Upper Functions for Positive Random Functionals. II. Application to the Empirical Processes Theory, Part 1

O Lepski

Upper functions for positive random functionals

Oleg Lepski

Laboratoire d'Analyse, Topologie, Probabilités
Université Aix-Marseille
39, rue F. Joliot-Curie
13453 Marseille, France
e-mail: lepski@cmi.univ-mrs.fr

Abstract: In this paper we are interested in finding upper functions for a collection real-valued random variables $\{\Psi(\chi_\theta), \theta \in \Theta\}$. Here $\{\chi_\theta, \theta \in \Theta\}$ is a family of continuous random mappings, Ψ is a given sub-additive positive functional and Θ is a totally bounded subset of a metric space. We seek a non-random function $U : \Theta \to \mathbb{R}_+$ such that $\sup_{\theta\in\Theta} \{\Psi(\chi_\theta) - U(\theta)\}_+$ is "small" with prescribed probability. We apply the results obtained in the general setting to the variety of problems related to gaussian random functions and empirical processes.

AMS 2000 subject classifications: Primary 60E15; secondary 62G07, 62G08.
Keywords and phrases: upper function, empirical processes, gaussian random function, metric entropy, doubling measure.

Contents

1. Introduction

The main objective of this paper is to look from a novel point of view at some phenomena arising in different areas of probability theory and mathematical statistics. We will try to understand what is common between classical probabilistic results, such as the law of iterated logarithm for example, and well-known problem in adaptive estimation called price to pay for adaptation. Why do two different kinds of this price exist? What relates exponential inequalities for M-estimators, so-called uniform-in-bandwidth consistency in density or regression model and the bounds for modulus of continuity of gaussian random functions defined on a metric space equipped with doubling measure?

It turned out that all these and many other problems can be reduced to the following one. Let $\mathfrak{T}$ be a set and let $(\Omega, \mathfrak{B}, \mathrm{P})$ be a complete probability space. Let χ defined on $\mathfrak{T} \times \Omega$ be a given $\mathfrak{B}$-measurable map into a linear metric space $\mathfrak{S}$ and let $\Psi : \mathfrak{S} \to \mathbb{R}_+$ be a given *continuous sub-additive* functional.
Let $\Theta \subset \mathfrak{T}$ and suppose that $\forall \theta \in \Theta$ and $\forall z > 0$ one can find non-random $U(\theta, z) > 0$ and $c > 0$ such that

$$\mathrm{P}\left\{[\Psi(\chi_\theta) - U(\theta, z)] > 0\right\} \le c e^{-z}. \tag{1.1}$$

Assuming additionally that $\lambda U(\cdot, z) \ge U(\cdot, \lambda z)$ for any $z > 0,\ \lambda \ge 1$ we also have for any $q \ge 1$

$$\mathrm{E}\left\{[\Psi(\chi_\theta) - U(\theta, z)]_+\right\}^q \le c\Gamma(q+1)\big[U(\theta, 1)\big]^q e^{-z}, \quad \forall z \ge 1, \tag{1.2}$$

where Γ is gamma-function and $[a]_+$ is the positive part of a.

The problem which we address now consists in a finding of $\mathbf{U}(\theta, z)$ and $\boldsymbol{U}_q(\theta, z)$ satisfying

$$\mathrm{P}\left\{\sup_{\theta\in\Theta}[\Psi(\chi_\theta) - \mathbf{U}(\theta, z)] > 0\right\} \le \mathbf{c} e^{-z}; \quad \forall z \ge 1 \tag{1.3}$$

$$\mathrm{E}\left\{\sup_{\theta\in\Theta}[\Psi(\chi_\theta) - \boldsymbol{U}_q(\theta, z)]_+\right\}^q \le \mathbf{c}_q \left[\inf_{\theta\in\Theta} \boldsymbol{U}_q(\theta, 1)\right]^q e^{-z}, \quad \forall z \ge 1, \tag{1.4}$$

where $\mathbf{c}$ and $\mathbf{c}_q$ are constants. If (1.3) and (1.4) hold we will say that $\mathbf{U}(\cdot, \cdot)$ and $\boldsymbol{U}_q(\cdot, \cdot)$ are *upper functions* for the collection of random variables $\{\Psi(\chi_\theta),\ \theta \in \Theta\}$.

The main questions which we would like to answer are the following.

- Do $\mathbf{U}(\cdot, \cdot)$ and $\boldsymbol{U}_q(\cdot, \cdot)$ coincide with $U(\cdot, \cdot)$ up to numerical constants or there is a "price to pay" for passing from pointwise results (1.1)–(1.2) to uniform ones given in (1.3)–(1.4)?

- Do $\mathbf{U}(\cdot,\cdot)$ and $\boldsymbol{U}_q(\cdot,\cdot)$ coincide up to numerical constants? In other words should one to pay the same price for the probability and moment's bounds?

We will show that a payment exists and in general $\boldsymbol{U}_q(\cdot,\cdot) \gg \mathbf{U}(\cdot,\cdot) \gg U(\cdot,\cdot)$. Thus, we will seek $\mathbf{U}(\cdot,\cdot)$ and $\boldsymbol{U}_q(\cdot,\cdot)$ satisfying (1.3) and (1.4) and "minimally" separated away from $U(\cdot,\cdot)$. We will realize this program under the following condition.

Assumption 1. *1. There exist* $A:\mathfrak{T}\to\mathbb{R}_+$, $B:\mathfrak{T}\to\mathbb{R}_+$ *and* $\mathrm{c}>0$ *such that* $\forall z>0$

$$\mathrm{P}\left\{\Psi(\chi_\theta)\geq z\right\}\leq \mathrm{c}\exp\left\{-\frac{z^2}{A^2(\theta)+B(\theta)z}\right\},\ \ \forall\theta\in\Theta. \tag{1.5}$$

2. There exist $\mathrm{a}:\mathfrak{T}\times\mathfrak{T}\to\mathbb{R}_+$ *and* $\mathrm{b}:\mathfrak{T}\times\mathfrak{T}\to\mathbb{R}_+$ *such that* $\forall z>0$

$$\mathrm{P}\left\{\Psi(\chi_{\theta_1}-\chi_{\theta_2})\geq z\right\}\leq \mathrm{c}\exp\left\{-\frac{z^2}{\mathrm{a}^2(\theta_1,\theta_2)+\mathrm{b}(\theta_1,\theta_2)z}\right\},\ \ \forall\theta_1,\theta_2\in\Theta. \tag{1.6}$$

Remark 1. *If Assumption 1 (1) holds on* $\mathfrak{T}$ *(not only on* Θ*),* $\mathfrak{T}$ *is linear space and if, additionally, the map* χ_t *is linear on* $\mathfrak{T}$ *, then the Assumption 1 (2) is automatically fulfilled since one can take* $\mathrm{a}(\mathfrak{t}_1,\mathfrak{t}_2)=A(\mathfrak{t}_1-\mathfrak{t}_2)$ *and* $\mathrm{b}(\mathfrak{t}_1,\mathfrak{t}_2)=B(\mathfrak{t}_1-\mathfrak{t}_2),\ \mathfrak{t}_1,\mathfrak{t}_2\in\mathfrak{T}$.

Remark 2. *We can easily deduce from (1.5) that for any* $\theta\in\Theta$

$$\mathrm{P}\Big\{\Psi(\chi_\theta)\geq A(\theta)\sqrt{z}+B(\theta)z\Big\}\leq \mathrm{c}\exp\left\{-z\right\},\ \ \forall z\geq 0; \tag{1.7}$$

$$\mathrm{E}\Big\{\Psi(\chi_\theta)-\Big[A(\theta)\sqrt{z}+B(\theta)z\Big]\Big\}_+^q\leq \mathrm{c}\Gamma(q+1)\Big[A(\theta)+B(\theta)\Big]^q\exp\left\{-z\right\},\ \ \forall z\geq 1. \tag{1.8}$$

Therefore, (1.1)–(1.2) hold with $U(\theta,z)=A(\theta)\sqrt{z}+B(\theta)z$.

Assumption 1 is not new. In particular, it can be found in slightly different form in van der Vaart and Wellner (1996), Talagrand (2005), where this assumption is used for deriving the bound for $\mathrm{E}\left[\sup_{\theta\in\Theta}\Psi(\chi_\theta)\right]$. The usual technique is based on the *chaining argument* available in view of (1.6). It is worth mentioning that uniform probability and moment bounds for $[\sup_{\theta\in\Theta}\Psi(\chi_\theta)]$ in the case where χ_θ is empirical or gaussian process are a subject of vast literature, see, e.g., Alexander (1984), Talagrand (1994), Lifshits (1995), van der Vaart and Wellner (1996), van de Geer (2000), Massart (2000), Bousquet (2002), Giné and Koltchinskii (2006) among many others. Such bounds play an important role in establishing the laws of iterative logarithm and central limit theorems [see, e.g., Alexander (1984) and Giné and Zinn (1984)].

However much less attention was paid to finding of upper functions. The majority of the papers, where such problems are considered, contains asymptotical results, see, i.e. Kalinauskaĭte (1966), Qualls and Watanabe (1972), Bobkov (1988), Shiryaev et al. (2002) and references therein. We would like especially mention the paper Egishyants and Ostrovskii (1996), where upper function satisfying the inequalities similar to (1.3), was obtained for the modulus of continuity of random fields satisfying the Cramer condition.

The researches carried out in the present paper complete the investigations done in Goldenshluger and Lepski (2011), where the upper functions as well as inequalities (1.3)–(1.4) were obtained under following condition: χ_t is linear and there are $A:\mathfrak{T}\to\mathbb{R}_+,\ B:\mathfrak{T}\to\mathbb{R}_+,\ V:\mathfrak{T}\to\mathbb{R}_+$ such that

$$\mathrm{P}\left\{\Psi(\chi_\mathfrak{t})-V(\mathfrak{t})\geq z\right\}\leq g\bigg(\frac{z^2}{A^2(\mathfrak{t})+B(\mathfrak{t})z}\bigg),\quad \forall\mathfrak{t}\in\mathfrak{T},$$

where $g : \mathbb{R}_+ \to \mathbb{R}_+$ is a strictly decreasing to zero function. We note that if $g(x) = e^{-x}$ and $V \equiv 0$ this assumption coincides with (1.5) and, since χ_t is linear (1.6) is automatically fulfilled, see Remark 1. In Goldenshluger and Lepski (2011) under additional assumption imposed on A, B, V and $\Theta \subset \mathfrak{T}$ the upper functions for the collection $\{\Psi(\chi_\theta),\ \theta \in \Theta\}$ were found. As it was shown that they coincide with the function V up to universal constants! The imposed assumptions do not admit the case $V \equiv 0$ that, as it was said above, leads to some "price to pay" for passing from pointwise results (1.1)–(1.2) to uniform ones given in (1.3)–(1.4).

To derive upper functions satisfying (1.3)–(1.4) we complete Assumption 1 by the following conditions.

Assumption 2. $\chi_\bullet : \mathfrak{T} \to \mathfrak{S}$ *is continuous* P*-a.s.*

Mappings a *and* b *are semi-metrics on* $\mathfrak{T}$ *and* Θ *is totally bounded with respect to* $\mathrm{a} \vee \mathrm{b}$.

$\overline{A}_\Theta := \sup_{\theta\in\Theta} A(\theta) < \infty,\ \overline{B}_\Theta := \sup_{\theta\in\Theta} B(\theta) < \infty$.

Denote by $\mathbb{S}$ the following set of real functions:

$$\mathbb{S} = \left\{ s : \mathbb{R} \to \mathbb{R}_+ \setminus \{0\} :\ \sum_{k=0}^{\infty} s\big(2^{k/2}\big) \le 1 \right\}.$$

For any $\widetilde{\Theta} \subseteq \Theta$ and any semi-metric d on $\mathfrak{T}$ let $\mathfrak{E}_{\widetilde{\Theta},\,\mathrm{d}}(\delta),\ \delta > 0$, denote the entropy of $\widetilde{\Theta}$ measured in d. For any $x > 0$, $\widetilde{\Theta} \subseteq \Theta$ and $s \in \mathbb{S}$ define the quantities

$$e_s^{(\mathrm{a})}\big(x, \widetilde{\Theta}\big) = \sup_{\delta>0} \delta^{-2} \mathfrak{E}_{\widetilde{\Theta},\,\mathrm{a}}\left(x(48\delta)^{-1} s(\delta)\right), \quad e_s^{(\mathrm{b})}\big(x, \widetilde{\Theta}\big) = \sup_{\delta>0} \delta^{-1} \mathfrak{E}_{\widetilde{\Theta},\,\mathrm{b}}\left(x(48\delta)^{-1} s(\delta)\right). \tag{1.9}$$

Assumption 3. *There exist* $s_1, s_2 \in \mathbb{S}$ *such that* $\forall x > 0$

$$e_{s_1}^{(\mathrm{a})}\big(x, \Theta\big) < \infty, \qquad e_{s_2}^{(\mathrm{b})}\big(x, \Theta\big) < \infty.$$

What is this paper about? In the next section we construct upper functions for $\{\Psi(\chi_\theta),\ \theta \in \Theta\}$ and prove for them the inequalities (1.3)–(1.4) under Assumptions 1–3. We show that they are completely determined by the functions A and B and by the entropies of their level sets measured in semi-metrics a and b. We will see that obtained upper functions do not coincide with $U(\theta, z) = A(\theta)\sqrt{z} + B(\theta)z$, see Remark 2, and provide with explicit expression for the "price to be paid for uniformity". In particular, if $\underline{A} := \inf_{\theta\in\Theta} A(\theta) > 0$ and $\underline{B} := \inf_{\theta\in\Theta} B(\theta) > 0$, we prove that this "price" can be expressed as a given function of $A(\theta)/\underline{A}$ and $B(\theta)/\underline{B}$.

In this context it is interesting to compare our results with the usual probability bounds for $\sup_{\theta\in\Theta} \Psi(\chi_\theta)$ above $\mathrm{E}\{\sup_{\theta\in\Theta} \Psi(\chi_\theta)\}$ obtained from Talagrand's or the Borel-Sudakov-Tsirelson inequality (when available), combined with uniform bounds for $\mathrm{E}\{\sup_{\theta\in\Theta} \Psi(\chi_\theta)\}$ proved in Talagrand (2005), Theorem 1.2.7, under condition close to our Assumption 1. Following this strategy we will come to upper functions which are constant in θ. The main question is then what can one gain using the technique developed in the paper with respect to the aforementioned approach? We do not think that the answer can be done under "abstract considerations", i.e. under Assumptions 1-3 since it would require to prove that the found "price to be paid for uniformity" is minimal. However, in concrete examples it seems to be possible, in particular for some problems studied in mathematical statistics. Let us mention some of them. First, we note that upper functions are used in all known constructions of adaptive procedures. Next, the use of upper functions being constant in θ will lead to adaptive estimators which are not optimal (remind, that the adaptive estimation theory is equipped with very developed criterion of optimality). Contrary to this, in all known to

the author examples the use of upper functions found in Propositions 2 and 3 allows to construct optimal adaptive procedures (for more details see discussion after Proposition 2). Also, we would like to emphasize that upper functions being constant in θ and the probability bounds related to them are similar to the construction and the results described in Proposition 1, which is, in its turn, the initial step for our considerations. This step as well as the Talagrand's bounds are obtained from chaining argument under, for instance, Assumption 1. In some sense one of our goals is to show that the use of concentration inequalities (which cannot be guaranteed by a condition similar to Assumption 1) in the construction of upper functions is not necessary. In particular, in Section 4.1 we derive an upper function for the $\mathbb{L}_p$-norms of Wiener integrals and deduce the corresponding probability bounds directly from Proposition 3 without passing to the concentration inequalities.

As it was mentioned above upper functions for random objects appear in various areas of mathematical statistics. To apply them in the construction of statistical procedures they have to be computed explicitly. In particular the study of the adaptive estimation in the density model requires to find upper functions for the empirical processes of different kind. For the majority of existed problems Assumption 1 follows from the Berstein's inequality. However the application of Propositions 2-3 requires to compute the functions $\mathcal{E}$ or $\widehat{\mathcal{E}}$ (involved in the description of upper functions) and there is no a general recipe how to do it. One of our main objectives is to provide with rather general assumptions under which the latter quantities can be computed explicitly. In particular, we provide with Assumption 4 and to the best of our knowledge the assumptions of such kind have not been appeared in the existing literature. Under this assumption the upper functions are found for the variety of particular problems.

Organization of the paper In Section 2 we construct upper functions for $\{\Psi(\chi_\theta),\ \theta\in\Theta\}$ and prove for them the inequalities (1.3)–(1.4) under Assumptions 1–3. In fact we present two different constructions which will be referred to upper functions of the first and second type (Propositions 2 and 3). We also derive some consequences related to the upper functions for modulus of continuity of random real-valued mappings (Propositions 4 and 5). Section 3 is devoted to the detailed consideration of generalized empirical processes. We provide with rather general assumption (Assumption 4) under which the upper functions admit the explicit expression, Section 3.2 (Theorem 1) and Section 3.3 (Theorem 3). We also establish non-asymptotical versions of the law of iterated logarithm (Theorem 2) and the law of logarithm (Theorem 4). Section 3.4 is devoted to the application of Theorems 1 and 3 to empirical processes possessing some special structure, Theorems 5–8. In Section 4 we apply Propositions 3 and 4 to gaussian random functions. In Section 4.1 we derive upper functions for $\mathbb{L}_p$-norm of some Wiener integrals (Theorem 9) and in Section 4.2 we study the local modulus of continuity of gaussian functions defined on a metric space satisfying doubling condition (Theorem 10). Proofs of main results are given in Sections 5–7 and technical lemmas are proven in Appendix.

2. General setting

Denote by $\mathcal{S}_{\mathrm{a,b}}$ the subset of $\mathbb{S}\times\mathbb{S}$ for which Assumption 3 holds and let A, B, a and b be any mappings for which Assumption 1 is fulfilled.

For any $\vec{s}=(s_1,s_2)\in\mathcal{S}_{\mathrm{a,b}}$, any $\varkappa=(\varkappa_1,\varkappa_2),\ \varkappa_1>0, \varkappa_2>0$, and any $\widetilde{\Theta}\subseteq\Theta$ put

$$e_{\vec{s}}\big(\varkappa,\widetilde{\Theta}\big)=e^{(\mathrm{a})}_{s_1}\big(\varkappa_1,\widetilde{\Theta}\big)+e^{(\mathrm{b})}_{s_2}\big(\varkappa_2,\widetilde{\Theta}\big). \tag{2.1}$$

2.1. Inequalities for the suprema

Put for any $\widetilde{\Theta} \subseteq \Theta$, any $\varepsilon > 0$ and any $y \geq 0$

$$U_{\vec{s}}^{(\varepsilon)}\big(y, \varkappa, \widetilde{\Theta}\big) = \varkappa_1 \sqrt{2\big[1+\varepsilon^{-1}\big]^2 e_{\vec{s}}\big(\varkappa, \widetilde{\Theta}\big) + y} + \varkappa_2 \Big(2\big[1+\varepsilon^{-1}\big]^2 e_{\vec{s}}\big(\varkappa, \widetilde{\Theta}\big) + y\Big).$$

Proposition 1. *Let Assumptions 1-3 hold and let $\widetilde{\Theta} \subseteq \Theta$ be fixed. Then for any $\widetilde{\varkappa}$ satisfying $\widetilde{\varkappa}_1 \geq \sup_{\theta\in\widetilde{\Theta}} A(\theta)$ and $\widetilde{\varkappa}_2 \geq \sup_{\theta\in\widetilde{\Theta}} B(\theta)$, any $\vec{s} \in \mathcal{S}_{\mathrm{a,b}}$, $\varepsilon \in \big(0, \sqrt{2}-1\big]$ and $y \geq 1$,*

$$\mathrm{P}\left\{\sup_{\theta\in\widetilde{\Theta}} \Psi\left(\chi_\theta\right) \geq U_{\vec{s}}^{(\varepsilon)}\big(y, \widetilde{\varkappa}, \widetilde{\Theta}\big)\right\} \leq 2\mathrm{c}\exp\big\{-y/(1+\varepsilon)^2\big\}.$$

Moreover, for any $q \geq 1$

$$\mathrm{E}\left\{\sup_{\theta\in\widetilde{\Theta}} \Psi\left(\chi_\theta\right) - U_{\vec{s}}^{(\varepsilon)}\big(y, \widetilde{\varkappa}, \widetilde{\Theta}\big)\right\}_+^q \leq 2\mathrm{c}\Gamma(q+1)\left[(1+\varepsilon)^2 y^{-1} U_{\vec{s}}^{(\varepsilon)}\big(y, \widetilde{\varkappa}, \widetilde{\Theta}\big)\right]^q \exp\big\{-y/(1+\varepsilon)^2\big\}.$$

We remark that $\sup_{\theta\in\widetilde{\Theta}} \Psi\left(\chi_\theta\right)$ is $\mathfrak{B}$-measurable for any $\widetilde{\Theta} \subseteq \Theta$ since Ψ is continuous, the mapping $\theta \mapsto \chi_\theta$ is continuous P-a.s., Θ is a totally bounded set and considered probability space is complete (see, e.g. Lemma 1 below).

Discussion We will see that the Proposition 1 is crucial technical tool for deriving upper functions. It contains the main ingredient of our future construction the quantity $e_{\vec{s}}$. The important issue in this context is the choice of $\vec{s} \in \mathcal{S}_{\mathrm{a,b}}$. For many particular problems it is sufficient to choose $\vec{s} = (s^*, s^*)$, where

$$s^*(x) = (6/\pi^2)\big(1 + [\ln x]^2\big)^{-1},\ x \geq 0. \tag{2.2}$$

This choice is explained by two simple reasons: its explicit description allowing to compute the quantity $e_{\vec{s}}$ in particular problem and the logarithmical decay of this function when $x \to \infty$. In view of the latter remark we can consider the set Θ those entropy obeys the restriction which is closer to the minimal one (c.f. Sudakov lower bound for gaussian random functions Lifshits (1995)). We note, however, that there exist examples where $\vec{s}$ has to be chosen on a more special way (see Theorem 9).

Let us now discuss the role of the parameter ε. In most particular problems considered in the paper we will not be interested in optimization of the numerical constants involved in the description of upper functions. If so, the choice of this parameter can be done in arbitrary way and we will put $\varepsilon = \sqrt{2}-1$ to simplify the notations and computations. Note, however, that there are some problems (see, for instance Section 4.2), where ε must be chosen carefully. The typical requirements to this choice is $\varepsilon = \varepsilon(y)$ and

$$\varepsilon(y) \to 0, \quad y\varepsilon^2(y) \to 0,\ \ y \to \infty.$$

The bounds similar to those presented in Proposition 1 are the subject of vast literature see, for instance, the books Lifshits (1995), van der Vaart and Wellner (1996) or van de Geer (2000). Note, however, that the results presented in the proposition may have an independent interest, at least, for the problems where the quantity $e_{\vec{s}}$ can be expressed explicitly. In this case under rather general conditions it is possible, putting $\widetilde{\Theta} = \Theta$ and $\widetilde{\varkappa} = \big(\overline{A}_\Theta, \overline{B}_\Theta\big)$, to compute the tail probability as well as the expected value of the suprema of random mappings. Note also that Assumptions 1-3 guarantee that $\mathrm{E}\left\{\sup_{\theta\in\Theta} \Psi\left(\chi_\theta\right)\right\}^q$ is finite for any $q \geq 1$.

2.2. Upper functions of the first and second type

We will now use Proposition 1 in order to derive the upper functions for $\Psi(\chi_\theta)$ on Θ. Denote $\underline{A} = \inf_{\theta\in\Theta} A(\theta)$ and $\underline{B} = \inf_{\theta\in\Theta} B(\theta)$.

We present two kinds of upper functions for $\Psi(\chi_\theta)$ on Θ which we will refer to upper functions of the first and second type. The first construction is completely determined by the functions A, B and by the semi-metrics a and b. It requires however the additional condition $\underline{A} > 0, \underline{B} > 0$. We will use corresponding results for the particular problems studied in Section 3.

The second construction is related to some special structure imposed on the set Θ. Namely we will suppose that $\Theta = \cup_\alpha \Theta_{\alpha\in\mathfrak{A}}$, where $\{\Theta_\alpha,\ \alpha\in\mathfrak{A}\}$ is a given collection of sets. Here we will be interested in a finding of upper function for $\sup_{\theta\in\Theta_\alpha}\Psi(\chi_\theta)$ on $\mathfrak{A}$, which can be also viewed as an upper function for $\Psi(\chi_\theta)$ on Θ. The corresponding results are used in order to obtain rather precise inequalities for the modulus of continuity of random functions, Section 4.2. Moreover we apply this bound for deriving of an upper function for the $\mathbb{L}_p$-norms of Wiener integrals, Section 4.1.

We finish this short introduction with the following remark. In order to establish the inequalities (1.3)–(1.4) for the upper functions presented below we will need to prove that corresponding supremum is a random variable. The result below is sufficient for all problems considered in the paper and before we start the proofs we will not discuss the measurability issue.

Lemma 1. *Let $\mathfrak{T}$ be the set equipped with the metric $\mathfrak{d}$, $(\mathbf{\Omega}, \mathfrak{B}, \mathbf{P})$ be a complete probability space and $\zeta : \mathbf{\Omega}\times\mathfrak{T}\to\mathbb{R}$ be $\mathbf{P}$-a.s. continuous. Let $\mathfrak{Z}$ be a set, $g : \mathfrak{Z}\to\mathbb{R}$ be a given function and $\{\mathfrak{T}_\mathfrak{z}\subseteq\mathfrak{T},\ \mathfrak{z}\in\mathfrak{Z}\}$ be an arbitrary sequence of sets. If $\mathfrak{T}$ is totally bounded then* $\sup_{\mathfrak{z}\in\mathfrak{Z}}\left[\sup_{\mathfrak{t}\in\mathfrak{T}_\mathfrak{z}}\zeta(\mathfrak{t},\cdot) - g(\mathfrak{z})\right]$ *is $\mathfrak{B}$-measurable.*

The proof of the lemma is given in Appendix. We would like to emphasize that there is no any assumption imposed on the function g, index set $\mathfrak{Z}$ and on the collection $\{\mathfrak{T}_\mathfrak{z}\subseteq\mathfrak{T},\ \mathfrak{z}\in\mathfrak{Z}\}$.

Putting $\mathfrak{Z} = \mathfrak{T}$ and $\mathfrak{T}_\mathfrak{t} = \{\mathfrak{t}\}$ we come to the following consequence of Lemma 1.

Corollary 1. *Under assumptions of Lemma 1* $\sup_{\mathfrak{t}\in\mathfrak{T}}\left[\zeta(\mathfrak{t},\cdot) - g(\mathfrak{t})\right]$ *is $\mathfrak{B}$-measurable.*

Upper functions of the first type As it was said above throughout this section we will suppose that $\underline{A} > 0, \underline{B} > 0$. Put for any $t > 0$

$$\Theta_A(t) = \Big\{\theta\in\Theta:\ A(\theta)\le t\Big\},\quad \Theta_B(t) = \Big\{\theta\in\Theta:\ B(\theta)\le t\Big\}.$$

For any $\vec{s}\in\mathcal{S}_{\mathrm{a,b}}$ introduce the function

$$\mathcal{E}_{\vec{s}}(u,v) = e^{(\mathrm{a})}_{s_1}\Big(\underline{A}u, \Theta_A\big(\underline{A}u\big)\Big) + e^{(\mathrm{b})}_{s_2}\Big(\underline{B}v, \Theta_B\big(\underline{B}v\big)\Big),\quad u,v\ge 1. \tag{2.3}$$

Denote also $\ell(u) = \ln\{1+\ln(u)\} + 2\ln\{1+\ln\{1+\ln(u)\}\}$ and set for any θ and $\varepsilon > 0, r\ge 0$

$$P_\varepsilon(\theta) = 2\big[1+\varepsilon^{-1}\big]^2\mathcal{E}_{\vec{s}}\big(\mathcal{A}_\varepsilon(\theta),\mathcal{B}_\varepsilon(\theta)\big) + (1+\varepsilon)^2\big[\ell\big(\mathcal{A}_\varepsilon(\theta)\big) + \ell\big(\mathcal{B}_\varepsilon(\theta)\big)\big]; \tag{2.4}$$

$$M_{\varepsilon,r}(\theta) = (1+\varepsilon)^2\left\{2\big[1+\varepsilon^{-1}\big]^2\mathcal{E}_{\vec{s}}\big(\mathcal{A}_\varepsilon(\theta),\mathcal{B}_\varepsilon(\theta)\big) + (\varepsilon+r)\ln\big[\mathcal{A}_\varepsilon(\theta)\mathcal{B}_\varepsilon(\theta)\big]\right\}, \tag{2.5}$$

where $\mathcal{A}_\varepsilon(\theta) = (1+\varepsilon)\big[A(\theta)/\underline{A}\big]$ and $\mathcal{B}_\varepsilon(\theta) = (1+\varepsilon)\big[B(\theta)/\underline{B}\big]$. Define for any $z\ge 0$

$$\mathrm{V}^{(z,\varepsilon)}(\theta) = (1+\varepsilon)^2\left(A(\theta)\sqrt{P_\varepsilon(\theta) + (1+\varepsilon)^2 z} + B(\theta)\Big[P_\varepsilon(\theta) + (1+\varepsilon)^2 z\Big]\right); \tag{2.6}$$

$$\mathrm{U}^{(z,\varepsilon,r)}(\theta) = (1+\varepsilon)^2\left(A(\theta)\sqrt{M_{\varepsilon,r}(\theta) + (1+\varepsilon)^2 z} + B(\theta)\Big[M_{\varepsilon,r}(\theta) + (1+\varepsilon)^2 z\Big]\right). \tag{2.7}$$

In the proposition below we prove that the functions defined in (2.6) and (2.7) are upper functions for $\Psi(\chi_\theta)$ on Θ. We remark that they are completely determined by the functions A and B and by the entropies of their level sets measured in semi-metrics a and b. The number ε and the couple of functions $\vec{s}$ can be viewed as tuning parameters allowing either to weaken assumptions or to obtain sharper bounds but they are not related to the random functional $\Psi(\chi_\theta)$ itself.

Proposition 2. *Let Assumptions 1-3 be fulfilled. Then* $\forall \vec{s} \in \mathcal{S}_{\mathrm{a,b}}$, $\forall \varepsilon \in \left(0, \sqrt{2}-1\right]$ *and* $\forall z \geq 1$

$$\mathrm{P}\left\{\sup_{\theta\in\Theta}\left[\Psi(\chi_\theta) - \mathrm{V}^{(z,\epsilon)}(\theta)\right] \geq 0\right\} \leq 2\mathrm{c}\left[1 + \left[\ln\left\{1+\ln(1+\varepsilon)\right\}\right]^{-2}\right]^2 \exp\{-z\};$$

$$\mathrm{E}\left\{\sup_{\theta\in\Theta}\left[\Psi(\chi_\theta) - \mathrm{U}^{(z,\epsilon,q)}(\theta)\right]\right\}_+^q \leq \mathrm{c}2^{(5q/2)+2}\Gamma(q+1)\,\varepsilon^{-q-4}\left[\underline{A}\vee\underline{B}\right]^q \exp\{-z\}.$$

It is obvious that the assertions of the proposition remain valid if one replaces the function $\mathcal{E}_{\vec{s}}$ by any its upper bound. It is important since the exact computation of this function is too complicated in general. We note that the role of the latter function in our construction is similar to those which Dudley integral plays in the computations of the expectation of the suprema of gaussian or sub-gaussian processes Lifshits (1995), Talagrand (2005).

Price to be paid for "uniformity" We remark that in view of (1.7) and (1.8), the function $U^{(z)}(\theta) := A(\theta)\sqrt{z} + B(\theta)z$ can be viewed as "pointwise upper function" for $\Psi(\chi_\theta)$, i.e. for fixed θ. Comparing the inequalities (1.7) and (1.8) with those given in Proposition 2 we conclude that they differ from each other by numerical constants only. In this context, the functions $P_\varepsilon(\cdot)$ and $M_{\varepsilon,r}(\cdot)$ given by (2.4) and (2.5) can be viewed as price to be paid for "uniformity". That means that in order to pass from "pointwise" result to the "uniform" one we need, roughly speaking, to multiply $A(\cdot)$ by $\sqrt{P_\varepsilon(\cdot)}$ or $\sqrt{M_{\varepsilon,r}(\cdot)}$ and $B(\cdot)$ by $P_\varepsilon(\cdot)$ or $M_{\varepsilon,r}(\cdot)$. The question, arising naturally: is such payment necessary or minimal? In this context it is worth mentioning the relation between well-known phenomenon in adaptive estimation, called *price to be paid for adaptation* Lepski (1991), Lepski and Spokoiny (1997) and Spokoiny (1996), and what we call here *price to be paid for uniformity*. We have no place here to describe this relation in detail and mention only several facts.

First let us remark almost all constructions of adaptive estimators (model selection Barron et al. (1999), risk hull minimization Cavalier and Golubev (2006), Lepski method Lepski (1991), or recently developed universal estimation routine Goldenshluger and Lepski (2008, 2009)) involve the upper functions for stochastic objects of different kinds. Next, it is known that there are two types of *price to be paid for adaptation*: (ln)-price, Lepski (1991) and (ln ln)-price, Spokoiny (1996). The (ln)-price appears in the problems where the risk of estimation procedures is described by a power loss-functions and it corresponds to the function $M_{\varepsilon,r}(\cdot)$, where the parameter r is a power. The (ln ln)-price appears in the case of bounded losses that corresponds to the function $P_\varepsilon(\cdot)$. Since the theory of adaptive estimation is equipped with very developed criteria of optimality, Lepski (1991), Tsybakov (1998), Kluchnikoff (2005), we might assert that the payment for uniformity is optimal if the use of corresponding upper function leads to optimal adaptive estimators.

We finish the discussion concerning the statements of Proposition 2 with the following remark. Comparing the result given in (1.8) with the second assertion of Proposition 2 we can state that the inequality obtained there is very precise since, remind, $\underline{A} = \inf_{\theta\in\Theta} A(\theta)$ and $\underline{B} = \inf_{\theta\in\Theta} B(\theta)$.

Upper functions of the second type Suppose that we are given by the collection $\left\{\Theta_\alpha,\ \alpha\in\mathfrak{A}\right\}$, satisfying $\Theta = \cup_{\alpha\in\mathfrak{A}}\Theta_\alpha$, and by two mappings $\tau_1 : \mathfrak{A} \to \left(0, \overline{\tau}_1\right]$, $\tau_2 : \mathfrak{A} \to \left(0, \overline{\tau}_2\right]$, where $\overline{\tau}_1, \overline{\tau}_2 < \infty$.

For any $u > 0$ put

$$\Theta_1'(u) = \bigcup_{\alpha:\, \tau_1(\alpha) \le u} \Theta_\alpha, \quad g_A^*(u) = \sup_{\theta \in \Theta_1'(u)} A(\theta);$$
$$\Theta_2'(u) = \bigcup_{\alpha:\, \tau_2(\alpha) \le u} \Theta_\alpha, \quad g_B^*(u) = \sup_{\theta \in \Theta_2'(u)} B(\theta),$$

and let g_A and g_B be arbitrary chosen increasing functions, satisfying $g_A \ge g_A^*$ and $g_B \ge g_B^*$ (we note that obviously g_A^* and g_B^* are increasing).

Since $\Theta_1'(\cdot), \Theta_2'(\cdot) \subseteq \Theta$, in view of Assumption 3 for any $u, v > 0$ one can find the functions $s_1(u, \cdot)$ and $s_2(v, \cdot)$ for which the latter assumption is fulfilled on $\Theta_1'(u)$ and $\Theta_2'(v)$ respectively. Let us suppose additionally that

$$\lambda_1 := \sup_{t \in [1, \sqrt{2}]} \sup_{x > \underline{\tau}_1} \sup_{\delta > 0} \frac{s_1(xt, \delta)}{s_1(x, \delta)} < \infty, \quad \lambda_2 := \sup_{t \in [1, \sqrt{2}]} \sup_{x > \underline{\tau}_2} \sup_{\delta > 0} \frac{s_2(xt, \delta)}{s_2(x, \delta)} < \infty, \tag{2.8}$$

where $\underline{\tau}_1 = \inf_\alpha \tau_1(\alpha)$ and $\underline{\tau}_2 = \inf_\alpha \tau_2(\alpha)$.

We remark that if the functions $s_1(u, \cdot)$ and $s_2(v, \cdot)$ are chosen independently of u, v then $\lambda_1 = \lambda_2 = 1$. It is also obvious that $\lambda_1, \lambda_2 \ge 1$.

The condition (2.8) allows us to define the function:

$$\mathcal{E}'(u, v) = e^{(\mathrm{a})}_{s_1(u,\cdot)}\Big(\lambda_1^{-1} g_A(u), \Theta_1'(u)\Big) + e^{(\mathrm{b})}_{s_2(v,\cdot)}\Big(\lambda_2^{-1} g_B(v), \Theta_2'(v)\Big), \quad u, v > 0. \tag{2.9}$$

We note that the function $\mathcal{E}'$ is constructed similarly to the function $\mathcal{E}$ used in the previous section, but now the functions s_1 and s_2 can be chosen in accordance with considered level sets.

At last, for any $\alpha \in \mathfrak{A}$ and any $\varepsilon > 0$ set

$$\widehat{\mathcal{E}}^{(\varepsilon)}(\alpha) = \mathcal{E}'\Big((1+\varepsilon)\tau_1(\alpha), (1+\varepsilon)\tau_2(\alpha)\Big).$$

Put $\delta_j = (1+\varepsilon)^{-j}, j \ge 0$, and let $R_r : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+, r \ge 0$, be an arbitrary family of *increasing (or decreasing)* in both arguments functions, satisfying for any $\varepsilon \in \big(0, \sqrt{2} - 1\big]$

$$\sum_{j=0}^{J} \sum_{k=0}^{K} \big[g_A\big(\overline{\tau}_1 \delta_j\big) \vee g_B\big(\overline{\tau}_2 \delta_k\big)\big]^r \exp\big\{-R_r\big(\overline{\tau}_1 \delta_j, \overline{\tau}_2 \delta_k\big)\big\} =: \mathcal{R}^{(\varepsilon, r)} < \infty. \tag{2.10}$$

Here integers J, K are defined as follows.

$$J = \big\lfloor \ln_{1+\epsilon}\big(\overline{\tau}_1 / \underline{\tau}_1\big) \big\rfloor + 1, \quad K = \big\lfloor \ln_{1+\epsilon}\big(\overline{\tau}_2 / \underline{\tau}_2\big) \big\rfloor + 1.$$

If $\underline{\tau}_{\mathrm{i}} = 0$, $\mathrm{i} = 1, 2$, the corresponding quantity is put equal to infinity.

Set $\widehat{R}_r^{(\varepsilon)}(\alpha) = R_r\Big(r_\varepsilon \tau_1(\alpha), r_\varepsilon \tau_2(\alpha)\Big)$, where $r_\varepsilon = (1+\varepsilon)$ if R_r is increasing and $r_\varepsilon = (1+\varepsilon)^{-1}$ if R_r is decreasing, and define

$$\begin{aligned} \widehat{\mathrm{U}}^{(z,\varepsilon,r)}(\alpha) &= (1+\varepsilon) g_A\Big([1+\varepsilon]^2 \tau_1(\alpha)\Big) \sqrt{2\big[1+\varepsilon^{-1}\big]^2\, \widehat{\mathcal{E}}^{(\varepsilon)}(\alpha) + \widehat{R}_r^{(\varepsilon)}(\alpha) + z} \\ &+ (1+\varepsilon)^2 g_B\Big([1+\varepsilon]^2 \tau_2(\alpha)\Big)\Big(2\big[1+\varepsilon^{-1}\big]^2\, \widehat{\mathcal{E}}^{(\varepsilon)}(\alpha) + \widehat{R}_r^{(\varepsilon)}(\alpha) + z\Big). \end{aligned}$$

Below we assert that $\widehat{U}^{(z,\varepsilon,r)}, r=0, r=q$, are upper functions for $\Big[\sup_{\theta\in\Theta_\alpha}\Psi\left(\chi_\theta\right)\Big]$ on $\mathfrak{A}$. However, before to present exact statements, let us briefly discuss some possible choices of the functions R_r. We would like to emphasize that the opportunity to select these functions allows to obtain quite different and precise results. First possible choice is given by

$$R_0(u,v)=\ell\Big(\overline{\tau}_1u^{-1}\Big)+\ell\Big(\overline{\tau}_2v^{-1}\Big),\quad R_r(u,v)=\varepsilon\Big[\ln\Big(\overline{\tau}_1u^{-1}\Big)+\ln\Big(\overline{\tau}_2v^{-1}\Big)\Big],\ r>0. \tag{2.11}$$

These functions are used in the problems in which $\widehat{\mathcal{E}}^{(\varepsilon)}(\cdot)$ is bounded by some absolute constant independent of all quantities involved in the description of the problem, assumptions etc.

This choice leads to the following values of the constants in (2.10):

$$\mathcal{R}^{(\varepsilon,0)}\le\left[2+\Big[\ln\big\{1+\ln\left(1+\varepsilon\right)\big\}\Big]^{-2}\right]^2,\quad \mathcal{R}^{(\varepsilon,r)}<4\Big[g_A\big(\overline{\tau}_1\big)\vee g_B\big(\overline{\tau}_2\big)\Big]^r\varepsilon^{-4}. \tag{2.12}$$

Another important choice is given by $R_r=\mathcal{E}'$ independently of r, see, for instance, Theorem 9. In view of (2.10), this choice corresponds to the case when the function $\mathcal{E}'$ increases to infinity.

Proposition 3. *Let Assumptions 1-3 be fulfilled. Then for any s_1,s_2 satisfying (2.8) and any $R_r,\ r\ge0$, satisfying (2.10), for any $\varepsilon\in\big(0,\sqrt{2}-1\big]$ and any $z\ge1, q\ge1$*

$$\mathrm{P}\left\{\sup_{\alpha\in\mathfrak{A}}\left[\sup_{\theta\in\Theta_\alpha}\Psi\left(\chi_\theta\right)-\widehat{\mathrm{U}}^{(z,\varepsilon,0)}(\alpha)\right]\ge0\right\}\le2\mathrm{c}\mathcal{R}^{(\varepsilon,0)}\exp\left\{-z\right\};$$

$$\mathrm{E}\left\{\sup_{\alpha\in\mathfrak{A}}\left[\sup_{\theta\in\Theta_\alpha}\Psi\left(\chi_\theta\right)-\widehat{\mathrm{U}}^{(z,\varepsilon,q)}(\alpha)\right]\right\}_+^q\le\mathrm{c}2^{(5q/2)+1}\Gamma(q+1)\mathcal{R}^{(\varepsilon,q)}\,\varepsilon^{-q}\exp\left\{-z\right\}.$$

Remark 3. *We note that the results of the proposition is very general. Indeed, there are no assumptions imposed on the collection $\Theta_\alpha,\ \alpha\in\mathfrak{A}$, and the functions τ_1,τ_2 can be chosen arbitrary. Moreover, the condition (2.10) is very mild, so the choice of functions R_r is quite flexible.*

2.3. Upper functions for the modulus of continuity of random mappings

In this section we apply Proposition 3 in order to derive upper functions for the local and global modulus of continuity of real-valued random mappings. It is worth mentioning that in this circle of problems the upper functions are actively exploited, see e.g. Egishyants and Ostrovskii (1996) and the references therein. We will suppose that Assumption 1 (*2*), Assumption 2 and Assumption 3 are verified, χ_{t} is real-valued random mapping defined on the metric space $\mathfrak{T}$, d is a semi-metric on $\mathfrak{T}$ and $\Psi(\cdot)=|\cdot|$.

Upper function for local modulus of continuity Let θ_0 be a fixed element of Θ and set for any $\Delta\in\big(0,D_{\mathrm{d}}(\Theta)\big]$, where $D_{\mathrm{d}}(\Theta)$ is the diameter of Θ measured in the semi-metric d,

$$\mathfrak{m}_\Delta(\theta_0)=\sup_{\theta\in\Theta_\Delta}\big|\chi_\theta-\chi_{\theta_0}\big|,\quad \Theta_\Delta=\Big\{\theta\in\Theta:\ \ \mathrm{d}\big(\theta,\theta_0\big)\le\Delta\Big\}.$$

Thus, $\mathfrak{m}_\Delta(\theta_0),\ \Delta\in\big(0,D_{\mathrm{d}}(\Theta)\big]$, is the local modulus of continuity of χ_θ in θ_0 measured in d.

If we put $\widetilde{\chi}_\theta=\chi_\theta-\chi_{\theta_0},\ \theta\in\Theta$, we assert first that Assumption 1 (*2*) can be viewed as Assumption 1 (*1*) for $\widetilde{\chi}_\theta$ on Θ with $A(\cdot)=\mathrm{a}(\cdot,\theta_0)$ and $B(\cdot)=\mathrm{b}(\cdot,\theta_0)$. Next, noting that $\widetilde{\chi}_{\theta_1}-\widetilde{\chi}_{\theta_2}=\chi_{\theta_1}-\chi_{\theta_2}$ for any $\theta_1,\theta_2\in\Theta$ we conclude that Assumption 1 (*2*) is verified for $\widetilde{\chi}_\theta$ on Θ with a and b.

Thus, we can apply Proposition 3 with $\alpha = \Delta$, $\Theta_\alpha = \Theta_\Delta$, $\mathfrak{A} = \big(0, D_{\rm d}(\Theta)\big]$ and we choose $\tau_1(\Delta) = \tau_2(\Delta) = \Delta$. This choice implies obviously for any $u \le D_{\rm d}(\Theta)$

$$\Theta_1'(u) = \Theta_2'(u) = \Theta_u, \quad g_A(u) = \sup_{\theta:\, {\rm d}\left(\theta,\theta_0\right)\le u} {\rm a}\big(\theta,\theta_0\big), \quad g_B(u) = \sup_{\theta:\, {\rm d}\left(\theta,\theta_0\right)\le u} {\rm b}\big(\theta,\theta_0\big).$$

Fix $\vec{s} \in \mathcal{S}_{\rm a,b}$ and put for any $\Delta \in \big(0, D_{\rm d}(\Theta)\big]$ and any $\varepsilon \in \big(0, \sqrt{2}-1\big]$

$$\widehat{\mathcal{E}}^{(\varepsilon)}(\Delta,\theta_0) = e^{({\rm a})}_{s_1}\Big(g_A\big([1+\varepsilon]\Delta\big), \Theta_{[1+\varepsilon]\Delta}\Big) + e^{({\rm b})}_{s_2}\Big(g_B\big([1+\varepsilon]\Delta\big), \Theta_{[1+\varepsilon]\Delta}\Big).$$

Here $e^{({\rm a})}_{s_1}$ and $e^{({\rm b})}_{s_2}$ are defined by (1.9). We also set $\lambda_1 = \lambda_2 = 1$ since the functions s_1, s_2 are chosen independently of the collection $\big\{\Theta_\Delta,\ \Delta \in \big(0, D_{\rm d}(\Theta)\big]\big\}$.

Choose also $R_0(u,v) = \ell\Big(\overline{\tau}_1 u^{-1}\Big) + \ell\Big(\overline{\tau}_2 v^{-1}\Big)$ and define

$$\begin{aligned}
\widehat{V}^{(z,\varepsilon)}_{\vec{s}}(\Delta,\theta_0) &= (1+\varepsilon) g_{\mathcal{A}}\Big([1+\varepsilon]^2\Delta\Big)\sqrt{2\big[1+\varepsilon^{-1}\big]^2\widehat{\mathcal{E}}^{(\varepsilon)}(\Delta,\theta_0) + 2\ell\Big((1+\varepsilon)D_{\rm d}(\Theta)/\Delta\Big) + z} \\
&+ (1+\varepsilon)^2 g_{\mathcal{B}}\Big([1+\varepsilon]^2\Delta\Big)\Big\{2\big[1+\varepsilon^{-1}\big]^2\widehat{\mathcal{E}}^{(\varepsilon)}(\Delta,\theta_0) + 2\ell\Big((1+\varepsilon)D_{\rm d}(\Theta)/\Delta\Big) + z\Big\}.
\end{aligned}$$

Then, applying Proposition 3 and taking into account (2.12) we come to the following result.

Proposition 4. *Let Assumptions 1-3 be fulfilled. Then for any $\vec{s} \in \mathcal{S}_{\rm a,b}$, $\varepsilon \in \big(0, \sqrt{2}-1\big]$ and $z \ge 1$*

$$\mathrm{P}\left\{\sup_{\Delta\in\left(0,D_{\rm d}(\Theta)\right]}\Big[\mathfrak{m}_\Delta - \widehat{V}^{(z,\varepsilon)}_{\vec{s}}(\Delta,\theta_0)\Big] \ge 0\right\} \le 2{\rm c}\left[2 + \Big[\ln\big\{1+\ln(1+\varepsilon)\big\}\Big]^{-2}\right]^2 \exp\{-z\}.$$

In Section 4 we apply Proposition 4 to gaussian random functions defined on a metric space satisfying so-called *doubling condition*.

Remark 4. *If* ${\rm b} \equiv 0$, ${\rm d} = {\rm a}$ *and* $\sup_{\Delta\in\left(0,D_{\rm d}(\Theta)\right]} \widehat{\mathcal{E}}^{(\varepsilon)}(\Delta,\theta_0) =: \widehat{\mathcal{E}}^{(\varepsilon)}(\theta_0) < \infty$, *the upper function* $\widehat{V}^{(z,\varepsilon)}_{\vec{s}}$ *has a very simple form*

$$\widehat{V}^{(z,\varepsilon)}_{\vec{s}}(\Delta,\theta_0) = (1+\varepsilon)^3\Delta\sqrt{2\big[1+\varepsilon^{-1}\big]^2\widehat{\mathcal{E}}^{(\varepsilon)}(\theta_0) + \ell\Big((1+\varepsilon)D_{\rm d}(\Theta)/\Delta\Big) + z}. \tag{2.13}$$

Hence, the result of Proposition 4 can be viewed as the non-asymptotical version of the law of iterated logarithm for sub-gaussian processes defined on some totaly bounded subset of metric space. In this context it is worth mentioning the paper Egishyants and Ostrovskii (1996) where the upper functions for local and global modulus of continuity were found for the stochastic processes satisfying Cramer's condition.

Remark 5. *We also note that we replaced in (2.13) the factor* $2\ell\big(D_{\rm d}(\Theta)/\Delta\big)$ *appeared in the upper function used in Proposition 4 by* $\ell\big(D_{\rm d}(\Theta)/\Delta\big)$. *It is explained by the fact that* $\overline{\tau}_2 = 0$ *in (2.11) since* $B, {\rm b} \equiv 0$. *By the same reason, the probability bound in this case is given by* $2{\rm c}\left[2 + \Big[\ln\big\{1+\ln(1+\varepsilon)\big\}\Big]^{-2}\right]\exp\{-z\}$.

Upper function for global modulus of continuity Set $\Theta^{(2)} = \Theta \times \Theta$ and let for any $\vartheta = (\theta_1, \theta_2) \in \Theta^{(2)}$ and any $\Delta \in \big(0, D_{\rm d}(\Theta)\big]$,

$$\zeta(\vartheta) = \chi_{\theta_1} - \chi_{\theta_2}, \quad \mathfrak{m}_\Delta = \sup_{\vartheta \in \Theta^{(2)}_\Delta} \big|\zeta_\vartheta\big|, \quad \Theta^{(2)}_\Delta = \Big\{\vartheta \in \Theta^{(2)} : \ {\rm d}\big(\theta_1, \theta_2\big) \le \Delta\Big\}.$$

Thus, $\mathfrak{m}_\Delta$, $\Delta \in \big(0, D_{\rm d}(\Theta)\big]$, is the global modulus of continuity of χ_θ on Θ measured in d.

Put $\mathbf{A}(\vartheta) = {\rm a}(\theta_1, \theta_2)$, $\mathbf{B}(\vartheta) = {\rm b}(\theta_1, \theta_2)$, $\vartheta = (\theta_1, \theta_2) \in \Theta^{(2)}$, and equip $\Theta^{(2)}$ with the following semi-metrics: $\vartheta = (\theta_1, \theta_2), \varsigma = (\varsigma_1, \varsigma_2) \in \Theta^{(2)}$

$${\rm a}^{(2)}(\vartheta, \varsigma) = 2\left[{\rm a}(\theta_1, \varsigma_1) \vee {\rm a}(\theta_2, \varsigma_2)\right], \ \ {\rm b}^{(2)}(\vartheta, \varsigma) = 2\left[{\rm b}(\theta_1, \varsigma_1) \vee {\rm b}(\theta_2, \varsigma_2)\right].$$

Some remarks are in order. We note first that Assumption 1 (*2*) can be viewed as Assumption 1 (*1*) for $\zeta(\vartheta)$ on $\Theta^{(2)}$ with $A = \mathbf{A}$ and $B = \mathbf{B}$.

Next we obtain in view of Assumption 1 (*2*) $\forall \vartheta, \varsigma \in \Theta^{(2)}$ and $\forall z > 0$

$$\begin{aligned}
&{\rm P}\Big\{|\zeta(\vartheta) - \zeta(\varsigma)| \ge z\Big\} \le {\rm P}\Big\{\big|\chi_{\theta_1} - \chi_{\varsigma_1}\big| \ge z/2\Big\} + {\rm P}\Big\{\big|\chi_{\theta_2} - \chi_{\varsigma_2}\big| \ge z/2\Big\} \\
&\le {\rm c}\exp\left\{-\frac{z^2}{4\big[{\rm a}(\theta_1, \varsigma_1)\big]^2 + 2{\rm b}(\theta_1, \varsigma_1) z}\right\} + {\rm c}\exp\left\{-\frac{z^2}{4\big[{\rm a}(\theta_2, \varsigma_2)\big]^2 + 2{\rm b}(\theta_2, \varsigma_2) z}\right\} \\
&\le {\rm c}^{(2)}\exp\left\{-\frac{z^2}{\big[{\rm a}^{(2)}(\vartheta, \varsigma)\big]^2 + {\rm b}^{(2)}(\vartheta, \varsigma) z}\right\}
\end{aligned}$$

We conclude that Assumption 1 (*2*) holds for $\zeta(\vartheta)$ on $\Theta^{(2)}$ with ${\rm a} = {\rm a}^{(2)}$, ${\rm b} = {\rm b}^{(2)}$ and ${\rm c}^{(2)} = 2{\rm c}$.

Since obviously

$$\mathfrak{E}_{{\rm a}^{(2)}, \Theta^{(2)}}(\varsigma) \le 2\mathfrak{E}_{{\rm a}, \Theta}(\varsigma/2), \quad \mathfrak{E}_{{\rm b}^{(2)}, \Theta^{(2)}}(\varsigma) \le 2\mathfrak{E}_{{\rm b}, \Theta}(\varsigma/2), \ \ \varsigma > 0, \tag{2.14}$$

we assert that Assumptions 2 and 3 are fulfilled on $\Theta^{(2)}$ with ${\rm a} = {\rm a}^{(2)}$ and ${\rm b} = {\rm b}^{(2)}$.

Put $\overline{\Theta}^{(2)} = \cup_{\Delta > 0} \Theta^{(2)}_\Delta$. Since $\overline{\Theta}^{(2)} \subset \Theta^{(2)}$ we can apply Proposition 3 with $\alpha = \Delta$, $\Theta_\alpha = \Theta^{(2)}_\Delta$, $\mathfrak{A} = \big(0, D_{\rm d}(\Theta)\big]$ and we choose $\tau_1(\Delta) = \tau_2(\Delta) = \Delta$.

The latter choice implies obviously for any $u \le D_{\rm d}(\Theta)$

$$\Theta'_1(u) = \Theta'_2(u) = \Theta^{(2)}_u, \quad g_{\mathbf{A}}(u) = \sup_{\vartheta \in \Theta^{(2)}_u} \mathbf{A}(\vartheta), \ \ g_{\mathbf{B}}(u) = \sup_{\vartheta \in \Theta^{(2)}_u} \mathbf{B}(\vartheta).$$

Fix $\vec{s} \in \mathcal{S}_{\rm a,b}$ and put for any $\Delta \in \big(0, D_{\rm d}(\Theta)\big]$ and any $\varepsilon \in \big(0, \sqrt{2} - 1\big]$

$$\widehat{\mathcal{E}}^{(\varepsilon)}(\Delta) = e^{({\rm a}^{(2)})}_{s_1}\Big(g_A\big([1+\varepsilon]\Delta\big), \Theta^{(2)}_{[1+\varepsilon]\Delta}\Big) + e^{({\rm b}^{(2)})}_{s_2}\Big(g_B\big([1+\varepsilon]\Delta\big), \Theta^{(2)}_{[1+\varepsilon]\Delta}\Big).$$

Here $e^{({\rm a}^{(2)})}_{s_1}$ and $e^{({\rm b}^{(2)})}_{s_2}$ are defined by (1.9), where ${\rm a}, {\rm b}$ are replaced by ${\rm a}^{(2)}$ and ${\rm b}^{(2)}$ respectively. We also set $\lambda_1 = \lambda_2 = 1$ since the functions s_1, s_2 are chosen independently of the collection $\Big\{\Theta^{(2)}_\Delta, \ \Delta \in \big(0, D_{\rm d}(\Theta)\big]\Big\}$. Choose $R_0(u, v) = \ell\Big(\overline{\tau}_1 u^{-1}\Big) + \ell\Big(\overline{\tau}_2 v^{-1}\Big)$ and define

$$\begin{aligned}
\widehat{V}^{(z,\varepsilon)}_{\vec{s}}(\Delta) \ &= \ (1+\varepsilon) g_{\mathcal{A}}\Big([1+\varepsilon]^2\Delta\Big)\sqrt{2\big[1+\varepsilon^{-1}\big]^2\widehat{\mathcal{E}}^{(\varepsilon)}(\Delta) + 2\ell\Big((1+\varepsilon)D_{\rm d}(\Theta)/\Delta\Big) + z} \\
&+ \ (1+\varepsilon)^2 g_{\mathcal{B}}\Big([1+\varepsilon]^2\Delta\Big)\Big\{2\big[1+\varepsilon^{-1}\big]^2\widehat{\mathcal{E}}^{(\varepsilon)}(\Delta) + 2\ell\Big((1+\varepsilon)D_{\rm d}(\Theta)/\Delta\Big) + z\Big\}.
\end{aligned}$$

Then, applying Proposition 3 and taking into account (2.12) we come to the following result.

Proposition 5. *Let Assumptions 1-3 be fulfilled. Then for any* $\vec{s} \in \mathcal{S}_{\mathrm{a,b}}$, $\varepsilon \in \left(0, \sqrt{2}-1\right]$ *and* $z \geq 1$

$$\mathrm{P}\left\{\sup_{\Delta \in \left(0, D_{\mathrm{d}}(\Theta)\right]} \left[\mathfrak{m}_{\Delta} - \widehat{V}_{\vec{s}}^{(z,\varepsilon)}(\Delta)\right] \geq 0\right\} \leq 4\mathrm{c}\left[2 + \left[\ln\left\{1 + \ln\left(1+\varepsilon\right)\right\}\right]^{-2}\right]^{2} \exp\left\{-z\right\}.$$

The obtained inequality allows, in particular, to prove that the families of probabilities measures generated by χ_{θ} is dense. This, in its turn, is crucial step in proving of the weak convergence of probabilities measures.

3. Application to empirical processes theory

Let $\left(\mathcal{X}, \mathfrak{X}, \nu\right)$ be σ-finite space and let $(\Omega, \mathfrak{A}, \mathsf{P})$ be a commplete probability space. Let $X_i,\ i \geq 1$, be a the collection of $\mathcal{X}$-valued *independent* random variables defined on $(\Omega, \mathfrak{A}, \mathsf{P})$ and having the densities f_i with respect to measure ν. Furthermore, $\mathbb{P}_{\mathrm{f}},\ \mathrm{f} = (f_1, f_2, \ldots)$, denotes the probability law of $(X_1, X_2, \ldots)$ and $\mathbb{E}_{\mathrm{f}}$ is mathematical expectation with respect to $\mathbb{P}_{\mathrm{f}}$.

Let $G : \mathfrak{H} \times \mathcal{X} \to \mathbb{R}$ be a given mapping, where $\mathfrak{H}$ is a set. Put $\forall n \in \mathbb{N}^{*}$

$$\xi_{\mathfrak{h}}(n) = n^{-1}\sum_{i=1}^{n}\left[G\big(\mathfrak{h}, X_i\big) - \mathbb{E}_{\mathrm{f}} G(\mathfrak{h}, X_i)\right], \quad \mathfrak{h} \in \mathfrak{H}. \tag{3.1}$$

We will say that $\xi_{\mathfrak{h}}(n),\ \mathfrak{h} \in \mathfrak{H}$, is a generalized empirical process. Note that if $\mathfrak{h} : \mathcal{X} \to \mathbb{R}$ and $G(\mathfrak{h}, x) = \mathfrak{h}(x),\ \mathfrak{h} \in \mathfrak{H}, x \in \mathcal{X}$, then $\xi_{\mathfrak{h}}(n)$ is the standard empirical process parameterized by $\mathfrak{H}$.

Throughout this section we will suppose that

$$\overline{\mathbf{G}}_{\infty}(\mathfrak{h}) := \sup_{x \in \mathcal{X}}\left|G(\mathfrak{h}, x)\right| < \infty, \quad \forall \mathfrak{h} \in \mathfrak{H}, \tag{3.2}$$

and it will be referred to *bounded case*. Some generalizations concerning the situations where this assumption fails are discussed in Section 3.1.

The condition (3.2) implies that the random variables $G(\mathfrak{h}, X_i), \mathfrak{h} \in \mathfrak{H}$, and $G(\mathfrak{h}_1, X_i) - G(\mathfrak{h}_2, X_j)$, $\mathfrak{h}_1, \mathfrak{h}_2 \in \mathfrak{H}$, $i = \overline{1, n}$, are bounded, and we obtain in view of Bernstein's inequality $\forall z > 0$

$$\mathbb{P}_{\mathrm{f}}\left\{\left|\xi_{\mathfrak{h}}(n)\right| > z\right\} \leq 2\exp\left\{-\frac{z^2}{A_{\mathrm{f}}^2(\mathfrak{h}) + zB_{\infty}(\mathfrak{h})}\right\}; \tag{3.3}$$

$$\mathbb{P}_{\mathrm{f}}\left\{\left|\xi_{\mathfrak{h}_1}(n) - \xi_{\mathfrak{h}_2}(n)\right| > z\right\} \leq 2\exp\left\{-\frac{z^2}{\mathrm{a}_{\mathrm{f}}^2(\mathfrak{h}_1, \mathfrak{h}_2) + z\mathrm{b}_{\infty}(\mathfrak{h}_1, \mathfrak{h}_2)}\right\}, \tag{3.4}$$

where

$$A_{\mathrm{f}}^2(\mathfrak{h}) = 2n^{-2}\sum_{i=1}^{n}\mathbb{E}_f G^2(\mathfrak{h}, X_i), \qquad \mathrm{a}_{\mathrm{f}}^2(\mathfrak{h}_1, \mathfrak{h}_2) = 2n^{-2}\sum_{i=1}^{n}\mathbb{E}_f\big(G(\mathfrak{h}_1, X_i) - G(\mathfrak{h}_2, X_i)\big)^2; \tag{3.5}$$

$$B_{\infty}(\mathfrak{h}) = (4/3)n^{-1}\sup_{x \in \mathcal{X}}\left|G(\mathfrak{h}, x)\right|, \quad \mathrm{b}_{\infty}(\mathfrak{h}_1, \mathfrak{h}_2) = (4/3)n^{-1}\sup_{x \in \mathcal{X}}\left|G(\mathfrak{h}_1, x) - G(\mathfrak{h}_2, x)\right|. \tag{3.6}$$

We conclude that *Assumption 1 is fulfilled* with $\Psi(\cdot) = |\cdot|,\ A = A_{\mathrm{f}},\ B = B_{\infty},\ \mathrm{a} = \mathrm{a}_{\mathrm{f}},\ \mathrm{b} = \mathrm{b}_{\infty}$ and $\mathrm{c} = 2$.

It is easily seen that a_{f} and b_{∞} are semi-metrics on $\mathfrak{H}$. We note also that $\xi_{\bullet} : \mathfrak{H} \to \mathbb{R}$ is P-a.s continuous in the topology generated by b_{∞}. Thus, if $\mathfrak{H} \subseteq \mathfrak{H}$ is totally bounded with respect to

$\mathrm{a}_f \vee \mathrm{b}_\infty$ and such that $\overline{A}_{\mathrm{f}} := \sup_{\mathfrak{h}\in\mathfrak{H}} A_{\mathrm{f}}(\mathfrak{h}) < \infty$, $\overline{B}_\infty := \sup_{\mathfrak{h}\in\mathfrak{H}} B_\infty(\mathfrak{h}) < \infty$, then we conclude that *Assumption 2 is verified.*

Thus, in the problems for which Assumption 3 is verified the machinery developed in Propositions 2-3 can be applied for $\bigl|\xi_{\mathfrak{h}}(n)\bigr|, \mathfrak{h} \in \mathfrak{H}$. We would like to emphasize, however, that problems studied below are not always related to the consideration of $\bigl|\xi_{\mathfrak{h}}(n)\bigr|, \mathfrak{h} \in \mathfrak{H}$ with $\mathfrak{H}$ being totally bounded, although such problems are also studied. The idea is to reduce them (if necessary) to those for which one of Propositions 2-3 can be used. For instance, we will be interested in finding upper functions for $\bigl|\xi_{\mathfrak{h}}(n)\bigr|$ on $\mathfrak{h} \in \mathfrak{H}$ not only for given n but mostly on $\mathbf{N} \times \mathfrak{H}$, where $\mathbf{N}$ is a given subset of $\mathbb{N}^*$. It will allow, in particular, to study generalized empirical processes with random number of summands.

However the application of Propositions 2-3 requires to compute the functions $\mathcal{E}$ or $\widehat{\mathcal{E}}$ and there is no a general recipe how to do it. The main goal of this section is to provide with rather general assumptions under which the latter quantities can be computed explicitly. As it was already mentioned in Introduction upper functions for random objects appear in various areas of probability theory and mathematical statistics. As the consequence the different nature of problems requires to specify the imposed assumptions. The assumptions presented below are oriented mostly to the problems arisen in mathematical statistics that definitely reflects author's scientific interests. However, some pure probabilistic results like the law of iterated logarithm and the law of logarithm will be established as well.

3.1. Problem formulation and examples. Main condition

In this section we find upper functions for several functionals of the generalized empirical process $\xi_{\mathfrak{h}}(n)$ defined in (3.1) under condition (3.2). We remark that the parameter $\mathfrak{h}$ may possess a composite structure and its components may have very different nature. In order to treat such situations it will be convenient for us to assume that for some $m \geq 1$

$$\mathfrak{H} = \mathfrak{H}_1 \times \cdots \times \mathfrak{H}_m, \tag{3.7}$$

where $\mathfrak{H}_j$, $j = \overline{1,m}$, be given sets. We will use the following notations. For any given $k = \overline{0,m}$ put

$$\mathfrak{H}_1^k = \mathfrak{H}_1 \times \cdots \times \mathfrak{H}_k, \quad \mathfrak{H}_{k+1}^m = \mathfrak{H}_{k+1} \times \cdots \times \mathfrak{H}_m,$$

with the agreement that $\mathfrak{H}_1^0 = \emptyset$, $\mathfrak{H}_{m+1}^m = \emptyset$. The elements of $\mathfrak{H}_1^k$ and $\mathfrak{H}_{k+1}^m$ will be denoted by $\mathfrak{h}^{(k)}$ and $\mathfrak{h}_{(k)}$ respectively. We will suppose that for any $j = \overline{k+1,m}$ the set $\mathfrak{H}_j$ is endowed with the semi-metric ϱ_j and the Borel measure $\varkappa_j$.

In the next two sections we find upper functions for $|\xi_{\mathfrak{h}}(n)|$ on on some subsets of $\mathfrak{H}$ (possibly depending on n!) and we will consider two cases.

Totally bounded case. In this case we will suppose that $\mathfrak{H}_j$ is totally bounded with respect to ϱ_j for any $j = \overline{k+1,m}$.

Partially totally bounded case. Here we first suppose that for some $p \geq 1$

$$(\mathcal{X}, \nu) = \bigl(\mathcal{X}_1 \times \cdots \times \mathcal{X}_p,\ \nu_1 \times \cdots \times \nu_p\bigr), \tag{3.8}$$

where $(\mathcal{X}_l, \nu_l)$ $l = \overline{1,p}$, are of measurable spaces and ν is the product measure.

Next we will assume that $\mathfrak{H}_m = \mathcal{X}_1$. As the consequence, the assumption, that $\mathfrak{H}_m$ is totally bounded, is too restrictive. In particular, it does not verified in the case $\mathcal{X} = \mathcal{X}_1 = \mathbb{R}^d$ which

appears in many examples. Before to start with the presentation of the results let us consider several examples.

Example 1. Density model. Let $K : \mathbb{R}^d \to \mathbb{R}$ be a given function and let

$$K_h(\cdot) = \left[\prod_{i=1}^{d} h_i\right]^{-1} K\left(\cdot/h_1, \ldots, \cdot/h_d\right), \quad h = (h_1, \ldots, h_d) \in (0,1]^d,$$

where, as previously, for two vectors $u, v \in \mathbb{R}^d$ the notation u/v denotes the coordinate-vice division.

Put $p = 1$, $m = d+1$, $k = d$, $\mathcal{X}_1 = \mathfrak{H}_{d+1} = \mathbb{R}^d$, $\mathfrak{H}_i = (0,1]$, $i = \overline{1,d}$ and consider for any $\mathfrak{h} = (h, x) \in \mathfrak{H} := (0,1]^d \times \mathbb{R}^d$

$$\xi_{\mathfrak{h}}(n) = \widehat{\xi}_{h,x}(n) := n^{-1} \sum_{i=1}^{n} \left[K_h\left(X_i - x\right) - \mathbb{E}_{\mathrm{f}}\left\{K_h\left(X_i - x\right)\right\}\right].$$

We have come to the well-known in nonparametric statistics kernel density estimation process. Here the function K is a kernel and the vector h is a multi-bandwidth.

Example 2. Regression model. Let $\varepsilon_i, i = \overline{1,n}$, be independent real random variables distributed on $\mathcal{I} \subseteq \mathbb{R}$ and such that $\mathbb{E}\varepsilon_i = 0$ for any $i = \overline{1,n}$. Let $Y_i, i = \overline{1,n}$, be independent d-dimension random vectors. The sequences $\left\{\varepsilon_i, i = \overline{1,n}\right\}$ and $\left\{Y_i, i = \overline{1,n}\right\}$ are assumed independent. Let $\mathcal{M}$ be a given set of $d \times d$ invertible matrices and let $\mathcal{I} \subseteq \mathbb{R}$ and $\mathcal{X}_1 \subseteq \mathbb{R}^d$ be given interval.

Put $p = 2$, $m = d+2$, $k = d$, $\mathcal{X}_1 = \mathfrak{H}_{d+2} = \mathbb{R}^d$, $\mathcal{X}_2 = \mathcal{I}$, $\mathfrak{H}_j = (0,1]$, $j = \overline{1,d}$ and $\mathfrak{H}_{d+1} = \mathcal{M}$. Consider for any $\mathfrak{h} = (h, M, x) \in \mathfrak{H} := (0,1]^d \times \mathcal{M} \times \mathbb{R}^d$

$$\xi_{\mathfrak{h}}(n) = \widetilde{\xi}_{h,M,x}(n) := n^{-1} |\det(M)| \sum_{i=1}^{n} K_h\Big[M(Y_i - x)\Big]\varepsilon_i.$$

The family of random fields $\left\{\widetilde{\xi}_{h,M,x}(n),\ x, h, M \in (0,1]^d \times \mathcal{M} \times \mathbb{R}^d\right\}$ appears in non-parametric regression under single index hypothesis, Stone (1985).

If $\mathcal{I}$ is bounded interval, i.e. ε_i are bounded random variables, then (3.5) and (3.6) hold and the results from Section 2 are applicable. However this assumption is too restrictive and it does not satisfied even in the classical gaussian regression. At the first glance it is seemed that if $\mathcal{I} = \mathbb{R}$ Propositions 2-3 are not applicable here. Although the aforementioned problem lies beyond of the scope of the paper, let us briefly discuss how to reduce it to the problem in which the machinery developed in Propositions 2-3 can be applied.

Some generalizations. Let $\left(\varepsilon_i, i = \overline{1,n}\right)$ be the sequence of independent real-valued random variables such that $\mathbb{E}\varepsilon_i = 0$ (later on for simplicity we assume that ε_i has symmetric distribution) and $\mathbb{E}\varepsilon_i^2 =: \sigma_i^2 < \infty$. Let $\bar{X}_i,\ i = \overline{1,n}$, be a the collection of $\bar{\mathcal{X}}$-valued *independent* random elements and suppose also that $\left(\bar{X}_i, i = \overline{1,n}\right)$ and $\left(\varepsilon_i, i = \overline{1,n}\right)$ are independent. Consider the generalized empirical process

$$\bar{\xi}_{\mathfrak{h}}(n) = n^{-1} \sum_{i=1}^{n} \bar{G}\big(\mathfrak{h}, \bar{X}_i\big)\varepsilon_i, \quad \mathfrak{h} \in \mathfrak{H},$$

where, as previously, $\bar{G} : \mathfrak{H} \times \mathcal{X} \to \mathbb{R}$ be a given mapping satisfying (3.2). For any $y > 0$ define

$$\bar{\xi}_{\mathfrak{h}}(n, y) = n^{-1} \sum_{i=1}^{n} \bar{G}\big(\mathfrak{h}, \bar{X}_i\big)\varepsilon_i 1_{[-y,y]}(\varepsilon_i), \quad \eta_n(y) = \sup_{i=\overline{1,n}} \big|\varepsilon_i\big|\big[1 - 1_{[-y,y]}(\varepsilon_i)\big].$$

Obviously, for any $y>0$

$$\bar{\xi}_{\mathfrak{h}}(n,y)=n^{-1}\sum_{i=1}^{n}\Big[G_y\big(\mathfrak{h},X_i\big)-\mathbb{E}_{\mathrm{f}}G_y(\mathfrak{h},X_i)\Big],\quad X_i=\big(\bar{X}_i,\varepsilon_i\big),$$

where $G_y(\mathfrak{h},x)=\bar{G}(\mathfrak{h},\bar{x})u1_{[-y,y]}(u)$, $x=(\bar{x},u)\in\mathcal{X}:=\bar{\mathcal{X}}\times\mathbb{R}$, $\mathfrak{h}\in\mathfrak{H}$. Since G_y is bounded for any $y>0$ the inequalities (3.3) and (3.4) hold and, analogously to (3.5) and (3.6), we have

$$A^2_{\mathrm{f}}(\mathfrak{h})=2n^{-2}\sum_{i=1}^{n}\sigma_i^2\mathbb{E}_f\bar{G}^2(\mathfrak{h},\bar{X}_i),\qquad \mathrm{a}^2_{\mathrm{f}}(\mathfrak{h}_1,\mathfrak{h}_2)=2n^{-2}\sum_{i=1}^{n}\sigma_i^2\mathbb{E}_f\big(\bar{G}(\mathfrak{h}_1,\bar{X}_i)-G(\mathfrak{h}_2,\bar{X}_i)\big)^2;$$
$$B_{\infty}(\mathfrak{h})=(4y/3)n^{-1}\sup_{x\in\mathcal{X}}\big|\bar{G}(\mathfrak{h},\bar{x})\big|,\qquad \mathrm{b}_{\infty}(\mathfrak{h}_1,\mathfrak{h}_2)=(4/3)yn^{-1}\sup_{x\in\mathcal{X}}\big|\bar{G}(\mathfrak{h}_1,\bar{x})-\bar{G}(\mathfrak{h}_2,\bar{x})\big|.$$

Let also $\mathfrak{H}\subseteq\mathfrak{H}$ be such that the results obtained in Propositions 2-3 are applicable to $\big|\bar{\xi}_{\mathfrak{h}}(n,y)\big|$ on $\mathfrak{H}$ for any $y>0$. It is extremely important to emphasize that neither $A_{\mathrm{f}}(\cdot)$ nor $\mathrm{a}_{\mathrm{f}}(\cdot,\cdot)$ depend on y.

This yields, in view of Theorems 1 and 3 below, that upper functions for $\big|\bar{\xi}_{\mathfrak{h}}(y)\big|$, $\mathfrak{h}\in\mathfrak{H}$ (for brevity $V(\mathfrak{h},y)$ and $U_q(\mathfrak{h},y),q\geq1$) can be found in the form:

$$V(\mathfrak{h},y)=V_1(\mathfrak{h})+yV_2(\mathfrak{h}),\quad U_q(\mathfrak{h},y)=U_{q,1}(\mathfrak{h})+yU_{q,2}(\mathfrak{h}).$$

It means that we are able to bound from above any $y>0$

$$\mathbb{P}_f\left\{\sup_{\mathfrak{h}\in\mathfrak{H}}\big[\big|\bar{\xi}_{\mathfrak{h}}(n,y)\big|-V(\mathfrak{h},y)\big]>0\right\},\quad \mathbb{E}_f\left\{\sup_{\mathfrak{h}\in\mathfrak{H}}\big[\big|\bar{\xi}_{\mathfrak{h}}(n,y)\big|-U_q(\mathfrak{h},y)\big]\right\}_+^q$$

Moreover, we obviously have for any $y>0$

$$\mathbb{P}_f\left\{\sup_{\mathfrak{h}\in\mathfrak{H}}\big[\big|\bar{\xi}_{\mathfrak{h}}(n)\big|-V(\mathfrak{h},y)\big]>0\right\}\leq\mathbb{P}_f\left\{\sup_{\mathfrak{h}\in\mathfrak{H}}\big[\big|\bar{\xi}_{\mathfrak{h}}(n,y)\big|-V(\mathfrak{h},y)\big]>0\right\}+\mathbb{P}_f\left\{\eta_n(y)>0\right\};$$
$$\mathbb{E}_f\left\{\sup_{\mathfrak{h}\in\mathfrak{H}}\big[\big|\bar{\xi}_{\mathfrak{h}}(n)\big|-U_q(\mathfrak{h},y)\big]\right\}_+^q\leq\mathbb{E}_f\left\{\sup_{\mathfrak{h}\in\mathfrak{H}}\big[\big|\bar{\xi}_{\mathfrak{h}}(n,y)\big|-U_q(\mathfrak{h},y)\big]\right\}_+^q+\left(\sup_{\mathfrak{h}\in\mathfrak{H}}\overline{G}_{\infty}(\mathfrak{h})\right)^q\mathbb{E}\left(\eta_n(y)\right)^q.$$

Typically, $V(\cdot,y)=V^{(n)}(\cdot,y)$ and $U_q(\cdot,y)=U_q^{(n)}(\cdot,y)$ and $V_2^{(n)}(\cdot)\ll V_1^{(n)}$ and $U_{q,2}^{(n)}(\cdot)\ll U_{q,1}^{(n)}$ for all n large enough. It allows to choose $y=y_n$ in optimal way, i.e. to balance both terms in latter inequalities, that usually leads to sharp upper functions $V_1^{(n)}(\cdot)+y_nV_2^{(n)}(\cdot)$ and $U_{q,1}^{(n)}(\cdot)+y_nU_{q,2}^{(n)}(\cdot)$.

Main Assumption Now let us come back to the consideration of generalized empirical processes obeying (3.2). Assumption 4 below is the main tool allowing us to compute explicitly upper functions. Introduce the following notation: for any $\mathfrak{h}^{(k)}\in\mathfrak{H}_1^k$

$$\mathbf{G}_{\infty}\big(\mathfrak{h}^{(k)}\big)=\sup_{\mathfrak{h}_{(k)}\in\mathfrak{H}_{k+1}^m}\sup_{x\in\mathcal{X}}|G(\mathfrak{h},x)|,$$

and let $G_{\infty}:\mathfrak{H}_1^k\to\mathbb{R}_+$ be any mapping satisfying

$$\mathbf{G}_{\infty}\big(\mathfrak{h}^{(k)}\big)\leq G_{\infty}\big(\mathfrak{h}^{(k)}\big),\ \ \forall\mathfrak{h}^{(k)}\in\mathfrak{H}_1^k. \tag{3.9}$$

Let $\{\mathfrak{H}_j(n)\subset\mathfrak{H}_j,\ n\geq1\},j=\overline{1,k}$, be a sequence of sets and denote $\mathfrak{H}_1^k(n)=\mathfrak{H}_1(n)\times\cdots\mathfrak{H}_k(n)$. Set for any $n\geq1$

$$\underline{G}_n=\inf_{\mathfrak{h}^{(k)}\in\mathfrak{H}_1^k(n)}G_{\infty}\big(\mathfrak{h}^{(k)}\big),\quad \overline{G}_n=\sup_{\mathfrak{h}^{(k)}\in\mathfrak{H}_1^k(n)}G_{\infty}\big(\mathfrak{h}^{(k)}\big).$$

For any $n \geq 1$, $j = \overline{1,k}$ and any $\mathfrak{h}_j \in \mathfrak{H}_j(n)$ define

$$G_{j,n}(\mathfrak{h}_j) = \sup_{\mathfrak{h}_1 \in \mathfrak{H}_1(n), \ldots, \mathfrak{h}_{j-1} \in \mathfrak{H}_{j-1}(n), \mathfrak{h}_{j+1} \in \mathfrak{H}_{j+1}(n), \ldots, \mathfrak{h}_k \in \mathfrak{H}_k(n)} G_\infty\big(\mathfrak{h}^{(k)}\big), \qquad \underline{G}_{j,n} = \inf_{\mathfrak{h}_j \in \mathfrak{H}_j(n)} G_{j,n}(\mathfrak{h}_j).$$

Noting that $\big|\ln(t_1) - \ln(t_2)\big|$ is a metric on $\mathbb{R}_+ \setminus \{0\}$, we equip $\mathfrak{H}_1^k(n)$ with the following semi-metric. For any $n \geq 1$ and any $\hat{\mathfrak{h}}^{(k)}, \bar{\mathfrak{h}}^{(k)} \in \mathfrak{H}_1^k(n)$ set

$$\varrho_n^{(k)}\Big(\hat{\mathfrak{h}}^{(k)}, \bar{\mathfrak{h}}^{(k)}\Big) = \max_{j=\overline{1,k}} \Big|\ln\big\{G_{j,n}(\hat{\mathfrak{h}}_j)\big\} - \ln\big\{G_{j,n}(\bar{\mathfrak{h}}_j)\big\}\Big|,$$

where $\hat{\mathfrak{h}}_j, \bar{\mathfrak{h}}_j,\ j = \overline{1,k}$, are the coordinates of $\hat{\mathfrak{h}}^{(k)}$ and $\bar{\mathfrak{h}}^{(k)}$ respectively.

Assumption 4. (**i**) $0 < \underline{G}_n \leq \overline{G}_n < \infty$ *for any* $n \geq 1$ *and for any* $j = \overline{1,k}$

$$\frac{G_\infty\big(\mathfrak{h}^{(k)}\big)}{\underline{G}_n} \geq \frac{G_{j,n}(\mathfrak{h}_j)}{\underline{G}_{j,n}}, \quad \forall \mathfrak{h}^{(k)} = (\mathfrak{h}_1, \ldots, \mathfrak{h}_k) \in \mathfrak{H}_1^k(n), \ \ \forall n \geq 1;$$

(**ii**) *There exist functions* $L_j : \mathbb{R}_+ \to \mathbb{R}_+,\ D_j : \mathbb{R}_+ \to \mathbb{R}_+,\ j = 0, k+1, \ldots, m$, *satisfying* L_j *non-decreasing and bounded on each bounded interval,* $D_j \in \mathbb{C}^1\big(\mathbb{R}\big),\ D(0) = 0$, *and such that*

$$\begin{aligned} \big\|G(\mathfrak{h}, \cdot) - G(\overline{\mathfrak{h}}, \cdot)\big\|_\infty \ &\leq \ \Big\{G_\infty\Big(\mathfrak{h}^{(k)}\Big) \vee G_\infty\Big(\overline{\mathfrak{h}}^{(k)}\Big)\Big\} D_0\Big\{\varrho_n^{(k)}\Big(\mathfrak{h}^{(k)}, \overline{\mathfrak{h}}^{(k)}\Big)\Big\} \\ &+ \ \sum_{j=k+1}^{m} L_j\Big\{G_\infty\Big(\mathfrak{h}^{(k)}\Big) \vee G_\infty\Big(\overline{\mathfrak{h}}^{(k)}\Big)\Big\} D_j\Big(\varrho_j\big(\mathfrak{h}_j, \mathfrak{h}'_j\big)\Big), \end{aligned}$$

for any $\mathfrak{h}, \overline{\mathfrak{h}} \in \mathfrak{H}_1^k(n) \times \mathfrak{H}_{k+1}^m$ *and* $n \geq 1$.

We remark that Assumption 4 (**i**) is automatically fulfilled if $k = 1$.

Remark 6. *If* $n \geq 1$ *is fixed or* $\mathfrak{H}_j(n),\ j = \overline{1,k}$, *are independent on* n, *for example* $\mathfrak{H}_j(n) = \mathfrak{H}_j,\ j = \overline{1,k}$, *for all* $n \geq 1$ *then upper functions for* $|\xi_h(n)|$ *can be derived under Assumption 4. However, if we are interested in finding of upper functions for* $|\xi_h(n)|$ *when* n *is varying, we cannot do it in general without specifying the dependence of* $\mathfrak{H}_j(n),\ j = \overline{1,k}$, *on* n.

In view of latter remark we will seek upper functions for $|\xi_h(n)|$ when $\mathfrak{h} \in \widetilde{\mathfrak{H}}(n) := \widetilde{\mathfrak{H}}_1^k(n) \times \mathfrak{H}_{k+1}^m$. Here $\widetilde{\mathfrak{H}}_1^k(n) = \widetilde{\mathfrak{H}}_1(n) \times \cdots \widetilde{\mathfrak{H}}_k(n)$ and $\Big\{\widetilde{\mathfrak{H}}_j(n) \subset \mathfrak{H}_j(n),\ n \geq 1\Big\}, j = \overline{1,k}$, be a sequence of sets satisfying additional restriction. We will not be tending here to the maximal generality and complete Assumption 4 by the following condition.

Assumption 5. *For any* $\mathbf{m} \in \mathbb{N}^*$ *there exists* $n[\mathbf{m}] \in \{\mathbf{m}, \mathbf{m}+1, \ldots, 2\mathbf{m}\}$ *such that*

$$\bigcup_{n \in \{\mathbf{m}, \mathbf{m}+1, \ldots, 2\mathbf{m}\}} \widetilde{\mathfrak{H}}_1^k(n) \subseteq \mathfrak{H}_1^k\big(n[\mathbf{m}]\big).$$

We note that Assumption 5 obviously holds if for any $j = \overline{1,k}$ the sequence $\Big\{\widetilde{\mathfrak{H}}_j(n),\ n \geq 1\Big\}$ is increasing/decreasing sequence of sets.

3.2. Totally bounded case

The objective is to find upper functions for $|\xi_{\mathfrak{h}}(n)|$ under Assumption 4 enforced, if necessary, by Assumption 5 and the condition imposed on the entropies of the sets $\mathfrak{H}_j,\ j = \overline{k+1, m}$.

3.2.1. Assumptions and main result

The following condition will be additionally imposed in this section.

Assumption 6. *Suppose that (3.7) holds and there exist $N, R < \infty$ such that for any $\varsigma > 0$ and any $j = \overline{k+1, m}$*

$$\mathfrak{E}_{\mathfrak{H}_j, \varrho_j}(\varsigma) \le N \left[\log_2 \left\{R/\varsigma\right\}\right]_+ ,$$

where, as previously, $\mathfrak{E}_{\mathfrak{H}_j, \varrho_j}$ denotes the entropy of $\mathfrak{H}_j$ measured in ϱ_j.

We remark that Assumption 6 is fulfilled, in particular, when $\big(\mathfrak{H}_j, \varrho_j, \varkappa_j\big)$, $j = \overline{k+1, m}$, are bounded and satisfy doubling condition. Note also that this assumption can be considerably weakened, see discussion after Theorem 1.

Notations Let $3 \le \mathbf{n_1} \le \mathbf{n_2} < 2\mathbf{n_1}$ be fixed and set $\widetilde{\mathbf{N}} = \{\mathbf{n_1}, \dots, \mathbf{n_2}\}$. For any $\mathfrak{h} \in \mathfrak{H}$ set

$$F_{\mathbf{n_2}}(\mathfrak{h}) = \begin{cases} \sup_{i=\overline{1,\mathbf{n_2}}} \mathbb{E}_{\mathrm{f}} \big|G(\mathfrak{h}, X_i)\big|, & \mathbf{n_1} \ne \mathbf{n_2}; \\ (\mathbf{n_2})^{-1} \sum_{i=1}^{\mathbf{n_2}} \mathbb{E}_{\mathrm{f}} \big|G(\mathfrak{h}, X_i)\big|, & \mathbf{n_1} = \mathbf{n_2}, \end{cases}$$

and remark that if additionally X_i, $i \ge 1$, are identically distributed then we have the same definition of $F_{\mathbf{n_2}}(\cdot)$ in both cases. We note that

$$F_{\mathbf{n_2}} := \sup_{n \in \widetilde{\mathbf{N}}} \sup_{\mathfrak{h} \in \widetilde{\mathfrak{H}}(n)} F_{\mathbf{n_2}}(\mathfrak{h}) \le \sup_{n \in \widetilde{\mathbf{N}}} \overline{G}_n < \infty$$

in view of Assumption 4 (**i**). Let $\boldsymbol{b} > 1$ be fixed and put

$$\mathbf{n} = \begin{cases} \mathbf{n_1}, & \mathbf{n_1} = \mathbf{n_2}; \\ n[\mathbf{n_1}], & \mathbf{n_1} \ne \mathbf{n_2}, \end{cases} \qquad \beta = \begin{cases} 0, & \mathbf{n_1} = \mathbf{n_2}; \\ \boldsymbol{b}, & \mathbf{n_1} \ne \mathbf{n_2}, \end{cases}$$

where, remind, that $n[\cdot]$ is defined in Assumption 5.

Define $\widehat{L}_j(z) = \sup_{u \le z} \max\left\{u^{-1} L_j(u), 1\right\}$ and $\mathcal{L}^{(k)}(z) = \sum_{j=k+1}^{m} \log_2 \left\{\widehat{L}_j\,(2z)\right\}$ and introduce the following quantities: for any $\mathfrak{h}^{(k)} \in \mathfrak{H}_1^k$ and any $q > 0$

$$P\left(\mathfrak{h}^{(k)}\right) = (36k\delta_*^{-2} + 6) \ln\left(1 + \ln\left\{2\underline{G}_{\mathbf{n}}^{-1} G_\infty\big(\mathfrak{h}^{(k)}\big)\right\}\right) + 36N\delta_*^{-2} \mathcal{L}^{(k)}\Big(G_\infty\big(\mathfrak{h}^{(k)}\big)\Big) + 18 C_{N,R,m,k};$$

$$M_q\left(\mathfrak{h}^{(k)}\right) = \big(72k\delta_*^{-2} + 2.5q + 1.5\big) \ln\Big(2\underline{G}_{\mathbf{n}}^{-1} G_\infty\big(\mathfrak{h}^{(k)}\big)\Big) + 72N\delta_*^{-2} \mathcal{L}^{(k)}\Big(G_\infty\big(\mathfrak{h}^{(k)}\big)\Big) + 36 C_{N,R,m,k}.$$

Here δ_* it is the smallest solution of the equation $(48\delta)^{-1} s^*(\delta) = 1$, where, remind, $s^*(\delta) = (6/\pi^2)\big(1 + [\ln \delta]^2\big)^{-1}$, $\delta \ge 0$. The quantities N, R are defined in Assumption 6.

The explicit expression of the constant $C_{N,R,m,k}$, as well as explicit expressions of the constants λ_1, λ_2 and $C_{D,\boldsymbol{b}}$ used in the description of the results below, are given in Section 6.1.2 which precedes the proof of Theorem 1.

Result For any $\mathbf{r} \in \overline{\mathbb{N}}$ put $F_{\mathbf{n_2},\mathbf{r}}(\mathfrak{h}) = \max\left[F_{\mathbf{n_2}}(\mathfrak{h}), e^{-\mathbf{r}}\right]$ and define for any $\mathfrak{h} \in \mathfrak{H}$, $u \geq 0$ and $q > 0$

$$\begin{aligned}\mathcal{V}_{\mathbf{r}}^{(u)}(n,\mathfrak{h}) &= \lambda_1\sqrt{G_\infty\big(\mathfrak{h}^{(k)}\big)\Big(F_{\mathbf{n_2},\mathbf{r}}(\mathfrak{h})n^{-1}\Big)\Big(P\big(\mathfrak{h}^{(k)}\big)+2\ln\left\{1+\left|\ln\left(F_{\mathbf{n_2},\mathbf{r}}(\mathfrak{h})\right)\right|\right\}+u\Big)}\\ &\quad+\lambda_2 G_\infty\big(\mathfrak{h}^{(k)}\big)\Big(n^{-1}\ln^\beta(n)\Big)\Big(P\big(\mathfrak{h}^{(k)}\big)+2\ln\left\{1+\left|\ln\left(F_{\mathbf{n_2},\mathbf{r}}(\mathfrak{h})\right)\right|\right\}+u\Big);\\ \mathcal{U}_{\mathbf{r}}^{(u,q)}(n,\mathfrak{h}) &= \lambda_1\sqrt{G_\infty\big(\mathfrak{h}^{(k)}\big)\Big(F_{\mathbf{n_2},\mathbf{r}}(\mathfrak{h})n^{-1}\Big)\Big(M_q\big(\mathfrak{h}^{(k)}\big)+2\ln\left\{1+\left|\ln\left(F_{\mathbf{n_2},\mathbf{r}}(\mathfrak{h})\right)\right|\right\}+u\Big)}\\ &\quad+\lambda_2 G_\infty\big(\mathfrak{h}^{(k)}\big)\Big(n^{-1}\ln^\beta(n)\Big)\Big(M_q\big(\mathfrak{h}^{(k)}\big)+2\ln\left\{1+\left|\ln\left(F_{\mathbf{n_2},\mathbf{r}}(\mathfrak{h})\right)\right|\right\}+u\Big).\end{aligned}$$

Theorem 1. *Let Assumptions 4 and 6 be fulfilled. If* $\mathbf{n_1} \neq \mathbf{n_2}$ *suppose additionally that Assumption 5 holds. Then for any* $\mathbf{r} \in \mathbb{N}$, $\boldsymbol{b} > 1$ $u \geq 1$ *and* $q \geq 1$

$$\mathbb{P}_{\mathrm{f}}\left\{\sup_{n\in\widetilde{\mathbf{N}}}\sup_{\mathfrak{h}\in\widetilde{\mathfrak{H}}(n)}\Big[\big|\xi_{\mathfrak{h}}(n)\big|-\mathcal{V}_{\mathbf{r}}^{(u)}(n,\mathfrak{h})\Big]\geq 0\right\}\leq 2419\,e^{-u};$$

$$\mathbb{E}_{\mathrm{f}}\left\{\sup_{n\in\widetilde{\mathbf{N}}}\sup_{\mathfrak{h}\in\widetilde{\mathfrak{H}}(n)}\Big[\big|\xi_{\mathfrak{h}}(n)\big|-\mathcal{U}_{\mathbf{r}}^{(u,q)}(n,\mathfrak{h})\Big]\right\}_+^q\leq c_q\left[\sqrt{(\mathbf{n_1})^{-1}F_{\mathbf{n_2}}\underline{G}_{\mathbf{n}}}\vee\Big((\mathbf{n_1})^{-1}\ln^\beta(\mathbf{n_2})\underline{G}_{\mathbf{n}}\Big)\right]^q e^{-u},$$

where $c_q = 2^{(7q/2)+5}3^{q+4}\Gamma(q+1)(C_{D,\boldsymbol{b}})^q$.

Remark 7. *The inspection of the proof of the theorem allows us to assert that Assumption 6 can be weakened. The condition that is needed in view of the used technique: for some* $\alpha \in (0,1)$, $L < \infty$

$$\sup_{\varsigma>0}\varsigma^{-\alpha}\mathfrak{E}_{\mathfrak{H}_j,\varrho_j}(\varsigma)\leq L,\quad j=\overline{k+1,m}. \tag{3.10}$$

In particular, it allows to consider the generalized empirical processes indexed by the sets of smooth functions. However the latter assumption does not permit to express upper functions explicitly as it is done in Theorem 1. This explains why we prefer to state our results under Assumption 6.

Several other remarks are in order.

1^0. First we note that the results presented in the theorem are obtained without any assumption imposed on the densities f_i, $i \geq 1$. In particular, found upper functions remain finite even if the densities f_i, $i \geq 1$ are unbounded.

2^0. Next, putting $\mathbf{r} = +\infty$ we get the results of the theorem with $F_{\mathbf{n_2},\mathbf{r}}(\cdot) = F_{\mathbf{n_2}}(\cdot)$. It improves the first terms in the expressions of $\mathcal{V}_{\mathbf{r}}^{(u)}(\cdot,\cdot)$ and $\mathcal{U}_{\mathbf{r}}^{(u,q)}(\cdot,\cdot)$, however the second terms may explode if $F_{\mathbf{n_2}}(\mathfrak{h}) = 0$ for some $\mathfrak{h} \in \mathfrak{H}$. The latter fact explains the necessity to "truncate" $F_{\mathbf{n_2}}(\cdot)$ from below, i.e. to consider $F_{\mathbf{n_2},\mathbf{r}}(\cdot)$ instead of $F_{\mathbf{n_2}}(\cdot)$.

3.2.2. Law of iterated logarithm

Our goal here is to use the first assertion of Theorem 1 in order to establish a non-asymptotical version of the law of iterated logarithm for

$$\eta_{\mathfrak{h}^{(k)}}(n) := \sup_{\mathfrak{h}_{(k)}\in\mathfrak{H}_{k+1}^m}\big|\xi_{\mathfrak{h}}(n)\big|.$$

Let us suppose that for some $\mathfrak{c} > 0$, $\mathfrak{b} > 0$

$$\mathfrak{c} \leq \underline{G}_n \leq \overline{G}_n \leq \mathfrak{c}n^{\mathfrak{b}},\quad \forall n \geq 1. \tag{3.11}$$

We would like to emphasize that the restriction $\underline{G}_n \geq \mathfrak{c}$ is imposed for the simplicity of the notations and the results presented below are valid if $\underline{G}_n$ decreases to zero polynomially in n.

Moreover we will assume that

$$\sup_{n\geq 1}\sup_{\mathfrak{h}\in\widetilde{\mathfrak{H}}(n)}\sup_{i\geq 1}\mathbb{E}_{\mathrm{f}}\big|G(\mathfrak{h},X_i)\big| =: \mathbf{F} < \infty. \tag{3.12}$$

We will see that the latter condition is checked in various particular problems if the densities f_i, $i \geq 1$ are uniformly bounded. Suppose finally that for some $\mathfrak{a} > 0$

$$\mathcal{L}^{(k)}(z) \leq \mathfrak{a}\ln\big\{1+\ln(z)\big\}, \quad \forall z \geq 3. \tag{3.13}$$

For any $a > 0$ and $n \geq 3$ define

$$\overline{\mathfrak{H}}_1^k(n,a) = \widetilde{\mathfrak{H}}_1^k(n) \cap \Big\{\mathfrak{h}^{(k)} : \ G_\infty\big(\mathfrak{h}^{(k)}\big) \leq n\big[\ln(n)\big]^{-a}\Big\}.$$

Theorem 2. *Let Assumptions 4, 5 and 6 be fulfilled and suppose additionally that (3.11), (3.12) and (3.13) hold. Then there exists $\Upsilon > 0$ such that for any $\mathbf{j} \geq 3$ and any $a > 2$*

$$\mathbb{P}_{\mathrm{f}}\left\{\sup_{n\geq\mathbf{j}}\sup_{\mathfrak{h}^{(k)}\in\overline{\mathfrak{H}}_1^k(n,a)}\left[\frac{\sqrt{n}\,\eta_{\mathfrak{h}^{(k)}}(n)}{\sqrt{G_\infty\big(\mathfrak{h}^{(k)}\big)\ln\big(1+\ln(n)\big)}}\right] \geq \Upsilon\right\} \leq \frac{2419}{\ln(\mathbf{j})}.$$

The explicit expression of the constant Υ can be easily derived but it is quite cumbersome and we omit its derivation.

Remark 8. *The inspection of the proof of the theorem shows that for any $y \geq 0$ one can find $0 < \Upsilon(y) < \infty$ such that the assertion of the theorem remains true if one replaces Υ by $\Upsilon(y)$ and the right hand side of the obtained inequality by $2419\big[\ln(\mathbf{j})\big]^{-(1+y)}$. It makes reasonable the consideration of small values of $\mathbf{j}$.*

The simple corollary of Theorem 2 is the law of iterated logarithm:

$$\limsup_{n\to\infty}\sup_{\mathfrak{h}^{(k)}\in\overline{\mathfrak{H}}_1^k(n,a)}\left[\frac{\sqrt{n}\,\eta_{\mathfrak{h}^{(k)}}(n)}{\sqrt{G_\infty\big(\mathfrak{h}^{(k)}\big)\ln\ln(n)}}\right] \leq \Upsilon, \quad \mathbb{P}_{\mathrm{f}}-\text{a.s.} \tag{3.14}$$

3.3. Partially totally bounded case

We begin this section with the following definition used in the sequel. Let $\mathbb{T}$ be a set equipped with a semi-metric $\mathfrak{d}$ and let $\mathfrak{n} \in \mathbb{N}^*$ be fixed.

Definition 1. *We say that $\{\mathbb{T}_{\mathbf{i}} \subset \mathbb{T},\ \mathbf{i} \in \mathbf{I}\}$ is $\mathfrak{n}$-totally bounded cover of $\mathbb{T}$ if*

- $\mathbb{T} = \cup_{\mathbf{i}\in\mathbf{I}}\mathbb{T}_{\mathbf{i}}$ *and* $\mathbf{I}$ *is countable;*
- $\mathbb{T}_{\mathbf{i}}$ *is totally bounded for any* $\mathbf{i} \in \mathbf{I}$*;*
- $card\Big(\{\mathbf{k} \in \mathbf{I} : \ \mathbb{T}_{\mathbf{i}} \cap \mathbb{T}_{\mathbf{k}} \neq \emptyset\}\Big) \leq \mathfrak{n}$ *for any* $\mathbf{i} \in \mathbf{I}$.

Let us illustrate the above definition by some examples.

Let $\mathbb{T} = \mathbb{R}^d,\ d \geq 1$. Then any countable partition of $\mathbb{R}^d$ consisted of bounded sets forms 1-totally bounded cover of $\mathbb{R}^d$. Note, however, that the partitions will not be suitable choice for particular problems studied later. We will be mostly interested in $\mathfrak{n}$-totally bounded covers satisfying the following separation property: there exists $\mathfrak{r} > 0$ such that for all $\mathbf{i}, \mathbf{k} \in \mathbf{I}$ satisfying $\mathbb{T}_\mathbf{i} \cap \mathbb{T}_\mathbf{k} = \emptyset$

$$\inf_{t_1 \in \mathbb{T}_\mathbf{i}, t_2 \in \mathbb{T}_\mathbf{k}} \mathfrak{d}(t_1, t_2) \geq \mathfrak{r}. \tag{3.15}$$

Let us return to $\mathbb{R}^d$ that we equip with the metric generated by the supremum norm. Denote by $\mathbb{B}_r(t),\ t \in \mathbb{R}^d, r > 0$, the closed ball in this metric with the radius r and the center t. For given $\mathfrak{r} > 0$ consider the collection $\left\{\mathbb{B}_{\frac{\mathfrak{r}}{2}}(\mathfrak{r}\mathbf{i}),\ \mathbf{i} \in \mathbb{Z}^d\right\}$, where we understand $\mathfrak{r}\mathbf{i}$ as coordinate-wise multiplication. It is easy to check that this collection is 3^d-totally bounded cover of $\mathbb{R}^d$ satisfying (3.15).

We would like to emphasize that $\mathfrak{n}$-totally bounded covers satisfying the separation property can be often constructed when $\mathbb{T}$ is a homogenous metric space endowed with the Borel measure obeying doubling condition. Some useful results for this construction can be found in the recent paper Coulhon et al. (2011), where such spaces were scrutinized.

We finish the discussion about $\mathfrak{n}$-totally bounded covers with the following notation: for any $t \in \mathbb{T}$ put

$$\mathbb{T}(t) = \bigcup_{\mathbf{i} \in \mathbf{I}:\ \mathbb{T}_\mathbf{i} \ni t} \ \bigcup_{\mathbf{k} \in \mathbf{I}:\ \mathbb{T}_\mathbf{i} \cap \mathbb{T}_\mathbf{k} \neq \emptyset} \mathbb{T}_\mathbf{k}.$$

3.3.1. Assumptions and main result

Throughout this section we will assume that the representation (3.8) holds and the elements of $\mathcal{X}_l,\ l = \overline{1,p}$, will be denoted by x_l . We keep all notations from previous section and replace Assumption 6 by the following conditions.

Assumption 7. (**i**) *Let (3.7) and (3.8) hold with $\mathcal{X}_1 = \mathfrak{H}_m$ and for some $\mathfrak{n} \in \mathbb{N}^*$ there exists a collection $\left\{\mathrm{H}_{m,\mathbf{i}},\ \mathbf{i} \in \mathbf{I}\right\}$ being the $\mathfrak{n}$-totally bounded cover of $\mathfrak{H}_m$ satisfying for some $N, R < \infty$*

$$\mathfrak{E}_{\mathrm{H}_{m,\mathbf{i}}, \varrho_m}(\varsigma) \leq N \left[\log_2 \left\{R/\varsigma\right\}\right]_+, \quad \forall \mathbf{i} \in \mathbf{I},\ \ \forall \varsigma > 0.$$

(**ii**) *For any $\varsigma > 0$*

$$\mathfrak{E}_{\mathfrak{H}_j, \varrho_j}(\varsigma) \leq N \left[\log_2 \left\{R/\varsigma\right\}\right]_+, \quad \forall j = \overline{k+1, m-1}.$$

Usually one can construct many $\mathfrak{n}$-totally bounded covers satisfying Assumption 7 (**i**). The condition below restricts this choice and relates it to properties of the mapping $G(\cdot,\cdot)$ describing generalized empirical process.

Assumption 8. *For any $n \geq 1$ and any $\mathfrak{h} = (\mathfrak{h}_1, \ldots, \mathfrak{h}_m) \in \mathfrak{H}(n)$*

$$\sup_{x \in \mathcal{X}:\ x_1 \notin \mathfrak{H}_m(\mathfrak{h}_m)} |G(\mathfrak{h}, x)| \leq n^{-1} G_\infty\big(\mathfrak{h}^{(k)}\big).$$

We would like to emphasize that in order to satisfy Assumption 8 in particular examples, the $\mathfrak{n}$-totally bounded cover $\left\{\mathrm{H}_{m,\mathbf{i}},\ \mathbf{i} \in \mathbf{I}\right\}$ should usually possess the separation property. Indeed, one of the typical examples, where Assumption 8 is fulfilled, is the following: there exist $\gamma > 0$ such that for $G(x, \mathfrak{h}) = 0$ for any $x \in \mathcal{X},\ \mathfrak{h} \in \mathfrak{H}$, satisfying $\rho_m(x_1, \mathfrak{h}_m) \geq \gamma$.

Result For any $i=\overline{1,n}$ we denote $X_i=\big(X_{1,i},\ldots,X_{p,i}\big)$,

$$f_{1,i}(x_1)=\int_{\mathcal{X}_2\times\cdots\times\mathcal{X}_p} f_i(x_1,\ldots,x_p)\prod_{l=2}^{p}\nu_l\big(\mathrm{d}x_l\big).$$

and if $\mathcal{X}=\mathcal{X}_1$ ($p=1$) then we put $X_{1,i}=X_i$ and $f_{1,i}=f_i$.

Put for any $n\geq 1$, $v>0$ and any $\mathfrak{h}_m\in\mathfrak{H}_m$

$$\mathfrak{L}_{n,v}(\mathfrak{h}_m)=-\ln\left(\left[n^{-1}\sum_{i=1}^{n}\int_{\mathfrak{H}_m(\mathfrak{h}_m)}f_{1,i}(x)\nu_1\big(\mathrm{d}x\big)\right]\vee n^{-v}\right).$$

Note that obviously $0\leq\mathfrak{L}_{n,v}(\mathfrak{h}_m)\leq v\ln(n)$, $\forall\mathfrak{h}_m\in\mathfrak{H}_m$. Put for any $\mathfrak{h}\in\boldsymbol{\mathfrak{H}}$

$$\widetilde{P}(\mathfrak{h})=P\big(\mathfrak{h}^{(k)}\big)+\mathfrak{L}_{n,v}\big(\mathfrak{h}_m\big)+2\ln\{1+|\ln(F_{\mathbf{n_2},\mathbf{r}}(\mathfrak{h}))|\};$$

$$\widetilde{M}_q(\mathfrak{h})=M_q\big(\mathfrak{h}^{(k)}\big)+\mathfrak{L}_{n,v}\big(\mathfrak{h}_m\big)+2\ln\{1+|\ln(F_{\mathbf{n_2},\mathbf{r}}(\mathfrak{h}))|\}.$$

Define for any $\mathfrak{h}\in\boldsymbol{\mathfrak{H}}$, $\mathbf{r}\in\overline{\mathbb{N}}$, $z\geq 0$ and $q>0$

$$\widetilde{\mathcal{V}}_{\mathbf{r}}^{(v,z)}(n,\mathfrak{h})=\lambda_1\sqrt{G_\infty\big(\mathfrak{h}^{(k)}\big)\Big(F_{\mathbf{n_2},\mathbf{r}}(\mathfrak{h})n^{-1}\Big)\big(\widetilde{P}\big(\mathfrak{h}\big)+z\big)}+\lambda_2 G_\infty\big(\mathfrak{h}^{(k)}\big)\Big(n^{-1}\ln^{\beta}(n)\Big)\big(\widetilde{P}\big(\mathfrak{h}\big)+z\big);$$

$$\widetilde{\mathcal{U}}_{\mathbf{r}}^{(v,z,q)}(n,\mathfrak{h})=\lambda_1\sqrt{G_\infty\big(\mathfrak{h}^{(k)}\big)\Big(F_{\mathbf{n_2},\mathbf{r}}(\mathfrak{h})n^{-1}\Big)\big(\widetilde{M}_q(\mathfrak{h})+z\big)}+\lambda_2 G_\infty\big(\mathfrak{h}^{(k)}\big)\Big(n^{-1}\ln^{\beta}(n)\Big)\big(\widetilde{M}_q(\mathfrak{h})+z\big).$$

Theorem 3. *Let Assumptions 4, 7 and 8 hold. If* $\mathbf{n_1}\neq\mathbf{n_2}$ *suppose additionally that Assumption 5 holds. Then for any* $\mathbf{r}\in\mathbb{N}$, $v\geq 1$, $z\geq 1$ *and* $q\geq 1$

$$\mathbb{P}_{\mathrm{f}}\left\{\sup_{n\in\widetilde{\mathbf{N}}}\sup_{\mathfrak{h}\in\widetilde{\mathfrak{H}}(n)}\Big[\big|\xi_{\mathfrak{h}}(n)\big|-\widetilde{\mathcal{V}}_{\mathbf{r}}^{(v,z)}(n,\mathfrak{h})\Big]\geq 0\right\}\leq\mathfrak{n}^5\Big\{4838e^{-z}+4{\mathbf{n_1}}^{2-v}\Big\};$$

$$\mathbb{E}_{\mathrm{f}}\left\{\sup_{n\in\widetilde{\mathbf{N}}}\sup_{\mathfrak{h}\in\widetilde{\mathfrak{H}}(n)}\Big[\big|\xi_{\mathfrak{h}}(n)\big|-\widetilde{\mathcal{U}}_{\mathbf{r}}^{(v,z,q)}(n,\mathfrak{h})\Big]\right\}_{+}^{q}\leq 2\mathfrak{n}^5c_q\left[\sqrt{(\mathbf{n_1})^{-1}F_{\mathbf{n_2}}\underline{G}_{\mathbf{n}}}\vee\Big((\mathbf{n_1})^{-1}\ln^{\beta}(\mathbf{n_2})\underline{G}_{\mathbf{n}}\Big)\right]^q e^{-z}$$
$$+2^{q+2}\mathfrak{n}^5(\overline{G}_{\mathbf{n}})^q\,{\mathbf{n_1}}^{2-v}.$$

Although the assertions of the theorem are true whenever $v\geq 1$ the presented results are obviously reasonable only if $v>2$. For example (as we will see later) the typical choice of this parameter for the "moment bound" is $v=q+2$.

In spite of the fact that upper functions presented in Theorem 3 are found explicitly their expressions are quite cumbersome. In particular, it is unclear how to compute the function $\mathfrak{L}_{n,v}(\cdot)$. Of course, since $\mathfrak{L}_{n,v}(\mathfrak{h}_m)\leq v\ln(n)$, $\forall\mathfrak{h}_m\in\mathfrak{H}_m$, one can replace it by $v\ln(n)$ in the definition of $\widetilde{P}(\cdot)$ and $\widetilde{M}_q(\cdot)$, but the corresponding upper functions are not always sufficiently tight.

Our goal now is to simplify the expressions for upper functions given in Theorem 3. Surprisingly, that if n is fixed, i.e. $\mathbf{n_1}=\mathbf{n_2}$, it can be done without any additional assumption.

Set for any $v>0$ and $\mathfrak{h}\in\boldsymbol{\mathfrak{H}}$

$$\widehat{P}_v\big(\mathfrak{h}^{(k)}\big)=P\big(\mathfrak{h}^{(k)}\big)+2v\left|\ln\Big(2G_\infty\big(\mathfrak{h}^{(k)}\big)\Big)\right|,\quad\widehat{M}_{q,v}\big(\mathfrak{h}^{(k)}\big)=M_q\big(\mathfrak{h}^{(k)}\big)+2v\left|\ln\Big(2G_\infty\big(\mathfrak{h}^{(k)}\big)\Big)\right|,$$

and let $\widehat{F}_{\mathbf{n_2}}(\mathfrak{h})=\max[F_{\mathbf{n_2}}(\mathfrak{h}),{\mathbf{n_2}}^{-1}]$.

Corollary 2. *Let the assumptions of Theorem 3 hold. If* $\mathbf{n_1} \neq \mathbf{n_2}$ *suppose additionally that* $X_{i,1},\ i \geq 1,$ *are identically distributed.*

Then, the results of Theorem 3 remain valid if one replaces $\widetilde{\mathcal{V}}_{\mathbf{r}}^{(v,z)}(n,\mathfrak{h})$ *and* $\widetilde{\mathcal{U}}_{\mathbf{r}}^{(v,z,q)}(n,\mathfrak{h})$ *by*

$$\begin{aligned}\widehat{\mathcal{V}}^{(v,z)}(n,\mathfrak{h}) &= \lambda_1\sqrt{G_\infty\big(\mathfrak{h}^{(k)}\big)\Big(\widehat{F}_{\mathbf{n_2}}(\mathfrak{h})n^{-1}\Big)\Big(\widehat{P}_v\big(\mathfrak{h}^{(k)}\big)+2(v+1)\big|\ln\big\{\widehat{F}_{\mathbf{n_2}}(\mathfrak{h})\big\}\big|+z\Big)}\\ &\quad+\lambda_2 G_\infty\big(\mathfrak{h}^{(k)}\big)\Big(n^{-1}\ln^\beta(n)\Big)\Big(\widehat{P}_v\big(\mathfrak{h}^{(k)}\big)+2(v+1)\big|\ln\big\{\widehat{F}_{\mathbf{n_2}}(\mathfrak{h})\big\}\big|+z\Big);\\ \widehat{\mathcal{U}}^{(v,z,q)}(n,\mathfrak{h}) &= \lambda_1\sqrt{G_\infty\big(\mathfrak{h}^{(k)}\big)\Big(\widehat{F}_{\mathbf{n_2}}(\mathfrak{h})n^{-1}\Big)\Big(\widehat{M}_{q,v}\big(\mathfrak{h}^{(k)}\big)+2(v+1)\big|\ln\big\{\widehat{F}_{\mathbf{n_2}}(\mathfrak{h})\big\}\big|+z\Big)}\\ &\quad+\lambda_2 G_\infty\big(\mathfrak{h}^{(k)}\big)\Big(n^{-1}\ln^\beta(n)\Big)\Big(\widehat{M}_{q,v}\big(\mathfrak{h}^{(k)}\big)+2(v+1)\big|\ln\big\{\widehat{F}_{\mathbf{n_2}}(\mathfrak{h})\big\}\big|+z\Big).\end{aligned}$$

We would like to emphasize that we do not require that $X_i,\ i \geq 1,$ would be identically distributed. In particular, coming back to the generalized empirical process considered in Example 2, Section 3.1, where $X_i = (Y_i, \varepsilon_i)$, the design points $Y_i,\ i \geq 1,$ are often supposed to be uniformly distributed on some bounded domain of $\mathbb{R}^d$. As to the noise variables $\varepsilon_i,\ i \geq 1,$ the restriction that they are identically distributed cannot be justified in general.

3.3.2. Law of logarithm

Our goal here is to use the first assertion of Corollary 2 in order to establish the result referred later to *the law of logarithm*. Namely we show that for some $\boldsymbol{\Upsilon} > 0$

$$\limsup_{n\to\infty}\ \sup_{\mathfrak{h}^{(k)}\in\overline{\mathfrak{H}}_1^k(n,a)}\frac{\sqrt{n}\,\eta_{\mathfrak{h}^{(k)}}(n)}{\sqrt{G_\infty\big(\mathfrak{h}^{(k)}\big)\Big[\ln\big\{G_\infty\big(\mathfrak{h}^{(k)}\big)\big\}\vee\ln\ln(n)\Big]}}\leq\boldsymbol{\Upsilon}\quad\mathbb{P}_{\mathrm{f}}-\text{a.s.}\tag{3.16}$$

As previously we will first provide with the non-asymptotical version of (3.16).

We will suppose that (3.11) and (3.12) are fulfilled and replace (3.13) by the following assumption. For some $\mathfrak{a} > 0$

$$\mathcal{L}^{(k)}(z)\leq\mathfrak{a}\ln(z),\quad\forall z\geq 2.\tag{3.17}$$

Theorem 4. *Let Assumptions 4, 5, 7 and 8 be fulfilled. Suppose also that (3.11), (3.12) and (3.17) hold and assume that* $X_{i,1},\ i \geq 1,$ *are identically distributed.*

Then there exits $\boldsymbol{\Upsilon} > 0$ *such that for any* $\mathbf{j} \geq 3$ *and any* $a > 4$

$$\mathbb{P}_{\mathrm{f}}\left\{\sup_{n\geq\mathbf{j}}\ \sup_{\mathfrak{h}^{(k)}\in\overline{\mathfrak{H}}_1^k(n,a)}\frac{\sqrt{n}\,\eta_{\mathfrak{h}^{(k)}}(n)}{\sqrt{G_\infty\big(\mathfrak{h}^{(k)}\big)\Big[\ln\big\{G_\infty\big(\mathfrak{h}^{(k)}\big)\big\}\vee\ln\ln(n)\Big]}}\geq\boldsymbol{\Upsilon}\right\}\leq\frac{4840\mathfrak{n}^5}{\ln(\mathbf{j})}.$$

Some remarks are in order. The explicit expression of the constant $\boldsymbol{\Upsilon}$ is available and the generalization , similar to one announced in Remark 8, is possible. Also, (3.16) is an obvious consequence of Theorem 4. At last, we note that in view of (3.11) the factor $\Big[\ln\big\{G_\infty\big(\mathfrak{h}^{(k)}\big)\big\}\vee\ln\ln(n)\Big]$ can be replaced by $\ln(n)$ which is, up to a constant, its upper estimate. The corresponding result is, of course, rougher than one presented in the theorem, but its derivation does not require $X_{i,1},\ i \geq 1,$ to be identically distributed. This result is deduced directly from Theorem 3. Its proof is almost the same as the proof of Theorem 4 and based on the trivial bound $\mathfrak{L}_{n,v}(\mathfrak{h}_m) \leq v\ln(n),\ \forall\mathfrak{h}_m\in\mathfrak{H}_m$.

3.4. Application to localized processes

Let $\big(\mathbb{X}_l, \mu_l, \rho_l\big)$, $l = \overline{1, d+1}$, $d \in \mathbb{N}$, be the collection of measurable metric spaces. Throughout this section we will suppose that (3.8) holds with $p = 2$,

$$\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2, \quad \big(\mathcal{X}_1, \nu_1\big) = \big(\mathbb{X}_1 \times \cdots \times \mathbb{X}_d, \mu_1 \times \cdots \times \mu_d\big) =: \big(\mathbb{X}_1^d, \mu^{(d)}\big), \quad \big(\mathcal{X}_2, \nu_2\big) = \big(\mathbb{X}_{d+1}, \mu_{d+1}\big),$$

x_j denotes the element of $\mathbb{X}_j$, $j = \overline{1, d+1}$, and $x^{(d)}$ will denotes the element of $\mathbb{X}_1^d$. We equip the space $\mathbb{X}_1^d$ with the semi-metric $\rho^{(d)} = \max_{l=\overline{1,d}} \rho_l$.

Problem formulation This section is devoted to the application of Theorems 1 and 3 in the following case:

- $\mathfrak{H}_1^d := \mathfrak{H}_1 \times \cdots \times \mathfrak{H}_d = (0,1] \times \cdots \times (0,1] = (0,1]^d$, (i.e. $k = d$);
- $\mathfrak{H}_{d+1}^{d+2} = \mathfrak{H}_{d+1} \times \mathfrak{H}_{d+2} := \mathcal{Z} \times \bar{\mathbb{X}}_1^d$, i.e. $m = d+2$, where $\bar{\mathbb{X}}_1^d := \bar{\mathbb{X}}_1 \times \cdots \times \bar{\mathbb{X}}_d$ be a given subset of $\mathbb{X}_1^d$ and $(\mathcal{Z}, \mathfrak{d})$ is a given metric space.
- The function $G(\cdot, \cdot)$ obeys some structural assumption described below and for any $\mathfrak{h} := \big(r, \mathfrak{z}, \bar{x}^{(d)}\big) \in (0,1]^d \times \mathcal{Z} \times \bar{\mathbb{X}}_1^d$ the function $G(\mathfrak{h}, \cdot)$ "decrease rapidly " outside of the set $\Big\{x_1 \in \mathbb{X}_1 : \ \rho_1\big(x_1, \bar{x}_1\big) \le r_1\Big\} \times \cdots \times \Big\{x_d \in \mathbb{X}_d : \ \rho_d\big(x_d, \bar{x}_d\big) \le r_d\Big\} \times \mathbb{X}_{d+1}$.

Let $K : \mathbb{R}^d \to \mathbb{R}$ be a given function, $\big(\gamma_1, \ldots, \gamma_d\big) \in \mathbb{R}_+^d$ be given vector and set for any $r \in (0,1]^d$

$$K_r(\cdot) = V_r^{-1} K\left(\cdot / r_1, \ldots, \cdot / r_d\right), \quad V_r = \prod_{l=1}^{d} r_l^{\gamma_l}.$$

where, as previously, for $u, v \in \mathbb{R}^d$ the notation u/v denotes the coordinate-wise division. Let

$$G(\mathfrak{h}, x) = g\big(\mathfrak{z}, x\big) K_r\left(\vec{\rho}\big(x^{(d)}, \bar{x}^{(d)}\big)\right), \quad \mathfrak{h} = \left(r, \mathfrak{z}, \bar{x}^{(d)}\right) \in (0,1]^d \times \mathcal{Z} \times \bar{\mathbb{X}}_1^d =: \mathfrak{H}, \tag{3.18}$$

where $g : \mathcal{Z} \times \mathcal{X} \to \mathbb{R}$ is a given function those properties will be described later and

$$\vec{\rho}\big(x^{(d)}, \bar{x}^{(d)}\big) = \big(\rho_1\big(x_1, \bar{x}_1\big), \ldots, \rho_d\big(x_d, \bar{x}_d\big)\big)\,.$$

The corresponding generalized empirical process is given by

$$\xi_{\mathfrak{h}}(n) = n^{-1} \sum_{i=1}^{n} \left[g\big(\mathfrak{z}, X_i\big) K_r\left(\vec{\rho}\Big(\left[X_i\right]^{(d)}, \bar{x}^{(d)}\Big)\right) - \mathbb{E}_{\mathrm{f}}\left\{ g\big(\mathfrak{z}, X_i\big) K_r\left(\vec{\rho}\Big(\left[X_i\right]^{(d)}, \bar{x}^{(d)}\Big)\right)\right\} \right].$$

We will seek upper functions for the random field $\zeta_r\big(n, \bar{x}^{(d)}\big) := \sup\limits_{\mathfrak{z} \in \mathcal{Z}} \left|\xi_{r,\mathfrak{z},\bar{x}^{(d)}}(n)\right|$ in two cases: $\bar{\mathbb{X}}_1^d = \mathbb{X}_1^d$ and $\bar{\mathbb{X}}_1^d = \left\{\bar{x}^{(d)}\right\}$ for a fixed $\bar{x}^{(d)} \in \mathbb{X}_1^d$.

To realize this program we will apply Theorems 1 and 3 to $\xi_{\mathfrak{h}}(n)$, $\mathfrak{h} = \left(r, \mathfrak{z}, \bar{x}^{(d)}\right)$. It is worth mentioning that corresponding upper functions can be used for constructing of estimation procedures in different areas of mathematical statistics: M-estimation with locally polynomial fitting (non-parametric regression), kernel density estimation and many others.

Moreover, we apply Theorem 2 for establishing a non-asymptotical version of the law of iterated logarithm for $\zeta_r\big(\bar{x}^{(d)}, n\big)$ in the case where $\bar{\mathbb{X}}_1^d = \left\{\bar{x}^{(d)}\right\}$. We also apply Theorem 4 for deriving a non-asymptotical version of the law of logarithm for $\left\|\zeta_r(n)\right\|_\infty := \sup_{\bar{x}^{(d)} \in \mathbb{X}_1^d} \left|\zeta_r\big(\bar{x}^{(d)}, n\big)\right|$. Our study here generalizes in several directions the existing results Einmahl and Mason (2000), Giné and Guillou (2002), Einmahl and Mason (2005), Dony et al. (2006), Dony et Einmahl (2009).

Assumptions and notations

Assumption 9. (**i**) $\|K\|_\infty < \infty$ *and for some* $L_1 > 0$

$$|K(t) - K(s)| \le \frac{L_1|t-s|}{1+|t| \wedge |s|}, \quad \forall t, s \in \mathbb{R}^d,$$

where $|\cdot|$ *denotes supremum norm on* $\mathbb{R}^d$.

(**ii**) $\|g\|_\infty := \sup\limits_{\mathfrak{z}\in\mathcal{Z},\, x\in\mathcal{X}} \big|g\big(\mathfrak{z}, x\big)\big| < \infty$, *and for some* $\alpha \in (0,1]$, $L_\alpha > 0$,

$$\sup_{x\in\mathcal{X}} \big|g\big(\mathfrak{z}, x\big) - g\big(\mathfrak{z}', x\big)\big| \le L_\alpha \left[\mathfrak{d}(\mathfrak{z}, \mathfrak{z}')\right]^\alpha, \ \forall \mathfrak{z}, \mathfrak{z}' \in \mathcal{Z};$$

The conditions (**i**) and (**ii**) are quite standards. In particular (**i**) holds if K is compactly supported and lipschitz continuous. If $g\big(\mathfrak{z}, \cdot\big) = \bar{g}(\cdot)$, for any $\mathfrak{z} \in \mathcal{Z}$, then (**ii**) is verified for any bounded $\bar{g}$.

Let $0 < r_l^{(\min)}(n) \le r_l^{(\max)}(n) \le 1$, $l = \overline{1,d}$, $n \ge 1$, be given decreasing sequences and let

$$\mathfrak{H}(n) = \mathcal{R}(n) \times \mathcal{Z} \times \bar{X}_1^d, \qquad \mathcal{R}(n) = \prod_{l=1}^d \big[r_l^{(\min)}(2n), r_l^{(\max)}(n)\big];$$

$$\widetilde{\mathfrak{H}}(n) = \widetilde{\mathcal{R}}(n) \times \mathcal{Z} \times \bar{X}_1^d, \qquad \widetilde{\mathcal{R}}(n) = \prod_{l=1}^d \big[r_l^{(\min)}(n), r_l^{(\max)}(n)\big].$$

We note that $\widetilde{\mathfrak{H}}(n) \subseteq \mathfrak{H}(n)$ for any $n \ge 1$ since $r_l^{(\min)}(\cdot), r_l^{(\max)}(\cdot)$, $l = \overline{1,d}$, are decreasing, and obviously $\widetilde{\mathfrak{H}}(n) \subseteq \mathfrak{H}(\mathbf{m})$ for any $n \in \{\mathbf{m}, \ldots, \mathbf{2m}\}$ and any $\mathbf{m} \ge 1$.

Remark 9. *Assumption 5 is fulfilled with* $n[\mathbf{m}] = \mathbf{m}$.

Lemma 2. *Suppose that Assumption 9 is fulfilled and let* $\bar{\mathbb{X}}_1^d \subseteq \mathbb{X}_1^d$ *be an arbitrary subset. Then, for arbitrary sequences* $0 < r_l^{(\min)}(n) \le r_l^{(\max)}(n) \le 1$, $l = \overline{1,d}$, $n \ge 1$, *Assumption 4 holds with*

$$\varrho_n^{(d)}\left(r, r'\right) = \max_{l=\overline{1,d}} \left|\gamma_l \ln\left(r_l / r_l'\right)\right|, \quad \varrho_{d+1} = \left[\mathfrak{d}\right]^\alpha, \quad \varrho_{d+2} = \max_{l=\overline{1,d}} \rho_l;$$

$$D_0(z) = \exp\{dz\} - 1 + (L_1/\|K\|_\infty)\Big(\exp\left\{\gamma^{-1}z\right\} - 1\Big), \quad \gamma = \min_{l=\overline{1,l}} \gamma_l;$$

$$D_{d+1}(z) = \big(L_\alpha/\|g\|_\infty\big)z, \quad D_{d+2}(z) = L_1\big(\|g\|_\infty \|K\|_\infty^2\big)^{-1} z, \quad L_{d+1}(z) = z, \quad L_{d+2}(z) = z^2.$$

Additionally, if $\bar{\mathbb{X}}_1^d$ *consists of a single point* $\bar{x}^{(d)} \in \mathbb{X}_1^d$ *then* $L_{d+2} \equiv 0$.

The proof of the lemma is postponed to Appendix. We remark that ϱ_{d+1} is a semi-metric, since $\alpha \in (0,1]$, and the semi-metric $\varrho_n^{(d)}$ is independent on n. In view of latter remark all quantities involved in Assumption 4 are independent on the choice of $r_l^{(\min)}(\cdot), r_l^{(\max)}(\cdot)$, $l = \overline{1,d}$,. We want to emphasize nevertheless that the assertion of the lemma is true for an arbitrary but *a priory* chosen $r_j^{(\min)}(\cdot), r_l^{(\max)}(\cdot)$, $l = \overline{1,d}$.

Thus, Lemma 2 guarantees the verification of the main assumption of Section 3.1, that makes possible the application of Theorems 1–4. Hence, we have to match the notations of these theorems to the notations used in the present section.

Since $k=d$ and $\mathfrak{H}_1^k=(0,1]^d$ we have $\mathfrak{h}^{(k)}=r$ and, therefore, in view of Assumption 9

$$\mathbf{G}_\infty(r):=\sup_{(\mathfrak{z},\bar{x}^{(d)})\in\mathcal{Z}\times\bar{\mathbb{X}}_1^d}\sup_{x\in\mathbb{X}_1^{d+1}}\left|G\left(\left\{r,\mathfrak{z},\bar{x}^{(d)}\right\},x\right)\right|\le V_r^{-1}\|g\|_\infty\|K\|_\infty=:G_\infty(r).$$

$$\underline{G}_n:=\inf_{r\in\mathcal{R}(n)}G_\infty(r)=V_{r^{(\max)}(n)}^{-1}\|g\|_\infty\|K\|_\infty,\quad\forall n\ge 1.$$

We remark that the function $G_\infty(\cdot)$ is independent of the choice of $\bar{\mathbb{X}}_1^d$. Define

$$f_i^{(d)}\big(x^{(d)}\big)=\int_{\mathbb{X}_{d+1}}f_i(x)\mu_{d+1}\big(\mathrm{d}x_{d+1}\big),\ \ i\ge 1,$$

and let $3\le\mathbf{n_1}\le\mathbf{n_2}\le\mathbf{2n_1}$ be fixed. Set for any $\big(r,\bar{x}^{(d)}\big)\in(0,1]^d\times\bar{\mathbb{X}}_1^d$

$$F_{\mathbf{n_2}}\left(r,\bar{x}^{(d)}\right)=\begin{cases}\|g\|_\infty(\mathbf{n_2})^{-1}\sum_{i=1}^{\mathbf{n_2}}\int_{\mathbb{X}_1^d}\left|K_r\left(\vec{\rho}\big(x^{(d)},\bar{x}^{(d)}\big)\right)\right|f_i^{(d)}\big(x^{(d)}\big)\mu^{(d)}(\mathrm{d}x^{(d)}), & \mathbf{n_1}=\mathbf{n_2};\\ \|g\|_\infty\sup\limits_{i=\overline{1,\mathbf{n_2}}}\displaystyle\int_{\mathbb{X}_1^d}\left|K_r\left(\vec{\rho}\big(x^{(d)},\bar{x}^{(d)}\big)\right)\right|f_i^{(d)}\big(x^{(d)}\big)\mu^{(d)}(\mathrm{d}x^{(d)}), & \mathbf{n_1}\ne\mathbf{n_2},\end{cases}$$

and note that in view of Assumption 9 (**ii**)

$$F_{\mathbf{n_2}}(\mathfrak{h})\le F_{\mathbf{n_2}}\left(r,\bar{x}^{(d)}\right),\quad\forall\mathfrak{h}\in(0,1]^d\times\mathcal{Z}\times\bar{\mathbb{X}}_1^d.$$

We remark that the function $F_{\mathbf{n_2}}(\cdot,\cdot)$ is independent of the choice of $\mathcal{Z}$. Put also

$$F_{\mathbf{n_2}}:=\sup_{n\in\widetilde{\mathbf{N}}}\sup_{\left(r,\bar{x}^d\right)\in\mathcal{R}(n)\times\bar{\mathbb{X}}_1^d}F_{\mathbf{n_2}}\big(r,\bar{x}^d\big)<\infty,$$

where, remind, $\widetilde{\mathbf{N}}=\{\mathbf{n_1},\dots,\mathbf{n_2}\}$. Finally for any $\mathbf{r}\in\overline{\mathbb{N}}$ set $F_{\mathbf{n_2},\mathbf{r}}(\cdot,\cdot)=\max\left[F_{\mathbf{n_2}}(\cdot,\cdot),e^{-\mathbf{r}}\right]$.

3.4.1. Pointwise results

Here we will consider the case, where $\bar{\mathbb{X}}_1^d=\left\{\bar{x}^{(d)}\right\}$ and $\bar{x}^{(d)}$ is a fixed element of $\mathbb{X}_1^d$. Note that in view of Lemma 2 $L_{d+1}(z)=z$ and $L_{d+2}\equiv 0$ that implies $\mathcal{L}^{(k)}\equiv 0$.

We will suppose that Assumption 6 holds with $k=d,m=d+1$ and $(\mathfrak{H}_{d+1},\varrho_{d+1})=(\mathcal{Z},[\mathfrak{d}]^\alpha)$. It is equivalent obviously to assume that Assumption 6 holds with $(\mathfrak{H}_{d+1},\varrho_{d+1})=(\mathcal{Z},\mathfrak{d})$ and with the constants $\tilde{N}=\alpha N$ and $\tilde{R}=R^{1/\alpha}$.

Let β and $C_{N,R,m,k}$ be the constants defined in Theorem 1. Set for any $r\in(0,1]^d$ and $q>0$

$$P(r)=(36d\delta_*^{-2}+6)\ln\left(1+\sum_{l=1}^d\gamma_l\ln\left\{\frac{2r_l^{(\max)}(n)}{r_l}\right\}\right)+18C_{N,R,d+1,d};$$

$$M_q(r)=\big(72d\delta_*^{-2}+2.5q+1.5\big)\sum_{l=1}^d\gamma_l\ln\left(\frac{2r_l^{(\max)}(n)}{r_l}\right)+36C_{N,R,d+1,d}.$$

and define for $\mathbf{r} \in \overline{\mathbb{N}}$ and $u > 0$

$$\begin{aligned}
\mathcal{V}_{\mathbf{r}}^{(u)}\big(n, r, \bar{x}^d\big) &= \boldsymbol{\lambda_1}\sqrt{\Big[F_{\mathbf{n_2},\mathbf{r}}\big(r, \bar{x}^d\big)(nV_r)^{-1}\Big]\Big[P(r) + 2\ln\big\{1 + \big|\ln\big\{F_{\mathbf{n_2},\mathbf{r}}\big(r, \bar{x}^d\big)\big\}\big|\big\} + u\Big]} \\
&\quad + \boldsymbol{\lambda_2}\Big[(nV_r)^{-1}\ln^{\beta}(n)\Big]\Big[P(r) + 2\ln\Big\{1 + \Big|\ln\Big\{F_{\mathbf{n_2},\mathbf{r}}\big(r, \bar{x}^d\big)\Big\}\Big|\Big\} + u\Big]; \\
\mathcal{U}_{\mathbf{r}}^{(u,q)}\big(n, r, \bar{x}^d\big) &= \boldsymbol{\lambda_1}\sqrt{\Big[F_{\mathbf{n_2},\mathbf{r}}\big(r, \bar{x}^d\big)(nV_r)^{-1}\Big]\Big[M_q(r) + 2\ln\big\{1 + \big|\ln\big\{F_{\mathbf{n_2},\mathbf{r}}\big(r, \bar{x}^d\big)\big\}\big|\big\} + u\Big]} \\
&\quad + \boldsymbol{\lambda_2}\Big[(nV_r)^{-1}\ln^{\beta}(n)\Big]\Big[M_q(r) + 2\ln\Big\{1 + \Big|\ln\Big\{F_{\mathbf{n_2},\mathbf{r}}\big(r, \bar{x}^d\big)\Big\}\Big|\Big\} + u\Big],
\end{aligned}$$

where $\boldsymbol{\lambda_1} = \sqrt{\|g\|_\infty \|K\|_\infty}\lambda_1$, $\boldsymbol{\lambda_2} = \|g\|_\infty\|K\|_\infty\lambda_2$ and $\lambda_1,\ \lambda_2$ are defined in Theorem 1.

The result below is the direct consequence of Theorem 1 and Lemma 2. We remark that defined above quantities are functions of r and n since $\bar{x}^d$ is fixed. Since they do not depend on the variable $\mathfrak{z}$, these quantities will be automatically upper functions for

$$\zeta_r\big(n, \bar{x}^{(d)}\big) := \sup_{\mathfrak{z}\in\mathcal{Z}}\Big|\xi_{r,\mathfrak{z},\bar{x}^{(d)}}\big(\bar{x}^{(d)}\big)\Big|.$$

Theorem 5. *Let Assumption 9 be fulfilled and suppose that Assumption 6 holds with $k = d, m = d+1$ and $(\mathfrak{H}_{d+1}, \varrho_{d+1}) = (\mathcal{Z}, [\mathfrak{d}]^\alpha)$.*

Then for any given decreasing sequences $0 < r_l^{(\min)}(n) \le r_l^{(\max)}(n) \le 1,\ l = \overline{1,d},\ n \ge 1$, any $\bar{x}^d \in \bar{\mathbb{X}}_1^d$ any $\mathbf{r} \in \mathbb{N}$, $\boldsymbol{b} > 1$ $u \ge 1$ and $q \ge 1$

$$\mathbb{P}_{\mathrm{f}}\left\{\sup_{n\in\widetilde{\mathbf{N}}}\sup_{r\in\widetilde{\mathcal{R}}(n)}\Big[\zeta_r\big(n, \bar{x}^{(d)}\big) - \mathcal{V}_{\mathbf{r}}^{(u)}\big(n, r, \bar{x}^d\big)\Big] \ge 0\right\} \le 2419\, e^{-u};$$

$$\mathbb{E}_{\mathrm{f}}\left\{\sup_{n\in\widetilde{\mathbf{N}}}\sup_{r\in\widetilde{\mathcal{R}}(n)}\Big[\zeta_r\big(n, \bar{x}^{(d)}\big) - \mathcal{U}_{\mathbf{r}}^{(u,q)}\big(n, r, \bar{x}^d\big)\Big]\right\}_+^q \le c_q'\left[\sqrt{\frac{F_{\mathbf{n_2}}}{\mathbf{n_1}V_{r^{(\max)}(\mathbf{n_1})}}} \vee \left(\frac{\ln^{\beta}(\mathbf{n_2})}{V_{r^{(\max)}(\mathbf{n_1})}\mathbf{n_1}}\right)\right]^q e^{-u},$$

where $c_q' = 2^{(7q/2)+5}3^{q+4}\Gamma(q+1)\left(C_{D,\boldsymbol{b}}\max\left[\sqrt{\|g\|_\infty\|K\|_\infty}, \|g\|_\infty\|K\|_\infty\right]\right)^q$.

The explicit expression for $C_{D,\boldsymbol{b}}$ can be also found in Theorem 1. In the case considered here it is completely determined by $(\gamma_1, \dots \gamma_d)$, L_1, L_α and $\boldsymbol{b}$.

As well as the assertions of Theorem 1 the latter theorem is proved without any assumption imposed on the densities $f_i,\ i = \overline{1,n}$. The choice of $r_l^{(\min)}(n), r_l^{(\max)}(n),\ l = \overline{1,d},\ n \ge 1$, is also assumption free. Additionally, Assumption 6 can be replaced by (3.10), see Remark 7.

Note also that if $g\big(\mathfrak{z}, \cdot\big) = \bar{g}(\cdot)$, for any $\mathfrak{z} \in \mathcal{Z}$, then Assumption 6 is not needed anymore and, moreover, Assumption 9 (**ii**) is verified for an arbitrary bounded $\bar{g}$. Hence, in this case the assertions of Theorem 5 are established under very mild Assumption 9 (**i**) imposed on the function K.

Remark 10. *We note that the discussed in Introduction so-called price to pay for uniformity disappears if $r = r^{(\max)}$. Indeed, $P\left(r^{(\max)}\right)$ and $M_q\left(r^{(\max)}\right)$ are absolute constants. This property is crucial, in particular, for constructing statistical procedures used in the estimation of functions possessing inhomogeneous smoothness, see Lepski et al. (1997), Kerkyacharian et al. (2001).*

Some additional assumptions and their consequences To apply Theorem 5 to specific problems one needs to find an efficient upper bound for the quantity $F_{\mathbf{n_2}}(\cdot,\cdot)$. Below we provide with sufficient condition allowing to solve this problem under general consideration and we will not

be tending here to the maximal generality. We impose some additional restrictions on the densities $f_i,\ i=\overline{1,n}$, and on the measures μ_l of ρ_l-balls in the spaces $\mathbb{X}_l,\ l=\overline{1,d}$. Moreover, we should precise the behavior of the function K at infinity. Then, we will use these assumptions for establishing of the law of iterated logarithm.

Introduce the following notations. For any $t\in\mathbb{R}^d_+$ define

$$\check{K}(t)=\sup_{|u|\notin\Pi_t}|K(u)|,\quad \Pi_t=[0,t_1]\times\cdots\times[0,t_d].$$

For any $l=\overline{1,d},\ x_l\in\mathbb{X}_l$ and $\mathrm{r}>0$ set $\mathbb{B}_l\big(\mathrm{r},x_l\big)=\big\{y\in\mathbb{X}_l:\ \ \rho_l\big(y,x_l\big)\le\mathrm{r}\big\}$.

Assumption 10. *There exists $L_2>0$ such that*

$$\sup_{t\in\mathbb{R}^d_+}\left[\Big(\prod_{l=1}^d t_l^{1+\gamma_l}\Big)\check{K}(t)\right]\le L_2; \tag{3.19}$$

For any $l=\overline{1,d}$ and any $x_l\in\mathbb{X}_l$ one has $\mathbb{X}_l=\cup_{\mathrm{r}>0}\big(\mathbb{B}_l\big(\mathrm{r},x_l\big)\big)$ and there exist $L^{(l)}>0$

$$\mu_l\Big(\mathbb{B}_l\big(\mathrm{r},x_l\big)\Big)\le L^{(l)}\mathrm{r}^{\gamma_l},\quad\forall\mathrm{r}>0; \tag{3.20}$$

Moreover,

$$\sup_{i\ge1}\sup_{x^{(d)}\in\mathbb{X}_1^d}f_i^{(d)}\big(x^{(d)}\big)=:\mathrm{f}_\infty<\infty. \tag{3.21}$$

The condition (3.19) is obviously fulfilled if K is compactly supported on $[0,1]^d$. It is also satisfied in the case of Gaussian or Laplace kernel.

The condition (3.20) can be easily checked if $\mathbb{X}_l,\ l=\overline{1,d}$ are doubling metric spaces. In particular, if $\mathbb{X}_l=\mathbb{R}$ and $\mu_l,\ l=\overline{1,d}$, are the Lebesgue measures than (3.20) holds with $L^{(l)}=1,\ \gamma_l=1,\ l=\overline{1,d}$. If $\mathbb{X}_l=\mathbb{R}^{d_l},\ \ l=\overline{1,d}$, then (3.20) holds with $\gamma_l=d_l$ and the constants $L^{(l)}$ depending on the choice of the distances ρ_l.

As to condition (3.21) we remark that the boundedness of the entire density f_i is not required. For example, under independence structure, i.e. $f_i(x)=f_i^{(d)}\big(x^{(d)}\big)p_i\big(x_{d+1}\big)$, the densities p_i may be unbounded.

Lemma 3. *The following bound holds under Assumption 10:*

$$\sup_{\mathbf{n_2}\ge1}\sup_{r\in(0,1]^d}\sup_{\bar{x}^{(d)}\in\mathbb{X}_1^d}F_{\mathbf{n_2}}\left(r,\bar{x}^{(d)}\right)\le 2^d\mathrm{f}_\infty\|g\|_\infty L_2\prod_{l=1}^d 2^{\gamma_l}L^{(l)}.$$

The proof of lemma is postponed to Appendix. Our goal now is to deduce the law of iterated logarithm for $\zeta_r\big(n,\bar{x}^{(d)}\big)$ from Theorem 2. Set for any $n\in\mathbb{N}^*$ and $a>0$

$$\overline{\mathcal{R}}_a(n)=\Big\{r\in(0,1]^d:\ \ V_r\ge\ n^{-1}(\ln n)^a\Big\}.$$

and choose $h^{(\max)}=(1,\dots,1)$ and $h^{(\min)}=(1/n,\dots,1/n)$.

Remark 11. 1^0. *Note that $\overline{\mathcal{R}}_a(n)\subset[n^{-1},1]^d=:\widetilde{\mathcal{R}}(n)$ for any $n\ge3$ and any $a>0$ and, therefore, the assertion of Lemma 2 holds.*

2^0. *We have $\underline{G}_n=\|K\|_\infty\|g\|_\infty,\quad\overline{G}_n=\|K\|_\infty\|g\|_\infty n^{-d}$ for any $n\ge1$ and, therefore, (3.11) is verified with $\mathfrak{c}=\|K\|_\infty\|g\|_\infty$ and $\mathfrak{b}=d$.*

3^0. *Lemma 3 implies that the condition (3.12) holds with* $\mathbf{F} \le 2^d \mathrm{f}_\infty \|g\|_\infty L_2 \prod_{l=1}^d 2^{\gamma_l} L^{(l)}$.

4^0. *In view of Lemma 2* $L_{d+1}(z) = z$ *and* $L_{d+2} \equiv 0$, *that implies* $\mathcal{L}^{(k)} \equiv 0$. *Hence, the condition (3.13) is fulfilled for any* $\mathfrak{a} > 0$.

Thus, all assumptions of Theorem 2 are checked and we come to the following statement.

Theorem 6. *Let Assumptions 9 and 10 be fulfilled and suppose that Assumption 6 holds with* $k = d, m = d+1$ *and* $(\mathfrak{H}_{d+1}, \varrho_{d+1}) = (\mathcal{Z}, [\mathfrak{d}]^\alpha)$. *Then there exists* $\Upsilon > 0$ *such that for any* $\bar{x}^d \in \bar{\mathbb{X}}_1^d$ *and any* $a > 2$

$$\mathbb{P}_{\mathrm{f}} \left\{ \sup_{n \ge \mathbf{j}} \sup_{r:\, n^{-1}(\ln n)^a \le V_r \le 1} \left[\frac{\sqrt{nV_r}\, \zeta_r\big(n, \bar{x}^{(d)}\big)}{\sqrt{\ln\big(1 + \ln(n)\big)}} \right] \ge \Upsilon \right\} \le \frac{2419}{\ln(\mathbf{j})}.$$

Remark 12. *The inspection of the proof of Theorem 2 together with Lemma 3 allows us to assert that the statement of Theorem 6 is* uniform *over the set of bounded densities.*

More precisely, for any $\mathfrak{f} > 0$ *there exists* $\Upsilon(\mathfrak{f})$ *such that*

$$\sup_{\mathrm{f} \in \mathcal{F}_\mathfrak{f}} \mathbb{P}_{\mathrm{f}} \left\{ \sup_{n \ge \mathbf{j}} \sup_{r:\, n^{-1}(\ln n)^a \le V_r \le 1} \left[\frac{\sqrt{nV_r}\, \zeta_r\big(n, \bar{x}^{(d)}\big)}{\sqrt{\ln\big(1 + \ln(n)\big)}} \right] \ge \Upsilon(\mathfrak{f}) \right\} \le \frac{2419}{\ln(\mathbf{j})}, \tag{3.22}$$

where $\mathcal{F}_\mathfrak{f} = \big\{(f_i, i \ge 1) :\ \mathrm{f}_\infty \le \mathfrak{f}\big\}$. *As before the explicit expression of* $\Upsilon(\cdot)$ *is available.*

The following consequence of Theorem 6 is straightforward.

$$\limsup_{n \to \infty} \sup_{r:\, n^{-1}(\ln n)^a \le V_r \le 1} \left[\frac{\sqrt{nV_r}\, \zeta_r\big(n, \bar{x}^{(d)}\big)}{\sqrt{\ln\big(1 + \ln(n)\big)}} \right] \le \Upsilon \quad \mathbb{P}_f - \mathbf{a.s.} \tag{3.23}$$

Theorem 6 generalizes the existing results, see for example Dony et Einmahl (2009), in the following directions.

1. *Structural assumption.* The structural condition (3.18) is imposed in cited papers but with additional restriction: either $g(\mathfrak{z}, x) \equiv$ const ("density case") or $g(\mathfrak{z}, x) = \bar{g}(x)$ ("regression case"). It excludes, for instance, the problems appearing in robust estimation. We note that Assumption 9 (**ii**) is fulfilled here if $\bar{g}$ is bounded function and Assumption 6 is not needed anymore, since $\bar{g}$ is independent of $\mathfrak{z}$.
2. *Anisotropy.* All known to the author results treat the case where $\mathbb{X}_l = \mathbb{R},\ l = \overline{1,d}$, and $\mathcal{R}(n) = \big\{(r_1, \ldots, r_d) \in (0,1]^d :\ \ r_l = \mathrm{r},\ \forall l = \overline{1,d},\ \ \mathrm{r} \in \big[\mathrm{r}^{(min)}(n), \mathrm{r}^{(max)}(n)\big]\big\}$ (isotropic case). We remark that (3.20) is automatically fulfilled with $\gamma_l = 1, L^{(l)} = 1,\ l = \overline{1,d}$, and $V_r = \mathrm{r}^d$. Note also that we consider independent but not necessarily identically distributed random variables. This is important, in particular, for various estimation problems arising in non-parametric regression model.
3. *Kernel.* We do not suppose that the function K is compactly supported. For instance, one can use the gaussian or laplace kernel. It allows, for instaince, to consider the problems where X_1^d is not linear space. In particular, it can be some manifold satisfying doubling condition.
4. *Non-asymptotic nature.* The existing results are presented as in (3.23). Note, however, that the random field $\zeta_r\big(n, \bar{x}^d\big)$ appears in various areas of nonparametric estimation (density estimation, regression). As the consequence a.s. convergence has no much sense since there is no a unique probability measure (see, also Remark 12).

3.4.2. Sup-norm results

Here we consider $\bar{\mathbb{X}}_1^d = \mathbb{X}_1^d$. We assume that there exists $\{\mathrm{X}_{\mathbf{i}},\ \mathbf{i} \in \mathbf{I}\}$ which is $\mathfrak{n}$-totally bounded cover of $\left(\mathbb{X}_1^d, \rho^{(d)}\right)$ satisfying Assumption 7 (**i**) and possessing the separation property.

Assumption 11. *There exists* $\mathfrak{t} > 0$ *such that for any* $\mathbf{i}, \mathbf{k} \in \mathbf{I}$ *satisfying* $\mathrm{X}_{\mathbf{i}} \cap \mathrm{X}_{\mathbf{k}} = \emptyset$

$$\inf_{x^{(d)} \in \mathrm{X}_{\mathbf{i}}} \inf_{y^{(d)} \in \mathrm{X}_{\mathbf{k}}} \rho^{(d)}\left(x^{(d)}, y^{(d)}\right) > \mathfrak{t}.$$

Also we suppose that Assumption 7 (**ii**) holds with $k = d, m = d+1$ and $(\mathfrak{H}_{d+1}, \varrho_{d+1}) = (\mathcal{Z}, [\mathfrak{d}]^{\alpha})$. We remark that in the considered case this assumption coincides with Assumption 6.

Let, as previously, $0 < r_l^{(\min)}(n) \le r_l^{(\max)}(n) \le 1,\ l = \overline{1,d},\ n \ge 1$, be given decreasing sequences,

$$\mathfrak{H}(n) = \mathcal{R}(n) \times \mathcal{Z} \times X_1^d, \qquad \mathcal{R}(n) = \prod_{l=1}^{d} \left[r_l^{(\min)}(2n), r_l^{(\max)}(n)\right];$$

$$\widetilde{\mathfrak{H}}(n) = \widetilde{\mathcal{R}}(n) \times \mathcal{Z} \times X_1^d, \qquad \widetilde{\mathcal{R}}(n) = \prod_{l=1}^{d} \left[r_l^{(\min)}(n), r_l^{(\max)}(n)\right].$$

Our last condition relates the choice of the vector $r^{(\max)}(n),\ n \ge 1$ and the kernel K with the parameter $\mathfrak{t}$ appearing in Assumption 11. Let us assume that for any $n \ge 1$

$$\sup_{r \in \mathcal{R}(n)} \sup_{|u| \notin (0,\mathfrak{t}]^d} |K(u/r)| \le \|K\|_{\infty} n^{-1}. \tag{3.24}$$

Note that (3.24) holds if K is compactly supported on $[-\mathfrak{t}, \mathfrak{t}]^d$ and $r^{(\max)}(n) \in (0, \mathfrak{t})^d$ for any $n \ge 1$.

Lemma 4. *Assumption 11 and (3.24) imply Assumption 8.*

The proof of lemma is given in Appendix. Set for any $r \in (0,1]^d$ and $v > 0$

$$\widehat{M}_{q,v}(r) = \left([72d + 108N]\delta_*^{-2} + 2.5q + 2v + 1.5\right) \ln\left(2V_r^{-1}\right) + \mathbf{C},$$

where we have put $\mathbf{C} = 72N\delta_*^{-2} \left|\log_2\left(\|g\|_{\infty}\|K\|_{\infty}\right)\right| + 36C_{N,R,d+1,d}$.

Let $3 \le \mathbf{n_1} \le \mathbf{n_2} \le 2\mathbf{n_1}$ be fixed. Set $\widehat{F}_{\mathbf{n_2}}\left(r, \bar{x}^{(d)}\right) = \max\left[F_{\mathbf{n_2}}\left(r, \bar{x}^{(d)}\right), \mathbf{n_2}^{-1}\right]$ and define

$$\begin{aligned}\widehat{\mathcal{U}}^{(v,z,q)}\left(n, r, \bar{x}^{(d)}\right) &= \boldsymbol{\lambda_1}\sqrt{\left[\widehat{F}_{\mathbf{n_2}}\left(r, \bar{x}^{(d)}\right)(nV_r)^{-1}\right]\left[\widehat{M}_{q,v}(r) + 2(v+1)\left|\ln\left\{\widehat{F}_{\mathbf{n_2}}\left(r, \bar{x}^{(d)}\right)\right\}\right| + z\right]} \\ &\quad + \boldsymbol{\lambda_2}\left[(nV_r)^{-1}\ln^{\beta}(n)\right]\left[\widehat{M}_{q,v}(r) + 2(v+1)\left|\ln\left\{\widehat{F}_{\mathbf{n_2}}\left(r, \bar{x}^{(d)}\right)\right\}\right| + z\right].\end{aligned}$$

Theorem 7 below is the direct consequence of Lemma 2, Lemma 4 and Corollary 2. Remind that $\zeta_r\left(n, \bar{x}^{(d)}\right) := \sup_{x_{d+1} \in \mathbb{X}_{d+1}} \left|\xi_{r,\mathfrak{d},\bar{x}^{(d)}}\left(\bar{x}^{(d)}\right)\right|$ and $\widetilde{\mathbf{N}} = \{\mathbf{n_1}, \ldots, \mathbf{n_2}\}$.

Theorem 7. *Let Assumption 9 be verified and suppose that Assumption 7 (***ii***) holds with* $k = d+1, m = d+2$ *and* $(\mathfrak{H}_{d+1}, \varrho_{d+1}) = (\mathcal{Z}, [\mathfrak{d}]^{\alpha})$. *Suppose also that Assumption 7 (***i***) is fulfilled with* $(\mathfrak{H}_{d+2}, \varrho_{d+2}) = \left(\mathbb{X}_1^d, \rho^{(d)}\right)$ *and* $\mathrm{H}_{d+2,\mathbf{i}} = \mathrm{X}_{\mathbf{i}},\ \mathbf{i} \in \mathbf{I}$, *satisfying Assumption 11. Assume that (3.24) holds as well and if* $\mathbf{n_1} \ne \mathbf{n_2}$ *let* $\left(X_i\right)^d,\ i \ge 1$, *be identically distributed.*

Then for any given decreasing sequences $0 < r_l^{(\min)}(n) \le r_l^{(\max)}(n) \le 1,\ l = \overline{1,d},\ n \ge 1,$ *any* $\boldsymbol{b} > 1,\ q \ge 1,\ v \ge 1$ *and* $z \ge 1$

$$\mathbb{P}_{\mathrm{f}}\left\{ \sup_{n\in\widetilde{\mathbf{N}}} \sup_{\left(r,\bar{x}^{(d)}\right)\in\widetilde{\mathcal{R}}(n)\times\mathbb{X}_1^d} \left[\zeta_r\left(n,\bar{x}^{(d)}\right) - \widehat{\mathcal{U}}^{(v,z,q)}\left(n,r,\bar{x}^{(d)}\right)\right] \ge 0 \right\} \le \mathfrak{n}^5\left\{4838e^{-z} + 2{\mathbf{n_1}}^{2-v}\right\};$$

$$\mathbb{E}_{\mathrm{f}}\left\{ \sup_{n\in\widetilde{\mathbf{N}}} \sup_{\left(r,\bar{x}^{(d)}\right)\in\widetilde{\mathcal{R}}(n)\times\mathbb{X}_1^d} \left[\zeta_r\left(n,\bar{x}^{(d)}\right) - \widehat{\mathcal{U}}^{(v,z,q)}\left(n,r,\bar{x}^{(d)}\right)\right] \right\}_+^q$$

$$\le 2\mathfrak{n}^5 c_q' \left[\sqrt{\frac{\widehat{F}_{\mathbf{n_2}}}{\mathbf{n_1}V_{r^{(\max)}(\mathbf{n_1})}}} \vee \left(\frac{\ln^{\beta}(\mathbf{n_2})}{V_{r^{(\max)}(\mathbf{n_1})}\mathbf{n_1}}\right)\right]^q e^{-z} + 2^{q+1}\mathfrak{n}^5\left(V_{r^{(\min)}(\mathbf{n_1})}\right)^{-q}\ {\mathbf{n_1}}^{2-v}.$$

Remind that $\widehat{F}_{\mathbf{n_2}} = \sup_{n\in\widetilde{\mathbf{N}}} \sup_{\left(r,\bar{x}^{(d)}\right)\in\widetilde{\mathcal{R}}(n)\times\mathbb{X}_1^d} \widehat{F}_{\mathbf{n_2}}\left(r,\bar{x}^{(d)}\right)$ and the expression for the constant c_q' can be found in Theorem 5. We also note that the first assertion of the theorem remains valid if one replaces the quantity $\widehat{M}_{q,v}(r)$ by the smaller quantity $\left([36d+54N]\delta_*^{-2} + 2v + 6\right)\ln\left(2V_r^{-1}\right) + \mathbf{C}/2$. But the corresponding upper function will differ from $\widehat{\mathcal{U}}^{(v,z,q)}$ only by numerical constant.

We also remark that $\widehat{F}_{\mathbf{n_2}} \le 2^d \mathrm{f}_\infty \|g\|_\infty L_2 \prod_{l=1}^d 2^{\gamma_l} L^{(l)}$ for any $\mathbf{n_2} \ge 3$ under Assumption 10 in view of Lemma 3. Moreover, if $V_{r^{(\min)}(n)} \ge n^{-p}$ for some $p > 0$ then $\widehat{M}_{q,v}(r)$ can be bound from above by $\left([72d+108N]\delta_*^{-2} + 2.5q + 2v + 1.5\right)p\ln(2n)$ which is independent on r. Hence, if both restrictions are fulfilled the upper function $\widehat{\mathcal{U}}^{(v,z,q)}$ in Theorem 7 takes rather simple form, namely

$$\boldsymbol{\lambda_1}(q)\sqrt{\frac{\ln(n)+z}{nV_r}} + \frac{\boldsymbol{\lambda_2}(q)\left[\ln^{\beta+1}(n)+z\right]}{nV_r},$$

where the constant $\boldsymbol{\lambda_1}(q)$ and $\boldsymbol{\lambda_2}(q)$ can be easily computed.

Law of logarithm In this paragraph we will additionally suppose that Assumption 10 holds. Then, we remark first that statements $1^0 - 3^0$ of Remark 11 hold. Next, we note that $L_{d+1}(z) = z$ and $L_{d+2}(z) = z^2$ in view of Lemma 2 that implies $\mathcal{L}^{(k)}(z) = \ln(z)$ for any $z \ge 1$. Hence, the condition (3.17) is fulfilled with $\mathfrak{a} = 1$.

Thus, all assumptions of Theorem 4 are checked and, taking into account that in our case

$$\eta_{\mathfrak{h}^{(k)}}(n) = \left\|\zeta_r(n)\right\|_\infty := \sup_{\bar{x}^{(d)}\in\mathbb{X}_1^d} \zeta_r\left(n,\bar{x}^{(d)}\right),$$

we come to the following statement.

Theorem 8. *Let assumptions of Theorem 7 be fulfilled and suppose additionally that that Assumption 10 holds. Then there exists* $\boldsymbol{\Upsilon}$ *such that for any* $a > 4$

$$\mathbb{P}_{\mathrm{f}}\left\{ \sup_{n\ge\mathbf{j}} \sup_{r:\, n^{-1}(\ln n)^a \le V_r \le 1} \frac{\sqrt{nV_r}\,\left\|\zeta_r(n)\right\|_\infty}{\sqrt{\ln\left(V_r^{-1}\right) \vee \ln\ln(n)}} \ge \boldsymbol{\Upsilon} \right\} \le \frac{4840\mathfrak{n}^5}{\ln(\mathbf{j})}.$$

The uniform version over the set of bounded densities, similar to (3.22), holds as well.

The immediate consequence of the latter theorem is so-called "*uniform-in-bandwidth consistency*":

$$\limsup_{n\to\infty} \sup_{r:\, n^{-1}(\ln n)^a \le V_r \le 1} \frac{\sqrt{nV_r}\, \left\|\zeta_r(n)\right\|_\infty}{\sqrt{\ln\left(V_r^{-1}\right) \vee \ln\ln(n)}} \le \boldsymbol{\Upsilon} \quad \mathbb{P}_{\mathrm{f}} - \mathbf{a.s} \tag{3.25}$$

The assertion of Theorem 8 and its corollary (3.25) generalizes in several directions the existing results Einmahl and Mason (2000), Giné and Guillou (2002), Einmahl and Mason (2005), Dony et al. (2006) (see, the discussion after Theorem 6).

We would like to conclude this section with the following remark. If K is compactly supported and $g\big(\mathfrak{z},\cdot\big) = \bar{g}(\cdot)$ for any $\mathfrak{z} \in \mathcal{Z}$, where $\bar{g}$ is a bounded function, then all results of this section remain true under Assumptions 7 (**i**), 9 (**i**), 11, (3.20) and (3.21).

4. Gaussian random functions

In this section we apply Propositions 2-4 to the family of zero-mean gaussian random functions. Thus, let $\chi_\theta, \theta \in \Theta$, is a real valued continuous gaussian random function such that $\mathrm{E}\chi_\theta = 0$, $\forall \theta \in \Theta$. We are interested first in finding an upper function for $\left|\chi_\theta\right|, \theta \in \Theta$. Let

$$V(\theta) = \sqrt{\mathrm{E}\left|\chi_\theta\right|^2}, \quad \rho(\theta_1,\theta_2) = \sqrt{\mathrm{E}\left|\chi_{\theta_1} - \chi_{\theta_2}\right|^2}$$

We remark that Assumption 1 holds with $\mathrm{c} = 2$, $B \equiv 0$ and $\mathrm{b} \equiv 0$ and $\forall A \ge \sqrt{2}V$, $\forall \mathrm{a} \ge \sqrt{2}\rho$. Since $\mathrm{b} \equiv 0$ Assumption 3 is reduced to

Assumption 3 [Gaussian case]. *There exist $s \in \mathbb{S}$ such that for any $x > 0$*

$$\sup_{\delta>0} \delta^{-2} \mathfrak{E}_{\widetilde{\Theta},\,\mathrm{a}}\left(x(48\delta)^{-1} s(\delta)\right) < \infty.$$

Thus, if the latter assumption holds, Propositions 2-4 can be applied.

The aim of this section is to find uppers functions for quite different functionals of various gaussian processes. We would like to emphasize that the original problem is not always related to the consideration of $\left|\chi_\theta\right|, \theta \in \Theta$, although such problems are also studied. The idea is to reduce it (if necessary) to those for which one of Propositions 2-4 can be used. Without special mentionning we will always consider a separable modification of $\chi_\theta, \theta \in \Theta$.

4.1. Upper functions for $\mathbb{L}_p$-norms of Wiener integrals

Let $K : \mathbb{R}^d \to \mathbb{R}$ be a continuous compactly supported function such that $\|K\|_\infty < \infty$. Without loss of generality we will assume that the support of K is $[-1/2, 1/2]^d$. Let $0 < h^{(\min)} \le h^{(\max)} \le 1$ be given numbers. Put $\mathcal{H} = \left[h^{(\min)}, h^{(\max)}\right]$ and let $K_h(u) = h^{-d} K\left(u_1/h, \ldots, u_d/h\right)$, $h \in \mathcal{H}$, $u \in \mathbb{R}^d$. Let $b(\mathrm{d}t)$ is white noise on $\mathbb{R}^d$ and consider the family of gaussian random fields

$$\xi_h(t) = \int_{\mathbb{R}^d} K_h\left(t-u\right) b(\mathrm{d}u), \quad h \in \mathcal{H}.$$

Let $\mathbb{K}_\mu = [-\mu/2, \mu/2]^d$, $\mu \ge 1$, be a given cube and let for any $1 \le p < \infty$

$$\left\|\xi_h\right\|_p = \left(\int_{\mathbb{K}_\mu} \left|\xi_h(t)\right|^p \mathrm{d}t\right)^{\frac{1}{p}}.$$

The objective is to find an upper function for $\left\|\xi_h\right\|_p$ on $\mathcal{H}$ and later on $C_1, C_2 \ldots$, denote the constants completely determined by $d, p, \mu,\ \gamma$ and K. It is worth mentioning that the explicit values of these constants can be found and some of them are given in the proof of the theorem.

We will be interested only the case $2 \le p < \infty$, since for any $p \in [1, 2)$ we obviously have

$$\left\|\xi_h\right\|_p \le (\mu)^{\frac{d(2-p)}{2p}} \left\|\xi_h\right\|_2$$

and, therefore, we can use the upper function found for $p = 2$ for any $p \in [1, 2)$.

Let $\mathbb{B}^s_{q,r}, s > 0,\ 1 \le q, r \le \infty$, denote the Besov space on $\mathbb{R}^d$, see e.g. Edmunds and Triebel (1996), and later on $\mathbb{H}_q(s, L)$ denote the the ball of the radius $L > 0$ in $\mathbb{B}^s_{q,\infty}$.

Suppose that $K \in \mathbb{H}_\infty(\gamma, L)$ and without loss of generality we assume that $L = 1$ that implies in particular that $\|K\|_\infty \le 1$.

Theorem 9. *Assume that $\gamma > d/2$. Then for any $2 \le p < \infty$, $h^{(\min)}, h^{(\max)} \in (0, 1)$, and $q \ge 1$*

$$\begin{aligned}
\mathrm{P}\left\{ \sup_{h\in\mathcal{H}} \left[\left\|\xi_h\right\|_p - C_1 h^{-d/2} \right] \ge 0 \right\} &\le C_2 \exp\left\{ -2^{-3/2} \left(h^{(\max)} \right)^{-2d/p} \right\} \\
\mathrm{E}\left\{ \sup_{h\in\mathcal{H}} \left[\left\|\xi_h\right\|_p - C_1 h^{-d/2} \right] \right\}_+^q &\le C_3(q) \left(h^{(\max)} \right)^{\frac{qd(2-p)}{2p}} \exp\left\{ -2^{-3/2} \left(h^{(\max)} \right)^{-2d/p} \right\}.
\end{aligned}$$

The proof of the theorem is given in Section 7. The constant C_1 involved in the description of found upper functions is bounded function of p on any bounded interval. Thus, the upper functions are independent of p if $p \in [2, p_0]$ for any given $p_0 \ge 2$.

Also it is important to mention that the obtained upper functions are sharp. Indeed, it is not difficult to prove that for any $h > 0$

$$C_4 h^{-d/2} \le \mathrm{E}\left\|\xi_h\right\|_p \le C_5 h^{-d/2}.$$

This, together with the concentration inequality for gaussian processes, Talagrand (1994), yields in particular for any given $h > 0$

$$\mathrm{P}\left\{ \left\|\xi_h\right\|_p \ge \widetilde{C}_1 h^{-d/2} \right\} \le \widetilde{C}_2 \exp\left\{ -\widetilde{C}_3 h^{-2d/p} \right\}.$$

This inequality coincides, up to numerical constants, with the first inequality in Theorem 9 in particular case when $h^{(\min)} = h^{(\max)} = h$.

4.2. Upper functions for local modulus of continuity under doubling condition

Let metric space $(\mathfrak{T}, \mathrm{d})$ be equipped with Borel measure $\varkappa$. This measure is doubling if $\exists Q \ge 1$ such that

$$\varkappa\big\{\mathbb{B}_\mathrm{d}(\mathfrak{t}, 2r)\big\} \le Q\varkappa\big\{\mathbb{B}_\mathrm{d}(\mathfrak{t}, r)\big\},\ \ \forall \mathfrak{t} \in \mathfrak{T},\ \ \forall r > 0,$$

where $\mathbb{B}_\mathrm{d}(\mathfrak{t}, r)$ is the closed ball with center $\mathfrak{t}$ and radius r. For example, if $\mathfrak{T} = \mathbb{R}^d$ and $\varkappa$ is Lebesgue measure then $Q = 2^d$.

As it was proved in Coifman and Weiss (1971) the existence of a doubling measure on $\mathfrak{T}$ implies that the space $\mathfrak{T}$ is doubling. It means that there exists $N_\mathrm{d} \in \mathbb{N}^*$ depending only on Q such that for any $r > 0$ each closed ball in $\mathfrak{T}$ of radius r can be covered by at most N_d closed balls of radius $r/2$. This yields that $\mathbb{B}(\mathfrak{t}, r)$ is totally bounded for any $\mathfrak{t} \in \mathfrak{T}$, $r > 0$ and, moreover,

$$\mathfrak{E}_{\mathbb{B}_\mathrm{d}(\mathfrak{t},r),\,\mathrm{d}}(\delta) \le \ln\left(N_\mathrm{d}\right) \left[\log_2\left(2r/\delta\right)\right]_+,\ \ \forall \delta > 0.$$

Let $\mathfrak{t} \in \mathfrak{T}$ and $r > 0$ be fixed. In this section, using Proposition 4 we establish the upper function for local modulus of continuity of χ_θ on $\Theta := \mathbb{B}_{\mathrm{d}}(\mathfrak{t}, r)$. The simplest consequence of this result will be the law of iterated logarithm (LIL) for $|\chi_\theta - \chi_{\mathfrak{t}}|$ as well as its non-asymptotical version. Studying the local modulus of continuity we are obviously interested in the case when r is small even $r \to 0$. Thus, without loss of generality we will assume that $r \le 1$.

To apply Proposition 4 we need to define the function g_A, $A(\cdot) = \sqrt{2}\rho(\cdot - \mathfrak{t})$, to choose the function $s_1 \in \mathcal{S}_{\mathrm{a},0}$, and to compute the function $\widehat{\mathcal{E}}^{(\varepsilon)}(\Delta, \mathfrak{t})$, $\Delta \in (0, r]$, $\varepsilon \in \bigl(0, \sqrt{2} - 1\bigr]$, given by

$$\widehat{\mathcal{E}}^{(\varepsilon)}(\Delta, \mathfrak{t}) = e^{(\mathrm{a})}_{s_1}\Bigl(g_A\bigl([1+\varepsilon]\Delta\bigr), \mathbb{B}_{\mathrm{d}}\bigl(\mathfrak{t}, [1+\varepsilon]\Delta\bigr)\Bigr), \quad \mathrm{a} = \sqrt{2}\rho.$$

Introduce the function

$$\psi(x) = \sqrt{2} \sup_{\substack{\theta_1, \theta_2 \in \mathbb{B}_{\mathrm{d}}\left(\mathfrak{t}, 1\right): \\ \mathrm{d}\left(\theta_1, \theta_2\right) \le x}} \rho\bigl(\theta_1, \theta_2\bigr), \quad x \in (0, 2],$$

and suppose that $\psi(2) < \infty$. Note that obviously $\psi(0) = 0$, since d is a metric, and ψ is increasing. Moreover, for any $u \in (0, r]$

$$g^*_A(u) := \sqrt{2} \sup_{\theta:\, \mathrm{d}(\theta, \mathfrak{t}) \le u} \rho(\theta, \mathfrak{t}) \le \psi(u),$$

that allows us to put $g_A = \psi$. Denoting ψ^{-1} the inverse function of ψ we assert that $\forall u \in (0, r]$

$$\mathfrak{E}_{\mathbb{B}_{\mathrm{d}}(\mathfrak{t}, u),\, \sqrt{2}\rho}(\delta) \le \mathfrak{E}_{\mathbb{B}_{\mathrm{d}}(\mathfrak{t}, u),\, \mathrm{d}}\Bigl(\psi^{-1}(\delta)\Bigr) \le \ln\left(N_{\mathrm{d}}\right) \bigl[\log_2\bigl\{2u/\psi^{-1}(\delta)\bigr\}\bigr]_+, \ \ \forall \delta > 0. \tag{4.1}$$

Hence, if the function ψ is such that Assumption 4 is fulfilled then Proposition 4 is applicable and that provides us with the upper function for $|\chi_\theta - \chi_{\mathfrak{t}}|$ on $\mathbb{B}_{\mathrm{d}}(\mathfrak{t}, r)$. However, this upper function does not admit an explicit expression, in particular its dependence on the variable Δ cannot be analyzed in general. So, we prefer to impose an additional assumption on the function ψ that allows us to obtain the explicit expression of the upper function for $|\chi_\theta - \chi_{\mathfrak{t}}|$ and analyze it as well as the corresponding probability bound for small values of the radius r. We will not be tending here to the maximal generality and suppose that there exist $0 < \underline{c} \le \overline{c} < \infty$ and $\beta > 0$ such that

$$\underline{c} u^\beta \le \psi(u) \le \overline{c} u^\beta, \quad \forall u \in (0, 1]. \tag{4.2}$$

For example, if $\mathrm{d} = \rho$ one has $\psi(u) = \sqrt{2}u$, and, therefore, (4.2) holds. Under (4.2) obviously $\bigl(\delta/\overline{c}\bigr)^{\frac{1}{\beta}} \le \psi^{-1}(\delta) \le \bigl(\delta/\underline{c}\bigr)^{\frac{1}{\beta}}$ and we get from (4.1)

$$\mathfrak{E}_{\mathbb{B}_{\mathrm{d}}(\mathfrak{t}, u),\, \sqrt{2}\rho}(\delta) \le \ln\left(N_{\mathrm{d}}\right) \Bigl[\log_2\Bigl\{2u\bigl(\delta/\overline{c}\bigr)^{-\frac{1}{\beta}}\Bigr\}\Bigr]_+, \ \ \forall \delta > 0. \tag{4.3}$$

Taking into account that $g_A(u) = \psi(u) \ge \underline{c} u^\beta$ and choosing $s(x) = s^*(x) = (6/\pi^2)\bigl(1 + [\ln x]^2\bigr)^{-1}$, we obtain from (4.3) for any $\Delta \in (0, r]$

$$\begin{aligned}
\widehat{\mathcal{E}}^{(\varepsilon)}(\Delta, \mathfrak{t}) &= \sup_{\delta > 0} \delta^{-2} \mathfrak{E}_{\mathbb{B}_{\mathrm{d}}\left(\mathfrak{t}, [1+\varepsilon]\Delta\right),\, \sqrt{2}\rho}\left(g_A\bigl([1+\varepsilon]\Delta\bigr)(48\delta)^{-1} s(\delta)\right) \\
&\le \sup_{\delta > 0} \delta^{-2} \mathfrak{E}_{\mathbb{B}_{\mathrm{d}}\left(\mathfrak{t}, [1+\varepsilon]\Delta\right),\, \sqrt{2}\rho}\left(\underline{c}\bigl([1+\varepsilon]\Delta\bigr)^\beta (48\delta)^{-1} s(\delta)\right) \\
&\le \beta^{-1} \ln\left(N_{\mathrm{d}}\right) \sup_{\delta > 0} \delta^{-2} \left[\log_2\left(\frac{96\overline{c}}{\underline{c}}\right) + \log_2\left(\frac{\delta}{s^*(\delta)}\right)\right]_+ =: \mathrm{C}(\beta, \underline{c}, \overline{c}, \mathrm{d}).
\end{aligned}$$

As we see $\widehat{\mathcal{E}}^{(\varepsilon)}(\Delta,\mathfrak{t})$ is independent of $\Delta,\mathfrak{t}$ and bounded from above by the constant which is completely determined by the triplet $(\mathfrak{T},\mathrm{d},\varkappa)$ and by the quantities $\underline{c},\overline{c}$ and β. Thus, in view of Proposition 4 the upper function for has the following form (see also Remark 5).

$$\widehat{V}^{(z,\varepsilon)}_{s^*}(\Delta,\mathfrak{t})=\overline{c}(1+\varepsilon)^{1+2\beta}\Delta^{\beta}\sqrt{2\big[1+\varepsilon^{-1}\big]^2\mathrm{C}(\beta,\underline{c},\overline{c},\mathrm{d})+\Big[\ell\Big(2(1+\varepsilon)r/\Delta\Big)+z\Big]},$$

where, remind, $\ell(y)=\ln\{1+\ln(y)\}+2\ln\{1+\ln\{1+\ln(y)\}\}$, $y>0$.

Choose $z=z(r)=\ln\big\{1+\ln\big\{1+\big|\ln(r)\big|\big\}\big\}$ and $\varepsilon=\varepsilon(r):=z^{-1}(r)$ and define

$$\begin{aligned}\mathfrak{a}(r)&=(1+\varepsilon(r))^{1+2\beta}\left(\sup_{\Delta\in(0,r]}\sqrt{\frac{2\big[1+\varepsilon^{-1}(r)\big]^2\mathrm{C}(\beta,\underline{c},\overline{c},\mathrm{d})+\Big[\ell\Big(2[1+\varepsilon(r)]r/\Delta\Big)+z(r)\Big]}{\ln\big\{1+\big|\ln(\Delta)\big|\big\}}}\right);\\ \mathfrak{p}(r)&=\frac{2+\Big[\ln\{1+\ln(1+\varepsilon(r))\}\Big]^{-2}}{1+\ln\big\{1+\big|\ln(r)\big|\big\}}.\end{aligned}$$

We note that if $r\to0$ then

$$\mathfrak{a}(r)=1+\mathcal{O}\left(\frac{\ln\big\{1+\ln\big\{1+\big|\ln(r)\big|\big\}\big\}}{\ln\big\{1+\big|\ln(r)\big|\big\}}\right),\quad \mathfrak{p}(r)=\mathcal{O}\left(\frac{\big[\ln\big\{1+\ln\big\{1+\big|\ln(r)\big|\big\}\big\}\big]^2}{\ln\big\{1+\big|\ln(r)\big|\big\}}\right). \tag{4.4}$$

The following result is immediate consequence of Proposition 4. Put $\mathfrak{m}(\Delta)=\sup_{\theta\in B_{\mathrm{d}}(\mathfrak{t},\Delta)}|\chi_\theta-\chi_\mathfrak{t}|$.

Theorem 10. *Let $\mathfrak{T}$ be doubling space and suppose that (4.2) holds. Then, we have for any $\mathfrak{t}\in\mathfrak{T}$ and any $r\in(0,1)$*

$$\mathrm{P}\left\{\sup_{\Delta\in\big(0,r\big]}\left[\frac{\mathfrak{m}(\Delta)}{\overline{c}\Delta^{\beta}\sqrt{\ln\big\{1+\big|\ln(\Delta)\big|\big\}}}\right]\geq\mathfrak{a}(r)\right\}\leq8\mathfrak{p}(r).$$

The first consequence of Theorem 10 is the law of iterated logarithm. Indeed, taking into account that $\mathfrak{p}(r)\to0$, $\mathfrak{a}(r)\to1$, $r\to0$, we come to the following assertion.

Corollary 3. *Let $\mathfrak{T}$ be doubling space and suppose that (4.2) holds. Then for any $\mathfrak{t}\in\mathfrak{T}$ $\mathrm{P}-a.s.$*

$$\limsup_{\Delta\to0+}\left[\frac{\mathfrak{m}(\Delta)}{\Delta^{\beta}\sqrt{\ln\big\{1+\big|\ln(\Delta)\big|\big\}}}\right]\leq\overline{c}.$$

We note, that although the statement of Corollary 3 is traditional in probability theory, the non-asymptotical statement of Theorem 10 is much more informative.

The next consequence of Theorem 10 seems more curious. We remark that if $\mathfrak{T}$ is doubling with respect to the intrinsic semi-metric ρ, then the normalizing sequence appeared in the theorem is independent of ρ. Moreover, the function $\mathfrak{a}(\cdot)$ depends only on ρ.

Indeed, if $\mathrm{d}=\rho$ then $\phi(u)=\sqrt{2}u$, and therefore, $\beta=1$ and $\underline{c}=\overline{c}=\sqrt{2}$. It yields, in particular, that $\mathrm{C}(\beta,\underline{c},\overline{c},\rho)=\mathrm{C}_{N_\rho}$, where

$$\mathrm{C}_{N_\rho}=\ln(N_\rho)\sup_{\delta>0}\delta^{-2}\left(\big[3+2\log_2(\pi)+\log_2\big\{\delta\big(1+[\ln\delta]^2\big)\big\}\big]_+1\right),$$

and, therefore, $\mathfrak{a}(\cdot)=\mathfrak{a}_{N_\rho}(\cdot)$, where $\mathfrak{a}_{N_\rho}(\cdot)$ is completely determined by ρ via the quantity N_ρ.

Corollary 4. *Let $\mathfrak{T}$ be doubling space with respect to* $\mathrm{d}=\rho$. *Then, for any* $\mathfrak{t}\in\mathfrak{T}$ *and* $r\in(0,1)$

$$\mathrm{P}\left\{\sup_{\Delta\in\left(0,r\right]}\left[\frac{\mathfrak{m}(\Delta)}{\Delta\sqrt{2\ln\left\{1+\left|\ln\left(\Delta\right)\right|\right\}}}\right]\geq\mathfrak{a}_{N_\rho}(r)\right\}\leq 8\mathfrak{p}(r).$$

We note that if $\mathfrak{R}^*$ be the set of metrics ρ for which $\mathfrak{T}$ is doubling and such that $N_\rho\leq N^*$ for some fixed $N^*\in\mathbb{N}^*$ then the function $\mathfrak{a}_{N_\rho}(\cdot)$ in the assertion of Corollary 4 can be replaced by the universal on $\mathfrak{R}^*$ function $\mathfrak{a}_{N^*}(\cdot)$. The corresponding inequality becomes "metric free".

We finish this section by the consideration of several examples, where the condition (4.2) is verified. In these examples $\mathfrak{T}$ is either $\mathbb{R}^d$ or $\mathbb{R}^d_+, d\geq 1$, $\varkappa$ is Lebesgue measure and d is the euclidian distance.

Example 1. Lévy function Here $\chi_{\mathfrak{t}},\ \mathfrak{t}\in\mathbb{R}^d$, is zero-mean gaussian random field with $\rho=\sqrt{d}$. Hence, (4.2) holds with $\underline{c}=\overline{c}=\sqrt{2}$ and $\beta=1/2$. We deduce from Theorem 10 that

$$\mathrm{P}\left\{\sup_{\Delta\in\left(0,r\right]}\left[\frac{\mathfrak{m}(\Delta)}{\sqrt{2\Delta\ln\left\{1+\left|\ln\left(\Delta\right)\right|\right\}}}\right]\geq\mathfrak{a}(r)\right\}\leq 8\mathfrak{p}(r),\quad r\in(0,1).$$

Example 2. Fractional brownian motion Here $\chi_{\mathfrak{t}},\ \mathfrak{t}\in\mathbb{R}_+$, is zero-mean gaussian random process with $\rho=d^{\alpha/2},\alpha\in(0,2]$. Hence, (4.2) holds with $\underline{c}=\overline{c}=\sqrt{2}$ and $\beta=\alpha/2$. We get from Theorem 10

$$\mathrm{P}\left\{\sup_{\Delta\in\left(0,r\right]}\left[\frac{\mathfrak{m}(\Delta)}{\sqrt{2\Delta^\alpha\ln\left\{1+\left|\ln\left(\Delta\right)\right|\right\}}}\right]\geq\mathfrak{a}(r)\right\}\leq 8\mathfrak{p}(r),\quad r\in(0,1).$$

Example 3. Ornstein-Uhlenbeck process Here $\chi_{\mathfrak{t}},\ \mathfrak{t}\in\mathbb{R}_+$, is given by

$$\chi_t=\frac{\sigma}{\sqrt{2\lambda}}b\left(e^{2\theta t}\right)e^{-\lambda t},\quad \lambda,\sigma>0,$$

where b is the standard Wiener process. In this case $\rho=\lambda^{-1/2}\sigma\sqrt{1-\exp\left\{-\lambda d\right\}}$. Since we consider $r\in(0,1)$ then (4.2) holds with $\underline{c}=\sigma\sqrt{2}e^{-1}e^{-\lambda/2}$, $\overline{c}=\sigma\sqrt{2}$, $\beta=1/2$ and we have

$$\mathrm{P}\left\{\sup_{\Delta\in\left(0,r\right]}\left[\frac{\mathfrak{m}(\Delta)}{\sigma\sqrt{2\Delta\ln\left\{1+\left|\ln\left(\Delta\right)\right|\right\}}}\right]\geq\mathfrak{a}(r)\right\}\leq 8\mathfrak{p}(r),\quad r\in(0,1).$$

5. Proof of Propositions 1-3

We start this section with establishing an auxiliary result. Let $\mathfrak{L}$ be a set, d_1 and d_2 be semi-metrics on $\mathfrak{L}$ and let $\mathcal{L}$ be a totally bounded subset of $\mathfrak{L}$ with respect to d_1 and d_2 simultaneously. Let $N_{\mathrm{i}}(\delta),\ \delta>0$, denote the minimal number of balls of the radius δ in the metric $\mathrm{d}_{\mathrm{i}},\ \mathrm{i}=1,2$ needed to cover the set $\mathcal{L}$.

Lemma 5. *Let* $l\in\mathbb{N}^*$ *and* $\delta_{1,j},\delta_{2,j}>0,\ j=\overline{1,l}$ *be an arbitrary numbers. One can construct the finite subset* $L\Big(\left[\delta_{1,j},\delta_{2,j}\right],j=\overline{1,l}\Big):=\{\ell_1,\ldots,\ell_N\}\subset\mathcal{L}$ *with* $N\leq\prod_{j=1}^{l}N_1\Big(\delta_{1,j}/2\Big)N_2\Big(\delta_{2,j}/2\Big)$ *and such that*

$$\forall\ell\in\mathcal{L}\ \ \exists\tilde{\ell}\in L(\left[\delta_{1,j},\delta_{2,j}\right],j=\overline{1,l}\Big):\ \ \mathrm{d}_1\big(\ell,\tilde{\ell}\big)\leq\min_{j=\overline{1,l}}\delta_{1,j},\quad \mathrm{d}_2\big(\ell,\tilde{\ell}\big)\leq\min_{j=\overline{1,l}}\delta_{2,j}.$$

Proof of Lemma 5 Set $N_{\mathrm{i},j} = N_{\mathrm{i}}\big(\delta_{\mathrm{i},j}/2\big),\ \mathrm{i} = 1,2,\ j = \overline{1,l}$. Since $\mathcal{L}$ is totally bounded in $\mathrm{d}_i, i = 1,2$ there exist $L^{(\mathrm{i},j)} = \Big\{\ell_1^{(\mathrm{i},j)}, \ldots, \ell_{N_{\mathrm{i},j}}^{(\mathrm{i},j)}\Big\} \subset \mathcal{L}$ such that

$$\forall \ell \in \mathcal{L}\ \exists\, \tilde{\ell} \in L^{(\mathrm{i},j)} :\ \mathrm{d}_{\mathrm{i}}\big(\ell, \tilde{\ell}\big) \le 2^{-1}\delta_{\mathrm{i},j},\ \mathrm{i} = 1,2,\ j = \overline{1,l}.$$

For any $k_{\mathrm{i}} = 1, \ldots N_{\mathrm{i},j},\ \mathrm{i} = 1,2$, put $\mathcal{L}_{k_{\mathrm{i}}}^{(\mathrm{i},j)} = \Big\{\ell \in \mathcal{L} :\ \mathrm{d}_{\mathrm{i}}\Big(\ell, \ell_{k_{\mathrm{i}}}^{(\mathrm{i},j)}\Big) \le 2^{-1}\delta_{\mathrm{i},j}\Big\}$ and let $\mathcal{L}_{k_1,k_2}^{(j)} = \mathcal{L}_{k_1}^{(1,j)} \cap \mathcal{L}_{k_2}^{(2,j)}$. First we note that for any $j = \overline{1,l}$

$$\mathcal{L} = \bigcup_{k_1=1}^{N_{1,j}} \bigcup_{k_2=1}^{N_{2,j}} \mathcal{L}_{k_1,k_2}^{(j)} \tag{5.1}$$

Moreover, the construction of $\mathcal{L}_{k_1}^{(1,j)}$ and $\mathcal{L}_{k_2}^{(2,j)}$ implies that

$$\mathrm{d}_{\mathrm{i}}\big(l_1, l_2\big) \le \delta_{\mathrm{i},j},\ \mathrm{i} = 1,2,\ \ \forall l_1, l_2 \in \mathcal{L}_{k_1,k_2}^{(j)}. \tag{5.2}$$

Put $\mathcal{N} = \otimes_{j=1}^{l}\Big[\,\{1, \ldots N_{1,j}\} \times \{1, \ldots N_{2,j}\}\,\Big]$ and define for any $\big(\mathrm{k}^{(1)}, \ldots, \mathrm{k}^{(l)}\big) \in \mathcal{N}$

$$\mathcal{L}_{\mathrm{k}^{(1)},\ldots,\mathrm{k}^{(l)}} = \bigcap_{j=1}^{l} \mathcal{L}_{\mathrm{k}^{(j)}}^{(j)}.$$

The choice of an arbitrary point in each $\mathcal{L}_{\mathrm{k}^{(1)},\ldots,\mathrm{k}^{(l)}}$ leads to the construction of $L\Big(\big[\delta_1^{(j)}, \delta_2^{(j)}\big], j = \overline{1,l}\Big)$ in view of (5.1) and (5.2).

It remains to note that the cardinality of $\mathcal{N}$ is equal to $\prod_{j=1}^{l} N_1\Big(\delta_{1,j}/2\Big)N_2\Big(\delta_{2,j}/2\Big)$.

■

5.1. Proof of Proposition 1

I. Probability bound Fix $\vec{s} \in \mathcal{S}_{\mathrm{a,b}}$ and put $s_{1,k} = s_1\Big(2^{k/2}\Big)$ and $s_{2,k} = s_2\left(2^k\right),\ k \ge 0$. For any $k \ge 0$ put $\delta_1(k) = (24)^{-1}\widetilde{\varkappa}_1 2^{-k/2}s_{1,k},\ \delta_2(k) = (24)^{-1}\widetilde{\varkappa}_2 2^{-k}s_{2,k}$ and note that $\delta_1(k), \delta_2(k) \to 0,\ k \to \infty$ since $s_1, s_2 \in \mathbb{S}$.

Let $Z_k = L\Big(\big[\delta_1(k), \delta_2(k)\big]\Big),\ k \ge 0$, be the set constructed in Lemma 5 with $\mathrm{d}_1 = \mathrm{a}$, $\mathrm{d}_2 = \mathrm{b}$, $\mathcal{L} = \widetilde{\Theta}$ and $l = 1$. By $N_k,\ k \ge 0$, we denote the cardinality of Z_k.

Fix $\varepsilon > 0$ and put $\epsilon = \varepsilon/(1+\varepsilon)$, $k_0 = \big\lfloor 2\ln_2(1/\epsilon)\big\rfloor + 1$. Let $\theta_m,\ m = 1, \ldots, N_{k_0}$, be the elements of Z_{k_0}. For any $m = 1, \ldots, N_{k_0}$ define

$$\Theta^{(m)} = \Big\{\theta \in \widetilde{\Theta} :\ \mathrm{a}(\theta, \theta_m) \le \delta_1(k_0),\ \ \mathrm{b}(\theta, \theta_m) \le \delta_2(k_0)\Big\},$$

and remark that the definition of the sets Z_{k_0} implies that $\widetilde{\Theta} = \bigcup_{m=1}^{N_{k_0}} \Theta^{(m)}$.

In view of the last remark we get

$$\mathrm{P}\left\{\sup_{\theta\in\widetilde{\Theta}} \Psi\left(\chi_\theta\right) \ge U_{\vec{s}}^{(\varepsilon)}(y, \widetilde{\varkappa}, \widetilde{\Theta})\right\} \le \sum_{m=1}^{N_{k_0}} \mathrm{P}\left\{\sup_{\theta\in\Theta^{(m)}} \Psi\left(\chi_\theta\right) \ge U_{\vec{s}}^{(\varepsilon)}(y, \widetilde{\varkappa}, \widetilde{\Theta})\right\}. \tag{5.3}$$

For any $\theta \in \widetilde{\Theta}$ let $z_k(\theta)$ be an arbitrary element of Z_k satisfying

$$\mathrm{a}\big(\theta, z_k(\theta)\big) \le \delta_1(k), \quad \mathrm{b}\big(\theta, z_k(\theta)\big) \le \delta_2(k). \tag{5.4}$$

Fix $m = 1, \ldots, N_{k_0}$. The continuity of the mapping $\theta \mapsto \chi_\theta$ guarantees that P-a.s.

$$\chi_\theta = \chi_{\theta_m} + \sum_{k=k_0+1}^{\infty} \left[\chi_{z_k(\theta)} - \chi_{z_{k-1}(\theta)}\right], \quad \forall \theta \in \Theta^{(m)}, \tag{5.5}$$

where $z_{k_0}(\theta) = \theta_m$, $\forall \theta \in \Theta^{(m)}$. Note also that independently of θ for all $k \ge k_0 + 1$

$$\mathrm{a}\big(z_k(\theta), z_{k-1}(\theta)\big) \le \delta_1(k) + \delta_1(k-1) =: \widetilde{\delta}_1(k), \tag{5.6}$$

$$\mathrm{b}\big(z_k(\theta), z_{k-1}(\theta)\big) \le \delta_2(k) + \delta_2(k-1) =: \widetilde{\delta}_2(k). \tag{5.7}$$

This is the simplest consequence of triangle inequality and (5.4). Introduce the sequence $c_k, k \ge 1$:

$$c_k = 4^{-1} \max\left\{s_{1,k}, s_{1,k-1}, s_{2,k}, s_{2,k-1}\right\}$$

and remark that $\sum_{k\ge1} c_k \le 1$ that follows from the assumption $s_1, s_2 \in \mathbb{S}$.

We get from sub-additivity of Ψ, (5.5), (5.6) and (5.7) for any $\theta \in \Theta^{(m)}$

$$\Psi\left(\chi_\theta\right) \le \Psi\left(\chi_{\theta_m}\right) + \sup_{k\ge k_0+1} \sup_{\substack{(u,v)\in Z_k\times Z_{k-1}:\\ \mathrm{a}(u,v)\le\widetilde{\delta}_1(k),\, \mathrm{b}(u,v)\le\widetilde{\delta}_2(k)}} c_k^{-1} \Psi\left(\chi_u - \chi_v\right), \tag{5.8}$$

To simplify the notations we will write U instead of $U_{\vec{s}}^{(\varepsilon)}(y, \widetilde{\varkappa}, \widetilde{\Theta})$ and $\mathcal{E}$ instead of $e_{\vec{s}}(\widetilde{\varkappa}, \widetilde{\Theta})$.

We obtain from (5.8)

$$\mathrm{P}\left\{\sup_{\theta\in\Theta^{(m)}} \Psi\left(\chi_\theta\right) \ge U\right\} \le \mathrm{P}\Big\{\Psi\left(\chi_{\theta_m}\right) \ge U(1+\varepsilon)^{-1}\Big\} \tag{5.9}$$

$$+ \sum_{k=k_0+1}^{\infty} \sum_{\substack{(u,v)\in Z_k\times Z_{k-1}:\\ \mathrm{a}(u,v)\le\widetilde{\delta}_1(k),\, \mathrm{b}(u,v)\le\widetilde{\delta}_2(k)}} \mathrm{P}\Big\{\Psi\left(\chi_u - \chi_v\right) \ge \epsilon U c_k\Big\} =: I_1 + I_2.$$

We have in view of Assumption 1 (*1*)

$$\begin{aligned}
\mathrm{P}\left\{\Psi\left(\chi_{\theta_m}\right) \ge U(1+\varepsilon)^{-1}\right\} &\le \mathrm{c}\exp\left\{-\frac{(1+\varepsilon)^{-2}U^2}{\left\{A\left(\theta_m\right)\right\}^2 + (1+\varepsilon)^{-1}UB\left(\theta_m\right)}\right\} \\
&\le \mathrm{c}\exp\left\{-\frac{(1+\varepsilon)^{-2}U^2}{\widetilde{\varkappa}_1^2 + U\widetilde{\varkappa}_2}\right\} \le \mathrm{c}\exp\left\{-(1+\varepsilon)^{-2}\big(y + 2\epsilon^{-2}\mathcal{E}\big)\right\} \\
&\le \mathrm{c}\exp\left\{-(1+\varepsilon)^{-2}y - \epsilon^{-2}\mathcal{E}\right\}.
\end{aligned} \tag{5.10}$$

In order to get (5.10) we have first used that $\widetilde{\varkappa}_1 \ge \sup_{\theta\in\widetilde{\Theta}} A(\theta)$, $\widetilde{\varkappa}_2 \ge \sup_{\theta\in\widetilde{\Theta}} B(\theta)$. Next, we have used that $U \ge v$, where v is the maximal root of the equation

$$\frac{u^2}{\widetilde{\varkappa}_1^2 + u\widetilde{\varkappa}_2} = y + 2\epsilon^{-2}\mathcal{E}. \tag{5.11}$$

We also have used that $(1+\varepsilon)^{-2} \geq 1/2$.

In view of Assumptions 1 (*2*), 2, (5.6) and (5.7) for any $u, v \in Z_k \times Z_{k-1}$ satisfying $\mathrm{a}(u,v) \leq \widetilde{\delta}_1(k),\ \mathrm{b}(u,v) \leq \widetilde{\delta}_2(k)$ we have

$$\begin{aligned} \mathrm{P}\left\{\Psi\left(\chi_u - \chi_v\right) \geq \epsilon U c_k\right\} &\leq \mathrm{c}\exp\left\{-\frac{\left(\epsilon U c_k\right)^2}{\left\{\mathrm{a}(u,v)\right\}^2 + \left(\epsilon U c_k\right)\,\mathrm{b}(u,v)}\right\} \\ &\leq \mathrm{c}\exp\left\{-\frac{(\epsilon U)^2}{\left\{\widetilde{\delta}_1(k)c_k^{-1}\right\}^2 + U\left\{\widetilde{\delta}_2(k)c_k^{-1}\right\}}.\right\} \end{aligned} \tag{5.12}$$

Here we have used that $\epsilon < 1$. Let us remark that

$$\begin{aligned} \widetilde{\delta}_1(k)c_k^{-1} &\leq 4(24)^{-1}\widetilde{\varkappa}_1\left(2^{-k/2}s_{1,k} + 2^{-(k-1)/2}s_{1,k-1}\right)\min\left\{s_{1,k}^{-1}, s_{1,k-1}^{-1}\right\} \leq \widetilde{\varkappa}_1 2^{-k/2-1}; \\ \widetilde{\delta}_2(k)c_k^{-1} &\leq 4(24)^{-1}\widetilde{\varkappa}_2\left(2^{-k}s_{2,k} + 2^{-k+1}s_{2,k-1}\right)\min\left\{s_{2,k}^{-1}, s_{2,k-1}^{-1}\right\} \leq \widetilde{\varkappa}_2 2^{-k-1}. \end{aligned}$$

Thus, continuing (5.12) we obtain

$$\mathrm{P}\left\{\Psi\left(\chi_u - \chi_v\right) \geq \epsilon U c_k\right\} \leq \mathrm{c}\exp\left\{-\frac{2^{k+1}\epsilon^2\,U^2}{\widetilde{\varkappa}_1^2 + U\widetilde{\varkappa}_2}\right\} \leq \mathrm{c}\exp\left\{-2^{k+1}\epsilon^2\big(y + 2\epsilon^{-2}\mathcal{E}\big)\right\}. \tag{5.13}$$

Here we have used (5.11). Noting that the right hand side of (5.13) does not depend on u, v we get

$$\begin{aligned} I_2 &\leq \mathrm{c}\sum_{k=k_0+1}^{\infty} N_k N_{k-1}\exp\left\{-2^{k+1}\epsilon^2\big(y + 2\epsilon^{-2}\mathcal{E}\big)\right\} \\ &\leq \mathrm{c}\exp\left(-y\right)\sum_{k=k_0+1}^{\infty} N_k N_{k-1}\exp\left\{-2^{k+2}\mathcal{E} - 2^{k-k_0}\right\}. \end{aligned} \tag{5.14}$$

Here we have used the definition of k_0 and that $y \geq 1$.

Let us make several remarks. First we note that in view of Lemma 5

$$\ln\big(N_k\big) \leq \mathfrak{E}_{\widetilde{\Theta},\,\mathrm{a}}\Big((24)^{-1}\widetilde{\varkappa}_1 2^{-1-k/2}s_{1,k}\Big) + \mathfrak{E}_{\widetilde{\Theta},\,\mathrm{b}}\Big((24)^{-1}\widetilde{\varkappa}_2 2^{-k-1}s_{2,k}\Big).$$

Taking into account that $s_{1,k} = s_1\big(2^{k/2}\big)$ and denoting $\delta_1 = 2^{k/2}$ we obtain from (1.9)

$$\mathfrak{E}_{\widetilde{\Theta},\,\mathrm{a}}\Big((24)^{-1}\widetilde{\varkappa}_1 2^{-1-k/2}s_{1,k}\Big) = \mathfrak{E}_{\widetilde{\Theta},\,\mathrm{a}}\Big(\widetilde{\varkappa}_1(48\delta_1)^{-1}s_1(\delta_1)\Big) \leq \delta_1^2 e_{s_1}^{(a)}\big(\widetilde{\varkappa}_1, \widetilde{\Theta}\big) = 2^k e_{s_1}^{(a)}\big(\widetilde{\varkappa}_1, \widetilde{\Theta}\big).$$

Taking into account that $s_{2,k} = s_1\big(2^k\big)$ and denoting $\delta_2 = 2^k$ we obtain from (1.9)

$$\mathfrak{E}_{\widetilde{\Theta},\,\mathrm{b}}\Big((24)^{-1}\widetilde{\varkappa}_2 2^{-1-k}s_{2,k}\Big) = \mathfrak{E}_{\widetilde{\Theta},\,\mathrm{b}}\Big(\widetilde{\varkappa}_2(48\delta_2)^{-1}s_2(\delta_2)\Big) \leq \delta_2 e_{s_2}^{(b)}\big(\widetilde{\varkappa}_2, \widetilde{\Theta}\big) = 2^k e_{s_2}^{(b)}\big(\widetilde{\varkappa}_2, \widetilde{\Theta}\big).$$

Thus, we have for any $k \geq 1$

$$\ln\big(N_k N_{k-1}\big) \leq 2^{k+1}\left[e_{s_1}^{(a)}(\widetilde{\Theta}) + e_{s_2}^{(b)}(\widetilde{\Theta})\right] = 2^{k+1}\mathcal{E} \tag{5.15}$$

and, therefore, $\forall k \geq k_0 + 1$

$$\ln\big(N_k N_{k-1}\big) - 2^{k+2}\mathcal{E} \leq -2^{k_0+2}\mathcal{E} \leq -4\epsilon^{-2}\mathcal{E}.$$

It yields together with (5.14)

$$I_2 \;\leq\; \mathrm{c}\exp\left\{-y-4\epsilon^{-2}\mathcal{E}\right\}. \tag{5.16}$$

We get from (5.9), (5.10) and (5.16) for any $m=1,\ldots,N_{k_0}$

$$\mathrm{P}\left\{\sup_{\theta\in\Theta^{(m)}}\Psi\left(\chi_\theta\right)\geq U\right\}\leq 2\mathrm{c}\exp\left\{-y(1+\varepsilon)^{-2}-\epsilon^{-2}\mathcal{E}\right\}. \tag{5.17}$$

The last bound is independent of m and we have from (5.3) and (5.17)

$$\mathrm{P}\left\{\sup_{\theta\in\widetilde{\Theta}}\Psi\left(\chi_\theta\right)\geq U\right\}\leq 2\mathrm{c}N_{k_0}\exp\left\{-y(1+\varepsilon)^{-2}-\epsilon^{-2}\mathcal{E}\right\}.$$

It remains to note that $(1-3\epsilon)^2=(1+\varepsilon)^{-1}$ and that, similarly to (5.15),

$$\ln\left(N_{k_0}\right)\leq 2^{k_0}\mathcal{E}\leq\epsilon^2\mathcal{E},$$

and we come to the first assertion of the proposition.

II. Moment's bound We get for any $y\geq 1$

$$\begin{aligned} E \;&:=\; \mathrm{E}\left(\sup_{\theta\in\widetilde{\Theta}}\Psi\left(\chi_\theta\right)-U\right)_+^q = q\int_0^\infty x^{q-1}\mathrm{P}\left\{\sup_{\theta\in\widetilde{\Theta}}\Psi\left(\chi_\theta\right)\geq U+x\right\}\mathrm{d}x \\ &=\; q\left[U\right]^q\int_0^\infty v^{q-1}\mathbb{P}\left\{\sup_{\theta\in\widetilde{\Theta}}\Psi\left(\chi_\theta\right)\geq(1+v)U\right\}\mathrm{d}v. \end{aligned} \tag{5.18}$$

Note that $(1+v)U\geq U_{\vec{s}}^{(\varepsilon)}\Big((1+v)y,\widetilde{\varkappa},\widetilde{\Theta}\Big)$. Therefore, applying the first statement of the proposition, where y is replaced by vy we obtain from (5.18)

$$E\leq 2\mathrm{c}\Gamma(q+1)\left[(1+\varepsilon)^2y^{-1}U\right]^q\exp\left\{-y/(1+\varepsilon)^2\right\}.$$

■

5.2. Proof of Proposition 2

We start with establishing some technical results used in the sequel.

Preliminaries 1^0. First we formulate the simple consequence of Proposition 1.

Let Θ_1,Θ_2 be given subsets of Θ. For any $\vec{s}=(s_1,s_2)\in\mathcal{S}_{\mathrm{a,b}}$ and any $\varkappa=(\varkappa_1,\varkappa_2)\in\mathbb{R}_+^2\setminus\{0\}$ introduce the following quantity

$$e_{\vec{s}}\big(\varkappa,\Theta_1,\Theta_2\big)=e_{s_1}^{(\mathrm{a})}\big(\varkappa_1,\Theta_1\big)+e_{s_2}^{(\mathrm{b})}\big(\varkappa_2,\Theta_2\big).$$

Put any $\varepsilon>0$ and any $y\geq 0$

$$U_{\vec{s}}^{(\varepsilon)}\big(y,\varkappa,\Theta_1,\Theta_2\big)=\varkappa_1\sqrt{2\big[1+\varepsilon^{-1}\big]^2e_{\vec{s}}\big(\varkappa,\Theta_1,\Theta_2\big)+y}+\varkappa_2\Big(2\big[1+\varepsilon^{-1}\big]^2e_{\vec{s}}\big(\varkappa,\Theta_1,\Theta_2\big)+y\Big).$$

Lemma 6. *Let Assumptions 1-3 hold and let Θ_1, Θ_2 be given subsets of Θ. Let $\varkappa$ be chosen such that $\varkappa_1 \geq \sup_{\theta\in\Theta_1} A(\theta)$ and $\varkappa_2 \geq \sup_{\theta\in\Theta_2} B(\theta)$. Then $\forall \vec{s} \in \mathcal{S}_{\mathrm{a,b}}$, $\forall \varepsilon \in \left(0, \sqrt{2}-1\right]$ and $\forall y \geq 1$,*

$$\mathrm{P}\left\{\sup_{\theta\in\Theta_1\cap\Theta_2} \Psi\left(\chi_\theta\right) \geq U_{\vec{s}}^{(\varepsilon)}\big(y, \varkappa, \Theta_1, \Theta_2\big)\right\} \leq 2\mathrm{c}\exp\left\{-y/(1+\varepsilon)^2\right\}.$$

Moreover, for any $q \geq 1$, putting $C_{\varepsilon,q} = 2\mathrm{c}\Gamma(q+1)(1+\varepsilon)^{2q}$, one has

$$\mathrm{E}\left\{\sup_{\theta\in\Theta_1\cap\Theta_2} \Psi\left(\chi_\theta\right) - U_{\vec{s}}^{(\varepsilon)}\big(y, \varkappa, \Theta_1, \Theta_2\big)\right\}_+^q \leq C_{\varepsilon,q}\left[y^{-1}U_{\vec{s}}^{(\varepsilon)}\big(y, \varkappa, \Theta_1, \Theta_2\big)\right]^q \exp\left\{-\frac{y}{(1+\varepsilon)^2}\right\}.$$

To prove the lemma it suffices to note the following simple facts. In view of the assumptions imposed on $\varkappa$ and obvious inclusions $\Theta_1 \cap \Theta_2 \subseteq \Theta_1$, $\Theta_1 \cap \Theta_2 \subseteq \Theta_2$ we have

$$\varkappa_1 \geq \sup_{\theta\in\Theta_1\cap\Theta_2} A(\theta), \quad \varkappa_2 \geq \sup_{\theta\in\Theta_1\cap\Theta_2} B(\theta), \quad e_{\vec{s}}\big(\varkappa, \Theta_1, \Theta_2\big) \geq e_{\vec{s}}\big(\varkappa, \Theta_1 \cap \Theta_2\big).$$

It yields $U_{\vec{s}}^{(\varepsilon)}\big(y, \varkappa, \Theta_1, \Theta_2\big) \geq U_{\vec{s}}^{(\varepsilon)}\big(y, \varkappa, \Theta_1 \cap \Theta_2\big)$ and to get the assertion of the lemma we apply Proposition 1 with $\widetilde{\Theta} = \Theta_1 \cap \Theta_2$.

2^0. Note that $\Psi\big(\chi_\bullet\big) : \Theta \to \mathbb{R}_+$ is obviously P-a.s. continuous in $\mathfrak{a} \vee \mathfrak{b}$ as a composition of two continuous mappings between metric spaces. Hence Corollary 1 is applicable with $\mathfrak{T} = \Theta$, $\mathfrak{d} = \mathfrak{a}\vee\mathfrak{b}$, $(\boldsymbol{\Omega}, \boldsymbol{\mathfrak{B}}, \mathbf{P}) = (\Omega, \mathfrak{B}, \mathrm{P})$ $\zeta(\mathfrak{t}, \cdot) = \Psi\big(\chi_\theta(\cdot)\big)$ and $g(\mathfrak{t})$ is either $\mathrm{V}_{\vec{s}}^{(z,\varepsilon)}(\theta)$ or $\mathrm{U}_{\vec{s}}^{(z,\varepsilon,q)}(\theta)$, $\mathfrak{t} = \theta$.

Proof of the proposition Put $\delta_l = (1+\varepsilon)^l$, $l \geq 0$, and introduce the following sets

$$\Theta_A^{(l)} = \{\theta \in \Theta : \ \underline{A}\delta_{l-1} \leq A(\theta) \leq \underline{A}\delta_l\}, \quad \Theta_B^{(l)} = \{\theta \in \Theta : \ \underline{B}\delta_{l-1} \leq B(\theta) \leq \underline{B}\delta_l\}, \ l \in \mathbb{N}^*.$$

The idea is to apply Lemma 6 with $\Theta_1 = \Theta_A^{(j)}$ and $\Theta_2 = \Theta_B^{(k)}$ for any given $j, k \geq 1$. To do that we will need to bound from below $\mathrm{V}_{\vec{s}}^{(z,\varepsilon)}(\theta)$ and $\mathrm{U}_{\vec{s}}^{(z,\varepsilon,q)}(\theta)$ on $\Theta_A^{(j)} \cap \Theta_B^{(k)}$. We will consider only j, k such that $\Theta_A^{(j)} \cap \Theta_B^{(k)} \neq \emptyset$ and supremum over empty set is assumed to be zero. Also we will accept the following agreement: if $B \equiv 0$, $b \equiv 0$ then $k \equiv 0$ and $\Theta_B^{(0)} = \widehat{\Theta}$.

Probability bound Let $\theta \in \Theta_A^{(j)} \cap \Theta_B^{(k)}$ be fixed and put $u = \mathcal{A}_\varepsilon(\theta)$, $v = \mathcal{B}_\varepsilon(\theta)$. Note that

$$e_{s_1}^{(\mathrm{a})}\Big(\underline{A}u, \Theta_A\big(\underline{A}u\big)\Big) \geq e_{s_1}^{(\mathrm{a})}\Big(\underline{A}u, \Theta_A\big(\underline{A}\delta_j\big)\Big) \geq e_{s_1}^{(\mathrm{a})}\Big(\underline{A}u, \Theta_A^{(j)}\Big) \geq e_{s_1}^{(\mathrm{a})}\Big(\underline{A}\delta_{j+1}, \Theta_A^{(j)}\Big). \tag{5.19}$$

To get the first two inequalities in (5.19) we have used that $\Theta_A^{(j)} \subseteq \Theta_A\big(\underline{A}\delta_j\big) \subseteq \Theta_A\big(\underline{A}u\big)$ in view of $\delta_j \leq u$ since $\theta \in \Theta_A^{(j)}$. To get the last inequality in (5.19) we have used that the entropy is decreasing function of its argument and that $\delta_{j+1} \geq u$ since $\theta \in \Theta_A^{(j)}$.

By the same reasons

$$e_{s_2}^{(\mathrm{b})}\Big(\underline{B}v, \Theta_B\big(\underline{B}v\big)\Big) \geq e_{s_2}^{(\mathrm{b})}\Big(\underline{B}v, \Theta_B\big(\underline{B}\delta_k\big)\Big) \geq e_{s_2}^{(\mathrm{b})}\Big(\underline{B}v, \Theta_B^{(k)}\Big) \geq e_{s_2}^{(\mathrm{b})}\Big(\underline{B}\delta_{k+1}, \Theta_B^{(k)}\Big). \tag{5.20}$$

Taking into account that left hand sides in (5.19) and (5.20) are independent of θ, whenever $\theta \in \Theta_A^{(j)} \cap \Theta_B^{(k)}$, we deduce from (2.3), (5.19) and (5.20)

$$\begin{aligned}\mathcal{E}_{\vec{s}}\Big(\mathcal{A}_\varepsilon(\theta), \mathcal{B}_\varepsilon(\theta)\Big) &:= e_{s_1}^{(\mathrm{a})}\Big(\underline{A}u, \Theta_A\big(\underline{A}u\big)\Big) + e_{s_2}^{(\mathrm{b})}\Big(\underline{B}v, \Theta_B\big(\underline{B}v\big)\Big) \\ &\geq e_{s_1}^{(\mathrm{a})}\Big(\underline{A}\delta_{j+1}, \Theta_A^{(j)}\Big)\Big) + e_{s_2}^{(\mathrm{b})}\Big(\underline{B}\delta_{k+1}, \Theta_B^{(k)}\Big) = e_{\vec{s}}\left(\varkappa, \Theta_A^{(j)}, \Theta_B^{(k)}\right), \end{aligned} \tag{5.21}$$

for any $\theta \in \Theta_A^{(j)} \cap \Theta_B^{(k)}$, where we put $\varkappa = \big(\underline{A}\delta_{j+1}, \underline{B}\delta_{k+1}\big)$.

We obtain from (5.21), putting $y = (1+\varepsilon)^2\Big[z + \ell\big(\delta_j\big) + \ell\big(\delta_k\big)\Big]$,

$$\mathrm{V}_{\vec{s}}^{(z,\varepsilon)}(\theta) \geq U_{\vec{s}}^{(\varepsilon)}\Big(y, \varkappa, \Theta_A^{(j)}, \Theta_B^{(k)}\Big), \ \ \forall \theta \in \Theta_A^{(j)} \cap \Theta_B^{(k)}. \tag{5.22}$$

Here we have also used that obviously $(1+\varepsilon)^2 A(\theta) \geq \underline{A}\delta_{j+1} =: \varkappa_1$ and $(1+\varepsilon)^2 B(\theta) \geq \underline{B}\delta_{k+1} =: \varkappa_2$ for any $\theta \in \Theta_A^{(j)} \cap \Theta_B^{(k)}$. Moreover, we have used that $2\big[1+\varepsilon^{-1}\big]^2 \geq 4\varepsilon^{-2}$ for any $\varepsilon \in \big(0, \sqrt{2}-1\big]$.

Therefore, we obtain $\forall j, k \geq 1$ in view of (5.22)

$$\Psi_{j,k}^*(z) := \sup_{\theta \in \Theta_A^{(j)} \cap \Theta_B^{(k)}} \Big\{\Psi\left(\chi_\theta\right) - \mathrm{V}_{\vec{s}}^{(z,\epsilon)}(\theta)\Big\} \leq \sup_{\theta \in \Theta_A^{(j)} \cap \Theta_B^{(k)}} \Psi\left(\chi_\theta\right) - U_{\vec{s}}^{(\varepsilon)}\Big(y, \varkappa, \Theta_A^{(j)}, \Theta_B^{(k)}\Big).$$

Since $\varkappa_1 := \underline{A}\delta_{j+1} \geq \sup_{\theta \in \Theta_A^{(j)}} A(\theta)$ and $\varkappa_2 := \underline{B}\delta_{k+1} \geq \sup_{\theta \in \Theta_B^{(k)}} B(\theta)$, Lemma 6 is applicable with $\Theta_1 = \Theta_A^{(j)}$ and $\Theta_2 = \Theta_B^{(k)}$. Thus, applying it we obtain $\forall j, k \geq 1$ and $\forall z \geq 1$

$$\mathrm{P}\left\{\Psi_{j,k}^*(z) \geq 0\right\} \leq 2\mathrm{c}\exp\left\{-z\right\} w_j w_k, \tag{5.23}$$

where we put $w_m = \Big[1 + m\ln\left(1+\epsilon\right)\Big]^{-1}\Big[1 + \ln\left\{1 + m\ln\left(1+\epsilon\right)\right\}\Big]^{-2}$. Noting that

$$\sum_{m=1}^{\infty} w_m \leq 1 + \Big[\ln\left\{1 + \ln\left(1+\varepsilon\right)\right\}\Big]^{-2},$$

taking into account that the union of $\Big\{\Theta_A^{(j)} \cap \Theta_B^{(k)},\ j, k \geq 1, \Big\}$ covers Θ and summing up the right hand side in (5.23) over j, k, we arrive at the first assertion of the proposition.

Moment's bound Using the same arguments having led to (5.22) and taking into account (5.21) we obtain for any $\theta \in \Theta^{(j,k)}$ and any $q \geq 1$

$$\mathrm{U}_{\vec{s}}^{(z,\varepsilon,q)}(\theta) \geq U_{\vec{s}}^{(\varepsilon)}\Big(y, \varkappa, \Theta_A^{(j)}, \Theta_B^{(k)}\Big),$$

where $\ y = (1+\varepsilon)^2\Big[z + \tau_\varepsilon e_{\vec{s}}\Big(\varkappa, \Theta_A^{(j)}, \Theta_B^{(k)}\Big) + (\varepsilon+q)\ln\big(\delta_j\delta_k\big)\Big],\ \tau_\varepsilon = 2(2+\varepsilon)\varepsilon^{-1}$.

Thus, applying the second assertion of Lemma 6, $\forall j, k \geq 1$ and $\forall z \geq 1$

$$\begin{aligned} E_{j,k} \ &:= \ \mathrm{E}\left(\sup_{\theta \in \Theta^{(j,k)}} \Psi\left(\chi_\theta\right) - U_{\vec{s}}^{(\varepsilon)}\Big(y, \varkappa, \Theta_A^{(j)}, \Theta_B^{(k)}\Big)\right)_+^q \\ &\leq \ 2\mathrm{c}\Gamma(q+1)\left[\frac{U_{\vec{s}}^{(\varepsilon)}\Big(y, \varkappa, \Theta_A^{(j)}, \Theta_B^{(k)}\Big)}{z + \tau_\varepsilon e_{\vec{s}}\Big(\varkappa, \Theta_A^{(j)}, \Theta_B^{(k)}\Big) + (\varepsilon+q)\ln\big(\delta_j\delta_k\big)}\right]^q \big(\delta_j\delta_k\big)^{-\varepsilon-q}\exp\left\{-z\right\}. \end{aligned} \tag{5.24}$$

Putting for brevity $e_{j,k} = e_{\vec{s}}\Big(\varkappa, \Theta_A^{(j)}, \Theta_B^{(k)}\Big)$ and noting that for any $j, k \geq 1$ and $z \geq 1$

$$U_{\vec{s}}^{(\varepsilon)}\Big(y, \varkappa, \Theta_A^{(j)}, \Theta_B^{(k)}\Big) \leq 2\big[\varkappa_1 \vee \varkappa_2\big]\left[4\varepsilon^{-2}e_{j,k} + y\right] \leq 2\big[\underline{A} \vee \underline{B}\big]\left[\delta_{j+1} \vee \delta_{k+1}\right]\big[2(1+\varepsilon^{-1})^2 e_{j,k} + y\big],$$

and that for any $\varepsilon \in \big(0, \sqrt{2}-1\big]$

$$\frac{2(1+\varepsilon^{-1})^2 e_{j,k} + (1+\varepsilon)^2\Big[z + \tau_\varepsilon e_{j,k} + (\varepsilon+q)\ln\big(\delta_j\delta_k\big)\Big]}{z + \tau_\varepsilon e_{j,k} + (\varepsilon+q)\ln\big(\delta_j \vee \delta_k\big)} \le (1+\varepsilon)^4(2\varepsilon)^{-1},$$

we get finally from (5.24) for $j,k \ge 1$

$$E_{j,k} \le 2\mathrm{c}\Gamma(q+1)\big[\varepsilon^{-1}(1+\varepsilon)^5\big]^q\big[\underline{A} \vee \underline{B}\big]^q\big(\delta_j\delta_k\big)^{-\varepsilon}\exp\{-z\}. \tag{5.25}$$

Here we have used that $(\delta_j \vee \delta_k)/(\delta_j\delta_k) \le 1$ since $\delta_j, \delta_k \ge 1$. Taking into account that

$$\mathrm{E}\left(\sup_{\theta\in\Theta}\Big\{\Psi\left(\chi_\theta\right) - \mathrm{U}_{\vec{s}}^{(z,\epsilon,q)}(\theta)\Big\}\right)_+^q \le \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} E_{j,k}$$

and that $\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}\big(\delta_j\delta_k\big)^{-\varepsilon} = \big[(1+\varepsilon)^\varepsilon - 1\big]^{-2}$, we come to the second assertion of the proposition in view of (5.25), where one can replace $(1+\varepsilon)^5$ par $2^{5/2}$ since $\varepsilon \in \big(0, \sqrt{2}-1\big]$. We have also used that $\big[(1+\varepsilon)^\varepsilon - 1\big]^{-2} \le 2\varepsilon^{-4}$ since $\varepsilon \in \big(0, \sqrt{2}-1\big]$.

∎

5.3. Proof of Proposition 3

First we discuss the measurability issue. Note that $\Psi\big(\chi_\bullet\big) : \Theta \to \mathbb{R}_+$ is obviously P-a.s. continuous in $\mathfrak{a} \vee \mathfrak{b}$ as a composition of two continuous mappings between metric spaces. Hence Lemma 1 is applicable with $\mathfrak{T} = \Theta$, $\mathfrak{d} = \mathfrak{a} \vee \mathfrak{b}$, $(\boldsymbol{\Omega}, \boldsymbol{\mathfrak{B}}, \mathbf{P}) = (\Omega, \mathfrak{B}, \mathrm{P})$ $\zeta(\mathfrak{t}, \cdot) = \Psi\big(\chi_\theta(\cdot)\big)$, $\mathfrak{t} = \theta$, $\mathfrak{Z} = \mathfrak{A}$, $\mathfrak{T}_\zeta = \Theta_\alpha$ and $g(\mathfrak{z}) = \widehat{\mathrm{U}}^{(z,\varepsilon,r)}(\alpha)$, $\mathfrak{z} = \alpha$.

The proof of the proposition is similar to those of Proposition 2. Let $\varepsilon \in \big(0, \sqrt{2}-1\big]$ be fixed and put $\delta_l = (1+\varepsilon)^{-l}$, $l \ge 0$, $\delta_l = 1$, $l < 0$. Introduce the following sets:

$$\mathfrak{A}_1^{(l)} = \{\alpha \in \mathfrak{A} : \ \overline{\tau}_1\delta_{l+1} \le \tau_1(\alpha) \le \overline{\tau}_1\delta_l\}, \quad \mathfrak{A}_2^{(l)} = \{\alpha \in \mathfrak{A} : \ \overline{\tau}_2\delta_{l+1} \le \tau_2(\alpha) \le \overline{\tau}_2\delta_l\}, \ \ l \in \mathbb{N}.$$

Fix $j,k \ge 0$, $\alpha \in \mathfrak{A}_1^{(j)} \cap \mathfrak{A}_2^{(k)}$ and put $u = (1+\varepsilon)\tau_1(\alpha)$, $v = (1+\varepsilon)\tau_2(\alpha)$. Note that

$$\begin{aligned} e^{(\mathrm{a})}_{s_1(u,\cdot)}\Big(\lambda_1^{-1}g_A(u), \Theta_1'(u)\Big) &\ge e^{(\mathrm{a})}_{s_1(u,\cdot)}\Big(\lambda_1^{-1}g_A(u), \Theta_1'\big(\overline{\tau}_1\delta_j\big)\Big) \\ &\ge e^{(\mathrm{a})}_{s_1(u,\cdot)}\Big(\lambda_1^{-1}g_A\big(\overline{\tau}_1\delta_{j-1}\big), \Theta_1'\big(\overline{\tau}_1\delta_j\big)\Big). \end{aligned} \tag{5.26}$$

To get the first inequality in (5.26) we have used that $\Theta_1'\big(\overline{\tau}_1\delta_j\big) \subseteq \Theta_1'(u)$ in view of $\overline{\tau}_1\delta_j \le u$ since $\alpha \in \mathfrak{A}_1^{(j)}$. To get the second inequality in (5.26) we have used that $g_A\big(\overline{\tau}_1\delta_{j-1}\big) \ge g_A(u)$, since g_A is increasing and $\overline{\tau}_1\delta_{j-1} \ge u$ for $\alpha \in \mathfrak{A}_1^{(j)}$. Moreover we have used that the entropy is decreasing function of its argument.

Remembering the definition of $e^{(\mathrm{a})}_{s_1(u,\cdot)}\Big(\lambda_1^{-1}g_A\big(\overline{\tau}_1\delta_{j-1}\big),\Theta_1'\big(\overline{\tau}_1\delta_j\big)\Big)$, see (1.9), we have

$$
\begin{aligned}
&e^{(\mathrm{a})}_{s_1(u,\cdot)}\Big(\lambda_1^{-1}g_A\big(\overline{\tau}_1\delta_{j-1}\big),\Theta_1'\big(\overline{\tau}_1\delta_j\big)\Big)=\sup_{\delta>0}\delta^{-2}\mathfrak{E}_{\Theta_1'\left(\overline{\tau}_1\delta_j\right),\,\mathrm{a}}\left(\lambda_1^{-1}g_A\big(\overline{\tau}_1\delta_{j-1}\big)(48\delta)^{-1}s_1(u,\delta)\right)\\
&=\sup_{\delta>0}\delta^{-2}\mathfrak{E}_{\Theta_1'\left(\overline{\tau}_1\delta_j\right),\,\mathrm{a}}\left(\lambda_1^{-1}g_A\big(\overline{\tau}_1\delta_{j-1}\big)(48\delta)^{-1}s_1\left(\overline{\tau}_1\delta_j,\delta\right)\left[\frac{s_1(u,\delta)}{s_1\left(\overline{\tau}_1\delta_j,\delta\right)}\right]\right)\\
&\geq\sup_{\delta>0}\delta^{-2}\mathfrak{E}_{\Theta_1'\left(\overline{\tau}_1\delta_j\right),\,\mathrm{a}}\left(g_A\big(\overline{\tau}_1\delta_{j-1}\big)(48\delta)^{-1}s_1\left(\overline{\tau}_1\delta_j,\delta\right)\right)\\
&=:e^{(\mathrm{a})}_{s_1\left(\overline{\tau}_1\delta_j,\cdot\right)}\Big(g_A\big(\overline{\tau}_1\delta_{j-1}\big),\Theta_1'\big(\overline{\tau}_1\delta_j\big)\Big).
\end{aligned}
\tag{5.27}
$$

To get (5.27) we have used that $1\leq u/\overline{\tau}_1\delta_j\leq 1+\varepsilon\leq\sqrt{2}$, the definition of λ_1 and, as previously, that the entropy is decreasing function of its argument. We obtain from (5.26) and (5.27)

$$
e^{(\mathrm{a})}_{s_1(u,\cdot)}\Big(\lambda_1^{-1}g_A(u),\Theta_1'(u)\Big)\geq e^{(\mathrm{a})}_{s_1\left(\overline{\tau}_1\delta_j,\cdot\right)}\Big(g_A\big(\overline{\tau}_1\delta_{j-1}\big),\Theta_1'\big(\overline{\tau}_1\delta_j\big)\Big).
\tag{5.28}
$$

By the same reasons we have

$$
e^{(\mathrm{b})}_{s_2(v,\cdot)}\Big(\lambda_2^{-1}g_B(v),\Theta_2'(v)\Big)\geq e^{(\mathrm{b})}_{s_2\left(\overline{\tau}_2\delta_k,\cdot\right)}\Big(g_B\big(\overline{\tau}_2\delta_{k-1}\big),\Theta_2'\big(\overline{\tau}_2\delta_k\big)\Big),
\tag{5.29}
$$

and, we get from (5.28) and (5.29) for any $\alpha\in\mathfrak{A}_1^{(j)}\cap\mathfrak{A}_2^{(k)}$

$$
\begin{aligned}
\widehat{\mathcal{E}}^{(\varepsilon)}(\alpha)\;&:=\;\mathcal{E}'(u,v)\geq e^{(\mathrm{a})}_{s_1\left(\overline{\tau}_1\delta_j,\cdot\right)}\Big(g_A\big(\overline{\tau}_1\delta_{j-1}\big),\Theta_1'\big(\overline{\tau}_1\delta_j\big)\Big)+e^{(\mathrm{b})}_{s_2\left(\overline{\tau}_2\delta_k,\cdot\right)}\Big(g_B\big(\overline{\tau}_2\delta_{k-1}\big),\Theta_2'\big(\overline{\tau}_2\delta_k\big)\Big)\\
&=\;e_{\vec{s}}\Big(\varkappa,\Theta_1'\big(\overline{\tau}_1\delta_j\big),\Theta_2'\big(\overline{\tau}_2\delta_k\big)\Big),
\end{aligned}
\tag{5.30}
$$

where we have put $\varkappa=\Big(g_A\big(\overline{\tau}_1\delta_{j-1}\big),g_B\big(\overline{\tau}_2\delta_{k-1}\big)\Big)$ and $\vec{s}=\Big(s_1\big(\overline{\tau}_1\delta_j,\cdot\big),s_2\big(\overline{\tau}_2\delta_k,\cdot\big)\Big)$.

Note also that for any $\alpha\in\mathfrak{A}_1^{(j)}\cap\mathfrak{A}_2^{(k)}$ in view of monotonicity of functions g_A and g_B

$$
g_A\left([1+\varepsilon]^2\tau_1(\alpha)\right)\geq g_A\big(\overline{\tau}_1\delta_{j-1}\big)=:\varkappa_1,\quad g_B\left([1+\varepsilon]^2\tau_2(\alpha)\right)\geq g_A\big(\overline{\tau}_2\delta_{k-1}\big)=:\varkappa_2.
\tag{5.31}
$$

Moreover, the definition of sets $\Theta_1'(\cdot)$ and $\Theta_2'(\cdot)$ implies that for any $\alpha\in\mathfrak{A}_1^{(j)}\cap\mathfrak{A}_2^{(k)}$

$$
\Theta_\alpha\subseteq\Theta_1'\big(\overline{\tau}_1\delta_j\big)\cap\Theta_2'\big(\overline{\tau}_2\delta_k\big)
$$

and, therefore, for any $\alpha\in\mathfrak{A}_1^{(j)}\cap\mathfrak{A}_2^{(k)}$

$$
\sup_{\theta\in\Theta_\alpha}\Psi\left(\chi_\theta\right)\leq\sup_{\theta\in\Theta_1'\left(\overline{\tau}_1\delta_j\right)\cap\Theta_2'\left(\overline{\tau}_2\delta_k\right)}\Psi\left(\chi_\theta\right).
\tag{5.32}
$$

Probability bound We get from (5.30) and (5.31) for any $\alpha\in\mathfrak{A}_1^{(j)}\cap\mathfrak{A}_2^{(k)}$

$$
\widehat{\mathrm{U}}^{(z,\varepsilon,0)}(\alpha)\geq U^{(\varepsilon)}_{\vec{s}}\Big(y,\varkappa,\Theta_1'\big(\overline{\tau}_1\delta_j\big),\Theta_2'\big(\overline{\tau}_2\delta_k\big)\Big),\;\;y=(1+\varepsilon)^2\Big[z+R_0\big(\overline{\tau}_1\delta_j,\overline{\tau}_2\delta_k\big)\Big],
$$

where we have used that R_0 is increasing/deacreasing, in both arguments, function.

It yields together with (5.32) $\forall j,k\geq 0$

$$\begin{aligned}\Psi_{j,k}^{(V)}(z) &:= \sup_{\alpha\in\mathfrak{A}_1^{(j)}\cap\mathfrak{A}_2^{(k)}}\left(\sup_{\theta\in\Theta_\alpha}\left\{\Psi\left(\chi_\theta\right)-\widehat{\mathrm{U}}^{(z,\epsilon,0)}(\alpha)\right\}\right)\\ &\leq \sup_{\theta\in\Theta_1'\left(\overline{\tau}_1\delta_j\right)\cap\Theta_2'\left(\overline{\tau}_2\delta_k\right)\Big)}\Psi\left(\chi_\theta\right)-U_{\vec{s}}^{(\varepsilon)}\Big(y,\varkappa,\Theta_1'\big(\overline{\tau}_1\delta_j\big),\Theta_2'\big(\overline{\tau}_2\delta_k\big)\Big).\end{aligned}$$

Let us remark that the definition of the sets $\Theta_A'(\cdot)$, $\Theta_B'(\cdot)$ and the functions g_A and g_B as well as their monotonicity imply that

$$\varkappa_1 := g_A\big(\overline{\tau}_1\delta_{j-1}\big) > g_A\big(\overline{\tau}_1\delta_j\big)\geq g_A^*\big(\overline{\tau}_1\delta_j\big) =: \sup_{\theta\in\Theta_1'\left(\overline{\tau}_1\delta_j\right)}A(\theta); \tag{5.33}$$

$$\varkappa_2 := g_B\big(\overline{\tau}_2\delta_{k-1}\big) > g_B\big(\overline{\tau}_2\delta_k\big)\geq g_B^*\big(\overline{\tau}_2\delta_k\big) =: \sup_{\theta\in\Theta_2'\left(\overline{\tau}_2\delta_k\right)}B(\theta), \tag{5.34}$$

and, therefore, Lemma 6 is applicable with $\Theta_1=\Theta_1'\big(\overline{\tau}_1\delta_j\big)$ and $\Theta_2=\Theta_2'\big(\overline{\tau}_2\delta_k\big)$.

Thus, applying the first assertion of Lemma 6, we obtain $\forall j,k\geq 0$ and $\forall z\geq 1$

$$\mathrm{P}\left\{\Psi_{j,k}^{(V)}(z)\geq 0\right\}\leq 2\mathrm{c}\exp\left\{-z-R_0\big(\overline{\tau}_1\delta_j,\overline{\tau}_2\delta_k\big)\right\}.$$

Noting that the union of $\left\{\mathfrak{A}_1^{(j)}\cap\mathfrak{A}_2^{(k)},\ j=\overline{0,J},k=\overline{0,K},\right\}$ covers $\mathfrak{A}$, summing up the right hand side in the last inequality over j,k we come to the first statement of the proposition.

Moment's bound We get from (5.30) and (5.31) for any $\alpha\in\mathfrak{A}_1^{(j)}\cap\mathfrak{A}_2^{(k)}$

$$\widehat{U}^{(z,\varepsilon,q)}(\alpha)\geq U_{\vec{s}}^{(\varepsilon)}\Big(y,\varkappa,\Theta_1'\big(\overline{\tau}_1\delta_j\big),\Theta_2'\big(\overline{\tau}_2\delta_k\big)\Big),$$

where $y=(1+\varepsilon)^2\Big[z+\tau_\varepsilon e_{\vec{s}}\Big(\varkappa,\Theta_1'\big(\overline{\tau}_1\delta_j\big),\Theta_2'\big(\overline{\tau}_2\delta_k\big)\Big)+R_q\big(\overline{\tau}_1\delta_j,\overline{\tau}_2\delta_k\big)\Big]$, $\tau_\varepsilon=2(2+\varepsilon)\varepsilon^{-1}$.

It yields together with (5.32) $\forall j,k\geq 0$

$$\begin{aligned}\Phi_{j,k}^{(U)}(z) &:= \left\{\sup_{\alpha\in\mathfrak{A}_1^{(j)}\cap\mathfrak{A}_2^{(k)}}\left(\sup_{\theta\in\Theta_\alpha}\left\{\Psi\left(\chi_\theta\right)-\widehat{U}^{(z,\epsilon,q)}(\alpha)\right\}\right)\right\}_+\\ &\leq \left\{\sup_{\theta\in\Theta_1'\left(\overline{\tau}_1\delta_j\right)\cap\Theta_2'\left(\overline{\tau}_2\delta_k\right)\Big)}\left[\Psi\left(\chi_\theta\right)-U_{\vec{s}}^{(\varepsilon)}\Big(y,\varkappa,\Theta_1'\big(\overline{\tau}_1\delta_j\big),\Theta_2'\big(\overline{\tau}_2\delta_k\big)\Big)\right]\right\}_+.\end{aligned}$$

Taking into account (5.33), (5.34) and applying the second assertion of Lemma 6, we have, analogously to (5.24), $\forall j,k\geq 0$ and $\forall z\geq 1$

$$\begin{aligned}&\mathrm{E}\left(\Phi_{j,k}^{(U)}(z)\right)^q\\ &\leq 2\mathrm{c}\Gamma(q+1)\left[\frac{U_{\vec{s}}^{(\varepsilon)}\Big(y,\varkappa,\Theta_1'\big(\overline{\tau}_1\delta_j\big),\Theta_2'\big(\overline{\tau}_2\delta_k\big)\Big)}{z+\tau_\varepsilon e_{\vec{s}}\Big(\varkappa,\Theta_1'\big(\overline{\tau}_1\delta_j\big),\Theta_2'\big(\overline{\tau}_2\delta_k\big)\Big)+R_q\big(\overline{\tau}_1\delta_j,\overline{\tau}_2\delta_k\big)}\right]^q\exp\left\{-z-R_q\big(\overline{\tau}_1\delta_j,\overline{\tau}_2\delta_k\big)\right\}.\end{aligned}$$

Putting for brevity $e_{j,k}=e_{\vec{s}}\big(\varkappa,\Theta_1'(\overline{\tau}_1\delta_j),\Theta_2'(\overline{\tau}_2\delta_k)\big)$ we note that for any $j,k\ge0$ and $z\ge1$

$$U_{\vec{s}}^{(\varepsilon)}\Big(y,\varkappa,\Theta_1'(\overline{\tau}_1\delta_j),\Theta_2'(\overline{\tau}_2\delta_k)\Big)\le2\big[g_A\big(\overline{\tau}_1\delta_{j-1}\big)\vee g_B\big(\overline{\tau}_2\delta_{k-1}\big)\big]\,\big[4\varepsilon^{-2}e_{j,k}+y\big]$$

Repeating the computation done after (5.24) we come to the second assertion of the proposition. ∎

6. Proof of Theorems 1–4

6.1. Proof of Theorem 1

6.1.1. Preliminaries

We start the proof with several technical results used in the sequel. Put for any $i=\overline{1,\mathbf{n_2}}$, $y\in\big[\mathbf{n_1}/\mathbf{n_2},1\big]$ and $\alpha=\boldsymbol{b}\big[\ln(\mathbf{n_2})\big]^{-1}$,

$$\mathcal{Q}_i(y)=1_{(i/\mathbf{n_2},1]}\big(y\big)+\big(\mathbf{n_2}y-i+1\big)^{\alpha}1_{\Delta_i}(y),\quad Q_i(y)=y^{-1}\mathcal{Q}_i(y).$$

Here we have denoted $\Delta_i=\big((i-1)/\mathbf{n_2},i/\mathbf{n_2}\big]$, $i=\overline{3,\mathbf{n_2}}$ and $\Delta_2=\big[1/\mathbf{n_2},2/\mathbf{n_2}\big]$.

For any $a\ge1$ let $\lceil a\rceil$ be the smallest integer larger or equal to a. It implies, in particular, $y\in\Delta_{\lceil\mathbf{n_2}y\rceil}$. First we note that for any $y,\bar{y}\in[\mathbf{n_1}/\mathbf{n_2},1]$ and any $i=\overline{1,\mathbf{n_2}}$

$$\mathcal{Q}_i(y)\le1,\qquad\mathcal{Q}_i(y)=0,\ \forall i>\lceil\mathbf{n_2}y\rceil,\qquad|\mathcal{Q}_i(y)-\mathcal{Q}_i(\bar{y})|\le1\wedge|\mathbf{n_2}(y-\bar{y})|^{\alpha}\,.\tag{6.1}$$

The first and third inequalities imply for any $i=\overline{1,\mathbf{n_2}}$ and any $y,\bar{y}\in[\mathbf{n_1}/\mathbf{n_2},1]$

$$|Q_i(y)-Q_i(\bar{y})|\le(y\wedge\bar{y})^{-1}\left[|\mathbf{n_2}(y-\bar{y})|^{\alpha}+\left(1-\frac{y\wedge\bar{y}}{y\vee\bar{y}}\right)\right].\tag{6.2}$$

For any $z,z'\in\mathbb{R}_+$ denote $\mathfrak{w}\big(z,z'\big)=\left(1-\sqrt{\frac{z\wedge z'}{z\vee z'}}\right)^{1/2}$, and remark that $\mathfrak{w}$ is a metric on $\mathbb{R}_+$. It follows from the relation $\sqrt{2}\mathfrak{w}\big(z,z'\big)=\left[\mathbb{E}\left(\frac{b(z)}{\sqrt{z}}-\frac{b(z')}{\sqrt{z'}}\right)^2\right]^{1/2}$, where b is the standard Wiener process. Taking into account that $y,\bar{y}\ge1/2$ and that $\mathfrak{w}(y\wedge\bar{y})\le1$ we obtain from (6.2)

$$|Q_i(y)-Q_i(\bar{y})|\le8e^{\boldsymbol{b}}\big[\mathfrak{w}(y,\bar{y})\big]^{\alpha}.\tag{6.3}$$

Here we have also used the definition of α. Taking into account that for any $\boldsymbol{a}\le\boldsymbol{c}$

$$\sup_{p\in(0,1]}p^{\boldsymbol{a}}\,(1-\ln(p))^{\boldsymbol{c}}=e^{\boldsymbol{a}-\boldsymbol{c}}\big[\boldsymbol{c}/\boldsymbol{a}\big]^{\boldsymbol{c}},$$

we obtain from (6.3) for any $\boldsymbol{b}>0$, $y,\bar{y}\in[\mathbf{n_1}/\mathbf{n_2},1]$ and $\mathbf{n_2}\ge3$

$$\sup_{i=\overline{1,\mathbf{n_2}}}|Q_i(y)-Q_i(\bar{y})|\le8e\left[\frac{\ln(\mathbf{n_2})}{1-\ln\big(\mathfrak{w}(y,\bar{y})\big)}\right]^{\boldsymbol{b}}\tag{6.4}$$

Next, for any $y,\bar{y}\in[\mathbf{n_1}/\mathbf{n_2},1]$

$$|\mathcal{Q}_i(y)-\mathcal{Q}_i(\bar{y})|=0,\quad i\notin\{\lceil\mathbf{n_2}(y\wedge\bar{y})\rceil,\dots,\lceil\mathbf{n_2}(y\vee\bar{y})\rceil\}\,.\tag{6.5}$$

We have for any $y, \bar{y} \in \left[\mathbf{n_1}/\mathbf{n_2}, 1\right]$ in view of the first and third bounds in (6.1) and (6.5)

$$\sum_{i=1}^{\mathbf{n_2}} |\mathcal{Q}_i(y) - \mathcal{Q}_i(\bar{y})|^2 \le \begin{cases} 2\mathbf{n_2}\,|y-\bar{y}|\,, & \lceil \mathbf{n_2}(y\vee\bar{y})\rceil - \lceil \mathbf{n_2}(y\wedge\bar{y})\rceil \ge 3; \\ 3\,|\mathbf{n_2}(y-\bar{y})|^{2\alpha}\,, & \lceil \mathbf{n_2}(y\vee\bar{y})\rceil - \lceil \mathbf{n_2}(y\wedge\bar{y})\rceil \le 2. \end{cases}$$

To get the first inequality we have also used that $\lceil \mathbf{n_2}(y\vee\bar{y})\rceil - \lceil \mathbf{n_2}(y\wedge\bar{y})\rceil \ge 3$ implies $\mathbf{n_2}(y\vee\bar{y} - y\wedge\bar{y}) > 2$ and, therefore, $\lceil \mathbf{n_2}(y\vee\bar{y})\rceil - \lceil \mathbf{n_2}(y\wedge\bar{y})\rceil + 1 \le \mathbf{n_2}(y\vee\bar{y} - y\wedge\bar{y}) + 2 \le 2\mathbf{n_2}(y\vee\bar{y} - y\wedge\bar{y}) = 2\mathbf{n_2}\,|y-\bar{y}|$. Thus, we have for any $y, \bar{y} \in \left[\mathbf{n_1}/\mathbf{n_2}, 1\right]$

$$\sqrt{\sum_{i=1}^{\mathbf{n_2}} |\mathcal{Q}_i(y) - \mathcal{Q}_i(\bar{y}|^2} \le \sqrt{2\mathbf{n_2}\,|y-\bar{y}|} + \sqrt{3}\,|\mathbf{n_2}(y-\bar{y})|^{\alpha} \le 2\sqrt{\mathbf{n_2}}\mathfrak{w}\big(y,\bar{y}\big) + 2\sqrt{3}e^{\boldsymbol{b}}\big[\mathfrak{w}\big(y,\bar{y}\big)\big]^{\alpha}.$$

Hence we get

$$\begin{aligned} \sqrt{\sum_{i=1}^{\mathbf{n_2}} |Q_i(y) - Q_i(\bar{y})|^2} &\le 8\sqrt{\mathbf{n_2}}\mathfrak{w}\big(y,\bar{y}\big) + 4\sqrt{3}e^{\boldsymbol{b}}\big[\mathfrak{w}\big(y,\bar{y}\big)\big]^{\alpha} \\ &\le 8\sqrt{\mathbf{n_2}}\mathfrak{w}\big(y,\bar{y}\big) + 4\sqrt{3}e\left[\frac{\ln(\mathbf{n_2})}{1-\ln\big(\mathfrak{w}(y,\bar{y})\big)}\right]^{\boldsymbol{b}}. \end{aligned}$$

Taking into account that $\sup_{z\ge 1} z^{-1/2}\left[\ln(2ez)\right]^{\boldsymbol{b}} \le (2\boldsymbol{b}/e)^{\boldsymbol{b}}$ we obtain

$$\sqrt{\sum_{i=1}^{\mathbf{n_2}} |Q_i(y) - Q_i(\bar{y})|^2} \le 8\sqrt{\mathbf{n_2}}\left[\mathfrak{w}\big(y,\bar{y}\big) + \sqrt{3/4}e(2\boldsymbol{b}/e)^{\boldsymbol{b}}\left\{1-\ln\big(\mathfrak{w}(y,\bar{y})\big)\right\}^{-\boldsymbol{b}}\right].$$

Finally we get for any $y, \bar{y} \in [\mathbf{n_1}/\mathbf{n_2}, 1]$ and any $\boldsymbol{b} > 1$

$$\sqrt{\sum_{i=1}^{\mathbf{n_2}} |Q_i(y) - Q_i(\bar{y})|^2} \le 8\left[2^{\boldsymbol{b}}+1\right](\boldsymbol{b})^{\boldsymbol{b}}\sqrt{\mathbf{n_2}}\Big[1-\ln\big(\mathfrak{w}(y,\bar{y})\big)\Big]^{-\boldsymbol{b}}. \tag{6.6}$$

6.1.2. Constants

The following constants appeared in the description of upper functions and inequalities found in Theorem 1. Let $\chi = 0$ if $\mathbf{n_1} = \mathbf{n_2}$ and $\chi = 1$ if $\mathbf{n_1} \ne \mathbf{n_2}$.

$$C_{N,R,m,k} = C^{(1)}_{N,R,m,k} + C^{(2)}_{N,R,m,k} + 2\chi\mathbf{a}_{\boldsymbol{b}}, \qquad \mathbf{a}_{\boldsymbol{b}} = 2\delta_*^{-2}\ln(2) + 2\sup_{\delta>\delta_*}(\delta^2\wedge\delta)^{-1}\big(96\delta/s^*(\delta)\big)^{\frac{1}{\boldsymbol{b}}};$$

$$C^{(1)}_{N,R,m,k} = \sup_{\delta>\delta_*}\delta^{-2}\left\{k\left[1+\ln\left(\frac{9216m\delta^2}{[s^*(\delta)]^2}\right)\right]_+ + N(m-k)\left[\log_2\left\{\left(\frac{4608mR\delta^2}{[s^*(\delta)]^2}\right)\right\}\right]_+\right\};$$

$$C^{(2)}_{N,R,m,k} = \sup_{\delta>\delta_*}\delta^{-1}\left\{k\left[1+\ln\left(\frac{9216m\delta}{s^*(\delta)}\right)\right]_+ + N(m-k)\left[\log_2\left\{\left(\frac{4608mR\delta}{s^*(\delta)}\right)\right\}\right]_+\right\}.$$

Put also $C_D := \left[\sup_{j=0,k+1,\dots,m}\ \sup_{z\in[0,1]} D_j'(z)\right]\vee 2$, where D_j' is the first derivative of the function D_j. Set at last, $\boldsymbol{c_b} = 4\sqrt{2}\left[2^{\boldsymbol{b}}+1\right]\boldsymbol{b}^{\boldsymbol{b}}$ and let

$$\lambda_1 = 4\sqrt{2e}\big(\sqrt{C_D}\vee[\chi\boldsymbol{c_b}]\big), \quad \lambda_2 = (16/3)\big(C_D\vee 8e\big), \quad C_{D,\boldsymbol{b}} = \big(\sqrt{2C_D}\vee[\chi\boldsymbol{c_b}]\big)\vee\big[(2/3)\big(C_D\vee 8e\big)\big].$$

6.1.3. Proof of the theorem

1⁰. Put for any $i = \overline{1,n}$

$$\varepsilon\big(\mathfrak{h}, X_i\big) = G\big(\mathfrak{h}, X_i\big) - \mathbb{E}_{\mathrm{f}} G(\mathfrak{h}, X_i),$$

and define for any $y \in (\mathbf{n_1}/\mathbf{n_2}, 1]$ and any $\mathfrak{h} \in \boldsymbol{\mathfrak{H}}$ the random function

$$\boldsymbol{\xi}(y, \mathfrak{h}) = \mathbf{n_2}^{-1} \sum_{i=1}^{\mathbf{n_2}} \varepsilon\big(\mathfrak{h}, X_i\big) Q_i(y). \tag{6.7}$$

We remark that $\xi_{\mathfrak{h}}(p) = \boldsymbol{\xi}\big(p/\mathbf{n_2}, \mathfrak{h}\big)$ for any $p \in \widetilde{\mathbf{N}}$ and any $\mathfrak{h} \in \boldsymbol{\mathfrak{H}}$. Thus, in order to get the assertions of the theorem it suffices to find upper functions for $|\boldsymbol{\xi}(\cdot,\cdot)|$ on $\big[\mathbf{n_1}/\mathbf{n_2}, 1\big] \times \mathfrak{H}\big(\mathbf{n}\big)$ in view of Assumption 5 and the definition of the number $\mathbf{n}$.

In view of Bernstein inequality Assumption 1 is fulfilled with $\theta = \mathrm{h} =: (y, \mathfrak{h})$ and $\bar{\theta} = \bar{\mathrm{h}} =: (\bar{y}, \bar{\mathfrak{h}})$

$$A^2(\theta) = A^2_{\mathrm{f}}(\mathrm{h}) := 2\mathbf{n_2}^{-2} \sum_{i=1}^{\mathbf{n_2}} Q_i^2(y) \mathbb{E}_f G^2(\mathfrak{h}, X_i); \tag{6.8}$$

$$\mathrm{a}^2(\theta, \bar{\theta}) = \mathrm{a}^2_{\mathrm{f}}(\mathrm{h}, \bar{\mathrm{h}}) := 2\mathbf{n_2}^{-2} \sum_{i=1}^{\mathbf{n_2}} \mathbb{E}_f \Big[Q_i(y) G(\mathfrak{h}, X_i) - Q_i(\bar{y}) G(\bar{\mathfrak{h}}, X_i) \Big]^2; \tag{6.9}$$

$$B(\theta) = B_\infty(\mathrm{h}) = (4/3)\mathbf{n_2}^{-1} \Big[\sup_{i=\overline{1,\mathbf{n_2}}} Q_i(y) \Big] \sup_{x \in \mathcal{X}} \big| G(\mathfrak{h}, x) \big|. \tag{6.10}$$

$$\mathrm{b}(\theta, \bar{\theta}) = \mathrm{b}_\infty(\mathrm{h}, \bar{\mathrm{h}}) := (2/3)\mathbf{n_2}^{-1} \sup_{i=\overline{1,n}} \sup_{x \in \mathcal{X}} \big| \varepsilon(\mathfrak{h}, x) Q_i(y) - \varepsilon(\bar{\mathfrak{h}}, x) Q_i(\bar{y}) \big|. \tag{6.11}$$

Note that $\mathrm{a_f}$ and b_∞ are semi-metrics on $\big[\mathbf{n_1}/\mathbf{n_2}, 1\big] \times \boldsymbol{\mathfrak{H}}$ and $\boldsymbol{\xi}(\cdot,\cdot)$ is obviously continuous on $\big[\mathbf{n_1}/\mathbf{n_2}, 1\big] \times \mathfrak{H}\big(\mathbf{n}\big)$ in the topology generated by b_∞. Moreover, A_{f} and B_∞ are bounded and, therefore, Assumption 2 is fulfilled.

Later on we will use the following notation: for any $\mathfrak{Q} : \mathcal{X} \to \mathbb{R}$ put $\|\mathfrak{Q}\|_\infty = \sup_{x \in \mathcal{X}} |\mathfrak{Q}(x)|$.

We obtain from (6.8)–(6.11) and (6.4) for any $\mathrm{h}, \bar{\mathrm{h}} \in [\mathbf{n_1}/\mathbf{n_2}, 1] \times \mathfrak{H}\big(\mathbf{n}\big)$

$$A^2_{\mathrm{f}}(\mathrm{h}) \le 2(\mathbf{n_1})^{-1} F_{\mathbf{n_2}}(\mathfrak{h}) G_\infty\big(\mathfrak{h}^{(k)}\big), \qquad B_\infty(\mathrm{h}) \le (4/3)(\mathbf{n_1})^{-1} G_\infty\big(\mathfrak{h}^{(k)}\big); \tag{6.12}$$

$$\mathrm{b}_\infty(\mathrm{h}, \bar{\mathrm{h}}) \le \frac{4 \ln^\beta (\mathbf{n_2})}{3\mathbf{n_1}} \Big\{ \big\| G(\mathfrak{h}, \cdot) - G(\bar{\mathfrak{h}}, \cdot) \big\|_\infty + \gamma 8e G_\infty\big(\bar{\mathfrak{h}}^{(k)}\big) \Big[1 - \ln\big(\mathfrak{w}(y, \bar{y})\big) \Big]^{-\boldsymbol{b}} \Big\}, \tag{6.13}$$

where, remind, $\gamma = 0$ if $\mathbf{n_1} = \mathbf{n_2}$ and $\gamma = 1$ if $\mathbf{n_1} \ne \mathbf{n_2}$. Here we have used that if $\mathbf{n_1} = \mathbf{n_2}$ the second term in the last inequality disappears.

We also get using (6.1) and (6.6)

$$\begin{aligned}
\mathrm{a_f}(\mathrm{h}, \bar{\mathrm{h}}) \;&\le\; \sqrt{2}\mathbf{n_2}^{-1} \Bigg\{ \sqrt{\sum_{i=1}^{\mathbf{n_2}} Q_i^2(y) \mathbb{E}_f \Big[G(\mathfrak{h}, X_i) - G(\bar{\mathfrak{h}}, X_i) \Big]^2} \\
&\qquad + \sqrt{F_{\mathbf{n_2}}(\bar{\mathfrak{h}}) G_\infty\big(\bar{\mathfrak{h}}^{(k)}\big) \sum_{i=1}^{\mathbf{n_2}} \big(Q_i(y) - Q_i(\bar{y}) \big)^2} \Bigg\}; \\
&\le\; \sqrt{2}(\mathbf{n_1})^{-1/2} \Bigg\{ \sqrt{\big(F_{\mathbf{n_2}}(\mathfrak{h}) + F_{\mathbf{n_2}}(\bar{\mathfrak{h}}) \big) \big\| G(\mathfrak{h}, \cdot) - G(\bar{\mathfrak{h}}, \cdot) \big\|_\infty} \\
&\qquad + \chi \boldsymbol{c_b} \sqrt{2 F_{\mathbf{n_2}}(\bar{\mathfrak{h}}) G_\infty\big(\bar{\mathfrak{h}}^{(k)}\big)} \Big[1 - \ln\big(\mathfrak{w}(y, \bar{y})\big) \Big]^{-\boldsymbol{b}} \Bigg\},
\end{aligned} \tag{6.14}$$

where we have put $\boldsymbol{c_b} = 4\sqrt{2}\left[2^{\boldsymbol{b}}+1\right](\boldsymbol{b})^{\boldsymbol{b}}$. Here we have used that if $\mathbf{n_1} = \mathbf{n_2}$ the second term in the last inequality disappears.

For any $\tau > 0$ put $\mathfrak{H}\big(\mathbf{n},\tau\big) = \big\{\mathfrak{h} \in \mathfrak{H}\big(\mathbf{n}\big) : \ \ F_{\mathbf{n_2}}(\mathfrak{h}) \le \tau\big\}$. Our first step consists in establishing an upper function for $|\boldsymbol{\xi}(\cdot,\cdot)|$ on $\mathrm{H}(\tau) := \big[\mathbf{n_1}/\mathbf{n_2}, 1\big] \times \mathfrak{H}\big(\mathbf{n},\tau\big)$. As always the supremum over empty set is supposed to be zero.

$\mathbf{2^0}$. Note that in view of (6.12) and (6.14) for any $\mathrm{h}, \overline{\mathrm{h}} \in \mathrm{H}(\tau)$

$$A^2_{\mathrm{f}}(\mathrm{h}) \le 2\tau(\mathbf{n_1})^{-1} G_\infty\big(\mathfrak{h}^{(k)}\big), \qquad B_\infty(\mathrm{h}) \le \frac{4\ln^\beta(\mathbf{n_2})}{3\mathbf{n_1}} G_\infty\big(\mathfrak{h}^{(k)}\big); \tag{6.15}$$

$$\mathrm{a_f}(\mathrm{h},\bar{\mathrm{h}}) \le 2\sqrt{\tau}(\mathbf{n_1})^{-1/2}\Big\{\sqrt{\big\|G(\mathfrak{h},\cdot) - G(\bar{\mathfrak{h}},\cdot)\big\|_\infty} + \chi\boldsymbol{c_b}\sqrt{G_\infty\big(\bar{\mathfrak{h}}^{(k)}\big)}\Big[1 - \ln\big(\mathfrak{w}(y,\bar{y})\big)\Big]^{-\boldsymbol{b}}\Big\}. \tag{6.16}$$

Moreover, in view of triangle inequality we obviously have for any $\mathrm{h}, \overline{\mathrm{h}} \in \mathrm{H}(\tau)$

$$\mathrm{a_f}(\mathrm{h},\overline{\mathrm{h}}) \le A_{\mathrm{f}}(\mathrm{h}) + A_{\mathrm{f}}(\overline{\mathrm{h}}) \le \sqrt{8\tau(\mathbf{n_1})^{-1}\left[G_\infty\Big(\mathfrak{h}^{(k)}\Big) \vee G_\infty\Big(\overline{\mathfrak{h}}^{(k)}\Big)\right]}; \tag{6.17}$$

$$\mathrm{b}_\infty(\mathrm{h},\overline{\mathrm{h}}) \le B_\infty(\mathrm{h}) + B_\infty(\overline{\mathrm{h}}) \le \frac{8\ln^\beta(\mathbf{n_2})}{3\mathbf{n_1}}\left[G_\infty\Big(\mathfrak{h}^{(k)}\Big) \vee G_\infty\Big(\overline{\mathfrak{h}}^{(k)}\Big)\right]. \tag{6.18}$$

Set

$$\mathcal{G}\left(\mathfrak{h}^{(k)},\overline{\mathfrak{h}}^{(k)}\right) = G_\infty\Big(\mathfrak{h}^{(k)}\Big) \vee G_\infty\Big(\overline{\mathfrak{h}}^{(k)}\Big).$$

We get for any $\mathfrak{h}, \overline{\mathfrak{h}}$, satisfying $\varrho^{(k)}\Big(\mathfrak{h}^{(k)},\overline{\mathfrak{h}}^{(k)}\Big) \vee \sup\limits_{j=\overline{k+1,m}} \varrho_j\big(\mathfrak{h}_j,\overline{\mathfrak{h}}_j\big) \le 1$ in view of Assumption 4 (**ii**)

$$\big\|G(\mathfrak{h},\cdot) - G(\bar{\mathfrak{h}},\cdot)\big\|_\infty \le C_D\Big\{\mathcal{G}\left(\mathfrak{h}^{(k)},\overline{\mathfrak{h}}^{(k)}\right)\varrho^{(k)}\Big(\mathfrak{h}^{(k)},\overline{\mathfrak{h}}^{(k)}\Big) + \sum_{j=k+1}^{m} L_j\left\{\mathcal{G}\left(\mathfrak{h}^{(k)},\overline{\mathfrak{h}}^{(k)}\right)\right\}\varrho_j\big(\mathfrak{h}_j,\overline{\mathfrak{h}}_j\big)\Big\}.$$

On the other hand, putting $\widetilde{L}_j(y) = L_j(y) \vee y,\ j = 0, k+1, \dots m$, we have for any $\mathfrak{h}, \overline{\mathfrak{h}}$, satisfying $\Big[\varrho^{(k)}\Big(\mathfrak{h}^{(k)},\overline{\mathfrak{h}}^{(k)}\Big) \vee \sup_{j=\overline{k+1,m}} \varrho_j\big(\mathfrak{h}_j,\mathfrak{h}'_j\big)\Big] > 1$

$$\begin{aligned}\big\|G(\mathfrak{h},\cdot) - G(\bar{\mathfrak{h}},\cdot)\big\|_\infty &\le \big\|G(\mathfrak{h},\cdot)\big\|_\infty + \big\|G(\bar{\mathfrak{h}},\cdot)\big\|_\infty \le 2\mathcal{G}\left(\mathfrak{h}^{(k)},\overline{\mathfrak{h}}^{(k)}\right)\\ &\le C_D\Big\{\mathcal{G}\left(\mathfrak{h}^{(k)},\overline{\mathfrak{h}}^{(k)}\right)\varrho^{(k)}\Big(\mathfrak{h}^{(k)},\overline{\mathfrak{h}}^{(k)}\Big) + \sum_{j=k+1}^{m} \widetilde{L}_j\left\{\mathcal{G}\left(\mathfrak{h}^{(k)},\overline{\mathfrak{h}}^{(k)}\right)\right\}\varrho_j\big(\mathfrak{h}_j,\overline{\mathfrak{h}}_j\big)\Big\}.\end{aligned}$$

Here we have also used that $C_D \ge 2$. Thus, finally we have for any $\mathfrak{h}, \overline{\mathfrak{h}}$

$$\big\|G(\mathfrak{h},\cdot) - G(\bar{\mathfrak{h}},\cdot)\big\|_\infty \le C_D\Big\{\mathcal{G}\left(\mathfrak{h}^{(k)},\overline{\mathfrak{h}}^{(k)}\right)\varrho^{(k)}\Big(\mathfrak{h}^{(k)},\overline{\mathfrak{h}}^{(k)}\Big) + \sum_{j=k+1}^{m} \widetilde{L}_j\left\{\mathcal{G}\left(\mathfrak{h}^{(k)},\overline{\mathfrak{h}}^{(k)}\right)\right\}\varrho_j\big(\mathfrak{h}_j,\overline{\mathfrak{h}}_j\big)\Big\}.$$

The latter inequality together with (6.13) and (6.16) yields for any $\mathrm{h}, \overline{\mathrm{h}} \in \mathrm{H}(\tau)$

$$\begin{aligned}\mathrm{a_f}(\mathrm{h},\overline{\mathrm{h}}) \ \ \le \ \ \mathfrak{a}\Big\{&\left(\mathcal{G}\left(\mathfrak{h}^{(k)},\overline{\mathfrak{h}}^{(k)}\right)\varrho^{(k)}\Big(\mathfrak{h}^{(k)},\overline{\mathfrak{h}}^{(k)}\Big) + \sum_{j=k+1}^{m} \widetilde{L}_j\left\{\mathcal{G}\left(\mathfrak{h}^{(k)},\overline{\mathfrak{h}}^{(k)}\right)\right\}\varrho_j\big(\mathfrak{h}_j,\overline{\mathfrak{h}}_j\big)\right)^{1/2}\\ &+\chi\sqrt{\mathcal{G}\left(\mathfrak{h}^{(k)},\overline{\mathfrak{h}}^{(k)}\right)}\Big[1 - \ln\big(\mathfrak{w}(y,\bar{y})\big)\Big]^{-\boldsymbol{b}}\Big\};\end{aligned} \tag{6.19}$$

$$\mathrm{b}_\infty(\mathrm{h},\overline{\mathrm{h}}) \;\le\; \mathfrak{b}\Big\{\mathcal{G}\left(\mathfrak{h}^{(k)},\overline{\mathfrak{h}}^{(k)}\right)\varrho^{(k)}\left(\mathfrak{h}^{(k)},\overline{\mathfrak{h}}^{(k)}\right)+\sum_{j=k+1}^{m}\widetilde{L}_j\left\{\mathcal{G}\left(\mathfrak{h}^{(k)},\overline{\mathfrak{h}}^{(k)}\right)\right\}\varrho_j\big(\mathfrak{h}_j,\overline{\mathfrak{h}}_j\big)$$
$$+\chi\mathcal{G}\left(\mathfrak{h}^{(k)},\overline{\mathfrak{h}}^{(k)}\right)\Big[1-\ln\big(\mathfrak{w}(y,\bar{y})\big)\Big]^{-\boldsymbol{b}}\Big\}, \tag{6.20}$$

where we have put $\mathfrak{a}=2\sqrt{\tau}(\mathbf{n_1})^{-1/2}\big(\sqrt{C_D}\vee[\chi\boldsymbol{c_b}]\big)$, $\mathfrak{b}=\frac{4\left(C_D\vee 8e\right)\ln^{\beta}(\mathbf{n_2})}{3\mathbf{n_1}}$.

3^0. We note that in view of (6.15)) Assumption 1 (*1*) is verified on $\mathrm{H}(\tau)$ with

$$A(\theta)=A\,(\mathrm{h}):=\mathfrak{a}\sqrt{G_\infty\big(\mathfrak{h}^{(k)}\big)},\quad B(\theta)=B\,(\mathrm{h}):=\mathfrak{b}G_\infty\big(\mathfrak{h}^{(k)}\big),\;\;\theta=\mathfrak{h}.$$

The idea now is to apply Proposition 2 with $\Theta=\mathrm{H}(\tau)$. Put

$$\underline{G}_{\mathbf{n}}[\tau]=\inf_{\mathfrak{h}\in\mathfrak{H}\left(\mathbf{n},\tau\right)}G_\infty\big(\mathfrak{h}^{(k)}\big),$$

that yields $\underline{A}=\mathfrak{a}\sqrt{\underline{G}_{\mathbf{n}}[\tau]}$ and $\underline{B}=\mathfrak{b}\underline{G}_{\mathbf{n}}[\tau]$. Choose $s_1=s_2=s^*$. To apply Proposition 2 one has to bound from above the function

$$\mathcal{E}_{\vec{s}}(u,v)=e^{(\mathrm{a})}_{s_1}\Big(\underline{A}u,\Theta_A\big(\underline{A}u\big)\Big)+e^{(\mathrm{b})}_{s_2}\Big(\underline{B}v,\Theta_B\big(\underline{B}v\big)\Big),\quad u,v\ge 1,$$

defined in (2.3). Here, in our case, $\mathrm{a}=\mathrm{a_f}$, $\mathrm{b}=\mathrm{b}_\infty$ and

$$\Theta_A\big(\underline{A}u\big) \;=\; \Big\{\mathfrak{h}\in\mathfrak{H}\big(\mathbf{n},\tau\big):\;\;G_\infty\big(\mathfrak{h}^{(k)}\big)\le u^2\underline{G}_{\mathbf{n}}[\tau]\Big\}\times\big[\mathbf{n_1}/\mathbf{n_2},1\big];$$
$$\Theta_B\big(\underline{B}v\big) \;=\; \Big\{\mathfrak{h}\in\mathfrak{H}\big(\mathbf{n},\tau\big):\;\;G_\infty\big(\mathfrak{h}^{(k)}\big)\le v\underline{G}_{\mathbf{n}}[\tau]\Big\}\times\big[\mathbf{n_1}/\mathbf{n_2},1\big].$$

To compute the function $\widetilde{\mathcal{E}}$ let us make several remarks.

3^0a. First remind that

$$e^{(\mathrm{a_f})}_{s^*}\Big(\underline{A}u,\Theta_A\big(\underline{A}u\big)\Big)=\sup_{\delta>0}\delta^{-2}\mathfrak{E}_{\Theta_A\left(\underline{A}u\right),\,\mathrm{a_f}}\left(\underline{A}u(48\delta)^{-1}s^*(\delta)\right);$$
$$e^{(\mathrm{b}_\infty)}_{s^*}\Big(\underline{B}v,\Theta_B\big(\underline{B}v\big)\Big)=\sup_{\delta>0}\delta^{-1}\mathfrak{E}_{\Theta_B\left(\underline{B}v\right),\,\mathrm{b}_\infty}\left(\underline{B}v(48\delta)^{-1}s^*(\delta)\right).$$

We have in view of (6.17) and (6.18) that for any $\mathrm{h},\overline{\mathrm{h}}\in\mathrm{H}(\tau)$

$$\mathrm{a_f}\big(\mathrm{h},\overline{\mathrm{h}}\big)\le\mathfrak{a}\sqrt{\Big[G_\infty\Big(\mathfrak{h}^{(k)}\Big)\vee G_\infty\Big(\overline{\mathfrak{h}}^{(k)}\Big)\Big]},\quad \mathrm{b}_\infty\big(\mathrm{h},\overline{\mathrm{h}}\big)\le\mathfrak{b}\left[G_\infty\Big(\mathfrak{h}^{(k)}\Big)\vee G_\infty\Big(\overline{\mathfrak{h}}^{(k)}\Big)\right],$$

where we have also used again that $C_D\ge 2$. Therefore,

$$\sup_{\mathrm{h},\overline{\mathrm{h}}\in\Theta_A\left(\underline{A}u\right)}\mathrm{a_f}\big(\mathrm{h},\overline{\mathrm{h}}\big)\le\underline{A}u,\qquad\sup_{\mathrm{h},\overline{\mathrm{h}}\in\Theta_B\left(\underline{B}v\right)}\mathrm{b}_\infty\big(\mathrm{h},\overline{\mathrm{h}}\big)\le\underline{B}v.$$

It yields for any $\delta\le\delta_*$, where remind δ_* be the smallest solution of the equation $(48\delta)^{-1}s^*(\delta)=1$,

$$\mathfrak{E}_{\Theta_A\left(\underline{A}u\right),\,\mathrm{a_f}}\left(\underline{A}u(48\delta)^{-1}s^*(\delta)\right)=0,\quad\mathfrak{E}_{\Theta_B\left(\underline{B}v\right),\,\mathrm{b}_\infty}\left(\underline{B}v(48\delta)^{-1}s^*(\delta)\right)=0$$

and, therefore

$$e_{s^*}^{(\mathrm{a_f})}\Big(\underline{A}u,\Theta_A\big(\underline{A}u\big)\Big)=\sup_{\delta>\delta_*}\delta^{-2}\mathfrak{E}_{\Theta_A\left(\underline{A}u\right),\,\mathrm{a_f}}\left(\underline{A}u(48\delta)^{-1}s^*(\delta)\right); \tag{6.21}$$

$$e_{s^*}^{(\mathrm{b}_\infty)}\Big(\underline{B}v,\Theta_B\big(\underline{B}v\big)\Big)=\sup_{\delta>\delta_*}\delta^{-1}\mathfrak{E}_{\Theta_B\left(\underline{B}v\right),\,\mathrm{b}_\infty}\left(\underline{B}v(48\delta)^{-1}s^*(\delta)\right). \tag{6.22}$$

3^0b. For any $t\geq 1$ put $\mathfrak{H}_1^k(t,\mathbf{n})=\big\{\mathfrak{h}^{(k)}\in\mathfrak{H}_1^k(\mathbf{n}):\ G_\infty\big(\mathfrak{h}^{(k)}\big)\leq\underline{G}_{\mathbf{n}}t\big\}$ and note that the following obvious inclusions hold:

$$\Theta_A\big(\underline{A}u\big)\subseteq\mathfrak{H}_1^k\left(u^2\underline{G}_{\mathbf{n}}[\tau]\underline{G}_{\mathbf{n}}^{-1},\mathbf{n}\right)\times\mathfrak{H}_{k+1}^m\times\big[\mathbf{n_1}/\mathbf{n_2},1\big]; \tag{6.23}$$

$$\Theta_B\big(\underline{B}v\big)\subseteq\mathfrak{H}_1^k\left(v\underline{G}_{\mathbf{n}}[\tau]\underline{G}_{\mathbf{n}}^{-1},\mathbf{n}\right)\times\mathfrak{H}_{k+1}^m\times\big[\mathbf{n_1}/\mathbf{n_2},1\big]. \tag{6.24}$$

For any $\varepsilon>0$ denote by $\mathfrak{N}_t^{(k)}(\varepsilon)$ the minimal number of $\varrho_{\mathbf{n}}^{(k)}$-balls of radius ε needed to cover $\mathfrak{H}_1^k(t,\mathbf{n})$, $\mathfrak{N}_j(\varepsilon),\ j=\overline{k+1,m}$, the minimal number of ϱ_j-balls of radius ε needed to cover $\mathfrak{H}_j$ and let $\mathfrak{N}(\varepsilon)$ be the minimal number of $\mathfrak{w}$-balls of radius ε needed to cover $\big[\mathbf{n_1}/\mathbf{n_2},1\big]$.

Let $\mathbb{H}$ be an arbitrary subset of $\mathfrak{H}_1^k\left(t,\mathbf{n}\right)\times\mathfrak{H}_{k+1}^m\times[1/2,1]$. It is evident that for any given $\epsilon^{(k)}>0,\ \epsilon_j>0,\ j=\overline{k+1,m}$ and $\epsilon>0$ one can construct a net $\Big\{\mathrm{h}(\mathbf{i}),\ \mathbf{i}=\overline{1,\mathbf{I}\big[\mathbb{H}\big]}\Big\}\subset\mathbb{H}$ such that $\forall\mathrm{h}=(\mathfrak{h},y)\in\mathbb{H}\quad\exists\mathbf{i}\in\big\{1,\ldots,\mathbf{I}\big[\mathbb{H}\big]\big\}$

$$\varrho_{\mathbf{n}}^{(k)}\left(\mathfrak{h}^{(k)},\mathfrak{h}^{(k)}(i)\right)\leq\epsilon^{(k)},\quad\varrho_j\Big(\mathfrak{h}_j,\mathfrak{h}_j(i)\Big)\leq\epsilon_j,\ j=\overline{k+1,m},\quad\mathfrak{w}(y,y(\mathbf{i}))\leq\epsilon; \tag{6.25}$$

$$\mathbf{I}\big[\mathbb{H}\big]\leq\mathfrak{N}(\epsilon/2)\,\mathfrak{N}_t^{(k)}\left(\epsilon^{(k)}/2\right)\prod_{j=k+1}^{m}\mathfrak{N}_j\big(\epsilon_j/2\big),\qquad\forall\mathbb{H}\subseteq\mathfrak{H}_1^k\left(t,\mathbf{n}\right)\times\mathfrak{H}_{k+1}^m\times\big[\mathbf{n_1}/\mathbf{n_2},1\big]. \tag{6.26}$$

Moreover we obtain from (6.19) and (6.20) for any $u,v\geq 1$

$$\begin{aligned}\mathrm{a_f}(\mathrm{h},\overline{\mathrm{h}})\ &\leq\ \mathfrak{a}\Bigg\{\left(\underline{G}_{\mathbf{n}}[\tau]u^2\ \varrho_{\mathbf{n}}^{(k)}\Big(\mathfrak{h}^{(k)},\overline{\mathfrak{h}}^{(k)}\Big)+\sum_{j=k+1}^{m}\widetilde{L}_j\left(\underline{G}_{\mathbf{n}}[\tau]u^2\right)\varrho_j\big(\mathfrak{h}_j,\overline{\mathfrak{h}}_j\big)\right)^{1/2}\\&\qquad+\chi u\sqrt{\underline{G}_{\mathbf{n}}[\tau]}\Big[1-\ln\big(\mathfrak{w}(y,\bar{y})\big)\Big]^{-\boldsymbol{b}}\Bigg\},\quad\forall\mathrm{h},\overline{\mathrm{h}}\in\Theta_A\big(\underline{A}u\big);\\ \mathrm{b}_\infty(\mathrm{h},\overline{\mathrm{h}})\ &\leq\ \mathfrak{b}\Bigg\{\underline{G}_{\mathbf{n}}[\tau]v\ \varrho_{\mathbf{n}}^{(k)}\Big(\mathfrak{h}^{(k)},\overline{\mathfrak{h}}^{(k)}\Big)+\sum_{j=k+1}^{m}\widetilde{L}_j\left(\underline{G}_{\mathbf{n}}[\tau]v\right)\varrho_j\big(\mathfrak{h}_j,\overline{\mathfrak{h}}_j\big)\\&\qquad+\chi\underline{G}_{\mathbf{n}}[\tau]v\ \Big[1-\ln\big(\mathfrak{w}(y,\bar{y})\big)\Big]^{-\boldsymbol{b}}\Bigg\},\qquad\forall\mathrm{h},\overline{\mathrm{h}}\in\Theta_B\big(\underline{B}v\big).\end{aligned}$$

Thus, putting $t=t_1:=u^2\underline{G}_{\mathbf{n}}[\tau]\underline{G}_{\mathbf{n}}^{-1}$ and choosing for any $\varsigma>0$

$$\epsilon^{(k)}=\frac{\varsigma^2}{2\mathfrak{a}^2m\underline{G}_{\mathbf{n}}[\tau]u^2},\quad\epsilon_j=\frac{\varsigma^2}{2\mathfrak{a}^2m\widetilde{L}_j\Big(\underline{G}_{\mathbf{n}}(\tau)u^2\Big)},\quad\epsilon=e^{-\left(\frac{2u\mathfrak{a}\sqrt{\underline{G}_{\mathbf{n}}[\tau]}}{\varsigma}\right)^{1/\boldsymbol{b}}}.$$

we obtain in view of (6.23) and (6.25) with $\mathbb{H}=\Theta_A\big(\underline{A}u\big)$

$$\forall\mathrm{h}\in\Theta_A\big(\underline{A}u\big)\quad\exists\mathbf{i}\in\Big\{1,\ldots,\mathbf{I}\Big[\Theta_A\big(\underline{A}u\big)\Big]\Big\}:\ \ \mathrm{a_f}\left(\mathrm{h},\mathrm{h}(\mathbf{i})\right)\leq\varsigma. \tag{6.27}$$

Putting $t = t_2 := v\underline{G}_{\mathbf{n}}[\tau]\underline{G}_{\mathbf{n}}^{-1}$ and choosing

$$\epsilon^{(k)} = \frac{\varsigma}{2\mathfrak{b}m\underline{G}_{\mathbf{n}}[\tau]v}, \quad \epsilon_j = \frac{\varsigma}{2\mathfrak{b}m\widetilde{L}_j\Big(\underline{G}_{\mathbf{n}}[\tau]v\Big)}, \quad \epsilon = e^{-\left(\frac{2v\mathfrak{b}\underline{G}_{\mathbf{n}}[\tau]}{\varsigma}\right)^{1/\boldsymbol{b}}}$$

we obtain in view of (6.23) and (6.25) with $\mathbb{H} = \Theta_B\big(\underline{B}v\big)$

$$\forall \mathrm{h} \in \Theta_B\big(\underline{B}v\big) \quad \exists \mathbf{i} \in \Big\{1, \ldots, \mathbf{I}\Big[\Theta_B\big(\underline{B}v\big)\Big]\Big\} : \ \ \mathrm{b}_\infty\left(\mathrm{h}, \mathrm{h}(\mathbf{i})\right) \le \varsigma. \tag{6.28}$$

We get from (6.26), (6.27) and (6.28) for any $\varsigma > 0$

$$\begin{aligned}\mathfrak{E}_{\Theta_A\left(\underline{A}u\right),\mathrm{a_f}}(\varsigma) \le \mathfrak{E}_{\mathfrak{H}_1^k(t_1,\mathbf{n}),\varrho_{\mathbf{n}}^{(k)}}\left(\frac{\varsigma^2}{4m\mathfrak{a}^2\underline{G}_{\mathbf{n}}[\tau]u^2}\right) + \sum_{j=k+1}^{m}\mathfrak{E}_{\mathfrak{H}_j,\varrho_j}\left(\frac{\varsigma^2}{4m\mathfrak{a}^2\widetilde{L}_j\Big(\underline{G}_{\mathbf{n}}[\tau]u^2\Big)}\right) \\ + \mathfrak{E}_{[\mathbf{n_1}/\mathbf{n_2},1],\mathfrak{w}}\left(2^{-1}\exp\Big\{-\Big(2u\mathfrak{a}\sqrt{\underline{G}_{\mathbf{n}}[\tau]}\varsigma^{-1}\Big)^{1/\boldsymbol{b}}\Big\}\right);\end{aligned} \tag{6.29}$$

$$\begin{aligned}\mathfrak{E}_{\Theta_B\left(\underline{B}v\right),\mathrm{b}_\infty}(\varsigma) \le \mathfrak{E}_{\mathfrak{H}_1^k(t_2,\mathbf{n}),\varrho_{\mathbf{n}}^{(k)}}\left(\frac{\varsigma}{4m\mathfrak{b}\underline{G}_{\mathbf{n}}[\tau]v}\right) + \sum_{j=k+1}^{m}\mathfrak{E}_{\mathfrak{H}_j,\varrho_j}\left(\frac{\varsigma}{4m\mathfrak{b}\widetilde{L}_j\Big(\underline{G}_{\mathbf{n}}[\tau]v\Big)}\right) \\ + \mathfrak{E}_{[\mathbf{n_1}/\mathbf{n_2},1],\mathfrak{w}}\Big(2^{-1}\exp\Big\{-\big(2v\mathfrak{b}\underline{G}_{\mathbf{n}}[\tau]\varsigma^{-1}\big)^{1/\boldsymbol{b}}\Big\}\Big).\end{aligned} \tag{6.30}$$

4⁰. We get in view of Assumption 6

$$\sum_{j=k+1}^{m}\mathfrak{E}_{\mathfrak{H}_j,\varrho_j}\left(\frac{\varsigma^2}{4m\mathfrak{a}^2\widetilde{L}_j\Big(\underline{G}_{\mathbf{n}}[\tau]u^2\Big)}\right) \le N\sum_{j=k+1}^{m}\Big[\log_2\Big\{4\mathfrak{a}^2mR\widetilde{L}_j\Big(\underline{G}_{\mathbf{n}}[\tau]u^2\Big)\varsigma^{-2}\Big\}\Big]_+; \tag{6.31}$$

$$\sum_{j=k+1}^{m}\mathfrak{E}_{\mathfrak{H}_j,\varrho_j}\left(\frac{\varsigma}{4m\mathfrak{b}\widetilde{L}_j\Big(\underline{G}_{\mathbf{n}}[\tau]v\Big)}\right) \le N\sum_{j=k+1}^{m}\Big[\log_2\Big\{4\mathfrak{b}mR\widetilde{L}_j\Big(\underline{G}_{\mathbf{n}}[\tau]v\Big)\varsigma^{-1}\Big\}\Big]_+. \tag{6.32}$$

Taking into account that $\mathfrak{E}_{[\mathbf{n_1}/\mathbf{n_2},1],\mathfrak{w}}(\cdot) \equiv 0$, if $\mathbf{n_1} = \mathbf{n_2}$, and $\mathfrak{E}_{[\mathbf{n_1}/\mathbf{n_2},1],\mathfrak{w}}(\varepsilon) \le \ln\left(2/\varepsilon^2\right)$ for any $\varepsilon \in (0,1]$ and any $\mathbf{n_2} \le 2\mathbf{n_1}$, we have

$$\mathfrak{E}_{[\mathbf{n_1}/\mathbf{n_2},1],\mathfrak{w}}\Big(2^{-1}\exp\Big\{-\Big(2u\mathfrak{a}\sqrt{\underline{G}_{\mathbf{n}}[\tau]}\varsigma^{-1}\Big)^{1/\beta}\Big\}\Big) = \chi\Big(2\ln(2) + 2\Big(2u\mathfrak{a}\sqrt{\underline{G}_{\mathbf{n}}[\tau]}\varsigma^{-1}\Big)^{\frac{1}{\boldsymbol{b}}}\Big); \tag{6.33}$$

$$\mathfrak{E}_{[\mathbf{n_1}/\mathbf{n_2},1],\mathfrak{w}}\Big(2^{-1}\exp\Big\{-\big(2v\mathfrak{b}\underline{G}_{\mathbf{n}}[\tau]\varsigma^{-1}\big)^{1/\beta}\Big\}\Big) = \chi\Big(2\ln(2) + 2\big(2v\mathfrak{b}\underline{G}_{\mathbf{n}}[\tau]\varsigma^{-1}\big)^{\frac{1}{\boldsymbol{b}}}\Big). \tag{6.34}$$

Let us now bound from above $\mathfrak{E}_{\mathfrak{H}_1^k(t,\mathbf{n}),\varrho_{\mathbf{n}}^{(k)}}$. First we note that in view of Assumption 4 (**i**)

$$\mathfrak{H}_1^k(t,\mathbf{n}) \subseteq \big\{\mathfrak{h}_1 \in \mathfrak{H}_1(\mathbf{n}) : \ G_{1,\mathbf{n}}(\mathfrak{h}_1) \le t\underline{G}_{1,\mathbf{n}}\big\} \times \cdots \times \big\{\mathfrak{h}_k \in \mathfrak{H}_k(\mathbf{n}) : \ G_{k,\mathbf{n}}(\mathfrak{h}_k) \le t\underline{G}_{k,\mathbf{n}}\big\}. \tag{6.35}$$

Consider the hyper-rectangle $\mathcal{Z}(t) = \left[\underline{G}_{1,\mathbf{n}}, t\underline{G}_{1,\mathbf{n}}\right] \times \cdots \times \left[\underline{G}_{k,\mathbf{n}}, t\underline{G}_{k,\mathbf{n}}\right]$, $t \ge 1$, which we equip with the metrics

$$\mathfrak{m}^{(k)}\left(z, z'\right) = \max_{i=\overline{1,k}}\left|\ln\left(z_i\right) - \ln\left(z_i'\right)\right|, \quad z, z' \in \mathcal{Z}(t),$$

where $z_i, z_i',\ i=\overline{1,k}$ are the coordinates of z, z' respectively. It easily seen that for any $\varsigma>0$

$$\mathfrak{E}_{\mathcal{Z}(t),\mathfrak{m}^{(k)}}(\varsigma)\le k\left[\ln\ln t-\ln\ln\left(1+\varsigma\right)\right]_+\le k\left(\ln\left(1+\ln t\right)+\left[1+\ln\left(1/\varsigma\right)\right]_+\right).$$

It yields together with (6.35) in view of obvious inequality $\mathfrak{E}_{\mathfrak{H}_1^k(t,\mathbf{n}),\varrho_{\mathbf{n}}^{(k)}}(\varsigma)\le\mathfrak{E}_{\mathcal{Z}(t),\mathfrak{m}^{(k)}}(\varsigma/2)$

$$\mathfrak{E}_{\mathfrak{H}_1^k(t,\mathbf{n}),\varrho_{\mathbf{n}}^{(k)}}(\varsigma)\le k\left(\ln\left(1+\ln t\right)+\left[1+\ln\left(2/\varsigma\right)\right]_+\right). \tag{6.36}$$

We obtain from (6.36)

$$\mathfrak{E}_{\mathfrak{H}_1^k(t_1,\mathbf{n}),\varrho_{\mathbf{n}}^{(k)}}\left(\frac{\varsigma^2}{4m\mathfrak{a}^2\underline{G}_{\mathbf{n}}[\tau]u^2}\right)\le k\left(\ln\left(1+\ln t_1\right)+\left[1+\ln\left(8m\mathfrak{a}^2\underline{G}_{\mathbf{n}}[\tau]u^2\varsigma^{-2}\right)\right]_+\right); \tag{6.37}$$

$$\mathfrak{E}_{\mathfrak{H}_1^k(t_2,\mathbf{n}),\varrho_{\mathbf{n}}^{(k)}}\left(\frac{\varsigma}{4m\mathfrak{b}\underline{G}_{\mathbf{n}}[\tau]v}\right)\le k\left(\ln\left(1+\ln t_2\right)+\left[1+\ln\left(8m\mathfrak{b}\underline{G}_{\mathbf{n}}[\tau]v\varsigma^{-1}\right)\right]_+\right). \tag{6.38}$$

Putting $\widehat{L}_j(z)=z^{-1}\widetilde{L}_j(z)=\max\left\{z^{-1}L_j(z),1\right\}$, we get from (6.21), (6.29), (6.31), (6.33) and (6.37)

$$\begin{aligned}
e_{s^*}^{(\mathrm{a_f})}\left(\underline{A}u,\Theta_A\left(\underline{A}u\right)\right)&\le k\delta_*^{-2}\ln\left(1+\ln\left(u^2\underline{G}_{\mathbf{n}}[\tau]\underline{G}_{\mathbf{n}}^{-1}\right)\right)+N\delta_*^{-2}\sum_{j=k+1}^{m}\log_2\left\{\widehat{L}_j\left(\underline{G}_{\mathbf{n}}[\tau]u^2\right)\right\}\\
&+\sup_{\delta>\delta_*}\delta^{-2}\left\{k\left[1+\ln\left(\frac{9216m\delta^2}{[s^*(\delta)]^2}\right)\right]_+ +N(m-k)\left[\log_2\left\{\left(\frac{4608mR\delta^2}{[s^*(\delta)]^2}\right)\right\}\right]_+\right\}\\
&+\chi\left(2\delta_*^{-2}\ln(2)+2\sup_{\delta>\delta_*}\delta^{-2}\left(96\delta/s^*(\delta)\right)^{\frac{1}{\boldsymbol{b}}}\right)\\
&=k\delta_*^{-2}\ln\left(1+\ln\left(u^2\underline{G}_{\mathbf{n}}[\tau]\underline{G}_{\mathbf{n}}^{-1}\right)\right)+N\delta_*^{-2}\sum_{j=k+1}^{m}\log_2\left\{\widehat{L}_j\left(\underline{G}_{\mathbf{n}}[\tau]u^2\right)\right\}+C_{N,R,m,k}^{(1)}+\chi\mathbf{a}_{\boldsymbol{b}},
\end{aligned}\tag{6.39}$$

where, remind, $\mathbf{a}_{\boldsymbol{b}}=2\delta_*^{-2}\ln(2)+2\sup_{\delta>\delta_*}(\delta^2\wedge\delta)^{-1}\left(96\delta/s^*(\delta)\right)^{\frac{1}{\boldsymbol{b}}}$. Note that $\mathbf{a}_{\boldsymbol{b}}<\infty$ since $\boldsymbol{b}>1$.

Repeating these computations we get from (6.22), (6.30), (6.32), (6.34) and (6.38)

$$\begin{aligned}
e_{s^*}^{(\mathrm{b}_\infty)}\left(\underline{B}v,\Theta_B\left(\underline{B}v\right)\right)&\le k\delta_*^{-1}\ln\left(1+\ln\left(v\underline{G}_{\mathbf{n}}[\tau]\underline{G}_{\mathbf{n}}^{-1}\right)\right)\\
&+N\delta_*^{-1}\sum_{j=k+1}^{m}\log_2\left\{\widehat{L}_j\left(\underline{G}_{\mathbf{n}}[\tau]v\right)\right\}+C_{N,R,m,k}^{(2)}+\chi\mathbf{a}_{\boldsymbol{b}},
\end{aligned}\tag{6.40}$$

We deduce from (6.39) and (6.40) that $\widetilde{\mathcal{E}}_{\vec{s}},\ \vec{s}=(s^*,s^*)$ is bounded from above by the function

$$\begin{aligned}
\mathcal{E}(u,v)&\le k\delta_*^{-2}\ln\left\{\left(1+\ln\left(u^2\underline{G}_{\mathbf{n}}[\tau]\underline{G}_{\mathbf{n}}^{-1}\right)\right)\left(1+\ln\left(v\underline{G}_{\mathbf{n}}[\tau]\underline{G}_{\infty}^{-1}\right)\right)\right\}\\
&+N\delta_*^{-2}\sum_{j=k+1}^{m}\log_2\left[\left\{\widehat{L}_j\left(\underline{G}_{\mathbf{n}}[\tau]u^2\right)\right\}\left\{\widehat{L}_j\left(\underline{G}_{\mathbf{n}}[\tau]v\right)\right\}\right]+C_{N,R,m,k}.
\end{aligned}\tag{6.41}$$

Here we have used that $\delta_*<1$. We note that (6.41) implies in particular Assumption 3 and, therefore, Proposition 2 is applicable with $\Theta=\mathrm{H}(\tau)$.

5⁰. To apply Proposition 2 on $\Theta = \mathrm{H}(\tau)$ we choose $\varepsilon = \sqrt{2}-1$ and bound from above the quantities

$$\begin{aligned}
P_{\sqrt{2}-1}(\mathrm{h}) &:= 4\big[\sqrt{2}-1\big]^{-2}\mathcal{E}\left(\sqrt{2\underline{G}_{\mathbf{n}}^{-1}[\tau]G_{\infty}\big(\mathfrak{h}^{(k)}\big)}, \sqrt{2}\underline{G}_{\mathbf{n}}^{-1}[\tau]G_{\infty}\big(\mathfrak{h}^{(k)}\big)\right)\\
&\quad +2\ell\left(\sqrt{2\underline{G}_{\mathbf{n}}^{-1}[\tau]G_{\infty}\big(\mathfrak{h}^{(k)}\big)}\right)+2\ell\left(\sqrt{2}\underline{G}_{\mathbf{n}}^{-1}[\tau]G_{\infty}\big(\mathfrak{h}^{(k)}\big)\right);\\
M_{\sqrt{2}-1,q}(\mathrm{h}) &:= 8\big[\sqrt{2}-1\big]^{-2}\mathcal{E}\left(\sqrt{2\underline{G}_{\mathbf{n}}^{-1}[\tau]G_{\infty}^{-1}\big(\mathfrak{h}^{(k)}\big)}, \sqrt{2}\underline{G}_{\mathbf{n}}^{-1}[\tau]G_{\infty}\big(\mathfrak{h}^{(k)}\big)\right)\\
&\quad +2\big(\sqrt{2}-1+q\big)\ln\left(\sqrt{2\underline{G}_{\mathbf{n}}^{-1}[\tau]G_{\infty}\big(\mathfrak{h}^{(k)}\big)}\sqrt{2}\underline{G}_{\mathbf{n}}^{-1}[\tau]G_{\infty}\big(\mathfrak{h}^{(k)}\big)\right),
\end{aligned}$$

where remind $\ell(u)=\ln\{1+\ln(u)\}+2\ln\{1+\ln\{1+\ln(u)\}\}$.

Taking into account that $\ell(u)\le 3\ln\{1+\ln(u)\}, u\ge 1$, $\big[\sqrt{2}-1\big]^{-2}\le 9$ and that $\underline{G}_{\mathbf{n}}[\tau]\ge\underline{G}_{\mathbf{n}}$ for any τ, we obtain from (6.41)

$$\begin{aligned}
P_{\sqrt{2}-1}(\mathrm{h}) &\le \big[72k\delta_*^{-2}+12\big]\ln\left\{1+\ln\left(2G_{\infty}\big(\mathfrak{h}^{(k)}\big)\underline{G}_{\mathbf{n}}^{-1}\right)\right\}\\
&\quad +72N\delta_*^{-2}\sum_{j=k+1}^{m}\log_2\left\{\widehat{L}_j\left(2G_{\infty}\big(\mathfrak{h}^{(k)}\big)\right)\right\}+36C_{N,R,m,k}=:2P\big(\mathfrak{h}^{(k)}\big);\\
M_{\sqrt{2}-1,q}(\mathrm{h}) &\le \big[144k\delta_*^{-2}+3(1+q)\big]\ln\left(2G_{\infty}\big(\mathfrak{h}^{(k)}\big)\underline{G}_{\mathbf{n}}^{-1}\right)\\
&\quad +144N\delta_*^{-2}\sum_{j=k+1}^{m}\log_2\left\{\widehat{L}_j\left(2G_{\infty}\big(\mathfrak{h}^{(k)}\big)\right)\right\}+72C_{N,R,m,k}=:2M_q\big(\mathfrak{h}^{(k)}\big).
\end{aligned}$$

We remark that P et M_q are independent of τ and y.

Put for any $z\ge 0$ and any $\mathrm{h}\in\mathrm{H}(\tau)$

$$\begin{aligned}
\check{\mathrm{V}}_{\tau}^{(z)}\big(\mathfrak{h}^{(k)}\big) &= 2\sqrt{2}\mathfrak{a}\sqrt{G_{\infty}\big(\mathfrak{h}^{(k)}\big)\big[P\big(\mathfrak{h}^{(k)}\big)+z\big]}+4\mathfrak{b}G_{\infty}\big(\mathfrak{h}^{(k)}\big)\Big[P\big(\mathfrak{h}^{(k)}\big)+z\Big];\\
\check{\mathrm{U}}_{\tau}^{(z,q)}\big(\mathfrak{h}^{(k)}\big) &= 2\sqrt{2}\mathfrak{a}\sqrt{G_{\infty}\big(\mathfrak{h}^{(k)}\big)\big[M_q\big(\mathfrak{h}^{(k)}\big)+z\big]}+4\mathfrak{b}G_{\infty}\big(\mathfrak{h}^{(k)}\big)\Big[M_q\big(\mathfrak{h}^{(k)}\big)+z\Big].
\end{aligned}$$

where remind $\mathfrak{a}=2\sqrt{\tau}(\mathbf{n_1})^{-1/2}\big(\sqrt{C_D}\vee[\chi\boldsymbol{c_b}]\big)$, $\mathfrak{b}=\frac{4\big(C_D\vee 8e\big)\ln^{\beta}(\mathbf{n_2})}{3\mathbf{n_1}}$.

We conclude that Proposition 2 is applicable with $\check{\mathrm{V}}_{\tau}^{(z)}$ and $\check{\mathrm{U}}_{\tau}^{(z,q)}$. Put for any $n\in\{\mathbf{n_1},\mathbf{n_1}+1,\ldots,\mathbf{n_2}\}$

$$\mathfrak{a}(n)=2\sqrt{2\tau}(n)^{-1/2}\big(\sqrt{C_D}\vee[\chi\boldsymbol{c_b}]\big),\ \ \mathfrak{b}(n)=\frac{8\big(C_D\vee 8e\big)\ln^{\beta}(2n)}{3n}$$

and define

$$\begin{aligned}
\mathrm{V}_{\tau}^{(z)}\big(n,\mathfrak{h}^{(k)}\big) &= 2\sqrt{2}\mathfrak{a}(n)\sqrt{G_{\infty}\big(\mathfrak{h}^{(k)}\big)\big[P\big(\mathfrak{h}^{(k)}\big)+z\big]}+4\mathfrak{b}(n)G_{\infty}\big(\mathfrak{h}^{(k)}\big)\Big[P\big(\mathfrak{h}^{(k)}\big)+z\Big];\\
\mathrm{U}_{\tau}^{(z,q)}\big(n,\mathfrak{h}^{(k)}\big) &= 2\sqrt{2}\mathfrak{a}(n)\sqrt{G_{\infty}\big(\mathfrak{h}^{(k)}\big)\big[M_q\big(\mathfrak{h}^{(k)}\big)+z\big]}+4\mathfrak{b}(n)G_{\infty}\big(\mathfrak{h}^{(k)}\big)\Big[M_q\big(\mathfrak{h}^{(k)}\big)+z\Big].
\end{aligned}$$

It is easily seen that $\mathfrak{a}(n) \geq \mathfrak{a}$, $\mathfrak{b}(n) \geq \mathfrak{b}$ for any $n \in \{\mathbf{n_1}, \ldots, \mathbf{n_2}\}$ since $\mathbf{n_2} \leq 2\mathbf{n_1}$. Therefore,

$$\mathrm{V}_{\tau}^{(z)}\bigl(n, \mathfrak{h}^{(k)}\bigr) \geq \check{\mathrm{V}}_{\tau}^{(z)}\bigl(\mathfrak{h}^{(k)}\bigr), \quad \mathrm{U}_{\tau}^{(z,q)}\bigl(n, \mathfrak{h}^{(k)}\bigr) \geq \check{\mathrm{U}}_{\tau}^{(z,q)}\bigl(\mathfrak{h}^{(k)}\bigr).$$

It remains to remind that $\xi_{\mathfrak{h}}(n) = \boldsymbol{\xi}\bigl(n/\mathbf{n_2}, \mathfrak{h}\bigr)$ for any $n \in \{\mathbf{n_1}, \mathbf{n_1}+1, \ldots, \mathbf{n_2}\}$ and any $\mathfrak{h} \in \boldsymbol{\mathfrak{H}}$. All saying above allows us to assert that Proposition, 2 is applicable to $|\xi_{\mathfrak{h}}(n)|$ on $\mathbf{H}(\tau) := \{\mathbf{n_1}, \mathbf{n_1}+1, \ldots, \mathbf{n_2}\} \times \mathfrak{H}(\mathbf{n}, \tau)$ for any $\tau > 0$ with $\mathrm{V}_{\tau}^{(z)}(\cdot,\cdot)$ and $\mathrm{U}_{\tau}^{(z,q)}(\cdot,\cdot)$.

Thus, putting $\mathbf{h} = (n, \mathfrak{h})$ we obtain for any $\tau > 0$, any $z \geq 1$ and any $q \geq 1$

$$\mathbb{P}_{\mathrm{f}}\left\{\sup_{\mathbf{h} \in \mathbf{H}(\tau)} \Bigl[|\xi_{\mathfrak{h}}(n)| - \mathrm{V}_{\tau}^{(z)}\bigl(n, \mathfrak{h}^{(k)}\bigr)\Bigr] \geq 0\right\} \leq 4\left[1 + \Bigl[\ln\bigl\{1 + 2^{-1}\ln 2\bigr\}\Bigr]^{-2}\right]^{2} \exp\{-z\}; \tag{6.42}$$

$$\mathbb{E}_{\mathrm{f}}\left\{\sup_{\mathbf{h} \in \mathbf{H}(\tau)} \Bigl[|\xi_{\mathfrak{h}}(n)| - \mathrm{U}_{\tau}^{(z,q)}\bigl(n, \mathfrak{h}^{(k)}\bigr)\Bigr]\right\}_{+}^{q} \leq 2^{(5q/2)+3} 3^{q+4} \Gamma(q+1) \ \bigl[\underline{A} \vee \underline{B}\bigr]^{q} \exp\{-z\}, \tag{6.43}$$

where, remind, $\underline{A} = \mathfrak{a}\sqrt{\underline{G}_{\mathbf{n}}[\tau]}$ and $\underline{B} = \mathfrak{b}\underline{G}_{\mathbf{n}}[\tau]$.

To get the statements of the theorem we will have to choose z. This, in its turn, will be done for V_{τ} and U_{τ} differently in dependence on the values of the parameter τ.

6^0. Let $\mathbf{r} \in \mathbb{N}$ be fixed and for any $r \in \mathbb{N}$ put $\tau_r = e^{r-\mathbf{r}}$. For any $r \in \mathbb{N}^*$ denote $\widehat{\mathfrak{H}}(r) = \mathfrak{H}\bigl(\mathbf{n}, \tau_r\bigr) \setminus \mathfrak{H}\bigl(\mathbf{n}, \tau_{r-1}\bigr)$, $\widehat{\mathfrak{H}}(0) = \mathfrak{H}\bigl(\mathbf{n}, \tau_0\bigr)$ and let $\widehat{\mathbf{H}}(r) := \{\mathbf{n_1}, \mathbf{n_1}+1, \ldots, \mathbf{n_2}\} \times \widehat{\mathfrak{H}}(r)$.

Probability bound For any $u \geq 1$ put $z_r(u) = u + 2\ln\bigl(1 + |r - \mathbf{r}|\bigr)$ and remark that

$$z_r(u) = \begin{cases} u + 2\ln\bigl(|\ln(\tau_{r-1})|\bigr), & r \leq \mathbf{r}; \\ u + 2\ln\bigl(1 + |\ln(\tau_r)|\bigr), & r \geq \mathbf{r}. \end{cases}$$

We have for any $r \in \mathbb{N}$ and any $\mathfrak{h} \in \widehat{\mathfrak{H}}(r)$

$$\begin{aligned} \tau_0 = F_{\mathbf{n_2},\mathbf{r}}(\mathfrak{h}) \quad &\Rightarrow \quad z_0(u) = u + 2\ln\Bigl\{1 + \bigl|\ln\bigl(F_{\mathbf{n_2},\mathbf{r}}(\mathfrak{h})\bigr)\bigr|\Bigr\}; \\ \tau_{r-1} \leq F_{\mathbf{n_2}}(\mathfrak{h}) = F_{\mathbf{n_2},\mathbf{r}}(\mathfrak{h}) \quad &\Rightarrow \quad z_r(u) \leq u + 2\ln\Bigl\{\bigl|\ln\bigl(F_{\mathbf{n_2},\mathbf{r}}(\mathfrak{h})\bigr)\bigr|\Bigr\}, \ \ 1 \leq r \leq \mathbf{r}-1; \\ \tau_r \geq F_{\mathbf{n_2}}(\mathfrak{h}) = F_{\mathbf{n_2},\mathbf{r}}(\mathfrak{h}) \quad &\Rightarrow \quad z_r(u) \leq u + 2\ln\Bigl\{1 + \bigl|\ln\bigl(F_{\mathbf{n_2},\mathbf{r}}(\mathfrak{h})\bigr)\bigr|\Bigr\}, \ \ r \geq \mathbf{r}. \end{aligned}$$

Hence, we have for any $r \in \mathbb{N}$

$$z_r(u) \leq u + 2\ln\Bigl\{1 + \bigl|\ln\bigl(F_{\mathbf{n_2},\mathbf{r}}(\mathfrak{h})\bigr)\bigr|\Bigr\}, \quad \forall \mathfrak{h} \in \widehat{\mathfrak{H}}(r), \tag{6.44}$$

that yields for any $r \in \mathbb{N}$

$$\mathrm{V}_{\tau_r}^{(z_r(u))}\bigl(n, \mathfrak{h}^{(k)}\bigr) \leq \mathcal{V}_{\mathbf{r}}^{(u)}(n, \mathfrak{h}), \ \ \forall (n, \mathfrak{h}) \in \widehat{\mathbf{H}}(r). \tag{6.45}$$

Here we have also taken into account that $\tau_r \leq eF_{\mathbf{n_2},\mathbf{r}}(\mathfrak{h})$, $\forall \mathfrak{h} \in \widehat{\mathfrak{H}}(r)$ for any $r \in \mathbb{N}$.

Thus, we get for any $r \in \mathbb{N}$ and $u \geq 0$, taking into account (6.45), the inclusion $\widehat{\mathbf{H}}(r) \subseteq \mathbf{H}(\tau_r)$ and applying (6.42) with $\tau = \tau_r$,

$$\mathbb{P}_{\mathrm{f}}\left\{\sup_{(n,\mathfrak{h}) \in \widehat{\mathbf{H}}(r)} \Bigl[|\xi_{\mathfrak{h}}(n)| - \mathcal{V}_{\mathbf{r}}^{(u)}(n, \mathfrak{h})\Bigr] \geq 0\right\} \leq \frac{4\left[1 + \Bigl[\ln\bigl\{1 + 2^{-1}\ln 2\bigr\}\Bigr]^{-2}\right]^{2} \exp\{-u\}}{[1 + |r - \mathbf{r}|]^{2}}. \tag{6.46}$$

Since obviously $\widetilde{\mathbf{N}} \times \mathfrak{H}(\mathbf{n}) = \cup_{r=0}^{\infty} \widehat{\mathbf{H}}(r)$, summing up the right hand side of (6.46) over r, we come to the first assertion of the theorem. Here we have also used that $16 \left[1 + \left[\ln\left\{1 + 2^{-1}\ln 2\right\}\right]^{-2}\right]^2 \leq 2419$ and the fact that $\widetilde{\mathfrak{H}}(n) \subseteq \mathfrak{H}(\mathbf{n})$ for any $n \in \widetilde{\mathbf{N}}$ in view of Assumption 5 and the definition of the number $\mathbf{n}$.

Moment's bound For any $u \geq 1$ put

$$z_r(u) = u + 2\ln\left(1 + |r - \mathbf{r}|\right) + q\ln\left(\underline{G}_{\mathbf{n}}[\tau_r]\underline{G}_{\mathbf{n}}^{-1}\right).$$

Similarly to (6.44) we have for any $r \in \mathbb{N}$ and any $\mathfrak{h} \in \widehat{\mathfrak{H}}(r)$

$$z_r(u) \leq u + 2\ln\left\{1 + \left|\ln\left(F_{\mathbf{n_2},r}(\mathfrak{h})\right)\right|\right\} + q\ln\left(\underline{G}_{\mathbf{n}}[\tau_r]\underline{G}_{\mathbf{n}}^{-1}\right).$$

Moreover, for any $r \in \mathbb{N}$ by definition

$$\underline{G}_{\mathbf{n}}[\tau_r] := \inf_{\mathfrak{h} \in \mathfrak{H}(\mathbf{n},\tau_r)} G_{\infty}\left(\mathfrak{h}^{(k)}\right)$$

and, therefore, for any $\mathfrak{h} \in \widehat{\mathfrak{H}}(r)$

$$z_r(u) \leq u + 2\ln\left\{1 + \left|\ln\left(F_{\mathbf{n_2},r}(\mathfrak{h})\right)\right|\right\} + q\ln\left\{G_{\infty}\left(\mathfrak{h}^{(k)}\right)\underline{G}_{\mathbf{n}}^{-1}\right\}.$$

Similarly to (6.45), it yields for any $r \in \mathbb{N}$

$$\mathrm{U}_{\tau_r}^{(z_r(u),q)}\left(n, \mathfrak{h}^{(k)}\right) \leq \mathcal{U}_{\mathbf{r}}^{(u,q)}(n,\mathfrak{h}), \ \ \forall (n,\mathfrak{h}) \in \widehat{\mathbf{H}}(r). \tag{6.47}$$

Note that for any $r \in \mathbb{N}$

$$\underline{A} \vee \underline{B} \leq 2C_{D,\boldsymbol{b}}\left[\sqrt{(\mathbf{n_1})^{-1}F_{\mathbf{n_2}}\underline{G}_{\mathbf{n}}} \vee \left((\mathbf{n_1})^{-1}\ln^{\beta}(\mathbf{n_2})\underline{G}_{\mathbf{n}}\right)\right]\left[\underline{G}_{\mathbf{n}}[\tau_r]\underline{G}_{\mathbf{n}}^{-1}\right],$$

where $C_{D,\boldsymbol{b}} = \left(\sqrt{2C_D} \vee [\gamma \boldsymbol{c_b}]\right) \vee \left[(2/3)\left(C_D \vee 8e\right)\right]$. We get from (6.43) and (6.47), similarly to (6.46),

$$\mathbb{E}_{\mathrm{f}}\left\{\sup_{(n,\mathfrak{h}) \in \widehat{\mathbf{H}}(r)}\left[|\xi_{\mathfrak{h}}(n)| - \mathcal{U}_{\mathbf{r}}^{(u,q)}(n,\mathfrak{h})\right]\right\}_{+}^{q} \leq \frac{K_q\left[\sqrt{(\mathbf{n_1})^{-1}F_{\mathbf{n_2}}\underline{G}_{\mathbf{n}}} \vee \left((\mathbf{n_1})^{-1}\ln^{\beta}(\mathbf{n_2})\underline{G}_{\mathbf{n}}\right)\right]^q e^{-u}}{[1 + |r - \mathbf{r}|]^2},$$

where $K_q = 2^{(7q/2)+3}3^{q+4}\Gamma(q+1)(C_{D,\boldsymbol{b}})^q$.

Summing up the right hand side of the last inequality over r we come to the second assertion of the theorem.

∎

6.2. Proof of Theorem 2

For any $l \in \mathbb{N}^*$ set $n_l = \mathbf{j}2^l$, $\mathbf{N}_l = \left\{n_l, n_l + 1, \ldots, n_{l+1}\right\}$ and let

$$\zeta_{\mathbf{j}} = \sup_{n \geq \mathbf{j}} \sup_{\mathfrak{h}^{(k)} \in \overline{\mathfrak{H}}_1^k(n,a)}\left[\frac{\sqrt{n}\ \eta_{\mathfrak{h}^{(k)}}(n)}{\sqrt{G_{\infty}\left(\mathfrak{h}^{(k)}\right)\ln\left(1 + \ln(n)\right)}}\right].$$

We obviously have

$$\begin{aligned}\mathbb{P}_{\mathrm{f}}\left\{\zeta_{\mathbf{j}}\geq\Upsilon\right\} &\leq \sum_{l=1}^{\infty}\mathbb{P}_{\mathrm{f}}\left\{\sup_{n\in\mathbf{N}_l}\sup_{\mathfrak{h}^{(k)}\in\overline{\mathfrak{H}}_1^k(n,a)}\left[\frac{\sqrt{n}\,\eta_{\mathfrak{h}^{(k)}}(n)}{\sqrt{G_{\infty}\left(\mathfrak{h}^{(k)}\right)\ln\left(1+\ln\left(n\right)\right)}}\right]\geq\Upsilon\right\}\\ &= \sum_{l=1}^{\infty}\mathbb{P}_{\mathrm{f}}\left\{\sup_{n\in\mathbf{N}_l}\sup_{\mathfrak{h}^{(k)}\in\overline{\mathfrak{H}}_1^k(n,a)}\left[\eta_{\mathfrak{h}^{(k)}}(n)-\Upsilon\sqrt{n^{-1}G_{\infty}\left(\mathfrak{h}^{(k)}\right)\ln\left(1+\ln\left(n\right)\right)}\right]>0\right\}.\end{aligned}$$

Let $l\in\mathbb{N}^*$ be fixed and later on $\boldsymbol{\Upsilon}_r$, $r=1,2,3$ denote the constants independent on l and n.

Note that in view of (3.11), (3.12) and (3.13) for any $n\in\mathbf{N}_l$

$$\begin{aligned}\mathcal{V}_0^{(2\ln(1+\ln(n_l)))}(n,\mathfrak{h}) &\leq \lambda_1\sqrt{\left(\mathbf{F}n^{-1}\right)G_{\infty}\left(\mathfrak{h}^{(k)}\left(P_n+2\ln\left\{1+\left|\ln\left(\mathbf{F}\right)\right|\right\}+2\ln\left(1+\ln\left(n\right)\right)\right)\right.}\\ &\quad+\lambda_2\left(n^{-1}\ln^{\boldsymbol{b}}\left(n\right)\right)G_{\infty}\left(\mathfrak{h}^{(k)}\right)\left(P_n+2\ln\left\{1+\left|\ln\left(\mathbf{F}\right)\right|\right\}+2\ln\left(1+\ln\left(n\right)\right)\right);\end{aligned}$$

where we have put

$$P_n=(36k\delta_*^{-2}+6)\ln\left(1+\mathfrak{b}\ln\left(2n\right)\right)+36N\delta_*^{-2}\mathfrak{a}\ln\left(1+\ln\left(2n^{\mathfrak{b}}\mathfrak{c}\right)\right)+18C_{N,R,m,k}(\boldsymbol{b}).$$

Hence, for any $n\in\mathbf{N}_l$ and any $\mathfrak{h}\in\widetilde{\mathfrak{H}}(n)$

$$\mathcal{V}_0^{(2\ln(1+\ln(n_l)))}(n,\mathfrak{h})\leq\Upsilon_1\sqrt{\frac{G_{\infty}\left(\mathfrak{h}^{(k)}\right)\ln\left(1+\ln\left(n\right)\right)}{n}}+\Upsilon_2\left[\frac{G_{\infty}\left(\mathfrak{h}^{(k)}\right)\ln^{\boldsymbol{b}}\left(n\right)\ln\left(1+\ln\left(n\right)\right)}{n}\right].$$

Since $\boldsymbol{b}>1$ can be arbitrary chosen and $a>2$ let $1<\boldsymbol{b}<a/2$. It yields for any $n\geq3$ and any $\mathfrak{h}^{(k)}\in\overline{\mathfrak{H}}_1^k(n,a)$

$$\frac{G_{\infty}\left(\mathfrak{h}^{(k)}\right)\ln^{\boldsymbol{b}}\left(n\right)\ln\left(1+\ln\left(n\right)\right)}{n}\leq\Upsilon_3\sqrt{\frac{G_{\infty}\left(\mathfrak{h}^{(k)}\right)\ln\left(1+\ln\left(n\right)\right)}{n}}$$

and, therefore, putting $\Upsilon=\Upsilon_1+\Upsilon_2\Upsilon_3$ we get for any $n\in\mathbf{N}_l$

$$\mathcal{V}_0^{(2\ln(1+\ln(n_l)))}(n,\mathfrak{h})\leq\Upsilon\sqrt{\frac{G_{\infty}\left(\mathfrak{h}^{(k)}\right)\ln\left(1+\ln\left(n\right)\right)}{n}}.$$

Noting that right hand side of the latter inequality is independent of $\mathfrak{h}_{(k)}$ and applying the first assertion of Theorem 1 with $\widetilde{\mathbf{N}}=\mathbf{N}_l$, $\mathbf{r}=0$ and $u=2\ln\left(1+\ln\left(n_l\right)\right)$ we have

$$\mathbb{P}_{\mathrm{f}}\left\{\zeta_{\mathbf{j}}\geq\Upsilon\right\}\leq2419\sum_{l=1}^{\infty}\left(l+\ln\left(\mathbf{j}\right)\right)^{-2}\leq\frac{2419}{\ln\left(\mathbf{j}\right)}.$$

∎

6.3. Proof of Theorem 3

1^0. We start the proof with establishing some simple facts used in the sequel.

For any $\mathbf{i} \in \mathbf{I}$ let $\mathfrak{n}(\mathbf{i}) \in \mathbb{N}^*$ and $\tilde{\pi}_j(\mathbf{i}) \in \mathbf{I},\ j = 1, \ldots, \mathfrak{n}(\mathbf{i})$, be the pairwise disjoint collection which is determined by the condition: $\mathrm{H}_{m,\mathbf{i}} \cap \mathrm{H}_{m,\mathbf{k}} = \emptyset, \quad \forall \mathbf{k} \notin \left\{\tilde{\pi}_1(\mathbf{i}), \ldots, \tilde{\pi}_{\mathfrak{n}(\mathbf{i})}(\mathbf{i})\right\}$. First we have

$$1 \le \mathfrak{n}(\mathbf{i}) \le \mathfrak{n}, \quad \forall \mathbf{i} \in \mathbf{I},$$

and we always put $\tilde{\pi}_{\mathfrak{n}(\mathbf{i})}(\mathbf{i}) = \mathbf{i}$. It yields, in particular, that we can construct another collection of indices $\pi(\mathbf{i}) := \left\{\pi_j(\mathbf{i}) \in \mathbf{I},\ j = \overline{1, \mathfrak{n}}\right\}$ given by

$$\pi_j(\mathbf{i}) = \begin{cases} \tilde{\pi}_j(\mathbf{i}), & 1 \le j \le \mathfrak{n}(\mathbf{i}); \\ \mathbf{i}, & \mathfrak{n}(\mathbf{i}) + 1 \le j \le \mathfrak{n}. \end{cases}$$

Note also that for any $1 \le j \le \mathfrak{n}$

$$\mathrm{card}\Big(\ \{\mathbf{i} \in \mathbf{I} :\ \ \pi_j(\mathbf{i}) = \mathbf{p}\}\ \Big) \le \mathfrak{n}, \quad \forall \mathbf{p} \in \mathbf{I}. \tag{6.48}$$

Indeed, if $\mathrm{card}\Big(\ \{\mathbf{i} \in \mathbf{I} :\ \ \pi_j(\mathbf{i}) = \mathbf{p}\}\ \Big) \ge \mathfrak{n} + 1$ for some $\mathbf{p} \in \mathbf{I}$, then

$$\mathrm{card}\Big(\ \{\mathbf{i} \in \mathbf{I} :\ \ \mathrm{H}_{m,\mathbf{p}} \cap \mathrm{H}_{m,\mathbf{i}} \ne \emptyset\}\ \Big) \ge \mathfrak{n} + 1,$$

that contradicts to the definition of a $\mathfrak{n}$-totally bounded cover. For any $\mathbf{i} \in \mathbf{I}$ define

$$\mathrm{H}_m(\mathbf{i}) = \bigcup_{\mathbf{k} \in \mathbf{I}:\ \mathrm{H}_{m,\mathbf{k}} \cap \mathrm{H}_{m,\mathbf{i}} \ne \emptyset} \ \bigcup_{\mathbf{j} \in \mathbf{I}:\ \mathrm{H}_{m,\mathbf{j}} \cap \mathrm{H}_{m,\mathbf{k}} \ne \emptyset} \mathrm{H}_{m,\mathbf{j}} = \bigcup_{l=1}^{\mathfrak{n}} \bigcup_{j=1}^{\mathfrak{n}} \mathrm{H}_{m,\pi_j\left(\pi_l(\mathbf{i})\right)}.$$

First we note that the definition of the set $\mathfrak{H}_m(\cdot)$ implies the following inclusion: for any $\mathbf{i} \in \mathbf{I}$

$$\mathfrak{H}_m(\mathfrak{h}_m) \subseteq \mathrm{H}_m(\mathbf{i}), \quad \forall \mathfrak{h}_m \in \mathrm{H}_{m,\mathbf{i}}. \tag{6.49}$$

Next, taking into account that $\sum_{\mathbf{q} \in \mathbf{I}} 1_{\mathrm{H}_{m,\mathbf{q}}}(\mathfrak{h}_m) \le \mathfrak{n}$ for any $\mathfrak{h}_m \in \mathfrak{H}_m$ in view of the definition of a $\mathfrak{n}$-totally bounded cover, we obtain in view of (6.48)

$$\begin{aligned} \sum_{\mathbf{i} \in \mathbf{I}} 1_{\mathrm{H}_m(\mathbf{i})}(\mathfrak{h}_m) &\le \sum_{\mathbf{i} \in \mathbf{I}} \sum_{j=1}^{\mathfrak{n}} \sum_{l=1}^{\mathfrak{n}} 1_{\mathrm{H}_{m,\pi_j\left(\pi_l(\mathbf{i})\right)}}(\mathfrak{h}_m) = \sum_{j=1}^{\mathfrak{n}} \sum_{l=1}^{\mathfrak{n}} \sum_{\mathbf{p} \in \mathbf{I}} \sum_{\mathbf{i}:\pi_l(\mathbf{i}) = \mathbf{p}} 1_{\mathrm{H}_{m,\pi_j(\mathbf{p})}}(\mathfrak{h}_m) \\ &\le \mathfrak{n} \sum_{j=1}^{\mathfrak{n}} \sum_{l=1}^{\mathfrak{n}} \sum_{\mathbf{p} \in \mathbf{I}} 1_{\mathrm{H}_{m,\pi_j(\mathbf{p})}}(\mathfrak{h}_m) = \sum_{j=1}^{\mathfrak{n}} \sum_{l=1}^{\mathfrak{n}} \sum_{\mathbf{q} \in \mathbf{I}} \sum_{\mathbf{p}:\pi_j(\mathbf{p}) = \mathbf{q}} 1_{\mathrm{H}_{m,\mathbf{q}}}(\mathfrak{h}_m) \\ &\le \mathfrak{n}^2 \sum_{j=1}^{\mathfrak{n}} \sum_{l=1}^{\mathfrak{n}} \sum_{\mathbf{q} \in \mathbf{I}} 1_{\mathrm{H}_{m,\mathbf{q}}}(\mathfrak{h}_m) \le \mathfrak{n}^5, \quad \forall \mathfrak{h}_m \in \mathfrak{H}_m. \end{aligned} \tag{6.50}$$

Define finally for any $\mathbf{i} \in \mathbf{I}$

$$\mathrm{f}_{\mathbf{i}} := {\mathbf{n_1}}^{-1} \sum_{i=1}^{\mathbf{n_2}} \int_{\mathrm{H}_m(\mathbf{i})} f_{1,i}(x) \nu_1\big(\mathrm{d}x\big)$$

and let $\mathbf{I_1} = \{\mathbf{i} \in \mathbf{I} : \ \mathrm{f_i} \geq (\mathbf{n_1})^{-v}\}$ and $\mathbf{I_2} = \mathbf{I} \setminus \mathbf{I_1}$.

2^0. Let us fix $\mathbf{i} \in \mathbf{I_1}$ and for any $n \geq 1$ define $\mathrm{H_i}(n) := \widetilde{\mathfrak{H}}_1^k(n) \times \mathfrak{H}_{k+1}^{m-1} \times \mathrm{H}_{m,\mathbf{i}}$, $\mathbf{i} \in \mathbf{I}$. The idea is to apply Theorem 1 to $\{\mathrm{H_i}(n),\ n \geq 1\}$ that is possible in view of Assumptions 7 (**i**) and 5. To do it we first note that the definition of $\mathbf{I_1}$ together with (6.49) implies for any $n \in \widetilde{\mathbf{N}}$

$$\mathfrak{L}_{n,v}\big(\mathfrak{h}_m\big) \geq \ln\big(1/\mathrm{f_i}\big), \ \ \forall \mathfrak{h}_m \in \mathrm{H}_{m,\mathbf{i}}.$$

It yields for any $n \in \widetilde{\mathbf{N}}$ and any $\mathfrak{h} \in \mathrm{H_i}(n)$

$$\widetilde{\mathcal{V}}_{\mathbf{r}}^{(v,z)}(n,\mathfrak{h}) \geq \mathcal{V}_{\mathbf{r}}^{(u)}(n,\mathfrak{h}), \qquad \widetilde{\mathcal{U}}_{\mathbf{r}}^{(v,z,q)}(n,\mathfrak{h}) \geq \mathcal{U}_{\mathbf{r}}^{(u,q)}(n,\mathfrak{h}),$$

where $u = \ln(1/\mathrm{f_i}) + z$. We deduce from Theorem 1 for any $\mathbf{i} \in \mathbf{I_1}$

$$\mathbb{P}_{\mathrm{f}}\left\{\sup_{n\in\widetilde{\mathbf{N}}}\sup_{\mathfrak{h}\in\mathrm{H_i}(n)}\Big[\big|\xi_{\mathfrak{h}}(n)\big| - \widetilde{\mathcal{V}}_{\mathbf{r}}^{(v,z)}(n,\mathfrak{h})\Big] \geq 0\right\} \leq 2419\ \mathrm{f_i}\ e^{-z}; \tag{6.51}$$

$$\mathbb{E}_{\mathrm{f}}\left\{\sup_{n\in\widetilde{\mathbf{N}}}\sup_{\mathfrak{h}\in\mathrm{H_i}(n)}\Big[\big|\xi_{\mathfrak{h}}(n)\big| - \widetilde{\mathcal{U}}_{\mathbf{r}}^{(v,z,q)}(n,\mathfrak{h})\Big]\right\}_+^q \leq \mathrm{f_i}\ \Lambda_q\big(\mathbf{n_1},\mathbf{n_2}\big)e^{-z}, \tag{6.52}$$

where we have put $\Lambda_q\big(\mathbf{n_1},\mathbf{n_2}\big) = c_q\left[\sqrt{(\mathbf{n_1})^{-1}F_{\mathbf{n_2}}\underline{G}_{\mathbf{n}}} \vee \big((\mathbf{n_1})^{-1}\ln^{\beta}(\mathbf{n_2})\underline{G}_{\mathbf{n}}\big)\right]^q$.

We have in view of (6.50), taking into account that $\mathbf{n_2} \leq 2\mathbf{n_1}$,

$$\sum_{\mathbf{i}\in\mathbf{I}}\mathrm{f_i} = (\mathbf{n_1})^{-1}\sum_{i=1}^{\mathbf{n_2}}\int f_{1,i}(x)\Big[\sum_{\mathbf{i}\in\mathbf{I}}1_{\mathrm{H}_m(\mathbf{i})}(x)\Big]\nu_1\big(\mathrm{d}x\big) \leq 2\mathfrak{n}^5. \tag{6.53}$$

Putting $\widetilde{\mathfrak{H}}^{(1)}(n) = \bigcup_{\mathbf{i}\in\mathbf{I_1}}\mathrm{H_i}(n)$, $n \geq 1$, we obtain from (6.51), (6.52) and (6.53)

$$\mathbb{P}_{\mathrm{f}}\left\{\sup_{n\in\widetilde{\mathbf{N}}}\sup_{\mathfrak{h}\in\widetilde{\mathfrak{H}}^{(1)}(n)}\Big[\big|\xi_{\mathfrak{h}}(n)\big| - \widetilde{\mathcal{V}}_{\mathbf{r}}^{(v,z)}(n,\mathfrak{h})\Big] \geq 0\right\} \leq 4838\ \mathfrak{n}^5\ e^{-z}; \tag{6.54}$$

$$\mathbb{E}_{\mathrm{f}}\left\{\sup_{n\in\widetilde{\mathbf{N}}}\sup_{\mathfrak{h}\in\widetilde{\mathfrak{H}}^{(1)}(n)}\Big[\big|\xi_{\mathfrak{h}}(n)\big| - \widetilde{\mathcal{U}}_{\mathbf{r}}^{(v,z,q)}(n,\mathfrak{h})\Big]\right\}_+^q \leq 2\Lambda_q\big(\mathbf{n_1},\mathbf{n_2}\big)\mathfrak{n}^5\ e^{-z}. \tag{6.55}$$

To get (6.55) we have used obvious equality: $[\sup_\alpha Q(\alpha)]_+^q = \sup_\alpha [Q(\alpha)]_+^q$.

3^0. Fix $\mathbf{i} \in \mathbf{I_2}$ and note that in view of Assumption 8 for any $n \geq 1$, any $\mathfrak{h} \in \mathrm{H_i}(n)$ and $i \geq 1$

$$\begin{aligned}\mathbb{E}_{\mathrm{f}}\big|G\big(\mathfrak{h},X_i\big)\big| &= \mathbb{E}_{\mathrm{f}}\Big\{\big|G\big(\mathfrak{h},X_i\big)\big|\,\mathbf{1}_{\mathfrak{H}_m\left(\mathfrak{h}_m\right)}(X_{1,i})\Big\} + \mathbb{E}_{\mathrm{f}}\Big\{\big|G\big(\mathfrak{h},X_i\big)\big|\,\mathbf{1}_{\mathfrak{H}_m\setminus\mathfrak{H}_m\left(\mathfrak{h}_m\right)}(X_{1,i})\Big\}\\ &\leq G_\infty\big(\mathfrak{h}^{(k)}\big)\Big[\mathbb{P}_{\mathrm{f}}\Big\{X_{1,i}\in\mathfrak{H}_m\big(\mathfrak{h}_m\big)\Big\} + n^{-1}\Big] \leq G_\infty\big(\mathfrak{h}^{(k)}\big)\Big[\mathbb{P}_{\mathrm{f}}\Big\{X_{1,i}\in\mathrm{H}_m(\mathbf{i})\Big\} + n^{-1}\Big].\end{aligned} \tag{6.56}$$

The last inequality follows from (6.49). It yields for any $n \in \widetilde{\mathbf{N}}$ and any $\mathfrak{h} \in \mathrm{H_i}(n)$

$$n^{-1}\sum_{i=1}^{n}\mathbb{E}_{\mathrm{f}}\big|G\big(\mathfrak{h},X_i\big)\big| \ \ \leq \ \ G_\infty\big(\mathfrak{h}^{(k)}\big)\big[\mathrm{f_i} + n^{-1}\big] \leq 2(\mathbf{n_1})^{-1}G_\infty\big(\mathfrak{h}^{(k)}\big), \tag{6.57}$$

since $\mathrm{f}_{\mathbf{i}} \le (\mathbf{n_1})^{-v}$ for any $\mathbf{i} \in \mathbf{I_2}$ and $v \ge 1$.

Introduce random events

$$\mathcal{C}_{\mathbf{i}} = \left\{ \sum_{i=1}^{\mathbf{n_2}} \mathbf{1}_{\mathrm{H}_m(\mathbf{i})}(X_{1,i}) \ge 2 \right\}, \quad \mathbf{i} \in \mathbf{I_2}, \qquad \mathcal{C} = \bigcup_{\mathbf{i} \in \mathbf{I_2}} \mathcal{C}_{\mathbf{i}}.$$

Note that if the random event $\bar{\mathcal{C}}$ holds (where, as usual, $\bar{\mathcal{C}}$ is complementary to $\mathcal{C}$) then for any $n \in \widetilde{\mathbf{N}}$ and any $\mathfrak{h} \in \mathrm{H}_{\mathbf{i}}(n)$ in view of Assumption 8 and (6.49)

$$n^{-1} \sum_{i=1}^{n} \left| G\big(\mathfrak{h}, X_i\big) \right| \ \le \ 2n^{-1} G_{\infty}\big(\mathfrak{h}^{(k)}\big) \le 2(\mathbf{n_1})^{-1} G_{\infty}\big(\mathfrak{h}^{(k)}\big). \tag{6.58}$$

Taking into account that bounds found in (6.57) and (6.58) are independent of $\mathbf{i}$ we get for any $n \in \widetilde{\mathbf{N}}$ and any $\mathfrak{h} \in \widetilde{\mathfrak{H}}^{(2)}(n) := \widetilde{\mathfrak{H}}(n) \setminus \in \widetilde{\mathfrak{H}}^{(1)}(n)$

$$\left| \xi_{\mathfrak{h}}(n) \right| \mathbf{1}_{\bar{\mathcal{C}}} \ \le \ 4(\mathbf{n_1})^{-1} G_{\infty}\big(\mathfrak{h}^{(k)}\big).$$

Noting that for any $\mathfrak{h} \in \boldsymbol{\mathfrak{H}}$, $z \ge 1$ and $n \in \widetilde{\mathbf{N}}$

$$\begin{aligned} \widetilde{\mathcal{V}}_{\mathbf{r}}^{(v,z)}(n, \mathfrak{h}) \ &> \ 8n^{-1} G_{\infty}\big(\mathfrak{h}^{(k)}\big) \ge 4(\mathbf{n_1})^{-1} G_{\infty}\big(\mathfrak{h}^{(k)}\big), \\ \widetilde{\mathcal{U}}_{\mathbf{r}}^{(v,z,q)}(n, \mathfrak{h}) \ &> \ 8n^{-1} G_{\infty}\big(\mathfrak{h}^{(k)}\big) \ge 4(\mathbf{n_1})^{-1} G_{\infty}\big(\mathfrak{h}^{(k)}\big). \end{aligned}$$

and, therefore, if the random event $\bar{\mathcal{C}}$ is realized we have

$$\sup_{n \in \widetilde{\mathbf{N}}} \sup_{\mathfrak{h} \in \widetilde{\mathfrak{H}}^{(2)}(n)} \left[\left| \xi_{\mathfrak{h}}(n) \right| - \widetilde{\mathcal{V}}_{\mathbf{r}}^{(v,z)}(n, \mathfrak{h}) \right] < 0, \quad \sup_{n \in \widetilde{\mathbf{N}}} \sup_{\mathfrak{h} \in \widetilde{\mathfrak{H}}^{(2)}(n)} \left[\left| \xi_{\mathfrak{h}}(n) \right| - \widetilde{\mathcal{U}}_{\mathbf{r}}^{(v,z,q)}(n, \mathfrak{h}) \right] < 0.$$

It yields, first,

$$\mathbb{P}_{\mathrm{f}} \left\{ \sup_{n \in \widetilde{\mathbf{N}}} \sup_{\mathfrak{h} \in \widetilde{\mathfrak{H}}^{(2)}(n)} \left[\left| \xi_{\mathfrak{h}}(n) \right| - \widetilde{\mathcal{V}}_{\mathbf{r}}^{(v,z)}(n, \mathfrak{h}) \right] \ge 0 \right\} \le \mathbb{P}_{\mathrm{f}} \left\{ \mathcal{C} \right\} \le \sum_{\mathbf{i} \in \mathbf{I_2}} \mathbb{P}_{\mathrm{f}} \left\{ \mathcal{C}_{\mathbf{i}} \right\}. \tag{6.59}$$

Next, taking into account the trivial bound $\left| \xi_{\mathfrak{h}}(n) \right| \le 2\overline{G}_{\mathbf{n}}$ for any $n \in \widetilde{\mathbf{N}}$ and any $\mathfrak{h} \in \mathfrak{H}(n)$, we get

$$\mathbb{E}_{\mathrm{f}} \left\{ \sup_{n \in \widetilde{\mathbf{N}}} \sup_{\mathfrak{h} \in \widetilde{\mathfrak{H}}^{(2)}(n)} \left[\left| \xi_{\mathfrak{h}}(n) \right| - \widetilde{\mathcal{U}}_{\mathbf{r}}^{(v,z,q)}(n, \mathfrak{h}) \right] \right\}_{+}^{q} \le \left(2\overline{G}_{\mathbf{n}} \right)^{q} \mathbb{P}_{\mathrm{f}} \left\{ \mathcal{C} \right\} \le \left(2\overline{G}_{\mathbf{n}} \right)^{q} \sum_{\mathbf{i} \in \mathbf{I_2}} \mathbb{P}_{\mathrm{f}} \left\{ \mathcal{C}_{\mathbf{i}} \right\}. \tag{6.60}$$

For any $\mathbf{i} \in \mathbf{I_2}$ put $\mathrm{p}_{i,\mathbf{i}} = \mathbb{P}_f \Big\{ X_{1,i} \in \mathrm{H}_m(\mathbf{i}) \Big\}$. Since $X_{1,i}$, $i \ge 1$, are independent random elements we have for any $\mathbf{i} \in \mathbf{I_2}$ and any $\lambda > 0$ in view of exponential Markov inequality

$$\mathbb{P}_{\mathrm{f}} \left\{ \mathcal{C}_{\mathbf{i}} \right\} \le \exp \left\{ -2\lambda + (e^{\lambda} - 1) \sum_{i=1}^{\mathbf{n_2}} \mathrm{p}_{i,\mathbf{i}} \right\} = \exp \Big\{ -2\lambda + \mathbf{n_1}(e^{\lambda} - 1) \mathrm{f}_{\mathbf{i}} \Big\}.$$

Minimizing the right hand side in λ we obtain for any $\mathbf{i} \in \mathbf{I_2}$

$$\mathbb{P}_{\mathrm{f}} \left\{ \mathcal{C}_{\mathbf{i}} \right\} \le (e/2)^2 (\mathbf{n_1} \mathrm{f}_{\mathbf{i}})^2 \le 2 \mathrm{f}_{\mathbf{i}} \ {\mathbf{n_1}}^{2-v}.$$

The last inequality follows from the definition of $\mathbf{I_2}$. We obtain finally in view of (6.53)

$$\sum_{\mathbf{i} \in \mathbf{I_2}} \mathbb{P}_{\mathrm{f}} \left\{ \mathcal{C}_{\mathbf{i}} \right\} \le 4 \mathbf{n}^5 \ {\mathbf{n_1}}^{2-v}. \tag{6.61}$$

The assertions of the theorem follow now from (6.54), (6.55), (6.59), (6.60) and (6.61).

∎

6.4. Proof of Corollary 2

To prove the assertion of the corollary it suffices to bound from above the function $\mathfrak{L}_{n,v}(\cdot)$. Remind that we proved, see (6.56), for any $n \geq 1$, any $\mathfrak{h} \in \mathfrak{H}(n)$ and $i \geq 1$

$$\mathbb{E}_{\mathrm{f}} \left| G\big(\mathfrak{h}, X_i\big) \right| \leq G_{\infty}\big(\mathfrak{h}^{(k)}\big) \left[\mathbb{P}_{\mathrm{f}}\Big\{ X_{1,i} \in \mathfrak{H}_m\big(\mathfrak{h}_m\big) \Big\} + n^{-1} \right].$$

It yields for any $n \in \widetilde{\mathbf{N}}$ and any $\mathfrak{h} \in \mathfrak{H}(n)$

$$F_{\mathbf{n_2}}(\mathfrak{h}) \leq G_{\infty}\big(\mathfrak{h}^{(k)}\big) \left[A_n(\mathfrak{h}_m) + n^{-1} \right], \quad A_n(\mathfrak{h}_m) = n^{-1} \sum_{i=1}^{n} \int_{\mathfrak{H}_m(\mathfrak{h}_m)} f_{1,i}(x) \nu_1\big(\mathrm{d}x\big). \tag{6.62}$$

Indeed, if $\mathbf{n_1} = \mathbf{n_2}$ then $n = \mathbf{n_2}$ and (6.62) is obvious. If $\mathbf{n_1} \neq \mathbf{n_2}$ then $\mathbb{P}_{\mathrm{f}}\Big\{ X_{1,i} \in \mathfrak{H}_m\big(\mathfrak{h}_m\big) \Big\}$ is independent of i since we supposed that $X_{1,i}$, $i \geq 1$ are identically distributed. Hence, $A_n(\cdot)$ is independent of n and (6.62) holds. Let $n \in \widetilde{\mathbf{N}}$ be fixed and let $\mathfrak{h} \in \mathfrak{H}(n)$ be such that $F_{\mathbf{n_2}}(\mathfrak{h}) \geq n^{-1/2}$.

If $A_n(\mathfrak{h}_m) \leq n^{-1}$ we have $G_{\infty}\big(\mathfrak{h}^{(k)}\big) \geq 2^{-1}\sqrt{n}$ and, therefore,

$$2v \Big| \ln\big\{ 2G_{\infty}\big(\mathfrak{h}^{(k)}\big) \big\} \Big| \geq v \ln(n) \geq \mathfrak{L}_{n,v}(\mathfrak{h}_m).$$

If $A_n(\mathfrak{h}_m) > n^{-1}$ we have $\widehat{F}_{\mathbf{n_2}}(\mathfrak{h}) = F_{\mathbf{n_2}}(\mathfrak{h}) \leq 2G_{\infty}\big(\mathfrak{h}^{(k)}\big) A_n(\mathfrak{h}_m)$ and, therefore,

$$\begin{aligned} \mathfrak{L}_{n,v}(\mathfrak{h}_m) \leq \ln\big(A_n^{-1}(\mathfrak{h}_m)\big) &\leq \ln\Big(2G_{\infty}\big(\mathfrak{h}^{(k)}\big)\widehat{F}^{-1}_{\mathbf{n_2}}(\mathfrak{h})\Big) = \Big|\ln\Big(2G_{\infty}\big(\mathfrak{h}^{(k)}\big)\widehat{F}^{-1}_{\mathbf{n_2}}(\mathfrak{h})\Big)\Big| \\ &\leq \Big|\ln\Big(2G_{\infty}\big(\mathfrak{h}^{(k)}\big)\Big)\Big| + \Big|\ln\Big(\widehat{F}_{\mathbf{n_2}}(\mathfrak{h})\Big)\Big|. \end{aligned}$$

Here we have also used that $A_n(\mathfrak{h}_m) \leq 1$. Thus, if $F_{\mathbf{n_2}}(\mathfrak{h}) \geq n^{-1/2}$ for any $v \geq 1$

$$\mathfrak{L}_{n,v}(\mathfrak{h}_m) \leq 2v \Big|\ln\Big(2G_{\infty}\big(\mathfrak{h}^{(k)}\big)\Big)\Big| + \Big|\ln\Big(\widehat{F}_{\mathbf{n_2}}(\mathfrak{h})\Big)\Big|. \tag{6.63}$$

If now $\mathfrak{h} \in \mathfrak{H}(n)$ be such that $F_{\mathbf{n_2}}(\mathfrak{h}) < n^{-1/2}$ then obviously $\widehat{F}_{\mathbf{n_2}}(\mathfrak{h}) < n^{-1/2}$ and, therefore,

$$2v \Big|\ln\Big(\widehat{F}_{\mathbf{n_2}}(\mathfrak{h})\Big)\Big| \geq v \ln(n) \geq \mathfrak{L}_{n,v}(\mathfrak{h}_m).$$

The latter inequality together with (6.63) yields for any $n \in \widetilde{\mathbf{N}}$, any $\mathfrak{h} \in \mathfrak{H}(n)$ and $v \geq 1$

$$\mathfrak{L}_{n,v}(\mathfrak{h}_m) \leq 2v \Big[\Big|\ln\Big(2G_{\infty}\big(\mathfrak{h}^{(k)}\big)\Big)\Big| + \Big|\ln\Big(\widehat{F}_{\mathbf{n_2}}(\mathfrak{h})\Big)\Big| \Big]. \tag{6.64}$$

Hence, choosing $\mathbf{r} = \ln(\mathbf{n_2})$ and replacing $\mathfrak{L}_{n,v}(\cdot)$ in the expressions of $\widetilde{\mathcal{V}}_{\mathbf{r}}^{(v,z)}(\cdot,\cdot)$ and $\widetilde{\mathcal{U}}_{\mathbf{r}}^{(v,z,q)}(\cdot,\cdot)$ by its upper bound found in (6.64) we come to the assertion of the corollary.

∎

6.5. Proof of Theorem 4

For any $l \in \mathbb{N}^*$ set $n_l = \mathbf{j}2^l$, $\mathbf{N}_l = \{n_l, n_l+1, \ldots, n_{l+1}\}$ and let

$$\zeta_{\mathbf{j}} = \sup_{n \geq \mathbf{j}} \sup_{\mathfrak{h}^{(k)} \in \overline{\mathfrak{H}}_1^k(n,a)} \frac{\sqrt{n}\, \eta_{\mathfrak{h}^{(k)}}(n)}{\sqrt{G_\infty\big(\mathfrak{h}^{(k)}\big)\Big[\ln\big\{G_\infty\big(\mathfrak{h}^{(k)}\big)\big\} \vee \ln\ln(n)\Big]}}.$$

We obviously have for any $y \geq 0$

$$\begin{aligned}
&\mathbb{P}_{\mathrm{f}}\{\zeta_{\mathbf{j}} \geq \boldsymbol{\Upsilon}\} \\
&\leq \sum_{l=1}^{\infty} \mathbb{P}_{\mathrm{f}}\left\{ \sup_{n \in \mathbf{N}_l} \sup_{\mathfrak{h}^{(k)} \in \overline{\mathfrak{H}}_1^k(n,a)} \left[\eta_{\mathfrak{h}^{(k)}}(n) - \boldsymbol{\Upsilon}\sqrt{n^{-1} G_\infty\big(\mathfrak{h}^{(k)}\big)\Big[\ln\big\{G_\infty\big(\mathfrak{h}^{(k)}\big)\big\} \vee \ln\ln(n)\Big]} \right] \geq 0 \right\}.
\end{aligned}$$

Remind, that for any $3 \leq \mathbf{n_1} \leq \mathbf{n_2} \leq 2\mathbf{n_1}$ and any $n \in \widetilde{\mathbf{N}}$

$$\begin{aligned}
\widehat{\mathcal{V}}^{(v,z)}(n,\mathfrak{h}) \;&=\; \lambda_1 \sqrt{\Big(\widehat{F}_{\mathbf{n_2}}(\mathfrak{h}) n^{-1}\Big) G_\infty\big(\mathfrak{h}^{(k)}\big)\Big(\widehat{P}_v(\mathfrak{h}^{(k)}) + 2(v+1)\big|\ln\big\{\widehat{F}_{\mathbf{n_2}}(\mathfrak{h})\big\}\big| + z\Big)} \\
&\quad + \lambda_2 \Big(n^{-1}\ln^{\beta}(n)\Big) G_\infty\big(\mathfrak{h}^{(k)}\big)\Big(\widehat{P}_v(\mathfrak{h}^{(k)}) + 2(v+1)\big|\ln\big\{\widehat{F}_{\mathbf{n_2}}(\mathfrak{h})\big\}\big| + z\Big);
\end{aligned}$$

Let $l \in \mathbb{N}^*$ be fixed and choose $v = 3$ and $z = 2\ln(1 + \ln(n_l))$. Later on $\boldsymbol{\Upsilon}_r$, $r = 1,2,3,4$ denote the constants independent on l and n.

We have in view of (3.11), (3.12) and (3.17) for any $n \in \mathbf{N}_l$ and $\mathfrak{h} \in \widetilde{\mathfrak{H}}(n)$

$$\widehat{\mathcal{V}}^{(3,\, 2\ln(1+\ln(n_l)))}(n,\mathfrak{h}) \leq \boldsymbol{\Upsilon}_1 \sqrt{\frac{G_\infty\big(\mathfrak{h}^{(k)}\big)\Big[\ln\big\{G_\infty\big(\mathfrak{h}^{(k)}\big)\big\} \vee \ln\ln(n)\Big]}{n}} + \boldsymbol{\Upsilon}_2 \left[\frac{G_\infty\big(\mathfrak{h}^{(k)}\big)\ln^{\boldsymbol{b}+1}(n)}{n}\right].$$

To get the latter inequality we have used, first, that

$$\widehat{F}_{\mathbf{n_2}}(\mathfrak{h})\big|\ln\big\{\widehat{F}_{\mathbf{n_2}}(\mathfrak{h})\big\}\big| \leq \sup_{x \in (0,\mathbf{F}]} x\big|\ln(x)\big| =: c(\mathbf{F}) < \infty, \quad \forall \mathbf{F} < \infty.$$

Next, to get the second term, we have used that for any $n \in \mathbf{N}_l$ and $\mathfrak{h} \in \widetilde{\mathfrak{H}}(n)$

$$\widehat{P}_3(\mathfrak{h}^{(k)}) \leq \boldsymbol{\Upsilon}_3 \ln(n), \qquad \big|\ln\big\{\widehat{F}_{\mathbf{n_2}}(\mathfrak{h})\big\}\big| \leq \max\big[\big|\ln\big\{\mathbf{F}\big\}\big|, \ln(n_{l+1})\big] \leq \max\big[\big|\ln\big\{\mathbf{F}\big\}\big|, \ln(2n)\big].$$

Since $\boldsymbol{b} > 1$ can be arbitrary chosen and $a > 4$ let $1 < \boldsymbol{b} < a/2 - 1$. It yields for any $n \geq 3$ and any $\mathfrak{h}^{(k)} \in \overline{\mathfrak{H}}_1^k(n,a)$

$$\frac{G_\infty\big(\mathfrak{h}^{(k)}\big)\ln^{\boldsymbol{b}+1}(n)}{n} \leq \boldsymbol{\Upsilon}_4 \sqrt{\frac{G_\infty\big(\mathfrak{h}^{(k)}\big)\Big[\ln\big\{G_\infty\big(\mathfrak{h}^{(k)}\big)\big\} \vee \ln\ln(n)\Big]}{n}}$$

and, therefore, putting $\boldsymbol{\Upsilon} = \boldsymbol{\Upsilon}_1 + \boldsymbol{\Upsilon}_2\boldsymbol{\Upsilon}_4$ we get for any $n \in \mathbf{N}_l$

$$\widehat{\mathcal{V}}^{(3,\, 2\ln(1+\ln(n_l)))}(n,\mathfrak{h}) \leq \boldsymbol{\Upsilon} \sqrt{\frac{G_\infty\big(\mathfrak{h}^{(k)}\big)\Big[\ln\big\{G_\infty\big(\mathfrak{h}^{(k)}\big)\big\} \vee \ln\ln(n)\Big]}{n}}.$$

Noting that right hand side of the latter inequality is independent of $\mathfrak{h}_{(k)}$ and applying the first assertion of Corollary 2 with $\widetilde{\mathbf{N}} = \mathbf{N}_l$ and $z = 2\ln(1+\ln(n_l))$ we obtain

$$\mathbb{P}_{\mathrm{f}}\left\{\zeta_{\mathbf{j}} \geq \Upsilon\right\} \leq 2\mathfrak{n}^5 \left\{ 2419 \sum_{l=1}^{\infty} (l+\ln(\mathbf{j}))^{-2} + \mathbf{j}^{-1} \sum_{l=1}^{\infty} 2^{-l} \right\} \leq 2\mathfrak{n}^5 \left\{ \frac{2419}{\ln(\mathbf{j})} + \mathbf{j}^{-1} \right\} \leq \frac{4840\mathfrak{n}^5}{\ln(\mathbf{j})}.$$

∎

7. Proof of Theorem 9

Below $c_1, c_2 \ldots$, denote the constants completely determined by d, p, μ and γ. We break the proof on several steps.

1^0. Let $\mathbb{B}_q$, $1 < q < \infty$, denote the set of functions vanishing outside $\mathbb{K}_\mu$ and those $\mathbb{L}_q$-norm is less or equal to 1. Later on the integration is always understood as the integration over $\mathbb{R}^d$.

Consider the set of functions

$$\Theta_h = \left\{ \theta : \mathbb{R}^d \to \mathbb{R} : \ \theta(x) = h^{-d} \int K_h(t-x)\, \ell(t) \mathrm{d}t, \ \ \ell \in \mathbb{B}_{\frac{p}{p-1}} \right\}, \ \ h \in \mathcal{H}.$$

Put also $\Theta = \cup_{h \in \mathcal{H}} \Theta_h$ and for any $\theta \in \Theta$ introduce the gaussian process

$$\chi_\theta = \int_{\mathbb{R}^d} \theta(x) b(\mathrm{d}x). \tag{7.1}$$

In view of Young inequality Folland (1999), Theorems 6.18 and 6.36, for any $\theta \in \Theta$

$$\|\theta\|_2 := \left(\int_{\mathbb{R}^d} \left|\theta(x)\right|^2 \mathrm{d}x \right)^{\frac{1}{2}} \leq h^{\frac{d(2-p)}{2p}} \|K\|_{\frac{2p}{p+2}} \leq h^{\frac{d(2-p)}{2p}} \leq \left(h^{(\min)}\right)^{\frac{d(2-p)}{2p}} < \infty, \tag{7.2}$$

since $2 \leq p < \infty$. Here we have also used that $\|K\|_q \leq \|K\|_\infty \leq 1$ in view of assumption imposed on the function K.

Thus, the stochastic integral in (7.1) is well-defined and χ_θ is zero-mean gaussian random function on Θ such that

$$V(\theta) = \|\theta\|_2, \quad \rho(\theta_1, \theta_2) = \|\theta_1 - \theta_2\|_2. \tag{7.3}$$

We equip Θ with the semi-metric ρ and in the next paragraph we compute the entropy of several subsets of Θ. This computation allows us, in particular, to assert that Dudley integral is finite on Θ. It yields Lifshits (1995) that $\chi_\bullet$ is P-a.s uniformly continuous on (Θ, ρ), therefore Assumption 2 holds. Moreover, we show that Assumption 3 is fulfilled.

We conclude that Proposition 3 are applicable to χ_θ, on Θ, with $\Theta_\alpha = \Theta_h$, $\mathfrak{A} = \mathcal{H}$, $\mathrm{a} = \sqrt{2}\rho$, $A = \sqrt{2}V$. It remains to note that in view of duality arguments for any $h \in \mathcal{H}$

$$\left\|\xi_h\right\|_p = \sup_{\ell \in \mathbb{B}_{\frac{p}{p-1}}} \int \ell(t) \xi_h(t) \mathrm{d}t,$$

and, therefore, for any $h \in \mathcal{H}$

$$\left\|\xi_h\right\|_p = \sup_{\theta \in \Theta_h} \chi_\theta. \tag{7.4}$$

2^0. In order to apply Proposition 3 we need to compute several quantities. First, we have to choose the function τ_1 (since $B, \mathrm{b} \equiv 0$, hence $\tau_2 \equiv 0$). Set $\tau_1(h) = h^{-1}$ and note that for any $u > 0$

$$\Theta_1'(u) := \bigcup_{h:\, \tau_1(h) \le u} \Theta_h = \bigcup_{h \ge u^{-1}} \Theta_h,$$

We note that computations given in (7.2) yield

$$g_A^*(u) := \sup_{\theta \in \Theta_1'(u)} A(\theta) \le \sqrt{2} u^{\frac{d(p-2)}{2p}} =: g_A(u), \;\; u \ge \left(h^{(\max)}\right)^{-1}.$$

3^0. Let $\mathfrak{E}^{(u)}(\delta),\ \delta > 0$, be the entropy of $\Theta_1'(u)$ computed with respect to semi-metric $\mathrm{a} = \sqrt{2}\rho$, where, remind $\rho(\cdot) = \|\cdot\|_2$. The following assertion is true: there exist c_0 completely determined by γ, d, p and μ such that for any $\forall \bar{\gamma} \in \left(d/2 - d/p, \gamma\right]$ and for any $u \ge \left(h^{(\max)}\right)^{-1}$

$$\mathfrak{E}^{(u)}(\delta) \le c_1 u^d\, \delta^{-d/\bar{\gamma}}, \;\; \forall\, \delta > 0. \tag{7.5}$$

where $c_1 = c_0 3^{4d} d^2$.

The proof of (7.5) is obtained by routine computations and it is postponed to the step 8^0.

4^0. Choosing $u = \left(h^{(\min)}\right)^{-1}$ (that yields $\Theta_1'(u) = \Theta$) and $\bar{\gamma} = \gamma$ we get for any $\delta > 0$

$$\mathfrak{E}_{\Theta,\mathrm{a}}(\delta) \le c_1 \left(h^{(\min)}\right)^{-d} \left(\frac{1}{\delta}\right)^{d/\gamma}.$$

In view the condition $\gamma > d/2$, Dudley integral is finite on Θ and, therefore, $\chi_\bullet$ is P-a.s uniformly continuous on (Θ, ρ). It complete the verification of Assumption 2.

The last inequality shows also that there exist $\tau > 0$ such that for any $s \in \mathbb{S}$, satisfying $\limsup_{a\to\infty} a^\tau s^{-1}(a) < \infty$, Assumption 3 is fulfilled.

5^0. Let us now choose the function $\mathfrak{s}$. Set $\delta(u) = u^{d/2-d/p}$ and for any $u \ge 1$ let $m(u) \in \mathbb{N}$ be such that $2^{m(u)} \le \delta(u) < 2^{m(u)+1}$. Define

$$\mathfrak{s}(u, \delta) = \left(3/4\pi^2\right) \left(1 + \left[\log_2\left(2^{-m(u)}\delta\right)\right]^2\right)^{-1}.$$

We remark that

$$\mathfrak{s}\left(u, 2^{k/2}\right) = (3/4\pi^2) \left[1 + \left((k/2) - m(u)\right)^2\right]^{-1}$$

and, therefore, $\mathfrak{s}(u, \cdot) \in \mathbb{S}$ for any $u \ge 1$. Moreover, we note that if $p = 2$ then $\mathfrak{s}$ does not depend on u and it is given by

$$\mathfrak{s}(\delta) = \left(3/4\pi^2\right) \left(1 + \left[\log_2 \delta\right]^2\right)^{-1}.$$

Obviously the factor $3/4\pi^2$ can be replaced here by $6/\pi^2$.

Now, let us compute the quantity λ_1 related to the function $\mathfrak{s}$. Remind that

$$\lambda_1 := \sup_{t \in \left[1, \sqrt{2}\right]} \sup_{x > \underline{\tau}_1} \sup_{\delta > 0} \frac{\mathfrak{s}(xt, \delta)}{\mathfrak{s}(x, \delta)},$$

where $\underline{\tau}_1 = \big(h^{(\max)}\big)^{-1} \geq 1$. It is evident that

$$\lambda_1 = \sup_{m\geq 0}\left\{1 \bigvee \sup_{\delta>0}\left[\frac{1+\Big(\log_2\big(2^{-m}\delta\big)\Big)^2}{1+\Big(\log_2\big(2^{-m-1}\delta\big)\Big)^2}\right]\right\} = 1 \bigvee \sup_{x>0}\left[\frac{1+\big(\log_2(x)\big)^2}{1+\big(\log_2(x/2)\big)^2}\right]$$

$$\leq 1+\sup_{x>0}\left|\frac{2\log_2(x)-1}{1+\big(\log_2(x)-1\big)^2}\right| = 1+\sup_{y\in\mathbb{R}}\left|\frac{2y+1}{1+y^2}\right| < 3.$$

6^0. Define

$$\bar{\gamma}(\delta) = \begin{cases} d/2-d/(2p), & 0<\delta<\delta(u); \\ \gamma, & \delta\geq\delta(u), \end{cases}$$

and note that $(d/2-d/p) < \gamma(\delta) \leq \gamma$ for any $\delta > 0$.

Putting $c_2 = c_1 4(144)^4$ we get for any $\delta > 0$ in view of (7.5)

$$\begin{aligned}\mathfrak{E}^{(u)}\Big(\lambda_1^{-1}\big[g_A(u)\mathfrak{s}(u,\delta)\big]/48\delta\Big) &\leq c_2 u^d\left[u^{\frac{d(p-2)}{2p}}\mathfrak{s}(u,\delta)\right]^{-d/\bar{\gamma}(\delta)}\delta^{d/\bar{\gamma}(\delta)} \\ &= c_2 u^{\frac{2d}{p}-\left(\frac{d(p-2)}{2p}\right)\left(\frac{d}{\bar{\gamma}(\delta)}-2\right)}\left[\mathfrak{s}(u,\delta)\right]^{-d/\bar{\gamma}(\delta)}\delta^{d/\bar{\gamma}(\delta)} \\ &= c_2 u^{\frac{2d}{p}}\big[\delta(u)\big]^{2-\frac{d}{\bar{\gamma}(\delta)}}\left[\mathfrak{s}(u,\delta)\right]^{-4}\delta^{d/\bar{\gamma}(\delta)}. \end{aligned} \tag{7.6}$$

To get the last inequality we have taken into account that $\mathfrak{s}(u,\delta) < 1$ for any $u \geq 1, \delta > 0$ and that $d/\bar{\gamma}(\delta) \leq 4$.

We obtain from (7.6) for any $\delta > 0$, putting $c_3 = (4\pi^2/3)^4 c_2$,

$$\begin{aligned}&\delta^{-2}\mathfrak{E}^{(u)}\Big(\lambda_1^{-1}\big[g_A(u)\mathfrak{s}(u,\delta)\big]/48\delta\Big) \leq c_2 u^{\frac{2d}{p}}\left(\frac{\delta}{\delta(u)}\right)^{d/\bar{\gamma}(\delta)-2}\left[\mathfrak{s}(u,\delta)\right]^{-4} \\ &\leq c_2\, u^{\frac{2d}{p}}\left[\Big(\delta\Big/\delta(u)\Big)^{2/(p-1)}1_{(0,\delta(u))}(\delta)+\big(\delta(u)/\delta\big)^{2-d/\gamma}1_{[\delta(u),\infty)}(\delta)\right]\left[\mathfrak{s}(u,\delta)\right]^{-4} \\ &\leq c_2\, u^{\frac{2d}{p}}\left[\Big(2^{-m(u)}\delta\Big)^{2/(p-1)}1_{(1,\delta(u))}(\delta)+4\Big(2^{m(u)}/\delta\Big)^{2-d/\gamma}1_{[\delta(u),\infty)}(\delta)\right]\left[\mathfrak{s}(u,\delta)\right]^{-4} \\ &= 5c_3\, u^{\frac{2d}{p}}\,2^{-\alpha\left|\log_2\left(2^{-m(u)}\delta\right)\right|}\left(1+\Big[\log_2\big(2^{-m(u)}\delta\big)\Big]^2\right)^4, \end{aligned} \tag{7.7}$$

where $\alpha = \min\big\{2/(p-1), 2-d/\gamma\big\}$. We obtain from (7.7)

$$\mathcal{E}'(u) := \sup_{\delta>0}\Big[\delta^{-2}\mathfrak{E}^{(u)}\Big(\lambda_1^{-1}\big[g_A(u)\mathfrak{s}(u,\delta)\big]/48\delta\Big)\Big] \leq c_4\, u^{\frac{2d}{p}}, \tag{7.8}$$

where $c_4 = 5c_3 \sup_{x\geq 0}\Big[2^{-\alpha x}\big(1+x^2\big)^4\Big]$.

7^0. Remind that $\tau_1(h) = h^{-1}$ and, in particular, $\overline{\tau}_1 = \big[h^{(\min)}\big]^{-1}$. Choosing $\varepsilon = \sqrt{2}-1$ we get from (7.8) putting $c_5 = c_4 2^{d/p}$

$$\widehat{\mathcal{E}}^{\left(\sqrt{2}-1\right)}(h) := \mathcal{E}'\Big(\sqrt{2}\tau_1(h)\Big) \leq c_5 h^{-\frac{2d}{p}}.$$

Choose also $R_r(t) = R(t) = t^{\frac{2d}{p}}$ (independent of r) that yields

$$\widehat{R}^{\left(\sqrt{2}-1\right)}(h) := R\Big(\sqrt{2}\tau_1(h)\Big) = 2^{d/p}h^{-\frac{2d}{p}}.$$

Choosing finally $z = 0$ and putting $\widehat{\mathrm{U}}(h) = \widehat{\mathrm{U}}^{(z,\varepsilon,r)}(h), \varepsilon = \sqrt{2}-1, z = 1$, we obtain that

$$\widehat{\mathrm{U}}(h) \le \sqrt{2}g_A\Big(2\tau_1(h)\Big)\sqrt{32c_5h^{-\frac{2d}{p}} + 2^{d/p}h^{-\frac{2d}{p}} + 1} \le c_6 h^{\frac{d(2-p)}{2p}-\frac{d}{p}} = c_6 h^{-\frac{d}{2}}.$$

Let us compute now the quantities $\mathcal{R}^{(\varepsilon,0)}, \mathcal{R}^{(\varepsilon,q)}$ defined in (2.10) with $\varepsilon = \sqrt{2}-1$.

Noting that $\underline{\tau}_1 = \left(h^{(\max)}\right)^{-1}$ and $2^{J/2}\overline{\tau}_1 \le 2^{-1/2}\underline{\tau}_1$ we get

$$\begin{aligned}
\mathcal{R}^{(\varepsilon,0)} &:= \sum_{j=0}^{J}\exp\left\{-\Big(\overline{\tau}_1 2^{-(j/2)}\Big)^{\frac{2d}{p}}\right\} \le c_8\exp\left\{-2^{-3/2}\left(h^{(\max)}\right)^{-2d/p}\right\};\\
\mathcal{R}^{(\varepsilon,q)} &:= \sum_{j=0}^{J}\Big[g_A\big(\overline{\tau}_1 2^{-j/2}\big)\Big]^q\exp\left\{-\Big(\overline{\tau}_1 2^{-(j/2)}\Big)^{\frac{2d}{p}}\right\}\\
&\le c_9\left(h^{(\max)}\right)^{\frac{qd(2-p)}{2p}}\exp\left\{-2^{-3/2}\left(h^{(\max)}\right)^{-2d/p}\right\}.
\end{aligned}$$

The assertions of the theorem follow now from Proposition 3.

8^0. It remains to prove (7.5). The proof is based on the following inclusion: for any $\bar{\gamma} \in (0, \gamma]$

$$\Theta_1'(u) \subset \mathbb{H}^*_{\frac{p}{p-1}}\left(\bar{\gamma}, 3^d\sqrt{d}u^{\bar{\gamma}}\right), \ \ \forall u > 0, \tag{7.9}$$

where $\mathbb{H}^*_q(\cdot,\cdot) \subset \mathbb{H}_q(\cdot,\cdot)$ consists of functions vanishing outside of $\mathbb{K}_{2\mu}$.

Let $\mathfrak{E}^*(\cdot)$ be the entropy of $\mathbb{H}^*_{\frac{p}{p-1}}(\bar{\gamma}, L), L > 0$, measured in $\|\cdot\|_2$. It is well-known Edmunds and Triebel (1996), that for any $p > 1$ there exist c_0 completely determined by γ, d, p and μ such that for any $(d/2 - d/p) < \bar{\gamma} \le \gamma$ and for any $L > 0$

$$\mathfrak{E}^*(\delta) \le c_0\left(L\delta^{-1}\right)^{d/\bar{\gamma}}, \ \ \forall\, \delta > 0.$$

Since $\bar{\gamma}$ takes only two values $d/2 - d/2p$ and γ, (7.5) follows immediately from (7.9). Thus, we shall prove (7.9).

Fix $\theta \in \Theta_1'(u)$. By its definition there exists $\ell \in \mathbb{B}_{\frac{p}{p-1}}$ and $h \ge u^{-1}$ such that $\theta = K_h * \ell$, where " $*$ " stands convolution operator on $\mathbb{R}^d$. First, we note that all functions belonging to Θ vanish outside the cube $\mathbb{K}_{2\mu}$ in view of assumption imposed on function K.

Next, for any $m = (m_1, \ldots, m_d) \in \mathbb{N}^d$ put $|m| = m_1 + \cdots + m_d$, and set $\gamma = l + \alpha$ and $\bar{\gamma} = \bar{l} + \bar{\alpha}$, where $l, \bar{l} \in \mathbb{N}$ and $0 < \alpha, \bar{\alpha} \le 1$.

Then, since $K \in \mathbb{H}_\infty(\gamma, 1)$ we have for any $m \in \mathbb{N}^d$ such that $|m| \le l$

$$\sup_{x\in\mathbb{R}^d}\left|\frac{\partial^{|m|}K_h(x)}{\partial^{m_1}x_1\cdots\partial^{m_d}x_d}\right| \le h^{-|m|} \le \left(h^{(\min)}\right)^{-\gamma} < \infty.$$

Above remarks allow us to assert that all partial derivatives $\theta^{(m)}$ exist whenever $|m| \le l$ and they are given by

$$\theta^{(m)}(x) = \int (K_h)^{(m)}(t-x)\ell(t)\mathrm{d}t, \ \ \forall x \in \mathbb{R}^d,$$

where for any function g the notation $g^{(m)}$ or (if it is more convenient) $(g)^{(m)}$ is used for its partial derivative.

We obtain in view of Young inequality for any $\Delta \in \mathbb{R}$ and any $m \in \mathbb{N}^d$ satisfying $|m| = \bar{l}$

$$\begin{aligned}\left\|\theta^{(m)}(\cdot+\Delta)-\theta^{(m)}(\cdot)\right\|_{\frac{p}{p-1}} &\leq \left\|(K_h)^{(m)}(\cdot+\Delta)-(K_h)^{(m)}(\cdot)\right\|_1 \\ &\leq h^{-\bar{l}}\left\|K^{(m)}(\cdot+[\Delta/h])-K^{(m)}(\cdot)\right\|_1.\end{aligned}$$

Here we have used that $\ell \in \mathbb{B}_{\frac{p}{p-1}}$.

We remark that if $h \leq |\Delta|$ then for any $u \in \mathbb{R}^d$ either $K^{(m)}(u+[\Delta/h]) = 0$ or $K^{(m)}(u) = 0$ in view of the assumption imposed on the support of K. Thus, if $h \leq |\Delta|$

$$\left\|K^{(m)}(\cdot+[\Delta/h])-K^{(m)}(\cdot)\right\|_1 \leq \left\|K^{(m)}\right\|_1 \leq \left\|K^{(m)}\right\|_1 \big(\Delta/h\big)^{\bar{\alpha}} \leq \big(\Delta/h\big)^{\bar{\alpha}},$$

since $h \leq |\Delta|$ and $\left\|K^{(m)}\right\|_1 \leq \left\|K^{(m)}\right\|_\infty \leq 1$ in view of assumption imposed on the function K.

If $h > |\Delta|$ then in view of the assumption imposed on the support of K we have

$$\left\|K^{(m)}(\cdot+[\Delta/h])-K^{(m)}(\cdot)\right\|_1 = \int_{\left[-\frac{3}{2},\frac{3}{2}\right]^d}\left|K^{(m)}(u+[\Delta/h])-K^{(m)}(u)\right|\mathrm{d}u$$

$$\leq 3^d\left|\sum_{i=1}^d(\Delta/h)^2\right|^{\bar{\alpha}/2} \leq 3^d\sqrt{d}\big(|\Delta|/h\big)^{\bar{\alpha}}.$$

Since $h \geq u^{-1}$ we conclude finally that $\forall \Delta \in \mathbb{R}$

$$\left\|\theta^{(m)}(\cdot+\Delta)-\theta^{(m)}(\cdot)\right\|_{\frac{p}{p-1}} \leq 3^d\sqrt{d}h^{-\bar{\gamma}}|\Delta|^{\bar{\alpha}} \leq 3^d\sqrt{d}u^{\bar{\gamma}}|\Delta|^{\bar{\alpha}}. \tag{7.10}$$

It means that $\theta \in \mathbb{H}_{\frac{p}{p-1}}\left(\bar{\gamma}, 3^d\sqrt{d}u^{\bar{\gamma}}\right)$. As it was mentioned above all function belonging to Θ vanish outside the cube $\mathbb{K}_{2\mu}$ that allows us to conclude that $\theta \in \mathbb{H}^*_{\frac{p}{p-1}}\left(\bar{\gamma}, 3^d\sqrt{d}u^{\bar{\gamma}}\right)$ and, therefore, (7.9) is proved.

∎

8. Appendix

Proof of Lemma 1 1^0. We start the proof with the following simple fact. Let $\widetilde{\mathfrak{T}}$ be an arbitrary subset of $\mathfrak{T}$. Then

$$\sup_{\mathfrak{t}\in\widetilde{\mathfrak{T}}}\zeta(\mathfrak{t},\cdot) \ \ is \ \ \mathfrak{B}-measurable. \tag{8.1}$$

Indeed, since $\mathfrak{T}$ is totally bounded $\widetilde{\mathfrak{T}}$ is totally bounded as well. Denote by $\widehat{\mathfrak{T}}$ the union of 2^{-l}-nets, $l \geq 0$, in $\widetilde{\mathfrak{T}}$. Let $\mathbf{\Omega}_0 = \big\{\omega \in \mathbf{\Omega}: \quad \zeta(\cdot,\omega) \text{ is continuous}\big\}$ and let $\overline{\mathbf{\Omega}}_0$ be the complementary to $\mathbf{\Omega}_0$. We have for any $x \in \mathbb{R}$

$$\left\{\omega \in \mathbf{\Omega}: \ \sup_{\mathfrak{t}\in\widetilde{\mathfrak{T}}}\zeta(\mathfrak{t},\omega) \leq x\right\} \cap \mathbf{\Omega}_0 = \left\{\omega \in \mathbf{\Omega}: \ \sup_{\mathfrak{t}\in\widehat{\mathfrak{T}}}\zeta(\mathfrak{t},\omega) \leq x\right\} \cap \mathbf{\Omega}_0 \in \mathfrak{B}$$

since $\widehat{\mathfrak{T}}$ is countable dense subset of $\widetilde{\mathfrak{T}}$. It remain to note that $\left\{\omega\in\mathbf{\Omega}:\ \sup_{\mathfrak{t}\in\widetilde{\mathfrak{T}}}\zeta(\mathfrak{t},\omega)\le x\right\}\cap\overline{\mathbf{\Omega}}_0\in\mathfrak{B}$ since $\mathbf{P}\big(\overline{\mathbf{\Omega}}_0\big)=0$ and the considered probability space is complete.

2^0. Set $\mathfrak{Z}(n,k)=\left\{\mathfrak{z}\in\mathfrak{Z}:\ \ g(\mathfrak{z})\in[k/n,(k+1)/n]\right\}$, $n\in\mathbb{N}^*$, $k\in\mathbb{Z}$, and let $\mathbf{K}(n)\subseteq\mathbb{Z}$, $n\in\mathbb{N}^*$, be defined from the relation if $k\in\mathbf{K}(n)\ \Leftrightarrow\ \mathfrak{Z}(n,k)\neq\emptyset$. Put also $\mathfrak{T}(k,n)=\cup_{\mathfrak{z}\in\mathfrak{Z}(k,n)}\mathfrak{T}_{\mathfrak{z}}$ and define

$$\xi_{k,n}(\omega)=\sup_{\mathfrak{t}\in\mathfrak{T}(k,n)}\zeta(\mathfrak{t},\omega)-(k+1)/n,\qquad \xi_n(\omega)=\sup_{k\in\mathbf{K}(n)}\xi_{k,n}(\omega).$$
$$\eta_{k,n}(\omega)=\sup_{\mathfrak{z}\in\mathfrak{Z}(k,n)}\Big[\sup_{\mathfrak{t}\in\mathfrak{T}_{\mathfrak{z}}}\zeta(\mathfrak{t},\omega)-g(\mathfrak{z})\Big],\qquad \eta(\omega)=\sup_{\mathfrak{z}\in\mathfrak{Z}}\Big[\sup_{\mathfrak{t}\in\mathfrak{T}_{\mathfrak{z}}}\zeta(\mathfrak{t},\omega)-g(\mathfrak{z})\Big].$$

Some remarks are in order. First, the definition of $\mathfrak{Z}(k,n)$ implies that for any $k\in\mathbf{K}(n)$, $n\in\mathbb{N}^*$

$$\xi_{k,n}(\cdot)\le\eta_{k,n}(\cdot)\le\xi_{k,n}(\cdot)+n^{-1}. \tag{8.2}$$

Next, taking into account that $\mathfrak{Z}=\cup_{k\in\mathbf{K}(n)}\mathfrak{Z}(k,n)$ for any $n\in\mathbb{N}^*$ we have

$$\eta(\cdot)=\sup_{k\in\mathbf{K}(n)}\eta_{k,n}(\cdot),\quad\forall n\in\mathbb{N}^*. \tag{8.3}$$

We obtain from (8.2) and (8.3) that for any $n\in\mathbb{N}^*$

$$0\le\eta(\cdot)-\xi_n(\cdot)\le\sup_{k\in\mathbf{K}(n)}[\eta_{k,n}(\cdot)-\xi_{k,n}(\cdot)]\le n^{-1},$$

and, therefore, $\eta(\cdot)=\lim_{n\to\infty}\xi_n(\cdot)$. It remains to note that $\xi_{k,n}(\cdot)$ are $\mathfrak{B}$-measurable for any $k\in\mathbf{K}(n)$, $n\in\mathbb{N}^*$ in view of (8.1), that implies obviously that $\xi_n(\cdot)$ is $\mathfrak{B}$-measurable for any $n\in\mathbb{N}^*$. Thus, $\eta(\cdot)$ is $\mathfrak{B}$-measurable as a pointwise limit of $\mathfrak{B}$-measurable functions.

∎

Proof of Lemma 2 1^0. Remind that

$$G_\infty(r)=V_r^{-1}\|g\|_\infty\|K\|_\infty,\qquad \underline{G}_n=V^{-1}_{r^{(\max)}(n)}\|g\|_\infty\|K\|_\infty,\ \ n\ge1. \tag{8.4}$$

Hence we have have for any $l=\overline{1,d}$ and any $\mathfrak{h}_l:=r_l\in\Big[r_l^{(\min)}(n),r_l^{(\max)}(n)\Big]$

$$G_{l,n}(r_l)=\|g\|_\infty\|K\|_\infty\left[V_{r^{(\min)}(n)}\right]^{-1}\left[r_l^{(\min)}(n)\Big/r_l\right]^{\gamma_l};$$
$$\underline{G}_{l,n}=\|g\|_\infty\|K\|_\infty\left[V_{r^{(\min)}(n)}\right]^{-1}\left[r_l^{(\min)}(n)\Big/r_l^{(\max)}(n)\right]^{\gamma_l}.$$

Thus, we get for any $n\ge1$, $r\in\mathcal{R}(n)$ and for any $j=\overline{1,d}$

$$\frac{G_\infty(r)}{\underline{G}_n}=\prod_{l=1}^d\left[\frac{r_l^{(\max)}}{r_l}\right]^{\gamma_l}\ge\left[\frac{r_j^{(\max)}}{r_j}\right]^{\gamma_j}=\frac{G_{j,n}(r_j)}{\underline{G}_{j,n}}.$$

We conclude that Assumption 4 (**i**) is fulfilled.

2^0. Remind that for any $r, r' \in \mathcal{R}(n)$

$$\varrho_n^{(d)}\Big(r, r'\Big) := \max_{l=\overline{1,d}} \mathfrak{m}_0\Big(G_{l,n}(r_l), G_{j,n}(r'_l)\Big) = \max_{l=\overline{1,d}} \gamma_l \Big|\ln(r_l) - \ln(r'_l)\Big| =: \varrho^{(d)}\Big(r, r'\Big), \tag{8.5}$$

3^0. Set $\big\|K_r - K_{r'}\big\|_\infty = \sup_{z\in\mathbb{R}^d}\big|K_r(z) - K_{r'}(z)\big|$ and note that for any $x \in \mathbb{X}_1^d \times \mathbb{X}_{d+1}$ and for any $\mathfrak{h} = \big(r, \mathfrak{z}, y^{(d)}\big)$, $\mathfrak{h}' = \big(r', \mathfrak{z}', z^{(d)}\big)$

$$\begin{aligned}
\big|G(\mathfrak{h}, x) - G(\mathfrak{h}', x)\big| &\le \|g\|_\infty \big\|K_r - K_{r'}\big\|_\infty + \|K\|_\infty \left[V_r \vee V_{r'}\right]^{-1} \big|g\big(\mathfrak{z}, x\big) - g\big(\mathfrak{z}', x\big)\big|, \\
&+ \|g\|_\infty V_{r'}^{-1} \left|K\left(\vec{\rho}\big(x^{(d)}, y^{(d)}\big)/r'\right) - K\left(\vec{\rho}\big(x^{(d)}, z^{(d)}\big)/r'\right)\right| \\
&\le \|g\|_\infty \big\|K_r - K_{r'}\big\|_\infty + L_\alpha \|K\|_\infty \left[V_r \vee V_{r'}\right]^{-1} \left[\mathfrak{d}(\mathfrak{z}, \mathfrak{z}')\right]^\alpha \\
&+ \|g\|_\infty V_{r'}^{-1} \left|K\left(\vec{\rho}\big(x^{(d)}, y^{(d)}\big)/r'\right) - K\left(\vec{\rho}\big(x^{(d)}, z^{(d)}\big)/r'\right)\right|.
\end{aligned}$$

The get the last inequality we have used Assumption 9 (**ii**). Using Assumption 9 (**i**) we have

$$\begin{aligned}
\left|K\left(\vec{\rho}\big(x^{(d)}, y^{(d)}\big)/r'\right) - K\left(\vec{\rho}\big(x^{(d)}, z^{(d)}\big)/r'\right)\right| &\le L_1 \max_{l=\overline{1,d}} \left[(r'_l)^{-1} \left|\rho_l\big(x_l, y_l\big) - \rho_l\big(x_l, z_l\big)\right|\right] \\
&\le L_1 \max_{l=\overline{1,d}} \left[(r'_l)^{-1} \rho_l\big(y_l, z_l\big)\right].
\end{aligned}$$

To get the last inequality we have taken into account that ρ_l, $l = \overline{1,d}$, are semi-metrics. Note also that $(r'_l)^{-1} \le V_{r'}^{-1}$ for any $l = \overline{1,d}$, since $r'_l \le 1$ and we obtain

$$\left|K\left(\vec{\rho}\big(x^{(d)}, y^{(d)}\big)/r'\right) - K\left(\vec{\rho}\big(x^{(d)}, z^{(d)}\big)/r'\right)\right| \le L_1 V_{r'}^{-1} \rho^{(d)}\big(y^{(d)}, z^{(d)}\big), \tag{8.6}$$

where we have put $\rho^{(d)} = \max_{l=\overline{1,d}} \rho_l$. Obviously,

$$\big\|K_r - K_{r'}\big\|_\infty \le \|K\|_\infty \big|V_r^{-1} - V_{r'}^{-1}\big| + \left[V_r \vee V_{r'}\right]^{-1} \big\|K\big(\,\cdot\,/r\big) - K\big(\,\cdot\,/r'\big)\big\|_\infty. \tag{8.7}$$

We have in view of Assumption 9 (**i**) and (8.5)

$$\begin{aligned}
\big\|K\big(\,\cdot\,/r\big) - K\big(\,\cdot\,/r'\big)\big\|_\infty &\le L_1 \sup_{u\in\mathbb{R}^d} \max_{l=\overline{1,d}} \left[\frac{|u_l|\,|1/r_l - 1/r'_l|}{1 + |u_l|\left(1/r_l \wedge 1/r'_l\right)}\right] \le L_1 \max_{l=\overline{1,d}} \left[\frac{r_l \vee r'_l}{r_l \wedge r'_l} - 1\right] \\
&= L_1\Big[\exp\Big\{\max_{l=\overline{1,d}} \Big|\ln(r_l) - \ln(r'_l)\Big|\Big\} - 1\Big] \le L_1\Big[\exp\Big\{\gamma^{-1}\varrho^{(d)}\Big(r, r'\Big)\Big\} - 1\Big],
\end{aligned}$$

we have put $\gamma = \min[\gamma_1, \ldots, \gamma_d]$. Moreover, we obviously have for any $r, r' \in (0,1]^d$

$$\frac{V_r \vee V_{r'}}{V_r \wedge V_{r'}} \le \frac{V_{r\vee r'}}{V_{r\wedge r'}} = \exp\left\{\sum_{l=1}^d \gamma_l \Big|\ln\big(r_l\big) - \ln\big(r'_l\big)\Big|\right\} \le \exp\Big\{d\varrho^{(d)}\big(r, r'\big)\Big\}.$$

Thus, we finally obtain from (8.7)

$$\big\|K_r - K_{r'}\big\|_\infty \le \left[V_r \vee V_{r'}\right]^{-1} \left[\|K\|_\infty \left(\exp\Big\{d\varrho^{(d)}\big(r, r'\big)\Big\} - 1\right) + L_1\Big[\exp\Big\{\gamma^{-1}\varrho^{(d)}\Big(r, r'\Big)\Big\} - 1\Big]\right].$$

This yields together with (8.6) for any $\mathfrak{h}=\bigl(r,\mathfrak{z},y^{(d)}\bigr)$ and $\mathfrak{h}'=\bigl(r',\mathfrak{z}',z^{(d)}\bigr)$

$$
\begin{aligned}
&\sup_{x\in\mathbb{X}_1^d\times\mathbb{X}_{d+1}}\bigl|G(\mathfrak{h},x)-G(\mathfrak{h}',x)\bigr| \qquad (8.8)\\
&\le\|g\|_\infty\bigl[V_r\vee V_{r'}\bigr]^{-1}\biggl[\|K\|_\infty\Bigl(\exp\Bigl\{d\varrho^{(d)}\bigl(r,r'\bigr)\Bigr\}-1\Bigr)+L_1\Bigl[\exp\Bigl\{\gamma^{-1}\varrho^{(d)}\Bigl(r,r'\Bigr)\Bigr\}-1\Bigr]\biggr]\\
&\qquad\qquad+L_\alpha\|K\|_\infty\bigl[V_r\vee V_{r'}\bigr]^{-1}\bigl[\mathfrak{d}(\mathfrak{z},\mathfrak{z}')\bigr]^\alpha+L_1\|g\|_\infty V_{r'}^{-2}\rho^{(d)}\bigl(y^{(d)},z^{(d)}\bigr)\\
&\le\|g\|_\infty\|K\|_\infty\bigl[V_r\wedge V_{r'}\bigr]^{-1}\biggl[D_0\Bigl(\varrho^{(d)}\Bigr)+D_{d+1}\bigl(\varrho_{d+1}\bigr)+\bigl[V_r\wedge V_{r'}\bigr]^{-1}D_{d+2}\Bigl(\rho^{(d)}\bigl(y^{(d)},z^{(d)}\bigr)\Bigr)\biggr],
\end{aligned}
$$

where we have put $\varrho_{d+1}=[\mathfrak{d}]^\alpha$, $D_{d+1}(z)=\bigl(L_\alpha/\|g\|_\infty\bigr)z$, $D_{d+2}(z)=L_1\bigl(\|g\|_\infty\|K\|_\infty^2\bigr)^{-1}z$ and

$$
D_0(z)=\exp\{dz\}-1+(L_1/\|K\|_\infty)\Bigl(\exp\bigl\{\gamma^{-1}z\bigr\}-1\Bigr).
$$

Putting $L_{d+1}(z)=z$ and $L_{d+2}(z)=z^2$ we obtain from (8.4) and (8.8) for any $\mathfrak{h}=\bigl(r,\mathfrak{z},y^{(d)}\bigr)$ and $\mathfrak{h}'=\bigl(r',\mathfrak{z}',z^{(d)}\bigr)$

$$
\begin{aligned}
&\sup_{x\in\mathbb{X}_1^d\times\mathbb{X}_{d+1}}\bigl|G(\mathfrak{h},x)-G(\mathfrak{h}',x)\bigr|\le G_\infty(r)\vee G_\infty(r')D_0\Bigl(\varrho^{(d)}\bigl(r,r'\bigr)\Bigr)\\
&+L_{d+1}\bigl(G_\infty(r)\vee G_\infty(r')\bigr)D_{d+1}\Bigl(\varrho_{d+1}\bigl(\mathfrak{z},\mathfrak{z}'\bigr)\Bigr)+L_{d+2}\bigl(G_\infty(r)\vee G_\infty(r')\bigr)D_{d+2}\Bigl(\rho^{(d)}\bigl(y^{(d)},z^{(d)}\bigr)\Bigr).
\end{aligned}
$$

We conclude that Assumption 4 (**ii**) is fulfilled. It remains to note that if $\bar{X}_1^d$ consists of a single element then last summand in the right hand side of the latter inequality disappears that correspond formally to $L_{d+2}\equiv0$. This completes the proof of the lemma.

∎

Proof of Lemma 3 In view of (3.21) for any $r\in(0,1]^d$

$$
F_{\mathbf{n_2}}\Bigl(r,\bar{x}^{(d)}\Bigr)\le\mathrm{f}_\infty\|g\|_\infty\int_{\mathbb{X}_1^d}\Bigl|K_r\Bigl(\vec{\rho}\bigl(x^{(d)},\bar{x}^{(d)}\bigr)\Bigr)\Bigr|\,\mu^{(d)}(\mathrm{d}x^{(d)})=:\mathrm{f}_\infty\|g\|_\infty\mathcal{I}_r. \qquad (8.9)
$$

Denote for any $l=\overline{1,d}$

$$
\mathfrak{R}_l(k_l,r_l)=\mathbb{B}_l\Bigl(2^{k_l+1}r_l,\bar{x}_l\Bigr)\setminus\mathbb{B}_l\Bigl(2^{k_l}r_l,\bar{x}_l\Bigr),\quad\mathfrak{R}_l(0,r_l)=\mathbb{B}_l\Bigl(r_l,\bar{x}_l\Bigr),\ \ k_l\in\mathbb{N}.
$$

and for any multi-index $\mathbf{k}=(k_1,\ldots,k_d)\in\mathbb{N}^d$ set $\mathfrak{R}_{\mathbf{k},r}=\mathfrak{R}_1(k_1,r_1)\times\cdots\times\mathfrak{R}_d(k_d,r_d)$. We get in view of Assumption 10 that $\mathbb{X}_1^d=\bigcup_{\mathbf{k}\in\mathbb{N}^d}\mathfrak{R}_{\mathbf{k},r}$ for any $r\in(0,1]^d$ and, therefore,

$$
\mathcal{I}_r=\sum_{\mathbf{k}\in\mathbb{N}^d}\int_{\mathfrak{R}_{\mathbf{k},r}}\Bigl|K_r\Bigl(\vec{\rho}\bigl(x^{(d)},\bar{x}^{(d)}\bigr)\Bigr)\Bigr|\,\mu^{(d)}(\mathrm{d}x^{(d)}).
$$

We note that for any $\mathbf{k}\in\mathbb{N}^d$ that for any $x^{(d)}\in\mathfrak{R}_{\mathbf{k},r}$

$$
\Bigl|K_r\Bigl(\vec{\rho}\bigl(x^{(d)},\bar{x}^{(d)}\bigr)\Bigr)\Bigr|=V_r^{-1}\left|K\left(\frac{\vec{\rho}\bigl(x^{(d)},\bar{x}^{(d)}\bigr)}{r}\right)\right|\le V_r^{-1}\sup_{|u|\notin\Pi_{\mathbf{t}(\mathbf{k})}}|K(u)|=V_r^{-1}\check{K}\bigl(\mathbf{t}(\mathbf{k})\bigr).
$$

where, we have put $\mathbf{t}(\mathbf{k}) = \big(2^{k_1}, \ldots, 2^{k_d}\big)$ and where, remind, $\Pi_t = [0, t_1] \times \cdots \times [0, t_d],\ \ t \in \mathbb{R}^d_+$.

Thus, we obtain from (3.20) of Assumption 10 (remind that $\mu^{(d)}$ is a product measure)

$$\begin{aligned}\mathcal{I}_r &\leq V_r^{-1} \sum_{\mathbf{k}\in\mathbb{N}^d} \check{K}\big(\mathbf{t}(\mathbf{k})\big)\mu^{(d)}\left(\Pi_{\mathbf{t}(\mathbf{k})}\right) \leq V_r^{-1} \sum_{\mathbf{k}\in\mathbb{N}^d} \check{K}\big(\mathbf{t}(\mathbf{k})\big) \left[\prod_{l=1}^{d} \mu_l\left(\mathbb{B}_l\big(2^{k_l+1} r_l, \bar{x}_l\big)\right)\right] \\ &\leq \left[\prod_{l=1}^{d} 2^{\gamma_l} L^{(l)}\right] \sum_{\mathbf{k}\in\mathbb{N}^d} \check{K}\big(\mathbf{t}(\mathbf{k})\big) \left[\prod_{l=1}^{d} 2^{\gamma_l k_l}\right]. \qquad (8.10)\end{aligned}$$

We get finally from (3.19) of Assumption 10 that for any $r \in (0,1]^d$

$$\mathcal{I}_r \leq 2^d L_2 \prod_{l=1}^{d} 2^{\gamma_l} L^{(l)}.$$

The assertion of the lemma follows now from (8.9).

∎

Proof of Lemma 4 Remind, that for the considered problem

$$\mathfrak{H}\big(\mathfrak{h}_{d+2}\big) = \mathbb{X}_1^d\big(\bar{x}^{(d)}\big) := \bigcup_{\mathbf{i}:\ \bar{x}^{(d)}\in \mathrm{X}_{\mathbf{i}}} \ \bigcup_{\mathbf{k}:\ \mathrm{X}_{\mathbf{k}}\cap \mathrm{X}_{\mathbf{i}}\neq\emptyset} \mathrm{X}_{\mathbf{k}}.$$

For any $\bar{x}^{(d)} \in \mathbb{X}_1^d$ and any $\mathrm{r} > 0$ denote $\mathbb{B}_{\rho^{(d)}}\big(\mathrm{r}, \bar{x}^{(d)}\big) = \big\{x^{(d)} \in \mathbb{X}_1^d :\ \ \rho^{(d)}\big(x^{(d)}, \bar{x}^{(d)}\big) \leq \mathrm{r}\big\}$ where, remind, $\rho^{(d)} = \max[\rho_1, \ldots, \rho_d]$. The following inclusion holds in view of Assumption 11

$$\mathbb{B}_{\rho^{(d)}}\big(\mathfrak{t}, \bar{x}^{(d)}\big) \subseteq \mathbb{X}_1^d\big(\bar{x}^{(d)}\big),\ \ \forall \bar{x}^{(d)} \in \mathbb{X}_1^d. \qquad (8.11)$$

Indeed, suppose that $\exists y^{(d)} \in \mathbb{B}_{\rho^{(d)}}\big(\mathfrak{t}, \bar{x}^{(d)}\big)$ such that $y^{(d)} \notin \mathbb{X}_1^d\big(\bar{x}^{(d)}\big)$. Then, the definition of $\mathbb{X}_1^d\big(\bar{x}^{(d)}\big)$ implies that for any $\mathbf{p}, \mathbf{q} \in \mathbf{I}$ such that $\bar{x}^{(d)} \in \mathrm{X}_{\mathbf{p}},\ \ y^{(d)} \in \mathrm{X}_{\mathbf{q}}$ necessarily

$$\mathrm{X}_{\mathbf{p}} \cap \mathrm{X}_{\mathbf{q}} = \emptyset.$$

Hence, in view of Assumption 11, $\rho^{(d)}\big(y^{(d)}, \bar{x}^{(d)}\big) > \mathfrak{t}$ and, therefore, $y^{(d)} \notin \mathbb{B}_{\rho^{(d)}}\big(\mathfrak{t}, \bar{x}^{(d)}\big)$. The obtained contradiction proves (8.11).

Note that in view of Assumption 9 (**ii**) for any $x \in \mathbb{X}_1^d \times \mathbb{X}_{d+1}$ and any $\mathfrak{h} = \big(r, \mathfrak{z}, \bar{x}^{(d)}\big)$

$$|G(\mathfrak{h}, x)| \leq \|g\|_\infty V_r^{-1} \left|K\left(\vec{\rho}\big(x^{(d)}, \bar{x}^{(d)}\big)/r\right)\right|$$

and, therefore, we get from (8.11) and (3.24)

$$\begin{aligned}\sup_{x\in\mathbb{X}_1^d\times\mathbb{X}_{d+1}:\ x^{(d)}\notin\mathbb{X}_1^d\left(\bar{x}^{(d)}\right)} |G(\mathfrak{h}, x)| &\leq \|g\|_\infty V_r^{-1} \sup_{x^{(d)}\notin\mathbb{B}_{\rho^{(d)}}\left(\mathfrak{t},\bar{x}^{(d)}\right)} \left|K\left(\vec{\rho}\big(x^{(d)}, \bar{x}^{(d)}\big)/r\right)\right| \\ &\leq \|g\|_\infty V_r^{-1} \sup_{r\in\mathcal{R}(n)} \sup_{u\notin[0,\mathfrak{t}]^d} |K(u/r)| \leq \|g\|_\infty \|K\|_\infty V_r^{-1} n^{-1} \\ &=: n^{-1} G_\infty(r) = n^{-1} G_\infty\big(\mathfrak{h}^{(d)}\big).\end{aligned}$$

∎

References

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

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

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

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

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

Coifman, R.R. and Weiss, G. (1971). *Analyse harmonique non-commutative sur certaines espaces homogènes.* Lectures Notes in Math. **242**, Springer-Verlag, Berlin.

Coulhon, T. , Kerkyacharian, G. and Petrushev, P. (2011). Heat kernel generated frames in the setting of Dirichlet spaces. *Manuscript.*

Dony, J. , Einmahl, U. and Mason, D. (2006). Uniform in bandwidt consistency of local polinomial regression function estimators. *Australian Journal if Stat.* **35**, 105-120.

Dony, J. and Einmahl, U. (2009). Uniform in bandwidt consistency of kernel regression estimators at a fixed point. *IMS Collection* **5**, (Highdimentional probability. The Lumini volume), 308–325.

Edmunds, D.E. and Triebel, H. (1996). Function Spaces, Entropy numbers and Differential Operators. *Cambridge University Press*

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

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.

Einmahl, U. and Mason, D. M. (2005). Uniform in bandwidt consistency of kernel-type function estimators. *Annals Stat.* **33, 3**, 1380–1403.

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

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

Giné, E. and ,Guillou A. (2002). Rate of strong uniform consistency for multivatiate kernel density estimaton. *Ann. Inst. H.Poincaré, Proba.Statist.* **38**, 907–921.

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

Goldenshluger, A. and Lepski, O. (2008). Universal pointwise selection rule in multivariate function estimation. *Bernoulli* **14**, 1150-1190.

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

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

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

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

KERKYACHARIAN, G. , LEPSKI, O.V. and PICARD, D. (2001). Non linear estimation in anisotropic multi-index denoising. *Probab. Theory and Relat. Fields* **121**, 137–170.

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

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

LEPSKI, O.V. , MAMMEN, E. and SPOKOINY V.G. (1997). Ideal spatial adaptation to inhomogeneous smoothness: an approach based on kernel estimates with variable bandwidth selection. *Ann. Statist.* **25**, 929–947.

LEPSKI, O. V. and SPOKOINY, V. G. (1997). Optimal pointwise adaptive methods in nonparametric estimation. *Annals Stat.* **25**, 2512–2546.

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

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

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

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

SPOKOINY, V. (1996). Adaptive hypothesis testing using wavelets. *Ann. Stat.* **24**, 2477–2498.

STONE, C.J. (1985). Additive regression and other nonparametric models. *Ann. Stat.* **13, 2**, 689–705.

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

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

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

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

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