Chapter 12
Uniform-in-Bandwidth Functional Limit Laws for Multivariate Empirical Processes

Paul Deheuvels

Abstract We provide uniform-in-bandwidth functional limit laws for multivariate local empirical processes. Statistical applications to kernel density estimation are given to motivate these results.

Keywords Functional limit laws · Kernel density estimation · Weak laws

AMS 2000 Subject Classification Primary 60F15, 60F17; Secondary 60G07

12.1 Introduction and Motivation

We establish uniform-in-bandwidth functional limit laws for local empirical processes in $\mathbb{R}^d$. Our main result, stated in Theorem 12.2.1, is motivated by statistical applications presented in Theorem 12.1.1. Let $\mathbf{X}^* = (\mathbf{X}, Y) \in \mathbb{R}^{d+1}$, with $\mathbf{X} := (X(1), \dots, X(d)) \in \mathbb{R}^d$ and $Y \in \mathbb{R}$, denote a random vector [rv], with continuous density $g_{\mathbf{X},Y}(\cdot,\cdot)$ on $\mathbb{R}^{d+1} = \mathbb{R}^d \times \mathbb{R}$, and support in $\mathbf{J} \times L$, where $\mathbf{J}$ and L are bounded open subsets of $\mathbb{R}^d$ and $\mathbb{R}$, respectively. Under these assumptions, the marginal density $f(\cdot)$ of $\mathbf{X}$ is continuous on $\mathbb{R}^d$, with $f(\mathbf{x}) = 0$ for $\mathbf{x} \notin \mathbf{J}$, and

$$f(\mathbf{x}) := \int_L g_{\mathbf{X},Y}(\mathbf{x}, y)\mathrm{d}y \quad \text{for} \quad \mathbf{x} \in \mathbb{R}^d. \tag{12.1}$$

Let $\mathcal{K}$ denote a family of *kernels* on $\mathbb{R}^d$, namely, of mappings $\mathbf{K} : \mathbb{R}^d \to \mathbb{R}$, fulfilling conditions $(K.1)$–$(K.4)$ below. For $\mathbf{u} := (u_1, \dots, u_d) \in \mathbb{R}^d$ and $\mathbf{v} := (v_1, \dots, v_d) \in \mathbb{R}^d$, we write $\mathbf{u} \le \mathbf{v}$ when $u_j \le v_j$ for $j = 1, \dots, d$. When this condition holds, we set $(\mathbf{u}, \mathbf{v}] := \prod_{j=1}^d (u_j, v_j]$, and define likewise, with obvious notation, $[\mathbf{u}, \mathbf{v}]$ and $(\mathbf{u}, \mathbf{v})$. In general, by an *interval* in $[r, s]^d$ will be

P. Deheuvels (✉)
LPSM, Sorbonne Université, Bourg-la-Reine, France
e-mail: paul.deheuvels@upmc.fr

N. Gozlan et al. (eds.), *High Dimensional Probability VIII*,
Progress in Probability 74, https://doi.org/10.1007/978-3-030-26391-1_12

meant a product of d subintervals of $[r, s]$. We set $\mathbf{0} := (0, \ldots, 0) \in \mathbb{R}^d$ and $\mathbf{1} := (1, \ldots, 1) \in \mathbb{R}^d$, and adopt a similar notation for $\boldsymbol{\infty} := (\infty, \ldots, \infty)$.

$(K.1)$ There exist an $A < \infty$, such that, for each $\mathbf{K} \in \mathcal{K}$, $\mathbf{K}(\mathbf{t}) = 0$ when $|\mathbf{t}| \geq A$ (with $|\cdot|$ denoting the Euclidian norm in $\mathbb{R}^d$);
$(K.2)$ There exists a $B < \infty$ such that each $\mathbf{K} \in \mathcal{K}$ has a Hardy-Krause variation $\mathcal{V}_{\mathrm{HK}}(\mathbf{K})$ in $\mathbb{R}^d$, fulfilling $\mathcal{V}_{\mathrm{HK}}(\mathbf{K}) \leq B$ (see Sect. 12.2.3 below for details);
$(K.3)$ Each $\mathbf{K}(\mathbf{t}) \in \mathcal{K}$ is a right-continuous function of $\mathbf{t} = (t_1, \ldots, t_d)$;
$(K.4)$ For all $\mathbf{K} \in \mathcal{K}$, $\int_{\mathbb{R}^d} \mathbf{K}(\mathbf{t})\mathrm{d}\mathbf{t} = 1$ (where $\mathrm{d}\mathbf{t}$ denotes Lebesgue measure).

Let $\psi : \mathbb{R} \to \mathbb{R}$ denote a right-continuous function of bounded variation $\|\mathrm{d}\psi\|_L$ on L. We will denote by $\|\mathrm{d}\psi\| := \|\mathrm{d}\psi\|_{\mathbb{R}}$ the total variation of ψ on $\mathbb{R}$. In most of our examples, ψ will be a linear combination of the identity mapping, $\mathcal{I}(y) = y$, and of the unit function, $\mathbb{I}(y) = 1$, for $y \in \mathbb{R}$. Consider a sequence of independent and identically distributed [iid] random replicæ $\mathbf{X}_i^* = (\mathbf{X}_i, Y_i)$, $i = 1, 2, \ldots$, of $\mathbf{X}^* = (\mathbf{X}, Y)$. Introduce the *kernel statistic* indexed by $\mathbf{K} \in \mathcal{K}$,

$$f_{\psi;n;h;\mathbf{K}}(\mathbf{x}) := (nh)^{-1} \sum_{i=1}^{n} \psi(Y_i)\mathbf{K}\left(h^{-1/d}\left(\mathbf{X}_i - \mathbf{x}\right)\right) \quad \text{for} \quad \mathbf{x} \in \mathbb{R}^d, \tag{12.2}$$

where $h > 0$ is a *bandwidth* parameter. In particular, $f_{n;h;\mathbf{K}}(\mathbf{x}) := f_{\mathbb{I};n;h;\mathbf{K}}(\mathbf{x})$ is the Parzen-Rosenblatt [29, 30] kernel estimator of $f(\mathbf{x})$, which, under $(K.1)$–$(K.4)$, fulfills $\int_{\mathbb{R}^d} f_{\mathbb{I};n;h;\mathbf{K}}(\mathbf{x})\mathrm{d}\mathbf{x} = 1$.

Let $\mathbf{I} := \prod_{j=1}^{d}[u_j, v_j] \subset \mathbf{J}$ with $-\infty < u_j < v_j < \infty$ for $j = 1, \ldots, d$, be such that $f(\mathbf{x}) > 0$ for all $\mathbf{x} \in \mathbf{I}$. The conditional expectation (or *regression*) of $\psi(Y)$, given that $\mathbf{X} = \mathbf{x}$, is continuous over $\mathbf{x} \in \mathbf{I}$, and defined by

$$\begin{aligned} m_\psi(\mathbf{x}) &:= \mathbb{E}(\psi(Y)|\mathbf{X} = \mathbf{x}) = \frac{f_\psi(\mathbf{x})}{f(\mathbf{x})} = \frac{f_\psi(\mathbf{x})}{f_{\mathbb{I}}(\mathbf{x})} \\ &= \frac{1}{f(\mathbf{x})} \int_L \psi(y) g_{\mathbf{X},Y}(\mathbf{x}, y)\mathrm{d}y \quad \text{for} \quad \mathbf{x} \in \mathbf{I}, \end{aligned} \tag{12.3}$$

where, for each measurable $\phi : \mathbb{R} \to \mathbb{R}$, rendering meaningful the expression below, we set

$$f_\phi(\mathbf{x}) := \int_L \phi(y) g_{\mathbf{X},Y}(\mathbf{x}, y)\mathrm{d}y \quad \text{for} \quad \mathbf{x} \in \mathbf{I}. \tag{12.4}$$

In view of (12.1) and (12.4), for $\phi = \mathbb{I}$, (12.4) reduces to $f_{\mathbb{I}}(\mathbf{x}) = f(\mathbf{x})$. Under the above assumptions, the conditional variance of $\psi(Y)$, given $\mathbf{X} = \mathbf{x}$, is continuous over $\mathbf{x} \in \mathbf{I}$, and given by

$$\begin{aligned} \sigma_\psi^2(\mathbf{x}) &:= \mathrm{Var}\left(\psi(Y)|\mathbf{X} = \mathbf{x}\right) \\ &= \frac{1}{f(\mathbf{x})} \int_L \left(\psi(y) - m_\psi(\mathbf{x})\right)^2 g_{\mathbf{X},Y}(\mathbf{x}, y)\mathrm{d}y \quad \text{for} \quad \mathbf{x} \in \mathbf{I}. \end{aligned} \tag{12.5}$$

The kernel estimator of the *regression function* $m_\psi(\mathbf{x}) = \mathbb{E}(\psi(Y)|\mathbf{X} = \mathbf{x})$ [25, 40], is then defined, for $\mathbf{x} \in \mathbf{I}$, by

$$m_{\psi;n;h;\mathbf{K}}(\mathbf{x}) := \begin{cases} \dfrac{f_{\psi;n;h;\mathbf{K}}(\mathbf{x})}{f_{\mathrm{I\!I};n;h;\mathbf{K}}(\mathbf{x})} & \text{when} \quad f_{\mathrm{I\!I};n;h;\mathbf{K}}(\mathbf{x}) > 0, \\ \overline{Y} := n^{-1} \sum_{i=1}^{n} Y_i & \text{when} \quad f_{\mathrm{I\!I};n;h;\mathbf{K}}(\mathbf{x}) \leq 0. \end{cases} \tag{12.6}$$

Introduce, whenever properly defined, the centering factor

$$\widehat{\mathbb{E}}\left(m_{\psi;n;h;\mathbf{K}}(\mathbf{x})\right) := \frac{\mathbb{E}\Big(\psi(Y)\mathbf{K}(h^{-1/d}(\mathbf{X}-\mathbf{x}))\Big)}{\mathbb{E}\Big(\mathbf{K}(h^{-1/d}(\mathbf{X}-\mathbf{x}))\Big)}. \tag{12.7}$$

Remark 12.1.1 Under $(K.1)$–$(K.4)$, for $\mathbf{x} \in \mathbf{I}$, we have $\mathbb{E}\left(f_{n;h;\mathbf{K}}(\mathbf{x})\right) \to f(\mathbf{x})$ and $\widehat{\mathbb{E}}\left(m_{\psi;n;h;\mathbf{K}}(\mathbf{x})\right) \to m_\psi(\mathbf{x})$, as $h \to 0$. (see, e.g., [9]). Thus, in the study of the consistency of $f_{n;h;\mathbf{K}}(\mathbf{x})$ and $m_{\psi;n;h;\mathbf{K}}(\mathbf{x})$, we will limit ourselves to the evaluation of the limiting behavior of the *random components* $f_{n;h;\mathbf{K}}(\mathbf{x}) - \mathbb{E}\left(f_{n;h;\mathbf{K}}(\mathbf{x})\right)$ and $m_{\psi;n;h;\mathbf{K}}(\mathbf{x}) - \widehat{\mathbb{E}}\left(m_{\psi;n;h;\mathbf{K}}(\mathbf{x})\right)$ of the estimators.

Let $0 < a_n \leq b_n$, for $n \geq 1$, be sequences of real constants, and set $\log_+ x := \log(x \vee e)$ for $x \in \mathbb{R}$. We have the following theorem.

Theorem 12.1.1 *Assume $(K.1)$–$(K.4)$, and let $0 < a_n \leq b_n$ be such that, as $n \to \infty$,*

$$na_n/\log n \to \infty \quad \textit{and} \quad b_n \to 0. \tag{12.8}$$

Then, with $\mathcal{H}_n := [a_n, b_n]$, we have, as $n \to \infty$,

$$\sup_{\mathbf{K}\in\mathcal{K}} \left(\sup_{h\in\mathcal{H}_n} \left| \left\{ \frac{nh}{2\log_+(1/h)} \right\}^{1/2} \sup_{\mathbf{x}\in\mathbf{I}} \pm \Big\{ f_{n;h;\mathbf{K}}(\mathbf{x}) - \mathbb{E}\left(f_{n;h;\mathbf{K}}(\mathbf{x})\right) \Big\} - \left\{ \sup_{\mathbf{x}\in\mathbf{I}} f(\mathbf{x}) \int_{\mathbb{R}^d} \mathbf{K}(\mathbf{t})^2 d\mathbf{t} \right\}^{1/2} \right| \right) = o_{\mathbb{P}}(1), \tag{12.9}$$

and

$$\sup_{\mathbf{K}\in\mathcal{K}} \left(\sup_{h\in\mathcal{H}_n} \left| \left\{ \frac{nh}{2\log_+(1/h)} \right\}^{1/2} \sup_{\mathbf{x}\in\mathbf{I}} \pm \Big\{ m_{\psi;n;h;\mathbf{K}}(\mathbf{x}) - \widehat{\mathbb{E}}\left(m_{\psi;n;h;\mathbf{K}}(\mathbf{x})\right) \Big\} - \left\{ \sup_{\mathbf{x}\in\mathbf{I}} \frac{\sigma_\psi^2(\mathbf{x})}{f(\mathbf{x})} \int_{\mathbb{R}^d} \mathbf{K}(\mathbf{t})^2 d\mathbf{t} \right\}^{1/2} \right| \right) = o_{\mathbb{P}}(1). \tag{12.10}$$

Remark 12.1.2

1°) When $\mathcal{K} = \{\mathbf{K}\}$ and $d = 1$, (12.9) in Theorem 12.1.1 reduces to Theorem 2 of Deheuvels and Ouadah [10]. This property does not hold for an arbitrary $f(\cdot)$, when (12.8) is not fulfilled (see Remark 1 in [10]).

2°) By Theorem 12.1.1, taken with $\mathcal{K} = \{\mathbf{K}\}$ and $h_n := a_n = b_n$, the condition

$$h_n \to 0 \quad \text{and} \quad nh_n/\log n \to \infty, \tag{12.11}$$

implies that, as $n \to \infty$,

$$\left\{\frac{nh_n}{2\log_+(1/h_n)}\right\}^{1/2} \sup_{\mathbf{x}\in\mathbf{I}} \pm \left\{f_{n;h_n;\mathbf{K}}(\mathbf{x}) - \mathbb{E}\left(f_{n;h_n;\mathbf{K}}(\mathbf{x})\right)\right\} \tag{12.12}$$
$$\overset{\mathbb{P}}{\to} \left\{\sup_{\mathbf{x}\in\mathbf{I}} f(\mathbf{x}) \int_{\mathbb{R}^d} \mathbf{K}(\mathbf{t})^2 d\mathbf{t}\right\}^{1/2},$$

and

$$\left\{\frac{nh_n}{2\log_+(1/h_n)}\right\}^{1/2} \sup_{\mathbf{x}\in\mathbf{I}} \pm \{m_{\psi;n;h_n;\mathbf{K}}(\mathbf{x}) - \widehat{\mathbb{E}}\left(m_{\psi;n;h_n;\mathbf{K}}(\mathbf{x})\right)\} \tag{12.13}$$
$$\overset{\mathbb{P}}{\to} \left\{\sup_{\mathbf{x}\in\mathbf{I}} \frac{\sigma_\psi^2(\mathbf{x})}{f(\mathbf{x})} \int_{\mathbb{R}^d} \mathbf{K}(\mathbf{t})^2 d\mathbf{t}\right\}^{1/2}.$$

The limiting statement (12.12) is due to Deheuvels [8] for $d = 1$, and [6] for $d \geq 1$ (see, e.g., Deheuvels and Einmahl [5], Deheuvels and Mason [9]). Earlier, Silverman [32] had established (12.12) for $d = 1$, under more stringent assumptions. Equation (12.13) is a particular case of Theorem 1.1 in Deheuvels and Mason [9] for $d = 1$, and of Theorem 1.2 in Deheuvels [7] for $d \geq 2$. The case where the rv Y has an unbounded support, will be considered elsewhere.

3°) The conclusion of Theorem 12.1.1 remains valid when $a_n \leq b_n$ are random sequences such that (12.8) holds in probability. As follows from the results of Deheuvels and Mason [8] and Deheuvels [5], additional conditions are required to obtain an almost sure [a.s.] version of this theorem.

4°) The properties of the estimators (12.2) and (12.6) have been extensively investigated since the seminal work of Rosenblatt [30], Parzen [29], Nadaraya [25] and Watson [40]. To allow data-dependent bandwidths, several authors (see, e.g., Mason et al. [24], Nolan and Marron [27], Deheuvels [4], Deheuvels and Mason [9]) have provided *uniform-in-bandwidth* limit laws for $f_{n,h}(\cdot)$, in the spirit of (12.9) and (12.10). Einmahl and Mason [16, 17] initiated the use of empirical processes indexed by functions to investigate this problem. For example, Theorem 1 of [17] shows that, for each $r > 0$,

$$\limsup_{n\to\infty} \left(\sup_{\frac{r\log n}{n} \leq h \leq 1} \left\{\frac{nh}{\log(1/h) \vee \log\log n}\right\}^{1/2} \right. \tag{12.14}$$
$$\left. \sup_{\mathbf{x}\in\mathbf{I}} |f_{n;h;\mathbf{K}}(\mathbf{x}) - \mathbb{E}\left(f_{n;h;\mathbf{K}}(\mathbf{x})\right)| \right) =: \mathcal{K}(I, r) < \infty,$$

a.s. for some $\mathcal{K}(I, r)$. We refer to Mason [22], Mason and Swanepoel [23], Dony [11, 13], Dony and Einmahl [12, 13], Dony et al. [15], Mason [21], Viallon [38], Varron [36, 37] and van Keilegom and Varron [35], for details on this methodology. In particular, an adaptation of the arguments of [16, 17] should allow us to prove that, under (12.8), as $n \to \infty$

$$\sup_{h\in\mathcal{H}_n}\left\{\frac{nh}{2\log_+(1/h)}\right\}^{1/2}\sup_{\mathbf{x}\in\mathbf{I}}\left|f_{n;h;\mathbf{K}}(\mathbf{x})-\mathbb{E}\left(f_{n;h;\mathbf{K}}(\mathbf{x})\right)\right| \tag{12.15}$$
$$-\left\{\sup_{\mathbf{x}\in\mathbf{I}} f(\mathbf{x})\int_{\mathbb{R}^d}\mathbf{K}(\mathbf{t})^2 d\mathbf{t}\right\} = o_{\mathbb{P}}(1).$$

It is not clear whether a proof of (12.9) (which is a stronger statement that (12.15)) can be achieved or not by these methods. Here, we make use of a different argument, based on the ideas of Deheuvels and Mason [8] and Deheuvels [5]. Further references are that of Dony and Mason [14] and Mason [20].

An outline of the remainder of our paper is as follows. We establish, in Theorem 12.2.1 below, a functional limit law for multivariate increments of a *non-uniform* empirical process (which is new, even for $d = 1$). To prove this theorem, we rely on classical arguments, to obtain, in the forthcoming Sect. 12.3.1, rough upper bounds for the modulus of continuity of multivariate empirical processes. Our proof then reduces to show that, for each fixed $M \geq 1$, the $N := M^d$ properly rescaled increments of the multivariate empirical process over sets of the form $\prod_{j=1}^{d}(\frac{k_j}{M}, \frac{k_j+1}{M}]$, cluster onto the unit ball of $\mathbb{R}^N$. To establish this property, we extend arguments of Deheuvels and Ouadah [10] to an dimension-free framework. The proof of Theorem 12.1.1 given Theorem 12.2.1 is captured in Sect. 12.2.4 below. The proofs being quite lengthy, we limit ourselves to the main arguments.

12.2 Functional Limit Laws

12.2.1 Main Result

For $d \geq 1$, let $(B([0,1]^d), \mathcal{U})$ denote the set $B([0,1]^d)$ of bounded functions on $[0,1]^d$, endowed with the topology $\mathcal{U}$, induced by the sup-norm $\|g\| := \sup_{\mathbf{u}\in[0,1]^d}|g(\mathbf{u})|$. Let $AC([0,1]^d)$ denote the set of absolutely continuous (with respect to the Lebesgue measure) functions on $[0,1]^d$, and set $AC_{\mathbf{0}}([0,1]^d) := \{g \in AC([0,1]^d) : g(\mathbf{0}) = 0\}$, with $\mathbf{0} := (0,\ldots,0) \in \mathbb{R}^d$. For each $\varepsilon > 0$ and $g \in B([0,1]^d)$, set $\mathcal{N}_\varepsilon(g) := \left\{\phi \in B([0,1]^d) : \|\phi - g\| < \varepsilon\right\}$, and for each

$A \subseteq [0,1]^d$, set $A^\varepsilon := \bigcup_{g\in A} \mathcal{N}_\varepsilon(g)$, with the convention that $\bigcup_\emptyset(\cdot) := \emptyset$. Define the sup-norm Hausdorff set-distance of $A, B \subseteq B([0,1]^d)$ by

$$\Delta(A,B) := \inf\{\theta > 0 : A \subseteq B^\theta \text{ and } B \subseteq A^\theta\},$$
$$\text{whenever such a } \theta \text{ exists, and}$$
$$\Delta(A,B) := \infty \quad \text{otherwise.}$$

Let $\dot{g}$ denote the Lebesgue derivative of $g \in AC([0,1]^d)$, and consider the Hilbert norm, defined on $B([0,1]^d)$ by

$$|g|_{\mathbb{H}} := \left\{\int_{[0,1]^d} \dot{g}(\mathbf{t})^2 d\mathbf{t}\right\}^{1/2} \text{ when } g \in AC_{\mathbf{0}}([0,1]^d),$$
$$|g|_{\mathbb{H}} := \infty \quad \text{otherwise.}$$

Set $\mathbb{S}_d = \{g \in B([0,1]^d) : |g|_{\mathbb{H}} \leq 1\}$. For $d = 1$, we will use this notation with subscripts omitted, and write, e.g., $\mathbb{S}$ for $\mathbb{S}_1$. The following relations follow readily from the Schwarz inequality and the definitions of $|\cdot|_{\mathbb{H}}$ and $\mathbb{S}_d$. For any $\psi \in B([0,1]^d)$, we have

$$\|\psi\| \leq |\psi|_{\mathbb{H}} \quad \text{and} \quad \sup_{g\in\mathbb{S}_d} \|g\| = 1. \tag{12.16}$$

Letting $\mathbf{X} := \mathbf{X}_1, \mathbf{X}_2, \ldots$ be as in Sect. 12.1, we denote the distribution function [df] of $\mathbf{X}$ by $\mathbb{F}(\mathbf{x}) := \mathbb{P}(\mathbf{X} \leq \mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^d$. Here, we write $\mathbf{x} \leq \mathbf{y}$, for $\mathbf{x} = (x(1),\ldots,x(d)) \in \mathbb{R}^d$ and $\mathbf{y} = (y(1),\ldots,y(d)) \in \mathbb{R}^d$, whenever $x(j) \leq y(j)$ for $j = 1,\ldots,d$. Denote the empirical df based upon $\mathbf{X}_1,\ldots,\mathbf{X}_n$, by

$$\mathbb{F}_n(\mathbf{x}) := n^{-1}\#\{\mathbf{X}_i \leq \mathbf{x} : 1 \leq i \leq n\} \quad \text{for} \quad \mathbf{x} \in \mathbb{R}^d, \tag{12.17}$$

where # denotes cardinality. Introduce the empirical process

$$a_n(\mathbf{x}) := n^{1/2}(\mathbb{F}_n(\mathbf{x}) - \mathbb{F}(\mathbf{x})) \quad \text{for} \quad \mathbf{x} \in \mathbb{R}^d. \tag{12.18}$$

Let $\mathbf{I} \subset \mathbf{J}$, with $\mathbf{I} = \prod_{j=1}^d [u_j, v_j]$, and $-\infty < u_j < v_j < \infty$ for $j = 1,\ldots,d$, be as in Sect. 12.1. We assume that the density $f(\cdot)$ of $\mathbf{X}$ is defined and continuous on $\mathbf{J}$, and bounded away from 0 on $\mathbf{I} \subset \mathbf{J}$. For $a > 0$, and $\mathbf{x} \in \mathbf{I}$, we consider the increment functions

$$\upsilon_n(a;\mathbf{x};\mathbf{u}) := \{a_n(\mathbf{x} + a^{1/d}\mathbf{u}) - a_n(\mathbf{x})\}/\sqrt{f(\mathbf{x})}, \tag{12.19}$$
$$\text{for} \quad \mathbf{u} \in [0,1]^d,$$

and set, for each $a > 0$, and $\mathbf{L} \subseteq \mathbf{I}$,

$$\mathcal{F}_{n;a;\mathbf{L}} = \left\{ \frac{\upsilon_n(a; \mathbf{x}; \cdot)}{\sqrt{2a \log_+(1/a)}} : \mathbf{x} \in \mathbf{L} \right\}. \tag{12.20}$$

Our main theorem may now be stated as follows.

Theorem 12.2.1 *Let* $0 < a_n \leq b_n$ *be such that, as* $n \to \infty$,

$$b_n \to 0 \quad \textit{and} \quad na_n / \log n \to \infty. \tag{12.21}$$

Then, with $\mathcal{H}_n = [a_n, b_n]$*, we have, as* $n \to \infty$,

$$\sup_{a \in \mathcal{H}_n} \Delta\left(\mathcal{F}_{n;a;\mathbf{I}}, \mathbb{S}_d\right) = o_{\mathbb{P}}(1). \tag{12.22}$$

Remark 12.2.1

1°) It will become obvious from our proofs that the conclusion of Theorem 12.2.1 remains valid if, in the definition (12.19) of $\upsilon(a; \mathbf{x}; \mathbf{u})$, $\mathbf{u}$ is assumed to vary in $\left[-\frac{1}{2}, \frac{1}{2}\right]$ (or in any specified bounded interval $[r, s]$, with $r < s$) instead of $[0, 1]$.

2°) To our best knowledge, the only version of Theorem 12.2.1 available up to now correspond to $d = 1$, and under the assumption that $\mathbf{X}$ uniformly distributed on $(0, 1)$ (see, e.g., Theorem 1(1) of Deheuvels and Ouadah [10]). When $a_n = b_n$ the problem has been considered by Deheuvels and Mason [8] and Deheuvels [5]) for $d = 1$, and by Mason [21] for $d \geq 1$. We note that the methods of [10] cannot be extended to $d \geq 2$, since the proofs rely on invariance principles for empirical processes, which are not presently available with the proper approximation rates.

The proof of Theorem 12.2.1 is postponed until Sect. 12.3. In the forthcoming Sect. 12.2.4, we shall provide a proof of Theorem 12.1.1 given Theorem 12.2.1.

12.2.2 A Limit Law for Local Empirical Processes Indexed by Functions

Let $\mathcal{K}$ denote a class of measurable functions defined on $\mathbb{R}^d$, with support in $\left[-\frac{1}{2}, \frac{1}{2}\right]^d$, and fulfilling $(K.1)$–$(K.3)$. Following (2.3)–(2.4) in Mason [21], for each $n \geq 1$, $h > 0$ and $\mathbf{x} \in \mathbb{R}^d$, denote the local empirical process at $\mathbf{x}$ indexed by $\mathbf{K} \in \mathcal{K}$ by

$$\mathcal{E}_n(h; \mathbf{x}; K) := (nh)^{-1/2} \sum_{i=1}^{n} \Big\{ \mathbf{K}(h^{-1/d}(\mathbf{x} - \mathbf{X}_i)) - \mathbb{E}\mathbf{K}(h^{-1/d}(\mathbf{x} - \mathbf{X}_i)) \Big\} = \sqrt{nh} \left\{ f_{n;h;\mathbf{K}}(\mathbf{x}) - \mathbb{E}\left(f_{n;h;\mathbf{K}}(\mathbf{x}) \right) \right\}, \tag{12.23}$$

and set, for $\mathbf{x} \in \mathbf{I}$,

$$\mathcal{L}_n(a; \mathbf{x}; K) = \frac{\mathcal{E}_n(a; \mathbf{x}; \mathbf{K})}{\sqrt{2 \log_+(1/a) f(\mathbf{x})}}. \tag{12.24}$$

Remark 12.2.2 Mason [21] make use of different conditions imposed upon $\mathcal{K}$. He assumes, namely that

$$\lim_{\|\mathbf{t}\| \to 0} \sup_{\mathbf{K} \in \mathcal{K}} \int_{\mathbb{R}^d} [\mathbf{K}(\mathbf{x} + \mathbf{t}) - \mathbf{K}(\mathbf{x})]^2 \, d\mathbf{x} = 0,$$

$$\lim_{\lambda \to 1} \sup_{\mathbf{K} \in \mathcal{K}} \int_{\mathbb{R}^d} [\mathbf{K}(\lambda \mathbf{x}) - \mathbf{K}(\mathbf{x})]^2 \, d\mathbf{x} = 0,$$

12.2.3 Properties of Kernels

We discuss here $(K.1)$–$(K.4)$. In $(K.1)$, the choice of the interval $[-A, A]^d \subset \mathbb{R}^d$ supporting the kernels $\mathbf{K} \in \mathcal{K}$, is a matter of convenience, so that we will work, without loss of generality, under the following variant of this assumption, for some $0 < \epsilon < \frac{1}{2}$.

$(K.1)^*$ Each $\mathbf{K} \in \mathcal{K}$ is such that $\mathbf{K}(\mathbf{t}) = 0$ for all $\mathbf{t} \notin \mathbf{I}_\epsilon := [\epsilon, 1 - \epsilon]^d$.

The condition $(K.2)$, requires each $\mathbf{K} \in \mathcal{K}$ to be of *Hardy-Krause bounded variation*. For functions of several variables, this notion is involved (see, e.g., Adams and Clarkson [1, 3], Niederreiter [26]), and some details must be given. The most common forms of variation [18, 19, 39], are as follows (see, e.g., Niederreiter [26, p. 22]). Set $\mathbf{I}_0 = [0, 1]^d$, and, for $1 \leq k \leq 1$ and $1 \leq i_1 < \ldots < i_k \leq d$, define a *face* of $\mathbf{I}_0$, by $\mathbf{I}_0(i_1, \ldots, i_k) := \{\mathbf{t} = (t_1, \ldots, t_d) \in \mathbf{I}_0 : t_j = 1 \text{ for } j \notin \{i_1, \ldots, i_k\}\}$. By an *interval* $\mathcal{J} \subseteq \mathbf{I}_0$, will be meant a product of d subintervals of [0, 1]. Denote the lower endpoint of $\mathcal{J}$ by $\mathbf{t}(\mathcal{J})$. For any function $\boldsymbol{\kappa}$ defined on $\mathbf{I}_0$, let $\Delta(\boldsymbol{\kappa}; \mathcal{J})$

denote the alternating sum of values of $\boldsymbol{\kappa}$ at vertices of $\mathcal{J}$, where $\boldsymbol{\kappa}(\mathbf{t}(\mathcal{J}))$ has coefficient 1. The *Vitali variation* of $\boldsymbol{\kappa}$ on $\mathbf{I}_0$ is then given by

$$\mathcal{V}_{\mathrm{V}}(\boldsymbol{\kappa};\mathbf{I}_0) := \sup_{\mathcal{P}(\mathbf{I}_0)} \sum_{\mathcal{J}\in\mathcal{P}(\mathbf{I}_0)} |\Delta(\boldsymbol{\kappa};\mathcal{J})|,$$

where the supremum is taken over all partitions $\mathcal{P}(\mathbf{I}_0)$ of $\mathbf{I}_0$ into subintervals $\mathcal{J} \subseteq \mathbf{I}_0$. The *Hardy-Krause variation* of $\boldsymbol{\kappa}$ on $\mathbf{I}_0$ is, in turn, defined by

$$\mathcal{V}_{\mathrm{HK}}(\boldsymbol{\kappa};\mathbf{I}_0) := \sum_{k=1}^{d} \Bigg\{ \sum_{1\le i_1<\ldots<i_k\le d} \mathcal{V}_{\mathrm{V}}(\kappa;\mathbf{I}_0(i_1,\ldots,i_k)) \Bigg\},$$

which sums, over all faces $\mathbf{I}_0(i_1,\ldots,i_k)$ of $\mathbf{I}_0$, the Vitali variation of the restriction of $\boldsymbol{\kappa}$ to $\mathbf{I}_0(i_1,\ldots,i_k)$. For $d = 1$, the Vitali and Hardy-Krause variations coincide with the usual *total variation*. In these definitions, we may replace $\mathbf{I}_0$ by other intervals of $\mathbb{R}^d$, via book-keeping arguments. In particular, we set, in $(K.2)$, $\mathcal{V}_{\mathrm{HK}}(\boldsymbol{\kappa}) := \mathcal{V}_{\mathrm{HK}}(\boldsymbol{\kappa};\mathbb{R}^d) := \sup_{m\ge1} \mathcal{V}_{\mathrm{HK}}(\boldsymbol{\kappa};[-m,m]^d)$.

Subject to the existence of continuous partial derivatives of $\boldsymbol{\kappa}$, the Vitali and Hardy-Krause variations of $\boldsymbol{\kappa}$ on $\mathbf{I}_0$ are given, respectively, by

$$\mathcal{V}_{\mathrm{V}}(\boldsymbol{\kappa};\mathbf{I}_0) = \int_{\mathbf{I}_0} \left| \frac{\partial^d \boldsymbol{\kappa}(\mathbf{t})}{\partial t_1 \ldots \partial t_d} \right| d\mathbf{t},$$

$$\mathcal{V}_{\mathrm{HK}}(\boldsymbol{\kappa};\mathbf{I}_0) = \sum_{k=1}^{d} \Bigg\{ \sum_{1\le i_1<\ldots<i_k\le d} \int_{\mathbf{I}_0(i_1,\ldots,i_k)} \left| \frac{\partial^k \boldsymbol{\kappa}(\mathbf{t})}{\partial t_{i_1} \ldots \partial t_{i_k}} \right| dt_{i_1}\ldots dt_{i_k} \Bigg\}.$$

In this case, an induction on d allows us to write, for each $\mathbf{0} \le \mathbf{u} \le \mathbf{v} \le \mathbf{1}$,

$$\boldsymbol{\kappa}(\mathbf{v}) - \boldsymbol{\kappa}(\mathbf{u}) = \sum_{k=1}^{d} \Bigg\{ \sum_{1\le i_1<\ldots<i_k\le d} \int_{\mathbf{t}\in\mathbf{I}_0(i_1,\ldots,i_k),\ \mathbf{u}<\mathbf{t}\le\mathbf{v}} (-1)^{k-d} \frac{\partial^k \boldsymbol{\kappa}(\mathbf{t})}{\partial t_{i_1} \ldots \partial t_{i_k}} \, \mathrm{d}t_{i_1}\ldots \mathrm{d}t_{i_k} \Bigg\}, \tag{12.25}$$

In general, subject to $\mathcal{V}_{\mathrm{HK}}(\boldsymbol{\kappa};\mathbf{I}_0) < \infty$, the totally bounded Lebesgue-Stieltjes signed measure $\boldsymbol{\nu} = \mathrm{d}\boldsymbol{\kappa}(\cdot)$, associated with $\boldsymbol{\kappa}$ and supported by $\mathbf{I}_0$, is defined by setting, for each continuous function ϕ on $\mathbf{I}_0$,

$$\int_{\mathbf{I}_0} \phi(\mathbf{t})\mathrm{d}\boldsymbol{\kappa}(\mathbf{t}) = \sum_{k=1}^{d} \Bigg\{ \sum_{1\le i_1<\ldots<i_k\le d} \lim_{|\mathcal{P}(\mathbf{I}_0(i_1,\ldots,i_k))|\to0} \sum_{\mathcal{J}\in\mathcal{P}(\mathbf{I}_0(i_1,\ldots,i_k))} (-1)^{k-d} \phi(\mathbf{t}(\mathcal{J}))\Delta(\boldsymbol{\kappa};\mathcal{J}) \Bigg\}. \tag{12.26}$$

Here, we set $|\mathcal{P}(\mathbf{I}_0(i_1,\ldots,i_k))| \to 0$, when the supremum vertice length of the intervals $\mathcal{J} \in \mathcal{P}(\mathbf{I}_0(i_1,\ldots,i_k))$ tends to 0. The kernel functions we consider have simple expressions in terms of $\boldsymbol{\nu} = \mathrm{d}\boldsymbol{\kappa}$. When $\boldsymbol{\kappa}$ is right-continuous, with $\boldsymbol{\kappa}(\mathbf{t}) = 0$ for $\mathbf{t} \notin \mathbf{I}_\varepsilon = [\epsilon, 1-\epsilon]^d$, $\boldsymbol{\kappa}(\mathbf{0}) = \boldsymbol{\kappa}(\mathbf{1}) = 0$, so that, by (12.26),

$$\boldsymbol{\kappa}(\mathbf{t}) = -\boldsymbol{\nu}((\mathbf{t},\mathbf{1}]) = \boldsymbol{\nu}((\mathbf{0},\mathbf{t}]) \quad \text{for} \quad \mathbf{t} \in \mathbf{I}_0. \tag{12.27}$$

Observe that $\boldsymbol{\nu} = \mathrm{d}\boldsymbol{\kappa}(\cdot)$ in (12.27) is a totally bounded signed measure with support in $\mathbf{I}_\varepsilon$. Letting $\boldsymbol{\nu} = \boldsymbol{\nu}_+ - \boldsymbol{\nu}_-$ denote the Hahn-Jordan decomposition of $\boldsymbol{\nu}$ into the difference of nonnegative bounded measures with supports in $\mathbf{I}_\varepsilon \subset \mathbf{I}_0$, we infer from (12.27) that these component measures fulfill

$$\boldsymbol{\kappa}(\mathbf{0}) = -\boldsymbol{\nu}((\mathbf{0},\mathbf{1}]) = -\boldsymbol{\nu}(\mathbf{I}_0) = -\boldsymbol{\nu}(\mathbf{I}_\varepsilon) = \boldsymbol{\nu}_-(\mathbf{I}_\varepsilon) - \boldsymbol{\nu}_+(\mathbf{I}_\varepsilon) = 0,$$

so that $0 \le \boldsymbol{\nu}_+(\mathbf{I}_\varepsilon) = \boldsymbol{\nu}_-(\mathbf{I}_\varepsilon) < \infty$. Following Bouleau [2] (see, e.g., p. 166 in Pagès and Xiao [28]), we define the *measure variation* of $\boldsymbol{\kappa}$ on $\mathbf{I}_0$, by

$$\mathcal{V}_{\mathrm{M}}(\boldsymbol{\kappa};\mathbf{I}_0) = \|\mathrm{d}\boldsymbol{\kappa}\|_{\mathrm{M}} := |\boldsymbol{\nu}|(\mathbf{I}_0) := \boldsymbol{\nu}_+(\mathbf{I}_0) + \boldsymbol{\nu}_-(\mathbf{I}_0). \tag{12.28}$$

The above-defined variations are related through the inequalities

$$\mathcal{V}_{\mathrm{V}}(\boldsymbol{\kappa};\mathbf{I}_0) \le \mathcal{V}_{\mathrm{M}}(\boldsymbol{\kappa};\mathbf{I}_0) \le \mathcal{V}_{\mathrm{HK}}(\boldsymbol{\kappa};\mathbf{I}_0) \le (2^d-1)\mathcal{V}_{\mathrm{M}}(\boldsymbol{\kappa};\mathbf{I}_0), \tag{12.29}$$

where $2^d - 1$ stands for the number of faces $\mathbf{I}_0(i_1,\ldots,i_k)$ of $\mathbf{I}_0$. In view of (12.29), under $(K.1)^*$–$(K.3)$, the assumption $(K.2)$ is equivalent to:

$(K.2)^*$ There exists a $B^* < \infty$ such that each $\mathbf{K} \in \mathcal{K}$ has a measure variation in $\mathbf{I}_0$ fulfilling $\mathcal{V}_{\mathrm{M}}(\mathbf{K};\mathbf{I}_0) \le B^*$.

Armed with these arguments, we establish, in Lemma 12.2.1 below, a useful integration by parts formula. We consider nonnegative bounded measures $\boldsymbol{\mu}_i$, $i = 1,2$ and $\boldsymbol{\nu}_i$, $i = 1,2$, with supports in $\mathbf{I}_\epsilon := [\epsilon, 1-\epsilon]^d$, and such that $\boldsymbol{\mu}_1(\mathbf{I}_\epsilon) = \boldsymbol{\mu}_2(\mathbf{I}_\epsilon)$, and $\boldsymbol{\nu}_1(\mathbf{I}_\epsilon) = \boldsymbol{\nu}_2(\mathbf{I}_\epsilon)$. Set, for $\mathbf{0} \le \mathbf{s} \le \mathbf{t} \le \mathbf{1}$,

$$\mathbf{M}_1(\mathbf{s},\mathbf{t}) = \Big\{\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2\Big\}((\mathbf{s},\mathbf{t}]) \quad \text{and} \quad \mathbf{M}_2(\mathbf{s},\mathbf{t}) = \Big\{\boldsymbol{\nu}_1 - \boldsymbol{\nu}_2\Big\}((\mathbf{s},\mathbf{t}]).$$

By (12.27), taken with $\boldsymbol{\nu} = d\{-\mathbf{M}_2(\mathbf{t},\mathbf{1})\}$ and $\boldsymbol{\kappa}(\mathbf{t}) = -\mathbf{M}_2(\mathbf{t},\mathbf{1})$, we see that $\boldsymbol{\nu}_1 - \boldsymbol{\nu}_2 = \mathrm{d}\{-\mathbf{M}_2(\mathbf{t},\mathbf{1})\}$ coincides with the Lebesgue-Stieltjes measure $\boldsymbol{\nu}$ induced by $-\mathbf{M}_2(\mathbf{t},\mathbf{1})$. Likewise, by (12.27), taken with $\boldsymbol{\nu} = \mathrm{d}\mathbf{M}_2(\mathbf{0},\mathbf{t})$ and $\boldsymbol{\kappa}(\mathbf{t}) = \mathbf{M}_1(\mathbf{0},\mathbf{t})$, we see that $\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 = \mathrm{d}\mathbf{M}_1(\mathbf{0},\mathbf{t})$ coincides with the Lebesgue-Stieltjes measure $\boldsymbol{\nu}$ induced by $\mathbf{M}_1(\mathbf{0},\mathbf{t})$.

Lemma 12.2.1 *Under the assumptions above, we have the integration by parts formula*

$$\int_{\mathbf{I}_0} \mathbf{M}_1(\mathbf{0}, \mathbf{t}) \mathrm{d}\mathbf{M}_2(\mathbf{t}, \mathbf{1}) = \int_{\mathbf{I}_0} \mathbf{M}_2(\mathbf{t}, \mathbf{1}) \mathrm{d}\mathbf{M}_1(\mathbf{0}, \mathbf{t}). \tag{12.30}$$

Proof We limit ourselves to the case where $\mathbf{M}_1(\mathbf{0}, \mathbf{t})$ and $\mathbf{M}_2(\mathbf{t}, \mathbf{1})$ have continuous partial derivatives of order d over $\mathbf{t} \in \mathbb{R}^d$. The proof in the general case is achieved by a smoothing argument which we omit. Observe that, for all $1 \leq k < d$ and $1 \leq i_1 < \ldots < i_k \leq d$, we have $\mathbf{M}_2(\mathbf{t}, \mathbf{1}) = 0$ for $\mathbf{t} \in \mathbf{I}_0(i_1, \ldots, i_k)$. Therefore, we may rewrite (12.25) into

$$\mathbf{M}_2(\mathbf{t}, \mathbf{1}) = (-1)^d \int_{\mathbf{s} \in \mathbf{I}_0,\ \mathbf{t} < \mathbf{s} \leq \mathbf{1}} \frac{\partial^d \mathbf{M}_2(\mathbf{s}, \mathbf{1})}{\partial s_1 \ldots \partial s_d} \mathrm{d}\mathbf{s}. \tag{12.31}$$

By a similar argument, with the formal replacement of $\mathbf{M}_2(\mathbf{t}, \mathbf{1})$ by $\mathbf{M}_1(\mathbf{0}, \mathbf{t})$, we may rewrite (12.25) into

$$\mathbf{M}_1(\mathbf{0}, \mathbf{t}) = \int_{\mathbf{s} \in \mathbf{I}_0,\ \mathbf{0} < \mathbf{s} \leq \mathbf{t}} \frac{\partial^d \mathbf{M}_1(\mathbf{0}, \mathbf{s})}{\partial s_1 \ldots \partial s_d} \mathrm{d}\mathbf{s}. \tag{12.32}$$

This shows that the signed measures $\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 = \mathrm{d}\mathbf{M}_1(\mathbf{0}, \mathbf{t})$ and $-\{\boldsymbol{\nu}_1 - \boldsymbol{\nu}_2\} = \mathrm{d}\mathbf{M}_2(\mathbf{t}, \mathbf{1})$ are absolutely continuous with respect to the Lebesgue measure in $\mathbb{R}^d$, with densities given, respectively, by

$$\mathbf{m}(\mathbf{t}) := \frac{\mathrm{d}\mathbf{M}_1(\mathbf{0}, \mathbf{t})}{\mathrm{d}\mathbf{t}} = \frac{\partial^d \mathbf{M}_1(\mathbf{0}, \mathbf{t})}{\partial t_1 \ldots \partial t_d},$$

and

$$\mathbf{n}(\mathbf{t}) := \frac{\mathrm{d}\mathbf{M}_2(\mathbf{t}, \mathbf{1})}{\mathrm{d}\mathbf{t}} = (-1)^d \frac{\partial^d \mathbf{M}_2(\mathbf{t}, \mathbf{1})}{\partial t_1 \ldots \partial t_d}.$$

Set $\mathbf{M}_{1;0}(\mathbf{t}) = \mathbf{m}(\mathbf{t})$, $\mathbf{M}_{2;0}(\mathbf{t}) = \mathbf{n}(\mathbf{t})$, and, for $1 \leq k \leq d$,

$$\mathbf{M}_{1;k}(\mathbf{t}) = \int_0^{t_1} \ldots \int_0^{t_k} \mathbf{m}(\mathbf{s}) \mathrm{d}s_1 \ldots \mathrm{d}s_k$$

and

$$\mathbf{M}_{2;k}(\mathbf{t}) = \int_{t_1}^1 \ldots \int_{t_k}^1 \mathbf{n}(\mathbf{s}) \mathrm{d}s_1 \ldots \mathrm{d}s_k.$$

Observe that $\mathbf{M}_{1;d}(\mathbf{t}) = \mathbf{M}_1(\mathbf{0}, \mathbf{t})$, $\mathbf{M}_{2;d}(\mathbf{t}) = \mathbf{M}_2(\mathbf{t}, \mathbf{1})$, and, for $1 \leq k \leq d$, $\frac{\partial}{dt_k} \mathbf{M}_{1;k}(\mathbf{t}) = \mathbf{M}_{1;k-1}(\mathbf{t})$ and $\frac{\partial}{dt_k} \mathbf{M}_{2;k}(\mathbf{t}) = -\mathbf{M}_{2;k-1}(\mathbf{t})$. In addition, for $1 \leq k \leq d$,

$\mathbf{M}_{1;k}(\mathbf{t}) = 0$ when $t_k = 0$ and $\mathbf{M}_{2;k}(\mathbf{t}) = 0$ when $t_k = 1$. We may therefore write the chain of equalities

$$\begin{aligned}
&\int_{[0,1]^d} \mathbf{M}_1(\mathbf{0},\mathbf{t}) d\mathbf{M}_2(\mathbf{t},\mathbf{1}) = (-1)^d \int_{[0,1]^d} \mathbf{M}_{1;d}(\mathbf{t})\mathbf{n}(\mathbf{t})\mathrm{d}\mathbf{t} \\
&= (-1)^d \int_{[0,1]^d} \mathbf{M}_{1;d}(\mathbf{t})\mathbf{M}_{2;0}(\mathbf{t}) d\mathbf{t} = (-1)^d \int_{[0,1]^d} \mathbf{M}_{1;d}(\mathbf{t}) \tfrac{\partial}{\partial t_1}\mathbf{M}_{2;1}(\mathbf{t})\mathrm{d}\mathbf{t} \\
&= (-1)^d \int_{[0,1]^{d-1}} \Big\{ \Big[\mathbf{M}_{1;d}(\mathbf{t})\mathbf{M}_{2;1}(\mathbf{t}) \Big]_{t_1=0}^{t_1=1} \\
&\qquad - \int_0^1 \tfrac{\partial}{\partial t_1}\mathbf{M}_{1;d}(\mathbf{t})\mathbf{M}_{2;1}(\mathbf{t})\mathrm{d}t_1 \Big\} \mathrm{d}t_2 \dots \mathrm{d}t_d \\
&= (-1)^{d-1} \int_{[0,1]^d} \mathbf{M}_{1;d-1}(\mathbf{t})\mathbf{M}_{2;1}(\mathbf{t})\mathrm{d}\mathbf{t} = \dots = \int_{[0,1]^d} \mathbf{M}_{1;0}(\mathbf{t})\mathbf{M}_{2;d}(\mathbf{t})\mathrm{d}\mathbf{t} \\
&= \int_{[0,1]^d} \mathbf{M}_{2;d}(\mathbf{t})\mathbf{m}(\mathbf{t})\mathrm{d}\mathbf{t} = \int_{[0,1]^d} \mathbf{M}_2(\mathbf{t},\mathbf{1})\mathrm{d}\mathbf{M}_1(\mathbf{0},\mathbf{t}),
\end{aligned}$$

which is (12.30). □

Remark 12.2.3 The version of (12.30) corresponding to $d = 1$, is readily checked, when $\mathbf{m}(\cdot)$ and $\mathbf{n}(\cdot)$ are continuous on [0, 1]. We obtain the relations

$$\begin{aligned}
&\int_0^1 \Big\{ \int_0^t \mathbf{m}(s)ds \Big\} \mathrm{d} \Big\{ \int_t^1 \mathbf{n}(s)ds \Big\} = \Big[\Big\{ \int_0^t \mathbf{m}(s)ds \Big\} \Big\{ \int_t^1 \mathbf{n}(s)ds \Big\} \Big]_{t=0}^{t=1} \\
&- \int_0^1 \Big\{ \int_t^1 \mathbf{n}(s)ds \Big\} \mathrm{d} \Big\{ \int_0^t \mathbf{m}(s)ds \Big\} = - \int_0^1 \Big\{ \int_t^1 \mathbf{n}(s)ds \Big\} \mathbf{m}(t)\mathrm{d}t.
\end{aligned}$$

12.2.4 Proof of Theorem 12.1.1

For each $\mathbf{K} \in \mathcal{K}$, set $\widetilde{\mathbf{K}}(\mathbf{u}) = \mathbf{K}(-\mathbf{u})$, and let $\widetilde{\mathcal{K}} = \{\widetilde{\mathbf{K}} : \mathbf{K} \in \mathcal{K}\}$. Following the arguments pp. 1278–1281 of [8], we may reduce the proof of (12.9) to the case where $\widetilde{K}$ fulfills $(K.1)^*$–$(K.2)^*$ and $(K.3)$, so that $\widetilde{\mathbf{K}}(\mathbf{u}) := \mathbf{K}(-\mathbf{u}) = 0$ for $\mathbf{u} \notin (0, 1)^d$. In view of (12.27), let $\mathrm{d}\widetilde{\mathbf{K}}(\cdot)$ be the Lebesgue-Stieltjes measure induced by $\widetilde{\mathbf{K}}$, in such a way that

$$-\widetilde{\mathbf{K}}(\mathbf{t}) = \int_{(\mathbf{t},\mathbf{1}]} \mathrm{d}\widetilde{\mathbf{K}}(\mathbf{u}).$$

Let $h_0 > 0$ be so small that $\mathbf{I} + h_0^{1/d}[0, 1]^d \subset \mathbf{J}$. By an application of Lemma 12.2.1, and making use of the definition (12.19) of $\upsilon_n(h; \mathbf{x}; \mathbf{u})$, we see that, for each $\mathbf{x} \in \mathbf{I}$, and $0 < h \le h_0$,

$$\left\{\frac{nh}{2\log_+(1/h)}\right\}^{1/2} \left(f_{n,h}(\mathbf{x}) - \mathbb{E}\left(f_{n,h}(\mathbf{x})\right)\right) \tag{12.33}$$
$$= \int_{[0,1]^d} \widetilde{\mathbf{K}}(\mathbf{u}) \left\{\frac{\mathrm{d}\{a_n(\mathbf{x} + h^{1/d}\mathbf{u}) - a_n(\mathbf{x})\}}{\sqrt{2h\log_+(1/h)}}\right\}$$
$$= -\int_{[0,1]^d} \left\{\frac{a_n(\mathbf{x} + h^{1/d}\mathbf{u}) - a_n(\mathbf{x})}{\sqrt{2h\log_+(1/h)}}\right\} \mathrm{d}\widetilde{\mathbf{K}}(\mathbf{u})$$
$$= -\sqrt{f(\mathbf{x})} \int_{[0,1]^d} \frac{\upsilon_n(h; \mathbf{x}; \mathbf{u})}{\sqrt{2h\log_+(1/h)}} \mathrm{d}\widetilde{\mathbf{K}}(\mathbf{u}).$$

We will need the following analytical result (see, e.g., Lemma 1 in [10]). Let $\mathcal{M}$ denote a subset of $B([0, 1]^d)$, such that $\mathbb{S}_d \subseteq \mathcal{M} \subseteq B([0, 1]^d)$, and let $\mathcal{T}$ denote a non-empty class of mappings $\Theta : \mathcal{M} \to \mathbb{R}$, continuous with respect to the uniform topology on $\mathcal{M}$. We assume that $\mathcal{T}$ has the following equicontinuity property. For each $\epsilon > 0$, there exists an $\eta(\epsilon) > 0$ such that, for each $\phi \in \mathcal{M}$ and $g \in \mathbb{S}_d$, we have

$$\|\phi - g\| < \eta(\epsilon) \;\Rightarrow\; \sup_{\Theta \in \mathcal{T}} |\Theta(\phi) - \Theta(g)| < \epsilon. \tag{12.34}$$

Lemma 12.2.2 *Under the assumptions above, for each* $\varepsilon > 0$*, there exists a* $\zeta(\varepsilon) > 0$*, such that, for any* $\mathcal{F} \subseteq \mathcal{M}$*, we have*

$$\Delta(\mathcal{F}, \mathbb{S}) < \zeta(\varepsilon) \;\Rightarrow\; \sup_{\Theta \in \mathcal{T}} \left| \sup_{\phi \in \mathcal{F}} \Theta(\phi) - \sup_{g \in \mathbb{S}_d} \Theta(g) \right| < \varepsilon. \tag{12.35}$$

Consider an arbitrary $\Theta \in \mathcal{T}$. By compactness of $\mathbb{S}_d$ and continuity of Θ, there exists a $g_\Theta \in \mathbb{S}_d$ such that $\Theta(g_\Theta) = \sup_{g \in \mathbb{S}_d} \Theta(g)$. Letting $\eta(\varepsilon)$ be as in (12.34), we see that, for each $\varepsilon > 0$, and $\phi \in \mathcal{M}$ such that $\|\phi - g_\Theta\| \le \eta(\varepsilon)$, we have $\sup_{\Theta \in \mathcal{T}} |\Theta(\phi) - \Theta(g_\Theta)| < \varepsilon$. In view of the implication $\Delta(\mathcal{F}, \mathbb{S}_d) \le \eta(\varepsilon) \Rightarrow \mathbb{S}_d \subseteq \mathcal{F}^{\eta(\varepsilon)}$, we see that $\Delta(\mathcal{F}, \mathbb{S}_d) \le \eta(\varepsilon)$ implies the existence of a $\phi_\Theta \in \mathcal{F}$ such that $\|\phi_\Theta - g_\Theta\| < \eta(\varepsilon)$. By an application of (12.34), we obtain therefore, that, whenever $\Delta(\mathcal{F}, \mathbb{S}_d) \le \eta(\varepsilon)$,

$$\forall\, \Theta \in \mathcal{T} :\; \sup_{\phi \in \mathcal{F}} \Theta(\phi) - \sup_{g \in \mathbb{S}_d} \Theta(g) \ge \Theta(\phi_\Theta) - \Theta(g_\Theta) \ge -\varepsilon. \tag{12.36}$$

Consider now the assumption

$$(H): \quad \Big\{\forall\,\eta>0, \exists\,\phi\in\mathcal{M}\cap\mathbb{S}_d^{\eta} : \sup_{\Theta\in\mathcal{T}}\Big\{\Theta(\phi)-\sup_{g\in\mathbb{S}_d}\Theta(g)\Big\}\geq\varepsilon\Big\}.$$

Under (H), there exists a sequence $(\phi_n,\Theta_n)\in(\mathcal{M}\cap\mathbb{S}_d^{1/n},\mathcal{T})$, $n=1,2,\ldots$, such that $\phi_n\in\mathcal{M}\cap\mathbb{S}_d^{1/n}$, and $\Theta_n(\phi_n)\geq\sup_{g\in\mathbb{S}_d}\Theta_n(g)+\varepsilon$, for all $n\geq 1$. The condition $\phi_n\in\mathbb{S}_d^{1/n}$ implies the existence, for each $n\geq 1$, of a $\psi_n\in\mathbb{S}$, such that $\|\phi_n-\psi_n\|\leq 1/n$. The compactness of $\mathbb{S}$ implies the existence of a convergent subsequence $\psi_{n_k}\to\psi\in\mathbb{S}_d$ as $k\to\infty$. Since then, $\|\phi_{n_k}-\psi\|\to 0$, as $k\to\infty$, an application of (12.34) shows that, as $k\to\infty$, $\sup_{\Theta\in\mathcal{T}}\big|\Theta(\phi_{n_k})-\Theta(\psi)\big|\to 0$. This entails that, for all k sufficiently large,

$$\Theta_{n_k}(\phi_{n_k})<\Theta_{n_k}(\psi)+\varepsilon\leq\sup_{g\in\mathbb{S}_d}\Theta_{n_k}(g)+\varepsilon,$$

which contradicts (H). The impossibility of (H) implies the existence of an $\eta_1(\varepsilon)$ such that whenever $\mathcal{F}\subseteq\mathcal{M}$ fulfills $\Delta(\mathcal{F},\mathbb{S}_d)\leq\eta_1(\varepsilon)$, and hence, $\mathcal{F}\subseteq\mathbb{S}_d^{\eta_1(\varepsilon)}$, we have

$$\forall\,\Theta\in\mathcal{T}: \sup_{\phi\in\mathcal{F}}\Theta(\phi)-\sup_{g\in\mathbb{S}_d}\Theta(g)\leq\varepsilon. \tag{12.37}$$

The conclusion (12.35) follows from (12.36) to (12.37), with $\zeta(\varepsilon):=\eta(\varepsilon)\wedge\eta_1(\varepsilon)$. □

Example 12.2.1

1°) Let $\mathcal{M}=B([0,1]^d)$, and $\mathcal{T}=\{\Theta_0\}$, with $\Theta_0(g):=\|g\|$. Since $\sup_{\Theta\in\mathcal{T}}|\Theta(\phi)-\Theta(g)|=\|\phi-g\|$, we see that (12.34) holds with $\eta(\varepsilon)=\varepsilon$, so that the assumptions of Lemma 12.2.2 are fulfilled.

2°) Let $\mathcal{K}$, where $\mathcal{K}$ fulfill $(K.1)^*$–$(K.2)^*$–$(K.3)$, and choose $\mathcal{M}$ as the set of all bounded measurable functions on $[0,1]^d$. The inclusions $\mathbb{S}_d\subseteq\mathcal{M}\subseteq B([0,1]^d)$ are then straightforward. Consider the functionals

$$g\in BV_{\mathbf{0};\mathrm{HK}}([0,1]^d)\to\Theta_K(g)=\int_{[0,1]^d}g(\mathbf{u})\mathrm{d}K(\mathbf{u}),$$

for $\mathbf{K}\in\mathcal{K}$. In view of the obvious inequality, for $g_1,g_2\in BV_0([0,1]^d)$,

$$|\Theta_K(g_1)-\Theta_K(g_2)|\leq\|g_1-g_2\|\times\mathcal{V}_{\mathrm{M}}(\mathbf{K},\mathbf{I}_0)\leq B^*\|g_1-g_2\|,$$

we see that (12.34) is fulfilled, with $\eta(\varepsilon)=\varepsilon/B^*$.

By a *rectangle* in $\mathbb{R}^d$ will be meant a product of d subintervals of $\mathbb{R}$. Below, we will denote by $|A|$ the Lebesgue measure of a measurable $A \subset \mathbb{R}^d$. Since $f(\cdot)$ is continuous on $\mathbf{J} \supset \mathbf{I}$, for each $0 < \epsilon < \nu := \inf_{\mathbf{x}\in\mathbf{I}} \sqrt{f(\mathbf{x})}$, we may partition the rectangle $\mathbf{I}$ into $\mathbf{I} = \mathbf{I}_1 \cup \ldots \cup \mathbf{I}_M$, where $\mathbf{I}_1, \ldots, \mathbf{I}_M \subset \mathbf{I}$ are disjoint rectangles in $\mathbb{R}^d$ such that, for $j = 1, \ldots, M$, $|\mathbf{I}_j| > 0$ and

$$m_j := \sup_{\mathbf{x}\in\mathbf{I}_j} \sqrt{f(\mathbf{x})} \geq \inf_{\mathbf{x}\in\mathbf{I}_j} \sqrt{f(\mathbf{x})} > m_j - \epsilon \geq \nu - \epsilon > 0.$$

By setting $\mathbf{L} = \mathbf{I}_j$, for $j = 1, \ldots, M$, and $a = h$ in (12.20), we may therefore write, for each $j = 1, \ldots, M$ and $0 < h \leq h_0$, the relations

$$\sup_{\mathbf{x}\in\mathbf{I}_j} \left| \left\{ m_j - \sqrt{f(\mathbf{x})} \right\} \int_{[0,1]^d} \frac{\upsilon_n(h;\mathbf{x};\mathbf{u})}{\sqrt{2h\log_+(1/h)}}\, d\widetilde{\mathbf{K}}(\mathbf{u}) \right| \tag{12.38}$$

$$\leq \epsilon \left\{ \sup_{g\in\mathcal{F}_{n;h;\mathbf{I}_j}} \|g\| \right\} \int_{[0,1]^d} |d\widetilde{\mathbf{K}}(\mathbf{u})| = \epsilon \left\{ \sup_{g\in\mathcal{F}_{n;h;\mathbf{I}_j}} \|g\| \right\} \|d\mathbf{K}\|,$$

where $\|d\mathbf{K}\| < \infty$ denotes the total variation of $\mathbf{K}(\cdot)$ on $\mathbb{R}^d$. Set now $\Theta(g) = \Theta_0(g) := \|g\|$ and $\mathcal{F} = \mathcal{F}_{n;h;\mathbf{I}_j}$. In view of (12.16) and (12.20), and by a repeated application of Theorem 12.2.1 with the formal replacement of $\mathbf{I}$ by $\mathbf{I}_j$, for $j = 1, \ldots, M$, we infer from (12.22) that, whenever $\mathcal{H}_n = [a_n, b_n]$ fulfills (12.21), we have, as $n \to \infty$,

$$\sup_{h\in\mathcal{H}_n} \left| \sup_{g\in\mathcal{F}_{n;h;\mathbf{I}_j}} \|g\| - \sup_{g\in\mathbb{S}_d} \|g\| \right| = \sup_{h\in\mathcal{H}_n} \left| \sup_{g\in\mathcal{F}_{n;h;\mathbf{I}_j}} \|g\| - 1 \right| = o_{\mathbb{P}}(1). \tag{12.39}$$

We infer readily from (12.38) and (12.39) that, as $n \to \infty$,

$$\mathbb{P}\Bigg(\max_{1\leq j\leq M} \sup_{h\in\mathcal{H}_n} \Bigg| \Bigg\{ \sup_{\mathbf{x}\in\mathbf{I}_j} \pm \left\{ \frac{nh}{2\log_+(1/h)} \right\}^{1/2} \left(f_{n,h}(\mathbf{x}) - \mathbb{E}\left(f_{n,h}(\mathbf{x}) \right) \right) \Bigg\}$$

$$- m_j \sup_{\mathbf{x}\in\mathbf{I}_j} \left\{ \pm (-1)^d \int_{[0,1]^d} \frac{\upsilon_n(h;\mathbf{x};\mathbf{u})}{\sqrt{2h\log_+(1/h)}}\, d\widetilde{\mathbf{K}}(\mathbf{u}) \right\} \Bigg| \geq 2\epsilon \|d\mathbf{K}\| \Bigg)$$

$$\leq \mathbb{P}\Bigg(\max_{1\leq j\leq M} \sup_{h\in\mathcal{H}_n} \Bigg\{ \sup_{\mathbf{x}\in\mathbf{I}_j} \Bigg| \left\{ m_j - \sqrt{f(\mathbf{x})} \right\}$$

$$\int_{[0,1]^d} \frac{\upsilon_n(h;\mathbf{x};\mathbf{u})}{\sqrt{2h\log_+(1/h)}}\, d\widetilde{\mathbf{K}}(\mathbf{u}) \Bigg| \Bigg\} \geq 2\epsilon \|d\mathbf{K}\| \Bigg) \to 0. \tag{12.40}$$

Set now

$$\Theta(g) = \Theta_1(g) := \pm \int_{[0,1]^d} g(\mathbf{u})\widetilde{\mathbf{K}}(\mathbf{u})d\mathbf{u}.$$

We may rewrite (12.40) into

$$\mathbb{P}\Bigg(\max_{1\le j\le M} \sup_{h\in\mathcal{H}_n} \Bigg| \Bigg\{ \sup_{\mathbf{x}\in\mathbf{I}_j} \pm \left\{ \frac{nh}{2\log_+(1/h)} \right\}^{1/2} \left(f_{n,h}(\mathbf{x}) - \mathbb{E}\left(f_{n,h}(\mathbf{x})\right)\right) \Bigg\}$$

$$-m_j \sup_{g\in\mathcal{F}_{n;h;\mathbf{I}_j}} \Theta(g) \Bigg| \ge 2\epsilon \|d\mathbf{K}\| \Bigg) \to 0. \tag{12.41}$$

After integrating by parts, we combine the definition of $\mathbb{S}_d$ with the Schwarz inequality, to obtain that

$$\sup_{g\in\mathbb{S}_d} \Theta(g) = \sup_{g\in\mathbb{S}_d} \left\{ \mp \int_{[0,1]^d} g(\mathbf{u}) d\widetilde{\mathbf{K}}(\mathbf{u}) \right\} \tag{12.42}$$

$$= \sup_{g\in\mathbb{S}_d} \left\{ \pm \int_{[0,1]^d} \dot{g}(\mathbf{u})\widetilde{\mathbf{K}}(\mathbf{u})d\mathbf{u} \right\} = \left\{ \int_{[0,1]^d} \mathbf{K}(\mathbf{u})^2 d\mathbf{u} \right\}^{1/2}.$$

For $j = 1,\ldots,M$, set $\mathcal{F} = \mathcal{F}_{n;h;\mathbf{I}_j}$. In view of (12.16)–(12.20), and by an application of Theorem 12.2.1, with $\mathbf{I} = \mathbf{I}_j$, for $j = 1,\ldots,M$, we infer from (12.22) that, whenever $\mathcal{H}_n = [a_n, b_n]$ fulfills (12.21), we have, as $n \to \infty$,

$$\max_{1\le j\le M} \sup_{h\in\mathcal{H}_n} \left| \sup_{g\in\mathcal{F}_{n;h;\mathbf{I}_j}} \Theta(g) - \sup_{g\in\mathbb{S}_d} \Theta(g) \right|$$

$$= \max_{1\le j\le M} \sup_{h\in\mathcal{H}_n} \left| \sup_{g\in\mathcal{F}_{n;h;\mathbf{I}_j}} \Theta(g) - \left\{ \int_{[0,1]^d} \mathbf{K}(\mathbf{u})^2 d\mathbf{u} \right\}^{1/2} \right| = o_{\mathbb{P}}(1).$$

This, when combined with (12.41), implies that, as $n \to \infty$,

$$\mathbb{P}\Bigg(\sup_{h\in\mathcal{H}_n} \Bigg| \Bigg\{ \sup_{\mathbf{x}\in\mathbf{I}} \pm \left\{ \frac{nh}{2\log_+(1/h)} \right\}^{1/2} \left(f_{n,h}(\mathbf{x}) - \mathbb{E}(f_{n,h}(\mathbf{x}))\right) \Bigg\} \tag{12.43}$$

$$-\left\{ \sup_{\mathbf{x}\in\mathbf{I}} \sqrt{f(\mathbf{x})} \right\} \left\{ \int_{[0,1]^d} \mathbf{K}(\mathbf{u})^2 d\mathbf{u} \right\}^{1/2} \Bigg| \ge \epsilon + 2\epsilon \|d\mathbf{K}\| \Bigg) \to 0.$$

Since $\epsilon \in (0, h_0]$ in (12.43) may be chosen arbitrarily small, we infer (12.9) from (12.43). This, together with routine arguments completes the proof of (12.9), given Theorem 12.2.1.

12.3 Proof of Theorem 12.2.1

12.3.1 A Bound for the Oscillation Modulus

In Proposition 12.3.1 below, we establish a rough bound for the oscillation modulus of the multivariate empirical process. This result will be instrumental in the proof of Theorem 12.2.1. We will work under the assumption that the support of the distribution of $\mathbf{X}$ is equal to $[0,1]^d$, and that the density $f(\cdot)$ of $\mathbf{X}$ is continuous and bounded away from 0 on $[0,1]^d$. This implies the existence of constants C_1, C_2, such that

$$0 < C_1 \leq f(\mathbf{x}) \leq C_2 < \infty \quad \text{for} \quad \mathbf{x} \in [0,1]^d. \tag{12.44}$$

The assumption that $\int_{[0,1]^d} f(\mathbf{x})d\mathbf{x} = 1$, implies that C_1, C_2 in (12.44) fulfill

$$0 < C_1 \leq 1 \leq C_2 < \infty. \tag{12.45}$$

Moreover, we may extend the definition of $f(\cdot)$ to $\overline{\mathbb{R}}^d := [-\infty, \infty]^d$, by setting

$$f(\mathbf{x}) = 0 \quad \text{for} \quad \mathbf{x} \notin [0,1]^d. \tag{12.46}$$

This entails that the *distribution function* [df] $\mathbb{F}(\mathbf{x}) := \mathbb{P}(\mathbf{X} \leq \mathbf{x})$ of $\mathbf{X} = (X(1), \ldots, X(d)) \in \mathbb{R}^d$, is continuous on $\overline{\mathbb{R}}^d$. For each $j = 1, \ldots, d$, set $\mathbf{x}^{[j]} := (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_d)$ and $d\mathbf{x}^{[j]} := dx_1 \ldots dx_{j-1}dx_{j+1} \ldots dx_d$. As follows from (12.44)–(12.46), for each $j = 1, \ldots, d$, the j-th coordinate $X(j)$ of $\mathbf{X}$ has a continuous density $f^{[j]}(\cdot)$ on $[0,1]$, fulfilling, for all $x_j \in [0,1]$,

$$C_1 \leq f^{[j]}(x_j) = \int_{\mathbf{x}^{[j]} \in [0,1]^{d-1}} f(\mathbf{x})d\mathbf{x}^{[j]} \leq C_2. \tag{12.47}$$

This, in turn, implies that for each $j = 1, \ldots, d$, the j-th marginal df of $\mathbb{F}(\cdot)$, denoted by $F^{[j]}(x) := \mathbb{P}(X(j) \leq x)$, $x \in \overline{\mathbb{R}}$, is continuous on $\overline{\mathbb{R}}$, and such that $U(j) := F^{[j]}(X(j))$ is uniformly distributed on $[0,1]$. For $j = 1, \ldots, d$, let $Q^{[j]}(t) := \inf\{x : F^{[j]}(x) \geq t\}$, $0 < t < 1$, $Q^{[j]}(0) := \inf\{x : F^{[j]}(x) > 0\}$, $Q^{[j]}(1) := \sup\{x : F^{[j]}(x) < 1\}$, denote the *quantile function* pertaining to $F^{[j]}(\cdot)$. For $j = 1, \ldots, d$, we have, almost surely [a.s.], $X(j) = Q^{[j]}(U(j))$. Without loss of generality, will therefore work on the set of probability 1 on which these relations hold. It is noteworthy that, unless $f(\mathbf{x}) = \prod_{j=1}^d f^{[j]}(x_j)$ for all $\mathbf{x} = (x_1, \ldots, x_d) \in [0,1]^d$, the components $U(1), \ldots, U(d)$ of $\mathbf{U} := (U(1), \ldots, U(d))$ are not independent. Their joint df, $\mathbb{C}(\mathbf{u}) := \mathbb{P}(\mathbf{U} \leq \mathbf{u})$, $\mathbf{u} \in \mathbb{R}^d$, is the *copula function* of $F(\cdot)$ (see, e.g., Schweizer and Wolff [31]). We have the reciprocal relations

$$\mathbb{F}(\mathbf{x}) = \mathbb{C}(F^{[1]}(x_1), \ldots, F^{[d]}(x_d)) \quad \text{for} \quad \mathbf{x} = (x_1, \ldots, x_d) \in \overline{\mathbb{R}}^d, \tag{12.48}$$

and

$$\mathbb{C}(\mathbf{u}) = \mathbb{F}(Q^{[1]}(u_1), \ldots, Q^{[d]}(u_d)) \quad \text{for} \quad \mathbf{u} = (u_1, \ldots, u_d) \in [0, 1]^d. \quad (12.49)$$

We infer from (12.47) that, for each $j = 1, \ldots, d$, the j-th quantile density function $q^{[j]}(t) := \frac{d}{dt} Q^{[j]}(t)$, $t \in (0, 1)$, is defined and continuous on $(0, 1)$, and fulfills, for $0 < t < 1$,

$$0 < \frac{1}{C_2} \leq q^{[j]}(t) = \frac{d}{dt} Q^{[j]}(t) = \frac{1}{f_j(Q^{[j]}(t))} \leq \frac{1}{C_1} < \infty. \quad (12.50)$$

The relations (12.44), (12.47)–(12.50), readily imply that the copula function $\mathbb{C}(\cdot)$ has a density $c(\cdot)$ on $(0, 1)^d$, fulfilling the relations, for $\mathbf{x} = (x_1, \ldots, x_d) \in (0, 1)^d$ and $\mathbf{u} = (u_1, \ldots, u_d) \in (0, 1)^d$

$$\begin{aligned} 0 < C_1 \leq f(\mathbf{x}) &= \frac{\partial^d}{\partial x_1 \ldots \partial x_d} \mathbb{F}(x_1, \ldots, x_d) \\ &= c(F^{[1]}(x_1), \ldots, F^{[d]}(x_d)) \prod_{j=1}^{d} f^{[j]}(x_j) \leq C_2 < \infty, \quad (12.51) \end{aligned}$$

$$\begin{aligned} 0 < \frac{C_1}{C_2^d} \leq c(\mathbf{u}) &= \frac{\partial^d}{\partial u_1 \ldots \partial u_d} \mathbb{C}(u_1, \ldots, u_d) \\ &= f(Q^{[1]}(u_1), \ldots, Q^{[d]}(u_d)) \prod_{j=1}^{d} q^{[j]}(u_j) \leq \frac{C_2}{C_1^d} < \infty. \quad (12.52) \end{aligned}$$

Let now $\mathbf{X}_i = (X_i(1), \ldots, X_i(d))$, $i \geq 1$, be iid random copies of $\mathbf{X}$, and set $\mathbf{U}_i = (U_i(1), \ldots, U_i(d)) := (F^{[1]}(X_i(1)), \ldots, F^{[d]}(X_i(d)))$, $i \geq 1$. In agreement with the notation of Sect. 12.2.1, the empirical df's based, respectively, upon $\mathbf{U}_1, \ldots, \mathbf{U}_n$ and $\mathbf{X}_1, \ldots, \mathbf{X}_n$, are denoted by

$$\mathbb{C}_n(\mathbf{u}) := n^{-1}\#\{\mathbf{U}_i \leq \mathbf{u} : 1 \leq i \leq n\}, \quad \mathbf{u} \in \mathbb{R}^d,$$

and

$$\mathbb{F}_n(\mathbf{x}) := n^{-1}\#\{\mathbf{X}_i \leq \mathbf{x} : 1 \leq i \leq n\}, \quad \mathbf{x} \in \mathbb{R}^d.$$

The corresponding empirical processes are denoted by

$$a_{n;\mathbb{C}}(\mathbf{u}) := n^{1/2} \{\mathbb{C}_n(\mathbf{u}) - \mathbb{C}(\mathbf{u})\}, \quad \mathbf{u} \in \overline{\mathbb{R}}^d,$$

and

$$a_{n;\mathbb{F}}(\mathbf{x}) := n^{1/2} \{\mathbb{F}_n(\mathbf{x}) - \mathbb{F}(\mathbf{x})\}, \quad \mathbf{x} \in \overline{\mathbb{R}}^d.$$

Denote the set of all rectangles in $[0, 1]^d$ by $\mathcal{R}_d$. The empirical measures indexed by $\mathcal{R}_d$, based, respectively, upon $\mathbf{U}_1, \dots, \mathbf{U}_n$ and $\mathbf{X}_1, \dots, \mathbf{X}_n$, are denoted by

$$\mu_{n;\mathbb{C}}(I) = n^{-1}\#\{\mathbf{U}_i \in I : 1 \leq i \leq n\}, \quad I \in \mathcal{R}_d,$$

and

$$\mu_{n;\mathbb{F}}(I) = n^{-1}\#\{\mathbf{X}_i \in I : 1 \leq i \leq n\}, \quad I \in \mathcal{R}_d,$$

with expectations, given, respectively, by

$$\mu_{\mathbb{C}}(I) = \int_I c(\mathbf{u})d\mathbf{u} \quad \text{and} \quad \mu_{\mathbb{F}}(I) = \int_I f(\mathbf{x})d\mathbf{x} \quad \text{for} \quad I \in \mathcal{R}_d.$$

The corresponding empirical processes indexed by $\mathcal{R}_d$ are denoted by

$$a_{n;\mathbb{C}}(I) := n^{1/2}\left\{\mu_{n;\mathbb{C}}(I) - \mu_{\mathbb{C}}(I)\right\} \quad \text{for} \quad I \in \mathcal{R}_d,$$

and

$$a_{n;\mathbb{F}}(I) := n^{1/2}\left\{\mu_{n;\mathbb{F}}(I) - \mu_{\mathbb{F}}(I)\right\} \quad \text{for} \quad I \in \mathcal{R}_d.$$

For $0 \leq u, v \leq 1$, consider the modulus of continuity of $a_{n;\mathbb{C}}$ and $a_{n;\mathbb{F}}$, defined, respectively, by

$$\omega_{n;\mathbb{C}}(v) = \sup\Big\{|a_{n;\mathbb{C}}(\mathbf{t} + vI)| : I \in \mathcal{R}_d, \quad \mathbf{t} \in [0, 1]^d, \mathbf{t} + vI \subseteq [0, 1]^d\Big\}, \tag{12.53}$$

$$\omega_{n;\mathbb{F}}(u) = \sup\Big\{|a_{n;\mathbb{F}}(\mathbf{x} + uI)| : I \in \mathcal{R}_d, \mathbf{x} \in \mathbb{R}^d\Big\}. \tag{12.54}$$

Recall the definition (12.44) of the constant C_2.

Lemma 12.3.1 *For all* $0 \leq u \leq 1/C_2$*, we have the inequality*

$$\omega_{n;\mathbb{F}}(u) \leq \omega_{n;\mathbb{C}}(C_2 u). \tag{12.55}$$

Proof Denote by $\overline{\mathcal{R}}_d$ the set of all closed rectangles of $\mathcal{R}_d$. Since (12.55) is trivial for $u = 0$, we assume that $0 < u \leq 1$, and set, for $\mathbf{x} := (x_1, \dots, x_d) \in [0, 1]^d$ and $I := \prod_{j=1}^d \left[y_j, z_j\right] \subseteq [0, 1]^d$, $I \in \overline{\mathcal{R}}_d$, such that $\mathbf{x} + uI \in [0, 1]^d$,

$$\mathbf{x} + uI = \prod_{j=1}^d [r_j(u, \mathbf{x}), s_j(u, \mathbf{x})] \subseteq [0, 1]^d, \tag{12.56}$$

where, for $j = 1, \dots, d$, $r_j(u, \mathbf{x})$ and $s_j(u, \mathbf{x})$ are such that

$$0 \leq r_j(u, \mathbf{x}) := x_j + uy_j \leq s_j(u, \mathbf{x}) := x_j + uz_j \leq 1, \tag{12.57}$$

and

$$0 \le s_j(u, \mathbf{x}) - r_j(u, \mathbf{x}) = u(z_j - y_j) \le u \le 1. \tag{12.58}$$

It is noteworthy that the mappings $\mathcal{F}$ and $\mathcal{Q}$, defined by

$$\begin{aligned} \mathbf{x} = (x(1), \dots, x(d)) \in [0, 1]^d \qquad & (12.59)\\ \to \mathcal{F}(\mathbf{x}) := (F^{[1]}(x(1)), \dots, F^{[d]}(x(d)) \in [0, 1]^d, & \\ \mathbf{u} = (u(1), \dots, u(d)) \in [0, 1]^d \qquad & (12.60)\\ \to \mathcal{Q}(\mathbf{u}) := (Q^{[1]}(u(1)), \dots, Q^{[d]}(u(d)) \in [0, 1]^d, & \end{aligned}$$

are continuous mappings of $[0, 1]^d$ onto itself, fulfilling $\mathcal{F} \circ \mathcal{Q} = \mathcal{Q} \circ \mathcal{F} = \mathbb{I}$, where $\mathbb{I}$ denotes identity. Therefore, for each $i \ge 1$, and $I \in \overline{\mathcal{R}}_d$, the event $\{\mathbf{X}_i \in \mathbf{x} + uI\}$ is identical to the event $\{\mathcal{F}(\mathbf{X}_i) = \mathbf{U}_i \in \mathcal{F}(\mathbf{x} + uI)\}$. Now we infer from (12.56), (12.57)–(12.58) and (12.59)–(12.60), that, with $\mathbf{x}$, u and I as above,

$$\mathcal{F}(\mathbf{x} + uI) = \prod_{j=1}^{d} \Big[F^{[j]}(r_j(u, \mathbf{x})), F^{[j]}(s_j(u, \mathbf{x})) \Big] = \mathbf{t} + vJ,$$

where $\mathbf{t} \in [0, 1]^d$, $v \in (0, 1]$ and $J \in \overline{\mathcal{R}}_d$ are such that

$$\mathbf{t} := \Big(F^{[1]}(r_1(u, \mathbf{x})), \dots, F^{[d]}(r_d(u, \mathbf{x})) \Big),$$

$$vJ := \prod_{j=1}^{d} \Big[0, F^{[j]}(s_j(u, \mathbf{x})) - F^{[j]}(r_j(u, \mathbf{x})) \Big],$$

with

$$v := C_2 u \quad \text{and} \quad J := \prod_{j=1}^{d} \left[0, \frac{F^{[j]}(s_j(u, \mathbf{x})) - F^{[j]}(r_j(u, \mathbf{x}))}{C_2 u} \right].$$

By (12.47) and (12.57)–(12.58), we see that, for $j = 1, \dots, d$ and $0 < u \le 1$,

$$\begin{aligned} 0 &\le F^{[j]}(s_j(u, \mathbf{x})) - F^{[j]}(r_j(u, \mathbf{x})) \\ &\le \Big\{ \sup_{0 \le x \le 1} f^{[j]}(x) \Big\} \big(s_j(u, \mathbf{x}) - r_j(u, \mathbf{x}) \big) \le C_2 u. \end{aligned}$$

Thus, we see that $J \subseteq [0, 1]^d$, whereas the inequality $0 < v \le 1$ is implied by the assumption $0 < u \le 1/C_2$. By all this, whenever $\mathbf{x} \in [0, 1]^d$, $I \in \overline{\mathcal{R}}_d$ and $0 < u \le 1/C_2$ are such that $\mathbf{x} + uI \subseteq [0, 1]^d$, then $\mathcal{F}(\mathbf{x} + uI) \subseteq [0, 1]^d$ is of the form $\mathbf{t} + vJ$, for some $\mathbf{t} \in [0,1]^d$, $J \in \overline{\mathcal{R}}_d$, and $0 < v = C_2 u \le 1$. In view of the definitions (12.53)–(12.54) of $\omega_{n;\mathbb{F}}(\cdot)$ and $\omega_{n;\mathbb{C}}(\cdot)$, and, making use of a similar argument for non-closed rectangles of $\mathcal{R}^d$, we readily obtain (12.55). □

The following fact is a special case of Theorem 1.5 in Stute [34].

Fact 12.3.1 *For each* $0 < \delta < \frac{1}{2}$*, there exist constants* $0 < c_1(\delta), c_2(\delta) < \infty$ *and* $C_3(\delta) > 0$*, such that, for all*

$$u^d \le c_2(\delta) \quad \text{and} \quad 0 < t \le c_1(\delta)\sqrt{\frac{nu^d}{2\log(1/u^d)}}\,,$$

we have

$$\mathbb{P}\left(\frac{\omega_{n;\mathbb{C}}(u)}{\sqrt{2u^d \log_+(1/u^d)}} \ge t\sqrt{\sup_{\mathbf{x}\in[0,1]^d} c(\mathbf{x})} \right) \le C_3(\delta) u^{d((1-\delta)t^2-1)}. \tag{12.61}$$

Lemma 12.3.2 *There exist constants* $c_3 > 0$*,* $c_4 > 0$*,* $C_4 > 0$ *and* $C_5 > 0$*, such that, whenever* $0 < a_n \le b_n < \infty$ *fulfill*

$$\frac{na_n^d}{\log(1/a_n^d)} \ge c_3 \quad \text{and} \quad b_n \le c_4, \tag{12.62}$$

we have, with $\mathcal{H}_n = [a_n, b_n]$*, as* $n \to \infty$*,*

$$\mathbb{P}\left(\sup_{a\in\mathcal{H}_n} \frac{\omega_{n;\mathbb{C}}(a)}{\sqrt{2a^d \log_+(1/a^d)}} \ge C_4 \right) \le C_5 b_n^{2d}. \tag{12.63}$$

Proof First, we observe that, for any $\frac{1}{2} \le \lambda \le 1$ and $h^d \le 1/e$, $\log_+(1/h^d) = \log(1/h^d) \le \log_+(1/(\lambda h)^d) = \log(1/(\lambda h)^d)$, and, therefore,

$$\begin{aligned} &\frac{\omega_{n;\mathbb{C}}(\lambda h)}{\sqrt{2(\lambda h)^d \log_+(1/(\lambda h)^d)}} \qquad\qquad (12.64)\\ &\le \frac{\omega_{n;\mathbb{C}}(h)}{\sqrt{2h^d \log(1/h^d)}} \times \frac{\sqrt{\log(1/h^d)}}{\sqrt{\lambda^d \log(1/(\lambda h)^d)}} \le \frac{2^{d/2}\omega_{n;\mathbb{C}}(h)}{\sqrt{2h^d \log(1/h^d)}}\,. \end{aligned}$$

Let now $0 < a \le 1$ be such that $a^d \le 1/e$ and select any $N \ge 0$. By a repeated application of (12.64) for $h = 2^{-k}a$ for $k = 0, \ldots, N$, we readily obtain that, for each $N \ge 0$,

$$A_N := \mathbb{P}\left(\sup_{2^{-N-1}a \le h \le a} \frac{\omega_{n;\mathbb{C}}(h)}{\sqrt{2h^d \log_+(1/h^d)}} \ge 2^{1+d/2} \sqrt{\sup_{\mathbf{x}\in[0,1]^d} c(\mathbf{x})} \right) \quad (12.65)$$

$$\mathbb{P}\left(\bigcup_{k=0}^{N} \left\{ \sup_{2^{-k-1}a \le h \le 2^{-k}a} \frac{2^{-d/2}\omega_{n;\mathbb{C}}(h)}{\sqrt{2h^d \log_+(1/h^d)}} \ge 2 \sqrt{\sup_{\mathbf{x}\in[0,1]^d} c(\mathbf{x})} \right\} \right)$$

$$\le \sum_{k=0}^{N} \mathbb{P}\left(\sup_{\frac{1}{2}\le\lambda\le 1} \frac{2^{-d/2}\omega_{n;\mathbb{C}}(\lambda 2^{-k}a)}{\sqrt{2(\lambda 2^{-k}a)^d \log_+(1/(\lambda 2^{-k}a)^d)}} \ge 2 \sqrt{\sup_{\mathbf{x}\in[0,1]^d} c(\mathbf{x})} \right)$$

$$\le \sum_{k=0}^{N} \mathbb{P}\left(\frac{\omega_{n;\mathbb{C}}(2^{-k}a)}{\sqrt{2(2^{-k}a)^d \log_+(1/(2^{-k}a)^d)}} \ge 2 \sqrt{\sup_{\mathbf{x}\in[0,1]^d} c(\mathbf{x})} \right).$$

Let now $0 < a \le 1$ and $N \ge 0$ be such that

$$a^d \le c_2(\tfrac{1}{4}) \wedge \{1/e\} \quad \text{and} \quad 2 \le c_1(\tfrac{1}{4}) \sqrt{\frac{n(2^{-N}a)^d}{2\log(1/((2^{-N}a)^d))}}.$$

By combining (12.65) with a repeated application of Fact 12.3.1, taken with $\delta = \frac{1}{4}$, $t = 2$ (so that $(1-\delta)t^2 - 1 = 2$) and $u = 2^{-k}a$ for $k = 0, \ldots, N$, we readily obtain that

$$A_N \le C_3(\tfrac{1}{4}) \sum_{k=0}^{N} \left(2^{-k}a\right)^{2d} \quad (12.66)$$

$$\le C_3(\tfrac{1}{4}) a^{2d} \sum_{k=0}^{\infty} \left(2^{-2d}\right)^k \le \tfrac{4}{3} C_3(\tfrac{1}{4}) a^{2d},$$

where we have used the fact that, independently of $d \ge 1$,

$$\sum_{k=0}^{\infty} \left(2^{-2d}\right)^k = \frac{1}{1-2^{-2d}} \le \tfrac{4}{3}.$$

We now set $a = b_n$ and choose $N \ge 0$ in such a way that

$$2^{-N-1}a \le a_n \le 2^{-N}a, \quad (12.67)$$

so that $a_n \leq 2^{-N}a < 2a_n$. Next, we observe that the function $\psi(t) := t/\log(1/t)$ is increasing on $(0, e]$. Thus, if we assume that $(2b_n)^d \leq e$, we obtain that $(2a_n)^d \leq e$, and $(a_n)^d \leq (2^{-N}a)^d \leq e$. We get therefore

$$\frac{na_n^d}{2\log(1/a_n^d)} \leq \frac{n(2^{-N}a)^d}{2\log(1/((2^{-N}a)^d))}.$$

By setting $\mathcal{H}_n = [a_n, b_n]$, we infer from (12.65)–(12.67) that, whenever $a_n \leq b_n$ fulfill

$$b_n^d \leq c_2(\tfrac{1}{4}) \wedge \{1/e\} \wedge \left\{2^{-d}e\right\} \quad \text{and} \quad \frac{na_n^d}{2\log(1/a_n^d)} \geq \frac{4}{c_1(\frac{1}{4})^2},$$

we have

$$\mathbb{P}\left(\sup_{a\in\mathcal{H}_n} \frac{\omega_{n;\mathbb{C}}(a)}{\sqrt{2a^d\log_+(1/a^d)}} \geq 2^{1+1/d}\sqrt{\sup_{\mathbf{x}\in[0,1]^d} c(\mathbf{x})}\right) \tag{12.68}$$
$$\leq A_N \leq \tfrac{4}{3}C_3(\tfrac{1}{4})b_n^{2d}.$$

Recalling (12.52), we set

$$C_4 := 2^{1+1/d}C_2^{1/2}C_1^{-d/2} \geq 2^{1+1/d}\sqrt{\sup_{\mathbf{x}\in[0,1]^d} c(\mathbf{x})}.$$

We therefore infer from (12.68) that (12.63) holds under (12.62), when the constants c_3, c_4 and C_5 are defined by

$$c_3 := 8/c_1(\tfrac{1}{4})^2,$$
$$c_4 := \left(c_2(\tfrac{1}{4}) \wedge \{1/e\} \wedge \left\{2^{-d}e\right\}\right)^{1/d},$$

and $C_5 := \frac{4}{3}C_3(\frac{1}{4})$. □

Proposition 12.3.1 *There exist constants $c_5 > 0$, $c_6 > 0$, $C_6 > 0$ and $C_7 > 0$, such that, whenever $0 < a_n \leq b_n < \infty$ fulfill*

$$\frac{na_n^d}{\log(1/a_n^d)} \geq c_5 \quad \textit{and} \quad b_n \leq c_6,$$

we have, with $\mathcal{H}_n = [a_n, b_n]$, as $n \to \infty$,

$$\mathbb{P}\left(\sup_{a\in\mathcal{H}_n} \frac{\omega_{n;\mathbb{F}}(a)}{\sqrt{2a^d\log_+(1/a^d)}} \geq C_6\right) \leq C_7 b_n^{2d}. \tag{12.69}$$

Proof We infer from (12.44) and (12.45) that $0 < 1/C_2 \le 1$. Thus, by (12.55), we have, for all $0 \le a \le \{1/C_2\} \wedge 1 = 1/C_2$,

$$\frac{\omega_{n;\mathbb{F}}(a)}{\sqrt{2a^d \log_+(1/a^d)}} \le \frac{\omega_{n;\mathbb{C}}(C_2 a)}{\sqrt{2(C_2 a)^d \log_+(1/(C_2 a)^d)}} \times C_2^{d/2} \left\{ \frac{\log(1/(C_2 a)^d)}{\log(1/a^d)} \right\}^{1/2}$$

$$= \frac{\omega_{n;\mathbb{C}}(C_2 a)}{\sqrt{2(C_2 a)^d \log_+(1/(C_2 a)^d)}} \times C_2^{d/2} \left\{ 1 + \frac{\log(1/C_2)}{\log(1/a)} \right\}^{1/2}$$

$$\le \frac{\omega_{n;\mathbb{C}}(C_2 a)}{\sqrt{2(C_2 a)^d \log_+(1/(C_2 a)^d)}} \times C_2^{d/2}.$$

□

12.3.2 Basic Arguments

For convenience, in the proof of Theorem 12.2.1 below, we will set $\mathbf{I} := \mathbf{I}_0 := [0, 1]^d$. The adaptation of our arguments to a general $\mathbf{I}$ is readily achieved, at the price of heavier notation. Letting $\mathbb{F}(\cdot)$ and $\mathbb{F}_n(\cdot)$ be as in Sect. 12.2.1, we denote by $d\mathbb{F}_n(\cdot)$ (resp. $d\mathbb{F}(\cdot)$) the empirical (resp. underlying) measure pertaining to $\{\mathbf{X}_i : 1 \le i \le n\}$, and write $da_n(\cdot) = n^{1/2}(d\mathbb{F}_n(\cdot) - d\mathbb{F}(\cdot))$, where $a_n(\cdot)$ is as in (12.18). For $N \ge 1$, we denote by $\mathbf{B}_N := \{\mathbf{z} \in \mathbb{R}^N : \|\mathbf{z}\| \le 1\}$ the unit ball of the Euclidian norm $\|\mathbf{z}\| := (\mathbf{z}'\mathbf{z})^{1/2}$ in $\mathbb{R}^N$. For each $\mathbf{z} \in \mathbb{R}^N$ and $\varepsilon > 0$, we set $\mathcal{N}_\varepsilon(\mathbf{z}) := \{\mathbf{y} \in \mathbb{R}^N : \|\mathbf{y} - \mathbf{z}\| < \varepsilon\}$, and for each $E \subseteq \mathbb{R}^N$, $E^\varepsilon := \bigcup_{\mathbf{z} \in E} \mathcal{N}_\varepsilon(\mathbf{z})$. For any $E, F \subseteq \mathbb{R}^N$, we write

$$\Delta(E, F) := \inf\{\theta > 0 : E \subseteq F^\theta \text{ and } F \subseteq E^\theta\},$$
$$\text{whenever such a } \theta \text{ exists, and}$$
$$\Delta(E, F) := \infty \quad \text{otherwise.}$$

Fix an integer $M \ge 1$, and select an $0 < a_0 < 1$ such that, for all $0 < a \le a_0$ and $\mathbf{x} \in \mathbf{I}_0 = [0, 1]^d$, we have $\mathbf{x} + a^{1/d}\mathbf{I}_0 \subseteq \mathbf{J}$. Let $\mathbf{i} := (i_1, \ldots, i_d) \in \mathbb{N}^d$ be such that $\mathbf{0} \le \mathbf{i} \le (M - 1) \times \mathbf{1}$, where $\mathbf{0} := (0, \ldots, 0) \in \mathbb{R}^d$ and $\mathbf{1} := (1, \ldots, 1) \in \mathbb{R}^d$. Consider the array of $N := M^d$ random variables, defined, for $\mathbf{0} \le \mathbf{i} \le (M - 1) \times \mathbf{1}$, by

$$Z_{n;\mathbf{x};\mathbf{i}}(a) := \frac{\sqrt{N}}{\sqrt{2af(\mathbf{x}) \log_+(1/a)}} \int_{\mathbf{x} + (a/M)^{1/d}(\mathbf{i} + \mathbf{I}_0)} da_n(\mathbf{t}). \qquad (12.70)$$

For each $\mathbf{x} \in \mathbf{I}_0$ and $0 < a \leq a_0$, denote by $Z_{n;\mathbf{x}}(a) \in \mathbb{R}^N$ the random vector of $\mathbb{R}^N$ obtained by sorting the array $\{Z_{n;\mathbf{x};\mathbf{i}}(a) : \mathbf{0} \leq \mathbf{i} \leq (M-1) \times \mathbf{1}\}$ in lexicographic order. For each $0 < a \leq a_0$ and $0 < \lambda < 1$ set

$$\mathbf{I}(a;\lambda) = \left\{\mathbf{x} \in \mathbf{I}_0 : \mathbf{x} = \lambda \mathbf{j}\, a^{1/d} \quad \text{for some} \quad \mathbf{j} \in \mathbb{N}^d\right\}.$$

Consider the set defined by

$$\mathcal{E}_{n;a;N}(\lambda) := \left\{Z_{n;\mathbf{x}}(a) : \mathbf{x} \in \mathbf{I}(a;\lambda)\right\}.$$

We note for further use that, for $0 < a \leq a_0$ and $0 < \lambda < 1$,

$$\#I(a;\lambda) = \#\{\mathbf{j} \in \mathbb{N}^d : \mathbf{0} \leq \mathbf{j} \leq \lfloor (1/(\lambda a^{1/d}) \rfloor \times \mathbf{1}\} \leq 2^d \lambda^{-d} a.$$

Observe that, for each $\mathbf{x} \in \mathbf{I}_0$, there exists a $\widetilde{\mathbf{x}} = \widetilde{\mathbf{x}}(\mathbf{x}) \in \mathbf{I}(a;\lambda)$ such that

$$\widetilde{\mathbf{x}} \leq \mathbf{x} \leq \widetilde{\mathbf{x}} + \lambda a^{1/d} \mathbf{1}.$$

We will show that Theorem 12.1.1 is equivalent to the following statement.

Theorem 12.3.1 *Set $\mathcal{H}_n = [a_n, b_n]$, where $0 < a_n \leq b_n$ fulfill, as $n \to \infty$,*

$$b_n \to 0 \quad \textit{and} \quad na_n/\log n \to \infty. \tag{12.71}$$

Then, for each $N = M^d \geq 1$ and $0 < \lambda \leq 1$, we have, as $n \to \infty$,

$$\sup_{a \in \mathcal{H}_n} \Delta\left(\mathcal{E}_{n;a;N}(\lambda), \mathbf{B}_N\right) = o_{\mathbb{P}}(1). \tag{12.72}$$

Proof of Theorem 12.3.1 To prove Theorem 12.1.1, we use of a discretization argument due to Deheuvels and Ouadah [10]. For each $0 < \rho < 1$ and $\mathcal{H}_n = [a_n, b_n]$, set

$$\mathcal{H}_n(\rho) = \left\{\rho^m b_n \in [a_n, b_n] : m \in \mathbb{N}\right\}.$$

We note that $\mathcal{H}_n(\rho)$ is never void, as long as $0 < a_n \leq b_n$. Given this notation, the proof of Theorem 12.1.1 reduces to show that, under (12.21), we have, for each $0 < \rho < 1$,

$$\sup_{a \in \mathcal{H}_n(\rho)} \Delta\left(\mathcal{F}_{n;a;\mathbf{I}}, \mathbb{S}_d\right) = o_{\mathbb{P}}(1). \tag{12.73}$$

The details of this argument are given in [10] for $d = 1$. However, it is easy to see that the same methods apply to an arbitrary $d \geq 1$, so that we omit details.

In a second step, we show that Theorem 12.1.1 is equivalent to Theorem 12.3.1. In view of the above preliminaries, this amounts to show that, under (12.71), the property that the assertion (12.73) holds for each $0 < \rho < 1$, is equivalent to the property that, for each $0 < \rho < 1$ and $N = M^d \geq 1$, as $n \to \infty$,

$$\sup_{a \in \mathcal{H}_n(\rho)} \Delta\left(\mathcal{E}_{n;a;N}, \mathbf{B}_N\right) = o_{\mathbb{P}}(1). \tag{12.74}$$

To show the equivalence between (12.73) and (12.74), we follow the discretization method used by Strassen [33] to establish his law of the iterated logarithm. The corresponding details are given in the forthcoming Sect. 12.3.4 for $d = 1$. Their extension to an arbitrary $d \geq 1$ is mostly a matter of book-keeping, with tedious notation for higher dimensions. We will therefore limit ourselves to the essential part of the argument. Consider the modulus of continuity of $a_n(\cdot)$, defined, for $0 < h \leq 1$, by

$$\omega_n(h) := \sup_{R \in \mathcal{R}} \left| \int_{h^{1/d} R} da_n(\mathbf{x}) \right|, \tag{12.75}$$

where $\mathcal{R}$ denotes the set of all *rectangles* in $\mathbf{I} = [0, 1]^d$. Given these preliminaries, the proof of the equivalence between (12.73) and (12.74) boils down to show that, under (12.72), for each $\varepsilon > 0$, there exists an $N = M^d$ such that

$$\mathbb{P}\left(\sup_{a \in \mathcal{H}_n} \frac{\omega_n(a/M)}{\sqrt{2a \log_+(1/a)}} \geq \varepsilon \right) \to 0. \tag{12.76}$$

This, in turn, will follow directly from Proposition 12.3.1 in the sequel. Given the above arguments, the proof of the equivalence between Theorems 12.1.1 and 12.3.1 is now complete.

It remains to show that (12.74) holds for each choice of $0 < \rho < 1$ and $N = M^d \geq 1$. This property turns out to be a consequence of the limiting results (12.77) and (12.78) below, which must hold, for each choice of $\varepsilon > 0$, $0 < \rho < 1$ and $N = M^d$. In the first place, we have, under (12.72),

$$\sum_{k: \rho^k b_n \in \mathcal{H}_n} \mathbb{P}\left(\mathcal{E}_{n;\rho^k b_n;N} \not\subseteq \mathbf{B}_N^{\varepsilon}\right) \to 0. \tag{12.77}$$

In the second place, we have, for each $0 \leq \|\mathbf{z}\| < 1$,

$$\sum_{k: \rho^k b_n \in \mathcal{H}_n} \mathbb{P}\left(\exists \mathbf{y} \in \mathcal{E}_{n;\rho^k b_n;N} : \mathbf{y} \in \mathcal{N}_{\varepsilon}(\mathbf{z})\right) \to 0. \tag{12.78}$$

The only remaining part of our proof is to obtain the appropriate probabilistic bounds allowing us to establish (12.77) and (12.78). Here, we use a simple trick.

Since the probabilities in (12.77) and (12.78) evaluate deviations of centered and rescaled *multinomial* random vectors in $\mathbb{R}^N$, for a *specified* $N \geq 1$, we may construct these multinomial laws in a space of arbitrary dimension d. This allows us to make use of the probabilistic inequalities obtained by Deheuvels and Ouadah [10] for $d = 1$. We note that the latter inequalities rely on strong invariance principles whose extension in higher dimensions is not presently available. Fortunately, the use of multinomial distributions allows us to avoid this technical difficulty. The proof of (12.77) and (12.78), follows directly from the forthcoming Propositions 12.3.2 and 12.3.3. In view of these arguments, the proofs of Theorems 12.1.1 and 12.3.1 is now completed. □

In the remainder of our paper, we outline the proofs of the key properties (12.76)–(12.78), on which rely the above-given proofs of Theorems 12.1.1 and 12.3.1.

12.3.3 Multinomial Inequalities

Let $N \geq 1$ be an integer which will be specified later on. Let $\mathbf{p} := (p_1, \ldots, p_N) \in \mathbb{R}_+^N$ fulfill $p_j > 0$ for $j = 1, \ldots, N$ and $p_{N+1} := 1 - |\mathbf{p}| := 1 - \sum_{j=1}^N p_j > 0$. For each $n \geq 1$, we denote the fact that the random vector $\mathbf{Z}_{n;\mathbf{p};N} := (Z_{n;\mathbf{p};1}, \ldots, Z_{n;\mathbf{p};N}) \in \mathbb{R}^N$ follows a multinomial distribution with parameters n and $\mathbf{p}$, by $\mathbf{Z}_{n;\mathbf{p};N} \overset{d}{=} \text{Mult}(n; \mathbf{p})$. This holds whenever, for any N-uple of nonnegative integers $\mathbf{k} := (k_1, \ldots, k_N)$, such that $k_{N+1} := n - |\mathbf{k}| := n - \sum_{j=1}^N k_j \geq 0$, we have

$$\mathbb{P}\left(\mathbf{Z}_{n;\mathbf{p};N} = \mathbf{k}\right) = \frac{n!}{k_1! \ldots k_{N+1}!}\, p_1^{k_1} \ldots p_{N+1}^{k_{N+1}}.$$

For each $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_N) \in \mathbb{R}_+^N$, set $|\boldsymbol{\delta}| := \sum_{j=1}^N \delta_j$, and consider

$$\mathcal{D}_N = \left\{\boldsymbol{\delta} := (\delta_1, \ldots, \delta_N) \in \mathbb{R}^N : \delta_j > 0,\ j = 1, \ldots, N;\ |\boldsymbol{\delta}| = N\right\}. \tag{12.79}$$

Whenever $\boldsymbol{\delta} \in \mathcal{D}_N$, set

$$0 < \boldsymbol{\delta}_{\min} := \min_{1 \leq j \leq N} \delta_j \leq 1 \leq \boldsymbol{\delta}_{\max} := \max_{1 \leq j \leq N} \delta_j. \tag{12.80}$$

We will set $\mathbf{p} = a\boldsymbol{\delta}/N$ for some $0 < a \leq 1$, so that $|\mathbf{p}| = aN^{-1}|\boldsymbol{\delta}| = a \leq 1$, and consider the random vector

$$\boldsymbol{\zeta}_{n;a;\boldsymbol{\delta}} := \frac{\sqrt{N}}{\sqrt{2na\log_+(1/a)}} \begin{bmatrix} Z_{n;a\boldsymbol{\delta}/N;1} - na\delta_1/N \\ \vdots \\ Z_{n;a\boldsymbol{\delta}/N;N} - na\delta_N/N \end{bmatrix} \in \mathbb{R}^N. \tag{12.81}$$

Denote by $\mathbf{B}_N := \{\mathbf{z} \in \mathbb{R}^N : \|\mathbf{z}\| \le 1\}$, the unit ball of the Euclidian norm $\|\mathbf{z}\| := (\mathbf{z}'\mathbf{z})^{1/2}$ in $\mathbb{R}^N$. Let, for each $\mathbf{z} \in \mathbb{R}^N$ and $\varepsilon > 0$, $\mathcal{N}_\varepsilon(\mathbf{z}) := \{\mathbf{y} \in \mathbb{R}^N : \|\mathbf{y} - \mathbf{z}\| < \varepsilon\}$, and set, for each $A \subseteq \mathbb{R}^N$, $A^\varepsilon := \bigcup_{\mathbf{z}\in A} \mathcal{N}_\varepsilon(\mathbf{z})$. We will need the following two propositions.

Proposition 12.3.2 *There exists a constant C_0 such that the following holds. For each $0 < \varepsilon \le 1$, there exist constants $0 < a_0(\varepsilon) \le 1/e$ and $0 < c_0(\varepsilon) < \infty$, together with an $n_0(\varepsilon) < \infty$, such that, for all $n \ge n_0(\varepsilon)$ and $a > 0$ fulfilling*

$$na/\log n \ge c_0(\varepsilon) \quad \text{and} \quad a \le a_0(\varepsilon), \tag{12.82}$$

and for all $N \ge 1$ and $\boldsymbol{\delta} \in \mathcal{D}_N$ fulfilling

$$\sqrt{\boldsymbol{\delta}_{min}} \ge \frac{1 + \frac{1}{2}\varepsilon}{1+\varepsilon}, \tag{12.83}$$

we have

$$\mathbb{P}\left(\boldsymbol{\zeta}_{n;a;\boldsymbol{\delta}} \notin \mathbf{B}_N^\varepsilon\right) \le C_0 a^{1+\varepsilon/(8N)}. \tag{12.84}$$

The proof of Proposition 12.3.2 is captured in Sects. 12.3.4 and 12.3.5 below.

For the next proposition, we will need the following additional notation. We consider a sequence $\boldsymbol{\delta}(k) = (\delta_1(k), \ldots, \delta_N(k)) \in \mathcal{D}_N$, $k = 1, \ldots, K$, and set $\mathbf{p}(k) = (p_1(k), \ldots, p_N(k)) := a\boldsymbol{\delta}(k)/N$, for $k = 1, \ldots, K$ and $0 < a \le 1/K$, so that $\sum_{k=1}^K |\mathbf{p}(k)| = aN^{-1}\sum_{k=1}^K |\boldsymbol{\delta}_k| = Ka \le 1$. Given $\{\boldsymbol{\delta}(k) : k = 1, \ldots, K\}$, we consider a sequence of random vectors

$$\mathbf{Z}^{(k)}_{n;\mathbf{p}(k);N} := (Z^{(k)}_{n;p_1(k);1}, \ldots, Z^{(k)}_{n;p_N(k);N}) \in \mathbb{R}^N, \quad k = 1, \ldots, K,$$

such that, with obvious notation,

$$(\mathbf{Z}^{(1)}_{n;\mathbf{p}(1);N}, \ldots, \mathbf{Z}^{(K)}_{n;\mathbf{p}(K);N}) \overset{d}{=} \text{Mult}(n; \mathbf{p}(1), \ldots, \mathbf{p}(K)).$$

In view of (12.81), we consider the random vectors, for $k = 1, \ldots, K$,

$$\boldsymbol{\zeta}^{(k)}_{n;a;\boldsymbol{\delta}(k)} := \frac{\sqrt{N}}{\sqrt{2na\log_+(1/a)}} \begin{bmatrix} Z^{(k)}_{n;p_1(k);1} - np_1(k) \\ \vdots \\ Z^{(k)}_{n;p_N(k);N} - np_N(k) \end{bmatrix} \in \mathbb{R}^N. \tag{12.85}$$

Proposition 12.3.3 *Fix any $\mathbf{z} \in \mathbf{B}_N$ such that $0 < \|\mathbf{z}\| < 1$. For each ε such that*

$$0 < \varepsilon < \left\{\frac{1}{2}\|\mathbf{z}\|\right\} \wedge \frac{1}{2N},$$

there exist an $a_2(\varepsilon, \mathbf{z})$*, together with* $n_2(\varepsilon) < \infty$ *and* $c_2(\varepsilon)$ *depending upon* ε *only, such that the following holds. For each* $\boldsymbol{\delta}(1), \ldots, \boldsymbol{\delta}(K) \in \mathcal{D}_N$*, and* $a_1, \ldots, a_k$*, whenever*

$$n \geq n_2(\varepsilon), \quad c_2(\varepsilon)n^{-1} \log n \leq a_1, \ldots, a_k \leq a_2(\varepsilon, g), \quad \sum_{k=1}^{K} a_k \leq \frac{1}{2}, \tag{12.86}$$

we have, for all $\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_K$*, fulfilling*

$$\frac{1}{\sqrt{\boldsymbol{\delta}_{max}}} \geq 1 - N\varepsilon \quad \textit{and} \quad \frac{1}{\sqrt{\boldsymbol{\delta}_{min}}} \leq 1 + N\varepsilon, \tag{12.87}$$

$$\mathbb{P}\left(\bigcap_{k=1}^{K} \left\{\boldsymbol{\zeta}_{n;a_k;\boldsymbol{\delta}(k)}^{(k)} \notin \mathcal{N}_{9N\varepsilon}(\mathbf{z})\right\}\right) \leq 2 \exp\left(-\tfrac{1}{4} \sum_{k=1}^{K} a_k^{1-\varepsilon/2}\right). \tag{12.88}$$

The proof of Proposition 12.3.3 is postponed until Sect. 12.3.6.

12.3.4 Outer Bounds

Let $U_1, U_2, \ldots$ be iid rv's with a uniform (0, 1) distribution. For $n \geq 1$ and $t \in \mathbb{R}$, denote by $\mathbb{U}_n(t) := n^{-1}\#\{U_i \leq t : 1 \leq i \leq n\}$ the empirical df based upon $U_1, \ldots, U_n$, and by $\alpha_n(t) := n^{1/2}(\alpha_n(t) - t)$, the uniform empirical process. For $n \geq 1$, $a > 0$, $t \in [0, 1]$ and $u \in \mathbb{R}$, set

$$\xi_n(a; t; u) = \alpha_n(t + au) - \alpha_n(t). \tag{12.89}$$

The following fact is Proposition 2 of Deheuvels and Ouadah [10].

Fact 12.3.2 *There exists a constant* C_2 *such that the following holds. For each* $0 < \varepsilon \leq 1$*, there exist constants* $0 < a_1(\varepsilon) \leq 1/e$ *and* $0 < c_1(\varepsilon) < \infty$*, together with an* $n_1(\varepsilon) < \infty$*, such that, for all* $n \geq n_1(\varepsilon)$ *and* $a > 0$ *fulfilling*

$$na/\log n \geq c_1(\varepsilon) \quad \textit{and} \quad a \leq a_1(\varepsilon), \tag{12.90}$$

we have, for all $t \in [0, 1 - a]$*,*

$$\mathbb{P}\left(\frac{\xi_n(a; t; \cdot)}{\sqrt{2a \log_+(1/a)}} \notin \mathbb{S}^\varepsilon\right) \leq C_2 a^{1+\varepsilon}. \tag{12.91}$$

The following lemmas are oriented towards the proof of Proposition 12.3.2.

Lemma 12.3.3 *For any* $g \in B([0,1])$ *and* $0 \le s, t \le 1$, *we have*

$$|g(t) - g(s)| \le |g|_{\mathbb{H}} \sqrt{|t-s|}\,, \tag{12.92}$$

and, for any $0 \le t \le t + h \le 1$, *we have*

$$\sup_{0 \le u \le 1} |g(t+hu) - g(t) - u(g(t+h) - g(t))| \le |g|_{\mathbb{H}} \sqrt{\tfrac{1}{2} h}\,, \tag{12.93}$$

Proof When $g \notin AC_0([0,1])$, $|g|_{\mathbb{H}} = \infty$ and (12.92)–(12.93) are trivial. Therefore, we limit ourselves to $g \in AC_0[0,1]$. The Schwarz inequality enables us to write the relations

$$|g(t) - g(s)| = \left| \int_s^t \dot{g}(u)du \right| \le \left| \int_s^t du \right|^{1/2} \left| \int_s^t \dot{g}(u)^2 du \right|^{1/2} \le |g|_{\mathbb{H}} \sqrt{|t-s|}\,,$$

which yield (12.92).

For $g \in AC_0([0,1])$, the function $\phi(u) := g(t+hu) - g(t) - u(g(t+h) - g(t))$, for $0 \le u \le 1$, is such that

$$\phi(0) = \phi(1) = \int_0^1 \dot{\phi}(u)du = 0.$$

Moreover, setting $\psi(u) := h\dot{g}(t+hu)$, for $0 \le u \le 1$, we get

$$\dot{\phi}(u) = h\dot{g}(t+hu) - (g(t+h) - g(t)) = \psi(u) - \int_0^1 \psi(t)dt.$$

Observe that

$$\int_0^1 \dot{\phi}(u)^2 du = \int_0^1 \psi(u)^2 du - \left\{ \int_0^1 \psi(t)dt \right\}^2$$
$$\le \int_0^1 \psi(u)^2 du = h \int_t^{t+h} \dot{g}(s)^2 ds \le h|g|_{\mathbb{H}}^2.$$

An easy argument shows that the supremum of $|\varphi(c)| = |\int_0^c \dot{\varphi}(u)du|$ subject to the constraints $0 \le c \le 1$, $\varphi(0) = 0$, $\int_0^1 \dot{\varphi}(u)du = 0$ and $\int_0^1 \dot{\varphi}(u)^2 du \le \lambda$, is equal to $\frac{1}{2}\sqrt{\lambda}$, and reached when $c = \frac{1}{2}$ and $\dot{\varphi}(u) = \sqrt{\lambda}$, $0 < u < \frac{1}{2}$, $\dot{\varphi}(u) = -\sqrt{\lambda}$, $\frac{1}{2} < u < 1$. Since $\varphi = \phi$ fulfills these conditions with $\lambda := h|g|_{\mathbb{H}}^2$, it follows that the maximal possible value of ϕ on $[0,1]$ is less than or equal to $|g|_{\mathbb{H}} \sqrt{\frac{1}{2} h}$. We so obtain (12.93). □

Fix $N \geq 1$, and let $\mathcal{D}_N$ be as in (12.79). For any $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_N) \in \mathcal{D}_N$, set $t_j(\boldsymbol{\delta}) = N^{-1} \sum_{k=1}^{j} \delta_j$, for $j = 0, \ldots, N$, with the convention that $\sum_{\emptyset}(\cdot) := 0$. As in (12.80), set $\boldsymbol{\delta}_{\min} = \min_{1 \leq j \leq N} \delta_j$, and $\boldsymbol{\delta}_{\max} = \max_{1 \leq j \leq N} \delta_j$. Consider the linear maps $\mathcal{P}_{N;\boldsymbol{\delta}}(\cdot)$ and $\mathcal{Q}_{N;\boldsymbol{\delta}}(\cdot)$, defined by

$$g \in B[0,1] \tag{12.94}$$

$$\to \mathcal{P}_{N;\boldsymbol{\delta}}(g) := \begin{bmatrix} \sqrt{\frac{N}{\delta_1}}\,(g(t_1(\boldsymbol{\delta})) - g(t_0(\boldsymbol{\delta}))) \\ \vdots \\ \sqrt{\frac{N}{\delta_N}}\,(g(t_N(\boldsymbol{\delta})) - g(t_{N-1}(\boldsymbol{\delta}))) \end{bmatrix} \in \mathbb{R}^N,$$

$$\mathbf{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_N \end{bmatrix} \in \mathbb{R}^N \to \mathcal{Q}_N(\mathbf{z}) \in AC[0,1], \tag{12.95}$$

where we define $\mathcal{Q}_{N,\boldsymbol{\delta}}(\mathbf{z})$ for $\mathbf{z} = (z_1, \ldots, z_N) \in \mathbb{R}^N$, by setting $z_0 = 0$, $\sum_{\emptyset}(\cdot) = 0$, and, for $k = 1, \ldots, N$,

$$\mathcal{Q}_{N,\boldsymbol{\delta}}(\mathbf{z})(t) = \sum_{j=1}^{k-1} \sqrt{\frac{\delta_j}{N}}\, z_j + \sqrt{\frac{N}{\delta_k}}\, z_k\,(t - t_{k-1}(\boldsymbol{\delta})) \tag{12.96}$$
$$\text{when } t_{k-1}(\boldsymbol{\delta}) \leq t \leq t_k(\boldsymbol{\delta})\,.$$

Lemma 12.3.4 *For $N \geq 1$, $\boldsymbol{\delta} \in \mathcal{D}_N$, $\mathbf{z} \in \mathbb{R}^N$ and $g \in B([0,1])$, we have*

$$\mathcal{P}_{N,\boldsymbol{\delta}}(\mathcal{Q}_{N,\boldsymbol{\delta}}(\mathbf{z})) = \mathbf{z}; \tag{12.97}$$

$$\big\| \mathcal{Q}_{N,\boldsymbol{\delta}}(\mathcal{P}_{N,\boldsymbol{\delta}}(g)) - g \big\| \leq (2N)^{-1/2} |g|_{\mathbb{H}} \sqrt{\boldsymbol{\delta}_{max}}\,; \tag{12.98}$$

$$\|\mathcal{P}_{N,\boldsymbol{\delta}}(g)\| \leq |g|_{\mathbb{H}} \quad \textit{and} \quad |\mathcal{Q}_{N,\boldsymbol{\delta}}(\mathbf{z})|_{\mathbb{H}} = \|\mathbf{z}\|; \tag{12.99}$$

$$\|\mathcal{P}_{N,\boldsymbol{\delta}}(g)\| \leq 2N \|g\| / \sqrt{\boldsymbol{\delta}_{min}}\,; \tag{12.100}$$

$$\mathcal{P}_{N,\boldsymbol{\delta}}(\mathbb{S}) = \mathbf{B}_N := \{\mathbf{t} \in \mathbb{R}^N : \mathbf{t}'\mathbf{t} \leq 1\}; \tag{12.101}$$

$$\mathcal{Q}_{N,\boldsymbol{\delta}}(\mathbf{B}_N) \subseteq \mathbb{S} \subseteq \mathcal{Q}_{N,\boldsymbol{\delta}}(\mathbf{B}_N)^{\sqrt{\boldsymbol{\delta}_{max}/(2N)}}\,. \tag{12.102}$$

Proof By (12.96), $\mathcal{Q}_{N,\boldsymbol{\delta}}(\mathbf{z})(t_j(\boldsymbol{\delta})) - \mathcal{Q}_{N,\boldsymbol{\delta}}(\mathbf{z})(t_{j-1}(\boldsymbol{\delta})) = z_j \sqrt{\delta_j/N}$ for $j = 1, \ldots, N$. Thus, by (12.94), we have $\mathcal{P}_{N,\boldsymbol{\delta}}(\mathcal{Q}_{N,\boldsymbol{\delta}}(\mathbf{z})) = \mathbf{z}$, which is (12.97). Since $|g|_{\mathbb{H}} = \infty$ when $g \notin AC_0([0,1])$, there is no loss of generality to assume in our proofs of (12.98)–(12.99) that $g \in AC_0([0,1])$. To establish (12.98) we observe that, for $j = 0, \dots, N$, $\mathcal{Q}_{N,\boldsymbol{\delta}}(\mathcal{P}_N(g))(t_j(\boldsymbol{\delta})) = g(t_j(\boldsymbol{\delta}))$, so that, by applying (12.93), for $j = 1, \dots, N$, with $h = \delta_j/N$, we get

$$\|\mathcal{Q}_{N,\boldsymbol{\delta}}(\mathcal{P}_{N,\boldsymbol{\delta}}(g)) - g\| \le \max_{1\le j\le N}\left(\sup_{0\le u\le 1}\left|g\left(t_{j-1}(\boldsymbol{\delta}) + u\frac{\delta_j}{N}\right) - g(t_{j-1}(\boldsymbol{\delta}))\right.\right.$$
$$\left.\left.-u\left\{g\left(t_{j-1}(\boldsymbol{\delta}) + \frac{\delta_j}{N}\right) - g\left(\frac{\delta_j}{N}\right)\right\}\right|\right) \le |g|_{\mathbb{H}} \max_{1\le j\le N}\sqrt{\frac{\delta_j}{2N}},$$

which yields (12.98). To establish the first half of (12.99), we select a $g \in AC_0[0,1]$ and set $\mathbf{z} = (z_1, \dots, z_d) = \mathcal{P}_{N,\boldsymbol{\delta}}(g)$. It follows from (12.94) that $z_j = \sqrt{\frac{N}{\delta_j}}\,(g(t_j(\boldsymbol{\delta})) - g(t_{j-1}(\boldsymbol{\delta})))$, for $j = 1, \dots, d$. Making use of the Schwarz inequality, we get, in turn,

$$\|\mathcal{P}_{N,\boldsymbol{\delta}}(g)\|^2 = \mathbf{z}'\mathbf{z} = \sum_{j=1}^N z_j^2 = N\sum_{j=1}^N \frac{1}{\delta_j}\left(\int_{t_{j-1}(\boldsymbol{\delta})}^{t_j(\boldsymbol{\delta})} \dot{g}(u)du\right)^2$$
$$\le \sum_{j=1}^N \frac{N}{\delta_j}\left(\int_{t_{j-1}(\boldsymbol{\delta})}^{t_j(\boldsymbol{\delta})} du\right)\left(\int_{t_{j-1}(\boldsymbol{\delta})}^{t_j(\boldsymbol{\delta})} \dot{g}(u)^2du\right) = \int_0^1 \dot{g}(u)^2du = |g|_{\mathbb{H}}^2,$$

as sought. Next, we choose a $\mathbf{z} \in \mathbb{R}^N$, and set $g = \mathcal{Q}_{N,\boldsymbol{\delta}}(\mathbf{z})$. We infer from (12.96) that, for $j = 1, \dots, N$,

$$\dot{g}(t) = z_j\sqrt{\frac{N}{\delta_j}} \quad \text{for} \quad t_{j-1}(\boldsymbol{\delta}) \le t \le t_j(\boldsymbol{\delta}),$$

whence

$$|\mathcal{Q}_{N,\boldsymbol{\delta}}(\mathbf{z})|_{\mathbb{H}}^2 = \sum_{j=1}^N \int_{t_{j-1}(\boldsymbol{\delta})}^{t_j(\boldsymbol{\delta})} \frac{Nz_j^2}{\delta_j}\,du = \sum_{j=1}^N z_j^2 = \|\mathbf{z}\|^2,$$

which yields the second half of (12.99). To establish (12.100), we infer from (12.94) that, for an arbitrary $g \in B([0,1])$,

$$\|\mathcal{P}_{N,\boldsymbol{\delta}}(g)\|^2 = \sum_{j=1}^d \frac{N}{\delta_j}(g(t_j(\boldsymbol{\delta})) - g(t_{j-1}(\boldsymbol{\delta})))^2$$
$$\le 4N\|g\|^2\sum_{j=1}^n \frac{1}{\delta_j} \le \frac{(2N\|g\|)^2}{\boldsymbol{\delta}_{\min}}.$$

To establish (12.101), we first infer from (12.99) that $\mathcal{P}_{N,\boldsymbol{\delta}}(g) \in \mathbf{B}_N$ for each $g \in \mathbb{S}$, so that $\mathcal{P}_{N,\boldsymbol{\delta}}(\mathbb{S}) \subseteq \mathbf{B}_N$. Conversely, by (12.99), for any $\mathbf{z} \in \mathbf{B}_N$, we have $g := \mathcal{Q}_{N,\boldsymbol{\delta}}(\mathbf{z}) \in \mathbb{S}$. This, in turn, implies, via (12.97), that $\mathcal{P}_{N,\boldsymbol{\delta}}(g) = \mathbf{z}$, whence $\mathbf{B}_N \subseteq \mathcal{P}_{N,\boldsymbol{\delta}}(\mathbb{S})$. We so obtain (12.101). Next, we infer from (12.99) that, for each $\mathbf{z} \in \mathbf{B}_N$, $\mathcal{Q}_{N,\boldsymbol{\delta}}(\mathbf{z}) \in \mathbb{S}$. This, in turn, implies that $\mathcal{Q}_{N,\boldsymbol{\delta}}(\mathbf{B}_N) \subseteq \mathbb{S}$. Finally, we infer from (12.98) and (12.99) that, for each $g \in \mathbb{S}$, we have $\mathbf{y} := \mathcal{P}_{N,\boldsymbol{\delta}}(g) \in \mathbf{B}_N$ and $\|\mathcal{Q}_{N,\boldsymbol{\delta}}(\mathbf{y}) - g\| \leq (2N)^{-1/2}|g|_{\mathbb{H}}\sqrt{\boldsymbol{\delta}_{\max}} \leq (2N)^{-1/2}\sqrt{\boldsymbol{\delta}_{\max}}$. This completes the proof of (12.102). □

Armed with Fact 12.3.1 and Lemmas 12.3.3–12.3.4, we recall (12.79), (12.89), (12.91), and fix an $N \geq 1$. For $n \geq 1$, $0 < a < 1$, $t \in [0, 1-a]$ and $\boldsymbol{\delta} \in \mathcal{D}_N$, we set

$$\mathbf{z}_{n,\boldsymbol{\delta}}(a;t) = \mathcal{P}_{N,\boldsymbol{\delta}}\left(\frac{\xi_n(a;t;\cdot)}{\sqrt{2a\log_+(1/a)}}\right) \in \mathbb{R}^N. \tag{12.103}$$

By combining (12.89) with (12.94) and (12.103), we observe that

$$\mathbf{z}_{n,\boldsymbol{\delta}}(a;t) = \frac{\sqrt{N}}{\sqrt{2na\log_+(1/a)}} \times \begin{bmatrix} \{\alpha_n(t+at_1(\boldsymbol{\delta})) - \alpha_n(t+at_0(\boldsymbol{\delta}))\}/\sqrt{\delta_1} \\ \vdots \\ \{\alpha_n(t+at_N(\boldsymbol{\delta})) - \alpha_n(t+at_{N-1}(\boldsymbol{\delta}))\}/\sqrt{\delta_N} \end{bmatrix}. \tag{12.104}$$

Set, for convenience,

$$\mathbf{z}^*_{n,\boldsymbol{\delta}}(a;t) = \frac{\sqrt{N}}{\sqrt{2a\log_+(1/a)}} \times \begin{bmatrix} \alpha_n(t+at_1(\boldsymbol{\delta})) - \alpha_n(t+at_0(\boldsymbol{\delta})) \\ \vdots \\ \alpha_n(t+at_N(\boldsymbol{\delta})) - \alpha_n(t+at_{N-1}(\boldsymbol{\delta})) \end{bmatrix}. \tag{12.105}$$

Recall the definition (12.81) of $\zeta_{n;a;\boldsymbol{\delta}}$. In view of (12.105), we may write, for each $0 < a < 1$ and $t \in [0, 1-a]$, the distributional equality

$$\zeta_{n;a;\boldsymbol{\delta}} \stackrel{d}{=} \mathbf{z}^*_{n,\boldsymbol{\delta}}(a;t). \tag{12.106}$$

We infer from (12.104) and (12.105) the inequality

$$\|\mathbf{z}^*_{n,\boldsymbol{\delta}}(a;t)\| \leq \|\mathbf{z}_{n,\boldsymbol{\delta}}(a;t)\|/\sqrt{\boldsymbol{\delta}_{\min}}\,. \tag{12.107}$$

Below, we let C_2, $n_1(\cdot)$, $c_1(\cdot)$ and $a_1(\cdot)$ be as in Fact 12.3.2.

Lemma 12.3.5 *For each* $0 < \varepsilon \leq 1$*, and for all* $n \geq n_1(\varepsilon)$ *and* $a > 0$ *fulfilling*

$$na/\log n \geq c_1(\varepsilon) \quad \textit{and} \quad a \leq a_1(\varepsilon), \tag{12.108}$$

we have, for all $t \in [0, 1 - a]$*,*

$$\mathbb{P}\left(\mathbf{z}_{n,\delta}(a;t) \notin \mathbf{B}_N^{\varepsilon}\right) \leq C_2 a^{1+\left(\varepsilon\sqrt{\delta_{min}}\right)/(2N)}. \tag{12.109}$$

Proof By (12.100), for any $\phi \in B([0, 1])$, $g \in \mathbb{S}$ and $\epsilon > 0$, we have the implication

$$\|\phi - g\| \leq \epsilon \;\Rightarrow\; \|\mathcal{P}_{N,\delta}(\phi) - \mathcal{P}_{N,\delta}(g)\| = \|\mathcal{P}_{N,\delta}(\phi - g)\| \leq 2N\epsilon/\sqrt{\delta_{\min}},$$

which is equivalent to the implication

$$\|\mathcal{P}_{N,\delta}(\phi) - \mathcal{P}_N(g)\| > 2N\epsilon/\sqrt{\delta_{\min}} \;\Rightarrow\; \|\phi - g\| > \epsilon. \tag{12.110}$$

We recall from (12.101) that $\mathcal{P}_{N,\delta}(\mathbb{S}) = \mathbf{B}_N$. Thus, by setting $\mathbf{z} = \mathcal{P}_{N,\delta}(g)$ in (12.110), and letting g vary in $\mathbb{S}$ we obtain the implication

$$\left\{\|\mathcal{P}_{N,\delta}(\phi) - \mathbf{z}\| > 2N\epsilon/\sqrt{\delta_{\min}} \;:\forall\, \mathbf{z} \in \mathbf{B}_N\right\} \;\Rightarrow\; \left\{\|\phi - g\| > \epsilon \;:\forall\, g \in \mathbb{S}\right\},$$

which may be rewritten into

$$\left\{\mathcal{P}_{N,\delta}(\phi) \notin \mathbf{B}_N^{2N\epsilon/\sqrt{\delta_{\min}}}\right\} \;\Rightarrow\; \left\{\phi \notin \mathbb{S}^{\epsilon}\right\}. \tag{12.111}$$

Recalling the definition (12.103) of $\mathbf{z}_{n,\delta}(a;t)$, by setting $\varepsilon = 2N\epsilon/\sqrt{\delta_{\min}}$ and $\phi = \xi_n(a;t;\cdot)/\sqrt{2a\log_+(1/a)}$ in (12.111), we conclude our proof by an application of Fact 12.3.2. □

12.3.5 Proof of Proposition 12.3.2

Fix an $0 < \varepsilon \leq 1$. In view of (12.106) and (12.33), whenever

$$\sqrt{\delta_{\min}} \geq \frac{1 + \frac{1}{2}\varepsilon}{1 + \varepsilon}, \tag{12.112}$$

we have, for $0 < a < 1$ and $0 \leq t \leq 1 - a$,

$$\begin{aligned} &\mathbb{P}\left(\boldsymbol{\zeta}_{n;a;\delta} \notin \mathbf{B}_N^{\varepsilon}\right) = \mathbb{P}\left(\|\boldsymbol{\zeta}_{n;a;\delta}\| > 1 + \varepsilon\right) \qquad (12.113)\\ &= \mathbb{P}\left(\|\mathbf{z}_{n;\delta}^*(a;t)\| > 1 + \varepsilon\right) \leq \mathbb{P}\left(\|\mathbf{z}_{n;\delta}(a;t)\| > (1 + \varepsilon)\sqrt{\delta_{\min}}\right)\\ &\leq \mathbb{P}\left(\|\mathbf{z}_{n;\delta}(a;t)\| > 1 + \tfrac{1}{2}\varepsilon\right) = \mathbb{P}\left(\mathbf{z}_{n;\delta}(a;t) \notin \mathbf{B}_N^{\varepsilon/2}\right). \end{aligned}$$

The assumption that $0 < \varepsilon \leq 1$, when combined with (12.112) implies that

$$\sqrt{\boldsymbol{\delta}_{\min}} \geq \frac{3}{4} > \frac{1}{2}.$$

By an application of Lemma 12.3.5 with the formal replacement of ε by $\varepsilon/2$, we see that, for all $n \geq n_0(\varepsilon) := n_1(\varepsilon/2)$ and $a > 0$ fulfilling

$$na/\log n \geq c_0(\varepsilon) := c_1(\varepsilon/2) \quad \text{and} \quad a \leq a_0(\varepsilon) := a_1(\varepsilon/2), \tag{12.114}$$

we have, for all $t \in [0, 1-a]$,

$$\mathbb{P}\left(\mathbf{z}_{n,\boldsymbol{\delta}}(a;t) \notin \mathbf{B}_N^{\varepsilon/2}\right) \leq C_0 a^{1+(\varepsilon\sqrt{\boldsymbol{\delta}_{\min}})/(4N)} \leq C_0 a^{1+\varepsilon/(8N)}. \tag{12.115}$$

By (12.113), this yields (12.84), with $C_0 := C_2$, and completes the proof of Proposition 12.3.2. □

12.3.6 Inner Bounds

The following fact is a version of Proposition 3 of Deheuvels and Ouadah [10], taken with $|\mathcal{I}| = \sum_{k=1}^{K} a_k$.

Fact 12.3.3 *For each $g \in \mathbb{S}$ such that $0 < |g|_{\mathbb{H}} < 1$, and $0 < \varepsilon < \frac{1}{2}|g|_{\mathbb{H}}$, there exist an $a_2(\varepsilon, g)$, together with $n_2(\varepsilon) < \infty$ and $c_2(\varepsilon)$, depending upon ε only, such that the following holds. Let, for $K \geq 1$, $t_1, \ldots, t_K \in [0,1]$, and $0 < a_1, \ldots, a_k < 1$, be such that the intervals $(t_k, t_k + a)$, $k = 1, \ldots, K$, are disjoint and in $[0,1]$, with $\sum_{k=1}^{K} a_k \leq \frac{1}{2}$. Then, whenever*

$$n \geq n_2(\varepsilon), \quad c_2(\varepsilon) n^{-1} \log n \leq a_1 \ldots, a_K \leq a_2(\varepsilon, g), \tag{12.116}$$

we have

$$\mathbb{P}\left(\bigcap_{k=1}^{K} \left\{ \frac{\xi_n(a_k; t_k; \cdot)}{\sqrt{2a_k \log_+(1/a_k)}} \notin \mathcal{N}_\varepsilon(g) \right\}\right) \leq 2\exp\left(-\frac{1}{4}\sum_{k=1}^{K} a_k^{1-\varepsilon/2}\right). \tag{12.117}$$

Fix any $\mathbf{z} \in \mathbf{B}_N$, such that $0 < \|\mathbf{z}\| < 1$, and set $g := \mathcal{Q}_{N;\boldsymbol{\delta}}(\mathbf{z})$. Fix $a > 0$ and $t \in [0, 1-a]$, and set, as in (12.103),

$$\phi := \frac{\xi_n(a;t;\cdot)}{\sqrt{2a\log_+(1/a)}} \quad \text{and} \quad \mathbf{z}_{n,\boldsymbol{\delta}}(a;t) = \mathcal{P}_{N,\boldsymbol{\delta}}(\phi) \in \mathbb{R}^N. \tag{12.118}$$

As follows from (12.99) and (12.99), we have $\mathcal{P}_{N;\boldsymbol{\delta}}(g) = \mathbf{z}$ and

$$0 < |g|_{\mathbb{H}} = \|\mathbf{z}\| < 1.$$

Therefore, we infer from the linearity of $\mathcal{P}_{N;\boldsymbol{\delta}}$ and (12.100) that

$$\|\mathbf{z}_{n,\boldsymbol{\delta}}(a;t) - \mathbf{z}\| = \left\|\mathcal{P}_{N,\boldsymbol{\delta}}(\phi) - \mathcal{P}_{N,\boldsymbol{\delta}}(g)\right\| = \left\|\mathcal{P}_{N,\boldsymbol{\delta}}(\phi - g)\right\|$$
$$\leq \frac{2N}{\sqrt{\boldsymbol{\delta}_{\min}}}\|\phi - g\| = \frac{2N}{\sqrt{\boldsymbol{\delta}_{\min}}}\left\|\frac{\xi_n(a;t;\cdot)}{\sqrt{2a\log_+(1/a)}} - g\right\|.$$

We have therefore the implication, for an arbitrary $\varepsilon > 0$,

$$\left\|\frac{\xi_n(a;t;\cdot)}{\sqrt{2a\log_+(1/a)}} - g\right\| \leq \varepsilon \Rightarrow \|\mathbf{z}_{n,\boldsymbol{\delta}}(a;t) - \mathbf{z}\| \leq \frac{2N\varepsilon}{\sqrt{\boldsymbol{\delta}_{\min}}}$$

which is readily shown to be equivalent to

$$\left\{\|\mathbf{z}_{n,\boldsymbol{\delta}}(a;t) - \mathbf{z}\| > \frac{2N\varepsilon}{\sqrt{\boldsymbol{\delta}_{\min}}}\right\} \subseteq \left\{\frac{\xi_n(a;t;\cdot)}{\sqrt{2a\log_+(1/a)}} \notin \mathcal{N}_\varepsilon(g)\right\}. \tag{12.119}$$

Recalling (12.104), and the definition (12.105) of $\mathbf{z}^*_{n,\boldsymbol{\delta}}(a;t)$, set, for $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_N)$,

$$\mathbf{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_N \end{bmatrix}, \quad \mathbf{z}^*_{n,\boldsymbol{\delta}}(a;t) = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} \quad \text{and} \quad \mathbf{z}_{n,\boldsymbol{\delta}}(a;t) = \begin{bmatrix} y_1/\sqrt{\delta_1} \\ \vdots \\ y_N/\sqrt{\delta_N} \end{bmatrix}.$$

By combining the triangle inequality with $\|\mathbf{z}\| < 1$, we see that

$$\|\mathbf{z}_{n,\boldsymbol{\delta}}(a;t) - \mathbf{z}\| = \left\{\sum_{j=1}^N (y_j/\sqrt{\delta_j} - z_j)^2\right\}^{1/2} \tag{12.120}$$
$$\geq \left\{\sum_{j=1}^N (y_j/\sqrt{\delta_j} - z_j/\sqrt{\delta_j})^2\right\}^{1/2} - \left\{\sum_{j=1}^N (z_j\sqrt{\delta_j} - z_j)^2\right\}^{1/2}$$
$$\geq \frac{1}{\sqrt{\boldsymbol{\delta}_{\max}}}\|\mathbf{z}^*_{n,\boldsymbol{\delta}}(a;t) - \mathbf{z}\| - \|\mathbf{z}\|\left\{\left(1 - \frac{1}{\sqrt{\boldsymbol{\delta}_{\max}}}\right) \vee \left(\frac{1}{\sqrt{\boldsymbol{\delta}_{\min}}} - 1\right)\right\}$$
$$\geq \frac{1}{\sqrt{\boldsymbol{\delta}_{\max}}}\|\mathbf{z}^*_{n,\boldsymbol{\delta}}(a;t) - \mathbf{z}\| - \left\{\left(1 - \frac{1}{\sqrt{\boldsymbol{\delta}_{\max}}}\right) \vee \left(\frac{1}{\sqrt{\boldsymbol{\delta}_{\min}}} - 1\right)\right\}.$$

Thus, if we assume that

$$\frac{1}{\sqrt{\delta_{\max}}} \geq 1 - N\varepsilon \quad \text{and} \quad \frac{1}{\sqrt{\delta_{\min}}} \leq 1 + N\varepsilon, \tag{12.121}$$

we infer from (12.120) that

$$\|\mathbf{z}_{n,\delta}(a;t) - \mathbf{z}\| \geq \frac{1}{\sqrt{\delta_{\max}}} \|\mathbf{z}^*_{n,\delta}(a;t) - \mathbf{z}\| + N\varepsilon.$$

This, when combined with (12.119), shows that

$$\left\{ \|\mathbf{z}^*_{n,\delta}(a;t) - \mathbf{z}\| > 3N\varepsilon\sqrt{\frac{\delta_{\max}}{\delta_{\min}}} \right\} \subseteq \left\{ \frac{\xi_n(a;t;\cdot)}{\sqrt{2a\log_+(1/a)}} \notin \mathcal{N}_\varepsilon(g) \right\}. \tag{12.122}$$

In view of (12.106), we infer from (12.122) the relation

$$\begin{aligned} &\bigcap_{k=1}^{K} \left\{ \|\boldsymbol{\zeta}^{(k)}_{n;a_k;\delta_k} - \mathbf{z}\| > 3N\varepsilon\sqrt{\frac{\delta_{\max}}{\delta_{\min}}} \right\} \\ &\subseteq \bigcap_{k=1}^{K} \left\{ \frac{\xi_n(a_k;t_k;\cdot)}{\sqrt{2a_k\log_+(1/a_k)}} \notin \mathcal{N}_\varepsilon(g) \right\} \end{aligned} \tag{12.123}$$

Now, we infer from (12.121) that, whenever $N\varepsilon \leq \frac{1}{2}$,

$$\sqrt{\frac{\delta_{\max}}{\delta_{\min}}} \leq \frac{1+N\varepsilon}{1-N\varepsilon} \leq 3.$$

Thus, by (12.123), we have

$$\begin{aligned} &\mathbb{P}\left(\bigcap_{k=1}^{K} \left\{ \|\boldsymbol{\zeta}^{(k)}_{n;a_k;\delta_k} - \mathbf{z}\| > 9N\varepsilon \right\} \right) \\ &\leq \mathbb{P}\left(\bigcap_{k=1}^{K} \left\{ \frac{\xi_n(a_k;t_k;\cdot)}{\sqrt{2a_k\log_+(1/a_k)}} \notin \mathcal{N}_\varepsilon(g) \right\} \right) \leq 2\exp\left(-\tfrac{1}{4}\sum_{k=1}^{K} a_k^{1-\varepsilon/2} \right). \end{aligned} \tag{12.124}$$

The remainder of the proof is given by routine arguments which we omit. □

References

1. C.R. Adams, J.A. Clarkson, On definitions of bounded variation for functions of two variables. Trans. Am. Math. Soc. **35**, 824–854 (1933)
2. N. Bouleau, On effective computation of expectations in large or infinite dimensions. J. Comput. Appl. Math. **31**, 23–34 (1990)
3. J.A. Clarkson, C. Raymond Adams, Properties of functions $f(x, y)$ of bounded variation. Trans. Am. Math. Soc. **36**, 711–730 (1933)
4. P. Deheuvels, Laws of the iterated logarithm for density estimators, in *Nonparametric Functional Estimation and Related Topics* (Kluwer, Dordrecht, 1991), pp. 19–29
5. P. Deheuvels, Functional laws of the iterated logarithm for large increments of empirical and quantile processes. Stoch. Process. Appl. **43**, 133–163 (1992)
6. P. Deheuvels, One bootstrap suffices to generate sharp uniform bounds in functional estimation. Kybernetika **47**, 855–865 (2011)
7. P. Deheuvels, J.H.J. Einmahl, Functional limit laws for the increments of Kaplan-Meier product-limit processes and applications. Ann. Probab. **28**, 1301–1335 (2000)
8. P. Deheuvels, D.M. Mason, Functional laws of the iterated logarithm for the increments of empirical and quantile processes. Ann. Probab. **20**, 1248–1287 (1992)
9. P. Deheuvels, D.M. Mason, General asymptotic confidence bands based on kernel-type function estimators. Stat. Infer. Stoch. Process **7**, 225–277 (2004)
10. P. Deheuvels, S. Ouadah, Uniform-in-bandwidth functional limit laws. J. Theor. Probab. **26**(3), 697–721 (2013)
11. J. Dony, Nonparametric regression estimation-An empirical process approach to uniform in bandwidth consistency of kernel-type estimators and conditional U-statistics. Doctoral Dissertation. Vrije Universiteit Brussel, Brussels, 2008
12. J. Dony, U. Einmahl, Weighted uniform consistency of kernel density estimators with general bandwidth sequences. Electron. J. Probab. **11**, 844–859 (2006)
13. J. Dony, U. Einmahl, Uniform in bandwidth consistency of kernel-type estimators at a fixed point. Inst. Math. Stat. Collect. **5**, 308–325 (2009)
14. J. Dony, D.M. Mason, Uniform in bandwidth consistency of conditional U-statistics. Bernoulli **14**(4), 1108–1133 (2008)
15. J. Dony, U. Einmahl, D.M. Mason, Uniform in bandwidth consistency of local polynomial regression function estimators. Aust. J. Stat. **35**, 105–120 (2006)
16. U. Einmahl, D.M. Mason, An empirical process approach to the uniform consistency of kernel-type function estimators. J. Theor. Probab. **13**, 1–37 (2000)
17. U. Einmahl, D.M. Mason, Uniform in bandwidth consistency of kernel-type function estimators. Ann. Stat. **33**, 1380–1403 (2005)
18. G.H. Hardy, On double Fourier series, and especially those which represent the double zeta-function with real and incommensurable parameters. Q. J. Math. **37**, 53–89 (1905)
19. M. Krause, Über Mittelwertsätze in Gebiete der Doppelsummen und Doppelintegrale. Leipziger Ber. **55**, 239–263 (1903)
20. D.M. Mason, A strong limit theorem for the oscillation modulus of the uniform empirical process. Stoch. Process. Appl. **17**, 127–136 (1984)
21. D.M. Mason, A uniform functional law of the logarithm for the local empirical process. Ann. Probab. **32**, 1391–1418 (2004)
22. D.M. Mason, Proving consistency of non-standard kernel estimators. Stat. Infer. Stoch. Process. **20**(2), 151–176 (2012)
23. D.M. Mason, J. Swanepoel, A general result on the uniform in bandwidth consistency of kernel-type function estimators. Test **20**, 72–94 (2011)
24. D.M. Mason, G.R. Shorack, J.A. Wellner, Strong limit theorems for oscillation moduli of the empirical process. Z. Wahrscheinlichkeitstheorie und Verwandte Geb. **65**, 93–97 (1983)
25. E.A. Nadaraya, On estimating regression. Theor. Probab. Appl. **9**, 141–142 (1964)

26. H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*. CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 63 (SIAM, Philadelphia, 1992)
27. D. Nolan, J.S. Marron, Uniform consistency and location adaptive delta-sequence estimators. Probab. Theory Relat. Fields **80**, 619–632 (1989)
28. G. Pagès, Y.-J. Xiao, Sequences with low discrepancy and pseudo-random numbers: theoretical results and numerical tests. J. Stat. Comput. Simul. **56**, 163–188 (1997)
29. E. Parzen, On the estimation of a probability density function and mode. Ann. Math. Stat. **33**, 1065–1076 (1962)
30. M. Rosenblatt, Remarks on some nonparametric estimates of a density function. Ann. Math. Stat. **27**, 832–837 (1956)
31. B. Schweizer, E.F. Wolff, On nonparametric measure of dependence for random variables. Ann. Stat. **6**, 177–184 (1981)
32. B. Silverman, Weak and strong consistency of the kernel estimate of a density and its derivatives. Ann. Stat. **6**, 177–184 (1978) (Addendum: (1980). **8**, 1175–1176)
33. V. Strassen, An invariance principle for the law of the iterated logarithm. Z. Wahrscheinlichkeitstheorie und Verwandte Geb. **3**, 211–226 (1964)
34. W. Stute, The oscillation behavior of empirical processes: the multivariate case. Ann. Probab. **12**, 361–379 (1984)
35. I. van Keilegom, D. Varron, Uniform in bandwidth exact rates for a class of kernel estimators. Ann. Inst. Stat. Math. **63**(6), 1077–1102 (2011)
36. D. Varron, Lois fonctionnelles uniforme du logarithme itéré pour les accroissements du processus empirique généralisé, in *Lois limites de type Chung-Mogulskii pour le processus empirique uniforme local*. Doctoral Dissertation, Université Pierre et Marie Curie, Paris, Dec. 17, 2004
37. D. Varron, A limited in bandwidth uniformity for the functional limit law of the increments of the empirical process. Electron. J. Stat. **2**, 1043–1064 (2008)
38. V. Viallon, Functional limit laws for the increments of the quantile process with applications. Electron. J. Stat. **1**, 496–518 (2007)
39. G. Vitali, Sui gruppi di punti e sulle funzioni di variabili reali. Atti Accad. Sci. Torino. **43**, 229–246 (1908)
40. G.S. Watson, Smooth regression analysis. Sankhyā Indian J. Stat. A **26**, 359–372 (1964)