Weak Limit Theorems for Stochastic Integrals and Stochastic Differential Equations

by

Thomas G. Kurtz
University of Wisconsin-Madison

and

Philip Protter
Purdue University

Technical Report #90-54

Department of Statistics
Purdue University

October 1990

Weak limit theorems for stochastic integrals and stochastic differential equations

Thomas G. Kurtz
Departments of Mathematics and Statistics
University of Wisconsin-Madison
Madison, WI 53706

Philip Protter
Department of Mathematics
Purdue University
West Lafayette, IN 47907

Abstract

Assuming that $\{(X_n,Y_n)\}$ is a sequence of cadlag processes converging in distribution to (X,Y) in the Skorohod topology, conditions are given under which the sequence $\{\int X_n\,dY_n\}$ converges in distribution to $\int X\,dY$. Examples of applications are given drawn from statistics and filtering theory. In particular, assuming that $(U_n,Y_n) \Rightarrow (U,Y)$ and that $F_n \to F$ in an appropriate sense, conditions are given under which solutions of a sequence of stochastic differential equations $dX_n = dU_n + F_n(X_n)\,dY_n$ converge to a solution of $dX = dU + F(X)\,dY$ where F_n and F may depend on the past of the solution. As is well-known from work of Wong and Zakai, this last conclusion fails if Y is Brownian motion and the Y_n are obtained by linear interpolation; however, the present theorem may be used to derive a generalization of the results of Wong and Zakai and their successors.

AMS 1980 *Subject Classification.* Primary 60H05,60F17; secondary 60G44
Key words and phrases. Stochastic integrals, stochastic differential equations, weak convergence, Skorohod topology, filtering, symmetric statistics, Wong-Zakai correction

1. Introduction. For $n = 1,2,\ldots$ let $\{Y_k^n : k \geq 0\}$ be a Markov chain. The classical assumptions leading to a diffusion approximation for such a sequence are that the increments of the chain satisfy

$$E[Y_{k+1}^n - Y_k^n | \mathcal{F}_k^n] = b(Y_k^n)\tfrac{1}{n} + o(\tfrac{1}{n}) \tag{1.1}$$

and

$$E[(Y_{k+1}^n - Y_k^n)^2 | \mathcal{F}_k^n] = a(Y_k^n)\tfrac{1}{n} + o(\tfrac{1}{n}) \tag{1.2}$$

Using these assumptions we can write

$$Y_k^n = Y_0^n + \sum_{i=0}^{k-1}(Y_{i+1}^n - Y_i^n) \tag{1.3}$$

$$= Y_0^n + \sum_{i=0}^{k-1} b(Y_i^n)\tfrac{1}{n} + \sum_{i=0}^{k-1}\sqrt{a(Y_i^n)}\,Z_{i+1}^n \tfrac{1}{\sqrt{n}} + \text{error}$$

where

$$Z_{k+1}^n = \frac{Y_{k+1}^n - Y_k^n - E[Y_{k+1}^n - Y_k^n|\mathcal{F}_t^n]}{\sqrt{E[(Y_{k+1}^n - Y_k^n - E[Y_{k+1}^n - Y_k^n|\mathcal{F}_t^n])^2|\mathcal{F}_t^n]}} \tag{1.4}$$

are martingale differences with conditional variance 1. If we define $X_n(t) = Y_{[nt]}^n$ and

$$W_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} Z_i^n \tag{1.5}$$

then

$$X_n(t) = X_n(0) + \int_0^{\frac{[nt]}{n}} b(X_n(s))\,ds + \int_0^t \sqrt{a(X_n(s-))}\,dW_n(s) + \text{error} \tag{1.6}$$

(Note that X_n is constant on intervals of length $\frac{1}{n}$, so the first sum in (1.3) equals the first integral in (1.6).) Under mild additional assumptions, the martingale central limit theorem implies $W_n \Rightarrow W$, (throughout $\Rightarrow$ will denote convergence in distribution) where W is a standard Brownian motion. This convergence suggests that X_n should converge to a solution of the obvious limiting stochastic differential equation. This approach to deriving diffusion approximations has been taken by many authors (see, for example, Skorohod (1965), Chapter 6, Kushner (1974), and Strasser (1986)) although in recent years it has been largely replaced by methods which exploit the characterization of a Markov process as a solution of a martingale problem.

A key step in the application of the stochastic differential equation approach is to show that the sequence of stochastic integrals in the approximating equation converges to the corresponding stochastic integral in the limit. That there is a difficulty to be overcome is well-known from the work of Wong and Zakai (1965). See also Protter (1985).

Growing interest in stochastic differential equations driven by martingales (and more generally semimartingales) other than Brownian motion has led to renewed interest in this approach to the derivation of approximating processes. In addition, functionals of stochastic processes which can be represented by stochastic integrals arise in many areas of application including filtering and statistics. Limit theorems in these settings require conditions under which convergence of the integrand and integrator in a stochastic integral implies convergence of the integral.

Throughout, we will be considering cadlag processes (that is, processes X whose sample paths are right continuous and for which the left limit $X(t\text{-})$ exists at each $t > 0$). This restriction to cadlag processes allows us to define stochastic integrals as limits of Riemann-Stieltjes-like sums, that is,

$$\int_0^t X(s\text{-})\,dY(s) = \lim \sum X(t_i)(Y(t_{i+1}) - Y(t_i)) \tag{1.7}$$

where $\{t_i\}$ is a partition of $[0,t]$ and the limit is taken as the maximum of $t_{i+1} - t_i$ tends to zero. The integral exists if the limit exists in probability. Recall that the choice of the left end-point of $[t_i,t_{i+1})$ as the argument of X is critical even when Y is a Brownian motion. Indeed in the Brownian differential case, if we take the argument of X to be the midpoint, we obtain the Stratonovich integral. (We will, of course, assume that X is adapted (and hence the left continuous process $X(\cdot\text{-})$ is predictable) and that Y is a semimartingale for the same filtration, but the uninitiated reader can follow much of what is going on without a thorough knowledge of these matters.) Throughout, we will use Protter (1990) as our basic reference for material on semimartingales and stochastic integration. See this volume for details and further references.

The following two examples will help motivate the assumptions of the main theorem.

<u>1.1 Example</u> Let $X = Y = X_n = \chi_{[1,\infty)}$ and $Y_n = \chi_{[1+\frac{1}{n},\infty)}$. Then for $t > 1+\frac{1}{n}$,

$$\int_0^t X_n(s\text{-})\,dY_n(s) = 1 \tag{1.8}$$

but the limiting integral gives

$$\int_0^t X(s\text{-})\,dY(s) = 0 \tag{1.9}$$

<u>1.2 Example</u> Let W be standard Brownian motion, and define W_n so that

$$\frac{d}{dt}W_n(t) = n(W(\tfrac{k+1}{n}) - W(\tfrac{k}{n})),\ \ t \in [\tfrac{k}{n},\tfrac{k+1}{n}) \tag{1.10}$$

Then

$$(1.11)\quad \int_0^t W_n(s-)\,dW_n(s)$$

$$= \int_0^t W_n(\tfrac{[ns]}{n})\,dW_n(s) + \int_0^t (W_n(s) - W_n(\tfrac{[ns]}{n}))\,dW_n(s)$$

$$= \sum W(\tfrac{k}{n})(W(\tfrac{k+1}{n}) - W(\tfrac{k}{n})) + \sum \int_0^{\frac{1}{n}} (W(\tfrac{k}{n}+s) - W(\tfrac{k}{n}))(W(\tfrac{k+1}{n}) - W(\tfrac{k}{n}))\,ds$$

$$\to \int_0^t W(s)\,dW(s) + \tfrac{1}{2}t$$

Example 1.1 is indicative of problems that will arise whenever the integrand and the integrator have discontinuities which "coalesce" in the wrong way. We will avoid these difficulties by requiring that the pair of processes (X_n, Y_n) converge in the Skorohod topology on $D_{\mathbf{R}^2}[0,\infty)$ which is stronger than assuming convergence of each component in $D_{\mathbf{R}}[0,\infty)$. For future reference, let Λ denote the collection of continuous, strictly increasing functions mapping $[0,\infty)$ onto $[0,\infty)$. Recall that for any metric space E a sequence of cadlag, E-valued functions $\{x_n\}$ converges in the Skorohod topology to x, if there exists a sequence $\{\lambda_n\} \subset \Lambda$ such that $x_n \circ \lambda_n(t) \to x(t)$ and $\lambda_n(t) \to t$ uniformly for t in bounded intervals. Note that in Example 1.1, Y_n converges in the Skorohod topology with $E = \mathbf{R}$, but the pair (X_n, Y_n) does not converge in the Skorohod topology with $E = \mathbf{R}^2$, and in general, convergence in the Skorohod topology with $E = \mathbf{R}^2$ excludes the possibility of the type of coalescence of jumps that causes the problem in that example. In particular, for each n, let y_n be piecewise constant, and suppose the number of discontinuities of y_n in a bounded time interval is uniformly bounded in n. Then if $(x_n, y_n) \to (x, y)$ in the Skorohod topology on $D_{\mathbf{R}^2}[0,\infty)$,

$$(1.12)\qquad \int_0^{\cdot} x_n(s-)\,dy_n(s) \to \int_0^{\cdot} x(s-)\,dy(s)$$

and

$$(1.13)\qquad \int_0^{\cdot} y_n(s-)\,dx_n(s) \to \int_0^{\cdot} y(s-)\,dx(s)$$

in the Skorohod topology on $D_{\mathbf{R}}[0,\infty)$. (Actually, the quadruple consisting of x_n, y_n, and the two integrals converges in $D_{\mathbf{R}^4}[0,\infty)$).

Example 1.2 points to more subtle problems, and we will come back to it when we discuss the hypotheses of the main theorem.

We will formulate the main theorem, Theorem 2.2, in Section 2. This theorem is essentially the same as that given by Jakubowski, Mémin, and Pages (1988), but we believe that our formulation and proof are more readily accessible to researchers without extensive expertise in the theory of semimartingales and stochastic integration. Section 3 will be devoted to further examples and applications. Section 4 contains some relative compactness results for stochastic integrals and some variations on the main theorem. Applications to stochastic differential equations will be discussed in Section 5. In particular, we generalize results of Slomiński (1989). Some technical results will be given in Section 6.

2. Weak convergence of stochastic integrals. Throughout we will be making various transformations of the processes involved. We will need to have information about the continuity properties of these transformations, and the following lemma will be useful in obtaining this information.

2.1 Lemma Let E_1 and E_2 be metric spaces, and let $F:D_{E_1}[0,\infty) \to D_{E_2}[0,\infty)$. Suppose $F(x \circ \lambda) = F(x) \circ \lambda$ for all $x \in D_{E_1}[0,\infty)$ and all $\lambda \in \Lambda$. Suppose $x_n(t) \to x(t)$ uniformly for t in bounded intervals implies $F(x_n) \to F(x)$ in the Skorohod topology. Then $x_n \to x$ in the Skorohod topology implies that $F(x_n) \to F(x)$ in the Skorohod topology. If $x_n(t) \to x(t)$ uniformly on bounded intervals implies $F(x_n)(t) \to F(x)(t)$ uniformly on bounded intervals, then $x_n \to x$ in the Skorohod topology implies $(x_n,F(x_n)) \to (x,F(x))$ in the Skorohod topology on $D_{E_1 \times E_2}[0,\infty)$.

Proof Suppose $x_n \to x$ in the Skorohod topology. Then there exist $\lambda_n \in \Lambda$ such that $x_n \circ \lambda_n(t) \to x(t)$ and $\lambda_n(t) \to t$ uniformly on bounded intervals. It follows that $F(x_n \circ \lambda_n) \to F(x)$ in the Skorohod topology, so there exist $\eta_n \in \Lambda$ such that $\eta_n(t) \to t$ and $F(x_n \circ \lambda_n) \circ \eta_n(t) \to F(x)(t)$ uniformly on bounded intervals. Since $\lambda_n \circ \eta_n(t) \to t$ and $F(x_n) \circ \lambda_n \circ \eta_n(t) = F(x_n \circ \lambda_n) \circ \eta_n(t) \to F(x)(t)$ uniformly on bounded intervals, it follows that $F(x_n) \to F(x)$ in the Skorohod topology. The last statement is immediate from the definition of the Skorohod topology. □

The following functional gives a good example of an application of the lemma. Fix m, and define $h_\delta:[0,\infty) \to [0,\infty)$ by $h_\delta(r) = (1 - \delta/r)^+$. Define $J_\delta:D_{\mathbf{R}^m}[0,\infty) \to D_{\mathbf{R}^m}[0,\infty)$ by

(2.1)
$$J_\delta(x)(t) = \sum_{s \le t} h_\delta(|x(s) - x(s-)|)(x(s) - x(s-))$$

Lemma 2.1 shows that $x \to J_\delta(x)$ and $x \to x - J_\delta(x)$ are continuous. Consequently, by (1.12), if $(x_n,y_n) \to (x,y)$, then

(2.2)
$$\int_0^\cdot x_n(s-)\,dJ_\delta(y_n)(s) \to \int_0^\cdot x(s-)\,dJ_\delta(y)(s)$$

Let $\{\mathcal{F}_t\}$ be a filtration. A cadlag, $\{\mathcal{F}_t\}$-adapted process Y is a semimartingale if it can be decomposed as $Y = M + A$ where M is an $\{\mathcal{F}_t\}$-local martingale and the sample paths of A have finite variation on bounded time intervals, that is, there exists a sequence of $\{\mathcal{F}_t\}$-

stopping times, τ_k, such that $\tau_k \to \infty$ a.s and for each k, $M^{\tau_k} \equiv M(\cdot\wedge\tau_k)$ is a uniformly integrable martingale, and for every $t > 0$, $T_t(A) = \sup \sum|A(t_{i+1}) - A(t_i)| < \infty$ a.s (where the supremum is over partitions of $[0,t]$).

An $\mathbb{R}^m$-valued process is an $\{\mathcal{F}_t\}$-semimartingale, if each component is a semimartingale. Let $\mathbb{M}^{km}$ denote the real-valued, k×m matrices. Throughout, $\int X\,dY$ will denote $\int X(s-)\,dY(s)$.

<u>2.2 Theorem</u> For each n, let (X_n,Y_n) be an $\{\mathcal{F}^n_t\}$-adapted process with sample paths in $D_{\mathbb{M}^{km}\times\mathbb{R}^m}[0,\infty)$, and let Y_n be an $\{\mathcal{F}^n_t\}$-semimartingale. Fix $\delta > 0$ (allowing $\delta = \infty$), and define $Y^\delta_n = Y_n - J_\delta(Y_n)$. (Note that Y^δ_n will also be a semimartingale.) Let $Y^\delta_n = M^\delta_n + A^\delta_n$ be a decomposition of Y^δ_n into an $\{\mathcal{F}^n_t\}$-local martingale and a process with finite variation. Suppose

C2.2(i) For each $\alpha > 0$, there exist stopping times $\{\tau^\alpha_n\}$ such that $P\{\tau^\alpha_n \le \alpha\} \le \frac{1}{\alpha}$ and $\sup_n E[[M^\delta_n]_{t\wedge\tau^\alpha_n} + T_{t\wedge\tau^\alpha_n}(A^\delta_n)] < \infty$.

If $(X_n,Y_n) \Rightarrow (X,Y)$ in the Skorohod topology on $D_{\mathbb{M}^{km}\times\mathbb{R}^m}[0,\infty)$, then Y is a semimartingale with respect to a filtration to which X and Y are adapted, and $(X_n,Y_n,\int X_n\,dY_n) \Rightarrow (X,Y,\int X\,dY)$ in the Skorohod topology on $D_{\mathbb{M}^{km}\times\mathbb{R}^m\times\mathbb{R}^k}[0,\infty)$. If $(X_n,Y_n) \to (X,Y)$ in probability, then the triple converges in probability.

<u>2.3 Remark</u> If there exist decompositions of $\{Y^\delta_n\}$ such that C2.2(i) holds, we will simply say that $\{Y_n\}$ satisfies C2.2(i) for δ. For $c > 0$, define $\tau^c_n = \inf\{t:|M^\delta_n(t)|\vee|M^\delta_n(t-)| \ge c$ or $T_t(A^\delta_n) \ge c\}$. Suppose the following conditions hold.

C2.2(ii) $\{T_t(A^\delta_n)\}$ is stochastically bounded for each $t > 0$.

C2.2(iii) For each $c > 0$, $\sup_n E[M^\delta_n(t\wedge\tau^c_n)^2 + T_{t\wedge\tau^c_n}(A^\delta_n)] < \infty$

Since convergence in distribution of $\{Y_n\}$ in the Skorohod topology implies stochastic boundedness for $\{\sup_{t\le\alpha} Y_n(t)\}$, $\sup_{t\le\alpha}|M^\delta_n(t)| = \sup_{t\le\alpha}|Y^\delta_n(t) - A^\delta_n(t)| \le \sup_{t\le\alpha}|Y_n(t)| + T_\alpha(A^\delta_n)$ is stochastically bounded in n for each α, and hence there exists c_α so that $P\{\tau^{c_\alpha}_n \le \alpha\} \le \frac{1}{\alpha}$. In addition $E[[M^\delta_n]_{t\wedge\tau^{c_\alpha}_n}] = E[(M^\delta_n(t\wedge\tau^{c_\alpha}_n)^2]$, and C2.2(i) is satisfied with $\tau^\alpha_n = \tau^{c_\alpha}_n$.

For $\delta < \infty$, C2.2(iii) will usually be immediate since the discontinuities of Y_n^δ are bounded in magnitude by δ (making Y_n^δ a special semimartingale) and there will exist a decomposition with the discontinuities of each term bounded by 2δ (see Jacod and Shiryaev (1987), Lemma I.4.24).

2.4 Remark To see that Y is a semimartingale it is enough to show that Y^δ is a semimartingale. Without loss of generality, we can assume that for $\alpha = 1,2,\ldots$, $\tau_n^\alpha \le \tau_n^{\alpha+1}$. Let $Y_n^{\delta\alpha} = Y_n^\delta(\cdot\wedge\tau_n^\alpha)$. Then $\{(X_n,Y_n,Y_n^\delta,Y_n^{\delta 1},Y_n^{\delta 2},\ldots,\tau_n^1,\tau_n^2,\ldots)\}$ is relatively compact in $D_{\mathbf{M}^{km}\times\mathbf{R}^m\times\mathbf{R}^m}[0,\infty)\times D_{\mathbf{R}^m}[0,\infty)^\infty\times[0,\infty]^\infty$. Let $(X,Y,Y^\delta,Y^{\delta 1},Y^{\delta 2},\ldots,\tau^1,\tau^2\ldots)$ be some limit point, and let $\{\mathcal{F}_t\}$ be the filtration generated by the limiting processes and random times. For each $T > 0$, let

$$V_T(Y_n^{\delta\alpha}) \equiv \sup E[\sum|E[Y_n^{\delta\alpha}(t_{i+1}) - Y_n^{\delta\alpha}(t_i)|\mathcal{F}_t^n]|] \tag{2.3}$$

where the supremum is over all partitions of $[0,T]$. Then

$$\sup_n V_T(Y_n^{\delta\alpha}) \le \sup_n E[T_{T\wedge\tau_n^\alpha}(A_n^\delta)] < \infty \tag{2.4}$$

and hence $V_T(Y^{\delta\alpha}) < \infty$ (V_T defined using $\{\mathcal{F}_t\}$). (See for example Meyer and Zheng (1984) Theorem 4 or Kurtz (1990) Theorem 5.8) It follows that $Y^{\delta\alpha}$ is a local $\{\mathcal{F}_t\}$-quasi-martingale and hence an $\{\mathcal{F}_t\}$-semimartingale. But

$$Y^\delta(t\wedge\tau^\alpha) = Y^{\delta\alpha}(t) + (Y^\delta(\tau^\alpha) - Y^{\delta\alpha}(\tau^\alpha))\chi_{\{\tau^\alpha\le t\}} \tag{2.5}$$

so Y^δ is a local $\{\mathcal{F}_t\}$-semimartingale and hence an $\{\mathcal{F}_t\}$-semimartingale.

2.5 Remark If $Y_n \equiv Y$ for each n, then $\{Y_n\}$ satisfies C2.2(i) for all finite δ. If $\{Y_n\}$ is relatively compact in the Skorohod topology and satisfies C2.2(i) for some $\delta \in (0,\infty]$, then $\{Y_n\}$ satisfies C2.2(i) for all finite δ. If $\{(X_n,Y_n)\}$ is relatively compact in the Skorohod topology and $\{Y_n\}$ satisfies C2.2(i) for some $\delta \in (0,\infty]$, then $\{\int X_n\,dY_n\}$ satisfies C2.2(i) for all finite δ.

2.6 Remark With reference to Example 1.2, note that $T_t(W_n) = O(\sqrt{n})$.

Proof Let $Z_n = (X_n,Y_n,J_\delta(Y_n),Y_n^\delta)$. Z_n has sample paths in $D_E[0,\infty)$ for $E = \mathbb{M}^{km}\times\mathbb{R}^m\times\mathbb{R}^m\times\mathbb{R}^m$. The limit in (1.13) suggests attempting to approximate X_n by a piecewise constant process. The problem is to find such an approximation that converges in distribution along with X_n (in fact, along with Z_n). Furthermore, the approximation must be adapted to a filtration with respect to which Y_n is a semimartingale. By Lemma 6.1, there exists a (random) mapping $I_\epsilon:D_E[0,\infty) \to D_E[0,\infty)$ such that $|z(t) - I_\epsilon(z)(t)| \le \epsilon$ for all $z \in D_E[0,\infty)$ and $t \ge 0$, $I_\epsilon(z)$ is a step function, and the mapping $z \to (z,I_\epsilon(z))$ is continuous at z a.s for each $z \in D_E[0,\infty)$. Furthermore, $I_\epsilon(Z_n)$ is adapted to a filtration $\mathcal{G}_t^n = \mathcal{F}_t^n \vee \mathcal{H}$, where $\mathcal{H}$ is independent of $\{\mathcal{F}_t^n\}$ (and hence Y_n will be a $\{\mathcal{G}_t^n\}$-semimartingale). Let X_n^ϵ denote the first, $\mathbb{M}^{km}$-valued component of $I_\epsilon(Z_n)$. Then $|X_n - X_n^\epsilon| \le \epsilon$, and $(X_n,Y_n,J_\delta(Y_n),Y_n^\delta,X_n^\epsilon) \Rightarrow (X,Y,J_\delta(Y),Y^\delta,X^\epsilon)$.

Define $U_n = \int X_n\,dY_n$ and $U_n^\epsilon = \int X_n^\epsilon\,dY_n^\delta + \int X_n\,dJ_\delta(Y_n)$ with similar definitions for U and U^ϵ. Then it follows as in (1.12) and (1.13) that $(X_n,Y_n,U_n^\epsilon) \Rightarrow (X,Y,U^\epsilon)$ in $D_{\mathbb{M}^{km}\times\mathbb{R}^m\times\mathbb{R}^k}[0,\infty)$. Observing that

$$(2.6)\qquad R_n^\epsilon \equiv U_n - U_n^\epsilon = \int(X_n - X_n^\epsilon)\,dY_n^\delta = \int(X_n - X_n^\epsilon)\,dM_n^\delta + \int(X_n - X_n^\epsilon)\,dA_n^\delta$$

we see that for any stopping time τ

$$(2.7)\qquad E[\sup_{s\le t\wedge\tau}|R_n^\epsilon(s)|] \le \epsilon\Big(2E[[M_n^\delta]_{t\wedge\tau}]^{\frac{1}{2}} + E[T_{t\wedge\tau}(A_n^\delta)]\Big)$$

with similar estimates holding for $U - U^\epsilon$. Applying C2.2(i), it follows that $(X_n,Y_n,U_n) \Rightarrow (X,Y,U)$.

A review of the proof shows that if convergence in distribution is replaced by convergence in probability in the hypotheses, then convergence in probability will hold in the conclusion. □

The transformation J_δ provides a convenient, continuous way to eliminate the large jumps from Y_n in Theorem 2.2. Occasionally, however, it may be useful to apply some other truncation of the large jumps. For example, if Y_n is a martingale it may be possible to truncate the large jumps in such a way that the truncated process is still a martingale, simplifying the verification of the hypotheses of the theorem. With these possibilities in mind, we state a slightly more general version of the theorem.

2.7 Theorem For each n, let (X_n,Y_n) be an $\{\mathcal{F}_t^n\}$-adapted process with sample paths in $D_{\mathbb{M}^{km}\times\mathbb{R}^m}[0,\infty)$, and let Y_n be an $\{\mathcal{F}_t^n\}$-semimartingale. Suppose that $Y_n = M_n + A_n + Z_n$, where M_n is a local $\{\mathcal{F}_t^n\}$-martingale, A_n is an $\{\mathcal{F}_t^n\}$-adapted, finite variation process, and Z_n is constant except for finitely many discontinuities in any finite time interval. Let $N_n(t)$ denote the number of discontinuities of Z_n in the interval $[0,t]$. Suppose $\{N_n(t)\}$ is stochastically bounded for each $t > 0$, and

C2.7 For each $\alpha > 0$, there exist stopping times $\{\tau_n^\alpha\}$ such that $P\{\tau_n^\alpha \le \alpha\} \le \frac{1}{\alpha}$ and $\sup_n E[[M_n]_{t\wedge\tau_n^\alpha} + T_{t\wedge\tau_n^\alpha}(A_n)] < \infty$.

If $(X_n,Y_n,Z_n) \Rightarrow (X,Y,Z)$ in the Skorohod topology on $D_{\mathbb{M}^{km}\times\mathbb{R}^m\times\mathbb{R}^m}[0,\infty)$, then Y is a semimartingale with respect to a filtration to which X and Y are adapted, and $(X_n,Y_n,\int X_n\,dY_n) \Rightarrow (X,Y,\int X\,dY)$ in the Skorohod topology on $D_{\mathbb{M}^{km}\times\mathbb{R}^m\times\mathbb{R}^k}[0,\infty)$. If $(X_n,Y_n,Z_n) \to (X,Y,Z))$ in probability, then convergence in probability holds in the conclusion.

3. Examples and applications

3.1 Example As a simple first example, we consider limit theorems for sums of products of independent random variables which arise in the study of U-statistics. Let $\{\xi_i\}$ be i.i.d. real-valued random variables with mean zero and variance σ^2. Define

$$W_n^{(k)}(t) = \frac{1}{n^{k/2}} \sum_{1\le i_1<\cdots i_k\le[nt]} \xi_{i_1}\cdots\xi_{i_k} \tag{3.1}$$

and $Z_n = (W_n^{(1)},\ldots,W_n^{(m)})$. Note that $W_n^{(1)} \Rightarrow \sigma W$, where W is standard Brownian motion, and observe that we can write

$$W_n^{(k)}(t) = \int_0^t W_n^{(k-1)}(s-)\,dW_n^{(1)}(s) \tag{3.2}$$

It follows (by induction) that $Z_n \Rightarrow Z = (W^{(1)},\ldots,W^{(m)})$, where $W^{(1)} = \sigma W$ and $W^{(k)}$ is the corresponding interated integral. (Note that $X_n \Rightarrow X$ in $D_E[0,\infty)$ implies that $(X_n,X_n) \Rightarrow (X,X)$ in $D_{E\times E}[0,\infty)$).

3.2 Example (Bobkoski (1983)) Let $\{\xi_i\}$ be as above. For a constant ϕ, let $\{Y_k\}$ satisfy

$$Y_{k+1} = \phi Y_k + \xi_{k+1} \tag{3.3}$$

Given $Y_1,\ldots,Y_m$, the least squares estimate $\hat{\phi}$ for an unknown ϕ is the value of ϕ minimizing $\sum(Y_{k+1} - \phi Y_k)^2$, that is, the solution of

$$\sum Y_k(Y_{k+1} - \phi Y_k) = 0 \tag{3.4}$$

given by

$$\hat{\phi} = \frac{\sum Y_k Y_{k+1}}{\sum Y_k^2} \tag{3.5}$$

Now consider a sequence of such processes $\{Y_k^n\}$ in which the true $\phi_n = (1 - \frac{\beta}{n})$. If we define $X_n(t) = \frac{1}{\sqrt{n}} Y_{[nt]}^n$

$$X_n(t) = \phi_n^{[nt]} X_n(0) + \int_0^t \phi_n^{[nt]-1-[ns-]} dW_n(s) \tag{3.6}$$

where $W_n = W_n^{(1)}$, and if $X_n(0) \to X(0)$, it follows that $X_n \Rightarrow X$ given by

$$X(t) = e^{-\beta t} X(0) + \int_0^t e^{-\beta(t-s)} \sigma\, dW(s) \tag{3.7}$$

Note that X is an Ornstein-Uhlenbeck process satisfying $dX = -\beta X\, dt + \sigma\, dW$. For the least squares estimate of ϕ_n at time t, we have

$$\sum_{k=0}^{[nt]-1} Y_k^n\Big((\phi_n - \hat{\phi}_n) Y_k^n + \xi_{k+1}\Big) = 0 \tag{3.8}$$

which implies

$$n(\phi_n - \hat{\phi}_n) \int_0^{\frac{[nt]}{n}} X_n(s)^2\, ds = \int_0^t X_n(s-)\, dW_n(s) \tag{3.9}$$

and it follows that

$$n(\phi_n - \hat{\phi}_n) \Rightarrow \frac{\int_0^t \sigma X(s)\, dW(s)}{\int_0^t X(s)^2\, ds} \tag{3.10}$$

More general results along these lines have been given by Llatas (1987), Chan and Wei (1988), and Cox and Llatas (1989).

3.3 Example Work on approximation of nonlinear filters, DiMasi and Rungaldier (1981, 1982), Johnson (1983), Goggin (1988), involves studying the limiting behavior of a sequence of Girsanov-type densities, each of which typically includes the exponential of a stochastic integral. For example, let $\{X_n\}$ be a sequence of processes with sample paths in $D_E[0,\infty)$, such that $X_n \Rightarrow X$. Let N be a unit Poisson process independent of the X_n, let the observation process Y_n be given by

$$Y_n(t) = N\Big(n \int_0^t \big(\lambda + n^{-\frac{1}{2}} h(X_n(s))\big) ds\Big) \tag{3.11}$$

and define

(3.12) $$U_n(t) = n^{-\frac{1}{2}}(Y_n(t) - \lambda nt)$$

Note that $\mathcal{F}_t^{Y_n} = \mathcal{F}_t^{U_n}$ and observe that $(X_n,U_n) \Rightarrow (X,U)$ where for a standard Brownian motion W independent of X

(3.13) $$U(t) = \sqrt{\lambda}W(t) + \int_0^t h(X(s))\,ds$$

Suppose that (X_n,U_n) is defined on a probability space $(\Omega,\mathcal{F},P_n)$. Then there exists a probability measure Q_n on the same measurable space, $(\Omega,\mathcal{F})$, under which X_n has the same distribution as under P_n, Y_n is independent of X_n and is a Poisson process with parameter $n\lambda$, and $P_n \ll Q_n$ on $\mathcal{G}_t^n = \sigma(X_n(s),U_n(s):s\leq t)$ with

(3.14) $$L_n(t) = \frac{dP_n}{dQ_n}\bigg|_{\mathcal{G}_t^n}$$

$$= \exp\Big\{\int_0^t \ln\Big(1+n^{-\frac{1}{2}}\lambda^{-1}h(X_n(s-))\Big)dY_n(s) - \int_0^t n^{\frac{1}{2}}h(X_n(s))\,ds\Big\}$$

$$= \exp\Big\{\int_0^t n^{\frac{1}{2}}\ln\Big(1+n^{-\frac{1}{2}}\lambda^{-1}h(X_n(s-))\Big)dU_n(s)$$

$$+ \int_0^t\Big(n\lambda\ln\Big(1+n^{-\frac{1}{2}}\lambda^{-1}h(X_n(s-))\Big) - n^{\frac{1}{2}}h(X_n(s))\Big)ds\Big\}$$

Similarly, if (X,U) is defined on a probability space $(\Omega,\mathcal{F},P)$, there exists a measure Q on $(\Omega,\mathcal{F})$ such that, under Q, X has the same distribution as under P, U is independent of X with the same distribution as $\sqrt{\lambda}W$, and $P \ll Q$ on $\mathcal{G}_t = \sigma\{X(s),U(s):s\leq t\}$ with

(3.15) $$L(t) = \frac{dP}{dQ}\bigg|_{\mathcal{G}_t} = \exp\Big\{\int_0^t \lambda^{-1}h(X(s))\,dU(s) - \int_0^t \tfrac{1}{2}\lambda^{-1}h^2(X(s))\,ds\Big\}$$

Expanding the logarithm in (3.14) in a Taylor series and applying Theorem 2.2, we see that $L_n \Rightarrow L$ under $\{P_n\}, P$ and under $\{Q_n\}, Q$.

Results of Goggin (1988) can then be applied to show that the conditional distribution $\mu_n(t)$ of $X_n(t)$ given $\mathcal{F}_t^{Y_n}$ converges in distribution to the conditional distribution $\mu(t)$ of $X(t)$ given $\mathcal{F}_t^U$ as a process in $D_{\mathcal{P}(E)}[0,\infty)$.

3.4 Example (Meyer (1989), Emery (1989)) Next we consider the problem of showing existence of solutions of the structure equation arising in the study of chaotic representations formulated by Meyer. Given $F \in C(\mathbb{R})$, the problem is to show existence of a martingale X satisfying

$$d[X]_t = dt + F(X(t-))\,dX(t) \tag{3.16}$$

or, equivalently,

$$X(t)^2 - X(0)^2 - 2\int_0^t X(s-)\,dX(s) = t + \int_0^t F(X(s-))\,dX(s) \tag{3.17}$$

Of course, if X is standard Brownian motion, then (3.16) is satisfied for $F(x) = 0$. If X is a martingale with $|X(t)| = \sqrt{t}$, then, obviously from (3.17), (3.16) holds with $F(x) = -2x$. See Protter and Sharpe (1979) and Emery (1989) for a construction of such a martingale. For Azema's martingale (Protter (1990) §IV.6), $F(x) = -x$.

Following Meyer (1989), we define a sequence of discrete time martingales and show that the sequence is relatively compact and that the limit satisfies (3.16). Setting $\Delta Y_n(k) = Y_n(k+1) - Y_n(k)$ and assuming for simplicity that $Y_n(0) = 0$, the discrete time analogue of (3.16) becomes

$$\Delta Y_n(k)^2 = \tfrac{1}{n} + F(Y_n(k))\,\Delta Y_n(k) \tag{3.18}$$

Consequently,

$$\Delta Y_n(k) = \frac{F(Y_n(k)) \pm \sqrt{F(Y_n(k))^2 + \frac{4}{n}}}{2} \equiv \Delta_n^{\pm}(k) \tag{3.19}$$

and since we want Y_n to be a martingale, we must have

$$P\{\Delta Y_n(k) = \Delta_n^{+}(k)\} = 1 - P\{\Delta Y_n(k) = \Delta_n^{-}(k)\} = \frac{\Delta_n^{-}(k)}{\Delta_n^{-}(k) - \Delta_n^{+}(k)} \tag{3.20}$$

Define $X_n(t) = Y_n([nt])$. Note that $E[X_n(t)^2] = \frac{[nt]}{n}$ and more generally

(3.21) $$E[(X_n(t+h) - X_n(t))^2 | \mathcal{F}_t^{X_n}] = \frac{[n(t+h)]}{n} - \frac{[nt]}{n}$$

The relative compactness of $\{X_n\}$ (and hence for $\{(X_n, F \circ X_n)\}$) follows easily. (See, for example, Ethier and Kurtz (1986), Remark 3.8.7.) Since X_n satisfies

(3.22) $$X_n(t)^2 - X_n(0)^2 - 2\int_0^t X_n(s-)\,dX_n(s) = \frac{[nt]}{n} + \int_0^t F(X_n(s-))\,dX_n(s)$$

we see that any limit point of the sequence $\{X_n\}$ satisfies (3.17).

More generally, the above construction will give solutions of

(3.23) $$d[X]_t = dt + F(X,t-)\,dX(t)$$

for any $F: D_{\mathbb{R}}[0,\infty) \to D_{\mathbb{R}}[0,\infty)$ satisfying C5.4(ii) and C5.4(iii) below and $F(x,t) = F(x^t,t)$ for all $x \in D_{\mathbb{R}}[0,\infty)$ and $t \geq 0$ where $x^t = x(\cdot \wedge t)$.

<u>3.5 Example</u> (Neuhaus (1977)) Let $\xi_1, \xi_2, \ldots$ be i.i.d. uniform-[0,1] random variables, and let h be a measurable, symmetric function defined on $[0,1]\times[0,1]$ satisfying

(3.24) $$\int_0^1\int_0^1 h^2(x,y)\,dx\,dy < \infty$$

and

(3.25) $$\int_0^1 h(x,y)\,dx = \int_0^1 h(x,y)\,dy = 0$$

Define

(3.26) $$Z_n^h = \frac{1}{n} \sum_{1\leq i<j\leq n} h(\xi_i, \xi_j)$$

Then $\{Z_n^h\}$ is asymptotically Gaussian. To see that this is the case and to identify the limit, we follow a suggestion of Lajos Horvath and represent (3.26) in terms of the empirical distribution function F_n

(3.27) $$F_n(t) = \frac{1}{n} \sum_{i=1}^n \chi_{[\xi_i,\infty)}(t)$$

In terms of F_n, Z_n^h can be written

(3.28)
$$Z_n^h = n \iint_{s<t} h(s,t)\,dF_n(s)\,dF_n(t)$$

and defining $B_n(t) = \sqrt{n}\,(F_n(t) - t)$, the symmetry of h and (3.25) give

(3.29)
$$Z_n^h = \iint_{s<t} h(s,t)\,dB_n(s)\,dB_n(t)$$

If g satisfies the same conditions as h, then

(3.30)
$$E[(Z_n^h - Z_n^g)^2] = \frac{n(n-1)}{2n^2}\int_0^1\int_0^1 (h(x,y) - g(x,y))^2\,dx\,dy$$

Since any $h \in L^2([0,1]\times[0,1])$ can be approximated by smooth g, we may as well assume that h is continuously differentiable. Under this assumption we can write

(3.31)
$$X_n(t) = \int_0^t h(s,t)\,dB_n(s) = h(t,t)\,B_n(t) - \int_0^t h_s(s,t)B_n(s)\,ds$$

and, since $B_n \Rightarrow B$, the Brownian bridge, (see, for example, Billingsley (1968), §13 and §19, or Protter (1990) §V.6), the continuous mapping theorem implies that $X_n \Rightarrow X$ given by

(3.32)
$$X(t) = \int_0^t h(s,t)\,dB(s)$$

More precisely, $(X_n,B_n) \Rightarrow (X,B)$ in $D_{\mathbf{R}\times\mathbf{R}}[0,\infty)$.

The process B_n is a semimartingale with decomposition

(3.33)
$$B_n(t) = \sqrt{n}\,(F_n(t) - t) = \sqrt{n}\Big(F_n(t) - \int_0^t \frac{1-F_n(s)}{1-s}\,ds\Big) - \sqrt{n}\int_0^t \frac{F_n(s)-s}{1-s}\,ds$$
$$= M_n(t) - \int_0^t \frac{1}{1-s}B_n(s)\,ds$$

Note that $E[M_n(t)^2] = E[[M_n]_t] = t$. In fact, $[M_n]_t \to t$, implying, by the martingale central theorem, that $M_n \Rightarrow W$ and yielding, in the limit, the classical stochastic differential equation for B. For this decomposition we have

$$(3.34)\qquad E\left[T_t\left(\int_0^{\cdot}\frac{1}{1-s}B_n(s)\,ds\right)\right] = E\left[\int_0^t \frac{1}{1-s}|B_n(s)|\,ds\right] \le \int_0^t \frac{1}{1-s}\sqrt{E[B_n(s)^2]}\,ds = \int_0^t \sqrt{\frac{s}{1-s}}\,ds < \infty$$

for $t \le 1$. Consequently, the conditions of Theorem 2.2 are satisfied, and Z_n^h converges in distribution to

$$(3.35)\qquad Z^h = \int_0^1\int_0^t h(s,t)\,dB(s)\,dB(t)$$

For related results see Hall (1979). Rubin and Vitale (1980) and Dynkin and Mandelbaum (1983) consider more general symmetric statistics. Rubin and Vitale represent the limiting random variables as series of products of Hermite polynomials of Gaussian random variables. Dynkin and Mandelbaum represent the limits as multiple Wiener integrals. These higher order limit theorems can also be obtained by the techniques used above with the limiting random variables represented as multiple integrals of B. Filippova (1961) obtained limits represented as multiple integrals of Brownian bridge in special cases. □

3.6 Example (Duffie and Protter (1989)) Theorem 2.2 is useful in the derivation and justification of models in continuous time finance theory as limiting cases of discrete time models. For example, let the sequence of random variables $\xi_1^n, \xi_2^n, \ldots$ denote the periodic rate of return on a security with initial price S_0. After k periods the price of the security will be

$$(3.36)\qquad S_k^n = S_0^n \prod_{i=1}^{k} (1 + \xi_i^n)$$

Let $Y_n(t) = \sum_{i\le[nt]}\xi_i^n$ and $S_n(t) = S_{[nt]}^n$. Noting that $S_{k+1}^n - S_k^n = S_k^n\xi_k^n$, we can write

$$(3.37)\qquad S_n(t) = S_n(0) + \int_0^t S_n(s-)\,dY_n(s)$$

If θ_k^n units of the security are held during the $(k+1)$th period, the financial gain for the period is $\theta_k^n(S_{k+1}^n - S_k^n)$, and the cumulative gain up to time t can be written

$$(3.38)\qquad G_n(t) = \int_0^t \theta_n(s-)\,dS_n(s)$$

where $\theta_n(t) = \theta^n_{[nt]}$. Suppose that $\{Y_n\}$ satisfies C2.2(i) for some δ and that $(Y_n,\theta_n,S_n(0)) \Rightarrow (Y,\theta,S(0))$ (in $D_{\mathbb{R}^2}[0,\infty)\times\mathbb{R}$). Then the limiting equation

(3.39) $$S(t) = S(0) + \int_0^t S(s-)\,dY(s)$$

has a (locally) unique global solution, so by Theorem 5.4 below (see also Avram (1988)), $S_n \Rightarrow S$. (More precisely $(Y_n,\theta_n,S_n) \Rightarrow (Y,\theta,S)$.) It follows that $\{S_n\}$ also satisfies C2.2(i), so that $G_n \Rightarrow G$ given by

(3.40) $$G(t) = \int_0^t \theta(s-)\,dS(s)$$

The solution of (3.39) with $S(0) = 1$ is called the stochastic or Doléans-Dade exponential and is denoted $\mathcal{E}(X)$. The general solution is then given by $S = S(0)\mathcal{E}(X)$. (Protter (1990) §II.8.) □

3.7 Example For each n, let Y_n be an $\{\mathcal{F}^n_t\}$-semimartingale, and let $\{\tau^n_k\}$ be a sequence of $\{\mathcal{F}^n_t\}$-stopping times with $\tau^n_0 = 0$ and $\lim_{t\to\infty}\tau^n_k = \infty$. Define

(3.41) $$\tilde{Y}_n(t) = Y_n(\tau^n_k), \quad \tau^n_k \le t < \tau^n_{k+1}$$

Suppose that $Y_n \Rightarrow Y$ and that $\{Y_n\}$ satisfies C2.2(i) for some $\delta \in (0,\infty]$. (In particular this last statement holds if $Y_n = Y$ for all n.) Then $\{\tilde{Y}_n\}$ is relatively compact in the Skorohod topology and satisfies C2.2(i). If $\sup_k \tau^n_{k+1} - \tau^n_k \to 0$, $\tilde{Y}_n \Rightarrow Y$. Since the increments of $\tilde{Y}_n$ can be estimated in terms of the increments of Y_n, the relative compactness of $\{\tilde{Y}_n\}$ follows easily (see Theorem 3.7.2 of Ethier and Kurtz (1986)). The convergence assertion follows from Proposition 3.6.5 of Ethier and Kurtz (1986). To see that C2.2(i) is satisfied, define

(3.42) $$\tilde{A}^\delta_n(t) = A^\delta_n(\tau^n_k), \quad \tilde{M}^\delta_n(t) = M^\delta_n(\tau^n_k), \quad \tau^n_k \le t < \tau^n_{k+1}$$

If $\delta = \infty$, then $\tilde{Y}_n = \tilde{M}^\infty_n + \tilde{A}^\infty_n$, $E[[\tilde{M}^\infty_n]_{t\wedge\tau^\alpha_n}] \le E[[M^\infty_n]_{t\wedge\tau^\alpha_n}]$ and $E[T_{t\wedge\tau^\alpha_n}(\tilde{A}^\infty_n)] \le E[T_{t\wedge\tau^\alpha_n}(A^\infty_n)]$ so C2.2(i) holds. If $\delta < \infty$, then

(3.43) $$\tilde{Y}_n = J_\delta(\tilde{Y}_n) + \tilde{M}^\delta_n + \tilde{A}^\delta_n + (\tilde{J}_\delta(Y_n) - J_\delta(\tilde{Y}_n))$$

where $\tilde{J}_\delta(Y_n) = J_\delta(Y_n)(\tau_k^n)$ for $\tau_k^n \le t < \tau_{k+1}^n$. As before, $E[[\tilde{M}_n^\delta]_{t\wedge\tau_n^\alpha}] \le E[[M_n^\delta]_{t\wedge\tau_n^\alpha}]$, and we claim that there exist stopping times $\tilde{\tau}_n^\alpha$ satisfying $P\{\tilde{\tau}_n^\alpha \le \alpha\} \le \frac{1}{\alpha}$ and $\sup_n E[T_{t\wedge\tilde{\tau}_n^\alpha}(\tilde{A}_n^\delta + \tilde{J}_\delta(Y_n) - J_\delta(\tilde{Y}_n))] < \infty$. The relative compactness of $\{Y_n\}$ and $\{\tilde{Y}_n\}$ implies that $\{T_t(\tilde{J}_\delta(Y_n) - J_\delta(\tilde{Y}_n))\}$ is stochastically bounded for each t. For each $\alpha > 0$, select c_α such that $P\{T_\alpha(\tilde{J}_\delta(Y_n) - J_\delta(\tilde{Y}_n)) \ge c_\alpha\} \le \frac{1}{2\alpha}$, and define $\eta_n^\alpha = \inf\{t : T_t(\tilde{J}_\delta(Y_n) - J_\delta(\tilde{Y}_n)) \ge c_\alpha\}$ and $\tilde{\tau}_n^\alpha = \tau_n^{2\alpha}\wedge\eta_n^\alpha$. Then, noting that the magnitude of the discontinuities of $\tilde{M}_n^\delta + \tilde{A}_n^\delta + (\tilde{J}_\delta(Y_n) - J_\delta(\tilde{Y}_n))$ is at most δ,

$$
\begin{aligned}
(3.44)\qquad & E[T_{t\wedge\tilde{\tau}_n^\alpha}(\tilde{A}_n^\delta + \tilde{J}_\delta(Y_n) - J_\delta(\tilde{Y}_n))] \\
&\le E[T_{t\wedge\tau_n^{2\alpha}}(A_n^\delta)] + c_\alpha \\
&\qquad + E[|(\tilde{J}_\delta(Y_n) - J_\delta(\tilde{Y}_n))(t\wedge\tilde{\tau}_n^\alpha) - (\tilde{J}_\delta(Y_n) - J_\delta(\tilde{Y}_n))(t\wedge\tilde{\tau}_n^\alpha\text{-})|] \\
&\le E[T_{t\wedge\tau_n^{2\alpha}}(A_n^\delta)] + c_\alpha + \delta \\
&\qquad + E[|\tilde{M}_n^\delta(t\wedge\tilde{\tau}_n^\alpha) - \tilde{M}_n^\delta(t\wedge\tilde{\tau}_n^\alpha\text{-})| + |\tilde{A}_n^\delta(t\wedge\tilde{\tau}_n^\alpha) - \tilde{A}_n^\delta(t\wedge\tilde{\tau}_n^\alpha\text{-})|] \\
&\le 2E[T_{t\wedge\tau_n^{2\alpha}}(A_n^\delta)] + c_\alpha + \delta + \sqrt{E[[M_n^\delta]_{t\wedge\tau_n^{2\alpha}}]}
\end{aligned}
$$

and C2.2(i) follows. □

<u>4. Relative compactness and additional convergence results</u> The conditional variation on $[0,t]$ of a process X with respect to a filtration $\{\mathcal{F}_t\}$ is defined by $V_t(X) = \sup E[\sum_i |E[X(t_{i+1}) - X(t_i)|\mathcal{F}_{t_i}]|]$ where the supremum is over all partitions of $[0,t]$. (For vector-valued X we take $|x| = \sum |x_i|$.) For a stopping time τ, X^τ will denote the stopped process given by $X^\tau(t) = X(t\wedge\tau)$.

<u>4.1 Lemma</u> For $n = 1,2,\ldots$, let X_n and Y_n be $\{\mathcal{F}_t^n\}$-adapted, X_n in $D_{\mathbb{M}^{km}}[0,\infty)$ and Y_n in $D_{\mathbb{R}^m}[0,\infty)$. Assume the following condition

C4.1 For each $\alpha > 0$, there exist stopping times $\{\tau_n^\alpha\}$ with $P\{\tau_n^\alpha \leq \alpha\} \leq \frac{1}{\alpha}$ such that for each $t \geq 0$, $\sup_n E[|Y_n^{\tau_n^a}(t)|] < \infty$ and $\sup_n V_t(Y_n^{\tau_n^a}) < \infty$ (where the conditional variation for Y_n is with respect to $\{\mathcal{F}_t^n\}$).

Let $H_n(t) = \sup_{s\leq t}|X_n(s)|$, and suppose that $\{H_n(t)\}$ is stochastically bounded for each t. Define

$$Z_n(t) = \int_0^t X_n(s-)\,dY_n(s) \tag{4.1}$$

Then $\{Z_n\}$ satisfies C4.1, and there exist strictly increasing, $\{\mathcal{F}_t^n\}$-adapted processes C_n, with $C_n(0) = 0$, $C_n(t+h) - C_n(t) \geq h$ and $\{C_n(t)\}$ stochastically bounded for all $t, h \geq 0$, such that, defining $\gamma_n = C_n^{-1}$, $\hat{Y}_n(t) = Y_n(\gamma_n(t))$ and $\hat{Z}_n(t) = Z_n(\gamma_n(t))$, $\{(\hat{Y}_n,\hat{Z}_n,\gamma_n)\}$ is relatively compact in $D_{\mathbb{R}^m\times\mathbb{R}^k\times\mathbb{R}}[0,\infty)$.

<u>4.2 Remark</u> a) Note that $Z_n(t) = \hat{Z}_n(\gamma_n^{-1}(t))$.

b) Theorem 3.5 of Kurtz (1990) gives conditions on the sequence $\{C_n\}$ which imply relative compactness for $\{(Y_n,Z_n)\}$. This theorem is an extension of Theorem 2.3 of Jacod, Mémin, and Métivier (1983).

c) The result will also hold under the assumption that X_n is predictable and H_n is a right continuous, adapted, increasing process satisfying $|X_n(s)| \leq H_n(t)$ for $s \leq t$, with the usual extension of the stochastic integral to predictable integrands.

d) Let

(4.2) $$Y_n = \sum_{k=0}^{n-1} \frac{(-1)^k}{2n} \chi_{[1+k/n^2,1+(k+1)/n^2)}$$

and $X_n = -\operatorname{sign}(Y_n)$. Then the conditions of the lemma are satisfied and

(4.3) $$Z_n = \sum_{k=1}^{n-1} \frac{1}{n}\chi_{[1+k/n^2,\infty)} + \frac{1}{2n}\chi_{[1+1/n,\infty)}$$

The γ_n can be selected so that $\dot{\gamma}_n = \frac{1}{n}$ on the interval $[1,2)$ and $\dot{\gamma}_n = 1$ otherwise. The sequence $\{\hat{Z}_n\}$ then converges in $D_{\mathbb{R}^k}[0,\infty)$ to a continuous, piecewise linear function. Note that $\{Z_n\}$ does not converge in the Skorohod topology. (We thank J. Mémin and L. Slominski for bringing this example to our attention and pointing out a serious error in an earlier version of this paper.)

Proof Each Y_n has a unique decomposition $Y_n = M_n + B_n$, where M_n is a local martingale and B_n is a predictable finite variation process satisfying $E[T_{t\wedge\tau_n^\alpha}(B_n)] \le V_t(Y_n^{\tau_n^a})$. (See Kurtz (1990), Proposition 5.1.) If we write

(4.4) $$Z_n(t) = \int_0^t X_n(s-)\,dM_n(s) + \int_0^t X_n(s-)\,dB_n(s)$$

the first term is a local martingale and the total variation of the second term up to time t is bounded by $H_n(t-)T_t(B_n)$. Consequently, if σ_n is a stopping time so that the first term stopped is a martingale, $H_n(\sigma_n-) \le c$, and $\sigma_n \le \tau_n^\alpha$, then $V_t(Z_n^{\sigma_n}) \le cV_t(Y_n^{\tau_n^a})$. But for any $\beta > 0$, c, α, and σ_n can be selected so that $P\{\sigma_n \le \beta\} \le \frac{1}{\beta}$, and it follows that $\{Z_n\}$ satisfies C4.1. This in turn implies that $\{(Y_n,Z_n)\}$ satisfies C4.1. Corollary 1.3 of Kurtz (1990) then gives the other conclusions. □

4.3 Proposition Let $\{(U_n,Y_n)\}$ be relatively compact (in the sense of convergence in distribution) in $D_{\mathbb{R}^k\times\mathbb{R}^m}[0,\infty)$ with (U_n,Y_n) adapted to $\{\mathcal{F}_t^n\}$, and $\{Y_n\}$ **satisfying** C2.2(i) for some $\delta > 0$. Suppose that X_n has sample paths in $D_{\mathbb{M}^{km}}[0,\infty)$ and is adapted to $\{\mathcal{F}_t^n\}$. Define

(4.5) $$Z_n(t) = U_n(t) + \int_0^t X_n(s-)\,dY_n(s)$$

Suppose there exist strictly increasing, $\{\mathcal{F}_t^n\}$-adapted processes C_n, with $C_n(t+h) - C_n(t) \geq h$ and $\{C_n(t)\}$ stochastically bounded for all $t, h \geq 0$ such that, defining $\gamma_n = C_n^{-1}$, $\hat{U}_n(t) = U_n(\gamma_n(t))$ etc., $\{(\hat{U}_n,\hat{X}_n,\hat{Y}_n,\gamma_n)\}$ is relatively compact in $D_{\mathbf{R}^k\times\mathbf{M}^{km}\times\mathbf{R}^m\times\mathbf{R}}[0,\infty)$. Then $\{(Z_n,U_n,Y_n,)\}$ is relatively compact in $D_{\mathbf{R}^k\times\mathbf{R}^k\times\mathbf{R}^m}[0,\infty)$.

Proof For technical reasons, we extend the definition of the processes to the time interval $[-1,\infty)$ by setting $U_n(t) = U_n(0)$, $X_n(t) = X_n(0)$, $Y_n(t) = Y_n(0)$, and $C_n(t) = t$ for $-1 \leq t < 0$. These definitions ensure that $\{(\hat{U}_n,\hat{X}_n,\hat{Y}_n,\gamma_n)\}$ is relatively compact in $D_{\mathbf{R}^k\times\mathbf{M}^{km}\times\mathbf{R}^m\times\mathbf{R}}[-1,\infty)$.

The fact that $\{Y_n\}$ satisfies C2.2(i) implies that $\{\hat{Y}_n\}$ satisfies C2.2(i). Consequently, selecting a convergent subsequence from $\{(\hat{U}_n,\hat{X}_n,\hat{Y}_n,\gamma_n)\}$ with limit $(\hat{U},\hat{X},\hat{Y},\gamma)$, by Theorem 2.2, $\{(\hat{Z}_n,\hat{U}_n,\hat{X}_n,\hat{Y}_n,\gamma_n)\}$ converges to $(\hat{Z},\hat{U},\hat{X},\hat{Y},\gamma)$ where

$$\hat{Z}(t) = \hat{U}(t) + \int_{-1}^{t} \hat{X}(s\text{-})\,d\hat{Y}(s) = \hat{U}(t) + \int_{0}^{t} \hat{X}(s\text{-})\,d\hat{Y}(s) \tag{4.6}$$

We may assume that (U_n,Y_n) converges along the same subsequence, and the limit must be $(U,Y) = (\hat{U}\circ\gamma^{-1},\hat{Y}\circ\gamma^{-1})$ where $\gamma^{-1}(t) \equiv \inf\{u:\gamma(u) > t\}$. (Note that γ^{-1} is defined so that it is right continuous, and that the conditions on C_n imply $\gamma^{-1}(t) = t$ for $t \leq 0$.) Lemma 2.3 of Kurtz (1990) (with the obvious modification for the time interval $[-1,\infty)$) implies that $(U_n,Y_n) \Rightarrow (U,Y)$ in the Skorohod topology if and only if on any interval on which γ is constant, $(\hat{U},\hat{Y})$ is constant except for at most one jump. But on any interval on which $(\hat{U},\hat{Y})$ is constant, $\hat{Z}$ is constant, and $\hat{Z}$ jumps only when $\hat{U}$ or $\hat{Y}$ jumps. Consequently, on any interval on which γ is constant, $(\hat{Z},\hat{U},\hat{Y})$ is constant except for at most one jump. Applying the cited lemma again, we have that, along the subsequence, $(Z_n,U_n,Y_n) \Rightarrow (Z,U,Y)$ in the Skorohod topology on $D_{\mathbf{R}^k\times\mathbf{M}^{km}\times\mathbf{R}^m\times\mathbf{R}}[-1,\infty)$. But (Z,U,Y) must be continuous at 0, so the convergence holds in $D_{\mathbf{R}^k\times\mathbf{M}^{km}\times\mathbf{R}^m\times\mathbf{R}}[0,\infty)$ as well, and the proposition follows. □

4.4 Corollary Let $\{(U_n,X_n,Y_n)\}$ have sample path in $D_{\mathbf{R}^k\times\mathbf{M}^{km}\times\mathbf{R}^m}[0,\infty)$ and be adapted to $\{\mathcal{F}_t^n\}$, and let Z_n be given by (4.5). Suppose that $\{(U_n,Y_n)\}$ is relatively compact in $D_{\mathbf{R}^k\times\mathbf{R}^m}[0,\infty)$, that $\{Y_n\}$ satisfies C2.2(i) for some $\delta > 0$, and that $\{X_n\}$ satisfies C4.1. Then $\{(Z_n,U_n,Y_n)\}$ is relatively compact in $D_{\mathbf{R}^k\times\mathbf{R}^k\times\mathbf{R}^m}[0,\infty)$.

Proof The relative compactness of $\{(U_n,Y_n,\int X_n\,dJ_\delta(Y_n))\}$ is immediate. Since the stochastic integral on the right of (4.5) has a discontinuity only when Y_n has a discontinuity, and $\{(U_n,Y_n)\}$ is relatively compact, the proposition will follow if we show that $\{\int X_n\,dY_n^\delta\}$ is relatively compact. (See, for example, Kurtz (1990), Lemma 2.2).

C2.2(i) implies C4.1 for $\{Y_n^\delta\}$. Consequently, Corollary 1.3 of Kurtz (1990) implies the existence of strictly increasing, $\{\mathcal{F}_t^n\}$-adapted processes C_n, with $C_n(0) = 0$, $C_n(t+h) - C_n(t) \geq h$ and $\{C_n(t)\}$ stochastically bounded for all $t, h \geq 0$, such that, defining $\gamma_n = C_n^{-1}$, $\hat{Y}_n^\delta(t) = Y_n^\delta(\gamma_n(t))$ and $\hat{X}_n(t) = X_n(\gamma_n(t))$, $\{(\hat{Y}_n^\delta,\hat{X}_n,\gamma_n)\}$ is relatively compact in $D_{\mathbb{R}^m\times\mathbb{R}^k\times\mathbb{R}}[0,\infty)$. Defining

$$V_n(t) = \int_0^t X_n(s-)\,dY_n^\delta(s) \tag{4.7}$$

Proposition 4.3 implies $\{(Y_n^\delta,V_n)\}$ is relatively compact in $D_{\mathbb{R}^m\times\mathbb{R}^k}[0,\infty)$. □

4.5 Corollary Suppose $\{(U_n,Y_n,X_n)\}$ is relatively compact in $D_{\mathbb{R}^k\times\mathbb{R}^m}[0,\infty)\times D_{M^{km}}[0,\infty)$, $\{Y_n\}$ satisfies C2.2(i) for some $\delta > 0$, and Z_n is given by (4.5). Then $\{(U_n,Y_n,Z_n)\}$ is relatively compact in $D_{\mathbb{R}^k\times\mathbb{R}^m\times\mathbb{R}^k}[0,\infty)$.

Proof Let $W_n \equiv (U_n,Y_n,X_n)$. The idea of the proof is to define a positive function $h(r,s)$ which is nondecreasing in r and nonincreasing in s such that

$$C_n(t) = t + \sum_{s\leq t} h(|W_n(s) - W_n(s-)|,s) \tag{4.8}$$

satisfies the hypotheses of Proposition 4.3. Note that C_n is designed so that the successive discontinuities of $W_n\circ C_n^{-1}$ are separated by a deterministic function of the size of the first discontinuity. Lemma 2.2 of Kurtz (1990) then gives the relative compactness. The difficulty arises in ensuring that $\{C_n(t)\}$ is stochastically bounded for each t. For $k = 1,2,\ldots$, let $N_n^k(t)$ be the number of discontinuities of W_n before time t satisfying $\frac{1}{k} \leq |W_n(s) - W_n(s-)| < \frac{1}{k-1}$. The relative compactness of $\{W_n\}$ in $D_{\mathbb{R}^k\times\mathbb{R}^m}[0,\infty)\times D_{M^{km}}[0,\infty)$ ensures that $\{N_n^k(t)\}$ is stochastically bounded for each t and that $\lim_{s\to 0}\sup_n P\{N_n^k(s) > 0\} = 0$. Consequently, there exist $a_k(t) > 0$ independent of n such that

(4.9) $$\sup_n P\{a_k(t)N_n^k(t) > \frac{1}{2^k}\} \le \frac{1}{2^k}$$

Without loss of generality, we can take $a_k(t)$ to be nonincreasing in t and k. Define

(4.10) $$C_n(t) = t + \sum_{m=0}^{\infty}\sum_{k=1}^{\infty} a_k(m+1)(N_n^k((m+1)\wedge t) - N_n^k(m\wedge t))$$

Note that the first sum in (4.10) is in fact finite and that (4.9) implies by Borel-Cantelli that only finitely many terms in the second sum exceed $\frac{1}{2^k}$. To check the stochastic boundedness of $\{C_n(t)\}$ it is enough to check the stochastic boundedness of

(4.11) $$K_n^m \equiv \sum_{k=1}^{\infty} a_k(m)N_n^k(m)$$

for each m. We have

(4.12) $$P\{K_n^m > a+1\} \le \sum_{k=1}^{\ell} P\{a_k(m)N_n^k(m) > \frac{a}{\ell}\} + P\{\sum_{k=\ell+1}^{\infty} a_k(m)N_n^k(m) > \frac{1}{2^\ell}\}$$

$$\le \sum_{k=1}^{\ell} P\{a_k(m)N_n^k(m) > \frac{a}{\ell}\} + \frac{1}{2^\ell}$$

and the stochastic boundedness of $\{K_n^m\}$ follows easily from the stochastic boundedness of the $N_n^k(m)$. □

These relative compactness results lead to the problem of identifying the limit under more general assumptions on the limiting behavior of $\{X_n\}$ than in Theorem 2.2. First assume that $(X_n,Y_n) \Rightarrow (X,Y)$ in $D_{\mathbb{M}^{km}}[0,\infty)\times D_{\mathbb{R}^m}[0,\infty)$ (rather than in $D_{\mathbb{M}^{km}\times\mathbb{R}^m}[0,\infty)$) and that $\{Y_n\}$ satisfies C2.2(i). For all but countably many $\epsilon > 0$, $(X_n(\cdot-\epsilon),Y_n) \Rightarrow (X(\cdot-\epsilon),Y)$ in $D_{\mathbb{M}^{km}\times\mathbb{R}^m}[0,\infty)$). Consequently, for each such ϵ,

(4.13) $$\int_0^t X_n(s-\epsilon-)\,dY_n(s) \Rightarrow \int_0^t X(s-\epsilon-)\,dY(s)$$

and hence there exists a sequence $\epsilon_n \to 0$ slowly enough such that

(4.14) $$\int_0^t X_n(s-\epsilon_n-)\,dY_n(s) \Rightarrow \int_0^t X(s-)\,dY(s)$$

Noting that $\{\int X_n\,dY_n\}$ is relatively compact by Corollary 4.5, assume that $\int X_n\,dY_n \Rightarrow Z$. Consequently,

(4.15) $$\int_0^{\cdot} (X_n(s-) - X_n(s-\epsilon_n-))\,dY_n(s) \Rightarrow Z(\cdot) - \int_0^{\cdot} X(s-)\,dY(s)$$

Note that the sequence on the left in (4.15) is relatively compact by an argument similar to that used in the proof of Corollary 4.5.

Let $J_\delta(X_n)$ denote the M^{km}-valued process whose ijth component is $J_\delta(X_n^{ij})$ where X_n^{ij} is the ijth component of X_n, and let $X_n^\delta = X_n - J_\delta(X_n)$. Let $V_n^\delta(t) = \sup_{s\le t}|X_n^\delta(s) - X_n^\delta(s-\epsilon_n)|$. Then $V_n^\delta \Rightarrow V^\delta$ given by $V^\delta(t) = \sup_{s\le t}|X^\delta(s) - X^\delta(s-)| \le \sqrt{km}\,\delta$. By the same type of estimate as in (2.7), to identify the right side of (4.15) it is enough to identify the limit of

(4.16) $$U_n^\delta(t) = \int_0^t \Big(J_\delta(X_n)(s-) - J_\delta(X_n)(s-\epsilon_n-)\Big)dY_n(s)$$

(along a subsequence if necessary) and then to let $\delta \to 0$. Let $\{\tau_{in}^\delta\}$ denote the times of discontinuity of $J_\delta(X_n)$ with $\tau_{0n}^\delta = 0$. Note that $\{\tau_{in}^\delta\}$ are just the times when at least one component of X_n has a discontinuity larger than δ. Then U_n^δ can be written

(4.17) $$\sum_{\tau_{in}^\delta \le t} (Y_n(\tau_{in}^\delta+\epsilon_n) - Y_n(\tau_{in}^\delta))(J_\delta(X_n)(\tau_{in}^\delta) - J_\delta(X_n)(\tau_{in}^\delta-))$$

and any limit point U^δ of $\{U_n^\delta\}$ satisfies

(4.18) $$U^\delta(t) = \sum_{\beta_i^\delta \le t} (J_\delta(X)(\beta_i^\delta) - J_\delta(X)(\beta_i^\delta-))(Y(\beta_i^\delta) - Y(\beta_i^\delta-))$$

where $\{\beta_i^\delta\}$ is some <u>subset</u> of the times at which some component of X has a discontinuity larger that δ. Letting $\delta \to 0$, we see that

(4.19) $$U(t) \equiv Z(t) - \int_0^t X(s-)\,dY(s) = \sum_{\beta_i^\delta \le t} (Y(\beta_i) - Y(\beta_i-))(X(\beta_i) - X(\beta_i-))$$

where $\{\beta_i\}$ is some subset of the times at which both Y and X have discontinuities. From (4.17) it is clear that $\{\beta_i\}$ is empty unless some discontinuities of Y_n "coalesce" with discontinuities of X_n from above. The following theorem gives conditions under which no such coalescence occurs.

4.6 Theorem For each n, let (X_n,Y_n) be an $\{\mathcal{F}_t^n\}$-adapted process with sample paths in $D_{\mathbb{M}^{km}\times\mathbb{R}^m}[0,\infty)$, and let Y_n be an $\{\mathcal{F}_t^n\}$-semimartingale. Suppose that for some $0 < \delta \leq \infty$, C2.2(i) holds and that for all $T > 0$ and $\eta > 0$ there exist random variables $\{\gamma_n^T(\eta)\}$ such that

$$(4.20) \qquad E[1\wedge|Y_n(t+u) - Y_n(t)||\mathcal{F}_t^n] \leq E[\gamma_n^T(\eta)|\mathcal{F}_t^n], \qquad 0\leq u\leq\eta,\ 0\leq t\leq T$$

and $\lim_{\eta\to 0}\overline{\lim}_{n\to\infty} E[\gamma_n^T(\eta)] = 0.$

If $(X_n,Y_n) \Rightarrow (X,Y)$ in $D_{\mathbb{M}^{km}}[0,\infty)\times D_{\mathbb{R}^m}[0,\infty)$, then Y is a semimartingale with respect to a filtration to which X and Y are adapted, and $(X_n,Y_n,\int X_n\,dY_n) \Rightarrow (X,Y,\int X\,dY)$ in $D_{\mathbb{M}^{km}\times\mathbb{R}^m\times\mathbb{R}^k}[0,\infty)$. If $(X_n,Y_n) \to (X,Y)$ in probability, then the triple converges in probability.

4.7 Remark See Ethier and Kurtz (1986) Theorem 3.8.6 and Remark 3.8.7 for the connection of (4.20) to conditions for the relative compactness of $\{Y_n\}$. These conditions imply a type of uniform quasi-left continuity on the sequence $\{Y_n\}$. Consequently, this theorem is related to Theorem 5.1 of Jakubowski, Mémin, and Pages (1989).

Proof We need only show that $U \equiv 0$ in (4.19). The inequality in (4.20) holds with t replaced by a stopping time. Consequently we have (with reference to (4.17)) for $\epsilon_n \leq \eta$

$$(4.21) \quad E\left[\sum_{i=1}^m 1\wedge\left|(Y_n(\tau_{in}^\delta\wedge T+\epsilon_n) - Y_n(\tau_{in}^\delta\wedge T))\right|1\wedge\left|(J_\delta(X_n)(\tau_{in}^\delta\wedge T) - J_\delta(X_n)(\tau_{in}^\delta\wedge T-))\right|\right]$$

$$\leq E\left[\sum_{i=1}^m \gamma_n^T(\eta)1\wedge\left|(J_\delta(X_n)(\tau_{in}^\delta\wedge T) - J_\delta(X_n)(\tau_{in}^\delta\wedge T-))\right|\right]$$

$$\leq mE[\gamma_n^T(\eta)]$$

Since the number of discontinuities of $J_\delta(X_n)$ in any finite time interval is stochastically bounded in n, it follows that $U^\delta(t) = 0$ for each $t > 0$. Consequently, $U \equiv 0$ and the theorem follows.

Noting that if a sequence $\{U_n\}$ is defined on a single sample space and $U_n \Rightarrow 0$, then $U_n \to 0$ in probability, we see that convergence in distribution can be replaced by convergence in probability in the statement of the theorem. $\square$

In the next theorem we weaken the assumption that the integrands converge in the Skorohod topology at the cost of adding the requirement that the limiting integrator be continuous. $M_E[0,\infty)$ denotes the space of (equivalence classes of) measurable E-valued functions topologized by convergence in measure.

4.7 Theorem For each n, let (X_n,Y_n) be an $\{\mathcal{F}_t^n\}$-adapted process with sample paths in $D_{\mathbb{M}^{km}\times\mathbb{R}^m}[0,\infty)$. Suppose that $\{Y_n\}$ satisfies C2.2(i) for some $0 < \delta \le \infty$, and that $\{X_n\}$ satisfies C4.1. If $(X_n,Y_n) \Rightarrow (X,Y)$ in $M_{\mathbb{M}^{km}}[0,\infty)\times D_{\mathbb{R}^m}[0,\infty)$ and Y is continuous, then X has a version with sample paths in $D_{\mathbb{M}^{km}}[0,\infty)$, Y is a semimartingale with respect to a filtration to which X and Y are adapted, and $(X_n,Y_n,\int X_n\,dY_n) \Rightarrow (X,Y,\int X\,dY)$ in $M_{\mathbb{M}^{km}}[0,\infty)\times D_{\mathbb{R}^m\times\mathbb{R}^k}[0,\infty)$. If $(X_n,Y_n) \to (X,Y)$ in $M_{\mathbb{M}^{km}}[0,\infty)\times D_{\mathbb{R}^m}[0,\infty)$ in probability, then the triple converges in probability.

Proof Let C_n, γ_n, $\hat{X}_n$, and $\hat{Y}_n$ be as in Corollary 4.4 and Proposition 4.3, and set $Z_n = \int X_n\,dY_n$ and $\hat{Z}_n = Z_n\circ\gamma_n = \int \hat{X}_n\,d\hat{Y}_n$. Then $\{(X_n,Y_n,Z_n,\hat{X}_n,\hat{Y}_n,\hat{Z}_n,\gamma_n)\}$ is relatively compact in $M_{\mathbb{M}^{km}}[0,\infty)\times D_{\mathbb{R}^m\times\mathbb{R}^k}[0,\infty)\times D_{\mathbb{M}^{km}\times\mathbb{R}^m\times\mathbb{R}^k\times\mathbb{R}}[0,\infty)$. If $(X,Y,Z,\hat{X},\hat{Y},\hat{Z},\gamma)$ is a limit point, then $X = \hat{X}\circ\gamma^{-1}$, $Y = \hat{Y}\circ\gamma^{-1}$, and $Z = \hat{Z}\circ\gamma^{-1}$. Since $\hat{Y}$ is continuous and $\{Y_n\}$ converges in the Skorohod topology, $\hat{Y}$ must be constant on any interval on which γ is constant, which implies

$$Z(t) = \hat{Z}\circ\gamma^{-1}(t) = \int_0^t \hat{X}\circ\gamma^{-1}(s-)\,d\hat{Y}\circ\gamma^{-1}(s) = \int_0^t X(s-)\,dY(s) \tag{4.22}$$

and the theorem follows. $\square$

The above theorem still is not optimal even in the case of continuous integrands. For example, if each Y_n is a standard Brownian motion and $(X_n,Y_n) \Rightarrow (X,Y)$ in $L^2_{\mathbb{R}}[0,\infty)\times D_{\mathbb{R}}[0,\infty)$, then $\int X_n\,dY_n \Rightarrow \int X\,dY$. The following theorem comes close to covering this situation at the cost of placing strong conditions on the relationship between X_n and Y_n. Of course, other approximations of X_n could be used in place of X_n^h defined below.

4.8 Theorem Let $Y_n = M_n + A_n + Z_n$, where $\{(M_n,A_n,Z_n)\}$ satisfies the conditions of Theorem 2.7. Let $H_n(t) = \sup_{s\le t}|X_n(s)|$, and suppose that $\{H_n(t)\}$ is stochastically bounded for each t. Define X_n^h by

(4.23)
$$X_n^h(t) = h^{-1}\int_{t-h}^{t} X_n(s)\,ds$$

Suppose that for each $t > 0$ and $\epsilon > 0$

(4.24)
$$\lim_{h\to 0}\overline{\lim_{n\to\infty}} P\left\{\int_0^t |X_n^h(s-)-X_n(s-)|^2 d[M_n]_s + \int_0^t |X_n^h(s-)-X_n(s-)|\,d(T_s(A_n)+T_s(Z_n)) \ge \epsilon\right\} = 0$$

If $(X_n,Y_n,Z_n) \Rightarrow (X,Y,Z)$ in $M_{\mathbb{M}^{km}}[0,\infty)\times D_{\mathbb{R}^m\times\mathbb{R}^m}[0,\infty)$, then

(4.25)
$$U(t) \equiv \lim_{h\to 0}\int_0^t X^h\,dY$$

exists, and $(X_n,Y_n,\int X_n\,dY_n) \Rightarrow (X,Y,U)$ in $M_{\mathbb{M}^{km}}[0,\infty)\times D_{\mathbb{R}^m\times\mathbb{R}^k}[0,\infty)$. If $(X_n,Y_n,Z_n) \to (X,Y,Z)$ in $M_{\mathbb{M}^{km}}[0,\infty)\times D_{\mathbb{R}^m\times\mathbb{R}^m}[0,\infty)$ in probability, then $(X_n,Y_n,\int X_n\,dY_n) \to (X,Y,U)$ converges in probability.

Proof Since X_n^h is locally Lipschitz, the conditions on H_n ensure that $(X_n^h,Y_n,Z_n) \Rightarrow (X^h,Y,Z)$ in $D_{\mathbb{M}^{km}\times\mathbb{R}^m\times\mathbb{R}^m}[0,\infty)$ and hence that $\int X_n^h\,dY_n \Rightarrow \int X^h\,dY$. Consequently, estimating as in (2.7), (4.24) implies the result. □

5. Stochastic differential equations In this section we generalize results of Slomiński (1989) concerning convergence of sequences of solutions of stochastic differential equations. (See also Hoffman (1989) for results assuming the limiting semimartingale is continuous.) Note that Slomiński also considers Stratonovich equations. Avram (1988) considered the special case of stochastic exponentials, that is solutions of equations of the form ($k = m = 1$)

$$X(t) = 1 + \int_0^t X(s-)\,dY(s) \tag{5.1}$$

For $n = 1,2,\ldots$ let $F_n:D_{\mathbb{R}^k}[0,\infty) \to D_{\mathbb{M}^{km}}[0,\infty)$, let U_n and Y_n be processes with sample paths in $D_{\mathbb{R}^k}[0,\infty)$ and $D_{\mathbb{R}^m}[0,\infty)$ respectively, adapted to a filtration $\{\mathcal{F}_t^n\}$. Suppose Y_n is a semimartingale and that F_n is nonanticipating in the sense that $F_n(x,t) = F_n(x^t,t)$ for all $t \geq 0$ and $x \in D_{\mathbb{R}^k}[0,\infty)$, where $x^t(\cdot) = x(\cdot \wedge t)$. Let X_n be adapted to $\{\mathcal{F}_t^n\}$ and satisfy

$$X_n(t) = U_n(t) + \int_0^t F_n(X_n,s-)\,dY_n(s) \tag{5.2}$$

In order to apply Theorem 2.2 to the study of the weak convergence of solutions of this sequence of equations to the solution of a limiting equation

$$X(t) = U(t) + \int_0^t F(X,s-)\,dY(s) \tag{5.3}$$

we need conditions under which weak convergence of the pair $(X_n,Y_n) \Rightarrow (X,Y)$ implies $(Y_n,F_n(X_n)) \Rightarrow (Y,F(X))$. We could, of course, simply assume that $(x_n,y_n) \to (x,y)$ in $D_{\mathbb{R}^k\times\mathbb{R}^m}[0,\infty)$ implies $(x_n,y_n,F_n(x_n)) \to (x,y,F(x))$ in $D_{\mathbb{R}^k\times\mathbb{R}^m\times\mathbb{M}^{km}}[0,\infty)$, and under that assumption we have the following proposition.

5.1 Proposition Suppose that (U_n,X_n,Y_n) satisfies (5.2), that $\{(U_n,X_n,Y_n)\}$ is relatively compact in $D_{\mathbb{R}^k\times\mathbb{R}^k\times\mathbb{R}^m}[0,\infty)$, that $(U_n,Y_n) \Rightarrow (U,Y)$, and that $\{Y_n\}$ satisfies C2.2(i) for some $0 < \delta \leq \infty$. Assume that $\{F_n\}$ and F satisfy

C5.1 If $(x_n,y_n) \to (x,y)$ in the Skorohod topology, then $(x_n,y_n,F_n(x_n)) \to (x,y,F(x))$ in the Skorohod topology.

Then any limit point of the sequence $\{X_n\}$ satisfies (5.3).

Proof First note that if a subsequence of $\{X_n\}$ converges in distribution, then along a further subsequence the triple will converge in distribution to a process (U,X,Y). Theorem 2.2 then implies that (5.3) is satisfied. □

The following lemma, a generalization of Lemma 2.1, shows that the assumption on the sequence $\{F_n\}$ is valid for many interesting examples. Let Λ^1 be the subset of absolutely continuous functions in Λ for which $\gamma(\lambda) \equiv \|\ln\dot{\lambda}\|_\infty$ is finite.

5.2 Lemma Suppose that $\{F_n\}$ and F satisfy the following conditions:

C5.2(i) For each compact subset $\mathcal{K} \subset D_{\mathbb{R}^k}[0,\infty)$ and $t > 0$, $\sup_{x\in\mathcal{K}}\sup_{s\le t}|F_n(x,s) - F(x,s)| \to 0$.

C5.2(ii) For $\{x_n\}$ and x in $D_{\mathbb{R}^k}[0,\infty)$ and each $t > 0$, $\sup_{s\le t}|x_n(s) - x(s)| \to 0$ implies $\sup_{s\le t}|F(x_n,s) - F(x,s)| \to 0$.

C5.2(iii) For each compact subset $\mathcal{K} \subset D_{\mathbb{R}^k}[0,\infty)$ and $t > 0$, there exists a continuous function $\omega:[0,\infty) \to [0,\infty)$ with $\omega(0) = 0$ such that for all $\lambda \in \Lambda^1$, $\sup_{x\in\mathcal{K}}\sup_{s\le t}|F(x\circ\lambda,s) - F(x,\lambda(s))| \le \omega(\gamma(\lambda))$.

Then $(x_n,y_n) \to (x,y)$ in the Skorohod topology implies $(x_n,y_n,F_n(x_n)) \to (x,y,F(x))$ in the Skorohod topology.

Proof If $(x_n,y_n) \to (x,y)$ in the Skorohod topology, then there exist $\lambda_n \in \Lambda^1$ such that $\gamma(\lambda_n) \to 0$ and $(x_n\circ\lambda_n, y_n\circ\lambda_n) \to (x,y)$ uniformly on bounded time intervals. Consequently,

(5.4) $F_n(x_n,\lambda_n(s)) - F(x,s) =$

$$F_n(x_n,\lambda_n(s)) - F(x_n,\lambda_n(s)) + F(x_n,\lambda_n(s)) - F(x_n\circ\lambda_n,s) + F(x_n\circ\lambda_n,s) - F(x,s)$$

goes to zero uniformly in s on bounded intervals. □

5.3 Examples Let $g:\mathbb{R}^k \times [0,\infty) \to \mathbb{M}^{km}$ and $h:[0,\infty) \to [0,\infty)$ be continuous. The following functions satisfy C5.2(ii) and C5.2(iii).

a) $F(x,t) = g(x(t),t)$

b) $F(x,t) = \int_0^t h(t-s)g(x(s),s)\,ds$

For $k = m = 1$

c) $F(x,t) = \sup_{s \le t} h(t-s)g(x(s),s)$

d) $F(x,t) = \sup_{s \le t} h(t-s)g(x(s)-x(s-),s)$

One shortcoming of Proposition 5.1 is the apriori assumption that the sequence of solutions is relatively compact. (See Theorem 2.3 of Jacod, Mémin, and Métivier (1983) for general conditions on $\{Y_n\}$ under which the desired relative compactness will hold.) We can avoid this assumption by localizing the result and applying Proposition 4.3. We say that (X,τ) is a local solution of (5.3) if there exists a filtration $\{\mathcal{F}_t\}$ to which X, U, and Y are adapted, Y is an $\{\mathcal{F}_t\}$-semimartingale, τ is an $\{\mathcal{F}_t\}$-stopping time, and

(5.5) $$X(t\wedge\tau) = U(t\wedge\tau) + \int_0^{t\wedge\tau} F(X,s-)\,dY(s)$$

We say that strong local uniqueness holds for (5.3) if any two local solutions (X_1,τ_1), (X_2,τ_2) satisfy $X_1(t) = X_2(t)$, $t \le \tau_1\wedge\tau_2$, a.s. To define a notion of weak local uniqueness (that is, uniqueness of distributions), we need to require that the stopping time associated with the solution be a measurable function of the solution. We say that $(\hat{U},\hat{Y},\hat{X},\hat{\tau})$ is a weak local solution of (5.3) if $(\hat{U},\hat{Y})$ is a version of (U,Y) and (5.5) holds with (U,Y,X,τ) replaced by $(\hat{U},\hat{Y},\hat{X},\hat{\tau})$. We say that weak local uniqueness holds for (5.3) if for any two weak local solutions (U_1,Y_1,X_1,τ_1) and (U_2,Y_2,X_2,τ_2) with $\tau_1 = h_1(X_1)$ and $\tau_2 = h_2(X_2)$ for measurable functions h_1, h_2 on $D_{\mathbb{R}^k}[0,\infty)$, $(X_1,h_1\wedge h_2(X_1))$ and $(X_2,h_1\wedge h_2(X_2))$ have the same distribution. See Protter (1990), Chapter V, for sufficient conditions for uniqueness.

In order to apply Proposition 4.3, we need assumptions on the properties of $F_n(x)$ and $F(x)$ under transformations of the time scale. Let $T_1[0,\infty)$ denote the collection of nondecreasing mappings λ of $[0,\infty)$ onto $[0,\infty)$ (in particular $\lambda(0) = 0$) such that $\lambda(t+h) - \lambda(t) \le h$ for all $t, h \ge 0$. Let ι denote the identity map $\iota(s) = s$. We will assume that there exist mappings G_n, $G:D_{\mathbb{R}^k}[0,\infty)\times T_1[0,\infty) \to D_{M^{km}}[0,\infty)$ such that $F_n(x)\circ\lambda = G_n(x\circ\lambda,\lambda)$ and $F(x)\circ\lambda = G(x\circ\lambda,\lambda)$ for $(x,\lambda) \in D_{\mathbb{R}^k}[0,\infty)\times T_1[0,\infty)$. We need the following strengthening of C5.2

C5.4(i) For each compact subset $\mathcal{K} \subset D_{\mathbb{R}^k}[0,\infty)\times T_1[0,\infty)$ and $t > 0$, $\sup_{(x,\lambda)\in\mathcal{K}} \sup_{s\le t} |G_n(x,\lambda,s) - G(x,\lambda,s)| \to 0$.

C5.4(ii) For $\{(x_n,\lambda_n)\} \in D_{\mathbb{R}^k}[0,\infty)\times T_1[0,\infty)$, $\sup_{s\le t}|x_n(s) - x(s)| \to 0$ and $\sup_{s\le t}|\lambda_n(s) - \lambda(s)| \to 0$ for each $t > 0$ implies $\sup_{s\le t}|G(x_n,\lambda_n,s) - G(x,\lambda,s)| \to 0$.

We note that each of the examples in 5.3 has a representation in terms of a G satisfying C5.4(ii) and that C5.4 implies C5.2.

5.4 Theorem Suppose that (U_n,X_n,Y_n) satisfies (5.2), $(U_n,Y_n) \Rightarrow (U,Y)$ in the Skorohod topology and that $\{Y_n\}$ satisfies C2.2(i) for some $0 < \delta \le \infty$. Assume that $\{F_n\}$ and F have representations in terms of $\{G_n\}$ and G satisfying C5.4. For $b > 0$, define $\eta_n^b = \inf\{t:|F_n(X_n,t)| \vee |F_n(X_n,t-)| \ge b\}$ and let X_n^b denote the solution of

$$X_n^b(t) = U_n(t) + \int_0^t \chi_{[0,\eta_n^b)}(s-)F_n(X_n^b,s-)\,dY_n \tag{5.6}$$

that agrees with X_n on $[0,\eta_n^b)$. Then $\{(U_n,X_n^b,Y_n)\}$ is relatively compact and any limit point, (U,X^b,Y), gives a local solution (X^b,τ) of (5.3) with $\tau = \eta^c \equiv \inf\{t:|F(X^b,t)| \vee |F(X^b,t-)| \ge c\}$ for any $c < b$. If there exists a global solution X of (5.3) and weak local uniqueness holds, then $(U_n,X_n,Y_n) \Rightarrow (U,X,Y)$.

Proof By Lemma 4.1, there exist γ_n such that $\{(U_n\circ\gamma_n,X_n^b\circ\gamma_n,Y_n\circ\gamma_n)\}$ is relatively compact in $D_{\mathbb{R}^k\times\mathbb{R}^k\times\mathbb{R}^m}[0,\infty)$. C5.4 then implies that $\{\chi_{[0,\eta_n^b)}\circ\gamma_n F_n(X_n^b)\circ\gamma_n\} = \{\chi_{[0,\eta_n^b)}\circ\gamma_n G_n(X_n^b\circ\gamma_n,\gamma_n)\}$ is relatively compact in $D_{\mathbb{R}^k}[0,\infty)$. The relative compactness of $\{(U_n,X_n^b,Y_n)\}$ then follows by Corollary 4.5 and Proposition 4.3. The sequence $\{(U_n,X_n^b,Y_n,\eta_n^b)\}$ will be relatively compact in $D_{\mathbb{R}^k\times\mathbb{R}^k\times\mathbb{R}^m}[0,\infty)\times[0,\infty]$. Let (U,X^b,Y,η_0^b) denote a weak limit point. To simplify notation, assume that the original sequence converges and (with reference to the Skorohod representation theorem) assume that the convergence is almost sure rather than in distribution. Note that $\eta^b \le \eta_0^b$.

It follows that $U_n + \int F_n(X_n^b)dY_n \to U + \int F(X^b)dY$ and since

(5.7) $$X_n^b(t) = U_n(t) + \int_0^t F_n(X_n^b,s-)\,dY_n(s)$$

for $t \le \eta_n^b$,

(5.8) $$X^b(t) = U(t) + \int_0^t F(X^b,s-)\,dY(s)$$

for $t < \eta_0^b$. Let $c < b$. If $\eta^c < \eta^b$, then (5.8) holds for $t \le \eta^c$. If $\eta^c = \eta^b$, then $F(X^b)$ has a discontinuity at η^c with $|F(X^b,\eta^c-)| \le c$ and $|F(X^b,\eta^c)| \ge b$. It follows that for $c < d < b$, $(U_n(\eta_n^d),X_n^b(\eta_n^d),Y_n(\eta_n^d),Y_n(\eta_n^d-),F_n(X_n^b,\eta_n^d),F_n(X_n^b,\eta_n^d-),\eta_n^d)$ converges to $(U(\eta^d),X^b(\eta^d),Y(\eta^d),Y(\eta^d-),F(X^b,\eta^d),F(X^b,\eta^d-),\eta^d)$ and

(5.9) $$X^b(\eta^d) = U(\eta^d) + \int_0^{\eta^d} F(X^b,s-)\,dY(s)$$

so that (5.8) holds for $t \le \eta^c$ $(= \eta^d)$. Consequently, (X^b,η^c) is a local solution of (5.3).

Note that η^c is a measurable function of X^b (say $h_c(X^b)$). Consequently, if weak local uniqueness holds for (5.3) and there exists a global weak solution $\hat{X}$, then (X^b,η^c) must have the same distribution as $(\hat{X},h_c(\hat{X}))$ for all c and b with $c < b$. Since $\hat{X}$ is a global solution, $h_c(\hat{X}) \to \infty$ as $c \to \infty$. Convergence in distribution of (U_n,X_n,Y_n) follows. □

Unlike Theorem 2.2, Theorem 5.4 does not immediately hold with convergence in distribution replaced by convergence in probability. In particular, we must assume a strong uniqueness for the limiting equation (5.3) or convergence in probability could fail to hold even with $(U_n,Y_n) \equiv (U,Y)$. (If X and $\hat{X}$ are solutions of (5.3) which are not almost surely equal and $\{\xi_n\}$ are i.i.d. with $P\{\xi_n = 1\} = P\{\xi_n = 0\} = \frac{1}{2}$, then take $X_n = \xi_n X + (1-\xi_n)\hat{X}$.) We need the following lemma.

<u>5.5 Lemma</u> Assume that F has a representation in terms of a G satisfying C5.4(ii). Suppose that there exists a global (weak) solution of (5.3) and that strong local uniqueness holds for (5.3) for any version of (U,Y). Then any solution of (5.3) is a measurable function of (U,Y) (that is, if X satisfies (5.3), then there exists a measurable mapping $g: D_{\mathbb{R}^k \times \mathbb{R}^m}[0,\infty) \to D_{\mathbb{R}^k}[0,\infty)$ such that $X = g(U,Y)$ a.s.).

Proof Define (U_n,Y_n) by $(U_n(t),Y_n(t)) = (U(\frac{[nt]}{n}),Y(\frac{[nt]}{n}))$ and let X_n satisfy (5.2). Then X_n is a measurable function of (U_n,Y_n) and hence of (U,Y). Let $n_k \to \infty$ and $m_k \to \infty$. Then by Theorem 5.4, $(U_{n_k},U_{m_k},X_{n_k},X_{m_k},Y_{n_k},Y_{m_k})$ converges in distribution to $(U,U,X,\tilde{X},Y,Y)$ where X and $\tilde{X}$ satisfy (5.3). But, strong local uniqueness implies $X = \tilde{X}$ a.s. Consequently, if d is a metric for $D_{\mathbb{R}^k}[0,\infty)$, then

$$\lim_{k\to\infty} E[1\wedge d(X_{n_k},X_{m_k})] = E[1\wedge d(X,X)] = 0 \tag{5.10}$$

and hence $\{X_n\}$ is a Cauchy sequence for convergence in probability. Since X_n is a measurable function of (U,Y), the lemma follows. □

5.6 Corollary If in Theorem 5.4, we assume that (U_n,Y_n) converges in probability to (U,Y), that there exists a global solution of (5.3), and that strong local uniqueness holds for (5.3) for any version of (U,Y), then X_n converges in probability.

Proof Let f be a bounded, continuous function on $D_{\mathbb{R}^k}[0,\infty)$ and g be a bounded continuous function on $D_{\mathbb{R}^k\times\mathbb{R}^m}[0,\infty)$. Then since $(U_n,X_n,Y_n) \Rightarrow (U,X,Y)$

$$\lim_{n\to\infty} E[f(X_n)g(U_n,Y_n)] = E[f(X)g(U,Y)] \tag{5.11}$$

The convergence in probability of (U_n,Y_n) then implies

$$\lim_{n\to\infty} E[f(X_n)g(U,Y)] = E[f(X)g(U,Y)] \tag{5.12}$$

and L^1-approximation of measurable functions by continuous functions implies that (5.12) holds for all bounded, measurable g. Lemma 5.5 ensures the existence of a bounded measurable g such that $f(X) = g(U,Y)$ a.s. Consequently,

$$\lim_{n\to\infty} E[(f(X_n) - f(X))^2] = \lim_{n\to\infty} (E[f(X_n)^2] - 2E[f(X_n)f(X)] + E[f(X)^2]) = 0 \tag{5.13}$$

and convergence in probability for $\{X_n\}$ follows. □

Theorem 5.4 perhaps makes the theory look more simple and benign than it really is. Example 1.2 of the introduction reveals a pathology originally discovered by Wong and Zakai (1965): that certain naive approximations of semimartingale differentials lead to a lack of continuity of the corresponding solutions of stochastic differential equations. Indeed, it was this pathology that led E. J. McShane to develop his integral and to his proposal of a "canonical form" (McShane (1975)), though these can now be recognized as special cases of the semimartingale integral. The Wong-Zakai pathology has also led people to pay increased attention to Stratonovich and more generally symmetrized integrals and their differential equations (e.g., Mackevicius (1987)).

Examples 5.7, 5.8, and 5.9, that follow, motivate Theorem 5.10, which extends the results of Wong and Zakai. This extension is by no means the first; however, almost all the previous ones (e.g., Nakao and Yamato (1976), Doss (1977), Sussman (1978), Krener (1979), Ikeda and Watanabe (1981), Marcus (1981), Konecny (1983), Protter (1985), Mackevicius (1987), Picard (1989), Bally (1989)) are concerned with L^p, almost sure, or in probability convergence, always on only one probability space. The one exception is Słomiński (1989), who deals with weak convergence. Moreover, the level of generality in previous work is only that of Proposition 5.12 (albeit for more general approximation schemes; we have not bothered with the obvious modifications needed to include all of the previous results), and hence Theorem 5.10 is new even on the level of convergence in probability.

<u>5.7 Example</u> Let W_n be as in Example 1.2. Then clearly $\{W_n\}$ does not satisfy C2.2(i) (otherwise $\int W_n\,dW_n$ would converge to $\int W\,dW$); however, if we define $Y_n(t) = W(\frac{[nt]+1}{n})$ and $Z_n = W_n - Y_n$, then $\{Y_n\}$ satisfies C2.2(i) (Y_n is a martingale with respect to the filtration defined by $\mathcal{F}_t^n = \sigma\{W(s): s \leq \frac{[nt]+1}{n}\}$) and $Z_n \Rightarrow 0$. Furthermore, we observe that

$$\int_0^t Z_n\,dZ_n \to -\tfrac{1}{2}t \tag{5.14}$$

$$[Z_n]_t = -[Y_n, Z_n] = Z_n^2(t) - 2\int_0^t Z_n\,dZ_n \to t \tag{5.15}$$

and (noting that $Z_n(t\text{-}) = 0$ at each discontinuity of Z_n)

$$T_t(\int Z_n\,dZ_n) = \int_0^t |Z_n(s)|\,|\dot{W}(s)|\,ds \to Ct \tag{5.16}$$

where $C = E[|W(1)| \int_0^t |W(1) - W(s)| ds]$. Setting $H_n \equiv \int Z_n dZ_n$ and $I_n = [Z_n]$, it follows that $\{H_n\}$ and $\{I_n\}$ satisfiy C2.2(i). □

5.8 Example Let V be the Ornstein-Uhlenbeck process satisfiying

(5.17)
$$dV = dW - V dt$$

where W is a standard Brownian motion. Let

(5.18)
$$W_n(t) = \frac{1}{n}\int_0^{tn^2} V(s)\, ds = \int_0^t nV(n^2 s)\, ds$$

It follows that

(5.19)
$$W_n(t) = \frac{1}{n} W(n^2 t) - \frac{1}{n} V(n^2 t)$$

and defining $Y_n(t) = \frac{1}{n} W(n^2 t)$ and $Z_n(t) = -\frac{1}{n} V(n^2 t)$ we see that, as in Example 5.7, $\{Y_n\}$ satisfies C2.2(i) (each Y_n is a standard Brownian motion) and $Z_n \Rightarrow 0$. Again, setting $H_n = \int Z_n dZ_n$, $K_n = [Y_n, Z_n]$, and $I_n = [Z_n]$, we see that

(5.20)
$$H_n(t) = \frac{1}{n^2}\int_0^{n^2 t} V(s)\, dW(s) - \frac{1}{n^2}\int_0^{n^2 t} V(s)^2\, ds$$

The first term on the right of (5.20) is a martingale with quadratic variation

(5.21)
$$\frac{1}{n^4}\int_0^{n^2 t} V(s)^2\, ds$$

while the second term obviously has finite variation. It follows that $\{H_n\}$ satisfies C2.2(i), and $H_n(t) \to -\frac{1}{2}t$. Note, in addition, that $I_n(t) = -K_n(t) = t$. □

5.9 Example Let $\{U_k, k \geq 0\}$ be a finite, irreducible Markov chain with transition matrix $P = ((p_{ij}))$. Let $\pi = (\pi_1, \ldots, \pi_M)$ give the stationary distribution, and let f be a function satisfying

(5.22)
$$\sum_m f(m)\pi_m = 0$$

Define

$$W_n(t) = \frac{1}{\sqrt{n}}\sum_{k=1}^{[nt]} f(U_k) \tag{5.23}$$

Letting $Pg(i) \equiv \sum_j g(j)p_{ij}$, by (5.22) there exists a function h such that $Ph - h = f$. Substituting in (5.23), we obtain

$$\begin{aligned} W_n(t) &= \frac{1}{\sqrt{n}}\sum_{k=1}^{[nt]} (Ph(U_k) - h(U_k)) \\ &= \frac{1}{\sqrt{n}}\sum_{k=1}^{[nt]} (Ph(U_{k-1}) - h(U_k)) + \frac{1}{\sqrt{n}}\Big(Ph(U_{[nt]}) - Ph(U_0)\Big) \\ &\equiv Y_n(t) + Z_n(t) \end{aligned} \tag{5.24}$$

As in Examples 5.7 and 5.8, $\{Y_n\}$ is a sequence of martingales satisfying C2.2(i) which, by the martingale central limit theorem (see, for example, Ethier and Kurtz (1986), Theorem 7.1.4), converges in distribution to σW where

$$\sigma^2 = \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{[nt]} (Ph(U_{k-1}) - h(U_k))^2 = \sum_{i,j}\pi_i p_{ij}(Ph(i) - h(j))^2 \tag{5.25}$$

Again $Z_n \Rightarrow 0$, $[Z_n] \Rightarrow Ct$, $\int Z_n\,dZ_n \Rightarrow -\frac{1}{2}Ct$, and $[Y_n,Z_n] \Rightarrow Dt$ where

$$C = \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{[nt]} (Ph(U_k) - Ph(U_{k-1}))^2 = \sum_{i,j}\pi_i p_{ij}(Ph(j) - Ph(i))^2 \tag{5.26}$$

$$\begin{aligned} D &= \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{[nt]} (Ph(U_{k-1}) - h(U_k)(Ph(U_k) - Ph(U_{k-1})) \\ &= \sum_{i,j}\pi_i p_{ij}(Ph(i) - h(j))(Ph(j) - Ph(i))^2 \end{aligned} \tag{5.27}$$

and $\{\int Z_n\,dZ_n\}$ and $\{[Z_n]\}$ satisfy C2.2(i). □

Clearly Theorem 5.4 does not apply directly to

$$X_n(t) = U_n(t) + \int_0^t F(X_n, s-)\,dW_n(s) \tag{5.28}$$

for $\{W_n\}$ as in any of the above examples; however, if we specialize to

(5.29)
$$X_n(t) = X_n(0) + \int_0^t F(X_n(s-))\,dW_n(s)$$
$$= X_n(0) + \int_0^t F(X_n(s-))\,dY_n(s) + \int_0^t F(X_n(s-))\,dZ_n(s)$$

we can apply Theorem 5.4 to obtain the following extension of the classical results of Wong and Zakai (1965).

5.10 Theorem Let Y_n and Z_n be $\{\mathcal{F}_t^n\}$-semimartingales, and let $X_n(0)$ be $\mathcal{F}_0^n$-measurable. Let $F:\mathbb{R}^k \to \mathbb{M}^{km}$ in (5.28) be bounded and have bounded first and second order derivatives. Define $H_n = ((H_n^{\beta\gamma}))$ and $K_n = ((K_n^{\beta\gamma}))$ by

(5.30)
$$H_n^{\beta\gamma}(t) = \int_0^t Z_n^\beta(s-)\,dZ_n^\gamma(s)$$

and

(5.31)
$$K_n^{\beta\gamma}(t) = [Y_n^\beta, Z_n^\gamma]_t$$

Suppose that $\{Y_n\}$ and $\{H_n\}$ satisfy C2.2(i) and that $(X_n(0),Y_n,Z_n,H_n,K_n) \Rightarrow (X(0),Y,0,H,K)$. Then $\{(X_n(0),Y_n,Z_n,H_n,K_n,X_n)\}$ is relatively compact, and any limit point $(X(0),Y,0,H,K,X)$ satisfies

(5.32)
$$X(t) = X(0) + \int_0^t F(X(s-))\,dY(s) + \sum_{\alpha,\beta,\gamma} \int_0^t \partial_\alpha F_\beta(X(s-))\,F_{\alpha\gamma}(X(s-))\,d(H^{\gamma\beta}(s) - K^{\gamma\beta}(s))$$

where ∂_α denotes the partial derivative with respect to the αth variable and F_β denotes the βth column of F.

5.11 Remark a) The boundedness assumptions on F and its derivatives may be dropped to obtain a localized result with a statement analogous to that of Theorem 5.4.

b) The theorem can be extended to equations of the form

$$(5.33)\qquad X_n(t) = U_n(t) + \int_0^t F(X_n(s-))\,dY_n(s) + \int_0^t F(X_n(s-))\,dZ_n(s)$$

by writing $U_n = \hat{Y}_n + \hat{Z}_n$ and forming the system

$$(5.34)\qquad \begin{pmatrix} U_n(t) \\ X_n(t) \end{pmatrix} = \begin{pmatrix} U_n(0) \\ U_n(0) \end{pmatrix} + \int_0^t \begin{bmatrix} I & 0 \\ 0 & F(X_n(s-) \end{bmatrix} d\begin{pmatrix} \hat{Y}_n(s) \\ Y_n(s) \end{pmatrix} + \int_0^t \begin{bmatrix} I & 0 \\ 0 & F(X_n(s-) \end{bmatrix} d\begin{pmatrix} \hat{Z}_n(s) \\ Z_n(s) \end{pmatrix}$$

b) Note that since $Z_n \Rightarrow 0$, $\sup_{s\leq t}|Z_n(s)-Z_n(s-)| \Rightarrow 0$ and H and K must be continuous. We are not, however, assuming that Y is continuous.

c) Let $I_n^{\beta\gamma}(t) = [Z_n^\beta, Z_n^\gamma]$. Since

$$(5.35)\qquad [Z_n^\beta, Z_n^\gamma]_t = Z_n^\beta(t)Z_n^\gamma(t) - Z_n^\beta(0)Z_n^\gamma(0) - \int_0^t Z_n^\beta(s-)\,dZ_n^\gamma(s) - \int_0^t Z_n^\gamma(s-)\,dZ_n^\beta(s)$$

it follows that $I_n^{\beta\gamma} \Rightarrow -(H^{\beta\gamma} + H^{\gamma\beta})$. Since $I_n^{\beta\beta}$ is nondecreasing and converges in distribution to a continuous process, it follows that $\{I_n^{\beta\beta}\}$ satisfies C2.2(i) (at least for any finite δ) and, by estimating the increments of $I_n^{\beta\gamma}$ by the increments of $I_n^{\beta\beta}$ and $I_n^{\gamma\gamma}$, that $\{I_n^{\beta\gamma}\}$ satisfies C2.2(i).

d) Since

$$(5.36)\qquad |K_n^{\beta\gamma}(t+h) - K_n^{\beta\gamma}(t)| \leq \tfrac{1}{2}\Big([Y_n^\beta]_{t+h} + [Z_n^\gamma]_{t+h} - [Y_n^\beta]_t - [Z_n^\gamma]_t\Big)$$

it follows that $\{K_n\}$ satisfies C2.2(i).

Proof The result is obtained by integrating the second term on the right of (5.28) by parts. Note that by Ito's formula

$$(5.37)\qquad F_{i\beta}(X_n(t)) = F_{i\beta}(X_n(0)) + \sum_\alpha \int_0^t \partial_\alpha F_{i\beta}(X_n(s-))\,dX_n^\alpha(s) + R_n^{i\beta}(t)$$

where the increments of $R_n^{i\beta}$ are dominated by a linear combination of the increments of $[Y_n^\alpha]$ and $[Z_n^\alpha]$ (which implies that $\{R_n\}$ satisfies C2.2(i)). Integrating by parts we obtain

$$(5.38)\quad \int_0^t F_{i\beta}(X_n(s-))\,dZ_n^\beta(s)$$

$$= F_{i\beta}(X_n(t))Z_n^\beta(t) - F_{i\beta}(X_n(0))Z_n^\beta(0) - \sum_\alpha \int_0^t \partial_\alpha F_{i\beta}(X_n(s-))Z_n^\beta(s-)\,dX_n^\alpha(s)$$

$$- \int_0^t Z_n^\beta(s-)\,dR_n^{i\beta}(t) - \sum_\alpha \int_0^t \partial_\alpha F_{i\beta}(X_n(s-))\,d[X_n^\alpha,Z_n^\beta]_s + [R_n^{i\beta},Z_n^\beta]_t$$

$$= \eta_n(t) - \sum_{\alpha,\gamma} \int_0^t \partial_\alpha F_{i\beta}(X_n(s-))F_{\alpha\gamma}(X_n(s-)\,Z_n^\beta(s-)\,dZ_n^\gamma(s)$$

$$- \sum_{\alpha,\gamma} \int_0^t \partial_\alpha F_{i\beta}(X_n(s-))F_{\alpha\gamma}(X_n(s-))\,d([Y_n^\gamma,Z_n^\beta]_s + [Z_n^\gamma,Z_n^\beta]_s)$$

$$= \eta_n(t) - \sum_{\alpha,\gamma} \int_0^t \partial_\alpha F_{i\beta}(X_n(s-))F_{\alpha\gamma}(X_n(s-)\,d(H_n^{\beta\gamma}(s) + K_n^{\gamma\beta}(s) + I_n^{\gamma\beta}(s))$$

where $\eta_n \Rightarrow 0$. Substituting (5.38) into (5.29), the theorem follows from Theorem 5.4. □

Much of the work on approximation of solutions of stochastic differential equations has been concerned with linear interpolations of the integrator. The next result shows that Theorem 5.10 applies to these approximations.

5.12 Proposition For each n, let V_n be an $\{\mathcal{F}_t^n\}$-semimartingale and $\{\tau_k^n\}$ be a sequence of $\{\mathcal{F}_t^n\}$-stopping times with $\tau_0^n = 0$ and $\lim_{k\to\infty} \tau_k^n = \infty$. Suppose that $\lim_{n\to\infty} \sup_k \tau_{k+1}^n - \tau_k^n = 0$. Define the linear interpolation

$$(5.39)\qquad \hat{V}_n(t) = \frac{\tau_{k+1}^n - t}{\tau_{k+1}^n - \tau_k^n} V_n(\tau_k^n) + \frac{t - \tau_k^n}{\tau_{k+1}^n - \tau_k^n} V_n(\tau_{k+1}^n), \quad \tau_k^n \le t < \tau_{k+1}^n$$

and define

$$(5.40)\qquad Y_n(t) = V_n(\tau_{k+1}^n) = \hat{V}_n(\tau_{k+1}^n), \quad \tau_k^n \le t < \tau_{k+1}^n$$

and

$$(5.41)\quad Z_n(t) = \hat{V}_n(t) - Y_n(t) = \frac{\tau_{k+1}^n - t}{\tau_{k+1}^n - \tau_k^n}\Big(V_n(\tau_k^n) - V_n(\tau_{k+1}^n)\Big), \quad \tau_k^n \le t < \tau_{k+1}^n$$

Define H_n and K_n as in Theorem 5.10. Suppose that $V_n \Rightarrow Y$ where Y is continuous, and that $\{V_n\}$ satisfies C2.2(i) for some $\delta \in (0,\infty]$. (In particular, this last statement holds if $V_n = Y$ for all n.) Then $Y_n \Rightarrow Y$ and $Z_n \Rightarrow 0$, $\{Y_n\}$ and $\{H_n\}$ satisfy C2.2(i) for some $\delta \in (0,\infty]$,

$$(5.42)\qquad H_n \Rightarrow -\tfrac{1}{2}(([Y^\beta,Y^\gamma]))$$

and

$$(5.43)\qquad K_n \Rightarrow -(([Y^\beta,Y^\gamma]))$$

Proof Let $\tau_n(t) = \min\{\tau_k^n : \tau_k^n > t\}$. Then $\tau_n(t)$ is an $\{\mathcal{F}_t^n\}$-stopping time for each t and $\hat{V}_n$, Y_n and Z_n are adapted to the filtration $\{\mathcal{G}_t^n\}$ given by $\mathcal{G}_t^n = \mathcal{F}_{\tau_n(t)}^n$. The proof that $\{Y_n\}$ satisfies C2.2(i) is essentially the same as the proof that $\{\tilde{Y}_n\}$ satisfies C2.2(i) in Example 3.7. The fact that $Y_n \Rightarrow Y$ follows from the convergence of V_n and Proposition 3.6.5 of Ethier and Kurtz (1986). The convergence of Z_n follows from the continuity of Y and the continuous mapping theorem.

Note that

$$(5.44)\quad H_n^{\beta\gamma}(t) = -\tfrac{1}{2}\sum \frac{(\tau_{k+1}^n \wedge t - \tau_k^n \wedge t)^2}{(\tau_{k+1}^n - \tau_k^n)^2}(V_n^\beta(\tau_{k+1}^n) - V_n^\beta(\tau_k^n))(V_n^\gamma(\tau_{k+1}^n) - V_n^\gamma(\tau_k^n))$$

$$\approx -\tfrac{1}{2}\sum_{\tau_{k+1}^n \le t}(V_n^\beta(\tau_{k+1}^n) - V_n^\beta(\tau_k^n))(V_n^\gamma(\tau_{k+1}^n) - V_n^\gamma(\tau_k^n))$$

$$= -\tfrac{1}{2}[Y_n^\beta, Y_n^\gamma]_t$$

$$= -\tfrac{1}{2}\Big(Y_n^\beta(t)Y_n^\gamma(t) - \int_0^t Y_n^\beta(s-)\,dY_n^\gamma(s) - \int_0^t Y_n^\gamma(s-)\,dY_n^\beta(s)\Big)$$

and (5.42) follows by Theorem 2.2. A similar calculation gives (5.43). The fact that $\{H_n\}$ satisfies C2.2(i) follows from the monotonicity and convergence of $H_n^{\beta\beta}$ and the fact that the total variation of $H_n^{\beta\gamma}$ can be estimated in terms of the total variation of $H_n^{\beta\beta}$ and $H_n^{\gamma\gamma}$. □

6. Technical results

Uniform approximation by step functions Let E be a metric space with metric r. Let $\{\theta_k\}$ be a sequence of independent random variables, uniformly distributed on the interval $[\frac{1}{2},1]$. Fix $\epsilon > 0$, and for $z \in D_E[0,\infty)$ define $\tau_0(z) = 0$ and $\tau_{k+1}(z) = \inf\{t > \tau_k(z): r(z(t),z(\tau_k(z)))\vee r(z(t-),z(\tau_k(z))) \geq \epsilon\theta_k\}$ and set $\gamma_k(z) = z(\tau_k(z))$. Finally, define $I_\epsilon(z)$ by $I_\epsilon(z)(t) = \gamma_k(z)$ for $\tau_k(z) \leq t < \tau_{k+1}(z)$. Note that $r(z(t),I_\epsilon(z)(t)) \leq \epsilon$ for all t. Let $U_1 = \{u:u = r(z(t),z(0))$ or $r(z(t-),z(0))$ for some t such that $z(t) \neq z(t-)\}$, and defining $m(t) = \sup_{s\leq t} r(z(s),z(0))$, let $U_2 = \{m(t):m$ is not strictly increasing at $t\}$. U_1 and U_2 are countable, so with probability one, $\epsilon\theta_0 \notin U_1\cup U_2$. Let $z_n \to z$, and assume that $\epsilon\theta_0 \notin U_1\cup U_2$. Either m is strictly increasing at $\tau_1(z)$ or $r(z(\tau_1(z)-),z(0)) < \epsilon\theta_0 < r(z(\tau_1(z)),z(0))$, and it follows that $\tau_1(z_n) \to \tau_1(z)$. Either z is continuous at $\tau_1(z)$ or $r(z(\tau_1(z)-),z(0)) < \epsilon\theta_0 < r(z(\tau_1(z)),z(0))$, and it follows that $\gamma_1(z_n) \to \gamma_1(z)$. In general, if $z_n \to z$ in the Skorohod topology, $t_n \to t$ and $z_n(t_n) \to z(t)$, then $z_n(t_n+\cdot) \to z(t+\cdot)$ in the Skorohod topology. Consequently, $z_n \to z$ implies $z_n(\tau_1(z_n)+\cdot) \to z(\tau_1(z)+\cdot)$ a.s. An induction argument then shows that $z_n \to z$ implies $\tau_k(z_n) \to \tau_k(z)$ and $\gamma_k(z_n) \to \gamma_k(z)$ a.s for all k. With these observations, we can prove the following lemma.

6.1 Lemma Let I_ϵ be defined as above. If $z_n \to z$ in the Skorohod topology on $D_E[0,\infty)$, then $(z_n,I_\epsilon(z_n)) \to (z,I_\epsilon(z))$ a.s. in the Skorohod topology on $D_{E\times E}[0,\infty)$.

To carry out the proof, we need the following. (See Proposition 3.6.5 of Ethier and Kurtz (1986). Note that the third condition in that proposition is implied by the other two.)

6.2 Lemma For an arbitrary metric space (E',r'), $v_n \to v$ in the Skorohod topology on $D_{E'}[0,\infty)$ if and only if the following conditions hold:

C6.2(i) If $t_n \to t$, then $\lim_{n\to\infty} r'(v_n(t_n),v(t))\wedge r'(v_n(t_n),v(t-)) = 0$

C6.2(ii) If $s_n \geq t_n$, $s_n,t_n \to t$, and $v_n(t_n) \to v(t)$, then $v_n(s_n) \to v(t)$.

Proof of Lemma 6.1 Suppose $z_n \to z$ in $D_E[0,\infty)$ and $t_n \to t$. If $\tau_k(z) < t < \tau_{k+1}(z)$, then $I_\epsilon(z)$ is continuous at t, $I_\epsilon(z_n)(t_n) \to \gamma_k(z) = I_\epsilon(z)(t)$, and C6.2(i) and (ii) follow for $\{(z_n,I_\epsilon(z_n))\}$ by the analogous conditions for $\{z_n\}$. If $t = \tau_k(z)$, we can assume that either

z is continuous at $\tau_k(z)$ or $r(z(\tau_k(z)-),z(\tau_{k-1}(z))) < \epsilon\theta_{k-1} < r(z(\tau_k(z)),z(\tau_{k-1}(z)))$. The convergence of $\tau_{k-1}(z_n)$, $\tau_k(z_n)$, $\gamma_{k-1}(z_n)$, and $\gamma_k(z_n)$ implies C6.2(i) and (ii) for $\{I_\epsilon(z_n)\}$, and if z is continuous at $\tau_k(z)$, C6.2(i) and (ii) follow for $\{(z_n,I_\epsilon(z_n))\}$. If $r(z(\tau_k(z)-),z(\tau_{k-1}(z))) < \epsilon\theta_{k-1} < r(z(\tau_k(z)),z(\tau_{k-1}(z)))$, then, with probability one, for n sufficiently large the same inequality holds with z replaced by z_n. Consequently, if $t_n \geq \tau_k(z_n)$ and $t_n \to \tau_k(z)$, then $z_n(t_n)$ and $I_\epsilon(z_n)(t_n)$ both converge to $\gamma_k(z)$, and if $t_n < \tau_k(z_n)$ and $t_n \to t$, then $z_n(t_n)$ converges to $z(\tau_k(z)-)$ and $I_\epsilon(z_n)(t_n)$ converges to $\gamma_{k-1}(z) = I_\epsilon(z)(\tau_k(z)-)$. C6.2(i) and (ii) follow for $\{(z_n,I_\epsilon(z_n))\}$. □

Uniform tightness Jakubowski, Mémin, and Pages (1989) and Slomiński (1989) develop their results under a "uniform tightness" condition. We discuss this condition for a sequence of one-dimensional semimartingales $\{Y_n\}$ satisfying $Y_n(0) = 0$. The results below are essentially contained in Lemma 3.1 of Jakubowski, Mémin, and Pages (1989).

Let $\mathcal{H}_n$ denote the collection of cadlag $\{\mathcal{F}_t^n\}$-adapted, $\mathbf{R}$-valued processes satisfying $|H_n(t)| \leq 1$ for all $t \geq 0$. Then $\{Y_n\}$ is uniformly tight if for each $t > 0$

(6.1)
$$\left\{\int_0^t H_n(s-)dY_n(s): H_n \in \mathcal{H}_n,\ n = 1,2,\ldots\right\}$$

is stochastically bounded.

Assume that $\{Y_n\}$ is uniformly tight. Let $\mathcal{T}_n$ denote the collection of $\{\mathcal{F}_t^n\}$-stopping times. For $\tau \in \mathcal{T}_n$ and $\epsilon > 0$, let $H_n = \chi_{[0,\tau)}$. Then the integral in (6.1) gives $Y_n(t\wedge\tau)$, and we see that for each $t > 0$, $\{Y_n(t\wedge\tau): \tau \in \mathcal{T}_n,\ n = 1,2,\ldots\}$ is stochastically bounded. Considering the collection of stopping times of the form $\tau = \inf\{s: |Y_n(s)| \geq c\}$, it follows that $\{\sup_{s\leq t}|Y_n(s)|: n = 1,2,\ldots\}$ is stochastically bounded. Recalling that

(6.2)
$$[Y_n]_t = Y_n(t)^2 - \int_0^t 2Y_n(s-)dY_n(s)$$

and using the stochastic boundedness of the suprema, we see that $\{[Y_n]_t: n = 1,2,\ldots\}$ is stochastically bounded.

The stochastic boundedness of the quadratic variations ensures that the uniform tightness of $\{Y_n\}$ implies uniform tightness of $\{Y_n^\delta\}$ for each $0 < \delta < \infty$. Fix $0 < \delta < \infty$ and let Y_n^δ

$= M_n^\delta + A_n^\delta$ be the canonical decomposition of Y_n^δ (Protter (1990) §III.5). Then the discontinuities of M_n^δ and A_n^δ are bounded by 2δ, and $E[[Y_n^\delta]_\tau] = E[[M_n^\delta]_\tau] + E[[A_n^\delta]_\tau]$ for any stopping time τ (with the possibility of $\infty = \infty$) (Protter (1990) §IV.2.) Let $\gamma_n^c = \inf\{s:[Y_n^\delta]_s \geq c\}$. Fix t and for $k = 1,2,\ldots$, let $\{t_i^k\}$ be a partition of $[0,t]$ with $\lim_{k\to\infty} \max_i(t_{i+1}-t_i) = 0$. Define

$$(6.3) \qquad H_n^k = \sum_i \operatorname{sign}\Big(E[A_n^\delta(t_{i+1}^k \wedge \gamma_n^c) - A_n^\delta(t_i^k \wedge \gamma_n^c) | \mathcal{F}_{t_i}^n]\Big)\chi_{[t_i^k\wedge\gamma_n^c, t_{i+1}^k\wedge\gamma_n^c)}$$

The first term on the right of

$$(6.4) \qquad \int_0^u H_n^k(s-)\,dY_n^\delta(s) = \int_0^u H_n^k(s-)\,dM_n^\delta(s) + \int_0^u H_n^k(s-)\,dA_n^\delta(s) \equiv U_n^k(u) + V_n^k(u)$$

satisfies

$$(6.5) \qquad E[\sup_{s\leq t} U_n^k(s)^2] \leq 4\,E[M_n^\delta(t\wedge\gamma_n^c)^2] \leq 4\,(c+(2\delta)^2)$$

so $\{U_n^k(t):k,n = 1,2,\ldots\}$ is stochastically bounded which, by the stochastic boundedness of (6.1) (with Y_n replaced by Y_n^δ), implies the stochastic boundedness of $\{V_n^k(t):k,m = 1,2,\ldots\}$. But the predictability of A_n^δ implies

$$(6.6) \qquad T_{t\wedge\gamma_n^c}(A_n^\delta)$$

$$= \lim_{k\to\infty} \sum \operatorname{sign}\Big(E[A_n^\delta(t_{i+1}^k\wedge\gamma_n^c) - A_n^\delta(t_i^k\wedge\gamma_n^c)|\mathcal{F}_{t_i}^n]\Big)\Big(A_n^\delta(t_{i+1}^k\wedge\gamma_n^c) - A_n^\delta(t_i^k\wedge\gamma_n^c)\Big)$$

$$= \lim_{k\to\infty} V_n^k(t)$$

(see Dellacherie and Meyer (1982), page 423) so $\{T_{t\wedge\gamma_n^c}(A_n^\delta)\}$ is stochastically bounded for each c. But the stochastic boundedness of $\{[Y_n^\delta]_t\}$ for each t implies that for each $\epsilon > 0$, there exists a c such that $P\{\gamma_n^c \leq t\} \leq \epsilon$ and hence there exists an $a > 0$ such that $P\{T_t(A_n^\delta) \geq a\} \leq P\{T_{t\wedge\gamma_n^c}(A_n^\delta) \geq a\} + P\{\gamma_n^c \leq t\} \leq 2\epsilon$, verifying the stochastic boundedness of $\{T_t(A_n^\delta)\}$ and C2.2(ii). C2.2(iii) is immediate, so C2.2(i) holds.

If there exists a δ for which $\{J_\delta(Y_n)\}$ is stochastically bounded and C2.2(i) holds, then $\{Y_n^\delta\}$ satisfies C4.1 and Lemma 4.1 implies $\{Y_n\}$ is uniformly tight.

7. References

Avram, F. (1988). Weak convergence of the variations, iterated integrals, and Doléans-Dade exponentials of sequences of semimartingales. *Ann. Probab.* 16, 246-250.

Bally, V. (1989). Approximation for the solutions of stochastic differential equations. I: L^p-convergence. *Stochastics* 28, 209-246.

Bobkoski, Mark J. (1983). Hypothesis testing in nonstationary time series. PhD Dissertation, University of Wisconsin-Madison.

Dellacherie, Claude and Meyer, Paul-André (1982). *Probabilities and Potential B.* North Holland, Amsterdam.

DiMasi, G. B. and Runggaldier, W. J. (1981). Continuous-time approximations for the nonlinear filtering problem. *Appl. Math. Optim.* 7, 233-245.

DiMasi, G. B. and Runggaldier, W. J. (1982). On approximation methods for nonlinear filtering. *Nonlinear Filtering and Stochastic Control, Lect. Notes Math.* 972. S. K. Mitter and A. Moro eds. Springer Verlag, Berlin, 249-259.

Duffie, Darrell and Protter, Philip (1989). From discrete to continuous time finance: Weak convergence of the financial gain process. (to appear)

Dynkin, E. B. and Mandelbaum, A. (1983). Symmetric statistics, Poisson point processes, and multiple Wiener integrals. *Ann. Statist.* 11, 739-745.

Billingsley, Patrick (1968). *Convergence of Probability Measures.* Wiley, New York.

Chan, N. H. and Wei, C. Z. (1988). Limiting distributions of least squares estimates of unstable autoregressive processes. *Ann. Statist.* 16, 367-401.

Cox, Dennis D. and Llatas, Isabel (1989). M-estimation for nearly nonstationary autoregressive time series. (to appear)

Doss, H. (1977). Liens entre équations différentielles stochastiques et ordinaires. *Ann. Inst. Henri Poincaré* 13, 99-125.

Emery, M. (1989). On the Azéma martingales. Séminaire de Probabilités XXIII (to appear).

Ethier, Stewart N. and Kurtz, Thomas G. (1986). *Markov Processes: Characterization and Convergence.* Wiley, New York.

Filippova, A. A. (1961). Mises' theorem on the asymptotic behavior of functionals of empirical distribution functions and its statistical applications. *Theor. Probab. Appl.* 7, 24-56.

Goggin, Eimear Mary (1988). Weak convergence of conditional probabilities. PhD dissertation, University of Wisconsin-Madison.

Hall, Peter (1979). On the invariance principle for U-statistics. *Stochastic Process. Appl.* 9, 163-174.

Hoffman, Karin (1989). Approximation of stochastic integral equations by martingale difference arrays. (preprint).

Ikeda, N. and Watanabe, S. (1981). *Stochastic Differential Equations and Diffusion Processes.* North-Holland, Amsterdam.

Jacod, J., Mémin, J. and Métivier, M. (1983). On tightness and stopping times. *Stochastic Process. Appl.* 14, 109-146.

Jacod, Jean and Shiryaev, Albert (1987). *Limit Theorems for Stochastic Processes.* Springer-Verlag, Berlin.

Jakubowski, A, Mémin, J., and Pages, G. (1989). Convergence en loi des suites d'integrales stochastique sur l'espace D^1 de Skorokhod. *Probab. Th. Rel. Fields* 81, 111-137.

Johnson, Daniel P. (1983). Diffusion approximations for optimal filtering of jump processes and for queuing networks. PhD dissertation, University of Wisconsin-Madison.

Konecny, F. (1983). On Wong-Zakai approximation of stochastic differential equations. *J. Multivariate Anal.* 13, 605-611.

Krener, A. J. (1979). A formal approach to stochastic integration and differential equations. *Stochastics* 3, 105-125.

Kurtz, Thomas G. (1990). Random time changes and convergence in distribution under the Meyer-Zheng conditions. *Ann. Probab.* (to appear).

Kushner, Harold J. (1974). On the weak convergence of interpolated Markov chains to a diffusion. *Ann. Probab.* 2, 40-50.

Llatas, Isabel (1987). Asymptotic inference for nearly non-stationary time series. PhD Dissertation, University of Wisconsin-Madison.

Marcus, S. I. (1981). Modeling and approximation of stochastic differential equations driven by semimartingales. *Stochastics* 4, 223-245.

Mackevicius, V. (1987). S^p stability of solutions of symmetric stochastic differential equations with discontinuous driving semimartingales. *Ann. Inst. Henri Poincaré* 23, 575-592.

McShane, E. J. (1975). Stochastic differential equations. *J. Multivariate Anal.* 5, 121-177.

Meyer, Paul-André (1989). Note sur les martingales d'Azema. Séminaire de Probabilités XXIII (to appear).

Meyer, Paul-André and Zheng, W. A. (1984). Tightness criteria for laws of semimartingales. *Ann. Inst. Henri Poincaré* 20, 353-372.

Nakao, S. and Yamato, Y. (1976). Approximation theorem on stochastic differential equations. *Proc. International Symp. S.D.E. Kyotô*, 283-296.

Neuhaus, G. (1977). Functional limit theorems for U-statistics in the degenerate case. *J. Multivariate Anal.* 7, 424-439.

Picard, Jean (1989). Convergence in probability for perturbed stochastic integral equations. *Probab. Th. Rel. Fields* 81, 383-451.

Protter, Philip (1985). Approximations of solutions of stochastic differential equations driven by semimartingales. *Ann. Probab.* 13, 716-743.

Protter, Philip (1990). *Stochastic Integration and Differential Equations: A New Approach.* Springer-Verlag, New York.

Protter, Philip and Sharpe, Michael (1979). Martingales with given absolute value. *Ann. Probab.* 7, 1056-1058.

Rubin, H. and Vitale, R. A. (1980). Asymptotic distribution of symmetric statistics. *Ann. Statist.* 8, 165-170.

Skorohod, A. V. (1965). *Studies in the Theory of Random Processes.* Addison-Wesley, Reading, MA.

Slomiński, Leszek (1989). Stability of strong solutions of stochastic differential equations. *Stochastic Process. Appl.* 31,173-202.

Strasser, H. (1986). Martingale difference arrays and stochastic integrals. *Probab. Th. Rel. Fields* 72, 83-98.

Sussman, H. (1978). On the gap between deterministic and stochastic ordinary differential equations. *Ann. Probab.* 6, 19-41.

Wong, Eugene, and Zakai, Moshe (1965). On the convergence of ordinary integrals to stochastic integrals. *Ann. Math. Statist.* 36, 1560-1564.

Acknowledgements The research of both authors is supported in part by the National Science Foundation. This work was completed while the first author was visiting the Department of Mathematics at the University of Utah. The hospitality of that department is appreciated and gratefully acknowledged.