Stepsize control for mean-square numerical methods for stochastic differential equations with small noise

Werner Römisch and Renate Winkler
Humboldt-Universität Berlin
Institut für Mathematik
10099 Berlin, Germany

Abstract

A strategy for controlling the stepsize in the numerical integration of stochastic differential equations (SDEs) is presented. It is based on estimating the p-th mean of local errors. The strategy leads to deterministic stepsize sequences that are identical for all paths. For the family of Euler schemes for SDEs with small noise we derive computable estimates for the dominating term of the p-th mean of local errors and show that the strategy becomes efficient for reasonable stepsizes. Numerical experience is reported for test examples including scalar SDEs and a stochastic circuit model.

1 Introduction

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

$$x(t) = x_0 + \int_{t_0}^{t} f(x(s),s)ds + \int_{t_0}^{t} G(x(s),s)dw(s), \quad t \in \mathcal{J}, \tag{1}$$

where $\mathcal{J} = [t_0, T]$, $f : \mathbb{R}^n \times \mathcal{J} \to \mathbb{R}^n$, $G : \mathbb{R}^n \times \mathcal{J} \to \mathbb{R}^{n\times m}$ are continuous functions, and, moreover, f has continuous derivatives with respect to x. w is an 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. It is assumed that there exists a pathwise unique, strong solution $x(\cdot)$.

We study mean-square and, more generally, p-th mean convergent numerical methods for solving (1) based on time discretization. Our work is motivated by practical SDE models in circuit simulation [23, 27] that do not satisfy the commutativity condition for G and are large scale with respect to n and m, respectively. As function calls are costly, we look at variable stepsize methods of low order and propose a new strategy for selecting stepsizes. More precisely, we consider p-th mean convergent one-step methods and present a stepsize control that is based on estimating the p-th mean of the local discretization error and of the local error constants, respectively. As in the deterministic case, the strategy is justified by the fact that p-th mean global errors can be estimated by the corresponding local ones provided that the method is stable. The proposed stepsize control requires to determine an ensemble of approximate solution paths simultaneously and uses stepsize sequences that are identical for all paths. Since the strategy is based on estimating local error constants, we develop representations of local errors for the family of Euler schemes in terms of multiple stochastic integrals. In case of small noise, a case of particular interest in applications [27], all terms containing multiple stochastic integrals become small such that they are negligible for realistic stepsizes that are not too small relative to the smallness of the noise. For the dominating term we provide computable estimates by using available local information. The stepsize strategy was implemented for the implicit Euler method and extensively tested on a set of test examples. The choice of the method allows us to study the effects of the stepsize selection on the accuracy, i.e., the global discretization error, and on the convergence behaviour of Newton's method for solving the implicit nonlinear equations simultaneously. In case of step rejections, the method described in [18] is used to ensure that the correct Wiener paths are followed. Our numerical experience of the stepsize strategy confirms its usefulness and potential for SDEs with small noise, and also provides some information on its limitations. It turns out that the stepsize control works successfully for ranges of stepsizes that lead to reasonably accurate results, but still are not too small such that the asymptotic order of convergence $\mathcal{O}(h^{\frac{1}{2}})$ dominates the error. The latter phenomenon for SDEs with small noise was experimentally observed in [1, 6].

Several variable stepsize strategies for SDEs were developed during the last few years. Most of them are based on pathwise arguments and lead to pathwise different stepsize sequences. Such approaches often require a separate convergence analysis, as the available convergence theory for SDEs (e.g., in the mean square or weak sense) is based on properties of certain expectations rather than paths which are typically non-smooth objects. The strategies for pathwise controlling stepsizes differ for each approach. The classical paper [7] proposes a pathwise strategy by comparing results of a given integration scheme with those of a higher order method. Hence, at least the higher order method requires the (approximate) computation of multiple Itô-integrals. The approaches in [17, 18] and [4] are also based on a comparison of two Runge-Kutta schemes of different order. In [14] conditions are provided that imply mean square convergence of the Euler-Maruyama scheme with pathwise different stepsize sequences. A different approach was developed in [11, 12, 22]. The authors obtain stepsize sequences that are (mean square and p-th mean) optimal for asymptotically small stepsizes.

Our paper is organized as follows. In Section 2 we introduce the class of discretization schemes that will be considered in this paper. We recall basic p-th mean stability results and conditions for p-th mean convergence stated in terms of the p-th mean of the local discretization error and of its rate of convergence as the stepsize tends to zero. Starting from this observation we present, in Section 3, a strategy for selecting pathwise identical stepsize sequences by estimating the p-th mean of the local error. For the special case of integrating SDEs with small noise by the family of Euler schemes we provide computable estimates for the local errors in Section 4. Finally, in Section 5 we report on numerical experience of test runs of an implementation of the stepsize control for the implicit Euler scheme.

2 Numerical stability, consistency and convergence of discretization methods for SDEs

We consider the drift-implicit discretization scheme of the form

$$x_\ell = x_{\ell-1} + \varphi(x_{\ell-1}, x_\ell; t_{\ell-1}, h_\ell) + \psi(x_{\ell-1}; t_{\ell-1}, h_\ell, I_{t_{\ell-1},h_\ell}),\ \ell = 1, \ldots, N, \tag{2}$$

for solving (1) on the deterministic grid $t_0 < t_1 < \ldots < t_N = T$ with stepsizes $h_\ell := t_\ell - t_{\ell-1}$, $\ell = 1, \ldots, N$. Here, φ and ψ are functions defined on $\mathbb{R}^n \times \mathbb{R}^n \times \mathcal{T}$ and $\mathbb{R}^n \times \mathcal{T} \times \mathbb{R}^{m_I}$ with $\mathcal{T} := \{(t,h) : t, t+h \in \mathcal{J}, h \in \mathbb{R}_+\}$, respectively, and mapping to $\mathbb{R}^n$. By $I_{t,h}$ we denote a vector of m_I multiple stochastic integrals being of the the form

$$I_{i_1,\ldots,i_k;t,h} = \int_t^{t+h} \int_t^{s_1} \cdots \int_t^{s_{k-1}} dw_{i_1}(s_k) dw_{i_2}(s_{k-1}) \cdots dw_{i_k}(s_1),$$

where the indices $i_1, \ldots, i_k$ are in $\{0, 1, \ldots, m\}$, k is bounded by a certain finite order k_{max}, and $dw_0(s)$ corresponds to ds.
For example, the family of drift-implicit Euler schemes, sometimes also called stochastic θ-methods, is of the form

$$x_\ell := x_{\ell-1} + h_\ell(\theta f(x_\ell, t_\ell) + (1-\theta) f(x_{\ell-1}, t_{\ell-1})) + G(x_{\ell-1}, t_{\ell-1})\Delta w_\ell\,,\ \ell = 1, \ldots, N, \tag{3}$$

where $\theta \in [0,1]$, and $\Delta w_\ell := w(t_\ell) - w(t_{\ell-1}) = (I_{i;t_{\ell-1},h_\ell})_{i=1}^m$. Hence, one has $k_{max} := 1$, $m_I := m$, and

$$\begin{aligned} \varphi(z, x; t, h) &:= h(\theta f(x, t+h) + (1-\theta) f(z,t)), \\ \psi(z; t, h, I_{t,h}) &:= G(z,t)(w(t+h) - w(t)) = \sum_{j=1}^m g_j(z,t) \int_t^{t+h} dw_j(s), \end{aligned}$$

where $g_j(z,t)$, $j = 1, \ldots, m$, are the columns of the matrix $G(z,t)$.
The family of drift-implicit Milstein schemes differs from the Euler schemes by an additional correction term for the stochastic part. The Milstein schemes are described by the same function φ, and $k_{max} := 2, m_I := m + m^2$, and

$$\psi(z; t, h, I_{t,h}) := G(z,t)\Delta w_{t,h} + \sum_{j=1}^m ({g_j}_x G)(z,t) I_{(j);t,h}, \tag{4}$$

where $\Delta w_{t,h} := w(t+h) - w(t) = (I_{i;t,h})_{i=1}^m$, and $I_{(j);t,h} := (I_{j,i;t,h})_{i=1}^m$.
For measuring errors in the discretization scheme we use the norm for p-th order integrable random variables $z \in L_p(\Omega, \mathbb{R}^n)$, i.e., $\|z\|_{L_p} := (\mathbb{E}|z|^p)^{1/p}$, where the exponent $p \geq 1$ is properly chosen in the sense that the initial value x_0 has a p-th order moment and that it fits to the unknown statistical parameters of $x(\cdot)$, which have to be computed approximately. We start our analysis by stating a result on p-th mean stability of (2), which extends the fundamental result of Milstein [19, 20]. Its proof is given in [23, 27].

Theorem 2.1 *Let $p \geq 1$ and x_0 have finite p-th mean. Assume that the scheme (2) satisfies the following properties:*

- *for all $z, \tilde{z}, x, \tilde{x} \in \mathbb{R}^n$, $(t,h) \in \mathcal{T}$, $h \le h^1$ we have*

  (A1) $\quad |\varphi(z,x;t,h) - \varphi(\tilde{z},\tilde{x};t,h)| \le h(L_1|z-\tilde{z}| + L_2|x-\tilde{x}|)$

  for some positive constants h^1, L_1, L_2.

- *for all $(t,h) \in \mathcal{T}$, $h \le h^1$, and $\mathcal{F}_t$-measurable random vectors y, $\tilde{y}$ we have*

  (A2) $\quad \mathbb{E}(\psi(y;t,h,I_{t,h}) - \psi(\tilde{y};t,h,I_{t,h})|\mathcal{F}_t) = 0,$

  (A3) $\quad \mathbb{E}(|\psi(y;t,h,I_{t,h}) - \psi(\tilde{y};t,h,I_{t,h})|^p|\mathcal{F}_t) \le h^{\frac{p}{2}} L_3^p |y-\tilde{y}|^p,$

  (A4) $\quad \mathbb{E}|\psi(0;t,h,I_{t,h})|^p < \infty,$

  for some constant $L_3 > 0$.

Then there exist constants $a \ge 1$, $h^0 > 0$ and a stability constant $S > 0$ such that the following holds true for each grid $\{t_0, t_1, \ldots, t_N\}$ having the property $h := \max_{\ell=1,\ldots,N} h_\ell \le h^0$ and $h \cdot N \le a \cdot (T - t_0)$:
For all $\mathcal{F}_{t_0}$-measurable random vectors x_0^*, $\tilde{x}_0$ having finite p-th mean, for all $\ell \in \{1,\ldots,N\}$ and $\mathcal{F}_{t_\ell}$-measurable perturbations d_ℓ^*, $\tilde{d}_\ell$ having finite p-th mean, the perturbed discrete system

$$\tilde{x}_\ell = \tilde{x}_{\ell-1} + \varphi(\tilde{x}_{\ell-1}, \tilde{x}_\ell; t_{\ell-1}, h_\ell) + \psi(\tilde{x}_{\ell-1}; t_{\ell-1}, h_\ell, I_{t_{\ell-1},h_\ell}) + \tilde{d}_\ell,\ \ell = 1,\ldots,N, \tag{5}$$

has a unique solution $\{\tilde{x}_\ell\}_{\ell=0}^N$, and the following estimates are valid for any two solutions $\{x_\ell^*\}_{\ell=1}^N$ and $\{\tilde{x}_\ell\}_{\ell=0}^N$ of the perturbed discrete systems with perturbations $\{d_\ell^*\}_{\ell=1}^N$ and $\{\tilde{d}_\ell\}_{\ell=1}^N$:

$$\mathbb{E} \max_{\ell=1,\ldots,N} |x_\ell^* - \tilde{x}_\ell|^p \le S^p \Big(\mathbb{E}|x_0^* - \tilde{x}_0|^p + \frac{\max_{\ell=1,\ldots,N} \mathbb{E}|s_\ell|^p}{h^{\frac{p}{2}}} + \frac{\mathbb{E} \max_{\ell=1,\ldots,N} |r_\ell|^p}{h^p} \Big), \tag{6}$$

$$\max_{\ell=1,\ldots,N} \mathbb{E}|x_\ell^* - \tilde{x}_\ell|^p \le S^p \Big(\mathbb{E}|x_0^* - \tilde{x}_0|^p + \frac{\max_{\ell=1,\ldots,N} \mathbb{E}|s_\ell|^p}{h^{\frac{p}{2}}} + \frac{\max_{\ell=1,\ldots,N} \mathbb{E}|r_\ell|^p}{h^p} \Big), \tag{7}$$

where $d_\ell := d_\ell^* - \tilde{d}_\ell$ is splitted such that $d_\ell = s_\ell + r_\ell$ with $\mathbb{E}(s_\ell|\mathcal{F}_{t_{\ell-1}}) = 0$.

Extracting the p-th root from (7) yields the stability inequality

$$\max_{\ell=1,\ldots,N} \|x_\ell^* - \tilde{x}_\ell\|_{L_p} \le S\Big(\|x_0^* - \tilde{x}_0\|_{L_p} + \max_{\ell=1,\ldots,N} \|s_\ell\|_{L_p}/h^{\frac{1}{2}} + \max_{\ell=1,\ldots,N} \|r_\ell\|_{L_p}/h \Big). \tag{8}$$

The scheme (2) is called *p-th mean stable* if it satisfies the properties (6) and (7), respectively, in the above result. Furthermore, it is called *p-th mean consistent* of order $\gamma > 0$ if the *local discretization error l_ℓ* at step ℓ, i.e.,

$$l_\ell := x(t_\ell) - x(t_{\ell-1}) - \varphi(x(t_{\ell-1}), x(t_\ell); t_{\ell-1}, h_\ell) - \psi(x(t_{\ell-1}); t_{\ell-1}, h_\ell, I_{t_{\ell-1},h_\ell}), \tag{9}$$

satisfies the properties

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

with constants $c, \bar{c} > 0$ only depending on the SDE and its solution. Clearly, the local discretization error arises by inserting the exact solution $x(\cdot)$ into the numerical scheme. By the global errors e_ℓ we denote the difference

$$e_\ell := x(t_\ell) - x_\ell$$

of the exact and approximate solution at time t_ℓ. The discretization scheme (2) for (1) is called *p-th mean convergent* with order $\gamma > 0$ if the global error $e_\ell := x(t_\ell) - x_\ell$ satisfies the property

$$\max_{\ell=1,\ldots,N} \|e_\ell\|_{L_p} \leq C \cdot h^\gamma\,, \text{ where } \quad h := \max_{\ell=1,\ldots,N} h_\ell\,,$$

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

Theorem 2.2 *If the discretization scheme (2) for (1) is p-th mean consistent with order $\gamma > 0$ and the assumptions of Theorem 2.1 are satisfied, then (2) is p-th mean convergent with order γ. For the difference of the solution $x(t_\ell)$ at the discrete time-points and the solution $\tilde{x}_\ell$ of the perturbed numerical scheme (5) we have the estimate*

$$\max_{\ell=1,\ldots,N} \|x(t_\ell) - \tilde{x}_\ell\|_{L_p} \leq S((c+\bar{c})h^\gamma + \max_{\ell=1,\ldots,N} \delta_\ell / h^{1/2} + \max_{\ell=1,\ldots,N} \bar{\delta}_\ell / h), \tag{10}$$

$$\text{where} \quad \delta_\ell := \|\tilde{d}_\ell\|_{L_p}\,, \quad \bar{\delta}_\ell := \|\mathbb{E}(\tilde{d}_\ell|\mathcal{F}_{t_{\ell-1}})\|_{L_p} \quad , \quad \text{with} \quad \tilde{d}_\ell \quad \text{from (5).}$$

Proof: The assertion follows by applying the triangle inequality

$$\max_{\ell=1,\ldots,N} \|x(t_\ell) - \tilde{x}_\ell\|_{L_p} \leq \max_{\ell=1,\ldots,N} \|x(t_\ell) - x_\ell\|_{L_p} + \max_{\ell=1,\ldots,N} \|x_\ell - \tilde{x}_\ell\|_{L_p}$$

and the stability estimate (6) once to $x(t_\ell)$ related to the perturbations l_ℓ and once again to $\tilde{x}_\ell$ related to the perturbations $\tilde{d}_\ell$.
The p-th mean convergence follows as a special case of (10) for $\tilde{d}_\ell = 0$. □

The general results immediately apply to well-known schemes for SDEs. We illustrate this for the families of drift-implicit Euler and Milstein schemes. Both schemes fit into the frame of (2). By checking the assumptions of Theorem 2.1 we observe that both are p-th mean stable: (A1) follows from the Lipschitz continuity of the drift coefficient f, (A2) is satisfied due to the explicit, non-anticipative discretization of the diffusion term, (A3) follows from the Lipschitz continuity of the diffusion coefficient G (and in case of the Milstein scheme of the functions $(g_j)'_x \cdot G$), and (A4) is a more technical assumption, which is satisfied since the function $G(0,\cdot)$ (and the functions $(g_j)'_x \cdot G(0,\cdot)$) is bounded on the compact interval $\mathcal{J}$. Summarizing we have:

Proposition 2.3 *Let the functions f and G be Lipschitz-continuous with respect to x. Then the Euler schemes (3) are p-th mean stable. If, in addition, the partial derivatives $(g_j)'_x$, $j = 1,\ldots,m$, exist and the functions $(g_j)'_x \cdot G$ are Lipschitz continuous with respect to x, then the Milstein schemes (4) are p-th mean stable.*

From the literature (see e.g. [20]) it is known that the Euler schemes (3) are mean-square consistent with order 1/2 if, in addition, the coefficients are Hölder continuous with exponent 1/2 with respect to t. The Milstein schemes are mean-square consistent with order 1 if the functions f and G are sufficiently smooth. Appealing to Theorem 2.2 then provides the known mean square convergence of the Euler schemes with order 1/2 and of the Milstein schemes with order 1.

3 Stepsize control based on the p-th mean of local errors

For the efficient numerical integration of applied nonlinear SDEs a reasonable stepsize control is indispensable. The problem was addressed in a number of recent papers, e.g. [3, 4, 7, 15, 17, 18]. Most of them suggest individual stepsize sequences for every path. Our approach is rigorously based on the stability and convergence results presented in the previous section. It leads to adaptive stepsize sequences that are uniform for all paths.
By means of the stability inequality (8) we know that the p-th mean of the global errors $e_\ell := x(t_\ell) - x_\ell$ can be estimated by a term that is proportional to the p-th mean of the local errors l_ℓ. Therefore, a natural approach consists in controlling the local error term

$$\eta_\ell := \max\{\|s_\ell\|_{L_p}/h_\ell^{1/2}\,,\ \|r_\ell\|_{L_p}/h_\ell\},\ \text{ where }\ l_\ell = s_\ell + r_\ell,\ I\!\!E(s_\ell|\mathcal{F}_{t_{\ell-1}}) = 0, \qquad (11)$$

according to a given tolerance. In order to develop a strategy for variable stepsize selection, computable estimates for η_ℓ are needed that are based on some insight into its behaviour. Clearly, we have $\eta_\ell = \mathcal{O}(h_\ell^\gamma)$ for a method that is p-th mean consistent with order γ. However, for problems with small noise, η_ℓ may even be dominated by a term $\kappa_\ell \cdot h_\ell^{\bar\gamma}$, where $\bar\gamma \geq \gamma$ and κ_ℓ is a slowly varying factor, for an interesting range of stepsizes h_ℓ that cannot be considered asymptotically small (cf. Sect. 4).
Next, we present the outline of an algorithm for computing an ensemble of M approximate paths $\{x_\ell^i\}_{\ell=1}^N$, $i = 1, \ldots, M$, of $x(\cdot)$ simultaneously. We assume that, for an interesting range of stepsizes h_ℓ, the local error term η_ℓ is dominated by

$$\eta_\ell \leq \kappa_\ell \cdot h_\ell^{\bar\gamma}, \quad \kappa_\ell = \|k_\ell\|_{L_p}, \qquad (12)$$

and that approximate realizations $\tilde k_\ell^i$ of k_ℓ are available for each path $i \in \{1, \ldots, M\}$.

Algorithm 3.1
Given initial values t_0; $x_0^1, \ldots, x_0^M$, an initial stepsize h_1 and a tolerance tol*; set $\ell := 1$.*

1) *Solve the M systems*

$$x_\ell^i = x_{\ell-1}^i + \varphi(x_{\ell-1}^i, x_\ell^i; t_{\ell-1}, h_\ell) + \psi(x_{\ell-1}^i; t_{\ell-1}, h_\ell, I_\ell^i)\,,\ i = 1, \ldots, M,$$

for determining $x_\ell^1, \ldots, x_\ell^M$.

2) *Compute the approximate error constants $\tilde k_\ell^i$, $i = 1, \ldots, M$, and set*

$$\eta_\ell \approx \tilde\eta_\ell := h_\ell^{\bar\gamma}\Big(\frac{1}{M}\sum_{i=1}^M |\tilde k_\ell^i|^p\Big)^{\frac{1}{p}}.$$

3) *Propose a new stepsize h_{new} such that the tolerance multiplied by a safety factor, say 0.8, is matched: The elementary control then leads to*

$$\frac{h_{new}}{h_\ell} := \Big(\frac{0.8 \cdot \mathrm{tol}}{\tilde\eta_\ell}\Big)^{1/\bar\gamma},$$

the proportional integral control PI34 (cf. [9, 25]) leads to

$$\frac{h_{new}}{h_\ell} := \Big(\frac{0.8 \cdot \mathrm{tol}}{\tilde\eta_\ell}\Big)^{0.3/\bar\gamma}\Big(\frac{\tilde\eta_{\ell-1}}{\tilde\eta_\ell}\Big)^{0.4/\bar\gamma}.$$

4) If $\tilde{\eta}_\ell \le \mathrm{tol}$, accept the step.

If $t_\ell \ge T$, stop, else set $\ell := \ell + 1$, $h_\ell := h_{new}$ and go to 1).

If $\tilde{\eta}_\ell > \mathrm{tol}$, reject the step and repeat it with the smaller stepsize $h_\ell := h_{new}$. Store the computed information of the Wiener path and compute intermediate values according to the strategy in [17, 18].

In case of a scheme (2), which uses only the Wiener increments $\Delta w_\ell = w(t_\ell) - w(t_{\ell-1})$, the intermediate values of a Wiener path are computed as follows (cf. [17, 18]):
Let w be an m-dimensional Wiener process, $\Delta w_h := w(t+h) - w(t)$ for some $t \in I\!R_+$ and each $h > 0$, and $h = h_1 + h_2$, $h_1 > 0$, $h_2 > 0$. Then the Wiener increments

$$\Delta w_{h_1} = w(t+h_1) - w(t) \quad \text{and} \quad \Delta w_{h_2} = w(t+h_1+h_2) - w(t+h_1),$$

are simulated according to the formulas

$$\Delta w_{h_1} = \frac{h_1}{h}\Delta w_h + \sqrt{\frac{h_1 h_2}{h}}\nu, \quad \Delta w_{h_2} = \frac{h_2}{h}\Delta w_h - \sqrt{\frac{h_1 h_2}{h}}\nu, \quad \nu \sim N(0, I_m).$$

4 Local error estimates for the family of Euler schemes for SDEs with small noise

There are important applications of SDEs with small noise, where the diffusion coefficients are orders of magnitude smaller than the drift coefficients. For such problems the asymptotic order of convergence is too pessimistic for a reasonable range of stepsizes. Special numerical methods are constructed in [21], taking into account the smallness of the stochastic parts in such systems. Here, we will deal with the family of Euler schemes and present an efficient stepsize control based on the p-th mean of local errors.
Following the lines of [21] we let the SDE with small noise be of the form

$$x(t) = x_0 + \int_{t_0}^{t} f(x(s), s)ds + \int_{t_0}^{t} \epsilon \tilde{G}(x(s), s)dw(s), \quad t \in \mathcal{J}, \tag{13}$$

where $f : I\!R^n \times \mathcal{J} \to I\!R^n$, $\tilde{G} : I\!R^n \times \mathcal{J} \to I\!R^{n\times m}$ are functions satisfying the assumptions introduced in Section 1 for f and G, and ϵ is a small parameter.
The family of drift-implicit Euler schemes with parameter $\theta \in [0, 1]$ for solving (13) on the deterministic grid $t_0 < t_1 < \ldots < t_N = T$ with stepsizes $h_\ell := t_\ell - t_{\ell-1}$, $\ell = 1, \ldots, N$, has the form

$$x_\ell = x_{\ell-1} + h_\ell\big(\theta f(x_\ell, t_\ell) + (1-\theta) f(x_{\ell-1}, t_{\ell-1})\big) + \epsilon \tilde{G}(x_{\ell-1}, t_{\ell-1})\Delta w_\ell, \ \ell = 1, \ldots, N, \tag{14}$$

where $\Delta w_\ell = w(t_\ell) - w(t_{\ell-1}) \sim N(0, h_\ell I_m)$.
In order to derive estimates for the local error l_ℓ of (14), we first establish, similarly as in [21], a representation of l_ℓ in terms of certain multiple stochastic integrals obtained by the Itô-Taylor expansion. The L_p-norm of these stochastic integrals is then characterized in terms of $\mathcal{O}(h_\ell^{k/2}\epsilon^q)$ for some $k, q \in I\!N \cup \{0\}$. Finally, we discuss which terms are dominating for interesting ranges of stepsizes and present computable estimates for the dominating terms.

4.1 Estimating local errors by Itô-Taylor expansion

In order to characterize the conditions on f and $\tilde{G}$ that are needed in the following, we introduce the classes C_L and $C^{s,s-1}$, $s \in \mathbb{N}$, of functions from $\mathbb{R}^n \times \mathcal{J}$ to $\mathbb{R}^n$. The class C_L contains all continuous functions that are Lipschitz continuous with a uniform constant with respect to the first variable. $C^{s,s-1}$ is the class of all functions having continuous partial derivatives up to order $s-1$ and, in addition, continuous partial derivatives of order s with respect to the first variable.
Let $x(\cdot)$ be a solution of the SDE (13) and y be a function in $C^{2,1}$. Then Itô's formula provides the expansion

$$\begin{aligned} y(x(t),t) - y(x_0,t_0) &= \int_{t_0}^{t} \big(y_t + y_x f + \epsilon^2 \frac{1}{2} \sum_{r=1}^{m} \sum_{i,j=1}^{n} y_{x_i x_j} \tilde{g}_{ri} \tilde{g}_{rj}\big)(x(s),s)ds \\ &\quad + \epsilon \sum_{r=1}^{m} \int_{t_0}^{t} (y_x \tilde{g}_r)(x(s),s)dw_r(s),\ t \in \mathcal{J}. \quad (15) \end{aligned}$$

Following [21] we introduce $m+1$ operators Λ_0 and Λ_r, $r = 1, \ldots, m$, defined on $C^{2,1}$ and $C^{1,0}$, respectively, by

$$\Lambda_0 y = y_t + y_x f + \epsilon^2 \frac{1}{2} \sum_{r=1}^{m} \sum_{i,j=1}^{n} y_{x_i x_j} \hat{g}_{ri} \tilde{g}_{rj}\ , \quad \Lambda_r y = y_x \tilde{g}_r\ ,\ r = 1, \ldots, m.$$

Then the Itô formula (15) reads

$$y(x(t),t) - y(x_0,t_0) = \int_{t_0}^{t} \Lambda_0 y(x(s),s)ds + \epsilon \sum_{r=1}^{m} \int_{t_0}^{t} \Lambda_r y(x(s),s)dw_r(s),\ t \in \mathcal{J}. \quad (16)$$

For $y \in C_L$ and similarly as in Section 2 we denote multiple stochastic integrals over intervals $[t, t+h] \subseteq \mathcal{J}$ by

$$I_{i_1 \ldots i_j;t,h}(y) = \int_t^{t+h} \int_t^{s_1} \ldots \int_t^{s_{j-1}} y(x(s_j),s_j)dw_{i_1}(s_j) \ldots dw_{i_{j-1}}(s_2)dw_{i_j}(s_1),$$

where $i_1, \ldots, i_j$ take values in $\{0, \ldots, m\}$, and $dw_0(s)$ is understood to mean ds. As the function y has linear growth with respect to the first variable, the stochastic integrals are well defined.

Lemma 4.1 *For any $p \geq 1$ such that x_0 has finite p-th mean, any $(t,h) \in \mathcal{T}$ and $i_j \in \{1, \ldots, m\}$, $j = 1, \ldots, k$, we have for any function $y \in C_L$ that*

$$\begin{aligned} \mathbb{E}(I_{i_1 \ldots i_k;t,h}(y)|\mathcal{F}_t) &= 0 \quad \text{if } i_j \neq 0 \text{ for some } j \in \{1, \ldots, k\}, \\ \|\mathbb{E}(I_{i_1,\ldots,i_k;t,h}(y)|\mathcal{F}_t)\|_{L_p} &\leq \|I_{i_1,\ldots,i_k;t,h}(y)\|_{L_p} = \mathcal{O}(h^{\sum_{j=1}^{k} \nu_j}),\ \text{where}\ \nu_j = \begin{cases} 1, & i_j = 0, \\ \frac{1}{2}, & i_j \neq 0. \end{cases} \end{aligned}$$

Proof: The first property is well known. The first estimate in the second assertion is due to properties of the conditional expectation. For $p = 2$ the second part is proved in [20,

Lemma 2.1]. For $1 \le p < 2$ it is a consequence of the estimate $\|\cdot\|_{L_p} \le \|\cdot\|_{L_2}$.
Now, let $p > 2$. For $i_1 = 0$ we obtain by Hölder's inequality that

$$\begin{aligned}\|I_{0,i_2,\ldots,i_k;t,h}(y)\|^p_{L_p} &= \mathbb{E}(|I_{0,i_2,\ldots,i_k,0;t,h}(y)|^p) \le \Big(\mathbb{E}\int_t^{t+h} |I_{i_2,\ldots,i_k;t,s_1-t}(y)|ds_1\Big)^p \\ &\le h^{\frac{p}{q}} \int_t^{t+h} \mathbb{E}(|I_{i_2,\ldots,i_k;t,s_1-t}(y)|^p)ds_1,\end{aligned}$$

where $\frac{1}{q} + \frac{1}{p} = 1$. Hence, for $\|I_{0,i_2,\ldots,i_k;t,h}(y)\|^p_{L_p}$ we obtain the order $\mathcal{O}(h^{\frac{p}{q}+1}) = \mathcal{O}(h^p)$.
For $i_1 \ne 0$ we make use of estimates for the p-th mean of stochastic integrals (see [8, Section 1.4, Theorem 6], [16, Section 1.7, Theorem 7.1]) and have

$$\|I_{i_1,\ldots,i_k;t,h}(y)\|^p_{L_p} \le (\frac{1}{2}p(p-1))^{\frac{p}{2}} h^{\frac{p-2}{2}} \int_t^{t+h} \mathbb{E}(|I_{i_2,\ldots,i_k;t,t-s_1}(y)|^p)ds_1.$$

Hence, in this case we obtain the order $\mathcal{O}(h^{\frac{p-2}{2}+1}) = \mathcal{O}(h^{\frac{p}{2}})$.
Repeating these arguments successively and using that the function y has linear growth and, thus, that $y(x(\cdot),\cdot)$ has finite p-th mean completes the proof. □

Proposition 4.2 *Assume that $f \in C^{4,3}$ and $\tilde{g}_r \in C^{2,1}$, $r = 1,\ldots,m$ and that the functions $\Lambda_0\Lambda_0 f$, $\Lambda_0\tilde{g}_r$, $\Lambda_r\Lambda_0 f$, $\Lambda_r f$ and $\Lambda_k\tilde{g}_r$, $k,r = 1,\ldots,m$, belong to C_L.*
Then the local discretization error l_ℓ (see (9)) of the family of drift-implicit Euler schemes (14) at step ℓ allows a decomposition

$$l_\ell = s_\ell + r_\ell \quad \textit{with} \quad \mathbb{E}(s_\ell|\mathcal{F}_{t_{\ell-1}}) = 0$$

and its two components r_ℓ and s_ℓ satisfy the estimates

$$\begin{aligned}\|r_\ell\|_{L_p}/h_\ell &= h_\ell|\theta - \frac{1}{2}|\,\|(\Lambda_0 f)_{t_{\ell-1}}\|_{L_p} + \mathcal{O}(h_\ell^2) \\ \|s_\ell\|_{L_p}/h_\ell^{1/2} &= \epsilon\mathcal{O}(h_\ell) + \epsilon^2\mathcal{O}(h_\ell^{1/2}).\end{aligned}$$

Proof: For $y \in C_L$ we make use of the following abbreviations

$$y_s := y(x(s),s), \quad I_{i_1\ldots i_j}(y) = I_{i_1\ldots i_j;t_{\ell-1},h_\ell}(y).$$

By reformulating the local error and by expanding all of its components at the pair $(x(t_{\ell-1}),t_{\ell-1})$ using (16) and the smoothness properties f, $\tilde{g}_r \in C^{2,1}$, $r = 1,\ldots,m$, we obtain

$$\begin{aligned}l_\ell &= x(t_\ell) - x(t_{\ell-1}) - h_\ell\big(\theta f(x(t_\ell),t_\ell) + (1-\theta)f(x(t_{\ell-1}),t_{\ell-1})\big) - \epsilon\tilde{G}(x(t_{\ell-1}),t_{\ell-1})\Delta w_\ell \\ &= \int_{t_{\ell-1}}^{t_\ell} f_s ds - h_\ell\big(\theta f_{t_\ell} + (1-\theta)f_{t_{\ell-1}}\big) + \int_{t_{\ell-1}}^{t_\ell} \epsilon\tilde{G}_s dw(s) - \epsilon\tilde{G}_{t_{\ell-1}}\Delta w_\ell \\ &= \int_{t_{\ell-1}}^{t_\ell} \Big\{ f_{t_{\ell-1}} + \int_{t_{\ell-1}}^{s} (\Lambda_0 f)_\tau d\tau + \epsilon\sum_{r=1}^m \int_{t_{\ell-1}}^{s} (\Lambda_r f)_\tau dw_r(\tau)\Big\} ds \\ &\qquad -\theta h_\ell\Big\{ f_{t_{\ell-1}} + \int_{t_{\ell-1}}^{t_\ell} (\Lambda_0 f)_\tau d\tau + \epsilon\sum_{r=1}^m \int_{t_{\ell-1}}^{t_\ell} (\Lambda_r f)_\tau dw_r(\tau)\Big\} - (1-\theta)h_\ell f_{t_{\ell-1}} \\ &\qquad +\epsilon\sum_{r=1}^m \int_{t_{\ell-1}}^{t_\ell} \Big\{ \int_{t_{\ell-1}}^{s} (\Lambda_0\tilde{g}_r)_\tau d\tau + \epsilon\sum_{k=1}^m \int_{t_{\ell-1}}^{s} (\Lambda_k\tilde{g}_r)_\tau dw_k(\tau)\Big\} dw_r(\tau)\end{aligned}$$

$$
\begin{aligned}
&= \int_{t_{\ell-1}}^{t_\ell} \Big\{ \int_{t_{\ell-1}}^{s} (\Lambda_0 f)_\tau d\tau + \epsilon \sum_{r=1}^{m} \int_{t_{\ell-1}}^{s} (\Lambda_r f)_\tau dw_r(\tau) \Big\} ds \\
&\quad -\theta h_\ell \Big\{ \int_{t_{\ell-1}}^{t_\ell} (\Lambda_0 f)_\tau d\tau + \epsilon \sum_{r=1}^{m} \int_{t_{\ell-1}}^{t_\ell} (\Lambda_r f)_\tau dw_r(\tau) \Big\} \\
&\quad +\epsilon \sum_{r=1}^{m} \int_{t_{\ell-1}}^{t_\ell} \Big\{ \int_{t_{\ell-1}}^{s} (\Lambda_0 \tilde{g}_r)_\tau d\tau + \epsilon \sum_{k=1}^{m} \int_{t_{\ell-1}}^{s} (\Lambda_k \tilde{g}_r)_\tau dw_k(\tau) \Big\} dw_r(\tau) \\
&= I_{00}(\Lambda_0 f) - \theta h_\ell I_0(\Lambda_0 f) + \epsilon \sum_{r=1}^{m} \big(I_{0r}(\Lambda_r f) - \theta h_\ell I_r(\Lambda_r f) + I_{r0}(\Lambda_0 \tilde{g}_r) \big) + \epsilon^2 \sum_{r,k=1}^{m} I_{rk}(\Lambda_k \tilde{g}_r),
\end{aligned}
$$

and, hence, a representation of the local error in terms of (multiple) stochastic integrals. Next, we study the leading term $I_{00}(\Lambda_0 f) - \theta h_\ell I_0(\Lambda_0 f)$ of this representation. Since $\Lambda_0 f$ belongs to $C^{2,1}$, we may use the Itô formula (16) again and obtain

$$
\Lambda_0 f_\tau = \Lambda_0 f_{t_{\ell-1}} + \int_{t_{\ell-1}}^{\tau} (\Lambda_0 \Lambda_0 f)_s ds + \epsilon \sum_{r=1}^{m} \int_{t_{\ell-1}}^{\tau} (\Lambda_r \Lambda_0 f)_s dw_r(s)), \quad \tau \in [t_{\ell-1}, t_\ell]. \tag{17}
$$

The latter equation (17) is taken to compute the desired (multiple) stochastic integrals and the whole leading term, respectively, i.e.,

$$
\begin{aligned}
I_{00}(\Lambda_0 f) &= \frac{1}{2} h_\ell^2 (\Lambda_0 f)_{t_{\ell-1}} + I_{000}(\Lambda_0 \Lambda_0 f) + \epsilon \sum_{r=1}^{m} I_{00r}(\Lambda_r \Lambda_0 f), \\
I_0(\Lambda_0 f) &= h_\ell (\Lambda_0 f)_{t_{\ell-1}} + I_{00}(\Lambda_0 \Lambda_0 f) + \epsilon \sum_{r=1}^{m} I_{0r}(\Lambda_r \Lambda_0 f), \\
I_{00}(\Lambda_0 f) - \theta h_\ell I_0(\Lambda_0 f) &= h_\ell^2 (\frac{1}{2} - \theta)(\Lambda_0 f)_{t_{\ell-1}} + I_{000}(\Lambda_0 \Lambda_0 f) - \theta h_\ell I_{00}(\Lambda_0 \Lambda_0 f) \\
&\quad + \epsilon \sum_{r=1}^{m} \big(I_{00r}(\Lambda_r \Lambda_0 f) - \theta h_\ell I_{0r}(\Lambda_r \Lambda_0 f) \big).
\end{aligned}
$$

Now, we split $l_\ell = s_\ell + r_\ell$, where s_ℓ is composed by all integral terms with at least one nonzero index. Then we have $\mathbb{E}(s_\ell | \mathcal{F}_{t_{\ell-1}}) = 0$, and

$$
\begin{aligned}
r_\ell &= h_\ell^2 (\frac{1}{2} - \theta)(\Lambda_0 f)_{t_{\ell-1}} + I_{000}(\Lambda_0 \Lambda_0 f) - \theta h_\ell I_{00}(\Lambda_0 \Lambda_0 f), \\
s_\ell &= \epsilon \sum_{r=1}^{m} \big(I_{00r}(\Lambda_r \Lambda_0 f) - \theta h_\ell I_{0r}(\Lambda_r \Lambda_0 f) \big) + \\
&\quad \epsilon \sum_{r=1}^{m} \big(I_{0r}(\Lambda_r f) - \theta h_\ell I_r(\Lambda_r f) + I_{r0}(\Lambda_0 \tilde{g}_r) \big) + \epsilon^2 \sum_{r,k=1}^{m} I_{rk}(\Lambda_k \tilde{g}_r).
\end{aligned}
$$

Since all the functions $\Lambda_0 \Lambda_0 f$, $\Lambda_0 \tilde{g}_r$, $\Lambda_r \Lambda_0 f$, $\Lambda_r f$ and $\Lambda_k \tilde{g}_r$, $k, r = 1, \ldots, m$, appearing as integrands of multiple stochastic integrals satisfy the assumptions of Lemma 4.1, we may use the lemma repeatedly and obtain

$$
\begin{aligned}
\|r_\ell\|_{L_p} / h_\ell &= h_\ell |\theta - \frac{1}{2}| \, \|(\Lambda_0 f)_{t_{\ell-1}}\|_{L_p} + \mathcal{O}(h_\ell^2), \\
\|s_\ell\|_{L_p} / h_\ell^{1/2} &= \epsilon \mathcal{O}(h_\ell) + \epsilon^2 \mathcal{O}(h_\ell^{1/2}).
\end{aligned}
$$

□

The previous result enables us to study the relation between ϵ and the stepsizes h_ℓ. Unless $\theta = 1/2$ for the trapezoidal rule, the dominating term of $\|r_\ell\|_{L_p}/h_\ell$ is

$$h_\ell|\theta - \frac{1}{2}|\,\|(\Lambda_0 f)_{t_{\ell-1}}\|_{L_p} = h_\ell|\theta - \frac{1}{2}|\,\|(f_t + f_x f)_{t_{\ell-1}}\|_{L_p} + \epsilon^2\mathcal{O}(h_\ell).$$

It clearly dominates the terms of order $\epsilon\mathcal{O}(h_\ell)$ in $\|s_\ell\|_{L_p}/h_\ell^{1/2}$. We assume the remaining terms of $\|s_\ell\|_{L_p}/h_\ell^{1/2}$ to satisfy the relation

$$\epsilon^2\|\sum_{r,k=1}^{m} I_{rk}(\Lambda_k\tilde{g}_r)\|_{L_p}/h_\ell^{1/2} << h_\ell\|(f_t + f_x f)_{t_{\ell-1}}\|_{L_p}. \tag{18}$$

The latter condition is valid if

$$\epsilon^2 h_\ell^{1/2} << h_\ell, \quad \text{i.e.,} \quad \epsilon^4 << h_\ell, \tag{19}$$

provided that the values of the functions $\tilde{g}_{rx}\tilde{g}_k$, $r,k = 1,\ldots,m$, and $f_t + f_x f$ are moderate. Then the local error term

$$\eta_\ell := \max\{\|s_\ell\|_{L_p}/h_\ell^{1/2}\,,\ \|r_\ell\|_{L_p}/h_\ell\}$$

is dominated by $\|r_\ell\|_{L_p}/h_\ell$, which itself is dominated by $h_\ell|\theta - \frac{1}{2}|\|(\Lambda_0 f)_{t_{\ell-1}}\|_{L_p}$ or $h_\ell|\theta - \frac{1}{2}|\|(f_t + f_x f)_{t_{\ell-1}}\|_{L_p}$. Hence, for $\theta \neq \frac{1}{2}$ we have

$$\eta_\ell \approx \bar{\eta}_\ell \ := \ h_\ell|\theta - \frac{1}{2}|\,\|(f_t + f_x f)(x(t_{\ell-1}), t_{\ell-1})\|_{L_p}\,. \tag{20}$$

Under the assumptions of Proposition 4.2 we also obtain the expansion

$$f_{t_\ell} - f_{t_{\ell-1}} = h_\ell(\Lambda_0 f)_{t_{\ell-1}} + I_{00}(\Lambda_0\Lambda_0 f) + \epsilon\sum_{r=1}^{m} I_{r0}(\Lambda_r\Lambda_0 f) + \epsilon\sum_{r=1}^{m} I_r(\Lambda_r f),$$

by inserting (17) into Itô's formula (16) for $y = f$ and, thus,

$$h_\ell\|(\Lambda_0 f)_{t_{\ell-1}}\|_{L_p} = \|f_{t_\ell} - f_{t_{\ell-1}}\|_{L_p} + \mathcal{O}(h_\ell^2 + \epsilon h_\ell^{3/2} + \epsilon h_\ell^{1/2}).$$

Under the additional assumption

$$\epsilon\|\sum_{r=1}^{m} I_r(\Lambda_r f)\|_{L_p} << h_\ell\|(\Lambda_0 f)_{t_{\ell-1}}\|_{L_p}, \tag{21}$$

which is true for

$$\epsilon h_\ell^{1/2} << h_\ell, \quad \text{i.e.,} \quad \epsilon^2 << h_\ell, \tag{22}$$

and moderate coefficients $\Lambda_r f$, $r = 0,\ldots,m$, we may proceed to

$$\eta_\ell \approx \bar{\bar{\eta}}_\ell \ := \ |\theta - \frac{1}{2}|\,\|f(x(t_\ell), t_\ell) - f(x(t_{\ell-1}), t_{\ell-1})\|_{L_p}\,. \tag{23}$$

Remark 4.3 *If the length $T-t_0$ of the considered time interval differs considerably from 1, the more detailed stability estimate*

$$\max_{\ell=1,\ldots,N}\|x_\ell^*-\tilde{x}_\ell\|_{L_p}\le\tilde{S}\Big(\|x_0^*-\tilde{x}_0\|_{L_p}+\max_{\ell=1,\ldots,N}\|s_\ell\|_{L_p}(\frac{T-t_0}{h})^{\frac{1}{2}}+\max_{\ell=1,\ldots,N}\|r_\ell\|_{L_p}\frac{T-t_0}{h}\Big)\quad(24)$$

(cf. [23]) should be used as a starting point. Since the length of the interval affects the local error terms differently, condition (18) modifies to

$$\epsilon^2\|\sum_{r,k=1}^{m}I_{rk}(\Lambda_k\hat{g}_r)\|_{L_p}(\frac{T-t_0}{h_\ell})^{\frac{1}{2}}<<(T-t_0)h_\ell\|(f_t+f_xf)_{t_{\ell-1}}\|_{L_p},\tag{25}$$

which is true for

$$\epsilon((T-t_0)h_\ell)^{\frac{1}{2}}<<(T-t_0)h_\ell,\qquad i.e.,\quad \epsilon^2<<(T-t_0)h_\ell,\tag{26}$$

and moderate coefficients $\tilde{g}_{rx}\tilde{g}_k$, $r,k=1,\ldots,m$, and (f_t+f_xf).

Remark 4.4 *The conditions (25), (26), (21), (22) are independent of the used time-scale. A transformation of the time scale $t\in[t_0,T]$ to $\tau\in[0,1]$ via*

$$\tau=(t-t_0)/(T-t_0),\ t(\tau)=t_0+(T-t_0)\tau,\quad y(\tau):=x(t(\tau)),\ \hat{w}(\tau):=(T-t_0)^{-\frac{1}{2}}w(t(\tau))$$

leads to the transformed SDE

$$y(s)\Big|_0^\tau=(T-t_0)\int_0^\tau f(y(s),t(s))ds+\epsilon(T-t_0)^{\frac{1}{2}}\int_{t_0}^\tau\tilde{G}(y(s),t(s))d\hat{w}(s),\ \tau\in[0,1].$$

The conditions (25), (26), (21), (22) in terms of the transformed functions and variables

$$\hat{f}(y,\tau)=(T-t_0)f(y,t(\tau)),\quad \hat{\tilde{G}}(y,\tau)=(T-t_0)^{\frac{1}{2}}\tilde{G}(y,t(\tau)),\quad \hat{h}=h/(T-t_0),\quad \hat{T}=1,$$

coincide with the original conditions. □

Remark 4.5 *The simple conditions (19), (26), and (22) together with the condition of moderate function values describe rather rules of thumb for the used stepsizes. We specify them for $p=2$, a scalar Wiener process, and the diffusion coefficient $G=(g)$. Neglecting higher order terms in (18),(25) and (21) we then obtain the conditions*

$$h_\ell>>\frac{\|(g_xg)_{t_{\ell-1}}\|_{L_2}^2}{2\|(f_t+f_xf)_{t_{\ell-1}}\|_{L_2}^2},\tag{27}$$

$$(T-t_0)h_\ell>>\frac{\|(g_xg)_{t_{\ell-1}}\|_{L_2}^2}{2\|(f_t+f_xf)_{t_{\ell-1}}\|_{L_2}^2},\tag{28}$$

and

$$h_\ell>>\frac{\|(f_xg)_{t_{\ell-1}}\|_{L_2}^2}{\|(f_t+f_xf+\frac{1}{2}f_{xx}gg)_{t_{\ell-1}}\|_{L_2}^2},\tag{29}$$

in terms of the ratio of the coefficients.

Proof: *For $m=1$ the condition (18) simplifies to $h_\ell^{-\frac{1}{2}}\|I_{11}(g_xg)\|_{L_p}<<h_\ell\|(f_t+f_xf)_{t_{\ell-1}}\|_{L_p}$. Choosing $p=2$ and neglecting higher order terms one obtains*

$$h_\ell^{-\frac{1}{2}}h_\ell 2^{-\frac{1}{2}}\|(g_xg)_{t_{\ell-1}}\|_{L_2}<<h_\ell\|(f_t+f_xf)_{t_{\ell-1}}\|_{L_2},$$

i.e., (27). Analogous arguments apply to (25) and (21). □

4.2 Computable estimates for the local error

When estimating the dominating local error term $\bar{\eta}_\ell$ from (20) in a variable stepsize implementation, one has to approximate the L_p-norm $\|(f_t + f_x f)(x(t_{\ell-1}), t_{\ell-1})\|_{L_p}$ by using a finite number of approximate solution paths.
Let us assume that an ensemble of M approximate solution paths is computed simultaneously and that $x_l^1, \dots, x_l^M$, $l = 1, \dots, \ell - 1$, is given after $\ell - 1$ steps.
If the derivatives f_t, f_x of the drift coefficient f with respect to time t and state x are available, we may form

$$\eta_\ell \approx \tilde{\eta}_\ell := h_\ell |\theta - \frac{1}{2}| \Big(\frac{1}{M} \sum_{i=1}^{M} |(f_t + f_x f)(x^i_{\ell-1}, t_{\ell-1})|^p \Big)^{\frac{1}{p}}, \tag{30}$$

and use Algorithm 3.1 with $\bar{\gamma} = 1$,$\tilde{k}^i_\ell = |\theta - \frac{1}{2}|(f_t + f_x f)(x^i_{\ell-1}, t_{\ell-1})$. The estimate (30) and the corresponding tests in Algorithm 3.1 have the great advantage that they can be performed a priori, before computing x^i_ℓ, $i = 1, \dots, M$, at the next time step. The only possible step rejections are then due to failures in solving the nonlinear equations for determining x^i_ℓ, $i = 1, \dots, M$.
Of course, it is more easy to compute the derivative-free estimate

$$\eta_\ell \approx \bar{\bar{\eta}}_\ell := |\theta - \frac{1}{2}| \Big(\frac{1}{M} \sum_{i=1}^{M} |f(x^i_\ell, t_\ell) - f(x^i_{\ell-1}, t_{\ell-1})|^p \Big)^{\frac{1}{p}} \tag{31}$$

according to (23) and to use Algorithm 3.1 with $\bar{\gamma} = 1$, $\tilde{k}^i_\ell = \frac{\frac{1}{2} - \theta}{h_\ell} \big(f(x^i_\ell, t_\ell) - f(x^i_{\ell-1}, t_{\ell-1}) \big)$.

5 Test results

The stepsize strategy proposed in Section 4 has been implemented for the drift-implicit Euler scheme and tested extensively for $p = 2$ on a set of SDEs with small noise. We report results for two scalar SDEs with known analytic solution and for a low-dimensional electric circuit model. We draw some conclusions on the potential and on the limitations of the strategy.

Example 5.1 *(linear homogeneous SDE with constant coefficients)*
We consider the linear scalar SDE in complex arithmetic

$$x(t) = x_0 + \int_0^t \alpha x(s) ds + \int_0^t i\beta x(s) dw(s), \; t \in [0, 1], \tag{32}$$

with coefficients $f(x,t) := \alpha x$, $G(x,t) = (g(x,t)) = (i\beta x)$, initial value x_0, parameters $\alpha, \beta \in I\!R$ and a scalar Wiener process w. Its solution is given by the geometric Brownian motion

$$x(t) = x_0 \exp\big((\alpha + \frac{1}{2}\beta^2)t + i\beta w(t) \big)$$

and the conditions (27) and (29) are equivalent to

$$\frac{\beta^4}{2\alpha^4} << h_\ell \quad \text{and} \quad \frac{\beta^2}{\alpha^2} << h_\ell\,,$$

respectively. Here, the ratio $|\frac{\beta}{\alpha}|$ plays the role of the small parameter ϵ. As long as stepsizes with $\frac{\beta^4}{2\alpha^4} << h_\ell$ are used, the Euler scheme shows order 1 of convergence. As long as even $\frac{\beta^2}{\alpha^2} << h_\ell$ is satisfied, the proposed stepsize control should act properly. In regions where the first condition is satisfied, but the second one is violated, the controlled quantity $\|f_\ell - f_{\ell-1}\|_{L_2}$ is dominated by

$$\| \int_{t_{\ell-1}}^{t_\ell} f_x g(x(s), s) dw(s) \|_{L_2} \approx |\alpha\beta| \, \|x_{\ell-1}\|_{L_2} h^{\frac{1}{2}}$$

instead of

$$\| \int_{t_{\ell-1}}^{t_\ell} f_x f(x(s), s) ds \|_{L_2} \approx \alpha^2 \|x_{\ell-1}\|_{L_2} h.$$

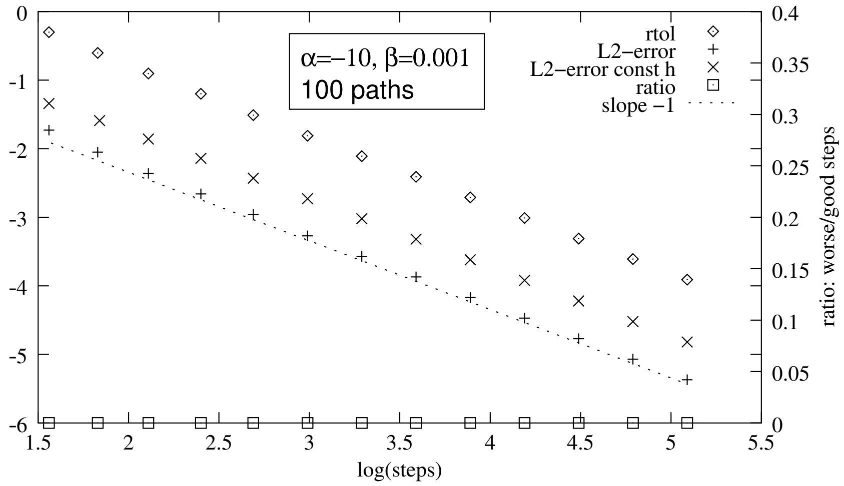
If sufficiently many paths are computed to approximate the expectations that form the norms in L_2, the proposed control leads to stepsizes that match

$$\alpha^2 \|x_{\ell-1}\|_{L_2} h \approx \mathrm{tol}^2 / (|\beta| \, \|x_{\ell-1}\|_{L_2}). \tag{33}$$

In the following we present results for different values of the parameters α, β. We plot the used relative tolerance ($\diamond$) and the achieved accuracy with adaptively chosen stepsizes ($+$) and with constant stepsizes ($\times$) versus the number of steps in logarithmic scale. The accuracy is measured as the maximum approximate L_2-norm of the global errors in the time-interval $[0, 1]$:

$$\max_{\ell=1,\ldots,N} \left(\frac{1}{M} \sum_{j=1}^{M} |x(t_\ell, \omega_j) - x_\ell(\omega_j)|^2 \right)^{1/2},$$

where N denotes the number of steps and M the number of computed paths. Lines with slopes -1 and -0.5 are provided to enable comparisons with convergence of order 1 or 1/2. The scale at the right-hand sides of the diagrams measures the ratio of rejected to accepted steps ($\square$). Large values of this ratio indicate that the proposed stepsize control is inadequate for the problem and the used tolerance.



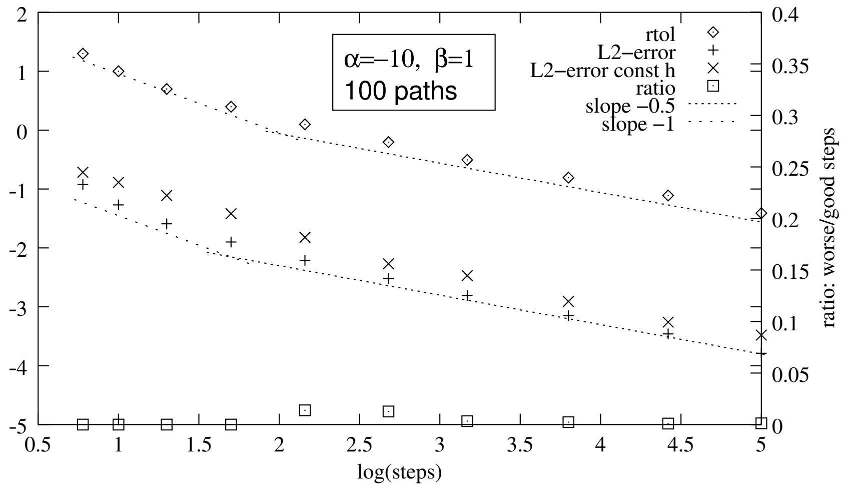
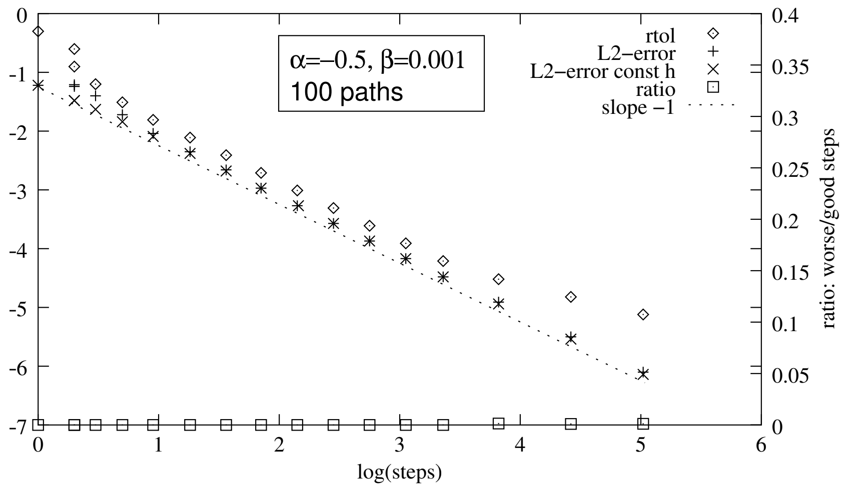


Figure 1: Simulation results for the SDE (32) for $\alpha = -10$ and $\beta = 10^{-3}$ and 1

Figure 1 shows simulation results for the parameter $\alpha = -10$ and an ensemble of 100 simultaneously computed paths. We observe that the error estimate is too pessimistic for both values of β. This is due to the damping of the solution by $\alpha = -10$, since the errors are damped as well. Further we observe that, for the small noise parameter $\beta = 10^{-3}$, the proposed stepsize control works well up to the accuracy 10^{-5}, and gives more accurate results than the computation with the same number of constant steps. This accuracy range decreases to approximately 10^{-2} as the noise parameter increases to $\beta = 1$.
Figure 2 shows simulation results for the parameter $\alpha = -0.5$ once for an ensemble of 100 simultaneously computed paths and once for a single path only. The achieved accuracies show a good correspondence with the tolerances as long as they are not too small. For 100 simultaneously computed paths one clearly observes the three different regimes, though not all for one choice of the parameters: For low accuracy we observe order 1 convergence and properly acting stepsize control. For medium accuracy we observe order 1 convergence but stepsizes that match (33). For high accuracy we observe convergence of order 1/2. If only a single path is computed, only the results for low accuracies are acceptable. For medium accuracy the large ratio of rejected to accepted steps indicates that the stepsize control does not act properly, for high accuracy or large noise the convergence gets even lost completely.



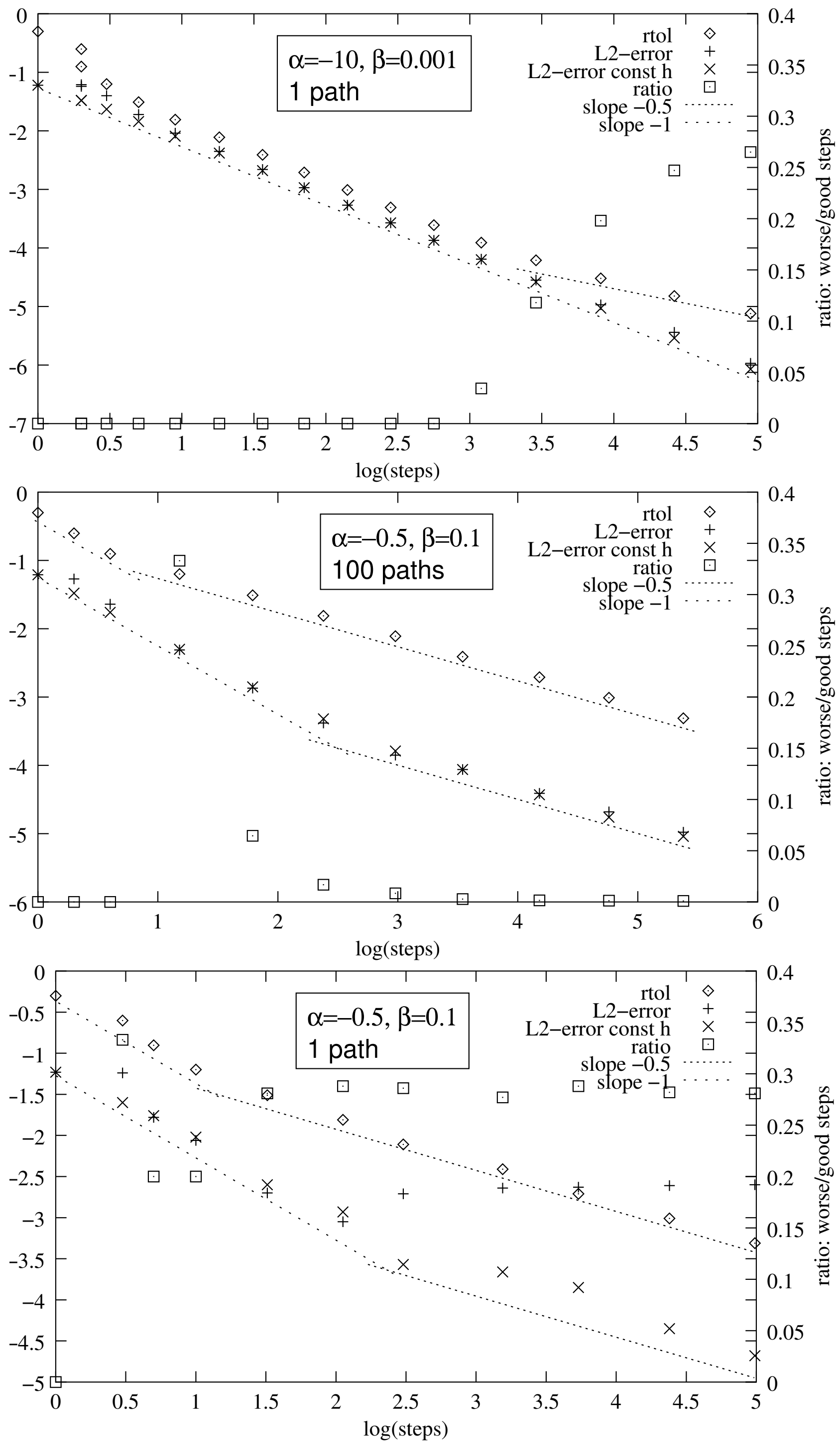


Figure 2: Simulation results for the SDE (32) for $\alpha = -0.5$ and $\beta = 10^{-3}$ and 1

Example 5.2 *(SDE with polynomial drift and diffusion)*
We consider the scalar SDE

$$x(t) = \int_0^t -(\alpha + \beta^2 x(s))(1 - x(s)^2)ds + \int_0^t \beta(1 - x(s)^2)dw(s)\,,\ t \in [0, T], \tag{34}$$

with coefficients $f(x,t) := -(\alpha + \beta^2 x)(1 - x^2)$, $G(x,t) = (g(x,t)) = (\beta(1 - x^2))$, real parameters α, β and a scalar Wiener process w. The solution is given by (cf. [13, (4.46)])

$$x(t) = \frac{\exp(-2\alpha t + 2\beta w(t)) - 1}{\exp(-2\alpha t + 2\beta w(t)) + 1}.$$

Due to the nonlinearity of the coefficients the conditions (27), (28), (29) are solution-dependent. Another effect of the nonlinearity are restrictions of the stepsize to ensure the convergence of Newton's method for solving the nonlinear equations of the drift-implicit Euler scheme in every step. Failures of Newton's method may also cause step rejections. In Figure 3 we present simulation results for the parameters $\alpha = -10$ and $\beta = 10^{-2}$ or $\beta = 10^{-1}$. We have computed 100 paths simultaneously and plotted the same quantities as already used in Example 5.1. For $\beta = 10^{-2}$ we observe order 1 behaviour up to accuracies of 10^{-5}, whereas this range decreases to accuracies of 10^{-3} for $\beta = 10^{-1}$. The parameter $\alpha = -10$ causes a damping in the solution and a prediction of the global error that is much too pessimistic. In both cases the proposed stepsize control provides considerably more accurate results than the computation with the same number of constant steps.
The observed stepsize rejections for larger stepsizes are mainly due to Newton failures. Failures of Newton's methods also prevented simulation results for larger stepsizes in case of larger noise $\beta = 1$.

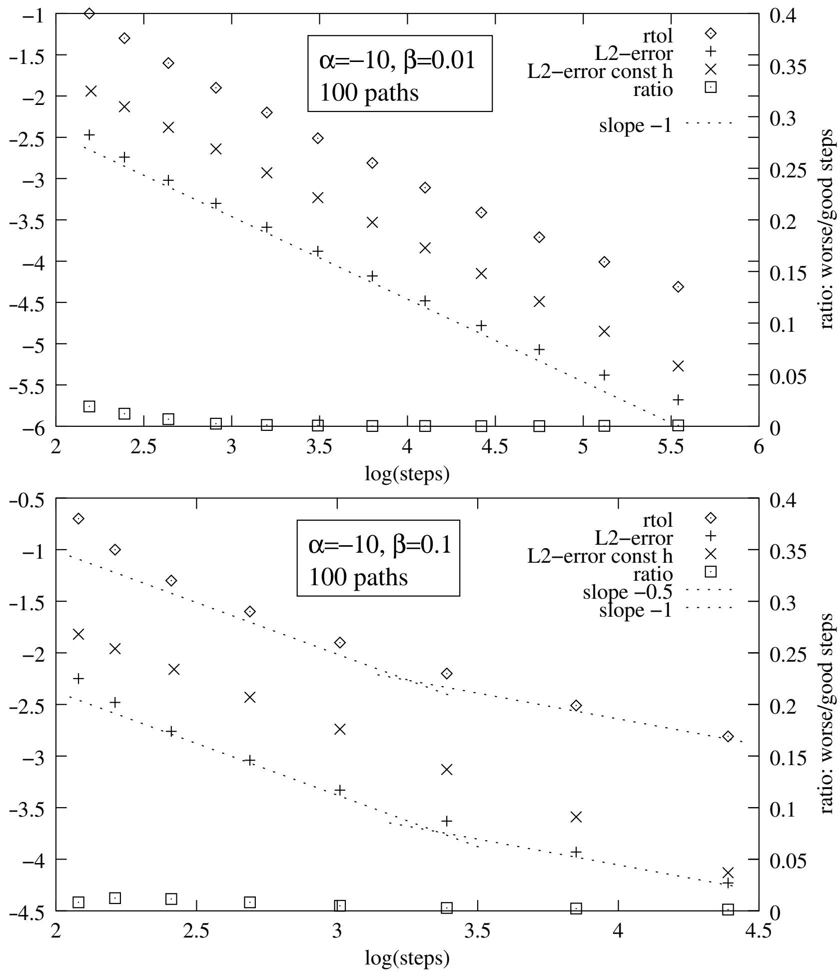


Figure 3: Simulation results for the SDE (34) for $\alpha = -10$ and $\beta = 10^{-2}$ and 10^{-1}

Example 5.3 *(linear electrical circuit)*
We consider a linear RC-oscillator circuit under the influence of thermal noise of the resistor (cf. [24] and Fig. 4). The thermal noise of the resistor is modelled as an additional current source in parallel to the linear resistor with the noise intensity given by Nyquist's

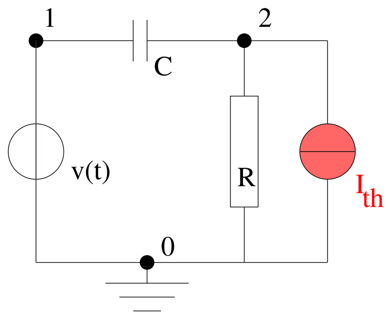


Figure 4: Linear RC-oscillator with thermal noise

rule (cf. [2, 5, 26]):

$$I_{th} = \sigma_{th} \cdot \xi(t) = \sqrt{\tfrac{2k\text{Temp}}{R}} \cdot \xi(t)\ ,$$

where $\xi(t)$ is a standard Gaussian white noise process, $k = 1.38066 \cdot 10^{-23}$ $[JK^{-1}]$ is Boltzmann's constant, Temp is the absolute temperature, R is the resistance. For the simulations we considered the time interval

$$t \in [0, 2.5 \cdot 10^{-8}]$$

and fixed the parameters to

$$v(t) = \sin(2 \cdot 10^8 \cdot t)\ [V], \quad C = 10^{-12}\ [F], \quad R = 10^4\ [\Omega], \quad \text{Temp} = 300\ [^oK].$$

We denote the nodal potentials in node i by e_i, $i = 0, 1, 2$, and by combining Kirchhoff's Current Law and the element characteristics we obtain

$$\begin{aligned} e_0 &= 0 \\ e_1 &= v(t) \\ -C(\dot{e}_1 - \dot{e}_2) + \frac{1}{R} e_2 + \sigma_{th}\xi(t) &= 0\,. \end{aligned}$$

While the nodal potentials e_0, e_1 are given explicitly, e_2 is given by the SDE

$$\begin{aligned} e_2(t) &= \int_0^t \big(-\frac{1}{RC} e_2(s) + \dot{v}(s) \big) ds + \int_0^t -\frac{\sigma_{th}}{C} dw(s) \\ &= \int_0^t \big(-10^8 e_2(s) + 2 \cdot 10^8 \cos(2 \cdot 10^8 s) \big) ds - \int_0^t \frac{\sqrt{2 \cdot 1.38066 \cdot 10^{-23} \cdot 300}}{10^{-12}\sqrt{10^4}} dw(s), \\ &= 10^8 \int_0^t \big(-e_2(s) + 2\cos(2 \cdot 10^8 s) \big) ds + \sqrt{0.6 \cdot 1.38066} \int_0^t dw(s), t \in [0, 2.5 \cdot 10^{-8}]. \end{aligned}$$

As the noise in this SDE model is too small, we present numerical results for a modified model where the noise intensity was scaled artificially by a factor of 10^3 in order to cause visible differences from the deterministic solution. Since the exact solution is not known to us, we use a numerical approximation that was obtained with the implicit Euler scheme and constant small stepsizes as a reference solution for estimating the global errors. The simulation results are presented in Fig. 5. The first picture shows one path and the mean of 100 computed paths of the solution e_2 vs. time t. The second one shows the used relative tolerance ($\diamond$) and the achieved accuracy with variable ($+$) and with constant stepsizes ($\times$) vs. the number of steps in logarithmic scale. The accuracy is measured as the maximum approximate L_2-norm of the global errors that was achieved at 20 equally distributed reference points in the time interval $[0, 2.5{\cdot}10^{-8}]$. In this example, the differences between the accuracy for constant and variable stepsizes are negligible. In principle, Fig. 5 shows a good accordance of the used relative tolerances and the achieved accuracies for the variable stepsize strategy. The accuracy is only slightly worse than the tolerances. To achieve this accordance it was essential to include the length of the time-interval $T = 2.5 \cdot 10^{-8}$ into the estimate $\tilde{\eta}_\ell$ for the local error per unit step (cf. Remark 4.3).

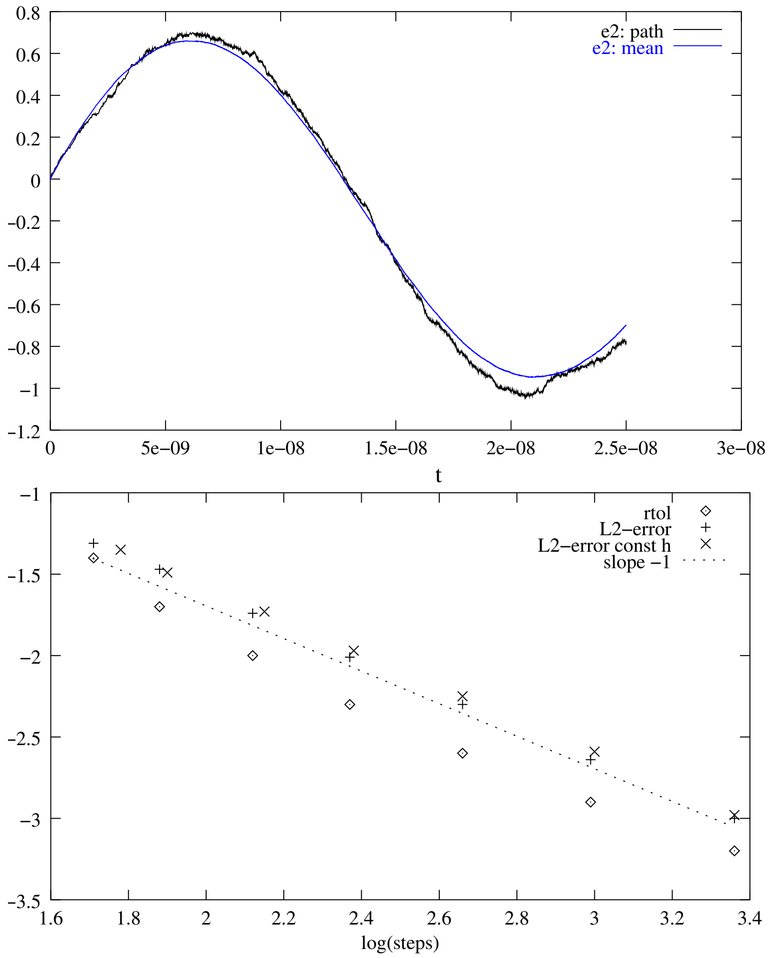


Figure 5

We conclude that the proposed stepsize strategy will work very well for SDEs with small noise if the ratio of the relative tolerances to the smallness of the noise is reasonable in the sense that the inequality $\epsilon^2 << h_\ell$ is valid. In that case, the stepsize strategy provides more accurate results compared to using equidistant grids. It also turns out that the approximate computation of the L_2-norm based on an ensemble of paths compared to computing only one path is important at least in regions of larger noise.

Acknowledgment: The research of the authors was supported by the German Ministry for Education, Science, Research and Technology under the grant 03ROM3B3.

References

[1] T. A. Averina, S.S. Artiemiev, and H. Schurz. Simulation of stochastic auto-oscillating systems through variable stepsize algorithms with small noise. Weierstraß-Institut für Angewandte Analysis und Stochastik, Berlin, Preprint 116, 1994.

[2] A. Blum. *Elektronisches Rauschen.* Teubner, 1996.

[3] K. Burrage, P. Burrage, and T. Mitsui. Numerical solutions of stochastic differential equations - implementation and stability issues. *J. Comp. Appl. Math.* 125 (2002), pp. 171–182.

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

[5] A. Demir and A. Sangiovanni-Vincentelli. *Analysis and Simulation of Noise in Nonlinear Electronic Circuits and Systems.* Kluwer, Boston 1998.

[6] G. Denk and S. Schäffler. Adams methods for the efficient solution of stochastic differential equations with additive noise, *Computing* 59 (1997), pp. 153–161.

[7] J.G. Gaines and T.J. Lyons. Variable stepsize control in the numerical solution of stochastic differential equations. *SIAM J. Appl. Math.* 57 (1997), pp. 1455–1484.

[8] I.I. Gichman and A.V. Skorohod. *Stochastic Differential Equations* (in Russian). Naukova Dumka, Kiev, 1968.

[9] K. Gustafsson, M. Lundh, and G. Söderlind. A PI stepsize control for the numerical integration of ordinary differential equations. *BIT* 28 (1988), pp. 270–287.

[10] E. Hairer, S.P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I.* Springer, Berlin, 1987.

[11] N. Hofmann, T. Müller-Gronbach and K. Ritter. Optimal approximation of stochastic differential equations by adaptive step-size control. *Math. Comp.* 69 (2000), pp. 1017–1034.

[12] N. Hofmann, T. Müller-Gronbach and K. Ritter. Step-size control for the uniform approximation of stochastic differential equations with additive noise. *Ann. Appl. Prob.* 10 (2000), pp. 616–633.

[13] P.E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations.* Springer, Berlin 1992.

[14] H. Lamba, J. Mattingly, A. Stuart. Strong convergence of an adaptive Euler-Maruyama scheme for stochastic differential equations. Manuscript, submitted, 2003.

[15] J. Lehn, A. Rößler and O. Schein. Adaptive schemes for the numerical solution of SDEs – a comparison. *J. Comp. Appl. Math.* 138 (2002), pp. 297-308.

[16] X. Mao. *Stochastic Differential Equations and Their Applications.* Horwood Publishing, Chichester, 1997.

[17] S. Mauthner. Step size control in the numerical solution of stochastic differential equations. *J. Comput. Appl. Math.* 100 (1998), pp. 93–109.

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

[19] G.N. Milstein. Theorem on the order of convergence for mean-square approximations of solutions of stochastic differential equations. *Theory Probab. Appl.* 32 (1987), pp. 738-741.

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

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

[22] T. Müller-Gronbach. *Strong approximation of systems of stochastic differential equations.* Habilitation thesis, Technische Universität Darmstadt, 2002.

[23] W. Römisch and R. Winkler. Stochastic DAEs in circuit simulation. Preprint 03-03, Institut für Mathematik, Humboldt-Universität Berlin, 2003 and in Proceedings of the Oberwolfach Conference *Mathematical Modelling, Simulation and Optimization of Integrated Electrical Circuits*, Birkhäuser (to appear).

[24] O. Schein. *Stochastic differential algebraic equations in circuit simulation*, Ph.D. thesis, Technische Universität Darmstadt, 1999.

[25] G. Söderlind. The automatic control of numerical integration. *CWI Quarterly* 11 (1998), pp. 55–74.

[26] L. Weiß and W. Mathis. A thermodynamical approach to noise in nonlinear networks. *Int. J. Circ. Theor. Appl.* 26 (1998), 147–165.

[27] R. Winkler. Stochastic differential algebraic equations and applications in circuit simulation. *J. Comp. Appl. Math.* 157 (2003), pp. 477-505.