A CENTRAL LIMIT THEOREM FOR SET-INDEXED STRONG MARTINGALES

Dean Slonowsky *
The Fields Institute
222 College Street
Toronto, CANADA M5T 3J1
dslonows@fields.utoronto.ca

B. Gail Ivanoff †
Department of Mathematics and Statistics
University of Ottawa
P.O. Box 450, Stn. A
Ottawa, CANADA K1N 6N5
givanoff@uottawa.ca

July 1999

*The results in this paper represent part of the author's Ph.D. dissertation under the direction of B. G. Ivanoff at the University of Ottawa. Research partially supported by an NSERC PGS-A Scholarship and an Ontario Graduate Scholarship.

†Research supported by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC).

Abstract

We define a generic form of quadratic variation for set-indexed strong martingales and define a new mode of convergence for set-indexed processes which places no explicit restrictions on the the size of the indexing class. Under this mode, we derive a Central Limit Theorem (CLT) for set-indexed strong martingales. As applications, we obtain a general CLT for sequences of d-dimensional strong martingales as well as a CLT for set-indexed weighted empirical processes.

AMS 1991 subject classification: Primary 60F05, 60G48; Secondary 60F17, 60G60.

Keywords and Phrases: set-indexed processes, strong martingales, quadratic variation, flows, central limit theorem, weighted empirical processes.

Running Head: Strong Martingale Central Limit Theorems

1 Introduction.

In [16], Ivanoff and Merzbach introduced the notion of a set-indexed strong martingale, a generalization of the multiparameter strong martingales defined by Cairoli and Walsh in [10]. Among other things, a set-indexed strong martingale is a set-indexed process, that is, a collection $X = \{X_A : A \in \mathcal{A}\}$ of random variables where $\mathcal{A}$ is a suitable class of compact subsets of a fixed topological space T. In this paper, we give conditions under which a sequence of set-indexed strong martingales converges in some sense to a set-indexed Gaussian process.

Employing the general theory of weak convergence in metric spaces, various authors have studied functional (i.e., weak) convergence for sequences of set-indexed partial-sum processes, $X_n = \{X_n(A) : A \in \mathcal{A}\}$, $n \in \mathbf{N}$. We highlight several such attempts. In [3] and [22], conditions were given ensuring functional convergence of smoothed versions of the X_n to an $\mathcal{A}$-indexed Gaussian process with respect to the uniform topology. When each X_n possessed a certain set-indexed form of cadlaguity, Bass and Pyke (cf. [5, 6]) gave conditions implying functional convergence to an $\mathcal{A}$-indexed Lévy process with respect to a topology which was weaker than the uniform topology.

In the above works, $\mathcal{A}$ had to be sufficiently small with respect to the notion of metric entropy (for example, $\mathcal{A}$ a Vapnik Cĕrvonenkis class) so as to ensure the existence of path-regular versions of the limiting process. Indeed, when $\mathcal{A}$ is the class of *lower layers* in $T = [0,1]^d$, $d \geq 2$, and W is any $\mathcal{A}$-indexed Gaussian process, it has been shown in [1] that $\sup_{A \in \mathcal{A}} |W_A| = \infty$ a.s., ensuring failure of the cited methods to yield functional central limit theorems (CLTs). For this reason, we devise an entirely new mode of convergence for set-indexed processes termed *flow-wise convergence.* This mode circumvents the said size restrictions, permitting CLTs for strong martingales indexed by a class $\mathcal{A}$ as large as the lower layers (cf. Proposition 7.2).

The paper is divided into six sections. Section 2 presents the general framework. The core assumptions on $\mathcal{A}$ found therein are close to those in [17]. This section also contains the definition of set-indexed strong martingales and the limiting Gaussian processes. Section 3 introduces *flows* and *simple flows.* Flows, which are functions of the form $f : [0,1] \to \{$finite unions in $\mathcal{A}\}$, have appeared earlier in [17]. Our interest in flows lies in their ability to transform set-indexed objects into continuous parameter objects (i.e., indexed by $[0,1]$ or $\mathbf{R}_+$), and thereby eliminating the need for a metric on T.

Section 4 introduces the flow-wise mode of convergence. In this mode, rather than establishing functional convergence over the entire class $\mathcal{A}$, one establishes functional convergence over each range $\{f(t) : t \in [0, 1]\}$ where f is a simple flow. Since each such range is trivially a Vapnik Cěrvonenkis class, it is no longer necessary to place size restrictions on $\mathcal{A}$, nor is T required to be metrizable. Furthermore, since the union of all such ranges captures $\mathcal{A}$ (cf. Lemma 3.3), flow-wise convergence to a Gaussian process implies convergence in finite dimensional distribution (cf. Proposition 4.3).

Section 5 introduces the notion of **-quadratic variation* for set-indexed strong martingales, a generalization of the usual quadratic variation for continuous parameter martingales. Unlike its continuous parameter counterpart, a *-quadratic variation is not required to be adapted. This omission increases the flexibility of the theory as demonstrated by Proposition 5.6 wherein a *-quadratic variation for a set-indexed weighted empirical process is calculated. The role of *-quadratic variation in flow-wise CLTs is highlighted below.

The main result of the paper, a flow-wise CLT for set-indexed strong martingales, is presented in Section 6. The general principle is this: given a sequence (X_n) of strong martingales with corresponding *-quadratic variations (X_n^*), under certain moment conditions, asymptotic rarefaction of jumps in the sample paths of the X_n plus convergence of (X_n^*) to a continuous, deterministic limit implies flow-wise convergence of (X_n) to an appropriately scaled Gaussian process W. As an application, a flow-wise CLT for weighted empirical processes indexed by the class of lower layers is derived. In this particular example, almost every sample path of the limiting Gaussian process is discontinuous everywhere (cf. Remark 7.5). In addition, the general flow-wise CLT yields a new functional CLT for d-dimensional strong martingales with sample paths in the Skorokhod space $D([0, 1]^d)$ (cf. Proposition 7.1), which appears to be new even in the case $d = 1$.

2 Preliminaries.

In this paper, $\subset$ will denote strict inclusion whereas *increasing* or *decreasing* sequences of sets are understood to be monotone non-decreasing, respectively monotone non-increasing with respect to $\subseteq$. Fix a Hausdorff topological space T. Given any $S \subseteq T$, denote the closure of S by $\overline{S}$, the interior of

S by S° and the boundary of S by ∂S. A sequence (S_n) of subsets of T is said to be *bounded* if for some compact subset K of T, $S_n \subseteq K$ for every n. Throughout, $\mathcal{A}$ will denote a collection of compact (hence closed) subsets of T satisfying the following conditions:

(a) $\phi \in \mathcal{A}$,

(b) there exists an increasing sequence (B_n) in $\mathcal{A}$ with $\bigcup_n B_n = T$ such that for any $A \in \mathcal{A}$, $A \subseteq B_n$ all n sufficiently large,

(c) $\mathcal{A}$ is closed under countable intersections,

(d) (A_n) increasing and bounded in $\mathcal{A}$ implies $\overline{\bigcup_n A_n} \in \mathcal{A}$ and

(e) if $A, B \in \mathcal{A}$ are such that $A, B \neq \phi$, then $A \cap B \neq \phi$.

To any such $\mathcal{A}$ we can associate three additional families: $\mathcal{A}(u)$ which consists of all finite unions in $\mathcal{A}$, $\mathcal{C}$ which consists of all set differences of the form $A \backslash B$ ($A \in \mathcal{A}, B \in \mathcal{A}(u)$) and $\mathcal{C}(u)$ which consists of all finite unions in $\mathcal{C}$. $\mathcal{A}$ is a semilattice under the partial order relation induced by set-inclusion and for this reason, any finite subcollection of $\mathcal{A}$ which is closed under intersections will be referred to as a finite sub-semilattice. Note that $\mathcal{C}$ is a semiring of subsets of T so that $\mathcal{C}(u)$ constitutes an ring. Clearly, $\mathcal{A} \subseteq \mathcal{C}$ and $\mathcal{A}(u) \subseteq \mathcal{C}(u)$.

The following separability assumption has appeared in [17]. It will be in force throughout the paper.

Assumption 2.1 There exists an increasing sequence $(\mathcal{A}_n)$ of finite sub-semilattices of $\mathcal{A}$ and a sequence (g_n) of functions of the form $g_n : \mathcal{A} \to \mathcal{A}_n(u) \cup \{T\}$ ($\mathcal{A}_n(u)$ denotes the collection of all finite unions in $\mathcal{A}_n$) such that given $A, A' \in \mathcal{A}$,

(A1) $(g_n(A))$ is decreasing with $\bigcap_n g_n(A) = A$,

(A2) $A \subseteq [g_n(A)]^\circ$,

(A3) $A \subset A'$ implies $A \subset g_n(A) \cap A'$,

(A4) $A \cup A' \in \mathcal{A}$ implies $g_n(A \cup A') = g_n(A) \cup g_n(A')$ and

(A5) g_n preserves countable intersections (i.e., $g_n(\bigcap_{i=1}^{\infty} A_i) = \bigcap_{i=1}^{\infty} g_n(A_i)$ for any (A_i) in $\mathcal{A}$).

Example 2.2 (a) When $T = \mathbf{R}_+^d = [0,\infty)^d$, some $d \in \mathbf{N}$, examples of $\mathcal{A}$ satisfying Assumption 2.1 include the *lower rectangles* $\{R_z : z \in \mathbf{R}_+^d\}$ where $R_z := \prod_{i=1}^d [0, z_i]$ for $z = (z_1, \cdots, z_d) \in \mathbf{R}_+^d$ and the collection $\mathcal{LL}_d$ of *lower layers* in $\mathbf{R}_+^d$. A compact set $L \subseteq \mathbf{R}_+^d$ is a lower layer if $z \in L$ implies $R_z \subseteq L$. Both examples can be extended to the case of $T = \mathbf{R}^d$ (cf. [17]).

(b) Let T be a tree with a finite number of edges. Embed T in $\mathbf{R}^2$ so that T is rooted at the origin and each edge is a line segment. $\mathcal{A}$ can be taken to be the collection of all closed connected subsets of T, i.e., $\mathcal{A}$ consists of all "continuous subtrees" of T. Alternately, $\mathcal{A}$ can be the collection of all $A_t \subseteq T$ $(t \in T)$ where A_t is the unique path in T from the origin to t.

(c) An example where T is a function space can be found in [11].

Given $n \in \mathbf{N}$, define the function $\tilde{g}_n : \mathcal{A}(u) \to \mathcal{A}_n(u) \cup \{T\}$ by $\tilde{g}_n(B) = \bigcup_{A \in \mathcal{A},\, A \subseteq B} g_n(A)$, $B \in \mathcal{A}(u)$. Since $\tilde{g}_n(A) = g_n(A)$ for every $A \in \mathcal{A}$, we denote this extension by g_n as well. By (A4), g_n preserves finite unions in $\mathcal{A}(u)$ so that

$$g_n(\cup_{i=1}^k A_i) = \cup_{i=1}^k g_n(A_i) \tag{1}$$

for any $k \in \mathbf{N}$ and $A_1, \cdots, A_k \in \mathcal{A}$.

To play the role of 0 in the classical theory, we define

$$\phi' = \bigcap_{A \in \mathcal{A},\, A \neq \phi} A\,.$$

By (A1), (e), and the compactness of elements in $\mathcal{A}$, ϕ' is a non-empty element of $\mathcal{A}$. In fact, (A1) implies $\mathcal{A}$ is closed under arbitrary intersections. Without loss of generality, we can assume all finite sub-semilattices of $\mathcal{A}$ contain ϕ' but do not contain ϕ.

Given any finite sub-semilattice $\mathcal{A}'$ of $\mathcal{A}$, the sets in $\mathcal{A}'$ can always be numbered $A_0, \cdots, A_k$ for some $k \in \mathbf{N}$ so that $A_0 = \phi'$ and, given any $1 \le i \le k$, $A_j \subset A_i$ implies $0 \le j \le i-1$. Following [17], we refer to any such numbering as being *consistent with the strong past.*

Given a finite sub-semilattice $\mathcal{A}'$ of $\mathcal{A}$ and an element $A \in \mathcal{A}'$, the *left-neighborhood* of A in $\mathcal{A}'$ is defined to be the set

$$C_A := A \setminus \cup_{A' \in \mathcal{A}',\, A \not\subseteq A'} A' \tag{2}$$

so that $C_{\phi'} = \phi'$. Since $\mathcal{A}'$ is finite and closed under intersections, $C_A = A \setminus \bigcup_{A' \in \mathcal{A}', A' \subset A} A'$ for each $A \in \mathcal{A}'$. Furthermore, if $\{A_0, \cdots, A_k\}$ is a numbering of $\mathcal{A}'$ consistent with the strong past, we clearly have $C_{A_i} = A_i \setminus \bigcup_{j=1}^{i-1} A_j$ for each $1 \le i \le k$.

We now introduce a probability structure. Given a complete probability space $(\Omega, \mathcal{F}, P)$, an $\mathcal{A}$-*indexed filtration* is any increasing family $(\mathcal{F}_A) = \{\mathcal{F}_A : A \in \mathcal{A}\}$ of complete sub-σ-algebras of $\mathcal{F}$ for which

$$\mathcal{F}_A = \bigcap_n \mathcal{F}_{A_n} \tag{3}$$

for any decreasing sequence (A_n) in $\mathcal{A}$ with $\bigcap_n A_n = A$. (3) determines a form of right-continuity for the filtration. Any $\mathcal{A}$-indexed filtration can be extended to an $\mathcal{A}(u)$-indexed family by defining

$$\mathcal{F}_B = \bigcap_k \mathcal{F}^{\circ}_{g_k(B)} \tag{4}$$

for any $B \in \mathcal{A}(u)$ where $\mathcal{F}^{\circ}_B = \bigvee_{A \in \mathcal{A}, A \subseteq B} \mathcal{F}_A$. Given a set $C \in \mathcal{C}(u)$, the *strong past* at C is then defined to be the sub-σ-algebra

$$\mathcal{G}^*_C = \bigvee_{\substack{B \in \mathcal{A}(u) \\ B \cap C = \phi}} \mathcal{F}_B \tag{5}$$

for $C \notin \mathcal{A}(u)$ and $\mathcal{G}^*_C = \mathcal{F}_{\phi'}$ for $C \in \mathcal{A}(u)$.

Any collection $X = \{X_A : A \in \mathcal{A}\}$ of random variables is referred to as a *set-indexed process*. When more convenient, we will write $X(A)$ for X_A. In many cases, such a process can be uniquely extended to a finitely additive $\mathcal{C}(u)$-indexed process, i.e., $X_{C \cup D} = X_C + X_D$ for any two disjoint sets $C, D \in \mathcal{C}(u)$. A necessary condition for such an extension is $X_\phi = 0$. Sufficient conditions can be found in [25]. Note that any process X possessing a finitely additive extension to $\mathcal{C}(u)$ necessarily satisfies the inclusion-exclusion formula,

$$X_C = \sum_{I \subseteq \{1, \cdots, k\}} (-1)^{|I|} \cdot X(A \cap \bigcap_{i \in I} A_i) \tag{6}$$

for any $C = A \setminus \bigcup_{i=1}^k A_i$ where $A, A_i \in \mathcal{A}$ and $k \in \mathbf{N}$. In the sequel, unless otherwise stated, we will only work with set-functions and set-indexed processes which possess unique finitely additive extensions to $\mathcal{C}(u)$.

Given an $\mathcal{A}$-indexed process X and $1 \le p \le \infty$, if $X_A \in L_p \; \forall A \in \mathcal{A}$ we say X is in L_p. For the case of $p = 1$, such an X is said to be *integrable*. As in the continuous parameter setting, a process $X = \{X_A : A \in \mathcal{A}\}$ is adapted to an $\mathcal{A}$-indexed filtration $(\mathcal{F}_A)$ if X_A is $\mathcal{F}_A$-measurable $\forall A \in \mathcal{A}$.

Definition 2.3 *An adapted integrable process $X = \{X_A : A \in \mathcal{A}\}$ is a strong martingale if $E[X_C|\mathcal{G}_C^*] = 0$ for every $C \in \mathcal{C}$.*

Comments. (i) The definition of a set-indexed strong martingale first appeared in [16] and is a generalization of the planar strong martingales introduced in [10]. Examples will be given in Section 5.

(ii) As mentioned in [17], if X is a strong martingale, then the relation $E[X_C|\mathcal{G}_C^*] = 0$ extends to all $C \in \mathcal{C}(u)$.

In the sequel, we will make extensive use of the theory of continuous parameter martingales, $X = (X_t) = \{X_t : t \in [0,1]\}$. Throughout, a *filtration* $(\mathcal{F}_t) = \{\mathcal{F}_t : t \in [0,1]\}$ is understood to be increasing, complete and right-continuous in the sense: $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s \ \forall t \in [0,1)$. We let $C[0,1]$ denote the collection of all continuous functions $u : [0,1] \to \mathbf{R}$. A process with sample paths in $C[0,1]$ is termed *continuous*. In addition, we let $D[0,1]$ denote the collection of all functions $x : [0,1] \to \mathbf{R}$ which are right-continuous with left limits on $[0,1]$ and define $\Delta x(t) = x(t) - x(t-) \ \forall t \in [0,1]$ with the convention $x(0-) = x(0)$. The elements of $D[0,1]$ (or processes with paths in $D[0,1]$) will occasionally be termed *cadlag*. The *jump functional*, $J : D[0,1] \to [0,\infty)$ defined by

$$J(x) = \sup_{0 \le t \le 1} |\Delta x(t)|, \quad x \in D[0,1]$$

is measurable with respect to the Skorokhod J_1 topology and is continuous at every $x \in D[0,1]$ for which $\Delta x(1) = 0$ (cf. [19], p.303). An integrable process $X = (X_t)$ will be called *increasing* if its sample paths are non-negative, right-continuous and increasing. Set-indexed analogues of these sample path properties are defined below.

Definition 2.4 *Let $X = \{X_A : A \in \mathcal{A}\}$ be a process.*

(i) *X is monotone outer-continuous if there is a set Ω' of full P-measure such that (A_n) decreasing in $\mathcal{A}$ implies $X(A_n) \to X(\bigcap_n A_n)$ on Ω'.*

(ii) *X is monotone inner-continuous if there is a set Ω' of full P-measure such that for any increasing and bounded sequence (A_n) in $\mathcal{A}$, $X(A_n) \to X(\overline{\bigcup_n A_n})$ on Ω'.*

(iii) *X is increasing if it is integrable, monotone outer-continuous and $X_C \geq 0$ a.s. $\forall C \in \mathcal{C}$.*

Comment. When $T = \mathbf{R}_+$ and $\mathcal{A} = \{[0,t] : t \in \mathbf{R}_+\}$, outer-continuity reduces to right-continuity and inner-continuity reduces to left-continuity under the identification $t \leftrightarrow [0,t]$.

The key result of this paper, Theorem 6.1, gives conditions under which a sequence of set-indexed strong martingales converges to an appropriately scaled set-indexed Gaussian process. The class of limiting Gaussian processes is defined below.

Definition 2.5 $\Lambda : \mathcal{A} \to [0,\infty)$ *is said to be a variance function on $\mathcal{A}$ if it is increasing, monotone inner- and outer-continuous and $\Lambda(\phi') = 0$. Given a variance function Λ, a process $W = \{W_A : A \in \mathcal{A}\}$ is said to be a Gaussian white noise based on Λ provided*

(g1) $W_C \sim N(0, \Lambda(C))$ $\forall C \in \mathcal{C}(u)$,

(g2) $E(W_A W_B) = \Lambda(A \cap B)$ $\forall A, B \in \mathcal{A}$ *and*

(g3) $C, D \in \mathcal{C}(u)$ *disjoint implies W_C and W_D are independent.*

Comments. (i) As mentioned on p.910 of [17], the existence of set-indexed Gaussian white noise with a specified variance function is guaranteed by Kolmogorov's extension theorem.

(ii) By (4) and (g3), an $\mathcal{A}$-indexed Gaussian white noise W is a strong martingale with respect to its minimal filtration if $W(A_n) \to W(\bigcap_n A_n)$ in L_1 for every decreasing sequence (A_n) in $\mathcal{A}$.

3 Flows.

The present section concerns special paths in $\mathcal{A}(u)$ termed *flows* which will play a key role in the upcoming limit theorems. Our definition of flow originates from [17].

Definition 3.1 *Given $a < b$ in $\mathbf{R}$, a function $f : [a,b] \to \mathcal{A}(u)$ is said to be a flow provided*

(i) $a \leq t \leq s \leq b$ *implies* $f(t) \subseteq f(s)$,

(ii) $f(s) = \bigcap_{v>s} f(v) \ \forall\, a \leq s < b$ *and*

(iii) $f(s) = \overline{\bigcup_{v<s} f(v)} \ \forall\, a < s \leq b$.

Condition (i) states that f is increasing on $[a, b]$ with respect to $\subseteq$ whereas conditions (ii) and (iii) determine a form of right- (respectively, left-)continuity for f. In the present paper, we will be primarily interested in the following class of flows which are locally $\mathcal{A}$-valued modulo a fixed set in $\mathcal{A}(u)$.

Definition 3.2 *A flow $f : [0, 1] \to \mathcal{A}(u)$ is simple provided there is a finite partition $0 = t_0 < t_1 < \cdots < t_k = 1$ of $[0, 1]$ and corresponding flows*

$$f_i : [t_{i-1}, t_i] \to \mathcal{A}\,, \quad i = 1, \cdots, k$$

such that $f(0) = \phi'$ and for each fixed $1 \leq i \leq k$,

$$f(t) \ = \ [\textstyle\bigcup_{j=1}^{i-1} f_j(t_j)] \cup f_i(t) \quad \forall\, t \in [t_{i-1}, t_i]\,. \tag{7}$$

The collection of all simple flows is denoted $S(\mathcal{A})$.

Since the domain of any flow can be linearly rescaled, we can always take $t_i = i/k$ in Definition 3.2. The next result, which appears as Lemma 3 in [17], illustrates the richness of $S(\mathcal{A})$.

Lemma 3.3 *If $\mathcal{A}' = \{A_0, \cdots, A_k\}$ is a finite sub-semilattice of $\mathcal{A}$ numbered in a manner consistent with the strong past, then there exists $f \in S(\mathcal{A})$ such that*

(a) $f(i/k) = \bigcup_{j=0}^{i} A_j \ \forall\, 0 \leq i \leq k$ *and*

(b) $C_{A_i} = f(i/k) \setminus f((i-1)/k) \ \forall\, 1 \leq i \leq k$.

The following three lemmas illustrate the effect of simple flows on various set-indexed elements. These properties will be required in the sequel.

Lemma 3.4 *Fix $f \in S(\mathcal{A})$. If $(\mathcal{F}_A)$ is an $\mathcal{A}$-indexed filtration, then $(\mathcal{F}_{f(t)}) = \{\mathcal{F}_{f(t)} : t \in [0, 1]\}$ is a filtration. Moreover, given an $\mathcal{A}$-indexed process X,*

(a) *X adapted to $(\mathcal{F}_A)$ implies $X \circ f$ is adapted to $(\mathcal{F}_{f(t)})$ and*

(b) *X a strong martingale implies $X \circ f$ is a martingale.*

Here, $X \circ f$ denotes the process $\{X_{f(t)} : t \in [0,1]\}$.

Proof. Clearly, the family $(\mathcal{F}_{f(t)})$ is increasing and complete. For right-continuity, take $t \in [(i-1)/k, i/k)$ for some $1 \leq i \leq k$. For a fixed n, (1) and (7) imply $g_n(f(s)) = \bigcup_{j=1}^{i-1} g_n(f_j(j/k)) \cup g_n(f_i(s))\ \ \forall s \in (t, i/k)$. However, since $f_i(s) \downarrow f_i(t)$ as $s \downarrow t$ and g_n is decreasing, preserves countable intersections in $\mathcal{A}$ and has a finite range, there exists $s_0 \in (t, i/k)$ such that $g_n(f_i(s)) = g_n(f_i(t))\ \ \forall s \in (t, s_0)$ so that $g_n(f(s)) = g_n(f(t))\ \ \forall s \in (t, s_0)$. Therefore, since $\{\mathcal{F}_{f(t)} : t \in [0,1]\}$ is increasing,

$$\bigcap_{t<s} \mathcal{F}_{f(s)} = \bigcap_{t<s<s_0} \mathcal{F}_{f(s)} = \bigcap_{n} \bigcap_{t<s<s_0} \mathcal{F}^{\circ}_{g_n(f(s))} = \bigcap_{n} \mathcal{F}^{\circ}_{g_n(f(t))} = \mathcal{F}_{f(t)}.$$

Part (a) follows from the inclusion $\mathcal{F}^{\circ}_{f(t)} \subseteq \mathcal{F}_{f(t)}$, $t \in [0,1]$, and the finite additivity of X on $\mathcal{C}(u)$. A simple proof for (b) can be found in [17]. □

Lemma 3.5 *Given a process $X = \{X_A : A \in \mathcal{A}\}$ and any $f \in S(\mathcal{A})$,*

(i) *X monotone outer- (inner-)continuous implies $X \circ f$ is right- (respectively, left-)continuous and*

(ii) *X increasing implies $X \circ f$ is increasing.*

Proof. A straighforward application of Definitions 2.4 and 3.1. □

We call a process $Y = \{Y_t : t \in [0,1]\}$ a *stretched-out Brownian motion* if there exists a standard Brownian motion $B = \{B_t : t \in \mathbf{R}_+\}$ and a continuous increasing function $\lambda : [0,1] \to [0,\infty)$ with $\lambda(0) = 0$ such that $Y_t = B_{\lambda(t)}\ \forall t \in [0,1]$. The proof of the following result is straightforward.

Lemma 3.6 *If W is an $\mathcal{A}$-indexed Gaussian white noise based on Λ, then for any $f \in S(\mathcal{A})$, $W \circ f$ has a modification which is a standard Brownian motion stretched out by the continuous increasing function $\Lambda \circ f : [0,1] \to [0,\infty)$.*

4 Flow-Wise Convergence.

A framework for functional convergence of set-indexed processes was introduced in [5] for the case of $T = [0, 1]^d$ $(d \in \mathbf{N})$ and in [15] for the case of a general compact metric space T. Central to this framework was the definition of a function space $D(\mathcal{A})$ which served as an analogue to the classical Skorokhod space $D[0, 1]$. A typical element $x : \mathcal{A} \to \mathbf{R}$ of $D(\mathcal{A})$ possessed a set-indexed cadlaguity property while $D(\mathcal{A})$ itself was equipped with a Skorokhod J_2-like metric. Following classical lines, functional convergence of $\mathcal{A}$-indexed processes with sample paths in $D(\mathcal{A})$ was then defined to be weak convergence of the induced measures on the metric space $D(\mathcal{A})$.

However, this approach required a metric on T and when $\mathcal{A}$ is large with respect to the notion of metric entropy, an $\mathcal{A}$-indexed Gaussian process may not possess a version in $D(\mathcal{A})$. Indeed, if W is any $\mathcal{LL}_d$-indexed Gaussian white noise, some $d \geq 2$, then with probability 1, no path of W lies in $D(\mathcal{LL}_d)$—a consequence of Proposition 1.3 on p.9 of [1]. This rules out the possibility of functional CLTs in $D(\mathcal{LL}_d)$ which is unfortunate since the lower layers are a rather natural indexing class. In particular, $\mathcal{LL}_d$ is the smallest collection of subsets of $\mathbf{R}^d_+$ which contains all finite unions of lower rectangles and is closed under countable intersections. For this reason, we introduce an entirely new mode of convergence for set-indexed processes termed *flow-wise convergence* which does not place explicit bounds on the size of $\mathcal{A}$. As a consequence, one can obtain non-trivial CLTs for processes indexed by the lower layers of arbitrarily high dimension (Proposition 7.2). Furthermore, the framework for flow-wise convergence does not require T to be metrizable. The class of admissible processes is defined below.

Definition 4.1 *$D[S(\mathcal{A})]$ is the collection of all $\mathcal{A}$-indexed processes X for which $X \circ f$ has a modification in $D[0, 1]$ for any $f \in S(\mathcal{A})$.*

Elements of $D[S(\mathcal{A})]$ include $\mathcal{A}$-indexed strong martingales (Lemma 3.4), any X which is either increasing or is monotone inner- and outer-continuous (Lemma 3.5) and any $\mathcal{A}$-indexed Gaussian white noise (Lemma 3.6). Given $X \in D[S(\mathcal{A})]$ and $f \in S(\mathcal{A})$ we will write $M_f(X)$ for the unique (up to indistinguishability) cadlag modification of $X \circ f$.

Throughout this paper, unless otherwise mentioned, weak convergence of processes in $D[0, 1]$ is defined with respect to the Skorokhod J_1 topology.

Definition 4.2 *Given processes $X, X_1, X_2, \cdots$ in $D[S(\mathcal{A})]$, (X_n) is said to converge flow-wise to X if $M_f(X_n) \to_{\mathcal{D}} M_f(X)$ in $D[0,1]$ for every fixed $f \in S(\mathcal{A})$.*

Comments. (i) Flow-wise convergence is distributional and hence it is not necessary for $X, X_1, X_2, \cdots$ to be defined on a common probability space.

(ii) While there is generally no relation between flow-wise and functional convergence, if $T = \mathbf{R}_+$ and $\mathcal{A}^1 = \{[0,x] : x \in \mathbf{R}_+\}$, flow-wise convergence of processes in $D(\mathcal{A}^1)$ is stronger than functional convergence. In particular, if we identify $D(\mathcal{A}^1)$ with $D[0,1]$ via $x \mapsto [0,x]$, then flow-wise convergence is equivalent to convergence with respect to the Skorokhod J_1 topology whereas functional convergence in $D(\mathcal{A}^1)$ is equivalent to convergence with respect to the weaker Skorokhod J_2 topology (cf. [24, 15]).

When the limiting process is a Gaussian white noise, flow-wise convergence implies convergence in finite dimensional distribution. This is a special case of the following result.

Proposition 4.3 *If $\mathcal{A}$-indexed processes (X_n) converge flow-wise to an $\mathcal{A}$-indexed process X which admits a continuous modification $M_f(X)$ for each $f \in S(\mathcal{A})$, then for any finite subcollection $\{C_0, \cdots, C_k\}$ of $\mathcal{C}(u)$,*

$$(X_n(C_0), \cdots, X_n(C_k)) \to_{\mathcal{D}} (X(C_0), \cdots, X(C_k))$$

as random vectors.

Proof. We only consider the case in which $C_0, \cdots, C_k$ are the left-neighborhoods generated by a finite sub-semilattice $\mathcal{A}' = \{A_0, \cdots, A_k\}$ of $\mathcal{A}$ which is numbered in a manner consistent with the strong past. The extension to the general case is straightforward (cf. [25]).

By Lemma 3.3, there exists $f \in S(\mathcal{A})$ such that

$$C_i \;=\; f(i/k) \setminus f((i-1)/k) \quad \forall\, 1 \le i \le k \tag{8}$$

and $C_0 = \phi' = f(0)$. Thus, since $M_f(X_n) \to_{\mathcal{D}} M_f(X)$ in $D[0,1]$, the definition of modification yields

$$(X_{n,f(0)}, X_{n,f(1/k)}, \cdots, X_{n,f(1)}) \;\xrightarrow{\mathcal{D}}\; (X_{f(0)}, X_{f(1/k)}, \cdots, X_{f(1)})$$

where $X_{n,B} := X_n(B)$, $B \in \mathcal{A}(u)$. The result now follows by (8) and the continuous mapping theorem. $\square$

We note that the preceding proposition may be applied to any $\mathcal{A}$-indexed Gaussian white noise W since by Lemma 3.6, $M_f(W)$ is continuous.

5 Quadratic Variation

Let $Y = \{Y_t : t \in [0,1]\}$ be a cadlag martingale in L_2. By the Doob-Meyer decomposition theorem, there exists a unique (up to indistinguishability) increasing and predictable process $\langle Y \rangle$ termed the *predictable quadratic variation* of Y for which $Y^2 - \langle Y \rangle$ is a martingale and $\langle Y \rangle(0) = Y^2(0)$. The role of predictable quadratic variation in martingale CLTs is illustrated by the following result which appears on p.179 of [21].

Proposition 5.1 *Take a sequence* (Y_n) *of* L_2 *martingales in* $D[0,1]$ *with* $Y_n(0) = 0 \ \forall n$. *If*

(i) $E[J(Y_n)^2] \to 0$ *as* $n \to \infty$ *and*

(ii) $\langle Y_n \rangle(t) \to \lambda(t)$ *in probability as* $n \to \infty$, *any* $t \in [0,1]$,

where $\lambda : [0,1] \to [0,\infty)$ *is continuous and increasing with* $\lambda(0) = 0$, *then* $Y_n \to_{\mathcal{D}} B_\lambda$ *in* $D[0,1]$, B_λ *a standard Brownian motion stretched out by* λ.

By slightly altering the conditions in Proposition 5.1, one can replace $\langle Y_n \rangle$ in (ii) with the *optional quadratic variation* $[Y_n]$ (cf. [12]). In other words, $Y_n \to_{\mathcal{D}} B_\lambda$ follows if the Y_n asymptotically resemble B_λ in two respects: asymptotically continuous samples paths (condition (i)) and quadratic variation (condition (ii)).

In this section, we introduce a notion of quadratic variation for set-indexed processes. The role it will play in the upcoming strong martingale CLT will be similar to that played by predictable quadratic variation in Proposition 5.1.

Definition 5.2 *Given a strong martingale* $X = \{X_A : A \in \mathcal{A}\}$ *in* L_2, *an increasing process* $X^* = \{X^*_A : A \in \mathcal{A}\}$ *is a *-quadratic variation of* X *if* $E[(X_C)^2 | \mathcal{G}^*_C] = E[X^*_C | \mathcal{G}^*_C]$ *for every* $C \in \mathcal{C}(u)$.

Comments. (i) Unlike its continuous parameter counterpart, a *-quadratic variation X^* of a strong martingale X is not required to be adapted. Furthermore, even if X^* is adapted to the filtration to which X is adapted, $X^2 - X^*$ need not be a strong martingale since $(X_C)^2$ and X_C^2 may fail to coincide. Here, $(X_C)^2$ is the square of the increment of X at C while X_C^2 denotes the increment of X^2 at C.

(ii) Conditions ensuring existence, adaptedness and a form of predictablility for *-quadratic variation were given in [25] and will appear in a future publication. However, in this paper, no property of *-quadratic variation outside of Definition 5.2 is required.

By Lemma 3.4, composing a strong martingale with a simple flow yields a martingale on $[0,1]$. The effect of simple flows on *-quadratic variation is given by the following result.

Lemma 5.3 *Let X be an $\mathcal{A}$-indexed strong martingale in L_2 with *-quadratic variation X^*. Given $f \in S(\mathcal{A})$, if $V := X^* \circ f$ and Y is the cadlag modification of the martingale $X \circ f$, then for any $s < t$ in $[0,1]$,*

$$E(Y_t^2 - Y_s^2 \,|\, \mathcal{F}_{f(s)}) \;=\; E(V_t - V_s \,|\, \mathcal{F}_{f(s)})\,.$$

*Hence, V is a *-quadratic variation for Y and for any $s < t$ in $[0,1]$, $E(\langle Y\rangle_t - \langle Y\rangle_s \,|\, \mathcal{F}_{f(s)}) = E(V_t - V_s \,|\, \mathcal{F}_{f(s)})$.*

Proof. Since Y is a martingale with respect to $(\mathcal{F}_{f(t)})$,

$$\begin{aligned} E[\,(Y_t^2 - Y_s^2) &- (V_t - V_s) \,|\, \mathcal{F}_{f(s)}] \\ &= E[\,(Y_t - Y_s)^2 - (V_t - V_s) \,|\, \mathcal{F}_{f(s)}] + 2Y_s \cdot E(Y_t - Y_s \,|\, \mathcal{F}_{f(s)}) \\ &= E[\,(Y_t - Y_s)^2 - (V_t - V_s) \,|\, \mathcal{F}_{f(s)}] \\ &= E[\,(X_{f(t)\setminus f(s)})^2 - X^*_{f(t)\setminus f(s)} \,|\, \mathcal{F}_{f(s)}]\,. \end{aligned}$$

Therefore, since (5) implies $\mathcal{F}_{f(s)} \subseteq \mathcal{G}^*_{f(t)\setminus f(s)}$, the proof follows by the tower property and the definition of *-quadratic variation. □

In the remainder of this section, we study a special class of strong martingales for which *-quadratic variation can be explicitly calculated. Fix $d \in \mathbf{N}$ and take a continuous distribution function $F : \mathbf{R}^d \to [0,1]$. Using marginal transforms if necessary, we may assume that F is supported on $[0,1]^d$ (cf. [9]). Now, take two sequences: (Y_n) of i.i.d. random vectors distributed F and (Z_n) of i.i.d. random variables which is independent of (Y_n) with $E(Z_1) = 0$ and $E(Z_1^2) = 1$. For technical reasons, we also insist that $P(Z_1 = 0) = 0$. In addition to the conditions in Section 2, $\mathcal{A}$ is a collection of closed subsets of $T = [0,1]^d$ satisfying the following:

Assumption 5.4 Every $C = A \setminus B \in \mathcal{C}$ ($A \in \mathcal{A}$, $B \in \mathcal{A}(u)$) possesses a *maximal representation*, that is, there exists $n \in \mathbf{N}$ and $A_1, A_2, \cdots, A_n \in \mathcal{A}$ with $C = A \setminus \bigcup_{i=1}^n A_i$ such that $A_i \not\subseteq A_j \ \forall\, i \neq j$ and, given $B' \in \mathcal{A}(u)$, $B' \cap C = \phi$ implies $B' \subseteq \bigcup_{i=1}^n A_i$. Further,

(a) $\mathcal{A}$ contains all lower rectangles R_z, $z \in [0,1]^d$, and

(b) $g_n(A) = \bigcap\{B \in \mathcal{A}_n(u) : A \subseteq B^\circ\}$ for any $A \in \mathcal{A}$ where $\mathcal{A}_n$ consists of all lower rectangles R_z with $z_i = \frac{m_i}{2^n}$ for some $m_i \in \{0, 1, \cdots, 2^n\}$, each $i = 1, \cdots, d$.

This assumption appeared in [18] in connection with Poisson convergence of set-indexed empirical processes. Examples of $\mathcal{A}$ for which it is satisfied include the class of all lower rectangles and the class of all lower layers in $[0,1]^d$. By (a), any such $\mathcal{A}$ generates the Borel σ-algebra on $[0,1]^d$.

For each $k \in \mathbf{N}$, define an $\mathcal{A}$-indexed process M_k by setting

$$M_k(A) \;=\; \mathbf{1}_{[Y_k \in A]}\, Z_k\,, \;\; A \in \mathcal{A}$$

and let $U_n := n^{-1/2} \sum_{k=1}^n M_k$, the n-th $\mathcal{A}$-indexed weighted empirical process. For each $B \in \mathcal{A}(u)$, define $\mathcal{H}_B$ to be the completion of $\sigma(\{M_k(A) : k \in \mathbf{N},\ A \in \mathcal{A} \text{ and } A \subseteq B\})$ and let $\mathcal{F}_A := \bigcap_n \mathcal{H}_{g_n(A)} \ \forall\, A \in \mathcal{A}$. Clearly, $(\mathcal{F}_A)$ is an $\mathcal{A}$-indexed filtration to which both M_n and U_n are adapted. The finitely additive extensions of M_k and U_n to $\mathcal{C}(u)$ are the obvious ones.

Given $z, z' \in (0,1]^d$, define the intervals $[z, z'] = \prod_{i=1}^d [z_i, z_i']$ and $(z, z'] = \prod_{i=1}^d (z_i, z_i']$. In addition, given $z \in (0,1)^d$, define the sets $S_z = (z, 1]$, $S_{z-} = [z, 1]$ and $L_z = [0,1]^d \setminus S_z$. Note that $(z, z'] \in \mathcal{C}$ and $L_z \in \mathcal{A}(u)$. It is straightforward to show $\mathcal{G}^*_{(z,z']} = \mathcal{F}_{L_z} = \bigcap_n \mathcal{H}_{g_n(L_z)}$ for any $z \in (0,1)^d$.

Fix k and take a random variable W independent of Y_k. If $C = (z, z']$, then since $C \subseteq S_z$, the smoothing property of conditional expectation (cf. [4], Theorem 6.5.9 (a)) and a straightforward limiting argument yield

$$E[\mathbf{1}_{[Y_k \in C]}\, W | \mathcal{G}^*_C] \;=\; \mathbf{1}_{[Y_k \in S_z]}\, E(W)\, F(C)\, h_F(z) \tag{9}$$

where $h_F(z) = [F(S_{z-})]^{-1}$ whenever $F(S_{z-}) > 0$ and $h_F(z) = 0$ otherwise. In particular, $E[M_k(z,z']|\mathcal{G}^*_{(z,z']}] = 0$. Following [18], this relation can be extended to all sets $C \in \mathcal{C}(u)$; since $\{\mathcal{G}^*_C : C \in \mathcal{C}(u)\}$ is decreasing and M is finitely additive, conditioning implies $E[M_k(D)|\mathcal{G}^*_D] = 0$ for any finite disjoint union $D = \bigcup_j C_j$ where each C_j is of the form $(z,z']$. Furthermore, under Assumption 5.4, any $C \in \mathcal{C}(u)$ can be approximated by a sequence of such unions so that dominated convergence implies $E[M_k(C)|\mathcal{G}^*_C] = 0$ (cf. p.84 of [18]). Therefore, we have shown

Proposition 5.5 *Under Assumption 5.4, each M_k (and hence each U_n) is a strong martingale with respect to $(\mathcal{F}_A)$.*

In addition,

Proposition 5.6 *Under Assumption 5.4, M_k has *-quadratic variation*

$$M^*_k(A) \;=\; \int_A \mathbf{1}_{[0,Y_k]}\, h_F\, dF\,, \quad A \in \mathcal{A}.$$

*Furthermore, $U^*_n = n^{-1}\sum_{k=1}^n M^*_k$ is a *-quadratic variation of U_n.*

Proof. Clearly, M^*_k is an increasing process whose finitely additive extension to $\mathcal{C}(u)$ is the obvious one. As shown on p.83 of [18], when $C = (z,z']$,

$$E[M^*_k(C)|\mathcal{G}^*_C] \;=\; \mathbf{1}_{[Y_k \in S_z]}\, F(C)\, h_F(z). \tag{10}$$

Therefore, since $E(Z_k^2) = 1$, (9) implies $E[(M_k(C))^2 - M^*_k(C)|\mathcal{G}^*_C] = 0$. As described above, this relation can be extended to finite disjoint unions of sets of the form $(z,z']$ and then to general $C \in \mathcal{C}(u)$.

Given $C = (z,z']$, (9) can be generalized to yield

$$E[M_i(C)\, M_j(C)|\mathcal{G}^*_C] \;=\; \mathbf{1}_{[Y_i \in C,\, Y_j \in C]}\, E(Z_i)\, E(Z_j)\, [F(C)\, h_F(z)]^2 \;=\; 0.$$

Thus, since M^*_k is a *-quadratic variation of M_k, $E[(U_n(C))^2 - U^*_n(C)|\mathcal{G}^*_C] = 0$. As before, this relation extends to all $C \in \mathcal{C}(u)$. □

6 A CLT For Set-Indexed Martingales.

This section contains the main result of the paper, a flow-wise CLT for sequences of set-indexed strong martingales. Recall that any $\mathcal{A}$-indexed strong martingale and any $\mathcal{A}$-indexed Gaussian white noise lie in $D[S(\mathcal{A})]$ so that the flow-wise mode of convergence is permissible in this setting. For simplicity, we state and prove this CLT for the case of T compact with $T \in \mathcal{A}$. Using standard techniques, this can be extended the case of a general σ-compact T if conditions (i) and (11) hold with B_k in place of T, each k, where (B_k) is the sequence in $\mathcal{A}$ defined in Section 2.

Theorem 6.1 *Let $X_1, X_2, \cdots$ be $\mathcal{A}$-indexed strong martingales in L_2 with $X_n(\phi') = 0 \ \forall n$ such that*

(i) $\sup_n E[|X_n(T)|^{2+\delta}] < \infty$ *for some* $\delta > 0$ *and*

(ii) $J(M_f(X_n)) \xrightarrow{P} 0$ *as* $n \to \infty$ *for any fixed* $f \in S(\mathcal{A})$.

If for each $n \in \mathbf{N}$ there exists a quadratic variation X_n^* of X_n such that

$$\{X_n^*(T) : n \geq 1\} \text{ is uniformly integrable} \tag{11}$$

and there exists a variance function Λ on $\mathcal{A}$ such that

$$X_n^*(A) \to \Lambda(A) \text{ in probability } \forall A \in \mathcal{A}, \tag{12}$$

then (X_n) converges flow-wise to an $\mathcal{A}$-indexed Gaussian white noise W based on Λ.

Comments. (a) Theorem 6.1 generalizes the principle found in Proposition 5.1. Namely, under a certain moment condition, if the jumps in the paths of the X_n become asymptotically negligible in the sense of condition (ii) and if there exists *-quadratic variations X_n^* of X_n which converge to a continuous, increasing, deterministic limit Λ, then (X_n) converges flow-wise to a Gaussian white noise based on Λ.

(b) It has been shown in [25] that tightness plus convergence of finite-dimensional distributions implies functional convergence in $D(\mathcal{A})$, the function space mentioned in the introduction of Section 4. Therefore, Theorem 6.1 and Proposition 4.3 yield functional CLTs for tight sequences of strong martingales in $D(\mathcal{A})$. Generic tightness criteria for processes in $D(\mathcal{A})$ can be found in [25]. However, as mentioned in Section 4, to ensure the existence of a limiting Gaussian process in $D(\mathcal{A})$, $\mathcal{A}$ must be sufficiently small, for example a Vapnik-Červonenkis class on T.

(c) In [25], conditions on $\mathcal{A}$ and $(\mathcal{F}_A)$ were given under which (ii) implies the existence of *-quadratic variations $X_1^*, X_2^*, \cdots$ satisfying (11).

Proof. Let W be an $\mathcal{A}$-indexed Gaussian white noise based on Λ. For the sake of notation, we consider only the case in which each X_n is defined with respect to a common filtration $(\mathcal{F}_A)$ on a common $(\Omega, \mathcal{F}, P)$ and W is also defined on $(\Omega, \mathcal{F}, P)$—the proof for the non-homogeneous case is identical.

Fix a simple flow $f : [0,1] \to \mathcal{A}(u)$. Since $T \in \mathcal{A}$, we may assume without loss of generality that $f(1) = T$. For each $n \in \mathbf{N}$, define processes Y_n and V_n on $[0,1]$ by letting $Y_n = M_f(X_n)$ and $V_n = M_f(X_n^*) = X_n^* \circ f$. Note that assumptions (i), (ii) and Doob's inequality imply that $E[J(Y_n)^2] \to 0$ since $J(Y_n) \leq 2 \sup_{0 \leq s \leq 1} |Y_n(s)|$. Applying Lemmas 3.4 and 3.5 in that order,

(f1) Y_n is a cadlag L_2 martingale with respect to the filtration $(\mathcal{H}_t)$ where $\mathcal{H}_t := \mathcal{F}_{f(t)}$, $t \in [0,1]$, and

(f2) V_n is an increasing (but not necessarily adapted) process.

If we set $\lambda(t) := \Lambda(f(t))\ \forall t \in [0,1]$, then Lemma 3.5 implies λ is increasing and continuous on $[0,1]$ with $\lambda(0) = 0$. Therefore, if we can show

$$\langle Y_n \rangle(t) \xrightarrow{P} \lambda(t) \quad \forall t \in [0,1], \tag{13}$$

then $M_f(X_n) \to_{\mathcal{D}} M_f(W)$ in $D[0,1]$ will follow by Proposition 5.1 and Lemma 3.6.

Since Λ and each X_n^* have finitely additive extensions to $\mathcal{C}(u)$, assumption (12) can be extended to all sets in $\mathcal{A}(u)$ so that

$$V_n(t) \xrightarrow{P} \lambda(t) \quad \forall t \in [0,1]. \tag{14}$$

To obtain (13) from (14), it will first be shown that $\{\langle Y_n \rangle : n \geq 1\}$ is tight in $D[0,1]$. Then, it will be argued that any weak limit point Z of $\{\langle Y_n \rangle : n \geq 1\}$ is continuous and increasing with $Z - \lambda$ a martingale so that $Z - \lambda$ is a continuous martingale of finite variation. By Proposition IV-1.2 in [23], any such $Z - \lambda$ is necessarily indistinguishable from the zero process, i.e., $Z = \lambda$ so that tightness implies weak convergence in $D[0,1]$ of $\{\langle Y_n\rangle : n \geq 1\}$ to the continuous determistic limit λ. These details are the content of the next two lemmas. □

Lemma 6.2 *$\{\langle Y_n\rangle : n \geq 1\}$ is tight in $D[0,1]$ with respect to the Skorokhod J_1 topology.*

Lemma 6.3 *If Z is the limiting process of a weakly convergent subsequence of $\{\langle Y_n\rangle : n \geq 1\}$, then*

(a) *Z is continuous and increasing and*

(b) *$Z - \lambda$ is a martingale with respect to its minimal filtration.*

Proof of Lemma 6.2. We employ the 1-dimensional stopping time condition of Aldous (cf. [2], Theorem 1). To begin with, note that tightness of $\{\langle Y_n\rangle(t) : n \geq 1\}$ for any fixed $t \in [0,1]$ follows easily from Markov's inequality and assumption (i) since each $\langle Y_n\rangle$ is increasing.

Take a sequence (δ_n) of constants where $0 \leq \delta_n \leq 1 \;\forall n$ and $\delta_n \to 0$ and for each n, take a stopping time $\tau_n : \Omega \to [0,1]$ with respect to the filtration generated by $\langle Y_n\rangle$. We will verify condition (A) of Aldous,

$$\langle Y_n\rangle(\sigma_n) - \langle Y_n\rangle(\tau_n) \xrightarrow{P} 0 \quad \text{as } n \to \infty \tag{A}$$

where $\sigma_n := 1 \wedge (\tau_n + \delta_n) \;\forall n$ by establishing L_1 convergence.

Since $\langle Y_n\rangle$ is adapted to $(\mathcal{H}_t)$, τ_n and σ_n are both stopping times with respect to $(\mathcal{H}_t)$. However, since V_n is not necessarily adapted to $(\mathcal{H}_t)$, $\langle Y_n\rangle - V_n$ is not necessarily a martingale with respect to $(\mathcal{H}_t)$. Just the same, applying (f 2) and Lemma 5.3, the proof of Doob's optional stopping theorem can be easily extended to our situation so as to yield

$$E[\langle Y_n\rangle(\sigma_n) - \langle Y_n\rangle(\tau_n) \,|\, \mathcal{H}_{\tau_n}] \;=\; E[V_n(\sigma_n) - V_n(\tau_n) \,|\, \mathcal{H}_{\tau_n}] \quad \forall n.$$

Hence, it is sufficient to show $\lim_n E[V_n(\sigma_n) - V_n(\tau_n)] = 0$.

Given any n,

$$\begin{aligned} |V_n(\sigma_n) - V_n(\tau_n)| &\leq |V_n(\sigma_n) - \lambda(\sigma_n)| \;+\; |V_n(\tau_n) - \lambda(\tau_n)| \;+ \\ &\quad\; |\lambda(\sigma_n) - \lambda(\tau_n)| \\ &\leq 2 \cdot \sup_{0 \leq t \leq 1} |V_n(t) - \lambda(t)| \;+\; |\lambda(\sigma_n) - \lambda(\tau_n)|. \end{aligned}$$

Applying Lemma 1 in [20], (14) implies $\sup_{0\le t\le 1}|V_n(t)-\lambda(t)| \xrightarrow{P} 0$ so that $|V_n(\sigma_n)-V_n(\tau_n)| \xrightarrow{P} 0$ follows from the continuity of λ. Therefore, since V_n is increasing, $|V_n(\sigma_n)-V_n(\tau_n)| \le 2\,V_n(1) = 2\,X_n^*(T)$ so that (11) implies $\lim_n E[V_n(\sigma_n)-V_n(\tau_n)] = 0$ by dominated convergence, completing the proof of Lemma 6.2. □

Proof of Lemma 6.3 (a). Suppose $\langle Y_{n'}\rangle \to_{\mathcal{D}} Z$ in $D[0,1]$ along a subsequence (n'). For the sake of notation, take $n = n'$ and assume Z is also defined on $(\Omega, \mathcal{F}, P)$. Being the weak limit of increasing processes in $D[0,1]$, Z is necessarily increasing.

Next, we establish the sample path continuity of Z. Since $\{\langle Y_n\rangle : n \ge 1\}$ is tight in $D[0,1]$, it is sufficient by Proposition VI-3.26 in [19] to show

$$J(\langle Y_n\rangle) \xrightarrow{P} 0 \ \text{ as } n \to \infty. \tag{15}$$

Given $n \in \mathbf{N}$ and $\epsilon > 0$, define $S_n = \inf\{0 < t \le 1 : \Delta\langle Y_n\rangle(t) > \epsilon\}$ with the convention $\inf\phi = \infty$. (15) will follow if we can show $P(S_n \le 1) \to 0$ as $n \to \infty$ for every $\epsilon > 0$. Our argument will be close to that found on p.268 of [13].

Fix $\epsilon > 0$. Given any n, since $\langle Y_n\rangle$ is non-negative and increasing, $(\Delta\langle Y_n\rangle)_{S_n} > \epsilon$ on $[S_n \le 1]$ so that

$$P(S_n \le 1) \ \le\ \epsilon^{-1}\cdot E\left[\mathbf{1}_{[S_n\le 1]}\cdot(\Delta\langle Y_n\rangle)_{S_n}\right]. \tag{16}$$

Furthermore, since S_n is a predictable stopping time with respect to $(\mathcal{H}_t)$ (cf. [19], Proposition I-2.13) and since (f2) and Lemma 5.3 allow us to extend Doob's optional stopping theorem, we have

$$E[\,(\Delta\langle Y_n\rangle)_{S_n}\,|\,\mathcal{H}_{S_n-}\,] \ = \ E[\,(\Delta V_n)_{S_n}\,|\,\mathcal{H}_{S_n-}\,].$$

Therefore, conditioning and (16) yield

$$P(S_n \le 1) \ \le\ \epsilon^{-1}\cdot E\left[\mathbf{1}_{[S_n\le 1]}\cdot(\Delta V_n)_{S_n}\right] \ \le\ \epsilon^{-1}\cdot E\left[J(V_n)\right].$$

As mentioned in the proof of Lemma 6.2, (14) implies $\sup_t|V_n(t)-\lambda(t)| \xrightarrow{P} 0$ so that $J(V_n) \xrightarrow{P} 0$ by the continuous mapping theorem. Therefore, since $J(V_n) \le 2\,X_n^*(T)$, (11) implies $E[J(V_n)] \to 0$ as $n \to \infty$ so that

$P(S_n \le 1) \to 0$, hence completing the proof of Lemma 6.3 (a). $\square$

Proof of Lemma 6.3 (b). Once again, assume $\langle Y_n \rangle$ converges weakly in $D[0,1]$to a process Z on $(\Omega, \mathcal{F}, P)$. Let $M = Z - \lambda$. For each $0 \le t \le 1$, define $\mathcal{G}_t^\circ = \sigma(\{Z_u : 0 \le u \le t\})$ and let $(\mathcal{G}_t)$ denote the filtration generated by the family $(\mathcal{G}_t^\circ)$. Since $\lambda(t)$ is deterministic $\forall\, t \in [0,1]$, $(\mathcal{G}_t)$ is in fact the minimal filtration generated by M.

The following argument is analogous to that found on pp.260-261 of [13]. Let Γ_t denote the set of P-continuity points of Z_t. Given any $s \in [0,1]$, it is clear that

$$\mathcal{E}_s \;=\; \{\cap_{i=1}^m [Z_{s_i} \le x_i] : m \in \mathbf{N},\, 0 \le s_i \le s\,\forall\, i \text{ and } x_i \in \Gamma_{s_i}\,\forall\, i\}$$

is a π-class generating $\mathcal{G}_s^\circ$. Therefore, given $s < t$ in $[0,1]$, if we can show

$$\int_A (M_t - M_s)\, dP = 0 \quad \forall\, A \in \mathcal{E}_s, \tag{17}$$

then $E[M_t |\, \mathcal{G}_s^\circ] = M_s$ will follow. By Lemma 6.3 (a), replacing $\mathcal{G}_s^\circ$ by $\mathcal{G}_s$ in this relation is a simple application of dominated convergence since

$$E[M_t|\mathcal{G}_s] \;=\; E[E(M_t|\mathcal{G}_{s+1/n}^\circ)|\mathcal{G}_s] \;=\; E[M_{s+1/n}|\mathcal{G}_s]$$

for all sufficiently large n. With this reduction in mind, fix $s < t$ in $[0,1]$ and select a set $A = \bigcap_{i=1}^m [Z(s_i) \le x_i]$ in $\mathcal{E}_s$. For each n, define the set $A_n = \bigcap_{i=1}^m [\,\langle Y_n \rangle (s_i) \le x_i]$ and the process $M_n = \langle Y_n \rangle - V_n$. To save space, given any process or function U, we will write $U(s,t]$ for $U(t) - U(s)$.

Since $\langle Y_n \rangle \to Z$ in finite dimensional distribution, it is clear that $\mathbf{1}_{A_n} \to_{\mathcal{D}} \mathbf{1}_A$ so that (14) implies

$$V_n(s,t] \cdot \mathbf{1}_{A_n} \xrightarrow{\mathcal{D}} \lambda(s,t] \cdot \mathbf{1}_A. \tag{18}$$

Furthermore, since $P(\langle Y_n \rangle (s,t] \cdot \mathbf{1}_{A_n} > x) = P\left([\langle Y_n \rangle (s,t] > x] \cap A_n\right)$ for every $n \ge 1$ and $x \ge 0$, it is straightforward to show

$$\langle Y_n \rangle (s,t] \cdot \mathbf{1}_{A_n} \xrightarrow{\mathcal{D}} Z(s,t] \cdot \mathbf{1}_A\,. \tag{19}$$

Employing the discrete approximation of predictable quadratic variation, assumption (i) and Rosenthal's inequality imply $\{\langle Y_n \rangle (1) : n \ge 1\}$

is uniformly integrable. Thus, since $|\langle Y_n\rangle(s,t] \cdot \mathbf{1}_{A_n}| \le 2\,\langle Y_n\rangle(1)\ \ \forall\, n$, we have $\{\langle Y_n\rangle(s,t] \cdot \mathbf{1}_{A_n} : n \ge 1\}$ uniformly integrable. Likewise, (11) implies $\{V_n(s,t] \cdot \mathbf{1}_{A_n} : n \ge 1\}$ is uniformly integrable. Hence, (18), (19) and the dominated convergence theorem (cf. [8], Theorem 25.12)imply $\int_A M(s,t]\,dP = \lim_n \int_{A_n} M_n(s,t]\,dP$. Furthermore, since $\langle Y_n\rangle$ is adapted to $(\mathcal{F}_{f(t)})$, $A_n \in \mathcal{F}_{f(s)}\ \forall\, n$. Therefore, Lemma 5.3 implies $\int_{A_n} M_n(s,t]\,dP = 0\ \forall\, n$ which establishes (17) and completes the proof of Lemma 6.3 (b). □

We observe that in the course of proving the central limit theorem, it was shown that if *any* sequence of quadratic variation processes corresponding to a sequence of continuous parameter martingales is asymptotically deterministic, then the usual predictable quadratic variations will exhibit the same behaviour. This is summarized in the following lemma which may be of independent interest:

Lemma 6.4 *Let (Y_n) be a sequence of martingales on $[0,1]$ and let (V_n) be any sequence of corresponding *-quadratic variations. Assume that:*

(a) $\sup_n E[|Y_n(1)|^{2+\delta}] < \infty$ *for some* $\delta > 0$, *and*

(b) $(V_n(1))_n$ *is uniformly integrable.*

If there exists a continuous function λ such that

$$V_n(s) \xrightarrow{P} \lambda(s) \quad \forall s \in [0,1],$$

then

$$\langle Y\rangle_n(s) \xrightarrow{P} \lambda(s) \quad \forall s \in [0,1].$$

7 Applications

In this section, we present two consequences of Theorem 6.1. The first is a functional CLT for sequences of multidimensional strong martingales. The second is a flow-wise CLT for sequences of weighted empirical processes indexed by the lower layers in $[0,1]^d$, $d \ge 2$, a class too large to support a continuous (or even bounded) Gaussian white noise.

Before presenting the first application, we recall several concepts from the multiparameter theory. Given $z, z' \in [0,1]^d$, we write $z \le z'$ if $z_i \le z'_i$ for each $i = 1, \cdots, d$ and $z \prec z'$ if $z_i < z'_i$ for every $i = 1, \cdots, d$. Fix a complete probability space $(\Omega, \mathcal{F}, P)$. A family $\{\mathcal{F}_z : z \in [0,1]^d\}$ of sub-σ-algebras of $\mathcal{F}$ constitutes a *d-dimensional filtration* if: $\mathcal{F}_0$ contains all P-null subsets of Ω, $z \leq z'$ implies $\mathcal{F}_z \subseteq \mathcal{F}_{z'}$ and $\mathcal{F}_z = \bigcap_{z \prec z'} \mathcal{F}_{z'}$ for any $z \in [0,1)^d$. This differs slightly from the definition given in [10] in that we do not require their (F4) conditional independence assumption to hold. In this setting, the strong past at $z \in (0,1]^d$ is the sub-σ-algebra

$$\mathcal{F}^*_z \;=\; \bigvee_{z \not\preceq z'} \mathcal{F}_{z'} \,.$$

Given a process $Y = \{Y_z : z \in [0,1]^d\}$ and points $z \prec z'$ in $(0,1]^d$, the increment of Y over $(z, z']$ is defined to be $Y(z, z'] = \sum_J (-1)^{|J|} Y(z_J)$, the sum ranging over all $J \subseteq \{1, \cdots, d\}$ where $z_J(i) = z_i$ if $i \in J$ and $z_J(i) = z'_i$ otherwise. An adapted integrable process Y is a *d-dimensional strong martingale* if Y vanishes on the axes and for every $z \prec z' \in (0,1]^d$

$$E[Y(z, z'] | \mathcal{F}^*_z] \;=\; 0 \,.$$

We now introduce a general notion of quadratic variation for d-dimensional strong martingales. Given a strong martingale $Y = \{Y_z : z \in [0,1]^d\}$ in L_2, an integrable process $V = \{V_z : z \in [0,1]^d\}$ which vanishes on the axes constitutes a *general quadratic variation* of Y if

(G1) $z \leq z'$ in $(0,1]^d$ implies $V(z, z'] \geq 0$,

(G2) (z_k) decreasing to z in $[0,1]^d$ implies $V(z_k) \to V(z)$ on Ω and

(G3) $E[(Y(z, z'])^2 | \mathcal{F}^*_z] = E[V(z, z'] | \mathcal{F}^*_z]$ for every $z \prec z'$ in $(0,1]^d$.

As was the case for *-quadratic variation, a general quadratic variation is not required to be adapted or to possess any form of predictability. In the case of $d = 1$, both optional and predictable quadratic variation are examples of general quadratic variation.

Let $D_d = D([0,1]^d)$ denote the Skorokhod space of d-dimensional cadlag functions, $x : [0,1]^d \to \mathbf{R}$ (cf. [7]). Given $x \in D_d$ and $z \in (0,1]^d$, define $\Delta x(z) = x(z) - \lim_{z' \prec z} x(z)$ As a consequence of Theorem 6.1, we have the following functional CLT for multiparameter strong martingales.

Proposition 7.1 *For each $n \in \mathbf{N}$, let $Y_n = \{Y_n(z) : z \in [0,1]^d\}$ be a strong martingale in D_d such that*

(i) $\sup_n E[|Y_n(1)|^{2+\delta}] < \infty$ *for some* $\delta > 0$ *and*

(ii) $\sup_{z\in(0,1]^d} |\Delta Y_n(z)| \xrightarrow{P} 0$ *as* $n \to \infty$.

Assume that for each $n \in \mathbf{N}$ *there exists a general quadratic variation* V_n *of* Y_n *such that* $\{V_n(1) : n \geq 1\}$ *is uniformly integrable. If there exists an increasing function* $H : [0,1]^d \to [0,\infty)$ *which is continuous in the sense*

$$z_n \uparrow (\downarrow)\, z \implies H(z_n) \uparrow (\downarrow)\, H(z)$$

such that

$$V_n(z) \to H(z) \text{ in probability } \forall\, z \in [0,1]^d,$$

then $Y_n \to_{\mathcal{D}} B_H$ *in the* J_1*-topology on* D_d *where* B_H *is a d-dimensional Brownian motion with* $E[W_z W_{z'}] = H(z \wedge z')$ *for every* $z, z' \in [0,1]^d$.

Proof. It is straightforward to show that X_n defined by $X_n(R_z) := Y_n(z)$ is a strong martingale indexed by the collection $\mathcal{A} = \{R_z : z \in [0,1]^d\}$ of lower rectangles in $[0,1]^d$. Furthemore, if we define $X_n^*(R_z) := V_n(z)$, then X_n^* is a *-quadratic variation of X_n. In particular, since each $C \in \mathcal{C}(u)$ is a finite disjoint union of sets of the form $(z, z'] \in \mathcal{C}$ $(z \prec z')$ and $\mathcal{F}_z^* = \mathcal{G}_{(z,z']}^*$, a simple conditioning argument allows us to extend (G3) to all sets in $\mathcal{C}(u)$. Clearly, the conditions of Theorem 6.1 hold since

$$J(M_f(X_n)) \leq \sup{}_{z\in(0,1]^d} |\Delta Y_n(z)|$$

for any $f \in S(\mathcal{A})$. Therefore, by Proposition 4.3, the finite dimensional distributions of (X_n) converges to those of a Gaussian white noise $\{W(R_z) : z \in [0,1]^d\}$ based on the variance function Λ where $\Lambda(R_z) := H(z)$.

Define $B_H = \{B_H(z) : z \in [0,1]^d\}$ to be a path-continuous version of $\{W(R_z) : z \in [0,1]^d\}$ (cf. [1]). Then, (Y_n) is a sequence of strong martingales converging in finite dimensional distribution to the continuous process B_H. By Theorem 4 of [14], this is sufficient for $Y_n \to_{\mathcal{D}} B_H$ in D_d. □

Comment. Even for the case of $d = 1$, Theorem 6.1 is new in that the quadratic variations of the 1-dimensional martingales Y_n need not be adapted.

For the second application of Theorem 6.1, recall the sequence (U_n) of weighted empirical processes defined in Section 5. When $Z_1 \sim N(0,1)$ and

$\mathcal{A} = \{R_z : z \in [0,1]^d\}$, Burke in [9] established functional convergence in D_d of (U_n), each U_n viewed as a point-indexed process $U_n(z) := U_n(R_z)$, to an appropriately scaled d-dimensional Brownian motion. Although the mode of convergence may no longer be equivalent to the functional mode, the following flow-wise CLT applies to more general (Z_n) and $\mathcal{A}$.

Proposition 7.2 *Under Assumption 5.4, if $E(Z_1^4) < \infty$, then (U_n) converges flow-wise to an $\mathcal{A}$-indexed Gaussian white noise W based on F.*

As this proposition suggests, F restricted to $\mathcal{A}$ is a variance function. Monotone outer-continuity of F is immediate whereas monotone inner-continuity is implied by the following result.

Lemma 7.3 *Under Assumption 5.4, $\lim_n F(A_n) = F(\overline{\bigcup_n A_n})$ for any increasing sequence (A_n) in $\mathcal{A}(u)$.*

Proof. Take an increasing sequence (A_n) in $\mathcal{A}(u)$ and let $B = \overline{\bigcup_n A_n}$. If we define $g_n^-(B) = \bigcup\{B' \in \mathcal{A}_n(u) : B' \subseteq B^\circ\}$ for each $n \in \mathbf{N}$, then by a straightforward but tedious argument, $B \setminus \bigcup_n A_n \subseteq \partial B \subseteq g_n(B) \setminus g_n^-(B)$ $\forall n$ and $\lambda(g_n(B) \setminus g_n^-(B)) \to 0$ as $n \to \infty$ where λ denotes Lebesgue measure on $[0,1]^d$. Therefore, since F is continuous, $F(B) = F(\bigcup_n A_n) = \lim_n F(A_n)$. □

Toward the proof of Proposition 7.2, we verify the required asymptotic rarefaction of jumps condition for (U_n).

Lemma 7.4 *Under Assumption 5.4, for any $f \in S(\mathcal{A})$, $E[J(M_f(U_n))^2] \to 0$ as $n \to \infty$.*

Proof. Fix $f \in S(\mathcal{A})$ and define $F_f : [0,1] \to [0,1]$ by $F_f(t) := F(f(t))$. By the continuity of flows and Lemma 7.3, F_f determines a continuous distribution function. (Since $[0,1]^d \in \mathcal{A}$, it can be assumed without loss of generality that $f(1) = [0,1]^d$.) If for each $n \in \mathbf{N}$ we define $W_n = \inf\{t \in [0,1] : Y_n \in f(t)\}$, then (W_n) is i.i.d. with common distribution F_f.

Given $t \in (0,1]$, define $\Delta f(t) = f(t) \setminus \bigcup_{s<t} f(s)$. Clearly, $[W_i = t] = [Y_i \in \Delta f(t)]$ $\forall i$ so that

$$|\Delta U_n(f(t))| \;=\; |U_n(\Delta f(t))| \;\leq\; n^{-1/2} \max_{1 \leq i \leq n} |Z_i| \sum_{i=1}^n \mathbf{1}_{[W_i = t]}.$$

Furthermore, the continuity of F_f implies $[W_i \neq W_j]$ has zero measure for every $i \neq j$ so that

$$E[J(M_f(U_n))^2] \;=\; E\left[\max_{0\leq t\leq 1} \mid \Delta U_n(f(t))|^2\right] \;\leq\; n^{-1}E\left[\max_{1\leq i\leq n} Z_i^2\right].$$

But (Z_i) is i.i.d. with $E(Z_1^2) < \infty$. Therefore, $E[\max_{1\leq i\leq n} Z_i^2] = o(n)$ which completes the proof of Lemma 7.4. □

Proof of Proposition 7.2. We must verify the remaining conditions of Theorem 6.1. First, under Assumption 5.4, $\phi' = \{0\}$ so that the continuity of F implies $U_n(\phi') = 0$. Next, given $n \in \mathbf{N}$, observe that $(\sum_{i=1}^n Z_i)^4$ contains n terms of the form Z_i^4 and $3n(n-1)$ terms of the form $Z_i^2 Z_j^2$ so that the i.i.d. assumption on (Z_i) yields $E[(U_n([0,1]^d)^4] = n^{-1}E(Z_1^4) + 3(1-n^{-1})(E(Z_1^2))^2$. This implies condition (i) is satisfied with $\delta = 2$.

Let M_k^* and U_n^* be as defined in Proposition 5.6. Conditions (11) and (12) will follow if $E[U_n^*(A)] = F(A)$ and $U_n^*(A) \xrightarrow{P} F(A)$ for every $A \in \mathcal{A}$.

The fact that

$$E[U_n^*(C)] = F(C) \tag{20}$$

for every $C \in \mathcal{C}$ is a simple consequence of (10). We shall show that $U_n^*(A) \xrightarrow{P} F(A)$, for every set $A \in \mathcal{A}$ of the form $A = R_z$. If this is the case, then it follows that $U_n^*(A) \xrightarrow{P} F(A)$ for $A \in \mathcal{A}$ of the form $A = \cup_{h=1}^k R_{z_h}$. For an arbitrary set $A \in \mathcal{A}$, the absolute continuity of F and the fact that A can be approximated uniformly by finite unions of rectangles permits us to conclude that $U_n^*(A) \xrightarrow{P} F(A)$.

To avoid technicalities in what follows, we shall assume that the density f of F with respect to Lebesgue measure is strictly positive. We shall prove below that $U_n^*(A) \to F(A)$ in L_2 for every set $A = R_z$ where $z \prec (1, ..., 1)$. This suffices, since if $D = R_v$ where $v_i = 1$ for at least one i, a sequence $(z_m) = (z_{m,1}, ..., z_{m,d})_m$ exists with $z_m \prec 1$ $\forall m$ and such that $R_{z_m} \subset D$ and $F(D \setminus R_{z_m}) < 1/m$. Hence, by (20)

$$P(|U_n^*(D) - U_n^*(R_{z_m})| \geq \epsilon) \leq 1/m\epsilon$$

for every n. Now, if $U_n^*(R_{z_m}) \xrightarrow{P} F(R_{z_m})$ as $n \to \infty$ and since $F(R_{z_m}) \to F(D)$ as $m \to \infty$, it follows that $U_n^*(D) \xrightarrow{P} F(D)$ as $n \to \infty$.

Therefore it suffices to show that for $z \prec 1$ and $A = R_z$,

$$E[(U_n^*(A))^2] \to (F(A))^2. \tag{21}$$

Now,

$$U_n^*(A)^2 = n^{-2} \sum_{i=1}^{n} \left[\int_A I_{[0,Y_i]} h_F(u) dF(u) \right]^2 \tag{22}$$
$$+ n^{-2} \sum_{i \neq j} \int_A I_{[0,Y_i]} h_F(u) dF(u)$$
$$\cdot \int_A I_{[0,Y_j]} h_F(v) dF(v). \tag{23}$$

By independence, the expected value of (23) is

$$n^{-2} n(n-1)(F(A))^2 \to (F(A))^2.$$

Thus, (21) will follow if it can be shown that the expected value of (22) converges to 0. We may apply a d-dimensional integration by parts formula to show that for an integrable function $f : [0,1]^d \to \mathbf{R}$, $(f(z))^2$ may be expressed as a finite linear combination of integrals of the form $\int_{R_z} f(u') df(u)$, where u' is such that $u'_j = u_j \ \forall j \in J$ where J is some nonempty subset of $\{1, ..., d\}$, and $u'_j = z_j \ \forall j \notin J$. If we let

$$f(z) = \int_{R_z} I_{[0,Y_i]} h_F(u) dF(u),$$

then a straightforward calculation shows that

$$E\left[\int_{R_z} f(u') df(u) \right] \leq h_F(z) = (F(S_z))^{-1} < \infty \tag{24}$$

if $z \prec 1$, and so the expected value of (22) converges to 0 as required. □

Remark 7.5 When $d \geq 2$ and $\mathcal{A}$ is the class of lower layers in $[0,1]^d$, Proposition 7.2 implies flow-wise convergence of (U_n) to a Gaussian white noise, almost every sample path of which is necessarily discontinuous at every $A \in \mathcal{A}$ with respect to the Hausdorff metric on $\mathcal{A}$ (cf. [1], Proposition 1.3).

References

[1] Adler, R. J. (1990). *An Introduction to Continuity, Extrema, and Related Topics for General Gaussian Processes*, Institute of Mathematical Statistics, Lecture Notes-Monograph Series, **12**.

[2] Aldous, D. (1978). Stopping times and tightness, *The Annals of Probability* **6**, 335-340.

[3] Alexander, K. and Bass, R. F. (1986). A uniform central limit theorem for set-indexed partial-sum processes with finite variance, *The Annals of Probability* **14**, 582-597.

[4] Ash, R. (1972). *Real Analysis and Probability*, Academic Press.

[5] Bass, R. F. and Pyke, R. (1985). The space $D(\mathcal{A})$ and weak convergence for set-indexed processes, *The Annals of Probability* **13**, 860-884.

[6] Bass, R. F. and Pyke, R. (1987). A central limit theorem for $D(\mathcal{A})$-valued processes, *Stochastic Processes and their Applications* **24**, p.109-131.

[7] Bickel, P. J. and Wichura, M. J. (1971). Convergence criteria for multiparameter processes and some applications, *Annals of Mathematical Statistics* **42**, 1656-1670.

[8] Billingsley, P. (1986). *Probability and Measure*, 2nd. ed., Wiley.

[9] Burke, M. D. (1997). Multivariate tests-of-fit and uniform confidence bands using a Gaussian bootstrap (preprint).

[10] Cairoli, R. and Walsh, J. B. (1975). Stochastic integrals in the plane, *Acta Mathematica* **134**, 111-183.

[11] Dozzi, M., Ivanoff, B. G. and Merzbach, E. (1994). Doob-Meyer decomposition for set-indexed submartingales, *Journal of Theoretical Probability* **7**, 499-525.

[12] Helland I. S. (1982). Central limit theorems for martingales with discrete or continuous time, *Scandinavian Journal of Statistics* **9**, 79-94.

[13] Ivanoff, B. G. (1985). Poisson convergence for point processes on the plane, *Journal of the Australian Mathematical Society* **39**, 253-269.

[14] Ivanoff, B. G. (1996). Stopping times and tightness for multiparameter martingales, *Statistics and Probability Letters* **28**, 111-114.

[15] Ivanoff, B. G., Lin, Y. -X. and Merzbach, E. (1996). Weak convergence of set-indexed point processes and the Poisson processes, *Theory of Probability and Mathematical Statistics* **55**, 77-89.

[16] Ivanoff, B. G. and Merzbach, E. (1995). Stopping and set-indexed local martingales, *Stochastic Processes and their Applications* **57**, 83-98.

[17] Ivanoff, B. G. and Merzbach, E. (1996). A martingale characterization of the set-indexed Brownian motion, *Journal of Theoretical Probability* **9**, 903-913.

[18] Ivanoff, B. G. and Merzbach, E. (1997). Poisson convergence for set-indexed empirical processes, *Statistics and Probability Letters* **32**, 81-86.

[19] Jacod, J. and Shiryaev, A. N. (1987). *Limit Theorems for Stochastic Processes*, Springer-Verlag.

[20] McLeish, D. L. (1978). An extended martingale invariance principle, *The Annals of Probability*, **6**, 144-150.

[21] Pollard, D. (1984). *Convergence of Stochastic Processes*, Springer-Verlag.

[22] Pyke, R. (1983). A uniform CLT for partial-sum processes indexed by sets, *London Mathematical Society Lecture Notes* **79**, 219-240.

[23] Revuz, R. and Yor, M. (1991). *Continuous Martingales and Brownian Motion*, 2nd. ed., Springer-Verlag.

[24] Skorokhod, A. V. (1956). Limit theorems for stochastic processes, *Theory of Probability and its Applications* **1**, 261-289.

[25] Slonowsky, D. (1998), Central Limit Theorems for Set-Indexed Strong Martingales, Ph.D. thesis, University of Ottawa.