Probab. Th. Rel. Fields 77, 271–305 (1988)

Probability Theory and Related Fields

The Central Limit Theorem and the Law of Iterated Logarithm for Empirical Processes under Local Conditions

Niels T. Andersen[1,*,**], Evarist Giné[2,**], Mina Ossiander[3,*], and Joel Zinn[2,**]

[1] Department of Mathematics, University of Aarhus, DK-8000, Aarhus, C, Denmark
[2] Department of Mathematics, Texas A & M University, College Station, TX 77843-3368, USA
[3] Department of Mathematics, University of Washington, Seattle, WA, 98195, USA

Summary. A CLT and a LIL are proved under weak-L_2 Gaussian bracketing conditions (weaker than the usual ones). These results have wide applicability and in particular provide an improvement of the Jain-Marcus central limit theorem for $C(S)$-valued random variables.

1. Introduction

Let $(S, \mathscr{S}, P)$ be a probability space and let $\mathscr{F}$ be a set of measurable functions on S with an envelope F finite everywhere. If X_i, $i \in \mathbb{N}$ are the coordinate functions of $(S^{\mathbb{N}}, \mathscr{S}^{\mathbb{N}}, P^{\mathbb{N}})$, the empirical measures P_n based on P (or on $\{X_i\}$) are defined as $P_n = n^{-1} \sum_{i=1}^{n} \delta_{X_i}$. We say that $\mathscr{F} \in \mathrm{CLT}(P)$ or that $\mathscr{F}$ is P-Donsker if $\{n^{1/2}(P_n - P)\}_{n=1}^{\infty}$ converges "weakly" in $\ell^{\infty}(\mathscr{F})$ (precise definitions are given below) and that $\mathscr{F} \in \mathrm{CLIL}(P)$ ($\mathrm{BLIL}(P)$) if $\{n^{1/2}(\ln \ln n)^{-1/2}(P_n - P)\}_{n=1}^{\infty}$ is a.s. a relatively compact (bounded) sequence in $\ell^{\infty}(\mathscr{F})$. Here CLT stands for the central limit theorem and CLIL (BLIL) for the compact (bounded) law of iterated logarithm. Since Dudley's (1978) influential article on the central limit theorem for empirical measures, a considerable amount of work has been devoted to the problem of determining the classes $\mathscr{F}$ which are in $\mathrm{CLT}(P)$. Giné and Zinn (1984) and Talagrand (1985) (see also Giné and Zinn 1986) gave necessary and sufficient conditions, modulo measurability, for $\mathscr{F} \in \mathrm{CLT}(P)$ in terms of the size of $\mathscr{F}$ measured in the $L_2(P)$ and $L_r(P_n)$ pseudo-distances, $r = 1, 2$. However, although these random conditions provide significant understanding of the problem and are of considerable practical value, it is obviously desirable to have sufficient conditions for $\mathscr{F} \in \mathrm{CLT}(P)$ that are not random, i.e., that do not depend on P^n for $n > 1$. Two interesting non-random sufficient conditions for $\mathscr{F} \in \mathrm{CLT}(P)$

* The research of these authors was carried out while visiting Texas A & M University

** The research of these authors has been totally or partially supported by the Danish Natural Science Research Council and by the National Science Foundation Grants numbers DMS-83-18610 and DMS-83-01367

already appear in Dudley (1978), who obtained sharp results for classes of sets: the Vapnik-Červonenkis property and the entropy with bracketing condition. See e.g., Alexander (1985) for the definitive result on the CLT for Vapnik-Červonenkis-graph classes of functions (see also Dudley 1984 and Pollard 1982). The CLT under metric entropy with bracketing has been considered by Dudley (1978 and e.g., 1984), Giné and Zinn (1984, 1986) and Ossiander (1985). The law of the iterated logarithm has received somewhat less attention in this context; see however Alexander (1984), Borisov (1985), Dudley and Kuelbs (1980), and Yukich (1986), among others. If the envelope function F of $\mathscr{F}$ satisfies $P^* F^2/L_2 F < \infty$ and $\mathscr{F} \in \mathrm{CLT}(P)$ then, at least under appropriate measurability, $\mathscr{F} \in \mathrm{CLIL}(P)$ (Goodman, Kuelbs and Zinn 1981, Heinkel 1979). But examples of Kuelbs (1976) in c_0 and a result of Ledoux (1982) on the LIL for processes with Lipschitzian trajectories show however that there are many interesting classes of functions $\mathscr{F}$ which are uniformly bounded or with $P^* F^2 < \infty$ satisfying the CLIL but not the CLT.

The object of this article is to present a CLT and an LIL for empirical processes under integrability conditions on their local modulus (with respect to a Gaussian distance) with a wide scope of applicability and which are sharp for some important classes of examples.

To motivate and explain our results we recall first Ossiander's CLT which is the sharpest bracketing CLT for unbounded classes in the literature: Let, for $\varepsilon > 0$, $N_{[\,]}^{(2)}(\varepsilon, \mathscr{F}, P) = \min\{n\colon \exists u_1, \ldots, u_n, \ell_1, \ldots, \ell_n \in \mathscr{L}_0(P)$ such that for each $f \in \mathscr{F}$ there are $i, j \leqq n$ such that $\ell_i \leqq f \leqq u_j$, $P(u_j - \ell_i)^2 \leqq \varepsilon^2\}$ be the $L_2(P)$-bracketing number of $\mathscr{F}$; if

$$\int_0 [\ln N_{[\,]}^{(2)}(\varepsilon, \mathscr{F}, P)]^{1/2}\, d\varepsilon < \infty \tag{1.1}$$

then $\mathscr{F} \in \mathrm{CLT}(P)$. Condition (1.1) implies that $P^* F^2 < \infty$ and, if G_P is the limiting Gaussian process, that its associated distance $d_G(f, g) := [E(G_P(f) - G_P(g))^2]^{1/2}$ satisfies the metric entropy condition, $\int_0 [\ln N(\varepsilon, \mathscr{F}, d_G)]^{1/2}\, d\varepsilon < \infty$, where $N(\varepsilon, \mathscr{F}, d_G) := \min\{n\colon \exists h_1, \ldots, h_n$ such that $\sup_{f \in \mathscr{F}} \min_{j \leqq d} d_G(f, h_i) \leqq \varepsilon\}$ is the covering number of $\mathscr{F}$ by d_G. However, neither $P^* F^2 < \infty$ is necessary for $\mathscr{F} \in \mathrm{CLT}(P)$ (although $t^2 P^*(F > t) \to 0$ as $t \to \infty$ is) nor is the entropy condition necessary for G_P to have a version with bounded d_G-uniformly continuous trajectories. A statistically significant example of a CLT for empirical processes holding, but with these two conditions failing, is the Chibisov-O'Reilly CLT for weighted empirical processes. In this paper we improve Ossiander's (1985) result by taking balls with respect to a Gaussian distance as "brackets" and by measuring them in the weak-L_2 pseudo-metric. So, we obtain a CLT and an LIL under what we could call a *Gaussian weak-L_2 bracketing condition* or a *weak-L_2 condition on the local modulus of $\{f(X)\}_{f \in \mathscr{F}}$ with respect to a Gaussian pseudo-metric.* The results obtained apply in several different situations, among others, the invariance principle, the Chibisov-O'Reilly theorem, the CLT and LIL for certain c_0-valued random variables, an improvement of Ledoux's LIL for Lipschitz processes and a significant improvement of the well known Jain and Marcus (1975) CLT. Actually, our main result for the i.i.d. case can be thought of as an improved Jain-Marcus CLT which also applies to not necessarily sample continuous processes: see Theorem 4.4″ and Corollary 4.5 below.

About the method of proof, first we note that if ρ is a Gaussian distance on $\mathscr{F}$ (i.e., ρ corresponds to the L_2-distance for a Radon Gaussian measure) or is dominated by a Gaussian distance, then a recent important theorem of Talagrand implies that ρ satisfies the (discrete) majorizing measure condition (see Sect. 2). This is an analytic condition that allows the use of "chaining" very much like metric entropy does. (We learned chaining under the majorizing measure condition from M. Talagrand.) Then we handle the "links" of the chain by truncation at the largest level possible that can be controlled using pregaussianness plus Bernstein's inequality and, as in Ossiander (1985), by use of the bracketing condition in the remainder, after summation by parts (see Sect. 3). The main differences from the method in Ossiander (1985) is that the chaining here must be done simultaneously for the L_2-distance and for the bracketing and, more significantly, that the non-uniformity of majorizing measures requires the use of a *new inequality*: an exponential bound for sums of positive random variables truncated away from 0 and from ∞ which may be of independent interest (Lemma 2.16 below). In the LIL we combine these techniques with techniques developed in the separable, measurable case by Kuelbs and Zinn (1979) and by Goodman, Kuelbs and Zinn (1981). A new important characterization of LIL in Banach spaces of Ledoux and Talagrand (to appear) will simplify our proof, however one may not be able to apply their result without further measurability requirements.

In Sect. 2 we present all the "lemmas", old and new, needed to prove our results; in particular, some elementary but useful results on majorizing measures, the theorems of Fernique and Talagrand on Gaussian processes, the usual eventual equicontinuity conditions for the CLT and the LIL in a version adapted to our needs, and the above-mentioned exponential bound for sums of non-negative random variables. In Sect. 3 we isolate the main argument of our proofs in the form of an inequality to be applied both in the CLT and in the LIL. In Sect. 4 we obtain the CLT, Gaussian convergence, both for i.i.d. and for non i.i.d. random variables. Donsker's invariance principle is an immediate consequence of the result for triangular arrays and the i.i.d. case is illustrated by the examples already mentioned. Finally, Sect. 5 contains the LIL with applications to some interesting c_0-valued random variables and to Ledoux's LIL in $C(S)$.

Techniques similar to those in this paper also apply to the CLT with limits other than Gaussian, like e.g., stable (slightly generalizing the CLT's of Marcus and Pisier 1984) and degenerate (Marcinkiewicz type weak and strong laws of large numbers): see Andersen et al. (in preparation).

Here is the notation not explained in the main text: for all $f, g \in \mathscr{F}$

$$Pf := \int f\,dP = Ef(X) \quad \text{if } \mathscr{L}(X) = P;$$

$$e_P(f,g) := [P(f-g)^2]^{1/2}, \qquad \rho_P(f,g) = e_P(f-Pf, g-Pg);$$

$$\Lambda^P_{2,\infty}(f) := [\sup_{t>0} t^2 P\{|f|>t\}]^{1/2};$$

$$d_G(f,g) := [E(G(f)-G(g))^2]^{1/2}$$

where G is a centered Gaussian process indexed by $\mathscr{F}$;

$\Pr^*$, P^*, P^*_{nj}, etc., denote outer probabilities and expectations;

$$\|h(f)\|_{\mathscr{F}} = \sup_{f \in \mathscr{F}} |h(f)| \qquad \text{for any function } h \text{ on } \mathscr{F}.$$

Regarding measurability, all of our random elements are defined on product spaces of the form $(\Omega, \Sigma, \Pr) = (S^{\mathbb{N}}, \mathscr{S}^{\mathbb{N}}, P^{\mathbb{N}}) \times ([0,1], \mathscr{B}, \lambda)$ where λ is Lebesgue measure on the Borel sets $\mathscr{B}$ of $[0,1]$, and depend on the coordinates of $S^{\mathbb{N}}$ sometimes multiplied by random variables defined on $([0,1], \mathscr{B}, \lambda)$, usually Rademacher $\left(\text{e.g., } \sum_{i=1}^{n} \varepsilon_i \delta_{X_i}\right)$. We recall that a Rademacher sequence $\{\varepsilon_i\}$ is a sequence of i.i.d. random variables such that $\Pr(\varepsilon_i = 1) = \Pr(\varepsilon_i = -1) = 1/2$. In particular if $F\colon S \to \mathbb{R}$ is not necessarily measurable and X is a coordinate function of $S^{\mathbb{N}}$, then it is "perfect" on $(\Omega, \Sigma, \Pr)$ and therefore $F^*(X) = F(X)^*$ a.s., where F^* is the P-outer envelope of F and $F(X)^*$ is the Pr-outer envelope of $F(X)$, hence we will have e.g., $\Pr^*\{F(X) > a\} = P\{F^*(s) > a\}$ (see e.g., Andersen [1985], Theorem 2.2 and Propositions 3.1 and 2.6). This applies also in the non-i.i.d. situation where $P^{\mathbb{N}}$ is replaced by a product of different probability measures. In fact the use of product space is not strictly necessary: perfect random elements suffice. We also use two exponential inequalities without further reference. One is Bernstein's inequality (e.g., Bennett 1962): if ξ_i are independent, centered and uniformly bounded by c, then

$$P\left\{\sum_{i=1}^{n} \xi_i > t\right\} \leqq \exp\left\{-t^2 \Big/ \left(2 \sum_{i=1}^{n} E\,\xi_i^2 + \tfrac{2}{3} t c\right)\right\}. \tag{1.2}$$

The other one is an inequality for binomial probabilities (Giné and Zinn 1984): if $PA_j = p_j$, $j = 1, \ldots, n$, and the A_j are independent sets then

$$P\left\{\sum_{j=1}^{n} I_{A_j} \geqq \ell\right\} \leqq \left(e \sum_{j=1}^{n} p_j / \ell\right)^{\ell} \tag{1.3}$$

for all n. (Actually, modulo constants, both inequalities are contained in Prohorov's inequality.)

2. Some Basic Preliminaries

(a) Gaussian Processes and Majorizing Measures

Let (T, d) be a pseudometric space. A Borel (sub-)probability measure μ on (T, d) is a *majorizing measure* for (T, d) (Fernique 1974) if

$$\sup_{t \in T} \int_0^{\infty} [\ln(\mu\{(B_d(t, \varepsilon)\})^{-1}]^{1/2}\, d\varepsilon < \infty \tag{2.1}$$

where $B_d(t,\varepsilon)=\{s\in T\colon d(s,t)\leqq\varepsilon\}$. A (sub-)probability measure μ is a *discrete majorizing measure* if there exists a (countable) set $S\subset T$, $S=\{\pi_q t\colon t\in T,\ q\in\mathbb{N}\}$, which supports μ and satisfies

$$\text{(2.2)}\qquad \begin{aligned}&\text{(i)}\ d(t,\pi_q t)\leqq 2^{-q}, \qquad t\in T,\ q\in\mathbb{N}\\ &\text{(ii)}\ \sup_{t\in T}\sum_{q=1}^{\infty}2^{-q}[\ln(\mu\{\pi_q t\})^{-1}]^{1/2}<\infty.\end{aligned}$$

We present some properties of majorizing measures before describing their relationship to Gaussian processes. Lemma 2.1 is embedded in Talagrand's results (1986); here is a direct proof.

2.1. Lemma. *If (T,d) has a majorizing measure, then it also has a discrete majorizing measure. If the majorizing measure ν satisfies*

$$\text{(2.3)}\qquad \lim_{\delta\to 0}\sup_{t\in T}\int_0^{\delta}\ln(\nu\{B_d(t,\varepsilon)\})^{-1}]^{1/2}\,d\varepsilon=0$$

then the discrete majorizing measure μ can be chosen to satisfy

$$\text{(2.4)}\qquad \lim_{k\to\infty}\sup_{t\in T}\sum_{q=k}^{\infty}2^{-q}[\ln(\mu\{\pi_q t\})^{-1}]^{1/2}=0.$$

Proof. Fix $q\in\mathbb{N}$. Let $\{t_i\}_{i=1}^r\subset T$ be a finite set such that $\sup_{t\in T}\inf_{i\leqq r}d(t,t_i)\leqq 2^{-q-3}$. (Note that the existence of a majorizing measure ν implies that (T,d) is totally bounded: otherwise there exist $\varepsilon>0$ and $\{s_i\}_{i=1}^{\infty}$ such that $\nu\{B_d(s_i,\varepsilon)\}\to 0$, thus contradicting (2.1).) We may assume that $\nu\{B(t_i,2^{-q-2})\}$ decreases in i by reordering if necessary. Define disjoint sets $\{T_i\}_{i=1}^r$ inductively as

$$T_1=B(t_1,2^{-q-2})$$

$$T_i=\begin{cases}\emptyset & \text{if } \nu\left\{B(t_i,2^{-q-2})\setminus\bigcup_{j=1}^{i-1}T_j\right\}<2^{-1}\nu\{B(t_i,2^{-q-2})\}\\[2ex] B(t_i,2^{-q-2})\setminus\bigcup_{j=1}^{i-1}T_j & \text{otherwise}\end{cases}$$

for $i\leqq r$. Define a subprobability measure μ_q with support $\{t_i\colon T_i\neq\emptyset\}$ by

$$\mu_q\{t_i\}=\nu(T_i\}\geqq 2^{-1}\nu\{B(t_i,2^{-q-2})\},\qquad T_i\neq\emptyset.$$

For each $t\in T$ define $\pi_q t$ in the following way: if $\{t_i\colon d(t,t_i)\leqq 2^{-q-3},\ T_i\neq\emptyset\}\neq\emptyset$ let $\pi_q t$ be one of these t_i. Then $d(t,\pi_q t)<2^{-q}$ and

$$\mu_q\{\pi_q t\}=\mu_q\{t_i\}\geqq 2^{-1}\nu\{B(t_i,2^{-q-2})\}\geqq 2^{-1}\nu\{B(t,2^{-q-3})\}.$$

Otherwise, if t_i is such that $d(t,t_i)\leqq 2^{-q-3}$ (hence $T_i=\emptyset$), there is $j<i$ such that $T_j\cap B(t_i,2^{-q-2})\neq\emptyset$; define $\pi_q t$ to be t_j for one of these j. Then

$$d(t,\pi_q t)\leqq d(t,t_i)+d(t_i,t_j)\leqq 2^{-q-3}+2\cdot 2^{-q-2}<2^{-q}$$

and

$$\mu_q\{\pi_q t\} = \mu_q\{t_j\} \geqq 2^{-1} \nu\{B(t_j, 2^{-q-2})\}$$
$$\geqq 2^{-1} \nu\{B(t_i, 2^{-q-2})\} \geqq 2^{-1} \nu\{B(t, 2^{-q-3})\}.$$

We claim that the measure $\mu = \sum_{q=1}^{\infty} 2^{-q} \mu_q$ is a discrete majorizing measure, and it satisfies (2.4) if ν satisfies (2.3). The claim follows from the fact that for all $t \in T$ and all $k \in \mathbb{N}$

$$\sum_{q=k}^{\infty} 2^{-q} [\ln(\mu\{\pi_q t\})^{-1}]^{1/2} \leqq \sum_{q=k}^{\infty} 2^{-q} [\ln(2^{-q} \mu_q\{\pi_q t\})^{-1}]^{1/2}$$
$$\leqq \sum_{q=k}^{\infty} 2^{-q} [\ln(2 \cdot \mu_q\{\pi_q t\})^{-1} + \ln 2^{q+1}]^{1/2}$$
$$\leqq \sum_{q=k}^{\infty} 2^{-q} [\ln(\nu\{B(t, 2^{-q-3})\})^{-1}]^{1/2} + \sum_{q=k}^{\infty} 2^{-q} [\ln 2^{q+1}]^{1/2}. \quad \square$$

2.2. Remarks. 1. It is obvious from the last part of the proof that there exists a discrete majorizing measure μ on $\{\pi_q t\colon t \in T, q \in \mathbb{N}\}$ if and only if there exist (sub-)probability measures μ_q on $\{\pi_q t\colon t \in T\}$, $q \in \mathbb{N}$, such that

$$\text{(2.2) (ii)}' \qquad \sup_{t \in T} \sum_{q=1}^{\infty} 2^{-q} [\ln(\mu_q\{\pi_q t\})^{-1}]^{1/2} < \infty$$

and that furthermore, there exists a discrete majorizing measure satisfying (2.4) if and only if there exist (sub-)probability measures μ_q on $\{\pi_q t\colon t \in T\}$, $q \in \mathbb{N}$, such that

$$\text{(2.4)}' \qquad \lim_{k \to \infty} \sup_{t \in T} \sum_{q=k}^{\infty} 2^{-q} [\ln(\mu_q\{\pi_q t\})^{-1}]^{1/2} = 0.$$

2. It is obvious that (2.2)(ii), (ii)′, (2.4) and (2.4)′ hold if and only if they hold with $(\mu_q\{\pi_q t\})^{-1}$ replaced by $2^q/\mu_q\{\pi_q t\}$ in the logarithmic part.

3. It may conceivably be useful to have $\{\pi_q t\}$ satisfying particular properties. It is not hard to prove (e.g., with a proof similar to that of Lemma 2.1) that if a discrete majorizing measure μ exists on (T, d) then there exists a possibly different majorizing measure $\bar{\mu}$ on (T, d) with support on a set $S = \{\bar{\pi}_q t\} \subseteq T$ such that

(i) $d(t, \bar{\pi}_q t) \leqq 2^{-q}$, $\quad t \in T$, $q \in \mathbb{N}$,

(ii) $d(\bar{\pi}_q t, \bar{\pi}_q s) > 2^{-q}$ $\quad$ if $\bar{\pi}_q t \neq \bar{\pi}_q s$, $t, s \in T$, $q \in \mathbb{N}$,

(iii) $\bar{\pi}_q$ is the identity on $\bar{\pi}_{q-1}(T)$, $q \in \mathbb{N}$.

One can also choose $\{\bar{\pi}_q\}$ to satisfy

(ii)′ $\bar{\pi}_q \circ \bar{\pi}_r = \bar{\pi}_q$, $1 \leqq q \leqq r$, $r \in \mathbb{N}$

instead of (ii). (Andersen 1986.)

4. It is well known (Fernique 1974; Preston 1972) that if (T, d) satisfies the metric entropy condition $\int_0 [\ln N(\varepsilon, T, d)]^{1/2} d\varepsilon < \infty$ (see e.g., Dudley [1978] for the definition) then (T, d) admits a discrete majorizing measure which verifies (2.4), but the converse does not hold.

5. In Lemma 2.1 we use the function $[\ln x^{-1}]^{1/2}$, but the statement holds for all decreasing functions ϕ: $\mathbb{R}_+ \to \mathbb{R}_+$ satisfying $\int_0 \phi(x)\,dx < \infty$ and $\phi(x+y) \leqq \phi(x) + \phi(y)$ for all $x, y \in \mathbb{R}_+$.

The following is a simple but important observation of Talagrand (private communication) which allows the use of discrete majorizing measure in "chaining" arguments as efficiently as metric entropy.

2.3. Lemma. (A) *For every $q \in \mathbb{N}$ let $S_q \subset T$ be a countable (or finite) subset of T and let μ_q be a (sub-)probability measure on S_q. Define on S_q*

$$\gamma_q(s) = [\ln(2^q/\mu_q\{s\})]^{1/2}, \qquad s \in S_q.$$

Then for $r > 1$

$$\sum_{s \in S_q} \exp(-r\gamma_q(s)^2) \leqq 2^{-qr}.$$

(B) *The following are equivalent for (T, d):*

(i) *there exists a discrete majorizing measure μ on (T, d);*

(ii) *there exist $S = \{\pi_q t\colon q \in \mathbb{N},\ t \in T\}$ satisfying* (2.2)(i) *and strictly positive functions γ_q on $S_q = \{\pi_q t\colon t \in T\}$ such that*

(a) $\sup_{t \in T} \sum_{q=1}^{\infty} 2^{-q} \gamma_q(\pi_q t) < \infty,$

(b) $\lim_{r \to \infty} \sum_{q=1}^{\infty} \sum_{s \in S_q} \exp(-r\gamma_q(s)^2) = 0.$

Moreover, condition (2.4) *for μ is equivalent to*

(a′) $\lim_{k \to \infty} \sup_{t \in T} \sum_{q=k}^{\infty} 2^{-q} \gamma_q(\pi_q t) = 0.$

Proof. (A) Just note that since μ_q is a sub-probability

$$\sum_{s \in S_q} \exp\{-r\gamma_q^2(s)\} = \sum_{s \in S_q} \left(\frac{\mu_q\{s\}}{2^q}\right)^r \leqq 2^{-qr}.$$

(B) (ii)⇒(i). Take r such that $\sum_{s \in S_q} \exp(-r\gamma_q(s)^2) \leqq 1$ for all $q \in \mathbb{N}$. Then $\mu_q\{s\} := \exp(-r\gamma_q(s)^2)$, $s \in S_q$, $q \in \mathbb{N}$, defines a family of subprobability measures satisfying (2.2)(ii)′ ((2.4)′ if (a′) holds).

(i)⇒(ii) Take $\gamma_q(s) = [\ln(2^{-q}\mu\{s\})^{-1}]^{1/2}$, $s \in S_q$. Then (ii)(a) obviously holds and (ii)(b) follows from (A). □

Actually our γ_q's will depend on $\pi_r t$, $r \leqq q$, but the same argument applies to give:

2.4. Lemma. *Let μ be a discrete majorizing measure for (T, d) with support $\{\pi_q t\colon q\in\mathbb{N}, t\in T\}$. Let*

$$\bar{\gamma}_q(t)=\left[\ln\left(2^q\Big/\prod_{r=1}^{q}\mu\{\pi_q t\}\right)\right]^{1/2}.$$

Then

$$\text{(i)}\qquad \sup_{t\in T}\sum_{q=1}^{\infty}2^{-q}\bar{\gamma}_q(t)<\infty$$

and

$$\text{(ii)}\qquad \lim_{r\to\infty}\sum e^{-r\bar{\gamma}_q(t)^2}=0$$

where the sum extends over all the distinct $(q+1)$-tuples

$$(q, \pi_1 t, \ldots, \pi_q t)\qquad t\in T,\ q\geqq 1.$$

If moreover μ satisfies (2.4) *then also*

$$\text{(i)}'\qquad \lim_{k\to\infty}\sup_{t\in T}\sum_{q=k}^{\infty}2^{-q}\bar{\gamma}_q(t)=0.$$

Proof. Let $\gamma_q(s)$, $q\in\mathbb{N}$, $s\in S_q$ be as in Lemma 2.3. Since $\mu\times\overset{q)}{\ldots}\times\mu$ is a subprobability measure on $T\times\overset{q)}{\ldots}\times T$, the counting argument of Lemma 2.3 applies to give (ii). Proving (i) from (2.2)(ii) is similar (easier) to prove than the limiting version of (i) from (2.4), so we check only this last implication. Assume (2.4) holds. Then

$$b_k:=\sup_{t\in T}\sum_{q=k}^{\infty}2^{-q}\gamma_q(\pi_q t)\to 0\qquad\text{as } k\to\infty$$

and

$$c_r:=\sup_{t\in T}\gamma_r(\pi_r t)\leqq b_r 2^r<\infty.$$

Since

$$\bar{\gamma}_q(t)\leqq\sum_{r=1}^{q}\gamma_r(\pi_r t),$$

we have, for any $r_0<k$ and any $t\in T$

$$\begin{aligned}\sum_{q=k}^{\infty}2^{-q}\bar{\gamma}_q(t)&\leqq\sum_{r=1}^{\infty}\gamma_r(\pi_r t)\sum_{q=r\vee k}^{\infty}2^{-q}\\&\leqq 2\cdot\sum_{r=k}^{\infty}2^{-r}\gamma_r(\pi_r t)+2^{-k+1}\sum_{r=1}^{r_0}\gamma_r(\pi_r t)+2\cdot\sum_{r=r_0+1}^{k-1}2^{-r}\gamma_r(\pi_r t)\\&\leqq 2b_k+2^{-k+1}r_0\max_{r\leqq r_0}c_r+2b_{r_0+1}.\end{aligned}$$

So, taking limits first as $k\to\infty$ and then as $r_0\to\infty$, we obtain (i)′. □

As mentioned in the introduction sample continuity and sample boundedness of Gaussian processes are characterized in terms of majorizing measures. The sufficiency part of the following theorem is due to Fernique (1974) (see also the earlier work of Preston 1972). Necessity was proved by Fernique (1974) for stationary Gaussian processes, and in general by Talagrand (1986) (actually, Fernique proves that the entropy condition is necessary in the stationary case).

2.5. Theorem. *Let $G=\{G(t)\}_{t\in T}$ be a centered Gaussian process on a set T. Assume G is separable in (T, d_G). We then have:*

(i) If *(T, d_G) admits a majorizing measure, then G has bounded sample paths a.s.; if moreover the majorizing measure satisfies* (2.3), *then G has bounded and d_G-uniformly continuous sample paths a.s.*

(ii) *If G has bounded sample paths a.s. then (T, d_G) admits a discrete majorizing measure. If G has bounded and d_G-uniformly continuous sample paths a.s. then the discrete majorizing measure can be chosen to satisfy* (2.4).

2.6. Remark. As a first example of chaining with a discrete majorizing measure, illustrative of what follows, here is a proof of the boundedness part of statement (i) in Theorem 2.5, along Talagrand's ideas. If a majorizing measure for (T, d_G) exists, then there exists a *discrete* one by Lemma 2.1. Let $T_q=\{(\pi_q t, \pi_{q-1} t): t\in T\}$ and for $s\in T_q$ let s_q and s_{q-1} denote its components. Note that $(E(G(s_q)-G(s_{q-1}))^2)^{1/2}\leqq 3\cdot 2^{-q}$. Note also that if $\tilde{\gamma}_q(t)=\left(\ln \frac{2^q}{\mu\{\pi_q t\}\,\mu\{\pi_{q-1} t\}}\right)^{1/2}$, which depends only on $s=(\pi_q t, \pi_{q-1} t)$, then $\sum_{q=q_0+1}^{\infty} \sum_{s\in T_q} \exp(-r\tilde{\gamma}_q(s)^2)\leqq \sum_{q=q_0+1}^{\infty} 2^{-qr}$, as in Lemma 2.3(A). Finally set $\beta_{q_0}=\sup_{t\in T} \sum_{q=q_0+1}^{\infty} 2^{-q}\tilde{\gamma}_q(s)$, which is finite by hypothesis. We then have, for any $q_0\in\mathbb{N}$ and $u>0$,

$$P\{\sup_{t\in T}|G(t)-G(\pi_{q_0} t)|>u\}=P\left\{\sup_{t\in T}\left|\sum_{q=q_0+1}^{\infty} (G(\pi_q t)-G(\pi_{q-1} t))\right|>u\right\}$$

$$\leqq \sum_{q=q_0+1}^{\infty} P\left\{\sup_{s\in T_q} \frac{|G(s_q)-G(s_{q-1})|}{2^{-q}\tilde{\gamma}_q(s)}>\frac{u}{\beta_{q_0}}\right\}$$

$$\leqq \sum_{q=q_0+1}^{\infty} \sum_{s\in T_q} \exp\left\{-\frac{u^2\, 2^{-2q}\tilde{\gamma}_q(s)^2/\beta_{q_0}^2}{18\cdot 2^{-2q}}\right\}\leqq \sum_{q=q_0+1}^{\infty} 2^{-qu^2/18\beta_{q_0}^2}.$$

Now, sample boundedness of G follows from the fact that $\#\{\pi_{q_0} t: t\in T\}<\infty$.

The above results will be useful mainly for the CLT. Now we give a lemma on majorizing measures specific to the LIL. We write $Lx:=1\vee \ln x$ and $L_k x := L \overset{k)}{\dots} Lx$, $k\in\mathbb{N}$.

2.7. Lemma. *Let (T, d) be a pseudometric space. Then the following are equivalent:*

(i) *There exists a (sub-)probability measure μ on (T, d) such that*

$$\sup_{t\in T} \int_0^{\infty} [(\ln(\mu\{B_d(t,\varepsilon)\})^{-1})/(L_2 \ln(\mu\{B_d(t,\varepsilon)\})^{-1})]^{1/2}\, d\varepsilon<\infty$$

$$\left(\text{and } \lim_{\delta\to 0}\sup_{t\in T} \int_0^{\delta} [(\ln(\mu\{B_d(t,\varepsilon)\})^{-1})/(L_2 \ln(\mu\{B_d(t,\varepsilon)\})^{-1})]^{1/2}\, d\varepsilon=0\right).$$

(ii) *There exists a countable subset* $S=\{\pi_q t: t\in T, q\in\mathbb{N}\}\subset T$ *and a (sub-)probability measure* μ *on* S *such that*

(a) $d(t,\pi_q t)\leqq 2^{-q}$, $t\in T$, $q\in\mathbb{N}$,

(b) $\sup_{t\in T}\sum_{q=1}^{\infty} 2^{-q}(Lq)^{-1/2}[\ln(\mu\{\pi_q t\})^{-1}]^{1/2}<\infty$

$\left(\textit{and } \lim_{k\to\infty}\sup_{t\in T}\sum_{q=k}^{\infty} 2^{-q}(Lq)^{-1/2}[\ln(\mu\{\pi_q t\})^{-1}]^{1/2}=0\right)$.

(iii) *There exist* $\{\pi_q t\}$ *satisfying* (ii a) *such that the Gaussian process*

$$Z(t)=\sum_{q=1}^{\infty} 2^{-q}(Lq)^{-1/2} g_{q,\pi_q t}, \qquad t\in T,$$

where $\{g_{q,t}: q\in\mathbb{N}, t\in S\}$ *are i.i.d.* $N(0,1)$, *has a version with bounded (and* d_Z*-uniformly continuous) sample paths.*

(iv) *There exists a centered Gaussian process* G *on* T *with bounded sample paths such that*

$$d\leqq d_G[L_2 d_G^{-1}]^{1/2}$$

(G *has also* d_G*-uniformly continuous sample paths).*

Proof. (i)$\Rightarrow$(ii). The proof of Lemma 2.1 will give us (ii) with $L_2\ln\mu(\{\pi_q t\})^{-1}$ instead of Lq in (b) (see Remark 2.2(5)). Now using that $2^{-2q}(L_2\ln(\mu\{\pi_q t\})^{-1})^{-1}(\ln(\mu\{\pi_q t\})^{-1})$ is uniformly bounded in q and t one can show that there exists $c<\infty$, not depending on t or q, such that $L_2\ln_q(\mu\{\pi_q t\})^{-1}\leqq c\cdot Lq$. Hence we get (b) as stated.

(ii)$\Rightarrow$(iii). We give only the proof of boundedness since the proof of continuity rests on the same principles (see e.g., Giné and Zinn [1986] Theorem 3.1.2): Let $\gamma_q(t)=[\ln(2^q/\mu\{\pi_q t\})]^{1/2}$ and

$$\beta=\sup_{t\in T}\sum_{q=1}^{\infty} 2^{-q}(Lq)^{-1/2}\gamma_q(t)<\infty.$$

Then

$$\begin{aligned}
&\Pr\{\sup_{t\in T} Z(t)>M\}\\
&\leqq \Pr\left\{\sup_{t\in T}\left|\sum_{q=1}^{\infty} 2^{-q}(Lq)^{-1/2} g_{q,\pi_q t}\right|\left(\sum_{q=1}^{\infty} 2^{-q}(Lq)^{-1/2}\gamma_q(t)\right)^{-1}>M\beta^{-1}\right\}\\
&\leqq \sum_{q=1}^{\infty}\Pr\{\sup_{t\in T}|g_{q,\pi_q t}|\,\gamma_q(t)^{-1}>M\beta^{-1}\}\\
&\leqq \sum_{q=1}^{\infty}\sum_{s\in S_q}\exp(-2^{-1}M^2\beta^{-2}\gamma_q(s)^2)\to 0 \qquad \text{as } M\to\infty
\end{aligned}$$

by Lemma 2.3 (A).

(iii)$\Rightarrow$(iv). Let Z be the Gaussian process in (iii). If $2^{-q-1}(L(q+1))^{-1/2}\leqq d_Z(s,t)<2^{-q}(Lq)^{-1/2}$ then $\pi_r(s)=\pi_r(t)$ for all $r\leqq q$. Hence

$$d(s,t) \leqq 2^{-q+1} \leqq 4 \cdot d_Z(s,t) \cdot (L(q+1))^{1/2}$$
$$\leqq 8(Lq)^{1/2} d_Z(s,t)$$
$$\leqq 8(L_2(d_Z(s,t)^{-1}))^{1/2} d_Z(s,t).$$

Therefore a suitable multiple G of Z satisfies $d \leqq d_G(L_2 d_G^{-1})^{1/2}$.

(iv)$\Rightarrow$(i). Let $\phi(t)=t(L_2 t^{-1})^{1/2}$, and let G be as in (iv). Put $B_\rho(t,\varepsilon)=\{s\in T: \rho(s,t)\leqq\varepsilon\}$ for $\rho=d$, d_G and $\varepsilon>0$. Then $B_d(t,\varepsilon)\supset B_{d_G}(t\,\phi^{-1}(\varepsilon))$ for all $\varepsilon>0$ and all $t\in T$. Hence if μ is a majorizing measure for G, which exists by Talagrand's Theorem 2.5 (ii), and if we let

$$\tau_\rho(\varepsilon,t):=\tau_\rho(\varepsilon):=\ln(\mu\{B_\rho(t,\varepsilon)\})^{-1} \qquad \text{for } \rho=d, d_G,$$

and $a=\operatorname{diam}(T, d_G)$, then

$$\int_0^\infty [\tau_d(\varepsilon)/L_2\tau_d(\varepsilon)]^{1/2}\,d\varepsilon \leqq \int_0^\infty [\tau_{d_G}(\phi^{-1}\varepsilon)/L_2\tau_{d_G}(\phi^{-1}\varepsilon)]^{1/2}\,d\varepsilon$$
$$\leqq \int_0^a [\tau_{d_G}(u)/L_2\tau_{d_G}(u)]^{1/2}\cdot(L_2 u^{-1})^{1/2}\,du$$
$$\leqq \int_0^a (\tau_{d_G}(u))^{1/2} I_{[\tau_{d_G}(u)>u^{-1}]}\,du + \int_0^a u^{-1/2}\,du$$
$$\leqq \int_0^a (\tau_{d_G}(u))^{1/2}\,du + 2a^{1/2}.$$

Hence, since $\sup_{t\in T}\int_0^\infty (\tau_{d_G}(u,t))^{1/2}\,du<\infty$ if G has bounded sample paths, we obtain (i) in this case (if moreover the sample paths of G are d_G uniformly continuous, the same computations together with $\lim_{\delta\to 0}\sup_{t\in T}\int_0^\delta (\tau_{d_G}(u,t))^{1/2}\,du=0$, also give the parenthetical statement in (i)). $\square$

(b) The Central Limit Theorem

Let $(S,\mathscr{S})$ be a measure space. Let $\{P_{nj}: j=1,\ldots,n,\, n\in\mathbb{N}\}$ be probability measures on $(S,\mathscr{S})$ and let $\mathscr{F}\subset\bigcap_{n,j}\mathscr{L}_1(S,\mathscr{S},P_{nj})$ such that

$$\sup_{f\in\mathscr{F}}|f(s)|<\infty \qquad \text{for all } s\in S. \tag{2.5}$$

Let also $(\Omega_n,\Sigma_n,\Pr_n)=(S^n,\mathscr{S}^n,P_{n1}\otimes\ldots\otimes P_{nn})\times([0,1],\mathscr{B},\lambda)$, let $X_{nj}:\Omega_n\to S$ be the coordinate projections and let $\{a_n\}$ be a sequence of real positive numbers. Then:

2.8. Definition. $\mathscr{F}$ satisfies the *CLT with centering at expectations* with respect to $\{P_{nj}\}$ and $\{a_n\}$ – $\mathscr{F} \in \mathrm{CLT}\{P_{nj}; a_n\}$ for short – if there exists a (centered) Radon measure γ on $\ell^\infty(\mathscr{F})$ such that for all $H: \ell^\infty(\mathscr{F}) \to \mathbb{R}$ bounded and continuous,

$$E^* H\left(a_n^{-1} \sum_{j=1}^{n} (\delta_{X_{nj}} - P_{nj})\right) \underset{n\to\infty}{\to} \int H\, d\gamma.$$

(With some extra care we could define CLT with the usual centering at expectations of truncations, but this definition is enough for our purposes in this paper.)

The following theorem combines finite dimensional approximations and randomization. (We only state the "sufficiency" half of it, but a "necessity" part, in similar terms, can also be proved.) Finite dimensional approximation in the CLT in infinite dimensions goes back, in a sense, to Prohorov, and more specifically, to de Acosta, Pisier and Dudley and Phillip; Jain and Marcus introduced randomization in the present setting, and Andersen and Dobrić showed that distances other than L_2 may be considered.

2.9. Theorem. *Let* $\mathscr{F} \subset \bigcap_{n,j} \mathscr{L}_1(S, \mathscr{S}, P_{nj})$ *and satisfy* (2.5). *Assume that for all* $(f_1, \ldots, f_k) \in \mathscr{F}$, $k \in \mathbb{N}$, *the finite dimensional distributions* $\left\{\mathscr{L}\left[a_n^{-1} \sum_{j=1}^{n} (f_i(X_{nj}) - P_{nj} f_i)_{i=1}^{k}\right]\right\}_{n=1}^{\infty}$ *converge weakly in* $\mathbb{R}^k$. *Assume further that there are maps* π_q: $\mathscr{F} \to \mathscr{F}$, $g \in \mathbb{N}$, *such that*

(i) $\#\{\pi_q f: f \in \mathscr{F}\} < \infty$ *for all* $q \in \mathbb{N}$,

and for all $\varepsilon > 0$, *with* $\{\varepsilon_i\}$ *a Rademacher sequence defined on* $([0,1], \mathscr{B}, \lambda)$,

(ii) $\lim_{q\to\infty} \limsup_n \Pr_n^* \left\{\left\|a_n^{-1} \sum_{j=1}^{n} \varepsilon_j (f - \pi_q f)(X_{nj})\right\|_{\mathscr{F}} > \varepsilon\right\} = 0$

and

(iii) $\lim_{q\to\infty} \limsup_n \sup_{f\in\mathscr{F}} \Pr_n^* \left\{\left|a_n^{-1} \sum_{j=1}^{n} [(f - \pi_q f)(X_{nj}) - P_{nj}(f - \pi_q f)]\right| > \varepsilon\right\} = 0.$

Then $\mathscr{F} \in \mathrm{CLT}\{P_{nj}, a_n\}$.

Proof (Sketch). By (iii) and the symmetrization argument in Lemma 2.7(b) of Giné and Zinn (1984) (using probabilities instead of second moments) it follows from (ii) that

(ii)′ $\lim_{q\to\infty} \limsup_n Pr_n^* \left\{\left\|a_n^{-1} \sum_{j=1}^{n} [(f - \pi_q f)(X_{nj}) - P_{nj}(f - \pi_q f)]\right\|_{\mathscr{F}} > \varepsilon\right\} = 0.$

Now, this implies an "asymptotic equicontinuity condition" for the pseudodistance $\rho(f, g) = 2^{-\min\{q: \pi_q f \neq \pi_q g\}}$ and obviously $(\mathscr{F}, \rho)$ is totally bounded. From this point on the proof is completely analogous to that of Theorem 1.1.3 in Giné and Zinn (1986), (ii)⇒(i). □

A better statement of Theorem 2.9 has (ii)′ as a hypothesis instead of (ii) and (iii) (note that (iii) is only used as a centering condition). But the random elements $\varepsilon_i f(X_i)$ are easier to handle than $f(X_i) - Pf$, particularly if truncation and recentering is required.

Next we recall some known facts and notation about the "$\sqrt{n}$ – i.i.d." case, that will be used without further mention. If $P_{nj}=P$ and $a_n=n^{1/2}$, then the limit γ is Gaussian and we write $\mathscr{F}\in\mathrm{CLT}(P)$ instead of $\mathscr{F}\in\mathrm{CLT}(\{P_{nj};a_n\})$. Definition 2.8 in this case is Hoffmann-Jørgensen's and is equivalent to Dudley's definition of Donsker classes. In the case $P_{nj}=P$, $a_n=n^{1/2}$, the finite dimensional distributions always converge to those of G_P, the centered Gaussian process with difference variance $\rho_P^2(f,g)=P(f-Pf-g+Pg)^2$. If $\mathscr{F}\in\mathrm{CLT}(P)$, then the limit γ is the law of G_P, and since it is Radon, G_P, in particular, admits a version with bounded and ρ_P-uniformly continuous sample paths (Andersen and Dobrić 1987).

2.10. Definition. $\mathscr{F}\subset\mathscr{L}_2(P)$ and satisfying (2.5) is P-pregaussian if G_P admits a version with bounded ρ_P-uniformly continuous sample paths.

As a corollary to the definition, a necessary condition for $\mathscr{F}\in\mathrm{CLT}(P)$ is that $\mathscr{F}$ be P-pregaussian. From Talagrand's theorem, we obtain the first part of the following corollary. The second part is trivial and therefore its proof is omitted.

2.11. Corollary. *A class of functions $\mathscr{F}\subset\mathscr{L}_2(P)$ is P-pregaussian if and only if $(\mathscr{F},\rho_P)$ admits a (discrete) majorizing measure. If $\|Pf\|_{\mathscr{F}}<\infty$, ρ_P can be replaced by e_P in this statement, where $e_P(f,g)=(P(f-g)^2)^{1/2}$.*

Let $F(s)=\sup_{f\in\mathscr{F}}|f(s)|$ be the envelope function of $\mathscr{F}$ and let $\bar{F}(s)=\sup_{f\in\mathscr{F}}|f(s)-Pf|$. Then (see e.g., Giné and Zinn [1986], Proposition 1.2.7) a necessary condition for $\mathscr{F}\in\mathrm{CLT}(P)$ is that $\lim_{t\to\infty}t^2P^*(\bar{F}>t)=0$. If $\|Pf\|_{\mathscr{F}}<\infty$, then $\lim_{t\to\infty}t^2P^*(F>t)=0$ is necessary for $\mathscr{F}\in\mathrm{CLT}(P)$.

Since e_P and F are more convenient than ρ_P and $\bar{F}$, we will assume in what follows that $\|Pf\|_{\mathscr{F}}<\infty$, in the *i.i.d.* case.

(c) The Law of the Iterated Logarithm

Recall that $Lx=1\vee\ln x$, $x\geqq 0$, and $L_2x=L(Lx)$. Let for $n\in\mathbb{N}$

(2.6) $$a_n=(nL_2n)^{1/2}.$$

Let $(S,\mathscr{S},P)$ be a probability space and let $\mathscr{F}\subseteq\mathscr{L}_2(P)$ satisfy (2.5). The following definitions are analogous to the definitions in separable Banach spaces (see e.g., Goodman, Kuelbs, and Zinn [1981]).

2.12. Definition. Let X_i: $(\Omega,\Sigma,\Pr)\to S$ be the coordinate projections from $(\Omega,\Sigma,\Pr)=(S^{\mathbb{N}},\mathscr{S}^{\mathbb{N}},R^{\mathbb{N}})\times([0,1],\mathscr{B},\lambda)$ onto S. Then we say that $\mathscr{F}$ *satisfies the compact law of the iterated logarithm for* P – $\mathscr{F}\in\mathrm{CLIL}(P)$ – if $\left\{a_n^{-1}\sum_{i=1}^n(f(X_i)-Pf)\right\}_{n=1}^\infty$ is relatively compact in $\mathscr{L}^\infty(\mathscr{F})$ a.s. We say that $\mathscr{F}$ *satisfies the bounded law of the iterated logarithm for* P – $\mathscr{F}\in\mathrm{BLIL}(P)$ – if $\limsup_n a_n^{-1}\left\|\sum_{i=1}^n(f(X_i)-Pf)\right\|_{\mathscr{F}}<\infty$ a.s.

Next we give a sufficient condition for $\mathscr{F} \in \mathrm{LIL}$ which is based on ideas of Pisier (1975) (we only state a sufficiency result since this is all we need). As usual $\{\varepsilon_i\}$ is a Rademacher sequence on the $([0,1], \mathscr{B}, \lambda)$ part of $(\Omega, \Sigma, \mathrm{Pr})$.

2.13. Theorem. *Let $\mathscr{F} \in \mathscr{L}_2(P)$ and satisfy* (2.5). *Assume that for every $q \in \mathbb{N}$ there exists $\pi_q\colon \mathscr{F} \to \mathscr{F}$ such that $\#\{\pi_q f\colon f \in \mathscr{F}\} < \infty$,*

$$\lim_{q\to\infty} \|P(f-\pi_q f)^2\|_{\mathscr{F}} = 0 \tag{2.7}$$

and such that for all $\alpha > 0$ there exists $q_0 \geqq 1$ such that for all $q \geqq q_0$

$$\lim_{n_0\to\infty} \mathrm{Pr}^*\left\{\sup_{n\geqq n_0}\left\|\sum_{i=1}^{n} \varepsilon_i (f-\pi_q f)(X_i)\right\|_{\mathscr{F}} \Big/ a_n > \alpha\right\} = 0. \tag{2.8}$$

Then $\mathscr{F} \in \mathrm{CLIL}$. (And if instead of (2.8) *we only have that for some $\alpha > 0$*

$$\lim_{n_0\to\infty} \mathrm{Pr}^*\left\{\sup_{n\geqq n_0}\left\|\sum_{i=1}^{n} \varepsilon_i f(X_i)\right\|_{\mathscr{F}} \Big/ a_n > \alpha\right\} = 0, \tag{2.8$'$}$$

then $\mathscr{F} \in \mathrm{BLIL}$.)

Proof. If (2.7) and (2.8) hold, since

$$\sup_{n\geqq n_0}\sup_{f\in\mathscr{F}} a_n^{-2} E\left(\sum_{i=1}^{n} [(f-\pi_q f)(X_i) - P(f-\pi_q f)]^2\right)$$
$$\leqq (L_2 n_0)^{-1} \sup_{f\in\mathscr{F}} P(f-\pi_q f)^2 \to 0 \quad \text{as } n_0 \to \infty,$$

it follows by standard symmetrization techniques (e.g., Corollary 2.6 and a simple modification of the proof of Lemma 2.7(b), both in Giné and Zinn [1984]) that for all $\alpha > 0$ there exists $q_0 \geqq 1$ such that for all $q \geqq q_0$,

$$\lim_{n_0\to\infty} \mathrm{Pr}^*\left\{\sup_{n\geqq n_0}\left\|\sum_{i=1}^{n} [(f-\pi_q f)(X_i) - P(f-\pi_q f)]\right\|_{\mathscr{F}} > \alpha a_n\right\} = 0.$$

Hence

$$\limsup_{n} a_n^{-1}\left\|\sum_{i=1}^{n} [(f-\pi_q f)(X_i) - P(f-\pi_q f)]\right\|_{\mathscr{F}} \leqq \alpha \text{ a.s.}$$

From this it follows by the usual finite dimensional approximation (see e.g., Pisier [1975], Theorem 3.1) that $\mathscr{F} \in \mathrm{CLIL}(P)$. The proof for the BLIL is trivial. □

A way to check (2.8) is via the following.

2.14. Proposition. *Let $I_k = \{n\colon 2^k < n \leqq 2^{k+1}\}$, $k \in \mathbb{N}$. Let $\mathscr{H} \subset \mathscr{L}_2(P)$. If for some $\varepsilon > 0$*

$$\sum_{k=0}^{\infty} \mathrm{Pr}^*\{\|\sum_{i\in I_k} \varepsilon_i h(X_i)\|_{\mathscr{H}} > \varepsilon a_{2^k}\} < \infty \tag{2.9}$$

then

$$\lim_{n_0\to\infty} \Pr{}^*\left\{\sup_{n\ge n_0}\left\|\sum_{i=1}^{n}\varepsilon_i h(X_i)\right\|_{\mathscr{H}}\Big/ a_n > 8\varepsilon\right\}=0.$$

Proof. Follows as in Stout (1974) pp. 159–160 by keeping track of constants. Note that Levy's inequality holds with Pr* (see e.g., Dudley and Phillip 1983). □

We also require a way to evaluate (2.9) due to Kuelbs and Zinn (1979, Theorem 1). For the reader's convenience we will state the result the way we need it.

2.15. Proposition. *Let $\{Z_j\}_{j\in\mathbb{N}}$ be independent, symmetric B-valued random variables, $(B, \|\cdot\|)$ a separable Banach space. The conditions*

(2.10) $$j^{-1}\|Z_j\|\to 0 \quad a.s.,$$

(2.11) $$\sum_k (\Lambda(k))^2<\infty, \quad where\ \Lambda(k)=\sum_{j\in I_k} 4^{-k}E\|Z_j\|^2,$$

(2.12) $$\sum_{j=1}^{k} Z_j/k\to 0 \quad in\ probability,$$

imply that

(2.13) $$\sum_{k=0}^{\infty}\Pr\{\|\sum_{j\in I_k} Z_j\|>\delta 2^k\}<\infty \quad for\ all\ \delta>0.$$

As a final remark, it should be added that at least under separability if the CLIL(P) holds for $\mathscr{F}$, the set of limit points of the sequence $\left\{a_n^{-1}\sum_{i=1}^{n}(f(X_i)-Pf)\colon f\in\mathscr{F}\right\}$ is a.s. equal to the ball of center zero and radius $\sqrt{2}$ of the reproducing kernel Hilbert space of the covariance of $\delta_{X_i}-P$. In this case the rkhs is

$$\mathscr{B}_{\mathscr{F}}=\{f\to\int gf\,dP\colon f\in\mathscr{F},\ g\in L_2(S,\mathscr{S},P),\ Pg=0\}$$

(Dudley and Kuelbs 1980, Yukich 1986).

(d) A Lemma on Sums of Nonnegative Random Variables

In the proof of Theorem 3.1 we will need the following lemma, which may be of independent interest. See Marcus and Pisier (1984) and Marcus and Zinn (1984) for results of a similar nature.

2.16. Lemma. *Let $\{Z_i\}_{i=1}^{k}$, $k\le\infty$, be a sequence of independent nonnegative real random variables, and let*

$$\|\{Z_i\}\|_{2,\infty}:=\left(\sup_{a>0} a^2\sum_{i=1}^{k} I_{[Z_i>a]}\right)^{1/2},$$

which is equivalent to the $\ell_{2,\infty}$ *norm of* $\{Z_i\}$. *The following inequalities hold:*

a) $$\|\{Z_i\}\|^2_{2,\infty} \leqq \sup_{a>0} a \sum_{i=1}^{k} Z_i I_{[Z_i>a]} \leqq 2\|\{Z_i\}\|^2_{2,\infty}$$

and

b) *if* $$K := \sup_{t>0} t^2 \sum_{i=1}^{k} P\{Z_i>t\} < \infty$$

and

$$Z_i(\omega) \leqq b < \infty \qquad \omega-a.s., \ i=1,\dots,k.$$

then for $c > eK$,

$$P\{\|\{Z_i\}\|^2_{2,\infty} > c\} \leqq (1-eK/c)^{-1}\exp(-c b^{-2}\ln(c/eK)).$$

Proof. Let $\{Z_i^*\}_{i=1}^k$ be the nonincreasing order statistics of $\{Z_i\}_{i=1}^k$, i.e., $Z_i^* \geqq Z_j^*$ for $i \leqq j$, and let $Z_{k+1}^* := 0$. We have

$$\begin{aligned}
\sup_{a>0} a \cdot \sum_{i=1}^{k} Z_i I_{[Z_i>a]} &= \sup_{a>0} a \cdot \sum_{i=1}^{k} Z_i^* I_{[Z_i^*>a]} \\
&= \sup_{1\leqq j\leqq k} \ \sup_{Z_{j+1}^* \leqq a < Z_j^*} a \cdot \sum_{i=1}^{j} Z_i^* \\
&= \sup_{1\leqq j\leqq k} Z_j^* \sum_{i=1}^{j} Z_i^* \\
&\leqq (\sup_{1\leqq i\leqq k} i^{1/2} Z_i^*) \sup_{1\leqq j\leqq k} \left(\sum_{i=1}^{j} i^{-1/2}\right) Z_j^* \\
&\leqq 2 \cdot \sup_{1\leqq i\leqq k} i Z_i^{*2} = 2\|\{Z_i\}\|^2_{2,\infty},
\end{aligned}$$

and part (a) is proved since the first inequality in (a) is trivial. As for part (b) we have

$$\begin{aligned}
P\left(\sup_{a>0} a^2 \cdot \sum_{i=1}^{k} I_{[Z_i>a]} > c\right) &= P(\sup_{1\leqq i\leqq k} i Z_i^{*2} > c) \\
&= P(\sup_{cb^{-2}<i\leqq k} i Z_i^{*2} > c) \quad (\text{since} \sup_{i\leqq cb^{-2}} i Z_i^{*2} \leqq c) \\
&\leqq \sum_{cb^{-2}<i\leqq k} P\left(\sum_{j=1}^{k} I_{[Z_j>(c/i)^{1/2}]} \geqq i\right) \\
&\leqq \sum_{cb^{-2}<i\leqq k} \left[e \sum_{j=1}^{k} P(Z_j>(c/i)^{1/2})/i\right]^i \\
&\leqq \sum_{cb^{-2}<i\leqq k} (eK/c)^i \\
&\leqq (1-eK/c)^{-1}(eK/c)^{cb^{-2}} \\
&= (1-eK/c)^{-1}\exp(-c b^{-2}\ln(c/(eK))),
\end{aligned}$$

where we have used inequality (1.3). □

In Theorem 3.1 we will use both inequalities (a) and (b) together i.e., the exponential inequality for the variable $\sup_{a>0} a \sum_{i=1}^{k} Z_i I_{[Z_i > a]}$.

3. The Basic Inequality

In this section we will prove the basic inequality needed for the proofs of Theorems 4.1 and 5.1. See Remark 3.2 for a clarification of the hypotheses.

3.1. Theorem. *Let $(\Omega, \Sigma, \Pr)$ be a probability space, $(S, \mathscr{S})$ a measurable space and $\{X_i\}_{i=1}^k$ a sequence of independent S-valued random variables with laws $\mathscr{L}(X_i) = P_i$. Let furthermore $\mathscr{F} \subset \bigcap_{i=1}^{k} \mathscr{L}_0(S, \mathscr{S}, P_i)$. For some $q_0 \in \mathbb{N}$ assume that for each $q \geqq q_0$ there exist a set T_q, and functions t_q: $\mathscr{F} \to T_q$, $\tilde{\pi}_q$: $T_q \to \bigcap_{i=1}^{k} \mathscr{L}_0(S, \mathscr{S}, P_i)$, $\tilde{\gamma}_q$: $T_q \to \mathbb{R}_+ \setminus \{0\}$ and $\tilde{\Delta}_q^i$: $T_q \to \mathscr{L}_0(S, \mathscr{S}, P_i)$ for $1 \leqq i \leqq k$, such that, for $\pi_q = \tilde{\pi}_q \circ t_q$, $\gamma_q = \tilde{\gamma}_q \circ t_q$ and $\Delta_q^i = \tilde{\Delta}_q^i \circ t_q$, we have*

(3.1) $$|(f - \pi_q f)| \leqq \Delta_q^i f \quad \text{for } f \in \mathscr{F},\ 1 \leqq i \leqq k,$$

(3.2) $$\beta_{q_0} := \sup_{f \in \mathscr{F}} \sum_{q \geqq q_0} 2^{-q} \gamma_q f < \infty,$$

(3.3) $$\sum_{i=1}^{k} E([\pi_q f - \pi_{q-1} f](X_i))^2 \leqq K \cdot 2^{-2q}, \quad f \in \mathscr{F},$$

(3.4) $$\sup_{t>0} t^2 \sum_{i=1}^{k} \Pr\{\Delta_q^i f(X_i) > t\} \leqq K \cdot 2^{-2q}, \quad f \in \mathscr{F},$$

(3.5) $$\Delta_q^i f \downarrow \text{ as } q \uparrow \quad \text{for } f \in \mathscr{F},\ 1 \leqq i \leqq k,$$

(3.6) $$t_{q-1} f = t_{q-1} g \quad \text{for all } f, g \in \mathscr{F} \text{ such that } t_q f = t_q g,$$

where $K < \infty$ is a constant. Then for all $\alpha \in \mathbb{R}_+$ and $q_1 > q_0$ such that

(3.7) $$\alpha \geqq (K e^2 \beta_{q_0}) \vee ((K k)^{1/2} \cdot 2^{-q_1})$$

we have

(3.8) $$\Pr^*\left\{\left\|\sum_{i=1}^{k} \varepsilon_i (f - \pi_{q_0} f) I_{[\Delta_{q_0}^i f \leqq 2^{-q_0-1}/\gamma_{q_0+1} f]}(X_i)\right\|_{\mathscr{F}} > 12\alpha\right\}$$
$$\leqq 3 \cdot \sum_{q=q_0+1}^{q_1} \sum_{t \in T_q} \exp(-\tfrac{3}{14} \alpha \beta_{q_0}^{-1} \gamma_q(t)^2)$$

where $\{\varepsilon_i\}$ is a Rademacher sequence independent of $\{X_i\}$.

3.2. Remark. The hypotheses (3.1)–(3.6) may be difficult to grasp at first glance. So, at the risk of incurring some repetition (see Sects. 4, 5), let us relate these conditions to majorizing measures and Gaussian processes. Let ρ be the pseudo-distance associated to a Gaussian process on $\mathscr{F}$ whose law is Radon in $\ell^\infty(\mathscr{F})$. Suppose:

(i)
$$\left[\sum_{i=1}^{k} P_i(f-g)^2\right]^{1/2} \leqq \rho(f,g), \quad f,g\in\mathscr{F}$$

and

(ii)
$$\sup_{t>0} t^2 \sum_{i=1}^{k} \Pr\{[\sup_{g\in B_\rho(f,\varepsilon)} |f(X_i)-g(X_i)|]^* > t\} \leqq \varepsilon^2$$

for all $f\in\mathscr{F}$, $\varepsilon>0$.

([ii] is the *local modulus* or *bracketing condition* alluded to in the title of this article.) Then all the assumptions of Theorem 3.1 hold and moreover the right side of (3.8) tends to zero as $q_0\to\infty$. To be more precise, if μ is a discrete majorizing measure for $(\mathscr{F},\rho)$, take $\pi_q f$ as in (2.2); $T_q=\{t_q f:=(q,\pi_1 f,\ldots,\pi_q f)$: $f\in\mathscr{F}\}$ ($(q,\pi_{q_0}f,\ldots,\pi_q f)$ suffices); take the $\gamma_q f$'s to be the $\bar{\gamma}_q t$'s of Lemma 2.4; and let

$$\Delta_q^i f = \min_{1\leqq r\leqq q} [\sup_{g\in B_\rho(\pi_r f,\, 3\cdot 2^{-r})} |g-\pi_r f|]^*$$

where the $*$ denotes the upper measurable envelope with respect to P_i (assume the X_i are perfect).

Proof of Theorem 3.1. Let $\tau_i f:=\min\{q\geqq q_0: \Delta_q^i f>2^{-q-1}/\gamma_{q+1} f\}$, $f\in\mathscr{F}$, $1\leqq i\leqq k$. We have for $f\in\mathscr{F}$ and $1\leqq i\leqq k$ that

$$\begin{aligned} f-\pi_{q_0} f &= f-\pi_{q_1} f + \sum_{q=q_0+1}^{q_1} (\pi_q f-\pi_{q-1} f)\\ &= f-\pi_{q_1} f + \sum_{q=q_0+1}^{q_1} (\pi_q f-\pi_{q-1} f) I_{[\tau_i f<q]}\\ &\quad + \sum_{q=q_0+1}^{q_1} (\pi_q f-\pi_{q-1} f) I_{[\tau_i f\geqq q]}\\ &= (f-\pi_{q_0} f) I_{[\tau_i f=q_0]} + (f-\pi_{q_1} f) I_{[\tau_i f\geqq q_1]}\\ &\quad + \sum_{q=q_0+1}^{q_1-1} (f-\pi_q f) I_{[\tau_i f=q]} + \sum_{q=q_0+1}^{q_1} (\pi_q f-\pi_{q-1} f) I_{[\tau_i f\geqq q]}. \end{aligned}$$

So, by (3.9) and (3.1)

(3.10)
$$\begin{aligned} &\Pr^*\left\{\left\|\sum_{i=1}^{k} \varepsilon_i (f-\pi_{q_0} f) I_{[\tau_i f>q_0]}(X_i)\right\|_{\mathscr{F}} > 12\alpha\right\}\\ &\leqq \Pr^*\left\{\left\|\sum_{i=1}^{k} \Delta_{q_1}^i f\, I_{[\tau_i f\geqq q_1]}(X_i)\right\|_{\mathscr{F}} > 3\alpha\right\}\\ &\quad + \Pr^*\left\{\left\|\sum_{q=q_0+1}^{q_1-1} \sum_{i=1}^{k} \Delta_q^i f\, I_{[\tau_i f=q]}(X_i)\right\|_{\mathscr{F}} > 8\alpha\right\}\\ &\quad + \Pr^*\left\{\left\|\sum_{q=q_0+1}^{q_1} \sum_{i=1}^{k} \varepsilon_i (\pi_q f-\pi_{q-1} f) I_{[\tau_i f\geqq q]}(X_i)\right\|_{\mathscr{F}} > \alpha\right\}\\ &:= \mathrm{I}+\mathrm{II}+\mathrm{III}. \end{aligned}$$

Estimation of I. If $\{Z_i\}_{i=1}^k$ are non-negative random variables then for $c = \sup_{t>0} t^2 \sum_{i=1}^k \Pr\{Z_i > t\}$ and $a = (c/k)^{1/2}$ we have

$$\sum_{i=1}^k E Z_i \leqq k \cdot a + \int_a^\infty \sum_{i=1}^k \Pr(Z_i > t)\, dt$$

$$\leqq k \cdot a + c \cdot \int_a^\infty t^{-2}\, dt = 2(k c)^{1/2}.$$

So, by (3.4) and (3.7)

$$\sum_{i=1}^k E(\Delta_{q_1}^i f)(X_i) \leqq 2(k K)^{1/2} 2^{-q_1} \leqq 2\alpha, \quad f \in \mathscr{F}. \tag{3.11}$$

Since $\{\tau_i f \geqq q_1\} \subseteq \{\Delta_{q_1}^i f \leqq 2^{-q_1}(\gamma_{q_1} f)^{-1}\}$, (3.11) gives

$$\sum_{i=1}^k E(\Delta_{q_1}^i f I_{[\tau_i f \geqq q_1]}(X_i))^2 \leqq 2\alpha\, 2^{-q_1}(\gamma_{q_1} f)^{-1}. \tag{3.12}$$

By (3.6) $\Delta_{q_1}^i f I_{[\tau_i f \geqq q_1]}$ only depends on f through $t_{q_1} f$ so we get by Berntein's inequality after centering,

$$\begin{aligned} \mathrm{I} &\leqq \sum_{t \in T_{q_1}} \Pr\left\{ \sum_{i=1}^k \Delta_{q_1}^i(t) I_{[\tau_i(t) \geqq q_1]}(X_i) > 3\alpha \right\} \\ &\leqq \sum_{t \in T_{q_1}} 2 \cdot \exp(-\tfrac{3}{14}\alpha\, 2^{q_1} \gamma_{q_1}(t)) \\ &\leqq 2 \cdot \sum_{t \in T_{q_1}} \exp(-\tfrac{3}{14}\alpha \beta_{q_0}^{-1} \gamma_{q_1}(t)^2) \end{aligned} \tag{3.13}$$

since $\beta_{q_0} \geqq 2^{-q_1} \gamma_{q_1}(t)$, $t \in T_{q_1}$.

Estimation of II. Note that

$$\{\tau_i f = q\} \subseteq \{(2\gamma_{q+1} f)^{-1} < 2^q \Delta_q^i f \leqq (\gamma_q f)^{-1}\}$$

and that $\Delta_q^i f$ and $\gamma_q f$ only depend on f through $t_q f$. So

$$\begin{aligned} \mathrm{II} &= \Pr{}^*\left\{ \left\| \sum_{q=q_0+1}^{q_1-1} (2^{-q}\gamma_{q+1} f)(\gamma_{q+1} f)^{-1} \sum_{i=1}^k 2^q \Delta_q^i f I_{[\tau_i f = q]}(X_i) \right\|_{\mathscr{F}} > 8\alpha \right\} \\ &\leqq \sum_{q=q_0+1}^{q_1-1} \sum_{t \in T_q} \Pr\left\{ \sup_{\{f: t_q f = t\}} (\gamma_{q+1} f)^{-1} \sum_{i=1}^k 2^q \Delta_q^i(t) I_{[\tau_i f = q]}(X_i) > 4\alpha \beta_{q_0}^{-1} \right\} \\ &\leqq \sum_{q=q_0+1}^{q_1-1} \sum_{t \in T_q} \Pr\left\{ \sup_{a>0} a \cdot \sum_{i=1}^k 2^q \Delta_q^i(t) I_{[a < 2^q \Delta_q^i(t) \leqq \gamma_q(t)^{-1}]}(X_i) > 2\alpha \beta_{q_0}^{-1} \right\}. \end{aligned} \tag{3.14}$$

Now, Lemma 2.16, (3.4) and (3.7) yield

$$\Pr\left\{\sup_{a>0} a\cdot\sum_{i=1}^{k} 2^q \Delta_q^i(t)\, I_{[a<2^q\Delta_q^i(t)\leqq\gamma_q(t)^{-1}]}(X_i)>2\alpha\beta_{q_0}^{-1}\right\}$$

$$\leqq 2\cdot\exp(-\alpha\beta_{q_0}^{-1}\gamma_q(t)^2).$$

So, by (3.14)

$$\text{(3.15)}\qquad \mathrm{II}\leqq 2\cdot\sum_{q=q_0+1}^{q_1-1}\sum_{t\in T_q}\exp(-\alpha\beta_{q_0}^{-1}\gamma_q(t)^2).$$

Estimation of III. Here we note that $(\pi_q f-\pi_{q-1}f)\,I_{[\tau_i f\geqq q]}$ and $\pi_q f$ only depend on f through $t_q f$. Hence

$$\text{(3.16)}\quad \mathrm{III}=\Pr{}^*\left\{\left\|\sum_{q=q_0+1}^{q_1} 2^{-q}(\gamma_q f)(2^{-q}\gamma_q f)^{-1}\cdot\sum_{i=1}^{k}\varepsilon_i(\pi_q f-\pi_{q-1}f)\,I_{[\tau_i f\geqq q]}(X_i)\right\|_{\mathscr{F}}>\alpha\right\}$$

$$\leqq\sum_{q=q_0+1}^{q_1}\sum_{t\in T_q}\Pr\left\{\left|\sum_{i=1}^{k}\varepsilon_i(\pi_q(t)-\pi_{q-1}(t))\,I_{[\tau_i(t)\geqq q]}(X_i)\right|>\alpha\, 2^{-q}\gamma_q(t)\,\beta_{q_0}^{-1}\right\}.$$

Since (by (3.1) and (3.5)) $|\pi_q f-\pi_{q-1}f|\leqq 2^{-q+1}(\gamma_q f)^{-1}$ on $\{\tau_i f\geqq q\}$ Bernstein's inequality and (3.3) give

$$\Pr\left\{\left|\sum_{i=1}^{k}\varepsilon_i(\pi_q(t)-\pi_{q-1}(t))\,I_{[\tau_i(t)\geqq q]}(X_i)\right|>\alpha\, 2^{-q}\gamma_q(t)\,\beta_{q_0}^{-1}\right\}$$

$$\leqq\exp\{-\alpha^2(2K\beta_{q_0}+4/3\,\alpha)^{-1}\beta_{q_0}^{-1}\gamma_q(t)^2\}.$$

This, together with (3.16) and using (3.7), gives

$$\text{(3.17)}\qquad \mathrm{III}\leqq\sum_{q=q_0+1}^{q_1}\sum_{t\in T_q}\exp(-\tfrac{1}{2}\alpha\beta_{q_0}^{-1}\gamma_q(t)^2).$$

Now since $\{\tau_i f=q_0\}=\{\Delta_{q_0}^i f>2^{-q_0-1}/\gamma_{q_0+1}f\}$, (3.10), (3.13), (3.15) and (3.17) imply (3.8). □

4. The Central Limit Theorem

We obtain first a non-i.i.d. CLT, Gaussian convergence, that includes Donsker's invariance principle and then we specialize to i.i.d. random elements.

(a) A CLT for Non-I.I.D. Random Variables

Using Theorem 2.9 and Theorem 3.1 we obtain the following "bracketing" CLT for not necessarily i.i.d. random variables. Set-up and notation are as in Theorem 2.9.

4.1. Theorem. *Let* $\mathscr{F} \subset \bigcap_{n,j} \mathscr{L}_1(S, \mathscr{S}, P_{nj})$ *and satisfy* (2.5), *and let* F *be the envelope function of* $\mathscr{F}$. *Assume*

(i) *For every* $k \in \mathbb{N}$, *and* $f_1, \ldots, f_k \in \mathscr{F}$, *the finite dimensional distributions* $\left\{\mathscr{L}\left[a_n^{-1} \sum_{j=1}^{n} (f_i(X_{nj}) - P_{nj} f_i)_{i=1}^k\right]\right\}_{n=1}^{\infty}$ *converge weakly;*

(ii) $\sum_{j=1}^{n} P_{nj}^* \{F > t\, a_n\} \underset{n \to \infty}{\to} 0$ *for all* $t > 0$;

and that there exists a pseudo-distance ρ *on* $\mathscr{F}$ *dominated by the distance* d_G *of a centered Gaussian process* G *on* $\mathscr{F}$ *with bounded* d_G*-uniformly continuous paths, such that*

(iii) $a_n^{-2} \sum_{j=1}^{n} P_{nj}(f-g)^2 \leqq \rho^2(f, g)$ *for all* $f, g \in \mathscr{F}$

and

(iv) *for all* $f \in \mathscr{F}$ *and* $\varepsilon > 0$,

$$\sup_{t>0} t^2 \sum_{j=1}^{n} P_{nj}^* \{ \sup_{g \in B_\rho(f, \varepsilon)} |f-g| > t\, a_n\} \leqq \varepsilon^2.$$

Then $\mathscr{F} \in \mathrm{CLT}\{P_{nj}; a_n\}$ *and the limiting measure is Gaussian.*

Proof. Let $T := \{\pi_q f : f \in \mathscr{F}, q \in \mathbb{N}\}$ and let μ be a probability measure on T satisfying the conditions (2.2) and (2.4) for $d = d_G$. These exist by Talagrand's theorem 2.5(ii). Then the π_q's satisfy conditions (i) and (iii) in Theorem 2.9. By hypothesis (ii), the finite dimensional limits are Gaussian. So, only condition (ii) in Theorem 2.9 remains to be checked. For this, we will use Theorem 3.1. Let $\mathscr{F}_n = \{a_n^{-1} f : f \in \mathscr{F}\}$. For all $r, q, n, j \in \mathbb{N}$, $f \in \mathscr{F}$, $\bar{f} = a_n^{-1} f \in \mathscr{F}_n$ we let

(4.1) $$t_q f = t_q \bar{f} = \{\pi_r f\}_{r=1}^q, \qquad T_q = t_q \mathscr{F},$$

(4.2) $$\pi_q \bar{f} = a_n^{-1} \pi_q f,$$

(4.3) $$\gamma_q f := \gamma_q \bar{f} := \gamma_q(t_q f) := \left[\ln\left(2^q \Big/ \prod_{r=1}^{q} \mu\{\pi_r f\}\right)\right]^{1/2},$$

(4.4) $$\beta_q = \sup_{f \in \mathscr{F}} \sum_{r \geqq q} 2^{-r} \gamma_r f,$$

(4.5) $$\Delta_q^{nj} f := \Delta_q^{nj} \bar{f} := \Delta_q^{nj}(t_q f) := a_n^{-1} \min_{1 \leqq r \leqq q} [\sup_{g \in B_\rho(\pi_r f, 3 \cdot 2^{-r})} |g - \pi_r f|]^*$$

where the upper envelopes are taken with respect to P_{nj}. Obviously

(4.6) $$2^{-1} q^{1/2} \leqq \gamma_q f \uparrow \quad \text{and} \quad \Delta_q^{nj} f \downarrow \quad \text{as } q \uparrow,$$

and by Lemma 2.4,

(4.7) $$\beta_q \downarrow 0, \quad \beta_q < \infty \quad \text{for all } q \geqq 1$$

and

(4.8) $$\lim_{r\to\infty} \sum_{q=1}^{\infty} \sum_{t\in T_q} \exp(-r\gamma_q(t)^2)=0.$$

Let $\alpha>0$ be fixed and let $\mathrm{Pr}_n = P_{n1}\otimes\ldots\otimes P_{nn}$. Then

(4.9) $$\begin{aligned}\mathrm{Pr}_n^*\left\{\left\|\sum_{j=1}^n \varepsilon_j(f-\pi_{q_0}f)(X_{nj})\right\|_{\mathscr{F}} > 2\alpha a_n\right\} &= \mathrm{Pr}_n^*\left\{\left\|\sum_{j=1}^n \varepsilon_j(f-\pi_{q_0}f)\, I_{[\Delta_{q_0}^{nj} f > 2^{-q_0-1}/\gamma_{q_0+1}f]}(X_i)\right\|_{\mathscr{F}} > \alpha a_n\right\}\\ &\quad + Pr_n^*\left\{\left\|\sum_{j=1}^n \varepsilon_j(\bar f-\pi_{q_0}\bar f)\, I_{[\Delta_{q_0}^{nj} \bar f \leqq 2^{-q_0-1}/\gamma_{q_0+1}\bar f]}(X_i)\right\|_{\mathscr{F}_n} > \alpha\right\}\\ &:= \mathrm{I}+\mathrm{II}.\end{aligned}$$

Estimation of I. Since $\Delta_{q_0}^{nj}\bar f(X_{nj}) \leqq 2\cdot a_n^{-1} F(X_{nj})^*$ and

$b_{q_0} := 2^{-q_0-1}(\sup_{f\in\mathscr{F}} \gamma_{q_0+1} f)^{-1} > 0$ (by (4.7)) we have by (ii)

(4.10) $$\mathrm{I} \leqq \sum_{j=1}^n P_{nj}^*\{F > \tfrac{1}{2} b_{q_0} a_n\} \to 0 \quad \text{as } n\to\infty.$$

Estimation of II. The conditions (3.1)–(3.7) in Theorem 3.1 are satisfied for $K=9$ and for q_0 and $q_1(n) > q_0$ sufficiently large (use (2.2), (4.5)–(4.7), [iii and iv]). So, by (4.7) and (4.8)

(4.11) $$\lim_{q_0\to\infty} \limsup_n \mathrm{II} \leqq \lim_{q_0\to\infty} \sum_{q=q_0+1}^{\infty} \sum_{t\in T_q} \exp(-\tfrac{3}{14}\alpha\beta_{q_0}^{-1}\gamma_{q_0}(t)^2) = 0.$$

Now (ii) in Theorem 2.9 follows from (4.9), (4.10) and (4.11). □

4.2. Remark. Minor modifications of the above proof yield stochastic boundedness of the sequence $\left\{a_n^{-1}\left\|\sum_{j=1}^n (f(X_{nj}) - P_{nj}(f))\right\|_{\mathscr{F}}\right\}_{n=1}^{\infty}$ under the weaker assumptions

(ii)′ $$\sup_{t>0} \sum_{j=1}^n P_{nj}^*\{F > t\, a_n\} < \infty$$

and

(iii)′ and (iv)′: (iii) and (iv) with ρ dominated by the distance d_G of a centered Gaussian process G with bounded sample paths.

4.3. Example. Donsker's invariance principle. Let $\{\xi_j\}_{j=1}^{\infty}$ be i.i.d. real valued random variables with $E\xi_j^2=1$, $E\xi_j=0$. Define S-valued random variables, with $S=(\mathbb{Q}\cap[0,1])^2\times\mathbb{R}$, by $X_{nj}=((j-1)/n, j/n, \xi_j)$, $j=1,\ldots,n$, $n\in\mathbb{N}$, and let $\mathscr{F}=\{f_t\colon t\in[0,1]\}$, with $f_t(r_1,r_2,x)=0$ if $t\leqq r_1$, $=x$ if $t\geqq r_2$, and linear for t between r_1 and r_2, where $0\leqq r_1\leqq r_2\leqq 1$, $r_1, r_2\in\mathbb{Q}$, $x\in\mathbb{R}$ (and set $f_t(r_1,r_2,x)=0$ if $r_2<r_1$). Then the invariance principle for the sequence $\{\xi_i\}$ is equivalent to $\mathscr{F}\in\mathrm{CLT}\{P_{nj}; n^{1/2}\}$ with $P_{nj}=(X_{nj})$. But standard computations show that Theorem 4.1 gives

$\mathscr{F} \in \mathrm{CLT}\{P_{nj}; n^{1/2}\}$ just by taking $\rho(f_t, f_s) = 2|t-s|^{1/2}$ in (iii) and (iv). (Since $\max_{j \leq n} |\xi_j|/n^{1/2} \to 0$ a.s., here we mean the invariance principle for both the polygonal lines and the step functions constructed from the partial sums of $\xi_j/n^{1/2}$.)

(b) The Central Limit Theorem for I.I.D. Random Variables

The following theorem follows directly from Theorem 4.1.

4.4. Theorem. *Let $\mathscr{F} \subset \mathscr{L}_2(S, \mathscr{S}, P)$ and let F be its envelope function. Assume $\|Pf\|_{\mathscr{F}} < \infty$ and*

(i) $$\lim_{t \to \infty} t^2 P\{F^* > t\} = 0.$$

Assume further that

(ii) $\mathscr{F}$ *is P-pregaussian*

and that

(iii) *there exists a bounded and d_G-uniformly continuous centered Gaussian process G such that for all $\varepsilon > 0$ and all $f \in \mathscr{F}$*

$$\Lambda_{2,\infty}^P([\sup_{g \in B_{d_G}(f,\varepsilon)} |f-g|]^*) \leq \varepsilon.$$

Then $\mathscr{F} \in \mathrm{CLT}(P)$.

We now restate the theorem in a more standard "bracketing" form (although the examples that will follow seem to indicate that condition (iii) in Theorem 4.4 is more readily applicable than the equivalent bracketing condition that follows).

4.4′. Theorem. *Let $\mathscr{F} \subset \mathscr{L}_1(S, \mathscr{S}, P)$ and let F be its envelope function. Assume $\|Pf\|_{\mathscr{F}} < \infty$ and*

(i) $$\lim_{t \to \infty} t^2 P(F^* > t) = 0.$$

Assume further that

(ii) $\mathscr{F}$ *is P-pregaussian,*

and that

(iii) *for all $q \in \mathbb{N}$ and $f \in \mathscr{F}$ there exist measurable functions $\ell_q(f)$ and $u_q(f)$ on $(S, \mathscr{S})$ and there exists a finite measure μ on the pairs $\{(\ell_q(f), u_q(f)): q \in \mathbb{N}, f \in \mathscr{F}\}$ such that*

(a) *$\ell_q(f) \leq f \leq u_q(f)$ for all $f \in \mathscr{F}$ and $q \in \mathbb{N}$,*

(b) *$\sup_{f \in \mathscr{F}} \Lambda_{2,\infty}^P(u_q(f) - \ell_q(f)) \leq 2^{-q}$ for all $q \in \mathbb{N}$,*

(c) *$\lim_{r \to \infty} \sup_{f \in \mathscr{F}} \sum_{q=r}^{\infty} 2^{-q} [\ln(\mu\{(\ell_q(f), u_q(f))\})^{-1}]^{1/2} = 0$ and the sup is finite for $r = 1$.*

Then $F \in \mathrm{CLT}(P)$.

Proof. Assume the hypotheses of Theorem 4.4 hold. Let $\{\pi_q f\}$ and ν satisfy (2.2)(i), (ii) and (2.4) for ρ. Define

$$u_q(f)=\pi_{q+1}f+(\sup_{g\in B_\rho(\pi_{q+1}f,\,2^{-q-1})}|\pi_{q+1}f-g|)^*,$$

$$\ell_q(f)=\pi_{q+1}f-(\sup_{g\in B_\rho(\pi_{q+1}f,\,2^{-q-1})}|\pi_{q+1}f-g|)_*$$

and $\mu\{(\ell_q f, u_q f)\}=v\{\pi_{q+1}f\}$ for all $f\in\mathscr{F}$. Then $\ell_q(f)$, $u_q(f)$ and μ satisfy condition (iii) in Theorem 4.4′. Conversely, assume condition (iii) in Theorem 4.4′ holds. Define

$$h_q(f)=\{(\ell_r(f), u_r(f))\}_{r=1}^q \qquad \text{for all } f\in\mathscr{F} \text{ and } q\in\mathbb{N}$$

and $\bar{\mu}\{h_q(f)\}=2^{-q}\prod_{r=1}^{q}\mu\{(\ell_r(f), u_r(f))\}$. Then the argument in Lemma 2.4 shows that $\bar{\mu}$ verifies (2.4) with $\pi_q\, t$ replaced by $h_q(f)$. Define

$$\bar{\rho}(f,g)=2^{-r(f,g)+1} \qquad \text{where } r(f,g)=\inf\{q\colon h_q(f)\neq h_q(g)\}.$$

Choose a function h in each class $\{g\colon h_q(f)=h_q(g)\}$, $q\in\mathbb{N}$, $f\in\mathscr{F}$, and define $\pi_{q-1}f=h$. Let $v\{\pi_{q-1}f\}=\bar{\mu}\{h_q f\}$, $q\in\mathbb{N}$, $f\in\mathscr{F}$. Since $\bar{\rho}(f,\pi_q f)\leqq 2^{-q}$ and v verifies (2.4) it follows that v is a majorizing measure for $\bar{\rho}$. Hence by Remark 2.6(2), $\bar{\rho}$ is dominated by the pseudo-distance d_G of a Gaussian process G with bounded and ρ-uniformly continuous sample paths. Finally the $\Lambda_{2,\infty}$-condition in Theorem 4.4 is obvious from the definitions of ρ ($\bar{\rho}$) and the h_q's and from (iii)(c) in Theorem 4.4′. □

The improvements of Theorem 4.4 on the already sharp CLT under bracketing in Ossiander (1985) are apparent from 4.4′: the L_2-brackets are replaced by $\Lambda_{2,\infty}$-brackets and the entropy condition is replaced by the weaker majorizing measure condition. This weakening of the L_2 and entropy conditions considerably widens the applicability of bracketing.

In order to make more transparent the relationship between Theorem 4.4 and the Jain-Marcus CLT, we reformulate Theorem 4.4 for processes. We recall first that a centered stochastic process $\{X(t)\colon t\in T\}$ on a metric space (T,d) is *pregaussian* if its covariance coincides with the convariance of a centered Gaussian process G_X on T with bounded and uniformly d-continuous sample paths. If no metric d is specified for T, we take $d=d_X$ in this definition, with $d_X(s,t)\coloneqq(E(X(t)-X(s))^2)^{1/2}$, $s,t\in T$.

4.4″. Theorem. *Let $\{X(t)\colon t\in T\}$ be a sample bounded process on a set T such that $EX(t)=0$ and $EX^2(t)<\infty$ for all $t\in T$. Assume:*

(i) $u^2 P^*\{\|X\|_\infty>u\}\to 0$ *as* $u\to\infty$,

(ii) *X is pregaussian, and*

(iii) *there is pseudometric ρ on T dominated by the pseudometric d_G corresponding to a centered Gaussian process G on T with bounded and uniformly d_G-continuous path such that for some $K>0$ and for all $t\in T$ and $\varepsilon>0$,*

$$\Lambda_{2,\infty}((\sup_{s\in B_\rho(t,\varepsilon)}|X(t)-X(s)|)^*)\leqq K\varepsilon.$$

Then $X \in \mathrm{CLT}$ *as a* $\ell^\infty(T)$*-valued random element. Moreover, if* (T, d) *is compact metric and* X *is sample continuous, then conditions* (i)–(iii) *imply that* $X \in \mathrm{CLT}$ *in* $C(T, d)$.

4.5. Corollary. *Let* (T, d) *be a compact metric space and let* $\{X(t)\colon t \in T\}$ *be a centered stochastic process such that* $EX^2(t) < \infty$ *for all* $t \in T$. *Assume:*

(i) *X is pregaussian, and*

(ii) *there exist* $M \in \Lambda_{2,\infty}(\Omega, \mathscr{F}, \Pr)$ *and a pseudometric* ρ *dominated by* d_G *for some centered Gaussian process G with bounded and uniformly* d_G*-continuous paths, such that*

$$|X(t, \omega) - X(s, \omega)| \leq M(\omega)\, \rho(s, t), \qquad \text{for all } \omega \in \Omega,\ s, t \in T. \tag{4.12}$$

Then $X \in \mathrm{CLT}$ *as a* $C(T, d)$*-valued random variable.*

Proofs. The first part of Theorem 4.4″ is obvious: take $S = \ell^\infty(T)$, $\mathscr{S} =$ the cylindrical σ-algebra of S, $P = \mathscr{L}(X)$ and $\mathscr{F} = \{\delta_t\colon t \in T\}$. If X is sample continuous on (T, d), take instead $\mathscr{S} = C(T, d)$ and $\mathscr{S} =$ the Borel σ-algebra of $\mathscr{S}$. Then note that if $\mathscr{F} \in \mathrm{CLT}(P)$, since G_X has bounded d-continuous paths, the finite dimensional approximation condition of Theorem 2.9 (iii) holds for $\rho = d$, and therefore $X \in \mathrm{CLT}$ as a $C(T, d)$-valued random variable. This proves the last statement of Theorem 4.4″. To prove the corollary note first that condition (4.12) is much stronger than condition (iii) in Theorem 4.4″. Finally, it remains to check that $EX^2(t) < \infty$ for all $t \in T$ and condition (4.12) imply the tail condition $n^2 P\{\|X\|_\infty > n\} \to 0$ as $n \to \infty$. If $\varepsilon > 0$ and if $\{s_i\}$ is an ε-net for ρ then $\|X\|_\infty \leq \max_i |X(s_i)| + \varepsilon M$ and therefore $\limsup_{n\to\infty} n^2 P\{\|X\|_\infty > n\} < \varepsilon$. ε being arbitrary, the tail condition holds. □

4.6. Remarks. (1) In the Jain-Marcus CLT, the conditions on M and ρ are stronger than in Corollary 4.5: in their theorem $M \in L_2(\Omega, \Sigma, \Pr)$ and ρ satisfies a metric entropy condition that implies its domination by a d_G metric. (2) B. Heinkel (to appear) has recently obtained a direct proof of our Corollary 4.5. Heinkel (1977) had proved a weaker version of Corollary 4.5 (with $M \in L_2$).

Next we indicate how a variety of known interesting CLT's are direct consequences of Theorem 4.1.

4.7. Example. Let $\{\xi_n\}_{n\in\mathbb{N}}$ be random variables such that $\varepsilon_n := (Ln)^{1/2} \cdot \Lambda_{2,\infty}(\sup_{j\geq n} (Lj)^{-1/2} |\xi_j|) \to 0$ as $n \to \infty$ and the vector $X = ((Ln)^{-1/2} \xi_n : n \in \mathbb{N})$, which is obviously a.s. in c_0, is pregaussian. We show now that Theorem 4.4 immediately gives that $X \in CLT$. (Paulauskas 1980, and Heinkel 1983, considered a similar but less general situation.) We must prove that the class of coordinate functions on c_0, $\mathscr{F} = \{f_n\colon f_n(x) = x_n,\ x \in c_0\}$, satisfies $CLT(P)$ with $P = \mathscr{L}(X)$. The hypotheses readily imply that $t^2 P(F > t) \to 0$. Moreover, the pseudo-distance $\rho(f_n, f_m) = \varepsilon_n (Ln)^{-1/2}$ for $0 < n < m < \infty$, is dominated by a Gaussian distance (that of $G = \{\varepsilon_n (Ln)^{-1/2} g_n\}_{n\in\mathbb{N}}$, with g_i i.i.d. $N(0, 1)$). Hence, to apply Theorem 4.4 it suffices to verify (iii) for ρ. Given $\varepsilon > 0$, if $\varepsilon_n (Ln)^{-1/2} \leq \varepsilon < \varepsilon_{n-1} (L(n-1))^{-1/2}$ then $B_\rho(f_k, \varepsilon) = \{f_k\}$ for $k < n$ and $B_\rho(f_k, \varepsilon) \subseteq \{f_j\}_{j\geq n}$ for $k \geq n$. So, the quantity in (iii) is 0 for $k < n$ and is bounded by $\Lambda_{2,\infty}(2 \sup_{j\geq n} (|\xi_j|/(Lj)^{1/2})) \leq 2\varepsilon_n (Ln)^{1/2} \leq 2\varepsilon$ for $k \geq n$.

4.8. Example. Let $\{X(t)\}_{t\in T}$ be a centered uniformly bounded stochastic process with all its sample paths in $C[0,1]$ ($D[0,1]$). Giné and Zinn (1984) prove that if

$$E|X(t)-X(s)|\leqq K|t-s|, \qquad s,t\in[0,1]$$

for some $K<\infty$ then $X\in\mathrm{CLT}$ as a $C[0,1]$-valued ($D[0,1]$-valued) random vector. This is also an easy consequence of Theorem 4.4. To see this let us prove that $\mathscr{F}=\{\delta_t\colon t\in[0,1]\}\in\mathrm{CLT}(P)$, where $P=\mathscr{L}(X)$. Assume $\sup_{t,\omega}|X(t,\omega)|\leqq 1$. Obviously $\mathscr{F}$ is P-pregaussian. Take $d_G(s,t)=(2K|t-s|)^{1/2}$ in (iii) in Theorem 4.4 and observe that for t fixed

$$\sup_{|t-s|\leqq\varepsilon}|X(t)-X(s)|\leqq\lim_{n\to\infty}\sum_{k=-2^n}^{2^n-1}|X(t+\varepsilon k2^{-n})-X(t+\varepsilon(k+1)2^{-n})|$$

so that for each t

$$E\sup_{\{s:\,d_G(s,t)\leqq\varepsilon\}}|X(t)-X(s)|\leqq\varepsilon$$

and (iii) holds.

4.9. Example. Weighted empirical processes. The Chibisov-O'Reilly theorem without continuity assumptions on the weight (as in Csörgö et al. 1986 and Dudley 1985) can also be obtained from Theorem 4.4. Let P be the Lebesgue measure on $[0,1]$ and let $\mathscr{F}=\{w(t)I_{[0,t]}\colon t\in(0,1/2]\}$ where w is a measurable weight function for which there exists $\gamma\in(0,1/2)$ such that w is non-increasing on $(0,\gamma)$ and uniformly bounded on $[\gamma,1/2]$ (usually one considers $t\in(0,1)$ and w U-shaped; the problem reduces to this "half" by symmetry). The Chibisov-O'Reilly theorem assert that the following are equivalent:

(a) $\mathscr{F}\in\mathrm{CLT}(P)$,

(b) $\mathscr{F}$ is P-pregaussian and $t^{1/2}w(t)\to 0$ as $t\to 0$,

(c) $\int_0^\gamma t^{-1}\exp(-\varepsilon/tw^2(t))\,dt<\infty$ for all $\varepsilon>0$.

Since for $\mathscr{F}=\bigcup_{i=1}^n\mathscr{F}_i$, $\mathscr{F}\in\mathrm{CLT}(P)$ iff $\mathscr{F}_i\in\mathrm{CLT}(P)$ for all $i\leqq n$ (Alexander 1985) we may consider separately $\mathscr{F}_i$, $i=1,2$, where $\mathscr{F}_1=\{w(t)I_{[0,t]}\colon t\in(0,\gamma)\}$ and $\mathscr{F}_2=\mathscr{F}\setminus\mathscr{F}_1$. Trivially, $\mathscr{F}_2\in\mathrm{CLT}(P)$. So it is enough to consider $\mathscr{F}_1$. Condition (c) implies that $t^{1/2}w(t)\to 0$ as $t\to 0$, hence the envelope function $F=w$ of $\mathscr{F}_1$ satisfies (i) in Theorem 4.4 under either (c) or (b). Now, (b) holds iff $w(t)W(t)\to 0$ a.s. as $t\to 0$, where $W(t)$ is Brownian motion (a version with continuous sample paths), and this is equivalent to (c) as follows. Independent increments and the Borel-Cantelli lemma show that the integral condition is equivalent to $w(2^{-k})W(2^{-k})\to 0$ a.s. as $k\to\infty$; to show that this implies $w(t)W(t)\to 0$ a.s. we use Levy's inequality to get

$$\sum_k\Pr\{\sup_{t\in(2^{-k+1},2^k]}|w(t)W(t)-w(2^{-k})W(2^{-k})|>\varepsilon\}$$
$$\leqq 4\sum_k\exp\{-\varepsilon^2/8\cdot 2^k w(2^{-k-1})^2\}.$$

This last expression converges for all $\varepsilon>0$ iff the integral condition holds (see e.g., Dudley 1985 for details).

So it remains to show that, under the integral condition (c), (iii) in Theorem 4.4 holds. Assume without loss of generality that w is right continuous. Take as G the centered Gaussian process

$$G(t)=2^{3/2}(w(t)\,W_1(t)+w(t-)\,W_2(t)), \qquad t\in(0,\gamma)$$

where W_1 and W_2 are two independent Brownian motions, so that

$$\begin{aligned} d_G(s,t)^2 &= 8[(t-s)\,w(t)^2+s(w(s)-w(t))^2] \\ &\quad +8[(t-s)\,w(t-)^2+s(w(s-)-w(t-))^2] \qquad 0<s\leqq t\leqq\gamma. \end{aligned}$$

Let $M_s^+(u)=\sup_{t\in B_{d_G}(s,\varepsilon),\ t>s}|w(t)\,I_{[0,t]}(u)-w(s)\,I_{[0,s]}(u)|$, $u, s\in[0,1]$ and let M_s^- be the same sup for $t<s$. Since M_s^+ and M_s^- can be dealt with in similar ways, we only prove

$$\Lambda_{2,\infty}^2(M_s^+)\leqq\varepsilon^2/2 \qquad \text{for } \varepsilon>0,\ s\in[0,1].$$

Let $s\in[0,1]$ and $\varepsilon>0$ be fixed, let $t_1:=\sup\{t\in B_{d_G}(s,\varepsilon)\colon s\leqq t\leqq\gamma\}$ and

$$(s,t_1\}:=\begin{cases}(s,t_1] & \text{if } t_1\in B_{d_G}(s,\varepsilon)\\ (s,t_1) & \text{otherwise}\end{cases}.$$

For $u\in[0,\gamma]$ set $u':=\inf\{t\in B_{d_G}(s,\varepsilon)\colon t\geqq u\}$, $u'':=\sup\{t\in B_{d_G}(s,\varepsilon)\colon t\leqq u\}$ and note that if $u\in(s,t_1\}$ then $u'\in(s,t_1]$. Then for all $u>\gamma$, $M_s^+(u)=0$, and for all $u\in(0,\gamma]$, $M_s^+(u)=(w(s)-w(t_1))\,I_{[0,s]}(u)+w(u')\,I_{[s,t_1\}}(u)$. So, with $t_\alpha=\sup\{u\colon w(u)\geqq\alpha\}$, we have

$$\begin{aligned} \Lambda_{2,\infty}^2(M_s^+) &= \sup_{\alpha>0}\alpha^2 P\{u\colon M_s^+(u)\geqq\alpha\} \\ &= \sup_{\alpha>0}[\alpha^2 P\{u\leqq s\colon w(s)-w(t_1)\geqq\alpha\}+\alpha^2 P\{u\in(s,t_1\}\colon w(u')\geqq\alpha\}] \\ &\leqq (w(s)-w(t_1))^2 s+\sup_{\alpha>0}[w(t_\alpha-)^2 P\{u\in(s,t_1\}\colon u'\leqq t_\alpha\}] \\ &\leqq (w(s)-w(t_1))^2 s+\sup_{\alpha>0}[w((t''_\alpha\wedge t_1)-)^2 P\{u\in(s,t_1\}\colon u\leqq t''_\alpha\wedge t_1\}] \\ &\leqq (w(s)-w(t_1))^2 s+\sup_{t\in B_{d_G}(s,\varepsilon),\,t\geqq s}[w(t-)^2(t-s)] \\ &\leqq \sup_{t\in B_{d_G}(s,\varepsilon)} d_G(s,t)^2/2\leqq\tfrac{1}{2}\varepsilon^2 \end{aligned}$$

Hence, (iii) in Theorem 4.4 is fulfilled. □

Alexander (1985) proved the CLT for VC-graph classes, and his CLT also implies the CLT for weighted empiricals, in fact more directly than our Theorem 4.4.

The previous examples show a very wide range of applicability for Theorem 4.4. Note that 4.5, 4.7 and 4.9 cannot be obtained from bracketing entropy CLT's (Alexander, private communication, for 4.9), and that 4.5 and 4.9 satisfy $EF^2=\infty$, hence cannot be obtained from an L_2-bracketing theorem. All the above examples are sharp in one sense or other (for Example 4.8 see Rhee [1985]).

5. The Law of the Iterated Logarithm

In this section we prove:

5.1. Theorem. *Let* $\mathscr{F} \subset \mathscr{L}_2(S, \mathscr{S}, P)$ *and let* F *be its envelope function. Assume* $\|Pf\|_{\mathscr{F}} < \infty$ *and*

(i) $P^*(F^2/L_2 F) < \infty$.

Assume further that there is a bounded (resp. bounded and d_G*-uniformly continuous) centered Gaussian process* G *and a pseudometric* ρ *on* $\mathscr{F}$ *so that*

(ii) $\rho \vee e_P \leqq d_G(L_2 d_G^{-1})^{1/2}$

and

(iii) *for every* $\varepsilon > 0$ *and* $f \in \mathscr{F}$

$$\Lambda^P_{2,\infty}([\sup_{g \in B_\rho(f,\varepsilon)} |f-g|]^*) \leqq \varepsilon.$$

Then $\mathscr{F} \in \mathrm{BLIL}(P)$ *(resp.* $\mathscr{F} \in \mathrm{CLIL}(P)$*).*

Proof. By Lemma 2.7 there exists a subprobability measure μ on $T = \{\pi_q f\colon q \in N, f \in \mathscr{F}\} \subset \mathscr{F}$ with $\#(\pi_q \mathscr{F}) < \infty$ such that

$$e_P(f, \pi_q f) \leqq 2^{-q}, \qquad \rho(f, \pi_q f) \leqq 2^{-q} \tag{5.1}$$

and

$$\lim_{r\to\infty} \sup_{f \in \mathscr{F}} \sum_{q=r}^{\infty} 2^{-q}(Lq)^{-1/2}[\ln(\mu\{\pi_q f\})^{-1}]^{1/2} = 0 \tag{5.2}$$

if G has bounded and d_G-uniformly continuous sample paths, or

$$\sup_{f \in \mathscr{F}} \sum_{q=1}^{\infty} 2^{-q}(Lq)^{-1/2}[\ln(\mu\{\pi_q f\})^{-1}]^{1/2} < \infty \tag{5.3}$$

if G has only bounded sample paths. Without loss of generality, $\sup Pf^2 \leqq 1$, ρ-diam $\mathscr{F} \leqq 1$ and $\pi_0 f = \pi_{-1} f = 0$. Let

$$t_q f := \{\pi_r f\}_{r=1}^q, \qquad T_q := t_q \mathscr{F}, \tag{5.4}$$

$$\bar{\gamma}_q f := \bar{\gamma}_q(t_q f) := \left[\ln\left(2^q \Big/ \prod_{r=1}^q \mu\{\pi_r f\}\right)\right]^{1/2}, \tag{5.5}$$

$$\gamma_q f := \gamma_q(t_q f) := (Lq)^{-1/2}\,\bar{\gamma}_q(f), \tag{5.6}$$

$$\beta_{q_0} := \sup_{f \in \mathscr{F}} \sum_{q \geqq q_0} 2^{-q}\gamma_q f, \tag{5.7}$$

$$\Delta_q f := \Delta_q(t_q f) := \min_{r \leqq q}[\sup_{g \in B_\rho(\pi_r f, 3\cdot 2^{-r})} |g - \pi_r f|]^*. \tag{5.8}$$

Obviously

$$2^{-1} q^{1/2} \leqq \bar{\gamma}_q(f)\uparrow \text{ and } \Delta_q f \downarrow \text{ as } q\uparrow. \tag{5.9}$$

By a slight variation of Lemma 2.4 we have

(5.10) $$\beta_{q_0}\downarrow 0,\ \beta_{q_0}<\infty \quad \text{for all } q_0\geqq 1 \text{ if (5.2) holds}$$

or

(5.11) $$\beta_1<\infty \quad \text{if (5.3) holds}$$

Also, and by Lemma 2.3 (A), for all $q_0\geqq 1$, $r_q>1$

(5.12) $$\sum_{t\in T_q} \exp(-r_q\,\bar{\gamma}_q(t)^2)\leqq 2^{-qr_q}$$

By Theorem 2.13 and Proposition 2.14, in order to prove the compact LIL it suffices to show that for all $\alpha>0$ there exists $q\geqq 1$ such that

(5.13) $$\sum_{k=0}^{\infty} \Pr^*\{\|\sum_{i\in I_k} \varepsilon_i(f-\pi_{q_0}f)(X_i)\|_{\mathscr{F}}>\alpha a_{2^k}\}<\infty$$

for all $q_0\geqq q$ (and for the bounded LIL it suffices to show that there exists $\alpha<\infty$ such that

(5.13)′ $$\sum_{k=0}^{\infty} \Pr^*\{\|\sum_{i\in I_k} \varepsilon_i f(X_i)\|_{\mathscr{F}}>\alpha a_{2^k}\}<\infty).$$

Define

$$A_{q,k}f:=\{\Delta_q f\geqq(2^k/Lk)^{1/2}\,2^{-q-1}/\gamma_{q+1}f)\}.$$

For $q_0>1$ we set

(5.14) $$f-\pi_{q_0}f=(f-\pi_{q_0}f)\,I_{A_{q_0,k}f}+(f-\pi_{q_0}f)\,I_{(A_{q_0,k}f)^c}:=I+II$$

Estimates for I. First we note that, by the integrability condition (i), there is $\delta_n\to 0$, $\delta_n\leqq 1/2$, such that $\delta_n a_n\uparrow\infty$ and

(5.15) $$\sum_{n\geqq 1} P\{F^*>\delta_n a_n\}<\infty.$$

Therefore also

(5.16) $$n\cdot P\{F^*>\delta_n a_n\}\to 0$$

and

(5.17) $$\sum_{k=1}^{\infty} 2^k P\{F^*>\delta_{2^k} a_{2^k}\}<\infty.$$

Hence, for any $\alpha>0$

(5.18) $$\sum_{k=1}^{\infty} \Pr^*\{\|\sum_{i\in I_k} \varepsilon_i(f-\pi_{q_0}f)\,I_{A_{q_0,k}f\cap[F^*>\delta_{2^k}a_{2^k}]}(X_i)\|_{\mathscr{F}}>\alpha a_{2^k}\}$$
$$\leqq \sum_{k=1}^{\infty} 2^k P\{F^*>\delta_{2^k} a_{2^k}\}<\infty.$$

Next we consider the truncated sum. It is easy to check that for a nonnegative random variable ξ, $E\xi I_{[\xi>a]} \leqq \Lambda_{2,\infty}^2(\xi)/a$ for $a>0$ so that by (5.8) and (iii)

$$(5.19) \qquad P\Delta_{q_0} f I_{A_{q_0,k}f} \leqq 9\cdot 2^{-q_0+1}(\gamma_{q_0+1} f)(2^{-k}Lk)^{1/2}.$$

Therefore, using (5.19), centering and symmetrization (e.g., Giné and Zinn [1984], 2.7(b) with first instead of second moment),

$$(5.20) \qquad \sum_{k=0}^{\infty} \Pr^*\{\|\sum_{i\in I_k} \varepsilon_i(f-\pi_{q_0} f) I_{A_{q_0,k}f\cap[F^*\leqq\delta_{2^k}a_{2^k}]}(X_i)\|_{\mathscr{F}} > \alpha a_{2^k}\}$$

$$\leqq \sum_{k=0}^{\infty} \Pr\{\|\sum_{i\in I_k} \Delta_{q_0} f I_{A_{q_0,k}f\cap[F^*\leqq\delta_{2^k}a_{2^k}]}(X_i)\|_{\mathscr{F}} > \alpha a_{2^k}\}$$

$$\leqq 4\cdot\sum_{k=0}^{\infty} \Pr\{\|\sum_{i\in I_k} \varepsilon_i \Delta_{q_0} f I_{A_{q_0,k}f\cap[F^*\leqq\delta_{2^k}a_{2^k}]}(X_i)\|_{\mathscr{F}}$$

$$\geqq \tfrac{1}{2}[\alpha - 27\sup_{f\in\mathscr{F}} 2^{-q_0+1}\gamma_{q_0+1} f]\, a_{2^k}\}$$

$$\leqq 4\cdot\sum_{k=0}^{\infty} \Pr\{\|\sum_{i\in I_k} \varepsilon_i \Delta_{q_0} f I_{A_{q_0,k}f\cap[F^*\leqq\delta_{2^k}a_{2^k}]}(X_i)\|_{\mathscr{F}} > \tfrac{1}{4}\alpha a_{2^k}\}$$

provided $\alpha \geqq 54 \sup_{f\in\mathscr{F}} 2^{-q_0+1}\gamma_{q_0+1} f$.

In order to prove that the last sum in (5.20) is finite we will apply Proposition 2.15. To check the hypotheses of Proposition 2.15, we proceed as in Goodman, Kuelbs and Zinn (1981) (in short, GKZ) pp. 727–728. We define for $q_0\in\mathbb{N}$, $k\geqq 1$ and $j\in I_k$

$$(5.21) \qquad Z_j^{q_0}(f) = 2^k a_{2^k}^{-1}\varepsilon_j \Delta_{q_0} f I_{A_{q_0,k}f\cap[F^*\leqq\delta_{2^k}a_{2^k}]}(X_j).$$

Fix $q_0\in\mathbb{N}$. The $\ell^\infty(\mathscr{F})$-valued random variables $Z_j^{q_0}$, $j\in\mathbb{N}$ (these are true random variables since $\#\{\Delta_{q_0} f I_{A_{q_0 k}f}\}<\infty$) satisfy condition (2.10) (condition (4.11) in GKZ), actually, $j^{-1}\|Z_j\|\leqq 4\delta_{2^k}$ for $j\in I_k$. Now

$$\|Z_j^{q_0}(f)\|_{\mathscr{F}} \leqq 2^k a_{2^k}^{-1}\|\Delta_{q_0} f\|_{\mathscr{F}}\, I_{[\lambda_{q_0}(2^k/Lk)^{1/2}\leqq\|\Delta_{q_0} f\|_{\mathscr{F}}\leqq a_{2^k}]}$$

where $\lambda_{q_0} = 2^{-2(q_0+1)}$ because $\delta_n \leqq 1/2$, $\|\Delta_{q_0} f\|_{\mathscr{F}} \leqq 2F^*$, and we may assume $\gamma_{q_0+1}(f) \leqq 2^{q_0+1}$ by (5.7) and (5.10)). Since $\sup_{t>0} t^2 P\{\|\Delta_{q_0} f\|_{\mathscr{F}} > t\}$ is finite by hypothesis (iii), the proof on pages 727, 728 in GKZ applies almost verbatim (the only differences are that we have slightly different bounds – α_n and β_n in GKZ – and that D in (4.17) of GKZ has an extra summand of lower order, actually bounded, as is easy to check) to obtain condition (2.11) of Lemma 2.15 (condition (4.13) in GKZ). To prove (2.12) of Proposition 2.15 we note that by (5.19)

$$(5.22) \qquad P^2 \Delta_{q_0} f I_{A_{q_0,k}f\cap[F^*\leqq\delta_{2^k}a_{2^k}]} \leqq 2\delta_{2^k} a_{2^k} P\Delta_{q_0} f I_{A_{q_0,k}f}$$

$$\leqq 9\cdot 2^{-q_0+2}\gamma_{q_0+1} f\delta_{2^k} Lk.$$

Hence

$$\sum_{j=1}^{m} P^2 Z_j^{q_0} f \leqq \varepsilon_{q_0} \sum_{j=1}^{m} 2^{k_j} \delta_{2^{k_j}} \quad \text{for all } f \in \mathscr{F}$$

where $\varepsilon_{q_0} = 9 \cdot 2^{-q_0+2} \sup_{f \in \mathscr{F}} \gamma_{q_0+1} f < \infty$ (by (5.2) or (5.3)) where k_j is defined by the inequalities $2^{k_j-1} < j \leqq 2^{k_j}$. Obviously

$$\bar{\delta}_m := m^{-2} \sum_{j=1}^{m} 2^{k_j} \delta_{2^{k_j}} \to 0 \quad \text{as} \quad m \to \infty$$

since $\delta_m \to 0$. Hence by Chebyshev's inequality

$$\lim_m \Pr\left\{\left\| \sum_{j=1}^{m} Z_j^{q_0} f \right\|_{\mathscr{F}} > m\beta\right\} \leqq (\# T_{q_0}) \limsup_m \sup_{f \in \mathscr{F}} \Pr\left\{\left| \sum_{j=1}^{m} Z_j^{q_0} f \right| > m\beta\right\}$$

$$\leqq (\# T_{q_0})\, \varepsilon_{q_0} \beta^{-2} \lim_m \bar{\delta}_m = 0$$

for all $q_0 \in \mathbb{N}$ and all $\beta > 0$, i.e., (2.12) is proved. So, Proposition 2.15 shows that

$$\sum_{k=0}^{\infty} \Pr\{\| \sum_{j \in I_k} Z_j^{q_0} f \|_{\mathscr{F}} > \alpha 2^k\} < \infty$$

for all $\alpha > 0$, which by the definition (5.21) of $Z_j^{q_0}$ and the inequalities (5.18) and (5.20) gives

$$\sum_{k=0}^{\infty} \Pr^*\{\| \sum_{i \in I_k} \varepsilon_i (f - \pi_{q_0} f) I_{A_{q_0,k} f}(X_i) \|_{\mathscr{F}} > \alpha a_{2^k}\} < \infty \tag{5.23}$$

for $\alpha \geqq 108 \cdot 2^{-q_0+1} \sup_{f \in \mathscr{F}} \gamma_{q_0+1} f$.

Estimation of II. Let $\mathscr{F}_k = \{f \cdot 2^{-k/2} : f \in \mathscr{F}\}$ and let for $\bar{f} = f \cdot 2^{-k/2} \in \mathscr{F}_k$ and $q \geqq 1$

$$\gamma_q \bar{f} = (Lk)^{1/2} \gamma_q f, \quad t_q \bar{f} = t_q f, \quad \pi_q \bar{f} = 2^{-k/2} \pi_q f \quad \text{and}$$
$$\Delta_q^i \bar{f} = 2^{-k/2} \Delta_q f \quad \text{for } i \in I_k.$$

Then

$$\Pr^*\{\| \sum_{i \in I_k} \varepsilon_i (f - \pi_{q_0} f) I_{(A_{q_0,k} f)^c}(X_i) \|_{\mathscr{F}} > \alpha a_{2^k}\} \tag{5.24}$$
$$= \Pr^*\{\| \sum_{i \in I_k} \varepsilon_i (\bar{f} - \pi_{q_0} \bar{f}) I_{[\Delta_{q_0}^i \bar{f} \leqq 2^{-q_0-1}/\gamma_{q_0+1} \bar{f}]}(X_i) \|_{\mathscr{F}_k} > \alpha (Lk)^{1/2}\}.$$

We will use Theorem 3.1 to estimate the last quantity. Note that (3.1), (3.2), (3.5) and (3.6) are fulfilled and that

$$\sum_{i \in I_k} E((\pi_q \bar{f} - \pi_{q-1} \bar{f})(X_i))^2 \leqq 9 \cdot 2^{-2q}, \quad \bar{f} \in \mathscr{F}_k,$$

$$\sup_{t > 0} t^2 \sum_{i \in I_k} \Pr\{\Delta_q^i \bar{f}(X_i) > t\} \leqq 9 \cdot 2^{-2q}, \quad \bar{f} \in \mathscr{F}_k$$

by (5.1), (5.8) and (iii). Let $q_1 = k > q_0$ and

$$\alpha \geqq (108 e^2 \beta_{q_0}) \vee (36 \cdot 2^{-q_0/2}). \tag{5.25}$$

Then $\frac{1}{12}(Lk)^{1/2} \alpha \geqq (9 e^2 (Lk)^{1/2} \beta_{q_0}) \vee (3(Lk)^{1/2} 2^{k/2} 2^{-q_1})$. So, by Theorem 3.1, using (5.6) and (5.7),

$$\Pr{}^* \{ \| \sum_{i \in I_k} \varepsilon_i (\bar{f} - \pi_{q_0} \bar{f}) I_{[\Delta^i_q \bar{f} \leqq 2^{-q_0 - 1}/\gamma_{q_0+1} \bar{f})]}(X_i) \|_{\mathscr{F}_k} > \alpha (Lk)^{1/2} \}$$
$$\leqq 3 \cdot \sum_{q=q_0+1}^{k} \sum_{t \in T_q} \exp(-\tfrac{1}{56} \alpha (Lk)(Lq)^{-1} \beta_{q_0}^{-1} \bar{\gamma}_q(t)^2) \tag{5.26}$$

Now (5.24) together with (5.26) and (5.12) give that, for α satisfying (5.25) (which implies that $\frac{1}{56} \alpha \beta_{q_0}^{-1} > 2$),

$$\sum_{k \geqq q_0} \Pr{}^* \{ \| \sum_{i \in I_k} \varepsilon_i (f - \pi_{q_0} f) I_{(A_{q_0, k} f)^c}(X_i) \|_{\mathscr{F}} > \alpha a_{2^k} \}$$
$$\leqq 6 \cdot \sum_{k \geqq q_0} \sum_{q=q_0+1}^{k} 2^{-\left(\frac{1}{56} \alpha (Lk)(Lq)^{-1} \beta_{q_0}^{-1}\right) q}$$
$$\leqq 6 \cdot \sum_{k \geqq q_0} k \cdot k^{-\frac{1}{56} \alpha \beta_{q_0}^{-1} \ln 2} < \infty. \tag{5.27}$$

Let us now prove the compact LIL. Let $\alpha > 0$ and choose q such that $\frac{1}{2}\alpha \geqq (108 e^2 \beta_q) \vee (36 \cdot 2^{-q/2})$, which is possible by (5.10). Then (by (5.23), (5.27), (5.10) and (5.7)) (5.13) holds for all $q_0 \geqq q$. The bounded LIL is proved by letting $q_0 = 1$ and choosing α sufficiently large. □

We omit here the "bracketing" version of Theorem 5.1 as it is completely analogous to Theorem 4.4′. In view of Lemma 2.7, Theorem 5.1 contains Yukich's (1986, Theorem 7) LIL under bracketing.

Our result was inspired by Ledoux's (1982) LIL for Lipschitzian processes. The following, which contains Ledoux's theorem (see Lemma 2.7), is an immediate consequence of Theorem 5.1′ (5.1).

5.2. Theorem. *Let T be an index set and let $\{X(t)\}_{t \in T}$ be a centered stochastic process on T such that $EX^2(t) < \infty$ for all $t \in T$, satisfying*

(i) $E^*(\|X\|_\infty^2 / L_2 \|X\|_\infty) < \infty$,
(ii) $d_X \leqq d_G (L_2 d_G^{-1})^{1/2}$

for some bounded (bounded and d_G-uniformly continuous) centered Gaussian process G on T. Assume further that there exist $M \in \Lambda_{2,\infty}(\Omega, \Sigma, P)$ and a pseudometric ρ on T satisfying

(ii)′ $\rho \leqq d_G (L_2 d_G^{-1})^{1/2}$

such that

(iii) $|X(t, \omega) - X(s, \omega)| \leqq M(\omega) \rho(s, t), \quad \omega \in \Omega, \ s, t \in T.$

Then $X \in$ BLIL ($X \in$ CLIL) as a $C_u(T, \rho)$-valued random variable.

Finally we will show that even the condition on ρ in Theorem 5.1 is sharp, at least in a particular instance.

5.3. Example. Let $\{\xi_n\}_{n\in\mathbb{N}}$ be random variables such that $E\xi_n^2\to 0$, let $\eta_n=\xi_n(L_3 n)^{1/2}/(Ln)^{1/2}$, $n\in\mathbb{N}$, and $F=\sup_{n\in\mathbb{N}}|\eta_n|$. Assume $(L_3 n)^{-1/2}\cdot(LN)^{1/2}\Lambda_{2,\infty}(\sup_{j\geqq n}|\eta_j|)\to 0$ as $n\to\infty$, and $EF^2/L_2F<\infty$. Then the c_0-valued random variable $X=(\eta_n:n\in\mathbb{N})$ satisfies the CLIL (the BLIL if the two zero limits are replaced by uniform boundedness). This follows from Theorem 5.1 exactly in the same way as the CLT is proved in Example 4.7 from Theorem 4.4. Kuelbs (1976) proves the CLIL for ess sup $|\xi_k|=(Lk)^{-1/2}$. We thus have additional examples of $X\in$CLIL, $X\notin$CLT, ess sup $\|X\|<\infty$. This example contains Theorem 2 of Heinkel (1983).

Example 5.3 is sharp:

5.4. Proposition. *Let $\{\varepsilon_k\}_{k\in\mathbb{N}}$ be a Rademacher sequence. The c_0-valued random variables $X=\{(L_3 k)^{1/2}(Lk)^{-1/2}\varepsilon_k\}_{k\in\mathbb{N}}$ does not satisfy the* CLIL. *If $c_k\to\infty$ the c_0-valued random variable $X=\{c_k(L_3 k)^{1/2}(Lk)^{-1/2}\varepsilon_k\}_{k\in\mathbb{N}}$ does not satisfy the* BLIL.

Proof. Let $b_k=(L_3 k)^{1/2}(Lk)^{-1/2}$ and $X_i=\{b_k\varepsilon_{ki}\}_{k\in\mathbb{N}}$, where $\{\varepsilon_{ki}: i\, k\in\mathbb{N}\}$ is a (double) Rademacher sequence. By Kuelbs (1977), in order to prove the first statement, it is enough to show $a_n^{-1}\sum_{i=1}^n X_i\not\to 0$ in probability. We have

$$P\left\{\left\|\sum_{i=1}^n X_i\right\|>a_n\delta\right\}=P\left\{\sup_k b_k\left|\sum_{i=1}^n \varepsilon_{ki}\right|>\delta a_n\right\}$$
$$\geqq\left(\sum_{k=1}^\infty P\left\{b_k\left|\sum_{i=1}^n\varepsilon_{ki}\right|>\delta a_n\right\}\right)\left(1+\sum_{k=1}^\infty P\left\{b_k\left|\sum_{i=1}^n\varepsilon_{ki}\right|>\delta a_n\right\}\right)^{-1}.$$

Since $x(1+x)^{-1}$ is decreasing, it suffices to show that the numerator of this fraction does not tend to zero. For this we use Kolmogorov's exponential lower bound (Kolmogorov [1929]; see e.g., Stout [1974), p. 262, Theorem 5.2.2 (iii)), which specialized to our case is as follows: for every $\gamma>0$ there exist positive constants $\alpha(\gamma)$ and $\beta(\gamma)$ such that if $n^{1/2}\beta(\gamma)\geqq\alpha\geqq\alpha(\gamma)$ then

$$P\left\{\sum_{i=1}^n\varepsilon_i>n^{1/2}\alpha\right\}\geqq\exp\left(-\tfrac{1}{2}\alpha^2(1+\gamma)\right).$$

We choose $\gamma>0$ and $\delta\leqq(1+\gamma)^{-1/2}$. Then we let $\alpha_n=(L_2 n)^{1/2}\delta$ and $a=\frac{1}{2}\beta^2(\gamma)(1+\gamma)$. Choose $n_0(\gamma)$ such that $b_k^{-1}\alpha_n\geqq\alpha_n\geqq\alpha(\gamma)$ and $L_2(n)/L_2(an)\leqq 2$ for $n\geqq n_0(\gamma)$. Hence, for $k\leqq\exp(an)$, $b_k^{-1}\alpha_n n^{-1/2}\leqq(2a)^{1/2}\delta\leqq\beta(\gamma)$ for $n\geqq n_0(\gamma)$. Therefore, for $n\geqq n_0(\gamma)$ and $k\leqq\exp(an)$, Kolmogorov's inequality gives

$$P\left\{\sum_{i=1}^n\varepsilon_{ki}>b_k^{-1}\delta a_n\right\}=P\left\{\sum_{i=1}^n\varepsilon_{ki}>b_k^{-1}\alpha_n n^{1/2}\right\}$$
$$\geqq\exp\left(-\tfrac{1}{2}b_k^{-2}\alpha_n^2(1+\gamma)\right)\geqq\exp\left(-(Lk)(L_2 n)(2L_3 k)^{-1}\right)$$
$$\geqq\exp\left(-an(L_2 n)(2L_2(an))^{-1}\right)\geqq\exp\{-an\}.$$

Hence

$$\sum_{k=1}^\infty P\left\{b_k\left|\sum_{i=1}^n\varepsilon_{ki}\right|>\delta a_n\right\}\geqq\sum_{k\leqq\exp(an)}P\left\{\sum_{i=1}^n\varepsilon_{k_i}>b_k^{-1}\delta a_n\right\}\geqq 1.$$

This shows that for $\delta \leqq (1+\gamma)^{-1/2}$, $P\left\{\left\|\sum_{i=1}^{n} X_i\right\| > a_n \delta\right\}$ is bounded from below by 1/2 i.e., $a_n^{-1} \sum_{i=1}^{n} X_i \nrightarrow 0$ in probability. The proof of $\bar{X} \notin$ BLIL is entirely similar (take $\delta = c_k(1+\gamma)^{-1/2}$) and is omitted.

For an example in the same direction see Yukich (1986), Theorem 8.

5.5. *Remark.* M. Ledoux and M. Talagrand have recently informed us that they can improve upon Theorem 5.1 (and hence its corollary) by relaxing condition (ii), at least in the separable case. Their result contains Borisov's (1985) whereas ours doesn't. We would like to thank M. Lacey for bringing Borisov's paper to our attention.

References

1. Alexander, K.: Probability inqualities for empirical processes and a law of the iterated logarithm. Ann. Probab. **12**, 1041–1067 (1984)
2. Alexander, K.: The central limit theorem for empirical processes on Vapnik-Červonenkis classes. Ann. Probab. **15**, 178–203 (1987)
3. Andersen, N.T.: The calculus of non-measurable functions and sets. Aarhus Univ., Math. Inst. Various Publ. Series No. 36 (1985)
4. Andersen, N.T.: Lectures on Talagrand's result about continuous Gaussian processes. Texas A & M preprint (1986)
5. Andersen, N.T., Dobrić, V.: The central limit theorem for stochastic processes. Ann. Probab. **15**, 164–177 (1987)
6. Reference deleted
7. Bennett, G.W.: Probability inqualitites for the sum of bounded random variables. J. Amer. Stat. Assoc. **57**, 33–45 (1962)
8. Borisov, I.S.: Exponential inequalities for the distributions of sums of independent random fields. Sib. Math. J. **26**, 12–22 (1985)
9. Csörgö, M., Csörgö, S, Horvath, L., Mason, D.L.: Weighted empirical and quantile processes. Ann. Probab. **14**, 31–85 (1986)
10. Dudley, R.M.: Central limit theorems for empirical processes. Ann. Probab. **6**, 899–929 (1978)
11. Dudley, R.M.: A course on empirical processes. (Lect. Notes Math. Vol. 1097, pp. 2–142) Berlin Heidelberg New York: Springer 1984
12. Dudley, R.M.: An extended Wichura theorem, definitions of Donsker class, and weighted empirical distributions. (Lect. Notes Math. Vol. 1153, pp. 141–178) Berlin Heidelberg New York: Springer 1985
13. Dudley, R.M., Kuelbs, J.: Log log laws for empirical measures. Ann. Probab. **8**, 405–418 (1980)
14. Dudley, R.M., Philipp, W.: Invariance principles for sums of Banach space valued random elements and empirical processes. Z. Wahrscheinlichkeitstheor. verw. Geb. **62**, 509–552 (1983)
15. Fernique, X.: Regularité des trajectoires des fonctions aléatoires gaussiennes. (Lect. Notes Math. Vol. 480, pp. 1–96) Berlin Heidelberg New York: Springer 1974
16. Giné, E., Zinn, J.: Some limit theorems for empirical processes. Ann. Probab. **12**, 929–989 (1984)
17. Giné, E., Zinn, J.: Lectures on the central limit theorem for empirical processes. (Lect. Notes Math. vol. 1221, pp. 50–113) Berlin, Heidelberg, New York: Springer 1986
18. Goodman, V., Kuelbs, J., Zinn, J.: Some results on the LIL in Banach space with applications to weighted empirical processes. Ann. Probab. **9**, 713–752 (1981)
19. Heinkel, B.: Mésures majorantes et théorème de la limite centrale dans $C(S)$. Z. Wahrscheinlichkeitstheor. verw. Geb. **38**, 339–351 (1977)
20. Heinkel, B.: Relation entre théorème central limite et loi du logarithme iteré dans les espaces de Banach. Z. Wahrscheinlichkeitstheor. verw. Geb. **49**, 211–220 (1979)
21. Heinkel, B.: Majorizing measures and limit theorems for c_0-valued random variables. (Lect. Notes Math. Vol. 990, pp. 136–149) Berlin Heidelberg New York: Springer 1983

22. Jain, N., Marcus, M.B.: Central limit theorems for $C(S)$-valued random variables. J. Funct. Anal. **19**, 216–231 (1975)
23. Kolmogorov, A.: Über das Gesetz des iterierten Logarithmus. Math. Ann. **101**, 126–135 (1929)
24. Kuelbs, J.: A counterexample for Banach space valued random variables. Ann. Probab. **4**, 684–689 (1976)
25. Kuelbs, J.: Kolmogorov law of the iterated logarithm for Banach space valued random variables. Ill. J. Math. **21**, 784–800 (1977)
26. Kuelbs, J., Zinn, J.: Some stability results for vector and valued random variables. Ann. Probab. **7**, 75–84 (1979)
27. Ledoux, M.: Loi du logarithm itéré dans $C(S)$ et fontion caractéristique empirique. Z. Wahrscheinlichkeitstheor. verw. Geb. **60**, 425–435 (1982)
28. Marcus, M.B., Pisier, G.: Characterization of a.s. continuous ρ-stable random Fourier series and strongly stationary processes. Acta Math. **152**, 245–301 (1984)
29. Marcus, M.B., Pisier, G.: Some results on the continuity of stable processes and the domain of attraction of continuous stable processes. Ann. Inst. Henri Poincaré Nouv. Ser., Sect. B **20**, 177–199 (1984)
30. Marcus, M.B., Zinn, J.: The bounded law of the iterated logarithm for the weighted empirical distribution process in the non-i.i.d. case. Ann. Probab. **12**, 335–360 (1984)
31. Ossiander, M.: A central limit theorem under metric entropy with L_2-bracketing. Ann. Probab. **15**, 897–919 (1987)
32. Paulauskas, V.: On the central limit theorem in the Banach space c_0. Soviet Math., Dokl. **22–2**, 349–351 (1980)
33. Pisier, G.: Le Théorème limite centrale et la loi du logarithme itéré dans les espaces de Banach. Séminaire Maurey-Schwartz 1975–76, Exp. 3 and 4. Paris: Ecole Polytechnique 1975
34. Pollard, D.: A central limit theorem for empirical processes. J. Aust. Math. Soc., Ser. A **33**, 235–248 (1982)
35. Preston, C.: Continuity properties of some Gaussian processes. Ann. Math. Stat. **43**, 285–292 (1972)
36. Rhee, W.T.: Central limit theorem and increment conditions. Stat. Prob. Letters, **4**, 191–195 (1986)
37. Stout, W.F.: Almost sure convergence. New York: Academic Press, 1974
38. Talagrand, M.: Donsker classes and random geometry, Ann. Probab. **15**, 1327–1338 (1987)
39. Talagrand, M.: Regularity of Gaussian processes, Acta Math. **159**, 99–149 (1987)
40. Reference deleted

Received August 14, 1986; in revised form July 7, 1987