Weak approximation of stochastic differential delay equations

Evelyn Buckwar* Tony Shardlow†

April 1, 2004

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

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

for initial conditions $Y_S \in \mathbf{R}^d$ and $Y_D \in C([-\tau,0],\mathbf{R}^d)$, where $a(\cdot)$ is a $d\times d$ matrix valued measure 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). Denote the floor function by $\lfloor t \rfloor$, which equals the greatest integer less than or equal to t. 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`.

†Department of Mathematics, Manchester Univeristy, Oxford Road, Manchester M13 9PL, `shardlow@maths.man.ac.uk`. This work was supported in part by the Nuffield Foundation NUF-NAL-00.

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 [24, 25, 26, 27, 28]. Recently this topic has gained more attention, see [1], [2], [13], and [15]. 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). The theoretical grounding developed for the Euler method in this paper should make it possible to understand higher order weak approximation methods for stochastic differential delay equations. There are two basic approaches to achieving this goal. As developed in [14, 16] for SDEs, we can look for higher order methods. The drawback is the difficulty in implementing higher order methods for practical problems. The second approach is to use Richardson extrapolation based on a lower order numerical method such as forward Euler. An understanding of the behaviour of the error as developed in [23] (where the error in weak approximation is expanded in Δt power series) is needed to justify this rigorously. We leave these issues as open problems.

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 that $a(ds)$ has a C^1 density $a : [-\tau, 0] \to \mathbf{R}$.

For an integer $p \geq 4$, introduce the spaces $\mathcal{G}_p$ of test functions $\phi\colon \mathbf{R}^d \to \mathbf{R}$ that are four times continuously differentiable and satisfy $\|\phi^{(n)}(h)\|_{\mathcal{L}(\mathbf{R}^{d\times n},\mathbf{R})} \leq K(1 + \|h\|^{p-n}_{\mathbf{R}^d})$, for $h \in \mathbf{R}^d$ and some constant K, for $n = 0, 1, 2, 3, 4$ ($\|\cdot\|_{\mathcal{L}(\mathbf{R}^m,\mathbf{R})}$ is the standard norm induced on linear operators from $\mathbf{R}^m$ to $\mathbf{R}$ by the Euclidean norm). In one dimension, this space is sufficiently general to include polynomials.

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

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

Theorem 1.2 *Let Hypothesis 1.1 hold. Consider $Y_S \in \mathbf{R}^d$ and a Lipschitz function $Y_D : [-\tau, 0] \to \mathbf{R}^d$. Let $Y(t)$ (respectively, Y_n) denote the solution of (1.1) (resp., (1.2)) corresponding to initial data $x = (Y_S, Y_D)^T$. For $T > 0$ and $\phi \in \mathcal{G}_p$, $p \geq 4$, there exists a constant $K_x > 0$ such that*

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

and a constant K independent of the initial data such that

$$K_x \leq K(1 + \|x\|^p) + K(1 + \|x\|^{p-1})\|Y_D\|_{Lip}. \tag{1.3}$$

This is the main result of the present paper. The theorem makes a number of assumptions on the regularity of the problem. Compared to the results available for SDEs, the hypothesis on f, b, and ϕ come as no surprise as four derivatives are required to derive the analogous result for SDEs. The assumptions can be relaxed for SDEs under a non-degeneracy assumption on the noise by use of the Malliavin calculus [4], but such techniques are not used in this paper.

The assumption on the delay term is more restrictive and excludes the important case of discrete delays, $a(ds) = \sum \delta_{\tau_i}(ds)$.

The main motivation for considering weak approximations is the computation of the expectation of functionals of the solutions of equation (1.1). This problem arises for example in stochastic finance theory for the fair pricing of options. The standard model there is that of Black and Scholes. Recently, though, several papers have appeared, where the authors propose generalisations of this model by including (some part of) the history of the evolution of the price of the security. Equation (1.1) fits into the framework of [6, Remark 2.3] and is similar to the model in [12]. Optimal harvesting strategies are considered in [9]; equation (1.1) is a special case of the type of equations investigated. Another purpose for weak approximations is the computation of Lyapunov exponents of systems described by stochastic functional differential equations. Lyapunov exponents for stochastic functional differential equations have been considered in [18, 19, 11]; equation (1.1) fits into the class of equations treated. The use of weak approximations for computing Lyapunov exponents has been suggested by Milstein and Tretyakov in [17].

The proof of our main theorem is built by developing the delay equation (1.1) as a stochastic evolution equation on an infinite dimensional space in order to achieve a Markov process and a Kolmogorov equation. We review the theory in §2 for (1.1) and develop the numerical method on this space in §3. Two corollaries of the Itô calculus are established in §4 concerning certain functionals of the solutions. The Kolmogorov equation is introduced in §5 and developed for the regularised delay equation. It is important to establish sufficient time and spatial regularity of $v^k(t,x) := \mathbf{E}\phi(Y^k(t))$, where $Y^k(t)$ is a regularised version of 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 (ii) was introduced. The proof of Theorem 1.2 is completed in §6.

The convergence of weak approximations has been established for many numerical approximations of SDEs by looking at the Kolmogorov equation. The argument given in this paper follows closely [14] and appeared first in [22]. The difference in the present case is the introduction of a delay term. A similar technique has been applied to study weak approximation of a linear stochastic heat equation [21].

The Kolmogorov equation is difficult for evolution equations forced by a Wiener process. The drift term 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 of a continuous linear operator between normed vector spaces H_1 to H_2 is denoted by $\|\cdot\|_{\mathcal{L}(H_1,H_2)}$. Let

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

Then $|\cdot|_\star$ is a well defined semi-norm on H and, for a constant K, $|X|_\star \le K\|X\|$, all $X \in H$. Further define $H' := \mathbf{R}^d \times L_\infty([-\tau,0],\mathbf{R}^d)$ with

$$\|X\|_{H'} := \max\{\|\pi_S X\|_{\mathbf{R}^d}, \|\pi_D X\|_{L_\infty([-\tau,0],\mathbf{R}^d)}\}, \qquad X \in H'. \tag{1.4}$$

For an orthonormal basis e_i of $\mathbf{R}^d$, a Hilbert space H_1 with norm $\|\cdot\|_{H_1}$, and $\mathcal{B} \in \mathcal{L}(\mathbf{R}^d, H_1)$, define the Hilbert-Schmidt norm

$$\|\mathcal{B}\|^2_{L_2^0(\mathbf{R}^d,H_1)} := \sum_{i=1}^d \|\mathcal{B}e_i\|^2_{H_1}.$$

Let $L_2^0(\mathbf{R}^d, H_1)$ equal the space $\mathcal{L}(\mathbf{R}^d, H_1)$ taken with the Hilbert-Schmidt norm. 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 $\hat{s} := \Delta t\lfloor s/\Delta t\rfloor$.

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 $X \in H$

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

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

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

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

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

For further details see [7, p. 123] and [5, 10]. 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 [8].

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 $L_2^0(\mathbf{R}^d, H)$ valued process such that $\int_0^t \|\mathcal{B}(s)\|^2_{L_2^0(\mathbf{R}^d,H)}\, ds$ 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 [7].

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,*

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

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:*

$$\mathbf{E}\Big[\Big\|\int_0^T \mathcal{B}(s)\, dW(s)\Big\|^2\Big] = \int_0^T \mathbf{E}\|\mathcal{B}(s)\|^2_{L_2^0(\mathbf{R}^d,H)}\, ds.$$

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

$$\mathbf{E}\Big[\sup_{0\le t\le T}\Big\|\int_0^t \mathcal{B}(s)\, dW(s)\Big\|^p\Big] \le c_p \mathbf{E}\Big|\int_0^T \|\mathcal{B}(s)\|^2_{L_2^0(\mathbf{R}^d,H)}\, ds\Big|^{p/2}.$$

2.3 Regularity of solutions

Denote by $L^p(\Omega, H)$ the set of H valued random variables on the sample space Ω with finite pth moment, $\mathbf{E}\|X\|^p < \infty$.

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). For fixed $t > 0$, the solution map from $x \in H$ to $X(t;x) \in L^p(\Omega, H)$ is four times continuously Frechet differentiable, with derivatives denoted by $X_x(t,x)$, $X_{xx}(t;x)$, $X_{xxx}(t;x)$, $X_{xxxx}(t;x)$. The 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 $0 \le t \le T$*

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

and for $h_i \in H$ for $i = 1, 2, 3, 4$,

$$\begin{aligned}
(\mathbf{E}\|X_x(t;x)h_1\|^p)^{1/p} \leq & K\|h_1\|, \\
(\mathbf{E}\|X_{xx}(t;x)(h_1,h_2)\|^p))^{1/p} \leq & K\|h_1\| \cdot \|h_2\|, \\
(\mathbf{E}\|X_{xxx}(t;x)(h_1,h_2,h_3)\|^p)^{1/p} \leq & K\|h_1\| \cdot \|h_2\| \cdot \|h_3\|, \\
(\mathbf{E}\|X_{xxxx}(t;x)(h_1,h_2,h_3,h_4)\|^p)^{1/p} \leq & K\|h_1\| \cdot \|h_2\| \cdot \|h_3\| \cdot \|h_4\|.
\end{aligned}$$

Proof See Da Prato–Zabczyk [7] Theorem 9.4, which gives existence of the first and second derivatives and shows for example that $X_x(t;x)h$ is a mild solution of the variational equation

$$X_x(t;x)h = S(t)h + \int_0^t S(t-s)F_x(X(s;x))X_x(s;x)h\,ds + \int_0^t S(t-s)B_x(X(s;x))X_x(s;x)h\,dW(s).$$

Thus, using the Burkholder-Davis-Gundy inequality and Hypothesis 1.1(i),

$$\mathbf{E}\|X_x(t;x)h\|^p \leq K\|h\|^p + K\int_0^t \mathbf{E}\|X_x(s;x)h\|^p\,ds + K\mathbf{E}\Big[\int_0^t \|X_x(s;x)h\|^2\,ds\Big]^{p/2}$$

(using $(\int_0^t \phi(s)\,ds)^p \leq K(\int_0^t \phi(s)^p\,ds)$ for $p > 1$)

$$\leq K\|h\|^p + K\int_0^t \mathbf{E}\|X_x(s;x)h\|^p\,ds + K\mathbf{E}\Big[\int_0^t (\|X_x(s;x)h\|^p)\,ds\Big].$$

This leads to the quoted bounds on $\mathbf{E}\|X_x(t;x)h\|^p$. 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 $\phi \in \mathcal{G}_p$ for $p \geq 4$ and let $v(t,x) := \mathbf{E}\phi(\pi_S X(t;x))$. 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 $0 \leq t \leq T$,*

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

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 $K(1 + \|x\|^{p-1})$ on the interval $[0,T]$.

Proof Clearly, $|v(t,x)| \leq K\mathbf{E}(1 + \|X(t;x)\|^p) \leq K(1 + \|x\|^p)$ 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 $\|x\|, \|x'\| \leq M$ and choose $\epsilon > 0$. Choose R sufficiently large that $\mathbf{P}(\|X(t;x)\| \leq R,\ 0 \leq t \leq T) \geq 1 - \epsilon$. Then, as ϕ is locally Lipschitz, for a constant K_R,

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

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 [20, 7].

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 on $[-\tau, 0]$ of

$$kX_D = kh - \frac{d}{dt}h, \quad \text{for } h(0) = \pi_S X. \tag{2.3}$$

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, $\|S_k(t)h - S_\ell(t)h\| \le K\|A_kh - A_\ell h\|$ for $k, \ell = 1, 2, \dots$ and $0 \le t \le T$.

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

(iv) For $X \in H$, $\|\mathcal{P}_kX\|_{L_2([-\tau,0],\mathbf{R}^d)} \le K\|X\|$ and for $X \in H'$,

$$\|\mathcal{P}_kX\|_{L_\infty([-\tau,0],\mathbf{R}^d)} \le \max\{\|\pi_S X\|_{\mathbf{R}^d}, \|\pi_D X\|_{L_\infty([-\tau,0],\mathbf{R}^d)}\}.$$

Proof The first two properties are standard results from C_0 semigroups (see §1.5 of [20]). The third property follows from property (i), if $\|\pi_S A_k\|_{\mathcal{L}(H,\mathbf{R}^d)}$ is bounded. But $\pi_S A_k = C\mathcal{P}_k$, a product of two operators, both of which are bounded for k large.

To understand the fourth property, one can show that

$$\begin{aligned}
\mathcal{P}_k\begin{pmatrix} 0 \\ \cos(2\pi ks/\tau) \end{pmatrix} &= \frac{2\pi kn\tau}{4\pi^2n^2 + k^2\tau^2}\sin(2\pi ns/\tau) + \frac{k^2\tau^2}{4\pi^2n^2 + k^2\tau^2}\cos(2\pi ns/\tau),\\
\mathcal{P}_k\begin{pmatrix} 0 \\ \sin(2\pi ks/\tau) \end{pmatrix} &= \frac{k^2\tau^2}{4\pi^2n^2 + k^2\tau^2}\sin(2\pi ns/\tau) + \frac{-2\pi kn\tau}{4\pi^2n^2 + k^2\tau^2}\cos(2\pi ns/\tau).
\end{aligned}$$

Note that the coefficients of the Fourier modes are less than one in magnitude. Hence by expanding X_D in Fourier series we can show that $\|\mathcal{P}_k(X_S, X_D)\|_{L_2([-\tau,0],\mathbf{R}^d)} \le K\|(X_S, X_D)\|$. For the $L_\infty([-\tau, 0], \mathbf{R}^d)$ bound,

$$(\mathcal{P}_kX)(t) = \int_t^0 ke^{k(t-s)}(\pi_D X)(s)\, ds + e^{kt}\pi_S X, \qquad -\tau \le t \le 0,$$

so that

$$\begin{aligned}
\Big\|(\mathcal{P}_kX)(t)\Big\|_{\mathbf{R}^d} &\le \|\pi_D X\|_{L_\infty([-\tau,0],\mathbf{R}^d)} \int_t^0 ke^{k(t-s)}\, ds + e^{kt}\|\pi_S X\|_{\mathbf{R}^d},\\
&= \|\pi_D X\|_{L_\infty([-\tau,0],\mathbf{R}^d)}(1 - e^{kt}) + e^{kt}\|\pi_S X\|_{\mathbf{R}^d},
\end{aligned}$$

which completes the proof as $0 \le e^{kt} \le 1$. *QED*

Lemma 2.6 *Let Hypothesis 1.1(i) hold. 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*

$$dX^k = \Big[A_k X^k + F(X^k)\Big]\,dt + B(X^k)\,dW, \quad X^k(0) = x. \tag{2.4}$$

Then,

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

Proof Proposition 7.5 [7]. *QED*

3 The numerical method on H

To perform the convergence analysis, we find an interpolant of the numerical solution Y_n in H that can be represented as a stochastic integral. We carefully write down the regularisation steps to define the interpolant rigorously. The main difficulty as before is working with the unbounded part of A.

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 $n\Delta t \le t < (n+1)\Delta t$ and we remind the reader of the notation $\hat s = \Delta t\lfloor s/\Delta t\rfloor$, defined in Section 1.1. Then, define $X^{\Delta t} = (X_S^{\Delta t}, X_D^{\Delta t})^T$ by

$$\begin{aligned} X_S^{\Delta t}(t;x) :=& Y_n + \Big[\sum_{i=-\lfloor\tau/\Delta t\rfloor}^{-1} a_i Y_{n+i} + f(Y_n)\Big](t-\hat t) + b(Y_n)\,(\bar W(t) - \bar W(\hat t))\\ =& X_S^{\Delta t}(\hat t;x) + \Big[\sum_{i=-\lfloor\tau/\Delta t\rfloor}^{-1} a_i X_D(\,\hat t;x)(i\Delta t) + f(X_S(\,\hat t;x))\Big](t-\hat t)\\ &+ b(X_S(\,\hat t;x))\,(\bar W(t) - \bar W(\hat t\,))\\ X_D^{\Delta t}(t;x)(s) :=& \begin{cases} X_S(t+s;x), & t+s\ge 0,\\ Y_D(t+s), & -\tau\le t+s<0, \end{cases} \qquad -\tau\le s\le 0. \end{aligned} \tag{3.1}$$

It is necessary to develop this equation as a well defined H valued stochastic integral. To simplify calculations later on, 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 $C^{\Delta t}\mathcal{P}_k$ is a well defined operator from H to $\mathbf{R}^d$. Introduce

$$\tilde A := \begin{pmatrix} 0 & 0\\ 0 & \frac{d}{dt}\end{pmatrix} \tag{3.2}$$

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

$$\begin{aligned} dX^{\Delta t,k}(t;x) =& \Big[\tilde{A}_k X^{\Delta t,k}(t;x) + \begin{pmatrix} C^{\Delta t} \\ 0 \end{pmatrix} \mathcal{P}_k X^{\Delta t,k}(\,\hat{t};x) + F(X^{\Delta t,k}(\,\hat{t}\,;x))\Big]\, dt \\ &+ B(X^{\Delta t,k}(\,\hat{t}\,;x))\, d\bar{W}(t), \qquad X^{\Delta t,k}(0;x) = x. \end{aligned} \tag{3.3}$$

All terms on the right hand side are evaluated at $\hat{t}$, except the first which is evaluated at t. 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 3.1 *Let Hypothesis 1.1(i) hold. The solution $X^{\Delta t,k}(t;x)$ of (3.3) converges to the interpolant $X^{\Delta t}(t;x)$ defined in (3.1) in the sense that*

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

Proof Note that $X^{\Delta t}(t;x)$ is the mild solution of

$$dX^{\Delta t}(t;x) = \Big[\tilde{A} X^{\Delta t}(t;x) + \begin{pmatrix} C^{\Delta t} \\ 0 \end{pmatrix} X^{\Delta t}(\,\hat{t};x) + F(X^{\Delta t}(\,\hat{t}\,;x))\Big]\, dt \tag{3.4}$$

$$+ B(X^{\Delta t}(\,\hat{t}\,;x))\, d\bar{W}(t), \qquad X^{\Delta t}(0;x) = x. \tag{3.5}$$

Now the result can be established as in Lemma 2.6. *QED*

For $X \in H$, let

$$\|X\|_{\Delta t} = \Big(\|X\|^2 + \sum_{i=-\lfloor\tau/\Delta t\rfloor}^{-1} \Delta t \|\mathcal{P}_k X(i\Delta t)\|^2_{\mathbf{R}^d}\Big)^{1/2}. \tag{3.6}$$

Lemma 3.2 *Let Hypothesis 1.1(i)-(ii) hold. Suppose that $x = (Y_S, Y_D)$, where Y_D is a continuous function on $[-\tau, 0]$. Then, for each $T > 0$ and $2 \le p < \infty$, there exists $K > 0$ such that*

$$\Big(\mathbf{E}\|X^{\Delta t,k}(t;x)\|^p_{\Delta t}\Big)^{1/p} \le K\Big(1 + \|Y_S\| + \sup_{-\tau\le s\le 0} \|Y_D(s)\|_{\mathbf{R}^d}\Big), \qquad 0 \le t \le T. \tag{3.7}$$

Further for any $\mathcal{Q} \in \mathcal{L}(H', \mathbf{R}^d)$ (recall the definition of H' in (1.4))

$$\Big(\mathbf{E}\|\mathcal{Q}X^{\Delta t,k}(t;x)\|^p_{\mathbf{R}^d}\Big)^{1/p} \le K(1 + \|Y_S\| + \sup_{-\tau\le s\le 0} \|Y_D(s)\|_{\mathbf{R}^d}), \qquad 0 \le t \le T$$

and

$$\mathbf{E}\|X^{\Delta t,k}(\Delta t;x) - \mathbf{E}X^{\Delta t,k}(\Delta t;x)\|^2 \le K\Delta t \tag{3.8}$$

$$\mathbf{E}\|\mathcal{Q}(X^{\Delta t,k}(\Delta t;x) - \mathbf{E}X^{\Delta t,k}(\Delta t;x))\|^2_{\mathbf{R}^d} \le K\|\mathcal{Q}\|_{\mathcal{L}(H',\mathbf{R}^d)}\Delta t. \tag{3.9}$$

For $0 \le t \le T$ and $x, x' \in H$ and $p \ge 2$ and $\mathcal{Q} = \mathcal{C}^{\Delta t}\mathcal{P}_k$,

$$\begin{aligned} \mathbf{E}\Big[\|\mathcal{Q}(X^{\Delta t,k}(t;x) - X^{\Delta t,k}(t;x'))\|^2_{\mathbf{R}^d} + \|\pi_S(X^{\Delta t,k}(t;x) - X^{\Delta t,k}(t;x'))\|^2_{\mathbf{R}^d}\Big] \\ \le K\Big[\|\mathcal{Q}\tilde{S}_k(t)(x - x')\|^2_{\mathbf{R}^d} + \|\pi_S(x - x')\|^2\Big]. \end{aligned} \tag{3.10}$$

Proof Recall that, $\tilde{A}_k = k^2 R(k\colon \tilde{A}) - kI$, where $R(k\colon \tilde{A}) = (kI - \tilde{A})^{-1}$, or otherwise written

$$R(k:\tilde{A})X = \begin{pmatrix} k^{-1}X_S \\ k^{-1}\mathcal{P}_k X \end{pmatrix}.$$

Hence, by Proposition 2.5(iv), $\|R(k\colon \tilde{A})\|_{\mathcal{L}(H',H')} \le 1/k$. Let $\tilde{S}_k$ denote the semigroup on H generated by $\tilde{A}$. Because $\tilde{S}_k(t) = e^{-kt}e^{k^2 R(k\colon \tilde{A})t}$, we conclude that $\|\tilde{S}_k(t)\|_{\mathcal{L}(H',H')} \le K$ for $0 \le t \le T$.

Let $\Psi_0(t) = \tilde{S}_k(t)x$ and note that $\|\Psi_0(t)\|_{H'} \le K\|x\|_{H'}$. Let

$$\Psi_1(s) = \tilde{S}_k(s)\begin{pmatrix} C^{\Delta t} \\ 0 \end{pmatrix}\mathcal{P}_k X^{\Delta t,k}(\hat{s};x)$$

and note that under Hypothesis 1.1(ii)

$$\|\Psi_1(s)\|_{H'} \le K\|C^{\Delta t}\mathcal{P}_k X^{\Delta t,k}(\hat{s};x)\|_{\mathbf{R}^d} \le K\|X^{\Delta t,k}(\hat{s};x)\|_{\Delta t}.$$

Let $\Psi_2(s) = \tilde{S}_k(s)F(X^{\Delta t,k}(\hat{s};x))$ and note that

$$\|\Psi_2(s)\|_{H'} \le K(1 + \|\pi_S X^{\Delta t,k}(\hat{s};x)\|_{\mathbf{R}^d}).$$

For $h \in \mathbf{R}^d$, let $\Psi_3(s)h = \tilde{S}_k(s)B(X^{\Delta t,k}(\hat{s};x))h$ and note that

$$\|\Psi_3(s)h\|_{H'} \le K\|h\|_{\mathbf{R}^d}.$$

For $-\tau \le r \le 0$, let $\Psi_3(s,r)h = \pi_D(\Psi_3(s)h)(r)$ so that $\|\Psi_3(s,r)h\|_{\mathbf{R}^d} \le K\|h\|_{\mathbf{R}^d}$.

Because $\|X\|_{\Delta t} \le K\|X\|_{H'}$, we have $\|\Psi_0(t)\|_{\Delta t} \le K\|x\|_{H'}$ and

$$\|\Psi_1(s)\|_{\Delta t} \le K\|X^{\Delta t,k}(\hat{s};x)\|_{\Delta t}, \qquad \|\Psi_2(s)\|_{\Delta t} \le K(1 + \|X^{\Delta t,k}(\hat{s};x)\|_{\Delta t}),$$
$$\|\Psi_3(s)\|_{L_2^0(\mathbf{R}^d,H)} \le K, \qquad \|\Psi_3(s,r)\|_{L_2^0(\mathbf{R}^d,\mathbf{R}^d)} \le K.$$

Note that

$$\begin{aligned}
&\mathbf{E}\Big\|\int_0^t \Psi_3(t-s)\,d\bar{W}(s)\Big\|_{\Delta t}^p \\
\le& K\mathbf{E}\Big\|\int_0^t \Psi_3(t-s)\,d\bar{W}(s)\Big\|^p + K\mathbf{E}\Big[\sum_{i=-\lfloor\tau/\Delta t\rfloor}^{-1}\Delta t\Big\|\int_0^t \Psi_3(t-s,i\Delta t)\,d\bar{W}(s)\Big\|_{\mathbf{R}^d}^2\Big]^{p/2} \\
\le& K\mathbf{E}\Big[\int_0^t \|\Psi_3(t-s)\|_{L_2^0(\mathbf{R}^d,H)}^2\,ds\Big]^{p/2} + K\mathbf{E}\sum_{i=-\lfloor\tau/\Delta t\rfloor}^{-1}\Delta t\Big\|\int_0^t \Psi_3(t-s,i\Delta t)\,d\bar{W}(s)\Big\|_{\mathbf{R}^d}^p
\end{aligned}$$

(using that $(\sum_{i=-\lfloor\tau/\Delta t\rfloor}^{-1}\Delta t\phi_i)^q \le K\Delta t\sum_{i=-\lfloor\tau/\Delta t\rfloor}^{-1}\phi_i^q$ for $q \ge 1$)

$$\le K\mathbf{E}\Big[\int_0^t \|\Psi_3(t-s)\|_{L_2^0(\mathbf{R}^d,H)}^2\,ds\Big]^{p/2} + K\sum_{i=-\lfloor\tau/\Delta t\rfloor}^{-1}\Delta t\,\mathbf{E}\Big[\int_0^t \|\Psi_3(t-s,i\Delta t)\|_{L_2^0(\mathbf{R}^d,\mathbf{R}^d)}^2\,ds\Big]^{p/2},$$

which is bounded uniformly in Δt and k. From the Variation of Constants formula for (3.3),

$$X^{\Delta t,k}(t;x) = \Psi_0(s) + \int_0^t \Psi_1(t-s)\,ds + \int_0^t \Psi_2(t-s)\,ds + \int_0^t \Psi_3(t-s)\,d\bar{W}(s),$$

we gain

$$\Big(\mathbf{E}\|X^{\Delta t,k}(t;x)\|_{\Delta t}^{p}\Big)^{1/p} \leq K\|x\|_{H'} + \int_0^t (\mathbf{E}\|X^{\Delta t,k}(\hat{s};x)\|_{\Delta t}^{p})^{1/p}\,ds + \int_0^t K(1+(\mathbf{E}\|X^{\Delta t,k}(\hat{s};x)\|_{\Delta t}^{p})^{1/p})\,ds + K,$$

and the Gronwall inequality implies that

$$\Big(\mathbf{E}\|X^{\Delta t,k}(t;x)\|_{\Delta t}^{p}\Big)^{1/p} \leq K(\|x\|_{H'}+1), \qquad 0 \leq t \leq T. \tag{3.11}$$

This certainly implies (3.7).

As $\mathcal{Q} \in \mathcal{L}(H', \mathbf{R}^d)$, we see $\|\mathcal{Q}\Psi_0(s)\|_{\mathbf{R}^d} \leq K\|x\|_{H'}$, $\|\mathcal{Q}\Psi_1(s)\|_{\mathbf{R}^d} \leq K\|X^{\Delta t,k}(\hat{s};x)\|_{\Delta t}$, $\|\mathcal{Q}\Psi_2(s)\|_{\mathbf{R}^d} \leq K(1+\|\pi_S X^{\Delta t,k}(\hat{s};x)\|_{\mathbf{R}^d})$, and

$$\|\mathcal{Q}\Psi_3(s)\|_{L_2^0(\mathbf{R}^d,\mathbf{R}^d)} \leq K.$$

We conclude using (3.11) that

$$\begin{aligned}\Big(\mathbf{E}\Big\|\mathcal{Q}X^{\Delta t,k}(t;x)\Big\|_{\mathbf{R}^d}^{p}\Big)^{1/p} &\leq K\|x\|_{H'} + K\int_0^t (\mathbf{E}\|X^{\Delta t,k}(t;x)\|_{\Delta t}^{p})^{1/p}\,ds \\ &+ \int_0^t K(1+(\mathbf{E}\|\pi_S X^{\Delta t,k}(\hat{s};x)\|_{\mathbf{R}^d}^{p})^{1/p})\,ds + \Big(\mathbf{E}\Big[\int_0^t \|\mathcal{Q}\Psi_3(t-s,r)\|_{L_2^0(\mathbf{R}^d,\mathbf{R}^d)}^2\,ds\Big]^{p/2}\Big)^{1/p} \\ &\leq K(\|x\|_{H'}+1).\end{aligned}$$

We now deal with the final two statements. For $0 \leq t \leq \Delta t$, $X^{\Delta t,k}(\hat{t};x) = x$ and

$$X^{\Delta t,k}(t;x) = \tilde{S}_k(t)x + \int_0^t \tilde{S}_k(t-s)\begin{pmatrix} C^{\Delta t} \\ 0 \end{pmatrix}\mathcal{P}_k x\,ds + \int_0^{\Delta t} \tilde{S}_k(t-s)F(x)\,ds + \int_0^{\Delta t} \tilde{S}_k(t-s)B(x)\,d\bar{W}(s).$$

Consequently,

$$X^{\Delta t,k}(\Delta t;x) - \mathbf{E}X^{\Delta t,k}(\Delta t;x) = \int_0^{\Delta t} \tilde{S}_k(\Delta t - s)B(x)\,d\bar{W}(s).$$

With Lemma 2.2, this implies (3.8)–(3.9). Second, note that

$$\begin{aligned}&X^{\Delta t,k}(t;x) - X^{\Delta t,k}(t;x') \\ =&\tilde{S}_k(t)(x-x') + \int_t^0 \tilde{S}_k(t-s)\begin{pmatrix} C^{\Delta t} \\ 0 \end{pmatrix}\mathcal{P}_k\Big[X^{\Delta t,k}(\hat{s};x) - X^{\Delta t,k}(\hat{s};x')\Big]\,ds \\ &+ \int_0^{\Delta t} \tilde{S}_k(t-s)\Big[F(X^{\Delta t,k}(\hat{s};x)) - F(X^{\Delta t,k}(\hat{s};x'))\Big]\,ds \\ &+ \int_0^{\Delta t} \tilde{S}_k(t-s)\Big[B(X^{\Delta t,k}(\hat{s};x) - B(X^{\Delta t,k}(\hat{s};x'))\Big]\,d\bar{W}(s).\end{aligned}$$

Let $\mathcal{Q} = C^{\Delta t}\mathcal{P}_k$; then

$$\begin{aligned}&\mathbf{E}\|\mathcal{Q}(X^{\Delta t,k}(t;x) - X^{\Delta t,k}(t;x'))\|_{\mathbf{R}^d}^2 \leq K\|\mathcal{Q}\tilde{S}_k(t)(x-x')\|_{\mathbf{R}^d}^2 \\ &+ K\int_0^t \mathbf{E}\|\mathcal{Q}(X^{\Delta t,k}(\hat{s};x) - X^{\Delta t,k}(\hat{s};x'))\|_{\mathbf{R}^d}^2\,ds + K\int_0^t \mathbf{E}\|\pi_S(X^{\Delta t,k}(\hat{s};x) - X^{\Delta t,k}(\hat{s};x'))\|_{\mathbf{R}^d}^2\,ds \\ &+ K\int_0^t \mathbf{E}\|\pi_S(X^{\Delta t,k}(\hat{s};x) - X^{\Delta t,k}(\hat{s};x'))\|_{\mathbf{R}^d}^2\,ds.\end{aligned}$$

Hence, using $\|X\|_{\#} = (\|\mathcal{Q}X\|^2_{\mathbf{R}^d} + \|\pi_S X\|^2_{\mathbf{R}^d})^{1/2}$ and $\pi_S \tilde{S}_k(t)x = \pi_S x$,

$$\begin{aligned}\mathbf{E}\|X^{\Delta t,k}(t;x) - X^{\Delta t,k}(t;x')\|^2_{\#} \le & K\|\tilde{S}_k(t)(x-x')\|^2_{\#} + K\int_0^t \mathbf{E}\|X^{\Delta t,k}(\hat{s};x) - X^{\Delta t,k}(\hat{s};x')\|^2_{\#}\,ds\\ &+ K\int_0^t \mathbf{E}\|X^{\Delta t,k}(\hat{s};x) - X^{\Delta t,k}(\hat{s};x')\|^2_{\#}\,ds.\end{aligned}$$

From which,

$$\mathbf{E}\|X^{\Delta t,k}(t;x) - X^{\Delta t,k}(t;x')\|^2_{\#} \le K\|\tilde{S}_k(t)(x-x')\|^2_{\#}, \qquad 0 \le t \le T.$$

This completes the proof. *QED*

We next state properties of the interpolant and explain two lemmas that will be used to understand the approximation of integrals with respect to the measure a. 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 $k \to \infty$.

Lemma 3.3 *Let Hypothesis 1.1(i)-(ii) hold. For $0 \le t \le T$,*

$$\mathbf{E}\|\pi_S(X^{\Delta t,k}(t;x) - X^{\Delta t,k}(\hat{t};x))\|^2_{\mathbf{R}^d} \le K(1+\|x\|)^2\Delta t. \tag{3.12}$$

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

$$\begin{aligned}I := \mathbf{E}\Big[\Big\langle &\int_{i\Delta t}^{(i+1)\Delta t} a(dr)\pi_S(X^{\Delta t,k}(\,\hat{t}+r;x) - X^{\Delta t,k}(\hat{t}+\hat{r};x)),\\ &\int_{j\Delta t}^{(j+1)\Delta t} a(dr)\pi_S(X^{\Delta t,k}(\hat{t}+r;x) - X^{\Delta t,k}(\hat{t}+\hat{r};x))\Big\rangle\Big] \le K(1+\|x\|)^2\Delta t^4.\end{aligned} \tag{3.13}$$

Proof The process $X^{\Delta t,k}$ solves (3.3) and hence satisfies

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

where $\tilde{S}_k$ is the semigroup with infinitesimal generator $\tilde{A}_k$. As $\pi_S\tilde{S}_k(t)x = \pi_S x$ and $\|\tilde{S}_k\|_{\mathcal{L}(H,H)}$ is bounded and $|t-\hat{t}| \le \Delta t$, this implies (3.12).

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 $X^{\Delta t,k}(\hat{t}+j\Delta t + r)$ 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,

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

Let $X^{\Delta t,k}(t_2,t_1;x)$ be the solution to (3.3) 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

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

Let $\mathcal{Q} = C^{\Delta t}\mathcal{P}_k$. From (3.10) and Hypothesis 1.1(ii), for $0 \le t_2 - t_1 \le T$,

$$\begin{aligned}\|\Gamma_{t_2,t_1}(x) - \Gamma_{t_2,t_1}(x')\| \le& K\Big[\|\mathcal{Q}(X^{\Delta t,k}(\hat{t}_2,\hat{t}_1;x) - X^{\Delta t,k}(\hat{t}_2,\hat{t}_1;x'))\|_{\mathbf{R}^d}\,\Delta t^2\\ &+ \|\pi_S(X^{\Delta t,k}(\hat{t}_2,\hat{t}_1;x) - X^{\Delta t,k}(\hat{t}_2,\hat{t}_1;x'))\|_{\mathbf{R}^d}\,\Delta t^2\Big]\\ \le& K\|X^{\Delta t,k}(\hat{t}_2,\hat{t}_1;x) - X^{\Delta t,k}(\hat{t}_2,\hat{t}_1;x')\|_{\#}\Delta t^2\end{aligned}$$

where $\|\cdot\|_{\#} = (\|\mathcal{Q}\cdot\|_{\mathbf{R}^d}^2 + \|\pi_S\cdot\|_{\mathbf{R}^d}^2)^{1/2}$. Now, from (3.10),

$$(\mathbf{E}\|\Gamma_{t_2,t_1}(x) - \Gamma_{t_2,t_1}(x')\|^2)^{1/2} \le K\Big[\|\mathcal{Q}\tilde{S}_k(t)(x-x')\|_{\mathbf{R}^d}^2 + \|\pi_S(x-x')\|_{\mathbf{R}^d}^2\Big]^{1/2}\Delta t^2.$$

Consider the case $\Delta t = t_1 \le t_2$ and let $y = X^{\Delta t,k}(\Delta t, x)$ and $y' = \mathbf{E}y$. Then from (3.8)–(3.9) with $\mathcal{Q} = C^{\Delta t}\mathcal{P}_k\tilde{S}_k(t)$,

$$(\mathbf{E}\|\Gamma_{t_2,t_1}(y) - \Gamma_{t_2,t_1}(y')\|^2)^{1/2} \le K\Delta t^{5/2}. \tag{3.14}$$

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 (3.14),

$$\begin{aligned}|I_{hard}| \le& \Big(\mathbf{E}\Big\|\Gamma_{(i-j)\Delta t,\Delta t}(X^{\Delta t,k}(\Delta t;x)) - \Gamma_{(i-j)\Delta t,\Delta t}(\mathbf{E}X^{\Delta t,k}(\Delta t;x))\Big\|_{\mathbf{R}^d}^2\\ &\times \mathbf{E}\Big\|\int_0^{\Delta t} a(dr)\int_0^r \pi_S B(x)\,d\bar{W}(s)\Big\|_{\mathbf{R}^d}^2\Big)^{1/2}\\ \le& K\Big(\Delta t^5\Delta t^3\Big)^{1/2} = K\Delta t^4.\end{aligned}$$

QED

Lemma 3.4 *Let Hypothesis 1.1(i)–(ii) hold. Suppose that the delay function of the initial data is 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* $-\tau \le s \le 0$ *and for* $-\lfloor \tau/\Delta t \rfloor \le i \ne j \le -1$,

$$\mathbf{E}\Big\| \int_{i\Delta t}^{(i+1)\Delta t} a(dr)\alpha(\hat{s},r;x)\Big\|^2_{\mathbf{R}^d} \le K(1+\|x\|+\|\pi_D x\|_{Lip})^2 \Delta t^3 + o(k^{-1}). \tag{3.15}$$

$$\begin{aligned}\mathbf{E}\Big[\Big\langle \int_{i\Delta t}^{(i+1)\Delta t} a(dr)\alpha(\hat{t},r;x), \int_{j\Delta t}^{(j+1)\Delta t} a(dr)\alpha(\hat{t},r;x)\Big\rangle\Big] \\ \le K(1+\|x\|+\|\pi_D x\|_{Lip})^2\Delta t^4 + o(k^{-1}).\end{aligned} \tag{3.16}$$

Proof To prove the lemma, we interpret the inequalities in Lemma 3.3 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 (3.1). 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 $X^{\Delta t}_S(t+s;x) = X^{\Delta t}_D(t;x)(s)$ 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 follow immediately from Lemma 3.3 using

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

for (3.15).

By applying Proposition 2.5, $\mathcal{P}_k \to \pi_D$ in $\mathcal{L}(H, L_2([-\tau,0],\mathbf{R}^d))$. Now, from Lemma 3.1, we have $\mathcal{P}_k X^{\Delta t,k}(t;x)(r) \to X^{\Delta t}_S(t+r;x)$. The left hand side of both inequalities in the lemma are continuous with respect to perturbation in $L_2([-\tau,0],\mathbf{R}^d)$ of $\alpha(\hat{s},\cdot;x)$. Consequently, the introduction of the $\mathcal{P}_k$ term for π_D 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*

3.1 Derivatives in the $\tilde{A}_k h$ direction

We provide two lemmas giving boundedness of spatial derivatives of $X^k(t;x)$ in the direction $\tilde{A}_k h$, which are uniform in k and improve on the bounds given in Theorem 2.3.

Lemma 3.5 *Let Hypothesis 1.1(i)–(ii) hold and consider* $p \ge 2$. *For* $h, x \in H$, *the following holds*

$$(\mathbf{E}\sup_{0\le t\le T} \|\pi_S X^k_x(t;x)\tilde{A}_k h\|^p_{\mathbf{R}^d})^{1/p} \le K\|h\|. \tag{3.17}$$

Proof Let $\xi^{k,h}(t;x) := X^k_x(t;x)h$, where $h \in H$. This is a strong solution of

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

with initial condition $\xi^{k,h}(0) = h$. We are interested in $\xi^{k,\tilde{A}_k h}(t;x)$, in this case the Variation of Constants formula states

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

We first look at $S_k(t)\tilde{A}_k h$. As A_k and S_k commute and as, under Hypothesis 1.1(ii), the first component of A_k is bounded uniformly from H to $\mathbf{R}^d$, we see

$$\|\pi_S S_k(t)A_k h\|_{\mathbf{R}^d} = \|\pi_S A_k S_k(t)h\|_{\mathbf{R}^d} \leq K\|S_k(t)h\|,$$

for a constant K independent of k. Further,

$$\|\pi_S S_k(t)\tilde{A}_k h\|_{\mathbf{R}^d} \leq \|\pi_S S_k(t)A_k h\|_{\mathbf{R}^d} + \|\pi_S S_k(t)(\tilde{A}_k - A_k)h\|_{\mathbf{R}^d}.$$

Because $\tilde{A}_k - A_k$ is bounded in $\mathcal{L}(H,H)$ uniformly in k, we conclude that $\|\pi_S S_k(t)\tilde{A}_k h\|_{\mathbf{R}^d} \leq K\|S_k(t)\|_{\mathcal{L}(H,H)}\|h\|$. Together with the Burkholder-Davis-Gundy inequality, this gives

$$\begin{aligned}&(\mathbf{E} \sup_{0\leq s\leq t} \|\pi_S \xi^{k,\tilde{A}_k h}(s;x)\|^p_{\mathbf{R}^d})^{1/p} \\ &\leq \sup_{0\leq s\leq t} K\|S_k(s)\|_{\mathcal{L}(H,H)}\|h\| + \Big(\mathbf{E}\Big[\int_0^{t'} \|\pi_S S_k(t'-s)F_x(X^k(s;x))\xi^{k,\tilde{A}_k h}(s;x)\|_{\mathbf{R}^d}\, ds\Big]^p\Big)^{1/p} \\ &\quad + K\Big(\mathbf{E}\Big[\int_0^{t'} \|S_k(t'-s)B_x(X^k(s;x))\xi^{k,\tilde{A}_k h}(s;x)\|^2_{HS}\, ds\Big]^{p/2}\Big)^{1/p}\end{aligned}$$

(using the boundedness of F_x, B_x, $S_k(t)$, and $p \geq 2$)

$$\leq K\|h\| + \Big(\mathbf{E}\Big[\sup_{0\leq t'\leq t}\int_0^{t'} \|\pi_S\xi^{k,\tilde{A}_k h}(s;x)\|_{\mathbf{R}^d}\, ds\Big]^p\Big)^{1/p} + K\Big(\mathbf{E}\sup_{0\leq t'\leq t}\int_0^{t'} \|\pi_S\xi^{k,\tilde{A}_k h}(s;x)\|^p_{\mathbf{R}^d}\, ds\Big)^{1/p}.$$

Thus,

$$(\mathbf{E}\sup_{0\leq s\leq t}\|\pi_S\xi^{k,\tilde{A}_k h}(s;x)\|^p_{\mathbf{R}^d})^{1/p} \leq K\|h\| + K\int_0^t \Big(\mathbf{E}\sup_{0\leq s'\leq s}\|\pi_S\xi^{k,\tilde{A}_k h}(s';x)\|^p_{\mathbf{R}^d}\Big)^{1/p} ds.$$

By applying the Gronwall Lemma, for each $T > 0$, there exists $K > 0$ such that for each k

$$\Big(\mathbf{E}\sup_{0\leq t\leq T}\|\pi_S\xi^{k,\tilde{A}_k h}(t;x)\|^p_{\mathbf{R}^d}\Big)^{1/p} \leq K\|h\|, \quad 0\leq t\leq T. \tag{3.18}$$

With this inequality in hand, the result (3.17) follows.

QED

Lemma 3.6 *Let Hypothesis 1.1(i)–(ii) hold and consider $p \geq 2$. For $h, g \in H$ and $0 \leq t \leq T$,*

$$(\mathbf{E}\|\pi_S X^k_{xx}(t;x)(h,\tilde{A}_k g)\|^p_{\mathbf{R}^d})^{1/p} \leq K\|g\|\,\|h\|.$$

Proof Denote $X^k_x(t;x)h$ by $\xi^{k,h}(t;x)$ and $X^k_{xx}(t;x)(h,g)$ by $\eta^{k,(h,g)}(t;x)$. Then $\eta^{k,(h,\tilde{A}_k g)}$ satisfies the following variational equation:

$$
\begin{aligned}
&d\eta^{k,(h,\tilde{A}_k g)}(t;x)\\
=&\Big[A_k\eta^{k,(h,\tilde{A}_k g)}(t;x)+F_{xx}(X^k(t;x))(\xi^{k,h}(t;x),\xi^{k,\tilde{A}_k g}(t;x))+F_x(X^k(t;x))\eta^{k,(h,\tilde{A}_k g)}(t;x)\Big]\,dt\\
&+(B_{xx}(X^k(t;x))(\xi^{k,h}(t;x),\xi^{k,\tilde{A}_k g}(t;x))+B_x(X^k(t;x))\eta^{k,(h,\tilde{A}_k g)}(t;x))\,dW(t),
\end{aligned}
$$

where $\eta^{k,(h,\tilde{A}_k g)}(0)=0$. Again the Variation of Constants formula yields a bound on $\eta^{k,(h,\tilde{A}_k g)}$:

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

Thus, arguing as in the previous lemma,

$$
\begin{aligned}
\Big(\mathbf{E}\Big[\sup_{0\le s\le t}\|\pi_S\eta^{k,(h,\tilde{A}_k g)}(s;x)\|^p_{\mathbf{R}^d}\Big]\Big)^{1/p}\le& K\Big(\mathbf{E}\Big[\int_0^t\|\pi_S\xi^{k,h}(s;x)\|^p_{\mathbf{R}^d}\,\|\pi_S\xi^{k,\tilde{A}_k g}(s;x)\|^p_{\mathbf{R}^d}\,ds\Big]\Big)^{1/p}\\
&+K\Big(\mathbf{E}\Big[\int_0^t\|\pi_S\eta^{k,(h,\tilde{A}_k g)}(s;x)\|^p_{\mathbf{R}^d}\,ds\Big]\Big)^{1/p}\\
&+K\Big(\mathbf{E}\Big[\int_0^t\|\pi_S\xi^{k,h}(s;x)\|^p_{\mathbf{R}^d}\|\pi_S\xi^{k,\tilde{A}_k g}(s;x)\|^p_{\mathbf{R}^d}\,ds\Big]\Big)^{1/p}\\
&+K\Big(\mathbf{E}\Big[\int_0^t\|\pi_S\eta^{k,(h,\tilde{A}_k g)}(s;x)\|^p_{\mathbf{R}^d}\,ds\Big]\Big)^{1/p}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\Big(\mathbf{E}\sup_{0\le s\le t}\|\pi_S\eta^{k,(h,\tilde{A}_k g)}(s;x)\|^p_{\mathbf{R}^d}\Big)^{1/p}\le& K\int_0^t\Big(\mathbf{E}\sup_{0\le s'\le t}\|\pi_S\xi^{k,h}(s';x)\|^p_{\mathbf{R}^d}\,\|\pi_S\xi^{k,\tilde{A}_k g}(s';x)\|^p_{\mathbf{R}^d}\Big)^{1/p}\,ds\\
&+K\int_0^t\Big(\mathbf{E}\sup_{0\le s'\le s}\|\pi_S\eta^{k,(h,\tilde{A}_k g)}(s';x)\|^p_{\mathbf{R}^d}\Big)^{1/p}\,ds.
\end{aligned}
$$

Now, by Cauchy-Schwarz,

$$
\begin{aligned}
\Big(\mathbf{E}\sup_{0\le s'\le t}\|\pi_S\xi^{k,h}(s';x)\|^p_{\mathbf{R}^d}\,\|\pi_S\xi^{k,\tilde{A}_k g}(s';x)\|^p_{\mathbf{R}^d}\Big)^{1/p}\le&\Big(\mathbf{E}\sup_{0\le s'\le t}\|\pi_S\xi^{k,h}(s';x)\|^{2p}_{\mathbf{R}^d}\Big)^{1/2p}\\
&\times\Big(\mathbf{E}\sup_{0\le s'\le t}\|\pi_S\xi^{k,\tilde{A}_k g}(s';x)\|^{2p}_{\mathbf{R}^d}\Big)^{1/2p}.
\end{aligned}
$$

Apply the estimate on $\xi^{k,h}$ in Lemma 2.3 and on $\xi^{k,\tilde{A}_k h}$ in Lemma 3.5, to derive for $0\le t\le T$

$$
(\mathbf{E}\sup_{0\le s\le t}\|\pi_S\eta^{k,(h,\tilde{A}_k g)}(t;x)\|^p_{\mathbf{R}^d})^{1/p}\le K\,\|h\|\,\|g\|+K\int_0^t\Big(\mathbf{E}\sup_{0\le s'\le s}\|\pi_S\eta^{k,(h,g)}(s';x)\|^p_{\mathbf{R}^d}\,ds\Big)^{1/p}.
$$

Apply Gronwall's inequality to complete the proof. *QED*

4 Corollaries of the Itô calculus

We formulate and prove two corollaries of the Itô calculus, concerning the time regularity for certain functionals of solutions of a class of stochastic evolution equations.

The first corollary is set up for an abstract equation, but we have in mind $Z(t;x) = (X^k(t;x), X^k_x(t;x)h)$, which obeys

$$\begin{aligned} dZ_1 =& \Big[AZ_1 + F(Z_1)\Big]\, dt + B(Z_1)\, dW(t), \quad Z_1(0) = x \\ dZ_2 =& \Big[AZ_2 + F_x(Z_1)Z_2\Big]\, dt + B_x(Z_1)Z_2\, dW(t), \quad Z_2(0) = h. \end{aligned} \tag{4.1}$$

A similar equation can be written down for the second derivative $X^k_{xx}(t;x)(h,g)$ involving four equations. Let H^m denote the product space $H \times H \times \cdots \times H$ (m times).

Corollary 4.1 *Consider locally Lipschitz functions $\bar F_i : H^m \to H$ and $\bar B_i : H^m \to L^0_2$ 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*

$$dZ^k_i = \Big[A_k Z^k_i + \bar F_i(Z^k)\Big]\, dt + \bar B_i(Z^k)\, dW, \quad Z^k_i(0) = z^k_i(x). \tag{4.2}$$

Suppose further that for some $p \ge 2$ and for each $0 \le t \le T$ and $i = 1,\dots,m$, we have

$$\mathbf{E}|Z^k_i(t;x)|^p_\star \le K(1 + \|x\|^p) \tag{4.3}$$

and

$$\mathbf{E}\|\bar F_i(Z^k(t;x))\|^2 \le K(1 + \|x\|)^2, \qquad \mathbf{E}\|\bar B_i(Z^k(t;x))\|^2_{HS} \le K. \tag{4.4}$$

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

$$|G_i(Z)| \le K\Big(1 + \sum_{\ell=1}^m \|\pi_S Z_\ell\|^{p-1}_{\mathbf{R}^d}\Big), \qquad |G_{ij}(Z)| \le K\Big(1 + \sum_{\ell=1}^m \|\pi_S Z_\ell\|^{p-2}_{\mathbf{R}^d}\Big). \tag{4.5}$$

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

$$|w^k_t(t,x)| \le K(1 + \|x\|^p), \qquad 0 \le t \le T.$$

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

$$\begin{aligned} w^k(t,x) - w^k(0,x) =& \mathbf{E}\sum_{i=1}^m \int_0^t G_i(Z^k(s;x))(A_k Z^k_i(s;x) + \bar F_i(Z^k(s;x)))\, ds \\ &+ \tfrac12 \sum_{i,j=1}^m \mathbf{E}\int_0^t \operatorname{Tr} G_{ij}(Z^k(s;x))\bar B_i(Z^k(s;x))\bar B_j(Z^k(s;x))^*\, ds. \end{aligned}$$

We attain limits from the dominated convergence theorem because, under (4.3) and (4.5),

$$\begin{aligned} \mathbf{E}\Big|\int_0^t G_i(Z^k(s;x))(A_k Z^k_i(s;x) + \bar F_i(Z^k(s;x)))\, ds\Big| \le& K\, t\, (1 + \|x\|^p) \\ \tfrac12\mathbf{E}\Big|\int_0^t \operatorname{Tr} G_{ij}(Z^k(s))\bar B_i(Z^k(s;x))\bar B_j(Z^k(s;x))^*\, ds\Big| \le& K\, t\, (1 + \|x\|^{p-2}). \end{aligned}$$

Thus,

$$\begin{aligned} w_t^k(t,x) =&\mathbf{E}\sum_{i=1}^m G_i(Z^k(t;x))(A_k Z_i^k(t;x)+\bar{F}_i(Z^k(t;x)))\\ &+\tfrac{1}{2}\sum_{i,j=1}^m \operatorname{Tr} G_{ij}(Z^k(t;x))\bar{B}_i(Z^k(t;x))\bar{B}_j(Z^k(t;x))^*. \end{aligned} \tag{4.6}$$

With this expression, it is easy to derive the required growth bound on $w_t^k(t,x)$ in $\|x\|$ uniform in $k\to\infty$.

To establish uniform continuity of $w_t^k(t,x)$ with respect to time, consider

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

Hence,

$$\|S_k(t)x-x\| \le K\|A_k\|_{\mathcal{L}(H,H)}\|x\|, \qquad 0\le t\le T.$$

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

$$\mathbf{E}\|Z_i(t;x)-Z_i(t';x)\|^2 \le K_k|t-t'|, \quad 0\le t,t'\le T, \quad \|x\|\le R. \tag{4.7}$$

We use here and below K_k to denote a generic constant that is independent of the initial data x but may diverge in the limit $k\to\infty$. In the above inequality, the constant K_k diverges because it depends on $\|A_k\|$.

Fix R the radius of a ball in H and consider $x\in H$ with $\|x\|\le R$. For any $\delta>0$, there exists L large by (4.3) and the Chebyshev inequality so that if $\mathcal{O}:=\{\|\pi_S Z_i(t;x)\|_{\mathbf{R}^d}\le L, 0\le t\le T, i=1,\dots,m\}$, the probability $\mathbf{P}(\mathcal{O})\ge 1-\delta$. Consider the expectations defining w_t^k in (4.6) split as a sum over $\mathcal{O}$ and $\mathcal{O}^c$. On the set $\mathcal{O}$, we have that $G_i, G_{ij}, \bar{F}, \bar{B}$ are all Lipschitz and the expectations in the difference $w_t^k(t,x)-w_t^k(t',x)$ may be bounded by $K_k|t-t'|^{1/2}$ using (4.7). By using (4.3), the expectations on the set $\mathcal{O}^c$ are bounded by $\delta|w_t^k(t,x)|\le \delta K(1+R^p)$. Thus to show uniform continuity on the bounded subset 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 $|t-t'|\le \epsilon^2/2K_k^2$ and $0\le t,t'\le T$,

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

This gives uniform continuity of w_t^k 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 4.2 *Let Hypothesis 1.1(i)–(ii) hold. Consider the strong solution $X^{\Delta t,k}(t;x)$ of (3.3) under the condition that F is globally Lipschitz, that B is bounded, and that the initial data $x=(Y_S,Y_D)$ where Y_D is continuous. 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 $\mathbf{R}^+\times H$. Fix $p\ge 2$. Further suppose, for $0\le t\le T$ and for a constant K, that*

$$|w_t(t,x)| \le K(1+\|x\|^p), \tag{4.8}$$

$$\|w_x(t,x)\|_{\mathcal{L}(H,\mathbf{R})}, \quad \|w_{xx}(t,x)\|_{\mathcal{L}(H\times H,\mathbf{R})} \le K(1+\|x\|^{p-1}), \tag{4.9}$$

and that, for $h \in \mathcal{D}(A)$ and some $\mathcal{Q} \in \mathcal{L}(H', \mathbf{R}^d)$,

$$|w_x(t,x)\tilde{A}_k h| \le K(1+\|x\|^{p-1}) \left[\|h\| + \|\mathcal{Q}h\|_{\mathbf{R}^d}\right], \tag{4.10}$$

Then,

$$\left|\mathbf{E}\left[w(s, X^{\Delta t,k}(s;x)) - w(\hat{s}, X^{\Delta t,k}(\hat{s};x))\right]\right| \le K_x \Delta t, \quad 0 \le s \le T.$$

where the constant K_x is independent of k but depends on x as follows

$$K_x \le K(1+\|x\|^p) + K(1+\|x\|^{p-1}) \sup_{-\tau \le s \le 0} \|Y_D(s)\|_{\mathbf{R}^d}.$$

Proof This is Lemma 14.1.6 of [14]. Apply the Itô formula to the strong solution $X^{\Delta t,k}$:

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

Now using (4.8)–(4.10) with the boundedness of B and the Lipschitz property of F, we have

$$\begin{aligned}
&\left|\mathbf{E}\left[w(s, X^{\Delta t,k}(s;x)) - w(\hat{s}, X^{\Delta t,k}(\hat{s};x))\right]\right|\\
&\quad\le \mathbf{E}\Big[\int_{\hat{s}}^{s} K(1+\|X^{\Delta t,k}(s';x)\|^p) + K(1+\|X^{\Delta t,k}(s';x)\|^{p-1})\left[\|X^{\Delta t,k}(\hat{s};x)\| + \|\mathcal{Q}X^{\Delta t,k}(\hat{s};x)\|_{\mathbf{R}^d}\right]\\
&\qquad + K(1+\|X^{\Delta t,k}(s';x)\|^{p-1})(1+\|X^{\Delta t,k}(\hat{s};x)\|)\, ds'\Big]\\
&\quad\le \int_{\hat{s}}^{s} K(1+\mathbf{E}\|X^{\Delta t,k}(s';x)\|^p) + K\left[\mathbf{E}\|X^{\Delta t,k}(\hat{s};x)\| + \mathbf{E}\|\mathcal{Q}X^{\Delta t,k}(\hat{s};x)\|_{\mathbf{R}^d}\right]\\
&\qquad + K(\mathbf{E}\|X^{\Delta t,k}(s';x)\|^{2(p-1)})^{1/2}\left[(\mathbf{E}\|X^{\Delta t,k}(\hat{s};x)\|^2)^{1/2} + (\mathbf{E}\|\mathcal{Q}X^{\Delta t,k}(\hat{s};x)\|^2_{\mathbf{R}^d})^{1/2}\right]\\
&\qquad + K\mathbf{E}\left[(1+\|X^{\Delta t,k}(s';x)\|^{p-1})^2\right]^{1/2}\mathbf{E}\left[(1+\|X^{\Delta t,k}(\hat{s};x)\|)^2\right]^{1/2} ds'.
\end{aligned}$$

By using Lemma 3.2, we have

$$\begin{aligned}
&\left|\mathbf{E}\left[w(s, X^{\Delta t,k}(s;x)) - w(\hat{s}, X^{\Delta t,k}(\hat{s};x))\right]\right|\\
&\quad\le K\int_{\hat{s}}^{s} (1+\|x\|^p) + (1+\|x\|^{p-1})(1+\|x\| + \sup_{-\tau \le s \le 0}\|Y_D(s)\|_{\mathbf{R}^d})\, ds'.
\end{aligned}$$

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

5 The Kolmogorov equation

We introduce the Kolmogorov equation for the regularised delay equation (2.4). The background theory is developed in Da Prato and Zabczyk [7], where further references are also given. The Kolmogorov equation is described in Theorem 5.1. We also discuss the regularity of the terms in the equation so that the Itô formula applies to $v(t,x) = \mathbf{E}\phi(\pi_S X(t;x))$ and to the terms in the Kolmogorov equation.

Theorem 5.1 *Let $\phi \in \mathcal{G}_p$ for $p \geq 4$ and set $v^k(t,x) := \mathbf{E}\phi(\pi_S X^k(t;x))$. Then v^k satisfies for $x \in H$ and $0 \leq t \leq T$*

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

The functional v^k has two spatial derivatives and one time derivative, which are uniformly continuously differentiable on bounded subsets of $\mathbf{R}^+ \times H$.

Proof Because A_k is a bounded operator, the Kolmogorov equation is a special case of that derived in [7]. The spatial regularity is described in Corollary 2.4. To establish time regularity, apply Corollary 4.1 with $Z^k(t;x) = X^k(t;x)$. *QED*

Proposition 5.2 *Let Hypothesis 1.1(i). Let $v^k(t,x) := \mathbf{E}\phi(\pi_S X^k(t;x))$ where $\phi \in \mathcal{G}_p$, $p \geq 4$.*

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

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

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

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

Proof The differentiability and bounds on 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 $v_x^k(t,x) = \mathbf{E}\phi'(\pi_S X^k(t;x))\pi_S X_x^k(t;x)$. Thus,

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

where $G(Z_1,Z_2) = \phi'(\pi_S Z_1)\pi_S Z_2$, $Z_1^k(t;x) = X^k(t;x)$, and $Z_2^k(t;x) = X_x^k(t;x)\psi(x)$.

Note that (Z_1^k, Z_2^k) satisfies (4.1) with $h = \psi(x)$. The growth condition (4.3) is given by Theorem 2.3. The drift and diffusion functions in (4.1) are locally Lipschitz and obey (4.4) by using the boundedness of the derivatives given in Hypothesis 1.1(i). The regularity of the test functional G is easily derived from the conditions on ϕ. Hence, Corollary 4.1 applies in this situation. Thus, we conclude that $v_x^k(t,x)\psi(x)$ is uniformly continuously differentiable in time on bounded subsets of $\mathbf{R}^+ \times H$.

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

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

Thus,

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

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

$$\begin{aligned}\operatorname{Tr} v^k_{xx}(t,x)\Psi(x)\Psi^*(x) =&\mathbf{E}\Big[\sum_{i=1}^d \phi''(\pi_S X^k(t;x))(\pi_S X^k_x(t;x)\Psi(x)e_i, \pi_S X^k_x(t;x)\Psi(x)e_i)\\ &\quad + \phi'(\pi_S X^k(t;x))\pi_S X^k_{xx}(t)(\Psi(x)e_i,\Psi(x)e_i)\Big]\\ =&\sum_{i=1}^d \mathbf{E}G(Z^k_1(t;x), Z^k_{2,i}(t;x), Z^k_{3,i}(t;x))\end{aligned}$$

for $G(Z_1,Z_2,Z_3) := \phi''(\pi_S Z_1)(\pi_S Z_2,\pi_S Z_2) + \phi'(\pi_S Z_1)\pi_S Z_3$ and $Z^k_1(t;x) := X^k(t;x)$ and

$$Z^k_{2,i}(t;x) := X^k_x(t;x)\Psi(x)e_i, \quad Z^k_{3,i}(t;x) := X^k_{xx}(t;x)(\Psi(x)e_i,\Psi(x)e_i).$$

Again, it can be shown that the processes $Z^k_1, Z^k_{2,i}, Z^k_{3,i}$ and the test function G satisfy the hypothesis of Corollary 4.1 (for each i, case $m=3$). The sum is finite, which means regularity of $\mathbf{E}G(Z_1,Z_{2,i},Z_{3,i})$ for each i gives the same regularity for $\operatorname{Tr} v^k_{xx}(t,x)\Psi(x)\Psi^*(x)$. *QED*

6 Proof of Theorem 1.2

The following argument gives weak convergence of order Δt of the forward Euler method. The argument follows that of Kloeden-Platen [14], Theorem 14.1.5.

Proof (of Theorem 1.2) Consider $v^k(t,x) := \mathbf{E}(\phi(\pi_S X^k(T-t;x))$ 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 5.1, we have that $\mathcal{L}^k v^k(t,x) = 0$ and that v^k satisfies the hypothesis of Itô's formula. Apply the Itô formula to the approximations $X^{\Delta t,k}$ defined in (3.3):

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

(subtracting off $0 = \mathcal{L}^k v^k(s, X^{\Delta t,k}(s;x)))$

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

Define for $h_1, h_2 \in H$

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

Clearly,

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

and hence

$$\begin{aligned}
&\Big|\mathbf{E}\phi(\pi_S X^{\Delta t,k}(T;x)) - \mathbf{E}\phi(\pi_S X^k(T;x))\Big| \\
\le& \int_0^T \sum_{i=1}^4 |\mathbf{E}w_i^k(s, X^{\Delta t,k}(s;x); X^{\Delta t,k}(\hat{s};x)) - \mathbf{E}w_i^k(s, X^{\Delta t,k}(\hat{s};x); X^{\Delta t,k}(\hat{s};x))|\, ds \qquad (6.1)\\
&+ \Big|\int_0^T \mathbf{E}v_x(s, X^{\Delta t,k}(s;x)) \begin{pmatrix}(C - C^{\Delta t})\mathcal{P}_k X^{\Delta t,k}(\hat{s};x) \\ 0\end{pmatrix} ds\Big|.
\end{aligned}$$

By definition of $\mathcal{C}$ and $\mathcal{C}^{\Delta t}$, the modulus of the integrand of the last term 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 (3.7). 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 3.4, the cross terms ($i \neq j$) obey

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

and by the first part of Lemma 3.4 and Hypothesis 1.1(ii) the diagonal terms

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

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 (6.1) is bounded by $K(1+\|x\|^p+\|x\|^{p-1}\|\pi_D x\|_{\mathrm{Lip}})\Delta t+o(k^{-1})$.

We wish to apply Lemma 4.2 to show that each pair of terms in w_i in (6.1) is order Δt. Because $s>\hat{s}$, it is sufficient to apply the lemma to $w^k(t,x)=w_i^k(t,x;h_2)$. We now verify the hypothesis of Lemma 4.2.

We require that w^k, w_t^k, w_x^k, w_{xx}^k exist, be uniformly continuous on bounded subsets of $\mathbf{R}^+\times H$, and obey the growth bounds (4.8)–(4.9) specified in Lemma 4.2. In each case, this is a consequence of Proposition 5.2. Part (i) covers w_1^k and w_2^k, while part (ii) covers w_3^k and w_4^k. To establish the hypothesis of this proposition, we use the Lipschitz continuity of F, the boundedness of B, and the continuity of $[C\mathcal{P}_k h, 0]$ implied by Hypothesis 1.1(ii).

We further require that for some operator $\mathcal{Q}\in\mathcal{L}(H',\mathbf{R}^d)$ that

$$|w_x^k(t,x)\tilde{A}_k h|\le K(1+\|x\|^{p-1})\left[\|h\|+\|\mathcal{Q}h\|_{\mathbf{R}^d}\right],\quad \text{for } h\in\mathcal{D}(A).$$

We look at $w^k(t,x)=w_2^k(t,x;x)$ in detail; the others are similar. Note

$$\begin{aligned} w_x^k(t,x)\tilde{A}_k h =& v_x^k(t,x)[C\mathcal{P}_k\tilde{A}_k h,0]^T+v_x^k(t,x)F_x(x)\tilde{A}_k h\\ &+v_{xx}^k(t,x)([C\mathcal{P}_k x,0]^T,\tilde{A}_k h)+v_{xx}^k(t,x)(F(x),\tilde{A}_k h). \end{aligned} \tag{6.2}$$

Now, by Corollary 2.4,

$$|v_x^k(t,x)[C\mathcal{P}_k\tilde{A}_k h,0]^T|\le K(1+\|x\|^{p-1})\|C\mathcal{P}_k\tilde{A}_k h\|_{\mathbf{R}^d}.$$

Let $\Psi(t):=\int_0^t(\mathcal{P}_k\tilde{A}_k h)(s)\,ds$; then $C\Psi'=C\mathcal{P}_k\tilde{A}_k h$. Using Hypothesis 1.1(ii),

$$C\mathcal{P}_k\tilde{A}_k h=\int_{-\tau}^0 a(s)\Psi'(s)\,ds=\Big[a(s)\Psi(s)\Big]_{-\tau}^0-\int_{-\tau}^0 a'(s)\Psi(s)\,ds.$$

Note that $\Psi=\mathcal{P}_k X^*$, where $(\pi_D X^*)(t)=\int_0^t(\pi_D\tilde{A}_k h)(s)\,ds=(\mathcal{P}_k h)(t)-(\mathcal{P}_k h)(0)$ and $\pi_S X^*=0$. Thus, using smoothness of the density $a(s)$ and $\|\mathcal{P}_k h\|_{L_2([-\tau,0],\mathbf{R}^d)}\le\|h\|$,

$$\begin{aligned} \|C\mathcal{P}_k\tilde{A}_k h\|_{\mathbf{R}^d}\le& K\Big\|\Psi(-\tau)-\Psi(0)\Big\|_{\mathbf{R}^d}+K\|\Psi\|_{L_2([-\tau,0],\mathbf{R}^d)}\\ \le& K\|\mathcal{Q}h\|_{\mathbf{R}^d}+K\|h\|, \end{aligned}$$

where $\mathcal{Q}h=(\mathcal{P}_k X^*)(-\tau)$. Notice that

$$\|\mathcal{Q}h\|_{\mathbf{R}^d}\le K\|X^*\|_{H'}\le K\|h\|_{H'}$$

and hence $Q\in\mathcal{L}(H',\mathbf{R}^d)$. We conclude that the first term in (6.2) is bounded by $K(1+\|x\|^{p-1})\,(\|h\|+\|\mathcal{Q}h\|_{\mathbf{R}^d})$.

The second term vanishes because $F_x(x)\tilde{A}_k=0$, using the fact that F is independent of the delay and the definition of $\tilde{A}_k$ (see (3.2)).

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

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

Theorem 2.3 gives bounds on $\xi^{k,h}$, Lemma 3.5 on $\xi^{k,\tilde{A}_k h}$ and Lemma 3.6 on $\eta^{k,(h,\tilde{A}_k g)}$. There results a bound on both terms of the form $(1+\|x\|^{p-1})\|h\|$. We conclude that the required bound on $|w_x(t,x)\tilde{A}_k h|$ holds.

Thus, Lemma 4.2 applies to the terms in the summation in (6.1), giving bounds of the form

$$K\Big[1 + \|x\|^p + (1 + \|x\|^{p-1}) \sup_{-\tau \le s \le 0} \|Y_D(s)\|_{\mathbf{R}^d}\Big]\Delta t.$$

In our case,

$$\sup_{-\tau \le s \le 0} \|Y_D(s)\|_{\mathbf{R}^d} \le \|x\| + \tau \|Y_D\|_{\mathrm{Lip}}.$$

Taking this observation with the bound for the last term in (6.1), we have a bound on the weak error in the Yosida approximation of the form $K_x \Delta t + o(k^{-1})$, where K_x is described in (1.3).

We have shown that

$$|\mathbf{E}\phi(\pi_S X^{\Delta t,k}(T;x)) - \mathbf{E}\phi(\pi_S X^k(T;x))| \le K_x \Delta t + o(k^{-1}). \tag{6.3}$$

Note that $X^k(t;x) \to X(t;x)$ (resp., $X^{\Delta t,k}(t;x) \to X(t;x)$) almost surely by Lemma 2.6 (resp., 3.1) and that $\mathbf{E}|\phi(\pi_S X^k(t;x))|$ can be bounded uniformly in k by using the properties of ϕ with Theorem 2.3 (resp., Lemma 3.2). Consequently, dominated convergence applies and $\mathbf{E}\phi(\pi_S X^k(t;x)) \to \mathbf{E}\phi(\pi_S X(t;x))$ (resp., $\mathbf{E}\phi(\pi_S X^{\Delta t,k}(t;x)) \to \mathbf{E}\phi(\pi_S X^{\Delta t}(t;x))$) as $k \to \infty$. Hence, taking the limit in (6.3) as $k \to \infty$ completes the proof.

QED

7 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 moments, that is we compute $\mathbf{E}\phi(Y(T))$ for $\phi(Y) = Y$ where $Y(T)$ denotes a solution of (1.1).

Example 7.1 Consider

$$dY(t) = \Big[\int_{t-1}^{t} Y(s)\ ds + \exp(-1)Y(t)\Big]\ dt + (\sigma_1\ + \sigma_2\ Y(t))dW(t), \tag{7.1}$$

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

$$m'(t) = \int_{t-1}^{t} m(s)\ ds + \exp(-1)m(t), \tag{7.2}$$

with initial condition

$$m(s) = \exp(s) \text{ for } -1 \le s \le 0. \tag{7.3}$$

Equation (7.2) subject to (7.3) 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

$$Y_{n+1}^{\Delta t} = Y_n^{\Delta t} + \Delta t\Big(\exp(-1)Y_n^{\Delta t} + \Delta t \sum_{j=n-k}^{n-1} Y_j^{\Delta t}\Big) + (\sigma_1 + \sigma_2 Y_n^{\Delta t})\ \Delta W_n \tag{7.4}$$

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 30,000 sample trajectories with each of the step-sizes $\Delta t = 2^{-3}, 2^{-4}, .., 2^{-8}$ and computed the error

$$\mu^{\Delta t}(T) = |\mathbf{E}Y_N^{\Delta t} - \mathbf{E}Y(T)| \tag{7.5}$$

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

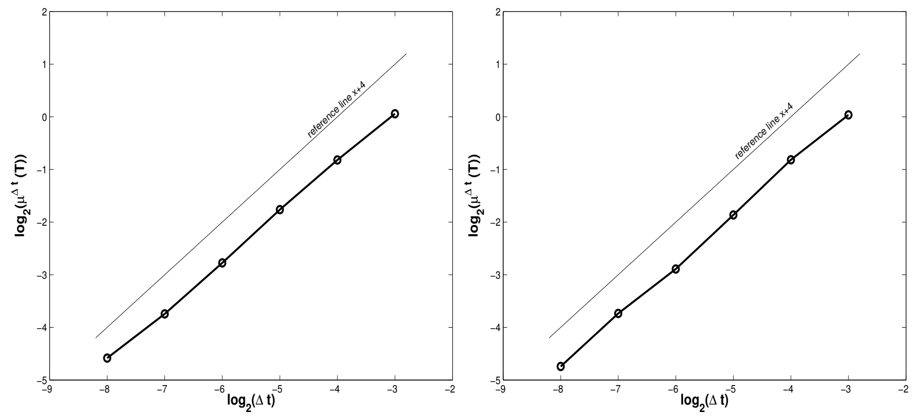

Figure 1: $\log_2(\mu^{\Delta t}(T))$ versus $\log_2(\Delta t)$ for (7.1) with left: $\sigma_1 = 0.2,\ \sigma_2 = 0$, right: $\sigma_1 = 0.0,\ \sigma_2 = 0.2$.

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 [14]). 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.

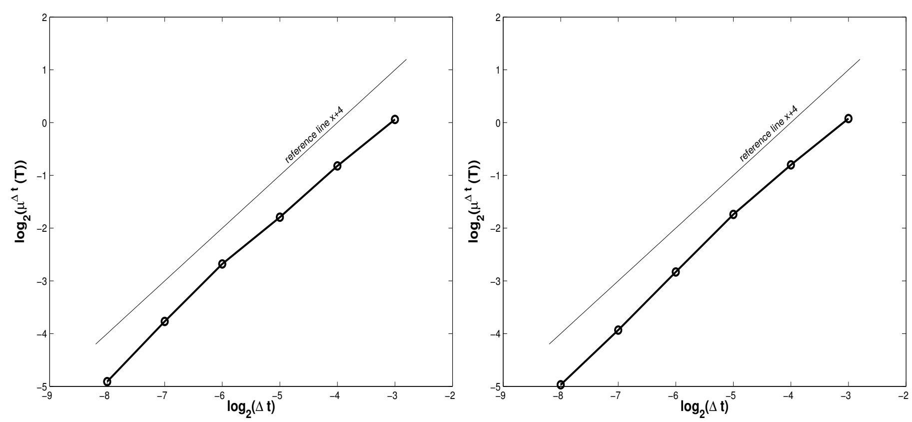

Figure 2: $\log_2(\mu^{\Delta t}(T))$ versus $\log_2(\Delta t)$ for (7.1) with left: $\sigma_1 = 0.2,\ \sigma_2 = 0$, right: $\sigma_1 = 0.0,\ \sigma_2 = 0.2$.

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].

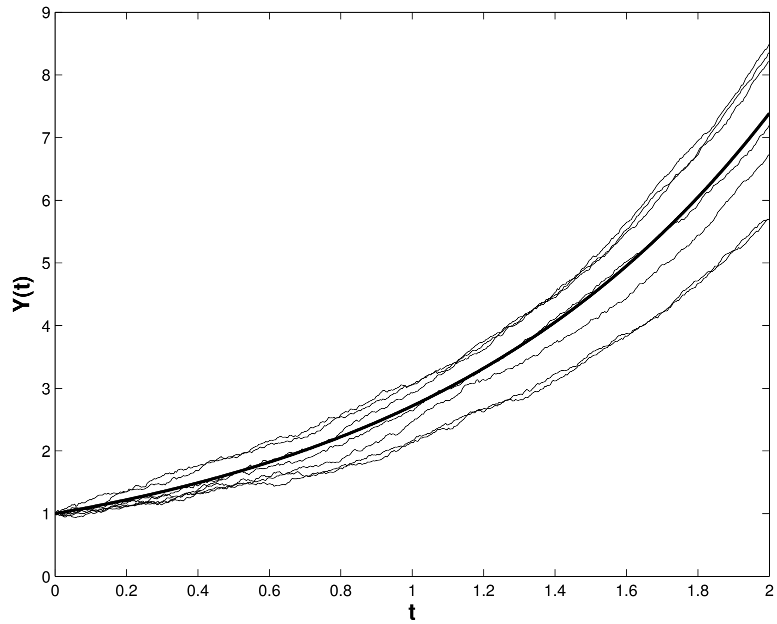

Figure 3: Trajectories of (7.1) with $\sigma_1 = 0.2,\ \sigma_2 = 0$.

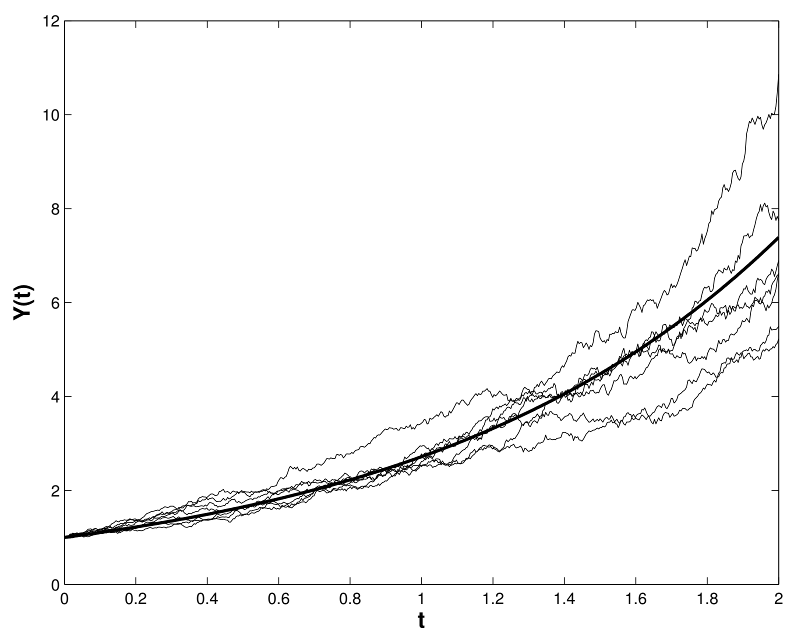

Figure 4: Trajectories of (7.1) with $\sigma_1 = 0.0,\ \sigma_2 = 0.2$.

8 Acknowledgements

We thank S. Mohammed, K. Ritter, A. Stuart, and J. Zabczyk for useful discussions, and also the IMA referees for many helpful comments.

References

[1] C. T. H. Baker and E. Buckwar, *Continuous Θ-methods for the stochastic pantograph equation.*, 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] 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), pp. 43–60.

[5] 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.

[6] M.-H. Chang and R. K. Youree, *The European option with hereditary price structures: Basic theory*, Applied Mathematics and Computation, 102 (1999), pp. 279–296.

[7] 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.

[8] 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.

[9] I. Elsanosi, B. Øksendal and A. Sulem, *Some solvable stochastic control problems with delay.*, Stochastics Stochastics Rep. , 71(1–2) (2000), pp. 69–89.

[10] 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.

[11] M. Gundlach and W. Kliemann, *Random Dynamical Systems, Volume in Ludwig Arnold's Honor*, Springer, 1998.

[12] D. G. Hobson and L. C. G. Rogers, *Complete models with stochastic volatility*, Mathematical Finance, 8(1) (1998), pp. 27–48.

[13] Y. Hu, S. E. A. Mohammed and F. Yan, *Discrete-time Approximations of Stochastic Delay Equations: The Milstein Scheme*, Accepted in *The Annals of Probability*, (Preprint URL: `http://salah.math.siu.edu/`).

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

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

[16] G. N. Milstein, *The Numerical Integration of Stochastic Differential Equations*, Kluwer, 1995.

[17] 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), pp. 2142–2167.

[18] 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), pp. 69–105.

[19] 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),pp. 1210–1240.

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

[21] T. Shardlow, *Weak convergence of a numerical method for a stochastic heat equation*, BIT 43(1), (2003).

[22] D. Talay, *Efficient numerical schemes for the approximation of expectations of functionals of the solution of a SDE and applications*, in Filtering and Control of Random Processes, J. Szpirglas, H. Korezlioglu, and G. Mazziotto, eds., vol. 61 of Lecture Notes in Control and Information Sciences, Springer–Verlag, 1984, pp. 294–313.

[23] D. Talay and L. Tubaro, *Expansion of the global error for numerical schemes solving stochastic differential equations*, Stochastic Anal. Appl., 8 (1990), pp. 483–509 (1991).

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

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

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

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

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