MEAN-SQUARE CONVERGENCE OF STOCHASTIC MULTI-STEP METHODS WITH VARIABLE STEP-SIZE

THORSTEN SICKENBERGER*

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 $h-\varepsilon$ approach and we construct some stochastic linear multi-step methods with variable step-size that have order 2 behavior 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.

AMS subject classifications. 60H35, 65C30, 65L06, 60H10, 65L20

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

$$X(s)|_{t_0}^{t} = \int_{t_0}^{t} f(X(s),s)\mathrm{d}s + \int_{t_0}^{t} G(X(s),s)\mathrm{d}W(s), \quad X(t_0) = X_0, \tag{1.1}$$

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 of the Wiener process and with finite second moments. It is assumed that there exists a path-wise unique strong solution $X(\cdot)$ of (1.1).

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 (1.1). Although there is a well-developed convergence analysis for discretization 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, 5]. 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.

*Humboldt-Universität zu Berlin, Institut für Mathematik, Unter den Linden 6, 10099 Berlin (sickenberger@mathematik.hu-berlin.de). Support of BMBF through project 03RONAVN is gratefully acknowledged.

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 fulfill 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

$$X(s)|_{t_0}^{t} = \int_{t_0}^{t} f(X(s),s)\mathrm{d}s + \varepsilon \int_{t_0}^{t} \hat{G}(X(s),s)\mathrm{d}W(s), \quad X(t_0) = X_0, \tag{1.2}$$

for $t \in \mathcal{J}$, where $\varepsilon \ll 1$ is a small parameter. The appendix contains the proof of Theorem (3.2).

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 discretization $t_0 < t_1 < \ldots < t_N = T$ 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

$$\sum_{j=0}^{k} \alpha_{\ell,j}\, X_{\ell-j} = h_\ell \sum_{j=0}^{k} \beta_{\ell,j}\, f(X_{\ell-j},t_{\ell-j}) + \sum_{j=1}^{k} \Gamma_{\ell,j}(X_{\ell-j},t_{\ell-j})\, I^{t_{\ell-j},t_{\ell-j+1}}. \tag{2.1}$$

The coefficients $\alpha_{\ell,j}$, $\beta_{\ell,j}$ and the diffusion terms $\Gamma_{\ell,j}$ actually depend on the ratios κ_j for $j = \ell-k+1,\ldots,\ell$. 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 on 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

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

A general multiple Wiener integral is given by

$$I_{r_1,r_2,\ldots,r_j}^{t,t+h}(y) = \int_t^{t+h}\int_t^{s_1}\ldots\int_t^{s_{j-1}} y(X(s_j),s_j)\mathrm{d}W_{r_1}(s_j)\ldots\mathrm{d}W_{r_j}(s_1), \tag{2.2}$$

where $r_i \in \{0, 1, \dots, m\}$ and $\mathrm{d}W_0(s) = \mathrm{d}s$. If $y \equiv 1$ we write $I^{t,t+h}_{r_1,r_2,\dots,r_j}$. 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 [7] that the multiple integrals have the properties

$$\mathbb{E}(I^{t,t+h}_{r_1,\dots,r_j}(\cdot)|\mathcal{F}_t) = 0 \quad \text{if at least one of the indices } r_i \neq 0, \tag{2.3}$$

$$\|\mathbb{E}(I^{t,t+h}_{r_1,\dots,r_j}(\cdot)|\mathcal{F}_t)\|_{L_2} = \mathcal{O}(h^{l_1+l_2/2}), \tag{2.4}$$

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 $\beta_{\ell,0} = 0,\ \ell = k, \dots, N$, the SLMM (2.1) is explicit, otherwise it is drift-implicit. For the diffusion term we use an explicit discretization.

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, 7, 6, 9]. Note that in the literature the term *strong* convergence is sometimes used synonymously for our expression *mean-square* convergence.

DEFINITION 3.1. *We call the SLMM (2.1) for the approximation of the solution of the SDE (1.1)* **mean-square convergent** *if the global error $e_\ell := X(t_\ell) - X_\ell$ satisfies*

$$\max_{\ell=1,\dots,N} \|e_\ell\|_{L_2} \to 0 \quad \textit{as} \quad \mathbf{h} \to 0,$$

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

$$\max_{\ell=1,\dots,N} \|e_\ell\|_{L_2} \leq C \cdot \mathbf{h}^\gamma,$$

with a grid-independent constant $C > 0$.

The mean-square convergence follows almost immediately with the notion of numerical stability in the mean-square sense together with mean-square consistency.

3.1. Numerical stability in the mean-square sense. We assume that the scheme (2.1) for the SDE (1.1) satisfies the following properties:

(A1) the function $f : R^n \times J \to R^n$ satisfies a **uniform Lipschitz condition** with respect to x:

$$|f(x,t) - f(y,t)| \leq L_f |x - y|, \quad \forall x, y \in \mathbb{R}^n, t \in \mathcal{J}, \tag{3.1}$$

where L_f is a positive constant;

(A2) the functions $\Gamma_{\ell,j} : \mathbb{R}^n \times J \to \mathbb{R}^{n \times m_\Gamma}$ satisfies a **uniform Lipschitz condition** with respect to x:

$$|\Gamma_{\ell,j}(x,t) - \Gamma_{\ell,j}(y,t)| \leq L_{\Gamma_{\ell,j}} |x - y|, \quad \forall x, y \in \mathbb{R}^n, t \in \mathcal{J}, \tag{3.2}$$

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

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

$$|\Gamma_{\ell,j}(x,t)| \leq K_{\Gamma_{\ell,j}} (1 + |x|^2)^{\frac{1}{2}}, \quad \forall x \in \mathbb{R}^n, t \in \mathcal{J}. \tag{3.3}$$

(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 (2.1)

$$\rho(\zeta) = \alpha_0(1, \dots, 1)\zeta^k + \alpha_1(1, \dots, 1)\zeta^{k-1} + \dots \alpha_k(1, \dots, 1) \tag{3.4}$$

  lie on or within the unit circle and
  (ii) the roots on the unit circle are simple.

Conditions (A1) - (A3) are standard assumptions for analyzing stochastic differential systems, condition (A4) is known [4] in the context of deterministic variable step-size multi-step methods. We now formulate and prove our main theorem on numerical stability. Additionally to the properties in the context of equidistant grids we have to fulfill conditions for the maximum step-size on the grid and for the step-size ratios of the sequence.

THEOREM 3.2. *Assume that (A1) - (A4) hold for the scheme (2.1). Then there exists constants κ, K ($\kappa < 1 < K$), $a \geq 0$, $h^0 > 0$ and a stability constant $S > 0$ such that the following holds true for each grid $\{t_0, t_1, \dots, t_N\}$ having the property $\mathbf{h} := \max_{\ell=1,\dots,N} h_\ell \leq h^0$, $\mathbf{h} \cdot N \leq a \cdot (T - t_0)$ and $\kappa \leq h_\ell / h_{\ell-1} \leq K$ 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 (2.1) and the perturbed discrete system

$$\sum_{j=0}^{k} \alpha_{\ell,j}\, \tilde{X}_{\ell-j} = h_\ell \sum_{j=0}^{k} \beta_{\ell,j}\, f(\tilde{X}_{\ell-j}, t_{\ell-j}) + \sum_{j=1}^{k} \Gamma_{\ell,j}(\tilde{X}_{\ell-j}, t_{\ell-j}) I^{t_{\ell-j}, t_{\ell-j+1}} + D_\ell, \tag{3.5}$$

$\ell = k, \dots, N$, 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 estimate by

$$\max_{\ell=1,\dots,N} \|e_\ell\|_{L_2} \leq S \Big\{ \max_{\ell=0,\dots,k-1} \|e_\ell\|_{L_2} + \max_{\ell=k,\dots,N} \Big(\frac{\|R_\ell\|_{L_2}}{\mathbf{h}} + \frac{\sqrt{\sum_{j=1}^{k} \|S_{j,\ell-j+1}\|_{L_2}^2}}{\sqrt{\mathbf{h}}} \Big) \Big\}, \tag{3.6}$$

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 $\mathbb{E}(S_{j,\ell-j+1} | F_{t_{i-j}}) = 0$ 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 (2.1) for the SDE (1.1) fulfils the assertion of Theorem 3.2, we call it *numerically stable in the mean-square sense.*

3.2. Mean-square consistency. Different notions of errors for pathwise approximation are studied in the literature. 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 (2.1) for the approximation of the solution of the SDE (1.1) is given as

$$L_\ell := \sum_{j=0}^{k} \alpha_{\ell,j}\, X(t_{\ell-j}) - h_\ell \sum_{j=0}^{k} \beta_{\ell,j}\, f(X(t_{\ell-j}), t_{\ell-j}) - \sum_{j=1}^{k} \Gamma_{\ell,j}(X(t_{\ell-j}), t_{\ell-j})\, I^{t_{\ell-j}, t_{\ell-j+1}}, \qquad \ell = k, \dots, N, \tag{3.7}$$

$$L_\ell := X(t_\ell) - X_\ell, \qquad \ell = 0, \dots, k-1. \tag{3.8}$$

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

$$L_\ell = R_\ell + S_\ell =: R_\ell + \sum_{j=1}^{k} S_{j,\ell-j+1}, \quad \ell = k, \dots, N, \tag{3.9}$$

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 $j = 1, \dots, k$ as in [2]. Note that the representation (3.9) is not unique.

Definition 3.3. *We call the SLMM (2.1) for the approximation of the solution of the SDE (1.1)* **mean-square consistent** *if the local error L_ℓ satisfies*

$$h_\ell^{-1} \|\mathbb{E}(L_\ell|\mathcal{F}_{t_{\ell-k}})\|_{L_2} \to 0 \text{ for } h_\ell \to 0, \text{ and } h_\ell^{-1/2} \|L_\ell\|_{L_2} \to 0 \text{ for } h_\ell \to 0; \tag{3.10}$$

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

$$\|\mathbb{E}(L_\ell|\mathcal{F}_{t_{\ell-k}})\|_{L_2} \leq \bar{c} \cdot h_\ell^{\gamma+1} \quad \text{and} \quad \|L_\ell\|_{L_2} \leq c \cdot h_\ell^{\gamma+\frac{1}{2}}, \quad \ell = k, \dots, N, \tag{3.11}$$

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

Subsequently we assume that the conditions of theorem 3.2 are fulfilled. In order to prove mean-square convergence of order γ it is then sufficient to find a representation (3.9) of the local error L_ℓ such that

$$\|\mathbb{E}(R_\ell)\|_{L_2} \leq \bar{c} \cdot h_\ell^{\gamma+1} \quad \text{and} \quad \|S_\ell\|_{L_2} \leq c \cdot h_\ell^{\gamma+\frac{1}{2}}, \quad \ell = k, \dots, N, \tag{3.12}$$

with constants $c, \bar{c} > 0$ only depending on the SDE and its solution. Together the condition (3.12) imply the estimates

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

4. Local error analysis. To analyse the local error L_ℓ of a discretization scheme for the SDE (1.1) and to achieve a suitable representation (3.9) 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.

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.
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, $r=1,\dots,m$, defined on $C^{2,1}$ and $C^{1,0}$, respectively, by

$$\Lambda_0 y = y'_t + y'_x f + \frac{1}{2}\sum_{r=1}^{m}\sum_{i,j=1}^{n} y''_{x_i x_j} g_{ri} g_{rj}\ , \quad \Lambda_r y = y'_x g_r\ ,\ r=1,\dots,m, \tag{4.1}$$

and remind the reader of the notation for multiple Wiener integrals (2.2). Using these operators the Itô formula for a function y in $C^{2,1}$ and the solution X of (1.1) reads

$$y(X(t),t) = y(X(t_0),t_0) + I_0^{t_0,t}(\Lambda_0 y) + \sum_{r=1}^{m} I_r^{t_0,t}(\Lambda_r y), \quad t \in \mathcal{J}. \tag{4.2}$$

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

$$\sum_{j=0}^{2} \alpha_{\ell,j} X_{\ell-j} = h_\ell \sum_{j=0}^{2} \beta_{\ell,j} f(X_{\ell-j},t_{\ell-j}) + \sum_{j=1}^{2}\gamma_{\ell,j}\sum_{r=1}^{m} g_r(X_{\ell-j},t_{\ell-j})\, I_r^{t_{\ell-j},t_{\ell-j+1}}, \tag{4.3}$$

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

$$f(X(t_{\ell-1}),t_{\ell-1}) = f(X(t_{\ell-2}),t_{\ell-2}) + I_0^{t_{\ell-2},t_{\ell-1}}(\Lambda_0 f) + \sum_{r=1}^{m} I_r^{t_{\ell-2},t_{\ell-1}}(\Lambda_r f), \tag{4.4}$$

$$\begin{aligned} f(X(t_\ell),t_\ell) = f(X(t_{\ell-2}),t_{\ell-2}) &+ I_0^{t_{\ell-2},t_{\ell-1}}(\Lambda_0 f) + I_0^{t_{\ell-1},t_\ell}(\Lambda_0 f) \\ &+ \sum_{r=1}^{m} I_r^{t_{\ell-2},t_{\ell-1}}(\Lambda_r f) + \sum_{r=1}^{m} I_r^{t_{\ell-1},t_\ell}(\Lambda_r f). \end{aligned} \tag{4.5}$$

For the general SDE (1.1) we have the following result.

LEMMA 4.1. *Assume that the coefficients f, $g_r, r=1,\dots,m$ of the SDE (1.1) 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 (3.7) of the stochastic 2-step scheme (4.3) allows the representation*

$$L_\ell = R_\ell^\circ + S_{1,\ell}^\circ + S_{2,\ell-1}^\circ, \quad \ell = 2,\dots,N, \tag{4.6}$$

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

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

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

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

with

$$\|\tilde{R}_\ell^\circ\|_{L_2} = O(h_\ell^2), \quad \|\tilde{S}_{1,\ell}^\circ\|_{L_2} = O(h_\ell), \quad \|\tilde{S}_{2,\ell-1}^\circ\|_{L_2} = O(h_\ell). \tag{4.7}$$

COROLLARY 4.2. *Let the coefficients f, g_r, $r=1,\dots,m$, of the SDE (1.1) satisfy the assumptions of Lemma 4.1 and suppose they are Lipschitz continuous with respect to their first variable. Let the stochastic linear two-step scheme with variable step-size (4.3) are stable and the coefficients satisfy the consistency conditions*

$$\sum_{j=0}^{2}\alpha_{\ell,j}=0,\quad \alpha_{\ell,0}+\frac{1}{\kappa_\ell}(\alpha_{\ell,0}+\alpha_{\ell,1})=\sum_{j=0}^{2}\beta_{\ell,j},\quad \alpha_{\ell,0}=\gamma_{\ell,1},\quad \alpha_{\ell,0}+\alpha_{\ell,1}=\gamma_{\ell,2}. \tag{4.8}$$

Then the global error of the scheme (4.3) applied to (1.1) allows the expansion

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

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

Proof. (of Corollary 4.2) By Lemma 4.1 we have the representation (4.6) for the local error. Applying the consistency conditions (4.8) 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 (4.3) satisfies the conditions of Theorem 3.2, it is numerically stable in the mean-square sense. Now the assertion follows from the estimates (4.7) by means of the stability inequality. ∎

Proof. (of Lemma 4.1) To derive a representation of the local error in the form (4.6) 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 (3.7) as

$$\begin{aligned}L_\ell=\;&\alpha_{\ell,0}\big(X(t_\ell)-X(t_{\ell-1})\big)+(\alpha_{\ell,0}+\alpha_{\ell,1})\big(X(t_{\ell-1})-X(t_{\ell-2})\big)+\sum_{j=0}^{2}\alpha_{\ell,j}X(t_{\ell-2})\\&-h_\ell\sum_{j=0}^{2}\beta_{\ell,j}f(X(t_{\ell-j}),t_{\ell-j})-\sum_{j=1}^{2}\gamma_{\ell,j}G(X(t_{\ell-j}),t_{\ell-j})\Delta W_{\ell-j+1}.\end{aligned}$$

The SDE (1.1) implies the identities

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

and, additionally using (4.4),

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

Inserting this and the expansions (4.4), (4.5) 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

$$\tilde{R}_\ell^\circ = \alpha_{\ell,0}\big\{h_\ell I_0^{t_{\ell-2},t_{\ell-1}}(\Lambda_0 f) + I_{00}^{t_{\ell-1},t_\ell}(\Lambda_0 f)\big\} + (\alpha_{\ell,0} + \alpha_{\ell,1}) I_{00}^{t_{\ell-2},t_{\ell-1}}(\Lambda_0 f)$$
$$-h_\ell \beta_{\ell,0}\big\{I_0^{t_{\ell-2},t_{\ell-1}}(\Lambda_0 f) + I_0^{t_{\ell-1},t_\ell}(\Lambda_0 f)\big\} - h_\ell \beta_{\ell,1} I_0^{t_{\ell-2},t_{\ell-1}}(\Lambda_0 f), \tag{4.9}$$

$$\tilde{S}_{1,\ell}^\circ = \sum_{r=1}^{m} \Big(\alpha_{\ell,0} I_{r0}^{t_{\ell-1},t_\ell}(\Lambda_r f) - h_\ell \beta_{\ell,0} I_r^{t_{\ell-1},t_\ell}(\Lambda_r f)\Big) + \alpha_{\ell,0} \sum_{r=1}^{m} I_{0r}^{t_{\ell-1},t_\ell}(\Lambda_0 g_r)$$
$$+ \alpha_{\ell,0} \sum_{r,q=1}^{m} I_{qr}^{t_{\ell-1},t_\ell}(\Lambda_q g_r)), \tag{4.10}$$

$$\tilde{S}_{2,\ell-1}^\circ = h_\ell(\alpha_{\ell,0} - \beta_{\ell,0} - \beta_{\ell,1}) \sum_{r=1}^{m} I_r^{t_{\ell-2},t_{\ell-1}}(\Lambda_r f) + (\alpha_{\ell,0} + \alpha_{\ell,1}) \sum_{r=1}^{m} I_{r0}^{t_{\ell-2},t_{\ell-1}}(\Lambda_r f)$$
$$+(\alpha_{\ell,0} + \alpha_{\ell,1}) \sum_{r=1}^{m} I_{0r}^{t_{\ell-2},t_{\ell-1}}(\Lambda_0 g_r) + (\alpha_{\ell,0} + \alpha_{\ell,1}) \sum_{r,q=1}^{m} I_{qr}^{t_{\ell-2},t_{\ell-1}}(\Lambda_q g_r). \tag{4.11}$$

Finally, the estimates (4.7) are derived by means of (2.3) and (2.4), where the last terms in (4.10) and (4.11) determine the order $\mathcal{O}(h_\ell)$. ∎

Example 4.3. As examples we give stochastic variants of the trapezoidal rule, the two-step Adams-Bashforth (AB) and the backward differential formulae (BDF) 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

$$X_\ell - X_{\ell-1} = h_\ell \frac{1}{2}\big(f(X_\ell,t_\ell)+f(X_{\ell-1},t_{\ell-1})\big) + \sum_{r=1}^{m} g_r(X_{\ell-1},t_{\ell-1})\ I_r^{t_{\ell-1},t_\ell}. \tag{4.12}$$

The Adams-Bashforth scheme is given as

$$\begin{aligned} X_\ell - X_{\ell-1} = h_\ell\bigg(\frac{\kappa_\ell+2}{2}f(X_{\ell-1},t_{\ell-1}) - \frac{\kappa_\ell}{2}f(X_{\ell-2},t_{\ell-2})\bigg)& \\ +\sum_{r=1}^{m} g_r(X_{\ell-1},t_{\ell-1})\ I_r^{t_{\ell-1},t_\ell}& \end{aligned} \tag{4.13}$$

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 two-step BDF takes the form

$$\begin{aligned} X_\ell - \frac{(\kappa_\ell+1)^2}{2\kappa_\ell+1}X_{\ell-1} + \frac{\kappa_\ell^2}{2\kappa_\ell+1}X_{\ell-2} = h_\ell\ \frac{\kappa_\ell+1}{2\kappa_\ell+1}f(X_\ell,t_\ell)& \\ +\sum_{r=1}^{m} g_r(X_{\ell-1},t_{\ell-1})\ I_r^{t_{\ell-1},t_\ell} - \frac{\kappa_\ell^2}{2\kappa_\ell+1}\sum_{r=1}^{m} g_r(X_{\ell-2},t_{\ell-2})\ I_r^{t_{\ell-2},t_{\ell-1}}.& \end{aligned} \tag{4.14}$$

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 $\gamma_{\ell,1}=1,\ \gamma_{\ell,2}=-\frac{\kappa_\ell^2}{2\kappa_\ell+1}$.

4.2. Consistency of two-step-Maruyama schemes for small noise SDEs. 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 $\hat\Lambda_r,\ r=1,\dots,m$ defined on $C^{2,1}$ and $C^{1,0}$, respectively, by

$$\Lambda_0^f y := y'_t + y'_x f, \quad \hat\Lambda_0 y := \frac{1}{2}\sum_{r=1}^{m}\sum_{i,j=1}^{n} y''_{x_i x_j}\hat g_{ri}\hat g_{rj}, \quad \hat\Lambda_r y := y'_x \hat g_r\ . \tag{4.15}$$

In terms of the original definition (4.1) we have

$$\Lambda_0 y = \Lambda_0^f y + \epsilon^2\hat\Lambda_0 y \quad \text{ and } \quad \Lambda_r y = \epsilon\hat\Lambda_r y. \tag{4.16}$$

Lemma 4.4. *Assume that the coefficients $f, \hat g_r, r=1,\dots,m$ of the small noise SDE (1.2), 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 $r,q=1,\dots,m$. Let the stochastic 2-step scheme with variable step-size (4.3) satisfy the consistency conditions (4.8). Then the local error (3.7) of the method (4.3) for the small noise SDE (1.2) allows the representation*

$$L_\ell = R_\ell^\diamond + S_{1,\ell}^\diamond + S_{2,\ell-1}^\diamond, \quad \ell=2,\dots,N, \tag{4.17}$$

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

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

where

$$\|\tilde{R}_\ell^\diamond\|_{L_2} = O(h_\ell^3 + \epsilon^2 h_\ell^2), \quad \|\tilde{S}_{1,\ell}^\diamond\|_{L_2} = O(\epsilon h_\ell^{5/2}), \quad \|\tilde{S}_{2,\ell-1}^\diamond\|_{L_2} = O(\epsilon h_\ell^{5/2}). \quad (4.18)$$

The terms $\tilde{S}_{1,\ell}^\circ, \tilde{S}_{2,\ell-1}^\circ$ are given by (4.10, 4.11) in the proof of Lemma 4.1 and satisfy here

$$\|\tilde{S}_{1,\ell}^\circ\|_{L_2} = \mathcal{O}(\epsilon^2 h_\ell + \epsilon h_\ell^{3/2}), \quad \|\tilde{S}_{2,\ell}^\circ\|_{L_2} = \mathcal{O}(\epsilon^2 h_\ell + \epsilon h_\ell^{3/2}). \quad (4.19)$$

Proof. We have from Lemma 4.1, if the consistency conditions (4.8) are satisfied, the representation

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

where $\tilde{R}_\ell^\circ$, $\tilde{S}_{1,\ell}^\circ$, $\tilde{S}_{2,\ell-1}^\circ$ are given by (4.9, 4.10, 4.11). Splitting $\Lambda_0 f = \Lambda_0^f f + \epsilon^2 \hat{\Lambda}_0 f$ immediately yields $\tilde{R}_\ell^\circ = \tilde{R}_\ell^{\circ f} + \epsilon^2 \hat{R}_\ell^\circ$ with

$$\begin{aligned}\tilde{R}_\ell^{\circ f} &:= (\alpha_{\ell,0} - \beta_{\ell,0} - \beta_{\ell,1}) h_\ell I_0^{t_{\ell-2},t_{\ell-1}}(\Lambda_0^f f) + (\alpha_{\ell,0} + \alpha_{\ell,1}) I_{00}^{t_{\ell-2},t_{\ell-1}}(\Lambda_0^f f) \\ &\quad + \alpha_{\ell,0} I_{00}^{t_{\ell-1},t_\ell}(\Lambda_0^f f) - h_\ell \beta_{\ell,0} I_0^{t_{\ell-1},t_\ell}(\Lambda_0^f f) \end{aligned} \quad (4.20)$$

$$\begin{aligned}\hat{R}_\ell^\circ &:= (\alpha_{\ell,0} - \beta_{\ell,0} - \beta_{\ell,1}) h_\ell I_0^{t_{\ell-2},t_{\ell-1}}(\hat{\Lambda}_0 f) + (\alpha_{\ell,0} + \alpha_{\ell,1}) I_{00}^{t_{\ell-2},t_{\ell-1}}(\hat{\Lambda}_0 f) \\ &\quad + \alpha_{\ell,0} I_{00}^{t_{\ell-1},t_\ell}(\hat{\Lambda}_0 f) - h_\ell \beta_{\ell,0} I_0^{t_{\ell-1},t_\ell}(\hat{\Lambda}_0 f). \end{aligned} \quad (4.21)$$

We note that (4.21) appears with the factor ϵ^2 in the local error representation, thus yielding the $\mathcal{O}(\epsilon^2 h_\ell^2)$ term in the estimate of $\|\tilde{R}_\ell^\diamond\|_{L_2}$ in (4.18) . We concentrate on developing $\tilde{R}_\ell^{\circ f}$ in more detail. Applying the Itô-formula (4.2) 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 (4.20) . Integrating again we obtain for the second integral in (4.20)

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

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 (4.20)

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

where

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

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

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

We arrive at $\tilde{R}_\ell^{\diamond} = \tilde{R}_\ell^{\diamond f} + \epsilon^2\hat{R}_\ell^{\circ}$. Finally, the estimates (4.18) are derived by means of (2.3) and (2.4). ◻

Corollary 4.5. *Let the coefficients f, $\hat{g}_r$, $r = 1, \ldots, m$, of the SDE (1.2) satisfy the assumptions of Lemma 4.4 and suppose they are Lipschitz continuous with respect to their first variable. Let the stochastic linear two-step scheme with variable step-size (4.3) are stable, the coefficients satisfy the consistency conditions (4.8) and*

$$\left(\frac{1}{\kappa_\ell^2} + \frac{2}{\kappa_\ell} + 1\right)\alpha_{\ell,0} + \frac{1}{\kappa_\ell^2}\ \alpha_{\ell,1} - \left(\frac{2}{\kappa_\ell} + 2\right)\beta_{\ell,0} - \frac{2}{\kappa_\ell}\ \beta_{\ell,1} = 0. \tag{4.22}$$

Then the global error of the scheme (4.3) applied to (1.2) allows the expansion

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

Proof. Lemma 4.4 stated the representation (4.17) for the local error. Applying the consistency condition (4.22) yields $R_\ell^{\diamond} = \tilde{R}_\ell^{\diamond}$ and by (4.18) we have $\|R_\ell^{\diamond}\|_{L_2} = \mathcal{O}(h_\ell^3 + \epsilon^2 h_\ell^2)$. 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 $\mathcal{O}(\epsilon^2 h_\ell + \epsilon h_\ell^{3/2})$. As the scheme (4.3) satisfies the conditions (A1) - (A4), it is numerically stable in the mean-square sense. Applying the stability inequality (3.6) to the representation (4.17) of the local error yields the assertion. ◻

We remark that the schemes (4.12), (4.13) and (4.14) satisfy the assumptions and the consistency conditions of corollary 4.5. Thus, these schemes are numerically stable in the mean-square sense and we can expect order 2 behavior if the term $\mathcal{O}(\mathbf{h}^2)$ of the global error dominates the term $\mathcal{O}(\varepsilon\mathbf{h} + \varepsilon^2\mathbf{h}^{1/2})$.

Appendix A. For the proof of Theorem (3.2) we need a discrete version of Gronwall's lemma.

LEMMA A.1. *Let a_ℓ, $\ell = 1,\dots,N$, and C_1, C_2 be nonnegative real numbers and assume that the inequalities*

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

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

Proof. (of Theorem (3.2))
Part1 (Existence of a solution $\tilde{X}_\ell$): We consider scheme (3.5). 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 (3.5) only implicity as the solution of the fixed point equation

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

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

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

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

Thus, $\eta_\ell(\cdot; z_1,\dots,z_k,b)$ has a globally unique fixed point $x = \xi_\ell(z_1,\dots,z_k,b)$, and $\xi_\ell(\tilde{X}_{\ell-1},\dots,\tilde{X}_{\ell-k},B_\ell)$ gives the unique solution $\tilde{X}_\ell$ of (3.5). Moreover, ξ_ℓ depends Lipschitz-continuously on $z_1,\dots,z_k$ and b since

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

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

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 $j = 1,\dots,k$. We compare $\tilde{X}_\ell = \xi_\ell(\tilde{X}_{\ell-1},\dots,\tilde{X}_{\ell-k},B_\ell)$ 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][4, Chap.III.4][8, Chap.8.2.1]).
For X_ℓ and $\tilde X_\ell$ being the solutions of (2.1) and (3.5), 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 trivial identities $E_{\ell-1} = E_{\ell-1},\ \dots\ E_{\ell-k+1} = E_{\ell-k+1}$ we obtain

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

or, in compact form

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

where $\mathcal{E}_\ell \in L_2(\Omega,\mathbb{R}^{k\times n}),\ \ell = k-1,\dots,N$. 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 $\ell = k,\dots,N$ 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. [4, 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 $\mathcal{A}_j$. The eigenvalues of the companion matrix $\mathcal{A}$ of the constant step-size formula are the roots of the characteristic polynomial ρ (3.4) 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 $\langle\mathcal{X},\mathcal{Y}\rangle_* = c^*_{11}\langle x,y\rangle_2 = c^*_{11}x^Ty$, 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 $|\big(\prod_{j=k}^{\ell}\mathcal{A}_j\big)\mathcal{E}_{k-1}|_*^2 \le |\mathcal{E}_{k-1}|_*^2$, and thus

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

For the term labelled 2) we have

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

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

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

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

$$\Delta\Psi_i = ((\Delta\psi^i)^T,0,\dots,0)^T = ((\sum_{j=1}^{k}\Delta\Gamma_{j,i-j}\ I^{t_{i-j},t_{i-j+1}})^T,0,\dots,0)^T = \sum_{j=1}^{k}\Delta\Psi_{j,i-j}$$

and

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

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

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

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 Gronwalls Lemma with $a_\ell := 0,\ \ell = 1, \dots, k-2$ and $a_\ell := \mathbb{E}|\mathcal{E}_\ell|_*^2,\ \ell = k-1, \dots, N$, and obtain the intermediate result

$$\max_{\ell=k-1,\dots,N} \mathbb{E}|\mathcal{E}_\ell|_*^2 \le \hat{S}\Big\{\mathbb{E}|\mathcal{E}_{k-1}|_*^2 + \mathbb{E}|\sum_{i=k}^{\ell}\Big(\prod_{j=i+1}^{\ell}\mathcal{A}_j\Big)\mathcal{D}_i|_*^2\Big\},\ \hat{S} := 8\exp(8LaT)\ . \tag{A.4}$$

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 $i_1 \ne i_2$, 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

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

which completes the proof. □

Acknowlegdements. The author thanks Renate Winkler for helpful discussions introducing to the subject. 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.*, 3:315–335, 2000.

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

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

[4] 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.

[5] S. Mauthner. *Schrittweitensteuerung bei der numerischen Lösung stochastischer Differentialgleichungen.* PhD thesis, Technische Universität Darmstadt, 1999.

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

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

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

[9] R. Winkler. Stochastic differential algebraic equations of index 1 and applications in circuit simulation. *J. Comput. Appl. Math.*, 157(2):477–505, 2003.