|_\infty)^{1-2/p} (\|\phi[\zeta_1] - \phi[\zeta_2]\|_2)^{2/p}.$$

Applying Lemma 6 with $q = \infty$ and $p = 2$, we have that $\sup_{\zeta \in \mathbb{Z}_2(b)} \|\phi[\zeta]\|_\infty \leq b\alpha_* \mu^{-1/2}$ for all $b \in [\underline{w}_2, \bar{w}_2]$. Then in view of Assumption (L)

$$\sup_{\zeta_1, \zeta_2 \in \mathbb{Z}_2(b)} \frac{n^{1/p} \|\phi[\zeta_1] - \phi[\zeta_2]\|_p}{[d(\zeta_1, \zeta_2)]^{2/p}} \leq b(2\alpha_*)^{1-2/p} \mu^{1/m-1/2} \quad \forall b \in [\underline{w}_2, \bar{w}_2],$$

as claimed. \square

LEMMA 8. *Let Assumptions (W) and (L) hold; then for any $\zeta \in \mathbb{Z}_a$*

$$(8.8) \quad \sqrt{n} \|\phi[\zeta]\|_p \leq \mu_*^{1/p-1/2} a \quad \forall p \in [1, 2),$$

$$(8.9) \quad n^{1/p} \|\phi[\zeta]\|_p \leq \alpha_* \mu^{1/p-1/2} a \quad \forall p > 2,$$

$$(8.10) \quad \sqrt{n} \|\phi^2[\zeta]\|_p \leq \alpha_* \mu^{-1/2} \mu_*^{1/p-1/2} a^2 \quad \forall p \in [1, 2).$$

PROOF. By the Hölder inequality $\|\phi[\zeta]\|_p \leq \mu_*^{1/p-1/2} \|\phi[\zeta]\|_2$; then (8.8) holds by definition of \mathbb{Z}_a . Inequality (8.9) follows Lemma 6. In order to prove (8.10), we write $\|\phi^2[\zeta]\|_p \leq \|\phi[\zeta]\|_\infty \|\phi[\zeta]\|_p$, note that by Lemma 6 $\|\phi[\zeta]\|_\infty \leq \alpha_* \mu^{-1/2} \sqrt{n} \|\phi[\zeta]\|_2$ and use (8.9). \square

1⁰. *Notation.* Now we establish some notation. Recall that $U_\xi(w, f) = c_1(s)[\sqrt{n}\Sigma_s(w, f) + 2n^{1/s}M_s(w)]$ and $\overline{U}_\xi(w, f)$ is given by (3.28). It follows from the definition of $\overline{U}_\xi(\cdot, f)$, (3.9) and (3.6) that $\overline{U}_\xi(w, f) \geq \sqrt{n}\|w\|_2$ and

$$\overline{U}_\xi(w, f) \leq c_1(s)[f_\infty^{1/2-1/s}\sqrt{n}\|w\|_2 + 2n^{1/s}\|w\|_s] \leq c_1(s)\alpha_*[f_\infty^{1/2} + 2]\sqrt{n}\|w\|_2,$$

where the last inequality is a consequence of Lemma 6. Therefore, we put

$$r_\xi = \sqrt{n}\underline{w}_2, \quad R_\xi = c_1(s)\alpha_*[f_\infty^{1/2} + 2]\sqrt{n}\overline{w}_2,$$

where \underline{w}_p and \overline{w}_p are defined in (3.24). Recall also that $\mathbb{Z}_a = \{\zeta : a/2 < \overline{U}_\xi(\phi[\zeta], f) \leq a\}$. By definition of $\overline{U}_\xi(w, f)$ and by the fact that $M_p(w) = \|w\|_p$ for all $p \geq 1$, we have that $\mathbb{Z}_a \subseteq \mathbb{Z}_2(a)$ for all $a \in [r_\xi, R_\xi]$; see (3.25). Define the distance

$$(8.11) \quad d_*(\zeta_1, \zeta_2) = k_* \times \begin{cases} d(\zeta_1, \zeta_2) \vee [d(\zeta_1, \zeta_2)]^{m_s}, & s \in [1, 4), \\ d(\zeta_1, \zeta_2) \vee [d(\zeta_1, \zeta_2)]^{m_s} \\ \vee [d(\zeta_1, \zeta_2)]^{m_{s/2}}, & s > 4, \end{cases}$$

where k_* is given in (8.5). Note that $d_*(\cdot, \cdot)$ is indeed a distance because by definition $m_p \leq 1$ for all $p \geq 2$.

2⁰. *Verification of condition (3.17).* It follows from definition of $\overline{U}_\xi(\cdot, f)$, (3.9) and (3.6) that

$$(8.12) \quad \begin{aligned} \sqrt{n}\|\phi[\zeta]\|_2 &\leq \overline{U}_\xi(\phi[\zeta], f) \\ &\leq c_1(s)[f_\infty^{1/2-1/s}\sqrt{n}\|\phi[\zeta]\|_2 + 2n^{1/s}\|\phi[\zeta]\|_s]. \end{aligned}$$

Therefore, by (8.12), Assumption (L) and (3.27) for any $\zeta_1, \zeta_2 \in \mathbb{Z}_a$

$$\begin{aligned} \overline{U}_\xi(\phi[\zeta_1] - \phi[\zeta_2], f) &\leq c_1(s)[f_\infty^{1/2-1/s}\sqrt{n}\|\phi[\zeta_1] - \phi[\zeta_2]\|_2 + 2n^{1/s}\|\phi[\zeta_1] - \phi[\zeta_2]\|_s] \\ &\leq c_1(s)[f_\infty^{1/2} + 2C_s]a\{d(\zeta_1, \zeta_2) \vee [d(\zeta_1, \zeta_2)]^{m_s}\}. \end{aligned}$$

Thus

$$\sup_{\zeta_1, \zeta_2} \frac{\overline{U}_\xi(\phi[\zeta_1] - \phi[\zeta_2], f)}{d_*(\zeta_1, \zeta_2)} \leq a \quad \forall a \in [r_\xi, R_\xi],$$

and (3.17) is valid, because $k_* \geq c_1(s)[f_\infty^{1/2} + 2C_s]$; see (8.5) and (8.11).

3⁰. *Computation of $\varkappa_{\tilde{U}}$ and verification of (3.18).*

We start with bounds on $\sup_{\zeta \in \mathbb{Z}_a} \tilde{U}(\phi^2[\zeta])$. Recall that

$$(8.13) \quad \tilde{U}(\phi^2[\zeta]) = \begin{cases} 4n^{2/s}M_{s/2}(\phi^2[\zeta]), & s \in (2, 4), \\ c_1(s/2)[f_\infty^{1/2}\sqrt{n}M_2(\phi^2[\zeta]) \\ + 2n^{2/s}M_{s/2}(\phi^2[\zeta])], & s \geq 4. \end{cases}$$

By (8.9), for any $\zeta \in \mathbb{Z}_a$,

$$(8.14) \quad \begin{aligned} n^{2/s} \|\phi^2[\zeta]\|_{s/2} &= (n^{1/s} \|\phi[\zeta]\|_s)^2 \leq \alpha_*^2 \mu^{(2/s)-1} n \|\phi[\zeta]\|_2^2 \\ &\leq \alpha_*^2 \mu^{(2/s)-1} a^2 \quad \forall s > 2, \end{aligned}$$

$$(8.15) \quad \sqrt{n} \|\phi^2[\zeta]\|_2 = (n^{1/4} \|\phi[\zeta]\|_4)^2 \leq \alpha_*^2 \mu^{-1/2} n \|\phi[\zeta]\|_2^2 \leq \alpha_*^2 \mu^{-1/2} a^2.$$

Substituting these bounds in the expression for $\tilde{U}(\phi^2[\zeta])$ and taking into account that $\mu \geq 1$ in view of (W3) we obtain for all $s > 2$

$$(8.16) \quad \sup_{\zeta \in \mathbb{Z}_a} \tilde{U}(\phi^2[\zeta]) \leq k_1 \mu^{2/(s \wedge 4)-1} a^2, \quad k_1 := 4\alpha_*^2 c_1(s/2)[f_\infty^{1/2} + 2].$$

Now we establish bounds on $\tilde{U}(\phi^2[\zeta_1] - \phi^2[\zeta_2])$, $\zeta_1, \zeta_2 \in \mathbb{Z}_a$.

(a) First, we consider the case $s \in (2, 4)$. By the Hölder and triangle inequalities, we have

$$\begin{aligned} n^{2/s} \|\phi^2[\zeta_1] - \phi^2[\zeta_2]\|_{s/2} &\leq n^{2/s-1/2} [\|\phi[\zeta_1]\|_{2s/(4-s)} + \|\phi[\zeta_2]\|_{2s/(4-s)}] \sqrt{n} \|\phi[\zeta_1] - \phi[\zeta_2]\|_2. \end{aligned}$$

Noting that $2s/(4-s) > 2$ and applying (8.9), we have

$$n^{2/s-1/2} \|\phi[\zeta]\|_{2s/(4-s)} \leq \alpha_* \mu^{2/s-1} \sqrt{n} \|\phi[\zeta]\|_2 \leq \alpha_* \mu^{2/s-1} a \quad \forall \zeta \in \mathbb{Z}_a.$$

Then using Assumption (L) we get

$$(8.17) \quad n^{2/s} \|\phi^2[\zeta_1] - \phi^2[\zeta_2]\|_{s/2} \leq 2\alpha_* \mu^{2/s-1} a^2 d(\zeta_1, \zeta_2) \quad \forall \zeta_1, \zeta_2 \in \mathbb{Z}_a.$$

This along with (8.13) implies that for $s \in (2, 4)$

$$(8.18) \quad \tilde{U}(\phi^2[\zeta_1] - \phi^2[\zeta_2]) \leq 8\alpha_* \mu^{2/s-1} a^2 d(\zeta_1, \zeta_2) \quad \forall \zeta_1, \zeta_2 \in \mathbb{Z}_a.$$

(b) Now assume that $s \geq 4$. We have for $\zeta_1, \zeta_2 \in \mathbb{Z}_a$

$$(8.19) \quad \begin{aligned} \sqrt{n} \|\phi^2[\zeta_1] - \phi^2[\zeta_2]\|_2 &\leq [\|\phi[\zeta_1]\|_\infty + \|\phi[\zeta_2]\|_\infty] \sqrt{n} \|\phi[\zeta_1] - \phi[\zeta_2]\|_2 \\ &\leq 2\alpha_* \mu^{-1/2} a^2 d(\zeta_1, \zeta_2), \end{aligned}$$

where we used Assumption (L), and (8.9) with $p = \infty$. Furthermore, we have for all $\zeta_1, \zeta_2 \in \mathbb{Z}$

$$(8.20) \quad \begin{aligned} n^{2/s} \|\phi^2[\zeta_1] - \phi^2[\zeta_2]\|_{s/2} &\leq [\|\phi[\zeta_1]\|_\infty + \|\phi[\zeta_2]\|_\infty] n^{2/s} \|\phi[\zeta_1] - \phi[\zeta_2]\|_{s/2} \\ &\leq 2C_{s/2} \alpha_* \mu^{-1/2} a^2 \{d(\zeta_1, \zeta_2)\}^{m_{s/2}}, \end{aligned}$$

where we have used (8.9) with $p = \infty$ and the definition of m_p [see (3.27)]. These inequalities lead to the following bound: for all $s > 4$

$$\begin{aligned} \tilde{U}(\phi^2[\zeta_1] - \phi^2[\zeta_2]) &\leq 2\alpha_* c_1(s/2)[f_\infty^{1/2} + 2C_{s/2}] \\ &\quad \times \mu^{-1/2} a^2 \{d(\zeta_1, \zeta_2) \vee [d(\zeta_1, \zeta_2)]^{m_{s/2}}\}. \end{aligned}$$

Combining this with (8.18), we obtain that for all $s > 2$

$$(8.21) \quad \tilde{U}(\phi^2[\zeta_1] - \phi^2[\zeta_2]) \leq k_2 \mu^{2/(s \wedge 4) - 1} a^2 \{d(\zeta_1, \zeta_2) \vee d^{m_{s/2}}(\zeta_1, \zeta_2)\},$$

where $k_2 := 8\alpha_* c_1(s/2)[f_\infty^{1/2} + 2C_{s/2}]$. Now using (8.16) and (8.21), we obtain

$$\varkappa_{\tilde{U}}(a) = \sup_{\zeta \in \mathbb{Z}_a} \frac{\tilde{U}(\phi^2[\zeta_1] - \phi^2[\zeta_2])}{d_*(\zeta_1, \zeta_2)} \vee \sup_{\zeta \in \mathbb{Z}_a} \tilde{U}(\phi^2[\zeta]) \leq \mu^{2/(s \wedge 4) - 1} a^2,$$

and the last bound holds because $k_* \geq k_1 \vee k_2$ [see (8.5)]. Thus, the condition (3.18) is valid with

$$(8.22) \quad \gamma = \mu^{1/(s \wedge 4) - 1/2}.$$

Note that condition of the theorem $\mu > [64c_1^2(s)]^{s \wedge 4 / (s \wedge 4 - 2)}$ ensures that $\gamma < [4c_1(1 + \epsilon)]^{-1}$ for any $\epsilon \in (0, 1)$ as required in Theorem 3.

4⁰. *Bounding Λ_{A_ξ} and Λ_{B_ξ} .* By the formula for $A_\xi^2(w)$ given immediately after (3.9), and by (8.8) and (8.9), we have for $\zeta \in \mathbb{Z}_a$

$$\begin{aligned} A_\xi^2(\phi[\zeta]) &\leq 2c_1(s) f_\infty^2 [n \|\phi[\zeta]\|_{2s/(s+2)}^2 + 4\sqrt{n} \|\phi[\zeta]\|_2 \|\phi[\zeta]\|_s + 8n^{1/s} \|\phi[\zeta]\|_s^2] \\ &\leq 2c_1(s) f_\infty^2 a^2 [\mu_*^{2/s} + 12\alpha_*^2 n^{-1/s}] \leq 24\alpha_*^2 c_1(s) f_\infty^2 a^2 [\mu_*^{2/s} + n^{-1/s}]. \end{aligned}$$

Here we have used that $\mu \geq 1$, $\alpha_* \geq 1$ and we write $c_1(s)$ instead of $c_3(s)$ in the definition of $A^2(\cdot)$ because for functions $w(t, x)$ depending on $t - x$ only the constant $c_2(s)$ equals one [see (3.2) and remark after Lemma 3 in Section 6]. Thus,

$$\sup_{\zeta \in \mathbb{Z}_a} A_\xi(\phi[\zeta]) \leq 5\sqrt{c_1(s)} \alpha_* f_\infty a [\mu_*^{1/s} + n^{-1/(2s)}].$$

In order to bound $A_\xi^2(\phi[\zeta_1] - \phi[\zeta_2])$, we note that for all $\zeta_1, \zeta_2 \in \mathbb{Z}_a$:

- by the Hölder inequality and by Assumption (L), $\sqrt{n} \|\phi[\zeta_1] - \phi[\zeta_2]\|_{2s/(s+2)} \leq a\mu_*^{1/s} d(\zeta_1, \zeta_2)$;
- by Assumption (L), $\sqrt{n} \|\phi[\zeta_1] - \phi[\zeta_2]\|_2 \leq ad(\zeta_1, \zeta_2)$;
- by (3.27), $n^{1/s} \|\phi[\zeta_1] - \phi[\zeta_2]\|_s \leq C_s a [d(\zeta_1, \zeta_2)]^{m_s}$.

Therefore,

$$\sup_{\zeta_1, \zeta_2 \in \mathbb{Z}_a} \frac{A_\xi(\phi[\zeta_1] - \phi[\zeta_2])}{d_*(\zeta_1, \zeta_2)} \leq 5\sqrt{c_1(s)} f_\infty (C_s \vee 1) a [\mu_*^{1/s} + n^{-1/(2s)}]$$

and $\Lambda_{A_\xi} \leq 5\sqrt{c_1(s)} \alpha_* f_\infty (C_s \vee 1) [\mu_*^{1/s} + n^{-1/(2s)}]$. Similarly, since $B_\xi(\phi[\zeta]) = \frac{4}{3}c_1(s) \|\phi[\zeta]\|_s$, we have by (8.9) that $\Lambda_{B_\xi} \leq \frac{4}{3}c_1(s) (C_s \vee 1) \alpha_* n^{-1/s}$. Thus, we have shown that

$$\Lambda_{A_\xi} \leq k_3 [\mu_*^{1/s} + n^{-1/(2s)}], \quad \Lambda_{B_\xi} \leq k_3 n^{-1/s}, \quad k_3 := 5c_1(s) \alpha_* f_\infty (C_s \vee 1).$$

These bounds on Λ_{A_ξ} and Λ_{B_ξ} lead to the definition of $C_\xi^*(y)$ in (3.29) [see also (2.9)]. Note that ϑ_0 in (3.29) satisfies $\vartheta_0 = k_3$.

5⁰. *Computation of $\lambda_{\tilde{A}}$, $\lambda_{\tilde{B}}$ and y_γ .*

(i) First, consider the case $s \in (2, 4)$. Recall that in this case $\tilde{A}^2(\phi^2[\zeta]) = 37n\|\phi^2[\zeta]\|_{s/2}^2 = 37n\|\phi[\zeta]\|_s^4$ and $\tilde{B}(\phi^2[\zeta]) = 0$. Hence, by (8.9)

$$\sup_{\zeta \in \mathbb{Z}_a} \tilde{A}(\phi^2[\zeta]) = \sup_{\zeta \in \mathbb{Z}_a} \sqrt{37n}\|\phi[\zeta]\|_s^2 \leq \sqrt{37}\alpha_*^2 \mu^{2/s-1} n^{1/2-2/s} a^2.$$

It follows from (8.17) that for any $\zeta_1, \zeta_2 \in \mathbb{Z}_a$

$$\sqrt{37n}\|\phi^2[\zeta_1] - \phi^2[\zeta_2]\|_{s/2} \leq 2\sqrt{37}\alpha_* \mu^{2/s-1} n^{1/2-2/s} a^2 d(\zeta_1, \zeta_2).$$

Combining these results, we obtain that $\lambda_{\tilde{A}} \leq 2\sqrt{37}\alpha_*^2 \mu^{2/s-1} n^{1/2-2/s}$ and $\lambda_{\tilde{B}} = 0$ which, in turn, by (3.19) and (8.22) implies that

$$y_\gamma = \gamma^4 \lambda_{\tilde{A}}^{-2} \geq (2\sqrt{37}\alpha_*^2)^{-2} n^{4/s-1} =: y_*.$$

This explains the definition of the constant ϑ_1 in (8.4).

(ii) Now let $s \geq 4$; here recall that

$$\begin{aligned} \tilde{A}^2(\phi^2[\zeta]) &= 2c_1(s/2) f_\infty^2 [n\|\phi^2[\zeta]\|_{2s/(s+4)}^2 + 4\sqrt{n}\|\phi^2[\zeta]\|_2 \|\phi^2[\zeta]\|_{s/2} \\ &\quad + 8n^{2/s} \|\phi^2[\zeta]\|_{s/2}^2]. \end{aligned}$$

Observing that for $\zeta \in \mathbb{Z}_a$:

- (a) $\sqrt{n}\|\phi^2[\zeta]\|_{2s/(s+4)} \leq \alpha_* \mu_*^{2/s} a^2$ by (8.10) and $\mu \geq 1$;
- (b) $n^{1/s} \|\phi^2[\zeta]\|_{s/2} \leq \alpha_*^2 n^{-1/s} a^2$ by (8.14);
- (c) $\sqrt{n}\|\phi^2[\zeta]\|_2 \|\phi^2[\zeta]\|_{s/2} \leq \alpha_*^2 \mu^{-1/2} n^{-2/s} a^4$ by (8.15) and (b),

we obtain

$$(8.23) \quad \sup_{\zeta \in \mathbb{Z}_a} \tilde{A}(\phi[\zeta]) \leq 5\sqrt{c_1(s/2)} f_\infty \alpha_* a [\mu_*^{2/s} + n^{-1/s}].$$

Similarly, for $\zeta_1, \zeta_2 \in \mathbb{Z}_a$ we have

$$\begin{aligned} \sqrt{n}\|\phi^2[\zeta_1] - \phi^2[\zeta_2]\|_{2s/(s+4)} &\leq \mu_*^{2/s} \sqrt{n}\|\phi^2[\zeta_1] - \phi^2[\zeta_2]\|_2 \\ &\leq 2\alpha_* \mu_*^{2/s} a^2 d(\zeta_1, \zeta_2), \\ n^{1/s} \|\phi^2[\zeta_1] - \phi^2[\zeta_2]\|_{s/2} &\leq 2C_{s/2} \alpha_* n^{-1/s} a^2 \{d(\zeta_1, \zeta_2)\}^{m_{s/2}}, \\ [\sqrt{n}\|\phi^2[\zeta_1] - \phi^2[\zeta_2]\|_2 \|\phi^2[\zeta_1] - \phi^2[\zeta_2]\|_{s/2}]^{1/2} \\ &\leq 2\sqrt{C_{s/2} \alpha_*} n^{-1/s} a^2 \{d(\zeta_1, \zeta_2) \vee [d(\zeta_1, \zeta_2)]^{m_{s/2}}\}, \end{aligned}$$

where the first line follows from the Hölder inequality and (8.19); the second one follows from (8.20); and the third line follows from the two previous inequalities.

This yields

$$\sup_{\zeta_1, \zeta_2 \in \mathbb{Z}_a} \frac{\tilde{A}(\phi^2[\zeta_1] - \phi^2[\zeta_2])}{d_*(\zeta_1, \zeta_2)} \leq 5f_\infty \sqrt{2c_1(s/2)} \alpha_* C_{s/2} a^2 [\mu_*^{2/s} + n^{-1/s}].$$

Combining the last inequality with (8.23), we obtain

$$\lambda_{\tilde{A}} \leq k_4[\mu_*^{2/s} + n^{-1/s}], \quad k_4 := 5f_\infty\sqrt{2c_1(s/2)\alpha_*C_{s/2}}.$$

Now in order to bound $\lambda_{\tilde{B}}$ we recall that $\tilde{B}(\phi^2[\zeta]) = \frac{4}{3}c_1(s/2)\|\phi^2[\zeta]\|_{s/2}$. Then (8.14) gives $\sup_{\zeta \in \mathbb{Z}_a} \tilde{B}(\phi^2[\zeta]) \leq \frac{4}{3}c_1(s/2)\alpha_*^2n^{-2/s}a^2$. This along with (8.20) leads to

$$\lambda_{\tilde{B}} \leq k_5n^{-2/s}, \quad k_5 := \frac{8}{3}c_1(s/2)\alpha_*^2C_{s/2}.$$

Combining these results with (3.19) and taking into account that, by (8.22), $\gamma = \mu^{-1/4} \leq 1$ for $s \geq 4$, we have

$$\mu^{-1/4} = \sqrt{y_\gamma}\lambda_{\tilde{A}} + y_\gamma\lambda_{\tilde{B}} \leq [\sqrt{y_\gamma} + y_\gamma](k_4 \vee k_5)[\mu_*^{2/s} + n^{-1/s}],$$

and an elementary calculation shows that

$$y_\gamma \geq \mu^{-1/2}(k_4 \vee k_5)^{-2}[\mu_*^{2/s} + n^{-1/s}]^{-2} =: y_*.$$

This inequality yields the constant ϑ_2 appearing in (8.4).

6^0 . *Application of Theorem 3.* In order to apply Theorem 3 with the distance $d_*(\cdot, \cdot)$ given in (8.11), we need to compute the quantity

$$L_{\text{exp}}^{(\epsilon)} = \sum_{k=1}^{\infty} \exp\{2\mathcal{E}_{\mathbb{Z}, d_*}(\epsilon 2^{-k}) - (9/16)2^k k^{-2}\}.$$

Note that the entropy number $\mathcal{E}_{\mathbb{Z}, d_*}(\cdot) = \ln\{N_{\mathbb{Z}, d_*}(\cdot)\}$ is computed with respect to the distance d_* . Therefore, we first express the entropy $\mathcal{E}_{\mathbb{Z}, d_*}(\cdot)$ in terms of the original distance d and then, using Assumption (W4), we derive a bound for $L_{\text{exp}}^{(\epsilon)}$.

By the definition of the distance d_* , for all $\delta \in (0, 1)$ and $\zeta_1, \zeta_2 \in \mathbb{Z}$,

$$d(\zeta_1, \zeta_2) \leq [k_*^{-1}\delta]^{1/m} \Rightarrow d_*(\zeta_1, \zeta_2) \leq \delta,$$

where $m := 1 \wedge m_s$ if $s \in (2, 4)$ and $m := 1 \wedge m_s \wedge m_{s/2}$ if $s \geq 4$. Therefore, $N_{\mathbb{Z}, d_*}(\delta) \leq N_{\mathbb{Z}, d}([k_*^{-1}\delta]^{1/m})$. In view of Assumption (W4), this yields

$$\begin{aligned} \sup_{\delta \in (0,1)} \{\mathcal{E}_{\mathbb{Z}, d_*}(\delta) - [k_*^{-1}\delta]^{-\beta/m}\} &\leq \sup_{\delta \in (0,1)} \{\mathcal{E}_{\mathbb{Z}, d}([k_*^{-1}\delta]^{1/m}) - [k_*^{-1}\delta]^{-\beta/m}\} \\ &\leq \sup_{x \in (0,1)} \{\mathcal{E}_{\mathbb{Z}, d}(x) - x^{-\beta}\} = C_{\mathbb{Z}}(\beta). \end{aligned}$$

Thus, we obtain that

$$\begin{aligned} L_{\text{exp}}^{(\epsilon)} &\leq \exp\{2C_{\mathbb{Z}}(\beta)\} \sum_{k=1}^{\infty} \exp\{2^{1+k\beta/m}(k_*^{-1}\epsilon)^{-\beta/m} - (9/16)2^k k^{-2}\} \\ &= \exp\{2C_{\mathbb{Z}}(\beta)\} L_*^{(\epsilon)}(\beta). \end{aligned}$$

Now the result of the theorem follows from the bounds of Theorem 3. The constants $T_{5,\epsilon}$ and $T_{6,\epsilon}$ given in the beginning of the proof are obtained from the expressions for $T_{1,\epsilon}$ and $T_{2,\epsilon}$ and bounds of Theorem 3. In particular, we used that in view of Lemma 6 $\sqrt{n} \sup_{w \in \mathcal{V}_0} \|w\|_s \leq n^{(s-2)/(2s)}\alpha_*\mu^{1/s-1/2}\bar{w}_2$.

9. Proof of Theorem 7. The proof is based on verification of conditions and application of Theorem 5. First, we establish auxiliary results that provide the basis for verification of Assumptions (L) and (W). Then, based on these results, we show that all conditions of Theorem 5 are fulfilled. This will yield the required result.

Let K and K' be any functions satisfying Assumptions (K1) and (K2), and let h, h' be given vectors from \mathcal{H} . Let $\zeta = (K, h), \zeta' = (K', h')$, and recall that $\phi_1[\zeta]$ is the mapping $(K, h) \mapsto n^{-1}K_h$.

Similarly, if K, Q, K', Q' are any functions satisfying Assumptions (K1) and (K2), and if $h, h', \mathfrak{h}, \mathfrak{h}'$ are vectors from \mathcal{H} then $z = [(K, h), (Q, \mathfrak{h})], z' = [(K', h'), (Q', \mathfrak{h}')]$, and $\phi_2[z]$ is the mapping $[(K, h), (Q, \mathfrak{h})] \mapsto n^{-1}(K_h * Q_{\mathfrak{h}})$.

1⁰. Auxiliary results. We begin with auxiliary results about properties of the mappings $\phi_1[\zeta]$ and $\phi_2[z]$. The proofs of these results are given in the Appendix.

Define the function

$$(9.1) \quad D(x) := e^{dx} \left[x + \frac{1}{2} L_{\mathcal{K}} \sqrt{d} (e^x - 1) + k_{\infty} (e^{dx} - 1) \right], \quad x \geq 0,$$

and put

$$(9.2) \quad \theta_1 := [k_{\infty}/k_1] D'(2), \quad \theta_2 := 2^{2d+2} k_{\infty}^4 k_1^{-2} D'(4),$$

where D' is the first derivative of the function D .

The next lemma states that Assumption (L) is fulfilled for the mappings $\zeta \mapsto \phi_1[\zeta]$ and $z \mapsto \phi_2[z]$.

LEMMA 9. *Let Assumption (K) hold, and $s \geq 1$. If the sets $\mathbb{Z}^{(i)}, i = 1, 2$, are equipped with the distances $d_{\theta_i}^{(i)}(\cdot, \cdot)$ then Assumption (L) is valid for the mappings $\zeta \mapsto \phi_1[\zeta]$ and $z \mapsto \phi_2[z]$.*

The next three statements provide a basis for verification of Assumption (W). For any $h, h' \in \mathcal{H}$, let $h \vee h' = (h_1 \vee h'_1, \dots, h_d \vee h'_d)$ and $h \wedge h' = (h_1 \wedge h'_1, \dots, h_d \wedge h'_d)$.

LEMMA 10. *Let Assumptions (K1) and (K2) hold; then for any $p \in [1, \infty]$*

$$(9.3) \quad \|\phi_1[\zeta]\|_p = n^{-1} V_h^{-1+1/p} \|K\|_p \quad \forall \zeta \in \mathbb{Z}^{(1)},$$

$$(9.4) \quad \|\phi_1[\zeta] - \phi_1[\zeta']\|_p \leq n^{-1} (V_{h \vee h'})^{-1+1/p} D(d_1^{(1)}(\zeta, \zeta')) \quad \forall \zeta, \zeta' \in \mathbb{Z}^{(1)},$$

$$(9.5) \quad \|\phi_2[z] - \phi_2[z']\|_p \leq 2n^{-1} k_{\infty} [(V_{h \vee h'}) \vee (V_{\mathfrak{h} \vee \mathfrak{h}'])]^{-1+1/p} \\ \times D(2d_1^{(2)}(z, z')) \quad \forall z, z' \in \mathbb{Z}^{(2)}.$$

Observe that Lemma 10 implies that Assumption (A2) of Section 3.2 is fulfilled for the mappings $\zeta \mapsto \phi_1[\zeta]$ and $z \mapsto \phi_2[z]$.

LEMMA 11. Let $w \in \mathbb{H}_d(1, P)$ with some $P > 0$, and let $\tilde{x} \in \mathbb{R}^d$ be a point such that $w(\tilde{x}) = \|w\|_\infty > 0$; then

$$\left\{x \in \mathbb{R}^d : |w(x)| \geq \frac{1}{2} \|w\|_\infty\right\} \supseteq \bigotimes_{i=1}^d \left[\tilde{x}_i - \frac{\|w\|_\infty}{2P\sqrt{d}}, \tilde{x}_i + \frac{\|w\|_\infty}{2P\sqrt{d}}\right].$$

LEMMA 12. Under Assumption (K) for any $p \geq 1$:

- (i) $\|\phi_2[z]\|_p \leq 2^{d/p} k_\infty^2 n^{-1} (V_{h \vee \mathfrak{h}})^{-1+1/p}$,
- (ii) $\|\phi_2[z]\|_p \geq 2^{d(1-p)/p} k_1^2 n^{-1} (V_{h \vee \mathfrak{h}})^{-1+1/p}$,
- (iii) $\text{mes}\{\text{supp}(\phi_2[z])\} \geq (V_{h \vee \mathfrak{h}}) \left[\frac{k_1^2}{2^{d+1}\sqrt{d}L_{\mathcal{K}}k_\infty}\right]^d$,
- (iv) $\text{mes}\left\{t : \phi_2[z](t) \geq \frac{1}{2} \|\phi_2[z]\|_\infty\right\} \geq \left[\frac{k_1^2}{2^{d+2}\sqrt{d}L_{\mathcal{K}}k_\infty}\right]^d \text{mes}\{\text{supp}(\phi_2[z])\}$.

2⁰. Verification of conditions of Theorem 5. We check Assumption (W) for the classes of weights $\mathcal{W}^{(1)}$ and $\mathcal{W}^{(2)}$ given by the parametrization $\phi_1[\zeta]$ and $\phi_2[z]$.

First, we note that (W1) is fulfilled both for $\phi_1[\zeta]$ and $\phi_2[z]$ in view of Assumption (K1). Furthermore, Assumptions (K1) and (K2) together with Lemma 11 imply (W2) for $\phi_1[\zeta]$ with

$$(9.6) \quad \alpha_1 = \frac{1}{2}, \quad \alpha_2 = \alpha_{2,1} := \left[\frac{k_1}{L_{\mathcal{K}}\sqrt{d}}\right]^d,$$

while the statement (iv) of Lemma 12 yield (W2) for $\phi_2[z]$ with the constants

$$(9.7) \quad \alpha_1 = \frac{1}{2}, \quad \alpha_2 = \alpha_{2,2} := \left[\frac{k_1^2}{2^{d+2}\sqrt{d}L_{\mathcal{K}}k_\infty}\right]^d.$$

Clearly, $\text{mes}\{\text{supp}(\phi_1[\zeta])\} \geq V_{h \min}$; hence the condition

$$nV_{h \min} > [64c_1^2(s)]^{(s \wedge 4)/(s \wedge 4 - 2)}$$

implies (W3) for $\phi_1[\zeta]$ with $\mu = nV_{h \min}$. It follows from the statement (iii) of Lemma 12 that Assumption (W3) holds for $\phi_2[z]$ with $\mu = nV_{h \min}$ if

$$nV_{h \min} > \alpha_{2,2}^{-1} [64c_1^2(s)]^{(s \wedge 4)/(s \wedge 4 - 2)}.$$

Finally, a standard calculation shows that if $\mathcal{E}_{\mathcal{H}}(\cdot)$ is the entropy number of the set \mathcal{H} measured in the distance $\Delta_{\mathcal{H}}$ [see (3.31)] then for any $\delta \in (0, 1]$

$$(9.8) \quad \mathcal{E}_{\mathcal{H}}(\delta) \leq d \ln(3/\delta) + \sum_{i=1}^d (\ln \ln [h_i^{\max} / h_i^{\min}])_+.$$

This result together with (K3) guarantees that Assumption (W4) is fulfilled for the both parametrizations.

Now we compute the quantities m_p and C_p appearing in (3.27). Although Lemma 7 shows that we always can set $m_p = 2/p$, it turns out that under Assumption (K) we can put $m_p = 1$ for all $p \geq 2$ both for $\phi_1[\zeta]$ and for $\phi_2[z]$. This leads to weaker conditions on the entropy $\mathcal{E}_{\mathcal{K}}(\cdot)$ (see formulation of Theorem 5).

First, consider the mapping $\phi_1[\zeta]$; here following (3.25), we set

$$\mathbb{Z}_2^{(1)}(b) := \{\zeta = (K, h) : n^{1/2} \|\phi_1[\zeta]\|_2 \leq b\} = \{\zeta = (K, h) : (nV_h)^{-1/2} \|K\|_2 \leq b\}$$

for $b \in [\underline{w}_2^{(1)}, \overline{w}_2^{(1)}]$ where by (3.24)

$$(9.9) \quad \underline{w}_2^{(1)} \geq k_1(nV_{h_{\max}})^{-1/2}, \quad \overline{w}_2^{(1)} \leq k_{\infty}(nV_{h_{\min}})^{-1/2}.$$

By (9.4) of Lemma 10, we have for any $p \geq 2$ and $\zeta_1 = (K, h)$, $\zeta_2 = (K', h')$

$$n^{1/p} \|\phi_1[\zeta_1] - \phi_1[\zeta_2]\|_p \leq (nV_{h \vee h'})^{-1+1/p} D(d_1^{(1)}(\zeta_1, \zeta_2)),$$

and if $\zeta_1, \zeta_2 \in \mathbb{Z}_2^{(1)}(b)$ are such that $d_1^{(1)}(\zeta_1, \zeta_2) \leq 2$ then by definition of $\mathbb{Z}_2^{(1)}(b)$

$$n^{1/p} \|\phi_1[\zeta_1] - \phi_1[\zeta_2]\|_p \leq [b/k_1]^{2-2/p} D'(2) d_1^{(1)}(\zeta_1, \zeta_2),$$

where we have used that $\|K\|_2 \geq \|K\|_1 \geq k_1$ for all $K \in \mathcal{K}$, $D(0) = 0$, and D is monotone increasing. If $\zeta_1, \zeta_2 \in \mathbb{Z}_2^{(1)}(b)$ and $d_1^{(1)}(\zeta_1, \zeta_2) > 2$, then by the triangle inequality, and (9.3) of Lemma 10

$$\begin{aligned} n^{1/p} \|\phi_1[\zeta_1] - \phi_1[\zeta_2]\|_p &\leq n^{1/p} \|\phi_1[\zeta_1]\|_p + n^{1/p} \|\phi_1[\zeta_2]\|_p \\ &\leq 2k_{\infty}(nV_h)^{-1+1/p} \leq k_{\infty}[b/k_1]^{2-2/p} d_1^{(1)}(\zeta_1, \zeta_2). \end{aligned}$$

These inequalities show that if $\mathbb{Z}^{(1)}$ is equipped with the distance $d_{\theta_1}^{(1)}(\cdot, \cdot)$ [see (9.2) for definition of θ_1] then (3.27) holds with

$$(9.10) \quad m_p = 1, \quad C_p = \theta_1^{-1} [k_{\infty}/k_1]^{2-2/p} D'(2) \leq 1,$$

because $nV_{h_{\min}} \geq 1$ (which implies $b \leq k_{\infty}$).

Now consider the mapping $\phi_2[z]$; following (3.25) we have here

$$\mathbb{Z}_2^{(2)}(b) := \{z = [(K, h), (Q, h)] : n^{1/2} \|\phi_2[z]\|_2 \leq b\}, \quad b \in [\underline{w}_2^{(2)}, \overline{w}_2^{(2)}],$$

where by the statements (i) and (ii) of Lemma 12

$$(9.11) \quad 2^{-d/2} k_1^2 (nV_{h_{\max}})^{-1/2} \leq \underline{w}_2^{(2)}, \quad \overline{w}_2^{(2)} \leq 2^{d/2} k_{\infty}^2 (nV_{h_{\min}})^{-1/2}.$$

Note that if $z = [(K, h), (Q, h)] \in \mathbb{Z}_2^{(2)}(b)$ then by the statement (ii) of Lemma 12 we have $(nV_{h \vee h})^{-1} \leq 2^d k_1^{-4} b^2$. By this fact and by (9.5) of Lemma 10, we have for $z_1 = [(K, h), (Q, h)]$, $z_2 = [(K', h'), (Q', h')]$ $\in \mathbb{Z}_2^{(2)}$ such that $d_1^{(2)}(z_1, z_2) \leq 2$

$$\begin{aligned} n^{1/p} \|\phi_2[z_1] - \phi_2[z_2]\|_p &\leq 2k_{\infty} [(nV_{h \vee h'}) \vee (nV_{h' \vee h})]^{-1+1/p} D(2d_1^{(2)}(z_1, z_2)) \\ &\leq 2^{d+2-d/p} k_{\infty} [b/k_1^2]^{2-2/p} D'(4) d_1^{(2)}(z_1, z_2). \end{aligned}$$

If $d_1^{(2)}(z_1, z_2) > 2$, then using the triangle inequality and Lemma 12(i) we have

$$n^{1/p} \|\phi_2[z_1] - \phi_2[z_2]\|_p \leq 2^{d+1} k_\infty^2 [b/k_1^2]^{2-2/p} \leq 2^d k_\infty^2 [b/k_1^2]^{2-2/p} d_1^{(2)}(z_1, z_2).$$

Combining these inequalities, we observe that if $\mathbb{Z}^{(2)}$ is equipped with the distance $d_{\theta_2}^{(2)}(\cdot, \cdot)$ [see (9.2)] then (3.27) holds with

$$(9.12) \quad m_p = 1, \quad C_p = \theta_2^{-1} 2^{2d+2-3d/p} k_\infty^2 [k_\infty^2/k_1]^{2-2/p} D'(4) \leq 1.$$

We have used that $b \leq 2^{d/2} k_\infty^2$ because $nV_{h_{\min}} \geq 1$. Thus (9.10) and (9.12) show that $m = 1$ and the condition $\beta < m$ of Theorem 5 holds if in Assumption (K3) $\beta_{\mathcal{K}} < 1$.

3⁰. *Application of Theorem 5.* First, note that $\vartheta_0^{(i)}, i = 1, 2$, defined in (3.38) satisfy

$$\vartheta_0^{(i)} := 5c_1(s) f_\infty \alpha_{*,i}, \quad \alpha_{*,i} := 2/\sqrt{\alpha_{2,i}}, \quad i = 1, 2,$$

where $\alpha_{2,i}, i = 1, 2$, are given in (9.6) and (9.7). This is in accordance with the definition of the constant ϑ_0 in (3.29) for the parametrizations $\phi_1[\zeta]$ and $\phi_2[z]$. Then the definition of $C_{\xi,i}^*(y)$ in (3.40) corresponds to (3.29). Following (8.4), we put

$$\vartheta_1^{(i)} := \alpha_{*,i}^{-4}/148, \quad \vartheta_2^{(i)} := 5\sqrt{2}c_1(s/2) f_\infty \alpha_{*,i}^2, \quad i = 1, 2.$$

Then the formula for $y_*^{(i)}$ appearing in the statement of the theorem is a version of (3.30).

Now we need to specify the constants $T_{5,\epsilon}$ and $T_{6,\epsilon}$; see (8.6), (8.7).

Following (8.5), we set for $i = 1, 2$

$$k_{*,i} := 8c_1(s) \alpha_{*,i}^2, \quad L_{*,i}^{(\epsilon)}(\beta) := \sum_{k=1}^{\infty} \exp\{2^{1+k\beta} (k_{*,i}^{-1}\epsilon)^{-\beta} - (9/16)2^k k^{-2}\}.$$

In view of (9.8) and Assumption (K3), we obtain for any $\beta \in (\beta_{\mathcal{K}}, 1)$ that

$$\begin{aligned} C_{\mathbb{Z}^{(1)}}(\beta) &= \sup_{\delta \in (0,1)} \{ \mathcal{E}_{\mathbb{Z}^{(1)}, d_{\theta_1}^{(1)}}(\delta) - \delta^{-\beta} \} \\ &\leq C_{\mathcal{K}} + C_{\beta,d} + \sum_{i=1}^d (\ln \ln [h_i^{\max}/h_i^{\min}])_+, \\ C_{\mathbb{Z}^{(2)}}(\beta) &= \sup_{\delta \in (0,1)} \{ \mathcal{E}_{\mathbb{Z}^{(2)}, d_{\theta_2}^{(2)}}(\delta) - \delta^{-\beta} \} \\ &\leq 2C_{\mathcal{K}} + 2C_{\beta,d} + 2 \sum_{i=1}^d (\ln \ln [h_i^{\max}/h_i^{\min}])_+, \end{aligned}$$

where we have taken into account that $\theta_1 \geq 1, \theta_2 \geq 1$ and denoted

$$C_{\beta,d} := \sup_{\delta \in (0,1)} [d \ln(3/\delta) + \delta^{-\beta_{\mathcal{K}}} - \delta^{-\beta}], \quad \beta \in (\beta_{\mathcal{K}}, 1).$$

Therefore for $i = 1, 2$

$$L_{*,i}^{(\epsilon)}(\beta) \exp\{2C_{\mathbb{Z}^{(i)}}(\beta)\} \leq [1 + A_{\mathcal{H}}]^i \exp\{2iC_{\mathcal{K}}\} \inf_{\beta \in (\beta_{\mathcal{K}}, 1)} [L_{*,i}^{(\epsilon)}(\beta) \exp\{2iC_{\beta,d}\}],$$

and, by Assumption (K3), (9.8) and (3.39)

$$N_{\mathbb{Z}^{(i)}, d_{\theta_i}^{(i)}}(k_{*,i}^{-1}\epsilon/8) \leq [1 + A_{\mathcal{H}}]^i [24k_{*,i}\theta_i/\epsilon]^{di} \exp\left\{i\left(\frac{8k_{*,i}\theta_i}{\epsilon}\right)^{\beta_{\mathcal{K}}}\right\} \exp\{iC_{\mathcal{K}}\}.$$

Finally, substituting these bounds in (8.6) and using (9.9), (9.11) and (3.39) we have that

$$T_{5,\epsilon}^{(i)} \leq (1 + A_{\mathcal{H}})^{2i} (1 + B_{\mathcal{H}}) \tilde{T}_1^{(i)},$$

where

$$\begin{aligned} \tilde{T}_{1,\epsilon}^{(i)} &:= I_{\epsilon}(q)(2^{1+d/2}u_{\epsilon}k_{*,i}k_{\infty}^2)^q [24k_{*,i}\theta_i\epsilon^{-1}]^{di} \exp\left\{i\left(\frac{8k_{*,i}\theta_i}{\epsilon}\right)^{\beta_{\mathcal{K}}}\right\} \exp\{3iC_{\mathcal{K}}\} \\ &\quad \times \log_2\left(\frac{2^d k_{\infty}^2 k_{*,i}}{k_1^2}\right) \left\{1 + \inf_{\beta \in (\beta_{\mathcal{K}}, 1)} [L_{*,i}^{(\epsilon)}(\beta) \exp\{2iC_{\beta,d}\}]\right\}, \quad i = 1, 2. \end{aligned}$$

This leads to the first statement of the theorem. The second statement of the theorem follows substitution of the above bounds in (8.7) which gives $T_{6,\epsilon}^{(i)} \leq (1 + A_{\mathcal{H}})^{2i} (1 + B_{\mathcal{H}}) \tilde{T}_2^{(i)}$, where

$$\begin{aligned} \tilde{T}_{2,\epsilon}^{(i)} &:= [c_1(s) + 2]^q (2^{d/2}\alpha_{*,i}k_{\infty}^2)^q [24k_{*,i}\theta_i\epsilon^{-1}]^{di} \exp\left\{i\left(\frac{8k_{*,i}\theta_i}{\epsilon}\right)^{\beta_{\mathcal{K}}}\right\} \exp\{3iC_{\mathcal{K}}\} \\ &\quad \times \log_2\left(\frac{2^d k_{\infty}^2 k_{*,i}}{k_1^2}\right) \left\{1 + \inf_{\beta \in (\beta_{\mathcal{K}}, 1)} [L_{*,i}^{(\epsilon)}(\beta) \exp\{2iC_{\beta,d}\}]\right\}, \quad i = 1, 2. \end{aligned}$$

10. Proofs of Theorems 8 and 9.

10.1. *Proof of Theorem 8.* Let $X' = (X, \varepsilon)$, and let $X'_i, i = 1, \dots, n$ be independent copies of X' . For any $l > 0, x' = (x, u) \in \mathcal{X} \times \mathbb{R}$ and $t \in \mathcal{T}$ define the function

$$w^{(l)}(t, x') = w(t, x)u\mathbf{1}_{[-l,l]}(u).$$

With this notation, we note that on the event $\{\max_{i=1,\dots,n} |\varepsilon_i| \leq l\}$

$$\eta_w(t) = \sum_{i=1}^n w(t, X_i)\varepsilon_i = \sum_{i=1}^n w^{(l)}(t, X'_i) = \xi_{w^{(l)}}(t),$$

and the last equality holds because $\mathbb{E}w^{(l)}(t, X') = 0$, for all $t \in \mathcal{T}$ and $l > 0$ because the distribution of ε is symmetric. Therefore for any $z > 0$,

$$\mathbb{P}\{\|\eta_w\|_{s,\tau} \geq z\} \leq \mathbb{P}\{\|\xi_{w^{(l)}}\|_{s,\tau} \geq z\} + n\mathbb{P}\{|\varepsilon| > l\}.$$

If Assumption (E1) is fulfilled, then for any $z > 0$

$$(10.1) \quad \mathbb{P}\{\|\eta_w\|_{s,\tau} \geq z\} \leq \mathbb{P}\{\|\xi_{w^{(l)}}\|_{s,\tau} \geq z\} + nv \exp\{-bl^\alpha\}.$$

If Assumption (E2) is fulfilled, then for any $z > 0$

$$(10.2) \quad \mathbb{P}\{\|\eta_w\|_{s,\tau} \geq z\} \leq \mathbb{P}\{\|\xi_{w^{(l)}}\|_{s,\tau} \geq z\} + nPl^{-p}.$$

In order to bound the first term on the right-hand side of (10.1) and (10.2), we repeat the steps in the proof of Theorem 1 with w replaced by $w^{(l)}$ and optimize with respect to the truncation level l .

For any $z > 0$, we define

$$\Upsilon_s(w, f, z) = \frac{z^2}{(1/3)\varpi_s^2(w, f) + (4/3)c_*(s)M_s(w)z},$$

where $c_*(s)$ is given in (3.2).

First, consider the case $s \geq 2$. Using the same reasoning as in the proof of Theorem 1, we have the following upper bound: for all $z > 0$

$$(10.3) \quad \mathbb{P}\{\|\xi_{w^{(l)}}\|_{s,\tau} \geq \varrho_s(w, f) + z\} \leq \exp\{-[1 \vee l]^{-1}\Upsilon_s(w, f, z)\}.$$

Under Assumption (E1), if we set

$$l = \begin{cases} [b^{-1}\Upsilon_s(w, f, z)]^{1/\alpha}, & b^{-1}\Upsilon_s(w, f, z) < 1, \\ [b^{-1}\Upsilon_s(w, f, z)]^{1/(1+\alpha)}, & b^{-1}\Upsilon_s(w, f, z) \geq 1, \end{cases}$$

then it follows from (10.1) and (10.3) that

$$\mathbb{P}\{\|\eta_w\|_{s,\tau} \geq \varrho_s(w, f) + z\} \leq G^{(1)}(\Upsilon_s(w, f, z)).$$

Thus, the first statement of the theorem is proved if $s \geq 2$.

If Assumption (E2) is fulfilled then we choose

$$l = \frac{\Upsilon_s(w, f, z)}{p \ln(1 + p^{-1}\Upsilon_s(w, f, z))}$$

and note that $l \geq 1$ for any value of $\Upsilon_s(w, f, z)$. Then (10.2) and (10.3) imply that

$$\begin{aligned} \mathbb{P}\{\|\eta_w\|_{s,\tau} \geq \varrho_s(w, f) + z\} &\leq \left[\frac{1}{(1 + p^{-1}\Upsilon_s(w, f, z))} \right]^p \\ &\quad + nP \left[\frac{p \ln(1 + p^{-1}\Upsilon_s(w, f, z))}{\Upsilon_s(w, f, z)} \right]^p. \end{aligned}$$

Using the trivial inequality $(1 + u)^{-1} \leq u^{-1} \ln(1 + u)$, $u \geq 0$ we get

$$\mathbb{P}\{\|\eta_w\|_{s,\tau} \geq \varrho_s(w, f) + z\} \leq [1 + nP] \left[\frac{p \ln(1 + p^{-1}\Upsilon_s(w, f, z))}{\Upsilon_s(w, f, z)} \right]^p$$

and, therefore, the second statement of the theorem is proved for the case $s \geq 2$.

If $s < 2$, then we have similarly to (10.3) that for all $z > 0$

$$\mathbb{P}\{\|\xi_{w^{(l)}}\|_{s,\tau} \geq \varrho_s(w, f) + z\} \leq \exp\{-[1 \vee l]^{-2}\Upsilon_s(w, f, z)\}.$$

The same computations as in the case $s \geq 2$ lead to the statement of the theorem when $s < 2$.

10.2. Proof of Theorem 9. Put

$$L_{\alpha,b}^{(\epsilon)} := \sum_{k=1}^{\infty} \exp\{\epsilon^{-\beta} 2^{\beta k+1}\} \sqrt{g_{\alpha,b}(9 \cdot 2^{k-3} k^{-2})},$$

$$J_{\alpha,b}^{(\epsilon)} := q \int_1^{\infty} (x-1)^{q-1} [g_{\alpha,b}(x)]^{1/4} dx,$$

$$T_{n,\epsilon} := (1+nv)[2^{2\epsilon}(1+\epsilon)a\bar{w}_2]^q [2^{q\epsilon} - 1]^{-1} \exp\{C_{\mathbb{Z}}(\beta) + (8/\epsilon)^\beta\} \\ \times (1 + \exp\{2C_{\mathbb{Z}}(\beta)\}) L_{\alpha,b}^{(\epsilon)} J_{\alpha,b}^{(\epsilon)}.$$

We note that $L_{\alpha,b}^{(\epsilon)} < \infty$ since $\beta < \alpha/(2 + \alpha)$ if $s < 2$, and $\beta < \alpha/(1 + \alpha)$ if $s \geq 2$. Note also that the quantity $J_g^{(\epsilon)}(\cdot)$ in the second inequality of Corollary 1 admits the following bound if $g = G_1$:

$$J_g^{(\epsilon)}(z) \leq (1+nv)[g_{\alpha,b}(z)]^{1/4} (1 + L_{\alpha,b}^{(\epsilon)}) J_{\alpha,b}^{(\epsilon)}, \quad z > 0.$$

If for any $\zeta \in \mathbb{Z}$, we let

$$U_\eta(\phi[\zeta]) = a\sqrt{n}\|\phi[\zeta]\|_2, \quad A_\eta(\phi[\zeta]) = b_n\sqrt{n}\|\phi[\zeta]\|_2, \\ B_\eta(\phi[\zeta]) = c_n\sqrt{n}\|\phi[\zeta]\|_2,$$

then we have for $f \in \mathcal{F}$

$$\varrho_s(\phi[\zeta], f) \leq U_\eta(\phi[\zeta]), \quad \frac{1}{3}\varpi_s^2(\phi[\zeta], f) \leq A_\eta^2(\phi[\zeta]), \\ \frac{4}{3}c_*(s)M_s(\phi[\zeta]) \leq B_\eta(\phi[\zeta]).$$

Thus, in view of Theorem 8, Assumption 1 holds with $U = U_\eta$, $A = A_\eta$, $B = B_\eta$ and $g = G^{(1)}$. Then standard computations show that $\Lambda_{A_\eta} = b_n$ and $\Lambda_{B_\eta} = c_n$. The assertion of the theorem follows now from Corollary 1.

APPENDIX

Proof of Lemma 4. Let

$$\mathcal{X}^{(n)} = \underbrace{\mathcal{X} \times \dots \times \mathcal{X}}_{n\text{-times}}, \quad \bar{\mathcal{X}}^{(n)} = \underbrace{\bar{\mathcal{X}} \times \dots \times \bar{\mathcal{X}}}_{n\text{-times}}.$$

Obviously, $\bar{\mathcal{X}}^{(n)}$ is a countable dense subset of $\mathcal{X}^{(n)}$. For any $x^{(n)} \in \mathcal{X}^{(n)}$ and $t \in \mathcal{T}$, put

$$\xi(t, x^{(n)}) = \sum_{i=1}^n [w(t, x_i) - \mathbb{E}w(t, X)],$$

and let

$$\mathfrak{L} = \left\{ l_{\bar{x}^{(n)}} : \mathcal{T} \rightarrow \mathbb{R} : l_{\bar{x}^{(n)}}(t) = \frac{|\xi(t, \bar{x}^{(n)})|^{s-1} \text{sign}[\xi(t, \bar{x}^{(n)})]}{\|\xi(\cdot, \bar{x}^{(n)})\|_{s,\tau}^{s-1}}, \bar{x}^{(n)} \in \bar{\mathcal{X}}^{(n)} \right\}.$$

Note that \mathcal{L} is countable and $\mathcal{L} \subset \mathbb{B}_{s/(s-1)}$ since, obviously

$$\|l_{\bar{x}^{(n)}}\|_{s/(s-1),\tau} = 1 \quad \forall \bar{x}^{(n)} \in \bar{\mathcal{X}}^{(n)}.$$

Note that $\xi_w(\cdot) = \xi(\cdot, X^{(n)})$, $X^{(n)} = (X_1, \dots, X_n)$, and therefore, in order to prove the assertion of the lemma it is sufficient to show that

$$(A.1) \quad \|\xi(\cdot, x^{(n)})\|_{s,\tau} = \sup_{l \in \mathcal{L}} \int l(t)\xi(t, x^{(n)})\tau(dt) \quad \forall x^{(n)} \in \mathcal{X}^{(n)}.$$

First, let us note that Assumption (A1) implies that for every $\varepsilon > 0$ and every $x^{(n)} \in \mathcal{X}^{(n)}$ there exists $\bar{x}^{(n)} \in \bar{\mathcal{X}}^{(n)}$ such that

$$(A.2) \quad \|\xi(\cdot, x^{(n)}) - \xi(\cdot, \bar{x}^{(n)})\|_{s,\tau} \leq \varepsilon.$$

Taking into account that $\mathcal{L} \subset \mathbb{B}_{s/(s-1)}$ and using the Hölder inequality, we obtain from (A.2) that

$$(A.3) \quad \left| \sup_{l \in \mathcal{L}} \int l(t)\xi(t, x^{(n)})\tau(dt) - \sup_{l \in \mathcal{L}} \int l(t)\xi(t, \bar{x}^{(n)})\tau(dt) \right| \leq \varepsilon.$$

Obviously

$$\|\xi(\cdot, \bar{x}^{(n)})\|_{s,\tau} = \int l_{\bar{x}^{(n)}}(t)\xi(t, \bar{x}^{(n)})\tau(dt).$$

It implies in view of the duality argument that

$$(A.4) \quad \|\xi(\cdot, \bar{x}^{(n)})\|_{s,\tau} = \sup_{l \in \mathcal{L}} \int l(t)\xi(t, \bar{x}^{(n)})\tau(dt).$$

Using the triangle inequality, we obtain from (A.2), (A.3) and (A.4) that for every $\varepsilon > 0$ and every $x^{(n)} \in \mathcal{X}^{(n)}$

$$\left| \|\xi(\cdot, x^{(n)})\|_{s,\tau} - \sup_{l \in \mathcal{L}} \int l(t)\xi(t, x^{(n)})\tau(dt) \right| \leq 2\varepsilon,$$

which completes the proof of (A.1) because $\varepsilon > 0$ can be chosen arbitrary small.

Proof of Lemma 5. First, note that for any $p \geq 1$ and $x \in \mathcal{X}$

$$\begin{aligned} \|\bar{w}(\cdot, x)\|_{p,\tau} &\leq 2^{1-1/p} \left[\int |w(t, x)|^p \tau(dt) + \int \mathbb{E}|w(t, X)|^p \tau(dt) \right]^{1/p} \\ &\leq 2 \sup_{x \in \mathcal{X}} \|w(\cdot, x)\|_{p,\tau}. \end{aligned}$$

Here, we have used the triangle inequality. Next, for any $p \geq 1$ and $t \in \mathcal{T}$,

$$\begin{aligned} &\left[\int |\bar{w}(t, x)|^p f(x)v(dx) \right]^{1/p} \\ &:= [\mathbb{E}|\bar{w}(t, X)|^p]^{1/p} \leq 2[\mathbb{E}|w(t, X)|^p]^{1/p} \\ &=: 2 \left[\int |w(t, x)|^p f(x)v(dx) \right]^{1/p}. \end{aligned}$$

Here, we used that $\mathbb{E}|\eta - \mathbb{E}\eta|^p \leq 2^p \mathbb{E}|\eta|^p$. Combining both inequalities, we have

$$M_p(\bar{w}) \leq 2M_p(w),$$

and the second statement of the lemma is proved.

Proof of Lemma 9. 1^0 . First, we establish statement of the lemma for the mapping $\zeta \mapsto \phi_1[\zeta]$. For any $s \geq 1$ let $\mathfrak{s} := s \wedge 2$. Following (3.24) and (3.25) and in view of (9.3), we have

$$\underline{w}_s^{(1)} \geq k_1(nV_{h\max})^{1/s-1}, \quad \bar{w}_s^{(1)} \leq k_\infty(nV_{h\min})^{1/s-1},$$

$$\mathbb{Z}_s^{(1)}(b) := \{\zeta = (K, h) \in \mathbb{Z}^{(1)} : (nV_h)^{1/s-1} \|K\|_s \leq b\}, \quad b \in [\underline{w}_s^{(1)}, \bar{w}_s^{(1)}].$$

We note that if $\zeta = (K, h) \in \mathbb{Z}_s^{(1)}(b)$ then

$$(A.5) \quad (nV_h)^{1/s-1} \leq k_1^{-1}b.$$

Let $\zeta_1, \zeta_2 \in \mathbb{Z}_s^{(1)}(b)$ be such that $d_1^{(1)}(\zeta_1, \zeta_2) \leq 2$. Applying (9.4) with $p = \mathfrak{s}$ and using (A.5), we get

$$(A.6) \quad n^{1/\mathfrak{s}} \|\phi_1[\zeta_1] - \phi_1[\zeta_2]\|_s \leq k_1^{-1}bD'(2)d_1^{(1)}(\zeta_1, \zeta_2) = bd_{\theta_1}^{(1)}(\zeta_1, \zeta_2).$$

Here we have taken into account that $D'(2) = \sup_{x \in [0, 2]} |D'(x)|$, where the function $D(\cdot)$ is given in (9.1). If $\zeta_1, \zeta_2 \in \mathbb{Z}_s^{(1)}(b)$ are such that $d_1^{(1)}(\zeta_1, \zeta_2) > 2$, then by the triangle inequality

$$(A.7) \quad n^{1/\mathfrak{s}} \|\phi_1[\zeta_1] - \phi_1[\zeta_2]\|_s \leq 2b \leq bd_1^{(1)}(\zeta_1, \zeta_2) \leq bd_{\theta_1}^{(1)}(\zeta_1, \zeta_2).$$

Thus, (A.6) and (A.7) imply that Assumption (L) holds if $\mathbb{Z}^{(1)}$ is equipped with the distance $d_{\theta_1}^{(1)}$, where we recall that $\theta_1 = k_\infty k_1^{-1} D'(2) \geq 1$ [see (9.2)].

2^0 . Now we prove the statement of the lemma for the mapping $z \mapsto \phi_2[z]$. By the statements (i) and (ii) of Lemma 12 applied with $p = \mathfrak{s}$ we have

$$2^{d(1-\mathfrak{s})/\mathfrak{s}} k_1^2 (nV_{h\max})^{1/s-1} \leq \underline{w}_s^{(2)}, \quad \bar{w}_s^{(2)} \leq 2^{d/\mathfrak{s}} k_\infty^2 (nV_{h\min})^{1/s-1}.$$

Recall that

$$\mathbb{Z}_s^{(2)}(b) := \{z = [(K, h), (Q, \mathfrak{h})] \in \mathbb{Z}^{(2)} : n^{1/\mathfrak{s}} \|\phi_2[z]\|_s \leq b\}, \quad b \in [\underline{w}_s^{(2)}, \bar{w}_s^{(2)}].$$

If $z = [(K, h), (Q, \mathfrak{h})] \in \mathbb{Z}_s^{(2)}(b)$ then by the statement (ii) of Lemma 12

$$(A.8) \quad (nV_{h \vee \mathfrak{h}})^{1/s-1} \leq 2^{d(s-1)/\mathfrak{s}} k_1^{-2} b \leq 2^{d/2} k_1^{-2} b.$$

Let $z_1, z_2 \in \mathbb{Z}_s^{(2)}(b)$ be such that $d_1^{(2)}(z_1, z_2) \leq 2$. Applying (9.5) with $p = \mathfrak{s}$ and using (A.8), we obtain

$$(A.9) \quad n^{1/\mathfrak{s}} \|\phi_2[z_1] - \phi_2[z_2]\|_s \leq b2^{2+d/2} k_\infty k_1^{-2} D'(4) d_1^{(2)}(z_1, z_2).$$

If $z_1, z_2 \in \mathbb{Z}_s^{(2)}(b)$ are such that $d_1^{(2)}(z_1, z_2) > 2$, then we have by the triangle inequality

$$(A.10) \quad n^{1/s} \|\phi_2[z_1] - \phi_2[z_2]\|_s \leq 2b \leq b d_1^{(2)}(z_1, z_2).$$

Thus, (A.9) and (A.10) imply that Assumption (L) is valid provided that $\mathbb{Z}^{(2)}$ is equipped with the distance $d_{\theta_2}^{(2)}(\cdot, \cdot)$, where $\theta_2 = 2^{2d+2} k_\infty^4 k_1^{-2} D'(4) \geq 2^{(d+4)/2} \times k_\infty k_1^{-2} D'(4) \geq 1$ [see (9.2)].

Proof of Lemma 10. 1^0 . Inequality (9.3) is immediate. We start with the proof of (9.4).

Since the required bound is symmetric in h and h' , without loss of generality we will assume that $V_h \geq V_{h'}$. By the triangle inequality in view of Assumption (K1), we get

$$(A.11) \quad \begin{aligned} \|K_h - K_{h'}\|_p &\leq \|K_h - K'_h\|_p + \|K'_h - K'_{h'}\|_p \\ &\leq V_h^{-1+1/p} \|K - K'\|_p + \|K'_h - K'_{h'}\|_p \\ &\leq V_h^{-1+1/p} \left[\|K - K'\|_\infty + k_\infty \left(\frac{V_h}{V_{h'}} - 1 \right) \right] \\ &\quad + V_{h'}^{-1} \|K'(\cdot/h) - K'(\cdot/h')\|_p \\ &\leq (V_{h \vee h'})^{-1+1/p} \left[\frac{V_{h \vee h'}}{V_{h \wedge h'}} \right] \left[\|K - K'\|_\infty + k_\infty \left(\frac{V_{h \vee h'}}{V_{h \wedge h'}} - 1 \right) \right] \\ &\quad + (V_{h \vee h'})^{-1+1/p} \left[\frac{V_{h \vee h'}}{V_{h \wedge h'}} \right] \\ &\quad \times \|K'(\cdot[h \vee h']/h) - K'(\cdot[h \vee h']/h')\|_p, \end{aligned}$$

where $h \wedge h' = (h_1 \wedge h'_1, \dots, h_d \wedge h'_d)$. The second term of the last inequality is obtained using the evident change-of-variables $t \mapsto t/[h \vee h']$ (the division is understood in the coordinate-wise sense).

Note that all coordinates of the vectors $[h \vee h']/h$ and $[h \vee h']/h'$ are greater or equal to 1. Therefore, in view of Assumption (K1) the integration (or supremum if $p = \infty$) over the whole \mathbb{R}^d in $\|K'(\cdot[h \vee h']/h) - K'(\cdot[h \vee h']/h')\|_p$ can be replaced by the integration (supremum) over the support of K' . Together with Assumption (K1), this yields

$$(A.12) \quad \begin{aligned} &\|K'(\cdot[h \vee h']/h) - K'(\cdot[h \vee h']/h')\|_p \\ &\leq L_{\mathcal{K}} \sqrt{\frac{1}{4} \sum_{j=1}^d \left[\frac{h_j \vee h'_j}{h_j \wedge h'_j} - 1 \right]^2} \leq 2^{-1} L_{\mathcal{K}} \sqrt{d} (\exp\{\Delta_{\mathcal{H}}(h, h')\} - 1). \end{aligned}$$

Noting that $V_{h \vee h'} / V_{h \wedge h'} \leq \exp\{d \Delta_{\mathcal{H}}(h, h')\}$ we obtain from (A.11) and (A.12) that

$$(A.13) \quad \begin{aligned} & \|K_h - K'_{h'}\|_p \\ & \leq (V_{h \vee h'})^{(1-p)/p} e^{d \Delta_{\mathcal{H}}(h, h')} \left[\|K - K'\|_{\infty} + k_{\infty} (e^{d \Delta_{\mathcal{H}}(h, h')} - 1) \right. \\ & \quad \left. + \frac{L_{\mathcal{K}} \sqrt{d}}{2} (e^{\Delta_{\mathcal{H}}(h, h')} - 1) \right]. \end{aligned}$$

Then (9.4) follows from the last inequality and the monotonicity of the function $D(\cdot)$.

2^0 . Now we turn to the proof of (9.5). Recall that $z = [(K, h), (Q, \mathfrak{h})]$ and $z' = [(K', h'), (Q', \mathfrak{h}')]]$. For brevity, we also write $\zeta_K = (K, h)$ and $\zeta_Q = (Q, \mathfrak{h})$ with evident changes in notation for ζ'_K and ζ'_Q .

By the triangle inequality, we have

$$(A.14) \quad \begin{aligned} \|K_h * Q_{\mathfrak{h}} - K'_{h'} * Q'_{\mathfrak{h}'}\|_p & \leq \|K_h * Q_{\mathfrak{h}} - K_h * Q'_{\mathfrak{h}'}\|_p \\ & \quad + \|K_h * Q'_{\mathfrak{h}'} - K'_{h'} * Q'_{\mathfrak{h}'}\|_p. \end{aligned}$$

Using the Young inequality (the first statement of Lemma 3), Assumption (K1) and (A.13) we obtain

$$\begin{aligned} \|K_h * Q_{\mathfrak{h}} - K_h * Q'_{\mathfrak{h}'}\|_p & \leq \|K_h\|_1 \|Q_{\mathfrak{h}} - Q'_{\mathfrak{h}'}\|_p \\ & \leq k_{\infty} (V_{\mathfrak{h} \vee \mathfrak{h}'})^{-1+1/p} D(d_1^{(1)}(\zeta_Q, \zeta'_Q)). \end{aligned}$$

On the other hand, applying the Young inequality and (A.13) with $p = 1$, we have

$$\begin{aligned} \|K_h * Q_{\mathfrak{h}} - K_h * Q'_{\mathfrak{h}'}\|_p & \leq \|K_h\|_p \|Q_{\mathfrak{h}} - Q'_{\mathfrak{h}'}\|_1 \leq k_{\infty} V_h^{-1+1/p} D(d_1^{(1)}(\zeta_Q, \zeta'_Q)) \\ & \leq k_{\infty} (V_{h \vee h'})^{-1+1/p} \exp\{d \Delta_{\mathcal{H}}(h, h')\} D(d_1^{(1)}(\zeta_Q, \zeta'_Q)) \\ & \leq k_{\infty} (V_{h \vee h'})^{-1+1/p} D(2d_1^{(2)}(z, z')), \end{aligned}$$

where we have used the definition of $\Delta_{\mathcal{H}}(\cdot, \cdot)$ and monotonicity of the function $D(\cdot)$. Combining the last two inequalities, we have

$$\|K_h * Q_{\mathfrak{h}} - K_h * Q'_{\mathfrak{h}'}\|_p \leq k_{\infty} [(V_{h \vee h'}) \vee (V_{\mathfrak{h} \vee \mathfrak{h}'})]^{-1+1/p} D(2d_1^{(2)}(z, z')).$$

Repeating the previous computations, we obtain the same bound for the second term on the right-hand side of (A.14), namely,

$$\|K_h * Q'_{\mathfrak{h}'} - K'_{h'} * Q'_{\mathfrak{h}'}\|_p \leq k_{\infty} [(V_{h \vee h'}) \vee (V_{\mathfrak{h} \vee \mathfrak{h}'})]^{-1+1/p} D(2d_1^{(2)}(z, z')).$$

Thus, we finally get

$$\|K_h * Q_{\mathfrak{h}} - K'_{h'} * Q'_{\mathfrak{h}'}\|_p \leq 2k_{\infty} [(V_{h \vee h'}) \vee (V_{\mathfrak{h} \vee \mathfrak{h}'})]^{-1+1/p} D(2d_1^{(2)}(z, z')),$$

as claimed.

Proof of Lemma 11. If $w \in \mathbb{H}_d(1, P)$, then for any

$$x \in \bigotimes_{i=1}^d \left[\tilde{x}_i - \frac{\|w\|_\infty}{2P\sqrt{d}}, \tilde{x}_i + \frac{\|w\|_\infty}{2P\sqrt{d}} \right]$$

we have by the triangle inequality

$$|w(x)| \geq |w(\tilde{x})| - |w(x) - w(\tilde{x})| \geq \|w\|_\infty - P|x - \tilde{x}| \geq \frac{1}{2}\|w\|_\infty.$$

This completes the proof.

Proof of Lemma 12. Recall that

$$\phi_2[z](t) = (K_h * Q_{\mathfrak{h}})(t) = \int K_h(t - y)Q_{\mathfrak{h}}(y) dy, \quad t \in \mathbb{R}^d.$$

1^0 . Let \mathcal{J} denote the set of indexes $j \in \{1, \dots, d\}$ such that $h_j \leq \mathfrak{h}_j$:

$$\mathcal{J} := \{j \in (1, \dots, d) : h_j \leq \mathfrak{h}_j\}.$$

Given two arbitrary vectors $u, v \in \mathbb{R}^d$, let $\Delta[u, v]$ and $\delta[u, v]$ denote the vectors in \mathbb{R}^d with the coordinates

$$\Delta_j[u, v] = \begin{cases} u_j, & j \in \mathcal{J}, \\ v_j, & j \notin \mathcal{J}, \end{cases} \quad \delta_j[u, v] = \begin{cases} u_j, & j \notin \mathcal{J}, \\ v_j, & j \in \mathcal{J}. \end{cases}$$

With this notation, we can write

$$(K_h * Q_{\mathfrak{h}})(t) = \frac{1}{V_h V_{\mathfrak{h}}} \int K \left(\Delta \left[\frac{t-v}{h}, \frac{v}{\mathfrak{h}} \right] \right) Q \left(\delta \left[\frac{t-v}{\mathfrak{h}}, \frac{v}{\mathfrak{h}} \right] \right) dv, \quad t \in \mathbb{R}^d.$$

Then changing the variables $v \mapsto u = (t - v)/(h \wedge \mathfrak{h})$ and setting for brevity $\eta = (h \wedge \mathfrak{h})/(h \vee \mathfrak{h})$ (as usual, all operations are understood in the coordinate-wise sense), we come to the formula

$$\begin{aligned} & (K_h * Q_{\mathfrak{h}})(t) \\ \text{(A.15)} \quad &= \frac{V_{h \wedge \mathfrak{h}}}{V_h V_{\mathfrak{h}}} \int K(\Delta[u, t/(h \vee \mathfrak{h}) - \eta u]) Q(\delta[t/(h \vee \mathfrak{h}) - \eta u, u]) du \\ &= \frac{1}{V_{h \vee \mathfrak{h}}} F \left(\frac{t}{h \vee \mathfrak{h}} \right), \end{aligned}$$

where we have denoted

$$\text{(A.16)} \quad F(t) := \int K(\Delta[u, t - \eta u]) Q(\delta[t - \eta u, u]) du, \quad t \in \mathbb{R}^d.$$

Now we note some properties of the function F that will be useful in the sequel. First, Assumption **(K1)** implies that the integration over \mathbb{R}^d in **(A.16)** can be replaced by the integration over $[-1/2, 1/2]^d$. Indeed, if at least one of the coordinates of u lies outside the interval $[-1/2, 1/2]$ then, in view of **(K1)**, one of

the functions K or Q vanishes. This fact along with Assumption (K2) and (A.16) imply that $\|F\|_\infty \leq k_\infty^2$; in addition,

$$(A.17) \quad \text{supp}(F) \subseteq [-1, 1]^d.$$

Taking into account these facts and using (A.15), we obtain

$$\|K_h * Q_h\|_p \leq (V_{h \vee h})^{-1+1/p} \|F\|_p \leq 2^{d/p} k_\infty^2 (V_{h \vee h})^{-1+1/p},$$

and the statement (i) of the lemma is proved.

To get the assertion (ii) of the lemma, we note that

$$\begin{aligned} \left| \int F(t) dt \right| &= \left| \iint K(\Delta[u, t - \eta u]) Q(\delta[t - \eta u, u]) du dt \right| \\ &= \left| \int K(x) dx \right| \left| \int Q(x) dx \right| \geq k_1^2. \end{aligned}$$

The second equality follows from the fact that functions K and Q are integrated over t and over u over disjoint sets of components; and the last inequality is a consequence of (K2). Therefore, invoking (A.17) we have

$$\begin{aligned} \|G\|_p &= (V_{h \vee h})^{-1+1/p} \|F\|_p \geq (2^d V_{h \vee h})^{-1+1/p} \|F\|_1 \\ &\geq 2^{d(1-p)/p} k_1^2 (V_{h \vee h})^{-1+1/p}, \end{aligned}$$

as claimed in the statement (ii) of the lemma.

2⁰. Now we turn to the proof of the statements (iii) and (iv) of the lemma. The idea in the proof of these statements is to show that F satisfies the Lipschitz condition and then to apply Lemma 11.

By (A.16) for any $x, y \in \mathbb{R}^d$, we have

$$\begin{aligned} |F(x) - F(y)| &\leq k_\infty \sup_{u \in [-1/2, 1/2]^d} |K(\Delta[u, x - \eta u]) - K(\Delta[u, y - \eta u])| \\ &\quad + k_\infty \sup_{u \in [-1/2, 1/2]^d} |Q(\delta[x - \eta u, u]) - Q(\delta[y - \eta u, u])| \\ (A.18) \quad &\leq L_{\mathcal{K}} k_\infty \left\{ \sqrt{\sum_{j \notin \mathcal{J}} (x_j - y_j)^2} + \sqrt{\sum_{j \in \mathcal{J}} (x_j - y_j)^2} \right\} \\ &\leq 2L_{\mathcal{K}} k_\infty |x - y|. \end{aligned}$$

The obtained inequality means that $F \in \mathbb{H}_d(1, P)$ with $P = 2L_{\mathcal{K}} k_\infty$; moreover, (A.17) implies that

$$(A.19) \quad \|F\|_\infty \geq 2^{-d} k_1^2.$$

Applying Lemma 11 and using (A.18), we obtain

$$\left\{ x \in \mathbb{R}^d : F(x) \geq \frac{1}{2} \|F\|_\infty \right\} \supseteq \bigotimes_{i=1}^d \left[\tilde{x}_i - \frac{\|F\|_\infty}{2P\sqrt{d}}, \tilde{x}_i + \frac{\|F\|_\infty}{2P\sqrt{d}} \right],$$

where, recall, $F(\tilde{x}) = \|F\|_\infty$. Using (A.19), we obviously deduce from (A.15) that

$$\left\{x : (K_h * Q_h)(x) \geq \frac{1}{2} \|K_h * Q_h\|_\infty\right\} \\ \supseteq \bigotimes_{i=1}^d \left[\tilde{x}_i(h \vee h)_i - \frac{k_1^2(h \vee h)_i}{2^{d+1} P \sqrt{d}}, \tilde{x}_i(h \vee h)_i + \frac{k_1^2(h \vee h)_i}{2^{d+1} P \sqrt{d}} \right],$$

which implies that

$$(A.20) \quad \text{mes} \left\{x : (K_h * Q_h)(x) \geq \frac{1}{2} \|K_h * Q_h\|_\infty\right\} \geq V_{h \vee h} \left[\frac{k_1^2}{2^{d+1} \sqrt{d} L_{\mathcal{K}} k_\infty} \right]^d.$$

Then the statement (iii) of the lemma follows because

$$\text{mes}\{\text{supp}(K_h * Q_h)\} \geq \text{mes} \left\{x : (K_h * Q_h)(x) \geq \frac{1}{2} \|K_h * Q_h\|_\infty\right\}.$$

It remains to note that (A.17) implies that $\text{mes}\{\text{supp}(K_h * Q_h)\} \leq 2^d V_{h \vee h}$. Therefore by (A.20),

$$\text{mes} \left\{x : (K_h * Q_h)(x) \geq \frac{1}{2} \|K_h * Q_h\|_\infty\right\} \\ \geq \left[\frac{k_1^2}{2^{d+2} \sqrt{d} L_{\mathcal{K}} k_\infty} \right]^d \text{mes}\{\text{supp}(K_h * Q_h)\}.$$

This completes the proof of the lemma.

Acknowledgments. We thank two anonymous referees for useful comments that led to significant improvements in the presentation.

REFERENCES

- ALEXANDER, K. S. (1984). Probability inequalities for empirical processes and a law of the iterated logarithm. *Ann. Probab.* **12** 1041–1067. [MR0757769](#)
- BARRON, A., BIRGÉ, L. and MASSART, P. (1999). Risk bounds for model selection via penalization. *Probab. Theory Related Fields* **113** 301–413. [MR1679028](#)
- BEIRLANT, J. and MASON, D. M. (1995). On the asymptotic normality of L_p -norms of empirical functionals. *Math. Methods Statist.* **4** 1–19. [MR1324687](#)
- BOUSQUET, O. (2002). A Bennett concentration inequality and its application to suprema of empirical processes. *C. R. Math. Acad. Sci. Paris* **334** 495–500. [MR1890640](#)
- CAVALIER, L. and GOLUBEV, Y. (2006). Risk hull method and regularization by projections of ill-posed inverse problems. *Ann. Statist.* **34** 1653–1677. [MR2283712](#)
- DEVROYE, L. and LUGOSI, G. (2001). *Combinatorial Methods in Density Estimation*. Springer, New York. [MR1843146](#)
- DÜMBGEN, L. and FATALOV, V. R. (2002). Asymptotics of the rate of convergence for nonparametric density estimators: A new approach based on the Laplace method. *Math. Methods Statist.* **11** 465–476 (2003). [MR1979745](#)
- EINMAHL, U. and MASON, D. M. (2000). An empirical process approach to the uniform consistency of kernel-type function estimators. *J. Theoret. Probab.* **13** 1–37. [MR1744994](#)

- FOLLAND, G. B. (1999). *Real Analysis: Modern Techniques and Their Applications*, 2nd ed. Wiley, New York. [MR1681462](#)
- GINÉ, E., KOLTCHINSKII, V. and ZINN, J. (2004). Weighted uniform consistency of kernel density estimators. *Ann. Probab.* **32** 2570–2605. [MR2078551](#)
- GINÉ, E. and KOLTCHINSKII, V. (2006). Concentration inequalities and asymptotic results for ratio type empirical processes. *Ann. Probab.* **34** 1143–1216. [MR2243881](#)
- GINÉ, E., MASON, D. M. and ZAITSEV, A. Y. (2003). The L_1 -norm density estimator process. *Ann. Probab.* **31** 719–768. [MR1964947](#)
- GINÉ, E. and NICKL, R. (2008). Uniform central limit theorems for kernel density estimators. *Probab. Theory Related Fields* **141** 333–387. [MR2391158](#)
- GINÉ, E. and ZINN, J. (1984). Some limit theorems for empirical processes. *Ann. Probab.* **12** 929–998. [MR0757767](#)
- GOLDENSHLUGER, A. and LEPSKI, O. (2008). Universal pointwise selection rule in multivariate function estimation. *Bernoulli* **14** 1150–1190. [MR2543590](#)
- GOLDENSHLUGER, A. and LEPSKI, O. (2009). Structural adaptation via L_p -norm oracle inequalities. *Probab. Theory Related Fields* **143** 41–71. [MR2449122](#)
- GOLUBEV, Y. and SPOKOINY, V. (2009). Exponential bounds for minimum contrast estimators. *Electron. J. Statist.* **3** 712–746. [MR2534199](#)
- JOHNSON, W. B., SCHECHTMAN, G. and ZINN, J. (1985). Best constants in moment inequalities for linear combinations of independent and exchangeable random variables. *Ann. Probab.* **13** 234–253. [MR0770640](#)
- LEDOUX, M. and TALAGRAND, M. (1991). *Probability in Banach Spaces: Isoperimetry and Processes*. Springer, Berlin. [MR1102015](#)
- MASSART, P. (2000). About the constants in Talagrand’s concentration inequalities for empirical processes. *Ann. Probab.* **28** 863–884. [MR1782276](#)
- PINELIS, I. F. (1990). Inequalities for distributions of the sums of independent random vectors and their application to the estimation of a density. *Theory Probab. Appl.* **35** 605–607. [MR1091221](#)
- PINELIS, I. (1994). Optimum bounds for the distributions of martingales in Banach spaces. *Ann. Probab.* **22** 1679–1706. [MR1331198](#)
- TALAGRAND, M. (1994). Sharper bounds for Gaussian and empirical processes. *Ann. Probab.* **22** 28–76. [MR1258865](#)
- VAN DE GEER, S. A. (2000). *Applications of Empirical Process Theory*. *Cambridge Series in Statistical and Probabilistic Mathematics* **6**. Cambridge Univ. Press, Cambridge. [MR1739079](#)
- VAN DER VAART, A. W. and WELLNER, J. A. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics*. Springer, New York. [MR1385671](#)
- VON BAHR, B. and ESSEEN, C.-G. (1965). Inequalities for the r th absolute moment of a sum of random variables, $1 \leq r \leq 2$. *Ann. Math. Statist.* **36** 299–303. [MR0170407](#)

DEPARTMENT OF STATISTICS
UNIVERSITY OF HAIFA
HAIFA 31905
ISRAEL
E-MAIL: goldensh@stat.haifa.ac.il

LABORATOIRE D’ANALYSE, TOPOLOGIE
AND PROBABILITÉS
UNIVERSITÉ AIX-MARSEILLE I
39, RUE F. JOLIOT-CURIE
13453 MARSEILLE
FRANCE
E-MAIL: lepski@cmi.univ-mrs.fr