Strong Martingales: Their Decompositions and Quadratic Variation

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

July 1999

Abstract

Set-indexed strong submartingales and a form of predictability for set-indexed processes are defined. Under a natural integrability condition, we show that any set-indexed strong submartingale can be decomposed in the Doob-Meyer sense. A form of predictable quadratic variation for square-integrable set-indexed strong submartingales is defined and sufficient conditions for its existence are given. Under a conditional independence assumption, these reduce to a simple moment condition and the resulting quadratic variation admits discrete approximations in the L_2-sense.

AMS 1991 subject classification: 60G48, 60G60, 60G07
Keywords and Phrases: set-indexed strong martingale, increasing process, *-predictablility, class $(D')^*$, Doob-Meyer decomposition, quadratic variation, conditional independence, discrete approximations.

Running Head: Strong martingale decompositions.

*The results in this paper represent part of the authors Ph.D. dissertation under the direction of B. G. Ivanoff at the University of Ottawa. Research partially supported by a Natural Science and Engineering Research Council of Canada PGS-A Scholarship and an Ontario Graduate Scholarship.

1 Introduction

A cornerstone in the theory of continuous parameter processes is the Doob-Meyer decomposition in which conditions are given ensuring a submartingale $X = \{X_t : t \in [0, \infty)\}$ can be uniquely decomposed into the sum of a martingale and a predictable increasing process. When applied to the square of a square-integrable martingale, the corresponding increasing process, termed the predictable quadratic variation, plays an important role in the theory of stochastic integration and martingale limit theorems. The Doob-Meyer decomposition has been extended by several authors to processes indexed by the points of the plane or, more generally, the points of a directed set. A consise history of such results can be found in the introduction of Dozzi et. al. (1994). In this paper we obtain such a decomposition and study a form of quadratic variation for the class of set-indexed strong submartingales, stochastic processes indexed by the elements of a subcollection $\mathcal{A}$ of closed subsets of a fixed topological space T.

This paper has five sections and one appendix. Section 2 begins with the necessary conditions on T and the class $\mathcal{A}$. The appropriate choice of indexing collection $\mathcal{A}$, which replaces the usual parameter spaces $[0, 1]$ or $[0, 1]^d$, is essential for the theory to work. Although our framework is similar to that found in many earlier set-indexed papers, it is worth mentioning that nowhere do we require $\mathcal{A}$ to satisfy the restrictive "shape property" and T is only required to be metrizable in the latter part of Section 5. After defining set-indexed processes $X = (X_A)_{A\in\mathcal{A}}$ and set-indexed filtrations $(\mathcal{F}_A)_{A\in\mathcal{A}}$, adapted processes are defined to be those for which X_A is $\mathcal{F}_A$-measurable $\forall A \in \mathcal{A}$. This section also contains the definition and examples of set-indexed strong submartingales.

In Section 3, the Doob-Meyer decomposition for set-indexed strong submartingales is given. After defining class $(D')^*$ strong submartingales—a set-indexed generalization of class D continuous parameter submartingales—and formulating a set-indexed concept of predictability termed *-predictability, the said decomposition is obtained in two phases. In the first, it is shown that any class $(D')^*$ strong submartingale X can be uniquely decomposed into a sum $X = M + V$ where V is a *-predictable process and M is a strong martingale "up to adaptedness" (Theorem 3.5). The proof, which is essentially that of Dozzi et. al. (1994), yields discrete approximations of V in the weak L_1-sense. (Unlike its continuous pararmeter counterpart, a *-predictable process is not necessarily adapted.) In the second phase, we show that under a generalized form of the F4 conditional independence assumption of Cairoli and Walsh (1975), both M and V are adapted (Proposition 3.6).

In contrast to the classical situation, the square of a square-integrable set-indexed strong martingale M is not necessarily a strong submartingale. Hence, one cannot define the quadratic variation of M to be the *-predictable process in the decomposition of M^2. In Section 4, by following the ideas of Gushchin (1982), we circumvent this shortcoming to define a suitable form of quadratic variation for set-indexed strong martingales termed *-predictable quadratic variation, a *-predictable process which is not necessarily adapted. In Theorem 4.1, sufficient conditions for the existence of *-predictable quadratic variation are given. In addition, under the generalized F4 conditional independence assumption, it is shown that any set-indexed strong martingale in $L_{2+\delta}$, some $\delta > 0$, has a unique, adapted *-predictable quadratic variation (Proposition 4.4). In both cases, our results yeild discrete weak L_1-approximants for the *-predictable quadratic variation. In Section 5, conditions are given under which these approximations can upgraded to the L_2-norm sense.

The paper closes with an appendix containing consequences of the conditional independence assumption found in Dozzi et. al. (1994). One such consequence, a set-indexed Rosenthal-type inequality, is used extensively in Section 5.

2 Preliminaries

Central to the theory is the choice of an appropriate indexing class $\mathcal{A}$. Similar to Slonowsky and Ivanoff (1999), we take $\mathcal{A}$ to be a collection of closed subsets of a fixed compact Hausdorff topological space T such that

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

(b) $\mathcal{A}$ is closed under countable intersections and

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

Modulo obvious modifications, the results obtained in the sequel will remain valid for T non-compact provided there exists an increasing sequence (B_n) of compact subsets of T such that $B_n \in \mathcal{A}\ \forall n$, $\bigcup_n B_n = T$ and, given any $A \in \mathcal{A}$, $A \subseteq B_n$ for all sufficiently large n.

Three natural extensions of $\mathcal{A}$ are

$$\begin{aligned} \mathcal{A}(u) &= \{\text{all finite unions in } \mathcal{A}\}, \\ \mathcal{C} &= \{A \setminus B : A \in \mathcal{A}, B \in \mathcal{A}(u)\}, \\ \mathcal{C}(u) &= \{\text{finite unions in } \mathcal{C}\}. \end{aligned}$$

$\mathcal{C}$ is a semi-algebra of subsets of T implying $\mathcal{C}(u)$ is the algebra generated by $\mathcal{A}$. Clearly, $\mathcal{A} \subseteq \mathcal{C}$ and $\mathcal{A}(u) \subseteq \mathcal{C}(u)$.

$\mathcal{A}$ constitutes a semilattice under $A \wedge B = A \cap B$, hence we refer to any subcollection of $\mathcal{A}$ which is closed under intersection as a sub-semilattice of $\mathcal{A}$. 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 the set $C_A \in \mathcal{C}$ defined by

$$C_A := A \setminus \bigcup_{A' \in \mathcal{A}',\, A \not\subseteq A'} A'. \tag{1}$$

Since $\mathcal{A}'$ is finite and closed under intersections, $C_A = A \setminus \bigcup_{A' \in \mathcal{A}',\, A' \subset A} A'$ for each $A \in \mathcal{A}'$.

In addition to conditions (a), (b) and (c), we will assume throughout that $\mathcal{A}$ satisifies the following two assumptions, variants of which have appeared in the earlier set-indexed martingale literature (for example, cf. Dozzi et. al. (1994) and Ivanoff and Merzbach (1996)).

Assumption 2.1 *(Separability.)* 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 $g_n : \mathcal{A} \to \mathcal{A}_n$ 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 $\mathcal{A}$).

Since $\mathcal{A}$ is closed under finite intersections, any $C \in \mathcal{C}$ can be written $A \setminus \bigcup_{i=1}^{n} A_i$ $(A, A_i \in \mathcal{A})$ where $A_i \not\subseteq A_j$ $\forall\, i \neq j$ and $A_i \subseteq A$ $\forall\, i$. Such representations are termed *minimal*. On the other hand, in general, the existence of *maximal representations* must be assumed.

Assumption 2.2 *(Existence of maximal representations.)* Given any $C = A \setminus B \in \mathcal{C}$ $(A \in \mathcal{A}$, $B \in \mathcal{A}(u))$, there exists $A_1, \cdots, A_n \in \mathcal{A}$ with $A_i \not\subseteq A_j$ $\forall\, i \neq j$ such that $C = A \setminus \bigcup_{i=1}^{n} A_i$ and, given any $B' \in \mathcal{A}(u)$, $B' \cap C = \phi$ implies $B' \subseteq \bigcup_{i=1}^{n} A_i$.

When $T = [0,1]^d$ for some $d \in \mathbf{N}$, examples of $\mathcal{A}$ satisfying Assumptions 2.1 and 2.2 include the *lower rectangles*, $\{R_z : z \in [0,1]^d\}$ where $R_z := \prod_{i=1}^{d} [0, z_i]$, $z \in [0,1]^d$, and the collection $\mathcal{LL}_d$ of *lower layers* in $[0,1]^d$. A closed set $L \subseteq [0,1]^d$ is a lower layer if $z \in L$ implies $R_z \subseteq L$. Both examples can be extended to the σ-compact case of $T = \mathbf{R}_+^d$ or $T = \mathbf{R}^d$ (cf. Ivanoff and Merzbach (1996)). Examples with T non-Euclidean are given in Slonowsky and Ivanoff (1999).

Since $\mathcal{A}$ is closed under countable intersection and each $\mathcal{A}_n$ is finite,

$$\bigcap_{i \in I} A_i \;=\; \bigcap_{n \in \mathbf{N}} \bigcap_{i \in I} g_n(A_i) \;\in \mathcal{A}$$

for any subcollecton $\{A_i : i \in I\}$ of $\mathcal{A}$. In particular, $\phi' \in \mathcal{A}$ where $\phi' = \bigcap_{A \in \mathcal{A},\, A \neq \phi} A$. The role played by ϕ' in the sequel will be similar to that of 0 in the continuous parameter theory. Without loss of generality, we can assume that every sub-semilattice of $\mathcal{A}$ contains ϕ' and T but not ϕ.

The elements of a finite sub-semilattice $\mathcal{A}'$ of $\mathcal{A}$ can always be numbered $A_1, \cdots, A_k$ for some $k \in \mathbf{N}$ so that $A_1 = \phi'$ and, given any $2 \leq i \leq k$, $A_j \subset A_i$ implies $j \leq i-1$. Following Ivanoff and Merzbach (1996), we refer to any such numbering as being *consistent with the strong past* in which case, given any $2 \leq i \leq k$,

$$C_{A_i} \;=\; A_i \setminus \textstyle\bigcup_{j=1}^{i-1} A_j\,. \tag{2}$$

Fix 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 $\mathcal{F}$ which is right-continuous in the following sense: $\mathcal{F}_{\bigcap_n A_n} = \bigcap_n \mathcal{F}_{A_n}$ for any decreasing sequence (A_n) in $\mathcal{A}$. We can extend any such $(\mathcal{F}_A)$ to an $\mathcal{A}(u)$-indexed family by defining $\mathcal{F}_B = \bigvee_{A \in \mathcal{A},\, A \subseteq B} \mathcal{F}_A$ for any $B \in \mathcal{A}(u)$. Given a set $C \in \mathcal{C}(u)$, the *strong past* at C is then defined to be the sub-σ-algebra

$$\mathcal{G}_C^* \;=\; \bigvee_{\substack{B \,\in\, \mathcal{A}(u) \\ B \,\cap\, C \,=\, \phi}} \mathcal{F}_B \tag{3}$$

for $C \notin \mathcal{A}(u)$ and $\mathcal{G}_C^* = \mathcal{F}_{\phi'}$ for $C \in \mathcal{A}(u)$. Since any finite sub-semilattice of $\mathcal{A}$ can be numbered in a manner consistent with the strong past, (2) and (3) yield the following result.

Lemma 2.3 *Given two distinct elements A and A' of a finite sub-semilattice $\mathcal{A}'$ of $\mathcal{A}$, either $\mathcal{F}_A \subseteq \mathcal{G}_{C'}^*$ or $\mathcal{F}_{A'} \subseteq \mathcal{G}_C^*$ where C and C' denote the left-neighborhoods of A and A' in $\mathcal{A}'$ respectively.*

Any collection $X = \{X_A : A \in \mathcal{A}\}$ of random variables on $(\Omega, \mathcal{F}, P)$ is referred to as a *set-indexed process*. In this paper, we only work with processes X which possess unique finitely additive extensions to $\mathcal{C}(u)$, i.e., $X_{C \cup D} = X_C + X_D$ for any two disjoint sets $C, D \in \mathcal{C}(u)$. Hence, given any $C = A \setminus \bigcup_{i=1}^{k} A_i$ where $A, A_i \in \mathcal{A}$ and $k \in \mathbf{N}$,

$$X_C \;=\; \sum_{I \subseteq \{1, \cdots, k\}} (-1)^{|I|} \cdot X(A \cap \textstyle\bigcap_{i \in I} A_i)\,. \tag{4}$$

Examples of processes possessing such extensions include purely atomic processes and processes indexed by a class $\mathcal{A}$ satisfying the so-called "shape property" ($A \subseteq \bigcup_{i=1}^{n} A_i$ with $A, A_i \in \mathcal{A}$ implies $A \subseteq A_i$). A necessary condition for such an extension is $X_\phi = 0$.

Given an $\mathcal{A}$-indexed process X and $1 \leq p \leq \infty$, 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*. An integrable process $X = \{X_A : A \in \mathcal{A}\}$ is said to be adapted to $(\mathcal{F}_A)$ if X_A is $\mathcal{F}_A$-measurable $\forall A \in \mathcal{A}$. In the absense of a metric on T, we define right-continuity for set-indexed processes as done in Dozzi et. al. (1994).

Definition 2.4 *Let $X = \{X_A : A \in \mathcal{A}\}$ be a process on $(\Omega, \mathcal{F}, P)$.*

(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) *Given $1 \leq p < \infty$, X is monotone L_p-outer-continuous if for any decreasing sequence (A_n) in $\mathcal{A}$, $X(A_n) \to X(\bigcap_n A_n)$ in L_p-norm.*

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

(iv) *X is increasing (with respect to a filtration $(\mathcal{F}_A)$ on $(\Omega, \mathcal{F}, P)$) if it is raw increasing and adapted.*

By a classic argument (cf. Dozzi et. al. (1994), Proposition 2.4),

Lemma 2.5 *Let $X = (X_A)_{A \in \mathcal{A}}$ and $Y = (Y_A)_{A \in \mathcal{A}}$ be two monotone outer-continuous processes. If X is a modification of Y in the sense that $X_A = Y_A$ a.s., any given $A \in \mathcal{A}$, then X and Y are indistinguishable, i.e., for a.e. ω, $X_A(\omega) = Y_A(\omega) \ \forall A \in \mathcal{A}$.*

Applying (A1) and (4),

Lemma 2.6 *If $X = (X_A)_{A \in \mathcal{A}}$ is a monotone L_p-outer-continuous process, then given any $C = A \setminus \bigcup_{i=1}^{k} A_i$ $(A, A_i \in \mathcal{A})$, $E[|X_{C \setminus C_n}|^p] \to 0$ as $n \to \infty$ when we take either $C_n = A \setminus \bigcup_{i=1}^{k} g_n(A_i)$ $\forall n$ or $C_n = g_n(A) \setminus \bigcup_{i=1}^{k} g_n(A_i)$ $\forall n$.*

The following terminology first appeared in Ivanoff and Merzbach (1995) and generalizes the planar strong submartingales introduced by Cairoli and Walsh (1975).

Definition 2.7 *An adapted integrable process $X = \{X_A : A \in \mathcal{A}\}$ is a strong (sub)martingale if $E[X_C | \mathcal{G}_C^*] (\geq) = 0$ for every $C \in \mathcal{C}$.*

Comments. (1) Since $(\mathcal{G}_C^*)_{C \in \mathcal{C}(u)}$ is a decreasing family, if X is a strong (sub)martingale, the tower property and the finite additivity of X imply $E[X_C | \mathcal{G}_C^*] (\geq) = 0 \ \forall C \in \mathcal{C}(u)$.

(2) Examples of strong martingales include processes with independent increments such as set-indexed Gaussian processes and, when $T = [0,1]^d$ for some $d \in \mathbf{N}$, the weighted empirical process (cf. Slonowsky and Ivanoff (1999)). Clearly, any increasing process is a strong submartingale.

(3) In Dozzi et. al. (1994), two additional notions of submartingale were introduced. An adapted integrable process $X = (X_A)_{A \in \mathcal{A}}$ was termed a *set-indexed submartingale* if $E[X_B | \mathcal{F}_A] \geq X_A$ for every $A \subseteq B$ whereas X was termed a *set-indexed weak submartingale* if $E[X_C | \mathcal{G}_C] \geq 0 \ \forall C \in \mathcal{C}$.

Here, $\mathcal{G}_C$, the *weak past* at $C \in \mathcal{C}$ is the σ-algebra $\mathcal{G}_C = \bigcap_{A\in\mathcal{A},\, A\cap C\neq\phi} \mathcal{F}_A$.

In general, X adapted to $(\mathcal{F}_A)$ does not necessarily imply that X_C is $\mathcal{G}^*_C$-measurable for every $C \in \mathcal{C}$. However,

Lemma 2.8 *Fix a finite sub-semilattice $\mathcal{A}'$ of $\mathcal{A}$ and two distinct elements A and A' of $\mathcal{A}'$. Let C and C' denote the left-neighborhoods of A and A' in $\mathcal{A}'$ respectively.*

(a) *If X is adapted to $(\mathcal{F}_A)$, then either X_C is $\mathcal{G}^*_{C'}$-measurable or $X_{C'}$ is $\mathcal{G}^*_C$-measurable.*

(b) *If X is a strong martingale in L_2 and $g \in L_\infty$ is $\mathcal{G}^*_C \cap \mathcal{G}^*_{C'}$-measurable, then $E[g\, X_C\, X_{C'}] = 0$.*

Proof. Without loss of generality, Lemma 2.3 implies $\mathcal{F}_A \subseteq \mathcal{G}^*_{C'}$ so that (a) follows by (4). Therefore, (b) follows by conditioning and the strong martingale property since $E[g\, X_C\, X_{C'}] = E[g\, X_C \cdot E(X_{C'}|\mathcal{G}^*_{C'})] = 0$. □

In Dozzi et. al. (1994), it was shown that any set-indexed martingale in L_p is monotone L_p-outer-continuous. Since strong martingales are automatically set-indexed martingales, we have

Lemma 2.9 *If $X = (X_A)_{A\in\mathcal{A}}$ is a strong martingale in L_p for some $1 \le p < \infty$, then X is monotone L_p-outer-continuous.*

We refer to any finite collection $\{X_i, \mathcal{H}_i : i = 1, \cdots, r\}$ of sub-σ-algebras $\mathcal{H}_1, \cdots, \mathcal{H}_r$ of $\mathcal{F}$ and random variables $X_1, \cdots, X_r \in L_1$ as a *martingale difference array* if $\mathcal{H}_i \subseteq \mathcal{H}_{i+1}\ \forall 1 \le i \le r-1$, X_i is $\mathcal{H}_i$-measurable $\forall 1 \le i \le r$ and $E[X_{i+1}|\,\mathcal{H}_i] = 0\ \forall 1 \le i \le r-1$.

Lemma 2.10 *Let M be a strong martingale and let $\mathcal{A}' = \{A_1, \cdots A_n\}$ be any finite sub-semilattice of $\mathcal{A}$ numbered in a manner consistent with the strong past. Then, given indices $1 \le k_1 < \cdots < k_r \le n$ with $C_i := C_{A_{k_i}} \neq \phi\ \forall i$, $\{X_i, \mathcal{H}_i : 1 \le i \le r\}$ forms a martingale difference array when we take $\mathcal{H}_i = \bigvee_{j=1}^{(k_{i+1})-1} \mathcal{F}_{A_j}$ $(1 \le i \le r-1)$, $\mathcal{H}_r = \mathcal{F}$ and $X_i = M(C_i)\ \forall 1 \le i \le r$.*

Proof. Since k_i increases in i, $\mathcal{H}_i \subseteq \mathcal{H}_{i+1}\ \forall 1 \le i \le r-1$. By (4), $M(C_i)$ is $\mathcal{F}_{A_{k_i}}$-measurable $\forall 1 \le i \le r$. Therefore, since $k_i \le k_{i+1} - 1\ \ \forall 1 \le i \le r-1$, it is clear that X_i is $\mathcal{H}_i$-measurable $\forall 1 \le i \le r-1$. (Trivially, X_r is $\mathcal{H}_r$-measurable.)

Fix $1 \le i \le r-1$. Since $\mathcal{A}'$ is numbered in a manner consistent with the strong past, (2) implies $C_{i+1} = A_{k_{i+1}} \setminus \bigcup_{j=1}^{(k_{i+1})-1} A_j$. If $B = \bigcup_{j=1}^{(k_{i+1})-1} A_j$, then $\mathcal{F}_B \subseteq \mathcal{G}^*_{C_{i+1}}$. Furthermore, since $(\mathcal{F}_B)_{B\in\mathcal{A}(u)}$ is increasing, $\mathcal{H}_i \subseteq \mathcal{F}_B$. Therefore, $E[X_{i+1}|\,\mathcal{H}_i] = E[E[M(C_{i+1})|\mathcal{G}^*_{C_{i+1}}]\,|\mathcal{H}_i] = 0$ when we apply the tower property and the strong martingale property respectively. □

In view of Lemma 2.10, Burkholder's inequality yields a set-indexed Burk- holder-type inequality.

Lemma 2.11 *Let M be a strong martingale in L_2. If $k_1, \cdots, k_r$ and $C_1, \cdots, C_r$ are as defined in Lemma 2.10, then given any $1 < p < \infty$,*

$$E\left[\left(\sum_{i=1}^{r} (M_{C_i})^2\right)^{p/2}\right] \ \le\ \kappa \cdot E\left[\,|M_{C_0}|^{\,p}\,\right] \tag{5}$$

where $C_0 = \bigcup_{i=1}^r C_i$ and κ is a positive constant depending only on p.

3 A Doob-Meyer Decompositon

Fix a stochastic base, $(\Omega, \mathcal{F}, P, (\mathcal{F}_A), \mathcal{A})$. As done in the continuous parameter theory and the theory of set-indexed weak submartingales (cf. Dozzi et. al. (1994)), we wish to determine a broad class of strong submartingales $X = (X_A)_{A\in\mathcal{A}}$ which can be uniquely written as

$$X_A = V_A + M_A \quad \forall\, A \in \mathcal{A} \tag{6}$$

where M is a strong martingale and V is an increasing process which is in some sense predictable.

In this section, after defining a suitable form of predictablility for set-indexed processes, we obtain such a decomposition in two stages. First, we obtain (6) where V is raw increasing and $E[M_C \,|\, \mathcal{G}^*_C] = 0 \;\forall\, C \in \mathcal{C}$ while M is not necessarily adapted. Then, sufficient conditions on $\mathcal{A}$ and $(\mathcal{F}_A)$ are given under which M and V are adapted to $(\mathcal{F}_A)$, resulting in a true Doob-Meyer decomposition.

In the classical theory, given a filtration $(\mathcal{F}_t)_{t\in[0,\infty)}$ on $(\Omega, \mathcal{F}, P)$, one defines the predictable σ-algebra Σ_p on $\Omega \times [0,\infty)$ to be that generated by all sets of the form $F \times \{0\}$, $F \in \mathcal{F}_0$, along with all sets of the form $F \times (s,t]$ where $s < t$ and $F \in \mathcal{F}_s$. By way of set-indexed analogue,

Definition 3.1 *Any set of the form $F \times C$ $(C \in \mathcal{C},\ F \in \mathcal{G}^*_C)$ is said to be a *-predictable rectangle. The collection of all *-predictable rectangles is denoted $\mathcal{P}^*_0$ and the *-predictable σ-algebra, $\mathcal{P}^*$ on $\Omega \times T$ is defined by $\mathcal{P}^* = \sigma(\mathcal{P}^*_0)$.*

In the continuous parameter setting, an adapted process $X : \Omega \times [0,\infty) \to \mathbf{R}$ is termed predictable if it is Σ_p-measurable. However, set-indexed processes have domain $\Omega\times\mathcal{A}$ and hence cannot themselves be $\mathcal{P}^*$-measurable. Instead, we define *-predictablility in an indirect fasion following that done in Dozzi et. al. (1994). This requires an additional assumption.

Assumption 3.2 For each $F \in \mathcal{F}$, there exists a collection $Y(F) = \{Y(F,t) : t \in T\}$ of random variables on $(\Omega, \mathcal{F}, P)$ such that

(i) the map $(\omega, t) \mapsto Y(F,t)(\omega)$ is $\mathcal{P}^*$-measurable and

(ii) for each $t \in T$, $Y(F,t)$ is a version of $E[\mathbf{1}_F \,|\, \mathcal{H}_t]$

where $\mathcal{H}_t = \bigvee_n \mathcal{G}^*_{C^n_t}$, C^n_t the left-neighbourhood of $A^n_t = \bigcap_{A\in\mathcal{A}_n, t\in A} A$ in $\mathcal{A}_n$. Moreover, the process $Y(F)$ is unique up to indistinguishability on T.

As commented in Dozzi et. al. (1994), Assumption 3.2 replaces the theorem of predictable projection of the classical theory and appears to be necessary for a Doob-Meyer decomposition of set-indexed strong submartingales.

Given an integrable but not necessarily adapted process $X = (X_A)_{A\in\mathcal{A}}$, the admissible function associated to X is the function μ_X defined by

$$\mu_X(F\times C) = E[\mathbf{1}_F\, X_C] \quad \forall\, C \in \mathcal{C}\,,\ F \in \mathcal{F}\,. \tag{7}$$

Clearly, a process $X = (X_A)_{A\in\mathcal{A}}$ is raw increasing if and only if $\mu_X = 0$ on $\{F\times C : C \in \mathcal{C},\ F \in \mathcal{F}\}$. Furthermore, if X is adapted, it is a strong (sub)martingale if and only if $\mu_X\ (\geq) = 0$ on $\mathcal{P}^*_0$.

Any admissible function is finitely additive on the semi-algebra $\mathcal{P}_0^*$ and therefore can be uniquely extended to a finitely additive function on the algebra $\mathcal{P}_0^*(u)$ consisting of all finite disjoint unions in $\mathcal{P}_0^*$. Furthermore, by Proposition 4.1 in Dozzi et. al. (1994), the admissible function of a raw increasing process necessarily extends to a unique measure on $\sigma(\mathcal{F} \times \mathcal{B}(\mathcal{A}))$ where $\mathcal{B}(\mathcal{A})$ denotes the σ-algebra on T generated by the sets in $\mathcal{A}$. This permits the following terminology.

Definition 3.3 *A raw increasing process $V = (V_A)_{A \in \mathcal{A}}$ is said to be *-predictable if for any $F \in \mathcal{F}$ and any $C \in \mathcal{C}$,*

$$\mu_V(F \times C) = \int_{\Omega \times C} Y(F,t)(\omega)\, d\mu_V(\omega, t) \tag{8}$$

Comments. (a) Our definition of *-predictability is motivated by Definition 4.1 of Dozzi et. al. (1994) in which a type of set-indexed predictability suited for weak submartingale decompositions was defined. In contrast, our definition does not require *-predictable processes to be adapted

(b) In the classic situation where $T = [0, \infty)$ and $\mathcal{A} = \{[0,t] : t \geq 0\}$, if $V = (V_t)_{t \in [0,\infty)}$ is adapted, then condition (8) is equivalent to V being equal to its dual predictable projection.

Definition 3.4 *A process $X = (X_A)_{A \in \mathcal{A}}$ is of class $(D')^*$ if*

$$\Big\{ \sum_{D \in \mathcal{N}_n} |E[X_D \mid \mathcal{G}_D^*]| : n \in \mathbf{N} \Big\}$$

is uniformly integrable where

$$\mathcal{N}_n = \{C_A : A \in \mathcal{A}_n\} \setminus \{\phi\}$$

denotes the collection of all non-empty disjoint left-neighborhoods generated by $\mathcal{A}_n$.

Comment. Our definition of class $(D')^*$ processes is intended to mimic the class D' processes from the continuous parameter theory. A similar class of set-indexed process was introduced in Dozzi et. al. (1994) except with the strong past $\mathcal{G}_C^*$ replaced by the weak past $\mathcal{G}_C$.

In Theorem 5.1 of Dozzi et. al. (1994), it has been shown that any monotone L_1-right-continuous weak submartingale X satisfying a close analogue of the class $(D')^*$ property possesses a unique decomposition $X = M + V$ where M is a weak martingale and V is an increasing process possessing a form of predictability similar to that found in Definition 3.3. With some minor modifications, by replacing "weak past" and "weak submartingale" by "strong past" and "strong submartingale" in the proof thereof, we obtain

Theorem 3.5 *Under Assumptions 2.1, 2.2 and 3.2, given a monotone L_1-right-continuous strong submartingale X of class $(D')^*$, there exists processes M and V, both unique up to indistinguishablity, such that*

(a) $X_A = V_A + M_A \;\; \forall A \in \mathcal{A}$,

(b) *V is *-predictable and*

(c) $E[M_C | \mathcal{G}_C^*] = 0 \;\; \forall C \in \mathcal{C}$

(M and V are not necessarily adapted). Moreover, given any $A \in \mathcal{A}$,

$$V_A \;=\; \lim_{m\to\infty} \lim_{n\to\infty,\, n\geq m} V^{(n)}_{g_m(A)} \quad \text{in the weak } L_1 \text{ topology} \tag{9}$$

where, given any $n \geq m$ and $B \in \mathcal{A}_m$, $V^{(n)}_B$ is the random variable $V^{(n)}_B = \sum_{D\in\mathcal{N}_n, D\subseteq B} E[X_D \,|\, \mathcal{G}^*_D]$.

In Dozzi et. al. (1994), adaptedness of the increasing process V in the Doob-Meyer decomposition for set-indexed weak submartingales followed automatically from (9), (4) and the inclusion $\mathcal{G}_C \subseteq \mathcal{F}_A$, $C = A \setminus B$ ($A \in \mathcal{A}$, $B \in \mathcal{A}(u)$). Whereas the inclusion $\mathcal{G}^*_C \subseteq \mathcal{F}_A$ is not valid in general, we have the following

Proposition 3.6 *Under (32) and Assumptions 2.1, 3.2, A.5, A.6 and A.7, the processes V and M in Theorem 3.5 are both adapted to $(\mathcal{F}_A)$.*

Proof. Given a sub-σ-algebra $\mathcal{G}$ of $\mathcal{F}$, let $L_1(\mathcal{G})$ denote the space of all $\mathcal{G}$-measurable random variables in L_1. Fix $A \in \mathcal{A}$. Since our filtration is right-continuous and subspaces of the form $L_1(\mathcal{G})$ are closed in the weak L_1 topology, it is sufficient by (9) to show $V^{(n)}_{g_m(A)} \in L_1(\mathcal{F}_{g_m(A)})$ for every $n, m \in \mathbf{N}$ such that $n \geq m$.

To this end, fix a pair $n \geq m$ and select $D \in \mathcal{N}_n$ such that $D \subseteq g_m(A)$. Since $D = A' \setminus \bigcup_{A''\in\mathcal{A}_n, A'\not\subseteq A''} A''$ for some $A' \in \mathcal{A}_n$, it must be that $A' \subseteq g_m(A)$. Therefore, since (4) implies X_D is $\mathcal{F}_{A'}$-measurable, the $\mathcal{F}_{g_m(A)}$-measurablility of $V^{(n)}_{g_m(A)}$ follows by Corollary A.10 and the inclusion $\mathcal{F}_{A'} \subseteq \mathcal{F}_{g_m(A)}$. □

4 The Existence of Quadratic Variation

In the continuous parameter theory, since the square of an L_2 martingale $M = (M_t)_{t\in[0,\infty)}$ is a class D submartingale, one defines the predictable quadratic variation $\langle M\rangle$ of M to be the unique increasing process for which $M^2 - \langle M\rangle$ is a martingale. In contrast, the square of a strong martingale M in L_2 need not be a strong submartingale (cf. Gushchin (1982)). Furthermore, in view of (4), $E[M^2_C|\mathcal{G}^*_C]$ is not necessarily non-negative for every $C \in \mathcal{C}$. For these reasons, we follow the lead of Gushchin (1982) and work with the collection of $\{(M_C)^2 : C \in \mathcal{C}\}$ of squared increments rather that the collection $\{M^2_C : C \in \mathcal{C}\}$ of increments of the square. In this way, we obtain the following,

Theorem 4.1 *Under Assumptions 2.1, 2.2 and 3.2, if M is a strong martingale in L_2 for which*

$$\Big\{ \sum_{D\in\mathcal{N}_n} E[\,(M_D)^2 \,|\, \mathcal{G}^*_D] \,:\, n \in \mathbf{N} \Big\} \tag{10}$$

*is uniformly integrable, then there exists a *-predictable process Q, unique up to indistinguishability, for which*

$$E[(M_C)^2|\mathcal{G}^*_C] \;=\; E[Q_C|\mathcal{G}^*_C] \quad \forall C \in \mathcal{C} \tag{11}$$

(Q is not necessarily adapted). Moreover,

$$Q_A \;=\; \lim_{m\to\infty} \lim_{n\to\infty,\, n\geq m} Q^{(n)}_{g_m(A)} \quad \text{in the weak } L_1 \text{ topology} \tag{12}$$

where, given any $n \geq m$ and $B \in \mathcal{A}_m$, $Q^{(n)}_B$ is the random variable given by $Q^{(n)}_B = \sum_{D\in\mathcal{N}_n, D\subseteq B} E[\,(M_D)^2 \,|\, \mathcal{G}^*_D]$.

Comments. (1) The process Q in the preceding theorem represents a form of predictable quadratic variation for the set-indexed strong martingale M. However, it is important to note that, unlike the continuous parameter setting, $M^2 - Q$ is not necessarily a strong martingale. First, as mentioned above, Q is not necessarily adapted. Secondly, even if Q is adapted, (11) does not necessarily imply $E[(M^2-Q)_C|\, \mathcal{G}^*_C] = 0 \ \forall C \in \mathcal{C}$. Such difficulties are unavoidable, even for the case of two-parameter strong martingales (cf. Gushchin (1982)).

(2) Since M^2_C does not coincide with $(M_C)^2$ in general, (10) is not necessarily equivalent to M^2 being of class $(D')^*$.

Define $\mu_{(M)^2} : \mathcal{P}^*_0 \to [0,\infty)$ by setting $\mu_{(M)^2}(F \times C) = E[\,\mathbf{1}_F\,(M_C)^2]$ for any $F \times C \in \mathcal{P}^*_0$. (Note that $\mu_{(M)^2}$ is not necessarily the admissible function of the process M^2.) The key step in the proof of Theorem 4.1 is the following,

Proposition 4.2 *Under Assumptions 2.1 and 2.2, if M is a strong martingale for which the family in* (10) *is uniformly integrable, then $\mu_{(M)^2}$ extends to a unique measure on $\mathcal{P}^*$.*

Proof. Assuming $\mu_{(M)^2}$ is finitely additive on $\mathcal{P}^*_0$, the existence of such an extension follows by an argument identical to that of Theorem 3.1 (iii) in Dozzi et. al. (1994). In particular, the uniform integrability of (10) plays the role of their class D' condition while the monontone L_2-outer-continuity of M (guaranteed by Lemma 2.9) replaces L_1-right-continuity. However, unlike the case for admissible functions, the finite additivity of $\mu_{(M)^2}$ on $\mathcal{P}^*_0$ is no longer trivial since $C \mapsto (M_C)^2$ $(C \in \mathcal{C}(u))$ is not necessarily finitely additive on $\mathcal{C}(u)$. Instead, taking disjoint sets $F_1 \times C_1, \cdots, F_n \times C_n$ in $\mathcal{P}^*_0$ such that $\bigcup_{i=1}^n F_i \times C_i = F \times C \in \mathcal{P}^*_0$, we argue by cases as done in Gushchin (1982).

First, consider the case in which $F = F_1 = \cdots = F_n$. Since $F \times C = \bigcup_{i=1}^n F_i \times C_i = F \times (\bigcup_{i=1}^n C_i)$, $C = \bigcup_{i=1}^n C_i$ a disjoint union, the finite additivity of M implies $\mu_{(M)^2}(F \times C) = E[\mathbf{1}_F\,(\sum_{i=1}^n M_{C_i})^2]$. In fact, $\mu_{(M)^2}(F \times C) = E[\mathbf{1}_F \sum_{i=1}^n (M_{C_i})^2]$ since

<u>Claim:</u> $E[\mathbf{1}_F\, M_{C_i}\, M_{C_j}] = 0 \ \forall\, i \neq j$.

<u>Proof:</u> To simplify notation, take $i = 1$ and $j = 2$. Clearly, there exists a finite sub-semilattice $\mathcal{A}'$ of $\mathcal{A}$ and disjoint left-neighborhoods $D_1, \cdots, D_n$ generated by elements in $\mathcal{A}'$ such that $C_1 = \bigcup_{i=1}^m D_i$ and $C_2 = \bigcup_{j=m+1}^r D_j$. Since $F \in \mathcal{G}^*_C$ and $(\mathcal{G}^*_D)_{D \in \mathcal{C}}$ is decreasing, $\mathbf{1}_F$ is $\mathcal{G}^*_{D_k}$-measurable $\forall\, 1 \le k \le r$. Therefore, by Lemma 2.8 (b), $E[\,\mathbf{1}_F\, M_{D_i}\, M_{D_j}] = 0 \ \forall 1 \le i \le m$ and $m+1 \le j \le r$ so that $E[\mathbf{1}_F\, M_{C_1}\, M_{C_2}] = 0$ follows by the finite additivity of M.

Next, consider the case in which any two of the C_i are either disjoint or equal. If we let $\Delta_1, \cdots, \Delta_k$ denote all distinct (hence disjoint) sets among the C_i, then since $F \times C = \bigcup_{i=1}^n F_i \times C_i$, it is clear that $F = \bigcup_{i\,:\,C_i = \Delta_j} F_i$ is a disjoint union $\forall\, 1 \le j \le k$. This implies

$$\sum_{i=1}^n \mu_{(M)^2}(F_i \times C_i) \;=\; \sum_{j=1}^k \sum_{i\,:\,C_i=\Delta_j} E[\mathbf{1}_{F_i}\,(M_{\Delta_j})^2] \;=\; \sum_{j=1}^k E[\mathbf{1}_F\,(M_{\Delta_j})^2]$$

so that finite additivity in this case follows by the previous case.

Finally, let $F \times C = \bigcup_{i=1}^{n} F_i \times C_i$ be any disjoint union in $\mathcal{P}_0^*$. Then, as in the above claim, there is a finite sub-semilattice $\mathcal{A}'$ of $\mathcal{A}$ such that for each $1 \le i \le n$, there exists disjoint left-neighborhoods $D_1^i, \cdots, D_{k(i)}^i$ generated by elements in $\mathcal{A}'$ such that $C_i = \bigcup_{k=1}^{k(i)} D_k^i$. Since $(\mathcal{G}_D^*)_{D \in \mathcal{C}}$ is a decreasing family, $F_i \times D_k^i \in \mathcal{P}_0^*$ for every i and k. Therefore, since $F \times C = \bigcup_{i=1}^{n} \bigcup_{k=1}^{k(i)} F_i \times D_k^i$ and the D_k^i are either disjoint or equal, $\mu_X(F \times C) = \sum_{i=1}^{n} \sum_{k=1}^{k(i)} \mu_X(F_i \times D_k^i) = \sum_{i=1}^{n} \mu_X(F_i \times C_i)$ when we apply the second and first cases respectively. This completes the proof of finite additivity of $\mu_{(M)^2}$ on $\mathcal{P}_0^*$. $\square$

Proof of Theorem 4.1. Since the proof is close to that of Theorem 5.1 in Dozzi et. al. (1994), we provide only a sketch and refer the reader to the said paper for details. We begin by constructing a process Q is such a way that (12) is satisfied. Take $m \in \mathbf{N}$ and select $A \in \mathcal{A}_m$. Given any $F \in \mathcal{F}$, if we define $\sigma_A(F) = \int_{\Omega \times A} Y(F)\, d\mu_{(M)^2}$ where Y is the T-indexed process in Assumption 3.2, it can be shown that

$$\sigma_A(F) \;=\; \lim_{\substack{n \to \infty \\ n \ge m}} E[\mathbf{1}_F\, Q_A^{(n)}]\,.$$

Therefore, by the Hahn-Vitali-Saks Theorem, there exists a random variable $\overline{Q}_A \in L_1$ such that $\sigma_A(F) = E[\mathbf{1}_F\, \overline{Q}_A]\ \forall\, F \in \mathcal{F}$, i.e., $(Q_A^{(n)})_{n \ge m}$ converges to $\overline{Q}_A$ in the weak L_1 topology.

Applying the above argument to each $A \in \mathcal{A}^* = \bigcup_m \mathcal{A}_m$, we obtain a collection $\{\overline{Q}_A : A \in \mathcal{A}^*\}$ of random variables which generates an $\mathcal{A}$-indexed process Q when we set

$$Q_A(\omega) \;=\; \inf_{\substack{A' \in \mathcal{A}^* \\ A \subseteq A'}} \overline{Q}_{A'}(\omega) \;=\; \lim_m \overline{Q}_{g_m(A)}(\omega)$$

for any $A \in \mathcal{A}$ and $\omega \in \Omega$. Furthermore, since the above limit is monotone in m, Q is monontone outer-continuous and convergence is in the weak L_1 topology. Following Dozzi et. al. (1994), since M is monotone L_2-outer-continuous (cf. Lemma 2.9) it is straightforward to show that $Q_C \ge 0$ $\forall\, C \in \mathcal{C}$. That is, $Q = (Q_A)_{A \in \mathcal{A}}$ is raw increasing.

To establish *-predictablility for Q, note that given $C = A \setminus \bigcup_{i=1}^{k} A_i$, if for each m we define $C^m = g_m(A) \setminus \bigcup_{i=1}^{k} g_m(A_i)$, then given any $F \in \mathcal{F}$, dominated convergence and the monotone outer-continuity of Q imply

$$\begin{aligned} E[\mathbf{1}_F\, Q_C] &= \lim_m E[\mathbf{1}_F\, \overline{Q}_{C^m}] \\ &= \lim_m \int_{\Omega \times C^m} Y(F)\, d\mu_{(M)^2} \\ &= \int_{\Omega \times C} Y(F)\, d\mu_{(M)^2}\,. \end{aligned}$$

Therefore, it is sufficient to show $\mu_{(M)^2} = \mu_Q$ on $\mathcal{P}^*$ which is in fact equivalent to showing

$$\mu_{(M)^2}(F \times C) \;=\; \mu_Q(F \times C) \quad \forall\, F \times C \in \mathcal{P}_0^* \tag{13}$$

since $\mathcal{P}_0^*$ is a π-system which generates $\mathcal{P}^*$. (13) clearly implies (11).

Fix $F \times C \in \mathcal{P}_0^*$. If $C = A \setminus \bigcup_{i=1}^{k} B_i$ is a maximal representation of C, then given any $B \in \mathcal{A}(u)$ such that $B \cap C = \phi$, $B \subseteq \bigcup_{i=1}^{k} B_i \subseteq \bigcup_{i=1}^{k} g_m(B_i)$ $\forall m$ so that $B \cap C^m = \phi$ $\forall m$. Therefore,

$F \in \mathcal{G}^*_C \subseteq \mathcal{G}^*_{C^m}$ $\forall m$. Since M is monotone L_2-outer-continuous, this implies

$$\begin{aligned} E[\mathbf{1}_F\,(M_C)^2] &= \lim_{m\to\infty} E[\mathbf{1}_F\,(M_{C^m})^2] \;=\; \lim_{m\to\infty} \lim_{\substack{n\,\to\,\infty\\ n\,\geq\, m}} E[\mathbf{1}_F\,Q^{(n)}_{C^m}] \\ &= \lim_{m\to\infty} E[\mathbf{1}_F\,Q_{C^m}] \;=\; E[\mathbf{1}_F\,Q_C] \end{aligned}$$

which establishes (13).

The argument for uniqueness of Q is identical to that found in the proof of Theorem 4.1 in Dozzi et. al. (1994). □

Theorem 4.1 motivates the following terminology.

Definition 4.3 *Given a strong martingale $M = (M_A)_{A\in\mathcal{A}}$ in L_2, a process $Q = (Q_A)_{A\in\mathcal{A}}$ is termed a *-predictable quadratic variation of M if Q is *-predictable and $E[(M_C)^2|\mathcal{G}^*_C] = E[Q_C|\mathcal{G}^*_C]$ $\forall\, C \in \mathcal{C}$.*

Comments. (1) By a straightforward argument, the conditioning relation in Definition 4.3 extends to all $C \in \mathcal{C}(u)$.

(2) As a consequence of *-predictablility, a *-predictable quadratic variation is unique if it exists. Once again, the proof is identical to that found in Theorem 4.1 of Dozzi et. al. (1994).

We close the section with a simple moment condition under which a strong martingale possesses an adapted *-predictable quadratic variation.

Proposition 4.4 *Under* (32) *and Assumptions 2.1, 3.2, A.5, A.6 and A.7, if M is a strong martingale in $L_{2+\delta}$ for some $\delta > 0$, then M possesses a unique *-predictable quadratic variation Q which is adapted to $(\mathcal{F}_A)$.*

Proof. By Lemma A.11, the $L_{1+\delta/2}$-norm of any element in (10) is bounded above by $(\kappa \cdot E[|M_T|^{2+\delta}])^{2/(2+\delta)} < \infty$, κ some positive constant depending only on δ. Therefore, by Theorem 4.1, M possesses a *-predictable quadratic variation Q. Adaptedness of Q follows by the argument for adaptedness of V in Proposition 3.6. □

5 L_2-Approximations of Quadratic Variation

Let $M = (M_A)_{A\in\mathcal{A}}$ be a strong martingale in L_4 so that under the conditions in Proposition 4.4, M possesses a unique *-predictable quadratic variation $Q = (Q_A)_{A\in\mathcal{A}}$. In this section, assuming that the underlying compact space T is metrizable, it will be shown that the quadratic variation Q is "calculable" whenever M has sample paths which are continuous in a certain sense. That is, Q_A can be approximated in L_2-norm by discrete sums of the form

$$Q^{(n)}_{g_m(A)} \;=\; \sum_{\substack{D\,\in\,\mathcal{N}_n\\ D\,\subseteq\, B}} E[\,(M_D)^2\,|\,\mathcal{G}^*_D\,]\,.$$

(This improves the weak L_1-approximations in Theorem 4.1.) We begin with a useful L_2-norm inequality.

Proposition 5.1 *Under (32) and Assumptions 2.1, 3.2, A.5, A.6 and A.7, if M is a strong martingale in L_4, then for any $k \in \mathbf{N}$ and $A \in \mathcal{A}_k$,*

$$\left\| Q_A^{(n)} - Q_A^{(m)} \right\|_2^2 \;\le\; C_1 \cdot \alpha_n \;+\; C_2 \cdot \alpha_m \tag{14}$$

for every $n, m \ge k$ where, for each $r \in \mathbf{N}$,

$$\alpha_r \;:=\; E\left[\max_{C \in \mathcal{N}_r} (M_C)^4 \right] \tag{15}$$

and C_1, C_2 are positive constants depending only on $E[M_T^4]$.

(**??**),

Proof. Take $k, m, n \in \mathbf{N}$ such that $m, n \ge k$. Without loss of generality, we may assume $n < m$ so that, given any $C \in \mathcal{N}_n$, the inclusion $\mathcal{A}_n \subseteq \mathcal{A}_m$ implies

$$E\left[(M_C)^2 \,\middle|\, \mathcal{G}_C^* \right] \;=\; \sum_{\substack{D_1 \in \mathcal{N}_m \\ D_1 \subseteq C}} \; \sum_{\substack{D_2 \in \mathcal{N}_m \\ D_2 \subseteq C}} E[\, M_{D_1} M_{D_2} \,|\, \mathcal{G}_C^* \,]\,. \tag{16}$$

Furthermore, by Lemma 2.8(a), given any $D_1 \ne D_2$ in $\mathcal{N}_m$, we may assume without loss of generality that M_{D_1} is $\mathcal{G}_{D_2}^*$-measurable. Therefore, since $\mathcal{G}_C^* \subseteq \mathcal{G}_D^*$ for all $D \in \mathcal{N}_m$ with $D \subseteq C$, a simple conditioning argument applied to (16) yields

$$E\left[(M_C)^2 \,\middle|\, \mathcal{G}_C^* \right] \;=\; \sum_{\substack{D \in \mathcal{N}_m \\ D \subseteq C}} E[\,(M_D)^2 \,|\, \mathcal{G}_C^* \,] \tag{17}$$

by the strong martingale property of M.

Now, select $A \in \mathcal{A}_k$. Applying (17) to each $C \in \mathcal{N}_n$ for which $C \subseteq A$,

$$Q_A^{(n)} - Q_A^{(m)} \;=\; \sum_{\substack{C \in \mathcal{N}_n \\ C \subseteq A}} \; \sum_{\substack{D \in \mathcal{N}_m \\ D \subseteq C}} E[\,(M_D)^2 \,|\, \mathcal{G}_C^*] \;-\; \sum_{\substack{D \in \mathcal{N}_m \\ D \subseteq A}} E[\,(M_D)^2 \,|\, \mathcal{G}_D^*]\,. \tag{18}$$

For each $D \in \mathcal{N}_m$ and $C \in \mathcal{N}_n$, define

$$d_D^{(C)} \;=\; E[\,(M_D)^2 \,|\, \mathcal{G}_C^*] \;-\; E[\,(M_D)^2 \,|\, \mathcal{G}_D^*]\,. \tag{19}$$

Since $\mathcal{A}_n \subseteq \mathcal{A}_m$, the definition of left-neighborhood implies

$$\bigcup_{\substack{D \in \mathcal{N}_m \\ D \subseteq A}} D \;=\; \bigcup_{\substack{C \in \mathcal{N}_n \\ C \subseteq A}} \; \bigcup_{\substack{D \in \mathcal{N}_m \\ D \subseteq C}} D$$

so that (18) becomes

$$Q_A^{(n)} - Q_A^{(m)} \;=\; \sum_{\substack{C \in \mathcal{N}_n \\ C \subseteq A}} \; \sum_{\substack{D \in \mathcal{N}_m \\ D \subseteq C}} d_D^{(C)}.$$

Therefore,

$$\left(Q_A^{(n)} - Q_A^{(m)}\right)^2 = \sum_{\substack{C \in \mathcal{N}_n \\ C \subseteq A}} \sum_{\substack{D \in \mathcal{N}_m \\ D \subseteq C}} \left(d_D^{(C)}\right)^2 + \sum_{\substack{C \in \mathcal{N}_n \\ C \subseteq A}} \sum_{\substack{D_1, D_2 \in \mathcal{N}_m \\ D_1, D_2 \subseteq C \\ D_1 \neq D_2}} d_{D_1}^{(C)} \cdot d_{D_2}^{(C)} + \sum_{\substack{C_1, C_2 \in \mathcal{N}_n \\ C_1, C_2 \subseteq A \\ C_1 \neq C_2}} \sum_{\substack{D_1, D_2 \in \mathcal{N}_m \\ D_i \subseteq C_i}} d_{D_1}^{(C_1)} \cdot d_{D_2}^{(C_2)}. \tag{20}$$

Employing this expansion, we will obtain the upper bound in (14) in three steps. In the first step, we dispose of the last sum in (20).

Step 1: *Given $C_1, C_2 \in \mathcal{N}_n$ such that $C_1 \neq C_2$ and $D_1, D_2 \in \mathcal{N}_m$ for which $D_i \subseteq C_i$ $(i = 1, 2)$, $E[\, d_{D_1}^{(C_1)} \cdot d_{D_2}^{(C_2)}\,] = 0$.*

Select such sets C_1, C_2, D_1 and D_2. By (19),

$$\begin{aligned} d_{D_1}^{(C_1)} \cdot d_{D_2}^{(C_2)} &= E[\,(M_{D_1})^2 \,|\, \mathcal{G}^*_{C_1}] \cdot E[\,(M_{D_2})^2 \,|\, \mathcal{G}^*_{C_2}] \\ &\quad - E[\,(M_{D_1})^2 \,|\, \mathcal{G}^*_{C_1}] \cdot E[\,(M_{D_2})^2 \,|\, \mathcal{G}^*_{D_2}] \\ &\quad - E[\,(M_{D_1})^2 \,|\, \mathcal{G}^*_{D_1}] \cdot E[\,(M_{D_2})^2 \,|\, \mathcal{G}^*_{C_2}] \\ &\quad + E[\,(M_{D_1})^2 \,|\, \mathcal{G}^*_{D_1}] \cdot E[\,(M_{D_2})^2 \,|\, \mathcal{G}^*_{D_2}]. \end{aligned} \tag{21}$$

Assume C_i is the left-neighborhood generated by $A_i \in \mathcal{A}_n$ $(i = 1, 2)$ and likewise, assume D_i is the left-neighborhood generated by $B_i \in \mathcal{A}_m$ $(i = 1, 2)$. By Lemma 2.3, we can assume without loss of generality that $\mathcal{F}_{A_1} \subseteq \mathcal{G}^*_{C_2}$. Moreover, $D_1 \subseteq C_1$ implies $B_1 \subseteq A_1$. (Otherwise $D_1 = D_1 \cap C_1 = \phi$, a contradiction.) Therefore, since $(\mathcal{F}_A)_{A \in \mathcal{A}}$ is increasing and $(\mathcal{G}^*_C)_{C \in \mathcal{C}}$ is decreasing,

$$\mathcal{F}_{B_1} \subseteq \mathcal{F}_{A_1} \subseteq \mathcal{G}^*_{C_2} \subseteq \mathcal{G}^*_{D_2}. \tag{22}$$

Furthermore, since D_1 is the left-neighborhood of B_1, (4) implies $(M_{D_1})^2$ is $\mathcal{F}_{B_1}$-measurable. Therefore, since $\mathcal{F}_{B_1} \subseteq \mathcal{F}_{A_1}$, Corollary A.10 implies

$$E[\,(M_{D_1})^2 \,|\, \mathcal{G}^*_{C_1}] \text{ is } \mathcal{F}_{A_1}\text{-measurable.} \tag{23}$$

In general, given sub-σ-algebras $\mathcal{H}_1$ and $\mathcal{H}_2$ of $\mathcal{F}$ and random variables $X, Y \in L_2$ such that X is $\mathcal{H}_i$-measurable $(i = 1, 2)$, conditioning yields $E[\, X \cdot E(Y \,|\, \mathcal{H}_1)\,] = E[\, X \cdot E(Y \,|\, \mathcal{H}_2)\,]$. Therefore, if we take $\mathcal{H}_1 = \mathcal{G}^*_{C_2}$, $\mathcal{H}_2 = \mathcal{G}^*_{D_2}$, $X = E[\,(M_{D_1})^2 \,|\, \mathcal{G}^*_{C_1}]$ and $Y = (M_{D_2})^2$, then (22) and (23) imply

$$E\left\{ E[\,(M_{D_1})^2 \,|\, \mathcal{G}^*_{C_1}] \cdot E[\,(M_{D_2})^2 \,|\, \mathcal{G}^*_{C_2}] \right\} = E\left\{ E[\,(M_{D_1})^2 \,|\, \mathcal{G}^*_{C_1}] \cdot E[\,(M_{D_2})^2 \,|\, \mathcal{G}^*_{D_2}\,] \right\}$$

and by an identical argument,

$$E\left\{ E[\,(M_{D_1})^2 \,|\, \mathcal{G}^*_{D_1}] \cdot E[\,(M_{D_2})^2 \,|\, \mathcal{G}^*_{C_2}] \right\} = E\left\{ E[\,(M_{D_1})^2 \,|\, \mathcal{G}^*_{D_1}] \cdot E[\,(M_{D_2})^2 \,|\, \mathcal{G}^*_{D_2}\,] \right\}.$$

In view of (21), this completes the proof of Step 1.
establishing Step 1.

Step 2: *There exists a positive constant* C_2 *depending only on* $E[M_T^4]$ *such that*

$$\sum_{\substack{C \in \mathcal{N}_n \\ C \subseteq A}} \sum_{\substack{D_1, D_2 \in \mathcal{N}_m \\ D_1, D_2 \subseteq C \\ D_1 \neq D_2}} E\left[d_{D_1}^{(C)} \cdot d_{D_2}^{(C)} \right] \;\leq\; C_2 \cdot \alpha_n \,.$$

Fix a set $C \in \mathcal{N}_n$ for which $C \subseteq A$ and let $D_1, D_2 \in \mathcal{N}_m$ be such that $D_1, D_2 \subseteq C$ and $D_1 \neq D_2$. Since $\mathcal{G}_C^* \subseteq \mathcal{G}_{D_i}^*$ $(i = 1, 2)$, we obtain

$$\begin{aligned} & E\Big\{ E[\,(M_{D_1})^2 \,|\, \mathcal{G}_C^*] \cdot E[\,(M_{D_2})^2 \,|\, \mathcal{G}_{D_2}^*] \Big\} \\ & \qquad\qquad = E\Big\{ E[\,(M_{D_1})^2 \,|\, \mathcal{G}_C^*] \cdot E[\,(M_{D_2})^2 \,|\, \mathcal{G}_C^*\,] \Big\} \end{aligned}$$

when we condition with respect to $\mathcal{G}_C^*$ and then apply the tower property.

Now, consider (21) with $C_1 = C_2 = C$. Since all four products in (21) are non-negative, the previous identity implies

$$E\left[d_{D_1}^{(C)} \cdot d_{D_2}^{(C)} \right] \;\leq\; E\Big\{ E[\,(M_{D_1})^2 \,|\, \mathcal{G}_{D_1}^*] \cdot E[\,(M_{D_2})^2 \,|\, \mathcal{G}_{D_2}^*\,]\Big\} \,.$$

Therefore,

$$\begin{aligned} \sum_{\substack{D_1, D_2 \in \mathcal{N}_m \\ D_1, D_2 \subseteq C \\ D_1 \neq D_2}} E\left[d_{D_1}^{(C)} \cdot d_{D_2}^{(C)} \right] \;&\leq\; E\left[\left(\sum_{\substack{D \in \mathcal{N}_m \\ D \subseteq C}} E\left[\,(M_D)^2 \,\middle|\, \mathcal{G}_D^* \right] \right)^2 \right] \\ &\leq\; \kappa \cdot E\Big\{ \Big[M\big(\textstyle\bigcup_{D \in \mathcal{N}_m,\, D \subseteq C} D\big) \Big]^4 \Big\} \\ &=\; \kappa \cdot E\Big[(M_C)^4 \Big] \end{aligned}$$

where κ is the universal constant in Lemma A.11. Furthermore, by Hölder's inequality,

$$\begin{aligned} E\Big\{ \sum (M_C)^4 \Big\} \;&=\; E\Big\{ \Big[\sup\,(M_C)^2 \Big] \cdot \Big[\sum (M_C)^2 \Big] \Big\} \\ &\leq\; \sqrt{E\,[\sup\,(M_C)^4]} \cdot \sqrt{E\Big\{ \Big[\sum (M_C)^2 \Big]^2 \Big\}} \end{aligned} \tag{24}$$

where all sums and supremums range over the sets $C \in \mathcal{N}_n$. Since Lemma 2.11 implies

$$E\Big\{ \Big[\textstyle\sum_{C \in \mathcal{N}_n} (M_C)^2 \Big]^2 \Big\} \;\leq\; \kappa' \cdot \sqrt{E[M_T^4]} \tag{25}$$

for some universal constant $\kappa' > 0$, Step 2 is completed by taking $C_2 = \kappa \cdot \kappa' \cdot \sqrt{E[M_T^4]}$.

Step 3: *There exists a positive constant C_1 depending only on $E[M_T^4]$ such that*

$$\sum_{\substack{C \in \mathcal{N}_n \\ C \subseteq A}} \sum_{\substack{D \in \mathcal{N}_m \\ D \subseteq C}} E\left[\left(d_D^{(C)}\right)^2\right] \;\leq\; C_1 \cdot \alpha_m.$$

Given any $C \in \mathcal{N}_n$ with $C \subseteq A$ and any $D \in \mathcal{N}_n$ with $D \subseteq C$,

$$\begin{aligned} E\left[\left(d_D^{(C)}\right)^2\right] &= E\left\{\left(E[\,(M_D)^2 \,|\, \mathcal{G}_C^*] - E[\,(M_D)^2 \,|\, \mathcal{G}_D^*]\,\right)^2\right\} \\ &\leq E\left(E[\,(M_D)^2 \,|\, \mathcal{G}_C^*]^2\right) + E\left(E[\,(M_D)^2 \,|\, \mathcal{G}_D^*]^2\right) \\ &\leq 2 \cdot E\left[(M_D)^4\right] \quad \text{(by Jensen's inequality).} \end{aligned}$$

(The second to last line follows from the deterministic inequality $(a-b)^2 \leq a^2 + b^2$ where $a, b \geq 0$.) Therefore,

$$\sum_{\substack{C \in \mathcal{N}_n \\ C \subseteq A}} \sum_{\substack{D \in \mathcal{N}_m \\ D \subseteq C}} E\left[\left(d_D^{(C)}\right)^2\right] \;\leq\; 2 \cdot \sum_{D \in \mathcal{N}_m} E\left[(M_D)^4\right] \tag{26}$$

so that Step 3, and hence the proof of Proposition 5.1, follows from (24) and (25) with $\mathcal{N}_m$ in place of $\mathcal{N}_n$. $\square$

The following assumption on $\mathcal{A}$ will be required in the sequel.

Assumption 5.2 There exists a universal constant K such that, given any $n \in \mathbf{N}$ and $A \in \mathcal{A}_n$, there is a subcollection $\{A_1, \cdots, A_k\} \subseteq \mathcal{A}_n$ with $k \leq K$ such that

(i) $\bigcup_{i=1}^k A_i = \bigcup\{A' \in \mathcal{A}_n : A' \subset A\}$,

(ii) $i \neq j$ implies $A_i \not\subseteq A_j$ and

(iii) if $A' \in \mathcal{A}_n$ is such that $A' \subset A$, then there exists $1 \leq j \leq k$ such that $A' \subseteq A_j$.

(Clearly, the set $A \setminus \bigcup_{i=1}^k A_i \in \mathcal{C}_n$ is a minimal representation of the left-neighborhood generated by A in $\mathcal{A}_n$.)

Examples satisfying Assumption 5.2 include the lower rectangles and the lower layers in $[0,1]^d$. In both, we can take $K = d$.

Lemma 5.3 *Define $\mathcal{A}^* = \bigcup_n \mathcal{A}_n$ and let M be an $\mathcal{A}$-indexed strong martingale for which* $\sup_{A \in \mathcal{A}^*} M_A \in L^4$. *Under Assumption 5.2 and those in Proposition 5.1, if* $\max_{C \in \mathcal{N}_n} |M_C| \to_P 0$ *as $n \to \infty$, then the sequence (α_n) defined in* (15) *converges to zero.*

Proof. By the dominated convergence theorem, it is sufficient to show that, for each n, $\max_{C \in \mathcal{N}_n} |M_C| \leq \beta \cdot (\sup_{A \in \mathcal{A}^*} M_A^4)$ for some universal constant $\beta > 0$. To this end, fix $n \in \mathbf{N}$. Given $C \in \mathcal{N}_n$, there exists $A \in \mathcal{A}_n$ such that C is the left-neighborhood generated by A in $\mathcal{A}_n$. If $A \setminus \bigcup_{i=1}^k A_i$ is the minimal representation of C described in Assumption 5.2 then by (4),

$$|M_C| \;\leq\; \sum_{I \subseteq \{1, \cdots, k\}} |M(A \cap \textstyle\bigcap_{i \in I} A_i)| \;\leq\; 2^K \cdot (\sup_{B \in \mathcal{A}^*} M_B\,)$$

so that the lemma follows by taking $\beta = 2^{4K}$. $\square$

Combining Lemma 5.3 and Proposition 5.1,

Corollary 5.4 *If M is as described in Lemma 5.3, then under* (32) *and Assumptions 2.1, 3.2, 5.2, A.5, A.6 and A.7,* $\{Q^{(n)}_{g_m(A)} : n \geq m\}$ *is Cauchy in L_2 for any fixed $m \in \mathbf{N}$ and $A \in \mathcal{A}$.*

Now, assume T is metrized by a metric d. As done in various papers (for example, cf. Ivanoff and Merzbach (1996)), we define the *Hausdorff metric* between any two closed sets, $A, B \in \mathcal{A}$ by $d_H(A, B) = \inf\{\epsilon > 0 : A \subseteq B^\epsilon \text{ and } B \subseteq A^\epsilon\}$ where $A^\epsilon = \{x \in T : d(x, A) < \epsilon\}$. A process $X = (X_A)_{A \in \mathcal{A}}$ will be termed *sample path continuous* if, on some event of full P-measure, all sample paths $A \mapsto X_A$ are d_H-continuous on $\mathcal{A}$.

In order to apply Corollary 5.4 to obtain L_2-norm approximations of the *-predictable quadratic variation of a sample path continuous strong martingale, it is necessary that $\max_{C \in \mathcal{N}_n} |X_C| \rightarrow_P 0$ for sample path continuous processes X. For this, we need the following assumption which we conjecture is a consequence of the compactness of T. In any case, it is satisfied by all known examples.

Assumption 5.5 *(Uniform Approximations.)* There exists a sequence (ϵ_n) of positive constants with $\epsilon_n \downarrow 0$ such that, given any n, $d_H(A, g_n(A)) \leq \epsilon_n$ for every $A \in \mathcal{A}$.

Comment. It has been shown in Slonowsky (1998) that any class $\mathcal{A}$ satisfying Assumptions 2.1 and 5.5 is compact with respect to the Hausdorff metric.

Under Assumptions 5.2 and 5.5,

Lemma 5.6 *Given n and $A \in \mathcal{A}_n$, if $\{A_1, \cdots, A_k\}$ is a subcollection of $\mathcal{A}_n$ satisfying the conditions of Assumption 5.2, then $d_H(\bigcap_{i \in I} A_i, \bigcap_{j \in J} A_j) \leq 2\,\epsilon_n$ for any $I, J \subseteq \{1, \cdots, k\}$ where (ϵ_n) is as defined in Assumption 5.5 and, by convention, we define $\bigcap_{i \in \phi} A_i = A$.*

Proof. First, note that $A \subseteq g_n(A_i)$ for every $1 \leq i \leq k$. Indeed, if $A \not\subseteq g_n(A_{i_0})$ for some $1 \leq i_0 \leq k$, then $g_n(A_{i_0}) \cap A \subset A$. In addition, by (A3) of Assumption 2.1, $A_{i_0} \subset A$ implies $A_{i_0} \subset g_n(A_{i_0}) \cap A$. If we define $A' = g_n(A_{i_0}) \cap A \in \mathcal{A}_n$, then, since $A' \subset A$, condition (iii) of Assumption 5.2 implies $A' \subseteq A_{j_0}$ for some $1 \leq j_0 \leq k$ so that $A_{i_0} \subset A_{j_0}$, contradicting condition (i) of Assumption 5.2.

Next, take $I \subseteq \{1, \cdots, k\}$. Since g_n preserves finite intersections,

$$A \;\subseteq\; g_n\left(\textstyle\bigcap_{i \in I} A_i\right) \;\subseteq\; \left(\textstyle\bigcap_{i \in I} A_i\right)^{\epsilon_n}$$

when we apply the above observation and Assumption 5.5 in that order. Since $\bigcap_{i \in I} A_i \subseteq A$, this implies $d_H(A, \bigcap_{i \in I} A_i) \leq \epsilon_n$ so that the present lemma follows by the triangle inequality. $\square$

Lemma 5.7 *If $X = (X_A)_{A \in \mathcal{A}}$ is a sample path continuous process, then under Assumptions 5.2 and 5.5,* $\max_{C \in \mathcal{N}_n} |X_C| \rightarrow_P 0$.

Proof. Given a d_H-continuous set-function $x : \mathcal{A} \to \mathbf{R}$, define the *modulus of continuity*, $w(x, \epsilon)$ of x ($\epsilon > 0$) by $w(x, \epsilon) = \sup\{|x(A) - x(B)| : d_H(A, B) \leq \epsilon,\ A, B \in \mathcal{A}\}$. Since $(\mathcal{A}, d_H)$ is compact, x is uniformly d_H-continuous on $\mathcal{A}$ so that $\lim_{\epsilon \to 0} w(x, \epsilon) = 0$. In particular, $\lim_{\epsilon \to 0} w(X, \epsilon) = 0$ a.e.

Now, fix $n \in \mathbf{N}$ and select $C \in \mathcal{N}_n$, the left-neighborhood of a set $A \in \mathcal{A}_n$. Let $A \setminus \bigcup_{i=1}^k A_i$ be the minimal representation of C described in Assumption 5.2. If we define

$$\mathcal{E} \ = \ \{I \subseteq \{1, \cdots k\} : |I| \text{ is even}\} \quad \text{and} \quad \mathcal{O} \ = \ \{I \subseteq \{1, \cdots k\} : |I| \text{ is odd}\},$$

then by the binomial theorem, there is a bijection $f : \mathcal{E} \to \mathcal{O}$. Clearly, $|\mathcal{E}| \leq 2^K$ where K is the universal constant defined in Assumption 5.2. Therefore, by Lemma 5.6,

$$|X_C| \ \leq \ \sum_{I \in \mathcal{E}} \Big| X(A \cap \textstyle\bigcap_{i \in I} A_i) \ - \ X(A \cap \bigcap_{j \in f(I)} A_j) \Big| \ \leq \ 2^K \cdot w\,(X,\, 2\,\epsilon_n)\ .$$

Taking the maximum over all $C \in \mathcal{N}_n$ and then letting $n \to \infty$, the proof is complete. □

The main result of the section is the following:

Theorem 5.8 *Let M be a strong martingale for which $\sup_{A \in \mathcal{A}^*} M_A^4 \in L_4$. Under (32) and Assumptions 2.1, 3.2, 5.2, 5.5, A.5, A.6 and A.7, if M is sample path continuous, then for every $A \in \mathcal{A}$,*

$$Q_A \ = \ \lim_{m \to \infty} \lim_{n \to \infty,\, n \geq m} Q^{(n)}_{g_m(A)} \ \text{ in } L_2 \tag{27}$$

*where Q is the unique *-predictable quadratic variation of M and $Q^{(n)}_{g_m(A)} = \sum_{D \in \mathcal{N}_n,\, D \subseteq g_m(A)} E[(M_D)^2 | \mathcal{G}^*_D]$.*

Proof. Fix $A \in \mathcal{A}$. By Proposition 4.4 and Theorem 4.1,

$$\lim_m \overline{Q}_{g_m(A)} \ = \ Q_A \ \text{ in the weak } L_1 \text{ topology} \tag{28}$$

where for every $m \in \mathbf{N}$,

$$\overline{Q}_{g_m(A)} \ = \ \lim_n Q^{(n)}_{g_m(A)} \ \text{ in the weak } L_1 \text{ topology}\,. \tag{29}$$

By the completeness of L_2 and Corollary 5.4, the convergence in (29) is also in L_2-norm. To establish L_2-norm convergence in (28), it is sufficient to show $(\overline{Q}_{g_m(A)})_m$ is Cauchy in L_2.

Toward this goal, note that

$$\left\| Q^{(n)}_{g_k(A)} - Q^{(n)}_{g_m(A)} \right\|_2 \ \leq \ \gamma \cdot \left\| M_{g_k(A)} - M_{g_m(A)} \right\|_4^2 \tag{30}$$

for any $k, m, n \in \mathbf{N}$ with $k, m \leq n$ where γ is a positive constant depending only on $E\,[M_T^4]$. Indeed, assuming $k \leq m$, since $g_k(A), g_m(A) \in \mathcal{A}_n$ and $g_m(A) \subseteq g_k(A)$,

$$\begin{aligned} Q^{(n)}_{g_k(A)} - Q^{(n)}_{g_m(A)} \ &= \ \sum_{\substack{C \,\in\, \mathcal{N}_n \\ C \,\subseteq\, g_k(A)}} E\Big[\,(M_C)^2\,|\,\mathcal{G}^*_C\Big] \ - \ \sum_{\substack{C \,\in\, \mathcal{N}_n \\ C \,\subseteq\, g_m(A)}} E\Big[\,(M_C)^2\,|\,\mathcal{G}^*_C\Big] \\ &= \ \sum_{\substack{C \,\in\, \mathcal{N}_n \\ C \,\subseteq\, g_k(A) \setminus g_m(A)}} E\Big[\,(M_C)^2\,|\,\mathcal{G}^*_C\Big] \end{aligned}$$

so that Lemma A.11 implies the existence of a positive constant κ depending only on $E\left[M_T^4\right]$ such that

$$\begin{aligned}\left\|Q^{(n)}_{g_k(A)} - Q^{(n)}_{g_m(A)}\right\|_2^2 &\le \kappa \cdot \left\|M\left(\bigcup\{C \in \mathcal{C}_n : C \subseteq g_k(A) \setminus g_m(A)\}\right)\right\|_4^4 \\ &= \kappa \cdot \left\|M_{g_k(A)\setminus g_m(A)}\right\|_4^4 .\end{aligned}$$

Now, to show that $(\overline{Q}_{g_m(A)})_m$ is Cauchy in L_2, fix $\epsilon > 0$. Given any $k, m, n \in \mathbf{N}$,

$$\begin{aligned}\left\|\overline{Q}_{g_k(A)} - \overline{Q}_{g_m(A)}\right\|_2 \le\ & \left\|\overline{Q}_{g_k(A)} - Q^{(n)}_{g_k(A)}\right\|_2 + \left\|Q^{(n)}_{g_k(A)} - Q^{(n)}_{g_m(A)}\right\|_2 + \\ & \left\|Q^{(n)}_{g_m(A)} - \overline{Q}_{g_m(A)}\right\|_2 .\end{aligned} \tag{31}$$

Since M is sample path continuous, $M_{g_m(A)} \to M_A$ in L_4 by dominated convergence. Therefore, by (30), there is an L such that $\|Q^{(n)}_{g_k(A)} - Q^{(n)}_{g_m(A)}\|_2 \le \epsilon/3$ whenever $n \ge k, m$ and $k, m \ge L$. Moreover, given any pair $k, m \ge L$, since the convergence in (29) was shown to be in L_2-norm, both of the remaining summands in (31) can be made less than $\epsilon/3$ by chosing such an n sufficiently large. This completes the proof. □

A Consequences of Conditional Independence

In this section, under the conditional independence assumption:

$$E[E(X|\mathcal{F}_A)|\mathcal{F}_B] = E[X|\mathcal{F}_{A\cap B}] \quad \forall\, A, B \in \mathcal{A},\ X \in L_1 \tag{32}$$

we will establish, among other things, a set-indexed generalization of Rosenthal's square function inequality. (32) has appeared earlier in Dozzi et. al. (1994) and reduces to the classic F4 property of Cairoli and Walsh (1975) when $\mathcal{A}$ is the collection of lower rectangles in $[0,1]^2$. In general, any filtration generated by a process $X = (X_A)_{A\in\mathcal{A}}$ with independent increments (e.g., an $\mathcal{A}$-indexed Gaussian process) will satisfy (32). Under (32),

$$\bigcap_{i=1}^n \mathcal{F}_{A_i} = \mathcal{F}_{\bigcap_{i=1}^n A_i} \tag{33}$$

for any $A_1, \cdots, A_n \in \mathcal{A}$ (cf. Dozzi et. al. (1994), Lemma 2.3).

Another form of conditional independence appearing in the literature states that sub-σ-algebras $\mathcal{F}_1$ and $\mathcal{F}_2$ of $\mathcal{F}$ are *conditionally independent* of a third sub-σ-algebra $\mathcal{F}_3$ if, given $\mathcal{F}_i$-measurable random variables X_i, $i = 1, 2$,

$$E[X_1 X_2 \,|\, \mathcal{F}_3] = E[X_1 \,|\, \mathcal{F}_3] \cdot E[X_2 \,|\, \mathcal{F}_3] . \tag{34}$$

In such a case, we write "$\mathcal{F}_1 \perp \mathcal{F}_2$ given $\mathcal{F}_3$". Clearly, it is sufficient to test (34) for the case in which $X_i = \mathbf{1}_{F_i}$, $F_i \in \mathcal{F}_i$ $(i = 1, 2)$. The relation between (32) and (34) is as follows.

Lemma A.1 *Under* (32)*, for any* $A_1, A_2 \in \mathcal{A}$*,* $\mathcal{F}_{A_1} \perp \mathcal{F}_{A_2}$ *given* $\mathcal{F}_{A_1\cap A_2}$*.*

Proof. Taking $X_i = \mathbf{1}_{F_i}$, $F_i \in \mathcal{F}_{A_i}$ $(i = 1, 2)$, (32) implies

$$E[X_1 X_2 \,|\, \mathcal{F}_{A_1 \cap A_2}] \;=\; E[\, E(X_1 X_2 \,|\, \mathcal{F}_{A_1}) \,|\, \mathcal{F}_{A_2}] \;=\; E[\, X_1 \, E(X_2 \,|\, \mathcal{F}_{A_1}) \,|\, \mathcal{F}_{A_2}].$$

Furthermore, since $F_i \in \mathcal{F}_{A_i}$, (32) implies $E[X_1 \,|\, \mathcal{F}_{A_2}] = E[X_1 \,|\, \mathcal{F}_{A_1 \cap A_2}]$ and $E[X_2 \,|\, \mathcal{F}_{A_1}] = E[X_2 \,|\, \mathcal{F}_{A_1 \cap A_2}]$ so that

$$\begin{aligned} E[\, X_1 \, E(X_2 \,|\, \mathcal{F}_{A_1}) \,|\, \mathcal{F}_{A_2}] &= E[X_2 \,|\, \mathcal{F}_{A_1 \cap A_2}] \cdot E[X_1 \,|\, \mathcal{F}_{A_2}] \\ &= E[X_2 \,|\, \mathcal{F}_{A_1 \cap A_2}] \cdot E[X_1 \,|\, \mathcal{F}_{A_1 \cap A_2}] \end{aligned}$$

which completes the proof. □

The following characterization of conditional independence can be found on p.36-II of Dellacherie Meyer (1978).

Lemma A.2 *If $\mathcal{F}_1$, $\mathcal{F}_2$ and $\mathcal{F}_3$ are sub-σ-algebras of $\mathcal{F}$, then $\mathcal{F}_1 \perp \mathcal{F}_2$ given $\mathcal{F}_3$ if and only if $E[X \,|\, \mathcal{F}_3] = E[X \,|\, \mathcal{F}_1 \vee \mathcal{F}_3]$ for all $\mathcal{F}_2$-measurable $X \in L_1$.*

We will make repeated use of the following technical results.

Lemma A.3 *Let $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3, \mathcal{G}, \mathcal{G}_1$ and $\mathcal{G}_2$ be sub-σ-algebras of $\mathcal{F}$.*

(a) *If $\mathcal{F}_1 \perp \mathcal{F}_2$ given $\mathcal{F}_3$, then $\mathcal{F}_2 \perp \mathcal{F}_1$ given $\mathcal{F}_3$. If, in addition, $\mathcal{G}_i \subseteq \mathcal{F}_i$ $(i = 1, 2)$, then $\mathcal{G}_1 \perp \mathcal{G}_2$ given $\mathcal{F}_3$.*

(b) *If $\mathcal{F}_1 \perp \mathcal{F}_2$ given $\mathcal{F}_3$, then $(\mathcal{F}_1 \vee \mathcal{F}_3) \perp (\mathcal{F}_2 \vee \mathcal{F}_3)$ given $\mathcal{F}_3$. If, in addition, $\mathcal{G} \subseteq \mathcal{F}_1 \vee \mathcal{F}_2$, then $\mathcal{F}_1 \perp \mathcal{F}_2$ given $\mathcal{F}_3 \vee \mathcal{G}$.*

Proof. (a) follows by definition whereas (b) is a special case of Lemma 2.2 in Carnal and Walsh (1991). □

Lemma A.4 *Let $\mathcal{F}_1, \mathcal{F}_2, \mathcal{G}$ and $\mathcal{H}$ be sub-σ-algebras of $\mathcal{F}$. If $\mathcal{F}_1 \perp \mathcal{F}_2$ given $\mathcal{H}$ and $\mathcal{G} \subseteq \mathcal{F}_2$, then $(\mathcal{F}_1 \vee \mathcal{G}) \perp \mathcal{F}_2$ given $\mathcal{H} \vee \mathcal{G}$.*

Proof. Since $\mathcal{G} \subseteq \mathcal{F}_1 \vee \mathcal{F}_2$ and $\mathcal{F}_1 \perp \mathcal{F}_2$ given $\mathcal{H}$, Lemma A.3 (b) implies $\mathcal{F}_1 \perp \mathcal{F}_2$ given $\mathcal{H} \vee \mathcal{G}$ and thus $(\mathcal{F}_1 \vee \mathcal{H} \vee \mathcal{G}) \perp (\mathcal{F}_2 \vee \mathcal{H} \vee \mathcal{G})$ given $\mathcal{H} \vee \mathcal{G}$. Therefore, by Lemma A.3 (a), $(\mathcal{F}_1 \vee \mathcal{G}) \perp \mathcal{F}_2$ given $\mathcal{H} \vee \mathcal{G}$. □

In the sequel, we will require three assumptions apart from (32). The first concerns $\mathcal{A}$ and is satisfied by all examples we have in mind.

Assumption A.5 There exists a binary operation $\vee$ on $\mathcal{A}$ such that $\mathcal{A}$ is a distributive lattice under $\vee$ and $\wedge = \cap$.

Assumption A.6 If $A \subseteq \bigcup_{i=1}^{n} A_i$ $(A, A_i \in \mathcal{A})$, then $\mathcal{F}_A \subseteq \bigvee_{i=1}^{n} \mathcal{F}_{A_i}$.

Trivially, if $A = \bigcup_{i=1}^n A_i$, $A_i \in \mathcal{A}$, Assumption A.6 implies $\mathcal{F}_A = \bigvee_{i=1}^n \mathcal{F}_{A_i}$. If $\mathcal{A}$ satisfies the shape property, then any $\mathcal{A}$-indexed filtration automatically satisfies Assumption A.6. Moreover, for any process $X = (X_A)_{A\in\mathcal{A}}$, the family $(\mathcal{H}_A)_{A\in\mathcal{A}}$ where $\mathcal{H}_A = \sigma(\{X_B : B \in \mathcal{A}, B \subseteq A\})$ can be easily shown to satisfy Assumption A.6.

Our final supplimentary assumption complements Assumption 2.2. Examples for which it is satisfied include both the lower rectangles and the lower layers in $[0, 1]^2$.

Assumption A.7 Given any $C = A \setminus \bigcup_{i=1}^n A_i \in \mathcal{C}$ $(A, A_i \in \mathcal{A})$, there exists a maximal representation $A \setminus \bigcup_{i=1}^m B_i$ of C such that $B_i \cap B_j \subseteq A \;\; \forall\, i \neq j$.

The key implication of (32) is the following:

Lemma A.8 *Under* (32) *and Assumption A.5, if A, $A_1, \cdots, A_n \in \mathcal{A}$ are such that $A_i \cap A_j \subseteq A$ $\forall\, i \neq j$, then $E[X \mid \bigvee_{i=1}^n \mathcal{F}_{A_i}] = E[X \mid \bigvee_{i=1}^n \mathcal{F}_{A\cap A_i}]$ for every $\mathcal{F}_A$-measurable random variable $X \in L_1$.*

Proof. Since $\vee$ is distributive over $\cap$ and $A_1 \cap A_i \subseteq A_1 \cap A \;\; \forall\, 2 \le i \le n$,

$$A_1 \cap [A \vee (\textstyle\bigvee_{i=2}^n A_i)] \;=\; (A_1 \cap A) \vee (\textstyle\bigvee_{i=2}^n A_1 \cap A_i) \;=\; A_1 \cap A \tag{35}$$

which by Lemma A.1 implies $\mathcal{F}_{A_1} \perp \mathcal{F}_{A\vee(\vee_{i=2}^n A_i)}$ given $\mathcal{F}_{A\cap A_1}$. Therefore, defining $\mathcal{G} = \bigvee_{i=2}^n \mathcal{F}_{A_i}$, Lemma A.4 implies $(\bigvee_{i=1}^n \mathcal{F}_{A_i}) \perp \mathcal{F}_{A\vee(\vee_{i=2}^n A_i)}$ given $\mathcal{F}_{A\cap A_1} \vee (\bigvee_{i=2}^n \mathcal{F}_{A_i})$ which by Lemma A.3 (a) implies

$$(\textstyle\bigvee_{i=1}^n \mathcal{F}_{A_i}) \perp \mathcal{F}_A \;\text{ given }\; \mathcal{F}_{A\cap A_1} \vee (\textstyle\bigvee_{i=2}^n \mathcal{F}_{A_i})\,. \tag{36}$$

Next, take any $2 \le j \le n-1$. By an argument similar to (35), $A_j \cap [A \vee (\bigvee_{i=j+1}^n A_i)] = A_j \cap A$ so that Lemma A.1 implies $\mathcal{F}_{A_j} \perp \mathcal{F}_{A\vee(\vee_{i=j+1}^n A_i)}$ given $\mathcal{F}_{A\cap A_j}$. If we let $\mathcal{G} = (\bigvee_{i=1}^{j-1} \mathcal{F}_{A\cap A_i}) \vee (\bigvee_{i=j+1}^n \mathcal{F}_{A_i})$, then by Lemma A.4,

$$(\textstyle\bigvee_{i=1}^{j-1} \mathcal{F}_{A\cap A_i}) \vee (\textstyle\bigvee_{i=j}^n \mathcal{F}_{A_i}) \perp \mathcal{F}_A \;\text{ given }\; (\textstyle\bigvee_{i=1}^{j} \mathcal{F}_{A\cap A_i}) \vee (\textstyle\bigvee_{i=j+1}^n \mathcal{F}_{A_i}) \tag{37}$$

when we replace $\mathcal{F}_{A\vee(\vee_{i=j+1}^n A_i)}$ by $\mathcal{F}_A$ via Lemma A.3 (a).

Finally, since $\mathcal{F}_{A_n} \perp \mathcal{F}_A$ given $\mathcal{F}_{A\cap A_n}$,

$$(\textstyle\bigvee_{i=1}^{n-1} \mathcal{F}_{A\cap A_i}) \vee \mathcal{F}_{A_n} \perp \mathcal{F}_A \;\text{ given }\; \textstyle\bigvee_{i=1}^n \mathcal{F}_{A\cap A_i} \tag{38}$$

when we add $\mathcal{G} = \bigvee_{i=1}^{n-1} \mathcal{F}_{A\cap A_i}$ via Lemma A.4.

Now, take any $\mathcal{F}_A$-measurable $X \in L_1$. By Lemma A.2, (36) implies

$$E[X \mid \textstyle\bigvee_{i=1}^n \mathcal{F}_{A_i}] \;=\; E[\,X \mid \mathcal{F}_{A\cap A_1} \vee (\textstyle\bigvee_{i=2}^n \mathcal{F}_{A_i})\,] \tag{39}$$

while for each $2 \le j \le n-1$, (37) implies

$$E[\,X \mid (\textstyle\bigvee_{i=1}^{j-1} \mathcal{F}_{A\cap A_i}) \vee (\textstyle\bigvee_{i=j}^n \mathcal{F}_{A_i})\,] \;=\; E[\,X \mid (\textstyle\bigvee_{i=1}^{j} \mathcal{F}_{A\cap A_i}) \vee (\textstyle\bigvee_{i=j+1}^n \mathcal{F}_{A_i})\,] \tag{40}$$

and (38) implies

$$E[\,X \mid (\textstyle\bigvee_{i=1}^{n-1} \mathcal{F}_{A\cap A_i}) \vee \mathcal{F}_{A_n}\,] \;=\; E[X \mid \textstyle\bigvee_{i=1}^n \mathcal{F}_{A\cap A_i}]\,. \tag{41}$$

Linking the n indentities from (39), (40) and (41) yields $E[X \mid \bigvee_{i=1}^n \mathcal{F}_{A_i}] = E[X \mid \bigvee_{i=1}^n \mathcal{F}_{A\cap A_i}]$ which completes the proof. □

As the next result illustrates, under certain conditions, conditioning with respect to the strong past $\mathcal{G}_C^*$ is equivalent to conditioning with respect to a much simpler σ-algebra.

Proposition A.9 *Take any* $C = A \setminus \bigcup_{i=1}^{n} A_i$ *(*$A, A_i \in \mathcal{A}$*). Under* (32) *and Assumptions A.5, A.6 and A.7, if* $X \in L_1$ *is* $\mathcal{F}_A$*-measurable, then* $E[X \mid \mathcal{G}_C^*] = E[X \mid \bigvee_{i=1}^{n} \mathcal{F}_{A_i}]$

Proof. Let $A \setminus \bigcup_{i=1}^{m} B_i$ be the maximal representation of C described in Assumption A.7. By maximality, $\mathcal{G}_C^* = \mathcal{F}_{\bigcup_{i=1}^{n} B_i}$ so that $\mathcal{G}_C^* = \bigvee_{i=1}^{n} \mathcal{F}_{B_i}$ by Assumption A.6. Furthemore, Lemma A.8 implies $E[X \mid \bigvee_{i=1}^{n} \mathcal{F}_{B_i}] = E[X \mid \bigvee_{i=1}^{n} \mathcal{F}_{A \cap B_i}]$. Therefore, by the tower property, it is sufficient to show $\bigvee_{i=1}^{m} \mathcal{F}_{A \cap B_i} \subseteq \bigvee_{i=1}^{n} \mathcal{F}_{A_i} \subseteq \mathcal{G}_C^*$. Indeed, the right-most inclusion is trivial since $A_i \cap C = \phi \ \forall\, i$ whereas, for the left-most inclusion, it is sufficient to note that $A \setminus \bigcup_{j=1}^{m} A \cap B_j = A \setminus \bigcup_{i=1}^{n} A \cap A_i$ implies $A \cap B_j \subseteq \bigcup_{i=1}^{n} A_i \ \ \forall\, j$. □

Since every $C \in \mathcal{C}$ possesses a minimal representation, Proposition A.9 yields

Corollary A.10 *Under* (32) *and Assumptions A.5, A.6 and A.7, given any* $C = A \backslash B \in \mathcal{C}$ *(*$A \in \mathcal{A}$*,* $B \in \mathcal{A}(u)$*), if* $X \in L_1$ *is* $\mathcal{F}_A$*-measurable, then* $E[X \mid \mathcal{G}_C^*]$ *is* $\mathcal{F}_A$*-measurable.*

Unlike the square function inequality in Lemma 2.11, our set-indexed Rosenthal-like inequality requires conditional independence of $(\mathcal{F}_A)$.

Lemma A.11 *Let* M *be a strong martingale in* L_2*. Under* (32) *and Assumptions A.5, A.6 and A.7, if* $k_1, \cdots, k_r$ *and* $C_1, \cdots, C_r$ *are as defined in Lemma 2.10, then given any* $2 \leq p < \infty$*,*

$$E\left[\left(\sum_{i=1}^{r} E[(M_{C_i})^2 | \mathcal{G}_{C_i}^*]\right)^{p/2}\right] \ \leq \ \kappa \cdot E\left[\,|M_{C_0}|^p\,\right] \tag{42}$$

where $C_0 = \bigcup_{i=1}^{r} C_i$ *and* κ *is a positive constant depending only on* p*.*

Proof. Let $\mathcal{H}_1 \subseteq \cdots \subseteq \mathcal{H}_r$ be as defined in Lemma 2.10, and let $\mathcal{H}_0 = \bigvee_{j=1}^{k_1-1} \mathcal{F}_{g_m(A_j)}$ if $k_1 \geq 2$ or $\mathcal{F}_{\phi'}$ if $k_1 = 1$. In either case, $\mathcal{H}_0 \subseteq \mathcal{H}_1$ so that Theorem A.8-V in Hall and Heyde (1980) yields

$$E\left[\left(\sum_{i=1}^{r} E[(M_{C_i})^2 | \mathcal{H}_{i-1}]\right)^{p/2}\right] \ \leq \ K \cdot E\left[\left(\sum_{i=1}^{r} (M_{C_i})^2\right)^{p/2}\right]$$

for some positive constant K. But by (2) and (4), Proposition A.9 implies

$$E[(M_{C_i})^2 | \mathcal{G}_{C_i}^*] \ = \ E[(M_{C_i})^2 | \mathcal{H}_{i-1}] \quad \forall\, 1 \leq i \leq r$$

so that (42) follows by Lemma 2.11. □

References

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

[2] Carnal, E. and Walsh, J. B. (1991). Markov properties for certain random fields, *Stochastic analysis*, Academic Press, 91-110.

[3] Dellacherie, C. and Meyer, P. A. (1978). *Probabilities and Potential*, Hermann.

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

[5] Gushchin, A. A. (1982). On the general theory of random fields in the plane, *Russian Mathematical Surveys* **37**:6, 55-80.

[6] Hall, P. and Heyde, C. C. (1980). *Martingale Limit Theory and its Applications*, Academic Press.

[7] 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.

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

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

[10] Ivanoff, B. G., Merzbach, E. and Schiopu-Kratina, I. (1993). Predictablility and stopping on lattices of sets, *Probability Theory and Related Fields* **97**, 433-446.

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

[12] Slonowsky, D. and Ivanoff, B. G. (1999). A central limit theorem for set-indexed strong martingales, preprint.