

***The Weak Euler Scheme for Stochastic
Differential Delay Equations***

Buckwar, Evelyn and Kuske, Rachel and Mohammed,
Salah-Eldin and Shardlow, Tony

2006

MIMS EPrint: **2006.2**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

Reports available from: <http://eprints.maths.manchester.ac.uk/>

And by contacting: The MIMS Secretary
School of Mathematics
The University of Manchester
Manchester, M13 9PL, UK

ISSN 1749-9097

THE WEAK EULER SCHEME FOR STOCHASTIC DIFFERENTIAL DELAY EQUATIONS

EVELYN BUCKWAR, RACHEL KUSKE, SALAH-ELDIN MOHAMMED, AND TONY SHARDLOW

ABSTRACT. We develop a weak numerical Euler scheme for non-linear stochastic delay differential equations (SDDEs) driven by multidimensional Brownian motion. The weak Euler scheme has order of convergence 1, as in the case of stochastic ordinary differential equations (SODEs) (i.e., without delay). The result holds for SDDEs with multiple finite fixed delays in the drift and diffusion terms. Although the set-up is *non-anticipating*, our approach uses the Malliavin calculus and the *anticipating* stochastic analysis techniques of Nualart and Pardoux.

1. INTRODUCTION

Stochastic differential delay equations serve as models of noisy physical processes whose time evolution depends on their past history. In physics, laser dynamics with delayed feedback is often investigated [?, ?], as well as the dynamics of noisy bi-stable systems with delay [?, ?]. Several authors have studied stochastic oscillator ensembles with delayed coupling [?, ?]. These can be interpreted as mean-field models of coupled biological oscillators, such as groups of chorusing crickets, flashing fireflies and cardiac pacemaker cells. In biophysics, stochastic delay equations are used to model delayed visual feedback systems [?, ?] or human postural sway [?].

Since the model equations are generally non-linear and do not allow for explicit solutions, there is a clear need for numerical approximation methods of solution. Until recently, emphasis on the numerical analysis of stochastic differential delay equations has been on *strong* convergence of the numerical schemes. Early investigations in this direction were initiated by Ahmed, Elsanousi and Mohammed [?] and C. and M. Tudor [Tud.1,Tud.2,Tud.3]. Recently, this topic has gained more attention. See for example [?, ?, ?, ?, ?].

Specific approximation methods studied include the Euler-Maruyama scheme and the Θ -method, with order of strong convergence $1/2$, and the Milstein-method, with strong order 1.

The main motivation for considering weak approximations is the computation of the expectation of functionals of solutions of stochastic differential equations. This problem arises, for example, in the fair pricing of options in mathematical finance. Weak approximations are also used in the computation of Lyapunov exponents of systems described by stochastic functional differential equations, as has been suggested by Milstein and Tretyakov in [?]. Lyapunov exponents for stochastic functional differential equations were studied by Mohammed and Scheutzow [?, ?].

Weak approximations for SODE's (without memory) are well-developed. To mention only a few references we quote Bally and Talay [B.T], Kloeden and Platen [K.P] and Kohatsu-Higa [K-H].

The question of weak convergence of numerical approximation methods for stochastic functional differential equations (SFDEs) was recently discussed by E. Buckwar and T. Shardlow in [?]. The result in [?] establishes weak convergence of order 1 for the Euler scheme applied to semi-linear SFDEs of the

2000 *Mathematics Subject Classification*. Primary 34K50, 60H07, 60H35. Secondary 60C30, 60H10, 37H10, 34K28.

Key words and phrases. Stochastic delay differential equation, weak Euler scheme, Malliavin Calculus, weak derivatives, Skorohod integral.

Buckwar's research is supported in part by DFG and the Canadian PIMS.

Kuske's research is supported by PIMS.

Mohammed's research is supported in part by NSF Grant DMS-0203368 and by PIMS. This author is grateful for financial support from the Centre of Mathematics for Applications (CMA), Oslo, Norway and Humboldt University, Berlin, Germany.

Shardlow's research is supported in part by PIMS and EPSRC grant GR/R78725/01.

form:

$$(1.1) \quad x(t) = \begin{cases} v + \int_{\sigma}^t \int_{-\tau}^0 x(u+s) d\mu(s) du + \int_{\sigma}^t f(x(u)) du + \int_{\sigma}^t g(x(u)) dW(u), & t \geq \sigma, \\ \eta(t-\sigma), & \text{a.e. } t \in [\sigma-\tau, \sigma]. \end{cases}$$

In the above SFDE, the memory term is *linear*, driven by a smooth measure μ , and there is *no delay* in the diffusion coefficient. The initial condition (v, η) belongs to the Delfour-Mitter Hilbert space $M_2 := \mathbf{R}^d \times L^2([-\tau, 0], \mathbf{R}^d)$, where τ is the length of the memory in the SFDE. The coefficients $f, g : \mathbf{R}^d \rightarrow \mathbf{R}^d$ satisfy appropriate regularity and linear growth conditions. The driving noise $W(t)$, $t \geq 0$, is one-dimensional standard Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$. The approach in [?] is based on embedding the SFDE (??) in an infinite-dimensional non-delay stochastic evolution equation in the Hilbert space $M_2 := \mathbf{R}^d \times L^2([-\tau, 0], \mathbf{R}^d)$. The weak numerical approximation is then performed at the level of the induced stochastic evolution equation in the Hilbert space M_2 .

In the present article, we prove weak convergence of order 1 of the Euler scheme for *fully non-linear* stochastic delay equations in \mathbf{R}^d , with multiple discrete (and continuous) delays and multidimensional Brownian noise. However, and for simplicity of exposition, we will focus on one-dimensional SDDEs with two delays and driven by a single Wiener process W . The appropriate extensions of our analysis to higher dimensions are straightforward. They are indicated in Section 4 of this article.

In order to describe the main result of the paper, we introduce some notation which will be used throughout the article.

Let $C([-\tau, 0], \mathbf{R})$ denote the Banach space of continuous paths $\eta : [-\tau, 0] \rightarrow \mathbf{R}$, given the supremum norm

$$\|\eta\|_C := \sup_{-\tau \leq s \leq 0} |\eta(s)|.$$

The symbol $H^{1,\infty}([-\tau, 0], \mathbf{R})$ denotes the Banach space of all continuous paths $\eta : [-\tau, 0] \rightarrow \mathbf{R}$, which are a.e. differentiable on $[-\tau, 0]$ and such that $\text{essup}_{s \in [-\tau, 0]} |\eta'(s)| < \infty$. The space $H^{1,\infty}([-\tau, 0], \mathbf{R})$ is furnished with the $H^{1,\infty}$ -norm

$$\|\eta\|_{1,\infty} := \sup_{-\tau \leq s \leq 0} |\eta(s)| + \text{essup}_{-\tau \leq s \leq 0} |\eta'(s)|.$$

The corresponding Banach spaces $C([-\tau, 0], \mathbf{R}^d)$ and $H^{1,\infty}([-\tau, 0], \mathbf{R}^d)$ of \mathbf{R}^d -valued mappings are defined analogously.

Consider the one-dimensional SDDE:

$$(1.2) \quad x(t) = \begin{cases} \eta(0) + \int_{\sigma}^t f(x(u-\tau_1), x(u)) du + \int_{\sigma}^t g(x(u-\tau_2), x(u)) dW(u), & \sigma \leq t \leq T, \\ \eta(t-\sigma) & \sigma - \tau \leq t < \sigma, \quad \tau := \tau_1 \vee \tau_2, \end{cases}$$

where $T > 0$ is fixed, the *initial instant* $\sigma \in [0, T]$, the coefficients $f, g : \mathbf{R}^2 \rightarrow \mathbf{R}$ satisfy suitable regularity and linear growth hypotheses, and the *initial path* $\eta \in H^{1,\infty}([-\tau, 0], \mathbf{R})$.

Let $\pi := \{-\tau = t_{-M} < t_{-M+1} < \dots < t_{-1} < t_0 = 0 < t_1 < t_2 \dots < t_{N-1} < t_N = T\}$ be a partition of $[-\tau, T]$, with mesh denoted by $|\pi| := \max\{(t_i - t_{i-1}) : -M+1 \leq i \leq N\}$. Let $\sigma \in [0, T]$, and for any $u \in [\sigma, T]$, define $\lfloor u \rfloor := t_{i-1} \vee \sigma$ whenever $u \in [t_{i-1}, t_i] \cap [\sigma, T]$. For each initial path $\eta \in H^{1,\infty}([-\tau, 0], \mathbf{R})$, define its *piecewise-linear approximation* $\eta^\pi \in H^{1,\infty}([-\tau, 0], \mathbf{R})$ by

$$\eta^\pi(s) := \left(\frac{t_i - s}{t_i - t_{i-1}} \right) \eta(t_{i-1}) + \left(\frac{s - t_{i-1}}{t_i - t_{i-1}} \right) \eta(t_i)$$

for $s \in [t_{i-1}, t_i]$, $-M+1 \leq i \leq 0$, and $\eta^\pi(0) := \eta(0)$. Define the continuous *Euler approximations* $y : [\sigma - \tau, T] \times \Omega \rightarrow \mathbf{R}$ of the solution $x : [\sigma - \tau, T] \times \Omega \rightarrow \mathbf{R}$ of (??) to be solutions of the SDDEs:

$$(1.3) \quad y(t) = \begin{cases} \eta(0) + \int_{\sigma}^t f(y(\lfloor u \rfloor - \tau_1), y(\lfloor u \rfloor)) du \\ \quad + \int_{\sigma}^t g(y(\lfloor u \rfloor - \tau_2), y(\lfloor u \rfloor)) dW(u), & \sigma \leq t \leq T, \\ \eta(t-\sigma), & \sigma - \tau \leq t < \sigma, \quad \tau := \tau_1 \vee \tau_2. \end{cases}$$

Denote solutions of (??) and (??) by $x(t; \sigma, \eta)$ and $y(t; \sigma, \eta)$, $\sigma - \tau \leq t \leq T$.

The main objective of this article is to show that, for a sufficiently regular test function $\phi : \mathbf{R} \rightarrow \mathbf{R}$, the following estimate holds:

$$(1.4) \quad |E\phi(x(t; \sigma, \eta)) - E\phi(y(t; \sigma, \eta^\pi))| \leq K(1 + \|\eta\|_{1,\infty}^q) |\pi|$$

for all $t \in [\sigma, T]$, $\sigma \in [0, T]$ and all $\eta \in H^{1,\infty}([-\tau, 0], \mathbf{R})$. In (1.4), $\|\cdot\|_{1,\infty}$ is the $H^{1,\infty}$ -norm of the initial path η , q is a positive integer, and K is a positive constant, independent of the partition π , the initial instant σ and the initial path η .

To establish the estimate (1.4), we employ an entirely different approach than that in [?], using the Malliavin calculus and anticipating stochastic analysis techniques developed by Nualart and Pardoux [?]. More specifically, we are able to handle SDDDEs with non-linear drift and diffusion coefficients. We further allow for multi-dimensional noise and several finite delays in both drift and diffusion terms, as well as *quasitame* dependence on the history in all the coefficients (cf. [?]).

As will be apparent in the sequel, the impetus for appealing to anticipating stochastic analysis is provided by the *tame Itô formula* developed in [?]. Malliavin calculus has recently become an important tool in the investigation of numerical methods for stochastic differential equations. See for example Bally and Talay [B.T], Kohatsu-Higa [?, ?] and Fournié, Lasry, Lebuchoux and Lions [?, ?]. Our approach is however different from these works as indicated at the end of this section.

We now give a brief outline of the proof of the estimate (1.4). We do this in several steps:

Step 1: For simplicity, take $\sigma = 0$, $t = t_n \in \pi \cap [0, T]$, $1 \leq n \leq N$. By exploiting the discrete delay structure of (1.3), it follows that the Euler approximation $y(t_n; t_i, \eta)$ is a (Fréchet) C^1 tame functional in $\eta \in C([-\tau, 0], \mathbf{R})$. Using a telescoping argument along the partition points t_i , Fréchet differentiability of the Euler approximation $y(t_n; t_i, \eta)$ in η , the Markov property for x_t and y_t together with the Mean Value Theorem, we write

$$\begin{aligned} E\phi(x(t_n; 0, \eta)) - E\phi(y(t_n; 0, \eta)) \\ = \sum_{i=1}^n E \int_0^1 D(\phi \circ y)(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i}) \cdot [x_{t_i} - y_{t_i}] d\lambda \end{aligned}$$

where $x_{t_i} := x_{t_i}(\cdot; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta))$ and $y_{t_i} := y_{t_i}(\cdot; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta))$.

Step 2: The main task is to show that each of the terms in the above sum is $O((t_i - t_{i-1})^2)$. This is realized through use of the *tame Itô formula*. The application of the latter formula to the differences $x_{t_i} - y_{t_i}$ generates multiple Skorohod integrals of the form

$$\begin{aligned} J_1^i &:= \int_{t_{i-1}-t_i}^0 Y(ds) \int_{t_{i-1}}^{t_i+s} \int_{t_{i-1}}^u \Sigma_1(v) dv dW(u), \\ J_2^i &:= \int_{t_{i-1}-t_i}^0 Y(ds) \int_{t_{i-1}}^{t_i+s} \int_{t_{i-1}}^u \Sigma_2(v) dW(v - \tau_2) dW(u), \\ J_3^i &:= \int_{t_{i-1}-t_i}^0 Y(ds) \int_{t_{i-1}}^{t_i+s} \int_{t_{i-1}}^u \Sigma_3(v) dW(v - \tau_1) du. \end{aligned}$$

In the above expressions, $Y(ds)$ is a random discrete measure on $[-\tau, 0]$ induced by Fréchet derivatives of the tame functionals $C([-\tau, 0], \mathbf{R}) \ni \eta \mapsto y(t_n; t_i, \eta) \in \mathbf{R}$. Thus $Y(ds)$ has Malliavin smooth random atoms. The processes Σ_j , $j = 1, 2, 3$, are Malliavin smooth and possibly anticipate the lagged Brownian motions $W(\cdot - \tau_i)$, $i = 1, 2$.

Step 3: To estimate the expectation of the terms containing Skorohod integrals in Step 2, we use the definition of the Skorohod integral as the adjoint of the weak differentiation operator, coupled with estimates on higher-order moments of the Malliavin derivatives of the Σ_j 's, $j = 1, 2, 3$. These higher order moment estimates are obtained using the corresponding higher moments of the Euler approximations y and their linearizations. This is a somewhat delicate computation which turns on the crucial fact that the Euler approximations $y(t_n; t_i, \eta)$ are tame in η . It yields

$$|EJ_j^i| = O((t_i - t_{i-1})^2), \quad j = 1, 2, 3.$$

Summing over $i = 1, \dots, n$, we get a positive constant K and a positive integer q such that the estimate

$$(1.5) \quad |E\phi(x(t; \sigma, \eta)) - E\phi(y(t; \sigma, \eta))| \leq K(1 + \|\eta\|_{1,\infty}^q) |\pi|$$

holds for all $t \in [\sigma, T]$, $\sigma \in [0, T]$ and all $\eta \in H^{1,\infty}([-\tau, 0], \mathbf{R})$.

Step 4: To complete the proof of the estimate (1.4), we need to replace η in the Euler approximation $y(t; \sigma, \eta)$ in (1.5) by its piecewise-linear approximation η^π . To do so, we use the triangle inequality

$$(1.6) \quad \begin{aligned} |E\phi(x(t; \sigma, \eta)) - E\phi(y(t; \sigma, \eta^\pi))| &\leq |E\phi(x(t; \sigma, \eta)) - E\phi(x(t; \sigma, \eta^\pi))| \\ &\quad + |E\phi(x(t; \sigma, \eta^\pi)) - E\phi(y(t; \sigma, \eta^\pi))| \end{aligned}$$

together with

$$(1.7) \quad |E\phi(x(t; \sigma, \eta)) - E\phi(x(t; \sigma, \eta^\pi))| \leq K\|\eta - \eta^\pi\|_C$$

and the elementary estimate

$$(1.8) \quad \|\eta - \eta^\pi\|_C \leq 2\|\eta'\|_\infty|\pi|.$$

See ([Mo.1], p. 41) for the estimate (1.7). The required estimate (1.4) now follows from (1.5)-(1.8).

The above outline highlights the following unique features of the analysis in this article:

- Although the original SDDE is *non-anticipating*, the tame Itô formula gives rise to *anticipating* terms containing Malliavin derivatives of the solution of the SDDE [?].
- The formulation and the implementation of the weak Euler scheme do not require the use of (or familiarity with) the Malliavin calculus.
- By contrast with the non-delay case (SODEs), SDDEs do not correspond to diffusions on Euclidean space. Thus techniques from deterministic PDEs do not apply.
- The use of anticipating calculus methods seems unavoidable in deriving the appropriate order of weak convergence (1) for the Euler scheme.

We conclude the introduction by giving an outline of the paper. In Section ?? we state the tame Itô formula used in Step 2 above. The main estimate (1.4) is proved in Section ?? in the one-dimensional case (Theorem 3.1). A multidimensional version of Theorem 3.1 with multiple discrete and quasitame delays is stated in Section 4 (Theorem 4.1). Section ?? provides numerical examples to illustrate our results.

2. THE TAME ITÔ FORMULA

Our proof of weak convergence of the Euler scheme depends crucially on an Itô formula for a certain class of functionals on $C([- \tau, 0], \mathbf{R}^d)$ called *tame functionals*. These tame functionals act on segments of sample-continuous random processes $[- \tau, \infty) \times \Omega \rightarrow \mathbf{R}^d$. We will refer to this formula as the *tame Itô formula*. Details of the proof of this formula are given in [?]. In the present context, we will only state the formula in one space dimension ($d = 1$), although the formula holds true in any Euclidean space for \mathbf{R}^d -valued processes driven by multi-dimensional Brownian motion.

First, we need some notation.

Suppose that $W(t)$, $t \geq 0$, is one-dimensional standard Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$. For simplicity of notation, set $W(t) := 0$ if $-\tau \leq t \leq 0$. Denote by \mathcal{D} the Malliavin differentiation operator associated with W .

Throughout the article, we will reserve the notation \mathcal{D} for the Malliavin differentiation operator, and D for the Fréchet differentiation operator.

For any $p > 1$ and any integer $k > 0$, let $\mathbb{L}^{k,p}$ stand for the space $L^p([0, T], \mathbb{D}^{k,p})$ where $\mathbb{D}^{k,p}$ is the closure of all real valued random variables Z having all weak derivatives $\mathcal{D}^j Z \in L^p(\Omega, H^{\otimes j})$ for $1 \leq j \leq k$, furnished with the norm

$$\|Z\|_{k,p} := (E|Z|^p)^{1/p} + \left(\sum_{j=1}^k E\|\mathcal{D}^j Z\|_{H^{\otimes j}}^p \right)^{1/p}.$$

In the above equation, H denotes the Hilbert space $L^2([0, T], \mathbf{R})$. The spaces $\mathbb{L}_{loc}^{k,p}$, $p > 4$, are defined to be the set of all processes X such that there is an increasing sequence of \mathcal{F} -measurable sets A_n , $n \geq 1$, and processes $X_n \in \mathbb{L}^{k,p}$, $n \geq 1$, such that $X = X_n$ a.s. on A_n for each $n \geq 1$, and $\bigcup_{n=1}^{\infty} A_n = \Omega$. By virtue of its local nature, the Malliavin differentiation operator \mathcal{D} extends unambiguously to the spaces $\mathbb{L}_{loc}^{k,p}$, $p > 4$. For further properties of weak derivatives and the spaces $\mathbb{L}^{k,p}$, the reader may refer to ([?, pp. 61, 151, 161] and [Nu.2]).

Define the *tame projection* $\Pi : C([-\tau, 0], \mathbf{R}) \rightarrow \mathbf{R}^k$ associated with $s_1, \dots, s_k \in [-\tau, 0]$ by

$$(2.1) \quad \Pi(\eta) := (\eta(s_1), \dots, \eta(s_k))$$

for all $\eta \in C([-\tau, 0], \mathbf{R}^d)$.

Denote by $C^{1,2}([0, T] \times \mathbf{R}^k, \mathbf{R})$ the space of all functions $\phi : [0, T] \times \mathbf{R}^k \rightarrow \mathbf{R}$ which are of class C^1 in the time variable $[0, T]$ and of class C^2 in the space variables \mathbf{R}^k .

A functional $\Psi : [0, T] \times C([-\tau, 0], \mathbf{R}) \rightarrow \mathbf{R}$ is called *tame*, if there exists a function $\phi \in C^{1,2}([0, T] \times \mathbf{R}^k, \mathbf{R})$ and a tame projection Π such that

$$\Psi(t, \eta) = \phi(t, \Pi(\eta))$$

for all $t \in [0, T]$ and $\eta \in C([-\tau, 0], \mathbf{R})$.

Consider a pathwise-continuous (not necessarily adapted) \mathbf{R} -valued process $X : [-\tau, \infty) \times \Omega \rightarrow \mathbf{R}$ given by

$$(2.2) \quad X(t) = \begin{cases} \eta(0) + \int_0^t u(s) dW(s) + \int_0^t v(s) ds, & t > 0, \\ \eta(t), & -\tau \leq t \leq 0, \end{cases}$$

where η belongs to $C([-\tau, 0], \mathbf{R})$ and is of bounded variation, $u \in \mathbb{L}_{loc}^{2,p}$, $p > 4$, and $v \in \mathbb{L}_{loc}^{1,4}$. The stochastic integral in (??) is understood in the Skorohod sense. Note that the processes u and v may not be adapted to the Brownian filtration $(\mathcal{F}_t)_{t \geq 0}$. For convenience, whenever $t \in [-\tau, 0)$, we set $u(t) := 0$, and $v(t) := \eta'(t)$, where η' is the (classical) derivative of η . Associate with X its *segment process* $X_t \in C([-\tau, 0], \mathbf{R})$, $t \in [0, T]$, defined by

$$(2.3) \quad X_t(s) := X(t + s), \quad s \in [-\tau, 0], \quad t \in [0, T].$$

We are now ready to state the *tame Itô formula*. It describes how the segment process $X_t \in C([-\tau, 0], \mathbf{R})$ transforms under tame functionals $\Psi : [0, T] \times C([-\tau, 0], \mathbf{R}) \rightarrow \mathbf{R}$.

Theorem 2.1. *Assume that X is a sample-continuous process defined by (??), where $\eta : [-\tau, 0] \rightarrow \mathbf{R}$ is of bounded variation, $u \in \mathbb{L}_{loc}^{2,4}$, and $v \in \mathbb{L}_{loc}^{1,4}$. Suppose $\phi \in C^{1,2}([0, T] \times \mathbf{R}^k, \mathbf{R})$, and let Π be the tame projection (2.1). Then for all $t \in [0, T]$ we have a.s.*

$$(2.4) \quad \begin{aligned} & \phi(t, \Pi(X_t)) - \phi(0, \Pi(X_0)) \\ &= \int_0^t \frac{\partial \phi}{\partial s}(s, \Pi(X_s)) ds \\ &+ \sum_{i=1}^k \int_0^t \frac{\partial \phi}{\partial x_i}(s, \Pi(X_s)) dX(s + s_i) \\ &+ \frac{1}{2} \sum_{i,j=1}^k \int_0^t \frac{\partial^2 \phi}{\partial x_i \partial x_j}(s, \Pi(X_s)) u(s + s_i) \nabla_{s_i, s_j}^+ X(s) ds \end{aligned}$$

where

$$\nabla_{s_i, s_j}^+ X(s) := \mathcal{D}_{s+s_i}^+ X(s + s_j) + \mathcal{D}_{s+s_i}^- X(s + s_j)$$

and

$$\mathcal{D}_{s+s_i}^+ X(s + s_j) := \lim_{\epsilon \rightarrow 0^+} \mathcal{D}_{s+s_i} X(s + s_j + \epsilon), \quad \mathcal{D}_{s+s_i}^- X(s + s_j) := \lim_{\epsilon \rightarrow 0^+} \mathcal{D}_{s+s_i} X(s + s_j - \epsilon).$$

Proof. See Hu, Mohammed and Yan [?], Theorem 2.3. □

Remark 2.2. Note the misprint in formula (2.7) of [?], where the factor $\mathcal{D}_{s+s_i} X(s + s_j)$ must be replaced by $\nabla_{s_i, s_j}^+ X(s)$.

The following corollary is an important special case of Theorem 2.1. In the proof of the weak Euler scheme (Theorem 3.1), it plays a crucial role in estimating the difference, across partition points, between the segments of the solution x of our SDDE (1.2) and its Euler approximation y .

Corollary 1. Let $\psi : \mathbf{R}^2 \rightarrow \mathbf{R}$ be of class C^2 , and suppose x solves the SDDE

$$(2.5) \quad x(t) = \begin{cases} \eta(0) + \int_0^t f(x(u - \tau_1), x(u)) \, du + \int_0^t g(x(u - \tau_2), x(u)) \, dW(u), & t > 0 \\ \eta(t) & -\tau < t < 0, \quad \tau := \tau_1 \vee \tau_2, \end{cases}$$

where the coefficients $f, g : \mathbf{R}^2 \rightarrow \mathbf{R}$ are of class C_b^2 , and $\eta \in C([-\tau, 0], \mathbf{R})$ is of bounded variation. Suppose $\delta > 0$. Then a.s.

$$(2.6) \quad \begin{aligned} & d\psi(x(t - \delta), x(t)) \\ &= \frac{\partial \psi}{\partial x_1}(x(t - \delta), x(t)) 1_{[0, \delta)}(t) \, d\eta(t - \delta) \\ &+ \frac{\partial \psi}{\partial x_1}(x(t - \delta), x(t)) 1_{[\delta, \infty)}(t) [f(x(t - \tau_1 - \delta), x(t - \delta)) \, dt \\ &\quad + g(x(t - \tau_2 - \delta), x(t - \delta)) \, dW(t - \delta)] \\ &+ \frac{\partial \psi}{\partial x_2}(x(t - \delta), x(t)) [f(x(t - \tau_1), x(t)) \, dt + g(x(t - \tau_2), x(t)) \, dW(t)] \\ &+ \frac{\partial^2 \psi}{\partial x_1 \partial x_2}(x(t - \delta), x(t)) g(x(t - \tau_2 - \delta), x(t - \delta)) 1_{[\delta, \infty)}(t) \mathcal{D}_{t-\delta}^- x(t) \, dt \\ &+ \frac{1}{2} \frac{\partial^2 \psi}{\partial x_1^2}(x(t - \delta), x(t)) g(x(t - \tau_2 - \delta), x(t - \delta))^2 1_{[\delta, \infty)}(t) \, dt \\ &+ \frac{1}{2} \frac{\partial^2 \psi}{\partial x_2^2}(x(t - \delta), x(t)) g(x(t - \tau_2), x(t))^2 \, dt, \end{aligned}$$

for all $t > 0$.

Proof. Suppose $t > \delta$. Apply Theorem ?? with $\psi := \psi(x_1, x_2)$, $X = x$, $\Pi(\eta) := (\eta(-\delta), \eta(0))$, $\eta \in C([-\tau, 0], \mathbf{R})$, where x solves the SDDE (??). This gives

$$\begin{aligned} d\psi(x(t - \delta), x(t)) &= \frac{\partial \psi}{\partial x_1}(x(t - \delta), x(t)) \, dx(t - \delta) + \frac{\partial \psi}{\partial x_2}(x(t - \delta), x(t)) \, dx(t) \\ &+ \frac{1}{2} \frac{\partial^2 \psi}{\partial x_1 \partial x_2}(x(t - \delta), x(t)) g(x(t - \delta - \tau_2), x(t - \delta)) 2\mathcal{D}_{t-\delta}^- x(t) \, dt \\ &+ \frac{1}{2} \frac{\partial^2 \psi}{\partial x_2 \partial x_1}(x(t - \delta), x(t)) g(x(t - \tau_2), x(t)) \underbrace{2\mathcal{D}_t x(t - \delta)}_0 \, dt \\ &+ \frac{1}{2} \frac{\partial^2 \psi}{\partial x_1^2}(x(t - \delta), x(t)) g(x(t - \delta - \tau_2), x(t - \delta)) \mathcal{D}_{t-\delta}^+ x(t - \delta) \, dt \\ &+ \frac{1}{2} \frac{\partial^2 \psi}{\partial x_2^2}(x(t - \delta), x(t)) g(x(t - \tau_2), x(t)) \mathcal{D}_t^+ x(t) \, dt, \end{aligned}$$

since $\mathcal{D}_t^- x(t) = 0$.

Now,

$$\mathcal{D}_t^+ x(t) = g(x(t - \tau_2), x(t))$$

and

$$\mathcal{D}_{t-\delta}^+ x(t - \delta) = g(x(t - \delta - \tau_2), x(t - \delta)).$$

Hence

$$\begin{aligned} d\psi(x(t - \delta), x(t)) &= \frac{\partial \psi}{\partial x_1}(x(t - \delta), x(t)) [f(x(t - \tau_1 - \delta), x(t - \delta)) \, dt \\ &\quad + g(x(t - \delta - \tau_2), x(t - \delta)) \, dW(t - \delta)] \\ &+ \frac{\partial \psi}{\partial x_2}(x(t - \delta), x(t)) [f(x(t - \tau_1), x(t)) \, dt + g(x(t - \tau_2), x(t)) \, dW(t)] \\ &+ \frac{\partial^2 \psi}{\partial x_1 \partial x_2}(x(t - \delta), x(t)) g(x(t - \delta - \tau_2), x(t - \delta)) \mathcal{D}_{t-\delta}^- x(t) \, dt \\ &+ \frac{1}{2} \frac{\partial^2 \psi}{\partial x_1^2}(x(t - \delta), x(t)) g(x(t - \delta - \tau_2), x(t - \delta))^2 \, dt \\ &+ \frac{1}{2} \frac{\partial^2 \psi}{\partial x_2^2}(x(t - \delta), x(t)) g(x(t - \tau_2), x(t))^2 \, dt. \end{aligned}$$

This proves statement (??) of the corollary when $t > \delta$.

If $0 < t < \delta$, then $x(t - \delta) = \eta(t - \delta)$ is of bounded variation; so (??) follows directly from the classical Itô formula. This completes the proof of the corollary. \square

Remark 2.3. Note that the second term in the right hand side of (??) contains the $(\mathcal{F})_{t \geq 0}$ -adapted factor $\frac{\partial \psi}{\partial x_1}(x(t - \delta), x(t))$ which anticipates the lagged Brownian motion $W(t - \delta)$. Although the process $[0, \infty) \ni t \mapsto (x(t - \delta), x(t)) \in \mathbf{R}^2$ is $(\mathcal{F})_{t \geq 0}$ -adapted, it is not a semimartingale with respect to any natural filtration. Therefore, it is not possible to infer (??) from the classical Itô formula for semi-martingales.

3. THE WEAK EULER SCHEME

In this section, we will establish weak convergence of order 1 of the Euler scheme for the non-linear SDDE

$$(3.1) \quad x(t) = \begin{cases} \eta(0) + \int_{\sigma}^t f(x(u - \tau_1), x(u)) \, du + \int_{\sigma}^t g(x(u - \tau_2), x(u)) \, dW(u), & t > \sigma, \\ \eta(t - \sigma) & \sigma - \tau < t < \sigma, \quad \tau := \tau_1 \vee \tau_2, \end{cases}$$

where the initial path $\eta \in H^{1, \infty}([-\tau, 0], \mathbf{R})$, $\sigma \geq 0$, and the coefficients $f, g : \mathbf{R}^2 \rightarrow \mathbf{R}$ are of class C_b^3 ; viz. f and g have all partial derivatives up to order 3 globally bounded on \mathbf{R}^2 . The space of all such functions is often denoted by $C_b^3(\mathbf{R}^2, \mathbf{R})$. Similarly $C_b^2(\mathbf{R}^2, \mathbf{R})$ denotes all functions $\mathbf{R}^2 \rightarrow \mathbf{R}$ with first and second-order partial derivatives globally bounded. As before, $W(t)$, $t \geq 0$, is one-dimensional standard Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$. Set $W(t) := 0$ if $-\tau \leq t \leq 0$.

Recall the partition $\pi := \{t_{-M} < t_{-M+1} < \dots < t_{-1} < t_0 = 0 < t_1 < t_2 \dots < t_{N-1} < t_N = T\}$ of $[-\tau, T]$, with mesh denoted by $|\pi|$, introduced in Section 1. The Euler approximations $y := y(\cdot; \sigma, \eta) : [\sigma - \tau, T] \times \Omega \rightarrow \mathbf{R}$ of the solution $x(\cdot; \sigma, \eta)$ of (??) satisfy the SDDEs

$$(3.2) \quad y(t) = \begin{cases} \eta(0) + \int_{\sigma}^t f(y(\lfloor u \rfloor - \tau_1), y(\lfloor u \rfloor)) \, du + \int_{\sigma}^t g(y(\lfloor u \rfloor - \tau_2), y(\lfloor u \rfloor)) \, dW(u), & t \geq \sigma, \\ \eta(t - \sigma), & \sigma - \tau < t < \sigma, \quad \tau := \tau_1 \vee \tau_2. \end{cases}$$

The following is the main result of this article. It establishes weak convergence of order 1 of the Euler scheme (??) to the solution of the SDDE (??).

Theorem 3.1. *Let π be a partition of $[-\tau, T]$ with mesh $|\pi|$, and $\phi : \mathbf{R} \rightarrow \mathbf{R}$ be of class C_b^3 . In the SDDE (??), assume that the coefficients f, g are C_b^3 . Let $x(\cdot; \sigma, \eta)$ be the unique solution of (??) with initial path $\eta \in H^{1, \infty}([-\tau, 0], \mathbf{R})$. Let $\eta^\pi \in H^{1, \infty}([-\tau, 0], \mathbf{R})$ be the piecewise-linear approximation of η along the partition π . Denote by $y(\cdot; \sigma, \eta)$ the Euler approximation to $x(\cdot; \sigma, \eta)$ associated with the partition π and defined by (??). Then there is a positive constant K and a positive integer q such that*

$$(3.3) \quad |E\phi(x(t; \sigma, \eta)) - E\phi(y(t; \sigma, \eta^\pi))| \leq K(1 + \|\eta\|_{1, \infty}^q) |\pi|$$

for all $t \in [\sigma - \tau, T]$, $\sigma \in [0, T]$, and all $\eta \in H^{1, \infty}([-\tau, 0], \mathbf{R})$. The constant K may depend on T, q and the test function ϕ , but is independent of $\pi, \eta, t \in [\sigma, T]$ and $\sigma \in [0, T]$.

For simplicity of presentation, we will only discuss the case of single delays τ_1, τ_2 in the drift and diffusion coefficients as in (??). It should be noted that the conclusion of Theorem ?? still holds if we allow for multi-dimensional noise and *several* delays in the drift and diffusion coefficients of the SDDE. The extension of Theorem ?? to this case is given in in Section ?? of this article.

The proof of Theorem ?? requires the following sequence of lemmas. In these lemmas and for the rest of the section, we will refer to the solutions of (3.1) and (3.2) by $x(\cdot; \sigma, \eta)$ and $y(\cdot; \sigma, \eta)$, respectively.

The first lemma establishes the tame character of the Euler approximation $y(t; \sigma, \eta)$ and its Fréchet derivative $Dy(t; \sigma, \eta)$ in the initial path η . This fact dictates that the telescoping argument in the proof of Theorem 3.1 is with respect to the Euler approximation y and *not* the solution x of the SDDE (??). This issue is especially important in view of the well-known fact that the solution $x(t; \sigma, \eta)$ of (??) is *almost surely extremely erratic* in the initial path η ([Mo.1], Chapter V, pp. 144-148).

Lemma 3.2. Fix a partition point $t_i \in \pi$ for some $i \in \{0, 1, \dots, N\}$. Then for a.a. $\omega \in \Omega$, the function

$$\begin{aligned} [t_i, T] \times C([- \tau, 0], \mathbf{R}) &\rightarrow \mathbf{R}, \\ (t, \eta) &\mapsto y(t, \omega; t_i, \eta) \end{aligned}$$

is a tame function. That is, there exists a deterministic function $F : \mathbf{R}^+ \times \mathbf{R}^{k+1} \times \mathbf{R}^h \times \mathbf{R}^l \rightarrow \mathbf{R}$ which is piece-wise continuous in the first variable (the time variable) and of class C_b^2 in all other variables (space variables), and there exist fixed numbers $t_1, t_2, \dots, t_k \leq t$, $s_1, s_2, \dots, s_h \leq t$, $\mu_1, \mu_2, \dots, \mu_l \in [-\tau, 0]$ such that a.s.

$$y(t; t_i, \eta) = F(t, W(t), W(t_1), W(t_2), \dots, W(t_k), s_1, s_2, \dots, s_h, \eta(\mu_1), \eta(\mu_2), \dots, \eta(\mu_l))$$

for all $\eta \in C([- \tau, 0], \mathbf{R})$ and all $t \in [t_i, T]$. In particular, for a.a. $\omega \in \Omega$ and each $t \in [t_i, T]$, the map

$$C([- \tau, 0], \mathbf{R}) \ni \eta \mapsto y(t, \omega; t_i, \eta) \in \mathbf{R}$$

is C^1 (in the Fréchet sense), and

$$(3.4) \quad \begin{aligned} Dy(t, \omega; t_i, \eta)(\xi) &= \sum_{m=1}^l \partial_m F(t, W(t, \omega), W(t_1, \omega), \dots, W(t_k, \omega), s_1, \dots, s_h, \eta(\mu_1), \\ &\quad \dots, \eta(\mu_m), \dots, \eta(\mu_l)) \xi(\mu_m) \end{aligned}$$

for all $\eta \in C([- \tau, 0], \mathbf{R})$ and for every bounded measurable function $\xi : [-\tau, 0] \rightarrow \mathbf{R}$. In the above formula, $\partial_m F$ denotes the partial derivative of F with respect to the variable $\eta(\mu_m)$.

Proof. The second and last assertions of the lemma are direct consequences of the first. So we will only prove the first assertion. The latter assertion is proved using forward steps along the partition points $\{0 = t_0, t_1, t_2, t_3, \dots, t_N = T\}$, and finite induction. More specifically, and with no loss of generality, suppose $i = 0$ and let $t \in [0, T]$. Consider the following cases:

$0 \leq t \leq t_1$:

$$\begin{aligned} y(t; 0, \eta) &= \eta(0) + f(\eta(-\tau_1), \eta(0))t + g(\eta(-\tau_2), \eta(0))W(t) \\ &= F_1(t, W(t), \eta(0), \eta(-\tau_1), \eta(-\tau_2)). \end{aligned}$$

This is clearly a tame function of η , satisfying the regularity properties stated in the lemma.

$t_1 < t \leq t_2$:

$$y(t; 0, \eta) = y(t_1) + f(y(t_1 - \tau_1), y(t_1))(t - t_1) + g(y(t_1 - \tau_2), y(t_1)) [W(t) - W(t_1)].$$

This is a tame function of η because $y(t_1; 0, \eta)$, $y(t_1 - \tau_1; 0, \eta)$ and $y(t_1 - \tau_2; 0, \eta)$ are all tame functions of η , and the composition of tame functions with real-valued functions is again tame. Indeed, suppose without loss of generality that $\tau_1 \leq \tau_2$ and let $t_1 - \tau_1 < 0$. Then

$$\begin{aligned} y(t; 0, \eta) &= \eta(0) + f(\eta(-\tau_1), \eta(0))t_1 + g(\eta(-\tau_2), \eta(0))W(t_1) \\ &+ f(\eta(t_1 - \tau_1), \eta(0) + f(\eta(-\tau_1), \eta(0))t_1 + g(\eta(-\tau_2), \eta(0))W(t_1))(t - t_1) \\ &+ g(\eta(t_1 - \tau_2), \eta(0) + f(\eta(-\tau_1), \eta(0))t_1 + g(\eta(-\tau_2), \eta(0))W(t_1)) [W(t) - W(t_1)] \\ &= F_2(t, W(t), W(t_1), t_1, \eta(0), \eta(t_1 - \tau_1), \eta(t_1 - \tau_2), \eta(-\tau_1), \eta(-\tau_2)). \end{aligned}$$

It is easy to see that F_2 is tame and fulfills the regularity requirements of the lemma. The other cases $\tau_1 \leq t_1 < \tau_2$, and $\tau_1 \leq \tau_2 \leq t_1$ can be treated similarly.

$t_2 < t \leq t_3$:

$$\begin{aligned} y(t; 0, \eta) &= y(t_2) + f(y(t_2 - \tau_1), y(t_2))(t - t_2) + g(y(t_2 - \tau_2), y(t_2)) [W(t) - W(t_2)] \\ &= F_3(t, W(t), W(t_1), W(t_2), W(t_2 - \tau_1), W(t_2 - \tau_2), t_1, t_2, \eta(0), \eta(-\tau_1), \\ &\quad \eta(-\tau_2), \eta(t_1 - \tau_1), \eta(t_1 - \tau_2)), \end{aligned}$$

with F_3 tame in η , as required.

$t_k < t \leq t_{k+1}$:

By induction, there are fixed numbers $s_1, s_2, \dots, s_k < t_k$, and $\mu_1, \mu_2, \dots, \mu_l \in [-\tau, 0]$ such that

$$y(t; 0, \eta) = F_{k+1}(t, W(t), W(s_i), t_j, 1 \leq j \leq k, 1 \leq i \leq k, W(t), \eta(\mu_m), 1 \leq m \leq l).$$

This is a tame function of η , continuous in the time variable t and of class C_b^2 in all space variables.

To complete the proof of the lemma, take $F(t, \cdot, \cdot, \cdot, \cdot) := \sum_{i=1}^{N-1} 1_{[t_i, t_{i+1})}(t) F_{i+1}(\cdot, \cdot, \cdot, \cdot)$, for $t \in [0, T]$. It is easy to check that F satisfies all the requirements of the lemma. \square

Warning:

The lemma is *false* if the Euler approximation y is replaced by the exact solution x of the SDDE (??). In fact, for a.a. $\omega \in \Omega$, every measurable version $x(t, \omega; 0, \eta)$ of the solution to (3.1) is locally unbounded in η ([Mo.1], pp. 144-147). As will be apparent later in the proof of Theorem 3.1, this fact will force the telescoping argument to be centered about the Euler approximation y rather than the solution x . On the other hand, the following statement is true: Let $\phi : \mathbf{R} \rightarrow \mathbf{R}$ be C_b^∞ , $h, g : \mathbf{R}^2 \rightarrow \mathbf{R}$ be C_b^∞ . Then the map

$$L^2(\Omega, C([-\tau, 0], \mathbf{R}); \mathcal{F}_\sigma) \ni \psi \xrightarrow{F} E\phi(x(t; \sigma, \psi)) \in \mathbf{R}$$

is globally Lipschitz and (Fréchet) C^1 , with Fréchet derivative

$$DF(\psi)(\eta) = E\{D\phi(x(t; \sigma, \psi))Dx(t; \sigma, \psi)(\eta)\}$$

for all $\psi, \eta \in L^2(\Omega, C([-\tau, 0], \mathbf{R}); \mathcal{F}_\sigma)$. To see this, note that the Lipschitz and C^1 properties of the map $L^2(\Omega, C([-\tau, 0], \mathbf{R}); \mathcal{F}_\sigma) \ni \psi \mapsto x(t; \sigma, \psi) \in L^2(\Omega, \mathbf{R})$ follow by arguments similar to the proofs of Theorems 3.1, 3.2 in ([Mo.1], pp. 41-45). Compose the above map with the C_b^∞ function $\phi : \mathbf{R} \rightarrow \mathbf{R}$ to establish the above assertion. Details are omitted.

The following lemma is key to the proof of Theorem ???. It involves the application of the tame Itô formula in Corollary 1, Section 2. See Step 2, Section 1.

Lemma 3.3. *Assume that f, g are C_b^2 . Fix $\eta \in C([-\tau, 0], \mathbf{R})$, of bounded variation. For each $1 \leq i \leq N$, define the process $\Lambda^i : [-\tau, 0] \times \Omega \rightarrow \mathbf{R}$ by*

$$\Lambda^i := x_{t_i}(\cdot; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta)) - y_{t_i}(\cdot; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta)),$$

viz.

$$\Lambda^i(s) := x(t_i + s; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta)) - y(t_i + s; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta)), \quad s \in [-\tau, 0].$$

For brevity of notation, set $x(u) := x(u; 0, \eta)$ and $y(u) := y(u; 0, \eta)$ for $u \in [-\tau, T]$. Then

$$\begin{aligned} \Lambda^i(s) &= \int_{t_{i-1}}^{(t_i+s) \vee t_{i-1}} [f(x(u - \tau_1), x(u)) - f(x(\lfloor u \rfloor - \tau_1), x(\lfloor u \rfloor))] du \\ &+ \int_{t_{i-1}}^{(t_i+s) \vee t_{i-1}} [g(x(u - \tau_2), x(u)) - g(x(\lfloor u \rfloor - \tau_2), x(\lfloor u \rfloor))] dW(u) \\ (3.5) \quad &:= \sum_{j=1}^{10} \Lambda_j^i(s), \quad s \in [-\tau, 0], \end{aligned}$$

where

$$\begin{aligned}
\Lambda_1^i(s) &:= \int_{t_{i-1}}^{(t_i+s) \vee t_{i-1}} \int_{[u]}^u \frac{\partial f}{\partial x_1}(x(v-\tau_1), x(v)) f(x(v-2\tau_1), x(v-\tau_1)) 1_{[\tau_1, \infty)}(v) \, dv \, du \\
&\quad + \int_{t_{i-1}}^{(t_i+s) \vee t_{i-1}} \int_{[u]}^u \frac{\partial f}{\partial x_1}(x(v-\tau_1), x(v)) g(x(v-\tau_1-\tau_2), x(v-\tau_1)) \times \\
&\quad \quad \quad \times 1_{[\tau_1, \infty)}(v) \, dW(v-\tau_1) \, du \\
&\quad + \int_{t_{i-1}}^{(t_i+s) \vee t_{i-1}} \int_{[u]}^u \frac{\partial f}{\partial x_1}(x(v-\tau_1), x(v)) 1_{[0, \tau_1)}(v) \, d\eta(v-\tau_1) \, du, \\
\Lambda_2^i(s) &:= \int_{t_{i-1}}^{(t_i+s) \vee t_{i-1}} \int_{[u]}^u \frac{\partial f}{\partial x_2}(x(v-\tau_1), x(v)) f(x(v-\tau_1), x(v)) \, dv \, du \\
&\quad + \int_{t_{i-1}}^{(t_i+s) \vee t_{i-1}} \int_{[u]}^u \frac{\partial f}{\partial x_2}(x(v-\tau_1), x(v)) g(x(v-\tau_2), x(v)) \, dW(v) \, du, \\
\Lambda_3^i(s) &:= \int_{t_{i-1}}^{(t_i+s) \vee t_{i-1}} \int_{[u]}^u \frac{\partial^2 f}{\partial x_1 \partial x_2}(x(v-\tau_1), x(v)) g(x(v-\tau_1-\tau_2), x(v-\tau_1)) \times \\
&\quad \quad \quad \times 1_{[\tau_1, \infty)}(v) \mathcal{D}_{v-\tau_1} x(v) \, dv \, du, \\
\Lambda_4^i(s) &:= \frac{1}{2} \int_{t_{i-1}}^{(t_i+s) \vee t_{i-1}} \int_{[u]}^u \frac{\partial^2 f}{\partial x_1^2}(x(v-\tau_1), x(v)) g(x(v-\tau_1-\tau_2), x(v-\tau_1))^2 \times \\
&\quad \quad \quad \times 1_{[\tau_1, \infty)}(v) \, dv \, du, \\
\Lambda_5^i(s) &:= \frac{1}{2} \int_{t_{i-1}}^{(t_i+s) \vee t_{i-1}} \int_{[u]}^u \frac{\partial^2 f}{\partial x_2^2}(x(v-\tau_1), x(v)) g(x(v-\tau_2), x(v))^2 \, dv \, du, \\
\Lambda_6^i(s) &:= \int_{t_{i-1}}^{(t_i+s) \vee t_{i-1}} \int_{[u]}^u \frac{\partial g}{\partial x_1}(x(v-\tau_2), x(v)) f(x(v-\tau_1-\tau_2), x(v-\tau_2)) \\
&\quad \quad \quad \times 1_{[\tau_2, \infty)}(v) \, dv \, dW(u) \\
&\quad + \int_{t_{i-1}}^{(t_i+s) \vee t_{i-1}} \int_{[u]}^u \frac{\partial g}{\partial x_1}(x(v-\tau_2), x(v)) 1_{[0, \tau_2)}(v) \, d\eta(v-\tau_2) \, dW(u), \\
\Lambda_7^i(s) &:= \int_{t_{i-1}}^{(t_i+s) \vee t_{i-1}} \int_{[u]}^u \frac{\partial g}{\partial x_1}(x(v-\tau_2), x(v)) g(x(v-2\tau_2), x(v-\tau_2)) \times \\
&\quad \quad \quad \times 1_{[\tau_2, \infty)}(v) \, dW(v-\tau_2) \, dW(u), \\
\Lambda_8^i(s) &:= \int_{t_{i-1}}^{(t_i+s) \vee t_{i-1}} \int_{[u]}^u \frac{\partial^2 g}{\partial x_1 \partial x_2}(x(v-\tau_2), x(v)) g(x(v-2\tau_2), x(v-\tau_2)) \times \\
&\quad \quad \quad \times 1_{[\tau_2, \infty)}(v) \mathcal{D}_{v-\tau_1} x(v) \, dv \, dW(u) \\
\Lambda_9^i(s) &:= \frac{1}{2} \int_{t_{i-1}}^{(t_i+s) \vee t_{i-1}} \int_{[u] \vee \tau_1}^u \frac{\partial^2 g}{\partial x_1^2}(x(v-\tau_2), x(v)) g(x(v-2\tau_2), x(v-\tau_2))^2 \times \\
&\quad \quad \quad \times 1_{[\tau_2, \infty)}(v) \, dv \, dW(u), \\
\Lambda_{10}^i(s) &:= \frac{1}{2} \int_{t_{i-1}}^{(t_i+s) \vee t_{i-1}} \int_{[u]}^u \frac{\partial^2 g}{\partial x_2^2}(x(v-\tau_2), x(v)) g(x(v-\tau_2), x(v))^2 \, dv \, dW(u)
\end{aligned}$$

for all $s \in [-\tau, 0]$ and $1 \leq i \leq N$.

Proof. Fix $1 \leq i \leq N$.

Recall that $x(u) := x(u; 0, \eta)$ and $y(u) := y(u; 0, \eta)$, $-\tau \leq u \leq T$. Suppose $t_{i-1} \leq u < t_i$. Then $[u] = t_{i-1}$ and $[u] - \tau_1 = t_{i-1} - \tau_1 < t_{i-1}$. Hence

$$\begin{aligned}
y([u]; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta)) &= y(t_{i-1}; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta)) = x_{t_{i-1}}(0; 0, \eta) \\
&= x(t_{i-1}; 0, \eta) = x([u])
\end{aligned}$$

and

$$\begin{aligned}
y([u] - \tau_1; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta)) &= y(t_{i-1} - \tau_1; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta)) \\
&= x_{t_{i-1}}(-\tau_1; 0, \eta) = x([u] - \tau_1).
\end{aligned}$$

Furthermore, if $t_{i-1} \leq u < t_i$, then by the Markov property for the solution segment x_t , it follows that

$$x(u; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta)) = x(u; 0, \eta) = x(u).$$

In addition, suppose that $u - \tau_1 > t_{i-1}$. Then we may replace u by $u - \tau_1$ in the above identity to obtain

$$x(u - \tau_1; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta)) = x(u - \tau_1; 0, \eta) = x(u - \tau_1).$$

On the other hand if $u - \tau_1 \leq t_{i-1}$, then

$$\begin{aligned} x(u - \tau_1; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta)) &= x_{t_{i-1}}(u - \tau_1 - t_{i-1}; 0, \eta) \\ &= x(u - \tau_1; 0, \eta) \\ &= x(u - \tau_1). \end{aligned}$$

Now suppose $s \in [-\tau, 0]$ and consider

$$\Lambda^i(s) := x(t_i + s; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta)) - y(t_i + s; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta))$$

in the following two cases:

$0 \leq t_{i-1} \leq t_i + s$:

$$\begin{aligned} \Lambda^i(s) &= \int_{t_{i-1}}^{t_i+s} f(x(u - \tau_1), x(u)) \, du + \int_{t_{i-1}}^{t_i+s} g(x(u - \tau_2), x(u)) \, dW(u) \\ &\quad - \int_{t_{i-1}}^{t_i+s} f(x(\lfloor u \rfloor - \tau_1), x(\lfloor u \rfloor)) \, du - \int_{t_{i-1}}^{t_i+s} g(x(\lfloor u \rfloor - \tau_2), x(\lfloor u \rfloor)) \, dW(u) \\ &= \int_{t_{i-1}}^{t_i+s} [f(x(u - \tau_1), x(u)) - f(x(\lfloor u \rfloor - \tau_1), x(\lfloor u \rfloor))] \, du \\ &\quad + \int_{t_{i-1}}^{t_i+s} [g(x(u - \tau_2), x(u)) - g(x(\lfloor u \rfloor - \tau_2), x(\lfloor u \rfloor))] \, dW(u). \end{aligned}$$

$t_i - \tau \leq t_i + s < t_{i-1} < t_i$:

$$\Lambda^i(s) = x_{t_{i-1}}(t_i + s - t_{i-1}; 0, \eta) - x_{t_{i-1}}(t_i + s - t_{i-1}; 0, \eta) = 0.$$

Putting the above two cases together, gives

$$(3.6) \quad \begin{aligned} \Lambda^i(s) &= \int_{t_{i-1}}^{(t_i+s) \vee t_{i-1}} [f(x(u - \tau_1), x(u)) - f(x(\lfloor u \rfloor - \tau_1), x(\lfloor u \rfloor))] \, du \\ &\quad + \int_{t_{i-1}}^{(t_i+s) \vee t_{i-1}} [g(x(u - \tau_2), x(u)) - g(x(\lfloor u \rfloor - \tau_2), x(\lfloor u \rfloor))] \, dW(u), \end{aligned}$$

a.s. for all $s \in [-\tau, 0]$.

Since f is in C_b^2 , we may now apply the tame Itô formula (??) (with $\psi = f$, $\delta = \tau_1$) to obtain:

$$\begin{aligned}
& f(x(u - \tau_1), x(u)) - f(x(\lfloor u \rfloor - \tau_1), x(\lfloor u \rfloor)) \\
&= \int_{\lfloor u \rfloor}^u \frac{\partial f}{\partial x_1}(x(v - \tau_1), x(v)) 1_{[\tau_1, \infty)}(v) [f(x(v - 2\tau_1), x(v - \tau_1)) \, dv \\
&\quad + g(x(v - \tau_1 - \tau_2), x(v - \tau_1)) \, dW(v - \tau_1)] \\
&\quad + \int_{\lfloor u \rfloor}^u \frac{\partial f}{\partial x_1}(x(v - \tau_1), x(v)) 1_{[0, \tau_1)}(v) \, d\eta(v - \tau_1) \\
&\quad + \int_{\lfloor u \rfloor}^u \frac{\partial f}{\partial x_2}(x(v - \tau_1), x(v)) [f(x(v - \tau_1), x(v)) \, dv \\
&\quad + g(x(v - \tau_2), x(v)) \, dW(v)] \\
&\quad + \int_{\lfloor u \rfloor}^u \frac{\partial^2 f}{\partial x_1 \partial x_2}(x(v - \tau_1), x(v)) g(x(v - \tau_1 - \tau_2), x(v - \tau_1)) \times \\
&\quad \quad \times 1_{[\tau_1, \infty)}(v) \mathcal{D}_{v - \tau_1} x(v) \, dv \\
&\quad + \frac{1}{2} \int_{\lfloor u \rfloor}^u \frac{\partial^2 f}{\partial x_1^2}(x(v - \tau_1), x(v)) g(x(v - \tau_1 - \tau_2), x(v - \tau_1))^2 1_{[\tau_1, \infty)}(v) \, dv \\
&\quad + \frac{1}{2} \int_{\lfloor u \rfloor}^u \frac{\partial^2 f}{\partial x_2^2}(x(v - \tau_1), x(v)) g(x(v - \tau_2), x(v))^2 \, dv.
\end{aligned} \tag{3.7}$$

Similarly, from (??) (with $\psi = g$, $\delta = \tau_2$), it follows that

$$\begin{aligned}
& g(x(u - \tau_2), x(u)) - g(x(\lfloor u \rfloor - \tau_2), x(\lfloor u \rfloor)) \\
&= \int_{\lfloor u \rfloor}^u \frac{\partial g}{\partial x_1}(x(v - \tau_2), x(v)) 1_{[\tau_2, \infty)}(v) [f(x(v - \tau_1 - \tau_2), x(v - \tau_2)) \, dv \\
&\quad + g(x(v - 2\tau_2), x(v - \tau_2)) \, dW(v - \tau_2)] + \int_{\lfloor u \rfloor}^u \frac{\partial g}{\partial x_1}(x(v - \tau_2), x(v)) 1_{[0, \tau_2)}(v) \, d\eta(v - \tau_2) \\
&\quad + \int_{\lfloor u \rfloor}^u \frac{\partial g}{\partial x_2}(x(v - \tau_2), x(v)) [f(x(v - \tau_1), x(v)) \, dv + g(x(v - \tau_2), x(v)) \, dW(v)] \\
&\quad + \int_{\lfloor u \rfloor}^u \frac{\partial^2 g}{\partial x_1 \partial x_2}(x(v - \tau_2), x(v)) g(x(v - 2\tau_2), x(v - \tau_2)) 1_{[\tau_2, \infty)}(v) \mathcal{D}_{v - \tau_1} x(v) \, dv \\
&\quad + \frac{1}{2} \int_{\lfloor u \rfloor}^u \frac{\partial^2 g}{\partial x_1^2}(x(v - \tau_2), x(v)) g(x(v - 2\tau_2), x(v - \tau_2))^2 1_{[\tau_2, \infty)}(v) \, dv \\
&\quad + \frac{1}{2} \int_{\lfloor u \rfloor}^u \frac{\partial^2 g}{\partial x_2^2}(x(v - \tau_2), x(v)) g(x(v - \tau_2), x(v))^2 \, dv.
\end{aligned} \tag{3.8}$$

Now substitute from (??) and (??) into (??) to obtain the last assertion (??) of the lemma. \square

Lemma 3.4. *Suppose $f, g \in C_b^2$. Then for any $p \geq 1$ there is a positive constant K such that*

$$\sup_{0 \leq \sigma \leq T} \sup_{\sigma - \tau \leq u, t \leq T} E |\mathcal{D}_u y(t; \sigma, \eta)|^{2p} < K \left(1 + E \|\eta\|_C^{2p} + \sup_{\sigma - \tau \leq s \leq \sigma} E \|\mathcal{D}_s \eta\|_\infty^{2p} \right) \tag{3.9}$$

$$\sup_{\substack{\|\xi\|_\infty \leq 1 \\ \xi \in L^\infty([-\tau, 0], \mathbf{R})}} \sup_{0 \leq \sigma \leq T} \sup_{\sigma - \tau \leq u, t \leq T} E |\mathcal{D}_u D y(t; \sigma, \eta)(\xi)|^{2p} < K \left(1 + E \|\eta\|_C^{4p} + \sup_{\sigma - \tau \leq s \leq \sigma} E \|\mathcal{D}_s \eta\|_\infty^{4p} \right)^{1/2} \tag{3.10}$$

and

$$\sup_{0 \leq \sigma \leq T} \sup_{\sigma - \tau \leq u, t \leq T} E |\mathcal{D}_u x(t; \sigma, \eta)|^{2p} < K \left(1 + E \|\eta\|_C^{2p} + \sup_{\sigma - \tau \leq s \leq \sigma} E \|\mathcal{D}_s \eta\|_\infty^{2p} \right) \tag{3.11}$$

for all $\eta \in L^{4p}(\Omega, C([-\tau, 0], \mathbf{R}); \mathcal{F}_\sigma)$ which are Malliavin smooth and such that $\sup_{\sigma - \tau \leq s \leq \sigma} E \|\mathcal{D}_s \eta\|_\infty^{4p} < \infty$. The constant K is independent of $t \in [\sigma - \tau, T]$, $\sigma \in [0, T]$ and η , but may depend on p, f, g and T .

Proof. We first establish the estimate (3.9). Let $\eta \in L^{2p}(\Omega, C([- \tau, 0], \mathbf{R}); \mathcal{F}_\sigma)$ be Malliavin smooth and such that $\sup_{\sigma - \tau \leq s \leq \sigma} E \|\mathcal{D}_s \eta\|_\infty^{2p} < \infty$. For brevity, denote by $y(t) := y(t; \sigma, \eta)$, $t \in [\sigma - \tau, T]$, the

solution of (3.2) with initial process η at σ .

Take Malliavin derivatives of the equation

$$(3.12) \quad y(t) = \begin{cases} \eta(0) + \int_\sigma^t f(y(\lfloor v \rfloor - \tau_1), y(\lfloor v \rfloor)) \, dv \\ \quad + \int_\sigma^t g(y(\lfloor v \rfloor - \tau_2), y(\lfloor v \rfloor)) \, dW(v), & t \geq \sigma \\ \eta(t - \sigma), & \sigma - \tau < t < \sigma, \end{cases}$$

to get

$$(3.13) \quad \begin{aligned} \mathcal{D}_u y(t) &= g(y(\lfloor u \rfloor - \tau_1), y(\lfloor u \rfloor)) + \int_u^t \frac{\partial f}{\partial x_1}(y(\lfloor v \rfloor - \tau_1), y(\lfloor v \rfloor)) \mathcal{D}_u y(\lfloor v \rfloor - \tau_1) \, dv \\ &+ \int_u^t \frac{\partial f}{\partial x_2}(y(\lfloor v \rfloor - \tau_1), y(\lfloor v \rfloor)) \mathcal{D}_u y(\lfloor v \rfloor) \, dv \\ &+ \int_u^t \frac{\partial g}{\partial x_1}(y(\lfloor v \rfloor - \tau_2), y(\lfloor v \rfloor)) \mathcal{D}_u y(\lfloor v \rfloor - \tau_2) \, dW(v) \\ &+ \int_u^t \frac{\partial g}{\partial x_2}(y(\lfloor v \rfloor - \tau_2), y(\lfloor v \rfloor)) \mathcal{D}_u y(\lfloor v \rfloor) \, dW(v) \end{aligned}$$

for $t \geq u$; and $\mathcal{D}_u y(t) = \mathcal{D}_u \eta(t - \sigma)$ for $\sigma - \tau < t < u$ ([B.M]).

The idea is to estimate the function

$$(3.14) \quad \theta(t') := \sup_{u \leq t \leq t'} E |\mathcal{D}_u y(t)|^{2p}, \quad u \leq t' \leq T,$$

for fixed $u \in [\sigma - \tau, T]$, using (??) and Gronwall's lemma.

Since f, g have linear growth, it is easy to see from (??) and Gronwall's lemma, that

$$(3.15) \quad \sup_{\sigma - \tau \leq t \leq T} E |y(t)|^{2p} \leq K (E \|\eta\|_C^{2p} + 1).$$

Since g has linear growth, then

$$(3.16) \quad \sup_{0 \leq u \leq T} E |g(y(\lfloor u \rfloor - \tau_1), y(\lfloor u \rfloor))|^{2p} \leq K (E \|\eta\|_C^{2p} + 1).$$

Fix $u \geq \sigma - \tau$. Take $E|\cdot|^{2p}$ on both sides of (??) (using the fact that f, g have bounded derivatives) to obtain

$$\begin{aligned} E |\mathcal{D}_u y(t)|^{2p} &\leq K \left(1 + E \|\eta\|_C^{2p} + \sup_{\sigma - \tau \leq s \leq \sigma} E \|\mathcal{D}_s \eta\|_\infty^{2p} \right) + K \int_u^t E |\mathcal{D}_u y(\lfloor v \rfloor - \tau_1)|^{2p} \, dv \\ &+ K \int_u^t E |\mathcal{D}_u y(\lfloor v \rfloor)|^{2p} \, dv + K \int_u^t E |\mathcal{D}_u y(\lfloor v \rfloor - \tau_2)|^{2p} \, dv \\ &+ K \int_u^t E |\mathcal{D}_u y(\lfloor v \rfloor)|^{2p} \, dv. \end{aligned}$$

Hence

$$\begin{aligned} \sup_{u \leq t \leq t'} E |\mathcal{D}_u y(t)|^{2p} &\leq K \left(1 + E \|\eta\|_C^{2p} + \sup_{\sigma - \tau \leq s \leq \sigma} E \|\mathcal{D}_s \eta\|_\infty^{2p} \right) \\ &+ K_1 \int_u^{t'} \sup_{u \leq v \leq v'} E |\mathcal{D}_u y(v)|^{2p} \, dv' + K_2 \int_u^{t'} \sup_{u \leq v \leq v'} E |\mathcal{D}_u y(v)|^{2p} \, dv' \\ &+ K_3 \int_u^{t'} \sup_{u \leq v \leq v'} E |\mathcal{D}_u y(v)|^{2p} \, dv' + K_4 \int_u^{t'} \sup_{u \leq v \leq v'} E |\mathcal{D}_u y(v)|^{2p} \, dv'. \end{aligned}$$

Thus

$$\theta(t') \leq K \left(1 + E \|\eta\|_C^{2p} + \sup_{\sigma - \tau \leq s \leq \sigma} E \|\mathcal{D}_s \eta\|_\infty^{2p} \right) + K \int_u^{t'} \theta(v') \, dv', \quad u \leq t' \leq T.$$

By Gronwall's lemma,

$$\theta(t') \leq K \left(1 + E \|\eta\|_C^{2p} + \sup_{\sigma - \tau \leq s \leq \sigma} E \|\mathcal{D}_s \eta\|_\infty^{2p} \right) \exp\{K(t' - u)\}, \quad u \leq t' \leq T.$$

This implies the first assertion (??) of the lemma.

To prove the second assertion (3.10) of the lemma, first linearize (??) (*pathwise*) with respect to any deterministic path $\eta \in C([-\tau, 0], \mathbf{R})$. This gives a.s.:

$$\begin{aligned}
(3.17) \quad Dy(t; \sigma, \eta)(\xi) &= \xi(0) + \int_{\sigma}^t \frac{\partial f}{\partial x_1}(y(\lfloor v \rfloor - \tau_1), y(\lfloor v \rfloor)) Dy(\lfloor v \rfloor - \tau_1; \sigma, \eta)(\xi) dv \\
&\quad + \int_{\sigma}^t \frac{\partial f}{\partial x_2}(y(\lfloor v \rfloor - \tau_1), y(\lfloor v \rfloor)) Dy(\lfloor v \rfloor; \sigma, \eta)(\xi) dv \\
&\quad + \int_{\sigma}^t \frac{\partial g}{\partial x_1}(y(\lfloor v \rfloor - \tau_2), y(\lfloor v \rfloor)) Dy(\lfloor v \rfloor - \tau_2; \sigma, \eta)(\xi) dW(v) \\
&\quad + \int_{\sigma}^t \frac{\partial g}{\partial x_2}(y(\lfloor v \rfloor - \tau_2), y(\lfloor v \rfloor)) Dy(\lfloor v \rfloor; \sigma, \eta)(\xi) dW(v), \quad t > \sigma, \\
Dy(t; \sigma, \eta)(\xi) &= \xi(t - \sigma), \quad \sigma - \tau < t < \sigma,
\end{aligned}$$

for all bounded measurable functions $\xi : [-\tau, 0] \rightarrow \mathbf{R}$. Secondly, replace the deterministic path η in the above integral equation by a Malliavin smooth random initial process (also denoted by) $\eta \in L^{4p}(\Omega, C([-\tau, 0], \mathbf{R}); \mathcal{F}_{\sigma})$ satisfying the hypotheses of the lemma. Thirdly, take Malliavin derivatives \mathcal{D}_u of the resulting equation to get the following integral equation for $\mathcal{D}_u Dy(t; \sigma, \eta)(\xi)$:

$$\begin{aligned}
(3.18) \quad \mathcal{D}_u Dy(t; \sigma, \eta)(\xi) &= \frac{\partial g}{\partial x_1}(y(\lfloor u \rfloor - \tau_2), y(\lfloor u \rfloor)) Dy(\lfloor u \rfloor - \tau_2; \sigma, \eta)(\xi) \\
&\quad + \frac{\partial g}{\partial x_2}(y(\lfloor u \rfloor - \tau_2), y(\lfloor u \rfloor)) Dy(\lfloor u \rfloor; \sigma, \eta)(\xi) \\
&\quad + \int_u^t \frac{\partial f}{\partial x_1}(y(\lfloor v \rfloor - \tau_1), y(\lfloor v \rfloor)) \mathcal{D}_u Dy(\lfloor v \rfloor - \tau_1; \sigma, \eta)(\xi) dv \\
&\quad + \int_u^t \frac{\partial f}{\partial x_2}(y(\lfloor v \rfloor - \tau_1), y(\lfloor v \rfloor)) \mathcal{D}_u Dy(\lfloor v \rfloor; \sigma, \eta)(\xi) dv \\
&\quad + \int_u^t \frac{\partial^2 f}{\partial x_1^2}(y(\lfloor v \rfloor - \tau_1), y(\lfloor v \rfloor)) \mathcal{D}_u y(\lfloor v \rfloor - \tau_1; \sigma, \eta) Dy(\lfloor v \rfloor - \tau_1; \sigma, \eta)(\xi) dv \\
&\quad \quad + \int_u^t \frac{\partial^2 f}{\partial x_1 \partial x_2}(y(\lfloor v \rfloor - \tau_1), y(\lfloor v \rfloor)) \mathcal{D}_u y(\lfloor v \rfloor; \sigma, \eta) Dy(\lfloor v \rfloor - \tau_1; \sigma, \eta)(\xi) dv \\
&\quad + \int_u^t \frac{\partial^2 f}{\partial x_2^2}(y(\lfloor v \rfloor - \tau_1), y(\lfloor v \rfloor)) \mathcal{D}_u y(\lfloor v \rfloor; \sigma, \eta) Dy(\lfloor v \rfloor; \sigma, \eta)(\xi) dv \\
&\quad \quad + \int_u^t \frac{\partial^2 f}{\partial x_1 \partial x_2}(y(\lfloor v \rfloor - \tau_1), y(\lfloor v \rfloor)) \mathcal{D}_u y(\lfloor v \rfloor - \tau_1; \sigma, \eta) Dy(\lfloor v \rfloor; \sigma, \eta)(\xi) dv \\
&\quad + \int_u^t \frac{\partial^2 g}{\partial x_1^2}(y(\lfloor v \rfloor - \tau_2), y(\lfloor v \rfloor)) \mathcal{D}_u y(\lfloor v \rfloor - \tau_2; \sigma, \eta) Dy(\lfloor v \rfloor - \tau_2; \sigma, \eta)(\xi) dv \\
&\quad \quad + \int_u^t \frac{\partial^2 g}{\partial x_1 \partial x_2}(y(\lfloor v \rfloor - \tau_2), y(\lfloor v \rfloor)) \mathcal{D}_u y(\lfloor v \rfloor; \sigma, \eta) Dy(\lfloor v \rfloor - \tau_2; \sigma, \eta)(\xi) dW(v) \\
&\quad \quad + \int_u^t \frac{\partial^2 g}{\partial x_2^2}(y(\lfloor v \rfloor - \tau_2), y(\lfloor v \rfloor)) \mathcal{D}_u y(\lfloor v \rfloor; \sigma, \eta) Dy(\lfloor v \rfloor; \sigma, \eta)(\xi) dW(v) \\
&\quad \quad + \int_u^t \frac{\partial g}{\partial x_1}(y(\lfloor v \rfloor - \tau_2), y(\lfloor v \rfloor)) \mathcal{D}_u Dy(\lfloor v \rfloor - \tau_2; \sigma, \eta)(\xi) dW(v) \\
&\quad \quad + \int_u^t \frac{\partial^2 g}{\partial x_1 \partial x_2}(y(\lfloor v \rfloor - \tau_2), y(\lfloor v \rfloor)) \mathcal{D}_u y(\lfloor v \rfloor - \tau_2; \sigma, \eta) Dy(\lfloor v \rfloor; \sigma, \eta)(\xi) dW(v) \\
&\quad \quad + \int_u^t \frac{\partial g}{\partial x_2}(y(\lfloor v \rfloor - \tau_2), y(\lfloor v \rfloor)) \mathcal{D}_u Dy(\lfloor v \rfloor; \sigma, \eta)(\xi) dW(v), \quad t > u > \sigma, \\
\mathcal{D}_u Dy(t; \sigma, \eta)(\xi) &= 0, \quad \sigma - \tau < t < \sigma,
\end{aligned}$$

for all bounded measurable functions $\xi : [-\tau, 0] \rightarrow \mathbf{R}$.

Since f and g have bounded first derivatives, it follows from (??) (with random η) and Gronwall's lemma that there is a positive constant K such that

$$(3.19) \quad \sup_{\xi \in L^\infty([- \tau, 0], \mathbf{R})} \sup_{\|\xi\|_\infty \leq 1} \sup_{\sigma - \tau \leq t \leq T} E|Dy(t; \sigma, \eta)(\xi)|^{2p} \leq K$$

for every $p \geq 1$. By the first assertion of this lemma, we have

$$(3.20) \quad \sup_{0 \leq \sigma \leq T} \sup_{\sigma - \tau \leq u, t \leq T} E|\mathcal{D}_u y(t; \sigma, \eta)|^{2p} \leq K \left(1 + E\|\eta\|_C^{2p} + \sup_{\sigma - \tau \leq s \leq \sigma} E\|\mathcal{D}_s \eta\|_\infty^{2p} \right)$$

for any $p \geq 1$, where $K > 0$ is independent of η .

Let $t > u \geq \sigma$ and assume that $\eta \in L^{4p}(\Omega, C([- \tau, 0], \mathbf{R}); \mathcal{F}_\sigma)$, is Malliavin smooth and satisfies the hypotheses of the lemma. Take $E|\cdot|^{2p}$ on both sides of (??). Using (??), (??) and the fact that f and g have bounded derivatives, this yields:

$$(3.21) \quad \begin{aligned} & E|\mathcal{D}_u Dy(t; \sigma, \eta)(\xi)|^{2p} \\ & \leq K_1 \|\xi\|_\infty^{2p} + K_2 \left(1 + E\|\eta\|_C^{4p} + \sup_{\sigma - \tau \leq s \leq \sigma} E\|\mathcal{D}_s \eta\|_\infty^{4p} \right)^{1/2} \cdot \|\xi\|_\infty^{2p} \\ & + K_3 \int_u^t E|\mathcal{D}_u Dy([v] - \tau_1; \sigma, \eta)(\xi)|^{2p} dv \\ & \leq K_4 \left(1 + E\|\eta\|_C^{4p} + \sup_{\sigma - \tau \leq s \leq \sigma} E\|\mathcal{D}_s \eta\|_\infty^{4p} \right)^{1/2} \cdot \|\xi\|_\infty^{2p} \\ & + K_3 \int_u^t \sup_{u \leq t' \leq v} E|\mathcal{D}_u Dy(t'; \sigma, \eta)(\xi)|^{2p} dv \end{aligned}$$

for $t \geq u \geq \sigma$ and $\xi \in L^\infty([- \tau, 0], \mathbf{R})$. Define

$$\Phi(t') := \sup_{u \leq t \leq t'} E|\mathcal{D}_u Dy(t; \sigma, \eta)(\xi)|^{2p}, \quad t' \geq u.$$

Then (??) says that

$$\Phi(t') \leq K_4 \left(1 + E\|\eta\|_C^{4p} + \sup_{\sigma - \tau \leq s \leq \sigma} E\|\mathcal{D}_s \eta\|_\infty^{4p} \right)^{1/2} \cdot \|\xi\|_\infty^{2p} + K_3 \int_u^{t'} \Phi(v) dv, \quad t' \geq u.$$

By Gronwall's lemma, it follows that

$$\Phi(t') \leq K_4 \left(1 + E\|\eta\|_C^{4p} + \sup_{\sigma - \tau \leq s \leq \sigma} E\|\mathcal{D}_s \eta\|_\infty^{4p} \right)^{1/2} \|\xi\|_\infty^{2p} \exp\{K_3(t' - u)\}, \quad t' \geq u.$$

Thus

$$\sup_{\sigma - \tau \leq u, t \leq T} E|\mathcal{D}_u Dy(t; \sigma, \eta)(\xi)|^{2p} < K_5 \left(1 + E\|\eta\|_C^{4p} + \sup_{\sigma - \tau \leq s \leq \sigma} E\|\mathcal{D}_s \eta\|_\infty^{4p} \right)^{1/2} \cdot \|\xi\|_\infty^{2p}$$

for all Malliavin smooth $\eta \in L^{4p}(\Omega, C([- \tau, 0], \mathbf{R}); \mathcal{F}_\sigma)$ satisfying the hypotheses of the lemma, and all $\xi \in L^\infty([- \tau, 0], \mathbf{R})$. This immediately implies the second assertion (3.10) of the lemma.

The last assertion (3.11) of the lemma follows by very similar argument to the proof of (3.9). Details are left to the reader. This completes the proof of the lemma. \square

Lemma 3.5. *Suppose $f, g \in C_b^3$. Then for any $p \geq 1$,*

$$(3.22) \quad \begin{aligned} & \sup_{0 \leq \sigma \leq T} \sup_{\sigma - \tau \leq u, w, t \leq T} E|\mathcal{D}_w \mathcal{D}_u y(t; \sigma, \eta)|^{2p} \\ & \leq K \left(1 + E\|\eta\|_C^{4p} + \sup_{\sigma - \tau \leq s \leq \sigma} E\|\mathcal{D}_s \eta\|_\infty^{4p} + \sup_{\sigma - \tau \leq s_1, s_2 \leq \sigma} E\|\mathcal{D}_{s_1} \mathcal{D}_{s_2} \eta\|_\infty^{4p} \right) \end{aligned}$$

and

$$(3.23) \quad \begin{aligned} & \sup_{\xi \in L^\infty([- \tau, 0], \mathbf{R})} \sup_{0 \leq \sigma \leq T} \sup_{\sigma - \tau \leq u, w, t \leq T} E|\mathcal{D}_w \mathcal{D}_u Dy(t; \sigma, \eta)(\xi)|^{2p} \\ & \leq K \left(1 + E\|\eta\|_C^{8p} + \sup_{\sigma - \tau \leq s \leq \sigma} E\|\mathcal{D}_s \eta\|_\infty^{8p} + \sup_{\sigma - \tau \leq s_1, s_2 \leq \sigma} E\|\mathcal{D}_{s_1} \mathcal{D}_{s_2} \eta\|_\infty^{8p} \right), \end{aligned}$$

for all $\eta \in L^{8p}(\Omega, C([- \tau, 0], \mathbf{R}); \mathcal{F}_\sigma)$ which are Malliavin smooth with the right hand side of (3.23) finite. The positive constant K is independent of $t \in [\sigma - \tau, T]$, $\sigma \in [0, T]$, and η .

Proof. Assume that the random process η satisfies the hypotheses of the lemma. By Proposition 3.3 in [?], it follows that the left hand side of (??) is finite. To complete the proof of (??), we take Malliavin derivatives \mathcal{D}_w on both sides of the integral equation (??). This yields the following integral equation for $\mathcal{D}_w \mathcal{D}_u y$:

$$\begin{aligned}
\mathcal{D}_w \mathcal{D}_u y(t; \sigma, \eta) &= \frac{\partial g}{\partial x_1}(y(\lfloor u \rfloor - \tau_2), y(\lfloor u \rfloor)) \mathcal{D}_w y(\lfloor u \rfloor - \tau_2; \sigma, \eta) \\
&+ \frac{\partial g}{\partial x_2}(y(\lfloor u \rfloor - \tau_2), y(\lfloor u \rfloor)) \mathcal{D}_w y(\lfloor u \rfloor; \sigma, \eta) \\
&+ \int_u^t \frac{\partial f}{\partial x_1}(y(\lfloor v \rfloor - \tau_1), y(\lfloor v \rfloor)) \mathcal{D}_w \mathcal{D}_u y(\lfloor v \rfloor - \tau_1; \sigma, \eta) \, dv \\
&+ \int_u^t \frac{\partial f}{\partial x_2}(y(\lfloor v \rfloor - \tau_1), y(\lfloor v \rfloor)) \mathcal{D}_w \mathcal{D}_u y(\lfloor v \rfloor; \sigma, \eta) \, dv \\
&+ \int_u^t \frac{\partial^2 f}{\partial x_1^2}(y(\lfloor v \rfloor - \tau_1), y(\lfloor v \rfloor)) \mathcal{D}_w y(\lfloor v \rfloor - \tau_1; \sigma, \eta) \mathcal{D}_u y(\lfloor v \rfloor - \tau_1; \sigma, \eta) \, dv \\
&+ \int_u^t \frac{\partial^2 f}{\partial x_1 \partial x_2}(y(\lfloor v \rfloor - \tau_1), y(\lfloor v \rfloor)) \mathcal{D}_w y(\lfloor v \rfloor; \sigma, \eta) \mathcal{D}_u y(\lfloor v \rfloor - \tau_1; \sigma, \eta) \, dv \\
&+ \int_u^t \frac{\partial^2 f}{\partial x_2^2}(y(\lfloor v \rfloor - \tau_1), y(\lfloor v \rfloor)) \mathcal{D}_w y(\lfloor v \rfloor; \sigma, \eta) \mathcal{D}_u y(\lfloor v \rfloor; \sigma, \eta) \, dv \\
&+ \int_u^t \frac{\partial^2 f}{\partial x_1 \partial x_2}(y(\lfloor v \rfloor - \tau_1), y(\lfloor v \rfloor)) \mathcal{D}_w y(\lfloor v \rfloor - \tau_1; \sigma, \eta) \mathcal{D}_u y(\lfloor v \rfloor; \sigma, \eta) \, dv \\
&+ \int_u^t \frac{\partial^2 g}{\partial x_1^2}(y(\lfloor v \rfloor - \tau_2), y(\lfloor v \rfloor)) \mathcal{D}_w y(\lfloor v \rfloor - \tau_2; \sigma, \eta) \mathcal{D}_u y(\lfloor v \rfloor - \tau_2; \sigma, \eta) \, dv \\
&+ \int_u^t \frac{\partial^2 g}{\partial x_1 \partial x_2}(y(\lfloor v \rfloor - \tau_2), y(\lfloor v \rfloor)) \mathcal{D}_w y(\lfloor v \rfloor; \sigma, \eta) \mathcal{D}_u y(\lfloor v \rfloor - \tau_2; \sigma, \eta) \, dW(v) \\
&+ \int_u^t \frac{\partial^2 g}{\partial x_2^2}(y(\lfloor v \rfloor - \tau_2), y(\lfloor v \rfloor)) \mathcal{D}_w y(\lfloor v \rfloor; \sigma, \eta) \mathcal{D}_u y(\lfloor v \rfloor; \sigma, \eta) \, dW(v) \\
&+ \int_u^t \frac{\partial g}{\partial x_1}(y(\lfloor v \rfloor - \tau_2), y(\lfloor v \rfloor)) \mathcal{D}_w \mathcal{D}_u y(\lfloor v \rfloor - \tau_2; \sigma, \eta) \, dW(v) \\
&+ \int_u^t \frac{\partial^2 g}{\partial x_1 \partial x_2}(y(\lfloor v \rfloor - \tau_2), y(\lfloor v \rfloor)) \mathcal{D}_w y(\lfloor v \rfloor - \tau_2; \sigma, \eta) \mathcal{D}_u y(\lfloor v \rfloor; \sigma, \eta) \, dW(v) \\
&+ \int_u^t \frac{\partial g}{\partial x_2}(y(\lfloor v \rfloor - \tau_2), y(\lfloor v \rfloor)) \mathcal{D}_w \mathcal{D}_u y(\lfloor v \rfloor; \sigma, \eta) \, dW(v), \quad t > u > \sigma \\
\mathcal{D}_w \mathcal{D}_u y(t; \sigma, \eta) &= 0, \quad \sigma - \tau < t < \sigma.
\end{aligned} \tag{3.24}$$

Observe that (??) is a *linear* SDDE in $\mathcal{D}_w \mathcal{D}_u y$. So taking $E|\cdot|^{2p}$ on both sides of (??), using the estimates (??), the fact that f, g are C_b^2 and Gronwall's lemma, a lengthy but straightforward argument yields (??). Details are left to the reader. Note that the estimate (??) requires that f, g be C_b^2 (rather than the stronger requirement that $f, g \in C_b^3$ which is needed for (??)).

The proof of (??) is similar to (but lengthier than) that of (??): Start by taking Malliavin derivatives \mathcal{D}_w on both sides of (??). This yields a linear integral equation for $\mathcal{D}_w \mathcal{D}_u y$. Using the fact that $f, g \in C_b^3$, the estimates (??), (??) and Gronwall's lemma in the latter integral equation, one obtains (??). \square

Proof of Theorem ??. Let $t \in [\sigma, T]$ and $\pi := \{t_{-M} < t_{-M+1} < \dots < t_{-1} < t_0 = 0 < t_1 < t_2 \dots < t_{N-1} < t_N = T\}$ be a partition of $[-\tau, T]$. Without loss of generality assume that $\sigma = 0$ and $t = t_n \in \pi$ for some $0 \leq n \leq N$. Suppose $\eta \in H^{1, \infty}([-\tau, 0], \mathbf{R})$.

Using the Markov property for the segments x_t and y_t ([?], [?]), we may rewrite

$$\begin{aligned}
& E\phi(x(t_n; 0, \eta)) - E\phi(y(t_n; 0, \eta)) \\
&= E\phi(y(t_n; t_n, x_{t_n}(\cdot; 0, \eta))) - E\phi(y(t_n; 0, \eta)) \\
&= \sum_{i=1}^n \{E\phi(y(t_n; t_i, x_{t_i}(\cdot; 0, \eta))) - E\phi(y(t_n; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta)))\} \\
&= \sum_{i=1}^n \{E\phi(y(t_n; t_i, x_{t_i}(\cdot; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta)))) \\
&\quad - E\phi(y(t_n; t_i, y_{t_i}(\cdot; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta))))\} \\
&= \sum_{i=1}^n E \int_0^1 D(\phi \circ y)(t_n; t_i, \lambda x_{t_i}(\cdot; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta)) \\
&\quad + (1 - \lambda)y_{t_i}(\cdot; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta))) \, d\lambda \\
&\quad \cdot [x_{t_i}(\cdot; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta)) - y_{t_i}(\cdot; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta))].
\end{aligned}$$

Our main objective is to show that each of the terms in the above sum is $O((t_i - t_{i-1})^2)$. The rest of the computations in this proof are directed towards this purpose.

In view of Lemma 3.2 and the chain rule for Fréchet derivatives, observe that each expression

$$\{D(\phi \circ y)(t_n; t_i, \lambda x_{t_i}(\cdot; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta)) + (1 - \lambda)y_{t_i}(\cdot; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta)))\}$$

corresponds to a purely atomic random measure on $[-\tau, 0]$. We will denote each such measure simply by

$$D(\phi \circ y)_i(\lambda, ds)$$

for each $\lambda \in [0, 1]$.

To further simplify the notation, we denote

$$(3.25) \quad x_{t_i} := x_{t_i}(\cdot; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta)), \quad y_{t_i} := y_{t_i}(\cdot; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta))$$

for the rest of this proof.

Using (??) of Lemma (??), where we have applied the tame Itô-formula, we obtain

$$(3.26) \quad E\phi(x(t_n; 0, \eta)) - E\phi(y(t_n; 0, \eta)) = \sum_{i=1}^n \sum_{j=1}^{10} E \int_0^1 \int_{-\tau}^0 \{D(\phi \circ y)_i(\lambda, ds) \Lambda_j^i(s)\} \, d\lambda.$$

Thus, by Fubini's theorem, we obtain

$$(3.27) \quad E\phi(x(t_n; 0, \eta)) - E\phi(y(t_n; 0, \eta)) = \sum_{j=1}^{10} \sum_{i=1}^n \int_0^1 \int_{-\tau}^0 E D(\phi \circ y)_i(\lambda, ds) \Lambda_j^i(s) \, d\lambda.$$

We now show how to estimate each of the 10 terms

$$\sum_{i=1}^n \int_0^1 \int_{-\tau}^0 E \{D(\phi \circ y)_i(\lambda, ds) \Lambda_j^i(s)\} \, d\lambda, \quad j = 1, 2, \dots, 10,$$

on the right hand side of (??), for any fixed $\lambda \in [0, 1]$.

Let $j = 10$ and fix any $\lambda \in [0, 1]$. Then by definition of the Skorohod integral (as adjoint of the Malliavin derivative), and using Lemma 3.2, we get

$$\begin{aligned}
I_{10}^i &:= \int_{-\tau}^0 E \{ D(\phi \circ y)_i(\lambda, ds) \Lambda_{10}^i(s) \} \\
&= \frac{1}{2} \int_{-\tau}^0 E \left\{ D(\phi \circ y)_i(\lambda, ds) \int_{t_{i-1}}^{(t_i+s) \vee t_{i-1}} \left(\int_{[u]}^u \frac{\partial^2 g}{\partial x_2^2}(x(v - \tau_2), x(v)) \right. \right. \\
&\quad \left. \left. \times g(x(v - \tau_2), x(v))^2 dv \right) dW(u) \right\} \\
&= \frac{1}{2} \sum_{m=1}^l \int_{t_{i-1}}^{(t_i+\mu_m) \vee t_{i-1}} E \left\{ \mathcal{D}_u [D(\phi \circ y)_i(\lambda, 1_{\{\mu_m\}})] \int_{[u]}^u \frac{\partial^2 g}{\partial x_2^2}(x(v - \tau_2), x(v)) \right. \\
&\quad \left. \times g(x(v - \tau_2), x(v))^2 dv \right\} du.
\end{aligned}$$

In the above formula, $1_{\{\mu_m\}}$ denotes the indicator function of $\{\mu_m\}$ for $1 \leq m \leq l$, where the μ_m 's are defined in Lemma 3.2.

Since f and g are C_b^2 , then $\sup_{0 \leq v \leq T} E|x(v)|^{2p} < \infty$ for all $p > 0$. Thus

$$\sup_{0 \leq v \leq T} E \left| \frac{\partial^2 g}{\partial x_2^2}(x(v - \tau_2), x(v)) g(x(v - \tau_2), x(v))^2 \right|^2 < \infty.$$

Recall the notation (3.25). Then by the chain rule for Fréchet derivatives and the product rule for Malliavin derivatives, we obtain

$$\begin{aligned}
&\mathcal{D}_u [D(\phi \circ y)_i(\lambda, \cdot)] \\
&= \mathcal{D}_u \{ D\phi(y(t_n; t_i, \lambda x_{t_i} + (1 - \lambda)y_{t_i})) Dy(t_n; t_i, \lambda x_{t_i} + (1 - \lambda)y_{t_i}) \} \\
&= \mathcal{D}_u \{ D\phi(y(t_n; t_i, \lambda x_{t_i} + (1 - \lambda)y_{t_i})) \} Dy(t_n; t_i, \lambda x_{t_i} + (1 - \lambda)y_{t_i}) \\
&\quad + D\phi(y(t_n; t_i, \lambda x_{t_i} + (1 - \lambda)y_{t_i})) \cdot \mathcal{D}_u \{ Dy(t_n; t_i, \lambda x_{t_i} + (1 - \lambda)y_{t_i}) \}.
\end{aligned}$$

By the chain rule for Malliavin derivatives [?],

$$\begin{aligned}
&\mathcal{D}_u [D(\phi \circ y)_i(\lambda, \cdot)] \\
&= D^2\phi(y(t_n; t_i, \lambda x_{t_i} + (1 - \lambda)y_{t_i})) \cdot \mathcal{D}_u y(t_n; t_i, \lambda x_{t_i} + (1 - \lambda)y_{t_i}) \\
&\quad \cdot Dy(t_n; t_i, \lambda x_{t_i} + (1 - \lambda)y_{t_i})(\cdot) \\
&\quad + D\phi(y(t_n; t_i, \lambda x_{t_i} + (1 - \lambda)y_{t_i})) \cdot \mathcal{D}_u \{ Dy(t_n; t_i, \lambda x_{t_i} + (1 - \lambda)y_{t_i}) \}.
\end{aligned}$$

(Note that $Dy(t_n; t_i, \lambda x_{t_i} + (1 - \lambda)y_{t_i})$ stands for the Fréchet derivative $Dy(\cdot \cdot \cdot) \in C([-\tau, 0], \mathbf{R})^*$).

Thus, using the above relation, Lemma 3.2 and its notation, we obtain

$$\begin{aligned}
I_{10}^i &:= \int_{-\tau}^0 E D(\phi \circ y)_i(\lambda, ds) \Lambda_{10}^i(s) \\
&= \int_{-\tau}^0 E D\phi(y(t_n; t_i, \lambda x_{t_i} + (1 - \lambda)y_{t_i})) \\
&\quad Dy(t_n; t_i, \lambda x_{t_i} + (1 - \lambda)y_{t_i})(ds) \Lambda_{10}^i(s) \\
&= E D\phi(y(t_n; t_i, \lambda x_{t_i} + (1 - \lambda)y_{t_i})) \sum_{m=1}^l \partial_m F(t, W(t), W(t_2), \dots, \\
&\quad W(t_k), s_1, \dots, s_h, \dots, \lambda x(t_i + \mu_m) + (1 - \lambda)y(t_i + \mu_m), \dots, \dots) \Lambda_{10}^i(\mu_m) \\
&= \sum_{m=1}^l E D\phi(y(t_n; t_i, \lambda x_{t_i} + (1 - \lambda)y_{t_i})) \partial_m F(t, W(t), \dots, \lambda x(t_i + \mu_m) \\
&\quad + (1 - \lambda)y(t_i + \mu_m), \dots) \Lambda_{10}^i(\mu_m).
\end{aligned}$$

Recall that F and $\mu_m, m = 1, 2, \dots, l$, in the above relation are defined as in Lemma 3.2.

By the definition of the Skorohod integral as adjoint of the Malliavin derivative, we may write each summand in the above relation as

$$\begin{aligned}
& E\{D\phi(y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i}))\partial_m F(t, W(t), \dots, \lambda x(t_i + \mu_m) \\
& \quad + (1-\lambda)y(t_i + \mu_m); \dots)\Lambda_{10}^i(\mu_m)\} \\
(3.28) \quad & = \frac{1}{2}E\{D\phi(y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i}))\partial_m F(t, W(t), \dots, \lambda x(t_i + \mu_m) \\
& \quad + (1-\lambda)y(t_i + \mu_m); \dots) \\
& \quad \times \int_{t_{i-1}}^{(t_i + \mu_m) \vee t_{i-1}} \int_{[u]}^u \frac{\partial^2 g}{\partial x_2^2}(x(v - \tau_2), x(v))g(x(v - \tau_2), x(v))^2 dv dW(u)\} \\
& = R_{1,m} + R_{2,m}
\end{aligned}$$

where

$$\begin{aligned}
(3.29) \quad R_{1,m} & := \frac{1}{2} \int_{t_{i-1}}^{(t_i + \mu_m) \vee t_{i-1}} E\left\{ \mathcal{D}_u D\phi(y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i}))\partial_m F(t, W(t), \dots, \lambda x(t_i + \mu_m) \right. \\
& \quad \left. + (1-\lambda)y(t_i + \mu_m); \dots) \times \int_{[u]}^u \frac{\partial^2 g}{\partial x_2^2}(x(v - \tau_2), x(v))g(x(v - \tau_2), x(v))^2 dv \right\} du
\end{aligned}$$

and

$$\begin{aligned}
(3.30) \quad R_{2,m} & := \frac{1}{2} \int_{t_{i-1}}^{(t_i + \mu_m) \vee t_{i-1}} E\left\{ D\phi(y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i}))\mathcal{D}_u \partial_m F(t, W(t), \dots, \lambda x(t_i + \mu_m) \right. \\
& \quad \left. + (1-\lambda)y(t_i + \mu_m); \dots) \times \int_{[u]}^u \frac{\partial^2 g}{\partial x_2^2}(x(v - \tau_2), x(v))g(x(v - \tau_2), x(v))^2 dv \right\} du.
\end{aligned}$$

Using the chain rule for Malliavin derivatives, and Lemma 3.4, we obtain

$$\begin{aligned}
& E|\mathcal{D}_u D\phi(y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i}))|^{2p} \\
& = E|D^2\phi(y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i}))\mathcal{D}_u y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})|^{2p} \\
& \leq K_1 E|\mathcal{D}_u y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})|^{2p} \\
& \leq K_2 \left(1 + E\|x_{t_i}\|_C^{2p} + E\|y_{t_i}\|_C^{2p} + \sup_{\sigma - \tau \leq s \leq \sigma} E\|\mathcal{D}_s x_{t_i}\|_\infty^{2p} + \sup_{\sigma - \tau \leq s \leq \sigma} E\|\mathcal{D}_s y_{t_i}\|_\infty^{2p} \right).
\end{aligned}$$

Since x and y both satisfy SDDs with coefficients having linear growth, then

$$E\|x_{t_i}(\cdot; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta))\|_C^{2p} = E\|x_{t_i}(\cdot, 0, \eta)\|_C^{2p} \leq K_3(1 + \|\eta\|_C^{2p})$$

and

$$E\|y_{t_i}(\cdot; t_{i-1}, x_{t_{i-1}}(\cdot; 0, \eta))\|_C^{2p} \leq K_4(1 + E\|x_{t_{i-1}}(\cdot, 0, \eta)\|_C^{2p}) \leq K_5(1 + \|\eta\|_C^{2p}).$$

Similar estimates hold for $\mathcal{D}_s x_{t_i}$ and $\mathcal{D}_s y_{t_i}$. Therefore,

$$(3.31) \quad E|\mathcal{D}_u D\phi(y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i}))|^{2p} \leq K_6(1 + \|\eta\|_C^{2p})$$

for all $0 \leq i \leq N$. The constants K_j , $j = 1, \dots, 6$, are independent of the partition π .

Now, using the tame representation of Dy in Lemma 3.2, we have

$$\begin{aligned}
(3.32) \quad \sum_{m=1}^l R_{1,m} & = \int_{t_{i-1}}^{t_i} EX(u) \cdot \sum_{m=1}^l \partial_m F(t, W(t), \dots, \lambda x(t_i + \mu_m) \\
& \quad + (1-\lambda)y(t_i + \mu_m); \dots)\xi^u(\mu_m) du \\
& = \int_{t_{i-1}}^{t_i} EX(u) Dy(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})(\xi^u) du
\end{aligned}$$

where

$$(3.33) \quad X(u) := \frac{1}{2} \mathcal{D}_u D\phi(y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})) \int_{[u]}^u \frac{\partial^2 g}{\partial x_2^2}(x(v - \tau_2), x(v))g(x(v - \tau_2), x(v))^2 dv,$$

for all $u \in [0, T]$, and $\xi^u \in L^\infty([-\tau, 0], \mathbf{R})$ is given by

$$\xi^u(s) := 1_{[t_{i-1}, (t_i+s) \vee t_{i-1}]}(u), \quad s \in [-\tau, 0], \quad u \in [0, T].$$

Using (3.19), we get

$$(3.34) \quad E|Dy(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})(\xi^u)|^2 \leq K$$

where K is a positive constant independent of η, u and the partition π .

Using the definition of $X(u)$ in (3.33), relation (3.31), the boundedness of $\frac{\partial^2 g}{\partial x_2^2}$, and the linear growth property of g , we obtain

$$(3.35) \quad \begin{aligned} E|X(u)|^2 &\leq K(1 + \|\eta\|_C^4)^{1/2} \\ &\quad \cdot \left\{ E \left| \int_{[u]}^u \left| \frac{\partial^2 g}{\partial x_2^2}(x(v-\tau_2), x(v)) \right| g(x(v-\tau_2), x(v))^2 dv \right|^4 \right\}^{1/2} \\ &\leq K(1 + \|\eta\|_C^2) \left\{ E \left| \int_{[u]}^u (1 + |x(v-\tau_2)|^2 + |x(v)|^2) dv \right|^4 \right\}^{1/2} \\ &\leq K(1 + \|\eta\|_C^2)(u - [u]) \left\{ E \left(\int_{[u]}^u (1 + |x(v-\tau_2)|^4 + |x(v)|^4) dv \right)^2 \right\}^{1/2} \\ &\leq K(1 + \|\eta\|_C^2)(u - [u])^{3/2} \left\{ \int_{[u]}^u (1 + E|x(v-\tau_2)|^8 + E|x(v)|^8) dv \right\}^{1/2} \\ &\leq K(1 + \|\eta\|_C^2)(1 + \|\eta\|_C^8)^{1/2}(u - [u])^2 \\ &\leq K(1 + \|\eta\|_C^6)(u - [u])^2 \end{aligned}$$

for all $u \in [0, T]$. In the above inequalities, and throughout this proof, K stands for a generic constant which may change from line to line, and is independent of $\eta \in C([-\tau, 0], \mathbf{R})$ and the partition π . Combining (3.32), (3.34) and (3.35), gives

$$(3.36) \quad \begin{aligned} \left| \sum_{m=1}^l R_{1,m} \right| &\leq \int_{t_{i-1}}^{t_i} [E|X(u)|^2]^{1/2} \left\{ E|Dy(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})(\xi^u)|^2 \right\}^{1/2} du \\ &\leq K(1 + \|\eta\|_C^3) \int_{t_{i-1}}^{t_i} (u - t_{i-1}) du \\ &\leq K(1 + \|\eta\|_C^3)(t_i - t_{i-1})^2. \end{aligned}$$

Next, we estimate $\sum_{m=1}^l R_{2,m}$ using (??) and Lemma 3.2. Rewrite the latter sum in the following form:

$$(3.37) \quad \begin{aligned} \sum_{m=1}^l R_{2,m} &:= \int_{t_{i-1}}^{t_i} E \sum_{m=1}^l \mathcal{D}_u \partial_m F(t, W(t), \dots, \lambda x(t_i + \mu_m) + (1-\lambda)y(t_i + \mu_m); \dots) \times \\ &\quad \left\{ 1_{[t_{i-1}, (t_i + \mu_m) \vee t_{i-1}]}(u) \cdot D\phi(y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})) \times \right. \\ &\quad \left. \int_{[u]}^u \frac{\partial^2 g}{\partial x_2^2}(x(v-\tau_2), x(v)) g(x(v-\tau_2), x(v))^2 dv \right\} du \\ &= \int_{t_{i-1}}^{t_i} E \mathcal{D}_u Dy(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})(\xi^u) Y(u) du \end{aligned}$$

where

$$(3.38) \quad Y(u) := D\phi(y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})) \int_{[u]}^u \frac{\partial^2 g}{\partial x_2^2}(x(v-\tau_2), x(v)) g(x(v-\tau_2), x(v))^2 dv$$

and, as before,

$$(3.39) \quad \xi^u(s) := 1_{[t_{i-1}, (t_i+s) \vee t_{i-1}]}(u), \quad s \in [-\tau, 0],$$

for all $u \in [0, T]$.

Since $D\phi$ and Dg are globally bounded, a similar computation to (??) gives

$$(3.40) \quad E|Y(u)|^2 \leq K(1 + \|\eta\|_C^4)(u - [u])^2, \quad u \in [0, T].$$

From (??), (??), (??) and Lemma ??, we get

$$\begin{aligned}
(3.41) \quad \left| \sum_{m=1}^l R_{2,m} \right| &\leq \int_{t_{i-1}}^{t_i} \{E|\mathcal{D}_u D y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})(\xi^u)|^2\}^{1/2} \cdot \{E|Y(u)|^2\}^{1/2} du \\
&\leq K(1 + \|\eta\|_C^4)^{1/2} (1 + E\|\eta\|_C^4)^{1/2} \int_{t_{i-1}}^{t_i} (u - \lfloor u \rfloor) du \\
&\leq K(1 + \|\eta\|_C^4) \int_{t_{i-1}}^{t_i} (u - t_{i-1}) du \\
&\leq K(1 + \|\eta\|_C^4)(t_i - t_{i-1})^2.
\end{aligned}$$

In the above inequalities, K is a generic positive constant which may change from line to line and is independent of η and the partition π . Using (??), (??) and (??), it follows that

$$\begin{aligned}
(3.42) \quad \left| \sum_{i=1}^n I_{10}^i \right| &= \left| \sum_{i=1}^n \int_0^1 \int_{-\tau}^0 E\{D(\phi \circ y)_i(\lambda, ds) \Lambda_{10}^i(s)\} d\lambda \right| \\
&\leq K(1 + \|\eta\|_C^3) \sum_{i=1}^n (t_i - t_{i-1})^2 + K(1 + \|\eta\|_C^4) \sum_{i=1}^n (t_i - t_{i-1})^2 \\
&\leq K((1 + \|\eta\|_C^4) |\pi|).
\end{aligned}$$

Our next task is to develop estimates similar to (??) for the cases $1 \leq j \leq 9$. As a sample computation, we will examine in detail the case $j = 7$. The rest of the computations are left to the reader. They are similar either to the case $j = 7$ or $j = 10$.

Consider the case $j = 7$:

$$I_7^i := \int_{-\tau}^0 E\{D(\phi \circ y)_i(\lambda, ds) \Lambda_7^i(s)\},$$

where (from the notation in Lemma ??)

$$\begin{aligned}
(3.43) \quad \Lambda_7^i(s) &:= \int_{t_{i-1}}^{(t_i+s) \vee t_{i-1}} \int_{\lfloor u \rfloor}^u \frac{\partial g}{\partial x_1}(x(v - \tau_2), x(v)) g(x(v - 2\tau_2), x(v - \tau_2)) \\
&\quad \times 1_{[\tau_2, \infty)}(v) dW(v - \tau_2) dW(u)
\end{aligned}$$

for all $s \in [-\tau, 0]$. For simplicity, denote

$$(3.44) \quad h(v) := \frac{\partial g}{\partial x_1}(x(v - \tau_2), x(v)) g(x(v - 2\tau_2), x(v - \tau_2)) 1_{[\tau_2, \infty)}(v), \quad v \geq 0.$$

Then, by the chain rule for Fréchet derivatives and Lemma 3.2, we get

$$\begin{aligned}
(3.45) \quad I_7^i &= \int_{-\tau}^0 E\{D(\phi \circ y)_i(\lambda, ds) \int_{t_{i-1}}^{(t_i+s) \vee t_{i-1}} \int_{\lfloor u \rfloor}^u h(v) dW(v - \tau_2) dW(u)\} \\
&= \sum_{m=1}^l E\{D\phi(y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})) \partial_m F(t, W(t), \dots, \lambda x(t_i + \mu_m) \\
&\quad + (1-\lambda)y(t_i + \mu_m); \dots) \\
&\quad \times \int_{t_{i-1}}^{(t_i+\mu_m) \vee t_{i-1}} \int_{\lfloor u \rfloor}^u h(v) dW(v - \tau_2) dW(u)\} \\
&= \sum_{m=1}^l (T_{1,m} + T_{2,m}),
\end{aligned}$$

where

$$\begin{aligned}
(3.46) \quad T_{1,m} &:= \int_{t_{i-1}}^{(t_i+\mu_m) \vee t_{i-1}} E\{\mathcal{D}_u [D\phi(y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i}))] \partial_m F(t, W(t), \dots, \lambda x(t_i + \mu_m) \\
&\quad + (1-\lambda)y(t_i + \mu_m); \dots) \int_{\lfloor u \rfloor}^u h(v) dW(v - \tau_2)\} du
\end{aligned}$$

and

$$(3.47) \quad T_{2,m} := \int_{t_{i-1}}^{(t_i+\mu_m)\vee t_{i-1}} E \left\{ D\phi(y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})) \mathcal{D}_u [\partial_m F(t, W(t), \dots, \lambda x(t_i + \mu_m) + (1-\lambda)y(t_i + \mu_m); \dots)] \int_{[u]}^u h(v) dW(v - \tau_2) \right\} du.$$

The above expressions (??) and (??) are obtained using the definition of the Skorohod integral (as adjoint of the Malliavin differentiation operator) together with the product rule for Malliavin derivatives ([?]).

Again, using the definition of the Skorohod integral once more in (??), yields

$$(3.48) \quad T_{1,m} := \int_{t_{i-1}}^{t_i} \int_{[u]}^u E \left\{ \mathcal{D}_v [\mathcal{D}_u \{ D\phi(y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})) \}] \times \partial_m F(t, W(t), \dots, \lambda x(t_i + \mu_m) + (1-\lambda)y(t_i + \mu_m); \dots)] h(v) \right\} \times 1_{[t_{i-1}, (t_i+\mu_m)\vee t_{i-1}]}(u) dv du$$

and

$$(3.49) \quad T_{2,m} := \int_{t_{i-1}}^{t_i} \int_{[u]}^u E \left\{ \mathcal{D}_v [D\phi(y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})) \times \mathcal{D}_u \partial_m F(t, W(t), \dots, \lambda x(t_i + \mu_m) + (1-\lambda)y(t_i + \mu_m); \dots)] h(v) \right\} \times 1_{[t_{i-1}, (t_i+\mu_m)\vee t_{i-1}]}(u) dv du$$

Using the product and chain rules for Malliavin derivatives, we may rewrite the expression (??) in the form:

$$\begin{aligned} T_{1,m} &:= \int_{t_{i-1}}^{t_i} \int_{[u]}^u E \left\{ \mathcal{D}_v [D^2\phi(y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})) \mathcal{D}_u (y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})) \times \partial_m F(t, W(t), \dots, \lambda x(t_i + \mu_m) + (1-\lambda)y(t_i + \mu_m); \dots)] h(v) \right\} \\ &\quad \times 1_{[t_{i-1}, (t_i+\mu_m)\vee t_{i-1}]}(u) dv du \\ &= \int_{t_{i-1}}^{t_i} \int_{[u]}^u E \left\{ [D^3\phi(y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})) \mathcal{D}_v (y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})) \cdot \mathcal{D}_u (y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})) \partial_m F(t, W(t), \dots, \lambda x(t_i + \mu_m) + (1-\lambda)y(t_i + \mu_m); \dots) + D^2\phi(y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})) \mathcal{D}_v \mathcal{D}_u (y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})) \times \partial_m F(t, W(t), \dots, \lambda x(t_i + \mu_m) + (1-\lambda)y(t_i + \mu_m); \dots) + D^2\phi(y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})) \mathcal{D}_u (y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})) \times \mathcal{D}_v \partial_m F(t, W(t), \dots, \lambda x(t_i + \mu_m) + (1-\lambda)y(t_i + \mu_m); \dots))] h(v) \right\} 1_{[t_{i-1}, (t_i+\mu_m)\vee t_{i-1}]}(u) dv du. \end{aligned}$$

As before, recall the definition of $\xi^u(s)$ in (??): That is

$$\xi^u(s) := 1_{[t_{i-1}, (t_i+s)\vee t_{i-1}]}(u), \quad s \in [-\tau, 0], \quad u \in [0, T].$$

Hence

$$(3.50) \quad \sum_{m=1}^l T_{1,m} = \sum_{m=1}^l T_{1,m}^1 + \sum_{m=1}^l T_{1,m}^2 + \sum_{m=1}^l T_{1,m}^3,$$

where

$$(3.51) \quad \sum_{m=1}^l T_{1,m}^1 := \int_{t_{i-1}}^{t_i} \int_{[u]}^u E \left\{ [D^3\phi(y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})) \mathcal{D}_v y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i}) \times \mathcal{D}_u y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i}) Dy(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i}) (\xi^u)] h(v) \right\} dv du,$$

$$(3.52) \quad \sum_{m=1}^l T_{1,m}^2 := \int_{t_{i-1}}^{t_i} \int_{[u]}^u E \left\{ [D^2\phi(y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})) \mathcal{D}_v \mathcal{D}_u y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i}) \times Dy(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i}) (\xi^u)] h(v) \right\} dv du$$

and

$$(3.53) \quad \sum_{m=1}^l T_{1,m}^3 := \int_{t_{i-1}}^{t_i} \int_{[u]}^u E \left\{ [D^2 \phi(y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})) \mathcal{D}_u y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i}) \right. \\ \left. \times \mathcal{D}_v \mathcal{D}_y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})(\xi^u)] h(v) \right\} dv du.$$

Using the fact that $\phi \in C_b^3$, Lemma ??, its proof ((?)) and (?), it follows that there exists a positive generic constant K such that

$$(3.54) \quad \begin{aligned} & \left| \sum_{m=1}^l T_{1,m}^1 \right| \\ & \leq K \int_{t_{i-1}}^{t_i} \int_{[u]}^u [E |\mathcal{D}_v y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})|^4]^{1/4} \\ & \quad \times [E |\mathcal{D}_u y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})|^4]^{1/4} \\ & \quad \times [E |Dy(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})(\xi^u)|^4]^{1/4} \times [E |h(v)|^4]^{1/4} dv du \\ & \leq K(1 + \|\eta\|_C^4) \int_{t_{i-1}}^{t_i} \int_{[u]}^u dv du \\ & \leq K(1 + \|\eta\|_C^4) \int_{t_{i-1}}^{t_i} (u - t_{i-1}) du \\ & = \frac{1}{2} K(1 + \|\eta\|_C^4) (t_i - t_{i-1})^2. \end{aligned}$$

Similarly, since $\phi \in C_b^2$, it follows from (?), (?) and lemmas 3.4, 3.5, that there is a positive generic constant K such that

$$(3.55) \quad \left| \sum_{m=1}^l T_{1,m}^2 \right| \leq K(1 + \|\eta\|_C^4) (t_i - t_{i-1})^2$$

and

$$(3.56) \quad \left| \sum_{m=1}^l T_{1,m}^3 \right| \leq K(1 + \|\eta\|_C^4) (t_i - t_{i-1})^2.$$

Using (?), (?), (?) and (?), we get

$$(3.57) \quad \begin{aligned} \left| \sum_{i=1}^n \sum_{m=1}^l T_{1,m} \right| & \leq K(1 + \|\eta\|_C^4) \sum_{i=1}^n (t_i - t_{i-1})^2 \\ & \leq K(1 + \|\eta\|_C^4) |\pi|. \end{aligned}$$

Next, from (?), the chain and product rules for Malliavin derivatives, it is easy to see that

$$(3.58) \quad T_{2,m} = T_{2,m}^1 + T_{2,m}^2,$$

where

$$(3.59) \quad \begin{aligned} \sum_{m=1}^l T_{2,m}^1 & = \int_{t_{i-1}}^{t_i} \int_{[u]}^u E \left\{ D^2 \phi(y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})) \mathcal{D}_v y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i}) \right. \\ & \quad \left. \times \mathcal{D}_u \mathcal{D}_y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})(\xi^u) h(v) \right\} dv du \end{aligned}$$

and

$$(3.60) \quad \begin{aligned} \sum_{m=1}^l T_{2,m}^2 & := \int_{t_{i-1}}^{t_i} \int_{[u]}^u E \left\{ D\phi(y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})) \times \right. \\ & \quad \left. \times [\mathcal{D}_v \mathcal{D}_u \mathcal{D}_y(t_n; t_i, \lambda x_{t_i} + (1-\lambda)y_{t_i})(\xi^u)] \cdot h(v) \right\} dv du. \end{aligned}$$

Using (?), (?), lemmas 3.3, 3.4 and the fact that $\phi \in C_b^2$, we obtain the following estimates

$$(3.61) \quad \left| \sum_{m=1}^l T_{2,m}^1 \right| \leq K(1 + \|\eta\|_C^4) (t_i - t_{i-1})^2$$

and

$$(3.62) \quad \left| \sum_{m=1}^l T_{2,m}^2 \right| \leq K(1 + \|\eta\|_C^3)(t_i - t_{i-1})^2,$$

where $K > 0$ is a generic constant independent of η and the the partition π . From (??)-(??), it is easy to see that

$$(3.63) \quad \left| \sum_{i=1}^n \sum_{m=1}^l T_{2,m} \right| \leq K(1 + \|\eta\|_C^4) |\pi|.$$

Now combine (??), (??) and (??) to get

$$(3.64) \quad \left| \sum_{i=1}^n I_7^i \right| \leq K(1 + \|\eta\|_C^4) |\pi|.$$

Estimates of the remaining 8 terms

$$\sum_{i=1}^n \int_0^1 \int_{-\tau}^0 E\{D(\phi \circ y)_i(\lambda, ds) \Lambda_j^i(s)\} d\lambda, \quad j = 1, 2, 3, 4, 5, 6, 8, 9,$$

on the right hand side of (??), may be obtained by very similar arguments to those used for developing (??) and (??). Note that in the cases $j = 3, 8$, one needs to employ moment estimates for the Malliavin derivatives $\mathcal{D}_u x(t)$ given in (3.11) of Lemma 3.4. Thus one obtains a generic constant $K > 0$ and a positive integer q such that

$$(3.65) \quad \left| \sum_{i=1}^n \int_0^1 \int_{-\tau}^0 E\{D(\phi \circ y)_i(\lambda, ds) \Lambda_j^i(s)\} d\lambda \right| \leq K(1 + \|\eta\|_C^q) |\pi|$$

for $j = 2, 3, 4, 5, 8, 9$; and

$$(3.66) \quad \left| \sum_{i=1}^n \int_0^1 \int_{-\tau}^0 E\{D(\phi \circ y)_i(\lambda, ds) \Lambda_j^i(s)\} d\lambda \right| \leq K(1 + \|\eta\|_{1,\infty}^q) |\pi|$$

for $j = 1, 6$. Note that the last two cases involve estimates on the derivatives of the initial process η . As before, the constant K may depend on T, q and the test function ϕ , but is independent of π, η and the choice of $t \in [\sigma, T], \sigma \in [0, T]$. Details of the computations in (??) and (??) are left to the reader.

Putting together the above inequalities, (??), (??) and (??), shows that there is a positive constant K and a positive integer q such that

$$(3.67) \quad |E\phi(x(t; \sigma, \eta)) - E\phi(y(t; \sigma, \eta))| \leq K(1 + \|\eta\|_{1,\infty}^q) |\pi|$$

for all $t \in [\sigma - \tau, T], \sigma \in [0, T]$, and all $\eta \in H^{1,\infty}([-\tau, 0], \mathbf{R})$.

To complete the proof of the weak convergence estimate (3.3), we proceed as follows:

The estimate (3.3) will hold if we can replace η in the Euler approximation $y(t; \sigma, \eta)$ in (3.67) by its piecewise-linear approximation η^π . To do so, we first appeal to the triangle inequality

$$(3.68) \quad |E\phi(x(t; \sigma, \eta)) - E\phi(y(t; \sigma, \eta^\pi))| \leq |E\phi(x(t; \sigma, \eta)) - E\phi(x(t; \sigma, \eta^\pi))| \\ + |E\phi(x(t; \sigma, \eta^\pi)) - E\phi(y(t; \sigma, \eta^\pi))|.$$

Using [Mo.1] (Theorem 3.1, p. 41), we have

$$(3.69) \quad |E\phi(x(t; \sigma, \eta)) - E\phi(x(t; \sigma, \eta^\pi))| \leq K\|\eta - \eta^\pi\|_C.$$

From the definition of η^π and the fundamental theorem of calculus, the following two elementary estimates hold

$$(3.70) \quad \|\eta - \eta^\pi\|_C \leq 2\|\eta'\|_\infty |\pi|$$

$$(3.71) \quad \|\eta^\pi\|_{1,\infty} \leq \|\eta\|_{1,\infty},$$

for all $\eta \in H^{1,\infty}([-\tau, 0], \mathbf{R})$.

Now in (3.67), replace η by η^π and use (3.71) to get

$$(3.72) \quad |E\phi(x(t; \sigma, \eta^\pi)) - E\phi(y(t; \sigma, \eta^\pi))| \leq K(1 + \|\eta\|_{1,\infty}^q) |\pi|.$$

The required estimate (3.3) now follows from (3.68), (3.69), (3.70) and (3.72).

The proof of Theorem ?? is now complete. \square

In this section, we present a generalization of Theorem ???. This generalization covers the case of \mathbf{R}^d -valued stochastic functional differential equations (SFDEs), driven by multi-dimensional Brownian motion and having several discrete delays in the drift and diffusion coefficients, as well as (smooth) quasitame dependence on the history of the solution in all coefficients.

We begin by setting up appropriate notation.

Let $W(t) := (W_1(t), W_2(t), \dots, W_m(t))$, $t \geq 0$, be m -dimensional standard Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$.

Consider a finite number of delays $\{\tau_1^i : 1 \leq i \leq k_1\}$, $\{\tau_2^{j,l} : 1 \leq j \leq k_{2,l}, 1 \leq l \leq m\}$, with maximum delay $\tau := \max\{\tau_1^i, \tau_2^{j,l} : 1 \leq i \leq k_1, 1 \leq j \leq k_{2,l}, 1 \leq l \leq m\}$. We will designate the memory in our SFDE by a collection of tame projections

$$\Pi^1 : C := C([- \tau, \mathbf{R}^d]) \rightarrow \mathbf{R}^{d_1}, \quad \Pi^{2,l} : C \rightarrow \mathbf{R}^{d_{2,l}}$$

$$\Pi^1(\eta) := (\eta(\tau_1^1), \eta(\tau_1^2), \dots, \eta(\tau_1^{k_1})), \quad \Pi^{2,l}(\eta) := (\eta(\tau_2^{1,l}), \eta(\tau_2^{2,l}), \dots, \eta(\tau_2^{k_{2,l},l}))$$

for all $\eta \in C$, and *quasitame projections*

$$\Pi_q^1 : C \rightarrow \mathbf{R}^{d_1^q}, \quad \Pi_q^{2,l} : C \rightarrow \mathbf{R}^{d_{2,l}^q}$$

where $d_1 = k_1 d$, $d_1^q = k_1^q d$, $d_{2,l} = k_{2,l} d$, $d_{2,l}^q = k_{2,l}^q d$ are integer multiples of d , for $1 \leq l \leq m$. The quasitame projections are of the form:

$$\begin{aligned} \Pi_q^1(\eta) &:= \left(\int_{-\tau}^0 \sigma_1^1(\eta(s)) \mu_1^1(s) ds, \int_{-\tau}^0 \sigma_2^1(\eta(s)) \mu_2^2(s) ds, \dots, \int_{-\tau}^0 \sigma_{k_2}^1(\eta(s)) \mu_{k_2}^1(s) ds \right) \\ \Pi_q^{2,l}(\eta) &:= \left(\int_{-\tau}^0 \sigma_1^2(\eta(s)) \mu_1^2(s) ds, \int_{-\tau}^0 \sigma_2^2(\eta(s)) \mu_2^2(s) ds, \dots, \int_{-\tau}^0 \sigma_{k_{2,l}}^2(\eta(s)) \mu_{k_{2,l}}^2(s) ds \right) \end{aligned}$$

for all $\eta \in C$. The functions $\sigma_i^1, \sigma_j^2, \mu_i^1, \mu_j^2$ are smooth.

Let

$$f : \mathbf{R}^+ \times \mathbf{R}^{d_1} \times \mathbf{R}^{d_1^q} \rightarrow \mathbf{R}^d, \quad g_l : \mathbf{R}^+ \times \mathbf{R}^{d_{2,l}} \times \mathbf{R}^{d_{2,l}^q} \rightarrow \mathbf{R}^d, \quad 1 \leq l \leq m,$$

be functions of class C^1 in the first variable and C_b^3 in all space variables.

Consider the SFDE

$$(4.1) \quad x(t) = \eta(0) + \int_0^t f(u, \Pi^1(x_u), \Pi_q^1(x_u)) du + \sum_{l=1}^m \int_0^t g_l(u, \Pi^{2,l}(x_u), \Pi_q^{2,l}(x_u)) dW_l(u),$$

for $\sigma \leq t \leq T$, with initial path

$$x_\sigma = \eta \in H^{1,\infty}([- \tau, 0], \mathbf{R}^d).$$

Associate with any partition $\pi := \{t_{-M} < t_{-M+1} < \dots < t_{-1} < t_0 = 0 < t_1 < t_2 < \dots < t_{N-1} < t_N = T\}$ of $[- \tau, T]$, the Euler approximations y of the solution x of (4.1) which satisfy the following SFDE:

$$(4.2) \quad y(t) = \eta(0) + \int_0^t f(\lfloor u \rfloor, \Pi^1(y_{\lfloor u \rfloor}), \Pi_q^1(y_{\lfloor u \rfloor})) du + \sum_{l=1}^m \int_0^t g_l(\lfloor u \rfloor, \Pi^{2,l}(y_{\lfloor u \rfloor}), \Pi_q^{2,l}(y_{\lfloor u \rfloor})) dW_l(u),$$

for $\sigma \leq t \leq T$, with initial path

$$y_\sigma = \eta \in H^{1,\infty}([- \tau, 0], \mathbf{R}^d).$$

Under sufficient regularity hypotheses on the coefficients of (??), one obtains weak convergence of order 1 of the Euler approximations y to the exact solution x .

Theorem 4.1. *Let $\phi : \mathbf{R}^d \rightarrow \mathbf{R}$ be of class C_b^3 . Assume that the coefficients $f, g_l, 1 \leq l \leq m$, in (??) are C^1 in the time variable and C_b^3 in all space variables. Let $x(\cdot; \sigma, \eta)$ be the unique solution of (??) with initial path $\eta \in H^{1,\infty}([- \tau, 0], \mathbf{R}^d)$ starting at $\sigma \in [0, T]$. Let η^π be the piecewise-linear approximation of η along the partition π . Denote by $y(\cdot; \sigma, \eta)$ the Euler approximation to $x(\cdot; \sigma, \eta)$ defined by (??). Then there is a positive constant K and a positive integer q such that*

$$(4.3) \quad |E\phi(x(t; \sigma, \eta)) - E\phi(y(t; \sigma, \eta^\pi))| \leq K(1 + \|\eta\|_{1,\infty}^q) |\pi|$$

for all $t \in [\sigma - \tau, T]$ and all $\eta \in H^{1,\infty}([- \tau, 0], \mathbf{R}^d)$. The constant K may depend on T, q and the test function ϕ , but is independent of $\pi, \eta, t \in [\sigma - \tau, T]$ and $\sigma \in [0, T]$.

The proof of Theorem ?? follows along very similar lines to that of Theorem ?. The main difference is a straightforward application of the classical Itô formula combined with the tame Itô formula in section 2 (cf. Lemma 4.2 below). Details of the proofs of Lemma 4.2 and Theorem 4.1 are left to the reader.

Lemma 4.2. *Let $\psi : \mathbf{R} + \times \mathbf{R}^3 \rightarrow \mathbf{R}$ be of class C^1 in the time-variable and C_b^2 in the three space variables x_1, x_2, x_3 . Suppose x solves the SFDE (4.1) (for $d = 1$) with coefficients satisfying the hypotheses of Theorem 4.1. Assume that h, μ are smooth functions. Let $\delta > 0$. Then*

$$\begin{aligned}
& d\psi(t, x(t-\delta), x(t), \int_{-\delta}^0 h(x(t+s))\mu(s) ds) \\
&= \frac{\partial\psi}{\partial t}(x(t-\delta), x(t), \int_{-\delta}^0 h(x(t+s))\mu(s) ds) dt \\
&+ \frac{\partial\psi}{\partial x_1}(x(t-\delta), x(t), \int_{-\delta}^0 h(x(t+s))\mu(s) ds) 1_{[0,\delta)}(t) d\eta(t-\delta) \\
&+ \frac{\partial\psi}{\partial x_1}(x(t-\delta), x(t), \int_{-\delta}^0 h(x(t+s))\mu(s) ds) 1_{[\delta,\infty)}(t) \times \\
&\times [f(t-\delta, \Pi^1(x_{t-\delta}), \Pi_q^1(x_{t-\delta})) dt + \sum_{l=1}^m g_l(t-\delta, \Pi^{2,l}(x_{t-\delta}), \Pi_q^{2,l}(x_{t-\delta})) dW_l(t-\delta)] \\
&+ \frac{\partial\psi}{\partial x_2}(x(t-\delta), x(t), \int_{-\delta}^0 h(x(t+s))\mu(s) ds) \times \\
&\times [f(t, \Pi^1(x_t), \Pi_q^1(x_t)) dt + \sum_{l=1}^m g_l(t, \Pi^{2,l}(x_t), \Pi_q^{2,l}(x_t)) dW_l(t)] \\
&+ \frac{\partial^2\psi}{\partial x_1 \partial x_2}(x(t-\delta), x(t), \int_{-\delta}^0 h(x(t+s))\mu(s) ds) \times \\
&\times \sum_{l=1}^m g_l(t-\delta, \Pi^{2,l}(x_{t-\delta}), \Pi_q^{2,l}(x_{t-\delta})) 1_{[\delta,\infty)}(t) \mathcal{D}_{t-\delta}^l x(t) dt \\
(4.4) \quad &+ \frac{1}{2} \frac{\partial^2\psi}{\partial x_1^2}(x(t-\delta), x(t), \int_{-\delta}^0 h(x(t+s))\mu(s) ds) \sum_{l=1}^m g_l(t-\delta, \Pi^{2,l}(x_{t-\delta}), \Pi_q^{2,l}(x_{t-\delta}))^2 \\
&\times 1_{[\delta,\infty)}(t) dt \\
&+ \frac{1}{2} \frac{\partial^2\psi}{\partial x_2^2}(x(t-\delta), x(t), \int_{-\delta}^0 h(x(t+s))\mu(s) ds) \sum_{l=1}^m g_l(t, \Pi^{2,l}(x_t), \Pi_q^{2,l}(x_t))^2 dt,
\end{aligned}$$

for all $t > 0$. The symbol \mathcal{D}^l stands for Malliavin derivative with respect to the Brownian component W_l , $1 \leq l \leq m$.

Appropriate generalizations of Lemma 4.2 hold for higher dimensional versions of the SFDE (4.1) ($d > 1$).

5. NUMERICAL EXPERIMENTS

We present results of numerical experiments corresponding to an example of (?). Our objective is to illustrate the weak convergence of the Euler-Maruyama method with respect to decreasing step-size by computing first moments, that is we compute $\mathbb{E}\phi(x(T))$ for $\phi(x) = x$ where $x(T)$ denotes a solution of (?).

Example 5.1. Consider the scalar SDDE

$$(5.1) \quad x(t) = \begin{cases} \eta(0) + \int_0^t 3e^{-1}x(u-1) - 4x(u) + 4 - 3e^{-1} du \\ + \int_0^t 0.01x(u-\frac{1}{2}) + 0.1x(u) dW(u), & 0 \leq t \leq 2, \\ \eta(t) = 1 + e^{-1t}, & -1 \leq t < 0, \end{cases}$$

where $W(t)$ is a one dimensional Wiener process.

Let $m(t) := \mathbb{E}x(t)$ for $t \geq 0$. Then, $m(t)$ satisfies the delay differential equation

$$(5.2) \quad m'(t) = 3e^{-1}m(t-1) - 4m(t) + 4 - 3e^{-1},$$

with initial condition

$$(5.3) \quad m(t) = 1 + e^{-1t} \quad \text{for } -1 \leq t \leq 0.$$

Equation (??) subject to (??) has the solution $m(t) = 1 + e^{-1t}$, the solution is a continuation of the initial function.

To illustrate the convergence of the method, we have simulated 150000 sample trajectories with each of the (equidistant) step-sizes $|\pi| = 2^{-3}, 2^{-4}, \dots, 2^{-7}$ and computed the error

$$(5.4) \quad \mu^{|\pi|}(T) = |\mathbb{E}(x(T)) - \mathbb{E}(y(T))|$$

at the final time $T = 2$. In Figure ??, we have plotted $\log_2(\mu^{|\pi|}(T))$ versus $\log_2(|\pi|)$.

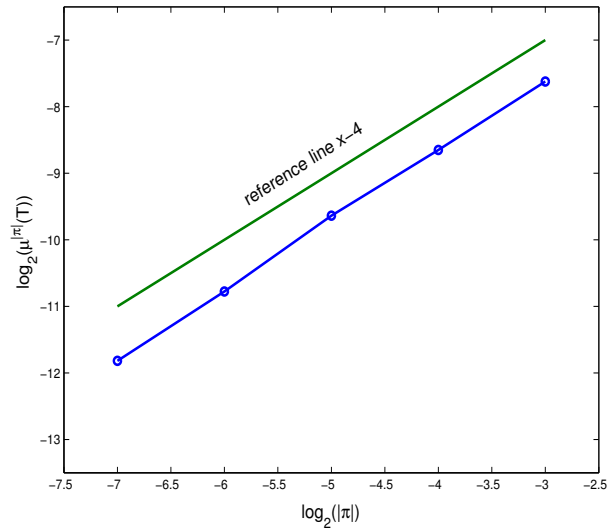


FIGURE 1. $\log_2(\mu^{|\pi|}(T))$ versus $\log_2(|\pi|)$ for the solution of (??).

For illustration purposes we also include some trajectories in the following figure, the thick line corresponds to $m(t) = 1 + e^{-1t}$.

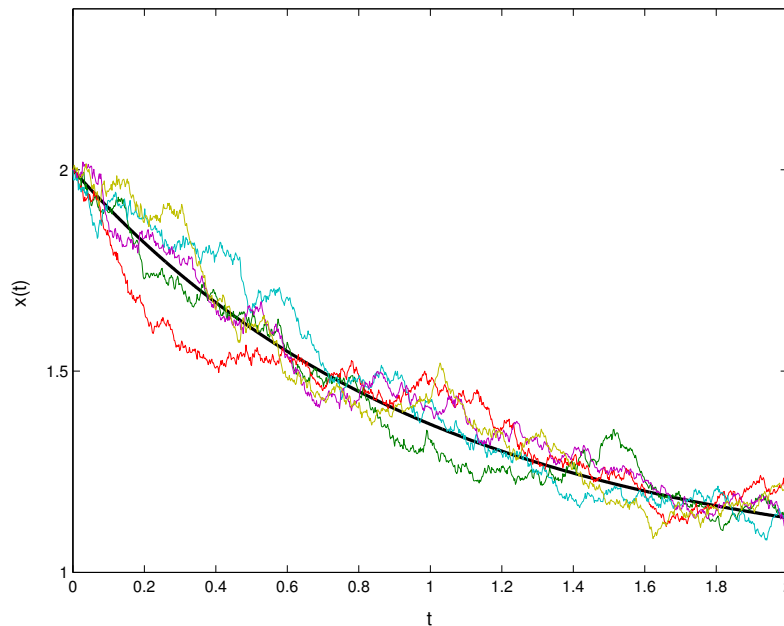


FIGURE 2. Trajectories of the solution of (?).

REFERENCES

- [A] T. A. Ahmed, *Stochastic Functional Differential Equations with Discontinuous Initial Data*. M.Sc. Thesis, University of Khartoum, Sudan (1983).
- [B.B.1] C. T. H. Baker and E. Buckwar, *Numerical analysis of explicit one-step methods for stochastic delay differential equations*. LMS J. Comput. Math. **3** (2000), 315–335.
- [B.B.2] C. T. H. Baker and E. Buckwar, *Continuous Θ -methods for the stochastic pantograph equation*. Electron. Trans. Numer. Anal. **11** (2000), 131–151.
- [B.M] D. Bell and S.-E. A. Mohammed, *The Malliavin calculus and stochastic delay equations*. Journal of Functional Analysis **99**(1) (1991), 75–99.
- [B.V] A. Beuter and K. Vasilakos, *Effects of noise on a delayed visual feedback system*. J. Theor. Biology **165** (1993), 389–407.
- [B.S] E. Buckwar and T. Shardlow, *Weak approximation of stochastic differential delay equations*. IMA J. Numer. Anal. **25** (2005), 57–86.
- [B.T] V. Bally, and D. Talay, *The law of the Euler scheme for stochastic differential equations. I. Convergence rate of the distribution function*, Probab. Theory Related Fields, 104 (1996), 43–60.
- [B.G.M.T.S] J. M. Buldú, J. Garcia-Ojalvo, C. R. Mirasso, M. C. Torrent and J. M. Sancho, *Effect of external noise correlation in optical coherence resonance*. Phys. Rev. E **64** (2001), 051109-1–051109-4.
- [E.M] C. W. Eurich and J. G. Milton, *Noise-induced transitions in human postural sway*. Phys. Rev. E **54**(6) (1996), 6681–6684.
- [F.L.L.L.1] E. Fournié, J.-M. Lasry, J. Lebuchoux and P.-L. Lions, *Applications of Malliavin calculus to Monte Carlo methods in finance*. Finance Stoch. **3**(4) (1999), 391–412.
- [F.L.L.L.2] E. Fournié, J.-M. Lasry, J. Lebuchoux and P.-L. Lions, *Applications of Malliavin calculus to Monte-Carlo methods in finance. II.*, Finance Stoch. **5**(2) (2001), 201–236.
- [G.P.R] D. Goldobin, M. Rosenblum and A. Pikovsky, *Coherence of noisy oscillators with delayed feedback* Physica A **327** (2003), 124–128.
- [Hu] Y. Hu, *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.
- [H.M.Y] Y. Hu, S.-E. A. Mohammed and F. Yan, *Discrete-time approximations of stochastic delay equations: The Milstein scheme*. The Annals of Probability, **32**(1A), (2004), 265–314.
- [H.T] D. Huber and L. S. Tsimring, *Dynamics of an ensemble of noisy bistable elements with global time delayed coupling*. Phys. Rev. Lett. **91** (2003), 260601.
- [K.P] P. E. Kloeden and E. Platen, *Numerical Solutions of Stochastic Differential Equations*, Springer-Verlag, New York-Berlin (1995).
- [K.P.S] P. E. Kloeden, E. Platen and H. Schurz, *Numerical Solutions of SDE Through Computer Experiment*, Springer-Verlag, New York-Berlin (1997).
- [K-H] A. Kohatsu-Higa, *Weak approximations. A Malliavin calculus approach*. Math. Comput. **70**(233) (2001), 135–172.
- [K-H.P] A. Kohatsu-Higa and R. Pettersson, *Variance reduction methods for simulation of densities on Wiener space*. SIAM J. Numer. Anal. **40**(2) (2002), 431–450.

- [L.M.B.M] A. Longtin and J. G. Milton and J. Bos and M. C. Mackey, *Noise and critical behavior of the pupil light reflex at oscillation onset*. Phys. Rev. A **41** (1990), 6992–7005.
- [Mao.1] X. Mao and S. Sabanis, *Numerical solutions of stochastic differential delay equations under local Lipschitz condition*. J. Comput. Appl. Math. **151**(1) (2003), 215–227.
- [Mao.2] X. Mao, *Numerical solutions of stochastic functional differential equations*. LMS J. Comput. Math. **6** (2003), 141–161.
- [Ma.1] C. Masoller, *Numerical investigation of noise-induced resonance in a semiconductor laser with optical feedback*., Physica D *168-169* (2002), 171–176.
- [Ma.2] C. Masoller, *Distribution of residence times of time-delayed bistable systems driven by noise*. Phys. Rev. Lett. **90**(2) (2003), 020601-1–020601-4.
- [M.T] G. N. MILSTEIN AND M. V. TRET'YAKOV, *Numerical methods in the weak sense for stochastic differential equations with small noise*. SIAM J. Numer. Anal. **34**(6), (1997), 2142–2167.
- [Mo.1] S.-E. A. Mohammed, *Stochastic Functional Differential Equations*. Pitman 99 (1984).
- [Mo.2] S.-E. A. Mohammed, *Stochastic differential systems with memory: Theory, examples and applications*. In: Stochastic Analysis, Decreusefond L. Gjerde J., Øksendal B., Ustunel A.S., eds., *Progress in Probability 42*, Birkhauser (1998), 1–77.
- [Mo.S.1] S.-E. A. Mohammed and M. K. R. Scheutzow, *Lyapunov exponents of linear stochastic functional differential equations driven by semimartingales. I: The multiplicative ergodic theory*. Ann. Inst. Henri Poincaré **32**(1), (1996), 69–105.
- [Mo.S.2] S.-E. A. Mohammed and M. K. R. Scheutzow, *Lyapunov exponents of linear stochastic functional differential equations. II: Examples and case studies*. Ann. Probab. **25**(3), (1997), 1210–1240.
- [N.P] D. Nualart and E. Pardoux, *Stochastic calculus with anticipating integrands*. Probability Theory and Related Fields **78** (1988), 535–581.
- [Nu.1] D. Nualart, *The Malliavin Calculus and Related Topics*. Springer-Verlag (1995).
- [Nu.2] D. Nualart, *Analysis on Wiener Space and Anticipating Stochastic Calculus*. 1995?
- [P.T] A. Pikovsky and L. S. Tsimring, *Noise-induced dynamics in bistable systems with delay*., Phys. Rev. Lett., **87**(25) (2001), 250602-1–250602-4.
- [Tud.1] C. Tudor and M. Tudor, *On approximation of solutions for stochastic delay equations*., Stud. Cercet. Mat., 39 (1987), pp. 265–274.
- [Tud.2] C. Tudor and M. Tudor, *Approximation of linear stochastic functional equations*., Rev. Roum. Math. Pures Appl., 35 (1990), pp. 81–99.
- [Tud.3] M. Tudor, *Approximation schemes for stochastic equations with hereditary argument*., Stud. Cercet. Mat., 44 (1992), pp. 73–85.

EVELYN BUCKWAR: DEPARTMENT OF MATHEMATICS, HUMBOLDT-UNIVERSITY BERLIN, UNTER DEN LINDEN 6, 10099 BERLIN, GERMANY

E-mail address: buckwar@mathematik.hu-berlin.de

RACHEL KUSKE: DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BRITISH COLUMBIA, VANCOUVER, BRITISH COLUMBIA, CANADA

E-mail address: rachel@math.ubc.ca

SALAH-ELDIN MOHAMMED: DEPARTMENT OF MATHEMATICS, SOUTHERN ILLINOIS UNIVERSITY AT CARBONDALE, CARBONDALE, ILLINOIS 62901, U.S.A.

E-mail address: salah@sfde.math.siu.edu

TONY SHARDLOW: SCHOOL OF MATHEMATICS, UNIVERSITY OF MANCHESTER, MANCHESTER, UK

E-mail address: shardlow@maths.man.ac.uk