ISSN 1066-5307, Mathematical Methods of Statistics, 2013, Vol. 22, No. 2, pp. 83–99. © Allerton Press, Inc., 2013.

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

O. Lepski[1*]

[1] *Laboratoire d'Analyse, Topologie, Probab., Univ. Aix-Marseille, France*

Received January 20, 2013; in final form, February 9, 2013

Abstract—In this part of the paper we apply the results obtained in Lepski (2013) to the variety of problems related to empirical processes.

Keywords: upper function, empirical processes, metric entropy.

2000 Mathematics Subject Classification: primary 60E15; secondary 62G07, 62G08.

DOI: 10.3103/S1066530713020014

INTRODUCTION

This Part II of the paper (consisting, in turn, of parts 1 and 2, the latter to appear in the next issue) deals with the application of the results obtained in Proposition 2, Lepski (2013), to the construction of upper functions for generalized empirical processes. The main difficulty in the realization of our approach consists in the following. To apply Proposition 2 to particular problems one has to compute the function $\mathcal{E}$ and there is no general recipe how to do this. The main goal of this and the next part is to provide rather general assumptions under which this quantity can be calculated explicitly. As was discussed in Lepski (2013), upper functions for random objects appear in various areas of probability theory and mathematical statistics. Therefore the problems of different nature require different assumptions. The assumptions presented below are oriented mostly to the problems arising in mathematical statistics, which definitely reflects author's interests. However, some purely probabilistic results like the law of iterated logarithm and the law of logarithm will be established as well.

1. GENERALIZED EMPIRICAL PROCESSES

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

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

$$\xi_{\mathfrak{h}}(n) = n^{-1} \sum_{i=1}^{n} \big[G(\mathfrak{h}, X_i) - \mathbb{E}_{\mathrm{f}} G(\mathfrak{h}, X_i)\big], \qquad \mathfrak{h} \in \mathfrak{H}. \tag{1.1}$$

We will say that $\xi_{\mathfrak{h}}(n)$, $\mathfrak{h} \in \mathfrak{H}$, is a generalized empirical process. Note that if $\mathfrak{h}\colon \mathcal{X} \to \mathbb{R}$ 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

$$\overline{\mathbf{G}}_\infty(\mathfrak{h}) := \sup_{x \in \mathcal{X}} \big|G(\mathfrak{h}, x)\big| < \infty, \qquad \forall \mathfrak{h} \in \mathfrak{H}, \tag{1.2}$$

and it will be referred to as the *bounded case*. Some generalizations concerning the situations, where this assumption fails, are discussed in Section 1.1.

[*] E-mail: lepski@cmi.univ-mrs.fr

Condition (1.2) implies that the random variables $G(\mathfrak{h}, X_i), \mathfrak{h} \in \boldsymbol{\mathfrak{H}}$, and $G(\mathfrak{h}_1, X_i) - G(\mathfrak{h}_2, X_j)$, $\mathfrak{h}_1, \mathfrak{h}_2 \in \boldsymbol{\mathfrak{H}}$, $i = 1, \dots, n$, are bounded, and we obtain in view of Bernstein's inequality $\forall z > 0$

$$\mathbb{P}_{\mathrm{f}}\big\{|\xi_{\mathfrak{h}}(n)| > z\big\} \le 2\exp\bigg\{-\frac{z^2}{A_{\mathrm{f}}^2(\mathfrak{h}) + zB_\infty(\mathfrak{h})}\bigg\}, \tag{1.3}$$

$$\mathbb{P}_{\mathrm{f}}\big\{|\xi_{\mathfrak{h}_1}(n) - \xi_{\mathfrak{h}_2}(n)| > z\big\} \le 2\exp\bigg\{-\frac{z^2}{\mathrm{a}_{\mathrm{f}}^2(\mathfrak{h}_1, \mathfrak{h}_2) + z\mathrm{b}_\infty(\mathfrak{h}_1, \mathfrak{h}_2)}\bigg\}, \tag{1.4}$$

where

$$A_{\mathrm{f}}^2(\mathfrak{h}) = 2n^{-2}\sum_{i=1}^n \mathbb{E}_f G^2(\mathfrak{h}, X_i), \qquad \mathrm{a}_{\mathrm{f}}^2(\mathfrak{h}_1, \mathfrak{h}_2) = 2n^{-2}\sum_{i=1}^n \mathbb{E}_f\big(G(\mathfrak{h}_1, X_i) - G(\mathfrak{h}_2, X_i)\big)^2, \tag{1.5}$$

$$B_\infty(\mathfrak{h}) = (4/3)n^{-1}\sup_{x\in\mathcal{X}}\big|G(\mathfrak{h}, x)\big|, \qquad \mathrm{b}_\infty(\mathfrak{h}_1, \mathfrak{h}_2) = (4/3)n^{-1}\sup_{x\in\mathcal{X}}\big|G(\mathfrak{h}_1, x) - G(\mathfrak{h}_2, x)\big|. \tag{1.6}$$

We conclude that Assumption 1, Lepski (2013), 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 $\boldsymbol{\mathfrak{H}}$. We note also that $\xi_\bullet\colon \boldsymbol{\mathfrak{H}} \to \mathbb{R}$ is P-a.s. continuous in the topology induced by b_∞. Thus, if $\mathfrak{H} \subseteq \boldsymbol{\mathfrak{H}}$ 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, Lepski (2013), is fulfilled.

Thus, in the problems for which Assumption 3, Lepski (2013), holds, the machinery developed in Proposition 2, Lepski (2013), can be applied for $|\xi_{\mathfrak{h}}(n)|$, $\mathfrak{h} \in \mathfrak{H}$. We would like to emphasize, however, that problems studied below are not always related to the consideration of $|\xi_{\mathfrak{h}}(n)|$, $\mathfrak{h} \in \mathfrak{H}$, with $\mathfrak{H}$ being totally bounded, although such problems are also studied. The idea is to reduce them (if necessary) to those for which Proposition 2, Lepski (2013), can be used. For instance, we will be interested in finding upper functions for $|\xi_{\mathfrak{h}}(n)|$ 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 $\mathbb{N}^*$. This will allow us, in particular, to study generalized empirical processes with random number of summands.

1.1. Problem Formulation and Examples

In this section we find upper functions for several functionals of the generalized empirical process $\xi_{\mathfrak{h}}(n)$ defined in (1.1) under condition (1.2). 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 to assume that for some $m \ge 1$

$$\boldsymbol{\mathfrak{H}} = \mathfrak{H}_1 \times \cdots \times \mathfrak{H}_m, \tag{1.7}$$

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

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

with the convention 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 $j = k+1, \dots, m$ the set $\mathfrak{H}_j$ is endowed with the semi-metric ϱ_j and the Borel measure κ_j.

In the next two sections we find upper functions for $|\xi_{\mathfrak{h}}(n)|$ 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 $j = k+1, \dots, m$.

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

$$(\mathcal{X}, \nu) = \big(\mathcal{X}_1 \times \cdots \times \mathcal{X}_p,\ \nu_1 \times \cdots \times \nu_p\big), \tag{1.8}$$

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

Next we will assume that $\mathfrak{H}_m = \mathcal{X}_1$. As a consequence, the assumption that $\mathfrak{H}_m$ is totally bounded is too restrictive. In particular, it does not hold in the case $\mathcal{X} = \mathcal{X}_1 = \mathbb{R}^d$, which appears in many examples. Before presenting the results, let us consider several examples.

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

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

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

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

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

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 ε_i, $i = 1, \ldots, n$, be independent real random variables distributed on $\mathcal{I} \subseteq \mathbb{R}$ and such that $\mathbb{E}\varepsilon_i = 0$ for any $i = 1, \ldots, n$. Let Y_i, $i = 1, \ldots, n$, be independent d-dimensional random vectors. The sequences $\{\varepsilon_i, i = 1, \ldots, n\}$ and $\{Y_i, i = 1, \ldots, n\}$ 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 a given interval.

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

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

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

If $\mathcal{I}$ is a bounded interval, i.e., the ε_i are bounded random variables, then (1.5) and (1.6) hold and the results from Lepski (2013) are directly applicable. However this assumption is too restrictive and it is not satisfied even in the classical Gaussian regression. At the first glance it seems that if $\mathcal{I} = \mathbb{R}$, then Proposition 2, Lepski (2013), is inapplicable here. Although the aforementioned problem lies beyond the scope of the paper, let us briefly discuss how to reduce it to the problem in which the machinery developed in Proposition 2, Lepski (2013), can be applied.

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

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

where, as before, $\bar{G}\colon \boldsymbol{\mathfrak{H}} \times \mathcal{X} \to \mathbb{R}$ is a given mapping satisfying (1.2). For any $y > 0$ define

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

Obviously, for any $y > 0$

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

where $G_y(\mathfrak{h},x)=\bar{G}(\mathfrak{h},\bar{x})u\mathbf{1}_{[-y,y]}(u)$, $x=(\bar{x},u)\in\mathcal{X}:=\bar{\mathcal{X}}\times\mathbb{R}$, $\mathfrak{h}\in\mathfrak{H}$. Since G_y is bounded for any $y>0$, inequalities (1.3) and (1.4) hold and, analogously to (1.5) and (1.6), we have

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

Let also $\mathfrak{H}\subseteq\mathfrak{H}$ be such that the results obtained in Proposition 2, Lepski (2013), are applicable to $|\bar{\xi}_{\mathfrak{h}}(n,y)|$ 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 implies, in view of Theorem 1 below, that upper functions for $|\bar{\xi}_{\mathfrak{h}}(y)|$, $\mathfrak{h}\in\mathfrak{H}$ (for brevity $V(\mathfrak{h},y)$ and $U_q(\mathfrak{h},y)$, $q\geq 1$) can be found in the form:

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

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

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

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

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

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. This allows us to choose $y=y_n$ in an optimal way, i.e., to balance both terms in the last inequalities, which usually leads to sharp upper functions $V_1^{(n)}(\cdot)+y_nV_2^{(n)}(\cdot)$ and $U_{q,1}^{(n)}(\cdot)+y_nU_{q,2}^{(n)}(\cdot)$.

1.2. Main Assumption

Now let us return to the consideration of generalized empirical processes obeying (1.2). Assumption 1 below is the main tool allowing us to compute upper functions *explicitly*. Introduce the following notation: for any $\mathfrak{h}^{(k)}\in\mathfrak{H}_1^k$

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

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

$$\mathbf{G}_\infty\big(\mathfrak{h}^{(k)}\big)\leq G_\infty\big(\mathfrak{h}^{(k)}\big),\quad\forall\mathfrak{h}^{(k)}\in\mathfrak{H}_1^k.\tag{1.9}$$

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

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

For any $n\geq1$, $j=1,\dots,k$, and any $\mathfrak{h}_j\in\mathfrak{H}_j(n)$ define

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

Noting that $\big|\log(t_1)-\log(t_2)\big|$ is a metric on $\mathbb{R}_+\setminus\{0\}$, we equip $\mathfrak{H}_1^k(n)$ with the following semi-metric. For any $n\geq1$ and any $\hat{\mathfrak{h}}^{(k)},\bar{\mathfrak{h}}^{(k)}\in\mathfrak{H}_1^k(n)$ set

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

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

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

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

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

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

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

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

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

In view of the 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 $\widetilde{\mathfrak{H}}_1^k(n) = \widetilde{\mathfrak{H}}_1(n) \times \dots \times \widetilde{\mathfrak{H}}_k(n)$ and let $\big\{\widetilde{\mathfrak{H}}_j(n) \subset \mathfrak{H}_j(n),\ n \ge 1\big\}$, $j = 1, \dots, k$, be a sequence of sets satisfying an additional restriction. We will not be tending here to the maximal generality and complete Assumption 1 by the following condition.

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

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

We note that Assumption 2 obviously holds if for any $j = 1, \dots, k$ the sequence $\big\{\widetilde{\mathfrak{H}}_j(n),\ n \ge 1\big\}$ is an increasing/decreasing sequence of sets.

1.3. Totally Bounded Case

The objective is to find upper functions for $|\xi_{\mathfrak{h}}(n)|$ under Assumption 1 enforced, if necessary, by Assumption 2 and the condition imposed on the entropies of the sets $\mathfrak{H}_j$, $j = k+1, \dots, m$.

1.3.1. Assumptions and main result. The following condition will be additionally imposed in this section.

Assumption 3. *Suppose that* (1.7) *holds and there exist* $N, R < \infty$ *such that for any* $\varsigma > 0$ *and any* $j = k+1, \dots, m$

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

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

We remark that Assumption 3 is fulfilled, in particular, when $(\mathfrak{H}_j, \varrho_j, \kappa_j)$, $j = k+1, \dots, m$, are bounded and satisfy the doubling condition. Note also that this assumption can be considerably weakened, see discussion after Theorem 1.

Notation. 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 $\mathfrak{h} \in \boldsymbol{\mathfrak{H}}$ set

$$F_{\mathbf{n}_2}(\mathfrak{h}) = \begin{cases} \sup_{i=1,\dots,\mathbf{n}_2} \mathbb{E}_{\mathrm{f}}\big|G(\mathfrak{h}, X_i)\big|, & \mathbf{n}_1 \ne \mathbf{n}_2, \\ (\mathbf{n}_2)^{-1} \sum_{i=1}^{\mathbf{n}_2} \mathbb{E}_{\mathrm{f}}\big|G(\mathfrak{h}, X_i)\big|, & \mathbf{n}_1 = \mathbf{n}_2, \end{cases}$$

and remark that if additionally X_i, $i \geq 1$, are identically distributed then we have the same definition of $F_{\mathbf{n}_2}(\cdot)$ in both cases. We note that

$$F_{\mathbf{n}_2} := \sup_{n\in\widetilde{\mathbf{N}}} \sup_{\mathfrak{h}\in\widetilde{\mathfrak{H}}(n)} F_{\mathbf{n}_2}(\mathfrak{h}) \leq \sup_{n\in\widetilde{\mathbf{N}}} \overline{G}_n < \infty$$

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

$$\mathbf{n} = \begin{cases} \mathbf{n}_1, & \mathbf{n}_1 = \mathbf{n}_2, \\ n[\mathbf{n}_1], & \mathbf{n}_1 \neq \mathbf{n}_2, \end{cases} \qquad \beta = \begin{cases} 0, & \mathbf{n}_1 = \mathbf{n}_2; \\ \boldsymbol{b}, & \mathbf{n}_1 \neq \mathbf{n}_2, \end{cases}$$

where, recall, $n[\cdot]$ is defined in Assumption 2.

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

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

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

Here δ_* is the smallest solution of the equation $(48\delta)^{-1}s^*(\delta) = 1$, where, recall, $s^*(\delta) = (6/\pi^2)\big(1 + [\log\delta]^2\big)^{-1}$, $\delta \geq 0$. The quantities N, R are defined in Assumption 3.

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 2.1.2, which precedes the proof of Theorem 1.

Result. For any $\mathbf{r} \in \overline{\mathbb{N}}$ put $F_{\mathbf{n}_2,\mathbf{r}}(\mathfrak{h}) = \max[F_{\mathbf{n}_2}(\mathfrak{h}), e^{-\mathbf{r}}]$ and define for any $\mathfrak{h} \in \mathfrak{H}$, $u \geq 0$ and $q > 0$

$$\mathcal{V}_{\mathbf{r}}^{(u)}(n,\mathfrak{h}) = \lambda_1\sqrt{G_\infty(\mathfrak{h}^{(k)})\big(F_{\mathbf{n}_2,\mathbf{r}}(\mathfrak{h})n^{-1}\big)\Big(P(\mathfrak{h}^{(k)}) + 2\log\big\{1 + \big|\log\big(F_{\mathbf{n}_2,\mathbf{r}}(\mathfrak{h})\big)\big|\big\} + u\Big)}$$
$$+ \lambda_2 G_\infty(\mathfrak{h}^{(k)})\big(n^{-1}\log^\beta(n)\big)\Big(P(\mathfrak{h}^{(k)}) + 2\log\big\{1 + \big|\log\big(F_{\mathbf{n}_2,\mathbf{r}}(\mathfrak{h})\big)\big|\big\} + u\Big),$$

$$\mathcal{U}_{\mathbf{r}}^{(u,q)}(n,\mathfrak{h}) = \lambda_1\sqrt{G_\infty(\mathfrak{h}^{(k)})\big(F_{\mathbf{n}_2,\mathbf{r}}(\mathfrak{h})n^{-1}\big)\Big(M_q(\mathfrak{h}^{(k)}) + 2\log\big\{1 + \big|\log\big(F_{\mathbf{n}_2,\mathbf{r}}(\mathfrak{h})\big)\big|\big\} + u\Big)}$$
$$+ \lambda_2 G_\infty\big(\mathfrak{h}^{(k)}\big)\big(n^{-1}\log^\beta(n)\big)\Big(M_q(\mathfrak{h}^{(k)}) + 2\log\big\{1 + \big|\log\big(F_{\mathbf{n}_2,\mathbf{r}}(\mathfrak{h})\big)\big|\big\} + u\Big).$$

Theorem 1. *Let Assumptions* 1 *and* 3 *be fulfilled. If* $\mathbf{n}_1 \neq \mathbf{n}_2$ *suppose additionally that Assumption* 2 *holds. Then for any* $\mathbf{r} \in \mathbb{N}$, $\boldsymbol{b} > 1$, $u \geq 1$ *and* $q \geq 1$

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

$$\mathbb{E}_{\mathrm{f}}\Big\{\sup_{n\in\widetilde{\mathbf{N}}} \sup_{\mathfrak{h}\in\widetilde{\mathfrak{H}}(n)} \big[|\xi_{\mathfrak{h}}(n)| - \mathcal{U}_{\mathbf{r}}^{(u,q)}(n,\mathfrak{h})\big]\Big\}_+^q \leq c_q\Big[\sqrt{(\mathbf{n}_1)^{-1}F_{\mathbf{n}_2}\underline{G}_{\mathbf{n}}} \vee \big((\mathbf{n}_1)^{-1}\log^\beta(\mathbf{n}_2)\underline{G}_{\mathbf{n}}\big)\Big]^q e^{-u},$$

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

Remark 2. The inspection of the proof of the theorem allows us to assert that Assumption 3 can be weakened. The condition that is needed in view of the used technique: for some $\alpha \in (0,1)$, $L < \infty$

$$\sup_{\varsigma>0} \varsigma^{-\alpha}\mathfrak{E}_{\mathfrak{H}_j,\varrho_j}(\varsigma) \leq L, \qquad j = k+1,\ldots,m. \tag{1.10}$$

In particular, this allows us to consider the generalized empirical processes indexed by sets of smooth functions. However the latter assumption does not permit us to express upper functions explicitly as this is done in Theorem 1. This explains why we prefer to state our results under Assumption 3.

Several other remarks are in order.

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

2°. Next, putting $\mathbf{r} = +\infty$ we get the results of the theorem with $F_{\mathbf{n}_2,\mathbf{r}}(\cdot) = F_{\mathbf{n}_2}(\cdot)$. This 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 $F_{\mathbf{n}_2}(\cdot)$.

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

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

Suppose that for some $\mathfrak{c} > 0$, $\mathfrak{b} > 0$

$$\mathfrak{c} \leq \underline{G}_n \leq \overline{G}_n \leq \mathfrak{c} n^{\mathfrak{b}}, \qquad \forall n \geq 1. \tag{1.11}$$

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

Moreover we will assume that

$$\sup_{n\geq 1} \sup_{\mathfrak{h} \in \widetilde{\mathfrak{H}}(n)} \sup_{i \geq 1} \mathbb{E}_{\mathrm{f}} |G(\mathfrak{h}, X_i)| =: \mathbf{F} < \infty. \tag{1.12}$$

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

$$\mathcal{L}^{(k)}(z) \leq \mathfrak{a} \log\{1 + \log(z)\}, \qquad \forall z \geq 3. \tag{1.13}$$

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

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

Theorem 2. *Let Assumptions* 1, 2 *and* 3 *be fulfilled and suppose additionally that* (1.11)*,* (1.12)*, and* (1.13) *hold. Then there exists* $\Upsilon > 0$ *such that for any* $\mathbf{j} \geq 3$ *and any* $a > 2$

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

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

Remark 3. An 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 $2419[\log(\mathbf{j})]^{-(1+y)}$. This makes sensible to consider small values of $\mathbf{j}$.

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

$$\limsup_{n\to\infty} \sup_{\mathfrak{h}^{(k)} \in \overline{\mathfrak{H}}_1^k(n,a)} \left[\frac{\sqrt{n}\, \eta_{\mathfrak{h}^{(k)}}(n)}{\sqrt{G_\infty\big(\mathfrak{h}^{(k)}\big) \log\log(n)}}\right] \leq \Upsilon, \qquad \mathbb{P}_{\mathrm{f}}\text{-a.s.} \tag{1.14}$$

2. PROOF OF THEOREMS 1 AND 2

2.1. Proof of Theorem 1

2.1.1. Preliminaries. We start the proof with several technical results used in the sequel. Put for any $i = 1, \dots, \mathbf{n}_2$, $y \in [\mathbf{n}_1/\mathbf{n}_2, 1]$ and $\alpha = \boldsymbol{b}\big[\log(\mathbf{n}_2)\big]^{-1}$,

$$\mathcal{Q}_i(y) = \mathbf{1}_{(i/\mathbf{n}_2,1]}(y) + (\mathbf{n}_2 y - i + 1)^{\alpha}\mathbf{1}_{\Delta_i}(y), \qquad Q_i(y) = y^{-1}\mathcal{Q}_i(y).$$

Here $\Delta_i = \big((i-1)/\mathbf{n}_2, i/\mathbf{n}_2\big]$, $i = 3, \dots, \mathbf{n}_2$, and $\Delta_2 = \big[1/\mathbf{n}_2, 2/\mathbf{n}_2\big]$.

For any $a \geq 1$ let $\lceil a \rceil$ be the smallest integer larger or equal to a. This implies, in particular, that $y \in \Delta_{\lceil \mathbf{n}_2 y \rceil}$. First we note that for any $y, \bar{y} \in [\mathbf{n}_1/\mathbf{n}_2, 1]$ and any $i = 1, \dots, \mathbf{n}_2$

$$\mathcal{Q}_i(y) \leq 1, \qquad \mathcal{Q}_i(y) = 0, \quad \forall i > \lceil \mathbf{n}_2 y \rceil, \qquad |\mathcal{Q}_i(y) - \mathcal{Q}_i(\bar{y})| \leq 1 \wedge |\mathbf{n}_2(y - \bar{y})|^{\alpha}. \tag{2.1}$$

The first and the third inequalities imply for any $i = 1, \dots, \mathbf{n}_2$ and any $y, \bar{y} \in [\mathbf{n}_1/\mathbf{n}_2, 1]$

$$|Q_i(y) - Q_i(\bar{y})| \leq (y \wedge \bar{y})^{-1}\left[|\mathbf{n}_2(y - \bar{y})|^{\alpha} + \left(1 - \frac{y \wedge \bar{y}}{y \vee \bar{y}}\right)\right]. \tag{2.2}$$

For any $z, z' \in \mathbb{R}_+$ denote $\mathfrak{w}(z, z') = \big(1 - \sqrt{\frac{z \wedge z'}{z \vee z'}}\big)^{1/2}$, and remark that $\mathfrak{w}$ is a metric on $\mathbb{R}_+$. This follows from the relation

$$\sqrt{2}\mathfrak{w}(z, z') = \left[\mathbb{E}\left(\frac{b(z)}{\sqrt{z}} - \frac{b(z')}{\sqrt{z'}}\right)^2\right]^{1/2},$$

where b is the standard Wiener process. Taking into account that $y, \bar{y} \geq 1/2$ and that $\mathfrak{w}(y \wedge \bar{y}) \leq 1$ we obtain from (2.2)

$$|Q_i(y) - Q_i(\bar{y})| \leq 8e^{\boldsymbol{b}}\big[\mathfrak{w}(y, \bar{y})\big]^{\alpha}. \tag{2.3}$$

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

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

we obtain from (2.3) for any $\boldsymbol{b} > 0$, $y, \bar{y} \in [\mathbf{n}_1/\mathbf{n}_2, 1]$ and $\mathbf{n}_2 \geq 3$

$$\sup_{i=1,\dots,\mathbf{n}_2} \big|Q_i(y) - Q_i(\bar{y})\big| \leq 8e\left[\frac{\log(\mathbf{n}_2)}{1 - \log\big(\mathfrak{w}(y, \bar{y})\big)}\right]^{\boldsymbol{b}}. \tag{2.4}$$

Next, for any $y, \bar{y} \in [\mathbf{n}_1/\mathbf{n}_2, 1]$

$$|\mathcal{Q}_i(y) - \mathcal{Q}_i(\bar{y})| = 0, \qquad i \notin \big\{\lceil \mathbf{n}_2(y \wedge \bar{y}) \rceil, \dots, \lceil \mathbf{n}_2(y \vee \bar{y}) \rceil\big\}. \tag{2.5}$$

We have for any $y, \bar{y} \in \big[\mathbf{n}_1/\mathbf{n}_2, 1\big]$ in view of the first and the third bounds in (2.1) and (2.5)

$$\sum_{i=1}^{\mathbf{n}_2} |\mathcal{Q}_i(y) - \mathcal{Q}_i(\bar{y})|^2 \leq \begin{cases} 2\mathbf{n}_2|y - \bar{y}|, & \lceil \mathbf{n}_2(y \vee \bar{y}) \rceil - \lceil \mathbf{n}_2(y \wedge \bar{y}) \rceil \geq 3, \\ 3|\mathbf{n}_2(y - \bar{y})|^{2\alpha}, & \lceil \mathbf{n}_2(y \vee \bar{y}) \rceil - \lceil \mathbf{n}_2(y \wedge \bar{y}) \rceil \leq 2. \end{cases}$$

To get the first inequality we have also used the fact that $\lceil \mathbf{n}_2(y \vee \bar{y}) \rceil - \lceil \mathbf{n}_2(y \wedge \bar{y}) \rceil \geq 3$ implies $\mathbf{n}_2(y \vee \bar{y} - y \wedge \bar{y}) > 2$ and therefore

$$\lceil \mathbf{n}_2(y \vee \bar{y}) \rceil - \lceil \mathbf{n}_2(y \wedge \bar{y}) \rceil + 1 \leq \mathbf{n}_2(y \vee \bar{y} - y \wedge \bar{y}) + 2 \leq 2\mathbf{n}_2(y \vee \bar{y} - y \wedge \bar{y}) = 2\mathbf{n}_2|y - \bar{y}|.$$

Thus we have for any $y, \bar{y} \in [\mathbf{n}_1/\mathbf{n}_2, 1]$

$$\sqrt{\sum_{i=1}^{\mathbf{n}_2} |\mathcal{Q}_i(y) - \mathcal{Q}_i(\bar{y}|^2} \leq \sqrt{2\mathbf{n}_2|y - \bar{y}|} + \sqrt{3}|\mathbf{n}_2(y - \bar{y})|^{\alpha} \leq 2\sqrt{\mathbf{n}_2}\mathfrak{w}(y, \bar{y}) + 2\sqrt{3}e^{\boldsymbol{b}}\big[\mathfrak{w}(y, \bar{y})\big]^{\alpha}.$$

Hence we get

$$\sqrt{\sum_{i=1}^{\mathbf{n}_2}|Q_i(y)-Q_i(\bar y)|^2} \le 8\sqrt{\mathbf{n}_2}\mathfrak{w}(y,\bar y) + 4\sqrt{3}e^{\boldsymbol{b}}\big[\mathfrak{w}(y,\bar y)\big]^{\alpha}$$

$$\le 8\sqrt{\mathbf{n}_2}\mathfrak{w}(y,\bar y) + 4\sqrt{3}e\left[\frac{\log(\mathbf{n}_2)}{1-\log\big(\mathfrak{w}(y,\bar y)\big)}\right]^{\boldsymbol{b}}.$$

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

$$\sqrt{\sum_{i=1}^{\mathbf{n}_2}|Q_i(y)-Q_i(\bar y)|^2} \le 8\sqrt{\mathbf{n}_2}\Big[\mathfrak{w}(y,\bar y) + \sqrt{3/4}e(2\boldsymbol{b}/e)^{\boldsymbol{b}}\big\{1-\log(\mathfrak{w}(y,\bar y))\big\}^{-\boldsymbol{b}}\Big].$$

Finally we get for any $y,\bar y\in[\mathbf{n}_1/\mathbf{n}_2,1]$ and any $\boldsymbol{b}>1$

$$\sqrt{\sum_{i=1}^{\mathbf{n}_2}|Q_i(y)-Q_i(\bar y)|^2} \le 8[2^{\boldsymbol{b}}+1](\boldsymbol{b})^{\boldsymbol{b}}\sqrt{\mathbf{n}_2}\big[1-\log\big(\mathfrak{w}(y,\bar y)\big)\big]^{-\boldsymbol{b}}. \tag{2.6}$$

2.1.2. Constants. The following constants appeared in the description of upper functions and inequalities obtained in Theorem 1. Let $\chi=0$ if $\mathbf{n}_1=\mathbf{n}_2$ and $\chi=1$ if $\mathbf{n}_1\neq\mathbf{n}_2$. Then

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

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

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

Put also $C_D := \big[\sup_{j=0,k+1,\ldots,m}\ \sup_{z\in[0,1]} D'_j(z)\big]\vee 2$, where D'_j is the first derivative of the function D_j. Set, finally, $\boldsymbol{c}_{\boldsymbol{b}} = 4\sqrt{2}\big[2^{\boldsymbol{b}}+1\big]\boldsymbol{b}^{\boldsymbol{b}}$ and let

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

2.1.3. Proof of the theorem. 1°. Put for any $i=1,\ldots,n$

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

and define for any $y\in(\mathbf{n}_1/\mathbf{n}_2,1]$ and any $\mathfrak{h}\in\mathfrak{H}$ the random function

$$\boldsymbol{\xi}(y,\mathfrak{h}) = \mathbf{n}_2^{-1}\sum_{i=1}^{\mathbf{n}_2}\varepsilon\big(\mathfrak{h},X_i\big)Q_i(y). \tag{2.7}$$

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 $\mathfrak{h}\in\mathfrak{H}$. 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 2 and the definition of the number $\mathbf{n}$.

In view of Bernstein's inequality Assumption 1 in Lepski (2013) is fulfilled with $\theta=\mathrm{h}=:(y,\mathfrak{h})$ and $\bar\theta=\bar{\mathrm{h}}=:(\bar y,\bar{\mathfrak{h}})$,

$$A^2(\theta) = A^2_{\mathrm{f}}(\mathrm{h}) := 2\mathbf{n}_2^{-2}\sum_{i=1}^{\mathbf{n}_2}Q_i^2(y)\mathbb{E}_fG^2(\mathfrak{h},X_i), \tag{2.8}$$

$$\mathrm{a}^2(\theta,\bar\theta) = \mathrm{a}^2_{\mathrm{f}}(\mathrm{h},\bar{\mathrm{h}}) := 2\mathbf{n}_2^{-2}\sum_{i=1}^{\mathbf{n}_2}\mathbb{E}_f\big[Q_i(y)G(\mathfrak{h},X_i) - Q_i(\bar y)G(\bar{\mathfrak{h}},X_i)\big]^2, \tag{2.9}$$

$$B(\theta) = B_\infty(\mathrm{h}) = (4/3)\mathbf{n}_2^{-1}\big[\sup_{i=1,\dots,\mathbf{n}_2} Q_i(y)\big]\sup_{x\in\mathcal{X}}\big|G(\mathfrak{h},x)\big|, \tag{2.10}$$

$$\mathrm{b}(\theta,\bar\theta) = \mathrm{b}_\infty(\mathrm{h},\bar{\mathrm{h}}) := (2/3)\mathbf{n}_2^{-1}\sup_{i=1,\dots,n}\sup_{x\in\mathcal{X}}\big|\varepsilon(\mathfrak{h},x)Q_i(y) - \varepsilon(\bar{\mathfrak{h}},x)Q_i(\bar y)\big|. \tag{2.11}$$

Note that $\mathrm{a_f}$ and b_∞ are semi-metrics on $[\mathbf{n}_1/\mathbf{n}_2,1]\times\boldsymbol{\mathfrak{H}}$ and $\boldsymbol{\xi}(\cdot,\cdot)$ is obviously continuous on $[\mathbf{n}_1/\mathbf{n}_2,1]\times\mathfrak{H}(\mathbf{n})$ in the topology generated by b_∞. Moreover, A_f and B_∞ are bounded and therefore Assumption 2 in Lepski (2013) is fulfilled.

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

We obtain from (2.8)–(2.11) and (2.4) for any $\mathrm{h},\bar{\mathrm{h}}\in[\mathbf{n}_1/\mathbf{n}_2,1]\times\mathfrak{H}(\mathbf{n})$

$$A_\mathrm{f}^2(\mathrm{h}) \le 2(\mathbf{n}_1)^{-1}F_{\mathbf{n}_2}(\mathfrak{h})G_\infty(\mathfrak{h}^{(k)}), \qquad B_\infty(\mathrm{h}) \le (4/3)(\mathbf{n}_1)^{-1}G_\infty(\mathfrak{h}^{(k)}), \tag{2.12}$$

$$\mathrm{b}_\infty(\mathrm{h},\bar{\mathrm{h}}) \le \frac{4\log^\beta(\mathbf{n}_2)}{3\mathbf{n}_1}\Big\{\big\|G(\mathfrak{h},\cdot)-G(\bar{\mathfrak{h}},\cdot)\big\|_\infty + \gamma 8eG_\infty(\bar{\mathfrak{h}}^{(k)})\big[1-\log\big(\mathfrak{w}(y,\bar y)\big)\big]^{-\boldsymbol{b}}\Big\}, \tag{2.13}$$

where, recall, $\gamma=0$ if $\mathbf{n}_1=\mathbf{n}_2$ and $\gamma=1$ if $\mathbf{n}_1\neq\mathbf{n}_2$. Here we have used that if $\mathbf{n}_1=\mathbf{n}_2$ the second term in the last inequality disappears.

We also get using (2.1) and (2.6)

$$\begin{aligned}\mathrm{a_f}(\mathrm{h},\bar{\mathrm{h}}) &\le \sqrt{2}\mathbf{n}_2^{-1}\Bigg\{\sqrt{\sum_{i=1}^{\mathbf{n}_2}Q_i^2(y)\mathbb{E}_f\big[G(\mathfrak{h},X_i)-G(\bar{\mathfrak{h}},X_i)\big]^2} \\ &\quad + \sqrt{F_{\mathbf{n}_2}(\bar{\mathfrak{h}})G_\infty(\bar{\mathfrak{h}}^{(k)})\sum_{i=1}^{\mathbf{n}_2}\big(Q_i(y)-Q_i(\bar y)\big)^2}\Bigg\} \\ &\le \sqrt{2}(\mathbf{n}_1)^{-1/2}\Big\{\sqrt{\big(F_{\mathbf{n}_2}(\mathfrak{h})+F_{\mathbf{n}_2}(\bar{\mathfrak{h}})\big)\big\|G(\mathfrak{h},\cdot)-G(\bar{\mathfrak{h}},\cdot)\big\|_\infty} \\ &\quad + \chi\boldsymbol{c_b}\sqrt{2F_{\mathbf{n}_2}(\bar{\mathfrak{h}})G_\infty(\bar{\mathfrak{h}}^{(k)})}\big[1-\log\big(\mathfrak{w}(y,\bar y)\big)\big]^{-\boldsymbol{b}}\Big\},\end{aligned} \tag{2.14}$$

where we have put $\boldsymbol{c_b} = 4\sqrt{2}[2^{\boldsymbol{b}}+1](\boldsymbol{b})^{\boldsymbol{b}}$. 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 $\mathfrak{H}\big(\mathbf{n},\tau\big)=\big\{\mathfrak{h}\in\mathfrak{H}(\mathbf{n})\colon F_{\mathbf{n}_2}(\mathfrak{h})\le\tau\big\}$. Our first step consists in establishing an upper function for $|\boldsymbol{\xi}(\cdot,\cdot)|$ on $\mathrm{H}(\tau) := \big[\mathbf{n}_1/\mathbf{n}_2,1\big]\times\mathfrak{H}(\mathbf{n},\tau)$. As usual, the supremum over empty set is supposed to be zero.

2°. Note that in view of (2.12) and (2.14) for any $\mathrm{h},\bar{\mathrm{h}}\in\mathrm{H}(\tau)$

$$A_\mathrm{f}^2(\mathrm{h}) \le 2\tau(\mathbf{n}_1)^{-1}G_\infty(\mathfrak{h}^{(k)}), \qquad B_\infty(\mathrm{h}) \le \frac{4\log^\beta(\mathbf{n}_2)}{3\mathbf{n}_1}G_\infty(\mathfrak{h}^{(k)}), \tag{2.15}$$

$$\mathrm{a_f}(\mathrm{h},\bar{\mathrm{h}}) \le 2\sqrt{\tau}(\mathbf{n}_1)^{-1/2}\Big\{\sqrt{\|G(\mathfrak{h},\cdot)-G(\bar{\mathfrak{h}},\cdot)\|_\infty} + \chi\boldsymbol{c_b}\sqrt{G_\infty(\bar{\mathfrak{h}}^{(k)})}\big[1-\log\big(\mathfrak{w}(y,\bar y)\big)\big]^{-\boldsymbol{b}}\Big\}. \tag{2.16}$$

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

$$\mathrm{a_f}(\mathrm{h},\overline{\mathrm{h}}) \le A_\mathrm{f}(\mathrm{h}) + A_\mathrm{f}(\overline{\mathrm{h}}) \le \sqrt{8\tau(\mathbf{n}_1)^{-1}[G_\infty(\mathfrak{h}^{(k)})\vee G_\infty(\overline{\mathfrak{h}}^{(k)})]}, \tag{2.17}$$

$$\mathrm{b}_\infty(\mathrm{h},\overline{\mathrm{h}}) \le B_\infty(\mathrm{h}) + B_\infty(\overline{\mathrm{h}}) \le \frac{8\log^\beta(\mathbf{n}_2)}{3\mathbf{n}_1}[G_\infty(\mathfrak{h}^{(k)})\vee G_\infty(\overline{\mathfrak{h}}^{(k)})]. \tag{2.18}$$

Set

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

We get for any $\mathfrak{h},\bar{\mathfrak{h}}$ satisfying $\varrho^{(k)}(\mathfrak{h}^{(k)},\bar{\mathfrak{h}}^{(k)})\vee\sup_{j=k+1,\dots,m}\varrho_j\big(\mathfrak{h}_j,\bar{\mathfrak{h}}_j\big)\leq 1$ in view of Assumption 1 (i)

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

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

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

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

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

The latter inequality together with (2.13) and (2.16) yields for any $\mathrm{h},\bar{\mathrm{h}}\in\mathrm{H}(\tau)$

$$\begin{aligned}\mathrm{a_f}(\mathrm{h},\bar{\mathrm{h}})\leq\mathfrak{a}\bigg\{&\bigg(\mathcal{G}(\mathfrak{h}^{(k)},\bar{\mathfrak{h}}^{(k)})\varrho^{(k)}(\mathfrak{h}^{(k)},\bar{\mathfrak{h}}^{(k)})+\sum_{j=k+1}^{m}\widetilde{L}_j\{\mathcal{G}(\mathfrak{h}^{(k)},\bar{\mathfrak{h}}^{(k)})\}\varrho_j(\mathfrak{h}_j,\bar{\mathfrak{h}}_j)\bigg)^{1/2}\\&+\chi\sqrt{\mathcal{G}(\mathfrak{h}^{(k)},\bar{\mathfrak{h}}^{(k)})}\big[1-\log\big(\mathfrak{w}(y,\bar{y})\big)\big]^{-\boldsymbol{b}}\bigg\},\end{aligned}\tag{2.19}$$

$$\begin{aligned}\mathrm{b}_\infty(\mathrm{h},\bar{\mathrm{h}})\leq\mathfrak{b}\bigg\{&\mathcal{G}(\mathfrak{h}^{(k)},\bar{\mathfrak{h}}^{(k)})\varrho^{(k)}(\mathfrak{h}^{(k)},\bar{\mathfrak{h}}^{(k)})+\sum_{j=k+1}^{m}\widetilde{L}_j\{\mathcal{G}(\mathfrak{h}^{(k)},\bar{\mathfrak{h}}^{(k)})\}\varrho_j(\mathfrak{h}_j,\bar{\mathfrak{h}}_j)\\&+\chi\mathcal{G}(\mathfrak{h}^{(k)},\bar{\mathfrak{h}}^{(k)})\big[1-\log\big(\mathfrak{w}(y,\bar{y})\big)\big]^{-\boldsymbol{b}}\bigg\},\end{aligned}\tag{2.20}$$

where we have put $\mathfrak{a}=2\sqrt{\tau}(\mathbf{n}_1)^{-1/2}\big(\sqrt{C_D}\vee[\chi\boldsymbol{c_b}]\big)$, $\mathfrak{b}=\frac{4\big(C_D\vee 8e\big)\log^\beta(\mathbf{n}_2)}{3\mathbf{n}_1}$.

3°. We note that in view of (2.15) Assumption 1 (1) in Lepski (2013) is verified on $\mathrm{H}(\tau)$ with

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

The idea now is to apply Proposition 2, Lepski (2013), with $\Theta=\mathrm{H}(\tau)$. Put

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

which 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, Lepski (2013), one has to bound from above the function

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

defined in this proposition. Here, in our case, $\mathrm{a}=\mathrm{a_f}$, $\mathrm{b}=\mathrm{b}_\infty$ and

$$\begin{aligned}\Theta_A\big(\underline{A}u\big)&=\big\{\mathfrak{h}\in\mathfrak{H}(\mathbf{n},\tau)\colon G_\infty(\mathfrak{h}^{(k)})\leq u^2\underline{G}_{\mathbf{n}}[\tau]\big\}\times[\mathbf{n}_1/\mathbf{n}_2,1],\\\Theta_B\big(\underline{B}v\big)&=\big\{\mathfrak{h}\in\mathfrak{H}(\mathbf{n},\tau)\colon G_\infty(\mathfrak{h}^{(k)})\leq v\underline{G}_{\mathbf{n}}[\tau]\big\}\times[\mathbf{n}_1/\mathbf{n}_2,1].\end{aligned}$$

Let us make several remarks.

3°a. First recall that

$$e^{(\mathrm{a_f})}_{s^*}\big(\underline{A}u,\Theta_A(\underline{A}u)\big)=\sup_{\delta>0}\delta^{-2}\mathfrak{E}_{\Theta_A(\underline{A}u),\,\mathrm{a_f}}\big(\underline{A}u(48\delta)^{-1}s^*(\delta)\big),$$

$$e^{(\mathrm{b}_\infty)}_{s^*}\big(\underline{B}v,\Theta_B(\underline{B}v)\big)=\sup_{\delta>0}\delta^{-1}\mathfrak{E}_{\Theta_B(\underline{B}v),\,\mathrm{b}_\infty}\big(\underline{B}v(48\delta)^{-1}s^*(\delta)\big).$$

We have in view of (2.17) and (2.18) that for any $\mathrm{h},\overline{\mathrm{h}}\in\mathrm{H}(\tau)$

$$\mathrm{a_f}\big(\mathrm{h},\overline{\mathrm{h}}\big)\leq\mathfrak{a}\sqrt{\big[G_\infty(\mathfrak{h}^{(k)})\vee G_\infty(\overline{\mathfrak{h}}^{(k)})\big]},\qquad \mathrm{b}_\infty\big(\mathrm{h},\overline{\mathrm{h}}\big)\leq\mathfrak{b}\big[G_\infty(\mathfrak{h}^{(k)})\vee G_\infty(\overline{\mathfrak{h}}^{(k)})\big],$$

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

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

This yields for any $\delta\leq\delta_*$, where, recall, δ_* is the smallest solution of the equation $(48\delta)^{-1}s^*(\delta)=1$,

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

and therefore

$$e^{(\mathrm{a_f})}_{s^*}\big(\underline{A}u,\Theta_A(\underline{A}u)\big)=\sup_{\delta>\delta_*}\delta^{-2}\mathfrak{E}_{\Theta_A(\underline{A}u),\,\mathrm{a_f}}\big(\underline{A}u(48\delta)^{-1}s^*(\delta)\big),\tag{2.21}$$

$$e^{(\mathrm{b}_\infty)}_{s^*}\big(\underline{B}v,\Theta_B(\underline{B}v)\big)=\sup_{\delta>\delta_*}\delta^{-1}\mathfrak{E}_{\Theta_B(\underline{B}v),\,\mathrm{b}_\infty}\big(\underline{B}v(48\delta)^{-1}s^*(\delta)\big).\tag{2.22}$$

3°b. For any $t\geq 1$ put $\mathfrak{H}^k_1(t,\mathbf{n})=\big\{\mathfrak{h}^{(k)}\in\mathfrak{H}^k_1(\mathbf{n})\colon G_\infty(\mathfrak{h}^{(k)})\leq\underline{G}_{\mathbf{n}}t\big\}$ and note that the following obvious inclusions hold:

$$\Theta_A\big(\underline{A}u\big)\subseteq\mathfrak{H}^k_1\big(u^2\underline{G}_{\mathbf{n}}[\tau]\underline{G}^{-1}_{\mathbf{n}},\mathbf{n}\big)\times\mathfrak{H}^m_{k+1}\times\big[\mathbf{n}_1/\mathbf{n}_2,1\big],\tag{2.23}$$

$$\Theta_B\big(\underline{B}v\big)\subseteq\mathfrak{H}^k_1\big(v\underline{G}_{\mathbf{n}}[\tau]\underline{G}^{-1}_{\mathbf{n}},\mathbf{n}\big)\times\mathfrak{H}^m_{k+1}\times\big[\mathbf{n}_1/\mathbf{n}_2,1\big].\tag{2.24}$$

For any $\varepsilon>0$ denote by $\mathfrak{N}^{(k)}_t(\varepsilon)$ the minimal number of $\varrho^{(k)}_{\mathbf{n}}$-balls of radius ε needed to cover $\mathfrak{H}^k_1(t,\mathbf{n})$, let $\mathfrak{N}_j(\varepsilon)$, $j=k+1,\ldots,m$, be 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 $\mathfrak{H}^k_1(t,\mathbf{n})\times\mathfrak{H}^m_{k+1}\times[1/2,1]$. It is evident that for any given $\epsilon^{(k)}>0$, $\epsilon_j>0$, $j=k+1,\ldots,m$, and $\epsilon>0$ one can construct a net $\big\{\mathrm{h}(\mathbf{i}),\ \mathbf{i}=1,\ldots,\mathbf{I}[\mathbb{H}]\big\}\subset\mathbb{H}$ such that $\forall\mathrm{h}=(\mathfrak{h},y)\in\mathbb{H}\ \ \exists\mathbf{i}\in\{1,\ldots,\mathbf{I}\big[\mathbb{H}\big]\}$

$$\varrho^{(k)}_{\mathbf{n}}(\mathfrak{h}^{(k)},\mathfrak{h}^{(k)}(i))\leq\epsilon^{(k)},\qquad \varrho_j(\mathfrak{h}_j,\mathfrak{h}_j(i))\leq\epsilon_j,\quad j=k+1,\ldots,m,\qquad \mathfrak{w}(y,y(\mathbf{i}))\leq\epsilon,\tag{2.25}$$

$$\mathbf{I}\big[\mathbb{H}\big]\leq\mathfrak{N}(\epsilon/2)\ \mathfrak{N}^{(k)}_t(\epsilon^{(k)}/2)\prod_{j=k+1}^{m}\mathfrak{N}_j(\epsilon_j/2),\quad \forall\mathbb{H}\subseteq\mathfrak{H}^k_1(t,\mathbf{n})\times\mathfrak{H}^m_{k+1}\times[\mathbf{n}_1/\mathbf{n}_2,1].\tag{2.26}$$

Moreover we obtain from (2.19) and (2.20) for any $u,v\geq 1$

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

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

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

we obtain in view of (2.23) and (2.25) with $\mathbb{H} = \Theta_A\big(\underline{A}u\big)$

$$\forall \mathrm{h} \in \Theta_A(\underline{A}u) \quad \exists \mathbf{i} \in \big\{1, \dots, \mathbf{I}\big[\Theta_A(\underline{A}u)\big]\big\} : \mathrm{a_f}\big(\mathrm{h}, \mathrm{h}(\mathbf{i})\big) \le \varsigma. \tag{2.27}$$

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

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

we obtain in view of (2.23) and (2.25) with $\mathbb{H} = \Theta_B\big(\underline{B}v\big)$

$$\forall \mathrm{h} \in \Theta_B\big(\underline{B}v\big) \quad \exists \mathbf{i} \in \big\{1, \dots, \mathbf{I}\big[\Theta_B(\underline{B}v)\big]\big\} : \mathrm{b}_\infty\big(\mathrm{h}, \mathrm{h}(\mathbf{i})\big) \le \varsigma. \tag{2.28}$$

We get from (2.26), (2.27) and (2.28) for any $\varsigma > 0$

$$\mathfrak{E}_{\Theta_A(\underline{A}u),\mathrm{a_f}}(\varsigma) \le \mathfrak{E}_{\mathfrak{H}_1^k(t_1,\mathbf{n}),\varrho_{\mathbf{n}}^{(k)}}\bigg(\frac{\varsigma^2}{4m\mathfrak{a}^2\underline{G}_{\mathbf{n}}[\tau]u^2}\bigg) + \sum_{j=k+1}^{m} \mathfrak{E}_{\mathfrak{H}_j,\varrho_j}\bigg(\frac{\varsigma^2}{4m\mathfrak{a}^2\widetilde{L}_j\big(\underline{G}_{\mathbf{n}}[\tau]u^2\big)}\bigg)$$
$$+ \mathfrak{E}_{[\mathbf{n}_1/\mathbf{n}_2,1],\mathfrak{w}}\Big(2^{-1}\exp\Big\{-\Big(2u\mathfrak{a}\sqrt{\underline{G}_{\mathbf{n}}[\tau]}\varsigma^{-1}\Big)^{1/\boldsymbol{b}}\Big\}\Big), \tag{2.29}$$

$$\mathfrak{E}_{\Theta_B(\underline{B}v),\mathrm{b}_\infty}(\varsigma) \le \mathfrak{E}_{\mathfrak{H}_1^k(t_2,\mathbf{n}),\varrho_{\mathbf{n}}^{(k)}}\bigg(\frac{\varsigma}{4m\mathfrak{b}\underline{G}_{\mathbf{n}}[\tau]v}\bigg) + \sum_{j=k+1}^{m} \mathfrak{E}_{\mathfrak{H}_j,\varrho_j}\bigg(\frac{\varsigma}{4m\mathfrak{b}\widetilde{L}_j\big(\underline{G}_{\mathbf{n}}[\tau]v\big)}\bigg)$$
$$+ \mathfrak{E}_{[\mathbf{n}_1/\mathbf{n}_2,1],\mathfrak{w}}\Big(2^{-1}\exp\Big\{-\big(2v\mathfrak{b}\underline{G}_{\mathbf{n}}[\tau]\varsigma^{-1}\big)^{1/\boldsymbol{b}}\Big\}\Big). \tag{2.30}$$

4°. We get in view of Assumption 3

$$\sum_{j=k+1}^{m} \mathfrak{E}_{\mathfrak{H}_j,\varrho_j}\bigg(\frac{\varsigma^2}{4m\mathfrak{a}^2\widetilde{L}_j\big(\underline{G}_{\mathbf{n}}[\tau]u^2\big)}\bigg) \le N \sum_{j=k+1}^{m} \Big[\log_2\big\{4\mathfrak{a}^2mR\widetilde{L}_j(\underline{G}_{\mathbf{n}}[\tau]u^2)\varsigma^{-2}\big\}\Big]_+, \tag{2.31}$$

$$\sum_{j=k+1}^{m} \mathfrak{E}_{\mathfrak{H}_j,\varrho_j}\bigg(\frac{\varsigma}{4m\mathfrak{b}\widetilde{L}_j\big(\underline{G}_{\mathbf{n}}[\tau]v\big)}\bigg) \le N \sum_{j=k+1}^{m} \Big[\log_2\big\{4\mathfrak{b}mR\widetilde{L}_j(\underline{G}_{\mathbf{n}}[\tau]v)\varsigma^{-1}\big\}\Big]_+. \tag{2.32}$$

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 $\mathfrak{E}_{[\mathbf{n}_1/\mathbf{n}_2,1],\mathfrak{w}}(\varepsilon) \le \log(2/\varepsilon^2)$ for any $\varepsilon \in (0,1]$ and any $\mathbf{n}_2 \le 2\mathbf{n}_1$, we have

$$\mathfrak{E}_{[\mathbf{n}_1/\mathbf{n}_2,1],\mathfrak{w}}\Big(2^{-1}\exp\Big\{-\Big(2u\mathfrak{a}\sqrt{\underline{G}_{\mathbf{n}}[\tau]}\varsigma^{-1}\Big)^{1/\beta}\Big\}\Big) = \chi\Big(2\log(2) + 2\Big(2u\mathfrak{a}\sqrt{\underline{G}_{\mathbf{n}}[\tau]}\varsigma^{-1}\Big)^{\frac{1}{\boldsymbol{b}}}\Big), \tag{2.33}$$

$$\mathfrak{E}_{[\mathbf{n}_1/\mathbf{n}_2,1],\mathfrak{w}}\Big(2^{-1}\exp\Big\{-\big(2v\mathfrak{b}\underline{G}_{\mathbf{n}}[\tau]\varsigma^{-1}\big)^{1/\beta}\Big\}\Big) = \chi\Big(2\log(2) + 2\big(2v\mathfrak{b}\underline{G}_{\mathbf{n}}[\tau]\varsigma^{-1}\big)^{\frac{1}{\boldsymbol{b}}}\Big). \tag{2.34}$$

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 1 (i)

$$\mathfrak{H}_1^k(t,\mathbf{n}) \subseteq \big\{\mathfrak{h}_1 \in \mathfrak{H}_1(\mathbf{n}) : G_{1,\mathbf{n}}(\mathfrak{h}_1) \le t\underline{G}_{1,\mathbf{n}}\big\} \times \cdots \times \big\{\mathfrak{h}_k \in \mathfrak{H}_k(\mathbf{n}) : G_{k,\mathbf{n}}(\mathfrak{h}_k) \le t\underline{G}_{k,\mathbf{n}}\big\}. \tag{2.35}$$

Consider the hyper-rectangle $\mathcal{Z}(t) = [\underline{G}_{1,\mathbf{n}}, t\underline{G}_{1,\mathbf{n}}] \times \cdots \times [\underline{G}_{k,\mathbf{n}}, t\underline{G}_{k,\mathbf{n}}]$, $t \ge 1$, which we equip with the metric

$$\mathfrak{m}^{(k)}(z,z') = \max_{i=1,\dots,k} \big|\log(z_i) - \log(z'_i)\big|, \qquad z, z' \in \mathcal{Z}(t),$$

where z_i, z'_i, $i = 1, \dots, k$, are the coordinates of z, z' respectively. It is easily seen that for any $\varsigma > 0$

$$\mathfrak{E}_{\mathcal{Z}(t),\mathfrak{m}^{(k)}}(\varsigma) \le k\big[\log\log t - \log\log(1+\varsigma)\big]_+ \le k\big(\log(1+\log t) + [1 + \log(1/\varsigma)]_+\big).$$

This yields together with (2.35) in view of the obvious inequality $\mathfrak{E}_{\mathfrak{H}_1^k(t,\mathbf{n}),\varrho_{\mathbf{n}}^{(k)}}(\varsigma) \le \mathfrak{E}_{\mathcal{Z}(t),\mathfrak{m}^{(k)}}(\varsigma/2)$

$$\mathfrak{E}_{\mathfrak{H}_1^k(t,\mathbf{n}),\varrho_{\mathbf{n}}^{(k)}}(\varsigma) \le k\big(\log(1+\log t) + [1 + \log(2/\varsigma)]_+\big). \tag{2.36}$$

We obtain from (2.36)

$$\mathfrak{E}_{\mathfrak{H}_1^k(t_1,\mathbf{n}),\varrho_{\mathbf{n}}^{(k)}}\left(\frac{\varsigma^2}{4m\mathfrak{a}^2\underline{G}_{\mathbf{n}}[\tau]u^2}\right) \le k\Big(\log\left(1+\log t_1\right)+\left[1+\log\left(8m\mathfrak{a}^2\underline{G}_{\mathbf{n}}[\tau]u^2\varsigma^{-2}\right)\right]_+\Big), \qquad (2.37)$$

$$\mathfrak{E}_{\mathfrak{H}_1^k(t_2,\mathbf{n}),\varrho_{\mathbf{n}}^{(k)}}\left(\frac{\varsigma}{4m\mathfrak{b}\underline{G}_{\mathbf{n}}[\tau]v}\right) \le k\Big(\log\left(1+\log t_2\right)+\left[1+\log\left(8m\mathfrak{b}\underline{G}_{\mathbf{n}}[\tau]v\varsigma^{-1}\right)\right]_+\Big). \qquad (2.38)$$

Putting $\widehat{L}_j(z)=z^{-1}\widetilde{L}_j(z)=\max\left\{z^{-1}L_j(z),1\right\}$, we get from (2.21), (2.29), (2.31), (2.33) and (2.37)

$$\begin{aligned}
e_{s^*}^{(\mathrm{a_f})}\big(\underline{A}u,\Theta_A(\underline{A}u)\big) &\le k\delta_*^{-2}\log\big(1+\log(u^2\underline{G}_{\mathbf{n}}[\tau]\underline{G}_{\mathbf{n}}^{-1})\big)+N\delta_*^{-2}\sum_{j=k+1}^{m}\log_2\big\{\widehat{L}_j(\underline{G}_{\mathbf{n}}[\tau]u^2)\big\} \\
&\quad+\sup_{\delta>\delta_*}\delta^{-2}\bigg\{k\bigg[1+\log\bigg(\frac{9216m\delta^2}{[s^*(\delta)]^2}\bigg)\bigg]_+ + N(m-k)\bigg[\log_2\bigg\{\bigg(\frac{4608mR\delta^2}{[s^*(\delta)]^2}\bigg)\bigg\}\bigg]_+\bigg\} \\
&\quad+\chi\Big(2\delta_*^{-2}\log(2)+2\sup_{\delta>\delta_*}\delta^{-2}\big(96\delta/s^*(\delta)\big)^{\frac{1}{b}}\Big) \\
&= k\delta_*^{-2}\log\big(1+\log\left(u^2\underline{G}_{\mathbf{n}}[\tau]\underline{G}_{\mathbf{n}}^{-1}\right)\big)+N\delta_*^{-2}\sum_{j=k+1}^{m}\log_2\big\{\widehat{L}_j(\underline{G}_{\mathbf{n}}[\tau]u^2)\big\} \\
&\quad+C_{N,R,m,k}^{(1)}+\chi\mathbf{a_b},
\end{aligned} \qquad (2.39)$$

where, recall, $\mathbf{a_b}=2\delta_*^{-2}\log(2)+2\sup_{\delta>\delta_*}(\delta^2\wedge\delta)^{-1}\big(96\delta/s^*(\delta)\big)^{\frac{1}{b}}$. Note that $\mathbf{a_b}<\infty$ since $\boldsymbol{b}>1$.

Repeating these computations we get from (2.22), (2.30), (2.32), (2.34) and (2.38)

$$\begin{aligned}
e_{s^*}^{(\mathrm{b}\infty)}\big(\underline{B}v,\Theta_B(\underline{B}v)\big) &\le k\delta_*^{-1}\log\big(1+\log\left(v\underline{G}_{\mathbf{n}}[\tau]\underline{G}_{\mathbf{n}}^{-1}\right)\big) \\
&\quad+N\delta_*^{-1}\sum_{j=k+1}^{m}\log_2\big\{\widehat{L}_j(\underline{G}_{\mathbf{n}}[\tau]v)\big\}+C_{N,R,m,k}^{(2)}+\chi\mathbf{a_b}.
\end{aligned} \qquad (2.40)$$

We deduce from (2.39) and (2.40) that $\widetilde{\mathcal{E}}_{\vec{s}}$, $\vec{s}=(s^*,s^*)$, is bounded from above by the function

$$\begin{aligned}
\mathcal{E}(u,v) &\le k\delta_*^{-2}\log\Big\{\big(1+\log\left(u^2\underline{G}_{\mathbf{n}}[\tau]\underline{G}_{\mathbf{n}}^{-1}\right)\big)\big(1+\log\left(v\underline{G}_{\mathbf{n}}[\tau]\underline{G}_{\infty}^{-1}\right)\big)\Big\} \\
&\quad+N\delta_*^{-2}\sum_{j=k+1}^{m}\log_2\Big[\big\{\widehat{L}_j(\underline{G}_{\mathbf{n}}[\tau]u^2)\big\}\big\{\widehat{L}_j(\underline{G}_{\mathbf{n}}[\tau]v)\big\}\Big]+C_{N,R,m,k}.
\end{aligned} \qquad (2.41)$$

Here we have used that $\delta_*<1$. We note that (2.41) implies in particular Assumption 3 in Lepski (2013) and therefore Proposition 2, Lepski (2013), is applicable with $\Theta=\mathrm{H}(\tau)$.

5°. To apply Proposition 2, Lepski (2013), on $\Theta=\mathrm{H}(\tau)$ we choose $\varepsilon=\sqrt{2}-1$ and bound from above the quantities

$$\begin{aligned}
P_{\sqrt{2}-1}(\mathrm{h}) &:= 4\big[\sqrt{2}-1\big]^{-2}\mathcal{E}\Big(\sqrt{2\underline{G}_{\mathbf{n}}^{-1}[\tau]G_\infty(\mathfrak{h}^{(k)})},\sqrt{2}\underline{G}_{\mathbf{n}}^{-1}[\tau]G_\infty(\mathfrak{h}^{(k)})\Big) \\
&\quad+2\ell\Big(\sqrt{2\underline{G}_{\mathbf{n}}^{-1}[\tau]G_\infty(\mathfrak{h}^{(k)})}\Big)+2\ell\Big(\sqrt{2}\underline{G}_{\mathbf{n}}^{-1}[\tau]G_\infty(\mathfrak{h}^{(k)})\Big),
\end{aligned}$$

$$\begin{aligned}
M_{\sqrt{2}-1,q}(\mathrm{h}) &:= 8\big[\sqrt{2}-1\big]^{-2}\mathcal{E}\Big(\sqrt{2\underline{G}_{\mathbf{n}}^{-1}[\tau]G_\infty^{-1}(\mathfrak{h}^{(k)})},\sqrt{2}\underline{G}_{\mathbf{n}}^{-1}[\tau]G_\infty(\mathfrak{h}^{(k)})\Big) \\
&\quad+2\big(\sqrt{2}-1+q\big)\log\Big(\sqrt{2\underline{G}_{\mathbf{n}}^{-1}[\tau]G_\infty(\mathfrak{h}^{(k)})}\sqrt{2}\underline{G}_{\mathbf{n}}^{-1}[\tau]G_\infty(\mathfrak{h}^{(k)})\Big),
\end{aligned}$$

where, recall, $\ell(u)=\log\big\{1+\log(u)\big\}+2\log\big\{1+\log\{1+\log(u)\}\big\}$.

Taking into account that $\ell(u) \le 3\log\{1+\log(u)\}$, $u \ge 1$, $\big[\sqrt{2}-1\big]^{-2} \le 9$ and that $\underline{G}_{\mathbf{n}}[\tau] \ge \underline{G}_{\mathbf{n}}$ for any τ, we obtain from (2.41)

$$P_{\sqrt{2}-1}(\mathrm{h}) \le [72k\delta_*^{-2} + 12]\log\big\{1+\log\big(2G_\infty\big(\mathfrak{h}^{(k)}\big)\underline{G}_{\mathbf{n}}^{-1}\big)\big\} + 72N\delta_*^{-2}\sum_{j=k+1}^{m}\log_2\big\{\widehat{L}_j\big(2G_\infty(\mathfrak{h}^{(k)})\big)\big\} + 36C_{N,R,m,k} =: 2P(\mathfrak{h}^{(k)}),$$

$$M_{\sqrt{2}-1,q}(\mathrm{h}) \le [144k\delta_*^{-2} + 3(1+q)]\log\big(2G_\infty(\mathfrak{h}^{(k)})\underline{G}_{\mathbf{n}}^{-1}\big) + 144N\delta_*^{-2}\sum_{j=k+1}^{m}\log_2\big\{\widehat{L}_j\big(2G_\infty(\mathfrak{h}^{(k)})\big)\big\} + 72C_{N,R,m,k} =: 2M_q(\mathfrak{h}^{(k)}).$$

We remark that P and M_q are independent of τ and y.

Put for any $z \ge 0$ and any $\mathrm{h} \in \mathrm{H}(\tau)$

$$\check{\mathrm{V}}_\tau^{(z)}(\mathfrak{h}^{(k)}) = 2\sqrt{2}\mathfrak{a}\sqrt{G_\infty(\mathfrak{h}^{(k)})\big[P(\mathfrak{h}^{(k)})+z\big]} + 4\mathfrak{b}G_\infty(\mathfrak{h}^{(k)})\big[P(\mathfrak{h}^{(k)})+z\big],$$

$$\check{\mathrm{U}}_\tau^{(z,q)}(\mathfrak{h}^{(k)}) = 2\sqrt{2}\mathfrak{a}\sqrt{G_\infty(\mathfrak{h}^{(k)})\big[M_q(\mathfrak{h}^{(k)})+z\big]} + 4\mathfrak{b}G_\infty(\mathfrak{h}^{(k)})\big[M_q(\mathfrak{h}^{(k)})+z\big],$$

where, recall, $\mathfrak{a} = 2\sqrt{\tau}(\mathbf{n}_1)^{-1/2}\big(\sqrt{C_D}\vee[\chi \boldsymbol{c}_{\mathfrak{b}}]\big)$, $\mathfrak{b} = \frac{4(C_D\vee 8e)\log^\beta(\mathbf{n}_2)}{3\mathbf{n}_1}$.

We conclude that Proposition 2, Lepski (2013) is applicable with $\check{\mathrm{V}}_\tau^{(z)}$ and $\check{\mathrm{U}}_\tau^{(z,q)}$. Put for any $n \in \{\mathbf{n}_1, \mathbf{n}_1+1, \dots, \mathbf{n}_2\}$

$$\mathfrak{a}(n) = 2\sqrt{2\tau}(n)^{-1/2}\big(\sqrt{C_D}\vee[\chi\boldsymbol{c}_{\mathfrak{b}}]\big), \qquad \mathfrak{b}(n) = \frac{8\big(C_D\vee 8e\big)\log^\beta(2n)}{3n}$$

and define

$$\mathrm{V}_\tau^{(z)}\big(n,\mathfrak{h}^{(k)}\big) = 2\sqrt{2}\mathfrak{a}(n)\sqrt{G_\infty\big(\mathfrak{h}^{(k)}\big)\big[P(\mathfrak{h}^{(k)})+z\big]} + 4\mathfrak{b}(n)G_\infty\big(\mathfrak{h}^{(k)}\big)\big[P(\mathfrak{h}^{(k)})+z\big],$$

$$\mathrm{U}_\tau^{(z,q)}\big(n,\mathfrak{h}^{(k)}\big) = 2\sqrt{2}\mathfrak{a}(n)\sqrt{G_\infty\big(\mathfrak{h}^{(k)}\big)\big[M_q(\mathfrak{h}^{(k)})+z\big]} + 4\mathfrak{b}(n)G_\infty\big(\mathfrak{h}^{(k)}\big)\big[M_q(\mathfrak{h}^{(k)})+z\big].$$

It is easily seen that $\mathfrak{a}(n) \ge \mathfrak{a}$, $\mathfrak{b}(n) \ge \mathfrak{b}$ for any $n \in \big\{\mathbf{n}_1,\dots,\mathbf{n}_2\big\}$ since $\mathbf{n}_2 \le 2\mathbf{n}_1$. Therefore

$$\mathrm{V}_\tau^{(z)}\big(n,\mathfrak{h}^{(k)}\big) \ge \check{\mathrm{V}}_\tau^{(z)}\big(\mathfrak{h}^{(k)}\big), \qquad \mathrm{U}_\tau^{(z,q)}\big(n,\mathfrak{h}^{(k)}\big) \ge \check{\mathrm{U}}_\tau^{(z,q)}\big(\mathfrak{h}^{(k)}\big).$$

It remains to recall that $\xi_{\mathfrak{h}}(n) = \boldsymbol{\xi}\big(n/\mathbf{n}_2,\mathfrak{h}\big)$ for any $n \in \{\mathbf{n}_1,\mathbf{n}_1+1,\dots,\mathbf{n}_2\}$ and any $\mathfrak{h} \in \mathfrak{H}$. All what was said above allows us to assert that Proposition 2, Lepski (2013), is applicable to $|\xi_{\mathfrak{h}}(n)|$ on $\mathbf{H}(\tau) := \{\mathbf{n}_1,\mathbf{n}_1+1,\dots,\mathbf{n}_2\}\times\mathfrak{H}(\mathbf{n},\tau)$ 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 $z \ge 1$ and any $q \ge 1$

$$\mathbb{P}_{\mathrm{f}}\Big\{\sup_{\mathbf{h}\in\mathbf{H}(\tau)}\big[|\xi_{\mathfrak{h}}(n)| - \mathrm{V}_\tau^{(z)}\big(n,\mathfrak{h}^{(k)}\big)\big] \ge 0\Big\} \le 4\Big[1+\big[\log\{1+2^{-1}\log 2\}\big]^{-2}\Big]^2\exp\{-z\}, \tag{2.42}$$

$$\mathbb{E}_{\mathrm{f}}\Big\{\sup_{\mathbf{h}\in\mathbf{H}(\tau)}\big[|\xi_{\mathfrak{h}}(n)| - \mathrm{U}_\tau^{(z,q)}\big(n,\mathfrak{h}^{(k)}\big)\big]Big\}_+^q \le 2^{(5q/2)+3}3^{q+4}\Gamma(q+1)\,\big[\underline{A}\vee\underline{B}\big]^q\exp\{-z\}, \tag{2.43}$$

where, recall, $\underline{A} = \mathfrak{a}\sqrt{\underline{G}_{\mathbf{n}}[\tau]}$ 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°. 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 $\widehat{\mathfrak{H}}(r) = \mathfrak{H}\big(\mathbf{n},\tau_r\big)\setminus\mathfrak{H}\big(\mathbf{n},\tau_{r-1}\big)$, $\widehat{\mathfrak{H}}(0) = \mathfrak{H}\big(\mathbf{n},\tau_0\big)$ and let $\widehat{\mathbf{H}}(r) := \{\mathbf{n}_1,\mathbf{n}_1+1,\dots,\mathbf{n}_2\}\times\widehat{\mathfrak{H}}(r)$.

Probability bound. For any $u \geq 1$ put $z_r(u) = u + 2\log\big(1 + |r - \mathbf{r}|\big)$ and remark that

$$z_r(u) = \begin{cases} u + 2\log\left(|\log(\tau_{r-1})|\right), & r \leq \mathbf{r}, \\ u + 2\log\left(1 + |\log(\tau_r)|\right), & r \geq \mathbf{r}. \end{cases}$$

We have for any $r \in \mathbb{N}$ and any $\mathfrak{h} \in \widehat{\mathfrak{H}}(r)$

$$\tau_0 = F_{\mathbf{n}_2,\mathbf{r}}(\mathfrak{h}) \Rightarrow z_0(u) = u + 2\log\big\{1 + |\log(F_{\mathbf{n}_2,\mathbf{r}}(\mathfrak{h}))|\big\},$$

$$\tau_{r-1} \leq F_{\mathbf{n}_2}(\mathfrak{h}) = F_{\mathbf{n}_2,\mathbf{r}}(\mathfrak{h}) \Rightarrow z_r(u) \leq u + 2\log\big\{|\log(F_{\mathbf{n}_2,\mathbf{r}}(\mathfrak{h}))|\big\}, \quad 1 \leq r \leq \mathbf{r} - 1,$$

$$\tau_r \geq F_{\mathbf{n}_2}(\mathfrak{h}) = F_{\mathbf{n}_2,\mathbf{r}}(\mathfrak{h}) \Rightarrow z_r(u) \leq u + 2\log\big\{1 + |\log(F_{\mathbf{n}_2,\mathbf{r}}(\mathfrak{h}))|\big\}, \quad r \geq \mathbf{r}.$$

Hence we have for any $r \in \mathbb{N}$

$$z_r(u) \leq u + 2\log\big\{1 + |\log(F_{\mathbf{n}_2,\mathbf{r}}(\mathfrak{h}))|\big\}, \qquad \forall \mathfrak{h} \in \widehat{\mathfrak{H}}(r), \tag{2.44}$$

which yields for any $r \in \mathbb{N}$

$$\mathrm{V}_{\tau_r}^{(z_r(u))}\big(n, \mathfrak{h}^{(k)}\big) \leq \mathcal{V}_{\mathbf{r}}^{(u)}(n, \mathfrak{h}), \quad \forall (n, \mathfrak{h}) \in \widehat{\mathbf{H}}(r). \tag{2.45}$$

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 (2.45), the inclusion $\widehat{\mathbf{H}}(r) \subseteq \mathbf{H}(\tau_r)$ and applying (2.42) with $\tau = \tau_r$,

$$\mathbb{P}_{\mathrm{f}}\Big\{\sup_{(n,\mathfrak{h}) \in \widehat{\mathbf{H}}(r)} \big[|\xi_{\mathfrak{h}}(n)| - \mathcal{V}_{\mathbf{r}}^{(u)}(n, \mathfrak{h})\big] \geq 0\Big\} \leq \frac{4\big[1 + [\log\{1 + 2^{-1}\log 2\}]^{-2}\big]^2 \exp\{-u\}}{\big[1 + |r - \mathbf{r}|\big]^2}. \tag{2.46}$$

Since obviously $\widetilde{\mathbf{N}} \times \mathfrak{H}(\mathbf{n}) = \bigcup_{r=0}^{\infty} \widehat{\mathbf{H}}(r)$, summing up the right-hand sides of (2.46) over r, we come to the first assertion of the theorem. Here we have also used that $16\big[1 + [\log\{1 + 2^{-1}\log 2\}]^{-2}\big]^2 \leq 2419$ and the fact that $\widetilde{\mathfrak{H}}(n) \subseteq \mathfrak{H}(\mathbf{n})$ for any $n \in \widetilde{\mathbf{N}}$ in view of Assumption 2 and the definition of the number $\mathbf{n}$.

Moment's bound. For any $u \geq 1$ put

$$z_r(u) = u + 2\log\big(1 + |r - \mathbf{r}|\big) + q\log\big(\underline{G}_{\mathbf{n}}[\tau_r]\underline{G}_{\mathbf{n}}^{-1}\big).$$

Similarly to (2.44) we have for any $r \in \mathbb{N}$ and any $\mathfrak{h} \in \widehat{\mathfrak{H}}(r)$

$$z_r(u) \leq u + 2\log\big\{1 + |\log(F_{\mathbf{n}_2,r}(\mathfrak{h}))|\big\} + q\log\big(\underline{G}_{\mathbf{n}}[\tau_r]\underline{G}_{\mathbf{n}}^{-1}\big).$$

Moreover, for any $r \in \mathbb{N}$ by definition

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

and, therefore, for any $\mathfrak{h} \in \widehat{\mathfrak{H}}(r)$

$$z_r(u) \leq u + 2\log\big\{1 + |\log(F_{\mathbf{n}_2,r}(\mathfrak{h}))|\big\} + q\log\big\{G_\infty(\mathfrak{h}^{(k)})\underline{G}_{\mathbf{n}}^{-1}\big\}.$$

Similarly to (2.45), this yields for any $r \in \mathbb{N}$

$$\mathrm{U}_{\tau_r}^{(z_r(u),q)}(n, \mathfrak{h}^{(k)}) \leq \mathcal{U}_{\mathbf{r}}^{(u,q)}(n, \mathfrak{h}), \qquad \forall (n, \mathfrak{h}) \in \widehat{\mathbf{H}}(r). \tag{2.47}$$

Note that for any $r \in \mathbb{N}$

$$\underline{A} \vee \underline{B} \leq 2C_{D,\boldsymbol{b}}\Big[\sqrt{(\mathbf{n}_1)^{-1}F_{\mathbf{n}_2}\underline{G}_{\mathbf{n}}} \vee \big((\mathbf{n}_1)^{-1}\log^{\beta}(\mathbf{n}_2)\underline{G}_{\mathbf{n}}\big)\Big]\big[\underline{G}_{\mathbf{n}}[\tau_r]\underline{G}_{\mathbf{n}}^{-1}\big],$$

where $C_{D,\boldsymbol{b}} = \big(\sqrt{2C_D} \vee [\gamma \boldsymbol{c}_{\boldsymbol{b}}]\big) \vee \big[(2/3)\big(C_D \vee 8e\big)\big]$. We get from (2.43) and (2.47), similarly to (2.46),

$$\mathbb{E}_{\mathrm{f}}\Big\{\sup_{(n,\mathfrak{h}) \in \widehat{\mathbf{H}}(r)} \big[|\xi_{\mathfrak{h}}(n)| - \mathcal{U}_{\mathbf{r}}^{(u,q)}(n, \mathfrak{h})\big]\Big\}_+^q \leq \frac{K_q\big[\sqrt{(\mathbf{n}_1)^{-1}F_{\mathbf{n}_2}\underline{G}_{\mathbf{n}}} \vee \big((\mathbf{n}_1)^{-1}\log^{\beta}(\mathbf{n}_2)\underline{G}_{\mathbf{n}}\big)\big]^q e^{-u}}{\big[1 + |r - \mathbf{r}|\big]^2},$$

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

Summing up the right-hand sides of the last inequality over r we come to the second assertion of the theorem. □

2.2. Proof of Theorem 2

For any $l \in \mathbb{N}^*$ set $n_l = \mathbf{j}2^l$, $\mathbf{N}_l = \{n_l, n_l + 1, \dots, n_{l+1}\}$ and let

$$\zeta_{\mathbf{j}} = \sup_{n \geq \mathbf{j}} \sup_{\mathfrak{h}^{(k)} \in \overline{\mathfrak{H}}_1^k(n,a)} \left[\frac{\sqrt{n}\, \eta_{\mathfrak{h}^{(k)}}(n)}{\sqrt{G_\infty\big(\mathfrak{h}^{(k)}\big) \log\big(1 + \log(n)\big)}} \right].$$

We obviously have

$$\mathbb{P}_f\{\zeta_{\mathbf{j}} \geq \Upsilon\} \leq \sum_{l=1}^{\infty} \mathbb{P}_f\left\{ \sup_{n \in \mathbf{N}_l} \sup_{\mathfrak{h}^{(k)} \in \overline{\mathfrak{H}}_1^k(n,a)} \left[\frac{\sqrt{n}\, \eta_{\mathfrak{h}^{(k)}}(n)}{\sqrt{G_\infty\big(\mathfrak{h}^{(k)}\big) \log\big(1 + \log(n)\big)}} \right] \geq \Upsilon \right\}$$

$$= \sum_{l=1}^{\infty} \mathbb{P}_f\left\{ \sup_{n \in \mathbf{N}_l} \sup_{\mathfrak{h}^{(k)} \in \overline{\mathfrak{H}}_1^k(n,a)} \left[\eta_{\mathfrak{h}^{(k)}}(n) - \Upsilon \sqrt{n^{-1} G_\infty\big(\mathfrak{h}^{(k)}\big) \log\big(1 + \log(n)\big)} \right] > 0 \right\}.$$

Let $l \in \mathbb{N}^*$ be fixed and later on $\boldsymbol{\Upsilon}_r$, $r = 1, 2, 3$, denote constants independent of l and n.

Note that in view of (1.11), (1.12) and (1.13) for any $n \in \mathbf{N}_l$

$$\mathcal{V}_0^{(2\log(1+\log(n_l)))}(n, \mathfrak{h}) \leq \lambda_1 \sqrt{(\mathbf{F}n^{-1}) G_\infty\big(\mathfrak{h}^{(k)}\big(P_n + 2\log\{1 + |\log(\mathbf{F})|\} + 2\log(1 + \log(n))\big)\big)}$$
$$+ \lambda_2 (n^{-1} \log^{\boldsymbol{b}}(n)) G_\infty(\mathfrak{h}^{(k)})\big(P_n + 2\log\{1 + |\log(\mathbf{F})|\} + 2\log(1 + \log(n))\big),$$

where we have put

$$P_n = (36k\delta_*^{-2} + 6)\log\big(1 + \mathfrak{b}\log(2n)\big) + 36N\delta_*^{-2}\mathfrak{a}\log\big(1 + \log(2n^{\mathfrak{b}}\mathfrak{c})\big) + 18C_{N,R,m,k}(\boldsymbol{b}).$$

Hence, for any $n \in \mathbf{N}_l$ and any $\mathfrak{h} \in \widetilde{\mathfrak{H}}(n)$,

$$\mathcal{V}_0^{(2\log(1+\log(n_l)))}(n, \mathfrak{h}) \leq \Upsilon_1 \sqrt{\frac{G_\infty\big(\mathfrak{h}^{(k)}\big)\log\big(1 + \log(n)\big)}{n}} + \Upsilon_2 \left[\frac{G_\infty\big(\mathfrak{h}^{(k)}\big)\log^{\boldsymbol{b}}(n)\log\big(1 + \log(n)\big)}{n} \right].$$

Since $\boldsymbol{b} > 1$ can be chosen arbitrarily and $a > 2$, let $1 < \boldsymbol{b} < a/2$. This yields for any $n \geq 3$ and any $\mathfrak{h}^{(k)} \in \overline{\mathfrak{H}}_1^k(n, a)$

$$\frac{G_\infty\big(\mathfrak{h}^{(k)}\big)\log^{\boldsymbol{b}}(n)\log\big(1 + \log(n)\big)}{n} \leq \Upsilon_3 \sqrt{\frac{G_\infty\big(\mathfrak{h}^{(k)}\big)\log\big(1 + \log(n)\big)}{n}}$$

and therefore putting $\Upsilon = \Upsilon_1 + \Upsilon_2\Upsilon_3$ we get for any $n \in \mathbf{N}_l$

$$\mathcal{V}_0^{(2\log(1+\log(n_l)))}(n, \mathfrak{h}) \leq \Upsilon \sqrt{\frac{G_\infty\big(\mathfrak{h}^{(k)}\big)\log\big(1 + \log(n)\big)}{n}}.$$

Noting that the right-hand side of this 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 $u = 2\log(1 + \log(n_l))$ we have

$$\mathbb{P}_f\{\zeta_{\mathbf{j}} \geq \Upsilon\} \leq 2419 \sum_{l=1}^{\infty} \big(l + \log(\mathbf{j})\big)^{-2} \leq \frac{2419}{\log(\mathbf{j})}.$$

□

REFERENCES

1. O. Lepski, "Upper Functions for Positive Random Functionals. I. General Setting and Gaussian Random Functions", Math. Methods Statist. **22** (1), 1–27 (2013).
2. C. J. Stone, "Additive Regression and Other Nonparametric Models", Ann. Statist. **13** (2), 689–705 (1985).