The Annals of Probability
1994, Vol. 22, No. 3, 1619–1661

FUNCTIONAL LAWS OF THE ITERATED LOGARITHM FOR LOCAL EMPIRICAL PROCESSES INDEXED BY SETS

BY PAUL DEHEUVELS[1] AND DAVID M. MASON[2]

Université Paris VI and University of Delaware

We introduce the notion of a multivariate local (and tail) empirical process indexed by sets and establish a number of functional laws of the iterated logarithm for such processes. This leads to a unified methodology to study the almost sure behavior of various statistics which are local functionals of the empirical distribution. Such statistics include density estimators and the Bahadur–Kiefer representation.

1. Introduction and outline of results. Let $U_1, U_2, \ldots$ be independent random variables uniformly distributed on $[0,1]^d$ with $d \geq 1$. Let $\mathbb{B}$ be the class of Borel subsets of $[0,1]^d$. Denote by $\lambda(\cdot)$ the Lebesgue measure in $\mathbb{R}^d$, and $\lambda_0(\cdot)$ the restriction of $\lambda(\cdot)$ to $[0,1]^d$. Consider the *uniform empirical measure indexed by* $\mathbb{B}$, defined by

$$\lambda_n(B) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(U_i \in B), \qquad B \in \mathbb{B}, \quad \text{for } n \geq 1, \tag{1.1}$$

with $\mathbf{1}$ denoting the indicator function. For any given subclass $\mathbb{D}$ of $\mathbb{B}$, introduce the *uniform empirical process indexed by* $\mathbb{D}$, defined by

$$\alpha_n(D) = n^{1/2}\{\lambda_n(D) - \lambda(D)\}, \qquad D \in \mathbb{D}. \tag{1.2}$$

We now consider the special class $\mathbb{D}$ defined as follows. Let $\mathbf{t} \in [0,1]^d$ be fixed and let $\mathbb{C}$ be a class of Borel subsets of $[a,b]^d$ for some $a < b$ with $b - a = 1$. Let $\mathbf{a} = (a, \ldots, a) \in \mathbb{R}^d$, set $\mathbb{D} = \{\mathbf{t} + C\colon C \in \mathbb{C}\}$ and assume further that:

(C.1) $\mathbf{t} + C \subseteq [0,1]^d$ for all $C \in \mathbb{C}$ and, for all $h > 0$ sufficiently small, $\mathbf{t} + h^{1/d}[a,b]^d \subseteq [0,1]^d$.
(C.2) $\mathbb{C}$ is *countably generated* (see Section 2 for definition).
(C.3) $\mathbb{C} - \mathbf{a}$ is a λ_0-*Donsker class* (see Section 2 for definition).
(C.4) For every $\frac{1}{2} \leq h_1 < h_2 \leq 1$, $h_1\mathbb{C} \subseteq h_2\mathbb{C} \subseteq \mathbb{C}$.

We note for future reference that, for every $C \in \mathbb{C}$,

$$\lim_{h \uparrow 1} d_\lambda(hC, C) = 0, \tag{1.3}$$

Received September 1992.
[1]Partially supported by a U.S.–France NSF–CNRS grant.
[2]Partially supported by an NSF grant.
AMS 1991 *subject classifications.* 60F05, 60F15, 62E20, 62G30.
Key words and phrases. Empirical measures, strong laws, functional laws of the iterated logarithm, Donsker classes, tail empirical processes, density estimation, Bahadur–Kiefer representation.

where d_λ denotes the *pseudometric* defined on $\mathbb{C}$ by $d_\lambda(A,B) = \lambda(A \triangle B)$, with $A \triangle B = (A - B) \cup (B - A)$ denoting the *symmetric difference* between A and B.

(C.5) For every finite subset $\{C_1, \dots, C_M\} \subseteq \mathbb{C}$ with each $\lambda(C_i) > 0$, the class $\mathbb{C}$ can be enlarged, if necessary, to include a finite set of disjoint $\{D_1, \dots, D_N\} \subseteq \mathbb{C}$ with each $\lambda(D_i) > 0$ such that for each C_i there exists a subset $J \subseteq \{1, \dots, N\}$ for which $\cup_{j \in J} D_j = C_i$.

We postpone until Section 2 a discussion of the meaning and implications of these assumptions.

Also, we choose a sequence of positive constants $\{k(n),\ n \geq 1\}$, let $\log^+ u = \log(\max(u, e))$ and let $\log_2 u = \log^+(\log^+ u)$. We assume that this sequence satisfies assumptions among the following, listed below.

(K.1) $0 < k(n) \leq n,\ k(n) \uparrow \infty$ and $n^{-1}k(n) \downarrow 0$ as $n \uparrow \infty$.
(K.2) $(\log_2 n)^{-1}k(n) \to \infty$ as $n \to \infty$.
(K.3) $(\log_2 n)^{-1}k(n) \to c \in (0, \infty)$ as $n \to \infty$.

By the *local empirical process at* **t** *indexed by* $\mathbb{C}$, we shall mean the process

$$\begin{aligned}\Theta_n(C) &= \Theta_n\left(C, \frac{k(n)}{n}\right) = \Theta_n\left(C, \mathbf{t}, \frac{k(n)}{n}\right) = \left(\frac{k(n)}{n}\right)^{-1/2} \alpha_n\left(\mathbf{t} + \left(\frac{k(n)}{n}\right)^{1/d} C\right) \\ &= \left(k(n)\right)^{-1/2}\left\{\sum_{i=1}^{n} \mathbf{1}\left(U_i \in \mathbf{t} + \left(\frac{k(n)}{n}\right)^{1/d} C\right) - k(n)\lambda(C)\right\}, \qquad C \in \mathbb{C}.\end{aligned} \tag{1.4}$$

The limiting properties of the local empirical process have been investigated at length in the special case when $d = 1$, $\mathbf{t} = 0$ and with $\mathbb{C}$ being the class of all intervals of the form $[0, s]$ for $0 \leq s \leq 1$. When $\mathbf{t} = 0$, this process is usually called the *tail empirical process*. We begin by briefly reviewing the main strong laws which have been obtained in this case. Set $\xi_n(s) = \Theta_n([0, s], 0, k(n)/n)$, $\eta_n(s) = \lambda_n([0, n^{-1}k(n)s])$ for $0 \leq s \leq 1$ and denote by $B(0, 1)$ the space of bounded functions on $[0, 1]$. Introduce the function h defined by

$$h(x) = \begin{cases} x \log x - x + 1, & \text{for } x > 0, \\ 1, & \text{for } x = 0, \\ \infty, & \text{for } x < 0. \end{cases} \tag{1.5}$$

We have the following results.

THEOREM A [Mason (1988)]. *Under* (K.1) *and* (K.2), *the sequence of functions* $\{(2\log_2 n)^{-1/2}\xi_n,\ n \geq 1\}$ *is almost surely relatively compact in* $B(0, 1)$ *with respect to the uniform metric, with limit set* $\mathbb{S}$ *composed of all absolutely continuous functions* f *on* $[0, 1]$ *with* $f(0) = 0$, *and having Lebesgue derivative* $\dot{f}$ *satisfying*

$$\int_0^1 \dot{f}(t)^2\, dt \leq 1. \tag{1.6}$$

THEOREM B [Deheuvels and Mason (1990)]. *Under* (K.1) *and* (K.3), *the sequence of functions* $\{(\log_2 n)^{-1} n\eta_n,\ n \geq 1\}$ *is almost surely relatively compact in* $B(0,1)$ *with respect to the uniform metric, with limit set composed of all absolutely continuous functions* f *on* $[0,1]$ *with* $f(0) = 0$, *and having Lebesgue derivative* $\dot{f}$ *satisfying*

$$\int_0^1 ch\big(c^{-1}\dot{f}(s)\big)\,ds \leq 1. \tag{1.7}$$

Results such as those given in Theorems A and B proved to be very useful to describe the limiting behavior of *tail statistics*, that is, statistics based on the upper (respectively lower) extreme values of a univariate sample. We refer to Deheuvels and Mason (1991), and the references therein, for examples of such applications.

The first purpose of this paper is to provide a unified approach to the study of the almost sure behavior of multivariate *tail* or *local statistics* (such local statistics include as examples the multivariate density estimators which we consider in Section 3).

Toward this end, we first establish in the more general setting of local empirical processes indexed by subsets of $\mathbb{R}^d$ a greatly extended version of the functional law of the iterated logarithm stated in Theorem A. We introduce the class $B(\mathbb{C})$ of all bounded functions defined on $\mathbb{C}$ and endow this space with the topology of uniform convergence on $\mathbb{C}$. We have the following theorem.

THEOREM 1.1. *Under* (K.1), (K.2) *and* (C.1), (C.2), (C.3), (C.4) *and* (C.5), *the set of functions* $\{(2\log_2 n)^{-1/2}\Theta_n(C),\ C \in \mathbb{C}\}$ *is almost surely relatively compact in* $B(\mathbb{C})$, *with limit set equal to*

$$\mathbb{S}(\mathbb{C}) := \left\{ f \in B(\mathbb{C}) \colon f(C) = \int_C \phi(\mathbf{s})\,d\lambda(\mathbf{s}) \text{ with } \int_{\mathbb{R}^d} \phi^2(\mathbf{s})\,d\lambda(\mathbf{s}) \leq 1 \right\}. \tag{1.8}$$

In Section 2 we discuss the regularity conditions that we must impose upon $\mathbb{C}$ for the validity of Theorem 1.1, and then we provide a proof of this theorem. Our methods are entirely based on the general theory of empirical processes and do not make use of strong invariance principles. We present applications of these results in Section 3. In particular, we establish laws of the iterated logarithm for the pointwise strong consistency of nonparametric density estimators which improve upon the results of Hall (1981).

The second purpose of this paper is to investigate extensions of Theorem 1.1 obtained by considering the limiting behavior of the empirical process in the neighborhood of more than one fixed point. We now present the corresponding results.

Let $\mathbf{t}_0 = \mathbf{0} = (0,\dots,0)$ and set $k_0(n) = n$ for $n \geq 1$. Assume that $\mathbf{t}_1,\dots,\mathbf{t}_N \in [0,1]^d$ are fixed and distinct. Consider N sequences of positive constants

$\{k_j(n),\ n \geq 1\}$, $j = 1, \ldots, N$, and $N+1$ classes $\mathbb{C}_j$ of Borel subsets of $[a_j, b_j]^d$ with $b_j - a_j = 1$ for $j = 0, \ldots, N$. Set

$$
\begin{aligned}
w_{n,j}(C) &= (2\log_2 n)^{-1/2}\Theta_n\left(C, \mathbf{t}_j, \frac{k_j(n)}{n}\right) \\
&= \left(2\frac{k_j(n)}{n}\log_2 n\right)^{-1/2}\alpha_n\left(\mathbf{t}_j + \left(\frac{k_j(n)}{n}\right)^{1/d} C\right), \\
&\qquad \text{for } C \in \mathbb{C}_j,\ j = 0, \ldots, N.
\end{aligned}
\tag{1.9}
$$

Our second main result may be now stated as follows.

THEOREM 1.2. *Assume that* $\mathbb{C}_0$ *satisfies* (C.2) *and* (C.3), *that, for each* $j = 1, \ldots, N$, *the sequence* $k(n) = k_j(n)$ *satisfies assumptions* (K.1) *and* (K.2) *and that, for each* $j = 1, \ldots, N$, $\mathbb{C} = \mathbb{C}_j$ *and* $\mathbf{t} = \mathbf{t}_j$ *satisfy conditions* (C.1), (C.2), (C.3), (C.4) *and* (C.5). *Then the sequence* $\{(w_{n,0}, \ldots, w_{n,N}),\ n \geq 1\}$ *is almost surely relatively compact in* $\mathbb{S}(\mathbb{C}_0) \times \cdots \times \mathbb{S}(\mathbb{C}_N)$ *with limit set equal to*

$$
\begin{aligned}
&\Bigg\{(w_0, \ldots, w_N) \in \prod_{j=0}^{N} \mathbb{S}(\mathbb{C}_j)\colon\ w_j(C) = \int_C \phi_j(\mathbf{s})\, d\lambda(\mathbf{s}),\ j = 0, \ldots, N, \\
&\qquad \int_{[0,1]^d} \phi_0(\mathbf{s})\, d\lambda(\mathbf{s}) = 0,\ \sum_{j=0}^{N} \int_{\mathbb{R}^d} \phi_j^2(\mathbf{s})\, d\lambda(\mathbf{s}) \leq 1\Bigg\}.
\end{aligned}
\tag{1.10}
$$

The description of the limiting behavior of $w_{n,0}(C) = (2\log_2 n)^{-1/2}\alpha_n(C)$, $C \in \mathbb{C}_0$, which follows from Theorem 1.2 is a special case of the results of Kuelbs and Dudley (1980). The first functional law of this type for $w_{n,0}(\cdot)$, $d = 1$ and $\mathbb{C}_0 = \{[0, s],\ 0 \leq s \leq 1\}$ is due to Finkelstein (1971).

We will prove Theorem 1.2 in Section 2. Among the applications of this theorem given in Section 3, we will obtain a simple derivation of the pointwise Bahadur–Kiefer representation of the empirical process [Bahadur (1967) and Kiefer (1970)], extending the methodology of Deheuvels and Mason (1992).

In view of Theorem B, the limiting behavior of $\Theta_n(\cdot)$ under assumptions (K.1) and (K.3) is much different. A complete study of this case is presented in Deheuvels and Mason (1994).

2. Strassen-type functional LILs for local empirical processes.

2.1. *Preliminary results and notation.* We inherit the notation of Section 1 and assume that the random variables $U_1, U_2, \ldots$ are defined on the same probability space $(\Omega, \mathbb{F}, P)$. We recall that $\mathbb{C}$ is a class of Borel subsets of $[a, b]^d$ for some $a < b$ with $b - a = 1$, and satisfying

(C.1) $\mathbf{t} + C \subseteq [0,1]^d$ for all $C \in \mathbb{C}$ and, for all $h > 0$ sufficiently small, $\mathbf{t} + h^{1/d}[a,b]^d \subseteq [0,1]^d$.

(C.2) $\mathbb{C}$ is *countably generated*; that is, there exists a countable subclass $\mathbb{G}$ of $\mathbb{C}$ such that, for any $D \in \mathbb{C}$, there exists a sequence $\{D_n,\ n \geq 1\}$ of $\mathbb{G}$ satisfying, for all $x \in \mathbb{R}^d$,

$$\lim_{n \to \infty} \mathbf{1}(x \in D_n) = \mathbf{1}(x \in D). \tag{2.1}$$

In the sequel we will say that a class $\mathbb{G}$ of this type is a *countable generating subclass* of $\mathbb{C}$.

(C.3) $\mathbb{C} - \mathbf{a}$ is a λ_0-*Donsker class.*

Following Dudley (1978), Kuelbs and Dudley [(1980), page 406] and Gaenssler [(1983), pages 46, 65 and 113], a class $\mathbb{D}$ is said to be a λ_0-Donsker class if the normalized empirical measures $\{\alpha_n(D),\ D \in \mathbb{D}\}$ as defined in (1.2) converge in law [in the sense of Dudley (1978), Section 1] to a Gaussian measure $\{G_\lambda(D),\ D \in \mathbb{D}\}$. We refer to Theorem B, page 113 in Gaenssler (1982), and to Giné and Zinn (1986), Talagrand (1988) and Alexander (1987) for further details concerning Donsker classes.

(C.4) For every $\frac{1}{2} \leq h_1 < h_2 \leq 1$, $h_1\mathbb{C} \subseteq h_2\mathbb{C} \subseteq \mathbb{C}$.

(C.5) For every finite subset $\{C_1, \dots, C_M\} \subseteq \mathbb{C}$ with each $\lambda(C_i) > 0$, the class $\mathbb{C}$ can be enlarged, if necessary, to include a finite set of disjoint $\{D_1, \dots, D_M\} \subseteq \mathbb{C}$ with each $\lambda(D_i) > 0$ such that for each C_i there exists a subset $J \subseteq \{1, \dots, N\}$ for which $\cup_{j \in J} D_j = C_i$.

We start by recalling some useful facts concerning the implications of the assumptions (C.1)–(C.5). Denote by $B_0(\mathbb{C})$ the linear space of all bounded real-valued set functions, defined on $\mathbb{C}$ and uniformly continuous with respect to the pseudometric d_λ. Notice that $C \to \lambda(C)$ belongs to $B_0(\mathbb{C})$, since for any two measurable sets C_1 and C_2,

$$|\lambda(C_1) - \lambda(C_2)| \leq d_\lambda(C_1, C_2). \tag{2.2}$$

Denote by $B_1(\mathbb{C})$ the linear space composed of all real-valued set functions ϕ, defined on $\mathbb{C}$ and of the form

$$\phi(C) = \phi_1(C) + \sum_{i=1}^{k} a_i \mathbf{1}(\mathbf{x}_i \in C), \tag{2.3}$$

with $a_i \in \mathbb{R}$, $\mathbf{x}_i \in \mathbb{R}^d$, $1 \leq i \leq k$, $k \in \mathbb{N}$, and $\phi_1 \in B_0(\mathbb{C})$. Recalling (1.4), we see that $\Theta_n \in B_1(\mathbb{C})$ for each $n \geq 1$. Observe that $B_0(\mathbb{C}) \subseteq B_1(\mathbb{C})$. Endow these two linear spaces with the supremum norm defined by

$$\|\phi\|_C = \sup_{C \in \mathbb{C}} |\phi(C)|. \tag{2.4}$$

For each $C \in \mathbb{C}$, let π_C: $B_1(\mathbb{C}) \to \mathbb{R}$ be defined by $\pi_C(\phi) = \phi(C)$. Denote by $\mathbb{A} := \sigma(\{\pi_C: C \in \mathbb{C}\})$ the σ-algebra generated in $B_1(\mathbb{C})$ by $\{\pi_C: C \in \mathbb{C}\}$, and by $\mathbb{B}_U$

the σ-algebra generated by the open balls with respect to $\|\cdot\|_{\mathbb{C}}$. The following useful result holds.

FACT 1. *Under* (C.2), *we have*

$$\mathbb{B}_U \subseteq \mathbb{A}. \tag{2.5}$$

Moreover, if $\mathbb{G}$ *is a countable generating subclass of* $\mathbb{C}$, *then, for any* $\phi \in B(\mathbb{C})$ *and* $\varepsilon > 0$,

$$\{f \in B(\mathbb{C}): \|\phi - f\|_{\mathbb{C}} \leq \varepsilon\} = \bigcap_{D \in \mathbb{G}} \{f \in B(\mathbb{C}): |\phi(D) - f(D)| \leq \varepsilon\}. \tag{2.6}$$

In addition, for each $n \geq 1$, *the set-function which maps* $C \in \mathbb{C}$ *to* $\lambda_n(\mathbf{a} + C)$ [*resp.* $\Theta_n(C)$] *is* $(\mathbb{F}, \mathbb{A})$-*measurable, and each sample path of* $\{\lambda_n(\mathbf{a} + C),\ C \in \mathbb{C}\}$ (*resp.* $\{\Theta_n(C),\ C \in \mathbb{C}\}$) *is uniquely determined by its values on* $\mathbb{G}$.

PROOF. See, for example, Lemma 20, page 108 of Gaenssler (1982). □

For any $A \in \mathbb{B}$, consider for $n \geq 1$,

$$\alpha_n(A) = n^{1/2}\big(\lambda_n(A) - \lambda(A)\big) = n^{-1/2} \sum_{i=1}^{n} \{\mathbf{1}(U_i \in A) - \lambda(A)\},$$

and, for $\varepsilon > 0$ and $n \geq 1$, set

$$\theta_n(\varepsilon) = \sup\{|\alpha_n(A) - \alpha_n(B)|: A, B \in \mathbb{C} - \mathbf{a},\ d_\lambda(A, B) < \varepsilon\}. \tag{2.7}$$

In view of (1.3), set for $h > 0,\ \varepsilon > 0,\ C \in \mathbb{B}$, and $n \geq 1$,

$$\Theta_n(C, h) = \frac{1}{\sqrt{nh}} \sum_{i=1}^{n} \Big\{\mathbf{1}\big(U_i \in \mathbf{t} + h^{1/d}C\big) - h\lambda(C)\Big\} = \frac{1}{\sqrt{h}}\alpha_n\big(\mathbf{t} + h^{1/d}C\big) \tag{2.8}$$

and

$$\omega_n(\varepsilon, h) = \sup\{|\Theta_n(C, h) - \Theta_n(D, h)|: C, D \in \mathbb{C},\ d_\lambda(C, D) < \varepsilon\}. \tag{2.9}$$

FACT 2. *Assumptions* (C.2) *and* (C.3) *jointly imply that, for all* $\varepsilon > 0$ *and* $\gamma > 0$,

$$\lim_{\varepsilon \downarrow 0} \left\{ \limsup_{n \to \infty} P\big(\theta_n(\varepsilon) > \gamma\big) \right\} = 0 \tag{2.10}$$

and

$$\mathbb{C} \textit{ is totally bounded with respect to } d_\lambda. \tag{2.11}$$

PROOF. See, for example, Dudley (1978), and Theorem B, page 113 of Gaenssler (1982). Note here that (C.2) implies, by Fact 1, that $\theta_n(\varepsilon)$ is $\mathbb{F}$-measurable. □

FACT 3. *We have the following inequalities, for any bounded* $C, C', D, D' \in \mathbb{B}$:

$$d_\lambda(C \cap D, C' \cap D') \le d_\lambda(C, C') + d_\lambda(D, D'), \tag{2.12}$$

$$d_\lambda(C - D, C' - D') \le d_\lambda(C, C') + d_\lambda(D, D'). \tag{2.13}$$

PROOF. We have

$$(C \cap D)\Delta(C' \cap D') \subseteq (C\Delta C') \cup (D\Delta D')$$

and

$$(C - D)\Delta(C' - D') \subseteq (C\Delta C') \cup (D\Delta D'),$$

which obviously imply (2.12) and (2.13) . □

2.2. *The oscillation modulus of the local empirical process.* In this section we will obtain a uniform upper bound for the oscillation of $\Theta_n(C)$ with respect to the pseudometric d_λ. This result, stated in Lemma 2.4 in the sequel, will turn out to be an essential tool in the proofs of our main theorems. We will frequently use the following fact.

FACT 4. *Let* $U_1, U_2, \dots$ *be i.i.d. uniform* $[0,1]^d$ *random variables. Then, for* $1 \le m < n$, *the conditional distribution of*

$$\sum_{i=1}^{m} \mathbf{1}\big(U_i \in \mathbf{t} + h^{1/d}C\big), \qquad C \in \mathbb{C},$$

given that $U_1, \dots, U_m \in \mathbf{t}+h^{1/d}[a,b]^d$ *and* $U_{m+1}, \dots, U_n \notin \mathbf{t}+h^{1/d}[a,b]^d$, *is equal to the distribution of*

$$\sum_{i=1}^{m} \mathbf{1}(U_i \in C - \mathbf{a}), \qquad C \in \mathbb{C}.$$

PROOF. Straightforward. □

LEMMA 2.1. *Whenever the sequence* $\{k(n), n \ge 1\}$ *is such that* $0 < k(n) \le n$, $k(n) \to \infty$ *and* $n^{-1}k(n) \to 0$ *as* $n \to \infty$ *and* $\mathbb{C}$ *satisfies* (2.10), *then, for all* $\eta > 0$,

$$\lim_{\varepsilon \downarrow 0} \left\{ \limsup_{n \to \infty} P\left(\omega_n\left(\varepsilon, \frac{k(n)}{n}\right) > \eta \right) \right\} = 0. \tag{2.14}$$

PROOF. For any $\eta > 0$ and $\varepsilon > 0$, set $\delta(\varepsilon, \eta) = \limsup_{n \to \infty} P(\theta_n(\varepsilon) > \eta/4)$. For any $0 < \varepsilon < 1/2$, let $z(\varepsilon)$ be defined by $P(|Z| > z(\varepsilon)) = \varepsilon$, where Z is a standard normal random variable. We note, by (2.10), that, for each $\eta > 0$,

$$\lim_{\varepsilon \downarrow 0} \delta(\varepsilon, \eta) = 0. \tag{2.15}$$

Since $P(|Z| > z) = (1 + o(1))\sqrt{(2/\pi)}z^{-1}\exp(-\frac{1}{2}z^2)$ as $z \to \infty$, an easy calculation shows that

$$\lim_{\varepsilon \downarrow 0} \varepsilon z(\varepsilon) = 0. \tag{2.16}$$

We will establish that, for all $0 < \varepsilon < 1/2$ such that $\eta - \varepsilon z(\varepsilon) > \eta/2$,

$$\limsup_{n \to \infty} P\Big(\omega_n\big(\varepsilon, n^{-1}k(n)\big) > \eta\Big) \le \delta(\varepsilon, \eta) + \varepsilon, \tag{2.17}$$

which, when combined with (2.15) and (2.16), implies (2.14).

Let

$$N_n = \#\left\{U_1, \dots, U_n \in \mathbf{t} + \left(\frac{k(n)}{n}\right)^{1/d} [a, b]^d\right\}.$$

Now N_n follows a binomial $\mathrm{Bin}(n, k(n)/n)$ distribution which satisfies, under (K.1) [observe that this assumption entails $n(k(n)/n)(1 - k(n)/n) = \{1 + o(1)\}k(n) \to \infty$ as $n \to \infty$],

$$\frac{N_n - k(n)}{\sqrt{k(n)}} \to_d N(0, 1) \quad \text{as } n \to \infty, \tag{2.18}$$

where "$\to_d$" denotes convergence in distribution.

We have the following easy bound:

$$\begin{aligned} &P\Big(\omega_n\big(\varepsilon, n^{-1}k(n)\big) > \eta\Big) \\ &\le \sum_{m:\ |m - k(n)| \le z(\varepsilon)\sqrt{k(n)}} P\Big(\omega_n\big(\varepsilon, n^{-1}k(n)\big) > \eta,\ N_n = m\Big) \\ &\quad + P\big(|N_n - k(n)| > z(\varepsilon)\sqrt{k(n)}\big). \end{aligned} \tag{2.19}$$

We see that

$$\begin{aligned} &P\Big(\omega_n\big(\varepsilon, n^{-1}k(n)\big) > \eta,\ N_n = m\Big) \\ &\quad = \binom{n}{m}\left(1 - \frac{k(n)}{n}\right)^{n-m}\left(\frac{k(n)}{n}\right)^m \\ &\quad\quad \times P\Bigg(\omega_n\big(\varepsilon, n^{-1}k(n)\big) > \eta \,\Big|\, U_i \in \mathbf{t} + \left(\frac{k(n)}{n}\right)^{1/d}[a, b]^d,\ i = 1, \dots, m, \\ &\quad\quad\quad U_i \notin \mathbf{t} + \left(\frac{k(n)}{n}\right)^{1/d}[a, b]^d,\ i = m + 1, \dots, n\Bigg). \end{aligned} \tag{2.20}$$

This, in turn, by the distributional identity given in Fact 4, and making use of the fact that $\lambda(C \triangle D) < \varepsilon$ and $|m - k(n)| \le z(\varepsilon)\sqrt{k(n)}$ together imply $|\lambda(C) - \lambda(D)| \times |m - k(n)| \le \varepsilon z(\varepsilon)\sqrt{k(n)}$, is less than or equal to

$$\binom{n}{m}\left(1 - \frac{k(n)}{n}\right)^{n-m}\left(\frac{k(n)}{n}\right)^m P\left(\theta_m(\varepsilon) > \sqrt{\frac{k(n)}{m}}\big(\eta - \varepsilon z(\varepsilon)\big)\right).$$

Observe that, for all large n, we have $\sqrt{k(n)/m} > 1/2$ uniformly over $|m-k(n)| \le z(\varepsilon)\sqrt{k(n)}$. Hence, by our choice of ε, our bound is, uniformly over $|m-k(n)| \le z(\varepsilon)\sqrt{k(n)}$, less than or equal to

$$P(N_n = m)P\big(\theta_m(\varepsilon) > \eta/4\big),$$

which entails (2.17). The proof of Lemma 2.1 is now complete. □

Let $s_1, s_2, \dots$ be a sequence of i.i.d. random signs, independent of $U_1, U_2, \dots$ and such that $P(s_n = 1) = P(s_n = -1) = 1/2$. Let, for any $C \in \mathbb{C}$,

$$S_n(C) = \sum_{i=1}^{n} s_i \mathbf{1}(U_i \in \mathbf{t} + C). \tag{2.21}$$

Now, (C.3) implies that for all $\eta > 0$ [see Theorem 2.14 of Giné and Zinn (1984)]

$$\lim_{\varepsilon \downarrow 0} \left\{ \limsup_{n\to\infty} P\left(\sup\left\{ \frac{|S_n(C) - S_n(D)|}{\sqrt{n}} \colon d_\lambda(C,D) < \varepsilon \right\} > \eta \right) \right\} = 0. \tag{2.22}$$

For any $\varepsilon > 0$ and $C \in \mathbb{C}$, let

$$\Delta_n(C,\varepsilon) = \sup\big\{ |S_n(C) - S_n(D)| \colon D \in \mathbb{C},\ d_\lambda(C,D) < \varepsilon \big\}.$$

Assertion (2.22) along with the symmetric distribution of $S_n(\cdot)$ allows us to follow exactly the steps used to prove Lemma 3.5 of Kuelbs and Dudley (1980) to produce the following inequality. For every $\varepsilon > 0$ there is an integer $M = M(\varepsilon) > 0$ such that, for every $k = 1, 2, \dots, \lambda$, and $\lambda > 0$, $b > 0$,

$$P\big(\Delta_{kM}(C,\varepsilon) \ge 2\lambda b\big) \le \exp\big(-\lambda^2 + \lambda^2 \phi(M,k,\lambda,b,\varepsilon)\big), \tag{2.23}$$

where

$$\phi(M,k,\lambda,b,\varepsilon) := \frac{1}{2}\left\{ \frac{1}{\lambda b}(17\varepsilon^2 kM)^{1/2} + \exp\left(\frac{2\lambda M}{b}\right)\frac{17\varepsilon^2 kM}{b^2} \right\}. \tag{2.24}$$

This inequality, in turn, yields, for all $n \ge M$, $\lambda > 0$ and $b > 0$,

$$P\big(\Delta_n(C,\varepsilon) \ge 2\lambda b + M\big) \le \exp\left(-\lambda^2 + \lambda^2 \phi\left(M, \frac{n}{M}, \lambda, b, \varepsilon\right)\right), \tag{2.25}$$

where we make use of the fact that $\Delta_n(C,\varepsilon) \le \Delta_{kM}(C,\varepsilon) + M$, with $k = \lfloor n/M \rfloor$ and $\lfloor u \rfloor$ denoting the integer part of u.

For any $C \in \mathbb{C}$ and $h > 0$, set

$$T_n(C,h) = \sum_{i=1}^{n} \left\{ \mathbf{1}\big(U_i \in \mathbf{t} + h^{1/d}C\big) - \mathbf{1}\big(U_i' \in \mathbf{t} + h^{1/d}C\big) \right\}, \tag{2.26}$$

where $U_1', U_2', \dots$ is a sequence of i.i.d. uniform $[0,1]^d$ random variables, independent of $U_1, U_2, \dots$. Furthermore, set

$$D_n(C,h,\varepsilon) = \sup\big\{ |T_n(C,h) - T_n(D,h)| \colon D \in \mathbb{C},\ d_\lambda(C,D) < \varepsilon \big\}. \tag{2.27}$$

LEMMA 2.2. *Assume that* (C.1)–(C.4), (K.1) *and* (K.2) *hold, that is, that*

(2.28)(i) $k(n)\uparrow$;

(2.28)(ii) $n^{-1}k(n)\downarrow$;

(2.28)(iii) $(\log_2 n)^{-1}k(n)\to\infty$ *as* $n\uparrow\infty$.

Then there exists a constant $K\le 480$ *such that, for all* $0<\varepsilon<1$ *and* $C\in\mathbb{C}$,

$$\limsup_{n\to\infty}\frac{D_n(C,k(n)/n,\varepsilon/2)}{\sqrt{k(n)\log_2 n}}\le K\varepsilon\quad a.s. \tag{2.29}$$

PROOF. Set $n_r=\lfloor\rho^r\rfloor$, $r=0,1,\dots$, for some $1<\rho<2$. By (2.28)(i), (2.28)(ii), (C.4) and (1.3), in combination with the inequality $\lambda((hC)\triangle D)\le\lambda((hC)\triangle C)+\lambda(C\triangle D)$, we obtain that, for some $1<\rho<2$ depending on $\varepsilon>0$,

$$\max_{n_r<n\le n_{r+1}}\frac{D_n(C,k(n)/n,\varepsilon/2)}{\sqrt{k(n)\log_2 n}}\le\max_{n_r<n\le n_{r+1}}\frac{2D_n(C,k(n_r)/n_r,\varepsilon)}{\sqrt{k(n_r)\log_2 n_r}}.$$

Next, by a general version of Lévy's inequality [see Lemma 3.1 of Kuelbs and Dudley (1980)], for any $z>0$ and $\varepsilon>0$,

$$\begin{aligned}&P\left(\max_{n_r<n\le n_{r+1}}\frac{D_n(C,k(n_r)/n_r,\varepsilon)}{\sqrt{k(n_r)\log_2 n_r}}>8z\varepsilon\right)\\&\quad\le 2P\left(\frac{D_{n_{r+1}}(C,k(n_r)/n_r,\varepsilon)}{\sqrt{k(n_r)\log_2 n_r}}>8z\varepsilon\right)=:2P(r).\end{aligned}\tag{2.30}$$

Set, for $r=1,2,\dots$,

$$\overline{N}_r=\#\left\{U_1,\dots,U_{n_{r+1}}\in\mathbf{t}+\left(\frac{k(n_r)}{n_r}\right)^{1/d}[a,b]^d\right\}$$

and

$$\overline{N}'_r=\#\left\{U'_1,\dots,U'_{n_{r+1}}\in\mathbf{t}+\left(\frac{k(n_r)}{n_r}\right)^{1/d}[a,b]^d\right\}.$$

Notice that $\overline{N}_r$ and $\overline{N}'_r$ are independent binomial Bin $(n_{r+1},k(n_r)/n_r)$ random variables. Set, for any $h>0$ and $C\in\mathbb{C}$,

$$\overline{T}_n(C,h)=\sum_{i=1}^n s_i\left\{\mathbf{1}\left(U_i\in t+h^{1/d}C\right)-\mathbf{1}\left(U'_i\in t+h^{1/d}C\right)\right\},$$

where $s_1,s_2,\dots$ is an i.i.d. sequence of random signs as above, independent of $U_1,U_2,\dots$ and $U'_1,U'_2,\dots$. Since

$$\left\{\overline{T}_n(C,h),C\in\mathbb{C}\right\}=_d T_n(C,h),C\in\mathbb{C}\right\},$$

with "$=_d$" denoting equality in distribution, we get immediately that

$$\{\overline{D}_n(C,h,\varepsilon),\ C\in\mathbb{C}\} =_d \{D_n(C,h,\varepsilon),\ C\in\mathbb{C}\},$$

where

$$\overline{D}_n(C,h,\varepsilon) := \sup\Big\{\big|\overline{T}_n(C,h)-\overline{T}_n(D,h)\big|:\ D\in\mathbb{C},\ d_\lambda(C,D)<\varepsilon\Big\}, \tag{2.31}$$

Now set, for $\{i_1,\dots,i_m,j_1,\dots,j_{m'}\}\subseteq\{1,\dots,n_{r+1}\}$,

$$\begin{aligned} P_r'(\varepsilon) := P\Bigg(\overline{D}_{n_{r+1}}\Big(C,\frac{k(n_r)}{n_r},\varepsilon\Big) \\ > 8z\varepsilon\sqrt{k(n_r)\log_2 n_r}\,\Big|\,A(i_1,\dots,i_m,j_1,\dots,j_{m'})\Bigg),\end{aligned} \tag{2.32}$$

where $A(i_1,\dots,i_m,j_1,\dots,j_{m'})$ is the event that

$$\begin{aligned} &U_i\in\mathbf{t}+\Big(\frac{k(n_r)}{n_r}\Big)^{1/d}[a,b]^d \quad \text{for } i\in\{i_1,\dots,i_m\},\\ &U_i\notin\mathbf{t}+\Big(\frac{k(n_r)}{n_r}\Big)^{1/d}[a,b]^d \quad \text{for } i\in\{1,\dots,n_{r+1}\}-\{i_1,\dots,i_m\},\\ &U_j'\in\mathbf{t}+\Big(\frac{k(n_r)}{n_r}\Big)^{1/d}[a,b]^d \quad \text{for } j\in\{j_1,\dots,j_{m'}\},\\ &U_j'\notin\mathbf{t}+\Big(\frac{k(n_r)}{n_r}\Big)^{1/d}[a,b]^d \quad \text{for } j\in\{1,\dots,n_{r+1}\}-\{j_1,\dots,j_m'\},\end{aligned}$$

from which it follows that

$$\begin{aligned} P_r'(\varepsilon) &= P\Bigg(\sup\Bigg\{\Bigg|\sum_{k=1}^m s_{i_k}\big\{\mathbf{1}(U_{i_k}\in C)-\mathbf{1}(U_{i_k}\in D)\big\}\\ &\qquad -\sum_{p=1}^{m'} s_{j_p}\big\{\mathbf{1}(U_{j_p}\in C)-\mathbf{1}(U_{j_p}\in D)\big\}\Bigg|\\ &\qquad > 8z\varepsilon\sqrt{k(n_r)\log_2 n_r}:\ D\in\mathbb{C},\ d_\lambda(C,D)<\varepsilon\Bigg\}\Bigg)\\ &\le P\big(\Delta_m(C,\varepsilon)>4z\varepsilon\sqrt{k(n_r)\log_2 n_r}\big)\\ &\qquad + P\big(\Delta_{m'}(C,\varepsilon)>4z\varepsilon\sqrt{k(n_r)\log_2 n_r}\big) =: P(r,m)+P(r,m').\end{aligned} \tag{2.33}$$

By (2.30), (2.31), (2.32) and (2.33), it follows that

$$\begin{aligned} P(r) \le& \sum_{k(n_r)\le m,\,m'\le 3k(n_r)} \{P(r,m)+P(r,m')\}P(\overline{N}_r=m)P(\overline{N}_r'=m')\\ &+2P\big(\overline{N}_r>3k(n_r)\text{ or }\overline{N}_r<k(n_r)\big) =: P_1(r)+P_2(r).\end{aligned} \tag{2.34}$$

Consider first $P_2(r)$ and recall that $\overline{N}_r$ is $\mathrm{Bin}(n_{r+1}, k(n_r)/n_r)$. By Inequality 1, page 440 in Shorack and Wellner (1986), whenever Y is $\mathrm{Bin}(n,p)$, we have

$$P(|Y-np| \geq \gamma\sqrt{n}) \leq 2\exp\left(-\frac{\gamma^2}{2p}\psi\left(\frac{\gamma}{p\sqrt{n}}\right)\right), \tag{2.35}$$

for all $\gamma > 0$ and $0 \leq p \leq 1/2$, with $\psi(u) = 2u^{-2}\{(1+u)\log(1+u) - u\}$.

Set in (2.35) $n = n_{r+1}$, $p = k(n_r)/n_r$ and, recalling that $\rho = n_{r+1}/n_r \in (1,2)$,

$$\gamma = n_{r+1}^{-1/2} k(n_r) \min\left\{(3 - (n_{r+1}/n_r)), ((n_{r+1}/n_r) - 1)\right\} = n_{r+1}^{-1/2} k(n_r)(\rho - 1).$$

We see that

$$\frac{\gamma^2}{2p} = \frac{(\rho-1)^2}{2\rho} k(n_r) \quad \text{and} \quad \frac{\gamma}{p\sqrt{n_{r+1}}} = \frac{\rho - 1}{\rho}.$$

Hence we have

$$P_2(r) \leq 4\exp\left(-\frac{(\rho-1)^2}{2\rho} k(n_r)\psi\left(\frac{\rho-1}{\rho}\right)\right) =: 4\exp(-c_1 k(n_r)). \tag{2.36}$$

By (2.28)(iii) we see that for all r sufficiently large, $k(n_r) \geq (3/c_1)\log_2 n_r \geq 2\log r$, which, in turn, implies by (2.36) that $P_2(r) \leq 4r^{-2}$. Hence we have

$$\sum_r P_2(r) < \infty. \tag{2.37}$$

Next, let $M = M(\varepsilon)$ be as in (2.25). Since $\sqrt{k(n_r)\log_2 n_r} \to \infty$, we have, for all r sufficiently large and $k(n_r) \leq m \leq 3k(n_r)$,

$$\begin{aligned} P(r,m) &= P\big(\Delta_m(C,\varepsilon) > 4z\varepsilon\sqrt{k(n_r)\log_2 n_r}\big) \\ &\leq P\big(\Delta_m(C,\varepsilon) > 2z\varepsilon\sqrt{k(n_r)\log_2 n_r} + M\big). \end{aligned} \tag{2.38}$$

By choosing $\lambda = \lambda_r = \sqrt{z\log_2 n_r}$, $k = m/M$ and $b = b_r = \varepsilon\sqrt{zk(n_r)}$ in (2.24), we see that, for all r sufficiently large and $k(n) \leq m \leq 3k(n_r)$,

$$\phi\left(M, \frac{m}{M}, \lambda_r, b_r, \varepsilon\right) \leq \frac{1}{2}\left\{\frac{\sqrt{51}}{z\sqrt{\log_2 n_r}} + \exp\left(\frac{2M}{\varepsilon}\sqrt{\frac{\log_2 h_r}{k(n_r)}}\right)\frac{51}{z}\right\}, \tag{2.39}$$

which by (2.28)(iii) is less than $26/z$ for all large r. Now choose $z = 30$. By combining (2.25), (2.38) and (2.39), we obtain that, for all r sufficiently large and $k(n_r) \leq m \leq 3k(n_r)$,

$$\begin{aligned} P(r,m) &\leq \exp\left(\lambda_r^2\left(-1 + \phi\left(M, \frac{m}{M}, \lambda_r, b_r, \varepsilon\right)\right)\right) \\ &\leq \exp\big((26 - z)\log_2 n_r\big) = \exp(-4\log_2 n_r) \leq r^{-3}. \end{aligned} \tag{2.40}$$

By (2.34) and (2.40), $P_1(r) \leq 2r^{-2}$ for all large r, from which we get

$$\sum_r P_1(r) < \infty.$$

Therefore, on account of (2.30), (2.34) and (2.37), we conclude from the Borel–Cantelli lemma that (2.29) holds with $K = 16z = 480$. This completes the proof of the lemma. □

We will make use of the symmetrization procedure given in Lemma 2.1 of Ledoux and Talagrand (1990) [see also Lemma 3.16 of Kuelbs and Dudley (1976)], which we state in a more general form as follows.

LEMMA 2.3. *Let* $\mathbf{X}$ *be a Hausdorff topological vector space and let* $\mathbf{B}$ *denote the corresponding* σ*-algebra of Borel sets. Let* $\|\cdot\|$ *be a* $\mathbf{B}$*-measurable semi-norm on* $\mathbf{X}$. *Let* $\{Z_n, n \geq 1\}$ *and* $\{Z'_n, n \geq 1\}$ *be identically distributed independent sequences of* $\mathbf{X}$*-valued random variables such that the sequence* $\{Z_n - Z'_n,\ n \geq 1\}$ *is almost surely bounded* (*resp. convergent to* 0) *and* $\{Z_n, n \geq 1\}$ *is bounded* (*resp. convergent to* 0) *in probability. Then* $\{Z_n, n \geq 1\}$ *is almost surely bounded* (*resp. convergent to* 0). *Moreover, if, for some numbers* M *and* A,

$$\limsup_{n \to \infty} \|Z_n - Z'_n\| \leq M \quad a.s. \tag{2.41}$$

and

$$\limsup_{n \to \infty} P\big(\|Z_n\| > A\big) < 1, \tag{2.42}$$

then

$$\limsup_{n \to \infty} \|Z_n\| \leq 2M + A \quad a.s. \tag{2.43}$$

PROOF. Assume that (2.43) does not hold. Then there exists almost surely a sequence $1 \leq n(1) < n(2) < \cdots$, depending on $\{Z_n\}$ only, and an $\varepsilon > 0$ such that

$$\liminf_{k \to \infty} \|Z_{n(k)}\| > 2M + A + \varepsilon.$$

On the other hand, (2.42), in combination with the fact that $\{Z'_n\}$ is independent of $\{n(k)\}$, implies

$$\liminf_{k \to \infty} P\big(\|Z'_{n(k)}\| \leq A\big) > 0,$$

so that we may extract from $\{n(k), k \geq 1\}$ a subsequence $\{m(k), k \geq 1\}$ such that

$$\limsup_{k \to \infty} \|Z'_{m(k)}\| \leq A \quad \text{a.s.}$$

Thus we have

$$\liminf_{k\to\infty} \|Z_{m(k)} - Z'_{m(k)}\| \geq M + \varepsilon \quad \text{a.s.},$$

a contradiction. The proofs of the other statements are similar and will be omitted. □

Let $\Theta_n(C,h)$ be as in (2.8) and set, for $C \in \mathbb{C}$, $\varepsilon > 0, h > 0$ and $n \geq 1$,

$$W_n(C,h,\varepsilon) = \sup\{|\Theta_n(C,h) - \Theta_n(D,h)|: D \in \mathbb{C},\ d_\lambda(C,D) < \varepsilon\}. \tag{2.44}$$

LEMMA 2.4. *Assume that* (C.1)–(C.4), (K.1) *and* (K.2) *hold, that is*,

(i) $k(n) \uparrow$;
(ii) $n^{-1}k(n) \downarrow$;
(iii) $(\log_2 n)^{-1}k(n) \to \infty$ *as* $n \uparrow \infty$.

Then there exists an $\varepsilon_0 > 0$ *such that, for each* $0 < \varepsilon < \varepsilon_0$,

$$\limsup_{n\to\infty} \left\{ \sup_{C\in\mathbb{C}} \frac{W_n\left(C, \frac{k(n)}{n}, \frac{\varepsilon}{2}\right)}{\sqrt{\log_2 n}} \right\} \leq 2K\varepsilon \quad a.s. \tag{2.45}$$

PROOF. We will apply Lemma 2.3 to the sequences

$$Z_n = \frac{1}{\sqrt{k(n)\log_2 n}} \sum_{i=1}^n \left\{ \mathbf{1}\left(U_i \in \mathbf{t} + \left(\frac{k(n)}{n}\right)^{1/d} C \right) - k(n)\lambda_n(C) \right\}$$

and

$$Z'_n = \frac{1}{\sqrt{k(n)\log_2 n}} \sum_{i=1}^n \left\{ \mathbf{1}\left(U'_i \in \mathbf{t} + \left(\frac{k(n)}{n}\right)^{1/d} C \right) - k(n)\lambda_n(C) \right\}.$$

Observe that Z_n and Z'_n as defined above are $\mathbb{B}(\mathbb{C})$-valued random variables. For any $\varepsilon > 0$ and $\phi \in \mathbb{B}(\mathbb{C})$, set

$$\|\phi\|_\varepsilon = \sup\{|\phi(C) - \phi(D)|: C \in \mathbb{C}, D \in \mathbb{C}, d_\lambda(C,D) < \varepsilon\}. \tag{2.46}$$

In view of Fact 1, (C.1), (C.2) and (C.4) imply that $\|\cdot\|_\varepsilon$ is a measurable seminorm on $\mathbb{B}(C)$. We see from (2.26), (2.27) and (2.46) that

$$\limsup_{n\to\infty} \|Z_n - Z'_n\|_\varepsilon = \limsup_{n\to\infty} \left\{ \sup_{C\in\mathbb{C}} \frac{D_n(C, k(n)/n, \varepsilon)}{\sqrt{k(n)\log_2 n}} \right\}. \tag{2.47}$$

Making use of Fact 1, and then of Lemma 2.2, we see that if $\mathbb{G}$ is a countable generating subclass of $\mathbb{C}$, then, with probability 1,

$$\begin{aligned} \limsup_{n\to\infty} \left\{ \sup_{C\in\mathbb{C}} \frac{D_n(C, k(n)/n, \varepsilon)}{\sqrt{k(n)\log_2 n}} \right\} &= \limsup_{n\to\infty} \left\{ \sup_{C\in\mathbb{G}} \frac{D_n(C, k(n)/n, \varepsilon)}{\sqrt{k(n)\log_2 n}} \right\} \\ &= \sup_{C\in\mathbb{G}} \left\{ \limsup_{n\to\infty} \frac{D_n(C, k(n)/n, \varepsilon)}{\sqrt{k(n)\log_2 n}} \right\} \leq K\varepsilon. \end{aligned} \tag{2.48}$$

Thus, by combining (2.47) and (2.48), we obtain

$$\limsup_{n\to\infty} \|Z_n - Z_n'\|_\varepsilon \le K\varepsilon \quad \text{a.s.} \tag{2.49}$$

On the other hand, we see from (2.8) and (2.9) that

$$\|Z_n\|_\varepsilon = \frac{\omega_n(\varepsilon, k(n)/n)}{\sqrt{k(n)\log_2 n}}. \tag{2.50}$$

Let $A > 0$ and $\eta > 0$ be arbitrary. In view of (2.50), Lemma 2.1 implies that, by choosing, independently of A, an $\varepsilon(\eta) > 0$ sufficiently small, we obtain that, for all $0 < \varepsilon \le \varepsilon(\lambda)$,

$$\begin{aligned}\limsup_{n\to\infty} P(\|Z_n\|_\varepsilon > A) &= \limsup_{n\to\infty} P\left(\omega_n\left(\varepsilon, \frac{k(n)}{n}\right) > A\sqrt{k(n)\log_2 n}\right)\\ &\le \limsup_{n\to\infty} P\left(\omega_n\left(\varepsilon, \frac{k(n)}{n}\right) > \eta\right) < 1.\end{aligned} \tag{2.51}$$

In view of (2.49), (2.51) and Lemma 2.4, it follows that, for each $0<\varepsilon \le \varepsilon_0 := \varepsilon(\eta)$,

$$\limsup_{n\to\infty}\left\{\sup_{C\in\mathbb{C}} \frac{W_n(C, k(n)/n, \varepsilon/2)}{\sqrt{\log_2 n}}\right\} = \limsup_{n\to\infty} \|Z_n\|_\varepsilon \le 2K\varepsilon + A \quad \text{a.s.} \tag{2.52}$$

Since $A > 0$ is arbitrary, we obtain readily (2.45) by applying (2.52) to a sequence $A_n \downarrow 0$. This completes the proof of Lemma 2.4. □

2.3. *Proof of the theorems.* We start by proving a special case of Theorem 1.1 obtained by considering the joint limiting behavior of $\{(2\log_2 n)^{-1/2}\Theta_n(C_i),\ 1 \le i \le M\}$, where $M \ge 1$ and $C_1, \dots, C_M \in \mathbb{C}$ are fixed, with $\lambda(C_i) > 0$ for $i = 1, \dots, M$. This result is given in Lemma 2.9 below. Before stating this lemma, we need the following preliminary results. We assume from now on that the assumptions of Theorem 1.1 hold.

Choose $D_1, \dots, D_N$ as in (C.5). Let

$$X_n(j) = \frac{\nu_j^{-1/2}\Theta_n(D_j)}{\sqrt{2\log_2 n}} = \left(\frac{n}{2\nu_j k(n)\log_2 n}\right)^{1/2} \alpha_n\left(\mathbf{t} + \left(\frac{k(n)}{n}\right)^{1/d} D_j\right) \tag{2.53}$$

and set $\mathbb{X}_n = (X_n(1), \dots, X_n(N)) \in \mathbb{R}^N$. We will describe the limiting behavior of $\mathbb{X}_n$ in Lemmas 2.7 and 2.8 in the sequel. For the proof of these two lemmas, we will need two technical results stated in Lemmas 2.5 and 2.6 below. Introduce the following notation.

Let $\{\Pi(t), t \ge 0\}$ denote a right-continuous Poisson process with $E(\Pi(t)) = t$. Let, for $j = 1, \dots, N$,

$$Y_n(j) = (2\nu_j k(n)\log_2 n)^{-1/2}\left\{\Pi\left(k(n)\sum_{i=1}^{j}\nu_i\right) - \Pi\left(k(n)\sum_{i=1}^{j-1}\nu_i\right) - k(n)\nu_j\right\} \tag{2.54}$$

and set $\mathbb{Y}_n = (Y_n(1), \dots, Y_n(N)) \in \mathbb{R}^N$.

LEMMA 2.5. *Whenever $n \geq 5$ and*

$$\frac{k(n)}{n}\sum_{i=1}^{N}\nu_i \leq \frac{1}{2},$$

we have, for any Borel subset A of $\mathbb{R}^N$,

$$P(\mathbb{X}_n \in A) \leq 2P(\mathbb{Y}_n \in A). \tag{2.55}$$

PROOF. Making use of the fact that the distribution of $\mathbb{Y}_n = (Y_n(1), \dots, Y_n(N))$, given that $\Pi(n) = n$, is equal to the distribution of $\mathbb{X}_n$, the proof is similar to that of Lemmas 2.1 and 3.1 in Deheuvels and Mason (1992a) and will be therefore omitted. □

LEMMA 2.6. *Let $N \geq 1$ be fixed and let $L = L(n)$ and $n \geq 1$ be positive integers. Let $(R_1, \dots, R_N)$ follow a multinomial distribution of the form*

$$P(R_1 = r_1, \dots, R_N = r_N) = \frac{L!}{r_1! \dots r_N!(L - \Sigma_{j=1}^N r_j)!} p_1^{r_1} \cdots p_N^{r_N}\left(1 - \sum_{j=1}^{N} p_j\right)^{L - \Sigma_{j=1}^N r_j},$$

for integers $r_1 \geq 0, \dots, r_N \geq 0$ and $\Sigma_{j=1}^N r_j \leq L$. Further, let

$$p_j = \frac{k(n)}{n}\nu_j \quad \text{and} \quad r_j = Lp_j + \delta_j\sqrt{2np_j \log_2 n} \quad \text{for } j = 1, \dots, N,$$

and assume that $n/L(n)$ and $\delta_j = \delta_{j,n}$, for $j = 1, \dots, N$, are bounded away from 0 and ∞ as $n \to \infty$. Then, for any $\varepsilon > 0$, there exists a constant $C > 0$ such that, ultimately as $n \to \infty$, we have

$$P(R_1 = r_1, \dots, R_N = r_N) \geq Ck(n)^{-N/2}\exp\left(-(1-\varepsilon)\frac{n}{L}\left\{\sum_{j=1}^{N}\delta_j^2\right\}\log_2 n\right). \tag{2.56}$$

PROOF. It follows from Stirling's formula and straightforward expansions. We omit the details. □

LEMMA 2.7. *Assume that* (C.1)–(C.5), (K.1) *and* (K.2) *hold. Then the sequence $\{\mathbb{X}_n, n \geq 1\}$ is almost surely relatively compact in $\mathbb{R}^N$, with limit set included in*

$$\mathbb{L}_N := \left\{\mathbb{x} = (x_1, \dots, x_N): \sum_{j=1}^{N} x_j^2 \leq 1\right\}. \tag{2.57}$$

PROOF. The proof is decomposed into the following sequence of steps.

Step 1. Fix a $\gamma \in (0,1)$ and introduce the sequence $n(m) = \lfloor (1+\gamma)^m \rfloor$. Consider the sets defined, for $n(m-1) < n \leq n(m)$ and $m = 0, 1, \dots,$ by

$$D_j'(n) = \left(\frac{n}{k(n)} \frac{k(n(m))}{n(m)} \right)^{1/d} D_j =: \rho_n D_j. \tag{2.58}$$

Let accordingly, for $j = 1, \dots, N$ and $n \geq 1$,

$$X_n'(j) = \left(\frac{n}{2\nu_j k(n) \log_2 n} \right)^{1/2} \alpha_n \left(\mathbf{t} + \left(\frac{k(n)}{n} \right)^{1/d} D_j'(n) \right) \tag{2.59}$$

and $\mathbb{X}_n' = (X_n'(1), \dots, X_n'(N))$.

We will show that, for any $\varepsilon > 0$, there exists a $\gamma_0 > 0$ such that, for all $0 < \gamma \leq \gamma_0$,

$$\limsup_{n \to \infty} \left\{ \max_{1 \leq j \leq N} \left| X_n(j) - X_n'(j) \right| \right\} < \varepsilon \quad \text{a.s.} \tag{2.60}$$

First, observe that, for all m large enough and $n(m-1) < n \leq n(m)$,

$$1 - \gamma \leq \frac{n}{n(m)} \leq 1, \tag{2.61}$$

which by $k(n) \uparrow 0$ and $n^{-1}k(n) \downarrow 0$ gives, for all large enough m and $n(m-1) < n \leq n(m)$,

$$(1-\gamma) k\big(n(m)\big) \frac{\log_2 n(m)}{\log_2 n} \leq \frac{n}{n(m)} k\big(n(m)\big) \leq k(n) \leq k\big(n(m)\big). \tag{2.62}$$

Thus, by (2.58), (2.61) and (2.62), for all large m and $n(m-1) < n \leq n(m)$, we have $D_j'(n) = \rho_n D_j$, with

$$(1-\gamma)^{1/d} \leq \rho_n \leq 1. \tag{2.63}$$

By (1.3) applied to $\{D_j,\ 1 \leq j \leq N\}$, it follows that, for any $\varepsilon > 0$, we may choose $\gamma = \gamma(\varepsilon) > 0$ so small that for all m large enough

$$\max_{n(m-1) < n \leq n(m)} \left\{ \max_{1 \leq j \leq N} d_\lambda\big(D_j, D_j'(n)\big) \right\} < \frac{\varepsilon}{2} \min\left(1, \frac{1}{K}\right), \tag{2.64}$$

where K is as in (2.45). We now apply Lemma 2.4, which, when combined with (2.64), suffices for (2.60).

Step 2. Let, for $j = 1, \dots, N$, $m \geq 1$ and $n(m-1) < n \leq n(m)$,

$$\begin{aligned} X_n''(j) &= \left(\frac{n}{2\nu_j k\big(n(m)\big) \log_2 n(m)} \right)^{1/2} \alpha_n \left(\mathbf{t} + \left(\frac{k\big(n(m)\big)}{n(m)} \right)^{1/d} D_j \right) \\ &= \left(\frac{k(n) \log_2 n}{k\big(n(m)\big) \log_2 n(m)} \right)^{1/2} X_n'(j) \leq \left(\frac{k(n)}{k\big(n(m)\big)} \right)^{1/2} X_n'(j), \end{aligned} \tag{2.65}$$

and $\mathbb{X}_n'' = (X_n''(1), \dots, X_n''(N))$. Observe from (2.59) that

$$X_{n(m)}''(j) = X_{n(m)}'(j) = X_{n(m)}(j), \tag{2.66}$$

and that, for $n(m-1) < n \le n(m)$,

$$\begin{aligned} &X_{n(m)}''(j) - X_n''(j) \\ &\quad =_d \left(\frac{n(m)-n}{2\nu_j k(n(m)) \log_2 n(m)} \right)^{1/2} \alpha_{n(m)-n}\left(\mathbf{t} + \left(\frac{k(n(m))}{n(m)} \right)^{1/d} D_j \right). \end{aligned} \tag{2.67}$$

Let $\boldsymbol{\delta} = (\delta_1, \dots, \delta_N)$, $\boldsymbol{\eta} = (\eta_1, \dots, \eta_N)$ and introduce the events

$$C_m(\boldsymbol{\delta} + \boldsymbol{\eta}) = \left\{ \bigcap_{1 \le j \le N} |X_n''(j)| \ge \delta_j + \eta_j \text{ for some } n(m-1) < n \le n(m) \right\},$$

$$D_m(\boldsymbol{\delta}) = \left\{ \bigcap_{1 \le j \le N} \left|X_{n(m)}''(j)\right| \ge \delta_j \right\} = \left\{ \bigcap_{1 \le j \le N} \left|X_{n(m)}(j)\right| \ge \delta_j \right\} \text{ by (2.66).}$$

We will prove that if $\delta_j > 0$ and $\eta_j > 0$ or $\delta_j = \eta_j = 0$ for $j = 1, \dots, N$, then there exists an m_0 such that, for all $m \ge m_0$,

$$P\big(C_m(\boldsymbol{\delta} + \boldsymbol{\eta})\big) \le 2P\big(D_m(\boldsymbol{\delta})\big). \tag{2.68}$$

Toward a proof of (2.68), set

$$\begin{aligned} E_n(\boldsymbol{\delta} + \boldsymbol{\eta}) &= \left\{ \bigcap_{1 \le j \le N} |X_n''(j)| \ge \delta_j + \eta_j \right\}, \\ F_{m,n}(\boldsymbol{\eta}) &= \left\{ \bigcap_{1 \le j \le N} \left|X_{n(m)}''(j) - X_n''(j)\right| < \eta_j \right\}. \end{aligned} \tag{2.69}$$

In the remainder of our proof, we assume, without loss of generality, that

$$\min(\delta_1, \dots, \delta_N) \ge 0 \quad \text{and} \quad \eta := \min(\eta_1, \dots, \eta_N) > 0.$$

First, we notice that, for $n(m-1) < n \le n(m)$, the events $\{E_q(\boldsymbol{\delta} + \boldsymbol{\eta}),\ n(m-1) < q \le n\}$ are independent of $F_{m,n}(\boldsymbol{\eta})$. Denoting by $\overline{E}$ the complement of the event E. It follows that

$$P\big(C_m(\boldsymbol{\delta} + \boldsymbol{\eta})\big) = \sum_{r=n(m-1)+1}^{n(m)} P\left(E_r(\boldsymbol{\delta} + \boldsymbol{\eta}) \cap \bigcap_{q=n(m-1)+1}^{r-1} \overline{E}_q(\boldsymbol{\delta} + \boldsymbol{\eta}) \right),$$

and hence, that

$$\begin{aligned} &\inf_{n(m-1) < n \le n(m)} P\big(F_{m,n}(\boldsymbol{\eta})\big) P\big(C_m(\boldsymbol{\delta} + \boldsymbol{\eta})\big) \\ &\quad \le \sum_{r=n(m-1)+1}^{n(m)} P\left(E_r(\boldsymbol{\delta} + \boldsymbol{\eta}) \cap F_{m,r}(\boldsymbol{\eta}) \cap \bigcap_{q=n(m-1)+1}^{r-1} \overline{E}_q(\boldsymbol{\delta} + \boldsymbol{\eta}) \right) \\ &\quad \le P\big(D_m(\boldsymbol{\delta})\big). \end{aligned} \tag{2.70}$$

Next, we observe by (2.67) and the Chebyshev inequality that, uniformly over $n(m-1) < n \le n(m)$, we have, as $m \to \infty$,

$$P(\overline{F}_{m,n}(\eta)) \le \eta^{-2} \sum_{j=1}^{N} E\Big(\big|X''_{n(m)}(j) - X''_n(j)\big|^2\Big) \le \frac{N(n(m)-n)}{2\eta^2 n(m)\log_2 n(m)} \to 0.$$

Hence there exists an m_0 such that, for $m \ge m_0$, $\inf_{n(m-1)<n\le n(m)} P(F_{m,n}(\eta)) > \frac{1}{2}$, which, by (2.70), entails (2.68).

Step 3. Denote by $|\cdot|$ the Euclidean norm in $\mathbb{R}^N$. For each subset E of $\mathbb{R}^N$ and $\varepsilon > 0$, let $E^\varepsilon = \{\mathbb{x} \in \mathbb{R}^N : |\mathbb{x} - \mathbb{y}| < \varepsilon \text{ for some } \mathbb{y} \in E\}$.

Fix any $0 < \varepsilon < 1$. We first prove that there exists a constant $K_1(\varepsilon)$ such that, for all large n,

$$\begin{aligned} P_m(\varepsilon) &:= P\big(\mathbb{X}''_n \notin \mathbb{L}_N^{\varepsilon/3} \text{ for some } n(m-1) < n \le n(m)\big)\\ &\le P\big(|\mathbb{X}''_n|^2 \ge 1+\varepsilon \text{ for some } n(m-1) < n \le n(m)\big) \qquad (2.71)\\ &\le K_1(\varepsilon)\exp\bigg(-\Big(1+\frac{\varepsilon}{4}\Big)\log_2 n(m)\bigg). \end{aligned}$$

Let $Q = \lceil 16N/\varepsilon \rceil$. Observe that if

$$x_1^2 + \cdots + x_N^2 = 1 + \varepsilon < 2,$$

then, for $|\delta_1| \le 2/Q, \ldots, |\delta_N| \le 2/Q$,

$$(x_1+\delta_1)^2 + \cdots + (x_N+\delta_N)^2 \ge 1+\varepsilon - 4\sum_{i=1}^{N} |\delta_i| \ge 1 + \frac{\varepsilon}{2}. \qquad (2.72)$$

Assume that $|\mathbb{X}''_n|^2 \ge 1+\varepsilon$, and set, for $j = 1, \ldots, N$,

$$x_j = X''_n(j)\bigg(\frac{\sqrt{1+\varepsilon}}{|\mathbb{X}''_n|}\bigg), \qquad m_j = \lfloor Q|x_j| \rfloor \quad \text{and} \quad \delta_j = \begin{cases} \dfrac{m_j}{Q} - x_j, & \text{if } m_j = 0,\\[2mm] \dfrac{m_j - 1}{Q} - x_j, & \text{otherwise.} \end{cases}$$

Since $|\delta_j| \le 2/Q$ for $j = 1, \ldots, N$, it follows from (2.72) that $\mathbb{M} := (m_1, \ldots, m_N)$ varies in the set

$$S_N(\varepsilon) = \bigg\{\mathbb{M} \in \{0, \ldots, 2Q\}^N : \sum_{1 \le j \le N : m_j \ge 2} \Big(\frac{m_j - 1}{Q}\Big)^2 \ge 1 + \frac{\varepsilon}{2}\bigg\}. \qquad (2.73)$$

Therefore, letting E_n be as in (2.69), we have the inclusions of events

$$\begin{aligned} \{|\mathbb{X}''_n|^2 \ge 1+\varepsilon\} &\subseteq \bigcup_{\mathbb{M} \in S_N(\varepsilon)} E_n(\mathbb{M}/Q) \qquad (2.74)\\ &\subseteq \bigcup_{\mathbb{M} \in S_N(\varepsilon)} \bigg\{ \bigcap_{1 \le j \le N : m_j \ge 2} \Big\{ |X''_n(j)| \ge \frac{m_j - 1}{Q} \Big\} \bigg\}. \end{aligned}$$

Set $M_j = \max(0, m_j - 1)$ for $j = 1, \dots, N$. We obtain from (2.68) and (2.74) that

$$\begin{aligned} &P\big(|\mathbb{X}_n''|^2 \ge 1+\varepsilon \quad \text{for some } n(m-1) < n \le n(m)\big) \\ &\quad \le \sum_{\mathbb{M}\in S_N(\varepsilon)} P\left(C_m\left(\frac{m_1}{Q},\dots,\frac{m_N}{Q}\right)\right) \\ &\quad \le 2\sum_{\mathbb{M}\in S_N(\varepsilon)} P\left(D_m\left(\frac{M_1}{Q},\dots,\frac{M_N}{Q}\right)\right) \\ &\quad \le 2\sum_{\mathbb{M}\in S_N(\varepsilon)} P\left(\bigcap_{1\le j\le N: M_j\ge 1}\left\{|X''_{n(m)}(j)| \ge \frac{M_j}{Q}\right\}\right). \end{aligned} \tag{2.75}$$

By Lemma 2.5 in combination with the fact that the $Y_{n(m)}(j)$ are independent for $j = 1, \dots, N$, we have

$$\begin{aligned} &P\left(\bigcap_{1\le j\le N: M_j\ge 1}\left\{|X''_{n(m)}(j)| \ge \frac{M_j}{Q}\right\}\right) \\ &\qquad \le 2\prod_{1\le j\le N: M_j\ge 1} P\left(|Y_{n(m)}(j)| \ge \frac{M_j}{Q}\right). \end{aligned} \tag{2.76}$$

By Shorack and Wellner [(1986), (14), page 441, and inequality 1, page 485], we obtain from (2.54) the upper bound, for $M_j \ge 1$,

$$P\left(|Y_n(j)| \ge \frac{M_j}{Q}\right) \le 2\exp\left(-(\log_2 n)\left(\frac{M_j}{Q}\right)^2\left(1 - \frac{M_j}{Q}\sqrt{\frac{2\log_2 n}{\nu_j k(n)}}\right)\right). \tag{2.77}$$

By (K.1) and (K.2), the expression above is, for all large n, $1 \le j \le N$ and all $1 \le M_j \le Q$, less than

$$2\exp\left(-\left(\frac{1+\varepsilon/4}{1+\varepsilon/2}\right)(\log_2 n)\left(\frac{M_j}{Q}\right)^2\right).$$

Let $K_0(\varepsilon) := \#S_N(\varepsilon)$. Recalling (2.71), (2.73), (2.75), (2.76) and (2.77), we obtain, in turn, that, for all large n,

$$\begin{aligned} P_m(\varepsilon) &\le 2^{N+2}\sum_{m\in S_N(\varepsilon)} \exp\left(-\left(\frac{1+\varepsilon/4}{1+\varepsilon/2}\right)(\log_2 n(m))\sum_{j=1}^N\left(\frac{M_j}{Q}\right)^2\right) \\ &\le 2^{N+2}K_0(\varepsilon)\exp\left(-\left(1+\frac{\varepsilon}{4}\right)\log_2 n(m)\right). \end{aligned} \tag{2.78}$$

By choosing $K_1(\varepsilon) = 2^{N+2}K_0(\varepsilon)$, we obtain (2.71) by combining (2.75) and (2.78).

Step 4. We observe from (2.71) that for any $\varepsilon > 0$, $\sum_m P_m(\varepsilon) < \infty$, so that, by the Borel–Cantelli lemma,

$$P\big(|\mathbb{X}_n''|^2 \ge 1+\varepsilon \text{ i.o.}\big) = 0. \tag{2.79}$$

Observe from (2.65) that, whenever (2.62) holds, we have $(1-\gamma)|\mathbb{X}_n'|^2 \le |\mathbb{X}_n''|^2$. Thus, by (2.79), we obtain

$$P\left(|\mathbb{X}_n'|^2 \ge \frac{1+\varepsilon}{1-\gamma}\text{ i.o.}\right) = 0. \tag{2.80}$$

Since $\varepsilon > 0$ and $\gamma > 0$ may be chosen arbitrarily small in (2.60) and (2.80), we complete the proof of Lemma 2.7 by the triangle inequality. □

LEMMA 2.8. *Assume that* (C.1)–(C.5), (K.1) *and* (K.2) *hold. Then the sequence* $\{\mathbb{X}_n, n \ge 1\}$ *is almost surely relatively compact in* $\mathbb{R}^N$, *with limit set containing*

$$\mathbb{L}_N := \left\{\mathbf{x} = (x_1,\dots,x_N)\colon \sum_{j=1}^N x_j^2 \le 1\right\}. \tag{2.81}$$

PROOF. The proof is decomposed into the following sequence of steps.

Step 1. Let $\gamma > 0$ be the fixed and set $n(m) = \lfloor (1+\gamma)^m \rfloor$ for $m = 0, 1, \dots$. Consider, for $j = 1,\dots,N$,

$$Z_m'(j) = \left(\frac{n(m)}{2\nu_j k(n(m))\log_2 n(m)}\right)^{1/2} \alpha_{n(m-1)}\left(\mathbf{t} + \left(\frac{k(n(m))}{n(m)}\right)^{1/d} D_j\right) \tag{2.82}$$

and

$$\begin{aligned} Z_m''(j) &= \left(\frac{n(m)}{2\nu_j k(n(m))\log_2 n(m)}\right)^{1/2} \alpha_{n(m)}\left(\mathbf{t} + \left(\frac{k(n(m))}{n(m)}\right)^{1/d} D_j\right) \\ &\quad - \left(\frac{n(m-1)}{2\nu_j k(n(m))\log_2 n(m)}\right)^{1/2} \alpha_{n(m-1)}\left(\mathbf{t} + \left(\frac{k(n(m))}{n(m)}\right)^{1/d} D_j\right) \\ &= X_{n(m)}(j) - \left(\frac{n(m-1)}{n(m)}\right)^{1/2} Z_m'(j). \end{aligned} \tag{2.83}$$

Observe that the $Z_m''(j)$ are independent for $m = 1, 2, \dots$, with

$$\begin{aligned} Z_m''(j) =_d &\left(\frac{n(m)-n(m-1)}{2\nu_j k(n(m))\log_2 n(m)}\right)^{1/2} \\ &\times \alpha_{n(m)-n(m-1)}\left(\mathbf{t} + \left(\frac{k(n(m))}{n(m)}\right)^{1/d} D_j\right). \end{aligned} \tag{2.84}$$

By repeating the arguments of steps 3 and 4 of the proof of Lemma 2.6, we obtain readily that, for $j = 1,\dots,N$,

$$\limsup_{m\to\infty} |Z_m'(j)| \le 1 \quad \text{a.s.} \tag{2.85}$$

Let $\mathbb{Z}''_m = (Z''_m(1), \dots, Z''_m(N))$. Since $n(m-1)/n(m) \to 1/\gamma$ as $m \to \infty$, (2.83) and (2.85) imply that, for any $\varepsilon > 0$, there exists almost surely a $\gamma_1 = \gamma_1(\varepsilon) > 0$ such that, for any $\gamma \geq \gamma_1$, we have for all m sufficiently large

$$|\mathbb{X}_{n(m)} - \mathbb{Z}''_m| < \varepsilon. \tag{2.86}$$

Step 2. Let $Q \geq 1$, $(m_1, \dots, m_N) \in \{1, \dots, Q\}$ be fixed, $(e_1, \dots, e_N) \in \{-1, 1\}^N$ and consider

$$P''_m := P\left(e_1 Z''_m(1) \in \left(\frac{m_1 - 1}{Q}, \frac{m_1}{Q}\right], \dots, e_N Z''_m(N) \in \left(\frac{m_N - 1}{Q}, \frac{m_N}{Q}\right]\right). \tag{2.87}$$

In this step we will show that, whenever

$$\sum_{j=1}^{N} \left(\frac{m_j}{Q}\right)^2 < 1, \tag{2.88}$$

we may choose $\gamma > 0$ so large that

$$\sum_m P''_m = \infty. \tag{2.89}$$

Since $Q \geq 1$ may be chosen arbitrarily large in (2.87) and $\varepsilon > 0$ arbitrarily small in (2.86), the Borel–Cantelli lemma, in combination with (2.88), (2.89) and the observation that the $\mathbb{Z}''_m$ are independent, will then readily imply the conclusion of the lemma.

Toward proving (2.89), we will make use of Lemma 2.5. Note from (2.84) that, for $j = 1, \dots, N$, the values of

$$R_j := \left(2\nu_j k(n(m)) \log_2 n(m)\right)^{1/2} Z''_m(j) + \frac{n(m) - n(m-1)}{n(m)} k(n(m)) \nu_j$$

are integers. Therefore, the number of possible distinct values of $e_j Z''_m(j)$ in the interval $(m_j - 1/Q, m_j/Q]$ is for all large m greater than $k(n(m))^{1/2}$. In view of (2.84), (2.87) and (2.54), it follows from Lemma 2.6, taken with $n = n(m)$, $\delta_j = e_j(m_j/Q)$ and $L = n(m) - n(m-1)$, that, for any $\varepsilon > 0$ and all large m,

$$P''_m \geq C \exp\left(-(1-\varepsilon)\frac{n(m)}{n(m) - n(m-1)} \left\{\sum_{j=1}^{N} \left(\frac{m_j}{Q}\right)^2\right\} \log_2 n(m)\right). \tag{2.90}$$

Finally, we see from (2.90) that if we choose $\varepsilon > 0$ so small and $\gamma > 0$ so large that

$$\frac{1-\varepsilon}{1 - (1/\gamma)} \sum_{j=1}^{N} \left(\frac{m_j}{Q}\right)^2 < 1,$$

then (2.89) holds. The proof of Lemma 2.8 is now complete. □

REMARK 2.1. A simple argument based on step 2 of the proof of Lemma 2.7, in combination with the just-given proof of Lemma 2.8, shows that, for any random sequence $\{\nu(n),\ n \geq 1\}$ of positive integers such that almost surely $n^{-1}\nu(n) \to \rho \in (0,\infty)$ as $n \to \infty$, the sequence $\{\mathbb{X}_{\nu(n)},\ n \geq 1\}$ has almost surely a limit set containing $\mathbb{L}_n$. Since Lemma 2.7 also implies that this limit set is almost surely included in $\mathbb{L}_n$, we see that we may formally replace n by $\nu(n)$ in the statement of Lemmas 2.7 and 2.8.

LEMMA 2.9. *The sequence* $\{((2\log_2 n)^{-1/2}\Theta_n(C_1),\dots,(2\log_2 n)^{-1/2}\Theta_n(C_M)),\ n \geq 1\}$ *is almost surely relatively compact in* $\mathbb{R}^M$, *with limit set equal to*

$$\left\{\left(\int_{C_1}\phi(\mathbf{s})\,d\lambda(\mathbf{s}),\dots,\int_{C_M}\phi(\mathbf{s})\,d\lambda(\mathbf{s})\right):\ \int_{\mathbb{R}^d}\phi^2(\mathbf{s})\,d\lambda(\mathbf{s}) \leq 1\right\}. \tag{2.91}$$

PROOF. By combining Lemma 2.7 and 2.8, we obtain readily that the sequence $\{((2\log_2 n)^{-1/2}\Theta_n(D_1),\dots,(2\log_2 n)^{-1/2}\Theta_n(D_N)),\ n \geq 1\}$ is almost surely relatively compact in $\mathbb{R}^N$, with limit set equal to

$$\left\{(y_1,\dots,y_N):\ \sum_{i=1}^N \frac{y_i^2}{\lambda(D_i)} \leq 1\right\}. \tag{2.92}$$

Let

$$\psi(\mathbf{s}) = \sum_{i=1}^N \frac{y_i}{\lambda(D_i)}\mathbf{1}(\mathbf{s}\in D_i).$$

Observe that, whenever (2.92) holds, we have

$$y_i = \int_{D_i}\psi(\mathbf{s})\,d\lambda(\mathbf{s}) \quad \text{for } i = 1,\dots,N \quad \text{and} \quad \int_{\mathbb{R}^d}\psi^2(\mathbf{s})\,d\lambda(\mathbf{s}) \leq 1. \tag{2.93}$$

Assume, conversely, that ϕ is an arbitrary measurable function on $\mathbb{R}^d$, with

$$\int_{\mathbb{R}^d}\phi^2(\mathbf{s})\,d\lambda(\mathbf{s}) \leq 1, \tag{2.94}$$

and let $y_i = \int_{D_i}\phi(\mathbf{s})\,d\lambda(\mathbf{s})$ for $i = 1,\dots,N$. We obtain by the Schwarz inequality and disjointness of the D_i's that

$$\begin{aligned}\sum_{i=1}^N \frac{y_i^2}{\lambda(D_i)} &= \sum_{i=1}^N \frac{1}{\lambda(D_i)}\left\{\int_{D_i}\phi(\mathbf{s})\,d\lambda(\mathbf{s})\right\}^2\\ &\leq \sum_{i=1}^N \frac{1}{\lambda(D_i)}\left\{\int_{D_i}\phi^2(\mathbf{s})\,d\lambda(\mathbf{s})\right\}\left\{\int_{D_i}d\lambda(\mathbf{s})\right\}\\ &\leq \int_{\mathbb{R}^d}\phi^2(\mathbf{s})\,d\lambda(\mathbf{s}) \leq 1.\end{aligned}$$

By all this, and in view of (2.57), it follows that the set defined by (2.92) coincides with

$$\left\{\left(\int_{D_1}\phi(\mathbf{s})\,d\lambda(\mathbf{s}),\dots,\int_{D_N}\phi(\mathbf{s})\,d\lambda(\mathbf{s})\right):\int_{\mathbb{R}^d}\phi^2(\mathbf{s})\,d\lambda(\mathbf{s})\le 1\right\}. \tag{2.95}$$

Since each C_i, $i=1,\dots,M$, is, λ_0-a.e., the disjoint union of sets taken among $\{D_1,\dots,D_N\}$, (2.91) follows readily from (2.95). □

PROOF OF THEOREM 1.1. We have now in hand all the ingredients necessary for proving this theorem, following the lines of the final part of the proof of Theorem 2.1 of Kuelbs and Dudley (1980), pages 415 and 416. Fix an arbitrary $\varepsilon>0$. If we choose $\varepsilon_1=\frac{1}{2}\min(\varepsilon_0,\varepsilon/8K,\varepsilon^2/4)$ in Lemma 2.4, we obtain from this lemma that there exists almost surely an $n_1<\infty$ such that, for all $n\ge n_1$,

$$\sup_{C\in\mathbb{C}}\frac{W_n(C,k(n)/n,\varepsilon_1/2)}{\sqrt{2\log_2 n}}\le\frac{\varepsilon}{2}. \tag{2.96}$$

By (2.11) there exists a finite sequence $\{C_i,\ 1\le i\le M\}\subset\mathbb{C}$ with the following property. For any $C\in\mathbb{C}$, there exists an $i\in\{1,\dots,M\}$, with $d_\lambda(C,C_i)<\varepsilon_1/2$. By (2.40) and (2.96) this implies, in turn, that

$$|(2\log_2 n)^{-1/2}\Theta_n(C)-(2\log_2 n)^{-1/2}\Theta_n(C_i)|\le\frac{\varepsilon}{2}. \tag{2.97}$$

Let $1\le n_1'<\cdots<n_m'<\cdots$ be a sequence such that, along $\{n_m'\}$, the sequence $\{(2\log_2 n)^{-1/2}\Theta_n(C_i)\}$ converges to a limit for each $1\le i\le M$. Then, by Lemma 2.9, there exists a measurable function ϕ satisfying (2.94) and such that, ultimately along $\{n_m'\}$,

$$\max_{1\le i\le M}\left|(2\log_2 n)^{-1/2}\Theta_n(C_i)-\int_{C_i}\phi(\mathbf{s})\,d\lambda(\mathbf{s})\right|\le\frac{\varepsilon}{4}. \tag{2.98}$$

On the other hand, the Schwarz inequality entails

$$\begin{aligned}&\left|\int_C\phi(\mathbf{s})\,d\lambda(\mathbf{s})-\int_{C_i}\phi(\mathbf{s})\,d\lambda(\mathbf{s})\right|\\&\quad\le\left\{\int_{C\,\triangle\,C_i}\phi^2(\mathbf{s})\,d\lambda(\mathbf{s})\right\}^{1/2}d_\lambda(C,C_i)^{1/2}\le\frac{\varepsilon}{4},\end{aligned} \tag{2.99}$$

uniformly over $\phi\in\mathbb{S}(\mathbb{C})$ and $d_\lambda(C,C_i)\le\varepsilon^2/16$.

Thus, by combining (2.97), (2.98) and (2.99), we obtain that, ultimately along $\{n_m'\}$, we have

$$\sup_{C\in\mathbb{C}}\left|(2\log_2 n)^{-1/2}\Theta_n(C)-\int_C\phi(\mathbf{s})\,d\lambda(\mathbf{s})\right|\le\varepsilon. \tag{2.100}$$

We will now show that, almost surely,

$$\limsup_{n\to\infty}\left\{\inf_{f\in\mathbb{S}(\mathbb{C})}\left(\sup_{C\in\mathbb{C}}|(2\log_2 n)^{-1/2}\Theta_n(C)-f(C)|\right)\right\}\le\varepsilon. \tag{2.101}$$

To prove this statement, we observe that if (2.101) did not hold, then there would exist an $\varepsilon'>0$ and a sequence $1\le n_1<n_2<\cdots$ such that, along $\{n_m\}$,

$$\inf_{f\in\mathbb{S}(\mathbb{C})}\left(\sup_{C\in\mathbb{C}}|(2\log_2 n)^{-1/2}\Theta_n(C)-f(C)|\right)\ge\varepsilon+\varepsilon'.$$

On the other hand, Lemma 2.9 enables us to extract from $\{n_m\}$ a sequence $\{n'_m\}$ along which $\{(2\log_2 n)^{-1/2}\Theta_n(C_i)\}$ converges for each $1\le i\le N$. In view of (2.100), we obtain a contradiction. Since $\varepsilon>0$ may be chosen as small as desired in (2.101), it follows that, almost surely,

$$\lim_{n\to\infty}\left\{\inf_{f\in\mathbb{S}(\mathbb{C})}\|(2\log_2 n)^{-1/2}\Theta_n(C)-f(C)\|_{\mathbb{C}}\right\}=0. \tag{2.102}$$

We now observe that $\mathbb{S}(\mathbb{C})$ as defined in (1.7) is a compact subset of $B(\mathbb{C})$. To prove that this property holds, it suffices to show that $\mathbb{S}(\mathbb{C})$ is totally bounded and complete with respect to the uniform topology. In view of the equality of the sets given in (2.92) and (2.95), and by compactness of the unit ball in $\mathbb{R}^M$, the first part is an easy consequence of (2.11) in combination with the existence, for each $\varepsilon>0$, of $\{C_i,\ 1\le i\le M\}\in\mathbb{C}$, such that (2.99) holds. The second part follows readily from the observation that $\mathbb{S}(\mathbb{C})$ is a closed subset of $B(\mathbb{C})$.

By (2.102) it follows that the limit set of $\{(2\log_2 n)^{-1/2}\Theta_n(C),\ C\in\mathbb{C}\}$ is almost surely included in $\mathbb{S}(\mathbb{C})$. To show that we have equality, we choose any $\phi\in\mathbb{S}(\mathbb{C})$. By Lemma 2.9, there exists a sequence $1\le n'_1<n'_2<\cdots$, depending on $\varepsilon>0$, along which (2.98) and (2.100) hold ultimately. Therefore, there exists an $n=n(\varepsilon,\phi)$ such that (2.100) holds. By repeating this argument for a sequence $\varepsilon=\varepsilon_k\downarrow 0$, we obtain readily the existence of an increasing sequence $1\le n''_1<n''_2<\cdots$, along which

$$\|(2\log_2 n)^{-1/2}\Theta_n(C)-\phi\|_{\mathbb{C}}\to 0.$$

This, in turn, implies that ϕ belongs to the limit set of $\{(2\log_2 n)^{-1/2}\Theta_n,\ n\ge 1\}$, and completes the proof of Theorem 1.1. □

We now turn to the proof of Theorem 1.2, which will be obtained by combining the following three main ingredients: (i) the just-given proof of Theorem 1.1, (ii) the classical proof of the functional law of the iterated logarithm for empirical processes indexed by sets, due to Kuelbs and Dudley (1980) and (iii) a general argument describing the LIL behavior of independent sequences, when the corresponding behavior is known for each of these sequences. In order not to repeat the unnecessary details of these proofs, we will limit ourselves presently to a streamlined description of the main ideas which underline the result. We start with three propositions.

PROPOSITION 2.1. *Let* $\mathbf{t}_1, \dots, \mathbf{t}_N \in [0,1]^d$ *be fixed. Consider* N *sequences of positive constants* $\{k_j(n),\ n \geq 1\}$, $j = 1, \dots, N$, *and* N *classes* $\mathbb{C}_j$ *of Borel subsets of* $[a_j, b_j]^d$ *with* $b_j - a_j = 1$ *for* $j = 1, \dots, N$. *Assume that, for each* $j = 1, \dots, N$, *the sequence* $k(n) = k_j(n)$ *satisfies assumptions* (K.1) *and* (K.2), *and that, for each* $j = 1, \dots, N$, *conditions* (C.1), (C.2), (C.3), (C.4) *and* (C.5) *are satisfied by* $\mathbb{C} = \mathbb{C}_j$ *and* $\mathbf{t} = \mathbf{t}_j$. *Consider* N *independent copies of the uniform empirical process indexed by* $\mathbb{B}$, *denoted by* $\alpha_n^{(j)}(\cdot)$, $j = 1, \dots, N$, *and set*

$$w_n^{(j)}(C) = \left(2\frac{k_j(n)}{n}\log_2 n\right)^{-1/2} \times \alpha_n^{(j)}\left(\mathbf{t}_j + \left(\frac{k_j(n)}{n}\right)^{1/d} C\right), \qquad C \in \mathbb{C}_j,\ j = 1, \dots, N. \tag{2.103}$$

Then the sequence $\{(w_n^{(1)}, \dots, w_n^{(N)}),\ n \geq 1\}$ *is almost surely relatively compact in* $\mathbb{S}(\mathbb{C}_1) \times \cdots \times \mathbb{S}(\mathbb{C}_N)$ *with limit set equal to*

$$\left\{(w_1, \dots, w_N) \in \prod_{j=1}^N \mathbb{S}(\mathbb{C}_j)\colon w_j(C) = \int_C \phi_j(\mathbf{s})\, d\lambda(\mathbf{s}),\ j = 1, \dots, N,\right.$$
$$\left.\sum_{j=1}^N \int_{\mathbb{R}^d} \phi_j^2(\mathbf{s})\, d\lambda(\mathbf{s}) \leq 1\right\}.$$

PROOF. The proof of Proposition 2.1 is obtained by repeating step by step the arguments of the just-given proof of Theorem 1.1 for $j = 1, \dots, N$, in combination with the argument of the proof of Lemma 1 in Finkelstein (1971). We omit the details for the sake of conciseness. □

PROPOSITION 2.2. *Assume that the assumptions of Theorem* 1.1. *hold. Let* $\{\kappa(n),\ n \geq 1\}$ *be a random sequence of positive numbers and let* $\{\nu(n),\ n \geq 1\}$ *be a random sequence of positive integers such that*

$$\kappa(n)/k(n) \to \Lambda \in (0, \infty) \quad \textit{and} \quad \nu(n)/n \to \rho \in (0, \infty) \quad \textit{a.s. as } n \to \infty. \tag{2.104}$$

Let

$$\Theta_n^{(1)}(C) = \left(\frac{\kappa(n)}{n}\right)^{-1/2} \alpha_{\nu(n)}\left(\mathbf{t} + \left(\frac{\kappa(n)}{n}\right)^{1/d} C\right), \qquad C \in \mathbb{C}. \tag{2.105}$$

Then the sequence of functions $\{(2\log_2 n)^{-1/2}\Theta_n^{(1)}(C), C \in \mathbb{C}\}$, $n = 1, 2, \dots$, *is almost surely relatively compact in* $B(\mathbb{C})$ *with limit set equal to* $\mathbb{S}(\mathbb{C})$.

PROOF. First, observe that if the sequence $\{k(n),\ n \geq 1\}$ satisfies (K.1) and (K.2), then it is also the case for the sequence $\{\Lambda\rho \widetilde{k}(n/\rho),\ n \geq 1\}$, where

$\widetilde{k}(n+u) := (1-u)k(n) + uk(n+1)$ for $0 \le u \le 1$. Let

$$\Theta_n^{(2)}(C) = \left(\frac{\Lambda\widetilde{k}(\rho^{-1}n)}{\rho^{-1}n}\right)^{-1/2} \alpha_n\left(\mathbf{t} + \left(\frac{\Lambda\widetilde{k}(\rho^{-1}n)}{\rho^{-1}n}\right)^{1/d} C\right), \qquad C \in \mathbb{C},$$

and

$$\Theta_n^{(3)}(C) = \left(\frac{\Lambda\widetilde{k}(\rho^{-1}\nu(n))}{\rho^{-1}\nu(n)}\right)^{-1/2} \alpha_{\nu(n)}\left(\mathbf{t} + \left(\frac{\Lambda\widetilde{k}(\rho^{-1}\nu(n))}{\rho^{-1}\nu(n)}\right)^{-1/d} C\right), \qquad C \in \mathbb{C}.$$

The fact that, for $i = 2$, $\{(2\log_2 n)^{-1/2}\Theta_n^{(i)}(C),\ C \in \mathbb{C}\}$ is almost surely relatively compact in $B(\mathbb{C})$ with limit set equal to $\mathbb{S}(\mathbb{C})$ is an obvious consequence of Theorem 1.1 taken with the formal replacement of $k(n)$ by $\Lambda\rho\widetilde{k}(n/\rho)$. In view of (2.104) combined with Remark 2.1, this implies, in turn, that the same statement holds for $i = 3$. By (2.104) we get

$$\left(\frac{\Lambda\widetilde{k}(\rho^{-1}\nu(n))}{\rho^{-1}\nu(n)}\right) \Big/ \left(\frac{\kappa(n)}{n}\right) \to 1 \quad \text{a.s. as } n \to \infty.$$

This, in combination with Lemma 2.4, suffices to show that

$$\sup_{C \in \mathbb{C}} |\Theta_n^{(1)}(C) - \Theta_n^{(3)}(C)| \to 0 \quad \text{a.s.},$$

which completes the proof of the proposition. □

Note for further use that the random time change considered in Proposition 2.2 could have been used without modification in the setting of Proposition 2.1.

Let $M \ge 1$ be fixed and let $B_1, \dots, B_M \in \mathbb{B}$ be disjoint and such that

(2.106) $$[0,1]^d = \bigcup_{1 \le j \le M} B_j \quad \text{and} \quad \lambda(B_j) > 0 \quad \text{for } j = 1, \dots, M.$$

Observe that the distribution of $\{\alpha_n(B_i),\ i = 1, \dots, M,\ n \ge 1\}$ is identical to that of $\{\xi_n(\sum_{j=1}^{i}\lambda(B_j)) - \xi_n(\sum_{j=1}^{i-1}\lambda(B_j)),\ i = 1, \dots, M,\ n \ge 1\}$, where ξ_n is as in Theorem A. Let further

(2.107) $$z_{n,j} = (2\log_2 n)^{-1/2}\alpha_n(B_j) \quad \text{for } j = 1, \dots, M.$$

We have the following result due to Finkelstein (1971).

PROPOSITION 2.3. *The sequence* $\{(z_{n,1}, \dots, z_{n,M}), n \ge 1\}$ *is almost surely relatively compact in* $\mathbb{R}^M$, *with limit set equal to*

(2.108) $$\left\{(z_1, \dots, z_M) \in \mathbb{R}^M: \sum_{j=1}^{M} z_j = 0, \sum_{j=1}^{M} \frac{z^2}{\lambda(B_j)} \le 1\right\}.$$

PROOF. See Lemma 1 and Theorem 1 of Finkelstein (1971). □

Note that a direct proof of this proposition may be obtained by similar subsequence arguments and probability inequalities as those used in the proof of Theorem 1.1. This observation is of interest in view of the proof of Theorem 1.2.

PROOF OF THEOREM 1.2. Let $M \geq N$ and let $B_1, \dots, B_M \in \mathbb{B}$ be disjoint, satisfying (2.106) and such that there exists an $h_0 > 0$ with

$$\mathbf{t}_j + hC \subseteq B_j \quad \text{for all } C \in \mathbb{C}_j,\ 0 < h \leq h_0 \text{ and } j = 1, \dots, N. \tag{2.109}$$

Note here that it is always possible, by eventually adding to $\mathbb{C}_0$ a finite number of sets (up to a constant of proportionality), to assume that there exist $B_1, \dots, B_M \in \mathbb{C}_0$ satisfying (2.109).

Set, for $j = 1, \dots, M$,

$$N_j(n) = n\lambda_n(B_j) =: n\lambda(B_j) + (2n \log_2 n)^{1/2} z_{n,j} \quad \text{for } j = 1, \dots, M, \tag{2.110}$$

and

$$N_j^{-1}(n) = \inf\{m \geq 0: N_j(m) = n\} \quad \text{for } j = 1, \dots, M. \tag{2.111}$$

Also, introduce the *conditional empirical processes* indexed by $\mathbb{B}$, defined for $j = 1, \dots, M$ by

$$\alpha_{n,j}(B) = n^{1/2}\left\{\frac{1}{n}\sum_{i=1}^{N_j^{-1}(n)} \mathbf{1}\big(U_i \in B \cap B_j\big) - \frac{\lambda(B \cap B_j)}{\lambda(B_j)}\right\}, \qquad B \in \mathbb{B}. \tag{2.112}$$

Observe that the processes $\alpha_{n,1}, \dots, \alpha_{n,M}$ are mutually independent and independent of the sequence $\{N_1(n), \dots, N_M(n)\}$. Moreover, since $N_j^{-1}(N_j(n)) = n$, we have, for $j = 1, \dots, M$,

$$\begin{aligned} \{N_j(n)\}^{1/2}\alpha_{N_j(n),j}(B) &= \sum_{i=1}^{n} \mathbf{1}(U_i \in B \cap B_j) - N_j(n)\frac{\lambda(B \cap B_j)}{\lambda(B_j)} \\ &= n^{1/2}\alpha_n(B \cap B_j) + n\lambda(B \cap B_j)\left(1 - \frac{N_j(n)}{n\lambda(B_j)}\right). \end{aligned} \tag{2.113}$$

Recall (1.8). By letting $B = \mathbf{t}_j + (k_j(n)/n)^{1/d}C$ in (2.113), it follows from (2.109) that, for all $j = 1, \dots, N$ and $C \in \mathbb{C}_j$, we have for n sufficiently large

$$\begin{aligned} w_{n,j}(C) &= \left(2\frac{k_j(n)}{n}\log_2 n\right)^{-1/2} \alpha_n\left(\mathbf{t}_j + \left(\frac{k_j(n)}{n}\right)^{1/d} C\right) \\ &= \left(\frac{N_j(n)}{n}\right)^{1/2}\left(2\frac{k_j(n)}{n}\log_2 n\right)^{-1/2} \alpha_{N_j(n),j}\left(\mathbf{t}_j + \left(\frac{k_j(n)}{n}\right)^{1/d} C\right) \\ &\quad -\left(\frac{k_j(n)}{2\log_2 n}\right)^{1/2}\lambda(C)\left(1 - \frac{N_j(n)}{n\lambda(B_j)}\right) \\ &=: w'_{n,j}(C) + w''_{n,j}(C). \end{aligned} \tag{2.114}$$

By the law of the iterated logarithm, we have $N_j(n) - n\lambda(B_j) = O((n \log_2 n)^{1/2})$, almost surely as $n \to \infty$. Thus, by (2.114), we have

$$\sup_{C \in \mathbb{C}_j} \left|w''_{n,j}(C)\right| = O\Big(\big(n^{-1}k_j(n)\big)^{1/2}\Big) \to 0 \quad \text{a.s. as } n \to \infty. \tag{2.115}$$

Since the processes $\alpha_{n,j}$ are mutually independent, and independent of the $N_j(n)$, by the same lines as used to establish Propositions 2.1, 2.2 and 2.3, we obtain readily, after a simple but lengthy argument, that the sequence $\{(z_{n,1}, \dots, z_{n,M};\, w'_{n,1}, \dots, w'_{n,N}), n \geq 1\}$ is almost surely relatively compact in $\mathbb{R}^M \times B(\mathbb{C})^N$, with limit set equal to

$$\begin{aligned} \Bigg\{ &(z_1, \dots, z_M;\, w_1, \dots, w_N) \in \mathbb{R}^M \times B(\mathbb{C})^N \colon \sum_{i=1}^{M} z_i = 0, \\ &w_j(C) = \int_C \phi_j(\mathbf{s})\, d\lambda(\mathbf{s}), j = 1, \dots, N, \sum_{i=1}^{M} \frac{z_i^2}{\lambda(B_i)} \\ &\qquad + \int_{\mathbb{R}^d} \sum_{j=1}^{N} \phi_j^2(\mathbf{s})\, d\lambda(\mathbf{s}) \leq 1 \Bigg\}. \end{aligned} \tag{2.116}$$

To complete the proof of Theorem 1.2, it suffices now to use a straightforward modification of the arguments of Kuelbs and Dudley (1980), pages 415 and 416, to show that if the preceding result holds for an arbitrary disjoint sequence $B_1, \dots, B_M \in \mathbb{C}_0$ satisfying (2.109), then the conclusion of the theorem is true. We omit the details of this argument for the sake of brevity. □

3. Applications.

3.1. *Introduction.* The results of this section are given to illustrate how our methods can be applied to describe the limiting behavior of statistics depending locally on the empirical process. We start in subsection 3.2 by extending our theorems to sequences of independent random variables with a common density on $\mathbb{R}^d$. We then investigate multivariate density estimators in subsection 3.3 and the Bahadur–Kiefer representation in subsection 3.4.

3.2. *Local theorems for nonuniform samples.* Let $X_n = (X_n(1), \dots, X_n(d))$, $n = 1, 2, \dots$, be an i.i.d. sequence of $\mathbb{R}^d$-valued random variables with common density f with respect to the Lebesgue measure λ. The following multivariate quantile transformation lemma, in the spirit of Rosenblatt (1952), will be instrumental in the applications of our theorems.

LEMMA 3.1. *Let $\mathbb{X}^* = (X^*(1), \dots, X^*(d)), d \geq 1$, be a random vector taking values in $\mathbb{R}^d$ with distribution function F^* and let $\mathbb{U}^* = (U^*(1), \dots, U^*(d))$ be a uniform $(0,1)^d$ random vector. Let F_1^* denote the distribution function of $X^*(1)$ and let $F_j^*(\cdot \mid x_1, \dots, x_{j-1})$ be a regular conditional distribution function of $X^*(j)$*

given $(X^*(1),\dots,X^*(j-1)) = (x_1,\dots,x_{j-1})$ *for* $2 \le j \le d$. *Define the function of* $\mathbf{x} = (x_1,\dots,x_d)$:

$$H^*(\mathbf{x}) = \Big(F_1^*(x_1),\dots,F_d^*(x_d \mid x_1,\dots,x_{d-1})\Big). \tag{3.1}$$

For $t \in (0,1)$, *set* $\operatorname{inv} F_1^*(t) = \inf\{x\colon F_1^*(x) \ge t\}$, *and, for* $2 \le j \le d$ *whenever* $d \ge 2$, *define* $\operatorname{inv} F_j^*(t \mid x_1,\dots,x_{j-1}) = \inf\{x\colon F_j^*(x \mid x_1,\dots,x_{j-1}) \ge t\}$. *Also, set for* $\mathbf{u} = (u_1,\dots,u_d) \in (0,1)^d$, $G_1^*(u_1) = \operatorname{inv} F_1^*(u_1)$ *and, for* $2 \le j \le d$ *whenever* $d \ge 2$, *define* $G_j^*(u_1,\dots,u_j) = \operatorname{inv} F_j^*(u_j \mid G_1^*(u_1),\dots,G_{j-1}^*(u_1,\dots,u_{j-1}))$. *Let further*

$$G^*(\mathbf{u}) = \big(G_1^*(u_1),\dots,G_d^*(u_1,\dots,u_d)\big). \tag{3.2}$$

Then

$$\mathbb{X}^* =_d G^*(\mathbb{U}^*). \tag{3.3}$$

Moreover, if H^* *is continuously (resp. twice continuously) differentiable in a neighborhood of* $\mathbf{x} \in \mathbb{R}^d$ *with a nonzero Jacobian at* $\mathbf{x}$, *then* G^* *is continuously (resp. twice continuously) differentiable in the neighborhood of* $\mathbf{t} := H^*(\mathbf{x})$ *and such that*

$$|G^*(\mathbf{u}) - G^*(\mathbf{t}) - DG^*(\mathbf{t})(\mathbf{u}-\mathbf{t})| = o\big(|\mathbf{u}-\mathbf{t}|\big) \quad \textit{and} \quad DG^*(\mathbf{t}) = \big(DH^*(x)\big)^{-1}, \tag{3.4}$$

where $DG^*(\mathbf{t})$ [*resp.* $DH^*(\mathbf{x})$] *denotes the differential of* G^* *at* $\mathbf{t}$ (*resp. of* H^* *at* $\mathbf{x}$). *If* $\mathbb{X}^*$ *has a continuous density* f^* *on* $\mathbb{R}^d$, *then the Jacobian of* G^* *at* $\mathbf{t}$ *is equal to*

$$1/f^*\big(G^*(\mathbf{t})\big). \tag{3.5}$$

PROOF. That (3.3) holds is Theorem 6 of Einmahl (1989), while (3.4) is a consequence of the inverse function theorem. For (3.5), note that $DH^*(\mathbf{x})$ is a lower diagonal matrix, and thus its determinant is the product of its diagonal elements which is easily checked to be equal to $f^*(\mathbf{x})$. Since $DG^*(\mathbf{t}) = (DH^*(\mathbf{x}))^{-1}$, the assertion follows. □

Our next lemma gives a local version of Lemma 3.1, where we assume that the density f of $\mathbb{X} \in \mathbb{R}^d$ is continuously differentiable in a neighborhood of $\mathbf{x} \in \mathbb{R}^d$. This ensures, via a suitable definition of F^*, that the function H^* defined in (3.1) is continuously differentiable in a neighborhood of $\mathbf{x}$. This assumption could be weakened at the price of lengthy technicalities. Since our aim is to present applications of our main theorems, we will limit ourselves to the present setting which simplifies greatly the exposition. Let $\mathbb{X}$ denote a random variable with the same distribution as X_n, $n = 1,2,\dots$, and let $\mathbb{U} = (U(1),\dots,U(d))$ denote a uniform $(0,1)^d$ random vector.

LEMMA 3.2. *Assume that* f *is continuously (resp. twice continuously) differentiable in a neighborhood of* $\mathbf{x} \in \mathbb{R}^d$ *and that* $f(\mathbf{x}) > 0$. *Then there exists a* *point* $\mathbf{t} \in (0,1)^d$, *a neighborhood* V *of* $\mathbf{x}$, *a neighborhood* N *of* $\mathbf{t}$ *and a one-to-one mapping* G *of* N *onto* V *such that* $\mathbf{x} = G(\mathbf{t})$ *and*

$$\mathbb{X}\mathbf{1}(\mathbb{X} \in V) =_d G(\mathbb{U})\mathbf{1}(\mathbb{U} \in N). \tag{3.6}$$

Moreover the function G *is of the form*

$$G(\mathbf{u}) = \big(G_1(u_1), G_2(u_1, u_2), \dots, G_d(u_1, \dots, u_d)\big), \tag{3.7}$$

with $\mathbf{u} = (u_1, \dots, u_d)$ *and such that, for each* $1 \le j \le d$, $G_i(u_1, \dots, u_j)$ *is an increasing function of* u_j. *In addition*, G *is continuously (resp. twice continuously) differentiable on* N *and such that*

$$\big|G(\mathbf{u}) - G(\mathbf{t}) - \{f(\mathbf{x})\}^{-1/d}(\mathbf{u} - \mathbf{t})\big| = o\big(|\mathbf{u} - \mathbf{t}|\big). \tag{3.8}$$

PROOF. Let $V^* = \mathbf{x} + (-c, c)^d$, where $c > 0$ is chosen in such a way that f is continuously (resp. twice continuously) differentiable and bounded away from 0 on V^*. Denote by $\mathbb{X}^*$ a random variable following the conditional distribution of $\mathbb{X}$ given $\mathbb{X} \in V^*$, with density equal to $f^* = f/P(\mathbb{X} \in V^*) =: f/q$. Let F^* denote the distribution function of $\mathbb{X}^*$ and let H^* be as in (3.1). Making use of Lebesgue's theorem, it is readily verified that H^* is continuously (resp. twice continuously) differentiable on V^* with nonzero Jacobian. Thus, by an application of Lemma 3.1, we obtain that (3.2)–(3.5) hold. Now $\mathbf{u} = \mathbf{u}(\mathbf{v})$ and $\mathbf{v} = \mathbf{v}(\mathbf{u})$ be related via the formula

$$\mathbf{v} = \mathbf{t} + \Big\{f^*\big(G^*(\mathbf{t})\big)\Big\}^{1/d} DG^*(\mathbf{t})(\mathbf{u} - \mathbf{t}). \tag{3.9}$$

Recalling from (3.5) that the Jacobian of $DG^*(\mathbf{t})$ equals $1/f^*(G^*(\mathbf{t}))$, we see that the Jacobian of the linear mapping $\mathbf{u} \to \mathbf{u}(\mathbf{v})$ equals 1. Thus, if $\mathbb{U}^*$ is a uniform $(0,1)^d$ random vector, then the random vector $\mathbb{V} := \mathbf{v}(\mathbb{U})$ has density equal to 1 on the set $E := \mathbf{v}((0,1)^d)$. Set, for convenience, $\overline{G}(\mathbf{v}) = G^*(\mathbf{u}(\mathbf{v}))$ and notice that $\overline{G}(\mathbf{t}) = \mathbf{t}$. By (3.9) we may rewrite (3.4) as

$$\left|\overline{G}(\mathbf{v}) - \overline{G}(\mathbf{t}) - \left\{\frac{1}{q}f\big(\overline{G}(\mathbf{t})\big)\right\}^{-1/d}(\mathbf{v} - \mathbf{t})\right| = o\big(|\mathbf{v} - \mathbf{t}|\big). \tag{3.10}$$

By (3.2) $DG^*(\mathbf{t})$ is lower triangular and therefore $\{DG^*(\mathbf{t})\}^{-1}$ is also lower triangular. This, in combination with (3.9), implies the existence of constants $\{a_{ij}\}$ and $\{b_j\}$ such that, with $\mathbf{v} = (v_1, \dots, v_d)$,

$$\mathbf{u}(\mathbf{v}) = (a_{11}v_1 + b_1, \dots, a_{d1}v_1 + \cdots + a_{dd}v_d + b_d). \tag{3.11}$$

Recalling that $\overline{G}(\mathbf{v}) = G^*(\mathbf{u}(\mathbf{v}))$, we deduce from (3.2) and (3.11) the existence of functions $\overline{G}_1, \dots, \overline{G}_d$ such that

$$\overline{G}(\mathbf{v}) = \big(\overline{G}_1(v_1), \dots, \overline{G}_d(v_1, \dots, v_d)\big). \tag{3.12}$$

We may now rewrite (3.3) into

$$\mathbb{X}^* =_d \overline{G}(\mathbb{V}). \tag{3.13}$$

By (3.9) $\mathbf{v}(\mathbf{t}) = \mathbf{t} \in (0, 1)^d$, and therefore the set $E = \mathbf{v}((0, 1)^d)$ is a neighborhood of $\mathbf{t}$. This, in turn, implies that $E_1 := E \cap (0, 1)^d$ is also a neighborhood of $\mathbf{t}$. Therefore, the open mapping theorem implies that $V := \overline{G}(E_1) \subseteq V^* = \overline{G}(E)$ is a neighborhood of $\mathbf{x} = \overline{G}(\mathbf{t})$. By (3.13) it follows that

$$\mathbb{X}^*\mathbf{1}(\mathbb{X}^* \in V) =_d \overline{G}(\mathbb{V})\mathbf{1}(\mathbb{V} \in E_1). \tag{3.14}$$

Since $\mathbb{V}$ has a density equal to 1 on E_1, we may replace $\mathbb{V}$ in (3.14) by any other random variable with density equal to 1 on E_1. Now let $\mathbb{U}$ denote a uniform $(0, 1)^d$ random variable and set

$$\mathbb{V}_1 = \mathbf{t} + q^{-1/d}(\mathbb{U} - \mathbf{t}). \tag{3.15}$$

Observe that $q = P(\mathbb{X} \in V) = P(\mathbb{V}_1 \in (0, 1)^d)$. Thus, by (3.15), the distribution of $\mathbb{V}_1$ conditional on $\mathbb{V}_1 \in (0,1)^d$ is uniform on $(0,1)^d$ (with density equal to 1 on E_1). Moreover, the conditional distribution of $\mathbb{X}$ given $\mathbb{X} \in V$ is equal to the distribution of $\mathbb{X}^*$. It follows therefore from (3.14) that

$$\mathbb{X}\mathbf{1}(\mathbb{X} \in V) =_d \overline{G}(\mathbb{V}_1)\mathbf{1}(\mathbb{V}_1 \in E_1) = G(\mathbb{U})\mathbf{1}(\mathbb{U} \in N), \tag{3.16}$$

where we set

$$G(\mathbf{u}) := \overline{G}\big(\mathbf{t} + q^{-1/d}(\mathbf{u} - \mathbf{t})\big) \quad \text{and} \quad N := \mathbf{t} + q^{1/d}\big(E_1 - \mathbf{t}\big). \tag{3.17}$$

By combining (3.10) and (3.17), we obtain (3.8), while (3.16) yields (3.6). Finally, (3.7) follows from (3.12) and (3.17). The proof of Lemma 3.2 is now complete. □

Our next result gives an example of how Theorem 1.1 may be applied to sequences of random variables with a nonuniform distribution.

THEOREM 3.1. *Let $X_n = (X_n(1), \dots, X_n(d))$, $n = 1, 2, \dots$, be a sequence of independent random vectors with common probability measure μ on $\mathbb{R}^d$. Assume that μ has a density $f = d\mu/d\lambda$, twice continuously differentiable in a neighborhood of $\mathbf{x} = (x(1), \dots, x(d)) \in \mathbb{R}^d$ and such that $f(\mathbf{x}) > 0$. Let $\{k(n), n \geq 1\}$ satisfy* (K.1) *and* (K.2). *Then, for any fixed $\mathbf{c} = (c_1, \dots, c_d) \in \mathbb{R}^d$, the sequence of functions of $\mathbf{s} = (s_1, \dots, s_d) \in [0, 1]^d$ defined by*

$$\begin{aligned} \Psi_n(\mathbf{s}) = {} & \big(2k(n)f(\mathbf{x})\log_2 n\big)^{-1/2} \\ & \times \sum_{i=1}^{n} \Bigg\{ \mathbf{1}\Bigg(X_n(j) - x(j) \in \left(\frac{k(n)}{n}\right)^{1/d}(c_j, c_j + s_j)\ for\ j = 1, \dots, d\Bigg) \\ & \qquad - \mu\Bigg(\prod_{j=1}^{n}\Bigg[x(j) + \left(\frac{k(n)}{n}\right)^{1/d} c_j, x(j) + \left(\frac{k(n)}{n}\right)^{1/d}(c_j + s_j)\Bigg]\Bigg)\Bigg\} \end{aligned} \tag{3.18}$$

is almost surely relatively compact in the set of bounded functions on $[0,1]^d$, *endowed with the uniform topology. The corresponding limit set consists of all functions of the form*

$$\psi(\mathbf{s}) = \int_0^{s_1}\cdots\int_0^{s_d}\phi(\mathbf{u})\,d\lambda(\mathbf{u}) \quad \textit{with} \int_0^1\cdots\int_0^1 \phi(\mathbf{u})\,d\lambda(\mathbf{u}) \le 1. \tag{3.19}$$

PROOF. Let $\mathbf{t}$, G, N and V be as in Lemma 3.2, with $G(\mathbf{t}) = \mathbf{x}$. For the sake of notational simplicity, let $\mathbf{x} = \mathbf{0} = (0,\dots,0)$, $\mathbf{t} = (\frac{1}{2},\dots,\frac{1}{2})$ and $\mathbf{c} = \mathbf{a} = (-\frac{1}{2},\dots,-\frac{1}{2})$, where $\mathbf{a}$ is as in Section 1. Further let $\mathbf{u} = (u_1,\dots,u_d)$, $\mathbf{v} = (v_1,\dots,v_d)$, $\mathbf{z} = (z_1,\dots,z_d)$ and define the class $\mathbb{C}' = \{C'(z): \mathbf{z} \in [-\frac{1}{2},\frac{1}{2}]^d\}$, of subsets of $[-\frac{1}{2},\frac{1}{2}]^d$, where

$$C'(\mathbf{z}) := \left\{\mathbf{v} \in K: v_1\{f(\mathbf{0})\}^{-1/d} \le z_1,\dots,v_d\{f(\mathbf{0})\}^{-1/d} \le z_d\right\} \tag{3.20}$$

and $K := [-\frac{1}{2}\{f(\mathbf{0})\}^{1/d}, \frac{1}{2}\{f(\mathbf{0})\}^{1/d}]^d$. Consider likewise the class of $\mathbb{C}''$ of subsets of $[-\frac{1}{2},\frac{1}{2}]^d$ defined by $\mathbb{C}'' = \{\gamma C''(\mathbf{z},\rho): \mathbf{z} \in [-\frac{1}{2},\frac{1}{2}]^d,\ \frac{1}{2} \le \gamma \le 1,\ 0 \le \rho \le \rho_0\}$, where

$$C''(\mathbf{z},\rho) = \begin{cases} \left\{\mathbf{v} \in K: \dfrac{G(\mathbf{t}+\rho\mathbf{v})}{\rho\{f(\mathbf{0})\}^{-1/d}} \in C'(\mathbf{z})\right\}, & \text{for } 0 < \rho \le \rho_0, \\ C'(\mathbf{z}), & \text{for } \rho = 0. \end{cases} \tag{3.21}$$

Here, $\rho_0 > 0$ is chosen in such a way that $\mathbf{t} + 2\rho_0[-\frac{1}{2},\frac{1}{2}]^d \subseteq N$. Recalling that $G(\mathbf{t}) = \mathbf{0}$, we infer from (3.8) that $|G(\mathbf{t}+\rho\mathbf{v}) - \rho\{f(\mathbf{0})\}^{-1/d}\mathbf{v}| = o(|\rho|)$ uniformly over $\mathbf{v} \in K$ as $\rho \to 0$, and obtain readily from (3.20) and (3.21) that, uniformly over $\mathbf{z} \in [-\frac{1}{2},\frac{1}{2}]^d$, we have

$$\lim_{\rho\downarrow 0} d_\lambda\big(C'(\mathbf{z}), C''(\mathbf{z},\rho)\big) = 0. \tag{3.22}$$

Letting $a = -\frac{1}{2}$ and $b = \frac{1}{2}$, it is easily checked from (3.20) and (3.21) that either of the classes $\mathbb{C} = \{f(\mathbf{0})\}^{-1/d}\mathbb{C}'$ or $\mathbb{C} = \{f(\mathbf{0})\}^{-1/d}\mathbb{C}''$ satisfies (C.1), (C.2) and (C.4). To check that (C.3) holds is more difficult. In the first place, we make use of the results of Donsker (1952) (for $d = 1$) and Dudley (1966) (for $d \ge 2$) who showed that the class of left orthants is a μ-Donsker class for any probability measure μ on $\mathbb{R}^d$. From there, it follows readily that $\{f(\mathbf{0})\}^{-1/d}\mathbb{C}' - \mathbf{a}$ is a λ_0-Donsker class [i.e., $\mathbb{C} = \{f(\mathbf{0})\}^{-1/d}\mathbb{C}'$ satisfies (C.3)]. Toward proving that a similar result holds for $\mathbb{C} = \{f(\mathbf{0})\}^{-1/d}\mathbb{C}''$, introduce the function of $\mathbf{v} \in K$ and $(\gamma,\rho,\mathbf{z}) \in L := [\frac{1}{2},1] \times [0,1] \times [-\frac{1}{2},\frac{1}{2}]^d$ defined by

$$\Phi(\mathbf{v};\gamma,\rho,\mathbf{z}) = \begin{cases} \min\limits_{1\le j\le d}\left(z_j - G_j\left(\mathbf{t} + \dfrac{\rho}{\gamma}\mathbf{v}\right)\Big/\left(\rho\{f(\mathbf{0})\}^{-1/d}\right)\right), & \text{for } 0 < \rho \le \rho_0, \\ \min\limits_{1\le j\le d}\left(z_j - \dfrac{1}{\gamma}v_j\right), & \text{for } \rho = 0, \end{cases} \tag{3.23}$$

where $G_j(\mathbf{u}) := G_j(u_1, \ldots, u_d)$, $1 \le j \le d$, is as in (3.7). Keeping in mind that $(\mathbf{v}, \gamma, \rho, \mathbf{z})$ varies in the compact set $K \times L$, we will show that Φ satisfies the following three conditions:

(i) For each fixed $(\gamma, \rho, \mathbf{z}) \in L$, $\Phi(\mathbf{v}; \gamma, \rho, \mathbf{z})$ is a continuous function of $\mathbf{v} \in K$;
(ii) Uniformly over $\mathbf{v} \in K$, the function $(\gamma, \rho, \mathbf{z}) \to \Phi(\mathbf{v}; \gamma, \rho, \mathbf{z})$ satisfies a Lipschitz condition of order 1;
(iii) $\lambda(\{\mathbf{v} \in K: |\Phi(\mathbf{v}; \gamma, \rho, \mathbf{z})| \le \varepsilon\}) = O(\varepsilon)$ uniformly over $(\gamma, \rho, \mathbf{z}) \in L$.

That (i) holds follows from the continuity of G. The fact, following from Lemma 3.2, that G is twice continuously differentiable on N is used for (ii), since then we may write $G(\mathbf{t} + (\rho/\gamma)\mathbf{v}) = (\rho/\gamma)\{f(0)\}^{-1/d}\mathbf{v} + R((\rho/\gamma)\mathbf{v})$, where R is twice continuously differentiable with differential at $\mathbf{0}$ equal to $DR(\mathbf{0}) = 0$. Observe that if a function g has continuous second derivative g'' and first derivative g' such that $g'(0) = 0$, then

$$\left|\frac{d}{dx}\left(\frac{g(x)}{x}\right)\right| = \left|\frac{g'(x)}{x} - \frac{g(x)}{x^2}\right|$$

is bounded in neighborhood of 0. From there, we obtain readily that $\rho^{-1}H((\rho/\gamma)\mathbf{v})$ is (uniformly over $\mathbf{v} \in K$) Lipschitz in (ρ, γ), so that (ii) holds. The proof of (iii) is obtained through similar arguments.

Conditions (i)–(iii) are required to apply Theorem 2.3 of Gaenssler (1984) which implies that the class of sets $\gamma C''(\mathbf{z}, \rho) = \{\mathbf{v} \in K: \Phi(\mathbf{v}; \gamma, \rho, \mathbf{z}) \ge 0\}$ [recall (3.21) and (3.23)] and is a μ-Donsker class, where here μ denotes the uniform distribution on K. This, in turn, implies that $\mathbb{C} = \{f(0)\}^{-1/d}\mathbb{C}'$ satisfies (C.3).

Now let $\{U_n, n \ge 1\}$ be an i.i.d. sequence of uniform $(0,1)^d$ random variables. In view of (3.6), set, without loss of generality for $n \ge 1$ and if $U_n \in N$ (or equivalently if $X_n \in V$),

$$X_n = G(U_n). \tag{3.24}$$

Observe from (3.20), (3.21) and (3.24) that the following equalities of events hold for each $\mathbf{z} \in K$ and $0 \le \rho \le \rho_0$. We have

$$\begin{aligned}\left\{X_n \in \rho\{f(\mathbf{0})\}^{-1/d}C'(\mathbf{z})\right\} &= \left\{G(U_n) \in \rho\{f(\mathbf{0})\}^{-1/d}C'(\mathbf{z})\right\}\\ &= \left\{U_n \in \mathbf{t} + \rho C''(\mathbf{z}, \rho)\right\}.\end{aligned} \tag{3.25}$$

Recalling (1.4) and (3.9) and setting $\rho = \{k(n)/n\}^{1/d}$ in (3.25), we see that, for all n sufficiently large and $\mathbf{s} \in [0,1]^d$,

$$\begin{aligned}\Psi_n(\mathbf{s}) &= \left(2k(n)f(\mathbf{0})\log_2 n\right)^{-1/2}\sum_{i=1}^{n}\left\{\mathbf{1}\left(X_n \in \left(\frac{k(n)}{n}\right)^{1/d}(f(\mathbf{0}))^{-1/d}C'(\mathbf{s}+\mathbf{c})\right)\right.\\ &\qquad\left. - P\left(X_n \in \left(\frac{k(n)}{n}\right)^{1/d}(f(\mathbf{0}))^{-1/d}C'(\mathbf{s}+\mathbf{c})\right)\right\}\\ &= \left(2f(\mathbf{0})\log_2 n\right)^{-1/2}\Theta_n\left(C''\left(\mathbf{s}+\mathbf{c}, \left\{\frac{k(n)}{n}\right\}^{1/d}\right), \left\{\frac{k(n)}{n}\right\}^{1/d}\right).\end{aligned} \tag{3.26}$$

Making use of the just-proven fact that both $\mathbb{C} = \{f(\mathbf{0})\}^{-1/d}\mathbb{C}'$ and $\mathbb{C} = \{f(\mathbf{0})\}^{-1/d}$ $\mathbb{C}''$ satisfy (C.1)–(C.4), we can apply Lemma 2.4 to $\mathbb{C} = \{f(\mathbf{0})\}^{-1/d}(\mathbb{C}' \cup \mathbb{C}'')$. In view of (3.26), this lemma, in combination with (3.22), implies that almost surely

$$\lim_{n\to\infty}(2\log_2 n)^{-1/2}\sup_{\mathbf{s}\in[0,1]^d}\left|\Psi_n(\mathbf{s}) - f(\mathbf{0})^{-1/2}\Theta_n\left(C'(s+\mathbf{c}), \left\{\frac{k(n)}{n}\right\}^{1/d}\right)\right|$$

$$\begin{aligned}&= \lim_{n\to\infty}\left(2f(\mathbf{0})\log_2 n\right)^{-1/2}\\ &\quad\times\sup_{\mathbf{s}\in[0,1]^d}\left|\Theta_n\left(C''\left(\mathbf{s}+\mathbf{c}, \left\{\frac{k(n)}{n}\right\}^{1/d}\right), \left\{\frac{k(n)}{n}\right\}^{1/d}\right)\right.\\ &\qquad\left. - \Theta_n\left(C'(\mathbf{s}+\mathbf{c}), \left\{\frac{k(n)}{n}\right\}^{1/d}\right)\right| = 0.\end{aligned} \tag{3.27}$$

Now it is easy to enlarge the class $\mathbb{C} = \{f(\mathbf{0})\}^{-1/d}\mathbb{C}'$ to also satisfy (C.5). From Theorem 1.1 applied to this enlarged class, we readily obtain from (3.27) that the sequence of functions

$$\left\{\left(2f(\mathbf{0})\log_2 n\right)^{-1/2}\Theta_n\left(C''\left(\mathbf{s}+\mathbf{c}, \left\{\frac{k(n)}{n}\right\}^{1/d}\right), \left\{\frac{k(n)}{n}\right\}^{1/d}\right), \mathbf{s}\in[0,1]^d\right\},$$
$$n = 1, 2, \dots,$$

is a.s. relatively compact in the set of bounded functions on $[0,1]^d$. By (1.7) the limit set is composed of all functions of the form

$$\begin{aligned}\Psi(\mathbf{s}) &= \int_{[f(\mathbf{0})]^{-1/d}C'(\mathbf{s}+\mathbf{c})}\phi(\mathbf{u}+\mathbf{c})\,d\lambda(\mathbf{u}+\mathbf{c})\\ &= \int_0^{s_1}\cdots\int_0^{s_d}\phi(\mathbf{u})d\lambda(\mathbf{u}) \quad\text{with}\quad \int_0^1\cdots\int_0^1\phi^2(\mathbf{u})d\lambda(\mathbf{u}) \le 1.\end{aligned} \tag{3.28}$$

Since (3.28) coincides with (3.19), the proof of Theorem 3.1 is completed by (3.27). □

REMARK 3.1. In Theorem 3.1 the assumption that f is twice continuously differentiable in a neighborhood of $\mathbf{x} \in \mathbb{R}^d$ is only required for $d \ge 2$. When $d = 1$ we can use a direct argument to show that the conclusion of the theorem remains valid under the much weaker condition that the distribution function F of X_n, $n = 1, 2, \dots$, is differentiable at $x = \mathbf{x}$ with differential $f(\mathbf{x}) > 0$. This is achieved as follows. To avoid technicalities, we limit ourselves to the case where F is continuous and strictly increasing on $\mathbb{R}$. Let $U_n = F(X_n)$, $n = 1, 2, \dots$, and observe that the random variables $\{U_n,\ n \ge 1\}$ are i.i.d. uniform $(0, 1)$. Next, recalling (1.3) and setting $s = s_1$, $\mathbf{x} = x = x_1$ and $c = c_1$ in (3.18), we see that

$$\Psi_n(s) = (2\log_2 n)^{-1/2}\Theta_n\Bigg(\bigg(\frac{k(n)}{n}f(x)\bigg)^{-1}\bigg(F\bigg(x+\frac{k(n)}{n}c\bigg)-F(x), \quad (3.29)$$
$$F\bigg(x+\frac{k(n)}{n}(c+s)\bigg)-F(x)\bigg), F(x), \frac{k(n)}{n}f(x)\Bigg).$$

Making use of the fact [see, e.g., Donsker (1952) for (C.3)] that the class $\mathbb{C} = \{(y,z]: -\frac{1}{2} \le y \le z \le \frac{1}{2}\}$ satisfies (C.1)–(C.5), we can apply Theorem 1.1 to $\mathbb{C}$. We so obtain that, for any $A > 0$, the sequence of functions of (y,z) defined by

$$\overline{\Psi}_n(y,z) = (2\log_2 n)^{-1/2}\Theta_n\bigg((y,z], F(x), \frac{k(n)}{n}f(x)\bigg)$$

for $-A \le y \le z \le A$ is almost surely relatively compact with respect to the uniform topology. The limit set consists of all functions of the form

$$\overline{\Psi}(y,z) = \int_y^z \phi(t)\,dt \quad \text{with} \int_{-\infty}^{\infty} \phi^2(t)\,dt \le 1. \tag{3.30}$$

An easy argument based on the uniform equicontinuity of the functions Ψ in (3.30) enables one to show that the sequence $\overline{\Psi}_n(g_n(y), g_n(z))$ is also relatively compact with limit set characterized by (3.30), whenever g_n is a sequence of nondecreasing functions such that, for any $M > 0$,

$$\lim_{n\to\infty}\Big(\sup_{|z|\le M}|g_n(z) - z|\Big) = 0. \tag{3.31}$$

Now let

$$g_n(z) = \bigg(\frac{k(n)}{n}f(x)\bigg)^{-1}\bigg(F\bigg(x+\frac{k(n)}{n}z\bigg)-F(x)\bigg).$$

Since this function obviously satisfies (3.31), we see that the conclusion of Theorem 3.1 holds.

REMARK 3.2. The argument of Remark 3.1 can also be applied for an arbitrary $d \ge 1$ if $F(x_1, \dots, x_d) = \prod_{j=1}^{d} F_j(x_j)$, in which case we let $G_j(s) = \inf\{x: F_j(x) \ge s\}$ for $0 < s < 1$. By replacing formally the function G of (3.7) by

$$G(\mathbf{u}) = \big(G_1(u_1), \dots, G_d(u_d)\big), \tag{3.32}$$

with G_j, $1 \le j \le d$, as above, we see that the conclusion of Theorem 3.1 holds whenever F_j is differentiable at x_j with differential $f_j(x_j) > 0$ for $j = 1, \dots, d$.

3.3. *Density estimation.* Let $X, X_1, X_2, \ldots$ be a sequence of independent and identically distributed $\mathbb{R}^d$-valued random variables with common distribution function F and density f (with respect to the Lebesgue measure λ). Much attention has been directed to the problem of estimating f by nonparametric techniques [see, e.g., Devroye and Györfi (1985) for a survey]. We will consider here the Parzen–Rosenblatt kernel density estimator [Parzen (1962) and Rosenblatt (1956)], defined by

$$f_n(\mathbf{x}) = \frac{1}{nh_n^d} \sum_{i=1}^{n} K\left(\frac{\mathbf{x} - X_i}{h_n}\right), \tag{3.33}$$

where $\{h_n,\ n \geq 1\}$ is a sequence of positive constants, and the kernel $K(\cdot)$ is a function such that:

(A.1) $\int_{\mathbb{R}^d} K(\mathbf{u})\, d\lambda(\mathbf{u}) = 1$.

(A.2) $K(\cdot)$ is of bounded variation on $\mathbb{R}^d$ in the sense of Hardy and Krause [see, e.g., Hobson (1937). In particular, this condition is fulfilled when (A.3) below holds, and K has bounded partial derivatives of order 2].

(A.3) There exists a $c < \infty$ such that $K(\mathbf{u}) = 0$ for $|\mathbf{u}| \geq c/2$.

Hall (1981) proved the following theorem for $d = 1$.

THEOREM C. *Let $d = 1$. Assume that $F(x) = P(X_1 \leq x)$ satisfies a Lipschitz condition of order* 1 *in a neighborhood of* $\mathbf{x}$ *and that $F'(x) = f(x)$ is defined at* $\mathbf{x}$. *Assume further that*

$$h_n \downarrow 0 \quad \text{and} \quad nh_n \uparrow \tag{3.34}$$

and that

$$\frac{nh_n \log_2 n}{\log^4 n} \to \infty \quad \text{as } n \to \infty. \tag{3.35}$$

Let $K(\cdot)$ satisfy (A.1), (A.2) *and* (A.3). *Then*

$$\begin{aligned} \limsup_{n\to\infty} \pm\Big(f_n(\mathbf{x}) - E\big(f_n(\mathbf{x})\big)\Big)\left(\frac{nh_n}{2\log_2 n}\right)^{1/2} \\ = \left(f(\mathbf{x}) \int_{-\infty}^{\infty} K^2(\mathbf{u})\, d\mathbf{u}\right)^{1/2} \quad a.s. \end{aligned} \tag{3.36}$$

Observe that the conditions of Theorem C imply the existence of the Lebesgue derivative f of F in a neighborhood of $\mathbf{x}$.

In the section we will prove the following result.

THEOREM 3.2. *When $d \geq 2$ assume that f is twice continuously differentiable in a neighborhood of $\mathbf{x} \in \mathbb{R}^d$, and when $d = 1$ assume that F is differentiable at*

$\mathbf{x} \in \mathbb{R}$ *with positive differential* $f(\mathbf{x})$. *Assume further that* (3.34) *holds and that*

$$\frac{nh_n^d}{\log_2 n} \to \infty \quad \text{as } n \to \infty. \tag{3.37}$$

Let $K(\cdot)$ *satisfy* (A.1), (A.2) *and* (A.3). *Then*

$$\begin{aligned}\limsup_{n\to\infty} \pm\Big(f_n(\mathbf{x}) - E\big(f_n(\mathbf{x})\big)\Big)\left(\frac{nh_n^d}{2\log_2 n}\right)^{1/2} \\ = \left(f(\mathbf{x})\int_{\mathbb{R}^d} K^2(\mathbf{u})\,d\lambda(\mathbf{u})\right)^{1/2} \quad a.s.\end{aligned} \tag{3.38}$$

PROOF. We limit ourselves to the main ideas underlining the proof, which we give for $d = 1$, in order to avoid technicalities related to integration by parts. Further details concerning this argument are given in proof of Theorem 4.1 of Deheuvels and Mason (1992a). Let $x = \mathbf{x}$,

$$F_n(x) = n^{-1}\#\{X_i \le x:\ 1 \le i \le n\} \quad \text{and} \quad F(x) = P(X_1 \le x). \tag{3.39}$$

We have, by (A.3),

$$f_n(x) - E\big(f_n(x)\big) = \frac{1}{h_n}\int_{-c}^{c} K(u)\,d\big\{F_n(x + h_n u) - F(x + h_n u)\big\}. \tag{3.40}$$

Integrating by parts, we get from (3.40)

$$\begin{aligned}&\pm\left(\frac{nh_n}{2\log_2 n}\right)^{1/2}\Big(f_n(x) - E\big(f_n(x)\big)\Big) \\ &\quad = \pm\int_{-c}^{c}\left\{\frac{\sqrt{n}\big(F_n(x+h_n u) - F_n(x - h_n c) - F(x + h_n u) + F(x - h_n c)\big)}{\sqrt{2h_n \log_2 n}}\right\} \\ &\qquad \times d\big(-K(u)\big).\end{aligned} \tag{3.41}$$

Next, an application of Theorem 3.1 shows that the set of limit points on the right-hand side of (3.41) is almost surely equal to

$$\left\{\pm\sqrt{f(x)}\int_{-c}^{c}\left\{\int_{-c}^{u}\phi(s)\,ds\right\}d\big(-K(u)\big):\ \int_{-c}^{c}\phi^2(u)\,du \le 1\right\} \quad \text{a.s.}$$

We conclude by integrating by parts once more to obtain that this limit set equals

$$\left\{\pm\sqrt{f(x)}\int_{-c}^{c}\phi(u)K(u)\,du:\ \int_{-c}^{c}\phi^2(u)\,du \le 1\right\}.$$

The conclusion (3.38) now follows by Schwarz's inequality. □

REMARK 3.3. The assumption that f is twice continuously differentiable in a neighborhood of $\mathbf{x} \in \mathbb{R}^d$ is only required for $d \geq 2$, since our proof then relies on an application of Theorem 3.1. For $d = 1$, we see from Remark 3.1 that the conclusion of Theorem 3.2 holds under the weaker assumption that F is differentiable at $x = \mathbf{x}$ with positive differential $f(\mathbf{x})$. Likewise, by Remark 3.2, the same result holds if X has independent coordinates whose distributions follow this condition. In particular, these requirements are satisfied under the assumptions of Theorem C, that is, for $d = 1$, when $F(x) = P(X_1 \leq x)$ satisfies a Lipschitz condition of order 1 in a neighborhood of $\mathbf{x}$, and $F'(x) = f(x)$ is defined at $\mathbf{x}$. This shows that Theorem C is valid when (3.35) is replaced by (3.37). In view of the results of Deheuvels (1974) [see also Devroye (1979)], the latter condition is sharp.

REMARK 3.4. By repeating the preceding arguments with Theorem 1.2 replacing Theorem 1.1, we obtain the following extension of Theorem 3.2. Let $\mathbf{x}_i, i = 1, \dots, N$, be fixed, distinct points of $\mathbb{R}^d$ such that the assumptions of Theorem 3.2 hold for $\mathbf{x} = \mathbf{x}_i$ and $i = 1, \dots, N$. Then, under (3.34) and (3.37), the limit set of the sequence of random vectors of $\mathbb{R}^N$ defined for $n = 1, 2, \dots$, by

$$\left\{ \left(\frac{f_n(\mathbf{x}_i) - E(f_n(\mathbf{x}_i))}{\sqrt{f(\mathbf{x}_i)}} \right) \left(\frac{nh_n^d}{2 \log_2 n} \right)^{1/2} \left(\int_{\mathbb{R}^d} K^2(u)\, du \right)^{-1/2}, \ i = 1, \dots, N \right\},$$

is almost surely equal to the (Euclidean) unit ball of $\mathbb{R}^N$.

3.4. *Bahadur–Kiefer representations.* Let $U_1, U_2, \dots$ be an i.i.d. sequence of uniform $(0, 1)$ random variables. Set further

$$U_n(x) = n^{-1} \#\{U_i \leq x\colon\ 1 \leq i \leq n\}, x \in \mathbb{R},$$
$$V_n(s) = \inf\{x \geq 0\colon\ U_n(x) \geq s\}, 0 \leq s \leq 1, \tag{3.42}$$

and consider the *uniform empirical process* $\{\alpha_n(t), 0 \leq t \leq 1\}$ and the *uniform empirical quantile process* $\{\beta_n(t), 0 \leq t \leq 1\}$ defined respectively by

$$\alpha_n(x) = n^{1/2}\big(U_n(x) - x\big), \quad x \in \mathbb{R},$$
$$\beta_n(s) = n^{1/2}\big(V_n(s) - s\big), \quad 0 \leq s \leq 1. \tag{3.43}$$

The *Bahadur–Kiefer process* [see, e.g., Deheuvels and Mason (1990a)] is then given by

$$R_n(t) = \alpha_n(t) + \beta_n(t) \quad \text{for } 0 \leq t \leq 1. \tag{3.44}$$

The process $\{R_n(t),\ 0 \leq t \leq 1\}$ was introduced by Bahadur (1966) and later studied by Kiefer (1967, 1970). Kiefer (1967) proved that, for any fixed $0 \leq t \leq 1$, we have

$$\limsup_{n \to \infty} \pm n^{1/4} (2 \log_2 n)^{-3/4} R_n(t) = \big(t(1-t)\big)^{1/4} 2^{1/2} 3^{-3/4} \quad \text{a.s.} \tag{3.45}$$

Recently, Deheuvels (1992) and Deheuvels and Mason (1992b) have provided new proofs of (3.45), the latter two authors also giving a simple explanation of the mechanism which generates the unusual limiting constant in this expression. Theorem 3.3 below extends these investigations by describing the joint limiting behavior of $\{(R_n(t_1), \dots, R_n(t_N)),\ n \geq 1\}$ for arbitrary fixed $0 < t_1 < \cdots < t_N < 1$. The proof of this theorem, obtained by a direct argument based on Theorem 1.2, gives a new proof of the main result of Deheuvels and Mason (1992b), corresponding to $N = 1$.

Throughout, we will use the following notation and assumptions. We assume that $t_0 = 0 < t_1 < \cdots < t_N < 1 = t_{N+1}$ are fixed and set, for $j = 1, \dots, N$,

$$x_{n,j} = (2 \log_2 n)^{-1/2} \alpha_n(t_j), \tag{3.46}$$

and, for $|u| \leq 1$,

$$\begin{aligned} f_{n,j}(u) &= \left(2\left\{\frac{2}{n}\log_2 n\right\}^{1/2} \log_2 n\right)^{-1/2} \\ &\quad \times \left(\alpha(t_j) - \alpha_n\left(t_j - \left\{\frac{2}{n}\log_2 n\right\}^{1/2} u\right)\right). \end{aligned} \tag{3.47}$$

The following lemma relates these expressions to the Bahadur–Kiefer process R_n.

LEMMA 3.3. *We have almost surely, for each* $j = 1, \dots, N$,

$$\left|n^{1/4}(2\log_2 n)^{-3/4} R_n(t_j) - f_{n,j}(x_{n,j})\right| \to 0 \quad \textit{a.s. as } n \to \infty. \tag{3.48}$$

PROOF. See, e.g., (2.6) and Fact 2 of Deheuvels and Mason (1992b). □

In view of Lemma 3.3, we see that the limiting properties of $\{R_n(t_j),\ j = 1, \dots, N\}$ are governed by that of $(x_{n,1}, \dots, x_{n,N}) \times (f_{n,1}, \dots, f_{n,N})$. Our next lemma gives the appropriate description of the strong limiting behavior of this sequence.

Denote by $B(-1,1)$ the set of bounded functions on $[-1,1]$, endowed with the topology defined by the uniform metric, and by $\mathrm{AC}(-1,1)$ the set of all absolutely continuous functions on $[-1,1]$.

LEMMA 3.4. *The sequence* $\{(x_{n,1}, \dots, x_{n,N}) \times (f_{n,1}, \dots, f_{n,N}),\ n \geq 1\}$ *is almost surely relatively compact in* $\mathbb{R}^N \times B(-1,1)^N$, *with limit set equal to*

$$\begin{aligned} \mathbb{K}_N := \Bigg\{ &(x_1, \dots, x_N) \times (f_1, \dots, f_N) \in \mathbb{R}^N \times \mathrm{AC}(-1,1)^N: \\ &\sum_{j=1}^{N+1} \frac{(x_j - x_{j-1})^2}{t_j - t_{j-1}} + \sum_{j=1}^{N} \int_{-1}^{1} \dot{f}_j(u)^2\, du \leq 1, \textit{ with } x_0 = x_{N+1} = 0 \Bigg\}. \end{aligned} \tag{3.49}$$

PROOF. It follows readily from Theorem 1.2. □

We may now state our main result concerning the Bahadur–Kiefer representation.

THEOREM 3.3. *The sequence* $\{n^{1/4}(2\log_2 n)^{-3/4}(R_n(t_1),\dots,R_n(t_N)),\ n \geq 1\}$ *is almost surely relatively compact in* $\mathbb{R}^N$, *with limit set equal to*

$$\Big\{(f_1(x_1),\dots,f_N(x_N)):\ (x_1,\dots,x_N)\times(f_1,\dots,f_N)\in\mathbb{K}_N\Big\}. \tag{3.50}$$

PROOF. (3.50) is a direct consequence of Lemmas 3.3 and 3.4. □

REMARK 3.5. For $N = 1$, (3.49) becomes, with $x = x_1$, $t = t_1$ and $f = f_1$,

$$\mathbb{K}_1 = \left\{(x,f)\in\mathbb{R}\times B(-1,1):\ \frac{x^2}{t(1-t)} + \int_{-1}^{1}\dot{f}(u)^2\,du \leq 1\right\}. \tag{3.51}$$

It is readily verified in this case [see, e.g., Deheuvels and Mason (1992b)] that, given any x such that $|x| \leq \sqrt{t(1-t)}$, an arbitrary function f with $(x,f)\in\mathbb{K}_1$ satisfies

$$|f(u)| \leq \left\{|u|\left(1-\frac{x^2}{t(1-t)}\right)\right\}^{1/2}. \tag{3.52}$$

By (3.52), on $\mathbb{K}_1$, $|f(x)|$ is less than or equal to

$$\big(t(1-t)\big)^{1/4}\sup_{0\leq s\leq 1}\sqrt{s(1-s^2)} = \big(t(1-t)\big)^{1/4}2^{1/2}3^{-3/4}. \tag{3.53}$$

As shown by Deheuvels and Mason (1992b), there exists a function f for which this supremum is reached. Thus (3.53) provides a simple interpretation of how the constant in the right-hand-side of (3.45) is generated. Given Theorem 3.3, similar evaluations as given above can be achieved for an arbitrary $N \geq 1$ by routine analysis.

4. Conclusion. The preceding examples show the power of the methods based on the asymptotic theory of local empirical processes. A systematic study of the applications of our results will be developed in forthcoming publications.

Acknowledgment. We thank the referee for insightful comments.

REFERENCES

ALEXANDER, K. S. (1987). The central limit theorem for empirical processes on Vapnik–Červonenkis classes. *Ann. Probab.* **15** 178–203.

BAHADUR, R. R. (1966). A note on quantiles in large samples. *Ann. Math. Statist.* **37** 577–580.

CHUNG, K. L. (1949). An estimate concerning the Kolmogorov limit distributions. *Trans. Amer. Math. Soc.* **67** 36–50.

DEHEUVELS, P. (1974). Conditions nécessaires et suffisantes de convergence ponctuelle presque sûre et uniforme presque sûre des estimateurs de la densité. *C.R. Acad. Sci. Paris.* **278A** 1217–1220.

DEHEUVELS, P. (1992). Pointwise Bahadur–Kiefer theorems (I). In *Probability Theory and Applications—Essays to the Memory of József Mogyoródi* (J. Galambos and I. Kátai, eds.) 235–256. Kluwer, Dordrecht.

DEHEUVELS, P. and MASON, D. M. (1990a). Bahadur–Kiefer-type processes. *Ann. Probab.* **18** 669–697.

DEHEUVELS, P. and MASON, D. M. (1990b). Nonstandard functional laws of the iterated logarithm for tail empirical and quantile processes. *Ann. Probab.* **18** 1693–1722.

DEHEUVELS, P. and MASON, D. M. (1991). A tail empirical process approach to some nonstandard laws of the iterated logarithm. *J. Theoret. Probab.* **4** 53–85.

DEHEUVELS, P. and MASON, D. M. (1992a). Functional laws of the iterated logarithm for the increments of empirical and quantile processes. *Ann. Probab.* **20** 1248–1287.

DEHEUVELS, P. and MASON, D. M. (1992b). A functional L.I.L. approach to pointwise Bahadur–Kiefer theorems. In *Probability in Banach Spaces* (R. M. Dudley, M. Hahn and J. Kuelbs, eds.) **8** 255–266. Birkhäuser, Boston.

DEHEUVELS, P. and MASON, D. M. (1994). Nonstandard local empirical processes indexed by sets. *J. Statist. Plann. Inference.* To appear.

DEVROYE, L. (1979). On the pointwise and the integral convergence of recursive kernel estimates of probability densities. *Utilitas Math.* **15** 113–128.

DEVROYE, L. and GYÖRFI, L. (1985). *Nonparametric Density Estimation: The L_1 View.* Wiley, New York.

DONSKER, M. D. (1952). Justification and extension of Doob's heuristic approach to the Kolmogorov–Smirnov theorems. *Ann. Math. Statist.* **23** 277–281.

DUDLEY, R. M. (1966). Weak convergence of probabilities on nonseparable metric spaces and empirical measures on Euclidean spaces. *Illinois J. Math.* **10** 190–126.

DUDLEY, R. M. (1978). Central limit theorems for empirical measures. *Ann. Probab.* **6** 899–929. (Correction: *Ann. Probab.* (1979) **7** 909–911.)

EINMAHL, U. (1989). Extensions of results of Komlós, Major, and Tusnády to the multivariate case. *J. Multivariate Anal.* **28** 20–68.

FINKELSTEIN, H. (1971). The law of the iterated logarithm for empirical distributions. *Ann. Math. Statist.* **42** 607–615.

GAENSSLER, P. (1982). *Empirical Processes*. IMS, Hayward, CA.

GAENSSLER, P. (1984). Limit theorems for empirical processes indexed by classes of sets allowing a finite-dimensional parametrization. *Probab. Math. Statist.* **4** 1–12.

GINÉ, E. and ZINN, J. (1984). Some limit theorems for empirical processes. *Ann. Probab.* **12** 929–989.

GINÉ E. and ZINN. J. (1986). Lectures on the central limit theorem for empirical processes. *Probability in Banach Spaces. Lecture Notes in Math.* **1221** 50–113. Springer, Berlin.

HALL, P. (1981). Laws of the iterated logarithm for nonparametric density estimators. *Z. Wahrsch. Verw. Gebiete* **54** 47–61.

HOBSON, E. W. (1927). *The Theory of Functions of a Real Variable and the Theory of Fourier Series* **1**, 3rd ed. Cambridge Univ. Press.

KIEFER, J. (1967). On Bahadur's representation of sample quantiles. *Ann. Math. Statist.* **38** 1323–1342.

KIEFER, J. (1970). Deviations between the sample quantile process and the sample d.f. In *Nonparametric Techniques in Statistical Inference* (M. Puri, ed.) 299–319. Cambridge Univ. Press.

KUELBS, J. and DUDLEY, R. M. (1980). Log log laws for empirical measures. *Ann. Probab.* **8** 405–418.

LEDOUX, M. and TALAGRAND, M. (1990). Some applications of isoperimetric methods to strong limit theorems for sums of independent random variables. *Ann. Probab.* **18** 754–789.

MASON, D. M. (1988). A strong invariance theorem for the tail empirical process. *Ann. Inst. H. Poincaré Probab. Statist.* **24** 491–506.

PARZEN, E. (1962). On estimation of a probability density function and mode. *Ann. Math. Statist.* **33** 1065–1076.

ROSENBLATT, M. (1952). Remarks on a multivariate transformation. *Ann. Math. Statist.* **23** 470–472.

ROSENBLATT, M. (1956). Remarks on some non-parametric estimates of a density function. *Ann. Math. Statist.* **27** 832–837.

SHORACK, G. R. and WELLNER, L. A. (1986). *Empirical Processes with Applications to Statistics.* Wiley, New York.

TALAGRAND, M. (1988). Donsker classes of sets. *Probab. Theory Related Fields* **78** 169–191.

L.S.T.A.
UNIVERSITÉ PARIS VI
4 PLACE JUSSIEU
75252 PARIS CEDEX 05
FRANCE

DEPARTMENT OF MATHEMATICAL SCIENCES
501 EWING HALL
UNIVERSITY OF DELAWARE
NEWARK, DELAWARE 19716