SIAM J. NUMER. ANAL.
Vol. 48, No. 3, pp. 922–952

RUNGE–KUTTA METHODS FOR THE STRONG APPROXIMATION OF SOLUTIONS OF STOCHASTIC DIFFERENTIAL EQUATIONS*

ANDREAS RÖßLER†

This paper is dedicated to Edith Rößler

Abstract. Some new stochastic Runge–Kutta (SRK) methods for the strong approximation of solutions of stochastic differential equations (SDEs) with improved efficiency are introduced. Their convergence is proved by applying multicolored rooted tree analysis. Order conditions for the coefficients of explicit and implicit SRK methods are calculated. As the main novelty, order 1.0 strong SRK methods with significantly reduced computational complexity for Itô as well as for Stratonovich SDEs with a multidimensional driving Wiener process are presented where the number of stages is independent of the dimension of the Wiener process. Further, an order 1.0 strong SRK method customized for Itô SDEs with commutative noise is introduced. Finally, some order 1.5 strong SRK methods for SDEs with scalar, diagonal, and additive noise are proposed. All introduced SRK methods feature significantly reduced computational complexity compared to well-known schemes.

Key words. stochastic Runge–Kutta method, stochastic differential equation, multicolored rooted tree analysis, strong approximation, numerical method, commutative noise, diagonal noise, additive noise

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

DOI. 10.1137/09076636X

1. Introduction. In recent years, numerical methods for the strong and weak approximation of solutions of stochastic differential equations (SDEs) have been developed; see, e.g., [8, 12, 20] and the references therein. Especially, derivative free Runge–Kutta-type schemes have been proposed for weak approximations [12, 14, 15, 16, 17, 18, 24, 26, 28, 29] as well as for strong approximations [3, 6, 11, 12, 21, 27]. Up to now, the Euler–Maruyama scheme is often the method of choice for strong approximations because there has been no efficient way to simulate multiple stochastic integrals needed by higher order methods which causes an order reduction to strong order 0.5 in the case of noncommutative SDEs [5]. However, the simulation method recently proposed by Wiktorsson [30] provides a very efficient way of simulating multiple stochastic integrals (see also [7, 13]) needed for higher order methods.

The aim of the present paper is to introduce new stochastic Runge–Kutta (SRK) methods for the strong pathwise approximation with considerably reduced computational complexity and higher efficiency compared to well-known methods such as the Euler–Maruyama scheme. Burrage and Burrage [1, 3] introduced colored trees for the calculation of order conditions for SRK methods for the strong approximation of Stratonovich SDEs. This colored tree approach is extended to Itô SDEs in [25]. Therefore, in sections 2–4 the multicolored rooted tree analysis for both Itô and Stratonovich SDEs with a multidimensional driving Wiener process is applied to a very general class of SRK methods [23]. A general convergence result giving order conditions for the class of SRK methods is presented in section 5. This result is applied in section 6.1, where strong order 1.0 SRK methods for possibly noncommutative

*Received by the editors July 27, 2009; accepted for publication (in revised form) April 5, 2010; published electronically June 29, 2010.
http://www.siam.org/journals/sinum/48-3/76636.html
†Fachbereich Mathematik, Technische Universität Darmstadt, Schlossgartenstr. 7, D-64289 Darmstadt, Germany (roessler@mathematik.tu-darmstadt.de).

Itô and Stratonovich SDEs with an m-dimensional driving Wiener process are introduced. As the main innovation, the number of stages in these new SRK methods does not depend on the dimension m of the Wiener process. Thus, especially for high dimensional problems, the complexity of these methods is considerably reduced compared to well-known approximation methods such as the Milstein scheme [19, 20] or the explicit order 1.0 strong SRK scheme due to Kloeden and Platen [12] where the number of stages depends linearly on m. A similar reduction of complexity has been achieved for weak order two SRK methods proposed in [26, 24] which are more efficient than well-known weak order two SRK schemes such as proposed in, e.g., [12, 14, 15, 16, 29]. As yet, strong order 1.0 derivative-free approximation methods taking advantage of commutative noise are only known for Stratonovich SDEs [2, 12]. Therefore, in section 6.2 we introduce strong order 1.0 SRK methods especially designed for Itô SDEs with commutative noise. In the case of a one-dimensional driving Wiener process, Newton [21], Kloeden and Platen [12], and Rümelin [27] introduced order 1.0 strong SRK methods, whereas Burrage and Burrage [1, 3, 4], Kaneko [11], and Kloeden and Platen [12] proposed order 1.5 strong SRK methods for Stratonovich and Itô SDEs, respectively. In section 6.3 we introduce order 1.5 strong SRK methods for Itô SDEs with scalar noise, in section 6.4 for diagonal noise, and in section 6.5 for Itô and Stratonovich SDE systems with additive noise which possess lower computational complexity compared to the well-known methods. Finally, all newly introduced SRK methods possess considerably reduced computational complexity compared to the well-known approximation methods, which is confirmed by numerical results in section 7.

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

$$X_t = X_{t_0} + \int_{t_0}^{t} a(s, X_s)\,\mathrm{d}s + \sum_{j=1}^{m} \int_{t_0}^{t} b^j(s, X_s) * \mathrm{d}W_s^j \tag{1.1}$$

with an m-dimensional driving Wiener process $(W_t)_{t\geq 0} = ((W_t^1, \dots, W_t^m)^T)_{t\geq 0}$ w.r.t. $(\mathcal{F}_t)_{t\geq 0}$ for $d, m \geq 1$ and $t \in \mathcal{I}$. 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 $a : \mathcal{I} \times \mathbb{R}^d \to \mathbb{R}^d$ and $b : \mathcal{I} \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ 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 $\mathcal{F}_{t_0}$-measurable initial value. In the following, we suppose that the conditions of the existence and uniqueness theorem [12] are fulfilled for SDE (1.1), and we denote by $\|\cdot\|$ the Euclidean norm. Let a discretization $\mathcal{I}_h = \{t_0, t_1, \dots, t_N\}$ with $t_0 < t_1 < \dots < t_N = T$ of the time interval $\mathcal{I} = [t_0, T]$ with step sizes $h_n = t_{n+1} - t_n$ for $n = 0, 1, \dots, N-1$ be given. Further, let $h = \max_{0\leq n<N} h_n$ denote the maximum step size.

DEFINITION 1.1. *A sequence of time discrete approximations Y^h converges in the mean-square sense with order $p > 0$ to the solution X of SDE* (1.1) *at time T if there exists a constant $C > 0$, which is independent of h, and some $\delta_0 > 0$ such that for each $h \in]0, \delta_0]$*

$$(\mathrm{E}(\|X_T - Y^h(T)\|^2))^{1/2} \leq C\, h^p\,. \tag{1.2}$$

2. A general class of SRK methods. We consider a universal class of SRK methods which has been introduced in [23]: Let $\mathcal{M}$ be an arbitrary finite set of multi-indices with $\kappa = |\mathcal{M}|$ elements, let $\theta_\iota^{(k)}(h) \in L^2(\Omega)$, $\iota \in \mathcal{M}$, $0 \leq k \leq m$, be some suitable random variables, and let $b^0(t,x) := a(t,x)$. For the strong approximation of the solution of SDE (1.1) an s-stage SRK method is given by $Y_0 = X_{t_0}$ and

$$Y_{n+1} = Y_n + \sum_{i=1}^{s} \sum_{k=0}^{m} \sum_{\nu \in \mathcal{M}} z_i^{(k),(\nu)} \; b^k \left(t_n + c_i^{(\nu)} h_n, H_i^{(\nu)} \right) \tag{2.1}$$

for $n = 0, 1, \dots, N-1$ with $Y_n = Y(t_n)$, $t_n \in \mathcal{I}_h$, 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 $i = 1, \dots, s$ and $\nu \in \mathcal{M}$. Here, let $0 \in \mathcal{M}$, and for $i, j = 1, \dots, s$ let

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

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 with 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

$$c^{(\nu)} = \sum_{\mu \in \mathcal{M}} {C_{ij}^{(0)}}^{(\nu),(0),(\mu)} e \tag{2.2}$$

with $e = (1, \dots, 1)^T \in \mathbb{R}^s$. If ${C_{ij}^{(\iota)}}^{(\nu),(l),(\mu)} = 0$ for $j \ge i$, then (2.1) 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

$$\mathrm{E}\left(\prod_{k=0}^{m} ((\theta_{\iota_1}^{(k)}(h))^{p_1^k} \cdot \ldots \cdot (\theta_{\iota_\kappa}^{(k)}(h))^{p_\kappa^k}) \right) = O\big(h^{p_1^0 + \cdots + p_\kappa^0 + \sum_{k=1}^m (p_1^k + \cdots + p_\kappa^k)/2}\big) \tag{2.3}$$

for all $p_i^k \in \mathbb{N}_0$, $k = 0, 1, \dots, m$, 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 [23, 20]). Remark that for a deterministic ODE, i.e., SDE (1.1) with $b \equiv 0$, the SRK method covers the well-known deterministic Runge–Kutta method [10], so the introduced class of SRK methods is a generalization of deterministic Runge–Kutta methods.

3. Colored rooted tree analysis.

We extend the colored rooted tree theory proposed for Stratonovich SDEs by Burrage and Burrage [1, 3] such that it can be also applied to Itô SDEs with a multidimensional driving Wiener process. We have to point out that in contrast to the analysis of weak approximation methods, here the required colored trees have a different structure [22]. Further, the analysis of order conditions for the strong convergence of an SRK method involves multiple stochastic integrals which are not explicitly needed in the case of the analysis of weak order conditions [23, 26]. Finally, the error estimates for the truncation terms are more

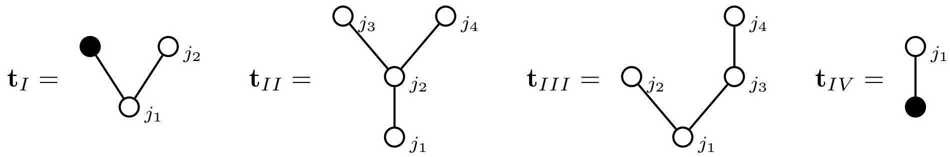

FIG. 3.1. *Four elements of TS with some $j_1, j_2, j_3, j_4 \in \{1, \dots, m\}$.*

sophisticated [25]. Therefore, we briefly provide some essential definitions and results which are necessary for the presentation of the main results in section 6. Without loss of generality, we restrict our considerations to autonomous SDE systems in this section.

Let TS denote the set of all S-trees (stochastic trees) which can be composed of deterministic nodes $\tau_0 = \bullet$ and stochastic nodes $\tau_j = \bigcirc_j$ with some $j \in \{1, \dots, m\}$. The index j is associated with the jth component of the m-dimensional driving Wiener process of the considered SDE. Further, let $\gamma \in TS$ denote the empty tree, i.e., the tree without any nodes. Figure 3.1 contains some trees of TS as an example. For $\mathbf{t} \in TS$, 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 of the tree $\mathbf{t}$. The order $\rho(\mathbf{t})$ of the tree $\mathbf{t} \in TS$ is defined as $\rho(\gamma) = 0$ and $\rho(\mathbf{t}) = d(\mathbf{t}) + \frac{1}{2} s(\mathbf{t})$. Thus, the order of the trees given in Figure 3.1 can be calculated as $\rho(\mathbf{t}_I) = \rho(\mathbf{t}_{II}) = \rho(\mathbf{t}_{III}) = 2$ and $\rho(\mathbf{t}_{IV}) = 1.5$. Every tree can be written by a combination of brackets: If $\mathbf{t}_1, \dots, \mathbf{t}_k \in TS$ are colored trees, 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 $j \in \{0, 1, \dots, m\}$. Therefore, proceeding recursively, for the trees in Figure 3.1 we obtain $\mathbf{t}_I = [\tau_0, \tau_{j_2}]_{j_1}$, $\mathbf{t}_{II} = [[\tau_{j_3}, \tau_{j_4}]_{j_2}]_{j_1}$, $\mathbf{t}_{III} = [\tau_{j_2}, [\tau_{j_4}]_{j_3}]_{j_1}$, and $\mathbf{t}_{IV} = [\tau_{j_1}]_0$.

For every tree $\mathbf{t} \in TS$, we define the corresponding *elementary differential* recursively by $F(\gamma)(x) = x$, $F(\tau_0)(x) = a(x)$, and $F(\tau_j)(x) = b^j(x)$ for single nodes and for a tree $\mathbf{t}$ with more than one node by

$$F(\mathbf{t})(x) = \begin{cases} a^{(k)}(x) \cdot (F(\mathbf{t}_1)(x), \dots, F(\mathbf{t}_k)(x)) & \text{for } \mathbf{t} = [\mathbf{t}_1, \dots, \mathbf{t}_k]_0, \\ {b^j}^{(k)}(x) \cdot (F(\mathbf{t}_1)(x), \dots, F(\mathbf{t}_k)(x)) & \text{for } \mathbf{t} = [\mathbf{t}_1, \dots, \mathbf{t}_k]_j \end{cases} \tag{3.1}$$

for $j \in \{1, \dots, m\}$. Here $a^{(k)}$ and ${b^j}^{(k)}$ define a symmetric k-linear differential operator, and one can choose the sequence of the trees $\mathbf{t}_1, \dots, \mathbf{t}_k$ in an arbitrary order. For example, the Ith component of ${b^j}^{(k)} \cdot (F(\mathbf{t}_1), \dots, F(\mathbf{t}_k))$ can be written as

$$({b^j}^{(k)} \cdot (F(\mathbf{t}_1), \dots, F(\mathbf{t}_k)))^I = \sum_{J_1, \dots, J_k = 1}^{d} \frac{\partial^k b^{I,j}}{\partial x^{J_1} \cdots \partial x^{J_k}} (F^{J_1}(\mathbf{t}_1), \dots, F^{J_k}(\mathbf{t}_k)),$$

where the components of vectors are denoted by superscript indices, which are chosen as capitals. As an example, we calculate for the Ith component of the trees $\mathbf{t}_I$ and $\mathbf{t}_{II}$ in Figure 3.1 the elementary differentials

$$F(\mathbf{t}_I)^I = ({b^{j_1}}''(a, b^{j_2}))^I = \sum_{J_1, J_2 = 1}^{d} \frac{\partial^2 b^{I,j_1}}{\partial x^{J_1} \partial x^{J_2}} \, a^{J_1} \, b^{J_2, j_2} \,,$$

$$F(\mathbf{t}_{II})^I = ({b^{j_1}}'({b^{j_2}}''(b^{j_3}, b^{j_4})))^I = \sum_{J_1 = 1}^{d} \frac{\partial b^{I,j_1}}{\partial x^{J_1}} \sum_{K_1, K_2 = 1}^{d} \frac{\partial^2 b^{J_1, j_2}}{\partial x^{K_1} \partial x^{K_2}} \, b^{K_1, j_3} \, b^{K_2, j_4}.$$

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 $\mathbf{t} \in TS$ with $\mathbf{t} = [\mathbf{t}_1, \ldots, \mathbf{t}_k]_j$ for some $j \in \{0, 1, \ldots, m\}$ the corresponding *multiple stochastic integral* recursively by

$$I_{\mathbf{t};t_0,t}[Z.] = \left(\int_{t_0}^{t} \prod_{i=1}^{k} I_{\mathbf{t}_i;t_0,s} * \mathrm{d}W_s^j \right) [Z.] \tag{3.2}$$

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

$$\begin{aligned} &\left(\int_{t_0}^{t} \int_{t_0}^{s_n} \cdots \int_{t_0}^{s_2} *\mathrm{d}W_{s_1}^{j_1} * \mathrm{d}W_{s_2}^{j_2} \cdots * \mathrm{d}W_{s_n}^{j_n} \right) [Z.] = \mathcal{I}_{(j_1,j_2,\ldots,j_n)}[Z.]_{t_0,t} \\ &= \int_{t_0}^{t} \int_{t_0}^{s_n} \cdots \int_{t_0}^{s_2} Z_{s_1} * \mathrm{d}W_{s_1}^{j_1} * \mathrm{d}W_{s_2}^{j_2} \cdots * \mathrm{d}W_{s_n}^{j_n} \end{aligned} \tag{3.3}$$

in (3.2). Here, the product of two stochastic integrals can be written as a sum:

$$\begin{aligned} &\int_{t_0}^{t} X_s * \mathrm{d}W_s^i \int_{t_0}^{t} Y_s * \mathrm{d}W_s^j = \int_{t_0}^{t} X_s \, Y_s \, 1_{\{i=j\neq 0 \wedge * \neq \circ\}} \, \mathrm{d}s \\ &+ \int_{t_0}^{t} X_s \left(\int_{t_0}^{s} Y_u * \mathrm{d}W_u^j \right) * \mathrm{d}W_s^i + \int_{t_0}^{t} \left(\int_{t_0}^{s} X_u * \mathrm{d}W_u^i \right) Y_s * \mathrm{d}W_s^j \end{aligned} \tag{3.4}$$

for $0 \leq i, j \leq m$ [12], where the first summand on the right-hand side appears only in the case of Itô calculus. For example, we calculate for $\mathbf{t}_I$ and $\mathbf{t}_{II}$

$$\begin{aligned} I_{\mathbf{t}_I;t_0,t}[1] &= \int_{t_0}^{t} I_{\tau_0;t_0,s} \, I_{\tau_{j_2};t_0,s} * \mathrm{d}W_s^{j_1}[1] = \mathcal{I}_{(0,j_2,j_1)}[1]_{t_0,t} + \mathcal{I}_{(j_2,0,j_1)}[1]_{t_0,t} \,, \\ I_{\mathbf{t}_{II};t_0,t}[1] &= \int_{t_0}^{t} \int_{t_0}^{s} I_{\tau_{j_3};t_0,u} \, I_{\tau_{j_4};t_0,u} * \mathrm{d}W_u^{j_2} * \mathrm{d}W_s^{j_1} \\ &= \mathcal{I}_{(j_3,j_4,j_2,j_1)}[1]_{t_0,t} + \mathcal{I}_{(j_4,j_3,j_2,j_1)}[1]_{t_0,t} + \mathcal{I}_{(0,j_2,j_1)}[1_{\{j_3=j_4\neq 0 \wedge * \neq \circ\}}]_{t_0,t}, \end{aligned}$$

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

Let $\mathbf{t} \in TS$ with $\mathbf{t} = [\mathbf{t}_1, \ldots, \mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_2, \ldots, \mathbf{t}_k, \ldots, \mathbf{t}_k]_j = [\mathbf{t}_1^{n_1}, \mathbf{t}_2^{n_2}, \ldots, \mathbf{t}_k^{n_k}]_j$, $j \in \{0, 1, \ldots, m\}$, where $\mathbf{t}_1, \ldots, \mathbf{t}_k$ are distinct subtrees with multiplicities $n_1, \ldots, n_k$, respectively. Then we recursively define the *symmetry factor* by

$$\sigma(\mathbf{t}) = \prod_{i=1}^{k} n_i! \; \sigma(\mathbf{t}_i)^{n_i} \tag{3.5}$$

with $\sigma(\tau_j) = \sigma(\gamma) = 1$. For the trees presented in Figure 3.1, we obviously obtain $\sigma(\mathbf{t}_I) = \sigma(\mathbf{t}_{III}) = \sigma(\mathbf{t}_{IV}) = 1$. 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 calculate $\sigma(\mathbf{t}_{II}) = 2$. As an example, all trees up to order 1.5 and the corresponding multiple integrals are presented in Table 3.1.

Then, based on the introduced multicolored rooted trees, the following stochastic Taylor expansion for the solution of SDE (1.1) holds [25].

TABLE 3.1
All trees $\boldsymbol{t} \in TS$ of order $\rho(\boldsymbol{t}) \leq 1.5$ with $j_1, j_2, j_3 \in \{1, \ldots, m\}$ arbitrarily eligible.

$\mathbf{t}$	tree	$I_{\mathbf{t};t_0,t}$	$\sigma(\mathbf{t})$	$\rho(\mathbf{t})$
$\mathbf{t}_{0.1}$	γ	1	1	0
$\mathbf{t}_{0.5.1}$	τ_{j_1}	$\mathcal{I}_{(j_1)}[1]_{t_0,t}$	1	0.5
$\mathbf{t}_{1.1}$	τ_0	$\mathcal{I}_{(0)}[1]_{t_0,t}$	1	1
$\mathbf{t}_{1.3}$	$[\tau_{j_2}]_{j_1}$	$\mathcal{I}_{(j_2,j_1)}[1]_{t_0,t}$	1	1
$\mathbf{t}_{1.5.1}$	$[\tau_{j_1}]_0$	$\mathcal{I}_{(j_1,0)}[1]_{t_0,t}$	1	1.5
$\mathbf{t}_{1.5.2}$	$[\tau_0]_{j_1}$	$\mathcal{I}_{(0,j_1)}[1]_{t_0,t}$	1	1.5
$\mathbf{t}_{1.5.6}$	$[\tau_{j_2}, \tau_{j_3}]_{j_1}$	$\mathcal{I}_{(j_3,j_2,j_1)}[1]_{t_0,t} + \mathcal{I}_{(j_2,j_3,j_1)}[1]_{t_0,t}$ $+\mathcal{I}_{(0,j_1)}[1_{\{j_2=j_3\neq 0 \wedge * \neq \circ\}}]_{t_0,t}$	$1 + 1_{\{j_2=j_3\}}$	1.5
$\mathbf{t}_{1.5.7}$	$[[\tau_{j_3}]_{j_2}]_{j_1}$	$\mathcal{I}_{(j_3,j_2,j_1)}[1]_{t_0,t}$	1	1.5

THEOREM 3.1. *For the solution $(X_t)_{t \in \mathcal{I}}$ of SDE* (1.1) *and for $p \in \frac{1}{2}\mathbb{N}_0$ with $a, b^j \in C^{2p+1}(\mathbb{R}^d, \mathbb{R}^d)$ for $j = 1, \ldots, m$, we obtain the expansion*

$$X_t = \sum_{\substack{t \in TS \\ \rho(\boldsymbol{t}) \leq p}} F(\boldsymbol{t})(X_{t_0}) \frac{I_{\boldsymbol{t};t_0,t}}{\sigma(\boldsymbol{t})} + \mathcal{R}_p(t, t_0) \tag{3.6}$$

P-*a.s. with a remainder term $\mathcal{R}_p(t, t_0)$, provided all of the appearing multiple stochastic integrals exist. If for all $\boldsymbol{t} \in TS$ with $\rho(\boldsymbol{t}) = p + \frac{1}{2}$ or $\rho(\boldsymbol{t}) = p+1$ there exists some constant $C > 0$ such that for all $x \in \mathbb{R}^d$*

$$\|F(\boldsymbol{t})(x)\| \leq C(1 + \|x\|^2)^{1/2}, \tag{3.7}$$

then $\mathrm{E}(\|\mathcal{R}_p(t, t_0)\|^2) \leq C(1 + \mathrm{E}(\|X_{t_0}\|^2))(t - t_0)^{2p+1}$ *and* $\|\mathrm{E}(\mathcal{R}_p(t, t_0))\| \leq C(1 + \mathrm{E}(\|X_{t_0}\|^2))^{1/2}(t - t_0)^{p+\kappa}$, *where $\kappa = 1$ if $p \in \mathbb{N}_0$ and $\kappa = 1/2$ if $p \notin \mathbb{N}_0$.*

4. Rooted tree expansion for the SRK method. Next, we give an expansion for the approximation process $(Y(t))_{t \in \mathcal{I}_h}$ defined by the SRK method (2.1). Since the expansion for $(Y(t))_{t \in \mathcal{I}_h}$ contains the coefficients of the SRK method, we define a coefficient function Φ_S which assigns to every tree an *elementary weight*. For every $\mathbf{t} \in TS$ with $\mathbf{t} = [\mathbf{t}_1, \ldots, \mathbf{t}_u]_k$ for some $k \in \{0, 1, \ldots, m\}$ the function Φ_S is defined by $\Phi_S(\gamma) = 1$ and

$$\Phi_S(\mathbf{t}) = \sum_{\nu \in \mathcal{M}} z^{(k),(\nu)^T} \prod_{i=1}^{u} \Psi^{(\nu)}(\mathbf{t}_i), \tag{4.1}$$

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

$$\Psi^{(\nu)}(\mathbf{t}) = \sum_{\mu \in \mathcal{M}} Z^{(\nu),(l),(\mu)} \prod_{i=1}^{v} \Psi^{(\mu)}(\mathbf{t}_i). \tag{4.2}$$

Here $e = (1, \ldots, 1)^T$ and the product of vectors in the definition of $\Psi^{(\nu)}$ is defined by componentwise multiplication, i.e., with $(a_1, \ldots, a_n) \cdot (b_1, \ldots, b_n) = (a_1 b_1, \ldots, a_n b_n)$. In the following, we also write $\Phi_S(\mathbf{t}; t_0, t) = \Phi_S(\mathbf{t})$ in order to emphasize the dependency on the current time step with step size $h = t - t_0$.

Now, the local expansion for the approximation process $Y^h = (Y(t))_{t\in\mathcal{I}_h}$ by the SRK method (2.1) follows from [23, Corollary 5.6, Proposition 6.1].

THEOREM 4.1. *For the approximation process* $(Y(t))_{t\in\mathcal{I}_h}$ *defined by the SRK method* (2.1) *and for* $p \in \frac{1}{2}\mathbb{N}_0$ *we obtain the expansion*

$$Y(t) = \sum_{\substack{\boldsymbol{t}\in TS\\ \rho(\boldsymbol{t})\le p}} F(\boldsymbol{t})(Y(t_0))\frac{\Phi_S(\boldsymbol{t};t_0,t)}{\sigma(\boldsymbol{t})} + \mathcal{R}_p^{\Delta}(t,t_0) \tag{4.3}$$

P-*a.s. with a remainder term* $\mathcal{R}_p^{\Delta}(t,t_0)$, *provided that* $a, b^j \in C^{2p}(\mathbb{R}^d,\mathbb{R}^d)$ *for* $1 \le j \le m$. *Suppose that* (2.3) *holds and that for all* $\boldsymbol{t} \in TS$ *with* $\rho(\boldsymbol{t}) = p + \frac{1}{2}$ *there exists some constant* $C > 0$ *such that for all* $x \in \mathbb{R}^d$

$$\|F(\boldsymbol{t})(x)\| \le C(1+\|x\|^2)^{1/2}. \tag{4.4}$$

Then, $\mathrm{E}(\|\mathcal{R}_p^{\Delta}(t,t_0)\|^2) \le C(1+\mathrm{E}(\|Y(t_0)\|^2))(t-t_0)^{2p+1}$ *and* $\|\mathrm{E}(\mathcal{R}_p^{\Delta}(t,t_0))\| \le C(1+\mathrm{E}(\|Y(t_0)\|^2))^{1/2}(t-t_0)^{p+1/2}$.

5. Order conditions for SRK methods. Now, we calculate order conditions for the random variables and the coefficients of the SRK method (2.1) if it is applied to the nonautonomous SDE (1.1). Let $X^{t,x}$ denote a process X with initial condition $X_t = X(t) = x$ P-a.s. Then we consider the local mean and mean-square errors of the SRK method (2.1),

$$le^m(h;t,x) = \|\mathrm{E}(X^{t,x}(t+h) - Y^{t,x}(t+h))\|, \tag{5.1}$$

$$le^{ms}(h;t,x) = (\mathrm{E}(\|X^{t,x}(t+h) - Y^{t,x}(t+h)\|^2))^{1/2}, \tag{5.2}$$

and we apply a result due to Milstein [19] for one-step approximations.

THEOREM 5.1. *Let* $p \in \frac{1}{2}\mathbb{N}_0$ *and suppose that for arbitrary* $t \in [t_0, T-h]$ *and* $x \in \mathbb{R}^d$ *there exists a constant* $C > 0$ *such that for the approximations* $(Y(t))_{t\in\mathcal{I}_h}$

$$\begin{aligned} le^{ms}(h;t,x) &\le C(1+\|x\|^2)^{1/2}\, h^{p+1/2}\,, \\ le^{m}(h;t,x) &\le C(1+\|x\|^2)^{1/2}\, h^{p+1}\,. \end{aligned} \tag{5.3}$$

Then the one-step approximations $(Y(t))_{t\in\mathcal{I}_h}$ *converge with order* p *in the mean-square sense to the solution* X *of SDE* (1.1) *as* $h \to 0$, *i.e., for any* $t \in \mathcal{I}_h$

$$(\mathrm{E}(\|X^{t_0,X_{t_0}}(t) - Y^{t_0,X_{t_0}}(t)\|^2))^{1/2} \le C(1+\mathrm{E}(\|X(t_0)\|^2))^{1/2}\, h^p\,. \tag{5.4}$$

Suppose that the assumptions of Theorems 3.1 and 4.1 are fulfilled. Then we obtain conditions for the coefficients and the random variables of the class of SRK methods (2.1) for strong convergence with some order p.

PROPOSITION 5.2. *Let* $p \in \frac{1}{2}\mathbb{N}_0$ *and* $a, b^j \in C^{\lceil p\rceil,2p+1}(\mathcal{I}\times\mathbb{R}^d,\mathbb{R}^d)$ *for* $j = 1,\dots,m$. *Then the SRK method* (2.1) *has mean-square order of accuracy* p *if the following conditions are fulfilled for arbitrary* $t, t+h \in \mathcal{I}$ *and if* (2.2) *and* (2.3) *hold:*

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

$$I_{\boldsymbol{t};t,t+h} = \Phi_S(\boldsymbol{t};t,t+h) \qquad \text{P-}a.s.; \tag{5.5}$$

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

$$\mathrm{E}(I_{\boldsymbol{t};t,t+h}) = \mathrm{E}(\Phi_S(\boldsymbol{t};t,t+h)). \tag{5.6}$$

Proof. Let $p \in \frac{1}{2}\mathbb{N}_0$ and suppose that (5.5) and (5.6) are fulfilled for p. Then we obtain for $x \in \mathbb{R}^d$ with Theorems 3.1 and 4.1

$$
\begin{aligned}
le^{ms}(h;t,x) &= (\mathrm{E}^{t,x}(\|X_{t+h} - Y(t+h)\|^2))^{1/2} \\
&\leq 2\Bigg(\mathrm{E}\Bigg(\Bigg\| \sum_{\substack{\mathbf{t}\in TS \\ \rho(\mathbf{t})\leq p}} F(\mathbf{t})(x)\,\frac{1}{\sigma(\mathbf{t})}(I_{\mathbf{t};t,t+h} - \Phi_S(\mathbf{t};t,t+h))\Bigg\|^2\Bigg) \\
&\qquad + \mathrm{E}^{t,x}(\|\mathcal{R}_p(t,t+h)\|^2) + \mathrm{E}^{t,x}(\|\mathcal{R}_p^{\Delta}(t,t+h)\|^2)\Bigg)^{1/2} \\
&\leq C(1+\|x\|^2)^{1/2} h^{p+1/2}
\end{aligned}
\tag{5.7}
$$

for the local mean-square error, and we get for the local mean error

$$
\begin{aligned}
le^{m}(h;t,x) &= \|\,\mathrm{E}^{t,x}(X_{t+h}) - \mathrm{E}^{t,x}(Y(t+h))\| \\
&\leq \Bigg\| \sum_{\substack{\mathbf{t}\in TS \\ \rho(\mathbf{t})\leq p+1/2}} F(\mathbf{t})(x)\,\frac{1}{\sigma(\mathbf{t})}\,\mathrm{E}(I_{\mathbf{t};t,t+h} - \Phi_S(\mathbf{t};t,t+h))\Bigg\| \\
&\quad + \Bigg\| \sum_{\substack{\mathbf{t}\in TS \\ \rho(\mathbf{t})\in\{p,p+1/2\} \\ \mathrm{E}^{t,x}(I_{\mathbf{t};t,t+h}[1])\neq 0}} \frac{1}{\sigma(\mathbf{t})}\,\mathrm{E}^{t,x}\Bigg(I_{\mathbf{t};t,t+h}\Bigg[\int_t^{\cdot} \sum_{i=1}^{d} a^i(X_s)\frac{\partial F(\mathbf{t})(X_s)}{\partial x^i}\,\mathrm{d}s \\
&\qquad + \frac{1}{2}\int_t^{\cdot} \sum_{i,j=1}^{d}\sum_{k=1}^{m} b^{i,k}(X_s)\,b^{j,k}(X_s)\,\frac{\partial^2 F(\mathbf{t})(X_s)}{\partial x^i \partial x^j}\,\mathrm{d}s \\
&\qquad + \int_t^{\cdot} \sum_{i=1}^{d}\sum_{k=1}^{m} b^{i,k}(X_s)\,\frac{\partial F(\mathbf{t})(X_s)}{\partial x^i}\,\mathrm{d}W_s^k\Bigg]\Bigg)\Bigg\| + \|\,\mathrm{E}^{t,x}(\mathcal{R}_{p+1/2}^{\Delta}(t,t+h))\| \\
&\leq C(1+\|x\|^2)^{1/2} h^{p+1}
\end{aligned}
\tag{5.8}
$$

for some constant $C > 0$. Applying Theorem 5.1 completes the proof. ∎

6. SRK methods for the strong approximation. As in the deterministic setting, there exist strong Taylor schemes such as the well-known order 1.0 strong Milstein scheme [12, 19, 20] for the pathwise approximation of the solutions of SDE (1.1). However, their main disadvantage is that derivatives of the drift and diffusion coefficients have to be evaluated at each step. Therefore, derivative-free schemes have been proposed where derivatives are replaced by difference quotients. However, such approaches do not automatically result in efficient schemes where the number of evaluations of the drift and diffusion coefficients is minimized. Therefore, we propose much more efficient derivative-free SRK methods with significantly reduced computational complexity by a systematic rooted tree approach instead of the application of difference quotients.

In the following, for $0 \leq i,j \leq m$, $\mathrm{d}W_s^0 = \circ\mathrm{d}W_s^0 = \mathrm{d}s$, and $t_n, t_{n+1} \in \mathcal{I}_h$ 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, and let

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

denote the multiple Stratonovich stochastic integrals. For convenience we also write $I_{(i)} = I_{(i),n}$ if n is obvious from the context. The increments of the Wiener process $I_{(i),n} = J_{(i),n}$ are independent $N(0,h_n)$ distributed with $h_n = t_{n+1} - t_n$ for $1 \leq i \leq m$. From (3.4) it follows that $I_{(i,0)} = J_{(i,0)}$, $I_{(0,i)} = J_{(0,i)} = h_n I_{(i)} - I_{(i,0)}$, and that $I_{(i,j)} = J_{(i,j)}$ for $1 \leq i,j \leq m$ if $i \neq j$. In the case of $i = j$, formula (3.4) results in $I_{(i,i)} = \frac{1}{2}(I_{(i)}^2 - h_n)$, whereas $J_{(i,i)} = \frac{1}{2} J_{(i)}^2$ for $1 \leq i \leq m$. Further, let $I_{(i,i,i)} = \frac{1}{6}(I_{(i)}^3 - 3 I_{(0)}\, I_{(i)})$ for $i \in \{1, \ldots, m\}$. For the SRK schemes introduced in this section, 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 that are independent of $I_{(i)}$ for all $1 \leq i \leq m$ [12, 20]. Since the exact simulation of the multiple stochastic integrals $I_{(i,j)}$ for $1 \leq i,j \leq m$ with $i \neq j$ is not possible, one has to substitute them by sufficiently exact and efficient approximations as recently proposed by Wiktorsson [30]; see also [9] for implementation issues. In the following, let (p_D, p_S) with $p_D \geq p_S$ denote the order of convergence of the considered SRK scheme if it is applied to a deterministic or stochastic differential equation, respectively.

6.1. Order 1.0 strong SRK methods. First, we introduce some efficient order 1.0 strong SRK methods for Itô and Stratonovich SDEs, which may also possess noncommutative noise. To the best of the author's knowledge, as yet all 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 [12] needs one evaluation of the drift coefficient a, but $m+1$ evaluations of each diffusion coefficient b^j, $j = 1, \ldots, m$, for each step. Thus, the complexity grows quadratically in m, which is a significant drawback, especially for high dimensional problems. Therefore, we introduce some more efficient SRK methods, where the number of evaluations of each drift and each diffusion coefficient is independent of the dimension m of the driving Wiener process.

For the multidimensional Itô SDE (1.1) with $d \geq 1$ and $m \geq 1$, we propose the s-stages order 1.0 strong SRK method with $Y_0 = X_{t_0}$ and

$$\begin{aligned} Y_{n+1} = Y_n &+ \sum_{i=1}^{s} \alpha_i \, a(t_n + c_i^{(0)} h_n, H_i^{(0)})\, h_n \\ &+ \sum_{k=1}^{m} \sum_{i=1}^{s} \big(\beta_i^{(1)} I_{(k)} + \beta_i^{(2)} \sqrt{h_n} \big)\, b^k(t_n + c_i^{(1)} h_n, H_i^{(k)}) \end{aligned} \tag{6.1}$$

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

(6.2)

$$H_i^{(0)} = Y_n + \sum_{j=1}^{s} A_{ij}^{(0)} \, a(t_n + c_j^{(0)} h_n, H_j^{(0)})\, h_n + \sum_{l=1}^{m} \sum_{j=1}^{s} B_{ij}^{(0)} \, b^l(t_n + c_j^{(1)} h_n, H_j^{(l)})\, I_{(l)},$$

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

Table 6.1
Coefficients for the 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.*

$$
\begin{array}{c|ccc|ccc|ccc}
0 & & & & & & & & & \\
0 & 0 & & & 0 & & & & & \\
0 & 0 & 0 & & 0 & 0 & & & & \\
\hline
0 & & & & & & & & & \\
0 & 0 & & & 1 & & & & & \\
0 & 0 & 0 & & -1 & 0 & & & & \\
\hline
 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & \frac{1}{2} & -\frac{1}{2}
\end{array}
\qquad
\begin{array}{c|ccc|ccc|ccc}
0 & & & & & & & & & \\
1 & 1 & & & 0 & & & & & \\
0 & 0 & 0 & & 0 & 0 & & & & \\
\hline
0 & & & & & & & & & \\
1 & 1 & & & 1 & & & & & \\
1 & 1 & 0 & & -1 & 0 & & & & \\
\hline
 & \frac{1}{2} & \frac{1}{2} & 0 & 1 & 0 & 0 & 0 & \frac{1}{2} & -\frac{1}{2}
\end{array}
$$

for $i = 1, \ldots, s$ and $k = 1, \ldots, m$. The SRK method (6.1) can be characterized by its coefficients given by an extended Butcher tableau:

$$
\begin{array}{c|c|c|c}
c^{(0)} & A^{(0)} & B^{(0)} & \\
\hline
c^{(1)} & A^{(1)} & B^{(1)} & \\
\hline
 & \alpha^T & {\beta^{(1)}}^T & {\beta^{(2)}}^T
\end{array}
\tag{6.3}
$$

The application of the rooted tree analysis and Proposition 5.2 gives order conditions up to strong order 1.0 for the coefficients of the SRK method (6.1).

Theorem 6.1. *Let* $a, b^j \in C^{1,2}(\mathcal{I} \times \mathbb{R}^d, \mathbb{R}^d)$ *for* $j = 1, \ldots, m$. *If the coefficients of the SRK method* (6.1) *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* (1.1). *If* $a, b^j \in C^{1,3}(\mathcal{I} \times \mathbb{R}^d, \mathbb{R}^d)$ *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. ${\beta^{(2)}}^T (B^{(1)} (B^{(1)} e)) = 0$

are fulfilled and if $c^{(i)} = A^{(i)} e$ *for* $i = 0, 1$, *then the SRK method* (6.1) *attains order* 1.0 *for the strong approximation of the solution of the Itô SDE* (1.1).

For the proof of Theorem 6.1, we refer to Appendix A. The Euler–Maruyama (EM) scheme [12] is the basic explicit order 0.5 strong SRK scheme with $s = 1$ stage fulfilling the order 0.5 conditions of Theorem 6.1. It has the coefficients $\alpha_1 = \beta_1^{(1)} = 1$ 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 Table 6.1 define the order $(1.0, 1.0)$ strong SRK scheme SRI1 and the order $(2.0, 1.0)$ strong SRK scheme SRI2. For example, scheme SRI1 is then given by $Y_0 = X_{t_0}$ and

$$
\begin{aligned}
Y_{n+1} = Y_n + a(t_n, Y_n)\, h_n + \sum_{k=1}^{m} b^k(t_n, Y_n)\, I_{(k)} \\
+ \frac{1}{2} \sum_{k=1}^{m} \left(b^k(t_n, H_2^{(k)}) - b^k(t_n, H_3^{(k)}) \right) \sqrt{h_n}
\end{aligned}
\tag{6.4}
$$

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

$$H_2^{(k)} = Y_n + \sum_{l=1}^{m} b^l(t_n, Y_n) \, \frac{I_{(l,k)}}{\sqrt{h_n}} \, , \qquad H_3^{(k)} = Y_n - \sum_{l=1}^{m} b^l(t_n, Y_n) \, \frac{I_{(l,k)}}{\sqrt{h_n}} \tag{6.5}$$

for $k = 1, \ldots, m$.

It can be easily checked that the order 1.0 conditions from Theorem 6.1 are fulfilled. As the main advantage, the SRK scheme SRI1 needs one evaluation of the drift coefficient a and only three evaluations of each diffusion coefficient b^j, $j = 1, \ldots, m$, for each step. Thus, the number of evaluations of the drift and diffusion coefficients is independent of the dimension m of the Wiener process. Since at least $s = 3$ stages are needed in the case of an explicit SRK method (6.1), we recommend applying the order 1.0 strong SRK method presented in section 6.3 in the case of $m = 1$ where only $s = 2$ stages are necessary.

Similarly to the case of an Itô SDE (1.1), we can develop an efficient SRK method also for Stratonovich SDEs. For $d \geq 1$ and $m \geq 1$ the order 1.0 strong SRK method for a Stratonovich SDE (1.1) takes the form (6.1) with the stages (6.2), but with $I_{(k)}$ and $I_{(l,k)}$ replaced by $J_{(k)}$ and $J_{(l,k)}$ for $k, l \in \{1, \ldots, m\}$, respectively. The order 1.0 strong SRK method for Stratonovich SDEs is characterized by the same Butcher tableau (6.3). Calculating the order conditions up to strong order 1.0, we obtain the following theorem.

THEOREM 6.2. *Let $a, b^j \in C^{1,2}(\mathcal{I} \times \mathbb{R}^d, \mathbb{R}^d)$ for $j = 1, \ldots, m$. If the coefficients of the SRK method* (6.1)–(6.2) *with $I_{(k)}$ and $I_{(l,k)}$ replaced by $J_{(k)}$ and $J_{(l,k)}$ for $k, l \in \{1, \ldots, m\}$ fulfill equations* 1–5 *of Theorem* 6.1*, then the method attains order* 0.5 *for the strong approximation of the solution of the Stratonovich SDE* (1.1)*. If $a, b^j \in C^{1,3}(\mathcal{I} \times \mathbb{R}^d, \mathbb{R}^d)$ for $j = 1, \ldots, m$ and if in addition equations* 6–8 *of Theorem* 6.1 *are fulfilled and if $c^{(i)} = A^{(i)} e$ for $i = 0, 1$, then the SRK method attains order* 1.0 *for the strong approximation of the solution of the Stratonovich SDE* (1.1)*.*

Proof. The proof is analogous to the proof of Theorem 6.1. ∎

Here, it turns out that Theorem 6.2 contains the same strong order 1.0 conditions for the coefficients of the Stratonovich SRK method as Theorem 6.1 in the case of the Itô SRK method. Therefore, we define the explicit Stratonovich SRK schemes SRS1 of order $(1.0, 1.0)$ and SRS2 of order $(2.0, 1.0)$ with $s = 3$ stages for Stratonovich SDEs by the SRK method (6.1)–(6.2) with $I_{(k)}$ and $I_{(l,k)}$ replaced by $J_{(k)}$ and $J_{(l,k)}$ for $k, l \in \{1, \ldots, m\}$ applied with the coefficients presented in Table 6.1, respectively. Again, two evaluations of the drift coefficient a and only three evaluations of each diffusion coefficient b^j, $j = 1, \ldots, m$, are needed for each step in the case of scheme SRS1.

6.2. Order 1.0 strong SRK methods for SDEs with commutative noise. We consider the case of SDE (1.1) with commutative noise, i.e., if the diffusion satisfies the commutativity condition

$$\sum_{i=1}^{d} b^{i,j_1}(t,x) \frac{\partial b^{k,j_2}}{\partial x^i}(t,x) = \sum_{i=1}^{d} b^{i,j_2}(t,x) \frac{\partial b^{k,j_1}}{\partial x^i}(t,x) \tag{6.6}$$

for all $1 \leq j_1, j_2 \leq m$, $1 \leq k \leq d$, and $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$. Then order 1.0 strong approximation schemes making use only of increments of the Wiener process can be applied [5]. Thus, there is no need to simulate the multiple stochastic integrals $I_{(i,j)}$ anymore, which saves a lot of computational effort. However, to the best of the author's knowledge, as yet derivative-free order 1.0 strong approximation methods are known for Stratonovich SDEs with commutative noise (see, e.g., [2, 12]) or for linear Itô SDEs (see (1.2.30) in [20]) but not for general Itô SDEs. Therefore, we introduce for $d \geq 1$ and $m \geq 1$ the order 1.0 strong SRK method for the Itô SDE (1.1) with commutative noise which is defined by $Y_0 = X_{t_0}$ and

$$
\begin{aligned}
Y_{n+1} = Y_n &+ \sum_{i=1}^{s} \alpha_i \, a(t_n + c_i^{(0)} h_n, H_i^{(0)}) \, h_n \\
&+ \sum_{k=1}^{m} \sum_{i=1}^{s} \left(\beta_i^{(1)} I_{(k)} + \beta_i^{(2)} \sqrt{h_n} \right) b^k(t_n + c_i^{(1)} h_n, H_i^{(k)})
\end{aligned}
\tag{6.7}
$$

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

(6.8)

$$
\begin{aligned}
H_i^{(0)} &= Y_n + \sum_{j=1}^{s} A_{ij}^{(0)} \, a(t_n + c_j^{(0)} h_n, H_j^{(0)}) \, h_n + \sum_{l=1}^{m} \sum_{j=1}^{s} B_{ij}^{(0)} \, b^l(t_n + c_j^{(1)} h_n, H_j^{(l)}) \, I_{(l)}, \\
H_i^{(k)} &= Y_n + \sum_{j=1}^{s} A_{ij}^{(1)} \, a(t_n + c_j^{(0)} h_n, H_j^{(0)}) \, h_n + \sum_{l=1}^{m} \sum_{j=1}^{s} B_{ij}^{(1)} \, b^l(t_n + c_j^{(1)} h_n, H_j^{(l)}) \, \frac{I_{(k)} \, I_{(l)}}{2\sqrt{h_n}} \\
&\quad - \sum_{j=1}^{s} B_{ij}^{(1)} \, b^k(t_n + c_j^{(1)} h_n, H_j^{(k)}) \, \frac{\sqrt{h_n}}{2}
\end{aligned}
$$

for $i = 1, \dots, s$ and $k = 1, \dots, m$. Again, the SRK method (6.7) is characterized by the extended Butcher tableau (6.3).

THEOREM 6.3. *Let $a, b^j \in C^{1,2}(\mathcal{I} \times \mathbb{R}^d, \mathbb{R}^d)$ for $j = 1, \dots, m$. If the coefficients of the SRK method* (6.7) *fulfill equations* 1–3 *of Theorem* 6.1*, then the method attains strong order* 0.5 *if it is applied to the Itô SDE* (1.1)*. If $a, b^j \in C^{1,3}(\mathcal{I} \times \mathbb{R}^d, \mathbb{R}^d)$ for $j = 1, \dots, m$ and if in addition equations* 4–8 *of Theorem* 6.1 *are fulfilled and if $c^{(i)} = A^{(i)} e$ for $i = 0, 1$, then the SRK method* (6.7) *attains strong order* 1.0 *if it is applied to the Itô SDE* (1.1) *in the case of commutative noise.*

For the proof of Theorem 6.3, we refer to Appendix B. The coefficients in Table 6.1 can also be applied for the SRK method (6.7) which define the explicit SRK schemes SRIC1 and SRIC2 with $s = 3$ stages for Itô SDEs with commutative noise of strong order $(1.0, 1.0)$ and $(2.0, 1.0)$, respectively. Here, one evaluation of the drift coefficient a and 3 evaluations of each diffusion coefficient b^j, $j = 1, \dots, m$, are necessary each step.

6.3. Order 1.5 strong SRK methods for SDEs with scalar noise. Order 1.5 strong SRK methods for Stratonovich SDEs with scalar noise have been derived by Burrage and Burrage [1, 3, 4]. On the other hand, for Itô SDEs with scalar noise order 1.5 strong SRK methods have been proposed by Kaneko [11] and by Kloeden and Platen [12]. However, the scheme due to Kaneko [11] is not efficient because it needs 4 evaluations of the drift coefficient a, 12 evaluations of the diffusion coefficient b, and the simulation of 2 independent normally distributed random variables for each step. On the other hand, the scheme (11.2.1) in [12] due to Kloeden and Platen is derived from the order 1.5 strong Taylor scheme where the derivatives are replaced by finite differences. It needs 3 evaluations of the drift coefficient a, 5 evaluations of the diffusion b, and also the simulation of 2 independent normally distributed random variables for each step. In contrast to this, due to a systematic rooted tree approach, we introduce a new order 1.5 strong SRK method for Itô SDEs with scalar noise with less computational complexity compared to the well-known schemes.

For the Itô SDE (1.1) with $d \geq 1$ and $m = 1$ we propose the order 1.5 strong SRK method defined by $Y_0 = X_{t_0}$ and

(6.9)
$$\begin{aligned} Y_{n+1} = Y_n &+ \sum_{i=1}^{s} \alpha_i \, a(t_n + c_i^{(0)} h_n, H_i^{(0)}) \, h_n \\ &+ \sum_{i=1}^{s} \left(\beta_i^{(1)} I_{(1)} + \beta_i^{(2)} \frac{I_{(1,1)}}{\sqrt{h_n}} + \beta_i^{(3)} \frac{I_{(1,0)}}{h_n} + \beta_i^{(4)} \frac{I_{(1,1,1)}}{h_n} \right) b(t_n + c_i^{(1)} h_n, H_i^{(1)}) \end{aligned}$$

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

(6.10)
$$\begin{aligned} H_i^{(0)} &= Y_n + \sum_{j=1}^{s} A_{ij}^{(0)} \, a(t_n + c_j^{(0)} h_n, H_j^{(0)}) \, h_n + \sum_{j=1}^{s} B_{ij}^{(0)} \, b(t_n + c_j^{(1)} h_n, H_j^{(1)}) \, \frac{I_{(1,0)}}{h_n}, \\ H_i^{(1)} &= Y_n + \sum_{j=1}^{s} A_{ij}^{(1)} \, a(t_n + c_j^{(0)} h_n, H_j^{(0)}) \, h_n + \sum_{j=1}^{s} B_{ij}^{(1)} \, b(t_n + c_j^{(1)} h_n, H_j^{(1)}) \, \sqrt{h_n} \end{aligned}$$

for $i = 1, \ldots, s$. The SRK method (6.9) is characterized by the Butcher tableau

(6.11)
$$\begin{array}{c|c|c|c} c^{(0)} & A^{(0)} & B^{(0)} & \\ \hline c^{(1)} & A^{(1)} & B^{(1)} & \\ \hline & \alpha^T & {\beta^{(1)}}^T & {\beta^{(2)}}^T \\ \hline & & {\beta^{(3)}}^T & {\beta^{(4)}}^T \end{array}$$

Then we obtain order conditions up to strong order 1.5 by applying the colored rooted tree analysis with bicolored trees.

THEOREM 6.4. *Let* $a, b \in C^{1,2}(\mathcal{I} \times \mathbb{R}^d, \mathbb{R}^d)$. *If the coefficients of the SRK method* (6.9) *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* (1.1). *If* $a, b \in C^{1,3}(\mathcal{I} \times \mathbb{R}^d, \mathbb{R}^d)$ *and if in addition the equations*

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* (6.9) *attains order* 1.0 *for the strong approximation of the solution of the Itô SDE* (1.1) *with scalar noise. If* $a, b \in C^{2,4}(\mathcal{I} \times \mathbb{R}^d, \mathbb{R}^d)$ *and if in addition the equations*

10. $\alpha^T A^{(0)} e = \frac{1}{2}$ 11. $\alpha^T B^{(0)} e = 1$ 12. $\alpha^T (B^{(0)} e)^2 = \frac{3}{2}$

Table 6.2
SRK scheme SRI1W1 of order $(2.0, 1.5)$ *and SRI2W1 of order* $(3.0, 1.5)$.

0												
$\frac{3}{4}$	$\frac{3}{4}$				$\frac{3}{2}$							
0	0	0			0	0						
0	0	0	0		0	0	0					
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	$\frac{1}{2}$					
	$\frac{1}{3}$	$\frac{2}{3}$	0	0	-1	$\frac{4}{3}$	$\frac{2}{3}$	0	-1	$\frac{4}{3}$	$-\frac{1}{3}$	0
					2	$-\frac{4}{3}$	$-\frac{2}{3}$	0	-2	$\frac{5}{3}$	$-\frac{2}{3}$	1

0												
1	1				0							
$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$			1	$\frac{1}{2}$						
0	0	0	0		0	0	0					
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	$\frac{1}{2}$					
	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{2}{3}$	0	-1	$\frac{4}{3}$	$\frac{2}{3}$	0	1	$-\frac{4}{3}$	$\frac{1}{3}$	0
					2	$-\frac{4}{3}$	$-\frac{2}{3}$	0	-2	$\frac{5}{3}$	$-\frac{2}{3}$	1

13. ${\beta^{(1)}}^T A^{(1)} e = 1$
14. ${\beta^{(2)}}^T A^{(1)} e = 0$
15. ${\beta^{(3)}}^T A^{(1)} e = -1$
16. ${\beta^{(4)}}^T A^{(1)} e = 0$
17. ${\beta^{(1)}}^T (B^{(1)} e)^2 = 1$
18. ${\beta^{(2)}}^T (B^{(1)} e)^2 = 0$
19. ${\beta^{(3)}}^T (B^{(1)} e)^2 = -1$
20. ${\beta^{(4)}}^T (B^{(1)} e)^2 = 2$
21. ${\beta^{(1)}}^T (B^{(1)}(B^{(1)} e)) = 0$
22. ${\beta^{(2)}}^T (B^{(1)}(B^{(1)} e)) = 0$
23. ${\beta^{(3)}}^T (B^{(1)}(B^{(1)} e)) = 0$
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* (6.9) *attains order* 1.5 *for the strong approximation of the solution of the Itô SDE* (1.1) *in the case of scalar noise.*

Proof. We refer to the proof of Theorem 6.5 with $m = 1$ and $d \geq 1$, which contains all calculations for the proof of Theorem 6.4 as a special case. □

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 presented in Table 6.2. The SRK scheme SRI1W1 needs only 2 evaluations of the drift coefficient, 4 evaluations of the diffusion coefficient b, and the simulation of 2 independent normally distributed random variables for each step. Thus, the SRK method (6.9) with the coefficients SRI1W1 saves one evaluation of the drift as well as one evaluation of the diffusion in each step compared to the explicit order 1.5 strong scheme proposed by Kloeden and Platen [12]. Note that the explicit

2-stage SRK method (6.9) with coefficients $\alpha_1 = \beta_1^{(1)} = \beta_2^{(2)} = A_{2,1}^{(1)} = B_{2,1}^{(1)} = 1$, $\beta_1^{(2)} = -1$, and $\alpha_2 = \beta_2^{(1)} = A_{2,1}^{(0)} = B_{2,1}^{(0)} = \beta_1^{(3)} = \beta_2^{(3)} = \beta_1^{(4)} = \beta_2^{(4)} = 0$ coincides with the order 1.0 strong scheme (11.1.3) in [12] due to Kloeden and Platen.

6.4. Order 1.5 strong SRK methods for SDEs with diagonal noise. We consider the Itô SDE (1.1) with diagonal noise where $d = m \geq 1$ and the diagonal diffusion coefficient $b^{k,k}$ depends only on t and x^k, i.e., $b^{k,j}(t,x) \equiv 0$ and $\frac{\partial b^{j,j}}{\partial x^k}(t,x) \equiv 0$ for all $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$ and $1 \leq j,k \leq m$ with $k \neq j$. Here, each component X^k of the solution X is disturbed only by the corresponding component W^k of the driving Wiener process W. Thus, the components of the solution process are coupled only through the drift term. Then particular order 1.5 strong approximation methods can be applied which make use only of the increments $I_{(i)}$ of the Wiener process and the multiple integrals $I_{(i,0)}$, but which do not need the simulation of mixed multiple stochastic integrals such as $I_{(i,j)}$ for $1 \leq i,j \leq m$ with $i \neq j$. Kloeden and Platen proposed the derivative-free order 1.5 strong scheme (11.2.10) in [12] for Itô SDEs with diagonal noise. However, their scheme needs $2m+1$ evaluations of the drift coefficient a, 5 evaluations of each diffusion coefficient b^j, $j = 1,\ldots,m$, and the simulation of $2m$ independent normally distributed random variables for each step. Clearly, the computational complexity is increasingly large, especially for high dimensional problems. Therefore, we introduce a new order 1.5 strong SRK method where the number of evaluations per step of the drift coefficient a and each diffusion coefficient b^j, $j = 1,\ldots,m$, does not depend on the dimension m. Further, the proposed SRK method needs also only $2m$ independent normally distributed random variables which have to be simulated each step.

For the Itô SDE (1.1) with $d = m \geq 1$ and diagonal noise we propose the order 1.5 strong SRK method defined by $Y_0 = X_{t_0}$ and

(6.12)

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

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

(6.13)

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

for $i = 1,\ldots,s$. We can characterize the SRK method (6.12) by the Butcher tableau (6.11). Then we calculate the following order conditions.

THEOREM 6.5. *Let $a, b^j \in C^{1,2}(\mathcal{I} \times \mathbb{R}^d, \mathbb{R}^d)$ for $1 \leq j \leq m$. If the coefficients of the SRK method* (6.12) *fulfill equations* 1–5 *of Theorem* 6.4*, then the method attains order* 0.5 *for the strong approximation of the solution of the Itô SDE* (1.1)*. If $a, b^j \in C^{1,3}(\mathcal{I} \times \mathbb{R}^d, \mathbb{R}^d)$ for $1 \leq j \leq m$ and if in addition equations* 6–9 *of Theorem* 6.4 *are fulfilled and if $c^{(i)} = A^{(i)} e$ for $i = 0,1$, then the SRK method* (6.12) *attains order*

1.0 *for the strong approximation of the solution of the Itô SDE* (1.1) *with diagonal noise. If* $a, b^j \in C^{2,4}(\mathcal{I} \times \mathbb{R}^d, \mathbb{R}^d)$ *for* $1 \leq j \leq m$ *and if in addition equations* 10–25 *of Theorem* 6.4 *are fulfilled and if* $c^{(i)} = A^{(i)}e$ *for* $i = 0, 1$, *then the SRK method* (6.12) *attains order* 1.5 *for the strong approximation of the solution of the Itô SDE* (1.1) *in the case of diagonal noise.*

For the proof of Theorem 6.5, we refer to Appendix C. Since the order conditions for SRK method (6.12) coincide with the ones of SRK method (6.9), the coefficients presented in Table 6.2 define the schemes SRID1 of order $(2.0, 1.5)$ and SRID2 of order $(3.0, 1.5)$ for the SRK method (6.12), respectively. As the main novelty, the SRK scheme SRID1 needs only 2 evaluations of the drift coefficient a, 4 evaluations of each diffusion coefficient b^j, $j = 1, \ldots, m$, and the simulation of $2m$ independent normally distributed random variables for each step. This is a significant reduction of the computational complexity compared to the scheme due to Kloeden and Platen [12].

6.5. Order 1.5 strong SRK methods for SDEs with additive noise. We consider SDE (1.1) with m-dimensional additive noise, i.e., $b(t, x) \equiv b(t)$ for all $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$. Then the Itô SDE (1.1) coincides with the Stratonovich SDE (1.1) and thus only one approximation method is needed. To the best of the author's knowledge, essentially there exist two order 1.5 strong approximation methods avoiding derivatives in the literature. The scheme (1.5.53) in [20] due to Milstein is not completely derivative free and needs 3 evaluations of the drift coefficient a, 1 evaluation of each diffusion coefficient b^j, and 1 evaluation of the derivative of each b^j with respect to the time t for $j = 1, \ldots, m$ at each step. On the other hand, the scheme (11.2.19) in [12] due to Kloeden and Platen is completely derivative free; however, it needs $2m + 1$ evaluations of the drift coefficient a and 2 evaluations of each diffusion coefficient b^j, $j = 1, \ldots, m$, for each step. Further, both schemes need the simulation of $2m$ independent normally distributed random variables for each step. So, again the computational complexity depends on m, which is a drawback for high dimensional problems. Therefore, we propose a newly structured SRK method which is completely derivative free and where the number of evaluations of the drift and diffusion coefficients no longer depends on m.

For SDE (1.1) with additive noise we propose in the case of $d \geq 1$ and $m \geq 1$ the order 1.5 strong SRK method defined by $Y_0 = X_{t_0}$ and

$$\begin{aligned} Y_{n+1} = Y_n &+ \sum_{i=1}^{s} \alpha_i \, a(t_n + c_i^{(0)} h_n, H_i^{(0)}) \, h_n \\ &+ \sum_{k=1}^{m} \sum_{i=1}^{s} \left(\beta_i^{(1)} I_{(k)} + \beta_i^{(2)} \frac{I_{(k,0)}}{h_n} \right) b^k(t_n + c_i^{(1)} h_n) \end{aligned} \tag{6.14}$$

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

$$H_i^{(0)} = Y_n + \sum_{j=1}^{s} A_{ij}^{(0)} \, a(t_n + c_j^{(0)} h_n, H_j^{(0)}) \, h_n + \sum_{l=1}^{m} \sum_{j=1}^{s} B_{ij}^{(0)} \, b^l(t_n + c_j^{(1)} h_n) \, \frac{I_{(l,0)}}{h_n} \tag{6.15}$$

for $i = 1, \ldots, s$. The SRK method (6.14) is characterized by the Butcher tableau

$$\begin{array}{c|c|c|c} c^{(0)} & A^{(0)} & B^{(0)} & c^{(1)} \\ \hline & \alpha^T & {\beta^{(1)}}^T & {\beta^{(2)}}^T \end{array} \tag{6.16}$$

Table 6.3

*Coefficients for the SRK schemes SRA*1 *and SRA*2 *of order* $(2.0, 1.5)$ *at the top and in the middle, and SRA*3 *of order* $(3.0, 1.5)$ *at the bottom.*

$$\begin{array}{c|cc|cc|cc}
0 & & & & & 1 & \\
\frac{3}{4} & \frac{3}{4} & & \frac{3}{2} & & 0 & \\
\hline
 & \frac{1}{3} & \frac{2}{3} & 1 & 0 & -1 & 1
\end{array}$$

$$\begin{array}{c|cc|cc|cc}
0 & & & & & \frac{1}{3} & \\
\frac{3}{4} & \frac{3}{4} & & \frac{3}{2} & & 1 & \\
\hline
 & \frac{1}{3} & \frac{2}{3} & 0 & 1 & -\frac{3}{2} & \frac{3}{2}
\end{array}$$

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

Applying the colored rooted tree analysis results in the strong order conditions for the proposed SRK method (6.14).

Theorem 6.6. *Let* $a \in C^{1,3}(\mathcal{I} \times \mathbb{R}^d, \mathbb{R}^d)$ *and* $b^j \in C^1(\mathcal{I}, \mathbb{R}^d)$ *for* $1 \leq j \leq m$. *If the coefficients of the SRK method* (6.14) *fulfill the equations*

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

then the method attains order 1.0 *for the strong approximation of the solution of the Itô and the Stratonovich SDE* (1.1) *with additive noise. If* $a \in C^{2,4}(\mathcal{I} \times \mathbb{R}^d, \mathbb{R}^d)$ *and* $b^j \in C^2(\mathcal{I}, \mathbb{R}^d)$ *for* $1 \leq j \leq m$ *and if in addition the equations*

$$4.\quad \alpha^T B^{(0)} e = 1 \qquad 5.\quad \alpha^T A^{(0)} e = \frac{1}{2} \qquad 6.\quad \alpha^T (B^{(0)} e)^2 = \tfrac{3}{2}$$

$$7.\quad {\beta^{(1)}}^T c^{(1)} = 1 \qquad 8.\quad {\beta^{(2)}}^T c^{(1)} = -1$$

are fulfilled and if $c^{(0)} = A^{(0)} e$, *then the SRK method* (6.14) *attains order* 1.5 *for the strong approximation of the solution of the Itô and the Stratonovich SDE* (1.1) *in the case of additive noise.*

For the proof of Theorem 6.6, we refer to Appendix D. Coefficients for SRK method (6.14) defining the explicit 2-stage SRK schemes SRA1 and SRA2 of order $(2.0, 1.5)$ are presented in the top and middle tables, respectively, of Table 6.3 and for the explicit 3-stage SRK scheme SRA3 of order $(3.0, 1.5)$ in the bottom table of Table 6.3. As the main novelty, the complexity of the SRK scheme SRA1 is significantly reduced compared to the well-known schemes. This is due to the fact that only 2 evaluations of the drift coefficient a and also only 2 evaluations of each diffusion coefficient b^j, $j = 1, \ldots, m$, are necessary for each step. Finally, the SRK method (6.14) needs the simulation of $2m$ independent normally distributed random variables for each step.

7. Numerical results.

Now, we apply the introduced SRK methods to some test SDEs in order to compare their performance with some widely used approximations methods. Therefore, let EM denote the order 0.5 strong Euler–Maruyama scheme [12] (see also section 6.1) and let MIL denote the order 1.0 strong Milstein scheme [19] (see also scheme (10.3.3) in [12]). Further, we apply the following explicit derivative-free schemes proposed by Kloeden and Platen: the order 1.0 strong scheme (11.1.7) [12] denoted as SPLI, the order 1.0 strong scheme (11.1.11) [12] for Stratonovich SDEs with commutative noise denoted as SKPSC, the order 1.5 strong scheme (11.2.1) [12] called SPLIW1 for Itô SDEs with scalar noise, the order 1.5 strong scheme (11.2.10) [12] for Itô SDEs with diagonal noise denoted as SKPID, and the order 1.5 strong scheme (11.2.19) [12] for SDEs with additive noise denoted as SKPA. Notice that the scheme SKPID coincides with the scheme SPLIW1 in the case of $d = m = 1$. The same applies for the scheme SRID1, which coincides with SRI1W1 if $d = m = 1$. Therefore, we do not mention the schemes SKPID and SRID1 separately in these cases.

As a measure for the computational effort, we take the number of evaluations of the drift and diffusion functions as well as the number of realizations of independent random variables needed for the calculation of the approximation. In particular, if the approximation method applies the random variables $I_{(i,j)}$ for $1 \leq i, j \leq m$ with $i \neq j$, then $I_{(i,j)}$ has to be approximated because the distribution of $I_{(i,j)}$ is not known. Therefore, we denote by $c = c(C, m, h)$ the number of realizations of independent random variables needed for the approximation of $I_{(i,j)}$ for $i \neq j$.

For the presented numerical results, $I_{(i,j)}$ is approximated by the method proposed in [30]. Here, we need to simulate $c(C, m, h) = \frac{1}{2}m(m-1)+2m\,q$ independent normally distributed random variables each step with $q \geq \sqrt{5m^2(m-1)/(24\pi^2)}\,(C\,h)^{-1/2}$ in order to approximate $I_{(i,j)}$ with a mean-square error no larger than Ch^3 [30], provided that the m random variables $I_{(i)}$ are given. Thus, the additional computational effort required for each step depends on the step size h and increases with order $c(C, m, h) = O(h^{-1/2})$ as $h \to 0$.

The computational complexity of the considered schemes as well as of the order 1.5 strong SRK scheme G5 due to Burrage and Burrage [3] for Stratonovich SDEs with scalar noise is given in Table 7.1. For example, the computational complexity of the scheme MIL is $d + md + md^2 + m + c$, whereas for scheme SRI1 we have only the complexity $d + 3md + m + c$ for each step. Thus, the introduced SRK 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, in the case of $d \geq 1$ and $m > 2$ the scheme SRI1 also has lower complexity than the scheme SPLI1 due to Kloeden and Platen having computational complexity $d+m^2d+md+m+c$. Clearly, the improved efficiency of SRI1 is independent of c, i.e., of the method chosen to simulate $I_{(i,j)}$ for $i \neq j$. As a result of this, the SRK scheme SRI1 is a serious alternative for the Milstein scheme MIL and the SRK scheme SPLI1.

If not stated otherwise, for each considered step size h we simulate 2000 trajectories by each scheme under consideration and take the mean of the corresponding attained errors at $T = 1$ as an estimator for the expectation in (1.2). Then we analyze the mean-square errors versus the computational effort in log-log diagrams with base two. We denote by p_{eff} the effective order of convergence which is the slope of the resulting line in the mean-square errors versus effort log-log diagrams as $h \to 0$. If $\hat{e}(h)$ denotes the mean-square error and $\hat{c}(h)$ are the computational costs based on step size h, then

$$p_{\text{eff}} = \lim_{h \to 0} \left| \frac{\log(\hat{e}(h)) - \log(\hat{e}(h/2))}{\log(\hat{c}(h)) - \log(\hat{c}(h/2))} \right| .$$

TABLE 7.1
Computational complexity of various schemes per step for a d-dimensional SDE system with a m-dimensional Wiener process ($m = 1$ for SPLIW1 and SRI1W1).

		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	$m\,d$	–	+	–	–
MIL	1.0	d	$m\,d$	$m\,d^2$	+	–	+
SPLI	1.0	d	$(m^2+m)\,d$	–	+	–	+
SKPSC	1.0	d	$2\,m\,d$	–	+	–	–
SPLIW1	1.5	$3\,d$	$5\,d$	–	+	+	–
G5	1.5	$5\,d$	$5\,d$	–	+	+	–
SKPID	1.5	$(2m+1)\,d$	$5\,m\,d$	–	+	+	–
SKPA	1.5	$(2m+1)\,d$	$2\,m\,d$	–	+	+	–
SRI1	1.0	d	$3\,m\,d$	–	+	–	+
SRS1	1.0	d	$3\,m\,d$	–	+	–	+
SRIC1	1.0	d	$3\,m\,d$	–	+	–	–
SRI1W1	1.5	$2\,d$	$4\,d$	–	+	+	–
SRID1	1.5	$2\,d$	$4\,m\,d$	–	+	+	–
SRA1	1.5	$2\,d$	$2\,m\,d$	–	+	+	–

Considering the effective order may cause an order reduction if $\hat{c}(h)$ depends on the step size h. This is the case for any strong order 1.0 scheme in the case of noncommutative noise which attains the effective order $p_{\text{eff}} = 2/3$ as $h \to 0$. This is due to the effort $c(C, m, h)$ for the approximation of the multiple integrals $I_{(i,j)}$ for each step, which is $O(h^{-1/2})$ [30]. Clearly, if a more efficient approximation method for $I_{(i,j)}$ would be available, then p_{eff} may be improved up to the strong order 1.0 of the scheme. However, compared to the order 0.5 strong Euler–Maruyama scheme, which attains the effective order $p_{\text{eff}} = 0.5$, 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 confirmed by the simulation results.

Further, the mean-square error (1.2) at time $T = 1$ versus the step size is analyzed in log-log diagrams with base two for some examples now showing the strong order p. Therefore, dotted order lines with slope 0.5, 1.0, 2/3, and 1.5 are plotted in the figures as a reference.

First, we consider for $d = m = 1$ a nonlinear autonomous Itô SDE [12]

$$\mathrm{d}X_t = -\frac{1}{4}X_t(1-X_t^2)\,\mathrm{d}t + \frac{1}{2}(1-X_t^2)\,\mathrm{d}W_t\,, \quad X_0 = \frac{1}{2}\,, \tag{7.1}$$

with solution $X_t = \frac{(1+X_0)\,\exp(W_t)+X_0-1}{(1+X_0)\,\exp(W_t)+1-X_0}$. The corresponding numerical results for step sizes $h = 2^0, 2^{-1}, \ldots, 2^{-18}$ are presented on the left of Figure 7.1. Here, the scheme SRI1W1 attains the effective order 1.5 and the best performance, particularly compared to the order 1.5 scheme SPLIW1 due to Kloeden and Platen.

The second example for $d = m = 1$ is the nonlinear autonomous Itô SDE [12]

$$\mathrm{d}X_t = -\left(\frac{1}{10}\right)^2 \sin(X_t)\cos^3(X_t)\,\mathrm{d}t + \frac{1}{10}\cos^2(X_t)\,\mathrm{d}W_t\,, \quad X_0 = 1\,, \tag{7.2}$$

with solution $X_t = \arctan(\frac{1}{10}W_t + \tan(X_0))$. Here, for $h = 2^0, \ldots, 2^{-16}$ the results are plotted on the right of Figure 7.1. Again, the proposed SRK scheme SRI1W1 has

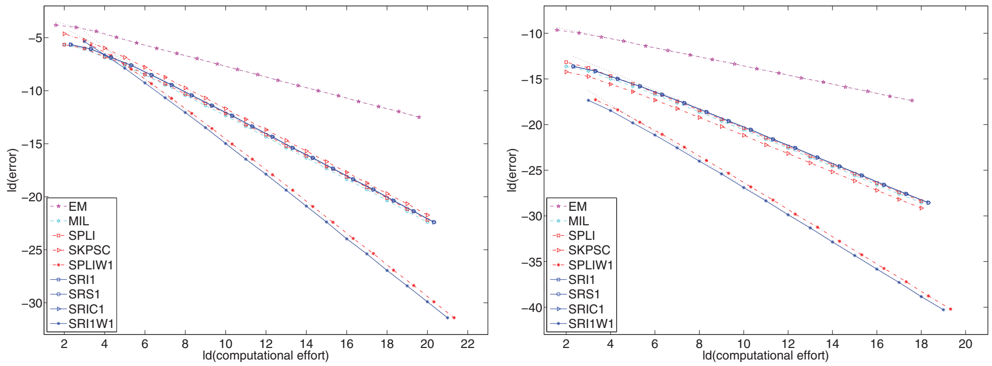

FIG. 7.1. *Errors vs. effort for SDE* (7.1) *and SDE* (7.2).

the effective order 1.5 and performs better than the other approximation schemes due to its reduced complexity.

As an example for additive noise, we analyze for $d = 1$ and $m \geq 1$ the SDE

$$\mathrm{d}X_t = \left(\frac{\beta}{\sqrt{1+t}} - \frac{1}{2(1+t)} X_t \right) \mathrm{d}t + \sum_{j=1}^{m} \alpha_j \frac{\beta}{\sqrt{1+t}} \, \mathrm{d}W_t^j \,, \quad X_0 = 1 \,, \tag{7.3}$$

with solution $X_t = \frac{1}{\sqrt{1+t}} X_0 + \frac{\beta}{\sqrt{1+t}} (t + \sum_{j=1}^m \alpha_j \, W_t^j)$. We consider SDE (7.3) with parameters $\alpha_j = \frac{1}{10}$ for $j = 1, \ldots, m$ and $\beta = \frac{1}{2}$. For the cases of $m = 1$ and $m = 4$ the corresponding results for $h = 2^0, \ldots, 2^{-16}$ are presented on the left and on the right at the top of Figure 7.2, respectively. Further, for $m = 10$ the effective order as well as the strong order are analyzed for $h = 2^0, \ldots, 2^{-14}$ on the left and on the right at the bottom of Figure 7.2, respectively. Here, we observe that the scheme SKPA attains only strong order 1.0 if $m > 1$. The introduced SRK scheme SRA1 designed for SDEs with additive noise has by far the best performance in all cases. Further, the effective as well as the strong order are both even higher than 1.5 for the considered example.

In order to consider also a multidimensional 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 $B_{ij}^k = \frac{1}{100}$ for $i \neq j$ and $B_{ii}^k = \frac{1}{5}$ for $1 \leq i, j \leq d$ and $k = 1, \ldots, m$. Then we consider the Itô SDE

$$\mathrm{d}X_t = A X_t \, \mathrm{d}t + \sum_{k=1}^{m} B^k X_t \, \mathrm{d}W_t^k \,, \quad X_0 = (1, \ldots, 1)^T \in \mathbb{R}^d \,, \tag{7.4}$$

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 cases of $m = 2$ with $h = 2^0, \ldots, 2^{-17}$ and $m = 4$ with $h = 2^0, \ldots, 2^{-16}$ the corresponding results are given on the left and on the right at the top of Figure 7.3. Further, for the case of $m = 10$ the effective and the strong orders are analyzed for $h = 2^0, \ldots, 2^{-15}$ on the left and on the right at the bottom of Figure 7.3. Here, we can see that the schemes MIL, SPLI, SRI1, and SRS1 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. Here, we have to point out that the introduced SRK schemes SRI1 and SRS1 show

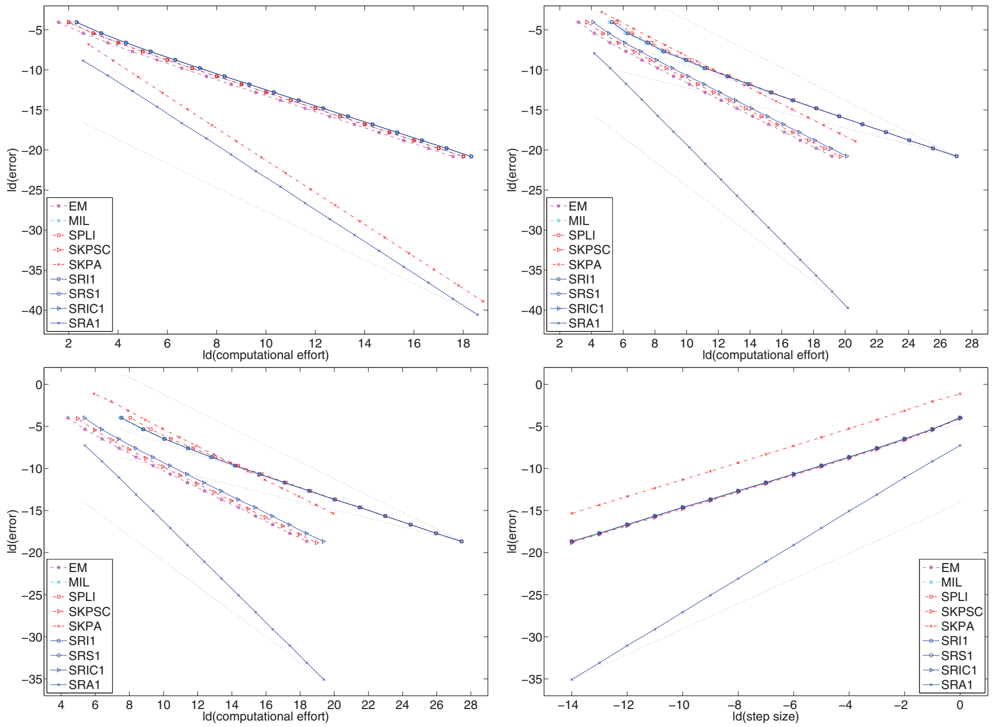

FIG. 7.2. *Errors vs. effort for SDE* (7.3) *with* $m = 1$ *and* $m = 4$ *(top). Errors vs. effort and errors vs. step sizes for SDE* (7.3) *with* $m = 10$ *(bottom).*

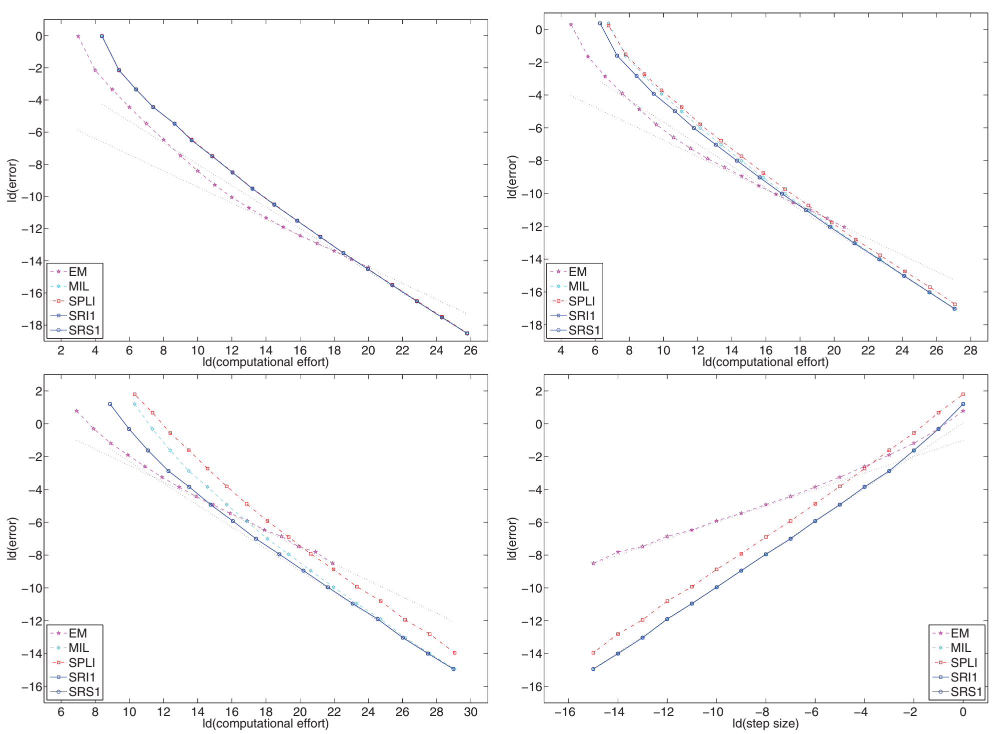

FIG. 7.3. *Errors vs. effort for SDE* (7.4) *with* $d = m = 2$ *and* $d = m = 4$ *(top). Errors vs. effort and errors vs. step sizes for SDE* (7.4) *with* $d = m = 10$ *(bottom).*

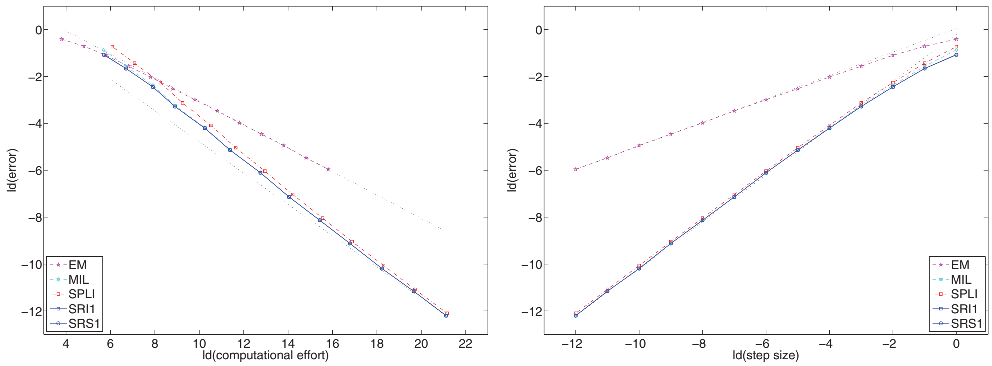

FIG. 7.4. *Errors vs. effort and errors vs. step sizes for SDE* (7.5) *with $d = 2$ and $m = 4$.*

the best performance, especially compared to the Milstein scheme MIL and the SRK scheme SPLI. This is confirmed by the increasing distance of the effective order lines of SRI1 and SRS1 from the order lines of the Milstein scheme and the scheme SPLI as d and m are increasing. Thus, the improvements by SRI1 and SRS1 become significant, particularly for high dimensional problems where the computational complexity of the schemes SRI1 and SRS1 is considerably lower than for the Milstein scheme and the scheme SPLI due to Kloeden and Platen [12].

As an example for a nonlinear multidimensional Itô SDE with $d = 2$ and $m = 4$ possessing noncommutative noise, we consider the two-dimensional stochastic flow on the manifold given by a torus [12]

$$\begin{aligned} \mathrm{d}\begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} &= \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix} \sin(X_t^1)\,\mathrm{d}W_t^1 + \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix} \cos(X_t^1)\,\mathrm{d}W_t^2 \\ &\quad + \begin{pmatrix} -\sin(\alpha) \\ \cos(\alpha) \end{pmatrix} \sin(X_t^2)\,\mathrm{d}W_t^3 + \begin{pmatrix} -\sin(\alpha) \\ \cos(\alpha) \end{pmatrix} \cos(X_t^2)\,\mathrm{d}W_t^4 \end{aligned} \tag{7.5}$$

with $\alpha = \frac{1}{2}$ and $X_0 = (2, 2)^T$. Here, a reference solution has been simulated with the Milstein scheme based on step size 2^{-14}. The results for $h = 2^0, \dots, 2^{-12}$ are presented in Figure 7.4. The strong order 1.0 schemes SRI1, SRS1, and MIL perform similar due to $d = 2$ and attain the strong order 1.0 and the effective order 2/3; however, the scheme MIL needs additionally the calculation of the Jacobian matrix, which has not been taken into account for the computational effort. The schemes SRI1, SRS1, and MIL perform better than the scheme SPLI due to $m = 4$ and much better than the scheme EM with strong and effective order 1/2.

As a high dimensional nonlinear example with noncommutative noise, we consider the stochastic Lotka–Volterra system, which is used as a multispecies model with interaction in population dynamics [12]. The corresponding d-dimensional SDE system with an m-dimensional Wiener process is given by

$$\mathrm{d}X_t^i = \left(\alpha^i X_t^i + \sum_{j=1}^{d} \beta^{i,j} X_t^j \, X_t^i \right) \mathrm{d}t + \sum_{j=1}^{m} b^{i,j}(X_t)\,\mathrm{d}W_t^j \tag{7.6}$$

for $i = 1, \dots, d$ with $X_0 = \frac{1}{10}(1, \dots, 1)^T \in \mathbb{R}^d$ and $m = d$. Here, we choose $\alpha^i = \frac{1}{5}$,

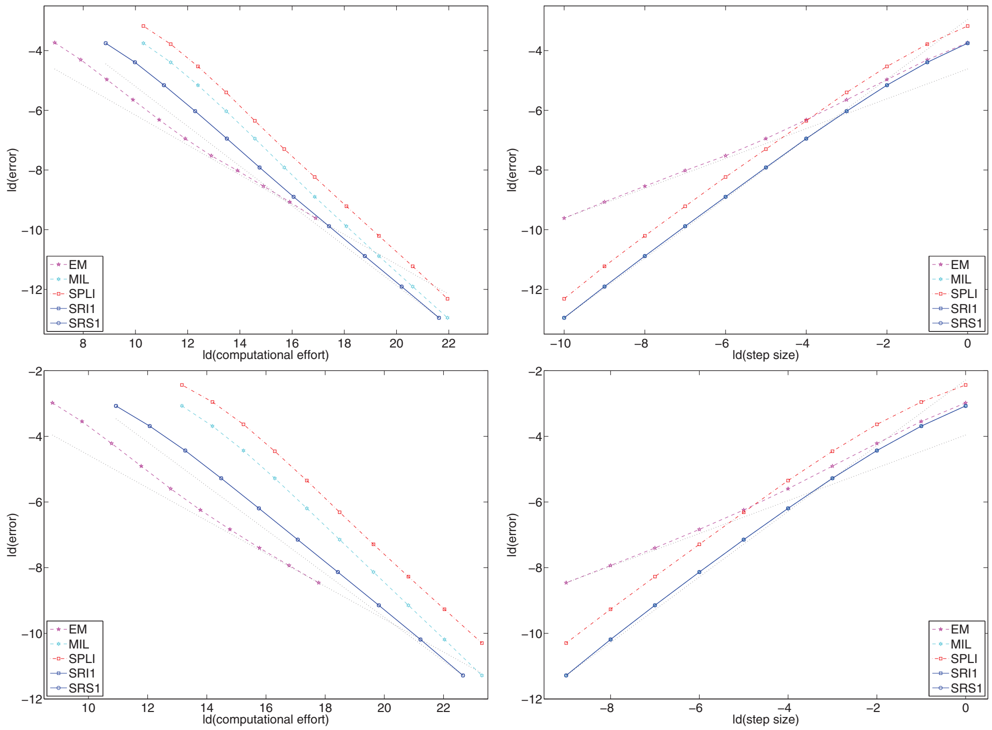

FIG. 7.5. *Errors vs. effort and errors vs. step sizes for SDE* (7.6) *with* $d = m = 10$ *(top) and with* $d = m = 20$ *(bottom).*

$\beta^{i,j} = \frac{1}{10} d^{-1}$ for $i \neq j$, $\beta^{i,j} = \frac{3}{2}$ for $i = j$, $b^{i,j}(X_t) = \frac{1}{5} X_t^j$ for $j-1 \leq i \leq j+1$, and $b^{i,j}(X_t) = 0$ for $i \notin \{j-1, j, j+1\}$. Explicit solutions are not known for SDE (7.6); therefore a reference solution is simulated using the Milstein scheme with step sizes 2^{-14} and 2^{-12}. The results for step sizes $h = 2^0, \dots, 2^{-10}$ in the case of $d = m = 10$ based on 250 trajectories and for $h = 2^0, \dots, 2^{-9}$ in the case of $d = m = 20$ based on 1000 trajectories are presented in Figure 7.5. In both cases, the schemes SRI1, SRS1, MIL, and SPLI attain strong order 1.0, while the scheme EM shows strong order 1/2 if we compare their errors with the used step sizes. Comparing their errors with the corresponding computational effort, the scheme EM attains the effective order 1/2, while the strong order 1.0 schemes attain the effective order 2/3. Further we can see that the new schemes SRI1 and SRS1 perform much better than the other schemes under consideration with increasing savings of computational effort as the dimension of the SDE system increases.

8. Conclusions. In the present paper, efficient order 1.0 strong SRK methods for Itô and Stratonovich SDEs with a multidimensional driving Wiener process are proposed. Compared to well-known schemes such as the SRK scheme SPLI [12, (11.1.7)] or the Milstein scheme MIL [19], the computational costs for the proposed SRK method (6.1) are significantly reduced; see Table 7.1. In the case of Itô SDEs with commutative noise, an order 1.0 strong SRK method (6.7) which needs only the simulation of increments of the driving Wiener process is introduced. Finally, order 1.5 strong SRK methods are proposed in the case of scalar noise (6.9), diagonal noise (6.12), and additive noise (6.14). Again, all of the proposed methods feature a significant reduction of computational costs compared to the schemes in the literature; see again Table 7.1. Full order conditions are calculated for all of these SRK methods based on the colored rooted tree analysis. Some coefficients fulfilling the corresponding order conditions are given for each of the methods. Clearly, there are some degrees of freedom in choosing the coefficients which may be used in order to minimize truncation terms. Further, an analysis of the space of admissible coefficients as well as coefficients for implicit schemes with good stability properties is an open question for future research.

Appendix A. Proof of Theorem 6.1. First, notice that the SRK method (6.1) is covered by the general class (2.1) if we choose $\mathcal{M} = \{\nu : 0 \leq \nu \leq m\}$ and if we assign

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

for $1 \leq k,l \leq m$ and if all other coefficients in (2.1) are set equal to zero. Now, we apply Proposition 5.2 with $p = 1$. Then (5.5) yields for $j_1, j_2 \in \{1, \ldots, m\}$ the following order conditions:

$$I_{\gamma;t,t+h} = \Phi_S(\gamma; t, t+h) \quad \Leftrightarrow \quad 1 = 1\,, \tag{A.2}$$

$$\begin{aligned} &I_{\tau_{j_1};t,t+h} = \Phi_S(\tau_{j_1}; t, t+h) \quad \Leftrightarrow \quad I_{(j_1)} = {z^{(j_1),(j_1)}}^T e \\ &\Leftrightarrow \quad I_{(j_1)} = {\beta^{(1)}}^T e\, I_{(j_1)} + {\beta^{(2)}}^T e \sqrt{h} \\ &\Leftrightarrow \quad {\beta^{(1)}}^T e = 1 \quad \wedge \quad {\beta^{(2)}}^T e = 0\,, \end{aligned} \tag{A.3}$$

$$\begin{aligned} &I_{\tau_0;t,t+h} = \Phi_S(\tau_0; t, t+h) \quad \Leftrightarrow \quad I_{(0)} = {z^{(0),(0)}}^T e \\ &\Leftrightarrow \quad I_{(0)} = \alpha^T e\, h \quad \Leftrightarrow \quad \alpha^T e = 1\,, \end{aligned} \tag{A.4}$$

$$\begin{aligned} &I_{[\tau_{j_2}]_{j_1};t,t+h} = \Phi_S([\tau_{j_2}]_{j_1}; t, t+h) \quad \Leftrightarrow \quad I_{(j_2,j_1)} = {z^{(j_1),(j_1)}}^T Z^{(j_1),(j_2),(j_2)} e \\ &\Leftrightarrow \quad I_{(j_2,j_1)} = ({\beta^{(1)}}^T I_{(j_1)} + {\beta^{(2)}}^T \sqrt{h}) B^{(1)} e\, \frac{I_{(j_2,j_1)}}{\sqrt{h}} \\ &\Leftrightarrow \quad {\beta^{(1)}}^T B^{(1)} e = 0 \quad \wedge \quad {\beta^{(2)}}^T B^{(1)} e = 1\,. \end{aligned} \tag{A.5}$$

Further, we calculate for $j_1, j_2, j_3 \in \{1, \ldots, m\}$ from (5.6) the conditions

$$\begin{aligned} &\mathrm{E}(I_{[\tau_{j_1}]_0;t,t+h}) = \mathrm{E}(\Phi_S([\tau_{j_1}]_0; t, t+h)) \\ &\Leftrightarrow \quad \mathrm{E}(I_{(j_1,0)}) = \mathrm{E}({z^{(0),(0)}}^T Z^{(0),(j_1),(j_1)} e) \\ &\Leftrightarrow \quad \mathrm{E}(I_{(j_1,0)}) = \alpha^T B^{(0)} e\, h\, \mathrm{E}(I_{(j_1)}) \quad \Leftrightarrow \quad 0 = 0\,, \end{aligned} \tag{A.6}$$

$$\begin{aligned} &\mathrm{E}(I_{[\tau_0]_{j_1};t,t+h}) = \mathrm{E}(\Phi_S([\tau_0]_{j_1}; t, t+h)) \\ &\Leftrightarrow \quad \mathrm{E}(I_{(0,j_1)}) = \mathrm{E}({z^{(j_1),(j_1)}}^T Z^{(j_1),(0),(0)} e) \\ &\Leftrightarrow \quad \mathrm{E}(I_{(0,j_1)}) = {\beta^{(1)}}^T A^{(1)} e\, h\, \mathrm{E}(I_{(j_1)}) + {\beta^{(2)}}^T A^{(1)} e\, h \sqrt{h} \\ &\Leftrightarrow \quad {\beta^{(2)}}^T A^{(1)} e = 0\,, \end{aligned} \tag{A.7}$$

$$\begin{aligned}
&\mathrm{E}(I_{[\tau_{j_2},\tau_{j_3}]_{j_1};t,t+h}) = \mathrm{E}(\Phi_S([\tau_{j_2},\tau_{j_3}]_{j_1};t,t+h)) \\
&\Leftrightarrow \quad 0 = \mathrm{E}(z^{(j_1),(j_1)^T}((Z^{(j_1),(j_2),(j_2)}e)(Z^{(j_1),(j_3),(j_3)}e))) \\
&\Leftrightarrow \quad 0 = \beta^{(1)^T}(B^{(1)}e)^2\ \mathrm{E}\left(I_{(j_1)}\frac{I_{(j_2,j_1)}}{\sqrt{h}}\frac{I_{(j_3,j_1)}}{\sqrt{h}}\right) \\
&\qquad\qquad + \beta^{(2)^T}(B^{(1)}e)^2\sqrt{h}\,\mathrm{E}\left(\frac{I_{(j_2,j_1)}}{\sqrt{h}}\frac{I_{(j_3,j_1)}}{\sqrt{h}}\right) \\
&\Leftrightarrow \quad 0 = 0 \quad \text{for } j_2 \neq j_3 \quad \text{and} \quad \beta^{(2)^T}(B^{(1)}e)^2 = 0 \quad \text{for } j_2 = j_3\,,
\end{aligned} \tag{A.8}$$

$$\begin{aligned}
&\mathrm{E}(I_{[[\tau_{j_3}]_{j_2}]_{j_1};t,t+h}) = \mathrm{E}(\Phi_S([[\tau_{j_3}]_{j_2}]_{j_1};t,t+h)) \\
&\Leftrightarrow \quad \mathrm{E}(I_{(j_3,j_2,j_1)}) = \mathrm{E}(z^{(j_1),(j_1)^T}(Z^{(j_1),(j_2),(j_2)}(Z^{(j_2),(j_3),(j_3)}e))) \\
&\Leftrightarrow \quad \mathrm{E}(I_{(j_3,j_2,j_1)}) = \beta^{(1)^T}(B^{(1)}(B^{(1)}e))\ \mathrm{E}\left(I_{(j_1)}\frac{I_{(j_2,j_1)}}{\sqrt{h}}\frac{I_{(j_3,j_2)}}{\sqrt{h}}\right) \\
&\qquad\qquad + \beta^{(2)^T}(B^{(1)}(B^{(1)}e))\sqrt{h}\,\mathrm{E}\left(\frac{I_{(j_2,j_1)}}{\sqrt{h}}\frac{I_{(j_3,j_2)}}{\sqrt{h}}\right) \\
&\Leftrightarrow \quad \beta^{(2)^T}(B^{(1)}(B^{(1)}e)) = 0 \quad \text{for } j_1 = j_2 = j_3 \quad \text{and} \quad 0 = 0 \quad \text{else.}
\end{aligned} \tag{A.9}$$

Summarizing the conditions (A.3)–(A.9) completes the proof. ∎

Appendix B. Proof of Theorem 6.3. The SRK method (6.7) is covered by the general class (2.1) if we choose $\mathcal{M} = \{\nu : 0 \leq \nu \leq m\}$ and use the assignment (A.1), but with

$$Z_{ij}^{(k),(l),(l)} = B_{ij}^{(1)}\,\frac{I_{(k)}I_{(l)}}{2\sqrt{h_n}} \qquad Z_{ij}^{(k),(k),(k)} = B_{ij}^{(1)}\,\frac{1}{2}\left(\frac{I_{(k)}I_{(k)}}{\sqrt{h_n}} - \sqrt{h_n}\right)$$

for $1 \leq k,l \leq m$ with $k \neq l$. For the trees $\mathbf{t}_{0.1}$, $\mathbf{t}_{0.5.1}$, $\mathbf{t}_{1.1}$, $\mathbf{t}_{1.5.1}$, and $\mathbf{t}_{1.5.2}$ the calculations are analogous to the proof of Theorem 6.1. For $\mathbf{t}_{1.3} = [\tau_{j_2}]_{j_1}$ we have to consider two cases: for $j_1 = j_2$ we obtain from condition (5.5)

$$\begin{aligned}
&I_{[\tau_{j_1}]_{j_1};t,t+h} = \Phi_S([\tau_{j_1}]_{j_1};t,t+h) \quad \Leftrightarrow \quad I_{(j_1,j_1)} = z^{(j_1),(j_1)^T}Z^{(j_1),(j_1),(j_1)}e \\
&\Leftrightarrow \quad I_{(j_1,j_1)} = (\beta^{(1)^T}I_{(j_1)} + \beta^{(2)^T}\sqrt{h})B^{(1)}e\,\frac{1}{2}\left(\frac{I_{(j_1)}I_{(j_1)}}{\sqrt{h}} - \sqrt{h}\right) \\
&\Leftrightarrow \quad \beta^{(1)^T}B^{(1)}e = 0 \quad \wedge \quad \beta^{(2)^T}B^{(1)}e = 1\,.
\end{aligned} \tag{B.1}$$

In the case of $j_1 \neq j_2$, due to the commutativity condition where $[\tau_{j_2}]_{j_1} = [\tau_{j_1}]_{j_2}$ and with the relation $I_{(j_2,j_1)} + I_{(j_1,j_2)} = I_{(j_1)}\,I_{(j_2)}$, we obtain that

$$\begin{aligned}
&I_{[\tau_{j_2}]_{j_1};t,t+h} + I_{[\tau_{j_1}]_{j_2};t,t+h} = \Phi_S([\tau_{j_2}]_{j_1};t,t+h) + \Phi_S([\tau_{j_1}]_{j_2};t,t+h) \\
&\Leftrightarrow \quad I_{(j_2,j_1)} + I_{(j_1,j_2)} = z^{(j_1),(j_1)^T}Z^{(j_1),(j_2),(j_2)}e + z^{(j_2),(j_2)^T}Z^{(j_2),(j_1),(j_1)}e \\
&\Leftrightarrow \quad I_{(j_1)}I_{(j_2)} = (\beta^{(1)^T}I_{(j_1)} + \beta^{(2)^T}\sqrt{h})B^{(1)}e\,\frac{I_{(j_1)}I_{(j_2)}}{2\sqrt{h}} \\
&\qquad + (\beta^{(1)^T}I_{(j_2)} + \beta^{(2)^T}\sqrt{h})B^{(1)}e\,\frac{I_{(j_2)}I_{(j_1)}}{2\sqrt{h}} \\
&\Leftrightarrow \quad \beta^{(1)^T}B^{(1)}e = 0 \quad \wedge \quad \beta^{(2)^T}B^{(1)}e = 1\,.
\end{aligned} \tag{B.2}$$

Finally, similar calculations for the tree $\mathbf{t}_{1.5.6} = [\tau_{j_2}, \tau_{j_3}]_{j_1}$, where the cases $j_2 \neq j_3$ and $j_2 = j_3$ have to be considered, and for $\mathbf{t}_{1.5.7}$ complete the proof. $\square$

Appendix C. Proof of Theorem 6.5. We choose $\mathcal{M} = \{\nu : 0 \leq \nu \leq m\}$ and

$$z_i^{(0),(0)} = \alpha_i\, h_n, \qquad z_i^{(k),(k)} = \beta_i^{(1)}\, I_{(k)} + \beta_i^{(2)}\, \frac{I_{(k,k)}}{\sqrt{h_n}} + \beta_i^{(3)}\, \frac{I_{(k,0)}}{h_n} + \beta_i^{(4)}\, \frac{I_{(k,k,k)}}{h_n},$$
$$Z_{ij}^{(0),(0),(0)} = A_{ij}^{(0)}\, h_n, \quad Z_{ij}^{(0),(k),(k)} = B_{ij}^{(0)}\, \frac{I_{(k,0)}}{h_n},$$
$$Z_{ij}^{(k),(0),(0)} = A_{ij}^{(1)}\, h_n, \quad Z_{ij}^{(k),(l),(l)} = B_{ij}^{(1)}\, \sqrt{h_n}$$

for $1 \leq k, l \leq m$, and we define all other coefficients in (2.1) equal to zero. Then the SRK method (6.12) is covered by the class (2.1) of SRK methods, and we can apply Proposition 5.2 with $p = 1.5$. Due to the diagonal noise the elementary differentials vanish for all trees which contain a sequence of nodes of type $[\ldots, [\ldots]_{j_2}]_{j_1}$ where at least one node τ_{j_2} directly succeeds a node τ_{j_1} for some $j_1, j_2 \in \{1, \ldots, m\}$ with $j_1 \neq j_2$. Thus, we do not need to consider such trees. As a result of this, we obtain with $I_{(0,j_1)} = h\, I_{(j_1)} - I_{(j_1,0)}$ for $j_1 \in \{1, \ldots, m\}$ the following conditions from (5.5):

$$\text{(C.1)} \quad I_{\gamma;t,t+h} = \Phi_S(\gamma; t, t+h) \quad \Leftrightarrow \quad 1 = 1\,,$$

$$I_{\tau_{j_1};t,t+h} = \Phi_S(\tau_{j_1}; t, t+h) \quad \Leftrightarrow \quad I_{(j_1)} = {z^{(j_1),(j_1)}}^T e$$
$$\Leftrightarrow \quad I_{(j_1)} = {\beta^{(1)}}^T e\, I_{(j_1)} + {\beta^{(2)}}^T e\, \frac{I_{(j_1,j_1)}}{\sqrt{h}} + {\beta^{(3)}}^T e\, \frac{I_{(j_1,0)}}{h} + {\beta^{(4)}}^T e\, \frac{I_{(j_1,j_1,j_1)}}{h}$$
$$\text{(C.2)} \quad \Leftrightarrow \quad {\beta^{(1)}}^T e = 1 \quad \wedge \quad {\beta^{(2)}}^T e = {\beta^{(3)}}^T e = {\beta^{(4)}}^T e = 0\,,$$

$$I_{\tau_0;t,t+h} = \Phi_S(\tau_0; t, t+h) \quad \Leftrightarrow \quad I_{(0)} = {z^{(0),(0)}}^T e$$
$$\text{(C.3)} \quad \Leftrightarrow \quad I_{(0)} = \alpha^T e\, h \quad \Leftrightarrow \quad \alpha^T e = 1\,,$$

$$I_{[\tau_{j_1}]_{j_1};t,t+h} = \Phi_S([\tau_{j_1}]_{j_1}; t, t+h) \quad \Leftrightarrow \quad I_{(j_1,j_1)} = {z^{(j_1),(j_1)}}^T Z^{(j_1),(j_1),(j_1)} e$$
$$\Leftrightarrow \quad I_{(j_1,j_1)} = \left({\beta^{(1)}}^T I_{(j_1)} + {\beta^{(2)}}^T \frac{I_{(j_1,j_1)}}{\sqrt{h}} + {\beta^{(3)}}^T \frac{I_{(j_1,0)}}{h} \right.$$
$$\left. + {\beta^{(4)}}^T \frac{I_{(j_1,j_1,j_1)}}{h} \right) B^{(1)} e\, \sqrt{h}$$
$$\text{(C.4)} \quad \Leftrightarrow \quad {\beta^{(1)}}^T B^{(1)} e = {\beta^{(3)}}^T B^{(1)} e = {\beta^{(4)}}^T B^{(1)} e = 0 \quad \wedge \quad {\beta^{(2)}}^T B^{(1)} e = 1\,,$$

$$I_{[\tau_{j_1}]_0;t,t+h} = \Phi_S([\tau_{j_1}]_0; t, t+h) \quad \Leftrightarrow \quad I_{(j_1,0)} = {z^{(0),(0)}}^T Z^{(0),(j_1),(j_1)} e$$
$$\text{(C.5)} \quad \Leftrightarrow \quad I_{(j_1,0)} = \alpha^T B^{(0)} e\, h\, \frac{I_{(j_1,0)}}{h} \quad \Leftrightarrow \quad \alpha^T B^{(0)} e = 1\,,$$

$$I_{[\tau_0]_{j_1};t,t+h} = \Phi_S([\tau_0]_{j_1}; t, t+h) \quad \Leftrightarrow \quad I_{(0,j_1)} = {z^{(j_1),(j_1)}}^T Z^{(j_1),(0),(0)} e$$
$$\Leftrightarrow \quad I_{(0,j_1)} = \left({\beta^{(1)}}^T I_{(j_1)} + {\beta^{(2)}}^T \frac{I_{(j_1,j_1)}}{\sqrt{h}} + {\beta^{(3)}}^T \frac{I_{(j_1,0)}}{h} \right.$$
$$\left. + {\beta^{(4)}}^T \frac{I_{(j_1,j_1,j_1)}}{h} \right) A^{(1)} e\, h$$
$$\text{(C.6)} \quad \Leftrightarrow \quad {\beta^{(1)}}^T A^{(1)} e = 1 \ \wedge \ {\beta^{(2)}}^T A^{(1)} e = {\beta^{(4)}}^T A^{(1)} e = 0 \ \wedge \ {\beta^{(3)}}^T A^{(1)} e = -1\,,$$

$$
\begin{aligned}
& I_{[\tau_{j_1},\tau_{j_1}]_{j_1};t,t+h} = \Phi_S([\tau_{j_1},\tau_{j_1}]_{j_1};t,t+h) \\
\Leftrightarrow \quad & 2I_{(j_1,j_1,j_1)} + I_{(0,j_1)} = {z^{(j_1),(j_1)}}^T (Z^{(j_1),(j_1),(j_1)} e)^2 \\
\Leftrightarrow \quad & 2I_{(j_1,j_1,j_1)} + I_{(0,j_1)} = \Bigg({\beta^{(1)}}^T I_{(j_1)} + {\beta^{(2)}}^T \frac{I_{(j_1,j_1)}}{\sqrt{h}} + {\beta^{(3)}}^T \frac{I_{(j_1,0)}}{h} \\
& \qquad\qquad + {\beta^{(4)}}^T \frac{I_{(j_1,j_1,j_1)}}{h} \Bigg) (B^{(1)} e)^2 \, h \\
\Leftrightarrow \quad & {\beta^{(1)}}^T (B^{(1)}e)^2 = 1 \ \wedge \ {\beta^{(2)}}^T (B^{(1)}e)^2 = 0 \ \wedge \ {\beta^{(3)}}^T (B^{(1)}e)^2 = -1 \\
& \qquad \wedge \ {\beta^{(4)}}^T (B^{(1)}e)^2 = 2 \, , \qquad \text{(C.7)}
\end{aligned}
$$

$$
\begin{aligned}
& I_{[[\tau_{j_1}]_{j_1}]_{j_1};t,t+h} = \Phi_S([[\tau_{j_1}]_{j_1}]_{j_1};t,t+h) \\
\Leftrightarrow \quad & I_{(j_1,j_1,j_1)} = {z^{(j_1),(j_1)}}^T (Z^{(j_1),(j_1),(j_1)} (Z^{(j_1),(j_1),(j_1)} e)) \\
\Leftrightarrow \quad & I_{(j_1,j_1,j_1)} = \Bigg({\beta^{(1)}}^T I_{(j_1)} + {\beta^{(2)}}^T \frac{I_{(j_1,j_1)}}{\sqrt{h}} + {\beta^{(3)}}^T \frac{I_{(j_1,0)}}{h} \\
& \qquad\qquad + {\beta^{(4)}}^T \frac{I_{(j_1,j_1,j_1)}}{h} \Bigg) (B^{(1)} (B^{(1)} e)) \, h \\
\Leftrightarrow \quad & {\beta^{(1)}}^T (B^{(1)}(B^{(1)}e)) = {\beta^{(2)}}^T (B^{(1)}(B^{(1)}e)) = {\beta^{(3)}}^T (B^{(1)}(B^{(1)}e)) = 0 \\
& \qquad \wedge \ {\beta^{(4)}}^T (B^{(1)}(B^{(1)}e)) = 1 \, . \qquad \text{(C.8)}
\end{aligned}
$$

Next, we apply (5.6) to the second order trees $\mathbf{t}_{2.1} = [\tau_0]_0$, $\mathbf{t}_{2.3} = [[\tau_{j_1}]_{j_1}]_0$, $\mathbf{t}_{2.4a} = [\tau_{j_1},\tau_{j_1}]_0$, $\mathbf{t}_{2.4b} = [\tau_{j_1},\tau_{j_2}]_0$, $\mathbf{t}_{2.20a} = [[\tau_{j_1}]_0]_{j_1}$, and $\mathbf{t}_{2.20b} = [[\tau_{j_2}]_0]_{j_1}$ with $j_1, j_2 \in \{1,\dots,m\}$ and $j_1 \neq j_2$, where we calculate the conditions

$$
\begin{aligned}
& \mathrm{E}(I_{[\tau_0]_0;t,t+h}) = \mathrm{E}(\Phi_S([\tau_0]_0;t,t+h)) \quad \Leftrightarrow \quad \mathrm{E}(I_{(0,0)}) = \mathrm{E}({z^{(0),(0)}}^T Z^{(0),(0),(0)} e) \\
\Leftrightarrow \quad & \frac{1}{2} h^2 = \alpha^T A^{(0)} e \, h^2 \quad \Leftrightarrow \quad \alpha^T A^{(0)} e = \frac{1}{2} \, , \qquad \text{(C.9)}
\end{aligned}
$$

$$
\begin{aligned}
& \mathrm{E}(I_{[[\tau_{j_1}]_{j_1}]_0;t,t+h}) = \mathrm{E}(\Phi_S([[\tau_{j_1}]_{j_1}]_0;t,t+h)) \\
\Leftrightarrow \quad & \mathrm{E}(I_{(j_1,j_1,0)}) = \mathrm{E}({z^{(0),(0)}}^T (Z^{(0),(j_1),(j_1)} (Z^{(j_1),(j_1),(j_1)} e))) \\
\Leftrightarrow \quad & \mathrm{E}(I_{(j_1,j_1,0)}) = \alpha^T (B^{(0)}(B^{(1)}e)) h \sqrt{h} \, \mathrm{E}\left(\frac{I_{(j_1,0)}}{h} \right) \quad \Leftrightarrow \quad 0 = 0 \, , \qquad \text{(C.10)}
\end{aligned}
$$

$$
\begin{aligned}
& \mathrm{E}(I_{[\tau_{j_1},\tau_{j_1}]_0;t,t+h}) = \mathrm{E}(\Phi_S([\tau_{j_1},\tau_{j_1}]_0;t,t+h)) \\
\Leftrightarrow \quad & \mathrm{E}(2I_{(j_1,j_1,0)} + I_{(0,0)}) = \mathrm{E}({z^{(0),(0)}}^T (Z^{(0),(j_1),(j_1)} e)^2) \\
\Leftrightarrow \quad & \mathrm{E}(2I_{(j_1,j_1,0)} + I_{(0,0)}) = \alpha^T (B^{(0)}e)^2 h \, \mathrm{E}\left(\frac{I_{(j_1,0)}}{h} \frac{I_{(j_1,0)}}{h} \right) \\
\Leftrightarrow \quad & \frac{1}{2} h^2 = \alpha^T (B^{(0)}e)^2 \frac{1}{3} h \quad \Leftrightarrow \quad \alpha^T (B^{(0)}e)^2 = \frac{3}{2} \, , \qquad \text{(C.11)}
\end{aligned}
$$

$$\begin{aligned}
&\mathrm{E}(I_{[\tau_{j_1},\tau_{j_2}]_0;t,t+h}) = \mathrm{E}(\Phi_S([\tau_{j_1},\tau_{j_2}]_0;t,t+h)) \\
\Leftrightarrow\quad &\mathrm{E}(I_{(j_1,j_2,0)} + I_{(j_2,j_1,0)}) = \mathrm{E}(z^{(0),(0)^T}((Z^{(0),(j_1),(j_1)}e)(Z^{(0),(j_2),(j_2)}e))) \\
\Leftrightarrow\quad &\mathrm{E}(I_{(j_1,j_2,0)} + I_{(j_2,j_1,0)}) = \alpha^T(B^{(0)}e)^2 h\,\mathrm{E}\left(\frac{I_{(j_1,0)}}{h}\frac{I_{(j_2,0)}}{h}\right) \quad\Leftrightarrow\quad 0=0\,,
\end{aligned}$$

(C.12)

$$\begin{aligned}
&\mathrm{E}(I_{[[\tau_{j_1}]_0]_{j_1};t,t+h}) = \mathrm{E}(\Phi_S([[\tau_{j_1}]_0]_{j_1};t,t+h)) \\
\Leftrightarrow\quad &\mathrm{E}(I_{(j_1,0,j_1)}) = \mathrm{E}(z^{(j_1),(j_1)^T}(Z^{(j_1),(0),(0)}(Z^{(0),(j_1),(j_1)}e))) \\
\Leftrightarrow\quad &\mathrm{E}(I_{(j_1,0,j_1)}) = \mathrm{E}\Bigg(\bigg(\beta^{(1)^T} I_{(j_1)} + \beta^{(2)^T}\frac{I_{(j_1,j_1)}}{\sqrt{h}} + \beta^{(3)^T}\frac{I_{(j_1,0)}}{h} \\
&\qquad\qquad + \beta^{(4)^T}\frac{I_{(j_1,j_1,j_1)}}{h}\bigg)(A^{(1)}(B^{(0)}e))h\frac{I_{(j_1,0)}}{h}\Bigg) \\
\Leftrightarrow\quad &\frac{1}{2}\beta^{(1)^T}(A^{(1)}(B^{(0)}e)) + \frac{1}{3}\beta^{(3)^T}(A^{(1)}(B^{(0)}e)) = 0\,,
\end{aligned}$$

(C.13)

$$\begin{aligned}
&\mathrm{E}(I_{[[\tau_{j_2}]_0]_{j_1};t,t+h}) = \mathrm{E}(\Phi_S([[\tau_{j_2}]_0]_{j_1};t,t+h)), \quad j_1 \neq j_2, \\
\Leftrightarrow\quad &\mathrm{E}(I_{(j_2,0,j_1)}) = \mathrm{E}(z^{(j_1),(j_1)^T}(Z^{(j_1),(0),(0)}(Z^{(0),(j_2),(j_2)}e))) \\
\Leftrightarrow\quad &\mathrm{E}(I_{(j_2,0,j_1)}) = \mathrm{E}\Bigg(\bigg(\beta^{(1)^T} I_{(j_1)} + \beta^{(2)^T}\frac{I_{(j_1,j_1)}}{\sqrt{h}} + \beta^{(3)^T}\frac{I_{(j_1,0)}}{h} \\
&\qquad + \beta^{(4)^T}\frac{I_{(j_1,j_1,j_1)}}{h}\bigg)(A^{(1)}(B^{(0)}e))h\frac{I_{(j_2,0)}}{h}\Bigg) \quad\Leftrightarrow\quad 0=0\,.
\end{aligned}$$

(C.14)

Next, we calculate that $\mathrm{E}(I_{\mathbf{t};t,t+h}) = \mathrm{E}(\Phi_S(\mathbf{t};t,t+h)) = 0$ for the remaining second order trees $\mathbf{t}_{2.9} = [[\tau_0]_{j_1}]_{j_1}$, $\mathbf{t}_{2.10} = [\tau_{j_1},\tau_0]_{j_1}$, $\mathbf{t}_{2.16} = [\tau_{j_1},\tau_{j_1},\tau_{j_1}]_{j_1}$, $\mathbf{t}_{2.17} = [\tau_{j_1},[\tau_{j_1}]_{j_1}]_{j_1}$, $\mathbf{t}_{2.18} = [[\tau_{j_1},\tau_{j_1}]_{j_1}]_{j_1}$, and $\mathbf{t}_{2.19} = [[[\tau_{j_1}]_{j_1}]_{j_1}]_{j_1}$ for $j_1 \in \{1,\ldots,m\}$; i.e., these trees contribute no additional order conditions. Summarizing the conditions (C.2)–(C.14) finally results in Theorem 6.5. ∎

Appendix D. Proof of Theorem 6.6. We choose $\mathcal{M} = \{\nu : 0 \le \nu \le m\}$ and

$$z_i^{(0),(0)} = \alpha_i\,h_n, \qquad z_i^{(k),(k)} = \beta_i^{(1)}\,I_{(k)} + \beta_i^{(2)}\,\frac{I_{(k,0)}}{h_n},$$
$$Z_{ij}^{(0),(0),(0)} = A_{ij}^{(0)}\,h_n, \qquad Z_{ij}^{(0),(k),(k)} = B_{ij}^{(0)}\,\frac{I_{(k,0)}}{h_n}$$

for $1 \le k \le m$ and define all other coefficients in (2.1) equal to zero. Then the SRK method (6.14) is contained in the class (2.1) and Proposition 5.2 applies with $p = 1.5$. For additive noise the elementary differentials vanish for all trees that contain a sequence of nodes where at least one node τ_0 or τ_{j_2} directly succeeds a stochastic node τ_{j_1} for some $j_1, j_2 \in \{1,\ldots,m\}$ except if the succeeding node τ_0 is the solely succeeding end node. Thus, we need not consider such trees for the analysis of order conditions. In the case of a singly succeeding end node τ_0 of a node τ_{j_1}, we have to replace the usually appearing coefficients $A^{(1)}e$ by $c^{(1)}$ because we have no stage values for the diffusion b in the case of additive noise, although there exists a derivative w.r.t. time. So we get new conditions for the coefficients $c^{(1)}$ instead of (2.2) while

(2.2) can be still applied for the definition of $c^{(0)}$. Due to $I_{(0,j_1)} = h\, I_{(j_1)} - I_{(j_1,0)}$ for $j_1 \in \{1,\dots,m\}$ we calculate from (5.5) the following conditions:

$$\text{(D.1)} \qquad I_{\gamma;t,t+h} = \Phi_S(\gamma;t,t+h) \quad \Leftrightarrow \quad 1 = 1\,,$$

$$I_{\tau_{j_1};t,t+h} = \Phi_S(\tau_{j_1};t,t+h) \quad \Leftrightarrow \quad I_{(j_1)} = {z^{(j_1),(j_1)}}^T e$$
$$\Leftrightarrow \quad I_{(j_1)} = {\beta^{(1)}}^T e\, I_{(j_1)} + {\beta^{(2)}}^T e\, \frac{I_{(j_1,0)}}{h}$$
$$\text{(D.2)} \qquad \Leftrightarrow \quad {\beta^{(1)}}^T e = 1 \quad \wedge \quad {\beta^{(2)}}^T e = 0\,,$$

$$I_{\tau_0;t,t+h} = \Phi_S(\tau_0;t,t+h) \quad \Leftrightarrow \quad I_{(0)} = {z^{(0),(0)}}^T e$$
$$\text{(D.3)} \qquad \Leftrightarrow \quad I_{(0)} = \alpha^T e\, h \quad \Leftrightarrow \quad \alpha^T e = 1\,,$$

$$I_{[\tau_{j_1}]_0;t,t+h} = \Phi_S([\tau_{j_1}]_0;t,t+h) \quad \Leftrightarrow \quad I_{(j_1,0)} = {z^{(0),(0)}}^T Z^{(0),(j_1),(j_1)} e$$
$$\text{(D.4)} \qquad \Leftrightarrow \quad I_{(j_1,0)} = \alpha^T B^{(0)} e\, h\, \frac{I_{(j_1,0)}}{h} \quad \Leftrightarrow \quad \alpha^T B^{(0)} e = 1\,,$$

$$I_{[\tau_0]_{j_1};t,t+h} = \Phi_S([\tau_0]_{j_1};t,t+h) \quad \Leftrightarrow \quad I_{(0,j_1)} = {z^{(j_1),(j_1)}}^T c^{(1)}$$
$$\Leftrightarrow \quad I_{(0,j_1)} = \left({\beta^{(1)}}^T I_{(j_1)} + {\beta^{(2)}}^T \frac{I_{(j_1,0)}}{\sqrt{h}} \right) c^{(1)}\, h$$
$$\text{(D.5)} \qquad \Leftrightarrow \quad {\beta^{(1)}}^T c^{(1)} = 1 \ \wedge \ {\beta^{(2)}}^T c^{(1)} = -1\,.$$

Next, we apply (5.6) to the second order trees $\mathbf{t}_{2.1} = [\tau_0]_0$, $\mathbf{t}_{2.4a} = [\tau_{j_1}, \tau_{j_1}]_0$, and $\mathbf{t}_{2.4b} = [\tau_{j_1}, \tau_{j_2}]_0$ with $j_1, j_2 \in \{1,\dots,m\}$ and $j_1 \neq j_2$:

$$\mathrm{E}(I_{[\tau_0]_0;t,t+h}) = \mathrm{E}(\Phi_S([\tau_0]_0;t,t+h)) \quad \Leftrightarrow \quad \mathrm{E}(I_{(0,0)}) = \mathrm{E}({z^{(0),(0)}}^T Z^{(0),(0),(0)} e)$$
$$\Leftrightarrow \quad \frac{1}{2}h^2 = \alpha^T A^{(0)} e\, h^2 \quad \Leftrightarrow \quad \alpha^T A^{(0)} e = \frac{1}{2}\,,$$
(D.6)

$$\mathrm{E}(I_{[\tau_{j_1},\tau_{j_1}]_0;t,t+h}) = \mathrm{E}(\Phi_S([\tau_{j_1},\tau_{j_1}]_0;t,t+h))$$
$$\Leftrightarrow \quad \mathrm{E}(2I_{(j_1,j_1,0)} + I_{(0,0)}) = \mathrm{E}({z^{(0),(0)}}^T (Z^{(0),(j_1),(j_1)} e)^2)$$
$$\Leftrightarrow \quad \mathrm{E}(2I_{(j_1,j_1,0)} + I_{(0,0)}) = \alpha^T (B^{(0)} e)^2 h\, \mathrm{E}\left(\frac{I_{(j_1,0)}}{h} \frac{I_{(j_1,0)}}{h} \right)$$
$$\Leftrightarrow \quad \frac{1}{2}h^2 = \alpha^T (B^{(0)} e)^2 \frac{1}{3} h \quad \Leftrightarrow \quad \alpha^T (B^{(0)} e)^2 = \frac{3}{2}\,,$$
(D.7)

$$\mathrm{E}(I_{[\tau_{j_1},\tau_{j_2}]_0;t,t+h}) = \mathrm{E}(\Phi_S([\tau_{j_1},\tau_{j_2}]_0;t,t+h)), \quad j_1 \neq j_2,$$
$$\Leftrightarrow \quad \mathrm{E}(I_{(j_1,j_2,0)} + I_{(j_2,j_1,0)}) = \mathrm{E}({z^{(0),(0)}}^T ((Z^{(0),(j_1),(j_1)} e)(Z^{(0),(j_2),(j_2)} e)))$$
$$\Leftrightarrow \quad \mathrm{E}(I_{(j_1,j_2,0)} + I_{(j_2,j_1,0)}) = \alpha^T (B^{(0)} e)^2 h\, \mathrm{E}\left(\frac{I_{(j_1,0)}}{h} \frac{I_{(j_2,0)}}{h} \right) \Leftrightarrow \ 0 = 0\,.$$
(D.8)

Summarizing the conditions (D.2)–(D.8) finally results in Theorem 6.6. $\square$

Acknowledgment. The author is very grateful for helpful suggestions of the referees.

REFERENCES

[1] K. Burrage and P. M. Burrage, *High strong order explicit Runge–Kutta methods for stochastic ordinary differential equations*, Appl. Numer. Math., 22 (1996), pp. 81–101.

[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., 28 (1998), pp. 161–177.

[3] K. Burrage and P. M. Burrage, *Order conditions of stochastic Runge–Kutta methods by B-series*, SIAM J. Numer. Anal., 38 (2000), pp. 1626–1646.

[4] P. M. Burrage, *Runge-Kutta Methods for Stochastic Differential Equations*, Ph.D. thesis, Dept. of Math., University of Queensland, Australia, 1999.

[5] J. M. C. Clark and R. J. Cameron, *The maximum rate of convergence of discrete approximations for stochastic differential equations*, in Stochastic Differential Systems, Proceedings of the IFIP-WG 7/1 Working Conference, Lecture Notes in Control Inform. Sci. 25, Springer, Berlin, 1980, pp. 162–171.

[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., 47 (2008), pp. 181–203.

[7] J. G. Gaines and T. J. Lyons, *Random generation of stochastic area integrals*, SIAM J. Appl. Math., 54 (1994), pp. 1132–1146.

[8] M. B. Giles, *Improved multilevel Monte Carlo convergence using the Milstein scheme*, in Monte Carlo and Quasi-Monte Carlo Methods 2006, Springer, Berlin, 2008, pp. 343–358.

[9] H. Gilsing and T. Shardlow, *SDELab: A package for solving stochastic differential equations in MATLAB*, J. Comput. Appl. Math., 205 (2007), pp. 1002–1018.

[10] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations* I, Springer, Berlin, 1993.

[11] J. Kaneko, *Explicit order* 1.5 *Runge-Kutta scheme for solutions of Itô stochastic differential equations*, in Various Problems in Stochastic Numerical Analysis II (Kyoto), Sūrikaisekikenkyūsho Kōkyūroku 932, Kyoto University Research Institute for Mathematical Sciences, Kyoto, Japan, 1995, pp. 46–60.

[12] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer, Berlin, 1999.

[13] P. E. Kloeden, E. Platen, and I. W. Wright, *The approximation of multiple stochastic integrals*, Stoch. Anal. Appl., 10 (1992), pp. 431–441.

[14] Y. Komori, *Multi-colored rooted tree analysis of the weak order conditions of a stochastic Runge-Kutta family*, Appl. Numer. Math., 57 (2007), pp. 147–165.

[15] Y. Komori, *Weak second-order stochastic Runge-Kutta methods for non-commutative stochastic differential equations*, J. Comput. Appl. Math., 206 (2007), pp. 158–173.

[16] Y. Komori, *Weak order stochastic Runge-Kutta methods for commutative stochastic differential equations*, J. Comput. Appl. Math., 203 (2007), pp. 57–79.

[17] Y. Komori, T. Mitsui, and H. Sugiura, *Rooted tree analysis of the order conditions of ROW-type scheme for stochastic differential equations*, BIT, 37 (1997), pp. 43-66.

[18] D. Küpper, J. Lehn, and A. Rössler, *A step size control algorithm for the weak approximation of stochastic differential equations*, Numer. Algorithms, 44 (2007), pp. 335–346.

[19] G. N. Milstein, *Numerical Integration of Stochastic Differential Equations*, Kluwer Academic, Dordrecht, The Netherlands, 1995.

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

[21] N. J. Newton, *Asymptotically efficient Runge–Kutta methods for a class of Itô and Stratonovich equations*, SIAM J. Appl. Math., 51 (1991), pp. 542–567.

[22] A. Rössler, *Stochastic Taylor expansions for the expectation of functionals of diffusion processes*, Stoch. Anal. Appl., 22 (2004), pp. 1553–1576.

[23] A. Rössler, *Rooted tree analysis for order conditions of stochastic Runge–Kutta methods for the weak approximation of stochastic differential equations*, Stoch. Anal. Appl., 24 (2006), pp. 97–134.

[24] A. Rössler, *Second order Runge–Kutta methods for Stratonovich stochastic differential equations*, BIT, 47 (2007), pp. 657–680.

[25] A. Rössler, *Stochastic Taylor expansions for functionals of diffusion processes*, Stoch. Anal. Appl., 28 (2010), pp. 415–429.

[26] A. Rössler, *Second order Runge–Kutta methods for Itô stochastic differential equations*, SIAM J. Numer. Anal., 47 (2009), pp. 1713–1738.
[27] W. Rümelin, *Numerical treatment of stochastic differential equations*, SIAM J. Numer. Anal., 19 (1982), pp. 604–613.
[28] D. Talay, *Second-order discretization schemes of stochastic differential systems for the computation of the invariant law*, Stoch. Stoch. Rep., 29 (1990), pp. 13–36.
[29] A. Tocino and J. Vigo-Aguiar, *Weak second order conditions for stochastic Runge–Kutta methods*, SIAM J. Sci. Comput., 24 (2002), pp. 507–523.
[30] M. Wiktorsson, *Joint characteristic function and simultaneous simulation of iterated Itô integrals for multiple independent Brownian motions*, Ann. Appl. Probab., 11 (2001), pp. 470–487.