
ABSTRACT

In the present work we have gone a step forward towards integration by part of higher order Malliavin derivatives by formulating and extending some formulas and results on Malliavin calculus and ordinary stochastic differential equations to include delay stochastic differential equations as well as ordinary SDE's. Here we have also stated clearly what we mean by the Malliavin derivatives and densities of distributions of the solutions process for delay stochastic differential equations which we are considering.

KEYWORDS: Stochastic Differential Equations, Malliavin Calculus, Euler Scheme for delay SDE's, Integration by Parts, Densities of Distributions.

INTRODUCTION, NOTATIONS AND DEFINITIONS

In Chapter 1 of the Ph.D. thesis of Ahmed we have proved the existence and uniqueness of a solution for certain types of delay (functional) stochastic differential equations (delay SDE's) with discontinuous initial data, see also , and the web cite www.sfde.math.siu.edu. See the delay SDE ([1.1.1]) in the present work. Here we establish an integration by parts formula involving solutions to such type of delay (functional) SDE's. The integration by parts formula which we establish can be used to extend the formulas in and to include delay SDE's as well as ordinary SDE's. In this work we also establish some other useful applications to delay SDE's. Generally speaking we can say that our work extends the first three chapters of the work by Norris to include delay SDE's as well as ordinary SDE's; see Theorems 2.3, 3.1 and 3.2 in . We will also show in a sequel paper to this work that the distribution of the solution process has smooth density. Also we will establish an integration by parts formula involving Malliavin derivatives of higher order.

Notations and Definitions

The following notations and definitions will be used throughout this work: $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space; T is a positive real number; $\{\mathcal{F}_t\}_{t \in [0, T]}$ is an increasing family of sub- σ algebras of \mathcal{F} , each of which contains all null subsets of Ω ; \mathbb{N} is the set of natural numbers; $W = (W^1, \dots, W^r): [0, T] \times \Omega \rightarrow \mathbb{R}^r$ is a r -dimensional normalized Brownian motion. If X is a topological space, then $\mathcal{B}(X)$ denotes its Borel field. The symbol λ refers to the Lebesgue measure on \mathbb{R}^d , and $|\cdot|$ denotes the Euclidean norm on \mathbb{R}^d , $d \in \mathbb{N}$.

Let G be a Banach space and let \mathcal{A} be a sub- σ algebra of \mathcal{F} containing all subsets of measure zero in \mathcal{F} , then $\mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P}; G)$ denotes the space of all functions $f: \Omega \rightarrow G$ which are \mathcal{A} - $\mathcal{B}(G)$ measurable and are such that $\int_{\Omega} \|f\|_G^2 d\mathbb{P} < \infty$.

The symbol $L^2(\Omega, \mathcal{A}, \mathbb{P}; G)$ denotes the Banach space (with norm determined by $\|f\|_{L^2}^2 = \int_{\Omega} \|f(\omega)\|_G^2 d\mathbb{P}$) of all equivalence classes of functions $f: \Omega \rightarrow G$ which are \mathcal{A} - $\mathcal{B}(G)$ measurable and which are such that $\int_{\Omega} \|f\|_G^2 d\mathbb{P} < \infty$. The symbol $L(\mathbb{R}^m, \mathbb{R}^n)$ ($m, n \in \mathbb{N}$) denotes the space of all linear maps from \mathbb{R}^m to \mathbb{R}^n . The symbol J refers to the interval $[-1, 0)$, and $\mathcal{H}(J)$ or $\mathcal{B}(J)$ refers to the Borel field on J .

If $X: [-1, T] \times \Omega \rightarrow \mathbb{R}^d$ is a process, then for each $t \in [0, T]$ and $\omega \in \Omega$ we define the map: $X_t: \Omega \rightarrow \mathcal{L}^2(J, \mathbb{R}^d)$ by $X_t(\omega)(s) = X(t+s, \omega)$ for all $s \in J$ and almost all ω . For each $0 \leq t \leq T$ we write $\|(X(t), X_t)\|^2 = \|X(t)\|^2 + \|X_t\|^2$. Let the function V belong to $\mathcal{L}^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d)$, θ belong to $\mathcal{L}^2(J \times \Omega, \mathcal{H}(J) \otimes \mathcal{F}_0, \lambda \otimes \mathbb{P}; \mathbb{R}^d)$, and for $\ell = 1, 2, \dots, r$ let f, g^ℓ be functions from $[0, T] \times \Omega \times \mathbb{R}^d \times \mathcal{L}^2(J, \mathbb{R}^d)$ to \mathbb{R}^d . Then a process $X: [-1, T] \times \Omega \rightarrow \mathbb{R}^d$ is called a solution of the delay SDE with integral form

$$X(t) = \begin{cases} V + \int_0^t f(u, X(u), X_u) du + \sum_{\ell=1}^r \int_0^t g^\ell(u, X(u), X_u) dW^\ell(u), & 0 \leq t \leq T, \\ \theta(t), & t \in J, \end{cases} \quad (1.1)$$

if

1. X is $\mathcal{B}([0, T]) \otimes \mathcal{F}$ - $\mathcal{B}(\mathbb{R}^d)$ measurable;
2. For each $t \in [0, T]$, the process $X(t, \cdot)$ is \mathcal{F}_t - $\mathcal{B}(\mathbb{R}^d)$ measurable, and for each $t \in J$, the process $X(t, \cdot)$ is \mathcal{F}_0 - $\mathcal{B}(\mathbb{R}^d)$ measurable;
3. $X \in \mathcal{L}^2([-1, T] \times \Omega, \mathcal{H} \times \mathcal{F}, \lambda \times \mathbb{P}; \mathbb{R}^d)$,
4. X satisfies the delay SDE ((1.1.1)).

The following conditions are sufficient for the existence of a unique solution to ((1.1)) (see and).

1. $V \in \mathcal{L}^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d)$.
2. $\theta \in \mathcal{L}^2(J \times \Omega, \mathcal{H} \otimes \mathcal{F}_0, \lambda \otimes \mathbb{P}, \mathbb{R}^d)$.
3. $f, g^\ell: [0, T] \times \Omega \times \mathbb{R}^d \times \mathcal{L}^2(J, \mathbb{R}^d) \rightarrow \mathbb{R}^d$ are such that
 1. f and g^ℓ are $\mathcal{B}([0, T]) \otimes \mathcal{F} \otimes \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(L^2(J, \mathbb{R}^d))$ - $\mathcal{B}(\mathbb{R}^d)$ measurable.
 2. For each $t \in [0, T]$, the stochastic variables $f(t, \cdot, \cdot, \cdot)$ and $g^\ell(t, \cdot, \cdot, \cdot)$ are $\mathcal{F}_t \otimes \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(L^2(J, \mathbb{R}^d))$ - $\mathcal{B}(\mathbb{R}^d)$ measurable.
 3. There exists a constant K and a function $\zeta \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ such that

$$|f(t, \omega, s, h)| + \sum_{\ell=1}^r |g^\ell(t, \omega, s, h)| \leq K(|s| + \|h\| + |\zeta(\omega)|) \quad (1.2)$$

4. for almost all ω and for all $t \in [0, T]$; $s \in \mathbb{R}^d$ and h belongs to $\mathcal{L}^2(J, \mathbb{R}^d)$.
5. There exists a constant K' such that, for almost all ω ,

$$|f(t, \omega, s, h_1) - f(t, \omega, u, h_2)| + \sum_{\ell=1}^r |g^\ell(t, \omega, s, h_1) - g^\ell(t, \omega, u, h_2)| \leq K'(|s - u| + \|h_1 - h_2\|) \quad (1.3)$$

6. for all $t \in [0, T]$; for all $s, u \in \mathbb{R}^d$, and for all $h_1, h_2 \in \mathcal{L}^2(J, \mathbb{R}^d)$.

INTEGRATION BY PARTS FORMULA

In the beginning of this section we recall the following five basic numbered equations and definitions, See . For

$(X(0), X_0) = (x, \xi) \in \mathbb{R}^d \times L^2(J, \mathbb{R}^d)$, let $v \mapsto D^v X^{x, \xi}(t)$, be the Malliavin derivative of the solution process $X^{x, \xi}(t)$. We write $D^v X_t^{x, \xi}(\vartheta) = D^v X^{x, \xi}(t + \vartheta)$ ($t \in [0, T]$, $\vartheta \in J = [-1, 0)$) for its time delay.

In the following definition we give a precise definition of the Malliavin derivative of a real-valued functional F of Brownian motion.

[D:Malliavin] Let $F((W(s))_{0 \leq s \leq T})$ be a functional of r -dimensional Brownian motion, and let

$$v(t) = (v^1(t), \dots, v^r(t))^* = \begin{pmatrix} v^1(t) \\ \vdots \\ v^r(t) \end{pmatrix} \text{ be a deterministic vector-valued function in } L^2([0, T], \mathbb{R}^r \otimes \mathbb{R}^d).$$

Then $D^v F((W(s))_{0 \leq s \leq T})$ is given by the limit:

$$D^v F((W(s))_{0 \leq s \leq T}) = \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \left(F \left((W(s) + \varepsilon \int_0^s v(\sigma) d\sigma)_{0 \leq s \leq T} \right) - F((W(s))_{0 \leq s \leq T}) \right). \quad (2.1)$$

The mapping $v \mapsto D^v F((W(s))_{0 \leq s \leq T})$ is a linear map (functional) from the space $L^2([0, T], \mathbb{R}^r \otimes \mathbb{R}^d)$ to \mathbb{R} . Here $\mathbb{R}^r \otimes \mathbb{R}^d$ denotes the space of all $r \times d$ -matrices (r rows, d columns).

Notice that, for $v(t) = (v^1(t), \dots, v^r(t))^T = \begin{pmatrix} v^1(t) \\ \vdots \\ v^r(t) \end{pmatrix}$ be a deterministic matrix-valued function in

$L^2([0, T], \mathbb{R}^r \otimes \mathbb{R}^d)$, $U^v(t)$ can be considered as a $d \times d$ -matrix where each entry is an \mathbb{R} -valued adapted stochastic process; U_t^v can be considered as a $d \times d$ -matrix where each entry is an $L^2(J, \mathbb{R})$ -valued adapted stochastic process. If $M = (m_{jk})_{1 \leq j \leq d, 1 \leq k \leq r}$ is a real $d \times r$ matrix, then $M^T = (m_{kj})_{1 \leq k \leq r, 1 \leq j \leq d}$ denotes its transposed: it is $r \times d$ matrix with entries m_{kj} .

The process $D^v X_t^{x, \xi}(\cdot)$ satisfies the following delay stochastic differential equation:

$$\begin{aligned}
 dD^\nu X_t(\vartheta) &= dD^\nu X(t + \vartheta) \\
 &= \left(\frac{\partial f}{\partial x}(t + \vartheta, X(t + \vartheta), X_{t+\vartheta}) D^\nu X(t + \vartheta) \right. \\
 &\quad \left. + \int_J \frac{\partial f}{\partial \xi}(t + \vartheta, X(t + \vartheta), X_{t+\vartheta})(\varphi) D^\nu X_{t+\vartheta}(\varphi) d\varphi \right) dt \\
 &\quad + \sum_{\ell=1}^r \frac{\partial g^\ell}{\partial x}(t + \vartheta, X(t + \vartheta), X_{t+\vartheta}) D^\nu X(t + \vartheta) dW^\ell(t + \vartheta) \\
 &\quad + \sum_{\ell=1}^r \int_J \frac{\partial g^\ell}{\partial \xi}(t + \vartheta, X(t + \vartheta), X_{t+\vartheta})(\varphi) D^\nu X_{t+\vartheta}(\varphi) d\varphi dW^\ell(t + \vartheta) \\
 &\quad + \sum_{\ell=1}^r g^\ell(t + \vartheta, X(t + \vartheta), X_{t+\vartheta}) v^\ell(t + \vartheta, X(t + \vartheta), X_{t+\vartheta}) dt,
 \end{aligned} \tag{2.2}$$

where ϑ belongs to J . If $t + \vartheta$ belongs to J we replace $t + \vartheta$ with 0 in ([E: SDEDP]). If $\vartheta = 0$ we obtain the delay stochastic differential equation for the process $D^\nu X(t)$:

$$\begin{aligned}
 dD^\nu X(t) &= \left(\frac{\partial f}{\partial x}(t, X(t), X_t) D^\nu X(t) + \int_J \frac{\partial f}{\partial \xi}(t, X(t), X_t)(\vartheta) D^\nu X_t(\vartheta) d\vartheta \right) dt \\
 &\quad + \sum_{\ell=1}^r \left(\frac{\partial g^\ell}{\partial x}(t, X(t), X_t) D^\nu X(t) + \int_J \frac{\partial g^\ell}{\partial \xi}(t, X(t), X_t)(\vartheta) D^\nu X_t(\vartheta) d\vartheta \right) dW^\ell(t) \\
 &\quad + \sum_{\ell=1}^r g^\ell(t, X(t), X_t) v^\ell(t, X(t), X_t) dt.
 \end{aligned} \tag{2.3}$$

We also write $U_{11}^{x,\xi}(t) = \frac{\partial}{\partial x} X^{x,\xi}(t)$, and $U_{12}^{x,\xi}(t) = \frac{\partial}{\partial \xi} X^{x,\xi}(t)$. In addition, we write $U_{21}^{x,\xi}(t) = \frac{\partial}{\partial x} X_t^{x,\xi} = U_{11,t}^{x,\xi}$ (the delay of $U_{11}^{x,\xi}(t)$), and $U_{22}^{x,\xi}(t) = \frac{\partial}{\partial \xi} X_t^{x,\xi} = U_{12,t}^{x,\xi}$, the delay of the process $U_{12}^{x,\xi}(t)$. The matrix $U_{11}^{x,\xi}(t)$ can be identified with an operator from \mathbb{R}^d to itself, the matrix $U_{12}^{x,\xi}(t)$ can be considered as a linear mapping from $L^2(J, \mathbb{R}^d)$ to \mathbb{R}^d , the matrix $U_{21}^{x,\xi}(t)$ as a mapping from \mathbb{R}^d to $L^2(J, \mathbb{R}^d)$, and, finally, $U_{22}^{x,\xi}(t)$ as a mapping from $L^2(J, \mathbb{R}^d)$ to itself. Notice that $U_{11}^{x,\xi}(t)$ can be considered as $d \times d$ -matrix where each entry is an \mathbb{R} -valued adapted stochastic process; $U_{12}^{x,\xi}(t)$ can be considered as $d \times d$ -matrix where each entry is an $L^2(J, \mathbb{R})$ -valued adapted stochastic process. To be precise, write the solution process as a d -vector $X^{x,\xi}(t) = (X_1^{x,\xi}(t), \dots, X_d^{x,\xi}(t))$, and consider the mapping ($1 \leq j, k \leq d$)

$$\xi_k \rightarrow X_j^{x, (\xi_1, \dots, \xi_{k-1}, \xi_k, \xi_{k+1}, \dots, \xi_d)}(t),$$

(2.4)

which is a mapping from $L^2(J, \mathbb{R})$ to \mathbb{R} , and where each variable $\xi_k, \forall k=1, \dots, d$, is a fixed function in $L^2(J, \mathbb{R})$.

The derivative of the function in ([E: mapping]) can be considered as a continuous linear functional on $L^2(J, \mathbb{R})$.

Therefore it can be represented as an inner-product with a function in $L^2(J, \mathbb{R})$, which is denoted by $\frac{\partial X_j^{x, \xi}(t)}{\partial \xi_k}$.

Consequently, we write

$$\frac{\partial X_j^{x, \xi}(t)}{\partial \xi_k}(\eta) = \lim_{h \rightarrow 0} \frac{X_j^{x, (\xi_1, \dots, \xi_{k-1}, \xi_k+h\eta, \xi_{k+1}, \dots, \xi_d)}(t) - X_j^{x, (\xi_1, \dots, \xi_{k-1}, \xi_k, \xi_{k+1}, \dots, \xi_d)}(t)}{h}$$

$$= \int_J \eta(\varphi) \frac{\partial X_j^{x, \xi}(t)}{\partial \xi_k}(\varphi) d\varphi, \quad \eta \in L^2(J, \mathbb{R}).$$

(2.5)

After giving a brief introduction to our work, we are now ready to continue the work that we have started in .

Here, and in the sequel, we write $f(t)$ and $g^\ell(t)$ instead of $f(t, X^{x, \xi}(t), X_t^{x, \xi})$ and $g^\ell(t, X^{x, \xi}(t), X_t^{x, \xi})$

respectively. For a concise formulation of the stochastic differential equation for the matrix-valued process $(U(t): t \geq 0)$ and its inverse we introduce the following *stochastic differentials*:

$$h_x(t) = \frac{\partial f}{\partial x}(t) dt + \sum_{\ell=1}^r \frac{\partial g^\ell}{\partial x}(t) dW^\ell(t);$$

$$h_\xi(t) = \frac{\partial f}{\partial \xi}(t) dt + \sum_{\ell=1}^r \frac{\partial g^\ell}{\partial \xi}(t) dW^\ell(t)$$

$$h_\xi(t, \vartheta) = \frac{\partial f}{\partial \xi}(t, \vartheta) dt + \sum_{\ell=1}^r \frac{\partial g^\ell}{\partial \xi}(t, \vartheta) dW^\ell(t)$$

(2.6) (2.7) & (2.8)

First we introduce the following definitions:

$$U_{11}(t) = \frac{\partial}{\partial x} X^{x, \xi}(t), \quad U_{12}(t) = \frac{\partial}{\partial \xi} X^{x, \xi}(t), \quad U_{21}(t) = \frac{\partial}{\partial x} X_t^{x, \xi}, \quad U_{22}(t) = \frac{\partial}{\partial \xi} X_t^{x, \xi}.$$

$$\tilde{U}(t) = \begin{pmatrix} U_{11}(t) & U_{12}(t) \\ U_{21}(t) & U_{22}(t) \end{pmatrix}$$

Next we consider the new process

satisfying the delay SDE

$$d\tilde{U}(t) = h_x(t)\tilde{U}(t) + \int_J h_\xi(t, \vartheta)\tilde{U}_t(\vartheta) d\vartheta,$$

(2.9)

The delay SDE ([E: Ito(6)]) can also be written as

$$d \begin{pmatrix} U_{11}(t) & U_{12}(t) \\ U_{21}(t) & U_{22}(t) \end{pmatrix} = \begin{pmatrix} h_x(t)U_{11}(t) & h_x(t)U_{12}(t) \\ h_x(t)U_{21}(t) & h_x(t)U_{22}(t) \end{pmatrix} + \begin{pmatrix} \int_J h_\xi(t, \vartheta)U_{11,t}(\vartheta) d\vartheta & \int_J h_\xi(t, \vartheta)U_{12,t}(\vartheta) d\vartheta \\ \int_J h_\xi(t, \vartheta)U_{21,t}(\vartheta) d\vartheta & \int_J h_\xi(t, \vartheta)U_{22,t}(\vartheta) d\vartheta \end{pmatrix} \quad (2.10)$$

We can thus split ([E: Umatrix]) into the following four delay SDE's

$$dU_{11}(t) = h_x(t)U_{11}(t) + \int_J h_{\xi, \vartheta}(t)U_{11,t}(\vartheta)d\vartheta; \quad (2.11)$$

$$dU_{12}(t) = h_x(t)U_{12}(t) + \int_J h_\xi(t, \vartheta)U_{12,t}(\vartheta) d\vartheta; \quad (2.12)$$

$$dU_{21}(t) = h_x(t)U_{21}(t) + \int_J h_\xi(t, \vartheta)U_{21,t}(\vartheta) d\vartheta; \quad (2.13)$$

$$dU_{22}(t) = h_x(t)U_{22}(t) + \int_J h_\xi(t, \vartheta)U_{22,t}(\vartheta) d\vartheta. \quad (2.14)$$

Recall that the process $(\tilde{U}(t), t \geq 0)$ satisfies

$$d\tilde{U}(t) = h_x(t)\tilde{U}(t) + \int_J h_\xi(t, \vartheta)\tilde{U}_t(\vartheta) d\vartheta, \quad (2.15)$$

Then it's delay $(\tilde{U}_t, t \geq 0)$ satisfies

$$d\tilde{U}_t(\cdot) = h_{x,t}(\cdot)\tilde{U}_t(\cdot) + \int_J h_{\xi,t}(\vartheta + \cdot)\tilde{U}_{t+\cdot}(\vartheta)d\vartheta, \quad (2.16)$$

where for, $\vartheta \in J$, (2.16) is equivalent to

$$d\tilde{U}(t + \vartheta) = h_x(t + \vartheta)\tilde{U}(t + \vartheta) + \int_J h_\xi(t + \vartheta + \varphi)\tilde{U}_{t+\vartheta}(\varphi)d\varphi, \quad (2.17)$$

Next we recall the delay SDE for $V(t)$, namely

$$\begin{aligned}
 dV(t) = & -V(t) \frac{\partial f}{\partial x}(t) dt - V(t) \sum_{\ell=1}^r \frac{\partial g^{\ell}}{\partial x}(t) dW^{\ell}(t) \\
 & - \sum_{\ell=1}^r \int_J V_t(\vartheta) \frac{\partial g^{\ell}}{\partial \xi}(t, \vartheta) d\vartheta dW^{\ell}(t) - V(t) \int_J \frac{\partial f}{\partial \xi}(t, \vartheta) U_t(\vartheta) d\vartheta V(t) dt \\
 & + V(t) \sum_{\ell=1}^r \left(\frac{\partial g^{\ell}}{\partial x}(t) + \int_J \frac{\partial g^{\ell}}{\partial \xi}(t, \vartheta) U_t(\vartheta) d\vartheta V(t) \right)^2 dt
 \end{aligned} \tag{2.18}$$

Observe that the delay SDE ([E: V 3]) is an extension of the SDE (3.3) in Norris to include delay SDE's as well as ordinary SDE's. We can see this by considering only the terms in ([E: V 3]) which include derivatives of the coefficients with respect to the space variable and in the same time it contain no derivative with respect to the delay variable. If we do this then we are automatically in the Norris case of SDE's.

Now we rewrite the four delay SDE's in (2.11), (2.12), (2.13), and (2.14) as in (2.19), (2.20), (2.21), and (2.22) respectively.

First we recall the following definitions:

$$U_{11}(t) = \frac{\partial}{\partial x} X^{x,\xi}(t), \quad U_{12}(t) = \frac{\partial}{\partial \xi} X^{x,\xi}(t), \quad U_{21}(t) = \frac{\partial}{\partial x} X_t^{x,\xi}, \quad U_{22}(t) = \frac{\partial}{\partial \xi} X_t^{x,\xi}.$$

We begin with a delay stochastic differential equation for the process $U_{11}(t)$:

$$\begin{aligned}
 dU_{11}(t) & = \left(\frac{\partial f}{\partial x}(t, X(t), X_t) U_{11}(t) + \int_J \frac{\partial f}{\partial \xi}(t, X(t), X_t)(\vartheta) U_{11,t}(\vartheta) d\vartheta \right) dt \\
 & + \sum_{\ell=1}^r \left(\frac{\partial g^{\ell}}{\partial x}(t, X(t), X_t) U_{11}(t) + \int_J \frac{\partial g^{\ell}}{\partial \xi}(t, X(t), X_t)(\vartheta) U_{11,t}(\vartheta) d\vartheta \right) dW^{\ell}(t).
 \end{aligned} \tag{2.19}$$

Next we also see that the process $U_{12}^{x,\xi}(t)$ satisfies the following delay stochastic differential equation:

$$\begin{aligned}
 dU_{12}(t) & = \left(\frac{\partial f}{\partial x}(t, X(t), X_t) U_{12}(t) + \int_J \frac{\partial f}{\partial \xi}(t, X(t), X_t)(\vartheta) U_{12,t}(\vartheta) d\vartheta \right) dt \\
 & + \sum_{\ell=1}^r \left(\frac{\partial g^{\ell}}{\partial x}(t, X(t), X_t) U_{12}(t) + \int_J \frac{\partial g^{\ell}}{\partial \xi}(t, X(t), X_t)(\vartheta) U_{12,t}(\vartheta) d\vartheta \right) dW^{\ell}(t),
 \end{aligned} \tag{2.20}$$

Then we also observe that the matrix-valued process $U_{21}^{x,\xi}(t)$ satisfies the following delay stochastic differential equation:

$$\begin{aligned}
 dU_{21}(t) & = \left(\frac{\partial f}{\partial x}(t, X(t), X_t) U_{21}(t) + \int_J \frac{\partial f}{\partial \xi}(t, X(t), X_t)(\vartheta) U_{21,t}(\vartheta) d\vartheta \right) dt \\
 & + \sum_{\ell=1}^r \left(\frac{\partial g^{\ell}}{\partial x}(t, X(t), X_t) U_{21}(t) + \int_J \frac{\partial g^{\ell}}{\partial \xi}(t, X(t), X_t)(\vartheta) U_{21,t}(\vartheta) d\vartheta \right) dW^{\ell}(t),
 \end{aligned} \tag{2.21}$$

where $U_{21,t}$ is the delay of $U_{21}(t)$, in other words the second delay of $U_{11}(t)$. Similarly the process $U_{22}^{x,\xi}(t)$

satisfies the following delay stochastic differential equation:

$$dU_{22}(t) = \left(\frac{\partial f}{\partial x}(t, X(t), X_t)U_{22}(t) + \int_J \frac{\partial f}{\partial \xi}(t, X(t), X_t)(\vartheta)U_{22,t}(\vartheta) d\vartheta \right) dt + \sum_{\ell=1}^r \left(\frac{\partial g^\ell}{\partial x}(t, X(t), X_t)U_{22}(t) + \int_J \frac{\partial g^\ell}{\partial \xi}(t, X(t), X_t)(\vartheta)U_{22,t}(\vartheta) d\vartheta \right) dW^\ell(t), \quad (2.22)$$

where $U_{22,t}$ is the delay of $U_{22}(t)$, in other words the second delay of $U_{12}(t)$. We shall recall the solutions to the equations (2.21) and (2.22), the delayed space flow and the delayed L^2 -flow respectively. Next we can see that the delay SDE's in (2.11),

(2.12), (2.13) and (2.14) are equivalent to the delay SDE's (2.19), (2.20), (2.21) and (2.22) respectively. In rewriting the above four delay SDE's (2.19), (2.20), (2.21)

and (2.22) we have used the fact that the following four pairs of differentials are equivalent where $i, j = 1, 2$:

$$\frac{\partial f}{\partial x}(t, X(t), X_t)U_{ij}(t)dt = \frac{\partial f}{\partial x}(t, X(t), X_t)dU_{ij}(t) \quad (2.23)$$

$$\frac{\partial g^\ell}{\partial x}(t, X(t), X_t)U_{ij}(t)dW^\ell(t) = \frac{\partial g^\ell}{\partial x}(t, X(t), X_t)dW^\ell(t)U_{ij}(t) \quad (2.24)$$

$$\int_J \frac{\partial f}{\partial \xi}(t, X(t), X_t)(\vartheta)U_{ij,t}(\vartheta) d\vartheta dt = \int_J \frac{\partial f}{\partial \xi}(t, X(t), X_t)(\vartheta) d\vartheta dU_{ij,t} \quad (2.25)$$

$$\int_J \frac{\partial g^\ell}{\partial \xi}(t, X(t), X_t)(\vartheta)U_{ij,t}(\vartheta) d\vartheta dW^\ell(t) = \int_J \frac{\partial g^\ell}{\partial \xi}(t, X(t), X_t)(\vartheta) dW^\ell(t)U_{ij,t}(\vartheta) d\vartheta \quad (2.26)$$

We notice that the j -th row, $1 \leq j \leq d$, of the matrix $\frac{\partial f}{\partial x}$ is given by:

$$\left(\frac{\partial f^j}{\partial x_1}(t, x, \xi), \dots, \frac{\partial f^j}{\partial x_d}(t, x, \xi) \right).$$

We also notice that the k -th columns, $1 \leq k \leq d$, of the matrices $U_{11}(t) = U_{11}^{x,\xi}(t)$ and $U_{21}(t) = U_{11,t}$ are respectively given by

$$\begin{pmatrix} \frac{\partial X^{1,x,\xi}(t)}{\partial x_k} \\ \vdots \\ \frac{\partial X^{d,x,\xi}(t)}{\partial x_k} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \frac{\partial X_t^{1,x,\xi}}{\partial x_k} \\ \vdots \\ \frac{\partial X_t^{d,x,\xi}}{\partial x_k} \end{pmatrix}.$$

Consequently the derivative $\frac{\partial f^j}{\partial x_k}(t, X^{x,\xi}(t), X_t^{x,\xi})$ is given by

$$\begin{aligned} \frac{\partial f^j}{\partial x_k}(t, X^{x,\xi}(t), X_t^{x,\xi}) &= \sum_{m=1}^d \frac{\partial f^j}{\partial x_m}(t, X^{x,\xi}(t), X_t^{x,\xi}) \frac{\partial X^{m,x,\xi}(t)}{\partial x_k} \\ &+ \sum_{m=1}^d \int_J \frac{\partial f^j}{\partial \xi_m}(t, X^{x,\xi}(t), X_t^{x,\xi})(\vartheta) \frac{\partial X_t^{m,x,\xi}}{\partial x_k}(\vartheta) d\vartheta \\ &= \sum_{m=1}^d \frac{\partial f^j}{\partial x_m}(t, X^{x,\xi}(t), X_t^{x,\xi}) U_{11}^{mk}(t) \\ &+ \sum_{m=1}^d \int_J \frac{\partial f^j}{\partial \xi_m}(t, X^{x,\xi}(t), X_t^{x,\xi})(\vartheta) U_{11,t}^{mk}(\vartheta) d\vartheta. \end{aligned} \tag{2.27}$$

It goes without saying that the derivative $\frac{\partial f^j(t, X^{x,\xi}(t), X_t^{x,\xi})}{\partial x_k}$ in the left side of (2.27) is the derivative of the function

$x_k \mapsto f^j(t, X^{x,\xi}(t), X_t^{x,\xi})$, with $x = (x_1, \dots, x_{k-1}, x_k, x_{k+1}, \dots, x_d)$, where the coordinates,

$x_j, j \neq k$, are frozen. By the same token the derivative $\frac{\partial f^j}{\partial x_m}(t, X^{x,\xi}(t), X_t^{x,\xi})$ in the right side of (2.27) is

the derivative of the function

$y_m \mapsto f^j(t, X^{1,x,\xi}(t), \dots, X^{m-1,x,\xi}(t), y_m, X^{m+1,x,\xi}(t), \dots, X^{d,x,\xi}(t), X_t^{x,\xi})$

at $y_m = X^{m,x,\xi}(t)$. Similar conventions are used for the $L^2(J)$ -valued derivatives:

$$\frac{\partial f^j}{\partial \xi_m}(t, X^{x,\xi}(t), X_t^{x,\xi}) \text{ and } \frac{\partial X^{m,x,\xi}(t)}{\partial \xi_k}.$$

Of course, for the derivatives of $g^\ell(t, X^{x,\xi}(t), X_t^{x,\xi})$ we also employ this kind of expressions. Consider the following delay SDE

$$dX^{x,\xi}(t) = f(t, X^{x,\xi}(t), X_t^{x,\xi}) dt + \sum_{\ell=1}^r g^\ell(t, X^{x,\xi}(t), X_t^{x,\xi}) dW^\ell(t) \tag{2.28}$$

We recall the delay SDE for $U(t)$

$$dU(t) = h_x(t)U(t) + \int_J h_\xi(t, \vartheta)U_t(\vartheta)d\vartheta, \tag{2.29}$$

Next we recall the delay SDE for $V(t)$, namely

$$\begin{aligned}
 dV(t) = & -V(t) \frac{\partial f}{\partial x}(t) dt - V(t) \sum_{\ell=1}^r \frac{\partial g^\ell}{\partial x}(t) dW^\ell(t) \\
 & - \sum_{\ell=1}^r \int_J V_t(\vartheta) \frac{\partial g^\ell}{\partial \xi}(t, \vartheta) d\vartheta dW^\ell(t) - V(t) \int_J \frac{\partial f}{\partial \xi}(t, \vartheta) U_t(\vartheta) d\vartheta V(t) dt \\
 & + V(t) \sum_{\ell=1}^r \left(\frac{\partial g^\ell}{\partial x}(t) \right)^2 dt + V(t) \sum_{\ell=1}^r \frac{\partial g^\ell}{\partial x}(t) \int_J \frac{\partial g^\ell}{\partial \xi}(t, \vartheta) U_t(\vartheta) d\vartheta V(t) dt \\
 & + V(t) \sum_{\ell=1}^r \int_J \frac{\partial g^\ell}{\partial \xi}(t, \vartheta) U_t(\vartheta) d\vartheta V(t) \frac{\partial g^\ell}{\partial x}(t) dt \\
 & + V(t) \sum_{\ell=1}^r \int_J \left(\frac{\partial g^\ell}{\partial \xi}(t, \vartheta) U_t(\vartheta) d\vartheta V(t) \right)^2 dt \\
 = & -V(t) \frac{\partial f}{\partial x}(t) dt - V(t) \sum_{\ell=1}^r \frac{\partial g^\ell}{\partial x}(t) dW^\ell(t) \\
 & - \sum_{\ell=1}^r \int_J V_t(\vartheta) \frac{\partial g^\ell}{\partial \xi}(t, \vartheta) d\vartheta dW^\ell(t) - V(t) \int_J \frac{\partial f}{\partial \xi}(t, \vartheta) U_t(\vartheta) d\vartheta V(t) dt \\
 & + V(t) \sum_{\ell=1}^r \left(\frac{\partial g^\ell}{\partial x}(t) + \int_J \frac{\partial g^\ell}{\partial \xi}(t, \vartheta) U_t(\vartheta) d\vartheta V(t) \right)^2 dt.
 \end{aligned}$$

(2.30)

If we have no delay variables (in other words if we have only state variables), then the three delay SDE's (2.28), (2.15), and (2.30) take the forms of the following three SDE's respectively

$$\begin{aligned}
 dX^x(t) &= f(t, X^x(t)) dt + \sum_{\ell=1}^r g^\ell(t, X^x(t)) dW^\ell(t); \\
 dU(t) &= h_x(t) U(t); \\
 dV(t) &= -V(t) \frac{\partial f}{\partial x}(t) dt - V(t) \sum_{\ell=1}^r \frac{\partial g^\ell}{\partial x}(t) dW^\ell(t) + V(t) \sum_{\ell=1}^r \left(\frac{\partial g^\ell}{\partial x}(t) \right)^2 dt.
 \end{aligned}$$

(2.31) (2.32) & (2.33)

Observe that the SDE's (2.31), (2.32) and (2.33) are equivalent to the SDE's (3.1), (3.2) and (3.3) in Norris respectively. Thus we can see that our delay stochastic differential equations (2.28), (2.15) and (2.30) in fact extend the SDE's (3.1), (3.2) and (3.3) in Norris; they include the case of delay as well as ordinary SDE's. Moreover, we have also proved that $U(t)$ and $V(t)$ are each others inverse, in the case of delay SDE's as well as ordinary SDE's.

Next we consider the delay versions of equations (2.28), (2.15), and (2.30) namely

$$dX_t = f_t dt + \sum_{\ell=1}^r g_t^\ell dW_t^\ell,$$

(2.34)

$$dU_t(\cdot) = h_{x,t}(\cdot)U_t(\cdot) + \int_J h_{\xi,t}(\cdot, \varphi)U_{t+\cdot}(\varphi)d\varphi, \quad (2.35)$$

And

$$\begin{aligned} dV_t &= -V_t \frac{\partial f_t}{\partial x} dt - V_t \sum_{\ell=1}^r \frac{\partial g_t^\ell}{\partial x} dW_t^\ell \\ &\quad - \sum_{\ell=1}^r \int_J V_{t+\cdot}(\vartheta) \frac{\partial g_{t+\cdot}^\ell}{\partial \xi}(\vartheta) d\vartheta dW_t^\ell - V_t \int_J \frac{\partial f_{t+\cdot}}{\partial \xi}(\vartheta) U_{t+\cdot}(\vartheta) d\vartheta V_t dt \\ &\quad + V_t \sum_{\ell=1}^r \left(\frac{\partial g_t^\ell}{\partial x} \right)^2 dt + V_t \sum_{\ell=1}^r \frac{\partial g_t^\ell}{\partial x} \int_J \frac{\partial g_{t+\cdot}^\ell}{\partial \xi} U_{t+\cdot}(\vartheta) d\vartheta V_t dt \\ &\quad + V_t \sum_{\ell=1}^r \int_J \frac{\partial g_{t+\cdot}^\ell}{\partial \xi}(\vartheta) U_{t+\cdot}(\vartheta) d\vartheta V_t \frac{\partial g_t^\ell}{\partial x} dt + V_t \sum_{\ell=1}^r \left(\int_J \frac{\partial g_{t+\cdot}^\ell}{\partial \xi} U_{t+\cdot}(\vartheta) d\vartheta V_t \right)^2 dt \\ &= -V_t \partial f_t \partial x dt - V_t \sum_{\ell=1}^r \frac{\partial g_t^\ell}{\partial x} dW_t^\ell \\ &\quad - \sum_{\ell=1}^r \int_J V_{t+\cdot}(\vartheta) \frac{\partial g_{t+\cdot}^\ell}{\partial \xi}(\vartheta) d\vartheta dW_t^\ell - V_t \int_J \frac{\partial f_{t+\cdot}}{\partial \xi}(\vartheta) U_{t+\cdot}(\vartheta) d\vartheta V_t dt \\ &\quad + V_t \sum_{\ell=1}^r \left(\frac{\partial g_t^\ell}{\partial x} + \int_J \frac{\partial g_{t+\cdot}^\ell}{\partial \xi} U_{t+\cdot}(\vartheta) d\vartheta V_t \right)^2 dt. \end{aligned} \quad (2.36)$$

Then we can see that it is most probable that U_t and V_t are also each others inverse. Now we can proceed to get our integration by parts formula by following steps similar to the ones in chapter 3 of Norris . As above $X(t)$ is a solution to the following delay SDE:

$$dX(t) = f(t, X(t), X_t)dt + \sum_{\ell=1}^r g^\ell(t, X(t), X_t) dW^\ell(t),$$

and the process $D^\nu X(t)$ satisfies the following delay SDE:

$$\begin{aligned}
 d D^v X(t) &= \frac{\partial}{\partial x} f(t, X(t), X_t) D^v X(t) dt + \int_J \frac{\partial}{\partial \xi} f(t, X(t), X_t)(\vartheta) D^v X_t(\vartheta) d\vartheta dt \\
 &\quad + \sum_{\ell=1}^r \frac{\partial}{\partial x} g^\ell(t, X(t), X_t) D^v X(t) dW^\ell(t) \\
 &\quad + \sum_{\ell=1}^r \int_J \frac{\partial}{\partial \xi} g^\ell(t, X(t), X_t)(\vartheta) D^v X_t(\vartheta) d\vartheta dW^\ell(t) \\
 &\quad + \sum_{\ell=1}^r g^\ell(t, X(t), X_t) v^\ell(t, X(t), X_t) dt.
 \end{aligned}
 \tag{2.37}$$

Then we can see that the delay SDE(2.37) in fact extends the SDE (2.8) in chapter two of the work of Norris [10] to include delay as well as ordinary SDE's.

1. All the results which we have established in this work can be extended by replacing the Brownian motion W by another process $Z: [0, a] \times \Omega \rightarrow \mathbf{R}^d$, ($d \in \mathbf{N}$) which is a continuous martingale adapted to $\{\mathcal{F}_t\}_{t \in [0, a]}$ and has independent increments and satisfies with some constant K the inequalities

$$\begin{aligned}
 &|\mathbf{E}[Z(t) - Z(s)] | \mathcal{F}_s| \leq K(t - s) \quad \text{and} \\
 &\mathbf{E} (|Z(t) - Z(s)|^2 | \mathcal{F}_s) \leq K(t - s) \quad \text{for } 0 \leq s \leq t \leq a.
 \end{aligned}$$

2. Observe that the above properties of Z which we have just mentioned are the only properties of W which we have used (in case of Brownian motion) to prove the results which we have obtained in this work. See and .
3. All the lemmas and theorems in this work hold for any delay interval $J' = [-r, 0)$ ($r \geq 0$) in place of $J = [-1, 0)$. See and .

REFERENCES

1. Ahmed, T.A. *Stochastic Functional Differential Equations with Discontinuous Initial Data*, M.Sc. Thesis, University of Khartoum, Khartoum, Sudan, (1983).
2. Bally, V. and Talay, D.: *The law of Euler scheme for stochastic differential equations (I): Convergence rate of the distribution function*. J. Probability Theory and related fields 104, 43-60 (1996)
3. Bally, V., Talay, D.: *The law of Euler scheme for stochastic differential equations (II): Convergence rate of the density*, Monte Carlo Methods and Appl., Vol.2, No.2, pp.93-128 (1996)
4. Elworthy, K.D., *Stochastic Differential Equations on Manifolds*, Cambridge University Press (1982).
5. Friedman, A., *Stochastic Differential Equations and Applications*, Academic press (1975).
6. Ikeda, N. and Watanabe, S, *Stochastic differential equations and Diffusion Processes*, Amsterdam: North-Holland 1981.
7. Kloeden, P.E and Platen, E., *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag Berlin Heidelberg (1992).
8. Kusuoka, S. and Stroock, D., *Applications of the Malliavin Calculus, Part II*. J. Fac. Sci. Univ. Tokyo 32, 1-76 (1985).
9. Mohammed, S.E.A., *Stochastic Functional Differential Equations*, Research Notes in Mathematics; Pitman Books Ltd., London (1984).
10. Norris, J.: *Simplified Malliavin Calculus*, Séminaire de Probabilités XX 1984-1985; Springer Lecture Notes in Mathematics 1204 (1996) 101-130.
11. Nualart, D.: *Malliavin Calculus and related topics*, Springer-Verlag 1995.
12. Øksendal, B., *Stochastic Differential Equations, Introduction with Applications*, Springer-Verlag Berlin Heidelberg (1985).

13. Talay, D. *Efficient Numerical Schemes for the Approximation of Expectations of functionals of the solution of a S.D.E., and applications*, In: Korezlioglu, H., Mazziotto, G. and Szpirglas, J. (editors): Filtering and control of random processes (Lect. Notes in Control and Information sciences, Vol. 61, Proceedings of E.N.S.T.-C.N.E.T. Coll., Paris, 1983) pp. 294–313, Berlin: Springer 1984.
14. Talay, D. and Tubaro, L.: *Expansion of the Global Error for Numerical Schemes Solving Stochastic differential equations*. Stochastic Anal. Appl. 8, 94-120 (1990).
15. Tagelsir A Ahmed, *Some approximation results for solution to delay stochastic differential equations*. PhD Thesis, University of Antwerp, Antwerp, Belgium, Congress Library, 2003.
16. Tagelsir A Ahmed and Van Casteren, J.A., *Densities of distributions of solutions to delay stochastic differential equations with discontinuous initial data(Part I)*. International Journal of Contemporary Applied Sciences Vol. 3, Issue 1, January 2016.