WEAK CONVERGENCE OF STOCHASTIC INTEGRALS AND DIFFERENTIAL EQUATIONS: II INFINITE DIMENSIONAL CASE

by

Thomas G. Kurtz
University of Wisconsin-Madison

and

Philip E. Protter
Purdue University

Technical Report #96-16

Department of Statistics
Purdue University
West Lafayette, IN USA

May 1996

WEAK CONVERGENCE OF STOCHASTIC INTEGRALS AND DIFFERENTIAL EQUATIONS II: INFINITE DIMENSIONAL CASE

by

Thomas G. Kurtz and Philip E. Protter

University of Wisconsin-Madison   Purdue University

Abstract

This is a semi-expository paper which nonetheless contains many new results. We treat the topics of stochastic integration, weak convergence, and stochastic differential equations in an infinite dimensional setting. Our results unify several approaches, and we give examples illustrating their power. The unification also leads to new results not covered with the usual techniques.

△

Weak convergence of stochastic integrals and differential equations II: Infinite dimensional case [1]

Thomas G. Kurtz[2]
Departments of Mathematics and Statistics
University of Wisconsin - Madison
Madison, WI 53706-1388

Philip E. Protter[3]
Departments of Mathematics and Statistics
Purdue University
West Lafayette, IN 47907-1395

January 28, 1996

Contents

[1]Lecture notes for the CIME school in probablity.

[3]Research supported in part by NSF grants DMS 92-04866 and DMS 95-04323
[3]Research supported in part by NSF grant INT 94-01109 and NSA grant MDR 904-94-H-2049

1 Introduction

In Part I, we discussed weak limit theorems for stochastic integrals, with the prinicple result being the following (cf. Part I, Section 7):

Theorem 1.1 *For each $n = 1, 2, \ldots$, let (X_n, Y_n) be an $\{\mathcal{F}_t^n\}$-adapted process with sample paths in $D_{\mathbb{M}^{km}\times\mathbb{R}^m}[0,\infty)$ such that Y_n is an $\{\mathcal{F}_t^n\}$-semimartingale. Let $Y_n = M_n + A_n$ be a decomposition of Y_n into an $\{\mathcal{F}_t^n\}$-martingale M_n and a finite variation process A_n. Suppose that one of the following two conditions hold:*

UT (Uniform tightness.) For $\mathcal{S}_0^n$, the collection of piecewise constant, $\{\mathcal{F}_t^n\}$-adapted processes

$$\mathcal{H}_t^0 = \cup_{n=1}^{\infty}\{|Z_- \cdot Y_n(t)| : Z \in \mathcal{S}_0^n, \sup_{s\le t}|Z(s)| \le 1\}$$

is stochastically bounded.

UCV (Uniformly controlled variations.) $\{T_t(A_n)\}$ is stochastically bounded for each $t > 0$, and there exist stopping times $\{\tau_n^\alpha\}$ such that $P\{\tau_n^\alpha \le \alpha\} \le \alpha^{-1}$ and

$$\sup_n E[[M_n]_{t\wedge\tau_n^\alpha}] < \infty$$

for each $t > 0$.

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 an $\{\mathcal{F}_t\}$-semimartingale for a filtration $\{\mathcal{F}_t\}$ with respect to which X is adapted and $(X_n, Y_n, X_{n-} \cdot Y_n) \Rightarrow (X, Y, X_- \cdot Y)$ 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 $(X_n, Y_n, X_{n-} \cdot Y_n) \to (X, Y, X_- \cdot Y)$ in probability.

In this part, we consider the analogous results for stochastic integrals with respect to infinite dimensional semimartingales. We are primarily concerned with integrals with respect to semimartingale random measures, in particular, worthy martingale measures as developed by Walsh (1986). We discover, however, that the class of semimartingale random measures is not closed under the natural notion of weak limit unlike the class of finite dimensional semimartingales (Part I, Theorem 7.3). Consequently, we work with a larger class of infinite dimensional semimartingales which we call $H^{\#}$-semimartingales. This class includes semimartingale random measures, Banach-space valued martingales, and cylindrical Brownian motion.

A summary of results on semimartingale random measures is given in Section 2. The definitions and results come primarily from Walsh (1986). $H^{\#}$-semimartingales are introduced in Section 3. The stochastic integral is defined through approximation by finite dimensional integrands. The basic assumption on the semimartingale is essentially the "good integrator" condition that defines a semimartingale in the sense of Section 1 of Part I. This approach allows us to obtain the basic stochastic integral covergence theorems in Sections 4 and 5 as an application of Theorem 1.1.

Previous general results on convergence of infinite dimensional stochastic integrals include work of Walsh (1986), Chapter 7, Cho (1994, 1995), and Jakubowski (1995). Walsh and Cho consider martingale random measures as distribution-valued processes converging weakly in $D_{\mathcal{S}'(\mathbf{R}^d)}[0,\infty)$. Walsh assumes all processes are defined on the same sample space (the canonical sample space) and requires a strong form of convergence for the integrands. Cho requires $(X_n, M_n) \Rightarrow (X, M)$ in $D_{L\times\mathcal{S}'(\mathbf{R}^d)}[0,\infty)$ where L is an appropriate function space. Both Walsh and Cho work under assumptions analogous to the UCV assumption. Jakubowski gives results for Hilbert space-valued semimartingales under the analogue of the UT condition. Our results are given under the UT condition, although estimates of the type used by Walsh and Cho are needed to verify that particular sequences satisfy the UT condition.

Section 6 contains a variety of technical results on the uniform tightness condition. Section 7 includes a uniqueness result for stochastic differential equations satisfying a Lipschitz condition with a proof that seems to be new even in the finite dimensional setting. As an example, a spin-flip model is obtained as a solution of a stochastic differential equation in sequence space. Convergence results for stochastic differential equations are given based on the results of Sections 4 and 5.

Section 8 briefly discusses stochastic differential equations for Markov processes and introduces L_1-estimates of Graham that are useful in proving existence and uniqueness, particularly for infinite systems. Infinite systems are the topic of Section 9. Existence and uniqueness results, similar to results of Shiga and Shimizu (1980) for systems of diffusions, are given for very general kinds of equations. Results of McKean-Vlasov type are given in Section 10 using the results of Section 9.

Stochastic partial differential equations are a natural area of application for the results discussed here. We have not yet developed these applications, but Section 11 summarizes some of the ideas that seem most useful in obtaining convergence theorems.

Section 12 includes several simple examples illustrating the methods of the paper. Diffusion approximations, an averaging theorem, limit theorems for jump processes, and error analysis for a simulation scheme are described.

Acknowledgements. Many people provided insight and important references during the writing of this paper. In particular, we would like to thank Don Dawson, Carl Graham, Peter Kotelenez, Jim Kuelbs, Luciano Tubaro, and John Walsh.

2 Semimartingale random measures.

Let (U, r_U) be a complete, separable metric space, let $U_1 \subset U_2 \subset \cdots$ be a sequence of sets $\{U_m\} \subset \mathcal{B}(U)$ satisfying $\cup_m U_m = U$, and let $\mathcal{A} = \{B \in \mathcal{B}(U) : B \subset U_m, \text{ some } m\}$. Frequently, $U_m = U$ and $\mathcal{A} = \mathcal{B}(U)$. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space and $\{\mathcal{F}_t\}$ a complete, right continuous filtration, and let Y be a stochastic process indexed by $\mathcal{A} \times [0, \infty)$ such that

- For each $B \in \mathcal{A}$, $Y(B, \cdot)$ is an $\{\mathcal{F}_t\}$-semimartingale with $Y(B, 0) = 0$.
- For each $t \geq 0$ and each disjoint sequence $\{B_i\} \subset \mathcal{B}(U)$, $Y(U_m \cap \cup_{i=1}^{\infty} B_i, t) = \sum_{i=1}^{\infty} Y(U_m \cap B_i, t)$ a.s.

Then Y is an *$\{\mathcal{F}_t\}$-semimartingale random measure*. We will say that Y is *standard* if $Y = M + V$ where $V(B, t) = \tilde{V}(B \times [0, t])$ for a random σ-finite signed measure $\tilde{V}$ on $U \times [0, \infty)$ satisfying $|\tilde{V}|(U_m \times [0, t]) < \infty$ a.s. for each $m = 1, 2, \ldots$ and $t \geq 0$ and M is a *worthy* martingale random measure in the sense of Walsh (1986), that is, $M(A. \cdot)$ is locally square integrable for each $A \in \mathcal{A}$, and there exists a (positive) random measure K on $U \times U \times [0, \infty)$ such that

$$|\langle M(A), M(B) \rangle_{t+s} - \langle M(A), M(B) \rangle_t| \leq K(A \times B \times (t, t+s]) , \quad A, B \in \mathcal{A},$$

and $K(U_m \times U_m \times [0, t]) < \infty$ a.s. for each $t > 0$. K is called the dominating measure. (Merzbach and Zakai (1993) define a slightly more general notion of *quasi-worthy* martingale which could be employed here. See also the defintion of *conditionally worthy* in Example 12.5.) Note that if U is finite, then every semimartingale random measure is standard.

M is *orthogonal* if $\langle M(A), M(B) \rangle_t = 0$ for $A \cap B = \emptyset$. If M is orthogonal, then $\pi(A \times (0, t]) \equiv \langle M(A) \rangle_t$ extends to a random measure on $U \times [0, \infty)$, and if we define $K(\Gamma) = \pi(f^{-1}(\Gamma))$ for $f(u, t) = (u, u, t)$, K is a dominating measure for M. In particular, if M is orthogonal, then M is worthy.

If φ is a simple function on U, that is $\varphi = \sum_{i=1}^{m} c_i I_{B_i}$ for disjoint $\{B_i\} \subset \mathcal{A}$, and $\{c_i\} \subset \mathbb{R}$, then we can define

$$Y(\varphi, t) = \sum_{i=1}^{m} c_i Y(B_i, t).$$

If Y is standard and $E[K(U \times U \times [0, t])] < \infty$ for each $t > 0$, then

$$E[M(\varphi, t)^2] \leq E[\int_{U \times U} |\varphi(u)\varphi(v)| K(du \times dv \times (0, t])] \leq \|\varphi\|_{\infty}^2 E[K(U \times U \times (0, t])] \quad (2.1)$$

so $Y(\varphi, t)$ can be extended uniquely at least to all $\varphi \in B(U)$ for which the integral against $\tilde{V}$ is defined, that is

$$Y(\varphi, t) = M(\varphi, t) + \int_U \varphi(u) \tilde{V}(du \times [0, t])$$

where $M(\varphi,t)$ is defined as the limit of $M(\varphi_n,t)$ for simple φ_n using (2.1).

More generally, for each $j=0,1,\ldots$, let $\{B_i^j\}\subset\mathcal{A}$ be disjoint, let $0=t_0<t_1<t_2<\cdots$, and let C_i^j be an $\mathcal{F}_{t_j}$-measurable random variable. Define

$$X(u,t)=\sum_{i,j}C_i^jI_{B_i^j}(u)I_{(t_j,t_{j+1}]}(t) \tag{2.2}$$

and define $X\cdot Y$ by

$$X\cdot Y(t)=\sum C_i^j(Y(B_i^j,t_{j+1}\wedge t)-Y(B_i^j,t_j\wedge t))\,.$$

Again, we can estimate the martingale part

$$\begin{aligned} E[(X\cdot M(t))^2] &= E\left[\sum_j\sum_{i_1,i_2}C_{i_1}^jC_{i_2}^j(\langle M(B_{i_1}^j),M(B_{i_2}^j)\rangle_{t_{j+1}\wedge t}-\langle M(B_{i_1}^j),M(B_{i_2}^j)\rangle_{t_j\wedge t})\right]\\ &\le E\left[\int_{U\times U\times[0,t]}|X(u,s)X(v,s)|K(du\times dv\times ds)\right], \end{aligned} \tag{2.3}$$

and if $E[K(U\times U\times[0,t])]<\infty$ for each $t>0$, we can extend the definition of $X\cdot M$ (and hence of $X\cdot Y$) to all bounded, $|\tilde{V}|$-integrable processes that can be approximated by simple processes of the form (2.2).

An alternative approach to defining $X_-\cdot M$ is to first consider

$$X(t,u)=\sum_{i=1}^m\xi_i(t)I_{B_i}(u), \tag{2.4}$$

for disjoint $\{B_i\}\subset\mathcal{A}$ and cadlag, adapted ξ_i. Set

$$X_-\cdot Y(t)=\sum_{i=1}^m\int_0^t\xi_i(s-)dY(B_i,s)\,,$$

where the integrals are ordinary semimartingale integrals. We then have

$$\begin{aligned} E[(X_-\cdot M(t))^2] &= E[\ \sum_{i,j}\int_0^t\xi_i(s-)\xi_j(s-)d\langle M(B_i,\cdot),M(B_j,\cdot)\rangle_s]\\ &\le E[\int_{U\times U\times(0,t]}|X(s-,u)X(s-,v)|K(du\times dv\times ds)]\,, \end{aligned} \tag{2.5}$$

which, of course, is the same as (2.3). By Doob's inequality, we have

$$E[\sup_{s\le t}(X_-\cdot M(s))^2]\le 4E[\int_{U\times U\times(0,t]}|X(s-,u)X(s-,v)|K(du\times dv\times ds)]. \tag{2.6}$$

For future reference, we also note the simple corresponding inequalities for V,

$$E[\sup_{s\le t}|X_-\cdot V(s)|]\le E[\int_{U\times(0,t]}|X(s-,u)||\tilde{V}|(du\times ds)] \tag{2.7}$$

and

$$E[\sup_{s\leq t}(X_{-}\cdot V(s))^2] \leq E[|\tilde{V}|(U\times[0,t])\int_{U\times(0,t]}|X(s-,u)|^2|\tilde{V}|(du\times ds)]\,. \tag{2.8}$$

Let $\mathcal{P}$ be the σ-algebra of subsets of $\Omega\times U\times[0,\infty)$ generated by sets of the form $A\times B\times(t,t+s]$ for $t,s\geq 0$, $A\in\mathcal{F}_t$, and $B\in\mathcal{B}(U)$. $\mathcal{P}$ is the σ-algebra of *predictable sets.* If $E[K(U\times U\times[0,t])]<\infty$ and $|\tilde{V}|(U\times[0,t])<\infty$ a.s. for each $t>0$, then the bounded $\mathcal{P}$-measurable functions gives the class of bounded processes X for which $X\cdot Y$ is defined. Of course, the estimate (2.3) also allows extension to unbounded X for which the right side is finite, provided X is also almost surely integrable with respect to $\tilde{V}$.

Note that if K satisfies

$$K(A\times B\times[0,t]) = K_1(A\cap B\times[0,t]) + K_2(A\times B\times[0,t]) \tag{2.9}$$

where K_1 is a random measure on $U\times[0,\infty)$ and we define

$$\hat{K}(A\times[0,t]) = K_1(A\times[0,t]) + \frac{1}{2}(K_2(A\times U\times[0,t]) + K_2(U\times A\times[0,t])), \tag{2.10}$$

then

$$E[\sup_{s\leq t}(X\cdot M(t))^2] \leq 4E[\int_{U\times[0,t]}|X(u,s)|^2\hat{K}(du\times ds)]\,. \tag{2.11}$$

For future reference, if $|\tilde{V}|(U_m\times[0,\cdot])$ is locally in L_1 for each m, that is, there is a sequence of stopping times $\tau_n\to\infty$ such that $E[|\tilde{V}|(U_m\times[0,t\wedge\tau_n])]<\infty$ for each $t>0$, then we can define $\hat{V}(A\times[0,\cdot])$ to be the predictable projection of $|\tilde{V}|(A\times[0,\cdot])$, and we have

$$E[\int_{U\times[0,t]}X(s-,u)|\tilde{V}|(du\times ds)] = E[\int_{U\times[0,t]}X(s-,u)\hat{V}(du\times ds)] \tag{2.12}$$

for all positive, cadlag, adapted X (allowing $\infty=\infty$).

Example 2.1 *Gaussian white noise.*

The canonical example of a martingale random measure is given by the Gaussian process indexed by $\mathcal{A}=\{A\in\mathcal{B}(U)\times\mathcal{B}([0,\infty)):\mu\times m(A)<\infty\}$ and satisfying $E[W(A)]=0$ and $E[W(A)W(B)]=\mu\times m(A\cap B)$, where m denotes Lebesgue measure and μ is a σ-finite measure on U. If we define $M(A,t)=W(A\times[0,t])$ for $A\in\mathcal{B}(U)$, $\mu(A)<\infty$, and $t\geq 0$, then M is an orthogonal martingale random measure with $K(A\times B\times[0,t])=t\mu(A\cap B)$, and for fixed A, $M(A,\cdot)$ is a Brownian motion.

Example 2.2 *Poisson random measures.*

Let ν be a σ-finite measure on U and let $h(u)$ be in $L^2(\nu)$. Let N be a Poisson random measure on $U\times[0,\infty)$ with mean measure $\nu\times m$, that is, for $A\in\mathcal{B}(U)\times\mathcal{B}([0,\infty))$, $N(A)$ has a Poisson distribution with expectation $\nu\times m(A)$ and $N(A)$ and $N(B)$ are independent if $A\cap B=\emptyset$. For $A\in\mathcal{B}(U)$ satisfying $\nu(A)<\infty$, define $M(A,t)=\int_A h(u)(N(du\times[0,t])-\nu(du)t)$. Noting that $E[M(A,t)^2]=t\int_A h(u)^2\nu(du)$ and that $\{M(A_i,t)\}$ are independent for disjoint $\{A_i\}$, we can extend M to all of $\mathcal{B}(U)$ by addition.

Suppose Z is a process with independent increments with generator

$$Af(z) = \int_{\mathbb{R}} (f(z+u) - f(z) - uI_{\{|u|\leq 1\}} f'(z))\nu(du).$$

Then ν must satisfy $\int_{\mathbb{R}} u^2 \wedge 1\nu(du) < \infty$. (See, for example, Feller (1971).) Let $U = \mathbb{R}$, and let N be the Poisson random measure with mean measure $\nu \times m$. Define $M(A,t) = \int_A uI_{\{|u|\leq 1\}}(N(du \times [0,t]) - \nu(du)t)$, $V(A,t) = \int_A uI_{\{|u|>1\}}N(du \times [0,t])$, and $Y(A,t) = M(A,t) + V(A,t)$. Then we can represent Z by $Z(t) = Y(\mathbb{R}, t)$.

Consider a sequence of Poisson random measures with mean measures $n\nu \times m$. Define

$$M_n(A,t) = \frac{1}{\sqrt{n}} \int_A h(u)(N_n(du \times [0,t]) - nt\nu(du)). \tag{2.13}$$

Then M_n is an orthogonal martingale random measure with

$$\langle M_n(A), M_n(B)\rangle_t = t\int_{A\cap B} h(u)^2\nu(du) = K(A \times B \times [0,t]).$$

By the central limit theorem, M_n converges (in the sense of finite dimensional distributions) to the Gaussian white noise martingale random measure outlined in Example 2.1 with $\mu(A) = \int_A h(u)^2\nu(du)$.

Example 2.3 *Empirical measures.*

Let $\xi_1, \xi_2, \ldots$ be iid U-valued random variables with distribution μ, and define

$$M_n(A,t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} (I_A(\xi_i) - \mu(A)) \,. \tag{2.14}$$

Then $\langle M_n(A), M_n(B)\rangle_t = \frac{[nt]}{n}(\mu(A\cap B) - \mu(A)\mu(B))$. Note that $K_0(A \times B) = \mu(A \cap B) + \mu(A)\mu(B)$ extends to a measure on $U \times U$ and $K_n(A \times B \times (0,t]) = K_0(A \times B)\frac{[nt]}{n}$ extends to a measure on $U \times U \times [0,\infty)$ which will be a dominating measure for M_n. Of course, M_n converges to a Gaussian martingale random measure with conditional covariation $\langle M(A), M(B)\rangle_t = t(\mu(A\cap B) - \mu(A)\mu(B))$ and dominating measure $K_0 \times m$.

2.1 Moment estimates for martingale random measures.

Suppose that M is an orthogonal martingale measure. If $A, B \in \mathcal{A}$ are disjoint, then $[M(A), M(B)]_t = 0$ and in particular, $M(A,\cdot)$ and $M(B,\cdot)$ have a.s. no simultaneous discontinuities. It follows that

$$\Pi(A \times [0,t]) = [M(A)]_t$$

determines a random measure on $U \times [0,\infty)$ as does

$$\Pi_k(A \times [0,t]) = \sum_{s\leq t} (M(A,s) - M(A,s-))^k \tag{2.15}$$

for even $k > 2$. For odd $k > 2$, (2.15) determines a random signed measure. For X of the form (2.4), it is easy to check that

$$[X_- \cdot M]_t = \int_{U\times[0,t]} X^2(s-,u)\Pi(du\times ds),$$

and setting $Z = X_- \cdot M$ and letting $\Delta Z(s) = Z(s) - Z(s-)$, we have

$$\begin{aligned} Z^k(t) &= \int_0^t kZ^{k-1}(s-)dZ(s) + \int_0^t \frac{k(k-1)}{2} Z^{k-2}(s)d[Z]_s \\ &\quad + \sum_{s\le t}\left(Z^k(s) - Z^k(s-) - kZ^{k-1}(s-)\Delta Z(s) - \binom{k}{2} Z^{k-2}(s-)\Delta Z(s)^2\right) \\ &= \int_{U\times[0,t]} kZ^{k-1}(s-)X(s-,u)M(du\times ds) + \int_{U\times[0,t]} \binom{k}{2} Z^{k-2}(s)X^2(s-,u)\Pi(du\times ds) \\ &\quad + \sum_{j=3}^{k}\binom{k}{j}\int_{U\times[0,t]} Z^{k-j}(s-)X^j(s-,u)\Pi_j(du\times ds) \end{aligned} \tag{2.16}$$

and can be extended to more general X under appropriate conditions.

Since $M(A,\cdot)$ is locally square integrable, $[M(A)]_t$ is locally in L_1, that is, there exists a sequence of stopping times $\{\tau_n\}$ such that $\tau_n \to \infty$ and $E[[M(A)]_{t\wedge\tau_n}] < \infty$ for each $t > 0$ and each n. In addition,

$$[M(A)]_t - \langle M(A)\rangle_t = \Pi(A\times[0,t]) - \pi(A\times[0,t])$$

is a local martingale. It follows from (2.16) and L_2 approximation that

$$E[(X_-\cdot M(t))^2] = E[\int_{U\times[0,t]} X^2(s-,u)\Pi(du\times ds)] = E[\int_{U\times[0,t]} X^2(s-,u)\pi(du\times ds)] \tag{2.17}$$

whenever either the second or third expression is finite. (Note that the left side may be finite with the other two expressions infinite.) We would like to obtain similar expressions for higher moments.

A discrete time version of the following lemma can be found in Burkholder (1971), Theorem 20.2. The continuous time version was given by Lenglart, Lepingle, and Pratelli (1980) (see Dellacherie, Maisonneuve and Meyer (1992), page 326). The proof we give here is from Ichikawa (1986), Theorem 1.

Lemma 2.4 *For $0 < p \le 2$ there exists a constant C_p such that for any locally square integrable martingale M with Meyer process $\langle M\rangle$ and any stopping time τ*

$$E[\sup_{s\le\tau}|M(s)|^p] \le C_p E[\langle M\rangle_\tau^{p/2}]$$

Proof. For $p = 2$ the result is an immediate consequence of Doob's inequality. Let $0 < p < 2$. For $x > 0$, let $\sigma_x = \inf\{t : \langle M\rangle_t > x^2\}$. Since σ_x is predictable there exists an increasing sequence of stopping times $\sigma_x^n \to \sigma_x$. Noting that $\langle M\rangle_{\sigma_x^n} \le x^2$, we have

$$\begin{aligned} P\{\sup_{s\le\tau}|M(s)|^p > x\} &\le P\{\sigma_x^n \le \tau\} + \frac{E[\langle M\rangle_{\tau\wedge\sigma_x^n}]}{x^2} \\ &\le P\{\sigma_x^n \le \tau\} + \frac{E[x^2\wedge\langle M\rangle_\tau]}{x^2}, \end{aligned}$$

and letting $n \to \infty$, we have

$$P\{\sup_{s\leq\tau} |M(s)|^p > x\} \leq P\{\langle M\rangle_\tau \geq x^2\} + \frac{E[x^2 \wedge \langle M\rangle_\tau]}{x^2}. \tag{2.18}$$

Using the identity

$$\int_0^\infty E[x^2 \wedge X^2]px^{p-3}dx = \frac{2}{2-p}E[|X|^p],$$

the lemma follows by multiplying both sides of (2.18) by px^{p-1} and integrating. □

Assume that for $2 < k \leq k_0$ and $A \in \mathcal{A}$, $|\Pi_k|(A \times [0,\infty])$ is locally in L_1 and there exist predictable random measures π_k and $\hat{\pi}_k$ such that

$$\Pi_k(A \times [0,t]) - \pi_k(A \times [0,t]) \tag{2.19}$$

and

$$|\Pi_k|(A \times [0,t]) - \hat{\pi}_k(A \times [0,t]) \tag{2.20}$$

are local martingales. Of course, for k even, $\pi_k = \hat{\pi}_k$. We define $\pi_2 = \pi$

If M is Gaussian white noise as in Example 2.1, then $\Pi_k = \pi_k = 0$ for $k > 2$. If M is as in Example 2.2, then

$$\Pi_k(A \times [0,t]) = \int_A h^k(u)N(du \times [0,t]),$$

$$\pi_k(A \times [0,t]) = t\int_A h^k(u)\nu(du),$$

and

$$\hat{\pi}_k(A \times [0,t]) = t\int_A |h|^k(u)\nu(du).$$

Theorem 2.5 *Let $k \geq 2$, and suppose that for $2 \leq j \leq k$*

$$H_{k,j} \equiv E\left[\left(\int_{U\times[0,t]} |X(s-,u)|^j\hat{\pi}_j(du \times ds)\right)^{\frac{k}{j}}\right] < \infty. \tag{2.21}$$

Then $E[\sup_{s\leq t}|Z(s)|^k] < \infty$ and

$$E[Z^k(t)] = \sum_{j=2}^k \binom{k}{j} E\left[\int_{U\times[0,t]} Z^{k-j}(s-)X^j(s-,u)\pi_j(du \times ds)\right] \tag{2.22}$$

Proof. For $k = 2$, the result follows by (2.17). Note that if (2.21) holds, then it holds with k replaced by $k' < k$. Consequently, proceeding by induction, suppose that $E[\sup_{s\leq t}|Z(s)|^{k-1}] < \infty$. Since

$$M_k(t) = \int_0^t kZ^{k-1}(r-)dZ(r)$$

is a local square integrable martingale with

$$\langle M_k\rangle_t = \int_0^t k^2Z^{2k-2}(s-)d\langle Z\rangle_s,$$

by Lemma 2.4, for any stopping time τ

$$E[\sup_{s\le t\wedge\tau}|\int_0^s kZ^{k-1}(r-)dZ(r)|]\le C_1kE[\sqrt{\sup_{s<t\wedge\tau}|Z(s)|^{2k-k}\langle Z\rangle_{t\wedge\tau}}],$$

and letting $\tau_c=\inf\{t:|Z(t)|>c\}$, it follows that

$$\begin{aligned}E[\sup_{s\le t\wedge\tau_c}|Z(s)|^k] \;\le\;& C_1kE[\sqrt{\sup_{s<t\wedge\tau_c}|Z(s)|^{2k-k}\langle Z\rangle_t}] \\ &+\sum_{j=2}^k\binom{k}{j}E[\sup_{s<t\wedge\tau_c}|Z(s)|^{k-j}\int_{U\times[0,t]}|X(s-,u)|^j\hat{\pi}_j(du\times ds)]\end{aligned}\tag{2.23}$$

which by the Hölder inequality implies

$$\begin{aligned}E[\sup_{s\le t\wedge\tau_c}|Z(s)|^k] \quad&\le C_1kE[\sup_{s<t\wedge\tau_c}|Z(s)|^k]^{\frac{k-1}{k}}E[\left(\int_{U\times[0,t]}|X(s-,u)|^2\pi_2(du\times ds)\right)^{\frac{k}{2}}]^{\frac{1}{k}} \\ &+\sum_{j=2}^k\binom{k}{j}E[\sup_{s<t\wedge\tau_c}|Z(s)|^k]^{\frac{k-j}{k}}E[\left(\int_{U\times[0,t]}|X(s-,u)|^j\hat{\pi}_j(du\times ds)\right)^{\frac{k}{j}}]^{\frac{j}{k}}\end{aligned}\tag{2.24}$$

where the right side is finite by (2.21) and the fact that $E[\sup_{s<t\wedge\tau_c}|Z(s)|^k]\le c^k$. The inequality then implies that $E[\sup_{s<t\wedge\tau_c}|Z(s)|^k]\le K_k$, where K_k is the largest number satisfying

$$K\le C_1kK^{\frac{k-1}{k}}H_{k,2}^{\frac{1}{k}}+\sum_{j=2}^k\binom{k}{j}K^{\frac{k-j}{k}}H_{k,j}^{\frac{j}{k}}.\tag{2.25}$$

(2.22) then follows from (2.16). □

2.2 A convergence theorem for counting measures.

For $n=1,2,\ldots$, let N_n be a random counting measure on $U\times[0,\infty)$ with the property that $N_n(A\times\{t\})\le 1$ for all $A\in\mathcal{B}(U)$ and $t\ge 0$. Let ν be a σ-finite measure on U, and let $F_1\subset F_2\subset\cdots$ be closed sets such that $\nu(F_k)<\infty$, $\nu(\partial F_k)=0$, and $\nu(A)=\lim_{k\to\infty}\nu(A\cap F_k)$ for each $A\in\mathcal{B}(U)$. Let Λ_n be a random measure also satisfying $\Lambda_n(A\times\{t\})\le 1$. Suppose that Λ_n and N_n are adapted to $\{\mathcal{F}_t^n\}$ in the sense that $N_n(A\times[0,t])$ and $\Lambda_n(A\times[0,t])$ are $\mathcal{F}_t^n$-measurable for all $A\in\mathcal{B}(U)$ and $t\ge 0$, and suppose that

$$N_n(A\cap F_k\times[0,t])-\Lambda_n(A\cap F_k\times[0,t])$$

is a local $\{\mathcal{F}_t^n\}$-martingale for each $A\in\mathcal{B}(U)$ and $k=1,2,\ldots$.

Theorem 2.6 *Let N be the Poisson random measure on $U\times[0,\infty)$ with mean measure $\nu\times m$. Suppose that for each $k=1,2,\ldots$, $f\in\bar{C}(U)$, and $t\ge 0$*

$$\lim_{n\to\infty}\int_{F_k}f(u)\Lambda_n(du\times[0,t])=t\int_{F_k}f(u)\nu(du)$$

in probability. Then $N_n \Rightarrow N$ in the sense that for any $A_1, \ldots, A_m$ such that for each i, $A_i \subset F_k$ for some k and $\nu(\partial A_i) = 0$,

$$(N_n(A_1 \times [0,\cdot]), \ldots, N_n(A_m \times [0,\cdot])) \Rightarrow (N(A_1 \times [0,\cdot]), \ldots, N(A_m \times [0,\cdot])).$$

It also follows that for $f \in \bar{C}(U)$,

$$\int_{F_k} f(u) N_n(du \times [0,\cdot]) \Rightarrow \int_{F_k} f(u) N(du \times [0,\cdot]).$$

Proof. The result is essentially a theorem of Brown (1978). Alternatively, assuming $\cup_{i=1}^m A_i \subset F_k$, let

$$\tau_n = \inf\{t : \Lambda_n(F_k \times [0,t]) > t\nu(F_k) + 1\}.$$

Note that $\tau_n \to \infty$ and that

$$N_n(A_i \times [0, t \wedge \tau_n]) - \Lambda_n(A_i \times [0, t \wedge \tau_n])$$

is an $\{\mathcal{F}_t^n\}$-martingale. For $T > 0$ and $\delta > 0$, let

$$\gamma_T^n(\delta) = \sup_{t \le T} \Lambda_n(F_k \times (t \wedge \tau_n, (t+\delta) \wedge \tau_n])$$

and observe that $\lim_{\delta \to 0} \limsup E[\gamma_T^n(\delta)] = 0$. It follows that for $0 \le t \le T$

$$E[N_n(A_i \times (t \wedge \tau_n, (t+\delta) \wedge \tau_n] | \mathcal{F}_t^n] \le E[\gamma_T^n(\delta) | \mathcal{F}_t^n]$$

and the relative compactness of $\{(N_n(A_1 \times [0,\cdot]), \ldots, N_n(A_m \times [0,\cdot]))\}$ follows from Theorem 3.8.6 of Ethier and Kurtz (1986). The theorem then follows from Theorem 4.8.10 of Ethier and Kurtz (1986). □

In addition to the conditions of Theorem 2.6, we assume that there exists $h \in \bar{C}(U)$ with $0 \le h \le 1$ such that $\int_U (1 - h(u))\nu(du) < \infty$ and for $f \in \bar{C}(U)$,

$$\int_U f(u)(1 - h(u))\Lambda_n(du \times [0,t]) \to t \int_U f(u)(1 - h(u))\nu(du)$$

in probability. Let D be a linear space of functions on U such that for each k and each $\varphi \in D$

$$M_n^k(\varphi, t) = \int_{F_k} \varphi(u)h(u)(N_n(du \times [0,t]) - \Lambda_n(du \times [0,t]))$$

$$\hat{M}_n^k(\varphi, t) = \int_{F_k^c} \varphi(u)h(u)(N_n(du \times [0,t]) - \Lambda_n(du \times [0,t]))$$

and

$$M_n(\varphi, t) = \int_U \varphi(u)h(u)(N_n(du \times [0,t]) - \Lambda_n(du \times [0,t]))$$

are local $\{\mathcal{F}_t^n\}$-martingales and

$$\int_U \varphi^2(u)h^2(u)\nu(du) < \infty. \tag{2.26}$$

Theorem 2.7 *Suppose that there exists $\alpha : D \to [0,\infty)$ and a sequence $m_n \to \infty$ such that for every sequence $k_n \to \infty$ with $k_n \leq m_n$ and each $t \geq 0$,*

$$\lim_{n\to\infty} E[\sup_{s\leq t} |\hat{M}_n^{k_n}(\varphi,s) - \hat{M}_n^{k_n}(\varphi,s-)|] = 0 \tag{2.27}$$

and

$$[\hat{M}_n^{k_n}(\varphi)]_t \to \alpha(\varphi)t \tag{2.28}$$

in probability. Then for $\varphi_1,\ldots,\varphi_m \in D$,

$$(M_n(\varphi_1,t),\ldots,M_n(\varphi_m,t)) \Rightarrow (M(\varphi_1,t),\ldots,M(\varphi_m,t))$$

for

$$M(\varphi,t) = W(\varphi,t) + \int_U \varphi(u)h(u)\tilde{N}(du\times[0,t])$$

where W is a continuous (in t), mean zero, Gaussian processes satisfying

$$E[W(\varphi_1,s)W(\varphi_2,t)] = s\wedge t\frac{1}{2}(\alpha(\varphi_1+\varphi_2)-\alpha(\varphi_1)-\alpha(\varphi_2)),$$

$\tilde{N}(A\times[0,t]) = N(A\times[0,t]) - t\nu(A)$, and W is independent of N.

Remark 2.8 *Note that the linearity of D and (2.28) implies*

$$[\hat{M}_n^{k_n}(\varphi_1),\hat{M}_n^{k_n}(\varphi_2)]_t \to \frac{t}{2}(\alpha(\varphi_1+\varphi_2)-\alpha(\varphi_1)-\alpha(\varphi_2)). \tag{2.29}$$

(2.27) and (2.29) verify the conditions of the martingale central limit theorem (see, for example, Ethier and Kurtz (1986), Theorem 7.1.4) and it follows that

$$(\hat{M}_n^{k_n}(\varphi_1,\cdot),\ldots,\hat{M}_n^{k_n}(\varphi_m,\cdot)) \Rightarrow (W(\varphi_1,\cdot),\ldots,W(\varphi_m,\cdot)).$$

Suppose that $A_n^k(\varphi,t)$ has the property that

$$(\hat{M}_n^k(\varphi,t))^2 - A_n^k(\varphi,t)$$

is a local $\{\mathcal{F}_t^n\}$-martingale for each $\varphi\in D$ and that for m_n and k_n as above, we replace (2.27) and (2.28) by the requirements that

$$\lim_{n\to\infty} E[\sup_{s\leq t} |\hat{M}_n^{k_n}(\varphi,s) - \hat{M}_n^{k_n}(\varphi,s-)|^2] = 0 \tag{2.30}$$

$$\lim_{n\to\infty} E[\sup_{s\leq t} |A_n^{k_n}(\varphi,s) - A_n^{k_n}(\varphi,s-)|] = 0$$

and

$$A_n^{k_n}(\varphi,t) \to \alpha(\varphi)t \tag{2.31}$$

in probability. Then the conclusion of the theorem remains valid. In particular, (2.30) and (2.31) verify alternative conditions for the martingale central limit theorem. Note that if $\Lambda_n(A\cap F_k\times[0,\cdot])$ is continuous for each $A\in\mathcal{B}(U)$ and $k=1,2,\ldots$, then we can take

$$A_n^k(\varphi,t) = \int_{F_k^c}\varphi^2(u)h^2(u)\Lambda_n(du\times[0,t]). \tag{2.32}$$

Proof. For simplicity, let $m = 1$. For each fixed k, Theorem 2.6 implies

$$M_n^k(\varphi, \cdot) \Rightarrow \int_{F_k} \varphi(u)h(u)\tilde{N}(du \times [0, \cdot])$$

and it follows from (2.26) that

$$\lim_{k\to\infty} \int_{F_k} \varphi(u)h(u)\tilde{N}(du \times [0,t]) = \int_U \varphi(u)h(u)\tilde{N}(du \times [0,t]).$$

Consequently, for $k_n \to \infty$ sufficiently slowly,

$$M_n^{k_n}(\varphi, \cdot) \Rightarrow \int_U \varphi(u)h(u)\tilde{N}(du \times [0,t]),$$

and since we can assume that $k_n \leq m_n$, for the same sequence, the martingale central limit theorem implies $\hat{M}_n^{k_n}(\varphi, \cdot) \Rightarrow W(\varphi, \cdot)$. The convergence in $D_{\mathbb{R}}[0, \infty)$ of each component implies the relative compactness of $\{(M_n^{k_n}(\varphi, \cdot), \hat{M}_n^{k_n}(\varphi, \cdot))\}$ in $D_{\mathbb{R}}\{0, \infty) \times D_{\mathbb{R}}[0, \infty)$. The fact that the second component is asymptotically continuous implies relative compactness in $D_{\mathbb{R}^2}[0, \infty)$. Consequently, at least along a subsequence $(M_n^{k_n}(\varphi, \cdot), \hat{M}_n^{k_n}(\varphi, \cdot))$ converges in $D_{\mathbb{R}^2}[0, \infty)$. To see that there is a unique possible limit and hence that there is convergence along the original sequence it is enough to check that W and N are independent. To verify this assertion, check that $W(\varphi, \cdot)$, $\varphi \in D$, and $\tilde{N}(A \cap F_k \times [0, \cdot])$, $A \in \mathcal{B}(U)$, $k = 1, 2, \ldots$ are all martingales with respect to the same filtration. Since trivially, $[W(\varphi), \tilde{N}(A \cap F_k)]_t = 0$, an application of Itô's formula verifies that W and N give a solution of the martingale problem that uniquely determines their joint distribution and implies their independence. It follows that

$$(M_n^{k_n}(\varphi, \cdot), \hat{M}_n^{k_n}(\varphi, \cdot)) \Rightarrow (\int_U \varphi(u)h(u)\tilde{N}(du \times [0, \cdot]), W(\varphi, \cdot))$$

and hence

$$M_n(\varphi, \cdot) \Rightarrow M(\varphi, \cdot).$$

□

3 $H^{\#}$-semimartingales.

We will, in fact, consider more general stochastic integrals than those corresponding to semimartingale random measures. As in most definitions of an integral, the first step is to define the integral for a "simple" class of integrands and then to extend the integral to a larger class by approximation. Since we already know how to define the semimartingale integral in finite dimensions, a reasonable approach is to approximate arbitrary integrands by finite-dimensional integrands.

3.1 Finite dimensional approximations.

We will need the following lemma giving a partition of unity. $\bar{C}(S)$ denotes the space of bounded continuous functions on S with the sup norm.

Lemma 3.1 *Let (S,d) be a complete, separable metric space, and let $\{x_k\}$ be a countable dense subset of S. Then for each $\epsilon > 0$, there exists a sequence $\{\psi_k^\epsilon\} \subset \bar{C}(S)$ such that $supp\{\psi_k^\epsilon\} \subset B_\epsilon(x_k)$, $0 \leq \psi_k^\epsilon \leq 1$, $|\psi_k^\epsilon(x) - \psi_k^\epsilon(y)| \leq \frac{4}{\epsilon}d(x,y)$, and for each compact $K \subset S$, there exists $N_K < \infty$ such that $\sum_{k=1}^{N_K} \psi_k^\epsilon(x) = 1$, $x \in K$. In particular, $\sum_{k=1}^\infty \psi_k^\epsilon(x) = 1$ for all $x \in S$.*

Proof. Fix $\epsilon > 0$. Let $\tilde{\psi}_k(x) = (1 - \frac{2}{\epsilon}d(x, B_{\epsilon/2}(x_k)) \vee 0$. Then $0 \leq \tilde{\psi}_k \leq 1$, $\tilde{\psi}_k(x) = 1$, $x \in B_{\epsilon/2}(x_k)$, and $\tilde{\psi}_k(x) = 0$, $x \notin B_\epsilon(x_k)$. Note also that $|\tilde{\psi}_k(x) - \tilde{\psi}_k(y)| \leq \frac{2}{\epsilon}d(x,y)$. Define $\psi_1^\epsilon = \tilde{\psi}_1$, and for $k > 1$, $\psi_k^\epsilon = \max_{i\leq k} \tilde{\psi}_i - \max_{i\leq k-1} \tilde{\psi}_i$. Clearly, $0 \leq \psi_k^\epsilon \leq \tilde{\psi}_k$ and $\sum_{i=1}^k \psi_i^\epsilon = \max_{i\leq k} \tilde{\psi}_i$. In particular, for compact $K \subset S$, there exists $N_K < \infty$ such that $K \subset \cup_{k=1}^{N_K} B_{\epsilon/2}(x_k)$ and hence $\sum_{k=1}^{N_K} \psi_k^\epsilon(x) = 1$ for $x \in K$. Finally,

$$|\psi_k^\epsilon(x) - \psi_k^\epsilon(y)| \leq 2\max_{i\leq k} |\tilde{\psi}_i(x) - \tilde{\psi}_i(y)| \leq \frac{4}{\epsilon}d(x,y)$$

□

Let U be a complete, separable metric space, and let H be a Banach space of functions on U. Let $\{\varphi_k\}$ be a dense subset of H. Fix $\epsilon > 0$, and let $\{\psi_k^\epsilon\}$ be as in Lemma 3.1 with $S = H$ and $\{x_k\} = \{\varphi_k\}$. The role of the ψ_k^ϵ is quite simple. Let $x \in D_H[0,\infty)$, and define $x^\epsilon(t) = \sum_k \psi_k^\epsilon(x(t))\varphi_k$. Then

$$\|x(t) - x^\epsilon(t)\|_H \leq \sum_k \psi_k^\epsilon(x(t))\|x(t) - \varphi_k\|_H \leq \epsilon. \tag{3.1}$$

Since x is cadlag, for each $T > 0$, there exists a compact $K_T \subset H$ such that $x(t) \in K_T$, $0 \leq t \leq T$. Consequently, for each $T > 0$, there exists $N_T < \infty$ such that $x^\epsilon(t) = \sum_{k=1}^{N_T} \psi_k^\epsilon(x(t))\varphi_k$ for $0 \leq t \leq T$. This construction gives a natural way of approximating any cadlag H-valued function (or process) by cadlag functions (processes) that are essentially finite dimensional. Let Y be an $\{\mathcal{F}_t\}$-semimartingale random measure, and suppose $Y(\varphi, \cdot)$ is defined for all $\varphi \in H$ (or at least for a dense subset of φ). Let X be a cadlag, H-valued, $\{\mathcal{F}_t\}$-adapted process, and let

$$X^\epsilon(t) = \sum_k \psi_k^\epsilon(X(t))\varphi_k. \tag{3.2}$$

Then $\|X - X^\epsilon\|_H \leq \epsilon$, and the integral $X_-^\epsilon \cdot Y$ is naturally (and consistently with the previous section) defined to be

$$X_-^\epsilon \cdot Y(t) = \sum_k \int_0^t \psi_k^\epsilon(X(s-))dY(\varphi_k, s).$$

We can then extend the integral to all cadlag, adapted processes by taking the limit provided we can make the necessary estimates. This approach to the definition of the stochastic integral is similar to that taken by Mikulevicius and Rozovskii (1994).

3.2 Integral estimates.

Definition 3.2 *Let S be the collection of H-valued processes of the form*

$$Z(t) = \sum_{k=1}^m \xi_k(t)\varphi_k$$

where the ξ_k are $\mathbb{R}$-valued, cadlag, and adapted.

Suppose that $Y = M + V$ is standard and M has dominating measure K. Then for $Z \in \mathcal{S}$, we define

$$Z_- \cdot Y(t) = \sum_{k=1}^{m} \int_0^t \xi_k(s-) dY(\varphi_k, s). \tag{3.3}$$

As in the previous section, we have

$$\begin{aligned} &E[\sup_{s \le t} |Z_- \cdot M(s)|^2] \\ &\quad \le 4E\left[\int_{U \times U \times [0,t]} |Z(u, s-)||Z(v, s-))|K(du \times dv \times ds)\right], \end{aligned} \tag{3.4}$$

and, letting $|\tilde{V}|$ denote the total variation measure for the signed measure $\tilde{V}$,

$$E[\sup_{s \le t} |Z_- \cdot V(s)|] \le E[\int_{U \times [0,t]} |Z(s,u)||\tilde{V}|(du \times ds)]. \tag{3.5}$$

If, for example, the norm on H is the sup norm and $\|Z(s)\|_H \le \epsilon$ for all $s \ge 0$, then

$$E[\sup_{s \le t} |Z_- \cdot M(s)|^2] \le 4\epsilon^2 E[K(U \times U \times [0,t])] \tag{3.6}$$

and

$$E[\sup_{s \le t} |Z_- \cdot V(s)|] \le \epsilon E[|\tilde{V}|(U \times [0,t])]. \tag{3.7}$$

If $H = L^p(\mu)$, for some $p \ge 2$, and $\hat{K}$ defined as in (2.10), has the representation

$$\hat{K}(du \times dt) = h(u,t)\mu(du)dt \tag{3.8}$$

and

$$\tilde{V}(du \times dt) = g(u,t)\mu(du)dt, \tag{3.9}$$

then for r satisfying $\frac{2}{p} + \frac{1}{r} = 1$ and q satisfying $\frac{1}{p} + \frac{1}{q} = 1$, we have for $\|Z\|_H \le \epsilon$

$$\begin{aligned} &E[\sup_{s \le t} |Z_- \cdot M(s)|^2] \\ &\quad \le 4E\left[\int_0^t \left(\int_U |Z(s,u)|^p \mu(du)\right)^{\frac{2}{p}} \left(\int_U |h(u,s)|^r \mu(du)\right)^{\frac{1}{r}} ds\right] \\ &\quad \le 4\epsilon^2 E\left[\int_0^t \|h(\cdot,s)\|_{L^r(\mu)} ds\right] \end{aligned} \tag{3.10}$$

and

$$\begin{aligned} &E[\sup_{s \le t} |Z_- \cdot V(s)|] \\ &\quad \le E\left[\int_0^t \left(\int_U |Z(s,u)|^p \mu(du)\right)^{\frac{1}{p}} \left(\int_U |g(u,s)|^q \mu(du)\right)^{\frac{1}{q}} ds\right] \\ &\quad \le \epsilon E\left[\int_0^t \|g(\cdot,s)\|_{L^q(\mu)} ds\right]. \end{aligned} \tag{3.11}$$

Note that either (3.6) and (3.7) or (3.10) and (3.11) give an inequality of the form

$$E[\sup_{s\leq t}|Z_-\cdot Y(s)|]\leq \epsilon C(t) \tag{3.12}$$

which in turn implies

$$\mathcal{H}_t=\{\sup_{s\leq t}|Z_-\cdot Y(s)|\;:\;Z\in\mathcal{S},\;\sup_{s\leq t}\|Z(s)\|_H\leq 1\} \tag{3.13}$$

is stochastically bounded. The following lemma summarizes the estimates made above in a form that will be needed later.

Lemma 3.3 *a) Let $\|\cdot\|_H$ be the* sup *norm, and suppose $E[K(U\times U\times[0,t])]<\infty$ and $E[|\tilde{V}|(U\times[0,t])]<\infty$ for all $t>0$. Then if $\sup_s\|Z(s)\|_H\leq 1$ and τ is a stopping time bounded by a constant c,*

$$\begin{aligned}E[\sup_{s\leq t}&|Z_-\cdot Y(\tau+s)-Z_-\cdot Y(\tau)|]\\&\leq 2\sqrt{E[K(U\times U\times(\tau,\tau+t])]}+E[|\tilde{V}|(U\times[\tau,\tau+t])]\end{aligned}$$

and

$$\lim_{t\to 0}\sqrt{E[K(U\times U\times(\tau,\tau+t])]}+E[|\tilde{V}|(U\times[\tau,\tau+t])]=0.$$

b) Let $H=L^p(\mu)$, for some $p\geq 2$, and for h and g as in (3.10) and (3.11), suppose $E[\int_0^t\|h(\cdot,s)\|_{L^r(\mu)}ds]<\infty$ and $E[\int_0^t\|g(\cdot,s)\|_{L^q(\mu)}ds]<\infty$ for all $t>0$. Then if $\sup_s\|Z(s)\|_H\leq 1$ *and τ is a stopping time bounded by a constant c,*

$$E[\sup_{s\leq t}|Z_-\cdot Y(\tau+s)-Z_-\cdot Y(\tau)|]\leq 2\sqrt{E[\int_\tau^{\tau+t}\|h(\cdot,s)\|_{L^r(\mu)}ds]}+E[\int_\tau^{\tau+t}\|g(\cdot,s)\|_{L^q(\mu)}ds]]$$

and

$$\lim_{t\to 0}\sqrt{E[\int_\tau^{\tau+t}\|h(\cdot,s)\|_{L^r(\mu)}ds]}+E[\int_\tau^{\tau+t}\|g(\cdot,s)\|_{L^q(\mu)}ds]]=0.$$

Proof. The probability estimates follow from the moment estimates (3.6) - (3.11), and the limits follow by the dominated convergence theorem, using the fact that $\tau\leq c$. □

We will see that for many purposes we really do not need the moment estimates of Lemma 3.3. Consequently, it suffices to assume stochastic boundedness for $|\tilde{V}|$ and to localize the estimate on K.

Lemma 3.4 *a) Let $\|\cdot\|_H$ be the* sup *norm. Let τ be a stopping time, and let σ be a random variable such that $P\{\sigma>0\}=1$, $\tau+\sigma$ is a stopping time, $E[K(U\times U\times(\tau,\tau+\sigma]]<\infty$, and $P\{|\tilde{V}|(U\times(\tau,\tau+\sigma])<\infty\}=1$. Then if $\sup_s\|Z(s)\|_H\leq 1$ and $\alpha>0$,*

$$\begin{aligned}P\{&\sup_{s\leq t}|Z_-\cdot Y(\tau+s)-Z_-\cdot Y(\tau)|\geq 2\alpha\}\\&\leq\frac{\sqrt{E[K(U\times U\times(\tau,\tau+t\wedge\sigma])]}}{\alpha}+P\{|\tilde{V}|(U\times[\tau,\tau+t\wedge\sigma])\geq\alpha\}+P\{\sigma<t\}\end{aligned}$$

and the right side goes to zero as $t \to 0$.

b) Let $H = L^p(\mu)$, *for some* $p \geq 2$, *and let* h *and* g *be as in (3.8) and (3.9). Let* τ *be a stopping time, and let* σ *be a random variable such that* $P\{\sigma > 0\} = 1$, $\tau + \sigma$ *is a stopping time,* $E[\int_\tau^{\tau+\sigma} \|h(\cdot,s)\|_{L^r(\mu)} ds] < \infty$ *and* $P\{\int_\tau^{\tau+\sigma} \|g(\cdot,s)\|_{L^q(\mu)} ds < \infty\} = 1$. *Then if* $\sup_s \|Z(s)\|_H \leq 1$

$$
\begin{aligned}
&P\{\sup_{s\leq t} |Z_- \cdot Y(\tau+s) - Z_- \cdot Y(\tau)| \geq 2\alpha\} \\
&\quad \leq \frac{\sqrt{E[\int_\tau^{\tau+t\wedge\sigma} \|h(\cdot,s)\|_{L^r(\mu)} ds]}}{\alpha} + P\{\int_\tau^{\tau+t\wedge\sigma} \|g(\cdot,s)\|_{L^q(\mu)} ds \geq \alpha\} + P\{\sigma < t\}
\end{aligned}
$$

and the right side goes to zero as $t \to 0$.

Proof. Observe that

$$
\begin{aligned}
&P\{\sup_{s\leq t} |Z_- \cdot Y(\tau+s) - Z_- \cdot Y(\tau)| \geq 2\alpha\} \\
&\quad \leq P\{\sup_{s\leq t\wedge\sigma} |Z_- \cdot M(\tau+s) - Z_- \cdot M(\tau)| \geq \alpha\} \\
&\qquad + P\{\sup_{s\leq t\wedge\sigma} |Z_- \cdot V(\tau+s) - Z_- \cdot V(\tau)| \geq \alpha\} + P\{\sigma < t\},
\end{aligned}
$$

and note that the first two terms on the right are bounded by the corresponding terms in the desired inequalities. □

3.3 $H^\#$-semimartingale integrals.

Now let H be an arbitrary, separable Banach space. With the above development in mind, we make the following definition.

Definition 3.5 *Y is an $\{\mathcal{F}_t\}$-adapted, $H^\#$-*semimartingale, *if Y is an $\mathbb{R}$-valued stochastic process indexed by $H \times [0,\infty)$ such that*

- *For each $\varphi \in H$, $Y(\varphi,\cdot)$ is a cadlag $\{\mathcal{F}_t\}$-semimartingale with $Y(\varphi,0) = 0$.*
- *For each $t \geq 0$, $\varphi_1,\ldots,\varphi_m \in H$, and $a_1,\ldots,a_m \in \mathbb{R}$, $Y(\sum_{i=1}^m a_i\varphi_i, t) = \sum_{i=1}^m a_i Y(\varphi_i,t)$ a.s.*

The definition of the integral in (3.3) extends immediately to this more general setting. Noting (3.12), (3.13), and their relationship to the assumption that the semimartingale measure is standard, we define:

Definition 3.6 *Y is a* standard *$H^\#$-semimartingale if $\mathcal{H}_t$ defined in (3.13) is stochastically bounded for each t.*

This stochastic boundedness is implied by an apparently weaker condition.

Definition 3.7 *Let $\mathcal{S}_0 \subset \mathcal{S}$ be the collection of processes*

$$Z(t) = \sum_{k=1}^{m} \xi_k(t)\varphi_k$$

in which the ξ_k are piecewise constant, that is,

$$\xi_k(t) = \sum_{i=0}^{j} \eta_i^k I_{[\tau_i^k, \tau_{i+1}^k)}(t)$$

where $0 = \tau_0^k \leq \cdots \leq \tau_j^k$ are $\{\mathcal{F}_t\}$-stopping times and η_i^k is $\mathcal{F}_{\tau_i^k}$-measurable.

Lemma 3.8 *If*

$$\mathcal{H}_t^0 = \{|Z_- \cdot Y(t)| : Z \in \mathcal{S}_0, \sup_{s \leq t} \|Z(s)\|_H \leq 1\}$$

is stochastically bounded, then $\mathcal{H}_t$ defined in (3.13) is stochastically bounded.

Remark 3.9 *If Y is real-valued, that is $H = \mathbb{R}$, then the definition of standard $H^\#$-semimartingale is equivalent to the definition of semimartingale given in Section II.1 of Protter (1990), that is, the process satisfies the "good integrator" condition.*

Proof. For each $\delta > 0$, there exists $K(t, \delta)$ such that

$$P\{|Z_- \cdot Y(t)| \geq K(t, \delta)\} \leq \delta \tag{3.14}$$

for all $Z \in \mathcal{S}_0$ satisfying $\sup_{s \leq t} \|Z(s)\|_H \leq 1$. We can assume, without loss of generality, that $K(t, \delta)$ is right continuous and strictly increasing in δ (so that the collection of random variables satisfying $P\{U \geq K(t, \delta)\} \leq \delta$ is closed under convergence in probability). Let $\tau = \inf\{s : |Z_- \cdot Y| \geq K(t, \delta)\}$ and $Z^\tau = I_{[0,\tau)} Z$. Then

$$P\{\sup_{s \leq t} |Z_- \cdot Y(s)| \geq K(t, \delta)\} = P\{|Z_- \cdot Y(t \wedge \tau)| \geq K(t, \delta)\} = P\{|Z_-^\tau \cdot Y(t)| \geq K(t, \delta)\} \leq \delta\,.$$

For $Z \in \mathcal{S}$ with $\sup_{s \leq t} \|Z(s)\|_H \leq 1$, there exists a sequence $\{Z_n\} \subset \mathcal{S}_0$ with $\sup_{s \leq t} \|Z_n(s)\| \leq 1$ such that $\sup_{s \leq t} \|Z(s) - Z_n(s)\|_H \to 0$. This convergence implies $Z_{n-} \cdot Y(t) \to Z_- \cdot Y$ in probability, and the lemma follows. □

The assumption of Lemma 3.8 holds if there exists a constant $C(t)$ such that for all $Z \in \mathcal{S}_0$ satisfying $\sup_{s \leq t} \|Z(s)\|_H \leq 1$,

$$E[|Z_- \cdot Y(t)|] \leq C(t).$$

The following lemma is essentially immediate. The observation it contains is useful in treating semimartingale random measures which can frequently be decomposed into a part (usually a martingale random measure) that determines an $H^\#$-semimartingale on an L^2 space and another part that determines an $H^\#$-semimartingale on an L^1-space. Note that if H_1 is a space of functions on U_1 and H_2 is a space of functions on U_2, then $H_1 \times H_2$ can be interpreted as a space of functions on $U = U_1 \cup U_2$ where, for example, $\mathbb{R} \cup \mathbb{R}$ is interpreted as the set consisting of two copies of $\mathbb{R}$.

Lemma 3.10 *Let Y_1 be a standard $H_1^\#$-semimartingale and Y_2 a standard $H_2^\#$-semimartingale with respect to the same filtration $\{\mathcal{F}_t\}$. Define $H = H_1 \times H_2$, with $\|\varphi\|_H = \|\varphi_1\|_{H_1} + \|\varphi_2\|_{H_2}$ and $Y(\varphi, t) = Y_1(\varphi_1, t) + Y_2(\varphi_2, t)$ for $\varphi = (\varphi_1, \varphi_2)$. Then Y is a standard $H^\#$-semimartingale.*

If Y is standard, then the definition of $Z_- \cdot Y$ can be extended to all cadlag, H-valued processes.

Theorem 3.11 *Let Y be a standard $H^\#$-semimartingale, and let X be a cadlag, adapted, H-valued process. Define X^ϵ by (3.2). Then*

$$X_- \cdot Y \equiv \lim_{\epsilon \to 0} X_-^\epsilon \cdot Y$$

exists in the sense that for each $t > 0$ $\lim_{\epsilon \to 0} P\{\sup_{s \le t} |X_- \cdot Y(s) - X_-^\epsilon \cdot Y(s)| > \eta\} = 0$ for all $\eta > 0$, and $X_- \cdot Y$ is cadlag.

Proof. Let $K(\delta, t) > 0$ be as in Lemma 3.8. Since $\|X^{\epsilon_1}(s) - X^{\epsilon_2}(s)\|_H \le \epsilon_1 + \epsilon_2$, we have that $P\{\sup_{s \le t} |X_-^{\epsilon_1} \cdot Y(s) - X_-^{\epsilon_2} \cdot Y(s)| \ge (\epsilon_1 + \epsilon_2) K(\delta, t)\} \le \delta$, and it follows that $\{X_-^\epsilon \cdot Y\}$ is Cauchy in probability and that the desired limit exists. Since $X_-^\epsilon \cdot Y$ is cadlag and the convergence is uniform on bounded intervals, it follows that $X_- \cdot Y$ is cadlag. □

The following corollary is immediate from the definition of the integral.

Corollary 3.12 *Let Y be a standard $H^\#$-semimartingale, and let X be a cadlag, adapted, H-valued process. Let τ be an $\{\mathcal{F}_t\}$-stopping time and define $X^\tau = I_{[0,\tau)} X$. Then*

$$X_- \cdot Y(t \wedge \tau) = X_-^\tau \cdot Y(t) .$$

For finite dimensional semimartingale integrals, the stochastic integral for cadlag, adapted integrands can be defined as a Riemann-type limit of approximating sums

$$X_- \cdot Y(t) = \lim \sum X(t_i \wedge t)(Y(t_{i+1} \wedge t) - Y(t_i \wedge t)) \tag{3.15}$$

where the limit is as $\max(t_{i+1} - t_i) \to 0$ for the partition of $[0, \infty)$, $0 = t_0 < t_1 < t_2 < \cdots$. Formally, the analogue for $H^\#$-semimartingale integrals would be

$$X_- \cdot Y(t) = \lim \sum (Y(X(t_i \wedge t), t_{i+1} \wedge t) - Y(X(t_i \wedge t), t_i \wedge t));$$

however, $Y(X, t)$ is not, in general, defined for random variables X. We can define an analog of the summands in (3.15) by first defining $X^{[t_i, t_{i+1})} = I_{[t_i, t_{i+1})} X(t_i)$ and then defining

$$\Delta_{[t_i, t_{i+1})} Y(X(t_i), t) = X_-^{[t_i, t_{i+1})} \cdot Y(t).$$

Similarly, we can define $\Delta_{[\tau_i, \tau_{i+1})} Y(X(\tau_i), t)$ for stopping times $\tau_i \le \tau_{i+1}$.

Proposition 3.13 *For each n let $\{\tau_i^n\}$ be an increasing sequence of stopping times $0 = \tau_0^n \leq \tau_1^n \leq \tau_2^n \leq \cdots$, and suppose that for each $t > 0$*

$$\lim_{n\to\infty} \max\{\tau_{i+1}^n - \tau_i^n : \tau_i^n < t\} = 0.$$

If X is a cadlag, adapted H-valued process and Y is a standard $H^{\#}$-semimartingale, then

$$X_- \cdot Y(t) = \lim_{n\to\infty} \sum \Delta_{[\tau_i^n, \tau_{i+1}^n)} Y(X(\tau_i^n), t) \tag{3.16}$$

where the convergence is uniform on bounded time intervals in probability.

Proof. Let

$$X_n = \sum I_{[\tau_i^n, \tau_{i+1}^n)} X(\tau_i^n).$$

Then the sum on the right of (3.16) is just $X_{n-} \cdot Y(t)$ and, with X_n^ϵ defined as above,

$$X_{n-}^\epsilon \cdot Y(t) = \sum_k \sum_i \psi_k^\epsilon(X(\tau_i^n))(Y(\varphi_k, \tau_{i+1}^n \wedge t) - Y(\varphi_k, \tau_i^n \wedge t)).$$

By the finite dimensional result, for each $\eta > 0$,

$$\lim_{n\to\infty} P\{\sup_{s\leq t} |X_{n-}^\epsilon \cdot Y(s) - X_-^\epsilon \cdot Y(s)| > \eta\} = 0$$

and by standardness

$$P\{\sup_{s\leq t} |X_{n-}^\epsilon \cdot Y(s) - X_{n-} \cdot Y(s)| \geq \epsilon K(\delta, t)\} + P\{\sup_{s\leq t} |X_-^\epsilon \cdot Y(s) - X_- \cdot Y(s)| \geq \epsilon K(\delta, t)\} \leq 2\delta$$

and the result follows. □

3.4 Predictable integrands.

Definition 3.14 *Let $\mathcal{S}^*$ be the collection of H-valued processes of the form*

$$Z(t) = \sum_{k=1}^m \xi_k(t)\varphi_k$$

where the ξ_k are $\{\mathcal{F}_t\}$-predictable and bounded.

$Z_- \cdot Y$ for $Z \in \mathcal{S}$, can be extended to $Z \in \mathcal{S}^*$ by setting

$$Z \cdot Y(t) = \sum_{k=1}^m \int_0^t \xi_k(s) dY(\varphi_k, s).$$

We will show that the condition that $\mathcal{H}_t$ is stochastically bounded implies that

$$\mathcal{H}_t^* = \{\sup_{s\leq t} |Z \cdot Y(s)| \ : \ Z \in \mathcal{S}^*, \sup_{s\leq t} \|Z(s)\|_H \leq 1\}$$

is also stochastically bounded, and hence $Z \cdot Y$ can be extended to all predictable, H-valued processes X that satisfy a compact range condition by essentially the same argument as in the proof of Theorem 3.11.

Lemma 3.15 *If $\mathcal{H}_t$ is stochastically bounded, then $\mathcal{H}_t^*$ is stochastically bounded.*

Proof. Let $K(t,\delta)$ be as in (3.14). Fix $\varphi_1,\ldots,\varphi_m \in H$ and let $C_\varphi = \{x \in \mathbb{R}^m : \|\sum_{i=1}^m x_i\varphi_i\|_H \le 1\}$. To simplify notation, let $Y_i(t) = Y(\varphi_i,t)$. We need to show that if $(\xi_1,\ldots,\xi_m)$ is predictable and takes values in C_φ, then

$$P\{\sup_{s\le t}|\sum_{i=1}^m \int_0^s \xi_i(s)dY_i(s)| \ge K(t,\delta)\} \le \delta. \tag{3.17}$$

Consequently, it is enough to show that there exists cadlag, adapted ξ_i^n such that $(\xi_1^n, \ldots, \xi_m^n) \in C_\varphi$ and $\lim_{n\to\infty}\sup_{s\le t}|\int_0^s(\xi_i(u)-\xi_i^n(u-))dY_i(u)| = 0$ in probability for each i. Assume for the moment that $Y_i = M_i + A_i$ where M_i is a square integrable martingale and $E[T_t(A_i)] < \infty$, $T_t(A_i)$ denoting the total variation up to time t. Let $\Gamma(t) = \sum_{i=1}^m [M_i]_t$ and $\Lambda(t) = \sum_{i=1}^m T_t(A_i)$. Then (see Protter (1990), Theorem IV.2) there exists a sequence of cadlag, adapted $\mathbb{R}^m$-valued processes $\tilde{\xi}^n$ such that

$$\lim_{n\to\infty}\left(\sqrt{E[\int_0^t |\xi(s)-\tilde{\xi}^n(s-)|^2 d\Gamma(s)]} + E[\int_0^t |\xi(s)-\tilde{\xi}^n(s-)|d\Lambda(s)]\right) = 0. \tag{3.18}$$

Letting π denote the projection onto the convex set C_φ, since $|\pi(x)-\pi(y)| \le |x-y|$ and $\xi \in C_\varphi$, if we define $\xi^n = \pi(\tilde{\xi}^n)$, $\xi^n \in C_\varphi$ and the limit in (3.18) still holds. Finally,

$$\begin{aligned}
&E[\sup_{s\le t}|\int_0^s (\xi_i(u)-\xi_i^n(u-))dY_i(u)|] \\
&\le 2\sqrt{E[\int_0^t |\xi_i(s)-\xi_i^n(s-)|^2 d[M_i]_s]} + E[\int_0^t |\xi_i(s)-\xi_i^n(s-)|dT_s(A_i)] \\
&\le 2\sqrt{E[\int_0^t |\xi(s)-\xi^n(s-)|^2 d\Gamma(s)]} + E[\int_0^t |\xi(s)-\xi^n(s-)|d\Lambda(s)]
\end{aligned}$$

so the stochastic integrals converge and the limit must satisfy (3.17).

For general Y_i, fix $\epsilon > 0$, and let $Y_i = M_i + A_i$ where M_i is a local martingale with discontinuities bounded by ϵ, that is, $\sup_t |M_i(t) - M_i(t-)| \le \epsilon$, and A_i is a finite variation process. (Such a decomposition always exists. See Protter (1990), Theorem III.13.) For $c > 0$, let $\tau_c = \inf\{t : \sum_{i=1}^m [M_i]_t + \sum_{i=1}^m T_t(A_i) \ge c\}$, and let $Y_i^{\tau_c} = Y_i(\cdot\wedge\tau_c)$. Then for cadlag, adapted ξ with values in C_φ, it still holds that

$$P\{\sup_{s\le t}|\sum_{i=1}^m \int_0^s \xi_i(s-)dY_i^{\tau_c}(s)| \ge K(t,\delta)\} \le \delta. \tag{3.19}$$

(replace ξ by $I_{[0,\tau_c)}\xi$). Define $\tilde{Y}_i^{\tau_c} = M_i^{\tau_c} + A_i^{\tau_c-}$ where $A_i^{\tau_c-}(t) = A_i(t)$, for $t < \tau_c$ and $A_i^{\tau_c-}(t) = A_i(\tau_c-)$ for $t \ge \tau_c$. It follows from (3.19) and the fact that

$$|\sum_{i=1}^m \xi_i(\tau_c-)(M_i(\tau_c)-M_i(\tau_c-))| \le \epsilon \sup_{x\in C_\varphi}\sum_{i=1}^m |x_i| \equiv \epsilon K_\varphi$$

that

$$P\{\sup_{s\leq t}|\sum_{i=1}^{m}\int_0^s \xi_i(s-)d\tilde{Y}_i^{\tau_c}(s)| \geq K(t,\delta)+\epsilon K_\varphi\} \leq \delta. \tag{3.20}$$

for all cadlag, adapted processes with values in C_φ. But $M_i^{\tau_c}$ is a square integrable martingale and $T_t(A_i^{\tau_c-}) \leq c$, so it follows that (3.20) holds with $\xi(s-)$ replaced by an arbitrary predictable process with values in C_φ. Letting $c \to \infty$ and observing that ϵ is arbitrary, we see that (3.17) holds and the lemma follows. □

Proposition 3.16 *Let Y be a standard $H^\#$-semimartingale, and let X be an H-valued predictable process such that for $t, \eta > 0$ there exists compact $K_{t,\eta} \subset H$ satisfying*

$$P\{X(s) \in K_{t,\eta},\ s \leq t\} \geq 1-\eta.$$

Then defining X^ϵ as in (3.2), $X \cdot Y \equiv \lim_{\epsilon\to 0} X^\epsilon \cdot Y$ exists.

Remark 3.17 *If estimates of the form (3.4) hold, then the definition of $X \cdot Y$ can be extended to locally bounded X, that is, the compact range condition can be dropped. (Approximate X be processes of the form $XI_K(X)$ where K is compact.) We do not know whether or not the compact range condition can be dropped for every standard $H^\#$-semimartingale.*

Proof. The proof is the same as for Theorem 3.11. □

3.5 Examples.

The idea of an $H^\#$-semimartingale is intended to suggest, but not be equivalent to, the idea of an H^*-semimartingale, that is, a semimartingale with values in H^*. In deed, any H^*-semimartingale will be an $H^\#$-semimartingale; however, there are a variety of examples of $H^\#$-semimartingales that are not H^*-semimartingales.

Example 3.18 *Poisson process integrals in L^p spaces.*

Let μ be a finite measure on U, and let $H = L^p(\mu)$ for some $1 \leq p < \infty$. Let N be a Poisson point process on $U \times [0,\infty)$, and for $\varphi \in H$, define $Y(\varphi,t) = \int_{U\times[0,t]} \varphi(u)N(du \times ds)$. Of course $Y(\varphi,\cdot)$ is just a compound Poisson process whose jumps have distribution given by $\nu(A) = \int I_A(\varphi(u))\mu(du)$. Since point evaluation is not a continous linear functional on L^p, Y is an $H^\#$-semimartingale, but not an H^*-semimartingale.

Example 3.19 *Cylindrical Brownian motion.*

Let H be a Hilbert space and let Q be a bounded, self-adjoint, nonnegative operator on H. Then there exists a Gaussian process W with covariance

$$E[W(\varphi_1,t)W(\varphi_2,s)] = t \wedge s\langle Q\varphi_1, \varphi_2\rangle.$$

If Q is nuclear, then W will be an $H^*(=H)$-valued martingale; however, in general, W will only be an $H^\#$-semimartingale. (See, for example, Da Prato and Zabczyk (1992), Section

4.3.) Note that if $X(t) = \sum_{i=1}^m \xi_i(t)\varphi_i$ is cadlag and adapted to the filtration generated by W, then

$$E[|X_- \cdot W(t)|^2] = \int_0^t E[\sum_{i,j} \xi_i(s)\xi_j(s)\langle Q\varphi_i, \varphi_j\rangle]ds \leq \|Q\| \int_0^t E[\|X(s)\|_H^2]ds$$

and it follows that W is standard.

4 Convergence of stochastic integrals.

Let H be a separable Banach space, and for each $n \geq 1$, let Y_n be an $\{\mathcal{F}_t^n\}$-$H^{\#}$-semimartingale. Note that the Y_n need not all be adapted to the same filtration nor even defined on the same probablity space. The minimal convergence assumption that we will consider is that for $\varphi_1, \ldots, \varphi_m \in H$, $(Y_n(\varphi_1, \cdot), \ldots, Y_n(\varphi_m, \cdot)) \Rightarrow (Y(\varphi_1, \cdot), \ldots, Y(\varphi_m, \cdot))$ in $D_{\mathbb{R}^m}[0, \infty)$ with the Skorohod topology.

Let $\{X_n\}$ be cadlag, H-valued processes. We will say that $(X_n, Y_n) \Rightarrow (X, Y)$, if $(X_n, Y_n(\varphi_1, \cdot), \ldots, Y_n(\varphi_m, \cdot)) \Rightarrow (X, Y(\varphi_1, \cdot), \ldots, Y(\varphi_m, \cdot))$ in $D_{H\times\mathbb{R}^m}[0, \infty)$ for each choice of $\varphi_1, \ldots, \varphi_m \in H$. We are interested in conditions on $\{(X_n, Y_n)\}$, under which $X_{n-} \cdot Y_n \Rightarrow X_- \cdot Y$. In the finite dimensional setting of Kurtz and Protter (1991a), convergence was obtained by first approximating by piecewise constant processes. This approach was also taken by Cho (1994, 1995) for integrals driven by martingale random measures. Here we take a slightly different approach, approximating the H-valued processes by finite dimensional H-valued processes in a way that allows us to apply the results of Kurtz and Protter (1991a) and Jakubowski, Mémin, and Pages, G. (1989).

Lemma 4.1 *Suppose that for each $\varphi \in H$, the sequence $\{Y_n(\varphi, \cdot)\}$ of $\mathbb{R}$-valued semimartingales satisfies the conditions of Theorem 1.1. Let $X_n^\epsilon(t) \equiv \sum_k \psi_k^\epsilon(X_n(t))\varphi_k$. If $(X_n, Y_n) \Rightarrow (X, Y)$, then $(X_n, Y_n, X_{n-}^\epsilon \cdot Y_n) \Rightarrow (X, Y, X_-^\epsilon \cdot Y)$. If $(X_n, Y_n) \to (X, Y)$ in probability, then $(X_n, Y_n, X_{n-}^\epsilon \cdot Y_n) \to (X, Y, X_-^\epsilon \cdot Y)$ in probability.*

Proof. By tightness, there exists a sequence of compact $K_m \subset H$ such that $P\{X_n(t) \in K_m, t \leq m\} \geq 1 - \frac{1}{m}$. Let $\tau_n^m = \inf\{t : X_n(t) \notin K_m\}$. Then $P\{\tau_n^m \geq m\} \geq 1 - \frac{1}{m}$, and $X_{n-}^\epsilon \cdot Y_n(t) = Z_n^m(t) \equiv \sum_{k=1}^{N_{K_m}} \int_0^t \psi_k^\epsilon(X_n(s-))dY_n(\varphi_k, s)$ for $t < \tau_n^m$. Theorem 1.1 implies $(X_n, Y_n, Z_n^m) \Rightarrow (X, Y, Z^m)$ for each m, where $Z^m(t) \equiv \sum_{k=1}^{N_{K_m}} \int_0^t \psi_k^\epsilon(X(s-))dY(\varphi_k, s)$. Since $Z_n^m(t) = X_{n-}^\epsilon \cdot Y_n(t)$ for $t \leq \tau_n^m$, using the metric of Ethier and Kurtz (1986), Chapter 3, Formula (5.2), we have $d(X_{n-}^\epsilon \cdot Y_n, Z_n^m) \leq e^{-\tau_n^m}$, and the convergence of $\{Z_n^m\}$ for each m implies the desired convergence for $X_{n-}^\epsilon \cdot Y_n$. □

In order to prove the convergence for $X_{n-} \cdot Y_n$, by Lemma 4.1, it is enough to show that $X_{n-}^\epsilon \cdot Y_n$ is a good approximation of $X_{n-} \cdot Y_n$, that is, we need to estimate $(X_{n-} - X_{n-}^\epsilon) \cdot Y_n$. If the Y_n correspond to semimartingale random measures, then (3.6) and (3.7) or (3.10) and (3.11) give estimates of the form

$$E[\sup_{s\leq t} |(X_{n-} - X_{n-}^\epsilon) \cdot Y(s)|] \leq \epsilon C_n(t).$$

If $\sup_n C_n(t) < \infty$ for each t, then defining

$$\mathcal{H}_{n,t} = \{\sup_{s \le t} |Z_- \cdot Y_n(s)| \;:\; Z \in \mathcal{S}^n, \sup_{s \le t} \|Z(s)\|_H \le 1\}, \tag{4.1}$$

$\hat{\mathcal{H}}_t = \cup_n \mathcal{H}_{n,t}$ is stochastically bounded for each t. This last assertion is essentially the *uniform tightness* (UT) condition of Jakubowski, Mémin, and Pages (1989). As in Lemma 3.8, the condition that $\cup_n \mathcal{H}_{n,t}$ be stochastically bounded is equivalent to the condition that

$$\hat{\mathcal{H}}_t^0 = \cup_n \mathcal{H}_{n,t}^0 = \cup_n \{|Z_- \cdot Y_n(t)| : Z \in \mathcal{S}_0^n, \sup_{s \le t} \|Z(s)\|_H\} \tag{4.2}$$

be stochastically bounded.

For finite dimensional semimartingales, the uniform tightness condition, Condition UT of Theorem 1.1, implies uniformly controlled variations, Condition UCV of Theorem 1.1. In the present setting, the relationship of the UT condition and some sort of "uniform worthiness" is not clear and certainly not so simple. Consequently, the following theorem is really an extension of the convergence theorem of Jakubowski, Mémin, and Pages (1989) rather than the results of Kurtz and Protter (1991a), although in the finite dimensional setting of those results, the conditions of the two theorems are equivalent.

Theorem 4.2 *For each $n = 1, 2, \ldots$, let Y_n be a standard $\{\mathcal{F}_t^n\}$-adapted, $H^\#$-semimartingale. Let $\mathcal{H}_{n,t}^0$ and $\hat{\mathcal{H}}_t^0$ be defined as in (4.2), and suppose that for each $t > 0$, $\hat{\mathcal{H}}_t^0$ is stochastically bounded. If $(X_n, Y_n) \Rightarrow (X, Y)$, then there is a filtration $\{\mathcal{F}_t\}$, such that Y is an $\{\mathcal{F}_t\}$-adapted, standard, $H^\#$-semimartingale, X is $\{\mathcal{F}_t\}$-adapted and $(X_n, Y_n, X_{n-} \cdot Y_n) \Rightarrow (X, Y, X_- \cdot Y)$. If $(X_n, Y_n) \to (X, Y)$ in probability, then $(X_n, Y_n, X_{n-} \cdot Y_n) \to (X, Y, X_- \cdot Y)$ in probability.*

Remark 4.3 *a) One of the motivations for introducing $H^\#$-semimartingales rather than simply posing the above result in terms of semimartingale random measures is that the Y_n may be given by standard semimartingale random measures while the limiting Y is not.*

b) Jakubowski (1995) proves the above theorem in the case of Hilbert space-valued semimartingales.

Proof. The stochastic boundedness condition implies that for each $t, \delta > 0$ there exists $K(\delta, t)$ such that for all $R \in \hat{\mathcal{H}}_t$, $P\{|R| \ge K(\delta, t)\} \le \delta$. Without loss of generality, we can assume that $K(\delta, t)$ is a nondecreasing, right continuous function of t. Note that this inequality will hold for $R = \sup_{s \le t} |Z_{n-} \cdot Y_n(s)|$ for any cadlag, $\{\mathcal{F}_t^n\}$-adapted Z_n satisfying $\sup_{s \le t} \|Z_n(s)\|_H \le 1$.

Let $\mathcal{F}_t = \sigma(X(s), Y(\varphi, s) : s \le t, \varphi \in H)$. Define

$$Z_n(t) = \sum_{i=1}^m f_i(X_n, Y_n(\hat{\varphi}_1, \cdot), \ldots, Y_n(\hat{\varphi}_d, \cdot), t)\varphi_i$$

where $(f_1, \ldots, f_m)$ is a continous function mapping $D_{H \times \mathbb{R}^d}[0, \infty)$ into $C_{A(\varphi_1, \ldots, \varphi_m)}[0, \infty)$, $A(\varphi_1, \ldots, \varphi_m) = \{\alpha \in \mathbb{R}^m : \|\sum_i \alpha_i \varphi_i\|_H \le 1\}$, in such a way that $f_i(x, y_1, \ldots, y_d, t)$ depends only on $(x(s), y_1(s), \ldots, y_d(s))$ for $s \le t$ (ensuring that Z_n is $\{\mathcal{F}_t^n\}$-adapted and $Z =$

$\sum_{i=1}^m f_i(X, Y(\hat{\varphi}_1,\cdot),\ldots,Y(\hat{\varphi}_d,\cdot))\varphi_i$ is $\{\mathcal{F}_t\}$-adapted). Theorem 1.1 implies $Z_{n-}\cdot Y_n \Rightarrow Z_-\cdot Y$, and it follows (using the right continuity of $K(\delta,t)$) that

$$P\{\sup_{s\leq t}|Z_-\cdot Y|\geq K(\delta,t)\}\leq\delta. \tag{4.3}$$

By approximation, one can see that (4.3) holds for any process Z of the form

$$Z(t)=\sum_{i=1}^m \xi_i(t)\varphi_i$$

where $\xi=(\xi_1,\ldots,\xi_m)$ is $\{\mathcal{F}_t\}$-adapted, cadlag and has values in $A(\varphi_1,\ldots,\varphi_m)$. By (4.3), it follows that $Y(\varphi,\cdot)$ is an $\{\mathcal{F}_t\}$-semimartingale for each φ and hence that Y is an $H^\#$-semimartingale. It also follows from (4.3) that Y is standard.

Finally, observing that $\|X_n(s)-X_n^\epsilon(s)\|_H/\epsilon$ is bounded by 1, we have that

$$P\{\sup_{s\leq t}|X_{n-}\cdot Y_n - X_{n-}^\epsilon\cdot Y_n|\geq\epsilon(K(\delta,t)\}\leq\delta$$

and similarly for X and Y. Consequently, the Theorem follows from Lemma 4.1. □

Example 4.4 *Many particle random walk.*

For each n let X_k^n, $k=1,\ldots,n$, be independent, continous-time, reflecting random walks on $E_n=\{\frac{i}{n}: i=0,\ldots,n\}$ with generator

$$B_nf(x)=\begin{cases}\frac{n^2}{2}(f(x+\frac{1}{n})+f(x-\frac{1}{n})-2f(x)), & 0<x<1\\ n^2(f(\frac{1}{n})-f(0)), & x=0\\ n^2(f(1-\frac{1}{n})-f(1), & x=1\end{cases}$$

and $X_k^n(0)$ uniformly distributed on E_n. Let $H=C^1([0,1])$ with $\|\varphi\|_H=\sup_{0\leq x\leq 1}(|\varphi(x)|+|\varphi'(x)|)$, and define

$$Y_n(\varphi,t)=\frac{1}{n}\sum_{k=1}^n\left(\varphi(n^{-1}X_k^n(t))-\varphi(n^{-1}X_k^n(0))-\int_0^t B_n\varphi(X_k^n(s))ds\right).$$

Note that Y_n corresponds to a martingale random measure and that

$$\begin{aligned}&E[(Y_n(\varphi,t_2)-Y_n(\varphi,t_1))^2|\mathcal{F}_{t_1}^n]\\ &=\frac{1}{\sqrt{n}}\sum_{k=1}^n E[\int_{t_1}^{t_2} n^2\frac{\left(\varphi(X_k^n(s)+\frac{1}{n})-\varphi(X_k^n(s))\right)^2+\left(\varphi(X_k^n(s)-\frac{1}{n})-\varphi(X_k^n(s))\right)^2}{2}|\mathcal{F}_t^n].\end{aligned}$$

It follows that for $Z\in\mathcal{S}_0^n$,

$$E[(Z_-\cdot Y_n(t))^2]\leq t\sup_{0\leq s\leq t}\sup_{0\leq x\leq 1}|Z'(s,x)|^2\leq t\sup_{0\leq s\leq t}\|Z(s)\|_H^2,$$

and hence $\{Y_n\}$ is uniformly tight. The martingale central limit theorem gives $Y_n\Rightarrow Y$ where Y is Gaussian and satisfies

$$\langle Y(\varphi_1,\cdot),Y(\varphi_2,\cdot)\rangle_t = t\int_0^1\varphi_1'(x)\varphi_2'(x)dx.$$

It follows that Y does not correspond to a standard martingale random measure.

5 Convergence in infinite dimensional spaces.

Theorem 4.2 extends easily to integrals with range in $\mathbb{R}^d$. The interest in semimartingales in infinite dimensional spaces, however, is frequently in relation to stochastic partial differential equations. Consequently, extension to function-valued integrals is desirable. For semimartingale random measures, this extension would be to integrals of the form

$$Z(t,x) = \int_{U\times[0,t]} X(s-,x,u)Y(du\times ds)$$

where X is a process with values in a function space on $E \times U$. We will take (E,r) to be a complete, separable metric space. More generally, let X be an H-valued process indexed by $[0,\infty)\times E$. If X is $\{\mathcal{F}_t\}$-adapted and $X(\cdot,x)$ is cadlag for each x, then

$$Z(t,x) = X(\cdot-,x)\cdot Y(t)$$

is defined for each x; however, the properties of Z as a function of x (even the measurability) are not immediately clear. Consequently, we construct the desired integral more carefully.

5.1 Integrals with infinite-dimensional range.

Let (E,r_E) and (U,r_U) be complete, separable metric spaces, let L be a separable Banach space of $\mathbb{R}$-valued functions on E, and let H be a separable Banach space of $\mathbb{R}$-valued functions on U. (We restrict our attention to function spaces so that for $f\in L$ and $\varphi\in H$, $f\varphi$ has the simple interpretation as a pointwise product. The restriction to function spaces could be dropped with the introduction of an appropriate definition of product.) Let $G_L=\{f_i\}\subset L$ be a sequence such that the finite linear combinations of the f_i are dense in L, and let $G_H=\{\varphi_j\}$ be a sequence such that the finite linear combinations of the φ_j are dense in H.

Definition 5.1 *Let $\hat{H}$ be the completion of the linear space $\{\sum_{i=1}^{l}\sum_{j=1}^{m}a_{ij}f_i\varphi_j\colon f_i\in G_L, \varphi_i\in G_H\}$ with respect to some norm $\|\cdot\|_{\hat{H}}$.*

For example, if

$$\begin{aligned}&\|\sum_{i=1}^{l}\sum_{j=1}^{m}a_{ij}f_i\varphi_j\|_{\hat{H}}\\ &\qquad = \sup\{\sum_{i=1}^{m}a_{ij}\langle\lambda,f_i\rangle\langle\eta,\varphi_i\rangle : \lambda\in L^*,\eta\in H^*,\|\lambda\|_{L^*}\le 1,\|\eta\|_{H^*}\le 1\}\end{aligned}$$

then we can interpret $\hat{H}$ as a subspace of bounded linear operators mapping H^* into L. Metivier and Pellaumail (1980) develop the stochastic integral in this setting.

We say that a norm $\|\cdot\|_G$ on a function space G is monotone, if $g\in G$ implies $|g|\in G$ and $|g_1|\le|g_2|$ implies $\|g_1\|_G\le\|g_2\|_G$. If $\|\cdot\|_L$ and $\|\cdot\|_H$ are both monotone, we may take

$$\|\sum_{i=1}^{l}\sum_{j=1}^{m}a_{ij}f_i\varphi_j\|_{\hat{H}} = \|\|\sum_{i=1}^{l}\sum_{j=1}^{m}a_{ij}f_i\varphi_j\|_L\|_H. \tag{5.1}$$

Note that in the above examples, the mapping $(f,\varphi) \in L \times H \to f\varphi \in \hat{H}$ is continuous, although in general, we do not require this continuity.

Let $\zeta_k = \sum a_{kij} f_i \varphi_j$, $k = 1, 2, \ldots$ be a dense sequence in $\hat{H}$, where each sum is finite, and let $\{\hat{\psi}_k^\epsilon\}$ be as in Lemma 3.1 with $\{x_k\}$ replaced by $\{\zeta_k\}$. Then for each $v \in D_{\hat{H}}[0,\infty)$, we can define $v^\epsilon(t) = \sum_k \hat{\psi}_k^\epsilon(v(t))\zeta_k$, and we have $\|v(t) - v^\epsilon(t)\|_{\hat{H}} \le \epsilon$. Furthermore, if we define

$$c_{ij}^\epsilon(v) = \sum_k \hat{\psi}_k^\epsilon(v) a_{kij} , \tag{5.2}$$

then $v^\epsilon(t) = \sum c_{ij}^\epsilon(v(t)) f_i \varphi_j$, and only finitely many of the $\{c_{ij}(v(t))\}$ are non-zero on any bounded interval $0 \le t \le T$.

With the above approximation in mind, let

$$X(t) = \sum_{i,j} \xi_{ij}(t) f_i \varphi_j$$

where the ξ_{ij} are $\mathbb{R}$-valued, cadlag, adapted processes and only finitely many of the ξ_{ij} are non-zero. If Y is an $H^\#$-semimartingale, we can define

$$X_- \cdot Y(t) = \sum_i f_i \sum_j \int_0^t \xi_{ij}(s-) dY(\varphi_j, s).$$

Then $X_- \cdot Y$ is in $D_L[0,\infty)$.

Definition 5.2 *Let $\mathcal{S}_{\hat{H}}^0$ be the collection of simple $\hat{H}$-valued processes of the form*

$$X = \sum \xi_{kij} I_{[t_k, t_{k+1})} f_i \varphi_j$$

where ξ_{kij} is an $\mathbb{R}$-valued, $\mathcal{F}_{t_k}$-measurable random variable and all but finitely many of the ξ_{kij} are zero, and let $\mathcal{S}_{\hat{H}}$ be the collection of $\hat{H}$-valued processes of the form

$$X(t) = \sum_{ij} \xi_{ij}(t) f_i \varphi_j,$$

where the ξ_{ij} are cadlag and adapted and all but finitely many are zero.

For $X \in \mathcal{S}_{\hat{H}}^0$

$$X_- \cdot Y(t) = \sum \xi_{kij} f_i (Y(\varphi_j, t \wedge t_{k+1}) - Y(\varphi_j, t \wedge t_k)),$$

and for $X \in \mathcal{S}_{\hat{H}}$

$$X_- \cdot Y(t) = \sum_{ij} f_i \int_0^t \xi_{ij}(s-) dY(\varphi_j, s).$$

As in the $\mathbb{R}$-valued case, we make the following definition.

Definition 5.3 *An $H^\#$-semimartingale Y is a* standard $(L,\hat{H})^\#$*-semimartingale if*

$$\mathcal{H}_t^0 = \{\|X_- \cdot Y(t)\|_L \, : \, X \in \mathcal{S}_{\hat{H}}^0, \ \sup_{s \le t} \|X(s)\|_{\hat{H}} \le 1\} \tag{5.3}$$

is stochastically bounded for each t, or equivalently (as in Lemma 3.8)

$$\mathcal{H}_t = \{\sup_{s \le t} \|X_- \cdot Y(s)\|_L \, : \, X \in \mathcal{S}_{\hat{H}}, \ \sup_{s \le t} \|X(s)\|_{\hat{H}} \le 1\}$$

is stochastically bounded for each t.

As before, this assertion holds if there exists a constant $C(t)$ such that for all $X \in \mathcal{S}^0_{\hat{H}}$ satisfying $\sup_{s\leq t}\|X(s)\|_{\hat{H}} \leq 1$,

$$E[\|X_- \cdot Y(t)\|_L] \leq C(t).$$

If Y is standard, then, as in the $\mathbb{R}$-valued case, the definition of $X_- \cdot Y$ can be extended to all cadlag, $\hat{H}$-valued processes X by approximating X by

$$X^\epsilon(t) = \sum_k \hat{\psi}^\epsilon_k(X(t))\zeta_k = \sum_{ij} c^\epsilon_{ij}(X(t))f_i\varphi_j. \tag{5.4}$$

In fact, Lemma 3.15 and Proposition 3.16 extend to the present setting and $X \cdot Y$ can be defined for all predictable $\hat{H}$-valued processes X satisfying the compact range condition. Note that $X \cdot Y$ will be cadlag.

Remark 5.4 Other approaches. *Metivier and Pellaumail (1980) develop an integral for, in our notation, H^*-semimartingales with integrands whose values are bounded linear operators from H^* to L. (Take $\|\cdot\|_{\hat{H}}$ to be given by (5.1).) They assume moment conditions that imply that the integrator is a standard $(L,\hat{H})^\#$-semimartingale. These conditions are sufficient to extend the integral to all locally bounded predictable integrands. (See Remark 3.17.) DePrato and Zabczyk (1992) is a recent account of the theory of stochastic integration and stochastic ordinary and partial differential equations driven by infinite dimensional Brownian motions of the form described in Example 3.19. Ustunel (1982) develops stochastic integration in nuclear spaces. Mikulevicius and Rozovskii (1994) consider integrals in more general linear topological spaces considering integrands in continously embedded Hilbert subspaces determined by the covariance operator of the martingale.*

5.2 Convergence theorem.

Let $\{Y_n\}$ be a sequence of standard $(L,\hat{H})^\#$-semimartingales and define

$$\mathcal{H}^0_{n,t} = \{\|X_{n-} \cdot Y_n(t)\|_L \;:\; X \in \mathcal{S}^0_{n,\hat{H}},\ \sup_{s\leq t}\|X_n(s)\|_{\hat{H}} \leq 1\} \tag{5.5}$$

If $\hat{\mathcal{H}}^0_t = \cup\mathcal{H}^0_{n,t}$ is stochastically bounded for each t, we will again say that $\{Y_n\}$ is uniformly tight. (Since bounded sets in L are not, in general, compact, this terminology is not entirely appropriate; however, see Lemma 6.14.) As in Lemma 3.8, uniform tightness implies

$$\hat{\mathcal{H}}_t = \cup_n\{\sup_{s\leq t}\|X_{n-} \cdot Y_n(s)\|_L \;:\; X_n \in \mathcal{S}_{n,\hat{H}}, \sup_{s\leq t}\|X_n(s)\|_{\hat{H}} \leq 1\}$$

is stochastically bounded for each t.

If $L = \mathbb{R}^k$ and $\hat{H} = H^k$ with $\|(\varphi_1,\ldots,\varphi_k)\|_{\hat{H}} = \sum_{i=1}^k \|\varphi_i\|_H$, then any uniformly tight sequence of $H^\#$-semimartingales is a uniformly tight sequence of $(L,\hat{H})^\#$-semimartingales.

Theorem 5.5 *For each $n = 1,2,\ldots$, let Y_n be a standard $\{\mathcal{F}^n_t\}$-adapted, $(L,\hat{H})^\#$-semimartingale, and assume that $\{Y_n\}$ is uniformly tight.*

If $(X_n, Y_n) \Rightarrow (X, Y)$, then there is a filtration $\{\mathcal{F}_t\}$, such that Y is an $\{\mathcal{F}_t\}$-adapted, standard, $(L, \hat{H})^\#$-semimartingale and X is $\{\mathcal{F}_t\}$-adapted, and $(X_n, Y_n, X_{n-} \cdot Y_n) \Rightarrow (X, Y, X_- \cdot Y)$. If $(X_n, Y_n) \to (X, Y)$ in probability, then $(X_n, Y_n, X_{n-} \cdot Y_n) \to (X, Y, X_- \cdot Y)$ in probability.

Proof. Noting that Lemma 4.1 extends to this setting, the proof is exactly the same as the proof of Theorem 4.2. □

5.3 Verification of standardness.

If $L = C[0,1])$ and $H = \mathbb{R}$ and we identify $\hat{H}$ with L, then a scalar semimartingale Y defines a standard $(L, L)^\#$-semimartingale if and only if Y is a finite variation process. In particular, for any $t_0 < t_1 < \cdots < t_m = t$ we can define $Z = \sum_{i=0}^{m-1} \xi_i I_{[t_i, t_{i+1})}$ so that

$$\|Z_- \cdot Y(t)\|_L = \sum_{i=0}^{m-1} |Y(t_{i+1}) - Y(t_i)|,$$

simply by ensuring that for each of the 2^m choices of $\theta_i = \pm 1$, $i = 0, \ldots, m-1$, there is some value of $x \in [0,1]$ such that $\xi_i(x) = \theta_i$.

Fortunately, more interesting examples of $(L, \hat{H})^\#$-semimartingales exist in other spaces. Let $L = L_2(\nu)$, and assume that $Y = M + V$ is a standard semimartingale random measure with dominating measure K. Then, as is (2.6),

$$\begin{aligned} &E\left[\sup_{s\le t}\left(\int_{U\times[0,s]} X(s-,x,u)M(du\times ds)\right)^2\right] \\ &\qquad \le 4E\left[\int_{U\times U\times[0,t]} |X(s-,x,u)||X(s-,x,v)|K(du\times dv\times ds)\right], \end{aligned} \tag{5.6}$$

and, since the integral of the sup is greater than the sup of the integral, it follows that

$$\begin{aligned} &E\left[\sup_{s\le t}\left\|\int_{U\times[0,s]} X(s-,\cdot,u)M(du\times ds)\right\|_L^2\right] \\ &\quad \le 4E\left[\int_{U\times U\times[0,t]}\int_E |X(s-,x,u)||X(s-,x,v)|\nu(dx)K(du\times dv\times ds)\right] \\ &\quad \le 4E\left[\int_{U\times U\times[0,t]} \|X(s-,\cdot,u\|_L\|X(s-,\cdot,v\|_L K(du\times dv\times ds)\right]. \end{aligned}$$

If the norm on H is the sup norm and $\|X(t)\|_{\hat{H}} \equiv \|\|X(t,\cdot,\cdot)\|_L\|_H \le 1$,

$$\begin{aligned} &E\left[\sup_{s\le t}\left\|\int_{U\times[0,s]} X(s-,\cdot,u)M(du\times ds)\right\|_L^2\right] \\ &\quad \le 4E[K(U\times U\times[0,t])] \end{aligned}$$

The analogous inequality for V is simply

$$
\begin{aligned}
E&\left[\sup_{s\le t}\left\|\int_{U\times[0,s]}X(s-,\cdot,u)V(du\times ds)\right\|_L\right]\\
&\le E\left[\int_{U\times[0,t]}\|X(s-,\cdot,u)\|_L|\tilde V|(du\times ds)\right]\\
&\le E[(|\tilde V|(U\times[0,t]))]\,.
\end{aligned}
$$

With $\hat K$ defined as in (2.10) and $\hat V$ as in (2.12), suppose that $\hat K(du\times dt)=h(u,t)\mu(du)dt$ and $\hat V(du\times dt)=g(u,t)\mu(du)dt$. Let $H=L_p(\mu)$ for $2\le p<\infty$ and again assume that $\|X(t)\|_{\hat H}\equiv\|\|X(t,\cdot,\cdot)\|_L\|_H\le 1$. Then with $\frac{2}{p}+\frac{1}{r}=1$ and $\frac{1}{p}+\frac{1}{q}=1$

$$
\begin{aligned}
E&\left[\sup_{s\le t}\left\|\int_{U\times[0,s]}X(s-,\cdot,u)M(du\times ds)\right\|_L^2\right]\\
&\le 4E\left[\int_0^t\int_U\|X(s-,\cdot,u)\|_L^2h(u,s)\mu(du)ds\right]\\
&\le 4E\left[\int_0^t\left(\int_U\|X(s-,\cdot,u)\|_L^p\mu(du)\right)^{\frac{2}{p}}\left(\int_U|h(u,s)|^r\mu(du)\right)^{\frac{1}{r}}ds\right]\\
&\le 4E\left[\int_0^t\left(\int_U|h(u,s)|^r\mu(du)\right)^{\frac{1}{r}}ds\right]
\end{aligned}
$$

(with the obvious modification if $r=\infty$) and for V

$$
\begin{aligned}
E&\left[\sup_{s\le t}\left\|\int_{U\times[0,s]}X(s-,\cdot,u)V(du\times ds)\right\|_L\right]\\
&\le E\left[\int_{U\times[0,t]}\|X(s-,\cdot,u)\|_L|\tilde V|(du\times ds)\right]\\
&\le E\left[\int_{U\times[0,t]}\|X(s-,\cdot,u)\|_L\hat V(du\times ds)\right]\\
&\le E\left[\int_0^t\|\|X(s,x,\cdot)\|_L\|_H\left(\int_U|g(u,s)|^q\mu(du)\right)^{\frac{1}{q}}ds\right]\\
&\le E\left[\int_0^t\left(\int_U|g(u,s)|^q\mu(du)\right)^{\frac{1}{q}}ds\right].
\end{aligned}
$$

From the above inequalities, we see that, at least in the Hilbert space setting, we can give conditions under which a semimartingale random measure gives a standard $(L,\hat H)^\#$-semimartingale and conditions under which a sequence of such $(L,\hat H)^\#$-semimartingales satisfies a uniform tightness condition. In particular, we have the following analog of Lemma 3.3. An analog of Lemma 3.4 will also hold.

Lemma 5.6 *Let $L=L_2(\nu)$ and $\|\cdot\|_{\hat H}=\|\|\cdot\|_L\|_H$.*

a) Let $\|\cdot\|_H$ be the sup *norm, and suppose $E[K(U\times U\times[0,t])]<\infty$ and $E[|\tilde V|(U\times[0,t])]<\infty$ for all $t>0$. Then if $\sup_s\|Z(s)\|_{\hat H}\le 1$ and τ is a stopping time bounded by a constant c,*

$$E[\sup_{s\le t}\|Z_-\cdot Y(\tau+s)-Z_-\cdot Y(\tau)\|_L]\le 2\sqrt{E[K(U\times U\times(\tau,\tau+t])]}+E[|\tilde V|(U\times[\tau,\tau+t])]$$

and

$$\lim_{t\to 0}E[K(U\times U\times(\tau,\tau+t])]+E[(|\tilde V|(U\times[\tau,\tau+t]))^2]=0. \tag{5.7}$$

b) Let $H=L_p(\mu)$, for some $p\ge 2$, and for h and g as in (3.10) and (3.11), suppose $E[\int_0^t\|h(\cdot,s)\|_{L_r(\mu)}ds]<\infty$ and $E[(\int_0^t\|g(\cdot,s)\|_{L_q(\mu)}ds)^2]<\infty$ for all $t>0$. Then if $\sup_s\|Z(s)\|_{\hat H}\le 1$ *and τ is a stopping time bounded by a constant c,*

$$E[\sup_{s\le t}\|Z_-\cdot Y(\tau+s)-Z_-\cdot Y(\tau)\|_L]\le 2\sqrt{E[\int_\tau^{\tau+t}\|h(\cdot,s)\|_{L_r(\mu)}ds]}+E[(\int_\tau^{\tau+t}\|g(\cdot,s)\|_{L_q(\mu)}ds)^2]\alpha^2$$

and

$$\lim_{t\to 0}E[\int_\tau^{\tau+t}\|h(\cdot,s)\|_{L_r(\mu)}ds]+E[(\int_\tau^{\tau+t}\|g(\cdot,s)\|_{L_q(\mu)}ds)^2]=0.$$

Now let $L=C([0,1]^d)$ with the sup norm. It is clear from the discussion at the beginning of this subsection, that we cannot let $\|f\|_{\hat H}=\sup_x\|f(x,\cdot)\|_H$ if we want interesting standard $(L,\hat H)^\#$-semimartingales. It is sufficient, however, to give $\hat H$ some kind of Hölder norm. In particular, let M be an orthogonal martingale random measure and suppose the $\hat\pi_j$ defined as in (2.20) satisfy

$$\hat\pi_j(du\times ds)=h_j(u,s)\nu_j(du)ds.$$

Let $\frac1p+\frac1q=1$, and suppose for each $t>0$,

$$C_{k,j}(t)\equiv E\left[\left(\int_0^t\left(\int_U|h_j(u,s)|^q\nu_j(du)\right)^{\frac1q}\right)^{\frac kj}\right]<\infty.$$

For k and α satisfying $0<\alpha\le 1$ and $k\alpha>d$, define

$$\|f\|_H=\sum_{j=2}^k\left(\int_U|f(u)|^{jp}\nu_j(du)\right)^{\frac1{jp}},$$

and

$$\|g\|_{\hat H}=\sup_x\|g(x,\cdot)\|_H+\sup_{x,y}\frac{\|g(x,\cdot)-g(y,\cdot)\|_H}{|x-y|^\alpha}.$$

Recall that if M is continuous, then $\nu_j=0$ for $j>2$ and H is just $L_{2p}(\nu_2)$. Now suppose $X\in\mathcal S^0_{\hat H}$ and $\|X\|_{\hat H}\le 1$. Fix $x,y\in[0,1]^d$, and set

$$\begin{aligned}Z(t,x,y)&=X(\cdot,x,\cdot)_-\cdot M(t)-X(\cdot,y,\cdot)_-\cdot M(t)\\&=\int_{U\times[0,t]}(X(s-,x,u)-Y(s-,y,u))M(du\times ds).\end{aligned}$$

With reference to (2.21),

$$\begin{aligned}
H_{k,j}(x,y) &= E\left[\left(\int_{U\times[0,t]}|X(s-,x,u)-X(s-,y,u)|^j\hat{\pi}_j(du\times ds)\right)^{\frac{k}{j}}\right]\\
&\le E\left[\left(\int_0^t\left(\int_U|X(s,x,u)-X(s,y,u)|^{jp}\nu_j(du)\right)^{\frac{1}{p}}\left(\int_U|h_j(u,s)|^q\nu_j(du)\right)^{\frac{1}{q}}\right)^{\frac{k}{j}}\right]\\
&\le |x-y|^{\alpha k}E\left[\left(\int_0^t\left(\int_U|h_j(u,s)|^q\nu_j(du)\right)^{\frac{1}{q}}\right)^{\frac{k}{j}}\right]\\
&= |x-y|^{\alpha k}C_{k,j}(t).
\end{aligned}$$

Then with reference to (2.25),

$$E[|Z(t,x,y)|^k]\le K|x-y|^{\alpha k} \tag{5.8}$$

where K is the largest number satisfying

$$K\le C_1kC_{k,2}^{\frac{1}{k}}K^{\frac{k-1}{k}}+\sum_{j=2}^{k}\binom{k}{j}C_{k,j}^{\frac{j}{k}}K^{\frac{k-j}{k}}.$$

The remainder of the proof of standardness is essentially the same as for the proof of the Kolmogorov continuity criterion. (Recall that the exponent αk on the right side of (5.8) is greater than d.) Note that any $x\in[0,1]^d$ can be represented as

$$x=\sum_{i=1}^{\infty}\frac{1}{2^i}(\theta_1^i(x),\ldots,\theta_d^i(x))$$

where $\theta_k^i(x)$ is 0 or 1. Let $x^0=0$ and

$$x^m=\sum_{i=1}^{m}\frac{1}{2^i}\theta^i(x)$$

It follows that

$$X(\cdot,x,\cdot)_-\cdot M(t)=X(\cdot,0,\cdot)\cdot M(t)+\sum_{m=0}^{\infty}Z(t,x^{m+1},x^m).$$

For each $\theta\in\{0,1\}^d$ and $m=1,2,\ldots$, let

$$\eta_m(\theta)=\sum_{\{y\in[0,1)^d:2^my\in\mathbf{Z}^d\}}|Z(t,y+\frac{1}{2^m}\theta,y)|^k.$$

Then

$$\sup_{x\in[0,1]^d}|X(\cdot,x,\cdot)_-\cdot M(t)|\le|X(\cdot,0,\cdot)\cdot M(t)|+\sum_{m=1}^{\infty}\sum_{\theta\in\{0,1\}^d}\eta_m(\theta)^{\frac{1}{k}},$$

and since by (5.8)

$$\begin{aligned} E\left[\sum_{m=1}^{\infty}\sum_{\theta\in\{0,1\}^d}\eta_m(\theta)^{\frac{1}{k}}\right] &\leq \sum_{m=1}^{\infty}\sum_{\theta\in\{0,1\}^d}E[\eta_m(\theta)]^{\frac{1}{k}} \\ &\leq \sum_{m=1}^{\infty}2^d\left[K\left(\frac{\sqrt{d}}{2^m}\right)^{\alpha k}2^{md}\right]^{\frac{1}{k}} \\ &\leq 2^dK^{\frac{1}{k}}d^{\frac{\alpha}{2}}\sum_{m=1}^{\infty}\left(\frac{1}{2^m}\right)^{\frac{\alpha k-d}{k}} < \infty, \end{aligned}$$

the stochastic boundedness of $\mathcal{H}_t^0$ follows.

5.4 Equicontinuity of stochastic integrals.

The fact that the estimates in Lemma 5.6 along with those in Lemmas 3.3 and 3.4 tend to zero as $t \to 0$ will be needed for the uniqueness theorem in Section 7. This fact holds very generally for standard $(L, \hat{H})^{\#}$-semimartingales, although we have not been able to prove that it always holds. The following lemma follows easily from Araujo and Giné (1980), Theorem 3.2.8.

Lemma 5.7 *Let $\{\theta_i\}$ be iid with $P\{\theta_i = 1\} = P\{\theta_i = -1\} = \frac{1}{2}$, let $\{x_i\} \subset L$, and define $S_k = \sum_{i=1}^k \theta_i x_i$. Then for $p \geq 1$,*

$$2P\{\|S_n\|_L > a\} \geq P\{\max_{k\leq n}\|S_k\|_L > a\} \geq 2^{1-p}\left(1 - \frac{a^p(1+3^p)}{E[\|S_n\|_L^p]}\right).$$

Let $f_{p,n}(x_1, \ldots, x_n) = E[\|S_n\|_L^p]$. If $X_1, X_2, \ldots$ are L-valued random variables, independent of the $\{\theta_i\}$, and $\tilde{S}_k = \sum_{i=1}^k \theta_i X_i$, then

$$P\{\max_{k\leq n}\|\tilde{S}_k\|_L > a\} \geq 2^{-p}P\{f_{p,n}(X_1, \ldots, X_n) \geq 2(1+3^p)a^p\}.$$

Remark 5.8 *By definition, if L is cotype p, then $f_{p,n}(x_1, \ldots, x_n) \geq c\sum_{i=1}^n \|x_i\|_L^p$. If L is uniformly convex, for each $\epsilon > 0$ there exists $\alpha(\epsilon) > 0$ such that*

$$f_{2,n}(x_1, \ldots, x_n) \geq \alpha(\epsilon)\sum_{i=1}^n I_{[\epsilon,\infty)}(\|x_i\|_L). \tag{5.9}$$

Note that the cotype inequality implies (5.9) (with 2 replaced by p). Since $L_1(\nu)$ is cotype 2 (see Araujo and Giné (1980), page 188) and $L_r(\nu)$, $1 < r < \infty$ is uniformly convex, (5.9) holds for all $L_r(\nu)$, $1 \leq r < \infty$. The inequality fails for $L_\infty(\nu)$.

Proof. If $\max_{1\leq i\leq n}\|x_i\|_L > 2a$, then $P\{\max_{k\leq n}\|S_k\|_L > a\} = 1$, so the bound c in Araujo and Giné (1980), Theorem 3.2.8, can, in the present setting, be replaced by $2a$, and the first inequality follows. The second inequality is an immediate consequence of the first. □

Lemma 5.9 *Suppose that L satisfies (5.9) (possibly with $f_{2,n}$ replaced by $f_{p,n}$). Let Y be a standard $(L, \hat{H})^{\#}$-semimartingale, and define*

$$K(t,\delta) = \inf\{K : P\{\sup_{s\le t} \|Z_- \cdot Y(s)\|_L \ge K\} \le \delta, Z \in \mathcal{S}_{\hat{H}}, \|Z\|_{\hat{H}} \le 1\}.$$

Then for each $\delta > 0$, $\lim_{t\to 0} K(t,\delta) = 0$. More generally, let τ be a stopping time bounded by a constant and define

$$K_\tau(t,\delta) = \inf\{K : P\{\sup_{s\le t} \|Z_- \cdot Y(\tau+s) - Z_- \cdot Y(\tau)\|_L \ge K\} \le \delta, Z \in \mathcal{S}_{\hat{H}}, \|Z\|_{\hat{H}} \le 1\}.$$

Then $\lim_{t\to 0} K_\tau(t,\delta) = 0$.

Proof. Consider the case $\tau = 0$. The general case is similar. Suppose that $\lim_{t\to 0} K(t,\delta) > 0$. Then there exists $t_n \to 0$, $K > 0$, and $Z_n \in \mathcal{S}_{\hat{H}}$ with $\|Z_n\|_{\hat{H}} \le 1$, such that

$$P\{\sup_{s\le t_n} \|Z_{n-} \cdot Y(s)\|_L \ge K\} \ge \frac{2}{3}\delta.$$

Since $Z_{n-} \cdot Y$ is right continuous and vanishes at zero, we can select $0 < r_n < t_n$ such that

$$P\{\sup_{r_n\le s\le t_n} \|Z_{n-} \cdot Y(s) - Z_{n-} \cdot Y(r_n)\|_L \ge \frac{K}{2}\} \ge \frac{\delta}{2}.$$

Let

$$\sigma_n = \inf\{s > r_n : \|Z_{n-} \cdot Y(s) - Z_{n-} \cdot Y(r_n)\|_L \ge \frac{K}{2}\},$$

and note that

$$P\{\|Z_{n-} \cdot Y(\sigma_n \wedge t_n) - Z_{n-} \cdot Y(r_n)\|_L \ge \frac{K}{2}\} \ge \frac{\delta}{2}.$$

Without loss of generality, we can assume that there is a sequence $\{\theta_i\}$, as in Lemma 5.7, that is independent of the Z_n and is $\mathcal{F}_0$-measurable. (If not, enlarge the sample space and the filtration to include such a sequence and note that the stochastic boundedness of $\mathcal{H}_t^0$ is unaffected.)

Select a subsequence satisfying $t_{n_1} = t_1$ and $t_{n_{k+1}} < r_{n_k}$, and define

$$Z^{(m)} = \sum_{k=1}^m \theta_k I_{[r_{n_k}, t_{n_k}\wedge\sigma_{n_k})} Z_{n_k}.$$

Then $Z^{(m)}$ is cadlag and adapted, $\|Z\|_{\hat{H}} \le 1$, and setting $X_k = Z_{n_k-} \cdot Y(\sigma_{n_k} \wedge t_{n_k}) - Z_{n_k-} \cdot Y(r_{n_k})$

$$Z_-^{(m)} \cdot Y(t_1) = \sum_{k=1}^m \theta_k X_k.$$

By Lemma 5.7 and (5.9)

$$\lim_{m\to\infty} P\{\|Z_-^{(m)} \cdot Y(t_1)\|_L > a\} \ge 2^{-p} P\{\|X_k\|_L > \frac{K}{4}, i.o.\} \ge 2^{-p-1}\delta.$$

Since a is arbitrary, this estimate violates the stochastic boundedness of $\mathcal{H}_t^0$, and the lemma follows. □

The following variation on the above lemma may be useful in proving uniqueness in spaces in which (5.9) fails.

Lemma 5.10 *Let Y be a standard $(L,\hat{H})^\#$-semimartingale, let τ be a stopping time bounded by a constant, and let $\Gamma \subset \hat{H}$ be compact. Define*

$$K_\tau(t,\delta,\Gamma) = \inf\{K : P\{\sup_{s\le t}\|Z_-\cdot Y(\tau+s) - Z_-\cdot Y(\tau)\|_L \ge K\} \le \delta, Z \in \mathcal{S}_{\hat{H}}, Z(s) \in \Gamma, s \le t\}.$$

Then for each $\delta > 0$, $\lim_{t\to 0} K_\tau(t,\delta,\Gamma) = 0$.

Proof. Take $\tau = 0$. The general case is similar. Let Z_n, t_{n_k}, and r_{n_k} be as in the proof of the previous lemma, and define

$$Z = \sum_{k=1}^{\infty} I_{(r_{n_k}, t_{n_k}\wedge\sigma_{n_k}]} Z_{n_k}.$$

Then Z is predictable and takes values in the compact set Γ. Consequently, $Z\cdot Y$ is defined and cadlag. In particular, $\lim_{t\to 0} Z\cdot Y(t) = 0$. But for $r_{n_k} \le t < t_{n_k}$, $Z\cdot Y(t) - Z\cdot Y(r_{n_k}) = Z_{n_k-}\cdot Y(t) - Z_{n_k-}\cdot Y(r_{n_k})$ so

$$\limsup_{t\to 0}\sup_{s\le t}\|Z\cdot Y(t) - Z\cdot Y(s)\|_L \ge \frac{K}{2}$$

with probability at least $\frac{\delta}{2}$, contradicting the right continuity at 0 and giving the result. □

6 Consequences of the uniform tightness condition.

Let $\{Y_n\}$ be a sequence of $H^\#$-semimartingales, and let $\hat{\mathcal{H}}_t^0$ be as in Theorem 4.2, that is,

$$\hat{\mathcal{H}}_t^0 = \cup_n\{|X_{n-}\cdot Y_n(t)| : X_n \in \mathcal{S}_0^n, \sup_{s\le t}\|X_n(s)\|_H \le 1\},$$

where $\mathcal{S}_0^n$ is the collection of simple, finite dimensional H-valued, $\{\mathcal{F}_t^n\}$-adapted processes. The sequence $\{Y_n\}$ is *uniformly tight (UT)* if $\hat{\mathcal{H}}_t^0$ is stochastically bounded for each $t > 0$. For real-valued semimartingales, this condition appears first in Stricker (1985) where it is shown to imply a type of relative compactness for the sequence of semimartingales previously studied by Meyer and Zheng (1984) under somewhat different conditions. Jakubowski (1995) develops the topological properties of the corresponding convergence, and Kurtz (1991) characterizes the convergence in terms of convergence in the Skorohod topology of a time-changed sequence. Uniform tightness is the basic condition in the stochastic integral convergence results of Jakubowski, Mémin, and Pages (1989).

As noted previously, uniform tightness is equivalent to the stochastic boundedness of

$$\hat{\mathcal{H}}_t = \cup_n\{\sup_{s\le t}|X_{n-}\cdot Y_n(s)| : X_n \in \mathcal{S}^n, \sup_{s\le t}\|X_n(s)\|_H \le 1\}, \tag{6.1}$$

for each $t > 0$. In particular, if $\{Y_n\}$ is uniformly tight, for each $t > 0$ and $\delta > 0$ there exists a $K(t, \delta)$ such that

$$P\{\sup_{s\leq t} |X_{n-} \cdot Y_n(s)| \geq K(t,\delta)\} \leq \delta \tag{6.2}$$

for each n and all $X_n \in \mathcal{S}^n$ satisfying $\|X_n(s)\|_H \leq 1$. The following lemma is an immediate consequence of (6.2).

Lemma 6.1 *Let $\{Y_n\}$ be as above and let X_n be an H-valued, $\{\mathcal{F}_t^n\}$-adapted, cadlag process. Let $\tau_n^M = \inf\{t : \|X_n\|_H \geq M\}$. Then for all $t, \delta > 0$,*

$$P\{\sup_{s\leq t\wedge\tau_n^M} |X_{n-} \cdot Y_n(s)| \geq MK(t,\delta)\} \leq \delta.$$

The next lemma gives conditions under which stochastic integrals of uniformly tight sequences define uniformly tight sequences.

Proposition 6.2 *Let H_0 be a Banach space of functions on U with the property that $g \in H_0$ and $h \in H$ implies $gh \in H$ and $\|gh\|_H \leq \|g\|_{H_0}\|h\|_H$. Let $\{Y_n\}$ be a uniformly tight sequence of $H^{\#}$-semimartingales, and for each n, let X_n be an H-valued, cadlag, $\mathcal{F}_t^n$-adapted process such that for each t, the sequence $\{\sup_{s\leq t} \|X_n(s)\|_H\}$ is stochastically bounded. Define*

$$Z_n(g,\cdot) = gX_{n-} \cdot Y_n$$

for $g \in H_0$. Then $\{Z_n\}$ is a uniformly tight sequence of $H_0^{\#}$-semimartingales. In particular (taking $H_0 = \mathbb{R}$), $\{X_{n-} \cdot Y_n\}$ is a uniformly tight sequence of $\mathbb{R}$-valued semimartingales.

Proof. Let $\tilde{X}_n$ be a simple, $\{\mathcal{F}_t^n\}$-adapted, H_0-valued process satisfying $\|\tilde{X}_n(s)\|_{H_0} \leq 1$. Then $\tilde{X}_{n-} \cdot Z_n = \tilde{X}_n X_{n-} \cdot Y_n$, and

$$P\{\sup_{s\leq t} |\tilde{X}_{n-} \cdot Z_n(s)| \geq MK(t,\delta)\} \leq \delta + P\{\sup_{s\leq t} \|X_n(s)\|_H \geq M\}.$$

In particular, select M such that $P\{\sup_{s\leq t} \|X_n(s)\|_H \geq M\} \leq \delta$, and define $\tilde{K}(t, 2\delta) = MK(t,\delta)$. □

We have the following analogue of Lemma 3.10.

Lemma 6.3 *Let $\{Y_n^1\}$ be a uniformly tight sequence of $H_1^{\#}$-semimartingales and $\{Y_n^2\}$ a uniformly tight sequence of $H_2^{\#}$-semimartingales, where the conditions on Y_n^1 and Y_n^2 are with respect to the same filtration $\{\mathcal{F}_t^n\}$. Define $H = \{(\varphi_1, \varphi_2) : \varphi_1 \in H_1, \varphi_2 \in H_2\}$, with $\|\varphi\|_H = \|\varphi_1\|_{H_1} + \|\varphi_2\|_{H_2}$ and $Y_n(\varphi, t) = Y_n^1(\varphi_1, t) + Y_n^2(\varphi_2, t)$ for $\varphi = (\varphi_1, \varphi_2)$. Then $\{Y_n\}$ is a uniformly tight sequence of $H^{\#}$-semimartingales.*

We will need a number of technical lemmas regarding convergence in distribution in the Skorohod topology. For any nonnegative, nondecreasing function a defined on $[0,\infty)$, we define $a^{-1}(t) = \inf\{u : a(u) > t\}$. Let $T_c[0,\infty)$ be the collection of continuous, nondecreasing functions γ satisfying $\gamma(0) = 0$ and $\lim_{t\to\infty} \gamma(t) = \infty$. (Note that γ^{-1} is right continuous and strictly increasing and that $\gamma(t) = \inf\{u : \gamma^{-1}(u) > t\}$.) If $\{X_n\}$ is a sequence of processes in $D_E[0,\infty)$ and $\{\gamma_n\}$ is a sequence of processes in $T_c[0,\infty)$, then we say that $\{\gamma_n\}$ *regularizes* $\{X_n\}$ if $\{\gamma_n^{-1}(t)\}$ is stochastically bounded for each $t > 0$ and the sequence $\{(X_n \circ \gamma_n, \gamma_n)\}$ is relatively compact in $D_{E\times\mathbb{R}}[0,\infty)$. Note that if $Y_n = X_n \circ \gamma_n$, then $X_n(t) = Y_n(\gamma_n^{-1}(t))$. The following lemma is a restatement of Theorem 1.1b,c of Kurtz (1991) using the above terminology.

Lemma 6.4 *Suppose that $\{\gamma_n\}$ regularizes $\{X_n\}$ and that $(X_n \circ \gamma_n, \gamma_n) \Rightarrow (Y, \gamma)$. Then (by the Skorohod representation theorem) there exists a proability space on which are defined processes $\{(\hat{Y}_n, \hat{\gamma}_n)\}$ converging almost surely to a process $(\hat{Y}, \hat{\gamma})$ in the Skorohod topology on $D_{E\times\mathbb{R}}[0,\infty)$ such that $(\hat{Y}_n, \hat{\gamma}_n)$ has the same distribution as $(X_n \circ \gamma_n, \gamma_n)$ (and hence $\hat{X}_n = \hat{Y}_n \circ \hat{\gamma}_n^{-1}$ has the same distribution as X_n) and with probability 1, $\hat{X}_n(t) \to \hat{X}(t) \equiv \hat{Y} \circ \hat{\gamma}^{-1}(t)$ for all but countably many $t \geq 0$.*

If γ is strictly increasing, then $X_n \Rightarrow Y \circ \gamma^{-1}$.

Remark 6.5 *Under the conditions of Lemma 6.4, there exists a countable set $D \subset [0,\infty)$ such that for $(t_1,\ldots,t_m) \subset [0,\infty) - D$, $(X_n(t_1),\ldots,X_n(t_m)) \Rightarrow (X(t_1),\ldots,X(t_m))$, for $X = Y \circ \gamma^{-1}$.*

It follows from results of Stricker (1985) and Kurtz (1991) that any uniformly tight sequence of real-valued semimartingales can be regularized. We will prove a corresponding result for $H^{\#}$-semimartingales. For a given $\mathbb{R}$-valued, cadlag process X and $u < v$, $N(u,v,t)$ will denote the number of upcrossings of the interval (u,v) by X before time t, and for $u > v$, $N(u,v,t)$ will denote the number of downcrossings of the interval (v,u) before time t. Note, for example, that $|N(u,v,t) - N(v,u,t)| \leq 1$.

Lemma 6.6 *Let $\{X_n\}$ be a sequence of cadlag processes with X_n adapted to $\{\mathcal{F}_t^n\}$, and let $N_n(u,v,t)$ count down/upcrossings for X_n. Suppose that for each $(u,v,t) \in \mathbb{R}^2 \times [0,\infty)$, $u \neq v$, $\{N_n(u,v,t)\}$ is stochastically bounded and that for each $t \geq 0$, $\{\sup_{s\leq t}|X_n(s)|\}$ is stochastically bounded. Then for each n there exists a strictly increasing, $\{\mathcal{F}_t^n\}$-adapted process C_n such that for each $t > 0$, $\{C_n(t)\}$ is stochastically bounded and $\{\gamma_n\}$ defined by $\gamma_n = C_n^{-1}$ regularizes $\{X_n\}$, that is $\{(X_n \circ \gamma_n, \gamma_n)\}$ is relatively compact in the Skorohod topology.*

Proof. Let $\{(u_l,v_l), l \geq 1\}$ be some ordering of $\{(u,v) : u,v \in \mathbb{Q}, u \neq v\}$. By the stochastic boundedness assumptions, for $k,l = 1,2,\ldots$ there exist $0 < c_{kl} \leq 1$ such that

$$\sup_n P\{c_{kl}N_n(u_l,v_l,k) > 2^{-(k+l)}\} \leq 2^{-(k+l)}.$$

Define

$$C_n(t) = 1 + t + \sum_{l=1}^{\infty} \int_0^t c_{[s]+1,l}\, dN_n(u_l,v_l,s).$$

Note that

$$H_{n,L}(t)=\sum_{l=L+1}^{\infty}\int_0^t c_{[s]+1,l}dN_n(u_l,v_l,s)\le\sum_{l=L+1}^{\infty}\sum_{k=1}^{[t]+1}c_{kl}N_n(u_l,v_l,k)$$

and

$$P\{H_{n,L}(t)>2^{-L}\}\le 2^{-L},$$

so

$$P\{C_n(t)>t+a+2\}\le 2^{-L}+P\{\sum_{l=1}^{L}N_n(u_l,v_l,t)>a\},$$

and the stochastic boundedness of $\{C_n(t)\}$ follows from the assumed stochastic boundedness of $\{N_n(u_l,v_l,t)\}$.

Let $\gamma_n=C_n^{-1}$, and define $Z_n(t)=X_n(\gamma_n(t))$. Let $\epsilon>0$ be rational. Note that each time X_n crosses an interval (u_l,v_l), C_n jumps and that each jump of C_n corresponds to an interval on which γ_n, and hence Z_n, is constant.

Define $\tau_0^{n,\epsilon}=0$,

$$\tau_{k+1}^{n,\epsilon}=\inf\{t>\tau_k^{n,\epsilon}:Z_n(t)\notin[[Z_n(\tau_k^{n,\epsilon})/\epsilon]\epsilon-\epsilon,[Z_n(\tau_k^{n,\epsilon})/\epsilon]+2\epsilon]\}$$

and

$$Z_n^\epsilon(t)=Z_n(\tau_k^{n,\epsilon})\,,\quad \tau_k^{n,\epsilon}\le t<\tau_{k+1}^{n,\epsilon}$$

so that $|Z_n(t)-Z_n^\epsilon(t)|\le 2\epsilon$. To show that $\{Z_n\}$ is relatively compact, it is enough to show that $\{Z_n^\epsilon\}$ is relatively compact. Since the Z_n^ϵ are piecewise constant, it is enough to show that $\{\sup_{s\le t}|Z_n^\epsilon(s)|\}$ is stochastically bounded (which follows from the fact that $\sup_{s\le t}|Z_n^\epsilon(s)|\le\sup_{s\le t}|X_n(s)|$) and that $\{\min\{\tau_{k+1}^{n,\epsilon}-\tau_k^{n,\epsilon}:\tau_k^{n,\epsilon}<t\},n=1,2,\ldots\}$ is stochastically bounded away from 0, that is, for each $\delta>0$ there exists an $\eta>0$ such that $P\{\min\{\tau_{k+1}^{n,\epsilon}-\tau_k^{n,\epsilon}:\tau_k^{n,\epsilon}<t\}\le\eta\}\le\delta$. Let $c(u,v,k)=c_{kl}$ if $(u,v)=(u_l,v_l)$. Define $\eta(a,t)=\min\{c(i\epsilon,j\epsilon,[t]+1):|i\epsilon|,|j\epsilon|\le a,j=i\pm 1\}$. Then

$$P\{\min\{\tau_{k+1}^{n,\epsilon}-\tau_k^{n,\epsilon}:\tau_k^{n,\epsilon}<t\}<\eta(a,t)\}\le P\{\sup_{s\le t}|Z_n(s)|>a\}$$

since $\tau_1^{n,\epsilon}\ge 1$ and at each time $\tau_k^{n,\epsilon}$ with $k>0$, Z_n finishes a downcrossing or an upcrossing of an interval of the form $[i\epsilon,(i+1)\epsilon]$ and is constant for an interval of length at least $c(i\epsilon,(i+1)\epsilon,[\tau_k^{n,\epsilon}]+1)$ (in the case of an upcrossing) after time $\tau_k^{n,\epsilon}$ implying $\tau_{k+1}^{n,\epsilon}-\tau_k^{n,\epsilon}\ge c(i\epsilon,(i+1)\epsilon,[\tau_k^{n,\epsilon}]+1)$. □

Following the proof of Theorem 2 of Stricker (1985), we have the following lemma.

Lemma 6.7 *Let $\{Y_n\}$ be a uniformly tight sequence of $\mathbb{R}$-valued semimartingales, and let $\{N_n\}$ be as in Lemma 6.6. Then for $u\ne v\in\mathbb{R}$ and $t>0$, $\{N_n(u,v,t)\}$ is stochastically bounded.*

Proof. Suppose $u<v$. Define $\tau_1^n=\inf\{s\ge 0:X_n(s)\le u\}$, $\sigma_k^n=\inf\{s>\tau_k^n:X_n(s)\ge v\}$, and for $k>1$, $\tau_k^n=\inf\{s>\sigma_k^n:X_n(s)\le u\}$. Note that $N_n(u,v,t)=\max\{k:\sigma_k^n\le t\}$. Define $X_n=\sum_k I_{[\tau_k^n,\sigma_k^n)}$. Then $X_{n-}\cdot Y_n(t)\ge(v-u)N_n(u,v,t)-2\sup_{s\le t}|Y_n(s)|$

and hence, $(v-u)N_n(u,v,t) \le |X_{n-} \cdot Y_n(t)| + 2\sup_{s\le t}|Y_n(s)|$. Since $\{|X_{n-} \cdot Y_n(t)|\}$ is stochastically bounded by the definition of the uniform tightness of $\{Y_n\}$ and $\{\sup_{s\le t}|Y_n(s)|\}$ is stochastically bounded by (6.1), the stochastic boundedness of $\{N_n(u,v,t)\}$ follows. The proof for $u > v$ is essentially the same. □

Lemmas 6.6 and 6.7 immediately give the following:

Lemma 6.8 *For each n, let Y_n be an $\mathbb{R}$-valued $\{\mathcal{F}_t^n\}$-semimartingale, and let the sequence $\{Y_n\}$ be uniformly tight. Then there exist $\{\mathcal{F}_t^n\}$-adapted C_n satisfying $C_n(t) - C_n(s) \ge t - s$, $t > s \ge 0$, and $\{C_n(t)\}$ stochastically bounded for each $t > 0$, such that for $\gamma_n = C_n^{-1}$, $\{Y_n \circ \gamma_n\}$ is relatively compact.*

The next lemma gives conditions under which a process with values in a product space can be regularized.

Lemma 6.9 *For each $k = 1, 2, \ldots$, let (E_k, r_k) be a complete, separable metric space, and for each $n = 1, 2, \ldots$, let X_k^n be an $\{\mathcal{F}_t^n\}$-adapted process in $D_{E_k}[0,\infty)$. Let E denote the product space $E_1 \times E_2 \times \cdots$. Suppose, for each k, that $\{X_k^n\}$ is relatively compact (in the sense of convergence in distribution in the Skorohod topology). (This assumption implies that the sequence of E-valued processes $\{(X_1^n, X_2^n, \ldots)\}$ is relatively compact in $D_{E_1}[0,\infty) \times D_{E_2}[0,\infty) \times \cdots$ but not necessarily in $D_E[0,\infty)$). Then there exist strictly increasing processes $\{C_n\}$, such that $C_n(0) \ge 0$; for $t > s$, $C_n(t) - C_n(s) \ge t - s$; for each $t > 0$, $\{C_n(t)\}$ is stochastically bounded; for each n, C_n is $\{\mathcal{F}_t^n\}$-adapted; and defining $\gamma_n = C_n^{-1}$, the sequence $\{(\hat{X}_1^n, \hat{X}_2^n, \ldots)\}$ obtained by setting $\hat{X}_k^n = X_k^n \circ \gamma_n$ is relatively compact in $D_E[0,\infty)$.*

Proof. Recall that a sequence converging in $D_{E_1}[0,\infty) \times D_{E_2}[0,\infty) \times \cdots$ fails to converge in $D_E[0,\infty)$ if discontinuities in two of the components "coalesce". With that in mind, C_n should be constructed to slow down the time scale after a jump in such a way that coalescence of discontinuities is prevented. Let $N_k^n(t,r)$, $t, r > 0$ be the cardinality of the set $\{s : s \le t, r_k(X_k^n(s), X_k^n(s-)) \ge r\}$. The relative compactness of $\{X_k^n\}$ ensures the stochastic boundedness of $\{N_k^n(t,r)\}$ for each choice of t, r and k. For $m = 0, 1, 2, \ldots$ and $l = 1, 2, \ldots$, let $c_k(l, 2^{-m}) = \sup\{c : c \le 1, \sup_n P\{cN_k^n(l, 2^{-m}) > 2^{-m}\} \le 2^{-m}\}$ and for $l - 1 \le s < l$, define $c_k(s,r) = c_k(l, 2^{-m})$ for $2^{-m} \le r < 2^{-(m-1)}$, $m \ge 1$, and $c_k(s,r) = c_k(l,1)$ for $r \ge 1$. Then $C_n^k(t) = \sum_{s\le t} c_k(s, r_k(X_k^n(s), X_k^n(s-)))$ converges and

$$P\{C_n^k(l) \ge x\} \le P\{N_k^n(l, 2^{-m}) \ge x - l\} + l2^{-m}. \tag{6.3}$$

It follows from the stochastic boundedness of $\{N_k^n(l, 2^{-m})\}$ and the fact that m is arbitrary, that $\lim_{x\to\infty} \sup_n P\{C_n^k(l) \ge x\} = 0$. Finally, let $0 < a_{kl} \le 1$ satisfy

$$\sup_n P\{a_{kl}(C_n^k(l) - C_n^k(l-1)) \ge 2^{-k-l}\} \le 2^{-k-l}.$$

Define $a_k(s) = a_{kl}$, $l - 1 < s \le l$ and

$$C_n(t) = t + \sum_k \sum_{s\le t} a_k(s) c_k(s, r_k(X_k^n(s), X_k^n(s-))). \tag{6.4}$$

Noting that $C_n(l) = l + \sum_{m=1}^{l} \sum_k a_{km}(C_n^k(m) - C_n^k(m-1))$, the stochastic boundedness of $\{C_n(t)\}$ follows, as in (6.3), from the stochastic boundedness of the $\{C_n^k(t)\}$ and the definition of the a_k.

Note that $\gamma_n = C_n^{-1}$ is continuous, in fact, absolutely continuous with $\gamma_n' \leq 1$. Setting $\hat{X}_k^n = X_k^n \circ \gamma_n$, we have $\{(\gamma_n, \hat{X}_1^n, \hat{X}_2^n, \ldots)\}$ relatively compact in $S \equiv D_{\mathbf{R}}[0,\infty) \times D_{E_1}[0,\infty) \times D_{E_2}[0,\infty) \times \cdots$. The proof of the lemma follows by showing that any subsequence that converges in S also converges in $D_{\mathbf{R}\times E}[0,\infty)$. The Skorohod representation theorem (for example, Ethier and Kurtz (1986), Theorem 2.1.8) and the characterization of convergence in the Skorohod topology given in Proposition 2.6.5 of Ethier and Kurtz (1986) can be used to complete the proof. □

Lemma 6.10 *Suppose that $\{X_n\}$ is relatively compact in $D_E[0,\infty)$ and that C_n is nonnegative and satisfies $C_n(t) - C_n(s) \geq \alpha(t-s)$, $t > s \geq 0$ for some $\alpha > 0$. Define $\gamma_n = C_n^{-1}$ and $\hat{X}_n = X_n \circ \gamma_n$. Then $\{\hat{X}_n\}$ is relatively compact.*

Proof. Note that γ_n is differentiable with $\gamma_n' \leq \frac{1}{\alpha}$. Let $w'(x,\delta,T)$ denote the modulus of continuity for $D_E[0,\infty)$ defined in (3.6.2) of Ethier and Kurtz (1986). Then

$$w'(\hat{X}_n, \delta, T) \leq w'(X_n, \delta/\alpha, T/\alpha).$$

This estimate and the relative compactness of $\{X_n\}$ implies the relative compactness of $\{\hat{X}_n\}$ by Theorem 7.2 of Ethier and Kurtz (1986). □

The next lemma says that if a sequence of time changes regularized a sequence of processes then a sequence of time changes that grows more slowly will also regularize the sequence of processes.

Lemma 6.11 *Let X_n be cadlag, E-valued, and $\{\mathcal{F}_t^n\}$-adapted. Let C_n^1 and C_n^2 be strictly increasing and $\{\mathcal{F}_t^n\}$-adapted. Define $\gamma_n^1(t) = \inf\{u : C_n^1(u) > t\}$ and $\gamma_n(t) = \inf\{u : C_n^1(u) + C_n^2(u) > t\}$. Suppose $\{X_n \circ \gamma_n^1\}$ is relatively compact. Then $\{X_n \circ \gamma_n\}$ is relatively compact.*

Proof. Let $0 = t_0 < t_1 < \cdots$. Then there exist $0 = s_0 < s_1 < \cdots$ satisfying $\gamma_n(s_k) = \gamma_n^1(t_k)$ and $t_{k+1} - t_k \leq s_{k+1} - s_k$. It follows that $w'(X_n \circ \gamma_n, \delta, t) \leq w'(X_n \circ \gamma_n^1, \delta, t)$, where w' is the modulus of continuity defined in (3.6.2) in Ethier and Kurtz (1986). Since $\gamma_n(t) \leq \gamma_n^1(t)$, for any compact set $K \subset E$, $P\{X_n \circ \gamma_n(s) \in K, s \leq t\} \geq P\{X_n \circ \gamma_n^1(s) \in K, s \leq t\}$. The lemma follows by Theorem 3.7.2 and Remark 3.7.3 of Ethier and Kurtz (1986). □

Lemma 6.12 *If $\gamma_n = C_n^{-1}$ regularizes $\{X_n\}$, then for $a > 0$, γ_n^a given by $\gamma_n^a(t) = \inf\{s : aC_n(s) > t\}$ regularizes $\{X_n\}$.*

Proof. Note that $\gamma_n^a(t) = \gamma_n(t/a)$ and that if $\{Z_n\}$ is relatively compact, then $\{Z_n(\cdot/a)\}$ is relatively compact. □

Lemma 6.13 *Suppose for each $k = 1,2,\ldots$ and $n = 1,2,\ldots$ C_n^k is $\{\mathcal{F}_t^n\}$-adapted, $C_n^k(t) - C_n^k(s) \geq t - s$, $t > s \geq 0$, for each $t > 0$, $\{C_n^k(t)\}$ is stochastically bounded, and $\gamma_n^k(t) = \inf\{u : C_n^k(u) > t\}$ regularizes $\{X_n^k\}$ in $D_{E^k}[0,\infty)$. Then there exists $\{\mathcal{F}_t^n\}$-adapted C_n such that $C_n(t) - C_n(s) \geq t - s$, $t > s \geq 0$, for each $t > 0$, $\{C_n(t)\}$ is stochastically bounded, and $\gamma_n = C_n^{-1}$ regularizes $\{(X_n^1, X_n^2, \ldots)\}$ in $D_E[0,\infty)$, $E = E_1 \times E_2 \times \cdots$.*

Proof. First construct a $\tilde{C}_n$ such that for $\tilde{\gamma}_n = \tilde{C}_n^{-1}$, $\{(X_n^1 \circ \tilde{\gamma}_n, X_n^2 \circ \tilde{\gamma}_n, \ldots)\}$ is relatively compact in $D_{E_1}[0,\infty) \times D_{E_2}[0,\infty) \times \cdots$ and then perturb $\tilde{C}_n$ by a process constructed as in (6.4) to obtain the desired C_n. □

The next result extends Lemma 6.8 to sequences of $H^\#$ semimartingales.

Proposition 6.14 *Let $\{Y_n\}$ be a uniformly tight sequence of $H^\#$-semimartingales, Y_n adapted to $\{\mathcal{F}_t^n\}$, and for each n, let U_n be an $\{\mathcal{F}_t^n\}$-adapted process in $D_E[0,\infty)$ where E is a complete, separable metric space. Suppose $\{U_n\}$ is relatively compact (in the sense of convergence in distribution in the Skorohod topology). Then there exist strictly increasing, $\{\mathcal{F}_t^n\}$-adapted processes C_n, with $C_n(0) \geq 0$, $C_n(t+h) - C_n(t) \geq h$, $t, h \geq 0$, and $\{C_n(t)\}$ stochastically bounded for all $t \geq 0$, such that, defining $\gamma_n = C_n^{-1}$ and $\hat{Y}_n(\varphi, t) = Y_n(\varphi, \gamma_n(t))$, $\{\hat{Y}_n\}$ is uniformly tight and $\{(U_n \circ \gamma_n, \hat{Y}_n)\}$ is relatively compact.*

Proof. For C_n with the desired properties, the uniform tightness of $\{\hat{Y}_n\}$ follows from the fact that $\gamma_n' \leq 1$. By the uniform tightness of $\{\hat{Y}_n\}$, to prove relative compactness, it is enough to prove relative compactness of $\{(U_n \circ \gamma_n, \hat{Y}_n(\varphi_1, \cdot), \hat{Y}_n(\varphi_2, \cdot), \ldots)\}$ in $D_{E\times\mathbf{R}^\infty}[0,\infty)$ for a dense sequence $\{\varphi_k\}$ or for a sequence whose finite linear combinations are dense. By Lemma 6.8, for each $k = 1,2,\ldots$ there exists a strictly increasing, $\{\mathcal{F}_t^n\}$-adapted process $\tilde{C}_n^k$ such that $\tilde{\gamma}_n^k$ defined by $\tilde{\gamma}_n^k(t) = \inf\{s : \tilde{C}_n^k(s) > t\}$ regularizes $\{Y_n(\varphi_k, \cdot)\}$. Letting $\iota(t) = t$, there exist $a_k > 0$ such that $\tilde{C}_n = \iota + \sum_k a_k \tilde{C}_n^k$ exists and has the property that for each $t > 0$, $\{\tilde{C}_n(t)\}$ is stochastically bounded. Setting $\tilde{\gamma}_n = \tilde{C}_n^{-1}$, it follows from Lemmas 6.11 and 6.12, that $\{(U_n \circ \tilde{\gamma}_n, Y_n(\varphi_1, \tilde{\gamma}_n(\cdot)), Y_n(\varphi_2, \tilde{\gamma}_n(\cdot)), \ldots)\}$ is relatively compact in $D_E[0,\infty) \times D_{\mathbf{R}}[0,\infty) \times D_{\mathbf{R}}[0,\infty) \times \cdots$. As in Lemma 6.9, we must modify $\tilde{C}_n$ to ensure that discontinuities for different components do not coalesce. Setting, $X_0^n = U_n$ and $X_k^n = Y_n(\varphi_k, \cdot)$, let $N_k^n(t,r)$ be as in the proof of Lemma 6.9. Note that $\{N_0^n(t,r)\}$ is stochastically bounded by the assumption that $\{U_n\}$ is relatively compact and that for $k > 0$, $\{N_k^n(t,r)\}$ is stochastically bounded by the assumption that $\{Y_n\}$ is uniformly tight. (In particular, this assertion follows from the fact that $\{[Y_n]_t\}$ is stochastically bounded for each $t > 0$.) Consequently, we can add the analogue of the right side of (6.4) to $\tilde{C}_n$ defined above to obtain a nonnegative, strictly increasing, $\{\mathcal{F}_t^n\}$-adapted process C_n such that for $\gamma_n = C_n^{-1}$, $\{(U_n \circ \gamma_n, Y_n(\varphi_1, \gamma_n(\cdot)), Y_n(\varphi_2, \gamma_n(\cdot)), \ldots)\}$ is relatively compact in $D_E[0,\infty) \times D_{\mathbf{R}}[0,\infty) \times D_{\mathbf{R}}[0,\infty) \times \cdots$ (by Lemma 6.11) and discontinuities do not coalesce, so that, in fact, relative compactness is in $D_{E\times\mathbf{R}^\infty}[0,\infty)$ as desired. □

Lemma 6.15 *For each n, let Z_n be an $\{\mathcal{F}_t^n\}$-adapted process in $D_L[0,\infty)$, and suppose that the compact containment condition holds, that is, for each $\epsilon > 0$ and $t > 0$, there exists a compact $K_{\epsilon,t} \subset L$ such that $P\{Z_n(s) \in K_{\epsilon,t}, s \leq t\} \geq 1 - \epsilon$. Let $\{\lambda_i\} \subset L^*$ satisfy $\sup_i \langle \lambda_i, x\rangle = \|x\|_L$ for all $x \in L$, and suppose for each i, the sequence $\{\langle \lambda_i, Z_n\rangle\}$ is uniformly tight. Then there exist $\{\mathcal{F}_t^n\}$-adapted C_n satisfying $C_n(t) - C_n(s) \geq t - s$, $t > s \geq 0$, and*

$\{C_n(t)\}$ stochastically bounded for each $t > 0$, such that for $\gamma_n = C_n^{-1}$, $\{Z_n \circ \gamma_n\}$ is relatively compact.

Proof. By Lemma 6.8 there is a C_n^i such that the corresponding γ_n^i regularizes $\{\langle \lambda_i, Z_n \rangle\}$, and by Lemma 6.13, there exists a C_n such that $\{(\langle \lambda_1, Z_n \circ \gamma_n \rangle, \langle \lambda_2, Z_n \circ \gamma_n \rangle, \ldots)\}$ is relatively compact in $D_{\mathbf{R}^\infty}[0,\infty)$. The relative compactness of $\{Z_n \circ \gamma_n\}$ then follows from Theorem 3.9.1 of Ethier and Kurtz (1986). □

The following lemma generalizes Proposition 4.3 of Kurtz and Protter (1991).

Lemma 6.16 *For each $n = 1,2,\ldots$, let U_n be an $\{\mathcal{F}_t^n\}$-adapted process in $D_L[0,\infty)$ and let Y_n be an $\{\mathcal{F}_t^n\}$-$(L,\hat{H})^\#$-semimartingale. Suppose that $\{Y_n\}$ is uniformly tight and that $\{(U_n, Y_n)\}$ is relatively compact in the sense that $\{(U_n, Y_n(\varphi_1,\cdot),\ldots,Y_n(\varphi_m,\cdot))\}$ is relatively compact in $D_{L\times\mathbf{R}^m}[0,\infty)$ for any finite collection $\varphi_1,\ldots,\varphi_m \in H$. Let X_n, be an $\{\mathcal{F}_t^n\}$-adapted process in $D_{\hat{H}}[0,\infty)$. Define*

$$Z_n(t) = U_n(t) + X_{n-} \cdot Y_n(t)\ .$$

Let C_n be a strictly increasing, $\{\mathcal{F}_t^n\}$-adapted process with $C_n(0) \geq 0$ and $C_n(t+h) - C_n(t) \geq h$, $t, h \geq 0$. Suppose that $\{C_n(t)\}$ is stochastically bounded for all $t \geq 0$. Define $\gamma_n = C_n^{-1}$, $\hat{U}_n = U_n \circ \gamma_n$, $\hat{Y}_n = Y_n \circ \gamma_n$, and $\hat{X}_n = X_n \circ \gamma_n$, and suppose that $\{(\hat{U}_n, \hat{X}_n, \hat{Y}_n, \gamma_n)$ is relatively compact in the sense that

$$\{(\hat{U}_n, \hat{X}_n, (\hat{Y}_n(\varphi_1,\cdot),\ldots,\hat{Y}_n(\varphi_m,\cdot),\gamma_n)\}$$

is relatively compact in $D_{L\times\hat{H}\times\mathbf{R}^m\times\mathbf{R}}[0,\infty)$ for $\varphi_1,\ldots,\varphi_m \in H$. Then $\{(Z_n, U_n, Y_n)\}$ is relatively compact.

Proof. First replace X_n by X_n^ϵ defined by $X_n^\epsilon(t) = \sum c_{ij}^\epsilon(X_n(t)) f_i \varphi_j$ where the c_{ij}^ϵ are defined as in (5.2) giving

$$Z_n^\epsilon(t) = U_n(t) + X_{n-}^\epsilon \cdot Y_n(t) = U_n(t) + \sum_{i,j} f_i \int_0^t c_{ij}^\epsilon(X_n(t)) dY_n(\varphi_j, s)$$

and apply Kurtz and Protter (1991), Lemma 4.3, to obtain the relative compactness of $\{(Z_n^\epsilon, U_n, Y_n)\}$. The lemma then follows from the uniform approximation of X_n by X_n^ϵ and the uniform tightness of $\{Y_n\}$. □

The following lemma shows that the "uniform tightness" terminology is appropriate even in the infinite dimensional setting, that is, if we restrict our attention to integrands taking values in a compact set, then the distributions of the values of the integrals are "tight" in the usual weak convergence sense.

Lemma 6.17 *Let $\{Y_n\}$ be a uniformly tight sequence of standard $(L,\hat{H})^\#$-semimartingales. Then for each compact $\Gamma_0 \subset \hat{H}$ and $\eta > 0$, there exists a compact $\Gamma_1 \subset L$ such that for any Γ_0-valued, $\mathcal{F}_t^n$-adapted, cadlag process X_n, $P\{X_{n-} \cdot Y_n(s) \in \Gamma_1, s \leq t\} \geq 1 - \eta$.*

Proof. With $\{\zeta_k\}$ as in Section 5, let $X_n^\epsilon(t) = \sum_k \hat{\psi}_k^\epsilon(X_n(t))\zeta_k = \sum_{i,j} c_{ij}^\epsilon(X_n(t)) f_i\varphi_j$. Then since $\sup_{x\in K_1} c_{ij}^\epsilon(x) = 0$ except for finitely many values of i and j, say $1 \le i \le I$ and $1 \le j \le J$, and the coefficient of f_i in the integral $X_{n-}^\epsilon \cdot Y_n(t)$ will be of the form

$$\sum_{j=1}^{J} \int_0^t c_{ij}^\epsilon(X_n(s-))dY_n(\varphi_j, s)$$

where the c_{ij}^ϵ are uniformly bounded on Γ_0, for any $\eta_0 > 0$ there will exist a compact set of the form

$$\Gamma = \{\sum_{i=1}^{I} \alpha_i f_i : |\alpha_i| \le a_i\} \tag{6.5}$$

such that

$$P\{X_{n-}^\epsilon \cdot Y_n(s) \in \Gamma, s \le t\} \ge 1 - \eta_0$$

for all Γ_0-valued, $\mathcal{F}_t^n$-adapted, cadlag processes X_n. Let $\Gamma^a = \{x : \|x - \Gamma\|_L < a\}$. Then for $\delta > 0$,

$$\begin{aligned} P\{X_{n-} \cdot Y_n(s) \notin \Gamma^{\epsilon K(t,\delta)}, \text{ some } s \le t\} &\le \eta_0 + P\{\sup_{s\le t} \|(X_{n-} - X_{n-}^\epsilon) \cdot Y_n(s)\|_L \ge \epsilon K(t,\delta)\} \\ &\le \eta_0 + \delta\,. \end{aligned}$$

For $m \ge 2$, let $\delta_m = \eta 2^{-m}$, let $\epsilon_m = 1/(mK(t,\delta_m))$, and select Γ_m of the form (6.5) so that

$$P\{X_{n-}^{\epsilon_m} \cdot Y_n(s) \in \Gamma_m, s \le t\} \ge 1 - \eta 2^{-m}$$

for all Γ_0-valued, $\mathcal{F}_t^n$-adapted, cadlag processes X_n. Then

$$P\{X_{n-} \cdot Y_n(s) \notin \Gamma_m^{1/m}, \text{ some } s \le t\} \le \eta 2^{-(m-1)},$$

and letting Γ_1 denote the closure of $\cap_{m=2}^\infty \Gamma_m^{1/m}$, we have

$$P\{X_{n-} \cdot Y_n(s) \notin \Gamma_1, \text{ some } s \le t\} \le \sum_{m=2}^{\infty} \eta 2^{-(m-1)} = \eta.$$

Note that Γ_1 is compact since it is complete and totally bounded. □

The following lemma is an immediate consequence of Lemma 6.17.

Lemma 6.18 *Let $\{Y_n\}$ be a uniformly tight sequence of standard $(L,\hat{H})^\#$-semimartingales. For each $n = 1, 2, \ldots$, let X_n be an $\{\mathcal{F}_t^n\}$-adapted process in $D_{\hat{H}}[0,\infty)$, and suppose that $\{X_n\}$ satisfies the compact containment condition. Then $\{X_{n-} \cdot Y_n\}$ satisfies the compact containment condition, that is, for $\eta > 0$, there exists a compact $\Gamma \subset L$ such that $P\{X_{n-} \cdot Y_n(s) \in \Gamma, s \le t\} \ge 1 - \eta$.*

7 Stochastic differential equations.

In this section we consider stochastic differential equations of the form

$$X(t) = U(t) + F(X, \cdot -) \cdot Y(t) , \tag{7.1}$$

where $F : D_L[0,\infty) \to D_{\hat{H}}[0,\infty)$ and F is nonanticipating in the sense that for $x \in D_L[0,\infty)$ and $x^t(\cdot) = x(\cdot \wedge t)$, $F(x,t) = F(x^t,t)$ for all $t \geq 0$. The usual Lipschitz condition on F implies uniqueness for equations driven by $(L,\hat{H})^\#$-semimartingales satisfying a condition slightly stronger than the assumption that Y is standard. Weak existence, under certain continuity assumptions, follows from a convergence theorem given below (see Corollary 7.7), and strong existence follows from weak existence and strong uniqueness.

7.1 Uniqueness for stochastic differential equations.

Theorem 7.1 *Let Y be an $(L,\hat{H})^\#$-semimartingale adapted to $\{\mathcal{F}_t\}$. Suppose that for each $\{\mathcal{F}_t\}$-stopping time τ, bounded by a constant, and $t, \delta > 0$, there exists $K_\tau(t,\delta)$ such that*

$$P\{\|Z_- \cdot Y(\tau + t) - Z_- \cdot Y(\tau)\|_L \geq K_\tau(t,\delta)\} \leq \delta \tag{7.2}$$

for all $Z \in S_{\hat{H}}$ satisfying $\sup_s \|Z(s)\|_{\hat{H}} \leq 1$, *and that* $\lim_{t\to 0} K_\tau(t,\delta) = 0$. *Suppose that there exists $M > 0$ such that*

$$\sup_{s\leq t} \|F(x,s) - F(y,s)\|_{\hat{H}} \leq M \sup_{s\leq t} \|x(s) - y(s)\|_L$$

for all $x, y \in D_L[0,\infty)$. Then there is at most one solution of (7.1).

Remark 7.2 *Note that if L and H are finite dimensional and Y is a finite dimensional semimartingale, then the hypothesized estimate (7.2) holds.*

More generally, if L satisfies the conditions of Lemma 5.9, then (7.2) will hold for any standard $(L,\hat{H})^\#$-semimartingale.

Finally, one can apply Lemma 5.10 to prove uniqueness for any standard $(L,\hat{H})^\#$-semimartingale under the additional condition on F that for each compact $\Gamma \subset L$ there exists a compact $K_\Gamma \subset \hat{H}$ such that $x(s) \in \Gamma$, $s \leq t$, implies $F(x,t) \in K_\Gamma$. (See Theorem 7.6 for another application of a related condition.) In particular, if $F(x,t) = f(x(t))$ for a Lipschitz continuous $f : L \to \hat{H}$, then uniqueness holds for all standard $(L,\hat{H})^\#$-semimartingales.

Proof. Without loss of generality, we can assume that $K_\tau(t,\delta)$ is a nondecreasing function of t.

Suppose X and $\tilde{X}$ satisfy (7.1). Let $\tau_0 = \inf\{t : \|X(t) - \tilde{X}(t)\|_L > 0\}$, and suppose $P\{\tau_0 < \infty\} > 0$. Select $r, \delta, t > 0$, such that $P\{\tau_0 < r\} > \delta$ and $MK_{\tau_0 \wedge r}(t,\delta) < 1$. Note that if $\tau_0 < \infty$, then

$$X(\tau_0) - \tilde{X}_0(\tau_0) = (F(X, \cdot -) - F(\tilde{X}, \cdot -)) \cdot Y(\tau_0) = 0. \tag{7.3}$$

Define

$$\tau_\epsilon = \inf\{s : \|X(s) - \tilde{X}(s)\|_L \geq \epsilon\}.$$

Noting that $\|X(s) - \tilde{X}(s)\|_L \le \epsilon$ for $s < \tau_\epsilon$, we have

$$\|F(X, s) - F(\tilde{X}, s)\|_{\hat{H}} \le \epsilon M,$$

for $s < \tau_\epsilon$, and

$$\|F(X, \cdot-) \cdot Y(\tau_\epsilon) - F(\tilde{X}, \cdot-) \cdot Y(\tau_\epsilon)\|_L = \|X(\tau_\epsilon) - \tilde{X}(\tau_\epsilon)\|_L \ge \epsilon.$$

Consequently, for $r > 0$, letting $\tau_0^r = \tau_0 \wedge r$, we have

$$\begin{aligned}
&P\{\tau_\epsilon - \tau_0^r \le t\} \\
&\quad \le P\{\sup_{s \le t \wedge (\tau_\epsilon - \tau_0^r)} \|F(X, \cdot-) \cdot Y(\tau_0^r + s) - F(\tilde{X}, \cdot-) \cdot Y(\tau_0^r + s)\|_L \ge \epsilon M K_{\tau_0^r}(t, \delta)\} \\
&\quad \le \delta.
\end{aligned}$$

Since the right side does not depend on ϵ and $\lim_{\epsilon \to 0} \tau_\epsilon = \tau_0$, it follows that $P\{\tau_0 - \tau_0 \wedge r < t\} \le \delta$ and hence that $P\{\tau_0 < r\} \le \delta$, contradicting the assumption on δ and proving that $\tau_0 = \infty$ a.s.

□

Example 7.3 *Equation for spin-flip models.*

A spin-flip model, for example on the lattice $\mathbb{Z}^d$, is a stochastic process whose state $\eta = \{\eta_i : i \in \mathbb{Z}^d\}$ assigns to each lattice point $i \in \mathbb{Z}^d$ the value ± 1. The model is prescribed by specifying for each $i \in \mathbb{Z}^d$ a flip rate c_i which determines the rate at which the associated state variable η_i changes sign. The rates c_i may depend on the full configuration η. We formulate a slightly more general model by tracking the cumulative number of sign changes X_i rather than just the current sign. Of course, if the initial configuration is known, then the current configuration can be recovered by the formula

$$\eta_i(t) = \eta_i(0)(-1)^{X_i(t) - X_i(0)}.$$

The model, then, consists of a collection of counting processes $X = \{X_i : i \in \mathbb{Z}^d\}$, and the specification of the rates c_i corresponds to the requirement that there exist a filtration $\{\mathcal{F}_t\}$ such that for each $i \in \mathbb{Z}^d$,

$$X_i(t) - \int_0^t c_i(X(s))ds. \tag{7.4}$$

is an $\{\mathcal{F}_t\}$-martingale. To completely specify the martingale problem corresponding to the $\{c_i\}$ we must also require that no two X_i have simultaneous jumps (that is, the martingales are orthogonal).

We formulate a corresponding system of stochastic differential equations for the X_i. Let $\{N_i : i \in \mathbb{Z}^d\}$ be independent Poisson random measures on $[0, \infty) \times [0, \infty)$ with Lebesgue mean measure. Then we require

$$X_i(t) = X_i(0) + \int_{[0,\infty) \times [0,t]} \sigma_i(X(s-), z) N_i(dz \times ds) \tag{7.5}$$

where

$$\sigma_i(x,z) = \begin{cases} 1 & 0 \le z \le c_i(x) \\ 0 & otherwise \end{cases}$$

Note that any solution of this system will be a family of counting process without simultaneous jumps (by the independence of the N_i). Letting $\tilde{N}_i(A) = N_i(A) - m(A)$, we can rewrite (7.5) as

$$X_i(t) = X_i(0) + \int_{[0,\infty)\times[0,t]} \sigma_i(X(s-),z)\tilde{N}_i(dz \times ds) + \int_0^t c_i(X(s))ds\,. \tag{7.6}$$

The second term on the right will be a martingale (at least, for example, if the c_i are bounded), and hence (7.4) will be a martingale.

We now give conditions under which the solution of (7.5) is unique. Let $J = \{\{k_i, i \in \mathbb{Z}^d\} : k_i \in \mathbb{Z}_+\}$. Assume

$$|c_i(x+e_j) - c_i(x)| \le a_{ij} \qquad |c_i(x)| \le b_i$$

for all $x, y \in J$ and $i, j \in \mathbb{Z}^d$ where e_j is the element of J such $e_{ji} = 0$ for $i \ne j$ and $e_{jj} = 1$. In addition, assume that there exist $\alpha_i > 0$ and $C > 0$ such that

$$\sum_i \alpha_i a_{ij} \le C\alpha_j \qquad \sum_i \alpha_i b_i < \infty. \tag{7.7}$$

These conditions are similar to those under which Liggett (1972) (see also Liggett (1985), Chapter III) proves uniqueness of the martingale problem for the spin-flip model. To give a complete proof of Liggett's theorem, we would need to show that any solution of the martingale problem is a weak solution of (7.5). We do not pursue this issue here, but see Kurtz (1980) for a closely related approach involving a different system of stochastic equations. Uniqueness of (7.5) can actually be proved under the same conditions as used in Kurtz (1980) using different methods; however, our point here is to illustrate the breadth of applicability of the theorem above.

In order to interpret the system as a solution of a single equation of the form (7.1), let $U = [0,\infty) \times \mathbb{Z}^d$ and $E = \mathbb{Z}^d$, and define $F_i(x,u) = \sigma_i(x,z)\delta_{ij}$ for $u = (z,j)$. Define

$$\begin{aligned} \|f\|_L &= \sum_i \alpha_i |f(i)| \\ \|\varphi\|_H &= \sum_j \int_0^\infty |\varphi(z,j)|dz \\ \|g\|_{\hat{H}} &= \sum_i \sum_j \alpha_i \int_{\mathbb{R}} |g(i,z,j)|du = \|\|g\|_H\|_L \end{aligned}$$

Then

$$\|F_\cdot(x,\cdot)\|_{\hat{H}} \le \sum_i \alpha_i \int_0^\infty |\sigma_i(x,z)|dz = \sum_i \alpha_i c_i(x) \le \sum_i \alpha_i b_i$$

and

$$\begin{aligned} \|F_\cdot(x,\cdot) - F_\cdot(y,\cdot)\|_{\hat{H}} &= \textstyle\sum_i \alpha_i \int_0^\infty |\sigma_i(x,z) - \sigma_i(y,z)|dz \\ &= \textstyle\sum_i \alpha_i |c_i(x) - c_i(y)| \\ &\le \textstyle\sum_i \alpha_i \sum_j a_{ij}|x_j - y_j| \\ &\le \textstyle\sum_j C\alpha_j |x_j - y_j| = C\|x-y\|_L \end{aligned}$$

so F is bounded and Lipschitz.

For many examples $a_{ij} = \rho(i-j)$, where ρ has bounded support and $b_i \equiv b$. One then can take $\alpha_i = \frac{1}{1+|i|^{d+1}}$ so that

$$\sum_i \frac{\rho(i-j)}{1+|i|^{d+1}} \leq \left(\sup_k \sum_i \frac{1+|k|^{d+1}}{1+|i|^{d+1}}\rho(i-k)\right)\frac{1}{1+|j|^{d+1}}$$

which gives (7.7).

To verify (7.2), suppose

$$\|X(s)\|_{\hat{H}} = \sum_i \sum_j \alpha_i \int_0^\infty |X(i,z,j)|dz \leq 1 .$$

Then

$$\begin{aligned} E[\|X_- \cdot N(t)\|_L] &= E[\sum_i \sum_j \alpha_i \int_{[0,\infty)\times[0,t]} X(s-,i,z,j)N_j(dz \times ds)] \\ &\leq E[\sum_i \sum_j \alpha_i \int_{[0,\infty)\times[0,t]} |X(s-,i,z)|dzds] \\ &\leq t \end{aligned}$$

7.2 Sequences of stochastic differential equations.

Consider a sequence of equations of the above form

$$X_n(t) = U_n(t) + F_n(X_n, \cdot-) \cdot Y_n(t). \tag{7.8}$$

The following analogue of Proposition 5.1 of Kurtz and Protter (1991) is an immediate consequence of Theorems 4.2 and 5.5.

Proposition 7.4 *Suppose (U_n, X_n, Y_n) satisfies (7.8), that $\{(U_n, X_n, Y_n)\}$ is relatively compact, that $(U_n, Y_n) \Rightarrow (U, Y)$, and that $\{Y_n\}$ is uniformly tight. Assume that F_n and F satisfy the following continuity condition:*

C.1 *If $(x_n, y_n) \rightarrow (x, y)$ in $D_{L\times\mathbb{R}^m}[0,\infty)$, then*

$$(x_n, y_n, F_n(x_n, \cdot)) \rightarrow (x, y, F(x, \cdot))$$

in $D_{L\times\mathbb{R}^m\times\hat{H}}[0,\infty)$.

Then any limit point of $\{X_n\}$ satisfies (7.1).

In many situations, the relative compactness of $\{X_n\}$ is easy to verify; however, with somewhat stronger conditions on the sequence $\{F_n\}$ (satisfied, for example, if $F_n(x,t) = f_n(x(t))$ for a sequence of continous mappings $f_n : \mathbb{R}^k \rightarrow H$ converging uniformly on compact subsets of $\mathbb{R}^k$ to $f : \mathbb{R}^k \rightarrow H$) we can drop the a priori assumption of relative compactness of $\{X_n\}$. Let $T_1[0,\infty)$ be the collection of nondecreasing mappings λ of $[0,\infty)$ onto $[0,\infty)$ satisfying $|\lambda(t)-\lambda(s)| \le |t-s|$. Let the topology on $T_1[0,\infty)$ be given by the metric

$$d_1(\lambda_1,\lambda_2) = \sup_{0\le t<\infty} \frac{|\lambda_1(t)-\lambda_2(t)|}{1+|\lambda_1(t)-\lambda_2(t)|}\,.$$

ι will denote the identity map, $\iota(t)=t$. We assume the following:

C.2a There exist $G_n, G : D_L[0,\infty)\times T_1[0,\infty) \to D_{\hat{H}}[0,\infty)$ such that $F_n(x)\circ\lambda = G_n(x\circ\lambda,\lambda)$ and $F(x)\circ\lambda = G(x\circ\lambda,\lambda)$, $x\in D_L[0,\infty)$, $\lambda\in T_1[0,\infty)$.

C.2b For each compact $\mathcal{K}\subset D_L[0,\infty)\times T_1[0,\infty)$ and $t>0$,

$$\sup_{(x,\lambda)\in\mathcal{K}} \sup_{s\le t} \|G_n(x,\lambda,s)-G(x,\lambda,s)\|_{\hat{H}} \to 0.$$

C.2c For $\{(x_n,\lambda_n)\}\in D_L[0,\infty)\times T_1[0,\infty)$, $\sup_{s\le t}\|x_n(s)-x(s)\|_L\to 0$ and $\sup_{s\le t}|\lambda_n(s)-\lambda(s)|\to 0$ for each $t>0$ implies $\sup_{s\le t}\|G(x_n,\lambda_n,s)-G(x,\lambda,s)\|_{\hat{H}}\to 0$.

Note that if $F(x,t)=f(x(t))$, then $G(x,\lambda,t)=f(x(t))$; if $F(x,t)=f(\int_0^t h(x(s))ds)$, then $G(x,\lambda,t)=f(\int_0^t h(x(s))\lambda'(s)ds)$. (See Kurtz and Protter (1991) for additional examples.)

The following theorem is the analogue of Theorem 5.4 of Kurtz and Protter (1991). For simplicity, we assume that F_n and F are uniformly bounded. The localization argument used in the earlier theorem can again be applied here to extend the result to the unbounded case.

Theorem 7.5 *Let $L=\mathbb{R}^k$. Suppose (U_n,X_n,Y_n) satisfies (7.8), that $(U_n,Y_n)\Rightarrow(U,Y)$, and that $\{Y_n\}$ is uniformly tight. Assume that $\{F_n\}$ and F satisfy Condition C.2 and that $\sup_n\sup_x\|F_n(x,\cdot)\|_{H^k}<\infty$. Then $\{(U_n,X_n,Y_n)\}$ is relatively compact and any limit point satisfies (7.1).*

Proof. Since Condition C.2 implies Condition C.1, it is enough to show the relative compactness of (U_n,X_n,Y_n). By Proposition 6.2 $\{F_n(X_n,\cdot-)\cdot Y_n\}$ is uniformly tight. By Proposition 6.14, there exists a γ_n such that $\{(U_n\circ\gamma_n, F_n(X_n,\cdot-)\cdot Y_n\circ\gamma_n)\}$ is relatively compact, which in turn implies $\{(X_n\circ\gamma_n, U_n\circ\gamma_n, F_n(X_n,\cdot-)\cdot Y_n\circ\gamma_n)\}$ is relatively compact. Setting $\hat{X}_n = X_n\circ\gamma_n$, etc., Condition C.2a implies

$$\hat{X}_n(t) = \hat{U}_n(t) + G_n(\hat{X}_n,\gamma_n,\cdot-)\cdot\hat{Y}_n(t)$$

and the relative compactness of $\{(\hat{X}_n,\hat{U}_n,\hat{Y}_n)\}$ implies the relative compactness of

$$\{(\hat{X}_n,\hat{U}_n,\hat{Y}_n,G_n(\hat{X}_n,\gamma_n,\cdot),\gamma_n)\}\,.$$

By Lemma 6.16, $\{(X_n,U_n,Y_n)\}$ is relatively compact, and the theorem follows from Proposition 7.4. □

Theorem 7.6 *Suppose (U_n, X_n, Y_n) satisfies (7.8), that $(U_n, Y_n) \Rightarrow (U, Y)$, and that $\{Y_n\}$ is uniformly tight. Assume that $\{F_n\}$ and F satisfy Condition C.2, that $\sup_n \sup_x \|F_n(x,\cdot)\|_{\hat{H}} < \infty$, and that for each $\kappa > 0$ there exists a compact $K_\kappa \subset \hat{H}$ such that $\sup_{s\le t} \|x(s)\|_L \le \kappa$ implies $F_n(x,t) \in K_\kappa$ for all n. Then $\{(U_n, X_n, Y_n)\}$ is relatively compact and any limit point satisfies (7.1).*

Proof. Again it is sufficient to verify the relative compactness of $\{(U_n, X_n, Y_n)\}$. The uniform tightness of $\{Y_n\}$ and the boundedness of the F_n imply the stochastic boundedness of $\{\sup_{s\le t} \|X_n(s)\|_L\}$. The compactness condition on the F_n then ensures that $\{F_n(X_n,\cdot)\}$ satisfies the compact containment condition. The sequence $\{F_n(X_n, \cdot-) \cdot Y_n\}$ satisfies the compact containment condition by Lemma 6.18 which in turn implies $\{X_n\}$ satisfies the compact containment condition. Lemma 6.15 then ensures the existence of γ_n as in the proof of Theorem 7.5, and the remainder of the proof is the same. □

Corollary 7.7 *a) Let $L = \mathbb{R}^k$. Suppose that F and G satisfy C.2a and C.2c, that*

$$\sup_x \|F(x,\cdot)\|_{H^k} < \infty,$$

and that Y is a standard $H^\#$-semimartingale. Then weak existence holds for (7.1).

b) For general L, suppose that F and G satisfy C.2a and C.2c, that $\sup_x \|F(x,\cdot)\|_{\hat{H}} < \infty$, and that F satisfies the compactness condition of Theorem 7.6, and that Y is a standard $(L,\hat{H})^\#$-semimartingale. Then weak existence holds for (7.1).

Proof. Let $F_n \equiv F$, and define $U_n(t) = U(\frac{[nt]}{n})$ and $Y_n(\varphi, t) = Y(\varphi, \frac{[nt]}{n})$. Let X_n be the solution of

$$X_n(t) = U_n(t) + F(X_n, \cdot-) \cdot Y_n(t)$$

which is easily seen to exist since U_n and Y_n are constant except for a discrete set of jumps. Then X_n, U_n, Y_n, and F_n satisfy the conditions of Theorem 7.5 in part (a) and of Theorem 7.6 in part (b). Consequently, a subsequence of X_n will converge in distribution to a process $\tilde{X}$ satisfying $\tilde{X}(t) = \tilde{U}(t) + F(\tilde{X}, \cdot-)\cdot\tilde{Y}(t)$, where $(\tilde{U}, \tilde{Y})$ has the same distribution as (U, Y), that is, $\tilde{X}$ is a weak solution of (7.1). □

The following corollary is the analogue, in the present setting, of a result of Yamada and Watanabe (1971) stating that weak existence and strong uniqueness imply strong existence. (See Engelbert (1991) for a more recent discussion.) The proof of the corollary is the same as that of Lemma 5.5 of Kurtz and Protter (1991). We say that strong uniqueness holds for (7.1) if any two solutions X_1 and X_2 satisfy $X_1 = X_2$ a.s.

Corollary 7.8 *Suppose, in addition to the conditions of Corollary 7.7, that strong uniqueness holds for (7.1) for any version of (U, Y) for which Y is a standard $H^\#$ (or $(L,\hat{H})^\#$)-semimartingale. Then any solution of (7.1) is a measurable function of (U, Y), that is, if X satisfies (7.1) and the finite linear combinations of $\{\varphi_k\}$ are dense in H, there exists a measurable mapping*

$$g : D_{L\times\mathbb{R}^\infty}[0,\infty) = D_L[0,\infty)$$

such that $X = g(U, Y(\varphi_1,\cdot), Y(\varphi_2,\cdot), \ldots)$. In particular, there exists a strong solution of (7.1).

Remark 7.9 *Note that under the conditions of Theorem 7.1 (see Remark 7.2), the strong uniqueness hypothesis of the present Corollary will hold.*

The proof of the following corollary is essentially the same as that of Corollary 5.6 of Kurtz and Protter (1991).

Corollary 7.10 *Suppose, in addition to the conditions of Theorem 7.5 or Theorem 7.6, that strong uniqueness holds for any version of (U,Y) for which Y is a standard $H^\#$ $(L,\hat{H})^\#$-semimartingale and that $(U_n,Y_n) \to (U,Y)$ in probability. Then $(U_n,Y_n,X_n) \to (U,Y,X)$ in probability.*

8 Markov processes.

An $H^\#$-semimartingale has *stationary independent increments* if $(Y(\varphi_1,\cdot),\ldots,Y(\varphi_m,\cdot))$ has stationary independent increments for each choice of $\varphi_1,\ldots,\varphi_m \in H$. If Y is a standard $(L,\hat{H})^\#$-semimartingale with stationary independent increments, $F: L \to \hat{H}$, and the equation

$$X(t) = x_0 + F(X(\cdot-)) \cdot Y(t) \tag{8.1}$$

has a unique solution for each $x_0 \in L$, then X is a temporally homogeneous Markov process. If Y is given by a standard semimartingale random measure, then (8.1) can be written

$$X(t) = x_0 + \int_{U\times[0,t]} F(X(s-),u)Y(du \times ds)\,.$$

For $\varphi_1,\ldots,\varphi_m \in H$, $(Y(\varphi_1,\cdot),\ldots,Y(\varphi_m,\cdot))$ has a generator of the form (see, for example, Ethier and Kurtz (1986), Theorem 8.3.4)

$$\begin{aligned} B(\varphi_1,\ldots,\varphi_m)f(x) &= \frac{1}{2}\sum_{i,j=1}^{m} a_{ij}(\varphi_1,\ldots,\varphi_m)\partial_i\partial_j f(x) + \sum_{i=1}^{m} b_i(\varphi_1,\ldots,\varphi_m)\partial_i f(x) \\ &\quad + \int_{\mathbb{R}^m}\left(f(x+y) - f(x) - \frac{y\cdot\nabla f(x)}{1+|y|^2}\right)\mu(\varphi_1,\ldots,\varphi_m,dy) \end{aligned}$$

where it is enough to consider $f \in C_c^\infty(\mathbb{R}^m)$, that is, the closure of $B(\varphi_1,\ldots,\varphi_m)$ defined on $C_c^\infty(\mathbb{R}^m)$, the space of infinitely differentiable functions with compact support, generates the strongly continuous contraction semigroup on $\hat{C}(\mathbb{R}^m)$ corresponding to the process $(Y(\varphi_1,\cdot),\ldots,Y(\varphi_m,\cdot))$. The assumption that Y is standard $H^\#$-semimartingale implies $B(\varphi_1,\ldots,\varphi_m)f$ is a continuous function of $(\varphi_1,\ldots,\varphi_m)$. (Note, however, that this continuity does not imply the continuity of a_{ij}.)

Suppose $L = \mathbb{R}^k$. Formally, at least, the solution of (8.1) should have a generator given by

$$Af(x) = B(F_1(x),\ldots,F_k(x))f(x),$$

that is, the solution of (8.1) is a solution of the martingale problem for A.

Consider a sequence of stochastic differential equations

$$X_n(t) = X_n(0) + \int_{U\times[0,t]} F_n(X(s-),u)Y_n(du \times ds)$$

where Y_n is an $H^{\#}$-semimartingale with stationary, independent increments, and let the generator for $(Y_n(\varphi_1,\cdot),\ldots,Y_n(\varphi_m,\cdot))$ be denoted $B_n(\varphi_1,\ldots,\varphi_m)$. The following theorem is an immediate consequence of Theorem 7.5 and Theorems 1.6.1 and 4.2.5 of Ethier and Kurtz (1986).

Theorem 8.1 *Let $L=\mathbb{R}^k$. Let F in (8.1) be bounded and Lipschitz. Assume that $\{Y_n\}$ is uniformly tight, Y_n is independent of $X_n(0)$, and $X_n(0)\Rightarrow X(0)$. Suppose for each $\varphi_1,\ldots,\varphi_m\in H$ and each $f\in C_c^\infty(\mathbb{R}^m)$*

$$\lim_{n\to\infty}\sup_x|B_n(\varphi_1,\ldots,\varphi_m)f(x)-B(\varphi_1,\ldots,\varphi_m)f(x)|=0, \tag{8.2}$$

and for each compact $K\subset\mathbb{R}^k$

$$\lim_{n\to\infty}\sup_{x\in K}\|F_n(x)-F(x)\|_{H^k}=0.$$

Then $(X_n,Y_n)\Rightarrow(X,Y)$, where Y is a standard $H^{\#}$-semimartingale, $(Y(\varphi_1,\cdot),\ldots,Y(\varphi_m,\cdot))$ has generator $B(\varphi_1,\ldots,\varphi_m)$, and X is the unique solution of

$$X(t)=X(0)+F(X(\cdot-))\cdot Y(t).$$

Proof. The convergence of B_n to B implies the convergence $(Y_n(\varphi_1,\cdot),\ldots,Y(\varphi_m,\cdot))\Rightarrow(Y(\varphi_1,\cdot),\ldots,Y(\varphi_m,\cdot))$, and the theorem follows from Theorem 7.5 □

For a specific form for (8.1), let W denote Gaussian white noise on $U_0\times[0,\infty)$ with $E[W(A,t)^2]=\mu(A)t$, let N_1 and N_2 be Poisson random measures on $U_1\times[0,\infty)$ and $U_2\times[0,\infty)$ with σ-finite mean-measures $\nu_1\times m$ and $\nu_2\times m$, respectively. Let $M_1(A,t)=N_1(A,t)-\nu_1(A)t$, and note that

$$\langle M_1(A,\cdot),M_1(B,\cdot)\rangle_t=t\nu_1(A\cap B).$$

Consider the equation for an $\mathbb{R}^k$-valued process X

$$\begin{aligned}X(t)=X(0)\ &+\ \int_{U_0\times[0,t]}F_0(X(s-),u)W(du\times ds)+\int_{U_1\times[0,t]}F_1(X(s-),u)M_1(du\times ds)\\ &+\ \int_{U_2\times[0,t]}F_2(X(s-),u)N_2(du\times ds)+\int_0^t F_3(X(s-))ds\,.\end{aligned} \tag{8.3}$$

Note that the above equation is essentially the same as that originally considered by Itô (1951). (The diffusion term in Itô's equation was driven by a finite dimensional standard Brownian motion rather than the Gaussian white noise.) The generality of this equation is demonstrated in the work of Çinlar and Jacod (1981).

Theorem 8.2 *Suppose that there exists $M>0$ such that*

$$\begin{aligned}&\sqrt{\int_{U_0}|F_0(x,u)-F_0(y,u)|^2\mu(du)}+\sqrt{\int_{U_1}|F_1(x,u)-F_1(y,u)|^2\nu_1(du)}\\ &\qquad+\int_{U_2}|F_2(x,u)-F_2(y,u)|\nu_2(du)+|F_3(x)-F_3(y)|\ \le\ M|x-y|\,.\end{aligned}$$

Then there exists a unique solution of (8.3).

Remark 8.3 *This result is essentially Theorem 1.2 of Graham (1992), the only difference being that we allow general Gaussian white noise and Graham only considers finite dimensional Brownian motion. His methods, however, which employ an L_1-martingale inequality instead of the typical application of Doob's L_2-martingale inequality (see Lemma 2.4 and additional discussion below), could also be used in the present setting. As Graham points out, the ability to use L_1 estimates is crucial for many applications. For example, consider the equation corresponding to the generator*

$$Af(x) = \int_{\mathbb{R}^d} \lambda(x,u)(f(x+u)-f(x))\nu(du)$$

given by

$$X(t) = X(0) + \int_{\mathbb{R}^d \times [0,\infty)} uI_{[0,\lambda(X(s-),u)]}(z)N(du \times dz \times ds),$$

where N is the Poisson random measure on $\mathbb{R}^d \times [0,\infty)$ with mean measure $\eta = \nu \times m$. The $L_1(\eta)$ Lipschitz condition becomes

$$\int_{\mathbb{R}^d} u|\lambda(x,u) - \lambda(y,u)|\nu(du) \leq K|x-y|$$

which will be satisfied under reasonable conditions on λ. Since the square of an indicator is the indicator, the corresponding L_2 condition would require

$$\int_{\mathbb{R}^d} u^2|\lambda(x,u) - \lambda(y,u)|\nu(du) \leq K|x-y|^2,$$

which essentially says that $\lambda(x,u)$ is constant in x. The classical conditions of Itô (1951) (see also Ikeda and Watanabe (1981)) as well as more recent work (for example, Kallianpur and Xiong (1994)) based on L_2 estimates do not cover this example. Roughly speaking, L_2 estimates will work if the jump sizes vary smoothly with the location of the process and the jump rates are constant, but fail if the jump rates vary.

Proof. Let $H_0 = L_2(\mu)$, $H_1 = L_2(\nu_1)$, $H_2 = L_1(\nu_2)$, and $H_3 = \mathbb{R}$, and let $Y_0(\varphi,t) = \int_{U_0} \varphi(u)W(du,t)$ for $\varphi \in L_2(\mu)$, $Y_1(\varphi,t) = \int_{U_1} \varphi(u)M_1(du,t)$ for $\varphi \in H_1$, $Y_2(\varphi,t) = \int_{U_2} \varphi(u)N_2(du,t)$ for $\varphi \in H_2$, and $Y_3(\varphi,t) = \varphi t$ for $\varphi \in \mathbb{R}$. Y_0 and Y_1 are standard $H_0^\#$- and $H_1^\#$-semimartingales, respectively, by Lemma 3.3. For $Z \in S_{H_2}$, we have

$$E[|Z_- \cdot Y_2(t)|] \leq E[\int_{U_2 \times [0,t]} |Z(s-,u)|N_2(du \times ds)] = E[\int_0^t \int_{U_2} |Z(s-,u)|\nu_2(du)ds]\ ,$$

which implies Y_2 is a standard $H_2^\#$-semimartingale. Trivially, Y_3 is a standard $H_3^\#$-semimartingale. Setting $H = H_0 \times H_1 \times H_2 \times H_3$, Y defined as in Lemma 3.10 is an $H^\#$-semimartingale, and the theorem follows by Theorem 7.1 and Corollaries 7.7 and 7.8. □

Let X and $\tilde{X}$ be solutions of (8.3). Graham's approach (cf. Remark 8.3) depends on the inequality

$$E[\sup_{s \leq t} |X(s) - \tilde{X}(s)|] \quad \leq \quad E[|X(0) - \tilde{X}(0)|]$$

$$
\begin{aligned}
&+C_1E\left[\left(\int_0^t\int_{U_0}|F_0(X(s-),u)-F_0(\tilde{X}(s-),u)|^2\mu(du)ds\right)^{1/2}\right]\\
&+C_1E\left[\left(\int_0^t\int_{U_1}|F_1(X(s-),u)-F_1(\tilde{X}(s-),u)|^2\nu_1(du)ds\right)^{1/2}\right]\\
&+E\left[\int_0^t\int_{U_2}|F_2(X(s-),u)-F_2(\tilde{X}(s-),u)|\nu_2(du)ds\right]\\
&+E\left[\int_0^t|F_3(X(s-))-F_3(\tilde{X}(s-))|ds\right]\\
\leq\ & E[|X(0)-\tilde{X}(0)|]+D(\sqrt{t}+t)E[\sup_{s\leq t}|X(s)-\tilde{X}(s)|] \qquad (8.4)
\end{aligned}
$$

where C_1 is the constant in Lemma 2.4 and D depends on C_1 and the Lipschitz constants of F_0 - F_3. Uniqueness follows by selecting t so that $D(\sqrt{t}+t)<1$. We will need estimates like this one in Section 9

9 Infinite systems.

Most uniqueness proofs for stochastic differential equations are based on L_2-estimates. Graham (1992) is one exception (see Remark 8.3), and as he notes, L_2-estimates may be completely inadequate in treating jump processes. Example 7.3 illustrates the problem. For that example, $E[(\sigma_i(X(t),z)-\sigma_i(\tilde{X}(t),z))^2]=E[|\sigma_i(X(t),z)-\sigma_i(\tilde{X}(t),z)|]$, since σ_i is an indicator function, and any attempt to estimate $E[(X_i(t)-\tilde{X}_i(t))^2]$ by the usual Gronwall argument will fail. One of the main advantages of the techniques developed in Theorem 7.1 is that a mixture of L_2 and L_1 (or other) estimates can be used to to obtain the fundamental estimate (7.2).

For finite dimensional problems or single equations in a Banach space, the techniques of Theorem 7.1 should work any time either L_2 or L_1 estimates work. (One exception appears to be the results of Yamada and Watanabe (1971) involving non-Lipschitz F.) For infinite systems, however, there are several examples for which L_2 or L_1 methods are effective but for which we have not found an analogue of the approach of Theorem 7.1. Example 7.3 is in fact one. The conditions under which we have applied Theorem 7.1, although they cover most of the standard examples of spin-flip models, are not the usual conditions employed. For example, the conditions in Liggett (1985), Chapter III, require $\sup_i\sum_j a_{ij}<\infty$. Kurtz (1980) considers the general L_p in i, L_q in j analogs of the L_1 in i, and L_∞ in j conditions covered in Example 7.3 and the L_∞ in i, L_1 in j conditions of Liggett.

9.1 Systems driven by Poisson random measures.

We consider a system more general than (7.5)

$$
X_i(t)=X_i(0)+\int_{V\times[0,t]}F_i(X(s-),v)N(dv\times ds) \qquad (9.1)
$$

where N is a Poisson random measures on $V\times[0,\infty)$ with mean measure $\nu\times m$. Note that to represent (7.5) as an equation of the form (9.1), let $V=\mathbb{Z}^d\times[0,\infty)$, $\nu=\gamma\times m$ where γ

is the measure with mass 1 at each point in $\mathbb{Z}^d$, and

$$F_i(x,(k,z)) = \sigma_i(x,z)\delta_{ik}.$$

We assume that $E[|X_i(0)|] < \infty$ and

$$\int_V |F_i(x,v)|\nu(dv) \le b_i$$

which ensures that $E[|X_i(t)|] < E[|X_i(0)|] + b_i t < \infty$. Observe that if X and Y are solutions of (9.1) with $X(0) = Y(0)$, then

$$E[|X_i(t) - Y_i(t)|] \le \int_0^t E[\int_V |F_i(X(s-),v) - F_i(Y(s-),v)|\nu(dv)]ds \, .$$

With this inequality and Example 7.3 in mind, we assume that

$$\int_V |F_i(x,v) - F_i(y,v)|\nu(dv) \le \sum_j a_{ij}|x_j - y_j| \, , \tag{9.2}$$

which implies

$$\begin{aligned} \alpha_i E[|X_i(t) - Y_i(t)|] &\le \int_0^t \sum_j \frac{\alpha_i a_{ij}}{\alpha_j} \alpha_j E[|X_j(s) - Y_j(s)|] \\ &\le \left\| \frac{\alpha_i a_{ij}}{\alpha_j} \right\|_{p,j} \int_0^t \|\alpha_j E[|X_j(s) - Y_j(s)|]\|_{q,j} \, ds \end{aligned} \tag{9.3}$$

where

$$\|f(i,j)\|_{p,j} = \left[\sum_j |f(i,j)|^p \right]^{1/p}$$

and similarly for $\|\cdot\|_{q,j}$. If

$$\|\alpha_j b_j\|_{q,j} < \infty \, , \tag{9.4}$$

then $\|\alpha_j E[|X_j(s) - Y_j(s)|]\|_{q,j} < \infty$, and (9.3) implies

$$\|\alpha_i E[|X_i(t) - Y_i(t)|]\|_{q,i} \le \left\| \left\| \frac{\alpha_i a_{ij}}{\alpha_j} \right\|_{p,j} \right\|_{q,i} \int_0^t \|\alpha_j E[|X_j(s) - Y_j(s)|]\|_{q,j} \, ds \, . \tag{9.5}$$

Consequently, if

$$\left\| \left\| \frac{\alpha_i a_{ij}}{\alpha_j} \right\|_{p,j} \right\|_{q,i} < \infty, \tag{9.6}$$

by Gronwall's inequality,

$$\|\alpha_i E[|X_i(t) - Y_i(t)|]\|_{q,i} = 0$$

and hence $X = Y$.

Note that with the appropriate definition of V, a system of the form

$$X_i(t) = X_i(0) + \sum_{k,l} \int_{[0,\infty)\times[0,t]} F_{ikl}(X(s-), z) N_{kl}(dz \times ds) \tag{9.7}$$

where the N_{kl} are independent Poisson random measures with mean measure $m \times m$, can be written in the form (9.1). Let

$$F_{ikl}(x,z) = (1-x_i)x_k\delta_{li}I_{[0,\lambda(k,l)]}(z) - x_i(1-x_l)\delta_{ki}I_{[0,\lambda(k,l)]}(z),$$

and assume that $X_i(0)$ is 0 or 1 for each i. Then (9.7) gives a simple exclusion model with jump rates $\lambda(k,l)$, that is, a particle at k attempts a jump to l at rate $\lambda(k,l)$; however, if l is occupied, the jump is rejected. Note that

$$\begin{aligned}
&\sum_{k,l} \int |F_{ikl}(x,z) - F_{ikl}(y,z)| dz \\
&\qquad \le \sum_k |(1-x_i)x_k - (1-y_i)y_k|\lambda(k,i) + \sum_l |x_i(1-x_l) - y_i(1-y_l)|\lambda(i,l)
\end{aligned}$$

We see that (9.2) is satisfied for

$$a_{ii} = \sum_{k\neq i} \lambda(k,i) + \sum_{l \neq i} \lambda(i,l) \qquad a_{ij} = \lambda(i,j) + \lambda(j,i), \ i \neq j \ .$$

If

$$\sup_i \sum_{j\neq i} (\lambda(i,j) + \lambda(j,i)) < \infty$$

then (9.6) is satisfied for $p = 1$ and $q = \infty$.

9.2 Uniqueness for general systems.

We now consider a general system of the form

$$X_i(t) = U_i(t) + F_i(X, \cdot -) \cdot Y_i(t) \tag{9.8}$$

where for each i, Y_i is an $(L, \hat{H})^\#$-semimartingale and $F_i : D_{L^\infty}[0,\infty) \to \hat{H}$ and is nonanticipating. Recall that while the product space L^∞ is not a Banach space, it is metrizable with a complete metric, for example,

$$d_{L^\infty}(x,y) = \sum_{k=1}^{\infty} \frac{\|x_k - y_k\|_L \wedge 1}{2^k}.$$

We could allow L and $\hat{H}$ to depend on i, but the notation is bad enough already. With (8.4) in mind, the following theorem employs L_1-estimates. A similar result could be stated using L_2-estimates.

Theorem 9.1 *Suppose that for each $T > 0$, there exists a nonnegative function C_T such that for every cadlag, adapted, $\hat{H}$-valued process Z and each stopping time $\tau \leq T$ and positive number $t \leq 1$*

$$E[\sup_{s\leq t} \|Z_- \cdot Y_i(\tau + s) - Z_- \cdot Y_i(\tau)\|_{\hat{H}}] \leq C_T(t) E[\sup_{s\leq t} \|Z(\tau + s)\|_{\hat{H}}] \tag{9.9}$$

and $\lim_{t\to 0} C_T(t) = 0$ (cf. (8.4)). Assume (cf. (9.2)) that for $x, y \in D_{L^\infty}[0,\infty)$ and all $t \geq 0$, $\|F_i(x,t)\| \leq b_i$ and

$$\|F_i(x,t) - F_i(y,t)\|_{\hat{H}} \leq \sum_j a_{ij} \sup_{s\leq t} \|x_j(s) - y_j(s)\|_L \tag{9.10}$$

and that for some positive sequence $\{\alpha_i\}$ and some p and q satisfying $p^{-1} + q^{-1} = 1$

$$\|\alpha_j b_j\|_{q,j} < \infty\ , \tag{9.11}$$

and

$$\left\| \left\| \frac{\alpha_i a_{ij}}{\alpha_j} \right\|_{p,j} \right\|_{q,i} < \infty. \tag{9.12}$$

Then there exists a unique solution of the system (9.8).

Remark 9.2 *a) Shiga and Shimizu (1980) give a similar result for systems of diffusions.*

b) One advantage of L_1 and L_2 estimates over the probability estimates used in Theorem 7.1 is that a direct, iterative approach to existence is possible.

Proof. To show uniqueness, let X and $\tilde{X}$ be solutions of (9.8) and let $\tau = \inf\{t : X(t) \neq \tilde{X}(t)\}$. Fix $T > 0$. As in (7.3), $X(\tau) = \tilde{X}(\tau)$, so by (9.9) and (9.10),

$$\begin{aligned} &E[\sup_{s\leq t} \|X_i(\tau \wedge T + s) - \tilde{X}_i(\tau \wedge T + s)\|_L] \\ &\quad = E[\sup_{s\leq t} \| \left(F_i(X, \cdot-) - F_i(\tilde{X}, \cdot-)\right) \cdot Y_i(\tau \wedge T + s)\|_L] \\ &\quad \leq C_T(t) \sum_j a_{ij} E[\sup_{s\leq t} \|X_j(\tau \wedge T + s) - \tilde{X}_j(\tau \wedge T + s)\|_L], \end{aligned}$$

and hence, as in (9.5),

$$\begin{aligned} &\left\| \alpha_i E[\sup_{s\leq t} \|X_i(\tau \wedge T + s) - \tilde{X}_i(\tau \wedge T + s)\|_L] \right\|_{q,i} \\ &\quad \leq C_T(t) \left\| \left\| \frac{\alpha_i a_{ij}}{\alpha_j} \right\|_{p,j} \right\|_{q,i} \left\| \alpha_j E[\sup_{s\leq t} \|X_j(\tau \wedge T + s) - \tilde{X}_j(\tau \wedge T + s)\|_L] \right\|_{q,j} \end{aligned} \tag{9.13}$$

and selecting $t > 0$ such that

$$C_T(t) \left\| \left\| \frac{\alpha_i a_{ij}}{\alpha_j} \right\|_{p,j} \right\|_{q,i} < 1, \tag{9.14}$$

we see that $\tau \geq \tau \wedge T + t$ a.s, which, in particular, implies $\tau > T$ a.s. Since T is arbitrary, $\tau = \infty$ a.s. and the uniqueness follows.

If t satisfies (9.14), then existence on the time interval $[0, t]$ follows by iteration using (9.13), that is, let $X^0 \equiv x^0$ and define X^n recursively by

$$X_i^{n+1}(t) = U_i(t) + F_i(X^n, \cdot -) \cdot Y_i(t).$$

Then as in (9.13),

$$\begin{aligned} &\left\| \alpha_i E[\sup_{s \leq t} \|X_i^{n+1}(s) - X_i^n(s)\|_L] \right\|_{q,i} \\ &\quad \leq C_T(t) \left\| \left\| \frac{\alpha_i a_{ij}}{\alpha_j} \right\|_{p,j} \right\|_{q,i} \left\| \alpha_j E[\sup_{s \leq t} \|X_j^n(s) - X_j^{n-1}(s)\|_L] \right\|_{q,j}, \end{aligned} \tag{9.15}$$

and it follows that $\{X^n\}$ is Cauchy. For $s < T$, if existence is known on $[0, s]$, then the solution can be extended to $[0, s+t]$ by the same interation argument. Consequently, existence holds on $[0, T]$, and since T is arbitrary, we have global existence of the solution. □

9.3 Convergence of sequences of systems.

We now consider a sequence of systems of the form

$$X_{n,i}(t) = U_{n,i}(t) + F_{n,i}(X_n, \cdot -) \cdot Y_{n,i}(t) \tag{9.16}$$

and extend the convergence theorems of Section 7 to this settings. (Note that the extension to finite systems is immediate.) We view $X_n = (X_{n,1}, X_{n,2}, \ldots)$ and $U_n = (U_{n,1}, U_{n,2}, \ldots)$ as processes in $D_{L^\infty}[0, \infty)$ and $F_n = (F_{n,1}, F_{n,2}, \ldots)$ as a mapping from $D_{L^\infty}[0, \infty)$ into $D_{\hat{H}^\infty}[0, \infty)$, and by $(U_n, Y_n) \Rightarrow (U, Y)$, we mean that

$$\begin{aligned} &(U_n, Y_{n,1}(\varphi_{1,1}), \ldots, Y_{n,1}(\varphi_{1,m_1}), \ldots, Y_{n,l}(\varphi_{l,1}), \ldots, Y_{n,l}(\varphi_{l,m_l})) \\ &\quad \Rightarrow (U, Y_1(\varphi_{1,1}), \ldots, Y_1(\varphi_{1,m_1}), \ldots, Y_l(\varphi_{l,1}), \ldots, Y_l(\varphi_{l,m_l})) \end{aligned}$$

in $D_{L^\infty \times \mathbb{R}^{m_1 + \cdots + m_l}}[0, \infty)$ for all choices of $\varphi_{i,j} \in H$. Recall that convergence in $D_{L^\infty}[0, \infty)$ is equivalent to convergence of the first k components in $D_{L^k}[0, \infty)$ for each k and that a sequence of processes $\{X^n\}$ in $D_{L^\infty}[0, \infty)$ satisfies the compact containment condition if and only if each component satisfies the compact containment condition. We need the following modification of Condition C2.

C.3a There exist $G_n, G : D_{L^\infty}[0, \infty) \times T_1[0, \infty) \to D_{\hat{H}^\infty}[0, \infty)$ such that $F_n(x) \circ \lambda = G_n(x \circ \lambda, \lambda)$ and $F(x) \circ \lambda = G(x \circ \lambda, \lambda)$, $x \in D_{L^\infty}[0, \infty)$, $\lambda \in T_1[0, \infty)$.

C.3b For each compact $\mathcal{K} \subset D_{L^\infty}[0, \infty) \times T_1[0, \infty)$ and $t > 0$,

$$\sup_{(x,\lambda) \in \mathcal{K}} \sup_{s \leq t} d_{\hat{H}^\infty}(G_n(x, \lambda, s), G(x, \lambda, s)) \to 0.$$

C.3c For $\{(x_n, \lambda_n)\} \in D_{L^\infty}[0,\infty) \times T_1[0,\infty)$, $\sup_{s\le t} d_{L^\infty}(x_n(s), x(s)) \to 0$ and $\sup_{s\le t} |\lambda_n(s) - \lambda(s)| \to 0$ for each $t > 0$ implies $\sup_{s\le t} d_{\hat{H}^\infty}(G(x_n, \lambda_n, s), G(x, \lambda, s)) \to 0$.

Theorem 9.3 *Let $L = \mathbb{R}^k$. Suppose that (U_n, X_n, Y_n) satisfies (9.16), that $(U_n, Y_n) \Rightarrow (U, Y)$, and that $\{F_n\}$ and F satisfy Condition C.3. For each i, assume that $\{Y_{n,i}\}$ is uniformly tight and that* $\sup_n \sup_x \|F_{n,i}(x,\cdot)\|_{H^k} < \infty$. *Then $\{(U_n, X_n, Y_n)\}$ is relatively compact and any limit point satisfies (9.8).*

Proof. As in the proof of Theorem 7.5, it is enough to show the relative compactness of (U_n, X_n, Y_n). By Proposition 6.2, $\{F_{n,i}(X_n, \cdot-) \cdot Y_{n,i}\}$ is uniformly tight. By Lemma 6.9 and Proposition 6.14, there exists a γ_n such that $\{(U_n \circ \gamma_n, F_n(X_n, \cdot-) \cdot Y_n \circ \gamma_n)\}$ is relatively compact, which in turn implies $\{(X_n \circ \gamma_n, U_n \circ \gamma_n, F_n(X_n, \cdot-) \cdot Y_n \circ \gamma_n)\}$ is relatively compact. Setting $\hat{X}_n = X_n \circ \gamma_n$, etc., Condition C.3a implies

$$\hat{X}_n(t) = \hat{U}_n(t) + G_n(\hat{X}_n, \gamma_n, \cdot-) \cdot \hat{Y}_n(t)$$

and the relative compactness of $\{(\hat{X}_n, \hat{U}_n, \hat{Y}_n)\}$ implies the relative compactness of

$$\{(\hat{X}_n, \hat{U}_n, \hat{Y}_n, G_n(\hat{X}_n, \gamma_n, \cdot), \gamma_n)\}\ .$$

Recalling that relative compactness in $D_{L^\infty}[0,\infty)$ is equivalent to relative compactness of the first k components in $D_{L^k}[0,\infty)$ for each k, by Lemma 6.16, $\{(X_n, U_n, Y_n)\}$ is relatively compact, and the theorem follows from the analog of Proposition 7.4. □

Theorem 9.4 *Suppose that (U_n, X_n, Y_n) satisfies (9.16), that $(U_n, Y_n) \Rightarrow (U, Y)$, and that $\{F_n\}$ and F satisfy Condition C.3. For each i, assume that $\{Y_{n,i}\}$ is uniformly tight, that* $\sup_n \sup_x \|F_{n,i}(x,\cdot)\|_{\hat{H}} < \infty$, *and that for each sequence of positive numbers $\{\kappa_j\}$, there exists a compact $K_{i,\{\kappa_j\}} \subset \hat{H}$ such that $\sup_{s\le t} \|x_j(s)\|_L \le \kappa_j$ for all j implies $F_{n,i}(x, s) \in K_{i,\{\kappa_j\}}$ for all $s \le t$ and n. Then $\{(U_n, X_n, Y_n)\}$ is relatively compact and any limit point satisfies (9.8).*

Proof. Again it is sufficient to verify the relative compactness of $\{(U_n, X_n, Y_n)\}$. The uniform tightness of $\{Y_n\}$ and the boundedness of the F_n imply the stochastic boundedness of $\{\sup_{s\le t} \|X_{n,i}(s)\|_L\}$ for each i. The compactness condition on the $F_{n,i}$ then ensures that $\{F_{n,i}(X_n, \cdot)\}$ satisfies the compact containment condition. It follows that the sequence $\{F_{n,i}(X_n, \cdot-) \cdot Y_n\}$ satisfies the compact containment condition by Lemma 6.18 which in turn implies $\{X_{n,i}\}$ satisfies the compact containment condition. Lemma 6.15 and Lemma 6.9 then ensure the existence of γ_n as in the proof of Theorem 9.3, and the remainder of the proof is the same. □

The proofs of the following corollaries are the same as for the corresponding results in Section 7.

Corollary 9.5 *a) Let $L = \mathbb{R}^k$. Suppose that F and G satisfy C.3a and C.3c and that for each i, $\sup_x \|F_i(x,\cdot)\|_{H^k} < \infty$ and Y_i is a standard $H^\#$-semimartingale. Then weak existence holds for (9.8).*

b) For general L, suppose that F and G satisfy C.3a and C.3c and that for each i, $\sup_x \|F_i(x,\cdot)\|_{\hat{H}} < \infty$, F_i satisfies the compactness condition of Theorem 9.4, and Y_i is a standard $(L,\hat{H})^\#$-semimartingale. Then weak existence holds for (9.8).

Corollary 9.6 *Suppose, in addition to the conditions of Corollary 9.5, that strong uniqueness holds for (9.8) for any version of (U,Y) for which each Y_i is a standard $H^\#$ (or $(L,\hat{H})^\#$)-semimartingale. Then any solution of (9.8) is a measurable function of (U,Y), that is, if X satisfies (9.8) and the finite linear combinations of $\{\varphi_k\}$ are dense in H, there exists a measurable mapping*

$$g : D_{L^\infty \times \mathbb{R}^\infty}[0,\infty) \to D_{L^\infty}[0,\infty)$$

such that $X = g(U, Y(\varphi_1,\cdot-), Y(\varphi_2,\cdot),\ldots)$. In particular, there exists a strong solution of (9.8).

Corollary 9.7 *Suppose, in addition to the conditions of Theorem 9.3 or Theorem 9.4, that strong uniqueness holds for (9.8) for any version of (U,Y) for which each Y_i is a standard $H^\#$ (or $(L,\hat{H})^\#$)-semimartingale and that $(U_n,Y_n) \to (U,Y)$ in probability. Then $(U_n,Y_n,X_n) \to (U,Y,X)$ in probability.*

10 McKean-Vlasov limits.

We now consider an infinite system indexed by $i \in \mathbb{Z}$

$$X_i(t) = U_i(t) + F_i(X,Z,\cdot-)\cdot Y_i(t)$$

which we assume to be shift invariant in the sense that $\{(U_i,Y_i)\}$ is a stationary sequence and $F_i(x,z,t) = F_0(x_{\cdot+i},z,t)$. We require that $\{(X_i,U_i,Y_i)\}$ also be stationary (which it will be if uniqueness holds) and that Z be the $\mathcal{P}(L)$-valued process given by

$$Z(t) = \lim_{k\to\infty} \frac{1}{2k+1}\sum_{i=-k}^{k} \delta_{X_i(t)}$$

where convergence is in the weak topology. Note that a.s. existence of the limit follows from the ergodic theorem and that $Z(t,\Gamma) = P\{X_i(t)\in\Gamma|\mathcal{I}\}$ a.s., independently of i, where $\mathcal{I}$ is the σ-algebra of invariant sets for the stationary sequence $\{(X_i,U_i,Y_i)\}$. Since $D_L[0,\infty)$ is a complete, separable metric space, there will exist a regular conditional distribution Q on $\mathcal{B}(D_L[0,\infty))$ for X_i given $\mathcal{I}$, and we can take $Z(t,\Gamma)$ to be $Q\{X_i(t)\in\Gamma\}$. Since for any probability measure on $D_L[0,\infty)$, the one dimensional distributions will be cadlag as $\mathcal{P}(L)$-valued functions, Z will be a cadlag process.

Typically, the driving processes in models of this type are assumed to be independent. (See, for example, Graham (1992), Kallianpur and Xiong (1994), and Méléard (1995) for further discussion and references.) In Section 11.1, we will see how a model of Kotelenez (1995) can be interpreted as a system of the present type in which the Y_i are identical.

Let $a_i, b > 0$ and $\sum_i a_i < \infty$. We assume

$$\|F_0(x,z,t) - F_0(\tilde{x},\tilde{z},t)\|_{\hat{H}} \leq \sum_i a_i \sup_{s\leq t} \|x_i(s) - \tilde{x}_i(s)\| + b \sup_{s\leq t} \rho_W(z(s),\tilde{z}(s))$$

where ρ_W is the Wasserstein metric on $\mathcal{P}(L)$, that is, letting $B_1 = \{f \in \bar{C}(L) : |f(x)| \leq 1, |f(x) - f(y)| \leq \|x-y\|_L, x, y \in L\}$,

$$\rho_W(\mu,\nu) = \sup_{f\in B_1} |\int f d\mu - \int f d\nu|.$$

In addition, we assume that Y_i satisfies (9.9) of Theorem 9.1 and that $\sup_{x,z,t} \|F_0(x,z,t)\|_{\hat{H}} < \infty$. Then uniqueness holds as in the proof of Theorem 9.1 with $p = 1$ and $q = \infty$. Observe that

$$\rho_W(Z(t),\tilde{Z}(t)) \leq \lim_{k\to\infty} \frac{1}{k} \sum_{i=1}^{k} \|X_i(t) - \tilde{X}_i(t)\|_L .$$

To see that existence holds, let $Z^0(t) \equiv \mu$ for some fixed $\mu \in \mathcal{P}(L)$, and define X^{n+1} recursively as the solution of

$$X_i^{n+1}(t) = U_i(t) + F_i(X^{n+1}, Z^n, \cdot-) \cdot Y_i(t)$$

where Z^n is defined by

$$Z^n(t,\Gamma) = \lim_{k\to\infty} \frac{1}{k} \sum_{i=1}^{k} \delta_{X_i^n(t)}(\Gamma) \ a.s.$$

Note that X^{n+1} exists and is (strongly) unique, given Z^n, as in Theorem 9.1. By the strong uniqueness, X^n is a functional of (U, Y), and we can take $Z^n(t,\Gamma) = P\{X_i^n(t) \in \Gamma|\mathcal{I}\}$, where $\mathcal{I}$ is the invariant σ-algebra for (U, Y). In particular, as noted above, we can take Z^n to be a cadlag process in $\mathcal{P}(L)$. As in the proof of Theorem 9.1, we have

$$\begin{aligned} &E[\sup_{s\leq t} \|X_i^{n+1}(s) - X_i^n(s)\|_L] \qquad (10.1) \\ &\leq C_T(t) \left(\sum_j a_j E[\sup_{s\leq t} \|X_{i+j}^{n+1}(s) - X_{i+j}^n(s)\|_L] + bE[\sup_{s\leq t} \rho_W(Z^n(s), Z^{n-1}(s))] \right) . \end{aligned}$$

Since by stationarity, $E[\sup_{s\leq t} \|X_i^{n+1}(s) - X_i^n(s)\|_L] = E[\sup_{s\leq t} \|X_j^{n+1}(s) - X_j^n(s)\|_L]$, we have

$$(1 - C_T(t) \sum_j a_j) E[\sup_{s\leq t} \|X_i^{n+1}(s) - X_i^n(s)\|_L] \leq C_T(t) b E[\sup_{s\leq t} \|X_i^n(s) - X_i^{n-1}(s)\|_L]$$

and we have convergence on $[0, t]$ provided

$$\frac{C_T(t)b}{1 - C_T(t)\sum_j a_j} < 1.$$

Existence for all t then follows.

We now obtain X as the limit of finite particle systems. In particular, we consider the system

$$X_i^n(t) = U_i(t) + F_i(X^n, Z^n, \cdot -) \cdot Y_i(t)$$

for $-n \leq i \leq n$ with "wrap-around" boundary conditions, that is, for $j \notin \{-n, \dots, n\}$, we set $X_j^n = X_{j(n)}^n$ where $-n \leq j(n) \leq n$ and $|j - j(n)|$ is a multiple of $2n+1$. Z^n is the empirical measure

$$Z^n(t) = \frac{1}{2n+1} \sum_{i=-n}^{n} \delta_{X_i^n(t)}.$$

Let $Z^{(n)}$ be defined by

$$Z^{(n)}(t) = \frac{1}{2n+1} \sum_{i=-n}^{n} \delta_{X_i(t)}$$

and note that $\lim_{n\to\infty} \rho_W(Z^{(n)}(t), Z(t)) = 0$ a.s. We want this convergence to be uniform. In general, for sequences $\{q_n\} \subset \mathcal{P}(D_L[0,\infty))$, $q_n \to q$ does not imply that the marginals converge, let alone uniformly, unless the marginals of the limit q are continuous. Consequently, we assume that Z is continuous which implies that

$$\lim_{n\to\infty} \sup_{s\leq t} \rho_W(Z^{(n)}(s), Z(s)) = 0 \ a.s. \tag{10.2}$$

for each $t > 0$. We suspect that in the current situation (10.2) will hold without the continuity assumption; however, the assumption is in fact rather mild and will hold, for example, if for all i, j, (U_i, Y_i) and (U_j, Y_j) have a.s. no common discontinuities.

As in (10.1), we have

$$\begin{aligned}
&E[\sup_{s\leq t} \|X_i(s) - X_i^n(s)\|_L] \\
&\quad \leq C_T(t)\left(\sum_j a_{j-i} E[\sup_{s\leq t} \|X_j(s) - X_j^n(s)\|_L] + bE[\sup_{s\leq t} \rho_W(Z(s), Z^n(s))]\right) \\
&\quad \leq C_T(t)\Bigg(\sum_j a_{j-i} E[\sup_{s\leq t} \|X_{j(n)}(s) - X_j^n(s)\|_L] + bE[\sup_{s\leq t} \rho_W(Z^{(n)}(s), Z^n(s))] \\
&\qquad\qquad + \sum_{|j|>n} a_{j-i} E[\sup_{s\leq t} \|X_j(s) - X_{j(n)}(s)\|_L] + bE[\sup_{s\leq t} \rho_W(Z(s), Z^{(n)}(s))]\Bigg)
\end{aligned}$$

Noting that

$$E[\sup_{s\leq t} \rho_W(Z^{(n)}(s), Z^n(s))] \leq \frac{1}{2n+1} \sum_{j=-n}^{n} E[\sup_{s\leq t} \|X_j(s) - X_j^n(s)\|_L],$$

if we define the matrix D^n by

$$D_{ij}^n = \sum_{j'(n)=j} a_{j'-i} + \frac{b}{2n+1},$$

and set $R_i^n(t) = E[\sup_{s\le t} \|X_i(s) - X_i^n(s)\|_L]$ and

$$S_i^n(t) = \sum_{|j|>n} a_{j-i} E[\sup_{s\le t} \|X_j(s) - X_{j(n)}(s)\|_L] + bE[\sup_{s\le t} \rho_W(Z(s), Z^{(n)}(s))]$$

for $-n \le i \le n$, then for t sufficiently small, $C_T(t)(\sum_j a_j + b) < 1$, and $(I - C_T(t)D^n)^{-1} = \sum_{k=0}^{\infty} (C_T(t)D^n)^k$ is a matrix with positive entries which implies

$$R^n(t) \le (I - C_T(t)D^n)^{-1} S^n(t) \,,$$

and the convergence follows.

11 Stochastic partial differential equations.

Let Y be an $(L, \hat{H})^{\#}$-semimartingale. We consider equations which, formally, can be written as

$$X(t) = X(0) + \int_0^t AX(s)ds + B(X(\cdot-)) \cdot Y(t) \tag{11.1}$$

where A is a linear, in general unbounded, operator on L. Typically, L is a Banach space of functions on a Euclidean space $\mathbb{R}^d$ and A is a differential operator (for example, $A = \Delta$), but for our purposes we will let L be arbitrary and assume that A is the generator of a strongly continuous semigroup on L. In most interesting cases (see Walsh (1986) and Da Prato and Zabczyk (1992) for systematic developments of the theory), (11.1) cannot hold rigorously in that no solution will exist taking values in the domain $\mathcal{D}(A)$ of the unbounded operator A. Consequently, (11.1) must be interpreted in a weak sense. For example, letting A^* denote the adjoint of A defined on some subspace $\mathcal{D}(A^*)$ of L^*, we can write

$$\langle h, X(t)\rangle = \langle h, X(0)\rangle + \int_0^t \langle A^*h, X(s)\rangle ds + \langle h, B(X(\cdot-)) \cdot Y(t)\rangle \tag{11.2}$$

for $h \in \mathcal{D}(A^*)$. For the right side of (11.2) to be defined, we need only require that X takes values in L and that $B(X)$ takes values in $\hat{H}$. More generally, it would be sufficient for $h \in \mathcal{D}(A^*)$ to be extended to the range of B in such a way that $\langle h, B(X(t))\rangle \in H$. Then (11.2) becomes

$$\langle h, X(t)\rangle = \langle h, X(0)\rangle + \int_0^t \langle A^*h, X(s)\rangle ds + \langle h, B(X(\cdot-))\rangle \cdot Y(t) \,. \tag{11.3}$$

A process X satisfying (11.2) is usually refered to as a *weak* solution of (11.1); however, note that this functional analytic notion of "weak" should not be confused with the "solution in distribution" notion of "weak" considered in Section 7.

A third notion of solution arises when A is the generator of an operator semigroup $\{S(t\}$ on L that can be extended to a semigroup on $\hat{H}$, starting with the obvious definition $S(t)(\sum a_{ij} f_i \varphi_j) = \sum a_{ij} \varphi_j S(t) f_i$. An infinite-dimensional, stochastic version of the variation of parameters formula leads to

$$X(t) = S(t)X(0) + S(t - \cdot)B(X(\cdot-)) \cdot Y(t) \,. \tag{11.4}$$

To clarify the meaning of the last term, if Y is given by a semimartingale random measure, then the last term can be written as

$$\int_{U\times[0,t]} S(t-s)B(X(s-),u)Y(du\times ds).$$

See Da Prato and Zabczyk (1992), Chapter 6, for a discussion of the relationships among these notions of solution in one setting.

A variety of weak convergence results for stochastic partial differential equations exist in the literature. See, for example, Bhatt and Mandrekar (1995), Blount (1991, 1995), Brezeźniak, Capiński, and Flandoli (1988), Fichtner and Manthey (1993), Jetschke (1991), Kallianpur and Pérez-Abreu (1989), Gyöngy (1988, 1989), and Twardowska (1993). We have not yet had the opportunity to explore the application of the convergence results developed in previous sections to stochastic partial differential equations. In this section, we collect a few of the results that may be useful in making that application.

11.1 Estimates for stochastic convolutions.

Let V be an adapted, cadlag, $\hat{H}$-valued process. If for each $t>0$, the mapping $s\in[0,t]\to S(t-s)V(s)\in\hat{H}$ is cadlag, then

$$Z(t)=S(t-\cdot)V(\cdot-)\cdot Y(t)$$

is well-defined for each t; however, the properties of Z as a process are less clear. In particular, for the solution of (11.4) to be cadlag, the stochastic integral must be cadlag. We begin with a discussion of a result of Kotelenez (1982).

Lemma 11.1 *Let Z be a cadlag, L-valued, $\{\mathcal{F}_t\}$-adapted process, and let $\psi: L\to[0,\infty)$. Suppose that*

$$E[\psi(Z(t+h))|\mathcal{F}_t]\leq E[\Lambda(t+h)-\Lambda(t)|\mathcal{F}_t]+e^{\beta(t+h)-\beta(t)}\psi(Z(t))$$

for $t,h\geq 0$, where Λ is a cadlag, nondecreasing $\{\mathcal{F}_t\}$-adapted process with $\alpha(t)=E[\Lambda(t)]<\infty$ and β is a nondecreasing function. Then for any stopping time τ and $\delta>0$,

$$\begin{aligned}P\{\sup_{s\leq\tau\wedge t}\psi(Z(s))>\delta\} &\leq \frac{1}{\delta}\left(E[\psi(Z(0))]e^{\beta(t)-\beta(0)}+E\left[\int_0^{t\wedge\tau}e^{\beta(t)-\beta(s)}d\Lambda(s)\right]\right)\\ &\leq \frac{1}{\delta}\left(E[\psi(Z(0))]e^{\beta(t)-\beta(0)}+\int_0^t e^{\beta(t)-\beta(s)}d\alpha(s)\right).\end{aligned}\qquad(11.5)$$

If, in addition, Λ is predictable and $\psi(Z(0)=0$, then for $0<p<1$,

$$E[(\sup_{s\leq\tau\wedge t}\psi(Z(s))^p]\leq\left(\frac{e^{\beta(t)}}{1-p}+1\right)E[(\Lambda(\tau\wedge t))^p].\qquad(11.6)$$

Proof. Observe that

$$
\begin{aligned}
&E[e^{-\beta(t+h)}\psi(Z(t+h) - \int_0^{t+h} e^{-\beta(s)}d\Lambda(s)|\mathcal{F}_t] \\
&\leq e^{-\beta(t+h)}E[\Lambda(t+h)-\Lambda(t)|\mathcal{F}_t] + e^{-\beta(t)}\psi(Z(t) - E[\int_0^{t+h} e^{-\beta(s)}d\Lambda(s)|\mathcal{F}_t] \\
&\leq e^{-\beta(t)}\psi(Z(t)) - \int_0^t e^{-\beta(s)}d\Lambda(s) \\
&\equiv U(t)
\end{aligned}
$$

so that U is a supermartingale. Consequently,

$$
\begin{aligned}
P\{\sup_{s\leq\tau\wedge t}\psi(Z(s)) \geq \delta\} &\leq P\{\sup_{s\leq\tau\wedge t} U(s) \geq e^{-\beta(t)}\delta\} \\
&\leq \frac{1}{\delta}e^{\beta(t)}(E[U(0)] + E[U^-(t)]) \\
&\leq \frac{1}{\delta}\left(e^{\beta(t)-\beta(0)}E[\psi(Z(0))] + E\left[\int_0^{t\wedge\tau} e^{\beta(t)-\beta(s)}d\Lambda(s)\right]\right)
\end{aligned}
$$

where $U^- = (-U)\vee 0$ and the last inequality follows from Doob's inequality, and (11.5) follows.

To prove (11.6), we follow an argument from Ichikawa (1986). Let $\sigma_x = \inf\{t : \Lambda(t) \geq x\}$. Since Λ is predictable, σ_x is predictable and can be approximated by an increasing sequence of stopping times σ_x^n satisfying $\sigma_x^n < \sigma_x$. Then

$$
\begin{aligned}
P\{\sup_{s\leq\tau\wedge t}\psi(Z(s)) > x\} &\leq \frac{1}{x}E\left[\int_0^{t\wedge\tau\wedge\sigma_x^n} e^{\beta(t)-\beta(s)}d\Lambda(s)\right] + P\{\tau\wedge t > \sigma_x^n\} \\
&\leq \frac{1}{x}e^{\beta(t)}E[x\wedge\Lambda(t\wedge\tau)] + P\{\tau > \sigma_x^n\}.
\end{aligned}
$$

Noting that $\lim_{n\to\infty} P\{\tau > \sigma_x^n\} = P\{\tau\wedge t \geq \sigma_x\} = P\{\Lambda(t\wedge\tau) \geq x\}$, we have

$$
P\{\sup_{s\leq\tau\wedge t}\psi(Z(s)) > x\} \leq \frac{1}{x}e^{\beta(t)}E[x\wedge\Lambda(t\wedge\tau)] + P\{\Lambda(t\wedge\tau)\geq x\},
$$

and multiplying both sides by px^{p-1} and integrating gives (11.6). □

Lemma 11.2 *Let $L = L_2(\nu)$, and let M be a worthy martingale measure with dominating measure K satisfying (2.9). Let $\hat{K}$ be given by (2.10), and suppose that M determines a standard $(L,\hat{H})^\#$-martingale. For $0 \leq s \leq t$, let $\tilde{\Gamma}(t,s)$ and $\Gamma(t,s)$ be bounded linear operators on L satisfying $\|\tilde{\Gamma}(t,s)\| \leq C(t)$ and $\|\Gamma(t,s)\| \leq e^{\beta(t)-\beta(s)}$, where C and β are nondecreasing. Suppose that for $V \in S^0_{\hat{H}}$, Z, defined by*

$$
Z(t,\cdot) = \int_{U\times[0,t]} \tilde{\Gamma}(t,s)V(s-,\cdot,u)M(du\times ds), \tag{11.7}
$$

is cadlag and

$$
\int_{U\times[0,t]} \tilde{\Gamma}(t+h,s)V(s-,\cdot,u)M(du\times ds) = \Gamma(t+h,t)\int_{U\times[0,t]} \tilde{\Gamma}(t,s)V(s-,\cdot,u)M(du\times ds).
$$

For $0 \leq t \leq T$, define

$$\Lambda(t) = C(T) \int_{U\times[0,t]} \|V(s-,\cdot,u\|_L^2 \hat{K}(du \times ds).$$

Then for each stopping time τ,

$$\begin{aligned} & P\{\sup_{s\leq\tau\wedge T} \|Z(s)\|_L^2 > \delta\} && (11.8)\\ & \leq \frac{1}{\delta} E\left[C(T) \int_{U\times[0,T\wedge\tau]} e^{\beta(T)-\beta(s)} \|V(s-,\cdot,u\|_L^2 \hat{K}(du \times ds)\right] \end{aligned}$$

and assuming $\hat{K}$ is predictable, for $0 < p < 2$,

$$E[\sup_{s\leq\tau\wedge T} \|Z(s)\|_L^p] \leq \left(\frac{2e^{\beta(t)}}{2-p} + 1\right) E[(C(T) \int_{U\times[0,\tau\wedge T]} \|V(s-,\cdot,u\|_L^2 \hat{K}(du \times ds))^{p/2}]. \quad (11.9)$$

Remark 11.3 *A number of closely related results exist in the literature, formulated in terms of Hilbert space-valued martingales. The inequality (11.8) is essentially Theorem 1 of Kotelenez (1982). The restriction to $0 < p < 2$ in (11.9) is not necessary. Kotelenez (1984) estimates the second moment of of $\sup_{0\leq s<\tau\wedge T} \|Z(s)\|_L$ in the setting of Hilbert space-valued martingales. Ichikawa (1986) gives moment estimates similar to those of Kotelenez (1984) for $0 < p \leq 2$, and under the assumption that the driving martingale is continuous, for $p > 2$. In the particular case of finite dimensional and Hilbert space-valued Wiener processes, Tubaro (1984), da Prato and Zabczyk (1992b), da Prato and Zabczyk (1992a, Theorem 7.3), and Zabczyk (1993) give estimates for $p > 2$. Walsh (1986, Theorem 7.13) also considers more general convolution integrals of the form $\int_{U\times[0,t]} g(t,s,u)M(du \times ds)$ under Hölder continuity assumptions on g.*

Proof. Observe that

$$\begin{aligned} E[Z(t+h,\cdot)^2|\mathcal{F}_t] &= E\left[\left(\int_{U\times(t,t+h]} \tilde{\Gamma}(t+h,s)V(s-,\cdot,u)M(du\times ds)\right)^2 \middle| \mathcal{F}_t\right] \\ &\quad + \left(\int_{U\times[0,t]} \tilde{\Gamma}(t+h,s)V(s-,\cdot,u)M(du\times ds)\right)^2 \\ &\leq E\left[\int_{U\times(t,t+h]} \left|\tilde{\Gamma}(t+h,s)V(s-,\cdot,u)\right|^2 \hat{K}(du\times ds) \middle| \mathcal{F}_t\right] \\ &\quad + \left(\Gamma(t+h,t)\int_{U\times[0,t]} \tilde{\Gamma}(t,s)V(s-,\cdot,u)M(du\times ds)\right)^2 \quad (11.10) \end{aligned}$$

Then for $0 \leq t < t+h \leq T$, integrating (11.10) with respect to ν, we have

$$E[\|Z(t+h)\|_L^2|\mathcal{F}_t] \leq E[\Lambda(t+h) - \Lambda(t)|\mathcal{F}_t] + e^{\beta(t+h)-\beta(t)}\|Z(t)\|_L^2$$

which, by Lemma 11.1 with $\psi(z) = \|z\|_L^2$, gives the result. □

For a filtration $\{\mathcal{F}_t\}$, let $\mathcal{T}$ denote the collection of $\{\mathcal{F}_t\}$-stopping times, let $\mathcal{S}^0_{\hat{H}}$ be given in Definition 5.2, let $\mathcal{D}_{\hat{H}}$ be the collection of cadlag, $\hat{H}$-valued, $\{\mathcal{F}_t\}$-adapted processes, and let $\mathcal{D}_L$ be the collection of cadlag, L-valued, $\{\mathcal{F}_t\}$-adapted processes. In the following development, the primary examples of interest are mappings of the form

$$G(X,t) = T(t-\cdot)X(\cdot)\cdot Y(t),$$

or in the case of semimartingale random measures,

$$G(X,t) = \int_{U\times[0,t]} T(t-s)X(s)Y(du\times ds).$$

Proposition 11.4 *Suppose that $G : \mathcal{S}^0_{\hat{H}} \to \mathcal{D}_L$ has the following properties:*

- *$G(X)$ is linear, that is $G(aX+bY) = aG(X)+bG(Y)$, $X,Y\in\mathcal{S}^0_{\hat{H}}$, $a,b\in\mathbb{R}$.*
- *For each $t>0$,*

  $$\mathcal{H}^0_t = \{\|G(X,\tau\wedge t)\|_L : X\in\mathcal{S}^0_{\hat{H}}, \sup_{s\le t}\|X(s)\|_{\hat{H}}\le 1, \tau\in\mathcal{T}\}$$

  is stochastically bounded.

Then G extends to a mapping $G:\mathcal{D}_{\hat{H}}\to\mathcal{D}_L$.

Remark 11.5 *Note that if Y is a standard, $(L,\hat{H})^\#$-semimartingale, then G defined by $G(X,t) = X_-\cdot Y(t)$ satisfies the conditions of the proposition. Recall that $X_-\cdot Y(\tau\wedge t) = X^\tau_-\cdot Y(t)$, where $X^\tau = I_{[0,\tau)}X$.*

Proof. The proof is essentially the same as in the definition of the stochastic integral. As before, for each $t,\delta>0$, there exists a $K(t,\delta)$ such that

$$P\{\|G(X,\tau\wedge t)\|_L \ge K(t,\delta)\}\le\delta$$

for all $X\in\mathcal{S}^0_{\hat{H}}$ with $\|X\|_{\hat{H}}\le 1$. For $X\in\mathcal{D}_{\hat{H}}$, let X^ϵ be given by (5.4). Then

$$P\{\sup_{s\le t}\|G(X^{\epsilon_1},s)-G(X^{\epsilon_2},s)\|_L\ge(\epsilon_1+\epsilon_2)K(t,\delta)\}\le\delta$$

and it follows that $\lim_{\epsilon\to 0}G(X^\epsilon)$ converges to a cadlag process which we define to be $G(X)$. □

Theorem 11.6 *Let G satisfy the conditions of Proposition 11.4. Suppose that for each stopping time τ, bounded by a constant, and $t,\delta>0$, there exists $K_\tau(t,\delta)$ with $\lim_{t\to 0}K_\tau(t,\delta)=0$ such that for all stopping times σ satisfying $\tau\le\sigma\le\tau+t$*

$$P\{\|G(X,\sigma)-G(X,\tau)\|_L\ge K(t,\delta)\}\le\delta$$

for all $X\in\mathcal{S}^0_{\hat{H}}$ with $\|X\|_{\hat{H}}\le 1$. Let $F: D_L[0,\infty)\to D_{\hat{H}}[0,\infty)$ satisfy $\sup_{s\le t}\|F(x,s)-F(y,s)\|_{\hat{H}}\le M\sup_{s\le t}\|x(s)-y(s)\|_L$. Then for $U\in\mathcal{D}_L$, there exists at most one solution of

$$X(t) = U(t)+G(F(X),t).$$

Proof. The proof is the same as for Theorem 7.1. □

11.2 Eigenvector expansions.

Suppose that L is spanned by a sequence $\{f_k\}$ of eigenvectors for A, that is, $Af_k = -\lambda_k f_k$, and that $h_k \in L^*$ satisfies $\langle h_k, f_l\rangle = \delta_{kl}$. Then $A_k^* h_k = -\lambda_k h_k$ and (11.11) implies

$$\langle h_k, X(t)\rangle = \langle h_k, X(0)\rangle - \lambda_k \int_0^t \langle h_k, X(s)\rangle ds + \langle h_k, B(X(\cdot -))\rangle \cdot Y(t)\,. \tag{11.11}$$

At least heuristically, if we define $V_k(t) = \langle h_k, X(t)\rangle$, then

$$X(t) = \sum_{k=1}^{\infty} V_k(t) f_k. \tag{11.12}$$

To study the convergence of (11.12), we need to be able to estimate V_k, and the following lemma of Blount (1991, 1995) gives a useful approach.

Lemma 11.7 *Let M be a continous, $\mathbb{R}$-valued martingale and Γ a constant with $[M]_t = \int_0^t U(s)ds$ and $|U(s)| \le \Gamma$, let C be an adapted process and C_0 a constant satisfying*

$$\sup_{s\le t} |C(s)| \le C_0,$$

and let $\lambda > 0$. Suppose V satisfies

$$V(t) = V(0) - \lambda \int_0^t V(s)ds + \int_0^t C(s)ds + M(t)\,.$$

Then for each $a > 0$,

$$P\{\sup_{s\le t} |V(s)| \ge a + |V(0)| + \frac{C_0}{\lambda}\} \le \frac{\lambda t}{\exp\{\frac{\lambda a^2}{4\Gamma}\} - 1}$$

Proof. The proof is based on a comparison with the Ornstein-Uhlenbeck process satisfying $d\tilde{V} = -\lambda \tilde{V} dt + \sqrt{\Gamma} dW$ (W a standard Brownian motion). See Lemma 3.19 of Blount (1991) and Lemma 1.1 of Blount (1995). □

11.3 Particle representations.

With the results of Section 10 in mind, consider the system $\{(X_i, A_i)\}$ with $X_i \in \mathbb{R}^d$ and $A_i \in \mathbb{R}$ satisfying

$$\begin{aligned} X_i(t) = X_i(0) \;&+\; \int_0^t \sigma(X_i(s), V(s))dW_i(s) + \int_0^t c(X_i(s), V(s))ds \\ &+\; \int_{U\times[0,t]} \alpha(X_i(s), V(s), u)W(du \times ds) \end{aligned}$$

and

$$\begin{aligned} A_i(t) = A_i(0) \;&+\; \int_0^t A_i(s)\gamma^T(X_i(s), V(s))dW_i(s) + \int_0^t A_i(s)d(X_i(s), V(s))ds \\ &+\; \int_{U\times[0,t]} A_i(s)\beta(X_i(s), V(s), u)W(du \times ds) \end{aligned}$$

where the W_i are independent, standard $\mathbb{R}^d$-valued Brownian motions and W is Gaussian white noise with $E[W(A,t)W(B,t)] = \mu(A \cap B)t$. Assume that $\{(A_i(0), X_i(0))\}$ are iid and independent of W_i and W. V is the signed measure-valued process obtained by setting

$$\langle \varphi, V(t)\rangle = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} A_i(t)\varphi(X_i(t)).$$

For simplicity, assume that σ, c, γ, α, β, and d are bounded. Letting Z be the $\mathcal{P}(\mathbb{R}^{d+1})$-valued process

$$Z(t) = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \delta_{(X_i(t),A_i(t))},$$

we have

$$\langle \varphi, V(t)\rangle = \int_{\mathbb{R}^{d+1}} a\varphi(x) Z(t, dx \times da),$$

and the uniqueness result of Section 10 can be translated into a uniqueness result here. Applying Itô's formula to $A_i(t)\varphi(X_i(t))$ we obtain

$$\begin{aligned} A_i(t)\varphi(X_i(t)) &= A_i(0)\varphi(X_i(0)) + \int_0^t A_i(s)\varphi(X_i(s))\gamma(X_i(s), V(s))dW_i(s) \\ &\quad + \int_0^t A_i(s)\varphi(X_i(s))d(X_i(s), V(s))ds \\ &\quad + \int_{U\times[0,t]} A_i(s)\varphi(X_i(s))\beta(X_i(s), V(s), u)W(du \times ds) \\ &\quad + \int_0^t A_i(s)L(V(s))\varphi(X_i(s))ds \\ &\quad + \int_0^t A_i(s)\nabla\varphi(X_i(s)) \cdot \sigma(X_i(s), V(s))dW_i(s) \\ &\quad + \int_{U\times[0,t]} A_i(s)\nabla\varphi(X_i(s))\alpha(X_i(s), V(s))W(du \times ds) \end{aligned}$$

where $L(v)\varphi(x) = \frac{1}{2}\sum a_{ij}(x,v)\partial_{x_i}\partial_{x_j}\varphi(x) + \sum b_i(x,v)\partial_{x_i}\varphi(x)$ with

$$b(x,v) = c(x,v) + \sigma(x,v)\gamma(x,v) + \int_U \beta(x,v,u)\alpha(x,v,u)\mu(du)$$

and

$$a(x,v) = \sigma(x,v)\sigma^T(x,v) + \int_U \alpha(x,v,u)\alpha^T(x,v,u)\mu(du).$$

Observing that the terms involving W_i and $\tilde{W}_i$ will average to zero, we have

$$\begin{aligned} \langle \varphi, V(t)\rangle &= \langle \varphi, V(0)\rangle + \int_0^t (\langle d(\cdot, V(s))\varphi, V(s)\rangle + \langle L(V(s))\varphi, V(s)\rangle)ds \\ &\quad + \int_{U\times[0,t]} \langle \beta(\cdot, V(s), u)\varphi + \alpha(\cdot, V(s), u)\cdot\nabla\varphi, V(s)\rangle W(du \times ds), \end{aligned}$$

and it follows that V is a weak solution of the stochastic partial differential equation

$$\begin{aligned} dv(x,t) &= (L^*(V(t))v(x,t) + d(x,V(t))v(x,t))\,dt \qquad (11.13) \\ &\quad + (\beta(x,V(t),u)v(x,t) - \mathrm{div}_x[\alpha(x,V(t),u)v(x,t)])\,W(du \times dt) \end{aligned}$$

where $V(t)$ is the signed measure given by $V(A,t) = \int_A v(x,t)dx$.

Remark 11.8 *The immediate motivation for the material in this subsection is Kotelenez (1995) who considers a model, formulated in a somewhat different way, which is essentially the case $\gamma = d = \beta = \sigma = 0$. In particular, the weights A_i are constant. Perkins (1995) uses time-varying weights for models based on historical Brownian motion (rather than an infinite system of stochastic differential equations) to obtain weak solutions for stochastic partial differential equations related to superprocesses. Donnelly and Kurtz (1995, 1996) obtain particle representations similar to those given here for a large class of measure-valued processes including many Fleming-Viot and Dawson-Watanabe processes. We suspect that methods used in Donnely and Kurtz (1996) to prove uniqueness for the martingale problem corresponding to the measure-valued process based on uniqueness for the particle model may extend to the present setting and give uniqueness of the weak solution of (11.13).*

12 Examples.

No attempt has been made at ultimate generality in the following examples. They are intended to illustrate the variety of models that can be represented as solutions of stochastic differential equations of the type considered here. In regard to technical points, note that for most of the examples there are many possible choices for the space H.

12.1 Averaging.

The study of the behavior of stochastic models with a rapidly varying component dates back at least to Khas'minskii (1966a,b). Characterization of the processes as solutions of martingale problems has proved to be an effective approach to proving limit theorems for these models. (See Kurtz (1992) for a discussion and additional references.) Formulating such a model as a solution of a stochastic differential equation, let W be a standard Brownian motion in $\mathbb{R}^\beta$, let $X_n(0)$ be independent of W, and let ξ be a stochastic process with state space U, independent of W and $X_n(0)$. Set $\xi_n(t) = \xi(nt)$, and for $\sigma : \mathbb{R}^\beta \times U \to \mathbb{M}^{\alpha\beta}$ and $b : \mathbb{R}^\beta \times U \to \mathbb{R}^\alpha$, let X_n satisfy

$$X_n(t) = X_n(0) + \int_0^t \sigma(X_n(s), \xi_n(s))dW(s) + \int_0^t b(X_n(s), \xi_n(s))ds\,.$$

We assume that

$$\frac{1}{t}\int_0^t f(\xi(s))ds \to \int_U f(u)\nu(du) \tag{12.1}$$

in probability for each $f \in \bar{C}(U)$. Define a sequence of orthogonal martingale random measures on $U \times \{1, \ldots, \beta\}$ by setting

$$M_n(A \times \{k\}, t) = \int_0^t I_A(\xi_n(s))dW_k(s)\,. \tag{12.2}$$

Observe that

$$\langle M_n(A \times \{k\}, \cdot), M_n(B \times \{l\}, \cdot)\rangle_t = \int_0^t \delta_{kl} I_{A\cap B}(\xi_n(s))ds$$

and

$$M_n(\varphi,t)=\sum_{k=1}^{\beta}\int_0^t \varphi(\xi_n(s),k)dW_k(s)$$

and that for $\varphi_1,\ldots,\varphi_m\in\bar{C}(U\times\{1,\ldots,\beta\})$, the martingale central limit theorem (see, for example, Ethier and Kurtz (1986), Theorem 7.1.4) implies

$$(M_n(\varphi_1,\cdot),\ldots,M_n(\varphi_\beta,\cdot))\Rightarrow(M(\varphi_1,\cdot),\ldots,M(\varphi_\beta,\cdot))$$

where M corresponds to Gaussian white noise with

$$\langle M(A\times\{k\},\cdot),M(B\times\{l\},\cdot)\rangle_t=\delta_{kl}t\nu(A\cap B).$$

(Approximation of martingale random measures by integrals against scalar Brownian motions, as in (12.2) has been studied by Méléard (1992) in the context of relaxed control theory.)

A variety of norms can be used to determine H. We will assume that U is locally compact. Let $\gamma\in C(U)$, $\gamma>0$, and assume that $\{u:\gamma(u)\le c\}$ is compact for each $c>0$. Define

$$\|\varphi\|_H\equiv\sup_{u,k}|\varphi(u,k)/\gamma(u)|$$

and define H to be the completion of $\bar{C}(U\times\{1,\ldots,\beta\})$ under $\|\cdot\|_H$. If

$$\sup_{t\ge0}\frac{1}{t}\int_0^t E\left[\gamma^2(\xi(s))\right]ds<\infty \tag{12.3}$$

then $\{M_n\}$ is uniformly tight, as is the sequence $\{V_n\}$ defined by

$$V_n(A,t)=\int_0^t I_A(\xi_n(s))ds,$$

which, by (12.1) converges to $\nu(du)ds$. Note, for example, that if ξ is stationary and ergodic, then there exists γ such that the above conditions are satisfied.

Theorem 12.1 *Let ξ and X_n be as above, and assume that (12.1) and (12.3) hold. If σ and b are bounded and continuous and $X_n(0)\Rightarrow X(0)$, then $\{X_n\}$ is relatively compact and any limit point satisfies*

$$X(t)=X(0)+\int_0^t\int_U\sigma(X(s),u)M(du\times ds)+\int_0^t\int_U b(X(s),u)\nu(du)ds\,.$$

Remark 12.2 *In order to apply Theorem 7.5 we need the mappings $x\to\sigma(x,\cdot)$ and $x\to b(x,\cdot)$ to be bounded and continous as mappings from $\mathbb{R}^\alpha$ to H. The assumption that σ and b are bounded and continous as mappings from $\mathbb{R}^\alpha\times U$ to $\mathbb{R}$ is a significantly stronger requirement.*

Proof. For $A\subset U$ and $B\subset\{0,\ldots,\beta\}$, define

$$Y_n(A\times B,t)=I_B(0)V_n(A,t)+\sum_{k=1}^{\beta}I_B(k)M_n(A\times\{k\},t)\,.$$

Then $\{Y_n\}$ is uniformly tight for H defined above, and the theorem follows immediately from Theorem 7.5. □

12.2 Diffusion approximations for Markov chains.

Any discrete-time Markov chain with stationary transition probabilities can be written as a recursion

$$X_{k+1} = F(X_k, \xi_{k+1})$$

where the $\{\xi_k\}$ are independent and identically distributed. Consider a sequence of such chains with values in $\mathbb{R}^\alpha$ satisfying

$$X_{k+1}^n = X_k^n + \sigma_n(X_k^n, \xi_{k+1})\frac{1}{\sqrt{n}} + b_n(X_k^n, \zeta_{k+1})\frac{1}{n}$$

where $\{(\xi_k, \zeta_k)\}$ is iid in $U_1 \times U_2$. We again assume that U_1 and U_2 are locally compact. Let μ be the distribution of ξ_k and ν the distribution of ζ_k, and suppose $\int_{U_2} G_n(x, u_2)\mu(du_2) = 0$ for all $x \in \mathbb{R}^\alpha$ and $n = 1, 2, \ldots$. Define $X_n(t) = X_{[nt]}^n$, $M_n(A,t) = \frac{1}{\sqrt{n}}\sum_{k=1}^{[nt]}(I_A(\xi_k) - \mu(A))$, and $V_n(B,t) = \frac{1}{n}\sum_{k=1}^{[nt]} I_B(\zeta_k)$. Note that $V_n(A,t) \to t\nu(A)$ and $M_n(A,t) \Rightarrow M(A,t)$ where M is Gaussian with covariance

$$E[M(A,t)M(B,s)] = t \wedge s(\mu(A \cap B) - \mu(A)\mu(B))$$

(see Example 2.3).

Theorem 12.3 *Let X_n be as above, and assume that* $\lim_n \sup_{(x,u_2)\in K} |b_n(x,u_2) - b(x,u_2)| = 0$ *for each compact $K \subset \mathbb{R}^\alpha \times U_2$ and* $\lim_{n\to\infty} \sup_{(x,u_1)\in K} |\sigma_n(x,u_1) - \sigma(x,u_1)| = 0$ *for each compact set $K \subset \mathbb{R}^\alpha \times U_1$. Suppose that* $\sup_{n,x,(u_1,u_2)}(|b_n(x,u_2)| + |\sigma_n(x,u_1)|) < \infty$, *that σ and b are bounded and continuous, and that $X_n(0) \Rightarrow X(0)$. Then $\{X_n\}$ is relatively compact and any limit point satisfies*

$$X(t) = X(0) + \int_0^t \int_{U_2} \sigma(X(s), u)M(du \times ds) + \int_0^t \int_{U_1} b(X(s), u)\nu(du)ds\ .$$

Proof. Let $U = U_1 \cup U_2$ and let H be the space of functions on U with norm

$$\|h\|_H = \sqrt{\int_{U_1} h^2(u)\mu(du)} + \int_{U_2} |h(u)|\nu(du).$$

Uniform tightness is again easy to check, and the result follows by Theorem 7.5. □

12.3 Feller diffusion approximation for Wright-Fisher model.

The Wright-Fisher model is a discrete generation genetic model for the evolution of a population of fixed size. For simplicity, we only consider the two-type case. Let N denote the size of the population and let X_n^N denote the fraction of the population in the nth generation that is of type I. Given the population in the nth generation, we assume that the population in the $(n+1)$st generation is obtained as follows: For each of the N individuals in generation $n+1$ a "parent" is selected at random (with replacement) from the population in generation n. If the parent is of type I, then with high probability the "offspring" will be of type I, but there is a small probability (which we will write as μ_1/N) of a "mutation" occuring and producing an offspring of type II. Similarly, if the parent is of type II, then the offspring is of type II with probability $(1-\mu_2/N)$ and of type I with probabilty μ_2/N. These "birth-events" are assumed to be independent conditioned on X_n^N. Note that if $X_n^N = x$, then the probabilty that an offspring is of type I is

$$\pi_N(x) = x(1-\mu_1/N) + (1-x)\mu_2/N = x + ((1-x)\mu_2 - x\mu_1)/N.$$

We can construct this model in the following way. Let $\{\xi_k^n, k = 1, \ldots, N, n = 1, 2, \ldots\}$ be iid uniform $[0,1]$ random variables. Then, given X_0^N, we can obtain X_n^N recursively by

$$X_{n+1}^N = \frac{1}{N}\sum_{k=1}^N I_{[0,\pi_N(X_n^N)]}(\xi_k^N) = \frac{1}{N}\sum_{k=1}^N \left(I_{[0,\pi_N(X_n^N)]}(\xi_k^N) - \pi_N(X_n^N)\right) + \pi_N(X_n^N)\,.$$

Define a martingale random measure with $U = [0,1]$ by

$$M_N(A,t) = \frac{1}{N}\sum_{n=1}^{[Nt]}\sum_{k=1}^N (I_A(\xi_k^n) - m(A))$$

and let $A_N(t) = \frac{[Nt]}{N}$. Then setting $X_N(t) = X_{[Nt]}^N$,

$$\begin{aligned} X_N(t) = X_N(0) \quad &+ \int_{U\times[0,t]} I_{[0,\pi_N(X_N(s-))]}(u) M_N(du \times ds) \\ &+ \int_0^t ((1 - X_N(s-))\mu_2 - X_N(s-)\mu_2)\, dA_N(s). \end{aligned}$$

Noting that

$$E[M_N(A,t)M_N(B,s)] = \frac{[N(t\wedge s)]}{N}(m(A\cap B) - m(A)m(B)),$$

it is easy to check that $M_N \Rightarrow M$ where M is the Gaussian martingale random measure in Example 2.3 (with $\mu = m$). We can take $H = L_2(m) \times \mathbb{R}$, and observing that $x \to F_N(x,\cdot) = (I_{[0,\pi_N(x)]}(\cdot), (1-x)\mu_2 - x\mu_1)$ is a continuous mapping from $[0,1] \to H$ and that $F_N(x,\cdot) \to F(x,\cdot) = (I_{[0,x]}, (1-x)\mu_2 - x\mu_1)$ uniformly in x, we can apply Theorem 7.5 to conclude that $\{X_N\}$ is relatively compact and that any limit point satisfies

$$X(t) = X(0) + \int_{[0,1]\times[0,t]} I_{[0,X(s)]}(u) M(du\times ds) + \int_0^t ((1 - X(s))\mu_2 - X(s)\mu_1) ds.$$

12.4 Limit theorems for jump processes.

The following results are essentially due to Kasahara and Yamada (1991). They treat some cases with discontinuous coefficients which we do not cover here. We include an additional parameter u that can be used to model different kinds of jump behavior. For simplicity, we assume the u takes values in a compact metric space U.

Let N_n be a point process on $\mathbb{R} \times U \times [0,\infty)$ such that if $A \subset \mathbb{R} - (-\epsilon,\epsilon)$ for some $\epsilon > 0$, $N_n(A \times U \times [0,t]) < \infty$ a.s. We assume that N_n is adapted to $\{\mathcal{F}_t^n\}$ in the sense that $N_n(A \times [0,\cdot])$, $A \in \mathcal{B}(\mathbb{R}) \times \mathcal{B}(U)$, is adapted, and we let Λ_n denote an adapted positive random measure such that $\tilde{N}_n(A \times [0,t]) = N_n(A \times [0,t]) - \Lambda_n(A \times [0,t])$ defines a σ-finite, orthogonal martingale random measure on $\mathbb{R} \times U \times [0,\infty)$ and $\Lambda_n(A \times [0,\cdot])$ is continuous for $A \subset (\mathbb{R} - (-\epsilon,\epsilon)) \times U$. Note that the orthogonality implies $N_n(A \times \{t\}) \leq 1$ for all A and t. We assume that

$$\int_{\mathbb{R}\times U} f(z,u)\Lambda_n(dz \times du \times [0,t]) \to t \int_{\mathbb{R}\times U} f(z,u)\nu(dz \times du) \tag{12.4}$$

in probability for all bounded continuous f that vanish for $|z| \leq \epsilon$ for some $\epsilon > 0$, and that

$$\int_{\mathbb{R}\times U} 1 \wedge z^2 \nu(dz \times du) < \infty.$$

Let N denote the Poisson point process on $\mathbb{R}\times U\times[0,\infty)$ with mean measure $\nu(dz\times du)dt$. The convergence in (12.4) implies that

$$\int_{\mathbb{R}\times U\times[0,\cdot]} f(z,u,s)N_n(dz \times du \times ds) \Rightarrow \int_{\mathbb{R}\times U\times[0,\cdot]} f(z,u,s)N(dz \times du \times ds)\ ,$$

for all bounded continous f on $\mathbb{R} \times U \times [0,\infty)$ that vanish for $|z| \leq \epsilon$ for some $\epsilon > 0$. (See Theorem 2.6).)

Let V_n be a positive random measure adapted to $\{\mathcal{F}_t^n\}$, and assume that there exists a finite measure μ such that

$$\int_{\mathbb{R}\times U\times[0,t]} f(z,u,s)V_n(dz \times du \times ds) \to \int_0^t \int_{\mathbb{R}\times U} f(z,u,s)\mu(dz \times du)ds$$

in probability for all bounded continuous f.

Let c satisfy $\nu(\{c\} \times U) = \nu(\{-c\} \times U) = 0$, and suppose X_n satisfies

$$\begin{aligned} X_n(t) = X_n(0) \quad &+ \int_{[-c,c]\times U\times[0,t]} \sigma_n(s,X_n(s),z,u)z\tilde{N}_n(dz \times du \times ds) \\ &+ \int_{(\mathbb{R}-[-c,c])\times U\times[0,t]} b_n^1(s,X_n(s),z,u)N_n(dz \times du \times ds) \\ &+ \int_{\mathbb{R}\times U\times[0,t]} b_n^2(s,X_n(s),z,u)V_n(dz \times du \times ds) \end{aligned}$$

Theorem 12.4 *Let X_n, N_n, and V_n be as above. Suppose that there exists a constant C such that $|\sigma_n| + |b_n^1| + |b_n^2| \leq C$ and that there are bounded and continous σ, b^1, and b^2, such that for each compact $K \subset [0,\infty) \times \mathbb{R}^\alpha \times \mathbb{R} \times U$*

$$\begin{aligned} \lim_{n\to\infty} \sup_{(s,x,u)\in K} (|\sigma(s,x,z,u) - \sigma_n(s,x,z,u)| \quad &+ \quad |b^1(s,x,z,u) - b_n^1(s,x,z,,u)| \\ &+ \quad |b^2(s,x,z,u) - b_n^2(s,x,z,u)|) = 0. \end{aligned}$$

Suppose that

$$\sup_n E[\int_{\mathbb{R}\times[0,t]} (z \wedge c)^2 \Lambda_n(dz \times du \times ds)] < \infty \tag{12.5}$$

and that there exist a positive measure ρ and positive $\epsilon_n \to 0$ such that $\delta_n \geq \epsilon_n$ and $\delta_n \to 0$ implies

$$\int_{(-\delta_n,\delta_n)\times U\times[0,t]} z^2 g(u)\Lambda_n(dz\times du\times ds) \to t\int_U g(u)\rho(du) \tag{12.6}$$

in probability for all $g \in \bar{C}(U)$. If $X_n(0) \to X(0)$, then $X_n \Rightarrow X$ satisfying

$$\begin{aligned} X(t) = X(0) \quad &+ \int_{U\times[0,t]} \sigma(s,X(s),0,u)W(du\times ds) \\ &+ \int_{[-c,c]\times U\times[0,t]} \sigma(s,X(s),z,u)z\tilde{N}(du\times ds) \\ &+ \int_{(\mathbb{R}-[-c,c])\times[0,t]} b^1(s,X(s),z,u)N(dz\times du\times ds) \\ &+ \int_0^t\int_{\mathbb{R}} b^2(s,X(s),z,u)\nu(dz\times du)ds \end{aligned}$$

where W is a Gaussian random measure on $U\times[0,\infty)$ satisfying

$$E[W(A\times[0,t])W(B\times[0,s])] = (t\wedge s)\rho(A\cap B).$$

Proof. It is enough to verify convergence for each bounded time interval $[0,T]$. Let N_n^c be N_n restricted to $(\mathbb{R}-[c,c])\times U\times[0,\infty)$, and $M_n(A\times B\times[0,t]) = \int_{A\cap[-c,c]\times B\times[0,t]} z\tilde{N}_n(dz\times du\times ds)$. The convergence of $\{N_n^c\}$ and $\{V_n\}$ to finite (random) measures implies, through Prohorov's theorem, a stochastic version of tightness which in turn implies the existence of $\gamma:\mathbb{R}\to[1,\infty)$ with $\lim_{|z|\to\infty}\gamma(z)=\infty$ such that $\{\int\gamma(z)N_n^c(dz\times U\times[0,T])\}$ and $\{\int\gamma(z)V_n(dz\times U\times[0,T])\}$ are stochastically bounded. Let $\|h\|_H = \sup_{z,u}|h(u,z)/\gamma(z)|$. The uniform tightness of $\{V_n\}$ and of $\{N_n^c\}$ as $H^\#$-semimartingales is immediate and that of M_n follows from (12.5). The convergence of M_n to M given by

$$M(A\times B\times[0,t]) = \delta_{\{0\}}(A)W(B\times[0,t]) + \int_{A\cap[-c,c]\times B\times[0,t]} z\tilde{N}(dz\times du\times ds)$$

follows from (12.4) and (12.6) and Theorem 2.7. □

12.5 An Euler scheme.

Consider

$$X(t) = X(0) + \int_0^t\int_U \sigma(X(s),u)W(du\times ds) + \int_0^t b(X(s))ds \tag{12.7}$$

where for definiteness we take $U=[0,1]$ and W to be Gaussian white noise determined by Lebesgue measure m on $U\times[0,\infty)$. The solution of this equation will be a diffusion with generator

$$Lf(x) = \frac{1}{2}\sum a_{ij}(x)\frac{\partial^2}{\partial x_i\partial x_j}f(x) + \sum b_i(x)\frac{\partial}{\partial x_i}b(x)$$

where $a(x) = \int_U \sigma(x,u)\sigma^T(x,u)du$. Of course, this diffusion could be rewritten as the solution of an Itô equation in its usual form and numerical schemes applied to approximate the solution of that equation; however, our interest here is in developing methods for approximating stochastic equations driven by martingale measures, and we use this simple case to explore the possibilities.

A simple simulation scheme for (12.7) might involve discretization in both time and in U to give an iteration of the form

$$\hat{X}(t_{i+1}) = \hat{X}(t_i) + \sum \sigma(\hat{X}(t_i), v_i) W((u_i, u_{i+1}] \times (t_i, t_{i+1}]) + b(\hat{X}(t_i))(t_{i+1} - t_i) \tag{12.8}$$

where the sum is over the partition $0 = u_0 < u_1 < \cdots < u_m = 1$ and $u_i \leq v_i \leq u_{i+1}$. This iteration gives the simplest Euler-type scheme for (12.7). Consistency for the scheme follows easily from Theorem 7.5. We are interested in analyzing the error of the scheme in a manner similar to that used in Kurtz and Protter (1991b) to study the Euler scheme for Itô equations. In particular, $\hat{X}$ defined in (12.8) can be extended to a solution of

$$\hat{X}(t) = X(0) + \int_0^t \int_U \sigma(\hat{X} \circ \eta(s), \gamma(u)) W(du \times ds) + \int_0^t b(X \circ \eta(s)) ds$$

where $\eta(s) = t_i$ for $s \in [t_i, t_{i+1})$ and $\gamma(u) = v_i$ for $u \in (u_i, u_{i+1}]$. Then

$$\begin{aligned} X(t) - \hat{X}(t) \;=\; & \int_0^t \int_U \left(\sigma(X(s), u) - \sigma(\hat{X}(s), u)\right) W(du \times ds) \\ & + \int_0^t \left(b(X(s)) - b(\hat{X}(s))\right) ds \\ & + \int_0^t \int_U \left(\sigma(\hat{X}(s), u) - \sigma(\hat{X} \circ \eta(s), u)\right) W(du \times ds) \\ & + \int_0^t \int_U \left(\sigma(\hat{X} \circ \eta(s), u) - \sigma(\hat{X} \circ \eta(s), \gamma(u))\right) W(du \times ds) \\ & + \int_0^t \left(b(\hat{X}(s)) - b(\hat{X} \circ \eta(s))\right) ds \, . \end{aligned}$$

We assume that σ and b are bounded and have two bounded continuous derivatives. Observing that

$$\hat{X}(t) - \hat{X} \circ \eta(t) = \int_U \sigma(\hat{X} \circ \eta(t), \gamma(u)) W(du \times (\eta(t), t]) + b(\hat{X} \circ \eta(t))(t - \eta(t)),$$

fix $\alpha > 0$ and define

$$\begin{aligned} U(t) &= \alpha(X(t) - \hat{X}(t)) \\ Y(A \times [0,t]) &= \int_0^t \alpha W(A \times (\eta(s), s]) ds \\ Z(A \times B \times [0,t]) &= \int_0^t \alpha W(A \times (\eta(s), s]) W(B \times ds) \\ V(A \times [0,t]) &= \int_0^t I_A(u) \alpha(u - \gamma(u)) W(du \times ds) \\ R(A \times [0,t]) &= \int_0^t \alpha(s - \eta(s)) W(A \times ds) \, . \end{aligned}$$

Assuming $d = 1$ to simplify notation,

$$\begin{aligned} U(t) &= \int_0^t \int_U \left(\frac{\sigma(X(s),u) - \sigma(\hat{X}(s),u)}{X(s) - \hat{X}(s)} \right) U(s) W(du \times ds) \\ &\quad + \int_0^t \left(\frac{b(X(s)) - b(\hat{X}(s))}{X(s) - \hat{X}(s)} \right) U(s) ds \\ &\quad + \int_0^t \int_U \sigma_x(\hat{X} \circ \eta(s), v) \sigma(\hat{X} \circ \eta(s), \gamma(u)) Z(du \times dv \times ds) \\ &\quad + \int_0^t \int_U \sigma_x(\hat{X} \circ \eta(s), v) b(\hat{X} \circ \eta(s)) R(dv \times ds) \\ &\quad + \int_0^t \int_U \left(\frac{\sigma(\hat{X} \circ \eta(s), u) - \sigma(\hat{X} \circ \eta(s), \gamma(u))}{u - \gamma(u)} \right) V(du \times ds) \\ &\quad + \int_0^t b_x(\hat{X} \circ \eta(s)) \sigma(\hat{X} \circ \eta(s), u) Y(du \times ds) + Err \end{aligned}$$

where Err will be negligible under our asymptotic assumptions on α and η.

Note that Z, V, and R are martingale random measures with

$$\langle Z(C \times \cdot), Z(D \times \cdot) \rangle_t = \int_0^t \int_U \alpha^2 W(C_v \times (\eta(s), s]) W(D_v \times (\eta(s), s]) dv ds,$$

where $C_v = \{u : (u, v) \in C\}$,

$$\langle V(A \times \cdot), V(B \times \cdot) \rangle_t = t \int_{A \cap B} \alpha^2 (u - \gamma(u))^2 du$$

and

$$\langle R(A \times \cdot), R(B \times \cdot) \rangle_t = m(A \cap B) \int_0^t \alpha^2 (s - \eta(s))^2 ds.$$

Let Z_n, V_n, R_n, U_n be defined by replacing η by $\eta_n(s) = \frac{[ns]}{n}$, α by $\sqrt{n}$, and $\gamma_n(u) = \frac{\beta k}{\sqrt{n}}$ if $\frac{\beta(k-\frac{1}{2})}{\sqrt{n}} < u \le \frac{\beta(k+\frac{1}{2})}{\sqrt{n}}$ for $k = 0, 1, \ldots$. Then as $n \to \infty$ $Z_n \Rightarrow \tilde{Z}$ with

$$\langle \tilde{Z}(C, \cdot), \tilde{Z}(D, \cdot) \rangle_t = \frac{1}{2} m_2(C \cap D) t$$

$V_n \Rightarrow \tilde{V}$ with

$$\langle \tilde{V}(A, \cdot), \tilde{V}(B, \cdot) \rangle_t = \frac{\beta^2}{12} m(A \cap B) t,$$

and $R_n \Rightarrow 0$.

We should have $U_n \Rightarrow \tilde{U}$ satisfying

$$\begin{aligned} \tilde{U}(t) &= \int_0^t \int_U \sigma_x(X(s), u) \tilde{U}(s) W(du \times ds) + \int_0^t b'(X(s)) \tilde{U}(s) ds \\ &\quad + \int_0^t \int_U \sigma_x(X(s), v) \sigma(X(s), u) \tilde{Z}(du \times dv \times ds) \\ &\quad + \int_0^t \int_U \sigma_u(X(s), u) \tilde{V}(du \times ds); \end{aligned}$$

but the conclusion does not follow from Theorem 7.5 since the sequence $\{Z_n\}$ is not uniformly tight. In particular, Z_n is not worthy.

The desired convergence does, however, hold. Note that the integrand for Z_n is adapted to the filtration $\{\mathcal{F}_t^n\}$ with $\mathcal{F}_t^n = \mathcal{F}_{\eta_n(t)} \subset \mathcal{F}_t$. This observation leads us to define the notion of a "conditionally worthy" martingale random measure. Let M be an $\{\mathcal{F}_t\}$-adapted, martingale random measure, and let $\mathcal{G}_t \subset \mathcal{F}_t$. Then M is $\{\mathcal{G}_t\}$*-conditionally worthy* if there exists a *dominating* random measure K on $U \times U \times [0, \infty)$ such that

$$|E[\langle M(A), M(B)\rangle_{t+r} - \langle M(A), M(B)\rangle_t|\mathcal{G}_t]| \leq E[K(A \times B \times (t, t+r])|\mathcal{G}_t].$$

In the present setting, we have

$$E[\langle Z_n(C), Z_n(D)\rangle_{t+r} - \langle Z_n(C), Z_n(D)\rangle_t|\mathcal{F}_t^n] = \int_t^{t+r}\int_U nm(C_v \cap D_v)(s - \eta_n(s))dvds$$

so Z_n is "conditionally orthogonal" uniformly in n. The convergence theorems extend to this setting and the convergence of U_n to $\tilde{U}$ follows.

13 References.

Araujo, Aloisio and Giné, Evarist (1980). *The Central Limit Theorem for Real and Banach Valued Random Variables.* Wiley, New York.

Bhatt, Abhay G. and Mandrekar, V. (1995). On weak solution of stochastic PDE's. (preprint)

Bichteler, K. and Jacod, J. (1983). Random measures and stochastic integration. *Lect. Notes in Control and Inform. Sci.* 49, Springer, New York, Berlin, 1-18.

Blount, Douglas (1991). Comparison of stochastic and deterministic models of a linear chemical reaction with diffusion. *Ann. Probab.* 19, 1440-1462.

Blount, Douglas (1995). A simple inequality with applications to SPDE's. (preprint)

Brzeźniak, Z., Capiński, M. and Flandoli, F. (1988). A convergence result for stochastic partial differential equations. *Stochastics.* 24, 423-445.

Brown, Timothy C. (1978). A martingale approach to the Poisson convergence of simple point processes. *Ann. Probab.* 6, 615-628.

Cho, Nahnsook (1994). *Weak convergence of stochastic integrals and stochastic differential equations driven by martingale measure and its applications.* PhD Dissertation, University of Wisconsin - Madison.

Cho, Nahnsook (1995). Weak convergence of stochastic integrals driven by martingale measure. *Stochastic Process. Appl.* (to appear).

Chow, Pao Liu and Jiang, Jing-Lin (1994). Stochastic partial differential equations in Hölder spaces. *Probab. Theory Related Fields* 99, 1-27.

Çinlar, E., and Jacod, J. (1981). Representation of semimartingale Markov processes in terms of Wiener processes and Poisson random measures. *Seminar on Stochastic Processes 1981.* E. Çinlar, K. L. Chung, R. K. Getoor, eds. Birkhäuser, Boston, 159-242.

Da Prato, Giuseppe and Zabczyk, Jerzy (1992a). *Stochastic Equations in Infinite Dimensions.* Cambridge University Press, Cambridge.

Da Prato, Giuseppe and Zabczyk, Jerzy (1992b). A note on stochastic convolution. *Stoch. Anal. Appl.* 10, 143-153.

Dellacherie, C., Maisonneuve, B. and Meyer, P. A. (1992). *Probabilites et Potentiel, Chapitres XVII a XXIV: Processus de Markov (fin); Complements de Calcul' Stochastique.* Hermann, Paris.

Donnelly, Peter E. and Kurtz, Thomas G. (1996a). A countable representation of the Fleming-Viot measure-valued diffusion. *Ann Probab.* (to appear)

Donnelly, Peter and Kurtz, Thomas G. (1996b). Particle representations for measure-valued population models. (in preparation)

El Karoui, N. and Méléard, S. (1990). Martingale measures and stochastic calculus. *Probab. Th. Rel. Fields* 84, 83-101.

Engelbert, H. J. (1991). On the theorem of T. Yamada and S. Watanabe. *Stochastics* 35, 205-216.

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

Feller, William (1971). *An Introduction to Probability Theory and Its Applications II,* 2nd ed. Wiley, New York.

Fellah, D. and Pardoux, E. (1992). Une formule d'Itô dans des espaces de Banach, et application. *Stochastic analysis and related topics (Silivri, 1990)*, 197-209, *Progr. Probab.* 31, Birkhäuser, Boston.

Fichtner, Karl-Heinz and Manthey, Ralf (1993). Weak approximation of stochastic equations. *Stochastics Stochastics Rep.* 43, 139-160.

Graham, Carl (1992). McKean-Vlasov Ito-Skorohod equations, and nonlinear diffusions with discrete jump sets. *Stochastic Process. Appl.* 40, 69-82.

Gyöngy, I. (1988, 1989). On the approximation of stochastic partial differential equations. I, II. *Stochastics* 25, 59-85, 26, 129-164.

Hadjiev, Dimitar I. (1985). Some remarks on Gaussian solutions and explicit filtering formulae. *Stochastic differential systems (Marseille-Luminy, 1984), Lecture Notes in Control and Information Sci. 69.* Springer, Berlin-New York, 207–216.

Hardy, G., Kallianpur, G., Ramasubramanian, S., and Xiong, J. (1994). The existence and uniqueness of solutions of nuclear space-valued equations driven by Poisson random measures. *Stochastics and Stochastics Reports.*

Ichikawa, Akira (1986). Some inequalities for martingales and stochastic convolutions. *Stoch. Anal. Appl.* 4, 329-339.

Ikeda, Nobuyuki and Watanabe, Shinzo (1981). *Stochastic Differential Equations and Diffusion Processes.* North Holland, Amsterdam.

Itô, Kiyosi (1951). On stochastic differential equations. *Mem. Amer. Math. Soc.* 4.

Jakubowski, Adam (1995). Continuity of the Ito stochastic integral in Hilbert spaces. (preprint)

Jakubowski, Adam (1995). A non-Skorohod topology on the Skorohod space.

Jakubowski, A. Mémin, J. and Pages, G. (1989). Convergence in loi des suites d'intégrales stochastique sur l'espace D^1 de Skorohod. *Probab. Theory Related Fields* 81, 111-137.

Jetschke, Gottfried (1991). Lattice approximation of a nonlinear stochastic partial differential equation with white noise. *Random Partial Differential Equations. Oberwolfach (1989).* Birkhäuser Verlag, Basel. 107-126.

Kallianpur, G. and Pérez-Abreu, V. (1989). Weak convergence of solutions of stochastic evolution equations on nuclear spaces. *Stochastic Partial Differential Equations and Applications II. Lect. Notes Math.* 1390, 119-131.

Kallianpur, G. and Xiong, J. (1994). Stochastic models of environmental pollution. *J. Appl. Probab.*

Kallianpur, G. and Xiong, J. (1994). Asymptotic behavior of a system of interacting nuclear space valued stochastic differential equations driven by Poisson random measures. *Appl. Math. Optim.* 30, 175-201.

Kallianpur, G. and Xiong, J. (1993). Diffusion approximations of nuclear space-valued stochastic differential equations driven by Poisson random measures. Center for Stochastic Processes Technical Report 399.

Kallianpur, G. and Xiong, J. (1992). A nuclear space-valued stochastic differential equation driven by Poisson random measures. *Stochastic PDE's and their Applications. Lect. Notes in Control and Information Sci.* 176. Springer, Berlin-New York, 135-143.

Kasahara, Yuji and Yamada, Keigo (1991). Stability theorem for stochastic differential equations with jumps. *Stochastic Process. Appl.* 38, 13-32.

Katzenberger, G. S. (1991). Solutions of a stochastic differential equation forced onto a manifold by a large drift. *Ann. Probab.* 19, 1587-1628.

Khas'minskii, R. Z. (1966a). On stochastic processes defined by differential equations with a small parameter. *Theory Probab. Appl.* 11, 211-228.

Khas'minskii, R. Z. (1966b). A limit theorem for the solutions of differential equations with random right-hand sides. *Theory Probab. Appl.* 11, 390-406.

Kotelenez, Peter (1982). A submartingale type inequality with applications to stochastic evolution equations. *Stochastics.* 8, 139-151.

Kotelenez, Peter (1984). A stopped Doob inequality for stochastic convolution integrals and stochastic evolution equations. *Stoch. Anal. Appl.* 2, 245-265.

Kotelenez, Peter (1995). A class of quasilinear stochastic partial differential equations of McKean-Vlasov type with mass conservation. *Probab. Theory Relat. Fields* 102, 159-188.

Kurtz, Thomas G. (1980). Representations of Markov processes as multiparameter time changes. *Ann. Probab.* 8, 682-715.

Kurtz, Thomas G. (1991). Random time changes and convergence in distribution under the Meyer-Zheng conditions. *Ann. Probab.* 19, 1010-1034.

Kurtz, Thomas G. (1992). Averaging for martingale problems and stochastic approximation. *Applied Stochastic Analysis. Proceedings of the US-French Workshop. Lect. Notes. Control. Inf. Sci.* 177, 186-209.

Kurtz, Thomas G. and Marchetti, Federico (1995). Averaging stochastically perturbed Hamiltonian systems. *Proceedings of Symposia in Pure Mathematics,* 57, 93-114.

Kurtz, Thomas G. and Protter, Philip (1991a). Weak limit theorems for stochastic integrals and stochastic differential quations. *Ann. Probab.* 19, 1035-1070.

Kurtz, Thomas G. and Protter, Philip (1991b). Wong-Zakai corrections, random evolutions, and simulation schemes for sde's. *Stochastic Analysis: Liber Amicorum for Moshe Zakai.* Academic Press, San Diego. 331-346.

Lebedev, V. A. (1990). Stochastic integration with respect to semimartingale random measures. *Theory Probab. Appl.* 34, 725-727.

Lebedev, V. A. (1990). Stochastic integrals with respect to semimartingale measures and change of the filtration. *Probability theory and mathematical statistics. Proceedings of the Fifth Vilnius Conference, Volume 2* VSP, Utrecht, The Netherlands; 70- 78.

Lenglart, E., Lepingle, D. and Pratelli, M. (1980). Presentation unifiee de certaines inegalites des martingales. *Seminares de probabilités XIV. Lect. Notes in Math.*, Springer, Berlin. 26-61.

Liggett, Thomas M. (1972). Existence theorems for infinite particle systems. *Trans. Amer. Math. Soc.* 165, 471-481.

Liggett, Thomas M. (1985). *Interacting Particle Systems.* Springer-Verlag, New York.

Méléard, Sylvie (1992). Representation and approximation of martingale measures. *Stochastic PDE's and their Applications. Lect. Notes in Control and Information Sci. 176.* Springer, Berlin-New York, 188-199.

Méléard, Sylvie (1995). Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. *CIME School in Probability Lecture Notes.* This volume.

Merzbach, Ely and Zakai, Moshe (1993). Worthy martingales and integrators. *Stat. Prob. Letters.* 16, 391-395.

Metivier, Michel and Pellaumail, J. (1980). *Stochastic Integration.* Academic Press, New York.

Metivier, Michel (1982). *Semimartingales.* de Gruyter, Berlin.

Meyer, P. A. and Zheng, W. A. (1984). Tightness criteria for laws of semimartingales. *Ann. Inst. H. Poincaré Probab. Statist. 20, 353-372.*

Mikulevicius, R. and Rozovskii, B. L. (1994). On stochastic integrals in topological vector spaces. *Stochastic Analysis. Proceedings of Symposia in Pure Mathematics.* 57, 593-602.

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

Perez-Abreu, V. and Kallianpur, G. (1989). Weak convergence of solutions of stochastic evolution equations in nuclear spaces. *Stochastic Partial Differential Equations and Applications. Lect. Notes in Math.* 1390. 133- 139.

Perkins, Edward (1995) On the martingale problem for interactive measure-valued branching diffusions. *Mem. Amer. Math. Soc.* 549.

Shiga, T. and Shimizu, A. (1980). Infinite-dimensional stochastic differential equations and their applications. *J. Mat. Kyoto Univ.* 20, 395-416.

Stricker, Cristophe (1985). Lois de semimartingales et critères de compacité. *Seminares de probabilités XIX. Lect. Notes in Math.* 1123, .

Trofimov, E. I. (1989). Semimartingales: L^2-construction of distributions. *Izv. Vyssh. Uchebn. Zaved. Mat.* 2, 68-78.

Tubaro, Luciano (1984). An estimate of Burkholder type for stochastic processes defined by the stochastic integral.

Twardowska, Krystyna (1993). Approximation theorems of Wong-Zakai type for stochastic differential equations in infinite dimensions. *Dissertationes Mathematicae CCXXV.*

Ustunel, S. (1982). Stochastic integration on nuclear spaces and its applications. *Ann. Inst. Henri Poincaré* 18, 165-200.

Walsh, John (1981). A stochastic model for neural response. *Adv. Appl. Probab.* 13, 231-281.

Walsh, John (1986). An introduction to stochastic partial differential equations. *Lect. Notes in Math.* 1180, 265-439.

Zabczyk, J. (1993). The fractional calculus and stochastic evolution equations. *Barcelona Seminar on Stochastic Analysis (St. Feliu de Goixols, 1991). Progr. Probab.* 32, 222-234. Birkhäuser, Basel.