DISCRETE-TIME APPROXIMATIONS OF STOCHASTIC DIFFERENTIAL SYSTEMS WITH MEMORY*

YAOZHONG HU*, SALAH-ELDIN A. MOHAMMED‡ AND FENG YAN

ABSTRACT. In this paper, we develop two discrete-time strong approximation schemes for solving stochastic differential systems with memory: strong Euler-Maruyama schemes for stochastic delay differential equations (SDDE's) and stochastic functional differential equations (SFDE's) with continuous memory, and a strong Milstein scheme for SDDE's. The convergence orders of the Euler-Maruyama and Milstein schemes are 0.5 and 1 respectively. In order to establish the Milstein scheme, we prove an infinite-dimensional Itô formula for "tame" functions acting on the segment process of the solution of an SDDE. It is interesting to note that the presence of the memory in the SDDE requires the use of the Malliavin calculus and the anticipating stochastic analysis of Nualart and Pardoux. Given the *non-anticipating* nature of the sfde's, the use of anticipating calculus methods appears to be novel.

1. Introduction

Discrete-time strong approximation schemes for stochastic ordinary differential equations (SODE's) are well developed. For an extensive study of these numerical schemes, one may refer to ([16]), ([17]), and ([19], Chapters 5 and 6). Some basic ideas of strong and weak orders of convergence are illustrated in ([11]).

If the rate of change of a physical system depends only on its *present state* and some noisy input, then the system can often be described by a stochastic *ordinary* differential equation (SODE). However, in many physical situations the rate of change of the state depends not only on the present but also on the past states of the system. In such

*The research of this author is supported in part by an NSF EPSCOR grant and the General Research Fund of the University of Kansas. April 27, 2001
‡The research of this author is supported in part by NSF grants DMS-9703596 and DMS-9975462.
AMS 1991 *subject classifications.* Primary 60H10, 60H20; secondary 60H25.
Key words and phrases. Euler-Maruyama scheme, Milstein scheme, Itô's formula, anticipating calculus, Malliavin calculus, weak derivatives.

cases, stochastic delay differential equations (SDDE's) or stochastic functional differential equations (SFDE's) provide an important tool to describe and analyze these systems. For various aspects of the qualitiative theory of SFDE's the reader may refer to ([20], [21]) and the references therein.

SDDE's and SFDE's arising in many applications cannot be solved explicitly, and hence the need for developing effective numerical techniques for such systems. Depending on the particular physical model, it may be necessary to design strong L^p (or almost sure) numerical schemes for *pathwise* solutions of the underlying SFDE. Strong approximation schemes for SFDE's may be used to simulate directly the a.s. stochastic dynamics of their trajectories or their random attractors. SFDE's are used to model population growth with incubation/gestation period ([21]). In such models, one is often interested in estimating the actual population rather than its distribution, and hence the need for strong approximation schemes.

In this article, we will not consider the order of convergence of *weak* numerical schemes, although such schemes are useful for some applications of SODE's (see [11], [16] and the references therein). In this connection, it is important to note that stochastic systems with memory *do not correspond to deterministic PDE's* (in finitely many space variables) ([20], [21]). Typically, a stochastic system with memory corresponds to an *infinite-dimensional* Feller diffusion whose principal coefficient degenerates on a hypersurface with *finite-codimension* ([20], Chapter IV, Theorem 3.2, [21], Theorem II.3). This aspect of SFDE's is in sharp contrast with the theory of SODE's where the latter theory has traditional ties to diffusions in Euclidean space. In a sense, the numerics of stochastic systems with memory resemble those of SPDE's in one space dimension.

A strong Cauchy-Maruyama scheme for a class of SFDE's with continuous memory, in the context of the Delfour-Mitter state space $\mathbf{R}^m \times L^2([-r,0],\mathbf{R}^m)$, was developed by T.A. Ahmed, S.A. Elsanousi and S.-E. A. Mohammed ([1]). See also [20], p. 227, and [13].

In sections 3 and 4 of this paper, we develop *strong Euler-Maruyama* schemes for SDDE's with several discrete delays and for stochastic functional differential equations (SFDE's) (with mixed discrete and continuous memory dependence). Our estimates are formulated using the supremum norm in the state space $C([-r,0],\mathbf{R}^m)$ (cf. [1]).

In sections 5-8, we establish the *strong Milstein scheme* for SDDE's with several delays. Although the solution of the SDDE is non-anticipating, methods from *anticipating* stochastic analysis and the Malliavin calculus are necessary in order to derive an Itô formula for the segment of the solution process. The Itô formula is essential for the development of the Milstein scheme.

In order to describe our set-up, we need the following notation.

Let $\mathbf{R}^m$ be m-dimensional Euclidean space with the Euclidean norm $|x| := \sqrt{x_1^2+\cdots+x_m^2}$, $x=(x_1,\cdots,x_m)\in\mathbf{R}^m$. Denote $T:=[0,a]$, $J:=[-r,0]$, $C:=C(J;\mathbf{R}^m)$, where m is a positive integer, $r>0$ and $a>0$. Furnish C with the supremum norm:

$$\|\eta\|_C := \sup_{-r\le s\le 0}|\eta(s)|$$

for all $\eta\in C$.

Define the projection $\Pi : C\to\mathbf{R}^{mk}$ associated with $s_1,\cdots,s_k\in[-r,0]$ by

$$\Pi(\eta) := (\eta(s_1),\cdots,\eta(s_k))\in\mathbf{R}^{mk} \tag{1.1}$$

for all $\eta\in C$.

Definition 1.1.

A function $\Phi\in C(T\times C(J;\mathbf{R}^m);\mathbf{R})$ is *tame* if there exist $\phi\in C(T\times\mathbf{R}^{mk},\mathbf{R})$ and a projection Π such that

$$\Phi(t,\eta)=\phi(t,\Pi(\eta)). \tag{1.2}$$

for all $t\in T$ and $\eta\in C$.

For any continuous m-dimensional process $\{X(t)\}_{t\in[-r,a]}$, define the segment process X_t, $t\in[0,a]$, by

$$X_t(u) = X(t+u), \quad t\in[0,a], \quad u\in[-r,0]. \tag{1.3}$$

Observe that $\{X_t\}$ may be considered as a C-valued or $L^2(J;\mathbf{R}^m)$-valued process.

It is important that one should distinguish between the finite-dimensional current state $X(t)$ and the infinite-dimensional segment X_t, $t\in[0,a]$.

Assume that $g : T\times\mathbf{R}^{mk_1}\to L(\mathbf{R}^d;\mathbf{R}^m)$ and $h : T\times\mathbf{R}^{mk_2}\to\mathbf{R}^m$ satisfy the following *Lipschitz condition* for all $t\in T$, $x,y\in\mathbf{R}^{mk_1}$ and $z,w\in\mathbf{R}^{mk_2}$:

$$|g(t,x)-g(t,y)|\le L|x-y|, \quad |h(t,z)-h(t,w)|\le L|z-w| \tag{1.4}$$

where $L>0$ is a constant, together with the *boundedness condition*:

$$\sup_{0\le t\le a}\big[|g(t,0)|+|h(t,0)|\big]<\infty. \tag{1.5}$$

Let Π_1 and and Π_2 be two projections associated with two sets of points $s_{1,1},\cdots,s_{1,k_1}\in[-r,0]$ and $s_{2,1},\cdots,s_{2,k_2}\in[-r,0]$, respectively. Suppose $\{W(t):=(W^1(t),\cdots,W^d(t)) : t\ge 0\}$ is a d-dimensional standard Brownian motion defined on a probability space $(\Omega,\mathcal{F},P)$. Let $\eta:\Omega\to C([-r,0];\mathbf{R}^m)$ be a random initial path independent of $\{W(t) : t\ge 0\}$.

We will first consider the following class of Itô SDDE's:

$$X(t) = \begin{cases} \eta(0)+\int_0^t g(s,\Pi_1(X_s))\,dW(s)+\int_0^t h(s,\Pi_2(X_s))\,ds, & t\ge 0\\ \eta(t), \quad -r\le t<0. \end{cases} \tag{1.6}$$

Under conditions (1.4) and (1.5) the SDDE (1.6) has a unique strong solution (c.f. [20], Theorem II.2.1, p. 36; and Theorem V.4.3, pp. 151-152). To see this, let $G(t,\eta):=g(t,\Pi_1(\eta))$ and $H(t,\eta):=h(t,\Pi_2(\eta))$ for $t\in[0,a],\eta\in C$. It is easy to check that G and H satisfy the Lipschitz and local boundedness conditions (with respect to the supremum norm on C) of Theorems II.2.1 and V.4.3 of [20]. Therefore, for each $m \geq 1$, there exists a constant $C = C(m, L, a) > 0$ such that

$$E||X_t||_C^{2m} \leq C(1 + E||\eta||_C^{2m}) \tag{1.7}$$

for all $\eta \in C, t \in [0, a]$.

First, we propose an *Euler-Maruyama scheme* for (1.6) as follows. Let $\pi : 0 = t_0 < t_1 < t_2 < \cdots < t_n$ be a partition of $[0, a]$ to be specified later. Denote by $|\pi| := \max_{0 \leq i \leq n-1} (t_{i+1} - t_i)$, the mesh of π. Define the Euler-Maruyama approximation X^π for the solution X of (1.6) by

(1.8)
$$X^\pi(t) = \begin{cases} X^\pi(t_i) + g(t_i, \Pi_1(X^\pi_{t_i}))(W(t) - W(t_i)) + h(t_i, \Pi_2(X^\pi_{t_i}))(t - t_i), t \in (t_i, t_{i+1}] \\ \eta(t), \qquad -r \leq t \leq 0, \end{cases}$$

where $X_t^\pi(s) = X^\pi(t + s)$, $s \in [-r, 0], t \geq 0$. It will be shown that under some regularity conditions on the coefficients, one has the error estimate

$$E \sup_{0 \leq t \leq a} ||X_t^\pi - X_t||_C^q \leq C(q)|\pi|^{\frac{q}{2}} \tag{1.9}$$

for any $q \geq 1$. As in the SODE case, the above estimate shows that the Euler-Maruyama scheme has 0.5 as a strong order of convergence. These results are presented in section 3.1.

There are many ways to partition an interval into subintervals. For example, when we graph a function $h : [0, a] \to \mathbf{R}$, we should evaluate it very frequently in those intervals where h changes dramatically. If a fixed number of evaluations are permitted, then there is a problem deciding exactly which points one should use for the above Euler-Maruyama scheme. See ([6]) and ([12]) for a discussion of this issue. In section 4, we shall consider this question for SDDE's. A non-negative function h with finitely many zeros is used to express a "way" of partitioning an interval into sub-intervals. An optimal way to achieve such a partition is also given in section 4 (Theorem 4.1). This result yields an *exact* convergence rate for the Euler-Maruyama scheme when applied to a one-dimensional linear SDDE with a single delay.

The second class of SFDE's that we will consider are those with mixed discrete and continuous memory:

(1.10)
$$X(t) = \begin{cases} \eta(0) + \int_0^t g(s, \Pi_1(X_s), Q_1(X_s))\, dW(s) + \int_0^t h(s, \Pi_2(X_s), Q_2(X_s))\, ds, \ t \geq 0\,, \\ \eta(t), \qquad -r \leq t \leq 0, \end{cases}$$

where Π_1 and Π_2 are two projections of "discrete type", Q_1 and Q_2 are two projections of "continuous type" defined by

$$Q_i(\eta) := (Q_{i,1}(\eta), \cdots, Q_{i,m_i}(\eta)),\ i = 1, 2,$$
$$Q_{ij}(\eta) := \int_{-1}^0 \phi_{ij}(\eta(s)) a_{ij}(s)\, ds\,, j = 1, \cdots, m_i,$$

where $m_1, m_2 \geq 1$ are integers, $a_{ij} \in C^{\frac{1}{2}}(J, \mathbf{R})$ and $\phi_{ij} : \mathbf{R}^m \to \mathbf{R}$, $i = 1, 2,\ j = 1, \cdots, m_i$, are functions satisfying Lipschitz and linear growth conditions.

For the SFDE (1.10), we can define the Euler-Maruyama approximations by

$$\begin{aligned} X^\pi(t) &= X^\pi(t_i) + g(t_i, \Pi_1(X^\pi_{t_i}), Q^\pi_1(X^\pi_{t_i}))(W(t) - W(t_i)) \\ &\qquad + h(t_i, \Pi_2(X^\pi_{t_i}), Q^\pi_2(X^\pi_{t_i}))(t - t_i), \quad t \in (t_i, t_{i+1}], \\ X^\pi(t) &= \eta^\pi(t), \qquad -r \leq t \leq 0, \end{aligned} \tag{1.11}$$

where $Q^\pi_i(\eta), i = 1, 2$, are approximations of $Q_i(\eta)$ to be specified in section 3. We prove in section 3.2 that the Euler-Maruyama scheme for (1.10) has strong order of convergence 0.5.

We then introduce the following Milstein scheme for the SDDE (1.6):

(1.12)
$$\begin{aligned} X^{i,\pi}(t) &= X^{i,\pi}(t_k) + h^i(t_k, \Pi_2(X^\pi_{t_k}))(t - t_k) + g^{il}(t_k, \Pi_1(X^\pi_{t_k}))(W^l(t) - W^l(t_k)) \\ &\quad + \frac{\partial g^{il}}{\partial x_{i_1 j_1}}(t_k, \Pi_1(X^\pi_{t_k})) u^{i_1 j_1, \pi}(t_k + s_{1,j_1}) I_{l,l_1}(t_k + s_{1,j_1}, t + s_{1,j_1}; s_{1,j_1}), \end{aligned}$$

for $t_k < t \leq t_{k+1}$, where

$$u^{i_1 j_1,\pi}(t) = \begin{cases} g^{i_1 j_1}(t, \Pi_1(X_t^\pi)), & t \geq 0, \\ 0, & -1 \leq t < 0, \end{cases} \tag{1.13}$$

and

$$I_{l,l_1}(t_0 + s_{i,j}, t + s_{i,j}; s_{i,j}) = \int_{t_0}^{t} \int_{t_0+s_{i,j}}^{t_1+s_{i,j}} \circ dW^l(t_2) \circ dW^{l_1}(t_1). \tag{1.14}$$

In (1.12), X^i, h^i and g^{il} denote coordinate representations of X, h and g with respect to standard bases in the underlying Euclidean spaces, and the Einstein summation convention is used for repeated indices.

In order to establish strong convergence of the above Milstein scheme for the SDDE (1.6), it turns out -surprisingly-that one requires the use of *anticipating* calculus techniques developed by Nualart and Pardoux ([22]). In particular, one needs to develop an infinite-dimensional Itô formula for "tame" functions acting on the segment X_t of the solution X of (1.6). Such an Itô formula is given in Section 3, Theorem 3.3. The formula is proved via anticipating calculus methods ([22]). To understand the need for anticipating calculus in such an intrinsically adapted setting, it is instructive to look at the following simple one-dimensional SDDE:

$$\begin{aligned} dX(t) &= g(X(t-1), X(t))\, dW(t), \quad t \geq 0 \\ X(t) &= W(t), \quad -1 \leq t < 0. \end{aligned}$$

where $g : \mathbf{R}^2 \to \mathbf{R}$ is a smooth function and $W(t), t \geq -1$, is one-dimensional Brownian motion. For a second-order scheme, we formally seek a stochastic differential of the coefficient $g(X(t-1), X(t))$ on the right hand side of the above SDDE. For $t \in (0, 1]$, this gives formally:

$$\begin{aligned} & d\{g(X(t-1), X(t))\} \\ &= d\{g(W(t-1), X(t))\} \\ &= \frac{\partial g}{\partial x}(W(t-1), X(t))\, dW(t-1) + \frac{\partial g}{\partial y}(W(t-1), X(t)) g(X(t-1), X(t))\, dW(t) \\ &\qquad + \text{second-order terms.} \end{aligned}$$

Note that although the coefficient $g(X(t-1), X(t))$ is $\mathcal{F}_t$-measurable, the first term $\frac{\partial g}{\partial x}(W(t-1), X(t))\, dW(t-1)$ in the right hand side of the last equality is an *anticipating* differential. Furthermore, it appears that the $(\mathcal{F}_t)_{0\le t\le 1}$-adapted process $[0,1] \ni t \to (X(t-1), X(t)) \in \mathbf{R}^2$ is not a semimartingale with respect to any natural filtration. In addition to this difficulty, the components $X(t-1)$ and $X(t)$ are not independent, so the existing anticipating versions of Itô's formula do not apply (cf. [2], [3] and [22]); hence the need for a new Itô formula for tame functions in order to justify the above computation. In section 5 (Theorem 5.3), we establish such a formula.

Using the Itô formula of section 5 and appropriate estimates on the weak Cameron-Martin derivatives of X, it is shown in section 8 that, under suitable regularity conditions on the coefficients of (1.6), one gets the following global error estimate for the Milstein approximations

$$E \sup_{0\le t\le a} ||X_t^\pi - X_t||_C^q \le C(q)|\pi|^q \tag{1.15}$$

for any $q \ge 1$. This says that the Milstein scheme has strong order of convergence 1.

2. Preliminary Results

Let $\eta : [-r, 0] \to \mathbf{R}^m$ be a given continuous initial path, and let W be a d-dimensional Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge 0}, P)$.

We shall use the notations introduced in section 1.

Assume that the functions $g : T \times \mathbf{R}^{mk_1} \to L(\mathbf{R}^d; \mathbf{R}^m)$ and $h : T \times \mathbf{R}^{mk_2} \to \mathbf{R}^m$ satisfy (1.4) and (1.5). Let Π_1 and Π_2 be two projections associated with two sets of points $s_{1,1}, \cdots, s_{1,k_1}$ and $s_{2,1}, \cdots, s_{2,k_2}$, respectively. Consider the SDDE

$$X(t) = \begin{cases} \eta(0) + \int_0^t g(s, \Pi_1(X_s))\, dW(s) + \int_0^t h(s, \Pi_2(X_s))\, ds, t \ge 0 \\ \eta(t), -r \le t < 0, \end{cases} \tag{2.1}$$

where $\eta \in C([-r, 0]; \mathbf{R}^m)$ almost surely and is independent of the Brownian motion $\{W(t) : t > 0\}$.

Recall that the SDDE (2.1) has a unique strong solution X, and for each integer $k \geq 1$, there exists a constant $C = C(k, L, a) > 0$ such that

$$E||X_t||_C^{2k} \leq C(1 + E||\eta||_C^{2k})] , \tag{2.2}$$

for all $\eta \in C$ and $t \in [0, a]$ ([20]).

Next we define *convergence, consistency and stability.*

Suppose that $X = \{X(t) : t \in T\}$ is the solution of some SDDE, and Y^π is a discrete-time approximation of X based on a partition $\pi := \{t_i : i = 1, \cdots, n\}$ of T.

Definition 2.1.

We say that a discrete-time approximation Y^π *converges strongly with order* $\gamma > 0$ at time t to X if there exists a positive constant C, independent of π, and a $\delta_0 > 0$ such that

$$E|Y^\pi(t) - X(t)| \leq C|\pi|^\gamma \tag{2.3}$$

whenever $|\pi| \in (0, \delta_0)$.

Definition 2.2.

We say that a discrete-time approximation Y^π is *strongly consistent* if there exists a nonnegative function $c = c(\delta)$ with

$$\lim_{\delta \downarrow 0} c(\delta) = 0 \tag{2.4}$$

such that

$$E\left|E\left(\frac{Y^\pi(t_{k+1}) - Y^\pi(t_k)}{\Delta_k}\middle|\mathcal{F}_{t_k}\right) - h(t_k, \Pi_2(Y^\pi_{t_k}))\right|^2 \leq c(|\pi|) \tag{2.5}$$

and

(2.6)

$$E(\left|\frac{1}{\Delta_k}\left[Y^\pi(t_{k+1}) - Y^\pi(t_k) - E(Y^\pi(t_{k+1}) - Y^\pi(t_k)|\mathcal{F}_{t_k})\right]\right.$$
$$\left. - g(t_k, \Pi_1(Y^\pi_{t_k}))(W(t_{k+1}) - W(t_k))\right|^2) \leq c(|\pi|)$$

for all fixed values $Y^\pi_{t_k} = \eta$ and where $\Delta_k := t_{k+1} - t_k$.

Definition 2.3.

Suppose that a numerical scheme for an SDDE gives rise to discrete-time approximations $Y^\pi, \bar{Y}^\pi$ starting at time t_0 at $Y^\pi_{t_0}, \bar{Y}^\pi_{t_0}$, respectively. We say that the numerical scheme is *(stochastically) numerically stable* if for any finite interval $[t_0, a]$ there exists a positive constant δ_0 such that for each $\epsilon > 0$ one has

$$\lim_{||Y^\pi_{t_0} - \bar{Y}^\pi_{t_0}||_C \to 0} \sup_{t_0 \le t \le a} P\{||Y^\pi_{t_{n_t}} - \bar{Y}^\pi_{t_{n_t}}||_C \ge \epsilon\} = 0, \tag{2.7}$$

whenever $|\pi| \in (0, \delta_0)$, and where $n_t := \max\{i : t_i \le t < t_{i+1}\}$.

To obtain the order of convergence, one needs to study the dependence of the solution of the SDDE (2.1) on the delays. Define the *distance* $d(\Pi_1, \Pi_2)$ between two projections Π_1 and Π_2 associated with two sets of points $s_{1,1} < \cdots < s_{1,k_1}$ and $s_{2,1} < \cdots < s_{2,k_2}$, by the formula

$$d(\Pi_1, \Pi_2) = \begin{cases} +\infty & \text{if } k_1 \neq k_2 \\ \max_{1 \le j \le k_1} |s_{1,j} - s_{2,j}| & \text{if } k_1 = k_2 \,. \end{cases}$$

Let us consider two SDDE's

$$X(t) = \begin{cases} \eta(0) + \int_0^t g(s, \Pi_{11}(X_s))\, dW(s) + \int_0^t h(s, \Pi_{12}(X_s))\, ds, \ t \in T, \\ \eta(t), \quad -r \le t < 0, \end{cases} \tag{2.8}$$

and

$$Y(t) = \begin{cases} \eta(0) + \int_0^t g(s, \Pi_{21}(Y_s))\, dW(s) + \int_0^t h(s, \Pi_{22}(Y_s))\, ds, \ t \in T, \\ \eta(t), \quad -r \le t < 0. \end{cases} \tag{2.9}$$

We shall estimate the difference between $X(t)$ and $Y(t)$. When $k_1 = k_2 = 1$, Bell and Mohammed ([5]) showed that if $s_{1,1} \to 0$ and $s_{2,1} \to 0$, then the solution $X(t)$ of (2.1) converges in $L^2(\Omega, \mathbf{R}^m)$ to the solution of the corresponding SODE.

The following lemma extends the result in [5] to the case of several delays. This extension will be useful in studying the order of convergence of the numerical schemes for SDDE's.

Lemma 2.4.

Suppose that g, h satisfy (1.4) and (1.5). Let $0 < \gamma \leq 1$, $\eta \in C^\gamma([-r,0], \mathbf{R}^m)$, and suppose that X and Y are solutions of (2.8) and (2.9), respectively. Then for each $q \geq 2$, there exists a constant $C(q) > 0$ such that

$$E \sup_{0 \leq t \leq a} ||Y_t - X_t||_C^q \leq C(q) \left\{ d(\Pi_{11}, \Pi_{21}) + d(\Pi_{12}, \Pi_{22}) \right\}^{q\gamma}. \tag{2.10}$$

Proof.

Note first that there exists a constant $M > 0$ such that

$$\sup_{-r \leq t_1 \leq z_1 < z_2 \leq t_2 \leq a} E|Y(z_1) - Y(z_2)|^q \leq M|t_2 - t_1|^{q\gamma}. \tag{2.11}$$

Let $\Pi_{11}, \Pi_{12}, \Pi_{21}, \Pi_{22}$ be associated with $\{s_{1,1} < s_{1,2} < \cdots < s_{1,k_1}\}$, $\{s_{2,1} < s_{2,2} < \cdots < s_{2,k_2}\}$, $\{r_{1,1} < r_{1,2} < \cdots < r_{1,k_1}\}$, and $\{r_{2,1} < r_{2,2} < \cdots < r_{2,k_2}\}$, respectively, with $r_{i,j}, s_{i,j} \in [-r, 0]$. Suppose $0 \leq t \leq a$. Then by the Burkholder-Davis-Gundy inequality and (1.7), we have

$$\begin{aligned}
&E \sup_{0 \leq u \leq t} |Y(u) - X(u)|^q \\
&\qquad \leq C_1(q) E \int_0^t |h(s, \Pi_{22}(Y_s)) - h(s, \Pi_{12}(X_s))|^q \, ds \\
&\qquad + C_2(q) E \int_0^t |g(s, \Pi_{21}(Y_s)) - g(s, \Pi_{11}(X_s))|^q \, ds \\
&\qquad \leq C_3(q) \int_0^t E(\sum_{i=1}^{k_2} |Y(s + r_{2,i}) - X(s + s_{2,i})|^q) \, ds \\
&\qquad + C_4(q) \int_0^t E(\sum_{i=1}^{k_1} |Y(s + r_{1,i}) - X(s + s_{1,i})|^q) \, ds \\
&\qquad \leq C_5(q) \int_0^t E \sum_{i=1}^{k_2} \{|Y(s + r_{2,i}) - Y(s + s_{2,i})|^q + |Y(s + s_{2,i}) - X(s + s_{2,i})|^q\} \, ds \\
&\qquad + C_6(q) \int_0^t E \sum_{i=1}^{k_1} \{|Y(s + r_{1,i}) - Y(s + s_{1,i})|^q + |Y(s + s_{1,i}) - X(s + s_{1,i})|^q\} \, ds
\end{aligned}$$

for all $t \in T$. Thus from the above inequality, (2.11) and the definition of $d(\Pi_1, \Pi_2)$, it follows that

$$E \sup_{0 \le u \le t} |Y(u) - X(u)|^q \le C_7(q) \left\{ M\delta^{q\gamma} + \int_0^t E \sup_{0 \le u \le s} ||Y_u - X_u||_C^q \, ds \right\}, \quad t \in T,$$

where $\delta := d(\Pi_{11}, \Pi_{21}) + d(\Pi_{12}, \Pi_{22})$. Hence

$$E \sup_{0 \le u \le t} ||Y_u - X_u||_C^q \le C_8(q) M \delta^{q\gamma} + C_9(q) \int_0^t E \sup_{0 \le u \le s} ||Y_u - X_u||_C^q \, ds, \tag{2.12}$$

for all $t \in T$. By Gronwall's lemma, this implies that

$$E \sup_{0 \le s \le t} ||Y_s - X_s||_C^q \le C(q) \delta^{q\gamma}, \quad t \in [0, a]. \tag{2.14}$$

The proof of the lemma is complete. □

Note that the constant $C(q)$ in Lemma 2.4 also depends on the process Y. Since the rationals $\mathbf{Q}$ are dense in $\mathbf{R}$, by Lemma 2.4, we need only deal with rational delays, i.e., we can assume that the delays $s_{i,j}$ in (2.1) are in $\mathbf{Q}$. This makes computer simulation possible, since one can then control the system error when the delays are irrational.

3. The strong Euler-Maruyama scheme

In this section, we shall develop Euler-Maruyama schemes for SFDE's with discrete and/or continuous memory. For simplicity, we assume that a is a positive integer, $T := [0, a]$ and $J := [-1, 0]$. We also assume rational delays:

$$\{s_{j,i} = -\frac{p_{j,i}}{q_{j,i}} : j = 1, 2, \ 1 \le i \le k_j, p_{j,i} \ge 0, p_{j,i} \in \mathbf{Z}, q_{j,i} \in \mathbf{N}\}.$$

We will adopt the following notation throughout this section.

Let N_0 be the least common multiple of $q_{j,i}, j = 1, 2, \ 1 \le i \le k_j$. Let $p \in \mathbf{N}$ and set $n := apN_0, l := pN_0$. Then l and n are positive integers. We define the rational partition points

$$t_i = \begin{cases} -1 + \frac{1}{pN_0}(i + pN_0), & -l \le i \le 0 \\ \frac{i}{pN_0}, & 0 < i \le n. \end{cases} \tag{3.1}$$

Note that for all $1 \leq i \leq k_j$ and $j = 1, 2$, $t_i + s_{j,i}$ belongs to the partition $\pi_p := \{t_i : -l \leq i \leq n\}$ of $[-1, a]$. Set $\delta_p := |\pi_p| = 1/(pN_0)$ and $n_t := \max\{n \in \mathbf{N} : t_n \leq t\}$. If $-1 \leq s \leq t$ define

$$\lfloor s \rfloor := \begin{cases} t_i, & \text{if } t_i \leq s < t_{i+1} \\ t_{n_t}, & \text{if } t_{n_t} \leq s \leq t. \end{cases}$$

for $-1 \leq i \leq n-1$. For each positive integer p, the superscript p will denote numerical quantities pertaining to the partition π_p, e.g. $X^p := X^{\pi_p}$.

3.1. The Euler-Maruyama scheme for SDDE's

Recall the SDDE

$$\text{(2.1)} \qquad X(t) = \begin{cases} \eta(0) + \int_0^t g(s, \Pi_1(X_s))\, dW(s) + \int_0^t h(s, \Pi_2(X_s))\, ds, & t \geq 0 \\ \eta(t), \quad -1 \leq t < 0, \end{cases}$$

with $r = 1$.

The *Euler-Maruyama* scheme for (2.1) is given by

(3.2)
$$X^p(t) = \begin{cases} X^p(t_i) + g(t_i, \Pi_1(X^p_{t_i}))(W(t) - W(t_i)) + h(t_i, \Pi_2(X^p_{t_i}))(t - t_i), t \in (t_i, t_{i+1}] \\ \eta^p(t), \qquad -1 \leq t \leq 0 \end{cases}$$

where the starting path $\eta^p \in C(J, \mathbf{R}^m)$ is prescribed (e.g. a piece-wise linear approximation of η using the partition points $\{t_{-l}, \cdots, t_0\}$). Define the error function Z^p by

$$\text{(3.3)} \qquad \begin{cases} Z^p(t) = X^p(t) - X(t), \quad 0 \leq t \leq a, \\ Z^p_0 = X^p_0 - X_0. \end{cases}$$

Theorem 3.1.

Assume that the coefficients g and h in (2.1) satisfy (1.4), (1.5) and the following condition

$$\text{(3.4)} \qquad \begin{cases} |g(s,x) - g(t,x)| \leq L_1(1 + |x|)|s - t|^\gamma, \ \textit{for all } x \in \mathbf{R}^{mk_1}, s, t \in T \\ |h(s,x) - h(t,x)| \leq L_1(1 + |x|)|s - t|^\gamma, \ \textit{for all } x \in \mathbf{R}^{mk_2}, s, t \in T \end{cases}$$

for some positive constant L_1. Fix any integer $q \geq 2$. Suppose that $\eta : [-1, 0] \to L^q(\Omega, \mathbf{R}^m)$ is Hölder continuous with exponent $\gamma \in (0, 1]$, i.e., there is a positive constant K such that

$$E|\eta(s) - \eta(t)|^q \leq K|s - t|^{\gamma q}$$

for all $s, t \in [-1, 0]$. Suppose also that there is a positive constant $C' := C'(q)$ such that

$$E||\eta^p - \eta||_C^q \leq C'\delta_p^{\gamma q}$$

Then there exists a constant $C'' := C''(q, a) > 0$, depending on a and q, such that

$$E \sup_{0 \leq s \leq a} ||Z_s^p||_C^q \leq C''\delta_p^{\tilde{\gamma} q}$$

where $\tilde{\gamma} := \gamma \wedge (1/2)$.

Proof.

Since the SDDE (2.1) is a special case of the SFDE (3.12), the reader may consult the proof of Theorem 3.4 in the next section. □

The requirement

$$E||\eta^p - \eta||_C^q \leq C'\delta_p^{\gamma q}$$

in the statement of Theorem 3.1 is fulfilled if one takes η^p to be the piecewise-linear approximation

$$\eta^p(s) := [(t_{i+1} - s)\eta(t_i) + (s - t_i)\eta(t_{i+1})](t_{i+1} - t_i)^{-1}, \quad s \in [t_i, t_{i+1}]$$

for $-l \leq i \leq 0$.

3.2. The Euler-Maruyama scheme for SFDE's with mixed discrete and continuous memory.

Let $m_1, m_2 \geq 1$, $a_{ij} \in C^{\frac{1}{2}}(J)$, and let $\phi_{ij} : \mathbf{R}^m \to \mathbf{R}$, $i = 1, 2$, $j = 1, \cdots, m_i$, satisfy Lipschitz and linear growth conditions. Consider the following SFDE with mixed discrete and continuous memories:

(3.12)

$$X(t) = \eta(0) + \int_0^t g(s, \Pi_1(X_s), Q_1(X_s))\, dW(s) + \int_0^t h(s, \Pi_2(X_s), Q_2(X_s))\, ds, \quad t \in [0, a],$$
$$X_0 = \eta \in C = C(J; \mathbf{R}^m)$$

where Π_1 and Π_2 are two projections of discrete type, Q_1 and Q_2 are two projections of continuous type defined by

$$Q_i(\eta) := (Q_{i,1}(\eta), \cdots, Q_{i,m_i}(\eta)), \quad i = 1, 2,$$
$$Q_{ij}(\eta) := \int_{-1}^{0} \phi_{ij}(\eta(s)) a_{ij}(s)\, ds\,, \quad j = 1, \cdots, m_i.$$

We assume that $g : T \times \mathbf{R}^{k_1 m + m_1} \to \mathbf{R}$ and $h : T \times \mathbf{R}^{k_2 m + m_2} \to \mathbf{R}$ satisfy the uniform Lipschitz condition:

$$\begin{cases} |g(t,x) - g(t,y)| \le L|x-y|\,, & x, y \in \mathbf{R}^{k_1 m + m_1} \\ |h(t,z) - h(t,w)| \le L|z-w|, & z, w \in \mathbf{R}^{k_2 m + m_2}\,,\ t \in [0,a], \end{cases}$$

and local boundedness condition:

$$\sup_{0 \le t \le a} \left[|g(t,0)| + |h(t,0)| \right] < \infty,$$

where L is a positive constant independent of $t \in [0, a]$.

Under the above conditions, the SFDE (3.12) has a unique strong solution (c.f. [20], Theorem II.2.1 and Theorem V.4.3).

Define the approximations Q_{ij}^p of Q_{ij} by

$$Q_{ij}^p(\eta) = \sum_{k=-l}^{-1} \phi_{ij}(\eta(s_k)) a_{ij}(s_k)(s_{k+1} - s_k). \tag{3.15}$$

Remark 3.2.

If $\eta : [-1, 0] \to L^q(\Omega, \mathbf{R}^m)$ is Hölder continuous with exponent γ, $0 < \gamma \le 1$, and $q \ge 2$, then it is easy to show that there is a constant $C(q) > 0$ such that

$$E|Q_{ij}^p(\eta) - Q_{ij}(\eta)|^q \le C(q) \delta_p^{\tilde{\gamma} q}, \tag{3.16}$$

where $\tilde{\gamma} := \gamma \wedge (1/2)$.

Lemma 3.3.

If $\eta : [-1,0] \to L^q(\Omega, \mathbf{R}^m)$, $(q \geq 2)$, is Hölder continuous with exponent γ, then there exists a constant $C(q) > 0$ such that

$$\sup_{0 \leq t \leq a} E|Q^p_{ij}(X_t) - Q_{ij}(X_t)|^q \leq C(q)\delta_p^{\tilde{\gamma}q} \tag{3.17}$$

for all $i = 1, 2$ and $j = 1, \cdots, m_i$, where $\tilde{\gamma} = \gamma \wedge (1/2)$.

Proof.

Fix i, j, where $i = 1, 2, 1 \leq j \leq m_i$, and let $t \in [0, a]$. Using the notation $\lfloor s \rfloor$, we may write

(3.18)
$$\begin{aligned} Q^p_{ij}(X_t) - Q_{ij}(X_t) &= \int_{-1}^0 [\phi_{ij}(X(t + \lfloor s \rfloor))a_{ij}(\lfloor s \rfloor) - \phi_{ij}(X(t+s))a_{ij}(s)]\, ds \\ &= \int_{-1}^0 a_{ij}(\lfloor s \rfloor)[\phi_{ij}(X(t + \lfloor s \rfloor)) - \phi_{ij}(X(t+s))]\, ds \\ &\quad + \int_{-1}^0 \phi_{ij}(X(t+s))[a_{ij}(\lfloor s \rfloor) - a_{ij}(s)]\, ds \\ &= I_1(t) + I_2(t), \end{aligned}$$

where

$$\begin{aligned} I_1(t) &:= \int_{-1}^0 a_{ij}(\lfloor s \rfloor)[\phi_{ij}(X(t + \lfloor s \rfloor)) - \phi_{ij}(X(t+s))]\, ds \\ I_2(t) &:= \int_{-1}^0 \phi_{ij}(X(t+s))[a_{ij}(\lfloor s \rfloor) - a_{ij}(s)]\, ds. \end{aligned}$$

It follows from (3.12) and standard properties of the Itô integral that there exists a constant $C_1(q) > 0$ such that

$$\sup_{-1 \leq r_1 \leq \alpha < \beta \leq r_2 \leq a} E(|X(\beta) - X(\alpha)|^q) \leq C_1(q)|r_2 - r_1|^{\tilde{\gamma}q} \tag{3.19}$$

for all $q \geq 1$. By (3.19) and the Lipschitz property of ϕ_{ij}, $i = 1, 2$, $j = 1, \cdots, m_i$, it follows that

$$\begin{aligned}\sup_{0 \leq t \leq a} E(|I_1(t)|^q) &\leq C_1(q)||a_{ij}||_C^q \int_{-1}^{0} \sup_{0 \leq t \leq a} E(|X(t + \lfloor s \rfloor) - X(t+s)|^q)\, ds \\ &\leq C_2(q)\|a_{ij}\|_C^q \delta_p^{\tilde{\gamma} q} \\ &\leq C_3(q)\delta_p^{\tilde{\gamma} q}.\end{aligned}$$

Using the Hölder continuity of a_{ij} and the linear growth property of ϕ_{ij}, I_2 can be estimated as follows:

$$\begin{aligned}\sup_{0 \leq t \leq a} E(|I_2(t)|^q) &\leq C_1(q) \int_{-1}^{0} \sup_{0 \leq t \leq a} E(|\phi_{ij}(X(t+s))|^q)|a_{ij}(\lfloor s \rfloor) - a_{ij}(s)|^q\, ds \\ &\leq C_4(q)\delta_p^{q/2} \int_{-1}^{0} \sup_{0 \leq t \leq a} E(|\phi_{ij}(X(t+s))|^q)\, ds \\ &\leq C_5(q)\delta_p^{q/2} \int_{-1}^{0} \sup_{0 \leq t \leq a} E(1 + |X(t+s)|^q)\, ds \\ &\leq C_6(q)\delta_p^{q/2} \int_{-1}^{0} E(1 + ||\eta||_C^q)\, ds \\ &\leq C_7(q)\delta_p^{q/2}.\end{aligned}$$

So there exists a constant $C(q) > 0$ such that

$$\sup_{0 \leq t \leq a} E|Q_{ij}^p(X_t) - Q_{ij}(X_t)|^q \leq C(q)\delta_p^{\tilde{\gamma} q}. \quad \square \tag{3.20}$$

We now introduce the *Euler-Maruyama* scheme for (3.12) as follows:

$$\begin{aligned}X^p(t) &= X^p(t_i) + h(t_i, \Pi_2(X_{t_i}^p), Q_2^p(X_{t_i}^p))(t - t_i) \\ &\quad + g(t_i, \Pi_1(X_{t_i}^p), Q_1^p(X_{t_i}^p))(W(t) - W(t_i)), \quad t \in (t_i, t_{i+1}], \\ X^p(t) &= \eta^p(t), \qquad -1 \leq t \leq 0,\end{aligned} \tag{3.21}$$

where $\eta^p \in C(J, \mathbf{R}^m)$ is prescribed subject to the conditions of Theorem 3.4 below.

Theorem 3.4.

Fix any $q \geq 2$. Assume that $\eta : [-1, 0] \to L^q(\Omega, \mathbf{R}^m)$, is Hölder continuous with exponent γ. Let $Z^p(t) := X^p(t) - X(t)$ denote the error function of the Euler-Maruyama scheme (3.21). Suppose that

$$E\|\eta^p - \eta\|_C^q \leq C'(q)\delta_p^{\tilde{\gamma} q}$$

for some constant $C'(q) > 0$, where $\tilde{\gamma} := \gamma \wedge (1/2)$. Assume also that the coefficients g and h satisfy the Lipschitz and boundedness conditions stated before Remark (3.2) together with the regularity condition

$$\begin{cases} |g(s,x) - g(t,x)| \leq L_1(1 + |x|)|s-t|^\gamma, \text{ for all } x \in R^{mk_1+m_1}, s, t \in T \\ |h(s,x) - h(t,x)| \leq L_1(1 + |x|)|s-t|^\gamma, \text{ for all } x \in R^{mk_2+m_2}, s, t \in T \end{cases}$$

for some positive constant L_1. Then there exists a constant $C(q) > 0$ such that

$$\sup_{-1 \leq s \leq a} E|Z^p(s)|^q \leq C(q)\delta_p^{\tilde{\gamma} q}.$$

Proof.

Let $t_i \leq t < t_{i+1}$. Then the global error $Z^p(t) := X^p(t) - X(t)$ may be written in the form:

$$Z^p(t) = Z^p(0) + I_1^p(t) + I_2^p(t) - U_1^p(t) - V_1^p(t) - U_2^p(t) - V_2^p(t),$$

where

$$I_1^p(t) = \int_0^t [h(\lfloor s \rfloor, \Pi_2(X^p_{\lfloor s \rfloor}), Q_2^p(X^p_{\lfloor s \rfloor})) - h(\lfloor s \rfloor, \Pi_2(X_{\lfloor s \rfloor}), Q_2^p(X_{\lfloor s \rfloor}))]\, ds,$$
$$I_2^p(t) = \int_0^t [g(\lfloor s \rfloor, \Pi_1(X^p_{\lfloor s \rfloor}), Q_1^p(X^p_{\lfloor s \rfloor})) - g(\lfloor s \rfloor, \Pi_1(X_{\lfloor s \rfloor}), Q_1^p(X_{\lfloor s \rfloor}))]\, dW(s),$$
$$U_1^p(t) = \int_0^t (h(s, \Pi_2(X_s), Q_2(X_s)) - h(\lfloor s \rfloor, \Pi_2(X_{\lfloor s \rfloor}), Q_2(X_{\lfloor s \rfloor})))\, ds,$$
$$V_1^p(t) = \int_0^t (g(s, \Pi_1(X_s), Q_1(X_s)) - g(\lfloor s \rfloor, \Pi_1(X_{\lfloor s \rfloor}), Q_1(X_{\lfloor s \rfloor})))\, dW(s),$$
$$U_2^p(t) = \int_0^t (h(\lfloor s \rfloor, \Pi_2(X_{\lfloor s \rfloor}), Q_2(X_{\lfloor s \rfloor})) - h(\lfloor s \rfloor, \Pi_2(X_{\lfloor s \rfloor}), Q_2^p(X_{\lfloor s \rfloor})))\, ds,$$
$$V_2^p(t) = \int_0^t (g(\lfloor s \rfloor, \Pi_1(X_{\lfloor s \rfloor}), Q_1(X_{\lfloor s \rfloor})) - g(\lfloor s \rfloor, \Pi_1(X_{\lfloor s \rfloor}), Q_1^p(X_{\lfloor s \rfloor})))\, dW(s).$$

By Hölder's inequality, we get

$$\begin{aligned}
&E \sup_{0\le s\le t} |I_1^p(s)|^q \\
&= \sup_{0\le s\le t} E\Big|\int_0^s [h(\lfloor r\rfloor, \Pi_2(X^p_{\lfloor r\rfloor}), Q_2^p(X^p_{\lfloor r\rfloor})) - h(\lfloor r\rfloor, \Pi_2(X_{\lfloor r\rfloor}), Q_2^p(X_{\lfloor r\rfloor}))]\, dr\Big|^q \\
&\le C_1(q)\int_0^t \sup_{0\le u\le s} (E|h(u, \Pi_2(X_u^p), Q_2^p(X_u^p)) - h(u, \Pi_2(X_u), Q_2^p(X_u))|^q)\, ds\,.
\end{aligned}$$

Therefore, by the Lipschitz property of h and ϕ_{ij}, it follows that

$$E \sup_{0\le s\le t} |I_1^p(s)|^q \le C_2(q)\int_0^t \sup_{-1\le u\le s} E|X^p(u) - X(u)|^q\, ds \tag{3.22}$$

for some constant $C_2(q) > 0$. The Burkholder-Davis-Gundy inequality implies that

$$\begin{aligned}
&E \sup_{0\le s\le t} |I_2^p(s)|^q \\
&\le C_3(q) E\Big(\int_0^t |g(\lfloor s\rfloor, \Pi_1(X^p_{\lfloor s\rfloor}), Q_1^p(X^p_{\lfloor s\rfloor})) - g(\lfloor s\rfloor, \Pi_1(X_{\lfloor s\rfloor}), Q_1^p(X_{\lfloor s\rfloor}))|^2\, ds\Big)^{\frac{q}{2}} \\
&\le C_4(q) E\Big(\int_0^t |g(\lfloor s\rfloor, \Pi_1(X^p_{\lfloor s\rfloor}), Q_1^p(X^p_{\lfloor s\rfloor})) - g(\lfloor s\rfloor, \Pi_1(X_{\lfloor s\rfloor}), Q_1^p(X_{\lfloor s\rfloor}))|^q\, ds\Big) \\
&\le C_4(q)\int_0^t \sup_{0\le u\le s} E(|g(u, \Pi_1(X_u^p), Q_1^p(X_u^p)) - g(u, \Pi_1(X_u), Q_1({}^pX_u))|^q)\, ds\,.
\end{aligned}$$

Using the Lipschitz property of g, we obtain

$$E \sup_{0\le s\le t} |I_2^p(s)|^q \le C_5(q)\int_0^t \sup_{-1\le u\le s} E|X^p(u) - X(u)|^q\, ds \tag{3.23}$$

for some constant $C_5(q) > 0$. Similarly,

$$\begin{aligned}
&E \sup_{0\le s\le t} |V_1^p(s)|^q \\
&\qquad\le C_6(q) E\Big(\int_0^t |g(s, \Pi_1(X_s), Q_1(X_s)) - g(\lfloor s\rfloor, \Pi_1(X_{\lfloor s\rfloor}), Q_1(X_{\lfloor s\rfloor}))|^2\, ds\Big)^{\frac{q}{2}} \\
&\qquad\le C_7(q) E\Big(\int_0^t |g(s, \Pi_1(X_s), Q_1(X_s)) - g(\lfloor s\rfloor, \Pi_1(X_{\lfloor s\rfloor}), Q_1(X_{\lfloor s\rfloor}))|^q\, ds\Big) \\
&\qquad\le C_7(q)\Big\{\sum_{i=1}^{n_t} E\int_{t_{i-1}}^{t_i} |g(s, \Pi_1(X_s), Q_1(X_s))
\end{aligned}$$

$$
\begin{aligned}
&- g(t_{i-1}, \Pi_1(X_{t_{i-1}}), Q_1(X_{t_{i-1}}))|^q\, ds \\
&+ E\int_{t_{n_t}}^t |g(s, \Pi_1(X_s), Q_1(X_s)) - g(t_{n_t}, \Pi_1(X_{t_{n_t}}), Q_1(X_{t_{n_t}}))|^q\, ds\} \\
&\le C_8(q)\{\sum_{i=1}^{n_t}(\int_{t_{i-1}}^{t_i}(\delta_p^{\tilde\gamma q} + \sup_{-1\le u\le 0} E|X(s+u) - X(t_{i-1}+u)|^q)\, ds \\
&+ E\int_{t_{n_t}}^t (\delta_p^{\tilde\gamma q} + \sup_{-1\le u\le 0} E|X(s+u) - X(t_{n_t}+u)|^q)\, ds\}.
\end{aligned}
$$

Hence

$$
E\sup_{0\le s\le t}|V_1^p(s)|^q \le C_9(q)(\delta_p^{\tilde\gamma q} + \int_0^t \sup_{-1\le u\le 0} E|X(s+u) - X(\lfloor s\rfloor + u)|^q\, ds) \tag{3.24}
$$

for some constant $C_9(q) > 0$. In a similar manner, U_1^p can be estimated as

$$
E\sup_{0\le s\le t}|U_1^p(s)|^q \le C_{10}(q)(\delta_p^{\tilde\gamma q} + \int_0^t \sup_{-1\le u\le 0} E|X(s+u) - X(\lfloor s\rfloor + u)|^q\, ds) \tag{3.25}
$$

for some constant $C_{10}(q) > 0$. Now use the inequality

$$
\sup_{-1\le r_1\le s<t\le r_2\le a} E(|X(s) - X(t)|^q) \le C_7|r_2 - r_1|^{\tilde\gamma q}. \tag{3.19}
$$

Therefore,

$$
\begin{cases} E\sup_{0\le s\le t}|V_1^p(s)|^q \le C_{11}(q)\delta_p^{\tilde\gamma q} \\ E\sup_{0\le s\le t}|U_1^p(s)|^q \le C_{11}(q)\delta_p^{\tilde\gamma q}. \end{cases} \tag{3.26}
$$

By Lemma 3.3, there exists a constant $C_{12}(q) > 0$ such that

$$
\begin{cases} \sup_{0\le s\le t} E|V_2^p(s)|^q \le C_{12}(q)\delta_p^{\tilde\gamma q} \\ \sup_{0\le s\le t} E|U_2^p(s)|^q \le C_{12}(q)\delta_p^{\tilde\gamma q}. \end{cases} \tag{3.27}
$$

From (3.22), (3.23), (3.26) and (3.27), we get

$$
\sup_{0\le s\le t} E|Z^p(s)|^q \le C_{13}(q)\int_0^t \sup_{-1\le u\le s} E|Z^p(u)|^q\, ds + C_{13}(q)\delta_p^{\tilde\gamma q}. \tag{3.28}
$$

Thus

$$
\begin{aligned}
\sup_{-1\le s\le t} E|Z^p(s)|^q &= C_{14}(q)(\sup_{-1\le s\le 0} E|Z^p(s)|^q) + \sup_{0\le s\le t} E|Z^p(s)|^q) \\
&\le C_{14}(q)E||Z_0^p||_C^q + C_{15}(q)\int_0^t \sup_{-1\le u\le s} E|Z^p(u)|^q\, ds + C_{15}(q)\delta_p^{\tilde\gamma q} \\
&\le C_{15}(q)\int_0^t \sup_{0\le u\le s} E|Z^p(u)|^q\, ds + C_{16}(q)\delta_p^{\tilde\gamma q}.
\end{aligned}
$$

By Gronwall's lemma, there exists a constant $C(q) > 0$ such that

$$\sup_{-1 \leq s \leq t} E|Z^p(s)|^q \leq C(q)\delta_p^{\tilde{\gamma}q}.$$

This completes the proof of the theorem. □

Remarks 3.6.

(i) The Cauchy Maruyama scheme can be extended to cover general SFDE's of the form

$$X(t) = \begin{cases} \eta(0) + \int_0^t G(s, X_s)\, dW(s) + \int_0^t H(s, X_s)\, ds, & t \geq 0, \\ \eta(t), \quad -r \leq t < 0. \end{cases} \tag{3.29}$$

Define the approximations to the solution X of (3.29) by

(3.30)
$$X^p(t) = \begin{cases} X^p(t_i) + G(t_i, X^p_{t_i})(W(t) - W(t_i)) + H(t_i, X^p_{t_i})(t - t_i), t \in (t_i, t_{i+1}] \\ \eta^p(t), \qquad -r \leq t \leq 0 \end{cases}$$

where the starting path $\eta^p \in C([-r, 0], \mathbf{R}^m)$ is prescribed so as to satisfy the requirements of Theorem 3.1 (e.g. a piece-wise linear approximation of η using the partition points $\{t_l, \cdots, t_0\}$). Then the conclusion of Theorem 3.4 holds under the following hypotheses on the functionals $G : T \times C([-r, 0], \mathbf{R}^m) \to L(\mathbf{R}^d; \mathbf{R}^m)$ and $H : T \times C([-r, 0], \mathbf{R}^m) \to \mathbf{R}^m$:

(3.31)
$$\|G(t, \eta) - G(t, \xi)\| + |H(t, \eta) - H(t, \xi)| \leq L\|\eta - \xi\|_C, \quad t \in T, \ \eta, \xi \in C([-r, 0], \mathbf{R}^m)$$

$$\sup_{0 \leq t \leq a} \left[\|G(t, 0)\| + |H(t, 0)|\right] < \infty. \tag{3.32}$$

$$\begin{cases} \|G(s, \eta) - G(t, \eta)\| \leq L_1(1 + \|\eta\|_C)|s - t|^\gamma, \text{ for all } \eta \in C(J, \mathbf{R}^m), s, t \in T, \\ |H(s, \eta) - H(t, \eta)| \leq L_1(1 + \|\eta\|_C)|s - t|^\gamma, \text{ for all } \eta \in C(J, \mathbf{R}^m), s, t \in T, \end{cases} \tag{3.33}$$

where L and L_1 are positive constants.

(ii) The Euler-Maruyama schemes (3.2) and (3.21) are strongly consistent (Definition 1.3) with control functions $C(\delta) \equiv 0$, and (stochastically) numerically stable (Definition 1.4). The numerical stability follows by similar arguments to those used in the above proof.

4. Exact convergence rate. An example.

In this section we consider regular partitions $\{\pi_n(h)\}$ of $[0,a]$ that are generated by a continuous positive (hence strictly positive) probability density function $h : [0,a] \to (0,\infty)$. More specifically, for each fixed sample size n and probability density function h the points $t_{k,n} \equiv t_k$ of the partition $\pi_n(h)$ in $[0,a]$ are chosen such that

$$t_0 = 0, \qquad \int_{t_k}^{t_{k+1}} h(s)\, ds = \frac{1}{n}, \qquad k = 0, 1, \cdots, n-1.$$

We thus subdivide the interval in such a way that the areas under h over each subinterval are all equal to $1/n$. It then follows that

$$\lim_{n\to\infty, t_k \to t} n(t_{k+1} - t_k) = 1/h(t). \tag{4.1}$$

Consider the following linear one-dimensional SDDE:

$$\begin{cases} dX(t) = b(t)X(t-1)dW(t)\,, & 0 \le t \le a \\ X(t) = \eta(t)\,, & -1 \le t \le 0. \end{cases} \tag{4.2}$$

The Euler-Maruyama scheme gives

$$X^{\pi_n}(t) = \begin{cases} X^{\pi_n}(t_k) + b(t_k)X^{\pi_n}(t_k - 1)(W(t) - W(t_k)), & t_k \le t < t_{k+1}, \\ \eta(t), & t \in J, \end{cases} \tag{4.3}$$

for $0 \le k \le n-1$. By Theorem 3.1, there is a positive constant C (independent of n) such that

$$nE \sup_{t\in[0,a]} |X(t) - X^{\pi_n}(t)|^2 \le C,$$

for all $n \ge 1$. The constant C is called a *leading coefficient* of the scheme and has various applications (see [6]). We shall show that as $n \to \infty$, the left hand side of the above inequality has a limit. We shall also determine the equation satisfied by this limit.

Theorem 4.1.

Suppose $\eta \in C^\gamma(J, \mathbf{R}^m), 1/2 < \gamma \le 1$. Let $a \ge 1$.. Suppose $b : [0,a] \to \mathbf{R}$ is a bounded continuous function such that

$$|b(t) - b(s)| \le K|t-s|^{(1/2)+\alpha}$$

for all $s, t \in [0,a]$ and some $K, \alpha > 0$. Let X be the solution of the SDDE (4.2), and X^{π_n} be its Euler approximation (4.3). Then $\mathcal{Z}(t) := \lim_{n\to\infty} n\ E|X(t) - X^{\pi_n}(t)|^2$ exists for each $t \in [0,a]$. Furthermore, $\mathcal{Z}(t)$ satisfies the following deterministic linear DDE

$$\mathcal{Z}'(t) = b^2(t)\mathcal{Z}(t-1) + b^2(t)b^2(t-1)EX^2(t-2)/h(t),\ 1 < t < a,$$

$$\mathcal{Z}(t) = 0, \quad -1 \le t \le 1, \tag{4.4}$$

where $EX^2(t)$ is given by the integral equation

$$EX^2(t) = \begin{cases} \eta(0)^2 + \int_0^t b^2(s)EX^2(s-1)\,ds, & t \in [0,a], \\ \eta(t)^2, & t \in [-1, 0). \end{cases} \tag{4.5}$$

Proof.

Rewrite (4.2) as

$$X(t) = X(t_k) + b(t_k)X(t_k - 1)(W(t) - W(t_k)) + I^1_{t_k,t}\,, \tag{4.6}$$

where $t_k \le t < t_{k+1}$ and

$$I^1_{t_k,t} = \int_{t_k}^t [b(s)X(s-1) - b(t_k)X(t_k-1)]\,dW(s)\,. \tag{4.7}$$

Set $Z^\pi(t) := X(t) - X^\pi(t)$, $t \in [-1,a]$. Then

$$Z^\pi(t) = Z^\pi(t_k) + b(t_k)Z^\pi(t_k - 1)(W(t) - W(t_k)) + I^1_{t_k,t}, \quad t_k \le t < t_{k+1}. \tag{4.8}$$

Since $E(Z^\pi(t) - Z^\pi(t_k)|Z^\pi_{t_k}) = 0$, for $t > t_k$, it follows that

$$E[Z^\pi(t) - Z^\pi(t_k)]^2 = E[(Z^\pi)^2(t)] - E[(Z^\pi)^2(t_k)].$$

Thus from (4.8), we obtain

$$E[(Z^\pi)^2(t)] = E[(Z^\pi)^2(t_k)] + b^2(t_k)E[(Z^\pi)^2(t_k-1)](t-t_k) + I^2_{t_k,t} + I^3_{t_k,t}\,. \tag{4.9}$$

where

$$I^2_{t_k,t} = \int_{t_k}^t E[b(s)X(s-1) - b(t_k)X(t_k-1)]^2\,ds \tag{4.10}$$

and

$$I^3_{t_k,t} = 2\int_{t_k}^t Eb(t_k)Z^\pi(t_k-1)[b(s)X(s-1) - b(t_k)X(t_k-1)]ds \tag{4.11}$$

Since $Z(t_k-1)$ is $\mathcal{F}_{t_k-1}$ measurable,

$$\begin{aligned} I^3_{t_k,t} &= 2\int_{t_k}^t b(t_k)\big(b(s)-b(t_k)\big)E[Z^\pi(t_k-1)X(t_k-1)]ds \\ &\le Cn^{-1/2}\int_{t_k}^t (s-t_k)^{\alpha+\frac{1}{2}}ds = Cn^{-1/2}(t-t_k)^{\alpha+\frac{3}{2}}, \end{aligned} \tag{4.12}$$

where we have used the fact that

$$E|Z^\pi(t_k-1)X(s-1)| \le \big\{E|Z^\pi(t_k-1)|^2\big\}^{1/2}\big\{E|X(s-1)|^2\big\}^{1/2} \le Cn^{-1/2}\,.$$

For the rest of the computation, we denote by $H_{t_k,t}$ a generic quantity satisfying the following type of estimate:

$$|H_{t_k,t}| \le Cn^{-1}(t-t_k)^{1+\alpha} \tag{4.13}$$

for some $C, \alpha > 0$ and all $n \ge 1$. With these notations, we may write $I^3_{t_k,t} = H_{t_k,t}$. It is easy to verify that

$$I^2_{t_k,t} = \int_{t_k}^t b(s)^2 E[X(s-1) - X(t_k-1)]^2\,ds + H_{t_k,t}\,.$$

Thus from (4.9) it follows that

$$\begin{aligned} E[(Z^\pi)^2(t)] &= E[(Z^\pi)^2(t_k)] + b^2(t_k)E[(Z^\pi)^2(t_k-1)](t-t_k) \\ &\quad + \int_{t_k}^t b^2(s)E[(X(s-1) - X(t_k-1)]^2\,ds + H_{t_k,t}. \end{aligned} \tag{4.14}$$

For each $n \geq 1$, define the process $J^n(t), 0 \leq t \leq a$, by

$$J^n(t) = J^n(t_k) + \int_{t_k}^{t} b^2(s)E[X(s-1) - X(t_k - 1)]^2 ds, \; t_k \leq t < t_{k+1}, \; k = 1, 2, \cdots n-1.$$

When $0 \leq t \leq 1$, $nJ^n(t) \to 0$. One can easily check that

$$E(X(s-1) - X(t_k - 1))^2 = \int_{t_k - 1}^{s-1} b(v-1)EX^2(v-2)dv\,.$$

Therefore

$$\begin{aligned} J^n(t) &= J^n(t_k) + \int_{t_k}^{t} b^2(s) \int_{t_k - 1}^{s-1} b^2(v-1)EX^2(v-2)dvds \\ &= J^n(t_k) + \frac{1}{2}b^2(t_k)b^2(t_k - 1)EX^2(t_k - 2)(t - t_k)^2 + H_{t_k, t}, \end{aligned}$$

for $t > 1, k \geq n_1$. Recall that $n_1 := \max\{n : t_n \leq 1\}$. By recursively applying the above computation, we obtain

$$\begin{aligned} J^n(t) = &\frac{1}{2}1_{[1,a]}(t)\sum_{k=1}^{n_t} b^2(t_{k-1})b^2(t_{k-1} - 1)EX^2(t_{k-1} - 2)(t_k - t_{k-1})^2 \\ &+ \frac{1}{2}1_{[1,a]}(t)b^2(t_{n_t})b^2(t_{n_t} - 1)E^2X(t_{n_t} - 2)(t - t_{n_t})^2 + \sum_{k=1}^{n_t} H_{t_{k-1}, t_k} + H_{t_{n_t}, t}, \end{aligned}$$

for all $t \in [0, a]$. This implies that

$$\lim_{n \to \infty} nJ^n(t) = 1_{[1,a]}(t) \int_0^t b^2(s)b^2(s-1)EX^2(s-2)/h(s)\, ds, \quad t \in [0, a].$$

Thus

$$\begin{aligned} nE[(Z^\pi)^2(t)] = &\sum_{k=1}^{n_t} b^2(t_{k-1})nE[(Z^\pi)^2(t_{k-1} - 1)](t_k - t_{k-1}) \\ &+ b^2(t_{n_t})nE[(Z^\pi)^2(t_{n_t} - 1)](t - t_{n_t})^2 + nJ^n(t), \end{aligned}$$

for all $t \in [0, a]$. Letting $n \to \infty$ in the above relation yields

$$\mathcal{Z}(t) = \int_0^t b^2(s)\mathcal{Z}(s-1)\, ds + 1_{[1,a]}(t) \int_0^t b^2(s)b^2(s-1)EX^2(s-2)/h(s)\, ds, \quad t \in [0, a].$$

In particular, $\mathcal{Z}(t) = 0$ for all $t \in [-1, 1]$.

From (4.2) it is easy to see that

$$EX^2(t) = \eta(0)^2 + \int_0^t b^2(s)EX^2(s-1)ds, \quad t \in [0, a].$$

Therefore $\mathcal{Z}$ satisfies the assertion of the theorem, and the proof is complete. □

Remark 4.1.

From Theorem 3.1, we know that under some reasonable conditions, the rate of convergence of $X^{\pi_n}(t)$ to $X(t)$ is $1/\sqrt{n}$ over the interval $[-1, a]$. The fact that $Z(t) \equiv 0$ for $0 \le t \le 1$ indicates that on the interval $[0, 1]$ the rate might be eventually higher.

5. Itô's formula for "tame" functions.

In order to derive higher order numerical schemes for SDDE's, we shall first prove an Itô formula for "tame" functions on $C(J, \mathbf{R}^m)$ (Definition 1.1).

Suppose that $(\Omega, \mathcal{F}, P)$ is a probability space and $W(t) := (W^1(t), \cdots, W^d(t))$, $t \ge 0$, is a d-dimensional standard Brownian motion on $(\Omega, \mathcal{F}, P)$. Denote by $D = (D_1, \cdots, D_d)$ the Malliavin differentiation operator associated with $\{W(t) : t \ge 0\}$. Assume

$$X(t) = \begin{cases} \eta(0) + \int_0^t u(s)\, dW(s) + \int_0^t v(s)\, ds, & t > 0, \\ \eta(t), & -r \le t \le 0, \end{cases} \tag{5.1}$$

where η belongs to C and is of bounded variation, $u = (u^1, \cdots, u^m)^T$, $u^i \in \mathbb{L}^{2,4}_{d,loc}$, $v = (v^1, \cdots, v^m)^T$, and $v^i \in \mathbb{L}^{1,4}_{loc}$. One can refer to ([23], pp. 61, 151, 161) for the definition of $\mathbb{L}^{k,p}_d$. Note that the processes u and v may not be adapted to the Brownian filtration $(\mathcal{F}_t)_{t \ge 0}$. For convenience, we define $u(t) = 0$ for $t < 0$ or $t > a$,

$$v(t) = \begin{cases} 0, & t > a \\ \eta'(t), & -r \le t \le 0. \end{cases}$$

We also set $W(t) = 0$ if $t < 0$ or $t > a$, and denote

$$U(t) := \int_0^t u(s)\, dW(s) \text{ and } V(t) := \begin{cases} \eta(0) + \int_0^t v(s)\, ds, & t > 0 \\ \eta(t), & -r \le t \le 0. \end{cases} \tag{5.2}$$

If $u \in \mathbb{L}^{2,p}_{loc}$ for some $p > 4$, then the indefinite Skorohod integral $\int_0^t u(s)\,dW(s)$ has a continuous version. Hence we may assume that the process $X(t)$ is continuous.

Let $T = [0,a]$, $J = [-r,0]$, $C = C(J;,\mathbf{R}^m)$ be as before, and let Π be the projection associated with $s_1, \cdots, s_k \in J$. Although there is a multi-dimensional Itô formula for $\phi(t, X(t))$ ([2], [3] and [22]), we can not apply it to $\phi(t, \Pi(X_t))$ because $\Pi(U_t)$ is of the form

$$\left(\int_0^t u(s+s_1)\,dW(s+s_1), \cdots, \int_0^t u(s+s_k)\,dW(s+s_k)\right), \tag{5.3}$$

and the components of the dk-dimensional process $(W(t+s_1), \cdots, W(t+s_k))$ are not independent. However, the ideas in Nualart and Pardoux ([22], section 6, [23], p. 161) can be used to derive an Itô formula for $\phi(t, \Pi(X_t))$.

We denote by

$$\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases} \tag{5.4}$$

the Kronecker delta.

Assume that $\phi \in C^{1,2}(T \times \mathbf{R}^{mk})$, $\vec{x} = (\vec{x}_1, \cdots, \vec{x}_m)$, $\vec{x}_i = (x_{i1}, \cdots, x_{ik}) \in \mathbf{R}^k$, Write

$$\phi(t, \vec{x}) = \phi(t, \vec{x}_1, \cdots, \vec{x}_m). \tag{5.5}$$

The next lemma follows from the independent increments property of Brownian motion. It will be needed in the proof of the Itô formula for tame functions (Theorem 5.3 below).

Lemma 5.2.

Assume that $\{\pi_n : 0 = t_0 < t_1 < \cdots < t_n = a\}$ *is a family of partitions of* $[0,a]$, *with* $\lim_{n\to\infty} |\pi_n| = 0$. *Let* $-r \leq s_1 \leq s_2 \leq 0$ *and denote by* $\Delta_{lk} W^i := W^i(t_l + s_k) - W^i(t_{l-1} + s_k)$, $1 \leq i \leq d, 1 \leq l \leq n, k = 1,2$, *the increments of Brownian motion. Then*

$$\lim_{n\to\infty} \sum_{l=1}^{n} \Delta_{l1} W^i \Delta_{l2} W^j = \begin{cases} a + s_1, & \text{if } i = j \text{ and } s_1 = s_2 \\ 0, & \text{otherwise}, \end{cases} \tag{5.6}$$

in $L^2(\Omega, \mathbf{R})$.

Proof.

We only need to consider the case $s_1 < s_2$ and $i = j$. Now

$$\left[\sum_{l=1}^{n} \Delta_{l1} W^i \Delta_{l2} W^i\right]^2 = \sum_{l=1}^{n} (\Delta_{l1} W^i)^2 (\Delta_{l2} W^i)^2 + 2 \sum_{l_1 < l_2} \Delta_{l_1 1} W^i \Delta_{l_1 2} W^i \Delta_{l_2 1} W^i \Delta_{l_2 2} W^i.$$

If n is sufficiently large, then $|\pi_n| < s_2 - s_1$. Hence $\Delta_{l_2 2} W^i$ is independent of $\Delta_{l_1 1} W^i \Delta_{l_1 2} W^i \Delta_{l_2 1} W^i$. Taking expectations in the above equality gives

$$E\left[\sum_{l=1}^{n} \Delta_{l1} W^i \Delta_{l2} W^i\right]^2 \le \sum_{l=1}^{n} (t_l - t_{l-1})^2 \le |\pi_n| a.$$

for sufficiently large n. Note that $a + s_1$ is the correct limit in (5.6) because of the convention that $W(t) = 0$ for $t < 0$. This completes the proof of the lemma. $\square$

We now state an *Itô's formula for "tame" functions.*

Theorem 5.3.

Assume that X *is a continuous process defined by (5.1), where* $\eta : J \to \mathbf{R}^m$ *is of bounded variation,* $u = (u^1, \cdots, u^m)^T$, $u^i \in \mathbb{L}^{2,4}_{d,loc}$, $v = (v^1, \cdots, v^m)^T$, *and* $v^i \in \mathbb{L}^{1,4}_{loc}$. *Suppose* $\phi \in C^{1,2}(T \times \mathbf{R}^{mk}, \mathbf{R})$. *Then*

(5.7)
$$\begin{aligned}
&\phi(t, \Pi(X_t)) - \phi(0, \Pi(X_0)) \\
&\qquad = \int_0^t \frac{\partial \phi}{\partial s}(s, \Pi(X_s))\, ds + \int_0^t \frac{\partial \phi}{\partial \vec{x}}(s, \Pi(X_s))\, d(\Pi(X_s)) \\
&\qquad + \frac{1}{2} \sum_{i,j=1}^{k} \sum_{i_1, j_1 = 1}^{m} \int_0^t \frac{\partial^2 \phi}{\partial x_{i,i_1} \partial x_{j,j_1}}(s, \Pi(X_s)) u^{i_1}(s + s_i) D_{s+s_i} X^{j_1}(s + s_j)\, ds
\end{aligned}$$

a.s. for all $t \in T$.

Remark 5.1.

(i) The Itô formula (5.7) may also be expressed in the form

(5.8)
$$\phi(t,\Pi(X_t))-\phi(0,\Pi(X_0))=\int_0^t \frac{\partial\phi}{\partial s}(s,\Pi(X_s))\,ds+\int_0^t \frac{\partial\phi}{\partial\vec{x}}(s,\Pi(X_s))\,d(\Pi(X_s))$$
$$+\frac{1}{2}\sum_{i,j=1}^{k}\int_0^t Tr\left[\frac{\partial^2\phi}{\partial\vec{x}_i\partial\vec{x}_j}(s,\Pi(X_s)(\Theta_s(s_i,s_j))\right]ds$$

where
$$\Theta_s(\alpha,\beta):=\frac{1}{2}\Big\{(u\Lambda)_sX_s(\alpha,\beta)+(u\Lambda)_sX_s(\beta,\alpha)\Big\},\quad \alpha,\beta\in[-r,0],$$

and the two-parameter process $(u\Lambda)_sX_s:\Omega\times J^2\to L(\mathbf{R}^m;\mathbf{R}^m)$ is defined by

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

for all $\alpha,\beta\in[-r,0]$.

(ii) Suppose $d=m=1$. Let us define a trace operator $\bigtriangledown$. For $1\le i,j\le k$, define

$$\bigtriangledown^{\pm}_{s_i,s_j}X(s):=\lim_{\epsilon\downarrow 0}(D_{s+s_i}X(s+s_j+\epsilon)\pm D_{s+s_i}X(s+s_j-\epsilon))\in\mathbf{R} \tag{5.9}$$

and $\bigtriangledown^{\pm}_{s_i}X(s):=(\bigtriangledown^{\pm}_{s_i,s_1}X(s),\cdots,\bigtriangledown^{\pm}_{s_i,s_k}X(s))\in\mathbf{R}^k$. Then the Itô formula for "tame" functions can be written as

(5.10)
$$\phi(t,\Pi(X_t))-\phi(0,\Pi(X_0))=\int_0^t \frac{\partial\phi}{\partial s}(s,\Pi(X_s))\,ds+\int_0^t \frac{\partial\phi}{\partial\vec{x}}(s,\Pi(X_s))\,d\Pi(W_s)$$
$$+\frac{1}{2}\sum_{i=1}^{k}\int_0^t\langle\frac{\partial^2\phi}{\partial x_i^2}(s,\Pi(X_s))\bigtriangledown^{+}_{s_i}X(s),\bigtriangledown^{-}_{s_i}X(s)\rangle_{\mathbf{R}^d}\,ds,$$

a.s. for all $t\in T$, where $\vec{x}:=(x_1,\cdots,x_k)$ and $\langle\cdot,\cdot\rangle_{\mathbf{R}^d}$ denotes the Euclidean inner product on $\mathbf{R}^d$. Cf. [22], Remark 7.6.

For simplicity, we shall prove the Itô formula for the case $d=m=1$. We thus assume in what follows that $d=m=1$.

Proof of Theorem 5.3.

By Taylor's Theorem, we may write

$$\begin{aligned}
\phi(t,\Pi(X_t)) &- \phi(t,\Pi(X_0)) \\
&= \sum_{l=1}^{n} [\phi(t_l,\Pi(X_{t_l})) - \phi(t_{l-1},\Pi(X_{t_l}))] + [\phi(t_{l-1},\Pi(X_{t_l})) - \phi(t_{l-1},\Pi(X_{t_{l-1}}))] \\
&= \sum_{l=1}^{n} \frac{\partial\phi}{\partial s}(\hat{t}_l,\Pi(X_{t_l}))\Delta t_l + \sum_{l=1}^{n}\{\sum_{i=1}^{k}\frac{\partial\phi}{\partial x_i}(t_{l-1},\Pi(X_{t_{l-1}}))\Delta_{li}X \\
&\qquad + \frac{1}{2}\sum_{i,j=1}^{k}\frac{\partial^2\phi}{\partial x_i\partial x_j}(t_{l-1},\Pi(\bar{X}_{t_l}))\Delta_{li}X\Delta_{lj}X\}, \quad t\in T,
\end{aligned}$$

where

$$\bar{X}_{t_l} = X_{t_{l-1}} + \alpha_l(X_{t_l} - X_{t_{l-1}}), \quad \bar{t}_l = t_{l-1} + \beta_l(t_l - t_{l-1}), \quad \hat{t}_l = t_{l-1} + \gamma_l(t_l - t_{l-1})$$

for some random variables $0 \le \alpha_l, \beta_l, \gamma_l \le 1$, $l = 1, \cdots, n$. The Itô formula (5.10) will then follow from Proposition 5.5 and Proposition 5.6 below. □

The rest of this section is devoted to the proofs of Propositions 5.4-5.6.

Proposition 5.4.

Suppose that $W(t)$ is a 1-dimensional Brownian motion. Let $u \in \mathbb{L}^{1,2}_{loc}$ be such that $u(t) = 0$ if $t > a$ or $t < 0$. Assume that $-r \le s_1, s_2 \le 0$, and let $\pi_n : 0 = t_0 < t_1 < \cdots < t_n = a$ be a family of partitions of $T = [0,a]$, with $|\pi_n| \to 0$ as $n \to \infty$. Then

$$\lim_{n\to\infty}\left[\sum_{l=1}^{n}\int_{t_{l-1}+s_1}^{t_l+s_1} u(s)\,dW(s)\right]^2 = \int_0^{a+s_1} u^2(s)\,ds \tag{5.11}$$

in probability. If $s_1 \neq s_2$, then

$$\lim_{n\to\infty}\sum_{l=1}^{n}\int_{t_{l-1}+s_1}^{t_l+s_1} u(s)\,dW(s)\int_{t_{l-1}+s_2}^{t_l+s_2} u(s)\,dW(s) = 0 \tag{5.12}$$

in probability. Furthermore, if $u \in \mathbb{L}^{1,2}$, then the above convergences are in $L^1(\Omega, \mathbf{R})$.

Proof.

We prove the proposition for $u \in \mathbb{L}^{1,2}$. The general case $u \in \mathbb{L}^{1,2}_{loc}$ follows by a standard localization argument ([23]).

If $u_i, u_j, v_i, v_j \in \mathbb{L}^{1,2}$ with $u_i(t) = v_i(t) = 0$ if $t < 0$ or $t > a + s_i$ and $u_j(t) = v_j(t) = 0$ if $t < 0$ or $t > a + s_j$. Set

$$\left\{ \begin{array}{l} U_i(t) := \int_0^t u_i(s)\,dW(s) \\ U_j(t) := \int_0^t u_j(s)\,dW(s) \end{array} \right. \qquad \left\{ \begin{array}{l} V_i(t) := \int_0^t v_i(s)\,dW(s) \\ V_j(t) := \int_0^t v_j(s)\,dW(s). \end{array} \right. \tag{5.13}$$

Then

$$\begin{aligned} E|\sum_{l=1}^n \Delta_{li} U_i \Delta_{lj} U_j - \sum_{l=1}^n \Delta_{li} V_i \Delta_{lj} V_j| \\ &= E|\sum_{l=1}^n \Delta_{li}(U_i - V_i)\Delta_{lj} U_j + \sum_{l=1}^n \Delta_{li} V_i \Delta_{lj}(U_j - V_j)| \\ &\le E|\sum_{l=1}^n \Delta_{li}(U_i - V_i)\Delta_{lj} U_j| + E|\sum_{l=1}^n \Delta_{li} V_i \Delta_{lj}(U_j - V_j)| \\ &\le (E\sum_{l=1}^n |\Delta_{li}(U_i - V_i)|^2)^{\frac{1}{2}} (E\sum_{l=1}^n |\Delta_{lj}(U_j)|^2)^{\frac{1}{2}} \\ &+ (E\sum_{l=1}^n |\Delta_{li}(V_i)|^2)^{\frac{1}{2}} (E\sum_{l=1}^n |\Delta_{lj}(U_j - V_j)|^2)^{\frac{1}{2}}. \end{aligned}$$

By an L^p estimate of the Skorohod integral ([22], Proposition 3.5; [23], p.158), we have

$$\begin{aligned} E\sum_{l=1}^n |\Delta_{lj} U_j|^2 &= E\sum_{l=1}^n |\int_{t_{l-1}+s_i}^{t_l+s_i} u_j(s)\,dW(s)|^2 \\ &= E\sum_{l=1}^n |\int_0^a I_{(t_{l-1}+s_i, t_l+s_i]}(s) u_j(s)\,dW(s)|^2 \\ &\le \sum_{l=1}^n \int_0^a I_{(t_{l-1}+s_i, t_l+s_i]}(s) E u_j^2(s)\,ds \\ &+ \sum_{l=1}^n \int_0^a \int_0^a I_{(t_{l-1}+s_i, t_l+s_i]}(s) E(D_t u_j(s))^2\,ds\,dt \\ &= \int_0^a E u_j^2(s)\,ds + \int_0^a \int_0^a E(D_t u_j(s))^2\,ds\,dt \\ &= ||u_j||_{1,2}^2. \end{aligned}$$

Hence we obtain the following inequality

$$(5.14)\quad E|\sum_{l=1}^n \Delta_{li}U_i\Delta_{lj}U_j - \sum_{l=1}^n \Delta_{li}V_i\Delta_{lj}V_j| \le ||u_i - v_i||_{1,2}||u_j||_{1,2} + ||v_i||_{1,2}||u_j - v_j||_{1,2}.$$

Since $\mathbb{L}^{1,2} \cap L^4(\Omega \times [0,a])$ is dense in $\mathbb{L}^{1,2}$, it suffices to prove (5.12) for the case $u \in \mathbb{L}^{1,2} \cap L^4(\Omega \times [0,a])$. Set

$$(5.15)\qquad u_i(t) := \begin{cases} u(t), & 0 \le t \le a + s_i \\ 0, & t < 0 \text{ or } t > a + s_i. \end{cases}$$

Define

$$(5.16)\qquad u_i^n(t) := \sum_{l=1}^n \frac{I_{(t_{l-1}+s_i, t_l+s_i]}(t)}{t_l - t_{l-1}} \int_{t_{l-1}+s_i}^{t_l+s_i} u(s)\, ds.$$

and u_j^n similarly. Let

$$(5.17)\qquad \begin{cases} U_i(t) := \int_0^t u_i(s)\, dW(s) \\ U_j(t) := \int_0^t u_j(s)\, dW(s) \end{cases} \qquad \begin{cases} U_i^n(t) := \int_0^t u_i^n(s)\, dW(s) \\ V_j^n(t) := \int_0^t u_j^n(s)\, dW(s). \end{cases}$$

Using (5.14) it is easy to check that

$$(5.18)\qquad \lim_{n\to\infty} E|\sum_{l=1}^n \Delta_{li}U_i^n\Delta_{lj}U_j^n - \sum_{l=1}^n \Delta_{li}U_i\Delta_{lj}U_j| = 0.$$

By the formula for the Skorohod integral of a process multiplied by a random variable ([22], Theorem 3.2), we get

$$\begin{aligned} \Delta_{li}U_i^n &= \int_{t_{l-1}+s_i}^{t_l+s_i} \sum_{k=1}^n \frac{I_{(t_{k-1}+s_i, t_k+s_i]}(t)}{t_k - t_{k-1}} \int_{t_{k-1}+s_i}^{t_k+s_i} u_i(s)\, ds\, dW(t) \\ &= \frac{1}{t_l - t_{l-1}} \int_{t_{l-1}+s_i}^{t_l+s_i} u_i(s)\, ds[W(t_l+s_i) - W(t_{l-1}+s_1)] \\ &+ \frac{1}{t_l - t_{l-1}} \int_{t_{l-1}+s_i}^{t_l+s_i} \int_{t_{l-1}+s_i}^{t_l+s_i} D_t u_i(s)\, ds\, dt \\ &= P_{li}\Delta_{li}W + Q_{li}. \end{aligned}$$

where

$$P_{li} := \frac{1}{t_l - t_{l-1}} \int_{t_{l-1}+s_i}^{t_l+s_i} u_i(s)\, ds, \quad Q_{li} := \frac{1}{t_l - t_{l-1}} \int_{t_{l-1}+s_i}^{t_l+s_i} \int_{t_{l-1}+s_i}^{t_l+s_i} D_t u_i(s)\, ds\, dt.$$

Therefore,

$$\begin{aligned}\sum_{l=1}^{n}\Delta_{li}U_i^n\Delta_{lj}U_j^n &= \sum_{l=1}^{n}(P_{li}\Delta_{li}W+Q_{li})(P_{lj}\Delta_{lj}W+Q_{lj})\\ &= \sum_{l=1}^{n}(P_{li}P_{lj})(\Delta_{li}W\Delta_{lj}W)+\sum_{l=1}^{n}(P_{li}Q_{lj})\Delta_{li}W\\ &+\sum_{l=1}^{n}(P_{lj}Q_{li})\Delta_{lj}W+\sum_{l=1}^{n}Q_{li}Q_{lj}.\end{aligned}$$

By Hölder's inequality,

$$\sum_{l=1}^{n}Q_{li}^2\le\sum_{l=1}^{n}\int_{t_{l-1}+s_i}^{t_l+s_i}\int_{t_{l-1}+s_i}^{t_l+s_i}|D_tu_i(s)|^2\,ds\,dt\,. \tag{5.19}$$

Thus $\lim_{n\to\infty}E\sum_{l=1}^{n}Q_{li}^2=0$. Now

$$\begin{aligned}\sum_{l=1}^{n}(P_{li}\Delta_{li}W)^2 &= \sum_{l=1}^{n}\frac{(\Delta_{li}W)^2}{(t_l-t_{l-1})^2}\left(\int_{t_{l-1}+s_i}^{t_l+s_i}u_i(s)\,ds\right)^2\\ &= \sum_{l=1}^{n}\frac{(\Delta_{li}W)^2}{t_l-t_{l-1}}\int_{t_{l-1}+s_i}^{t_l+s_i}(u_i^n(s))^2\,ds\,.\end{aligned}$$

It is easy to check that $E||(u_i^n)^2||_{L^2([0,a+s_i])}\le E||u_i^2||_{L^2([0,a+s_i])}$ and

$$\lim_{n\to\infty}E||(u_i^n)^2-u_i^2||_{L^2([0,a+s_i])}=0\,. \tag{5.20}$$

By an argument similar to the one used in the proof of Lemma A.1, we can show that $\{\sum_{l=1}^{n}(P_{li}\Delta_{li}W)^2, n\ge 1\}$ is uniformly integrable. Applying Lemma A.1, we have

$$\lim_{n\to\infty}E|\sum_{l=1}^{n}(P_{li}\Delta_{li}W)^2-\int_0^{a+s_i}u_i^2(s)\,ds|=0. \tag{5.21}$$

The Cauchy-Schwartz type inequality

$$E\left|\sum_{l=1}^{n}(P_{li}\Delta_{li}W)Q_{li}\right|\le\sqrt{E\sum_{l=1}^{n}(P_{li}\Delta_{li}W)^2E\sum_{l=1}^{n}Q_{li}^2} \tag{5.22}$$

together with (5.19) and (5.21) implies that $\lim_{n\to\infty} E|\sum_{l=1}^{n}(P_{li}\Delta_{li}W)Q_{li}| = 0$.

Now consider the case $i \neq j$. The Cauchy-Schwartz inequality implies

$$E|\sum_{l=1}^{n} Q_{lj}Q_{li}| \leq \sqrt{E\sum_{l=1}^{n} Q_{lj}^2 E\sum_{l=1}^{n} Q_{li}^2}\,. \tag{5.23}$$

We may write

$$\begin{aligned}\sum_{l=1}^{n}(P_{li}P_{lj})(\Delta_{li}W\Delta_{lj}W) &= \sum_{l=1}^{n}\frac{\Delta_{li}W\Delta_{lj}W}{(t_l-t_{l-1})^2}\int_{t_{l-1}+s_i}^{t_l+s_i} u_i(s)\,ds\int_{t_{l-1}+s_j}^{t_l+s_j} u_j(s)\,ds\\ &= \sum_{l=1}^{n}\frac{\Delta_{li}W\Delta_{lj}W}{t_l-t_{l-1}}\int_{t_{l-1}+s_i}^{t_l+s_i} u_i^n(s)\hat{u}_j^n(s)\,ds\,,\end{aligned} \tag{5.24}$$

where

$$\hat{u}_j^n(s) = \sum_{l=1}^{m}\frac{I_{(t_{l-1}+s_i,t_l+s_i]}(s)}{t_l-t_{l-1}}\int_{t_{l-1}+s_i}^{t_l+s_i} u_j(s'+s_j-s_i)\,ds'\,. \tag{5.25}$$

Similar to the case $i = j$, we have

$$\lim_{n\to\infty} E|\sum_{l=1}^{n}(P_{li}P_{lj})(\Delta_{li}W\Delta_{lj}W)| = 0. \tag{5.26}$$

This completes the proof of the proposition. □

Suppose that

$$\bar{X}_{t_l} = X_{t_{l-1}} + \alpha_l(X_{t_l} - X_{t_{l-1}}), \quad \bar{t}_l = t_{l-1} + \beta_l(t_l - t_{l-1})$$

for some random variables $0 \leq \alpha_l, \beta_l, \leq 1$, $l = 1, \cdots, n$. Denote

$$\Delta(\Pi(X_{t_l})) = (\Pi(\Delta X_{t_l})) = \Pi(X_{t_l}) - \Pi(X_{t_{l-1}}), \tag{5.27}$$

$$\Pi(\bar{X}_{t_l}) = \Pi(X_{t_{l-1}}) + \alpha_l\Delta\Pi(X_{t_l}), \tag{5.28}$$

$$\Delta_{li}X = X(t_l + s_i) - X(t_{l-1} + s_i), \text{ for } 1 \leq i \leq k \text{ and } 1 \leq l \leq n. \tag{5.29}$$

Proposition 5.5.

Suppose that $\phi \in C^{1,2}(T \times \mathbf{R}^k, \mathbf{R})$, and let $1 \leq i, j \leq k$. Under the hypotheses of Proposition 5.4, we have

$$\sum_{l=1}^{n} \frac{\partial^2 \phi}{\partial x_i \partial x_j}(t_{l-1}, \Pi(\bar{X}_{t_l}))\Delta_{li} X \Delta_{lj} X \to \begin{cases} \int_0^{t+s_i} \frac{\partial^2 \phi}{\partial x_i^2}(s, \Pi(X_s))u^2(s)\, ds, & i = j \\ 0, & i \neq j \end{cases} \tag{5.30}$$

as $n \to \infty$, in probability.

Proof.

For $0 \leq i, j \leq n$,

$$\begin{aligned} \Delta_{li} X \Delta_{lj} X &= (\Delta_{li} U + \Delta_{li} V)(\Delta_{lj} U + \Delta_{lj} V) \\ &= \Delta_{li} U \Delta_{lj} U + \Delta_{li} U \Delta_{lj} V + \Delta_{li} V \Delta_{lj} U + \Delta_{li} V \Delta_{lj} V, \end{aligned} \tag{5.31}$$

where U, V are defined by (5.2). Since U, V are continuous and V is of bounded variation, it follows that

$$\begin{cases} \lim_{n\to\infty} \sum_{l=1}^{n} \Delta_{li} U \Delta_{lj} V = 0 \\ \lim_{n\to\infty} \sum_{l=1}^{n} \Delta_{li} V \Delta_{lj} U = 0 \\ \lim_{n\to\infty} \sum_{l=1}^{n} \Delta_{li} V \Delta_{lj} V = 0 \end{cases} \tag{5.32}$$

in probability, for all $0 \leq i, j \leq n$. To handle the term $\sum_{l=1}^{n} \Delta_{li} U \Delta_{lj} U$, we adapt an approach by Nualart and Pardoux (c.f. [22] Theorem 5.4 or [23] Theorem 3.2.1).

Set $Y(s) := \frac{\partial^2 \phi}{\partial x_i^2}(s, \Pi(X_s)) I_{[0,t]}(s)$ and

$$Y^n(s) := Y(0) I_{\{0\}}(s) + \sum_{l=1}^{n} \frac{\partial^2 \phi}{\partial x_i^2}(t_{l-1}, \Pi(\bar{X}_{t_l})) I_{(t_{l-1}, t_l]}(s). \tag{5.33}$$

Then $Y^n(s) \to Y(s)$ as $n \to \infty$, uniformly in $s \in [0, t]$. Applying Proposition 5.4 and Lemma A.2, we get

$$\sum_{l=1}^{n} \frac{\partial^2 \phi}{\partial x_i \partial x_j}(t_{l-1}, \Pi(\bar{X}_{t_l}))\Delta_{li} X \Delta_{lj} X \to \delta_{ij} \int_0^{t+s_i} \frac{\partial^2 \phi}{\partial x_i^2}(s, \Pi(X_s)) u^2(s)\, ds \tag{5.34}$$

in probability as $n \to \infty$. $\square$

Proposition 5.6.

Suppose that $\phi \in C^{1,2}(T \times \mathbf{R}^k)$ and let $X(t)$ be a continuous stochastic process defined by (5.1), where $u \in \mathbb{L}^{2,4}_{loc}$, $v \in \mathbb{L}^{1,4}_{loc}$, and $\eta \in C([-r,0],\mathbf{R}^m)$ is of bounded variation. Assume that $\pi_n : -r = s_0 < \cdots < s_n = 0$ are partitions of $[-r,0]$ such that $|\pi_n| \to 0$ as $n \to \infty$. Then, for each $1 \le i \le k$ and each $t \in T$, we have

(5.35)

$$\begin{aligned}
&\lim_{n\to\infty} \sum_{l=1}^{n} \frac{\partial\phi}{\partial x_i}(t_{l-1},\Pi(X_{t_{l-1}}))\Delta_{li}X \\
&= \int_0^t \frac{\partial\phi}{\partial x_i}(s,\Pi(X_s))\,dX(s+s_i) + \sum_{j=i+1}^{k}\int_0^t \frac{\partial^2\phi}{\partial x_i\partial x_j}(s,\Pi(X_s))u^2(s+s_i)\,ds \\
&+ \sum_{j=1}^{k}\int_0^t \frac{\partial^2\phi}{\partial x_i\partial x_j}(s,\Pi(X_s))\left[\int_0^{s+s_j} D_{s+s_i}u(r)\,dW(r) + \int_0^{s+s_j} D_{s+s_i}v(r)\,dr\right]u(s+s_i)\,ds
\end{aligned}$$

in probability.

Proof.

By a localization argument, we may assume that $\phi \in C_b^{1,2}(T \times \mathbf{R}^k, \mathbf{R})$. Let $|\pi_n| < \min_{\{1\le i\le k\}} |s_i - s_{i-1}|$. Fix $1 \le i \le k, 1 \le l \le n$, and set

$$F_l := \frac{\partial\phi}{\partial x_i}(t_{l-1},\Pi(X_{t_{l-1}})). \tag{5.36}$$

Using an integration by parts formula ([22], Theorem 3.2), it follows that

$$F_l\Delta_{li}U = \int_{t_{l-1}+s_i}^{t_l+s_i} u(s)F_l\,dW(s) + \int_{t_{l-1}+s_i}^{t_l+s_i} D_r(F_l)u(r)\,dr, \tag{5.37}$$

where U is defined by (5.2). The chain rule yields

$$D_r(F_l) = \sum_{j=1}^{k} \frac{\partial^2\phi}{\partial x_i\partial x_j}(t_{l-1},\Pi(X_{t_{l-1}}))D_rX(t_{l-1}+s_j). \tag{5.38}$$

Now, taking the Malliavin derivative D_r in (5.1) gives

$$D_rX(t) = u(r)I_{\{r\le t\}} + \int_0^t D_ru(s)\,dW(s) + \int_0^t D_rv(s)\,ds, \tag{5.39}$$

Consequently

$$\sum_{l=1}^{n} \frac{\partial \phi}{\partial x_i}(t_{l-1}, \Pi(X_{t_{l-1}}))\Delta_{li} U = c_1 + c_2 + c_3 + c_4,$$

where

$$(5.40)\quad \begin{cases} c_1 := \sum_{l=1}^{n} \int_{t_{l-1}+s_i}^{t_l+s_i} \frac{\partial \phi}{\partial x_i}(t_{l-1}, \Pi(X_{t_{l-1}}))u(s)\, dW(s) \\ c_2 := \sum_{l=1}^{n} \int_{t_{l-1}+s_i}^{t_l+s_i} \sum_{j=1}^{k} \frac{\partial^2 \phi}{\partial x_i \partial x_j}(t_{l-1}, \Pi(X_{t_{l-1}}))I_{\{r \le t_{l-1}+s_j\}}u^2(r)\, dr \\ c_3 := \sum_{l=1}^{n} \int_{t_{l-1}+s_i}^{t_l+s_i} \sum_{j=1}^{k} \frac{\partial^2 \phi}{\partial x_i \partial x_j}(t_{l-1}, \Pi(X_{t_{l-1}})) \int_0^{t_{l-1}+s_j} D_r u(s)\, dW(s) u(r)\, dr \\ c_4 := \sum_{l=1}^{n} \int_{t_{l-1}+s_i}^{t_l+s_i} \sum_{j=1}^{k} \frac{\partial^2 \phi}{\partial x_i \partial x_j}(t_{l-1}, \Pi(X_{t_{l-1}})) \int_0^{t_{l-1}+s_j} D_r v(s)\, ds u(r)\, dr. \end{cases}$$

We will study the limits of the above expressions as $n \to \infty$.

Step 1. First we show that the limit of c_2 is given by

$$(5.41)\qquad c_2 \to \sum_{j=i+1}^{k} \int_0^{t+s_i} \frac{\partial^2 \phi}{\partial x_i \partial x_j}(r - s_i, \Pi(X_{r-s_i}))u^2(r)\, dr, \qquad \text{a.s.}$$

If $j \le i$, then $t_{l-1} + s_i \ge t_{l-1} + s_j$. So when $t_{l-1} + s_i < r < t_l + s_i$, $I_{\{r \le t_l + s_j\}} = 0$. We have

$$\begin{aligned} c_2 &= \sum_{j=i+1}^{k} \sum_{l=1}^{n} \int_{t_{l-1}+s_i}^{t_l+s_i} \frac{\partial^2 \phi}{\partial x_i \partial x_j}(t_{l-1}, \Pi(X_{t_{l-1}}))I_{\{r \le t_{l-1}+s_j\}}u^2(r)\, dr \\ &\to \sum_{j=i+1}^{k} \int_0^{t+s_i} \frac{\partial^2 \phi}{\partial x_i \partial x_j}(r - s_i, \Pi(X_{r-s_i}))u^2(r)\, dr \end{aligned}$$

a.s. as $n \to \infty$.

Step 2. Next we study the limit of c_3 as $n \to \infty$. We claim that

$$(5.42)\qquad c_3 \to \sum_{j=1}^{k} \int_0^{t+s_i} \frac{\partial^2 \phi}{\partial x_i \partial x_j}(r - s_i, \Pi(X_{r-s_i})) \int_0^{r-s_i+s_j} D_r u(s)\, dW(s) u(r)\, dr$$

as $k \to \infty$ in probability. In fact,

$$
\begin{aligned}
T_j^n &:= |\sum_{l=1}^n \int_{t_{l-1}+s_i}^{t_l+s_i} [\frac{\partial^2\phi}{\partial x_i \partial x_j}(t_{l-1}, \Pi(X_{t_{l-1}})) \int_0^{t_{l-1}+s_j} D_r u(s)\, dW(s) \\
&- \int_0^{t+s_i} \frac{\partial^2\phi}{\partial x_i \partial x_j}(r-s_i, \Pi(X_{r-s_i})) \int_0^{r-s_i+s_j} D_r u(s)\, dW(s)] u(r)\, dr| \\
&\le |\sum_{l=1}^n \int_{t_{l-1}+s_i}^{t_l+s_i} \frac{\partial^2\phi}{\partial x_i \partial x_j}(t_{l-1}, \Pi(X_{t_{l-1}})) \int_{t_{l-1}+s_j}^{r+s_j-s_i} D_r u(s)\, dW(s)) u(r)\, dr| \\
&+ |\sum_{l=1}^n \int_{t_{l-1}+s_i}^{t_l+s_i} [\frac{\partial^2\phi}{\partial x_i \partial x_j}(t_{l-1}, \Pi(X_{t_{l-1}})) - \frac{\partial^2\phi}{\partial x_i \partial x_j}(r-s_i, \Pi(X_{r-s_i}))] \\
&\times \int_0^{r-s_i+s_j} D_r u(s)\, dW(s) u(r)\, dr| \\
&\le \|\frac{\partial^2\phi}{\partial x_i \partial x_j}\|_\infty \sum_{l=1}^n \int_{t_{l-1}+s_i}^{t_l+s_i} |\int_{t_{l-1}+s_j}^{r+s_j-s_i} D_r u(s)\, dW(s))||u(r)|\, dr \\
&+ \sup_{1\le l\le n} \sup_{r\in[t_{l-1}+s_i, t_l+s_i]} |\frac{\partial^2\phi}{\partial x_i \partial x_j}(t_{l-1}, \Pi(X_{t_{l-1}})) - \frac{\partial^2\phi}{\partial x_i \partial x_j}(r-s_i, \Pi(X_{r-s_i}))| \\
&\times \int_0^{t_l+s_i} |\int_0^{r-s_i+s_j} D_r u(s)\, dW(s) u(r)|\, dr \\
&= T_{j1}^n + T_{j2}^n,
\end{aligned}
$$

where T_{j1}^n and T_{j2}^n denote the first and second term on the right hand side of the last inequality. Using the Cauchy-Schwartz inequality and the L^p inequality for Skorohod integral ([22] Proposition 3.5, [23] p.158), we have

$$
\begin{aligned}
ET_{j1}^n \le& \left\| \frac{\partial^2\phi}{\partial x_i \partial x_j} \right\|_\infty (E \int_0^{a+s_i} u^2(r)\, dr)^{\frac{1}{2}} \\
&\times \{E \sum_{l=1}^n \int_{t_{l-1}+s_i}^{t_l+s_i} \int_{t_{l-1}+s_j}^{r+s_j-s_i} |D_r u(s)|^2\, ds\, dr \\
&+ E \sum_{l=1}^n \int_{t_{l-1}+s_i}^{t_l+s_i} \int_{t_{l-1}+s_j}^{r+s_j-s_i} \int_0^a |D_\theta(D_r u(s))|^2\, d\theta\, ds\, dr\}^{\frac{1}{2}} \\
&\to 0
\end{aligned}
$$

as $n \to \infty$. The uniform continuity of $\frac{\partial^2\phi}{\partial x_i \partial x_j}$ implies $T_{j2}^n \to 0$ a.s. So as $n \to \infty$, $T_j^n \to 0$ in probability.

Step 3. Now we will show that

$$c_4 \to \sum_{j=1}^{k} \int_0^{t+s_i} \frac{\partial^2 \phi}{\partial x_i \partial x_j}(r - s_i, \Pi(X_{r-s_i})) \int_0^{r+s_j-s_i} D_r v(s)\, ds u(r)\, dr, \quad \text{a.s.} \tag{5.43}$$

As in Step 2, we have

$$\begin{aligned}
&|\sum_{l=1}^{n} \int_{t_{l-1}+s_i}^{t_l+s_i} \Big[\frac{\partial^2 \phi}{\partial x_i \partial x_j}(t_{l-1}, \Pi(X_{t_{l-1}})) \int_0^{t_{l-1}+s_j} D_r v(s)\, ds \\
&\qquad - \frac{\partial^2 \phi}{\partial x_i \partial x_j}(r - s_i, \Pi(X_{r-s_i})) \int_0^{r-s_i+s_j} D_r v(s)\, ds \Big] u(r)\, dr| \\
&\le \Big\| \frac{\partial^2 \phi}{\partial x_i \partial x_j} \Big\|_\infty \sum_{l=1}^{n} \int_{t_{l-1}+s_i}^{t_l+s_i} |\int_{t_{l-1}+s_j}^{r+s_j-s_i} D_r v(s)\, ds||u(r)|\, dr \\
&+ \sup_{1\le l\le n} \sup_{r\in[t_{l-1}+s_i, t_l+s_i]} \Big| \frac{\partial^2 \phi}{\partial x_i \partial x_j}(t_{l-1}, \Pi(X_{t_{l-1}})) - \frac{\partial^2 \phi}{\partial x_i \partial x_j}(r - s_i, \Pi(X_{r-s_i})) \Big| \\
&\times \int_0^{t_l+s_i} \Big| \int_0^{r-s_i+s_j} D_r v(s)\, ds \Big| |u(r)|\, dr \\
&\to 0 \quad \text{a.s as} \quad n \to \infty\,.
\end{aligned}$$

Step 4. Finally we study the limit of c_1 as $n \to \infty$. We shall show that

$$c_1 \to \int_0^{t+s_i} \frac{\partial \phi}{\partial x_i}(s - s_i, \Pi(X_{s-s_i}))u(s)\, dW(s) \tag{5.44}$$

in $L^2(\Omega, \mathbf{R})$ as $n \to \infty$. To see this, define

$$u^n(s) = u(s) \sum_{l=1}^{n} \frac{\partial \phi}{\partial x_i}(t_{l-1}, \Pi(X_{t_{l-1}}))I_{(t_{l-1}+s_i, t_l+s_i]}(s). \tag{5.45}$$

It suffices to show that

$$u^n(s) \to \frac{\partial \phi}{\partial x_i}(s - s_i, \Pi(X_{t_{s-s_i}}))u(s)I_{(0,t+s_i]}(s) \tag{5.46}$$

in $\mathbb{L}^{1,2}$ as $n \to \infty$. It is clear that the sequence $\{u^n(s)\}$ converges to $\frac{\partial \phi}{\partial x_i}(s - s_i, \Pi(X_{s-s_i}))u(s)I_{(0,t+s_i]}(s)$ in $L^2(\Omega \times T, \mathbf{R})$. It remains to show that the sequence

$\{D_r u^n(s)\}_{n=1}^\infty, r, s \in T^2$, converges in $L^2(\Omega\times T^2, \mathbf{R})$ to $D_r\left[\frac{\partial\phi}{\partial x_i}(s-s_i, \Pi(X_{s-s_i}))u(s)I_{(0,t+s_i]}(s)\right]$. Now

$$\begin{aligned}
&D_r u^n(s)\\
&= D_r u(s)\sum_{l=1}^n \frac{\partial\phi}{\partial x_i}(t_{l-1}, \Pi(X_{t_{l-1}}))I_{(t_{l-1}+s_i, t_l+s_i]}(s)\\
&+ u(s)\sum_{l=1}^n[\sum_{j=1}^k \frac{\partial^2\phi}{\partial x_i\partial x_j}(t_{l-1}, \Pi(X_{t_{l-1}}))\int_0^{t_{l-1}+s_j} D_r u(s')\,dW(s')]I_{(t_{l-1}+s_i, t_l+s_i]}(s)\\
&+ u(s)\sum_{l=1}^n[\sum_{j=1}^k \frac{\partial^2\phi}{\partial x_i\partial x_j}(t_{l-1}, \Pi(X_{t_{l-1}}))\int_0^{t_{l-1}+s_j} D_r v(s')\,ds']I_{(t_{l-1}+s_i, t_l+s_i]}(s)\\
&+ u(s)\sum_{l=1}^n[\sum_{j=1}^k \frac{\partial^2\phi}{\partial x_i\partial x_j}(t_{l-1}, \Pi(X_{t_{l-1}}))u(r)I_{[0,t_{l-1}+s_j]}(r)]I_{(t_{l-1}+s_i, t_l+s_i]}(s)\\
&= d_1 + d_2 + d_3 + d_4.
\end{aligned}$$

where d_1, d_2, d_3, d_4 stand for the first, second, third and fourth term, respectively, on the right hand side of the above equality. It is easy to see that

$$d_1 \to D_r u(s)\frac{\partial\phi}{\partial x_i}(\Pi(s-s_i, X_{s-s_i}))I_{(0,t+s_i]}(s)$$

in $L^2(\Omega, \mathbf{R})$. Since for all $1 \leq j \leq k$, $u(s)\int_0^{s+s_j-s_i} D_r v(\theta)\,d\theta$ belongs to $L^2(\Omega \times T^2, \mathbf{R})$, then by Lebesgue's dominated convergence theorem, we have

$$\begin{aligned}
d_3 :=& u(s)\sum_{j=1}^k\sum_{l=1}^n[\frac{\partial^2\phi}{\partial x_i\partial x_j}(t_{l-1}, \Pi(X_{t_{l-1}}))\int_0^{s+s_j-s_i} D_r v(\theta)\,d\theta]I_{(t_{l-1}+s_i, t_l+s_i]}(s)\\
&\to \sum_{j=1}^k u(s)[\frac{\partial^2\phi}{\partial x_i\partial x_j}(s-s_i, \Pi(X_{s-s_i}))\int_0^{s+s_j-s_i} D_r v(\theta)\,d\theta]I_{(0,t+s_i]}(s)
\end{aligned}$$

in $L^2(\Omega \times T^2, \mathbf{R})$. Moreover,

$$\begin{aligned}
&\sum_{l=1}^n \int_{t_{l-1}+s_i}^{t_l+s_i}\int_0^a u^2(s)\left[\frac{\partial^2\phi}{\partial x_i\partial x_j}(t_{l-1}, \Pi(X_{t_{l-1}}))\right]^2[\int_{t_{l-1}+s_j}^{s+s_j-s_i} D_r v(\theta)\,d\theta]^2\,dr\,ds\\
&\leq |\pi_n|\left\|\frac{\partial^2\phi}{\partial x_i\partial x_j}\right\|_\infty^2\int_0^a u^2(s)\,ds\int_0^a\int_0^a (D_r v(\theta))^2\,dr\,d\theta\\
&\to 0
\end{aligned}$$

as $n \to \infty$ in $L^1(\Omega, \mathbf{R})$, because $v \in \mathbb{L}^{1,4}$ and $u \in L^4(\Omega \times T, \mathbf{R})$. Hence,

$$d_3 \to \sum_{j=1}^{k} u(s)[\frac{\partial^2 \phi}{\partial x_i \partial x_j}(s - s_i, \Pi(X_{s-s_i})) \int_0^{s+s_j-s_i} D_r v(\theta)\, d\theta] I_{(0,t+s_i]}(s)$$

in $L^2(\Omega \times T^2, \mathbf{R})$. To find the limit of d_2, we need to check that for all j, the two parameter process $\big(u(s) \int_0^{s+s_j-s_i} D_r u(\theta)\, dW(\theta)\,, 0 \leq s, r \leq a\big)$ belongs to $L^2(\Omega \times T^2, \mathbf{R})$. This follows from the following estimates:

$$\begin{aligned}
&E \int_0^a \int_0^a u^2(s) [\int_0^{s+s_j-s_i} D_r u(\theta)\, dW(\theta)]^2\, ds\, dr \\
&\leq \left\{ E \int_0^a u^4(s)\, ds E \int_0^a \{ \int_0^a [\int_0^{s+s_j-s_i} D_r u(\theta)\, dW(\theta)]^2\, dr\}^2\, ds \right\}^{\frac{1}{2}} \\
&\leq C\{E \int_0^a u^4(s)\, ds [E(\int_0^a \int_0^a |D_r u(\theta)|^2\, d\theta\, dr)^2 \\
&+ E(\int_0^a \int_0^a \int_0^a D_\alpha(D_r u(\theta))\, d\theta\, dr\, d\alpha)^2]\}^{\frac{1}{2}}.
\end{aligned}$$

Here we have used a slight modification of the L^p estimate of the Skorohod integral for $p = 4$ (c.f. [23], Exercise 3.2.7). Using similar L^p estimates to the above, we obtain

(5.47)

$$\begin{aligned}
&\sum_{l=1}^{n} \int_{t_{l-1}+s_i}^{t_l+s_i} \int_0^a u^2(s) \left[\frac{\partial^2 \phi}{\partial x_i \partial x_j}(t_{l-1}, \Pi(X_{t_{l-1}}))\right]^2 [\int_{t_{l-1}+s_j}^{s+s_j-s_i} D_r u(\theta)\, dW(\theta)]^2\, dr\, ds \\
&\leq \left\| \frac{\partial^2 \phi}{\partial x_i \partial x_j} \right\|_\infty^2 \left(\int_0^a E u^4(s)\, ds \right)^{\frac{1}{2}} \times \left\{ \sum_{l=1}^{n} E \int_{t_{l-1}+s_i}^{t_l+s_i} \left[\int_0^a (\int_{t_{l-1}+s_j}^{s+s_j-s_i} D_r v(\theta)\, d\theta)^2\, dr \right]^2 ds \right\}^{\frac{1}{2}}.
\end{aligned}$$

Note that the right hand side of the above inequality tends to zero as $n \to \infty$. Thus

$$d_2 \to \sum_{j=1}^{k} u(s)[\frac{\partial^2 \phi}{\partial x_i \partial x_j}(s - s_i, \Pi(X_{s-s_i})) \int_0^{s+s_j-s_i} D_r u(\theta)\, dW(\theta)] I_{(0,t+s_i]}(s) \tag{5.48}$$

in $L^2(\Omega \times T^2, \mathbf{R})$ as $n \to \infty$.

It is easy to check that

$$d_4 \to \sum_{j=1}^{k} u(s) \frac{\partial^2 \phi}{\partial x_i \partial x_j}(s - s_i, \Pi(X_{s-s_i})) u(r) I_{[0,s+s_j-s_i]}(r) I_{(0,t+s_i]}(s)$$

as $n \to \infty$ in $L^2(\Omega, \mathbf{R})$. Therefore,

$$D_r u^n(s) \to D_r[u(s)\frac{\partial^2\phi}{\partial x_i \partial x_j}(s - s_i, \Pi(X_{s-s_i}))I_{(0,t+s_i]}(s)] \tag{5.49}$$

in $L^2(\Omega \times T^2, \mathbf{R})$. Finally it is easy to see that

$$c_1 \to \int_0^{t+s_i} \frac{\partial\phi}{\partial x_i}(s - s_i, \Pi(X_{s-s_i}))u(s)\, dW(s)$$

in $L^2(\Omega, \mathbf{R})$ as $n \to \infty$.

<u>*Step* 5.</u> The convergence

$$\sum_{l=1}^{n} \frac{\partial\phi}{\partial x_i}(t_{l-1}, \Pi(X_{t_{l-1}}))\Delta_{li} V \to \int_{s_i}^{t+s_i} \frac{\partial\phi}{\partial x_i}(s - s_i, \Pi(X_{s-s_i}))\, dV(s) \text{ a.s.} \tag{5.50}$$

as $n \to \infty$, is easy to verify. □

We complete the section by giving a Stratonovich version of the Itô formula (5.7).

Suppose that $k \geq 1$ and $p \geq 2$. The set $\mathbb{L}_{d,C}^{k,p}$ (c.f. [22] Definition 7.2, [23] p.167) is the class of processes $u \in L_d^{k,p}$ such that the mappings $s \hookrightarrow D_{s\wedge t}u(s \vee t)$ and $s \hookrightarrow D_{s\vee t}u(s \wedge t)$ are continuous in $L^p(\Omega)$, uniformly in t, and $\sup_{s,t\in T} E(|D_s u(t)|^p) < \infty$.

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

$$\begin{cases} D_t^+ u(t) = \lim_{\epsilon\downarrow 0} \sum_{i=1}^d D_t^i u^i(t+\epsilon) \\ D_t^- u(t) = \lim_{\epsilon\downarrow 0} \sum_{i=1}^d D_t^i u^i(t-\epsilon) \end{cases} \tag{5.52}$$

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

Consider the process

$$X(t) = \begin{cases} \eta(0) + \int_0^t u(s) \circ dW(s) + \int_0^t v(s)\, ds, & t > 0 \\ \eta(t), & -r \leq t \leq 0, \end{cases} \tag{5.53}$$

where η belongs to C and is of bounded variation, $u = (u^1, \cdots, u^m)^T$, $u^i \in \mathbb{L}_{d,C,loc}^{2,4}$, $(\bigtriangledown u) \in \mathbb{L}_{loc}^{1,4}$, $v = (v^1, \cdots, v^m)^T$, $v^i \in \mathbb{L}_{loc}^{1,4}$, and the stochastic integral is a Stratonovich one. Assume also that the process X is continuous.

Using the relationship between the Skorohod and Stratonovich integrals ([22], Theorem 7.3; [23], Theorem 3.11) and Theorem 3.3, we can easily obtain the following Stratonovich version of Itô's formula for the segment process X_t.

Corollary 5.8.

Suppose that the process $X(t)$ is defined by (5.53), and let $\phi \in C^{1,2}(T \times \mathbf{R}^{mk}, \mathbf{R})$. Then

(5.54)

$$\begin{aligned}\phi(t, \Pi(X_t)) - \phi(0, \Pi(X_0)) &= \int_0^t \frac{\partial \phi}{\partial s}(s, \Pi(X_s))\, ds + \sum_{i=1}^k \int_0^t \frac{\partial \phi}{\partial \vec{x}^i}(s, \Pi(X_s)) u(s+s_i) \circ dW(s+s_i) \\ &\quad + \sum_{i=1}^k \int_0^t \frac{\partial \phi}{\partial \vec{x}^i}(s, \Pi(X_s)) v(s+s_i)\, ds\,.\end{aligned}$$

for all $t \in T$ a.s..

6. Weak differentiability of solutions of SDDE's.

In this section, we will study the weak differentiability of the solution of the Itô SDDE (1.6). Bell and Mohammed ([4]) have applied the Malliavin calculus to study regularity of solutions of SDDE's with a single delay in the noise term. Their analysis relies on weak differentiability of the solution of the SDDE. In Section 8 of this article, the weak differentiability of the solution to the SDDE (1.6) together with the Itô formula (5.10) are used to develop higher order numerical schemes for solving the SDDE. The next three results (Proposition 6.1, Lemma 6.2, and Proposition 6.3) are analogous to those in Nualart ([23] Theorem 2.2.1, Lemma 2.2.2, and Theorem 2.2.2). Denote $\mathbb{D}_m^{k,\infty} := \cap_{p\geq 2}\mathbb{D}_m^{k,p}$, for $k \in N$. Recall that $D_r^l, 1 \leq l \leq d$, stand for weak differentiation with respect to the l-th component of W.

Proposition 6.1. *(c.f. [23], Proposition 1.2.3).*

In the Itô SDDE (1.6), assume that $g \in C_b^{0,1}(T \times \mathbf{R}^{k_1 m}; L(\mathbf{R}^d, \mathbf{R}^m)$ and $h \in C_b^{0,1}(T \times \mathbf{R}^{k_2 m}; \mathbf{R}^m)$. Let X be the solution of (1.6). Then $X(t) \in \mathbb{D}_m^{1,\infty}$ for all $t \in T$, and

$$\sup_{0\leq r\leq a} E(\sup_{r\leq s\leq a} |D_r X(s)|^p) < \infty \tag{6.1}$$

for all $p \geq 2$. Furthermore, the "partial" weak derivatives $D_r^l X^j(t)$ with respect to the l-th coordinate of W satisfy the following linear SDDE's a.s.:

$$D_r^l X^j(t) = \begin{cases} g^{jl}(r, \Pi_1(X_r^j)) + \int_r^t \sum_{i=1}^{k_1} \frac{\partial g^{jl}}{\partial \vec{x}_i}(s, \Pi_1(X_s)) D_r^l X^j(s + s_{1,i})\, dW^l(s) \\ + \int_0^t \sum_{i=1}^{k_2} \frac{\partial h^j}{\partial \vec{x}_i}(s, \Pi_2(X_s)) D_r^l X^j(s + s_{2,i})\, ds, \quad t \geq r, \\ = 0, \quad t < r, \end{cases} \tag{6.2}$$

for $l = 1, \cdots, d$, $j = 1, \cdots, m$. In (6.2), g^{jl} is the (j,l) entry of the $m \times d$ matrix g, and h^j is the j-th coordinate of h.

Proof.

For simplicity, we will only consider the one-dimensional case $d = m = 1$.

$$X^0(t) = \begin{cases} \eta(0), & t \geq 0 \\ \eta(t), & -1 \leq t < 0, \end{cases}$$

$$X^{n+1}(t) = \eta(0) + \int_0^t g(s, \Pi_1(X_s^n))\, dW(s) + \int_0^t h(s, \Pi_2(X_s^n))\, ds. \tag{6.3}$$

It is easy to see that

$$\begin{aligned} & D_r\Big(\int_0^t g(s, \Pi_1(X_s^n))\, dW(s)\Big) \\ &= g(r, \Pi_1(X_r^n)) + \int_{r-s_{1,k_1}}^t D_r(g(s, \Pi_1(X_s^n)))\, dW(s) \end{aligned} \tag{6.4}$$

and

$$D_r\Big(\int_0^t h(s, \Pi_2(X_s^n))\, ds\Big) = \int_{r-s_{2,k_2}}^t D_r(h(s, \Pi_2(X_s^n)))\, ds. \tag{6.5}$$

Since g and h have bounded space derivatives, it is easy to see that there is a positive constant K such that

$$\begin{cases} D_r(g(s, \Pi_1(X_s^n))) \leq K \sup_{r \leq u \leq s} |D_r X^n(u)| \\ D_r(h(s, \Pi_2(X_s^n))) \leq K \sup_{r \leq u \leq s} |D_r X^n(u)|, \end{cases} \tag{6.6}$$

almost surely. From the Burkholder-Davis-Gundy inequality and (6.3)-(6.6), it follows that $X^n(t) \in \mathbb{D}^{1,\infty}$ for all $t \in [0,a]$, and there are positive constants C_1, C_2 such that

$$E(\sup_{r\le u\le t}|D_rX^{n+1}(u)|^p) \le C_1(1+E\|X_r^n\|_C^p) + C_2\int_r^t E\left(\sup_{r\le u\le s}|D_rX^n(u)|^p\right)ds. \tag{6.7}$$

By induction on n, the above inequality implies that $E(\sup_{r\le s\le a}|D_rX^n(s)|^p)$ are uniformly bounded in n for all $p \ge 2$. By [23], proposition 1.5.5, it follows that $X(t) \in \mathbb{D}^{1,\infty}$ for all t. Applying the operator D to (1.6) (and using [23] Proposition 1.2.3), we obtain the linear SDDE (6.2) for the weak derivative of $X(t)$. The estimate (6.1) follows from (6.2), Burkholder-Davis-Gundy's inequality and Gronwall's lemma. □

The following lemma may be proved using similar ideas. Its proof is left to the reader.

Lemma 6.2.

Suppose that the real-valued process $\alpha = \{\alpha(r,t) : t \in [r,a]\}$ is adapted and continuous. Assume that the processes $a(t) = (a_1(t),\cdots,a_{k_1}(t)) \in \mathbf{R}^{k_1}$ and $b(t) = (b_1(t),\cdots,b_{k_2}(t)) \in \mathbf{R}^{k_2}$ are adapted, continuous and uniformly bounded. Furthermore, suppose that the random variables $\alpha(r,t)$, $a(t)$ and $b(t)$ belong to $\mathbb{D}^{1,\infty}$ and satisfy the conditions

$$\begin{cases} \sup_{0\le r\le a} E(\sup_{r\le t\le a}|\alpha(r,t)|^p) + \sup_{0\le r,s\le a} E(\sup_{s\le t\le a}|D_s\alpha(r,t)|^p) < \infty \\ \sup_{0\le s\le a}\left\{E(\sup_{s\le t\le a}|a(t)|^p) + E(\sup_{s\le t\le a}|D_sa(t)|^p)\right\} < \infty \\ \sup_{0\le s\le a}\left\{E(\sup_{s\le t\le a}|b(t)|^p) + E(\sup_{s\le t\le a}|D_sb(t)|^p)\right\} < \infty \end{cases} \tag{6.11}$$

for all $p \ge 2$. Let $Y = \{Y(t) : t \in [0,a]\}$ be the solution of the linear SDDE

$$Y(t) = \begin{cases} \alpha(r,t) + \int_r^t \langle a(s), \Pi_1(Y_s)\rangle_{\mathbf{R}^{k_1}}\,dW(s) + \int_r^t \langle b(s), \Pi_2(Y_s)\rangle_{\mathbf{R}^{k_2}}\,ds, & t \ge r, \\ 0, \quad 0 \le t \le r. \end{cases} \tag{6.12}$$

Then $Y(t)$ belongs to $\mathbb{D}^{1,\infty}$, and for all integers $p \ge 2$, we have

$$\begin{cases} \sup_{0\le s\le a} E(\sup_{s\le t\le a}|D_sY(t)|^p) < \infty \\ \sup_{0\le s\le a} E(\sup_{s\le t\le a}|Y(t)|^p) < \infty. \end{cases} \tag{6.13}$$

Furthermore, the weak derivative $D_sY(t)$ of $Y(t)$ satisfies the linear SDDE

(6.14)
$$\begin{aligned} D_sY(t) &= D_s\alpha(r,t) + \langle a(s), \Pi_1(Y_s)\rangle_{\mathbf{R}^{k_1}} I_{\{r\le s\le t\}} \\ &\quad + \int_r^t [\langle D_s a(v), \Pi_1(Y_v)\rangle_{\mathbf{R}^{k_1}} + \langle a(v), \Pi_1(D_sY_v)\rangle_{\mathbf{R}^{k_1}}]\, dW(v) \\ &\quad + \int_r^t [\langle D_s b(v), \Pi_2(Y_v)\rangle_{\mathbf{R}^{k_2}} + \langle b(v), \Pi_2(D_sY_v)\rangle_{\mathbf{R}^{k_2}}]\, dv, \quad s < t. \end{aligned}$$

The next proposition follows from Proposition 6.1 and Lemma 6.2.

Proposition 6.3.

Let $X = \{X(t) : t \in T = [0,a]\}$ be the solution of the SDDE (1.6), where $g \in C_b^{0,2}(T \times \mathbf{R}^{k_1 m}, L(\mathbf{R}^d, \mathbf{R}^m))$, $h \in C_b^{0,2}(T \times \mathbf{R}^{k_2 m}, \mathbf{R}^m)$ have bounded first and second partial derivatives in the space variables. Then $X(t) \in \mathbb{D}_m^{2,\infty}$ for all $t \in T$, and

(6.15)
$$\sup_{0\le r_1, r_2 \le a} E\Big(\sup_{r_1 \vee r_2 \le s \le a} |D_{r_1}^{l_1} D_{r_2}^{l_2} X(s)|^p\Big) < \infty$$

for $l_1, l_2 = 1, \cdots, d$, and all $p \ge 2$.

7. Strong approximation of multiple Stratonovich integrals.

The following iterated Stratonovich integrals are used in the Milstein scheme for the SDDE (1.6):

(7.1)
$$J_{i,j}(t_0, t_1; -b) := \int_{t_0+b}^{t_1+b} \int_{t_0}^{s-b} \circ dW^i(v) \circ dW^j(s),$$

where $0 < t_0 < t_1, b \ge 0$.

We will adopt the discretization scheme in [16] (section 5.8) in order to handle the above double stochastic integral. For alternative discretization approaches to iterated stochastic integrals, see [10] and [26].

Set

(7.2)
$$J(t_0, t_1; -b) := J_{1,1}(t_0, t_1; -b),$$

$t := t_1 - t_0$ and $r := 2\pi/t$. We choose a complete orthonormal basis of $L^2[0,t]$ as

$$\{\frac{1}{\sqrt{t}}\} \cup \{\sqrt{\frac{2}{t}}\sin nrs, \sqrt{\frac{2}{t}}\cos nrs : n = 1, 2, \cdots, 0 \le s \le t\}. \tag{7.3}$$

Set $\bar{W}^i(s) := W^i(s+t_0) - W^i(t_0)$ and $\bar{B}^j(s) := \bar{W}^j(s+b) - \bar{W}^j(b)$, $s \ge 0, 1 \le i,j \le d$. Using the Kahunen-Loève expansion technique, we have

$$\bar{W}^i(s) - \frac{s}{t}\bar{W}^i(t) = \frac{a_0^i(t_0)}{2} + \sum_{n=1}^{\infty}[a_n^i(t_0)\cos nrs + b_n^i(t_0)\sin nrs] \tag{7.4}$$

and

$$\bar{B}^j(s) - \frac{s}{t}\bar{B}^j(t) = \frac{a_0^{j,b}(t_0)}{2} + \sum_{n=1}^{\infty}[a_n^{j,b}(t_0)\cos nrs + b_n^{j,b}(t_0)\sin nrs] \tag{7.5}$$

where

$$\begin{cases} a_n^i(t_0) = \frac{2}{t}\int_0^t(\bar{W}^i(s) - \frac{s}{t}\bar{W}^i(t))\cos nrs\, ds \\ b_n^i(t_0) = \frac{2}{t}\int_0^t(\bar{W}^i(s) - \frac{s}{t}\bar{W}^i(t))\sin nrs\, ds \end{cases} \tag{7.6}$$

and

$$\begin{cases} a_n^{j,b}(t_0) = \frac{2}{t}\int_0^t(\bar{B}^j(s) - \frac{s}{t}\bar{B}^j(t))\cos nrs\, ds \\ b_n^{j,b}(t_0) = \frac{2}{t}\int_0^t(\bar{B}^j(s) - \frac{s}{t}\bar{B}^j(t))\sin nrs\, ds \end{cases}. \tag{7.7}$$

for $n \ge 1$. The convergences in (7.4) and (7.5) are in $L^2(\Omega \times [0,t])$. It is easy to see that if $n \ge 1$, $a_n^i(t_0)$, $b_n^i(t_0)$, $a_n^{j,b}(t_0)$ and $b_n^{j,b}(t_0)$ are normally distributed with mean 0 and variance $t/2\pi^2 n^2$ ([16], p.198). Furthermore, $\{a_n^i(t_0), b_n^i(t_0)\}$ and $\{a_n^{j,b}(t_0), b_n^{j,b}(t_0)\}$ are pairwise independent ([16], p. 198). One can use well-known random number generators to simulate these random coefficients (c.f. [9], section 3.1.2, [16], section 1.3, and [17], section 1.2).

Lemma 7.1.

Let $t_0, t \ge 0$. Then

$$\begin{aligned} J_{i,j}(t_0, t_0+t; -b) &= \frac{1}{2}(\bar{W}^i(t)\bar{B}^j(t)) - \frac{1}{2}(\bar{W}^i(t)a_0^{j,b}(t_0) - \bar{B}^j(t)a_0^i(t_0)) \\ &\quad + \pi\sum_{n=1}^{\infty} n[a_n^i(t_0)b_n^{j,b}(t_0) - b_n^i(t_0)a_n^{j,b}(t_0)], \quad 1 \le i,j \le d, \end{aligned} \tag{7.8}$$

where the infinite series converges in $L^2(\Omega, \mathbf{R})$.

Proof.

It suffices to show (7.8) for $t_0 = 0$. Fix $t > 0$. For simplicity of notation, we write

$$a_n^j = a_n^j(0),\ b_n^j = b_n^j(0),\ a_n^{j,b} = a_n^{j,b}(0),\ b_n^{j,b} = b_n^{j,b}(0) \tag{7.9}$$

and

$$W_N^i(s) = \frac{s}{t}W^i(t) + \frac{a_0^i}{2} + \sum_{n=1}^{N}(a_n^i \cos nrs + b_n^i \sin nrs), \tag{7.10}$$

It is easy to check that

$$\int_b^{t+b}\int_0^{s-b} \circ dW_N^i(v) \circ dW^j(s) \to \int_b^{t+b}\int_0^{s-b} \circ dW^i(v) \circ dW^j(s) \tag{7.11}$$

in $L^2(\Omega)$ as $N \to \infty$. Then we may write

$$\begin{aligned}
J_{i,j}(0,t;-b) &= \int_b^{t+b} W^i(s-b) \circ dW^j(s) \\
&= \int_b^{t+b} \frac{s-b}{t} W^i(t) \circ\ dW^j(s) + \frac{a_0^i}{2}\bar{B}^j(t) \\
&\quad + \sum_{n=1}^{\infty}[a_n^i \int_b^{t+b} \cos nr(s-b)\, dW^j(s) + b_n^i \int_b^{t+b} \sin nr(s-b)\, dW^j(s)].
\end{aligned}$$

For any $n \geq 1$, we have

$$\begin{aligned}
\int_b^{t+b} \cos nr(s-b)\, dW^j(s) &= \int_0^t \cos nrs\, d\bar{B}^j(s) \\
&= \int_0^t \cos nrs\, d(\bar{B}^j(s) - \frac{s}{t}\bar{B}^j(t)) + \int_0^t \cos nrs\, d(\frac{s}{t}\bar{B}^j(t)) \\
&= \cos nrs(\bar{B}^j(s) - \frac{s}{t}\bar{B}^j(t))|_0^t + nr\int_0^t (\bar{B}^j(s) - \frac{s}{t}\bar{B}^j(t)) \sin nrs\, ds \\
&\qquad + \frac{\bar{B}^j(t)}{t}\int_0^t \cos nrs\, ds \\
&= \frac{t}{2} nr b_n^{j,b}.
\end{aligned}$$

Similarly, we have

$$\int_b^{t+b} \sin nr(s-b)\, dW^j(s) = -\frac{t}{2} nr a_n^{j,b}. \tag{7.12}$$

So

$$J_{i,j}(0,t;-b) = \frac{W^i(t)}{t}\int_0^t s\,d\bar{B}^j(s) + \frac{a_0^i}{2}\bar{B}^j(t) + \frac{rt}{2}\sum_{n=1}^{\infty} n(a_n^i b_n^{j,b} - b_n^i a_n^{j,b}). \tag{7.13}$$

On the other hand,

$$\begin{aligned}\int_0^t s\,d\bar{B}^j(s) &= t\bar{B}^j(t) - \int_0^t \bar{B}^j(s)\,ds \\ &= \frac{t}{2}\bar{B}^j(t) - \int_0^t (\bar{B}^j(s) - \frac{s}{t}\bar{B}^j(t))\,ds \\ &= \frac{t}{2}(\bar{B}^j(t) - a_0^{j,b}).\end{aligned}$$

Therefore,

$$J_{i,j}(0,t;-b) = \frac{1}{2}W^i(t)\bar{B}^j(t) - \frac{1}{2}(W^i(t)a_0^{j,b} - \bar{B}^j(t)a_0^i) + \frac{rt}{2}\sum_{n=1}^{\infty} n(a_n^i b_n^{j,b} - b_n^i a_n^{j,b}). \tag{7.14}$$

□

The expansion of $J_{i,j}(0,t;-b)$ is a generalization of the expansion of

$$\begin{aligned}J(i,j) &:= \int_0^t \int_0^s \circ\, dW^i(v) \circ\, dW^j(s) \\ &= \frac{1}{2}(W^i(t)W^j(s)) - \frac{1}{2}(W^i(t)a_0^{j,b} - W^j(t)a_0^i) \\ &\quad + \frac{rt}{2}\sum_{n=1}^{\infty} n(a_n^i b_n^j - b_n^i a_n^j)\end{aligned} \tag{7.15}$$

(see [10], [16], and [17]). Set

$$\begin{aligned}J_{i,j}^p(t_0,t_1;-b) &:= \frac{1}{2}(\bar{W}^i(t)\bar{B}^j(t)) - \frac{1}{2}(\bar{W}^i(t)a_0^{j,b}(t_0) - \bar{B}^j(t)a_0^i(t_0)) \\ &\quad + \pi\sum_{n=1}^{p} n[a_n^i(t_0)b_n^{j,b}(t_0) - b_n^i(t_0)a_n^{j,b}(t_0)].\end{aligned} \tag{7.16}$$

Then $J_{i,j}^p(t_0,t_1;-b)$ can be used to approximate $J_{i,j}(t_0,t_1;-b)$ in the mean square. The rate of convergence is given in Lemma 7.2 below.

Lemma 7.2.

For any integer $p \geq 1$ and $t > 0$, we have

$$E|J^p_{i,j}(0,t;-b) - J_{i,j}(0,t;-b)|^2 \leq \frac{t^2}{2\pi^2 p}. \tag{7.17}$$

Proof.

Let $p \geq 1$ be any integer. Then

$$\sum_{n=p+1}^{\infty} \frac{1}{n^2} \leq \int_p^{\infty} \frac{1}{u^2}\, du = \frac{1}{p}. \tag{7.20}$$

Since $\{W^i(t)\}$ and $\{B^j(t)\}$ are independent, $E(a^i_n b^i_n) = 0$ and $E(a^{j,b}_n b^{j,b}_n) = 0$, we have

$$\begin{aligned} E|J^p_{i,j}(0,t;-b) - J_{i,j}(0,t;-b)|^2 &= \pi^2 \sum_{n=p+1}^{\infty} n^2 E(a^i_n b^{j,b}_n - b^i_n a^{j,b}_n)^2 \\ &= \pi^2 \sum_{n=p+1}^{\infty} n^2 [E(a^i_n b^{j,b}_n)^2 + E(b^i_n a^{j,b}_n)^2] \\ &= \frac{t^2}{2\pi^2} \sum_{n=p+1}^{\infty} \frac{1}{n^2} \leq \frac{t^2}{2\pi^2 p}. \quad \square \end{aligned}$$

8. The strong Milstein scheme.

In this section we construct a strong Milstein scheme of order 1 for the SDDE (1.6). Our construction relies heavily on the Itô formula for "tame" functions (Theorem 5.3).

Throughout this section, we assume that in (1.6) the coefficients $g \in C^{1,2}(T \times \mathbf{R}^{k_1 m}, L(\mathbf{R}^d, \mathbf{R}^m))$ and $h \in C^{1,2}(T \times \mathbf{R}^{k_2 m}, \mathbf{R}^m)$. For convenience, set $W(s) = W(0) = 0$, for all $s \leq 0$. We also define

$$u(t) := \begin{cases} g(t, \Pi_1(X_t)), & 0 \leq t \leq a, \\ 0, & t < 0, \end{cases} \quad \text{and} \quad v(t) := \begin{cases} h(t, \Pi_2(X_t)), & 0 \leq t \leq a, \\ \eta(t), & t < 0. \end{cases} \tag{8.1}$$

We first derive the Milstein scheme for the case $d = m = 1$.

8.1. Itô-Taylor expansion.

Assume that $0 < t_0 < t$, and $\vec{x} = (x_1, \cdots, x_{k_1}) \in \mathbf{R}^{k_1}$. Applying the Itô formula (5.10), we have

(8.2)

$$\begin{aligned} & g(t, \Pi_1(X_t)) - g(t_0, \Pi_1(X_{t_0})) \\ & = \int_{t_0}^{t} \frac{\partial g}{\partial s}(s, \Pi_1(X_s))\, ds + \sum_{i=1}^{k_1} \int_{t_0+s_{1,i}}^{t+s_{1,i}} \frac{\partial g}{\partial x_i}(s - s_{1,i}, \Pi_1(X_{s-s_{1,i}}))u(s)\, dW(s) \\ & + \sum_{i=1}^{k_1} \int_{t_0}^{t} [\frac{\partial g}{\partial x_i}(s, \Pi_1(X_s))v(s + s_{1,i}) + \frac{1}{2}\langle \frac{\partial^2 g}{\partial x_i^2}(s, \Pi_1(X_s)) \bigtriangledown^{+}_{s_{1,i}} X(s), \bigtriangledown^{-}_{s_{1,i}} X(s)\rangle]\, ds, \end{aligned}$$

where $\bigtriangledown^{\pm}_{s_{1,i}} X(s)$ are defined by (5.9). Applying the Itô formula (5.10) again and using similar notations for h, we obtain

(8.3)

$$\begin{aligned} & h(t, \Pi_2(X_t)) - h(t_0, \Pi_2(X_{t_0})) \\ & = \int_{t_0}^{t} \frac{\partial h}{\partial s}(s, \Pi_2(X_s))\, ds + \sum_{i=1}^{k_2} \int_{t_0+s_{2,i}}^{t+s_{2,i}} \frac{\partial h}{\partial x_i}(s - s_{2,i}, \Pi_2(X_{s-s_{2,i}}))\, u(s)\, dW(s) \\ & + \frac{1}{2} \sum_{i=1}^{k_2} \int_{t_0}^{t} [\frac{\partial h}{\partial x_i}(s, \Pi_2(X_s))v(s + s_{2,i}) + \frac{1}{2}\langle \frac{\partial^2 g}{\partial x_i^2}(s, \Pi_2(X_s)) \bigtriangledown^{+}_{s_{2,i}} X(s), \bigtriangledown^{-}_{s_{2,i}} X(s)\rangle]\, ds. \end{aligned}$$

Substituting (8.2) and (8.3) into (1.6), we get the following approximate (Itô-Taylor) expansion of (1.6):

(8.5)

$$\begin{aligned} X(t) = {} & X(t_0) + g(t_0, \Pi_1(X_{t_0}))[W(t) - W(t_0)] + h(t_0, \Pi_2(X_{t_0})(t - t_0) \\ & + \sum_{i=1}^{k_1} \frac{\partial g}{\partial x_i}(t_0, \Pi_1(X_{t_0}))u(t_0 + s_{1,i}) \int_{t_0}^{t} \int_{t_0+s_{1,i}}^{t_1+s_{1,i}} dW(t_2)\, dW(t_1) + R(t_0, t), \end{aligned}$$

where

(8.6)

$$
\begin{aligned}
R(t_0,t) := \sum_{i=1}^{k_1}\{\int_{t_0}^{t}\int_{t_0+s_{1,i}}^{t_1+s_{1,i}} & \left[\frac{\partial g}{\partial x_i}(t_2-s_{1,i},\Pi_1(X_{t_2-s_{1,i}}))u(t_2)-\frac{\partial g}{\partial x_i}(t_0,\Pi_1(X_{t_0}))\right.\\
&\left.\times u(t_0+s_{1,i})\right] dW(t_2)\,dW(t_1)\} + \int_{t_0}^{t}\int_{t_0}^{t_1}\sum_{i=1}^{k_1}\left[\frac{\partial g}{\partial x_i}(t_2,\Pi_1(X_{t_2}))v(t_2+s_{1,i})\right.\\
&\left.+\frac{1}{2}\langle\frac{\partial^2 g}{\partial x_i^2}(t_2,\Pi_1(X_{t_2}))\,\nabla^{+}_{s_{1,i}}X_{t_2},\nabla^{-}_{s_{1,i}}X_{t_2}\rangle\right] dt_2\,dW(t_1)\\
&+\sum_{i=1}^{k_2}\int_{t_0}^{t}\int_{t_0+s_{2,i}}^{t_1+s_{2,i}}\frac{\partial h}{\partial x_i}(t_2-s_{2,i},\Pi_2(X_{t_2-s_{2,i}}))u(t_2)\,dW(t_2)\,dt_1\\
&+\int_{t_0}^{t}\int_{t_0}^{t_1}\sum_{i=1}^{k_2}\left[\frac{\partial h}{\partial x_i}(t_2,\Pi_2(X_{t_2}))v(t_2+s_{2,i})\right.\\
&\left.+\frac{1}{2}\langle\frac{\partial^2 h}{\partial x_i^2}(t_2,\Pi_2(X_{t_2}))\,\nabla^{+}_{s_{2,i}}X_{t_2},\nabla^{-}_{s_{2,i}}X_{t_2}\rangle\right] dt_2\,dt_1\\
&+\int_{t_0}^{t}\int_{t_0}^{t_1}\left[\frac{\partial g}{\partial t_2}(t_2,\Pi_1(X_{t_2}))+\frac{\partial h}{\partial t_2}(t_2,\Pi_2(X_{t_2}))\right] dt_2\,dt_1.
\end{aligned}
$$

In the above expression, the stochastic integrals

$$\int_{t_0+s_{1,i}}^{t_1+s_{1,i}}\frac{\partial g}{\partial x_i}(t_2-s_{1,i},\Pi_1(X_{t_2-s_{1,i}}))u(t_2)\,dW(t_2)$$

and

$$\int_{t_0+s_{2,i}}^{t_1+s_{2,i}}\frac{\partial h}{\partial x_i}(t_2-s_{2,i},\Pi_2(X_{t_2-s_{2,i}}))u(t_2)\,dW(t_2)$$

are Skorohod integrals. Define

$$I(t_0+s_{i,j},t+s_{i,j};s_{i,j}) := \int_{t_0}^{t}\int_{t_0+s_{i,j}}^{t_1+s_{i,j}} dW(t_2)\,dW(t_1),$$

for $i=1,2$ and $j=1,\cdots,k_i$. Recall the definition of $J(t_0+s_{i,j},t+s_{i,j};s_{i,j})$ in (7.1). Note that if $s_{i,j}<0$, then

$$I(t_0+s_{i,j},t+s_{i,j};s_{i,j}) = \int_{t_0}^{t}\int_{t_0+s_{i,j}}^{t_1+s_{i,j}} \circ dW(t_2)\circ dW(t_1); \tag{8.7}$$

if $s_{i,j} = 0$, then

$$I(t_0 + s_{i,j}, t + s_{i,j}; s_{i,j}) = \int_{t_0}^{t} [W(t_1) - W(t_0)]\, dW(t_1) = \frac{(W(t) - W(t_0))^2}{2} - \frac{t - t_0}{2}. \tag{8.8}$$

8.2. The one-dimensional Milstein scheme ($d = m = 1$).

Assume $d = m = 1$. Recall the partition $\pi_p := -1 = t_{-l} < \cdots < t_0 = 0 < \cdots < t_n = t$ constructed in Section 3. We introduce the *Milstein* scheme for the SDDE (1.6) as follows:

$$\begin{aligned} X^p(t) =& X^p(t_k) + h(t_k, \Pi_2(X^p_{t_k}))(t - t_k) + g(t_k, \Pi_1(X^p_{t_k}))(W(t) - W(t_k)) \\ &+ \sum_{i=1}^{k_1} \frac{\partial g}{\partial x_i}(t_k, \Pi_1(X^p_{t_k}))u^p(t_k + s_{1,i})I(t_k + s_{1,i}, t + s_{1,i}; s_{1,i}), \end{aligned} \tag{8.9}$$

for $t_k < t \leq t_{k+1}$, where

$$u^p(t) = \begin{cases} g(t, \Pi_1(X^p_t)), & t \geq 0, \\ 0, & -1 \leq t < 0, \end{cases}$$

and

$$I(t_k + s_{1,i}, t + s_{1,i}; s_{1,i}) = \int_{t_k}^{t} \int_{t_k + s_{1,i}}^{t_1 + s_{1,i}} \circ dW(t_2) \circ dW(t_1).$$

Recall the notation

$$\lfloor s \rfloor := \begin{cases} t_k, & \text{if } t_k \leq s < t_{k+1} \\ t_{n_t}, & \text{if } t_{n_t} \leq s \leq t. \end{cases}$$

and introduce the following notation:

$$\lceil s \rceil = \begin{cases} t_{k+1}, & t_k < s \leq t_{k+1}, \\ t, & t_{n_t} < s \leq t. \end{cases}$$

In view of (8.7) and Lemma 7.2, we will use $J^p(t_i, t; s_{1,i})$ to approximate $I(t_i, t; s_{1,i})$.

Lemma 8.1.

In the SDDE (1.6), suppose that $g \in C_b^2(\mathbf{R}^{k_1}, \mathbf{R})$, $h \in C_b^2(\mathbf{R}^{k_2}, \mathbf{R})$, have bounded first and second derivatives. Then for each integer $m \geq 1$, there exists a constant $K(m) > 0$ such that

$$\begin{cases} E(\langle \frac{\partial^2 g}{\partial \vec{x}^2}(s, \Pi_1(X_s)) \bigtriangledown^+_{s_{1,i}} X_s, \bigtriangledown^-_{s_{1,i}} X_s \rangle^m) \leq K(m) \\ E(\langle \frac{\partial^2 h}{\partial \vec{x}^2}(s, \Pi_2(X_s)) \bigtriangledown^+_{s_{2,i}} X_s, \bigtriangledown^-_{s_{2,i}} X_s \rangle^m) \leq K(m) \end{cases} \tag{8.10}$$

for all $t \in [0, a]$.

Proof.

By the definition of $\bigtriangledown^{\pm}_{s_{2,i}} X(s)$ (see (5.9)), we have

$$\begin{aligned} \bigtriangledown^+_{s_{1,i}, s_{1,j}} X(s) &= 2u(s + s_{1,i}) I_{\{s_{1,i} < s_{1,j}\}} + u(s + s_{1,i}) \delta_{ij} \\ &+ 2 \int_0^{s+s_{1,j}} D_{s+s_{1,i}} u(r)\, dW(r) + 2 \int_0^{s+s_{1,j}} D_{s+s_{1,i}} v(r)\, dr, \end{aligned} \tag{8.11}$$

and

$$\bigtriangledown^-_{s_{1,i}, s_{1,j}} X(s) = u(s + s_{1,i}) \delta_{ij}\,. \tag{8.12}$$

Therefore,

$$\begin{aligned} &\langle \frac{\partial^2 g}{\partial \vec{x}^2}(s, \Pi_1(X_s)) \bigtriangledown^+_{s_{1,i}} X(s), \bigtriangledown^-_{s_{1,i}} X(s) \rangle \\ &= 2 \sum_{i=1}^{k_1} \{ \frac{\partial^2 g}{\partial x_i \partial x_j}(s, \Pi_1(X_s)) u(s + s_{1,i}) [u(s + s_{1,i}) I_{\{s_{1,i} < s_{1,j}\}} + \frac{1}{2} u(s + s_{1,i}) \delta_{ij} \\ &\quad + \int_0^{s+s_{1,j}} D_{s+s_{1,i}} u(r)\, dW(r) + \int_0^{s+s_{1,j}} D_{s+s_{1,i}} v(r)\, dr] \}. \end{aligned} \tag{8.13}$$

If $r > 0$, then

$$D_s u(r) = D_s g(\Pi_1(X_r)) = \sum_{i=1}^{k_1} \frac{\partial g}{\partial x_i}(r, \Pi_1(X_r)) D_s X(r + s_{1,i}), \tag{8.14}$$

and

$$D_t D_s u(r) = \sum_{i,j=1}^{k_1} \frac{\partial^2 g}{\partial x_i \partial x_j}(r, \Pi_1(X_r)) D_s X(r + s_{1,i}) D_t X(r + s_{1,j}) + \sum_{i=1}^{k_1} \frac{\partial g}{\partial x_i}(r, \Pi_1(X_r)) D_t D_s X(r + s_{1,i}). \tag{8.15}$$

By Proposition 6.1 and Proposition 6.3, there exists a constant $C_1 > 0$ such that

$$\begin{cases} \sup_{0 \le s \le a} E(\sup_{s \le r \le a} |D_s X(r)|^2) \le C_1 \\ \sup_{0 \le s,t \le a} E(\sup_{s \vee t \le r \le a} |D_t D_s X(r)|^2) \le C_1. \end{cases}$$

Since g has bounded first and second derivatives, then there is a positive constant C_2 such that

$$\sup_{0 \le s \le a} E(\sup_{s \le r \le a} |D_s u(r)|^2) \le C_2 k_1 \sup_{0 \le s \le a} E(\sup_{s \le r \le a} |D_s X(r)|^2) \le C_1 C_2 k_1,$$

and

$$\sup_{0 \le s,t \le a} E(\sup_{s \vee t \le r \le a} |D_t D_s u(r)|^2) \le C_1^2 C_2^2 k_1 + C_1 C_2 k_1.$$

If $r < s + s_{1,i}$, then

$$\begin{cases} D_{s+s_{1,i}} u(r) = 0 \\ D_{s+s_{1,i}} v(r) = 0. \end{cases}$$

Therefore,

$$\begin{aligned} &E(\int_{t+s_{1,i}}^{t+s_{1,j}} D_{t+s_{1,i}} u(r)\, dW(r))^2 \\ &\le \int_{t+s_{1,i}}^{t+s_{1,j}} \int_{t+s_{1,i}}^{t+s_{1,j}} E(D_s D_{t+s_{1,i}} u(r))^2\, dr\, ds + \int_{t+s_{1,i}}^{t+s_{1,j}} E(D_{t+s_{1,i}} u(r))^2\, dr \\ &\le C_2 k_1^2 C_1^2 + 2 C_2 k_1 C_1 =: K_1. \end{aligned}$$

Similarly, there exists a constant $K_2 > 0$ such that

$$E(\int_{t+s_{1,i}}^{t+s_{1,j}} D_{t+s_{1,i}} v(r)\, dr)^2 \le K_2.$$

So the first inequality of (8.10) follows from the above two inequalities and the Lipschitz and bounded conditions on h, g ((1.4),(1.5)). The second estimate of (8.10) is proved by a similar argument. □

Theorem 8.2.

Consider the Milstein scheme (8.9) for the SDDE (1.6) ($r = 1$).Let $0 < \gamma \leq 1$. Suppose that $\eta : [-1, 0] \to L^2(\Omega, \mathbf{R}^m)$ is Hölder continuous with exponent $\frac{\gamma}{2}$, i.e. there is a positive constant K such that

$$E|\eta(s) - \eta(t)|^2 \leq K|s - t|^\gamma$$

for all $s, t \in J$. Suppose that $g \in C^{1,2}(T \times \mathbf{R}^{k_1}, \mathbf{R})$, $h \in C^{1,2}(T \times \mathbf{R}^{k_2}, \mathbf{R})$ and have bounded first and second spatial derivatives. Assume that

$$\sup_{-1 \leq s \leq 0} E(|Z^p(s)|^2) \leq C' \delta_p^{2\gamma}$$

for some positive constant C', where $\delta_p := |\pi_p|$. Then there exists a constant $C > 0$ (depending on a and independent of π) such that

$$\sup_{-1 \leq s \leq a} E|Z^p(s)|^2 \leq C \delta_p^{2\gamma}$$

for any $p \geq 1$.

Proof.

As in the proof of Theorem 3.4, we express the global error in the form

$$Z^p(t) = Z^p(0) + I^p(t) - R^p(t),$$

where

$$\begin{aligned}
I^p(t) = &\sum_{i=1}^{n_t} [h(t_{i-1}, \Pi_2(X^p(t_{i-1}))) - h(t_{i-1}, \Pi_2(X(t_{i-1})))](t_i - t_{i-1}) \\
&+ \sum_{i=1}^{n_t} [g(t_{i-1}, \Pi_1(X^p(t_{i-1}))) - g(t_{i-1}, \Pi_1(X(t_{i-1})))](W(t_i) - W(t_{i-1})) \\
&+ [h(t_{n_t}, \Pi_2(X^p(t_{n_t}))) - h(t_{n_t}, \Pi_2(X(t_{n_t})))](t - t_{n_t}) \\
&+ [g(t_{n_t}, \Pi_1(X^p(t_{n_t}))) - g(t_{n_t}, \Pi_1(X(t_{n_t})))](W(t) - W(t_{n_t})) \\
&+ \sum_{i=1}^{n_t} \sum_{j=1}^{k_1} \Big\{ I(t_{i-1}, t_i; s_{1,j}) \Big[\frac{\partial g}{\partial x_j}(t_{i-1}, \Pi_1(X^p(t_{i-1})))
\end{aligned}$$

$$\times u^p(t_{i-1}+s_{1,j}) - \frac{\partial g}{\partial x_j}(t_{i-1},\Pi_1(X(t_{i-1})))u(t_{i-1}+s_{1,j})\Big]\Big]\Big\} + \sum_{j=1}^{k_1}\Big\{I_{t_{n_t},t;s_{1,j}}$$
$$\times\Big[\frac{\partial g}{\partial x_j}(t_{n_t},\Pi_1(X^p(t_{n_t})))u^p(t_{n_t}+s_{1,j}) - \frac{\partial g}{\partial x_j}(t_{n_t},\Pi_1(X(t_{n_t})))u(t_{n_t}+s_{1,j})\Big]\Big]\Big\},$$

and

$$R^p(t) = \sum_{i=1}^{n_t} R(t_{i-1},t_i) + R(t_{n_t},t).$$

We shall decompose $R^p(t)$ into five parts:

$$R^p(t) = R_1^p(t) + R_2^p(t) + R_3^p(t) + R_4^p(t) + R_5^p(t),$$

where

$$R_1^p(t) := \sum_{i=1}^{n_t}\sum_{j=1}^{k_1}\Big\{\int_{t_{i-1}}^{t_i}\int_{t_{i-1}+s_{1,j}}^{s+s_{1,j}}\Big[\frac{\partial g}{\partial x_j}(r-s_{1,j},\Pi_1(X_{r-s_{1,j}}))u(r)$$
$$-\frac{\partial g}{\partial x_j}(t_{i-1},\Pi_1(X_{t_{i-1}}))u(t_{i-1}+s_{1,j})\Big]\,dW(r)\,dW(s)\Big\} + \sum_{j=1}^{k_1}\Big\{\int_{t_{n_t}}^{t}\int_{t_{n_t}+s_{1,j}}^{s+s_{1,j}}$$
$$\Big[\frac{\partial g}{\partial x_j}(r-s_{1,j},\Pi_1(X_{r-s_{1,j}}))u(r) - \frac{\partial g}{\partial x_j}(t_{n_t},\Pi_1(X_{t_{n_t}}))u(t_{n_t}+s_{1,j})\Big]\,dW(r)\,dW(s)\Big\}$$
$$=\sum_{j=1}^{k_1}\Big\{\int_0^t\int_{\lfloor s\rfloor+s_{1,j}}^{s+s_{1,j}}\Big[\frac{\partial g}{\partial x_j}(r-s_{1,j},\Pi_1(X_{r-s_{1,j}}))u(r)$$
$$-\frac{\partial g}{\partial x_j}(P_{\lfloor s\rfloor,\underline{s_1}^{k_1}}(X_{\lfloor s\rfloor}))u(\lfloor s\rfloor+s_{1,j})\Big]\,dW(r)\,dW(s)\Big\},$$
$$R_2^p(t) := \sum_{j=1}^{k_1}\int_0^t\int_{\lfloor s\rfloor}^{s}\Big[\frac{\partial g}{\partial x_j}(r,\Pi_1(X_r))v(r+s_{1,j})$$
$$+\frac{1}{2}\langle\frac{\partial^2 g}{\partial\vec{x}^2}(r,\Pi_1(X_r))\,\bigtriangledown^+_{s_{1,j}}X_r,\bigtriangledown^-_{s_{1,j}}X_r\rangle\Big]\,dr\,dW(s),$$
$$R_3^p(t) := \sum_{j=1}^{k_2}\int_0^t\int_{\lfloor s\rfloor+s_{2,j}}^{s+s_{2,j}}\frac{\partial h}{\partial x_j}(r-s_{2,j},\Pi_2(X_{r-s_{2,j}}))u(r)\,dW(r)\,ds,$$
$$R_4^p(t) := \sum_{j=1}^{k_2}\int_0^t\int_{\lfloor s\rfloor}^{s}\Big[\frac{\partial h}{\partial x_j}(r,\Pi_2(X_r))v(r+s_{2,j})$$
$$+\frac{1}{2}\langle\frac{\partial^2 h}{\partial\vec{x}^2}(r,\Pi_2(X_r))\,\bigtriangledown^+_{s_{2,j}}X_r,\bigtriangledown^-_{s_{2,j}}X_r\rangle\Big]\,dr\,ds,$$

and

$$R_5^p(t) := \int_0^t \int_{\lfloor s \rfloor}^s \left\{ \frac{\partial h}{\partial r}(r, \Pi_2(X_r)) + \frac{\partial g}{\partial r}(r, \Pi_1(X_r)) \right\} dr\, ds.$$

By the Itô isometry and the formula for covariance between two Skorohod integrals ([23], Section 1.3.1), we have

$$\begin{aligned}
\sup_{0\le s\le t} E|R_1^p(s)|^2 \le k_1 \sum_{j=1}^{k_1} E \int_0^t \Bigg\{ \int_{\lfloor s \rfloor + s_{1,j}}^{s+s_{1,j}} \Bigg[& \frac{\partial g}{\partial x_j}(r - s_{1,j}, \Pi_1(X_{r-s_{1,j}}))u(r) \\
& - \frac{\partial g}{\partial x_j}(\lfloor s \rfloor, \Pi_1(X_{\lfloor s \rfloor}))u(\lfloor s \rfloor + s_{1,j}) \Bigg] dW(r) \Bigg\}^2 ds \\
& \le k_1 \sum_{j=1}^{k_1} \int_0^t \int_{\lfloor s \rfloor + s_{1,j}}^{s+s_{1,j}} E \Bigg[\frac{\partial g}{\partial x_j}(r - s_{1,j}, \Pi_1(X_{r-s_{1,j}}))u(r) \\
& - \frac{\partial g}{\partial x_j}(\lfloor s \rfloor, \Pi_1(X_{\lfloor s \rfloor}))u(\lfloor s \rfloor + s_{1,j}) \Bigg]^2 dr\, ds \\
& + k_1 \sum_{j=1}^{k_1} \int_0^t \int_{\lfloor s \rfloor + s_{1,j}}^{s+s_{1,j}} \int_{\lfloor s \rfloor + s_{1,j}}^{s+s_{1,j}} E \Bigg\{ D_{r_1} \Bigg[\frac{\partial g}{\partial x_j}(r - s_{1,j}, \Pi_1(X_{r-s_{1,j}}))u(r) \\
& - \frac{\partial g}{\partial x_j}(\lfloor s \rfloor, \Pi_1(X_{\lfloor s \rfloor}))u(\lfloor s \rfloor + s_{1,j}) \Bigg] \Bigg\}^2 dr_1\, dr_2\, ds \\
& = k_1 R_{11}^p(t) + k_1 R_{12}^p(t).
\end{aligned}$$

By assumption, the function $G_j(s,x,z) = \frac{\partial g}{\partial x_j}(s,x)g(s,z)$, ($x \in \mathbf{R}^{k_1}$ and $z \in \mathbf{R}^{k_1}$), is Lipschitz, i.e., there exists a constant $L_1 > 0$ such that

$$|G_j(s,z) - G_j(s,w)| \le L_1|z-w|, \quad \forall z, w \in \mathbf{R}^{2k_1} \text{ and } 1 \le j \le k_1.$$

Using

$$u(r) = \begin{cases} g(r, \Pi_1(X_r)), & r \ge 0 \\ 0, & r < 0, \end{cases}$$

and

$$\sup_{-1\le r_1\le\alpha<\beta\le r_2\le a} E|X(\beta) - X(\alpha)|^2 \le C_2|r_2 - r_1|^\gamma,$$

it follows that

$$\begin{aligned} R_{11}^p(t) \le & \sum_{j=1}^{k_1} \int_0^t \int_{\lfloor s\rfloor + s_{1,j}}^{s+s_{1,j}} E[G_j(r - s_{1,j}, \Pi_1(X_{r-s_{1,j}}), \Pi_1(X_r)) \\ & - G_j(\lfloor s\rfloor, \Pi_1(X_{\lfloor s\rfloor}), \Pi_1(X_{\lfloor s\rfloor + s_{1,j}}))]^2 I_{\{\lfloor s\rfloor + s_{1,j} \ge 0\}} \, dr\, ds \\ & \le 2k_1 L_1^2 \sum_{j=1}^{k_1} \int_0^t \int_{\lfloor s\rfloor + s_{1,j}}^{s+s_{1,j}} \sup_{\substack{-1 \le r_1 < r_2 \le a \\ |r_2 - r_1| \le \delta_p}} E|X(r_2) - X(r_1)|^2 \, dr\, ds \\ & \le 2(a+1) k_1^2 L_1^2 C_2 \delta_p^{2\gamma}. \end{aligned}$$

Now for all $r \ge 0$ and $1 \le j \le k_1$,

$$\begin{aligned} D_s(\frac{\partial g}{\partial x_j}&(r - s_{1,j}, \Pi_1(X_{r-s_{1,j}}))u(r)) \\ & = g(r, \Pi_1(X_r)) \sum_{i=1}^{k_1} \frac{\partial^2 g}{\partial x_j \partial x_i}(r - s_{1,j}, \Pi_1(X_{r-s_{1,j}})) D_s X(r - s_{1,j} + s_{1,i}) \\ & + \frac{\partial g}{\partial x_j}(r - s_{1,j}, \Pi_1(X_{r-s_{1,j}})) \sum_{i=1}^{k_1} \frac{\partial g}{\partial x_i}(r, \Pi_1(X_r)) D_s X(r + s_{1,i}). \end{aligned}$$

By Proposition 6.1, there exists a constant $C_3 > 0$ such that

$$\sup_{0 \le r \le a} E(\sup_{0 \le s \le a} |D_r X(s)|^2) \le C_3.$$

By (1.8), (1.10), and boundedness of the spatial derivatives of g , there exists a constant $C_4 > 0$ such that

$$\sup_{0 \le r \le a} \sup_{0 \le s \le a} E(|D_s(\frac{\partial g}{\partial x_j}(r - s_{1,j}, \Pi_1(X_{r-s_{1,j}}))u(r))|^2) \le 2C_4 k_1^2.$$

Therefore

$$\begin{aligned} R_{12}^p(t) & \le k_1 \sum_{j=1}^{k_1} \int_0^t \int_{\lfloor s\rfloor + s_{1,j}}^{s+s_{1,j}} \int_{\lfloor s\rfloor + s_{1,j}}^{s+s_{1,j}} E\{D_{r_1}[4C_4 k_1^2] \, dr_1\, dr_2\, ds \\ & \le 4(a+1) C_4 k_1^4 \delta_p^2. \end{aligned}$$

Hence there is a constant $C_5 > 0$ such that

$$\sup_{0 \le s \le t} E|R_1(s)|^2 \le C_5 \delta_p^{2\gamma}. \qquad (8.16)$$

Applying Fubini's theorem, we can rewrite $R_3^p(t)$ as

$$R_3^p(t) = \sum_{i=1}^{n_t}\sum_{j=1}^{k_2}\int_{t_{i-1}}^{t_i}\int_{t_{i-1}+s_{2,j}}^{s+s_{2,j}} \frac{\partial h}{\partial x_j}(r-s_{2,j},\Pi_2(X_{r-s_{2,j}}))u(r)\,dW(r)\,ds$$
$$+\sum_{j=1}^{k_2}\int_{t_{n_t}}^{t}\int_{t_{n_t}+s_{2,1}}^{s+s_{2,j}} \frac{\partial h}{\partial x_j}(r-s_{2,j},\Pi_2(X_{r-s_{2,j}}))u(r)\,dW(r)\,ds.$$

So we have

$$R_3^p(t) = \sum_{i=1}^{n_t}\sum_{j=1}^{k_2}\int_{t_{i-1}+s_{2,j}}^{t_i+s_{2,j}}\int_{r-s_{2,j}}^{t_i} \frac{\partial h}{\partial x_j}(r-s_{2,j},\Pi_2(X_{r-s_{2,j}}))u(r)\,ds\,dW(r)$$
$$+\sum_{j=1}^{k_2}\int_{t_{n_t}+s_{2,j}}^{t+s_{2,j}}\int_{r-s_{2,j}}^{t} \frac{\partial h}{\partial x_j}(r-s_{2,j},\Pi_2(X_{r-s_{2,j}}))u(r)\,ds\,dW(r)$$
$$=\sum_{i=1}^{n_t}\sum_{j=1}^{k_2}\int_{t_{i-1}+s_{2,j}}^{t_i+s_{2,j}}(t_i+s_{2,j}-r)\frac{\partial h}{\partial x_j}(r-s_{2,j},\Pi_2(X_{r-s_{2,j}}))u(r)\,dW(r)$$
$$+\sum_{j=1}^{k_2}\int_{t_{n_t}+s_{2,j}}^{t+s_{2,j}}(t+s_{2,j}-r)\frac{\partial h}{\partial x_j}(r-s_{2,j},\Pi_2(X_{r-s_{2,j}}))u(r)\,dW(r)$$
$$=\sum_{j=1}^{k_2}\int_{s_{2,j}}^{t+s_{2,j}}(\lceil r-s_{2,j}\rceil+s_{2,j}-r)\frac{\partial h}{\partial x_j}(r-s_{2,j},\Pi_2(X_{r-s_{2,j}}))u(r)\,dW(r).$$

Applying the formula for covariance between two Skorohod integrals ([23], Section 1.3.1) and Proposition 6.1, we can show that there exists a constant $C_6 > 0$ such that

$$\sup_{0\le s\le t} E|R_3(s)|^2 \le C_6\delta_p^2. \tag{8.17}$$

Similarly, by Lemma 8.1, we can easily show that there exist $C_7 > 0$ such that

$$\begin{cases} \sup_{0\le s\le t} E|R_2(s)|^2 \le C_7\delta_p^2, \\ \sup_{0\le s\le t} E|R_4(s)|^2 \le C_7\delta_p^2 \\ \sup_{0\le s\le t} E|R_5(s)|^2 \le C_7\delta_p^2 \end{cases} \tag{8.18}$$

By similar arguments to the ones used in the proof of Theorem 7.1, we obtain the following inequality

$$\sup_{0\le u\le t} E|I^p(u)|^2 \le C_1\int_0^t \sup_{-1\le u\le s} E(|Z^p(u)|^2)\,ds \tag{8.19}$$

for some constant $C_1 > 0$. From (8.16)–(8.19), there exist $C_8 > 0$ and $C_9 > 0$ such that

$$\sup_{0\le u\le t} E|Z^p(u)|^2 \le E|Z^p(0)|^2 + C_8\delta_p^{2\gamma} + C_9\int_0^t \sup_{-1\le u\le s} E|Z^p(u)|^2\, ds. \tag{8.20}$$

So

$$\sup_{-1\le u\le t} E|Z^p(u)|^2 \le (2C' + C_8)\delta_p^{2\gamma} + C_9\int_0^t \sup_{-1\le u\le s} E|Z^p(u)|^2\, ds. \tag{8.21}$$

By Gronwall's lemma, there exists a constant $C > 0$ such that

$$E\sup_{-1\le s\le t} |Z^p(s)|^2 \le C\delta_p^{2\gamma}. \quad \square$$

Let us consider a particular case when g and h are of the (linear) form

$$\begin{cases} g(s,\Pi_1(X_s)) = \sum_{i=1}^{k_1} a_i(s, X_s(s_{1,i})) \\ h(s,\Pi_2(X_s)) = \sum_{j=1}^{k_2} b_j(s, X_s(s_{2,j})), \end{cases} \tag{8.22}$$

where $a_i, b_j \in C_b^{1,2}(T\times R)$ for $1\le i\le k_1$ and $1\le j\le k_2$. In this case we can obtain a stronger estimate than the one given in Theorem 8.2.

Theorem 8.3.

Consider the Milstein scheme (8.9) for the SDDE (1.6) in the special case (8.22). Suppose that $0<\gamma\le 1$ and η is Hölder continuous (in $L^q(\Omega,\mathbf{R})$) with exponent $\frac{\gamma}{2}$, i.e.,

$$E|\eta(s)-\eta(t)|^q \le K|s-t|^{\frac{\gamma q}{2}} \tag{8.23}$$

for some constant $K > 0$. Suppose that g and h have bounded first and second space derivatives. Assume that

$$E||Z_0^p||_C^q \le C'(q)\delta_p^{(1+\frac{q}{2})\gamma}$$

for some constant $C'(q)$. Then there exists a constant $C(q) > 0$ (depending on a) such that

$$E\sup_{0\le s\le a} ||Z_s^p||_C^q \le C(q)\delta_p^{(1+\frac{q}{2})\gamma}$$

for any $p \geq 1$.

Proof.

The proof is analogous to that of Theorem 8.2. Instead of using the formula for covariance of two Skorohod integrals ([23], Section 1.3.1), we use the Burkholder-Davis-Gundy inequality to estimate the errors. One may also apply the non-anticipating Itô formula to

$$\begin{cases} a_i(t+s_{1,1}, X(t+s_{1,1})) - a_i(t_0+s_{1,1}, X(t_0+s_{1,1})) \\ a_i(t+s_{1,1}, X(t+s_{1,1})) - a_i(t_0+s_{1,1}, X(t_0+s_{1,1})) \end{cases}$$

in order to obtain the expressions (8.2) and (8.3). □

Remark 8.4.

It is easy to check that the Milstein scheme (8.9) is (stochastically) numerically stable (Definition 2.3). The criteria for strong consistence (Definition 2.2) may not suit the case of higher order ($\gamma \geq 1$) approximation of SDDE because anticipating stochastic integrals are involved.

We can rewrite the SDDE (1.6) in Stratonovich form, namely, if $t \geq 0$,

$$\begin{aligned} X(t) &= \eta(0) + \int_0^t g(s, \Pi_1(X_s)) \circ dW(s) \\ &\quad + \int_0^t [h(s, \Pi_2(X_s)) - \frac{1}{2}\frac{\partial g}{\partial x_{k_1}}(s, \Pi_1(X_s))g(s, \Pi_1(X_s))]\, ds, \end{aligned} \tag{8.24}$$

if $s_{k_1} = 0$. If $s_{k_1} < 0$, then the SDDE is of the same form as (1.6) except the Itô integral is replaced by Stratonovich integral, i.e.,

$$X(t) = \eta(0) + \int_0^t g(s, \Pi_1(X_s)) \circ dW(s) + \int_0^t h(s, \Pi_2(X_s))\, ds,$$

Bell and Mohammed ([5]) derived a similar result in the case of a single delay. From Corollary 5.8, we can obtain the following Stratonovich-Taylor expansion of $X(t)$

$$\begin{aligned} X(t) &= X(t_0) + g(t_0, \Pi_1(X_{t_0}))[W(t) - W(t_0)] + \bar{h}(t_0, \Pi_2(X_{t_0}))(t - t_0) \\ &\quad + \sum_{i=1}^{k_1} \frac{\partial g}{\partial x_i}(t_0, \Pi_1(X_{t_0}))u(t_0 + s_{1,i}) \int_{t_0}^t \int_{t_0+s_{1,i}}^{t_1+s_{1,i}} \circ dW(t_2) \circ dW(t_1) + \bar{R}(t_0, t), \end{aligned} \tag{8.25}$$

where

$$\begin{aligned}
\bar{R}(t_0,t) = &\sum_{i=1}^{k_1}\{\int_{t_0}^{t}\int_{t_0+s_{1,i}}^{t_1+s_{1,i}}[\frac{\partial g}{\partial x_i}(t_2-s_{1,i},\Pi_1(X_{t_2-s_{1,i}}))u(t_2) \\
&-\frac{\partial g}{\partial x_i}(t_0,\Pi_1(X_{t_0}))u(t_0+s_{1,i})]\circ dW(t_2)\circ dW(t_1)\} \\
&+\int_{t_0}^{t}\int_{t_0}^{t_1}\sum_{i=1}^{k_1}\frac{\partial g}{\partial x_i}(t_2,\Pi_1(X_{t_2}))\bar{v}(t_2+s_{1,i})\,dt_2\circ dW(t_1) \\
&+\sum_{i=1}^{k_2}\int_{t_0}^{t}\int_{t_0+s_{2,i}}^{t_1+s_{2,i}}\frac{\partial \bar{h}}{\partial \vec{x}_i}(t_2-s_{2,i},\Pi_2(X_{t_2-s_{2,i}}))u(t_2)\circ dW(t_2)\,dt_1 \\
&+\int_{t_0}^{t}\int_{t_0}^{t_1}\sum_{i=1}^{k_2}\frac{\partial \bar{h}}{\partial \vec{x}_i}(t_2,\Pi_2(X_{t_2}))\bar{v}(t_2+s_{2,i})\,dt_2\,dt_1.
\end{aligned}\tag{8.26}$$

and

$$\bar{h}=h-\frac{1}{2}g_{k_1}g, \text{ and } \bar{v}(t)=\begin{cases}\bar{h}(t,\Pi_2(X_t)), & 0\le t\le a\\ \eta(t), & t<0.\end{cases}\tag{8.27}$$

One can also derive the Milstein scheme for (8.24) using the Stratonovich-Taylor expansion (8.25) of $X(t)$ as follows: Let $t_k < t \le t_{k+1}$. Then

$$\begin{aligned}
X^p(t) = &X^p(t_k)+\bar{h}(t_k,\Pi_2(X^p_{t_k}))(t-t_k)+g(t_k,\Pi_1(X^p_{t_k}))(W(t)-W(t_k)) \\
&+\sum_{i=1}^{k_1}\frac{\partial g}{\partial x_i}(t_k,\Pi_1(X^p_{t_k}))u^p(t_k+s_{1,i})J(t_k+s_{1,i},t+s_{1,i};s_{1,i}),
\end{aligned}\tag{8.28}$$

where

$$u^p(t)=\begin{cases} g(t,\Pi_1(X^p_t)), & t\ge 0,\\ 0, & -1\le t<0.\end{cases}$$

8.3. The multi-dimensional Milstein scheme.

Write $h(s,x)=(h^1(s,x),\cdots,h^m(s,x))^T$, $\vec{x}\in\mathbf{R}^{mk_1}$,

$$\vec{x}=\begin{pmatrix} x_{11},\cdots,x_{1k_1}\\ \cdots \\ x_{m1},\cdots,x_{mk_1}\end{pmatrix}.$$

Denote by $g^{jl}(s,\vec{x})$ the (j,l) element of the $m\times d$ matrix $g(s,\vec{x})$. To simplify notation, we use below the summation convention on repeated indices. Recall the notations for the partition $-1 = t_{-y} < \cdots < t_0 = 0 < \cdots < t_n = t$ introduced in Section 2. We formulate the *Milstein* scheme for the SDDE (1.6) as follows: if $t_k < t \leq t_{k+1}$, the ith coordinate $X^i(t)$ of $X(t) = (X^1(t), \cdots, X^m(t))^T$ is approximated by

(8.29)
$$
\begin{aligned}
X^{i,p}(t) &= X^{i,p}(t_k) + h^i(t_k, \Pi_2(X^p_{t_k}))(t - t_k) + g^{il}(t_k, \Pi_1(X^p_{t_k}))(W^l(t) - W^l(t_k)) \\
&\quad + \frac{\partial g^{il}}{\partial x_{i_1 j_1}}(t_k, \Pi_1(X^p_{t_k}))u^{i_1 j_1, p}(t_k + s_{1,j_1})I_{l,l_1}(t_k + s_{1,j_1}, t + s_{1,j_1}; s_{1,j_1}),
\end{aligned}
$$

where

$$
u^{i_1 j_1, p}(t) = \begin{cases} g^{i_1 j_1}(t, \Pi_1(X^p_t)), & t \geq 0, \\ 0, & -1 \leq t < 0. \end{cases}
$$

Remark 8.5.

One may check that Lemma 8.1, Theorems 8.2 and 8.3 also hold in the multi-dimensional case. In fact, it is easy to extend these results to the multi-dimensional case, thanks to the weak differentiability results (Proposition 6.1, Lemma 6.2 and Proposition 6.3) and the results concerning strong approximation of double Stratonovich integrals (Lemma 7.1 and Lemma 7.2).

In comparison with SODE's, it seems very difficult to derive higher order strong approximation schemes for the SDDE (1.6). One may try to avoid involving the *differential operator* D and the *trace operator* $\bigtriangledown$ in the numerical scheme by attempting to employ multiple Stratonovich integrals instead of multiple Skorohod integrals. The idea is to use Stratonovich-Taylor expansions of the coefficients in the SDDE (1.6) (c.f. (8.4) and (8.5)) instead of Itô-Taylor expansions. However, this is difficult, because it is hard to estimate the order of the error via the remainder term. This is because a multiple (anticipating) Stratonovich integral can not be expressed in terms of multiple (non-anticipating) Itô integrals. The *Hu-Meyer Formula* gives the relationship between multiple Stratonovich and Skorohod integrals ([7], Theorem 3.1 (with non-deterministic kernels); [30], Theorem 3.1 and [28], Theorem 3.4 (with deterministic kernels)) (c.f. [25], [30] and [28]). However, the formula still involves the differential operator D and the trace operator $\bigtriangledown$, and hence it is hard to estimate the remainder term.

One may refer to Jolis and Sanz ([15]), Delgado and Sanz ([7]), Solé and Utzet ([28]), and Zakai ([30]) for multiple Skorohod and multiple Stratonovich integrals.

Appendix A.

The following lemma extends a result by Nualart and Pardoux ([22], Lemma C1).

Lemma A.1.

Suppose that $x = \{x(t) : t \in [0,a]\}$ *is a measurable real-valued process,* $x(t) = 0$ *if* $t > a$ *or* $t < 0$*, and* $x \in L^p([0,a], \mathbf{R})$ *a.s.,* $p > 1$.. *Assume that* $\{\pi_n : 0 = t_0 < t_1 < \cdots < t_n = a\}$ *is a family of partitions of* $[0,a]$*, with* $\lim_{n\to\infty} |\pi_n| = 0$*, and* $-r \le s_1, s_2 \le 0$*. Then*

$$\lim_{n\to\infty} \sum_{l=1}^{n} \frac{\Delta_{l1} W \Delta_{l2} W}{t_l - t_{l-1}} \int_{t_{l-1}+s_1}^{t_l+s_1} x(s)\,ds = \begin{cases} \int_0^{a+s_1} x(s)\,ds, & s_1 = s_2 \\ 0, & s_1 \neq s_2 \end{cases} \tag{A.1}$$

in probability. Moreover, if $x \in L^p(\Omega \times [0,a], \mathbf{R})$*, then the above convergences hold in* $L^1(\Omega, \mathbf{R})$.

Proof.

It clearly suffices to show that (A.1) holds in $L^1(\Omega, \mathbf{R})$ whenever $x \in L^p(\Omega \times [0,a], \mathbf{R})$. Fix $m \ge 1$, define

$$x^m = \sum_{l=1}^{m} \frac{I_{(t_{l-1}+s_1, t_l+s_1]}}{t_l - t_{l-1}} \int_{t_{l-1}+s_1}^{t_l+s_1} x(s)\,ds.$$

For $n \ge 1$, define

$$\alpha_n(x) = \sum_{l=1}^{n} \frac{\Delta_{l1} W \Delta_{l2} W}{t_l - t_{l-1}} \int_{t_{l-1}+s_1}^{t_l+s_1} x(s)\,ds.$$

Define $\alpha_n(X_m)$ similarly. It follows from Hölder's inequality that if $1/p + 1/q = 1$, then

$$E|\alpha_n(x)| \le \left\{ E \sum_{l=1}^{n} \frac{|\Delta_{l1} W \Delta_{l2} W|^q}{(t_l - t_{l-1})^{q-1}} \right\}^{\frac{1}{q}} \left\{ E \sum_{l=1}^{n} \frac{(\int_{t_{l-1}+s_1}^{t_l+s_1} |x(s)|\,ds)^p}{(t_l - t_{l-1})^{\frac{p}{q}}} \right\}^{\frac{1}{p}}, \tag{A.2}$$

i.e.

$$||\alpha_n(x)||_{L^1(\Omega)} \le C_p||x||_{L^p(\Omega\times[0,a+s_1])} \le C_p||x||_{L^p(\Omega\times[0,a])},$$

Therefore,

(A.3)

$$\begin{aligned} E|\alpha_n(x) - \int_0^{a+s_1} x(s)\,ds| &\le E|\alpha_n(x-x^m)| + E|\alpha_n(x^m) - \int_0^{a+s_1} x(s)\,ds| \\ &\le E|\alpha_n(x^m) - \int_0^{a+s_1} x(s)\,ds| + C_p||x-x^m||_{L^p(\Omega\times[0,a+s_1])}, \end{aligned}$$

since

$$\begin{aligned} \alpha_n(x^m) = &\sum_{i=1}^m \left\{ \sum_{\substack{(t_{l-1},t_l]\subseteq(t_{i-1},t_i]\\ 1\le l\le n}} \int_{t_{i-1}+s_1}^{t_i+s_1} \frac{I_{(t_{i-1}+s1,t_i+s_1]}(t)}{t_l - t_{l-1}}\,dt\Delta_{l1}W\Delta_{l2}W \right\} \\ &\times \frac{1}{t_i - t_{i-1}} \int_{t_{i-1}+s_1}^{t_i+s_1} x(s)\,ds \\ = &\sum_{i=1}^m \left\{ \sum_{\substack{(t_{l-1},t_l]\subseteq(t_{i-1},t_i]\\ 1\le l\le n}} \Delta_{l1}W\Delta_{l2}W \right\} \frac{1}{t_i - t_{i-1}} \int_{t_{i-1}+s_1}^{t_i+s_1} x(s)\,ds. \end{aligned}$$

Let k_m be the index such that $t_{k_m-1}+s_1 < 0 \le t_{k_m}+s_1$. If $s_1 = s_2$, then by Lemma 5.2, the following limit exists in probability

$$\begin{aligned} \lim_{n\to\infty} \alpha_n(x^m) &= \sum_{i=1}^m [(t_i+s_1)\wedge 0 - (t_{i-1}+s_1)\vee 0] \frac{1}{t_i - t_{i-1}} \int_{t_{i-1}+s_1}^{t_i+s_1} x(s)\,ds \\ &= \sum_{i=k_m+1}^m \int_{t_{i-1}+s_1}^{t_i+s_1} x(s)\,ds + \frac{t_{k_m}+s_1}{t_{k_m - t_{k_m-1}}} \int_0^{t_{k_m}+s_1} x(s)\,ds \\ &= \int_0^{a+s_1} x(s)\,ds + \frac{t_{k_m}+s_1}{t_{k_m - t_{k_m-1}}} \int_0^{t_{k_m}+s_1} x(s)\,ds. \end{aligned}$$

Equivalently,

$$\alpha_n(x^m) - \int_0^{a+s_1} x(s)\,ds - \frac{t_{k_m}+s_1}{t_{k_m - t_{k_m-1}}} \int_0^{t_{k_m}+s_1} x(s)\,ds \to 0$$

as $n\to\infty$ in probability.

A slight modification in the proof of (A.2) yields the estimate

$$||\alpha_n(x^m)||_{L^{p'}(\Omega)} \leq C(p,p')||x^m||_{L^p(\Omega\times[0,a+s_1])},$$

for all $p' \in (1,p)$. Therefore, the family $\{\alpha_n(x^m) : n \geq 1\}$ is uniformly integrable. From (A.3) we have

$$\begin{aligned}\lim_{n\to\infty} E|\alpha_n(x) - \int_0^{a+s_1} x(s)\,ds| \\ &\leq E|\frac{t_{k_m}+s_1}{t_{k_m-t_{k_m-1}}}\int_0^{t_{k_m}+s_1} x(s)\,ds| + C_p||x-x^m||_{L^p(\Omega\times[0,a+s_1])} \\ &\leq E\int_0^{t_{k_m}+s_1} |x(s)|\,ds + C_p||x-x^m||_{L^p(\Omega\times[0,a+s_1])}.\end{aligned}$$

Clearly, $x^m \to x$ in $L^p(\Omega\times[0,a+s_1])$ and $E\int_0^{t_{k_m}+s_1}|x(s)|\,ds \to 0$ as $m \to \infty$. So

$$\lim_{n\to\infty} E|\alpha_n(x^m) - \int_0^{a+s_1} x(s)\,ds| = 0.$$

Now consider the case $s_1 \neq s_2$. Since

$$\begin{aligned}E|\alpha_n(x)| &\leq E|\alpha_n(x^m)| + E|\alpha_n(x-x^m)| \\ &\leq E|\alpha_n(x^m)| + C_p||x-x^m||_{L^p(\Omega\times[0,a+s_1])}.\end{aligned}$$

A similar argument gives $\lim_{n\to\infty} E|\alpha_n(x)| = 0$. □

The following useful result is due to Föllmer ([8]), and Nualart and Pardoux ([22], Lemma C.2):

Lemma A.2.

Let $\{x^i(t) : 0 \leq t \leq a\}_{i=1}^2$ be two continuous processes, and $\{\pi_n : 0 = t_0 < t_1 < \cdots < t_n = a\}$ a family of partitions of $[0,a]$, with $\lim_{n\to\infty} |\pi_n| = 0$. For each n and $l = 1,\cdots,n$, let $x^i_{t_l,n}$ denote $x^i(t_l)$. Assume that

$$\sum_{l=1}^{n}(x^i_{t_l,n} - x^i_{t_{l-1},n})(x^j_{t_l,n} - x^j_{t_{l-1},n}) \to \int_0^a a^{ij}(s)\,ds \tag{A.4}$$

in probability as $n \to \infty$, *where* $\{a^{ij}(t) : 0 \le t \le a; i, j = 1, 2\}$ *are measurable processes such that a.s.*

$$\int_0^a |a^{ij}(s)|\, ds < \infty, \quad i, j = 1, 2. \tag{A.5}$$

Let $\{Y(t) : 0 \le t \le a\}$ *be a continuous process, and* $\{Y^n(t) : 0 \le t \le a\}_{n=1}^{\infty}$ *be measurable processes which converge a.s. to* $\{Y(t)\}$ *as* $n \to \infty$, *uniformly with respect to* $t \in [0, a]$. *Then*

$$\sum_{l=1}^{n} Y^n(t_{l-1})(x^i_{t_l,n} - x^i_{t_{l-1},n})(x^j_{t_l,n} - x^j_{t_{l-1},n}) \to \int_0^a a^{ij}(s) Y(s)\, ds \tag{A.6}$$

in probability as $n \to \infty$, *for* $i = 1, 2$.

REFERENCES

[1] Ahmed, T. A. *Stochastic Functional Differential Equations with Discontinuous Initial Data*, M.Sc. Thesis, University of Khartoum, Sudan (1983).

[2] Alòs E. and Nualart, D., *An extension of Itô's formula for anticipating processes,* Journal of Theoretical Probability **2** (1998), 493–514.

[3] Asch, J. and Potthoff, J., *Itô's lemma without non-anticipatory conditions,* Probability and Related Fields **88** (1991), 17–46.

[4] Bell, D. and Mohammed, S.-E. A., *The Malliavin calculus and stochastic delay equations,* Journal of Functional Analysis **99 No. 1** (1991), 75–99.

[5] Bell, D. and Mohammed, S.-E. A., *On the solution of stochastic ordinary differential equations via small delays,* Stochastics and Stochastics Reports **28 No. 4** (1989), 293–299.

[6] Cambanis, S. and Hu, Y., *The exact convergence rate of Euler-Maruyama scheme and application to sample design*, Stochastics and Stochastics Report, **59** (1996), 211-240.

[7] Delgado, R. and Sanz, M., *The Hu-Meyer Formula for non-deterministic kernels,* Stochastics and Stochastics Reports **38** (1992), 149–158.

[8] Föllmer, H., *Calcul d'Itô sans probabilités,* Séminaire de Probabilit és XV Lect. Notes Maths. **850,** 143–150, Berlin Heidelberg New York, 1981.

[9] Gentle, J., *Random number generation and Monte Carlo methods,* Statistics and Computing, Springer-Verlag, 1998.

[10] Gaines, J. G. and Lyons, T. J., *Random generation of stochastic area integrals.* SIAM J. Appl. Math. 54 (1994), 1132–1146.

[11] Hu, Y., *Strong and weak order of time discretization schemes of stochastic differential equations*, In Séminaire de Probabilités XXX, ed. by J. Azema, P.A. Meyer and M. Yor, Lecture Notes in Mathematics **1626,** Springer-Verlag, 1996, 218-227.

[12] Hu, Y., *Optimal times to observe in the Kalman-Bucy model*, Stochastics and Stochastic Reports **69** (2000), 123-140.

[13] Hu, Y. and Mohammed, S.-E. A., *Numerical simulation of stochastic delay equations,* (preprint) (January, 1997), pp. 11.

[14] Hu, Y. and Nualart, D., *Continuity of some anticipating integral processes,* Statistics and Probability Letters **37** (1998), 203–211.

[15] Jolis, M. and Sanz, M., *On generalized multiple stochastic integrals and multiparameter anticipative calculus,* Stochastic Analysis and Related Topics II, Lecture Notes in Mathematics **1444,** 141–182, Springer-Verlag, 1988.

[16] Kloeden P. and Platen, R., *Numerical Solution of Stochastic Differential Equations,* Springer-Verlag, 1992.

[17] Kloeden P., Platen, R., and Schurz, H., *Numerical Solution of SDE Through Computer Experiments,* Springer-Verlag, 1994.

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

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

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

[21] Mohammed, S.-E. A., *Stochastic Differential Systems with Memory: Theory, Examples and Application, Geilo Workshop 1996,* Pitman Advanced Publishing Program, 1984.

[22] Nualart, D. and Pardoux, E., *Stochastic calculus with anticipating integrands,* Probability Theory and Related fields **78** (1988), 535–581.

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

[24] Rosinski,J., *On stochastic integration by series of Wiener integrals,* Applied Mathematics and Optimization **19** (1989), 137–155.

[25] Russo, F. and Vallois P., *Forward, backward and symmetric stochastic integration,* Probability Theory and Related fields **97** (1993), 403–421.

[26] Ryden, T. and Wiktorsson, M., *On the simulation of iterated Ito integrals,* Stochastic Process and Appl. **91** (2001), 151-168.

[27] Scheutzow, M., *Qualitative behavior of stochastic delay equations with a bounded memory,* Stochastics **12 no. 1** (1984), 41–80.

[28] Solé, J. and Utzet, F., *Stratonovich integral and trace,* Stochastics and Stochastics Reports **29** (1990), 203–220.

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

[30] Zakai, M., *Stochastic integration, trace and the skeleton of Wiener functionals,* Stochastics and Stochastics Reports **32** (1990), 93–108.

Yaozhong Hu,

Department of Mathematics,

University of Kansas, Lawrence,

Kansas 66045-2142, USA.

Email: hu@math.ukans.edu

Salah-Eldin A. Mohammed

Department of Mathematics,

Southern Illinois University at Carbondale.

Carbondale, IL 62901, USA.

Email: salah@sfde.math.siu.edu

Feng Yan

Williams Energy Marketing and Trading

One Williams Center, WRC2-4,

Tulsa, OK 74119, USA.

Email: fyan1@yahoo.com