U N I V E R S I T Ä T   H A M B U R G

Strong and Weak Approximation Methods for Stochastic Differential Equations – Some Recent Developments

Andreas Rößler

Preprint–No. 2010-02 Februar 2010

DEPARTMENT MATHEMATIK

SCHWERPUNKT MATHEMATISCHE STATISTIK
UND STOCHASTISCHE PROZESSE

BUNDESSTR. 55, D – 20146 HAMBURG

Strong and Weak Approximation Methods for Stochastic Differential Equations – Some Recent Developments

Andreas Rößler

Abstract Some efficient stochastic Runge–Kutta (SRK) methods for the strong as well as for the weak approximation of solutions of stochastic differential equations (SDEs) with improved computational complexity are considered. Their convergence is analyzed by a concise colored rooted tree approach for both, Itô as well as Stratonovich SDEs. Further, order conditions for the coefficients of order 1.0 and 1.5 strong SRK methods as well as for order 2.0 weak SRK methods are given. As the main novelty, the computational complexity of the presented order 1.0 strong SRK method and the order 2.0 weak SRK method depends only linearly on the dimension of the driving Wiener process. This is a significant improvement compared to well known methods where the computational complexity depends quadratically on the dimension of the Wiener process.

1 Approximation of Solutions of Stochastic Differential Equations

Let $(\Omega, \mathscr{F}, \mathrm{P})$ be a complete probability space with a filtration $(\mathscr{F}_t)_{t \geq 0}$ fulfilling the usual conditions and let $\mathscr{I} = [t_0, T]$ for some $0 \leq t_0 < T < \infty$. We denote by $X = (X_t)_{t \in \mathscr{I}}$ the solution of the d–dimensional SDE system

$$X_t = X_{t_0} + \int_{t_0}^{t} a(s, X_s) \, \mathrm{d}s + \sum_{j=1}^{m} \int_{t_0}^{t} b^j(s, X_s) * \mathrm{d}W_s^j \tag{1}$$

with an m–dimensional driving Wiener process $(W_t)_{t \geq 0} = ((W_t^1, \ldots, W_t^m)^T)_{t \geq 0}$ w.r.t. $(\mathscr{F}_t)_{t \geq 0}$ for $d, m \geq 1$ and $t \in \mathscr{I}$. We write $* \mathrm{d}W_s^j = \mathrm{d}W_s^j$ in the case of an Itô stochastic integral and $* \mathrm{d}W_s^j = \circ \mathrm{d}W_s^j$ for a Stratonovich stochastic integral. Suppose that

Andreas Rößler, Universität Hamburg, Department Mathematik, Bundesstraße 55, D-20146 Hamburg, Germany, e-mail: andreas.roessler@math.uni-hamburg.de

$a : \mathscr{I} \times \mathbb{R}^d \to \mathbb{R}^d$ and $b : \mathscr{I} \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ are continuous functions which fulfill a global Lipschitz condition and denote by b^j the jth column of the $d \times m$-matrix function $b = (b^{i,j})$ for $j = 1,\dots,m$. Let $X_{t_0} \in L^2(\Omega)$ be the $\mathscr{F}_{t_0}$-measurable initial value. In the following, we suppose that the conditions of the Existence and Uniqueness Theorem (cf., e.g., Kloeden and Platen (1999)) are fulfilled for SDE (1) and we denote by $\|\cdot\|$ the Euclidean norm. Let $C_P^l(\mathbb{R}^d,\mathbb{R})$ denote the space of all $g \in C^l(\mathbb{R}^d,\mathbb{R})$ with polynomial growth, see e.g. Kloeden and Platen (1999) or Rößler (2006,2009) for details. Then g belongs to $C_P^{k,l}(\mathscr{I} \times \mathbb{R}^d,\mathbb{R})$ if $g \in C^{k,l}(\mathscr{I} \times \mathbb{R}^d,\mathbb{R})$ and $g(t,\cdot) \in C_P^l(\mathbb{R}^d,\mathbb{R})$ is fulfilled uniformly in $t \in \mathscr{I}$.

For the numerical approximation let a discretization $\mathscr{I}_h = \{t_0,t_1,\dots,t_N\}$ with $t_0 < t_1 < \ldots < t_N = T$ of the time interval $\mathscr{I}$ with step sizes $h_n = t_{n+1} - t_n$ for $n = 0,1,\dots,N-1$ be given. Further, let $h = \max_{0\le n<N} h_n$ denote the maximum step size. If one is interested in a good pathwise approximation of the solution of SDE (1), then strong approximation methods converging in the mean–square sense are applied. Note that mean–square convergence implies strong convergence.

Definition 1. A sequence of approximation processes $Y^h = (Y(t))_{t\in\mathscr{I}_h}$ converges in the mean–square sense with order p to the solution X of SDE (1) at time T if there exists a constant $C > 0$ and some $\delta_0 > 0$ such that for each $h \in]0,\delta_0]$

$$(\mathrm{E}(\|X_T - Y^h(T)\|^2))^{1/2} \le C\,h^p\,. \tag{2}$$

However, if one is interested in the approximation of some distributional characteristics of the solution of SDE (1), then weak approximation methods are applied.

Definition 2. A sequence of approximation processes $Y^h = (Y(t))_{t\in\mathscr{I}_h}$ converges in the weak sense with order p to the solution X of SDE (1) at time T if for each $f \in C_P^{2(p+1)}(\mathbb{R}^d,\mathbb{R})$ exists a constant C_f and some $\delta_0 > 0$ such that for each $h \in]0,\delta_0]$

$$|\,\mathrm{E}(f(X_T)) - \mathrm{E}(f(Y^h(T)))| \le C_f\,h^p\,. \tag{3}$$

2 A General Class of Stochastic Runge–Kutta Methods

For the approximation of the solution X of SDE (1) we consider the universal class of stochastic Runge–Kutta (SRK) methods introduced in Rößler (2006): Let $\mathscr{M}$ be an arbitrary finite set of multi-indices with $\kappa = |\mathscr{M}|$ elements, let $\theta_\iota^{(k)}(h) \in L^2(\Omega)$ for $\iota \in \mathscr{M}$ and $0 \le k \le m$ be some suitable random variables. Further, define $b^0(t,x) := a(t,x)$. Then, an s–stages SRK method is given by $Y_0 = X_{t_0}$ and

$$Y_{n+1} = Y_n + \sum_{i=1}^{s} \sum_{k=0}^{m} \sum_{\nu\in\mathscr{M}} z_i^{(k),(\nu)}\, b^k\left(t_n + c_i^{(\nu)} h_n, H_i^{(\nu)}\right) \tag{4}$$

for $n = 0,1,\dots,N-1$ with $Y_n = Y(t_n)$, $t_n \in \mathscr{I}_h$, and with stages

$$H_i^{(\nu)} = Y_n + \sum_{j=1}^{s} \sum_{l=0}^{m} \sum_{\mu \in \mathcal{M}} Z_{ij}^{(\nu),(l),(\mu)} \, b^l \left(t_n + c_j^{(\mu)} h_n, H_j^{(\mu)} \right)$$

for $i = 1, \ldots, s$ and $\nu \in \mathcal{M}$. Here, let $0 \in \mathcal{M}$ and let for $i, j = 1, \ldots, s$

$$z_i^{(k),(\nu)} = \sum_{\iota \in \mathcal{M}} \gamma_i^{(\iota)}{}^{(k),(\nu)} \, \theta_\iota^{(k)}(h_n) \,, \qquad Z_{ij}^{(\nu),(l),(\mu)} = \sum_{\iota \in \mathcal{M}} C_{ij}^{(\iota)}{}^{(\nu),(l),(\mu)} \, \theta_\iota^{(l)}(h_n)$$

with $\theta_0^{(0)}(h_n) = h_n$ and the coefficients $\gamma_i^{(\iota)}{}^{(k),(\nu)}, C_{ij}^{(\iota)}{}^{(\nu),(l),(\mu)} \in \mathbb{R}$ of the SRK method. In the following, we use the notation $z^{(k),(\nu)} = (z_i^{(k),(\nu)})_{1 \le i \le s} \in \mathbb{R}^s$ and $Z^{(\nu),(l),(\mu)} = (Z_{ij}^{(\nu),(l),(\mu)})_{1 \le i,j \le s} \in \mathbb{R}^{s \times s}$. The vector of weights can be defined by

$$c^{(\nu)} = \sum_{\mu \in \mathcal{M}} C_{ij}^{(0)}{}^{(\nu),(0),(\mu)} e \tag{5}$$

with $e = (1, \ldots, 1)^T \in \mathbb{R}^s$. If $C_{ij}^{(\iota)}{}^{(\nu),(l),(\mu)} = 0$ for $j \ge i$ then (4) is called an explicit SRK method, otherwise it is called implicit. We assume that the random variables $\theta_\iota^{(k)}(h)$ satisfy the moment condition

$$\mathrm{E}\Big(\prod_{k=0}^{m} ((\theta_{\iota_1}^{(k)}(h))^{p_1^k} \cdot \ldots \cdot (\theta_{\iota_\kappa}^{(k)}(h))^{p_\kappa^k})\Big) = O\big(h^{p_1^0 + \ldots + p_\kappa^0 + \sum_{k=1}^m (p_1^k + \ldots + p_\kappa^k)/2}\big) \tag{6}$$

for all $p_i^k \in \mathbb{N}_0$, $k = 0, 1, \ldots, m$, and $\iota_i \in \mathcal{M}$, $1 \le i \le \kappa$. Further, we assume that in the case of an implicit method each random variable can be expressed as $\theta_\iota^{(0)}(h) = h \cdot \vartheta_\iota^{(0)}$ and $\theta_\iota^{(k)}(h) = \sqrt{h} \cdot \vartheta_\iota^{(k)}$, $1 \le k \le m$, for $\iota \in \mathcal{M}$ with suitable bounded random variables $\vartheta_\iota^{(0)}, \vartheta_\iota^{(k)} \in L^2(\Omega)$ such that each stage can be solved w.r.t. $H_i^{(\nu)}$ for sufficiently small h. These conditions are not necessary in the case of explicit SRK methods (see also Rößler (2006) or Milstein and Tretyakov (2004)).

3 Colored Rooted Tree Analysis

In the following, we present a concise rooted tree analysis for the convergence of the general class of SRK methods (4). For simplicity, we restrict our investigations without loss of generality to the autonomous SDE (1) in this section. We denote by TS the set of all stochastic trees, see also Rößler (2004,2010), which have a root $\tau_\gamma = \otimes$ and which can furthermore be composed of deterministic nodes $\tau_0 = \bullet$ and stochastic nodes $\tau_j = \bigcirc_j$ with some $j \in \{1, \ldots, m\}$. The index j is associated with the jth component of the m-dimensional driving Wiener process of the considered SDE. Some examples of trees in TS are presented in Fig. 1. Let $d(\mathbf{t})$ denote the number of deterministic nodes τ_0 and let $s(\mathbf{t})$ denote the number of stochastic nodes τ_j with $j \in \{1, \ldots, m\}$ of the tree $\mathbf{t} \in TS$. The order $\rho(\mathbf{t})$ of the tree $\mathbf{t} \in TS$ is defined

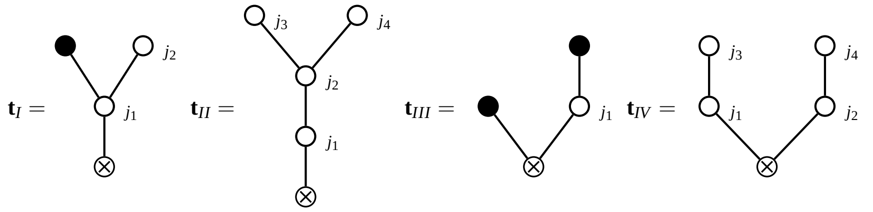


Fig. 1 Four elements of TS with some $j_1, j_2, j_3, j_4 \in \{1,\dots,m\}$.

as $\rho(\tau_\gamma) = 0$ and $\rho(\mathbf{t}) = d(\mathbf{t}) + \frac{1}{2}s(\mathbf{t})$. As an example, for the trees in Fig. 1 we have $\rho(\mathbf{t}_I) = \rho(\mathbf{t}_{II}) = \rho(\mathbf{t}_{IV}) = 2$ and $\rho(\mathbf{t}_{III}) = 2.5$.

Every tree can be written by a combination of brackets: If $\mathbf{t}_1,\dots,\mathbf{t}_k$ are colored subtrees then we denote by $[\mathbf{t}_1,\dots,\mathbf{t}_k]_j$ the tree in which $\mathbf{t}_1,\dots,\mathbf{t}_k$ are each joined by a single branch to the node τ_j for some $j \in \{\gamma,0,1,\dots,m\}$. Therefore proceeding recursively, for the trees in Fig. 1 we obtain $\mathbf{t}_I = [[\tau_0,\tau_{j_2}]_{j_1}]_\gamma$, $\mathbf{t}_{II} = [[[\tau_{j_3},\tau_{j_4}]_{j_2}]_{j_1}]_\gamma$, $\mathbf{t}_{III} = [\tau_0,[\tau_0]_{j_1}]_\gamma$ and $\mathbf{t}_{IV} = [[\tau_{j_3}]_{j_1},[\tau_{j_4}]_{j_2}]_\gamma$.

Next, we assign to each tree $\mathbf{t} \in TS$ an elementary differential which is defined recursively by $F(\tau_\gamma)(x) = f(x)$, $F(\tau_j)(x) = b^j(x)$ and

$$F(\mathbf{t})(x) = \begin{cases} f^{(k)}(x)\cdot(F(\mathbf{t}_1)(x),\dots,F(\mathbf{t}_k)(x)) & \text{for } \mathbf{t} = [\mathbf{t}_1,\dots,\mathbf{t}_k]_\gamma \\ b^{j\,(k)}(x)\cdot(F(\mathbf{t}_1)(x),\dots,F(\mathbf{t}_k)(x)) & \text{for } \mathbf{t} = [\mathbf{t}_1,\dots,\mathbf{t}_k]_j \end{cases} \tag{7}$$

for $j \in \{0,1,\dots,m\}$. Here $f^{(k)}$ and $b^{j\,(k)}$ define a symmetric k-linear differential operator, and one can choose the sequence of subtrees $\mathbf{t}_1,\dots,\mathbf{t}_k$ in an arbitrary order.

Finally, we assign to every tree a multiple stochastic integral. Let $(Z_t)_{t\geq t_0}$ be a progressively measurable stochastic process. Then, we define for $\mathbf{t} \in TS$ the corresponding multiple stochastic integral recursively by

$$I_{\mathbf{t};t_0,t}[Z_\cdot] = \begin{cases} \left(\prod_{i=1}^{k} I_{\mathbf{t}_i;t_0,t}\right)[Z_\cdot] & \text{if } \mathbf{t} = [\mathbf{t}_1,\dots,\mathbf{t}_k]_\gamma \\ \left(\int_{t_0}^{t} \prod_{i=1}^{k} I_{\mathbf{t}_i;t_0,s} * \mathrm{d}W_s^j\right)[Z_\cdot] & \text{if } \mathbf{t} = [\mathbf{t}_1,\dots,\mathbf{t}_k]_j,\ j \in \{0,1,\dots,m\} \end{cases} \tag{8}$$

with $*\mathrm{d}W_s^0 = \mathrm{d}s$, $I_{\tau_j;t_0,t}[Z_\cdot] = \int_{t_0}^t Z_s * \mathrm{d}W_s^j$, $I_{\tau_\gamma;t_0,t}[Z_\cdot] = Z_t$, $I_{\mathbf{t};t_0,t} = I_{\mathbf{t};t_0,t}[1]$ provided that the stochastic integral exists and by using the notation

$$\begin{aligned} &\left(\int_{t_0}^{t}\int_{t_0}^{s_n}\cdots\int_{t_0}^{s_2} *\mathrm{d}W_{s_1}^{j_1} * \mathrm{d}W_{s_2}^{j_2}\cdots * \mathrm{d}W_{s_n}^{j_n}\right)[Z_\cdot] = I_{(j_1,j_2,\dots,j_n)}[Z_\cdot]_{t_0,t} \\ &= \int_{t_0}^{t}\int_{t_0}^{s_n}\cdots\int_{t_0}^{s_2} Z_{s_1} * \mathrm{d}W_{s_1}^{j_1} * \mathrm{d}W_{s_2}^{j_2}\cdots * \mathrm{d}W_{s_n}^{j_n} \end{aligned} \tag{9}$$

in (8). The product of two stochastic integrals can be written as a sum (cf., e.g., Kloeden and Platen (1999))

$$\int_{t_0}^t X_s * \mathrm{d}W_s^i \int_{t_0}^t Y_s * \mathrm{d}W_s^j = \int_{t_0}^t X_s Y_s \mathbf{1}_{\{i=j\neq 0 \wedge * \neq \circ\}} \, \mathrm{d}s + \int_{t_0}^t X_s \big(\int_{t_0}^s Y_u * \mathrm{d}W_u^j \big) * \mathrm{d}W_s^i + \int_{t_0}^t \big(\int_{t_0}^s X_u * \mathrm{d}W_u^i \big) Y_s * \mathrm{d}W_s^j \tag{10}$$

for $0 \leq i, j \leq m$, where the first summand on the right hand side appears only in the case of Itô calculus. E.g., we calculate for $\mathbf{t}_I$ and $\mathbf{t}_{II}$

$$I_{\mathbf{t}_I;t_0,t}[1] = \int_{t_0}^t I_{\tau_0;t_0,s} I_{\tau_{j_2};t_0,s} * \mathrm{d}W_s^{j_1}[1] = I_{(0,j_2,j_1)}[1]_{t_0,t} + I_{(j_2,0,j_1)}[1]_{t_0,t} \,,$$

$$I_{\mathbf{t}_{II};t_0,t}[1] = \int_{t_0}^t \int_{t_0}^s I_{\tau_{j_3};t_0,u} I_{\tau_{j_4};t_0,u} * \mathrm{d}W_u^{j_2} * \mathrm{d}W_s^{j_1}$$
$$= I_{(j_3,j_4,j_2,j_1)}[1]_{t_0,t} + I_{(j_4,j_3,j_2,j_1)}[1]_{t_0,t} + I_{(0,j_2,j_1)}[\mathbf{1}_{\{j_3=j_4\neq 0 \wedge * \neq \circ\}}]_{t_0,t}$$

where the last summand for $I_{\mathbf{t}_{II};t_0,t}[1]$ only appears in the case of Itô calculus.

Table 1: All trees $\mathbf{t} \in TS$ of order $\rho(\mathbf{t}) \leq 1.5$ with $j_1, j_2, j_3 \in \{1, \ldots, m\}$ arbitrarily eligible.

$\mathbf{t}$	tree	$I_{\mathbf{t};t_0,t}$	$\sigma(\mathbf{t})$	$\rho(\mathbf{t})$
$\mathbf{t}_{0,1}$	τ_γ	1	1	0
$\mathbf{t}_{0.5,1}$	$[\tau_{j_1}]_\gamma$	$I_{(j_1)}[1]_{t_0,t}$	1	0.5
$\mathbf{t}_{1,1}$	$[\tau_0]_\gamma$	$I_{(0)}[1]_{t_0,t}$	1	1
$\mathbf{t}_{1,2}$	$[\tau_{j_1}, \tau_{j_2}]_\gamma$	$I_{(j_1,j_2)}[1]_{t_0,t} + I_{(j_2,j_1)}[1]_{t_0,t} + I_{(0)}[\mathbf{1}_{\{j_1=j_2 \wedge * \neq \circ\}}]_{t_0,t}$	$1 + \mathbf{1}_{\{j_1=j_2\}}$	1
$\mathbf{t}_{1,3}$	$[[\tau_{j_2}]_{j_1}]_\gamma$	$I_{(j_2,j_1)}[1]_{t_0,t}$	1	1
$\mathbf{t}_{1.5,1}$	$[[\tau_{j_1}]_0]_\gamma$	$I_{(j_1,0)}[1]_{t_0,t}$	1	1.5
$\mathbf{t}_{1.5,2}$	$[[\tau_0]_{j_1}]_\gamma$	$I_{(0,j_1)}[1]_{t_0,t}$	1	1.5
$\mathbf{t}_{1.5,3}$	$[\tau_0, \tau_{j_1}]_\gamma$	$I_{(0,j_1)}[1]_{t_0,t} + I_{(j_1,0)}[1]_{t_0,t}$	1	1.5
$\mathbf{t}_{1.5,4}$	$[\tau_{j_1}, \tau_{j_2}, \tau_{j_3}]_\gamma$	$I_{(j_1,j_2,j_3)}[1]_{t_0,t} + I_{(j_1,j_3,j_2)}[1]_{t_0,t} + I_{(j_2,j_1,j_3)}[1]_{t_0,t} + I_{(j_2,j_3,j_1)}[1]_{t_0,t} + I_{(j_3,j_1,j_2)}[1]_{t_0,t} + I_{(j_3,j_2,j_1)}[1]_{t_0,t} + (I_{(j_1,0)}[1]_{t_0,t} + I_{(0,j_1)}[1]_{t_0,t})\mathbf{1}_{\{j_2=j_3 \wedge * \neq \circ\}} + (I_{(j_2,0)}[1]_{t_0,t} + I_{(0,j_2)}[1]_{t_0,t})\mathbf{1}_{\{j_1=j_3 \wedge * \neq \circ\}} + (I_{(j_3,0)}[1]_{t_0,t} + I_{(0,j_3)}[1]_{t_0,t})\mathbf{1}_{\{j_1=j_2 \wedge * \neq \circ\}}$	$1 + \mathbf{1}_{\{j_1=j_2\neq j_3\}} + \mathbf{1}_{\{j_1=j_3\neq j_2\}} + \mathbf{1}_{\{j_2=j_3\neq j_1\}} + 5 \cdot \mathbf{1}_{\{j_1=j_2=j_3\}}$	1.5
$\mathbf{t}_{1.5,5}$	$[[\tau_{j_2}]_{j_1}, \tau_{j_3}]_\gamma$	$I_{(j_2,j_3,j_1)}[1]_{t_0,t} + I_{(j_3,j_2,j_1)}[1]_{t_0,t} + I_{(j_2,j_1,j_3)}[1]_{t_0,t} + I_{(0,j_1)}[\mathbf{1}_{\{j_2=j_3 \wedge * \neq \circ\}}]_{t_0,t} + I_{(j_2,0)}[\mathbf{1}_{\{j_1=j_3 \wedge * \neq \circ\}}]_{t_0,t}$	1	1.5
$\mathbf{t}_{1.5,6}$	$[[\tau_{j_2}, \tau_{j_3}]_{j_1}]_\gamma$	$I_{(j_3,j_2,j_1)}[1]_{t_0,t} + I_{(j_2,j_3,j_1)}[1]_{t_0,t} + I_{(0,j_1)}[\mathbf{1}_{\{j_2=j_3 \wedge * \neq \circ\}}]_{t_0,t}$	$1 + \mathbf{1}_{\{j_2=j_3\}}$	1.5
$\mathbf{t}_{1.5,7}$	$[[[\tau_{j_3}]_{j_2}]_{j_1}]_\gamma$	$I_{(j_3,j_2,j_1)}[1]_{t_0,t}$	1	1.5

Let $\mathbf{t} \in TS$ with $\mathbf{t} = [\mathbf{t}_1, \ldots, \mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_2, \ldots, \mathbf{t}_k, \ldots, \mathbf{t}_k]_j = [\mathbf{t}_1^{n_1}, \mathbf{t}_2^{n_2}, \ldots, \mathbf{t}_k^{n_k}]_j$, $j \in \{\gamma, 0, 1, \ldots, m\}$, where $\mathbf{t}_1, \ldots, \mathbf{t}_k$ are distinct subtrees with multiplicities $n_1, \ldots, n_k$, respectively. Then the symmetry factor σ is recursively defined by $\sigma(\tau_j) = 1$ and

$$\sigma(\mathbf{t}) = \prod_{i=1}^{k} n_i! \; \sigma(\mathbf{t}_i)^{n_i} . \tag{11}$$

For the trees in Fig. 1, we obtain $\sigma(\mathbf{t}_I) = \sigma(\mathbf{t}_{III}) = 1$. For the tree $\mathbf{t}_{II}$ we have to consider two cases: If $j_3 \neq j_4$ we have $\sigma(\mathbf{t}_{II}) = 1$. However, in the case of $j_3 = j_4$ we have some symmetry and thus we calculate $\sigma(\mathbf{t}_{II}) = 2$. Further, for tree $\mathbf{t}_{IV}$ we get $\sigma(\mathbf{t}_{IV}) = 2$ if $j_1 = j_2$ and $j_3 = j_4$ and $\sigma(\mathbf{t}_{IV}) = 1$ otherwise. E.g., all trees up to order 1.5 and the corresponding multiple integrals are presented in Tab. 1.

Next, we define the coefficient function Φ_S which assigns to every tree an elementary weight. For every $\mathbf{t} \in TS$ the function Φ_S is defined by $\Phi_S(\tau_\gamma) = 1$ and

$$\Phi_S(\mathbf{t}) = \begin{cases} \prod_{i=1}^{k} \Phi_S(\mathbf{t}_i) & \text{if } \mathbf{t} = [\mathbf{t}_1, \dots, \mathbf{t}_k]_\gamma \\ \sum_{\nu \in \mathscr{M}} z^{(j),(\nu)^T} \prod_{i=1}^{k} \Psi^{(\nu)}(\mathbf{t}_i) & \text{if } \mathbf{t} = [\mathbf{t}_1, \dots, \mathbf{t}_k]_j, \; j \in \{0,1,\dots,m\} \end{cases} \tag{12}$$

where $\Psi^{(\nu)}(\emptyset) = e$ with the representation $\tau_j = [\emptyset]_j$ and for each subtree $\mathbf{t} = [\mathbf{t}_1, \dots, \mathbf{t}_q]_l$ with some $l \in \{0,1,\dots,m\}$ we recursively define

$$\Psi^{(\nu)}(\mathbf{t}) = \sum_{\mu \in \mathscr{M}} Z^{(\nu),(l),(\mu)} \prod_{i=1}^{q} \Psi^{(\mu)}(\mathbf{t}_i). \tag{13}$$

Here $e = (1, \dots, 1)^T$ and the product of vectors in the definition of $\Psi^{(\nu)}$ is defined by component-wise multiplication, i.e. with $(a_1, \dots, a_n)(b_1, \dots, b_n) = (a_1 b_1, \dots, a_n b_n)$. In the following, we also write $\Phi_S(\mathbf{t}; t, t+h) = \Phi_S(\mathbf{t})$ in order to emphasize the dependency on the current time step with step size h.

Now, the following local Taylor expansions can be proved: For the solution X of SDE (1) and for $p \in \frac{1}{2}\mathbb{N}_0$ with $f \in C^{2p+2}(\mathbb{R}^d, \mathbb{R})$ and $a, b^j \in C^{2p+1}(\mathbb{R}^d, \mathbb{R}^d)$ for $j = 1, \dots, m$, we obtain the expansion (see Rößler (2004,2010))

$$f(X_t) = \sum_{\substack{\mathbf{t} \in TS \\ \rho(\mathbf{t}) \leq p}} F(\mathbf{t})(X_{t_0}) \frac{I_{\mathbf{t};t_0,t}}{\sigma(\mathbf{t})} + \mathscr{R}_p^*(t,t_0) \tag{14}$$

P-a.s. with remainder term $\mathscr{R}_p^*(t,t_0)$ provided all multiple Itô integrals exist. For the approximation Y by the SRK method (4) and for $p \in \frac{1}{2}\mathbb{N}_0$ with $f \in C^{2p+1}(\mathbb{R}^d, \mathbb{R})$ and $a, b^j \in C^{2p}(\mathbb{R}^d, \mathbb{R}^d)$, $j = 1, \dots, m$, we get the expansion (see Rößler (2006,2009))

$$f(Y(t)) = \sum_{\substack{\mathbf{t} \in TS \\ \rho(\mathbf{t}) \leq p}} F(\mathbf{t})(Y(t_0)) \frac{\Phi_S(\mathbf{t}; t_0, t)}{\sigma(\mathbf{t})} + \mathscr{R}_p^\Delta(t,t_0) \tag{15}$$

P-a.s. with remainder term $\mathscr{R}_p^\Delta(t,t_0)$.

4 Order Conditions for Stochastic Runge–Kutta Methods

Using the colored rooted tree analysis, we obtain order conditions for the random variables and the coefficients of the SRK method (4) if it is applied to SDE (1). The following results can be applied for the development of SRK methods for the Itô as well as the Stratonovich version of SDE (1). First, we consider conditions for strong convergence with some order $p \in \frac{1}{2}\mathbb{N}$ due to Rößler (2009). Therefore, let TS^* denote the set of trees $\mathbf{t} \in TS$ which have only one ramification at the root node τ_γ, i.e. which are of type $[[\ldots]_j]_\gamma$ for some $j \in \{0,1,\ldots,m\}$. The reason is, that we are interested in the approximation of X, thus we have to choose $f(x) = x$. However, in this case all elementary differentials vanish except for the trees in TS^*. For example, the trees $\mathbf{t}_{1,2}$, $\mathbf{t}_{1.5,3}$, $\mathbf{t}_{1.5,4}$ and $\mathbf{t}_{1.5,5}$ in Tab. 1 as well as the trees $\mathbf{t}_{III}$ and $\mathbf{t}_{IV}$ in Fig. 1 do not belong to TS^*. A comparison of the Taylor expansions (14) and (15) results in the following two theorems.

Theorem 1. *Let $p \in \frac{1}{2}\mathbb{N}_0$ and $a, b^j \in C^{\lceil p \rceil, 2p+1}(\mathscr{I} \times \mathbb{R}^d, \mathbb{R}^d)$ for $j = 1,\ldots,m$. Then, the SRK method (4) has mean–square order of accuracy p if the conditions*

a) for all $\boldsymbol{t} \in TS^*$ with $\rho(\boldsymbol{t}) \leq p$

$$I_{\boldsymbol{t};t,t+h} = \Phi_S(\boldsymbol{t};t,t+h) \qquad \text{P-}a.s., \tag{16}$$

b) for all $\boldsymbol{t} \in TS^*$ with $\rho(\boldsymbol{t}) = p + \frac{1}{2}$

$$\mathrm{E}(I_{\boldsymbol{t};t,t+h}) = \mathrm{E}(\Phi_S(\boldsymbol{t};t,t+h)), \tag{17}$$

are fulfilled for arbitrary $t, t+h \in \mathscr{I}$ and if (5) and (6) hold.

For the proof of Theorem 1 we refer to Rößler (2009). Next, we give conditions for the weak convergence of the SRK method (4) based on trees in TS having also multiple ramifications at the root node (see Theorem 6.4 in Rößler (2006)).

Theorem 2. *Let $p \in \mathbb{N}$ and $a, b^j \in C_P^{p+1,2p+2}(\mathscr{I} \times \mathbb{R}^d, \mathbb{R}^d)$ for $j = 1,\ldots,m$. Then the SRK method (4) is of weak order p if for all $\boldsymbol{t} \in TS$ with $\rho(\boldsymbol{t}) \leq p + \frac{1}{2}$ the order conditions*

$$\mathrm{E}(I_{\boldsymbol{t};t,t+h}) = \mathrm{E}(\Phi_S(\boldsymbol{t};t,t+h)) \tag{18}$$

are fulfilled for arbitrary $t, t+h \in \mathscr{I}$, provided that (5) and (6) apply and that the approximation Y has uniformly bounded moments w.r.t. the number N of steps.

For the proof of Theorem 2 we refer to Rößler (2006).

Remark 1. The approximation Y by the SRK method (4) has uniformly bounded moments if bounded random variables are used by the method, if (6) is fulfilled and if $\mathrm{E}(z^{(k,\nu)^T} e) = 0$ for $1 \leq k \leq m$ and $\nu \in \mathscr{M}$ (see Rößler (2006) for details). Further, Theorem 2 provides uniform weak convergence with order p in the case of a non-random time discretization $\mathscr{I}_h$.

5 Strong Approximation of SDEs

For higher order strong numerical approximation methods for SDEs, the simulation of multiple stochastic integrals is necessary in general. Therefore, for $t_n, t_{n+1} \in \mathscr{I}_h$ and $1 \le i, j \le m$ let

$$I_{(i),n} = \int_{t_n}^{t_{n+1}} \mathrm{d}W_s^i, \qquad I_{(i,j),n} = \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} \mathrm{d}W_u^i \, \mathrm{d}W_s^j,$$

denote the multiple Itô stochastic integrals. For convenience we write e.g. $I_{(i)} = I_{(i),n}$ if n is obvious from the context. The increments of the Wiener process $I_{(i),n}$ are independent $N(0,h_n)$ distributed with $h_n = t_{n+1} - t_n$. From (10) follows that $I_{(0,i)} = h_n I_{(i)} - I_{(i,0)}$. In the case of $i = j$, formula (10) results in $I_{(i,i)} = \frac{1}{2}(I_{(i)}^2 - h_n)$. Further, let $I_{(i,i,i)} = \frac{1}{6}(I_{(i)}^3 - 3I_{(0)} I_{(i)})$. In the following, the multiple integrals $I_{(i,0)}$ can be simulated by $I_{(i,0)} = \frac{1}{2} h_n (I_{(i)} + \frac{1}{\sqrt{3}} \zeta_i)$ with some independent $N(0,h_n)$ distributed random variables ζ_i which are independent from $I_{(j)}$ for all $1 \le j \le m$ (cf., e.g., Kloeden and Platen (1999) or Milstein (1995)). However, since the exact distribution and thus the exact simulation of the multiple stochastic integrals $I_{(i,j)}$ for $1 \le i, j \le m$ with $i \ne j$ is not known, we substitute them in our numerical experiments by sufficiently exact and efficient approximations as recently proposed by Wiktorsson (2001). Further, let (p_D, p_S) with $p_D \ge p_S$ denote the order of convergence of the considered SRK scheme if it is applied to a deterministic or stochastic differential equation, respectively.

5.1 Order 1.0 Strong SRK Methods

Firstly, we consider an efficient order 1.0 strong SRK method for Itô SDEs (1). Yet, known derivative free order 1.0 strong approximation methods suffer from an inefficiency in the case of an m–dimensional driving Wiener process. For example, the derivative free scheme (11.1.7) in Kloeden and Platen (1999) needs one evaluation of the drift coefficient a, however $m+1$ evaluations of each diffusion coefficient b^j, $j = 1, \ldots, m$, each step. Thus, the computational complexity grows quadratically in m which is a significant drawback especially for high dimensional problems. Therefore, efficient SRK methods were firstly proposed in Rößler (2009) where the number of necessary evaluations of each drift and each diffusion coefficient is independent of the dimension m of the driving Wiener process.

For the multi–dimensional Itô SDE (1) with $d, m \ge 1$, the efficient s–stages order 1.0 strong SRK method due to Rößler (2009) is given by $Y_0 = X_{t_0}$ and

$$
\begin{aligned}
Y_{n+1} = Y_n &+ \sum_{i=1}^{s} \alpha_i \, a(t_n + c_i^{(0)} h_n, H_i^{(0)}) \, h_n \\
&+ \sum_{k=1}^{m} \sum_{i=1}^{s} \big(\beta_i^{(1)} I_{(k)} + \beta_i^{(2)} \sqrt{h_n} \big) \, b^k(t_n + c_i^{(1)} h_n, H_i^{(k)})
\end{aligned}
\tag{19}
$$

for $n = 0, 1, \ldots, N-1$ with stages

$$
\begin{aligned}
H_i^{(0)} &= Y_n + \sum_{j=1}^{s} A_{ij}^{(0)} \, a(t_n + c_j^{(0)} h_n, H_j^{(0)}) \, h_n + \sum_{l=1}^{m} \sum_{j=1}^{s} B_{ij}^{(0)} \, b^l(t_n + c_j^{(1)} h_n, H_j^{(l)}) \, I_{(l)} \\
H_i^{(k)} &= Y_n + \sum_{j=1}^{s} A_{ij}^{(1)} \, a(t_n + c_j^{(0)} h_n, H_j^{(0)}) \, h_n + \sum_{l=1}^{m} \sum_{j=1}^{s} B_{ij}^{(1)} \, b^l(t_n + c_j^{(1)} h_n, H_j^{(l)}) \, \frac{I_{(l,k)}}{\sqrt{h_n}}
\end{aligned}
\tag{20}
$$

for $i = 1, \ldots, s$ and $k = 1, \ldots, m$. A modified version of the efficient SRK method (19) suitable for Stratonovich SDEs can be found in Rößler (2009). The SRK method (19) can be characterized by its coefficients given by an extended Butcher tableau:

$$
\begin{array}{c|c|c|c}
c^{(0)} & A^{(0)} & B^{(0)} & \\
\hline
c^{(1)} & A^{(1)} & B^{(1)} & \\
\hline
 & \alpha^T & {\beta^{(1)}}^T & {\beta^{(2)}}^T
\end{array}
\tag{21}
$$

Here, the class of SRK methods (4) is applied with $\mathscr{M} = \{\nu : 0 \le \nu \le m\}$ and

$$
\begin{aligned}
z_i^{(0),(0)} &= \alpha_i \, h_n , & Z_{ij}^{(0),(0),(0)} &= A_{ij}^{(0)} h_n , & Z_{ij}^{(0),(k),(k)} &= B_{ij}^{(0)} \, I_{(k)} , \\
z_i^{(k),(k)} &= \beta_i^{(1)} I_{(k)} + \beta_i^{(2)} \sqrt{h_n} , & Z_{ij}^{(k),(0),(0)} &= A_{ij}^{(1)} h_n , & Z_{ij}^{(k),(l),(l)} &= B_{ij}^{(1)} \frac{I_{(l,k)}}{\sqrt{h_n}} ,
\end{aligned}
$$

for $1 \le k, l \le m$ and all other coefficients in (4) are set equal to zero. Thus, the presented SRK method (19) belongs to the general class (4). The application of the rooted tree analysis and Theorem 1 gives order conditions up to strong order 1.0 for the coefficients of the SRK method (19), see also Rößler (2009).

Theorem 3. *Let $a, b^j \in C^{1,2}(\mathscr{I} \times \mathbb{R}^d, \mathbb{R}^d)$ for $j = 1, \ldots, m$. If the coefficients of the SRK method (19) fulfill the equations*

1. $\alpha^T e = 1$ 2. ${\beta^{(1)}}^T e = 1$ 3. ${\beta^{(2)}}^T e = 0$

then the method attains order 0.5 for the strong approximation of the solution of the Itô SDE (1). If $a, b^j \in C^{1,3}(\mathscr{I} \times \mathbb{R}^d, \mathbb{R}^d)$ for $j = 1, \ldots, m$ and if in addition the equations

4. ${\beta^{(1)}}^T B^{(1)} e = 0$ 5. ${\beta^{(2)}}^T B^{(1)} e = 1$ 6. ${\beta^{(2)}}^T A^{(1)} e = 0$
7. ${\beta^{(2)}}^T (B^{(1)} e)^2 = 0$ 8. ${\beta^{(2)}}^T (B^{(1)} (B^{(1)} e)) = 0$

0				0			
0	0	0		1	1	0	
0	0 0	0 0		0	0 0	0 0	
0				0			
0	0	1		1	1	1	
0	0 0	-1 0		1	1 0	-1 0	
	1 0 0	1 0 0	0 $\frac{1}{2}$ $-\frac{1}{2}$		$\frac{1}{2}$ $\frac{1}{2}$ 0	1 0 0	0 $\frac{1}{2}$ $-\frac{1}{2}$

Table 2 Coefficients for the strong SRK schemes SRI1 of order $(1.0,1.0)$ on the left hand side and SRI2 of order $(2.0,1.0)$ on the right hand side.

are fulfilled and if $c^{(i)} = A^{(i)}e$ for $i = 0,1$, then the SRK method (19) attains order 1.0 for the strong approximation of the solution of the Itô SDE (1).

For the detailed proof of Theorem 3 we refer to Rößler (2009). The Euler–Maruyama scheme EM is the basic explicit order 0.5 strong SRK scheme with $s = 1$ stage, $\alpha_1 = \beta_1^{(1)} = 1$ and $\beta_1^{(2)} = A_{1,1}^{(0)} = A_{1,1}^{(1)} = B_{1,1}^{(0)} = B_{1,1}^{(1)} = 0$. As an example for some explicit order 1.0 strong SRK schemes, the coefficients presented in Tab. 2 define the order $(1.0,1.0)$ strong SRK scheme SRI1 and the order $(2.0,1.0)$ strong SRK scheme SRI2. As the main advantage, the scheme SRI1 needs one evaluation of the drift coefficient a and only 3 evaluations of each diffusion coefficient b^j, $j = 1,\ldots,m$, each step. Thus, the number of evaluations of the drift and diffusion coefficients is independent of the dimension m of the Wiener process.

5.2 Order 1.5 Strong SRK Methods for SDEs with Scalar Noise

In contrast to the multi-dimensional Wiener process case, higher order 1.5 strong approximation methods can be applied if the driving Wiener process is scalar. E.g., order 1.5 strong SRK methods for Stratonovich SDEs with a scalar Wiener process have been proposed by Burrage and Burrage (1996,2000). On the other hand, for Itô SDEs with a scalar Wiener process order 1.5 strong SRK methods have been proposed by Kaneko (1995) and by Kloeden and Platen (1999). However, the scheme due to Kaneko (1995) is not efficient because it needs 4 evaluations of the drift coefficient a, 12 evaluations of the diffusion coefficient b and the simulation of two independent normally distributed random variables for each step. On the other hand, the scheme (11.2.1) in Kloeden and Platen (1999) due to Platen needs 3 evaluations of the drift coefficient a, 5 evaluations of the diffusion b and also the simulation of two independent normally distributed random variables each step. In contrast to this, we consider the order 1.5 strong SRK method for Itô SDEs with less computational complexity proposed in Rößler (2009).

For the Itô SDE (1) with $d \geq 1$ and $m = 1$ the efficient order 1.5 strong SRK method due to Rößler (2009) is defined by $Y_0 = X_{t_0}$ and

$$
\begin{aligned}
Y_{n+1} = Y_n &+ \sum_{i=1}^{s} \alpha_i \, a(t_n + c_i^{(0)} h_n, H_i^{(0)}) \, h_n \\
&+ \sum_{i=1}^{s} \big(\beta_i^{(1)} I_{(1)} + \beta_i^{(2)} \frac{I_{(1,1)}}{\sqrt{h_n}} + \beta_i^{(3)} \frac{I_{(1,0)}}{h_n} + \beta_i^{(4)} \frac{I_{(1,1,1)}}{h_n} \big) \, b(t_n + c_i^{(1)} h_n, H_i^{(1)})
\end{aligned}
\tag{22}
$$

for $n = 0, 1, \ldots, N-1$ with stages

$$
\begin{aligned}
H_i^{(0)} &= Y_n + \sum_{j=1}^{s} A_{ij}^{(0)} \, a(t_n + c_j^{(0)} h_n, H_j^{(0)}) \, h_n + \sum_{j=1}^{s} B_{ij}^{(0)} \, b(t_n + c_j^{(1)} h_n, H_j^{(1)}) \, \frac{I_{(1,0)}}{h_n} \\
H_i^{(1)} &= Y_n + \sum_{j=1}^{s} A_{ij}^{(1)} \, a(t_n + c_j^{(0)} h_n, H_j^{(0)}) \, h_n + \sum_{j=1}^{s} B_{ij}^{(1)} \, b(t_n + c_j^{(1)} h_n, H_j^{(1)}) \, \sqrt{h_n}
\end{aligned}
\tag{23}
$$

for $i = 1, \ldots, s$. A more general version of the order 1.5 strong SRK method (22) for SDEs with diagonal noise and a simplified version for additive noise can be found in Rößler (2009). The SRK method (22) is characterized by the Butcher tableau:

$$
\begin{array}{c|c|c|c}
c^{(0)} & A^{(0)} & B^{(0)} & \\
\hline
c^{(1)} & A^{(1)} & B^{(1)} & \\
\hline
 & \alpha^T & {\beta^{(1)}}^T & {\beta^{(2)}}^T \\
\hline
 & & {\beta^{(3)}}^T & {\beta^{(4)}}^T
\end{array}
\tag{24}
$$

For the SRK method (22) we choose $\mathscr{M} = \{0, 1\}$ and we then define

$$
\begin{aligned}
& z_i^{(0),(0)} = \alpha_i \, h_n \,, && z_i^{(1),(1)} = \beta_i^{(1)} \, I_{(1)} + \beta_i^{(2)} \, \frac{I_{(1,1)}}{\sqrt{h_n}} + \beta_i^{(3)} \, \frac{I_{(1,0)}}{h_n} + \beta_i^{(4)} \, \frac{I_{(1,1,1)}}{h_n} \,, \\
& Z_{ij}^{(0),(0),(0)} = A_{ij}^{(0)} \, h_n \,, && Z_{ij}^{(0),(1),(1)} = B_{ij}^{(0)} \, \frac{I_{(1,0)}}{h_n} \,, \\
& Z_{ij}^{(1),(0),(0)} = A_{ij}^{(1)} \, h_n \,, && Z_{ij}^{(1),(1),(1)} = B_{ij}^{(1)} \, \sqrt{h_n} \,,
\end{aligned}
$$

with all remaining coefficients in (4) defined equal to zero. Then, the SRK method (22) is also covered by the class (4) of SRK methods. Thus, we can apply Theorem 1 with $p = 1.5$ to obtain strong order 1.5 conditions, see Rößler (2009) for details.

Theorem 4. *Let $a, b \in C^{1,2}(\mathscr{I} \times \mathbb{R}^d, \mathbb{R}^d)$. If the coefficients of the SRK method (22) fulfill the equations*

1. $\alpha^T e = 1$ 2. ${\beta^{(1)}}^T e = 1$ 3. ${\beta^{(2)}}^T e = 0$
4. ${\beta^{(3)}}^T e = 0$ 5. ${\beta^{(4)}}^T e = 0$

then the method attains order 0.5 for the strong approximation of the solution of the Itô SDE (1). If $a, b \in C^{1,3}(\mathscr{I} \times \mathbb{R}^d, \mathbb{R}^d)$ and if in addition the equations

$$
\begin{array}{c|cccc|cccc|cccc}
0 & & & & & & & & & & & & \\
\frac{3}{4} & \frac{3}{4} & & & & \frac{3}{2} & & & & & & & \\
0 & 0 & 0 & & & 0 & 0 & & & & & & \\
0 & 0 & 0 & 0 & & 0 & 0 & 0 & & & & & \\
\hline
0 & & & & & & & & & & & & \\
\frac{1}{4} & \frac{1}{4} & & & & \frac{1}{2} & & & & & & & \\
1 & 1 & 0 & & & -1 & 0 & & & & & & \\
\frac{1}{4} & 0 & 0 & \frac{1}{4} & & -5 & 3 & \frac{1}{2} & & & & & \\
\hline
 & \frac{1}{3} & \frac{2}{3} & 0 & 0 & -1 & \frac{4}{3} & \frac{2}{3} & 0 & -1 & \frac{4}{3} & -\frac{1}{3} & 0 \\
\hline
 & & & & & 2 & -\frac{4}{3} & -\frac{2}{3} & 0 & -2 & \frac{5}{3} & -\frac{2}{3} & 1
\end{array}
\qquad
\begin{array}{c|cccc|cccc|cccc}
0 & & & & & & & & & & & & \\
1 & 1 & & & & 0 & & & & & & & \\
\frac{1}{2} & \frac{1}{4} & \frac{1}{4} & & & 1 & \frac{1}{2} & & & & & & \\
0 & 0 & 0 & 0 & & 0 & 0 & 0 & & & & & \\
\hline
0 & & & & & & & & & & & & \\
\frac{1}{4} & \frac{1}{4} & & & & -\frac{1}{2} & & & & & & & \\
1 & 1 & 0 & & & 1 & 0 & & & & & & \\
\frac{1}{4} & 0 & 0 & \frac{1}{4} & & 2 & -1 & \frac{1}{2} & & & & & \\
\hline
 & \frac{1}{6} & \frac{1}{6} & \frac{2}{3} & 0 & -1 & \frac{4}{3} & \frac{2}{3} & 0 & 1 & -\frac{4}{3} & \frac{1}{3} & 0 \\
\hline
 & & & & & 2 & -\frac{4}{3} & -\frac{2}{3} & 0 & -2 & \frac{5}{3} & -\frac{2}{3} & 1
\end{array}
$$

Table 3 Strong SRK scheme SRI1W1 of order $(2.0, 1.5)$ on the left hand side and SRI2W1 of order $(3.0, 1.5)$ on the right hand side.

6. $\beta^{(1)^T} B^{(1)} e = 0$ 7. $\beta^{(2)^T} B^{(1)} e = 1$

8. $\beta^{(3)^T} B^{(1)} e = 0$ 9. $\beta^{(4)^T} B^{(1)} e = 0$

are fulfilled and if $c^{(i)} = A^{(i)} e$ for $i = 0, 1$, then the SRK method (22) attains order 1.0 for the strong approximation of the solution of the Itô SDE (1) with scalar noise. If $a, b \in C^{2,4}(\mathscr{I} \times \mathbb{R}^d, \mathbb{R}^d)$ and if in addition the equations

10. $\alpha^T A^{(0)} e = \frac{1}{2}$ 15. $\beta^{(3)^T} A^{(1)} e = -1$ 20. $\beta^{(4)^T} (B^{(1)} e)^2 = 2$

11. $\alpha^T B^{(0)} e = 1$ 16. $\beta^{(4)^T} A^{(1)} e = 0$ 21. $\beta^{(1)^T} (B^{(1)} (B^{(1)} e)) = 0$

12. $\alpha^T (B^{(0)} e)^2 = \frac{3}{2}$ 17. $\beta^{(1)^T} (B^{(1)} e)^2 = 1$ 22. $\beta^{(2)^T} (B^{(1)} (B^{(1)} e)) = 0$

13. $\beta^{(1)^T} A^{(1)} e = 1$ 18. $\beta^{(2)^T} (B^{(1)} e)^2 = 0$ 23. $\beta^{(3)^T} (B^{(1)} (B^{(1)} e)) = 0$

14. $\beta^{(2)^T} A^{(1)} e = 0$ 19. $\beta^{(3)^T} (B^{(1)} e)^2 = -1$ 24. $\beta^{(4)^T} (B^{(1)} (B^{(1)} e)) = 1$

25. $\frac{1}{2} \beta^{(1)^T} (A^{(1)} (B^{(0)} e)) + \frac{1}{3} \beta^{(3)^T} (A^{(1)} (B^{(0)} e)) = 0$

are fulfilled and if $c^{(i)} = A^{(i)} e$ for $i = 0, 1$, then the SRK method (22) attains order 1.5 for the strong approximation of the solution of the Itô SDE (1) in the case of scalar noise.

For a proof of Theorem 4 we refer to Rößler (2009). Coefficients for the order 1.5 strong SRK schemes SRI1W1 of order $(2.0, 1.5)$ and SRI2W1 of order $(3.0, 1.5)$ are given in Tab. 3. The SRK scheme SRI1W1 needs only 2 evaluations of the drift coefficient, 4 evaluations of the diffusion coefficient b and the simulation of two independent normally distributed random variables for each step. Note that the explicit 2–stages SRK method (22) with coefficients $\alpha_1 = \beta_1^{(1)} = \beta_2^{(2)} = A_{2,1}^{(1)} =$

		number of evaluations			random variables		
scheme	order	a^k	$b^{k,j}$	$\frac{\partial b^{k,j}}{\partial x^l}$	$I_{(j)}$	$I_{(j,0)}$	$I_{(i,j)}$
EM	0.5	d	md	–	+	–	–
MIL	1.0	d	md	md^2	+	–	+
SPLI	1.0	d	$(m^2+m)d$	–	+	–	+
SPLIW1	1.5	$3d$	$5d$	–	+	+	–
SRI1	1.0	d	$3md$	–	+	–	+
SRI1W1	1.5	$2d$	$4d$	–	+	+	–

Table 4 Computational complexity of some schemes per step for a d–dimensional SDE system with a m–dimensional Wiener process ($m = 1$ for SPLIW1 and SRI1W1).

$B^{(1)}_{2,1} = 1$, $\beta^{(2)}_1 = -1$ and $\alpha_2 = \beta^{(1)}_2 = A^{(0)}_{2,1} = B^{(0)}_{2,1} = \beta^{(3)}_1 = \beta^{(3)}_2 = \beta^{(4)}_1 = \beta^{(4)}_2 = 0$ coincides with the order 1.0 strong scheme (11.1.3) in Kloeden and Platen (1999).

5.3 Numerical Results

The presented efficient SRK methods are applied to some test SDEs in order to analyze their performance. Let EM denote the order 0.5 strong Euler–Maruyama scheme and let MIL denote the order 1.0 strong Milstein scheme in Milstein (1995). Further, the order 1.0 strong scheme (11.1.7) denoted as SPLI and the order 1.5 strong scheme (11.2.1) called SPLIW1 for Itô SDEs with scalar noise in Kloeden and Platen (1999) are applied. As a measure for the computational effort, we take the number of evaluations of the drift and diffusion coefficients as well as the number of realizations of (normally distributed) random variables needed each step. If the approximation method needs the random variables $I_{(i,j)}$ for $1 \leq i, j \leq m$ with $i \neq j$, then $I_{(i,j)}$ is simulated by the method due to Wiktorsson (2001) and we need to simulate $\frac{1}{2}m(m-1) + 2mq$ independent normally distributed random variables each step with $q \leq \lceil \sqrt{5m^2(m-1)/(24\pi^2)}\, h^{-1/2} \rceil$ in the mean (see Wiktorsson (2001)), provided that the m random variables $I_{(i)}$ are given. Thus, the additional computational effort increases with order $O(h^{-1/2})$ as $h \rightarrow 0$. The computational complexity is given in Tab. 4. E.g., the computational complexity of the scheme MIL is $d + md + md^2 + m + \frac{1}{2}m(m-1) + 2mq$ whereas scheme SRI1 has only complexity $d + 3md + m + \frac{1}{2}m(m-1) + 2mq$ each step. Thus, the scheme SRI1 has lower computational complexity than the Milstein scheme MIL in the case of $d > 2$ and $m \geq 1$ even if we neglect the effort for the calculation of the derivatives of b^j needed by the Milstein scheme. Further, the scheme SRI1 has also lower computational complexity than the scheme SPLI1 due to Platen in the case of $d \geq 1$ and $m > 2$.

We simulate 2000 trajectories and take the mean of the attained errors at $T = 1$ as an estimator for the expectation in (2). Then, we analyze the mean–square errors versus the computational effort as well as versus the step size in log–log–diagrams with base two. We denote by p_{eff} the effective order of convergence which is the

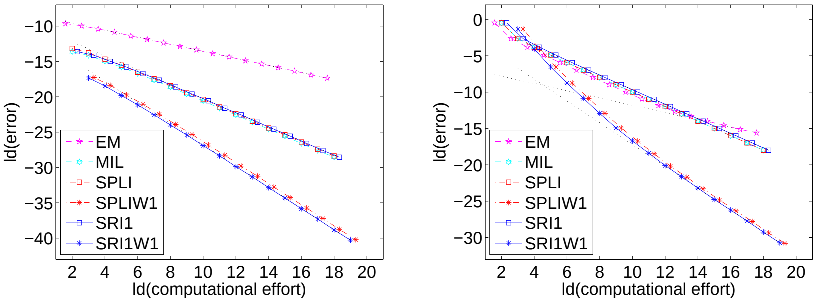


Fig. 2 Errors vs. effort for SDE (25) and SDE (26) with $d = m = 1$.

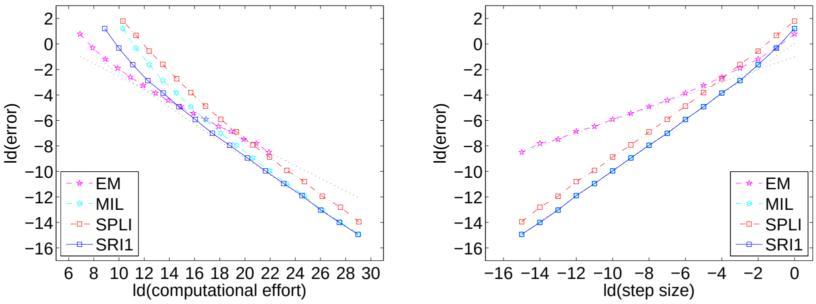


Fig. 3 Errors vs. effort for SDE (26) and errors vs. step sizes for SDE (26) with $d = m = 10$.

slope of the resulting line in the mean–square errors versus effort diagrams. Considering the effective order may cause an order reduction such that an strong order 1.0 scheme attains the effective order $p_{\text{eff}} = 2/3$ as $h \longrightarrow 0$. This is due to the effort for the simulation of the multiple integrals $I_{(i,j)}$ which depends on h. Dotted order lines with slope 0.5, 1.0, 2/3 and 1.5 are plotted as a reference. Clearly, a more efficient method to simulate the multiple integrals $I_{(i,j)}$ would result in a higher effective order. However, compared to the Euler–Maruyama scheme EM with $p_{\text{eff}} = 0.5$, there is still a significantly improved convergence for the order 1.0 methods. As a result of this, the order 1.0 strong approximation methods are superior to the order 0.5 strong Euler–Maruyama scheme, which is also confirmed by the simulation results.

As the first example, consider for $d = m = 1$ the nonlinear Itô SDE

$$dX_t = -\left(\frac{1}{10}\right)^2 \sin(X_t)\cos^3(X_t)\,dt + \frac{1}{10}\cos^2(X_t)\,dW_t\,, \quad X_0 = 1\,, \tag{25}$$

with solution $X_t = \arctan(\frac{1}{10}W_t + \tan(X_0))$ in Kloeden and Platen (1999). The results for $h = 2^0, \ldots, 2^{-16}$ are plotted on the left of Fig. 2. Scheme SRI1W1 has effective order 1.5 and performs better than the other schemes due to its reduced complexity.

In order to consider also a multi-dimensional Itô SDE with $d,m \geq 1$, we define $A \in \mathbb{R}^{d \times d}$ as a matrix with entries $A_{ij} = \frac{1}{20}$ if $i \neq j$ and $A_{ii} = -\frac{3}{2}$ for $1 \leq i,j \leq d$. Further, define $B^k \in \mathbb{R}^{d \times d}$ by $B^k_{ij} = \frac{1}{100}$ for $i \neq j$ and $B^k_{ii} = \frac{1}{5}$ for $1 \leq i,j \leq d$ and $k = 1,\dots,m$. Then, we consider the Itô SDE

$$\mathrm{d}X_t = AX_t\,\mathrm{d}t + \sum_{k=1}^{m} B^k X_t\,\mathrm{d}W_t^k\,, \quad X_0 = (1,\dots,1)^T \in \mathbb{R}^d\,, \tag{26}$$

with solution $X_t = X_0 \exp((A - \frac{1}{2}\sum_{k=1}^{m}(B^k)^2)\,t + \sum_{k=1}^{m} B^k W_t^k)$. For the case of $d = m = 1$ the numerical results for $h = 2^0,\dots,2^{-16}$ are presented on the right of Fig. 2 where the scheme SRI1W1 has the best performance. On the other hand, for the case of $d = m = 10$ the effective and the strong orders are analyzed for $h = 2^0,\dots,2^{-15}$ in Fig. 3. Here, the schemes MIL, SPLI, and SRI1 have strong order 1.0 while the Euler–Maruyama scheme EM has order 1/2. Further, due to the effort for the simulation of the multiple integrals, all order 1.0 strong schemes attain the effective order $2/3$ and thus perform significantly better than the Euler–Maruyama scheme EM with effective order 1/2. The scheme SRI1 shows the best performance, especially compared to the Milstein scheme MIL and the scheme SPLI.

6 Weak Approximation of SDEs

In contrast to strong approximation methods, we now consider methods which are designed for the approximation of distributional characteristics of the solution of SDEs. Numerical methods for the weak approximation do not need information about the driving Wiener process, their random variables can be simulated on a different probability space. Therefore, we can make use of random variables with distributions which are easy to simulate. In the following, we make use of random variables which are defined by

$$\hat{I}_{(k,l)} = \begin{cases} \frac{1}{2}(\hat{I}_{(k)}\hat{I}_{(l)} - \sqrt{h_n}\tilde{I}_{(k)}) & \text{if } k < l \\ \frac{1}{2}(\hat{I}_{(k)}\hat{I}_{(l)} + \sqrt{h_n}\tilde{I}_{(l)}) & \text{if } l < k \\ \frac{1}{2}(\hat{I}_{(k)}^2 - h_n) & \text{if } k = l \end{cases} \tag{27}$$

for $1 \leq k,l \leq m$ with independent random variables $\hat{I}_{(k)}$, $1 \leq k \leq m$, and random variables $\tilde{I}_{(k)}$, $1 \leq k \leq m-1$, possessing the moments

$$\mathrm{E}(\hat{I}_{(k)}^q) = \begin{cases} 0 & \text{for } q \in \{1,3,5\} \\ (q-1)h_n^{q/2} & \text{for } q \in \{2,4\} \\ O(h_n^{q/2}) & \text{for } q \geq 6 \end{cases}, \quad \mathrm{E}(\tilde{I}_{(k)}^q) = \begin{cases} 0 & \text{for } q \in \{1,3\} \\ h_n & \text{for } q = 2 \\ O(h_n^{q/2}) & \text{for } q \geq 4 \end{cases}. \tag{28}$$

Thus, only $2m-1$ independent random variables are needed for each step $n = 0,1,\dots,N-1$. For example, we can choose $\hat{I}_{(k)}$ as three point distributed random variables with $\mathrm{P}(\hat{I}_{(k)} = \pm\sqrt{3h_n}) = \frac{1}{6}$ and $\mathrm{P}(\hat{I}_{(k)} = 0) = \frac{2}{3}$. The random variables $\tilde{I}_{(k)}$ can be defined by a two point distribution with $\mathrm{P}(\tilde{I}_{(k)} = \pm\sqrt{h_n}) = \frac{1}{2}$.

6.1 Order 2.0 Weak SRK Methods

We consider the class of efficient SRK methods introduced in Rößler (2009) for the weak approximation of the solution of the Itô SDE (1) where the number of stages s is independent of the dimension m of the driving Wiener process. A similar class of second order SRK methods for the Stratonovich version of SDE (1) can be found in Rößler (2007). For the Itô SDE (1) the d-dimensional SRK approximation Y with $Y_n = Y(t_n)$ for $t_n \in \mathscr{I}_h$ due to Rößler (2009) is defined by $Y_0 = x_0$ and

$$\begin{aligned} Y_{n+1} = Y_n &+ \sum_{i=1}^{s} \alpha_i\, a(t_n + c_i^{(0)} h_n, H_i^{(0)})\, h_n \\ &+ \sum_{i=1}^{s}\sum_{k=1}^{m} \beta_i^{(1)}\, b^k(t_n + c_i^{(1)} h_n, H_i^{(k)}) \hat{I}_{(k)} + \sum_{i=1}^{s}\sum_{k=1}^{m} \beta_i^{(2)}\, b^k(t_n + c_i^{(1)} h_n, H_i^{(k)}) \frac{\hat{I}_{(k,k)}}{\sqrt{h_n}} \\ &+ \sum_{i=1}^{s}\sum_{k=1}^{m} \beta_i^{(3)}\, b^k(t_n + c_i^{(2)} h_n, \hat{H}_i^{(k)}) \hat{I}_{(k)} + \sum_{i=1}^{s}\sum_{k=1}^{m} \beta_i^{(4)}\, b^k(t_n + c_i^{(2)} h_n, \hat{H}_i^{(k)}) \sqrt{h_n} \end{aligned} \tag{29}$$

for $n = 0, 1, \dots, N-1$ with stage values

$$\begin{aligned} H_i^{(0)} &= Y_n + \sum_{j=1}^{s} A_{ij}^{(0)}\, a(t_n + c_j^{(0)} h_n, H_j^{(0)})\, h_n + \sum_{j=1}^{s}\sum_{l=1}^{m} B_{ij}^{(0)}\, b^l(t_n + c_j^{(1)} h_n, H_j^{(l)})\, \hat{I}_{(l)} \\ H_i^{(k)} &= Y_n + \sum_{j=1}^{s} A_{ij}^{(1)}\, a(t_n + c_j^{(0)} h_n, H_j^{(0)})\, h_n + \sum_{j=1}^{s} B_{ij}^{(1)}\, b^k(t_n + c_j^{(1)} h_n, H_j^{(k)})\, \sqrt{h_n} \\ \hat{H}_i^{(k)} &= Y_n + \sum_{j=1}^{s} A_{ij}^{(2)}\, a(t_n + c_j^{(0)} h_n, H_j^{(0)})\, h_n + \sum_{j=1}^{s}\sum_{\substack{l=1 \\ l \neq k}}^{m} B_{ij}^{(2)}\, b^l(t_n + c_j^{(1)} h_n, H_j^{(l)})\, \frac{\hat{I}_{(k,l)}}{\sqrt{h_n}} \end{aligned}$$

for $i = 1, \dots, s$ and $k = 1, \dots, m$. In the case of a scalar driving Wiener process, i.e. for $m = 1$, the SRK method (29) reduces to the SRK method proposed in Rößler (2006). The coefficients of the SRK method (29) can be represented by an extended Butcher array:

$$\begin{array}{c|c|c|c} c^{(0)} & A^{(0)} & B^{(0)} & \\ \hline c^{(1)} & A^{(1)} & B^{(1)} & \\ \hline c^{(2)} & A^{(2)} & B^{(2)} & \\ \hline & \alpha^T & {\beta^{(1)}}^T & {\beta^{(2)}}^T \\ \cline{2-4} & & {\beta^{(3)}}^T & {\beta^{(4)}}^T \end{array}$$

Applying the rooted tree analysis and Theorem 2 with $p = 2$, we obtain order two conditions for the SRK method (29) which were calculated in Rößler (2009).

Theorem 5. *Let $a^i, b^{i,j} \in C_P^{2,4}(\mathscr{I} \times \mathbb{R}^d, \mathbb{R})$ for $1 \le i \le d$, $1 \le j \le m$. If the coefficients of the stochastic Runge–Kutta method (29) fulfill the equations*

1. $\alpha^T e = 1$	2. ${\beta^{(4)}}^T e = 0$	3. ${\beta^{(3)}}^T e = 0$
4. $({\beta^{(1)}}^T e)^2 = 1$	5. ${\beta^{(2)}}^T e = 0$	6. ${\beta^{(1)}}^T B^{(1)} e = 0$
7. ${\beta^{(4)}}^T A^{(2)} e = 0$	8. ${\beta^{(3)}}^T B^{(2)} e = 0$	9. ${\beta^{(4)}}^T (B^{(2)} e)^2 = 0$

then the method attains order 1 for the weak approximation of the solution of the Itô SDE (1). Further, if $a^i, b^{i,j} \in C_P^{3,6}(\mathscr{I} \times \mathbb{R}^d, \mathbb{R})$ for $1 \le i \le d$, $1 \le j \le m$ and if in addition the equations

10. $\alpha^T A^{(0)} e = \frac{1}{2}$	11. $\alpha^T (B^{(0)} e)^2 = \frac{1}{2}$
12. $({\beta^{(1)}}^T e)(\alpha^T B^{(0)} e) = \frac{1}{2}$	13. $({\beta^{(1)}}^T e)({\beta^{(1)}}^T A^{(1)} e) = \frac{1}{2}$
14. ${\beta^{(3)}}^T A^{(2)} e = 0$	15. ${\beta^{(2)}}^T B^{(1)} e = 1$
16. ${\beta^{(4)}}^T B^{(2)} e = 1$	17. $({\beta^{(1)}}^T e)({\beta^{(1)}}^T (B^{(1)} e)^2) = \frac{1}{2}$
18. $({\beta^{(1)}}^T e)({\beta^{(3)}}^T (B^{(2)} e)^2) = \frac{1}{2}$	19. ${\beta^{(1)}}^T (B^{(1)} (B^{(1)} e)) = 0$
20. ${\beta^{(3)}}^T (B^{(2)} (B^{(1)} e)) = 0$	21. ${\beta^{(3)}}^T (B^{(2)} (B^{(1)} (B^{(1)} e))) = 0$
22. ${\beta^{(1)}}^T (A^{(1)} (B^{(0)} e)) = 0$	23. ${\beta^{(3)}}^T (A^{(2)} (B^{(0)} e)) = 0$
24. ${\beta^{(4)}}^T (A^{(2)} e)^2 = 0$	25. ${\beta^{(4)}}^T (A^{(2)} (A^{(0)} e)) = 0$
26. $\alpha^T (B^{(0)} (B^{(1)} e)) = 0$	27. ${\beta^{(2)}}^T A^{(1)} e = 0$
28. ${\beta^{(1)}}^T ((A^{(1)} e)(B^{(1)} e)) = 0$	29. ${\beta^{(3)}}^T ((A^{(2)} e)(B^{(2)} e)) = 0$
30. ${\beta^{(4)}}^T (A^{(2)} (B^{(0)} e)) = 0$	31. ${\beta^{(2)}}^T (A^{(1)} (B^{(0)} e)) = 0$
32. ${\beta^{(4)}}^T ((B^{(2)} e)^2 (A^{(2)} e)) = 0$	33. ${\beta^{(4)}}^T (A^{(2)} (B^{(0)} e)^2) = 0$
34. ${\beta^{(2)}}^T (A^{(1)} (B^{(0)} e)^2) = 0$	35. ${\beta^{(1)}}^T (B^{(1)} (A^{(1)} e)) = 0$
36. ${\beta^{(3)}}^T (B^{(2)} (A^{(1)} e)) = 0$	37. ${\beta^{(2)}}^T (B^{(1)} e)^2 = 0$
38. ${\beta^{(4)}}^T (B^{(2)} (B^{(1)} e)) = 0$	39. ${\beta^{(2)}}^T (B^{(1)} (B^{(1)} e)) = 0$
40. ${\beta^{(1)}}^T (B^{(1)} e)^3 = 0$	41. ${\beta^{(3)}}^T (B^{(2)} e)^3 = 0$
42. ${\beta^{(1)}}^T (B^{(1)} (B^{(1)} e)^2) = 0$	43. ${\beta^{(3)}}^T (B^{(2)} (B^{(1)} e)^2) = 0$
44. ${\beta^{(4)}}^T (B^{(2)} e)^4 = 0$	45. ${\beta^{(4)}}^T (B^{(2)} (B^{(1)} e))^2 = 0$
46. ${\beta^{(4)}}^T ((B^{(2)} e)(B^{(2)} (B^{(1)} e))) = 0$	47. $\alpha^T ((B^{(0)} e)(B^{(0)} (B^{(1)} e))) = 0$

48. $\beta^{(1)^T}((A^{(1)}(B^{(0)}e))(B^{(1)}e)) = 0$
49. $\beta^{(3)^T}((A^{(2)}(B^{(0)}e))(B^{(2)}e)) = 0$
50. $\beta^{(1)^T}(A^{(1)}(B^{(0)}(B^{(1)}e))) = 0$
51. $\beta^{(3)^T}(A^{(2)}(B^{(0)}(B^{(1)}e))) = 0$
52. $\beta^{(4)^T}((B^{(2)}(A^{(1)}e))(B^{(2)}e)) = 0$
53. $\beta^{(1)^T}(B^{(1)}(A^{(1)}(B^{(0)}e))) = 0$
54. $\beta^{(3)^T}(B^{(2)}(A^{(1)}(B^{(0)}e))) = 0$
55. $\beta^{(1)^T}((B^{(1)}e)(B^{(1)}(B^{(1)}e))) = 0$
56. $\beta^{(3)^T}((B^{(2)}e)(B^{(2)}(B^{(1)}e))) = 0$
57. $\beta^{(1)^T}(B^{(1)}(B^{(1)}(B^{(1)}e))) = 0$
58. $\beta^{(4)^T}((B^{(2)}e)(B^{(2)}(B^{(1)}(B^{(1)}e)))) = 0$
59. $\beta^{(4)^T}((B^{(2)}e)(B^{(2)}(B^{(1)}e)^2)) = 0$

are fulfilled and if $c^{(i)} = A^{(i)}e$ for $i = 0,1,2$, then the stochastic Runge–Kutta method (29) attains order 2 for the weak approximation of the solution of the Itô SDE (1).

Proof. We only give a sketch of the proof and refer to Rößler (2009) for the detailed proof. Calculating the order conditions by Theorem 2, it turns out that there are some trees which restrict the class of efficient SRK methods significantly and which give a deep insight to the necessary structure of such methods. Therefore, we concentrate our investigation to the trees

$$\mathbf{t}_{2,12} = [\tau_{j_1}, \tau_{j_2}, [\tau_{j_4}]_{j_3}]_\gamma, \qquad \mathbf{t}_{2,15} = [[\tau_{j_2}]_{j_1}, [\tau_{j_4}]_{j_3}]_\gamma, \tag{30}$$

with some $j_1, j_2, j_3, j_4 \in \{1,\dots,m\}$. Then, we have $l(\mathbf{t}_{2,12}) = l(\mathbf{t}_{2,15}) = 5$, $\rho(\mathbf{t}_{2,12}) = \rho(\mathbf{t}_{2,15}) = 2$ and $s(\mathbf{t}_{2,12}) = s(\mathbf{t}_{2,15}) = 4$. Now, for the SRK method (29) we choose $\mathscr{M} = \{(0),(\nu),(\nu,0),(\nu,1) : 1 \le \nu \le m\}$ and

$$z_i^{(0),(0)} = \alpha_i h_n, \quad z_i^{(k),(k,0)} = \beta_i^{(1)} \hat{I}_{(k)} + \beta_i^{(2)} \frac{\hat{I}_{(k,k)}}{\sqrt{h_n}}, \quad z_i^{(k),(k,1)} = \beta_i^{(3)} \hat{I}_{(k)} + \beta_i^{(4)} \sqrt{h_n},$$

$$\begin{aligned} Z_{ij}^{(0),(0),(0)} &= A_{ij}^{(0)} h_n, & Z_{ij}^{(k,0),(0),(0)} &= A_{ij}^{(1)} h_n, & Z_{ij}^{(k,1),(0),(0)} &= A_{ij}^{(2)} h_n, \\ Z_{ij}^{(0),(k),(k,0)} &= B_{ij}^{(0)} \hat{I}_{(k)}, & Z_{ij}^{(k,0),(k),(k,0)} &= B_{ij}^{(1)} \sqrt{h_n}, & Z_{ij}^{(k,1),(l),(l,0)} &= B_{ij}^{(2)} \frac{\hat{I}_{(k,l)}}{\sqrt{h_n}}, \end{aligned}$$

for $1 \le k,l \le m$ with $k \ne l$ and with $H_i^{(k,0)} = H_i^{(k)}$ and $H_i^{(k,1)} = \hat{H}_i^{(k)}$ for $1 \le i,j \le s$. Thus, the class of SRK methods is covered by the general class (4). Then, the coefficient function (12) yields

$$\begin{aligned} \Phi_S(\mathbf{t}_{2,12}) &= (z^{(j_1),(j_1,0)^T} e + z^{(j_1),(j_1,1)^T} e)(z^{(j_2),(j_2,0)^T} e + z^{(j_2),(j_2,1)^T} e) \\ &\quad \times (z^{(j_3),(j_3,0)^T} Z^{(j_3,0),(j_4),(j_4,0)} e + z^{(j_3),(j_3,1)^T} Z^{(j_3,1),(j_4),(j_4,0)} e), \\ \Phi_S(\mathbf{t}_{2,15}) &= (z^{(j_1),(j_1,0)^T} Z^{(j_1,0),(j_2),(j_2,0)} e + z^{(j_1),(j_1,1)^T} Z^{(j_1,1),(j_2),(j_2,0)} e) \\ &\quad \times (z^{(j_3),(j_3,0)^T} Z^{(j_3,0),(j_4),(j_4,0)} e + z^{(j_3),(j_3,1)^T} Z^{(j_3,1),(j_4),(j_4,0)} e), \end{aligned} \tag{31}$$

for $j_1, j_2, j_3, j_4 \in \{1,\dots,m\}$. Further, the multiple stochastic integrals are

$$\begin{aligned}
I_{\mathbf{t}_{2,12};t,t+h} = {} & I_{(j_4,j_3,j_2,j_1);t,t+h} + I_{(j_4,j_3,j_1,j_2);t,t+h} + I_{(j_1,j_4,j_3,j_2);t,t+h} + I_{(j_4,j_1,j_3,j_2);t,t+h} \\
& + I_{(0,j_3,j_2);t,t+h}[\mathbf{1}_{\{j_1=j_4\}}] + I_{(j_4,0,j_2);t,t+h}[\mathbf{1}_{\{j_1=j_3\}}] + I_{(j_4,j_3,0);t,t+h}[\mathbf{1}_{\{j_1=j_2\}}] \\
& + I_{(j_4,j_2,j_3,j_1);t,t+h} + I_{(j_4,j_2,j_1,j_3);t,t+h} + I_{(j_1,j_4,j_2,j_3);t,t+h} + I_{(j_4,j_1,j_2,j_3);t,t+h} \\
& + I_{(0,j_2,j_3);t,t+h}[\mathbf{1}_{\{j_1=j_4\}}] + I_{(j_4,0,j_3);t,t+h}[\mathbf{1}_{\{j_1=j_2\}}] + I_{(j_4,j_2,0);t,t+h}[\mathbf{1}_{\{j_1=j_3\}}] \\
& + I_{(j_2,j_4,j_3,j_1);t,t+h} + I_{(j_2,j_4,j_1,j_3);t,t+h} + I_{(j_1,j_2,j_4,j_3);t,t+h} + I_{(j_2,j_1,j_4,j_3);t,t+h} \\
& + I_{(0,j_4,j_3);t,t+h}[\mathbf{1}_{\{j_1=j_2\}}] + I_{(j_2,0,j_3);t,t+h}[\mathbf{1}_{\{j_1=j_4\}}] + I_{(j_2,j_4,0);t,t+h}[\mathbf{1}_{\{j_1=j_3\}}] \\
& + I_{(0,j_3,j_1);t,t+h}[\mathbf{1}_{\{j_2=j_4\}}] + I_{(j_1,0,j_3);t,t+h}[\mathbf{1}_{\{j_2=j_4\}}] + I_{(0,j_1,j_3);t,t+h}[\mathbf{1}_{\{j_2=j_4\}}] \\
& + I_{(j_4,0,j_1);t,t+h}[\mathbf{1}_{\{j_2=j_3\}}] + I_{(j_1,j_4,0);t,t+h}[\mathbf{1}_{\{j_2=j_3\}}] + I_{(j_4,j_1,0);t,t+h}[\mathbf{1}_{\{j_2=j_3\}}] \\
& + I_{(0,0);t,t+h}[\mathbf{1}_{\{j_2=j_4\}}\mathbf{1}_{\{j_1=j_3\}}] + I_{(0,0);t,t+h}[\mathbf{1}_{\{j_2=j_3\}}\mathbf{1}_{\{j_1=j_4\}}]
\end{aligned}$$

and

$$\begin{aligned}
I_{\mathbf{t}_{2,15};t,t+h} = {} & I_{(j_4,j_3,j_2,j_1);t,t+h} + I_{(j_4,j_2,j_3,j_1);t,t+h} + I_{(j_2,j_4,j_3,j_1);t,t+h} + I_{(j_2,j_1,j_4,j_3);t,t+h} \\
& + I_{(0,j_3,j_1);t,t+h}[\mathbf{1}_{\{j_2=j_4\}}] + I_{(j_4,0,j_1);t,t+h}[\mathbf{1}_{\{j_2=j_3\}}] + I_{(j_2,0,j_3);t,t+h}[\mathbf{1}_{\{j_1=j_4\}}] \\
& + I_{(j_2,j_4,j_1,j_3);t,t+h} + I_{(j_4,j_2,j_1,j_3);t,t+h} + I_{(0,0);t,t+h}[\mathbf{1}_{\{j_1=j_3\}}\mathbf{1}_{\{j_2=j_4\}}] \\
& + I_{(j_4,j_2,0);t,t+h}[\mathbf{1}_{\{j_1=j_3\}}] + I_{(j_2,j_4,0);t,t+h}[\mathbf{1}_{\{j_1=j_3\}}] + I_{(0,j_1,j_3);t,t+h}[\mathbf{1}_{\{j_2=j_4\}}] \,.
\end{aligned}$$

If we apply Theorem 2 to $\mathbf{t}_{2,12}$ and $\mathbf{t}_{2,15}$, then we have to consider the cases $j_k = j_l$ and $j_k \neq j_l$ for $1 \leq k < l \leq 4$. In the case of $j_1 = j_2 = j_3 = j_4$ we obtain $\sigma(\mathbf{t}_{2,12}) = 2$ and $\mathrm{E}(I_{\mathbf{t}_{2,12};t,t+h}) = h^2$. The order condition (18) yields that $\mathrm{E}(\Phi_S(\mathbf{t}_{2,12};t,t+h)) = h^2$ has to be fulfilled. Applying (31) and taking into account the order conditions ${\beta^{(4)}}^T e = 0$ and ${\beta^{(2)}}^T e = 0$ due to the trees $\mathbf{t}_{0.5,1} = [\tau_{j_1}]_\gamma$ and $\mathbf{t}_{1.5,4} = [\tau_{j_1}, \tau_{j_2}, \tau_{j_3}]_\gamma$ (see Rößler (2009) for details) yields

$$\begin{aligned}
\mathrm{E}(\Phi_S(\mathbf{t}_{2,12})) &= \mathrm{E}\Big(\Big(\Big({\beta^{(1)}}^T e \hat{I}_{(j_1)} + {\beta^{(2)}}^T e \frac{\hat{I}_{(j_1,j_1)}}{\sqrt{h}}\Big) + \Big({\beta^{(3)}}^T e \hat{I}_{(j_1)} + {\beta^{(4)}}^T e \sqrt{h}\Big)\Big)^2 \\
&\quad \times \Big({\beta^{(1)}}^T B^{(1)} e \hat{I}_{(j_1)} \sqrt{h} + {\beta^{(2)}}^T B^{(1)} e \frac{\hat{I}_{(j_1,j_1)}}{\sqrt{h}} \sqrt{h}\Big)\Big) \\
&= ({\beta^{(1)}}^T e + {\beta^{(3)}}^T e)^2 ({\beta^{(2)}}^T B^{(1)} e)\, \mathrm{E}(\hat{I}_{(j_1)}^2 \hat{I}_{(j_1,j_1)}) \,.
\end{aligned}$$

Due to $\mathrm{E}(\hat{I}_{(j_1)}^2 \hat{I}_{(j_1,j_1)}) = h^2$, the order condition is fulfilled if for the coefficients holds $({\beta^{(1)}}^T e + {\beta^{(3)}}^T e)^2 ({\beta^{(2)}}^T B^{(1)} e) = 1$. In the case of $j_1 = j_3 \neq j_2 = j_4$ we calculate with $\sigma(\mathbf{t}_{2,12}) = 2$ and $\mathrm{E}(I_{\mathbf{t}_{2,12};t,t+h}) = \frac{1}{2} h^2$ from (18) the order condition $\mathrm{E}(\Phi_S(\mathbf{t}_{2,12};t,t+h)) = \frac{1}{2} h^2$. Then, we obtain for the SRK method (29)

$$\begin{aligned}
\mathrm{E}(\Phi_S(\mathbf{t}_{2,12})) &= \mathrm{E}\Big(\big((\beta^{(1)^T} e \hat{I}_{(j_1)} + \beta^{(2)^T} e \tfrac{\hat{I}_{(j_1,j_1)}}{\sqrt{h}}) + (\beta^{(3)^T} e \hat{I}_{(j_1)} + \beta^{(4)^T} e \sqrt{h})\big) \\
&\quad \times \big((\beta^{(1)^T} e \hat{I}_{(j_2)} + \beta^{(2)^T} e \tfrac{\hat{I}_{(j_2,j_2)}}{\sqrt{h}}) + (\beta^{(3)^T} e \hat{I}_{(j_2)} + \beta^{(4)^T} e \sqrt{h})\big) \\
&\quad \times (\beta^{(3)^T} B^{(2)} e \hat{I}_{(j_1)} \tfrac{\hat{I}_{(j_1,j_2)}}{\sqrt{h}} + \beta^{(4)^T} B^{(2)} e \sqrt{h} \tfrac{\hat{I}_{(j_1,j_2)}}{\sqrt{h}})\Big) \\
&= (\beta^{(1)^T} e + \beta^{(3)^T} e)^2 (\beta^{(4)^T} B^{(2)} e)\, \mathrm{E}(\hat{I}_{(j_1)} \hat{I}_{(j_2)} \hat{I}_{(j_1,j_2)}) \,.
\end{aligned}$$

Now, we can calculate that $\mathrm{E}(\hat{I}_{(j_1)} \hat{I}_{(j_2)} \hat{I}_{(j_1,j_2)}) = \frac{1}{2} h^2$. Thus, the order condition is fulfilled if $(\beta^{(1)^T} e + \beta^{(3)^T} e)^2 (\beta^{(4)^T} B^{(2)} e) = 1$.

For $\mathbf{t}_{2,15}$, we calculate in the case of $j_1 = j_2 = j_3 = j_4$ with $\sigma(\mathbf{t}_{2,15}) = 2$ and $\mathrm{E}(I_{\mathbf{t}_{2,15};t,t+h}) = \frac{1}{2} h^2$ from (18) the order condition $\mathrm{E}(\Phi_S(\mathbf{t}_{2,15};t,t+h)) = \frac{1}{2} h^2$. Again, applying (31) results in

$$\begin{aligned}
\mathrm{E}(\Phi_S(\mathbf{t}_{2,15})) &= \mathrm{E}((\beta^{(1)^T} B^{(1)} e \hat{I}_{(j_1)} \sqrt{h} + \beta^{(2)^T} B^{(1)} e \tfrac{\hat{I}_{(j_1,j_1)}}{\sqrt{h}} \sqrt{h})^2) \\
&= (\beta^{(1)^T} B^{(1)} e)^2\, \mathrm{E}(\hat{I}^2_{(j_1)}) h + (\beta^{(2)^T} B^{(1)} e)^2\, \mathrm{E}(\hat{I}^2_{(j_1,j_1)}) \,.
\end{aligned}$$

Now, due to $\mathrm{E}(\hat{I}^2_{(j_1)}) = h$ and $\mathrm{E}(\hat{I}^2_{(j_1,j_1)}) = \frac{1}{2} h^2$ the order condition is $(\beta^{(1)^T} B^{(1)} e)^2 + \frac{1}{2} (\beta^{(2)^T} B^{(1)} e)^2 = \frac{1}{2}$. On the other hand, in the case of $j_1 = j_3 \neq j_2 = j_4$ with $\sigma(\mathbf{t}_{2,15}) = 2$ and $\mathrm{E}(I_{\mathbf{t}_{2,15};t,t+h}) = \frac{1}{2} h^2$, we get from (18) that $\mathrm{E}(\Phi_S(\mathbf{t}_{2,15};t,t+h)) = \frac{1}{2} h^2$ has to be fulfilled. Now, we obtain with (31) that

$$\begin{aligned}
\mathrm{E}(\Phi_S(\mathbf{t}_{2,15})) &= \mathrm{E}((\beta^{(3)^T} B^{(2)} e \hat{I}_{(j_1)} \tfrac{\hat{I}_{(j_1,j_2)}}{\sqrt{h}} + \beta^{(4)^T} B^{(2)} e \sqrt{h} \tfrac{\hat{I}_{(j_1,j_2)}}{\sqrt{h}})^2) \\
&= (\beta^{(3)^T} B^{(2)} e)^2\, \mathrm{E}(\hat{I}^2_{(j_1)} \hat{I}^2_{(j_1,j_2)}) h^{-1} + (\beta^{(4)^T} B^{(2)} e)^2\, \mathrm{E}(\hat{I}^2_{(j_1,j_2)}) \,.
\end{aligned}$$

Due to $\mathrm{E}(\hat{I}^2_{(j_1)} \hat{I}^2_{(j_1,j_2)}) = h^3$ and $\mathrm{E}(\hat{I}^2_{(j_1,j_2)}) = \frac{1}{2} h^2$, we finally get the order condition $(\beta^{(3)^T} B^{(2)} e)^2 + \frac{1}{2} (\beta^{(4)^T} B^{(2)} e)^2 = \frac{1}{2}$.

For all remaining cases of type $j_k = j_l$ or $j_k \neq j_l$ for $1 \leq k < l \leq 4$, we have $\mathrm{E}(I_{\mathbf{t}_{2,12};t,t+h}) = \mathrm{E}(I_{\mathbf{t}_{2,15};t,t+h}) = 0$ and we also calculate that $\mathrm{E}(\Phi_S(\mathbf{t}_{2,12};t,t+h)) = \mathrm{E}(\Phi_S(\mathbf{t}_{2,15};t,t+h)) = 0$. Therefore, (18) is fulfilled in these cases without any additional restrictions for the coefficients. Applying the rooted tree analysis and Theorem 2 to all remaining rooted trees up to order 2.5, we can calculate the complete order two conditions for the SRK method (29), see Rößler (2009). □

Remark 2. In the case of $m = 1$ and if we choose $A^{(2)}_{ij} = 0$ for $1 \leq i, j \leq s$ then the 59 conditions of Theorem 5 reduce to 28 conditions (see also Rößler (2006)). For an explicit SRK method of type (29) $s \geq 3$ is needed due to conditions 4., 6. and 17. Further, in the case of commutative noise significantly simplified SRK methods have been developed in Rößler (2004).

For example, the well known Euler-Maruyama scheme EM belongs to the introduced class of SRK methods having weak order 1 with $s = 1$ stage and with coeffi-

0									
1	1			$\frac{1}{3}$					
$\frac{5}{12}$	$\frac{25}{144}$	$\frac{35}{144}$		$-\frac{5}{6}$	0				
0									
$\frac{1}{4}$	$\frac{1}{4}$			$\frac{1}{2}$					
$\frac{1}{4}$	$\frac{1}{4}$	0		$-\frac{1}{2}$	0				
0									
0	0			1					
0	0	0		-1	0				
	$\frac{1}{10}$	$\frac{3}{14}$	$\frac{24}{35}$	1	-1	-1	0	1	-1
				$\frac{1}{2}$	$-\frac{1}{4}$	$-\frac{1}{4}$	0	$\frac{1}{2}$	$-\frac{1}{2}$

0									
1	1			1					
0	0	0		0	0				
0									
1	1			1					
1	1	0		-1	0				
0									
0	0			1					
0	0	0		-1	0				
	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	0	$\frac{1}{2}$	$-\frac{1}{2}$
				$-\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	0	$\frac{1}{2}$	$-\frac{1}{2}$

Table 5 Weak SRK scheme RI5 of order $p_D = 3$ and $p_S = 2$ and RI6 of order $p_D = p_S = 2$.

cients $\alpha_1 = \beta_1^{(1)} = 1$, $\beta_1^{(2)} = \beta_1^{(3)} = \beta_1^{(4)} = 0$, $A_{11}^{(0)} = A_{11}^{(1)} = 0$ and $B_{11}^{(0)} = B_{11}^{(1)} = 0$. We refer to Debrabant and Rößler (2009) for a detailed analysis of the solution space of the order conditions in Theorem 5 and for some coefficients which minimize the error constants of the SRK method (29). The SRK scheme RI5 presented on the left hand side of Table 5, is of order $p_S = 2$ and $p_D = 3$, while the SRK scheme RI6 on the right hand side of Table 5 is of order $p_D = p_S = 2$. Considering the computational complexity of the efficient SRK schemes RI5 and RI6, we take again the number of evaluations of the drift and diffusion functions and the number of random numbers needed in each step as a measure for the complexity of the schemes. Then, the SRK scheme RI5 needs 3 evaluations of the drift a while the SRK scheme RI6 needs 2 evaluations of a. Furthermore, we have to point out that only 5 evaluations of each diffusion function b^k for $k = 1, \dots, m$ are needed by both SRK schemes RI5 and RI6. This is due to the fact that the number of stages $s = 3$ does not depend on the dimension m of the driving Wiener process and because of $H_1^{(k)} = \hat{H}_1^{(k)}$, which saves one evaluation of each b^k in the case of explicit SRK schemes. As a further feature, only $2m-1$ independent random numbers have to be generated for the new SRK schemes in each step. Thus, the scheme RI6 has computational complexity $2d + 5md + 2m - 1$ while e.g. the order 2.0 weak SRK method PL1WM due to Platen (see Kloeden and Platen (1999) or Tocino and Vigo-Aguiar (2002)) has computational complexity $2d + (2m^2 + m)d + m + \frac{1}{2}m(m-1)$ which grows quadratically with the dimension m of the Wiener process. Thus, this is a significant reduction of complexity for the new SRK method (29) compared to well known SRK methods.

7 Numerical Results

We compare the schemes RI5 and RI6 with the order one Euler-Maruyama scheme EM, with the order 2.0 weak SRK scheme PL1WM due to Platen (see Kloeden and Platen (1999) or Tocino and Vigo-Aguiar (2002)) and with the extrapolated Euler-Maruyama scheme ExEu due to Talay and Tubaro (1990) attaining order two. In the following, we approximate $\mathrm{E}(f(X_T))$ for $f(x^1,\dots,x^d)=x^1$ by Monte Carlo simulation. Therefore, we estimate $\mathrm{E}(f(Y_T))$ by the sample average of M independently simulated realizations of the approximations $f(Y_{T,k})$, $k=1,\dots,M$, with $Y_{T,k}$ calculated by the scheme under consideration. The obtained errors at time $T=1.0$ are plotted versus the corresponding step sizes or the corresponding computational effort with double logarithmic scale in order to analyze the empirical order of convergence and the performance of the schemes, respectively.

The first test equation is a non-linear SDE system for $d=m=2$ with non-commutative noise given by

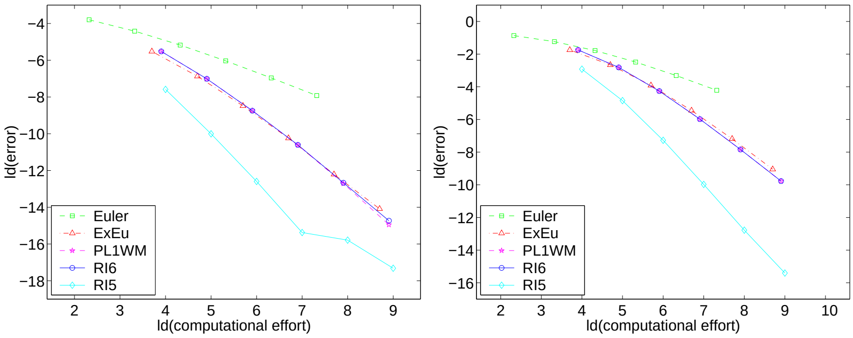
$$\mathrm{d}\begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2}X_t^1 + \frac{3}{2}X_t^2 \\ \frac{3}{2}X_t^1 - \frac{1}{2}X_t^2 \end{pmatrix} \mathrm{d}t + \begin{pmatrix} \sqrt{\frac{3}{4}(X_t^1)^2 - \frac{3}{2}X_t^1X_t^2 + \frac{3}{4}(X_t^2)^2 + \frac{3}{20}} \\ 0 \end{pmatrix} \mathrm{d}W_t^1$$
$$+ \begin{pmatrix} -\sqrt{\frac{1}{4}(X_t^1)^2 - \frac{1}{2}X_t^1X_t^2 + \frac{1}{4}(X_t^2)^2 + \frac{1}{20}} \\ \sqrt{(X_t^1)^2 - 2X_t^1X_t^2 + X_t^2 + \frac{1}{5}} \end{pmatrix} \mathrm{d}W_t^2, \tag{32}$$

with initial value $X_0=(\frac{1}{10},\frac{1}{10})^T$. Then, we calculate the first moments as $\mathrm{E}(X_t^i)=\frac{1}{10}\exp(t)$ for $i=1,2$. Here, we choose $M=10^9$ and the corresponding results are presented in Figure 4.

Next, we consider a non-linear SDE with non-commutative noise and some higher dimension $d=4$ which is given for $\lambda,\mu\in\{0,1\}$ as

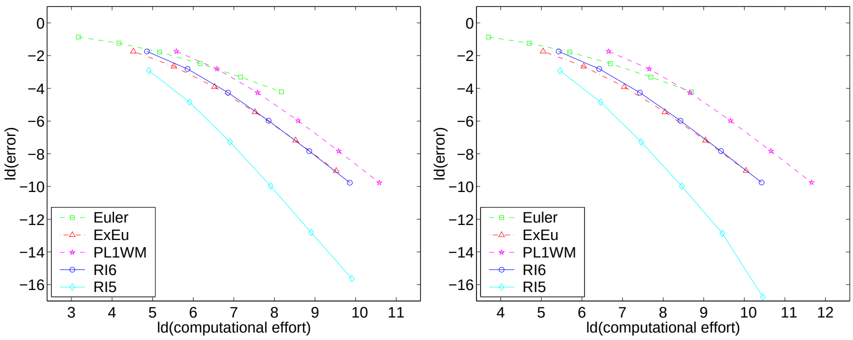
$$\mathrm{d}\begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \\ X_t^4 \end{pmatrix} = \begin{pmatrix} \frac{243}{154}X_t^1 - \frac{27}{77}X_t^2 + \frac{23}{154}X_t^3 - \frac{65}{154}X_t^4 \\ \frac{27}{77}X_t^1 - \frac{243}{154}X_t^2 + \frac{65}{154}X_t^3 - \frac{23}{154}X_t^4 \\ \frac{5}{154}X_t^1 - \frac{61}{154}X_t^2 + \frac{162}{77}X_t^3 - \frac{36}{77}X_t^4 \\ \frac{61}{154}X_t^1 - \frac{5}{154}X_t^2 + \frac{36}{77}X_t^3 - \frac{162}{77}X_t^4 \end{pmatrix} \mathrm{d}t$$
$$+ \frac{1}{9}\sqrt{(X_t^2)^2 + (X_t^3)^2 + \frac{2}{23}} \begin{pmatrix} \frac{1}{13} \\ \frac{1}{14} \\ \frac{1}{13} \\ \frac{1}{15} \end{pmatrix} \mathrm{d}W_t^1 + \frac{1}{8}\sqrt{(X_t^4)^2 + (X_t^1)^2 + \frac{1}{11}} \begin{pmatrix} \frac{1}{14} \\ \frac{1}{16} \\ \frac{1}{16} \\ \frac{1}{12} \end{pmatrix} \mathrm{d}W_t^2$$
$$+ \frac{\lambda}{12}\sqrt{(X_t^1)^2 + (X_t^2)^2 + \frac{1}{9}} \begin{pmatrix} \frac{1}{6} \\ \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{6} \end{pmatrix} \mathrm{d}W_t^3 + \frac{\lambda}{14}\sqrt{(X_t^3)^2 + (X_t^4)^2 + \frac{3}{29}} \begin{pmatrix} \frac{1}{8} \\ \frac{1}{9} \\ \frac{1}{8} \\ \frac{1}{9} \end{pmatrix} \mathrm{d}W_t^4$$
$$+ \frac{\mu}{10}\sqrt{(X_t^1)^2 + (X_t^3)^2 + \frac{1}{13}} \begin{pmatrix} \frac{1}{11} \\ \frac{1}{15} \\ \frac{1}{13} \\ \frac{1}{11} \end{pmatrix} \mathrm{d}W_t^5 + \frac{\mu}{11}\sqrt{(X_t^2)^2 + (X_t^4)^2 + \frac{2}{25}} \begin{pmatrix} \frac{1}{12} \\ \frac{1}{13} \\ \frac{1}{16} \\ \frac{1}{13} \end{pmatrix} \mathrm{d}W_t^6 \tag{33}$$

Fig. 4 Computational effort vs. error for the approximation of $\mathrm{E}(X_T^1)$ for SDE (32) in the left and for SDE (33) for $\lambda = \mu = 0$ with $m = 2$ in the right figure.



with initial value $X_0 = (\frac{1}{8}, \frac{1}{8}, 1, \frac{1}{8})^T$. Then, we have $m = 2 + 2\lambda + 2\mu$ as the dimension of the driving Wiener process. The moments of the solution can be calculated as $\mathrm{E}(X_T^i) = \frac{1}{8}\exp(2T)$ for $i = 1, 2, 4$ and $\mathrm{E}(X_T^3) = \exp(2T)$. We compare the performance of the considered schemes for the cases $m = 2$ with $\lambda = \mu = 0$, for $m = 4$ with $\lambda = 1$ and $\mu = 0$, and for $m = 6$ if $\lambda = \mu = 1$. Here, $M = 10^8$ independent trajectories are simulated and the results are presented in Figures 4–5. On the right

Fig. 5 Computational effort vs. error for the approximation of $\mathrm{E}(X_T^1)$ for SDE (33) for $\lambda = 1$, $\mu = 0$ with $m = 4$ in the left and for $\lambda = \mu = 1$ with $m = 6$ in the right figure.



hand side in Figure 4 and in Figure 5, we can see the performance of the considered schemes as the dimension m increases from 2 to 6. Comparing these results, we can see the significantly reduced complexity for the new SRK schemes RI5 and RI6 compared to the well known SRK scheme PL1WM in the case of $m > 2$. This benefit becomes more and more significant if we increase the dimension m of the driving Wiener process, which confirms our theoretical results. For the considered examples, we obtained very good results especially for the SRK scheme RI5 having order $p_D = 3$ and $p_S = 2$.

References

1. K. BURRAGE AND P. M. BURRAGE, *High strong order explicit Runge–Kutta methods for stochastic ordinary differential equations*, Appl. Numer. Math., Vol. 22, 81-101, 1996.
2. K. BURRAGE AND P. M. BURRAGE, *General order conditions for stochastic Runge-Kutta methods for both commuting and non-commuting stochastic ordinary differential equation systems*, Appl. Numer. Math., Vol. 28, 161-177, 1998.
3. K. BURRAGE AND P. M. BURRAGE, *Order conditions of stochastic Runge–Kutta methods by B-series*, SIAM J. Numer. Anal., Vol. 38, No. 5, 1626-1646, 2000.
4. P. M. BURRAGE, *Runge-Kutta Methods for Stochastic Differential Equations*, Ph.D. thesis, Dept. of Math., University of Qeensland, Australia, 1999.
5. J. M. C. CLARK AND R. J. CAMERON, *The maximum rate of convergence of discrete approximations for stochastic differential equations*, Lecture Notes in Control and Information Sciences, Springer Berlin, Vol. 25, 162–171, 1980.
6. K. DEBRABANT AND A. KVÆRNØ, *B-series analysis of stochastic Runge-Kutta methods that use an iterative scheme to compute their internal stage values*, SIAM J. Numer. Anal., Vol. 47, No. 1, 181-203, (2008/09).
7. K. DEBRABANT AND A. RÖSSLER, *Diagonally Drift-Implicit Runge-Kutta Methods of Weak Order One and Two for It SDEs and Stability Analysis*, Appl. Numer. Math., Vol. 59, No. 3-4, 595-607 (2009).
8. K. DEBRABANT AND A. RÖSSLER, *Classification of stochastic Runge–Kutta methods for the weak approximation of stochastic differential equations*, Math. Comput. Simulation, Vol. 77, No. 4, 408–420, 2008.
9. K. DEBRABANT AND A. RÖSSLER, *Families of efficient second order Runge–Kutta methods for the weak approximation of Itô stochastic differential equations*, Appl. Numer. Math., Vol. 59, No. 3-4, 582-594, 2009.
10. K. DEBRABANT AND A. RÖSSLER, *Continuous weak approximation for stochastic differential equations*, J. Comput. Appl. Math., Vol. 214, No. 1, 259-273 (2008).
11. J. G. GAINES AND T. J. LYONS, *Random generation of stochastic area integrals*, SIAM J. Appl. Math., Vol. 54, No. 4, 1132-1146, 1994.
12. M. B. GILES, *Improved multilevel Monte Carlo convergence using the Milstein scheme*, Monte Carlo and quasi-Monte Carlo methods 2006, 343-358, Springer, Berlin, 2008.
13. H. GILSING AND T. SHARDLOW, *SDELab: A package for solving stochastic differential equations in MATLAB*, J. Comput. Appl. Math., Vol. 205, No. 2, 1002-1018, 2007.
14. E. HAIRER, S. P. NØRSETT AND G. WANNER, *Solving Ordinary Differential Equations I*, Springer-Verlag, Berlin, 1993.
15. J. KANEKO, *Explicit order 1.5 Runge-Kutta scheme for solutions of Itô stochastic differential equations*, Various problems in stochastic numerical analysis II, (Kyoto), Suri-Kaiseki-Kenkyusho-Kokyuroku, No. 932, 46-60, 1995.
16. P. E. KLOEDEN AND E. PLATEN, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin, 1999.
17. P. E. KLOEDEN, E. PLATEN AND N. HOFMANN, *Extrapolation methods for the weak approximation of Ito diffusions*, SIAM J. Numer. Anal., Vol. 32, No. 5, 1519-1534, 1995.
18. P. E. KLOEDEN, E. PLATEN AND I. W. WRIGHT, *The approximation of multiple stochastic integrals*, Stochastic Anal. Appl., Vol. 10, No. 4, 431-441, 1992.
19. Y. KOMORI, *Weak second–order stochastic Runge–Kutta methods for non–commutative stochastic differential equations*, J. Comput. Appl. Math., Vol. 206, No. 1, 158–173, 2007.

20. Y. KOMORI, T. MITSUI AND H. SUGIURA, *Rooted tree analysis of the order conditions of ROW-type scheme for stochastic differential equations*, BIT, Vol. 37, No. 1, 43–66, 1997.
21. D. KÜPPER, J. LEHN AND A. RÖSSLER, *A step size control algorithm for the weak approximation of stochastic differential equations*, Numer. Algorithms, Vol. 44, No. 4, 335–346, 2007.
22. J. LEHN, A. RÖSSLER AND O. SCHEIN, *Adaptive schemes for the numerical solution of SDEs – a comparison*, J. Comput. Appl. Math., Vol. 138, (2), 297–308, 2002.
23. G. N. MILSTEIN, *Numerical Integration of Stochastic Differential Equations*, Kluwer Academic Publishers, Dordrecht, 1995.
24. G. N. MILSTEIN AND M. V. TRETYAKOV, *Stochastic Numerics for Mathematical Physics*, Springer-Verlag, Berlin, 2004.
25. K.-S. MOON, A. SZEPESSY, R. TEMPONE AND G. E. ZOURARIS, *Convergence rates for adaptive weak approximation of stochastic differential equations*, Stochastic Anal. Appl., Vol. 23, 511–558, 2005.
26. N. J. NEWTON, *Asymptotically efficient Runge–Kutta methods for a class of Itô and Stratonovich equations*, SIAM J. Appl. Math., Vol. 51, No. 2, 542–567, 1991.
27. A. RÖSSLER, *Runge–Kutta methods for Stratonovich stochastic differential equation systems with commutative noise*, J. Comput. Appl. Math., Vol. 164–165, 613–627, 2004.
28. A. RÖSSLER, *Stochastic Taylor Expansions for the Expectation of Functionals of Diffusion Processes*, Stochastic Anal. Appl., Vol. 22, No. 6, 1553-1576, 2004.
29. A. RÖSSLER, *Runge-Kutta methods for Itô stochastic differential equations with scalar noise*, BIT, Vol. 46, No. 1, 97–110, 2006.
30. A. RÖSSLER, *Rooted tree analysis for order conditions of stochastic Runge–Kutta methods for the weak approximation of stochastic differential equations*, Stochastic Anal. Appl., Vol. 24, No. 1, 97–134, 2006.
31. A. RÖSSLER, *Second Order Runge–Kutta Methods for Stratonovich Stochastic Differential Equations*, BIT, Vol. 47, No. 3, 657–680, 2007.
32. A. RÖSSLER, *Runge–Kutta Methods for the Strong Approximation of Solutions of Stochastic Differential Equations*, Preprint, 2009 (submitted).
33. A. RÖSSLER, *Stochastic Taylor Expansions for Functionals of Diffusion Processes*, To appear in: Stochastic Analysis and Applications (2010).
34. A. RÖSSLER, *Second Order Runge–Kutta Methods for Itô Stochastic Differential Equations*, SIAM J. Numer. Anal., Vol. 47, No. 3, 1713–1738, 2009.
35. W. RÜMELIN, *Numerical treatment of stochastic differential equations*, SIAM J. Numer. Anal., Vol. 19, No. 3, 604–613, 1982.
36. D. TALAY, *Second-order discretization schemes of stochastic differential systems for the computation of the invariant law*, Stoch. and Stoch. Rep., Vol. 29, No. 1, 13–36, 1990.
37. D. TALAY AND L. TUBARO, *Expansion of the global error for numerical schemes solving stochastic differential equations*, Stochastic Anal. Appl., Vol. 8, No. 4, 483-509, 1990.
38. A. TOCINO AND J. VIGO-AGUIAR, *Weak second order conditions for stochastic Runge–Kutta methods*, SIAM J. Sci. Comput., Vol. 24, No. 2, 507–523, 2002.
39. M. WIKTORSSON, *Joint characteristic function and simultaneous simulation of iterated Itô integrals for multiple independent Brownian motions*, Ann. Appl. Probab., Vol. 11, No. 2, 470-487, 2001.