Multi-step Maruyama methods for stochastic delay differential equations*

Evelyn Buckwar and Renate Winkler†

Abstract

In this paper the numerical approximation of solutions of Itô stochastic delay differential equations is considered. We construct stochastic linear multi-step Maruyama methods and develop the fundamental numerical analysis concerning their L_p-consistency, numerical L_p-stability and L_p-convergence. For the special case of two-step Maruyama schemes we derive conditions guaranteeing their mean-square consistency.

1 Introduction

We consider n-dimensional systems of Itô stochastic delay differential equations with m driving Wiener processes and Q commensurate discrete lags of the form

$$\begin{aligned} X(s)\Big|_0^t &= \int_0^t F(s, X(s), X(s-\tau_2), \cdots, X(s-\tau_Q))\,\mathrm{d}s \qquad (1)\\ &\quad + \int_0^t G(s, X(s), X(s-\tau_2), \cdots, X(s-\tau_Q))\,\mathrm{d}W(s)\,, \quad \text{for } t \in [0,T],\\ X(t) &= \Psi(t) \quad \text{for } t \in J, \quad \text{where } J := [-\tau M^*, 0]. \qquad (2) \end{aligned}$$

Here the lags satisfy $\tau_q = \tau M_q$, $\tau \in \mathbb{R}$, $M_q \in \mathbb{N}$, $q = 2, \dots, Q$, the value M^* in J is defined as $\max(M_2, \dots, M_Q)$ and $M_1 = 0$ and thus $\tau_1 = 0$.
In many areas of science there has been an increasing interest in the investigation of **s**tochastic **d**elay **d**ifferential **e**quations (SDDEs), in particular, in the combined effects of noise and delay in dynamical systems. The use of delay equations in modelling deterministic dynamical systems in the natural sciences has a long history, see e. g., [27] for various examples. References concerning the application

*This work was supported by the DFG grant 234499 (first author) and the DFG Research Center *Mathematics for Key Technologies* in Berlin (second author)

†Humboldt-Universität zu Berlin, Institut für Mathematik, (buckwar/winkler@mathematik.hu-berlin.de)

of SDDEs in **biology** are e. g. to infectious diseases [5] and plankton populations [47], the recent review article [7] contains a section on stochastic models. In **biophysics** SDDEs are used to model for example delayed visual feedback systems [6, 29] and human postural sway [13, 41]. Several authors have studied stochastic oscillator ensembles with delayed coupling [17, 23, 26, 50]. These can be interpreted as mean-field models of coupled biological oscillators, such as groups of chorusing crickets, flashing fireflies and cardiac pacemaker cells. In **physics** often laser dynamics with delayed feedback are investigated [10, 15, 32], as well as the dynamics of noisy bi-stable systems with delay [33, 48]. One of the reduced conceptual stochastic differential equation models for a qualitative understanding of the El Niño phenomenon in climate dynamics relies on a delay effect caused by Kelvin wave reflection [3],[11]. In **engineering science** SDDEs arise e. g. as a problem in ship stability [25, 46] and often as control problems [18, 40]. In **financial mathematics** SDDEs appear in volatility models of stock markets [12, 21]. From a (stochastic) dynamical systems point of view, SDDEs display a range of interesting properties, such as multi-stability [26], noise induced bifurcations [28] and oscillations [1] and stochastic resonance [14, 31, 39]. For some general background on (deterministic and stochastic) delay differential equations we refer to e. g., [20, 30, 36].

This work consists of two parts. In the first part (Sections 2 and 3) we present a convergence theory for **s**tochastic **l**inear **m**ulti-step **M**aruyama **m**ethods (SLMMMs) applied to SDDEs. In [9] we considered general **s**tochastic **l**inear **m**ulti-step **m**ethods (SLMMs) applied to **s**tochastic **o**rdinary **d**ifferential **e**quations (SODEs). We analysed the mean-square convergence properties of these methods and their relation to mean-square consistency and mean-square zero-stability. In the current work we restrict the class of methods to SLMMMs only containing the Wiener process increments, but extend the results in two directions: first, we apply the SLMMMs to SDDEs. Mean-square convergence analysis for *one-step methods* applied to SDDEs has been carried out in [2, 8]. Second, we prove that the results also hold in the L_p-norms with $p \geq 1$. For one-step methods applied to SODEs the corresponding result has been obtained in [44].

In the second part (Section 4) we derive conditions for the mean-square consistency, i.e. we choose $p = 2$, for stochastic linear two-step Maruyama schemes applied to a simple scalar SDDE with a single lag and one driving Wiener process. These conditions allow to determine the parameters for the stochastic part from the parameters of the deterministic scheme and reduce to those of the underlying deterministic schemes when there is no noise. The main tool in these computations is the Itô-formula for SDDEs derived in [22], which we provide for the simple case considered. Due to the appearance of stochastic integrals with non-adapted integrands, the derivation of the Itô-formula relies on Malliavin calculus. The former is an unexpected feature of the analysis of SDDEs, which are intrinsically adapted processes. For the convenience of the reader we thus give a heuristic sketch of the derivation of the Itô-formula in Appendix B. The Itô-formula derived in [22] holds for the general case of SDDEs considered in Sections 2 and 3 and the consistency conditions for stochastic linear two-step Maruyama schemes are valid for the general case, too.

2 Definitions and preliminary results

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P})$ be a complete probability space with the filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$ satisfying the usual conditions (that is, it is increasing and right-continuous, and each $\{\mathcal{F}_t\}$, $t \in [0,T]$ contains all $\mathbb{P}$-null sets in $\mathcal{F}$). Let $W(t) = (W_1(t), \ldots, W_m(t))^T$ be an m-dimensional standard Wiener process on that probability space.
Throughout the article let $|.|$ denotes the Euclidean norm in $\mathbb{R}^n$, $\langle\cdot,\cdot\rangle$ its induced scalar product and $\|\cdot\|$ the corresponding induced matrix norm. The L_p-norm of a vector-valued L_p-integrable random variable $Z \in L_p(\Omega, \mathbb{R}^n)$ will be denoted by $\|Z\|_{L_p} := (\mathbb{E}\,|Z|^p)^{1/p}$, where $\mathbb{E}$ is expectation with respect to $\mathbb{P}$. The drift and diffusion function are given as $F : [0,T] \times (\mathbb{R}^n)^Q \to \mathbb{R}^n$ and $G = (G_1, \ldots, G_m) : [0,T] \times (\mathbb{R}^n)^Q \to \mathbb{R}^{n\times m}$, respectively. The initial path $\Psi(t) : J \to \mathbb{R}^n$ is assumed to be a continuous and $\mathcal{F}_0$-measurable random variable such that $(\mathbb{E}\,\sup_{s\in J} |\Psi(s)|^p)^{1/p} < \infty$. We assume that there exists a path-wise unique strong solution $X(\cdot)$ of (1).
We define a family of meshes on the interval $\mathcal{T} := [0,T]$ with a uniform step h with

$$hN_\tau = \tau, \quad hN \le T, \quad h(N+1) > T, \qquad N_\tau\,,\ N \in \mathbb{N}, \tag{3}$$

such that

$$\mathcal{T}_h^N := \{t_0 < t_1 < t_2, \cdots < t_N\} \subseteq \mathcal{T}, \quad t_\ell = \ell \cdot h,\ \ell = 0, \ldots, N\,. \tag{4}$$

We denote by $\mathcal{J}_h$ the correspondingly discretized initial interval where $\mathcal{J}_h \subseteq J$.
In our discussion of numerical methods we will denote by $Y(t_n)$ the approximation of the solution $X(t_n)$ of (1) at some point t_n in $\mathcal{T}_h^N$. Further, as an abbreviation, $\{Y_{t_n}\}$ will denote the evaluation of the lag terms $(Y(t_n), Y(t_n-\tau_2), \cdots, Y(t_n-\tau_Q))$ at $t_n \in \mathcal{T}_h^N$. We require given initial values $Y(t_0), \ldots, Y(t_{k-1}) \in L_p(\Omega, \mathbb{R}^n)$ such that $Y(t_\ell)$ is $\mathcal{F}_{t_\ell}$-measurable for $\ell = 0\ldots, k-1$. For simplicity the values on the initial interval will be taken as $Y(t_\ell) := \Psi(t_\ell)$ for $t_\ell \in \mathcal{J}_h$.
Thus we consider a stochastic linear k-step Maruyama method (setting $\alpha_0 = 1$), which for $\ell = k, \ldots, N$, takes the form

$$\sum_{j=0}^{k} \alpha_j\, Y(t_{\ell-j}) = h \sum_{j=0}^{k} \beta_j\, F(t_{\ell-j}, \{Y_{t_{\ell-j}}\}) + \sum_{j=1}^{k} \gamma_j\, G(t_{\ell-j}, \{Y_{t_{\ell-j}}\})\, I^{t_{\ell-j}, t_{\ell-j+1}}. \tag{5}$$

Here $I^{t,t+h} = (I_r^{t,t+h})_{r=1}^m$ denotes the vector of increments of the Wiener process, where $I_r^{t,t+h} = \int_t^{t+h}\, \mathrm{d}W_r(s) = W_r(t+h) - W_r(t)$.

We emphasize that an explicit discretization is used for the diffusion term. For $\beta_0 = 0$, the SLMMM (5) is explicit, otherwise it is drift-implicit.
We give an example of the two-step Maruyama methods (5).

Example 2.1 This is a stochastic variant of the implicit two-step BDF method, which we have termed BDF2-Maruyama method, applied to Equation (1). For $\ell = 2, \ldots, N$, it takes the form

$$\begin{aligned} Y(t_\ell) - \frac{4}{3}Y(t_{\ell-1}) + \frac{1}{3}Y(t_{\ell-2}) &= h\ \frac{2}{3}F(t_\ell, Y(t_\ell), Y(t_{\ell-M_2\cdot N_\tau}), \ldots, Y(t_{\ell-M_Q\cdot N_\tau})) \\ &+ \sum_{r=1}^{m} G_r(t_{\ell-1}, Y(t_{\ell-1}), Y(t_{\ell-1-M_2\cdot N_\tau}), \ldots, Y(t_{\ell-1-M_Q\cdot N_\tau}))\ I_r^{t_{\ell-1},t_\ell} \\ &- \frac{1}{3}\sum_{r=1}^{m} G_r(t_{\ell-2}, Y(t_{\ell-2}), Y(t_{\ell-2-M_2\cdot N_\tau}), \ldots, Y(t_{\ell-2-M_Q\cdot N_\tau}))\ I_r^{t_{\ell-2},t_{\ell-1}}. \end{aligned}$$

One has the parameters $\alpha_0 = 1,\ \alpha_1 = -\frac{4}{3},\ \alpha_2 = \frac{1}{3},\ \beta_0 = \frac{2}{3},\ \beta_1 = \beta_2 = 0,$ $\gamma_1 = 1,\ \gamma_2 = -\frac{1}{3}$.

We will consider L_p-convergence of SLMMMs in the sense discussed in Milstein [34, 35] ($p = 2$) and others [43].

Definition 1 *We call the SLMMM (5) for the approximation of the solution of the SDDE (1) L_p-**convergent** if the global error $X(t_\ell) - Y(t_\ell)$ satisfies*

$$\max_{\ell=1,\ldots,N} \|X(t_\ell) - Y(t_\ell)\|_{L_p} \to 0 \text{ as } h \to 0,$$

*we say it is L_p-**convergent with order** γ ($\gamma > 0$) if the global error satisfies*

$$\max_{\ell=1,\ldots,N} \|X(t_\ell) - Y(t_\ell)\|_{L_p} \le C \cdot h^\gamma, \tag{6}$$

with a constant $C > 0$ which is independent of the step-size h, but may depend on the length of the interval $[0, T]$ and the initial data. In addition to (6) we also consider the stronger condition

$$(\mathbb{E} \max_{\ell=1,\ldots,N} |X(t_\ell) - Y(t_\ell)|^p\,)^{\frac{1}{p}} \le C \cdot h^\gamma. \tag{7}$$

In the following we will define what we understand by local errors and refer to [9] for a discussion of this concept in the area of numerics for SODEs.

Definition 2 *We define the* **local error** *of the SLMMM (5) for the approximation of the solution X of the SDDE (1), for $\ell = k, \ldots, N$, as*

$$\begin{aligned} L_\ell := \sum_{j=0}^{k} \alpha_j X(t_{\ell-j}) &- h\sum_{j=0}^{k} \beta_j\, F(t_{\ell-j}, X(t_{\ell-j}), X(t_{\ell-j} - \tau_2), \cdots, X(t_{\ell-j} - \tau_Q)) \\ - \sum_{j=1}^{k} \gamma_j\ \sum_{r=1}^{m} G_r(t_{\ell-j}, X(t_{\ell-j}), X(t_{\ell-j} - \tau_2), \cdots, X(t_{\ell-j} - \tau_Q))\ I_r^{t_{\ell-j},t_{\ell-j+1}}&. \end{aligned} \tag{8}$$

We aim to conclude L_p-convergence from local properties of the SLMMM by means of numerical stability in the L_p sense. Numerical stability concerns the influence of perturbations of the right-hand side of the discrete scheme on the global solution of that discrete scheme. Sources of perturbations may be the local error, round-off errors or defects in the approximate solution of implicit schemes. The L_p stability estimate of the global error is based on the L_p-norm **and** on the conditional p-th mean of the perturbations. In the case of one-step schemes and $p = 2$ this appears e.g. in [2, 49], we refer in particular to the discussion in [34, Chapter 1.4] or [35, Chapter 1.1.4]. We remark that in the case of k-step schemes the conditional p-th mean has to be taken with respect to the σ-algebra $\mathcal{F}_{t_{\ell-k}}$.

In the subsequent analysis we also consider the following discrete system, the perturbed form of (5), for $\ell = k, \ldots, N$

$$\sum_{j=0}^{k} \alpha_j \, \widetilde{Y}(t_{\ell-j}) \\ = h \sum_{j=0}^{k} \beta_j \, F(t_{\ell-j}, \{\widetilde{Y}_{t_{\ell-j}}\}) + \sum_{j=1}^{k} \gamma_j \sum_{r=1}^{m} G_r(t_{\ell-j}, \{\widetilde{Y}_{t_{\ell-j}}\}) \, I_r^{t_{\ell-j}, t_{\ell-j+1}} + D_\ell, \quad (9)$$

with values $\widetilde{Y}(t_\ell) = Y(t_\ell) + D_\ell$ for $\ell = 0, \ldots, k-1$. We suppose that the perturbations D_ℓ are $\mathcal{F}_{t_\ell}$-measurable and that $D_\ell \in L_p(\Omega, \mathbb{R}^n)$ for $\ell = 0, \ldots, k-1$. On the interval J we allow perturbations of the initial function such that $\widetilde{\Psi}(s) = \Psi(s) + D(s),\ s \in J$, where $\sup_{s \in J} \|D(s)\|_{L_p} < \infty$. Slightly abusing notation, we write D_ℓ for $D(t_\ell),\ t_\ell \in \mathcal{J}_h$.

Remark 2.2 *It is useful to represent the perturbations in the form*

$$\begin{array}{c} D_\ell = R_\ell + S_\ell =: R_\ell + \sum\limits_{j=1}^{k} S_{j,\ell-j+1}, \quad \ell = k, \ldots, N, \\ \textit{where each } \ S_{j,\ell} \ \textit{ is } \ \mathcal{F}_{t_\ell} \ \textit{ measurable with } \ \mathbb{E}(S_{j,\ell} | \mathcal{F}_{t_\ell}) = 0 \ . \end{array} \quad (10)$$

This repesentation is not unique. A particularly useful one is given by

$$\begin{array}{rcl} R_\ell^* &=& \mathbb{E}(D_\ell | \mathcal{F}_{t_{\ell-k}}), \qquad S_\ell^* = D_\ell - R_\ell^*, \\ S_{j,\ell-j+1}^* &=& \mathbb{E}(D_\ell - R_\ell^* - \sum_{i=j+1}^{k} S_{i,\ell-i+1}^* | \mathcal{F}_{t_{\ell-j+1}}), \ j = k, k-1, \ldots, 1 \ . \end{array} \quad (11)$$

For further discussions of these representations we refer to Remark 2.5 in [9].

Now we give the precise definition of L_p-stability and consistency that we consider in this paper.

Definition 3 *We call the SLMMM (5)* ***numerically*** L_p***-stable*** *if there exist constants $h_0 > 0$ and $S > 0$ such that for all step-sizes $h < h_0$ and for all $\mathcal{F}_{t_\ell}$*

measurable perturbations $D_\ell \in L_p(\Omega,\mathbb{R}^n)$ ($\ell = 0,\dots,N$) and all their representations (10), the following inequality holds

$$\max_{\ell=0,\dots,N} \|Y(t_\ell) - \widetilde{Y}(t_\ell)\|_{L_p}$$
$$\le S_p\Big\{\sup_{s\in J}\|D(s)\|_{L_p} + \max_{\ell=0,\dots,k-1}\|D_\ell\|_{L_p} + \max_{\ell=k,\dots,N}\big(\frac{\|R_\ell\|_{L_p}}{h} + \frac{\|S_\ell\|_{L_p}}{h^{1/2}}\big)\Big\}, \quad (12)$$

where $(Y(t_\ell))_{\ell=1}^N$ and $(\widetilde{Y}(t_\ell))_{\ell=1}^N$ are the solutions of the SLMMM (5) and the perturbed discrete system (9), respectively. In addition to (12), we consider the stronger inequality

$$(\mathbb{E}\max_{\ell=1,\dots,N}|Y(t_\ell) - \widetilde{Y}(t_\ell)|^p)^{\frac{1}{p}}$$
$$\le S_p\Big\{\sup_{s\in J}\|D(s)\|_{L_p} + \max_{\ell=0,\dots,k-1}\|D_\ell\|_{L_p} + \frac{(\mathbb{E}\max_{\ell=k,\dots,N}|R_\ell|^p)^{\frac{1}{p}}}{h} + \max_{\ell=k,\dots,N}\frac{\|S_\ell\|_{L_p}}{h^{1/2}}\Big\}, \quad (13)$$

We refer to S_p as the stability constant and to (12) and (13) as stability inequalities.

Definition 4 *We call the SLMMM (5) for the approximation of the solution of the SDDE (1) L_p-**consistent** if the local error L_ℓ satisfies*

$$h^{-1}\,\|\mathbb{E}(L_\ell|\mathcal{F}_{t_{\ell-k}})\|_{L_p} \to 0 \text{ for } h\to 0, \text{ and } h^{-1/2}\,\|L_\ell\|_{L_p} \to 0 \text{ for } h \to 0.$$

*We call the SLMMM (5) for the approximation of the solution of the SDDE (1) L_p-**consistent** of order γ ($\gamma > 0$), if the local error L_ℓ satisfies*

$$\|\mathbb{E}(L_\ell|\mathcal{F}_{t_{\ell-k}})\|_{L_p} \le \bar{c}\cdot h^{\gamma+1}, \text{ and } \quad \|L_\ell\|_{L_p} \le c\cdot h^{\gamma+\frac{1}{2}}, \quad \ell = 1,\dots,N,$$

with constants c, $\bar{c} > 0$ only depending on the SDDE and its solution.

We remind the reader that consistency is only concerned with the local error. In the case that we disregard other sources of errors in (9) we only have to deal with perturbations $D_\ell = L_\ell$.

We adapt the following Lemma from [9, Lemma 2.8]. Its proof consists of obvious modifications of the arguments in the proof of [9, Lemma 2.8].

Lemma 2.3 *The SLMMM (5) is L_p-consistent of order γ, if*

$$\|R_\ell\|_{L_p} \le \bar{c}\cdot h^{\gamma+1}, \text{ and } \quad \|S_\ell\|_{L_p} \le c\cdot h^{\gamma+\frac{1}{2}}, \quad \ell = 1,\dots,N,$$

for any representation (10) of the local error $D_\ell = L_\ell$. The SLMMM (5) is L_p-consistent of order γ, if and only if

$$\|R^*_\ell\|_{L_p} \le \bar{c}\cdot h^{\gamma+1}, \text{ and } \quad \|S^*_\ell\|_{L_p} \le c\cdot h^{\gamma+\frac{1}{2}}, \quad \ell = 1,\dots,N,$$

where the representation (11) is chosen for the local error $D_\ell = L_\ell$.

For further reference we state the following definitions and results.

Definition 5 *A function $F : [0,T] \times (\mathbb{R}^n)^Q \to \mathbb{R}^n$ satisfies a* **uniform Lipschitz condition** *with respect to ξ_q, $q = 1\ldots,Q$, if there exists a positive constant L_F, such that for all $t \in [0,T]$ and all ξ_q, $\eta_q \in \mathbb{R}^n$, $q = 1\ldots,Q$*

$$|F(t,\xi_1,\xi_2,\cdots,\xi_Q) - F(t,\eta_1,\eta_2,\cdots,\eta_Q)| \le L_F \sum_{q=1}^{Q} |\xi_q - \eta_q|. \tag{14}$$

A function $G : [0,T] \times (\mathbb{R}^n)^Q \to \mathbb{R}^{n\times m}$ satisfies a **uniform Lipschitz condition** *with respect to ξ_q, $q = 1\ldots,Q$, if there exists a positive constant L_G, such that for all $t \in [0,T]$ and all ξ_q, $\eta_q \in \mathbb{R}^n$, $q = 1\ldots,Q$*

$$\|G(t,\xi_1,\xi_2,\cdots,\xi_Q) - G(t,\eta_1,\eta_2,\cdots,\eta_Q)\| \le L_G \sum_{q=1}^{Q} |\xi_q - \eta_q|. \tag{15}$$

Let $C^{s-1,s}$ denote the class of all functions from $[0,T] \times (\mathbb{R}^n)^Q$ to $\mathbb{R}^n$ having continuous partial derivatives up to order $s-1$ with respect to the first variable and continuous partial derivatives up to order s with respect to the other Q variables. Let C^K denote the class of functions y from $[0,T] \times (\mathbb{R}^n)^Q$ to $\mathbb{R}^n$ that satisfy a **linear growth condition** *in the form*

$$|y(t,x_1,\ldots,x_Q)| \le K(1 + \sum_{j=1}^{Q} |x_j|^2)^{\frac{1}{2}},\ \forall t \in [0,T],\ \forall x_j \in \mathbb{R}^n, j = 1,\ldots,Q. \tag{16}$$

Definition 6 *The* **characteristic polynomial** *of (5) is given by*

$$\rho(\zeta) = \alpha_k \zeta^k + \alpha_{k-1}\zeta^{k-1} + \ldots + \alpha_0. \tag{17}$$

The SLMM (5) is said to fulfil **Dahlquist's root condition**, *if* i) *the roots of $\rho(\zeta)$ lie on or within the unit circle, and* ii) *the roots on the unit circle are simple.*

Lemma 2.4 *(A discrete version of Gronwall's lemma) Let a_ℓ, $\ell = 1,\ldots,N$, and C_1, C_2 be nonnegative real numbers and assume that the inequalities*

$$a_\ell \le C_1 + C_2 \frac{1}{N} \sum_{i=1}^{\ell-1} a_i, \quad \ell = 1,\ldots,N,$$

are valid. Then we have $\max\limits_{\ell=1,\ldots,N} a_\ell \le C_1 \exp(C_2)$.

3 Global properties of stochastic LMMs

In this section we will first establish the solvability of the recurrence equations (9) (and thus of (5)), then we will discuss numerical stability and L_p-convergence of the SLMMM (5). For a discussion of numerical stability in the deterministic and stochastic setting we refer to [9].
We now turn to the solvability of the recurrence equations. It is obvious that every iterate $Y(t_\ell)$, $\widetilde{Y}(t_\ell)$, $\ell \geq k$ of the recurrence equations (5) and (9), respectively, can be obtained explicitly, if the right-hand sides of (5) and (9) do not depend on $Y(t_\ell)$, $\widetilde{Y}(t_\ell)$. This happens if either the parameter $\beta_0 = 0$ or if the memory term in (1) is a pure delay term and does not itself depend on the current time instance. In these cases the recurrence equations (5) and (9) have unique solutions. In the case of implicit systems we need to consider the solvability of the systems of nonlinear equations (5) and (9). In addition, we have to verify that the L_p-norm of the iterates exists. (The straightforward extension to fully implicit systems would serve as an example were the L_p-norm of the iterates does not exist.)

Theorem 3.1 *Suppose that $\beta_0 \neq 0$ and the drift-coefficient F satisfies (14) and assume that $2\ h\ \beta_0\ L_F < 1$. Then the perturbed discrete scheme (9) and, in consequence, the SLMMM (5) have a unique solution. If, in addition the coefficient G satisfies (15), then the L_p-norm of the iterates exists.*

Proof: The proof of the theorem follows the line of proofs used in the deterministic analysis of numerical schemes for DDEs [4] and stochastic linear multi-step schemes [9].

We now formulate our main theorem on numerical stability.

Theorem 3.2 *The stochastic linear multi-step Maruyama method (5) is numerically L_p-stable for every continuous F and G satisfying (14) and (15), respectively, if and only if its characteristic polynomial $\rho(\zeta)$ (17) satisfies Dahlquist's root condition given in Definition 6.*

We postpone the proof to Appendix A.

With the powerful notion of numerical L_p-stability, together with L_p-consistency the L_p-convergence follows almost immediately.

Theorem 3.3 *An L_p-consistent SLMMM (5) for the approximation of the solution of SDDE (1) is L_p-convergent for all continuous F and G satisfying (14) and (15), respectively, if and only if it is numerically L_p-stable. If, in addition, it is L_p-consistent with order $\gamma > 0$, then the SLMMM (5) is L_p-convergent with order γ.*

The proof follows with obvious modifications the proof of Theorem 3.3 in [9].

4 Two-step-Maruyama schemes for scalar SDDEs

In this section we consider the scalar case of Equation (1) with a single lag τ, i.e. $Q = 2$, $M_1 = 0$, $M_2 = 1$ and a single driving Wiener process, i.e. Equation (1) becomes for $t \in [0, T]$

$$X(s)\Big|_0^t = \int_0^t F(s, X(s), X(s-\tau))\,\mathrm{d}s + \int_0^t G_1(s, X(s), X(s-\tau))\,\mathrm{d}W_1(s)\,. \quad (18)$$

The reason for using this basic equation is that we would like to keep the notation as simple as possible. In order to derive consistency conditions we will apply an Itô-formula to F and G_1, where the former is much more complex in the delayed case. We refer to [22] for the development of the Itô-formula in the delayed case and for several useful results. The methods we investigate are stochastic linear two-step-Maruyama schemes. Thus we have for $\ell = 2, \ldots, N$

$$\begin{aligned}\sum_{j=0}^{2} \alpha_j Y(t_{\ell-j}) = h \sum_{j=0}^{2} \beta_j\ F(t_{\ell-j}, Y(t_{\ell-j}), Y(t_{\ell-j} - \tau)) \\ + \sum_{j=1}^{2} \gamma_j\ G_1(t_{\ell-j}, Y(t_{\ell-j}), Y(t_{\ell-j} - \tau))\ I_1^{t_{\ell-j}, t_{\ell-j+1}}. \quad (19)\end{aligned}$$

For sufficiently smooth drift and diffusion coefficients F and G_1 Theorem 3.2 applies and, choosing $p = 2$, the two-step scheme (19) is mean square stable if the coefficients $\alpha_0,\ \alpha_1,\ \alpha_2$ satisfy Dahlquist's root condition. Then the scheme (19) is mean-square convergent of some order γ, if it is mean-square consistent of that order. Thus we will be concerned with mean-square consistency of the above scheme and derive order conditions in terms of the coefficients $\alpha_0, \alpha_1, \alpha_2, \beta_0, \beta_1, \beta_2, \gamma_1, \gamma_2$. In general, the mean-square order of convergence will be not higher than $\frac{1}{2}$, since the only information about the driving noise process that the Maruyama-type schemes include are the Wiener increments. We note that the simple Euler-Maruyama method would suffice to obtain the same order of convergence. However, convergence is an asymptotic property, i.e. it holds for $h \to 0$ and a result concerning the order of convergence may not provide sufficient information about the size of the actual error that arises for reasonable choices of the step-size. In particular when one considers equations with a small noise term as in [9], one may find that the influence of the noise is not dominant and properties of the methods in the deterministic setting are recovered to some extent.

In the first part of this section we present the Itô-formula for functions of the form of F and G_1. An informal sketch of its derivation can be found in Appendix B. In the second part we derive consistency conditions for the two-step scheme (19) applied to the SDDE (18). We establish a representation of the local error L_ℓ in terms of certain multiple stochastic integrals obtained by the Itô-Taylor expansion. It turns out that consistency is guaranteed under the usual conditions for deterministic order 1 and additional conditions that determine the stochastic method parameters γ_1 and γ_2.

4.1 An Itô-formula

For a function $\phi(t,x,z)$ in $C^{1,2,2}$ we denote with ϕ_t, ϕ_x, ϕ_z, ϕ_{xx}, etc. first and second order derivatives with respect to t, x, z. Further we abbreviate the evaluation of a function $\phi(s,X(s),X(s-\tau))$ at s by $\phi\big|_s$. The Itô formula for a function $\phi(t,x,z)$ in $C^{1,2,2}$ and the solution X of (18) for $a,b \in [0,T]$, $b>a$, reads

$$
\begin{aligned}
&\phi(b,X(b),X(b-\tau)) = \phi(a,X(a),X(a-\tau)) \\
&+ \int_a^b \phi_t\big|_s + \phi_x\big|_s \cdot F\big|_s + \frac{1}{2}\,\phi_{xx}\big|_s \cdot G_1^2\big|_s \,\mathrm{d}s + \int_a^b \phi_x\big|_s \cdot G_{!}\big|_s \,\mathrm{d}W_1(s) && (20)\\
&+ \frac{1}{2}\int_{a-\tau}^{b-\tau} \phi_{zz}\big|_s \cdot G_1^2\big|_s \,\mathrm{d}s + \Big\{ \int_{a-\tau}^{b-\tau} \phi_z\big|_{s+\tau} \cdot F\big|_s \,\mathrm{d}s && (21)\\
&+ \boxed{\int_{a-\tau}^{b-\tau} \phi_z\big|_{s+\tau} \cdot G_1\big|_s \,\delta W_1(s)} + \int_{a-\tau}^{b-\tau} \phi_{zx}\big|_{s+\tau} \cdot G_1^2\big|_s \,\mathrm{d}s && (22)\\
&+ \int_{a-\tau}^{b-\tau} \phi_{zx}\big|_{s+\tau}\, G_1\big|_s \Big(\int_0^{s+\tau} D_s F\big|_r \,\mathrm{d}r + \int_0^{s+\tau} D_s G_1\big|_r \,\mathrm{d}W_1(r)\Big) \,\mathrm{d}s && (23)\\
&+ \int_{a-\tau}^{b-\tau} \phi_{zz}\big|_{s+\tau}\, G_1\big|_s \Big(\int_0^{s} D_s F\big|_r \,\mathrm{d}r + \int_0^{s} D_s G_1\big|_r \,\mathrm{d}W_1(r)\Big) \,\mathrm{d}s \Big\}. && (24)
\end{aligned}
$$

The terms appearing in (20) are those present in the Itô-formula for an SODE. If the lower bound $a-\tau$ of some of the integrals is less than 0 then it is to be replaced by 0. In the second order terms D_s denotes the Malliavin derivative (see e.g. [37]). The framed integral in (22) is a Skorokhod integral due to the term $X(s+\tau)$ in the integrand. All terms in curly brackets are derived from the framed sum in (56) in Appendix B.
In the heuristic exposition in Appendix B we indicate the main steps in the derivation of the Itô-formula above. The mathematically rigorous derivation can be found in the fundamental work [22].
We introduce operators Λ_0, Λ_1, Λ_0^τ, Λ_S and Λ_M, applied to a function $\phi(t,x,z)$

by

$$\Lambda_0\phi = \phi_t + \phi_x F + \frac{1}{2}\,\phi_{xx} G^2 \,, \qquad \Lambda_1\phi = \phi_x G \,, \tag{25}$$

$$\Lambda_0^\tau \phi\big|_s = \phi_z\big|_{s+\tau}\cdot F\big|_s \;+\; \frac{1}{2}\,\phi_{zz}\big|_s\cdot G^2\big|_s \;+\; \phi_{zx}\big|_{s+\tau}\cdot G^2\big|_s \,, \tag{26}$$

$$\Lambda_S\phi\big|_s = \phi_z\big|_{s+\tau}\cdot G\big|_s \,, \tag{27}$$

$$\begin{aligned}\Lambda_M\phi\big|_s &= \phi_{zx}\big|_{s+\tau}\, G\big|_s \Big(\int_0^{s+\tau} D_s F\big|_r \,\mathrm{d}r + \int_0^{s+\tau} D_s G\big|_r \,\mathrm{d}W(r)\Big) \\ &\quad +\phi_{zz}\big|_{s+\tau}\, G\big|_s \Big(\int_0^{s} D_s F\big|_r \,\mathrm{d}r + \int_0^{s} D_s G\big|_r \,\mathrm{d}W(r)\Big) \,.\end{aligned} \tag{28}$$

We denote multiple Wiener integrals and the Skorokhod integral by

$$I_{r_1,r_2,\dots,r_j}^{a,b;\tau_1,\dots,\tau_j}(\phi) = \int_{a-\tau_1}^{b-\tau_1}\int_{a-\tau_2}^{s_1-\tau_2}\dots\int_{a-\tau_j}^{s_{j-1}-\tau_j}\phi\big|_{s_j}\,\mathrm{d}W_{r_1}(s_j)\dots\,\mathrm{d}W_{r_j}(s_1), \tag{29}$$

$$S_1^{a,b;\tau}(\phi) = \int_{a-\tau}^{b-\tau}\phi\big|_s\,\delta W_1(s)\,, \tag{30}$$

where $r_i \in \{0,1,\dots,m\}$ and $\mathrm{d}W_0(s) = \mathrm{d}s$. If $\phi \equiv 1$ we omit the argument (ϕ). Using the operators and the notation introduced above we can write the Itô-formula as

$$\begin{aligned}&\phi(b,X(b),X(b-\tau)) = \phi(a,X(a),X(a-\tau)) \\ &+\; I_0^{a,b;0}(\Lambda_0\phi) + I_1^{a,b;0}(\Lambda_1\phi) + I_0^{a,b;\tau}(\Lambda_0^\tau\phi) + S_1^{a,b;\tau}(\Lambda_S\phi) + I_0^{a,b;\tau}(\Lambda_M\phi)\,.\end{aligned} \tag{31}$$

Applying the Itô-formula (31) with ϕ taken as the drift coefficient F and the diffusion coefficient G_1 and $s \in [t_{\ell-j}, t_{\ell-j+1}]$, $j = 1,2$, yields

$$\begin{aligned}&F(s,X(s),X(s-\tau)) = F(t_{\ell-j},X(t_{\ell-j}),X(t_{\ell-j}-\tau)) + I_0^{t_{\ell-j},s;0}(\Lambda_0\,F) \\ &+\; I_1^{t_{\ell-j},s;0}(\Lambda_1\,F) + I_0^{t_{\ell-j},s;\tau}(\Lambda_0^\tau F) + S_1^{t_{\ell-j},s;\tau}(\Lambda_S F) + I_0^{t_{\ell-j},s;\tau}(\Lambda_M F),\end{aligned} \tag{32}$$

$$\begin{aligned}&G_1(s,X(s),X(s-\tau)) = G_1(t_{\ell-j},X(t_{\ell-j}),X(t_{\ell-j}-\tau)) + I_0^{t_{\ell-j},s;0}(\Lambda_0 G_1) \\ &+\; I_1^{t_{\ell-j},s;0}(\Lambda_1 G_1) + I_0^{t_{\ell-j},s;\tau}(\Lambda_0^\tau G_1) + S_1^{t_{\ell-j},s;\tau}(\Lambda_S G_1) + I_0^{t_{\ell-j},s;\tau}(\Lambda_M G_1).\end{aligned} \tag{33}$$

Over each interval $[t_{\ell-2}, t_\ell]$ we evaluate the drift function F in (19) three times, at $t_{\ell-2}$, $t_{\ell-1}$ and t_ℓ. We now trace back the values of the drift coefficient at $t_{\ell-1}$

and t_ℓ to the point $t_{\ell-2}$ and obtain

$$F(t_{\ell-1},X(t_{\ell-1}),X(t_{\ell-1}-\tau)) = F(t_{\ell-2},X(t_{\ell-2}),X(t_{\ell-2}-\tau)) + I_0^{t_{\ell-2},t_{\ell-1};0}(\Lambda_0 F) \\ +I_1^{t_{\ell-2},t_{\ell-1};0}(\Lambda_1 F) + I_0^{t_{\ell-2},t_{\ell-1};\tau}(\Lambda_0^\tau F) + S_1^{t_{\ell-2},t_{\ell-1};\tau}(\Lambda_S F) + I_0^{t_{\ell-2},t_{\ell-1};\tau}(\Lambda_M F), \quad (34)$$

$$\begin{aligned} F(t_\ell,X(t_\ell),X(t_\ell-\tau)) &= F(t_{\ell-2},X(t_{\ell-2}),X(t_{\ell-2}-\tau)) + I_0^{t_{\ell-2},t_{\ell-1};0}(\Lambda_0 F) \\ &+ I_1^{t_{\ell-2},t_{\ell-1};0}(\Lambda_1 F) + I_0^{t_{\ell-2},t_{\ell-1};\tau}(\Lambda_0^\tau F) + S_1^{t_{\ell-2},t_{\ell-1};\tau}(\Lambda_S F) + I_0^{t_{\ell-2},t_{\ell-1};\tau}(\Lambda_M F) \\ &+ I_0^{t_{\ell-1},t_\ell;0}(\Lambda_0 F) + I_1^{t_{\ell-1},t_\ell;0}(\Lambda_1 F) + I_0^{t_{\ell-1},t_\ell;\tau}(\Lambda_0^\tau F) + S_1^{t_{\ell-1},t_\ell;\tau}(\Lambda_S F) \\ &+ I_0^{t_{\ell-1},t_\ell;\tau}(\Lambda_M F). \end{aligned} \quad (35)$$

Further, the SDDE (18) implies the identities

$$\begin{aligned} X(t_{\ell-1}) - X(t_{\ell-2}) &= \int_{t_{\ell-2}}^{t_{\ell-1}} F(s,X(s),X(s-\tau))\,\mathrm{d}s + \int_{t_{\ell-2}}^{t_{\ell-1}} G_1(s,X(s),X(s-\tau))\,\mathrm{d}W_1(s) \\ &= h\,F(t_{\ell-2},X(t_{\ell-2}),X(t_{\ell-2}-\tau)) + I_{00}^{t_{\ell-2},t_{\ell-1};0,0}(\Lambda_0 F) + I_{10}^{t_{\ell-2}t_{\ell-1};0,0}(\Lambda_1 F) \\ &+ I_{00}^{t_{\ell-2},t_{\ell-1};\tau,0}(\Lambda_0^\tau F) + I_0^{t_{\ell-2},t_{\ell-1};0}(S_1^{t_{\ell-2},s;\tau}(\Lambda_S F)) + I_{00}^{t_{\ell-2},t_{\ell-1};\tau,0}(\Lambda_M F) \\ &+ G_1(t_{\ell-2},X(t_{\ell-2}),X(t_{\ell-2}-\tau))I_1^{t_{\ell-2},t_{\ell-1};0} + I_{01}^{t_{\ell-2},t_{\ell-1};0,0}(\Lambda_0 G_1) + I_{11}^{t_{\ell-2}t_{\ell-1};0,0}(\Lambda_1 G_1) \\ &+ I_{01}^{t_{\ell-2},t_{\ell-1};\tau,0}(\Lambda_0^\tau G_1) + I_1^{t_{\ell-2},t_{\ell-1};0}(S_1^{t_{\ell-2},s;\tau}(\Lambda_S G_1)) + I_{01}^{t_{\ell-2},t_{\ell-1};\tau,0}(\Lambda_M G_1)\,, \end{aligned} \quad (36)$$

and, using additionally (34),

$$\begin{aligned} X(t_\ell) - X(t_{\ell-1}) &= \int_{t_{\ell-1}}^{t_\ell} F(s,X(s),X(s-\tau))\,\mathrm{d}s + \int_{t_{\ell-1}}^{t_\ell} G_1(s,X(s),X(s-\tau))\,\mathrm{d}W_1(s) \\ &= h\,\big\{F(t_{\ell-2},X(t_{\ell-2}),X(t_{\ell-2}-\tau)) + I_0^{t_{\ell-2},t_{\ell-1};0}(\Lambda_0 F) + I_1^{t_{\ell-2},t_{\ell-1};0}(\Lambda_1 F) \\ &+ I_0^{t_{\ell-2},t_{\ell-1};\tau}(\Lambda_0^\tau F) + S_1^{t_{\ell-2},t_{\ell-1};\tau}(\Lambda_S F) + I_0^{t_{\ell-2},t_{\ell-1};\tau}(\Lambda_M F)\big\} \\ &+ I_{00}^{t_{\ell-1},t_\ell;0,0}(\Lambda_0 F) + I_{10}^{t_{\ell-1}t_\ell;0,0}(\Lambda_1 F) \\ &+ I_{00}^{t_{\ell-1},t_\ell;\tau,0}(\Lambda_0^\tau F) + I_0^{t_{\ell-1},t_\ell;0}(S_1^{t_{\ell-1},s;\tau}(\Lambda_S F)) + I_{00}^{t_{\ell-1},t_\ell;\tau,0}(\Lambda_M F) \\ &+ G_1(t_{\ell-1},X(t_{\ell-1}),X(t_{\ell-1}-\tau))I_1^{t_{\ell-1},t_\ell;0} + I_{01}^{t_{\ell-1},t_\ell;0,0}(\Lambda_0 G_1) + I_{11}^{t_{\ell-1},t_\ell;0,0}(\Lambda_1 G_1) \\ &+ I_{01}^{t_{\ell-1},t_\ell;\tau,0}(\Lambda_0^\tau G_1) + I_1^{t_{\ell-1},t_\ell;0}(S_1^{t_{\ell 1},s;\tau}(\Lambda_S G_1)) + I_{01}^{t_{\ell-1},t_\ell;\tau,0}(\Lambda_M G_1)\,. \end{aligned} \quad (37)$$

4.2 Consistency conditions

To analyse the local error L_ℓ of the scheme (19) for the SDDE (18) and to achieve a suitable representation (10) we want to derive appropriate Itô-Taylor expansions, where we take special care to separate the multiple stochastic integrals over the different subintervals of integration. For the SDDE (18) we have the following result.

Lemma 4.1 *Assume that the coefficients F and G_1 of the SDDE (18) belong to the class $C^{1,2,2}$ with $\Lambda_0 F, \Lambda_0 G_1, \Lambda_1 F, \Lambda_1 G_1, \Lambda_0^\tau F, \Lambda_0^\tau G_1, \Lambda_S F, \Lambda_S G_1, \Lambda_M F$ and $\Lambda_M G \in C^K$. Then the local error (8) of the stochastic 2-step scheme (19) allows the representation*

$$L_\ell = R_\ell^\circ + S_{1,\ell}^\circ + S_{2,\ell-1}^\circ, \quad \ell = 2, \ldots, N, \tag{38}$$

where $R_\ell^\circ, S_{j,\ell}^\circ, j = 1, 2$ are $\mathcal{F}_{t_\ell}$-measurable with $\mathbb{E}(S_{j,\ell}^\circ | \mathcal{F}_{t_{\ell-1}}) = 0$ and

$$\begin{aligned}
R_\ell^\circ &= \Big[\sum_{j=0}^{2} \alpha_j\Big] X(t_{\ell-2}) + \Big[2\alpha_0 + \alpha_1 - \sum_{j=0}^{2} \beta_j\Big] hF(t_{\ell-2}, X(t_{\ell-2}), X(t_{\ell-2} - \tau)) + \widetilde{R}_\ell^\circ, \\
S_{1,\ell}^\circ &= \Big[\alpha_0 - \gamma_1\Big] G_1(t_{\ell-1}, X(t_{\ell-1}), X(t_{\ell-1} - \tau)) I_1^{t_{\ell-1}, t_\ell} + \widetilde{S}_{1,\ell}^\circ, \\
S_{2,\ell-1}^\circ &= \Big[(\alpha_0 + \alpha_1) - \gamma_2\Big] G_1(t_{\ell-2}, X(t_{\ell-2}), X(t_{\ell-2} - \tau)) I_1^{t_{\ell-2}, t_{\ell-1}} + \widetilde{S}_{2,\ell-1}^\circ
\end{aligned}$$

with

$$\|\widetilde{R}_\ell^\circ\|_{L_2} = O(h^2), \quad \|\widetilde{S}_{1,\ell}^\circ\|_{L_2} = O(h), \quad \|\widetilde{S}_{2,\ell-1}^\circ\|_{L_2} = O(h). \tag{39}$$

Corollary 4.2 *Let the coefficients F and G_1 of the SDDE (18) satisfy the assumptions of Lemma 4.1 and suppose they are Lipschitz continuous in the sense of (14) and (15), respectively. Let the coefficients of the stochastic linear two-step Maruyama scheme (19) satisfy Dahlquist's root condition and the consistency conditions*

$$\sum_{j=0}^{2} \alpha_j = 0, \quad 2\alpha_0 + \alpha_1 = \sum_{j=0}^{2} \beta_j, \quad \alpha_0 = \gamma_1, \quad \alpha_0 + \alpha_1 = \gamma_2. \tag{40}$$

Then the global error of the scheme (19) applied to (18) allows the expansion

$$\max_{\ell=2,N} \|X(t_\ell) - Y(t_\ell)\|_{L_2} = \mathcal{O}(h^{1/2}) + \mathcal{O}(\max_{\ell=0,1} \|X(t_\ell) - Y(t_\ell)\|_{L_2}) \; .$$

Proof: *(of Corollary 4.2)* By Lemma 4.1 we have the representation (38) for the local error. Applying the consistency conditions (40) yields

$$R_\ell^\circ = \widetilde{R}_\ell^\circ, \quad S_{1,\ell}^\circ = \widetilde{S}_{1,\ell}^\circ, \quad S_{2,\ell-1}^\circ = \widetilde{S}_{2,\ell-1}^\circ, \quad \ell = 2, \ldots, N.$$

As the scheme (19) satisfies Dahlquist's root condition, it is numerically stable in the mean-square sense. Now the assertion follows from the estimates (39) by means of the stability inequality (12).

Proof: *(of Lemma 4.1)* To derive a representation of the local error in the form (38) we evaluate and resume the deterministic parts at the point $(X(t_{\ell-2}), t_{\ell-2})$ and separate the stochastic terms carefully over the different subintervals $[t_{\ell-2}, t_{\ell-1}]$ and $[t_{\ell-1}, t_\ell]$. This ensures the independence of the random variables. It does make the calculations more messy, though. By rewriting

$$\sum_{j=0}^{2} \alpha_j X(t_{\ell-j}) = \alpha_0\big(X(t_\ell)-X(t_{\ell-1})\big)+(\alpha_0+\alpha_1)\big(X(t_{\ell-1})-X(t_{\ell-2})\big)+\big(\sum_{j=0}^{2}\alpha_j\big)X(t_{\ell-2}),$$

we can express the local error (8) as

$$\begin{aligned} L_\ell \;=\;& \alpha_0\big(X(t_\ell) - X(t_{\ell-1})\big) + (\alpha_0+\alpha_1)\big(X(t_{\ell-1}) - X(t_{\ell-2})\big) + \sum_{j=0}^{2}\alpha_j X(t_{\ell-2}) \\ & - h\sum_{j=0}^{2}\beta_j\; F(t_{\ell-j}, X(t_{\ell-j}), X(t_{\ell-j}-\tau)) \\ & - \sum_{j=1}^{2}\gamma_j\; G_1(t_{\ell-j}, X(t_{\ell-j}), X(t_{\ell-j}-\tau))\; I_1^{t_{\ell-j},t_{\ell-j+1}}. \end{aligned}$$

Inserting the expansions (36), (37), (34) and (35) into the local error formula and reordering the terms, yields

$$\begin{aligned} L_\ell \;=\;& \Big[\sum_{j=0}^{2}\alpha_j\Big]X(t_{\ell-2}) + \Big[2\alpha_0+\alpha_1-\sum_{j=0}^{2}\beta_j\Big]h\,F(t_{\ell-2}, X(t_{\ell-2}), X(t_{\ell-2}-\tau)) + \widetilde{R}^\circ_\ell \\ & +\Big[\alpha_0-\gamma_1\Big]G_1(t_{\ell-1}, X(t_{\ell-1}), X(t_{\ell-1}-\tau))\, I_1^{t_{\ell-1},t_\ell;0} + \widetilde{S}^\circ_{1,\ell} \\ & +\Big[(\alpha_0+\alpha_1)-\gamma_2\Big]G_1(t_{\ell-2}, X(t_{\ell-2}), X(t_{\ell-2}-\tau))\, I_1^{t_{\ell-2},t_{\ell-1};0} + \widetilde{S}^\circ_{2,\ell-1}, \end{aligned}$$

where

$$\begin{aligned} \widetilde{R}^\circ_\ell \;=\;& \alpha_0\big\{hI_0^{t_{\ell-2},t_{\ell-1};0}(\Lambda_0 F) + I_{00}^{t_{\ell-1},t_\ell;0}(\Lambda_0 F) + h\,I_0^{t_{\ell-2},t_{\ell-1};\tau}(\Lambda_0^\tau F) \\ & +I_{00}^{t_{\ell-1},t_\ell;\tau,0}(\Lambda_0^\tau F) + h\,I_0^{t_{\ell-2},t_{\ell-1};\tau}(\Lambda_M F) + I_{00}^{t_{\ell-1},t_\ell;\tau,0}(\Lambda_M F)\big\} \\ +\;& (\alpha_0+\alpha_1)\big\{I_{00}^{t_{\ell-2},t_{\ell-1};0}(\Lambda_0 F) + I_{00}^{t_{\ell-2},t_{\ell-1};\tau,0}(\Lambda_0^\tau F) + I_{00}^{t_{\ell-2},t_{\ell-1};\tau,0}(\Lambda_M F)\big\} \\ -\;& h\,\beta_0\big\{I_0^{t_{\ell-2},t_{\ell-1};0}(\Lambda_0 F) + I_0^{t_{\ell-1},t_\ell;0}(\Lambda_0 F) + I_0^{t_{\ell-2},t_{\ell-1};\tau}(\Lambda_0^\tau F) \\ & +I_0^{t_{\ell-1},t_\ell;\tau}(\Lambda_0^\tau F) + I_0^{t_{\ell-2},t_{\ell-1};\tau}(\Lambda_M F) + I_0^{t_{\ell-1},t_\ell;\tau}(\Lambda_M F)\big\} \\ -\;& h\,\beta_1\big\{I_0^{t_{\ell-2},t_{\ell-1};0}(\Lambda_0 F) + I_0^{t_{\ell-2},t_{\ell-1};\tau}(\Lambda_0^\tau F) + I_0^{t_{\ell-2},t_{\ell-1};\tau}(\Lambda_M F)\big\}, \end{aligned} \tag{41}$$

$$\begin{aligned}
\widetilde{S}^{\circ}_{1,\ell} &= \alpha_0 \left\{ I_{10}^{t_{\ell-1},t_\ell;0,0}(\Lambda_1 F) + I_0^{t_{\ell-1},t_\ell;0}(S_1^{t_{\ell-1},s;\tau}(\Lambda_S F)) \right. \\
&+ I_{01}^{t_{\ell-1},t_\ell;0,0}(\Lambda_0 G_1) + \boxed{I_{11}^{t_{\ell-1},t_\ell;0,0}(\Lambda_1 G_1)} \\
&+ \left. I_{01}^{t_{\ell-1},t_\ell;\tau,0}(\Lambda_0^\tau G_1) + \boxed{I_1^{t_{\ell-1},t_\ell;0}(S_1^{t_{\ell-1},s;\tau}(\Lambda_S G_1))} + I_{01}^{t_{\ell-1},t_\ell;\tau,0}(\Lambda_M G) \right\} \\
&- h\,\beta_0 \left\{ I_1^{t_{\ell-1},t_\ell;0}(\Lambda_1 F) + S_1^{t_{\ell-1},t_\ell;\tau}(\Lambda_S F) \right\}, \qquad (42)
\end{aligned}$$

$$\begin{aligned}
\widetilde{S}^{\circ}_{2,\ell-1} &= h(\alpha_0 - \beta_0 - \beta_1) I_1^{t_{\ell-2},t_{\ell-1};0}(\Lambda_1 F) + h\,\alpha_0\, S_1^{t_{\ell-2},t_{\ell-1};\tau}(\Lambda_S F) \\
&+ (\alpha_0 + \alpha_1)\left\{ I_{10}^{t_{\ell-2},t_{\ell-1};0,0}(\Lambda_1 F) + I_0^{t_{\ell-2},t_{\ell-1};0}(S_1^{t_{\ell-2},s;\tau}(\Lambda_S F)) \right. \\
&+ I_{01}^{t_{\ell-2},t_{\ell-1};0,0}(\Lambda_0 G_1) + \boxed{I_{11}^{t_{\ell-2},t_{\ell-1};0,0}(\Lambda_1 G_1)} \\
&+ \left. I_{01}^{t_{\ell-2},t_{\ell-1};\tau,0}(\Lambda_0^\tau G_1) + \boxed{I_1^{t_{\ell-2},t_{\ell-1};0}(S_1^{t_{\ell-2},s;\tau}(\Lambda_S G_1))} + I_{01}^{t_{\ell-2},t_{\ell-1};\tau,0}(\Lambda_M G_1) \right\} \\
&- h\,(\beta_0 + \beta_1)\, S_1^{t_{\ell-2},t_{\ell-1};\tau}(\Lambda_S F). \qquad (43)
\end{aligned}$$

The estimates (39) are derived by means of Lemmata 2.1 and 2.2 in [34], applied to the integrals containing the operators Λ_0 and Λ_1. The main ingredients in the proofs of those Lemmata are the Hölder inequality and the Itô isometry. With obvious modifications these Lemmata can be adapted to estimate the integrals containing the operator Λ_0^τ. Combining these modifications and similar computations as in Lemma 5.1 in [22] yield the estimates for the integrals containing the operator Λ_M. In particular one obtains bounds on the integrand by applying the chain rule for weak derivatives ([37, Prop. 1.2.2]) and Propositions 3.1 and 3.3 in [22], which provide estimates on weak derivatives of the solution of (18). For the estimates of the Skorokhod integrals we proceed similarly, additionally using the formula for the covariance between two Skorokhod integrals ([37, Section 1.3.1]). The framed terms in (42) and (43) determine the order $\mathcal{O}(h)$.

Remark 4.3 *Corollary 4.2 holds, in particular, for the Θ-Maruyama method, which is of the form (19) with $k = 1$, $\alpha_0 = 1$, $\alpha_1 = -1$, $\beta_0 = \Theta$, $\beta_1 = 1 - \Theta$ and $\gamma_1 = \alpha_0 = 1$.*

A Proof of Theorem 3.2

Proof: Necessity: This part can be proved as in the deterministic case, i. e. we take the equation $X'(t) = 0$, then F and G satisfy obviously (14) and (15). We then follow in principle the proof of [16, Thm.6.3.3].

Sufficiency: Since the SLMM (5) contains the stochastic part related to the diffusion G, we can not rely on the theory of difference equations and the representations of their solutions. Instead, we will follow the route of rewriting the

k-step recurrence equation as a one-step recurrence equation in a higher dimensional space (see e.g. [19, Chap.III.4],[42, Chap.8.2.1]). The proof follows the proofs of Theorem 3.2 in [9] and Theorem 3.1 in [44].

For $Y(t_\ell)$ and $\widetilde{Y}(t_\ell)$ being the solutions of (5) and (9), respectively, let the n-dimensional vector E_ℓ be defined as the difference $Y(t_\ell) - \widetilde{Y}(t_\ell)$. We have with $E_0, \ldots, E_{k-1} \in L_p(\Omega, \mathbb{R}^n)$ for $\ell = k, \ldots, N$, the recursion

$$E_\ell = -\sum_{j=1}^{k} \alpha_j \; E_{\ell-j} + h \; \underbrace{\sum_{j=0}^{k} \beta_j \; \Delta F_{\ell-j}}_{=:\Delta\phi^\ell} + \underbrace{\sum_{j=1}^{k} \gamma_j \, \Delta G_{\ell-j} \, I^{t_{\ell-j}, t_{\ell-j+1}}}_{=:\Delta\psi^\ell} - D_\ell, \tag{44}$$

where

$$\begin{aligned} \Delta F_{\ell-j} &:= F(t_{\ell-j}, \{Y_{t_{\ell-j}}\}) - F(t_{\ell-j}, \{\widetilde{Y}_{t_{\ell-j}}\})\,, \\ \Delta G_{\ell-j} &:= G(t_{\ell-j}, \{Y_{t_{\ell-j}}\}) - G(t_{\ell-j}, \{\widetilde{Y}_{t_{\ell-j}}\}). \end{aligned}$$

We rearrange this k-step recursion in the space $L_p(\Omega, \mathbb{R}^n)$ to a one-step recursion in $L_p(\Omega, \mathbb{R}^{k\times n})$. Together with the trivial identities $E_{\ell-1} = E_{\ell-1}, \ \ldots \ E_{\ell-k+1} = E_{\ell-k+1}$ and using the n-dimensional unit matrix I_n and the $k \times k$-matrix A with

$$A = \begin{pmatrix} -\alpha_1 & \cdots\cdots & -\alpha_k \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix},$$

we obtain

$$\underbrace{\begin{pmatrix} E_\ell \\ E_{\ell-1} \\ \vdots \\ E_{\ell-k+1} \end{pmatrix}}_{=:\,\mathcal{E}_\ell} = \underbrace{\begin{pmatrix} -\alpha_1 I_n & \cdots\cdots & -\alpha_k I_n \\ I_n & 0 & & \\ & \ddots & \ddots & \\ & & I_n & 0 \end{pmatrix}}_{=:\,\mathcal{A} = A \otimes I_n} \underbrace{\begin{pmatrix} E_{\ell-1} \\ E_{\ell-2} \\ \vdots \\ E_{\ell-k} \end{pmatrix}}_{=:\mathcal{E}_{\ell-1}} + h \underbrace{\begin{pmatrix} \Delta\phi^\ell \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{=:\,\Delta\Phi_\ell} + \underbrace{\begin{pmatrix} \Delta\psi^\ell \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{=:\,\Delta\Psi_\ell} + \underbrace{\begin{pmatrix} -D_\ell \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{=:\,\mathcal{D}_\ell}$$

or, in compact form

$$\mathcal{E}_\ell = \mathcal{A}\mathcal{E}_{\ell-1} + h\,\Delta\Phi_\ell + \Delta\Psi_\ell + \mathcal{D}_\ell\;,\;\; \ell = k, \ldots, N$$

and $\mathcal{E}_{k-1} = (-D_{k-1}, -D_{k-2}, \ldots, -D_0)^T$, where $\mathcal{E}_\ell \in L_p(\Omega, \mathbb{R}^{k\times n}),\ \ell = k-1, \ldots, N$. The vector $\mathcal{E}_{k-1}$ consists of the perturbations to the initial values. We now trace back the recursion in $\mathcal{E}_\ell$ to the initial vector $\mathcal{E}_{k-1}$. For $\ell = k, \ldots, N$ we have

$$\begin{aligned}
\mathcal{E}_\ell &= \mathcal{A}\,\mathcal{E}_{\ell-1} + h\,\Delta\Phi_\ell + \Delta\Psi_\ell + \mathcal{D}_\ell \\
&= \mathcal{A}(\mathcal{A}\,\mathcal{E}_{\ell-2} + h\,\Delta\Phi_{\ell-1} + \Delta\Psi_{\ell-1} + \mathcal{D}_{\ell-1}) + h\,\Delta\Phi_\ell + \Delta\Psi_\ell + \mathcal{D}_\ell \\
&= \mathcal{A}^2\,\mathcal{E}_{\ell-2} + h\,(\Delta\Phi_\ell + \mathcal{A}\Delta\Phi_{\ell-1}) + (\Delta\Psi_\ell + \mathcal{A}\Delta\Psi_{\ell-1}) + (\mathcal{D}_\ell + \mathcal{A}\mathcal{D}_{\ell-1}) \\
&\ \ \vdots \\
&= \mathcal{A}^{\ell-k+1}\,\mathcal{E}_{k-1} + h\,\sum_{i=0}^{\ell-k}\mathcal{A}^i\,\Delta\Phi_{\ell-i} + \sum_{i=0}^{\ell-k}\mathcal{A}^i\,\Delta\Psi_{\ell-i} + \sum_{i=0}^{\ell-k}\mathcal{A}^i\,\mathcal{D}_{\ell-i} \\
&= \mathcal{A}^{\ell-k+1}\,\mathcal{E}_{k-1} + h\,\sum_{i=k}^{\ell}\mathcal{A}^{\ell-i}\,\Delta\Phi_i + \sum_{i=k}^{\ell}\mathcal{A}^{\ell-i}\,\Delta\Psi_i + \sum_{i=k}^{\ell}\mathcal{A}^{\ell-i}\,\mathcal{D}_i\,.
\end{aligned}$$

Due to the assumption that Dahlquist's root condition is satisfied and standard arguments (see e.g. [19, Chap.III.4,Lemma 4.4] or [24, Lemmas B.3,B.10] there exists a vector norm $|.|_*$ on $\mathbb{R}^{k\times n}$ with a subordinate matrix norm $\|.\|_*$ such that in this norm we have $\|\mathcal{A}\|_* = \|A\otimes I_n\|_* \le 1$. Further, due to the norm equivalence on finite-dimensional spaces there are constants $c_1, c_2 > 0$ such that

$$c_1|\mathcal{X}|_p^p \le |\mathcal{X}|_*^p \le c_2|\mathcal{X}|_p^p \qquad \forall \mathcal{X} \in \mathbb{R}^{k\times n}\,,$$

where $|\mathcal{X}|_p^p = \sum_{j=1}^k |x_j|^p$, for $\mathcal{X} = (x_1^T, \ldots, x_k^T)^T$ and $|.|$ denotes the Euclidean norm on $\mathbb{R}^n$. In the case of a vector of the form $\mathcal{X} = (x^T, 0, \ldots, 0)^T$ the expression $|\mathcal{X}|_p^p$ obviously reduces to $|x|^p$.

In the remainder of the proof all constants are properly chosen to fulfill the inequalities.

We now apply $|\boldsymbol{.}|_*^p$ to estimate $|\mathcal{E}_\ell|_*^p$ and, later, $\max\limits_{\nu=k,\ldots,\ell} \mathbb{E}\,(|\mathcal{E}_\nu|_*^p)$ and $\mathbb{E}\,(\max\limits_{\nu=k,\ldots,\ell} |\mathcal{E}_\nu|_*^p)$. We start with

$$|\mathcal{E}_\ell|_*^p \le 4^{p-1}\Big\{ \underbrace{|\mathcal{A}^{\ell-k+1}\mathcal{E}_{k-1}|_*^p}_{1)} + \underbrace{|\sum_{i=k}^{\ell}\mathcal{A}^{\ell-i}\Delta\Phi_i|_*^p}_{2)} + \underbrace{|\sum_{i=k}^{\ell}\mathcal{A}^{\ell-i}\Delta\Psi_i|_*^p}_{3)} + \underbrace{|\sum_{i=k}^{\ell}\mathcal{A}^{\ell-i}\mathcal{D}_i|_*^p}_{4)} \Big\}. \tag{45}$$

For the term labelled 1) we have for $\ell = k, \ldots, N$

$$|\mathcal{A}^{\ell-k+1}\mathcal{E}_{k-1}|_*^p \;\le\; |\mathcal{E}_{k-1}|_*^p\,. \tag{46}$$

For the term labelled 2) we have

$$
\begin{aligned}
h^p \Big|\sum_{i=k}^{\ell} \mathcal{A}^{\ell-i} \Delta\Phi_i\Big|_*^p &\le h^p(\ell-k+1)^{p-1} \sum_{i=k}^{\ell} |\mathcal{A}^{\ell-i}\Delta\Phi_i|_*^p \ \le\ h^p N^{p-1} \sum_{i=k}^{\ell} |\Delta\Phi_i|_*^p \\
&\le\ \frac{T^p}{N} c_2 \sum_{i=k}^{\ell} |\Delta\phi^i|^p \ \le\ \frac{T^p}{N}\, c_2\, (k+1)^{p-1} \sum_{i=k}^{\ell}\sum_{j=0}^{k} |\beta_j|^p\, |\Delta F_{i-j}|^p \\
&\le\ \frac{T^p}{N}\, c_2\, (k+1)^{p-1} \sum_{i=k}^{\ell}\sum_{j=0}^{k} |\beta_j|^p\, \big(L_F \sum_{q=1}^{Q} |E_{i-j-M_q\cdot N_\tau}|\big)^p \\
&\le\ \frac{T^p}{N}\, c_2\, (k+1)^{p-1}\, L_F^p\, Q^{p-1} \sum_{i=k}^{\ell}\sum_{j=0}^{k} |\beta_j|^p \ \sum_{q=1}^{Q} |E_{i-j-M_q\cdot N_\tau}|^p\,.
\end{aligned}
$$

To deal with the implicit term $|\beta_0|^p|E_\ell|^p$ we need to take it out of the above triple sum. Obviously this term is not present for explicit methods. The implicit term occurs exactly once for $i=\ell$ in the outer sum and $j=0$ in the middle sum and $q=1$ in the inner sum. When $i=\ell$, we can shift the indices, such that all other terms (other than $|\beta_0|^p|E_\ell|^p$) can be subsumed into terms with $i<\ell$. Further, with $M^*=\max(M_2,\ldots,M_Q)$, for every index $i=k,\ldots,\ell-1$ in the outer sum there can occur terms $|E_j|^p$ with the index j between i–k–M^*N_τ and i . However, every $|E_i|^p$ with $i<\ell$ occurs only a finite number of times, at maximum say $M_{k,Q}$ times. The indices emphasise that this number only depends on k and Q, not on the step-size h. With $c_\beta := \max_{j=0,\ldots,k} |\beta_j|^p$ and $C_1 := T^p\, c_2\, (k+1)^{p-1}\, L_F^p\, Q^{p-1}$ we then have

$$
\begin{aligned}
h^p\, \Big|\sum_{i=k}^{\ell} \mathcal{A}^{\ell-i}\Delta\Phi_i\Big|_*^p &\le \frac{1}{N}\, C_1 \Big\{ |\beta_0|^p\, |E_\ell|^p + c_\beta\, M_{k,Q} \sum_{i=-M^*N_\tau}^{\ell-1} |E_i|^p \Big\} \\
&=\ \frac{1}{N}\, C_1 \Big\{ |\beta_0|^p\, |E_\ell|^p + c_\beta\, M_{k,Q} \left(\sum_{i=-M^*N_\tau}^{k-1} |D_i|^p + \sum_{i=k}^{\ell-1} |E_i|^p \right) \Big\} \\
&\le\ \frac{1}{N}\, C_1 \Big\{ \frac{|\beta_0|^p}{c_1}\, |\mathcal{E}_\ell|_*^p + c_\beta\, M_{k,Q} \left(\sum_{i=-M^*N_\tau}^{k-1} |D_i|^p + \frac{1}{c_1} \sum_{i=k}^{\ell-1} |\mathcal{E}_i|_*^p \right) \Big\}. \qquad (47)
\end{aligned}
$$

After inserting this and the estimate (46) into inequality (45) we subtract the term containing $|\mathcal{E}_\ell|_*^p$ from both sides. If necessary we choose a bound h^0 on the step-size such that for all $h<h^0$, or rather $\frac{1}{N}<h^0$, we have

$$
\frac{1}{N}\, 4^{p-1}\, C_1 \frac{|\beta_0|^p}{c_1} \le \frac{1}{2}\,.
$$

Multiplying by 2 we obtain a new inequality for $|\mathcal{E}_\ell|_*^p$:

$$
\begin{aligned}
|\mathcal{E}_\ell|_*^p \;\le\; & 2\cdot 4^{p-1}\Big\{|\mathcal{E}_{k-1}|_*^p + \frac{1}{N}\, C_1 c_\beta\, M_{k,Q}\left(\sum_{i=-M^*N_\tau}^{k-1}|D_i|^p + \frac{1}{c_1}\sum_{i=k}^{\ell-1}|\mathcal{E}_i|_*^p\right) \\
& + \underbrace{|\sum_{i=k}^{\ell}\mathcal{A}^{\ell-i}\Delta\Psi_i|_*^p}_{3)} + \underbrace{|\sum_{i=k}^{\ell}\mathcal{A}^{\ell-i}\mathcal{D}_i|_*^p}_{4)}\Big\}. \qquad (48)
\end{aligned}
$$

We will now treat the term labelled 3). For that purpose we introduce the notation $\Delta\Psi_{j,i} := ((\gamma_j\,\Delta G_i\, I^{t_i,t_{i+1}})^T, 0, \ldots, 0)^T$. Using this we can write

$$
\Delta\Psi_i = ((\sum_{j=1}^{k}\gamma_j\,\Delta G_{i-j}\; I^{t_{i-j},t_{i-j+1}})^T, 0, \ldots, 0)^T = \sum_{j=1}^{k}\Delta\Psi_{j,i-j}\,.
$$

We now reorder the last term above such that we have a sum of terms where each term contains all multiple Wiener integrals over just one subinterval. For this, we subsume under $M_i,\ i = 0, \ldots, \ell-1$ all terms of the form $\mathcal{A}^\nu\Delta\Psi_{j,i}$. Then

$$
\sum_{i=k}^{\ell}\mathcal{A}^{\ell-i}\Delta\Psi_i = \sum_{\nu=0}^{\ell-1}M_\nu\,, \qquad (49)
$$

and every M_i consists of maximally k terms, e.g. for $i = \ell-k$ one has

$$
M_{\ell-k} = \mathcal{A}^0\Delta\Psi_{k,\ell-k} + \mathcal{A}^1\Delta\Psi_{k-1,\ell-k} + \ldots + \mathcal{A}^{k-1}\Delta\Psi_{1,\ell-k}
$$

For $i > \ell-k$ and $i < k-1$ there are even less terms. For the convenience of notation we set $\Delta\Psi_{j,i} =: 0$ for the terms that do not occur. Using Hölder's inequality and $\|\mathcal{A}\|_* = 1$ we can estimate

$$
|M_i|_*^p \;\le\; k^{p-1}\sum_{j=1}^{k}|\Delta\Psi_{j,i}|_*^p \;\le\; k^{p-1}c_2\sum_{j=1}^{k}|\Delta\Psi_{j,i}|_p^p\,.
$$

Every $\Delta\Psi_{j,i-j}$ is $\mathcal{F}_{t_{i-j+1}}$-measurable and $\mathbb{E}(\Delta\Psi_{j,i-j}|\mathcal{F}_{t_{i-j}}) = 0$. Thus every M_i is $\mathcal{F}_{t_{i+1}}$-measurable and $\mathbb{E}(M_i|\mathcal{F}_{t_i}) = 0$, for $i = 0, \ldots, \ell-1$. We now observe that the discrete parameter process $\{\mathcal{M}_\ell := \sum_{\nu=0}^{\ell-1}M_\nu, \mathcal{F}_{t_\ell}\}_{\ell=0}^{N}$ is a martingale and has finite p-th order moments. The Burkholder-Davis-Gundy inequality (cf. e.g. [45, VII.3] or [30, I.1.7]) then yields the following estimate

$$
\mathbb{E}\,(\max_{i=0,\ldots,\ell-1}|\sum_{\nu=0}^{i}M_\nu|_*^p) \;\le\; B_p\,\mathbb{E}\,(\sum_{i=0}^{\ell-1}|M_i|_*^2)^{\frac{p}{2}} \;\le\; B_p\,\ell^{\frac{p}{2}-1}\sum_{i=0}^{\ell-1}\mathbb{E}\,|M_i|_*^p
$$

for $\ell = 1, \ldots, N$ and with some universal constant B_p. We thus obtain

$$
\begin{aligned}
&\mathbb{E}\,(\max_{\mu=k,\dots,\ell} |\sum_{i=k}^{\mu} \mathcal{A}^{\ell-i}\Delta\Psi_i|_*^p) = \mathbb{E}\,(\max_{\mu=k,\dots,\ell} |\sum_{\nu=0}^{\mu-1} M_\nu|_*^p) \;\leq\; \mathbb{E}\,(\max_{i=0,\dots,\ell-1} |\sum_{\nu=0}^{i} M_\nu|_*^p) \\
&\leq\;\; B_p\,\ell^{\frac{p}{2}-1} \sum_{i=0}^{\ell-1} \mathbb{E}\,|M_i|_*^p \;\leq\; B_p\,\ell^{\frac{p}{2}-1}\,k^{p-1}\,c_2 \sum_{i=0}^{\ell-1}\sum_{j=1}^{k} \mathbb{E}\,|\Delta\Psi_{j,i}|_p^p \\
&=\;\; B_p\,\ell^{\frac{p}{2}-1}\,k^{p-1}\,c_2 \sum_{i=0}^{\ell-1}\sum_{j=1}^{k} \mathbb{E}\,|\gamma_j\,\Delta G_i\; I^{t_i,t_{i+1}}|^p \\
&=\;\; B_p\,\ell^{\frac{p}{2}-1}\,k^{p-1}\,c_2\,c_\gamma \sum_{i=0}^{\ell-1} \mathbb{E}\,|\Delta G_i\; I^{t_i,t_{i+1}}|^p \qquad\qquad (\text{with } c_\gamma := \sum_{j=i}^{k} |\gamma_j|^p) \\
&\leq\;\; B_p\,\ell^{\frac{p}{2}-1}\,k^{p-1}\,c_2\,c_\gamma \sum_{i=0}^{\ell-1} \mathbb{E}\,\|\Delta G_i\|^p\;\mathbb{E}|I^{t_i,t_{i+1}}|^p \\
&\leq B_p\,\ell^{\frac{p}{2}-1}\,k^{p-1}\,c_2\,c_\gamma L_G^p \sum_{i=0}^{\ell-1} \mathbb{E}\left(\sum_{q=1}^{Q} |Y(t_{i-M_q\cdot N_\tau}) - \widetilde{Y}(t_{i-M_q\cdot N_\tau})|\right)^p \;\mathbb{E}\,|I^{t_i,t_{i+1}}|^p \\
&\leq B_p\,\ell^{\frac{p}{2}-1}\,k^{p-1}\,c_2\,c_\gamma\,L_G^p Q^{p-1} \sum_{i=0}^{\ell-1}\sum_{q=1}^{Q} \mathbb{E}|Y(t_{i-M_q\cdot N_\tau}) - \widetilde{Y}(t_{i-M_q\cdot N_\tau})|^p\;\mathbb{E}\,|I^{t_i,t_{i+1}}|^p \\
&\leq B_p\,\ell^{\frac{p}{2}-1}\,k^{p-1}\,c_2\,c_\gamma\,L_G^p\,Q^{p-1} m^{\frac{p}{2}} (\frac{p(p-1)}{2})^{\frac{p}{2}} h^{\frac{p}{2}} \sum_{i=0}^{\ell-1}\sum_{q=1}^{Q} \mathbb{E}|E_{i-M_q\cdot N_\tau}|^p \\
&:=\;\; (\ell\,h)^{\frac{p}{2}-1} h\,C_2 \sum_{i=0}^{\ell-1}\sum_{q=1}^{Q} \mathbb{E}\,|E_{i-M_q\cdot N_\tau}|^p \\
&\leq\;\; T^{\frac{p}{2}-1}\,\frac{T}{N} C_2 \sum_{i=0}^{\ell-1}\sum_{q=1}^{Q} \mathbb{E}\,|E_{i-M_q\cdot N_\tau}|^p \;:=\; \frac{1}{N} C_3 \sum_{i=0}^{\ell-1}\sum_{q=1}^{Q} \mathbb{E}\,|E_{i-M_q\cdot N_\tau}|^p\,.
\end{aligned}
$$

Again, with $M^* = \max(M_2,\dots,M_Q)$, for every index $i = 0,\dots,\ell-1$ there can occur terms $\mathbb{E}|E_j|^p$ with the index j between $i - M^*\cdot N_\tau$ and i. However, every $\mathbb{E}|E_i|^p$ with $i < \ell$ occurs only a finite number of times, at maximum say again $M_{k,Q}$ times. We then have

$$
\begin{aligned}
&\mathbb{E}\,(\max_{\nu=k,\dots,\ell} |\sum_{i=k}^{\nu} \mathcal{A}^{\ell-i}\Delta\Psi_i|_*^p) \leq \frac{1}{N}\,C_3\,M_{k,Q} \sum_{i=-M^*\cdot N_\tau}^{\ell-1} \mathbb{E}\,|E_i|^p \\
&\leq\;\; \frac{1}{N}\,C_3\,M_{k,Q} \left(\sum_{i=-M^*\cdot N_\tau}^{k-1} \mathbb{E}\,|D_i|^p + \frac{1}{c_1}\sum_{i=k}^{\ell-1} \mathbb{E}\,|\mathcal{E}_i|_*^p \right)\,. \qquad (50)
\end{aligned}
$$

Applying $\mathbb{E}\max_{\nu=k,\dots,\ell}$ to both sides of (48) and inserting (50) we obtain

$$\mathbb{E}\max_{\nu=k,\dots,\ell}|\mathcal{E}_\nu|_*^p \le 2\cdot 4^{p-1}\Big\{\mathbb{E}|\mathcal{E}_{k-1}|_*^p \tag{51}$$
$$+\frac{1}{N}(C_1c_\beta+C_3)M_{k,Q}\left(\sum_{i=-M^*N_\tau}^{k-1}\mathbb{E}|D_i|^p+\frac{1}{c_1}\sum_{i=k}^{\ell-1}\mathbb{E}|\mathcal{E}_i|_*^p\right)+\mathbb{E}\max_{\nu=k,\dots,\ell}|\sum_{i=k}^{\nu}\mathcal{A}^{\nu-i}\mathcal{D}_i|_*^p\Big\}.$$

We have $|\mathcal{E}_{k-1}|_*^p \le \frac{1}{c_1}|\mathcal{E}_{k-1}|_p^p = \frac{1}{c_1}\sum\limits_{i=0}^{k-1}|D_i|^p$. Thus (51) becomes

$$\begin{aligned}
\mathbb{E}\max_{\nu=k,\dots,\ell}|\mathcal{E}_\nu|_*^p &\le 2\cdot 4^{p-1}\Big\{\sum_{i=0}^{k-1}\mathbb{E}|D_i|^p(\frac{1}{c_1}+\frac{1}{N}\left(C_1c_\beta+C_3\right)M_{k,Q})\\
&+\ \frac{1}{N}\left(C_1c_\beta+C_3\right)M_{k,Q}\left(\sum_{i=-M^*N_\tau}^{-1}\mathbb{E}|D_i|^p+\frac{1}{c_1}\sum_{i=k}^{\ell-1}\mathbb{E}|\mathcal{E}_i|_*^p\right)+\mathbb{E}\max_{\nu=k,\dots,\ell}|\sum_{i=k}^{\nu}\mathcal{A}^{\nu-i}\mathcal{D}_i|_*^p\Big\}\\
&\le\ 2\cdot 4^{p-1}\Big\{\sum_{i=0}^{k-1}\mathbb{E}|D_i|^p(\frac{1}{c_1}+\frac{1}{N}\left(C_1c_\beta+C_3\right)M_{k,Q})\\
&\quad+\frac{1}{N}\left(C_1c_\beta+C_3\right)M_{k,Q}\left(M^*N_\tau\sup_{s\in J}\mathbb{E}|D(s)|^p+\frac{1}{c_1}\sum_{i=k}^{\ell-1}\mathbb{E}|\mathcal{E}_i|_*^p\right)\\
&\quad+\mathbb{E}\max_{\nu=k,\dots,\ell}|\sum_{i=k}^{\nu}\mathcal{A}^{\nu-i}\mathcal{D}_i|_*^p\Big\}\\
&\le\ C_4\sum_{i=0}^{k-1}\mathbb{E}|D_i|^p+C_5\sup_{s\in J}\mathbb{E}|D(s)|^p+2\cdot 4^{p-1}\mathbb{E}\max_{\nu=k,\dots,\ell}|\sum_{i=k}^{\nu}\mathcal{A}^{\nu-i}\mathcal{D}_i|_*^p\\
&\quad+\frac{1}{N}C_6\sum_{i=k}^{\ell-1}\mathbb{E}\max_{\nu=k,\dots,i}|\mathcal{E}_\nu|_*^p
\end{aligned}$$

with properly chosen constants C_4, C_5, C_6. We now apply Lemma 2.4 with $a_\ell := 0,\ \ell = 1,\dots,k-1$ and $a_\ell := \mathbb{E}\max\limits_{\nu=k,\dots,\ell}|\mathcal{E}_\nu|_*^p,\ \ell = k,\dots,N$, and obtain the intermediate result

$$\begin{aligned}
&\mathbb{E}\max_{\nu=k,\dots,N}|\mathcal{E}_\nu|_*^p\\
&\quad\le e^{C_6}\Big\{C_4\sum_{i=0}^{k-1}\mathbb{E}|D_i|^p+C_5\sup_{s\in J}\mathbb{E}|D(s)|^p+2\cdot 4^{p-1}\mathbb{E}\max_{\nu=k,\dots,N}|\sum_{i=k}^{\nu}\mathcal{A}^{\nu-i}\mathcal{D}_i|_*^p\Big\}\\
&\qquad\le \hat{S}\Big\{\max_{i=0,\dots,k-1}\mathbb{E}|D_i|^p+\sup_{s\in J}\mathbb{E}|D(s)|^p+\mathbb{E}\max_{\nu=k,\dots,N}|\sum_{i=k}^{\nu}\mathcal{A}^{\nu-i}\mathcal{D}_i|_*^p\Big\}\ . \qquad (52)
\end{aligned}$$

It remains to deal with the term labelled 4) (in (48)), i.e. the perturbations D_i in $\mathcal{D}_i$. We decompose D_i, and, analogously, $\mathcal{D}_i$ into

$$D_i = R_i + S_i = R_i + \sum_{j=1}^{k} S_{j,i-j+1}, \qquad \mathcal{D}_i = \mathcal{R}_i + \mathcal{S}_i = \mathcal{R}_i + \sum_{j=1}^{k} \mathcal{S}_{j,i-j+1},$$

where $S_{j,i-j+1}$ is $\mathcal{F}_{t_{i-j+1}}$-measurable with $\mathbb{E}(S_{j,i-j+1}|\mathcal{F}_{t_{i-j}}) = 0$ for $i = k, \ldots, N$ and $j = 1, \ldots, k$. Proceeding in the same way as in (49) we subsume under $\tilde{M}_i$ all the terms of the form $\mathcal{A}^{\nu}\mathcal{S}_{j,i}$ occuring in 4.). Then we have

$$\sum_{i=k}^{\nu} \mathcal{A}^{\nu-i}\mathcal{D}_i = \sum_{i=1}^{\nu} \tilde{M}_i + \sum_{i=k}^{\nu} \mathcal{A}^{\nu-i}\mathcal{R}_i \ , \tag{53}$$

and $\{\sum_{i=1}^{\nu} \tilde{M}_i, \mathcal{F}_{t_\ell}\}_{\ell=0}^{N}$ is a martingale and has finite p-th order moments. For the martingale part of (53) we apply again the Burkholder-Davis-Gundy inequality (cf. e.g. [45, VII.3] or [30, I.1.7]) and obtain

$$\begin{aligned}
\mathbb{E}\,(\max_{\nu=k,\ldots,N} |\sum_{i=1}^{\nu} \tilde{M}_i|_*^p) &\leq B_p\,\mathbb{E}\,(\sum_{i=1}^{N} |\tilde{M}_i|_*^2)^{\frac{p}{2}} \leq B_p\,N^{\frac{p}{2}-1} \sum_{i=1}^{N} \mathbb{E}|\tilde{M}_i|_*^p \\
&\leq B_p\,N^{\frac{p}{2}-1} \sum_{i=1}^{N} k^{p-1} \sum_{j=1}^{k} \mathbb{E}|\mathcal{S}_{j,i}|_*^p \\
&\leq B_p\,N^{\frac{p}{2}-1} k^{p-1} c_2 \sum_{i=1}^{N} \sum_{j=1}^{k} \mathbb{E}|S_{j,i}|^p \\
&= B_p\,N^{\frac{p}{2}-1} k^{p-1} c_2 \sum_{i=k}^{N} \mathbb{E} \sum_{j=1}^{k} |S_{j,i-j+1}|^p \\
&\leq B_p\,N^{\frac{p}{2}-1} k^{p-1} c_2 C_p \sum_{i=k}^{N} \mathbb{E}|S_i|^p \\
&=: C_7\,N^{\frac{p}{2}-1} \sum_{i=k}^{N} \mathbb{E}|S_i|^p \\
&\leq C_7\,N^{\frac{p}{2}} \max_{i=k,\ldots,N} \mathbb{E}|S_i|^p.
\end{aligned}$$

For the other part of (53) we obtain

$$\begin{aligned}
\mathbb{E} \max_{\nu=k,\ldots,N} |\sum_{i=k}^{\nu} \mathcal{A}^{\nu-i}\mathcal{R}_i|_*^p &\leq N^{p-1}\,\mathbb{E} \sum_{i=k}^{N} |\mathcal{A}^{\nu-i}\mathcal{R}_i|_*^p = N^{p-1}\,\mathbb{E} \sum_{i=k}^{N} |\mathcal{R}_i|_*^p \\
&\leq N^p\,\mathbb{E} \max_{i=k,\ldots,N} |\mathcal{R}_i|_*^p \\
&\leq c_2 N^p\,\mathbb{E} \max_{i=k,\ldots,N} |R_i|^p \,.
\end{aligned}$$

Together, this yields

$$
\begin{aligned}
\mathbb{E} \max_{\nu=k,\dots,N} |\sum_{i=k}^{\nu} \mathcal{A}^{\nu-i}\mathcal{D}_i|_*^p &= \mathbb{E} \max_{\nu=k,\dots,N} |\sum_{i=1}^{\nu} \tilde{M}_i + \sum_{i=k}^{\nu} \mathcal{A}^{\nu-i}\mathcal{R}_i|_*^p \\
&\leq 2^{p-1} \left\{\mathbb{E} \max_{\nu=k,\dots,N} |\sum_{i=1}^{\nu} \tilde{M}_i|_*^p + \mathbb{E} \max_{\nu=k,\dots,N} |\sum_{i=k}^{\nu} \mathcal{A}^{\nu-i}\mathcal{R}_i|_*^p\right\} \\
&\leq 2^{p-1} \left\{C_7 N^{\frac{p}{2}} \max_{i=k,\dots,N} \mathbb{E}|S_i|^p + c_2 N^p \,\mathbb{E} \max_{i=k,\dots,N} |R_i|^p\right\} \\
&\leq 2^{p-1} \left\{C_7 \frac{T^{\frac{p}{2}}}{h^{\frac{p}{2}}} \max_{i=k,\dots,N} \mathbb{E}|S_i|^p + c_2 \frac{T^p}{h^p} \,\mathbb{E} \max_{i=k,\dots,N} |R_i|^p\right\}.
\end{aligned}
$$

Inserting this into the intermediate result (52) we have

$$
\begin{aligned}
\mathbb{E} \max_{\nu=k,\dots,N} |\mathcal{E}_\nu|_*^p &\leq \hat{S}\Big\{ \max_{i=0,\dots,k-1} \mathbb{E}|D_i|^p + \sup_{s\in J} \mathbb{E}|D(s)|^p + \mathbb{E} \max_{\nu=k,\dots,N} |\sum_{i=k}^{\nu} \mathcal{A}^{\nu-i}\mathcal{D}_i|_*^p \Big\} \\
&\leq \hat{S}\Big\{ \max_{i=0,\dots,k-1} \mathbb{E}|D_i|^p + \sup_{s\in J} \mathbb{E}|D(s)|^p \\
&\qquad +2^{p-1} \left\{C_7 \frac{T^{\frac{p}{2}}}{h^{\frac{p}{2}}} \max_{i=k,\dots,N} \mathbb{E}|S_i|^p + c_2 \frac{T^p}{h^p} \,\mathbb{E} \max_{i=k,\dots,N} |R_i|^p\right\} \Big\}.
\end{aligned}
$$

Finally, we estimate $|E_\nu|^p \leq |\mathcal{E}_\nu|_*^p$ and take the p-th root to obtain the final estimate

$$
\begin{aligned}
&(\mathbb{E} \max_{\ell=1,\dots,N} |E_\ell|^p)^{\frac{1}{p}} \\
&\leq S_p\Big\{ \sup_{s\in J} \|D(s)\|_{L_p} + \max_{\ell=0,\dots,k-1} \|D_\ell\|_{L_p} + \frac{(\mathbb{E} \max_{\ell=k,\dots,N} |R_\ell|^p)^{\frac{1}{p}}}{h} + \max_{\ell=k,\dots,N} \frac{\|S_\ell\|_{L_p}}{h^{1/2}} \Big\}. \qquad (54)
\end{aligned}
$$

We have now proved the validity of the stronger stability inequality (13). The proof of the weaker estimate (12) is similar.

B Heuristic justification of the Itô-formula

With heuristic we mean, that we omit all details concerning modes of convergence, conditions on the appearing functions and proofs. We refer to the major work [22] for these details. Here, within the limitations of this paper, we would like to explain the terms appearing in the Itô-formula in the lines (21) to (24).

Take a partition of the interval $[t_0, t]$ as $\{t_0, t_1, \dots, t_N\}$, with $t_0 < t_1 < \dots < t_N = t$ and set $\Delta t_n = t_n - t_{n-1}$, $\widetilde{t}_n = t_{n-1} + \delta_n(t_n - t_{n-1})$, $\delta_n \in (0,1)$, $\Delta X(t_n) = X(t_n) - X(t_{n-1})$ and $\widetilde{X}(t_n) = X(t_{n-1}) + \gamma_{t_n}(X(t_n) - X(t_{n-1}))$, $\gamma_{t_l} \in (0,1)$, γ_{t_l}

random. We continue to use the abbreviation $\phi\big|_s$ to indicate the evaluation of a function ϕ at s. If we evaluate at an intermediate value $\widetilde{t}_n$, $\widetilde{X}(t_n)$, we write $\phi(\widetilde{t},.,.)\big|_{t_{n-1}}$, $\phi(.,\widetilde{X},.)\big|_{t_{n-1}}$ or $= \phi(.,.,\widetilde{X}_\tau)\big|_{t_{n-1}}$. We then apply the deterministic Taylor-formula to

$$\phi(t,X(t),X(t-\tau)) - \phi(t_0,X(t_0),X(t_0-\tau))$$

$$= \sum_{n=1}^{N} \phi\big|_{t_n} - \phi(t_{n-1},X(t_n),X(t_n-\tau)) + \phi(t_{n-1},X(t_n),X(t_n-\tau)) - \phi\big|_{t_{n-1}}$$

$$= \sum_{n=1}^{N} \phi_t(\widetilde{t},.,.)\big|_{t_n}\Delta t_n + \sum_{n=1}^{N} \phi_x\big|_{t_{n-1}}\Delta X(t_n) + \sum_{n=1}^{N} \frac{1}{2}\phi_{xx}(.,\widetilde{X},.)\big|_{t_{n-1}}(\Delta X(t_n))^2 \tag{55}$$

$$+ \sum_{n=1}^{N} \frac{1}{2}\phi_{zz}(.,.,\widetilde{X}_\tau)\big|_{t_{n-1}}(\Delta X(t_n-\tau))^2 + \boxed{\sum_{n=1}^{N} \phi_z\big|_{t_{n-1}}\Delta X(t_n-\tau)} \tag{56}$$

$$+ \sum_{n=1}^{N} \frac{1}{2}\phi_{xz}(.,\widetilde{X},\widetilde{X}_\tau)\,(\Delta X(t_n))(\Delta X(t_n-\tau)) + \text{'higher order terms'}. \tag{57}$$

The terms in (55) are those appearing in the Taylor expansion for the non-delay Itô formula and they will converge to (20). The second order terms in (56) and (57) will converge to the first integral in (21) and to 0, respectively.
For the framed term in (56) we have

$$\sum_{n=1}^{N} \phi_z\big|_{t_{n-1}}\Delta X(t_n-\tau) = \sum_{n=1}^{N} \phi_z\big|_{t_{n-1}} \times \left\{ \int_{t_{n-1}-\tau}^{t_n-\tau} F\big|_s \,\mathrm{d}s + \int_{t_{n-1}-\tau}^{t_n-\tau} G\big|_s \,\mathrm{d}W(s) \right\}.$$

Now $\sum_{n=1}^{N} \phi_z\big|_{t_{n-1}} \left\{ \int_{t_{n-1}-\tau}^{t_n-\tau} F\big|_s \,\mathrm{d}s \right\}$ converges to $\int_{t_0-\tau}^{t-\tau} \phi_z\big|_{s+\tau}F\big|_s \,\mathrm{d}s$, the first integral within the curly brackets in (21).
The 'obvious' limit of the second term would be $\int_{t_0-\tau}^{t-\tau} \phi_z\big|_{s+\tau}G\big|_s \,\mathrm{d}W(s)$, however, then the integrand $\phi_z(s+\tau,\underline{X(s+\tau)},X(s))G\big|_s$ is not adapted to the filtration generated by the Wiener process on the interval $[t_{n-1}-\tau,t_n-\tau]$, thus the integral $\int_{t_{n-1}-\tau}^{t_n-\tau} \phi_z\big|_{s+\tau}G\big|_s \,\mathrm{d}W(s)$ would be a non-anticipating integral and not longer an Itô integral. At this point the application of the Malliavin calculus becomes necessary. The idea developed in [22] is to use a property of the Skorohod integral, i.e., a formula for the Skorohod integral of a process multiplied by a random variable ([38, Thm. 3.2], [37, Eq. 1.49 in Chapter1]) and an application of this formula yields

$$\phi_z\big|_{t_{n-1}} \int_{t_{n-1}-\tau}^{t_n-\tau} G\big|_s \,\mathrm{d}W(s) = \int_{t_{n-1}-\tau}^{t_n-\tau} G\big|_s\, \phi_z\big|_{t_{n-1}} \,\mathrm{d}W(s) + \int_{t_{n-1}-\tau}^{t_n-\tau} D_s(\phi_z\big|_{t_{n-1}})G\big|_s \,\mathrm{d}s.$$

With the chain rule for weak derivatives ([37, Prop. 1.2.2]) one obtains

$$D_s(\phi_z\big|_{t_{n-1}}) = \phi_{zx}\big|_{t_{n-1}} D_sX(t_{n-1}) + \phi_{zz}\big|_{t_{n-1}} D_sX(t_{n-1}-\tau)$$

where the Malliavin derivative of X is given by

$$D_sX(t) = G\big|_s \mathbb{1}_{\{s\le t\}} + \int_0^t D_sF\big|_r \,\mathrm{d}r + \int_0^t D_sG\big|_r \,\mathrm{d}W(r)\,.$$

Summarizing the last three formulas, we arrive at

$$\begin{aligned}
\sum_{n=1}^N \phi_z\big|_{t_{n-1}} \int_{t_{n-1}-\tau}^{t_n-\tau} G\big|_s \,\mathrm{d}W(s) \;=\; \sum_{n=1}^N \Bigg\{ & \underbrace{\int_{t_{n-1}-\tau}^{t_n-\tau} \phi_z\big|_{t_{n-1}} G\big|_s \,\mathrm{d}W(s)}_{1)} \\
& + \underbrace{\int_{t_{n-1}-\tau}^{t_n-\tau} \phi_{zx}\big|_{t_{n-1}} \mathbb{1}_{\{s\le t_{n-1}\}} G^2\big|_s \,\mathrm{d}s}_{2)} + \underbrace{\int_{t_{n-1}-\tau}^{t_n-\tau} \phi_{zz}\big|_{t_{n-1}} \mathbb{1}_{\{s\le t_{n-1}-\tau\}} G^2\big|_s \,\mathrm{d}s}_{3)} \\
& + \underbrace{\int_{t_{n-1}-\tau}^{t_n-\tau} \phi_{zx}\big|_{t_{n-1}} \int_0^{t_{n-1}} D_sF\big|_r \,\mathrm{d}r\, G\big|_s \,\mathrm{d}s}_{4)} + \underbrace{\int_{t_{n-1}-\tau}^{t_n-\tau} \phi_{zx}\big|_{t_{n-1}} \int_0^{t_{n-1}} D_sG\big|_r \,\mathrm{d}W(r)\, G\big|_s \,\mathrm{d}s}_{5)} \\
& + \underbrace{\int_{t_{n-1}-\tau}^{t_n-\tau} \phi_{zz}\big|_{t_{n-1}} \int_0^{t_{n-1}-\tau} D_sF\big|_r \,\mathrm{d}r\, G\big|_s \,\mathrm{d}s}_{6)} + \underbrace{\int_{t_{n-1}-\tau}^{t_n-\tau} \phi_{zz}\big|_{t_{n-1}} \int_0^{t_{n-1}-\tau} D_sG\big|_r \,\mathrm{d}W(r)\, G\big|_s \,\mathrm{d}s}_{7)} \Bigg\}.
\end{aligned}$$

The sum of integrals denoted by 1) converges to the framed Skorokhod integral in (22), the term 2) converges to the second integral in (22). The term 3) is equal to 0. The sums of the terms 4) and 5) converge to (23), those in 6) and 7) to (24).

References

[1] J. A. D. Appleby and E. Buckwar, *Noise induced oscillation in solutions of stochastic delay differential equations*, tech. report, Humboldt University Berlin, 2003. (submitted), Preprint: `http://www.mathematik.hu-berlin.de/~buckwar/publications`.

[2] C. T. H. Baker and E. Buckwar, *Numerical analysis of explicit one-step methods for stochastic delay differential equations*, LMS J. Comput. Math., 3 (2000), pp. 315–335.

[3] D. S. Battisti, *On the role of off-equatorial oceanic Rossby waves during ENSO*, J.Phys. Oceanography, 19 (1989), pp. 551–559.

[4] A. Bellen and M. Zennaro, *Numerical methods for delay differential equations.*, Numerical Mathematics and Scientific Computation. Oxford: Oxford University Press. , 2003.

[5] E. Beretta, V. B. Kolmanovskii, and L. Shaikhet, *Stability of epidemic model with time delays influenced by stochastic perturbations*, Math. Comput. Simulation, 45 (1998), pp. 269–277. Delay systems (Lille, 1996).

[6] A. Beuter and K. Vasilakos, *Effects of noise on a delayed visual feedback system*, J. Theor. Biology, 165 (1993), pp. 389–407.

[7] G. A. Bocharov and F. A. Rihan, *Numerical modelling in biosciences using delay differential equations*, J. Comput. Appl. Math., 125 (2000), pp. 183–199.

[8] E. Buckwar, *One-step approximations for stochastic functional differential equations*, tech. report, Humboldt University Berlin, 2004. (submitted), Preprint:
`http://www.mathematik.hu-berlin.de/∼buckwar/publications`.

[9] E. Buckwar and R. Winkler, *Multi-step methods for sdes and their application to problems with small noise*, (2003). submitted for publication.

[10] J. M. Buldú, J. Garcia-Ojalvo, C. R. Mirasso, M. C.Torrent, and J. M. Sancho, *Effect of external noise correlation in optical coherence resonance*, Phys. Rev. E, 64 (2001), pp. 051109–1–051109–4.

[11] M. A. Cane, M. Matthias, and S. E. Zebiak, *A study of self-excited oscillations in the tropical ocean-atmosphere system*, J. Atmos. Sci., 46 (1990), pp. 1562–1577.

[12] M.-H. Chang and R. K. Youree, *The European option with hereditary price structures: Basic theory*, Appl. Math. Comp., 102 (1999), pp. 279–296.

[13] C. W. Eurich and J. G. Milton, *Noise-induced transitions in human postural sway*, Phys. Rev. E, 54 (1996), pp. 6681–6684.

[14] M. Fischer and P. Imkeller, *A two-state model for noise induced resonance in bistable systems with delay.* Preprint.

[15] J. Garcia-Ojalvo and R. Roy, *Noise amplification in a stochastic ikeda model*, Phys. Lett. A, 224 (1996), pp. 51–56.

[16] W. Gautschi, *Numerical analysis. An introduction*, Birkhäuser, Boston, 1997.

[17] D. Goldobin, M. Rosenblum, and A. Pikovsky, *Coherence of noisy oscillators with delayed feedback*, Physica A, 327 (2003), pp. 124–128.

[18] M. Grigoriu, *Control of time delay linear systems with Gaussian white noise*, Probabilistic Engineering Mechanics, 12 (1997), pp. 89–96.

[19] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving ordinary differential equations. I: Nonstiff problems. 2. rev. ed.*, vol. 8 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 1993.

[20] J. Hale and S. M. V. Lunel, *Introduction to Functional Differential Equations*, Springer Verlag, New York, 1993.

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

[22] Y. Hu, S.-E. A. Mohammed, and F. Yan, *Discrete-time approximations of stochastic delay differential equations: The Milstein scheme*, Annals of Probability, 32 (2004), pp. 265–314.

[23] D. Huber and L. S. Tsimring, *Dynamics of an ensemble of noisy bistable elements with global time delayed coupling*, Phys. Rev. Lett., 91 (2003), p. 260601.

[24] A. R. Humphries and A. M. Stuart, *Dynamical Systems and Numerical Analysis*, Cambridge University Press, Cambridge, 1996.

[25] C. Jiang, A. W. Troesch, and S. W. Shaw, *Capsize criteria for ship models with memory-dependent hydrodynamics and random excitation*, Phil. Trans. R. Soc. Lond. A, 358 (2000), pp. 1761–1791.

[26] S. Kim, S. H. Park, and C. S. Ryu, *Multistability in coupled oscillator systems with time delay*, Phys. Rev. Lett., 79 (1997), pp. 2911–2914.

[27] V. B. Kolmanovskiĭ and A. Myshkis, *Introduction to the Theory and Applications of Functional Differential Equations*, Kluwer Academic Publishers, Dordrecht, 1999. ISBN 0-7923-5504-0.

[28] A. Longtin, *Noise-induced transitions at a Hopf bifurcation in a first order delay-differential equation*, Phys. Rev. A, 44 (1991), pp. 4801–4813.

[29] A. Longtin, J. G. Milton, J. Bos, and M. C. Mackey, *Noise and critical behavior of the pupil light reflex at oscillation onset*, Phys. Rev. A, 41 (1990), pp. 6992–7005.

[30] X. Mao, *Stochastic differential equations and their applications*, Horwood Publishing Limited, Chichester, 1997.

[31] C. Masoller, *Noise-induced resonance in delayed feedback systems*, Phys. Rev. Lett., 88 (2002), pp. 034102–1–034102–4.

[32] ——, *Numerical investigation of noise-induced resonance in a semiconductor laser with optical feedback.*, Physica D, 168-169 (2002), pp. 171–176.

[33] ——, *Distribution of residence times of time-delayed bistable systems driven by noise.*, Phys. Rev. Lett., 90 (2003), pp. 020601–1–020601–4.

[34] G. Milstein, *Numerical integration of stochastic differential equations*, Kluwer, 1995. Translation from the Russion original of 1988.

[35] G. N. Milstein and M. V. Tretyakov, *Stochastic Numerics for Mathematical Physics*, Springer Verlag, Berlin, 2004.

[36] S. E. A. Mohammed, *Stochastic functional differential equations*, Pitman (Advanced Publishing Program), Boston, Mass., 1984.

[37] D. Nualart, *The Malliavin calculus and related topics.*, Probability and Its Applications. New York, Springer-Verlag, 1995.

[38] D. Nualart and E. Pardoux, *Stochastic calculus with anticipating integrands.*, Probab. Theory Relat. Fields, 78 (1988), pp. 535–581.

[39] T. Ohira and Y. Sato, *Resonance with noise and delay*, Phys. Rev. Lett., 82 (1999), pp. 2811–2815.

[40] M. D. Paola and A. Pirrotta, *Time delay induced effects on control of linear systems under random excitation*, Probabilistic Engineering Mechanics, 16 (2001), pp. 43–51.

[41] R. J. Peterka, *Postural control model interpretation of stabilogram diffusion analysis*, Biol Cybern., 82 (2000), pp. 335–343.

[42] R. Plato, *Numerische Mathematik kompakt. Grundlagenwissen für Studium und Praxis*, Vieweg, Braunschweig, 2000.

[43] W. Römisch and R. Winkler, *Stepsize control for mean-square numerical methods for sdes with small noise.* Preprint 2003-08, Institut für Mathematik, Humboldt-Universität Berlin, 2003, and submitted.

[44] ——, *Stochastic DAEs in circuit simulation*, in Modeling, Simulation and Optimization of Integrated Circuits, A. G. K. Antreich, R. Bulirsch and P. Rentrop, eds., Birkhäuser, 2003, pp. 303–318.

[45] A. N. Shiryaev, *Probability (Second Edition)*, Springer, New York, 1966.

[46] W. Sichermann and E. Kreuzer, *Capsize risk assessment and designing against capsize*, PAMM - Proc. in Appl. Math. and Mech., 3 (2003), pp. 501–502.

[47] P. Tapaswi and A. Mukhopadhyay, *Effects of environmental fluctuation on plankton allelopathy.*, J. Math. Biol., 39 (1999), pp. 39–58.

[48] L. S. Tsimring and A. Pikovsky, *Noise-Induced Dynamics in Bistable Systems with Delay*, Phys. Rev. Lett., 87 (2001), pp. 250602–1–250602–4.

[49] R. Winkler, *Stochastic differential algebraic equations of index 1 and applications in circuit simulation*, J. Comput. Appl. Math., 157 (2003), pp. 477–505.

[50] M. K. S. Yeung and S. H. Strogatz, *Time delay in the Kuramoto model of coupled oscillators*, Phys. Rev. Lett., 82 (1999), pp. 648–651.