Mean-square convergence of stochastic multi-step methods with variable step-size

Thorsten Sickenberger [1]

Humboldt-Universität zu Berlin, Institut für Mathematik, Unter den Linden 6, 10099 Berlin, Germany

Abstract

We study mean-square consistency, stability in the mean-square sense and mean-square convergence of drift-implicit linear multi-step methods with variable step-size for the approximation of the solution of Itô stochastic differential equations. We obtain conditions that depend on the step-size ratios and that ensure mean-square convergence for the special case of adaptive two-step Maruyama schemes. Further, in the case of small noise we develop a local error analysis with respect to the $h-\epsilon$ approach and we construct some stochastic linear multi-step methods with variable step-size that have order 2 behaviour if the noise is small enough.

Key words: Stochastic linear multi-step methods, Adaptive methods, Mean-square convergence, Mean-square numerical stability, Mean-square consistency, Small noise, Two-step Maruyama methods

1991 MSC: 60H35, 65C30, 65L06, 60H10, 65L20

1 Introduction

We consider Itô stochastic differential equations (SDEs) of the form

$$X(s)\Big|_{t_0}^{t} = \int_{t_0}^{t} f(X(s), s)\mathrm{d}s + \int_{t_0}^{t} G(X(s), s)\mathrm{d}W(s), \quad X(t_0) = X_0, \qquad (1)$$

for $t \in \mathcal{J}$, where $\mathcal{J} = [t_0, T]$. The drift and diffusion functions are given as $f : \mathbb{R}^n \times \mathcal{J} \to \mathbb{R}^n$, $G = (g_1, \ldots, g_m) : \mathbb{R}^n \times \mathcal{J} \to \mathbb{R}^{n\times m}$. The process W is a m-dimensional Wiener process on a given probability space $(\Omega, \mathcal{F}, P)$ with a filtration $(\mathcal{F}_t)_{t\in\mathcal{J}}$ and X_0 is a given $\mathcal{F}_{t_0}$-measurable initial value, independent

Email address: sickenberger@math.hu-berlin.de (Thorsten Sickenberger).

[1] Support of BMBF through project 03RONAVN is gratefully acknowledged.

of the Wiener process and with finite second moments. It is assumed that there exists a path-wise unique strong solution $X(\cdot)$ of (1).

In this paper the mean-square convergence properties of, in general, drift-implicit linear multi-step methods with variable step-size (LMMs) are analysed w.r.t. the approximation of the solution of (1). Although there is a well-developed convergence analysis for discretisation schemes for SDEs, less emphasis has been put on a numerical stability analysis to estimate the effect of errors. Numerical stability allows to conclude convergence from consistency. So, we aim for a numerical stability inequality for such schemes with variable step-size. Our approach is based on techniques proposed in [2] in the context of equidistant grids.

Most common methods use fixed step-size and thus are not able to react to the characteristics of a solution path. It is clear that an efficient integrator must be able to change the step-size. However, changing the step-size with multi-step methods is difficult, so we have to construct methods which are adjusted to variable grid points. Only a few papers deal with adaptive step-size control; for an example for strong approximation see [3, 6]. Certainly, for an adaptive algorithm we have to explain the choice of suitable error estimates and step-size strategies. This will be the subject of a separate paper.

Our interest in stochastic multi-step methods (SLMMs) stems from applications with small noise in circuit simulation (see e.g. [4, 10, 13]), where especially the backward differential formulae (BDF) have proven valuable in the deterministic case. We act on a suggestion of Milstein and Tretyakov [9] and construct special numerical methods which are more effective and easier than in the general case. Some first simulation results that illustrate the performance of the presented methods can be found in [12].

The structure of the paper is as follows. In Section 2 we introduce the class of SLMMs considered and provide necessary definitions and useful facts. In Section 3 we deal with variable step-size and we focus upon our main result of consistency, stability and convergence in the mean-square sense. Additionally to the properties in the context of equidistant grids we have to fulfil conditions for the maximum step-size on the grid and for the step-size ratios of the sequence. In Section 4 we consider adaptive two-step-Maruyama methods. Both the coefficients of such a scheme and the conditions for their mean-square consistency actually depend on the step-size ratios. As an application, we get some of the properties of deterministic LMMs for the SDEs with small noise, i. e. SDEs that can be written in the form

$$X(s)\Big|_{t_0}^{t} = \int_{t_0}^{t} f(X(s),s)\mathrm{d}s + \epsilon \int_{t_0}^{t} \hat{G}(X(s),s)\mathrm{d}W(s), \quad X(t_0) = X_0, \qquad (2)$$

for $t \in \mathcal{J}$, where $\epsilon \ll 1$ is a small parameter. The appendix contains the proof of Theorem 2.

2 Definitions and preliminary results

We denote by $|\cdot|$ the Euclidian norm in $\mathbb{R}^n$ and by $\|\cdot\|$ the corresponding induced matrix norm. The mean-square norm of a vector-valued square-integrable random variable $Z \in L_2(\Omega, \mathbb{R}^n)$, with $\mathbb{E}$ the expectation with respect to P, will be denoted by

$$\|Z\|_{L_2} := (\mathbb{E}|Z|^2)^{1/2} .$$

Consider a discretisation $t_0 < t_1 < \ldots < t_N = T$ of $\mathcal{J}$ with step-sizes $h_\ell := t_\ell - t_{\ell-1}$, $\ell = 1, \ldots, N$. Let $\mathbf{h} := \max_{1\le\ell\le N} h_\ell$ be the maximal step-size of the grid and $\kappa_\ell = h_\ell / h_{\ell-1}$, $\ell = 2, \ldots, N$ the step-size ratios.

We discuss *mean-square convergence* of possibly drift-implicit stochastic linear multi-step methods (SLMM) with variable step-size, which for $\ell = k, \ldots, N$, takes the form

$$\sum_{j=0}^{k} \alpha_{\ell,j}\, X_{\ell-j} = h_\ell \sum_{j=0}^{k} \beta_{\ell,j}\, f(X_{\ell-j}, t_{\ell-j}) + \sum_{j=1}^{k} \Gamma_{\ell,j}(X_{\ell-j}, t_{\ell-j})\, I^{t_{\ell-j}, t_{\ell-j+1}}. \quad (3)$$

The coefficients $\alpha_{\ell,j}$, $\beta_{\ell,j}$ and the diffusion terms $\Gamma_{\ell,j}$ actually depend on the ratios κ_ℓ for $\ell = 2, \ldots, N$. We require given initial values $X_0, \ldots, X_{k-1} \in L_2(\Omega, \mathbb{R}^n)$ such that X_ℓ is $\mathcal{F}_{t_\ell}$-measurable for $\ell = 0, \ldots, k-1$. As in the deterministic case, usually only $X_0 = X(t_0)$ is given by the initial value problem and the values $X_1, \ldots, X_{k-1}$ need to be computed numerically. This can be done by suitable one-step methods, where one has to be careful to achieve the desired accuracy. Every diffusion term $\Gamma_{\ell,j}(x,t)\; I^{t_{\ell-j}, t_{\ell-j+1}}$ is a finite sum of terms each containing an appropriate function $\mathcal{G}_{\ell,j}$ of x and t multiplied by a multiple Wiener integral over $[t_{\ell-j}, t_{\ell-j+1}]$, i.e. it takes the general form

$$\Gamma_{\ell,j}(x,t)\; I^{t_{\ell-j}, t_{\ell-j+1}} = \sum_{r=1}^{m} \mathcal{G}^{r}_{\ell,j}(x,t)\; I^{t_{\ell-j}, t_{\ell-j+1}}_{r} + \sum_{\substack{r_1, r_2 = 0 \\ r_1 + r_2 > 0}}^{m} \mathcal{G}^{r_1, r_2}_{\ell,j}(x,t)\; I^{t_{\ell-j}, t_{\ell-j+1}}_{r_1, r_2} + \ldots\, .$$

A general multiple Wiener integral is given by

$$I^{t,t+h}_{r_1, r_2, \ldots, r_j}(y) = \int_t^{t+h} \int_t^{s_1} \ldots \int_t^{s_{j-1}} y(X(s_j), s_j) \mathrm{d}W_{r_1}(s_j) \ldots \mathrm{d}W_{r_j}(s_1), \quad (4)$$

where $r_i \in \{0, 1, \ldots, m\}$ and $\mathrm{d}W_0(s) = \mathrm{d}s$. If $y \equiv 1$ we write $I^{t,t+h}_{r_1, r_2, \ldots, r_j}$. Note that the integral $I^{t,t+h}_{r}$ is simply the increment $W_r(t+h) - W_r(t)$ of the scalar Wiener process W_r. The term $I^{t,t+h}$ denotes the collection of multiple Wiener integrals associated with the interval $[t, t+h]$. It is known [8] that the multiple integrals have the properties

$$\mathbb{E}(I^{t,t+h}_{r_1, \ldots, r_j}(\cdot)|\mathcal{F}_t) = 0 \quad \text{if at least one of the indices } r_i \neq 0, \quad (5)$$

$$\|\mathbb{E}(I^{t,t+h}_{r_1, \ldots, r_j}(\cdot)|\mathcal{F}_t)\|_{L_2} = O(h^{l_1 + l_2/2}), \quad (6)$$

where l_1 is the number of zero indices r_i and l_2 the number of non-zero indices r_i. We point out that for $\beta_{\ell,0} = 0$, $\ell = k, \ldots, N$, the SLMM (3) is explicit, otherwise it is drift-implicit. For the diffusion term we use an explicit discretisation.

3 Mean-square convergence of stochastic linear multi-step methods with variable step-size

We will consider mean-square convergence of SLMMs in the sense discussed in Milstein and others [1, 8, 7, 13]. Note that in the literature the term *strong* convergence is sometimes used synonymously for our expression *mean-square* convergence.

Definition 1 *We call the SLMM (3) for the approximation of the solution of the SDE (1)* **mean-square convergent** *if the global error* $X(t_\ell) - X_\ell$ *satisfies*

$$\max_{\ell=1,\ldots,N} \|X(t_\ell) - X_\ell\|_{L_2} \to 0 \quad \textit{as} \quad \mathbf{h} \to 0,$$

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

$$\max_{\ell=1,\ldots,N} \|X(t_\ell) - X_\ell\|_{L_2} \leq C \cdot \mathbf{h}^\gamma,$$

with a constant $C > 0$ *which is independent of the step-size* $\mathbf{h}$.

The aim is to conclude mean-square convergence from local properties of the SLMM by means of numerical stability in the mean-square sense, together with mean-square consistency.

3.1 Numerical stability in the mean-square 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 and should not be mistaken for properties like asymptotic stability. We assume that the scheme (3) for the SDE (1) satisfies the following properties:

(A1) the function $f : R^n \times J \to R^n$ satisfies a **uniform Lipschitz condition** with respect to x, if there exists a positive constant L_f, such that

$$|f(x,t) - f(y,t)| \leq L_f |x - y|, \quad \forall x, y \in \mathbb{R}^n, t \in \mathcal{J}, \tag{7}$$

(A2) the function $\Gamma_{\ell,j} : \mathbb{R}^n \times J \to \mathbb{R}^{n \times m_\Gamma}$ satisfies a **uniform Lipschitz condi-**

tion with respect to x:

$$|\Gamma_{\ell,j}(x,t) - \Gamma_{\ell,j}(y,t)| \le L_{\Gamma_{\ell,j}} |x-y|, \quad \forall x,y \in \mathbb{R}^n, t \in \mathcal{J}, \tag{8}$$

where $L_{\Gamma_{\ell,j}}$ is a positive constant;

(A3) and the function $\Gamma_{\ell,j} : \mathbb{R}^n \times J \to \mathbb{R}^{n\times m_\Gamma}$ satisfies a **linear growth condition** with a positive constant $K_{\Gamma_{\ell,j}}$ in the form

$$|\Gamma_{\ell,j}(x,t)| \le K_{\Gamma_{\ell,j}}(1+|x|^2)^{\frac{1}{2}}, \quad \forall x \in \mathbb{R}^n, t \in \mathcal{J}, \text{ and} \tag{9}$$

(A4) the coefficients $\alpha_{\ell,j} = \alpha_j(\kappa_{\ell-k+1},\dots,\kappa_\ell)$ are continuous in a neighbourhood of $(1,\dots,1)$, fulfil $1+\sum_{j=1}^k \alpha_{\ell,j} = 0$ for all ℓ and the underlying constant step-size formula satisfy Dahlquist's root condition, i.e.

(i) the roots of the characteristic polynomial of (3)

$$\rho(\zeta) = \alpha_0(1,\dots,1)\zeta^k + \alpha_1(1,\dots,1)\zeta^{k-1} + \dots \alpha_k(1,\dots,1) \tag{10}$$

lie on or within the unit circle and

(ii) the roots on the unit circle are simple.

Conditions (A1) - (A3) are standard assumptions for analysing stochastic differential systems, condition (A4) is known [5] in the context of deterministic variable step-size multi-step methods. We now formulate and prove our main theorem on numerical stability. The mean-square stability estimate of the global error is based on the mean-square norm and on the conditional mean of the perturbations. Additionally to the properties in the context of equidistant grids we need a new relation between the maximum step-size and the number of steps and we have to ensure that the coefficients $\alpha_{\ell,j}$, $\beta_{\ell,j}$ and $\gamma_{\ell,j}$ of (3) are bounded. Hence we have to fulfil conditions for the maximum step-size on the grid and for the step-size ratios of the sequence.

Theorem 2 *Assume that (A1) - (A4) hold for the scheme (3). Also there exists constants $a \ge 1$, $h^0 > 0$, a stability constant $S > 0$ and due to (A4) constants κ, K ($\kappa < 1 < K$) such that the following holds true for each grid $\{t_0, t_1, \dots, t_N\}$ having the property $\mathbf{h}{:=}\max_{\ell=1,\dots,N} h_\ell \le h^0$, $\mathbf{h}{\cdot}N \le a{\cdot}(T-t_0)$ and $\kappa \le h_\ell/h_{\ell-1} \le K$ for all ℓ:*

For all F_{t_ℓ}-measurable, square-integrable initial values $X_\ell, \tilde{X}_\ell$ for $\ell = 0,\dots,k-1$ and all F_{t_ℓ}-measurable perturbations D_ℓ having finite second moments the system (3) and the perturbed discrete system

$$\sum_{j=0}^k \alpha_{\ell,j}\, \tilde{X}_{\ell-j} = h_\ell \sum_{j=0}^k \beta_{\ell,j}\, f(\tilde{X}_{\ell-j}, t_{\ell-j}) + \sum_{j=1}^k \Gamma_{\ell,j}(\tilde{X}_{\ell-j}, t_{\ell-j}) I^{t_{\ell-j},t_{\ell-j+1}} + D_\ell, \tag{11}$$

$\ell = k,\dots,N$, have unique solutions $\{X_\ell\}_{\ell=0}^N$, $\{\tilde{X}_\ell\}_{\ell=0}^N$, and the mean-square

norm of their differences $e_\ell = X_\ell - \tilde{X}_\ell$ *can be estimated by*

$$\max_{\ell=1,\dots,N} \|e_\ell\|_{L_2} \le S\Big\{ \max_{\ell=0,\dots,k-1} \|e_\ell\|_{L_2} + \max_{\ell=k,\dots,N} \Big(\frac{\|R_\ell\|_{L_2}}{\mathbf{h}} + \frac{\sqrt{\sum_{j=1}^{k} \|S_{j,\ell-j+1}\|^2_{L_2}}}{\sqrt{\mathbf{h}}} \Big) \Big\}, \tag{12}$$

where $D_\ell = R_\ell + \sum_{j=1}^{k} S_{j,\ell-j+1}$ *and* $S_{j,\ell-j+1}$ *is* $F_{\ell-j+1}$*-measurable with* $\mathbb{E}(S_{j,\ell-j+1}|F_{t_{\ell-j}}) = 0$ *for* $\ell = k,\dots,N$ *and* $j = 1,\dots,k$.

The proof is divided into several parts and given in the appendix. First, we show the existence of unique solutions of the perturbed discrete system. Second, we show that the second moments of these solutions exists, and, third, we derive a stability inequality.

If scheme (3) for the SDE (1) fulfils the assertion of Theorem 2, we call it *numerically stable in the mean-square sense* and refer to S as the *stability constant* and to (12) as the *stability inequality.*

3.2 Mean-square consistency

Different notions of errors for pathwise approximation are studied in the literature. In the following we will define what we understand by local errors. We recall the notions from [2] and define *the local error* as the defect that is obtained when the exact solution values are inserted into the numerical scheme, i.e. the local error of SLMM (3) for the approximation of the solution of the SDE (1) is given as

$$\begin{aligned} L_\ell :=& \sum_{j=0}^{k} \alpha_{\ell,j}\, X(t_{\ell-j}) - h_\ell \sum_{j=0}^{k} \beta_{\ell,j}\, f(X(t_{\ell-j}), t_{\ell-j}) \\ & - \sum_{j=1}^{k} \Gamma_{\ell,j}(X(t_{\ell-j}), t_{\ell-j})\, I^{t_{\ell-j}, t_{\ell-j+1}}, \qquad \ell = k,\dots,N, \quad (13) \\ L_\ell :=& X(t_\ell) - X_\ell, \qquad \ell = 0,\dots,k-1. \quad (14) \end{aligned}$$

In order to exploit the adaptivity and independence of the stochastic terms arising on disjoint subintervals we refer to [2] and represent the local error in the form

$$L_\ell = R_\ell + S_\ell =: R_\ell + \sum_{j=1}^{k} S_{j,\ell-j+1}, \quad \ell = k,\dots,N, \tag{15}$$

where each $S_{j,\ell-j+1}$ is $F_{t_{\ell-j+1}}$-measurable with $\mathbb{E}(S_{j,\ell-j+1}|\mathcal{F}_{t_{\ell-j}}) = 0$ for $\ell = k,\dots,N$ and $j = 1,\dots,k$. Note that the representation (15) is not unique.

Definition 3 *We call the SLMM (3) for the approximation of the solution of* *the SDE (1)* **mean-square consistent** *if the local error L_ℓ satisfies*

$$h_\ell^{-1}\|\mathbb{E}(L_\ell|\mathcal{F}_{t_{\ell-k}})\|_{L_2}\to 0 \text{ for } h_\ell\to 0, \text{ and } h_\ell^{-1/2}\|L_\ell\|_{L_2}\to 0 \text{ for } h_\ell\to 0; \quad (16)$$

and **mean-square consistent of order** γ $(\gamma > 0)$, *if the local error L_ℓ satisfies*

$$\|\mathbb{E}(L_\ell|\mathcal{F}_{t_{\ell-k}})\|_{L_2} \le \bar{c}\cdot h_\ell^{\gamma+1} \quad \text{and} \quad \|L_\ell\|_{L_2} \le c\cdot h_\ell^{\gamma+\frac{1}{2}}\ , \quad \ell = k,\ldots,N\ , \quad (17)$$

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

Subsequently we assume that the conditions of Theorem 2 are fulfilled. Following Thm. 2.8 in [2] in order to prove mean-square consistency of order γ it is then sufficient to find a representation (15) of the local error L_ℓ such that

$$\|R_\ell\|_{L_2} \le \bar{c}\cdot h_\ell^{\gamma+1} \quad \text{and} \quad \|S_\ell\|_{L_2} \le c\cdot h_\ell^{\gamma+\frac{1}{2}}\ , \quad \ell = k,\ldots,N\ , \quad (18)$$

with constants c, $\bar{c} > 0$ only depending on the SDE and its solution. Together the condition (18) imply the estimates

$$\|\mathbb{E}(L_\ell|\mathcal{F}_{t_{\ell-k}})\|_{L_2} \le O(h_\ell^{\gamma+1}) \quad \text{and} \quad \|L_\ell\|_{L_2} \le O(h_\ell^{\gamma+\frac{1}{2}})\ , \quad \ell = k,\ldots,N\ ,$$

and the SLMM (3) is mean-square consistent.

4 Local error analysis

To analyse the local error L_ℓ of a discretisation scheme for the SDE (1) and to achieve a suitable representation (15) we want to derive appropriate Itô-Taylor expansions, where we take special care to separate the multiple stochastic integrals and the step-size over the different subintervals of integration.

Let $C^{s,s-1}$ denote the class of functions form $\mathbb{R}^n \times \mathcal{J}$ to $\mathbb{R}^n$ having continuous partial derivations up to order $s-1$ and, in addition, continuous partial derivations of order s with respect to the first variable; and let C^K denote the class of functions from $\mathbb{R}^n \times \mathcal{J}$ to $\mathbb{R}^n$ that satisfies a linear growth condition (A3).

We introduce operators Λ_0 and Λ_r, $r = 1,\ldots,m$, defined on $C^{2,1}$ and $C^{1,0}$, respectively, by

$$\Lambda_0 y = y'_t + y'_x f + \frac{1}{2}\sum_{r=1}^{m}\sum_{i,j=1}^{n} y''_{x_i x_j} g_{ri} g_{rj}\ , \quad \Lambda_r y = y'_x g_r\ ,\ r = 1,\ldots,m, \quad (19)$$

and remind the reader of the notation for multiple Wiener integrals (4). Using these operators the Itô formula for a function y in $C^{2,1}$ and the solution X of

(1) reads

$$y(X(t),t) = y(X(t_0),t_0) + I_0^{t_0,t}(\Lambda_0 y) + \sum_{r=1}^{m} I_r^{t_0,t}(\Lambda_r y), \quad t \in \mathcal{J}. \tag{20}$$

4.1 Two-step-Maruyama schemes for general SDEs

We consider linear two-step-Maruyama schemes with variable step-size, thus we have for $\ell = 2,\dots,N$

$$\sum_{j=0}^{2} \alpha_{\ell,j} X_{\ell-j} = h_\ell \sum_{j=0}^{2} \beta_{\ell,j} f(X_{\ell-j}, t_{\ell-j}) + \sum_{j=1}^{2} \gamma_{\ell,j} \sum_{r=1}^{m} g_r(X_{\ell-j}, t_{\ell-j})\ I_r^{t_{\ell-j},t_{\ell-j+1}}, \tag{21}$$

where the coefficients $\alpha_{\ell,j}$, $\beta_{\ell,j}$ and $\gamma_{\ell,j}$ actually depend on the ratio $\kappa_\ell = h_\ell/h_{\ell-1}$.

We apply the Itô-formula (20) on the corresponding intervals to the drift coefficient f and trace back the values to the point $t_{\ell-2}$ to obtain

$$f(X(t_{\ell-1}),t_{\ell-1}) = f(X(t_{\ell-2}),t_{\ell-2}) + I_0^{t_{\ell-2},t_{\ell-1}}(\Lambda_0 f) + \sum_{r=1}^{m} I_r^{t_{\ell-2},t_{\ell-1}}(\Lambda_r f), \tag{22}$$

$$\begin{aligned} f(X(t_\ell),t_\ell) &= f(X(t_{\ell-2}),t_{\ell-2}) + I_0^{t_{\ell-2},t_{\ell-1}}(\Lambda_0 f) + I_0^{t_{\ell-1},t_\ell}(\Lambda_0 f) \\ &\quad + \sum_{r=1}^{m} I_r^{t_{\ell-2},t_{\ell-1}}(\Lambda_r f) + \sum_{r=1}^{m} I_r^{t_{\ell-1},t_\ell}(\Lambda_r f). \end{aligned} \tag{23}$$

For the general SDE (1) we have the following result.

Lemma 4 *Assume that the coefficients f, $g_r, r = 1,\dots,m$ of the SDE (1) belong to the class $C^{2,1}$ with $\Lambda_0 f, \Lambda_0 g_r, \Lambda_r f, \Lambda_q g_r \in C^K$ for $r,q = 1,\dots,m$. Then the local error (13) of the stochastic 2-step scheme (21) allows the representation*

$$L_\ell = R_\ell^\circ + S_{1,\ell}^\circ + S_{2,\ell-1}^\circ, \quad \ell = 2,\dots,N, \tag{24}$$

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

$$\begin{aligned} R_\ell^\circ &= \Big[\sum_{j=0}^{2} \alpha_{\ell,j}\Big] X(t_{\ell-2}) \\ &\qquad + \Big[\alpha_{\ell,0} + \frac{1}{\kappa_\ell}(\alpha_{\ell,0} + \alpha_{\ell,1}) - \sum_{j=0}^{2} \beta_{\ell,j}\Big] h_\ell f(X(t_{\ell-2}),t_{\ell-2}) + \tilde{R}_\ell^\circ, \\ S_{1,\ell}^\circ &= \Big[\alpha_{\ell,0} - \gamma_{\ell,1}\Big] \sum_{r=1}^{m} g_r(X(t_{\ell-1}),t_{\ell-1}) I_r^{t_{\ell-1},t_\ell} + \tilde{S}_{1,\ell}^\circ, \end{aligned}$$

$$S^\circ_{2,\ell-1} = \Big[(\alpha_{\ell,0}+\alpha_{\ell,1}) - \gamma_{\ell,2}\Big] \sum_{r=1}^{m} g_r(X(t_{\ell-2}), t_{\ell-2}) I_r^{t_{\ell-2},t_{\ell-1}} + \tilde{S}^\circ_{2,\ell-1}$$

with

$$\|\tilde{R}^\circ_\ell\|_{L_2} = O(h_\ell^2), \quad \|\tilde{S}^\circ_{1,\ell}\|_{L_2} = O(h_\ell), \quad \|\tilde{S}^\circ_{2,\ell-1}\|_{L_2} = O(h_\ell). \tag{25}$$

Corollary 5 *Let the coefficients f, g_r, $r = 1,\dots,m$, of the SDE (1) satisfy the assumptions of Lemma 4 and suppose they are Lipschitz continuous with respect to their first variable. Let the stochastic linear two-step scheme with variable step-size (21) are stable and let the coefficients satisfy the consistency conditions*

$$\sum_{j=0}^{2}\alpha_{\ell,j}=0, \quad \alpha_{\ell,0}+\frac{1}{\kappa_\ell}(\alpha_{\ell,0}+\alpha_{\ell,1})=\sum_{j=0}^{2}\beta_{\ell,j}, \quad \alpha_{\ell,0}=\gamma_{\ell,1}, \quad \alpha_{\ell,0}+\alpha_{\ell,1}=\gamma_{\ell,2}. \tag{26}$$

Then the global error of the scheme (21) applied to (1) allows the expansion

$$\max_{\ell=0,\dots,N} \|X(t_\ell) - X_\ell\|_{L_2} = O(\mathbf{h}^{1/2}) + O(\max_{\ell=0,1} \|X(t_\ell) - X_\ell\|_{L_2})$$

where $\mathbf{h} := \max_{\ell=2,\dots,N} h_\ell$.

PROOF. *(of Corollary 5)* By Lemma 4 we have the representation (24) for the local error. Applying the consistency conditions (26) yields

$$R^\circ_\ell = \tilde{R}^\circ_\ell, \quad S^\circ_{1,\ell} = \tilde{S}^\circ_{1,\ell}, \quad S^\circ_{2,\ell-1} = \tilde{S}^\circ_{2,\ell-1}, \quad \ell = 2,\dots,N.$$

As the scheme (21) satisfies the conditions of Theorem 2, it is numerically stable in the mean-square sense. Now the assertion follows from the estimates (25) by means of the stability inequality.

PROOF. *(of Lemma 4)* To derive a representation of the local error in the form (24) 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

$$\sum_{j=0}^{2} \alpha_{\ell,j} X(t_{\ell-j}) = \alpha_{\ell,0}\Big(X(t_\ell) - X(t_{\ell-1})\Big) + (\alpha_{\ell,0}+\alpha_{\ell,1})\Big(X(t_{\ell-1}) - X(t_{\ell-2})\Big) + \Big(\sum_{j=0}^{2} \alpha_{\ell,j}\Big) X(t_{\ell-2}),$$

we can express the local error (13) as

$$L_\ell = \alpha_{\ell,0}\Big(X(t_\ell) - X(t_{\ell-1})\Big) + (\alpha_{\ell,0} + \alpha_{\ell,1})\Big(X(t_{\ell-1}) - X(t_{\ell-2})\Big) + \sum_{j=0}^{2} \alpha_{\ell,j} X(t_{\ell-2})$$
$$-h_\ell \sum_{j=0}^{2} \beta_{\ell,j} f(X(t_{\ell-j}), t_{\ell-j}) - \sum_{j=1}^{2} \gamma_{\ell,j} \sum_{r=1}^{m} g_r(X_{\ell-j}, t_{\ell-j})\ I_r^{t_{\ell-j}, t_{\ell-j+1}},$$

The SDE (1) implies the identities

$$X(t_{\ell-1}) - X(t_{\ell-2}) = \int_{t_{\ell-2}}^{t_{\ell-1}} f(X(s), s)\mathrm{d}s + \sum_{r=1}^{m} \int_{t_{\ell-2}}^{t_{\ell-1}} g_r(X(s), s)\mathrm{d}W_r(s)$$
$$= h_{\ell-1} f(X(t_{\ell-2}), t_{\ell-2}) + I_{00}^{t_{\ell-2}, t_{\ell-1}}(\Lambda_0 f) + \sum_{r=1}^{m} I_{r0}^{t_{\ell-2} t_{\ell-1}}(\Lambda_r f)$$
$$+ \sum_{r=1}^{m} g_r(X(t_{\ell-2}), t_{\ell-2}) I_r^{t_{\ell-2}, t_{\ell-1}} + \sum_{r=1}^{m} I_{0r}^{t_{\ell-2}, t_{\ell-1}}(\Lambda_0 g_r) + \sum_{r,q=1}^{m} I_{qr}^{t_{\ell-2}, t_{\ell-1}}(\Lambda_q g_r),$$

and, additionally using (22),

$$X(t_\ell) - X(t_{\ell-1}) = \int_{t_{\ell-1}}^{t_\ell} f(X(s), s)\mathrm{d}s + \sum_{r=1}^{m} \int_{t_{\ell-1}}^{t_\ell} g_r(X(s), s)\mathrm{d}W_r(s)$$
$$= h_\ell \Big\{ f(X(t_{\ell-2}), t_{\ell-2}) + I_0^{t_{\ell-2}, t_{\ell-1}}(\Lambda_0 f) + \sum_{r=1}^{m} I_r^{t_{\ell-2}, t_{\ell-1}}(\Lambda_r f) \Big\}$$
$$+ I_{00}^{t_{\ell-1}, t_\ell}(\Lambda_0 f) + \sum_{r=1}^{m} I_{r0}^{t_{\ell-1} t_\ell}(\Lambda_r f)$$
$$+ \sum_{r=1}^{m} g_r(X(t_{\ell-1}), t_{\ell-1}) I_r^{t_{\ell-1}, t_\ell} + \sum_{r=1}^{m} I_{0r}^{t_{\ell-1}, t_\ell}(\Lambda_0 g_r) + \sum_{r,q=1}^{m} I_{qr}^{t_{\ell-1}, t_\ell}(\Lambda_q g_r).$$

Inserting this and the expansions (22), (23) into the local error formula and reordering the terms, yields

$$L_\ell = \Big[\sum_{j=0}^{2} \alpha_{\ell,j}\Big] X(t_{\ell-2}) + \Big[h_\ell \alpha_{\ell,0} + h_{\ell-1}(\alpha_{\ell,0} + \alpha_{\ell,1}) - h_\ell \sum_{j=0}^{2} \beta_{\ell,j}\Big] f(X(t_{\ell-2}), t_{\ell-2})$$
$$+\ \tilde{R}_\ell^\circ + \Big[\alpha_{\ell,0} - \gamma_{\ell,1}\Big] \sum_{r=1}^{m} g_r(X(t_{\ell-1}), t_{\ell-1}) I_r^{t_{\ell-1}, t_\ell} + \tilde{S}_{1,\ell}^\circ$$
$$+ \Big[(\alpha_{\ell,0} + \alpha_{\ell,1}) - \gamma_{\ell,2}\Big] \sum_{r=1}^{m} g_r(X(t_{\ell-2}), t_{\ell-2}) I_r^{t_{\ell-2}, t_{\ell-1}} + \tilde{S}_{2,\ell-1}^\circ,$$

where

$$\tilde{R}_\ell^\circ = \alpha_{\ell,0}\Big\{h_\ell I_0^{t_{\ell-2}, t_{\ell-1}}(\Lambda_0 f) + I_{00}^{t_{\ell-1}, t_\ell}(\Lambda_0 f)\Big\} + (\alpha_{\ell,0} + \alpha_{\ell,1}) I_{00}^{t_{\ell-2}, t_{\ell-1}}(\Lambda_0 f)$$
$$-h_\ell \beta_{\ell,0}\Big\{I_0^{t_{\ell-2}, t_{\ell-1}}(\Lambda_0 f) + I_0^{t_{\ell-1}, t_\ell}(\Lambda_0 f)\Big\} - h_\ell \beta_{\ell,1} I_0^{t_{\ell-2}, t_{\ell-1}}(\Lambda_0 f), \quad (27)$$

$$\tilde{S}^{\circ}_{1,\ell} = \sum_{r=1}^{m} \Big(\alpha_{\ell,0} I_{r0}^{t_{\ell-1},t_\ell}(\Lambda_r f) - h_\ell \beta_{\ell,0} I_{r}^{t_{\ell-1},t_\ell}(\Lambda_r f) \Big) + \alpha_{\ell,0} \sum_{r=1}^{m} I_{0r}^{t_{\ell-1},t_\ell}(\Lambda_0 g_r)$$

$$+ \alpha_{\ell,0} \sum_{r,q=1}^{m} I_{qr}^{t_{\ell-1},t_\ell}(\Lambda_q g_r)), \tag{28}$$

$$\tilde{S}^{\circ}_{2,\ell-1} = h_\ell(\alpha_{\ell,0} - \beta_{\ell,0} - \beta_{\ell,1}) \sum_{r=1}^{m} I_{r}^{t_{\ell-2},t_{\ell-1}}(\Lambda_r f) + (\alpha_{\ell,0} + \alpha_{\ell,1}) \sum_{r=1}^{m} I_{r0}^{t_{\ell-2},t_{\ell-1}}(\Lambda_r f)$$

$$+ (\alpha_{\ell,0} + \alpha_{\ell,1}) \sum_{r=1}^{m} I_{0r}^{t_{\ell-2},t_{\ell-1}}(\Lambda_0 g_r) + (\alpha_{\ell,0} + \alpha_{\ell,1}) \sum_{r,q=1}^{m} I_{qr}^{t_{\ell-2},t_{\ell-1}}(\Lambda_q g_r). \tag{29}$$

Finally, the estimates (25) are derived by means of (5) and (6), where the last terms in (28) and (29) determine the order $O(h_\ell)$.

Example 6 As examples we give stochastic variants of the trapezoidal rule, the two-step Adams-Bashforth (AB) and the two-step backward differential formulae (BDF_2), which we term the BDF_2-Maruyama method, with variable step-sizes. The trapezoidal rule, also known as stochastic Theta method with $\theta = \frac{1}{2}$, is the one-step scheme with the coefficients $\alpha_{\ell,0} = 1$, $\alpha_{\ell,1} = -1$, $\beta_{\ell,0} = \beta_{\ell,1} = \frac{1}{2}$, $\gamma_{\ell,1} = 1$, $\alpha_{\ell,2} = \beta_{\ell,2} = \gamma_{\ell,2} = 0$ independent of the step-size ratio $\kappa_\ell = h_\ell / h_{\ell-1}$ and reads

$$X_\ell - X_{\ell-1} = h_\ell \frac{1}{2} \Big(f(X_\ell, t_\ell) + f(X_{\ell-1}, t_{\ell-1}) \Big) + \sum_{r=1}^{m} g_r(X_{\ell-1}, t_{\ell-1}) \; I_r^{t_{\ell-1},t_\ell}. \tag{30}$$

The Adams-Bashforth scheme is given as

$$X_\ell - X_{\ell-1} = h_\ell \left(\frac{\kappa_\ell + 2}{2} f(X_{\ell-1}, t_{\ell-1}) - \frac{\kappa_\ell}{2} f(X_{\ell-2}, t_{\ell-2}) \right)$$

$$+ \sum_{r=1}^{m} g_r(X_{\ell-1}, t_{\ell-1}) \; I_r^{t_{\ell-1},t_\ell}, \tag{31}$$

where $\alpha_{\ell,0} = 1$, $\alpha_{\ell,1} = -1$, $\beta_{\ell,0} = \frac{\kappa_\ell + 2}{2}$, $\beta_{\ell,1} = -\frac{\kappa_\ell}{2}$, $\gamma_{\ell,1} = 1$ and $\beta_{\ell,2} = \alpha_{\ell,2} = \gamma_{\ell,2} = 0$. The implicit BDF_2-Maruyama method takes the form

$$X_\ell - \frac{(\kappa_\ell + 1)^2}{2\kappa_\ell + 1} X_{\ell-1} + \frac{\kappa_\ell^2}{2\kappa_\ell + 1} X_{\ell-2} = h_\ell \; \frac{\kappa_\ell + 1}{2\kappa_\ell + 1} f(X_\ell, t_\ell)$$

$$+ \sum_{r=1}^{m} g_r(X_{\ell-1}, t_{\ell-1}) \; I_r^{t_{\ell-1},t_\ell} - \frac{\kappa_\ell^2}{2\kappa_\ell + 1} \sum_{r=1}^{m} g_r(X_{\ell-2}, t_{\ell-2}) \; I_r^{t_{\ell-2},t_{\ell-1}}. \tag{32}$$

Here one has $\alpha_{\ell,0} = 1$, $\alpha_{\ell,1} = -\frac{(\kappa_\ell+1)^2}{2\kappa_\ell+1}$, $\alpha_{\ell,2} = \frac{\kappa_\ell^2}{2\kappa_\ell+1}$, $\beta_{\ell,0} = \frac{\kappa_\ell+1}{2\kappa_\ell+1}$, $\beta_{\ell,1} = \beta_{\ell,2} = 0$, and $\gamma_{\ell,1} = 1$, $\gamma_{\ell,2} = -\frac{\kappa_\ell^2}{2\kappa_\ell+1}$. The BDF_2-Maruyama method is numerically stable (as in the deterministic case) if and only if

$$0 < \kappa_\ell < 1 + \sqrt{2} \qquad \text{for } \ell \geq 2.$$

Before concluding, we discuss the special case of small noise SDEs (2). To be able to exploit the effect of the small parameter ϵ in the expansions of the local error we introduce operators $\Lambda_0^f, \hat{\Lambda}_0$ and $\hat{\Lambda}_r,\ r=1,\ldots,m$ defined on $C^{2,1}$ and $C^{1,0}$, respectively, by

$$\Lambda_0^f y := y_t' + y_x' f, \quad \hat{\Lambda}_0 y := \frac{1}{2}\sum_{r=1}^{m}\sum_{i,j=1}^{n} y''_{x_i x_j}\hat{g}_{ri}\hat{g}_{rj}, \quad \hat{\Lambda}_r y := y_x'\hat{g}_r\,. \tag{33}$$

In terms of the original definition (19) we have

$$\Lambda_0 y = \Lambda_0^f y + \epsilon^2\hat{\Lambda}_0 y \quad \text{ and } \quad \Lambda_r y = \epsilon\hat{\Lambda}_r y. \tag{34}$$

Starting from the expression of Lemma 4 we will analyse the local error by expanding the term $\Lambda_0 f$ appearing in $\tilde{R}_\ell^\circ$ (27) and we show the potential of two-step Maruyama schemes for the special case of small noise SDEs.

Lemma 7 *Assume that the coefficients $f, \hat{g}_r, r=1,\ldots,m$ of the small noise SDE (2), as well as $\Lambda_0^f f = f_x' f + f_t'$ belong to the class $C^{2,1}$ with $\Lambda_0 f, \Lambda_0\hat{g}_r, \hat{\Lambda}_r f$, $\hat{\Lambda}_q\hat{g}_r, \Lambda_0\Lambda_0^f f$, $\hat{\Lambda}_r\Lambda_0^f f \in C^K$ for $r,q=1,\ldots,m$. Let the stochastic 2-step scheme with variable step-size (21) satisfy the consistency conditions (26). Then the local error (13) of the method (21) for the small noise SDE (2) allows the representation*

$$L_\ell = R_\ell^\diamond + S_{1,\ell}^\diamond + S_{2,\ell-1}^\diamond, \quad \ell = 2,\ldots,N, \tag{35}$$

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

$$\begin{aligned} R_\ell^\diamond &= \Big[(\frac{1}{\kappa_\ell^2}+\frac{2}{\kappa_\ell}+1)\alpha_{\ell,0} + \frac{1}{\kappa_\ell^2}\alpha_{\ell,1} - (\frac{2}{\kappa_\ell}+2)\beta_{\ell,0} - \frac{2}{\kappa_\ell}\beta_{\ell,1}\Big]\cdot \\ &\qquad\qquad\qquad\qquad \cdot\frac{h_\ell^2}{2}(\Lambda_0^f f)(X(t_{\ell-2}),t_{\ell-2}) + \tilde{R}_\ell^\diamond, \\ S_{1,\ell}^\diamond &= \tilde{S}_{1,\ell}^\circ + \tilde{S}_{1,\ell}^\diamond, \\ S_{2,\ell-1}^\diamond &= \tilde{S}_{2,\ell-1}^\circ + \tilde{S}_{2,\ell-1}^\diamond, \end{aligned}$$

where

$$\|\tilde{R}_\ell^\diamond\|_{L_2} = O(h_\ell^3+\epsilon^2 h_\ell^2), \quad \|\tilde{S}_{1,\ell}^\diamond\|_{L_2} = O(\epsilon h_\ell^{5/2}), \quad \|\tilde{S}_{2,\ell-1}^\diamond\|_{L_2} = O(\epsilon h_\ell^{5/2}). \tag{36}$$

The terms $\tilde{S}_{1,\ell}^\circ, \tilde{S}_{2,\ell-1}^\circ$ are given by (28, 29) in the proof of Lemma 4 and satisfy here

$$\|\tilde{S}_{1,\ell}^\circ\|_{L_2} = O(\epsilon^2 h_\ell + \epsilon h_\ell^{3/2}), \quad \|\tilde{S}_{2,\ell}^\circ\|_{L_2} = O(\epsilon^2 h_\ell + \epsilon h_\ell^{3/2}). \tag{37}$$

PROOF. We have from Lemma 4, if the consistency conditions (26) are satisfied, the representation

$$L_\ell = \tilde{R}^\circ_\ell + \tilde{S}^\circ_{1,\ell} + \tilde{S}^\circ_{2,\ell-1},\ \ell = 2,\dots,N,$$

where $\tilde{R}^\circ_\ell,\ \tilde{S}^\circ_{1,\ell},\ \tilde{S}^\circ_{2,\ell-1}$ are given by (27, 28, 29). Splitting $\Lambda_0 f = \Lambda_0^f f + \epsilon^2\hat{\Lambda}_0 f$ immediately yields $\tilde{R}^\circ_\ell = \tilde{R}^{\circ f}_\ell + \epsilon^2\hat{R}^\circ_\ell$ with

$$\begin{aligned}\tilde{R}^{\circ f}_\ell := &(\alpha_{\ell,0}-\beta_{\ell,0}-\beta_{\ell,1})h_\ell I_0^{t_{\ell-2},t_{\ell-1}}(\Lambda_0^f f) + (\alpha_{\ell,0}+\alpha_{\ell,1})I_{00}^{t_{\ell-2},t_{\ell-1}}(\Lambda_0^f f)\\ &+\alpha_{\ell,0}I_{00}^{t_{\ell-1},t_\ell}(\Lambda_0^f f) - h_\ell\beta_{\ell,0}I_0^{t_{\ell-1},t_\ell}(\Lambda_0^f f)\end{aligned} \tag{38}$$

$$\begin{aligned}\hat{R}^\circ_\ell := &(\alpha_{\ell,0}-\beta_{\ell,0}-\beta_{\ell,1})h_\ell I_0^{t_{\ell-2},t_{\ell-1}}(\hat{\Lambda}_0 f) + (\alpha_{\ell,0}+\alpha_{\ell,1})I_{00}^{t_{\ell-2},t_{\ell-1}}(\hat{\Lambda}_0 f)\\ &+\alpha_{\ell,0}I_{00}^{t_{\ell-1},t_\ell}(\hat{\Lambda}_0 f) - h_\ell\beta_{\ell,0}I_0^{t_{\ell-1},t_\ell}(\hat{\Lambda}_0 f).\end{aligned} \tag{39}$$

We note that (39) appears with the factor ϵ^2 in the local error representation, thus yielding the $O(\epsilon^2 h_\ell^2)$ term in the estimate of $\|\tilde{R}^\diamond_\ell\|_{L_2}$ in (36) . We concentrate on developing $\tilde{R}^{\circ f}_\ell$ in more detail. Applying the Itô-formula (20) to $\Lambda_0^f f(X(s),s)$ for $s\in[t_{\ell-2},t_{\ell-1}]$ and integrating yields

$$I_0^{t_{\ell-2},s}(\Lambda_0^f f) = (s-t_{\ell-2})\Lambda_0^f f(X(t_{\ell-2}),t_{\ell-2}) + I_{00}^{t_{\ell-2},s}(\Lambda_0\Lambda_0^f f) + \epsilon\sum_{r=1}^m I_{r0}^{t_{\ell-2},s}(\hat{\Lambda}_r\Lambda_0^f f).$$

For $s = t_{\ell-1}$ we obtain

$$I_0^{t_{\ell-2},t_{\ell-1}}(\Lambda_0^f f) = h_{\ell-1}\Lambda_0^f f(X(t_{\ell-2}),t_{\ell-2}) + I_{00}^{t_{\ell-2},t_{\ell-1}}(\Lambda_0\Lambda_0^f f) + \epsilon\sum_{r=1}^m I_{r0}^{t_{\ell-2},t_{\ell-1}}(\hat{\Lambda}_r\Lambda_0^f f).$$

for the first integral in (38) . Integrating again we obtain for the second integral in (38)

$$I_{00}^{t_{\ell-2},t_{\ell-1}}(\Lambda_0^f f) = \frac{h_{\ell-1}^2}{2}\Lambda_0^f f(X(t_{\ell-2}),t_{\ell-2}) + I_{000}^{t_{\ell-2},t_{\ell-1}}(\Lambda_0\Lambda_0^f f) + \epsilon\sum_{r=1}^m I_{r00}^{t_{\ell-2},t_{\ell-1}}(\hat{\Lambda}_r\Lambda_0^f f).$$

Both the other integrals are over the interval $[t_{\ell-1},t_\ell]$ with step-size h_ℓ. In the analogous expressions for these the term $\Lambda_0^f f(X(t_{\ell-1}),t_{\ell-1})$ has to be substituted by

$$\Lambda_0^f f(X(t_{\ell-1}),t_{\ell-1}) = \Lambda_0^f f(X(t_{\ell-2}),t_{\ell-2}) + I_0^{t_{\ell-2},t_{\ell-1}}(\Lambda_0\Lambda_0^f f) + \epsilon\sum_{r=1}^m I_r^{t_{\ell-2},t_{\ell-1}}(\Lambda_r\Lambda_0^f f).$$

Then we obtain from (38)

$$\tilde{R}_\ell^{\circ f} = \Big[(h_\ell h_{\ell-1} + \frac{h_{\ell-1}^2}{2} + \frac{h_\ell^2}{2})\alpha_{\ell,0} + \frac{h_{\ell-1}^2}{2}\alpha_{\ell,1} - (h_\ell h_{\ell-1} - h_\ell^2)\beta_{\ell,0} - h_\ell h_{\ell-1}\beta_{\ell,1}\Big] \cdot$$
$$\cdot\, \Lambda_0^f f(X(t_{\ell-2}), t_{\ell-2}) + \tilde{R}_\ell^{\diamond f} + \tilde{S}_{1,\ell}^{\diamond} + \tilde{S}_{2,\ell}^{\diamond}$$
$$= \Big[(\frac{1}{\kappa_\ell^2} + \frac{2}{\kappa_\ell} + 1)\alpha_{\ell,0} + \frac{1}{\kappa_\ell^2}\alpha_{\ell,1} - (\frac{2}{\kappa_\ell} + 2)\beta_{\ell,0} - \frac{2}{\kappa_\ell}\beta_{\ell,1}\Big]\frac{h_\ell^2}{2}\Lambda_0^f f(X(t_{\ell-2}), t_{\ell-2})$$
$$+ \tilde{R}_\ell^{\diamond f} + \tilde{S}_{1,\ell}^{\diamond} + \tilde{S}_{2,\ell}^{\diamond},$$

where

$$\tilde{R}_\ell^{\diamond f} = (\alpha_{\ell,0} - 2\beta_{\ell,0})\frac{h_\ell^2}{2} I_0^{t_{\ell-2},t_{\ell-1}}(\Lambda_0\Lambda_0^f f)$$
$$+ (\alpha_{\ell,0} - \beta_{\ell,0} - \beta_{\ell,1})h_\ell I_{00}^{t_{\ell-2},t_{\ell-1}}(\Lambda_0\Lambda_0^f f) - \beta_{\ell,0} h_\ell I_{00}^{t_{\ell-1},t_\ell}(\Lambda_0\Lambda_0^f f)$$
$$+ (\alpha_{\ell,0} + \alpha_{\ell,1}) I_{000}^{t_{\ell-2},t_{\ell-1}}(\Lambda_0\Lambda_0^f f) + \alpha_{\ell,0} I_{000}^{t_{\ell-1},t_\ell}(\Lambda_0\Lambda_0^f f),$$

$$\tilde{S}_{1,\ell}^{\diamond} = \alpha_{\ell,0}\epsilon \sum_{r=1}^m I_{r00}^{t_{\ell-1},t_\ell}(\hat{\Lambda}_r\Lambda_0^f f) - h_\ell\beta_{\ell,0}\epsilon\sum_{r=1}^m I_{r0}^{t_{\ell-1},t_\ell}(\hat{\Lambda}_r\Lambda_0^f f),$$
$$\tilde{S}_{2,\ell}^{\diamond} = (\alpha_{\ell,0} - 2\beta_{\ell,0})\frac{h_\ell^2}{2}\epsilon\sum_{r=1}^m I_r^{t_{\ell-2},t_{\ell-1}}(\hat{\Lambda}_r\Lambda_0^f f) + (\alpha_{\ell,0} - \beta_{\ell,0} - \beta_{\ell,1}) \cdot$$
$$\cdot\, h_\ell\, \epsilon\sum_{r=1}^m I_{r0}^{t_{\ell-2},t_{\ell-1}}(\hat{\Lambda}_r\Lambda_0^f f) + (\alpha_{\ell,0} + \alpha_{\ell,1})\epsilon\sum_{r=1}^m I_{r00}^{t_{\ell-2},t_{\ell-1}}(\hat{\Lambda}_r\Lambda_0^f f).$$

We arrive at $\tilde{R}_\ell^{\diamond} = \tilde{R}_\ell^{\diamond f} + \epsilon^2\hat{R}_\ell^{\circ}$. Finally, the estimates (36) are derived by means of (5) and (6).

Corollary 8 *Let the coefficients f, $\hat{g}_r$, $r = 1, \ldots, m$, of the SDE (2) satisfy the assumptions of Lemma 7 and suppose they are Lipschitz continuous with respect to their first variable. Let the stochastic linear two-step scheme with variable step-size (21) are stable, the coefficients satisfy the consistency conditions (26) and*

$$\left(\frac{1}{\kappa_\ell^2} + \frac{2}{\kappa_\ell} + 1\right)\alpha_{\ell,0} + \frac{1}{\kappa_\ell^2}\,\alpha_{\ell,1} - \left(\frac{2}{\kappa_\ell} + 2\right)\beta_{\ell,0} - \frac{2}{\kappa_\ell}\,\beta_{\ell,1} = 0. \qquad (40)$$

Then the global error of the scheme (21) applied to (2) allows the expansion

$$\max_{\ell=0,\ldots,N}\|X(t_\ell) - X_\ell\|_{L_2} = O(\mathbf{h}^2 + \epsilon\mathbf{h} + \epsilon^2\mathbf{h}^{1/2}) + O(\max_{\ell=0,1}\|X(t_\ell) - X_\ell\|_{L_2}).$$

PROOF. Lemma 7 stated the representation (35) for the local error. Applying the consistency condition (40) yields $R_\ell^{\diamond} = \tilde{R}_\ell^{\diamond}$ and by (36) we have $\|R_\ell^{\diamond}\|_{L_2} = O(h_\ell^3 + \epsilon^2 h_\ell^2)$. The stochastic terms $S_{1,\ell}^{\diamond}$, $S_{2,\ell-1}^{\diamond}$ are dominated by $\tilde{S}_{1,\ell}^{\circ}$, $\tilde{S}_{2,\ell-1}^{\circ}$ and thus are of order of magnitude $O(\epsilon^2 h_\ell + \epsilon h_\ell^{3/2})$. As the scheme

(21) satisfies the conditions (A1) - (A4), it is numerically stable in the mean-square sense. Applying the stability inequality (12) to the representation (35) of the local error yields the assertion.

We remark that the schemes (30), (31) and (32) satisfy the assumptions and the consistency conditions of corollary 8. Thus, these schemes are numerically stable in the mean-square sense and we can expect order 2 behaviour if the term $O(\mathbf{h}^2)$ of the global error dominates the term $O(\epsilon\mathbf{h} + \epsilon^2\mathbf{h}^{1/2})$. A first simulation result that illustrate that performance is given in the next section. If (40) is not fulfilled, the global error allows the expansion

$$\max_{\ell=0,\dots,N} \|X(t_\ell) - X_\ell\|_{L_2} = O(\mathbf{h} + \epsilon^2\mathbf{h}^{1/2}) + O(\max_{\ell=0,1} \|X(t_\ell) - X_\ell\|_{L_2}) .$$

which is similar to the error behaviour of the one-step Euler-Maruyama scheme.

5 Numerical experiments

Here, we illustrate the error behaviour by simulation results for the stochastic BDF_2 applied to a test problem. We consider a nonlinear SDE, where the exact solution is known such that we can access the actual errors. We present the tolerances and accuracies versus steps to show the order 2 behavior if the noise is small. Some first results in the context of circuit simulation can be found in [12].

Example 9 (Test-SDE with polynomial drift and diffusion.) *We consider a nonlinear scalar SDE with known solution and drift- and diffusion coefficients that are tunable by real parameters α, β, which we have chosen as $\alpha = -10$ and $\beta = 0.01$:*

$$X(t) = \int_0^t -(\alpha + \beta^2 X(s))(1 - X(s)^2)ds + \int_0^t \beta(1 - X(s)^2)dW(s), \quad t \in [0,1], \tag{41}$$

where W denotes a scalar Wiener process. The solution is given by

$$X(t) = \frac{\exp(-2\alpha t + 2\beta W(t)) - 1}{\exp(-2\alpha t + 2\beta W(t)) + 1} . \tag{42}$$

In Figure 1 we present a work-precision diagram. We plotted the tolerance ($\triangle$) and the mean-square norm of the errors for adaptively chosen (+) and constant ($\times$) step-sizes for 100 computed paths vs. the number of steps in logarithmic scale. Lines with slopes -2 and -0.5 are provided to enable comparisons with convergence of order 2 or 1/2. The accuracy is measured as the maximum *approximate L_2-norm of the global errors in the time-interval:*

$$\max_{\ell=1,\ldots,N} \left(\frac{1}{M} \sum_{j=1}^{M} |X(t_\ell,\omega_j) - X_\ell(\omega_j)|^2 \right)^{1/2} \approx \max_{\ell=1,\ldots,N} \|X(t_\ell) - X_\ell\|_{L_2}$$

where N denotes the number of steps and M the number of computed paths.

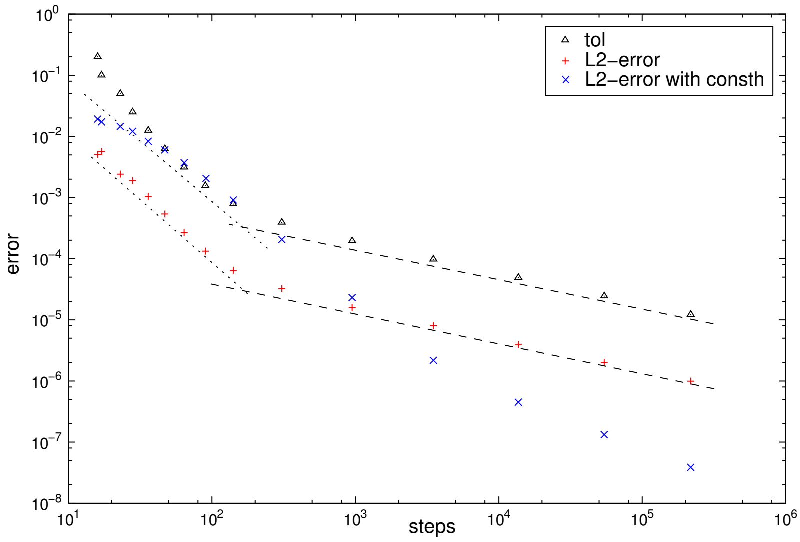

Fig. 1. Tolerance and accuracy versus steps for a test-SDE.

Let us add some observations. In general, all numerical schemes that include only information on the increments of the Wiener process have only an asymptotic order of strong convergence of $1/2$. However, when the noise is small, the error behaviour is much better. The error of the BDF_2-Maruyama scheme appears to be of the size $\max(c_1\mathbf{h}^2, c_2\epsilon\mathbf{h}, c_3\epsilon^2\mathbf{h}^{1/2})$, where c_1 is the error constant of the deterministic BDF_2, the constants c_2, c_3 appear to coincide for the scheme and the small parameter is given by $\epsilon = |\beta/\alpha| = 10^{-3}$.

In fact, we observe order 2 behaviour up to accuracies of 10^{-4}, then there is a very small region where the order of convergence is 1 and thereafter only $1/2$. So the results show that the errors are dominated by the deterministic terms as long as the step-size is large enough.

A Proof of Theorem (2)

For the proof of Theorem (2) we need a discrete version of Gronwall's lemma.

Lemma 10 *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 $\max_{\ell=1,\ldots,N} a_\ell \le C_1 \exp(C_2)$.

PROOF. (of Theorem (2))
Part 1 (Existence of a solution $\tilde{X}_\ell$): We consider scheme (11). If the right hand side does not depend on the variable X_ℓ, the new iterate $\tilde{X}_\ell$ is given explicitly. Otherwise, the new iterate $\tilde{X}_\ell$ is given by (11) only implicitly as the solution of the fixed point equation

$$X = h_\ell \beta_{\ell,0} f(X, t_\ell) + h_\ell \sum_{j=1}^{k} \beta_{\ell,j} f(\tilde{X}_{\ell-j}, t_{\ell-j}) + B_\ell =: \eta_\ell(X; \tilde{X}_{\ell-1}, \ldots, \tilde{X}_{\ell-k}, B_\ell),$$

$$\text{where} \qquad B_\ell = -\sum_{j=1}^{k} \alpha_{\ell,j}\, \tilde{X}_{\ell-j} + \sum_{j=1}^{k} \Gamma_{\ell,j}(\tilde{X}_{\ell-j}, t_{\ell-j}) I^{t_{\ell-j}, t_{\ell-j+1}} + D_\ell.$$

is a known F_{t_ℓ}-measurable random variable. The function $\eta_\ell(x; z_1, \ldots, z_k, b)$ is globally contractive with respect to x, since, due to the global Lipschitz condition (A1),

$$\begin{aligned} |\eta_\ell(x; z_1, \ldots, z_k, b) - \eta_\ell(\tilde{x}, z_1, \ldots, z_k, b)| &= |h_\ell\, \beta_{\ell,0}\, (f(x, t_\ell) - f(\tilde{x}, t_\ell))| \\ &\le h_\ell\, \beta_{\ell,0}\, L_f |x - \tilde{x}| \le \frac{1}{2}|x - \tilde{x}| \qquad \forall h_\ell \le \mathbf{h} \le h^0 \le \frac{1}{2\, \beta_{\ell,0}\, L_f} \end{aligned}$$

Thus, $\eta_\ell(\cdot; z_1, \ldots, z_k, b)$ has a globally unique fixed point $x = \xi_\ell(z_1, \ldots, z_k, b)$, and $\xi_\ell(\tilde{X}_{\ell-1}, \ldots, \tilde{X}_{\ell-k}, B_\ell)$ gives the unique solution $\tilde{X}_\ell$ of (11). Moreover, ξ_ℓ depends Lipschitz-continuously on $z_1, \ldots, z_k$ and b since

$$\begin{aligned} &|\xi_\ell(z_1, \ldots, z_k, b) - \xi_\ell(\tilde{z}_1, \ldots, \tilde{z}_k, \tilde{b})| \\ &\quad = |\eta_\ell(\xi_\ell(z_1, \ldots, z_k, b); z_1, \ldots, z_k, b) - \eta_\ell(\xi_\ell(\tilde{z}_1, \ldots, \tilde{z}_k, \tilde{b}), \tilde{z}_1, \ldots, \tilde{z}_k, \tilde{b})| \\ &\quad \le h_\ell L_f \sum_{j=1}^{k} \beta_{\ell,j} |z_j - \tilde{z}_j| + h_\ell \beta_{\ell,0} L_f |\xi_\ell(z_1, \ldots, z_k, b) - \xi_\ell(\tilde{z}_1, \ldots, \tilde{z}_k, \tilde{b})| + |b - \tilde{b}| \\ &\quad \le \mathbf{h}\, \beta_{\ell,*}\, L_f \sum_{j=1}^{k} |z_j - \tilde{z}_j| + \frac{1}{2} |\xi_\ell(z_1, \ldots, z_k, b) - \xi_\ell(\tilde{z}_1, \ldots, \tilde{z}_k, \tilde{b})| + |b - \tilde{b}| \\ &|\xi_\ell(z_1, \ldots, z_k, b) - \xi_\ell(\tilde{z}_1, \ldots, \tilde{z}_k, \tilde{b})| \\ &\quad \le 2\, \mathbf{h}\, \beta_{\ell,*}\, L_f \sum_{j=1}^{k} |z_j - \tilde{z}_j| + 2|b - \tilde{b}|, \qquad \text{where } \beta_{\ell,*} := \max_{j=1,\ldots,k} \beta_{\ell,j}. \end{aligned}$$

Part 2 (Existence of finite second moments $\mathbb{E}|\tilde{X}_\ell|^2 < \infty$): Assume that $\mathbb{E}|\tilde{X}_{\ell-j}|^2 < \infty$ for $j = 1, \ldots, k$. We compare $\tilde{X}_\ell = \xi_\ell(\tilde{X}_{\ell-1}, \ldots, \tilde{X}_{\ell-k}, B_\ell)$ with the deterministic value $X_\ell^0 := \xi_\ell(0,\dots,0,0)$. Using the Lipschitz continuity of the implicit function ξ_ℓ we obtain

$$|\tilde{X}_\ell - X_\ell^0| = |\xi_\ell(\tilde{X}_{\ell-1},\dots,\tilde{X}_{\ell-k},B_\ell) - \xi_\ell(0,\dots,0,0)| \le 2\mathbf{h}\beta_{\ell,*}L_f\sum_{j=1}^{k}|\tilde{X}_{\ell-j}| + 2|B_\ell|,$$

$$\|\tilde{X}_\ell\|_{L_2} \le \|\tilde{X}_\ell - X_\ell^0\|_{L_2} + \|X_\ell^0\|_{L_2} \le 2\ \mathbf{h}\ \beta_{\ell,*}\ L_f\sum_{j=1}^{k}\|\tilde{X}_{\ell-j}\|_{L_2} + 2\|B_\ell\|_{L_2} + \|X_\ell^0\|_{L_2}$$

It remains to show that $\|B_\ell\|_{L_2} < \infty$, which follows from

$$\begin{aligned}
&\|\sum_{j=1}^{k}\Gamma_{\ell,j}(\tilde{X}_{\ell-j},t_{\ell-j})I^{t_{\ell-j},t_{\ell-j+1}}\|_{L_2}\\
&\qquad \le h_\ell^{1/2}\sum_{j=1}^{k}L_{\Gamma_{\ell,j}}\|\tilde{X}_{\ell-j}\|_{L_2} + \|\sum_{j=1}^{k}\Gamma_{\ell,j}(0,t_{\ell-j})I^{t_{\ell-j},t_{\ell-j+1}}\|_{L_2} < \infty
\end{aligned}$$

Part 3 (Stability inequality): 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. [2][5, Chap.III.4][11, Chap.8.2.1]).

For X_ℓ and $\tilde{X}_\ell$ being the solutions of (3) and (11), respectively, let the n-dimensional vector E_ℓ be defined as the difference $X_\ell - \tilde{X}_\ell$. We have with $E_0,\dots,E_{k-1} \in L_2(\Omega,\mathbb{R}^n)$ for $\ell = k,\dots,N$, the recursion

$$E_\ell = -\sum_{j=1}^{k}\alpha_{\ell,j}\ E_{\ell-j} + \underbrace{h_\ell\ \sum_{j=0}^{k}\beta_{\ell,j}\ \Delta f_{\ell-j}}_{=:\Delta\phi^\ell} + \underbrace{\sum_{j=1}^{k}\Delta\Gamma_{\ell,j}\ I^{t_{\ell-j},t_{\ell-j+1}}}_{=:\Delta\psi^\ell} - D_\ell,$$

where

$$\begin{aligned}
\Delta f_{\ell-j} &:= f(X_{\ell-j},t_{\ell-j}) - f(\tilde{X}_{\ell-j},t_{\ell-j})\\
\Delta\Gamma_{\ell,j} &:= \Gamma_{\ell,j}(X_{\ell-j},t_{\ell-j}) - \Gamma_{\ell,j}(\tilde{X}_{\ell-j},t_{\ell-j}).
\end{aligned}$$

We rearrange this k-step recursion in the space $L_2(\Omega,\mathbb{R}^n)$ to a one-step recursion in $L_2(\Omega,\mathbb{R}^{k\times n})$. Together with the identities $E_{\ell-1} = E_{\ell-1}$, …, $E_{\ell-k+1} = E_{\ell-k+1}$ we obtain

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

or, in compact form

$$\begin{aligned}\mathcal{E}_\ell &= \mathcal{A}_\ell \mathcal{E}_{\ell-1} + \Delta\Phi_\ell + \Delta\Psi_\ell + \mathcal{D}_\ell\ ,\ \ \ell = k,\ldots,N \quad \text{and}\\ \mathcal{E}_{k-1} &= (-D_{k-1}, -D_{k-2}, \ldots, -D_0)^T,\end{aligned}$$

where $\mathcal{E}_\ell \in L_2(\Omega, \mathbb{R}^{k\times n})$, $\ell = k-1,\ldots,N$. 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

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

A crucial point for the subsequent calculations is to find a scalar product inducing a matrix norm such that this norm of the matrix product $\mathcal{A}_{\ell-i}\cdots\mathcal{A}_\ell$ for all ℓ and $i \geq 0$ is less than or equal to 1 (see e.g. [5, Chap.III.4,Lemma 4.4 and Chap.III.5, Theorem 5.5].
In [2] it is shown in detail for constant matrices $\mathcal{A}_j = \mathcal{A}$, that this is possible if the eigenvalues of the Frobenius matrix $\mathcal{A}$ lie inside the unit circle of the complex plane and are simple if their modulus is equal to 1. Assumption (A4) implies that this property holds true for each $\mathcal{A}_j$. Therefor it is necessary that all entries of $\mathcal{A}_j$ are bounded. Due to the fact that the entries only depend on the stepsize ratios, this holds true since $\kappa < h_\ell/h_{\ell-1} < K$ for $\ell \geq 2$. The eigenvalues of the companion matrix $\mathcal{A}$ of the constant step-size formula are the roots of the characteristic polynomial ρ (10) and due to the assumption that Dahlquist's root condition is satisfied they have the required property. Then there exists a non-singular matrix $\mathcal{C}$ with a block-structure like $\mathcal{A}$ such that $\|\mathcal{C}^{-1}\mathcal{A}\mathcal{C}\|_2 \leq 1$, where $\|\cdot\|_2$ denotes the spectral matrix norm that is induced by the Euclidian vector norm in $\mathbb{R}^{k\times n}$. And, by continuity, we have $\|\mathcal{C}^{-1}\mathcal{A}_j\mathcal{C}\|_2 \leq 1$ which implies that $\|\mathcal{C}^{-1}\mathcal{A}_\ell\cdots\mathcal{A}_{\ell-i}\mathcal{C}\|_2 \leq 1$ for all ℓ and $i = k-1,\ldots,\ell$, if $\kappa_\ell,\ldots,\kappa_{\ell-k}$ are sufficiently close to 1.

We can thus choose a scalar product for $\mathcal{X},\mathcal{Y} \in \mathbb{R}^{k\times n}$ as

$$\langle\mathcal{X},\mathcal{Y}\rangle_* := \langle\mathcal{C}^{-1}\mathcal{X},\mathcal{C}^{-1}\mathcal{Y}\rangle_2$$

and then have $|\cdot|_*$ as the induced vector norm on $\mathbb{R}^{k\times n}$ and $\|\cdot\|_*$ as the induced matrix norm with $\|\mathcal{A}_\ell\cdots\mathcal{A}_{\ell-i}\|_* = \|\mathcal{C}^{-1}\mathcal{A}_\ell\cdots\mathcal{A}_{\ell-i}\mathcal{C}\|_2 \leq 1$. We also have

$$\langle \mathcal{X}, \mathcal{Y} \rangle_* = \mathcal{X}^T \mathcal{C}^{-T} \mathcal{C}^{-1} \mathcal{Y} = \mathcal{X}^T \mathcal{C}^* \mathcal{Y} \quad \text{with} \quad \mathcal{C}^* = \mathcal{C}^{-T} \mathcal{C}^{-1} = (c^*_{ij} I_n)_{i,j=1,\ldots,k}.$$

Due to the norm equivalence there are constants $c^*, c_* > 0$ such that

$$|\mathcal{X}|_2^2 \leq c^* |\mathcal{X}|_*^2 \quad \text{and} \quad |\mathcal{X}|_*^2 \leq c_* |\mathcal{X}|_\infty^2 \quad \forall \mathcal{X} \in \mathbb{R}^{k \times n},$$

where $|\mathcal{X}|_2^2 = \sum_{j=1,\ldots,k} |x_j|^2$, $|\mathcal{X}|_\infty = \max_{j=1,\ldots,k} |x_j|$ for $\mathcal{X} = (x_1^T, \ldots, x_k^T)^T$. For the special vectors $\mathcal{X} = (x^T, 0, \ldots, 0)^T$ and $\mathcal{Y} = (y^T, 0, \ldots, 0)^T$ with $\mathcal{X}, \mathcal{Y} \in \mathbb{R}^{k \times n}$ and $x, y \in \mathbb{R}^n$, one has $\langle \mathcal{X}, \mathcal{Y} \rangle_* = c^*_{11} \langle x, y \rangle_2 = c^*_{11} x^T y$, where c^*_{11} is given by the matrix $\mathcal{C}^*$.

We now apply $|.|_*^2$ to estimate $|\mathcal{E}_\ell|_*^2$ and, later, $\mathbb{E}|\mathcal{E}_\ell|_*^2$. We start with

$$|\mathcal{E}_\ell|_*^2 \leq 4 \Big\{ \underbrace{|\Big(\prod_{j=k}^{\ell} \mathcal{A}_j\Big) \mathcal{E}_{k-1}|_*^2}_{1)} + \underbrace{|\sum_{i=k}^{\ell} \Big(\prod_{j=i+1}^{\ell} \mathcal{A}_j\Big) \Delta\Phi_i|_*^2}_{2)} + \underbrace{|\sum_{i=k}^{\ell} \Big(\prod_{j=i+1}^{\ell} \mathcal{A}_j\Big) \Delta\Psi_i|_*^2}_{3)}$$
$$+ \underbrace{|\sum_{i=k}^{\ell} \Big(\prod_{j=i+1}^{\ell} \mathcal{A}_j\Big) \mathcal{D}_i|_*^2}_{4)} \Big\}.$$

For the term labelled 1) we have $|\Big(\prod_{j=k}^{\ell} \mathcal{A}_j\Big) \mathcal{E}_{k-1}|_*^2 \leq |\mathcal{E}_{k-1}|_*^2$, and thus

$$\mathbb{E}|\Big(\prod_{j=k}^{\ell} \mathcal{A}_j\Big) \mathcal{E}_{k-1}|_*^2 \leq \mathbb{E}|\mathcal{E}_{k-1}|_*^2 . \tag{A.1}$$

For the term labelled 2) we have

$$\begin{aligned}
|\sum_{i=k}^{\ell} \Big(\prod_{j=i+1}^{\ell} \mathcal{A}_j\Big) \Delta\Phi_i|_*^2 &\leq (\ell - k + 1) \sum_{i=k}^{\ell} |\Big(\prod_{j=i+1}^{\ell} \mathcal{A}_j\Big) \Delta\Phi_i|_*^2 \;\leq\; N \sum_{i=k}^{\ell} |\Delta\Phi_i|_*^2 \\
&\leq \frac{aT}{\mathbf{h}} c^*_{11} \sum_{i=k}^{\ell} |\Delta\phi^i|^2 \leq \frac{aT}{\mathbf{h}} c^*_{11} \mathbf{h}^2 \sum_{i=k}^{\ell} |\sum_{j=0}^{k} \beta_j \Delta f_{i-j}|^2 \\
&= \mathbf{h} a T c^*_{11} \sum_{i=k}^{\ell} |\sum_{j=0}^{k} \beta_j \Delta f_{i-j}|^2 \leq \mathbf{h}\ a\ T\ c^*_{11}\ (k+1) \sum_{i=k}^{\ell} \sum_{j=0}^{k} |\beta_{i,j}\ \Delta f_{i-j}|^2 \\
&\leq \mathbf{h}\ a\ T\ c^*_{11}\ (k+1)\ L_f^2 \sum_{i=k}^{\ell} \sum_{j=0}^{k} \beta_{i,j}^2\ |E_{i-j}|^2
\end{aligned}$$

$$
\begin{aligned}
&= \mathbf{h}\ a\ T\ c_{11}^*(k{+}1)\ L_f^2 \sum_{i=k}^{\ell} \{\beta_{i,0}^2 |E_i|^2 + \beta_{i,1}^2 |E_{i-1}|^2 + \sum_{j=2}^{k} \beta_{i,j}^2 |E_{i-j}|^2\} \\
&\leq \mathbf{h}\ a\ T\ c_{11}^*\ (k{+}1)\ L_f^2 \Big\{\beta_{\ell,0}^2\ |E_\ell|^2 + \sum_{i=k}^{\ell} \{\beta_{i-1,0}^2\ |E_{i-1}|^2 + \sum_{j=1}^{k} \beta_{i,j}^2\ |E_{i-j}|^2\}\Big\} \\
&\leq \mathbf{h}\ a\ T\ c_{11}^*\ (k{+}1)\ L_f^2 \Big\{\beta_{\ell,0}^2 |E_\ell|^2 + \sum_{i=k}^{\ell} c_\beta^* |\mathcal{E}_{i-1}|_*^2\Big\} \\
&\leq \mathbf{h}\ a\ T\ c_{11}^*\ (k{+}1)\ L_f^2 \Big\{c^* \beta_{\ell,0}^2\ |\mathcal{E}_\ell|_*^2 + C_\beta\ c^* \sum_{i=k-1}^{\ell-1} |\mathcal{E}_i|_*^2\Big\},
\end{aligned}
$$

where $C_\beta = 2 \cdot \max_{j=0,\ldots,k;i=k,\ldots,N} \beta_{i,j}$ and $\mathbf{h} \cdot N \leq a\ T$. Hence,

$$
\mathbb{E}|\sum_{i=k}^{\ell} \Big(\prod_{j=i+1}^{\ell} \mathcal{A}_j\Big) \Delta\Phi_i|_*^2 \leq \mathbf{h}\ a\ T\ c_{11}^*\ (k{+}1)\ L_f^2\ \Big\{c^*\ \beta_{\ell,0}^2\ \mathbb{E}|\mathcal{E}_\ell|_*^2 + C_\beta\ c^* \sum_{i=k-1}^{\ell-1} \mathbb{E}|\mathcal{E}_i|_*^2\Big\}. \tag{A.2}
$$

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

$$
|\sum_{i=k}^{\ell} \Big(\prod_{j=i+1}^{\ell} \mathcal{A}_j\Big) \Delta\Psi_i|_*^2 \quad = \quad |\sum_{i=k}^{\ell} \Big(\prod_{j=i+1}^{\ell} \mathcal{A}_j\Big) \sum_{j=1}^{k} \Delta\Psi_{j,i-j}|_*^2.
$$

Every $\Delta\Psi_{j,i-j}$ is $\mathcal{F}_{t_{i-j+1}}$-measurable and $\mathbb{E}(\Delta\Psi_{j,i-j}|\mathcal{F}_{t_{i-j}}) = 0$. We can 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. The expectation of products of terms from different subintervals vanishes, hence we obtain

$$
\begin{aligned}
&\mathbb{E}|\sum_{i=k}^{\ell} \Big(\prod_{j=i+1}^{\ell} \mathcal{A}_j\Big) \Delta\Psi_i|_*^2 \quad = \quad \mathbb{E}|\sum_{i=k}^{\ell} \Big(\prod_{j=i+1}^{\ell} \mathcal{A}_j\Big) \sum_{j=1}^{k} \Delta\Psi_{j,i-j}|_*^2 \\
&= \mathbb{E}|\Big(\prod_{j=k+2}^{\ell} \mathcal{A}_j\Big) \Delta\Psi_{k,0}|_*^2 \\
&\quad + \Big(\prod_{j=k+2}^{\ell} \mathcal{A}_j\Big) \Delta\Psi_{k,1} + \Big(\prod_{j=k+1}^{\ell} \mathcal{A}_j\Big) \Delta\Psi_{k-1,1}|_*^2 \\
&\quad \vdots \\
&\qquad + \mathbb{E}|\mathcal{A}^{\ell-2k+1} \Delta\Psi_{k,k-1} + \mathcal{A}^{\ell-2k+2} \Delta\Psi_{k-1,k-1} + \ldots + \mathcal{A}^{\ell-k} \Delta\Psi_{1,k-1}|_*^2 \\
&\qquad \vdots
\end{aligned}
$$

$$
\begin{aligned}
&+\mathbb{E}|\mathcal{A}^0\Delta\Psi_{k,\ell-k}+\mathcal{A}^1\Delta\Psi_{k-1,\ell-k}+\ldots+\mathcal{A}^{k-1}\Delta\Psi_{1,\ell-k}|_*^2\\
&\vdots\\
&+\mathbb{E}|\Big(\prod_{j=\ell+1}^{\ell}\mathcal{A}_j\Big)\Delta\Psi_{2,\ell-2}+\Big(\prod_{j=\ell}^{\ell}\mathcal{A}_j\Big)\Delta\Psi_{1,\ell-2}|_*^2\\
&+\mathbb{E}|\Big(\prod_{j=\ell+1}^{\ell}\mathcal{A}_j\Big)\Delta\Psi_{1,\ell-1}|_*^2\\
&\le k\sum_{i=k}^{\ell}\sum_{j=1}^{k}\mathbb{E}|\Delta\Psi_{j,i-j}|_*^2 \;=\; k\;c_{11}^*\sum_{i=k}^{\ell}\sum_{j=1}^{k}\mathbb{E}|\Delta\Gamma_{j,i-j}\;I^{t_{i-j},t_{i-j+1}}|^2\\
&\le k\;c_{11}^*\sum_{i=k}^{\ell}\sum_{j=1}^{k}\mathbb{E}\|\Delta\Gamma_{j,i-j}\|^2\;\mathbb{E}|I^{t_{i-j},t_{i-j+1}}|^2\\
&\le \mathbf{h}\;k\;c_{11}^*\;L_\Gamma^2\sum_{i=k}^{\ell}\sum_{j=1}^{k}\mathbb{E}|E_{i-j}|^2\;\le\;\mathbf{h}\;k\;c_{11}^*\;L_\Gamma^2\;c^*\sum_{i=k}^{\ell}|\mathcal{E}_{i-1}|_*^2.
\end{aligned}
$$

Thus, for the term labelled 3), we obtain

$$
\mathbb{E}|\sum_{i=k}^{\ell}\Big(\prod_{j=i+1}^{\ell}\mathcal{A}_j\Big)\Delta\Psi_i|_*^2\le\;\mathbf{h}\;k\;c_{11}^*\;L_\Gamma^2\;c^*\sum_{i=k-1}^{\ell-1}|\mathcal{E}_i|_*^2. \tag{A.3}
$$

We will, for a shorter notation, deal with the term labelled 4), i. e. the perturbations D_i in $\mathcal{D}_i$, after obtaining an intermediate result. Using (A.1), (A.2) and (A.3) and setting $L_0 := a\;L_f^2\;(k+1)\;c_{11}^*\;T\;c^*\;\beta_0^2$ and $L := L_f^2\;(k+1)\;c_{11}^*\;T\;c_\beta^*\;+\;L_\Gamma^2\;k\;c_{11}^*\;c^*$, we have now arrived at

$$
\mathbb{E}|\mathcal{E}_\ell|_*^2\le 4\Big\{\mathbb{E}|\mathcal{E}_{k-1}|_*^2+\mathbf{h}\;L_0\mathbb{E}|\mathcal{E}_\ell|_*^2+\mathbf{h}\;L\sum_{i=k-1}^{\ell-1}\mathbb{E}|\mathcal{E}_i|_*^2+\mathbb{E}|\sum_{i=k}^{\ell}\mathcal{A}^{\ell-i}\mathcal{D}_i|_*^2\Big\},
$$

$\ell=k,\ldots,N$. If necessary we choose a bound h^0 on the step-size such that $4\cdot\mathbf{h}\cdot L_0<\frac{1}{2}$ holds for all $\mathbf{h}<h^0$ and conclude that

$$
\begin{aligned}
\mathbb{E}|\mathcal{E}_\ell|_*^2&\le 8\Big\{\mathbb{E}|\mathcal{E}_{k-1}|_*^2+\mathbf{h}\;L\sum_{i=k-1}^{\ell-1}\mathbb{E}|\mathcal{E}_i|_*^2+\mathbb{E}|\sum_{i=k}^{\ell}\Big(\prod_{j=i+1}^{\ell}\mathcal{A}_j\Big)\mathcal{D}_i|_*^2\Big\}\\
&=8\;\mathbb{E}|\mathcal{E}_{k-1}|_*^2+8\;\mathbb{E}|\sum_{i=k}^{\ell}\Big(\prod_{j=i+1}^{\ell}\mathcal{A}_j\Big)\mathcal{D}_i|_*^2+8\;\mathbf{h}\;L\sum_{i=k-1}^{\ell-1}\mathbb{E}|\mathcal{E}_i|_*^2\\
&\le 8\;\mathbb{E}|\mathcal{E}_{k-1}|_*^2+8\;\mathbb{E}|\sum_{i=k}^{\ell}\Big(\prod_{j=i+1}^{\ell}\mathcal{A}_j\Big)\mathcal{D}_i|_*^2+8\;L\;\frac{aT}{N}\sum_{i=k-1}^{\ell-1}\mathbb{E}|\mathcal{E}_i|_*^2.
\end{aligned}
$$

We now apply Gronwall's Lemma with $a_\ell:=0,\;\ell=1,\ldots,k-2$ and $a_\ell:=\mathbb{E}|\mathcal{E}_\ell|_*^2,\;\ell=k-1,\ldots,N$, and obtain the intermediate result

$$
\max_{\ell=k-1,\ldots,N}\mathbb{E}|\mathcal{E}_\ell|_*^2\le\hat{S}\Big\{\mathbb{E}|\mathcal{E}_{k-1}|_*^2+\mathbb{E}|\sum_{i=k}^{\ell}\Big(\prod_{j=i+1}^{\ell}\mathcal{A}_j\Big)\mathcal{D}_i|_*^2\Big\}, \tag{A.4}
$$

where $\hat{S} := 8\exp(8LaT)$. It remains to deal with the term labelled 4), 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 $i = k, \dots, N$ and $j = 1, \dots, k$. Then $\mathbb{E}\langle \mathcal{A}^{\ell_1}\mathcal{S}_{j_1,i_1}, \mathcal{A}^{\ell_2}\mathcal{S}_{j_2,i_2}\rangle_* = 0$ for $i_1 \neq i_2$, and by similar computations as above we obtain

$$\begin{aligned}
&\mathbb{E}|\sum_{i=k}^{\ell}\Big(\prod_{j=i+1}^{\ell}\mathcal{A}_j\Big)\mathcal{D}_i|_*^2 = \mathbb{E}|\sum_{i=k}^{\ell}\Big(\prod_{j=i+1}^{\ell}\mathcal{A}_j\Big)(\mathcal{R}_i + \sum_{j=1}^{k}\mathcal{S}_{j,i-j+1})|_*^2 \\
&\le 2\ \mathbb{E}|\sum_{i=k}^{\ell}\Big(\prod_{j=i+1}^{\ell}\mathcal{A}_j\Big)\mathcal{R}_i|_*^2 + 2\ \mathbb{E}|\sum_{i=k}^{\ell}\Big(\prod_{j=i+1}^{\ell}\mathcal{A}_j\Big)\sum_{j=1}^{k}\mathcal{S}_{j,i-j+1}|_*^2 \\
&\le 2\ (\ell-k+1)\sum_{i=k}^{\ell}\mathbb{E}|\Big(\prod_{j=i+1}^{\ell}\mathcal{A}_j\Big)\mathcal{R}_i|_*^2 + 2\ k\ \sum_{i=k}^{\ell}\sum_{j=1}^{k}\mathbb{E}|\Big(\prod_{j=i+1}^{\ell}\mathcal{A}_j\Big)\mathcal{S}_{j,i-j+1}|_*^2 \\
&\le 2\Big(\frac{a\ T}{\mathbf{h}}\sum_{i=k}^{\ell}\mathbb{E}|\mathcal{R}_i|_*^2 + k\sum_{i=k}^{\ell}\sum_{j=1}^{k}\mathbb{E}|\mathcal{S}_{j,i-j+1}|_*^2\Big) \\
&= 2\sum_{i=k}^{\ell}\Big(\frac{a\ T}{\mathbf{h}}\ \mathbb{E}|\mathcal{R}_i|_*^2 + k\sum_{j=1}^{k}\mathbb{E}|\mathcal{S}_{j,i-j+1}|_*^2\Big).
\end{aligned}$$

Inserting this into the intermediate result (A.4) we obtain

$$\max_{\ell=k-1,\dots,N}\mathbb{E}|\mathcal{E}_\ell|_*^2 \le \hat{S}\Big\{\mathbb{E}|\mathcal{E}_{k-1}|_*^2 + 2\sum_{i=k}^{\ell}\Big(\frac{a\ T}{\mathbf{h}}\ \mathbb{E}|\mathcal{R}_i|_*^2 + k\sum_{j=1}^{k}\mathbb{E}|\mathcal{S}_{j,i-j+1}|_*^2\Big)\Big\},$$

and thus $\max_{\ell=k-1,\dots,N}\mathbb{E}|E_\ell|^2$

$$\le c^*\hat{S}\Big\{c_*\max_{\ell=0,\dots,k-1}\mathbb{E}|E_\ell|^2 + 2\ c_{11}^*\max_{\ell=k,\dots,N}\Big(\frac{a^2T^2}{\mathbf{h}^2}\mathbb{E}|R_\ell|^2 + \frac{kaT}{\mathbf{h}}\sum_{j=1}^{k}\mathbb{E}|S_{j,\ell-j+1}|^2\Big)\Big\}.$$

Taking the square root yields the final estimate

$$\begin{aligned}
\max_{\ell=k-1,\dots,N}\|E_\ell\|_{L_2} &\le \sqrt{c^*\hat{S}}\Bigg\{\sqrt{c_*}\max_{\ell=0,\dots,k-1}\|E_\ell\|_{L_2} \\
&\qquad + \sqrt{2c_{11}^*}\max_{\ell=k,\dots,N}\Bigg(\frac{aT}{\mathbf{h}}\|R_\ell\|_{L_2} + \sqrt{\frac{kaT}{\mathbf{h}}\sum_{j=1}^{k}\|S_{j,\ell-j+1}\|_{L_2}^2}\Bigg)\Bigg\} \\
&\le S\left\{\max_{\ell=0,\dots,k-1}\|E_\ell\|_{L_2} + \max_{\ell=k,\dots,N}\left(\frac{\|R_\ell\|_{L_2}}{\mathbf{h}} + \frac{\sqrt{\sum_{j=1}^{k}\|S_{j,\ell-j+1}\|_{L_2}^2}}{\sqrt{\mathbf{h}}}\right)\right\},
\end{aligned}$$

which completes the proof.

Acknowledgements

The author thanks Renate Winkler for helpful discussions introducing to the subject and the anonymous referees for their valuable comments. Support of the German Ministry of Education and Research (BMBF) through project 03RONAVN is gratefully acknowledged.

References

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

[2] E. Buckwar and R. Winkler, Multi-step methods for SDEs and their application to problems with small noise, *SIAM J. Num. Anal.* **44**(2) (2006) 779–803.

[3] P.M. Burrage and K. Burrage, A variable stepsize implementation for stochastic differential equations, *SIAM J. Sci. Comp.* **24** (2002) 848–864.

[4] G. Denk and R. Winkler, Modeling and simulation of transient noise in circuit simulation, *Mathematical and Computer Modelling of Dynamical Systems*, to appear.

[5] E. Hairer, S.P. Nørsett and G. Wanner, *Solving Ordinary Differential Equations I: Nonstiff Problems, 2. Rev.Ed.*, volume **8** of *Springer Series in Computational Mathematics*, Springer Verlag, Berlin, 1993.

[6] S. Mauthner, Schrittweitensteuerung bei der numerischen Lösung stochastischer Differentialgleichungen, Ph.D. Thesis, Technische Universität Darmstadt, 1999.

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

[8] G.N. Milstein, *Numerical Integration of Stochastic Differential Equations*, Kluwer, 1995. Translation from the Russian original of 1988.

[9] G.N. Milstein and M.V. Tretyakov, Mean-square numerical methods for stochastic differential equations with small noise. *SIAM J.Sci. Comut.* **18** (1997) 1067–1087.

[10] C. Penski, A new numerical method for SDEs and its application in circuit simulation, *J. Comput. Appl. Math.* **115**(1-2) (2000) 461–470.

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

[12] T. Sickenberger and R. Winkler, Efficient transient noise analysis in ciruit simulation, Proceedings of the GAMM Annual Meeting 2006, Berlin, submitted to PAMM (Proc. Appl. Math. Mech.).

[13] R. Winkler, Stochastic differential algebraic equations of index 1 and

applications in circuit simulation, *J. Comput. Appl. Math.* **157** (2003) 477–505.