

# STEPWISE CONTROL FOR MEAN-SQUARE NUMERICAL METHODS FOR STOCHASTIC DIFFERENTIAL EQUATIONS WITH SMALL NOISE \*

WERNER RÖMISCH<sup>†</sup> AND RENATE WINKLER<sup>†</sup>

**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 stepsize sequences that are identical for all computed 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.

**Key words.** stepsize control, stochastic differential equations, small noise, mean-square numerical methods

**AMS subject classifications.** 65C30,60-80

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

$$x(t) = x_0 + \int_{t_0}^t f(x(s), s)ds + \int_{t_0}^t G(x(s), s)dw(s), \quad t \in \mathcal{J}, \quad (1.1)$$

where  $\mathcal{J} = [t_0, T]$ ,  $f : \mathbb{R}^n \times \mathcal{J} \rightarrow \mathbb{R}^n$ ,  $G : \mathbb{R}^n \times \mathcal{J} \rightarrow \mathbb{R}^{n \times m}$  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 (1.1) based on time discretization. Our work is motivated by practical SDE models in circuit simulation [24, 27, 28] 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.

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 [18, 19] and [4] are also based on a comparison of two Runge-Kutta schemes of different order. In [15] conditions are provided that imply mean square convergence of the Euler-Maruyama scheme with pathwise different stepsize

---

\*This work was supported by the German Ministry for Education, Science, Research and Technology under the grant 03ROM3B3 and by the DFG Research Center *MATHEON Mathematics for Key Technologies* in Berlin.

<sup>†</sup>Humboldt-Universität zu Berlin, Institut für Mathematik, 10099 Berlin, Germany

sequences. A different approach was developed in [12, 13, 23], where the authors obtain stepsize sequences that are (mean square and  $p$ -th mean, respectively) optimal for asymptotically small stepsizes.

In contrast to the above approaches we present a stepsize control that is based on estimates of the mean square or the  $p$ -th mean of the local discretization error. This 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 and the exact initial value is given. In particular, we analyze the errors for the family of Euler-Maruyama schemes in case of small noise. The local errors are represented in terms of stochastic integrals. The terms containing multiple stochastic integrals become so small that they are negligible for realistic stepsizes, and the low asymptotic order of convergence  $1/2$  of the Euler-Maruyama schemes is observed only for stepsizes that are far too small to be used. We provide estimates for the mean square or  $p$ -th mean of the dominating local error term that does not cost additional evaluations of the coefficients of the SDE or their derivatives.

Implementing a numerical scheme for the approximate integration of SDEs requires also a discretization of the sample space. One can compute only a finite number of paths. To implement the stepsize control we used a heuristic approach where the mean-square of the local terms was approximated by the information available from an ensemble of approximate solution paths that is computed simultaneously. This way our approach leads to stepsize sequences that are identical for all computed paths.

The stepsize strategy was implemented for the drift-implicit Euler method and extensively tested on a set of test examples. The choice of this drift-implicit method allows us to study the effects of the stepsize selection on the accuracy, i.e., the global discretization error, and on the convergence behavior of Newton's method for solving the implicit nonlinear equations simultaneously. In case of step rejections, the method described in [19] 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. As expected by the analysis 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]. In the test examples we have used deterministic initial values. In general we think of applications where the initial value has at least small variance. For deterministic differential equations with random initial values a pathwise stepsize-control should be more efficient.

Our paper is organized as follows. In §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 §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 §4. Finally, in §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

$$x_\ell = x_{\ell-1} + \varphi(x_{\ell-1}, x_\ell; t_{\ell-1}, h_\ell) + \psi(x_{\ell-1}; t_{\ell-1}, h_\ell, I_{t_{\ell-1}, h_\ell}), \quad \ell = 1, \dots, N, \quad (2.1)$$

for solving (1.1) on the deterministic grid  $t_0 < t_1 < \dots < t_N = T$  with stepsizes  $h_\ell := t_\ell - t_{\ell-1}$ ,  $\ell = 1, \dots, N$ . Here,  $\varphi$  and  $\psi$  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 form

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

where the indices  $i_1, \dots, i_k$  are in  $\{0, 1, \dots, m\}$ ,  $k$  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  $\theta$ -methods, is of the form

$$x_\ell := x_{\ell-1} + h_\ell(\theta f(x_\ell, t_\ell) + (1-\theta)f(x_{\ell-1}, t_{\ell-1})) + G(x_{\ell-1}, t_{\ell-1})\Delta w_\ell, \quad (2.2)$$

$\ell = 1, \dots, N$ , 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  $g_j(z, t)$ ,  $j = 1, \dots, m$ , 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  $\varphi$ , and  $k_{max} := 2$ ,  $m_I := m + m^2$ , and

$$\psi(z; t, h, I_{t,h}) := G(z, t)\Delta w_{t,h} + \sum_{j=1}^m (g_{j_x} G)(z, t) I_{(j); t, h}, \quad (2.3)$$

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 (2.1), which extends the fundamental result of Milstein [20, 21]. Its proof is given in [24, 27].

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

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

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

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

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

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

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

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

for some constant  $L_3 > 0$ .

Then for all  $a \geq 1$  there exist a maximal stepsize  $h^0 > 0$  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  $h := \max_{\ell=1, \dots, N} h_\ell \leq h^0$  and  $h \cdot N \leq a \cdot (T - t_0)$ :

For all  $\mathcal{F}_{t_0}$ -measurable random vectors  $x_0^*, \tilde{x}_0$  having finite  $p$ -th mean, for all  $\ell \in \{1, \dots, N\}$  and  $\mathcal{F}_{t_\ell}$ -measurable perturbations  $d_\ell^*, \tilde{d}_\ell$  having finite  $p$ -th mean, the perturbed discrete system

$$\tilde{x}_\ell = \tilde{x}_{\ell-1} + \varphi(\tilde{x}_{\ell-1}, \tilde{x}_\ell; t_{\ell-1}, h_\ell) + \psi(\tilde{x}_{\ell-1}; t_{\ell-1}, h_\ell, I_{t_{\ell-1}, h_\ell}) + \tilde{d}_\ell, \quad \ell = 1, \dots, N, \quad (2.4)$$

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  $\{\tilde{d}_\ell\}_{\ell=1}^N$  and any splittings of  $d_\ell := d_\ell^* - \tilde{d}_\ell$  such that  $d_\ell = s_\ell + r_\ell$  with  $\mathbb{E}(s_\ell | \mathcal{F}_{t_{\ell-1}}) = 0$ :

$$\mathbb{E} \max_{\ell=1, \dots, N} |x_\ell^* - \tilde{x}_\ell|^p \leq S^p \left( \mathbb{E}|x_0^* - \tilde{x}_0|^p + \frac{\max_{\ell=1, \dots, N} \mathbb{E}|s_\ell|^p}{h^{\frac{p}{2}}} + \frac{\mathbb{E} \max_{\ell=1, \dots, N} |r_\ell|^p}{h^p} \right), \quad (2.5)$$

$$\max_{\ell=1, \dots, N} \mathbb{E}|x_\ell^* - \tilde{x}_\ell|^p \leq S^p \left( \mathbb{E}|x_0^* - \tilde{x}_0|^p + \frac{\max_{\ell=1, \dots, N} \mathbb{E}|s_\ell|^p}{h^{\frac{p}{2}}} + \frac{\max_{\ell=1, \dots, N} \mathbb{E}|r_\ell|^p}{h^p} \right). \quad (2.6)$$

Extracting the  $p$ -th root from (2.6) yields the stability inequality

$$\max_{\ell=1, \dots, N} \|x_\ell^* - \tilde{x}_\ell\|_{L_p} \leq S \left( \|x_0^* - \tilde{x}_0\|_{L_p} + \max_{\ell=1, \dots, N} \|s_\ell\|_{L_p} / h^{\frac{1}{2}} + \max_{\ell=1, \dots, N} \|r_\ell\|_{L_p} / h \right). \quad (2.7)$$

The scheme (2.1) is called  *$p$ -th mean stable* if it satisfies the properties (2.5) and (2.6), respectively, in the above result. Furthermore, it is called  *$p$ -th mean consistent* of order  $\gamma > 0$  if the *local discretization error*  $l_\ell$  at step  $\ell$ , i.e.,

$$l_\ell := x(t_\ell) - x(t_{\ell-1}) - \varphi(x(t_{\ell-1}), x(t_\ell); t_{\ell-1}, h_\ell) - \psi(x(t_{\ell-1}); t_{\ell-1}, h_\ell, I_{t_{\ell-1}, h_\ell}), \quad (2.8)$$

satisfies the properties

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

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  $e_\ell$  we denote the difference

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

of the exact and approximate solution at time  $t_\ell$ . The discretization scheme (2.1) for (1.1) is called  *$p$ -th mean convergent* of order  $\gamma > 0$  if the global error  $e_\ell := x(t_\ell) - x_\ell$  satisfies the property

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

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

**THEOREM 2.2.** *If the discretization scheme (2.1) for (1.1) is  $p$ -th mean consistent of order  $\gamma > 0$  and the assumptions of Theorem 2.1 are satisfied, then (2.1) is  $p$ -th mean convergent of order  $\gamma$ . 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 (2.4) we have the estimate*

$$\max_{\ell=1,\dots,N} \|x(t_\ell) - \tilde{x}_\ell\|_{L_p} \leq S((c + \bar{c})h^\gamma + \max_{\ell=1,\dots,N} \delta_\ell/h^{1/2} + \max_{\ell=1,\dots,N} \bar{\delta}_\ell/h), \quad (2.9)$$

where  $\delta_\ell := \|\tilde{d}_\ell\|_{L_p}$ ,  $\bar{\delta}_\ell := \|\mathbb{E}(\tilde{d}_\ell | \mathcal{F}_{t_{\ell-1}})\|_{L_p}$ , with  $\tilde{d}_\ell$  from (2.4).

*Proof.* The assertion follows by applying the triangle inequality

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

and the stability estimate (2.5) once to  $x(t_\ell)$  related to the perturbations  $l_\ell$  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 (2.9) for  $d_\ell^* = l_\ell$ ,  $\tilde{d}_\ell = 0$ . By means of  $r_\ell = \mathbb{E}(l_\ell | \mathcal{F}_{t_{\ell-1}})$ , a suitable splitting  $d_\ell = l_\ell = s_\ell + r_\ell$  is realized.  $\square$

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 (2.1). 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  $g_{j_x} G$ ), and (A4) is a more technical assumption, which is satisfied since the function  $G(0, \cdot)$  (and the functions  $(g_{j_x} G)(0, \cdot)$ ) 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 (2.2) are  $p$ -th mean stable. If, in addition, the partial derivatives  $g_{j_x}$ ,  $j = 1, \dots, m$ , exist and the functions  $g_{j_x} \cdot G$  are Lipschitz continuous with respect to  $x$ , then the Milstein schemes (2.3) are  $p$ -th mean stable.*

From the literature (see e.g. [21]) it is known that the Euler schemes (2.2) are mean-square consistent of 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 of 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 of order 1/2 and of the Milstein schemes of 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, 16, 18, 19]. 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 computed paths.

By means of the stability inequality (2.7) 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  $l_\ell$ . Here, we assume that the initial value  $x_0$  is given

exactly. Therefore, a natural approach consists in controlling the local error term

$$\eta_\ell := \max\{\|s_\ell\|_{L_p}/h_\ell^{1/2}, \|r_\ell\|_{L_p}/h_\ell\}, \text{ where } l_\ell = s_\ell + r_\ell, \mathbb{E}(s_\ell|\mathcal{F}_{t_{\ell-1}}) = 0, \quad (3.1)$$

according to a given tolerance. Controlling this term requires some insight into its behavior. Clearly, we have  $\eta_\ell = \mathcal{O}(h_\ell^\gamma)$  for a method that is  $p$ -th mean convergent of order  $\gamma$ . However, for problems with small noise and stepsizes that cannot be considered asymptotically small and are of interest,  $\eta_\ell$  may even be dominated by a term  $\kappa_\ell \cdot h_\ell^{\bar{\gamma}}$ , where  $\bar{\gamma} \geq \gamma$  and  $\kappa_\ell$  is a slowly varying factor (cf. §4). The insight into the behavior of  $\eta_\ell$  should lead to expressions for the error constant in the form  $\kappa_\ell = \|k_\ell\|_{L_p}$  with  $k_\ell = \chi(x_{\ell-1}, x_\ell, t_{\ell-1}, t_\ell)$ .

Given that, for an interesting range of stepsizes  $h_\ell$ , the local error term  $\eta_\ell$  is dominated by

$$\eta_\ell \leq \kappa_\ell \cdot h_\ell^{\bar{\gamma}}, \quad \kappa_\ell = \|k_\ell\|_{L_p} = \|\chi(x_{\ell-1}, x_\ell, t_{\ell-1}, t_\ell)\|_{L_p}, \quad (3.2)$$

the stepsizes should be chosen according to the following conceptual algorithm:

ALGORITHM 3.1. Given the initial value  $x_0 \in L_p(\Omega)$  at  $t_0$ , an initial stepsize  $h_1$  and a tolerance  $\text{tol}$ ; set  $\ell := 1$ .

1) Solve the system

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

for  $x_\ell \in L_p(\Omega)$ .

2) Set

$$\bar{\eta}_\ell := h_\ell^{\bar{\gamma}} \|\chi(x_{\ell-1}, x_\ell, t_{\ell-1}, t_\ell)\|_{L_p}. \quad (3.3)$$

3) Apply a control strategy that aims at matching the tolerance multiplied by a safety factor, say 0.8, with the proposed new stepsize. For example, the elementary control leads to

$$\frac{h_{new}}{h_\ell} := \left(\frac{0.8 \cdot \text{tol}}{\bar{\eta}_\ell}\right)^{1/\bar{\gamma}}, \quad (3.4)$$

for  $\ell \geq 2$ , the proportional integral control PI34 (cf. [10, 25]) leads to

$$\frac{h_{new}}{h_\ell} := \left(\frac{0.8 \cdot \text{tol}}{\bar{\eta}_\ell}\right)^{0.3/\bar{\gamma}} \left(\frac{\bar{\eta}_{\ell-1}}{\bar{\eta}_\ell}\right)^{0.4/\bar{\gamma}}. \quad (3.5)$$

4) If  $\bar{\eta}_\ell \leq \text{tol}$ , accept the step.

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

If  $\bar{\eta}_\ell > \text{tol}$ , reject the step and repeat it with the smaller stepsize  $h_\ell := h_{new}$  that results from 3.4.

We emphasize that a stepsize sequence generated by Algorithm 3.1 is deterministic, since  $\|\chi(x_{\ell-1}, x_\ell, t_{\ell-1}, t_\ell)\|_{L_p}$  is deterministic, though  $x_{\ell-1}, x_\ell$  are random variables. Hence, the Theorems 2.1 and 2.2 apply. If Algorithm 3.1 is realized and the computed quantities  $\bar{\eta}_\ell$  are really dominating the local error term  $\eta_\ell$ , say, e.g.,  $\eta_\ell \leq 1.2 \bar{\eta}_\ell$ , then we have

$$\max_{\ell=1, \dots, N} \|x(t_\ell) - x_\ell\|_{L_p} \leq 2.4 S \cdot \text{tol}$$

by the  $p$ -th mean stability inequality (2.7) from Theorem 2.1. Even if the actual errors  $\eta_\ell$  show a smaller order than the used quantities  $\bar{\eta}_\ell$ , and only a relation like  $\eta_\ell \leq \text{const} \cdot \bar{\eta}_\ell^\beta$ ,  $\beta \in (0, 1)$ , is true, we still have

$$\max_{\ell=1, \dots, N} \|x(t_\ell) - x_\ell\|_{L_p} \leq 2\text{const} \cdot S \cdot \text{tol}^\beta.$$

Intending to make use of this theoretical result for practical implementations, one faces several questions.

First, how large is the stability constant  $S$ ? Of course, it is problem-dependent. A more detailed look into the proof of Theorem 2.1 together with Proposition 2.3 shows that  $S$  behaves essentially as  $e^{L(T-t_0)}$ , where  $L$  is a Lipschitz constant for the drift and diffusion coefficients  $f$  and  $G$  (and of  $(g_j)'_x \cdot G$  for the Milstein schemes) of the problem. By analogy with deterministic ODEs one may consider the problem as nonstiff and numerically well-posed, as long as  $e^{L(T-t_0)}$  is moderate. Modifications for stiff problems are desirable, but go beyond the scope of this paper.

Second, how can the observed order  $\bar{\gamma}$  and the error function  $\chi$  be determined and what conditions guarantee an estimate like  $\eta_\ell \leq 1.2 \bar{\eta}_\ell$ ? This question is discussed for the family of Euler-Maruyama schemes in the next section.

Finally, we discuss implementation issues of the conceptual Algorithm 3.1. Of course, any implementation requires a finite number of realizations of the random variables. Here, we consider an ensemble of  $M$  paths starting from  $M$  samples  $x_0^1, \dots, x_0^M$  of the initial value. Starting from  $\ell = 1$  the next elements  $x_\ell^1, \dots, x_\ell^M$  of the  $M$  paths are computed by solving the nonlinear equations

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

where  $I_{t_{\ell-1}, h_\ell}^i$  are samples of the corresponding stochastic integrals.

Next, the  $L_p$ -norm  $\|\chi(x_{\ell-1}, t_{\ell-1}, x_\ell, t_\ell)\|_{L_p}$  in step 2) of Algorithm 3.1 is estimated by using the  $M$  samples of  $x_{\ell-1}$  and  $x_\ell$ , namely, by

$$\bar{\kappa}_\ell := \left( \frac{1}{M} \sum_{i=1}^M |\chi(x_{\ell-1}^i, x_\ell^i; t_{\ell-1}, t_\ell)|^p \right)^{\frac{1}{p}}. \quad (3.6)$$

Set

$$\hat{\eta}_\ell := h_\ell^{\bar{\gamma}} \bar{\kappa}_\ell \quad (3.7)$$

and perform the next steps 3) and 4) as described in Algorithm 3.1 with  $\bar{\eta}_\ell$  replaced by  $\hat{\eta}_\ell$ . In case of step rejections, the information computed so far is stored and used to compute intermediate values according to the strategy in [18, 19]. For a scheme (2.1), which uses only the Wiener increments  $\Delta w_\ell = w(t_\ell) - w(t_{\ell-1})$ , the computation of the intermediate values of the Wiener process is done as follows: Given  $\Delta w_h := w(t+h) - w(t)$  for some  $t \in \mathbb{R}_+$  and  $h > 0$ , and  $h = h_1 + h_2$ ,  $h_1 > 0$ ,  $h_2 > 0$ , 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

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

In this way, the estimate (3.7) of (3.3) leads to estimates of the next stepsize and, hence, of the whole grid  $t_0, t_1, \dots, t_N$ . Though all computed paths  $(x_0^i, x_1^i, \dots, x_N^i)$ ,  $i = 1, \dots, M$ , are determined by using the same stepsizes  $h_1, \dots, h_N$ , the step-size sequence is no longer deterministic and due to possible step rejections the computed grid-points do not need to be stopping times. The estimate of the grid depends on the  $M$  paths and its quality clearly depends on the sample size  $M$  as well as on the smallness of the noise. The resulting implementable algorithm represents a theory-based heuristic stepsize control.

The numerical results in Section 5 show that a relatively small number  $M = 100$  already provides good results for SDEs with small noise.

**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 [22], 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 [22] we let the SDE with small noise be of the form

$$x(t) = x_0 + \int_{t_0}^t f(x(s), s)ds + \int_{t_0}^t \epsilon \tilde{G}(x(s), s)dw(s), \quad t \in \mathcal{J}, \quad (4.1)$$

where  $f : \mathbb{R}^n \times \mathcal{J} \rightarrow \mathbb{R}^n$ ,  $\tilde{G} : \mathbb{R}^n \times \mathcal{J} \rightarrow \mathbb{R}^{n \times m}$  are functions satisfying the assumptions introduced in Section 1 for  $f$  and  $G$ , and  $\epsilon$  is a small parameter.

The family of drift-implicit Euler schemes with parameter  $\theta \in [0, 1]$  for solving (4.1) on the deterministic grid  $t_0 < t_1 < \dots < t_N = T$  with stepsizes  $h_\ell := t_\ell - t_{\ell-1}$ ,  $\ell = 1, \dots, N$ , has the form

$$x_\ell = x_{\ell-1} + h_\ell(\theta f(x_\ell, t_\ell) + (1-\theta)f(x_{\ell-1}, t_{\ell-1})) + \epsilon \tilde{G}(x_{\ell-1}, t_{\ell-1})\Delta w_\ell, \quad \ell = 1, \dots, N, \quad (4.2)$$

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 discretization error  $l_\ell$  of (4.2), we first establish, similarly as in [22], a representation of  $l_\ell$  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  $k, q \in \mathbb{N} \cup \{0\}$ . 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 (4.1) and  $y$  be a function in  $C^{2,1}$ . Then Itô's



formula provides the expansion

$$\begin{aligned} y(x(t), t) - y(x_0, t_0) &= \int_{t_0}^t (y_t + y_x f + \epsilon^2 \frac{1}{2} \sum_{r=1}^m \sum_{i,j=1}^n y_{x_i x_j} \tilde{g}_{ri} \tilde{g}_{rj})(x(s), s) ds \\ &\quad + \epsilon \sum_{r=1}^m \int_{t_0}^t (y_x \tilde{g}_r)(x(s), s) dw_r(s), \quad t \in \mathcal{J}. \end{aligned} \quad (4.3)$$

Following [22] we introduce  $m + 1$  operators  $\Lambda_0$  and  $\Lambda_r$ ,  $r = 1, \dots, m$ , defined on  $C^{2,1}$  and  $C^{1,0}$ , respectively, by

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

Then the Itô formula (4.3) reads

$$y(x(t), t) - y(x_0, t_0) = \int_{t_0}^t \Lambda_0 y(x(s), s) ds + \epsilon \sum_{r=1}^m \int_{t_0}^t \Lambda_r y(x(s), s) dw_r(s), \quad t \in \mathcal{J}. \quad (4.4)$$

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

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

where  $i_1, \dots, i_j$  take values in  $\{0, \dots, m\}$ , 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  $i_j \in \{1, \dots, m\}$ ,  $j = 1, \dots, k$ , we have for any function  $y \in C_L$  that*

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

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

$$\begin{aligned} \|I_{0, i_2, \dots, i_k; t, h}(y)\|_{L_p}^p &= \mathbb{E}(|I_{0, i_2, \dots, i_k; 0; t, h}(y)|^p) \leq \left( \mathbb{E} \int_t^{t+h} |I_{i_2, \dots, i_k; t, s_1-t}(y)|^p ds_1 \right)^p \\ &\leq h^{\frac{p}{q}} \int_t^{t+h} \mathbb{E}(|I_{i_2, \dots, i_k; t, s_1-t}(y)|^p) ds_1, \end{aligned}$$

where  $\frac{1}{q} + \frac{1}{p} = 1$ . Hence, for  $\|I_{0, i_2, \dots, i_k; t, h}(y)\|_{L_p}^p$  we obtain the order  $\mathcal{O}(h^{\frac{p}{q}+1}) = \mathcal{O}(h^p)$ . For  $i_1 \neq 0$  we make use of estimates for the  $p$ -th mean of stochastic integrals (see [8, §1.4, Theorem 6], [17, §1.7, Theorem 7.1]) and have

$$\|I_{i_1, \dots, i_k; t, h}(y)\|_{L_p}^p \leq \left(\frac{1}{2} p(p-1)\right)^{\frac{p}{2}} h^{\frac{p-2}{2}} \int_t^{t+h} \mathbb{E}(|I_{i_2, \dots, i_k; t, t-s_1}(y)|^p) ds_1.$$

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.  $\square$

PROPOSITION 4.2. *Assume that  $f \in C^{4,3}$  and  $\tilde{g}_r \in C^{2,1}$ ,  $r = 1, \dots, m$  and that the functions  $\Lambda_0 \Lambda_0 f$ ,  $\Lambda_0 \tilde{g}_r$ ,  $\Lambda_r \Lambda_0 f$ ,  $\Lambda_r f$  and  $\Lambda_k \tilde{g}_r$ ,  $k, r = 1, \dots, m$ , belong to  $C_L$ . Then the local discretization error  $l_\ell$  (see (2.8)) of the family of drift-implicit Euler schemes (4.2) at step  $\ell$  allows a decomposition*

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

and

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

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

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

By reformulating the local error and by expanding all of its components at the pair  $(x(t_{\ell-1}), t_{\ell-1})$  using (4.4) and the smoothness properties  $f, \tilde{g}_r \in C^{2,1}$ ,  $r = 1, \dots, m$ , we obtain

$$\begin{aligned} l_\ell &= x(t_\ell) - x(t_{\ell-1}) - h_\ell(\theta f(x(t_\ell), t_\ell) + (1-\theta)f(x(t_{\ell-1}), t_{\ell-1})) - \epsilon \tilde{G}(x(t_{\ell-1}), t_{\ell-1}) \Delta w_\ell \\ &= \int_{t_{\ell-1}}^{t_\ell} f_s ds - h_\ell(\theta f_{t_\ell} + (1-\theta)f_{t_{\ell-1}}) + \int_{t_{\ell-1}}^{t_\ell} \epsilon \tilde{G}_s dw(s) - \epsilon \tilde{G}_{t_{\ell-1}} \Delta w_\ell \\ &= \int_{t_{\ell-1}}^{t_\ell} \left\{ f_{t_{\ell-1}} + \int_{t_{\ell-1}}^s (\Lambda_0 f)_\tau d\tau + \epsilon \sum_{r=1}^m \int_{t_{\ell-1}}^s (\Lambda_r f)_\tau dw_r(\tau) \right\} ds \\ &\quad - \theta h_\ell \left\{ f_{t_{\ell-1}} + \int_{t_{\ell-1}}^{t_\ell} (\Lambda_0 f)_\tau d\tau + \epsilon \sum_{r=1}^m \int_{t_{\ell-1}}^{t_\ell} (\Lambda_r f)_\tau dw_r(\tau) \right\} - (1-\theta) h_\ell f_{t_{\ell-1}} \\ &\quad + \epsilon \sum_{r=1}^m \int_{t_{\ell-1}}^{t_\ell} \left\{ \int_{t_{\ell-1}}^s (\Lambda_0 \tilde{g}_r)_\tau d\tau + \epsilon \sum_{k=1}^m \int_{t_{\ell-1}}^s (\Lambda_k \tilde{g}_r)_\tau dw_k(\tau) \right\} dw_r(s) \\ &= I_{00}(\Lambda_0 f) - \theta h_\ell I_0(\Lambda_0 f) + \epsilon \sum_{r=1}^m (I_{0r}(\Lambda_r f) - \theta h_\ell I_r(\Lambda_r f) + I_{r0}(\Lambda_0 \tilde{g}_r)) + \epsilon^2 \sum_{r,k=1}^m I_{rk}(\Lambda_k \tilde{g}_r), \end{aligned}$$

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 (4.4) again and obtain

$$\Lambda_0 f_\tau = \Lambda_0 f_{t_{\ell-1}} + \int_{t_{\ell-1}}^\tau (\Lambda_0 \Lambda_0 f)_s ds + \epsilon \sum_{r=1}^m \int_{t_{\ell-1}}^\tau (\Lambda_r \Lambda_0 f)_s dw_r(s), \quad \tau \in [t_{\ell-1}, t_\ell]. \quad (4.5)$$

The latter equation (4.5) is taken to compute the desired (multiple) stochastic inte-

grals and the whole leading term, respectively, i.e.,

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

Now, we split  $l_\ell = s_\ell + r_\ell$ , where  $s_\ell$  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

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

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

**4.2. Suggestions for local error estimates.** The previous result enables us to study the local error term  $\eta_\ell$  for special relations between  $\epsilon$  and the stepsizes  $h_\ell$ . Unless  $\theta = 1/2$  for the trapezoidal rule, the dominating term of  $\|r_\ell\|_{L_p}/h_\ell$  is

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

hence

$$\eta_\ell = \bar{\eta}_\ell^* + \mathcal{O}(\epsilon^2 h_\ell^{1/2} + \epsilon h_\ell + h_\ell^2), \quad \bar{\eta}_\ell^* := h_\ell |\theta - \frac{1}{2}| \|(f_t + f_x f)(x(t_{\ell-1}), t_{\ell-1})\|_{L_p}. \quad (4.6)$$

Substituting the exact solution value  $x(t_{\ell-1})$  by its numerical approximation  $x_{\ell-1}$  does not lead to perturbations that are larger than  $\mathcal{O}(\epsilon^2 h^{1/2} + \epsilon h + h^2)$ . This motivates the choice of  $\bar{\gamma} = 1$  and  $\chi(z, x, s, t) := |\theta - \frac{1}{2}| (f_t + f_x f)(z, s)$  for the definition of  $\bar{\eta}_\ell$  in (3.3). We assume now that  $\theta \neq \frac{1}{2}$  and that the relation of the small parameter  $\epsilon$  and the applied stepsize  $h_\ell$  is such that the global error term of order  $\mathcal{O}(h_\ell)$  dominates the error term of order  $\mathcal{O}(\epsilon^2 h_\ell^{1/2})$ . This is realized if  $h_\ell$  is much larger than  $\epsilon^2 h_\ell^{1/2}$ , which we express by

$$\epsilon^2 h_\ell^{1/2} \ll h_\ell, \quad \text{i.e.,} \quad \epsilon^4 \ll h_\ell. \quad (4.7)$$

More precisely, we need that

$$\epsilon^2 \left\| \sum_{r,k=1}^m I_{rk}(\Lambda_k \tilde{g}_r) \right\|_{L_p} / h_\ell^{1/2} < h_\ell |\theta - \frac{1}{2}| \|(f_t + f_x f)_{t_{\ell-1}}\|_{L_p}, \quad (4.8)$$

which follows from (4.7) provided that the values of the functions  $\tilde{g}_{rx}\tilde{g}_k$ ,  $r, k = 1, \dots, m$ , and  $f_t + f_x f$  are moderate. Then, the local error term  $\eta_\ell$  is indeed dominated by  $\bar{\eta}_\ell^*$  from (4.6). Note that the corresponding choice of  $\bar{\eta}_\ell$  leads to an a-priori known estimate of the local error, since  $\bar{\eta}_\ell$  involves only the knowledge of  $x_{\ell-1}$ , not that of  $x_\ell$ . That way step rejections do not occur.

One drawback of the choice (4.6) is its explicit use of the derivatives of the drift function  $f$ , which may not be available in practical problems. Therefore, we look for a derivative-free estimate of  $\eta_\ell$ . We use

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

which follows from the expansion

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

that is valid under the assumptions of Proposition 4.2 and is obtained by inserting (4.5) into Itô's formula (4.4) for  $y = f$ . Hence,

$$\eta_\ell = \bar{\eta}_\ell^* + \mathcal{O}(h_\ell^2 + \epsilon h_\ell + \epsilon h_\ell^{1/2}), \quad \bar{\eta}_\ell^* := |\theta - \frac{1}{2}| \|f(x(t_\ell), t_\ell) - f(x(t_{\ell-1}), t_{\ell-1})\|_{L_p}. \quad (4.9)$$

The estimate  $\bar{\eta}_\ell^*$  is dominating the local error term  $\eta_\ell$  under the more restrictive assumption

$$\epsilon h_\ell^{1/2} \ll h_\ell, \quad \text{i.e.,} \quad \epsilon^2 \ll h_\ell, \quad (4.10)$$

or, more precisely,

$$\epsilon \left\| \sum_{r=1}^m I_r(\Lambda_r f) \right\|_{L_p} < h_\ell \|(\Lambda_0 f)_{t_{\ell-1}}\|_{L_p}. \quad (4.11)$$

Then, the term  $\bar{\eta}_\ell^*$  behaves like an order 1 term. Hence, we represent this choice by  $\bar{\gamma} = 1$  and  $\chi(z, x, s, t) := |\theta - \frac{1}{2}| (f(x, t) - f(z, s)) / (t - s)$ . In Algorithm 3.1 this is realized by the choice

$$\bar{\gamma} := 1, \quad \bar{\eta}_\ell := |\theta - \frac{1}{2}| \|f(x_\ell, t_\ell) - f(x_{\ell-1}, t_{\ell-1})\|_{L_p}. \quad (4.12)$$

**REMARK 4.3.** Note that  $\bar{\eta}_\ell^*$  defined by (4.6) as well as  $\bar{\eta}_\ell^*$  defined by (4.9) vanishes for the trapezoidal rule. Here a more detailed discussion of the remaining terms is necessary. Then the part  $\|r_\ell\|_{L_p}/h_\ell$  of the local error is of order 2 and is dominated by  $\|(\Lambda_0 \Lambda_0 f)_{t_{\ell-1}}\|_{L_p} \cdot h_\ell^2/12$ . This term dominates the local error term  $\eta_\ell$  as long as  $\epsilon \ll h_\ell$ , which is more restrictive than (4.10). Details will be given in a separate paper.

**REMARK 4.4.** If the length  $T - t_0$  of the considered time interval differs considerably from 1, the more detailed stability estimate

$$\max_{\ell=1, \dots, N} \|x_\ell^* - \tilde{x}_\ell\|_{L_p} \leq \tilde{S} \left( \|x_0^* - \tilde{x}_0\|_{L_p} + \max_{\ell=1, \dots, N} \|s_\ell\|_{L_p} \left(\frac{T-t_0}{h}\right)^{\frac{1}{2}} + \max_{\ell=1, \dots, N} \|r_\ell\|_{L_p} \frac{T-t_0}{h} \right) \quad (4.13)$$

(cf. [24]) should be used as a starting point. Since the length of the interval affects the local error terms differently, condition (4.8) modifies to

$$\epsilon^2 \left\| \sum_{r,k=1}^m I_{rk}(\Lambda_k \tilde{g}_r) \right\|_{L_p} \left( \frac{T-t_0}{h_\ell} \right)^{\frac{1}{2}} < (T-t_0)h_\ell \left| \theta - \frac{1}{2} \right| \| (f_t + f_x f)_{t_{\ell-1}} \|_{L_p}, \quad (4.14)$$

which is true for

$$\epsilon^2 ((T-t_0)h_\ell)^{\frac{1}{2}} \ll (T-t_0)h_\ell, \quad \text{i.e.,} \quad \epsilon^4 \ll (T-t_0)h_\ell, \quad (4.15)$$

and moderate coefficients  $\tilde{g}_{r,x} \tilde{g}_k$ ,  $r, k = 1, \dots, m$ , and  $(f_t + f_x f)$ .

REMARK 4.5. The conditions (4.14), (4.15), (4.11), (4.10) are independent of the used time-scale.

Proof: A transformation of the time scale  $t \in [t_0, T]$  to  $\tau \in [0, 1]$  via

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

leads to the transformed SDE

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

The conditions (4.14), (4.15), (4.11), (4.10) 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.  $\square$

REMARK 4.6. The simple conditions (4.7), (4.15), and (4.10) 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 (4.8), (4.14) and (4.11) we then obtain the conditions

$$h_\ell > \frac{\| (g_x g)_{t_{\ell-1}} \|_{L_2}^2}{2 \left| \theta - \frac{1}{2} \right| \| (f_t + f_x f)_{t_{\ell-1}} \|_{L_2}^2}, \quad (4.16)$$

$$(T-t_0)h_\ell > \frac{\| (g_x g)_{t_{\ell-1}} \|_{L_2}^2}{2 \left| \theta - \frac{1}{2} \right| \| (f_t + f_x f)_{t_{\ell-1}} \|_{L_2}^2}, \quad (4.17)$$

and

$$h_\ell > \frac{\| (f_x g)_{t_{\ell-1}} \|_{L_2}^2}{\| (f_t + f_x f + \frac{1}{2} f_{xx} g g)_{t_{\ell-1}} \|_{L_2}^2}, \quad (4.18)$$

in terms of the ratio of the coefficients.

Proof: For  $m = 1$  the condition (4.8) simplifies to

$$h_\ell^{-\frac{1}{2}} \| I_{11}(g_x g) \|_{L_p} < h_\ell \left| \theta - \frac{1}{2} \right| \| (f_t + f_x f)_{t_{\ell-1}} \|_{L_p}.$$

Choosing  $p = 2$  and neglecting higher order terms one obtains

$$h_\ell^{-\frac{1}{2}} h_\ell 2^{-\frac{1}{2}} \| (g_x g)_{t_{\ell-1}} \|_{L_2} < h_\ell \left| \theta - \frac{1}{2} \right| \| (f_t + f_x f)_{t_{\ell-1}} \|_{L_2},$$

i.e., (4.16). Analogous arguments apply to (4.14) and (4.11).  $\square$

**5. Test results.** We implemented the stepsize strategy proposed in Section 3 for the drift-implicit Euler scheme (i.e. (4.2) with  $\theta = 1$ ). The error estimate (3.3) has been realized by the derivative-free choice (4.12), the control has been realized by (3.4) for  $\ell = 1$ , and (3.5) for  $\ell \geq 2$ . We tested the resulting algorithm extensively for  $p = 2$  on a set of SDEs with small noise. First, we report results for two scalar SDEs with known analytic solution where we can access the actual errors. The accuracy is measured by the empirical error quantity

$$\max_{\ell=1,\dots,N} \left( \frac{1}{M} \sum_{j=1}^M |x^j(t_\ell) - x_\ell^j|^2 \right)^{1/2}, \quad (5.1)$$

that is considered as an estimate of the maximum  $L_2$ -norm of the global errors in the considered time-interval. Here  $N$  denotes the number of steps and  $M$  the number of computed paths,  $x_\ell^j$  denotes the computed approximation of the  $j$ -th path at time  $t_\ell$ , while  $x^j(t_\ell)$  denotes the corresponding exact solution value.

Finally, we present results for a low-dimensional electronic 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

$$x(t) = x_0 + \int_0^t \alpha x(s) ds + \int_0^t i\beta x(s) dw(s), \quad t \in [0, 1], \quad (5.2)$$

with coefficients  $f(x, t) := \alpha x$ ,  $G(x, t) = (g(x, t)) = (i\beta x)$ , initial value  $x_0$ , parameters  $\alpha, \beta \in \mathbb{R}$  and a scalar Wiener process  $w$ . It was implemented as a two-dimensional system in real arithmetic. Its solution is given by the geometric Brownian motion  $x(t) = x_0 \exp\left(\left(\alpha + \frac{1}{2}\beta^2\right)t + i\beta w(t)\right)$ . The conditions (4.16) and (4.18) are equivalent to

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

respectively. Here, the ratio  $|\beta/\alpha|$  plays the role of the small parameter  $\epsilon$ . As long as stepsizes with  $\beta^4/(2\alpha^4) \ll h_\ell$  are used, the Euler scheme shows order 1 of convergence. As long as even  $\beta^2/\alpha^2 \ll h_\ell$  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

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

instead of

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

The proposed control leads to stepsizes that match

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

In the following we present results for different values of the parameters  $\alpha, \beta$ . The initial value was chosen to be  $x_0 = 1$ . We start with results for the parameters

$\alpha = -10, \beta = 10^{-2}$  with ratio  $|\beta/\alpha| = 10^{-3}$ . The solution shows an exponential decrease with the steepest gradients at the beginning of the integration interval. In Figure 5.1 we give the real part of the solution (+) together with the adaptively chosen stepsizes (left picture,  $\times$ ) and the observed global errors (right picture,  $\times$ ) for the relative tolerance  $\text{tol}=0.125$  in Algorithm 3.1. The total number of steps was 129 corresponding to an average stepsize of  $7.75 \cdot 10^{-3}$ , whereas the minimal stepsize was  $2.33 \cdot 10^{-3}$ .

One hundred paths are computed simultaneously. Figure 5.2 gives the tolerance and

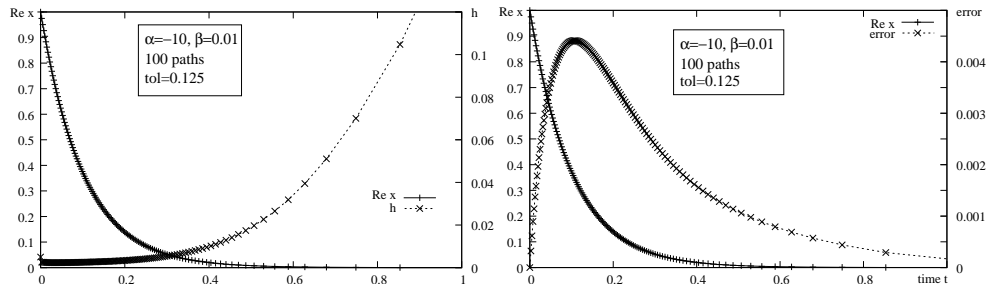


FIG. 5.1. A computed solution path and stepsizes (left) or global errors (right) for the SDE (5.2)

accuracy (5.1) versus the number of steps. We plot the relative tolerance ( $\Delta$ ) and the accuracy with adaptively chosen stepsizes (+) and with constant stepsizes ( $\times$ ) versus the number of steps in logarithmic scale with base 10. Lines with slopes  $-1$  and  $-0.5$  are provided to enable comparisons with convergence of order 1 or  $1/2$ .

To see how much the results are influenced by the statistical error due to only 100

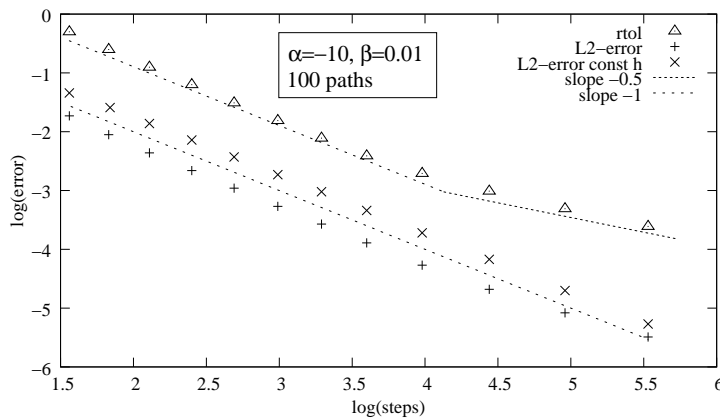


FIG. 5.2. Tolerance and accuracy versus steps for the SDE (5.2) for  $\alpha = -10$  and  $\beta = 10^{-2}$

randomly chosen paths we repeated the computations 100 times for selected values of the relative tolerance. In Table 5.1 the mean value and standard deviation (in percent) of the observed grid-points, of the rejected steps and of the accuracy (5.1) are reported.

Let us add some observations. The global errors do not increase over the whole time interval, but show an exponential decrease after a shorter phase of increase at the beginning. Therefore, the maximum of the global errors over the integration steps is more meaningful than just the global error at the end of the interval. Furthermore,

tolerance	successful steps		rejected steps		accuracy	
	mean	deviation	mean	deviation	mean	deviation
$2^{-5}$	494	0%	0	0%	$1.08 \cdot 10^{-3}$	0.073%
$2^{-10}$	27426.25	0.11%	17.96	21.43%	$2.11 \cdot 10^{-5}$	0.077%

TABLE 5.1

Variation of 100 repetitions of the simulation results for the SDE (5.2) with parameters  $\alpha = -10$ ,  $\beta = 0.01$  and 100 simultaneously computed paths

the error estimates are too pessimistic. Both effects are due to the damping of the solution by  $\alpha = -10$ , since the errors are damped as well. Further, in case of the small noise parameter  $|\beta/\alpha| = 10^{-3}$ , the stepsize control works well up to average stepsizes of  $10^{-4}$ , and provides more accurate results than solving the SDE with the same number of constant steps. For higher numbers of steps the smaller stepsizes at the beginning of the integration interval already lie below the threshold  $\beta^2/\alpha^2 \ll h_\ell$ . The results in Table 5.1 show that the proposed stepsize control works very well for stepsizes above this threshold and still quite reasonable for stepsizes slightly below the threshold. The achieved accuracy differs only by less than 0.1% from one repetition of the simulation to another.

Next, we present results for the parameters  $\alpha = -0.5$  and  $\beta = 0.01$ , i.e.,  $|\beta/\alpha| = 0.02$ . Here, the decrease of the solution is much slower and constant stepsizes are nearly optimal. Compared to the problem with  $\alpha = -10$  the noise intensity is larger. In Figures 5.3, 5.4 and Table 5.2 we present results of experiments that correspond to those described for  $\alpha = -10$ .

Due to the less small noise the range of accuracy, where the stepsize selection

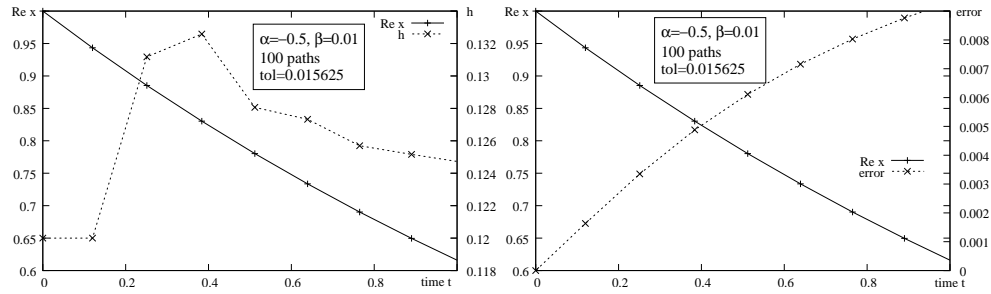


FIG. 5.3. A computed solution path and stepsizes (left) or global errors (right) for the SDE (5.2)

tolerance	successful steps		rejected steps		accuracy	
	mean	deviation	mean	deviation	mean	deviation
$2^{-6}$	8	0%	0	0%	$9.37 \cdot 10^{-3}$	0.018%
$2^{-11}$	2537.64	0.57%	9.64	34.53%	$3.26 \cdot 10^{-5}$	0.62%

TABLE 5.2

Variation of 100 repetitions of the simulation results for the SDE (5.2) with parameters  $\alpha = -0.5$ ,  $\beta = 0.01$  and 100 simultaneously computed paths

strategy works as intended, decreases to average stepsizes below  $10^{-2}$ . In this range



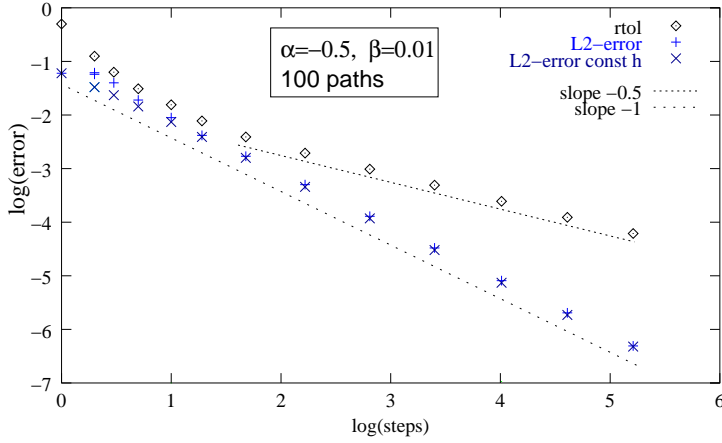


FIG. 5.4. Tolerance and accuracy versus steps for the SDE (5.2) for  $\alpha = -0.5$  and  $\beta = 10^{-2}$

one can see a good correspondence of the tolerance to the accuracy. For very low accuracy (see the very left side of Figure 5.4), constant stepsizes seem to perform even better. Here, one has to see that the adaptive stepsize control first had to find the suitable stepsizes that correspond to the tolerance. Further, Table 5.2 shows that the simulations for  $|\beta/\alpha| = 0.02$  and average stepsizes below  $10^{-2}$  are slightly more affected by the randomly chosen 100 paths than the simulations for  $|\beta/\alpha| = 0.001$  above. However the differences from one repetition of the experiment to another still lie below 1%.

Finally, we report results for  $\alpha = -0.5$ ,  $\beta = 0.1$ , i.e.,  $|\beta/\alpha| = 0.2$ , where one can hardly speak of small noise. To obtain reasonable approximations of the  $\|\cdot\|_{L_2}$ -norm of the local error estimates, we increased the number of simultaneously computed paths to  $M=1000$ . Nevertheless, we are in a situation, for which the stepsize control proposed in this article is not tailored. Again, we present corresponding results in Figures 5.5 and 5.6 and in Table 5.3.

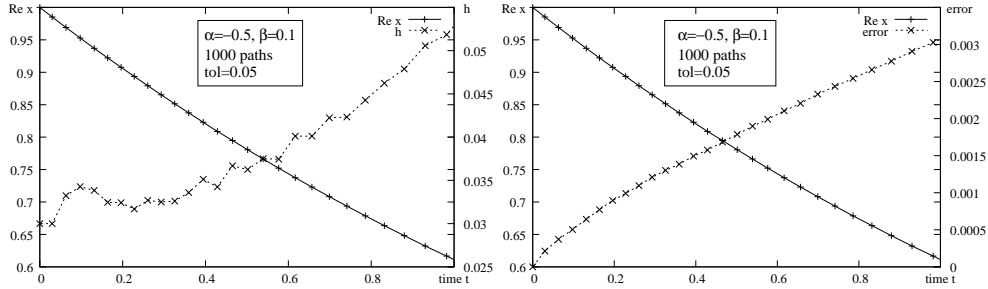


FIG. 5.5. A computed solution path and stepsizes (left) or global errors (right) for the SDE (5.2)

Here one observes the asymptotic order  $1/2$  of convergence already for medium accuracy. In this region the local error term  $\eta_\ell$  is dominated by  $\|I_{11}(\Lambda_1 g_1)_{\ell-1}\|_{L_2} = \beta^2 h^{1/2} \|x_{\ell-1}\|_{L_2}$ , and the stepsize control leads to stepsizes that match

$$\beta^2 h^{1/2} \|x_{\ell-1}\|_{L_2} \approx \text{tol} |\beta/\alpha|.$$

Though we have increased the number of simultaneously computed paths to 1000, the

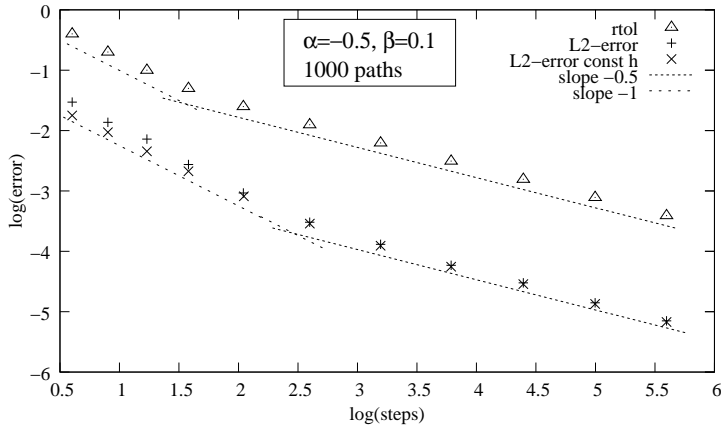


FIG. 5.6. Tolerance and accuracy versus steps for the SDE (5.2) for  $\alpha = -0.5$  and  $\beta = 0.1$

tolerance	successful steps		rejected steps		accuracy	
	mean	deviation	mean	deviation	mean	deviation
$0.1 \cdot 2^{-2}$	102.05	1.17%	0	0%	$9.21 \cdot 10^{-4}$	1.70%
$0.1 \cdot 2^{-6}$	24667.24	1.27%	8.46	60.94%	$2.86 \cdot 10^{-5}$	2.17%

TABLE 5.3

Variation of 100 repetitions of the simulation results for the SDE (5.2) with parameters  $\alpha = -0.5$ ,  $\beta = 0.1$  and 1000 simultaneously computed paths

results are more influenced by the randomness than in the previous simulations for smaller noise. We emphasize that the stepsize-selection algorithm relies on a sufficient approximation of the norm of the local error terms by the simultaneously computed paths.

**Example 5.2.** (SDE with polynomial drift and diffusion)

Here we consider a nonlinear scalar SDE with known solution and drift- and diffusion coefficients  $f(x, t) := -(\alpha + \beta^2 x)(1 - x^2)$ ,  $G(x, t) = (g(x, t)) = (\beta(1 - x^2))$  that are tunable by real parameters  $\alpha, \beta$ :

$$x(t) = \int_0^t -(\alpha + \beta^2 x(s))(1 - x(s)^2) ds + \int_0^t \beta(1 - x(s)^2) dw(s), \quad t \in [0, T], \quad (5.3)$$

where  $w$  denotes a scalar Wiener process. The solution is given by (cf. [14, (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 (4.16), (4.17), (4.18) 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. As in the deterministic setting, failures of Newton's method may also cause step rejections, especially for larger stepsizes, where the quality of the starting point for Newton's method may be worse. In such a case we halved the unsuccessful stepsize and forbade stepsize enlargements for the next five steps.

Though the problem is nonlinear, the principal observations we made, were the same as for the linear Example 5.1. Here, we restrict our presentation to simulation results for one set of parameters,  $\alpha = -10$  and  $\beta = 0.1$ , where 100 paths were computed simultaneously. In Figures 5.7, 5.8 and Table 5.4 we present results of experiments that correspond to those described for Example 5.1. The solution shows a transient

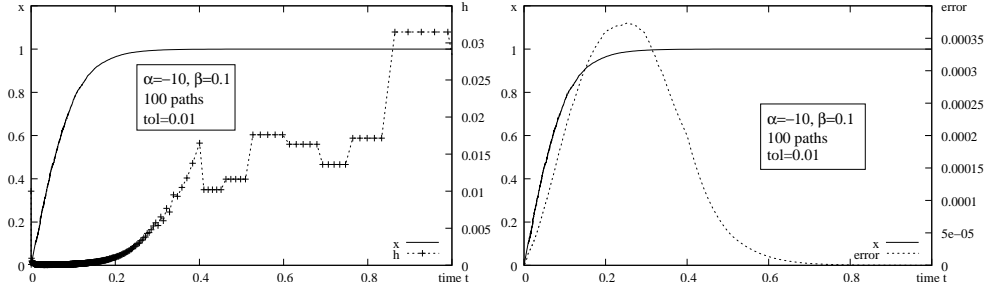


FIG. 5.7. A computed solution path and stepsizes (left) or global errors (right) for the SDE (5.3)

behavior at the beginning of the time-interval. At the end of the time-interval the stepsizes are bounded by the convergence behavior of Newton’s method. Failures of Newton’s methods also prevented simulation results for larger stepsizes in case of larger noise. In Table 5.4 additional columns report the number of step rejections due to failures of Newton’s method. In Figure 5.8 we observe order 1 behavior up to accuracies of  $10^{-3}$ . The parameter  $\alpha = -10$  causes a damping in the solution and a prediction of the global error that is much too pessimistic. The use of adaptive stepsizes provides considerably more accurate results than the computation with the same number of constant steps.

toler.	successful steps		rejected steps		Newton failures		accuracy	
	mean	deviat.	mean	deviat.	mean	deviat.	mean	deviat.
0.01	1316.0	0.20%	5.01	12.46%	17.36	12.94%	$3.75 \cdot 10^{-4}$	0.60 %
0.005	3399.4	0.22%	8.45	23.56%	15.63	12.78%	$1.88 \cdot 10^{-4}$	0.71%

TABLE 5.4

Variation of 100 repetitions of the simulation results for the SDE (5.3) with parameters  $\alpha = -10$ ,  $\beta = 0.1$  and 100 simultaneously computed paths

**Example 5.3.** (electronic circuit)

We consider a model of an inverter circuit with a MOSFET-transistor under the influence of thermal noise. The equivalent circuit diagram is given in Figure 4.

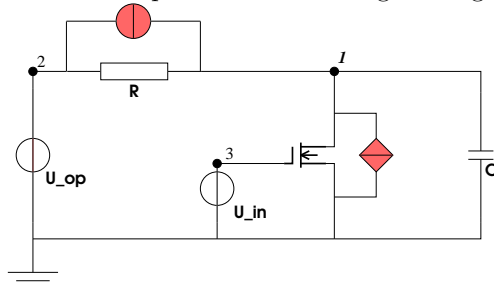


Figure 4: Thermal noise sources in a MOSFET inverter circuit

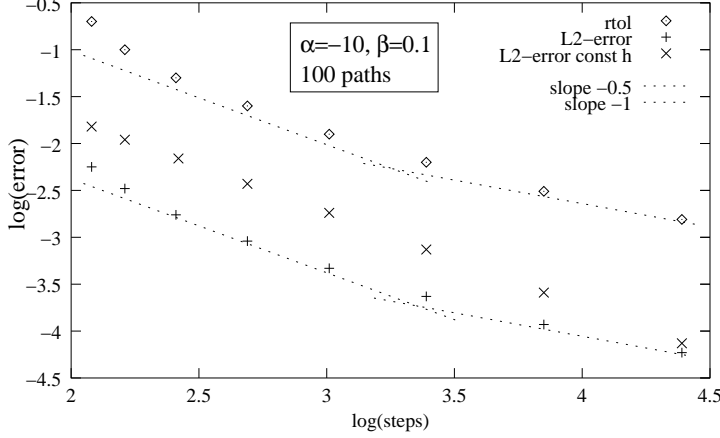


FIG. 5.8. Tolerance and accuracy versus steps for the SDE (5.3) for  $\alpha = -10$ ,  $\beta = 0.1$

The MOSFET is modelled as a current source from source to drain that is controlled by the nodal potentials at gate, source and drain:  $j_D = f_{mosfet}(e_{gate}, e_{drain}, e_{source})$ . In our example the current  $j_D$  through the MOSFET is controlled by the input voltage  $U_{in}$  and the nodal potential  $e_1$  at node 1:  $j_D(U_{in}, e_1) := f_{mosfet}(U_{in}, e_1, 0)$ . Often the models of MOSFETs are very sophisticated and involve hundreds of parameters. A first order model for MOSFETs is described in [9], where also further references are given. We simply used

$$j_D = K \cdot (\max(U_{GS} - V_{th}, 0)^2 - \max(U_{GD} - V_{th}, 0)^2)$$

where  $U_{GS} = e_{gate} - e_{source}$ ,  $U_{GD} = e_{gate} - e_{drain}$  and the threshold voltage  $V_{th}$  and the current amplification factor  $K$  are given parameters. The thermal noise of the resistor and of the MOSFET is modelled by additional white noise current sources that are shunt in parallel to the original, noise-free elements. The noise intensity is given by Nyquist's rule (cf. [2, 5, 26]):

$$I_{th} = \sigma_R \cdot \xi(t) = \sqrt{\frac{2kTemp}{R}} \cdot \xi(t),$$

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 thermal noise source of the MOSFET this formula is modified by considering a solution dependent conductance  $g_D = g_{mosfet}(e_{gate}, e_{drain}, e_{source})$  instead of  $g = 1/R$ . We used

$$\begin{aligned} g_D &= 0, & \text{if } U_{GS} \leq V_{th}, \\ g_D &= \beta \cdot (U_{GS} - V_{th}) \cdot (1 + \lambda U_{DS}) & \text{if } 0 < U_{GS} - V_{th} \leq U_{DS}, \\ g_D &= \beta \cdot (U_{DS}) \cdot (1 + \lambda U_{DS}) & \text{if } 0 < U_{GS} - V_{th} > U_{DS}, \end{aligned}$$

where  $U_{DS} = e_{drain} - e_{source}$ , and  $\beta, \lambda$  are parameters. Applying Kirchhoff's Current Law gives a model for the output voltage  $e_1$  at node 1:

$$C e_1' - (U_{op} - e_1)/R + j_D(U_{in}, e_1) - \sigma_R \xi_1 + \sigma_D(U_{in}, e_1) \xi_2 = 0, \quad (5.4)$$

where  $\xi_1, \xi_2$  are independent standard Gaussian white noise processes. We treat this system as an Itô-SDE with  $n = 1$ ,  $m = 2$ ,  $f(x, t) = (U_{op} - x)/(C \cdot R) - j_D(U_{in}, x)/C$ ,  $g_1(x, t) = \sigma_R/C$ ,  $g_2(x, t) = -\sigma_D(U_{in}, x)/C$ .

In this simple model nearly no differences between the solutions of the noisy and the deterministic problem could be seen. Therefore, we dealt with a system where the diffusion coefficients had been scaled by a factor thousand. In Figure 5 we present simulation results for the parameters  $C = 2 \cdot 10^{-13}$  [F],  $R = 5 \cdot 10^3$  [ $\Omega$ ],  $U_{op} = 5$  [V],  $Temp = 300$  [K] on the time-interval  $[0, 2.5 \cdot 10^{-8}]$  [s]. In the MOSFET model we chose  $V_{th} = 1$ ,  $K = 10^{-3}$ ,  $\beta = 0.3 \cdot 10^{-4}$ ,  $\lambda = 0.02$ , and the tolerance was  $10^{-2}$ . In solid lines we plotted the values of the nodal potential  $e_1$  and the given input voltage  $U_{in}$  versus time. Moreover, the applied stepsizes, suitably scaled, are shown by means of single crosses. We compare simulation results for the computation of a single path (in the left picture of Figure 5) with those for the computation of 100 simultaneously computed solution paths (in the right picture of Figure 5), where we additionally plotted the mean  $\mu(E e_1)$  in the figure) and the lines  $\mu \pm 3\sigma$  ( $+3\sigma$ ,  $-3\sigma$  in the figure), where  $\sigma$  was computed as the empirical estimate of the standard deviation for the output voltage. Using the information of an ensemble of simultaneously computed solution paths smoothes the step-size-sequence and considerably reduces the number of rejected steps (from 46 to 9). Even the number of needed successful steps is reduced from 188 to 166.

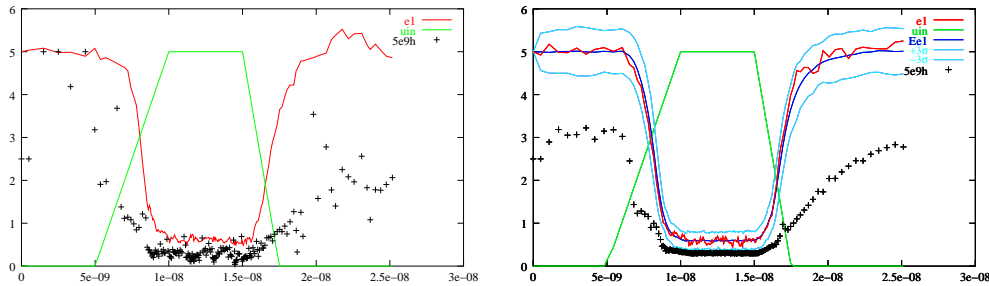


Figure 5: Simulation results for the noisy inverter circuit

1 path 188(+46 rejected) steps

100 paths 166(+9 rejected) steps

**Acknowledgment.** The authors would like to thank Uwe Feldmann and Georg Denk (Infineon Technologies) for invaluable discussions and the excellent cooperation and two anonymous referees for their remarks and suggestions which led to an improved presentation of the material.

#### 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, *em 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] M. GÜNTHER AND U.FELDMANN, *CAD-based electric-circuit modeling in industry II. Impact of circuit configurations and parameters*. Surv. Math. Ind. 8 (1999), pp. 131–157.
- [10] 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.
- [11] E. HAIRER, S.P. NØRSETT, AND G. WANNER, *Solving Ordinary Differential Equations I*, Springer, Berlin, 1987.
- [12] 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.
- [13] ———, *Step-size control for the uniform approximation of stochastic differential equations with additive noise*, Ann. Appl. Prob. 10 (2000), pp. 616–633.
- [14] P.E. KLOEDEN AND E. PLATEN, *Numerical Solution of Stochastic Differential Equations*, Springer, Berlin 1992.
- [15] H. LAMBA, J. MATTINGLY, A. STUART, *Strong convergence of an adaptive Euler-Maruyama scheme for stochastic differential equations*, manuscript, submitted, 2003.
- [16] 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.
- [17] X. MAO, *Stochastic Differential Equations and Their Applications*, Horwood Publishing, Chichester, 1997.
- [18] S. MAUTHNER, *Step size control in the numerical solution of stochastic differential equations*, J. Comput. Appl. Math. 100 (1998), pp. 93–109.
- [19] ———, *Schrittweitensteuerung bei der numerischen Lösung stochastischer Differentialgleichungen*, PhD thesis, Technische Universität Darmstadt, 1999.
- [20] 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.
- [21] ———, *Numerical Integration of Stochastic Differential Equations*. Kluwer, 1995.
- [22] G.N. MILSTEIN AND M.V. TRETAKOV, *Mean-square numerical methods for stochastic differential equations with small noise*, SIAM J. Sci. Comput. 18 (1997), pp. 1067–1087.
- [23] T. MÜLLER-GRONBACH, *Strong approximation of systems of stochastic differential equations*. Habilitation thesis, Technische Universität Darmstadt, 2002.
- [24] W. RÖMISCH AND R. WINKLER, *Stochastic DAEs in circuit simulation*, in: Modeling, Simulation and Optimization of Integrated Circuits, (Eds.) K.Antreich, R.Bulirsch, A. Gilg and P. Rentrop, Birkhäuser, 2003, 303–318.
- [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 of index 1 and applications in circuit simulation*, J. Comp. Appl. Math. 157 (2003), pp. 477–505.
- [28] ———, *Stochastic DAEs in transient noise simulation*, In: Scientific Computing in Electrical Engineering, Springer Series 'Mathematics in Industry', (Eds.) W. Schilders, J. ter Maten, S. Houben (2004), pp. 408–415.