Weak approximation of stochastic differential delay equations

Evelyn Buckwar* Tony Shardlow†

September 26, 2001

Abstract

A numerical method for a class of Itô stochastic differential equations with a finite delay term is introduced. The method is based on the forward Euler approximation and is parameterised by its time step. Weak convergence with respect to a class of smooth test functionals is established by using the infinite dimensional version of the Kolmogorov equation. With regularity assumptions on coefficients and initial data, the rate of convergence is shown to be proportional to the time step. Some computations are presented to demonstrate the rate of convergence.

Key words Theoretical approximation of solutions, Stochastic partial differential equations, Stochastic delay equations, Stability and convergence of numerical approximations.

AMS Subject Classifications 60H15, 34K50, 65L20, 34A45.

1 Introduction

Consider stochastic differential delay equations on $\mathbf{R}^d$ of the form

$$\begin{aligned} dY(t) =& \Big[\int_{-\tau}^{0} a(ds) Y(t+s) + f(Y(t)) \Big] \, dt + b(Y(t)) \, dW(t), \\ Y(0) =& Y_S, \quad Y(s) = Y_D(s) \text{ for } -\tau < s < 0, \end{aligned} \tag{1.1}$$

for initial conditions $Y_S \in \mathbf{R}^d$ and $Y_D \in L_2([-\tau, 0], \mathbf{R}^d)$, where $a(\cdot)$ is a $d \times d$ measure valued function on $[-\tau, 0]$, $f(\cdot) \colon \mathbf{R}^d \to \mathbf{R}^d$, $b(\cdot) \colon \mathbf{R}^d \to \mathbf{R}^{d\times d}$, and $W(\cdot)$ is a Brownian motion on $\mathbf{R}^d$ with covariance I. The delay is τ, which should be finite and positive. The equation should be interpreted in the sense of Itô.

We now define the forward Euler method for (1.1). Let

$$a_i := \int_{-\tau}^{0} a(ds) \mathbf{1}_{[i\Delta t, (i+1)\Delta t)}(s), \qquad i = -\lfloor \tau/\Delta t \rfloor, \dots, -1,$$

where $\mathbf{1}_{[t_1,t_2)}(s)$ is the $d \times d$ identity matrix on $[t_1, t_2)$ and is zero otherwise. Let $\Delta\beta_n$ be independent and normally distributed with mean zero and variance $\Delta t I$. Generate approximations Y_n to $Y(n\Delta t)$ for $n = 1, 2, \dots$ by

$$Y_{n+1} - Y_n = \Big[\sum_{i=-\lfloor \tau/\Delta t \rfloor}^{-1} a_i Y_{n+i} + f(Y_n) \Big] \Delta t + b(Y_n) \Delta\beta_n, \tag{1.2}$$

*Humboldt-Universität zu Berlin, Institut für Mathematik, Bereich Stochastik, Unter den Linden 6, 10099 Berlin, buckwar@mathematik.hu-berlin.de. Financial support was received by the Deutsche Forschungsgemeinschaft, SFB 373 ("Quantifikation und Simulation Ökonomischer Prozesse"), Humboldt-Universität zu Berlin.

†Department of Mathematical Sciences, Science Laboratories, South Road, Durham University, Durham DH1 3LE, Tony.Shardlow@durham.ac.uk.

with initial conditions $Y_i = Y_D(i\Delta t)$ for $i = -\lfloor \tau/\Delta t \rfloor, \dots, -1$ and $Y_0 = Y_S$.

In a series of papers, strong approximation methods for stochastic differential delay equations were considered by C. and M. Tudor [13, 14, 16, 15, 17]. Recently this topic has gained more attention, see [11], [2], [1], and [9]. The theory gives convergence rates of order $\Delta t^{1/2}$ for the forward Euler method, which is optimal, and applies to delay equations more general than (1.1). The aim of this work is to understand the weak convergence properties of the forward Euler method for (1.1). It is hoped that the theoretical grounding developed for the Euler method in this paper will make it possible to understand higher order weak approximation methods for stochastic differential delay equations. We now describe the hypothesis needed for our weak convergence analysis. The hypothesis are more restrictive than those needed for strong convergence, but give better convergence rates.

Hypothesis 1.1 (i) Suppose that $f\colon \mathbf{R}^d \to \mathbf{R}^d$ is four times continuously differentiable with f', f'', f''', f'''' bounded. Suppose that $b\colon \mathbf{R}^d \to \mathbf{R}^{d\times d}$ is bounded with four bounded derivatives.

(ii) Suppose there exists a strictly positive continuous density $\bar{a}(s)$ on $[-\tau, 0]$ such that for all $g \in L_2([-\tau,0], \mathbf{R}^d)$

$$\Big\| \int_{-\tau}^{0} a(ds) g(s) \Big\|_{\mathbf{R}^d} \le \int_{-\tau}^{0} \bar{a}(s) \|g(s)\|_{\mathbf{R}^d} \, ds.$$

(iii) Suppose that there exists $K > 0$ so that for all $g \in L_2([-\tau,0], \mathbf{R}^d)$,

$$\Big\| \int_{-\tau}^{0} a(ds) \frac{d}{ds} g(s) \, ds \Big\|_{\mathbf{R}^d} \le K \|g\|_{L_2([-\tau,0],\mathbf{R}^d)}.$$

For an integer $p \ge 0$, introduce the space $\mathcal{G}_p$ of test functions $\phi\colon \mathbf{R}^d \to \mathbf{R}$ that are four times continuously differentiable and satisfy $\|\phi^{(n)}(h)\|_{\mathcal{L}(\mathbf{R}^{d\times n}, \mathbf{R})} \le K(1 + \|h\|_{\mathbf{R}^d}^{p-n})$, for $h \in H$ and some constant K, for $n = 0, 1, 2, 3, 4$. Thus the derivatives of ϕ can be bounded like a polynomial.

For $x = (Y_S, Y_D)^T$, write $\|x\| := (\|Y_S\|_{\mathbf{R}^d}^2 + \|Y_D\|_{L_2([-\tau,0],\mathbf{R}^d)}^2)^{1/2}$. For a continuous function $Y_D\colon [-\tau, 0] \to \mathbf{R}^d$, let

$$\|Y_D\|_{\mathrm{Lip}} := \sup_{-\tau \le t, t' \le 0} \frac{\|Y_D(t) - Y_D(t')\|_{\mathbf{R}^d}}{|t - t'|}.$$

Theorem 1.2 *Let Hypothesis 1.1 hold. Consider $Y_S \in \mathbf{R}^d$ and a globally Lipschitz function $Y_D : [-\tau, 0] \to \mathbf{R}^d$. Let $Y(t)$ (respectively, Y_n) denotes the solution of (1.1) (resp., (1.2)) corresponding to initial data $x = (Y_S, Y_D)^T$. For $T > 0$ and $\phi \in \mathcal{G}_p$, $p \ge 1$, there exists a constant $K_x > 0$ such that*

$$\Big| \mathbf{E}\phi(Y(T)) - \mathbf{E}\phi(Y_N) \Big| \le K_x \Delta t, \quad N\Delta t = T$$

and a constant K independent of the initial data such that

$$K_x \le K(1 + \|x\|^p) + K(1 + \|x\|^{p-1})(1 + \|Y_D\|_{Lip}).$$

This is the main result of the present paper. The proof is built by developing the delay equation (1.1) as a stochastic evolution equation on an infinite dimensional space. We review the theory in §2. Two corollaries of the Itô calculus are established in §3 concerning certain functionals of the solutions. The Kolmogorov equation is introduced in §4 and developed in full for the delay equation (though in fact, only a regularised version of the Kolmogorov equation is used directly in the proof of Theorem 1.2). A number of regularity results are established. It is important to establish sufficient time and spatial regularity of $v(t,x) := \mathbf{E}\phi(Y(t))$, where $Y(t)$ is the solution of (1.1) for initial data $x := (Y_S, Y_D)^T$, and the terms in the Kolmogorov equation, to apply again the Itô formula. To gain the necessary regularity, Hypothesis 1.1 (iii) was introduced. This hypothesis excludes the important case of discrete delays, $a(ds) = \sum \delta_{\tau_i}(ds)$. The proof is completed in §5.

Weak approximation has been established for many numerical approximations of SDEs by looking at the Kolmogorov equation. The argument given in this paper follows closely [10]; an alternative that makes use of the Malliavin calculus is given in [8]. The difference in the present case is the introduction of a delay term so that the equation must be phrased on an infinite dimensional phase space to achieve a Markov process and a Kolmogorov equation. The authors are unaware of any previous use of the infinite dimensional Kolmogorov equations to analyse the weak convergence of numerical methods. It remains to be seen whether the technique can be more widely applied, for example to numerical methods for a heat equation forced by a Wiener process.

The Kolmogorov equation is difficult for evolution equations forced by a Wiener process. The drift terms in the underlying evolution equation frequently involve a differential operator A which is unbounded. Further, the covariance of the Wiener process may involve an infinite number of non-trivial eigenvalues. In our case, the Kolmogorov equation is simplified as there are only finitely many noise terms and the operator A has a nice structure. Though A is unbounded, we can take advantage of A being bounded in its first component. To do this, we have taken a particularly simple space of test functions by working over averages at the current time and keeping the test functions independent of the delay. The averages of these test functions carry no information about the correlation between the state variable over the delay interval, but are a natural space of functions to use in this situation.

1.1 Notation

We will work on the space $H := \mathbf{R}^d \times L_2([-\tau, 0], \mathbf{R}^d)$ with norm $\|(X_S, X_D)\| := (\|X_S\|^2_{\mathbf{R}^d} + \|X_D\|^2_{L_2([-\tau,0],\mathbf{R}^d)})^{1/2}$, which consists of the state variable and delay function. If $X = (X_S, X_D)^T$, let $\pi_S X := X_S$ and $\pi_D X := X_D$. The norm induced on a linear operator between normed vector spaces H_1 to H_2 is denoted by $\|\cdot\|_{\mathcal{L}(H_1,H_2)}$. Let $\|X\|_S := \|X_S\|_{\mathbf{R}^d}$ and

$$|X|_\star := \|X_S\|_{\mathbf{R}^d} + \left\| \int_{-\tau}^0 a(ds) X_D(s)\, ds \right\|_{\mathbf{R}^d}.$$

Then $|\cdot|_\star$ is a well defined semi-norm on H. Note that, for a constant K, we have $|X|_\star \le K\|X\|$, all $X \in H$. For an orthonormal basis e_i of $\mathbf{R}^d$ and $\mathcal{B} \in \mathcal{L}(\mathbf{R}^d, H)$, define the Hilbert-Schmidt norm

$$\|\mathcal{B}\|^2_{HS} := \sum_{i=1}^d \|\mathcal{B} e_i\|^2.$$

Let L_2^0 be the $\mathcal{L}(\mathbf{R}^d, H)$ valued operators with finite Hilbert-Schmidt norm $\|\cdot\|_{HS}$. Throughout the paper, we will make use of a generic constant K, which will be independent of the time interval $[0, T]$, the initial data x, and k, the parameter of the Yosida approximant A_k. Let

$\hat{s} := \Delta t\lfloor s/\Delta t \rfloor$, the largest multiple of Δt less than s. For $\phi \in \mathcal{G}_p$ and $X \in H$, we will write $\phi(X)$ for $\phi(\pi_S X)$.

2 Background

2.1 Stochastic Evolution Equations

For the analysis, it is convenient to present (1.1) as a stochastic evolution equation on the infinite dimensional space H as follows. Consider

$$dX(t) = \Big[AX(t) + F(X(t))\Big]\,dt + B(X(t))\,dW(t), \qquad X(0) = x := (Y_S, Y_D)^T, \tag{2.1}$$

where for $X = (X_S, X_D)^T$

$$F(X) = \begin{pmatrix} f(X_S) \\ 0 \end{pmatrix} \qquad B(X) = \begin{pmatrix} b(X_S) \\ 0 \end{pmatrix}$$

and A is a densely defined linear operator with domain $\mathcal{D}(A)$,

$$\mathcal{D}(A) := \Big\{(X_S, X_D)^T \in \mathbf{R}^d \times W^{1,2}([-\tau, 0]; \mathbf{R}^d) : \\ X_D \text{ absolutely continuous and } X_D(0) = X_S\Big\}$$

and for $X \in \mathcal{D}(A)$

$$AX := \begin{pmatrix} 0 & C \\ 0 & \frac{d}{dt} \end{pmatrix} X, \qquad CX_D := \int_{-\tau}^{0} a(ds)X_D(s).$$

$W^{1,2}([-\tau, 0], \mathbf{R}^d)$ is the Sobolev space with norm $(\|f\|^2_{L_2([-\tau,0],\mathbf{R}^d)} + \|f'\|^2_{L_2([-\tau,0],\mathbf{R}^d)})^{1/2}$. For further details see [12] and for delay equations [7] and [4]. The evolution equation (2.1) has a unique mild solution subject to Lipschitz conditions on f and b. That is, we can find $X(t;x)$, an adapted H valued process such that

$$X(t;x) = S(t)x + \int_0^t S(t-s)F(X(s;x))\,ds + \int_0^t S(t-s)B(X(s;x))\,dW(s),$$

where $S(t)$ is the semigroup with generator A. The solution $X(t;x)$ corresponds to the solution of (1.1), in the sense that $\pi_S X(t;x) = Y(t)$.

The process $X(t;x)$ is a Markov process [6]. Note that, under Hypothesis 1.1(iii), $C\frac{d}{dt}$ is a bounded operator from $L_2([-\tau, 0], \mathbf{R}^d)$ to $\mathbf{R}^d$.

2.2 Itô Calculus

For reference, we state two basic results of the Itô calculus on infinite dimensional spaces. Let $\mathcal{A}(t)$ be a H valued predictable process, Bochner integrable on $[0,T]$. Let $\mathcal{B}(t)$ be an L_2^0 valued process such that $\int_0^t \|\mathcal{B}(s)\|^2_{HS}\,ds$ is finite almost surely. Consider $X(t)$ such that

$$dX(t) = \mathcal{A}(t)\,dt + \mathcal{B}(t)\,dW(t),$$

where $W(t)$ is a Wiener process on $\mathbf{R}^d$ with covariance I. The next two results are dealt with by [5].

Theorem 2.1 (Itô Formula) *Consider a function* $\Phi\colon [0,T]\times H\to \mathbf{R}$. *Suppose that* Φ *and its partial derivatives* Φ_t, Φ_x, Φ_{xx} *are uniformly continuous on bounded subsets of* $[0,T]\times H$. *For* $0\le t\le T$, *almost surely,*

$$\Phi(t,X(t)) = \Phi(0,X(0)) + \int_0^t \Phi_x(s,X(s))\mathcal{B}(s)\,dW(s)$$
$$+\int_0^t \Big\{\Phi_t(s,X(s)) + \Phi_x(s,X(s))\mathcal{A}(s) + \tfrac{1}{2}\operatorname{Tr}\Phi_{xx}(s,X(s))\mathcal{B}(s)\mathcal{B}(s)^*\Big\}\,ds,$$

where (for an orthonormal basis e_i *of* $\mathbf{R}^d$*)*

$$\operatorname{Tr}\Phi_{xx}(s,X(s))\mathcal{B}(s)\mathcal{B}(s)^* = \sum_{i=1}^d \Phi_{xx}(s,X(s))(\mathcal{B}(s)e_i,\mathcal{B}(s)e_i).$$

Lemma 2.2 *The Itô Isometry:*

$$\mathbf{E}\Big[\Big(\int_0^T \mathcal{B}(s)\,dW(s)\Big)^2\Big] = \int_0^T \mathbf{E}\|\mathcal{B}(s)\|_{HS}^2\,ds.$$

The Burkholder-Davis-Gundy Inequality: for $p>0$, *there exists a constant* c_p *with*

$$\mathbf{E}\Big[\sup_{0\le t\le T}\Big(\int_0^t \mathcal{B}(s)\,dW(s)\Big)^p\Big] \le c_p\mathbf{E}\Big|\int_0^T \|\mathcal{B}(s)\|_{HS}^2\,ds\Big|^{p/2}$$

2.3 Regularity of solutions

Theorem 2.3 (dependence on initial condition) *Let Hypothesis 1.1(i) hold. There exists a unique mild solution* $X(t;x)$ *of (2.1), which is four times continuously differentiable in the initial condition* x *and whose derivatives are mild solutions of the corresponding variational equation (obtained by differentiating (2.1) with respect to the initial condition). For* $T>0$, *the solution* $X(t;x)$ *of (2.1) obeys for* $0\le t\le T$

$$\begin{aligned}\mathbf{E}\|X(t;x)\|^p \le& K(1+\|x\|^p)\\ (\mathbf{E}\|X(t;x)-X(t;x')\|^2)^{1/2} \le& K\|x-x'\|(1+\|x\|).\end{aligned}$$

Moreover the derivatives are bounded in the following sense. The quantities

$$\mathbf{E}(\|X_x(t;x)\|^p_{\mathcal{L}(H,H)}),\quad \mathbf{E}(\|X_{xx}(t;x)\|^p_{\mathcal{L}(H\times H,H)}),$$
$$\mathbf{E}(\|X_{xxx}(t;x)\|^p_{\mathcal{L}(H\times H\times H,H)}),\quad \mathbf{E}(\|X_{xxxx}(t;x)\|^p_{\mathcal{L}(H\times H\times H\times H,H)})$$

are bounded for $0\le t\le T$.

Proof See Da Prato–Zabczyk [5] Theorem 9.4. The higher order derivatives are understood by writing the appropriate variational equation. The bound is uniform in x because of the boundedness of the derivatives of f and b in Hypothesis 1.1. *QED*

Corollary 2.4 *Let Hypothesis 1.1(i) hold. Consider* $\phi\in\mathcal{G}_p$ *and let* $v(t,x) := \mathbf{E}\phi(X(t;x))$. *The function* v *and its derivatives* v_x, v_{xx}, v_{xxx}, *and* v_{xxxx} *are uniformly continuous in* x *on bounded subsets of* $\mathbf{R}^+\times H$. *For* $0\le t\le T$

$$|v(t,x)|\le K(1+\|x\|^p)$$

and

$$\|v_x\|_{\mathcal{L}(H,\mathbf{R})}, \quad \|v_{xx}\|_{\mathcal{L}(H\times H,\mathbf{R})}, \quad \|v_{xxx}\|_{\mathcal{L}(H\times H\times H,\mathbf{R})}, \quad \|v_{xxxx}\|_{\mathcal{L}(H\times H\times H\times H,\mathbf{R})}$$

are all bounded by a constant times $(1+\|x\|^{p-1})$ *on the interval* $[0,T]$.

Proof Clearly, $|v(t,x)| \le K\mathbf{E}(1+\|X(t;x)\|^p) \le K(1+\|x\|^p)$ from Theorem 2.3. Similar estimates follow for v_x, v_{xx}, v_{xxx}, and v_{xxxx} given the estimates on X_x, X_{xx}, X_{xxx} and X_{xxxx} in Theorem 2.3 and the hypothesis on ϕ.

To argue for uniform continuity, consider data x, x' with $\|x\|, \|x'\| \le M$ and choose $\epsilon > 0$. Choose R sufficiently large that $\mathbf{P}(\|X(t;x)\| \le R,\ 0 \le t \le T) \ge 1-\epsilon$. Then, as ϕ is locally Lipschitz, for a constant K_R,

$$\begin{aligned}|v(t,x)-v(t,x')| \le& \epsilon K(1+\|x\|^p) + K_R(\mathbf{E}\|X(t;x)-X(t;x')\|^2)^{1/2}\\ \le& \epsilon K(1+M^p) + K_R(1+M)\|x-x'\|.\end{aligned}$$

This can be made arbitrarily small by choosing ϵ small (viz. R large) and then $\|x-x'\|$ small, and implies uniform continuity of $v(t,x)$ in x on bounded subsets of $\mathbf{R}^+\times H$. The argument extends to v_x, v_{xx}, v_{xxx}, and v_{xxxx} given the continuity in the initial condition of X_x, X_{xx}, etc. described in Theorem 2.3. *QED*

2.4 Yosida approximations

The operator A is unbounded due to the differential operator in the second component. We will frequently approximate A by its Yosida approximant A_k (defined shortly). By use of the Yosida approximant, we find strong solutions of an SDE that converge to the mild solutions of (2.1) and that yield to the Itô formula. For a review of these ideas, see [12].

The Yosida approximant $A_k := kAR(k\colon A) = k^2R(k\colon A) - kI$, where the resolvent $R(k\colon A) := (kI-A)^{-1}$. A simple calculation shows that

$$A_kX = \begin{pmatrix}0 & Ck(kI-\frac{d}{dt})^{-1}\\ 0 & \frac{d}{dt}k(kI-\frac{d}{dt})^{-1}\end{pmatrix} X = A\begin{pmatrix}0\\ \mathcal{P}_kX\end{pmatrix}, \tag{2.2}$$

where $\mathcal{P}_kX = h$, the solution of $kX_D = kh - \frac{d}{dt}h$ on $[-\tau,0]$ for $h(0) = X_S$.

Define $S_k(t) = e^{A_kt}$ and $S(t) = e^{At}$, the semigroups generated by A_k and A. The following properties hold.

Proposition 2.5 (Yosida approximants) *(i)* $A_kh \to Ah$ *for* $h \in \mathcal{D}(A)$ *as* $k\to\infty$.

(ii) $S_k(t)h \to S(t)h$ *as* $k\to\infty$ *for* $h\in H$ *and* $S_k(t)$ *is bounded in* $\mathcal{L}(H,H)$ *uniformly in* k. *Moreover,* $\|S_k(t)x - S_\ell(t)x\| \le K\|A_kx - A_\ell x\|$ *for* $k,\ell = 1,2,\dots$ *and* $0\le t\le T$.

(iii) π_SA_k *is an operator from* H *to* $\mathbf{R}^d$ *uniformly bounded in* k. *Further* π_SA_kh *converges in* $\mathbf{R}^d$ *for every* $h\in H$ *to a limit, which we denote by* π_SAh. *In practice, for* $\phi\in\mathcal{G}_p$, *this means* $\phi'(X)Ah$ *is well defined as the limit of* $\phi'(X)A_kh$.

Proof The first two properties are standard results from C_0 semigroups (see §1.5 of [12]). The third property follows from property (i), if $\|\pi_SA_k\|_{\mathcal{L}(H,\mathbf{R}^d)}$ is bounded. But $\pi_SA_k = C\mathcal{P}_k$, a product of two operators, both of which are bounded for k large. *QED*

Lemma 2.6 *Consider the mild solution $X(t;x)$ of*

$$dX = \Big[AX + F(X)\Big]\,dt + B(X)\,dW, \quad X(0) = x,$$

and the strong solution $X^k(t;x)$ of

$$dX^k = \Big[A_k X^k + F(X^k)\Big]\,dt + B(X^k)\,dW, \quad X^k(0) = x. \tag{2.3}$$

Then,

$$\sup_{0\le t\le T} \mathbf{E}\|X(t;x) - X^k(t;x)\|^p \to 0, \qquad \text{as } k \to \infty.$$

Proof Proposition 7.5 [6]. *QED*

2.5 The numerical method on H

To perform the convergence analysis, we need an interpolant of the numerical solution Y_n in H. We will denote the interpolant by $X^{\Delta t}(t;x)$ and will also consider a smoothed process $X^{\Delta t,k}(t;x)$. Introduce $\bar{W}(t)$, an $\mathbf{R}^d$ valued Wiener process with covariance I such that the increments generate $\Delta\beta_n$ in (1.2). Thus, $\bar{W}((n+1)\Delta t) - \bar{W}(n\Delta t) = \Delta\beta_n$. Consider $n\Delta t \le t < (n+1)\Delta t$. Then, define $X^{\Delta t} = (X_S^{\Delta t}, X_D^{\Delta t})^T$ by

$$\begin{aligned}
X_S^{\Delta t}(t;x) :=& Y_n + \Big[\sum_{i=-\lfloor\tau/\Delta t\rfloor}^{-1} a_i Y_{n+i} + f(Y_n)\Big](t - n\Delta t) + b(Y_n)\,(\bar{W}(t) - \bar{W}(n\Delta t))\\
=& X_S^{\Delta t}(\hat{t};x) + \Big[\sum_{i=-\lfloor\tau/\Delta t\rfloor}^{-1} a_i X_D(\,\hat{t};x)(i\Delta t) + f(X_S(\,\hat{t};x))\Big](t-\hat{t})\\
&+ b(X_S(\,\hat{t};x))\,(\bar{W}(t) - \bar{W}(\hat{t}\,))
\end{aligned} \tag{2.4}$$

$$X_D^{\Delta t}(t;x)(s) := \begin{cases} X_S(t+s;x), & t+s \ge 0,\\ Y_D(t+s), & -\tau \le t+s < 0,\end{cases} \qquad -\tau \le s \le 0.$$

It is necessary to develop this equation as a well defined H valued stochastic integral. However, the delay term is not well defined for $X_D \in L_2([-\tau,0],\mathbf{R}^d)$. We smooth out the delay term by using $\mathcal{P}_k$ as in (2.2) and writing for a continuous function $X_D\colon [-\tau,0] \to \mathbf{R}^d$

$$C^{\Delta t} X_D := \sum_{i=-\lfloor\tau/\Delta t\rfloor}^{-1} a_i X_D(i\Delta t).$$

The expression $C^{\Delta t,k}\mathcal{P}_k$ is a well defined operator from H to $\mathbf{R}^d$. Introduce

$$\tilde{A} := \begin{pmatrix} 0 & 0\\ 0 & \frac{d}{dt}\end{pmatrix} \tag{2.5}$$

and denote the Yosida approximation of $\tilde{A}$ by $\tilde{A}_k$ (in fact, $\tilde{A}_k = \tilde{A}[0,\mathcal{P}_k]^T$). Let $X^{\Delta t,k}(t;x)$ solve

$$\begin{aligned}
dX^{\Delta t,k}(t;x) =& \Big[\tilde{A}_k X^{\Delta t,k}(t;x) + \begin{pmatrix} C^{\Delta t,k}\\ 0\end{pmatrix} \mathcal{P}_k X^{\Delta t}(\,\hat{t};x) + F(X^{\Delta t,k}(\,\hat{t}\,;x))\Big]\,dt\\
&+ B(X^{\Delta t,k}(\,\hat{t}\,;x))\,d\bar{W}(t), \qquad X^{\Delta t,k}(0;x) = x.
\end{aligned} \tag{2.6}$$

This equation admits a unique strong solution, which converges to $X^{\Delta t}$ as described in the following Lemma. Notice that the effects of smoothing and applying the numerical method to A is that the integral term acts on at the frozen function $X(\hat{t};x)$ rather than $X(t;x)$; the time derivative is smoothed as in (2.2).

Lemma 2.7 *The solution $X^{\Delta t,k}(t;x)$ of (2.6) converges to the interpolant $X^{\Delta t}(t;x)$ defined in (2.4) in the sense that*

$$\sup_{0\le t\le T} \mathbf{E}\|X^{\Delta t}(t;x)-X^{\Delta t,k}(t;x)\|^2 \to 0, \qquad \text{as } k\to\infty.$$

We now state some properties of the interpolant and then explain two Lemmas that will be used later to understand the approximation of the integral $\int a(dr)$. Let $\langle\cdot,\cdot\rangle$ denote the standard Euclidean inner product and $o(k^{-1})$ denote a real valued function that tends to zero as $k\to\infty$. The following estimates for the numerical solution are easily established: for $0\le t\le T$ and $x,x'\in H$ and $p\ge 2$,

$$\mathbf{E}\|X^{\Delta t,k}(t;x)\|^p \le K(1+\|x\|)^p \tag{2.7}$$

$$\mathbf{E}\|X^{\Delta t,k}(\Delta t;x)-\mathbf{E}X^{\Delta t,k}(\Delta t;x)\|^2 \le K(1+\|x\|)^2\,\Delta t \tag{2.8}$$

$$\mathbf{E}\|X^{\Delta t,k}(t;x)-X^{\Delta t,k}(t;x')\|^2 \le K(1+\|x\|)^2\,\|x-x'\|^2. \tag{2.9}$$

Lemma 2.8 *For $0\le t\le T$,*

$$\mathbf{E}\|\pi_S(X^{\Delta t,k}(t;x)-X^{\Delta t,k}(\hat{t};x))\|^2_{\mathbf{R}^d} \le K(1+\|x\|)^2\Delta t. \tag{2.10}$$

For $-\lfloor\tau/\Delta t\rfloor\le i\ne j\le -1$ and $\hat{t}+\min(i,j)\Delta t\ge 0$,

$$\begin{aligned} I := \mathbf{E}\Big[\Big\langle \int_{i\Delta t}^{(i+1)\Delta t} & a(dr)\pi_S(X^{\Delta t,k}(\hat{t}+r;x)-X^{\Delta t,k}(\hat{t}+\hat{r};x)),\\ \int_{j\Delta t}^{(j+1)\Delta t} & a(dr)\pi_S(X^{\Delta t,k}(\hat{t}+r;x)-X^{\Delta t,k}(\hat{t}+\hat{r};x))\Big\rangle\Big] \le K(1+\|x\|)^2\Delta t^4. \end{aligned} \tag{2.11}$$

Proof The process $X^{\Delta t,k}$ solves (2.6) and hence satisfies

$$\begin{aligned} X^{\Delta t,k}(t;x)-X^{\Delta t,k}(\hat{t};x) &= (\tilde{S}_k(t-\hat{t})-I)X^{\Delta t,k}(\hat{t};x)\\ &+\int_{\hat{t}}^t \tilde{S}_k(t-s)\begin{pmatrix}C^{\Delta t,k}\\0\end{pmatrix}\mathcal{P}_k X^{\Delta t}(\hat{t};x)\,ds+\int_{\hat{t}}^t \tilde{S}_k(t-s)F(X^{\Delta t,k}(\hat{t};x))\,ds\\ &+\int_{\hat{t}}^t \tilde{S}_k(t-s)B(X^{\Delta t,k}(\hat{t};x))\,d\bar{W}(s), \end{aligned}$$

where $\tilde{S}_k$ is the semigroup with infinitesimal generator $\tilde{A}_k$. As $\pi_S\tilde{S}_k(s)$ equals the identity matrix acting on $\mathbf{R}^d$ and $\|\tilde{S}_k\|_{\mathcal{L}(H,H)}$ is bounded and $|t-\hat{t}|\le\Delta t$, this implies (2.10).

Consider integers $j<i$ with $t+j\Delta t\ge 0$. Let $\mathcal{F}_t$ be the σ-algebra generated by $\{\bar{W}(s)\colon s\le t\}$. Because $X(\hat{t}+j\Delta t+r)$ for $0\le r\le\Delta t$ is $\mathcal{F}_{\hat{t}+i\Delta t}$ measurable,

$$\begin{aligned} I = \mathbf{E}\Big[\Big\langle \int_0^{\Delta t} & a(dr)\pi_S\mathbf{E}\Big[(X^{\Delta t,k}(\hat{t}+i\Delta t+r;x)-X^{\Delta t,k}(\hat{t}+i\Delta t;x))\Big|\mathcal{F}_{\hat{t}+i\Delta t}\Big],\\ \int_0^{\Delta t} & a(dr)\pi_S(X^{\Delta t,k}(\hat{t}+j\Delta t+r;x)-X^{\Delta t,k}(\hat{t}+j\Delta t;x))\Big\rangle\Big]. \end{aligned}$$

Now, almost surely,

$$\begin{aligned}
&\mathbf{E}\Big[X^{\Delta t,k}(\hat{t}+i\Delta t+r;x)-X^{\Delta t,k}(\hat{t}+i\Delta t;x)\Big|\mathcal{F}_{\hat{t}+i\Delta t}\Big]\\
=&(\tilde{S}_k(r)-I)X^{\Delta t,k}(\hat{t}+i\Delta t;x)+\int_0^r \tilde{S}_k(r-s)\begin{pmatrix}C^{\Delta t,k}\\ 0\end{pmatrix}\mathcal{P}_k X^{\Delta t}(\hat{t}+i\Delta t;x)\,ds\\
&+\int_0^r \tilde{S}_k(r-s)F(X^{\Delta t,k}(\hat{t}+i\Delta t;x))\,ds.
\end{aligned}$$

Let $X^{\Delta t,k}(t_2,t_1;x)$ be the solution to (2.6) at time t_2 with initial condition x at time t_1 for $0\le t_1\le t_2\le T$. Then, $X^{\Delta t,k}(t_2,0;x)=X^{\Delta t,k}(t_2,t_1;X^{\Delta t,k}(t_1,0;x))$ expresses the Markov property. Let

$$\begin{aligned}
\Gamma_{t_2,t_1}(x):=&\int_0^{\Delta t} a(dr)\pi_S\mathbf{E}\Big[X^{\Delta t,k}(\hat{t}_2+r,\hat{t}_1;x)-X^{\Delta t,k}(\hat{t}_2,\hat{t}_1;x)\Big|\mathcal{F}_{\hat{t}_2}\Big]\\
=&\int_0^{\Delta t} a(dr)\int_0^r C^{\Delta t,k}\mathcal{P}_k X^{\Delta t,k}(\hat{t}_2,\hat{t}_1;x)\,ds+\int_0^{\Delta t} a(dr)\int_0^r \pi_S F(X^{\Delta t,k}(\hat{t}_2,\hat{t}_1;x))\,ds.
\end{aligned}$$

It is easy to show from (2.9) that for $0\le t_2-t_1\le T$

$$(\mathbf{E}\|\Gamma_{t_2,t_1}(x)-\Gamma_{t_2,t_1}(x')\|^2)^{1/2}\le K(1+\|x\|)\,\|x-x'\|\,\Delta t^2. \tag{2.12}$$

We have, dropping two integrals which are easier to bound, $|I|\le|I_{hard}|+K(1+\|x\|)^2\Delta t^4$ and

$$I_{hard}:=\mathbf{E}\Big[\Big\langle\Gamma_{t+i\Delta t,0}(x),\int_0^{\Delta t}a(dr)\int_{\hat{t}+j\Delta t}^{\hat{t}+j\Delta t+r}\pi_S B(X^{\Delta t,k}(\hat{t}+j\Delta t;x))\,d\bar{W}(s)\Big\rangle\Big].$$

We consider the case $\hat{t}+j\Delta t=0$; the general case is similar.

$$\begin{aligned}
I_{hard}=&\mathbf{E}\Big[\Big\langle\Gamma_{(i-j)\Delta t,0}(x),\int_0^{\Delta t}a(dr)\int_0^r\pi_S B(x)\,d\bar{W}(s)\Big\rangle\Big]\\
=&\mathbf{E}\Big[\Big\langle\Gamma_{(i-j)\Delta t,\Delta t}(X^{\Delta t,k}(\Delta t;x))-\Gamma_{(i-j)\Delta t,\Delta t}(\mathbf{E}X^{\Delta t,k}(\Delta t;x)),\\
&\int_0^{\Delta t}a(dr)\int_0^r\pi_S B(x)\,d\bar{W}(s)\Big\rangle\Big],
\end{aligned}$$

because for all $h\in H$ the average $\mathbf{E}\langle\Gamma_{(i-j)\Delta t,\Delta t}(h),\int_0^{\Delta t}a(dr)\int_0^r\pi_S B(x)\,d\bar{W}(s)\rangle=0$ by the independent increment property. Now, from (2.7)-(2.8) and (2.12),

$$\begin{aligned}
|I_{hard}|\le&\Big(\mathbf{E}\Big\|\Gamma_{(i-j)\Delta t,\Delta t}(X^{\Delta t,k}(\Delta t;x))-\Gamma_{(i-j)\Delta t,\Delta t}(\mathbf{E}X^{\Delta t,k}(\Delta t;x))\Big\|^2_{\mathbf{R}^d}\\
&\times\mathbf{E}\Big\|\int_0^{\Delta t}a(dr)\int_0^r\pi_S B(x)\,d\bar{W}(s)\Big\|^2_{\mathbf{R}^d}\Big)^{1/2}\\
\le&\Big(\mathbf{E}K(1+\|X^{\Delta t,k}(t;x)\|^2)\|X^{\Delta t,k}(\Delta t;x)-\mathbf{E}X^{\Delta t,k}(\Delta t;x)\|^2\Delta t^4\Delta t^3\Big)^{1/2}\\
\le&K(1+\|x\|)^2\Delta t^4.
\end{aligned}$$

QED

Lemma 2.9 *Suppose that the delay function of the initial data is globally Lipschitz,* $\|Y_D\|_{Lip} = \|\pi_D x\|_{Lip} < \infty$. *Let* $\alpha(s,r;x) := \mathcal{P}_k X^{\Delta t,k}(s;x)(r) - \mathcal{P}_k X^{\Delta t,k}(s;x)(\hat{r})$. *For* $0 \le t \le T$ *and* $-\tau \le s \le 0$,

$$\mathbf{E}\|\alpha(\hat{t},s;x)\|^2_{\mathbf{R}^d} \le K(1+\|x\|+\|\pi_D x\|_{Lip})^2\Delta t + o(k^{-1})$$

and for $-\lfloor \tau/\Delta t\rfloor \le i \ne j \le -1$

$$\begin{aligned}\mathbf{E}\Big[\Big\langle \int_{i\Delta t}^{(i+1)\Delta t} a(dr)\alpha(\hat{t},r;x), \int_{j\Delta t}^{(j+1)\Delta t} a(dr)\alpha(\hat{t},r;x)\Big\rangle\Big] \\ \le K(1+\|x\|+\|\pi_D x\|_{Lip})^2\Delta t^4 + o(k^{-1}).\end{aligned} \tag{2.13}$$

Proof To prove the Lemma, we interpret the inequalities in Lemma 2.8 for the delay function $\pi_D X^{\Delta t,k}(t,x)(\cdot)$. For small time, the delay function carries information from the initial condition as in (2.4). The Lipschitz assumptions on the initial delay function can be used to derive the required estimates for small time. For larger time, the state variable translates into the delay function as described by $X_S^{\Delta t}(t+s;x) = X_D^{\Delta t}(t;x)(s)$ for $-\tau \le s < 0$ and $t+s \ge 0$. If this statement held for the smoothed process $X^{\Delta t,k}$ and $\mathcal{P}_k = \pi_D$, the Lemma would be immediate from Lemma 2.8. We see that $\mathcal{P}_k \to \pi_D$ in $\mathcal{L}(H, L_2([-\tau,0],\mathbf{R}^d))$ by examining its definition (2.2). Now, from Lemma 2.7, we have $\mathcal{P}_k X^{\Delta t,k}(t;x)(r) \to X_S^{\Delta t}(t+r;x)$. This introduces a small error that goes to zero as k goes to infinity, which accounts for the $o(k^{-1})$ term in the final result. *QED*

3 Corollaries of the Itô calculus

We wish to apply the following Corollary to gain time regularity of functionals of $X(t;x)$ and its spatial derivatives. The corollary is set up for an abstract equation, but we have in mind the application to say $Z(t;x) = (X(t;x), X_x(t;x)h)$, which obeys

$$\begin{aligned} dZ_1 &= \Big[AZ_1 + F(Z_1)\Big]\,dt + B(Z_1)\,dW(t), \quad Z_1(0) = x \\ dZ_2 &= \Big[AZ_2 + F_x(Z_1)Z_2\Big]\,dt + B_x(Z_1)Z_2\,dW(t), \quad Z_2(0) = h. \end{aligned} \tag{3.1}$$

A similar equation can be written down for the second derivative $X_{xx}(t;x)(h,g)$ involving four equations.

Corollary 3.1 *Consider locally Lipschitz functions* $\bar{F}_i : H^m \to H$ *and* $\bar{B}_i : H^m \to L_2^0$ *for* $i = 1,\dots,m$ *such that* $\bar{F}_i(Z_1,\dots,Z_m)$ *and* $\bar{B}_i(Z_1,\dots,Z_m)$ *are independent of* $\pi_D Z_i$. *Suppose that there exists a unique strong solution* $Z^k(t;x)$ *in* H^m *of*

$$dZ_i^k = \Big[A_k Z_i^k + \bar{F}_i(Z^k)\Big]\,dt + \bar{B}_i(Z^k)\,dW, \quad Z_i^k(0) = z_i^k(x), \tag{3.2}$$

and a mild solution to

$$dZ_i = \Big[AZ_i + \bar{F}_i(Z)\Big]\,dt + \bar{B}_i(Z)\,dW, \quad Z_i(0) = z_i(x), \tag{3.3}$$

where the initial data $z_i(x), z_i^k(x) \in H$ *are parameterised by* $x \in H$. *Suppose that*

$$\sup_{0\le t\le T}\ \sup_{i=1,\dots,m} \mathbf{E}|Z_i(t;x) - Z_i^k(t;x)|_\star^2 \to 0, \quad \text{as } k \to \infty. \tag{3.4}$$

Suppose further that, for $0 \le t \le T$ and $i = 1, \dots, m$,

$$\mathbf{E}|Z_i^k(t;x)|_\star^p \le K(1 + \|x\|^p) \tag{3.5}$$

and that $\bar{F}_i$ and $\bar{B}_i$ satisfy

$$\mathbf{E}\|\bar{F}_i(Z^k(t;x))\|^2 \le K(1 + \|x\|)^2, \qquad \mathbf{E}\|\bar{B}_i(Z^k(t;x))\|_{HS}^2 \le K. \tag{3.6}$$

Consider continuously differentiable $G\colon H^m \to \mathbf{R}$ such that $G(Z_1, \dots, Z_m)$ is independent of $\pi_D Z_i$ and the first derivatives G_i and second derivatives G_{ij} obey

$$|G_i(Z)| \le K\Big(1 + \sum_{\ell=1}^m \|Z_\ell\|_S^{p-1}\Big), \qquad |G_{ij}(Z)| \le K\Big(1 + \sum_{\ell=1}^m \|Z_\ell\|_S^{p-2}\Big). \tag{3.7}$$

Let $w(t,x) := \mathbf{E}G(Z(t;x))$ and $w^k(t,x) := \mathbf{E}G(Z^k(t;x))$. Then, w_t and w_t^k are uniformly continuously differentiable in time on bounded subsets of $\mathbf{R}^+ \times H$ and

$$|w_t^k(t,x)|, \quad |w_t(t,x)| \le K(1 + \|x\|^p), \qquad 0 \le t \le T.$$

Proof Let $w^k(t,x) := \mathbf{E}G(Z^k(t;x))$. Because G is continuously differentiable and $Z^k(t;x)$ is a strong solution, the Itô formula implies that

$$\begin{aligned} w^k(t,x) - w^k(0,x) =& \mathbf{E}\sum_{i=1}^m \int_0^t G_i(Z^k(s;x))(A_k Z_i^k(s;x) + \bar{F}_i(Z^k(s;x)))\,ds \\ &+ \tfrac{1}{2}\sum_{i,j=1}^m \mathbf{E}\int_0^t \operatorname{Tr} G_{ij}(Z^k(s;x))\bar{B}_i(Z^k(s;x))\bar{B}_j(Z^k(s;x))^*\,ds. \end{aligned}$$

We attain limits from the dominated convergence theorem because, under (3.5) and (3.7),

$$\begin{aligned} \mathbf{E}\Big|\int_0^t G_i(Z^k(s;x))(A_k Z_i^k(s;x) + \bar{F}_i(Z^k(s;x)))\,ds\Big| \le& K\,t\,(1 + \|x\|^p) \\ \tfrac{1}{2}\mathbf{E}\Big|\int_0^t \operatorname{Tr} G_{ij}(Z^k(s))\bar{B}_i(Z^k(s;x))\bar{B}_j(Z^k(s;x))^*\,ds\Big| \le& K\,t\,(1 + \|x\|^{p-2}). \end{aligned}$$

Thus,

$$\begin{aligned} w_t^k(t,x) =& \mathbf{E}\sum_{i=1}^m G_i(Z^k(t;x))(A_k Z_i^k(t;x) + \bar{F}_i(Z^k(t;x))) \\ &+ \tfrac{1}{2}\sum_{i,j=1}^m \operatorname{Tr} G_{ij}(Z^k(t;x))\bar{B}_i(Z^k(t;x))\bar{B}_j(Z^k(t;x))^*. \end{aligned}$$

The convergence of Z_i^k to Z_i in $|\cdot|_\star$ implies the convergence of each term in the limit $k \to \infty$. This convergence depends on G, $\bar{F}_i$, and $\bar{B}_i$ being independent of the the delay part of the space H^m so that Lemma 2.5 can be applied. The second component of $A_k Z^k(t;x)$ does not converge in H. We now have

$$\begin{aligned} w_t(t,x) =& \mathbf{E}\sum_{i=1}^m G_i(Z(t;x))(AZ_i(t;x) + \bar{F}_i(Z(t;x))) \\ &+ \tfrac{1}{2}\sum_{i,j=1}^m \operatorname{Tr} \phi_{ij}(Z(t;x))\bar{B}_i(Z(t;x))\bar{B}_j(Z(t;x))^*. \end{aligned} \tag{3.8}$$

¿From this expression, it is easy to derive the required growth bound on $w_t(t,x)$ in $\|x\|$. Similar estimates hold for w_t^k with bounds uniform in $k \to \infty$.

We now turn to establishing uniform continuity of $w_t(t,x)$ with respect to time (the analysis for w_t^k is similar). Consider

$$S(t)x - x = A \int_0^t S(s)x\, ds, \qquad x \in H.$$

Hence, using Hypothesis 1.1(iii),

$$\begin{aligned}|S(t)x - x|_\star \leq & K \Big\| \pi_D \int_0^t S(s)x\, ds \Big\|_{L_2([-\tau,0],\mathbf{R}^d)} + \Big\| \int_{-\tau}^0 a(ds) \frac{d}{ds} \Big(\int_0^t \pi_D S(s')x\, ds' \Big)(s) \Big\|_{\mathbf{R}^d} \\ \leq & K\, t\, \|x\|.\end{aligned}$$

It follows easily that Z is uniformly continuous in time in the following sense: for $R, T > 0$, there exists K with

$$\mathbf{E}|Z_i(t;x) - Z_i(t';x)|_\star^2 \leq K|t-t'|, \quad 0 \leq t, t' \leq T, \quad \|x\| \leq R. \tag{3.9}$$

In particular $\pi_S AZ(t;x)$ is uniformly continuous in time on bounded subsets of H.

Fix R the radius of a ball in H and consider $x \in H$ with $\|x\| \leq R$. For any $\delta > 0$, there exists L large by (3.5) and the Chebyshev inequality so that if $\mathcal{O} := \{\|Z_i(t;x)\|_S \leq L, 0 \leq t \leq T, i = 1, \dots, m\}$, the probability $\mathbf{P}(\mathcal{O}) \geq 1 - \delta$. Consider the expectations defining w_t in (3.8) split as a sum over $\mathcal{O}$ and $\mathcal{O}^c$. On the set $\mathcal{O}$, $G_i, G_{ij}, \bar{F}, \bar{B}$ are all Lipschitz and the expectations in the difference $w_t(t,x) - w_t(t',x)$ may be bounded by $K|t-t'|^{1/2}$ using (3.9). By using (3.5), the expectations on the set $\mathcal{O}^c$ are bounded by $\delta |w_t(t,x)| \leq \delta K(1+R^p)$. Thus to show uniform continuity on the bounded set of H of radius R, pick L large enough that $\delta K(1+R^p) < \epsilon/2$ (a bound on the integral over $\mathcal{O}^c$). Then, for $|t-t'| \leq \epsilon^2/2K^2$ and $0 \leq t, t' \leq T$,

$$|w_t(t,x) - w_t(t',x)| \leq \epsilon/2 + \epsilon/2, \text{ if } \|x\| \leq R.$$

This gives uniform continuity of w_t on bounded subsets of $\mathbf{R}^+ \times H$. *QED*

The following Lemma gives an order Δt estimate on a functional of the numerical interpolant $X^{\Delta t,k}$.

Lemma 3.2 *Consider the strong solution $X^{\Delta t,k}(t;x)$ of (2.6) under the condition that F is globally Lipschitz and that B is bounded. Consider a function $w\colon \mathbf{R}^+ \times H \to \mathbf{R}$ with one time and two spatial derivatives that are uniformly continuous on bounded subsets of $\mathbf{R}^+ \times H$. Further suppose, for $0 \leq t \leq T$ and for a constant K, that*

$$|w_t(t,x)| \leq K(1+\|x\|^p), \tag{3.10}$$

$$\|w_x(t,x)\|_{\mathcal{L}(H,\mathbf{R})}, \quad \|w_{xx}(t,x)\|_{\mathcal{L}(H\times H,\mathbf{R})} \leq K(1+\|x\|^{p-1}), \tag{3.11}$$

and that for $h \in H$ the following holds uniformly in k

$$|w_x(t,x)\tilde{A}_k h| \leq K(1+\|x\|^{p-1})\, \|h\|. \tag{3.12}$$

Then, the following bound holds uniformly in k,

$$\Big| \mathbf{E}\Big[w(s, X^{\Delta t,k}(s;x)) - w(\hat{s}, X^{\Delta t,k}(\hat{s};x)) \Big] \Big| \leq K(1+\|x\|^p)\Delta t, \quad 0 \leq s \leq T.$$

Proof This is Lemma 14.1.6 of [10]. Apply the Itô formula to the strong solution $X^{\Delta t,k}$:

$$\begin{aligned}
&\mathbf{E}\Big[w(s, X^{\Delta t,k}(s;x)) - w(\hat{s}, X^{\Delta t,k}(\hat{s};x))\Big]\\
&\quad=\mathbf{E}\Big[\int_{\hat{s}}^{s}\Big\{w_t(s', X^{\Delta t,k}(s';x))\\
&\qquad+ w_x(s', X^{\Delta t,k}(s';x))\Big(\tilde{A}_k X^{\Delta t,k}(\hat{s};x) + \begin{pmatrix} C^{\Delta t}\mathcal{P}_k X^{\Delta t}(\hat{s};x)\\ 0\end{pmatrix} + F(X^{\Delta t,k}(\hat{s};x))\Big)\\
&\qquad+ \tfrac{1}{2}\operatorname{Tr} w_{xx}(s', X^{\Delta t,k}(s';x))B(X^{\Delta t,k}(\hat{s};x))B(X^{\Delta t,k}(\hat{s};x))^*\Big\}\,ds'\Big].
\end{aligned}$$

Now using (3.10)–(3.12) with the boundedness of B and the Lipschitz property of F, we have

$$\begin{aligned}
&\Big|\mathbf{E}\Big[w(s, X^{\Delta t,k}(s;x)) - w(\hat{s}, X^{\Delta t,k}(\hat{s};x))\Big]\Big|\\
&\quad\le\mathbf{E}\Big[\int_{\hat{s}}^{s} K(1+\|X^{\Delta t,k}(s';x)\|^p) + K(1+\|X^{\Delta t,k}(s';x)\|^{p-1})(1+\|X^{\Delta t,k}(\hat{s};x)\|)\,ds'\Big].
\end{aligned}$$

By using (2.7), we have

$$\begin{aligned}
&\Big|\mathbf{E}\Big[w(s, X^{\Delta t,k}(s;x)) - w(\hat{s}, X^{\Delta t,k}(\hat{s};x))\Big]\Big|\\
&\quad\le\mathbf{E}\int_{\hat{s}}^{s} K(1+\|X^{\Delta t,k}(\hat{s};x)\|^{p-1})(1+\mathbf{E}\|X^{\Delta t,k}(\hat{s};x)\|)\,ds'\\
&\quad\le\int_{\hat{s}}^{s} K(1+\|x\|^p)\,ds'.
\end{aligned}$$

As $|s-\hat{s}| \le \Delta t$, this completes the proof. *QED*

4 The Kolmogorov equation

We introduce the Kolmogorov equation for the stochastic evolution equation (2.1). The background theory is developed in Da Prato and Zabczyk [5] (1.1), where further references are also given. The Kolmogorov equation is described in Theorem 4.2. We also discuss the regularity of the terms in the equation so that the Itô formula applies to $v(t,x) = \mathbf{E}\phi(X(t;x))$ and to the terms in the Kolmogorov equation. Throughout this section, we assume that Hypothesis 1.1 holds.

Lemma 4.1 *Let $\xi^k(t,x) := X^k_x(t;x)A_k h$, where $X^k(t;x)$ is the strong solution to (2.3). Then, for $p \ge 2$,*

$$\sup_{0\le t\le T}\mathbf{E}|\xi^{A_k h}(t;x) - \xi^{A_\ell h}(t;x)|^p_\star \to 0 \qquad \text{as } k,\ell\to\infty. \tag{4.1}$$

For $0 \le t \le T$,

$$\lim_{k\to\infty}(\mathbf{E}|\xi^{A_k h}(t;x)|^p_\star)^{1/p} \le K\|h\|. \tag{4.2}$$

Moreover, the limit of $\pi_S X^k_x(t;x)A_k^2 x$ exists with respect to $\|\cdot\|_S$ and

$$\lim_{k\to\infty}(\mathbf{E}\|X^k_x(t;x)A_k^2 x\|^p_S)^{1/p} \le K\|x\|. \tag{4.3}$$

Proof ξ^k is a strong solution of

$$d\xi^k(t;x) = \Big[A_k\xi^k(t;x) + F_x(X^k(t;x))\xi^k(t;x)\Big]\,dt + B_x(X^k(t;x))\xi^k(t;x)\,dW(t),$$

with initial condition $\xi^k(0) = A_k h$. The variation of constants formula:

$$\begin{aligned}\xi^k(t;x) =& S_k(t)A_k h + \int_0^t S_k(t-s)F_x(X^k(s;x))\xi^k(s;x)\,ds \\ &+ \int_0^t S_k(t-s)B_x(X^k(s;x))\xi^k(s;x)\,dW(s).\end{aligned}$$

Using the fact that $|X|_\star \le K\|X\|$ on the stochastic integral, the Burkholder-Davis-Gundy inequality gives

$$\begin{aligned}(\mathbf{E}|\xi^k(t;x)|_\star^p)^{1/p} \le& |A_k S_k(t)h|_\star + \Big(\mathbf{E}\Big[\int_0^t |S_k(t-s)F_x(X^k(s;x))\xi^k(s;x)|_\star^p\,ds\Big]\Big)^{1/p} \\ &+ \Big(\mathbf{E}\Big[\int_0^t \|S_k(t-s)B_x(X^k(s;x))\xi^k(s;x)\|_{HS}^2\,ds\Big]^{p/2}\Big)^{1/p} \\ \le& |A_k S_k(t)h|_\star + \Big(\mathbf{E}\Big[\int_0^t |S_k(t-s)F_x(X^k(s;x))\xi^k(s;x)|_\star^p\,ds\Big]\Big)^{1/p} \\ &+ K\Big(\mathbf{E}\int_0^t \|S_k(t-s)B_x(X^k(s;x))\xi^k(s;x)\|_{HS}^p\,ds\Big)^{1/p}.\end{aligned}$$

N.B., $A_k S_k(t) = \frac{d}{dt}S_k(t)$ so that by using Lemma 2.5(iii) and Hypothesis 1.1(iii)

$$\begin{aligned}|A_k S_k(t)h|_\star =& \|A_k S_k(t)h\|_S + \Big\|\int_{-\tau}^0 a(ds)(\pi_D A_k S_k(t)h)(s)\Big\|_{\mathbf{R}^d} \\ \le& K\|S_k(t)h\| + \Big\|\int_{-\tau}^0 a(ds)\frac{d}{ds}(\pi_D S_k(t)h)(s)\Big\|_{\mathbf{R}^d} \\ \le& K\|h\| + K\|\pi_D S_k(t)h\|_{L_2([-\tau,0],\mathbf{R}^d)} \\ \le& K\|h\|.\end{aligned}$$

Thus, using the boundedness of F_x and B_x,

$$(\mathbf{E}|\xi^k(t;x)|_\star^p)^{1/p} \le K\|h\| + K\Big(\int_0^t \mathbf{E}|\xi^k(s;x)|_\star^p\,ds\Big)^{1/p}.$$

Note in particular that the choice of K can be made independent of the particular Yosida approximation A_k. By applying the Gronwall Lemma, for each $T > 0$, there exists $K > 0$ such that for each k

$$\mathbf{E}|\xi_k(t;x)|_\star^p \le K\|h\|^p, \quad 0 \le t \le T. \tag{4.4}$$

Note that the is uniform in k and gives the estimate (4.2).

We show that the sequence in Cauchy with respect to $|\cdot|_\star$ for $p=2$. Consider

$$\begin{aligned}
&\mathbf{E}|\xi^k(t;x)-\xi^\ell(t;x)|_\star^2 \le K\alpha_{k\ell}(t;h)^2\\
&\quad + K\mathbf{E}\int_0^t |S_k(t-s)F_x(X^k(s;x))\xi^k(s;x)-S_\ell(t-s)F_x(X^k(s;x))\xi^k(s;x))|_\star^2\,ds\\
&\quad + K\mathbf{E}\int_0^t |S_\ell(t-s)F_x(X^k(s;x))\xi^k(s;x)-S_\ell(t-s)F_x(X^k(s;x))\xi^\ell(s;x))|_\star^2\,ds\\
&\quad + K\mathbf{E}\int_0^t |S_\ell(t-s)F_x(X^k(s;x))\xi^\ell(s;x)-S_\ell(t-s)F_x(X^\ell(s;x))\xi^\ell(s;x))|_\star^2\,ds\\
&\quad + K\mathbf{E}\int_0^t \|S_k(t-s)B_x(X^k(s;x))\xi^k(s;x)-S_\ell(t-s)B_x(X^k(s;x))\xi^k(s;x)\|_{HS}^2\,ds\\
&\quad + K\mathbf{E}\int_0^t \|S_\ell(t-s)B_x(X^k(s;x))\xi^k(s;x)-S_\ell(t-s)B_x(X^k(s;x))\xi^\ell(s;x)\|_{HS}^2\,ds\\
&\quad + K\mathbf{E}\int_0^t \|S_\ell(t-s)B_x(X^k(s;x))\xi^\ell(s;x)-S_\ell(t-s)B_x(X^\ell(s;x))\xi^\ell(s;x)\|_{HS}^2\,ds
\end{aligned}$$

where

$$\alpha_{k\ell}(t;h) := |A_kS_k(t)h - A_\ell S_\ell(t)h|_\star.$$

Note that

$$\begin{aligned}
&\alpha_{k\ell}(t;h) \le \|A_k(S_k(t)-S_\ell(t))h\|_S + \|(A_k-A_\ell)S_\ell(t)h\|_S\\
&+\Big\|\int_{-\tau}^0 a(ds)(\pi_D A_k(S_k(t)-S_\ell(t))h)(s)\Big\|_{\mathbf{R}^d} + \Big\|\int_{-\tau}^0 a(ds)(\pi_D(A_k-A_\ell)S_\ell(t)h)(s)\Big\|_{\mathbf{R}^d}.
\end{aligned}$$

By using Hypothesis 1.1(iii) and the definition of A_k in (2.2), it is possible to show that $\alpha_{k\ell}(t;h)\to 0$ as $k,\ell\to\infty$.

For $0\le t\le T$, we have that $S_k(t)$ is a bounded operator from H to H. Temporarily dropping the $(s;x)$ argument on X and ξ, we have

$$|S_\ell(t-s)F_x(X^k)\xi^k - S_\ell(t-s)F_x(X^k)\xi^\ell|_\star \le K\|\xi^k-\xi^\ell\|_S.$$

Similarly,

$$\|S_\ell(t-s)B_x(X^k)\xi^k - S_\ell(t-s)B_x(X^k)\xi^\ell\|_{HS} \le K\|\xi^k-\xi^\ell\|_S.$$

We also have that

$$|S_\ell(t-s)F_x(X^k)\xi^k - S_\ell(t-s)F_x(X^\ell)\xi^k|_\star \le K\|X^k-X^\ell\|_S\|\xi^k\|_S$$

and

$$|S_k(t-s)F_x(X^k)\xi^k - S_\ell(t-s)F_x(X^k)\xi^k|_\star \le \|(S_k(t-s)-S_\ell(t-s))F_x(X^k)\|_{\mathcal{L}(H,H)}\|\xi^k\|.$$

After writing the similar expressions involving B, we find that for $0\le t\le T$,

$$\begin{aligned}
&\mathbf{E}|\xi^k(t;x)-\xi^\ell(t;x)|_\star^2 \le K\alpha_{k\ell}(t;h)^2 + K\int_0^t \mathbf{E}\|\xi^k(s;x)-\xi^\ell(s;x)\|_P^2\,ds\\
&+ K\int_0^t\Big(\mathbf{E}\|X^k(s;x)-X^\ell(s;x)\|^4\Big)^{1/2}ds\\
&+ K\int_0^t\Big(\mathbf{E}\|(S_k(t-s)-S_\ell(t-s))F_x(X^k(s;x))\|_{\mathcal{L}(H,H)}^4\Big)^{1/2}ds\\
&+ K\int_0^t\Big(\mathbf{E}\|(S_k(t-s)-S_\ell(t-s))B_x(X^k(s;x))\|_{\mathcal{L}(H,L_2^0)}^4\Big)^{1/2}ds,
\end{aligned}$$

using boundedness of ξ^k in $\mathbf{E}\|\cdot\|^4$. Note that $\|(S_k(t)-S_\ell(t))F_x(h)\|_{\mathcal{L}(H,H)} \leq K\|(A_k - A_\ell)F_x(h)\|_{\mathcal{L}(H,H)} \to 0$ uniformly in $h \in H$ as $F_x(h)$ is bounded and equal to zero in the second component. Hence Gronwall's lemma and Lemma 2.6 applies to give convergence of $\xi^k(t;x)$ in the sense of (4.1).

In a similar way, it is easy to establish the limit for $X_x^k(t;x)A_k^2 x$ with respect to $\|\cdot\|_S$ by exploiting

$$\|A_k^2 S_k(t)h\|_S = \|A_k \tfrac{d}{dt} S_k(t)h\|_S \leq K\|C\tfrac{d}{dt}\|_{\mathcal{L}(L_2([-\tau,0],\mathbf{R}^d),\mathbf{R}^d)}\ \|\pi_D S_k(t)h\|_{L_2([-\tau,0],\mathbf{R}^d)}.$$

The bound $\mathbf{E}\|X_x^k(t;x)A^2x\|_S \leq K\|x\|$ follows easily.

QED

Theorem 4.2 *Let $\phi \in \mathcal{G}_p$ and set $v(t,x) := \mathbf{E}\phi(X(t;x))$. Then v satisfies for $x \in H$ and $0 \leq t \leq T$*

$$v_t(t,x) = \tfrac{1}{2}\operatorname{Tr}\Big[v_{xx}(t,x)B(x)B(x)^*\Big] + v_x(t,x)Ax + v_x(t,x)F(x)$$

where

$$v_x(t,x)Ax := \lim_{k\to\infty} \mathbf{E}\phi'(X(t;x))X_x^k(t;x)A_kx.$$

The functional v is two times in space and one time in time uniformly continuously differentiable on bounded subsets of $\mathbf{R}^+ \times H$.

Proof Apply Itô's formula to X^k with the function ϕ:

$$\begin{aligned}\mathbf{E}\phi(X^k(t;x)) =& \mathbf{E}\phi(X^k(s;x)) \\ &+ \mathbf{E}\Big[\int_s^t \phi'(X^k(s';x))A_kX^k(s';x) + \phi'(X^k(s;x))F(X^k(s';x))\,ds'\Big] \\ &+ \tfrac{1}{2}\mathbf{E}\Big[\int_s^t \operatorname{Tr}\phi''(X^k(s';x))B(X^k(s';x))B(X^k(s';x))^*\,ds'\Big].\end{aligned}$$

By hypothesis on ϕ, F, and B,

$$\mathbf{E}\Big|\int_s^t \lim_{k\to\infty} \phi'(X^k(s';x))(A_kX^k(s';x) + F(X^k(s';x)))\,ds'\Big| \leq K(1+\|x\|^p)\,(t-s)$$
$$\mathbf{E}\Big|\int_s^t \operatorname{Tr}\phi''(X(s';x))B(X^k(s';x))B^*(X^k(s';x))\,ds'\Big| \leq K(1+\|x\|^{p-2})\,(t-s).$$

Thus, dominated convergence applies, to give

$$\begin{aligned}\frac{\partial}{\partial t}\mathbf{E}\phi(X^k(t;x)) =& \mathbf{E}\phi'(X^k(t;x)), A_kX^k(t;x) + \mathbf{E}\phi'(X^k(t;x))F(X^k(t;x)) \\ &+ \tfrac{1}{2}\mathbf{E}\operatorname{Tr}\phi''(X^k(t;x))B(X^k(t;x))B(X^k(t;x))^*.\end{aligned}$$

Now standard arguments apply to give the Kolmogorov equation for the process $X^k(t;x)$: $v^k(t;x) := \mathbf{E}\phi(X^k(t;x))$ obeys

$$v_t^k(t,x) = \tfrac{1}{2}\operatorname{Tr}\Big[v_{xx}^k(t,x)B(x)B(x)^*\Big] + v_x^k(t,x)A_kx + v_x^k(t,x)F(x),$$

where

$$v_x^k(t,x)A_kx = \mathbf{E}\phi'(X^k(t;x))X_x^k(t;x)A_kx.$$

The limits in each term converges as $k \to \infty$. The only difficult convergence is that of v_x^k, which exists by Lemma 4.1. For convenience, we replace $X^k(t;x)$ by $X(t;x)$ in the $\phi'(\cdot)$ term in the definition $v_x(t;x)$.

The spatial regularity is described in Corollary 2.4. To establish time regularity, apply Corollary 3.1 with $Z(t;x) = X(t;x)$ and $Z^k(t;x) = X^k(t;x)$. Theorem 2.3 certainly gives convergence of X_k to X with respect to $|\cdot|_\star$. *QED*

Lemma 4.3 *Let* $v(t,x) := \mathbf{E}\phi(X^k(t;x))$ *where* $\phi \in \mathcal{G}_p$.

(i) Consider a function $\psi\colon H \to H$ *that is globally Lipschitz with two uniformly continuous derivatives. Let* $w(t,x) := v_x(t,x)\psi(x)$. *Then* w_t, w_x, *and* w_{xx} *exist and are uniformly continuous on bounded subsets of* $\mathbf{R}^+ \times H$ *such that, for a constant* K *independent of* k, $\|w_t(t,x)\|$ *is bounded by* $K(1+\|x\|^p)$ *and*

$$\|w_x(t,x)\|_{\mathcal{L}(H,\mathbf{R})}, \qquad \|w_{xx}(t,x)\|_{\mathcal{L}(H\times H,\mathbf{R})} \le K(1+\|x\|^{p-1}).$$

(ii) Consider a function $\Psi\colon H \to \mathcal{L}(\mathbf{R}^d, H)$ *that is bounded with two uniformly continuous derivatives. Let* $w(t,x) = \operatorname{Tr} v_{xx}(t,x)\Psi(x)\Psi^*(x)$. *Then* w_t, w_x, *and* w_{xx} *exist and are uniformly continuous on bounded subsets of* $\mathbf{R}^+ \times H$. *For a constant* K *independent of* k, $\|w_t(t,x)\|$ *is bounded by* $K(1+\|x\|^p)$ *and*

$$\|w_x(t,x)\|_{\mathcal{L}(H,\mathbf{R})}, \quad \|w_{xx}(t,x)\|_{\mathcal{L}(H\times H,\mathbf{R})} \le K(1+\|x\|^{p-1}).$$

Proof The differentiability and bounds of the derivatives in x follow from the hypothesis on ψ, Ψ together with Corollary 2.4. To understand the time derivative, argue as follows:

(i) First note that $v_x(t,x) = \mathbf{E}\phi'(X^k(t;x))X_x^k(t;x)$. Thus,

$$v_x(t,x)\psi(x) = \mathbf{E}\phi'(X^k(t;x))X_x^k(t;x)\psi(x) = \mathbf{E}G(Z_1^k(t;x), Z_2^k(t;x)),$$

where $G(Z_1, Z_2) = \phi'(Z_1)Z_2$, $Z_1^k(t;x) = X^k(t;x)$, and $Z_2^k(t;x) = X_x^k(t;x)\psi(x)$. Corollary 3.1 applies in this situation as (Z_1, Z_2) satisfies (3.1) with $h = \psi(x)$. The growth condition (3.5) is given by Theorem 2.3. The coefficients in (3.1) are locally Lipschitz and obey (3.6) by using the boundedness of the derivatives given in Hypothesis 1.1. The regularity of test functional G is easily derived from the conditions on ϕ. Thus, we conclude that $v_x(t,x)\psi(x)$ is uniformly continuously differentiable in time on bounded subsets of $\mathbf{R}^+ \times H$.

(ii) Similarly, for $h_1, h_2 \in H$,

$$\begin{aligned} v_{xx}(t,x)(h_1,h_2) =& \mathbf{E}\phi''(X^k(t;x))(X_x^k(t;x)h_1, X_x^k(t;x)h_2) \\ &+ \mathbf{E}\phi'(X^k(t;x))X_{xx}^k(t;x)(h_1,h_2). \end{aligned}$$

Thus,

$$\begin{aligned} v_{xx}(t,x)(\Psi(x)h_1, \Psi(x)h_2) =& \mathbf{E}\phi''(X^k(t;x))(X_x^k(t;x)\Psi(x)h_1, X_x^k(t;x)\Psi(x)h_2) \\ &+ \mathbf{E}\phi'(X^k(t;x))X_{xx}^k(t;x)(\Psi(x)h_1, \Psi(x)h_2). \end{aligned}$$

Let e_i be an orthonormal basis for $\mathbf{R}^d$, so that

$$\begin{aligned} \operatorname{Tr} v_{xx}(t,x)\Psi(x)\Psi^*(x) =& \mathbf{E}\Big[\sum_{i=1}^d \phi''(X^k(t;x))(X_x^k(t;x)\Psi(x)e_i, X_x^k(t;x)\Psi(x)e_i) \\ &+ \phi'(X^k(t;x))X_{xx}^k(t)(\Psi(x)e_i, \Psi(x)e_i)\Big] \\ =& \sum_{i=1}^d \mathbf{E}G(Z_1^k(t;x), Z_{2,i}^k(t;x), Z_{3,i}^k(t;x)) \end{aligned}$$

for $G(Z_1, Z_2, Z_3) := \phi''(Z_1)(Z_2, Z_2) + \phi'(Z_1)Z_3$ and

$$Z_1^k(t;x) = X^k(t;x), \quad Z_{2,i}^k(t;x) = X_x^k(t;x)\Psi(x)e_i, \quad Z_{3,i}^k(t;x) = X_{xx}^k(t;x)(\Psi(x)e_i, \Psi(x)e_i).$$

Again, it can be shown that the processes $Z_1^k, Z_{2,i}^k, Z_{3,i}^k$ satisfy the hypothesis of Corollary 3.1 (for each i, case $m = 3$). The sum is finite, which means regularity of $\mathbf{E}G$ for each i gives the same regularity for $\operatorname{Tr} v_{xx}\Psi\Psi^*$. *QED*

Lemma 4.4 *For $h, g \in H$ and $0 \le t \le T$,*

$$(\mathbf{E}\|X_{xx}(t;x)(h,g)\|_S^2)^{1/2} \le K\|g\|\,\|h\|$$

and for k large

$$(\mathbf{E}\|X_{xx}(t;x)(A_k h, A_k g)\|_S^2)^{1/2} \le K\|g\|\,\|h\|.$$

Proof Denote $X_x(t;x)h$ by $\xi^h(t;x)$ and $X_{xx}(t;x)(h,g)$ by $\eta^{h,g}(t;x)$. $\eta^{h,g}$ satisfies the following variational equation:

$$\begin{aligned} d\eta^{h,g} =& \Big[A\eta^{h,g} + F_{xx}(X)\xi^h\xi^g + F_x(X)\eta^{h,g}\Big]\,dt + (B_{xx}(X)\xi^h\xi^g + B_x(X)\eta^{h,g})\,dW \\ \eta^{h,g}(0) =& 0. \end{aligned}$$

Again the variation of constants formula can be studied to gain a bound on $\eta^{h,g}$ using bounds on ξ and X:

$$\begin{aligned} \eta^{h,g}(t;x) =& \int_0^t S(t-s)\Big(F_{xx}(X(t;x))(\xi^h(t;x), \xi^g(t;x)) + F_x(X(t;x))\eta^{h,g}(t;x)\Big)\,ds \\ &+ \int_0^t S(t-s)\Big(B_{xx}(X(t;x))(\xi^h(t;x), \xi^g(t;x)) + B_x(X(t;x))\eta^{h,g}(t;x)\Big)\,dW(s). \end{aligned}$$

Thus, by the Burkholder-Davis-Gundy inequality,

$$\begin{aligned} &\Big(\mathbf{E}\Big[\|\eta^{h,g}(t;x)\|_S^p\Big]\Big)^{1/p} \\ &\le K\Big(\mathbf{E}\Big[\int_0^t \|f_{xx}(\pi_S X(s;x))\|_{\mathcal{L}(\mathbf{R}^d\times\mathbf{R}^d,\mathbf{R}^d)}^p\,\|\xi^h(s;x)\|_S^p\,\|\xi^g(s;x)\|_S^p\,ds\Big]\Big)^{1/p} \\ &\quad + K\Big(\mathbf{E}\Big[\int_0^t \|f_x(\pi_S X(s;x))\|_{\mathcal{L}(\mathbf{R}^d,\mathbf{R}^d)}^p\,\|\eta^{h,g}(s;x)\|_S^p\,ds\Big]\Big)^{1/p} \\ &\quad + K\Big(\mathbf{E}\Big[\int_0^t \|b_{xx}(\pi_S X(s;x))\|_{\mathcal{L}(\mathbf{R}^d\times\mathbf{R}^d,L_2^0)}^p\,\|\xi^h(s;x)\|_S^p\|\xi^g(s;x)\|_S^p\,ds\Big]\Big)^{1/p} \\ &\quad + K\Big(\mathbf{E}\Big[\int_0^t \|b_x(\pi_S X(s;x))\|_{\mathcal{L}(\mathbf{R}^d,L_2^0)}^p\,\|\eta^{h,g}(s;x)\|_S^p\,ds\Big]\Big)^{1/p}. \end{aligned}$$

Hence,

$$\begin{aligned} \Big(\mathbf{E}\|\eta^{h,g}(t;x)\|_S^p\Big)^{1/p} \le& K\Big(\int_0^t \mathbf{E}\|\xi^h(s;x)\|_S^p\,\|\xi^g(s;x)\|_S^p\,ds\Big)^{1/p} \\ &+ K\Big(\int_0^t \mathbf{E}\|\eta^{h,g}(s;x)\|_S^p\,ds\Big)^{1/p} \\ &+ K\Big(\int_0^t \mathbf{E}\|\xi^h(s;x)\|_S^p\|\xi^g(s;x)\|_S^p\,ds\Big)^{1/p} \\ &+ K\Big(\int_0^t \mathbf{E}\|\eta^{h,g}(s;x)\|_S^p\,ds\Big)^{1/p}. \end{aligned}$$

Apply the estimate on ξ^h in Theorem 2.3 with $\mathbf{E}\|\xi^h\|^p\|\xi^g\|^p \le (\mathbf{E}\|\xi^h\|^{2p})^{1/2}(\mathbf{E}\|\xi^g\|^{2p})^{1/2}$, to derive (recalling that F_x, B_x, B_{xx} are bounded) for $0 \le t \le T$

$$(\mathbf{E}\|\eta^{h,g}(t;x)\|_S^p)^{1/p} \le K\ \|h\|_S\ \|g\|_S + K\Big(\int_0^t \mathbf{E}\|\eta^{h,g}(s;x)\|_S^p\ ds\Big)^{1/p}.$$

Apply Gronwall's inequality to prove that $\mathbf{E}\|\eta^{h,g}(t;x)\|_S^p \le K\|g\|_S^p\ \|h\|_S^p$.

By making use on the bound on $\xi^{A_k h}$ in Lemma 4.1, we may replace either h or g by $A_k g$ or $A_k h$ in the definition of η and repeat the argument to prove the second inequality in the Lemma. *QED*

5 Weak convergence

The following argument gives weak convergence of order Δt of the forward Euler method. The argument follows that of Kloeden-Platen [10], Theorem 14.1.5.

Proof (of Theorem 1.2) Consider $v^k(t,x) := \mathbf{E}(\phi(X^k(T-t;x))$ and

$$\mathcal{L}^k v(t,x) := v_t(t,x) + \tfrac{1}{2}\operatorname{Tr}\Big[v_{xx}(t,x)B(x)B(x)^*\Big] + v_x(t,x)A_k x + v_x(t,x)F(x).$$

As in Theorem 4.2, we have that $\mathcal{L}v^k(t,x) = 0$ and that v^k satisfies the hypothesis of Itô's formula. Apply the Itô formula to the approximations $X^{\Delta t,k}$ defined in (2.6):

$$\begin{aligned}
&v^k(T, X^{\Delta t,k}(T;x)) - v^k(0, X^{\Delta t,k}(0;x))\\
=&\mathbf{E}\Big[\int_0^T \Big\{v_x^k(s, X^{\Delta t,k}(s;x))\Big(\tilde{A}_k X^{\Delta t,k}(s;x) + \begin{pmatrix} C^{\Delta t}\\ 0\end{pmatrix}\mathcal{P}_k X^{\Delta t,k}(\hat{s};x)\Big)\\
&\quad + v_x^k(s, X^{\Delta t,k}(s;x))F(X^{\Delta t,k}(\hat{s};x))\\
&\quad + \tfrac{1}{2}\operatorname{Tr}\Big[v_{xx}^k(s, X^{\Delta t,k}(s;x))B(X^{\Delta t,k}(\hat{s};x))B(X^{\Delta t,k}(\hat{s};x))^*\Big]\\
&\quad + v_t^k(s, X^{\Delta t,k}(s;x))\Big\}\ ds\Big]
\end{aligned}$$

(subtracting off $0 = \mathcal{L}^k v^k$)

$$\begin{aligned}
=&\mathbf{E}\Big[\int_0^T \tfrac{1}{2}\operatorname{Tr}\Big[v_{xx}^k(s, X^{\Delta t,k}(s;x))B(X^{\Delta t,k}(\hat{s};x))B(X^{\Delta t,k}(\hat{s};x))^*\Big]\\
&\quad - \tfrac{1}{2}\operatorname{Tr}\Big[v_{xx}^k(s, X^{\Delta t,k}(s;x))B(X^{\Delta t,k}(s;x))B(X^{\Delta t,k}(s;x))^*\Big]\\
&\quad + v_x^k(s, X^{\Delta t,k}(s;x))\Big([C^{\Delta t}\mathcal{P}_k X^{\Delta t,k}(\hat{s};x), 0]^T + F(X^{\Delta t,k}(\hat{s};x))\Big)\\
&\quad - v_x^k(s, X^{\Delta t,k}(s;x))\Big([C\mathcal{P}_k X^{\Delta t,k}(s;x), 0]^T + F(X^{\Delta t,k}(s;x))\Big)\ ds\Big].
\end{aligned}$$

Define for $h_1, h_2 \in H$

$$\begin{aligned}
w_1(t, h_1; h_2) :=& v_x^k(t, h_1)[C\mathcal{P}_k h_2, 0]^T + v_x^k(t, h_1)F(h_2)\\
w_2(t, h_1; h_2) :=& v_x^k(t, h_1)[C\mathcal{P}_k h_1, 0]^T + v_x^k(t, h_1)F(h_1)\\
w_3(t, h_1; h_2) :=& \tfrac{1}{2}\operatorname{Tr}(v_{xx}^k(t, h_1)B(h_1)B(h_1)^*)\\
w_4(t, h_1; h_2) :=& \tfrac{1}{2}\operatorname{Tr}(v_{xx}^k(t, h_1)B(h_2)B(h_2)^*).
\end{aligned}$$

Clearly,

$$\mathbf{E}\phi(X^{\Delta t,k}(T;x)) - \mathbf{E}\phi(X^k(T;x)) = v^k(T, X^{\Delta t,k}(T;x)) - v^k(0, X^{\Delta t,k}(0;x))$$

and hence

$$\begin{aligned}
&\Big|\mathbf{E}\phi(X^{\Delta t,k}(T;x)) - \mathbf{E}\phi(X^k(T;x))\Big| \\
\le& \int_0^T \sum_{i=1}^4 |\mathbf{E}w_i(s, X^{\Delta t,k}(s;x); X^{\Delta t,k}(\hat{s};x)) - \mathbf{E}w_i(s, X^{\Delta t,k}(\hat{s};x); X^{\Delta t,k}(\hat{s};x))|\, ds \qquad (5.1)\\
&+ \Big|\int_0^T \mathbf{E}v_x(s, X^{\Delta t,k}(s;x)) \begin{pmatrix} (C - C^{\Delta t})\mathcal{P}_k X^{\Delta t,k}(\hat{s};x) \\ 0 \end{pmatrix} ds\Big|.
\end{aligned}$$

The modulus of the integrand of the last term in this estimate is

$$\begin{aligned}
\le&(\mathbf{E}\|v_x(s; X^{\Delta t,k}(s;x))\|^2_{\mathcal{L}(H,\mathbf{R})})^{1/2} \\
&\times \Big(\mathbf{E}\Big\| \int_{-\tau}^0 a(dr)\Big(\mathcal{P}_k X^{\Delta t,k}(\hat{s};x)(r) - \mathcal{P}_k X^{\Delta t,k}(\hat{s};x)(\hat{r})\Big)\Big\|^2_{\mathbf{R}^d}\Big)^{1/2}.
\end{aligned}$$

The term $\mathbf{E}\|v_x(s; X^{\Delta t,k}(s;x))\|^2_{\mathcal{L}(H,R)}$ is bounded by $K(1 + \|x\|^{p-1})$ by using Corollary 2.4 and (2.7). Let $\alpha(s,r;x) := \mathcal{P}_k X^{\Delta t,k}(s;x)(r) - \mathcal{P}_k X^{\Delta t,k}(s;x)(\hat{r})$. Using this notation and assuming that τ is an integer multiple of Δt,

$$\begin{aligned}
&\mathbf{E}\Big[\Big\| \int_{-\tau}^0 a(dr)\alpha(\hat{s},r;x)\Big\|^2_{\mathbf{R}^d}\Big] \\
=& \sum_{i=-\lfloor \tau/\Delta t\rfloor}^{-1} \sum_{j=-\lfloor \tau/\Delta t\rfloor}^{-1} \mathbf{E}\Big[\Big\langle \int_{i\Delta t}^{(i+1)\Delta t} a(dr)\alpha(\hat{s},r;x), \int_{j\Delta t}^{(j+1)\Delta t} a(dr)\alpha(\hat{s},r;x)\Big\rangle\Big].
\end{aligned}$$

By the second part of Lemma 2.9, the cross terms ($i \neq j$) obey

$$\begin{aligned}
\mathbf{E}\Big[\Big\langle \int_{i\Delta t}^{(i+1)\Delta t} a(dr)\alpha(\hat{s},r;x), \int_{j\Delta t}^{(j+1)\Delta t} a(dr)\alpha(\hat{s},r;x)\Big\rangle\Big]& \\
\le K(1 + \|x\| + \|\pi_D x\|_{\text{Lip}})^2 \Delta t^4 + o(k^{-1})&
\end{aligned}$$

and by the first part of Lemma 2.9 the diagonal terms

$$\begin{aligned}
\mathbf{E}\Big\| \int_{i\Delta t}^{(i+1)\Delta t} a(dr)\alpha(\hat{s},r;x)\Big\|^2_{\mathbf{R}^d} \le& \mathbf{E}\Big[\int_{i\Delta t}^{(i+1)\Delta t} \bar{a}(r)\|\alpha(\hat{s},r;x)\|_{\mathbf{R}^d}\, dr\Big]^2 \\
\le& \mathbf{E}\Big[\sup_{i\Delta t \le r < (i+1)\Delta t} \|\alpha(\hat{s},r;x)\|_{\mathbf{R}^d} \int_{i\Delta t}^{(i+1)\Delta t} \bar{a}(r)\, dr\Big]^2 \\
\le& K(1 + \|x\| + \|\pi_D x\|_{\text{Lip}})^2 \Delta t^3 + o(k^{-1}).
\end{aligned}$$

Consequently,

$$\begin{aligned}
&\mathbf{E}\Big[\Big\| \int_{-\tau}^0 a(dr)\alpha(\hat{s},r;x)\Big\|^2\Big] \\
\le& K(1 + \|x\| + \|\pi_D x\|_{\text{Lip}})^2 \Big(\lfloor \tau/\Delta t\rfloor \Delta t^3 + (\lfloor \tau/\Delta t\rfloor)^2 \Delta t^4\Big) + o(k^{-1}) \\
\le& K(1 + \|x\| + \|\pi_D x\|_{\text{Lip}})^2 \Delta t^2 + o(k^{-1}).
\end{aligned}$$

Thus the final term in (5.1) is bounded by $K(1 + \|x\|^p + \|x\|^{p-1}\|\pi_D x\|_{\text{Lip}})\Delta t + o(k^{-1})$.

We wish to apply Lemma 3.2 to show that each pair of terms in w_i in (5.1) is order Δt. Because $s > \hat{s}$, it is sufficient to apply the Lemma to $w(t,x) = w_i(t,x;h_2)$. We now verify the hypothesis of Lemma 3.2.

derivatives We require that w, w_t, w_x, w_{xx} exist, be uniformly continuous on bounded subsets of $\mathbf{R}^+ \times H$, and obey the growth bounds specified in Lemma 3.2. In each case, this is a consequence on Lemma 4.3. Part (i) of this lemma covers w_1 and w_2, while part (ii) covers w_3 and w_4. We make use of differentiability properties of F and B in applying this Lemma.

drift term in A We further require that in each case $|w_x(t,x)\tilde{A}_k h| \le K(1+\|x\|^{p-1})\,\|h\|$ for $h \in H$. We look at $w(t,x) = w_2(t,x;x)$ in detail; the others are similar. Note

$$\begin{aligned} w_x(t,x)\tilde{A}_k h =& v_x^k(t,x)[C\mathcal{P}_k\tilde{A}_k h, 0]^T + v_x^k(t,x)F_x(x)\tilde{A}_k h \\ &+ v_{xx}^k(t,x)([C\mathcal{P}_k x, 0]^T, \tilde{A}_k h) + v_{xx}^k(t,x)(F(x), \tilde{A}_k h). \end{aligned}$$

By Hypothesis 1.1(iii), $C\mathcal{P}_k\tilde{A}_k$ is bounded from H to $\mathbf{R}^d$ uniformly in k. Further $F_x(x)\tilde{A}_k = 0$ because F is independent of the delay and the definition of $\tilde{A}_k$ (see (2.5)). By using the bound on v_x in Corollary 2.4, we conclude that the first two terms are bounded by $K(1+\|x\|^p)\,\|h\|$.

Write out the terms in v_{xx}^k using $\xi^g = X_x^k(t;x)g$ and $\eta^{h,g} = X_{xx}^k(t;x)(g,h)$ and $Q = [C\mathcal{P}_k x, 0]^T$

$$\begin{aligned} v_{xx}^k(t,x)(Q, \tilde{A}_k h) =& \mathbf{E}\phi''(X^k(t;x))(\xi^Q, \xi^{\tilde{A}_k h}) + \mathbf{E}\phi'(X^k(t;x))\eta^{Q,\tilde{A}_k h} \\ v_{xx}^k(t,x)(F(x), \tilde{A}_k h) =& \mathbf{E}\phi''(X^k(t;x))(\xi^{F(x)}, \xi^{\tilde{A}_k h}) + \mathbf{E}\phi'(X^k(t;x))\eta^{F(x),\tilde{A}_k h}. \end{aligned}$$

Lemma 4.1 derives bounds on $\xi^{A_k h}$; clearly the same technique gives bounds for $\xi^{\tilde{A}_k h}$. Similarly, the techniques in Lemma 4.4 give bounds on the terms in η. We conclude that the required bound on $|w_x(t,x)\tilde{A}_k h|$ holds.

Thus, Lemma (3.2) applies to the terms in the summation in (5.1), giving bounds of the form $K(1+\|x\|^p)\Delta t$. Taking this observation with the bound for the last term in (5.1), we have a bound on the weak error in the Yosida approximation of the form $K(1+\|x\|^p)\Delta t + K(1+\|x\|^p+\|x\|^{p-1}\|\pi_D x\|_{\mathrm{Lip}})\Delta t + o(k^{-1})$, where K can be chosen independent of k. Take the limit in $k \to \infty$ to complete the proof.

QED

6 Numerical Experiments

We present results of numerical experiments corresponding to examples of (1.1). Our objective is to illustrate the convergence of the weak Euler method with respect to decreasing step-size by computing first or second moments, that is we compute $\mathbf{E}\phi(Y(T))$ for $\phi(Y) = Y$ where $Y(T)$ denotes a solution of (1.1).

Example 6.1 Consider

$$dY(t) = \Big[\int_{t-1}^{t} Y(s)\,ds + \exp(-1)Y(t)\Big]\,dt + (\sigma_1 + \sigma_2\,Y(t))dW(t), \tag{6.1}$$

for $t \in [0,T]$ and $Y(s) = \exp(s)$ for $-1 \le s \le 0$ and $W(t)$ is a one dimensional Wiener process.

Let $m(t) := \mathbf{E}Y(t)$ for $t \ge 0$. Then, $m(t)$ satisfies the delay-integro-differential equation

$$m'(t) = \int_{t-1}^{t} m(s)\,ds + \exp(-1)m(t), \tag{6.2}$$

with initial condition

$$m(s) = \exp(s) \text{ for } -1 \le s \le 0. \tag{6.3}$$

Equation (6.2) subject to (6.3) has the solution $m(t) = \exp(t)$.

With a step-size $\Delta t = T/N$ and $k = \tau/\Delta t = 1/\Delta t$, the weak Euler method takes the form

$$Y_{n+1}^{\Delta t} = Y_n^{\Delta t} + \Delta t\Big(\exp(-1)Y_n^{\Delta t} + \Delta t \sum_{j=n-k}^{n-1} Y_j^{\Delta t}\Big) + (\sigma_1 + \sigma_2 Y_n^{\Delta t})\ \Delta W_n \tag{6.4}$$

for $n = 0, \ldots, N-1$ and with $Y_j^{\Delta t} = \exp(j\Delta t)$ for $j \le 0$. The ΔW_n denote IID $\mathcal{N}(0, \Delta t)$ distributed random variables approximating $W((n+1)\Delta t) - W(n\Delta t)$. We have used the composite explicit Euler rule to approximate the integral. To illustrate the convergence of the method, we have simulated 10,000 sample trajectories with each of the step-sizes $\Delta t = 2^{-3}, 2^{-4}, ..$ and computed the error

$$\mu^{\Delta t}(T) = |\mathbf{E}Y_N^{\Delta t} - \mathbf{E}Y(T)| \tag{6.5}$$

at the final time $T = 2$. In Figure 1, we have plotted $\log_2(\mu^{\Delta t}(T))$ versus $\log_2(\Delta t)$.

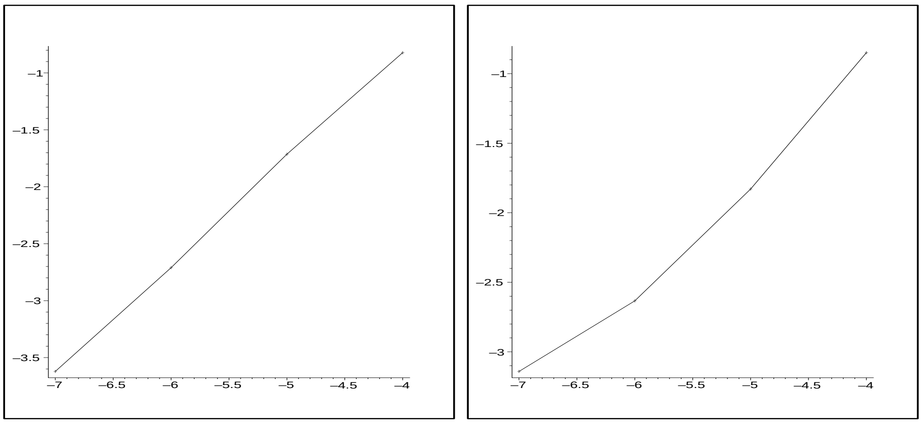

Figure 1: $\log_2(\mu^{\Delta t}(T))$ versus $\log_2(\Delta t)$ for (6.1) with left: $\sigma_1 = 0.2$, $\sigma_2 = 0$, right: $\sigma_1 = 0.0$, $\sigma_2 = 0.2$.

A well-known feature of weak approximation methods is that the Gaussian random numbers $\Delta\beta_n$ can be replaced by simpler random variables $\Delta\hat{\beta}_n$ (see [10]). We have performed numerical experiments with two-point distributed random variables with

$$\mathbf{P}(\Delta \hat{W}_n = \pm\sqrt{\Delta t}) = \tfrac{1}{2}.$$

In Figure 2 we present corresponding error-plots.

For illustration purposes we also include some trajectories in the following figure, the thick line corresponds to $m(t) = \exp(t)$.

The computations follow the exposition in [3].

7 Acknowledgements

We thank S. Mohammed, K. Ritter, A. Stuart, and J. Zabczyk for useful discussions.

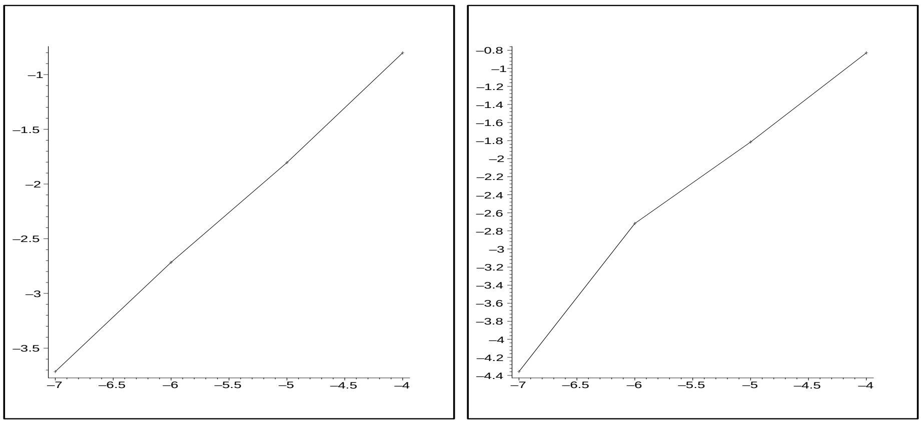

Figure 2: $\log_2(\mu^{\Delta t}(T))$ versus $\log_2(\Delta t)$ for (6.1) with left: $\sigma_1 = 0.2$, $\sigma_2 = 0$, right: $\sigma_1 = 0.0$, $\sigma_2 = 0.2$.

References

[1] C. T. Baker and E. Buckwar, *Continuous Θ-methods for the stochastic pantograph equation.*, ETNA, Electron. Trans. Numer. Anal., 11 (2000), pp. 131–151.

[2] ——, *Numerical analysis of explicit one-step methods for stochastic delay differential equations.*, LMS J. Comput. Math., 3 (2000), pp. 315–335.

[3] C. T. H. Baker and N. Ford, *Convergence of linear multistep methods for a class of delay-integro-differential equations*, in International Series of Numerical Mathematics, vol. 86, Birkhäuser Verlag Basel, 1988, pp. 47–59.

[4] C. Bernier and A. Manitius, *On semigroups in $R^n \times L^p$ corresponding to differential equations with delays*, Canadian Journal of Mathematics, XXX (1978), pp. 897–914. This paper derives the domain for semigroup generated by the linear delay system correctly, but states the definition of the domain without the absolutely continuous condition.

[5] G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*, vol. 44 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 1992.

[6] G. Da Prato and J. Zabczyk, *Ergodicity for infinite-dimensional systems*, vol. 229 of London Mathematical Society Lecture Note Series, Cambridge University Press, Cambridge, 1996.

[7] G. Gripenberg, S.-O. Londen, and O. Staffans, *Volterra Integral and Functional Equations*, vol. 34 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, 1990.

[8] Y. Hu, *Strong and weak order of the time discretization schemes of stochastic differential equations*, in Springer Lecture Notes in Mathematics, no. 1626, pp. 218–227.

[9] Y. Hu, S. E. A. Mohammed, and F. Yan, *Numerical solution of stochastic differential systems with memory*, tech. rep., Department of Mathematics, Southern Illinois University, Carbondale, Illinois,, 2001.

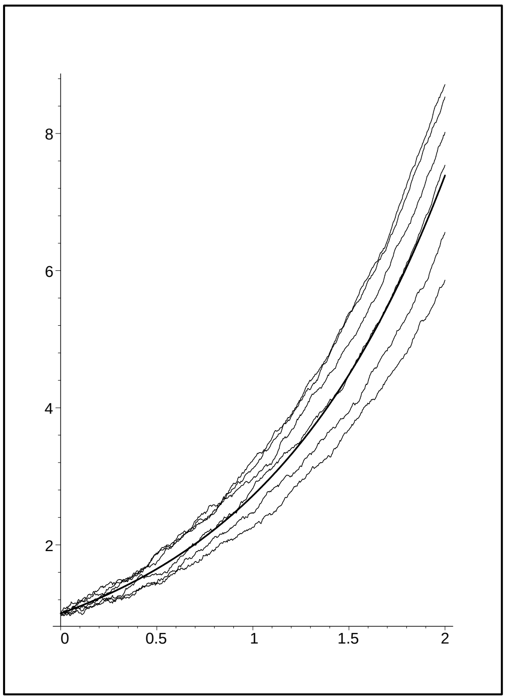

Figure 3: Trajectories of (6.1) with $\sigma_1 = 0.2,\ \sigma_2 = 0$.

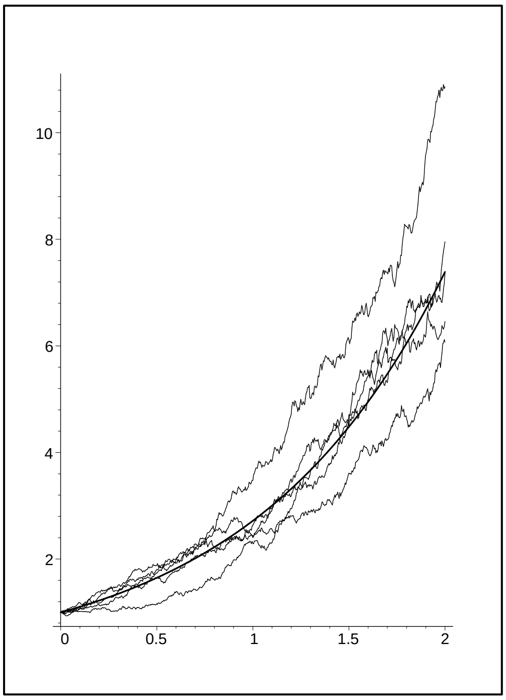

Figure 4: Trajectories of (6.1) with $\sigma_1 = 0.0,\ \sigma_2 = 0.2$.

[10] P. E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, vol. 23 of Applications of mathematics, Springer–Verlag, 1992.

[11] U. Küchler and E. Platen, *Strong discrete time approximation of stochastic differential equations with time delay*, Math. Comput. Simulation, 54 (2000), pp. 189–205.

[12] A. Pazy, *Semi groups of Linear Operators and Applications to Partial Differential Equations*, vol. 44 of Applied Mathematical Sciences, Springer–Verlag, 1983.

[13] C. Tudor and M. Tudor, *On approximation of solutions for stochastic delay equations.*, Stud. Cercet. Mat., 39 (1987), pp. 265–274.

[14] C. Tudor and M. Tudor, *Approximation of linear stochastic functional equations.*, Rev. Roum. Math. Pures Appl., 35 (1990), pp. 81–99.

[15] M. Tudor, *Approximation schemes for stochastic equations with hereditary argument.*, Stud. Cercet. Mat., 44 (1992), pp. 73–85.

[16] ——, *Some approximations of second order for stochastic equations with hereditary argument.*, Stud. Cercet. Mat., 44 (1992), pp. 147–158.

[17] ——, *Difference approximations for linear stochastic functional equations.*, Stud. Cercet. Mat., 45 (1993), pp. 351–362.