The Annals of Probability
1984, Vol. 12, No. 1, 13–34

FUNCTIONAL LAW OF THE ITERATED LOGARITHM AND UNIFORM CENTRAL LIMIT THEOREM FOR PARTIAL-SUM PROCESSES INDEXED BY SETS.

BY RICHARD F. BASS AND RONALD PYKE[1]
University of Washington

Let $\{X_{\mathbf{j}}: \mathbf{j} \in J^d\}$ be an array of independent random variables, where J^d denotes the d-dimensional positive integer lattice. The main purpose of this paper is to obtain a functional law of the iterated logarithm (LIL) for suitably normalized and smoothed versions of the partial-sum process $S(B) = \sum_{\mathbf{j}\in B} X_{\mathbf{j}}$. The method of proof involves the definition of a set-indexed Brownian process, and the embedding of the partial-sum process in this Brownian process. In addition, the LIL is derived for this Brownian process. The method is extended to yield a uniform central limit theorem for the partial-sum process.

1. Introduction. This paper focuses upon the asymptotic properties of partial-sum processes indexed by sets in Euclidean spaces. These processes are determined by an array of random variables (r.v.'s) $\{X_{\mathbf{j}}: \mathbf{j} \in J^d\}$ where J^d denotes the d-dimensional positive integer lattice. We view $X_{\mathbf{j}}$ as a random mass attached to the grid point $\mathbf{j}$. For any subset $B \subset R^d$, one may then define the partial-sum

$$S(B) = \sum_{\mathbf{j}\in B} X_{\mathbf{j}} \tag{1.1}$$

to represent the random measure of the region B. We assume throughout that $EX_{\mathbf{j}} = 0$ and $\text{var}(X_{\mathbf{j}}) = 1$ for every $\mathbf{j} \in J^d$ and that the $X_{\mathbf{j}}$ are independent and identically distributed.

The main purpose of the paper is to obtain a functional law of the iterated logarithm (LIL) for suitably normalized and smoothed versions of S. The technique used involves the construction of an equivalent version of $\{X_{\mathbf{j}}: \mathbf{j} \in J^d\}$ that is embedded in a suitable Brownian process. The general approach is then analogous to that introduced by Strassen (1964) for his proof of the first functional LIL for sequences of independent and identically distributed (IID) r.v.'s. The methods are then extended to obtain a uniform central limit theorem.

As motivation and potential application for our results, consider the situation where a sample is taken over an area (e.g., of insect larvae in a forest, of mineral deposits, of cells in tissue, etc.). What can one say about the properties of the sample if the area is large? That is, if $\{A_n\}$ is a (not necessarily nested) sequence of sets whose areas are increasing and $S(A_n)$ are the sample statistics, one would expect, under appropriate independence assumptions and normalizations, that the $S(A_n)$'s should satisfy a strong law of large numbers (SLLN), a central limit

Received October 1982; revised July 1983.
[1] This work was supported in part by the National Science Foundation, Grants MCS-83-00581 and MCS-82-02861.
AMS 1970 *subject classifications.* Primary, 60F15; secondary, 60B10, 60J65.
Key words and phrases. Functional law of the iterated logarithm, partial-sum processes indexed by sets, central limit theorem, embedding by stopping times, invariance principles.

theorem (CLT), and also a law of the iterated logarithm (LIL). If the sequence of sets $\{n^{-1}A_n\}$ is contained in a family $\mathscr{A}$ satisfying certain assumptions, then the results of this paper will provide the desired LIL and CLT. The SLLN, which requires completely different assumptions and methods, is derived in Bass and Pyke (1984).

By (1.1), S is defined for any subset of R^d. However, it will be necessary to restrict the domain of S in order to obtain the "uniform" results of this paper. It will also be necessary, due to the largeness of the index families (cf. Pyke (1983) and remark number 7 in Section 8 below), to work with a smoothed version of S, namely X, that is defined by

$$X(B) = \textstyle\sum_{\mathbf{j}\in J^d} |B \cap C_\mathbf{j}| X_\mathbf{j} \tag{1.2}$$

where $C_\mathbf{j}$ denotes the unit cube $(\mathbf{j} - \mathbf{1}, \mathbf{j}]$, and $|\cdot|$ and $\lambda(\cdot)$ are used interchangeably as convenient to denote Lebesgue measure. Thus X is a signed measure which is absolutely continuous with respect to Lebesgue measure λ, and for which the density $dX/d\lambda$ equals $X_\mathbf{j}$ on $C_\mathbf{j}$. We view $X(B)$ as the random measure of B that is obtained when the random mass $X_\mathbf{j}$ is uniformly spread over the cube $C_\mathbf{j}$, rather than being attached to the point $\mathbf{j}$ as it is in the definition of $S(B)$.

Before discussing further the scope of the results and the related literature, the following notation is needed. Let $\mathscr{A}$ denote a family of subsets of the unit d-dimensional cube $I^d = [0, 1]^d$. Define

$$L_2 n = \log \log n, \quad b_n = (2L_2 n)^{1/2}.$$

All logarithms are to base e, and we assume $n \geq 3$ to insure proper definition. For any set A and real x, let $xA = \{xa\colon a \in A\}$. Define $\{S_n(A)\colon A \in \mathscr{A}\}$ and $\{X_n(A)\colon A \in \mathscr{A}\}$ by

$$S_n(A) = n^{-d/2} b_n^{-1} S(nA), \quad X_n(A) = n^{-d/2} b_n^{-1} X(nA). \tag{1.3}$$

Let Z be a Brownian process indexed by a family $\mathscr{B}$ of Lebesgue measurable subsets of R^d. That is, for each $B \in \mathscr{B}$, $Z(B)$ is a mean zero normal r.v. with variance $|B|$ and

$$\operatorname{cov}(Z(B), Z(C)) = |B \cap C|, \quad \text{for} \quad B, C \in \mathscr{B}. \tag{1.4}$$

In order for the functional LIL to be formulated, it is necessary that $\mathscr{B}$ be a large family of subsets of $[0, \infty)^d$, determined by $\mathscr{A}$, on which it is possible for Z to have continuous sample paths with respect to the Lebesgue symmetric difference pseudo-metric. This requires $\mathscr{A}$ to satisfy certain structural properties that are described in Section 2. For now, assume only that the following normalized form of Z is well-defined, namely $\{Z_n(A)\colon A \in \mathscr{A}\}$ where

$$Z_n(A) = n^{-d/2} b_n^{-1} Z(nA). \tag{1.5}$$

In Section 3 we prove the a.s. relative compactness of $\{Z_n\}$ and identify the limit points to be

$$\mathscr{G}:=\left\{G\equiv\int_{(\cdot)}g\,d\lambda:\ \text{Domain}(G)=\mathscr{A},\right.\qquad(1.6)$$
$$\left.g\colon I^d\to R^1 \text{ such that } \int_{I^d}|g|^2\,d\lambda\le 1\right\},$$

the set of all absolutely continuous signed measures with total variation bounded by 1. (The elements $G \in \mathscr{G}$, though restricted here to $\mathscr{A}$, are definable on the family of Lebesgue measurable subsets of I^d.) It will be convenient to identify G and its density g whenever speaking of members of $\mathscr{G}$.

In Section 4, the embedding of the partial-sum process in the Brownian process Z is defined. Many embeddings can be suggested, but the one used here seems to be the simplest to describe and the most tractable. It was proposed by the second author in 1970 for the purpose of obtaining the functional LIL for partial-sum processes indexed by the orthants. However, the moment conditions required to use it were too strong for the orthants. The problem of this functional LIL (cf. Pyke, 1973) was solved by Wichura (1973) using the methods of Hartman and Wintner (1941) and only second moments were required. This embedding of an array of partial sums into Z was also proposed independently by Kiefer (1972) where it is central to his embedding of empirical processes in tied-down Brownian sheets; Kiefer's overall construction for empirical processes is necessarily much more difficult than for partial sums.

In order to use the embedding method for our set-indexed processes, it is necessary to have good tail estimates on the probability of deviations between the embedded process and the Brownian process in which it is embedded. These estimates are provided in Section 5. Some of these results may be of general interest as they make use of exponential bounds for martingales and stochastic integrals. These estimates are then applied in Section 6 to obtain the main functional LIL for the smoothed partial-sum processes.

In Section 7, the embedding methods and estimates of the preceding sections are used to obtain a Uniform CLT for the partial-sum processes. The first such Uniform CLT was derived by Pyke (1983) in which the index families were metrized by the stronger Hausdorff metric. The result obtained in this paper provides an improvement in the moment condition, though the order of the finite moment still increases with the size of the index family at the same rate.

In Section 8, some general remarks, including open problems and directions for future research suggested by our results, are discussed.

2. The index families $\mathscr{A}$ and Brownian process Z. As mentioned in Section 1, the Brownian process $Z:=\{Z(A): A\in\mathscr{B}\}$ is a mean-zero Gaussian process indexed by $\mathscr{B}$, whose covariances are given by (1.4). The index family $\mathscr{B}$ is a collection of subsets of $R_+^d=[0,\infty)^d$. It is necessary for the results of this paper that Z exist as a process whose sample paths are continuous with respect to the Lebesgue symmetric difference pseudometric $d_L(A, B):=|A\,\Delta\,B|$.

Sufficient conditions on an index family are known that will insure continuous sample paths when the index family comprises subsets of a compact set such as I^d. (Cf. Dudley, 1973.) In the case of Brownian Sheet, as in the case of standard Brownian Motion, the extension of a continuous process on I^d to one on R_+^d is straightforward. In the case studied here, the extension is not so obvious and in fact requires that additional structure be imposed.

The main focus in this paper is upon the normalized partial-sum and Brownian processes, X_n and Z_n, defined respectively by (1.3) and (1.5). These processes are random set functions defined on $\mathscr{A}$, a family of subsets of I^d. It is this family $\mathscr{A}$ that is central, and it is upon it that properties must be imposed. These properties will enable Z to be defined continuously on a family of subsets in R_+^d that is sufficiently large for the existence in Z of suitable embedded partial-sum processes; cf. Section 4.

The properties of $\mathscr{A}$ that will be used below are defined as follows:

(i) *Contraction closed*: $A \in \mathscr{A}$ implies $sA \in \mathscr{A}$ for all $0 < s < 1$.

(ii) *Interval closed*: $(\mathbf{s}, \mathbf{t}] \in \mathscr{A}$ for all $\mathbf{s}, \mathbf{t} \in I^d$.

(iii) *Totally bounded*: For every $\delta > 0$, there exists a finite δ-net $\mathscr{A}_\delta \subset \mathscr{A}$ such that for any $A \in \mathscr{A}$ there exists $A_\delta \in \mathscr{A}_\delta$ such that

$$d_L(A, A_\delta) = |A \,\Delta\, A_\delta| \leq \delta.$$

(iv) *Boundary smooth*: There exists a constant $\mathbb{c}$ such that for all $\varepsilon > 0$ and for all $A \in \mathscr{A}$, $|A(\varepsilon)| \leq \mathbb{c}\varepsilon$ where $A(\varepsilon) = A^\varepsilon \backslash A_\varepsilon$ denotes the ε-annulus about the boundary of A with respect to Euclidean distance.

(v) *Entropy integrable*: If $\nu(\delta) := \operatorname{card}(\mathscr{A}_\delta)$ and $H(\delta) := \log \nu(\delta)$, then

$$\int_0^1 \{H(u)/u\}^{1/2}\, du < \infty. \tag{2.1}$$

For the statements of Theorems 6.1 and 7.1 we will place a slightly stronger condition than (2.1) upon H, namely,

$$H(\delta) \leq K\delta^{-r}, \quad 0 < \delta \leq 1 \tag{2.2}$$

for some constants $K > 0$ and $0 < r < 1$. We refer to the resulting stronger property as

(v′) H satisfies (2.2) for $0 < r < 1$.

Assumptions (i) and (ii) may be assumed without loss of generality in the sense that if $\mathscr{A}$ were any family of subsets of I^d that satisfies (iii) and (v), and if $\mathscr{A}^+$ denotes the union of $\mathscr{A}$ with the set of all contractions sA for $0 < s < 1$ and $A \in \mathscr{A}$ and with the set of all intervals, then $\mathscr{A}^+$ also satisfies (iii) and (v). For the latter statement, note that the entropy of $\mathscr{A}^+$, ν^+ say, satisfies

$$\nu^+(\delta) \leq (\delta^{-1} + 1)^d \nu(\delta/2) + (\delta^{-1} + 1)^{2d}$$

since $\delta^{-1} + 1$ bounds the number of points in a Euclidean δ-net of $[0, 1]$ and $d_L(sA, [s/\delta]\delta A_\delta) \leq d_L(sA, sA_\delta) + d_L(sA_\delta, [s/\delta]\delta A_\delta) \leq s^d\delta + \delta \leq 2\delta$. Its logarithm, H^+ say, therefore satisfies (2.1).

The inclusion of the nonrestrictive assumptions (i) and (ii) ensures that our embedded processes (cf. Section 4) will always be well defined. However, as already mentioned, it is necessary to have the Brownian process Z defined over the entire positive orthant R_+^d. This involves additional difficulties that are not present when working with Brownian Sheets which are indexed by points (or equivalently, lower-left orthants) rather than sets. The context here is as follows. The functional LIL involves contracting and rescaling random measures of sets in R_+^d to sets in I^d. Specifically, we require the existence of a Brownian process, Z, that is defined on nA for each integer $n \geq 1$ and each $A \in \mathscr{A}$. For this purpose we must even impose an additional property on $\mathscr{A}$, namely,

(vi) *Origin sparse*: For every $\mathbf{j} \in J^d$, the family

$$\mathscr{A}^{\mathbf{j}} := \{nA \cap (\mathbf{j} + I^d); A \in \mathscr{A}, n \geq 1\}$$

satisfies (iii) and (v).

This property (vi) will be needed for the functional LIL for Z but not for X. Its purpose is to ensure that a continuous Brownian process can be defined on the family of all intersections of the cube $\mathbf{j} + I^d = C_{\mathbf{j}}$ with multiples of $\mathscr{A}$. If $Z^{\mathbf{j}}$ denotes this Brownian process, then the desired process Z that is defined over $\cup_{n=1}^{\infty} n\mathscr{A}$ can be constructed by patching together the $Z^{\mathbf{j}}$'s as follows:

$$Z(nA) := \sum_{\mathbf{j} \leq n\mathbf{1}} Z^{\mathbf{j}}(nA \cap C_{\mathbf{j}}). \tag{2.3}$$

To see that (vi) is indeed a restriction, consider, for example, a sequence $\{\mathscr{B}_n : n \geq 1\}$ of families of closed subsets of I^d such that every element of every $\mathscr{B}_n$ is contained in $I^d \backslash 2^{-1} I^d$ and such that the log-entropy of $\mathscr{B}_n$ is $H_n(\delta) = \delta^{-(1-1/n)}$. Now let

$$\mathscr{A} = \cup_{n=1}^{\infty} 2^{-n} \mathscr{B}_n. \tag{2.4}$$

For this choice of $\mathscr{A}$, it is not true that (vi) is satisfied. In particular, take $\mathbf{j} = \mathbf{0}$. Then

$$\mathscr{A}^{\mathbf{0}} \supset \cup_{n=1}^{\infty} 2^n (2^{-n} \mathscr{B}_n) = \cup_{n=1}^{\infty} \mathscr{B}_n$$

so that neither (iii) nor (v) need be satisfied. To simplify the illustration, one may specialize to $d = 1$ and $\mathscr{B}_n$ equal to the family of all finite unions of the intervals $2^{-1}[1 + (i-1)/n, 1 + i/n]$, $1 \leq i \leq n$. It is straightforward to check that the $\mathscr{A}$ of (2.4) satisfies (iii) and (v) but that $\mathscr{A}^0 = \cup_n \mathscr{B}_n$ does not.

THEOREM 2.1. *If $\mathscr{A}$ is a family of subsets of I^d that possesses Property* (vi), *then there exits a continuous Brownian process $\{Z(B): B \in \mathscr{B}\}$ with respect to d_L where $\mathscr{B} = \{nA : A \in \mathscr{A}, n \geq 1\}$.*

PROOF. By property (vi), the existence of a continuous $Z^{\mathbf{j}}$ on $\mathscr{A}^{\mathbf{j}}$ follows from know results; Dudley (1973), Section 1. The required Z follows from (2.3). □

3. Functional LIL for Z.

Recall the definition of $\mathscr{S}$ given in (1.6) and the construction of the Brownian process Z given in Theorem 2.1. For $f: \mathscr{A} \to R^1$, define the supremum norm $\| f \|_{\mathscr{A}}$ to be $\sup_{A \in \mathscr{A}} | f(A) |$.

THEOREM 3.1. *If $\mathscr{A}$ satisfies properties* (i)–(iii), (v), *and* (vi), *then with probability one*, $\{Z_n: n \geq 3\}$ *is relatively compact with respect to* $\|\cdot\|_{\mathscr{A}}$ *with limit points exactly equal to* $\mathscr{G}$.

PROOF. Let

$$\mathscr{G}^{\varepsilon} = \{f: \mathscr{A} \to R^1 \text{ for which } \|f - G\|_{\mathscr{A}} < \varepsilon \text{ for some } G \in \mathscr{G}\}$$

denote the open ε-sphere about $\mathscr{G}$ with respect to the supremum norm $\|\cdot\|_{\mathscr{A}}$. Note that for any $A \in \mathscr{A}$ and $G \in \mathscr{G}$, Holder's inequality implies that

$$|G(A)| \leq |A|^{1/2}. \tag{3.1}$$

For $\lambda > 1$, set $n_i = [\lambda^i]$, $i \geq 1$. Let Y_i be defined by

$$Y_i(A) = Z(n_i A)/n_i^{d/2}.$$

The Y_i are $\mathscr{C}(\mathscr{A})$-valued random variables that are Gaussian, mean 0, and equal in law to Z. If $\mathscr{B}_{\mathscr{A}}$ is the Borel σ-field with respect to the pseudo-metric d_L, and if μ is a signed measure on $(\mathscr{A}, \mathscr{B}_{\mathscr{A}})$ with total variation M, then

$$E \int Y_i(A)\mu(dA) \int Y_k(B)\mu(dB)$$

$$= \int\int |(n_i A) \cap (n_k B)|\, \mu(dA)\mu(dB)/n_i^{d/2} n_k^{d/2} \leq M^2 |n_i I^d| / n_i^{d/2} n_k^{d/2}$$

$$\leq M^2 n_i^{d/2}/n_k^{d/2} \to 0 \quad \text{as} \quad i \to \infty \quad \text{and} \quad k - i \to \infty.$$

By Theorem 4.1 of Carmona and Kôno (1976), this condition implies that $\{Y_i/\sqrt{2 \log i}, i \geq 3\}$ is relatively compact in $\mathscr{C}(\mathscr{A})$ with limit points exactly $\mathscr{K}$, the closed unit ball in $\mathscr{H}$, the reproducing kernel Hilbert space for Z. Here

$$\mathscr{H} = \{f: f \text{ maps } \mathscr{A} \to R^1 \text{ and for some finite signed measures } \mu$$
$$\text{on } (\mathscr{A}, \mathscr{B}_{\mathscr{A}}),\ f(A) = \int |A \cap B|\, \mu(dB)\}$$

is a Hilbert space when given the inner product

$$\left\langle \int |\cdot \cap B|\, \mu(dB), \int |\cdot \cap C|\, \varphi(dC) \right\rangle = \int\int |B \cap C|\, \mu(dB)\varphi(dC). \tag{3.2}$$

$\mathscr{H}$ has the reproducing property

$$\langle f, R(\cdot, A)\rangle = f(A) \quad \text{for} \quad f \in \mathscr{H}, A \in \mathscr{A},$$

where $R(B, A) = E(Z(B)Z(A)) = |B \cap A|$ is the covariance function of the process Z.

If μ is a finite signed measure on $(\mathscr{A}, \mathscr{B}_{\mathscr{A}})$ and $f(A) = \int |A \cap B| \, \mu(dB)$, set

$$g(\mathbf{t}) = \int 1_B(\mathbf{t})\mu(dB), \quad \mathbf{t} \in I^d. \tag{3.3}$$

Then,

$$\int_A g \, d\lambda = \int \int_A 1_B(\mathbf{t}) \, d\lambda \, \mu(dB) = \int |A \cap B| \, \mu(dB) = f(A). \tag{3.4}$$

Using (3.2),

$$\begin{aligned} \|f\|^2_{\mathscr{H}} &= \int\int |A \cap B| \, \mu(dA)\mu(dB) = \int f(A)\mu(dA) \\ &= \int\int 1_A(\mathbf{t})g(\mathbf{t}) \, d\lambda \, \mu(dA) = \int g^2(\mathbf{t}) \, d\lambda. \end{aligned} \tag{3.5}$$

From (3.4) and (3.5) we conclude $\mathscr{H} \subseteq \mathscr{G}$.

To show $\mathscr{G} \subseteq \mathscr{H}$, it suffices to show, in view of (3.5), that $f(A) = \int_A g \, d\lambda$ is in $\mathscr{H}$ if $\int g^2 d\lambda < \infty$. By standard approximation techniques and linearity, it suffices to consider g of the form $g = 1_C$, where C is an interval $(\mathbf{s}, \mathbf{t}] \in \mathscr{A}$. If μ is the measure assigning unit mass to $\{C\}$, then by (3.4), $\int_A g \, d\lambda = \int |A \cap B| \, \mu(dB) \in \mathscr{H}$.

Since $\log i \sim \log\log n_i$, the above discussion shows that every point of $\mathscr{G}$ is a limit point of $\{Z_{n_i}\}$, and hence of $\{Z_n\}$. Moreover, for any $\varepsilon > 0$, $P[Z_{n_i} \notin \mathscr{G}^\varepsilon \text{ i.o.}] = 0$. Since by construction $Z_n = n^{-d/2}b_n^{-1}Z(n\cdot)$, one obtains that for $n_{i-1} < n \le n_i$

$$Z_n(A) = a_{n,i}Z_{n_i}((n/n_i)A), \quad a_{n,i} = n_i^{d/2}b_{n_i}/n^{d/2}b_n. \tag{3.6}$$

Thus, to show that Z_n does not vary significantly from the terms of the subsequence, consider,

$$\begin{aligned} |Z_n(A) - Z_{n_i}(A)| &= |a_{n,i}[Z_{n_i}((n/n_i)A) - Z_{n_i}(A)] + (a_{n,i} - 1)Z_{n_i}(A)| \\ &\le a_{n,i}|Z_{n_i}((n/n_i)A) - Z_{n_i}(A)| + |a_{n,i} - 1|\,|Z_{n_i}(A)|. \end{aligned} \tag{3.7}$$

For given $\varepsilon_1 > 0$, one can choose λ sufficiently close to 1, to make $|a_{n,i} - 1| < \varepsilon_1$ and $|(n/n_i)A \,\Delta\, A| < \varepsilon_1$ for $n_{i-1} < n \le n_i$. Hence, whenever $Z_{n_i} \in \mathscr{G}^\varepsilon$ for $\varepsilon > 0$, we have by (3.7) and (3.1) that

$$\|Z_n - Z_{n_i}\|_{\mathscr{A}} \le (1 + \varepsilon_1)(2\varepsilon + 2\varepsilon^{1/2}) + \varepsilon_1(\varepsilon + 1). \tag{3.8}$$

This shows that for any $\varepsilon > 0$

$$P[Z_n \notin G^\varepsilon \text{ i.o.}] = 0.$$

The above argument therefore completes the proof since it establishes that with probability 1, $\mathscr{G}$ contains all of the limit points of $\{Z_n; n \ge 3\}$. □

4. The embedding: Definition and properties. The embedding of processes into Brownian Motion has often been used to prove weak convergence and iterated logarithm results. The idea of using embeddings based on stopping times originated with Skorokhod (1965) when he introduced a particular stopping time based on the two barrier problem. If $W = \{W_t: t \geq 0\}$ denotes a standard Brownian Motion and Y is any real-valued mean zero r.v., a stopping time embedding of Y in W is the existence of a stopping time, τ say, for which $Y =_L W_\tau$. Many methods for generating such τ's have been suggested since Skorokhod (1965), including among others those of Root (1969), based on hitting times, Dubins (1968), based on a sequence of two-barrier problems determined by conditional means, Monroe (1971), based on local times, and Bass (1983), based on martingale representations. The hitting time method of Root is known to minimize the variance of the stopping time τ among all others. (Cf. Rost, 1976.)

The embedding of processes, such as martingales, partial-sum and empirical processes, is possible by repeated usage of the above single r.v. embedding. In these applications, including the one to be described below, it does not matter very much which stopping time is used to generate the individual r.v.'s. The only restriction is that finite moments of τ of appropriate order should be implied by finite moments of Y; see Lemma 4.1 below.

The object of the embedding in this section is to use the Brownian process Z of Section 2 to generate each member of the given array $\{X_{\mathbf{j}}: \mathbf{j} \in J^d\}$ so that they will be independent with the proper distributions and so that the stopping times will be measurable relative to an appropriate "past" of Z. We proceed as follows.

Let $\mathcal{Q}$ be the partition of $R_+^{d-1}: = (0, \infty)^{d-1}$ into disjoint unit cubes of the form $(\mathbf{q} - \mathbf{1}, \mathbf{q}]$ for $\mathbf{q} \in J^{d-1}$. For each $Q \in \mathcal{Q}$, define Z^Q by

(4.1) $$Z^Q(t) = Z(Q \times (0, t]), \quad t \geq 0.$$

Observe that each Z^Q is a standard Brownian Motion and that Z^{Q_1} and Z^{Q_2} are independent for disjoint Q_1 and Q_2. Let $\mathbf{q} = (q_1, q_2, \ldots, q_{d-1})$ denote the upper integer vertex of Q, so that $Q = (\mathbf{q} - \mathbf{1}, \mathbf{q}]$. Now, use Z^Q to generate, by any of the available embedding methods, independent r.v.'s $X_{\mathbf{q},1}, X_{\mathbf{q},2}, \cdots$ with common distribution equal to that of $X_{\mathbf{j}}$. We will not hereafter distinguish between these constructed X's and the given array. Let $0 \leq T_{\mathbf{q},1} \leq T_{\mathbf{q},2} \leq \ldots$ be the successive stopping times of the embedding, so that in particular

$$Z^Q(T_{\mathbf{q},j}) = X_{\mathbf{q},1} + \cdots + X_{\mathbf{q},j}.$$

For any $\mathbf{j} = (\mathbf{q}, j) \in J^d$, define the random set

(4.2) $$\Gamma_{\mathbf{j}} = Q \times (T_{\mathbf{q},j-1}, T_{\mathbf{q},j}]$$

so that then $Z(\Gamma_{\mathbf{j}}) = X_{\mathbf{j}}$. In view of the definitions of the smoothed and unsmoothed partial-sum processes given in (1.3), we may then write

$$S_n(A) = n^{-d/2} b_n^{-1} Z(\Gamma_n^*(A))$$

and

(4.3) $$X_n(A) = n^{-d/2} b_n^{-1} \{Z(\Gamma_n(A)) + \textstyle\sum_{\mathbf{j}: C_{\mathbf{j}} \cap n(\partial A) \neq \varnothing} |C_{\mathbf{j}} \cap nA| Z(\Gamma_{\mathbf{j}})\}$$

where $C_{\mathbf{j}} = (\mathbf{j} - \mathbf{1}, \mathbf{j}]$,

$$\Gamma_n^*(A) = \cup \{\Gamma_{\mathbf{j}}: \mathbf{j} \in nA\} \quad \text{and} \quad \Gamma_n(A) = \cup \{\Gamma_{\mathbf{j}}: C_{\mathbf{j}} \subset nA\}.$$

Thus the Brownian measures of the random sets $\Gamma_n^*(A)$ and $\Gamma_n(A)$ are essentially equivalent to the partial-sum measures of A. Our task will be to show that for the smoothed case, the Z-measure of the random set $\Gamma_n(A)$ is sufficiently close to $Z(A)$, uniformly over $\mathscr{A}$, to relate the asymptotics of X_n to those of Z. Some additional notation will be needed. For $A \in \mathscr{A}$, define

$$C_n^+(A) = \cup \{C_{\mathbf{j}}: C_{\mathbf{j}} \cap (nA) \neq \varnothing\}, \quad C_n(A) = \cup \{C_{\mathbf{j}}: C_{\mathbf{j}} \subset nA\}$$

and

$$\Gamma_n^+(A) = \cup \{\Gamma_{\mathbf{j}}: C_{\mathbf{j}} \cap (nA) \neq \varnothing\}.$$

The sets $C_n^+(A)$, $C_n(A)$ and $\Gamma_n^+(A)$, $\Gamma_n(A)$ are respectively the outer and inner rectilinear fits of nA by the nonrandom cubes $C_{\mathbf{j}}$ and the random rectangles $\Gamma_{\mathbf{j}}$. In the next section, bounds on the Lebesgue measure of symmetric differences between these fits will be derived.

As a necessary preliminary to these bounds we need the following result relating the moments of $|\Gamma_{\mathbf{j}}|$ to those of $X_{\mathbf{j}}$. (Recall that in our embedding, $Z(\Gamma_{\mathbf{j}}) = X_{\mathbf{j}}$.) Under any of the usual embedding schemes including those referenced at the start of this section, it is easy to show that the stopping time τ satisfies, for $\beta > 0$ and $\alpha > 1$,

$$\sup_{0 \le t < \infty} E|W_{\tau\wedge t}|^2(\log^+|W_{\tau\wedge t}|)^{1+\beta} \le E|Y|^2(\log^+|Y|)^{1+\beta} \tag{4.4}$$

and

$$\sup_{0 \le t < \infty} E|W_{\tau\wedge t}|^{\alpha} \le E|Y|^{\alpha}. \tag{4.5}$$

LEMMA 4.1. a) *If the embedding stopping times satisfy* (4.4) *for* $\beta > 0$, *then*

$$E|X_{\mathbf{j}}|^2(\log^+|X_{\mathbf{j}}|)^{1+\beta} < \infty \textit{ implies } E|\Gamma_{\mathbf{j}}|\,(\log^+|\Gamma_{\mathbf{j}}|)^{1+\beta} < \infty.$$

b) *If the embedding stopping times satisfy* (4.5) *for* $\alpha > 1$, *then*

$$E|X_{\mathbf{j}}|^{\alpha} < \infty \textit{ implies } E|\Gamma_{\mathbf{j}}|^{\alpha/2} < \infty.$$

PROOF. a) Let Φ be convex, increasing, continuous, nonnegative, and, for $x \ge e$, equal to $x(\log^+ x)^{(1+\beta)/2}$. By applying Doob's inequality, with exponent $p = 2$, to the submartingale $\Phi(|W_t|)$, we get $E\sup_{t\le\tau}(|W_t|^2(\log^+|W_t|^{1+\beta}) < \infty$. To complete the proof, apply the inequalities of Burkholder, Davis and Gundy (Meyer, 1976, page 351).

b) The proof is similar, using $\Phi(x) = |x|^{(1+\alpha)/2}$ and Doob's inequality with exponent $p = 2\alpha/(1+\alpha)$. □

5. Tail estimates on the closeness of the embedding. The purpose of this section is to derive bounds on $|Z_n(A) - X_n(A)|$ and $|X_n(A)|$ for a fixed set A. Throughout this section, the letter c without subscripts will denote a generic constant whose value may change from line to line. Summations over $\mathbf{q} \in \{1, \dots, n\}^{d-1}$ and over $\mathbf{j} \in \{1, \dots, n\}^d$ will be abbreviated by $\sum_{\mathbf{q}}$ and $\sum_{\mathbf{j}}$, respectively.

We begin with the following proposition which is an extension of some results of Heyde and Rohatgi (1967). Let $M_{\mathbf{q}} = \max_{1\le i\le n}|T_{\mathbf{q},i} - i|$.

PROPOSITION 5.1. a) *Given* $\eta > 0$ *and* $\beta > 0$, *there exist* c_1 *and* n_1 *depending only on* β, η, *and* $E|\Gamma_j|(\log^+|\Gamma_j|)^{1+\beta}$ *such that if* $n \geq n_1$,

$$P(\textstyle\sum_{\mathbf{q}} M_{\mathbf{q}} \geq n^d\eta) \leq c_1/\eta(\log^+ n^d\eta)^{1+\beta}.$$

b) *For* $K > 0$, $\beta > 0$, $s \geq 1$, *and* $\varphi > 0$, *there exist* c_2 *and* n_2 *depending only on* K, β, s, φ, $E|\Gamma_{\mathbf{j}}|^s$, *and* $E|\Gamma_{\mathbf{j}}|^{s+\varphi}$ *such that if* $n \geq n_2$ *and* $E|T_{\mathbf{j}}|^s < K$,

$$P(\textstyle\sum_{\mathbf{j}} |\Gamma_{\mathbf{j}}|^s \geq n^d K) \leq c_2/(\log^+ n)^{1+\beta}.$$

PROOF. a) We suppose $d \geq 2$, the $d = 1$ case being somewhat simpler. In preparation for the central limit theorem in Section 7, we take $n \geq 3$ large enough so that $\eta \geq (\log n)^{-\beta/3}$.

Define $V_{\mathbf{j}}$ to be $|\Gamma_{\mathbf{j}}| - 1$ if $||\Gamma_{\mathbf{j}}| - 1| \leq \eta n^d$ and 0 otherwise and let $M_{\mathbf{q}}^* = \max_{1\leq i\leq n} \sum_{k=1}^i V_{\mathbf{q},k}$.

First of all, by Chebyshev's inequality and the monotoneity of $x(\log^+ x)^{1+\beta}$,

$$P(||\Gamma_j| - 1| > x) \leq E||\Gamma_j| - 1|(\log^+||\Gamma_j| - 1|)^{1+\beta}/x(\log^+ x)^{1+\beta}.$$

Then, since $E|\Gamma_{\mathbf{j}}| - 1 = EX_{\mathbf{j}}^2 - 1 = 0$,

$$|EV_{\mathbf{j}}| \leq \int_{\eta n^d}^{\infty} P(||\Gamma_{\mathbf{j}}| - 1| > x)\,dx \leq c/\log(\eta n^d)^{\beta}. \tag{5.1}$$

Secondly,

$$EV_{\mathbf{j}}^2 = 2\int_0^{\eta n^d} xP(||\Gamma_{\mathbf{j}}| - 1| > x)\,dx \leq c\eta n^d/(\log \eta n^d)^{1+\beta}. \tag{5.2}$$

Suppose for the moment that we have shown that for n sufficiently large, $n^{-1}EM_{\mathbf{q}}^* \leq \eta/2$. Then

$$\begin{aligned} &P(\textstyle\sum_{\mathbf{q}} M_{\mathbf{q}} > n^d\eta) \\ &\quad\leq P(|\Gamma_{\mathbf{j}}| - 1 \neq V_{\mathbf{j}} \text{ for some } \mathbf{j}) + P(\textstyle\sum_{\mathbf{q}}(M_{\mathbf{q}}^* - EM_{\mathbf{q}}^*) \geq n^d\eta/2). \end{aligned} \tag{5.3}$$

The first term on the right hand side of (5.3) is

$$\leq n^d P(||\Gamma_{\mathbf{j}}| - 1| > \eta n^d) \leq c/\eta(\log \eta n^d)^{1+\beta},$$

as desired. The second term on the right hand side of (5.3) is

$$\begin{aligned} &\leq 4n^{d-1}\mathrm{Var}\, M_{\mathbf{q}}^*/n^{2d}\eta^2 \leq cEM_{\mathbf{q}}^{*2}/n^{d+1}\eta^2 \leq 2c\eta^{-2}n^{-d-1} \\ &\quad\cdot(E[\max_{1\leq i\leq n}|\textstyle\sum_{k=1}^i (V_{\mathbf{q},k} - EV_{\mathbf{q},k})|]^2 + \max_{1\leq i\leq n}|E\sum_{k=1}^i V_{\mathbf{q},k}|^2) \\ &\leq 8c\eta^{-2}\mathrm{Var}\,\textstyle\sum_{k=1}^n V_{\mathbf{q},k}/n^{d+1} + 2c\eta^{-2}|EV_{\mathbf{q},1}|^2/n^{d-1}, \end{aligned}$$

the last line following by Doob's inequality. Using $\mathrm{Var}\sum_{k=1}^n V_{\mathbf{q},k} = n\,\mathrm{Var}\,V_{\mathbf{q},1} \leq n\,EV_{\mathbf{q},1}^2$, (5.1), and (5.2) gives the desired estimate.

It remains to show that $n^{-1}EM_{\mathbf{q}}^* \leq \eta/2$. This is not an immediate consequence of a law of large numbers since $V_{\mathbf{q},1}$ depends on n.

So suppose $n \geq 3$, $\lambda \geq (\log n)^{-\beta/2}$, fix $\mathbf{q}$, and let $\hat{V}_i = V_{\mathbf{q},i}$ if $|V_{\mathbf{q},i}| > n\lambda$ and 0 otherwise, and let $\hat{M}^* = \max_{1\leq i\leq n}|\sum_{k=1}^i \hat{V}_k|$.

As in (5.1) and (5.2), $|E\hat{V}_i| \le c/(\log n)^\beta$, which is $\le \lambda/2$ for n sufficiently large, and $E\hat{V}_i^2 \le cn\lambda/(\log n\lambda)^{1+\beta}$. Then

$$P(M_{\mathbf{q}}^* > n\lambda) \le P(|V_{\mathbf{q},i}| > n\lambda \quad \text{for some} \quad i) + P(\hat{M}^* > n\lambda)$$
$$\le n/(n\lambda(\log^+ n\lambda)^{1+\beta}) + P(\max_{1\le i\le n}|\textstyle\sum_{k=1}^i (\hat{V}_i - E\hat{V}_i)| > n\lambda/2).$$

By Kolmogorov's inequality, the last term on the right is

$$\le 4 \operatorname{Var}(\textstyle\sum_{k=1}^n \hat{V}_k)/n^2\lambda^2 \le 4\, E\hat{V}_1^2/n\lambda^2.$$

We thus get $P(M_q^*/n > \lambda) \le c/(\lambda(\log n\lambda)^{1+\beta})$, and so

$$E(M_{\mathbf{q}}^*/n) \le (\log n)^{-\beta/2} + \int_{(\log n)^{-\beta/2}}^\infty P(M_q^*/n > \lambda)\, d\lambda$$
$$\le (\log n)^{-\beta/2} + c\int_{n(\log n)^{-\beta/2}}^\infty (x(\log x)^{1+\beta})^{-1} dx$$
$$\le \eta/2 \text{ for } n \text{ sufficiently large.}$$

b) is proved in a very similar fashion, except that one works with $|\Gamma_{\mathbf{j}}|^s - E|\Gamma_{\mathbf{j}}|^s$ instead of $|\Gamma_{\mathbf{j}}| - 1$. □

We now use Proposition 5.1 to get bounds on how much a given set is transformed by the embedding.

PROPOSITION 5.2 a) $\sup_{A\in\mathscr{A}}|C_n^+(A)\backslash C_n(A)| \le ⓒd^{1/2}n^{d-1}$,
b) $P(\sup_{A\in\mathscr{A}}|\Gamma_n(A)\,\Delta\, C_n(A)| \ge n^d\lambda) \ge c_3/(\log^+ n)^{1+\beta}$,
c) $P(\sup_{A\in\mathscr{A}}|\Gamma_n^+(A)\,\Delta\, C_n^+(A)| \ge n^d\lambda) \le c_4/(\log^+ n)^{1+\beta}$,
where c_3 and c_4 depend on ⓒ, λ, and the constants c_1 and c_2 of Proposition 5.1.

PROOF. Fix $A \in \mathscr{A}$. First observe that any point in $C_n^+(A)\backslash C_n(A)$ must be within a distance $d^{1/2}$ from $\partial(nA)$. Then, by Property (iv) of Section 2, $|C_n^+(A)\backslash C_n(A)| \le |(nA)(d^{1/2})| = n^d|A(d^{1/2}/n)| \le ⓒn^{d-1}d^{1/2}$, which proves (a). Note in particular that the number of $\mathbf{j}$'s for which $C_{\mathbf{j}} \cap \partial(nA) \ne \varnothing$ can be at most $ⓒd^{1/2}n^{d-1}$.

Given λ, choose an integer $m \ge 1$ so that $2ⓒd^{1/2}/m < \lambda/2$. For each $\mathbf{q} \in \{1, \dots, n\}^{d-1}$, let $N_{\mathbf{q}}$ be the number of i's for which $C_{\mathbf{q},i} \cap \partial(nA) \ne \varnothing$. Let $Q_1 = \{\mathbf{q}\colon N_{\mathbf{q}} \ge m\}$, $Q_2 = \{\mathbf{q}\colon N_{\mathbf{q}} < m\}$. By the previous remark, $\sum_{\mathbf{q}} N_{\mathbf{q}} \le ⓒd^{1/2}n^{d-1}$, and so the cardinality of Q_1 is at most $ⓒd^{1/2}n^{d-1}/m$.

Let $D_{\mathbf{q}} = (\mathbf{q} - \mathbf{1}, \mathbf{q}] \times (0, n]$, the rectangular channel with base $(\mathbf{q} - \mathbf{1}, \mathbf{q}]$. If $\mathbf{q} \in Q_1$,

$$|\Gamma_n(A) \cap D_{\mathbf{q}}| + |C_n(A) \cap D_{\mathbf{q}}| \le T_{\mathbf{q},n} + n \le M_{\mathbf{q}} + 2n. \tag{5.4}$$

Now consider $\mathbf{q} \in Q_2$. We want to count the number of components of $D_{\mathbf{q}} \cap C_n(A)$. If for some i, either a) $C_{\mathbf{q},i} \subseteq C_n(A)$ but $C_{\mathbf{q},i+1} \not\subseteq C_n(A)$ or b) $C_{\mathbf{q},i} \not\subseteq C_n(A)$ but $C_{\mathbf{q},i+1} \subseteq C_n(A)$, then either $C_{\mathbf{q},i}$ or $C_{\mathbf{q},i+1}$ must intersect $\partial(nA)$. So the number of i's for which either a) or b) holds is at most $2N_{\mathbf{q}} + 2 \le 2m$. Thus there are $2m$

pairs of indices $i_{(1)} \le i^{(1)}, \ldots, i_{(2m)} \le i^{(2m)}$ such that $C_n(A) \cap D_{\mathbf{q}} = \cup \{C_{\mathbf{q},i} : i_{(k)} \le i < i^{(k)}$ for some $k \le 2m\}$.

Since

$$(\Gamma_n(A) \cap D_{\mathbf{q}}) \,\Delta\, (C_n(A) \cap D_{\mathbf{q}}) \subseteq \cup_{k=1}^{2m} \{(\cup_{i_{(k)} \le i < i^{(k)}} \Gamma_{\mathbf{q},i}) \,\Delta\, (\cup_{i_{(k)} \le i < i^{(k)}} C_{\mathbf{q},i})\},$$

then

$$\begin{aligned} |(\Gamma_n(A) \cap D_{\mathbf{q}}) \,\Delta\, (C_n(A) \cap D_{\mathbf{q}})| &\le \textstyle\sum_{k=1}^{2m} (|T_{\mathbf{q},i_{(k)}} - i_{(k)}| + |T_{\mathbf{q},i^{(k)}} - i^{(k)}|) \\ &\le 4mM_{\mathbf{q}}. \end{aligned} \tag{5.5}$$

We thus get by combining (5.4) and (5.5),

$$\begin{aligned} |\Gamma_n(A) \,\Delta\, C_n(A)| &\le \textstyle\sum_{\mathbf{q}} |(\Gamma_n(A) \cap D_{\mathbf{q}}) \,\Delta\, (C_n(A) \cap D_{\mathbf{q}})| \\ &\le \textstyle\sum_{\mathbf{q} \in Q_1} (M_{\mathbf{q}} + 2n) + \sum_{\mathbf{q} \in Q_2} 4mM_{\mathbf{q}} \\ &\le 2n\mathfrak{c}d^{1/2}n^{d-1}/m + 4m \textstyle\sum_{\mathbf{q}} M_{\mathbf{q}}, \end{aligned}$$

a bound independent of A. Then

$$P(\sup_{A \in \mathscr{A}} |\Gamma_n(A) \,\Delta\, C_n(A)| \ge n^d\lambda) \le P(\textstyle\sum_{\mathbf{q}} M_{\mathbf{q}} \ge n^d\lambda/4m)$$

and b) now follows from an application of Proposition 5.1 a). The proof of c) is virtually identical to that of b). □

Next we prove a proposition of the stochastic calculus which will enable us to get estimates on $|Z(\Gamma_n(A)) - Z(C_n(A))|$ from estimates on $|\Gamma_n(A) \,\Delta\, C_n(A)|$.

In the following, $\int H_s \, dW_s$ denotes the stochastic integral of H with respect to W. See Meyer (1976, Chapter 2) for further information about stochastic integrals.

Proposition 5.3. *Let W_t: $t \ge 0$ be a one-parameter Brownian motion, H_t: $t \ge 0$ a nonanticipating functional and t_0, λ, and L arbitrary positive constants. Then*

$$P\left(\left| \int_0^{t_0} H_s \, dW_s \right| > \lambda, \int_0^{t_0} H_s^2 \, ds \le L \right) \le 2e^{-\lambda^2/2L}.$$

Proof. Set $U = \inf\{t > 0 : \int_0^t H_s^2 \, ds > L\}$ and $M_t = \int_0^{t \wedge U} H_s \, dW_s$. It can be shown that M_t is a continuous martingale with quadratic variation $\int_0^{t \wedge U} H_s^2 \, ds \le L$. Then

$$\begin{aligned} P\left(\left| \int_0^{t_0} H_s \, dW_s \right| > \lambda, \int_0^{t_0} H_s^2 \, ds \le L \right) &= P\left(\left| \int_0^{t_0} H_s \, dW_s \right| > \lambda, U > t_0 \right) \\ &\le P(|M_{t_0}| > \lambda) \le 2e^{-\lambda^2}/2L. \end{aligned}$$

The last inequality is well-known; a proof may be found, for example, in Dellacherie, Doléans-Dade, and Meyer (1970), page 247. □

Suppose $E|\Gamma_{\mathbf{j}}|^{s+\varphi} < \infty$ for some $s > 1$, $\varphi > 0$. Let p satisfy $p^{-1} + s^{-1} = 1$. Let $\varepsilon > 0$, $k > (E|\Gamma_{\mathbf{j}}|^s)^{1/s}$. We define the event Λ by

$$\Lambda := (\sup_{A\in\mathscr{A}}|\Gamma_n(A)\ \Delta\ C_n(A)| \le \varepsilon n^d,\ \sup_{A\in\mathscr{A}}|\Gamma_n^+(A)\ \Delta\ C_n^+(A)| \le \varepsilon n^d,\ \textstyle\sum_{\mathbf{j}}|\Gamma_{\mathbf{j}}|^s \le k^s n^d). \tag{5.6}$$

We then obtain the following estimates:

PROPOSITION 5.4. *If, $A, B \in \mathscr{A}$, then*
(a) $P(|X_n(B) - X_n(A)| > \lambda, \Lambda) \le 2\exp(-\lambda^2 L_2 n/2k|A\ \Delta\ B|^{1/p})$, and
(b) *for $n \ge n_0$, independent of A,*

$$P(|Z_n(A) - X_n(A)| > \lambda, \Lambda) \le 4\exp(-\lambda^2 L_2 n/12\varepsilon).$$

PROOF. First of all, note that $C_n(I^d) = nI^d$ and $|\Gamma_n(I^d) \le |\Gamma_n(I^d)\ \Delta\ C_n(I^d)| + |C_n(I^d)| \le (1+\varepsilon)n^d$ on the event Λ. Secondly, on Λ, $T_{\mathbf{j}} \le |\Gamma_n(I^d)| \le (1+\varepsilon)n^d \le n^{d+1}$ for all $\mathbf{j}$.

Next we define the Brownian Motion W_t to which we will apply Proposition 5.3. The idea is to place end to end the Brownian Motions $Z^Q(t)$, $0 \le t \le n^{d+1}$ of Section 4. Let ψ be any one-to-one mapping of the integers $\{1, 2, \dots, n^{d-1}\}$ to $\{1, 2, \dots, n\}^{d-1}$. Now let $W_0 = 0$, and for $m = 1, 2, \dots, n^{d-1}$, let

$$W_{t+(m-1)n^{d+1}} = W_{(m-1)n^{d+1}} + Z((\psi(m) - \mathbf{1}, \psi(m)] \times (0, t]), \quad 0 < t \le n^{d+1}.$$

If $\mathbf{j} = (\mathbf{q}, i)$, we want to define $K_{\mathbf{j}}(t)$ so that $\int_0^{n^{2d}} K_{\mathbf{j}}(t)\,dW_t = Z(C_{\mathbf{j}})/n^{d/2}b_n$. We do this by setting $K_{\mathbf{j}}(t)$ equal to $1/n^{d/2}b_n$ if $(\psi^{-1}(\mathbf{q}) - 1)n^{d+1} + (i-1) < t \le (\psi^{-1}(\mathbf{q}) - 1)n^{d+1} + i$ and 0 otherwise.

Analogously, set $\xi_{\mathbf{j}}(t) = 1/n^{d/2}b_n$ if $(\psi^{-1}(\mathbf{q}) - 1)n^{d+1} + (T_{\mathbf{q},i-1} \wedge n^{d+1}) < t \le (\psi^{-1}(\mathbf{q}) - 1)n^{d+1} + (T_{\mathbf{q},i} \wedge n^{d+1})$ and 0 otherwise. $\xi_{\mathbf{j}}(t)$ chosen in this fashion satisfies

$$\int_0^{n^{2d}} \xi_{\mathbf{j}}(t)\,dW_t = Z(\Gamma_{\mathbf{j}} \cap \hat{I})/n^{d/2}b_n, \quad \text{where} \quad \hat{I} = (0, n)^{d-1} \times (0, n^{d+1}].$$

On Λ, $Z(\Gamma_{\mathbf{j}} \cap \hat{I}) = Z(\Gamma_{\mathbf{j}})$ since we argued that $T_{\mathbf{j}} \le n^{d+1}$ on Λ.

We are now ready to prove (a). Let $t_0 = n^{2d}$. On the event Λ,

$$X_n(A) = \int_0^{t_0} \textstyle\sum_{\mathbf{j}}|nA \cap C_{\mathbf{j}}|\,\xi_{\mathbf{j}}(t)\,dW_t,$$

and similarly for $X_n(B)$. If we let $H_t = \sum_{\mathbf{j}}\{|nA \cap C_{\mathbf{j}}| - |nB \cap C_{\mathbf{j}}|\}\xi_{\mathbf{j}}(t)$, (a) will follow from Proposition 5.3 provided we show $\int_0^{t_0} H_t^2\,dt \le k|A\ \Delta\ B|^{1/p}/b_n^2$ on Λ.

If we let $d_{\mathbf{j}} = |nA \cap C_{\mathbf{j}}| - |nB \cap C_{\mathbf{j}}|$,

$$|d_{\mathbf{j}}| \le |(nA\ \Delta\ nB) \cap C_{\mathbf{j}}| \le 1,$$

and

$$\textstyle\sum_{\mathbf{j}}|d_{\mathbf{j}}| \le |nA\ \Delta\ nB| = n^d|A\ \Delta\ B|.$$

Note that $\int_0^{t_0} \xi_j(t)\,dt = |\Gamma_j|$ on Λ and that the ξ_j's have disjoint support. Then on Λ,

$$\int_0^{t_0} H_t^2\,dt = \int_0^{t_0} \sum_j d_j^2\xi_j^2(t)\,dt = \sum_j d_j^2|\Gamma_j|/n^d b_n^2$$
$$\le \left(\sum_j d_j^{2p}\right)^{1/p}\left(\sum_j |\Gamma_j|^s\right)^{1/s}/n^d b_n^2$$
$$\le \left(\sum_j d_j\right)^{1/p}(k^s n^d)^{1/s}/n^d b_n^2$$
$$\le |A \,\Delta\, B|^{1/p}(n^d)^{1/p}k(n^d)^{1/s}/n^d b_n^2 = k|A \,\Delta\, B|^{1/p}/b_n^2$$

as required.

We now proceed to prove (b). Since A and $C_n(A)$ are deterministic sets, the usual tail bounds for the normal distribution give

$$P(|Z(nA) - Z(C_n(A))| > (\lambda/2)n^{d/2}b_n) \le 2\exp(-\lambda^2 n^d b_n^2/8\,|nA \,\Delta\, C_n(A)|).$$

By Proposition 5.2, $|nA \,\Delta\, C_n(A)| \le \text{c}d^{1/2}n^{d-1}$. It therefore suffices to get a bound on $P(|Z(C_n(A))/n^{d/2}b_n - X_n(A)| > \lambda/2)$.

To do this, let

$$H_t = \sum_j\{|nA \cap C_j|\,\xi_j(t) - |C_n(A) \cap C_j|\,K_j(t)\}/n^{d/2}b_n.$$

Again, we will get (b) from Proposition 5.3 provided we show that $\int_0^{t_0} H_t^2\,dt \le 3\varepsilon/b_n^2$ on Λ.

Note that

$$|H_t| \le \max\{|\textstyle\sum_{nA\cap C_j\ne\varnothing}\xi_j(t) - \sum_{C_j\subseteq nA}K_j(t)|,\ |\sum_{C_j\subseteq nA}K_j(t) - \sum_{C_j\subseteq nA}\xi_j(t)|\}.$$

So on Λ,

$$n^d b_n^2\int_0^{t_0} H_t^2\,dt \le |\Gamma_n^+(A) \,\Delta\, C_n(A)| + |\Gamma_n(A) \,\Delta\, C_n(A)|$$
$$\le |\Gamma_n^+(A) \,\Delta\, C_n^+(A)| + |\Gamma_n(A \,\Delta\, C_n(A)| + |C_n^+(A)\backslash C_n(A)|$$
$$\le n^d(2\varepsilon + \text{c}d^{1/2}/n)$$
$$\le 3\varepsilon n^d \quad \text{for} \quad n \ge n_0 \quad \text{sufficiently large.}\ \square$$

6. The functional LIL for partial-sum processes. We are now ready to prove the functional LIL for the smoothed partial-sum processes.

THEOREM 6.1. *Suppose that $\mathscr{A}$ satisfies Properties* (i), (ii), (iii), (iv) and (v′) *of Section* 4 (*but not necessarily assumption* (vi)). *Suppose also that* $E|X_j|^{2s} < \infty$ *where* $s > (1-r)^{-1}$. *We then have that with probability* 1, $\{X_n;\, n \ge 3\}$ *is relatively compact in* $\mathscr{C}(\mathscr{A})$ *with limit points exactly* $\mathscr{G}$.

PROOF. Observe first of all that we do not require property (vi), that $\mathscr{A}$ be origin sparse. In fact, this assumption is used only to give content to the statement of Theorem 3.1. Note, however, that $Z_n(\cdot)$ may be defined for $\mathscr{A}$ satisfying only (i)–(v), as long as n is fixed. Looking at the proof of Theorem 3.1 we note that what it tells us is that, for all $\varepsilon > 0$, $\lim_{n\to\infty}\lim_{m\to\infty}P[Z_k \notin \mathscr{G}^\varepsilon,\ n \le k \le m] = 0$.

We show below that

$$\| X_k - Z_k \|_{\mathscr{A}} \to 0 \quad \text{a.s. as} \quad k \to \infty,$$

with no use of assumption (vi). It will then follow that

$$\lim_{n\to\infty} \lim_{m\to\infty} P[X_k \notin \mathscr{G}^\varepsilon,\ n \le k \le m] = 0,$$

or $P[X_n \notin \mathscr{G}^\varepsilon \text{ i.o.}] = 0$. Thus, the X_n's will be relatively compact with probability 1 with limit points in $\mathscr{G}$.

Since the set of limit points of Z_n is exactly $\mathscr{G}$, analogous reasoning will show that the set of limit points of X_n is exactly $\mathscr{G}$ as desired, without the use of assumption (vi). It thus suffices to show that $\| \Delta_n(\cdot) \|_{\mathscr{A}}$ is $o(1)$ almost surely, where $\Delta_n := X_n - Z_n$.

By definition

$$\Delta_n(A) = X_n(A) - Z_n(A) = (n^{d/2} b_n)^{-1} \sum_{j\in J^d} \{ | C_j \cap nA | Z(\Gamma_j) - Z(C_j \cap nA) \}.$$

From Proposition 5.4 (ii), for each $\lambda > 0$, $\varepsilon > 0$, there exist n_0 such that for $n > n_0$,

$$P([| \Delta_n(A) | > \lambda] \cap \Lambda) \le 2 \exp\{-\lambda^2 L_2 n / 12\varepsilon\} \tag{6.1}$$

while by Propositions 5.1 and 5.2

$$P(\Lambda^c) \le c(\log^+ n)^{-1-\beta} \tag{6.2}$$

where Λ is defined in (5.6). Since

$$P(\| \Delta_n \|_{\mathscr{A}} > \lambda) \le P(\| \Delta_n \|_{\mathscr{A}} > \lambda,\ \Lambda) + P(\Lambda^c) \tag{6.3}$$

it suffices to bound $P(\| \Delta_n \|_{\mathscr{A}} > \lambda, \Lambda) \le c(\log^+ n)^{-1-\beta}$. For a fixed $\delta_0 > 0$, (6.1) implies

$$\begin{aligned} P(\| \Delta_n \|_{\mathscr{A}_{\delta_0}} < \lambda,\ \Lambda) &\le 2 \exp\{H(\delta_0) - \lambda^2 L_s n / 12\varepsilon\} \\ &= 2 \exp\left\{ -\lambda^2 (12\varepsilon)^{-1} L_2 n \left(1 - \frac{12\varepsilon H(\delta_0)}{\lambda^2 L_2 n} \right) \right\} \\ &\le (\log^+ n)^{-\rho} \end{aligned} \tag{6.4}$$

for some $\rho > 1$ and all sufficiently large n, provided only that ε is chosen sufficiently small.

To obtain the desired bound for $\| \Delta_n \|_{\mathscr{A}}$ rather than $\| \Delta_n \|_{\mathscr{A}_{\delta_0}}$, write

$$\Delta_n(A) = \Delta_n(A_{\delta_0}) + \sum_{i=1}^{\infty} \{ \Delta_n(A_{\delta_i}) - \Delta_n(A_{\delta_{i-1}}) \}$$

for $A_{\delta_i} \in \mathscr{A}_{\delta_i}$ and $| A_{\delta_i} \Delta A | < \delta_i$, $i \ge 0$. Then use

$$| \Delta_n(A) - \Delta_n(A_{\delta_0}) | \le \sum_{i=1}^{\infty} \{ | X_n(A_{\delta_i}) - X_n(A_{\delta_{i-1}}) | + | Z_n(A_{\delta_i}) - Z_n(A_{\delta_{i-1}}) | \}.$$

We continue to work on the event Λ. By Proposition 5.4 (i)

$$\begin{aligned} &P(\max_{A\in\mathscr{A}} | X_n(A_{\delta_i}) - X_n(A_{\delta_{i-1}}) > \lambda_i,\ \Lambda) \\ &\qquad \le 2 \exp\{2H(\delta_i) - \lambda_i^2 L_2 n / 2k(2\delta_{i-1})^{1/p}\} \end{aligned} \tag{6.5}$$

since $|A_{\delta_i} \Delta A_{\delta_{i-1}}| \le \delta_i + \delta_{i-1} \le 2\delta_{i-1}$. Recall that $1/p + 1/s = 1$. Choose $\delta_i = \delta_0\gamma^i$ and $\lambda_i = \lambda_0\gamma^{bi}$ for $b > 0$. Since $H(\delta) \le K\delta^{-r}$, $0 < r < 1$, the bound on the right hand side of (6.5) is bounded above by

$$(6.6)\qquad \begin{aligned} &2\exp\left\{-\lambda_i^2(2k(2\delta_{i-1})^{1/p})^{-1}L_2n\left(1 - \frac{4Kk(2\delta_{i-1})^{1/p}\delta_i^{-r}}{\lambda_i^2 L_2 n}\right)\right\} \\ &\qquad \le 2\exp\{-C_1\gamma^{(2b-p^{-1})i}L_2n(1 - C_2\gamma^{(p^{-1}-r-2b)i}/L_2n)\} \end{aligned}$$

where

$$C_1 = \lambda_0^2(\gamma/2\delta_0)^{1/p}/2k, \quad C_2 = 4Kk(2/\gamma)^{1/p}\delta_0^{1/p-r}/\lambda_0^2.$$

For $1 \le p < 1/r$, choose $b \le (p^{-1} - r)/2$ to insure that the bound of (6.6) is summable over i. Choose λ_0 so that $\sum_{i=1}^{\infty} \lambda_i \le \lambda$, and then choose δ_0 sufficiently small to make the sum of the bounds of (6.6) less than $(\log^+ n)^{-\rho}$ for some $\rho > 1$ and all n sufficiently large.

It remains to note that

$$P(\max_{A\in\mathscr{A}}|Z_n(A_{\delta_i}) - Z_n(A_{\delta_{i-1}})| > \lambda_i, \Lambda) \le 2\exp\{2H(\delta_i) - \lambda_i^2 L_2 n/2\delta_{i-1}\}$$

since Z_n is Normal with zero mean and variance $|A_{\delta_i} \Delta A_{\delta_{i-1}}| \le 2\delta_{i-1}$. With the same choices of constants δ_i and λ_i, this last term equals

$$(6.7)\qquad \begin{aligned} &2\exp\left\{-\lambda_i^2(2\delta_{i-1})^{-1}L_2n\left(1 - \frac{4\delta_{i-1}H(\delta_i)}{\lambda_i^2 L_2 n}\right)\right\} \\ &\qquad \le 2\exp\{-C_3\gamma^{-(1-2b)i}L_2n(1 - C_4\gamma^{(1-r-2b)i}/L_2n)\} \end{aligned}$$

where

$$C_3 = \lambda_0^2\gamma/2\delta_0 \quad \text{and} \quad C_4 = 4\delta_0 K/\gamma\lambda_0^2.$$

Again the bound in (6.7) is summable over i, and by possibly making δ_0 even smaller than heretofore required, the sum can be made less than $(\log^+ n)^{-\rho}$ for some $\rho > 1$ and all n sufficiently large.

In view of (6.3), the preceding proves that for all $\lambda > 0$, there exists $\rho > 1$ and n_0 such that

$$(6.8)\qquad P[\|X_n - Z_n\|_{\mathscr{A}} > 3\lambda] \le (\log^+ n)^{-\rho} \quad \text{for} \quad n > n_0.$$

This in turn shows that for the subsequence $n_k = [v^k]$, $v > 1$,

$$(6.9)\qquad \|X_{n_k} - Z_{n_k}\|_{\mathscr{A}} \to 0 \quad \text{a.s.}$$

By (3.8), for any $\varepsilon > 0$, there exists v sufficiently close to 1 and $k_0 = k_0(\omega)$ such that

$$\max_{n_{k-1}<n\le n_k}\|Z_n - Z_{n_k}\|_{\mathscr{A}}(\omega) < \varepsilon \quad \text{for} \quad k > k_0(\omega).$$

Also, as in (3.7)

$$\max_{n_{k-1}<n\le n_k} \| X_n - X_{n_k} \|_{\mathscr{A}}$$

$$\le \max_{n_{k-1}<n\le n_k} \{ a_{n,k} \| X_{n_k}((n/n_k)\cdot) - X_{n_k}(\cdot) \|_{\mathscr{A}} + | a_{n,k} - 1 | \, \| X_{n_k}(\cdot) \|_{\mathscr{A}} \}$$

which also can be made less than ε for $k > k_1(\omega)$ in view of (6.9). □

7. CLT for partial-sum processes. Just as a Skorokhod-type embedding may be used to give a simple proof of Donsker's invariance theorem in the case of one-parameter Brownian motion (see, for example, Breiman, 1968), our methods can be used to prove a uniform (functional) central limit theorem (CLT) for partial-sum processes indexed by sets. The CLT for these partial sum processes was originally proved by Pyke (1983), using different techniques, and requiring slightly stronger moment conditions than we need here, $2(1 + r)\cdot(1 - r)^{-1}$ instead of $2(1 - r)^{-1}$, where $H(\delta) \sim \delta^{-r}$.

Write

$$\tilde{X}_n(A) = n^{-d/2} \textstyle\sum_j | C_j \cap nA \,|\, Z(\Gamma_j) = b_n X_n(A)$$

and $\tilde{Z}_n(A) = n^{-d/2} Z(nA)$. We will show that $\| \tilde{X}_n - \tilde{Z}_n \|_{\mathscr{A}} \to 0$ in probability. Note that we no longer have the $(L_2 n)^{-1/2}$ to help us in the computations.

Proposition 7.1 *There exists a sequence $\eta_n \to 0$ such that*

$$P(\textstyle\sum_{\mathbf{q}} M_{\mathbf{q}} \ge n^d \eta_n) \to 0.$$

Proof. This follows from the proof of Proposition 5.1 provided we take the η_n's going to 0, yet slowly enough so that $\eta_n \ge (\log n)^{-\beta/3}$ and $\eta_n(\log^+ n^d \eta_n)^{1+\beta} \to \infty$. □

Proposition 7.2. *There exists a sequence $\lambda_n \to 0$ such that*

$$P(\sup_{A\in\mathscr{A}} | \Gamma_n(A) \,\Delta\, C_n(A) | \ge n^d \lambda_n) \to 0$$

and

$$P(\sup_{A\in\mathscr{A}} | \Gamma_n^+(A) \,\Delta\, C_n^+(A) | \ge n^d \lambda_n) \to 0 \quad as \quad n \to \infty.$$

Proof. It is only necessary to modify slightly the proof of Proposition 5.2. The proof of Proposition 5.2 shows that, for n fixed,

(7.1) $$P(\sup_{A\in\mathscr{A}} | \Gamma_n(A) \,\Delta\, C_n(A) | \ge n^d \lambda) \le P(\textstyle\sum_{\mathbf{q}} M_{\mathbf{q}} \ge n^d \lambda / 2m),$$

provided $2\mathbb{c} d^{1/2}/m < \lambda/2$. In (7.1), replace λ by λ_n, m by m_n, where $m_n = [4\mathbb{c} d^{1/2}/\lambda_n] + 1$ ([ ] means "integer part"). Now choose λ_n tending to 0 sufficiently slowly; then λ_n/m_n will tend to 0 slowly enough that Proposition 7.1 may be used. □

Define

$$\tilde{\Lambda}_n := (\sup_{A\in\mathscr{A}} |\Gamma_n(A)\ \Delta\ C_n(A)| \le \varepsilon_n n^d,$$

(7.2)

$$\sup_{A\in\mathscr{A}} |\Gamma_n^+(A)\ \Delta\ C_n(A)| \le \varepsilon_n n^d,\ \textstyle\sum_j |\Gamma_j|^s \le k^s n^d).$$

If $\varepsilon_n \to 0$ slowly enough, $P(\tilde{\Lambda}^c) \to 0$. With no changes in the proof of Proposition 5.4, other than eliminating the b_n's and L_{2n}'s and replacing ε by ε_n, we get

PROPOSITION 7.3. *If* $A, B \in \mathscr{A}$,
(i) $P(|\tilde{X}_n(B) - \tilde{X}_n(A)| > \lambda, \tilde{\Lambda}) \le 2e^{-\lambda^2/2k|A\Delta B|^{1/p}}$, *and*
(ii) *for* $n \ge n_0$ *(independent of* A*)*,

$$P(|\tilde{Z}_n(A) - \tilde{X}_n(A)| > \lambda, \tilde{\Lambda}) \le 4e^{-\lambda^2/12\varepsilon_n}.$$

THEOREM 7.1. *With* $\mathscr{A}$ *and* X_j *satisfying the same conditions as in Theorem* 6.1, *we then have* $\tilde{X}_n \to_L Z$, *where the convergence is in* $C(\mathscr{A})$.

PROOF. The proof is similar to that of Theorem 6.1. Letting $\tilde{\Delta}_n(A) = \tilde{Z}_n(A) - \tilde{X}_n(A)$, it will suffice to prove $\|\tilde{\Delta}_n(\cdot)\| \to 0$ in probability.

Let $\varepsilon_n \to 0$ slowly enough so that $P(\tilde{\Lambda}_n^c) \to 0$. Then, with λ fixed, by Proposition 7.3,

$$P(|\tilde{\Delta}_n(A)| > \lambda, \tilde{\Lambda}_n) \le 4e^{-\lambda^2/12\varepsilon_n} \to 0 \quad \text{as} \quad n \to \infty.$$

Choose $\delta_0(n) \to 0$ slowly enough so that $e^{H(\delta_0(n))}e^{-\lambda^2/12\varepsilon_n}$ still converges to 0. Then

$$P(\|\tilde{\Delta}_n\|_{\mathscr{A}_{\delta_0(n)}} > \lambda, \tilde{\Lambda}_n) \to 0.$$

Let $0 < \gamma < 1$ be fixed, $0 < b < \frac{1}{2}(1/p - r)$, $\delta_i(n) = \delta_0(n)\gamma^i$, and $\lambda_i = \lambda_0\gamma^{bi}$, where λ_0 is chosen so that $\sum_{i=0}^{\infty} \lambda_i \le \lambda$.

Writing $\tilde{\Delta}_n(A) = \tilde{\Delta}_n(A_{\delta_0(n)}) + \sum_{i=1}^{\infty} \{\tilde{\Delta}_n(A_{\delta_i(n)}) - \tilde{\Delta}_n(A_{\delta_{i-1}(n)})\}$ with $A_{\delta_i(n)} \in \mathscr{A}_{\delta_i(n)}$, $|A_{\delta_i(n)}\ \Delta\ A| < \delta_i$, it sufficies to show

$$\textstyle\sum_{i=1}^{\infty} P(\max_{A\in\mathscr{A}} |\tilde{X}_n(A_{\delta_i(n)}) - \tilde{X}_n(A_{\delta_{i-1}(n)})| > \lambda_i, \tilde{\Lambda}_n) \to 0 \quad \text{as} \quad n \to \infty, \tag{7.3}$$

and

$$\textstyle\sum_{i=1}^{\infty} P(\max_{A\in\mathscr{A}} |\tilde{Z}_n(A_{\delta_i(n)}) - \tilde{Z}_n(A_{\delta_{i-1}(n)})| > \lambda_i, \tilde{\Lambda}_n) \to 0 \quad \text{as} \quad n \to \infty. \tag{7.4}$$

We will prove (7.3), (7.4) being similar but simpler.

As in the proof of Theorem 6.1, the summands in (7.3) are bounded by (here we use Proposition 7.3)

$$\begin{aligned} &2\exp\{2H(\delta_i(n)) - \lambda_i^2/2k(2\delta_{i-1}(n))^{1/p}\} \\ &\quad = 2\exp\{-\lambda_i^2/2k(2\delta_{i-1}(n))^{1/p}[1 - 2k(2\delta_{i-1}(n))^{1/p}H(\delta_i(n))/\lambda_i^2]\} \qquad (7.5)\\ &\quad \le 2\exp\{-\lambda^2/4k(2\delta_{i-1}(n))^{1/p}\}, \end{aligned}$$

for n sufficiently large, since

$$2k(2\delta_{i-1}(n))^{1/p}H(\delta_i(n))/\lambda_i^2 \le kK2^{1+1/p}\gamma^{-1/p}(\delta_0(n))^{1/p}(\gamma^{1/p-2b-r})^i/\lambda_0^2$$

converges to 0 as $n \to \infty$, uniformly in i.

Using the fact that

$$\lambda_i^2/4k(2\delta_{i-1}(n))^{1/p} = \{\lambda_0^2/4k(2\delta_0(n))^{1/p}\gamma^{-1/p}\}(\gamma^{2b-1/p})^i,$$

that $\gamma^{2b-1/p} > 1$ (since $\gamma < 1$ and $2b - 1/p < 0$), and that

$$\gamma^{(2b-1/p)i} = (\gamma^{2b-1/p} - 1 + 1)^i \ge i(\gamma^{2b-1/p} - 1),$$

we may sum (7.5) over i to get

$$\begin{aligned}\sum_{i=1}^{\infty} P(\max_{A\in\mathscr{A}} |\tilde{X}_n(A_{\delta_i(n)}) - \tilde{X}_n(A_{\delta_{i-1}(n)})| > \lambda_i, \tilde{\Lambda}_n)\\ \le \sum_{i=1}^{\infty} 2\exp\{-i\lambda_0^2(\gamma^{2b-1/p} - 1)/4k(2\delta_0(n))^{1/p}\gamma^{-1/p}\},\end{aligned}$$

a geometric series whose sum tends to 0 as $n \to \infty$ since $\delta_0(n) \to 0$ as $n \to \infty$. □

8. Remarks. 1. Although the independence of the $X_{\mathbf{j}}$'s is crucial for the methods of this paper, the only place where we use the fact that the $X_{\mathbf{j}}$'s are identically distributed is in deriving some of the probability estimates of Section 5. For these results one really only requires that the tails of the $X_{\mathbf{j}}$'s be uniformly bounded by appropriate quantities, in which case the necessary modifications are quite straightforward. A much more challenging and useful modification would be to replace the independence assumption by a suitable mixing condition that would provide approximate independence to $X_{\mathbf{j}}$ and $X_{\mathbf{k}}$ when $\mathbf{j}$ and $\mathbf{k}$ are sufficiently far apart. The methods of Phillip and Stout (1975) should be applicable here.

2. One might ask what functions of n could replace $b_n = (2 \log \log n)^{1/2}$ so that Theorems 3.1 and 6.1 remain true. One would expect that there is an integral test that distinguishes between upper and lower class functions for the LIL in our context. The abstract by Bulinskiĭ (1978) might be relevant in this connection, as might the paper by Carmona and Kôno (1976).

3. The embedding described in Section 4 allowed us to prove invariance principles that were strong enough to yield both the functional LIL and the uniform CLT. Moreover, as by-products of the proof of Theorem 6.1 we obtain probability estimates on the tail of the distribution of $\| Z_n(\cdot) - X_n(\cdot) \|_{\mathscr{A}}$. One might ask whether one could get faster rates of convergence if one used a method of embedding that did not use stopping times. For example, some of the methods of Komlos, Major and Tusnday (see Csörgő and Revesz, 1981) might be applicable to this situation, at least for the LIL. Such application has been made by Dudley and Philipp (1983) for the case of linearly ordered sums of Banach space valued random elements.

A further question concerning embeddings is whether or not any other stopping-time embedding might require weaker moment conditions than those of this paper. For example, our embedding associates with each unit cube $C_{\mathbf{j}}$, a random rectangular interval $\Gamma_{\mathbf{j}}$, only one dimension of which differs from that of $C_{\mathbf{j}}$. One can easily visualize configurations of random sets, some of which may possibly yield closer fits.

4. Our theorems were for the real-valued case; that is, where the ranges of Z and X are $\mathbb{R}^1$. Theorem 3.1 could be extended to the case $Z: \mathscr{A} \to \mathbb{R}^m$ with no essential change. This is not true for Theorem 6.1 by the nature of the embedding. To prove Theorem 6.1 for vector-valued random variables, one could attempt to use more classical techniques such as those of Hartman and Wintner (1941), as was done in Wichura (1973).

5. According to property (v′) in Section 2, we have focused attention on large index families $\mathscr{A}$ that satisfy (2.2) with $0 < r < 1$. This leaves two cases of interest which we will loosely refer to as $r = 0$ and $r = 1$. The key entropy condition is the integrability condition (2.1). The $r = 1$ case involves those functions H for which $\{H(u)/u\}^{1/2}$ is integrable but (2.2) does not hold for $r < 1$. Since our moment condition involves $s = 2(1 - r)^{-1}$, it is clear that in the $r = 1$ case, our results imply that all moments of $X_{\mathbf{j}}$ must be finite in order for the LIL and CLT to hold. It would be interesting to determine whether this is sufficient, or whether stronger assumptions on the moment generating function are needed.

The $r = 0$ case includes the usual smaller families (e.g. orthants, spheres, polytopes with a bounded number of vertices). Although our results apply to this case to yield LIL and CLT, the interesting question is whether or not the moment condition can be weakened. Our methods require a $2 + \delta$ moment but the results probably require a condition much closer to the second moment only. In this regard, compare the LIL for orthants by Wichura (1973).

6. It should be observed that for the Central Limit theorem obtained in Pyke (1983), the Hausdorff metric was used, rather than the symmetric difference metric of this paper, to determine the entropy of $\mathscr{A}$. Also, recall that in Pyke (1983), a stronger moment condition based on $2(1 + r)(1 - r)^{-1}$ rather than $2(1 - r)^{-1}$ was imposed. It remains to find the optimal moment conditions that are possible for a fixed metric and entropy.

7. As stated at the outset, it is necessary to work with the smoothed partial-sum process X rather than the unsmoothed S. The reason for this can be seen by considering a family of smooth sets in I^2 which includes sets whose upper boundaries are slight perturbations of the line $y = 1/2$. In this way, every subsum of the masses $\{X_{\mathbf{j}}: j_2 = n/2\}$ can arise as $S(A \,\Delta\, B)$ for A and B very close together. This prevents one from having the desired continuity for most asymptotic results. This is discussed further in Pyke (1983) and Erickson (1981). By adding further structure to the index families, possibly making $C_N^*(A) = \cup \{n^{-1}C_{\mathbf{j}}: \mathbf{j} \in nA\}$, rather than $C_n(A)$ and $C_n^+(A)$, the key fit to A, one might be able to obtain results for S. Alternatively, the use of a metric different from d_L may suffice to relate adequately the smoothness of sets to the lattice locations of the $X_{\mathbf{j}}$.

8. We have focused in this paper upon set-indexed processes. It is of interest also to consider the case where the processes are indexed by functions. If F is a family of real-valued functions defined on R^d, one could define processes $\{T(f): f \in \mathscr{F}\}$ and $\{Y(f): f \in \mathscr{F}\}$ by

$$T(f) = \int f\, dS \text{ and } Y(f) = \int f\, dX$$

where both integrals are over $[0, \infty)^d$. One may also define the analogous integral processes on I^d by appropriately scaling back the S and X processes. The processes of this paper would be special cases of these in which $\mathscr{F}$ is a class of indicator functions. The general methods of this paper may be used to obtain corresponding results for the function-indexed case. For related results concerning empirical processes, see Strassen and Dudley (1969), Pollard (1981) and LeCam (1983).

Acknowledgement. We wish to thank the referees for their suggestions, and in particular, for the reference to the result of Carmona and Kôno (1976) which enabled us to shorten our original proof of Theorem 3.1. The latter proof involved an extension of the techniques of Strassen (1964) to this context, analogous to the extension to Brownian sheets by Pyke (1973); cf. Bass and Pyke (1982). A further benefit of the present approach is that assumption (iv) on boundary smoothness is no longer required for Theorem 3.1.

REFERENCES

[1] BASS, R. F. (1983). Skorokhod imbedding via stochastic integrals. Séminaire de Probabilités XVII, *Lecture Notes in Math.* **986** 221–224. Springer, New York.
[2] BASS, R. F. and PYKE, R. (1982). Functional law of the iterated logarithm and uniform central limit theorem for partial-sum processes indexed by sets. Technical Report No. 37, Department of Mathematics, University of Washington.
[3] BASS, R. F. and PYKE, R. (1984). A strong law of large numbers for partial-sum processes indexed by sets. *Ann. Probab.* **12** 13–34.
[4] BREIMAN, L. (1968). *Probability.* Addison-Wesley, Reading, Mass.
[5] BULINSKIĬ, A. V. (1978). On the functional law of the iterated logarithm. *Dokl. Akad. Nauk* SSSR **238** 1025–1027 (*Soviet Math. Dokl.* **19** 141–143).
[6] CARMONA, R. and KÔNO, K. (1976). Convergence en loi et lois du logarithme itéré pour les vecteurs gaussiens. *Z. Wahrsch. verw. Gebiete* **36** 241–267.
[7] CSÖRGŐ, M. and REVÉSZ, P. (1981). *Strong Approximations in Probability and Statistics.* Academic, New York.
[8] DELLACHERIE, C., DOLÉANS-DADE, C., and MEYER, P. A. (1970). Diffusions à coefficients continus d'après D. W. Stroock et S. R. S. Varadhan. Seminaire de Probabilités IV, *Lecture Notes in Math.* **124** 240–282. Springer, New York.
[9] DUBINS, L. E. (1968). On a theorem of Skorokhod. *Ann. Math. Statist.* **39** 2094–2097.
[10] DUDLEY, R. M. (1973). Sample functions of the Gaussian process. *Ann. Probab.* **1** 66–103.
[11] DUDLEY, R. M. and PHILIPP, WALTER (1983). Invariance principles for sums of Banach space valued random elements and empirical processes. *Z. Wahrsch. verw Gebiete* **62** 509–552.
[12] ERICKSON, R. V. (1981). Lipschitz smoothness and convergence with applications to the central limit theorem for summation processes. *Ann. Probab.* **9** 831–851.
[13] HARTMAN, P. and WINTNER, A. (1941). On the law of the iterated logarithm. *Amer. J. Math.* **63** 169–176.
[14] HEYDE, C. C. and ROHATGI, V. K. (1967). A pair of complementary theorems on convergence rates in the law of large numbers. *Proc. Camb. Phil. Soc.* **63** 73–82.
[15] KIEFER, J. (1972). Skorokhod embedding of multivariate RV's, and the sample DF. *Z. Wahrsch. verw. Gebiete* **24** 1–35.
[16] LECAM, L. (1983). A remark on empirical measures. *A Festschrift for Erich L. Lehmann in Honor of his 65th Birthday.* (Edited by P. Bickel, K. Doksum, and J. Hodges.) Wadsworth, Belmont, Calif.
[17] MEYER, P. A. (1976), Un cours sur les intégrales stochastiques. Séminaire de Probabilités X, *Lecture Notes in Math.* **511** 245–400. Springer, New York.

[18] MONROE, I. (1971). Using additive functionals to embed preassigned distributions in symmetric stable processes. *Trans. Amer. Math. Soc.* **163** 131–146.

[19] PHILIPP, W. and STOUT, W. (1975). Almost sure invariance principles for partial sums of weakly dependent random variables. *Memoirs A.M.S.* **2** No. 161, pages 140.

[20] POLLARD, D. (1981). A central limit theorem for empirical processes. *J. Australian Math. Soc.* Ser. A. **33** 235–248.

[21] PYKE, R. (1973). Partial sums of matrix arrays, and Brownian sheets. *Stochastic Analysis*, ed. by D. G. Kendall and E. F. Harding, 331–348. Wiley, London.

[22] PYKE, R. (1983). A uniform central limit theorem for partial-sum processes indexed by sets. Probab. Statist. and Anal. (Edited by J.F.C. Kingman and G.E.H. Reuter.) *London Math. Soc. Lect. Notes Series* **79** 219–240.

[23] ROOT, D. H. (1969). The existence of certain stopping times on Brownian motion. *Ann. Math. Statist.* **40** 715–718.

[24] ROST, H. (1976). Skorokhod stopping times of minimal variance. Seminaire de Probabilities X, *Lecture Notes in Math.* **511** 194–208, Springer, New York.

[25] SKOROKHOD, A. V. (1965). *Studies in the theory of random processes.* Addison, Reading, Mass.

[26] STRASSEN, V. (1964). An invariance principle for the law of iterated logarithm. *Z. Wahrsch. verw. Gebiete* **3** 211–226.

[27] STRASSEN, V. and DUDLEY, R. M. (1969). The central limit theorem and ε-entropy. *Lecture Notes in Math.* **89** 224–231. Springer, Berlin.

[28] WICHURA, M. J. (1973). Some Strassen-type laws of the iterated logarithm for multiparameter stochastic processes with independent increments. *Ann. Probab.* **1** 272–296.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON 98195