

## Applications of the Integration By Parts Formula II

Tagelsir A. Ahmed<sup>1</sup>, Van Casteren, Jan A.

<sup>1</sup>Department Of Pure Mathematics, Faculty Of Mathematical Science, University Of Khartoum, B.O.Box 321, Khartoum, Sudan

<sup>2</sup>Department Of Mathematics And Computer Science, Universityof Antwerp (UA), Middelheimlaan 1, 2020 Antwerp

**Abstract:** We have established an integration by parts formula involving higher order Malliavin derivatives. This integration by parts formula can be used to extend the formulas in the work by Bally and Talay to include delay SDE's as well as ordinary SDE's.

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

### I. Introduction

In Chapter 1 of the Ph.D. thesis of Ahmed [15] 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 [1], [9] and the web cite [www.sfde.math.siu.edu](http://www.sfde.math.siu.edu). See the delay SDE (1.1) in the present work. In [18] we have established integration by parts formula involving Malliavin derivatives of 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 [2] and [3] 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 [10]. In a sequel paper we will show that the distribution of the solution process has smooth densities. Moreover we will establish integration by part formula involving Malliavin derivatives of higher order.

#### 1.2 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- $\sigma$  algebras of  $\mathcal{F}$ , each of which contains all null subsets of  $\Omega$ ;  $\mathbb{N}$  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  $\lambda$  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- $\sigma$  algebra of  $\mathcal{F}$  containing all subsets of measure zero in  $\mathcal{F}$ , then  $L^2(\Omega, \mathcal{A}, \mathbb{P}; G)$  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:  $X_t: \Omega \rightarrow \mathcal{L}(J, \mathbb{R}^d)$  by  $X_t(\omega)(s) = X(t+s, \omega)$  for all  $s \in J$  and almost all  $\omega$ . 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  $L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d)$ ,  $\theta$  belong to  $L^2(J \times \Omega, \mathcal{H}(J) \otimes \mathcal{F}_0, \lambda \otimes \mathbb{P}; \mathbb{R}^d)$ , and for  $\ell = 1, 2, \dots, r$  let  $f, g^\ell$  be functions from  $[0, T] \times \Omega \times \mathbb{R}^d \times \mathcal{L}(J, \mathbb{R}^d)$  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

- (i)  $X$  is  $\mathcal{B}([0, T]) \otimes \mathcal{F}$ - $\mathcal{B}(\mathbb{R}^d)$  measurable;
- (ii) 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;
- (iii)  $X \in L^2([-1, T] \times \Omega, \mathcal{H} \times \mathcal{F}, \lambda \times \mathbb{P}; \mathbb{R}^d)$ ,
- (IV)  $X$  satisfies the delay SDE ((1.1.1)).

The following conditions are sufficient for the existence of a unique solution to (1.1) (see [1] and [15]).

- (i)  $V \in \mathcal{L}^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{R}^d)$ .
- (ii)  $\theta \in \mathcal{L}^2(J \times \Omega, \mathcal{H} \otimes \mathcal{F}_0, \lambda \otimes \mathbb{P}, \mathbb{R}^d)$ .
- (iii)  $f, g^\ell: [0, T] \times \Omega \times \mathbb{R}^d \times \mathcal{L}^2(J, \mathbb{R}^d) \rightarrow \mathbb{R}^d$  are such that
  - (a)  $f$  and  $g^\ell$  are  $\mathcal{B}([0, T]) \otimes \mathcal{F} \otimes \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathcal{L}^2(J, \mathbb{R}^d))$ - $\mathcal{B}(\mathbb{R}^d)$  measurable.
  - (b) For each  $t \in [0, T]$ , the stochastic variables  $f(t, \cdot, \cdot, \cdot)$  and  $g^\ell(t, \cdot, \cdot, \cdot)$  are  $\mathcal{F}_t \otimes \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathcal{L}^2(J, \mathbb{R}^d))$ - $\mathcal{B}(\mathbb{R}^d)$  measurable.
  - (c) There exists a constant  $K$  and a function  $\zeta \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$  such that
 
$$|f(t, \omega, s, h)| + \sum_{\ell=1}^r |g^\ell(t, \omega, s, h)| \leq K(|s| + \|h\| + |\zeta(\omega)|) \quad (1.2) \quad (1.2)$$
 for almost all  $\omega$  and for all  $t \in [0, T]$ ;  $s \in \mathbb{R}^d$  and  $h$  belongs to  $\mathcal{L}^2(J, \mathbb{R}^d)$ .
  - (d) There exists a constant  $K'$  such that, for almost all  $\omega$ ,
 
$$|f(t, \omega, s, h_1) - f(t, \omega, u, h_2)| + \sum_{\ell=1}^r |g^\ell(t, \omega, s, h_1) - g^\ell(t, \omega, u, h_2)| \leq K'(|s - u| + \|h_1 - h_2\|) \quad (1.3)$$
 for all  $t \in [0, T]$ ; for all  $s, u \in \mathbb{R}^d$ , and for all  $h_1, h_2 \in \mathcal{L}^2(J, \mathbb{R}^d)$ .

## II. Integration By Parts Formula

In the beginning of this section we recall the following eight basic numbered equations and definitions, See(16) and(17) . For  $(X(0), X_0) = (x, \xi) \in \mathbb{R}^d \times \mathcal{L}^2(J, \mathbb{R}^d)$ , 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.

**I.Definition:** Let  $F((W(s))_{0 \leq s \leq T})$  be a functional of  $r$ -dimensional Brownian motion, and let  $v(t) = (v^1(t), \dots, v^r(t))^* = \begin{pmatrix} v^1(t) \\ \vdots \\ v^r(t) \end{pmatrix}$  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  $\mathcal{L}^2(J, \mathbb{R})$ -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:**

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

where  $\vartheta$  belongs to  $J$ . If  $t + \vartheta$  belongs to  $J$  we replace  $t + \vartheta$  with 0 in (2.2). If  $\vartheta = 0$  we obtain the delay stochastic differential equation for the process  $D^v X(t)$ :

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

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  $(1 \leq j, k \leq d)$

$$\xi_k \rightarrow X_j^{x,(\xi_1, \dots, \xi_{k-1}, \xi_k, \xi_{k+1}, \dots, \xi_d)}(t), \tag{2.4}$$

which is a mapping from  $L^2(J, \mathbb{R})$  to  $\mathbb{R}$ , and where each variable  $\xi_\ell, \ell \neq k$ , is a fixed function in  $L^2(J, \mathbb{R})$ . The derivative of the function in (2.4) 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

$$\begin{aligned}
 \frac{\partial X_j^{x,\xi}(t)}{\partial \xi_k}(\eta) &= \lim_{h \rightarrow 0} \frac{X_j^{x,(\xi_1, \dots, \xi_{k-1}, \xi_k+h\eta, \xi_{k+1}, \dots, \xi_d)}(t) - X_j^{x,(\xi_1, \dots, \xi_{k-1}, \xi_k, \xi_{k+1}, \dots, \xi_d)}(t)}{h} \\
 &= \int_J \eta(\varphi) \frac{\partial X_j^{x,\xi}(t)}{\partial \xi_k}(\varphi) d\varphi, \quad \eta \in L^2(J, \mathbb{R}).
 \end{aligned} \tag{2.5}$$

After giving a brief introduction to our work, we are now ready to continue the work that we have started in (16).

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*:

$$h_x(t) = \frac{\partial f}{\partial x}(t) dt + \sum_{\ell=1}^r \frac{\partial g^\ell}{\partial x}(t) dW^\ell(t); \tag{2.6}$$

$$h_\xi(t) = \frac{\partial f}{\partial \xi}(t) dt + \sum_{\ell=1}^r \frac{\partial g^\ell}{\partial \xi}(t) dW^\ell(t) \tag{2.7}$$

$$h_\xi(t, \vartheta) = \frac{\partial f}{\partial \xi}(t, \vartheta) dt + \sum_{\ell=1}^r \frac{\partial g^\ell}{\partial \xi}(t, \vartheta) dW^\ell(t) \tag{2.8}$$

**Application of the Integration by Parts Formula:**

Relevant SDE's are  $(v(t))$  is a  $r \times d$  matrix-valued adapted process:  $d$  columns,  $r$  rows)

$$dD^v X(t) = h_x(t) D^v X(t) + \int_J h_\xi(t, \vartheta) D^v X_t(\vartheta) d\vartheta + \sum_{\ell=1}^r g^\ell(t) v^\ell(t)^\tau dt; \tag{2.9}$$

$$\begin{aligned}
 dV^v(t) &= -V^v(t) h_x(t) - V^v(t) \int_J h_\xi(t, \vartheta) D^v X_t(\vartheta) d\vartheta (D^v X(t))^{-1} \\
 &\quad + V^v(t) \sum_{\ell=1}^r \left( \frac{\partial g^\ell(t)}{\partial x} + \int_J \frac{\partial g^\ell(t, \vartheta)}{\partial \xi} D^v X_t(\vartheta) d\vartheta (D^v X(t))^{-1} \right)^2 dt; \tag{2.10}
 \end{aligned}$$

$$dU^v(t) = h_x(t) U^v(t) + \int_J h_\xi(t, \vartheta) D^v X_t(\vartheta) d\vartheta (D^v X(t))^{-1} U^v(t). \tag{2.11}$$

Next we come to higher order Malliavin derivatives. As above the first order Malliavin derivative  $D^v X(t)$ , which, by the way, is linear in the process  $v(t)$ , satisfies the following stochastic differential equation.

$$dD^v X(t) = h_x(t)D^v X(t) + \int_J h_\xi(t, \vartheta)D^v X_t(\vartheta) d\vartheta + \sum_{\ell=1}^r g^\ell(t)v^\ell(t)^\tau dt. \quad (2.12)$$

In this equation we replace  $W(t)$  with  $W^{v_1, h}(t) := W(t) + \int_0^t v_1(s)h(s)ds$ , where  $v_1(t)$  is a column vector of an  $r \times d$ -matrix:  $v_1(t) = \begin{pmatrix} v_1^1(t) \\ \vdots \\ v_1^r(t) \end{pmatrix}$ . Here each  $v_1^\ell(t)$ ,  $1 \leq \ell \leq r$  is row vector of length  $d$ . The corresponding

delay stochastic differential equation for the process  $D^{v, v_1, h} X(t)$  reads as follows

$$\begin{aligned} & dD^{v, v_1, h} X(t) \\ &= h_x(t)D^{v, v_1, h} X(t) + \int_J h_\xi^{v_1, h}(t, \vartheta)D^{v, v_1, h} X_t(\vartheta) d\vartheta + \sum_{\ell=1}^r g^\ell(t)v^\ell(t)^\tau dt \\ &= \frac{\partial}{\partial x} f(t, X(t), X_t)D^{v, v_1, h} X(t)dt + \int_J \frac{\partial f}{\partial \xi}(t, X(t), X_t, \vartheta)D^{v, v_1, h} X_t(\vartheta) d\vartheta dt \\ &\quad + \sum_{\ell=1}^r \int_J \frac{\partial}{\partial \xi} g^\ell(t, X(t), X_t, \vartheta)D^{v, v_1, h} X_t(\vartheta) d\vartheta dW^\ell(t) \\ &+ \sum_{\ell=1}^r \int_J \frac{\partial}{\partial \xi} g^\ell(t, X(t), X_t, \vartheta)D^{v, v_1, h} X_t(\vartheta) d\vartheta v_1^\ell(t)^\tau h(t)dt + \sum_{\ell=1}^r g^\ell(t)v^\ell(t)^\tau dt. \end{aligned} \quad (2.13)$$

Here we used the fact that  $h_\xi^{v_1, h}(t, \vartheta)$  is given by

$$h_\xi^{v_1, h}(t, \vartheta) = \frac{\partial f}{\partial \xi}(t, X(t), X_t, \vartheta) dt + \sum_{\ell=1}^r \frac{\partial}{\partial \xi} g^\ell(t, X(t), X_t, \vartheta) dW^{\ell, v_1, h}(t).$$

Then we differentiate equation (2.13) with respect to  $h$  to obtain a delay SDE for the process  $D^{v_1} D^v X(t) = D^{v_1}(D^v X(t))$ :

$$\begin{aligned} & dD^{v_1} D^v X(t) \\ &= \frac{\partial}{\partial x} f(t, X(t), X_t)D^{v_1} D^v X(t)dt + \int_J \frac{\partial f}{\partial \xi}(t, X(t), X_t, \vartheta)D^{v_1} D^v X_t(\vartheta) d\vartheta dt \\ &\quad + \sum_{\ell=1}^r \int_J \frac{\partial}{\partial \xi} g^\ell(t, X(t), X_t, \vartheta)D^{v_1} D^v X_t(\vartheta) d\vartheta dW^\ell(t) \\ &\quad + \sum_{\ell=1}^r \int_J \frac{\partial}{\partial \xi} g^\ell(t, X(t), X_t, \vartheta)D^{v_1} D^v X_t(\vartheta) d\vartheta v_1^\ell(t)^\tau dt \\ &= h_x(t)D^{v_1} D^v X(t) + \int_J h_\xi(t, \vartheta)D^{v_1} D^v X_t(\vartheta) d\vartheta \\ &+ \sum_{\ell=1}^r \int_J \frac{\partial}{\partial \xi} g^\ell(t, X(t), X_t, \vartheta)D^{v_1} D^v X_t(\vartheta) d\vartheta v_1^\ell(t)^\tau dt. \end{aligned} \quad (2.14)$$

$$\begin{aligned}
 & dD^{v_1}D^v X(t) \\
 = & \left(\frac{\partial}{\partial x}\right)^2 f(t)D^{v_1}X(t)D^v X(t)dt + \int_J \frac{\partial^2}{\partial \xi \partial x} f(t, \vartheta)D^{v_1}X_t(\vartheta)d\vartheta D^v X(t) dt \\
 & + \frac{\partial}{\partial x} f(t)D^{v_1}D^v X(t)dt + \int_J \frac{\partial^2}{\partial \xi \partial x} f(t, \vartheta)D^{v_1}X(t) D^v X_t(\vartheta)d\vartheta dt \\
 & + \int_J \int_J \frac{\partial^2}{\partial \xi^2} f(t, \vartheta, \vartheta_1)D^{v_1}X_t(\vartheta_1)D^v X_t(\vartheta)d\vartheta d\vartheta_1 dt \\
 & + \int_J \frac{\partial}{\partial \xi} f(t, \vartheta)D^{v_1}D^v X_t(\vartheta)d\vartheta dt \\
 & + \sum_{\ell=1}^r \left(\frac{\partial}{\partial x}\right)^2 g^\ell(t)D^{v_1}X(t)D^v X(t)dW^\ell(t) \\
 & + \sum_{\ell=1}^r \int_J \frac{\partial^2}{\partial \xi \partial x} g^\ell(t, \vartheta)D^{v_1}X_t(\vartheta)d\vartheta D^v X(t) dW^\ell(t) \\
 & + \sum_{\ell=1}^r \frac{\partial}{\partial x} g^\ell(t)D^{v_1}D^v X(t)dW^\ell(t) \\
 & + \sum_{\ell=1}^r \int_J \frac{\partial^2}{\partial \xi \partial x} g^\ell(t, \vartheta)D^{v_1}X(t) D^v X_t(\vartheta)d\vartheta dW^\ell(t) \\
 & + \sum_{\ell=1}^r \int_J \int_J \frac{\partial^2}{\partial \xi^2} g^\ell(t, \vartheta, \vartheta_1)D^{v_1}X_t(\vartheta_1)D^v X_t(\vartheta)d\vartheta d\vartheta_1 dW^\ell(t) \\
 & + \sum_{\ell=1}^r \int_J \frac{\partial}{\partial \xi} g^\ell(t, \vartheta)D^{v_1}D^v X_t(\vartheta)d\vartheta dW^\ell(t) \\
 & + \sum_{\ell=1}^r \frac{\partial}{\partial x} g^\ell(t)D^{v_1}X(t)v^\ell(t)^\tau dt + \sum_{\ell=1}^r \int_J \frac{\partial}{\partial \xi} g^\ell(t, \vartheta)d\vartheta v^\ell(t)^\tau dt \\
 & + \sum_{\ell=1}^r g^\ell(t) \frac{\partial}{\partial x} v^\ell(t)^\tau D^{v_1}X(t) dt + \sum_{\ell=1}^r g^\ell(t) \int_J \frac{\partial}{\partial \xi} v^\ell(t, \vartheta)D^{v_1}X_t(\vartheta)d\vartheta dt \\
 & + \sum_{\ell=1}^r \frac{\partial}{\partial x} g^\ell(t)D^v X(t)v_1^\ell(t)^\tau dt. \tag{2.15}
 \end{aligned}$$

We need to mention that the integration by parts formula in [18] can be used to extend the formulas in [2] to include delay SDE's as well as ordinary SDE's. For this purpose we need to use the integration by parts formula in [18] where  $v_1 = v_2 = \dots = v_{n+1}$ . See the proof of the following Theorem 3. We also need the following definition.

**2 Definition:** For a stochastic variable  $Y$  attaining its values in  $\mathbb{R}^m \times L^2(J, \mathbb{R}^n)$  the space  $\mathcal{D}(Y)$  consists of those functions  $F: \mathbb{R}^m \times L^2(J, \mathbb{R}^n) \rightarrow \mathbb{R}$  such that for some open subset  $O$  of  $\mathbb{R}^m \times L^2(J, \mathbb{R}^n)$

- I. The variable  $Y$  belongs to  $O$  almost surely;
- II. The restriction  $F|_O$  is  $C^\infty$ ;
- III. For every  $\alpha \in \mathbb{N}^m$  and  $\beta \in \mathbb{N}^n$  the Frechet derivative  $D_x^\alpha D_\xi^\beta F(Y)$  belongs to

$$L^2([0, T] \times \Omega, \mathcal{B} \otimes \mathcal{F}, \lambda \times \mathbb{P}).$$

**3 Theorem:** Let  $W(t)$  be  $r$ -dimensional Brownian motion and let  $X(t)$  be the solution of

$$dX(t) = f(t, X(t), X_t)dt + \sum_{\ell=1}^r g^\ell(t, X(t), X_t)dW^\ell(t), \tag{2.16}$$

with  $(X(0), X_0) = (x, \xi)$ , where  $f, g^1, g^2, \dots, g^r$  are maps from  $[0, T] \times \mathbb{R}^d \times L^2(J, \mathbb{R}^d)$  to  $\mathbb{R}^d$  and  $J = [-1, 0)$  and  $T$  is a positive real number. Let  $v$  and  $v^\ell: [0, T] \times \mathbb{R}^d \times L^2(J, \mathbb{R}^d) \rightarrow \mathbb{R}^d$ , ( $\ell = 1, 2, \dots, r$ ,  $j = 1, 2, \dots, n + 1$ ) be  $C^\infty$ -mappings, with all their derivatives of polynomial growth. Then the linear delay SDE

$$\begin{aligned}
 d D^v X(t) &= \frac{\partial}{\partial x} f(t, X(t), X_t) D^v X(t) dt + \int_J \frac{\partial}{\partial \xi} f(t, X(t), X_t)(\vartheta) D^v X_t(\vartheta) d\vartheta dt \\
 &\quad + \sum_{\ell=1}^r \frac{\partial}{\partial x} g^\ell(t, X(t), X_t) D^v X(t) dW^\ell(t) \\
 &\quad + \sum_{\ell=1}^r \int_J \frac{\partial}{\partial \xi} g^\ell(t, X(t), X_t)(\vartheta) D^v X_t(\vartheta) d\vartheta dW^\ell(t) \\
 &\quad + \sum_{\ell=1}^r g^\ell(t, X(t), X_t) v^\ell(t, X(t), X_t) dt. \tag{2.17}
 \end{aligned}$$

with  $(D^v X(0), D^v X_0) = 0 \in (\mathbb{R}^d \times \mathbb{R}^d \times L^2(J, \mathbb{R}^d \times \mathbb{R}^d))$  has a unique solution with  $\sup_{s \leq t} |D^v X(s)|$  and  $\sup_{s \leq t} |D^{v_j} X(s)|$

belonging to  $L^p(\Omega, \mathcal{F}_t, P)$ , and

$$\sup_{s \leq t} |D^v X_s| \quad \text{and} \quad \sup_{s \leq t} |D^{v_j} X_s|$$

belonging to  $L^p(\Omega \times J, \mathcal{F}_t \otimes \mathcal{B}, \mathbb{P} \times \lambda)$  for all  $t \geq 0$  and for all  $p < \infty$  and for all  $j = 1, 2, \dots, n + 1$ . The notation  $D^v X_t$  is also employed for the delay process  $\vartheta \mapsto D^v X(t + \vartheta)$ . Furthermore, for any function  $\Phi: U_0 \rightarrow \mathbb{R}$ , where  $U_0$  is an open subset of  $[0, T] \times \mathbb{R}^d \times L^2(J, \mathbb{R}^d)$  with  $(t, X(t), X_t) \in U_0$  a.s. such that  $\Phi$  is differentiable, and  $D\Phi(t, X(t), X_t)$  and  $\Phi(t, X(t), X_t)$  belonging to  $L^2(\Omega, \mathcal{F}_t, P)$ ,

$$\begin{aligned}
 &\mathbb{E} \left[ \left\{ \frac{\partial}{\partial x} \Phi(t, X(t), X_t) + \int_J \frac{\partial}{\partial \xi} \Phi(t, X(t), X_t)(\vartheta) D^{v_{n+1}} X_t(\vartheta) (D^{v_{n+1}} X(t))^{-1} d\vartheta \right\} \right. \\
 &\quad \left. \int D^{v_{n+1}} X(t) \Psi(D^{v_n} X(t)) G(y^{(n)}(t)) \right] \\
 &= \mathbb{E} [\Phi(t, X(t), X_t) A^{(v_n, \dots, v_1)} G(y^{(n+1)}(t))] \tag{2.18}
 \end{aligned}$$

where

$$\begin{aligned}
 A^{(v_n, \dots, v_1)} G(y^{(n+1)}(t)) &= \Psi(D^{v_n} X(t)) G(y^{(n)}(t)) \sum_{\ell=1}^r \int_0^t v_{n+1}^\ell(s, X(s), X_s)^\tau dW^\ell(s) \\
 &\quad - \Psi(D^{v_n} X(t)) D G(y^{(n)}(t)) D^{v_{n+1}}(y^{(n)}(t)) \\
 &\quad - D \Psi(D^{v_n} X(t)) (D^{v_{n+1}} D^{v_n} X(t)) G(y^{(n)}(t)); \tag{2.19}
 \end{aligned}$$

$$y^{(n)}(t) = (y^{(0)}(t), D^{v_1} y^{(0)}(t), D^{v_2} D^{v_1} y^{(0)}(t), \dots, D^{v_n} D^{v_{n-1}} \dots D^{v_1} y^{(0)}(t)); \tag{2.20}$$

$$y^{(0)}(t) = (X(t), D^v X(t), R^v(t)); \tag{2.21}$$

$$R^v(t) = \sum_{\ell=1}^r \int_0^t v^\ell(s)^\tau dW^\ell(s), \tag{2.22}$$

and  $G \in \mathcal{D}[y^{(n)}(t)]$  where a choice for  $v^\ell$  could be

$$v^\ell(s, X(s), X_s) = (V(s) g^\ell(s, X(s), X_s))^\tau. \tag{2.23}$$

As in (2.12) in [18] a specialization of ([E: int.formula2]) to the partial derivatives  $\frac{\partial}{\partial x_k}$  and  $\frac{\partial}{\partial \xi_k}$  yields:

$$\begin{aligned}
 &\mathbb{E} \left[ \left\{ \frac{\partial}{\partial x_k} \Phi(t, X(t), X_t) + \int_J \frac{\partial}{\partial \xi_k} \Phi(t, X(t), X_t)(\vartheta) D^{v_{n+1}} X_t(\vartheta) (D^{v_{n+1}} X(t))^{-1} d\vartheta \right\} \right. \\
 &\quad \left. D^{v_{n+1}} X(t) \Psi(D^{v_n} X(t)) G(y^{(n)}(t)) \right] \\
 &= \mathbb{E} [\Phi(t, X(t), X_t) A_k^{(v_n, \dots, v_1)} G(y^{(n+1)}(t))] \tag{2.24}
 \end{aligned}$$

where  $A_k^{(v_n, \dots, v_1)} G(y^{(n+1)}(t))$  is the  $k$ -th component of  $A^{(v_n, \dots, v_1)} G(y^{(n+1)}(t))$ , which is row vector of length  $d$ .

We shall prove this theorem by induction on the non-negative integers  $n = 0, 1, 2, \dots$

**Proof:** We first observe that the formula in (2.10) in [18] with  $y^{(n)}(t)$  replacing the pair  $(X(t), X_t)$ , and  $v_{n+1}$  replacing the vector  $v$ , implies

$$\begin{aligned}
 &\mathbb{E} \left[ D^{v_{n+1}} \left( \Phi(t, X(t), X_t) \Psi(D^{v_n} X(t)) G(y^{(n)}(t)) \right) \right] \\
 &= \sum_{\ell=1}^r \mathbb{E} \left[ \Phi(t, X(t), X_t) \Psi(D^v X(t)) G(y^{(n)}(t)) \int_0^t v_{n+1}^\ell(s)^\tau dW^\ell(s) \right] \tag{2.25}
 \end{aligned}$$

which in turn, by Leibniz rule, implies

$$\begin{aligned} & \mathbb{E}[D^{v_{n+1}}\Phi(t, X(t), X_t)\Psi(D^{v_n}X(t))G(y^{(n)}(t))] \\ & + \mathbb{E}[\Phi(t, X(t), X_t)D^{v_{n+1}}\Psi(D^{v_n}X(t))G(y^{(n)}(t))] \\ & + \mathbb{E}[\Phi(t, X(t), X_t)\Psi(D^{v_n}X(t))D^{v_{n+1}}G(y^{(n)}(t))] \\ = & \mathbb{E}\left[\Phi(t, X(t), X_t)\Psi(D^{v_n}X(t))G(y^{(n)}(t))\sum_{\ell=1}^r\int_0^t v_{n+1}^\ell(s)^\tau dW^\ell(s)\right] \end{aligned} \quad (2.26)$$

Again, by applying Leibniz rule to the Malliavin derivative  $D^{v_{n+1}}\Phi(t, X(t), X_t)$  we see that the first term of the right-hand side of equality [2.26] coincides with the left-hand side of the equality in [2.18]. Consequently [2.18] follows from [2.26] here we applied the chain rule for the Malliavin derivative  $D^{v_{n+1}}$  a number of times.

Then by choosing  $G = 1$  we can formulate the above equation as follows:

$$\begin{aligned} & \mathbb{E}\left[\frac{\partial}{\partial x}\Phi(t, X(t), X_t)D^{v_{n+1}}X(t)\Psi(D^{v_n}X(t))\right] \\ & + \mathbb{E}\left[\int_J\frac{\partial}{\partial \xi}\Phi(t, X(t), X_t)(\vartheta)D^{v_{n+1}}X_t(\vartheta)\Psi(D^{v_n}X(t))d\vartheta\right] \\ & + \mathbb{E}[\Phi(t, X(t), X_t)D^{v_{n+1}}\Psi(D^{v_n}X(t))] \\ = & \mathbb{E}\left[\Phi(t, X(t), X_t)\Psi(D^{v_n}X(t))\int_0^t v_{n+1}^\ell(s)^\tau dW^\ell(s)\right] \end{aligned} \quad (2.27)$$

Hence we can rewrite equation ([int.formula 3]) in the following form:

$$\begin{aligned} & \mathbb{E}\left[\left\{\frac{\partial}{\partial x}\Phi(t, X(t), X_t) + \int_J\frac{\partial}{\partial \xi}\Phi(t, X(t), X_t)(\vartheta)D^{v_{n+1}}X_t(\vartheta)(D^{v_{n+1}}X(t))^{-1}d\vartheta\right\}\right. \\ & \quad \left.\int_J D^{v_{n+1}}X(t)\Psi(D^{v_n}X(t))\right] \\ & = \mathbb{E}[\Phi(t, X(t), X_t)\Psi(D^{v_n}X(t))R^{v_{n+1}}(t)] \\ & + \mathbb{E}[\Phi(t, X(t), X_t)D\Psi(D^{v_n}X(t))D^{v_{n+1}}D^{v_n}X(t)] \\ = & \mathbb{E}[A^{(v_{n+1}, 1)}1(y^{(n)}(t))]. \end{aligned} \quad (2.28)$$

Hence the theorem is valid for all non-negative integers  $n = 0, 1, 2, \dots$

**4 Theorem:** Let  $W(t)$  be  $r$ -dimensional Brownian motion and let  $X(t)$  be the solution of equation (2.9) in [18], to wit:

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

with  $(X(0), X_0) = (x, \xi)$ , where  $f, g^1, g^2, \dots, g^r$  are maps  $[0, T] \times \mathbb{R}^d \times L^2(J, \mathbb{R}^d) \rightarrow \mathbb{R}^d$  and  $J = [-1, 0)$  and  $T$  is a positive real number. Let  $v^\ell: [0, T] \times \mathbb{R}^d \times L^2(J, \mathbb{R}^d) \rightarrow \mathbb{R}^d$ ,  $(\ell = 1, 2, \dots, r)$  be  $C^\infty$ , with all derivatives of polynomial growth. Suppose that for some  $t \in [0, T]$  the inverse  $(D^v X(t))^{-1}$  of the Malliavin derivative  $D^v X(t)$  belongs to  $L^p(\Omega, \mathcal{F}_t, \mathbb{P})$  for all  $1 \leq p < \infty$ . Then the law of  $X(t)$  has a  $C^\infty$ -density.

**proof As in Theorem 3.2 page 121 of Norris' paper [10] the present Theorem 3 implies:**

$$\begin{aligned} & \mathbb{E}[D_{k_1}^{v_1} \dots D_{k_n}^{v_n} \Phi(X(t))D^{v_{n+1}}X(t)\Psi(D^{v_n}X(t))G(y^{(n)}(t))] \\ = & \mathbb{E}[\Phi(X(t))A_{k_n}^{v_n, \dots, v_1} \circ \dots \circ A_{k_1}^{v_1} G(y^{(n+1)}(t))], \end{aligned} \quad (2.29)$$

**where  $\Phi$  is a bounded  $C^\infty$ -function. If  $v_{n+1} = v_n = \dots = v_1 = v$ , then from (2.29) we get**

$$\begin{aligned} & \mathbb{E}[D_{k_1}^v \dots D_{k_n}^v \Phi(X(t))G(y^{(n)}(t))] \\ = & \mathbb{E}\left[\Phi(X(t))A_{k_n}^{v, \dots, v} \circ \dots \circ A_{k_1}^v G(y^{(n+1)}(t))\right]. \end{aligned} \quad (2.30)$$

**Hence, with  $C(k_1, \dots, k_n; G) = \mathbb{E}\left[A_{k_n}^{(v, \dots, v)} \circ \dots \circ A_{k_1}^{(v)} G(y^{(n+1)}(t))\right]$  we obtain**

$$|\mathbb{E}[D_{k_1}^v \dots D_{k_n}^v \Phi(X(t))G(y^{(n)}(t))]| \leq C(k_1, \dots, k_n; G)\|\Phi\|_\infty.$$

### III. Remarks

- 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  $|Z(t) - Z(s)| \leq K\sqrt{t-s}$ .

$K(t - s)$  and  $\mathbf{E}(|Z(t) - Z(s)|^2 \mathcal{F}_s) \leq K(t - s)$  for  $0 \leq s \leq t \leq a$ . 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 [1], [15], [16], [17], [18], and [19].

2. 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 [1], [15], [16], [17], [18], and [19].

### 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., Mazzitot, 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.
- [17]. Tagelsir A Ahmed and Van Casteren, J.A., Densities of distributions of solutions to delay stochastic differential equations with discontinuous initial data (Part II). *International Journal of Engineering Sciences & research technology* Vol. 5, Issue 3, March 2016.
- [18]. Tagelsir A Ahmed and Van Casteren, J.A., Integration by Parts formula involving Malliavin derivatives and solutions to delay SDE's, *International Journal of Engineering Sciences & research technology* Vol. 5, Issue 4, April 2016.
- [19]. Tagelsir A Ahmed and Van Casteren, J.A., Applications of the Integration by Parts Formula I, *International Journal of Engineering Researches and Management Studies* Vol. 3, Issue 7, July 2016.