5-2005

A Stochastic Calculus for Systems with Memory

Feng Yan
Southern Illinois University Carbondale

Salah-Eldin A. Mohammed
Southern Illinois University Carbondale, salah@sfde.math.siu.edu

Published in *Stochastic Analysis and Applications*, 23(3), 613 - 657.

A Stochastic Calculus for Systems with Memory

Feng Yan * Salah Mohammed † ‡

January 23, 2005

Abstract

For a given stochastic process X, its *segment* X_t at time t represents the "slice" of each path of X over a fixed time-interval $[t-r,t]$, where r is the length of the "memory" of the process. Segment processes are important in the study of stochastic systems with memory (stochastic functional differential equations, SFDEs). The main objective of this paper is to study non-linear transforms of segment processes. Towards this end, we construct a stochastic integral with respect to the Brownian segment process. The difficulty in this construction is the fact that the stochastic integrator is infinite dimensional and is not a (semi)martingale. We overcome this difficulty by employing Malliavin (anticipating) calculus techniques. The segment integral is interpreted as a Skorohod integral via a stochastic Fubini theorem. We then prove Itô's formula for the segment of a continuous Skorohod-type process and embed the segment calculus in the theory of anticipating calculus. Applications of the Itô formula include the weak infinitesimal generator for the solution segment of a stochastic system with memory, the associated Feynman-Kac formula and the Black-Scholes PDE for stock dynamics with memory.

1 Introduction

The segment process of a continuous-time stochastic process is an important ingredient in the study and formulation of stochastic differential systems with memory ([13]). Such systems are described by stochastic functional differential equations (SFDEs) of the form

$$X(t) = \begin{cases} x + \int_0^t G(s, X_s, X(s))\, dW(s) + \int_0^t H(s, X_s, X(s))\, ds, & t \geq 0 \\ \eta(t), & -r \leq t < 0, \end{cases} \tag{1.1}$$

where η is an initial path in $V := L^2([-r,0], R^m)$, x an initial vector in R^m and $r \geq 0$ is the length of the system memory. The solution $\{X(t) : -r \leq t \leq a\}$ of the above SFDE is an m-dimensional stochastic process. Its *segment process* $\{X_t : 0 \leq t \leq a\}$ is defined by

$$X_t(s) := X(t+s), \quad t \geq 0, \quad s \in [-r, 0]. \tag{1.2}$$

*The research was supported by Dissertation Research Award of Southern Illinois University at Carbondale.

†Department of Mathematics, Southern Illinois University at Carbondale Carbondale, IL 62901, USA. Email: salah@sfde.math.siu.edu. Research supported in part by NSF grant DMS-0203368.

‡AMS 1991 *subject classifications.* Primary 60H05; secondary 60H10, 60H20.
Key words and phrases. Itô's formula, segment process, anticipating calculus, Malliavin calculus, weak derivatives.

Note that the process $\{X_t\}$ can be viewed as a $V-$valued process. In this context it is important to distinguish between the *finite-dimensional* R^m-valued process $\{X(t)\}$ and the *infinite-dimensional* segment process $\{X_t\}$, both of which appear in the right-hand side of the SFDE (1.1).

To complete the formulation of the SFDE (1.1), suppose (Ω, F, P) is a probability space, $T = [0, a]$, $a > 0$ and $J := [-r, 0]$. Denote by $||\cdot||_V$ and $\langle\cdot,\cdot\rangle_V$ the L^2-norm and inner product (respectively) on the Hilbert space $V = L^2(J, R^m)$. The SFDE (1.1) is driven by d-dimensional standard Brownian motion $\{W(t) = (W^1(t), W^2(t), \cdots, W^d(t)) : t \geq 0\}$ on (Ω, F, P). The drift and diffusion coefficients

$$\begin{cases} H : T \times V \times R^m \to R^m \\ G : T \times V \times R^m \to L(R^d, R^m) \end{cases} \tag{1.3}$$

are Lipschitz on bounded sets and satisfy linear growth conditions. Under these conditions, the SFDE (1.1) has a unique strong solution (c.f. [13], pp. 226 – 228; [14]). Qualitative properties of solutions of stochastic functional differential equations (SFDEs) and stochastic delay differential equations (SDDEs) have been studied by one of the authors in ([13], [14]).

An important aspect of the stochastic calculus of (1.1) is the study of non-linear transforms $f(X_t)$, $t \geq 0$, of the segment X_t where $f : V \to R$ belongs to a large class of twice Fréchet differentiable functionals on V. If one (formally) takes stochastic differentials of the real-valued process $\{f(X_t)\}$ with respect to t, one quickly sees the need for an Itô-type formula which will necessarily involve stochastic segment integrals of the form

$$\int_0^a \langle Y_t, dX_t\rangle_V \tag{1.4}$$

where Y_t is an $L^2(J, R^d)-$valued process. The goal of this paper is to develop the above stochastic segment integral and its calculus. In particular, we will establish Itô's formula for the segment process $\{X_t\}$ where X is a general Skorohod-type process of the form:

$$X(t) = \begin{cases} \eta(0) + \int_0^t u(s)\, dW(s) + \int_0^t v(s)\, ds, & t > 0 \\ \eta(t), & -r \leq t \leq 0. \end{cases} \tag{1.5}$$

with coefficients $u : T\times\Omega \to L(R^d; R^m)$ and $v : T\times\Omega \to R^m$ that may not be adapted to the Brownian filtration $(F_t)_{t\geq 0}$. A major difficulty in the construction of the integral (1.4) is the fact that the infinite-dimensional segment process $\{X_t\}$ is in general *not a semimartingale*. However, we will overcome this difficulty by appealing to Malliavin calculus techniques.

One possible application of the Itô formula is to study the convergence rates of (strong and weak) numerical schemes of stochastic delay equations. Other potential applications include developing new models in mathematical finance based upon SFDEs, and evaluating path dependent financial, energy and weather derivatives.

The paper is organized as follows. First, we define the stochastic integral with respect to the Brownian segment using the Skorohod integral (Sections 2 – 4). Secondly, in Sections 5-7, we study the weak derivatives of V-valued random variables and the L^2 approximation of the segment integral. In Section 8, we prove an Itô formula for processes of the form $\{f(t, X_t, X(t))\}$ where $f : T \times V \times R^m \to R$ is a sufficiently regular non-linear functional. Finally, in Section 9, we study the weak infinitesimal generator of a stochastic functional differential equation, establish the Feynman-Kac formula and derive the Black-Scholes PDE for the pricing of past-dependent financial assets.

We now introduce some notation which will be used throughout this article. Suppose E and F are two Banach spaces. Denote by $C_b(E; F)$ the space of all bounded functions from E to F, which are uniformly continuous on bounded sets, and by $C_b^1(E; F)$ the set of all functions $f \in C_b(E; F)$ which are *Fréchet differentiable*, with Fréchet derivative $f' \in C_b(V; L(E; F))$. Set $W(t) := 0$ if $t < 0$.

2 Difficulty in defining the Brownian segment integral.

Although we can define the stochastic integral with respect to an infinite dimensional martingale, ([4, 6, 20]), we can not apply this definition to the Brownian segment process because it is not a $L^2(J; R^d)-$valued (or $C(J; R^d)$-valued) martingale. As we shall show in Section 5, it is more difficult to define the segment integral for the $C(J; R^d)$-valued case than for the $L^2(J; R^d)$-valued case. This difficulty may be attributed to the fact that the Banach space $C(J; R^d)$ is not smooth.

One of the mild conditions (Condition A) for the existence of McShane's integral ([12], p. 102, [5], p. 23) is the following:

Condition A: There exist constants $K > 0$, and $\delta > 0$ such that if $0 \le s < t \le a$ and $t - s < \delta$, then almost everywhere

$$||E_s(W_t - W_s)||_V \le K(t - s), \tag{2.1}$$

where $V = L^2(J; R^d)$, $E_s(W_t - W_s) := E(W_t - W_s|F_s)$, and $\{F_t\}$ is the filtration of the Brownian segment $\{W_t\}$.

A V-valued martingale always satisfies Condition A. If $\{X(t)\}$ is a sample-continuous d-dimensional stochastic process adapted to a filtration $\{F_t\}$, then the $C(J; R^d)-$valued segment process $\{X_t\}$ is also adapted to $\{F_t\}$ (c.f. [13], p. 30). By continuity of the embedding $I : C(J; R^d) \to L^2(J; R^d)$, it follows that $\{X_t\}$ is also $\{F_t\}$-adapted as an $L^2(J; R^d)-$valued process.

Proposition 2.1 *Let $\{W(t) : t \geq -r\}$ denote d-dimensional Brownian motion. As an $L^2(J; R^d)$-valued process, the Brownian segment process $\{W_t\}$ satisfies Condition A for the existence of McShane's integral if and only if $r = 0$. In particular, $\{W_t\}$ is an an $L^2(J; R^d)$-valued martingale if and only if $r = 0$. Similar assertions hold if W_t is viewed as a $C(J; R^d)$-valued process.*

Proof Let $V = L^2(J; R^d)$, where $J := [-r, 0]$ and $r \geq 0$. Assume that $\{W_t, F_t\}$ satisfies Condition A as a V-valued process. Then there exist constants $K > 0$ and $\delta > 0$ such that if $0 \leq s < t \leq \min(\delta, r)$, then $||E_s(W_t - W_s)||_V \leq K(t - s)$, a.s.. Now for any $h \in [-r, s-t]$, we have

$$\begin{aligned} E_s(W_t - W_s)(h) &= E(W_t - W_s|F_s)(h) \\ &= E(W(t+h) - W(s+h)|F_s) = W(t+h) - W(s+h) \quad a.s. \end{aligned}$$

This holds because $W_t : \Omega \to C(J, R^d)$ is Bochner integrable, and the Bochner integral commutes with evaluations.

Now view $\{W_t\}$ as an $L^2(J, R^d)$-valued process. Using the above identity, we obtain

$$\begin{aligned} |K(t-s)|^2 &\geq ||E_s(W_t - W_s)||_V^2 = \int_{-r}^{0} |E_s(W(t) - W(s)(h)|^2 \, dh \\ &\geq \int_{-r}^{s-t} |W(t+h) - W(s+h)|^2 \, dh \end{aligned}$$

a.s.. Taking expectations in the above inequality, it follows that

$$K^2(t-s)^2 \geq \int_{-r}^{s-t} E|W(t+h) - W(s+h)|^2 \, dh = d(t-s)[(s-t) + r] \tag{2.2}$$

Dividing both sides of the above inequality by $(t-s)$ and letting $s \to t-$ gives $r = 0$. Conversely, suppose $r = 0$. Then $J = \{0\}$ and $L^2(J; R^d), \{W_t\}$ may be identified with $R^d, \{W(t)\}$ respectively. Since $\{W(t)\}$ is a martingale, then so is $\{W_t\}$.

The second assertion of the proposition follows easily from the second.

The last assertion for the $C(J; R^d)$-valued segment $\{W_t\}$ follows from the above argument because the embedding $I : C(J; R^d) \to L^2(J; R^d)$ is continuous. ∎

In view of Proposition 2.1, it is not possible to define the stochastic integral (1.4) in the classical Itô sense when the delay r is positive. However, we may formally rewrite the segment integral (1.4) as follows:

$$\int_0^a \langle Y_t, dW_t \rangle_V = \int_{-r}^0 \int_0^a Y_t(s)\, dW(t+s)\, ds = \int_{-r}^0 \int_s^{a+s} Y_{t-s}(s)\, dW(t)\, ds \tag{2.3}$$

Note that for any fixed $s \in [-r, 0)$, the R^d-valued process $\{Y_{t-s}(s), t \in [0, a+s]\}$ may not be (F_t)-adapted (even if $\{Y_t\}$ is). This suggests that we can overcome the difficulty by using the Skorohod integral. Indeed, in Section 4, we define the above segment integral as a Skorohod integral. To do this, we will need to impose appropriate generalized "smoothness" requirements on the integrand. In the following section, we will give a brief introduction to the basic concepts of anticipating stochastic calculus.

3 A Brief outline of anticipating calculus.

Anticipating stochastic calculus is used in the study of stochastic differential equations with non-adapted initial values ([17], Section 3.3, and [19]). Anticipating stochastic calculus has also been applied by Bell and Mohammed to establish the existence of smooth densities of solutions to stochastic delay differential equations (SDDE's) ([2]). In this paper, we shall use anticipating stochastic calculus to define the segment integral and prove Itô's formula for segments of solutions of stochastic functional differential equations (SFDE's) .

The following outline of anticipating calculus is adopted from Nualart and Pardoux ([16]) and Nualart ([17]). Cf. also ([18]), Malliavin ([11]) and Kuo ([10]).

We denote by D the Malliavin differentiation operator (c.f. [16], Section 2). Let F be a random variable which belongs to the domain of D. Its derivative DF is a stochastic process $\{D_t F, t \in T\}$. The derivative DF may be considered as a random variable taking values in the Hilbert space $H = L^2(T, R^d)$. More generally, the N-th derivative of F, $D^N F := D_{s_1}^{j_1} \ldots D_{s_N}^{j_N} F$ is an $H^{\hat{\otimes}_2 N}$-valued random variable (c.f. Appendix A). For any integer $N \geq 1$ and any real number $p > 1$ we denote by $\mathbb{D}^{N,p}$ the Banach space of all the random variables having all the i-th derivatives belonging to $L^p(\Omega, H^{\hat{\otimes}_2 i})$ $(1 \leq i \leq N)$, with the norm $||\cdot||_{N,p}$ defined by

$$||F||_{N,p} = ||F||_p + ||\ ||D^N F||_{(2)}||_p, \tag{3.1}$$

where $||\cdot||_{(2)}$ denotes the Hilbert-Schmidt norm in $H^{\hat{\otimes}_2 N}$, i.e.,

$$||D^N F||_{(2)}^2 = \sum_{j_1,\ldots,j_N=1}^{d} \int_{T^N} E[(D^N F)_{s_1,\ldots,s_N}^{j_1,\ldots,j_N}]^2\, ds_1 \ldots ds_N. \tag{3.2}$$

(Cf. [16], Section 2.)

For $N = 1$ and $p = 2$, the space $\mathbb{D}^{1,2}$ is a Hilbert space with the scalar product

$$\langle F, G\rangle_{1,2} = E(FG) + E(\langle DF, DG\rangle_H). \tag{3.3}$$

We have the chain rule for differentiation ([17], p.29): Let $\phi : R^m \to R$ be a continuously differentiable function with bounded partial derivatives, and fix $p > 1$. Suppose that $F = (F^1, \ldots, F^m)$ is a random vector whose components belong to the space $\mathbb{D}^{1,p}$. Then $\phi(F) \in \mathbb{D}^{1,p}$, and

$$D(\phi(F)) = \sum_{i=1}^{n} \frac{\partial \phi}{\partial x_i}(F) DF^i.$$

For a more detailed study of the differential operator D, the reader may refer to Nualart and Zakai ([18]), Malliavin ([11]) and Kuo ([10]).

We denote by δ the *divergence operator, Dom* δ the domain of δ, and $\delta(u)$ the *Skorohod stochastic integral* of the d-dimensional process $u \in Dom\,\delta$, i.e.,

$$\delta(u) = \int_T u(t) \cdot dW(t) = \sum_{i=1}^{d} u^i(t)\, dW^i(t). \tag{3.4}$$

The divergence operator transforms square integrable processes into random variables. Actually δ is the adjoint operator of D (c.f. [10], Section 9.4 and Section 13.4, [3] Section 5.5.8). In Section 4, we will define the segment integral (1.4) as an adjoint of a differential operator. Since adjoint operators of densely defined operators are always closed, the operator δ is closed. The Itô integral is a particular case of the Skorohod stochastic integral ([17], section 1.3.2). The set $Dom\,\delta$ is not easy to handle and it is more convenient to deal with processes belonging to some subset of $Dom\,\delta$.

We denote by $\mathbb{L}^{1,2}$ the class of all processes $u \in L^2(T \times \Omega)$ such that $u(t) \in \mathbb{D}^{1,2}$ for almost all t, and there exists a measurable version of the two-parameter process $D_s u(t)$ satisfying $E\int_T \int_T (D_s u(t))^2 ds dt < \infty$ ([16], Definition 3.3, [17], p. 38). It can be shown that $\mathbb{L}^{1,2}$ is a Hilbert space with the norm

$$||u||^2_{1,2} = ||u||^2_{L^2(T\times\Omega)} + ||Du||^2_{L^2(T^2\times\Omega)}, \quad u \in \mathbb{L}^{1,2}. \tag{3.5}$$

Note that $\mathbb{L}^{1,2}$ is isomorphic to $L^2(T; \mathbb{D}^{1,2})$. For every $p > 1$ and any positive integer k we denote by $\mathbb{L}^{k,p}$ the space $L^2(T; \mathbb{D}^{k,p})$. The operator δ is bounded from $\mathbb{L}^{k,p}$ into $\mathbb{D}^{k-1,p}$, for all $k \geq 1$, and $p \geq 1$ ([17], Proposition 1.5.4).

We denote by $\mathbb{L}^{k,p}_d$ the set of d-dimensional processes whose components are in $\mathbb{L}^{k,p}$. We also denote by $\mathbb{D}^{k,p}_{d,loc}$ ($\mathbb{L}^{k,p}_{d,loc}$) the set of random variables that are "locally" in $\mathbb{D}^{k,p}_d$ ($\mathbb{L}^{k,p}_d$)

(c.f. [17], p. 45). If $u \in \mathbb{L}^{k,p}_{d,loc}$, the Skorohod integral $\int_T u(t) \cdot dW(t)$ is also defined by using equation (3.4).

We say that a processes $u \in \mathbb{L}^{1,2}_d$ is in the class $\mathbb{L}^{1,2}_{d,C}$ if there exists a neighborhood A of the diagonal in $[0,a]^2$ such that

(1) there exists a version of Du, so that the mappings $t \hookrightarrow D_s u(t)$ is continuous from $[0,a]$ into $L^2(\Omega; L(R^d; R^d))$, uniformly with respect to s, on $A \cap \{s \leq t\}$;

(2) there exists a version of Du, so that the mappings $t \hookrightarrow D_s u(t)$ is continuous from $[0,a]$ into $L^2(\Omega; L(R^d; R^d))$, uniformly with respect to s, on $A \cap \{s \geq t\}$;

(3) $\sup_{(s,t)\in A} E(|D_s u(t)|^2) < \infty$. ([16], Definition 7.2.)

The space $\mathbb{L}^{1,2}_{d,C,loc}$ is the class of all processes that are locally in $\mathbb{L}^{1,2}_{d,C}$. For any $u \in \mathbb{L}^{1,2}_{d,C}$, the following limits

$$D^{\pm}_t u(t) = \lim_{\epsilon \downarrow 0} \sum_{i=1}^{d} D^i_t u^i(t \pm \epsilon) \tag{3.6}$$

exist in $L^2(\Omega)$ uniformly in t. We set $\bigtriangledown := D^+ + D^-$, i.e., $(\bigtriangledown u)(t) := D^+_t u(t) + D^-_t u(t)$.

Suppose that $u = \{u(t), 0 \leq t \leq a\}$ is a Skorohod integrable process. In general, a process of the form $uI_{(s,t]}$ may not be Skorohod integrable ([17], Exercise 3.2.1). Let us denote by $\mathbb{L}^s_d$ the set of all processes u such that $uI_{[0,t]}$ is Skorohod integrable for each $t \in [0,a]$. Notice that the space $\mathbb{L}^{1,2}_d$ is a subspace of $\mathbb{L}^s_d$. When u belongs to $\mathbb{L}^s_d$, we define its indefinite Skorohod integral by

$$\int_0^t u(s) \cdot dW(s) := \delta(uI_{[0,t]}). \tag{3.7}$$

4 Definition of the stochastic segment integral.

Recall that $J = [-r, 0]$ and $T = [0,a]$. Assume $V = L^2(J; R^m)$ and $H = L^2(T; R^m)$. Let $H \oplus V$ be the direct sum of H and V, $H \otimes_2 V$ be the tensor product of H and V (under the $\epsilon-$ topology, c.f. [21], Section 8), and $H \hat{\otimes}_2 V$ the completion of $H \otimes_2 V$. Then $H \oplus V \cong L^2([-r,a]; R^m)$ and $H \hat{\otimes}_2 V \cong L^2(J \times T; R^m)$. In Appendix A, we state some basic results on tensor product spaces.

We now extend the definition of $\mathbb{D}^{k,p}$ and the operator D to an infinite-dimensional setting (c.f. [17], p. 61). For a Banach space G, we denote by $S(G)$ the class of of all smooth G-valued random variables of the form

$$\Psi = \sum_{i=1}^{n} F_i \eta_i, \quad F_i \in S, \eta_i \in G, \tag{4.1}$$

where S is the class of all smooth random variables (c.f. [16], section 2). We define

$$D^k\Psi := \sum_{i=1}^{n}(D^k F_i)\otimes \eta_i, \quad k \geq 1. \tag{4.2}$$

For $k \in N$, $p \geq 1$, we define $\mathbb{D}^{k,p}(G)$ to be the completion of $S(G)$ with respect to the norm

$$||\Psi||_{k,p,G} = [E||\Psi||_G^p + \sum_{j=1}^{k} E(||D^j\Psi||^p_{L^2(T^j)\hat{\otimes}G})]^{\frac{1}{p}}. \tag{4.3}$$

Note that $\mathbb{D}^{1,2}(H) \cong \mathbb{L}_d^{1,2}$. Suppose $x \in \mathbb{D}^{1,2}(L^2(J;R)) \cong L^2(J;\mathbb{D}^{1,2})$, then for almost all $s \in J$, $x(s) \in \mathbb{D}^{1,2}$ and there is a measurable version of the two-parameter process $Dx := \{D_t x(s) : t \in T, s \in J\}$ such that $Dx \in L^2(\Omega \times T \times J)$.

Now let us define the *segment operator* $\Gamma : H \oplus V \to H\hat{\otimes}_2 V$ by

$$\Gamma\phi(t,s) := \phi(t+s), \quad s \in J, t \in T, \phi \in H \oplus V. \tag{4.4}$$

Denote $\Gamma_t\phi := \phi_t$, for $t \in T$ and $\phi \in H \oplus V$, where $\phi_t(s) := \phi(t+s)$ for $s \in J$. Clearly Γ is bounded linear. One may check that D and Γ commute on processes, i.e., if $\phi \in \mathbb{L}_d^{1,2}$, then

$$D_s\Gamma_t\phi = \Gamma_t D_s\phi. \tag{4.5}$$

We denote by $\Gamma^* : H\hat{\otimes}_2 V \to H \oplus V$ the adjoint of Γ, i.e., for $\eta \in H\hat{\otimes}_2 V$ and $\phi \in H \oplus V$,

$$\langle \Gamma^*\eta, \phi\rangle_{H\oplus V} = \langle \eta, \Gamma\phi\rangle_{H\hat{\otimes}_2 V}. \tag{4.6}$$

From (4.6) we can find the expression for $\Gamma^*\eta$. Actually, simple algebra yields

$$\langle \Gamma^*\eta, \phi\rangle_{H\oplus V} = \int_{-r}^{a} \phi(t)\cdot \int_{-r}^{0} \eta(t-s,s) I_{[t-a,t]}(s)\,ds\,dt. \tag{4.7}$$

Therefore $\Gamma^*\eta$ can be written as:

$$\Gamma^*\eta(t) = \int_{-r}^{0} \eta(t-s,s) I_{[t-a,t]}(s)\,ds = \int_{0}^{a} \eta(s,t-s) I_{[t,t+r]}(s)\,ds. \tag{4.8}$$

Denote by P_H and P_V the projections

$$\begin{cases} P_H : H \oplus V \to H \\ P_V : H \oplus V \to V. \end{cases} \tag{4.9}$$

and define $\Gamma_H^* := P_H \circ \Gamma^*$ and $\Gamma_V^* := P_V \circ \Gamma^*$. It is easy to show that $\Gamma(H \oplus V)$ is a closed subspace of $H\hat{\otimes}_2 V$.

The operator Γ_H^* is a bridge connecting SFDE to the Skorohod integral. It plays an important role in our definition of the segment integral. Γ_H^* and equation (4.5) allow us to study segment processes using anticipating stochastic calculus.

Definition 4.1 Suppose $W = \{W(t)\}$ is a m-dimensional standard Brownian motion. Denote by δ the divergence operator and $Dom\,\delta$ its domain, for a two parameter process $X \in (\Gamma_H^*)^{-1}(Dom\,\delta)$, the *Skorohod segment integral* of X with respect to the Brownian segment $\{W_t\}$ is defined by

$$\int_0^a \langle X_t, dW_t \rangle_V := \delta(\Gamma_H^* X). \tag{4.10}$$

For a V-valued stochastic process $X = \{X_t\}$, if the Stratonovich integral (c.f. [17] Definition 3.1.1) $\delta^s(\Gamma_H^*(X))$ of $\Gamma_H^*(X)$ exists, then we define the *Stratonovich segment integral* of X with respect to the Brownian segment $\{W_t\}$ as

$$\int_0^a \langle X_t, \circ dW_t \rangle_V := \delta^s(\Gamma_H^*(X)). \tag{4.11}$$

It is easy to see that $\mathbb{D}^{1,2}(H\hat{\otimes}_2 V) \subset (\Gamma_H^*)^{-1}(Dom\,\delta)$. The norm $||\cdot||_{1,2,V}$ on $\mathbb{D}^{1,2}(H\hat{\otimes}_2 V)$ is defined by (4.3), i.e.,

$$||X||_{1,2,V} = ||X||_{L^2(\Omega;H\hat{\otimes}_2 V)} + ||DX||_{L^2(\Omega;H\hat{\otimes}_2 H\hat{\otimes}_2 V)} \tag{4.12}$$

If we use D^* to denote δ (as the adjoint of D), then $D^*\Gamma_H^* = D^*P_H\Gamma^* = (\Gamma P_H^* D)^*$, and the operator $\Lambda := \Gamma P_H^* D$ is a continuous differential operator from $\mathbb{D}^{k,p}$ to $\mathbb{D}^{k-1,p}(H\hat{\otimes}_2 V)$.

The next result ([17] Exercise 3.2.8) is a stochastic Fubini Theorem.

Lemma 4.2 *Consider a random field* $\{u_t(x) : t \in [0,a], x \in G\}$, *where* G *is an open set in* R^m, *such that* $u \in L^2(\Omega \times [0,a] \times G)$. *Suppose that for each* $x \in R^m$, $u(x) \in Dom\,\delta$ *and* $E\int_G |\delta(u(x))|^2\,dx) < \infty$. *Then the process* $\{\int_G u_t(x)\,dx : t \in [0,a]\}$ *is Skorohod integrable and*

$$\delta(\int_G u_\cdot(x)\,dx) = \int_G \delta(u_\cdot(x))\,dx. \tag{4.13}$$

5 L^2 approximation.

Recall $T = [0,a]$, $J = [-r,0]$, $C = C(J;R^m)$, and $V = L^2(J;R^m)$. In this section we shall derive some useful results for approximating elements of the Hilbert space V. As we shall see later, these approximation results are crucial for the approximation of segment integrals and the proof of Itô's formula for segment processes.

We denote a vector by a vector header (i.e., $\vec{x}$), a matrix by a bold style letter (i.e., $\mathbf{X}$), and $L(R^m;R^k)$ (or R^{km}) the vector space of all k by m matrices. We skip the vector header if η is a vector-valued function. If $\mathbf{A}$ is a matrix, we denote by a_{ij} its entry at the i-th row and j-th column and by $\vec{a}^i$ its i-th row. We also write $\mathbf{X}$ as $(x_{ij})_{k\times m}$. If η is a matrix-valued function, we denote by η^{ij} its entry at the i-th row and j-th column and by η^i its i-th row. Let us adopt the ideas of "tame" function and "quasitame" function ([13] Definition 4.4.2) to define two series of linear approximations:

Suppose $\Pi_k : -r \le s_1 < \cdots < s_k \le 0$ is a partition of $[-r, 0]$, and $||\Pi_k|| := \max_{2\le i\le k}(s_i - s_{i-1})$. Denote by $\underline{s}^k$ the k-tuple $(s_1, \ldots, s_k)$. If $\eta : [-r, 0] \to R$ is defined everywhere in $[-r, 0]$, we denote by $P_{\underline{s}^k}$ the projection associated with Π_k:

$$P_{\underline{s}^k}(\eta) := (\eta(s_1), \ldots, \eta(s_k)). \tag{5.1}$$

With abusing the notation $\underline{s}^k$, we also define the L^2 projection $Q_{\underline{s}^k}$ for $\eta \in L^2(J; R)$ as

$$Q_{\underline{s}^k}(\eta) := (\frac{1}{s_1 - s_0}\int_{s_0}^{s_1} \eta(s)\,ds, \ldots, \frac{1}{s_k - s_{k-1}}\int_{s_{k-1}}^{s_k} \eta(s)\,ds). \tag{5.2}$$

If $\eta = (\eta_1, \ldots, \eta_m)' \in L^2(J; R^m)$, then we define

$$Q_{\underline{s}^k}(\eta) := (Q_{\underline{s}^k}(\eta_1), \ldots, Q_{\underline{s}^k}(\eta_m))'. \tag{5.3}$$

Suppose $\mathbf{X}$ is an m by k matrix with entries x_{ij}. We define the continuous linear embedding $I_k : R^{mk} \to L^2(J; R^m)$ associated with Π as the step function

$$I_k(\mathbf{X})(s) := (\sum_{i=1}^{k} x_{1i} I_{(s_{i-1},s_i]}(s), \ldots, \sum_{i=1}^{k} x_{mi} I_{(s_{i-1},s_i]}(s))'. \tag{5.4}$$

Thus we can define a linear map $L^2(J; R^m) \to L^2(J; R^m)$ by

$$\eta \hookrightarrow I_k \circ Q_{\underline{s}^k}(\eta).$$

Denote $J_i = I_{(s_{i-1},s_i]}$ and $\Delta_i = s_i - s_{i-1}$. Similarly, if $\eta \in L^2(J^2; R)$, we can define a linear map $L^2(J^2; R) \to L^2(J^2; R)$ by

$$\eta \hookrightarrow I_{k^2} Q_{\underline{s}^{k^2}}(\eta) := \sum_{i,j=1}^{k} \frac{J_i \otimes J_j}{\Delta_i \Delta_j} \int_{s_{i-1}}^{s_i} \int_{s_{j-1}}^{s_j} \eta(u, v)\,dv\,du.$$

If $\eta = (\eta^{ij})_{m\times m} \in L^2(J^2; L(R^m; R^m))$, then we define

$$I_{k^2} \circ Q_{\underline{s}^{k^2}}(\eta) = (I_{k^2} \circ Q_{\underline{s}^{k^2}}(\eta^{ij}))_{m\times m}. \tag{5.5}$$

For each $k \ge 1$, the operator $A_k := I_k \circ Q_{\underline{s}^k}$ has the kernel

$$a_k(s, t) = \sum_{i=1}^{k} \frac{1}{s_i - s_{i-1}} I_{(s_{i-1},s_i]}(s) I_{(s_{i-1},s_i]}(t). \tag{5.6}$$

It is a (symmetric) operator of trace class and satisfying $A_k^2 = A_k$. The following approximation result plays an important role in this paper.

Lemma 5.1 *(1)If $\eta \in V = L^2(J; R^m)$, then $||(I_k \circ Q_{\underline{s}^k})(\eta)||_V \leq ||\eta||_V$, and*

$$\lim_{k\to\infty} ||(I_k \circ Q_{\underline{s}^k})(\eta) - \eta||_V = 0. \tag{5.7}$$

(2)If $\eta \in L^2(J^2; L(R^m; R^m))$, then

$$||(I_{k^2} \circ Q_{\underline{s}^{k^2}})(\eta)||_{L^2(J^2;L(R^m;R^m))} \leq ||\eta||_{L^2(J^2;L(R^m;R^m))}, \text{ and}$$

$$\lim_{k\to\infty} ||(I_{k^2} \circ Q_{\underline{s}^{k^2}})(F) - F||_{L^2(J^2;L(R^m;R^m))} = 0. \tag{5.8}$$

Proof Fix $\epsilon > 0$. There exist a continuous function $g : [-r, 0] \to R^m$ such that $||\eta - g||_V^2 < \epsilon$. By uniform continuity of g, there exists $\delta > 0$ such that

$$\sup\{|g(x) - g(y)| : x, y \in [-r, 0], |x - y| \leq \delta\} \leq \sqrt{\epsilon}.$$

Now choose k_0 sufficiently large so that $|s_i - s_{i-1}| < \delta$ for all $k \geq k_0$. Then

$$\begin{aligned}
||(I_k \circ Q_{\underline{s}^k})(g) - g||_V^2 &= \int_{-r}^0 |\sum_{i=1}^k \frac{I_{(s_{i-1},s_i]}(s)}{s_i - s_{i-1}} \int_{s_{i-1}}^{s_i} g(v)\,dv - \sum_{i=1}^k I_{(s_{i-1},s_i]}(s) g(s)|^2\,ds \\
&= \int_{-r}^0 |\sum_{i=1}^k \frac{I_{(s_{i-1},s_i]}(s)}{s_i - s_{i-1}} \int_{s_{i-1}}^{s_i} g(v) - g(s)\,dv|^2\,ds \\
&\leq r\epsilon, \text{ for all } k \geq k_0.
\end{aligned}$$

Also for η,

$$\begin{aligned}
||(I_k \circ Q_{\underline{s}^k})(\eta)||_V^2 &= \int_{-r}^0 |\sum_{i=1}^k \frac{I_{(s_{i-1},s_i]}(s)}{s_i - s_{i-1}} \int_{s_{i-1}}^{s_i} \eta(v)\,dv|^2\,ds \\
&\leq \int_{-r}^0 \sum_{i=1}^k \frac{I_{(s_{i-1},s_i]}(s)}{s_i - s_{i-1}} \int_{s_{i-1}}^{s_i} |\eta(v)|^2\,dv\,ds \\
&= \sum_{i=1}^k \int_{s_{i-1}}^{s_i} |\eta(v)|^2\,dv = ||\eta||_V^2, \text{ for all } k \geq 1.
\end{aligned}$$

Therefore

$$\begin{aligned}
||(I_k \circ Q_{\underline{s}^k})(\eta) - \eta||_V &\leq ||(I_k \circ Q_{\underline{s}^k})(\eta - g)||_V + ||(I_k \circ Q_{\underline{s}^k})(g) - g||_V + ||g - \eta||_V \\
&\leq (\sqrt{r} + 2)\sqrt{\epsilon}, \text{ for all } k \geq k_0.
\end{aligned}$$

Assertion (2) of the lemma follows by a similar argument. ∎

6 Weak differential rule for functionals of an infinite dimensional random variable.

Bell and Mohammed ([2]) derived weak differential rule for stochastic delay differential equations (with single delay) by using Malliavin calculus. In this section we shall prove more general results concerning the differential rule of D for functions of infinite-dimensional random variables by using Malliavin calculus too. We shall apply the differential rule in the derivation of Itô's formula for segment processes later. Recall $J=[-r,0]$, $T=[0,a]$, $V=L^2(J;R^m)$, $H=L^2(T;R^m)$ and $W=\{W(t)\}$ is a d-dimensional standard Brownian motion.

For a differentiable function $f(x,y)$, denote by f_1 and f_2 the partial derivatives of f with respect to the first and second variable, respectively. Let $f'=(f_1,f_2)$ and f'' be the Hessian of f. Denote by $V\times R^m$ the product space of $V=L^2([-r,0];R^m)$ and R^m, with usual the addition and scalar product rules, i.e., if $k\in R$, $\vec{x},\vec{y}\in R^m$ and $\eta,\psi\in V$, then $(\eta,\vec{x})+(\psi,\vec{y})=(\eta+\psi,\vec{x}+\vec{y})$ and $k(\eta,\vec{x})=(k\eta,k\vec{x})$. $V\times R^m$ is a Hilbert space endowed with the inner product

$$\langle(\eta,\vec{x}),(\psi,\vec{y})\rangle_{V\times R^m}=\langle\eta,\psi\rangle_V+\vec{x}\cdot\vec{y}. \tag{6.1}$$

Lemma 6.1 *Suppose $p\geq 0, F,G\in\mathbb{D}^{1,p}(V)$. Then $\langle F,G\rangle_V$ belongs to $\mathbb{D}^{1,p}$ and for almost all $t\in T$, $D_t\langle F,G\rangle_V=\langle D_tF,G\rangle_V+\langle F,D_tG\rangle_V$.*

Proof Let $S(V)$ be the family of smooth V-valued random variables (c.f. ([17] P61)). First let us assume $F,G\in S(V)$, i.e.,

$$\begin{cases} F=\sum_{i=1}^n F_i\eta_i, & F_i\in S,\eta_i\in V,\\ G=\sum_{j=1}^m G_j\psi_j, & G_j\in S,\psi_j\in V.\end{cases}$$

Since $D_t(F_iG_j)=(D_tF_i)G_j+F_i(D_tG_j)$,

$$\begin{aligned} D_t\langle F,G\rangle_V &= \sum_{i=1}^n\sum_{j=1}^m D_t(F_iG_j)\langle\eta_i,\psi_j\rangle_V\\ &= \sum_{i=1}^n\sum_{j=1}^m((D_tF_i)G_j+F_i(D_tG_j))\langle\eta_i,\psi_j\rangle_V\\ &= \langle D_tF,G\rangle_V+\langle F,D_tG\rangle_V.\end{aligned}$$

Now for any $F,G\in\mathbb{D}^{1,p}(V)$, there exist $\{F^n\},\{G^n\}\subset S(V)$ such that

$$\begin{cases} \lim_{n\to\infty}||F^n-F||_{1,p,V}=0,\\ \lim_{n\to\infty}||G^n-G||_{1,p,V}=0,\end{cases}$$

where the norm $||\cdot||_{1,p,V}$ is defined by (4.3). It is immediate that $\langle F^n,G^n\rangle_V\to\langle F,G\rangle_V$ in $L^p(\Omega)$ as $n\to\infty$. Since

$$\lim_{n\to\infty}||(\langle DF^n,G^n\rangle_V+\langle F^n,DG^n\rangle_V)-(\langle DF,G\rangle_V+\langle F,DG\rangle_V)||_{L^2(\times T)}=0,$$

in $L^p(\Omega)$, the result follows. ∎

Proposition 6.2 *Suppose $p \geq 2$, $h \geq 1$, $n \geq 2$, $\psi \in \mathbb{D}^{n,p}(V)$, $\vec{F} = (F_1, \ldots, F_h)'$, $F_i \in \mathbb{D}^{n,p}$, $i = 1, \ldots, h$ and $f \in C_b^n(V \times R^h)$. Then $f(\psi, \vec{F}) \in \mathbb{D}^{n,p}$, and for almost all $t \in T$,*

$$D_s^{j_1} f(\psi, \vec{F}) = \langle f'(\psi, \vec{F}), (D_s^{j_1}\psi, D_s^{j_1}\vec{F})\rangle_{V \times R^h}, \; j_1 = 1, \ldots, d, \text{ and} \tag{6.2}$$

$$\begin{aligned} D_t^{j_1} D_s^{j_2} f(\psi, \vec{F}) &= \langle f''(\psi, \vec{F})(D_t^{j_1}\psi, D_t^{j_1}\vec{F}), (D_s^{j_2}\psi, D_s^{j_2}\vec{F})\rangle_{V \times R^h} \\ &+ \langle f'(\psi, \vec{F}), (D_t^{j_1} D_s^{j_2}\psi, D_t^{j_1} D_s^{j_2}\vec{F})\rangle_{V \times R^h}, \; 1 \leq j_1, j_2 \leq d. \end{aligned}$$

Proof For simplicity, we assume $h = 1$ and we will prove result only for the case $n = 2$. The case $n > 2$ is similar. Suppose $\Pi_k : -r = s_0 < s_1 < \cdots < s_k = 0$ is a family of partitions of $[-r, 0]$, with $||\Pi_k|| \to 0$ as $k \to \infty$, I_k is the linear embedding and $Q_{\underline{s}^k}$ is the projection defined in Section 5. Denote $J_i := I_{(s_{i-1}, s_i]}$ and $\Delta_i := s_i - s_{i-1}$, for $i = 1, \ldots, k$.

We can write $\psi = (\psi_1, \ldots, \psi_m)'$, where $\psi_j \in \mathbb{D}^{n,p}(L^2(J; R))$, $1 \leq j \leq m$. Set $\psi^k = I_k \circ Q_{\underline{s}^k}(\psi)$, and define $f^k : R^{mk} \times R \to R$ by

$$f^k(\mathbf{X}, x_{k+1}) := f(I_k(\mathbf{X}), x_{k+1}),$$

where $\mathbf{X}$ is an m by k matrix with entries x_{ji}, and $x_{k+1} \in R$. Then $f^k \in C_b^2(R^{mk+1})$, and $f(\psi^k, F) = f^k(Q_{\underline{s}^k}(\psi), F)$. If $s \in [0, a]$, then

$$\begin{aligned} D_s^{j_2} f(\psi^k, F) &= D_s^{j_2} f^k(Q_{\underline{s}^k}(\psi), F) \\ &= \sum_{j,i} \frac{1}{\Delta_i} \frac{\partial f^k}{\partial x_{ji}}(Q_{\underline{s}^k}(\psi), F)\langle D_s^{j_2}\psi_j, J_i\rangle_V + \frac{\partial f^k}{\partial x_{k+1}}(Q_{\underline{s}^k}(\psi), F) D_s^{j_2} F \\ &= \langle f_1(\psi^k, F), I_k \circ Q_{\underline{s}^k}(D_s^{j_2}\psi)\rangle_V + f_2(\psi^k, F) D_s^{j_2} F. \end{aligned}$$

By Lemma 5.1,

$$D_s^{j_2} f(\psi^k, F) - \langle f'(\psi, F), (D_s^{j_2}\psi, D_s^{j_2} F)\rangle_{V \times R} \to 0 \tag{6.3}$$

as $k \to \infty$, a.s. $P \otimes \mu$. Let $M = \sup\{||f'(\eta, x)|| : \eta \in V, x \in R\}$. By Lemma 5.1,

$$|D_s^{j_2} f(\psi^k, F)| \leq M(||D_s^{j_2}\psi||_V + |D_s^{j_2} F|), \text{ and} \tag{6.4}$$

$$|\langle f'(\psi, F), (D_s^{j_2}\psi, D_s^{j_2} F)| \leq M(||D_s^{j_2}\psi||_V + |D_s^{j_2} F|). \tag{6.5}$$

By applying the Dominated Convergence Theorem, and by the fact $\psi \in \mathbb{D}^{2,p}(V)$ and $F \in \mathbb{D}^{2,p}$, we have

$$\lim_{k \to \infty} E\Big(\int_0^a |D_s^{j_2} f(\psi^k, F) - \langle f'(\psi, F), (D_s^{j_2}\psi, D_s^{j_2} F)\rangle_{V \times R}|^2 \, ds\Big)^{\frac{p}{2}} = 0. \tag{6.6}$$

Now let us consider the second part, by the chain rule of the Malliavin differential operator D,

$$
\begin{aligned}
& D_t^{j_1} D_s^{j_2} f(\psi^k, F) \\
= \; & D_t^{j_1}[\sum_{j,i} \frac{1}{\Delta_i} \frac{\partial f^k}{\partial x_{ji}} (Q_{\underline{s}^k}(\psi), F) \langle D_s^{j_2} \psi_j, J_i \rangle_V] + D_t^{j_1}[\frac{\partial f^k}{\partial x_{k+1}} (Q_{\underline{s}^k}(\psi), F) D_s^{j_2} F] \\
= \; & \langle f''(\psi^k, F)(I_k \circ Q_{\underline{s}^k}(D_t^{j_1}\psi), D_t^{j_1} F), (I_k \circ Q_{\underline{s}^k}(D_s^{j_2}\psi), D_s^{j_2} F) \rangle_{V \times R} \\
+ \; & \langle I_k \circ Q_{\underline{s}^k}(D_t^{j_1} D_s^{j_2} \psi), f_1(\psi^k, F) \rangle_V + f_2(\psi^k, F) D_t^{j_1} D_s^{j_2} F.
\end{aligned}
$$

Analogous to the argument showing the convergence of (6.6), we can show that

$$
\begin{aligned}
& \lim_{k\to\infty} E(\int_0^a \int_0^a |D_t^{j_1} D_s^{j_2} f(\psi, F) - (\langle f''(\psi, F)(D_t^{j_1}\psi, D_t^{j_1} F), (D_s^{j_2}\psi, D_s^{j_2} F) \rangle_{V \times R} \\
+ \; & \langle f'(\psi, F), (D_t^{j_1} D_s^{j_2} \psi, D_t^{j_1} D_s^{j_2} F) \rangle_{V \times R})|^2 \, ds \, dt)^{\frac{p}{2}} = 0.
\end{aligned}
$$

Since $f(\psi^k, F) \to f(\psi, F)$ in $L^p(\Omega)$, $\{f(\psi^k, F)\}_{k=1}^\infty$ has a subsequence which is a Cauchy sequence in $\mathbb{D}^{2,p}$ (we write this subsequence as $\{f(\psi^k, F)\}_{k=1}^\infty$ itself). By the closeness of the operator D, $f(\psi, F) \in \mathbb{D}^{2,p}$ and $f(\psi^k, F) \to f(\psi, F)$ in $\mathbb{D}^{2,p}$. Thus the proposition follows. ∎

As an application of Proposition 6.2, we can easily rewrite the SFDE (1.1) in Stratonovich integral form instead of the Itô integral form by the relationship between the Stratonovich and Skorohod integrals (c.f. [17] Definition 3.1.1 and Theorem 3.1.1). We skip the detail because it is not in the scope of this paper.

By Lemma 6.1 and Proposition 6.2, if $p \geq 2$, a V-valued process $X = \{X_t : 0 \leq t \leq a\}$ belongs to the space $\mathbb{D}1, p(H\hat{\otimes}_2 V)$, then there exists a measurable two-parameter process $DX = \{D_s X_t : t \in [0, a]\}$, such that

(1) $E(\int_0^a ||DX_t||^2_{H\hat{\otimes}_2 V} \, dt)^{\frac{p}{2}} < \infty$,

(2) for all $\eta \in V$, the process $\langle X, \eta \rangle_V$ belongs to $\mathbb{L}^{1,p}$, and

(3) $D_s^i(\langle X_t, \eta \rangle_V) = \langle D_s^i X_t, \eta \rangle_V$, for all $1 \leq i \leq d$, $\eta \in V$, and almost all $s, t \in [0, a]$.

Note 6.3 *If* $X = \{X(t) : 0 \leq t \leq a\} \in \mathbb{L}_m^{1,p}$ *and* $X(t) = 0$ *if* $t > a$, *then* $\{X_t\} \in \mathbb{D}^{1,p}(H\hat{\otimes}_2 V)$.

Suppose $\{Y(t)\} \in \mathbb{L}_d^{1,p}$, $f \in C_b^1(T \times V; V)$, and $X_t = f(t, Y_t)$, we shall show that $\{X_t\} \in \mathbb{D}^{1,p}(H\hat{\otimes}_2 V)$. Denote by $f_2(t, \eta)$ the derivative of $f(t, \eta)$ with respect to the second variable η.

Set $z_k(t) := \langle f(t, I_k(Q_{\underline{s}^k}(Y_t))), I_k(Q_{\underline{s}^k}(\eta)) \rangle_V$ and $z(t) := \langle f(t, Y_t), \eta \rangle_V$, where $0 \leq t \leq a$, and $\eta \in V$.

Lemma 6.4 *For each $\eta \in V$, $\{z_k : k \geq 1\}$ converges to z in $\mathbb{L}^{1,p}$ and*

$$D_s^i z(t) = \langle f_2(t, Y_t)(D_s^i Y_t), \eta \rangle_V,$$

for all $1 \leq i \leq d$ and almost all $s, t \in [0, a]$. Furthermore, $\{f(t, Y_t)\}$ belongs to $\mathbb{D}^{1,p}(H \hat{\otimes}_2 V)$ and $D_s^i f(t, Y_t) = f_2(t, Y_t)(D_s^i Y_t)$ for all $1 \leq i \leq d$.

Proof Suppose $n > k$. By Lemma 5.1,

$$\begin{aligned}
|z_k(t) - z_n(t)| &\leq |\langle f(t, I_k(Q_{\underline{s}^k}(Y_t))) - f(t, I_n(Q_{\underline{s}^n}(Y_t))), I_k(Q_{\underline{s}^k}(\eta)) \rangle_V| \\
&+ |\langle f(t, I_n(Q_{\underline{s}^n}(Y_t))), I_k(Q_{\underline{s}^k}(\eta)) - I_n(Q_{\underline{s}^n}(\eta)) \rangle_V| \\
&\leq ||f(t, I_k(Q_{\underline{s}^k}(Y_t))) - f(t, I_n(Q_{\underline{s}^n}(Y_t)))||_V ||\eta||_V \\
&+ \sup_{\eta \in V} ||f(\eta)||_V ||I_k(Q_{\underline{s}^k}(\eta)) - I_n(Q_{\underline{s}^n}(\eta))||_V \\
&\to 0 \quad \text{as} \quad k \to \infty.
\end{aligned}$$

By the Bounded Convergence Theorem, $(\lim_{k\to\infty} E(\int_0^a |z_k(t) - z_n(t)|^2\, dt)^{\frac{p}{2}} = 0$. Similarly we can show that

$$\lim_{k\to\infty} E(\int_0^a |z_k(t) - z(t)|^2\, dt)^{\frac{p}{2}} = 0.$$

By applying the chain rule we have

$$\begin{aligned}
D_s^i(z_k(t)) &= \langle f_2(t, I_k(Q_{\underline{s}^k}(Y_t)))(I_k(D_s^i(Q_{\underline{s}^k}(Y_t))), I_k(Q_{\underline{s}^k}(\eta))) \rangle \\
&= \langle f_2(t, I_k(Q_{\underline{s}^k}(Y_t)))(I_k(Q_{\underline{s}^k}(D_s^i Y_t))), I_k(Q_{\underline{s}^k}(\eta))) \rangle
\end{aligned}$$

Analogous to the above argument, by using the Dominated Convergence Theorem, we can show that

$$\lim_{k\to\infty} E(\int_0^a \int_0^a |D_s^i(z_k(t) - z_n(t))|^2\, ds\, dt)^{\frac{p}{2}} = 0$$
$$\lim_{k\to\infty} E(\int_0^a \int_0^a |D_s^i z_k(t) - \langle f_2(t, Y_t)(D_s^i Y_t), \eta \rangle_V|^2\, ds\, dt)^{\frac{p}{2}} = 0.$$

Thus $||z_k - z_n||_{1,p}^p \to 0$ as $k \to \infty$. By Lemma 5.1, we conclude that $||z_k - z||_{1,p}^p \to 0$ and $D_s^i z(t) = \langle f_2(t, Y_t)(D_s^i Y_t), \eta \rangle_V$ for almost all $s, t \in [0, a]$. ∎

7 L^p approximation of the segment integral.

Using finite dimensional-valued sequences of variables to approximate infinite dimensional-valued variables is a powerful technique when we go into infinite dimensional spaces. Suppose

$p \geq 2$, in this section we shall give some L^p approximations of the segment integral. As we shall see in the following section, the approximation techniques and results derived in this section help us reach the Itô formula for segment processes. As before, we assume $J = [-r, 0]$, $T = [0, a]$, $V = L^2(J; R^m)$ and $H = L^2(T; R^m)$.

Suppose $W = \{W(t)\}$ is a m-dimensional standard Brownian motion, $u \in \mathbb{L}_m^{1,2}$, $-r \leq \alpha < 0$, we can define the integral

$$\int_0^t u(s+\alpha) \cdot dW(s+\alpha) := \int_\alpha^{t+\alpha} u(s) \cdot dW(s) \tag{7.1}$$

by change of variable or by using approximation scheme (c.f. [17] Section 3.1). We define $W(t) = 0$ if $t < 0$.

If $X \in \mathbb{D}^{1,p}(H \hat{\otimes}_2 V)$, and $\{\Pi_k : -r = s_0 < \cdots < s_k = 0\}_{k=1}^\infty$ is a sequence of partitions of $[-r, 0]$, we can use the Skorohod integrals

$$J_{\Pi_k}(a) := \int_0^a \langle Q_{\underline{s}^k}(X_t), \cdot dQ_{\underline{s}^k}(W_t) \rangle \tag{7.2}$$

to approximate the segment integral $\int_0^a \langle X_t, dW_t \rangle_V$.

If $v \in R$, we define a Heaviside-type function $H_v \in L^2[-r, 0]$ by

$$H_v(s) = \begin{cases} 1, & v \leq s \leq 0 \\ 0, & -r \leq s < v. \end{cases} \tag{7.3}$$

If $\psi = (\psi_1, \ldots, \psi_m) \in V = L^2(J; R^m)$ and $\eta \in L^2(J; R)$, then we define

$$\langle \psi, \eta \rangle_V := (\langle \psi_1, \eta \rangle_{L^2(J;R)}, \ldots, \langle \psi_m, \eta \rangle_{L^2(J;R)}). \tag{7.4}$$

By Lemma 4.2, we have

$$\begin{aligned} J_{\Pi_k}(a) &= \int_0^a \langle Q_{\underline{s}^k}(\hat{X}_t^{(k)})), Q_{\underline{s}^k}(H_{t-a}) \rangle \cdot dW(t) \\ &= \int_0^a \langle I_k(Q_{\underline{s}^k}(\hat{X}_t^{(k)})), I_k(Q_{\underline{s}^k}(H_{t-a})) \rangle_V \cdot dW(t), \end{aligned}$$

where

$$\hat{X}_t^{(k)}(s) := \sum_{i=1}^k \frac{I_{(s_{i-1}, s_i]}(s)}{s_i - s_{i-1}} \int_{s_{i-1}}^{s_i} X_{t-s}(v)\, dv, \quad s \in [-r, 0]. \tag{7.5}$$

If we can show that the sequence

$$\{\langle I_k(Q_{\underline{s}^k}(\hat{X}_t^{(k)})), I_k(Q_{\underline{s}^k}(H_{t-a})) \rangle_V\}_{k=1}^\infty \tag{7.6}$$

converge to $(\Gamma_H^* X)(t)$ under the norm of $\mathbb{L}_m^{1,p}$, then by boundedness of the operator $\delta : \mathbb{L}_m^{1,2} \to L^2(\Omega)$, the Skorohod integral $J_{\Pi_k}(a)$ converge in the space $L^2(\Omega)$ to $\int_0^a \langle X_t, dW_t \rangle_V$.

In the following we assume $X_t = 0$ if $t < 0$ or $t > a$. We also define a rotation $\hat{X}$ of X by

$$\hat{X}_t(s) := X_{t-s}(s). \tag{7.7}$$

We may think that the process $\hat{X}_t$ is simply an element of $L^2(\Omega \times J \times T; R^m)$. We can also write $\int_{-r}^0 X_{t-s}(s) I_{\{t \le a+s\}}\, ds$ as $\langle \hat{X}_t, H_{t-a}\rangle_V$, which is just $(\Gamma_H^* X)(t)$, or $\Gamma_H^*(X_.)(t)$.

Lemma 7.1 *Suppose* $X_t \in \mathbb{D}^{1,p}(H\hat{\otimes}_2 V)$, *and* $\{\Pi_k\}$ *is a sequence of partitions of* $[-r, 0]$, *with* $\lim_{k\to\infty} ||\Pi_k|| = 0$. *Then* $E(\int_0^a ||\hat{X}_t^{(k)}||_V^2\, dt)^{\frac{p}{2}} \le E(\int_0^a ||X_t||_V^2\, dt)^{\frac{p}{2}}$, *and*

$$\lim_{k\to\infty} E(\int_0^a \int_{-r}^0 |\hat{X}_t^{(k)}(s) - \hat{X}_t(s)|^2\, ds\, dt)^{\frac{p}{2}} = 0. \tag{7.8}$$

Proof First we show that the inequalities hold almost surely. Fix $\omega \in \Omega$ and $\epsilon = \epsilon(\omega) > 0$. There exists a continuous function $g : [0, a+r] \times [-r, 0] \to R^m$ such that

$$\int_0^{a+r} \int_{-r}^0 |X_t(s,\omega) - g(t,s)|^2\, ds\, dt < \epsilon. \tag{7.9}$$

Define

$$J_k(t,s,\omega) := \sum_{i=1}^k \frac{I_{(s_{i-1},s_i]}(s)}{s_i - s_{i-1}} \int_{s_{i-1}}^{s_i} [X_{t-s}(v,\omega) - X_{t-v}(v,\omega)]\, dv. \tag{7.10}$$

We can write $J_k(t,s)$ as $J_k(t,s,\omega) = J_{k1}(\omega) + J_{k2} + J_{k3}(\omega)$, where

$$\begin{aligned}
J_{k1}(\omega) &= \sum_{i=1}^k \frac{I_{(s_{i-1},s_i]}(s)}{s_i - s_{i-1}} \int_{s_{i-1}}^{s_i} [X_{t-s}(v,\omega) - g(t-s,v)]\, dv, \\
J_{k2} &= \sum_{i=1}^k \frac{I_{(s_{i-1},s_i]}(s)}{s_i - s_{i-1}} \int_{s_{i-1}}^{s_i} [g(t-s,v) - g(t-v,v)]\, dv, \\
J_{k3}(\omega) &= \sum_{i=1}^k \frac{I_{(s_{i-1},s_i]}(s)}{s_i - s_{i-1}} \int_{s_{i-1}}^{s_i} [g(t-v,v) - X_{t-v}(v,\omega)]\, dv.
\end{aligned}$$

Since g is uniformly continuous on $[0, a+r] \times [-r, 0]$, there exist $\delta > 0$, such that if $t_1, t_2 \in [0, a+r], s_1, s_2 \in [-r, 0], \max\{|t_1 - t_2|, |s_1 - s_2|\} < \delta$, then $|g(t_1, s_1) - g(t_2, s_2)| < \sqrt{\epsilon/ar}$. If $||\Pi_k|| < \delta$, then

$$\begin{aligned}
|J_{k2}| &\le \sum_{i=1}^k \frac{I_{(s_{i-1},s_i]}(s)}{s_i - s_{i-1}} \int_{s_{i-1}}^{s_i} |g(t-s,v) - g(t-v,v)|\, dv \\
&\le \sum_{i=1}^k \frac{I_{(s_{i-1},s_i]}(s)}{s_i - s_{i-1}} \int_{s_{i-1}}^{s_i} \sqrt{\frac{\epsilon}{ar}}\, dv = \sqrt{\frac{\epsilon}{ar}}.
\end{aligned}$$

We also have

$$\begin{aligned}
\int_0^a \int_{-r}^0 |J_{k1}(\omega)|^2 \, dv \, dt &\leq \int_0^a \sum_{i=1}^k \int_{s_{i-1}}^{s_i} |X_{t-s}(v,\omega) - g(t-s,v)|^2 \, dv \, dt \\
&= \sum_{i=1}^k \int_{s_{i-1}}^{s_i} \int_0^a |X_{t-s}(v,\omega) - g(t-s,v)|^2 \, dt \, dv \\
&\leq \sum_{i=1}^k \int_{s_{i-1}}^{s_i} \int_0^{a+r} |X_t(v,\omega) - g(t,v)|^2 \, dt \, dv \\
&= \int_{-r}^0 \int_0^{a+r} |X_t(v,\omega) - g(t,v)|^2 \, dt \, dv < \epsilon.
\end{aligned}$$

Analogous to the above argument we have

$$\int_0^a \int_{-r}^0 |\sum_{i=1}^k \frac{I_{(s_{i-1},s_i]}(s)}{s_i - s_{i-1}} \int_{s_{i-1}}^{s_i} X_{t-s}(v) \, dv|^2 \, ds \, dt \leq \int_0^a \int_{-r}^0 |X_t(s)|^2 \, ds \, dt, \tag{7.11}$$

which is equivalent to the inequality

$$E(\int_0^a ||\hat{X}_t^{(k)}||_V^2 \, dt \leq E \int_0^a ||X_t||_V^2 \, dt)^{\frac{p}{2}}. \tag{7.12}$$

Similarly, we can show that $\int_0^a \int_{-r}^0 |J_{k3}(\omega)|^2 \, ds \, dt < \epsilon$. Thus we have

$$\lim_{k\to\infty} \int_0^a \int_{-r}^0 |J_k(t,s)|^2 \, ds \, dt = 0 \quad \text{a.s.} \tag{7.13}$$

For each $t \in [0,a]$, $\int_{-r}^0 X_{t-s}^2(s) \, ds < \infty$ a.s. By Lemma 5.1, we have

$$\lim_{k\to\infty} \int_{-r}^0 \{\sum_{i=1}^k \int_{s_{i-1}}^{s_i} X_{t-s}(s) \, ds \frac{I_{(s_{i-1},s_i]}(s)}{s_i - s_{i-1}} - X_{t-s}(s)\}^2 \, ds = 0 \text{ a.s.} \tag{7.14}$$

By the Dominated Convergence Theorem,

$$\lim_{k\to\infty} \int_0^a \int_{-r}^0 \{\sum_{i=1}^k \int_{s_{i-1}}^{s_i} X_{t-s}(s) \, ds \frac{I_{(s_{i-1},s_i]}(s)}{s_i - s_{i-1}} - X_{t-s}(s)\}^2 \, ds \, dt = 0. \tag{7.15}$$

From (7.10), (7.13) and (7.15), it follows that

$$\lim_{k\to\infty} \int_0^a \int_{-r}^0 |\sum_{i=1}^k \frac{I_{(s_{i-1},s_i]}(s)}{s_i - s_{i-1}} \int_{s_{i-1}}^{s_i} X_{t-s}(v) \, dv - \hat{X}_t(s)|^2 \, ds \, dt = 0 \tag{7.16}$$

a.s. -P. By (7.11) and the Dominated Convergence Theorem, (7.8) follows. ∎

Let $D = (D^1, \ldots, D^d)'$ be the Malliavin differential operator associated with a d-dimensional standard Brownian motion. Analogous to (7.5), we define a $d\times m$-matrix-valued three parameter random process by

$$\widehat{D_u X}_t^{(k)}(s) := \sum_{i=1}^k \frac{I_{(s_{i-1},s_i]}(s)}{s_i - s_{i-1}} \int_{s_{i-1}}^{s_i} D_u X_{t-s}(v) \, dv. \quad s \in [-r, 0] \tag{7.17}$$

Denote by $D_u X_t^{j,(k)}$ the $j-$th row vector of the $d \times m$ matrix-valued process $D_u X_t^{(k)}$. We can prove the convergence result of the derivative process of $\{X_t\}$ in a similar way, i.e., for each $1 \le j \le d$,

$$E(\int_0^a \int_0^a ||\widehat{D_u X}_t^{j,(k)}||_V^2 \, du \, dt \le E \int_0^a \int_0^a ||D_u^j X_t||_V^2 \, ds \, dt)^{\frac{p}{2}}, \tag{7.18}$$

$$\lim_{k \to \infty} E(\int_0^a \int_0^a \int_{-r}^0 |\widehat{D_u X}_t^{j,(k)}(s) - D_u^j X_{t-s}(s)|^2 \, ds \, du \, dt)^{\frac{p}{2}} = 0. \tag{7.19}$$

Next we show that $\hat{X}_t$ belongs to $\mathbb{D}^{1,p}(H \hat{\otimes}_2 V)$ and $\int_{-r}^0 \hat{X}_t(s) \, ds$ is the limit of the sequence defined by (7.6).

Lemma 7.2 *For each $k \ge 1$, $\hat{X}^{(k)} \in \mathbb{D}^{1,2}(H \hat{\otimes}_2 V)$, and the derivative process $D_u \hat{X}_t^{(k)}$ is equal to $\widehat{D_u X}_t^{(k)}$.*

Proof If $\eta \in V$, then

$$\langle \hat{X}_t^{(k)}, \eta \rangle_V = \sum_{i=1}^k \frac{1}{s_i - s_{i-1}} \int_{s_{i-1}}^{s_i} \int_{s_{i-1}}^{s_i} X_{t-s}(v) \cdot \eta(v) \, dv \, ds. \tag{7.20}$$

For each $1 \le i \le k$ and $-r \le s \le 0$,

$$\int_{s_{i-1}}^{s_i} X_{t-s}(v) \, dv = \langle X_{t-s}, I_{(s_{i-1}, s_i]} \rangle_V \in \mathbb{L}_m^{1,p}.$$

Thus $\langle \hat{X}_t^{(k)}, \eta \rangle_V \in \mathbb{L}_m^{1,p}$. Since for each $1 \le j \le d$,

$$\begin{aligned} D_u^j \langle \hat{X}_t^{(k)}, \eta \rangle_V &= \sum_{i=1}^k \frac{1}{s_i - s_{i-1}} \int_{s_{i-1}}^{s_i} \eta(v) \langle D_u^j X_{t-v}, I_{(s_{i-1}, s_i]} \rangle_V \, dv \\ &= \langle \widehat{D_u X}_t^{j,(k)}, \eta \rangle_V. \end{aligned}$$

By Inequality (7.18) we have

$$E(\int_0^a \int_0^a ||\widehat{D_u X}_t^{j,(k)}||_V^2 \, du \, dt)^{\frac{p}{2}} \le E(\int_0^a \int_0^a ||D_u^j X_t||_V^2 \, ds \, dt)^{\frac{p}{2}} < \infty. \tag{7.21}$$

So $\hat{X}^{(k)} \in \mathbb{D}^{1,p}(H \hat{\otimes}_2 V)$, and $D_u \hat{X}_t^{(k)}$ is equal to $\widehat{D_u X}_t^{(k)}$. ∎

Now we obtain the main results of this section.

Theorem 7.3 *Under the same hypotheses as Lemma 7.1, $\hat{X} = \{\hat{X}_t\}$ belongs to the space $\mathbb{D}^{1,p}(H \hat{\otimes}_2 V)$, and for all $\eta \in V$ and $s, t \in [0, a]$,*

$$\int_{-r}^0 D_u \hat{X}_t(s) \eta(s) \, ds = \int_{-r}^0 D_u X_{t-s}(s) \eta(s) \, ds \text{ and} \tag{7.22}$$

$$\langle I_k(Q_{\underline{s}^k}(\hat{X}_t^{(k)})), I_k(Q_{\underline{s}^k}(H_{t-a})) \rangle_V \to \Gamma_H^*(X_\cdot)(t) \tag{7.23}$$

as $k \to \infty$ in $\mathbb{L}_m^{1,p}$.

Proof By Lemma 7.1 – 7.2, $\hat{X} = \{\hat{X}_t\}$ belongs to $\mathbb{D}^{1,p}(H\hat{\otimes}_2 V)$, and for all $\eta \in V$ and almost all $s, t \in [0, a]$, $\int_0^a D_u \hat{X}_t(s)\eta(s)\, ds = \int_0^a D_u X_{t-s}(s)\eta(s)\, ds$. Since

$$\lim_{k\to\infty} ||\langle \hat{X}^{(k)}, \zeta\rangle_V - \langle \hat{X}, \zeta\rangle_V||_{\mathbb{L}^{1,p}_m} = 0, \tag{7.24}$$

for all $\zeta \in L^2(J; R)$, (7.23) follows. ∎

Since δ is a bounded linear operator form $\mathbb{L}^{1,2}$ into $\mathbb{D}^{0,2} \cong L^2(\Omega)$, we have

Theorem 7.4 *For any sequence of partitions* $\{\Pi_k\}$,

$$\lim_{k\to\infty} E\{J_{\Pi_k}(a) - \int_0^a \int_{-r}^0 X_{t-s}(s) I_{\{t\le a+s\}}(s)\, ds\, dW(t)\}^2 = 0, \tag{7.25}$$

where $J_{\Pi_k}(a)$ *is defined by (7.5).*

Proof From the fact that $\int_{-r}^0 X_{t-s}(s) I_{\{t\le a+s\}}(s)\, ds$ belongs to $\mathbb{L}^{1,2}_m$, the Skorohod integral

$$\int_0^a \int_{-r}^0 X_{t-s}(s) I_{\{t\le a+s\}}\, ds \cdot dW(t)$$

belongs to $L^2(\Omega)$. ∎

By Definition 4.1 and Theorem 7.3, the Skorohod segment integral $\int_0^t \langle X_v, dW_v\rangle_V$ can also be defined in the following way:

$$\int_0^t \langle X_v, dW_v\rangle_V := \int_0^a \langle I_{[0,t]}(v) X_v, dW_v\rangle_V \tag{7.26}$$

We can rewrite the right hand side of (7.26) as

$$\begin{aligned}\int_0^a \langle I_{[0,t]}(v) X_v, dW_v\rangle_V &= \int_0^a \int_{-r}^0 X_{v-s}(s) I_{\{v-s\le t\}} I_{\{v\le a+s\}}\, ds \cdot dW(v)\\ &= \int_0^t \int_{-r}^0 X_{v-s}(s) I_{\{v\le t+s\}}\, ds \cdot dW(v).\end{aligned}$$

Suppose X_t is a $V-$valued stochastic process, and $U(t)$ is a m-dimensional process defined by

$$U(t) = \int_0^t u(s)\, dW(s), \tag{7.27}$$

where $u : \Omega \times T \to L(R^d; R^m)$. We define

$$\int_0^t \langle X_s, dU_s\rangle_V = \int_0^t [\int_{-r}^0 X_{\alpha-s}(s) I_{\{\alpha\le t+s\}}\, ds u(\alpha)]\, dW(\alpha) \tag{7.28}$$

if the Skorohod integral of the right hand side of (7.28) exists. Since

$$||I_k(Q_{\underline{s}^k}(H_{t-a})) - I_k(P_{\underline{s}^k}(H_{t-a}))||_V \le ||\Pi_k||,$$

we can also use the Skorohod integrals

$$J_{\Pi_k}(a) := \int_0^a \langle Q_{\underline{s}^k}(X_t), dP_{\underline{s}^k}(W_t)\rangle \tag{7.29}$$

to approximate the segment integral $\int_0^a \langle X_t, dW_t\rangle_V$. In this case, we can use the sequence

$$\{\langle I_k(Q_{\underline{s}^k}(X_t)), I_k(P_{\underline{s}^k}(H_{t-a}))\rangle_V\}_{k=1}^{\infty} \tag{7.30}$$

to approximate $\Gamma_H^*(X_\cdot)(t)$, for $t \in [0,a]$.

Theorem 7.5 *Under the same hypotheses as Lemma 7.1, we have*

$$\lim_{k\to\infty} \langle I_k(Q_{\underline{s}^k}(X_t)), I_k(P_{\underline{s}^k}(H_{t-a}))\rangle_V = \Gamma_H^*(X_\cdot)(t) \tag{7.31}$$

in $\mathbb{L}_m^{1,p}$.

Remark 7.6 *Suppose X_t is a $C[-r,0]-$valued stochastic process. Analogous to (7.5), we may want to use the sequence*

$$J_{\Pi_k}(a) = \int_0^a (I_k(P_{\underline{s}^k}(X_t)))(I_k(P_{\underline{s}^k}(H_{t-a})))\ \cdot dW(t)$$

to approximate the integral $\int_0^a X_t\, dW_t$. Since almost surely, $\{I_k(P_{\underline{s}^k}(H_{t-a}))\}$ is not a Cauchy sequence in $C[-r,0]$, and $I_k(P_{\underline{s}^k}(X_t))$ converge to X_t in $(C[-r,0])^*$), the sequence

$$\{(I_k(P_{\underline{s}^k}(X_t)))(I_k(P_{\underline{s}^k}(H_{t-a})))\}$$

is not a Cauchy sequence (even in Probability). Thus it is difficult to show that $J_{\Pi_k}(a)$ has a limit.

This is also the difficulty if we want to use the above approximation scheme to define the integral $\int_0^a X_t\, dW_t$ for a $(C[-r,0])^*$ -valued process X_t with respect to the Brownian segment.

8 Itô's formula for the segment process

Now we are ready to derive Itô's Formula for the segment associated with the "semimartingale" process defined by (1.5). Throughout this section, we assume

$$\{\Pi_k : -r = s_0 < s_1 < \cdots < s_k = 0\}_{k=1}^{\infty}$$

is a family of partitions of $[-r,0]$, with $||\Pi_k|| \to 0$ as $k \to \infty$, I_k is the linear embedding and $Q_{\underline{s}^k}$ is the projection defined in Section 5. Denote $J = [-r,0]$, $T = [0,a]$, $V = L^2(J;R^m)$, $H = L^2(T;R^m)$, $J_i := I_{(s_{i-1},s_i]}$ and $\Delta_i := s_i - s_{i-1}$, for $i = 1,\ldots,k$.

For an $m \times k$ matrix $\mathbf{X}$, we denote by $\vec{x}^i$ its i-th column vector and by x_{ij} its entry at the i-th row and j-th column. Suppose $f = f(t, \eta, \vec{x}^{m+1}) \in C_b^2(T \times V \times R^m)$. For each $k \geq 1$ we define a $C^2(T \times R^{mk} \times R^m; R)$ function by

$$f^k(t, \mathbf{X}, \vec{x}^{m+1}) := f(t, I_k(\mathbf{X}), \vec{x}^{m+1}) = f(t, I_k(\vec{x}^1), \dots, I_k(\vec{x}^m), \vec{x}^{m+1}), \tag{8.1}$$

where $\vec{x}^{m+1} = (x_{m+1,1}, \dots, x_{m+1,k}) \in R^k$. Suppose $\eta = (\eta_1, \dots, \eta_m)$, where $\eta_j \in L^2(J; R)$, $1 \leq j \leq m$. We write $f(t, \eta, \vec{x}^{m+1})$ as $f(t, \eta_1, \dots, \eta_m, \vec{x}^{m+1})$. In this case, denote by $f_j(t, \eta_1, \dots, \eta_m, \vec{x}^{m+1})$ the derivative of $f(t, \eta_1, \dots, \eta_m, \vec{x}^{m+1})$ with respect to the variable η_j (or to the vector $\vec{x}^{m+1}$, if $j = m+1$) and $\frac{\partial f}{\partial t}$ the derivative of f with respect to t. Similarly, we denote by $f_{j_1 j_2}$ the derivative of f_{j_1} with respect to the variable η_{j_2} (or $\vec{x}^{m+1}$, if $j_2 = m+1$). We will use similar notations for the derivatives of $f^k(t, \mathbf{X}, \vec{x}^{m+1}) = f^k(t, \vec{x}^1, \dots, \vec{x}^m, \vec{x}^{m+1})$. Suppose $\mathbf{Y}$ and $\mathbf{Z}$ are $m \times k$ matrices. If $1 \leq j \leq m$, then

$$f_j^k(t, \mathbf{X}, \vec{x}^{m+1}) \cdot \vec{y}^j = \langle f_j(t, I_k(\mathbf{X}), \vec{x}^{m+1}), I_k(\vec{y}^j) \rangle_{L^2(J;R)}. \tag{8.2}$$

If $1 \leq j_1, j_2 \leq m$, then

$$f_{j_1 j_2}^k(t, \mathbf{X}, \vec{x}^{m+1})(\vec{z}^{j_2} \otimes \vec{y}^{j_1}) = f_{j_1 j_2}(t, I_k(\mathbf{X}), \vec{x}^{m+1})(I_k(\vec{y}^{j_1}) \otimes I_k(\vec{z}^{j_2})). \tag{8.3}$$

Suppose H_1 and H_2 are two Hilbert spaces. Denote by $H_1 \hat{\otimes}_2 H_2$ and $H_1 \hat{\otimes}_1 H_2$ the completions of $H_1 \otimes H_2$ under the $\epsilon-$topology ([21] Section 43) and the $\pi-$topology (projective topology) ([21] Section 43) on $H_1 \otimes H_2$ respectively.

Suppose $A \in L^2(J; R^m) \otimes_1 L^2(J; R^m)$ and $\Phi \in (L^2(J; R^m) \otimes_1 L^2(J; R^m))^*$. We can "decompose" A and Φ into $m \times m$ matrices (still denoted by A and Φ) with entries in $L^2(J; R) \otimes_1 L^2(J; R)$ and $(L^2(J; R) \otimes_1 L^2(J; R))^*$ respectively. We then formally write $\Phi(A)$ as $Tr(\Phi \cdot A)$, where Tr is the trace of a square matrix.

For a twice differentiable function $f = f(t, \eta, \vec{x}) \in C^2(T \times V \times R^m)$, the derivative $\frac{\partial^2 f}{\partial \eta^2}(t, \eta, \vec{x})$ is in general an element of $(V \hat{\otimes}_1 V)^*$. Due to this, two alternative conditions must be satisfied in order to prove the Itô formula:

Hypotheses 8.A The map $(t, \eta, \vec{x}) \hookrightarrow \frac{\partial^2 f}{\partial \eta^2}(t, \eta, \vec{x})$ is uniformly continuous from $T \times V \times R^m$ into $(V \hat{\otimes}_2 V)^*$ on bounded sets.

Hypotheses 8.B The map $(t, \eta, \vec{x}) \hookrightarrow \frac{\partial^2 f}{\partial \eta^2}(t, \eta, \vec{x})$ is uniformly continuous from $T \times V \times R^m$ into $(V \hat{\otimes}_1 V)^*$ on bounded sets, and the linear operator on $H = L^2([0, a]; R^m)$ with kernel $u(s) D_s X(s)$ is of trace class almost surely, where u is the coefficient process in (1.5).

If the $m \times d$ matrix-valued coefficient process u in the SDE (1.5) is deterministic, then

$u(s)D_sX(t)$ is of trace class. One example of a function satisfying Hypotheses 8.A is $f(t,\eta) = h(t, \langle \zeta, \eta \rangle_V)$, where $\zeta \in V$ and $h \in C^2(T \times R)$.

Note that $V\hat{\otimes}_1 V \subsetneq V\hat{\otimes}_2 V$. Although Hypotheses 8.B allows a larger class of functions $f \in C_b^2(T \times V \times R^m)$, it requires an additional condition on the solution process $\{X(t)\}$.

When we are saying "a function $a(s,t)$ is of trace class", we mean that the operator associated with the kernel $a(s,t)$ is of trace class.

Lemma 8.1 *If u is an $m \times d$ matrix-valued process with row vectors $u^1, \ldots, u^m$, where $u^i \in L^2(T; R^d)$, and $v \in \mathbb{L}_m^{1,2}$, then the kernel $u(s)\int_0^t D_s v(r)\,dr$ (as an $m \times m$ matrix-valued function) is of trace class. (We skip the words "almost surely").*

Proof We write $v = (v_1, \ldots, v_m)'$. For all $1 \le i \le m$ and $1 \le j \le d$, the functions $u^{ij}(s)D_s^j v^i(r)$ and $I_{\{0 \le r \le t\}}$ belong to $L^2(T^2)$ (with the notation $D = (D^1, \ldots, D^d)'$), and

$$u^{ij}(s)\int_0^t D_s^j v_i(r)\,dr = \int_0^a u^{ij}(s)D_s^j v_i(r) I_{\{0 \le r \le t\}}\,dr, \tag{8.4}$$

it follows that the operator with kernel $u^{ij}(s)\int_0^t D_s^j v_i(r)\,dr$ is the composite of two Hilbert-Schmidt operators with kernel $u^{ij}(s)D_s^j v_i(r)$ and $I_{\{0 \le r \le t\}}$, and it must be of trace class. ∎

One can refer to Elworthy ([5] Section V.1), Metivier and Pellaumail ([15] Chapter 2), Da Prato and Zabczyk ([4] Section 9.45) for Itô's formula for infinite-dimensional semimartingales. Because the quadratic variation (operator-valued) process of a martingale process is always of trace class, the Itô formula is valid for any twice differentiable function $f \in C^2(B)$ satisfying the condition that f, f' and f'' are uniformly continuous on bounded sets of B, where B is a Banach space with a "smooth" norm.

Lemma 8.2 *Suppose $F(s,t)$ is a function in $L^2(T^2; L(R^m; R^m))$, which is the kernel of a trace class operator F on $H = L^2(T; R^m)$. Define $F(s,t) = 0$ if $s < 0$ or $t < 0$, and define*

$$F_s(\alpha, \beta) := F(s+\alpha, s+\beta), \text{ where } s \in T, \alpha, \beta \in J = [-r, 0]. \tag{8.5}$$

Then for each $s \in T$, f_s is the kernel of a trace class operator on $V = L^2(J; R^m)$.

Proof Since F can be written as BA, where A an B are two Hilbert-Schimidt operators on H ([8] Problem 572). Assume (matrix-valued functions) $A(s,t)$ and $B(s,t)$ are the kernels of A and B, then F can be written as

$$F(s,t) = \int_0^a A(s,u) \cdot B(u,t)\,du.$$

It follows that

$$F_s(\alpha,\beta) = \int_0^a A(s+\alpha,u)\cdot B(u,s+\beta)\,du.$$

For each $s\in T$, $A^s(\alpha,u) := A(s+\alpha,u)$ and $B^s(u,\beta) := B(u,s+\beta)$ are kernels of Hilbert-Schmidt operators $A^s : V\to H$ and $B^s : H\to V$, respectively. Let F^s be the linear operator on V with kernel F_s, then $F^s = B^sA^s$, which is of trace class. ∎

Now we are ready to prove Itô's formula for $f(t,X_t)$.

Theorem 8.3 *Suppose $X(t)$ is a continuous stochastic process defined by (1.5), where u has row vectors $u^1,\ldots,u^m$, $u^i\in\mathbb{L}^{2,4}_{d,loc}$, $v\in\mathbb{L}^{1,4}_{m,loc}$, and $\eta : J\to R^m$ is a function of bounded variation, Assume $f = f(t,\eta)\in C_b^1(T\times V)$ has bounded continuous second order partial derivative with respect to η (i.e., $\frac{\partial^2 f}{\partial\eta^2}\in C_b(T\times V)$) and Hypotheses 8.A, then we have*

$$\begin{aligned} f(t,X_t)-f(0,X_0) &= \int_0^t \frac{\partial f}{\partial s}(s,X_s)\,ds + \int_0^t \langle\frac{\partial f}{\partial\eta}(s,X_s),dX_s\rangle_V \\ &+ \int_0^t \frac{\partial^2 f}{\partial\eta^2}(s,X_s)(\Theta_s)\,ds, \end{aligned}$$

where $\Theta_s(\alpha,\beta) = \frac{1}{2}((u\Lambda)_sX_s(\alpha,\beta) + (u\Lambda)_sX_s(\beta,\alpha))$, and $(u\Lambda)_sX_s : \Omega\times J^2\to L(R^m;R^m)$ is the $m\times m$ matrix-valued process defined by

$$(u\Lambda)_sX_s(\alpha,\beta) = I_{\{0\le s+\alpha\wedge\beta\}}u(s+\alpha)D_{s+\alpha}X(s+\beta). \tag{8.6}$$

Note that the right hand side of (8.6) is equal to

$$I_{\{0\le s+\alpha\wedge\beta\}}u(s+\alpha)[u'(s+\alpha)I_{\{0\le s+\alpha\le s+\beta\}} + \int_0^{s+\beta} D_{s+\alpha}u(r)\,dW(r) + \int_0^{s+\beta} D_{s+\alpha}v(r)\,dr].$$

Proof For simplicity, we assume η is absolutely continuous and define $v(s) := \eta'(s)$ and $u(s) = 0$ for $s\in J$. We will identify a linear operator with its kernel (if it has a kernel) in the proof. We also write the $m\times d$ matrix u as $(u^{ij})_{m\times d}$.

By localization technique (c.f. [17] P 45), we can assume f, f' and f'' are uniformly continuous on $T\times V$, $u^j\in\mathbb{L}^{2,4}_d$, and $v\in\mathbb{L}^{1,4}_m$. Suppose $\Pi_k : -r = s_0 < s_1 < \cdots < s_k = 0$ is a family of partitions of $[-r,0]$, with $||\Pi_k||\to 0$ as $k\to\infty$, $J_i := I_{(s_{i-1},s_i]}$ and $\Delta_i := s_i - s_{i-1}$, for $i = 1,\ldots,k$. Write $X(t) = (X^1(t),\ldots,X^m(t))'$.

If $1\le j\le m$, then $\langle X_t^j, J_i\rangle = \int_{s_{i-1}}^{s_i} X^j(t+s)\,ds = \int_{t+s_{i-1}}^{t+s_i} X^j(s)\,ds$. It follows that

$$\begin{aligned} \frac{d}{dt}\langle X_t^j, J_i\rangle &= X^j(t+s_i) - X^j(t+s_{i-1}) \\ \frac{d}{dt}I_k\circ Q_{\underline{s}^k}(X_t^j) &= \sum_{i=1}^k \frac{1}{\Delta_i}J_i(X^j(t+s_i) - X^j(t+s_{i-1})). \end{aligned}$$

Denote by $\langle \cdot, \cdot \rangle$ the inner product of the Hilbert space $L^2(J; R)$, u^j the j-th row vector of u, and v^j the j-th entry of the vector v. Let

$$\hat{U}^j(t) = \int_0^t u^j(s) \cdot dW(s), \text{ and } \hat{V}^j(t) = \begin{cases} \eta(0) + \int_0^t v^j(s)\, ds, & t > 0 \\ \eta(t), & -r \le t \le 0, \end{cases} \tag{8.7}$$

$\hat{U} = (\hat{U}^1, \ldots, \hat{U}^m)'$ and $\hat{V} = (\hat{V}^1, \ldots, \hat{V}^m)'$. Applying the chain rule to the function $f(t, I_k \circ Q_{\underline{s}^k}(X_t))$, we have

$$\begin{aligned} & f(t, I_k \circ Q_{\underline{s}^k}(X_t)) - f(0, I_k \circ Q_{\underline{s}^k}(\eta)) - \int_0^t \frac{\partial f}{\partial s}(s, I_k \circ Q_{\underline{s}^k}(X_s))\, ds \\ = & \sum_{i=1}^k \sum_{j=1}^m \int_0^t \frac{1}{\Delta_i} \langle f_j(s, I_k \circ Q_{\underline{s}^k}(X_s)), J_i \rangle (X^j(s+s_i) - X^j(s+s_{i-1}))\, ds \\ = & \sum_{i,j} \int_0^t \frac{1}{\Delta_i} \langle f_j(s, I_k \circ Q_{\underline{s}^k}(X_s)), J_i \rangle (\hat{U}^j(s+s_i) - \hat{U}^j(s+s_{i-1}))\, ds \\ + & \sum_{i,j} \int_0^t \frac{1}{\Delta_i} \langle f_j(s, I_k \circ Q_{\underline{s}^k}(X_s)), J_i \rangle (\hat{V}^j(s+s_i) - \hat{V}^j(s+s_{i-1}))\, ds \\ = & \psi_1 + \psi_2. \end{aligned}$$

By Lemma 5.1 and the continuity of f',

$$\begin{aligned} \psi_2 &= \sum_{i,j} \int_0^t \frac{1}{\Delta_i} \langle f_j(s, I_k \circ Q_{\underline{s}^k}(X_s)), J_i \rangle \int_{s_{i-1}}^{s_i} v_s^j(r)\, dr\, ds \\ &= \int_0^t \langle \frac{\partial f}{\partial \eta}(s, I_k \circ Q_{\underline{s}^k}(X_s)), I_k \circ Q_{\underline{s}^k}(v_s) \rangle_V\, ds \\ &\to \int_0^t \langle \frac{\partial f}{\partial \eta}(s, X_s), v_s \rangle_V\, ds = \int_0^t \langle \frac{\partial f}{\partial \eta}(s, X_s),\, d\hat{V}_s \rangle_V. \end{aligned}$$

By the formula for the Skorohod integral of a process multiplied by a random variable ([16] Theorem 3.2),

$$\begin{aligned} \psi_1 &= \sum_{i,j} \int_0^t \frac{1}{\Delta_i} \int_{s+s_{i-1}}^{s+s_i} \langle f_j(s, I_k \circ Q_{\underline{s}^k}(X_s)), J_i \rangle u^j(r)\, dW(r)\, ds \\ &+ \sum_{i,j} \int_0^t \frac{1}{\Delta_i} \int_{s+s_{i-1}}^{s+s_i} u^j(r) \cdot D_r \langle f_j(s, I_k \circ Q_{\underline{s}^k}(X_s)), J_i \rangle\, dr\, ds \\ &= \psi_{11} + \psi_{12}. \end{aligned}$$

By the stochastic Fubini Theorem (Lemma 4.2),

$$\psi_{11} = \sum_{i,j} \int_0^t \frac{1}{\Delta_i} \int_0^t \langle f_j(s, I_k \circ Q_{\underline{s}^k}(X_s)), J_i \rangle I_{(s_{i-1}, s_i]}(\alpha - s) u^j(\alpha)\, ds\, dW(\alpha). \tag{8.8}$$

Let $g_1(\alpha) = (g_1^1(\alpha), \ldots, g_1^m(\alpha))$, $\alpha \in T = [0, a]$, where

$$g_1^j(\alpha) := \sum_{i=1}^k \frac{1}{\Delta_i} \int_0^t \langle f_j(s, I_k \circ Q_{\underline{s}^k}(X_s)), J_i \rangle I_{(s_{i-1}, s_i]}(\alpha - s)\, ds. \tag{8.9}$$

We will show that

$$g_1(\alpha)u(\alpha) \to \Gamma_H^*(\frac{\partial f}{\partial \eta}(\cdot, X_\cdot))(\alpha)u(\alpha) \tag{8.10}$$

in $\mathbb{L}_d^{1,2}$ as $k \to \infty$, where Γ_H^* is defined by (4.6), and

$$\Gamma_H^*(\frac{\partial f}{\partial \eta}(\cdot, X_\cdot))(t) = \int_{-r}^0 \frac{\partial f}{\partial \eta}(t-\beta, X_{t-\beta})(\beta) I_{[t-a,t]}(\beta)\, d\beta, \quad t \in [0,a]. \tag{8.11}$$

Then by the continuity of δ, we have

$$\psi_{11} \to \int_0^t \Gamma_H^*(\frac{\partial f}{\partial \eta}(\cdot, X_\cdot))(\alpha)u(\alpha)\, dW(\alpha) = \int_0^t \langle \frac{\partial f}{\partial \eta}(s, X_s), dU_s\rangle_V. \tag{8.12}$$

Set $g_2(\alpha) = (g_2^1(\alpha), \ldots, g_2^m(\alpha))$, $\alpha \in T = [0,a]$, where

$$g_2^j(\alpha) := \sum_{i=1}^k \frac{1}{\Delta_i} \int_0^t \langle f_j(s, X_s), J_i\rangle I_{(s_{i-1}, s_i]}(\alpha - s)\, ds. \tag{8.13}$$

First we will show that $g_1(\alpha) - g_2(\alpha) \to 0$ in $\mathbb{L}_m^{1,4}$ as $k \to \infty$. By Definition (5.4) and Lemma 5.1,

$$\begin{aligned}
& E(\int_0^a |g_1^j(\alpha) - g_2^j(\alpha)|^2\, d\alpha)^2 \\
\leq\ & t^2 E(\sum_{i=1}^k \int_0^a \int_0^t (\frac{1}{\Delta_i}\langle f_j(s, I_k \circ Q_{\underline{s}^k}(X_s)) - f_j(s, X_s), J_i\rangle I_{(s_{i-1}, s_i]}(\alpha - s))^2\, ds\, d\alpha)^2 \\
=\ & t^2 E(\sum_{i=1}^k \int_0^t \frac{1}{\Delta_i}\langle f_j(s, I_k \circ Q_{\underline{s}^k}(X_s)) - f_j(s, X_s), J_i\rangle^2\, ds)^2 \\
=\ & t^2 E(\int_0^t \langle I_k \circ Q_{\underline{s}^k}(f_j(s, I_k \circ Q_{\underline{s}^k}(X_s)) - f_j(s, X_s)), \\
& f_j(s, I_k \circ Q_{\underline{s}^k}(X_s)) - f_j(s, X_s)\rangle\, ds)^2 \\
\leq\ & t^2 E(\int_0^t ||f_j(s, I_k \circ Q_{\underline{s}^k}(X_s)) - f_j(s, X_s)||^2\, ds)^2 \to 0 \text{ as } k \to \infty.
\end{aligned}$$

Similarly,

$$\begin{aligned}
& E(\int_0^a \int_0^a |D_u^i(g_1^j(\alpha) - g_2^j(\alpha))|^2\, d\alpha\, du)^2 \\
\leq\ & t^2 E(\int_0^a \int_0^t ||D_u^i(f_j(s, I_k \circ Q_{\underline{s}^k}(X_s)) - f_j(s, X_s))||^2\, ds\, du)^2 \\
=\ & t^2 E(\sum_{j'=1}^m \int_0^a \int_0^t ||f_{jj'}(s, I_k \circ Q_{\underline{s}^k}(X_s))(I_k \circ Q_{\underline{s}^k}(D_u^i X_s^{j'})) \\
-\ & f_{jj'}(s, X_s)(D_u^i X_s^{j'})||^2\, ds\, du)^2 \to 0 \text{ as } k \to \infty.
\end{aligned}$$

Thus $g_1 - g_2 \to 0$ in $\mathbb{L}_d^{1,4}$ as $k \to \infty$. We can rewrite g_2^j as

$$g_2^j(\alpha) = \sum_{i=1}^k \frac{1}{\Delta_i} \int_{-r}^0 \langle f_j(\alpha - \beta, X_{\alpha-\beta}), J_i\rangle I_{(s_{i-1}, s_i]}(\beta) H_{\alpha - t}(\beta)\, d\beta, \tag{8.14}$$

where $H_v(s)$ is the Heaviside-type function defined by (7.3).

By Theorem 7.5,

$$g_2 \to \Gamma_H^*(\frac{\partial f}{\partial \eta}(\cdot, X_\cdot)) \text{ in } \mathbb{L}_m^{1,4} \text{ as } k \to \infty. \tag{8.15}$$

Since for all $1 \le j \le m$, $u^j \in \mathbb{L}_d^{2,4}$, By Holder's Inequality, it follows that

$$g_2(t)u(t) \to \Gamma_H^*(f_2(\cdot, X_\cdot))(t)u(t) \text{ in } \mathbb{L}_d^{1,2} \text{ as } k \to \infty. \tag{8.16}$$

Now let us consider ψ_{12}. Set $\Theta_s(\alpha,\beta) := \frac{1}{2}((u\Lambda)_s X_s(\alpha,\beta) + (u\Lambda)_s X_s(\beta,\alpha))$, where $(u\Lambda)_s X_s(\alpha,\beta)$ is the $m \times m$ matrix-valued process defined by

$$\begin{aligned}(u\Lambda)_s X_s(\alpha,\beta) \quad &:= \quad u(s+\alpha)D_{s+\alpha}X(s+\beta) = u(s+\alpha)[u'(s+\alpha)I_{\{0\le s+\alpha\le s+\beta\}} \\ &+ \quad \int_0^{s+\beta} D_{s+\alpha}u(r)\,dW(r) + \int_0^{s+\beta} D_{s+\alpha}v(r)\,dr].\end{aligned}$$

Write $\Theta_s = (\Theta_s^{ij})_{m\times m}$. Applying Lemma 6.1 and Proposition 6.2, we have

$$\begin{aligned}\psi_{12} \quad &= \quad \sum_{j=1}^{d}\sum_{i_1,i_2=1}^{k}\sum_{j_1,j_2=1}^{m}[\int_0^t \frac{1}{\Delta_{i_1}\Delta_{i_2}} f_{j_1j_2}(s, I_k\circ Q_{\underline{s}^k}(X_s))(J_{i_1}\otimes J_{i_2}) \\ &\times \quad \int_{s_{i_1-1}}^{s_{i_1}}\int_{s_{i_2-1}}^{s_{i_2}} D^j_{s+\alpha}X^{j_2}(s+\beta)u^{j_1j}(s+\alpha)\,d\alpha\,d\beta\ ,ds].\end{aligned}$$

By the commutativity of the operators D and δ ([16] Proposition 3.4),

$$\begin{aligned}D^j_{s+\alpha}X^{j_2}(s+\beta) \quad &= \quad u^{j_2j}(s+\alpha)I_{\{0\le s+\alpha\le s+\beta\}} \\ &+ \quad \int_0^{s+\beta} D^j_{s+\alpha}u^{j_2}(r)\,dW(r) + \int_0^{s+\beta} D^j_{s+\alpha}v^{j_2}(r)\,dr.\end{aligned}$$

Denote by $Tr(A)$ the trace of a square matrix and E_{ij} the $m \times m$ matrix whose entries are 0 except the entry at position (i,j) is 1. Since $\frac{\partial^2 f}{\partial \eta^2}(s,\eta)$ is symmetric for all $\eta \in V$, applying Lemma 5.1 (under Hypotheses 8.A), we have

$$\begin{aligned}\psi_{12} \quad &= \quad \sum_{j_1,j_2=1}^{m}\int_0^t f_{j_1j_2}(s, I_k\circ Q_{\underline{s}^k}(X_s))(I_{k^2}\circ Q_{\underline{s}^{k^2}}(\Theta_s^{j_1j_2}))\,ds \\ &= \quad \sum_{j_1,j_2=1}^{m}\int_0^t f_{j_1j_2}(s, I_k\circ Q_{\underline{s}^k}(X_s))(I_{k^2}\circ Q_{\underline{s}^{k^2}}(Tr(\Theta_s E_{j_2j_1})))\,ds \\ &= \quad \int_0^t \frac{\partial^2 f}{\partial \eta^2}(s, X_s)(I_{k^2}\circ Q_{\underline{s}^{k^2}}(\Theta_s))\,ds \\ &+ \quad \int_0^t (\frac{\partial^2 f}{\partial \eta^2}(s, I_k\circ Q_{\underline{s}^k}(X_s)) - \frac{\partial^2 f}{\partial \eta^2}(s, X_s))(I_{k^2}\circ Q_{\underline{s}^{k^2}}(\Theta_s))\,ds \\ &\to \quad \int_0^t \frac{\partial^2 f}{\partial \eta^2}(s, X_s)(\Theta_s)\,ds \text{ a.s. as } k\to\infty. \quad \blacksquare\end{aligned}$$

Corollary 8.4 *Theorem 8.3 still holds if Hypotheses 8.A is replaced by Hypotheses 8.B.*

Proof

We only need to re-compute the term ψ_{12} in the proof of Theorem 8.3. By Lemma 8.2, for each s, $(u\Lambda)_s X_s$ belongs to $V\hat{\otimes}_1 V$ almost surely. We can easily finish the proof by applying Lemma A.3 (in Appendix A) under Hypotheses 8.B. ∎

Note 8.5 *When X is brownian motion, the left hand side of equation (8.6) has a simple form:*

$$(u\Lambda)_s X_s(\alpha,\beta) = I_{\{0\le s+\alpha\}}(\alpha) I_{\{0\le s+\beta\}}(\beta).$$

We can easily see that it is of trace class.

Next we shall extend Itô's formula to functions of the form $f(t, X_t, X(t))$ and use it to study the weak infinitesimal generator of SFDE's and derive the Feynman-Kac formula in the next section.

Theorem 8.6 *Suppose $f = f(t,\eta,\vec{x}) \in C_b^1(T\times V\times R^m)$ has bounded continuous second order partial derivative with respect to $(\eta,\vec{x})$ (i.e., $\frac{\partial^2 f}{\partial(\eta,\vec{x})^2}\in C_b(T\times V)$), under the same hypotheses as Theorem 8.3 (or Corollary 8.4), we have*

$$\begin{aligned}
& f(t,X_t,X(t)) - f(0,X_0,X(0)) \\
= & \int_0^t \frac{\partial f}{\partial s}(s,X_s,X(s))\,ds + \int_0^t \langle \frac{\partial f}{\partial \eta}(s,X_s,X(s)), dX_s\rangle_V \\
+ & \int_0^t \frac{\partial f}{\partial \vec{x}}(s,X_s,X(s))\,dX(s) + \int_0^t \frac{\partial^2 f}{\partial \eta^2}(s,X_s,X(s))(\Theta_s)\,ds \\
+ & \int_0^t \frac{\partial^2 f}{\partial\eta\partial\vec{x}}(s,X_s,X(s))[(u\Lambda)_s X(s)]\,ds + \int_0^t \frac{\partial^2 f}{\partial\vec{x}\partial\eta}(s,X_s,X(s))[u(s)D_s X_s]\,ds \\
+ & \frac{1}{2}\sum_{i=1}^d \int_0^t \frac{\partial^2 f}{\partial \vec{x}^2}(s,X_s,X(s))[\bigtriangledown_+^i X)(s)\otimes u^{\cdot i}(s)]\,ds,
\end{aligned}$$

where Θ_s is defined in the statement of Theorem 8.3, $u^{\cdot i}$ is the $i-$th column vector of the $m\times d$ matrix u,

$$(\bigtriangledown_+^i X)(s) = \lim_{\epsilon\downarrow 0}(D_t^i X(t+\epsilon) + D_t^i X(t-\epsilon)),\ 1\le i\le d,\ \ (\text{in } L^2(\Omega;R^m)), \tag{8.17}$$

$(u\Lambda)_s X(s): \Omega\times J\to L(R^m;R^m)$ and $u(s)D_s X_s: \Omega\times J \to L(R^m;R^m)$ are $m\times m$ matrix-valued processes defined by

$$\begin{cases} (u\Lambda)_s X(s)(\alpha) := u(s+\alpha)D_{s+\alpha}X(s) I_{\{s+\alpha\ge 0\}} \\ u(s)D_s X_s(\alpha) := u(s)D_s X(s+\alpha) I_{\{s+\alpha\ge 0\}}. \end{cases} \tag{8.18}$$

Proof We will use the same notations as in the proof of Theorem 8.3. For simplicity, we assume $m = d = 1$, η is absolutely continuous and define $v(s) := \eta'(s)$ and $u(s) = 0$ for $s\in J$. We will identify a linear operator with its kernel (if it has a kernel) in the proof.

By localization technique (c.f. [17] P 45), we can assume f, f' and f'' are uniformly continuous on $T\times V$, $u\in\mathbb{L}^{2,4}$ and $v\in\mathbb{L}^{1,4}$. Define $V_i(s):=\frac{1}{\Delta_i}\langle X(t),J_i\rangle_V$. Since

$$\frac{d}{dt}\langle X_t,J_i\rangle_V = X(t+s_i)-X(t+s_{i-1}), \tag{8.19}$$

applying the multi-dimensional Itô formula due to Nualart and Pardoux ([16] Theorem 6.4) to derive the chain rule for $f^k(t,Q_{\underline{s}^k}(X_t),X(t))$, we have

$$\begin{aligned}
& f^k(t,Q_{\underline{s}^k}(X_t),X(t)) - f^k(0,Q_{\underline{s}^k}(X_0),X(0)) - \int_0^t \frac{\partial f^k}{\partial s}(s,I_k\circ Q_{\underline{s}^k}(X_s),X(s))\,ds \\
= \; & \sum_{i=1}^k \int_0^t \frac{\partial f^k}{\partial x_i}(s,Q_{\underline{s}^k}(X_s),X(s))\,dV_i(s) + \int_0^t \frac{\partial f^k}{\partial x_{k+1}}(s,Q_{\underline{s}^k}(X_s),X(s))u(s)\,dW(s) \\
+ \; & \frac{1}{2}\int_0^t \frac{\partial^2 f^k}{\partial x_i^2}(s,Q_{\underline{s}^k}(X_s),X(s))(\bigtriangledown_+X)(s)u(s)\,ds \\
+ \; & \int_0^t \sum_{i=1}^k \frac{\partial^2 f^k}{\partial x_{k+1}\partial x_i}(s,Q_{\underline{s}^k}(X_s),X(s))D_sV_i(s)u(s)\,ds \\
= \; & \int_0^t \sum_{i=1}^k \frac{1}{\Delta_i}\langle f_1(s,I_k\circ Q_{\underline{s}^k}(X_s),X(s)),J_i\rangle_V (X(s+s_i)-X(s+s_{i-1}))\,ds \\
+ \; & \int_0^t f_2(s,I_k\circ Q_{\underline{s}^k}(X_s),X(s))u(s)\,dW(s) \\
+ \; & \frac{1}{2}\int_0^t f_{22}(s,I_k\circ Q_{\underline{s}^k}(X_s),X(s))(\bigtriangledown_+X)(s)u(s)\,ds \\
+ \; & \int_0^t \sum_{i=1}^k \frac{1}{\Delta_i}\langle f_{21}(s,I_k\circ Q_{\underline{s}^k}(X_s),X(s)),J_i\rangle_V \langle D_sX_s,J_i\rangle_V\,ds \\
= \; & I_1(k)+I_2(k)+I_3(k)+I_4(k)
\end{aligned}$$

It is easy to see that

$$\lim_{k\to\infty} I_3(k) = \frac{1}{2}\int_0^t f_{22}(s,X_s,X(s))(\bigtriangledown_+X)(s)u(s)\,ds \quad \text{a.s.} \tag{8.20}$$

Similar argument as the proof of the convergence result (8.15) shows that

$$\lim_{k\to\infty} f_2(s,I_k\circ Q_{\underline{s}^k}(X_s),X(s)) = f_2(s,X_s,X(s)) \tag{8.21}$$

in $\mathbb{L}^{1,4}$. Thus

$$\begin{cases} \lim_{k\to\infty} f_2(s,I_k\circ Q_{\underline{s}^k}(X_s),X(s))u(s) = f_2(s,X_s,X(s))u(s) \text{ in } \mathbb{L}^{1,2}, \\ \lim_{k\to\infty} I_2(k) = \int_0^t f_2(s,X_s,X(s))u(s)\,dW(s) \text{ in } L^2(\Omega) \end{cases} \tag{8.22}$$

by continuity of the operator $\delta:\mathbb{L}^{1,2}\to L^2(\Omega)$. Since

$$I_4(k) = \int_0^t \langle I_k\circ Q_{\underline{s}^k}(f_{21}(s,I_k\circ Q_{\underline{s}^k}(X_s),X(s))),D_sX_s\rangle_V u(s)\,ds, \tag{8.23}$$

it follows that

$$\lim_{k\to\infty} I_4(k) = \int_0^t \langle f_{21}(s, X_s, X(s)), D_s X_s\rangle_V u(s)\, ds. \tag{8.24}$$

Now let us consider $I_1(k)$, since

$$\begin{aligned} I_1(k) &= \int_0^t \sum_{i=1}^k \frac{1}{\Delta_i} \langle f_1(s, I_k \circ Q_{\underline{s}^k}(X_s), X(s)), J_i\rangle_V \int_{s+s_{i-1}}^{s+s_i} u(r)\, dW(r)\, ds \\ &+ \int_0^t \sum_{i=1}^k \frac{1}{\Delta_i} \langle f_1(s, I_k \circ Q_{\underline{s}^k}(X_s), X(s)), J_i\rangle_V \int_{s+s_{i-1}}^{s+s_i} v(r)\, dr\, ds \\ &= I_{11}(k) + I_{12}(k). \end{aligned}$$

it follows that

$$\begin{aligned} I_{12}(k) &= \int_0^t \langle I_k \circ Q_{\underline{s}^k}(f_1(s, I_k \circ Q_{\underline{s}^k}(X_s), X(s))), v_s\rangle_V\, ds \\ &\to \int_0^t \langle f_1(s, X_s, X(s)), v_s\rangle_V\, ds = \int_0^t \langle f_1(s, X_s, X(s)),\, dV_s\rangle_V \text{ a.s.} \end{aligned}$$

as $k \to \infty$. By the formula for the Skorohod integral of a process multiplied by a random variable ([16] Theorem 3.2),

$$\begin{aligned} I_{11}(k) &= \int_0^t \sum_{i=1}^k \frac{1}{\Delta_i} \int_{s+s_{i-1}}^{s+s_i} \langle f_1(s, I_k \circ Q_{\underline{s}^k}(X_s), X(s)), J_i\rangle_V u(r)\, dW(r)\, ds \\ &+ \int_0^t \sum_{i=1}^k \frac{1}{\Delta_i} \int_{s+s_{i-1}}^{s+s_i} D_r \langle f_1(s, I_k \circ Q_{\underline{s}^k}(X_s), X(s)), J_i\rangle_V u(r)\, dr\, ds \\ &= I_{111}(k) + I_{112}(k). \end{aligned}$$

Analogous to the argument showing (8.12), we can show that

$$\lim_{k\to\infty} I_{111}(k) = \int_0^t \int_0^t \langle f_1(s, X_s, X(s)),\, dU_s\rangle_V \tag{8.25}$$

in $L^2(\Omega)$. By Proposition 6.2, we have

$$\begin{aligned} I_{112}(k) &= \int_0^t \sum_{i,j=1}^k \frac{1}{\Delta_i \Delta_j} \langle f_{11}(s, I_k \circ Q_{\underline{s}^k}(X_s), X(s)) J_i, J_j\rangle_V \\ &\times \int_{s_{i-1}}^{s_i} \int_{s_{j-1}}^{s_j} D_{s+\beta} X(s+\alpha) u(s+\beta)\, d\alpha\, d\beta\, ds \\ &+ \int_0^t \sum_{i=1}^k \frac{1}{\Delta_i} \langle f_{12}(s, I_k \circ Q_{\underline{s}^k}(X_s), X(s)), J_i\rangle_V \int_{s_{i-1}}^{s_i} D_{s+\beta} X(s) u(s+\beta)\, d\beta\, ds \\ &= I_{1121}(k) + I_{1122}(k). \end{aligned}$$

Analogous to the argument deriving the term ψ_{12} in the proof of Theorem 8.3, we can show that

$$\lim_{k\to\infty} I_{1121}(k) = \int_0^t f_{11}(s, X_s, X(s))(\Theta_s)\, ds. \tag{8.26}$$

Finally, we have

$$\begin{aligned} I_{1122}(k) &= \int_0^t \langle I_k \circ Q_{\underline{s}^k}(f_{12}(s, I_k \circ Q_{\underline{s}^k}(X_s), X(s))), (u\Lambda)_s X(s)\rangle_V \, ds \\ &\to \int_0^t \langle f_{12}(s, X_s, X(s)), (u\Lambda)_s X(s)\rangle_V \, ds \text{ a.s.} \end{aligned}$$

as $k \to \infty$. The Itô formula follows. ∎

As an application of Corollary 8.4, let us look at a simple example.

Example 8.7 *$f(\eta) = ||\eta||_V^2$, $\{W(t)\}$ is the standard Brownian motion, by (5.6), we have*

$$||W_t||_V^2 = 2\int_0^t \langle W_s,\, dW_s\rangle_V + \int_0^t f''(W_s)(\Theta_s)\, ds.$$

Since $f''(\eta)(\zeta_1, \zeta_2) = 2\langle \zeta_1, \zeta_2\rangle_V$, and

$$\begin{aligned} \Theta_s(\alpha, \beta) &= \tfrac{1}{2}(I_{\{\alpha+s\ge 0\}} I_{\{\alpha\le\beta\}} + I_{\{\beta+s\ge 0\}} I_{\{\beta\le\alpha\}}) \\ &= \tfrac{1}{2} I_{\{\alpha\wedge\beta\ge -s\}} = \tfrac{1}{2} I_{\{\alpha\ge -s\}} I_{\{\beta\ge -s\}}, \end{aligned}$$

it follows that

$$\begin{aligned} ||W_t||_V^2 &= 2\int_0^t \int_{-r}^0 W(v) I_{\{v\le t+s\}}\, ds\, dW(v) + \int_0^t \int_{-r}^0 I_{\{\alpha\ge -s\}}\, d\alpha\, ds \\ &= 2\int_{-r}^0 \int_0^{t+s} W(v)\, dW(v)\, ds + \int_0^t s\wedge r\, ds. \end{aligned}$$

Since

$$\int_0^t s\wedge r\, ds = \begin{cases} \int_0^t s\, ds, & t \le r \\ \int_0^t s\, ds + \int_r^t r\, ds, & t > r, \end{cases} \tag{8.27}$$

One may check it agrees with the formula

$$2\int_0^t W(s)\, dW(s) = W^2(t) - t. \qquad ∎$$

9 The weak infinitesimal generator, the Feynman-Kac formula, and the Black-Scholes PDE for SFDEs

As an application of Itô's formula for segment processes, we shall study the weak infinitesimal generator of SFDE's and derive the Feynman-Kac formula in this section. Suppose $J = [-r, 0]$, $T = [0, a]$, $V = L^2(J; R^m)$, $H = L^2(T; R^m)$ and $\{W(t) : t \ge 0\}$ is a d-dimensional standard Brownian motion, we denote by $W^i(t)$ the ith component of $W(t)$. Denote by $e_j = (0, \ldots, 1_j, \ldots, 0)$, then $\{e_1, \ldots, e_d\}$ is a basis of R^d. For $(\eta, \vec{x}) \in V \times R^m$, let

$$\widehat{\eta^x}(t) := \begin{cases} \vec{x}, & t \in [0, a] \\ \eta(t), & t \in [-r, 0). \end{cases} \tag{9.1}$$

Then for each $s \in J, t \in T$,

$$\widehat{\eta_t^x}(s) = \widehat{\eta^x}(t+s) = \begin{cases} \vec{x}, & t+s \geq 0 \\ \eta(t+s), & t+s < 0. \end{cases} \tag{9.2}$$

Denote by $\{S_t : t \in [0,a]\}$ the weakly continuous contraction semigroup of the shift operators defined on $C_b(V \times R^m)$ (c.f. [1] and [13] Chapter 4) by

$$S_t(\phi)(\eta, \vec{x}) := \phi(\widehat{\eta_t^x}, \vec{x}) \text{ for } \phi \in C_b(V \times R^m).$$

Denote by C_b^0 the set of all $\phi \in C_b(V \times R^m)$ such that $S_t\phi$ is strongly continuous, S the weak infinitesimal generator of S_t, and $D(S) \subset C_b^0$ the domain of S.

If $f \in C^2(V \times R^m)$, denote by f_1 the Fréchet derivative with respect to the first variable, f_2 the derivative with respect to the second variable, and f_{22} the derivative of f_2 with respect to the second variable. Assume

$$\begin{cases} G : V \times R^m \to L(R^d; R^m) \\ H : V \times R^m \to R^m \end{cases}$$

satisfying Lipshitz and linear growth condition. Let us consider a class of autonomous stochastic functional differential equations (SFDE's) of type (c.f. [1], [13] P 226)

$$X(t) = \begin{cases} \vec{x} + \int_0^t G(X_s, X(s))\, dW(s) + \int_0^t H(X_s, X(s))\, ds, & t \geq 0 \\ \eta(t), & -r \leq t < 0, \end{cases} \tag{9.3}$$

where $\eta \in V$ and $\vec{x} \in R^m$. We will write $H = (H^1, \ldots, H^m)$ and $G = (G^1, \ldots, G^m)$, where $G^i = (G^{i1}, \ldots, G^{id})$.

Under linear growth and Lipschitz condition the SFDE (9.3) has a strong unique solution(c.f. [13], pp. 226 – 228, [1] Chapter 2).

There is a *weakly* continuous contraction semigroup $\{P_t : P_t\psi(\eta, \zeta) := E\psi(X_t, X(t))\}$ associated with the solution ([1], section 3.3). The semigroup $\{P_t : t \geq 0\}$ is strongly continuous if and only if the delay r is zero (cf. [13], for the case of the state space $C(J, R^d)$.)

Denote by A the weak infinitesimal generator of $\{P_t\}$ (c.f. [1] Chapter 4) and by $D(A)$ its domain. The class of quasitame function is dense in the domain of A. The action of A on quasitame function is well studied (c.f. [13] Section 4.4, [1] Section 5.2).

Remarks 9.1 *Let us consider the SFDE (9.3), suppose $H \in C_b^1(V \times R^m; R^m)$ and $G^i \in C_b^2(V \times R^m; R^d)$, then one may check that $G^i(X_t, X(t))$ belong to $\mathbb{L}_d^{2,4}$ and $H(X_t, X(t))$ belong to $\mathbb{L}_d^{1,4}$ respectively.*

Remarks 9.2 *Suppose we consider the SFDE (9.3). Theorem 8.3 and 8.6 hold for all initial values* $\eta \in C([-r,0])$ *if the derivative* $\frac{\partial f}{\partial \eta}(t,\eta,\vec{x})$ *of* $f(t,\eta,\vec{x})$ *belongs to* $C^1([-r,0])$ *for all* $(t,\eta,\vec{x}) \in T \times V \times R^m$.

Proof Suppose $\psi \in C([-r,0])$, $\frac{\partial f}{\partial \psi}(s,\psi,\psi(0))$ belongs to $C^1([-r,0])$ and $\eta \in C([-r,0])$ is absolute continuous. By the change of integration formula,

$$\begin{aligned} & \int_{-r}^{0} \frac{\partial f}{\partial \psi}(s,\psi,\psi(0))(\alpha)\eta'(\alpha)\,d\alpha \\ = & \frac{\partial f}{\partial \psi}(s,\psi,\psi(0))(0)\eta(0) - \frac{\partial f}{\partial \psi}(s,\psi,\psi(-r))(-r)\eta(-r) \\ - & \int_{-r}^{0} \frac{\partial^2 f}{\partial \alpha \partial \psi}(s,\psi,\psi(0))(\alpha)\eta(\alpha)\,d\alpha. \end{aligned}$$

Now for any $\eta \in C([-r,0])$, let η_n be a sequence of absolute continuous functions that converge to η in $C([-r,0])$. We can define the integral

$$\int_{-r}^{0} \frac{\partial f}{\partial \psi}(s,\psi,\psi(0))(\alpha)\,d\eta(\alpha)$$

as the limit of

$$\begin{aligned} & \frac{\partial f}{\partial \psi}(s,\psi,\psi(0))(0)\eta_n(0) - \frac{\partial f}{\partial \psi}(s,\psi,\psi(-r))(-r)\eta_n(-r) \\ - & \int_{-r}^{0} \frac{\partial^2 f}{\partial \alpha \partial \psi}(s,\psi,\psi(0))(\alpha)\eta_n(\alpha)\,d\alpha. \end{aligned}$$

Let $X^n(t)$ be the solution of the SFDE (9.3) with initial value η_n. We can see Theorem 8.3 and 8.6 hold by letting $n \to \infty$. ∎

We now derive the generator A associated with the SFDE (9.3). Then we can express the operator S (hence A) as sum of differential operators by applying Itô's formula (Theorem 8.6),

just as the ordinary stochastic differential equation case.

Lemma 9.3 *Suppose* $f \in C_b^2(V \times R^m)$ *belongs to the domain of A,* $G^i \in C_b^2(V \times R^m; R^d)$ *and* $H \in C_b^1(V \times R^m; R^m)$. *Assume* $\eta \in V$, $\vec{x} \in R^m$, *and* $\{X(t)\}$ *is the solution of SFDE (9.3). Let* $\{e_j : j = 1, \ldots, d\}$ *be a normalized basis of* R^d. *Then for all* $0 \le t < a$,

$$\begin{aligned} Af(X_t, X(t)) &= Sf(X_t, X(t)) + f_2(X_t, X(t))H(X_t, X(t)) \\ &+ \frac{1}{2}\sum_{j=1}^{d} f_{22}(X_t, X(t))[(G(X_t, X(t))e_j) \otimes (G(X_t, X(t))e_j)]. \end{aligned}$$

Proof First we assume that η is absolutely continuous and $\eta(0) = x$. Set

$$u(t) = \begin{cases} G(X_t, X(t)), & t \ge 0 \\ 0, & -r \le t < 0, \end{cases} \quad \text{and } v(t) = \begin{cases} H(X_t, X(t)), & t \ge 0 \\ \eta'(t), & -r \le t < 0. \end{cases} \qquad (9.4)$$

Fix $0 \le t_0 < a$. Let $\phi := X_{t_0}$ and $\widehat{\phi_t^x}$ be defined by (9.2).

$$\begin{aligned}
& E(f(X_t, X(t)) - f(X_{t_0}, X(t_0))|X_{t_0}, X(t_0)) \\
= \quad & E(f(\widehat{\phi_{t_0}^x}, X(t_0)) - f(X_{t_0}, X(t_0))|X_{t_0}, X(t_0)) \\
+ \quad & E(f(\widehat{\phi_{t_0}^x}, X(t)) - f(\widehat{\phi_{t_0}^x}, X(t_0))|X_{t_0}, X(t_0)) \\
+ \quad & E(f(X_t, X(t)) - f(\widehat{\phi_{t_0}^x}, X(t))|X_{t_0}, X(t_0)) \\
= \quad & E_1 + E_2 + E_3.
\end{aligned}$$

By the definition of S, $E_1/(t-t_0)$ converges to $Sf(X_{t_0}, X(t_0))$ as $t \to t_0$. By classical Itô's formula, as $t \to t_0$, $E_2/(t-t_0)$ converges to

$$\begin{aligned}
& f_2(X_{t_0}, X(t_0))H(X_{t_0}, X(t_0)) \\
+ \quad & \frac{1}{2}\sum_{j=1}^{d} f_{22}(X_{t_0}, X(t_0))[(G(X_{t_0}, X(t_0))e_j) \otimes (G(X_{t_0}, X(t_0))e_j)].
\end{aligned}$$

By Taylor's formula,

$$f(X_t, X(t)) - f(\widehat{\phi_{t_0}^x}, X(t)) = f_1(\widehat{\phi_{t_0}^x}, X(t))(X_t - \widehat{\phi_{t_0}^x}) + R(t_0, t),$$

where

$$R(t_0, t) = \int_0^1 (1-\alpha) f_{11}(\widehat{\phi_{t_0}^x} + \alpha(X_t - \widehat{\phi_{t_0}^x}))(X_t - \widehat{\phi_{t_0}^x}, X_t - \widehat{\phi_{t_0}^x})\, du.$$

Thus $E_3/(t-t_0)$ converges to 0 as $t \to t_0$. We proved the lemma for the case the initial function η is absolutely continuous and the initial value $\vec{x}$ is $\eta(0)$.

For any $(\eta, \vec{x}) \in V \times R^m$, there exists a sequence of absolutely continuous functions $\{\eta_n\} \subset V$ such that

$$\lim_{n\to\infty} ||\eta_n - \eta||_V = 0.$$

Define a sequence of linear approximation of $(\eta, \vec{x})$ by

$$\bar{\eta}_n(s) = \begin{cases} \eta_n(s), & -r \le s \le -\frac{1}{n} \\ \vec{x} + ns(\vec{x} - \eta_n(s)), & -\frac{1}{n} < s \le 0. \end{cases} \tag{9.5}$$

Then $\bar{\eta}_n$ is absolutely continuous and

$$\lim_{n\to\infty} ||(\bar{\eta}_n, \eta_n(0)) - (\eta, \vec{x})||_{V\times R^m} = 0.$$

Let X^n be the solution of SFDE (9.3) with initial function η_n and initial value $\eta_n(0)$. By above argument, the lemma holds for each n, i.e.,

$$\begin{aligned}
& Af(X_t^n, X^n(t)) = Sf(X_t^n, X^n(t)) + f_2(X_t^n, X^n(t))H(X_t^n, X^n(t)) \\
+ \quad & \frac{1}{2}\sum_{j=1}^{d} f_{22}(X_t^n, X^n(t))[(G(X_t^n, X^n(t))e_j) \otimes (G(X_t^n, X^n(t))e_j)].
\end{aligned}$$

The proof is completed by letting $n \to \infty$. ∎

We can extend the definition of A and S to $f \in C_b^2(T \times V \times R^m)$. For $t \in T$, let $f^t(\eta, \vec{x}) := f(t, \eta, \vec{x})$. We define $Af(t, \eta, \vec{x}) := Af^t(\eta, \vec{x})$ and $Sf(t, \eta, \vec{x}) := Sf^t(\eta, \vec{x})$.

We can also prove the lemma by applying Itô's formula (Theorem 8.6) [22].

By Itô's formula (Theorem 8.6) and Lemma 9.3 we can express the operator S (hence A) as sum of differential operators. We can also extend Itô's formula for SFDE for any initial value $\eta \in V$. First we need to define the integral

$$\int_{-r}^{0} f_1(\eta, \vec{x})(\alpha)\, d\eta(\alpha),$$

where $\eta \in V$, $f \in C^1(V \times R^m)$, f_1 is the derivative of f with respect to the first variable. It turns out the above integral is the same as $Sf(\eta, \vec{x})$. To see this, we first assume η is absolute continuous and $\vec{x} = \eta(0)$.

$$\begin{aligned} Sf(\eta, \eta(0)) &= \lim_{t \to 0} (f(\widehat{\eta_t^x}, \vec{x}) - f(\eta, \vec{x}))/t \\ &= \lim_{t \to 0} \int_{-r}^{0} f_1(\tilde{\eta}, \eta(0)) \frac{\widehat{\eta_t^x}(\alpha) - \eta(\alpha)}{t}\, d\alpha \\ &= \int_{-r}^{0} f_1(\eta, \eta(0))\eta'(\alpha)\, d\alpha = \int_{-r}^{0} f_1(\eta, \eta(0))\, d\eta(\alpha), \end{aligned}$$

Let η_n be defined by (9.5). Then

$$Sf(\eta, \vec{x}) = \lim_{n \to \infty} Sf(\eta_n, \eta_n(0)).$$

We define

$$< f_1(\eta, \vec{x}),\, d\eta >_V = \int_{-r}^{0} f_1(\eta, \vec{x})(\alpha)\, d\eta(\alpha) := Sf(\eta, \vec{x}). \tag{9.6}$$

Since Sf is continuous, by (9.6), the following result holds:

$$\lim_{\phi \to \eta} < f_1(\phi, \vec{x}),\, d\phi >_V = < f_1(\eta, \vec{x}),\, d\eta >_V . \tag{9.7}$$

Corollary 9.4 *Suppose $X(t)$ is the solution of the SFDE (9.3), and $u(t)$ is defined by (9.4). Under the same assumption as Theorem 8.6, if $0 \le s \le a$,*

$$\begin{aligned} Sf(s, X_s, X(s)) &= \langle \frac{\partial f}{\partial \eta}(s, X_s, X(s)),\, dX_s \rangle_V \\ &+ \frac{\partial^2 f}{\partial \eta^2}(s, X_s, X(s))(\Theta_s) + \frac{\partial^2 f}{\partial \eta \partial \vec{x}}(s, X_s, X(s))[(u\Lambda)_s X(s)], \end{aligned}$$

where Θ_s and $(u\Lambda)_s X(s)$ are defined in the statements of Theorem 8.3 and Theorem 8.6, respectively, and $\langle \frac{\partial f}{\partial \eta}(s, X_s, X(s)),\, dX_s \rangle_V$ is defined by (9.6).

Proof Let

$$q(t) := \int_0^t \langle \frac{\partial f}{\partial \eta}(s, X_s, X(s)), dX_s \rangle_V. \tag{9.8}$$

By Itô's formula (Theorem 8.6) and Lemma 9.3, and the fact $X(t)$ is adaptive and $D_s X(s) = 0$ a.s., we only need to verify that for all $t_1 < t_2$,

$$\lim_{t_2 \downarrow t_1} E(\frac{q(t_2) - q(t_1)}{t_2 - t_1} | F_{t_1}) = \langle \frac{\partial f}{\partial \eta}(t_1, X_{t_1}, X(t_1)), dX_{t_1} \rangle_V. \tag{9.9}$$

By the stochastic Fubini Theorem (Lemma 4.2) and the definition of segment integral,

$$\begin{aligned}
q(t_2) - q(t_1) &= \int_{-r}^0 \int_{t_1}^{t_2} \frac{\partial f}{\partial \eta}(s, X_s, X(s))(\alpha)\, dX(s+\alpha)\, d\alpha \\
&= \int_{-r}^0 \int_{t_1}^{t_2} \frac{\partial f}{\partial \eta}(s, X_s, X(s))(\alpha) I_{\{s+\alpha \le t_1\}}\, dX(s+\alpha)\, d\alpha \\
&+ \int_{-r}^0 \int_{t_1}^{t_2} \frac{\partial f}{\partial \eta}(s, X_s, X(s))(\alpha) I_{\{s+\alpha > t_1\}}\, dX(s+\alpha)\, d\alpha \\
&= J_1 + J_2.
\end{aligned}$$

We can further decompose J_2:

$$\begin{aligned}
J_2 &= \int_{-r}^0 \int_{t_1}^{t_2} \frac{\partial f}{\partial \eta}(s, X_s, X(s))(\alpha) I_{\{s+\alpha > t_1\}} v(s+\alpha)\, ds\, d\alpha \\
&+ \int_{-r}^0 \int_{t_1}^{t_2} \frac{\partial f}{\partial \eta}(s, X_s, X(s))(\alpha) I_{\{s+\alpha > t_1\}} u(s+\alpha)\, dW(s+\alpha)\, d\alpha \\
&= J_{21} + J_{22}.
\end{aligned}$$

Let us consider the SFDE

$$Y(t) = \begin{cases} \vec{x} + \int_{t_1}^t G(Y_s, Y(s))\, dB(s) + \int_{t_1}^t H(Y_s, Y(s))\, ds, & t \ge t_1 \\ X(t), & -r \le t < t_1, \end{cases} \tag{9.10}$$

where the the Brownian motion B vanishes on $[-r, t_1]$. Hence $\langle \frac{\partial f}{\partial \eta}(t_1, X_{t_1}, X(t_1)), dX_{t_1} \rangle_V$ can be defined by (9.6). We can see that

$$\lim_{t_2 \downarrow t_1} E(\frac{J_1}{t_2 - t_1} | F_{t_1}) = \langle \frac{\partial f}{\partial \eta}(t_1, X_{t_1}, X(t_1)), dX_{t_1} \rangle_V, \tag{9.11}$$

$$\lim_{t_2 \downarrow t_1} E(\frac{J_{21}}{t_2 - t_1} | F_{t_1}) = 0, \tag{9.12}$$

and $E(J_{22} | F_{t_1}) = 0$. ∎

Note that if $S_{\delta_n} f(t, \eta, \vec{x}) \to S(f)(t, \eta, \vec{x})$ as $n \to \infty$ uniformly for $\eta \in V$ and $\vec{x} \in R^m$, where $\delta_n = t/n$. Then we can prove the Itô formula expressed using S (or A) directly ([22]). Next we shall derive the Feynman-Kac formula for SFDE's.

Theorem 9.5 *(The Feynman-Kac formula) Suppose g is a function defined on $V \times R^m$. If a function $f \in C_b^2(T \times V \times R^m)$ satisfies the partial functional differential equation:*

$$\begin{aligned} -\frac{\partial f}{\partial t}(t,\eta,\vec{x}) &= Sf(t,\eta,\vec{x}) + \frac{\partial f}{\partial \vec{x}}(t,\eta,\vec{x})G(\eta,\vec{x}) \\ &+ \frac{1}{2}\sum_{j=1}^{d} \frac{\partial^2 f}{\partial \vec{x}^2}(t,\eta,\vec{x}))[(G(\eta,\vec{x})e_j) \otimes (G(\eta,\vec{x})e_j)], \end{aligned}$$

with final value $f(a,\eta,\vec{x}) = g(\eta,\vec{x})$, then

$$f(t,\eta,\vec{x}) = E(g(X_a, X(a))|X_t = \eta, X(t) = \vec{x}) \qquad \forall 0 \le t \le a. \tag{9.13}$$

The reverse is also true if we assume $g \in C^2(V \times R^m)$.

Proof Without loss of generality we assume $m = 1$. If $g \in C^2(V \times R^m)$, then $f(t,\cdot,\cdot) \in C^2(V \times R^m)$ for all t and $f(\cdot,\eta,\vec{x}) \in C^1(T)$ for all $(\eta,\vec{x})$. By Itô's formula (8.17)

$$\begin{aligned} & f(t,X_t,X(t)) - f(t_0,X_{t_0},X(t_0)) \\ = & \int_{t_0}^{t} \frac{\partial f}{\partial s}(s,X_s,X(s))\,ds + \int_{t_0}^{t} \langle \frac{\partial f}{\partial \eta}(s,X_s,X(s)), dX_s \rangle_V \\ + & \int_{t_0}^{t} \frac{\partial f}{\partial \vec{x}}(s,X_s,X(s))\,dX(s) + \int_{t_0}^{t} \frac{\partial^2 f}{\partial \eta^2}(s,X_s,X(s))(\Theta_s)\,ds \\ + & \int_{t_0}^{t} \frac{\partial^2 f}{\partial \eta \partial \vec{x}}(s,X_s,X(s))[(u\Lambda)_s X(s)]\,ds + \int_{t_0}^{t} \frac{\partial^2 f}{\partial \vec{x} \partial \eta}(s,X_s,X(s))[u(s)D_s X_s]\,ds \\ + & \frac{1}{2}\sum_{i=1}^{d} \int_{t_0}^{t} \frac{\partial^2 f}{\partial \vec{x}^2}(s,X_s,X(s))[\bigtriangledown_+^i X)(s) \otimes u^{\cdot i}(s)]\,ds, \end{aligned}$$

By the property of Itô integral,

$$E(\int_{t_0}^{t} \frac{\partial f}{\partial \vec{x}}(s,X_s,X(s))u(s)\,dW(s)|X_{t_0} = \eta, X(t_0) = \vec{x}) = 0.$$

Suppose equation (9.13) is true. If $0 \le t < t_2 \le a$, by the Markov property of the segment process [14],

$$\begin{aligned} & E(f(t_2,X_{t_2},X(t_2)) - f(t,X_t,X(t))|X_t,X(t)) \\ = & E[E(g(X_a,X(a))|X_{t_2},X(t_2)) - E(g(X_a,X(a))|X_t,X(t))|X_t,X(t))] \\ = & E[E(g(X_a,X(a))|F_{t_2}) - E(g(X_a,X(a))|F_t)|F_t] = 0. \end{aligned}$$

By Corollary 9.4, for $(\eta,x) \in V \times R$, the Feynman-Kac PDE holds.

On the other hand, we need to show that f satisfies the PDE implies

$$f(t,\eta,\vec{x}) = E(g(X_a,X(a))|X_t = \eta, X(t) = \vec{x}) \qquad \forall 0 \le t \le a.$$

Fix $0 \leq t_0 < a$. Define for all $t_0 \leq t \leq a$

$$q(t) := E(f(t, X_t, X(t))|F_{t_0}). \tag{9.14}$$

We shall show that $q(t)$ is a constant a.s. on F_{t_0}. If $t_0 \leq t_1 < t_2 \leq a$,

$$q(t_2) - q(t_1) = E[E(f(t_2, X_{t_2}, X(t_2)) - f(t_1, X_{t_1}, X(t_1))|F_{t_1})|F_{t_0}]. \tag{9.15}$$

Let J_1, J_{21} and J_{22} be defined in the proof of Corrolary (9.4). Then $E(J_{22}|F_{t_1}) = 0$. It is easy to see that the following limits hold:

$$\lim_{t_2 \downarrow t_1} E(\frac{J_1}{t_2 - t_1}|F_{t_1}) = \langle \frac{\partial f}{\partial \eta}(t_1, X_{t_1}, X(t_1)), dX_{t_1} \rangle_V, \tag{9.16}$$

$$\lim_{t_2 \downarrow t_1} E(\frac{J_{21}}{t_2 - t_1}|F_{t_1}) = 0, \tag{9.17}$$

Thus the right derivative of q

$$q'_+(t_1) = \lim_{t_2 \downarrow t_1} \frac{q(t_2) - q(t_1)}{t_2 - t_1} = 0 \quad a.s. \quad on \quad F_{t_0}. \tag{9.18}$$

Thus the function q is continuous and has continuous right derivatives. By a well-known lemma ([23], P 239), q is differentiable and hence a constant. We conclude that

$$q(t_0) = q(a) = E(f(a, X_a, X(a))|F_{t_0}) = E(g(X_a, X(a))|F_{t_0}) \tag{9.19}$$

and complete the proof of the Feynman-Kac equation. ∎

Note 9.6 *It is unclear whether we can assume a weaker condition on the boundary function g than the condition $g \in C^2(V \times R^m)$ in Theorem 9.5. Unlike the finite dimensional case, it is hard to describe a "smooth" density function in an infinite dimensional space V since we can not define a Lebesgue measure on V.*

The functional PDE in Theorem 9.5 (the Feynman-Kac formula) can be solved numerically using forward finite difference scheme. On the other hand, it can be solved using Monte Carlo simulation (c.f. [7, 22]).

Because the Feynman-Kac equation is derived from a dynamic system with memory driven by finite dimensional Brownian motions, the variation of η is governed by the forward shift operator S. For example, suppose we try to solve the Feynman-Kac functional PDE using finite difference scheme, the variation of the infinite dimensional variable η is "finite-dimensional" instead of "infinite-dimensional", as defined by equation (9.2). This behavior is different than that of evolutionary equations.

As an application of the Feynman-Kac formula, we shall derive the functional Black-Scholes PDE under the SFDE setting. Suppose there is an European option on a portfolio of m assets whose prices movement follow the SFDE (9.3). The pay-off function is $g(X_a, X(a))$ when the option expires at $t = a$, where $g \in C^2(V \times R^m)$. Note that this option is path dependent if the function g is non-singular on the first variable. For simplicity, we assume interest rate is zero. The value of the option at time t is $f(t, \eta, \vec{x}) = E(g(X_a, X(a))|X_t = \eta, X(t) = \vec{x})$. From the Feynman-Kac formula, $f(t, \eta, \vec{x})$ satisfies the functional Black-Scholes PDE

$$\begin{aligned} -\frac{\partial f}{\partial t}(t,\eta,\vec{x}) &= Sf(t,\eta,\vec{x}) + \frac{\partial f}{\partial \vec{x}}(t,\eta,\vec{x})G(\eta,\vec{x}) \\ &+ \frac{1}{2}\sum_{j=1}^{d} \frac{\partial^2 f}{\partial \vec{x}^2}(t,\eta,\vec{x}))[(G(\eta,\vec{x})e_j) \otimes (G(\eta,\vec{x})e_j)], \end{aligned}$$

with final value $f(a, \eta, \vec{x}) = g(\eta, \vec{x})$.

This paper was inspired by the works of Mohammed ([13]), Nualart ([17]), Nualart and Pardoux ([16]), in stochastic functional differential equations and anticipating stochastic calculus.

Appendix A. Tensor products of Hilbert spaces.

Let H_1 and H_2 be two Hilbert spaces. We first consider the *algebraic* tensor product space of all formal finite sums represented by

$$z = \sum_{i=1}^{n} (x_i \otimes y_i), \qquad x_i \in H_1, y_i \in H_2. \tag{9.20}$$

We endow this space with the inner product

$$\langle \sum_{i=1}^{n_1} (x_i \otimes y_i), \sum_{j=1}^{n_2} (s_j \otimes t_j) \rangle = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \langle x_i, s_j \rangle \langle y_i, t_j \rangle. \tag{9.21}$$

The indices n_1 and n_2 may be taken to be same without loss of generality by adding zero entries. Note that the inner product is independent of the tensor representation. We denote this inner product space by $H_1 \otimes_2 H_2$, and the completion of this space by $H_1 \hat{\otimes}_2 H_2$ under the norm induced by the inner product. The n-fold tensor product $H_1 \hat{\otimes}_2 \ldots \hat{\otimes}_2 H_n$ is defined as the tensor product of $H_1 \hat{\otimes}_2 \ldots \hat{\otimes}_2 H_{n-1}$ and H_n.

Remark A.1 The inner product above induces a topology ($\epsilon-$topology) which is weaker than the projective topology ($\pi-$topology) on $H_1 \otimes H_2$ ([21] Section 43). If $H_1 \otimes H_2$ is endowed with the $\pi-$topology, we denote the space and its completion as $H_1 \otimes_1 H_2$ and

$H_1\hat{\otimes}_1 H_2$ respectively. If $\eta_1, \eta_2 \in H_1$, then (c.f. [21] Proposition 43.11)

$$||\eta_1 \otimes \eta_2||_{(2)} = ||\eta_1 \otimes \eta_2||_{(1)} = ||\eta_1||_{H_1}||\eta_2||_{H_1}. \tag{9.22}$$

Denote by $L_{(2)}(H)$ and $L_{(1)}(H)$ the spaces of all Hilbert-Schimdt operators and trace class operators on H, respectively. Clearly $L_{(1)}(H) \cong H\hat{\otimes}_1 H$ and $L_{(2)}(H) \cong H\hat{\otimes}_2 H$ ([21] Section 47). For every $A \in L_{(1)}(H)$, the sum

$$trace(A) := \sum_{n=1}^{\infty} \langle Ae_n, e_n\rangle_H \tag{9.23}$$

converges for every basis $\{e_n\}$ of H and independent of the choice of $\{e_n\}$. We call $trace(A)$ the trace of the operator A. The function $||A||_{(1)} := trace(A)$ is a norm on $L_{(1)}(H)$, with respect to which this space is a Banach space.

Remark A.2

(1) The composition of two Hilbert-Schmidt operator A, B belonging to $L_{(2)}(H)$ is of trace class.

(2) If $H = L^2(T)$ and A has kernel $a(s,t)$, then A^* has kernel $a(t,s)$.

Lemma A.3 Suppose $\Pi_k : -r = s_0 < \cdots < s_k = 0$ is a sequence of partitions of $[-r, 0]$, with $||\Pi_k|| \to 0$ as $k \to \infty$. Denote $J_i = (s_{i-1}, s_i]$ and $\Delta_i = s_i - s_{i-1}$. If Y belongs to $V\hat{\otimes}_1 V$ and Y is symmetric, set $Y_k := I_{k^2} \circ Q_{\underline{s}^{k^2}}(Y)$, then $||Y_k||_{(1)} \leq ||Y||_{(1)}$ and

$$\lim_{k\to\infty} ||Y_k - Y||_{(1)} = 0.$$

Proof We also use Y to denote the operator (of trace class) associated with kernel $Y = Y(s,t)$ on V. Let $\{\phi_n\}_{n=1}^{\infty}$ be the normalized eigenvectors of Y, ($||\phi_n||_V = 1$). Then $Y = \sum_{n=1}^{\infty}(Y, \phi_n \otimes \phi_n)\phi_n \otimes \phi_n$ and

$$\begin{aligned} Y_k &= \sum_{n=1}^{\infty}(Y, \phi_n \otimes \phi_n) I_{k^2} \circ Q_{\underline{s}^{k^2}}(\phi_n \otimes \phi_n) \\ &= \sum_{n=1}^{\infty}(Y, \phi_n \otimes \phi_n) I_k \circ Q_{\underline{s}^{k}}(\phi_n) \otimes I_k \circ Q_{\underline{s}^{k}}(\phi_n). \end{aligned}$$

Thus

$$\begin{aligned} ||Y_k||_{(1)} &\leq \sum_{n=1}^{\infty} |(Y, \phi_n \otimes \phi_n)| ||I_k \circ Q_{\underline{s}^{k}}(\phi_n)||_V^2 \\ &\leq \sum_{n=1}^{\infty} |(Y, \phi_n \otimes \phi_n)| ||\phi_n||_V^2 = ||Y||_{(1)}. \end{aligned}$$

Fix $\epsilon > 0$, there exist $N = N(\epsilon) > 0$ such that

$$\sum_{n>N} |(Y, \phi_n \otimes \phi_n)| < \frac{\epsilon}{4}.$$

There exists $K = K(\epsilon, N(\epsilon)) > 0$, such that if $k > K$, then

$$||I_k \circ Q_{\underline{s}^k}(\phi_n) - \phi_n||_V < \frac{\epsilon}{4||Y||_{(1)}}, \text{ for all } n \leq N.$$

If $k > K$, then

$$\begin{aligned}
& ||Y_k - Y||_{(1)} \\
\leq\ & \sum_{n=1}^{\infty} |(Y, \phi_n \otimes \phi_n)| ||I_k \circ Q_{\underline{s}^k}(\phi_n) \otimes I_k \circ Q_{\underline{s}^k}(\phi_n) - \phi_n \otimes \phi_n||_{(1)} \\
\leq\ & \sum_{n>N} |(Y, \phi_n \otimes \phi_n)| \{||I_k \circ Q_{\underline{s}^k}(\phi_n) \otimes I_k \circ Q_{\underline{s}^k}(\phi_n)||_{(1)} + ||\phi_n \otimes \phi_n||_{(1)}\} \\
+\ & \sum_{n=1}^{N} |(Y, \phi_n \otimes \phi_n)| ||I_k \circ Q_{\underline{s}^k}(\phi_n) \otimes I_k \circ Q_{\underline{s}^k}(\phi_n) - \phi_n \otimes \phi_n||_{(1)} \\
\leq\ & 2 \sum_{n>N} |(Y, \phi_n \otimes \phi_n)| + \sum_{n=1}^{N} |(Y, \phi_n \otimes \phi_n)| \\
\times\ & ||(I_k \circ Q_{\underline{s}^k}(\phi_n) - \phi_n) \otimes I_k \circ Q_{\underline{s}^k}(\phi_n) + \phi_n \otimes (I_k \circ Q_{\underline{s}^k}(\phi_n) - \phi_n)||_{(1)} \\
<\ & \frac{\epsilon}{2} + 2 \sum_{n=1}^{N} |(Y, \phi_n \otimes \phi_n)| ||I_k \circ Q_{\underline{s}^k}(\phi_n) - \phi_n||_V < \epsilon. \quad \blacksquare
\end{aligned}$$

References

[1] Arriojas, M., *A Stochastic Calculus for Functional Differential Equations*, Ph.D. Dissertation, Southern Illinois University at Carbondale,1997.

[2] Bell, D. and Mohammed, S., *The Malliavin Calculus and Stochastic Delay Equations*, Journal of Functional Analysis **99 No. 1** 75–99, 1991.

[3] Bogachev, V., *Gaussian Measures,* American Mathematical Society, 1998.

[4] Da Prato, G. and Zabczyk, J., *Stochastic Equations in Infinite Dimensions*, Cambridge University Press, 1992.

[5] Elworthy, K.D., *Stochastic Differential Equations on Manifolds,* Cambridge University Press, 1982.

[6] Grecksch, W., *Stochastic Evolution Equations,* Akademie Verlag, 1995.

[7] Hu,Y., Mohammed, S.-E. and Yan, F., *Discrete-Time Approximations of Stochastic Delay Equations: The Milstein Scheme,* The Annals of Probability **32** No. 1A, 265 – 314, 2004.

[8] Kirillov, A.A. and Gvishiami A.D., *Theorems and Problems in Functional Analysis,* Springer-Verlag, 1982.

[9] Karatzas, I. and Shreve, S., *Brownian Motion and Stochastic Analysis,* Springer-Verlag, 1991.

[10] Kuo, H., *White Noise Distribution Theory,* CRC Press, 1996.

[11] Malliavin, P., *Stochastic Analysis,* Springer-Verlag, 1997.

[12] McShane, E.J., *Stochastic Calculus and Stochastic Models,* Academic Press, 1974.

[13] Mohammed, S.-E.A., *Stochastic Functional Differential Equations,* Pitman Advanced Publishing Program, 1984.

[14] Mohammed, S.-E.A., *Stochastic Differential Systems with Memory: Theory, Examples and Applications*, in *Stochastic Analysis and Related Topics VI. The Geilo Workshop, 1996* , ed. L. Decreusefond, Jon Gjerde, B. Oksendal, A.S. Ustunel, Progress in Probability, 1-77, Birkhäuser, 1998

[15] Metivier, M. and Pellaumail, J., *Stochastic Integration,* Academic Press, 1980.

[16] Nualart, D. and Pardoux, E., *Stochastic Calculus with Anticipating Integrands,* Probability Theory and Related fields **78** 535–581, 1988.

[17] Nualart, D., *The Malliavin Calculus and Related Topics,* Springer-Verlag, 1995.

[18] Nualart, D. and Zakai, M., *Generalized Stochastic Integrals and the Malliavin Calculus,* Probability Theory and Related Fields **73** 255–280, 1986.

[19] Pardoux, E., *Applications of Anticipating Stochastic Calculus to Stochastic Differential Equations,* Stochastic Analysis and Related Topics II, Lecture Notes in Mathematics **1444,** 63–105, Springer-Verlag, 1988.

[20] Skorohod, A.V., *Random Linear Operators,* D.Reiel Publishing Company, Dordreht, Holland, 1984.

[21] Treves, F., *Topological Vector Spaces, Distributions and Kernels,* Academic Press, 1967.

[22] Yan, F., *Topics on Stochastic Delay Equations,* Ph.D. Thesis, Southern Illinois University at Carbondale, August, 1999.

[23] Yosida, K. *Functional Analysis,* 6th Edition, Springer-Verlag, 1995.